The SPARC Data Initiative: Assessment of stratospheric trace gas and aerosol climatologies from satellite limb sounders

March 2017

Prepared by the SPARC Data Initiative Team,
edited by M. I. Hegglin and S. Tegtmeier

SPARC Report No. 8, WCRP-05/2017

SPARC
Stratosphere-troposphere
Processes And their Role in Climate

WCRP
World Climate Research Programme

CSA ASC

esa

INTERNATIONAL
SPACE
SCIENCE
INSTITUTE

Citation
SPARC, 2017: The SPARC Data Initiative: Assessment of stratospheric trace gas and aerosol climatologies from satellite limb sounders. M. I. Hegglin and S. Tegtmeier (Eds.), SPARC Report No. 8, WCRP-05/2017, DOI 10.3929/ethz-a-010863911, available at www.sparc-climate.org/publications/sparc-reports/.

Electronic version available at http://www.sparc-climate.org/publications/sparc-reports/sparc-report-no8/

ISSN 2296-5785 (Print)

Editing, Design, and Layout: Petra Bratfisch, Carolin Arndt and Diane Pendlebury

Authors & Reviewers

Lead Authors / Editors / Co-leads of the SPARC Data Initiative

Michaela I. Hegglin	University of Reading	United Kingdom
Susann Tegtmeier	GEOMAR Helmholtz Centre for Ocean Research Kiel	Germany

Authors

John Anderson	Hampton University	USA
Adam Bourassa	University of Saskatchewan	Canada
Samuel Brohede	Chalmers University of Technology	Sweden
Doug Degenstein	University of Saskatchewan	Canada
Lucien Froidevaux	Jet Propulsion Laboratory, California Institute of Technology	USA
Bernd Funke	Instituto de Astrofísica de Andalucía, CSIC	Spain
John Gille	University of Colorado and NCAR	USA
Ashley Jones	University of Toronto	Canada
Yasuko Kasai	NICT	Japan
Erkki Kyrölä	Finnish Meteorological Institute	Finland
Jerry Lumpe	Computational Physics, Inc.	USA
Jessica Neu	Jet Propulsion Laboratory, California Institute of Technology	USA
Ellis Remsberg	NASA Langley Research Center	USA
Alexei Rozanov	University of Bremen	Germany
Matthew Toohey	GEOMAR Helmholtz Centre for Ocean Research Kiel	Germany
Joachim Urban	Deceased/Chalmers University of Technology	Sweden
Thomas von Clarmann	Karlsruhe Institute of Technology	Germany
Kaley A. Walker	University of Toronto	Canada
Ray Wang	Georgia Institute of Technology	USA

Contributing Authors

Gregory Bodeker	Bodeker Scientific	New Zealand
Juan A. Añel Cabanelas	Universidade de Vigo	Spain
Florian Ernst	University of Bremen	Germany
Ryan Fuller	Jet Propulsion Laboratory, California Institute of Technology	USA
Chris McLinden	Environment Canada	Canada
Gretchen Lingenfelser	NASA Langley Research Center	USA
Bruno Nardi	NCAR	USA
Diane Pendlebury	University of Toronto	Canada
Charles Robert	BIRA	Belgium
Chris Roth	University of Saskatchewan	Canada
Christian von Savigny	University of Bremen	Germany
Lesley Smith	University of Colorado	USA
Katja Weigel	University of Bremen	Germany

Reviewers

Peter Braesike	Karlsruhe Institute of Technology	Germany
Veronika Eyring	DLR	Germany
Jens-Uwe Grooß	Research Centre Juelich	Germany
Neil Harris	University of Cambridge	United Kingdom
Gloria L. Manney	NorthWest Research Associates (NWRA) and NMT	USA
Chris McLinden	Environment and Climate Change Canada	Canada
Gerald Nedoluha	Naval Research Laboratory	USA
Mijeong Park	NCAR	USA
William Randel	NCAR	USA
Karen Rosenlof	NOAA	USA
Michelle Santee	Jet Propulsion Laboratory, California Institute of Technology	USA
Robyn Schofield	University of Melbourne	Australia
Ted Shepherd	University of Reading	United Kingdom
Johannes Staehelin	ETH Zurich	Switzerland
Susan Strahan	NASA Goddard Research Center	USA
Kimberly Strong	University of Toronto	Canada
Michel Van Roozendael	BIRA	Belgium
Mark Weber	University of Bremen	Germany

In Memory of Joachim Urban
2 December 1964 - 14 August 2014

Acknowledgements

While the SPARC Data Initiative has been driven from a user perspective, the measurement partners have been critical to its success. These partners to whom the SPARC Data Initiative extends its thanks include the relevant instrument teams, the various space agencies (CSA, ESA, NASA, JAXA, SNSB, and other national agencies), and organisations such as CEOS-ACC and IGACO. The SPARC Data Initiative particularly thanks the International Space Science Institute in Bern (ISSI) who supported the activity through their ISSI International Team activity program and facilitated two successful team meetings in Bern. The SPARC Data Initiative also extends its thanks to the World Climate Research Programme (WCRP) and the Toronto SPARC office (sponsored by the CSA and the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) pogramme) for providing generous travel support for team and review meetings. The Zurich SPARC Office (through ETH Zurich) covered the costs of publication and dissemination, including support for the editing and layout of the report by Petra Bratfisch and Carolin Arndt. The Toronto SPARC Office (through CSA) supported the work on editing and layout of the report by Diane Pendlebury. Finally, Michaela I. Hegglin thanks the CSA and ESA for supporting her work within the SPARC Data Initiative through the CSA SSEP (9SCIGRA-29) and ESA STSE-SPIN (4000105291/12/I-NB) programmes, respectively. The work from Susann Tegtmeier was funded from the WGL project TransBrom and the EU project SHIVA (FP7-ENV-2007-1-226224). Work at the Jet Propulsion Laboratory, California Institute of Technology, was funded by NASA. Development of the ACE-FTS climatologies was supported by grants from the CFCAS and the CSA. Work at Instituto de Astrofísica de Andalucía, CSIC, was supported by the Spanish MCINN under grant ESP2014-54362-P and EC FEDER funds. Development of the SCIAMACHY climatologies was funded in part by the German Aerospace Agency, ESA, and by the University and State of Bremen.

Executive summary

Preface

The past 30 years have been a 'golden age' for vertically resolved satellite measurements and provided a wealth of knowledge regarding atmospheric constituents in the stratosphere. Datasets of chemical trace gases and aerosol are widely used for empirical studies of stratospheric climate, trends and variability, for process studies and for the evaluation of the representation of transport and chemistry in numerical models. The datasets available from the different satellite instruments vary in terms of measurement method, geographical coverage, spatial and temporal sampling and resolution, time period, and retrieval algorithm. Basic information on the availability, quality and consistency of the datasets is required for all scientific applications of the data, and in particular to evaluate Chemistry-Climate Models or to merge datasets from various sources into homogeneous data records suitable for trend studies. However, as pointed out by the SPARC Chemistry-Climate Model Validation (CCMVal) report [*SPARC*, 2010], knowledge on the availability and quality of satellite observations needed for meaningful model-measurement comparison exercises is not always readily available.

The Stratosphere-troposphere Processes And their Role in Climate (SPARC) core project of the World Climate Research Programme (WCRP) initiated the SPARC Data Initiative in 2009 with the mandate to coordinate an assessment of available, vertically resolved chemical trace gas and aerosol observations obtained from a multi-national suite of space-based satellite instruments. The SPARC Data Initiative assessed, in a first step, the current availability of such past and present vertically resolved, chemical trace gas and aerosol datasets. In a second step, chemical trace gas and aerosol monthly zonal mean time series were compiled in a common and simple-to-use data format. The focus thereby lay on the trace gas and aerosol products that were considered relatively mature, neglecting some of the research products available for minor trace gas species. It is also important to note that not all measurements of satellite instruments available could be included in the report due to the lack of resources to produce the monthly zonal mean climatologies. The monthly zonal mean time series of the observations are provided to the end user via the SPARC Data Center (**http://www.sparc-climate.org/data-center/data-access/sparc-data-initiative/**). In a third step, climatologies derived from the monthly zonal mean time series underwent detailed comparisons, which identified strengths and shortcomings of all datasets and differences between them.

The detailed comparisons of the chemical trace gas and aerosol climatologies are presented in the SPARC Data Initiative report. Basic information on quality and consistency of the various data products is provided. Differences between the climatologies and unphysical behaviour of individual datasets are identified and, where possible, an expert judgment on the source of those differences is given. In the report presented here the spread in the climatologies is used to provide an estimate of the overall systematic uncertainty in our knowledge of the true atmospheric state. Such an assessment of the relative uncertainty in the mean trace gas fields yields information on how well we know the global annual mean distribution of each gas and will help to identify regions where more detailed evaluations or more data are needed. The report provides recommendations for model-measurement inter-comparisons and points out particular diagnostics that would be especially appropriate for model evaluations. Additionally, the different instrument techniques and retrieval procedures are documented. The evaluations do not include trend studies or assess instrument drifts but provide valuable information to activities focused on characterizing long-term changes and data merging efforts. Note that the SPARC Data Initiative does currently not provide a merged dataset of the individual instruments' monthly zonal mean time series.

The report is targeted at various kinds of data users, including (1) scientists seeking an introduction into and overview of available stratospheric satellite datasets and their quality, (2) data analysts aiming at data merging exercises and trend evaluations, and (3) scientists working with Chemistry-Climate Models aiming at model-measurement comparisons. The report also aims at providing guidance and feedback to space agencies about required improvements in existing datasets and the need for future observations.

The objective of the SPARC Data Initiative report is an assessment of the atmospheric trace gas climatologies based on data versions available at the time. These data versions are (especially for the more recent satellite instruments) undergoing revisions regularly in order to account for known shortcomings in the instrument retrievals, so some of the findings in this report may become obsolete when moving to newer data versions (although the main characteristics of an instrument are not expected to change substantially). In fact, the SPARC Data Initiative's comparison results have already triggered major revisions of some of the datasets which are now included in their revised form. The SPARC Data Initiative climatologies will be updated with newer data versions as soon as they become available, and information on the changes in the data versions will be provided as appropriate in the future. While the SPARC Data Initiative report presented here focuses

Table ES.1: ***Atmospheric constituent climatologies available from the SPARC Data Initiative archive, listed by instrument.*** *Blue indicates the participating limb sounders, grey the nadir sounder Aura-TES, which was included for a comparison of ozone in the upper troposphere and lower stratosphere (UTLS).*

SPARC Data Initiative	O_3	H_2O	CH_4	N_2O	CCl_3F	CCl_2F_2	CO	HF	SF_6	NO	NO_2	NO_x	HNO_3	HNO_4	N_2O_5	$ClONO_2$	NO_y	HCl	ClO	HOCl	BrO	OH	HO_2	CH_2O	CH_3CN	aerosol
ACE-FTS	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x						x		
Aura-MLS	x	x		x			x						x					x	x	x		x	x			
GOMOS	x										x															x
HALOE	x	x	x					x		x	x	x						x								x
HIRDLS	x				x	x					x		x													
LIMS	x	x									x		x													
MAESTRO	x																									
MIPAS	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x		x	x				x		
OSIRIS	x										x	x_d					x_m				x					x
POAM II	x										x															x
POAM III	x	x									x															x
SAGE I	x																									
SAGE II	x	x									x															x
SAGE III	x	x									x															x
SCIAMACHY	x	x									x	x_d									x					x
SMILES	x												x					x	x	x	x		x		x	
Odin/SMR	x	x		x			x			x			x				x_m		x				x_{lc}			
UARS-MLS	x	x											x													
TES	x_t																									

x available climatology
x_t used in UTLS comparisons only
x_d derived with help of a chemical box model
x_m merged and derived from OSIRIS NO_2 and Odin/SMR HNO_3 data
x_{lc} with limited coverage

on vertically resolved global satellite measurements, other important but temporally and spatially limited observations from aircraft missions, ground-based stations, or balloons also exist that would be of great value for future comparisons.

The report was prepared by the scientists of the SPARC Data Initiative Team with contributions from a number of internationally recognised instrument experts and data analysts, and underwent several rounds of extensive peer review and revisions. This Executive Summary outlines the overall approach as well as the key findings and recommendations obtained from the evaluations in this report.

Overall approach

- The SPARC Data Initiative has performed the up-to-date most comprehensive comparison of satellite instrument observations obtained from the CSA, ESA, JAXA, NASA, SNSB and other national agencies. In particular, it is the first systematic comparison between the older and younger generation, and the North American and European satellite instruments.
- The SPARC Data Initiative provides the most comprehensive set of monthly zonal mean time series of vertically resolved atmospheric trace gases and aerosol on a predefined latitude-pressure grid and in a common format easily useable by the atmospheric science community. The time series are available from the SPARC Data Center and will be updated in the future as soon as new data versions become available. The data products include 25 different chemical trace gases and aerosol from 18 different satellite limb sounders (see **Table ES.1**).
- The trace gas and aerosol time series have been evaluated by a common approach, comparing (single- or multi-year) annual or monthly climatologies derived from the monthly zonal mean fields, allowing for maximum overlap between different instruments. The comparison results do not change substantially when changing the number of years going into a climatology or, in case of the longer-lived species, when calculating instrument differences for a month instead of a year. From this, it follows that the comparisons shown yield relatively robust conclusions on instrument/retrieval performance.
- By evaluating monthly zonal mean averages, we take 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 for all instrument comparisons, avoids sensitivity to chosen limits defining coincident measurements, and produces generally larger sample sizes, which should in theory minimise the random sampling error. The climatological approach, however, has the disadvantage that climatological means may be biased due to non-uniformity of sampling. The extent to which the monthly and annual zonal mean climatologies are representative of the true mean has therefore been evaluated by investigating the impact of each instrument's sampling patterns on the climatologies. This yields information on the potential sampling bias of each instrument's climatology and is particularly useful to users examining variability and trends, or comparisons with free-running models.
- Evaluations focus on different regions, the upper troposphere (UT, from 300 hPa to the tropopause), the lower stratosphere (LS, from the tropopause to 30 hPa), the middle stratosphere (MS, 30-5 hPa), the upper stratosphere (US, 5-1 hPa), and the lower mesosphere (LM, 1-0.1 hPa), illustrating where the various data products are consistent and where they are not. Data products exhibiting unphysical features or strong deviations from the other instruments are highlighted in the report.

- In the SPARC Data Initiative report, the term climatology is not used to refer to a time-averaged climate state (which should be reproduced by free-running models, averaged over many years) but to refer to year-by-year values (which free-running models would not be expected to match). Based on different satellite instruments measuring in the same year (or over identical longer time periods), the spread in those measurements is regarded as representing the systematic uncertainty of our knowledge of the true atmospheric state in any given year (or over this period).
- The SPARC Data Initiative evaluations are based on the use of the multi-instrument mean (MIM) as a common point of reference (see **Box 1** in the *Introduction*). It must be emphasised that the MIM is not considered to provide a best estimate since it is impacted, among other things, by changes of the composition of instruments over time and unphysical behaviour of individual instruments. The MIM is not a data product and will not be provided.

Overall key findings

- The findings of the trace gas climatology intercomparison for both, short-lived and long-lived species, are generally consistent with the results of previous validation efforts (where available) based on the classical validation approach using profile coincidences. However, the uncertainty of the climatological mean values (as given by the standard error of the mean) is generally smaller than that of a single profile or a set of coincident profiles, yielding statistically more significant results on the differences or agreement between the instruments. In addition, the climatological approach yields more comprehensive information on measurement uncertainty in terms of latitude-pressure range covered. The comparisons of the climatologies thereby have in many cases improved our knowledge of the quality of the available data products.
- A large number of limb-viewing satellite instruments have observed the stratospheric composition over the last 30 years. A total of 18 instruments provide O_3 monthly zonal mean time series to the SPARC Data Initiative. A second group of atmospheric constituents consists of NO_2, H_2O, HNO_3 and aerosol, which have been measured by a smaller but still substantial number of instruments. For these gases and aerosol, between 7 and 12 climatologies are available for the SPARC Data Initiative comparisons, while for most other atmospheric trace gases less than 5 climatologies have been contributed.
- The report provides an estimate of the systematic uncertainty in our knowledge of the measured fields' mean state derived from the inter-instrument spread of ±1σ. This range is given for all trace gases where

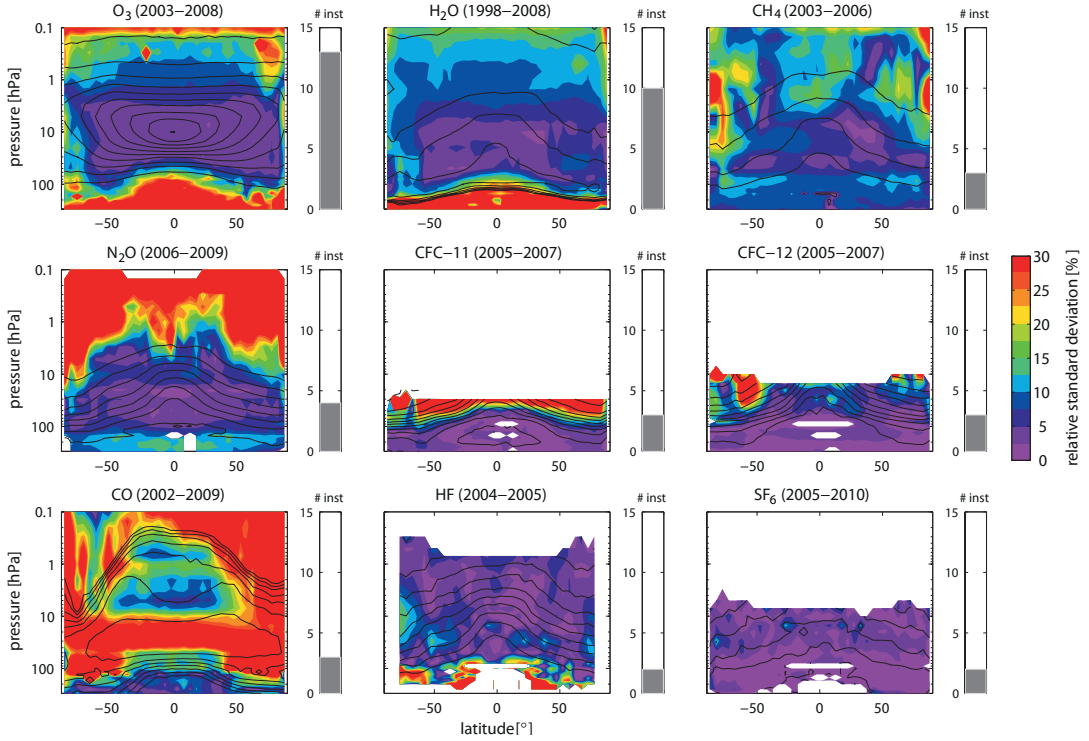


Figure ES.1: Synopsis of the uncertainty in the annual zonal mean state of longer-lived species. *The relative standard deviation over all instruments' multi-annual zonal mean datasets is presented for O_3, H_2O, CH_4, N_2O, CFC-11, CFC-12, CO, HF, and SF_6 (colour contours). The relative standard deviations are calculated by dividing the absolute standard deviations by the multi-instrument mean (MIM). The black contour lines in each panel represent the MIM trace gas distribution for each species. The number of instruments included is given by the right-hand grey bar, while detailed information on which instruments are included can be found in the respective trace gas sections. Note that the time periods used depend on the availability of the instruments included in the assessment and hence differ from trace gas to trace gas.*

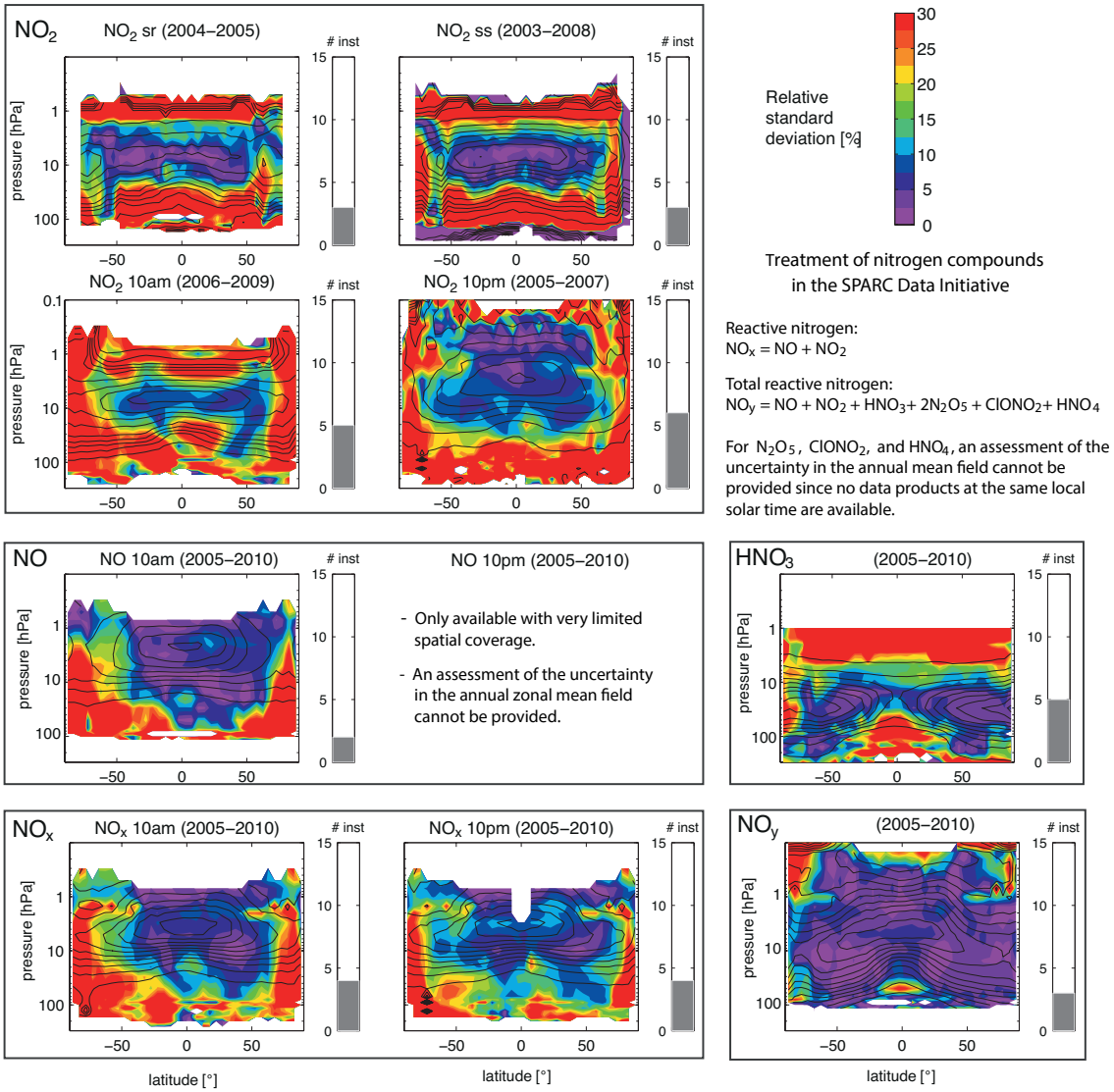


Figure ES.2: Synopsis of the uncertainty in the annual zonal mean state of nitrogen containing species. *As Figure ES.1, but for the nitrogen containing species. The assessment of the uncertainty in the annual mean state of NO, NO_x and NO_2 is based on climatologies corresponding to 10am and 10pm, and for the latter also on climatologies corresponding to local sunrise (sr) and local sunset (ss). Note that some of the included climatologies have been derived by scaling the individual measurements with a chemical box model to 10am/10pm local solar time (LST) (see individual chapters for detailed information).*

the available number of datasets allows for such an evaluation (but not aerosol) and is presented in the form of synopsis plots (**Figures ES.1**, **ES.2**, and **ES.3**). For a more detailed discussion of the individual trace gases see the *Summary by trace gas Section*.

- Agreement for the longer-lived trace gases O_3, H_2O, CH_4, N_2O, and HF is best in the tropical and mid-latitude MS and LS and worse towards the UTLS (in particular for O_3, H_2O, and HF), the US (N_2O and CH_4) and the LM (O_3 and H_2O) (**Figure ES.1**). In contrast, the trace gases CFC-11 (CCl_3F), CFC-12 (CCl_2F_2), and SF_6 show the best agreement in the UTLS and larger deviations in the MS. Climatologies of CO, which is a trace gas with an intermediate lifetime, are characterised by large relative differences throughout most of the measurement range. Nearly all trace gases show larger deviations in the polar regions than at lower latitudes.
- The agreement of the nitrogen species NO, NO_2, and HNO_3, as derived from the relative deviations between the climatologies, depends strongly on the atmospheric distribution of the respective gas with larger relative differences in regions of smaller mixing ratios (**Figure ES.2**). While NO and NO_x agree very well in the tropical and subtropical MS and US, NO_2 and HNO_3 have larger deviations in the US and show the best agreement in the tropical and mid-latitude MS and for HNO_3 also in the LS. All climatologies (except for HNO_3 and NO_y in the Northern Hemisphere (NH)) have considerably larger deviations in the polar regions. Finally, the NO_y climatologies show an excellent agreement throughout most of the measurement range except for the polar latitude LM.
- The nitrogen species $ClONO_2$, HNO_4, and N_2O_5 with mostly low atmospheric abundances and large diurnal

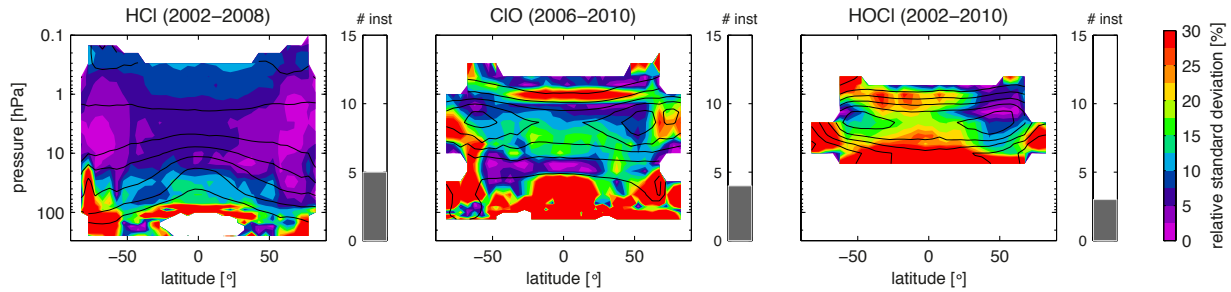


Figure ES.3: Synopsis of the uncertainty in the annual zonal mean state of chlorine containing species. *As Figure ES.1, but for the chlorine containing species. The assessment of the uncertainty in the annual mean state is based on ClO daytime climatologies and on HOCl night-time climatologies. Note that for ClO, climatologies from SMR are included which have been derived by scaling the individual measurements with a chemical box model to 1:30pm LST (see ClO chapter for detailed information).*

variations are measured by two satellite instruments only and are not included in the synopsis plots. Very good agreement is found for $ClONO_2$ in the MS where the diurnal cycle is not very pronounced. For HNO_4 and N_2O_5 the region of small diurnal variations, the MS and LS, respectively, coincides with low mixing ratios resulting in large relative differences.

- The agreement between climatologies of the bromine and chlorine compounds HCl, ClO, HOCl (see **Figure ES.3**), and BrO, depends strongly on the lifetime of the trace gas considered. The longer-lived HCl exhibits very good agreement and the daytime climatologies of the shorter-lived ClO show good to reasonable agreement in the MS and US where mixing ratios are highest. The short-lived HOCl shows mostly reasonable agreement in the US during night-time. Only little spatial overlap exists for BrO measurements, which show good agreement in the MS, but considerable disagreement in the LS.
- The short-lived species CH_2O, HO_2, OH, and CH_3CN are available from a small number of instruments only and are thus not included in the synopsis plots. For CH_2O, datasets overlap only in the LS, where mixing ratios are low and a large disagreement is found. HO_2 shows promising first results with mostly good agreement throughout the MS, US, and LM. OH and CH_3CN are each obtained from one instrument and shown for illustration purposes only to provide information on distribution and seasonality of these trace gases.
- The large deviations between the datasets of shorter-lived species stem partially from the difficulty of accounting for the strong diurnal cycles these trace gases exhibit. Scaling of the data to a common day/night-time using a chemical box model helped to improve the comparisons in some cases. However, it remains a challenge to estimate how much these deviations are related to errors introduced by the scaling procedures and how much of the deviations correspond to direct measurement differences.
- Comparisons of aerosol extinction values are complicated due to a strong wavelength dependency of the retrieved products. Evaluation of aerosol extinction products retrieved at similar wavelengths indicates variable agreement between them, with at least part of the biases attributable to differences in the wavelength. A new comparison approach based on scaling of different wavelength products towards a common standard reveals general consistency between the aerosol climatologies in terms of physical structures and, during periods of low aerosol loading, very good agreement also in terms of absolute values.
- Sampling bias, produced by the non-uniform sampling of temporally and spatially varying trace gas fields, is estimated through analysis of model fields. Climatologies from instruments with regular and uniform sampling patterns have generally small sampling bias. Climatologies from instruments whose latitudinal coverage varies with time can have strong sampling biases for certain months and locations. Monthly mean sampling biases for O_3 were found in some instances to be above 10%, primarily due to non-uniformity in day-of-month sampling. Throughout most of the stratosphere, sampling bias is much more important for O_3 than for H_2O, since the variability of O_3 is stronger.
- The knowledge of uncertainty and inter-instrument differences derived in the SPARC Data Initiative evaluations is used to improve model-measurement inter-comparisons. Particular diagnostics that are supported by a well-defined and small observational uncertainty range and would thus be especially appropriate for model evaluations are recommended. The suggested observational uncertainty range is derived from all available and suitable datasets instead of recommending one particular satellite dataset for the model-measurement comparison. The selection of the data points suitable for the construction of the uncertainty range is based on their agreement with the mean state of the atmosphere as given by all instruments and on the specific satellite characteristics such as sampling patterns and vertical resolution. For most diagnostics presented, the uncertainty range was reduced in comparison to values used in existing model evaluation efforts such as the CCMVal activity.

Overall recommendations

- Highlighted species for which further investigations are recommended:
  - H_2O and O_3 show particularly large uncertainties (with an inter-instrument spread of ±30%) in the UTLS, where satellite measurements can be affected by clouds and spatial smearing. At least part of the uncertainty could be reduced by accounting for geophysical variability in the comparisons.
  - For O_3, large inter-instrument differences are found at high latitudes (up to ±30%), which at least partially may be attributable to sampling issues. More detailed evaluations are required (especially for ozone hole conditions), including the use of coincident measurement comparisons, polar vortex coordinates and the incorporation of other correlative datasets (*e.g.*, in-situ measurements).
  - O_3 evaluations in the USLM are impacted by diurnal ozone variations, which may cause systematic biases between the instrument climatologies and need to be accounted for in future evaluation activities.
  - CO exhibits large differences in the annual zonal mean structure (±30% in the LS), an issue that should be addressed in forthcoming retrieval revisions.
  - The evaluations of some of the short-lived species can be improved by further efforts to scale the datasets to a common LST (*e.g.*, for HNO_4, N_2O_5, and $ClONO_2$ no detailed comparisons were possible due to the lack of scaled datasets) and to improve existing scaled datasets by removing outliers (*e.g.*, for NO).

- Highlighted species or regions for which more data are needed:
  - At present, there is a lack of correlative measurements that provide the necessary species to establish budgets of chemical families such as Br_y, and Cl_y. Current estimates of these families rely heavily on the use of chemical box models and hence independent validation is not possible.
  - The long-term monitoring of stratospheric HCl hinges on two instruments (ACE-FTS and Aura-MLS), which are both past their expected lifetimes. The abundance of HCl (scaled to yield Cl_y) in the stratosphere is an important measure to assess the effectiveness of the Montreal Protocol, which controls the use chlorine-containing substances that lead to the destruction of ozone. New HCl measurements are needed to be able to fulfill the obligations to the Montreal Protocol in the future.
  - There is a general need to improve trace gas observations (especially H_2O) in the tropical and extra-tropical UTLS. Higher spatial and temporal resolution measurements that also penetrate to lower altitudes (well into the upper troposphere) are required to overcome this issue.

- Instrument differences generally increase towards the UTLS, which is a critical region for chemistry-climate interactions. A detailed UTLS measurement intercomparison, using high spatial and temporal resolution measurements and diagnostic tools that minimise geophysical variability and differences in vertical resolution, is needed to fully characterise differences between satellite instruments. These comparisons should also be extended to include other correlative measurements such as from balloon, aircraft, or ground-based instrument platforms.
- A specific focus on instrument comparisons should also be carried out in the mesosphere where inter-instrument differences have been found to be large. Such comparisons will need to take into account differences in the datasets arising from tides and the diurnal cycle.
- The SPARC Data Initiative has provided data in an easily accessible format, basic knowledge on data quality, and some first model evaluation diagnostics. A mixed team of scientists from the SPARC Data Initiative and the model analysis community should further pursue on the generation of specific model evaluation diagnostics that include a 'best' estimate and its uncertainty range for ready use in model-measurement comparison exercises.
- Knowledge of bottom-up, absolute measurement uncertainty as identified in the SPARC Data Initiative should be improved. The uncertainties would need to include a range of error sources such as uncertainty in the spectroscopic data, calibration, pointing accuracy, and others. As much as possible the uncertainties would need to be derived consistently.
- The loss of expert knowledge on datasets from some past missions inhibited their careful assessment in this exercise. It is essential to maintain appropriate documentation, knowledge of data quality, and capacity to reprocess the data from historic measurements as necessary.
- Some evaluations suffer from too short overlap periods. Also, given the growing importance of long climate data records from multiple instruments, the lifetime of currently flying limb instruments should be extended as long as possible.
- In addition, the dearth of approved future limb satellite missions is likely to lead to a gap (chasm) in vertically resolved stratospheric and mesospheric trace gas measurements when the current missions end. There is therefore an urgent need to develop and fly new atmospheric limb sounder missions in order to remedy this situation.

Summary by trace gas

The overall findings for all trace gas species and aerosol are presented in the following summary with the systematic uncertainty in our knowledge of the atmospheric composition mean state shown in **Figures ES.1** (long-lived species), **ES.2** (nitrogen species), and **ES.3** (halogenated species). Additionally, the summary highlights similarities and differences between the individual datasets based on comparisons of zonal monthly mean latitude-height cross sections, seasonal cycles, and deseasonalised anomalies and other evaluations that test the physical consistency of the datasets (*e.g.*, tropical tape recorder, QBO).

Ozone (O_3)

A large number of satellite instruments have been measuring stratospheric ozone profiles over the past three decades. The comprehensive evaluation of vertically resolved monthly zonal mean ozone climatologies from the 18 limb-viewing 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, HIRDLS, and SMILES results in the following conclusions:

- The uncertainty in our knowledge of the ozone annual mean state (derived as the ±1σ inter-instrument spread) is smallest in the tropical MS and mid-latitude LS/MS. Nearly all instruments show very good agreement in those regions, with differences smaller than ±5%; some datasets even agree within ±2.5%.
- In the tropical UTLS, the spread between the datasets increases quickly with decreasing altitude, reaching ±30% at the tropical tropopause. In the mid-latitude UTLS, the various datasets show closer agreement, with a spread of ±10% at the tropopause. The poor agreement in the tropical UTLS is related to the small ozone abundances as well as instrumental limitations and demonstrates the need for further evaluation activities, including the use of existing in-situ measurements and nadir sounders.
- In the US, all datasets agree well, with deviations around ±10%. Identified inter-instrument deviations in the LM are not necessarily representative of real climatological differences due to the growing importance of the ozone diurnal cycle at altitudes above 1 hPa.
- At polar latitudes, the climatologies give a larger spread of the ozone mean state (±15%) compared to lower latitudes. Maximum variations (up to ±30%) are found during times of the ozone hole in the Antarctic LS, possibly related to the different sampling patterns of the individual instruments.
- Nearly all datasets show very good agreement in terms of interannual variability and are suitable 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. In best agreement with SAGE II are the datasets from Aura-MLS, OSIRIS, GOMOS (only in the MS/US) and MIPAS (not at altitudes above 10 hPa).
- To improve future model-measurement comparison activities, evaluations of natural variability presented here (seasonal cycle, interannual variability, and downward propagating QBO signal) are recommended. Depending on the application, individual instruments may need to be excluded from the comparison as demonstrated in Chapter 5.

Water vapour (H_2O)

In this report, we assessed the quality of 13 H_2O 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) that provide measurements within the time period from 1978 to 2010. SMR provides two data versions and we also treat MIPAS measurements before and after 2005 as two different datasets.

- Our knowledge of the atmospheric mean state of H_2O derived from the full set of instruments available between 1998 and 2008 is best in the LS and MS of the tropics and mid-latitudes, with a relative uncertainty (given by the ±1σ inter-instrument spread) of ± 2-6%.
- The relative uncertainty in the atmospheric mean state of H_2O increases toward the polar latitudes (± 10% and 15% for NH and the Southern Hemisphere (SH), respectively), the LM (±15%) and the UT (±30-50%).
- The H_2O minimum found just above the tropical tropopause shows annual zonal mean values ranging from approximately 2.5 to 4.5 ppmv, with a mean of 3.5 ± 0.5 ppmv (or ± 14%, 1σ uncertainty). The uncertainty is somewhat larger (± 15-20%) when looking at individual months. Our knowledge of the absolute amount of water vapour entering the stratosphere through the tropical tropopause as derived from satellite observations is hence relatively poor.
- Most sensors exhibit very good agreement on 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 for the recommendation of these data for use in studies of climate variability.
- Excellent agreement in interannual variability is typically observed between the older dataset from HALOE and the more recent ones from Aura-MLS, MIPAS(1) (MIPAS high spectral resolution measurements before 2005), MIPAS(2) (MIPAS high vertical resolution measurements from 2005 onwards) and ACE-FTS, indicating their potential 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. Also, HALOE shows consistently lower values (between -2.5 and -5%) than the MIM throughout the atmosphere, with increasing negative values when moving towards the tropical LS (-15%) and extratropical UTLS (values up to -50%) at altitudes below 100 hPa. The high quality of SAGE II data promises it to be a useful alternative to HALOE at these lower altitudes.
- Using a combination of the SPARC Data Initiative water vapour datasets shows great potential for improving past model-measurement comparisons based on the HALOE dataset only (*e.g.*, the seasonal cycle at the tropical tropopause or the tape recorder). However, careful choices have to be made when identifying the set of instruments for specific applications, depending on the region of the atmosphere.

Methane (CH_4)

CH_4 vertically resolved climatologies are compared from three instruments: HALOE, MIPAS, and ACE-FTS. The instruments overlap only in the year 2005.

- The uncertainty in our knowledge of the atmospheric CH_4 annual mean state is smallest in the LS and tropical/NH subtropical MS, with an inter-instrument spread of less than ±6%. The uncertainty is larger in the UT and lowermost stratosphere, with spread of around ±10%. The uncertainty increases also towards higher altitudes and latitudes, where relative uncertainties reach ±20% or more.
- HALOE shows consistently lower values than the MIM throughout the atmosphere and in fact shows lower values than all other instruments in the tropics and extratropical UTLS. MIPAS exhibits strong vertical oscillations around the MIM, which are approximately opposite between its high- and low-spectral resolution data version from before and after 2005. ACE-FTS shows features mostly consistent with the other instruments despite the strong impact of sampling (which results in somewhat noisy fields).

Nitrous oxide (N_2O)

N_2O vertically resolved climatologies are available from four instruments, SMR, MIPAS, ACE-FTS and Aura-MLS, with the earliest starting in 2001.

- The 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 extratropics, with inter-instrument spreads of less than ±4% and ±6%, respectively. Good knowledge is also obtained in the UT and extratropical LS at altitudes below 100 hPa, where the uncertainty is smaller than ±15%. The relative uncertainty increases moving towards the USLM (with values of more than ±50%).
- While the relative inter-instrument differences increase towards the USLM, the instruments show excellent agreement in terms of interannual variability.

Trichlorofluoromethane (CFC-11) and dichlorodifluoromethane (CFC-12)

CFC-11 and CFC-12 vertically resolved climatologies are available from three satellite instruments, MIPAS, ACE-FTS and HIRDLS, which overlap in 2005-2007.

- The uncertainty in our knowledge of the atmospheric CFC-11 annual mean state is small at altitudes below 100 hPa, with a relative uncertainty (given by the ±1σ inter-instrument spread) of less than ±5% in the tropics and mid-latitudes and less than ±10% at higher latitudes. In the tropical LS, the spread between the datasets increases quickly with increasing altitude to ±30% due to high ACE-FTS values. In the mid-latitudes LS, HIRDLS displays considerably lower values, and a large relative spread of up to ±50% exists.
- The uncertainty in our knowledge of the atmospheric CFC-12 annual mean state is very small at altitudes below 100 hPa, with a spread of less than ±5% and often even less than ±2.5%. In the LS, good agreement between all datasets exists in the tropics, in the NH, and in the SH subtropics (±10%). An exception to this good agreement are the SH extratropics, where considerable disagreement is found (±50%) between ACE-FTS and HIRDLS.
- Overall, there is better agreement of the CFC-12 climatologies than of the CFC-11 climatologies, in particular in the LS between 70 and 30 hPa. For CFC-12, largest discrepancies are found in the SH above 50 hPa resulting in pronounced discrepancies between the performance in the NH and SH extratropical regions.
- A large number of instrument-specific features can be observed for both trace gases. MIPAS CFC-11 and CFC-12 in the winter hemisphere have different meridional gradients at 200 hPa than the other two instruments. ACE-FTS has problems at its highest retrieval level in the tropics for both trace gases and shows in most regions no clear signals of seasonal cycle or interannual variability. HIRDLS climatologies of CFC-11 and CFC-12 show different gradients in the subtropics compared to the other instruments and large negative deviations in the mid-latitudes.

Carbon monoxide (CO)

The SPARC Data Initiative evaluated vertically resolved CO climatologies from four different instruments: SMR, MIPAS, ACE-FTS, and Aura-MLS.

- The uncertainty in our knowledge of the atmospheric CO annual mean state as derived from the four satellite instruments is smallest in the global UT, with an inter-instrument spread of less than ±6%. Good knowledge is obtained in the tropical MS, where the uncertainty is about ±10-15%. The uncertainty is largest in the global LS and also at high latitudes from the MS up to the LM (up to ±50%).

- The CO climatologies obtained from the four satellite instruments show large relative differences from the MIM, and do not agree on some key structures in the annual zonal mean distribution. Despite the shortcomings in reproducing the annual zonal mean distribution, the different datasets agree very well on deseasonalised anomalies in the tropical UTLS and MS.
- ACE-FTS and MIPAS show a very similar structure of the CO field; however, ACE-FTS exhibits consistently lower values than the MIM in all atmospheric regions. Aura-MLS exhibits an atypical isopleth structure in the LS, and SMR has values that are too high through most of the LS/MS.

Hydrogen fluoride (HF)

Vertically resolved HF climatologies are available from HALOE and ACE-FTS, which overlap in 2004-2005.

- The uncertainty in our knowledge of the atmospheric HF annual mean state as derived from the two satellite datasets is smallest at altitudes above 100 hPa, with a multi-instrument spread in this region of less than ±10% (±5% above 10 hPa). Larger deviations (±15%) are found in the SH high latitude MS mainly caused by the impact of the sampling bias on the annual mean datasets. At altitudes below 100 hPa, HF is less well determined, with a multi-instrument spread of ±30% or larger.
- ACE-FTS observes more HF than HALOE at altitudes above 50 hPa, while below 50 hPa HALOE detects more HF than ACE-FTS. For the 2-year long overlap period, both datasets agree roughly on the seasonal and interannual variability, with some differences found for month-to-month variations. Sampling issues are suspected to cause the noticable features found in the latitudinal structure of the HALOE and ACE-FTS annual mean cross sections.

Sulfur hexafluoride (SF_6)

Vertically resolved SF_6 climatologies are available from MIPAS and ACE-FTS, which overlap in 2005-2010.

- The differences between the two satellite datasets are overall very small (up to ±5% and at altitudes below 50 hPa up to ±2%), which implies a small uncertainty and good knowledge of the atmospheric mean state. The only exceptions are some individual grid points where the spread reaches values of ±12%. Note that ACE-FTS and MIPAS both measure SF_6 around the same spectral band, and it is therefore possible that the two datasets share systematic error components.
- MIPAS detects less SF_6 than ACE-FTS in most atmospheric regions, except for altitudes above 10 hPa and below 100 hPa in the SH. MIPAS SF_6 in the UTLS around 25°S/25°N shows some elevated mixing ratio peaks, which are most pronounced in the respective winter/spring hemisphere. ACE-FTS shows pronounced month-to-month variations, no clear seasonal cycle, as well as less steep and much noisier isopleths, likely as a result of its less dense sampling.

Nitrogen monoxide (NO)

The assessment of the atmospheric NO annual mean state is based on the climatologies from MIPAS, and ACE-FTS corresponding to 10am/10pm. Note that the latter have been derived by scaling the individual measurements with a chemical box model to 10am/10pm LST. Additionally, climatologies corresponding to local sunset/sunrise are available from the solar occultation instruments HALOE and ACE-FTS.

- The uncertainty in our knowledge of the atmospheric NO annual mean state, as estimated from the 10am MIPAS and scaled ACE-FTS climatologies, is smallest in the MS and US, with an inter-instrument spread of up to ±5% in the tropics, ±10% in the NH mid-latitudes, and ±20% in the SH mid-latitudes. In the LS, the two datasets agree very well in the tropics, but in the mid-latitudes deviations are larger (±20%).
- Due to the strong diurnal cycle with a near-zero NO abundance after sunset, the 10pm climatologies provide data only in the high summer latitudes and are not suitable for an assessment of the NO annual mean state.
- The local monthly mean sunset/sunrise climatologies from the solar occultation instruments HALOE and ACE-FTS agree well in the US (differences up to ±10%) but show a larger spread (up to ±50%) above and below this region. Their annual mean state for the overlap period 2004-2005 is strongly impacted by sampling and not suitable to derive information on the uncertainty in the annual mean field.

Nitrogen dioxide (NO_2)

Vertically resolved NO_2 solar occultation measurements are available from SAGE II, HALOE, POAM II, POAM III, SAGE III, and ACE-FTS and can be compared directly if separated into local sunrise and local sunset measurements. NO_2 measurements by limb emission and scattering techniques are available from LIMS, MIPAS, OSIRIS, SCIAMACHY, and HIRDLS, with the latter three scaled with a chemical box model to 10am/10pm LST in order to allow for a direct comparison of the different instruments. Additionally, ACE-FTS data scaled to 10am/10pm are available. GOMOS provides stellar occultation measurements at 10pm.

- The uncertainty in our knowledge of the atmospheric NO_2 annual mean state is estimated independently for local sunrise/sunset and 10am/10pm LST. The uncertainty is smallest in the tropical and mid-latitude MS, with an inter-instrument spread in this region of ±5% to ±10% (sunrise/sunset) and ±10% to ±20% (10am/10pm). In the LS, the NO_2 abundances decrease quickly, and for all climatologies a large spread (up to ±50%) exists. In the US, the best agreement is found for

the climatologies corresponding to 10 pm LST (±5% to ±10%). At high latitudes, the instruments show larger deviations (±50%) than at lower latitudes. Here, sampling issues during high-NO_x descent events may contribute substantially to the differences obtained.

- The solar occultation climatologies from SAGE II, HALOE and ACE-FTS show very good agreement in the MS, while above and below the differences increase steadily, reaching values of up to ±20% in the US and up to ±50% in the LS. For most regions the NO_2, sunrise and sunset evaluations give a consistent picture, however, some differences exist. All three solar occultation instruments, except for SAGE II sunrise data, display the tropical QBO signal. With the exception of the MS sunrise climatologies, SAGE II typically detects the largest NO_2 abundances, ACE-FTS resides in the middle range and HALOE is lowest.
- The 10am/10pm climatologies show good agreement in the MS, with mean differences of ±10%. In particular, MIPAS, GOMOS, OSIRIS and SCIAMACHY agree very well, with differences below ±5%. In the LS, overall mean differences can be as large as ±40% (GOMOS and scaled ACE-FTS), however, MIPAS, OSIRIS and SCIAMACHY are very close to each other (±5%) in most cases. All 10am/10pm climatologies show the tropical QBO signal, with the best agreement found between MIPAS, OSIRIS, SCIAMACHY and GOMOS. Interannual anomalies from GOMOS and HIRDLS are characterised by stronger month-to-month fluctuations than the ones from other instruments.
- In the tropical MS, scaled ACE-FTS agrees very well with the 10am/10pm climatologies and unscaled ACE-FTS agrees very well with the sunrise/sunset climatologies, with differences up to ±5%. If one were to assume no errors from the scaling, this agreement would suggest that all available measurements are consistent with each other in this region.
- Note that scaling with box model is problematic for data during night-time and users should not use scaled data that is based on unscaled values smaller than 0.5 ppbv.

Nitrogen oxides (NO_x)

The assessment of the atmospheric NO_x annual mean state is based on the climatologies from MIPAS, ACE-FTS, SCIAMACHY and OSIRIS corresponding to 10am and 10pm. For the latter three instruments, individual measurements have been scaled to 10am/10pm LST with the help of a chemical box model. Note that the OSIRIS and SCIAMACHY NO_x climatologies are compiled based on their NO_2 measurements and on NO profiles derived from a chemical box model. Additionally, climatologies corresponding to local sunset/sunrise are available from the solar occultation instruments HALOE and ACE-FTS.

- The uncertainty in our knowledge of the atmospheric NO_x annual mean state is smallest in the tropical and NH mid-latitude MS/US, with an inter-instrument spread in this region of up to ±10%. Although in the NH mid-latitude MS, the instruments agree very well (±5%), deviations increase in the SH mid-latitudes, in particular for the 10am climatologies (up to ±20%). In the LS, low NO_x abundances and large relative deviations (±30%) are found. In the high-latitude USLM, the sampling error is exacerbated by stronger gradients due to polar night NO_x descent causing an increase of the inter-instrument spread for altitudes above 1 hPa.
- The local sunrise/sunset climatologies from HALOE and ACE-FTS show excellent agreement in the US, with mean differences below ±2.5%. In the MS, HALOE detects slightly larger NO_x abundances than ACE-FTS (±5%), while in the LS, differences increase steadily (up to ±30%), with HALOE on the low and ACE-FTS on the high side. Overall, the NO_x local sunrise and sunset evaluations give a consistent picture, with the exception of the mid-latitude MS. Both solar occultation instruments display important signals of interannual variability like the tropical QBO cycle, but are characterised by stronger month-to-month fluctuations. Their annual mean state for the overlap period 2004-2005 is clearly impacted by sampling and not suitable to derive information on the uncertainty in the annual mean field.
- The climatologies corresponding to 10am/10pm show good agreement in the tropical and NH mid-latitude MS, with mean differences of ±5% to ±10%. In particular, the 10am climatologies from MIPAS, OSIRIS, and SCIAMACHY agree very well in the mid-latitude MS and US, with differences of less than ±5%. While scaled ACE-FTS agrees well with the other datasets in the tropical and NH mid-latitude MS, it is considerably lower in the SH mid-latitudes, with differences of up to -30%. This inconsistency between NH and SH mid-latitudes causes the larger inter-instrument spread in the latter region.

Nitric acid (HNO_3)

HNO_3 climatologies from ACE-FTS, Aura-MLS, HIRDLS, LIMS, MIPAS, SMILES, SMR, and UARS-MLS are evaluated as part of the SPARC Data Initiative.

- The uncertainty in our knowledge of the atmospheric HNO_3 annual mean state is smallest in the tropical MS and mid-latitude LS/MS, with an inter-instrument spread in this region of up to ±10%. In the tropical LS, mixing ratios are small and the relative differences reach ±50%. In the US, the situation is similar and additionally further complicated by the growing importance of the diurnal variations in HNO_3.
- At high SH latitudes, a large spread between the annual mean climatologies of up to ±30% exist. Deviations of the individual datasets are often of opposite sign when compared to lower and northern latitudes and are most pronounced during times of the ozone hole. Further evaluations of high-latitude HNO_3 need to include the use of coincident measurements and polar vortex coordinates, since the monthly mean comparisons can be impacted by the sampling patterns of the instruments.

- In the tropical MS, the climatologies from Aura-MLS, HIRDLS and SMR agree well on the positive side, while ACE-FTS and MIPAS show good agreement on the negative side of the mean over all datasets. In the mid-latitude LS/MS, the datasets from MIPAS, HIRDLS and SMR are very close to each other.

Peroxynitric acid (HNO_4), dinitrogen pentoxide (N_2O_5), chlorine nitrate ($ClONO_2$)

The nitrogen species HNO_4, N_2O_5, and $ClONO_2$ are part of the reactive nitrogen family and exhibit large diurnal variations in most parts of the atmosphere. All three species are measured at local sunrise/sunset by ACE-FTS and around 10am/10pm LST by MIPAS. The SPARC Data Initiative report presents the annual and monthly zonal mean state of HNO_4, N_2O_5 and $ClONO_2$ at local sunrise/sunset and 10am/10pm.

- Quantitative comparison of the ACE-FTS and MIPAS climatologies would require scaling one to the LST of the other. However, such scaling of HNO_4, N_2O_5 and $ClONO_2$ climatologies has not been performed and therefore the instrument comparisons focus on regions with a small diurnal cycle. Very good agreement is found for $ClONO_2$ in the MS where the diurnal cycle is not very pronounced. For HNO_4 and N_2O_5 the region of small diurnal variations, the MS and LS, respectively, coincides with low mixing ratios resulting in large relative differences.
- The evaluation of the sum of the three gases is implicitly included in the comparison of NO_y. In regions where the three gases contribute more than a negligible fraction to the reactive nitrogen family, good agreement between MIPAS and ACE-FTS NO_y, NO_x, and HNO_3 suggests that instrument differences for HNO_4, N_2O_5 and $ClONO_2$ are small.

Total reactive nitrogen (NO_y)

The assessment of the atmospheric NO_y annual mean state is based on the climatologies from ACE-FTS, MIPAS and Odin. The ACE-FTS and MIPAS climatologies include NO, NO_2, HNO_3, $ClONO_2$, HNO_4 and $2\times N_2O_5$. The Odin climatology is based on NO_2 from OSIRIS, HNO_3 from SMR and NO, $2\times N_2O_5$ and $ClONO_2$ taken from chemical box model simulations, while HNO_4 is not included.

- The three NO_y climatologies show very good agreement over large parts of the lower to upper stratosphere and mesosphere, with an inter-instrument spread of around ±5%.
- Exceptions to the overall good agreement are the tropical LS (±30%), with Odin on the high and ACE-FTS on the low side, and the high-latitude LM (±50%), with MIPAS on the high and ACE-FTS on the low side.

Hydrogen chloride (HCl)

The uncertainty in our knowledge of the atmospheric HCl annual mean state is derived from four satellite instruments including HALOE, ACE-FTS, Aura-MLS, and SMILES.

- The HCl climatologies from HALOE, ACE-FTS, Aura-MLS, and SMILES agree generally well with each other. The multi-instrument spread is smallest in the MS and US, and smaller in the polar regions (±4%) than in the tropics (±8%). Good knowledge is obtained in the LM and tropical LS, where the uncertainty is about ±10-15%. The uncertainty is largest in the SH polar vortex region and the UTLS (reaching more than ±50%). The uncertainty in these regions may be explained by the relatively small HCl abundances, in addition to potential sampling biases.
- The HCl climatologies from Aura-MLS and ACE-FTS agree within ±2.5% through most of the stratosphere and up to the LM. Slightly higher deviations from the MIM are found in the UTLS and at the edges of the polar vortices, where sampling bias may play a crucial role in determining a climatology correctly.
- Note that while HALOE and SMILES are on the low side of the MIM, a newer version of SMILES HCl exhibits values that are more consistent with Aura-MLS and ACE-FTS, indicating a low-bias in the HALOE HCl product.

Chlorine monoxide (ClO), hypochlorous acid (HOCl), and bromine oxide (BrO)

The halogenated species ClO, HOCl, and BrO exhibit large diurnal variations in most parts of the atmosphere. ClO and HOCl are both measured by Aura-MLS (at about 1:30am/pm), MIPAS (10am/pm; HOCl is restricted to the high-spectral measurement mode), and SMILES (resolving the full diurnal cycle). In addition, ClO is measured by SMR at 6:30am/pm and climatologies are also available scaled to 1:30am/pm. For ClO (HOCl), daytime (night-time) measurements are evaluated in the SPARC Data Initiative, since diurnal variations are smaller during the day (night) for the respective species. BrO climatologies from OSIRIS (at 6:30am/pm and scaled to 10am), SCIAMACHY (10am equator-crossing time and scaled to 10am), and SMILES (full diurnal cycle, with two products from different measurement bands) are evaluated as part of the SPARC Data Initiative.

- For ClO, Aura-MLS and SMILES show the most consistent results with differences relative to the MIM of ±10%. SMR (scaled and unscaled) is on the low side of the other instruments, with the scaled product showing differences mostly between -5% and -10% to the MIM, except in the autumn/winter hemisphere where differences increase. MIPAS (in the high-spectral measurement mode), on the other hand tends to lie on the high side, and MIPAS (in the high-vertical measurement mode) on the low side of the MIM.

- For HOCl, MIPAS lies on the high side, Aura-MLS on the low side, and SMILES in the middle of the three instruments with differences relative to the MIM of ±20% at pressures < 10 hPa. At pressures >10 hPa, SMILES looses sensitivity and shows differences increasing to +50% relative to the MIM.
- For BrO, OSIRIS and SCIAMACHY show good agreement (±10%) in the MS, however with increasing differences towards the high latitudes and down into the LS. SMILES overlaps with OSIRIS in a very limited altitude range only (5-10 hPa), where its band A measurements show good to very good agreement with OSIRIS.

Hydroxyl radical (OH), hydroperoxy radical (HO_2), formaldehyde (CH_2O), and acetonitrile (CH_3CN)

The SPARC Data Initiative evaluated HO_2 (short lifetime) from Aura-MLS, SMILES and SMR using daytime climatologies (when the impact of its diurnal cycle is smallest), and CH_2O (intermediate to short lifetime) from MIPAS and ACE-FTS using daily climatologies. In addition, mean distributions and their seasonal evolution are shown for OH from Aura-MLS and CH_3CN from SMILES to provide insight into the behaviour of these not very well known, minor species.

- HO_2 shows excellent to very good agreement between Aura-MLS and SMILES from the MS up to the LM. SMR climatologies (available for 2003 and 2004 only, unscaled) show very large negative biases, which could be better constrained by using a scaled product for comparison.
- CH_2O climatologies from MIPAS and ACE-FTS show very large differences in the limited region of overlap (UTLS). MIPAS measurements exhibit a very low vertical resolution, which may partially explain the differences between the two instruments.

Aerosol

Aerosol extinction climatologies from 8 satellite instruments (SAGE II, HALOE, POAM II, POAM III, OSIRIS, SAGE III, SCIAMACHY, and GOMOS) are evaluated within the SPARC Data Initiative. The instruments offer a total of 34 products, all retrieved at different wavelengths ranging between 350 and 5260 nm. Two alternative evaluation approaches were used to deal with the wavelength-dependency of the products.

- Comparison of aerosol extinction products at similar wavelengths show differences that are at least partially expected from the wavelength dependency. In some cases, indications for real differences could be revealed, including a high bias of the POAM II product at 603 nm, a low bias in AERGOM at 750 nm, and large differences of SCIAMACHY and AERGOM climatologies at 470 nm relative to the MIM. On the other hand, OSIRIS and SCIAMACHY at 750 nm agree well with each other throughout the LS and MS (±10%; except at higher latitudes), and OSIRIS agrees very well with SAGE III at 755 nm (±5%).
- Comparison of normalised products (using a scaling factor derived from a period with low aerosol loading) shows that most of the aerosol extinction products capture the physical structure of the evolving aerosol layer well. Exceptions are HALOE at 2450 and 5260 nm (exhibiting larger differences), POAM II at 352 nm (exhibiting the wrong structure), and AERGOM at 350, 600, and 750 nm (exhibiting more noise). The evaluations indicate generally very good agreement (±5-10%) between most aerosol products in the MS, and good to reasonable agreement (±10-20%) in the LS during a time period with relatively low aerosol loading. The differences between the products increase during time periods with higher aerosol loading, indicating sensitivities towards the assumptions on aerosol size distributions used in retrievals.

Ozone (O_3) evaluation in the UTLS using TES averaging kernels

A particular case study comparing the limb-viewing instruments with the nadir sounder TES participating in the SPARC Data Initiative has been carried out in order to cross-validate ozone distributions in the UTLS with an independent dataset that is frequently used for the evaluation of ozone in tropospheric models.

- Comparing climatologies from nadir- and limb-viewing instruments requires accounting for large differences in vertical resolution between the two types of instruments and is complicated by the fact that nadir-viewing UTLS retrievals can have a significant contribution from the troposphere and large altitudinal and latitudinal variations in sensitivity.
- In order to account for the differences, observations of the higher vertical resolution limb sounders have been smoothed using the observational operator of TES and then compared using the same evaluation diagnostics as used for the trace gas evaluations discussed above.
- There are large relative differences among the climatologies in both mean ozone abundance and in the timing and magnitude of ozone temporal variability in the tropical UTLS. Most of the limb sounders show large positive biases in this region of up to 50% when compared to TES. The climatologies are in much better agreement in the mid-latitude UTLS.

List of abbreviations

ACE-FTS	Atmospheric Chemistry Experiment - Fourier Transform Spectrometer
ACE-MAESTRO	Atmospheric Chemistry Experiment - Measurement of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation
ADEOS	Advanced Earth Observing Satellite
AERGOM	Aerosol retrieval from GOMOS satellite observations
ATLAS	Atmospheric Laboratory for Applications and Science
ATMOS	Atmospheric Trace Molecule Spectroscopy
CCM	Chemistry-Climate Model
CCMVal	Chemistry-Climate Model Validation Activity
CLAES	Cryogenic Limb Array Etalon Spectrometer
CMAM	Canadian Middle Atmosphere Model
CSA	Canadian Space Agency
ECMWF	European Centre for Medium-Range Weather Forecasts
Envisat	European environmental satellite
EOS	Earth Observing System
EPP	Energetic Particle Precipitation
ESA	European Space Agency
FOV	Field-Of-View
GEOS	Goddard Earth Observing System
GOMOS	Global Ozone Monitoring by Occultation of Stars
GOZCARDS	Global OZone Chemistry And Related trace gas Data records for the Stratosphere
HALOE	Halogen Occultation Experiment
HIRDLS	High Resolution Dynamics Limb Sounder
IGACO	Integrated Global Atmospheric Chemistry Observations
ILAS	Improved Limb Absorption Spectrometer
IO_3C	International Ozone Commission
ISAMS	Improved Stratospheric and Mesospheric Sounder
ISSI	International Space Science Institute
JAXA	Japan Aerospace EXploration Agency
JPL	Jet Propulsion Laboratory
LIDAR	LIght Detection And Ranging
LIMS	Limb Infrared Monitor of the Stratosphere
LM	Lower Mesosphere
LOS	Line-Of-Sight
LS	Lower Stratosphere

LST	Local Solar Time
MAD	Median Absolute Deviation
MAS	Microwave Atmospheric Sounder
MAX	Maximum
MEaSUREs	Making Earth Science data records for Use in Research Environments
MIM	Multi-Instrument Mean
MIN	Minimum
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLS	Microwave Limb Sounder
MS	Middle Stratosphere
MSIS	Mass-Spectrometer-Incoherent-Scatter model
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NDACC	Network for Detection of Atmospheric Composition Change
NH	Northern Hemisphere
O_3-CCI	Ozone - Climate Change Initiative
OSIRIS	Optical Spectrograph and InfraRed Imager System
POAM	Polar Ozone and Aerosol Measurement
PSC	Polar Stratospheric Clouds
QBO	Quasi-Biennial Oscillation
SAGE	Stratospheric Aerosol and Gas Experiment
SAO	Semi-Annual Oscillation
SAMS	Stratospheric and Mesospheric Sounder
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SD	Standard Deviation
SEM	Standard Error of the Mean
SH	Southern Hemisphere
SI^2N	SPARC, IO_3C, IGACO-O_3/UV and NDACC Ozone Profile Phase II Initiative
SMILES	Superconducting sub-MIllimeter-wave Limb-Emission Sounder
SMR	Sub-Millimetre Radiometer
SNSB	Swedish National Space Board
SPARC	Stratosphere-troposphere Processes And their Role in Climate
SR	SunRise
SS	SunSet
SSiRC	Stratospheric Sulfur and its Role in Climate
SZA	Solar Zenith Angle
TES	Tropospheric Emission Spectrometer
UARS	Upper Atmosphere Research Satellite
UKMO	United Kingdom Meteorological Office
US	Upper Stratosphere
UT	Upper Troposphere

VMR	Volume Mixing Ratio
WACCM	Whole Atmosphere Community Climate Model
WAVAS II	Water Vapour Assessment phase 2
WCRP	World Climate Research Programme

Table of contents

see Part 2

see Part 2

see Part 2

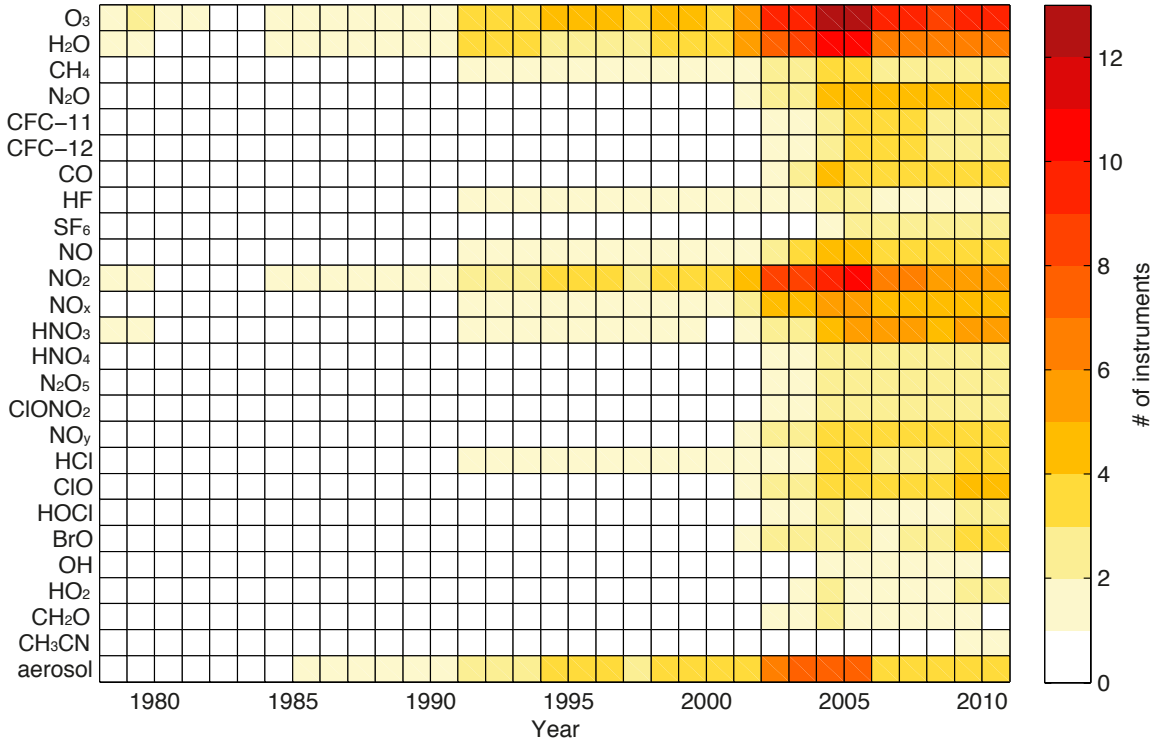
Chapter 1: Introduction

The past 30 years have been a 'golden age' for satellite measurements and have provided a wealth of knowledge regarding chemical trace gas abundances in the stratosphere. There is a danger that in the future the stratosphere will not be as well measured and it is therefore important to capture existing knowledge of current and recent instruments, retrievals and datasets before this knowledge is lost.

Satellite instruments from CSA, ESA, JAXA, NASA, SNSB, and other national space agencies provide a large number of trace gas datasets, which differ in terms of measurement method, geographical and seasonal coverage, spatial and temporal sampling and resolution, time period, and retrieval technique. These datasets of chemical trace gases are widely used for empirical studies of stratospheric climate, trends, and variability, and for the evaluation of the representation of transport and chemistry in numerical models. However, the validity of such studies strongly depends on the quality and representativeness of the datasets used, and it is often difficult for a user to determine which is the most reliable or useful dataset for a particular application. Hence, it is essential that the characteristics of the datasets be known prior to their use and prior to the interpretation of results. For example, comparing numerical model output to different chemical datasets can lead to conflicting results, which limits the value of model-measurement intercomparison studies.

Issues arising when using observational datasets for model evaluations have been identified in the SPARC CCMVal report [*SPARC*, 2010], which undertook a comprehensive assessment of model performance in the stratosphere. The report's recommendations directly motivated the work for the SPARC Data Initiative. The recommendations included: (1) *'Long-term vertically resolved datasets of constituent observations in the stratosphere are required to assess model behaviour and test model predictions. This includes ozone, but also other species that can be used to diagnose transport and chemistry. The current set of GCOS* [note at the time of writing] *Essential Climate Variables is not sufficient for process-oriented evaluation of CCMs.' (2) 'More global vertically resolved observations are required, particularly in the UTLS. As CCMs evolve towards including tropospheric chemistry, lack of observations in this region will become a major limitation on model evaluation.' (3) 'A systematic comparison of existing observations is required in order to underpin future model evaluation efforts, by providing more accurate assessments of measurement uncertainty.'*

Table 1.1: Number of instruments within the SPARC Data Initiative measuring a particular chemical trace gas species or aerosol in a given year.



There is also a strong need to characterise instrument differences as a prelude to data merging activities. These activities aim to merge various data sources into homogeneous climate data records suitable for trend studies, evaluation diagnostics, or climate forcings in global climate models. Merging of data for such purposes is only meaningful if differences between datasets are systematic and consistent.

Finally, the atmospheric trace gas datasets are not always available in a standard form, or with appropriate documentation. To enable the best possible use of the satellite datasets it is important to provide easy access to the datasets in a common format as well as to the information on the different instrument techniques and retrieval procedures.

The SPARC Data Initiative helps to address these issues by having performed the first comprehensive multi-instrument comparison of stratospheric chemical trace gas climatologies. It thereby provides a user guide to the different datasets, along with easy access to the data in a common format, and recommends future studies that would enhance the quality and usefulness of the existing data. In order to attain these goals, the SPARC Data Initiative assessed, *in a first step*, the current availability of vertically resolved, chemical trace gas and aerosol datasets from a suite of multi-national space-based instruments. *In a second step*, chemical trace gas and aerosol monthly zonal mean time series were compiled in a common and simple-to-use NetCDF data format. *In a third step*, these trace gas time series underwent detailed comparisons, which identified strengths and shortcomings of all datasets and differences between them. Where possible, an expert judgment on the source of those differences is provided.

Assessment of trace gas availability: Middle atmospheric trace gas observations are available from an international suite of satellite limb sounders, with the first measurements starting in 1979. Some of the instruments launched after 2000 are presently still taking regular measurements, despite being already past their expected lifetimes. All instruments have been measuring different sets of chemical species depending on the measurement technique applied. Earlier instruments were mostly based on the solar occultation technique, measured in the UV/VIS range and focused on ozone, water vapour and some nitrogen species. Instruments launched after 2000 were more often scattering and emission sounders, the latter extending measurements into microwave and sub-mm wavelengths, and covered a wider range of measured species. For each trace gas the number of satellite datasets within the SPARC Data Initiative is given as a function of time in **Table 1.1**.

Compilation of zonal monthly mean time series: The observational datasets have been compiled into a common data format, which is easy to handle by data users. To this end, zonal monthly mean time series of each trace gas species (in volume mixing ratio, VMR) and aerosol (as extinction ratio) have been calculated for each instrument on the SPARC Data Initiative climatology grid, using 5 degree latitude bins (with mid-points at 87.5°S, 82.5°S, 77.5°S, ..., 87.5°N) and 28 pressure levels (300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa) corresponding to the CCMVal pressure levels. The data therefore encompass the atmospheric region from the upper troposphere up to the lower mesosphere. Along with the monthly zonal mean value, the standard deviation and the number of averaged data values are given for each month, latitude bin and pressure level. Furthermore, the mean, minimum, and maximum local solar time (LST), the average day of the month, and the average latitude of the data within each bin for one selected pressure level are provided.

Evaluation diagnostics: In contrast to traditional data evaluation techniques based on coincident profiles, the SPARC Data Initiative compares climatologies in order to reduce geophysical variability and to obtain an assessment of our knowledge of the mean atmospheric state. Different standard evaluation diagnostics are used, such as single- or multi-year annual or monthly mean climatologies, vertical and meridional profiles, and seasonal cycles. In addition, time-latitude or time-altitude evolutions are assessed in order to test the physical consistency of the datasets. These include the tropical tape recorder in water vapour, polar dehydration, polar ozone loss, or the Quasi-Biennial Oscillation (QBO). The general approach taken is to compare the instruments to the multi-instrument mean, as explained in **Box 1**.

The notations for different atmospheric and geographical regions that are being used throughout this report are listed in **Tables 1.2** and **1.3**, respectively. **Table 1.4** defines the naming convention for the level of agreement between the instruments used in this report.

Table 1.2: Definitions and abbreviations of different atmospheric regions referred to in the report. *Note that the notations UTLS and USLM refer accordingly to the total extent of the sub-regions (i.e., 300-30 hPa and 5-0.1 hPa).*

Region	Abbreviation	Lower boundary	Upper boundary
Upper Troposphere	UT	300 hPa	Tropopause
Lower Stratosphere	LS	Tropopause	30 hPa
Middle Stratosphere	MS	30 hPa	5 hPa
Upper Stratosphere	US	5 hPa	1 hPa
Lower Mesosphere	LM	1	0.1 hPa

Table 1.3: Definitions of different geographical regions referred to in the report.

Region	Latitude range
Tropics	30°S-30°N
Subtropics	20°S-40°S and 20°N-40°N
Mid-latitudes	30°S-60°S and 30°N-60°N
High/polar latitudes	60°S-90°S and 60°N-90°N

Table 1.4: Definition of levels of agreement between a given climatology and the multi-instrument mean.

%-differences	Level of agreement
Up to ± 2.5%	Excellent agreement
Up to ± 5%	Very good agreement
Up to ± 10%	Good agreement
Up to ± 20%	Reasonable agreement
Up to ± 50%	Considerable disagreement
Up to ± 100%	Large disagreement

An approximate measure of random uncertainty in each climatological mean is the standard error of the mean (SEM); calculated from n measurements and a standard deviation, SD, as SEM=SD/√n. Due to its ease of computation and frequent use in past studies and despite its shortcomings (see *Chapter 3* for details), the SEM will be used as an approximate measure of uncertainty in each climatological mean, graphically illustrated by error bars of ±SEM, which can be loosely interpreted as a 68% confidence interval of the mean.

The analysis of O_3, aerosol and H_2O climatologies in the report is intended to support other ongoing SPARC activities focused on characterising long-term changes such as WAVAS II (for H_2O), SI^{2}N (for O_3), and SSiRC (for aerosol), and also to provide valuable information on data quality to "data merging" activities currently being carried out by NASA and ESA.

The zonal mean climatologies of the different chemical trace gas and aerosol products that were compiled during the SPARC Data Initiative can be downloaded from the SPARC Data Centre website (**http://www.sparc-climate.org/data-centre/**). In general, the results of this report depend on the specific level-2 data versions on which the climatologies are based, and future data versions might give different results. The goal is to provide updated climatologies whenever new data versions become available. The improvements achieved in moving to the next data version will be explained in meta-data or references provided. Interested users of the SPARC Data Initiative climatologies are asked to follow the data policy instructions posted in the same directory.

The report is structured as follows. *Chapter 2* comprises detailed information on the instruments participating in the SPARC Data Initiative, including measurement techniques and retrieval descriptions. *Chapter 3* gives an overview of the methodology used by the SPARC Data Initiative to create the climatologies and the approach used to evaluate them. *Chapter 4* features all comparisons of the chemical trace gases and aerosol, while *Chapter 5* summarises some general interpretation and higher-level conclusions of the results.

Box 1: *Multi-Instrument Mean Reference*

The approach of the SPARC Data Initiative is to use the multi-instrument mean (**MIM**) as a common point of reference. The choice of the MIM is by no means based on the assumption that it is the best estimate of the atmospheric trace gas field, but is motivated by the need for a reference that does not favor a certain instrument. It should be stated that the MIM *is not* a data product and *is not* provided as part of the SPARC Data Initiative datasets.

The MIM is calculated by taking the mean of all available instrument climatologies within a given time period of interest. The time periods can vary for the different trace gases and are chosen to ensure maximum spatial and temporal data coverage for each instrument and to limit the impact of sampling bias. In general, all available instrument datasets are included in the MIM regardless of their quality and without any weighting applied to them. Only if measurements from a particular instrument are deemed completely unrealistic, or if the same instrument is providing two versions of a specific trace gas data product, are they not included in the MIM.

The SPARC Data Initiative evaluations are based on relative differences between the trace gas mixing ratios of an instrument (X_i) and the MIM (X_{MIM}) given by:

$$\mathrm{diff}[\%] = 100 * (X_i - X_{MIM}) / X_{MIM}$$

One has to keep in mind when interpreting relative differences with respect to the MIM that the composition of instruments from which the MIM was calculated may have changed between time periods. Hence, changes in derived differences are not to be interpreted as changes in the performance (or drifts) of an individual instrument. Also, if there is unphysical behaviour in one instrument, the MIM and thus the differences of the other instruments with respect to the MIM will most certainly reflect this unphysical behaviour as well. Finally, if one instrument does not have global coverage for every month some sampling biases may be introduced into the MIM. A detailed assessment of the uncertainty introduced due to inhomogeneous temporal or spatial sampling in the SPARC Data Initiative climatologies is provided in *Chapter 3*.

Chapter 2: Satellite instruments

Satellite remote sensors are instruments designed to obtain information on the atmospheric composition through the analysis of data acquired without direct contact with the atmosphere. While remote sensors can also be employed from the ground, balloon or aircraft, on satellites they provide a unique global view with a more comprehensive geographical coverage and regular observations. Satellite instruments can offer total column or height-resolved measurements. For this purpose, satellite instruments take advantage of different interactions of radiation with the atmosphere (*e.g.*, absorption, emission or scattering) and detect wavelengths throughout the electromagnetic spectrum. Disadvantages of satellite instruments are that they are often expensive, can be high risk, require complex space-qualified instrumentation, and have limited lifetimes.

In this chapter, *Section 2.1* presents a general discussion of the satellite measurement techniques and orbit types relevant for the instruments participating in the SPARC Data Initiative. More detailed descriptions of the specific instruments, including information on retrieval processes, are given in *Section 2.2*.

2.1 Satellite measurement techniques

The satellite instruments participating in the SPARC Data Initiative are all passive sensors. Passive sensors detect natural radiation emitted from an external source (*i.e.,* the sun or stars) or by the atmosphere itself. Active sensors, on the other hand, emit high-energy radiation themselves and detect what is reflected back from the atmosphere (*e.g.*, LIDARs). In this section, general characteristics of various passive remote sensing techniques are described in terms of measurement geometry and wavelength coverage, however the scope is limited to concepts relevant to this study.

Table 2.1 provides a classification of the instruments participating in the SPARC Data Initiative according to both categories (observation geometry and wavelengths), which are explained in more detail in *Sections 2.1.1* and *2.1.2,* respectively.

2.1.1 Classification by observation geometry

Satellite instruments can be classified according to their observation geometry into limb-viewing or nadir-viewing sounders. Limb sounders look tangentially through the atmosphere, while nadir sounders have a downward-viewing observation geometry, pointing towards the Earth's surface. Limb geometries are the natural choice for stratospheric measurements because the signal is not masked by the denser tropospheric signal, the long ray-path through the atmosphere provides large sensitivity to species with low atmospheric concentrations, and the variation of the observation angle allows vertical scanning of the atmosphere. As a result, altitude information on the observed atmospheric state variables can be obtained at high vertical resolution, while the horizontal resolution is limited. For tropospheric observations, limb measurements are more challenging because of the saturation of measured radiances and the opaqueness of the troposphere due to the presence of clouds, humidity, and generally larger density. For many aspects of tropospheric sounding, nadir sounders are advantageous, due to their small horizontal footprint.

In the following, limb-viewing sounders are further classified according to their measurement modes, which are based on emission, scattering, solar occultation, and stellar occultation. In parts of the satellite observation community the term 'limb sounding' is reserved for limb emission and limb scattering measurements, but here the term is used in a more general sense, including the occultation geometry. A description of the nadir emission technique is also provided.

Limb emission

Emission measurements in limb geometry record the signal that is emitted along a horizontal path through the atmosphere and is partly absorbed on its way between the emitting air parcel and the observer (see **Figure 2.1**). Variation of the elevation angle of the line-of-sight (LOS) allows altitude-resolved temperature and composition measurements from approximately cloud-top height to the thermosphere. In turn, the horizontal resolution is limited to ~300 km unless corrections for LOS gradients are applied, or tomography is used. Since the Planck function at terrestrial temperatures is very low for wavelengths shorter than about 2.5 μm, limb emission measurements are, at least under conditions of local thermodynamic equilibrium, feasible only at wavelengths larger than this threshold, *i.e.,* in the mid-infrared to the microwave spectral region. At these wavelengths, atmospheric scattering is negligible except for clouds and large aerosol particles. Since, in contrast to occultation measurements (see

Table 2.1: Instruments classified according to observation geometry and wavelength categories. *Only instruments participating in the SPARC Data Initiative, and the measurement modes considered, are listed.*

	Microwave / Sub-mm 100 µm - 10 cm	Mid-IR 2.5 - 20 µm	Near-IR 0.8 - 2.5 µm	VIS / UV < 0.8 µm
Limb emission	UARS-MLS Aura-MLS SMR SMILES	MIPAS HIRDLS LIMS		
Solar occultation		ACE-FTS HALOE	POAM II/III SAGE I/II/III	POAM II/III SAGE I/II/III ACE-MAESTRO
Stellar occultation				GOMOS
Limb scattering			SCIAMACHY	SCIAMACHY OSIRIS
Nadir emission		TES		

below), no direct illumination source is needed, emission measurements can be obtained during both day and night. Depending on the orbit of the platform, measurements can be performed globally with dense spatial coverage, and the azimuth angle can be arbitrarily chosen as long as the Sun is avoided. A disadvantage of the emission technique compared to occultation measurements is the relatively small signal to noise ratio, which is caused by the faint signal of atmospheric emission. Calibration and determination of the exact elevation angle of the LOSs are crucial to avoid propagation of related errors onto the retrieved trace gas abundance profiles. Within the SPARC Data Initiative, the limb emission technique is used by Aura-MLS, HIRDLS, LIMS, MIPAS, SMILES, SMR, and UARS-MLS.

Solar occultation

Solar occultation instruments record radiance emitted by the Sun and attenuated along a horizontal ray-path through the atmosphere by extinction, *i.e.,* absorption and scattering (see **Figure 2.2**). Similar to the limb emission measurements, altitude-resolved information is obtained by variation of the elevation angle of the LOS. However, in contrast to limb emission where the measurement geometry can be freely chosen, the geometry is defined by the position of the Sun with respect to that of the satellite and the Earth. Measurements in occultation geometry can only be performed during the sunrise and sunset as seen from the satellite, *i.e.,* two times per orbit, which results in a limited global coverage and greatly reduced data density (compared to an emission sounder). On the other hand, the Sun provides a large radiance signal, allowing highly precise measurements even at shorter wavelengths. Occultation measurements are usually performed at wavelengths from the UV to the mid-IR. These measurements are self-calibrating in a sense that the division of atmospheric spectra by direct Sun (*e.g.*, exo-atmospheric) spectra yields transmission spectra. Within the SPARC Data Initiative, solar occultation is represented by ACE-FTS, ACE-MAESTRO, HALOE, POAM II/III, and SAGE I/II/III.

Stellar occultation

Stellar occultation measurements use the same concept as solar occultation measurements, except that stars act as the radiation source instead of the Sun (see **Figure 2.2**). Since multiple stars can be used, this results in a larger data density compared to that achieved by solar occultation. Night-time measurements are of better quality than daytime measurements because the scattered solar signal interferes with the target signal of the stars during daytime. The useful spectral range is limited to wavelengths below about 1 µm. At longer wavelengths terrestrial thermal emission interferes with the stellar signal. Weak stellar radiation and scintillations from atmospheric irregularities are particular challenges of stellar occultation techniques. Within the SPARC Data Initiative, stellar occultation is represented by GOMOS.

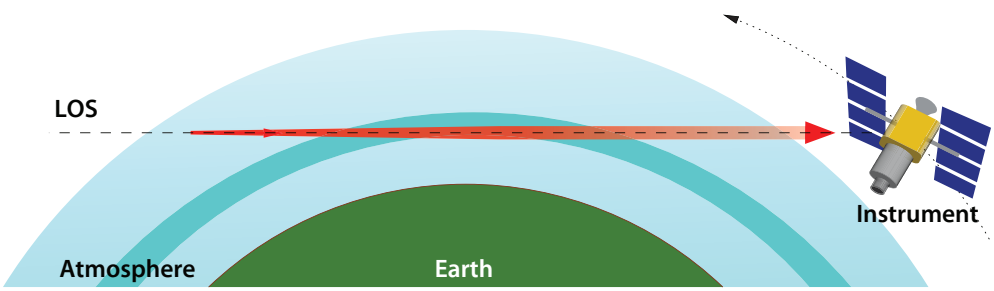


Figure 2.1: Limb emission observation geometry. *The instrument measures radiation emitted by the atmosphere along the LOS.*

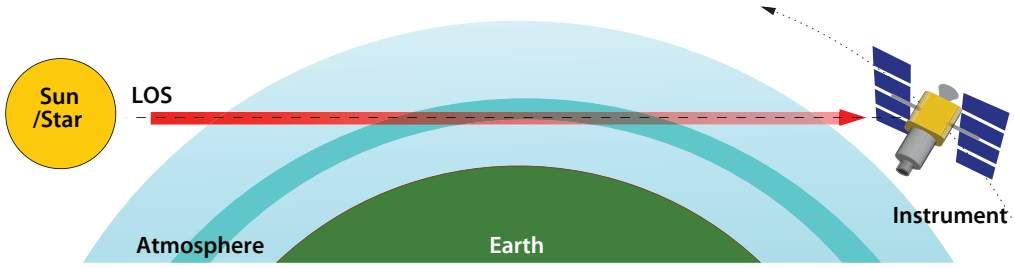


Figure 2.2: Occultation observation geometry. *The instrument points at the radiation source (the Sun or a star) and measures the radiation attenuated along the LOS.*

Limb scattering

The radiance received by limb scattering instruments consists of photons originating from the Sun and scattered into the field-of-view of the instrument (see **Figure 2.3**). The information on the atmospheric state is provided by the scattering itself, or by the absorption of scattered photons along their way through the atmosphere. In contrast to the measurement techniques discussed above, the ray-path is not defined by the measurement geometry, but is scattered by the atmosphere into the LOS of the instrument. As for all measurement techniques using the Sun as the source of the signal, measurements are only possible during daylight. On the sunlit part of the globe, good spatial coverage is achieved. The vertical resolution is similar to that of limb emission and solar occulation instruments. Measurements are made in the UV to the near-infrared range where scattering is relevant. Within the SPARC Data Initiative, limb scattering is represented by OSIRIS and SCIAMACHY.

Nadir emission

Nadir observations are measurements for which the LOS points down to the surface of the Earth. The signal received by nadir emission instruments can contain photons emitted by the Earth's surface or atmosphere and transmitted through the atmosphere. In contrast to limb measurements, for which vertically resolved measurements are achieved simply by variation of the elevation angle of the LOS, the altitude information of nadir observations is given by pressure broadening of spectral lines and by varying opacity at different wavelengths. While the altitude resolution is far inferior to that of limb sounders, the horizontal resolution is better, and allows more measurements between clouds that can penetrate lower into the troposphere. The LOS through the atmosphere is shorter than in limb sounding, which reduces sensitivity to low abundance species but also reduces opacity problems. Infrared nadir sounding is possible during both day and night, but thermal contrast has an impact on altitude resolution and sensitivity to the abundance of species in the lower troposphere. Nadir infrared measurements require on-board blackbody calibration and a space view for cold space calibration measurements. Uncertainties in surface emissivity can complicate the retrieval process. Within the SPARC Data Initiative, nadir emission measurements are represented by TES. Note that TES is the only nadir-viewing instrument considered by the SPARC Data Initiative. TES evaluations presented in this report account for the relatively broad averaging kernel of the instrument and serve as an example for the more comprehensive comparisons that would be needed when considering nadir instruments (such instruments include, for example SBUV, TOMS, and MOPITT).

2.1.2 Classification by wavelengths

The different instruments can, in addition to the classification by observation geometry, be classified according to the spectral range in which they operate. Wavelengths used for atmospheric composition measurements range from the microwave to the ultraviolet spectral region. Instruments contributing to the SPARC Data Initiative include both radiometers, which measure a signal spectrally integrated over certain frequency bands, and spectrometers, which provide spectrally resolved measurements. Better spectral resolution allows measurement of trace gas species with weaker spectral signatures. On the other hand, the advantage of lower spectral resolution is a higher signal-to-noise ratio for single measurements, which helps to provide better spatial resolution.

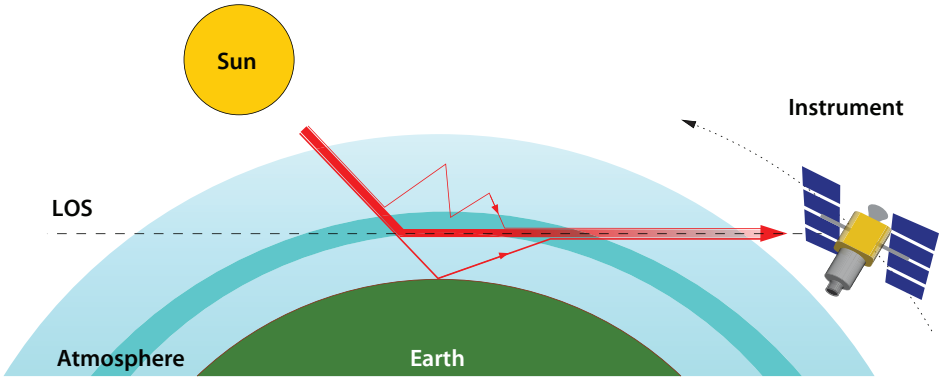


Figure 2.3: Limb scattering observation geometry. *The instrument measures radiation emitted by the Sun and scattered into the field-of-view.*

Table 2.2: Atmospheric constituents and the wavelength bands they are detected by the instruments used in this study.

Species / Wavelength bands	O_3	H_2O	CH_4 CCl_3F CCl_2F_2 HF SF_6 HNO_4 N_2O_5 $ClONO_2$ $BrONO_2$ CH_2O	N_2O CO NO HNO_3 HCl ClO HOCl	NO_2	HO_2 CH_3CN	BrO	Aerosol
Microwave/ Sub-mm	X	X		X		X	X	
Mid-IR	X	X	X	X	X			
Near-IR		X						X
VIS / UV	X				X		X	X

Table 2.2 lists the atmospheric constituents together with the wavelength bands in which they are observed by the instruments participating in the SPARC Data Initiative. In the following, the main characteristics of the different wavelength bands are briefly described.

Microwave and sub-millimeter

The microwave and sub-millimeter spectral region covers wavelengths from 10 cm to 100 μm. This corresponds to frequencies of about 3 GHz to 3 THz, respectively. The sources of radiation are rotational transitions of molecules with a permanent dipole moment. The temperature dependence of microwave and far-infrared emissions is lower than in the mid-infrared, and clouds are not as much of an interference in the former than in the latter. Measurements are not sensitive to aerosol particles or thin clouds. The typical measurement mode is emission sounding. Within the SPARC Data Initiative this spectral region is represented by Aura-MLS, SMILES, SMR, and UARS-MLS.

Mid-infrared

The mid-infrared spectral region covers wavelengths from about 2.5-20 μm, corresponding to wavenumbers from about 4000-500 cm^{-1}. The sources of the signal are rotational-vibrational transitions of molecules with a transitional dipole moment. The temperature dependence of these transitions is high and the frequency range covers the maximum of the Planck function at terrestrial temperatures. Clouds are less transparent in the mid-infrared than in the microwave spectral region. Radiative transfer is dominated by emission and absorption, while scattering is only an issue in the presence of clouds or elevated aerosol levels. Emission sounding is possible and often applied at wavelengths longer than about 4 μm (non-local thermodynamic equilibrium emission can also be detected at shorter wavelengths), but solar absorption measurements are common as well. Within the SPARC Data Initiative, this spectral region is represented by ACE-FTS, HALOE, HIRDLS, LIMS, MIPAS, and TES.

Near-infrared, visible and ultraviolet

In the near-infrared spectral region (wavelengths 0.8-2.5 μm, wavenumbers 12500-4000 cm^{-1}), overtone and combined vibrational transitions give rise to the signal, while in the visible (0.4-0.8 μm) and ultraviolet (below 0.4 μm) spectral regions, emission is caused by electronic transitions. The maximum of the Planck function of the effective temperature of the Sun's photosphere is in the visible spectral range, while emission at terrestrial temperatures at these wavelengths is negligible. Thus, remote sensing in this spectral region relies on absorption and scattering of signal emitted by hot background sources like the Sun or stars. Within the SPARC Data Initiative, this spectral region is represented by ACE-MAESTRO, GOMOS, OSIRIS, POAM II/III, SAGE I/II/III, and SCIAMACHY.

2.1.3 Satellite orbits

Stratospheric composition sounding is currently performed exclusively from low Earth orbit (LEO), platforms flying at altitudes between approximately 300 and 2000 km above the Earth's surface. The latitude coverage of the orbit is determined by its inclination, *i.e.,* the angle between the orbit plane and the equatorial plane. Polar orbits (*e.g.*, that of Envisat) with inclinations near 90° provide global coverage and allow observation of the polar regions. Many of these satellites are in sun-synchronous orbits, *i.e.,* orbits where measurements at a given geo-location on either the ascending or descending segments of the orbit have approximately the same local solar time (LST). Therefore, sun-synchronous satellites cannot provide information on the diurnal variation of the state of atmosphere at any fixed latitude. Instruments within the SPARC Data Initiative that were/are flying

on sun-synchronous satellites are Aura-MLS, GOMOS, HIRDLS, LIMS, MIPAS, OSIRIS, POAM II/III, SAGE III, SCIAMACHY, SMR, and TES. Non-sun-synchronous orbits allow Earth observation at different local times but lead to temporally varying datasets. This can be an issue when creating climatologies, particularly for species with pronounced diurnal variations. Instruments within the SPARC Data Initiative that were/are flying on non-sun-synchronous satellites/platforms are ACE-FTS, ACE-MAESTRO, HALOE, SAGE I/II, SMILES, and UARS-MLS.

2.2 Instrument and retrieval descriptions

The satellite instruments participating in the SPARC Data Initiative are all passive sensors using a limb viewing observation geometry with the exception of one nadir-viewing sounder used for particular upper troposphere/lower stratosphere (UTLS) studies (see *Section 4.27*). The measurement modes of the limb-viewing sounders (emission, scattering, solar occultation, and stellar occultation) determine data coverage and sampling density.

Retrieval processes include a so-called forward model and an inversion algorithm. The forward model computes radiances that would be observed given a state vector of atmospheric composition and temperature profiles. The inversion algorithm then "inverts" these calculations and solves for an atmospheric state from a given set of radiance observations. In many cases (ACE-FTS, Aura-MLS, HALOE, HIRDLS, LIMS, MIPAS, SMR, TES, UARS-MLS), initial retrievals of temperature and pressure are performed using observations of molecules whose abundances are well known (usually CO_2 in the infrared and O_2 in the microwave). Temperature and pressure can be retrieved as separate products if the emission lines are strong enough (*e.g.*, SMILES, SMR). Some instruments (*e.g.*, OSIRIS, SAGE II, SCIAMACHY) rely on meteorological analyses for temperature profile information. In either case, accurate knowledge of tangent altitude/pressure is required for limb measurements.

Uncertainties are typically provided by the operational retrieval systems, but they generally do not include systematic effects such as the propagation of spectroscopic uncertainties. Beyond such uncertainties, retrieval constraints (*e.g.*, smoothing) affect the altitude resolution and lead to an imperfect representation of the true atmospheric state. Available validation information is provided separately for each molecule in *Chapter 4*.

In the following, the different instruments together with their retrieval processes are described, in order of their launch date, with the earliest instrument first.

2.2.1 LIMS on Nimbus 7

Nimbus 7 was launched on October 24, 1978, and carried a number of instruments for making measurements of the state of the middle atmosphere. The Limb Infrared Monitor of the Stratosphere (LIMS) experiment was a limb-infrared sounder, focused on measurements of temperature, O_3, and those species that affect ozone (H_2O, NO_2 and HNO_3) [see *Gille and Russell*, 1984]. Nimbus 7 was in a sun-synchronous orbit with a noon and midnight equator crossing time. However, LIMS was designed to look off-plane, so that the measurements were made near 1pm and 11pm local time at equator crossing. The resulting sampling pattern can be found in Figure 10 of *Gille and Russell* [1984]. The temperature and ozone profiles extend from cloud-top to near the mesopause, while the profiles of H_2O, HNO_3, and NO_2 are restricted to the stratosphere, due to their signal-to-noise (S/N) limitations. The cryogen gases that were used to cool the detectors only lasted until May 28, 1979, as planned. Thus, the LIMS dataset extends for about 7.5 months and consists of daily, orbital profiles from about 64°S to 84°N latitude. The data were processed with a Version 5 algorithm and archived in 1982 at NASA Goddard. More recently, the algorithm was revised to Version 6, and new retrievals were conducted and archived at the Mirador site of the Goddard Earth Sciences and Data Information Services Center (GES DISC) or at **http://daac.gsfc.nasa.gov** and can be downloaded *via* ftp from there. A separate LIMS website exists at **http://www.gats-inc.com/projects.html#lims** for viewing daily plots of the data. Descriptions of the quality of the Version 6 temperature, O_3, H_2O, and HNO_3 and NO_2 can be found in *Remsberg et al.* [2004, 2007, 2009, and 2010], respectively.

Retrievals for the LIMS V6 temperature *versus* pressure (or $T(p)$) profiles are described in *Remsberg et al.* [2004] and references therein. The algorithm uses a top-down, onion-peeling approach and iterates to achieve a match of the calculated and measured radiances for its wide and narrow CO_2 radiometer channels in the 15-μm region. A constant CO_2 mixing ratio profile was assumed for the forward radiance models. Radiance profiles for the LIMS species channels are registered with pressure according to the associated $T(p)$ profiles, and their forward models account for the retrieved temperatures. Level 2 profiles of the temperature and species volume mixing ratio (VMR) are tabulated at 18 levels per decade of pressure or at a spacing of 0.88 km. They have an effective vertical resolution of 3.7 km. The retrieval algorithm for NO_2 accounts for interfering radiances from H_2O, CH_4, and the oxygen continuum in the 6-7 μm region. The algorithm for HNO_3 accounts for interfering radiances from the primary CFC molecules and from aerosol emissions in the 11-μm region.

2.2.2 SAGE I on AEM-B, SAGE II on ERBS, and SAGE III on Meteor-3M

The Stratospheric Aerosol and Gas Experiment (SAGE) series of instruments consists of four instruments including the Stratospheric Aerosol Measurement (SAM II) that span the period from 1978 through 2005. All of the instruments use solar occultation to measure attenuated solar radiation through the Earth's limb during satellite

sunrise and sunset. The first instrument in the series, SAM II on-board Nimbus 7 (1978-1993), consisted of a single 1000-nm aerosol channel with measurements restricted to high latitudes (>53° in both hemispheres). Note, SAM II is not included in the evaluations of this report. SAGE I on-board AEM-B (1979-1981) consisted of four measurement channels (corresponding to wavelengths of 385, 450, 600, and 1000 nm), which were used to infer aerosol extinction profiles at two wavelengths (450 and 1000 nm) and O_3 and NO_2 concentration profiles. SAGE II on-board ERBS (1984-2005) made measurements at seven wavelengths (385, 448, 452, 525, 600, 940, and 1020 nm) from which O_3, NO_2, H_2O and aerosol extinction at four wavelengths (385, 452, 525, and 1020 nm) were retrieved [*McCormick et al.*, 1989]. SAGE III on-board the Russian Meteor-3M satellite was launched on December 2001 and remained operational into December 2005. It used an 800 element Charged Coupled Device (CCD) linear array detector to provide continuous spectral coverage between 280 and 1040 nm. An additional single photodiode at 1550 nm was used for aerosol extinction measurements. The SAGE III measurements at 87 channels between 285 and 1545 nm were used to infer vertical profiles of O_3, NO_2, H_2O, and aerosol extinction at nine wavelengths (285, 448, 521, 602, 676, 755, 868, 1019, and 1545 nm) [*Thomason and Taha*, 2003].

Both SAGE I and II instruments were in inclined (~57°) orbits that permitted near-global coverage over the course of 30 to 40 days (see **Figure 2.4**). There are 15 sunrise and 15 sunset measurements each day that cover a narrow latitude band and are separated by ~24° in longitude. Unlike SAGE I and II, where sunrise and sunset measurements alternatively observe the Northern and Southern Hemispheres, all SAGE III sunrise measurements occur in the Southern Hemisphere (30°S to 60°S) while all sunset

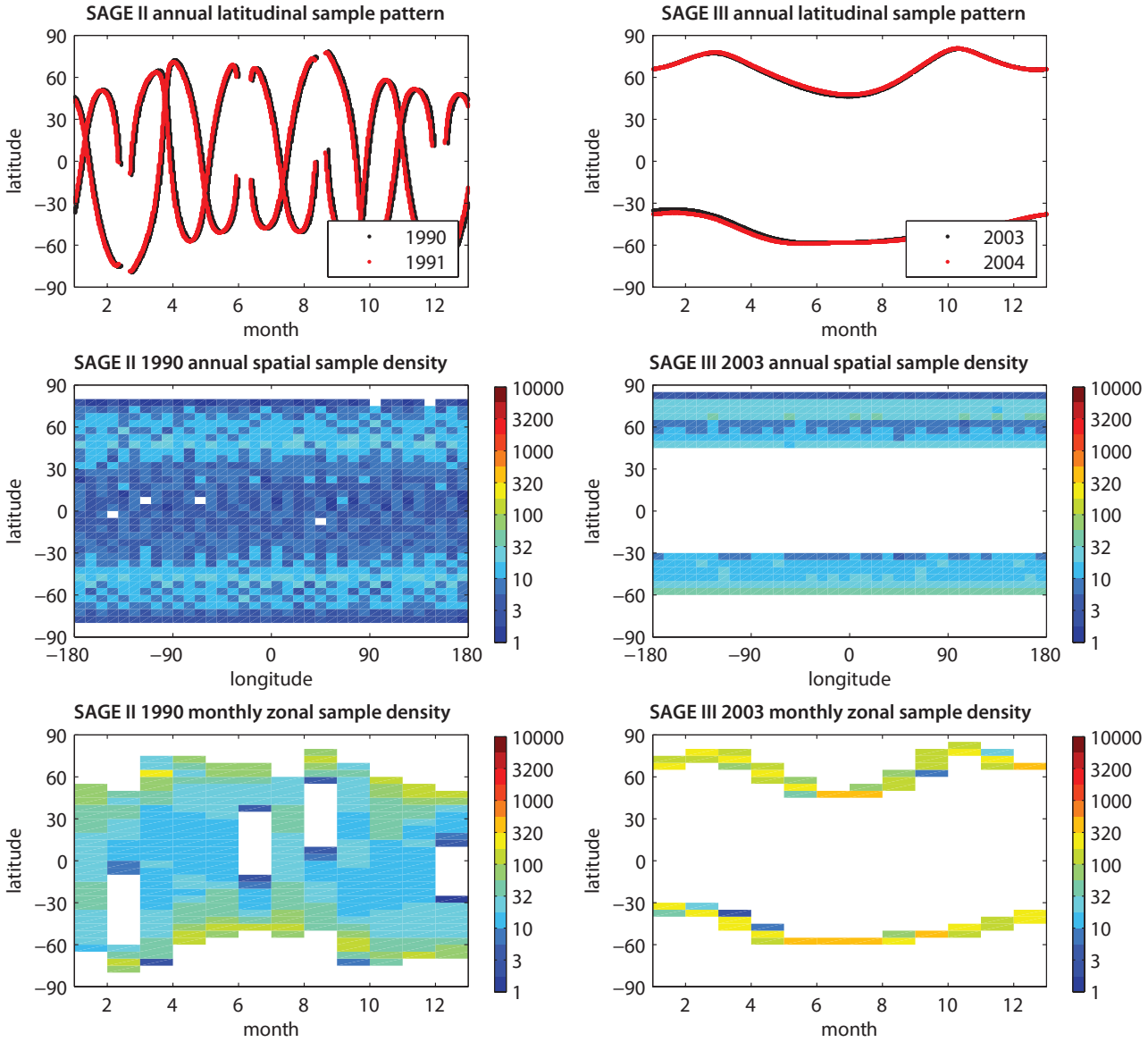


Figure 2.4: Sampling pattern and resulting sample density for SAGE II (left) and SAGE III (right). *Note, SAGE I provided similar geographical and temporal sampling as SAGE II. For SAGE III, sunrise measurements occur in the Southern Hemisphere, and sunset events occur in the Northern Hemisphere.*

measurements occur in the Northern Hemisphere (40°N to 80°N) due to its sun-synchronous orbit (see **Figure 2.4**).

SAGE III additionally operated in lunar occultation mode from which O_3, NO_2, NO_3, and OClO were derived. Currently no aerosol product is produced from lunar occultation measurements. Since there are fewer lunar occultation data from SAGE-III, only measurements from solar occultation are used to create the climatologies used in this report.

An aerosol climatology was developed by the SPARC Assessment of Stratospheric Aerosol Properties (ASAP) and is available on the SPARC Data Centre website (**http://www.sparc-climate.org/data-centre/**). Months during 2005 that are missing on this website are available by request from Larry Thomason (l.w.thomason@nasa.gov).

The retrieval of trace gas profiles from SAGE measurements is accomplished by taking the following major steps. First the solar radiance at all measured wavelengths along with spacecraft ephemeris data are processed to produce slant path optical depth profiles as a function of tangent height. The total slant path optical depth at a particular wavelength is a linear combination of Rayleigh scattering and other contributed trace gases (*e.g.*, O_3, NO_2, and aerosol). The contribution of Rayleigh scattering is first removed from the total slant path optical depth before an inversion algorithm is applied to optimally account for the contribution of other measured gases. Detailed descriptions of retrieval algorithms for SAGE I, SAGE II, and SAGE III can be found in *Chu and McCormick* [1979], *Chu et al.* [1989] and SAGE III ATBD [2002], respectively. The native data files can be found *via* the NASA LaRC data website **http://eosweb.larc.nasa.gov/**.

2.2.3 HALOE on UARS

The Halogen Occultation Experiment (HALOE) was launched on-board the Upper Atmosphere Research Satellite (UARS) on September 12, 1991. The HALOE instrument performed flawlessly over the UARS lifetime through November 2005. The UARS was in a 600-km near-circular orbit with a 57° inclination. HALOE used the solar occultation technique and the instrumental methods of gas-filter radiometry to measure vertical profiles of HF (2.45 μm), HCl (3.4 μm), CH_4 (3.46 μm) and NO (5.26 μm), and broadband radiometry to measure vertical profiles of NO_2 (6.25 μm), H_2O (6.6 μm), O_3 (9.6 μm), and temperature *versus* pressure with approximately 2.3 km vertical resolution. HALOE also measured aerosol extinction in the four gas-filter channels. The altitude coverage is species-dependent, but is limited to within the 10-150 km range. HALOE measured 15 sunrise and 15 sunset events per day and achieved near-global coverage in approximately a month. The daily measurement spacing was equal in longitude and varied seasonally in latitude. The HALOE measurement sampling was influenced over the lifetime of the mission by: 1) drifts in the UARS orbit; 2) the power-sharing mode among UARS instruments due to a malfunction of the solar array in May 1995; 3) reduced battery power in June 1997; and 4) difficulties with the spacecraft tape-recorder mechanism in October 1999. For a detailed description of the HALOE measurement and retrieval techniques, see *Russell et al.* [1993]. The sampling pattern and resulting measurement density from HALOE can be seen in **Figure 2.5**.

The HALOE temperature retrieval assumes a CO_2 concentration that varies based on the annual CO_2 increase rate determined form ground-based and *in situ* measurements. The observed 3570 cm^{-1} transmission is matched in an upward, hydrostatically-constrained process. This is iterated several times, with intervening profile registrations. Above ~85 km, temperatures from the MSIS model [*Hedin*, 1991] are assumed, and below ~35 km NCEP temperatures are used. The 1510, 1600 and 1015 cm^{-1} radiometer channels are used to retrieve NO_2, H_2O, and O_3, respectively, in an

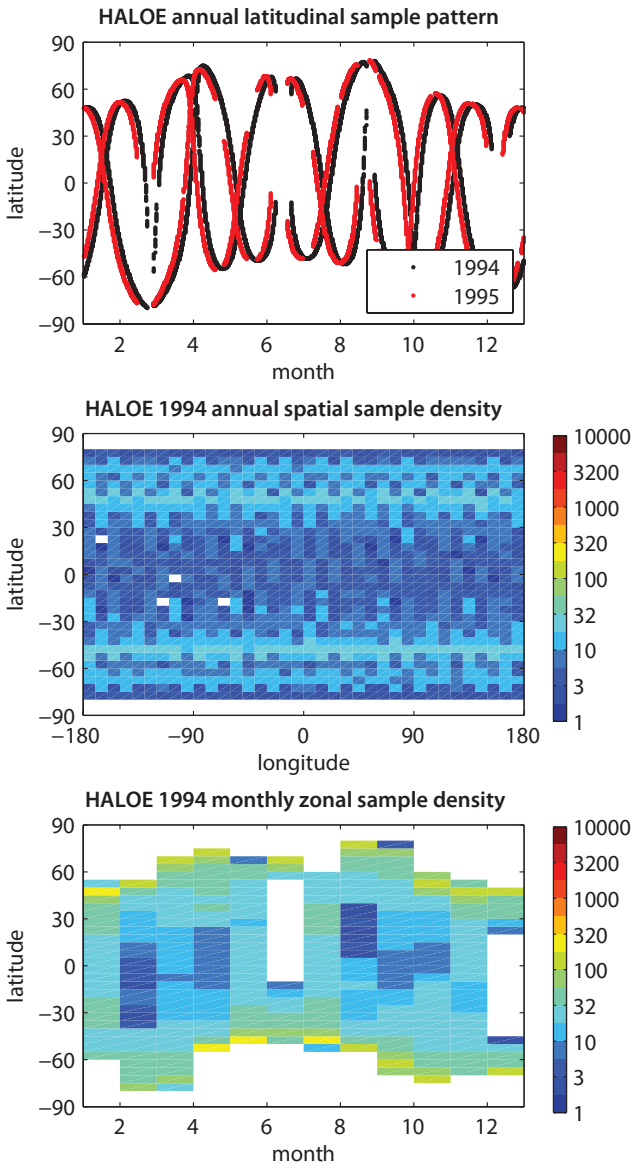


Figure 2.5: Sampling pattern and resulting sample density for HALOE. *Note that the sampling pattern shifts from year to year.*

onion-peeling fashion. The Gas Filter Radiometer differential technique is used to retrieve HF, HCl, CH_4, and NO from the 4080, 2940, 2890, and 1900 cm^{-1} channels. In these channels, the light is split. Half is sent through a cell filled with the target gas, and the other half through a vacuum path. The exo-atmospheric difference of these signals is balanced to within the noise levels. The difference-signal that develops when viewing through the atmosphere is highly sensitive to atmospheric absorption from the target gas, but virtually insensitive to aerosol absorption. The aerosol extinction is retrieved from the 1900 cm^{-1} (*i.e.,* NO channel) vacuum-path signal and extrapolated to the radiometer channels assuming a sulphate model to account for the sensitivity to aerosols at these wavelengths. The spectroscopy used in the HALOE forward model is based on HITRAN 1991–1992. The HALOE algorithm has gone through two major revisions. The initial HALOE validation results for each species were published in 1996 [*Russell et al.*, 1996a, 1996b; *Gordley et al.*, 1996; *Harries et al.,* 1996; *Hervig et al.*, 1996a, 1996b; *Park et al.*, 1996; *Brühl et al.*, 1996]. The HALOE processing version used in the SPARC Data Initiative is the third public release (V19) which can be obtained from the following website: **http://haloe.gats-inc.com/home/index.php**. Numerous satellite science teams have used HALOE V19 to compare and validate their instruments [*e.g.*, *Randall et al.*, 2003; *Froidevaux et al.*, 2006] and this version has been extensively used in previous SPARC reports [*e.g.*, *SPARC*, 2000]. In addition, a comprehensive stratospheric climatology of O_3, H_2O, NO_x, HF, HCl, and CH_4 was developed from HALOE V19 measurements by *Grooß and Russell* [2005].

2.2.4 MLS on UARS

UARS-MLS was one of ten instruments on the UARS platform, launched on 12 September 1991 as mentioned in *Section 2.2.3* [*Reber et al.*, 1993]. UARS-MLS (a predecessor to Aura-MLS) pioneered microwave limb sounding of the Earth's stratosphere and mesosphere from space. It was designed to measure stratospheric O_3, H_2O and ClO, but also provided stratospheric and mesospheric temperature, and stratospheric HNO_3 (as well as upper tropospheric humidity and other information not used in this report). UARS-MLS measured millimeter-wavelength thermal emission as the antenna was vertically scanned (every 65.54 s) from about 1 to 90 km through the atmospheric limb [*Barath et al.*, 1993; *Waters et al.*, 1993]. There were typically 26 limb views during each 65-s scan. The vertical resolution as constrained by the field-of-view is ~3 km, and the UARS-MLS data (for the data versions used here) are produced on a vertical grid with a resolution of ~2.7 km. The spatial resolution is about 400 km along the LOS, and about 7 km across. UARS-MLS used three radiometers to measure the microwave emission near 63, 205, and 183 GHz. The radiances in each band were measured by one of six identical spectrometer filter-banks, each consisting of 15 contiguous channels, covering up to ±255 MHz away from the line centre. The channels vary in width from 2 MHz near the line centre to 128 MHz in the wings.

The UARS orbit was inclined at 57° and the satellite performed a 180° yaw maneuver 10 times per year, at approximately 36 day intervals. The UARS-MLS measurements cover 34° on one side of the equator to 80° on the other side, with hemispheric coverage switching with each yaw maneuver. The orbit precession ensured that the measurements covered essentially all LSTs during each 36 day interval. Profiles were spaced ~3-4° along the orbit track and the average daily sampling in longitude was ~12°. Coverage was denser near the turn-around latitudes. The main operational events affecting the time series from UARS-MLS were the mid-April 1993 failure of the 183-GHz radiometer, resulting in the loss of stratospheric H_2O (and 183-GHz O_3 observations), and the mid-June 1997 cessation of 63-GHz observations in order to save spacecraft power, resulting in a loss of the temperature information. The frequency of MLS operational days generally decreased over the mission, from close to 100% from late 1991 through 1993, down to about 50% in late 1994, and only several tens of

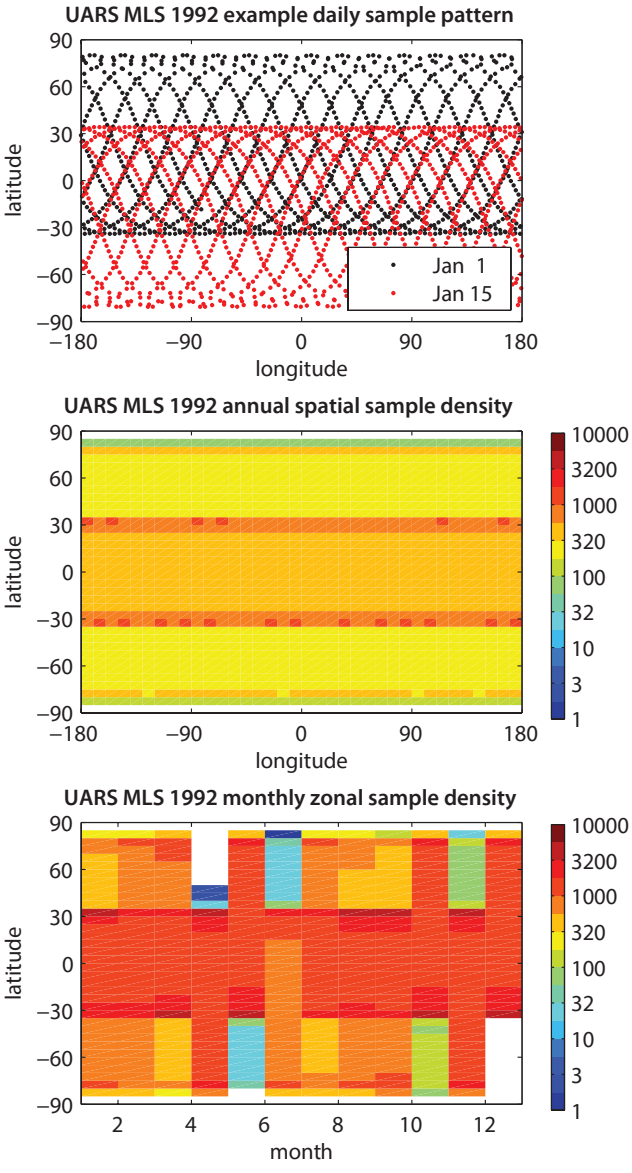


Figure 2.6: Sampling pattern and resulting sample density for UARS-MLS.

measurement days per year from 1995 onward; the last retrievals were obtained on 25 August 2001. The relevant (O_3, ClO, HNO_3) UARS-MLS data are therefore generally considered most robust for "long-term" series analyses until mid-June 1997; we have included data through 1999 for this report and the related database. The sampling pattern and resulting measurement density from UARS-MLS can be seen in **Figure 2.6**, for one of the early years (with best coverage).

The UARS-MLS retrieval algorithms are based on the optimal estimation approach [*Rodgers*, 1976, 2000]. These algorithms make use of two different forward models; one is a complete line-by-line radiative transfer model, and the other is based on a Taylor series computation using precomputed output from the full model. The standard UARS-MLS products are temperature, H_2O, O_3, HNO_3, ClO, and CH_3CN. The Version 5 data were the last major public release of UARS MLS data, however, updates and improvements were made available for H_2O and HNO_3 [see *Livesey et al.*, 2003], which is why we have used a Version 6 file label for these two species. For stratospheric H_2O, the work of *Pumphrey* [1999] and *Pumphrey et al.* [2000] demonstrated the value of using the originally-named V0104 dataset (also used here and referred to as V6), rather than V5 H_2O. UARS-MLS stratospheric H_2O mixing ratios are typically flagged as bad for pressures larger than 100 hPa. Moreover, there are no valid data after the month of April, 1993, as the radiometer measuring stratospheric H_2O failed that month.

The original data files used to produce the climatological files are the standard Level 3AT UARS MLS daily files. These files contain data on a subset of the standard "UARS" pressure surfaces, which are evenly spaced with six surfaces per decade change in pressure (or about 2.7 km), although the true resolution is typically somewhat coarser. In addition, Level 3TP "Parameter files" are produced for each day of MLS observations. These files contain information on the quality of the UARS-MLS data. The supplementary material from *Livesey et al.* [2003] gives more information on the implementation of the UARS-MLS retrieval algorithms, as well as data screening guidelines; the mixing ratio profiles (*versus* pressure) were screened accordingly, interpolated vertically, and averaged to obtain the monthly zonal means used here. The general guidelines for the proper use of UARS-MLS data (see *Livesey et al.* [2003]) have been followed, namely: 1) only data whose associated uncertainty is positive should be used; 2) only profiles where the MMAF_STAT diagnostic field is set to G, T, or t should be used; 3) only profiles where the appropriate QUALITY field is equal to 4 should be used; and 4) the spike information given on the MLS science team website should also be used for removing outliers. The official public distribution location for UARS-MLS data used here is (as for Aura-MLS) at the NASA GES-DISC Mirador website, namely **http://mirador.gsfc.nasa.gov**. Public information about both MLS instruments, data access, and MLS-related publications, can be found at the MLS website (**http://mls.jpl.nasa.gov**).

2.2.5 POAM II on SPOT-3 and POAM III on SPOT-4

The Polar Ozone and Aerosol Measurement II (POAM II) instrument was launched on-board the French SPOT-3 satellite on 26 September 1993 into a 98.7° inclination, sun-synchronous orbit at an altitude of 833 km. The instrument operated between October 1993 and November 1996 when the SPOT-3 satellite suffered a malfunction and contact with the instrument was terminated. POAM III was launched on the French SPOT-4 spacecraft on 24 March 1998 into an orbit identical to the one of SPOT-3. The instrument began taking data on 22 April 1998 and operated until 5 December 2005, when instrument failure terminated the mission. POAM III was functionally very similar to its predecessor, although it contained a number of design changes that improved sensitivity and accuracy. POAM II and III both used the solar occultation technique, measuring the extinction of solar radiation in nine narrow-band channels from approximately 350 to 1060 nm and 353 to 1018 nm, respectively, to retrieve the vertical distribution of atmospheric O_3, H_2O, NO_2, and aerosol extinction. Over their mission lifetimes, POAM II and III compiled datasets of approximately 21,000 and more than 43,000 good occultation profiles, respectively. POAM II and III made 14 measurements per day in each hemisphere, equally spaced in longitude around a circle of approximately constant latitude. Satellite sunrise measurements were made in the Northern Hemisphere and sunsets in the Southern Hemisphere. Sunrise measurements occur in a latitude band from 55-71°N while sunsets occur between 63-88°S. The latitude coverage changes slowly with season and is exactly periodic from year to year. The sampling patterns of POAM II and III are shown in **Figure 2.7**.

Vertical resolution of the POAM data products is approximately 1 to 1.5 km, depending on the species. The altitude range also varies by species and instrument version; for POAM II O_3 (15-50 km), NO_2 (20-40 km) and aerosols (10-30 km), and for POAM III O_3 (5-60 km), NO_2 (20-40 km), H_2O (5-45 km) and aerosols (5-25 km). Note that unlike POAM II, POAM III also provided a water vapour product that was thoroughly validated against a variety of correlative satellite-, aircraft- and balloon-borne datasets. Due to uncertainties in the optical filters for the differential water vapour channels, water vapour was never retrieved operationally from POAM II measurements.

A complete discussion of the POAM II instrument can be found in *Glaccum et al.* [1996]. The Version 6 algorithms, error analysis and data characterisation are described by *Lumpe et al.* [1997]. A discussion of the POAM III instrument can be found in *Lucke et al.* [1999]. The Version 4 algorithms, error analysis and data characterisation are described by *Lumpe et al.* [2002]. The final public release datasets for POAM II (V6.0) and POAM III (V4) are available at the NASA Langley Atmospheric Sciences Data Center (**http://www.eosweb.larc.nasa.gov**) and are also distributed by the Naval Research Laboratory *via* **https://www.nrl.navy.mil/rsd/7220/poam-ftp**.

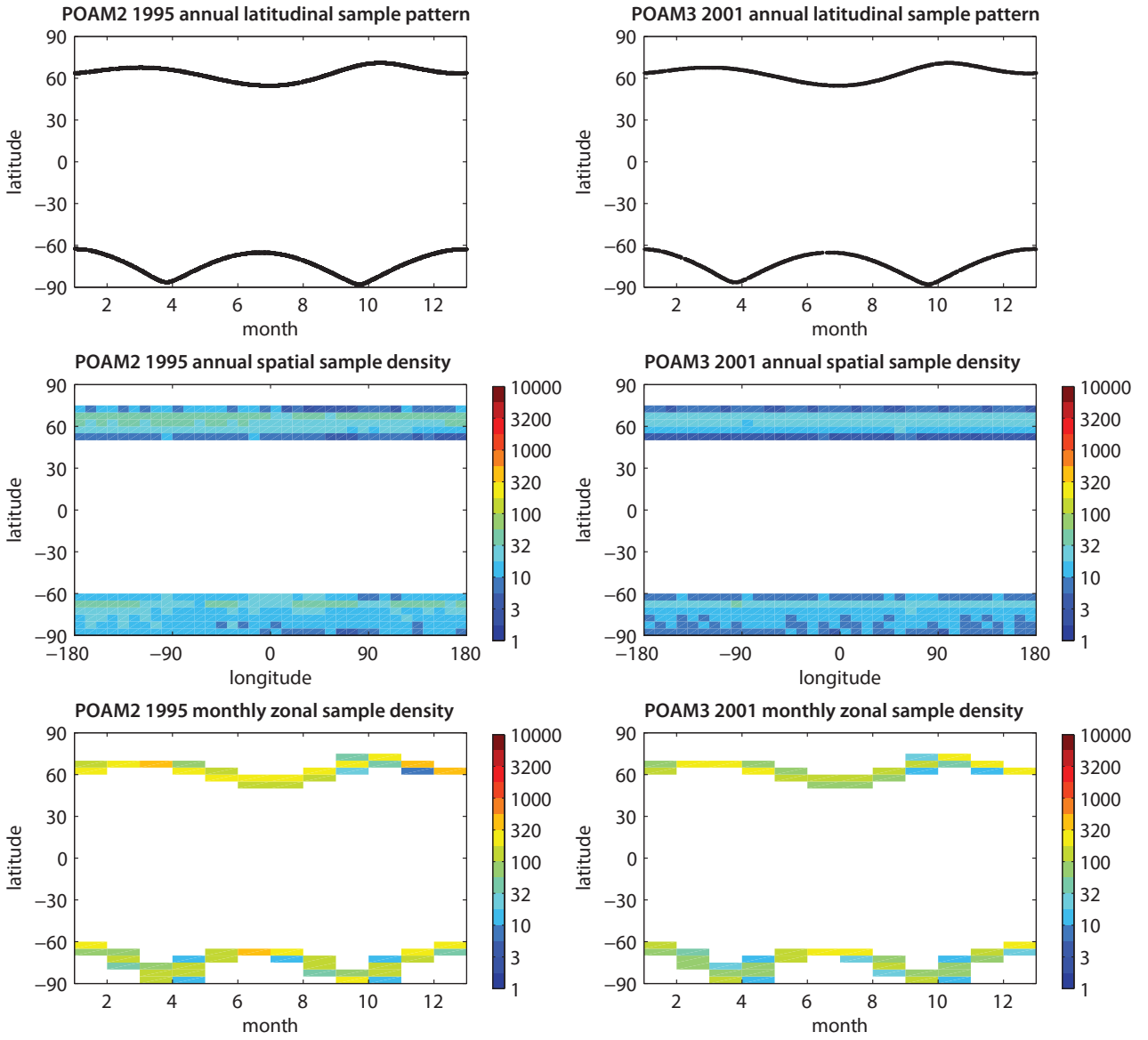


Figure 2.7: Sampling pattern and resulting sample density for POAM II (left) and III (right).

POAM measures limb profiles of slant-path transmission in nine spectral channels from roughly 350 to 1000 nm. Using this input data stream the algorithms retrieve vertical profiles of O_3, NO_2, H_2O, and O_2 (or total) density, as well as aerosol extinction between 350 and 1000 nm (POAM II did not retrieve H_2O or total density – see above). All atmospheric species are retrieved simultaneously using an optimal estimation algorithm (fixed, non-varying *a priori* for all species; constraints are tuned to minimise retrieval variability at the desired vertical resolution). The conversion of transmission data to geophysical profiles is achieved *via* a two-step process, beginning with a spectral inversion to partition the various gas and aerosol components of the measured total slant optical depth, followed by a spatial inversion to produce altitude profiles of gas density and aerosol extinction from the path integrated quantities. The NO_2 and H_2O retrievals use closely spaced differential absorption pair channels in the UV and Near-IR, respectively, while O_3 is retrieved from a single channel at the peak of the Chappuis band at 602 nm. Aerosols are retrieved at all wavelengths by constraining the spectral dependence to a quadratic in log-log space (optical depth *versus* wavelength).

Both instruments included an O_2 A-band channel designed to provide self-consistent temperature/pressure retrievals, however they were never made operational (the POAM II channel saturated, while POAM III had an unresolved systematic bias presumably due to bandpass characterisation errors). The POAM III retrievals used the Rayleigh scattering signal in the 350-nm channel to retrieve total density above 30 km and hence remove the background Rayleigh scattering self-consistently from all channels. Below 30 km the density is tightly constrained to the United Kingdom Meteorological Office (UKMO) analysis. The POAM II retrievals were constrained to fix the total density to the UKMO analysis (co-located in time and space) due to an unresolved overall altitude grid error.

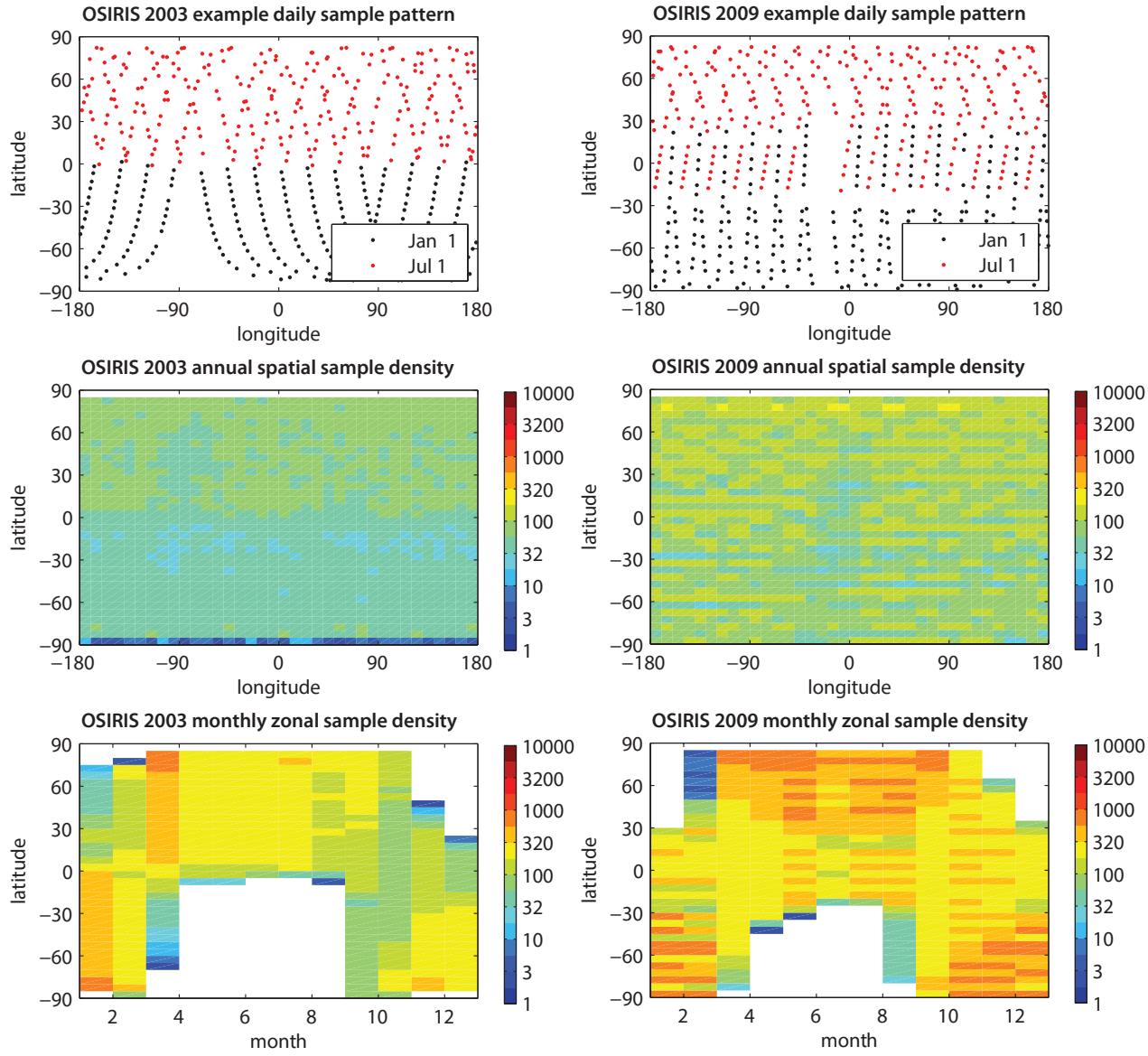


Figure 2.8: Sampling pattern and resulting sample density for Odin/OSIRIS for 2003 and 2009.

2.2.6 OSIRIS on Odin

The Odin satellite was launched on 20 February 2001 into a 600-km circular sun-synchronous near-terminator orbit with a 97.8° inclination [*Murtagh et al.*, 2002]. Odin carries two instruments: the Optical Spectrograph and InfraRed Imager System (OSIRIS) [*Llewellyn et al.*, 2004] and the Sub-Millimetre Radiometer (SMR; see *Section 2.2.7*) [*Frisk et al.*, 2003]. The instruments are co-aligned and scan the limb of the atmosphere through controlled nodding of the satellite over a tangent height range from 7 to 70 km in approximately 85 s (stratospheric mode, ~65 scans per orbit) or from 7-110 km in about 140 s (stratospheric-mesospheric mode, ~40 scans per orbit). Due to Odin's orbit, the data from both instruments are generally limited to between 82°N and 82°S except for occasional short periods of off-plane pointing at high latitudes during early polar spring. The LSTs of the observations are close to 6pm and 6am for low and mid-latitudes during the ascending and descending nodes respectively, but sweep quickly over local midnight and noon at the poles. Moreover, the equator crossing times are slowly drifting in LST during the Odin mission. A particularity of the Odin satellite is that observation times were initially equally shared between astronomical and atmospheric observation modes. The astronomy mission ended in April 2007 and since then Odin has been entirely dedicated to atmospheric sciences.

OSIRIS is a grating spectrometer that measures limb-scattered sunlight spectra in the spectral range from 280 nm to 800 nm at a resolution of about 1 nm. The scattered sunlight measurements are used to provide vertical profiles of minor stratospheric constituents including O_3, NO_2, BrO and aerosol. Additional datasets exist, but only the official products are mentioned here. Since OSIRIS observations are dependent on sunlight, the full latitude range is only covered around the equinoxes and hemispheric coverage is provided elsewhere. Examples of daily and annual sampling distributions are shown in **Figure 2.8.**

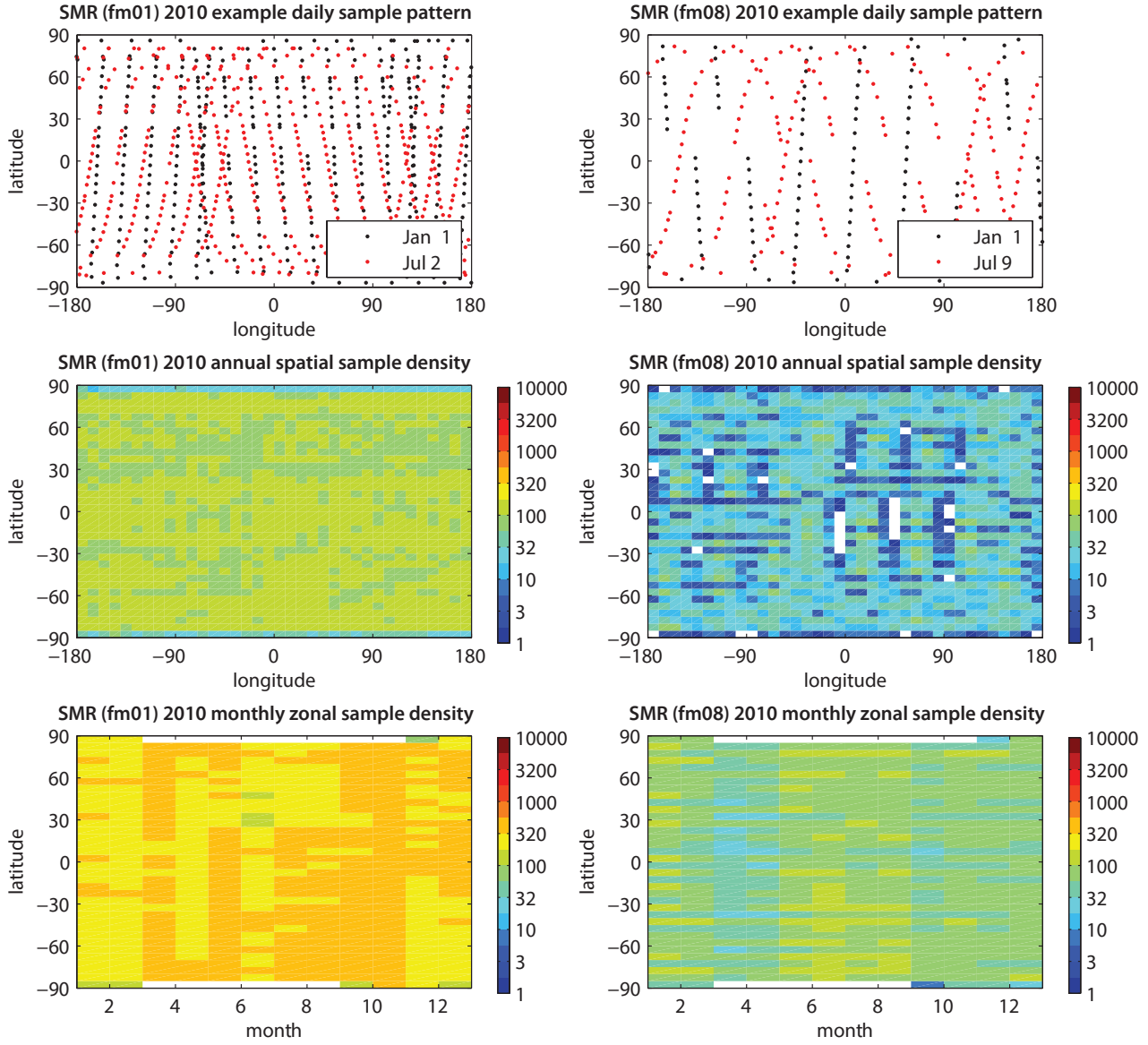


Figure 2.9: Typical sampling pattern and resulting sample density for Odin/SMR for 2010. *Left: stratospheric mode; Right: water isotope mode (H_2O-16). Note that the sampling density increased from April 2007 when the Odin astronomy mission ended.*

The NO_2 (V3.0) product is retrieved using a combination of DOAS and the log-space optimal estimation method using wavelengths between 435 and 451 nm [*Haley et al.*, 2004; *Brohede et al.*, 2007a; *Haley and Brohede* 2007]. BrO (V5) is also retrieved with optimal estimation, but on zonally-averaged OSIRIS spectra, in the 346-377 nm range [*McLinden et al.*, 2010]. Ozone (V5) is retrieved with a multiplicative algebraic reconstruction technique (MART) using a range of doublet/triplets in the Hartley and Huggins bands [*Degenstein et al.*, 2009]. OSIRIS ozone profile measurements show agreement with coincident SAGE II occultation measurements to within 2% from 18 to 53 km altitude over a large range of geo-locations and solar zenith angles. Stratospheric aerosol (V5) is also retrieved using a MART algorithm where the retrieval vector is designed to enhance the extra scattering, above the Rayleigh background, due to sulphate aerosols [*Bourassa et al.*, 2007]. For this vector a wavelength ratio of 750 nm to 470 nm is used to characterise the effect of the Mie scattering signal. Hydrated sulphuric acid particle microphysics, including a size distribution for typical background aerosol, are assumed to calculate the scattering cross section and phase functions that are required to retrieve the aerosol extinction. The altitude range and resolution vary for each species and profile but are usually limited to the stratosphere and a maximum of ~2 km vertical resolution.

Inferred NO_y, NO_x and Br_y data products are also compiled using OSIRIS data, combined with photochemical box-model simulations for each individual profile [*Brohede et al.*, 2008; *McLinden et al.*, 2010], although Br_y is not presented in this report. Note that HNO_3 observations from the Odin/SMR instrument are also used in the NO_y product (NO_2+NO+HNO_3+$ClONO_2$+2*N_2O_5). The NO_x dataset (NO_2+NO) is not explicitly described in the literature but is compiled using box-model scaling factors, following

the approach in *Brohede et al.* [2008]. Previous climatology studies and model inter-comparisons with OSIRIS data are described by *Brohede et al.* [2007b] for NO_2, *McLinden et al.* [2010] for BrO/Br_y and *Brohede et al.* [2008] for NO_y. See the OSIRIS official website for more information and data access: **http://osirus.usask.ca/.**

2.2.7 SMR on Odin

The Sub-Millimetre Radiometer (SMR) on-board the Odin satellite (for launch and orbit details, see *Section 2.2.6*) uses four sub-millimetre and 1-millimetre wave radiometer to measure thermal emission from the atmospheric limb in the 486-581 GHz spectral range and around 119 GHz [*Murtagh et al.*, 2002; *Frisk et al.*, 2003]. The signal is collected by a 1.1 m telescope and spectrally analysed by two auto-correlator spectrometers, each with 800 MHz bandwidth and 2 MHz effective resolution. Stratospheric mode observations of O_3, ClO, N_2O, HNO_3, and H_2O in the UTLS are performed using two bands around 501 and 544 GHz on every third observation day (on every other day since April 2007) [*e.g.*, *Urban et al.*, 2005a, 2006; *Urban*, 2008; *Ekström et al.*, 2008]. Other regular observation modes are dedicated to the measurements of target species in the middle atmosphere such as water and ozone isotopologues around 490 GHz [*Urban et al.*, 2007; *Jones et al.*, 2009], mesospheric and lower thermospheric H_2O at 557 GHz [*Urban et al.*, 2007; *Lossow et al.*, 2009; *Orsolini et al.*, 2010], stratospheric and mesospheric CO, O_3 and HO_2 around 576 GHz [*Dupuy et al.*, 2004; *Jin et al.*, 2009; *Baron et al.*, 2009], and H_2O-17, O_3, and NO in a band at 551 GHz [*Urban et al.*, 2007]. For example, water isotope mode observations of H_2O^{16} were performed on 1 day per week until 2007 (10 days per month since April 2007). The sampling pattern and resulting measurement density from SMR for the stratospheric mode and the water isotope mode can be seen in **Figure 2.9**.

Vertical profiles (Level-2 data) are retrieved from the calibrated spectral measurements of the limb scans (Level-1b data) by inverting the radiative transfer equation for a non-scattering atmosphere. Employed retrieval techniques for Odin/SMR Level-2 processing are based on the optimal estimation method (except for upper tropospheric humidity and ice) [*Urban et al.*, 2004; *Buehler et al.*, 2005; *Eriksson et al.*, 2005]. The altitude range and resolution varies for each species depending on the signal-to-noise ratio and frequency band employed. Currently recommended data versions are V2.0 for the 544 GHz band and V2.1 for all other modes.

Climatologies of several species (N_2O, HNO_3, NO_2, NO_y, CO, ClO, O_3), derived from Odin observations since 2001 and compiled in terms of altitude or equivalent latitude *versus* pressure, altitude, or potential temperature, are available from the Odin/SMR website (**http://odin.rss.chalmers.se**). For information on the climatologies of HNO_3, NO_2, and derived NO_y the reader is referred to *Urban et al.* [2009], *Brohede et al.* [2007a] and *Brohede et al.* [2008].

2.2.8 GOMOS on Envisat

GOMOS (Global Ozone Monitoring by Occultation of Stars) was a stellar occultation instrument on-board the European Space Agency's Environmental satellite, Envisat [*Bertaux et al.*, 2010; **http://envisat.esa.int/handbooks/gomos/**]. Envisat was launched into its sun-synchronous polar orbit of 98.55° inclination at about 800 km altitude on 1 March 2002. Contact to the satellite was lost on 8 April 2012. Its equator crossing time was 10am. For every occultation GOMOS first measured a star's reference spectrum when the star was seen above the atmosphere. This reference spectrum and the spectra measured through the atmosphere were used to calculate the horizontal transmission spectra through the atmosphere. Transmissions are the basis for spectral and vertical retrieval of species profiles. GOMOS performed 100-200 night occultations per day. The measurement coverage of night occultations was global, except in the

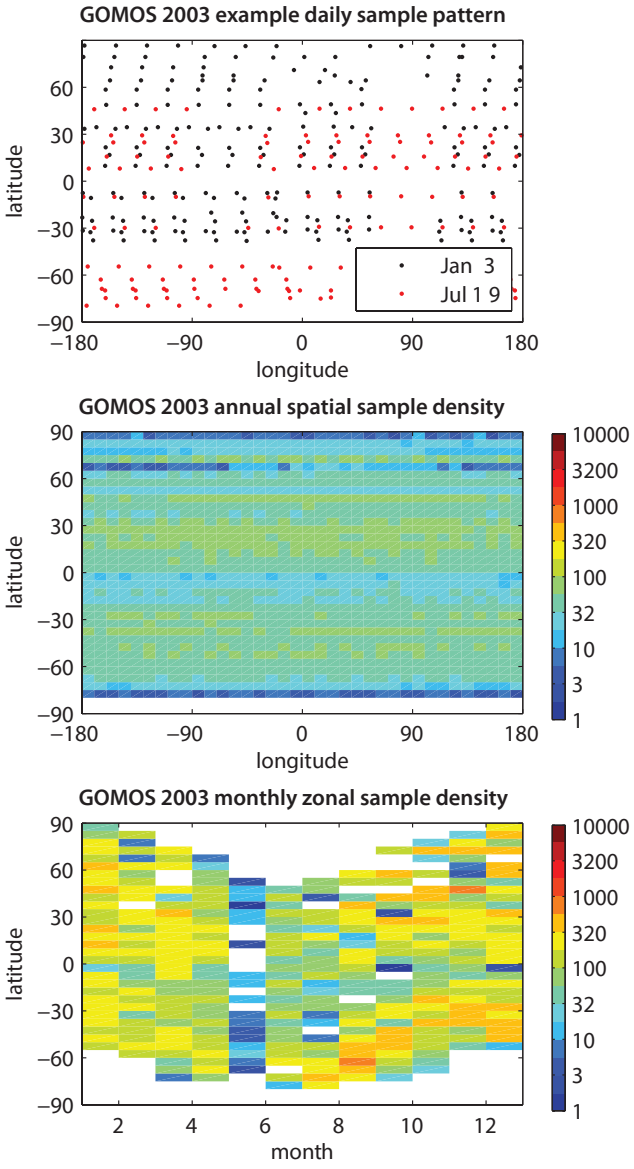


Figure 2.10: Sampling pattern and resulting sample density for GOMOS.

summer-time polar regions. Daytime occultations were also measured, but they are not used in the present work due their lower quality. Measurements start at 150 km and extend down to 5 km in cloudless conditions. The altitude-sampling resolution is 0.5-1.7 km and depends on the azimuth of the LOS with respect to the orbital plane. The nominal vertical resolution of the retrieved ozone profiles is 2 km below 30 km, 2-3 km between 30-40 km and 3 km above 40 km, and for other species about 4 km (see also *Section 3.1.3.8*). The instrument optical design was based on a 30-cm telescope that simultaneously fed UV-VIS and IR spectrometers, two fast photometers and two redundant star trackers. Spectra were recorded by CCD detectors. The UV-VIS spectrometer spectral range were 250-690 nm with 0.3 nm sampling and 0.9 nm resolution. The constituents retrieved are O_3, NO_2, NO_3, and aerosol. The IR spectrometer channels are 750-776 nm and 916-956 nm with 0.06 nm sampling and 0.1 nm resolution. IR data are used to retrieve O_2 and H_2O. Two fast (1 kHz) photometers at blue and red wavelengths were used to make the scintillation correction for the spectrometer data, retrieve high-resolution temperature profile and probe stratospheric turbulence.

The self-calibrating measurement principle with good vertical resolution and accurate vertical geo-location made GOMOS a good candidate to produce long time series and climatologies (see *Hauchecorne et al.* [2005], *Kyrölä et al.* [2006, 2010a, 2010b], *Vanhellemont* [2010]). However, difficulties with the pointing system in 2003, 2005 and 2009 have left some gaps in the data coverage. Noise levels of the CCDs increased steadily from the launch date, and this has led to a decrease in the quality of data over time. The sampling pattern and resulting measurement density from GOMOS can be seen in **Figure 2.10**.

The climatologies are constructed using GOMOS data from ESA processing Version IPF 5. The retrieval scheme is discussed in *Kyrölä et al.* [2010b]. The GOMOS constituent profile retrieval starts from the horizontal transmission spectra. Occultations are processed one at a time. The data processing is split into Level 1b and Level 2 stages. In Level 1b, dark charge removal and other instrumental corrections are performed and finally transmission spectra are constructed. Geo-location is determined starting from the satellite location and from the known direction of the star, and performing ray-tracing calculations with the atmosphere assumed to be the one given by the ECMWF data below 1 hPa and the MSIS90 climatology in the upper atmosphere. In Level 2 processing, the transmission spectra are first corrected for dilution caused by refraction and for modulations by scintillations. The fast photometer data are used in the scintillation correction. In case of off-orbital-plane occultations, the correction is not able to remove the scintillation modulation arising from isotropic turbulence in the LOS. The ozone retrieval, however, is only weakly sensitive to modulations by scintillations [*Sofieva et al.*, 2010]. Ozone as well as NO_2, NO_3, and aerosols are retrieved from the UV/VIS range 250-675 nm. The Rayleigh extinction is removed using the ECMWF+MSIS90 data. The UV/VIS retrieval is divided into two consecutive stages. In the spectral inversion the model transmission function is fitted by a non-linear Levenberg-Marquardt method to the transmissions. Because of perturbations caused by uncorrected isotropic scintillations, NO_2 and NO_3 retrievals are based on sub-iteration using the differential cross section method [see *Hauchecorne et al.*, 2005].

After spectral inversion the vertical inversion is performed using so-called onion-peeling method. The inversion is constrained using the target resolution Tikhonov method [*Sofieva et al.*, 2004]. For ozone the target vertical resolution is 2 km below 30 km and 3 km above 40 km. For other constituents the target vertical resolution is 4 km. An iteration loop over spectral and vertical inversion is performed in order to take into account the temperature dependence of the cross sections. The retrieval errors for constituent profiles depend on the brightness of the star measured. For ozone, the error depends also on the spectral type of the star. Data quality and error estimates of GOMOS are discussed in detail in *Tamminen et al.* [2010].

2.2.9 MIPAS on Envisat

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) was a mid-infrared Fourier transform limb emission spectrometer designed and operated for measurement of atmospheric trace species from space [*Fischer et al.*, 2008]. It was part of the instrumentation of Envisat (for launch and orbit details, see *Section 2.2.8*). MIPAS passed the equator in a southerly direction at 10am local time 14.3 times a day, observing the atmosphere during day and night with global coverage from pole to pole. The instrument's field of view was 30 km in the horizontal and approximately 3 km in the vertical direction. MIPAS covered the 4.3-15 μm region in five spectral bands: band A (685-970 cm^{-1}), AB (1020-1170 cm^{-1}), B (1215-1500 cm^{-1}), C (1570-1750 cm^{-1}), and D (1820-2410 cm^{-1}).

MIPAS operated during July 2002 – March 2004 at full spectral resolution of 0.035 cm^{-1} (unapodised) in terms of full width at half maximum. During this period, MIPAS recorded a rear-viewing limb sequence of 17 spectra each 90 seconds, corresponding to an along track sampling of approximately 500 km and providing about 1000 vertical profiles per day in its standard observation mode. Tangent heights covered then the altitude range from 68 down to 6 km with tangent altitudes at 68, 60, 52, 47, and then at 3 km steps from 42 to 6 km.

Due to problems with the interferometer-mirror-slide system, MIPAS performed few operations from April–December 2004. In January 2005 regular observations resumed, but with a reduced duty cycle and a reduced spectral resolution of 0.0625 cm^{-1}. These new measurements have the advantage that more spectra could be measured during the same time interval compared to the former "high"-spectral resolution observations. Tangent heights covered the range from 70 down to 6 km with tangent

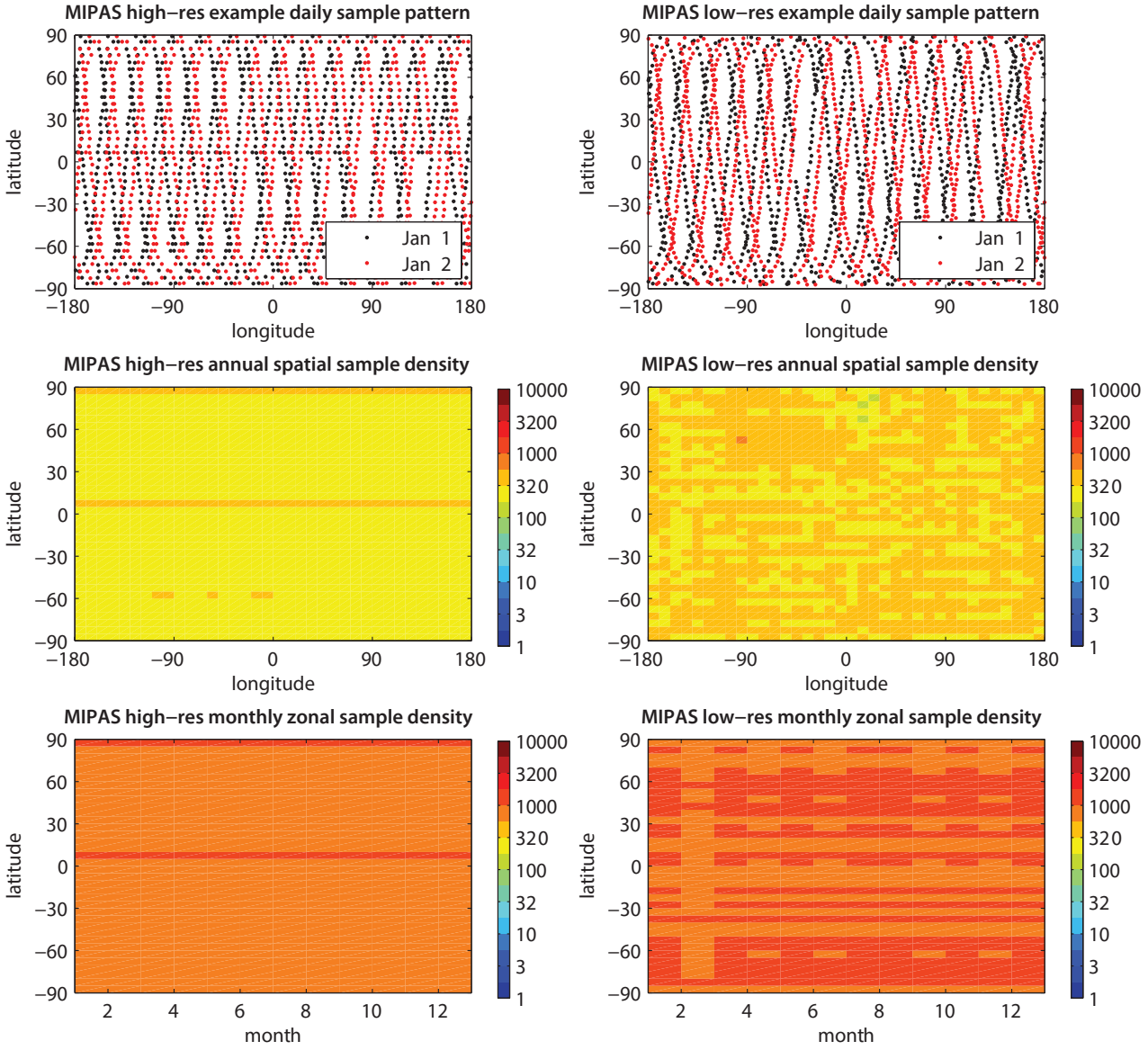


Figure 2.11: Sampling pattern and resulting sample density for MIPAS. *Left panels show results for the full (high)-spectral resolution mode from 2002-2004, right panels for the reduced (low)-spectral resolution mode from 2005-ongoing.*

altitudes at 70, 66, 62, 58, 54, 50, 46, 43, 40, 37, 34, 31, 29, 27, 25, 23, and then at 1.5 km steps from 21 to 6 km. Due to this modified measurement scenario the number of profiles increased by about 20%.

Trace gas profiles included in this climatology have been retrieved from calibrated geo-located limb emission spectra with the MIPAS Level 2 research processor developed and operated by the Institute of Meteorology and Climate Research (IMK) in Karlsruhe together with the Instituto de Astrofísica de Andalucía (IAA) in Granada. The general retrieval strategy, which is a constrained multi-parameter non-linear least squares fitting of measured and modelled spectra, is described in detail in *von Clarmann et al.* [2003c]. Its extension to retrievals under consideration of non-LTE (CO, NO, and NO_2) is described in *Funke et al.* [2001]. After wavenumber-recalibration, target quantities are retrieved sequentially, starting with temperature and LOS elevation (from CO_2 emissions around 15 μm), followed by the atmospheric main IR emitters H_2O, O_3, CH_4 and N_2O. Afterwards all other species are retrieved under consideration of the results of the preceding retrievals. Instead of the commonly used optimal estimation scheme, a Tikhonov-type first order regularisation is used [*Steck and von Clarmann*, 2001] because it does not constrain the column information but only how this information is distributed over altitude and, thus, does not push the mixing ratios towards *a priori* information. The strength of the regularisation is altitude dependent, with the aim of finding the best trade-off between the vertical resolution and the precision of the retrieved parameters. While trace gas abundances are retrieved in terms of VMR for most species, for some species (H_2O, NO_2, NO, CO), ln(VMR) is retrieved instead in order to better account for their pronounced temporal and spatial variability and reduce their dynamical range. Further, some target quantities (temperature and the trace gases NO, NO_2, and CO) are characterised by a pronounced spatial inhomogeneity, particularly close to transport barriers. In these cases, horizontal gradient profiles are taken into account within the retrieval [*Kiefer*, 2010]. In addition,

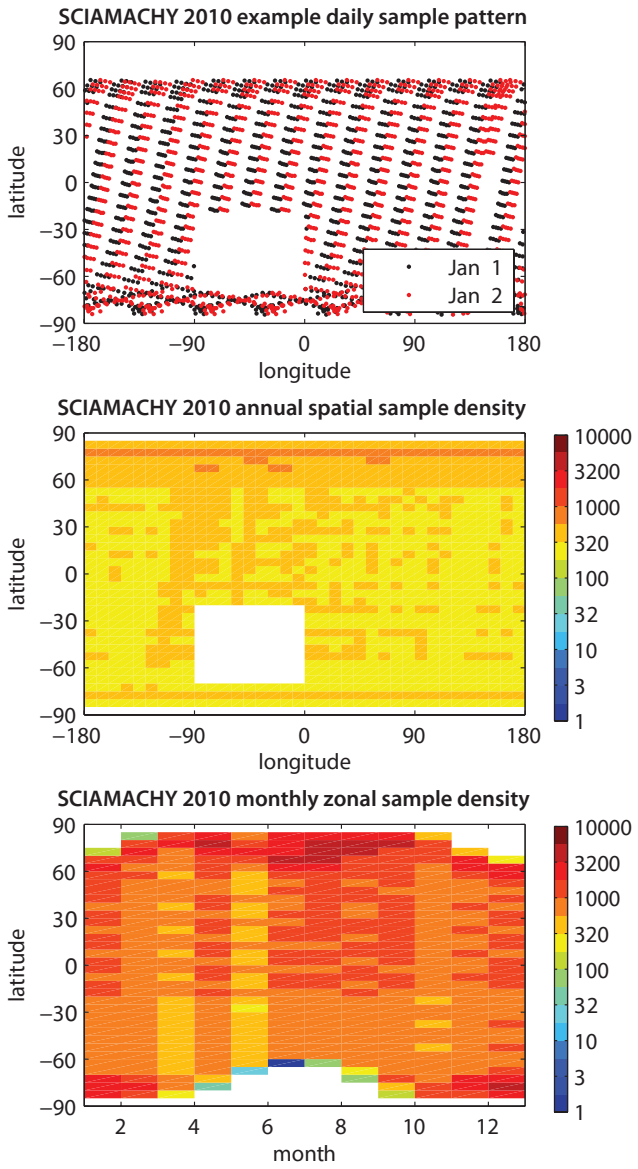


Figure 2.12: Sampling pattern and resulting sample density for SCIAMACHY.

a radiance offset and a continuum-like optical depth profile are fitted jointly for each microwindow in order to compensate for calibration errors and atmospheric contributions of weak wavenumber dependence not reproduced by the radiative transfer forward model [*von Clarmann et al.*, 2003c]. The MIPAS-IMK/IAA research data product, along with related diagnostics, is available to registered users *via* **http://www.imk-asf.kit.edu/english/308.php**. The sampling patterns and resulting measurement densities from MIPAS high and reduced spectral resolution measurement modes can be seen in **Figure 2.11**.

2.2.10 SCIAMACHY on Envisat

The Scanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) [*Burrows at al.*, 1995, *Bovensmann et al.*, 1999] was a payload on Envisat launched in March 2002 (for launch and orbit details, see *Section 2.2.8*). SCIAMACHY was one of the new-generation of space-borne instruments capable of performing spectrally-resolved measurements in several different modes: alternate nadir and limb observations of the solar radiation scattered by the atmosphere or reflected by the Earth's surface; and observations of the light transmitted through the atmosphere during solar or lunar occultation when feasible. The SCIAMACHY instrument was a passive imaging spectrometer comprised of eight spectral channels covering a wide spectral range from 214 to 2386 nm. Each spectral channel comprised a grating spectrometer, having a 1024-element diode array as a detector. Depending on the spectral channel the spectral sampling ranged from 0.11 to 0.74 nm and the spectral resolution from 0.22 to 1.48 nm.

This study uses SCIAMACHY measurements from scattered solar light in the limb-viewing geometry. In this geometry, the atmosphere was observed tangentially to the Earth's surface starting at about 4.5 km below the horizon (~1.5 km below the horizon since January 2011), *i.e.,* when the Earth's surface was still within the field-of-view of the instrument, and then scanning vertically up to the top of the neutral atmosphere (about 100 km tangent height). At each tangent height a horizontal scan of 1.5 s duration was performed followed by an elevation step of about 3.3 km. No measurements were performed during the vertical step. This results in a vertical sampling of 3.3 km. The vertical instantaneous field-of-view of the SCIAMACHY instrument was about 2.6 km at the tangent point. Although the horizontal instantaneous field-of-view of the instrument was about 110 km at the tangent point, the horizontal resolution was mainly determined by the integration time during the horizontal scan, reaching typically about 240 km. The entire distance at the tangent point covered by the horizontal scan was about 960 km. The along-track horizontal resolution was estimated to be about 400 km. In the nominal mode, about 100 measurements per orbit with 14 complete orbits per day were performed. Global coverage was achieved after six days. The sampling pattern and resulting data density for SCIAMACHY limb observations can be seen in **Figure 2.12**. The sampling pattern shown in **Figure 2.12** refers to standard retrievals with measurements at SZAs of up to 89°, resulting in a maximum latitude coverage of 65° in the winter hemisphere. This applies to all SCIAMACHY climatologies used in this study except for water vapour, for which only measurements at SZAs smaller than 85° are processed, resulting in a reduced latitude coverage of 55°. The gap in the sampling seen in the Southern Hemisphere is due to the South Atlantic anomaly. In this area the instrument electronics were exposed to an increased flux of energetic particles, which disturbed the measured signal resulting in a significant retrieval bias. This makes it necessary to reject the affected data when creating the climatologies (see *Section 3.1.3.10* for details).

Similar to other limb scattering instruments, the pointing uncertainty is a major error source. Currently, the accuracy of the pointing for the whole limb scan is estimated to be about 200 m. The relative pointing error between different tangent heights is negligible. The measurements at the lower tangent heights are affected by clouds; no retrievals

can be done in the presence of a cloud in the instrument field-of-view.

More general information on the SCIAMACHY instrument can be found at **http://envisat.esa.int/instruments/sciamachy/** and **http://www.iup.physik.uni-bremen.de/sciamachy/**.

Vertical profiles of atmospheric species and aerosol extinction coefficients included in this climatology are retrieved from SCIAMACHY limb measurements using the scientific processor developed and operated by the Institute of the Environmental Physics (IUP) at the University of Bremen. Depending on the species, several spectral sub-windows in UV, visible, or near-infrared spectral ranges are used. Retrievals of O_3 and aerosol extinction coefficients exploit radiance profiles averaged over several nanometer wide spectral windows, whereas NO_2, BrO, and H_2O algorithms gain information from the differential structure of the trace gas absorption bands (DOAS technique). All retrievals except for H_2O use the reference tangent height normalisation technique to reduce the influence of the solar Fraunhofer lines, instrument calibration errors, and radiation scattered in the lower troposphere or reflected from the underlying surface. The retrieval relies on the optimal estimation type technique including an additional smoothing constraint (first order Tikhonov term). The non-linearity of the inverse problem is accounted for by employing the Gauss-Newton iterative scheme.

For most species, the retrieval is done for number densities while for H_2O the logarithms of the number densities are retrieved. Details on the retrieval algorithms and validation results for different species can be found in *Rozanov et al.* [2005], *Ernst et al.* [2009], *Sonkaev et al.* [2009], *Bauer et al.* [2012], *Mieruch et al.* [2012], *Rozanov et al.* [2011a; 2011b]. The SCIAMACHY scientific products retrieved by IUP Bremen are available to registered users *via* **http://www.iup.physik.uni-bremen.de/scia-arc**. Except for the aerosol extinction coefficients, the results are provided along with the averaging kernels, retrieval precision, and cloud flags.

2.2.11 ACE-FTS on SCISAT-1

The Atmospheric Chemistry Experiment-Fourier Transform Spectrometer (ACE-FTS), on board the SCISAT-1 satellite, uses mid-infrared solar occultation to investigate the chemical composition of the atmosphere [*Bernath*, 2006]. The SCISAT-1 satellite was launched on 12 August 2003 and routine measurements began on 21 February 2004. The ACE-FTS instrument is a high-resolution (0.02 cm^{-1}) FTS measuring the full spectral range between 13.3 and 2.2 μm (750 and 4400 cm^{-1}) [*Bernath et al.*, 2005]. The ACE-FTS measures approximately 15 sunrise and 15 sunset occultations per day and achieves global latitude coverage over a period of three months (*e.g.*, one "season"). The sampling pattern and resulting measurement density from the ACE-FTS can be seen in **Figure 2.13**. These spectral measurements extend from

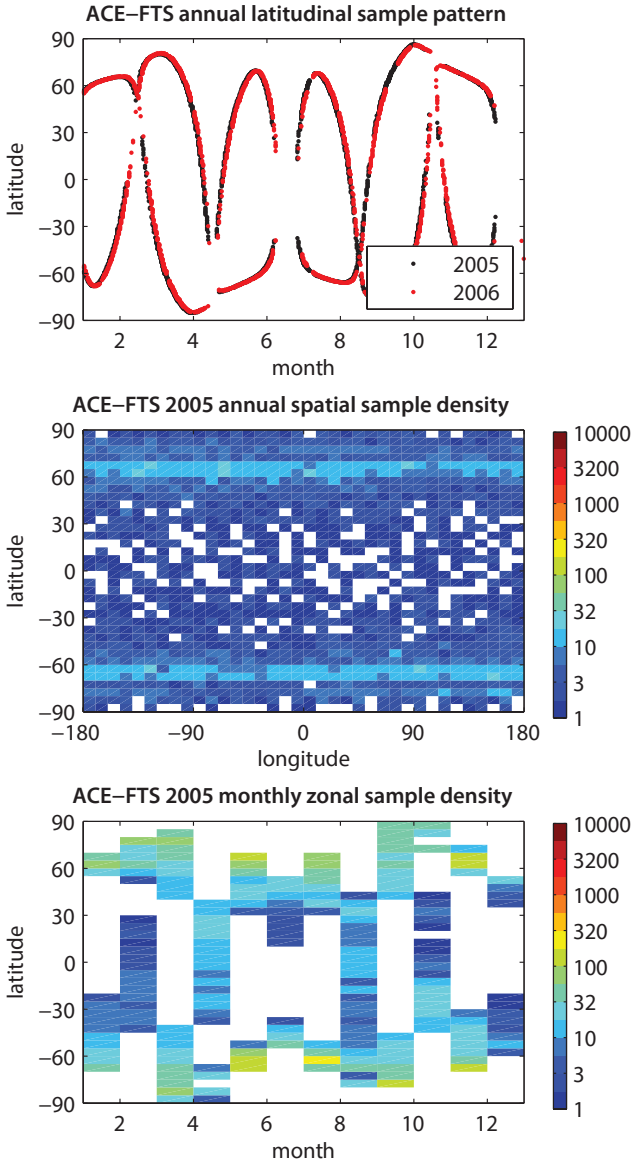


Figure 2.13: Sampling pattern and resulting sample density for the ACE-FTS. *The sampling corresponds to the year 2005, which is representative for all years.*

the cloud tops to 150 km. The vertical spacing between each 2-second ACE-FTS measurement varies between 1.5 and 6 km depending on the satellite's orbit geometry. The FOV of the instrument is approximately 3 km at the limb. Because of the high inclination of the SCISAT-1 orbit (74°), almost 50% of the occultation measurements made by the ACE-FTS are at latitudes of 60° and higher. The SCISAT-1 orbit was tuned to obtain a pattern of measurement latitudes that repeats each year. Thus, as noted below, the sampling pattern and density of measurements are representative for all years of the SCISAT-1 mission.

Exo-atmospheric and deep space spectra recorded during each occultation are used to calculate atmospheric transmission spectra from the ACE-FTS measurements. The use of transmission spectra provides "self-calibration" for these occultation measurements. It makes the ACE-FTS dataset less susceptible to changes over the mission and provides very good long-term stability in the measurements.

Level 2 constituent profiles are retrieved from the ACE-FTS transmission spectra in VMR using an unconstrained non-linear least squares global fitting approach [*Boone et al.*, 2005, and references therein]. In the first step, CO_2 lines in the spectra are used to determine the pressure and temperature as a function of altitude. The microwindows used for the retrieval cover the following wavenumber ranges: 932–937, 1890–1976, 2042–2073, 2277–2393, 2408–2448, 3301–3380, and 3570–3740 cm^{-1}. Temperature and pressure profiles are retrieved from the ACE-FTS spectra between 12 and 120 km. Below 12 km, meteorological results from the Canadian Meteorological Centre operational weather analysis and forecast system are used. Above 120 km, output from the Naval Research Laboratories MSISE-00 software is employed. The resulting temperature and pressure profiles are used to retrieve VMR profiles of over 30 trace gas species from sets of microwindows chosen to contain spectral features specific to each of the target molecules. The spectroscopic parameters used for these calculations are from the HITRAN 2004 linelist [*Rothman et al.*, 2005]. The retrieval algorithm uses first guess profiles taken from the four ATMOS missions on-board the Space Shuttle. However, the retrievals are not constrained by these first guess profiles. Currently, there is no error budget available for the ACE-FTS products. For each measurement, there is an associated fitting uncertainty provided. This one-sigma fitting uncertainty is the square root of the diagonal element of the covariance matrix obtained in the retrieval process [*Boone et al.*, 2005]. A document describing the microwindows used for the ACE-FTS retrievals is available from **http://www.ace.uwaterloo.ca** [ACE Report ACE-SOC-0020, Microwindow List for ACE-FTS retrievals – Version 2.2 + updates, Dec. 2006].

For the SPARC Data Initiative, the ACE-FTS Version 2.2 data products are used including updates for O_3 and N_2O_5. The validation results for these species and parametres are included in a special issue of Atmos. Chem. Phys. (**http://www.atmos-chem-phys.org/special_issue114.html**). In addition, two climatologies have been created for the 2004-2009 period using the Version 2.2 (plus updates) dataset: a climatology for O_3, H_2O, CH_4, N_2O, CO, NO, NO_2, N_2O_5, HNO_3, HCl, $ClONO_2$, CCl_3F, CCl_2F_2, and HF [*Jones et al.*, 2012] and an NO_y climatology derived from the ACE-FTS NO, NO_2, HNO_3, HNO_4, N_2O_5 and $ClONO_2$ products [*Jones et al.*, 2011]. Both are five-year zonal mean climatologies provided on a monthly and three-month basis. The Level 2 ACE-FTS data products are stored by occultation in ASCII format (main isotopologues and minor isotopologues are in separate files for each occultation). Further information about ACE-FTS and the ACE mission, including the Level 2 Version 2.2 data products, can be found at: **http://www.ace.uwaterloo.ca/**.

2.2.12 ACE-MAESTRO on SCISAT-1

The Measurement of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation (MAESTRO) is a dual UV/VIS/Near-IR spectrophotometer that is part of the Atmospheric Chemistry Experiment (ACE) mission on-board the SCISAT-1 satellite [*McElroy et al.*, 2007]. ACE-MAESTRO was designed to extend the ACE wavelength coverage to the 280-1030 nm spectral region using two spectrometers with overlapping coverage (280-550 nm, 500-1030 nm) to reduce stray light. Currently, it makes measurements of solar radiation between 450-1030 nm during each sunrise and sunset with a spectral resolution of 1-2 nm (depending on spectral region). The two ACE instruments take simultaneous measurements of the same air mass using a common sun-tracking mirror that is located within the ACE-FTS. During each occultation (sunrise or sunset measurement), approximately 60 spectra are measured by ACE-MAESTRO between the cloud tops and 100 km. The vertical spacing of these measurements varies from 300 m to 2 km at altitudes below 50 km and the spacing increases to every 5 km for altitudes above 50 km. The FOV of the instrument is approximately 1 km at the limb.

As noted in *Section 2.2.11*, the SCISAT-1 satellite was launched on 12 August 2003 and began routine measurements on 21 February 2004. The sampling pattern and resulting measurement density from ACE-MAESTRO is essentially identical to that of ACE-FTS (as shown in **Figure 2.13**). As can be seen in **Figure 2.13**, ACE-MAESTRO achieves global latitude coverage in its measurements over a period of three months.

Level 2 profiles of O_3, NO_2, and optical depth are retrieved from the ACE-MAESTRO measurements as a function of altitude using a differential optical method combined with an interactive Chahine relaxation inversion algorithm [*McElroy et al.*, 2007, and references therein]. In the first step, the raw data are converted to wavelength-calibrated spectra, corrected for stray light, dark current and other instrument parameters. Then, a non-linear least-squares spectral fitting routine is used to analyse the corrected spectra. These are used to calculate slant-path column densities for each spectrum and, from these, vertical profiles of O_3 and NO_2 VMRs are derived by adjusting an initial guess profile from a high-vertical-resolution model simulation. Pressure and temperature profiles used in the ACE-MAESTRO retrievals are obtained from the corresponding ACE-FTS occultation measurement and are used to fix the tangent heights for the ACE-MAESTRO retrievals. The retrieval algorithm does not require any *a priori* information or other constraints [*McElroy et al.*, 2007]. No error budget has been produced for the ACE-MAESTRO data products. For each measurement, there is an associated error provided. This is essentially the random error of the measurement and is produced by propagating the instrument noise through the spectral fitting and profile retrieval codes.

For this project, the ACE-MAESTRO Version 1.2 O_3 data products are used, which are available from: **http://www.ace.uwaterloo.ca**. The Level 2 ACE-MAESTRO data products for each species/parameter are stored individually by occultation in ASCII format. The validation of the ACE-MAESTRO ozone product was described by *Dupuy et al.* [2009] and is part of the special issue of Atmospheric

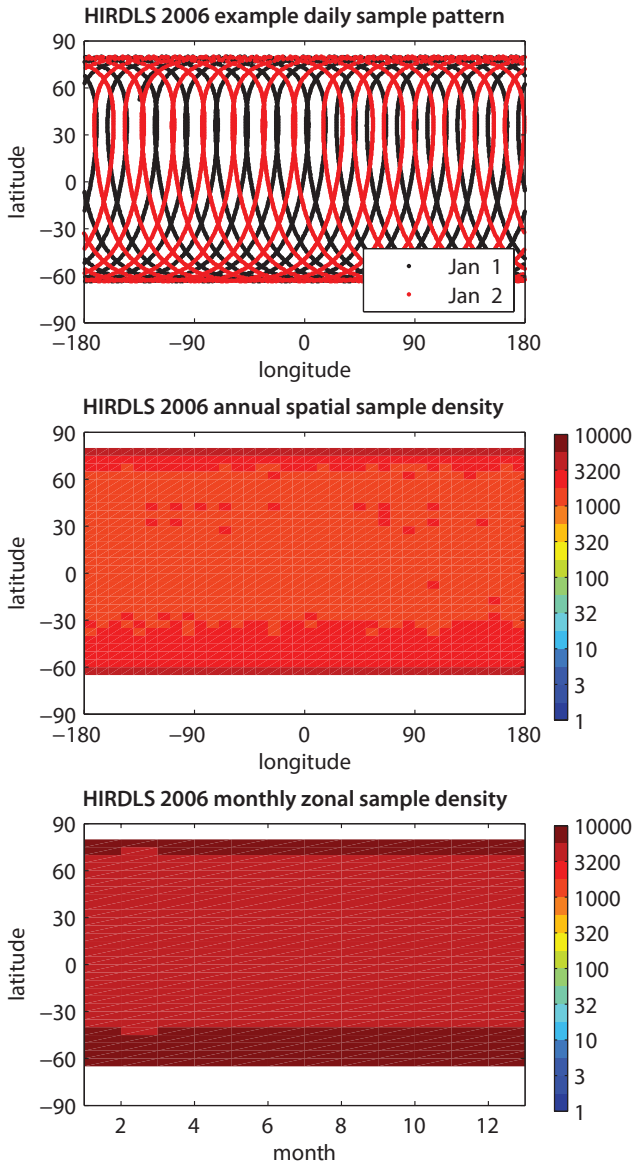


Figure 2.14: Sampling pattern and resulting sample density for HIRDLS.

Chemistry and Physics (described above: **http://www.atmos-chem-phys.org/special_issue114.html**).

2.2.13 HIRDLS on Aura

The High Resolution Dynamics Limb Sounder (HIRDLS) instrument is a 21-channel limb-scanning infrared radiometer, designed to scan from the upper troposphere into the mesosphere and provide data with 1-km vertical resolution. HIRDLS was launched on the Aura satellite into a polar orbit on July 15, 2004 (see *Section 2.2.15*). The original description of the experiment is by *Gille and Barnett* [1992]. Its channels cover the wavelength range from 6.12 to 17.76 μm, or 563 -1634 cm^{-1}, in order to measure emission features from 11 trace gases and from aerosols. Four channels measure emission by CO_2, from which temperature is recovered as a function of pressure. Three channels are dedicated to the retrieval of O_3, two to H_2O, and one each to CH_4, N_2O, NO_2, N_2O_5, HNO_3, $ClONO_2$, CFC-11 and CFC-12, with others for measurement of aerosols at four wavelengths. The large number of channels allows several to be dedicated to obtaining measurements in weaker parts of the bands, allowing sounding of the upper troposphere.

Unfortunately, HIRDLS was damaged during launch such that most of the aperture was covered. The blocking material is believed to be a thin film of plastic that became dislodged during launch and settled in the optical train, blocking 80-95% of the beams leading to the 21 detectors. In addition to blocking the aperture areas, this material gives off radiance signals that vary with scan angle and time on many scales. *Gille et al.* [2008] give a post-launch description of HIRDLS. Data coverage is from 63°S to 80°N, with profiles spaced every 100 km along the scan track, as shown in **Figure 2.14**. Because three of the retrieved species have diurnal variations, it is also important to know the local time of the retrievals on the ascending (northward) and descending portions of the orbit. *Gille et al.* [2008] also describes the initial corrections that resulted in Version 3 (V3) data. While details have changed, the procedure is the same for the V6 data discussed here.

Vertical coverage varies for each species, but the 1 km vertical resolution has been preserved. Persistent effort has resulted in successive improvements, leading to the release of the V6 data relevant to this report. They include temperature, O_3, HNO_3, CFC-11, CFC-12, and zonal means of day (1500 UT) and night (0 UT) NO_2, and night N_2O_5. All products for a single day are in a single file, on a grid of 24 pressure levels per decade of pressure, uniformly distributed in log pressure. More detailed characteristics of these data are included in the Data Description and Quality Document, Version 6, available from **http://archive-eos.acom.ucar.edu/hirdls/**, **http://disc.sci.gsfc.nasa.gov/data-holdings**, or **http://badc.nerc.ac.uk/browse/badc/hirdls**.

HIRDLS data are calibrated, corrected and retrieved in two major processors. In the first, the L1 processor, the conversion from raw counts to corrected radiances takes place, and the scans are geo-located. Subsequently they are corrected for the effects of the blocking material. The first correction is the removal of the small amplitude oscillations at ~1.8 Hz, which are initiated when the scan mirror contacts the plastic film during a scan. Next, the signal emitted by the film is removed, based on measurements made when Aura is pitched so that the complete HIRDLS scan is above the atmosphere, and only the film is viewed. Finally, the signal is corrected for the reduced effective aperture. Recent efforts have been made to model more closely the signal from the film, especially its change over the 3-year mission. After these corrections are made, the input radiances are on a nearly unifom elevation angle scale with a spacing that corresponds to ~200 m at the limb. They are then filtered to remove noise at spatial frequencies too high to be seen by HIRDLS, then splined onto an altitude grid with 1 km spacing. Channel 6 near 830 cm^{-1} is in the most transparent portion of the spectrum; it is used for the detection of clouds and aerosols. The altitude at which channel 6 radiances

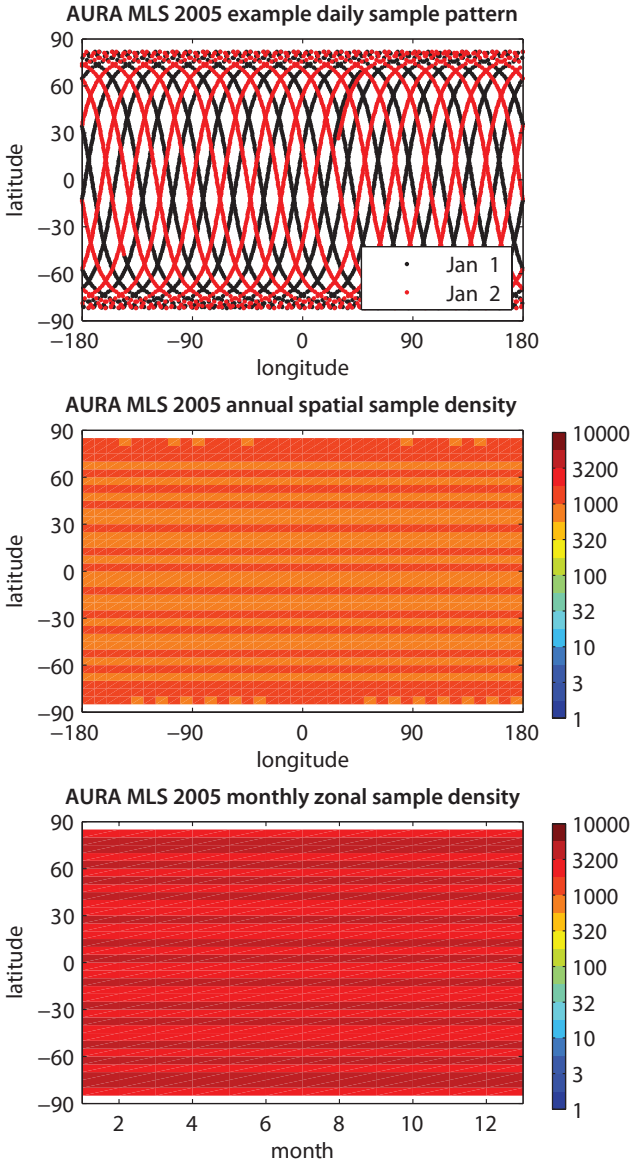


Figure 2.15: Sampling pattern and resulting sample density for Aura-MLS.

suddenly increase is tagged as the cloud top, but it is verified and possibly adjusted with data from channel 12. These are input to the second step, the L2 retrieval processor.

The retrieval algorithm is based on optimal estimation theory [*Rodgers*, 2000], using a modified Levenberg-Marquardt approach for the iterative solution. The application for HIRDLS is described in detail in *Khosravi et al.* [2009]. The L2 step accepts the conditioned radiance data from the L2CLD, where cloud top heights are determined, and performs the retrievals through a series of iterations. This code is designed to be flexible in handling combinations of radiance channels to retrieve the HIRDLS target species in a user-defined sequence. One of the major features is the use of ancillary data from the Goddard Earth Observing System Model (GEOS-5), produced by NASA's Global Modeling and Assimilation Office (GMAO) to determine temperature gradients along the LOS, which are incorporated to yield improved retrievals. This processor is described in detail in the L1-2 Algorithm Theoretical Basis Document (ATBD) available on the web at **http://archive-eos.acom.ucar.edu/hirdls/data/products/HIRDLS-DQD_V6-1.pdf**. GEOS-5 Version 5.01 data were used through January 2, 2008, after which Version 5.1 data were used.

2.2.14 MLS on Aura

Aura-MLS is a Microwave Limb Sounder (MLS) instrument, which is part of the Earth Observing System (EOS) and launched on the Aura satellite on 15 July 2004 (for orbit details, see *Section 2.2.15*). Aura-MLS, like its predecessor version on the UARS (see *Section 2.2.4*), measures microwave thermal emission day and night, simultaneously from several spectral regions, using an antenna that scans the Earth's atmospheric limb, in this case every 24.7 s.

Aura-MLS measures thermal emission from the limb in five broad spectral regions between 118 GHz and 2.5 THz. Aura-MLS views the atmosphere ahead of the Aura satellite, which is in a sun-synchronous near-polar orbit, with a ~1:45pm equatorial crossing time (ascending node). The MLS vertical scans are synchronised to the Aura orbit, leading to retrieved profiles at the same latitude every orbit, with a spacing of 1.5° great circle angle (about 165 km) along the sub-orbital track; the horizontal (along-track) resolution is limited by the smearing of sensitivity near the tangent point, including the impact of retrieval smoothing constraints, but typically ranges from 200 to 500 km in the stratosphere. The 240 limb scans per orbit provide almost 3500 profiles (per species) every day, from about 82°S to 82°N. The sampling pattern and resulting measurement density from Aura-MLS can be seen in **Figure 2.15**. The vertical retrievals are typically on a pressure grid with six levels (pressure surfaces) per decade change in pressure in the stratosphere and lower mesosphere; the main Version 3.3 exception relevant for (and used in) this report is the H_2O product, which is retrieved on a vertical grid that is twice as fine as that for most other species.

The Aura-MLS retrievals use the "optimal estimation" method [*Rodgers*, 1976; 2000]. This involves the nonlinear weighted least squares optimisation of a cost function describing the fit to observed radiance signals, including the use of *a priori* constraints for regularisation. Uncertainty estimates are provided as a result of the inversion process, based on input radiance uncertainties and *a priori* profile uncertainties. Gauss-Newton iteration is used, with a second order Tikhonov constraint [*Tikhonov*, 1963]; this constraint is applied to the profile second derivatives (vertically and horizontally). Specific retrieval aspects include adaptation to a two-dimensional system, using the LOS measurements from several scans to derive information about several profiles [see *Livesey and Read*, 2000]. The various species are retrieved from overlapping "chunks" of observations, typically consisting of a 15° span of great circle angle (about ten vertical scans). Several retrieval "phases" are performed in sequence, each using a different set of measured radiances; some phases retrieve temperature and pressure, and some include this information from an earlier phase. In

the MLS retrieval system, the state vector represents vertical profiles of mixing ratios in a piecewise-linear manner. The only exception is water vapour, where the representation is piecewise-linear in the logarithm of the VMR, nevertheless the retrieved quantity is VMR.

The Aura-MLS Level 2 data files include various screening flags to provide users with information about instrument and retrieval status (and also about quality of fit and retrieval convergence) based on various criteria described in the species-specific V2.2 validation papers. The updated V3.3 retrievals used to generate the SPARC Data Initiative climatologies for this report (except for ozone, for which V2.2 is recommended overall because of vertical oscillations that exist primarily in V3.3 in the UTLS at low latitudes) follow generally the same (V2.2) flagging/screening methodologies and recommendations for data usage, albeit with some changes in the screening flag threshold values. Also, the V3.3 data for CO and HNO_3 (used here) are more sensitive to cloud effects than previous versions, as a result of changes in the retrieval approach and the vertical range used in the UT. The resulting data screening methods include the removal of negative outliers (spikes of a certain size), as described in the V3.3 Aura MLS Data Quality documentation [*Livesey et al.*, 2011]. The impact of clouds depends on cloud thickness and altitude, and this mostly affects the species retrieved at low latitudes in the UTLS. The fraction of (daily) discarded profile values in these regions is typically 5 to 10%, and occasionally more than 20%. Different sensitivities to clouds can, in effect, lead to sampling biases between instruments; other satellite sensors are typically more affected by clouds and humidity than those in the microwave region. More details about the Aura-MLS retrieval approach are provided by *Livesey et al.* [2006] and calculation specifics of the Aura-MLS radiance model ('forward model') are described by *Read et al.* [2006] and *Schwartz et al.* [2006]. *Waters et al.* [2006] provide a detailed description of the Aura-MLS instrument's characteristics, spectral bands, and geophysical profile measurements.

Level 2 daily profiles are stored in Level 2 data files (one file per parameter) in Hierarchical Data Format (HDF-EOS 5 format type), and available from the NASA Goddard Spaceflight Center Distributed Active Archive Center (DAAC), specifically the Goddard Earth Sciences (GES) Data and Information Services Center (DISC), at the Mirador website, namely **http://mirador.gsfc.nasa.gov**. Information about MLS, data access, and MLS-related publications, can be found at the MLS website (**http://mls.jpl.nasa.gov**). The Aura-MLS data quality documentation [*Livesey et al.*, 2011] is available at **http://mls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf**.

2.2.15 TES on Aura

The Tropospheric Emission Spectrometer (TES) is a Fourier Transform Spectrometer that was launched on the NASA Earth Observing System (EOS) Aura satellite in 2004 [*Beer*, 2006; *Beer et al.*, 2001]. The Aura satellite has a 705 km

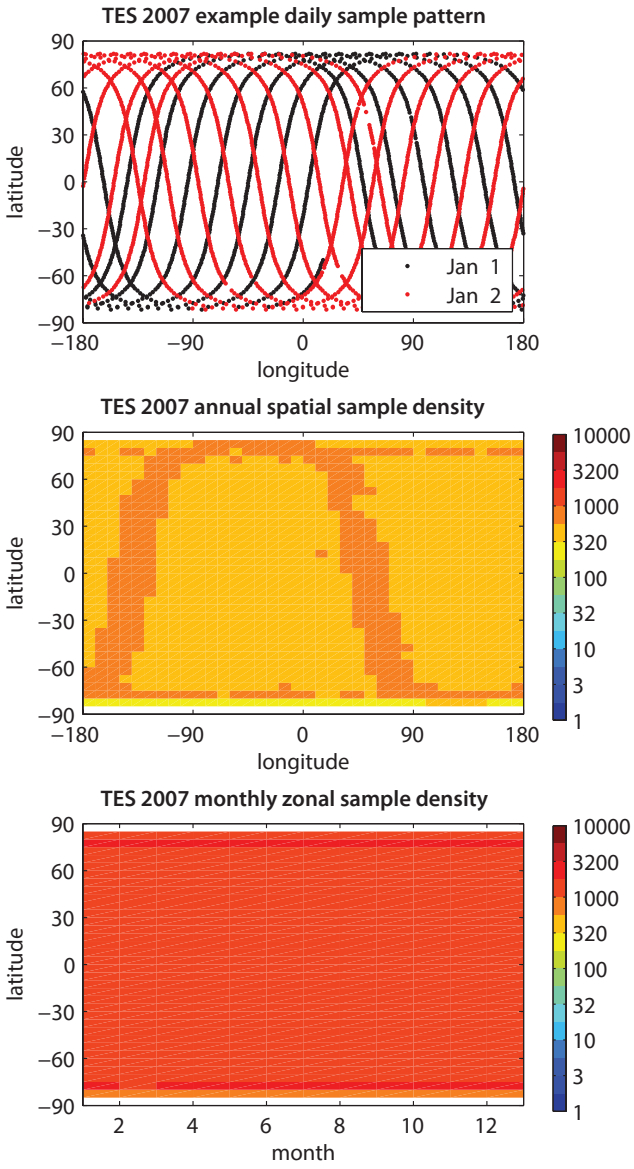


Figure 2.16: Sampling pattern and resulting sample density for TES before June 2008.

sun-synchronous polar orbit with an inclination of 98.21°, which provides global coverage from 82°S to 82°N with equator crossing times of 1:43pm (ascending node) and 1:43am (descending node) and a 16 day repeat cycle. TES measures spectrally-resolved thermal infrared radiation (650-3050 cm^{-1}) with a spectral resolution of 0.06 cm^{-1} (unapodised) in the nadir mode. TES covers this spectral range with four filters: 2B1 (650-900 cm^{-1}), 1B2 (950-1150 cm^{-1}), 2A1 (1100-1325 cm^{-1}), and 1A1 (1900-2250 cm^{-1}), and measures surface and atmospheric temperature as well as a variety of trace gases including O_3, CO, H_2O, HDO, CH_4, CO_2, NH_3, CH_3OH, and HCOOH, with greatest sensitivity in the troposphere. TES supports both nadir and limb scanning modes, but the limb measurements were discontinued in May 2005 in order to extend the life of the instrument. TES observes multiple spectra through a linear array of 16 pixels. At the nadir, the spatial resolution of each pixel is 0.5 x 5 km and is averaged to a footprint of 5.3 x 8.5 km, with a separation of ~182 km. TES is a pointable instrument and can access any target within 45° of the local vertical,

allowing for more tightly spaced measurements during Special Observations modes. Here we use only the standard nadir-viewing Global Survey O_3 measurements, with near-global coverage in 16 orbits (~26 hours) (see **Figure 2.16**). In cloud-free conditions, TES nadir O_3 profiles have approximately 4 degrees of freedom for signal, with ~2 in the troposphere and ~2 in the stratosphere (below ~5 hPa). This is equivalent to a vertical resolution of ~6-7 km.

TES sampling has changed over the mission lifetime in response to instrument aging. To extend the life of the instrument, the latitudinal coverage was reduced in June 2008 to 60°S-82°N and in July 2008, to 50°S-70°N. From January to April 2010, the instrument went offline due to problems with the scanning mechanism. When operations resumed in May 2010, the latitude coverage was further reduced to 30°S-50°N and the calibration strategy was changed from multiple black body scans per orbit to two sets of black body scans per day to reduce wear on the pointing mechanism of the instrument. This reduction in the number of calibration scans resulted in a 25% increase in the number of observations per global survey and regular but non-uniform spacing between the measurements (ground track separation cycles through 56 km, 195 km, 187 km, and 122 km and then returns to 56 km). A second data gap of approximately three weeks occurred in October 2010, with only two Global Surveys conducted that month. Since April 2011, data gaps became more common as the instrument continues to age.

TES retrievals and error estimation are described in *Worden et al.* [2004], *Bowman et al.* [2002, 2006], and *Kulawik et al.* [2006a]. The optimal estimation retrieval method that is used [*Rodgers*, 2000] is based on minimising the difference between observed radiances and a radiative transfer model subject to *a priori* constraints. Use of optimal estimation provides detailed characterisation of the smoothing, random, and systematic errors for the target parameters as well as important retrieval metrics such as degrees of freedom, information content, and vertical resolution. The radiative transfer model is referenced with respect to the logarithm of pressure (67 levels with a geometric layer thickness of 0.6-0.8 km from 100-1 hPa and 1.5 km above 1 hPa), with surface temperature, emissivity, and clouds included in the forward model. Spectral windows are selected to reduce the computational load and minimise systematic errors from non-retrieved atmospheric parameters. The TES retrieval strategy begins with updates to surface temperature and cloud parameters based on brightness temperature in a window region near 10 μm. Ozone is jointly retrieved with water vapour (both in ln(VMR) to account for their large dynamic range) following CO_2 and temperature on a subset of the 67-level forward model pressure grid. Each retrieval step includes the constituent of interest, interferents, and cloud and surface parameters, and the subset of vertical levels is chosen so as to capture the expected vertical variations of the retrieved trace gas. *A priori* profiles for temperature and water vapour are taken from the GEOS global circulation model of NASA's Global Modeling and Assimilation Office (GMAO), and initial profiles for O_3 are taken from the MOZART Chemistry-Transport Model [*Brasseur et al.*, 1998, *Park et al.*, 2004]. Constraint matrices are based on the altitude-dependent Tikhonov constraint and covariances from MOZART [*Kulawik et al.*, 2006b]. The least squares minimisation is based on the trust-region Levenberg-Marquardt algorithm [*Moré*, 1977] and subject to the constraint that the estimated state must be consistent with *a priori* probability distribution for that state. TES data, including averaging kernels and error covariance matrices, are publicly available. For more information, see **http://tes.jpl.nasa.gov/**.

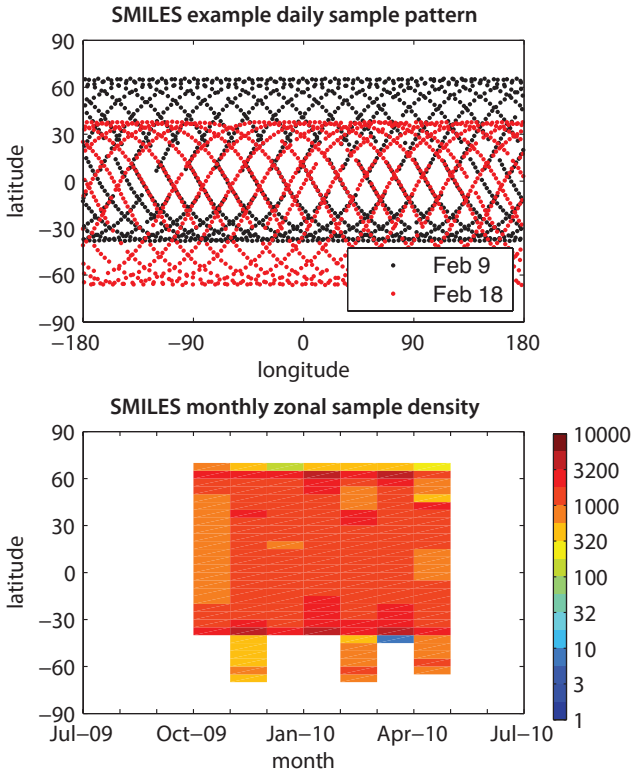


Figure 2.17: Sampling pattern and resulting sample density for SMILES.

2.2.16 SMILES on the ISS

SMILES (Superconducting Submillimeter-Wave Limb Emission Sounder) was selected as a first Earth observation mission for the Exposed Facility (EF) of the Japanese Experiment Module (JEM) on the International Space Station (ISS) in 1997, where it was installed on 25 September 2009. The purpose of the SMILES instrument was the demonstration of the ultra sensitive sub-mm limb emission observation with a 4-K cooled receiver system [*Kikuchi et al.*, 2010]. SMILES targeted atmospheric constituent observations such as for O_3, O_3 isotopomers, O_3 in the vibrational excited state, $H^{35}Cl$, $H^{37}Cl$, ClO, HNO_3, CH_3CN, HOCl, HO_2, and BrO in the stratosphere and mesosphere. Water vapour and ice clouds were observed in the UTLS. H_2O_2 and HOBr were also observed in the stratosphere and mesosphere although their spectrum signals are weak. The non-sun-synchronous orbit of the ISS allowed the instrument to observe the diurnal variation of these minor species. Observations

Table 2.3: Data versions of SMILES research products.

Data	Version	Availability	Comments
L1b	006	Feb 2011	Improved: AOS response function Problem: Calibration non-linearly
	007	June 2011	Improved: Calibration non-linearly Problem: Tangent height and latitude/longitude
L2r	201	Nov 2011	L1b data: Version 006 Altitude region: 24-90 km Problem: Spectrum calibration non-linearly
	215	Oct 2011	L1b data: Version 007 Altitude region: 12-90 km Improved: Spectrum calibration non-linearly Problem: Tangent height and latitude/longitude

were made by SMILES between 12 October 2009 and 21 April 2010, when SMILES stopped operations due to the failure of the local oscillator and the 4-K cooler. SMILES is a cooperative mission of the Japan Aerospace Exploration Agency (JAXA) and the National Institute of Information and Communications Technology (NICT). Details of the mission are described in the SMILES Mission Plan, Version 2.1, (**http://smiles.nict.go.jp/Mission_Plan/**).

The platform (ISS) altitudes are typically between 350 and 400 km. The altitude region of the antenna scan is from -10 to 120 km (nominal). The scanning altitude region of the observation was changed due to the change of the altitude, rotation, and vibration of the ISS platform. The antenna FOV is 0.009° (about 3-4 km). The SMILES instrument employs two superconductor-insulator-superconductor (SIS) mixers cooled at about 4 K and high-electron-mobility-transistor (HEMT) amplifiers at 20 and 100 K, cooled by a mechanical cryo-cooler. SMILES has two spectrometers. There are three observation frequency regions, band A: 624.32-625.52 GHz, band B: 625.12-626.32 GHz, and band C: 649.12-650.32 GHz. The transition of O_3 at 625.37 GHz was observed in both bands A and B for comparison/validation purposes. $H^{35}Cl$ was observed in band B, and $H^{37}Cl$ was observed in band A. Details of the frequency allocation are described in the SMILES Mission Plan, Version 2.1 (**http://smiles.nict.go.jp/Mission_Plan/**).

Since the ISS orbit is circular, with an inclination of 51.6°, the highest latitude reached by the ISS orbit is 52°N and S. To extend the latitudinal coverage to the northern higher latitudes, the SMILES antenna is mounted so that its FOV is 45° to the left of the orbital plane. The observed latitude region was between 38°S and 65°N (nominal). 1630 observation points were obtained per day, resulting in a sampling pattern as shown in **Figure 2.17**.

Trace gas profiles used for the SPARC Data Initiative climatologies have been retrieved from calibrated limb emission spectra with the SMILES Level 2 research processing system developed and operated by the National Institute of Information and Communications Technology (NICT). NICT developed an algorithm, named AMATERASU, to retrieve the vertical profiles of the atmospheric constituents from the calibrated limb emission spectra in the frequency region 624.32-625.52 GHz, 625.12-626.32 GHz, and 649.12-650.32 GHz. The maximum *a posteriori* (MAP) method with the Gauss-Newton interactive procedure modified by Levenberg has been adopted as the retrieval for O_3, HCl, ClO, HNO_3, CH_3CN, HOCl, HO_2, H_2O_2, BrO, and HOBr in the stratosphere and mesosphere, as well as H_2O and ice- cloud in the UTLS. Observations from the ISS generally suffer pointing problems. While the pointing information and temperature are commonly retrieved from molecular oxygen lines, there is no oxygen line in the SMILES spectral region. The retrievals of the LOS elevation angles and of temperature have been obtained from the strong ozone line at 625.371 GHz. For Version 2.1.5, it was pointed out by *Baron* [2011] that "The pointing parameters and the ozone profiles are retrieved from the line wings which are measured with high signal to noise ratio, whereas the temperature profile is retrieved from the optically thick line center. The main systematic component of the retrieval error was found to be the neglect of the non-linearity of the radiometric gain in the calibration procedure. This causes a temperature retrieval error of 5-10 K. Because of these large temperature errors, it is not possible to construct a reliable hydrostatic pressure profile. However, as a consequence of the retrieval of pointing parameters, pressure induced errors are significantly reduced if the retrieved trace gas profiles are represented on pressure levels instead of geometric altitude levels". The Level 2r Version 2.1.5 products for the SPARC Data Initiative suffer from the non-linearity problem of the radiometric gain in the calibration procedure of the spectrum. The error of the latitude-longitude position was estimated to be of the order of about 10-50 km.

The AMATERASU algorithm Version 2 series including V2.0.1 and V2.1.5 used only the clear-sky part of the radiative transfer calculation. The continuum component used the modified Pardo approach [*Pardo et al.*, 2001]. Although the continuum including H_2O is retrieved, the altitude region between about 16 and 90 km is maintained for the atmospheric composition. The "definitive window method" is used for the retrieval frequency range in order to obtain more accurate values in the stratosphere and mesosphere. The details of the retrieval method are described in *Baron et al.* [2011] for the Version 2 series of the SMILES research products, and the evaluation and validation status are discussed further in *Sato et al.* [2012] and *Kasai et al.* [2013].

The SMILES research data product is, along with related diagnostics, available to registered users *via* **https://data.smiles.nict.go.jp/products/research_latitude-longitude.jsf**.

Chapter 3: Climatology framework

This chapter discusses the datasets evaluated within the SPARC Data Initiative, including information on how the climatologies were constructed, and on the diagnostics used to evaluate them. Note that here we use the term 'climatology' for monthly mean zonal mean cross sections. The evaluations are based on single year cross sections, or on multi-year means compiled over particular reference periods. The resulting climatologies may be single-year or multi-year monthly or annual means.

Monthly zonal mean time series have been calculated for each trace gas species and aerosol listed in **Table ES.1** (*Executive Summary*) on the SPARC Data Initiative climatology grid, using 5° latitude bins (with mid-points at -87.5°, -82.5°, -77.5°, …, 87.5°) and 28 pressure levels (300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa). Trace gas species are reported as volume mixing ratios (VMR), and aerosols as extinction coefficients. The monthly zonal mean value and the 1σ standard deviation, along with the number of averaged data values are given for each month, latitude bin and pressure level. The mean, minimum, and maximum local solar time (LST), average day of the month, and average latitude of the data within each bin for one selected pressure level are also provided.

For species with large diurnal variations we separate the measurements based on LST (see detailed discussion in *Section 3.1.1*). Additional climatologies are built using a photochemical box model to scale the measurements to a common LST in order to enable direct comparison between products from different instruments with different sampling patterns. All satellite-based measurements of trace gas species are imperfect estimates of the truth characterised by measurement errors. The compilation of climatologies from these measurements can introduce additional errors such as sampling biases produced by non-uniform spatial or temporal sampling, or by the use of different filtering techniques. Biases can also be introduced by applying different averaging techniques.

The climatology construction, including common methodology and information specific to each instrument, is described in *Section 3.1*. A discussion of climatology uncertainties is provided in *Section 3.2*, while the diagnostics used to evaluate the trace gas climatologies are explained in *Section 3.3*.

3.1 Climatology construction

3.1.1 Methodology

The original data products are first interpolated to the SPARC Data Initiative pressure grid using a hybrid log-linear interpolation. For instruments providing data on an altitude grid, a conversion from altitude to pressure levels is performed using retrieved temperature/pressure profiles or meteorological analyses (ECMWF, GEOS-5, or NCEP, see **Table 3.1** for detailed information). The same pressure and temperature profiles are used to convert data products retrieved as number densities to VMR.

Original data have been carefully screened according to recommendations given in relevant quality documents, in the published literature, or according to the best knowledge of the involved instrument scientists. Monthly zonal mean products are calculated as the average of all of the measurements on a given pressure level within each latitude bin and month. An exception is MIPAS, for which measurements are interpolated to the centre of the latitude bin after averaging (see *Section 3.1.3.9* for details). For some species and instruments, averaging was done in $\log_{10}$(VMR) space. The 1σ standard deviation along with the number of averaged data values are also given for each month, latitude bin and pressure level. If not otherwise mentioned, a minimum of five measurements within the bin is required to calculate a monthly zonal mean for each instrument. The mean, minimum, and maximum LST, average day of the month, and average latitude of the data within each bin are provided for one selected pressure level for each latitude bin and month. Instrument-specific information for the calculation of the monthly zonal mean values is given in *Section 3.1.3*.

For species with large diurnal variations the monthly zonal mean climatologies cannot be compared directly since the LST of the measurements can differ from instrument to instrument, and between seasons and latitudes for the same instrument. Two types of climatologies are produced for diurnally varying species; climatologies from observations binned by LST (unscaled), and climatologies from observations scaled to a common LST. Most of the instruments measure two distinct LSTs per latitude. These instruments are in polar sun-synchronous orbits, with one LST for the ascending portion of the orbit and one for the descending portion, or in the case of sun-synchronous solar occultation

Table 3.1: Instrument specifications relevant for the climatology construction.

Instrument	Latitudinal coverage	LT at equator[1]	LT of measurement[2]	Inc.[3]	Vert. Grid[4]	Alternate grid[5]	Meas.[6]	Conversion to VMR[7]	Data density per day
LIMS on Nimbus 7	64°S–84°N (daily)	a: 11:51am d: 11:51pm	a: 1pm d: 11pm	99.3°	p	N/A	VMR	N/A	3000
SAGE I on AEM-B	75°S–75°N (~one month)	N/A	N/A	56°	z	NCEP	ND	NCEP	30
SAGE II on ERBS	75°S–75°N (~one month)	N/A	N/A	57°	z	NCEP	ND	NCEP	30
SAGE III on Meteor-3M	60°S–30°S 40°N–80°N (~over one season)	a: 9:30am d: 9:30pm	N/A	99.6°	z	NCEP	ND	NCEP	30
HALOE on UARS	75°S–75°N (~over one season)	N/A	N/A	57°	p	N/A	VMR	N/A	30
UARS-MLS on UARS	80°S–80°N (~over two months)	N/A	N/A	57°	p	N/A	VMR	N/A	1318
POAM II on SPOT-3	88°S–63°S 55°N–71°N (over one year)	a: 10:30pm d: 10:30am	N/A	98.7°	z	UKMO analysis	ND	UKMO analysis	30
POAM III on SPOT-4	88°S–63°S 55°N–71°N (over one year)	a: 10:30pm d: 10:30am	N/A	98.7°	z	UKMO analysis	ND	UKMO analysis	30
OSIRIS on Odin	82°S–82°N (daily, no winter hemisphere)	a: 6:30pm d: 6:30am	a: 6:30pm d: 6:30am	97.8°	z	ECMWF operational analysis	ND	ECMWF operational analysis	300–975
SMR on Odin	83°S–83°N (daily)	a: 6:30pm d: 6:30am	a: 6:30pm d: 6:30am	97.8°	p	N/A[8]	VMR	N/A	600–975
GOMOS on Envisat	90°S–90°N (daily, no summer poles for night)	a: 10:00pm d: 10:00am	a: 10-12pm d: 8-10:30am	98.55°	z	ECMWF operational analysis	ND	ECMWF operational analysis	100–300 (night measurements)
MIPAS on Envisat	90°S–90°N (daily)	a: 10:00pm d: 10:00am	a: 10:00pm d: 10:00am	98.55°	z	MIPAS	VMR	N/A	1000 (1300 since 2005)
SCIAMACHY on Envisat	85°S–85°N (65° for winter hemisphere)[9]	a: 10:00pm d: 10:00am	d: 10:00am	98.55°	z	ECMWF operational analysis	ND	ECMWF operational analysis	364–1456
ACE-FTS on SCISAT-1	85°S–85°N (~over one season)	N/A	N/A	74°	z	ACE-FTS	VMR	ACE-FTS	30
ACE-MAESTRO on SCISAT-1	85°S–85°N (~over one season)	N/A	N/A	74°	z	ACE-FTS	ND	ACE-FTS	30
HIRDLS on Aura	65°S–82°N (daily)	a: 1:43pm d: 1:43am	a: 2:57pm d: 0:30am	98.21°	p	N/A	VMR	N/A	5600
MLS on Aura	82°S–82°N (daily)	a: 1:43pm d: 1:43am	a: 1:25am d: 1:25pm	98.21°	p	N/A	VMR	N/A	3500
TES on Aura	82°S–82°N (daily) (50°S–70°N for 2008/09; 30°S–50°N for 2010)	a: 1:43pm d: 1:43am	a: 1:43pm d: 1:43am	98.21°	p	N/A	ln(VMR)	N/A	3145 (2126 for 2008/09; 1890 for 2010)
SMILES on ISS	38°S–65°N (daily)	N/A	N/A	51.6°	p	N/A[8]	VMR	N/A	1620

[1] Local time of equator crossing for satellites with sun-synchronous orbit (a=ascending, d=descending)
[2] Local time of measurement made at equator crossing for satellites with sun-synchronous orbit (a=ascending, d= descending)
[3] Inclination of the orbital plane
[4] Vertical grid used for retrieval of species (altitude '*z*' or pressure '*p*')
[5] Data used for conversion to alternate vertical grid
[6] Measure of species: volume mixing ratio (VMR) or number density (ND)
[7] Pressure/temperature data used for conversion from number density to volume mixing ratio
[8] For SMR and SMILES the tangent-pressure is retrieved but Level 2 data are provided on altitude grids. Conversion between *p* and *z* is done using ECMWF (for SMR) or GEOS-5 (for SMILES) data.
[9] 55° for winter hemisphere for water vapour climatologies

sounders, with measurements at sunrise and sunset as seen from the satellite. For the latter, the LSTs shift with the day of year. Climatologies of diurnally varying trace gases from instruments in a sun-synchronous orbit are generally based on measurements separated into ante meridiem (am) and post meridiem (pm) data. A representative LST can be assigned to each month and latitude bin. However, in some cases the LST variations between season and latitude bin must be considered. Instruments that observe from non-sun-synchronous orbits are characterised by drifting observation times with respect to LST. Climatologies for these instruments are generally separated into daytime and night-time measurements. Climatologies of diurnally varying trace gases from non-sun-synchronous solar occultation measurements are based on data separated into local sunrise and sunset measurements. Additional climatologies are compiled using a photochemical box model to scale the measurements to a common LST, as explained in more detail in *Section 3.1.2*. For chemical families (NO_x, *Section 4.1.12*, and NO_y, *Section 4.1.17*) the total family abundance is derived using all members of the family available from the instrument, supplemented with species derived from a photochemical box model if needed.

3.1.2 Local time scaling

For species with large diurnal variations additional climatologies are compiled by scaling the measurements with a photochemical box model to a common LST. The scaled climatologies enable a direct comparison between products from different instruments with different sampling patterns. For the diurnally varying species NO, NO_2, NO_x and BrO scaled climatologies are calculated for 10am and 10pm, the approximate local time of the MIPAS measurements at the equator. The ClO climatologies are scaled to 1:30am and 1:30pm, which is the approximate local time of the Aura-MLS measurements (for ~60°S-60°N).

A derivative of the University of California, Irvine photochemical box model [*Prather,* 1992; *McLinden et al.,* 2000; *McLinden et al.,* 2010] was applied to calculate the diurnal scaling factors used to map the VMR of a diurnally varying species from one local time (LST_1) to another (LST_2). This was done by scaling the measured $VMR(LST_1)$ by the model-calculated ratio $VMR(LST_2)/VMR(LST_1)$, which will be referred to as scaling factor in the following text. The VMR at the new local time is then derived as:

$$VMR(LST_2) = VMR(LST_1)\,[VMR(LST_2)/VMR(LST_1)]_{model}$$

The scaling factors are calculated with the photochemical box model based on LST, temperature, surface albedo and concentration of various trace gases (O_3, N_2O, NO_y, CH_4, Cl_y, Br_y). With these parameters specified, all remaining species are calculated to be in a 24-hour steady state by integrating the model for 30 days (fixed to the prescribed Julian day and latitude). The kinetic reaction rate coefficients and photochemical data used by the box model are based on JPL-06 and JPL-09 recommendations.

The model-calculated scaling factors were provided as a function of altitude, latitude, day of year, and LST as lookup tables. The calculations were based on the photochemical box model initialised with climatological inputs. Each table consists of 25 pressure-altitudes, from 10 to 58 km in 2 km increments, with pressure-altitude $z^*=-16\ \log_{10}(p/1000)$, p given in hPa, and z^* given in km. The latitude grid ranges from 77.5°S to 77.5°N in 2.5° increments. Tables are given for the 1st, 11th, and 21st of each month for 34 local times spanning 24-hours (fewer for polar regions). The input data includes O_3 and temperature from measurement-based climatologies and N_2O, NO_y, and CH_4 from three-dimensional model output. The Cl_y and Br_y families are prescribed using trace gas correlations. Surface albedo, which impacts the photodissociation rates, was set to 0.2.

OSIRIS uses a separate run of the photochemical model for each scan, initialised with OSIRIS-measured O_3 abundances and ECMWF temperatures. However, this process is computationally expensive. Thus, for most instruments, the scaling is done profile-by-profile with the pre-calculated lookup tables mentioned above.

The box model can likewise be used to supply information about an unmeasured species provided it is closely coupled to one that is measured. For example, the OSIRIS NO_y climatology was obtained from the box model using OSIRIS NO_2 and SMR HNO_3 measurements [*Brohede et al.*, 2008].

The box model was evaluated using measurements from the JPL Mk-IV FTIR interferometer [*Toon*, 1991] from 10 balloon flights between 1997 and 2005. A comparison of the partitioning of stratospheric NO_y is presented in *Brohede et al.* [2008] in which good overall agreement is found except for instances near the polar day-night boundary where air mass history becomes a dominant factor. Such studies indicate that when constrained by measurements of temperature, ozone and long-lived species, the box model is able to accurately simulate the radical species. This point, combined with the fact that the diurnal scaling approach has been used successfully in numerous validation studies of diurnally-varying species [*e.g.*, *Kerzenmacher et al.*, 2008], suggests that on average the error in the scaling factors is small. For any given profile, there may be significant errors if the assumed inputs to the model also have significant errors. However, this represents a random source of error, which is effectively minimised when averaging over a large number of profiles, as it is done in the compilation of the SPARC Data Initiative climatologies. While a rigorous error assessment has not been performed, the systematic error of these scaling factors is estimated to be less than 20% based on the above discussion.

For the scaled HIRDLS climatologies, the Specified Dynamics Whole Atmosphere Community Climate Model (SD-WACCM) is used to calculate the local time scaling factors to 10am and 10pm as a function of altitude, latitude, day of year, and LST. SD-WACCM is a global chemistry-climate model based on the Community Atmospheric Model (CAM) [*Collins et al.*, 2004] with temperature and

wind specified by the Goddard Earth Observing System (GEOS-5) reanalyses. The gravity wave drag and vertical diffusion parameterisations are described in *Garcia et al.* [2007] and the neutral chemistry modules in *Kinnison et al.* [2007].

3.1.3 Instrument-specific information

In the following, information relevant for the construction of the SPARC Data Initiative climatologies is described for each instrument. **Table 3.1** summarises specifications for all instruments.

3.1.3.1 LIMS climatologies

The LIMS Level 3 V6 combined node (ascending and descending) daily zonal mean Fourier coefficients for O_3, H_2O, HNO_3 and ascending and descending node daily zonal mean Fourier coefficients for NO_2 were used to obtain the monthly zonal mean data. Note that the ascending and descending measurements were taken at approximately 1pm and 11pm local time, respectively, for the low and mid-latitudes. The LIMS Level 3 product was the more appropriate data to use for the SPARC climatology because it has no missing data, while the LIMS Level 2 product is missing data for certain orbits or even complete days. LIMS species are given in VMR, and the profiles are first interpolated to the latitudes and then to the pressure levels used within the SPARC Data Initiative. These data were then averaged per month. The LIMS V6 data retrievals near tropopause levels may contain residual effects from cloud radiances, especially at low latitudes. The LST_MEAN, LST_MIN, LST_MAX, AVE_DOM, and AVE_LAT values provided by all other instrument climatologies are missing in the data files. Level 3 data and documentation (Level-3 README) reside at the GES DISC archive that is located at **http://disc.sci.gsfc.nasa.gov/acdisc/documentation/LIMS_dataset.gd.shtml**.

3.1.3.2 SAGE I/II/III climatologies

The SAGE climatologies are based on retrieved Level 2 products from SAGE I V5.9), SAGE II (V6.2) and SAGE III (V4.0). It is known that there are altitude errors in the original SAGE I (V5.9) data due to less reliable ephemeris information. An empirical altitude correction based on *Wang et al.* [1996] has therefore been applied to these data before their use in this study. All natively retrieved species from SAGE instruments are given in number density in altitude co-ordinates. In order to generate the SPARC Data Initiative climatologies, all number density profiles were first converted to VMR using NCEP temperature and pressure profiles, which are reported along with each individual number density profile in the SAGE Level 2 data files. A linear interpolation in $\log_{10}(p)$ was then used to derive VMRs on the SPARC Data Initiative pressure levels. Additional data screenings, as described in the following, were also applied before generating the final climatologies.

Only a few studies describing how to screen SAGE I data for anomalous values exist. The main uncertainty in retrieved O_3, H_2O and NO_2 is the interference of aerosol and clouds especially in the lower stratosphere below ~15 to 20 km. For SAGE I data, all O_3 measurements with corresponding aerosol extinctions at 1.0 μm ≥ 1.0×10^{-3} km^{-1}, are flagged (L. W. Thomason, personal communication). For SAGE II and SAGE III O_3 and NO_2 measurements, screenings follow the approach by *Wang et al.* [2002], which removes anomalously low values and those affected by "short events" or aerosols/clouds. Due to an instrument problem, the SAGE II NO_2 data from satellite sunrise measurements are not included in this study. The SAGE-retrieved H_2O is more sensitive to interferences from aerosol compared to O_3 and NO_2. More stringent criteria based on *Thomason et al.* [2004] and *Taha et al.* [2004] are therefore used to screen the H_2O data.

3.1.3.3 HALOE climatologies

The HALOE V19 measurements starting in October 1991 and extending through November 2005 are used to create climatologies for O_3, HCl, HF, H_2O, CH_4, NO, NO_2, NO_x ($NO+NO_2$), and aerosol extinction. Each individual profile is first screened for clouds and heavy aerosols. The O_3, NO_2, and NO profile data are further screened for anomalous values caused by an aerosol minimum. Each individual profile is then interpolated to the SPARC Data Initiative pressure levels. These screened and interpolated data are then averaged within each SPARC Data Initiative latitude bin to produce monthly zonal means and standard deviations of the trace gases and the aerosol extinction coefficients. The diurnally varying species NO_2, NO and NO_x are separated into local am and local pm climatological fields. The NO_x climatology is produced by first combining the screened and interpolated profiles of collocated NO and NO_2 measurements, and then zonally averaging them on the SPARC Data Initiative pressure-latitude grid. The aerosol extinction profiles were only screened for clouds before further processing.

3.1.3.4 UARS-MLS climatologies

UARS-MLS climatologies are based on Level 3AT data (similar to Aura-MLS Level 2 along-track profiles), using V5 for O_3, V6 for HNO_3, and V6 for H_2O. The main reference for the latest UARS data is *Livesey et al.* [2003]. The V6 HNO_3 files were a correction to the V5 dataset, to more properly account for emission from some of the HNO_3 excited vibrational states. The V6 H_2O dataset (originally named V0104) is described in *Pumphrey* [1999]. These source datasets are available from the GES DISC, and the H_2O dataset can also be accessed *via* the British Atmospheric Data Centre (BADC), see **http://badc.nerc.ac.uk/home/**. The above references and the UARS-MLS validation papers and data quality documentation (see individual species sections of this report) provide information about the recommended data screening for each species. The screening

methods were applied to each profile prior to the averaging and interpolation processes that were used to generate the climatological time series. This generally means that only profiles with good status values (meaning "G", "T", or "t" for the "MMAF_STAT" parameter) were considered. Other screening methods are described in *Livesey et al.* [2003]; in particular, associated UARS-MLS Level 3 Parameter files contain "QUALITY" parameters that should be (and were) considered for data screening. Also, when mixing ratios are flagged negative, this indicates that the *a priori* information is playing a non-negligible role in the retrieval process, so these values are not used in this report. Vertical profiles are retrieved as VMRs *versus* a fixed pressure grid (with spacing corresponding to 6 levels per decade change in pressure). The pressure ranges used here reflect the recommended levels for UARS-MLS profiles, although some additional information often exists beyond these ranges (mostly for higher altitude regions). If average monthly values are negative, they are not used for the SPARC climatological dataset, although small negative values may be within the calculated error. Note that UARS-MLS data after 14 June 1997 are considered slightly less reliable than for the earlier dates due to a change in UARS-MLS operations after that date (in order to conserve satellite power), whereby temperature information from the MLS retrievals was lost, and meteorological temperature fields were used instead. Therefore, some small discontinuities are to be expected at this date. Furthermore, the data become increasingly sparse after 1997. Nevertheless, this report includes UARS-MLS data after mid-June, 1997, as trend analysis is not the main focus of this report.

3.1.3.5 POAM II/III climatologies

The POAM climatologies were constructed using Level 2 data V6.0 (POAM II) and V4 (POAM III). POAM retrieves gas number density and aerosol extinction on a uniform altitude grid (0-60 km in 1-km increments). The conversion from density to VMR for the gases is done slightly differently for the two instruments. For POAM II, the UKMO total density profile, interpolated spatially and temporally to the POAM measurement, is used for the conversion. POAM III uses a total density profile retrieved directly from the measured Rayleigh scattering above 30 km, and tightly constrained to UKMO below this altitude. Each mixing ratio/aerosol extinction profile is interpolated from the POAM altitude grid to the SPARC Data Initiative pressure grid using the co-located UKMO pressure profile. The data are then binned by month and latitude bin by calculating the median value (VMR or aerosol extinction) at each standard pressure level. A minimum of 15 valid data points are required for each month and latitude bin. Data are only used within the recommended altitude range for each species, as described in the POAM documentation. The data are screened in the binning process according to the data quality flags provided with the POAM Level 2 data (described in detail in the POAM algorithm and error analysis papers, and in documentation provided with the POAM data archives). Any suspect data were eliminated before generating the climatologies. The quality flags screen data for a number of potential error sources. For gas species, the primary source of error is due to high aerosol loading in the presence of polar stratospheric clouds (PSCs), which can cause feedback noise in the gas retrievals. This is not an issue for O_3 but can be a significant source of error for NO_2 and H_2O. Both gas and aerosol retrievals can also be flagged due to the presence of sunspots in the POAM field of view. Again, these errors are species-dependent and more significant for NO_2, H_2O and aerosols. Finally, optically thick PSCs can cause the POAM scan to terminate at unusually high altitudes, resulting in higher than average retrieval noise in NO_2 and H_2O at the lowest 2-3 km of the scan. Since POAM measures at the terminator, the climatology of NO_2, which has a strong diurnal variation, was generated separately for local sunrise and sunset conditions.

3.1.3.6 OSIRIS climatologies

Climatologies from OSIRIS are based on the following Level 2 versions: BrO V5; O_3 V5.07; stratospheric aerosol V5.07; and NO_2 V3. The derived products (NO_x and NO_y) are based on the NO_2 V3 dataset but have no specific dataset number. Note that in the case of NO_y, SMR HNO_3 V2.0 data are also included (see *Section 2.2.7*). All quantities except aerosol are retrieved as number density on a fixed altitude grid and converted to VMR on pressure levels using temperature and pressure profiles from ECMWF operational analysis. The aerosol product is retrieved as extinction per km on a fixed altitude grid. OSIRIS can only provide daytime observations, (only profiles with solar zenith angles smaller than 92° are processed). In the Level 2 files, profiles with large pointing offsets, non-converging profiles and altitudes with clouds in the field-of-view have been filtered out. Note that due to low signal-to-noise ratios for BrO, only zonally averaged spectra (10° latitude bins) are used in the retrievals. The number of BrO profiles in each climatology bin will therefore be significantly less than for the other species and a true 5° latitude binning cannot be performed.

In the case of species retrieved using optimal estimation, *i.e.*, BrO and NO_2, only levels with a measurement response above 0.67 are included in the climatologies. Note that the measurement response cut-off is not applied to individual profiles but to the average values within each climatology bin. This is done in order to reduce a bias to the *a priori* profiles in the climatological averaging. Due to NO_2 log(VMR) retrievals, the climatology averaging for NO_2 (and the NO_2 derivative NO_x) is performed using the logarithm of the number densities. Other species are averaged in linear space.

The diurnal scaling of BrO uses lookup tables calculated from a photochemical box model initialised with climatological inputs (see *Section 3.1.2*). NO_2 scaling factors are obtained in a more sophisticated way from the (same) photochemical model initialised with measured OSIRIS O_3 abundances and temperature/pressure (from ECMWF) for each individual profile. Because of this scan-based ap-

proach, NO_2 (and NO_x) data can be scaled to any local time without large uncertainties. For BrO, however, only am data is used to scale to am local times and pm data to pm local times. The NO_x diurnal scaling factors are calculated simultaneous to the NO_2/NO ratios, used to calculate NO_x from NO_2.

3.1.3.7 SMR climatologies

SMR climatologies are based on Level 2 V2.1. The sole exception is HNO_3, which is based on Level 2 V2.0. In general, only 'good' quality profiles (Level 2 Quality flag = 0) have been used. Vertical profiles were retrieved as VMR or as $\log_{10}$(VMR) for CO, NO, and H_2O from the 544.6 GHz band on an altitude grid given by the refraction-corrected tangent altitudes. Conversion to pressure was done using ECMWF profiles. Retrieved VMRs with a measurement response smaller than 0.75 were rejected (0.8 for N_2O). Unphysical outliers were also filtered. The pressure range for some species was restricted: N_2O: $p \geq 170$ hPa; HNO_3: $p \leq 1$ hPa; H_2O (544.6 GHz band): 150 hPa $\geq p \geq$ 25 hPa. The minimum number of data values required per latitude bin and pressure level was set to a threshold of five; for H_2O (both from 488.9GHz and 544.6 GHz band) and NO at least ten values were demanded. For H_2O in the 544.6 GHz band, the median value was calculated instead of the mean in order to reduce the effect of unphysical outliers present in this dataset. SMR provides several Level 2 ozone data products. Ozone climatologies evaluated in this report are derived from the main stratospheric mode observations at 501.8 GHz. Climatologies have also been compiled for a second ozone product (measured in a band centred at 488.9 GHz) which has very similar characteristics compared to the 501.8 GHz SMR ozone product and is not shown in the following evaluations.

3.1.3.8 GOMOS climatologies

The GOMOS data used for the SPARC Data Initiative were produced by the ESA operational processor V5. GOMOS constituent data are number densities given at geographical altitudes. Data files also include ECMWF pressure and temperature data up to 1 hPa at GOMOS measurement locations. These data are used for ray tracing and estimating refractive effects. Above 1 hPa, the MSIS90 climatology is used in place of ECMWF. For the construction of the SPARC climatologies, VMRs and the altitude-to-pressure grid conversion are derived using these external data.

Here, we use GOMOS dark limb measurements only, requiring solar zenith angles greater than 107°. The solar zenith angle limit and the ability of GOMOS to follow and measure stars outside the orbital plane of Enivsat leads to a variation in the LST of the measurements. This is important for measurements of diurnally varying constituents NO_2, NO_3 and O_3 in the mesosphere/lower thermosphere. Envisat equator-crossing times were 10am and 10pm local time. GOMOS tangent-point local times covered about 1.5 h near the equator and 3 h at mid-latitudes.

GOMOS occultations that used stars with magnitudes weaker than 1.9 and temperatures less than 7000 K often failed to capture the whole ozone profile from 15-100 km beginning in 2003 [*Kyrölä et al.*, 2006; 2010]. After 2003, GOMOS signal-to-noise ratios decreased due to aging of the instrument. In order to guarantee ozone data quality and consistency over the whole time period we have applied the following specific filters on ozone profiles:

i. Estimated errors must be smaller than 50%;
ii. VMRs must be positive in the 25-45 km range;
iii. VMRs must be smaller than 15 ppm in the 20-45 km range;
iv. Occultations with cool stars (cooler than 6000 K) are rejected below 45 km; the same restriction applies to star numbers 170 and 178.

For NO_2, NO_3, and aerosols all stars were used regardless of their magnitude and temperature. In all datasets, we rejected occultations with the obliquity angle (the angle between the occultation plane and the orbital plane of Envisat) larger than 80°. To determine the monthly zonal mean climatologies, we have used the median as a statistical average since it is more robust against outliers than the mean. The uncertainty of the median value is estimated according to Equation 1 in *Kyrölä et al.* [2010].

3.1.3.9 MIPAS climatologies

MIPAS trace gas profiles included in the SPARC Data Initiative climatologies were retrieved on a fixed (*i.e.*, tangent altitude independent) altitude grid. Conversion to the pressure grid relies on hydrostatics and MIPAS temperature profiles. Averaging is always performed linearly in VMR, even for species retrieved in $\log_{10}$(VMR) (*cf. Funke and von Clarmann* [2011] for discussion of this specific issue). For the climatologies the unweighted mean of all measurements within a month and latitude bin is used. Note that weighting the mean by the inverse squared retrieval error would bias the mean towards warmer parts of the atmosphere. The sampling pattern, particularly from 2002-2004, is such that the measurements are not representative of the full latitude range within the latitude bins. The average values within each bin are interpolated to the centre latitude of the bin, as are the standard deviations and the number of measurements (see *von Clarmann et al.* [2012] for further details). Measurements affected by clouds were discarded from the analysis, and results where the diagonal element of the averaging kernel was below a given threshold were excluded, as well as results from non-converged retrievals. Level 2 data versions distinguish between the full spectral resolution measurements (2002-2004) and reduced resolution measurements (after 2004). Species dependent version numbers are listed in **Table 3.2**. 'FR' stands for full spectral resolution, the measurement mode MIPAS operated in from 2002 to 2004, while 'RR' stands for reduced spectral resolution as applied since 2005. Data version specifiers

are composed of a prefix indicating the version of the ESA Level 1 calibrated spectra used, a gas specifier, and a suffix indicating the version of the retrieval setup. In this report the climatologies for 2002-2004, when MIPAS operated in full spectral resolution, are referred to as MIPAS(1) while climatologies for 2005-2010, when MIPAS operated in reduced spectral resolution, are referred to as MIPAS(2). Note that the version numbers in the climatology file names and in the tables in *Chapter 4* are simplified, and only consist of the retrieval version.

3.1.3.10 SCIAMACHY climatologies

Each data product in the scientific retrieval dataset has its own version number, which is not related to the version number of the other species. The SCIAMACHY climatology is compiled using the following versions of the Level-2 products: V2.5 for O_3, V3.1 for NO_2, V3.2 for BrO, V3.1 for H_2O, and V1.0 for aerosol extinction coefficients.

Trace gas profiles and aerosol extinction coefficients are retrieved on an equidistant altitude grid. The retrieval is done in number density for all gases with except water vapour, which is retrived in logarithm of the number density. The results are then converted to VMR and interpolated to the SPARC Data Initiative pressure grid using pressure and temperature information from the ECMWF operational analysis model with a spatial resolution of 1.5° x 1.5° and a temporal resolution of 6 h. The mean value of VMR for each species in each month and latitude bin is calculated through linear averaging. Aerosol extinction coefficients, retrieved in km^{-1} are interpolated to the pressure grid and then averaged.

Because of the signal to noise ratio and radiative transfer modelling issues, only limb measurements at solar zenith angles smaller than 89° (or 85° for water vapour) are processed. Generally, these measurements are made on the dayside of the orbit (descending node, 10am equator crossing time). At high latitudes during the summer, there are also some observations on the night-side of the orbit (ascending node, 10pm equator crossing time) made at solar zenith angles smaller than 89°. However, results from these measurements are not included in the current climatology because of their substantially different local times. Furthermore, all data obtained when Envisat crosses the South Atlantic anomaly (see also *Section 2.2.10*) are excluded from the climatology. The rejected area is located between 20°S to 70°S and 0° to 90°W. For observations with clouds in the instrument field-of-view, the retrieved absorber amounts below the cloud top altitude are skipped.

3.1.3.11 ACE-FTS climatologies

The ACE-FTS climatology uses the Level 2 V2.2 dataset (including updates for O_3 and N_2O_5). The ACE-FTS VMR profiles are provided on an altitude grid with the pressures retrieved from the spectral measurements (as described in *Section 2.2.11*). The retrieved pressure information is used for the vertical co-ordinate of this climatology. The VMR measurements for each individual profile are vertically binned using the midpoints between the pressure levels (in log-pressure), which define the bins. Since no screening flags are provided with the ACE-FTS data, we use the following filtering methods: data are excluded if the fitting uncertainty value is 100% of its corresponding VMR value and where a given uncertainty value is 0.01% of its corresponding VMR value. This is the technique used for other ACE studies [*e.g.*, *Dupuy et al.*, 2009]. Binned data are subject to various criteria including statistical analysis (for further details, see *Jones et al.*, 2011; 2012). Observations that are larger than three median absolute deviations (MADs) from the median value in each grid cell are disregarded as they are deemed not a true representation (to a high probability, 95%) of the typical state of the atmosphere at a given time and place. Quality-controlled climatological fields are then created for each of the 17 species by considering the measurement uncertainties associated with each binned measurement. Each of the measurements in a bin is weighted by the inverse of the fitting uncertainty to calculate the mean. Furthermore, quality-controlled NO_x (combination of NO and NO_2) and NO_y (combination of NO, NO_2, HNO_3, $ClONO_2$, N_2O_5, and HNO_4) climatologies have also been derived using a linear combination of the individual atmospheric gas climatologies that contribute to each family. Moreover, these nitrogen species have strong diurnal features and thus climatologies based on separated local sunrise or local sunset measurements have been compiled, in addition to the combined sunrise and sunset climatological fields, using the LST information for each occultation. It should be noted that only one measurement is needed per bin from each individual contributing species in order to produce an eventual NO_x or NO_y value for that given

Table 3.2: MIPAS-IMK/IAA Level 2 data versions of different trace gases used in this report.

	FR (2002-2004)	RR (2005-2010)
H_2O	V3o_H2O_13	V4o_H2O_220
O_3	V3o_O3_9	V4o_O3_220
CH_4	V3o_CH4_11	V4o_CH4_220
N_2O	V3o_CH4_11	V4o_N2O_220
HNO_3	V3o_HNO3_9	V4o_HNO3_220
NO_2	V3o_NO2_15	V4o_NO2_220
NO	V3o_NO_15	V4o_NO_220
N_2O_5	V3o_N2O5_10	V4o_N2O5_220
HNO_4	V3o_HNO4_12	V4o_HNO4_220
$ClONO_2$	V3o_ClONO2_12	V4o_ClONO2_220
ClO	V3o_ClO_11	V4o_ClO_220
HOCl	V3o_HOCl_4	–
CCl_3F	V3o_CFC11_10	V4o_CFC11_220
CCl_2F_2	V3o_CFC12_10	V4o_CFC12_220
CH_2O	V3o_H2CO_2	–
CO	V3o_CO_12	V4o_CO_220
SF_6	–	V4o_SF6_221

bin. Scaled initial guess profiles are included as they allow for full altitude coverage to be obtained. This technique is described in detail in *Jones et al.* [2012]. A similar approach has been employed when producing the ACE-FTS climatological database [*Jones et al.*, 2011; 2012].

3.1.3.12 ACE-MAESTRO climatologies

The ACE-MAESTRO O_3 climatologies are produced using a similar methodology to that of ACE-FTS. ACE-MAESTRO VMR profiles are provided on an altitude grid, and converted to a pressure grid by linearly interpolating the ACE-FTS pressure profiles. Individual ACE-MAESTRO measurements are then binned (as described in *Section 3.1.3.11*) according to the SPARC Data Initiative pressures. Since no data screening flags are provided, data are only used if the uncertainty value is less than 100% of its given VMR value. Similar to the ACE-FTS climatology, we also apply a three median absolute deviation filter to the ACE-MAESTRO data so that outliers are identified and removed. Finally, a quality-controlled zonal mean average value is calculated using the measurement uncertainties associated with each individual binned measurement.

3.1.3.13 HIRDLS climatologies

All HIRDLS data for the SPARC Data Initiative are monthly zonal means created from the V6 Level 2 data. To minimise the impact of missing orbits or bad data points, the L3 processor is used to create a statistically best estimate for each day. These are then averaged to give the monthly mean. The L3 processor reads in all the L2 VMRs for a given product and pressure level over the entire mission and treats the data within 2° latitude bands as time series. Following a suggestion by *Rodgers* [1976], the data are represented as time-varying zonal means plus the amplitudes and phases of 6 zonal waves. A Kalman filter is used to make sequential estimates of all 13 values, with an estimate of their errors and the RMS difference between the estimated fit from the original measurements. This is done going forward and backward in time, and the estimates combined give the optimal values. *Kohri* [1981] and *Remsberg et al.* [1990] have described the method in more detail.

For quality control, parameters in each run limit the range of the data to physically reasonable values. In addition, each L2 value has an uncertainty on input, which is checked to make sure it is similar to the RMS differences from the fit. A spike detection is used so that data points that are 6σ from the estimated fit, as estimated from the covariance of the fit, have their weights reduced. This essentially means that these points have virtually no effect on the mapping or the zonal means presented here. Based on validation studies for V6 [*Gille and Gray*, 2011], the pressure level ranges for the resulting species have been restricted as shown in **Table 3.3**. It should be noted that data outside of the useful range have been eliminated from publicly released data, including the SPARC Data Initiative climatologies.

Table 3.3: Trace gas species given with their pressure level ranges for HIRDLS.

Species	Pressure range (hPa)
Ozone	422 – 0.1
Nitric Acid	100* – 10*
CFC 11	316 – 26.1
CFC 12	316 – 10.0
Daytime NO_2	56.2 – 1.0
Night-time NO_2	56.2 – 0.75

* Best range

3.1.3.14 Aura-MLS climatologies

Aura-MLS climatologies are based on Level 2 V3.3. The sole exception is O_3, which is based on Level 2 V2.2. This is mainly because of the more oscillatory (and poorer) UTLS tropical retrievals from the finer vertical resolution V3.3 data. The validation references and the Aura-MLS data quality documents provide information about the recommended data screening for each species (see individual species sections of this report). These screening methods have been applied for each profile prior to the averaging and interpolation processes that were used to generate the climatological time series used here. This generally means that only profiles with good "Status" and mixing ratios based on acceptable "Quality" and "Convergence" parameter values were included. An attempt to minimise cloud and outlier effects is also included per the MLS-recommended cloud screening methods, as well as other MLS data screening recommendations for each species (*e.g.*, removal of outliers). Also, when mixing ratio precision values are flagged negative, this indicates that the *a priori* information is playing a non-negligible role, and these values are typically not used for producing the averages. In general, only a small percentage of values is excluded *via* these screening methods, although this percentage can sometimes be larger than 20% for the tropical UTLS region (this applies to O_3, CO, and HNO_3). Vertical profiles are retrieved as VMRs *versus* a fixed pressure grid (typically with spacing corresponding to 6 levels per decade change in pressure, and double for H_2O). Also, H_2O is retrieved as $\log_{10}$(VMR). However, the Aura-MLS H_2O averages are performed in the same way as the other Aura-MLS averages, using mixing ratios, so as to compare most directly with the other climatologies using this averaging method. The pressure ranges used here reflect the recommended levels for Aura-MLS profiles although some additional information often clearly exists beyond these ranges (in particular, for higher altitude regions). Retrieved negative values are sometimes obtained due to the instrument measuring close to its detection limit. Where these measurements have resulted in negative monthly averaged values in the climatologies, the results have been flagged as bad, although it may be that some of the small negative values are within the error bars, and therefore not unreasonable.

3.1.3.15 TES climatologies

TES climatologies are based on Level 2 V4 data. Vertical profiles are retrieved as $\log_{10}$(VMR) on a 67-level pressure grid, and are interpolated in $\log_{10}(p)$ to the SPARC Data Initiative pressure grid. Only good quality retrievals have been used, and there is an additional screening to eliminate "C-curve" O_3 profiles. These profiles, which make up approximately 1-2% of TES V4 O_3 data, result from "jack-knifing" of the retrieval and convergence to an unphysical state in which the O_3 profile takes on a "C" shape under particular thermal conditions.

As stated in *Section 2.2.15*, TES measures in both Global Survey and Special Observations modes; only Global Survey data are used here. TES data are normally averaged using $\log_{10}$(VMR), but for proper comparison to the other SPARC Data Initiative climatologies, here we use linear averaging. Simple unweighted means of the available data are calculated for each month and latitude bin. A minimum of two observations per bin is required, but in practice the minimum number of profiles is 28 and in most cases the number is >1000. While the data are provided for the full range of pressures (300 to 0.1 hPa), the sensitivity of the TES O_3 retrievals drops off dramatically above 10 hPa. Data above this level should be treated with caution.

TES is a thermal instrument that measures radiances both day and night. Each global survey has measurements at two local solar times (equator crossing times of 1:43 and 13:43). The LST_MEAN value is therefore not provided because it does not reflect an average value for the measurements within the bin. Rather, the LST_MAX and LST_MIN variables represent the mean of the day and night LSTs, respectively, within each latitude bin. The variability around these values is small, ranging from ±55 minutes near the poles to ±15 minutes near the equator.

3.1.3.16 SMILES climatologies

SMILES climatologies are based on the Level 2 research (L2r) product V2.0.1. There are two O_3 products, Band-A O_3 and Band-B O_3 for the same O_3 transition at 625.37 GHz with a different receiver and spectrometer to check the spectrum calibration accuracy. Level 2 data were filtered according to the quality criteria specified for this release. Measurements that were deemed of good quality based on an acceptable "measurement response" and "convergence" parameter values were included. Only clear sky data was provided for the L2r V2.0.1 data product. In this climatology, retrieved VMRs with a measurement response smaller than 0.75 have been rejected and the minimum number of data values required per latitude bin and pressure level was set to five. The pressure range has been limited to ≥ 10 hPa for BrO, ≤ 1 hPa for HNO_3, and ≥ 25 hPa for HOCl. Water vapour was retrieved from the continuum but is not included as a product. The quality and sensitivity of each individual species used in this report, the recommended data screening for each species, and validation references are provided in the SMILES Mission Plan, Version 2.1, (**http://smiles.nict.go.jp/Mission_Plan/**), and in the SMILES L2r products guide, (**http://smiles.nict.go.jp/pub/data/products.html**). L2r V2.1.5 products have been used in this report where data were made available in time for processing. In the V2.1.5 data a known issue of non-linearity in the spectrum has been improved.

The instrument is on-board the International Space Station in a 51.6° inclined orbit and observations drift slowly with respect to LST, so that all LSTs are sampled for each latitude over a 2-month period. Climatologies of short-lived species are separated into daytime (solar zenith angle ≤87°) and night-time measurements (≥93°).

3.2 Climatology uncertainties

Measurements are imperfect estimates of the truth. Measurement error, defined as the difference between any measurement and the truth, can be decomposed into two parts; a random component that has, over large sample, a mean of zero, and a bias that has a non-zero mean. For satellite-based measurements of trace gas species, the magnitude of the error depends on many factors, including the measurement technique, the chemical species measured, and the time and location of the measurement.

Calculated climatological fields can be affected by the presence of errors in the measurements. Random errors, by definition, have little impact on climatological means. Measurement bias on the other hand will produce a difference between a measurement climatology and the true climatology. Measurement biases can come about due to a number of factors, including (but not limited to) retrieval errors (*e.g.*, the diurnal effect), errors in the input parameters of the retrieval that are assumed to be known but may have their own uncertainties (*e.g.*, spectroscopic data), and so-called smoothing errors related to the spatial resolution of the retrievals. Absolute bias determination for any one satellite instrument is quite difficult since the truth is rarely known, but inter-instrument biases can be deduced through validation exercises.

For limb sounders, one important aspect of the absolute measurement error is the degree to which vertical resolution can smooth the profile. This smoothing error differs between instruments, retrieval schemes and species [*cf.*, *Rodgers*, 2000 for details]. Therefore, the climatologies will have some instrument-specific characteristics that can be understood only by consideration of the averaging kernels (for example, instruments with better vertical resolution will see a drier hygropause). It should be noted that an instrument with poorer vertical resolution is not *per se* bad; its results are still useful, but the data user must take the instrument and retrieval characteristics properly into account when interpreting the data.

Wherever possible, differences in climatologies within the SPARC Data Initiative will be explained based on the results of prior validation work. However, in addition to the error in the raw measurements, the monthly mean climatologies contain errors introduced by the climatology production. This section will focus on highlighting important sources of climatology error, including added uncertainty due to instrument sampling (*Section 3.2.1*), and due to differences in averaging techniques (S*ection 3.2.2*). *Section 3.2.3* concludes with a description of the climatology error bars used in this report.

3.2.1 Uncertainties due to sampling

The monthly zonal mean SPARC Data Initiative climatologies are produced by binning measurements from each instrument in month and latitude bins. Each instrument obtains a finite sample of profile measurements in each bin, based on the space-time pattern of measurement locations for that instrument. The space-time sampling pattern may be dense and uniform, or sparse and highly non-uniform, or somewhere in between. The degree of non-uniformity of the sampling pattern, together with the space-time gradients in the measured field may lead to a difference between the sample mean and the true mean.

This sub-section briefly describes an exercise that aims to produce pseudo-quantitative estimates of sampling bias for a number of instruments participating in the SPARC Data Initiative. These sampling biases can be seen as example cases, and can be used to highlight regions and seasons of significant sampling bias, and its approximate magnitude. This information should help in the comparisons of instrument climatologies in other chapters.

Sampling patterns have been collected from each instrument team, and defined by day, latitude and longitude of measurement locations. For many instruments, a typical year of actual sampling locations has been used in the analysis, rather than, for instance, a time series of all possible measurements, which may differ because of *e.g.*, data download limitations. The time periods used to define each instruments' sampling pattern are the same as those used to produce the sampling density figures in the instrument descriptions of *Chapter 2.2*.

We have used output from the WACCM3, a fully coupled chemistry-climate model, spanning the range of altitude from the Earth's surface to the thermosphere [*Garcia et al.*, 2007]. The particular version of the model used here (3.4.58) is the same as that used for the last Chemistry-Climate Model Validation Activity [*SPARC CCMVal*, 2010], except that the number of vertical levels has increased to 102, and the number of chemical species included has increased to 125. The horizontal resolution is 1.9° by 2.5° (latitude by longitude). Here, we use model output with daily resolution at 0 UTC from one year of a transient simulation under current climate conditions.

Instrument sampling patterns for each month of the year are used to subsample the model data. For each sample, model fields from the corresponding Julian day are linearly interpolated in space to the latitude and longitude of the sample location. (Interpolation is not performed to the time-of-day of the measurements, since the effect of diurnal variability on SPARC Data Initiative climatologies is explicitly dealt with for short-lived species, for which the diurnal cycle is important.) Once model data have been interpolated to each sample location, the subsampled fields are binned according to the SPARC Data Initiative latitude grid, and the mean is calculated. The "true" model climatology, or population mean, is produced by first calculating the mean of all model fields on each latitude circle of the model's latitude grid, then linearly interpolating these mean values to the midpoint of each SPARC Data Initiative latitude bin. The difference between the instrument-sampling-pattern-based field mean and the full-model-resolution field mean gives the sampling bias. For each month and for each instrument, this bias is calculated for every latitude bin in which an instrument has measurements, and at all pressure levels of the model fields.

As an example result, the monthly zonal mean sampling bias for O_3 in March is shown for each instrument as a function of latitude and height in **Figure 3.1**. Monthly zonal mean climatology sampling bias estimates from the sampling exercise for O_3 for all months and for all instruments are available in *Appendix A*. The results of the sampling bias exercise can be very briefly summarised by categorizing instruments according to the severity of their sampling bias. We see:

i. A weak sampling bias (always <5%) for dense samplers Aura-MLS, HIRDLS, MIPAS, SMR and TES.
ii. Strong sampling bias (>5%) for occultation instruments ACE-FTS, HALOE, POAM II, POAM III, SAGE II, SAGE III, and GOMOS which is strongest at, but not limited to high latitudes.
iii. Occasionally (in time or space) strong (> 5%) sampling bias for OSIRIS, SCIAMACHY, SMILES and UARS-MLS.

The largest sampling biases can be understood to be a product of non-uniform sampling throughout the days of a month, as can be seen when one examines variations in ozone over a month and the correlation of these variations with instrument sampling patterns. **Figure 3.2** shows the time evolution of zonal mean O_3 in March from the model, at pressure levels 100, 10 and 1 hPa, as anomalies from the monthly zonal mean. Superimposed on the chemical fields are latitude *versus* time sampling patterns of ACE-FTS and MIPAS, as examples of the two extremes in types of sampling patterns.

The MIPAS sampling pattern contains measurements in all latitude bins for all days, *i.e.*, there is no variation in the sampling locations with time, and as a result the sampling bias is small. ACE-FTS, on the other hand, as a solar occultation instrument, samples each latitude band over only a few days of the month. For example, in the month of March, SH

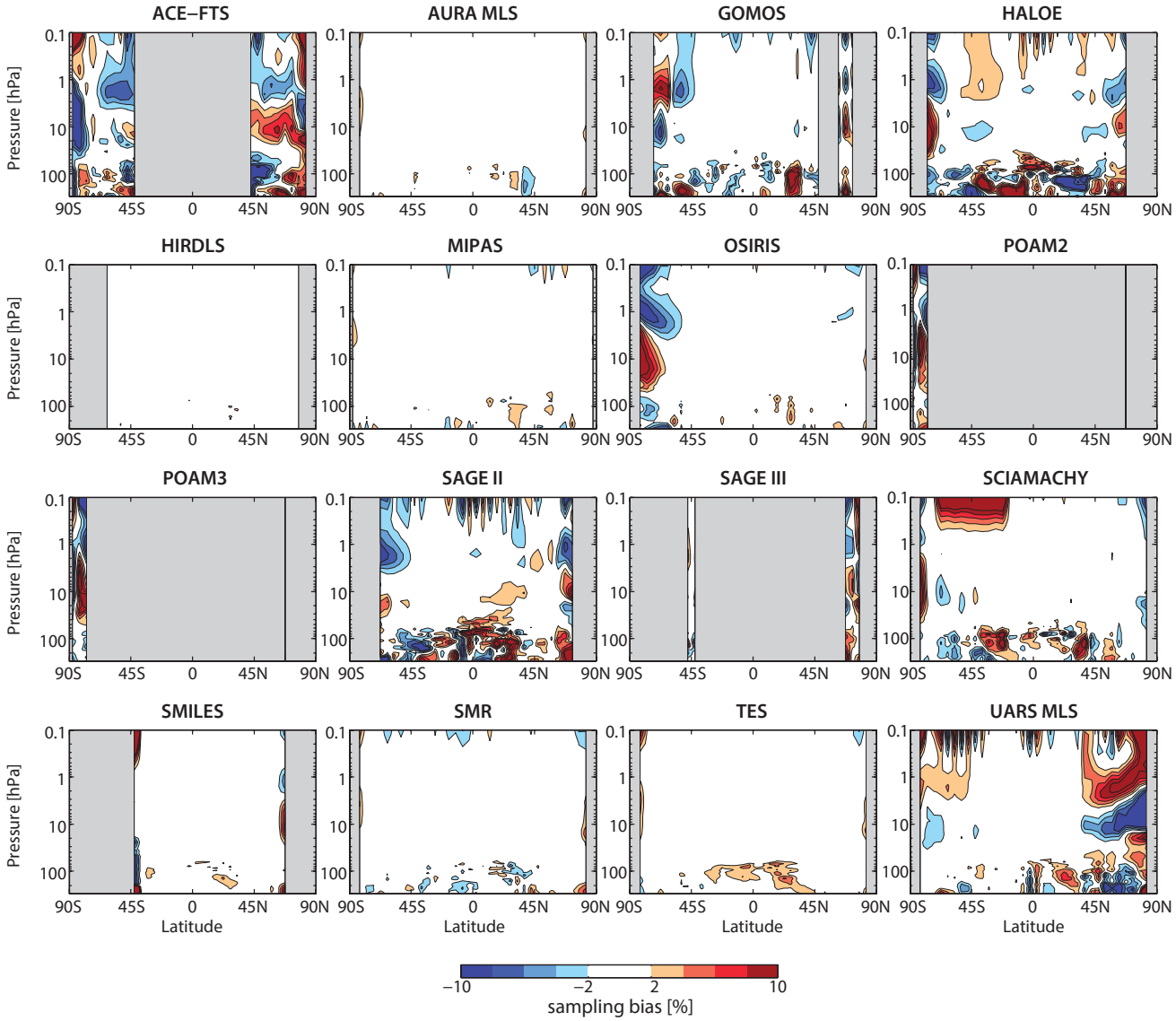


Figure 3.1: Latitude-height sections of calculated sampling error for O_3 in March, based on sampling patterns of instruments as labelled in each panel. *Grey regions denote regions of no measurements.*

mid-latitudes (45°S) are sampled only at the very beginning of the month, while SH high latitudes (80°S) are sampled only at the very end of the month. At 1 hPa, ozone mixing ratios are increasing through the month over this latitude range, therefore, the ACE-FTS sampling pattern leads to negative sampling bias around 45°S, and slightly positive sampling bias at the highest latitudes. The seasonal cycle of ozone is comparatively reversed at 10 hPa, leading to slightly positive bias in the SH mid-latitudes, and negative bias in the SH high latitudes. In this way, it can be seen that the sampling biases of ACE-FTS can be well explained by the instrument's sampling pattern and the intra-monthly variations in ozone, which depend strongly on height and latitude. At 100 hPa, intra-monthly O_3 variations are relatively noisy, and as a result the sampling bias is dependent on the sampling of the intra-monthly variability. We therefore can expect that in regions where the sampling bias is due to the non-uniform sampling of the slow seasonal variability through a month, that the sign and approximate magnitude of the sampling bias calculated through our model exercise to be a reasonably accurate estimate of the real sampling bias for each instrument. However, in regions where variability is dominated by short-term (intra-monthly) variations, limited sampling of such a chemical field will lead to a random sampling error. In this case the sign and magnitude of the sample error calculated through our model exercise serves as an example, and should be used only to identify regions where sampling error may be important.

The sampling biases for solar occultation instruments are similar to that of ACE-FTS, and are primarily a result of the non-uniform day-of-month sampling. The sampling biases of OSIRIS and UARS-MLS come from a similar source: while these instruments have dense sampling patterns, the latitudinal coverage of their measurements changes periodically, and as a result, certain latitudes (or in fact a whole hemisphere) are often sampled for less than the full month. Such is the case for OSIRIS in the SH and UARS-MLS in the NH in the sampling error exercise results shown in **Figure 3.1**.

In general, the sampling bias for all instruments is weak in the tropics where variability is weak on both intra-seasonal

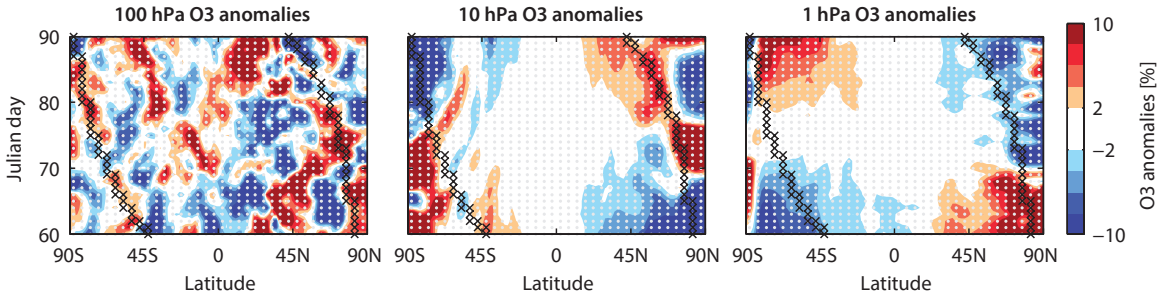


Figure 3.2: March O_3 anomalies from the March monthly mean shown as a function of latitude and Julian day for the 100, 10 and 1 hPa surfaces. *The locations of the latitude bins that contain measurements according to the MIPAS (grey dots) and ACE-FTS (black crosses) sampling patterns.*

and seasonal time scales. In the extra-tropics and polar regions, where variability is more pronounced, the sampling bias becomes much larger. Between 60°-65° in both hemispheres, sampling bias has a double-peak structure, with maximum values around 20 and 2 hPa. It is interesting to note that the solar occultation instruments ACE-FTS, HALOE and SAGE II, as well as OSIRIS, show similar sampling biases for March at around 1 hPa between 45°-65° in both hemispheres due to similarities in the seasonal progression of their sampling patterns. This is one example where close agreement between data climatologies from different instruments may not imply good agreement with the true climatological mean.

In order to assess how the sampling bias can affect annual mean climatologies, we calculate the annual mean sampling bias for each instrument by averaging the sampling biases for the 12 calendar months. These annual mean sampling biases are shown for each instrument in **Figure 3.3**.

The instruments with the highest sampling density (Aura-MLS, HIRDLS, MIPAS, and TES) show small annual mean sampling biases of only a few percent, as would be expected due to the small sampling biases in their monthly means. Due to the seasonal variability of the OSIRIS and UARS-MLS sampling patterns, their sampling bias somewhat cancels out in the annual average, with maximum values of a few percent. Finally, for the occultation instruments (ACE-FTS, GOMOS, HALOE, POAM II, POAM III, SAGE II, and SAGE III), the annual mean sampling biases are on the order of 5% at latitudes >50° in both hemispheres. The details of the sampling bias – its sign and magnitude – are generally different for the different instruments, however, some features are common to multiple instruments (*e.g.*, negative sampling bias at 1 hPa and ~60° in both hemispheres) and are related to similarities in the sampling patterns.

In summary, when constructing climatologies by averaging binned atmospheric measurements, sampling bias can arise due to non-uniform sampling in time or space. We have examined sampling biases produced by the sampling patterns of a number of instruments participating in the SPARC Data Initiative using ozone from WACCM. We find that:

- Climatologies based on measurements from instruments with high sample density generally have small sampling biases due to their highly uniform sampling of each latitude bin.
- Climatologies based on measurements from instruments whose latitudinal coverage varies with time can have strong sampling biases for certain months and locations. Sampling biases for O_3 were found in some instances to be above 10%. This is primarily due to non-uniformity in day-of-month sampling, and occurs whenever an instrument provides measurements in one month over only a portion of that month. Whenever the atmospheric variability is dominated by the seasonal cycle, this type of sampling error could in theory be reasonably well quantified or even corrected, however, when variability is dominated by intra-seasonal (short- term) variations, only the absolute magnitude of the sampling bias can be estimated from model studies. This type of sampling bias is most relevant for solar occultation instruments, but also for instruments with high sample density when the latitudinal coverage changes with time, such as OSIRIS, SMILES and UARS-MLS.
- Annual mean sampling bias can be on the order of 5% or larger for solar occultation instruments at high latitudes, and a few percent for instruments with varying latitudinal coverage such as OSIRIS, SMILES and UARS-MLS.
- In the UTLS region, intra-monthly variations and gradients in many trace gas species are large, therefore the sampling bias is more important. The sampling bias for O_3 in monthly mean climatologies is found to be often on the order of 10% (higher for H_2O; not shown), and still significant in annual mean climatologies. For precise monthly-mean or annual-mean climatologies in the UTLS, one requires a high sample density.

3.2.2 Uncertainties due to averaging technique

Averaging of data may lead to biases between climatologies in cases when different averaging procedures are used to generate the climatologies. Averages are typically defined as monthly zonal mean VMRs, but averages of $\log_{10}$(VMR) or of median values of the spatio-temporal distributions are also used. Under particular atmospheric conditions,

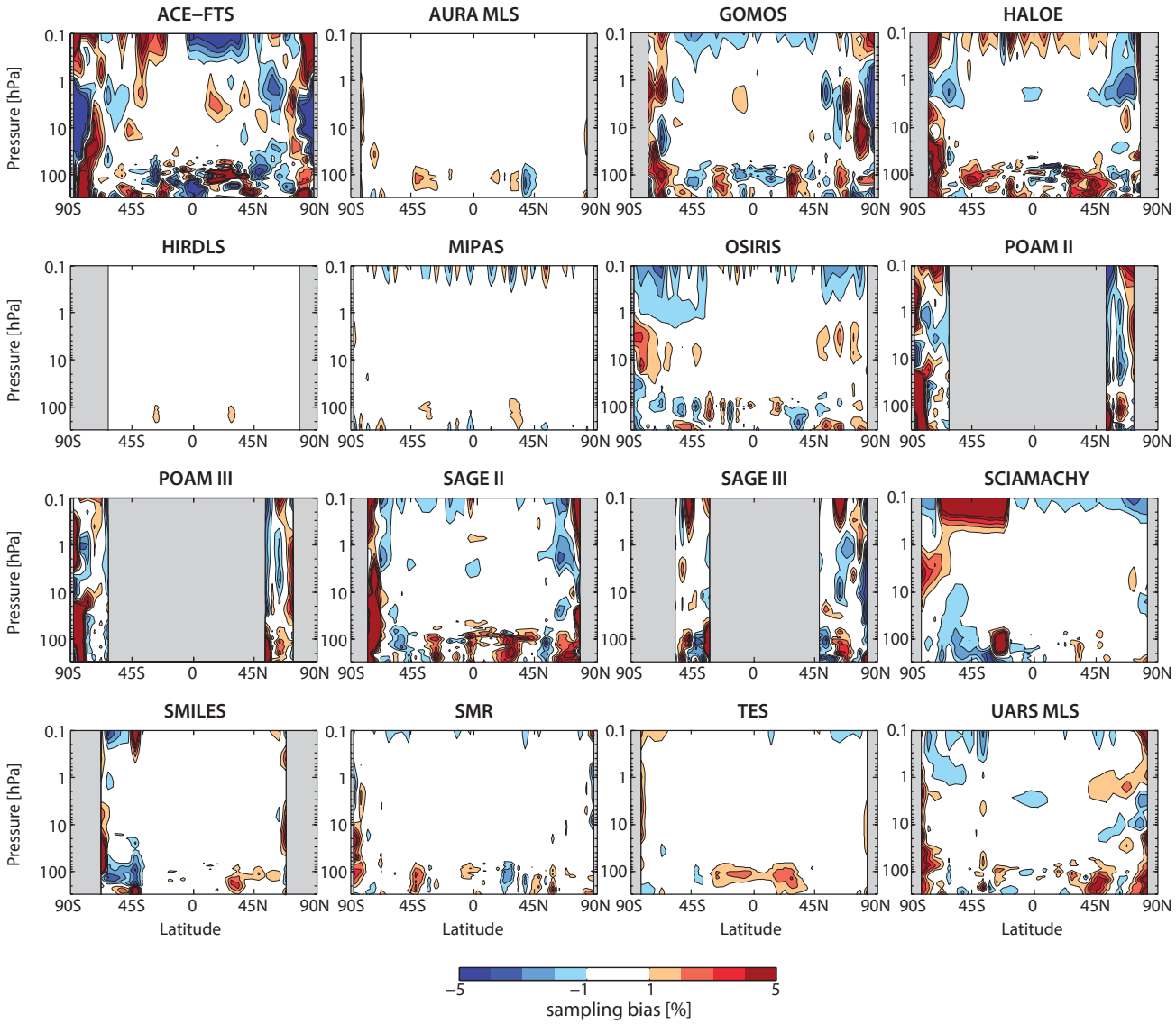


Figure 3.3: Latitude-height sections of calculated annual mean sampling error for O_3, based on sampling patterns of instruments as labelled in each panel. *Grey regions denote regions of no measurements.*

these averaging methods can lead to significantly different results for many trace gas species. As an example, we show in **Figure 3.4** monthly zonal mean distributions of H_2O, CO and O_3 and their standard deviations calculated from WACCM model simulations described in *Jackman et al.* [2008] for November 2003. The monthly zonal means are calculated from 10,000 modelled mixing ratios per species for each latitude-pressure grid point, and are compared to averages calculated in $\log_{10}$(VMR) space, as well as to their respective median values. The following conclusions can be drawn from this comparison:

- The bias between differently averaged zonal mean fields (*i.e.*, linear or logarithmic averages or median values) correlates spatially with the standard deviation of the distributions.
- Standard deviations and hence biases are most pronounced where spatial gradients are strongest, *e.g.*, in regions of transport barriers or strong vertical transport. In our example, this occurs for CO in the polar regions in the mid-stratosphere and is related to vertical transport by the meridional circulation. H_2O variability is highest in the UTLS. Additionally, averaging biases related to diurnal variations are found for O_3 in the mesosphere.
- Logarithmic averaging always yields smaller values than linear averaging.
- Median values can be higher or lower than linearly averaged zonal means.

The sign of the bias depends on the asymmetry of the distribution. This is particularly evident in the case of O_3 in the mesosphere where the O_3 distribution is bi-modal due to diurnal effects. In the summer hemisphere, where daytime population is dominant, the median yields values closer to the daytime VMR and hence is smaller than the linear average, while the opposite occurs in the winter hemisphere.

Most of the climatologies within the SPARC Data Initiative were built on the basis of linear monthly zonal means, though exceptions exist; *e.g.*, GOMOS O_3 and NO_2 climatologies and SMR H_2O from the 544.6 GHz band (SMR2) are

based on median values, while OSIRIS NO_2 and NO_x climatologies are based on $\log_{10}$(VMR). The comparison of these climatologies with those of other instruments (see *Chapter 4*) might therefore suffer from statistical averaging biases.

In the case of GOMOS O_3, however, such pronounced mesospheric biases resulting from the use of the median as seen in **Figure 3.4** are not expected since GOMOS measures only during the night-time and issues related to different diurnal populations do not play a role in the averaging technique. Remaining biases, most likely located in the UTLS region, are expected to be within 15%, which is considerably smaller than the inter-instrumental spread observed in this altitude range. GOMOS NO_2 median values are likely to be smaller than linear averages at the edge of NO_x-rich air masses descending in polar winter, as observed for CO. On the other hand, a slightly positive bias might occur in the core of these air masses. As in the case of O_3, averaging biases related to diurnal variations are not expected to occur. Regarding the SMR H_2O climatology obtained from the 544.6 GHz band (SMR2), biases related to the use of median values might be an issue. **Figure 3.4** indicates deviations on the order of ±20% in the altitude range 16-20 km (~100-60 hPa) where this data product is provided.

No important averaging biases are expected for the OSIRIS NO_2 and NO_x climatologies since they are restricted to sunlit conditions (*i.e.*, no diurnal issues) and do not cover the polar winter regions where averaging differences related to the mixing of NO_x-rich mesospheric and stratospheric air masses might occur.

Apart from these biases, which arise from the comparison of differently averaged climatologies, there exists an intrinsic source of statistical averaging errors for climatologies built from trace gas abundance data retrieved in the $\log_{10}$(VMR) space (*i.e.*, CO, NO, NO_2, and H_2O from MIPAS, SMR, OSIRIS, SCIAMACHY). A detailed discussion of this error source on basis of idealised retrieval simulations is given in *Funke and von Clarmann* [2011]. A quantitative evaluation of related errors in the context of this study is not feasible due to the complex dependence of their magnitude and sign on natural variability, measurement sensitivity, and retrieval constraints. However, efforts have been undertaken in the definition and optimisation of the instrument-specific retrieval algorithms operating in the $\log_{10}$(VMR) space in order to reduce these errors whenever possible.

3.2.3 Climatology error bars

The statistical uncertainty in a mean value, calculated from n measurements with a standard deviation σ, is commonly estimated through the standard error of the mean (SEM):

$$\mathrm{SEM}=\sigma/n^{1/2}. \qquad (3.1)$$

The SEM is an estimate of the standard deviation of all the possible mean values one would produce if one was able to re-sample the original population from which the sample is drawn. The formalism of the SEM assumes that individual samples are independent. This may not be the case within the SPARC Data Initiative, since, for example, the sampling patterns of some instruments may be dense enough that closely spaced measurements are autocorrelated. In fact, satellite data sorted into latitude bands may exhibit positive or negative autocorrelations, depending on the details of the sampling pattern and latitude grid [*Toohey and von Clarmann*, 2013]. It is therefore not possible to know whether the "classical" SEM, as calculated by Equation 3.1, is in general an over- or underestimate of the true uncertainty in the mean climatologies.

Standard deviations are also affected by the climatology production. The standard deviations are themselves a function of both the random measurement error and the natural variability sampled at the spatial and temporal resolution/pattern of the instrument. Thus, the magnitude of the natural variability present in the climatological standard deviation fields is also subject to sampling error compared to the true variability within a latitude bin. In some cases, it may be preferable or necessary to interpolate the standard deviation to latitude grid midpoints (see *von Clarmann et al.*, 2012). It should be noted that linear interpolation, as used to produce climatologies on a standard vertical grid, will decrease the variability of a field when the correlation between adjacent points is low (*i.e.*, when random measurement errors are large compared to natural variability). Due to this effect, the standard deviation of the climatologies will in some cases be less than the standard deviation calculated on an instrument's native retrieval grid. This reduction in standard deviation is artifical in that any interpolation between two data points on the original grid acts to reduce the uncertainty associated with the random measurement error, as when calculating the mean of multiple data points.

Despite its shortcomings, due to its ease of computation and its frequent use in past studies, the SEM as calculated *via* Equation 3.1 using the standard deviations provided in the climatology will be used in this report to indicate an approximate measure of uncertainty in each climatological mean. In particular, uncertainties in the mean will be graphically illustrated by 2×SEM error bars, which can be loosely interpreted as a 95% confidence interval of the mean.

It should be stressed that the statistical error in the mean is in many cases much smaller than the overall error of the climatology, which contains the systematic errors of both the measurements and the climatology construction. We have briefly explored the potential importance of two types of climatology error in this subsection, but this discussion is not exhaustive. For example, potential biases introduced through filtering of retrievals used in the climatology construction (*e.g.*, including only cloud-free measurements) are not addressed here. A complete characterisation of the systematic errors of each climatology is beyond the scope of this report and would require

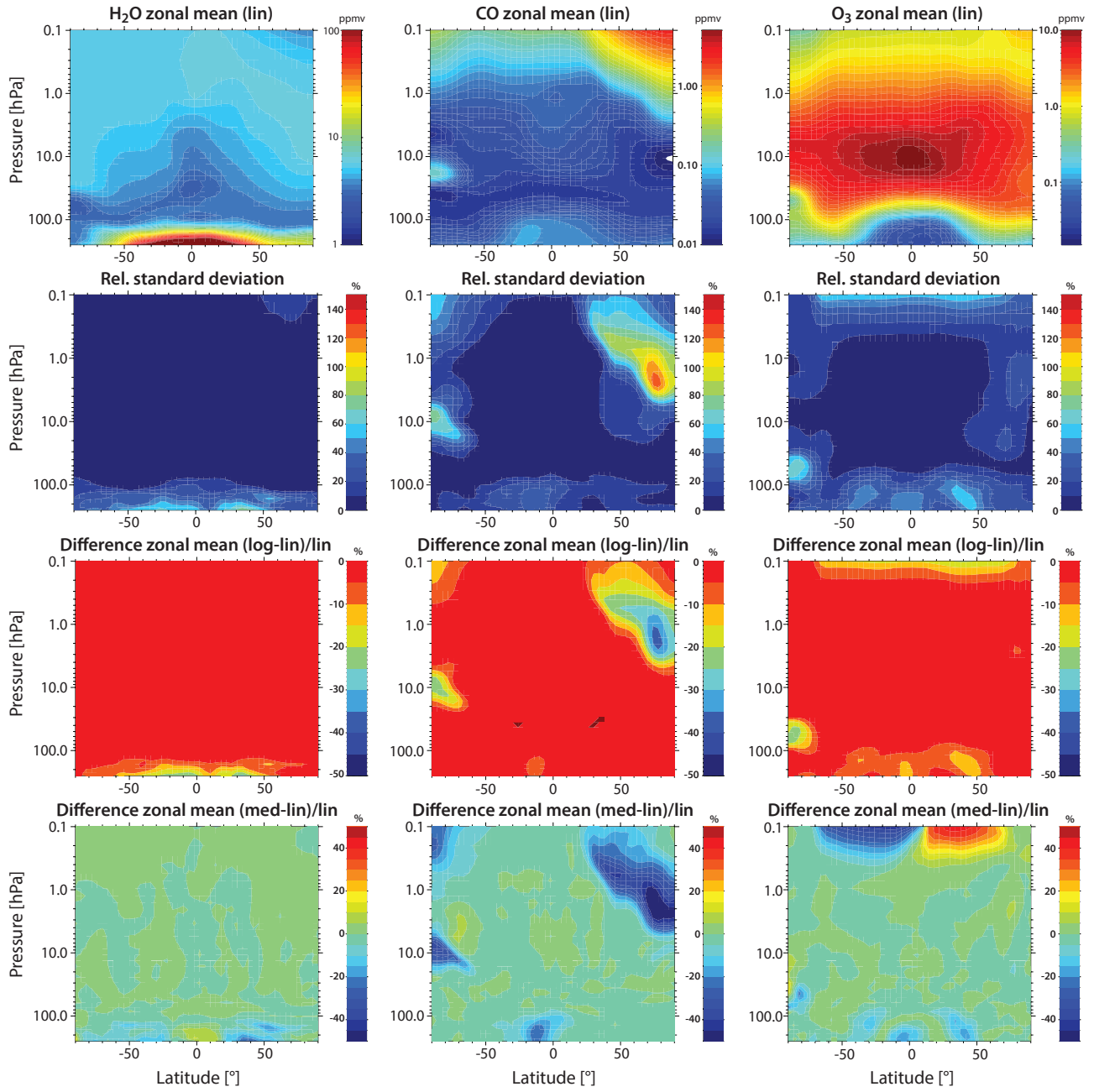


Figure 3.4: Monthly zonal mean distributions of H_2O, CO and O_3 and their standard deviations calculated from WACCM for November 2003 in the two upper rows. *Differences of linear mean values to averages calculated in $\log_{10}$(VMR) space, as well as to their respective median values are shown in the two lower rows.*

a precise knowledge of the absolute measurement uncertainties for all instruments. Since such knowledge is not available in a consistent way for all instruments, it is recommended that future efforts that focus on deriving absolute measurement uncertainties. The uncertainties would need to include a range of error sources such as uncertainty in the spectroscopic data, calibration, pointing accuracy, and others. The uncertainties would need to be derived consistently between the instruments according to a common standard so to allow for apple-to-apple comparisons. In the absence of such bottom up measurement uncertainties, we will use the inter-instrument spread of the climatologies to provide a measure of the overall uncertainty in the underlying chemical fields.

3.3 Climatology diagnostics

A set of standard diagnostics is used to investigate and test the differences between the trace gas time series obtained from each instrument. The diagnostics include annual and monthly zonal mean climatologies, vertical and meridional mean profiles, seasonal cycles, and interannual variability. In addition, trace gas-specific evaluations such as the tape recorder for H_2O and the quasi-biennial oscillation (QBO) for O_3, which test the physical consistency of a dataset, are carried out. Such diagnostics include the latitude-time or altitude-time evolution of trace gases that are sensitive to specific transport processes, such as descent within the polar vortex or the seasonal varition in the strength of the

Brewer-Dobson circulation. The evaluation methods are described in more detail in the following.

3.3.1 The multi-instrument mean (MIM)

We introduce the concept of the multi-instrument mean (MIM), which we use throughout the report as a common point of reference. The MIM is calculated by taking the mean of all available instrument climatologies within a given time period of interest. Note, that the MIM is not a data product and will not be provided with the instrument climatologies. By no means should the MIM be regarded as the best estimate of the atmospheric state, since all instruments are included in its calculation regardless of their quality and without any applied weighting applied. Where instruments offer more than one data product of a given trace gas species, only one data product is included in the MIM, so not to bias the MIM towards this instrument.

Throughout the report we calculate relative differences between the trace gas mixing ratios of an instrument ($\chi_{instrument}$) and the MIM (χ_{MIM}) using

$$100 \times (\chi_{instrument} - \chi_{MIM}) / \chi_{MIM} \qquad (3.2)$$

It should be emphasised that when interpreting relative differences with respect to the MIM, one must keep in mind that the set of instruments from which the MIM was calculated may have changed in between time periods. Also, if there is an unphysical behaviour in one instrument, the MIM and thus the differences with respect to the MIM of the other instruments will most certainly reflect this unphysical behaviour. Finally, if one instrument does not have global coverage for every month a non-physical structure may be introduced into the MIM that reflects this sampling issue. Despite its shortcomings, we have chosen to use the MIM throughout the report as a common point of reference for comparison between instruments, in order to avoid singling out any particular instrument as a benchmark.

3.3.2 Annual and monthly mean cross sections and profiles

For the annual and monthly mean cross sections, as well as the altitude and meridional profile evaluations, multi-annual means were produced in order to reduce potential sampling errors, and to limit the influence of interannual variability, *e.g.*, through the QBO. However, we also intended to compare a maximum number of available instruments for the same time period, so often a trade-off between number of instruments and length of the climatology had to be made. The monthly or annual zonal mean cross sections are analysed to investigate mean biases in the datasets. The vertical and meridional profiles help focus on particular height/latitude regions and months. This evaluation (along with other evaluations that follow) will also help to determine if biases between datasets are persistent over the entire year. The comparison of cross sections (or profiles) from individual instruments is based on the relative differences of each instrument to the MIM (see *Section 3.3.1*).

3.3.3 Seasonal cycles

For the seasonal cycles, the multi-year approach has been chosen. The seasonal cycle results include the MIM (see explanation above) together with its 1σ standard deviation, which is a measure of the range of mean values obtained from the different instruments. A combined annual and semi-annual fit has been applied to all the available monthly mean values of a single instrument, in order to yield a

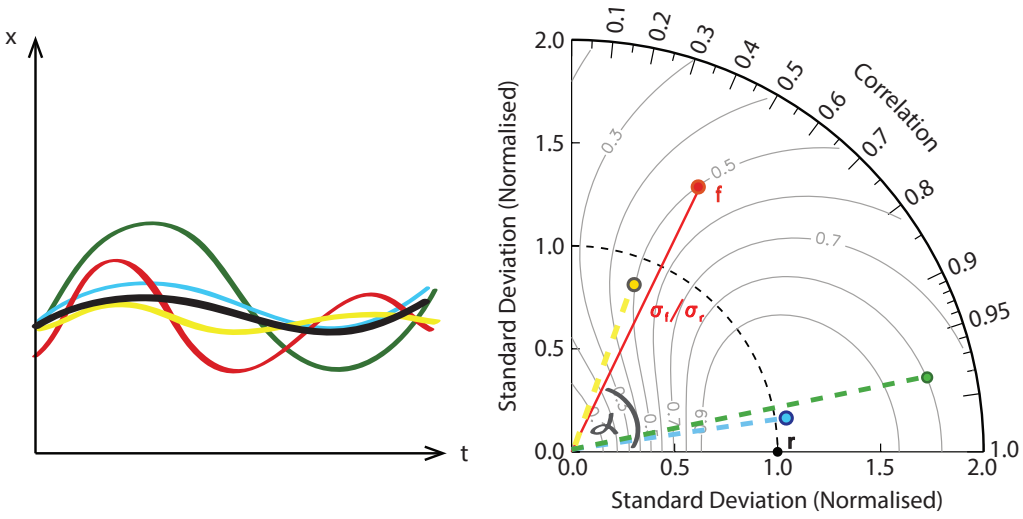


Figure 3.5: Left panel: Exemplary seasonal cycles corresponding in colour to the dots in the Taylor diagram. Right panel: Taylor diagram describing the agreement between the reference field (r) and a test field (f). The angle a represents the correlation between the fields. *The radial distance shows the amplitude in the seasonal cycle of the test field normalised by the standard deviation of the reference field (σ_f / σ_r). The grey thin lines indicate the skill score of the test field, which is an overall metric of the agreement (see text for explanation).*

seasonal cycle that is comparable even for instruments that do not measure for all months of the year. Finally, Taylor diagrams [*Taylor*, 2001] are used in order to compare the different instruments in a more quantitative way. Taylor diagrams offer a visual summary of the pattern statistics of how well a certain instrument's seasonal cycle reproduces the seasonal cycle of a reference field or a 'true' state. Three measures can be deduced from the Taylor plots as illustrated in **Figure 3.5**: the correlation on the azimuthal axis, which represents how well the phase of the true seasonal cycle is reproduced by the instrument; the normalised amplitude on the x- and y-axis; and the skill factor, indicated by the light grey lines, which summarises the overall performance of an instrument's field. The closer the instrument lies to the '1' on the x-axis, the better it agrees with the reference field. The Taylor diagram shown in **Figure 3.5** demonstrates that the blue seasonal cycle is closest to the reference field (*r*, black), with a skill score of about 0.97, green shows a similar phase, but too large an amplitude (resulting in a skill score of about 0.8), yellow shows the wrong phase but the right amplitude (skill score 0.5), and red shows the wrong phase with too large an amplitude (skill score 0.5).

Note, that the Taylor diagrams do not include information on the performance of how well the instruments reproduce the mean values of the seasonal cycles, so this measure needs to be examined in addition. Please see *Hegglin et al.* [2010] for an additional example of how to interpret Taylor diagrams.

3.3.4 Time series of latitude and altitude profiles

Time series of both the absolute values and deseasonalised anomalies are used to analyse intra-annual and interannual variability in the trace gas datasets. Examples of time series based on absolute values are the H_2O tape recorder or polar dehydration evaluations, which show the time-pressure evolution of absolute mean values over several years. In some instances, the latitude or altitude time series are averaged over several years so to yield a more robust estimate of the mean annual evolution of monthly zonal mean values.

Deseasonalised time series are shown for selected latitude bands and pressure levels or as an altitude-time evolution of the trace gas, *e.g.*, to analyse the QBO. For each month the anomalies are calculated by subtracting the multi-year mean value of the month of the respective instrument (averaged over all years taken into account for this diagnostic) from the monthly mean values.

3.3.5 Summary plots

We use two different types of summary plots in order to present an overview of the findings within each trace gas chapter: one highlighting the uncertainty in our knowledge of the atmospheric mean state; and the other highlighting specific inter-instrument differences.

For each trace gas species the first type of summary plot shows the inter-instrument spread of climatologies to give some measure of the overall uncertainty in the underlying chemical fields. Annual zonal MIM, multi-instrument minimum (MIN) and multi-instrument maximum (MAX) fields are provided, with the latter two based on the minimum and maximum over all instruments estimated separately for each grid point. The difference between MAX and MIN, as well as the standard deviation over all instruments, is presented in absolute and relative values to demonstrate the maximum spread and the variations from the MIM over all instruments. Again, the two quantities are estimated separately for each grid point.

In the second type of summary plot, average deviations of each instrument from the MIM are presented for different regions showing which datasets are consistent with each other and which not. The regions are divided into different altitude ranges (300-100 hPa; 100-30 hPa; 30-5 hPa; 5-1 hPa; 1-0.1 hPa) and into the extra-tropics (40°-80°S/N) and the tropics (20°S-20°N). The tropics show somewhat smaller variability than the extra-tropics, hence trace gas evaluations are generally less sensitive to sampling issues and give a cleaner estimate of the overall measurement error. In the extra-tropics, inter-instrument differences are expected to be larger due to larger dynamical variability and hence greater sensitivity to sampling issues. The average deviation of each instrument for a particular region is calculated as the median (MED) over all values the instrument exhibits in this region. The median is regarded to be more robust against outliers. Additionally, the median absolute deviation (MAD) is provided for each instrument and region. The MAD over the sample $x = (x_1, \ldots, x_n)$ is defined as:

$$\mathrm{MAD} = \mathrm{MED}\left(|x - \mathrm{MED}(x)| \right) \qquad (3.3)$$

and represents the interval around the median that contains 50% of the data [*Rousseeuw and Croux,* 1993]. For comparison, the range indicating the mean $\pm 1\sigma$ is also indicated.

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.

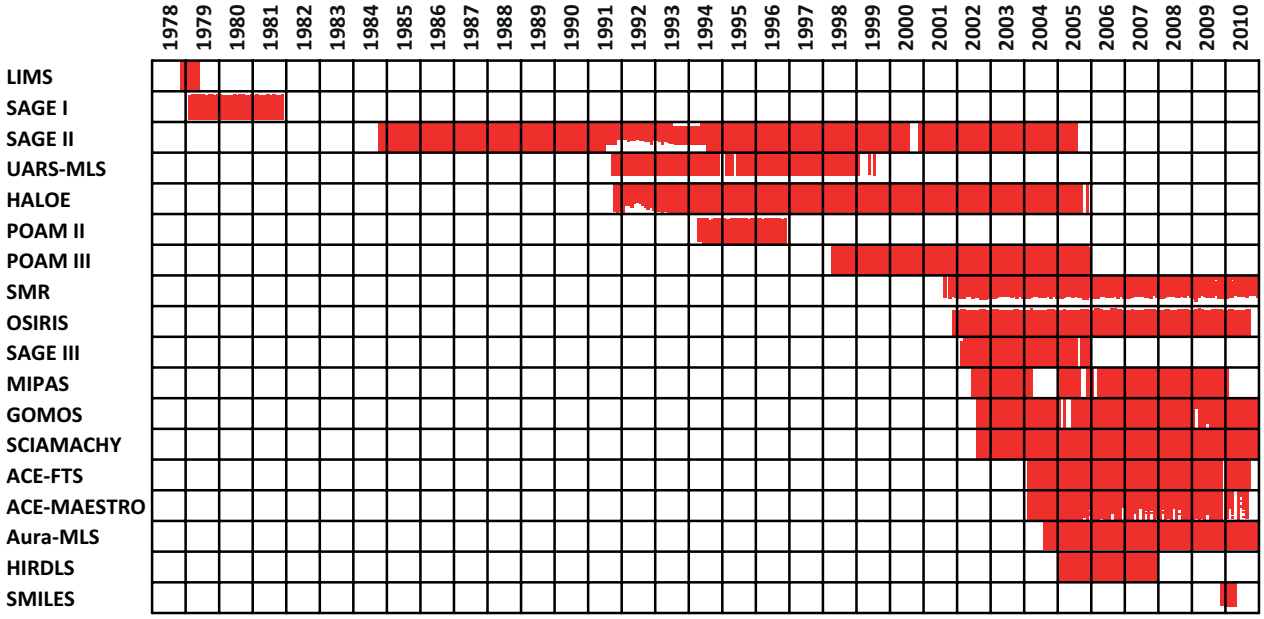
4.1 Ozone – O_3

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

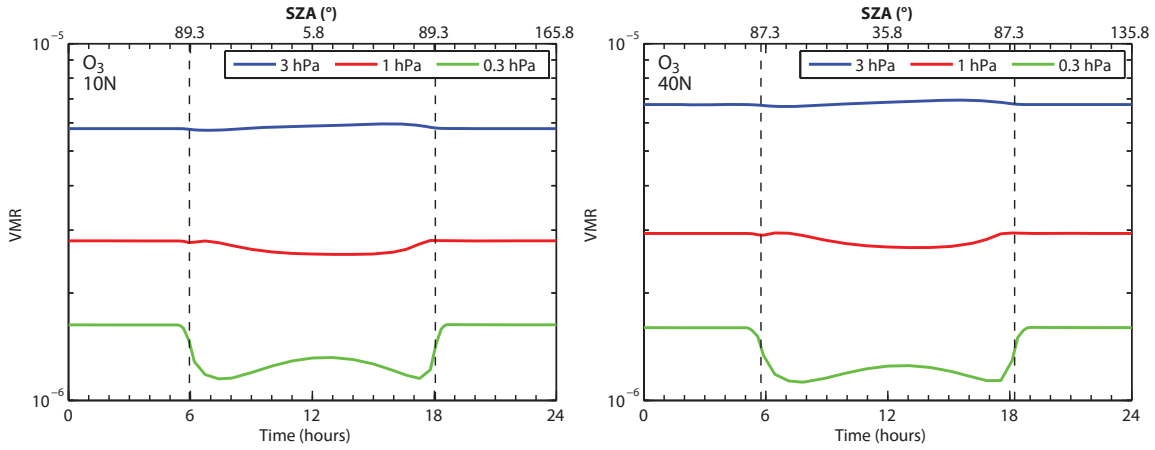


Figure 4.1.1: Diurnal ozone cycle. *Ozone variations as function of local solar time are shown at 10°N and 40°N at 3, 1 and 0.3 hPa for March 15.*

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

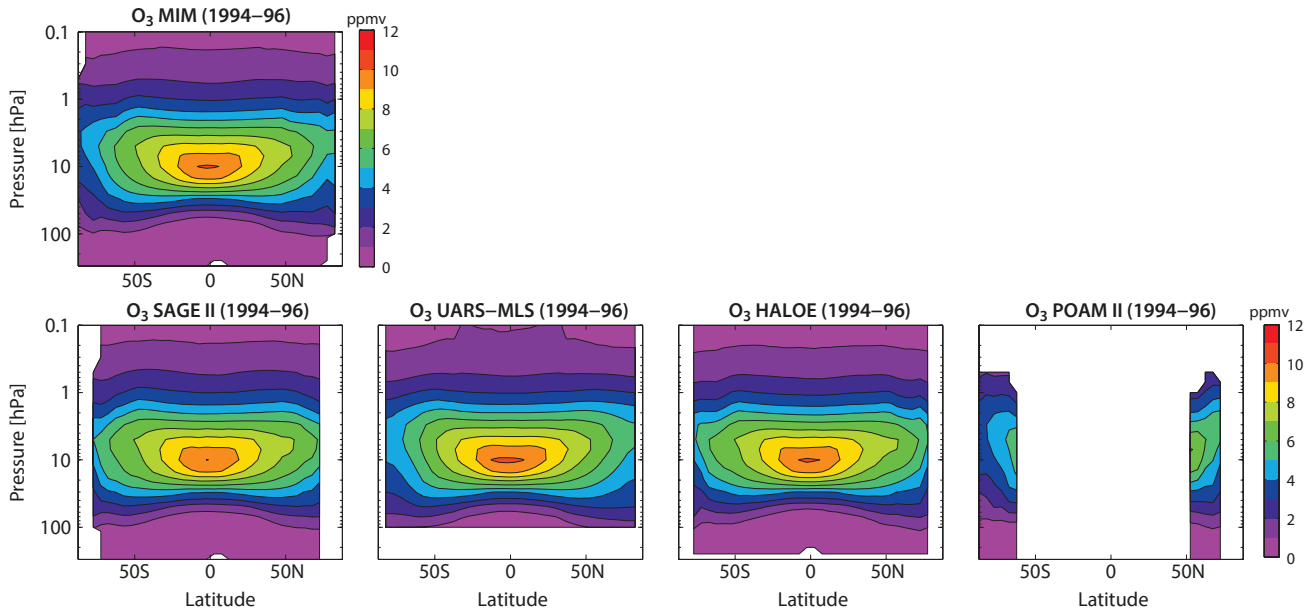


Figure 4.1.2: Cross sections of annual zonal mean ozone for 1994-1996. *Annual zonal mean ozone cross sections are shown for the MIM in the upper panel and for SAGE II, UARS-MLS, HALOE, and POAM II in the lower panels.*

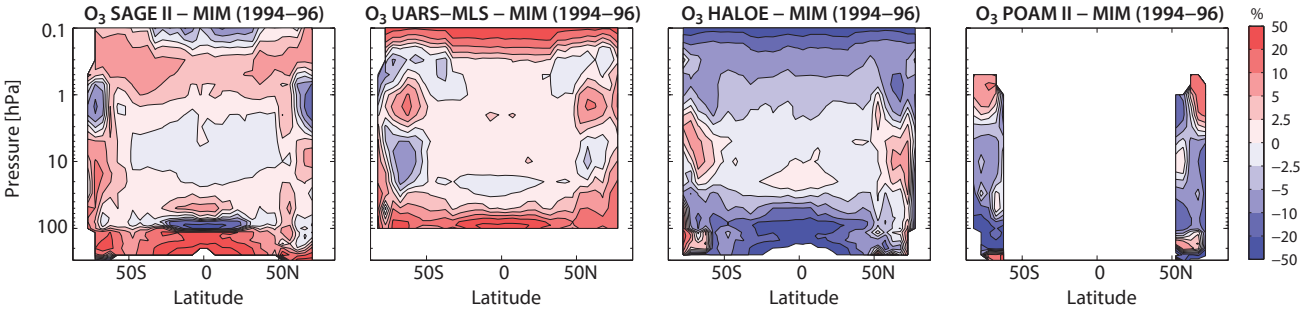


Figure 4.1.3: Cross sections of annual zonal mean ozone differences for 1994-1996. *Annual zonal mean ozone differences between the individual instruments and the MIM are shown.*

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

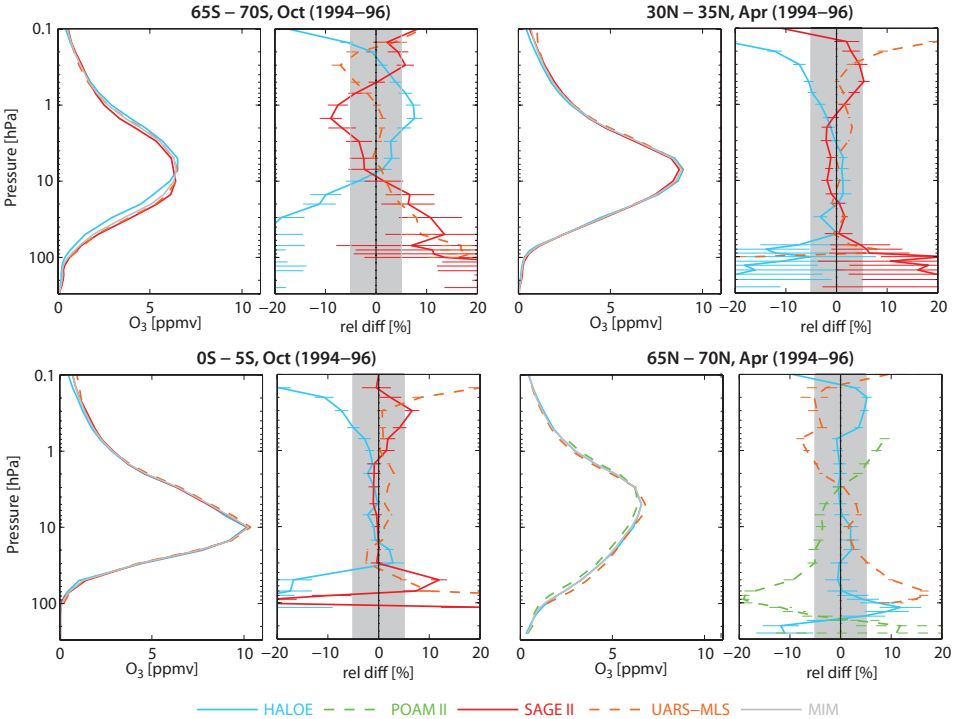


Figure 4.1.4: Vertical Profiles of monthly zonal mean ozone for 1994-1996. *Zonal mean ozone profiles for 65°S-70°S and 0°S-5°S for October in the left column and for 30°N-35°N and 65°N-70°N for April in the right column. Differences between the individual instruments and the MIM profiles are shown on 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%.*

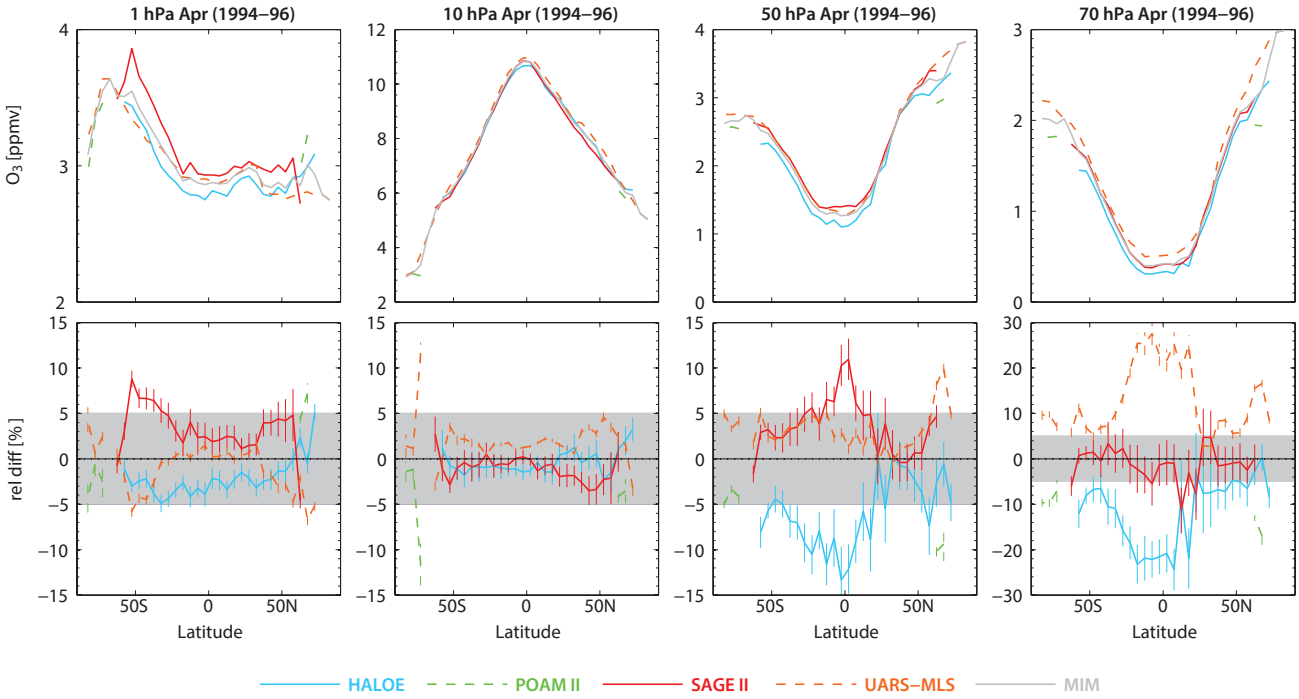


Figure 4.1.5: Meridional profiles of monthly zonal mean ozone for 1994-1996. *Meridional zonal mean ozone profiles at 1, 10, 50 and 70 hPa for April 1994-1996 are shown in the upper row. Differences between the individual instruments and the MIM profiles are shown in the lower row. Error bars indicate 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

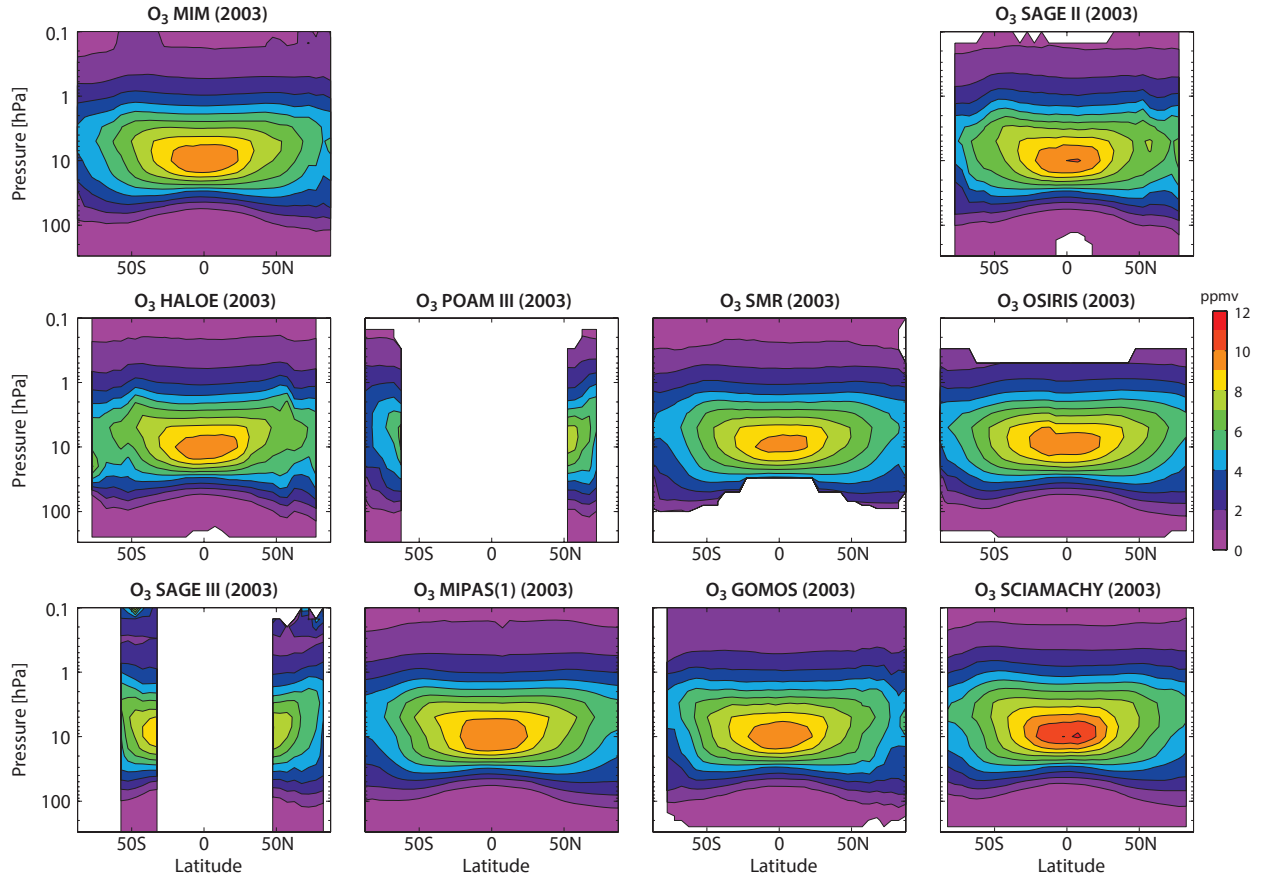


Figure 4.1.6: Cross sections of annual zonal mean ozone for 2003. *Annual zonal mean ozone cross sections for 2003 are shown for the MIM, SAGE II, HALOE, POAM III, SMR, OSIRIS, SAGE III, MIPAS(1), GOMOS, and SCIAMACHY.*

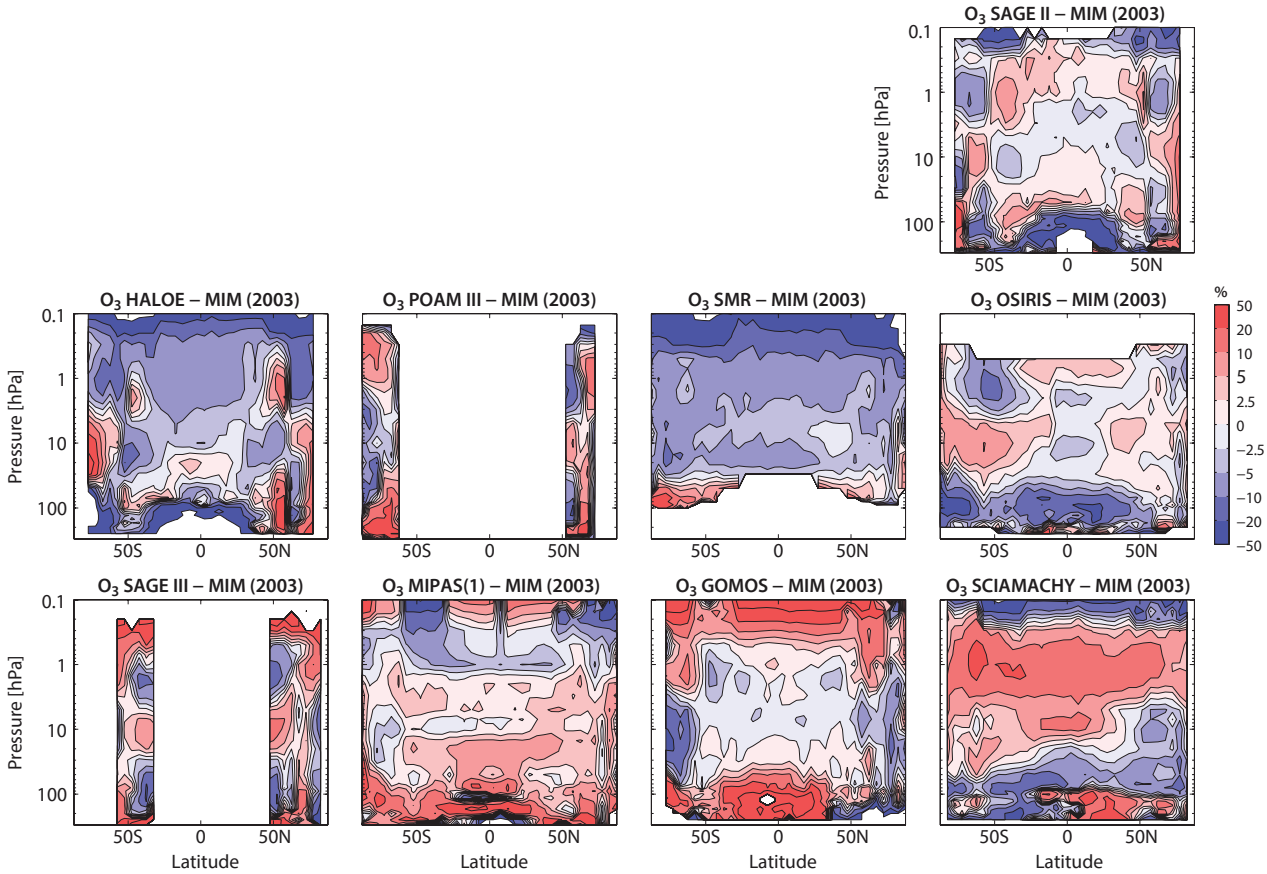


Figure 4.1.7: Cross sections of annual zonal mean ozone differences for 2003. *Annual zonal mean ozone differences for 2003 between the individual datasets and the MIM are shown.*

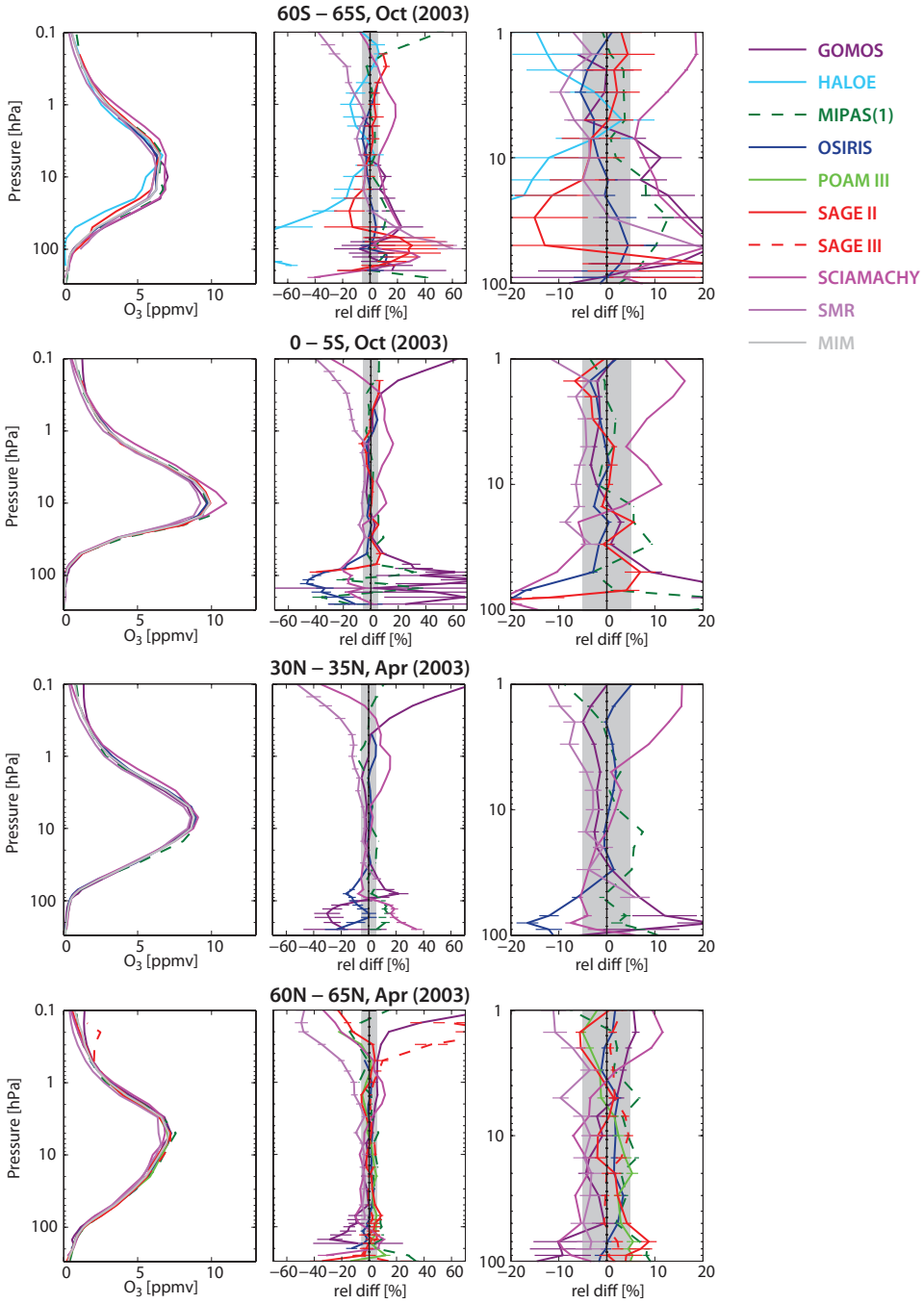


Figure 4.1.8: Vertical profiles of monthly zonal mean ozone for 2003. *Monthly zonal mean ozone profiles for 60°S-65°S and 0°S-5°S for October 2003 and for 30°N-35°N and 60°N-65°N for April 2003 are shown in the leftmost column. Differences between the individual instruments and the MIM profiles are 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

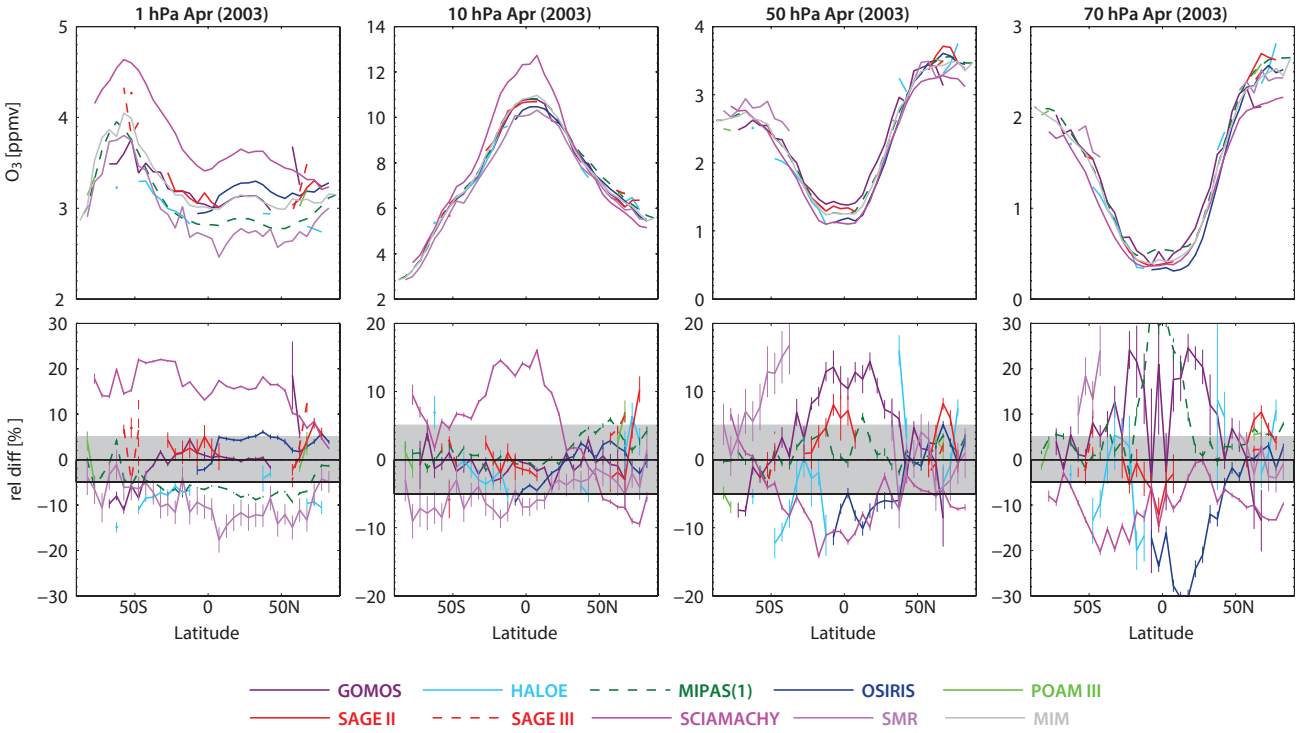


Figure 4.1.9: Meridional profiles of monthly zonal mean ozone for 2003. *Meridional zonal mean ozone profiles at 1, 10, 50 and 70 hPa for April 2003 are shown in the upper row. Differences between the individual instruments and the MIM profiles are shown in the lower row. Error bars indicate the uncertainties 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

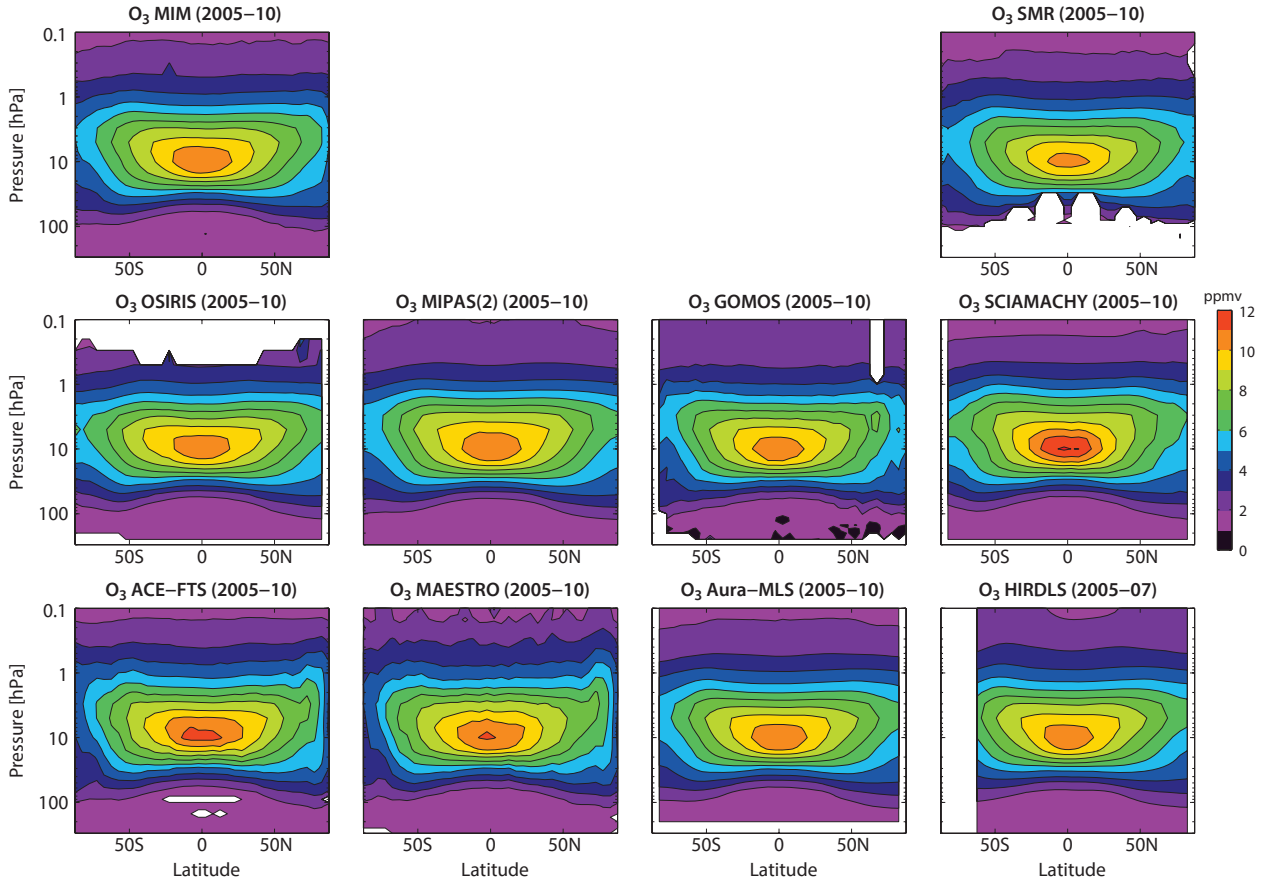


Figure 4.1.10: Cross sections of annual zonal mean ozone for 2005-2010. *Annual zonal mean ozone cross sections for 2005-2010 are shown for the MIM, SMR, OSIRIS, MIPAS(2), GOMOS, SCIAMACHY, ACE-FTS, ACE-MAESTRO, Aura-MLS and HIRDLS. Note that HIRDLS covers only 2005-2007 and is not included in the MIM.*

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

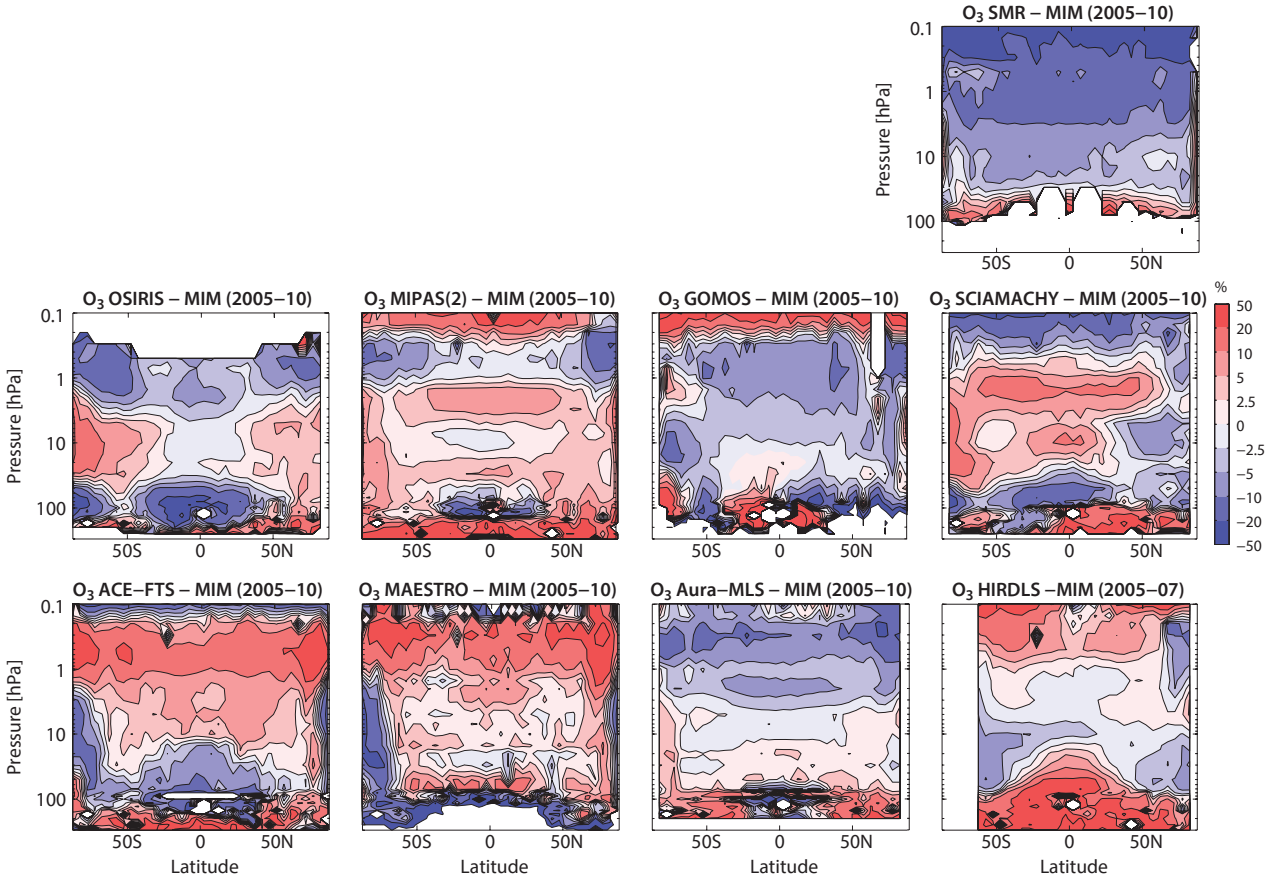


Figure 4.1.11: Cross sections of annual zonal mean ozone differences for 2005-2010. *Annual zonal mean ozone differences for 2005-2010 between the individual datasets and the MIM are shown. Note that HIRDLS covers only 2005-2007 and is not included in the MIM.*

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

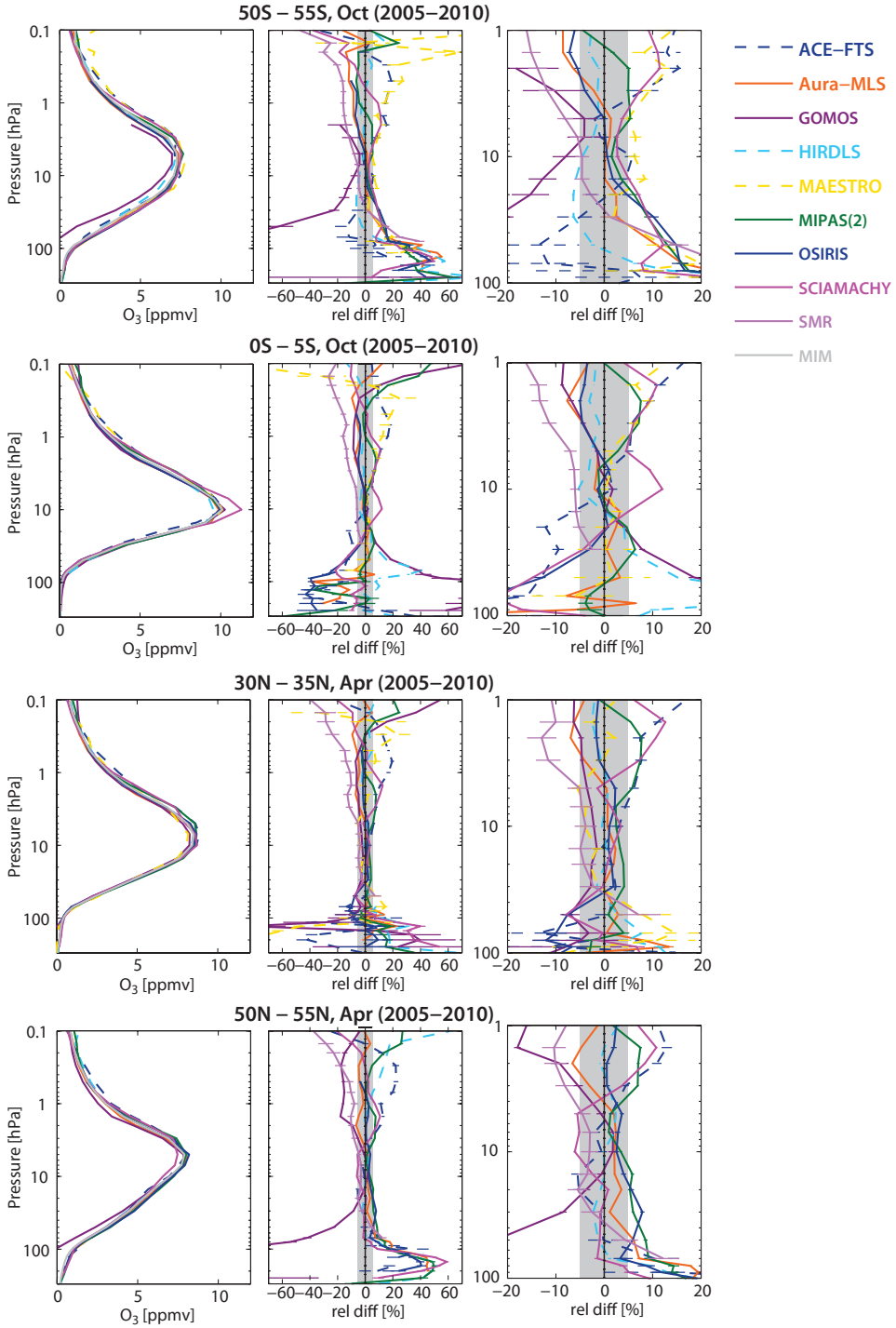


Figure 4.1.12: Vertical profiles of monthly zonal mean ozone for 2005-2010. *Zonal mean ozone profiles for 50°S-55°S and 0°S-5°S for October 2005-2010 and for 30°N-35°N and 50°N-55°N for April 2005-2010 are shown in the leftmost column. Differences between the individual instruments and the MIM profiles 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

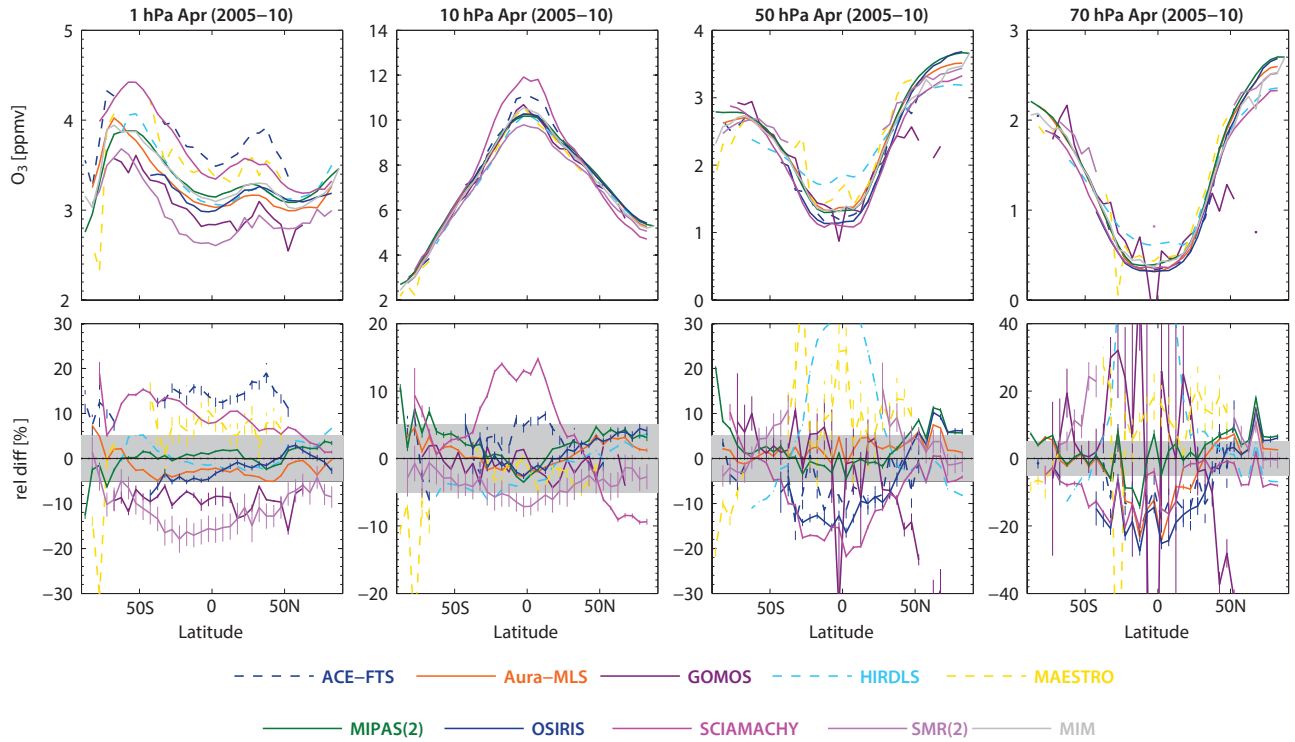


Figure 4.1.13: Meridional profiles of monthly zonal mean ozone for 2005-2010. *Meridional zonal mean ozone profiles at 1, 10, 50 and 70 hPa for April 2005-2010 are shown in the upper row. Differences between the individual instruments and the MIM profiles are shown in the lower row. Error bars indicate 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.*

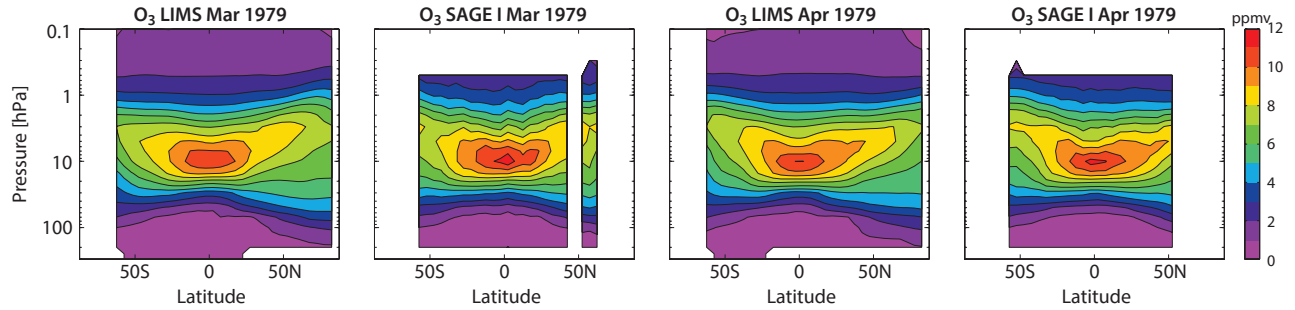


Figure 4.1.14: Cross sections of monthly zonal mean ozone for 1979. *Monthly zonal mean ozone cross sections for 1979 are shown for LIMS and SAGE I for March and April.*

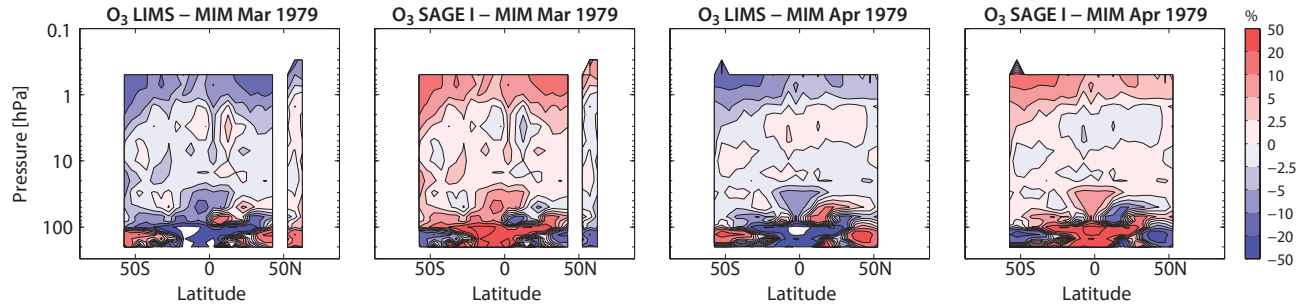


Figure 4.1.15: Cross sections of monthly zonal mean ozone differences for 1979. *Monthly zonal mean ozone differences for 1979 between LIMS, SAGE I and the MIM are shown for March and April.*

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

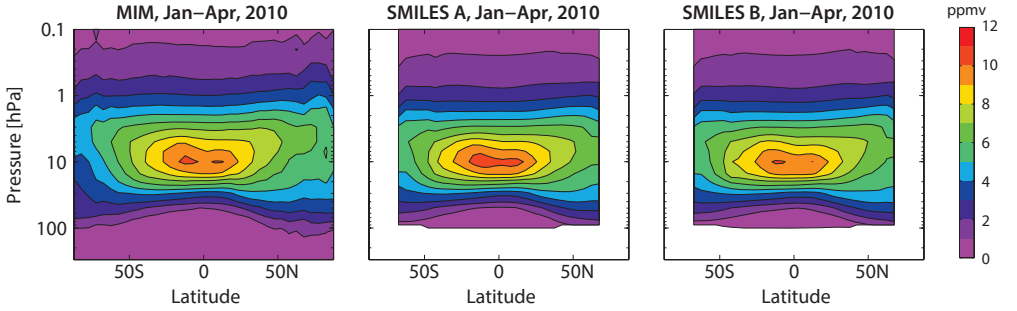


Figure 4.1.16: Cross sections of zonal mean ozone for 2010. *Zonal mean ozone cross sections for January-April 2010 are shown for the MIM, SMILES-A and SMILES-B.*

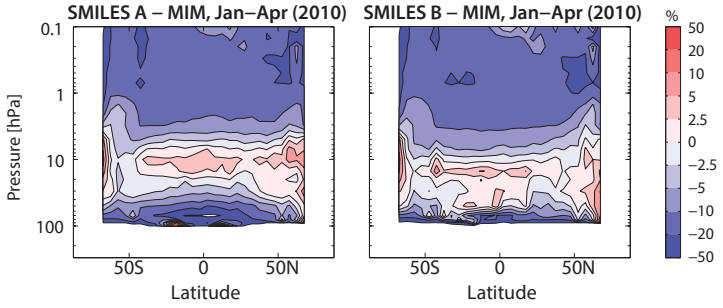


Figure 4.1.17: Cross sections of zonal mean ozone differences for 2010. *Zonal mean ozone differences for January-April 2010 between SMILES-A, SMILES-B and the MIM are shown.*

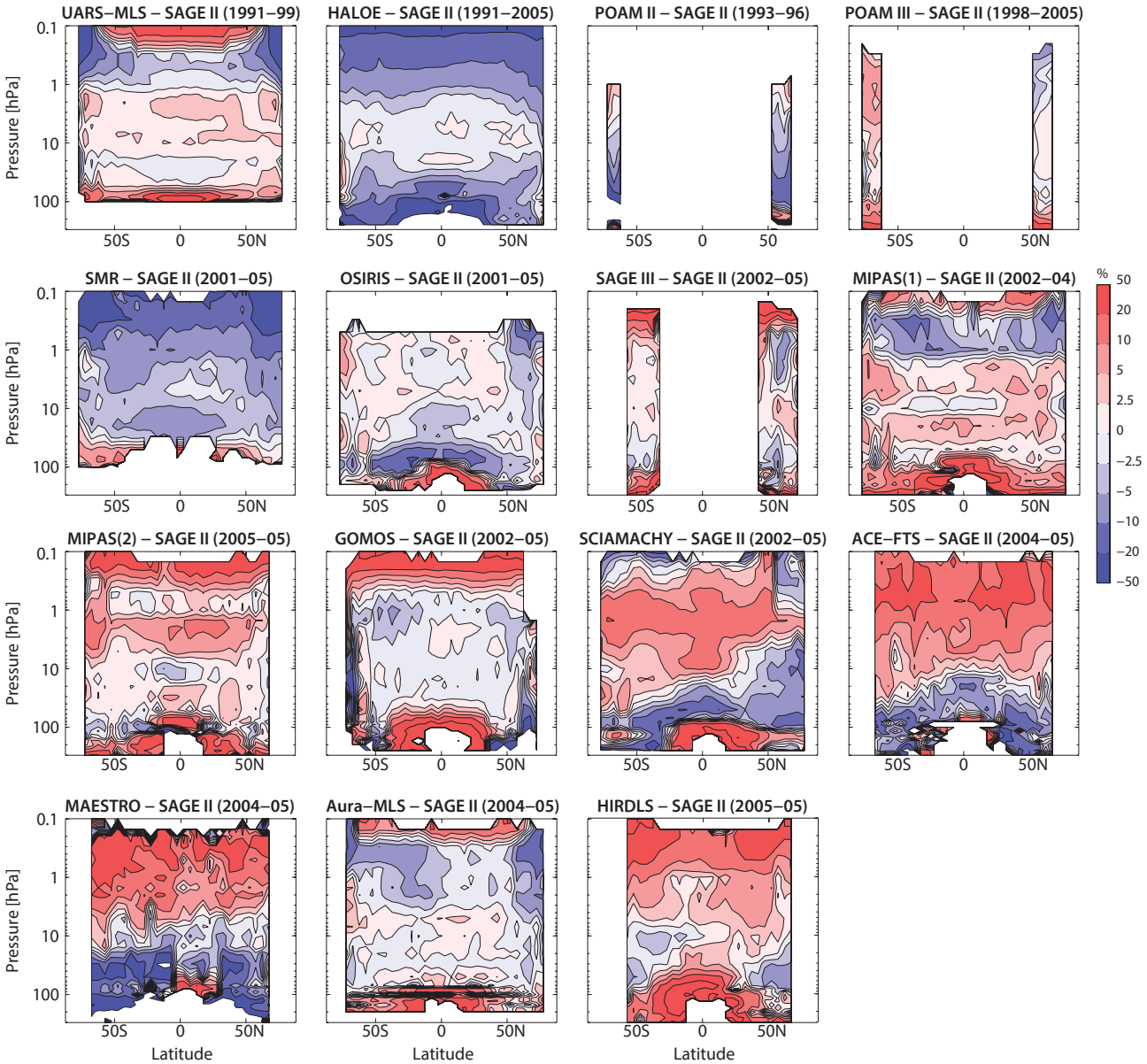


Figure 4.1.18: Cross sections of zonal mean ozone differences to SAGE II. *Zonal mean ozone differences between the individual instruments and SAGE II are shown for time periods of maximum overlap.*

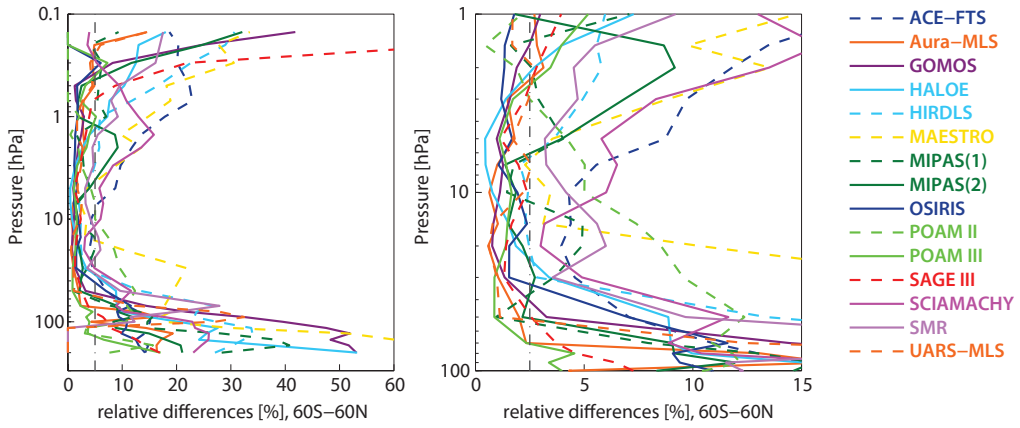


Figure 4.1.19: Vertical profiles of mean ozone differences to SAGE II between 60°S and 60°N. *Absolute values of the relative ozone differences averaged between 60°N and 60°S for time periods of maximum overlap between the individual instruments (HALOE, POAM III, SAGE III, ACE-FTS, Aura-MLS, GOMOS, MIPAS, 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.

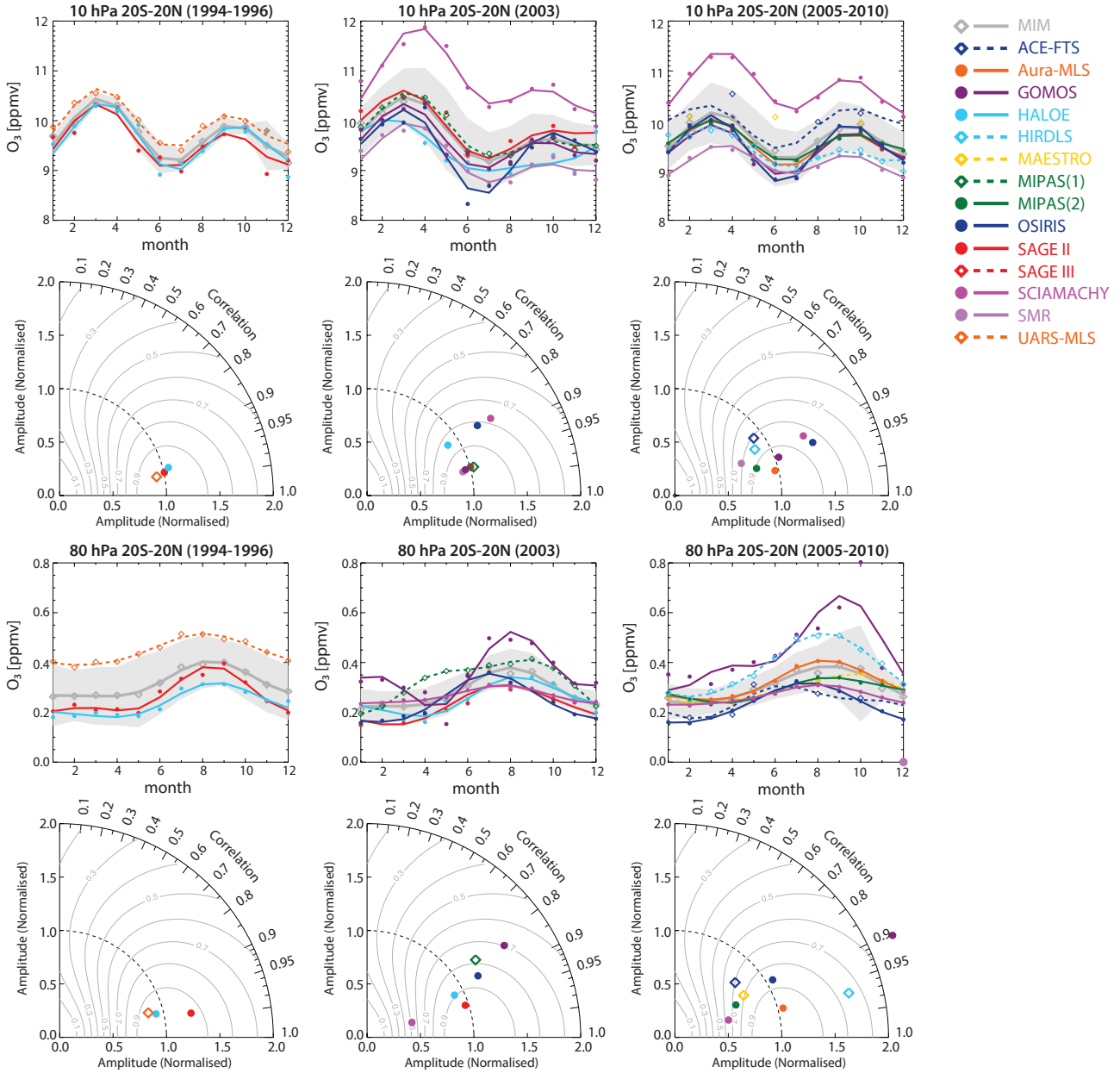


Figure 4.1.20: Seasonal cycle of ozone in the tropics at 10 and 80 hPa. *Seasonal cycle and corresponding Taylor diagram of monthly zonal mean ozone for 20°S-20°N at 10 hPa and 80 hPa. The seasonal cycle is shown for 1994-1996 (left column), 2003 (middle column) and 2005-2010 (right column). The grey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

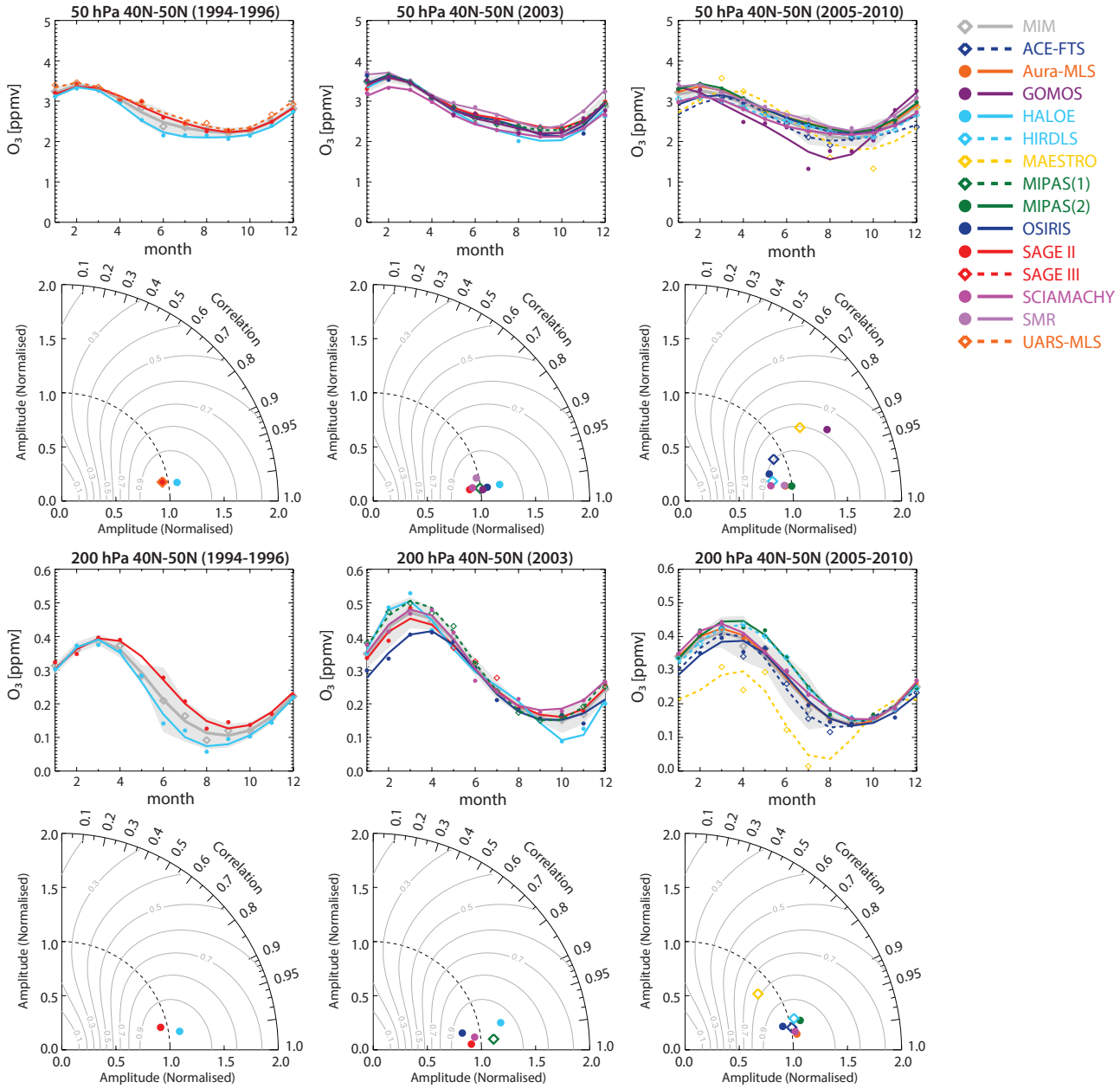


Figure 4.1.21: Seasonal cycle of ozone in the NH mid-latitudes at 50 and 200 hPa. *Seasonal cycle and corresponding Taylor diagram of monthly zonal mean ozone for 40°N-50°N at 50 hPa and 200 hPa. The seasonal cycle is shown for 1994-1996 (left column), 2003 (middle column) and 2005-2010 (right column).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

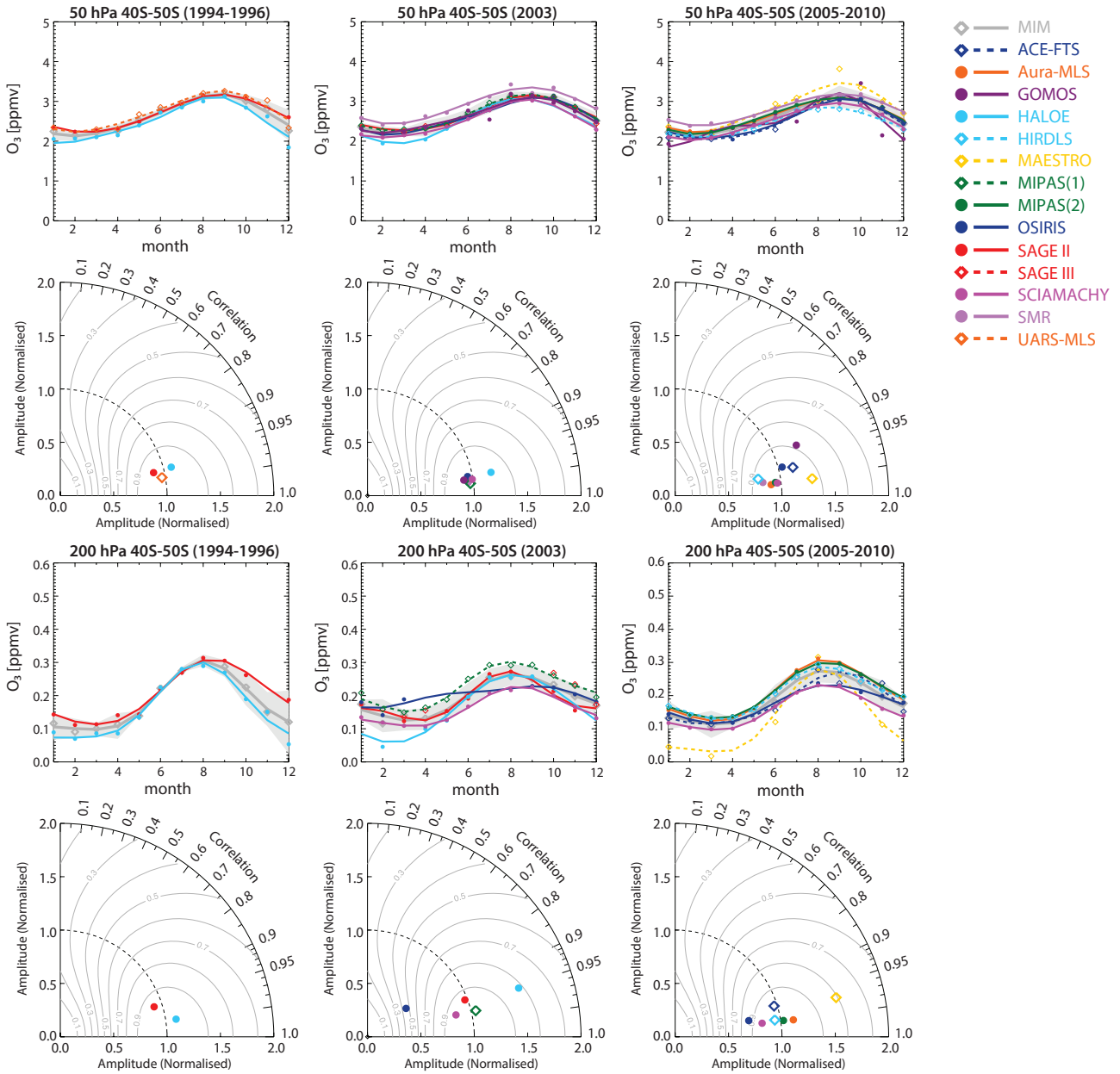


Figure 4.1.22: Seasonal cycle of ozone in the SH mid-latitudes at 50 and 200 hPa. *Seasonal cycle and corresponding Taylor diagram of monthly zonal mean ozone for 40°S-50°S at 50 hPa and 200 hPa. The seasonal cycle is shown for 1994-1996 (left column), 2003 (middle column) and 2005-2010 (right column).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

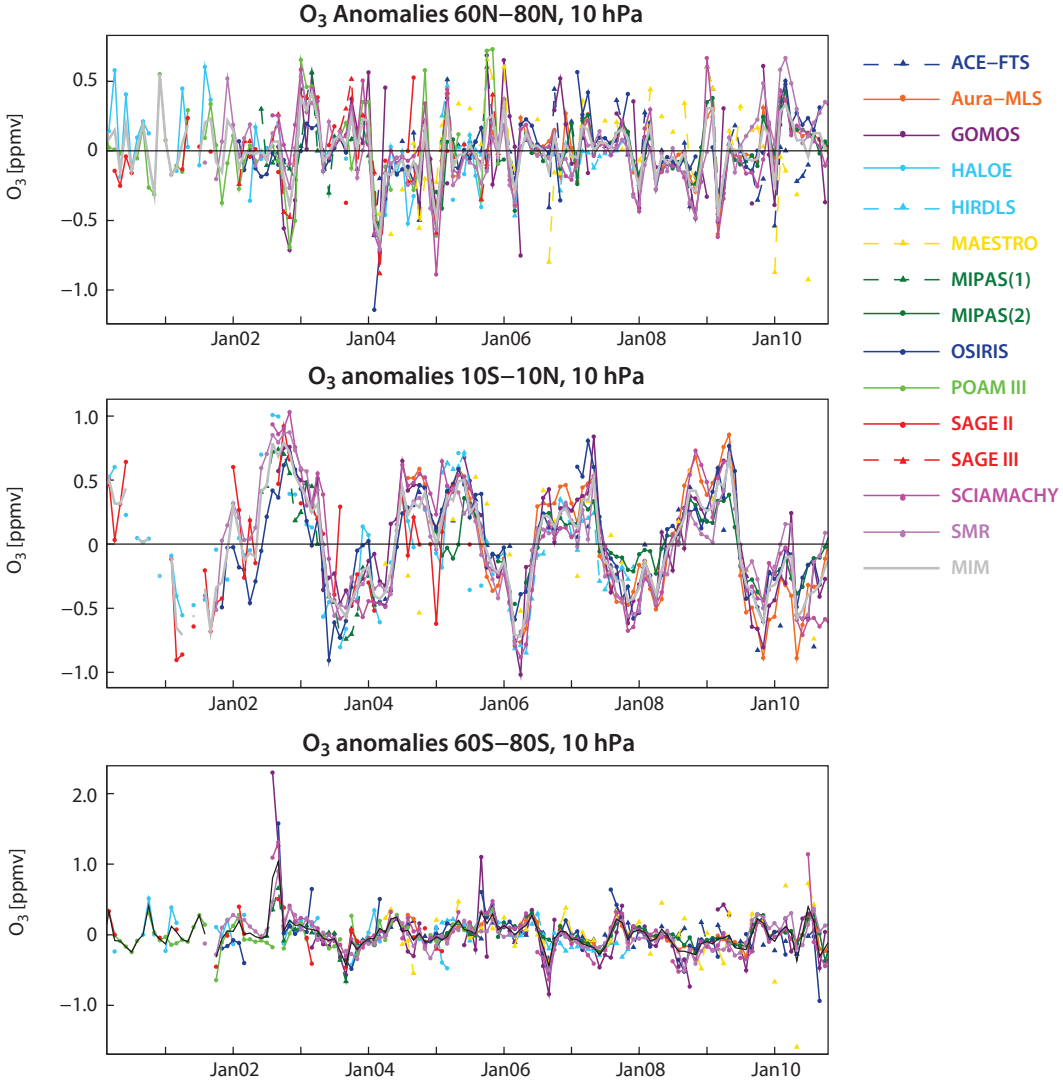


Figure 4.1.23: Time series of deseasonalised ozone anomalies at 10 hPa. *Deseasonalised ozone anomalies at 10 hPa between 60°N – 80°N (upper panel), 10°S – 10°N (middle panel) and 60°S – 80°S (lower panel).*

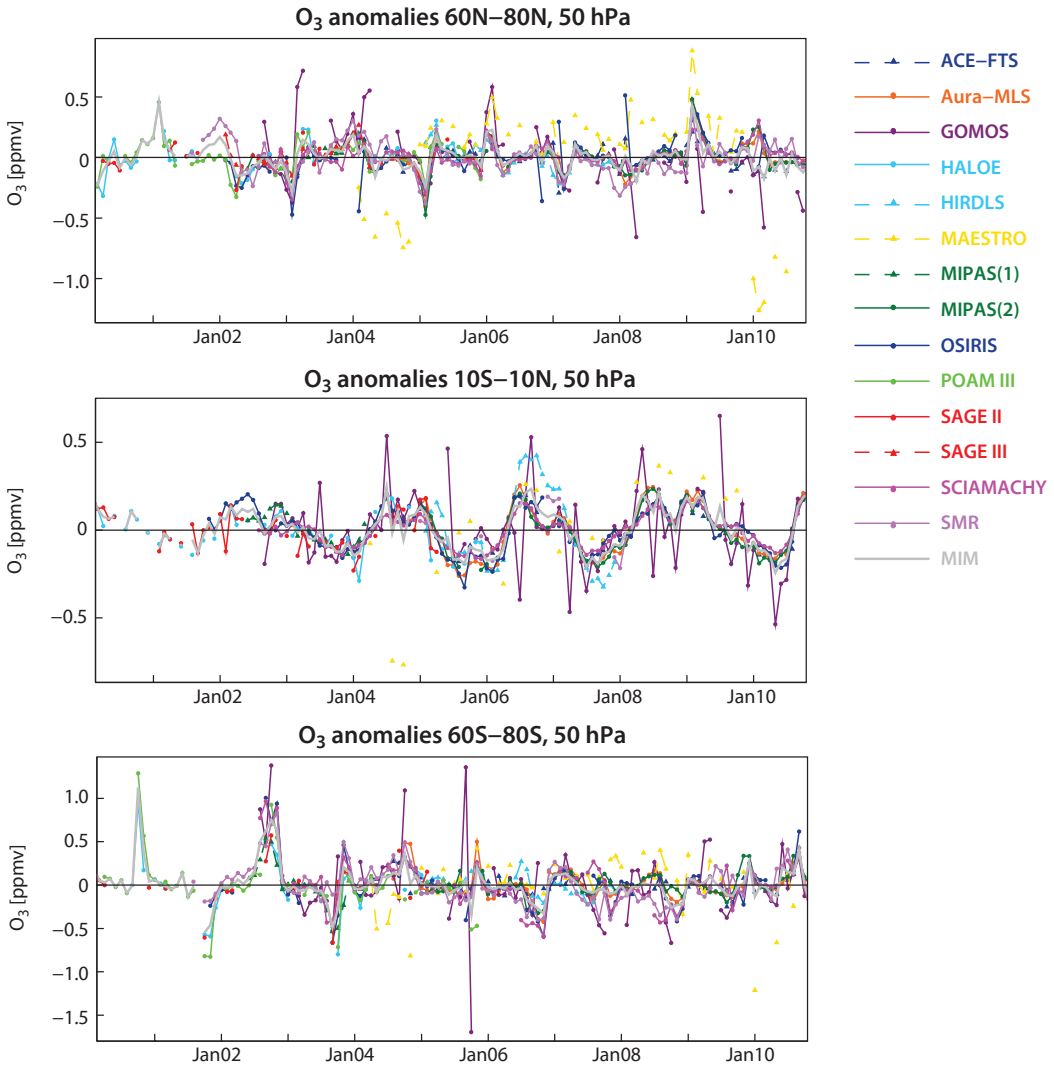


Figure 4.1.24: Time series of deseasonalised ozone anomalies at 50 hPa. *Deseasonalised ozone anomalies at 50 hPa between 60°N – 80°N (upper panel), 10°S – 10°N (middle panel) and 60°S – 80°S (lower panel).*

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

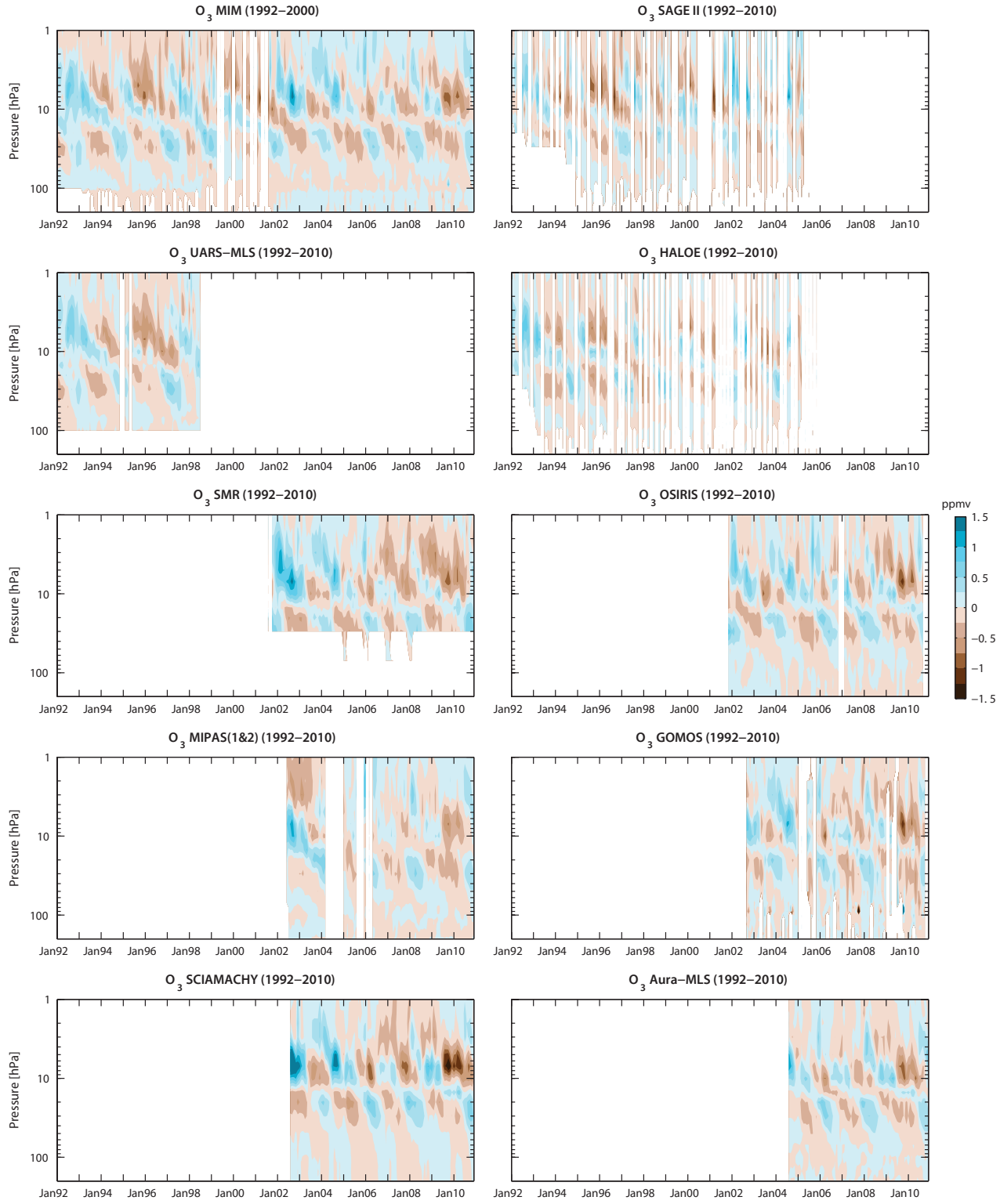


Figure 4.1.25: QBO signal for 1992-2010. *Altitude-time sections of deseasonalised ozone anomalies for 10°S-10°N from 1992 to 2010 are shown. The MIM is based on all displayed datasets.*

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

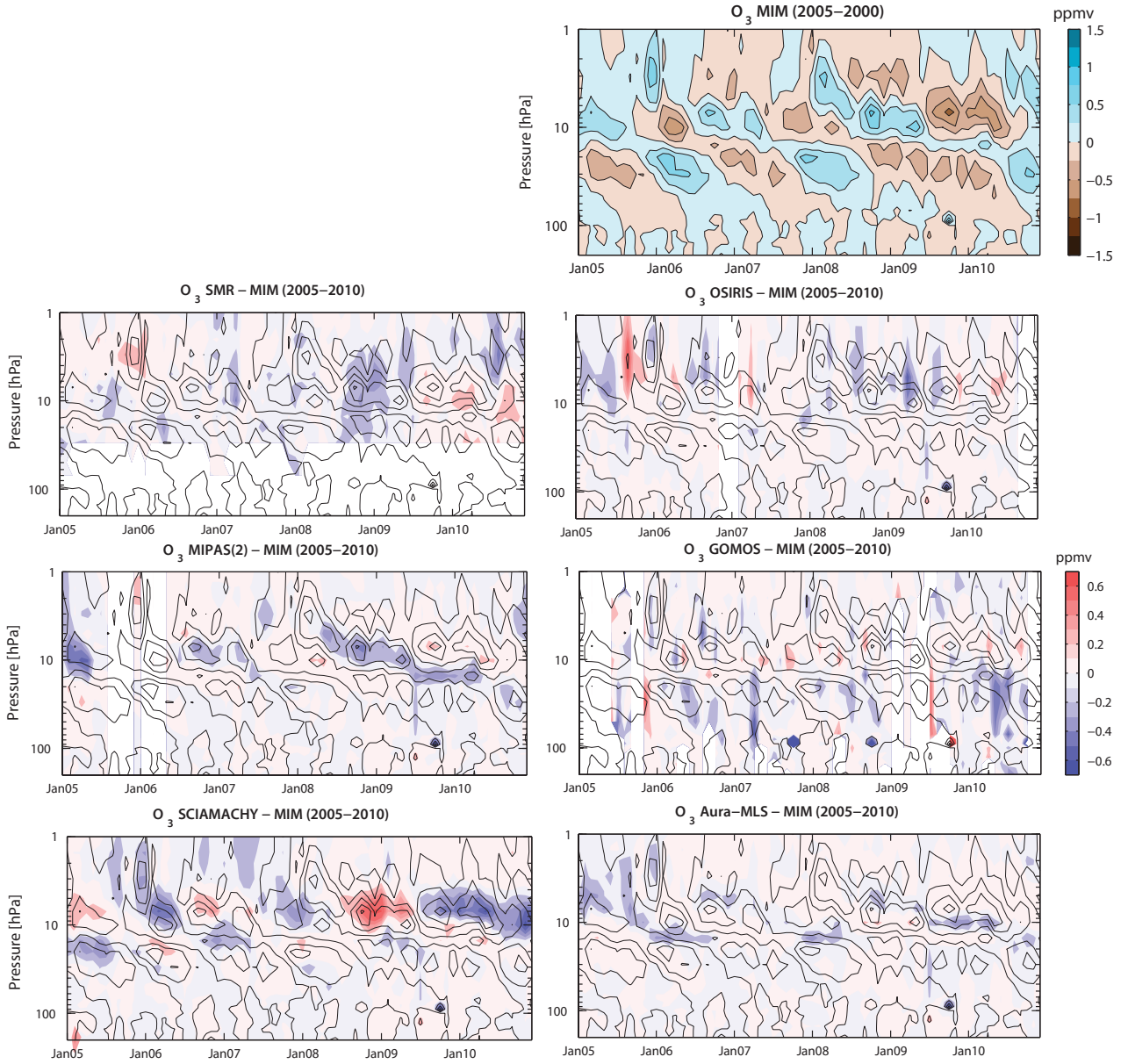


Figure 4.1.26: Differences with respect to MIM for QBO signal for 2005-2010. *Altitude-time section of MIM deseasonalised ozone anomalies for 10°S-10°N from 2005 to 2010 (upper left panel). Ozone anomaly differences between the individual datasets and the MIM are shown in the other panels by colour contours.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

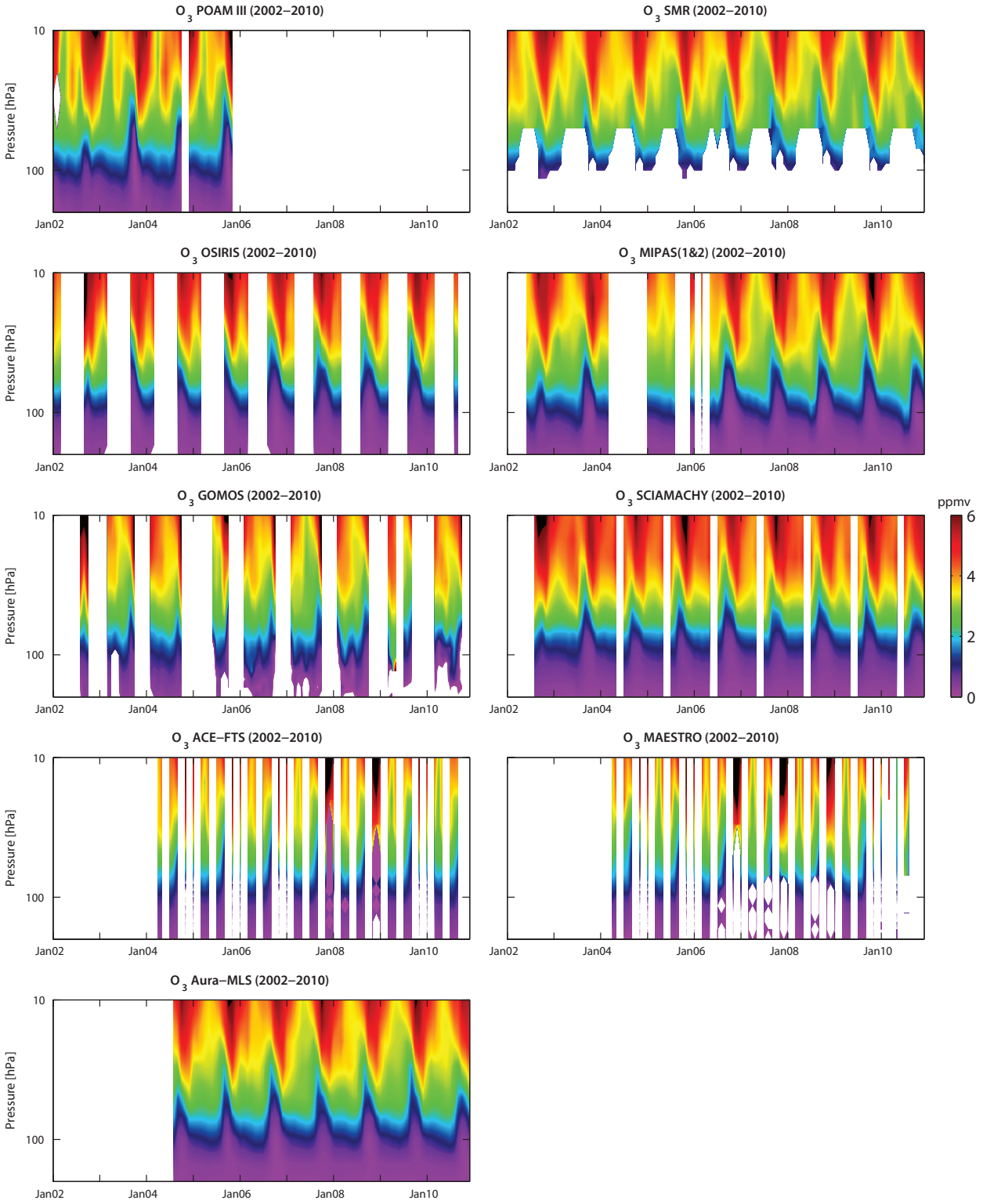


Figure 4.1.27: Antarctic ozone hole for 2002-2010. *Altitude-time sections of monthly zonal mean ozone for 60°S-90°S from 2002 to 2010 are shown.*

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

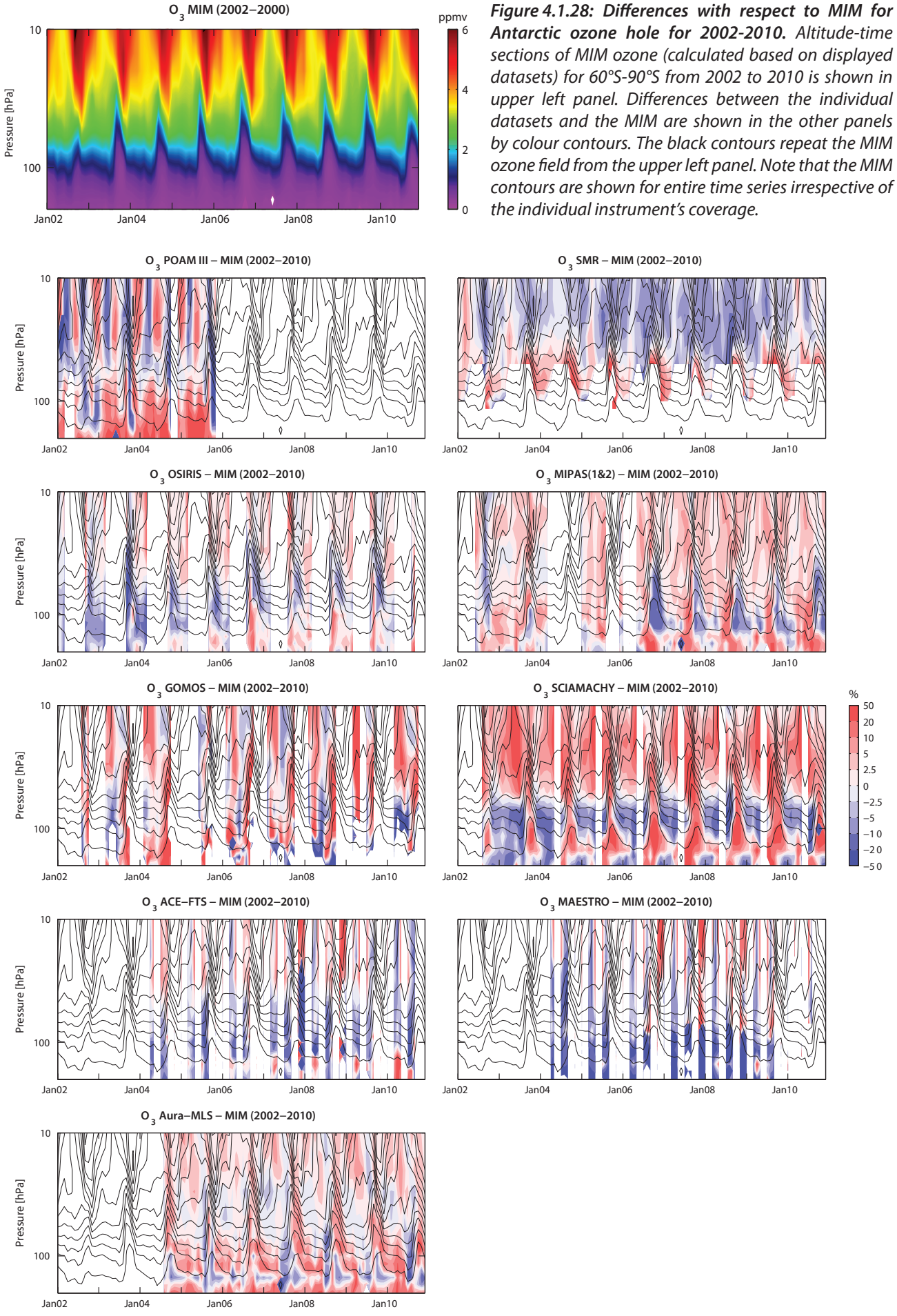


Figure 4.1.28: Differences with respect to MIM for Antarctic ozone hole for 2002-2010. *Altitude-time sections of MIM ozone (calculated based on displayed datasets) for 60°S-90°S from 2002 to 2010 is shown in upper left panel. Differences between the individual datasets and the MIM are shown in the other 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.*

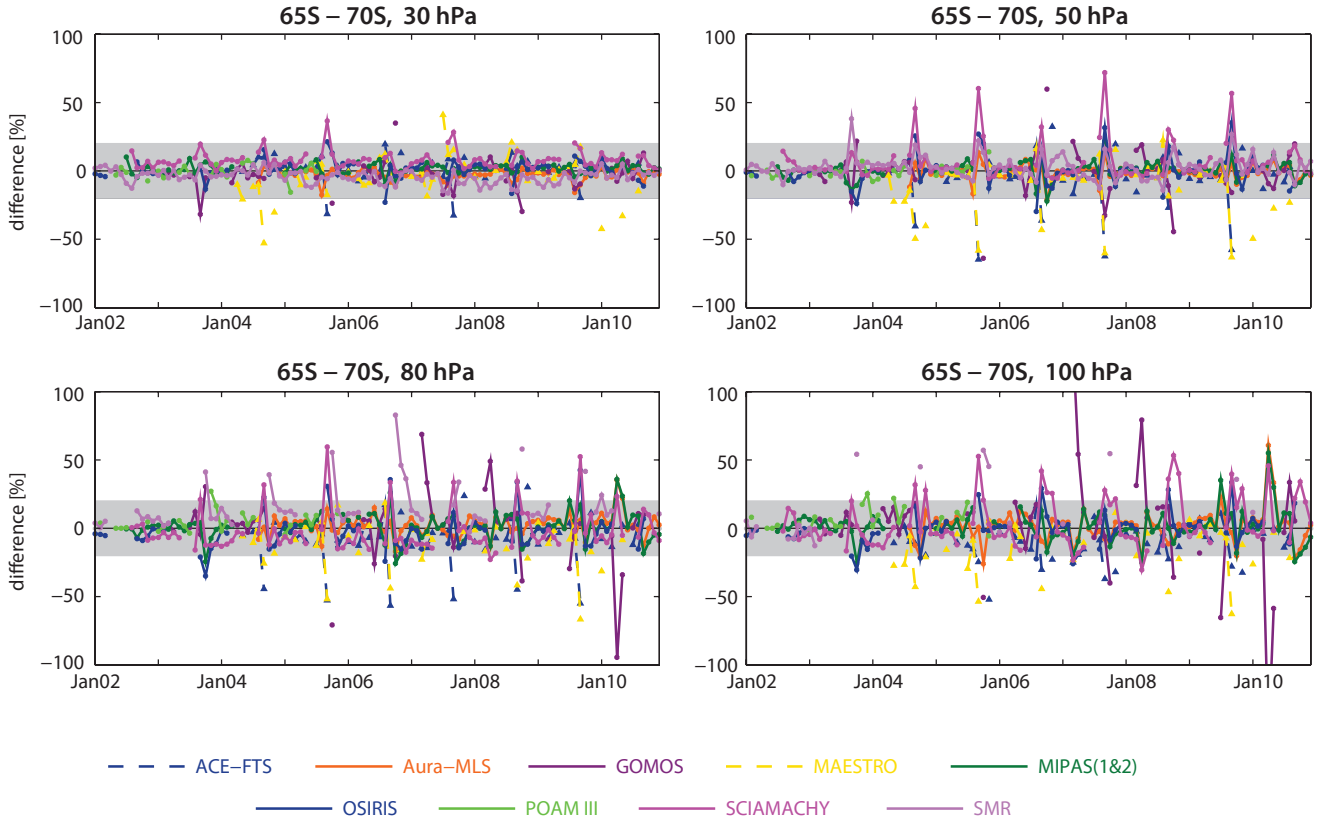


Figure 4.1.29: Time series of relative differences with respect to MIM for ozone at 65°S-70°S. *Time series of the relative differences between the individual instruments and the MIM at 30, 50, 80 and 100 hPa for 65°S-70°S are shown. The grey shaded area indicates where relative differences are smaller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

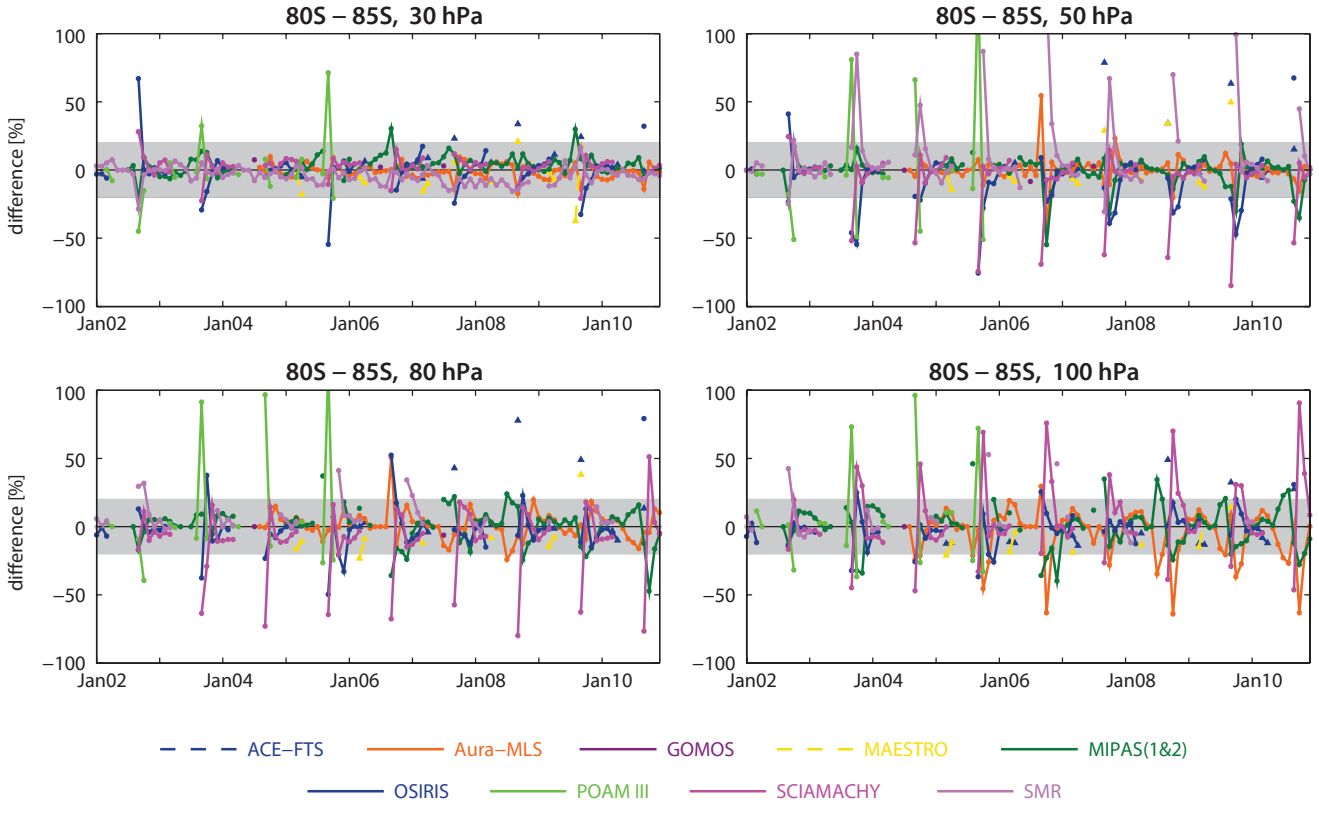


Figure 4.1.30: Time series of relative differences with respect to MIM for ozone at 80°S-85°S. *Time series of the relative differences between the individual instruments and the MIM at 30, 50, 80 and 100 hPa for 80°S-85°S are shown. The grey shaded area indicates where relative differences are smaller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

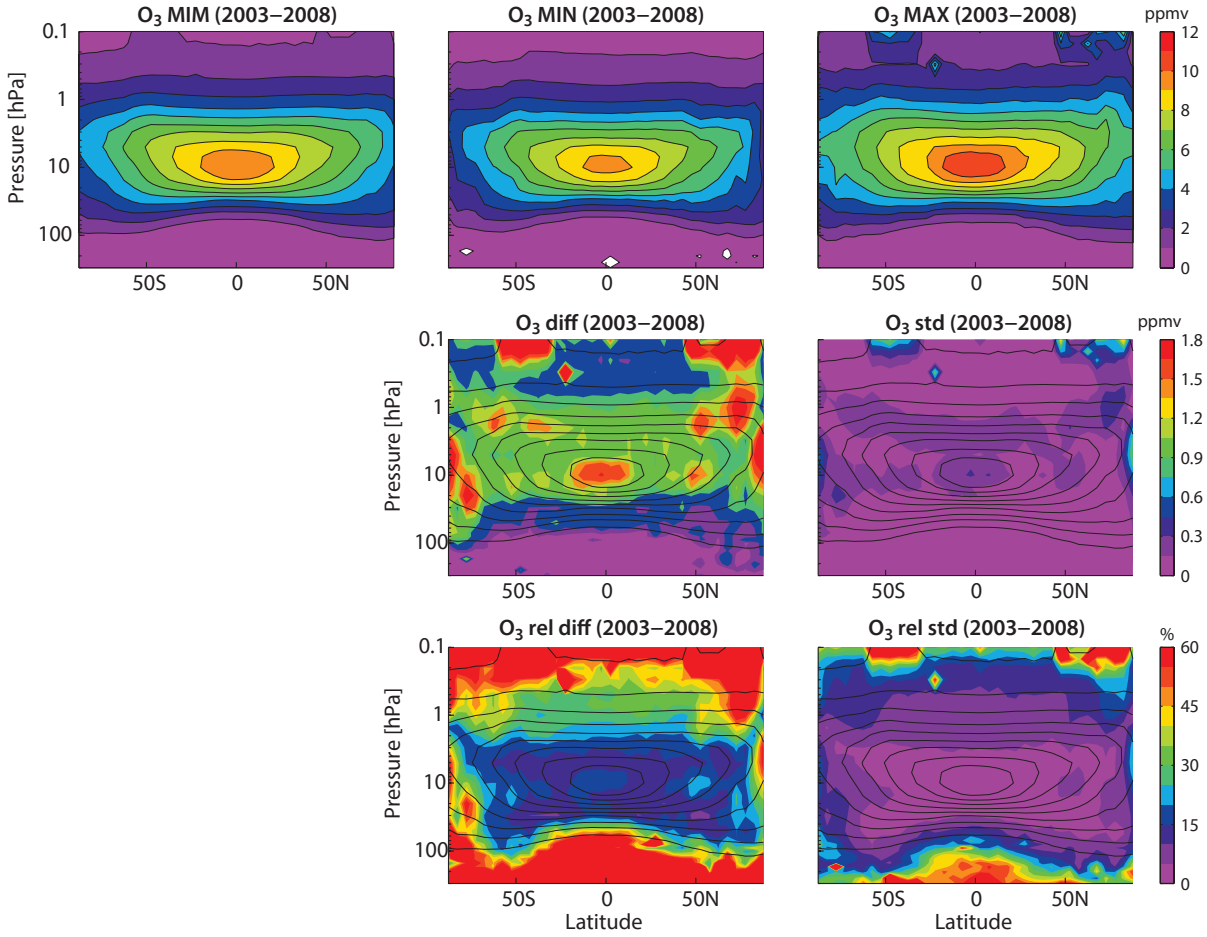


Figure 4.1.31: Summary of ozone annual zonal mean state for 2003-2008. *Shown are the annual zonal mean cross sections of the MIM, minimum (MIN), and maximum (MAX) ozone values in the upper row, the maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments in 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.

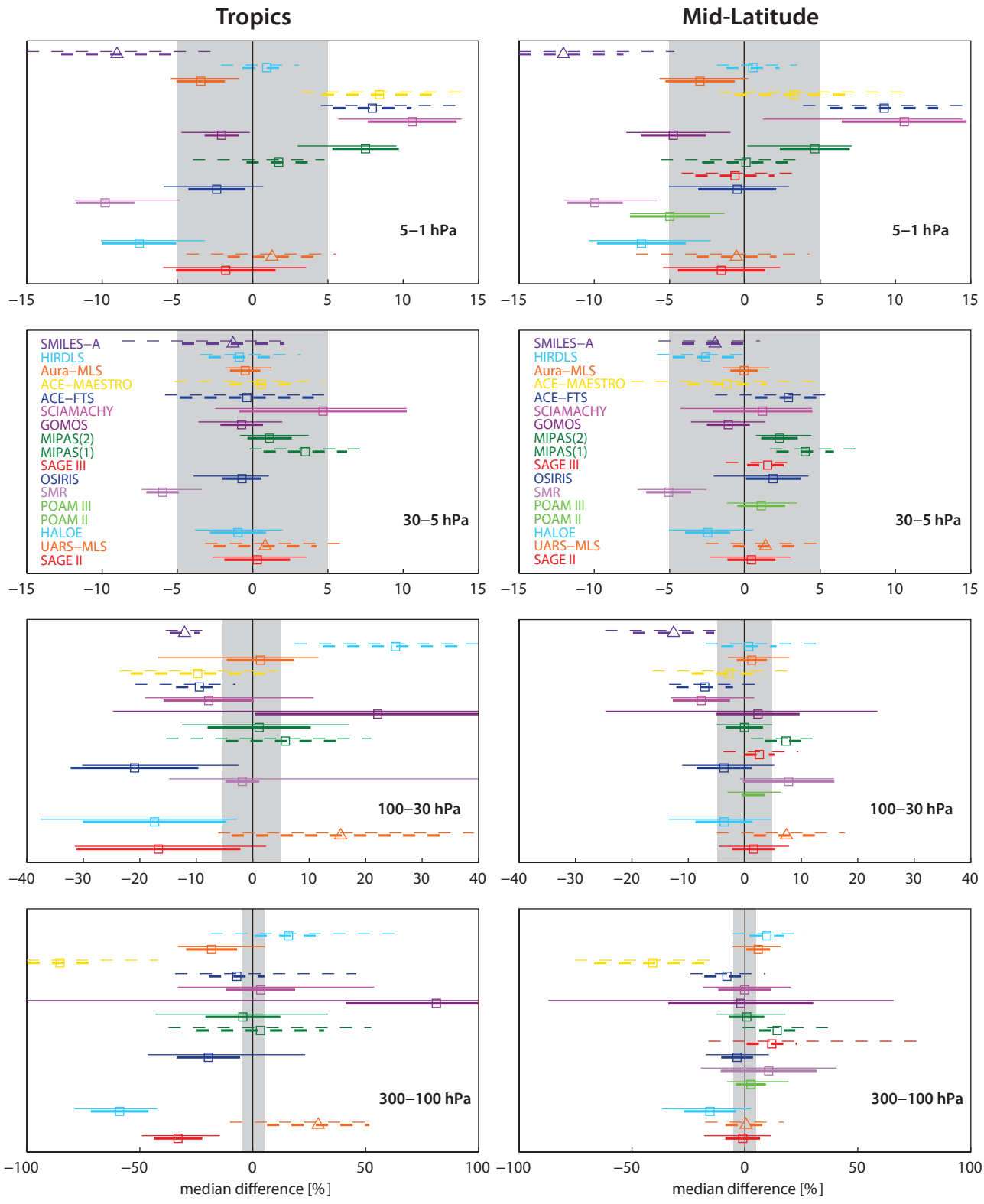


Figure 4.1.32: Summary plot of ozone differences for 2003-2008. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology 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.*

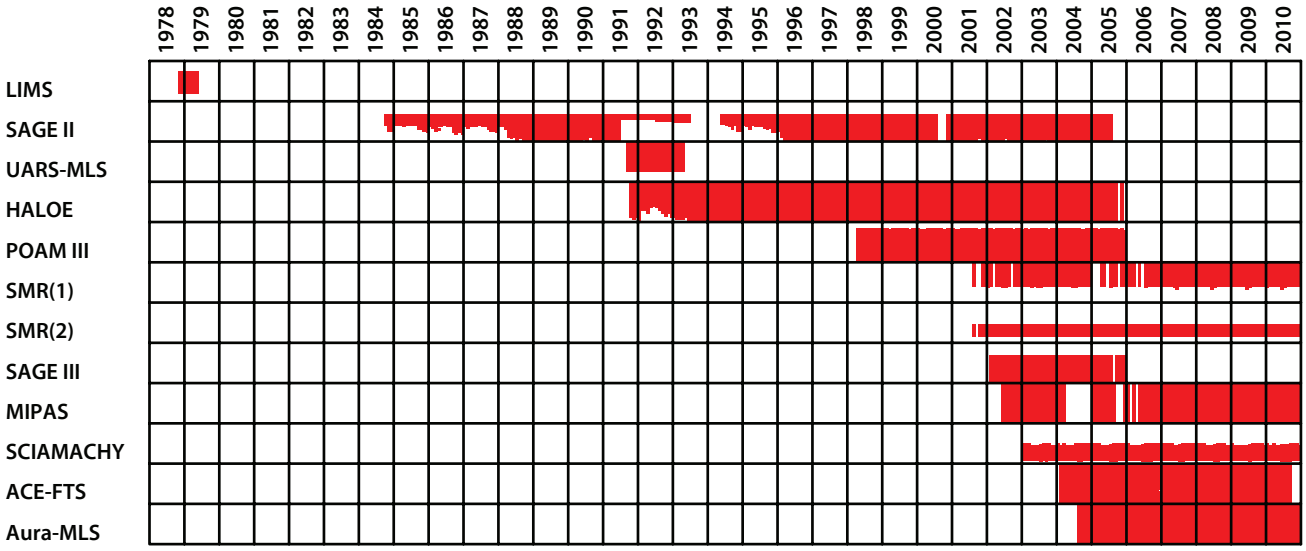


Table 4.2.2: Data version, time period, vertical range, vertical resolution, references and other comments for H_2O profile measurements used to generate the SPARC Data Initiative monthly zonal mean climatologies.

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

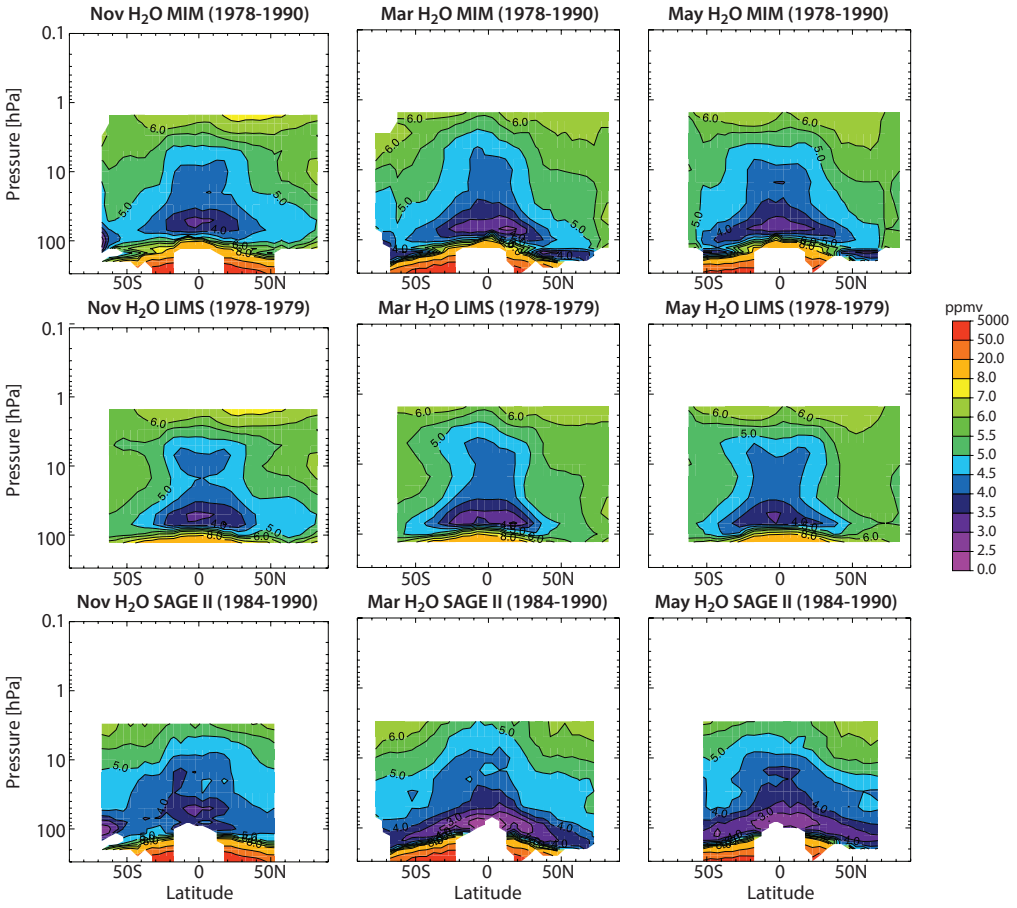


Figure 4.2.1a: Cross sections of monthly zonal mean H_2O for LIMS (1978-1979) and SAGE II (1984-1990). *November, March and May monthly zonal means (from left to right) are shown for the MIM, LIMS, and SAGE II (from top to bottom).*

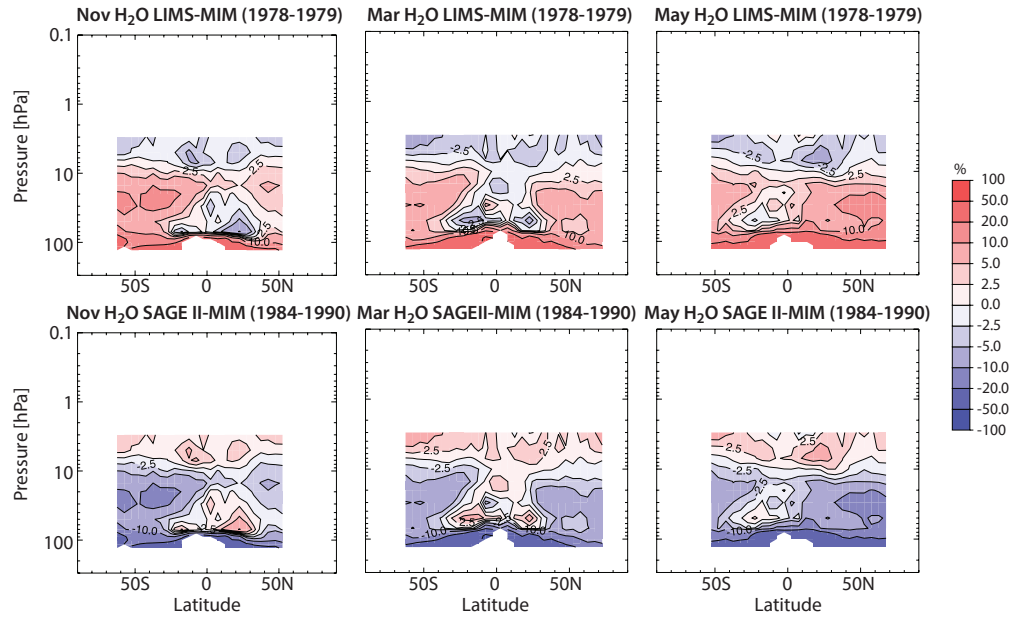


Figure 4.2.1b: Cross sections of monthly zonal mean H_2O differences relative to the MIM for LIMS (1978-1979) and SAGE II (1984-1990). *Same order as in Figure 4.2.1a.*

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

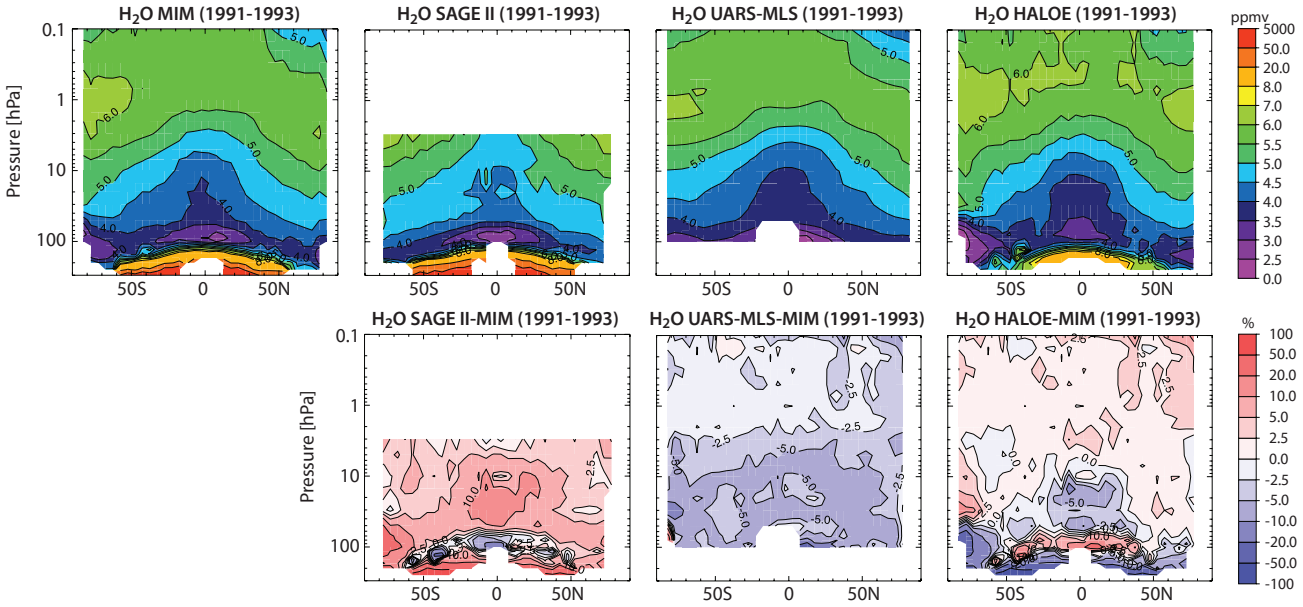


Figure 4.2.2: Cross sections of annual zonal mean H_2O and differences for 1991-1993. *Shown from left to right are the MIM, SAGE II, UARS-MLS, and HALOE. Upper panels show absolute values, lower panels the differences relative to the MIM.*

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

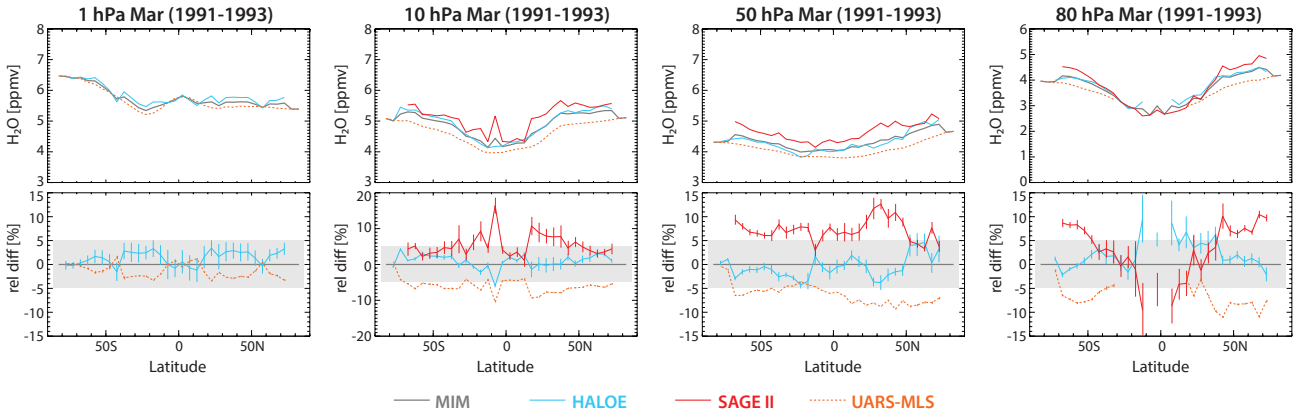


Figure 4.2.3: Meridional profiles of monthly zonal mean H_2O for 1991-1993. *Shown are meridional profiles for March at 1, 10, 50, and 80 hPa (from left to right). Upper panels show absolute values, lower panels relative differences between the individual instruments (SAGE II, HALOE, and UARS-MLS) and 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

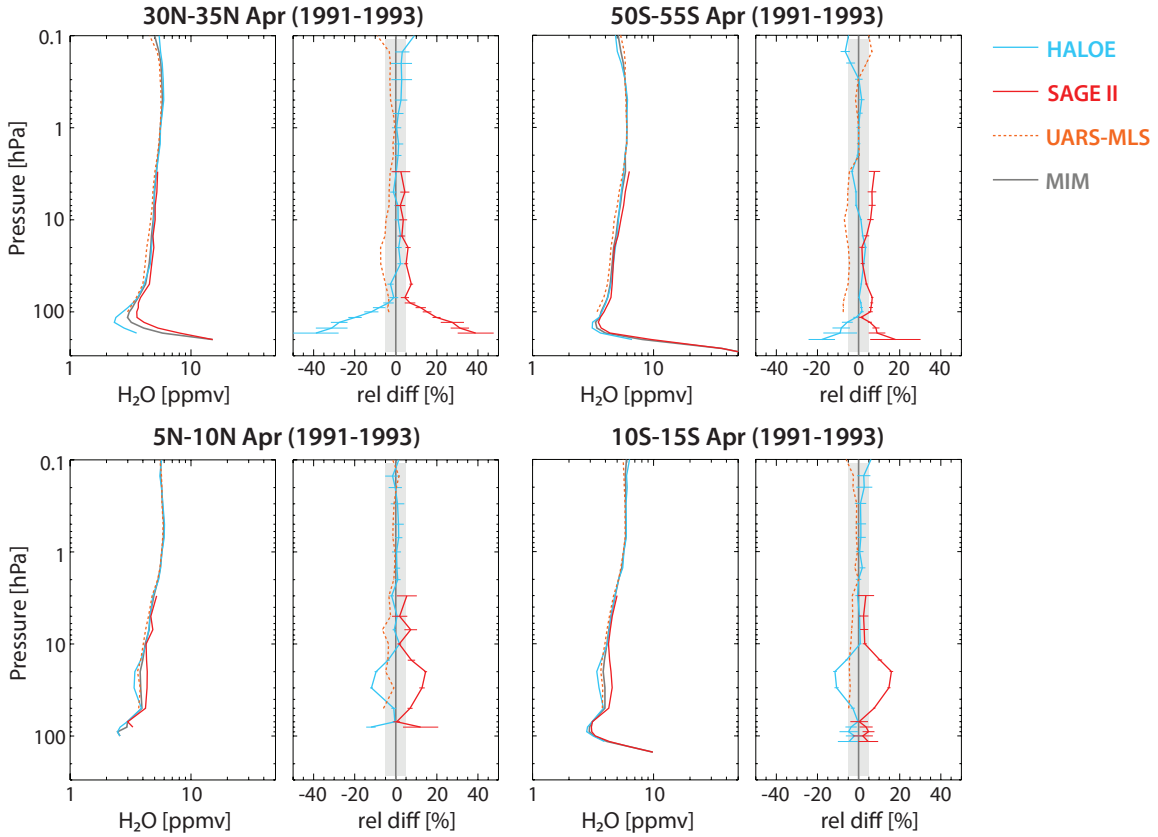


Figure 4.2.4: Vertical profiles of monthly zonal mean H_2O for 1991-1993. *The H_2O profiles are shown for 30°N-35°N and 50°S-55°S (upper panels), and 5°N-10°N and 10°S-15°S (lower panels) for April. The relative differences between the individual instruments (SAGE II, HALOE, and UARS-MLS) and the MIM 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%.*

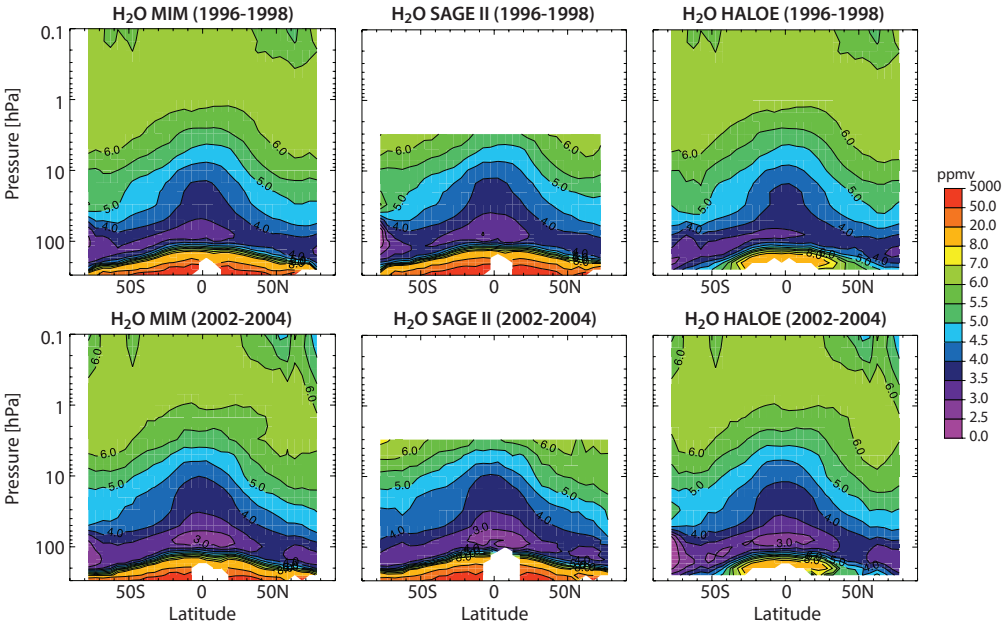


Figure 4.2.5a: Cross sections of annual zonal mean H_2O for 1996-1998 and 2002-2004. *Shown from left to right are the MIM, SAGE II, and HALOE. The upper (lower) panels show the climatologies for the earlier (later) time period.*

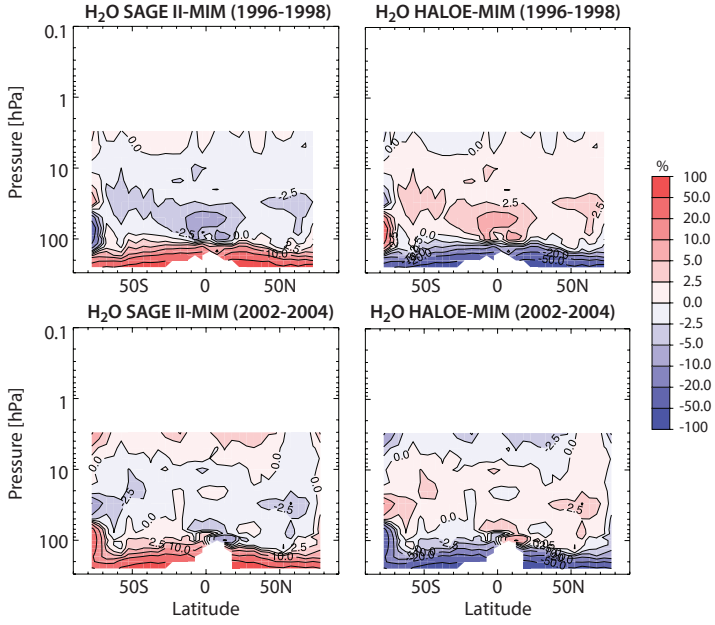


Figure 4.2.5b: Cross sections of annual zonal mean H_2O differences for 1996-1998 and 2002-2004. *Shown are the relative differences for each instrument (SAGE II and HALOE) with respect to the MIM. Same ordering as in Figure 4.2.5a.*

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.

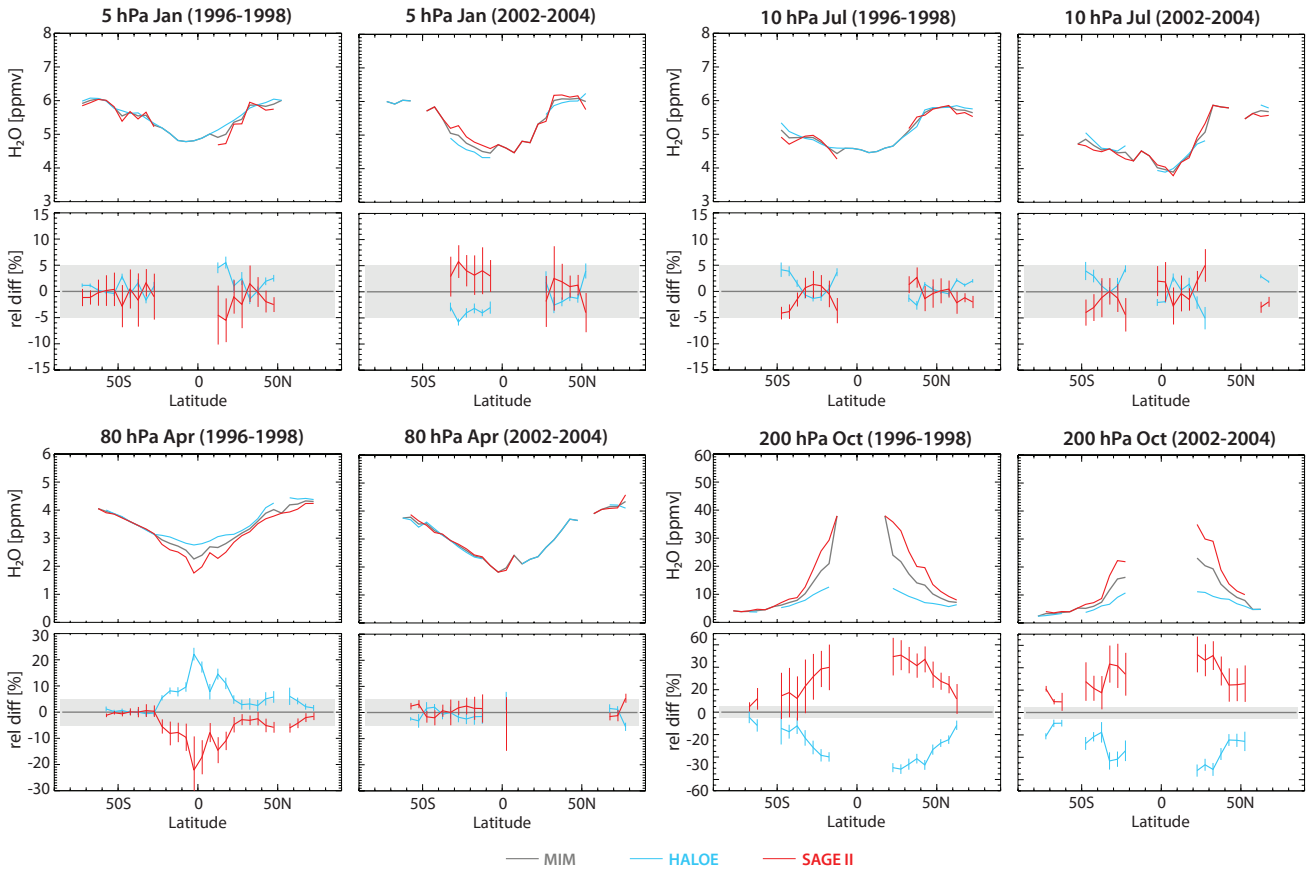


Figure 4.2.6: Meridional profiles of monthly zonal mean H_2O for 1996-1998 versus 2002-2004. *Meridional profiles are shown at 5 hPa for January, 10 hPa for July, 80 hPa for April, and 200 hPa for October (from upper left to lower right) for the two time periods. Upper panels show absolute values, lower 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.*

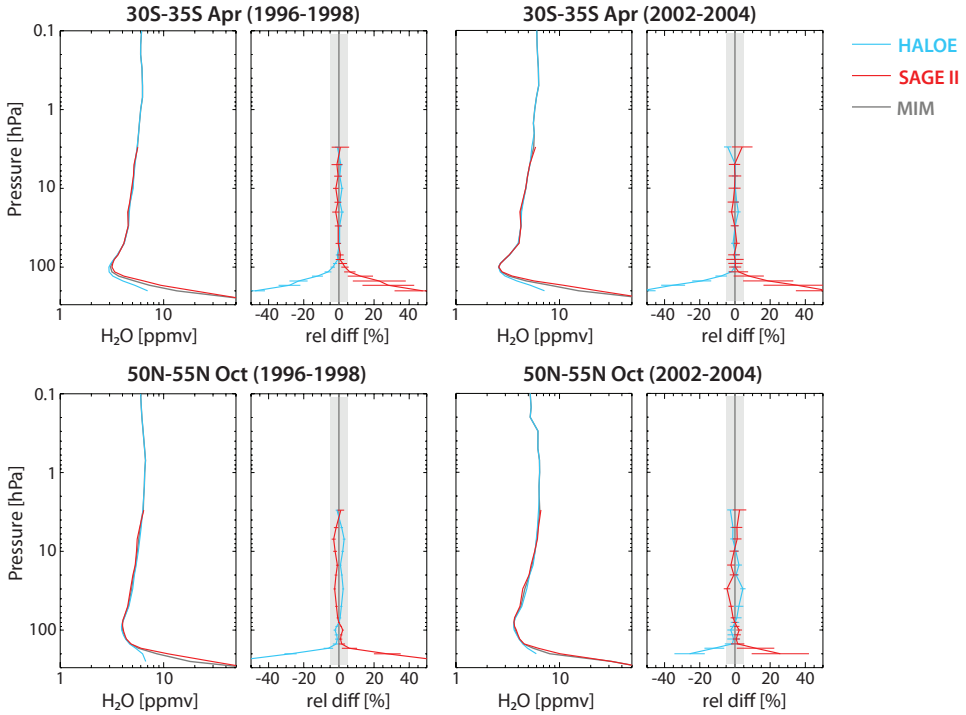


Figure 4.2.7: Vertical profiles of monthly zonal mean H_2O for 1996-1998 versus 2002-2004. *The H_2O profiles and their relative differences from the MIM are shown for April 30°S-35°S and October 50°N-55°N for the two time periods, respectively. Error bars indicate the uncertainty in the relative differences 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*.

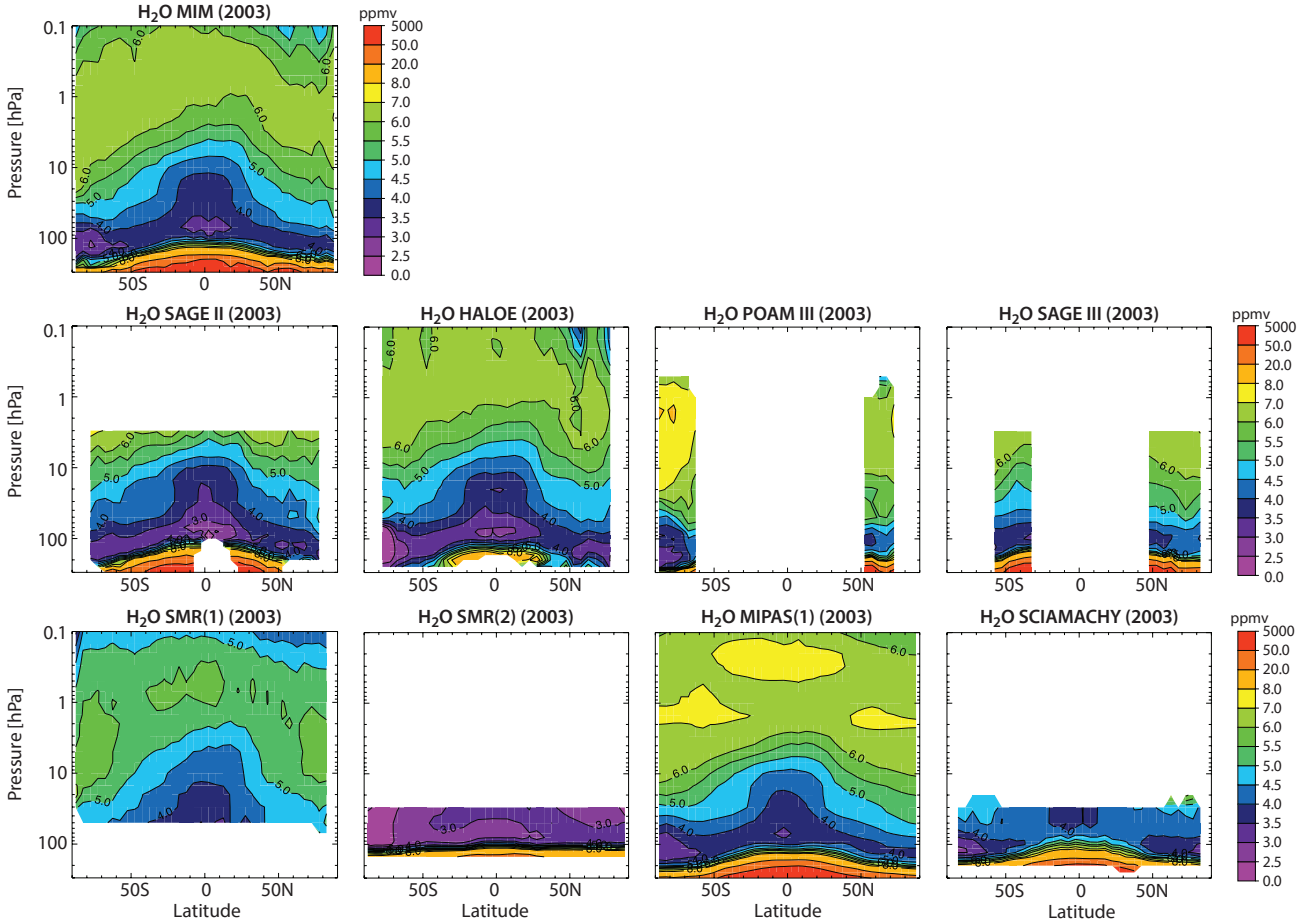


Figure 4.2.8a: Cross sections of annual zonal mean H_2O for 2003. *Shown are the annual mean cross sections for the MIM (upper row), SAGE II, HALOE, POAM III, SAGE III (middle row), and SMR(1), SMR(2), MIPAS(1), and SCIAMACHY (lower row). Note that SMR(2) is not included in the MIM.*

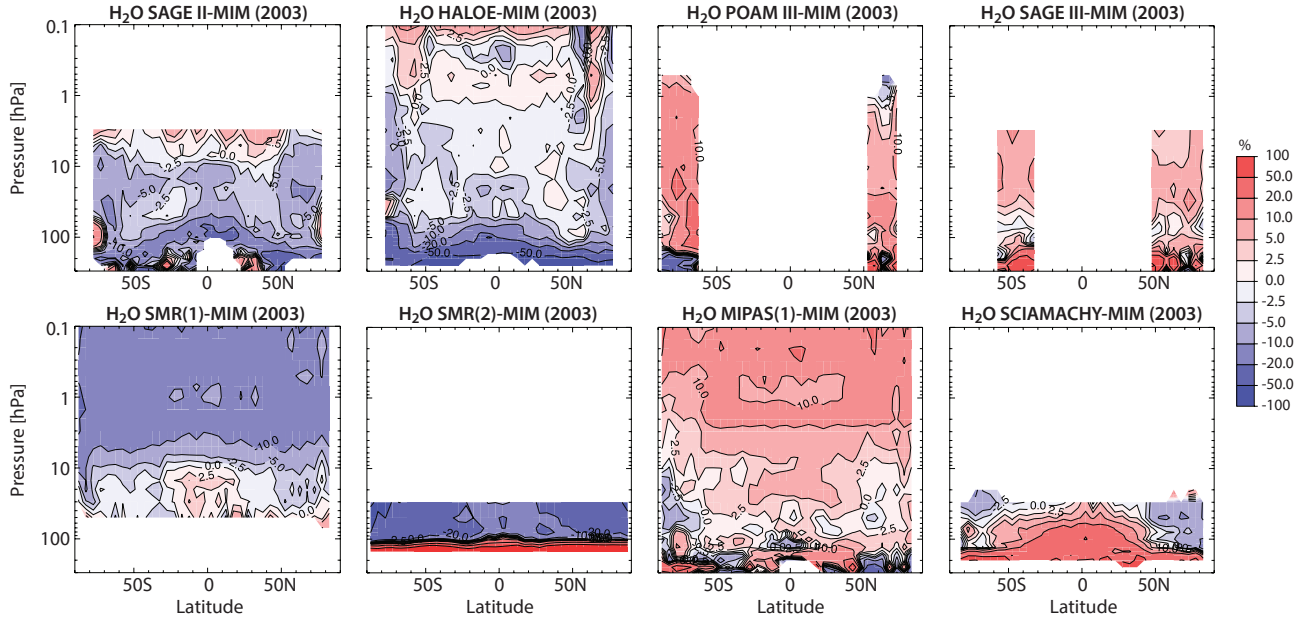


Figure 4.2.8b: Cross sections of annual zonal mean H_2O differences for 2003. *Relative differences with respect to the MIM are shown for the individual instruments shown in Figure 4.2.8a. Note that SMR(2) is not included in the MIM.*

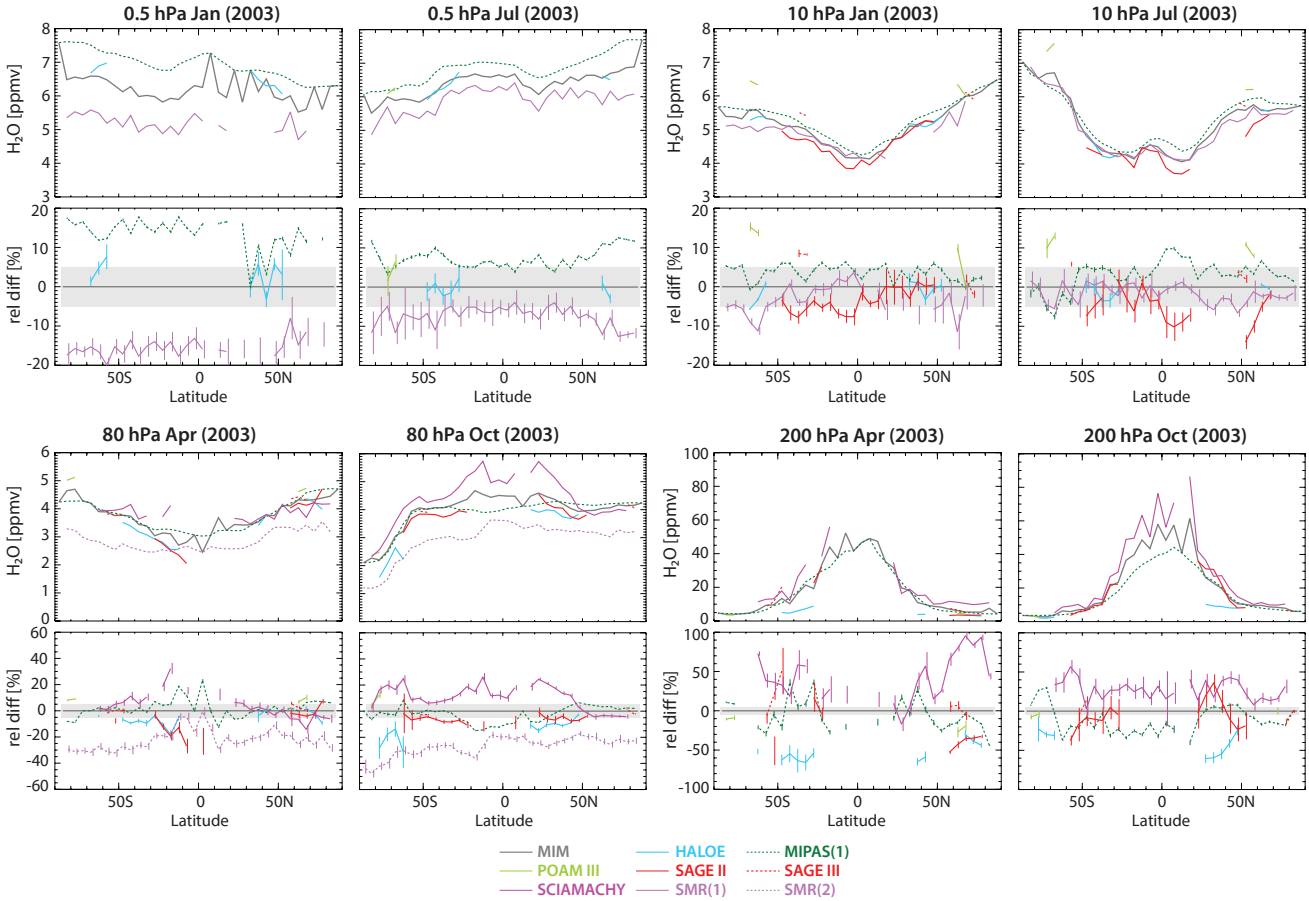


Figure 4.2.9: Meridional profiles of monthly zonal mean H_2O for 2003. *Meridional profiles are shown at 0.5 and 10 hPa for January and July (upper row), and at 80 and 200 hPa for April and October (lower row). Upper panels show absolute values, and lower panels show relative differences between the 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.*

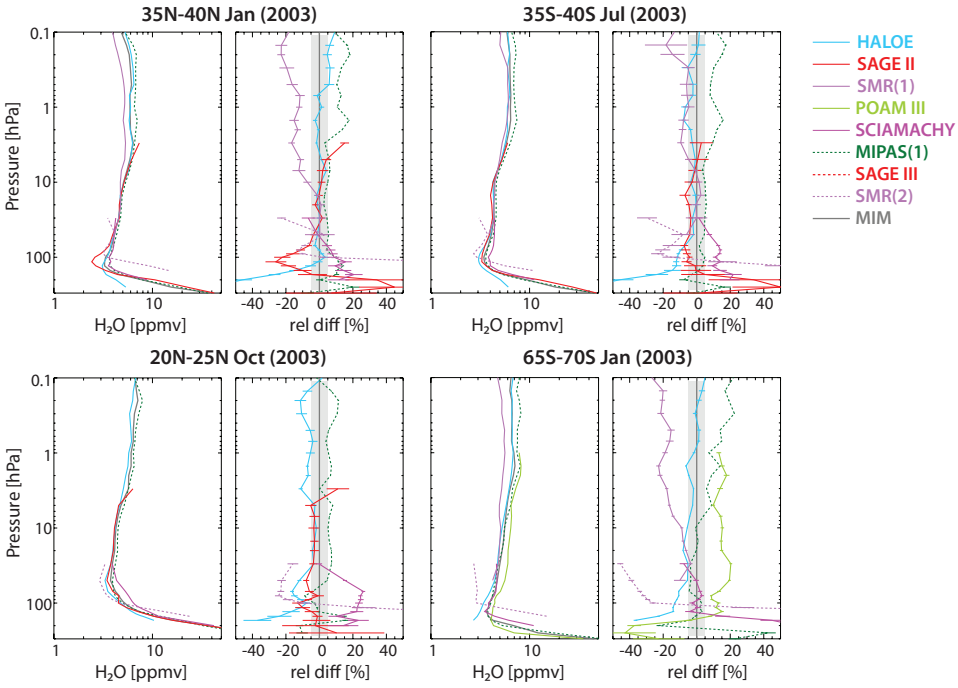


Figure 4.2.10: Vertical profiles of monthly zonal mean H_2O for 2003. *The H_2O profiles and their differences relative to the MIM are shown for January 35°N-40°N, July 35°S-40°S, January 65°S-70°S, and October 20°N-25°N. Error bars indicate the uncertainty in the relative differences based on the SEM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

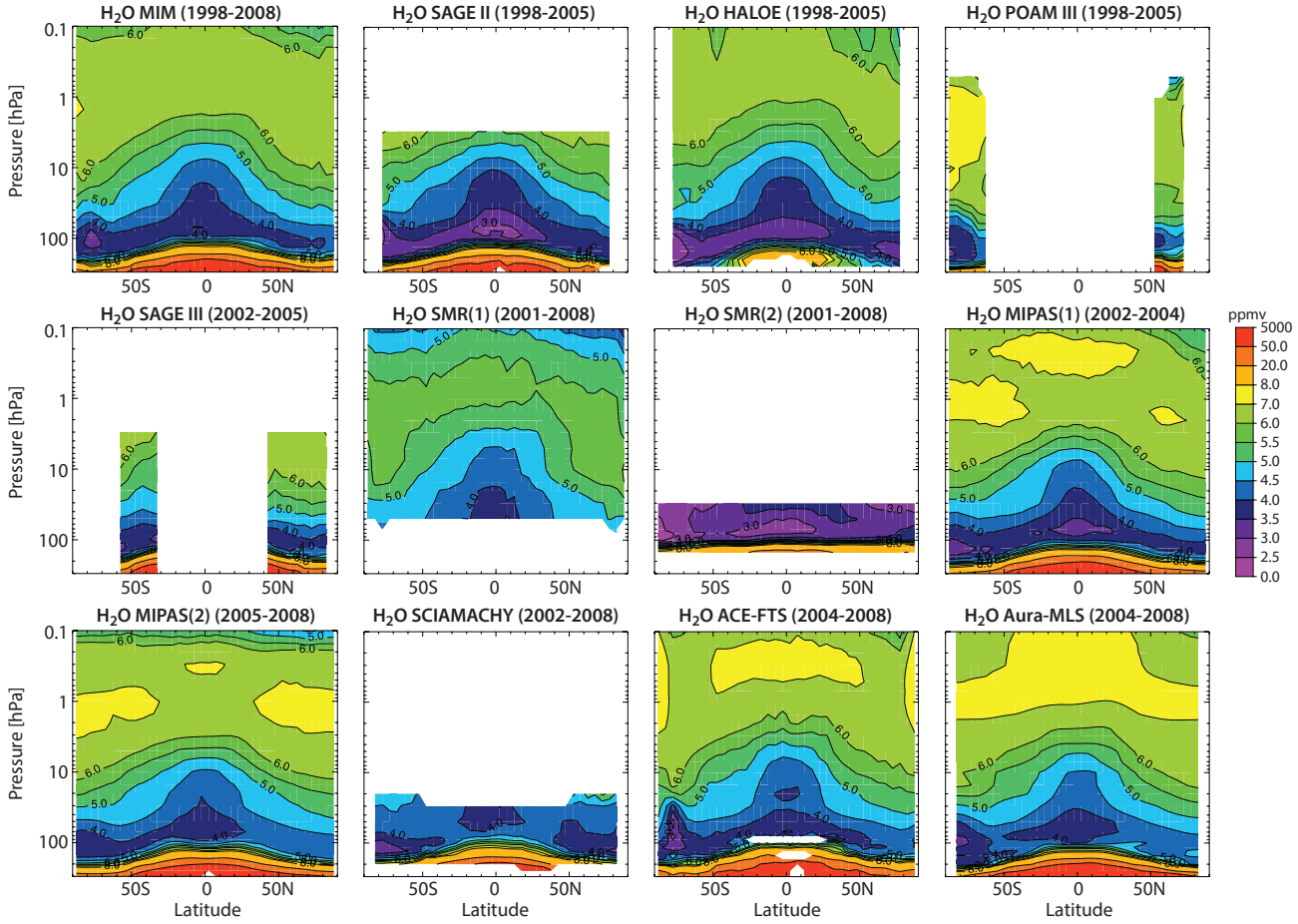


Figure 4.2.11a: Cross sections of annual zonal mean H_2O for 1998-2008. *Shown are the MIM, SAGE II, HALOE, POAM III (upper row), SAGE III, SMR(1), SMR(2), MIPAS(1) (middle row), and MIPAS(2), SCIAMACHY, ACE-FTS, and Aura-MLS (lower row). Note, SMR(2) and MIPAS(1) are not included in the MIM.*

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

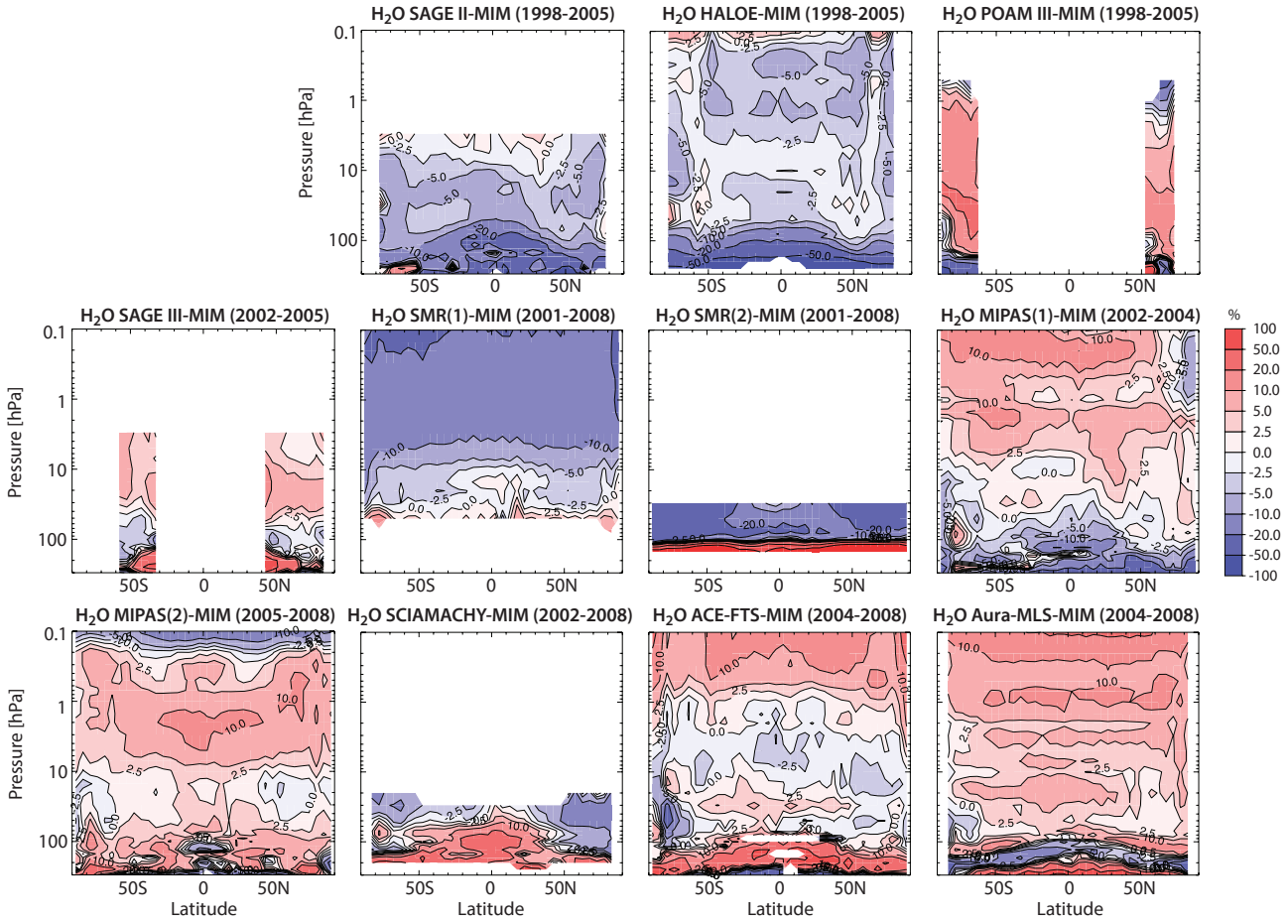


Figure 4.2.11b: Cross sections of monthly zonal mean H_2O differences for 1998-2008. *Differences relative to the MIM are shown for the individual instruments shown in Figure 4.2.11a. Note that SMR(2) and MIPAS(1) are not included in the MIM.*

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%.

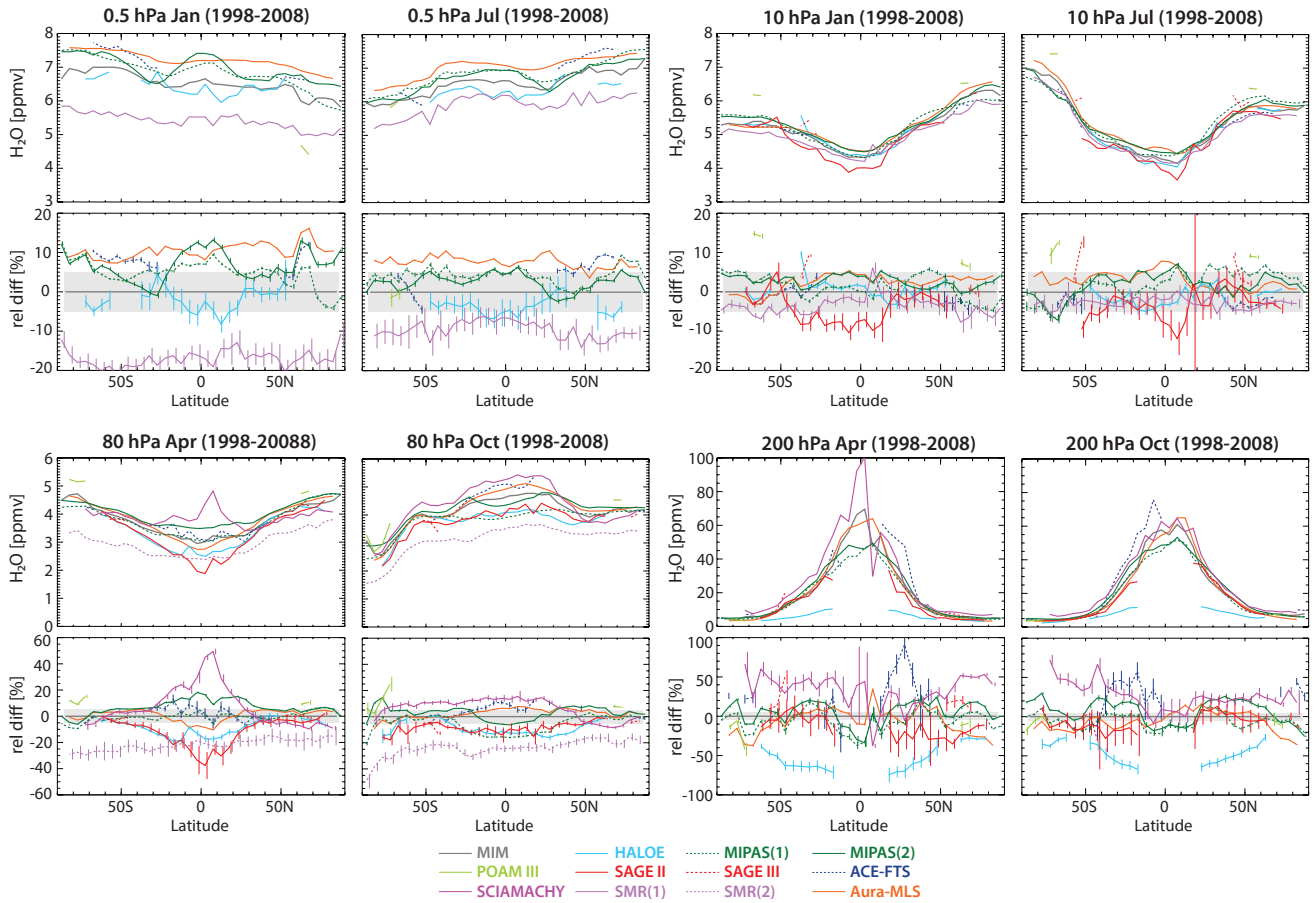


Figure 4.2.12: Meridional profiles of monthly zonal mean H_2O for 1998-2008. *Meridional profiles are shown at 0.5 and 10 hPa for January and July (upper row), and at 80 and 200 hPa for April and October (lower row). Upper panels show absolute values, lower panels relative differences between the individual 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

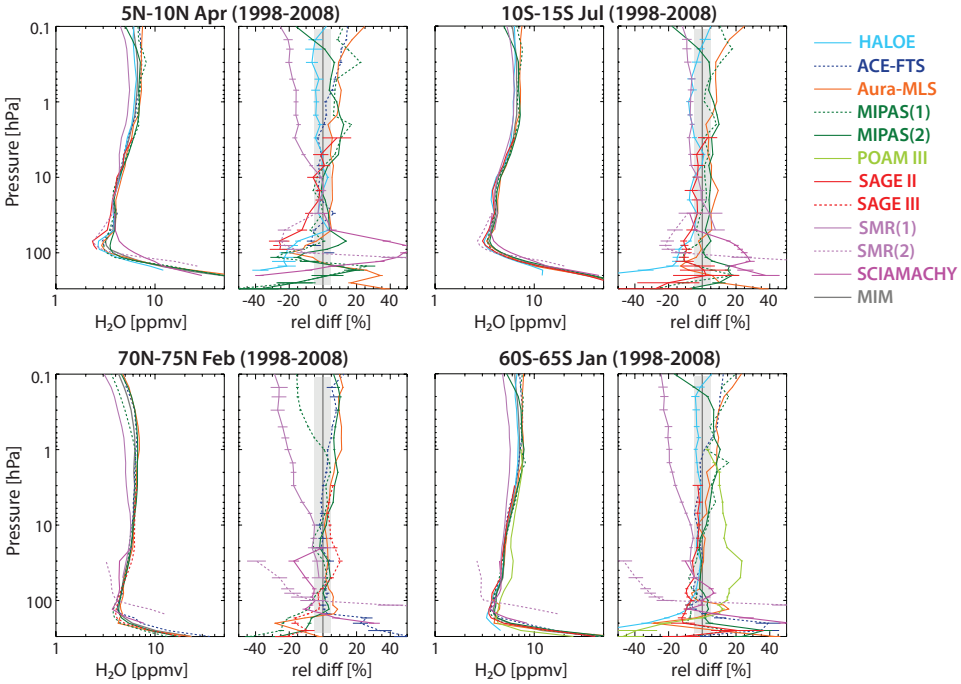


Figure 4.2.13: Vertical profiles of monthly zonal mean H_2O for 1998-2008. *The H_2O profiles and their relative differences to the MIM are shown for April 5°N-10°N, July 10°S-15°S, February 70°N-75°N, and January 60°S-65°S. Error bars indicate the uncertainties in the relative differences based on the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

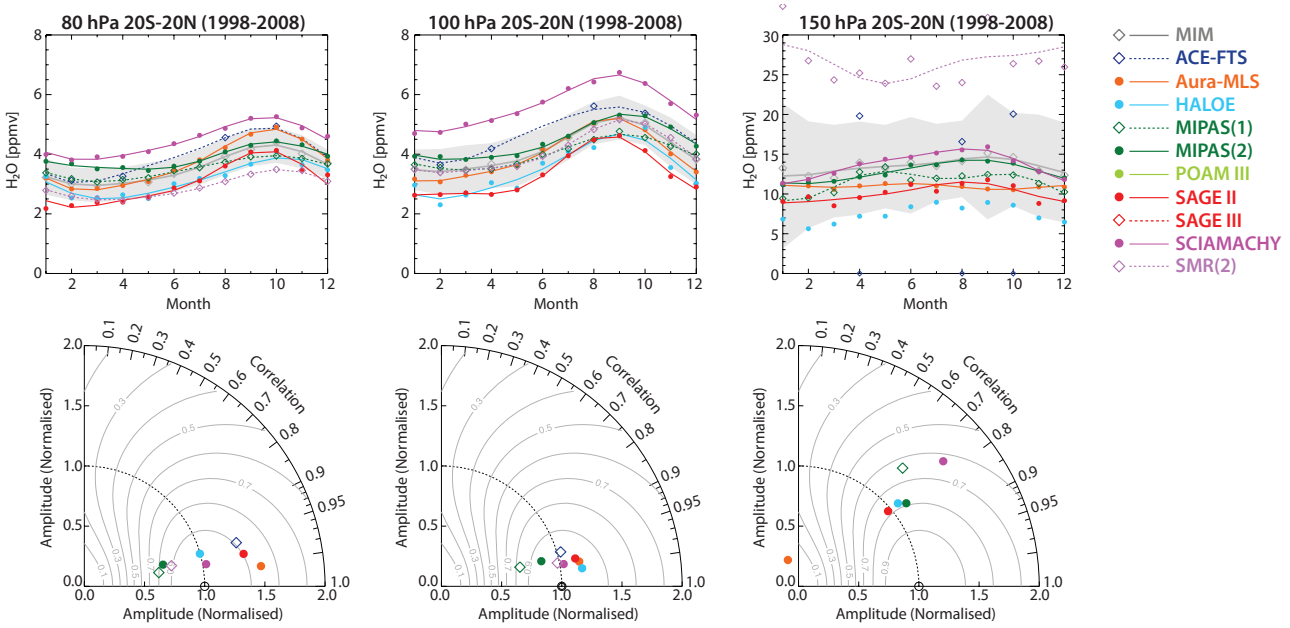


Figure 4.2.14: Seasonal cycles of H_2O in the tropics for 1998-2008. *Seasonal cycles and corresponding Taylor diagrams of monthly zonal mean H_2O averaged over 20°S-20°N are shown at 80 (left column), 100 hPa (middle column) and 150 hPa (right column). Coloured lines represent fits including an annual 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.*

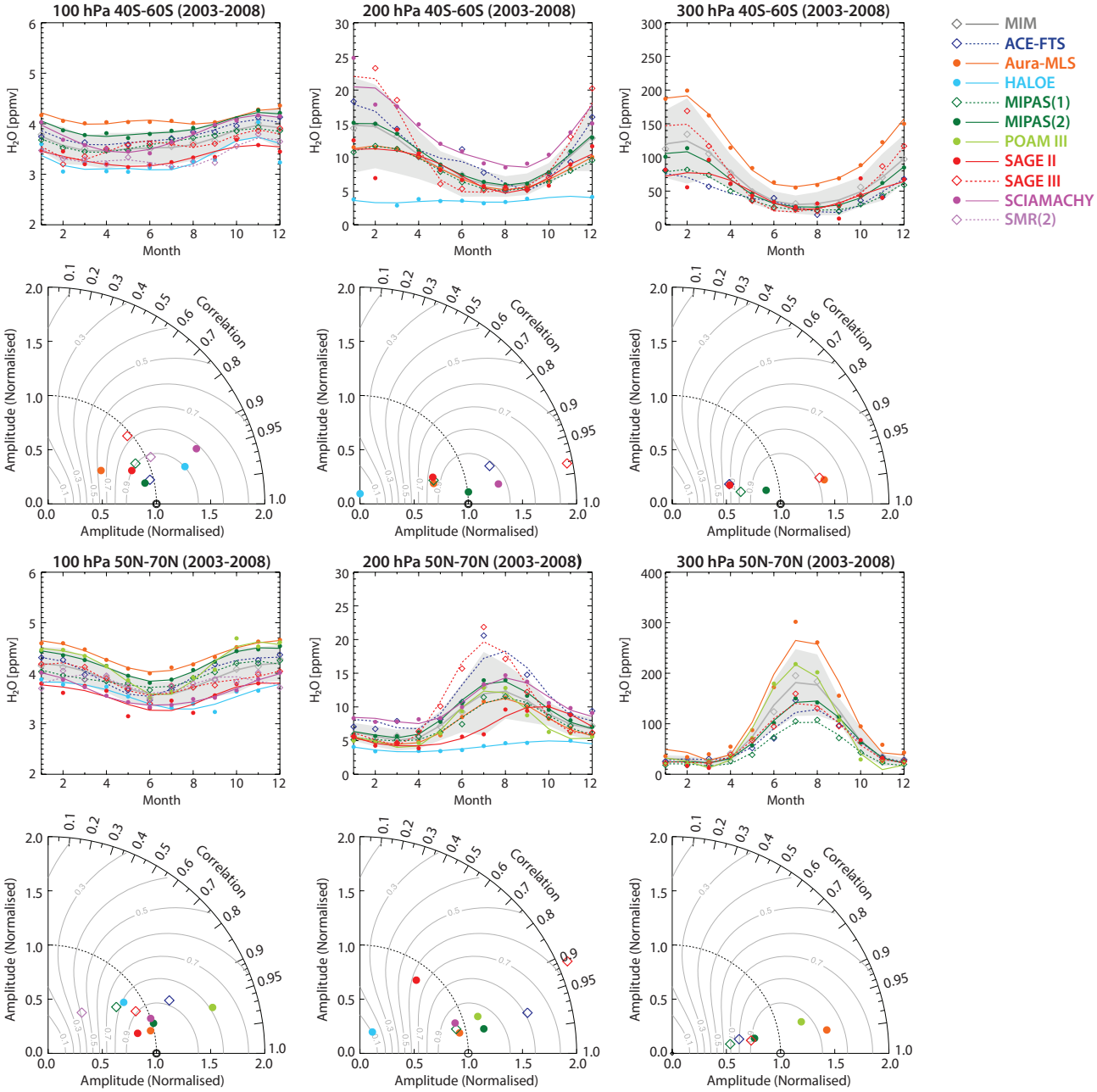


Figure 4.2.15: Seasonal cycles of H_2O in the SH and NH mid-latitudes for 1998-2008. *Seasonal cycles and corresponding Taylor diagrams of monthly zonal mean H_2O averaged over 40°S-60°S (upper two rows) and 50°N-70°N (lower two rows) are shown at 100, 200, and 300 hPa (from left to right). Coloured 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.

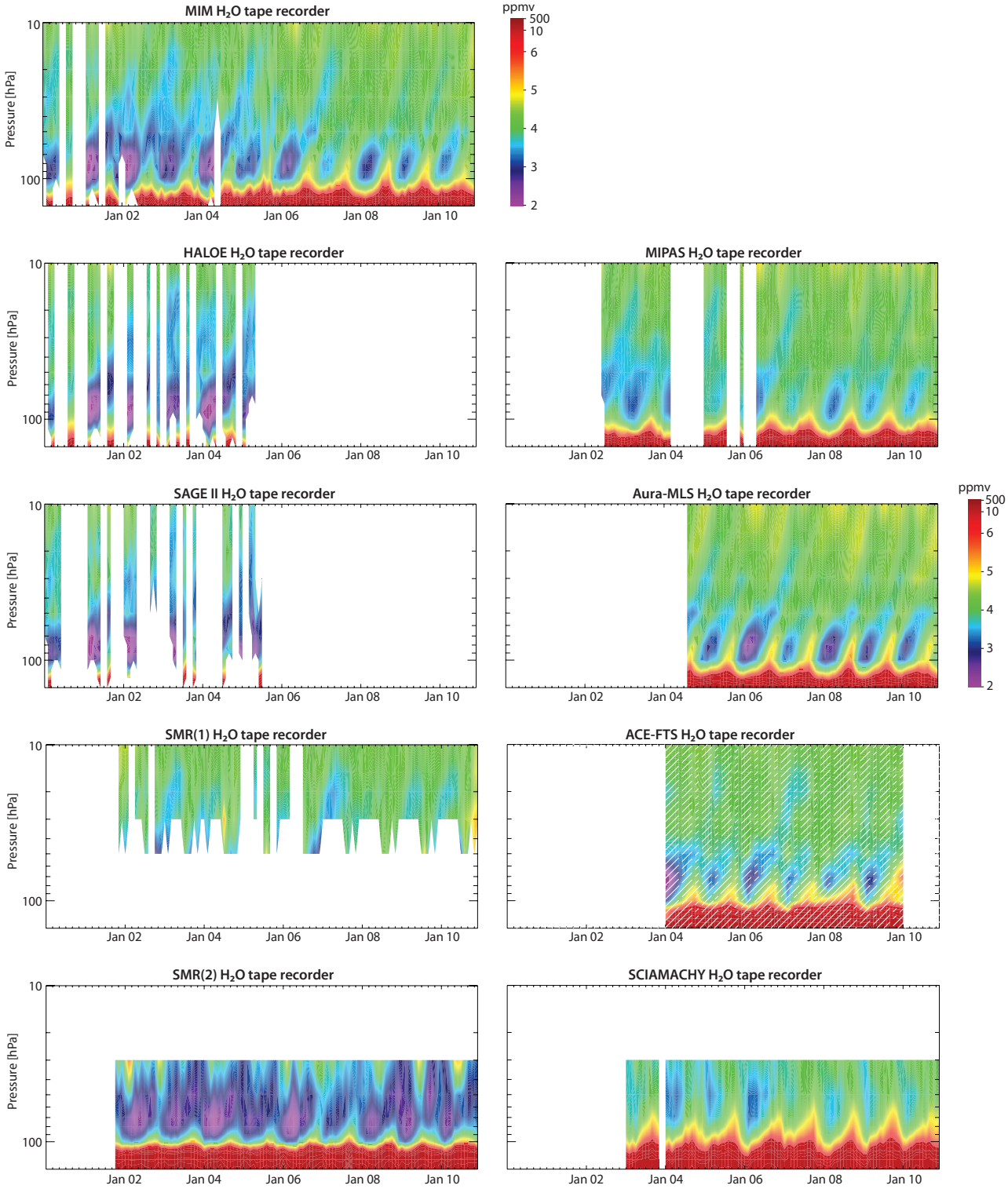


Figure 4.2.16a: H_2O tape recorder. *Shown is the altitude-time evolution of H_2O averaged over 15°S-15°N for the time period 2000-2010. The very limited tropical ACE-FTS data were interpolated in time and altitude; white hatching indicates regions that do not contain data. Note that the SMR(2) and SCIAMACHY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

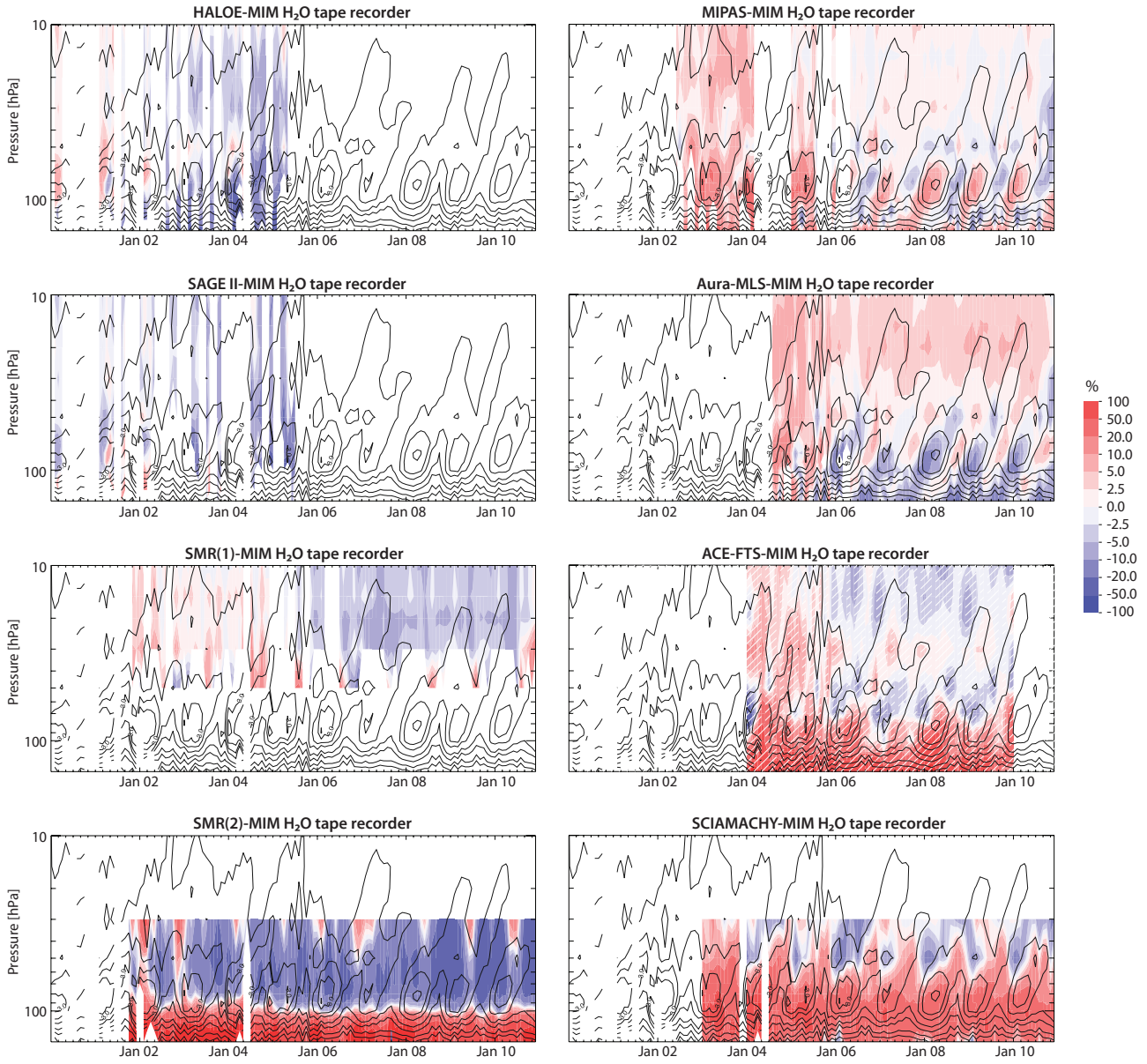
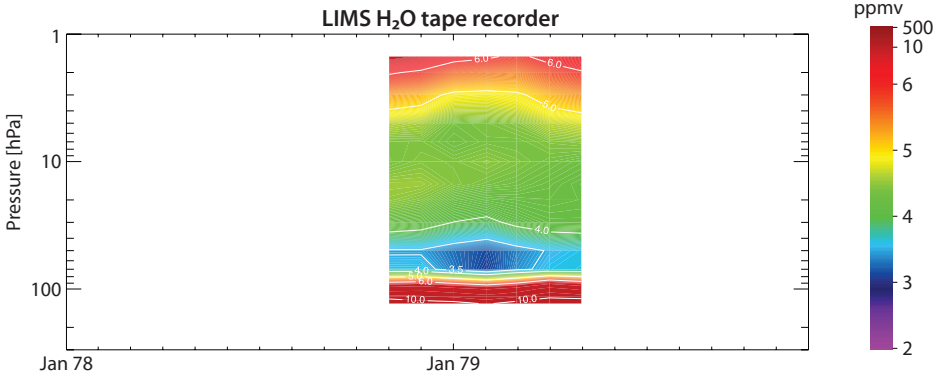


Figure 4.2.16b: Differences for H_2O tape recorder. *Altitude-time evolutions of H_2O differences relative to the MIM are shown for the time period 2000-2010 and each individual instrument (same ordering as in Figure 4.2.16a). Contour levels (2.5, 3, 3.5, 4, 5, 6, 8, 10, 50, 100 ppmv, with the 3-ppmv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



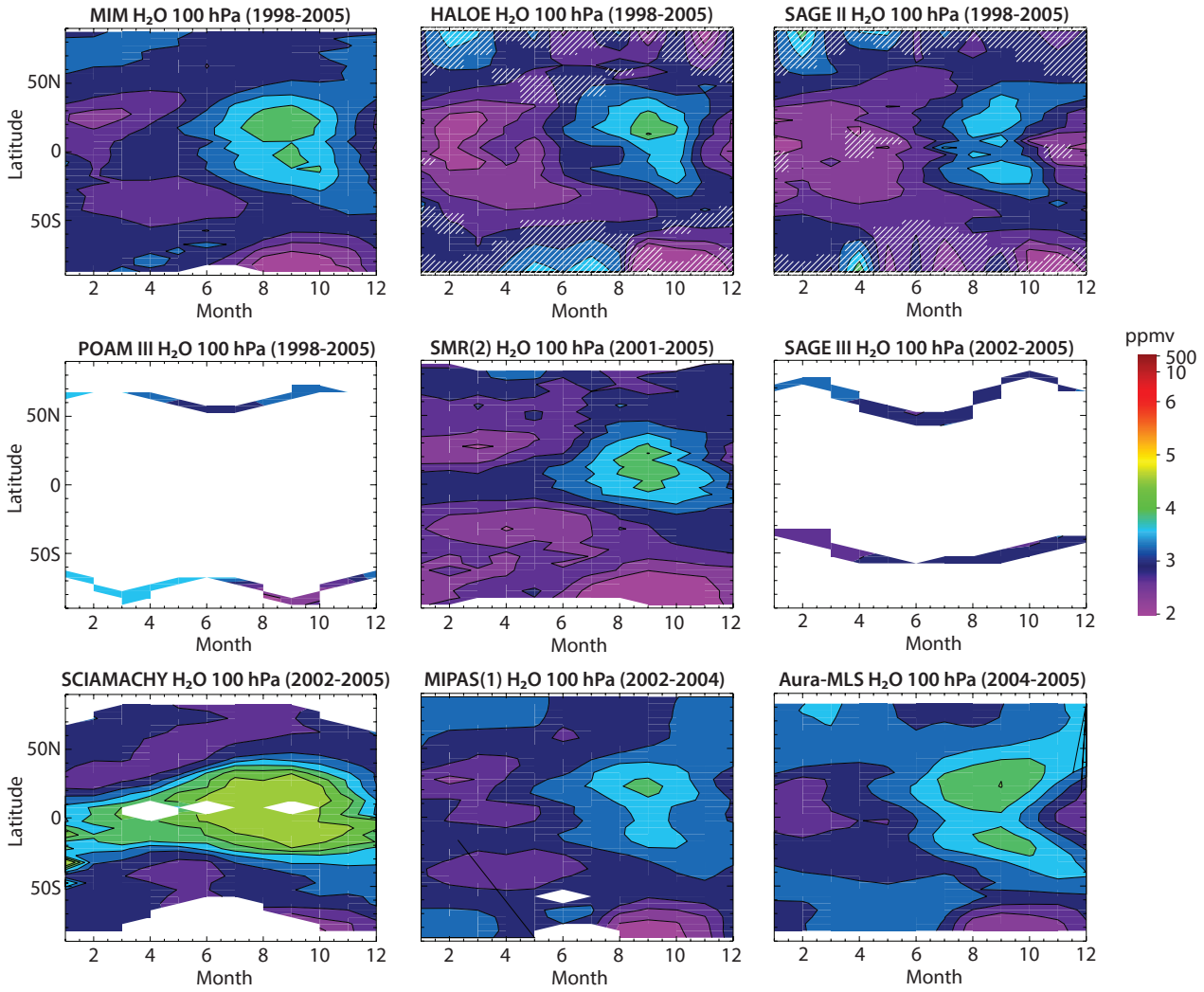


Figure 4.2.18: The horizontal tape recorder during 1998-2005. *Shown is a latitude-time evolution of H_2O at 100 hPa averaged over this period (or periods within this timeframe as indicated in the panel headers). HALOE and SAGE II show interpolated data; white hatching indicates the areas where no data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

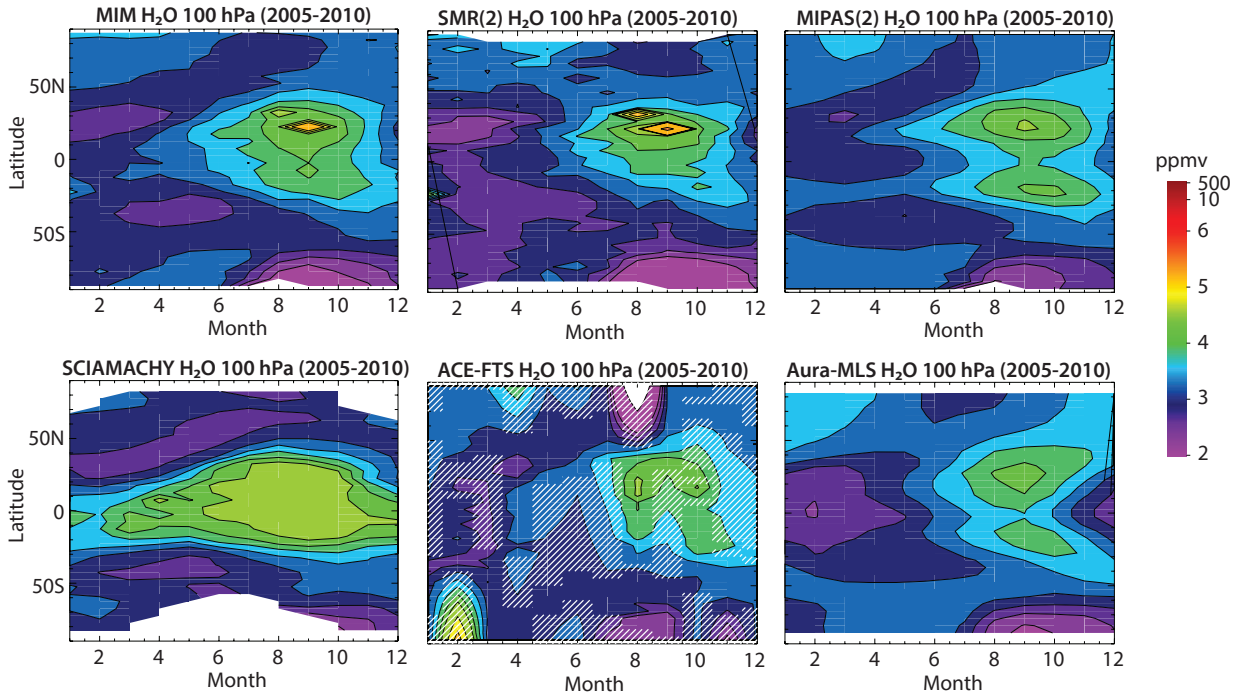


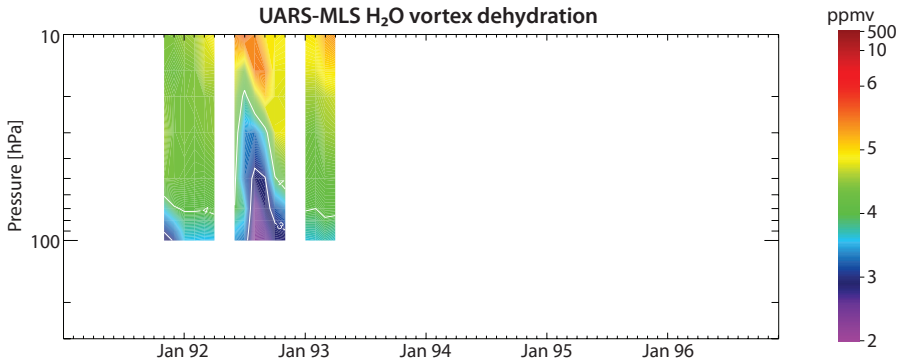
Figure 4.2.19: The horizontal tape recorder during 2005-2010. *Shown is a latitude-time evolution of H_2O at 100 hPa averaged over this period. ACE-FTS shows interpolated data; white hatching indicates the areas where no data was available. Note the differences in Aura-MLS, SMR(2), and SCIAMACHY when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



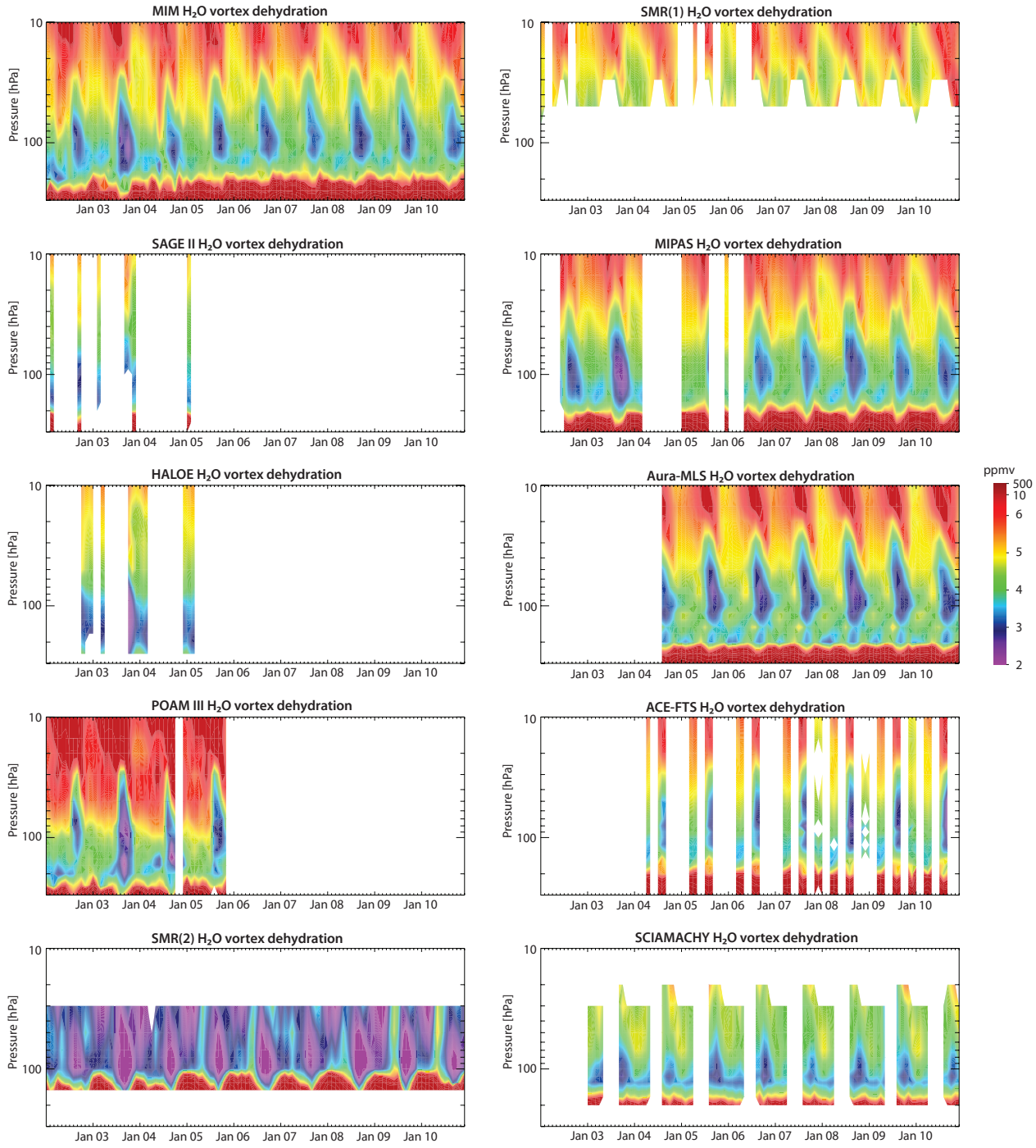


Figure 4.2.21a: Polar vortex dehydration. *The altitude-time evolution of Antarctic polar vortex descent and dehydration between 2002 and 2010 is shown for individual instruments and the MIM (uppermost left panel) using H_2O averaged over 60°S to 90°S. Note that SMR(2) is not included in the MIM.*

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

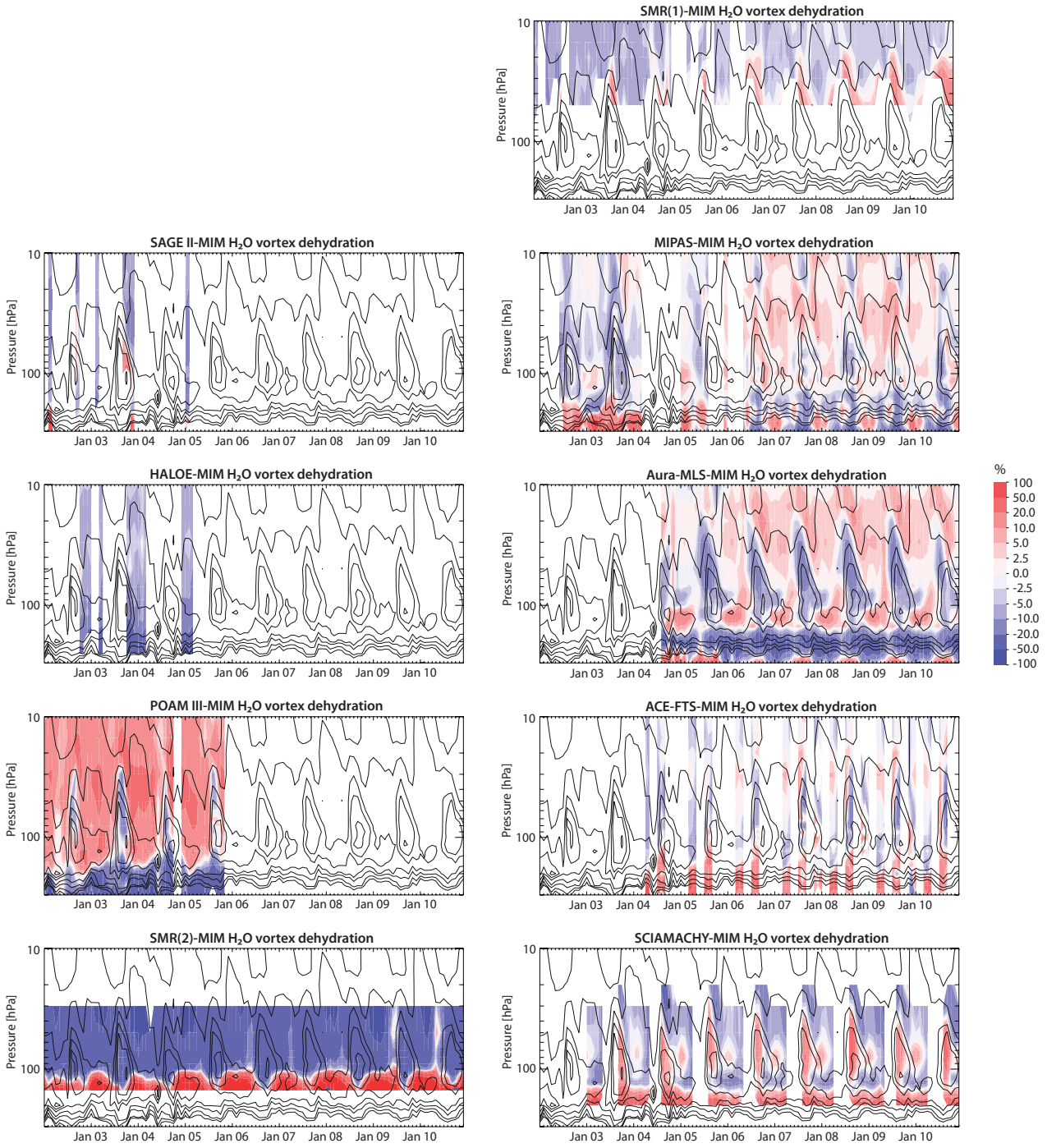


Figure 4.2.21b: Differences for polar vortex dehydration. *The time-altitude evolution of H_2O differences relative to the MIM between 2002 and 2010 is shown for each individual instrument (same ordering as in Figure 4.2.21a). Contour levels (2.5, 3, 3.5, 4, 5, 6, 8, 10, 50, 100 ppmv, with the 3-ppmv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

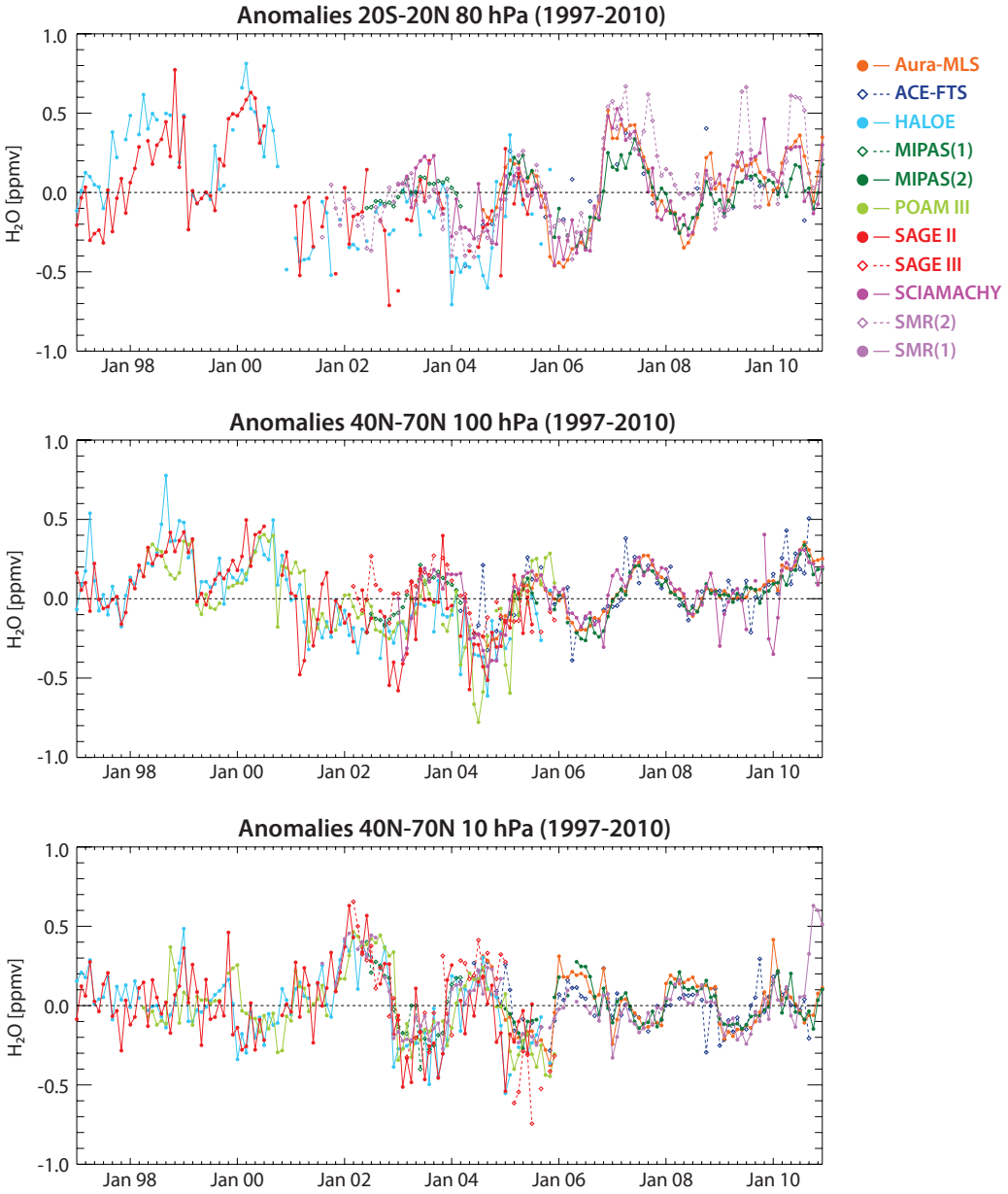


Figure 4.2.22: Time series of deseasonalised anomalies of H_2O for 1997-2010. *Time series of deseasonalised anomalies in H_2O at 80 hPa between 20°S and 20°N (upper panel), and at 100 hPa (middle panel) and 10 hPa (lower panel) between 40°N and 70°N, respectively.*

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

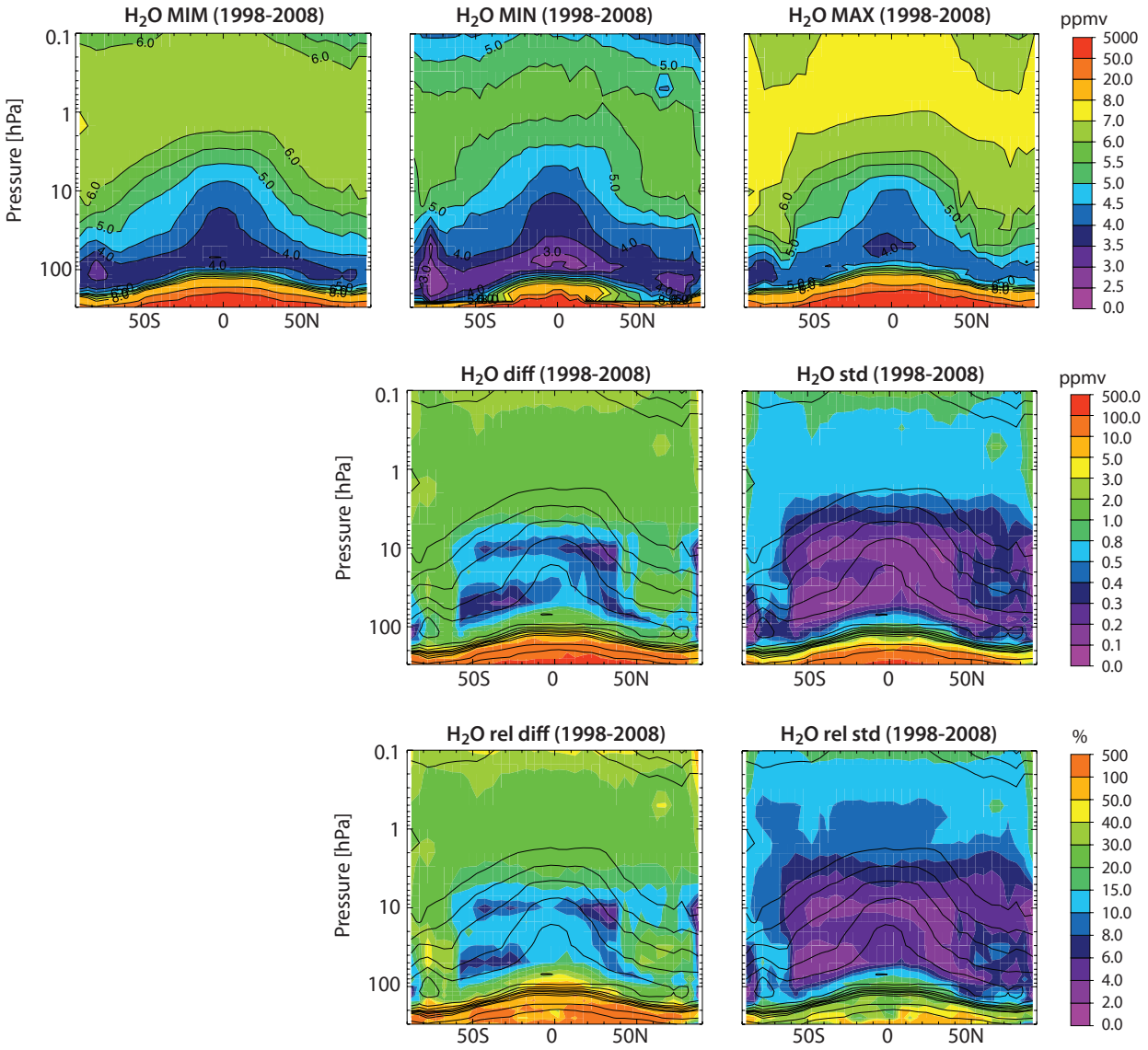


Figure 4.2.23: Summary of H_2O annual zonal mean state for 1998-2008. *Shown are the annual zonal mean cross sections of the MIM, minimum (MIN) and maximum (MAX) H_2O values (upper row), the absolute differences (MAX-MIN) and absolute standard deviations (middle row), and relative differences and relative 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

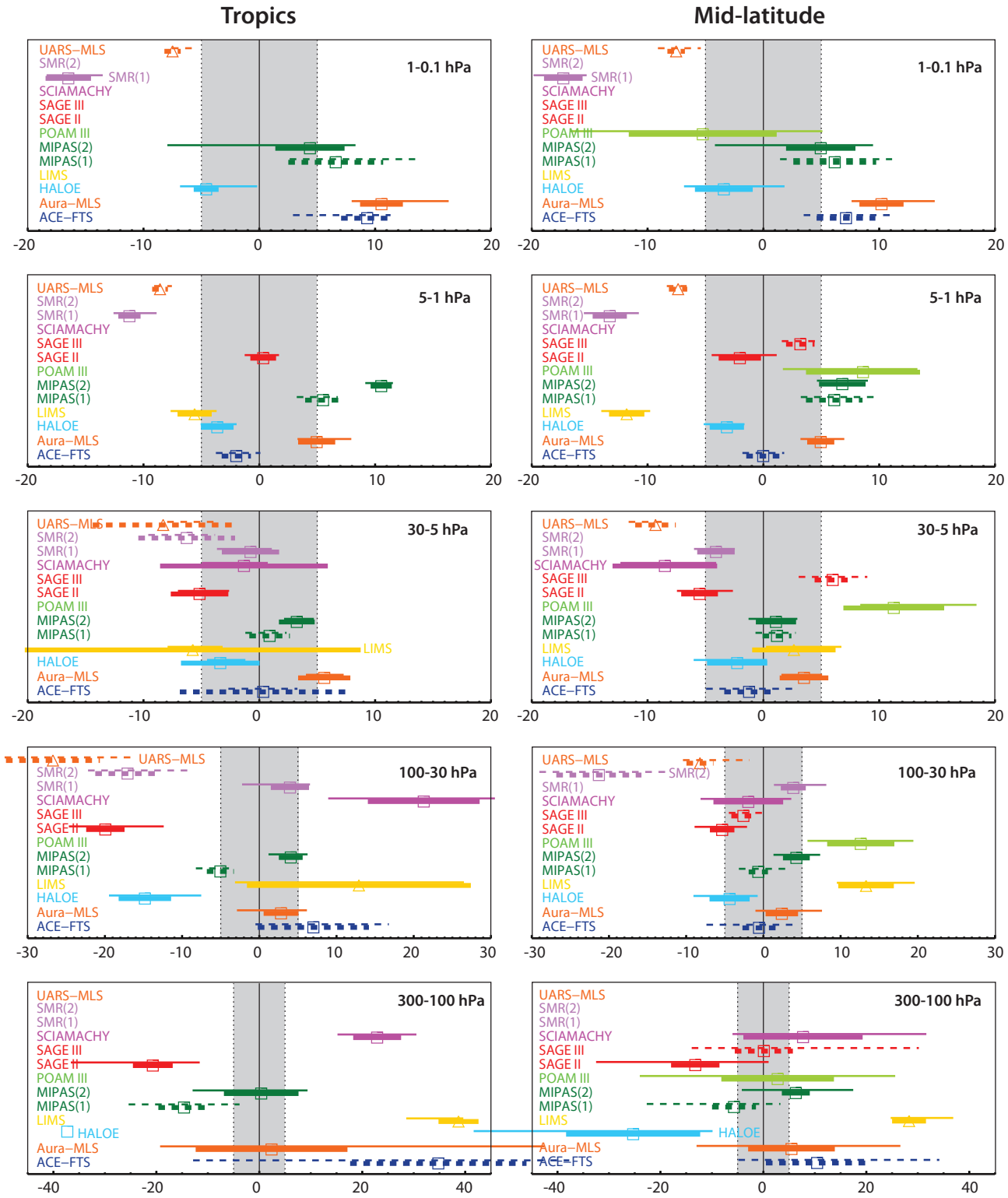


Figure 4.2.24: Summary of inter-instrument differences in H_2O for 1998-2008. *Results are calculated for the tropics 20°S-20°N (left) and extra-tropics 40°S-80°S and 40°N-80°N (right) and for 5 different altitude regions from the UT up to the LM between 300 and 0.1 hPa as defined in Table 0.1. Shown 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.*

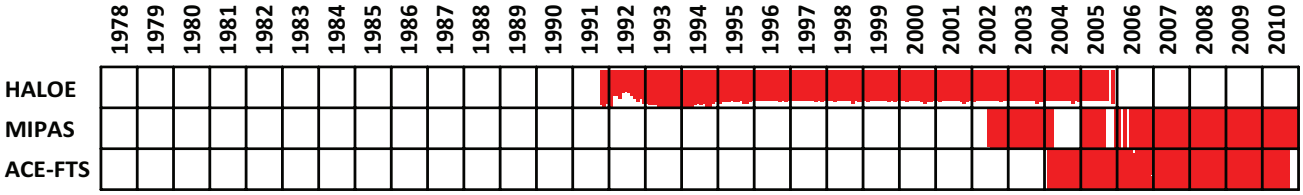


Table 4.3.2: Time period, vertical range, vertical resolution, references and other comments for CH_4 measurements.

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.

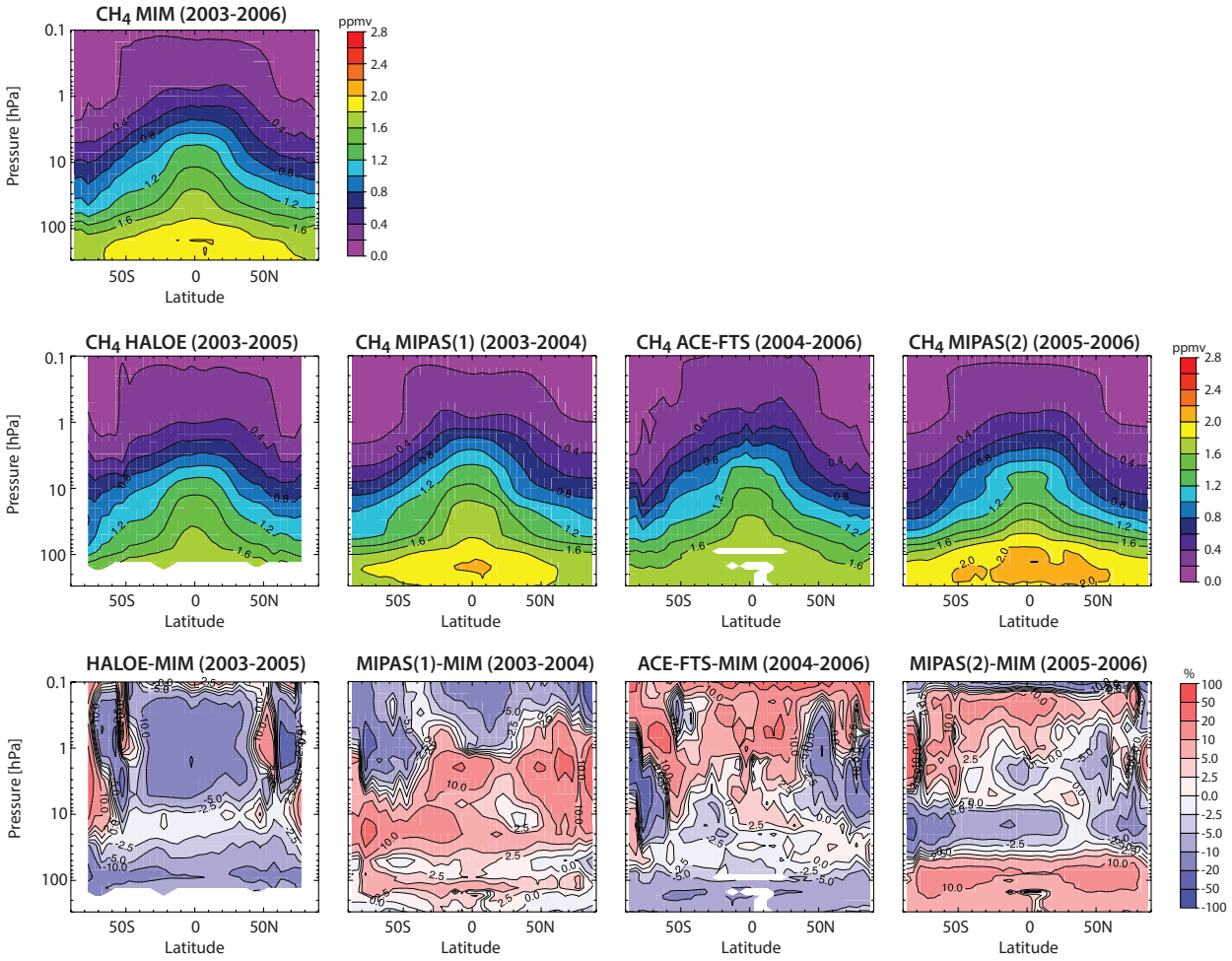


Figure 4.3.1: Cross sections of annual zonal mean CH_4 for 2003-2006. *Upper panel shows the CH_4 cross section for the MIM, middle panels show cross sections for the different instruments (HALOE, MIPAS(1), ACE-FTS, and MIPAS(2)), and lower panels show the relative differences between each instrument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

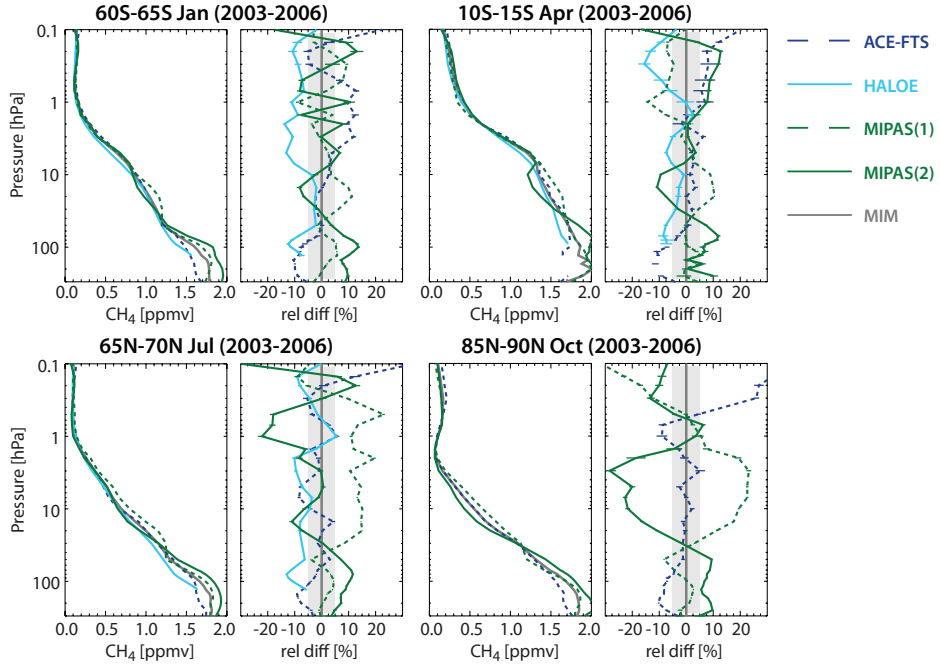


Figure 4.3.2: Vertical profiles of monthly zonal mean CH_4 for 2003-2006. *Vertical CH_4 profiles for 60°S-65°S January and 10°S-15°S April (upper panels) and for 65°N-70°N July and 85°N-90°N October (lower panels) are shown together with their differences from the MIM. HALOE, MIPAS(1), ACE-FTS, and 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

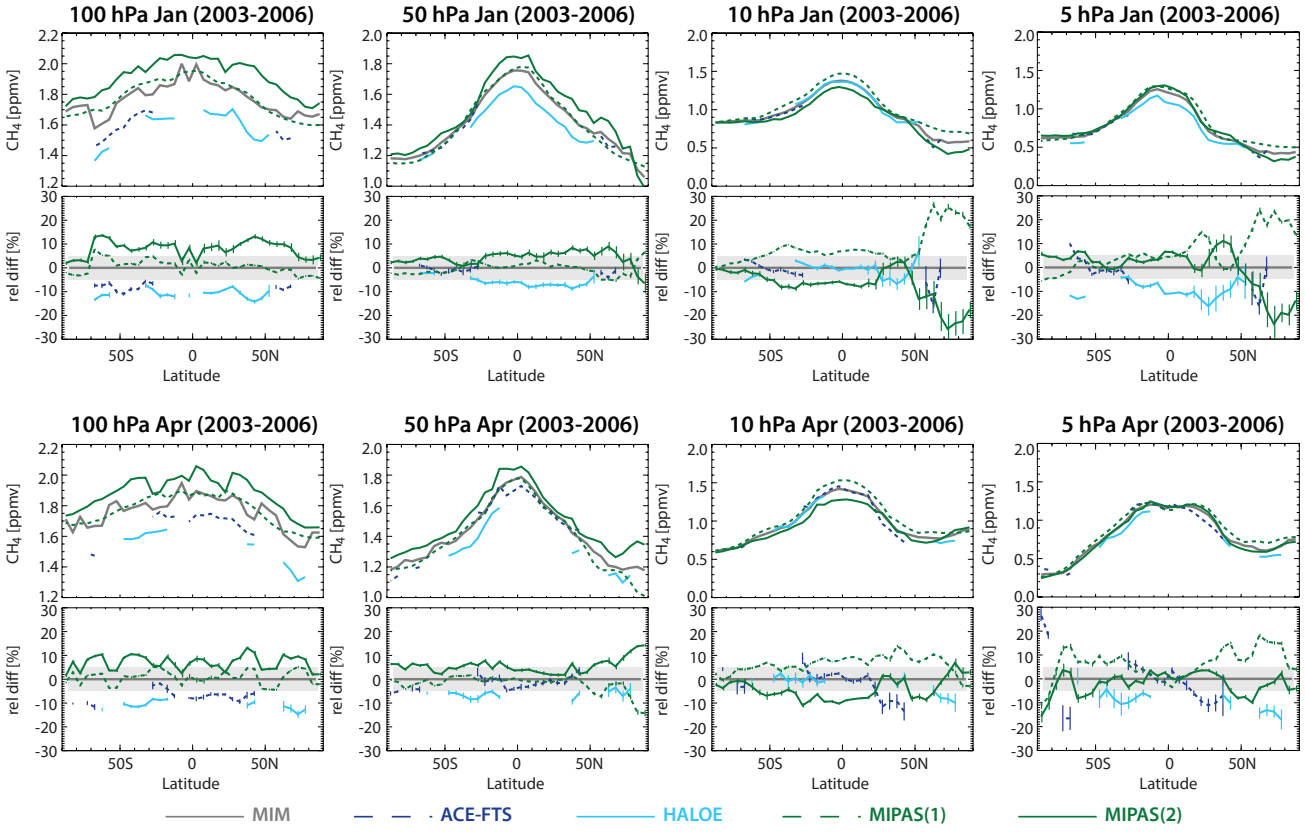


Figure 4.3.3: Meridional profiles of monthly zonal mean CH_4 for 2003-2006. *Meridional zonal mean CH_4 profiles for HALOE, MIPAS(1), ACE-FTS, and MIPAS (2) are shown at 100, 50, 10, and 5 hPa for January (upper row) and April (lower row), respectively. Differences between the individual instruments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

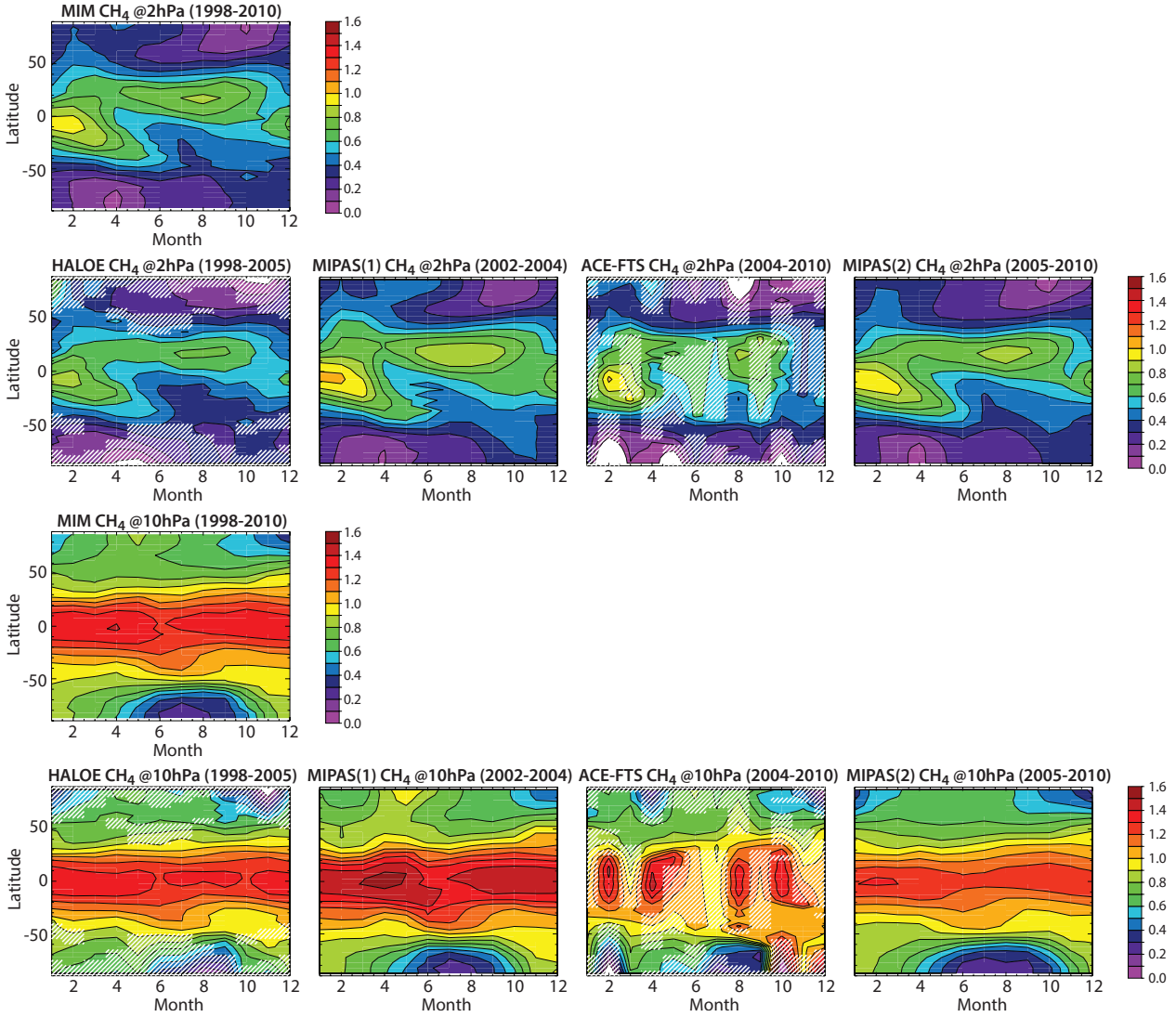


Figure 4.3.4a: Latitude–time evolution of CH_4. *The latitude-time evolution of montly zonal mean CH_4 at 2 hPa (top two rows), and 10 hPa (bottom two rows) are shown for the MIM (1998-2010), HALOE (1998-2005), MIPAS(1) (2002-2004), ACE-FTS (2004-2010), and MIPAS(2) (2005-2010) averaged over the time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

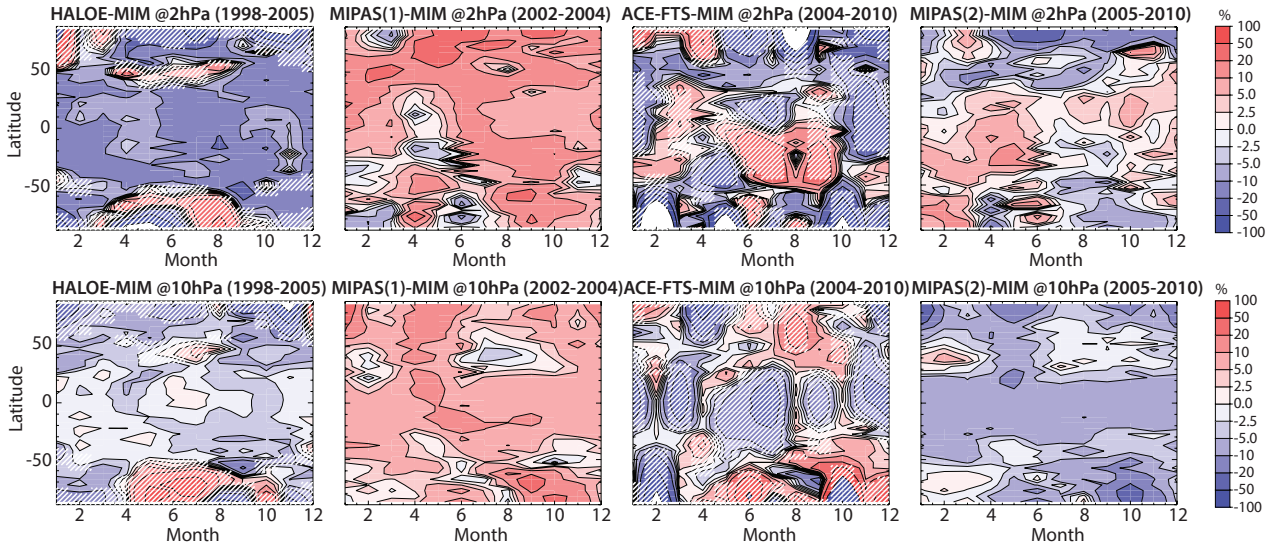


Figure 4.3.4b: Latitude–time evolution of differences in CH_4. *CH_4 differences with respect to the MIM at 2 hPa (top), and 10 hPa (bottom) are shown for HALOE, MIPAS(1), ACE-FTS, and MIPAS(2) over the time period as indicated in Figure 4.3.4a. For HALOE and ACE-FTS, hatched regions indicate where no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

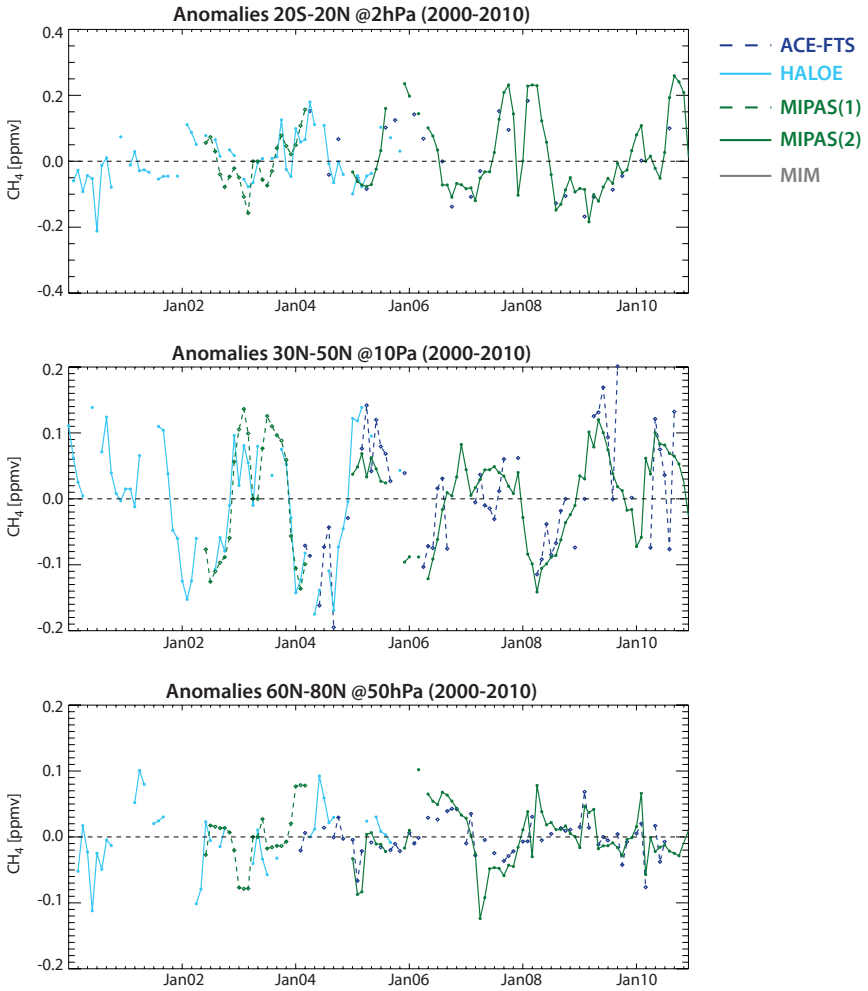


Figure 4.3.5: Time series of deseasonalised CH_4 anomalies between 2000 and 2010. *Deseasonalised CH_4 anomalies are shown for 2 hPa in the tropics (20°S-20°N; upper panel), 10 hPa at Northern mid-latitudes (30°N-50°N; middle panel), and 50 hPa at Northern high latitudes (60°N-80°N; lower panel).*

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

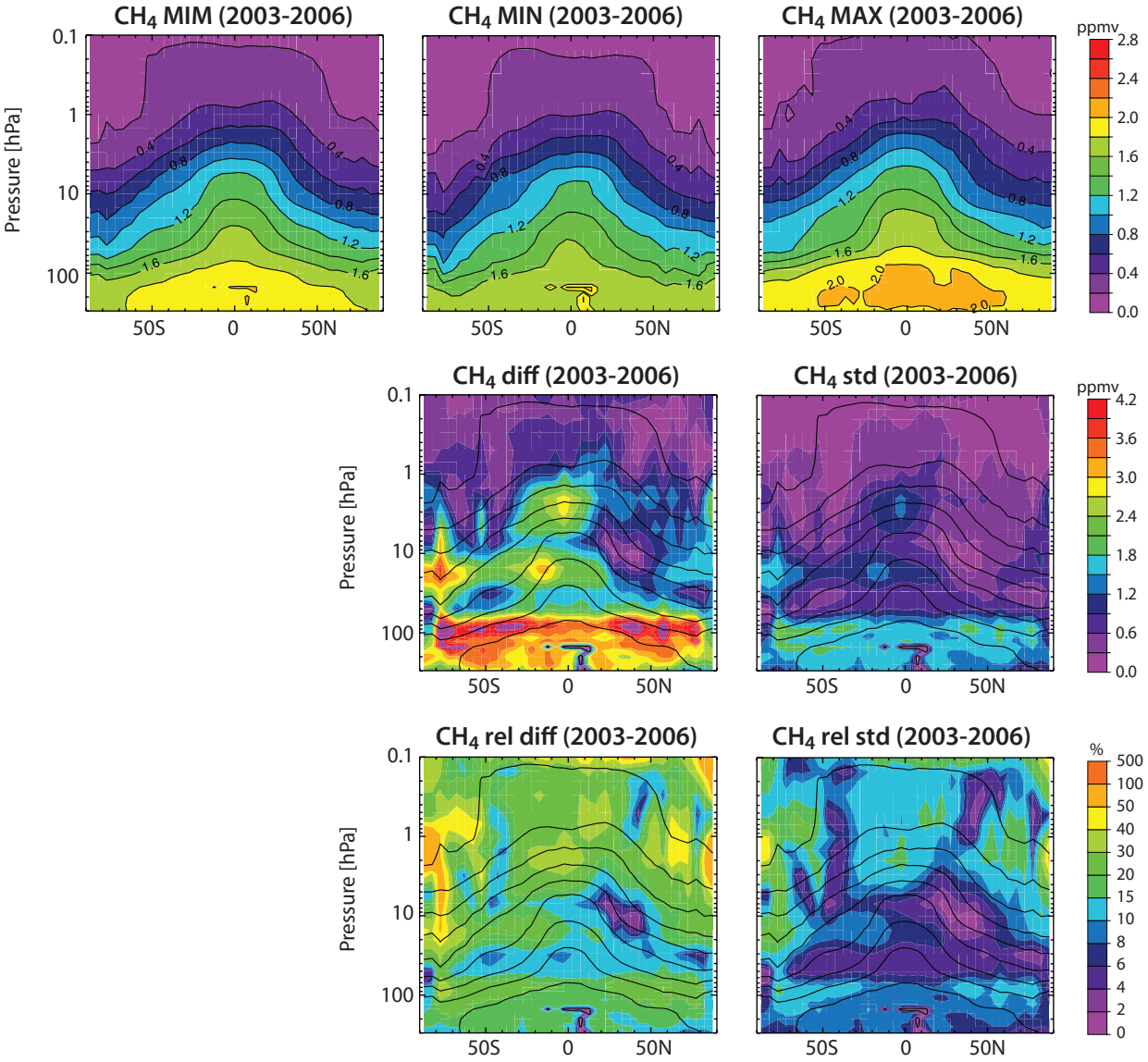


Figure 4.3.6: Summary of CH_4 annual zonal mean state for 2003-2006. *Annual zonal mean cross sections for 2003-2006 of the MIM, minimum (MIN), and maximum (MAX) CH_4 values are shown in the upper row. The maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments 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.

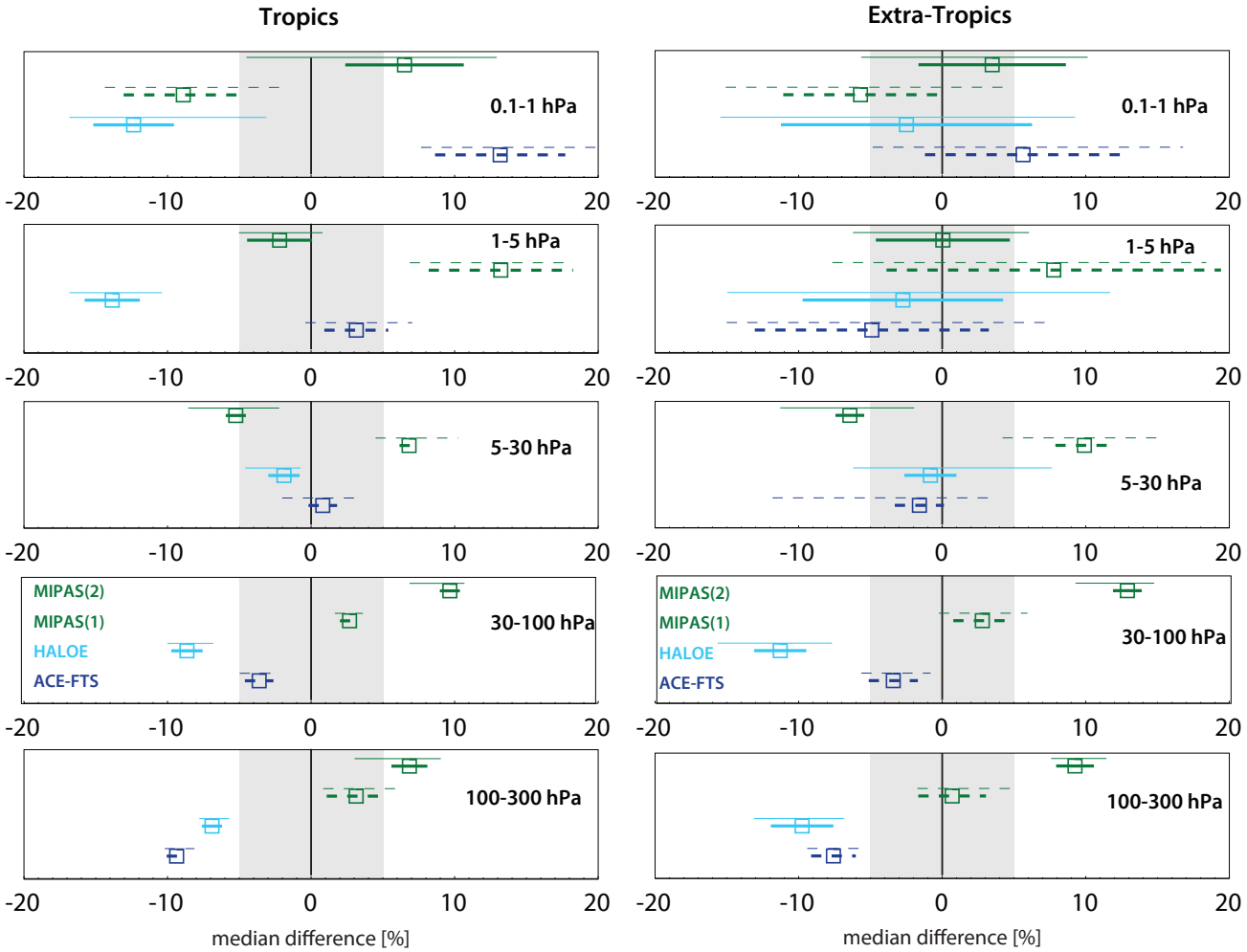


Figure 4.3.7: Summary plot of CH_4 inter-instrument differences for 2003-2006. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology 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.*

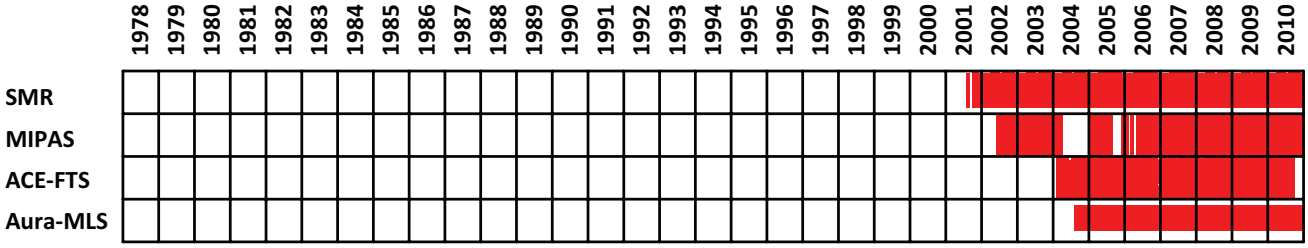


Table 4.4.2: Time period, vertical range, vertical resolution, references and other comments for N_2O measurements.

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

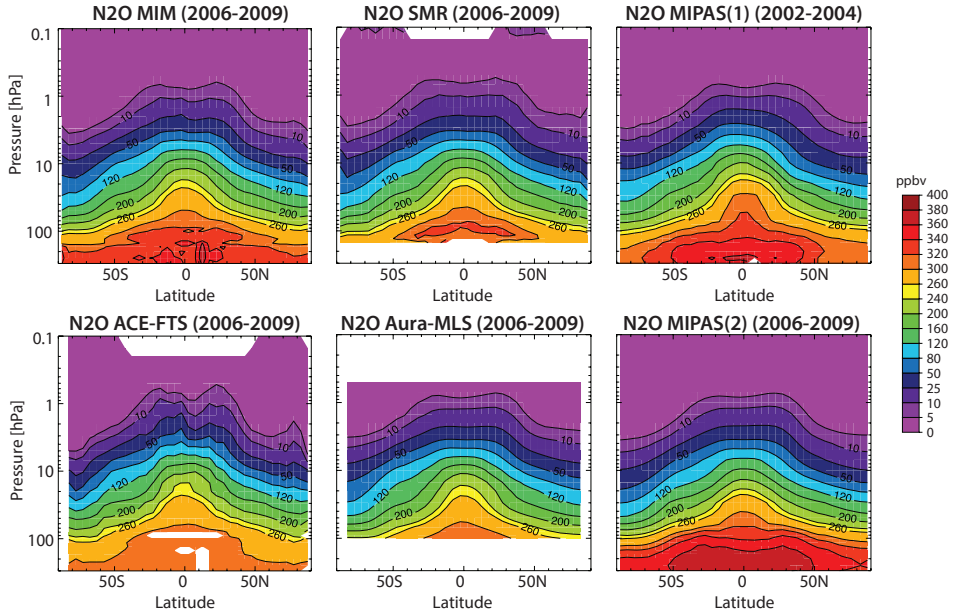


Figure 4.4.1a: Cross sections of annual zonal mean N_2O for 2006-2009. Annual zonal mean N_2O cross sections are shown for the MIM in the leftmost upper panel along with SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2). Note, MIPAS(1) is excluded from the MIM so not to bias the MIM towards this instrument.

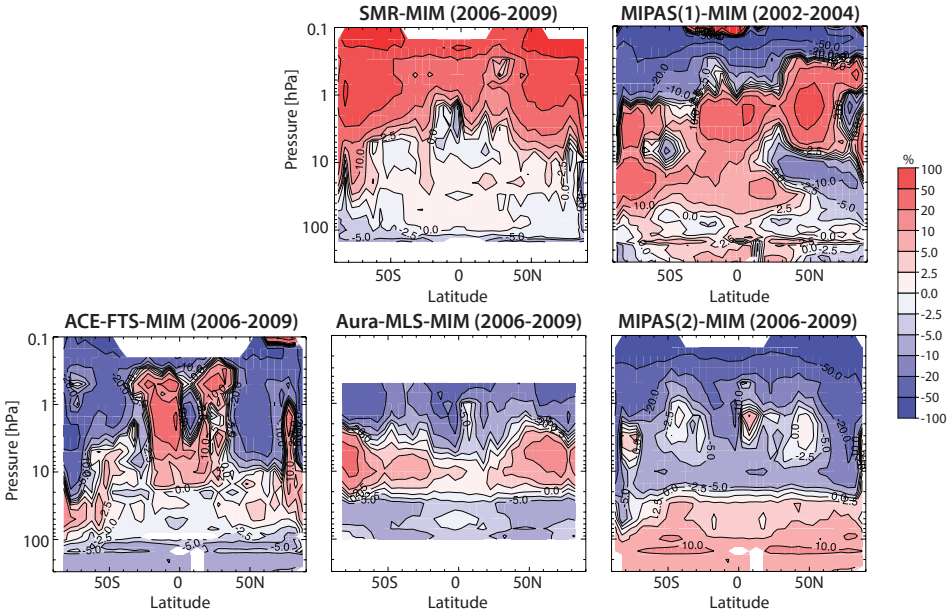


Figure 4.4.1b: Cross sections of annual zonal mean N_2O differences for 2006-2009. *Shown are the relative differences between the individual instruments' (SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2)) annual zonal mean N_2O distributions and the MIM.*

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.

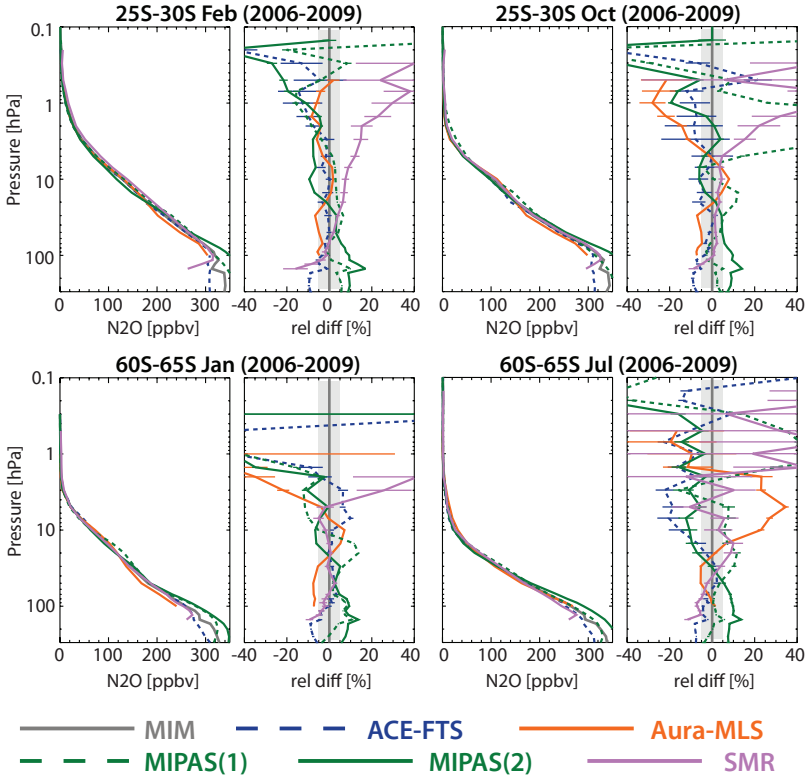


Figure 4.4.2: Vertical profiles of monthly zonal mean N_2O for the Southern Hemisphere. *Vertical N_2O profiles for 25°S-30°S February and October (upper panels), and for 60°S-65°S January and July (lower panels) are shown together with the instrument differences from the MIM for SMR, MIPAS(1), ACE-FTS,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

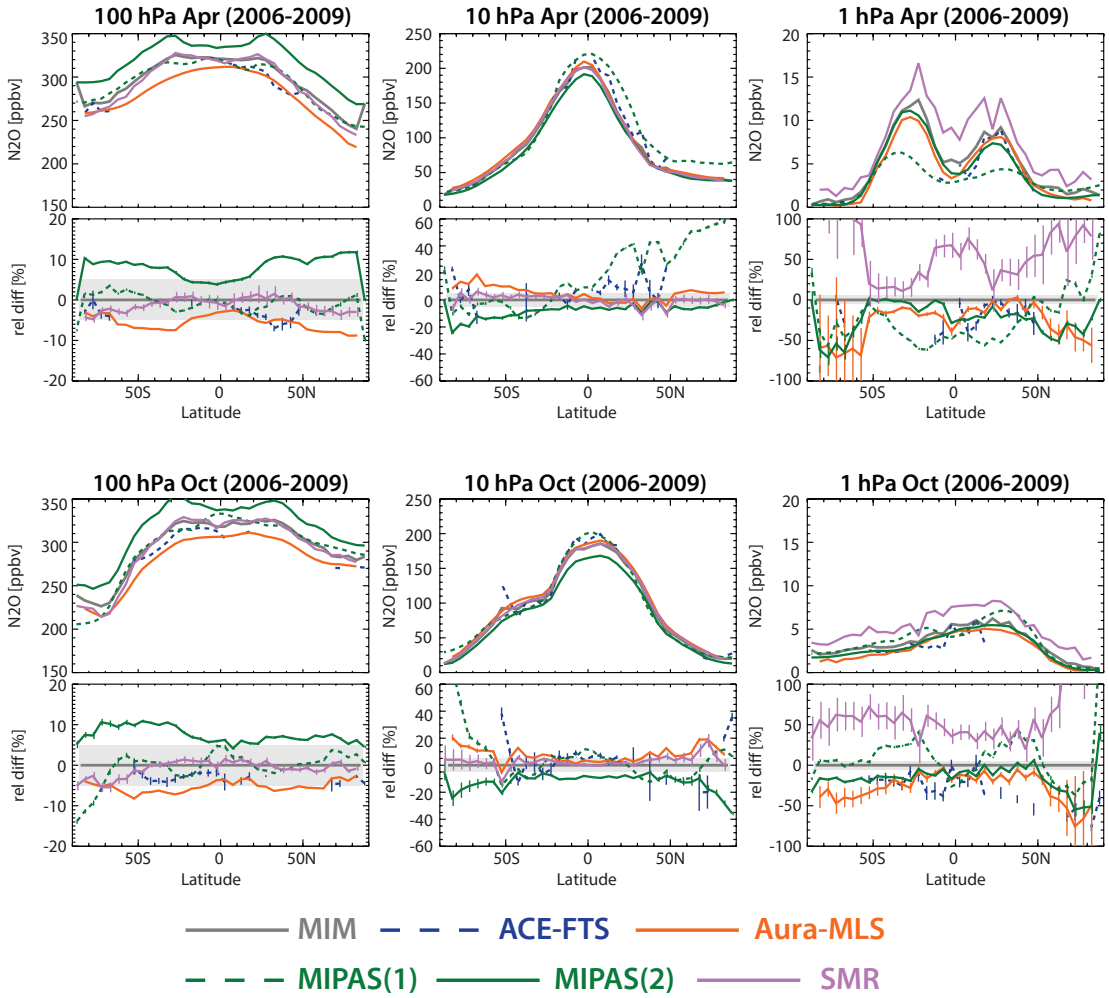


Figure 4.4.3: Meridional profiles of monthly zonal mean N_2O for 2006-2009. *Meridional N_2O profiles are shown at 100, 10, and 1 hPa for April (upper row) and October (lower row). Differences between the individual instruments (SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2)) and the MIM profiles are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

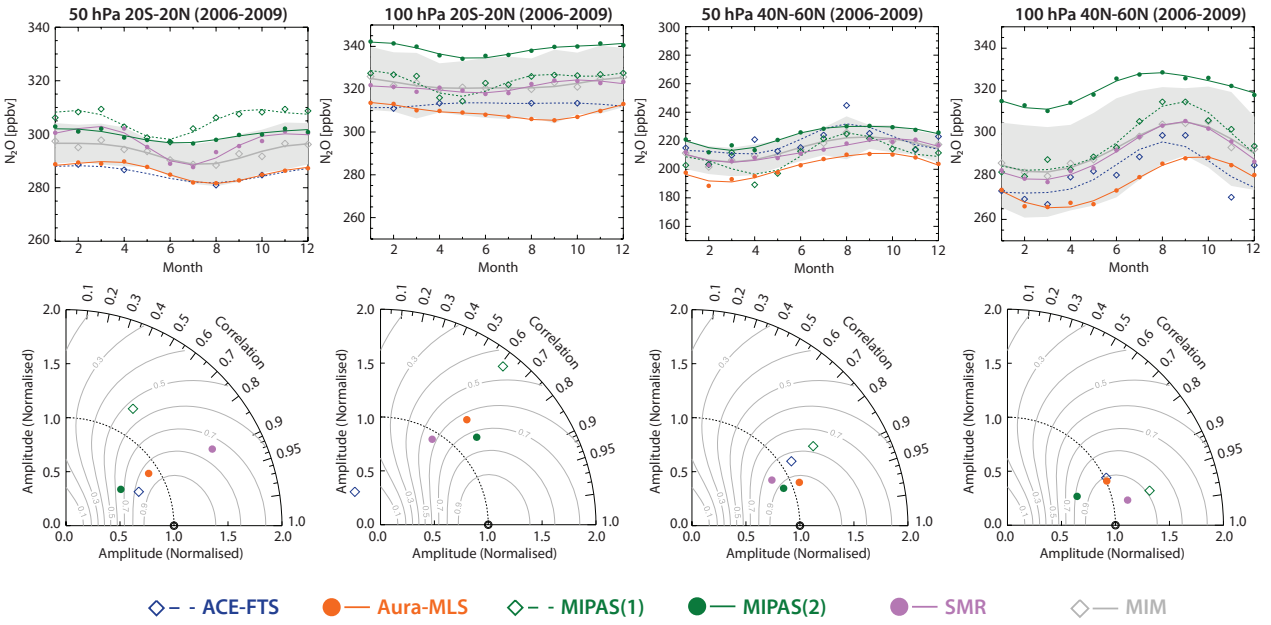


Figure 4.4.4: Seasonal cycle of N_2O in the tropics and at NH mid-latitudes at 100 and 50 hPa. *Seasonal cycles (upper panels) and corresponding Taylor diagrams (bottom panels) for monthly zonal mean N_2O are shown for 20°S-20°N (left two panels) and 40°N-60°N (right two panels) at 50 and 100 hPa, averaged 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%.

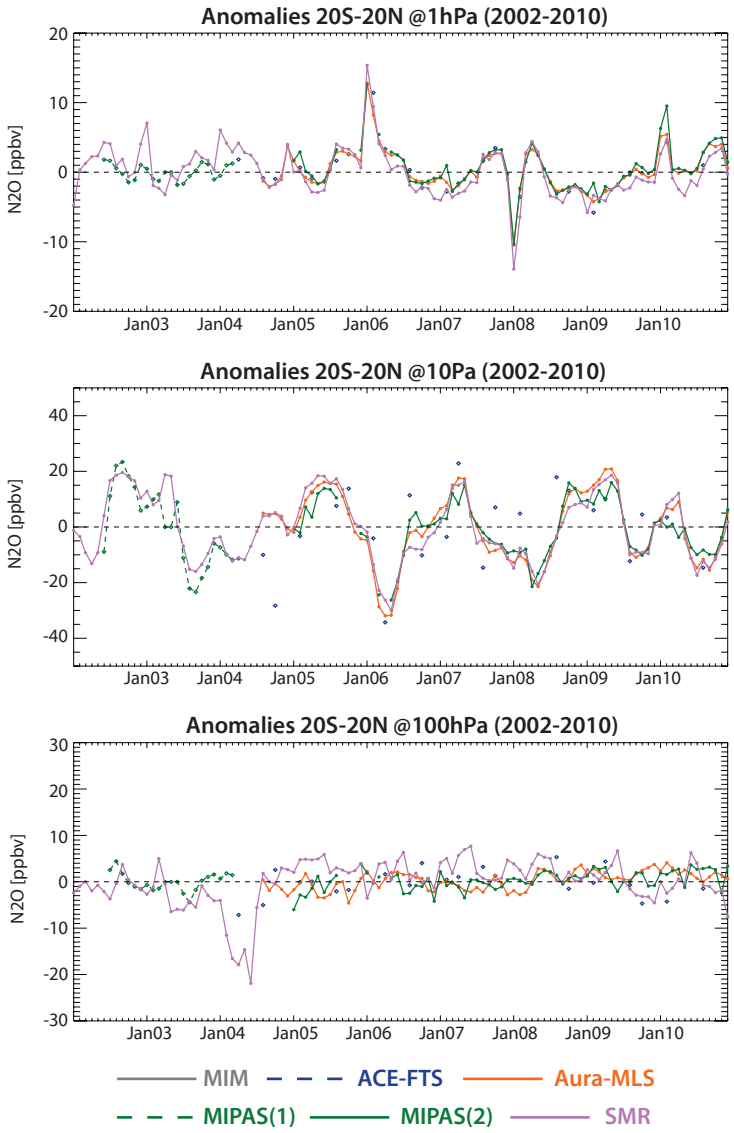


Figure 4.4.5: Time series of deseasonalised N_2O anomalies in the tropics. *Deseasonalised N_2O anomalies between 20°S-20°N are shown for the 1 hPa (upper panel), 10 hPa (middle panel), and 100 hPa (lower panel) levels.*

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.

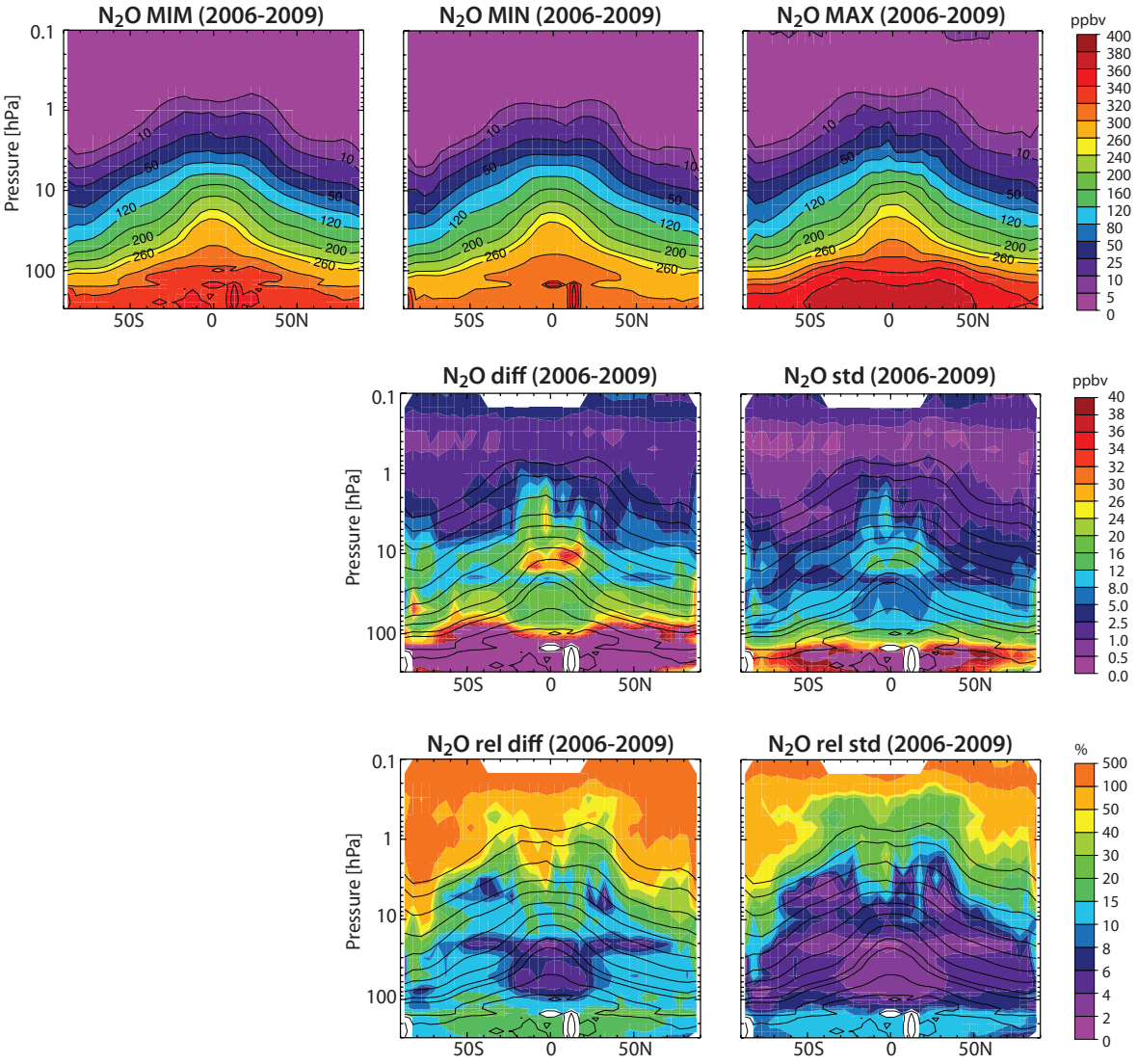


Figure 4.4.6: Summary of N_2O annual zonal mean state for 2006-2009. *Annual zonal mean cross sections for 2006-2009 of the MIM, minimum (MIN), and maximum (MAX) N_2O values are shown in the upper row. The maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments 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

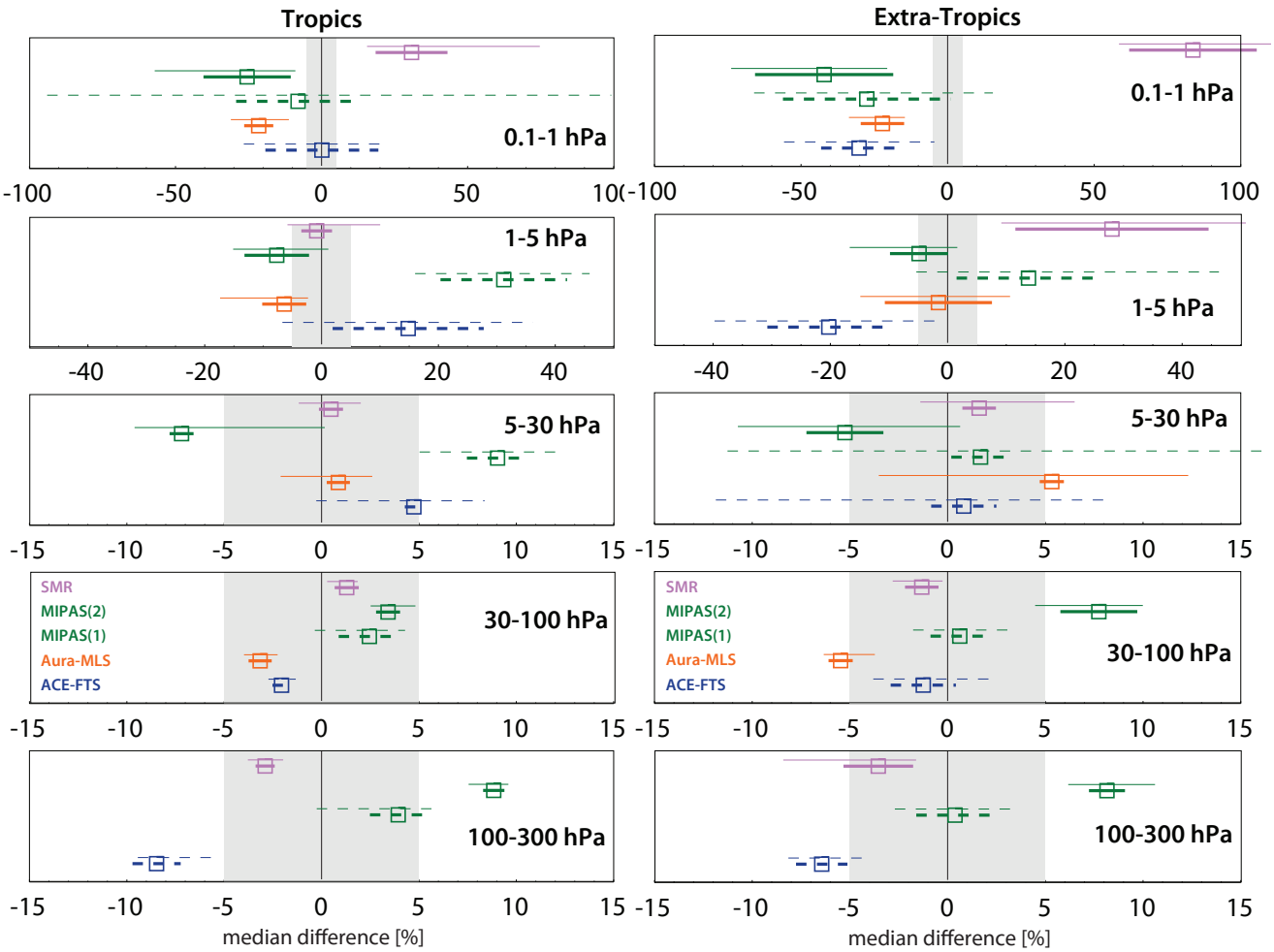


Figure 4.4.7: Inter-instrument differences in N_2O calculated for the tropics (left) (20°S–20°N) and (right) extra-tropics (40°S–80°S and 40°N–80°N) and for five different altitude regions from the UT up to the LM. *Shown are the median (squares), median absolute deviations (MAD, thick lines), and the 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.*

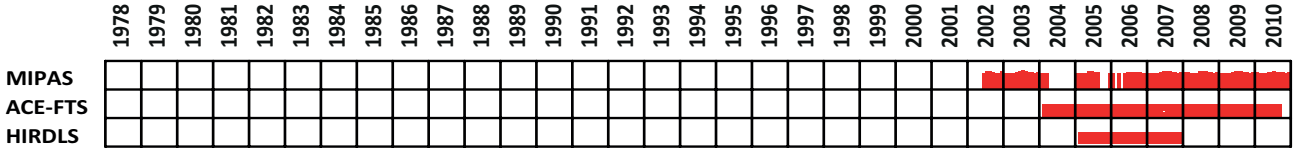


Table 4.5.2: Time period, vertical range, vertical resolution, references and other comments for CFC-11 measurements.

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	

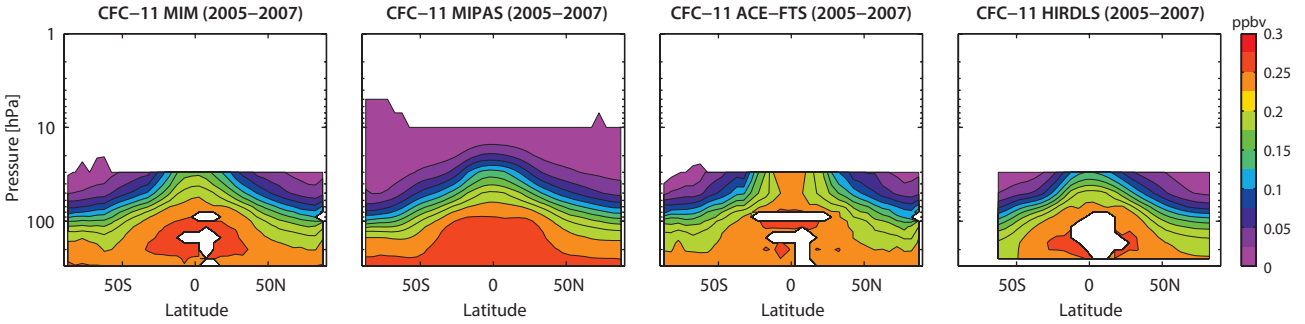


Figure 4.5.1: Cross sections of annual zonal mean CFC-11 for 2005-2007. *Annual zonal mean CFC-11 cross sections are shown for the MIM, MIPAS, ACE-FTS, and HIRDLS. The MIM is only displayed for regions where at least two instruments provide measurements.*

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

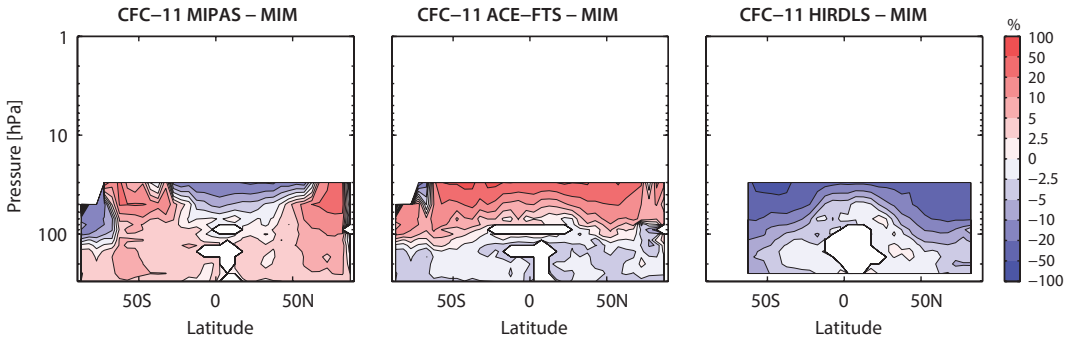


Figure 4.5.2: Cross sections of annual zonal mean CFC-11 differences for 2005-2007. *Annual zonal mean CFC-11 differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM are shown.*

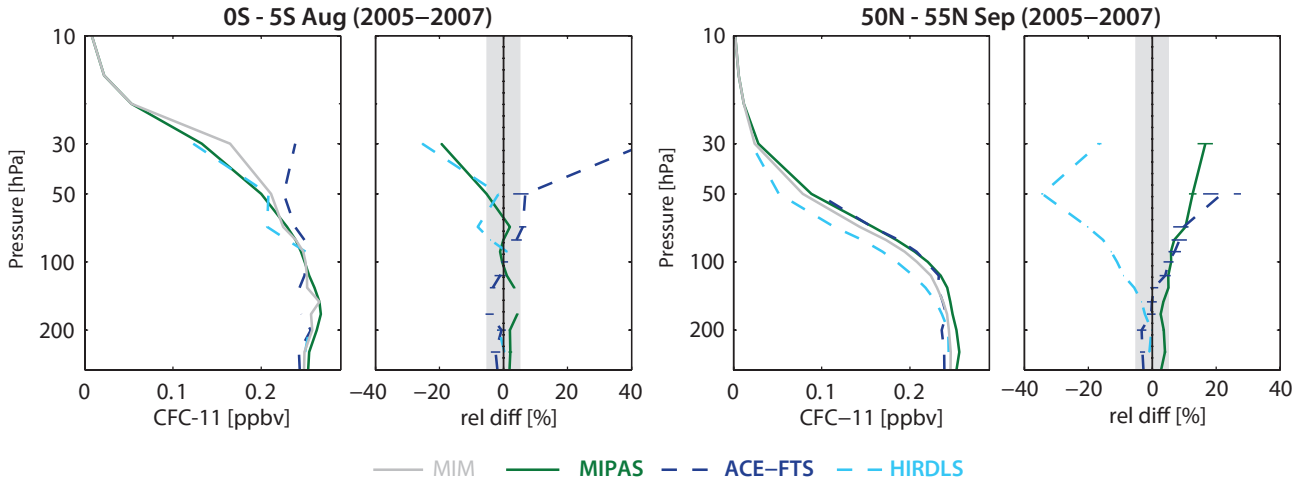


Figure 4.5.3: Profiles of monthly zonal mean CFC-11 for 2005-2007. *Vertical CFC-11 profiles for 0°S-5°S, August and 50°N-55°N September are shown together with their relative differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative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

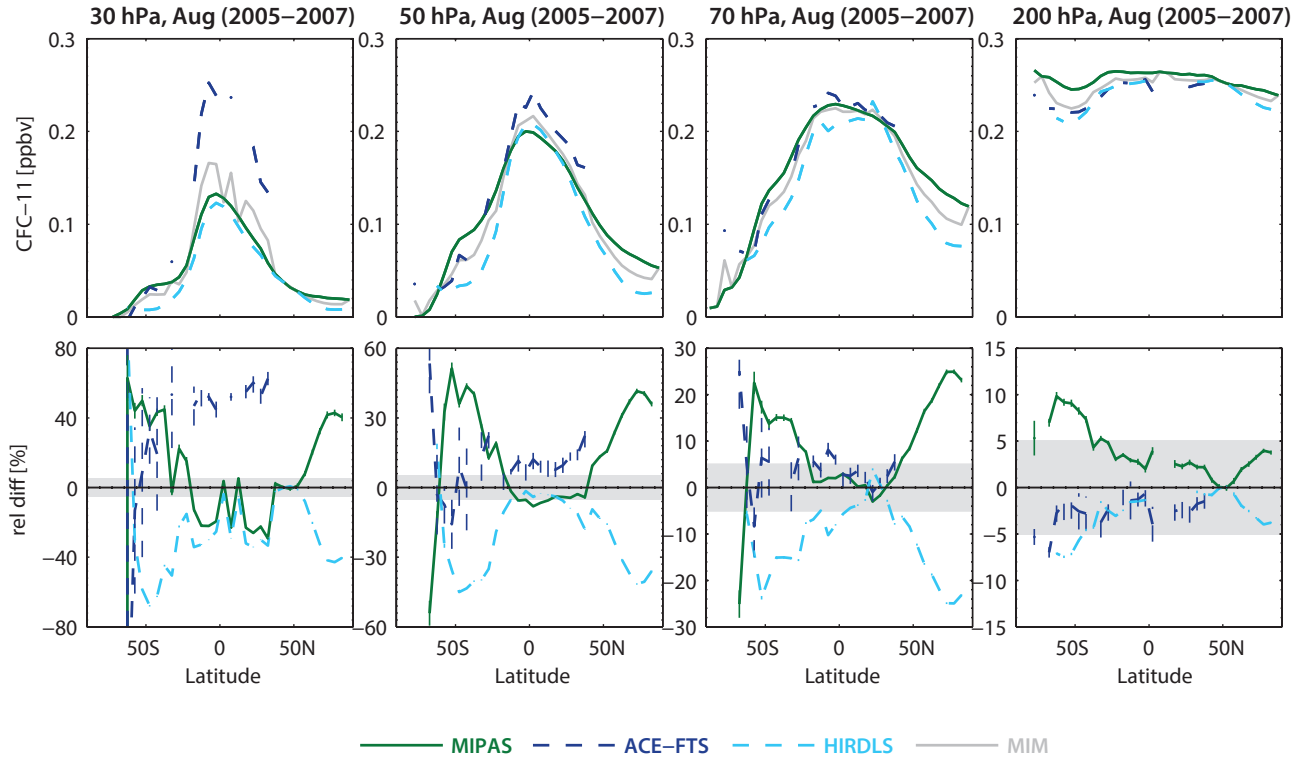


Figure 4.5.4: Meridional profiles of monthly zonal mean CFC-11 for 2005-2007. *Meridional CFC-11 profiles at 30, 50, 70 and 100 hPa for August are shown in the upper row. Relative differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM profiles are shown in the lower row.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

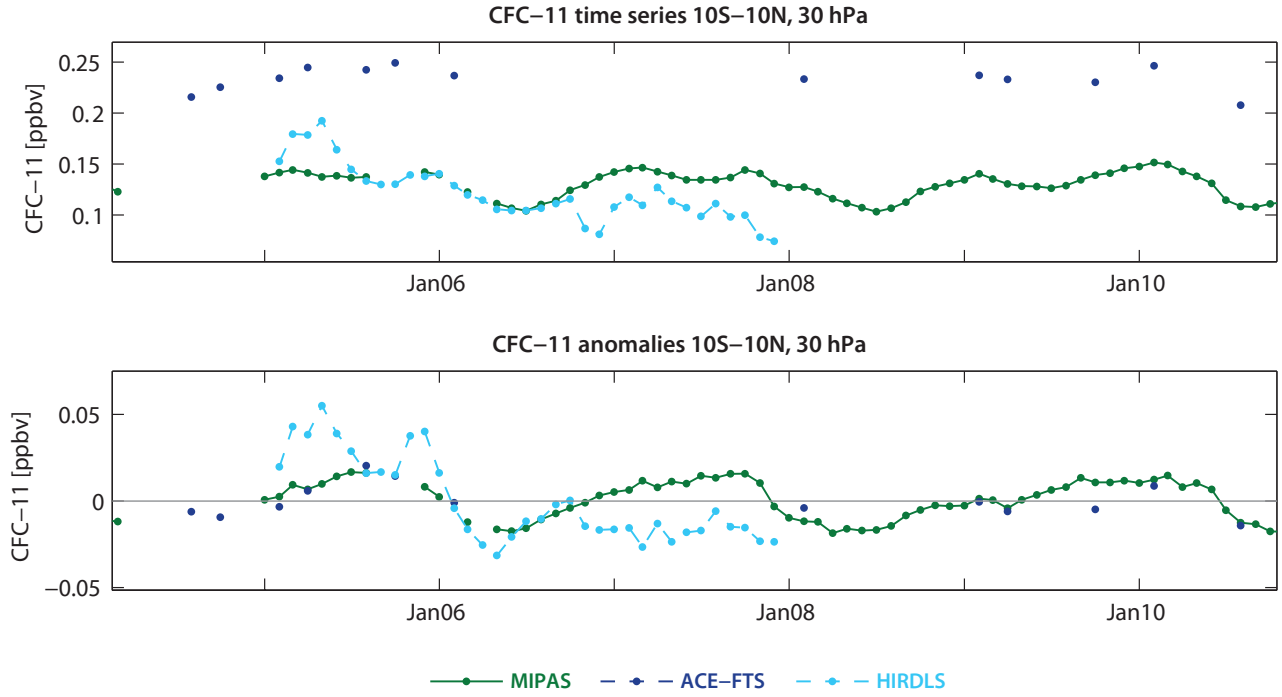


Figure 4.5.5: Time series of CFC-11 monthly zonal mean values and deseasonalised anomalies in the tropics. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-11 between 10°S – 10°N at 30 hPa.*

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.

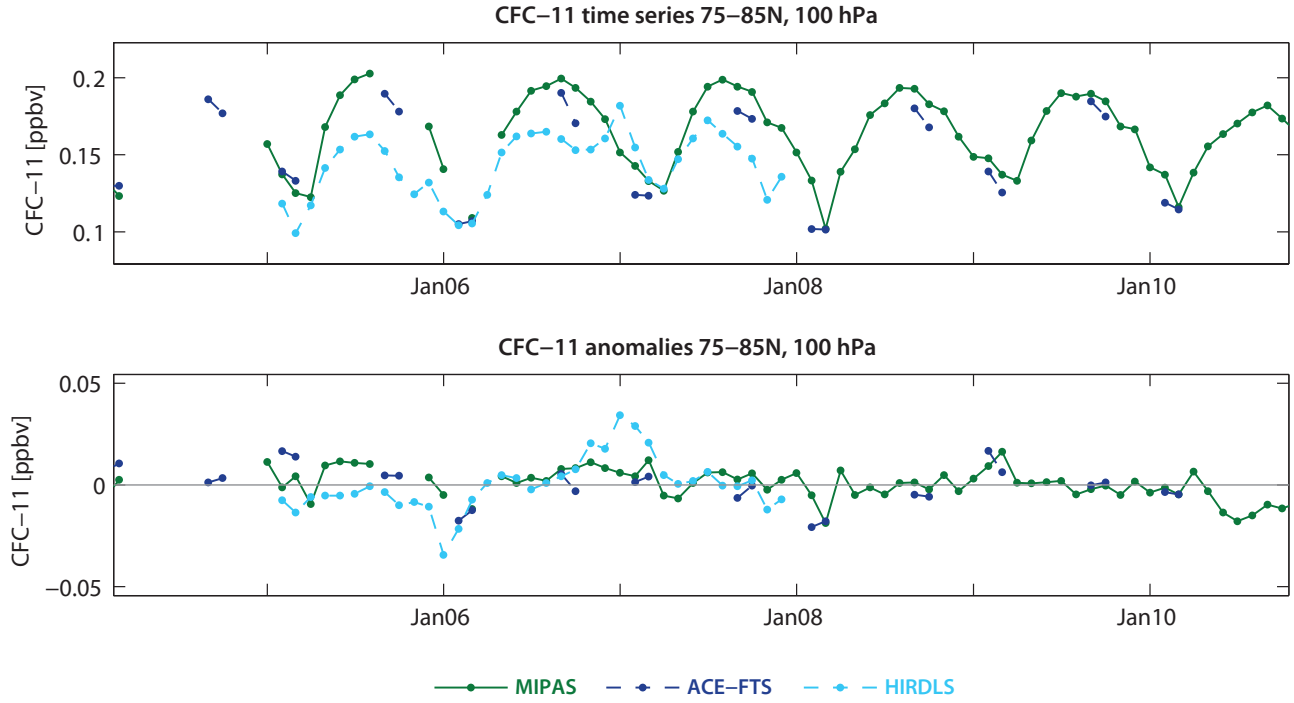


Figure 4.5.6: Time series of CFC-11 monthly zonal mean values and deseasonalised anomalies at NH high latitudes. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-11 between 75°N – 85°N at 100 hPa.*

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.

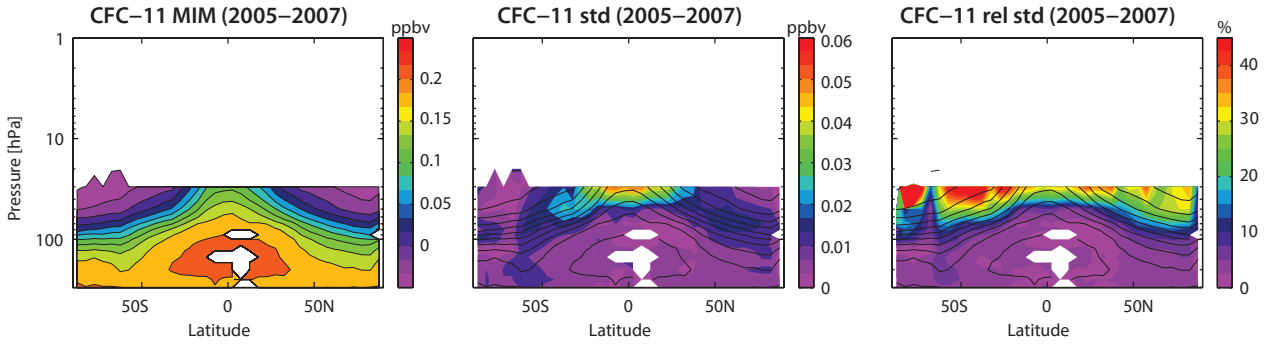


Figure 4.5.7: Summary of CFC-11 annual zonal mean state for 2005-2007. *Shown are the annual zonal mean cross section for the MIM of CFC-11 (left panel), the standard deviation over all three instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black 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.*

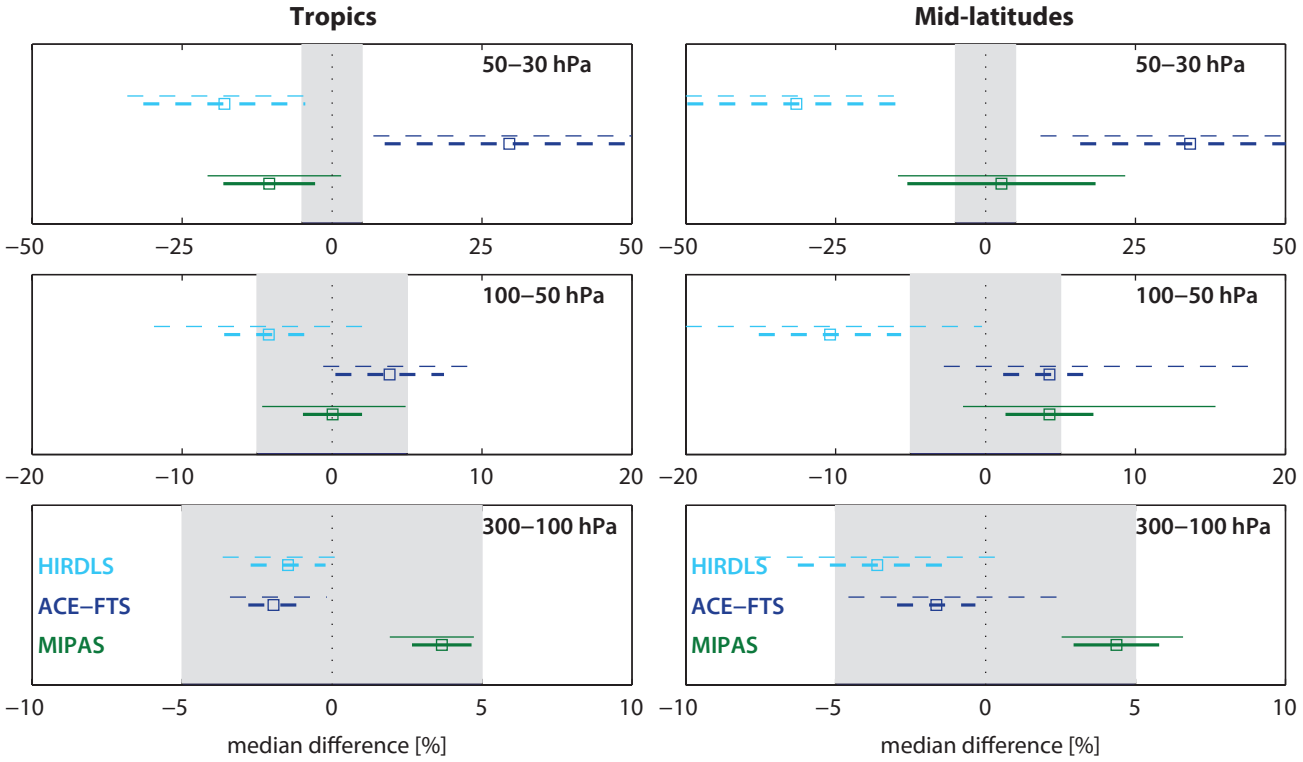


Figure 4.5.8: Summary CFC-11 differences for 2005-2007. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the 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.*

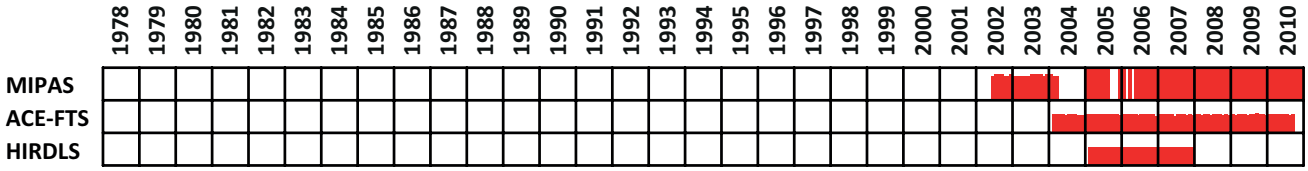


Table 4.6.2: Time period, vertical range, vertical resolution, references and other comments for CFC-12 measurements.

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

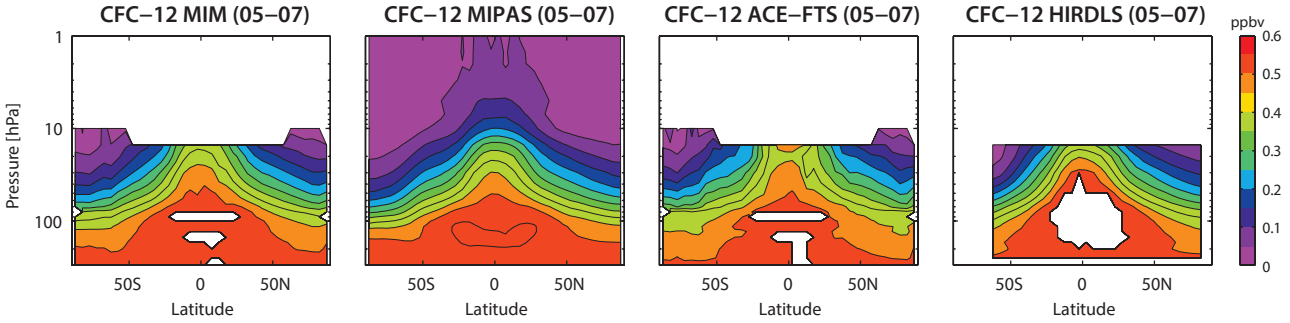


Figure 4.6.1: Cross sections of annual zonal mean CFC-12 for 2005-2007. *Annual zonal mean CFC-12 cross sections are shown for the MIM, MIPAS, ACE-FTS, and HIRDLS. The MIM is only displayed for regions where at least two instruments provide measurements.*

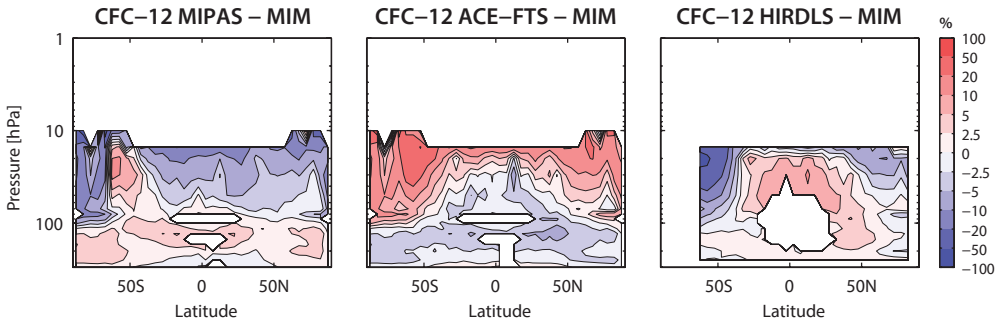


Figure 4.6.2: Cross sections of annual zonal mean CFC-12 differences for 2005-2007. *Annual zonal mean CFC-12 differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM are shown.*

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*).

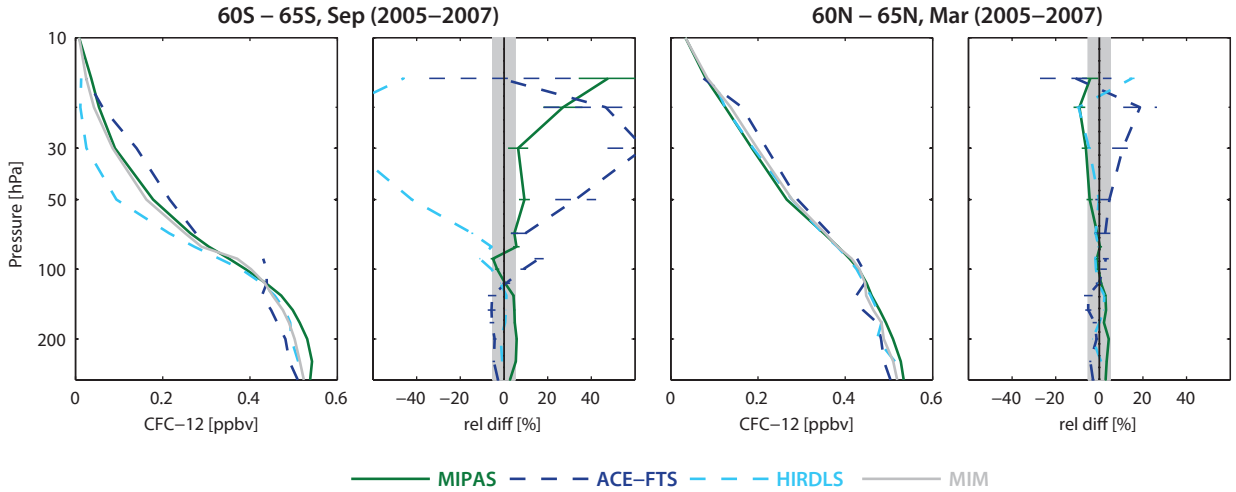


Figure 4.6.3: Profiles of zonal mean CFC-12 for 2005-2007. *Zonal mean CFC-12 profiles for 60°S-65°S in September and 60°N-65°N in March are shown together with their relative differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative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

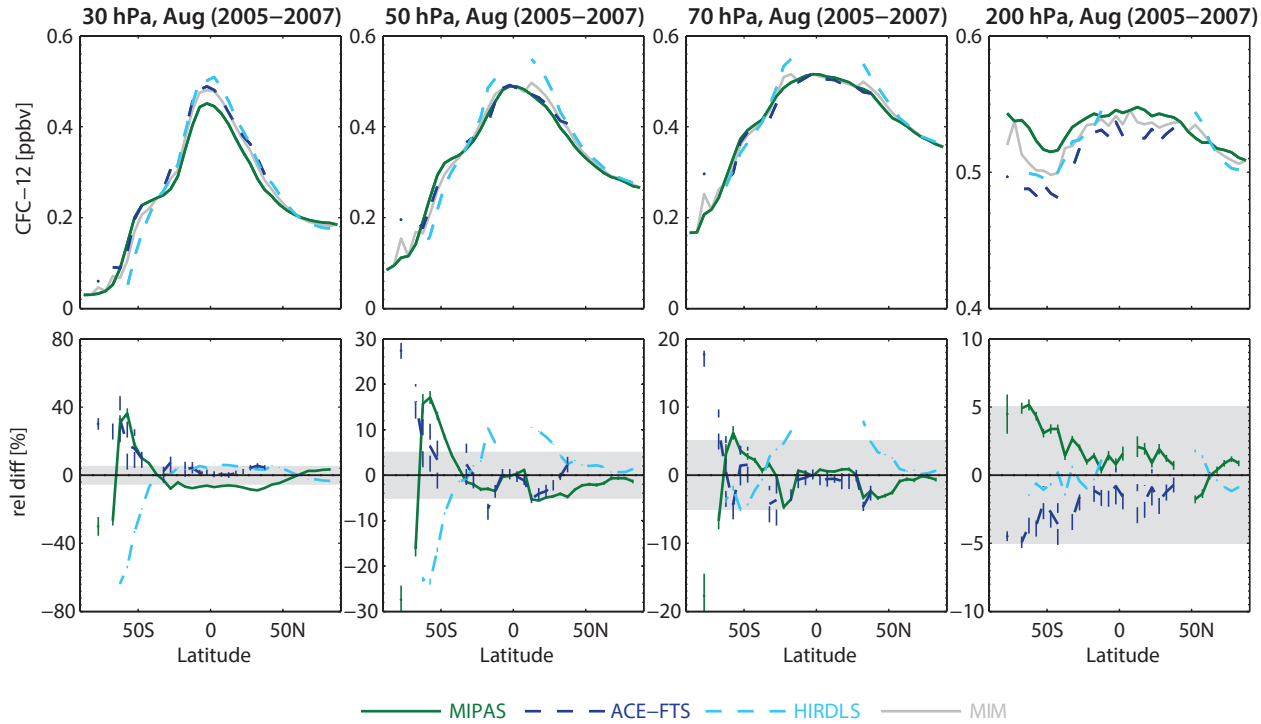


Figure 4.6.4: Meridional profiles of zonal mean CFC-12 for 2005-2007. *Meridional zonal mean CFC-12 profiles at 30, 50, 70, and 200 hPa for August are shown in the upper row. Relative differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM profiles are shown in the lower 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

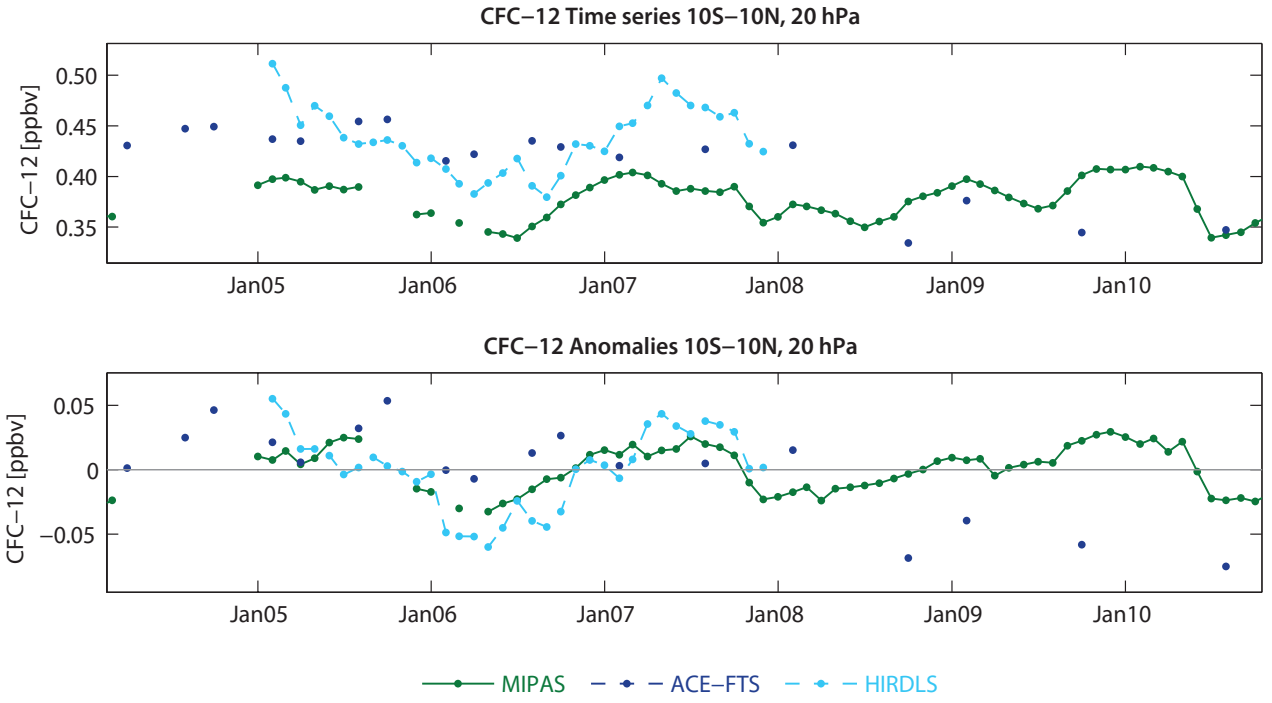


Figure 4.6.5: Time series of CFC-12 monthly mean values and deseasonalised anomalies in the tropics. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-12 between 10°S – 10°N at 20 hPa.*

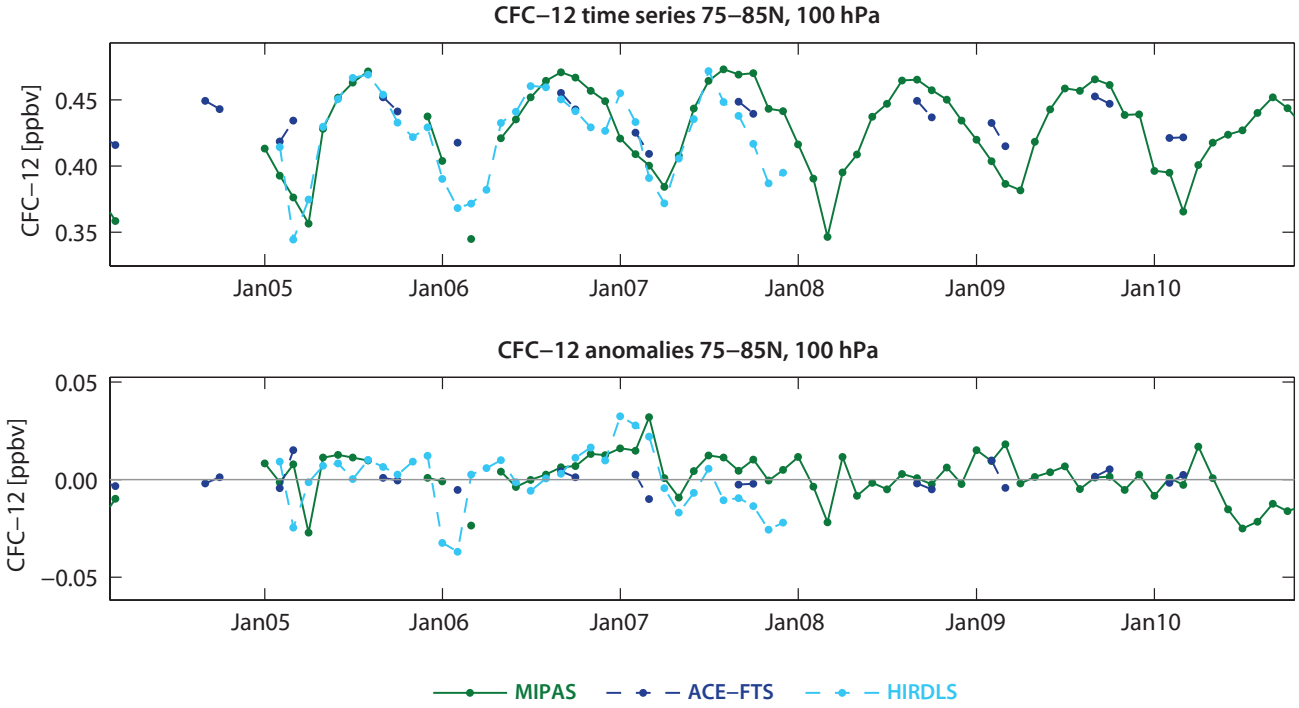


Figure 4.6.6: Time series of CFC-12 monthly mean values and deseasonalised anomalies at NH high latitudes. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-12 between 75°N – 85°N at 100 hPa.*

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

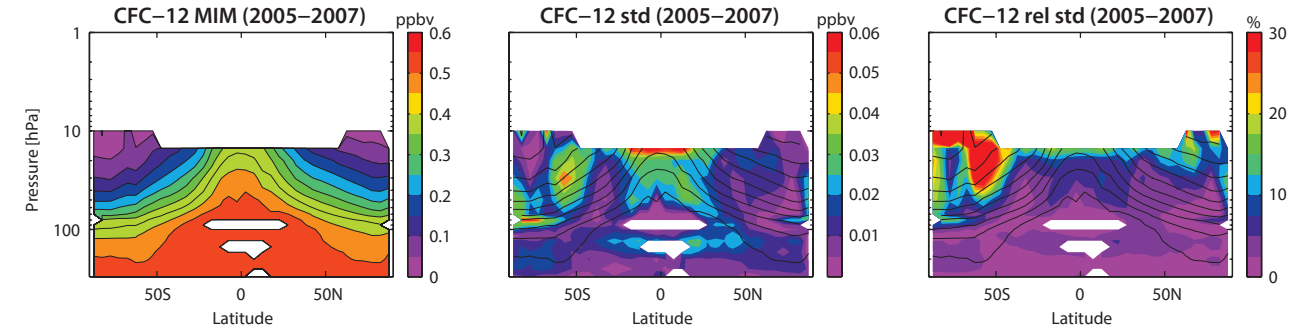


Figure 4.6.7: Summary of CFC-12 annual zonal mean state for 2005-2007. *Shown are the annual zonal mean cross section for the MIM of CFC-12 (left panel), the standard deviation over all three instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black 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.*

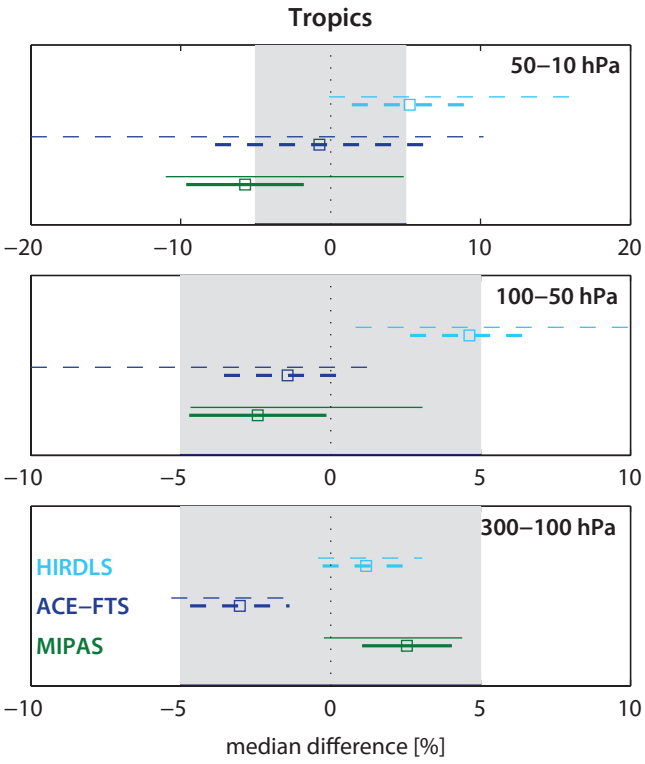


Figure 4.6.8a: Summary CFC-12 differences in the tropics for 2005-2007. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology 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

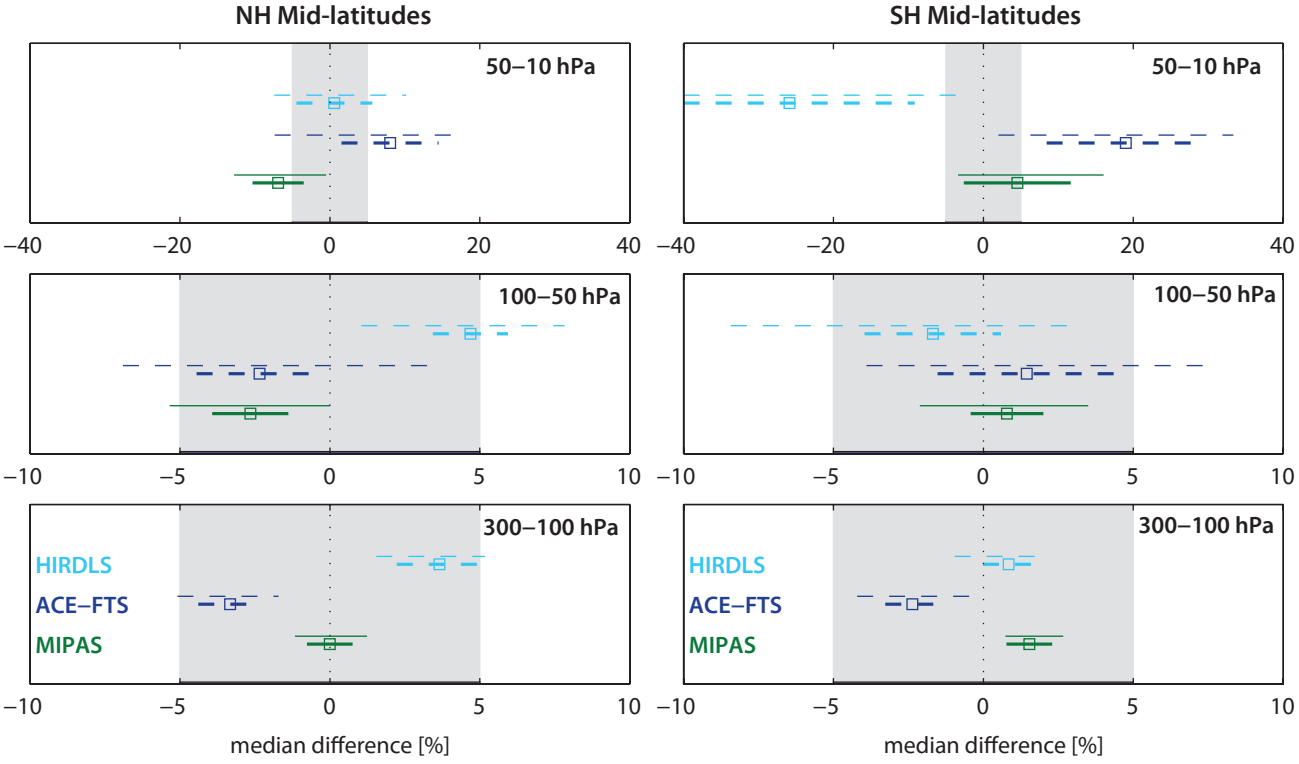


Figure 4.6.8b: Summary CFC-12 differences in mid-latitudes for 2005-2007. *Like Figure 4.6.8a but for NH mid-latitudes (30°N-60°N) and SH mid-latitudes (30°S-60°S).*

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.

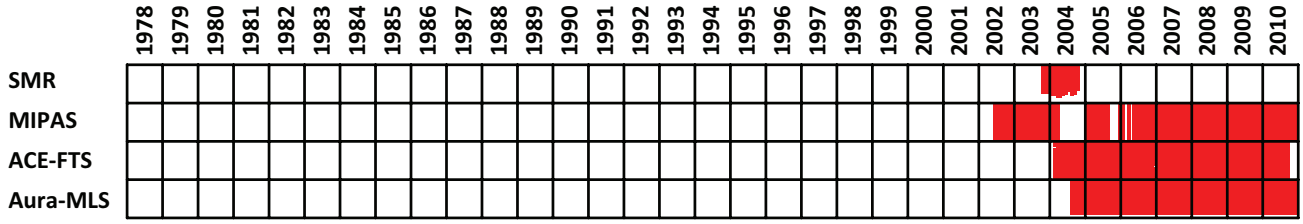
4.7.2 CO evaluations: Zonal mean cross sections, vertical and meridional profiles

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	

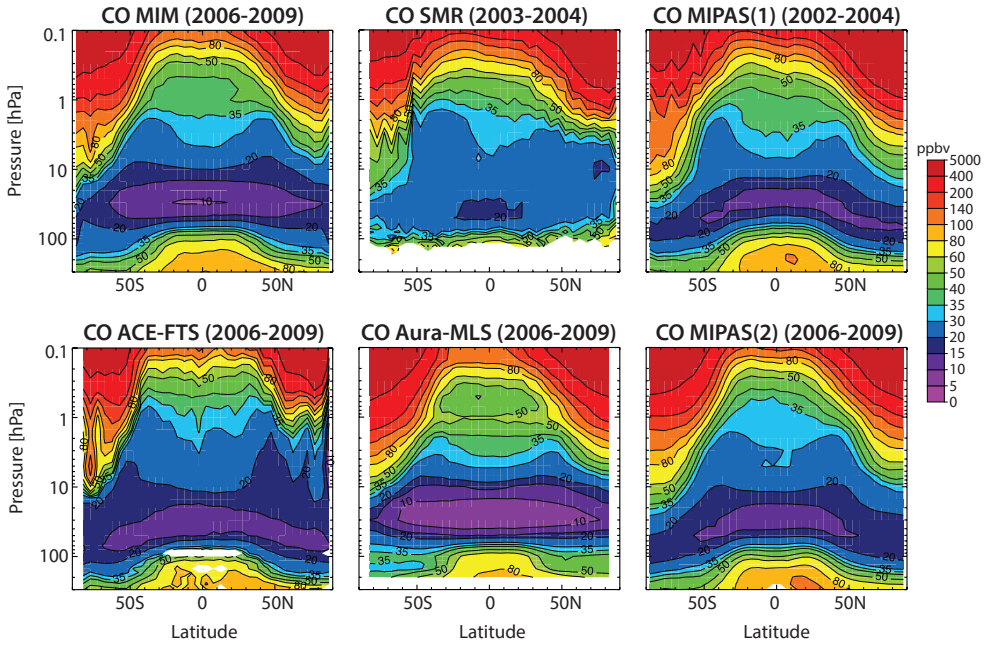


Figure 4.7.1a: Cross sections of annual zonal mean CO for 2006-2009. *Cross sections of CO are shown for the MIM, SMR, MIPAS(1), ACE-FTS, MIPAS(2), and Aura-MLS. Note that SMR is averaged over the period October 2003-October 2004 and MIPAS(1) over July 2002-March 2004. These datasets are not included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

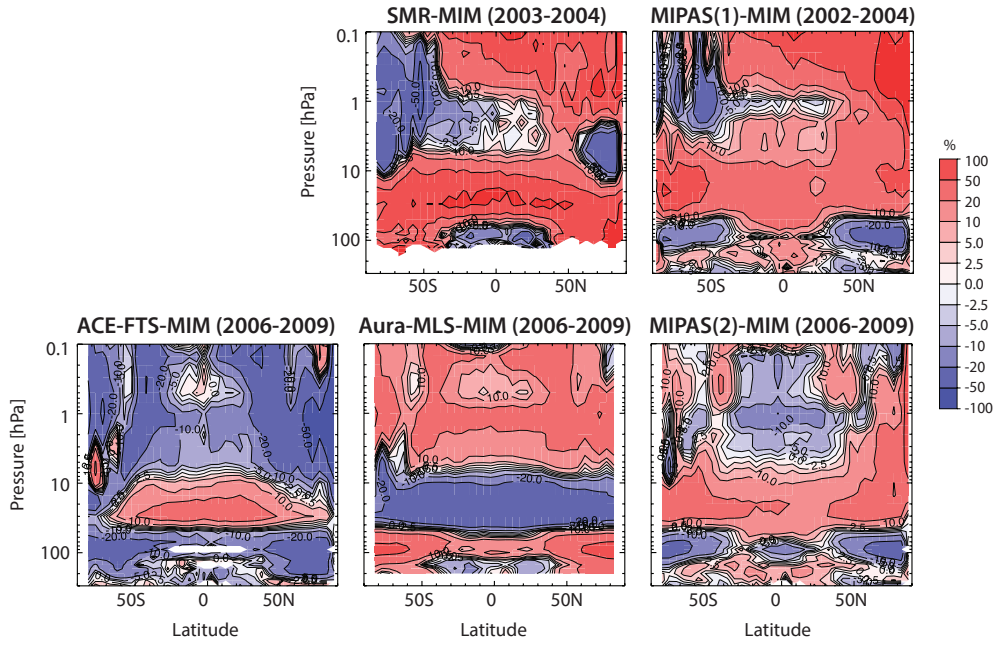


Figure 4.7.1b: Cross sections of annual zonal mean CO differences for 2006-2009. *Cross sections of CO relative differences between the individual instruments (MIPAS(1), ACE-FTS, Aura-MLS, MIPAS(2) and SMR) and the MIM are shown. Note, SMR (October 2003-October 2004) and MIPAS(1) (July 2002-March 2004)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

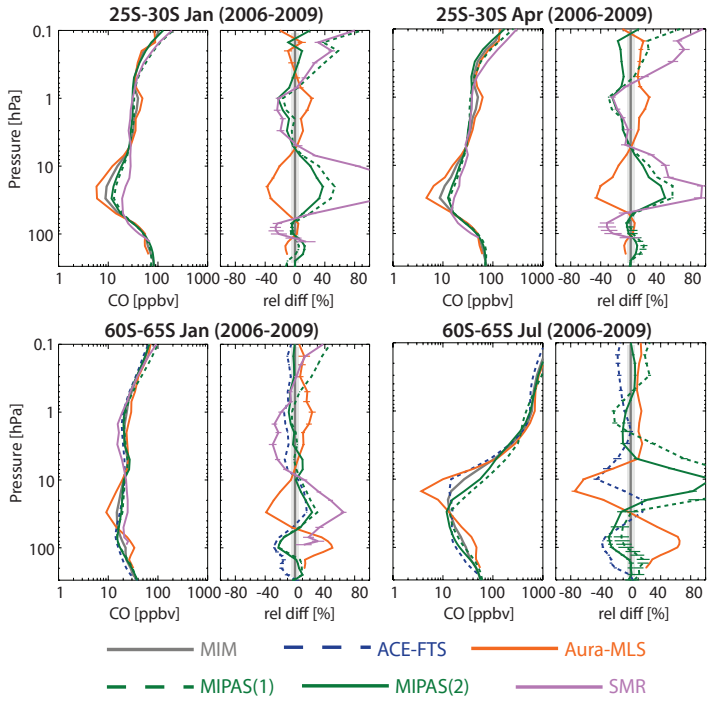


Figure 4.7.2a: Vertical profiles of monthly zonal mean CO for 2006-2009. *Vertical CO profiles for January and April 25°S-30°S (upper panels), and for January and July 60°S-65°S (lower panels) are shown together with instrument differences from the MIM. Note, SMR and MIPAS(1) measurements are taken in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

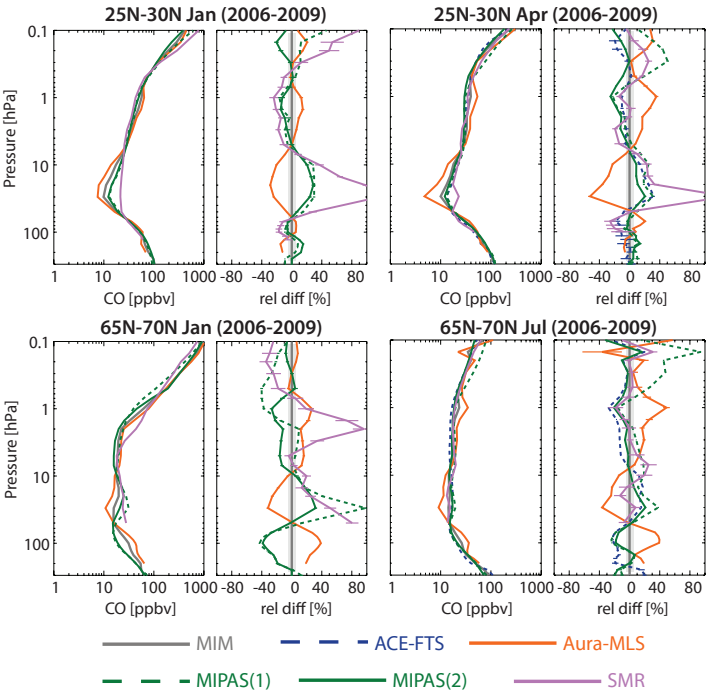


Figure 4.7.2b: Vertical profiles of monthly zonal mean CO for 2006-2009. *Vertical CO profiles for January and April 25°N-30°N (upper panels), and for January and July 65°N-70°N (lower panels) are shown together with instrument differences from the MIM. Note, SMR and MIPAS(1) measurements are taken in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

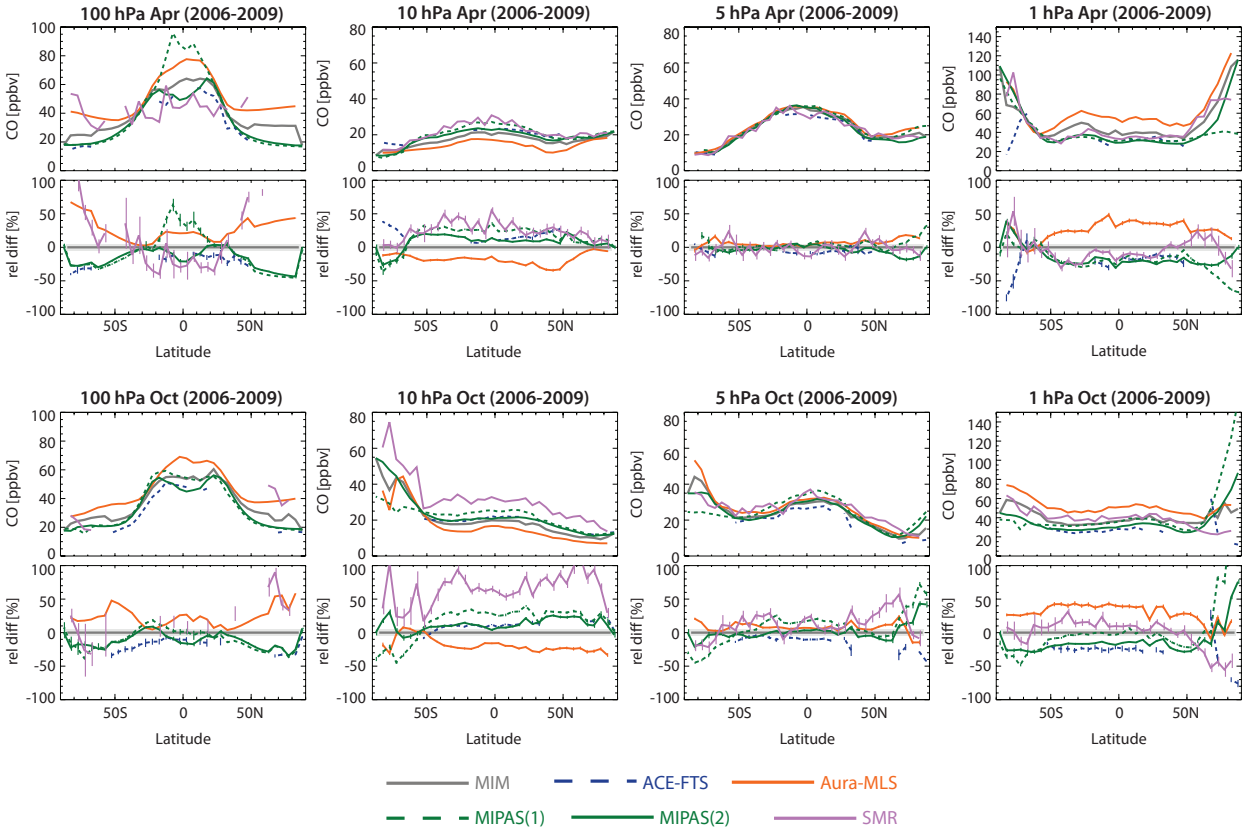


Figure 4.7.3: Meridional profiles of monthly zonal mean CO for 2006-2009. *Meridional CO profiles at 100, 10, 5, and 1 hPa for April (upper row) and October (lower row) averaged over 2006-2009. Differences between the individual instruments (SMR, MIPAS(1), MIPAS(2), ACE-FTS, and Aura-MLS) and the MIM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

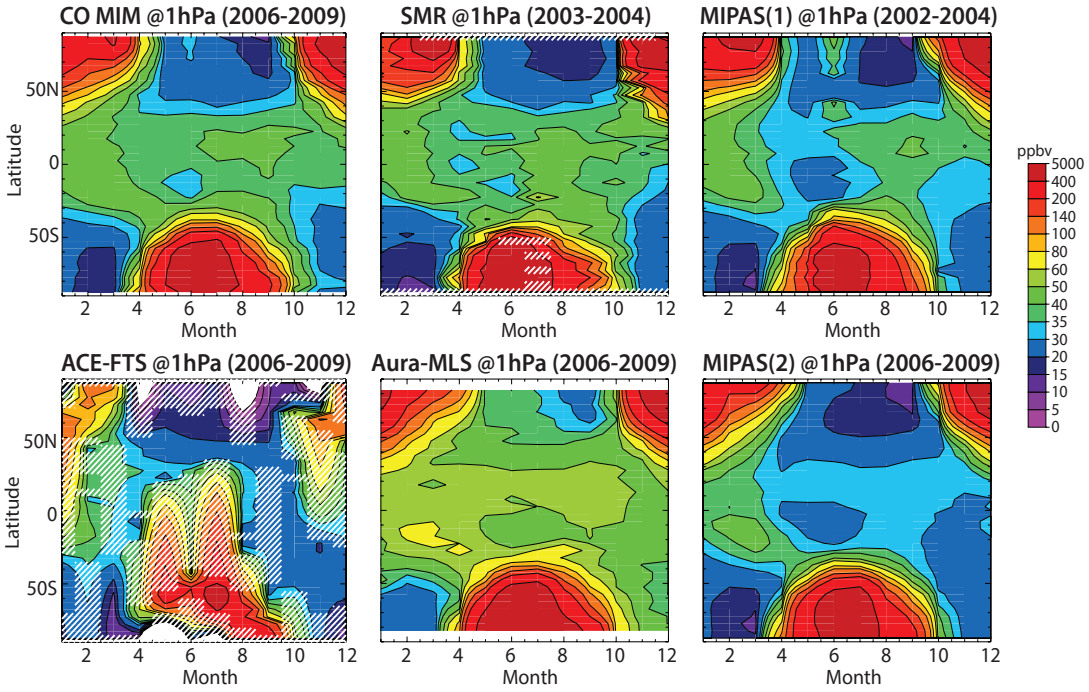


Figure 4.7.4a: Latitude–time evolution of CO at 1hPa. *The latitude-time evolution of CO at 1 hPa is shown for the MIM (2006-2009 average) in the upper leftmost panel and the instruments SMR, MIPAS(1), ACE-FTS, Aura-MLS and MIPAS(2). SMR and the ACE-FTS show interpolated fields, with hatched regions 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.

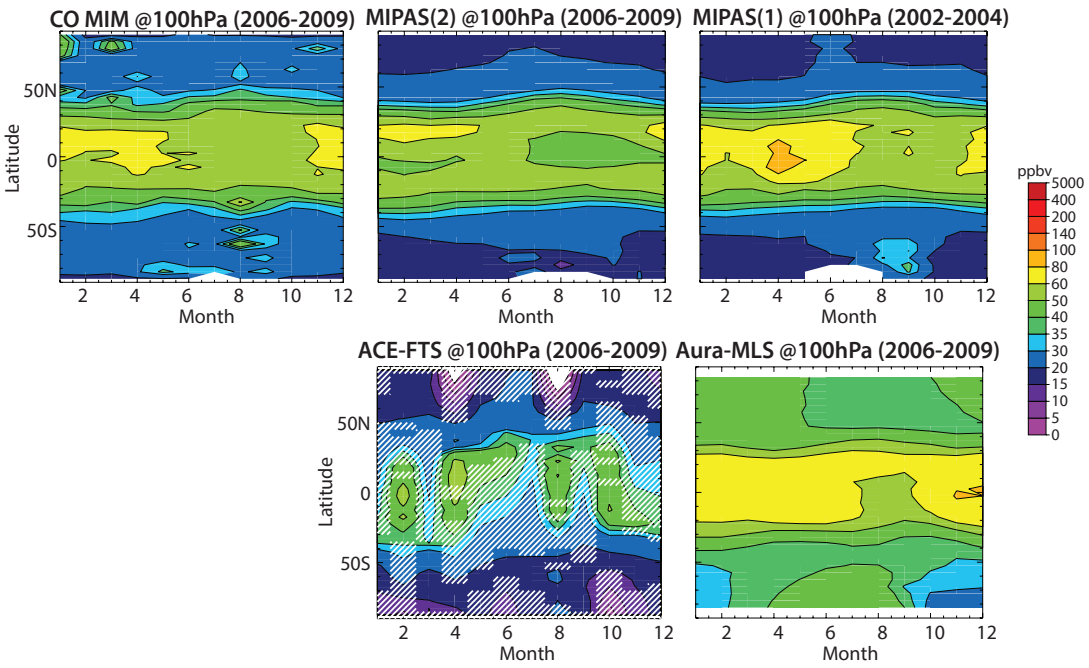


Figure 4.7.4b: Same as Figure 4.7.4a, but for 100 hPa.

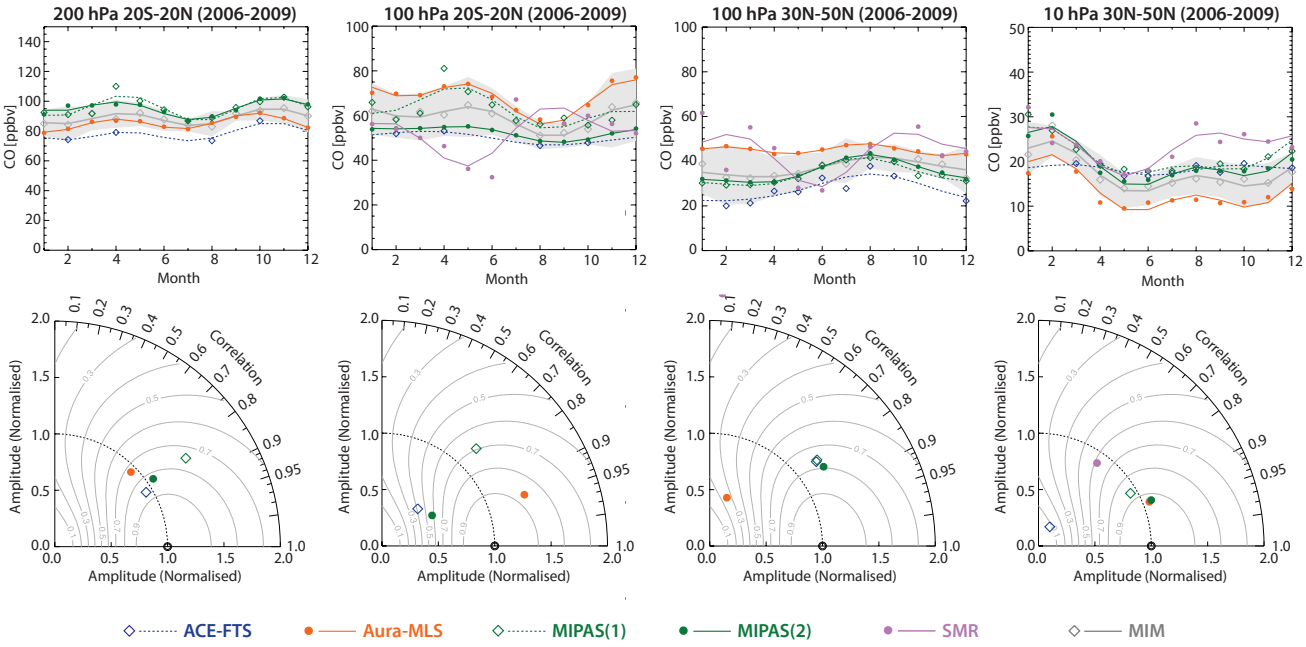


Figure 4.7.5: Seasonal cycle of CO. *Seasonal cycles (upper panels) and corresponding Taylor diagrams (lower panels) of monthly zonal mean CO are shown for the tropics (20°S-20°N) at 200 and 100 hPa (two left columns), and for the extra-tropics (30°N-50°N) at 100 and 10 hPa (two right columns).*

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

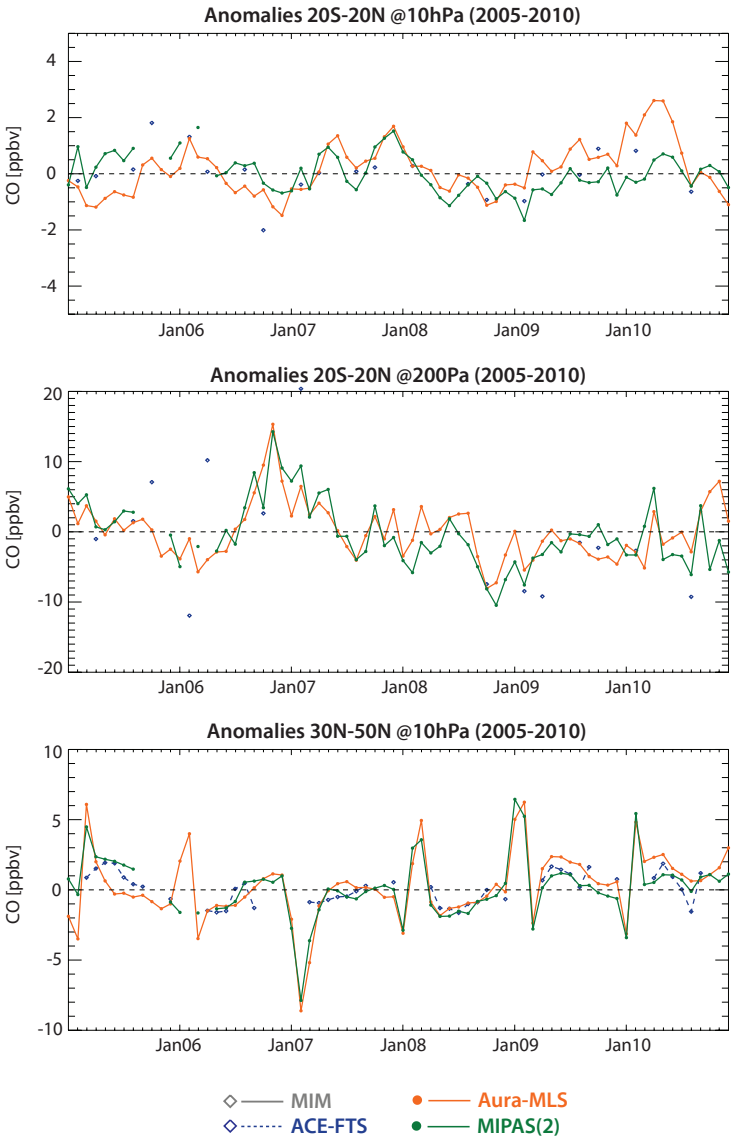


Figure 4.7.6: Time series of deseasonalised CO anomalies for 2005-2010. *Deseasonalised CO anomalies are shown for 20°S-20°N at 10 hPa (upper panel) and 200 hPa (middle panel), and for 30°N-50°N at 10 hPa (lower panel). Note that MIPAS here consists of MIPAS(2) data only.*

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

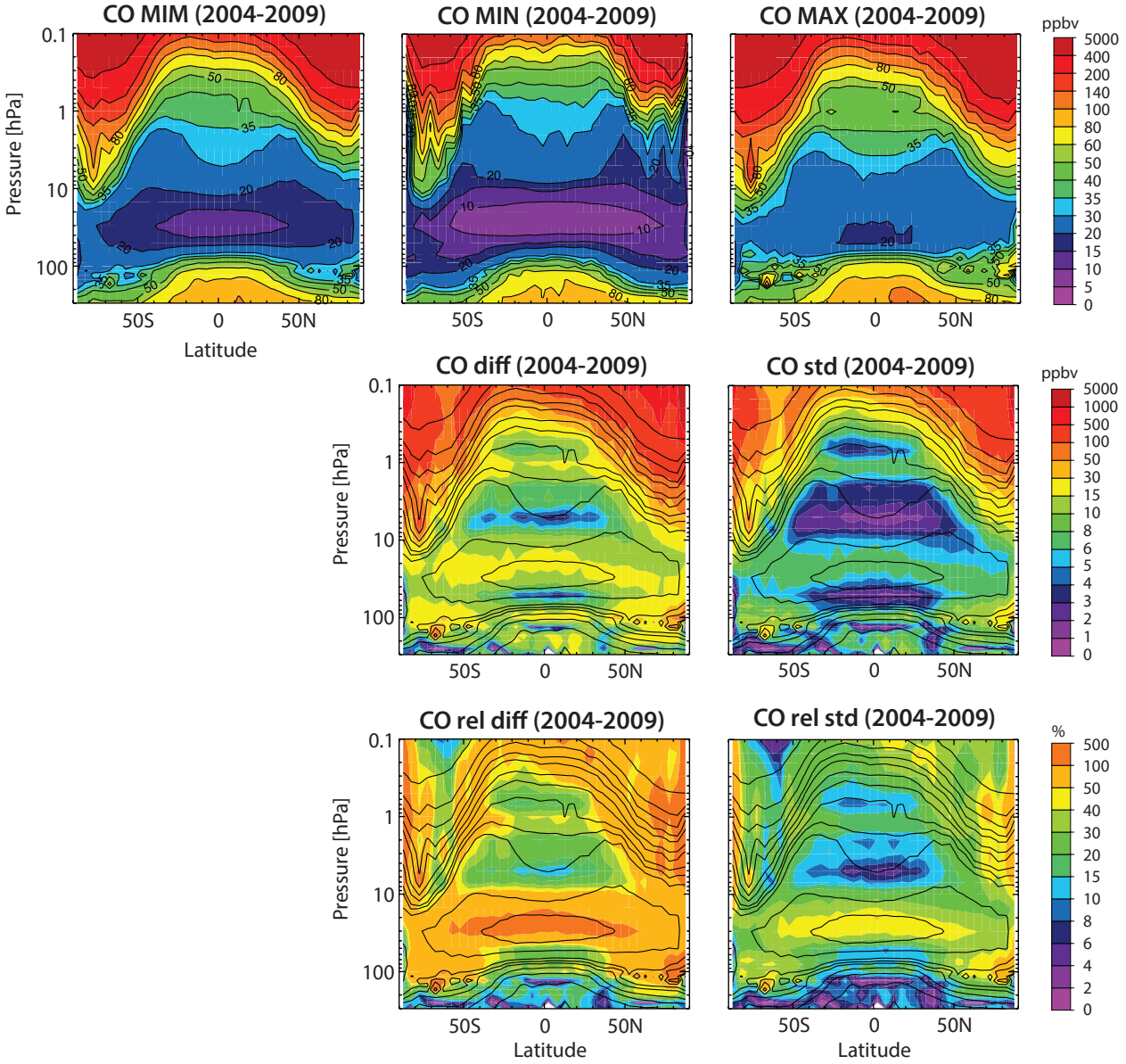


Figure 4.7.7: Summary of CO annual zonal mean state for 2002-2009. *Shown are the annual zonal mean cross-sections of the MIM, minimum (MIN), and maximum (MAX) CO values (upper row), the maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments (middle row), and 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.*

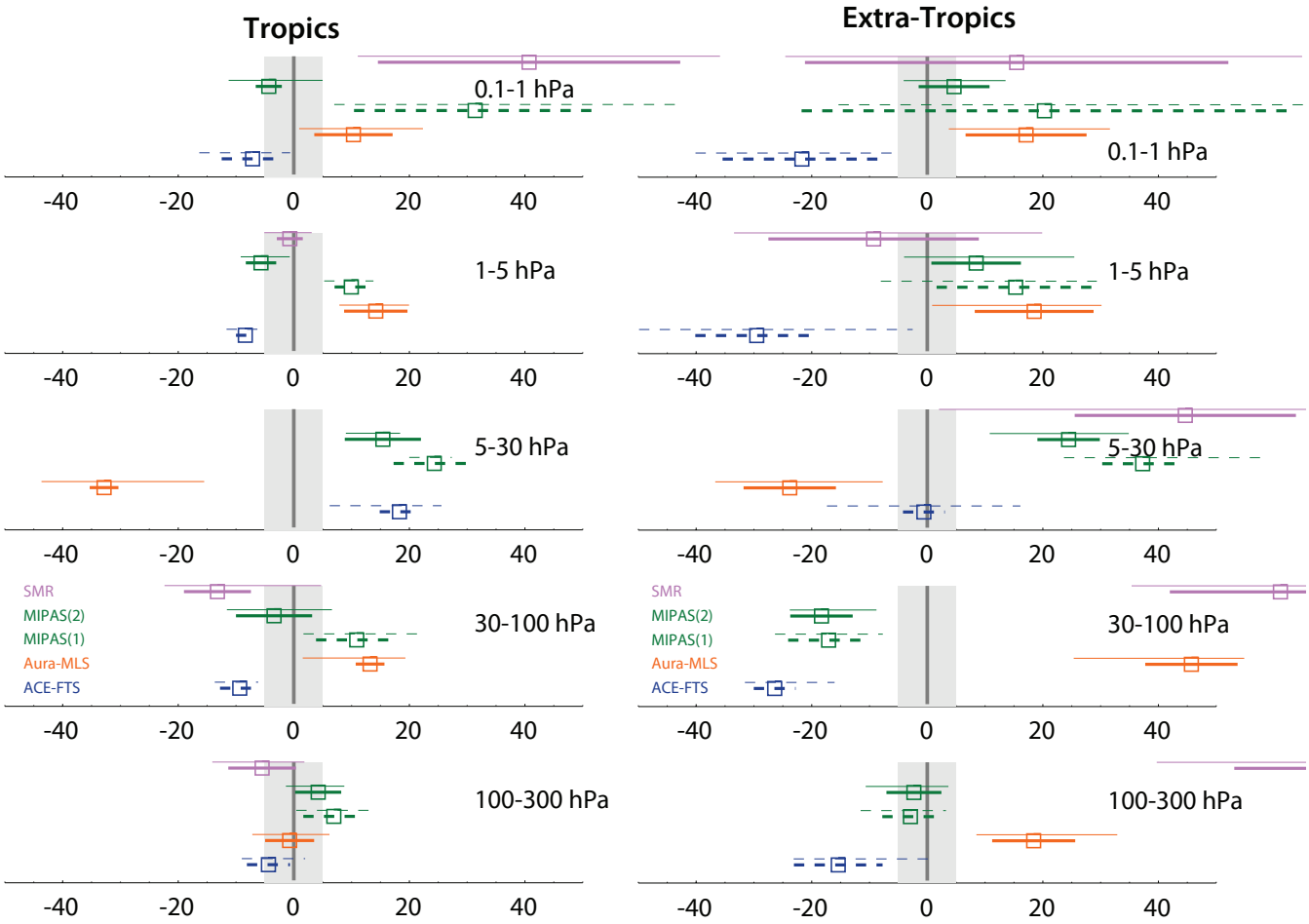


Figure 4.7.8: Inter-instrument differences in CO calculated for the tropics (left) (20°S–20°N) and extra-tropics (right) (40°S–80°S and 40°N–80°N), and for altitude regions from the UT up to the LM. *Shown are the median (squares), median absolute deviations (MAD, thick lines), and the mean ±1σ ranges (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.*

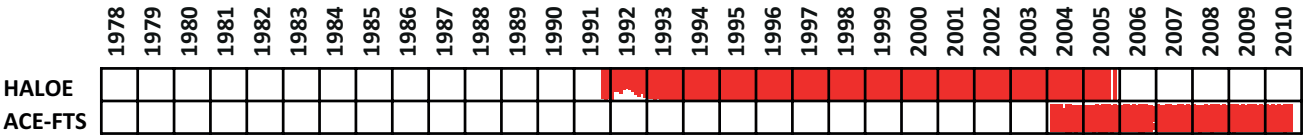


Table 4.8.2: Time period, vertical range, vertical resolution, references and other comments for HF measurements.

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	

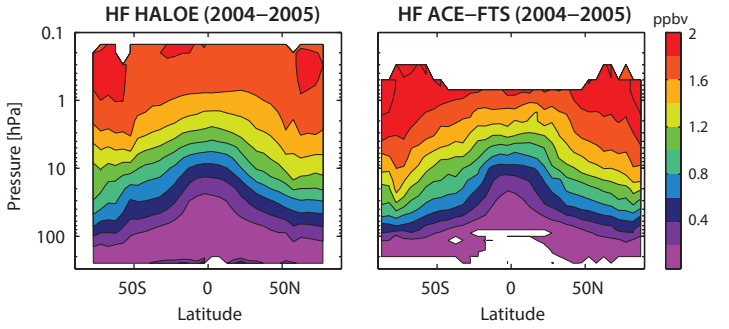


Figure 4.8.1: Cross sections of annual zonal mean HF for 2004-2005. *HF cross sections are shown for HALOE and ACE-FTS.*

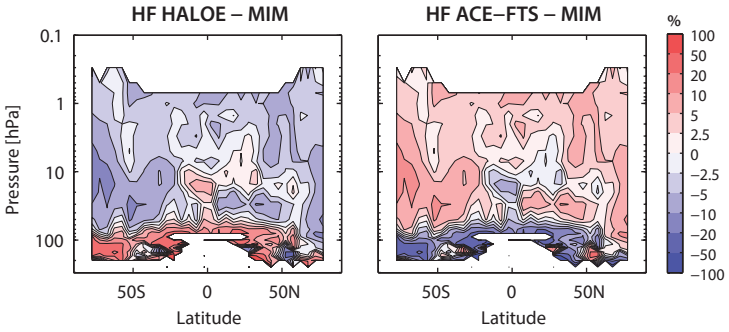


Figure 4.8.2: Cross sections of annual zonal mean HF differences for 2004-2005. *HF differences for HALOE and ACE-FTS with respect to the MIM are shown. Note that, since the MIM consists of only two instruments, any issue with one dataset will fully be reflected by the difference of the other dataset.*

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

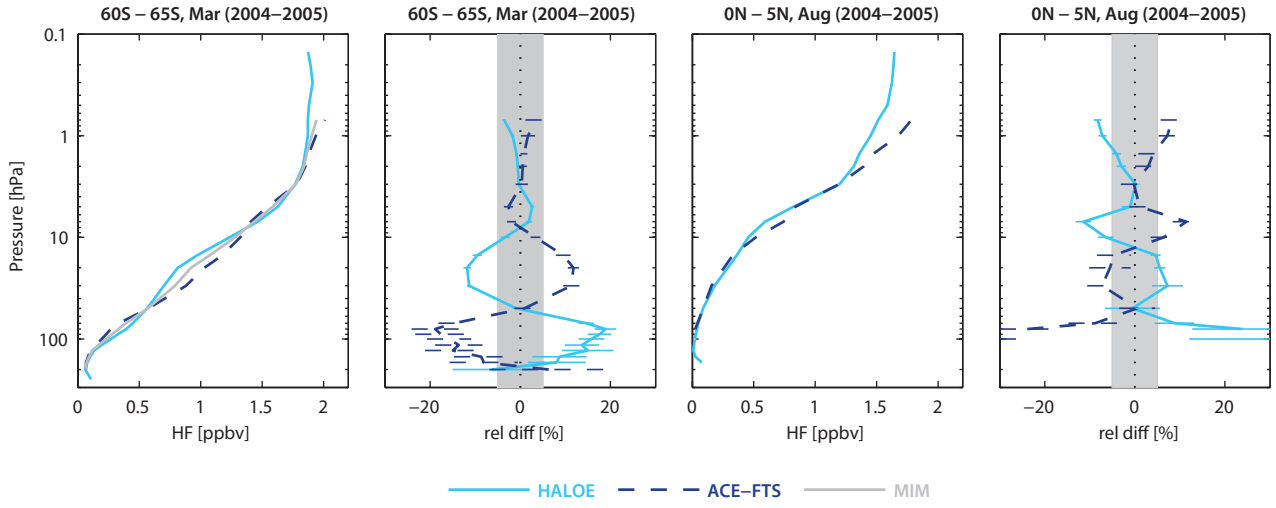


Figure 4.8.3: Vertical profiles of monthly zonal mean HF for 2004-2005. *Zonal mean HF profiles for 60°S-65°S, March (left panels) and 0°N-5°N, August (right panels) are shown together with their differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties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

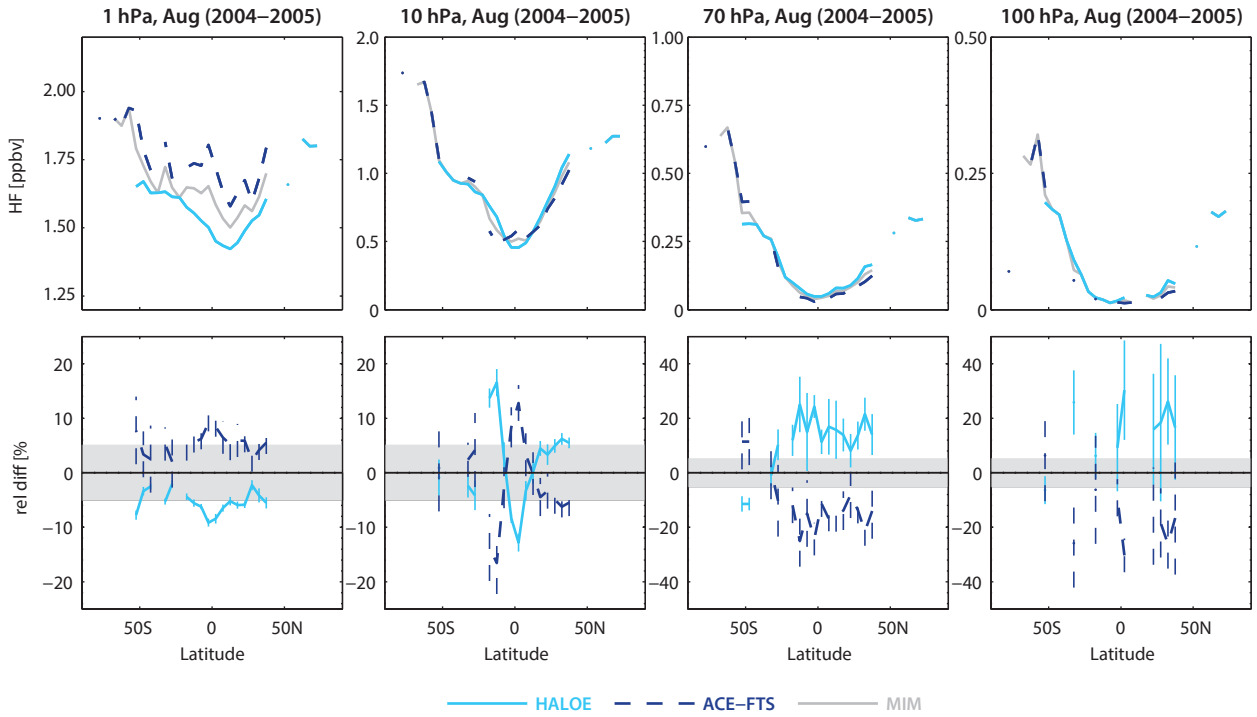


Figure 4.8.4: Meridional profiles of monthly zonal mean HF for 2004-2005. *HF profiles at 1, 10, 70 and 100 hPa for August are shown in the upper row. Differences of the individual instruments (HALOE and ACE-FTS) to the MIM are shown in the lower row. The grey shading indicates the ±5% difference range.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

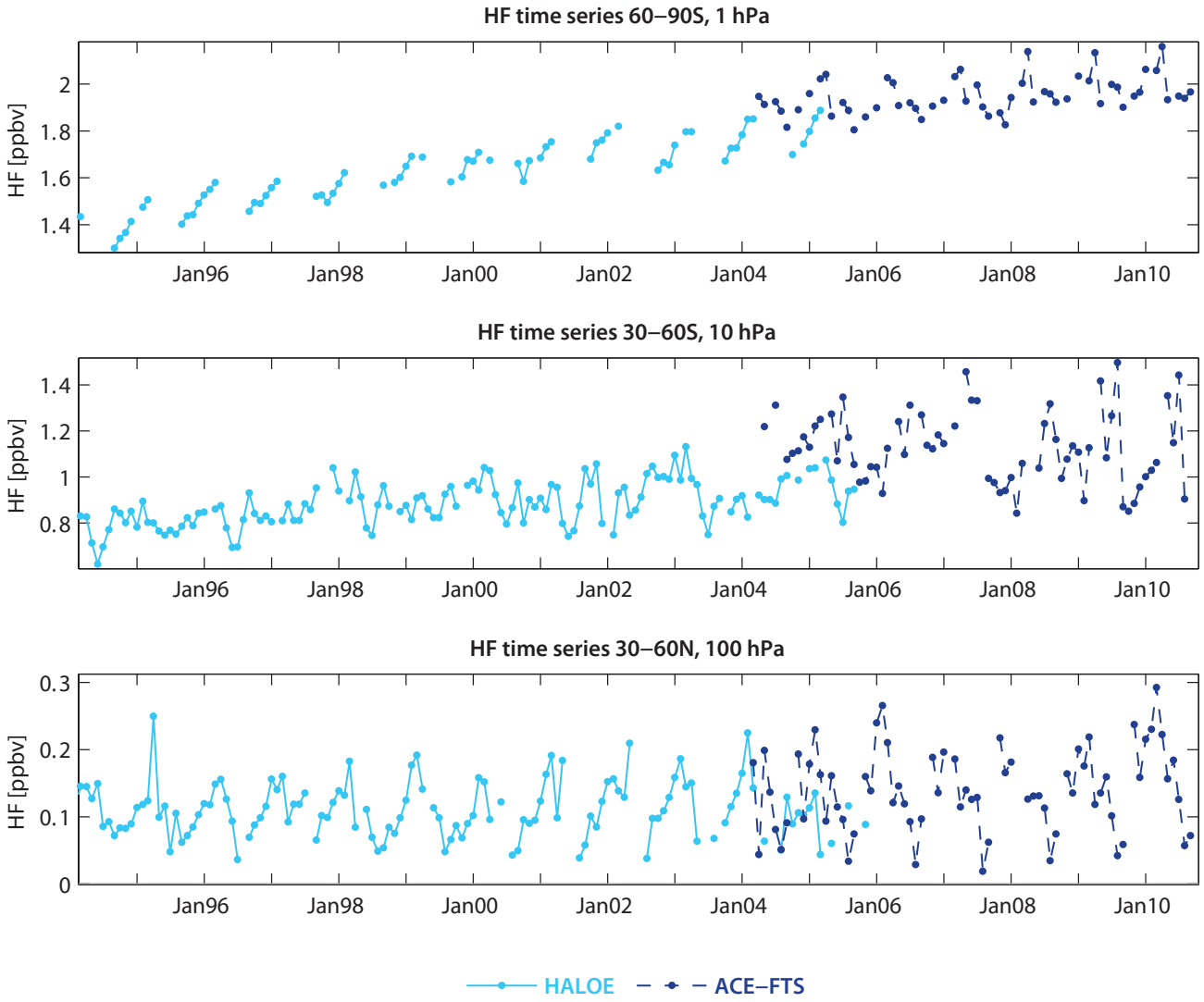


Figure 4.8.5: Time series of HF monthly mean values in mid- and high latitudes. *Monthly mean HF values between 60°S-90°S at 1 hPa (upper panel), 30°S-60°S at 10 hPa (middle panel), and 30°N-60°N at 100 hPa (lower panel).*

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

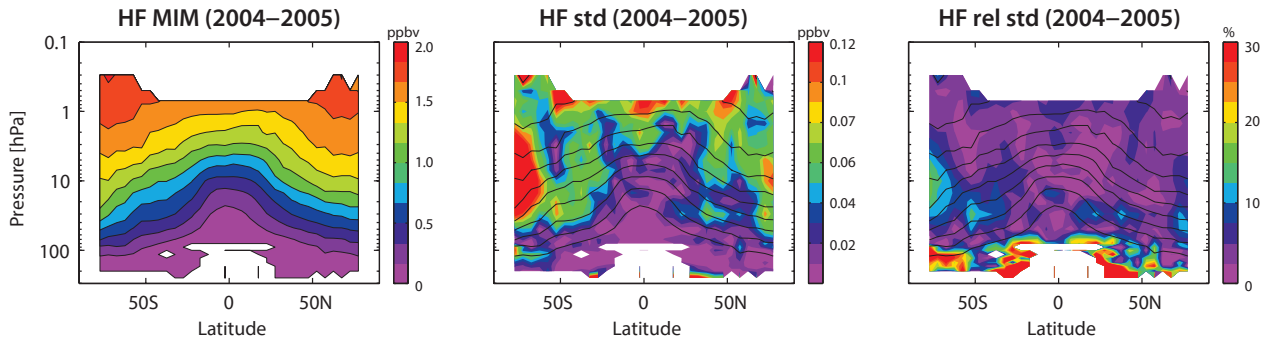


Figure 4.8.6: Summary of HF annual zonal mean state for 2004-2005. *Shown are the annual zonal mean cross section of the HF MIM (left panel), the standard deviation over both instruments (middle panel), and relative standard deviation with respect to the MIM (right panel). Black contour lines in the 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.*

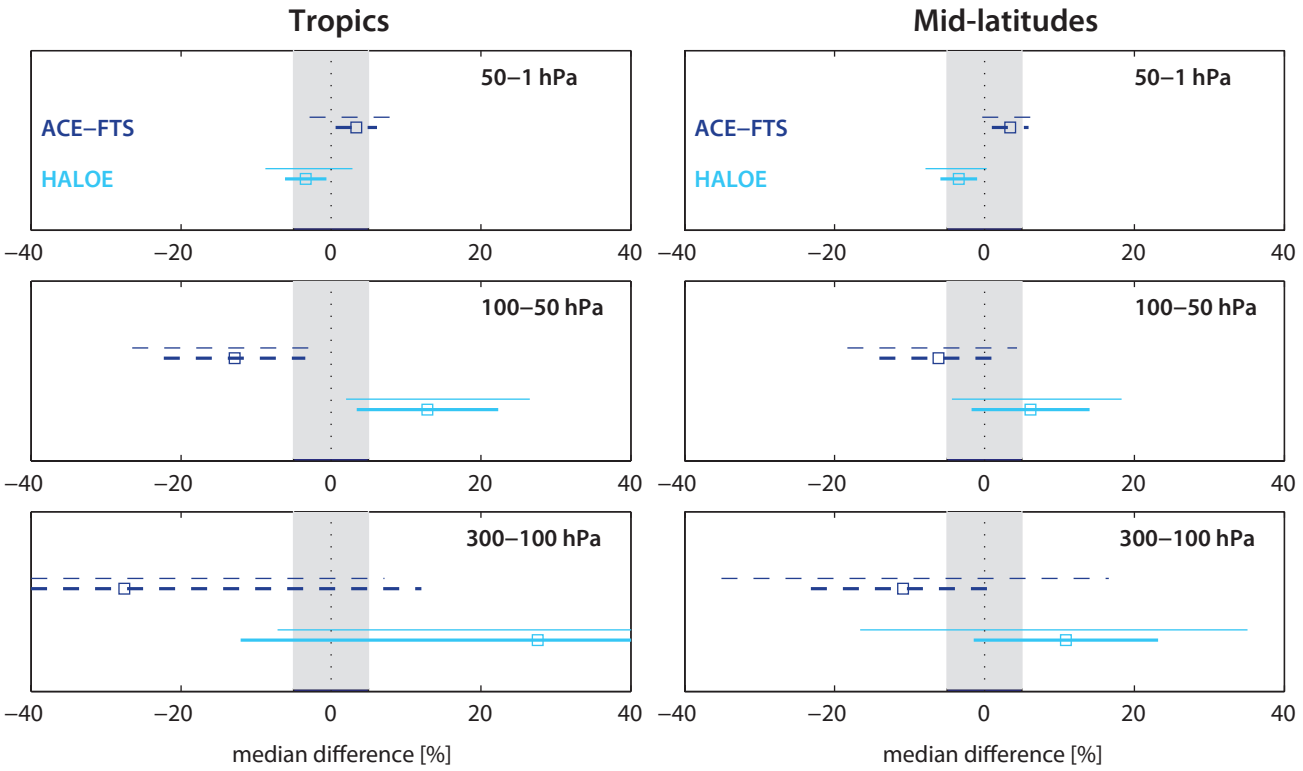


Figure 4.8.7: Summary HF differences for 2004-2005. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM 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.*

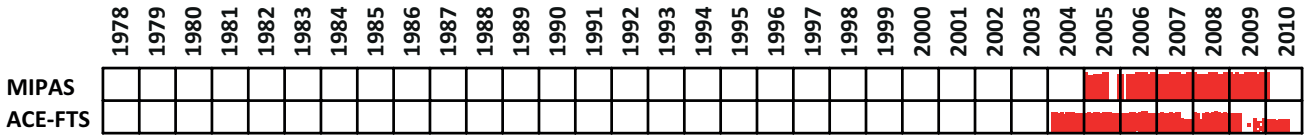


Table 4.9.2: Time period, vertical range, vertical resolution, references and other comments for SF_6 measurements.

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

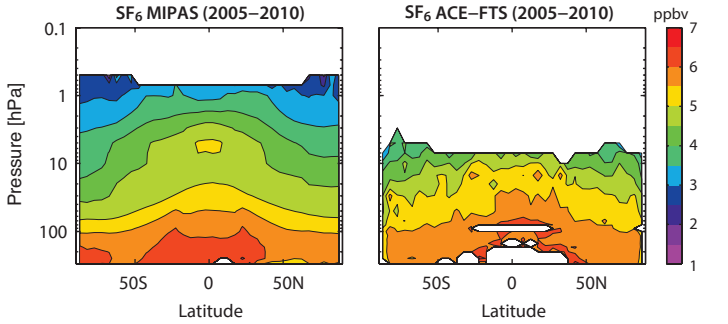


Figure 4.9.1: Cross sections of annual zonal mean SF_6 for 2005-2010. *SF_6 cross sections are shown for MIPAS and ACE-FTS.*

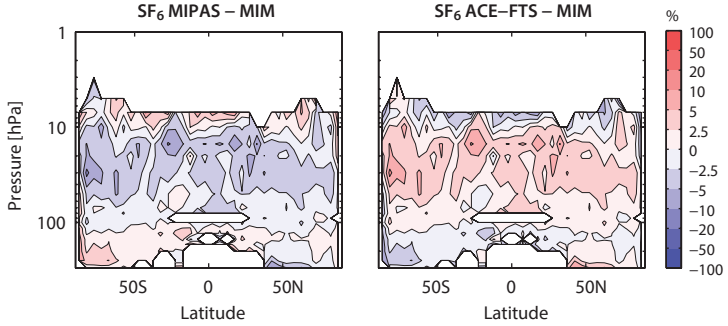


Figure 4.9.2: Cross sections of annual zonal mean SF_6 differences for 2005-2010. *SF_6 differences between MIPAS, ACE-FTS and their MIM are shown. Note that, since the MIM consists of only two instruments, any issue with one dataset will fully be reflected by the difference of the other dataset.*

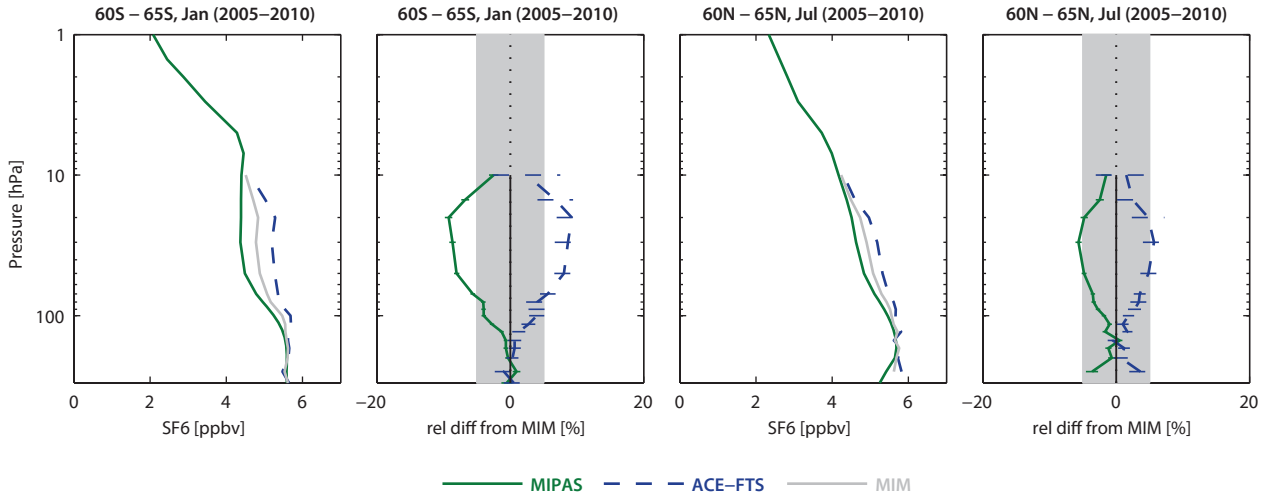


Figure 4.9.3: Vertical profiles of monthly zonal mean SF_6 for 2005-2010. *Zonal mean SF_6 profiles for 60°S-65°S, January (left panels) and 60°N-65°N, July (right panels) are shown together with their differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties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

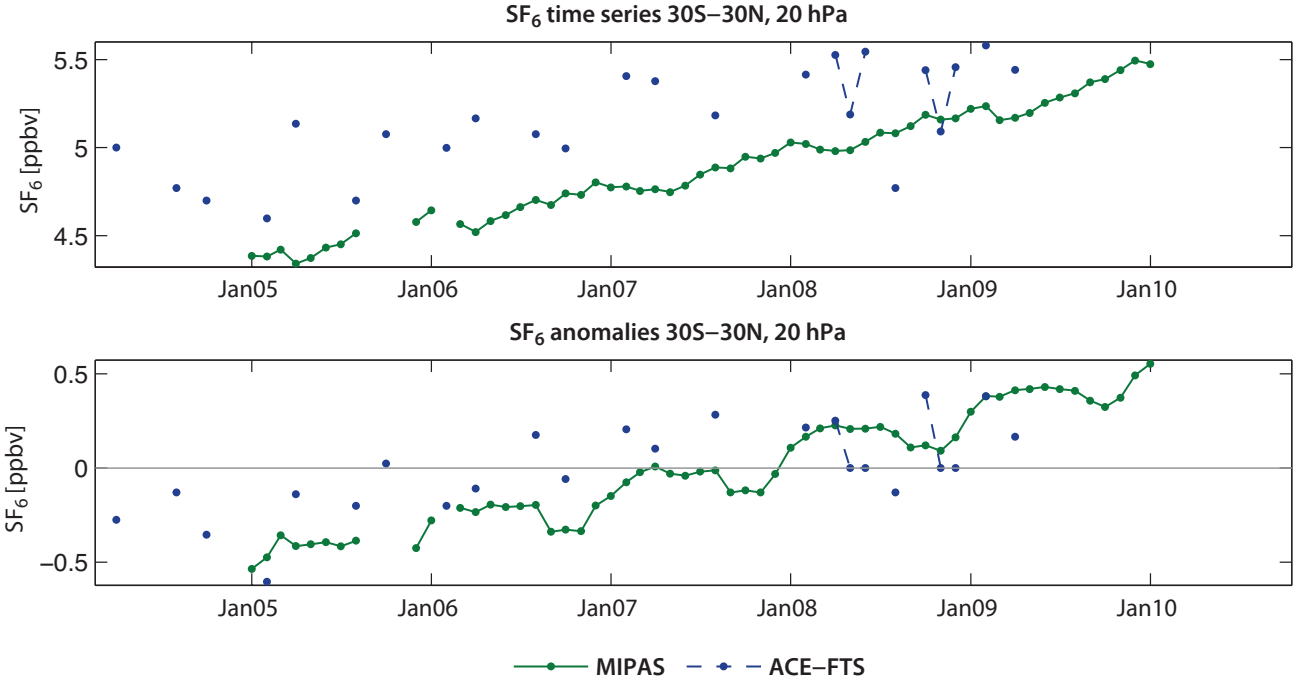


Figure 4.9.4: Time series of SF_6 monthly mean values and deseasonalised anomalies in the tropics. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of SF_6 between 30°S – 30°N at 20 hPa.*

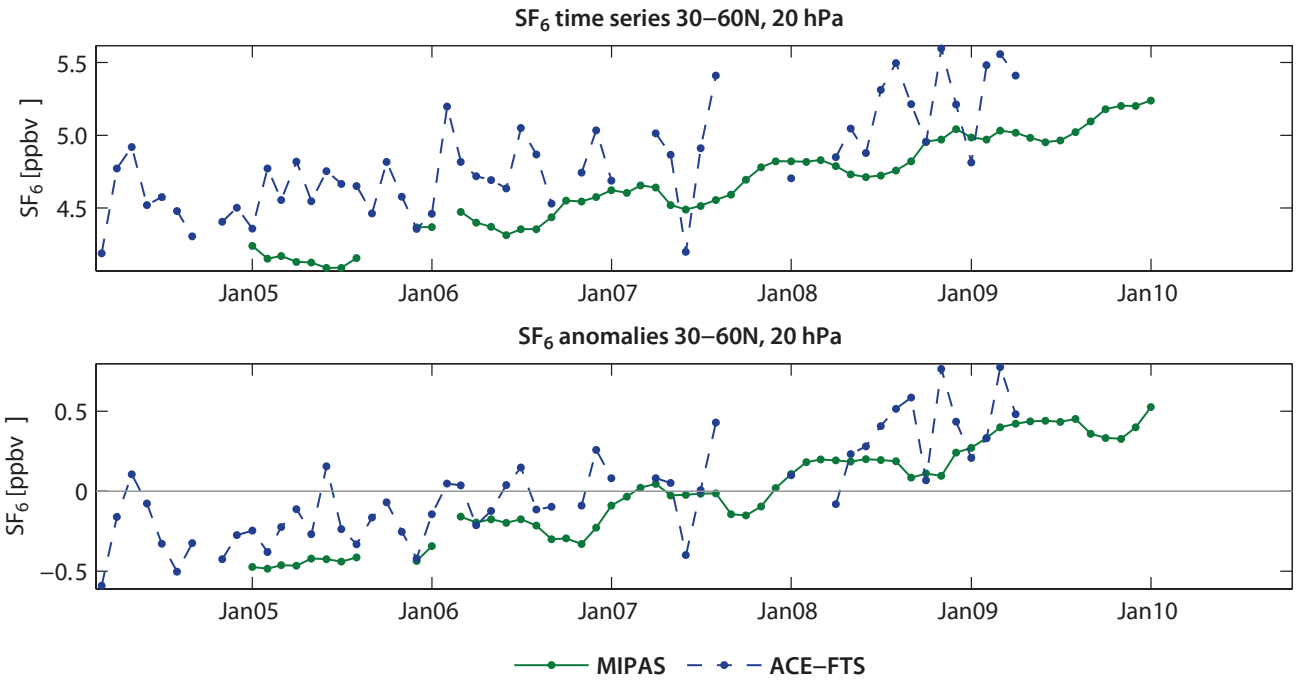


Figure 4.9.5: Time series of SF_6 monthly mean values and deseasonalised anomalies in the NH mid-latitudes. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of SF_6 between 30°N – 60°N at 20 hPa.*

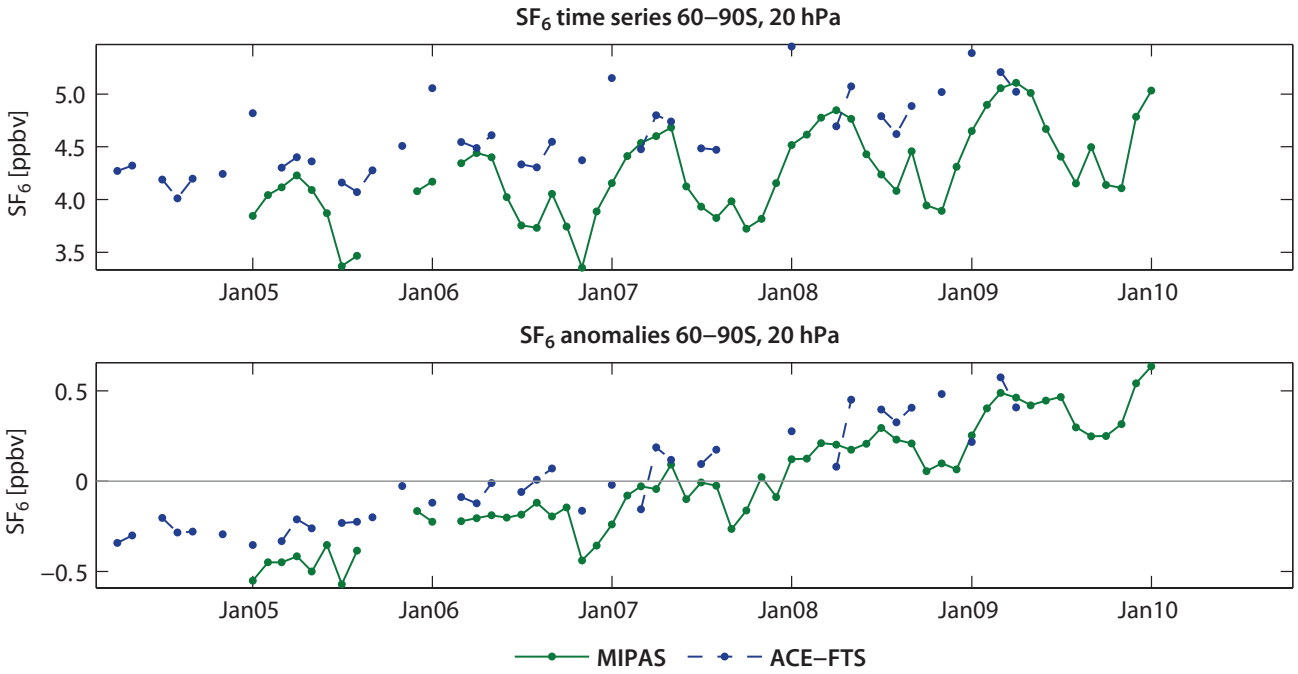


Figure 4.9.6: Time series of SF_6 monthly mean values and deseasonalised anomalies at SH high latitudes. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of SF_6 between 60°S – 90°S at 20 hPa.*

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

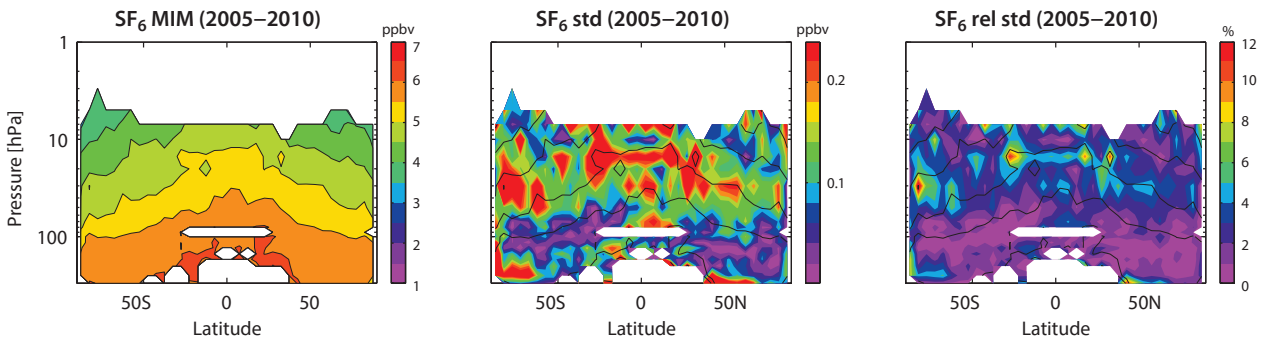


Figure 4.9.7: Summary of SF_6 annual zonal mean state for 2005-2010. *Shown are the annual zonal mean cross section of the SF_6 MIM (left panel), the standard deviation over both instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black contour lines 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.*

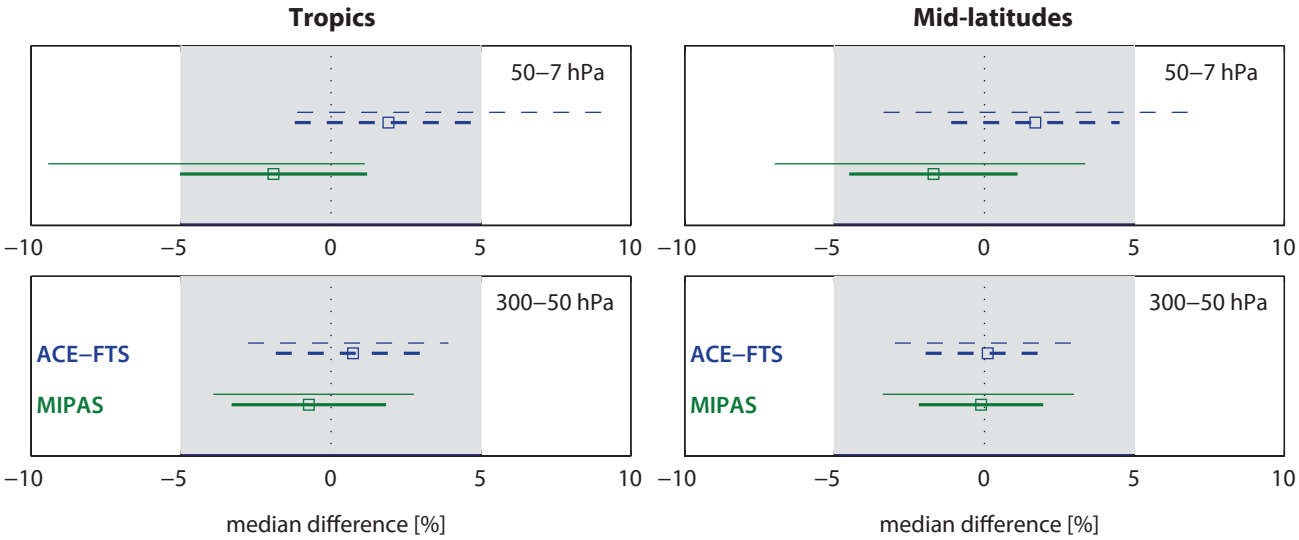


Figure 4.9.8: Summary SF_6 differences for 2005-2010. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM 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

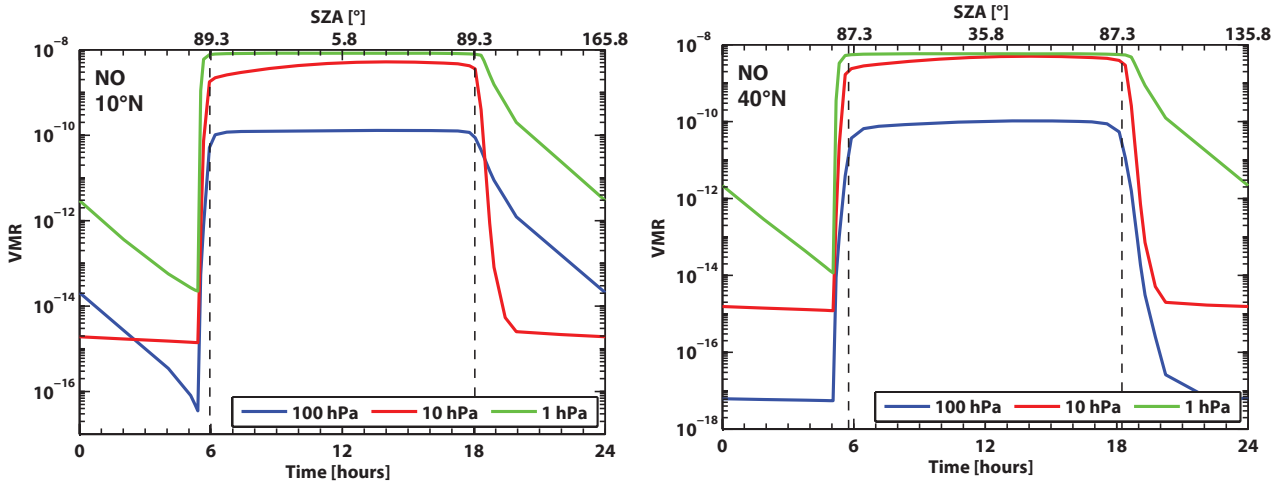
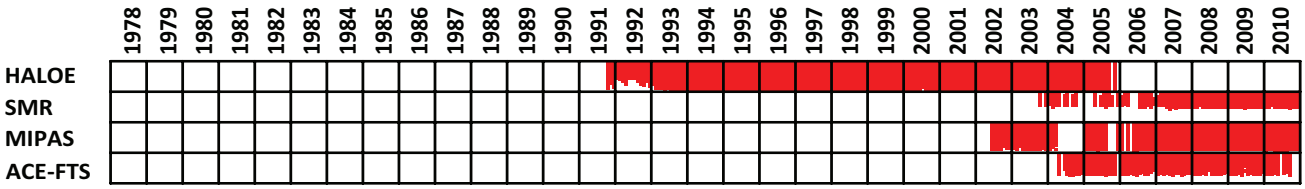


Figure 4.10.1: Diurnal NO cycle. *NO variations as function of LST are shown at 10°N and 40°N at 1, 10, and 100 hPa for March 15.*

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

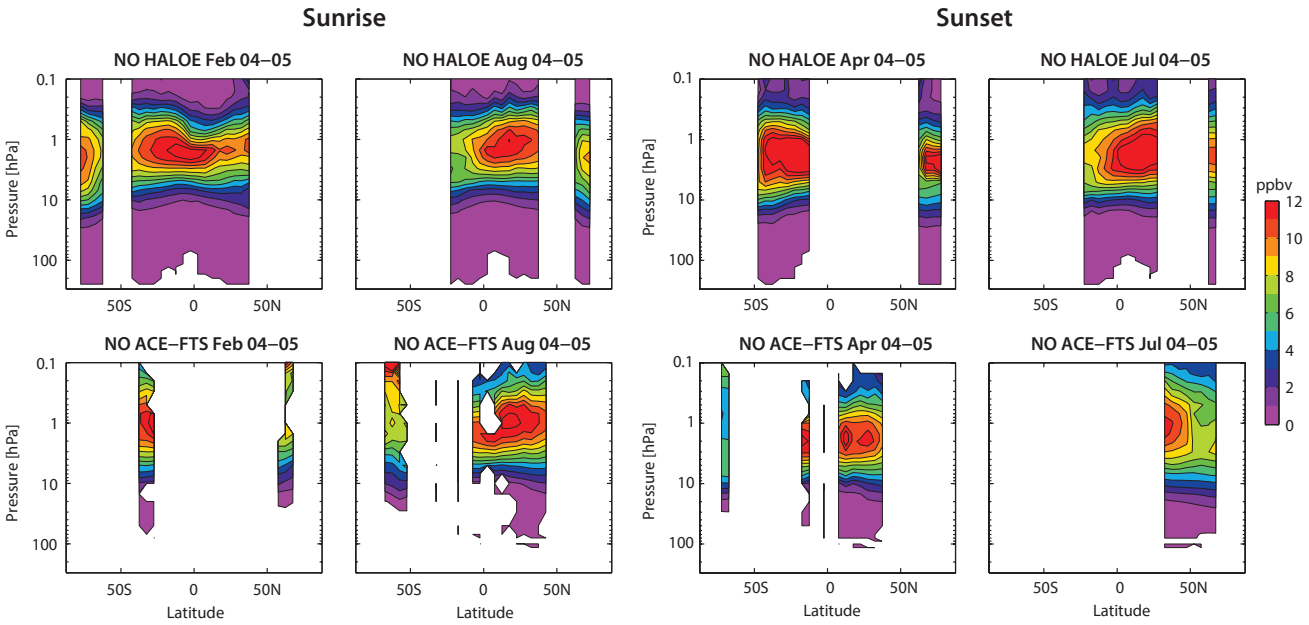


Figure 4.10.2: Cross sections of monthly zonal mean, local sunrise and sunset NO for 2004-2005. *Monthly zonal mean, local sunrise for February and August (column 1 and 2) and local sunset for April and July (column 3 and 4) NO cross sections are shown for HALOE (upper row) and ACE-FTS (lower row).*

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

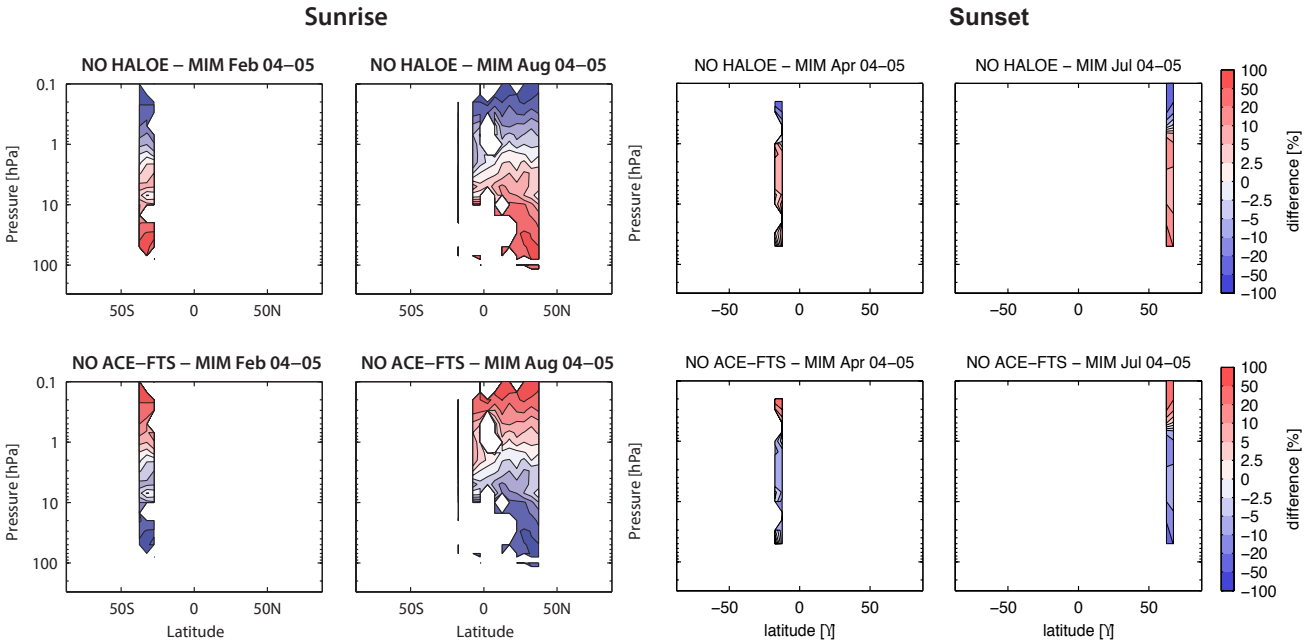


Figure 4.10.3: Cross sections of monthly zonal mean, local sunrise and sunset NO differences for 2004-2005. *Monthly zonal mean, local sunrise for February and August (column 1 and 2) and local sunset for April and July (column 3 and 4) NO differences between the individual instruments (HALOE and ACE-FTS)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

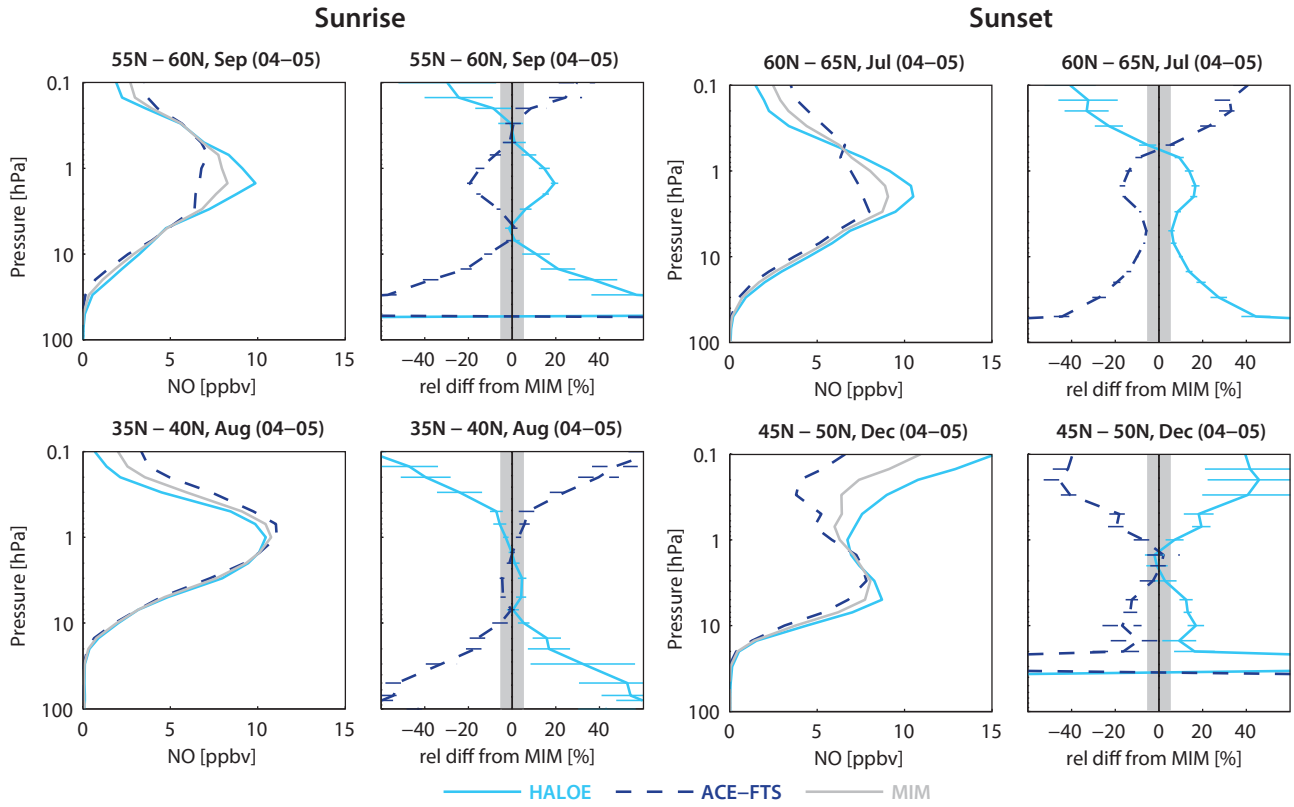


Figure 4.10.4: Profiles of monthly zonal mean, local sunrise and sunset NO for 2004-2005. *Zonal mean NO profiles are shown together with their differences from the MIM for local sunrise, 55°N-60°N, September and 35°N-40°N, August (column 1 and 2) and local sunset, 60°N-65°N, July and 45°N-50°N, December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

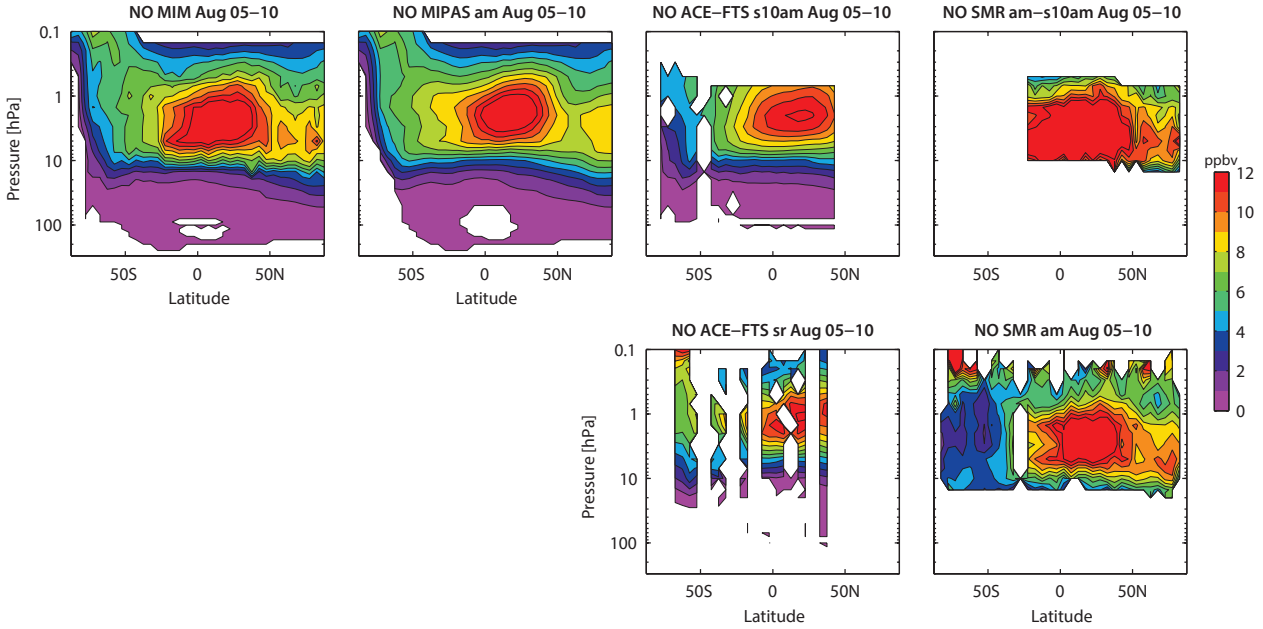


Figure 4.10.5: Cross sections of monthly zonal mean 10am NO for 2005-2010. *Monthly zonal mean 10am NO cross sections for August 2005-2010 are shown for the MIM, MIPAS (corresponding to 10am), ACE-FTS scaled to 10am (s10AM), and SMR am scaled to 10am data (am-s10am). Additionally, ACE-FTS local sunrise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

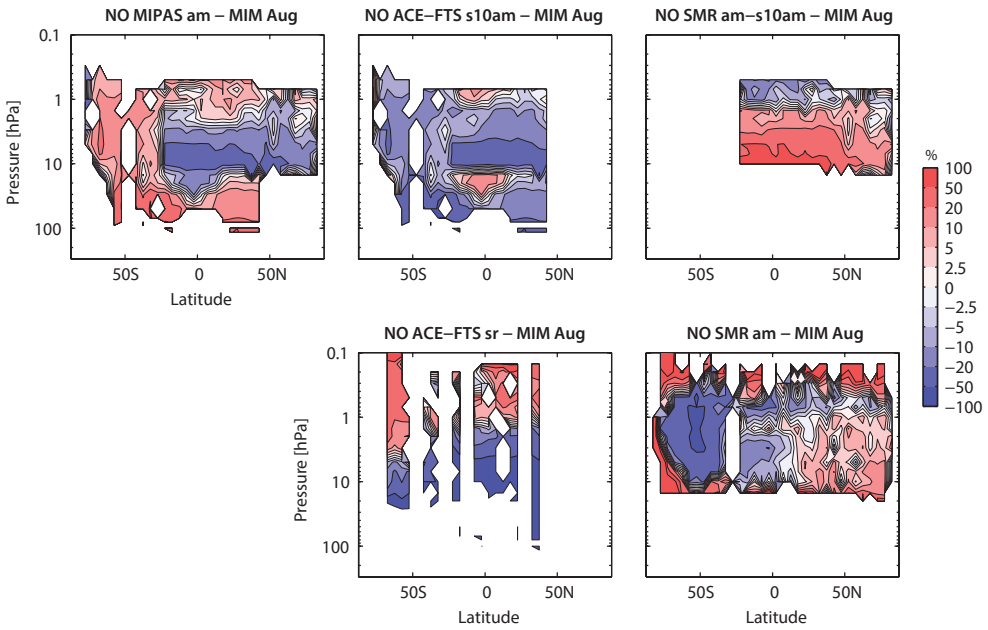


Figure 4.10.6: Cross sections of monthly zonal mean, 10am NO differences for 2005-2010. *Monthly zonal mean 10am NO differences for August 2005-2010 of MIPAS (corresponding to 10am), ACE-FTS scaled to 10am, and SMR am scaled to 10am with respect to their MIM are shown. Additionally, differences of ACE-FTS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

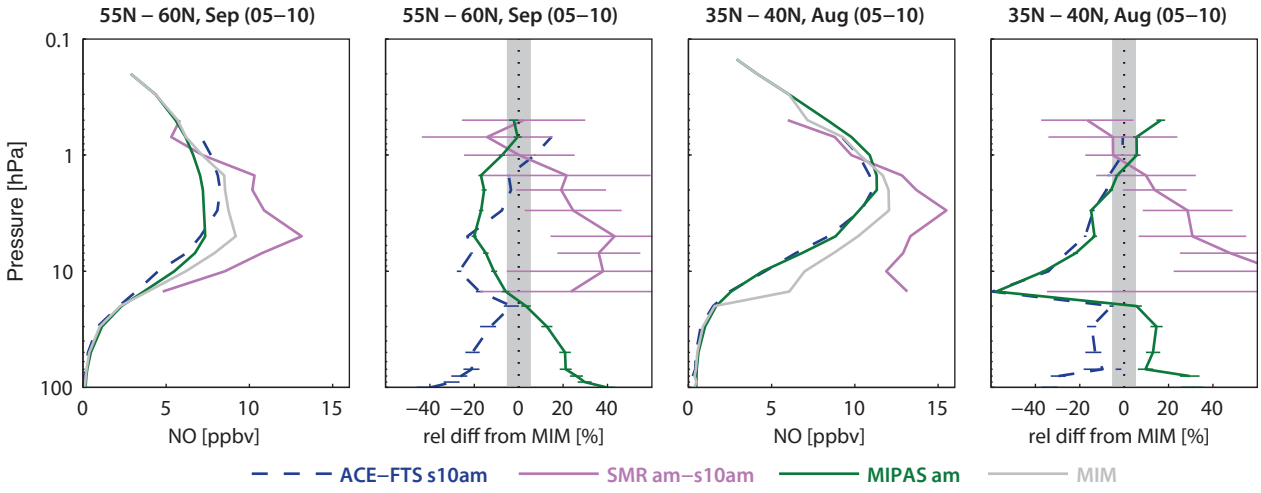


Figure 4.10.7: Vertical profiles of monthly zonal mean 10am NO for 2005-2010. *Zonal mean 10am NO profiles for 55°N-60°N, September and 35°N-40°N, August are shown together with their differences from the MIM.*

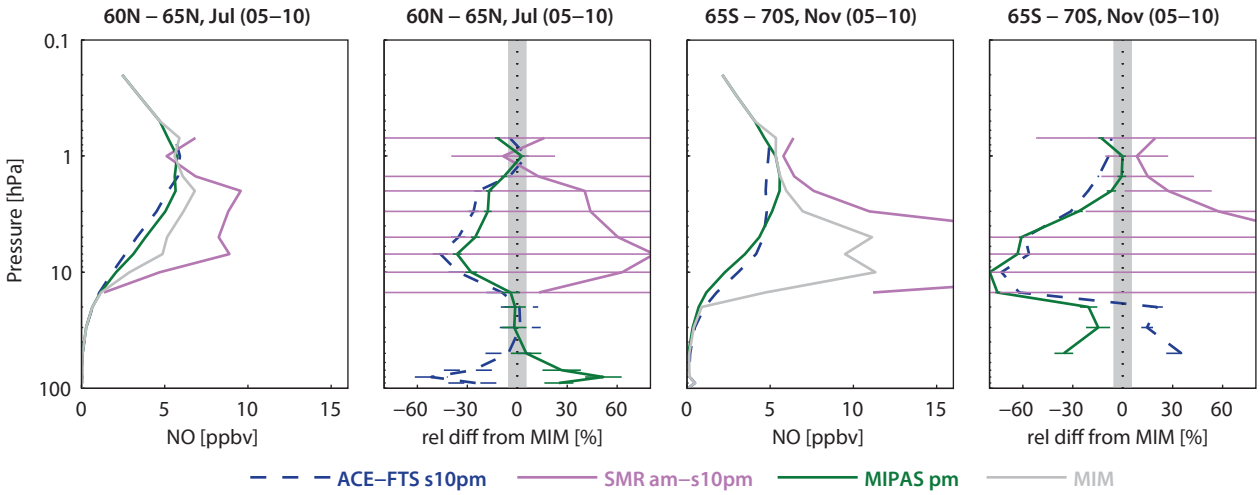


Figure 4.10.8: Vertical profiles of monthly zonal mean 10pm NO for 2005-2010. *Zonal mean 10pm NO profiles for 60°N-65°N, July and 65°S-70°S, November are shown together with their differences from the MIM.*

(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**

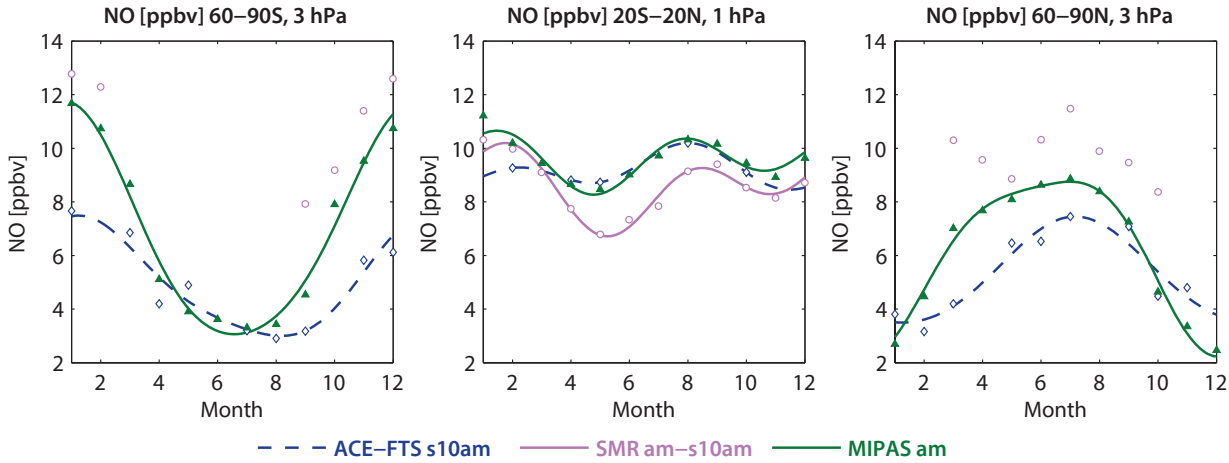


Figure 4.10.9: Seasonal cycle of 10am NO for 2005-2010. *Seasonal cycle of monthly zonal mean NO for 60°S-90°S, 3 hPa (left column), 20°S-20°N, 1hPa (middle column) and 60°N-90°N, 3 hPa (right column). Measurements correspond to 10am LST (MIPAS, filled symbols) or are scaled to 10am LST (ACE-FTS, SMR,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

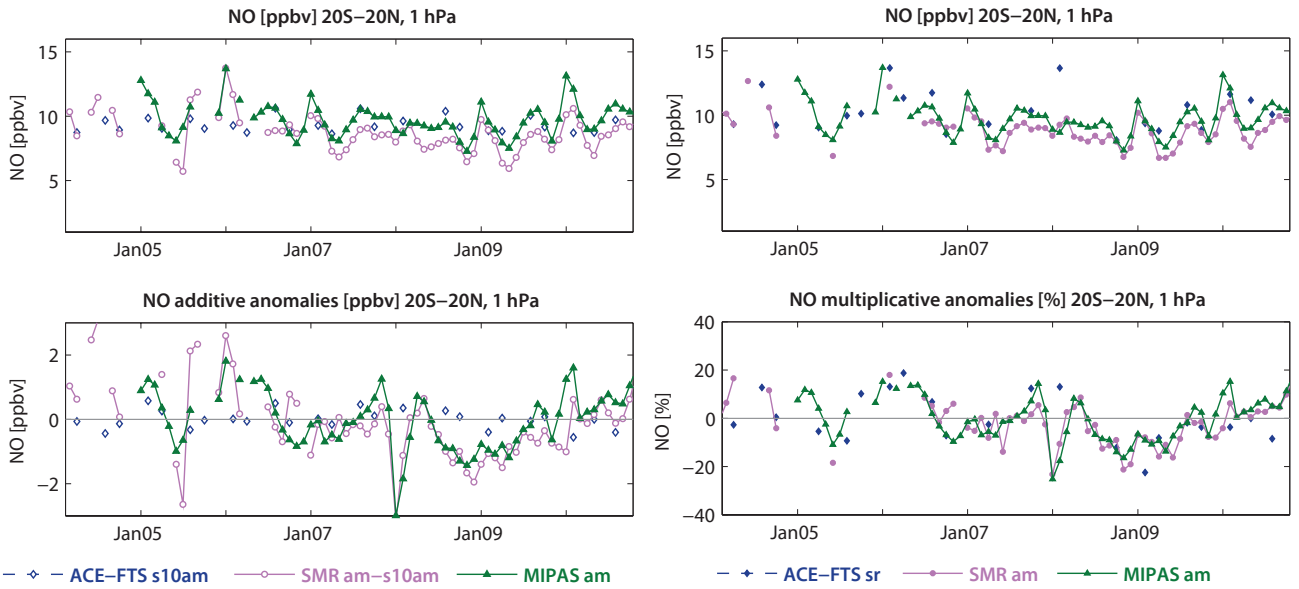


Figure 4.10.10: Time series of tropical NO mean values and anomalies for 2005-2010. *Monthly mean values (upper panels) and deseasonalised anomalies (lower panels) of NO between 20°S – 20°N at 1 hPa. The 10am climatologies (left panel) correspond directly to 10am LST (filled symbols) or are scaled to 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%.

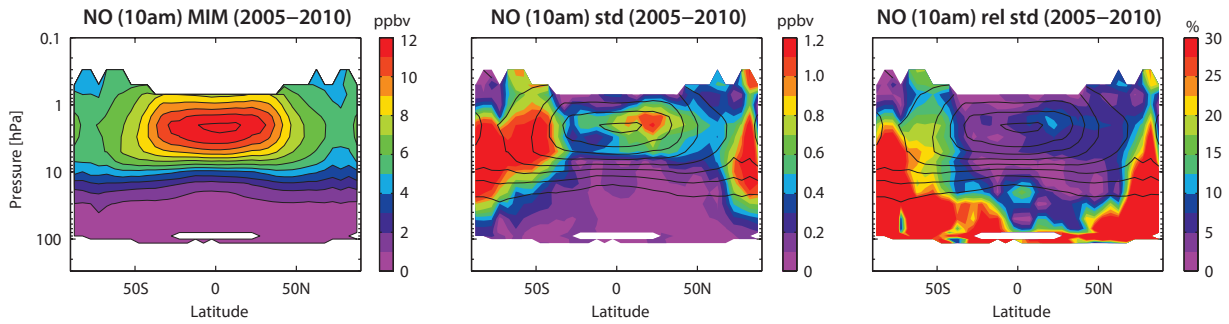


Figure 4.10.11: Summary of NO annual zonal mean state for 2005-2010. *Annual zonal mean cross section of the NO MIM for 10am is shown in the left panel. The NO mean state at 10am is based on MIPAS at 10am and ACE-FTS scaled to 10am with a chemical box model. Additionally, the standard deviation over 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%).

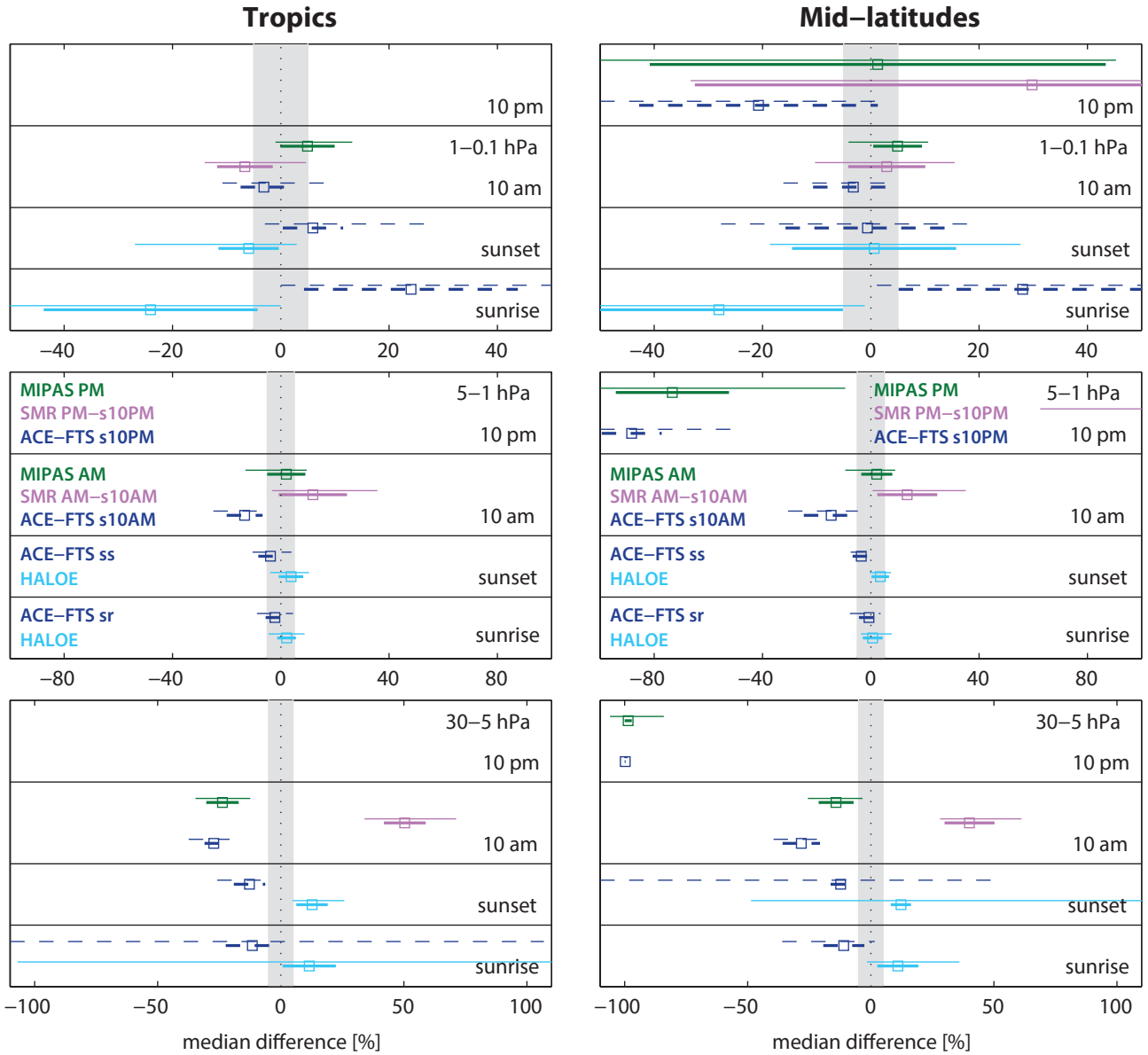


Figure 4.10.12: Summary NO differences for 2005-2010. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM 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.*

4.11 Nitrogen dioxide – NO_2

Nitrogen dioxide (NO_2) is a major air pollutant in the troposphere, and is produced mainly by fossil fuel burning (*via* production of NO, see *Section 4.10*). Natural sources of tropospheric NO_2 involve lightning, bacterial processes in soil and water, and biomass burning. The main source of NO_2 in the stratosphere is the oxidation of N_2O, also originating from soil emissions (see *Section 4.4*), followed by downward transport of upper atmospheric NO_x in the winter polar vortex (up to 9%). Stratospheric NO_2 is considered an ozone depleting substance through the catalytic NO_x cycle. At the same time, NO_2 acts as a buffer against halogen-catalyzed ozone loss by converting reactive chlorine, bromine, and hydrogen compounds into stable reservoir substances ($ClONO_2$, $BrONO_2$, HNO_3). The removal of nitrogen from the stratosphere, as observed in the SH polar vortex during the formation and sedimentation of polar stratospheric clouds (PSCs), is a key microphysical process in the formation of the Antarctic polar ozone hole during spring.

Stratospheric NO_2 variations are controlled by the sunlight-driven equilibrium between NO_x (NO, NO_2) on one hand and the reservoir substances (N_2O_5, HNO_3, $ClONO_2$) on the other hand. In particular, the strong diurnal NO_2 cycle complicates a comparison of satellite-based NO_2 measurements that correspond to different LSTs. **Figure 4.11.1** shows examples of the diurnal NO_2 cycle as a function of LST for three different pressure levels as derived with a chemical box model [*McLinden et al.*, 2010]. As seen in these plots there are large differences between day and night-time NO_2, with a steep gradient at sunrise and sunset. Solar occultation measurements are always made at SZA = 90° and can therefore be compared amongst each other. Limb scattering and emission measurements however, can correspond to different SZAs, and a direct comparison of the climatologies is not meaningful in most cases unless the dependence on the SZA is taken into account.

4.11.1 Availability of NO_2 measurements

The first vertically resolved satellite NO_2 measurements were made by LIMS in 1978/1979. SAGE II and HALOE provide the longest continuous NO_2 datasets, both ending in 2005. A number of more recent satellite missions have been measuring NO_2 from 2002 onwards. Solar occultation measurements are available from SAGE II, HALOE, POAM II, POAM III, SAGE III and ACE-FTS. Periods of maximum overlap are 1994-1996 (SAGE II, HALOE, and POAM II) and 2005 (SAGE II, HALOE, POAM III, SAGE III, and ACE-FTS). NO_2 measurements by limb emission and scattering techniques are available for OSIRIS, SCIAMACHY, and MIPAS from 2002 onward, while GOMOS provides stellar occultation measurements for the same time period. Additionally, HIRDLS data for 2005-2007 and LIMS data for 1978-1979 exist. **Table 4.11.1** summarises information on the available NO_2 measurement record including time period and vertical range.

The solar occultation climatologies can be compared directly if separated into local sunrise and local sunset measurements. Note that there is a difference between the sunrise/sunset as seen from a satellite and the local sunrise/sunset that determines the chemical state of the measured air mass, and is therefore used to categorise the measurements. The corresponding data files are labelled as am/pm with am LST generally corresponding to local sunrise and pm LST generally corresponding to local sunset. One deviation from the photochemical conditions at local sunrise can occur when a satellite crosses the terminator towards the polar day area at high latitudes. During such observations, the photochemical conditions of the atmosphere at local surise may be completely different than that of a typical sunrise observation because the area is continuously illuminated during polar day. The same situation can occur when the satellite crosses the terminator towards polar night. **Table 4.11.2** summarises the local sunrise/sunset climatologies.

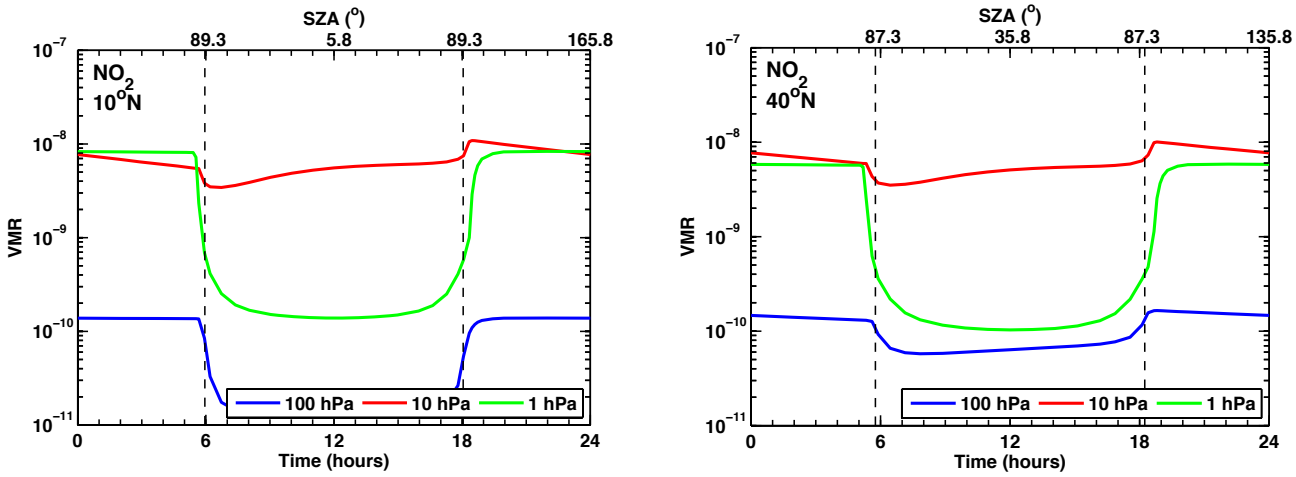


Figure 4.11.1: Diurnal NO_2 cycle. *NO_2 variations as function of LST are shown at 10°N and 40°N at 1, 10 and 100 hPa for March 15.*

NO_2 measurements by limb emission and scattering techniques correspond approximately to a fixed LST if the instrument has a sun-synchronous orbit (*e.g.*, MIPAS measurements correspond to 10am and 10pm LST). The measurement LST can vary from instrument to instrument, and may also vary for some instruments between seasons and latitudes. In order to compare other instruments (OSIRIS, SCIAMACHY, ACE-FTS) with MIPAS, their measurements have been scaled to LST 10am and 10pm with the help of a chemical box model [*McLinden et al.*, 2010].

Table 4.11.1: Available NO_2 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. Instruments are grouped according to their measurement LST into the group of solar occultation instruments (upper panel) and the group of limb emission/scattering and stellar occultation instruments (lower panel).*

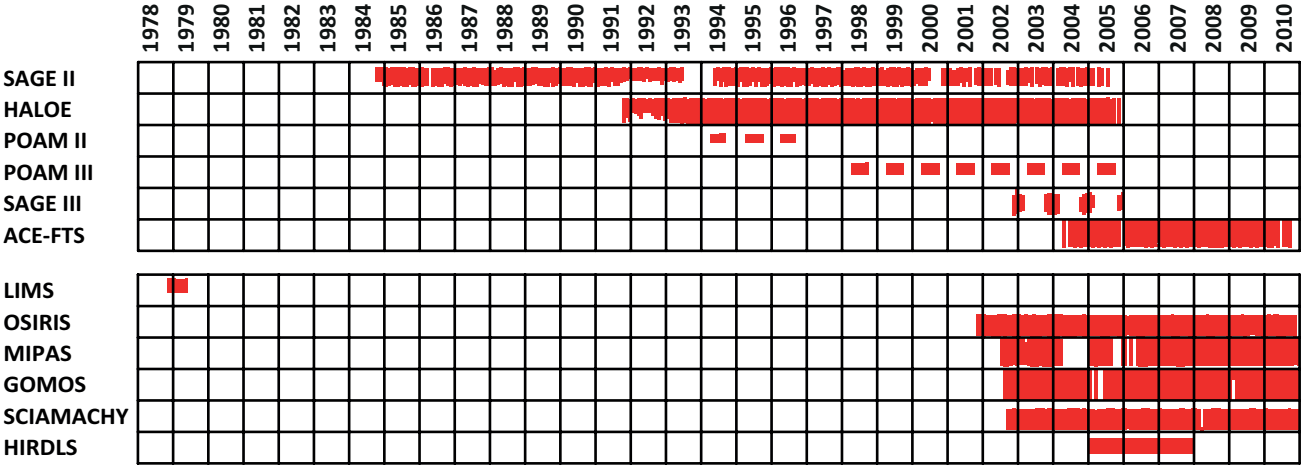


Table 4.11.2: Overview of available NO_2 local sunrise and local sunset climatologies. *For local sunrise/sunset climatologies, the instrument name, the method used to derive the climatology, the corresponding LST and the satellite orbit are given. Detailed information on the corresponding LST depending on month and latitude can be found in the data files.*

Instrument	**Local sunrise climatologies**		**Local sunset climatologies**		Satellite orbit
	Method	Corresponding LST	Method	Corresponding LST	
SAGE II	**Local sunrise**	am (mixed times)	**Local sunset**	pm (mixed times)	non-sun-synchronous
HALOE	**Local sunrise**	am (mixed times)	**Local sunset**	pm (mixed times)	non-sun-synchronous
POAM II	**Local sunrise**	am (mixed times)	**Local sunset**	pm (mixed times)	sun-synchronous
POAM III	**Local sunrise**	am (mixed times)	**Local sunset**	pm (mixed times)	sun-synchronous
SAGE III	**Local sunrise**	am (mixed times)	**Local sunset**	pm (mixed times)	sun-synchronous
ACE-FTS	**Local sunrise**	am (mixed times)	**Local sunset**	pm (mixed times)	non-sun-synchronous

Table 4.11.3: Overview of available NO_2 daytime and night-time climatologies including the 10am/pm climatologies. *For daytime and night-time climatologies, the instrument name, the method used to derive the climatology, i.e., scaling measurements or sorting measurements according to LST, orbit or SZA, and the corresponding LST range are given. Climatologies corresponding to 10am and 10pm LST are printed in bold face. Detailed information on the corresponding LST depending on month and latitude can be found in the data files.*

Instrument	**Daytime climatologies**		**Night-time climatologies**	
	Method	Corresponding LST	Method	Corresponding LST
LIMS	Ascending orbit	~1pm	Descending orbit	~11pm
OSIRIS	am LST	Mixed times	pm LST	Mixed times
	Scaled	**10am**	**Scaled**	**10pm**
SCIAMACHY	No adjustments	~8am - 2pm		
	Scaled	**10am**	**Scaled**	**10pm**
MIPAS	**am LST**	**10am**	**pm LST**	**10pm**
GOMOS			**No adjustments**	**10pm**
HIRDLS	SZA < 90°	~3pm	SZA > 90°	00:30am
	Scaled	**10am**	**Scaled**	**10pm**
ACE-FTS	**Scaled**	**10am**	**Scaled**	**10pm**

For OSIRIS, scaling is done profile-by-profile, with the initialisation of the model using the measured trace gas abundances and temperature. The scaling for SCIAMACHY and ACE-FTS on the other hand, is done based on lookup tables calculated from the photochemical box model initialised with climatological inputs (see *Section 3.2.1*). GOMOS provides stellar occultation measurements at 10pm. These are evaluated with the group of limb emission and scattering instruments. The ACE-FTS climatology derived from data scaled to 10am/pm provides an opportunity to compare one solar occultation dataset with the measurements based on emission and scattering techniques. HIRDLS data from June 2005 to May 2006 have been scaled to 10am/pm with the SD-WACCM Version 3548 (SD stands for Specified Dynamics using GEOS5.1) [*Garcia et al.*, 2007]. **Table 4.11.3** summarises all available daytime and night-time NO_2 climatologies including the original available daytime and night-time climatologies and the ones scaled to 10am/pm LST. All instruments listed in **Table 4.11.3**, except for ACE-FTS, are in sun-synchronous orbits. **Table 4.11.4** compiles information on all NO_2 measurements, including time period, vertical range and resolution, and references relevant for the data product used in this report.

4.11.2 NO_2 evaluations: Zonal monthly mean cross sections and vertical profiles of local sunrise/sunset climatologies

Local sunrise NO_2 climatologies from the solar occultation instruments SAGE II, HALOE, ACE-FTS, SAGE III, POAM II, and POAM III are displayed in **Figure 4.11.2**. The annual zonal mean climatologies are calculated over the respective lifetime of each instrument. In general, the zonal mean distribution shows the largest NO_2 abundances around 10 hPa, with maximum values at high latitudes. The three solar occultation instruments in a non-sun-synchronous orbit (SAGE II, HALOE, ACE-FTS) offer near-global coverage over the course of a year, while the three instruments in a sun-synchronous orbit (SAGE III, POAM II, and POAM III) provide measurements in a narrow range at high latitudes. The higher annual mean abundances reported by the first group of instruments when compared to the latter may be the result of varying seasonal coverage instead of an actual measurement bias. One example of the complications arising from the varying data coverage is the evaluation of 1994-1996 POAM II local sunrise monthly zonal means that show no direct overlap with any of the other datasets. (Note that the overlap of the annual means visible in **Figure 4.11.2** does not necessarily correspond to an overlap of the monthly zonal means.) Moreover, the instruments with near-global coverage over the course of a year or season (SAGE II, HALOE and ACE-FTS) have strongly varying data coverage from month to month. Therefore, the comparison of annual zonal means, as carried out for the long-lived trace gases, may not be meaningful, and instead a comparison of monthly zonal means will be presented. Note that the data coverage problem is intensified by the necessary separation into local sunrise and sunset data.

Table 4.11.4: Time period, vertical range, vertical resolution, references and other comments for NO_2 measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
LIMS V6.0	Nov 78 – May 79	Cloud top – 50 km (+ mesosphere for polar night)	3.7 km	*Remsberg et al.*, 2010	
SAGE II V6.2	Oct 84 – Aug 05	Cloud top – 50 km	1.0 km (< 38 km) 5.0 km (> 38 km)	*Cunnold et al.*, 1991	Only satellite SS data are used
HALOE V19	Oct 91 – Nov 05	up to 50 km	2.5 km	*Grooß and Russell*, 2005	
POAM II V6.0	Oct 93 – Nov 96	20 – 40 km	1.5 – 2.5 km	*Lumpe et al.*, 1997 *Randall et al.*, 1998	
POAM III V4.0	Apr 98 – Dec 05	20 – 40 km	1.5 – 2.5 km	*Lumpe et al.*, 1997 *Randall et al.*, 1998	
OSIRIS V3-0	Oct 01 –	13 – 45 km	2 km	*Brohede et al.*, 2007a *Brohede et al.*, 2007b	
SAGE III V4.0	May 02 – Dec 05	Cloud top – 50 km	0.5 ~ 1.0 km		Only solar occultation products used
MIPAS V15 V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 12 – 50/70 km (day/night)	 3 – 6 km 2.5 – 6 km	 *Funke et al.*, 2005a *Funke et al.*, 2005b	 Change in spectral resolution in 2005
GOMOS V5.0	Mar 02 – Apr 12	20 – 50/70 km	4 km	*Kyrölä et al.*, 2010a	
SCIAMACHY V3-1	Sep 02 – Apr 12	11 – 42 km	3 – 5 km	*Bauer et al.*, 2012	
ACE-FTS V2.2	Mar 04 –	7 – 52 km	3 – 4 km	*Kerzenmacher et al.*, 2008	
HIRDLS V6.0	Jan 05 – Jan 08	20 – 50 km	1 km	*Gille and Gray*, 2011	

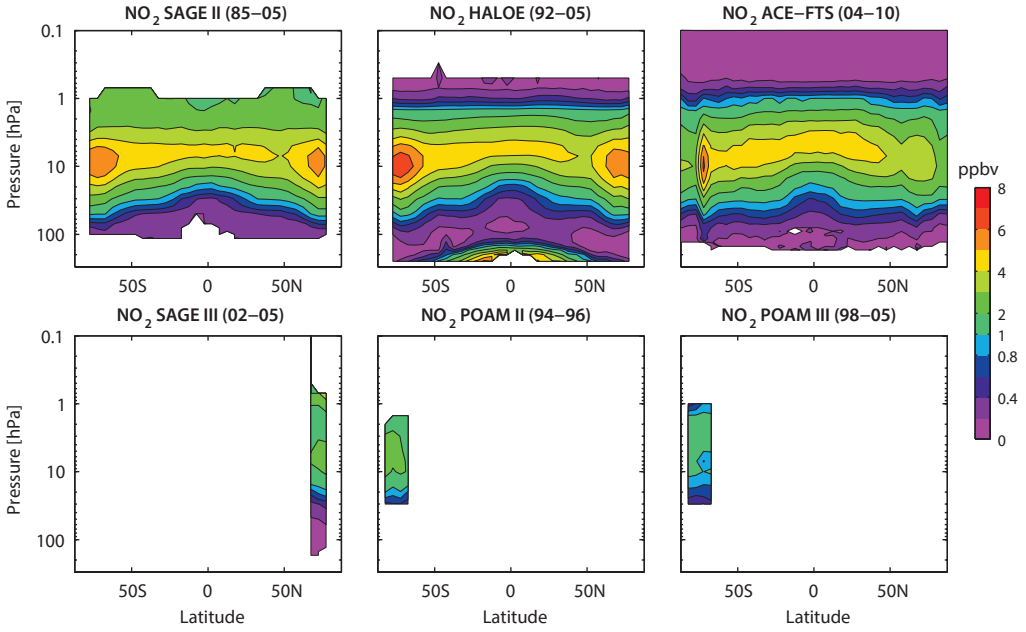


Figure 4.11.2: Cross sections of annual zonal mean, local sunrise NO_2 for solar occultation instruments. *Annual zonal mean, local sunrise NO_2 cross sections are shown for the SAGE II (1985-2005), HALOE (1992-2005), ACE-FTS (2004-2010), SAGE III (2002-2005), POAM II (1994-1996), and POAM III (1998-2005).*

As a first step, the evaluations of local sunrise and sunset climatologies will focus on the comparison of the two longer time series from SAGE II and HALOE during their overlap time period 1992-2005. The other instruments will be compared to SAGE II and HALOE for time periods that allow as many instruments as possible to be included (1994-1996 for POAM II, 2004-2005 for ACE-FTS, POAM III, SAGE III).

SAGE II and HALOE (1992-2005)

Figure 4.11.3 shows NO_2 local sunrise and sunset monthly mean climatologies for SAGE II and HALOE for April and October. The monthly multi-annual means of the two datasets overlap in both hemispheres. The sunset climatologies show notably more NO_2 than the sunrise climatologies as a result of N_2O_5 photolysis, which is driven by the available

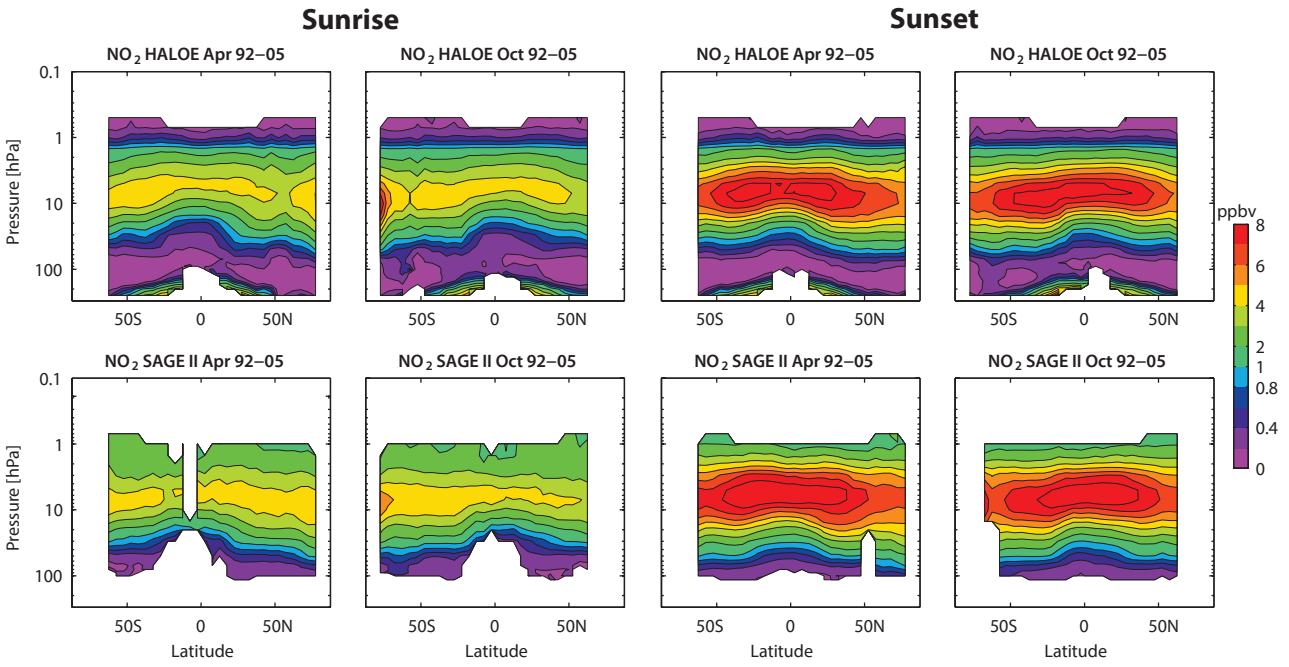


Figure 4.11.3: Cross sections of monthly zonal mean, local sunrise and sunset NO_2 for 1992-2005. *Monthly zonal mean, local sunrise (column 1 and 2) and local sunset (column 3 and 4) NO_2 cross sections for April and October are shown for HALOE (upper row) and SAGE II (lower row).*

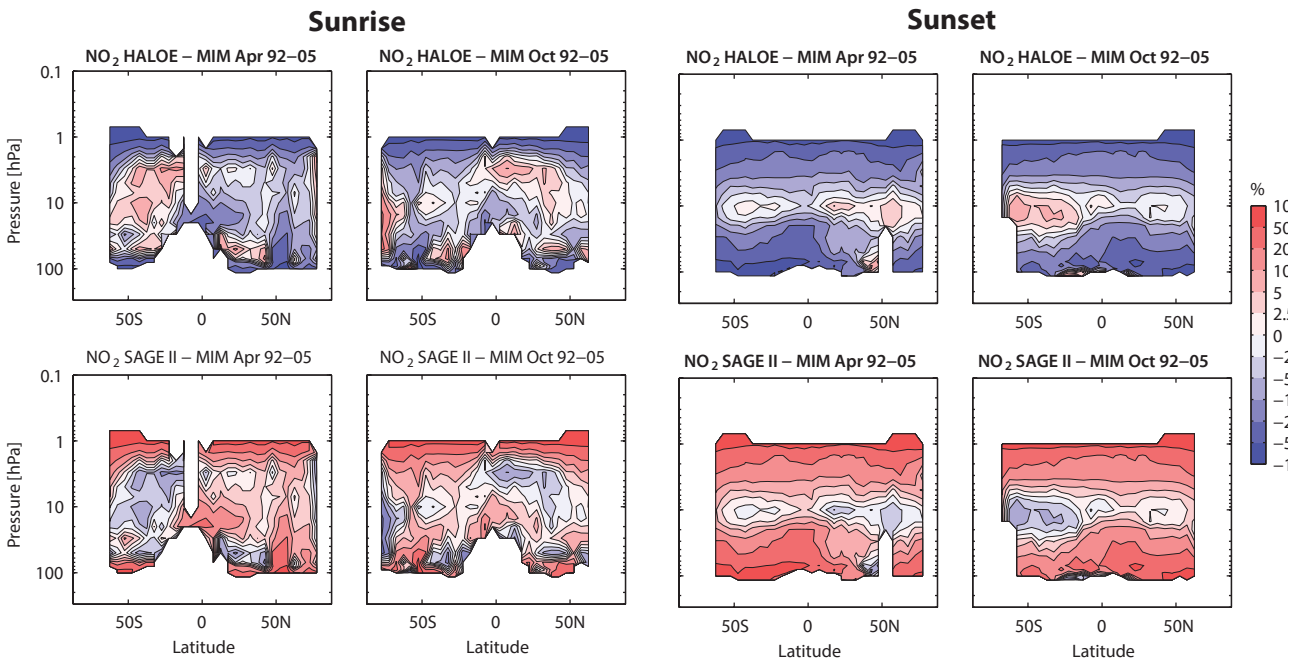


Figure 4.11.4: Cross sections of monthly zonal mean, local sunrise and sunset NO_2 differences for 1992-2005. *Monthly zonal mean, local sunrise (column 1 and 2) and local sunset (column 3 and 4) NO_2 differences for April, and October between the individual instruments (SAGE II, HALOE) and their MIM are shown.*

sunlight and reactivates NO_2 from its reservoir species. The N_2O_5 photolysis during daytime is fast enough that by sunset nearly all N_2O_5 is converted to NO_2, leading to a nearly symmetric distribution at the equator independent of the time of year. The N_2O_5 production during nighttime is somewhat slower, and as a result larger NO_2 sunrise abundances are observed for shorter nights (*e.g.*, in the SH during October).

The relative differences of SAGE II and HALOE from their MIM in April and October are displayed in **Figure 4.11.4**. At the level of maximum NO_2 abundance (10 hPa), the two datasets agree well with differences from the MIM of up to ±10% (corresponding to differences of 20% between the two instruments). Above and below this level, the relative differences increase steadily reaching values of up to ±50% at 2 hPa and 20/50 hPa for local sunset/sunrise climatologies. The steadily increasing differences below and above the maximum are related to the smaller vertical gradients in SAGE II NO_2 compared to HALOE. At around 10 hPa, HALOE detects more NO_2 than SAGE II, but above and below the maximum it quickly reaches lower values due to its stronger gradients and shows mostly negative differences when compared to SAGE II. Exceptions to this pattern are the local sunrise climatologies in the summer (**Figure A4.11.2** in *Appendix A4*) and autumn hemispheres (**Figure 4.11.4**) when HALOE shows positive differences from the MIM everywhere below 2 hPa. Overall, the NO_2 local sunrise and sunset evaluations give a consistent picture, however, some small differences exist (*e.g.*, sunset differences

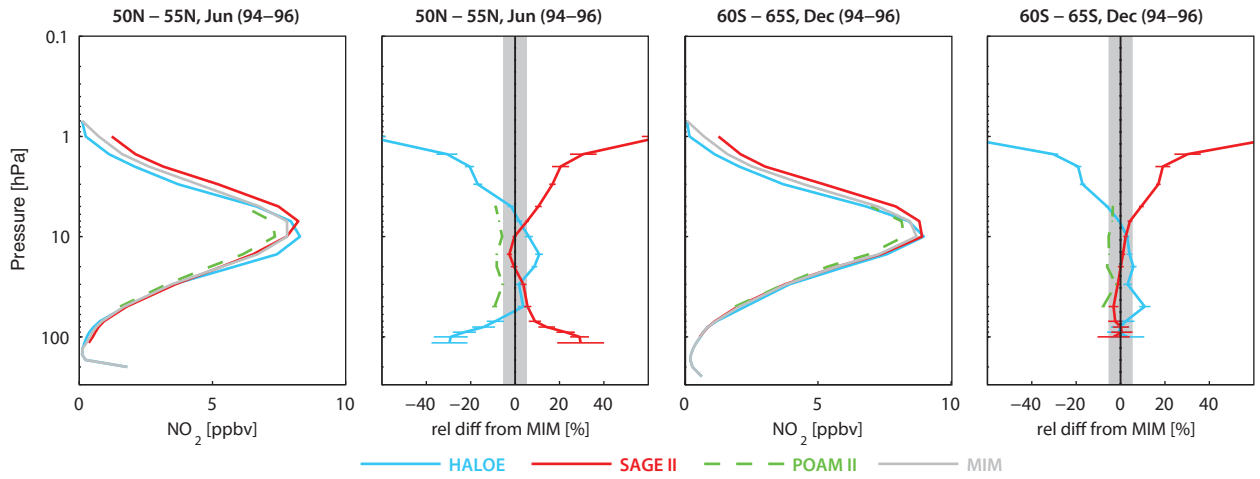


Figure 4.11.5: Profiles of zonal mean, local sunset NO_2 for 1994-1996. *Zonal mean, local sunset NO_2 profiles for 50°N-55°N for June and 60°S-65°S for December are shown with their differences from the MIM. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.*

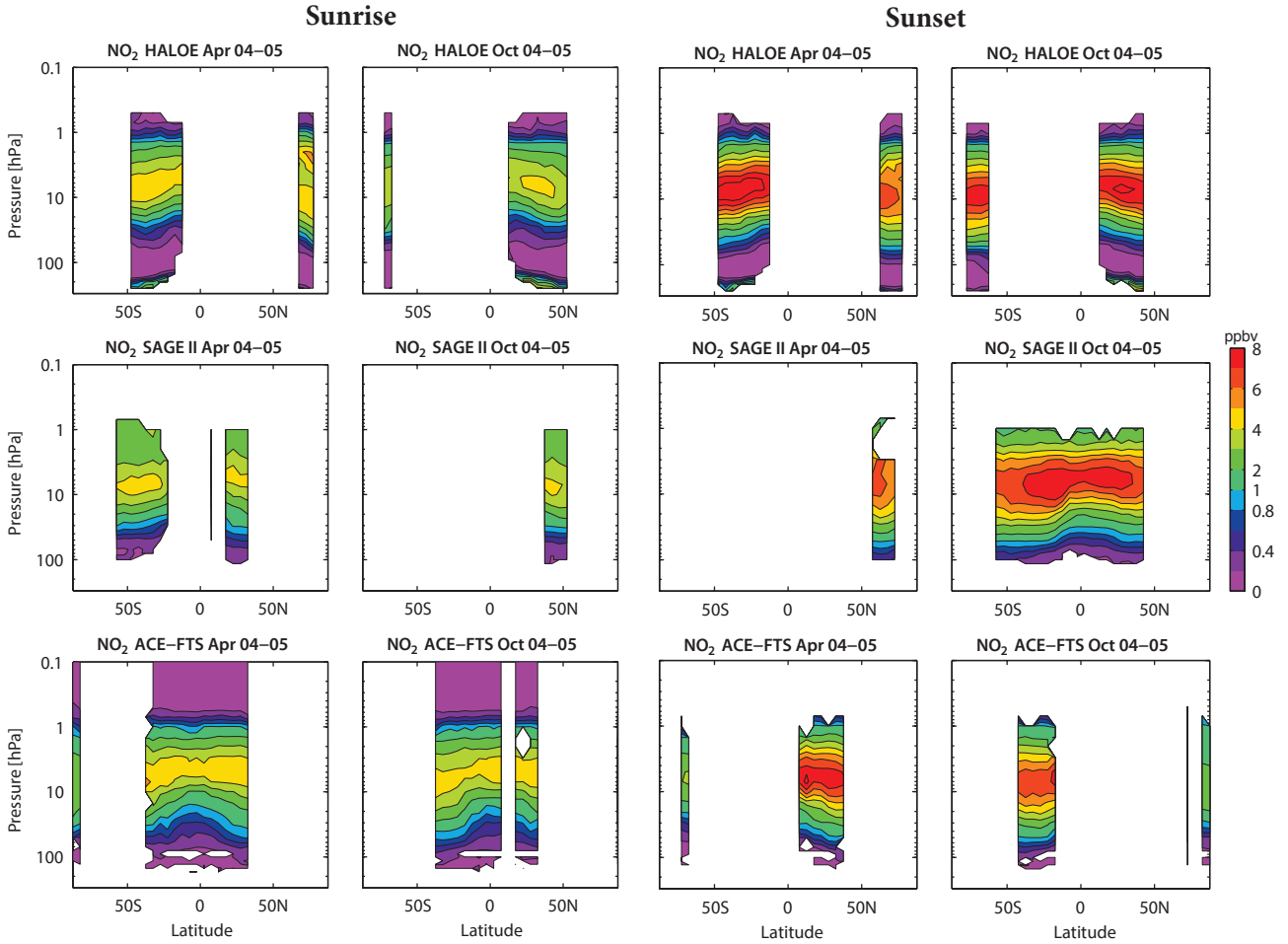


Figure 4.11.6: Cross sections of monthly zonal mean, local sunrise and sunset NO_2 for 2004-2005. *Monthly zonal mean, local sunrise (column 1 and 2) and local sunset (column 3 and 4) NO_2 cross sections for April and October are shown for the HALOE, SAGE II, and ACE-FTS.*

show a stronger hemispheric symmetry consistent with the same feature observed for the sunset mixing ratios).

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

As a consequence of the gaps in temporal and spatial coverage of POAM II, a direct comparison of local sunrise NO_2 climatologies from POAM II and the other two instruments is not possible. However, for NO_2 sunset measurements, some months and latitude bands exist where all three datasets overlap, allowing for a direct comparison of POAM II, SAGE II and HALOE. Vertical profiles from all three instrument local sunset climatologies and their relative differences from the MIM are displayed in **Figure 4.11.5**. In both latitude bands (NH and SH mid-latitudes), the POAM II climatology reports smaller values than SAGE II and HALOE, with differences with respect to the MIM of -10%. Above 7 hPa, POAM II agrees better with HALOE than with SAGE II, while below 7 hPa the situation is reversed.

SAGE II, HALOE, POAM III, SAGE III, and ACE-FTS (2004-2005)

For the period 2004-2005, the long NO_2 time series from SAGE II and HALOE overlap with the more recent data from ACE-FTS. Also available during this period are datasets from SAGE III and POAM III, which focus on narrow latitude ranges. A comparison of SAGE III and POAM III zonal monthly mean cross sections is not possible, therefore the evaluation of these two climatologies will be based on vertical profiles.

Figure 4.11.6 shows NO_2 local sunrise and sunset monthly mean climatologies for SAGE II, HALOE, and ACE-FTS for April and October. Note that SAGE II and HALOE have less latitudinal coverage for 2004-2005 (the last two years of their lifetimes) when compared to earlier time periods (*e.g.*, **Figure 4.11.3**). The respective relative differences between the individual datasets and their MIM are presented in **Figure 4.11.7**. Overall, the magnitude and spread of the relative differences are similar to those already discussed for the long-term comparison of SAGE II and HALOE, with the smallest differences (up to ±10%) around 10 hPa, and larger differences (up to ±50%) above 2 hPa and below 50 hPa. Note that for some cases the differences are below ±20% for the complete altitude range from 1 to 100 hPa (*e.g.*, sunset differences of SAGE II and ACE-FTS to their MIM for SH October). For the local sunrise climatologies, ACE-FTS shows the largest, and HALOE the lowest NO_2 abundances. For the local sunset climatologies, SAGE II show the largest and HALOE the lowest NO_2 abundances, with ACE-FTS values between the other two

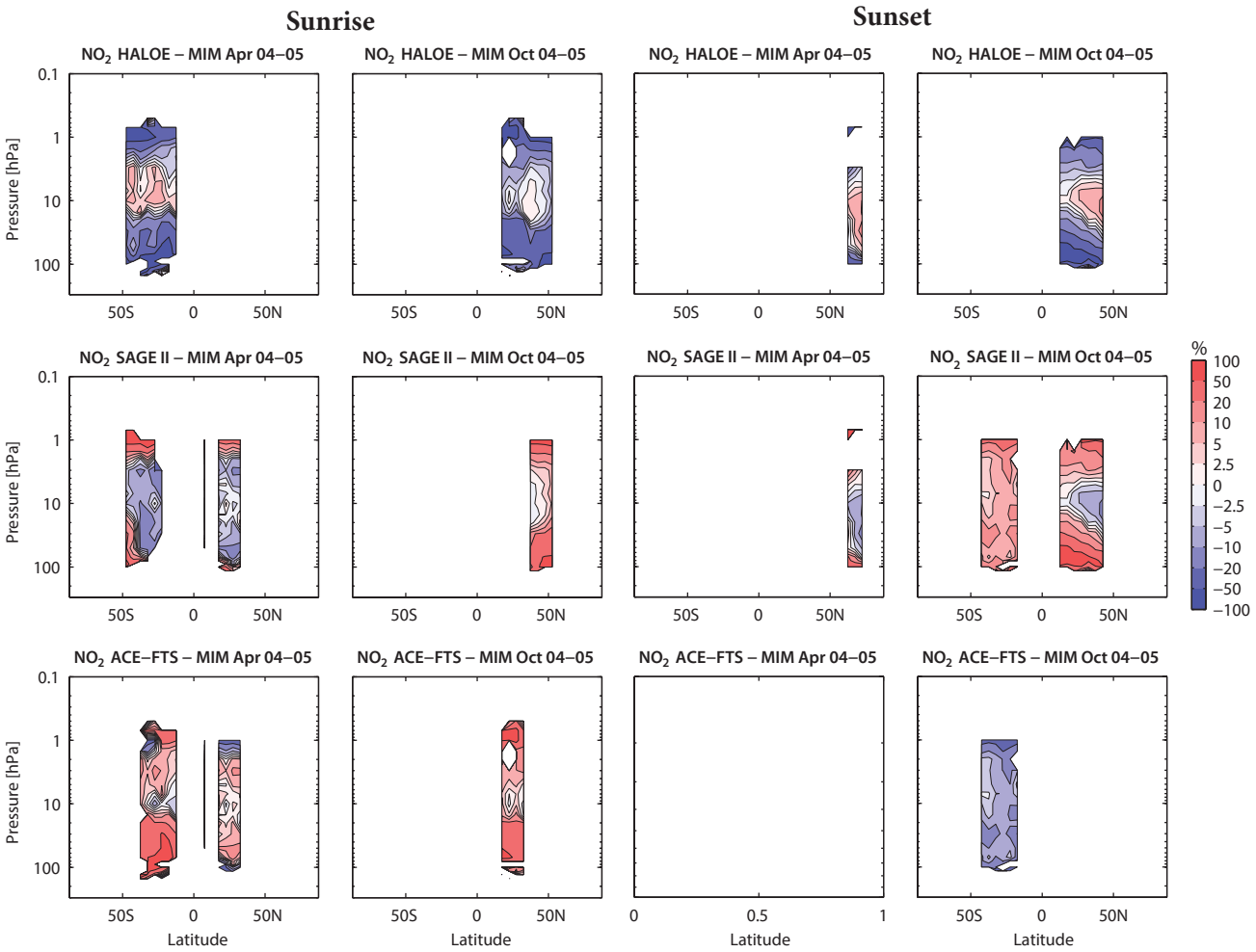


Figure 4.11.7: Cross sections of monthly zonal mean, local sunrise and sunset NO_2 differences for 2004-2005. *Monthly zonal mean, local sunrise (column 1 and 2) and local sunset (column 3 and 4) NO_2 differences for April and October between the individual instruments (HALOE, SAGE II, ACE-FTS) and their MIM are shown.*

instruments. The only exception is found at 10 hPa where HALOE detects a larger NO_2 peak than the other two datasets, consistent with the evaluations of the earlier time period. Overall, ACE-FTS agrees better with SAGE II than with HALOE.

A comparison of POAM III and SAGE III with the other datasets is shown in **Figure 4.11.8** as vertical profiles and their relative differences from the MIM. The months and latitude bands used for this comparison are chosen because the coverage includes the maximum number of datasets available. The profiles confirm the good agreement of all instruments

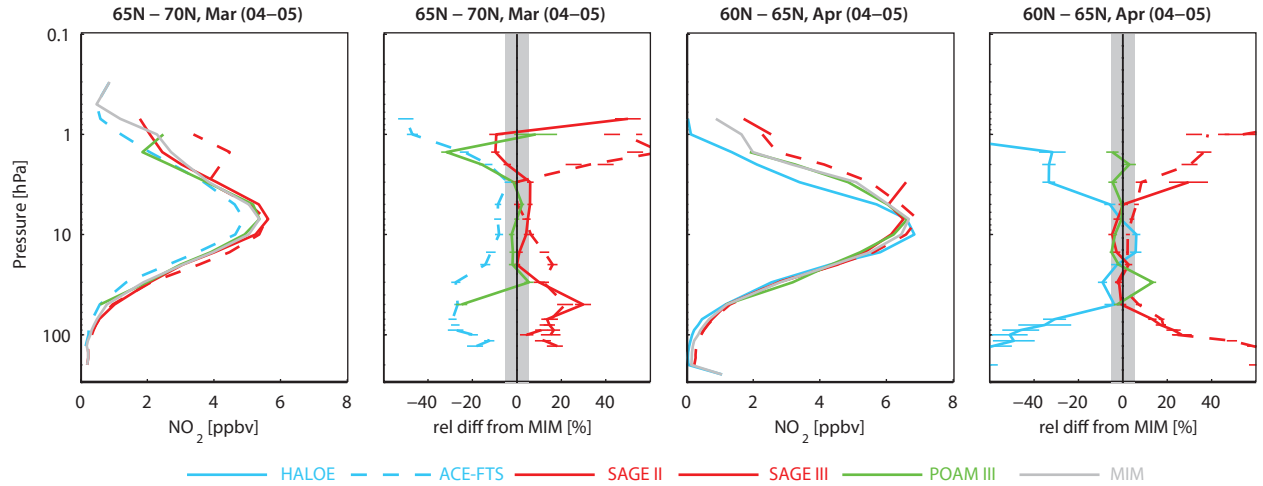


Figure 4.11.8: Profiles of zonal mean, local sunset NO_2 for 2004-2005. *Zonal mean, local sunset NO_2 profiles for 65°N-70°N for March and 60°N-65°N for April are shown together with their differences from the MIM. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.*

in the MS with differences often below ±5%, except for a divergence between ACE-FTS and SAGE III in March which show differences from the MIM of up to ±10%. Differences are large in the LS and USLM consistent with low NO_2 abundances. However, there is a much better agreement between 50 and 10 hPa (below the NO_2 VMR peak) compared to between 5 and 1 hPa (above the NO_2 peak) for similar NO_2 background abundances. In general, SAGE III is similar to SAGE II and shows larger NO_2 values than the other datasets, while POAM III values reside mostly in the middle.

4.11.3 NO_2 evaluations: Zonal monthly mean cross sections of 10am/pm climatologies

NO_2 measurements by limb emission and scattering techniques and from stellar occultation have been sorted according to LST or scaled with the help of chemical box models (see *Section 4.11.1* for details). Additionally, the solar occultation dataset from ACE-FTS has been scaled to allow for a first-step comparison between the two groups of instruments. In the following, climatologies corresponding to 10am and 10pm will be compared with each other. For a better understanding of the scaling effects, the initial climatologies are also shown. Note that the 10am/pm climatologies are part of the larger groups of daytime/night-time climatologies as explained in **Table 4.11.3**.

OSIRIS, SCIAMACHY, MIPAS, GOMOS, HIRDLS, and ACE-FTS (2005-2007)

Figure 4.11.9 shows the NO_2 10am climatologies for October 2005-2007 that can be directly compared to each other: MIPAS (corresponding to 10am LST) and ACE-FTS, OSIRIS, SCIAMACHY, and HIRDLS, all scaled to 10am LST (labelled as s10am in the figure titles). Additionally, unscaled daytime data from SCIAMACHY and OSIRIS (corresponding to a range of LSTs), HIRDLS daytime measurements (corresponding to 3pm), and ACE-FTS local sunrise measurements are shown. The overall NO_2 distribution shows flat isopleths and maximum abundances in the subtropics between 5 and 10 hPa. Note that the elevated values observed by MIPAS at very high NH latitudes are related to small changes of the LST from 10am and the timing of sunsets/sunrises (near 10am) during October. Similarly, MIPAS shows elevated values at high SH latitudes for April (see **Figure A4.11.6** in *Appendix A4*).

The MIM is calculated based on MIPAS and scaled ACE-FTS, OSIRIS, SCIAMACHY, and HIRDLS climatologies

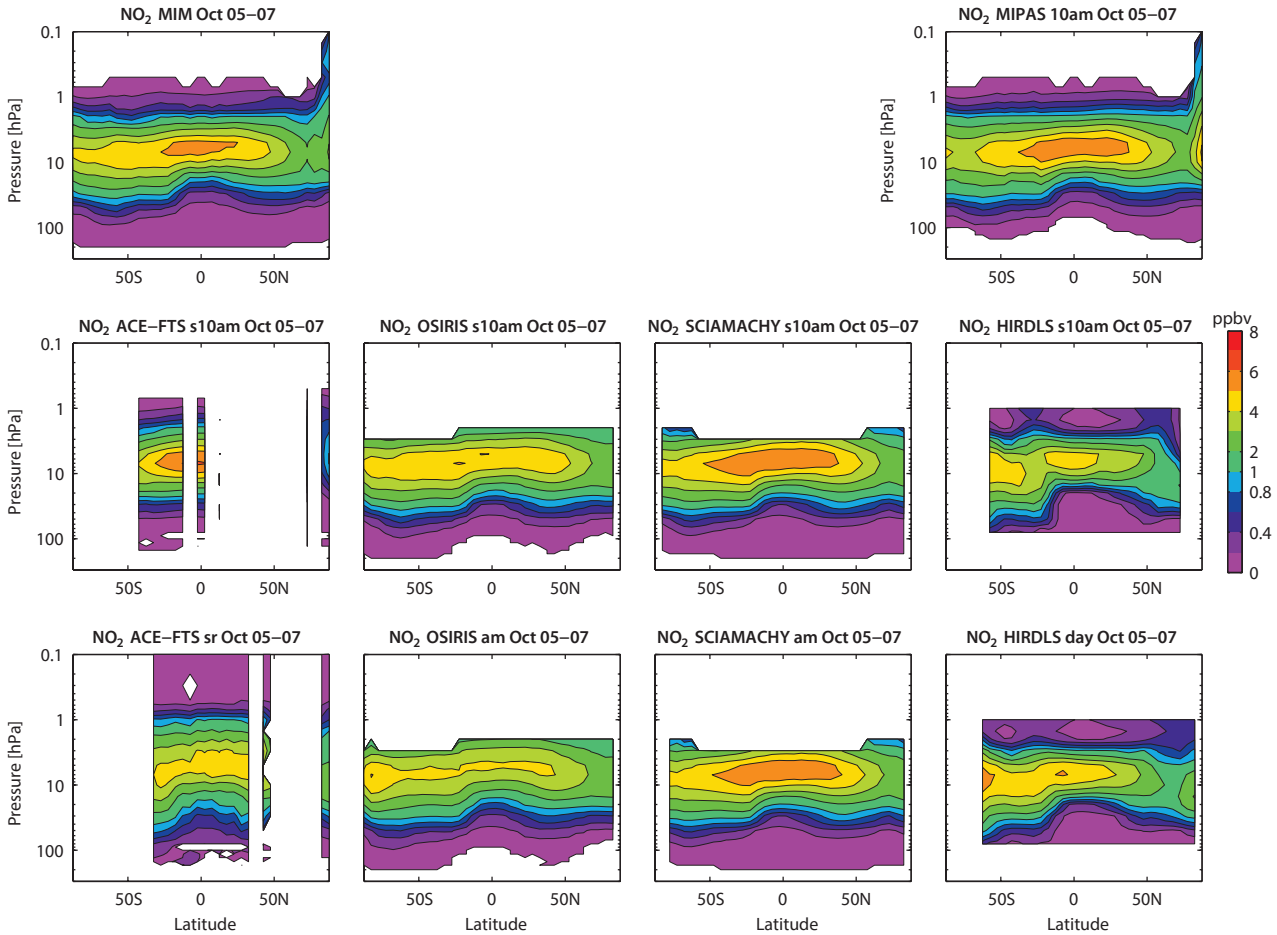


Figure 4.11.9: Cross sections of zonal mean daytime NO_2 for October 2005-2007. *Monthly zonal mean NO_2 cross sections for October 2005-2007 are shown for the MIM (upper left panel) and the individual instruments. Measurements correspond to 10am LST (MIPAS) or are scaled to 10am LST. Note that scaled HIRDLS data are only available for June 2005 – May 2006. In addition, unscaled data from ACE-FTS, OSIRIS, SCIAMACHY, and HIRDLS are shown.*

for 2005-2007. Differences of the individual datasets from the MIM for October 2005-2007 are shown in **Figure 4.11.10**. In general, the differences in the MS vary strongly, from ±5% in some regions to ±50% in others. MIPAS and SCIAMACHY show similar features and observe larger NO_2 abundances than the other instruments, leading to deviations from the MIM in the MS of around +20%, and locally +50%. One exception is the SH extra-tropics where both instruments show negative differences from the MIM of up to -20%. While MIPAS and SCIAMACHY are on the high side, HIRDLS observes values at the lower end of the range, producing large negative deviations from the MIM of up to -50% (and locally -100%). An exception to this behaviour is again the SH extra-tropics where HIRDLS detects larger NO_2 abundances than the other instruments. OSIRIS is mostly in the middle range; deviations from the MIM change sign depending on the latitude band and exceed ±50% only occasionally. ACE-FTS also observes deviations of mixed signs, mostly opposite to HIRDLS with particularly large negative deviations where HIRDLS shows a positive difference. Note that at high SH latitudes, where only MIPAS, OSIRIS and SCIAMACHY measurements are available, the three instruments agree considerably better, with deviations from their MIM of only up to ±20%, while in regions where all five instruments provide measurements, deviations can reach ±50% to ±100%. Evaluations of January, April, and July climatologies (see **Figure A4.11.6** in *Appendix A4* for April evaluations) give a consistent picture for all four seasons. The main difference from the October evaluations discussed above is that MIPAS and HIRDLS show more differences of mixed signs. While in the SH spring and summer, MIPAS is mostly positive and HIRDLS mostly negative, MIPAS deviations from the MIM in the NH spring and summer are negative almost everywhere below 10 hPa. The pattern of OSIRIS deviations from the MIM changes from month to month. Note that scaled HIRDLS data are only available for June 2005 – May 2006, and that the evaluation of this one-year long time period (not shown) gives very similar results to the evaluation presented above.

For a better understanding of the scaling effects, the initial climatologies are also shown. The unscaled OSIRIS and ACE-FTS datasets have a considerably larger bias compared to the scaled datasets, indicating that scaling is necessary in order to compare the climatologies. SCIAMACHY shows similar differences for the scaled and unscaled datasets, which can be explained by the fact that the LST of the original SCIAMACHY dataset is 10am at the equator and only changes slightly when moving to higher latitudes. For some regions (*e.g.*, SH in March) the scaled dataset shows somewhat larger differences from the MIM than the unscaled

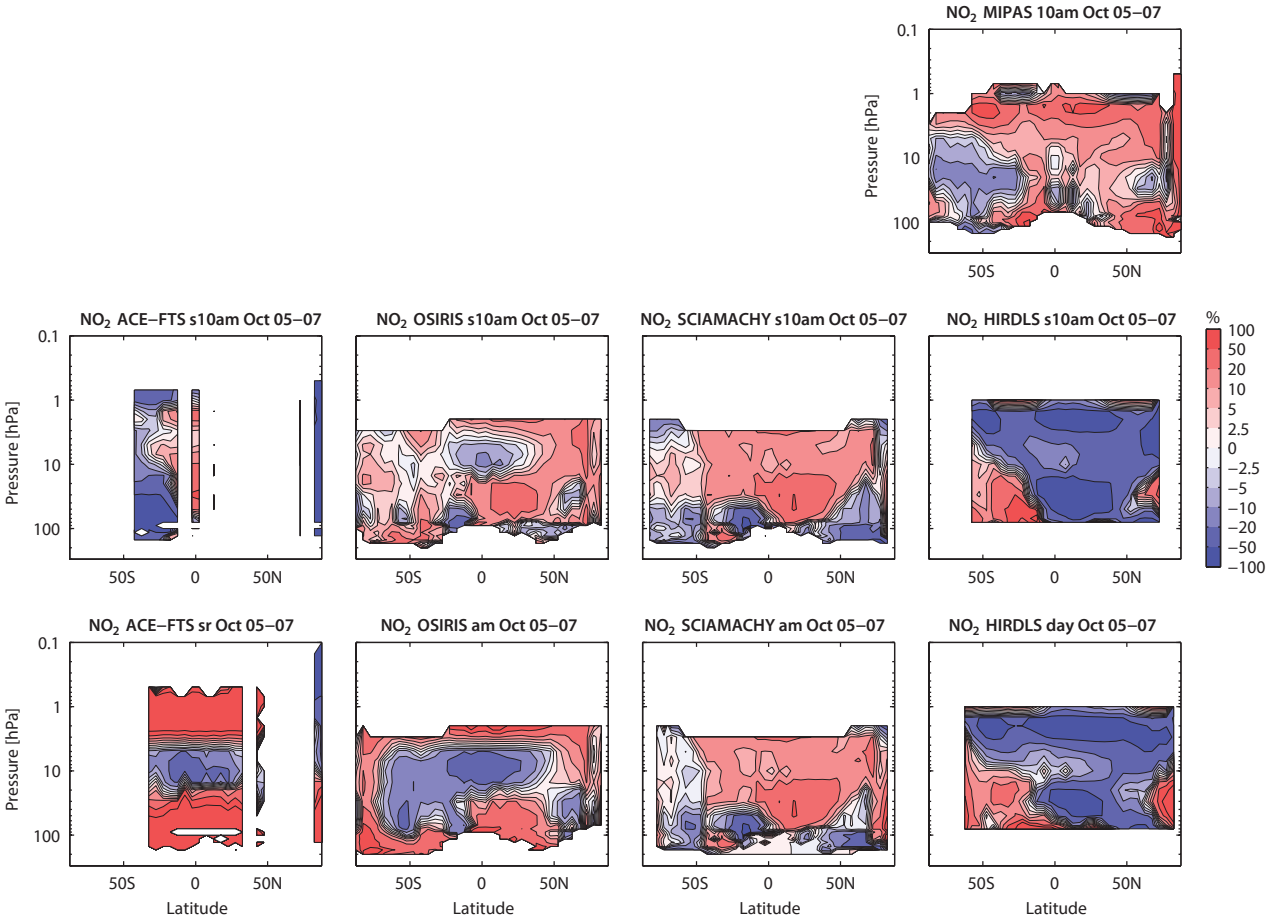


Figure 4.11.10: Cross sections of zonal mean daytime NO_2 differences for October 2005-2007. *Monthly zonal mean NO_2 differences from the MIM for October 2005-2007 are shown. Measurements correspond to 10am LST (MIPAS) or are scaled to 10am LST. Note that scaled HIRDLS data are only available for June 2005 – May 2006. In addition, unscaled data from ACE-FTS, OSIRIS, SCIAMACHY, and HIRDLS are shown.*

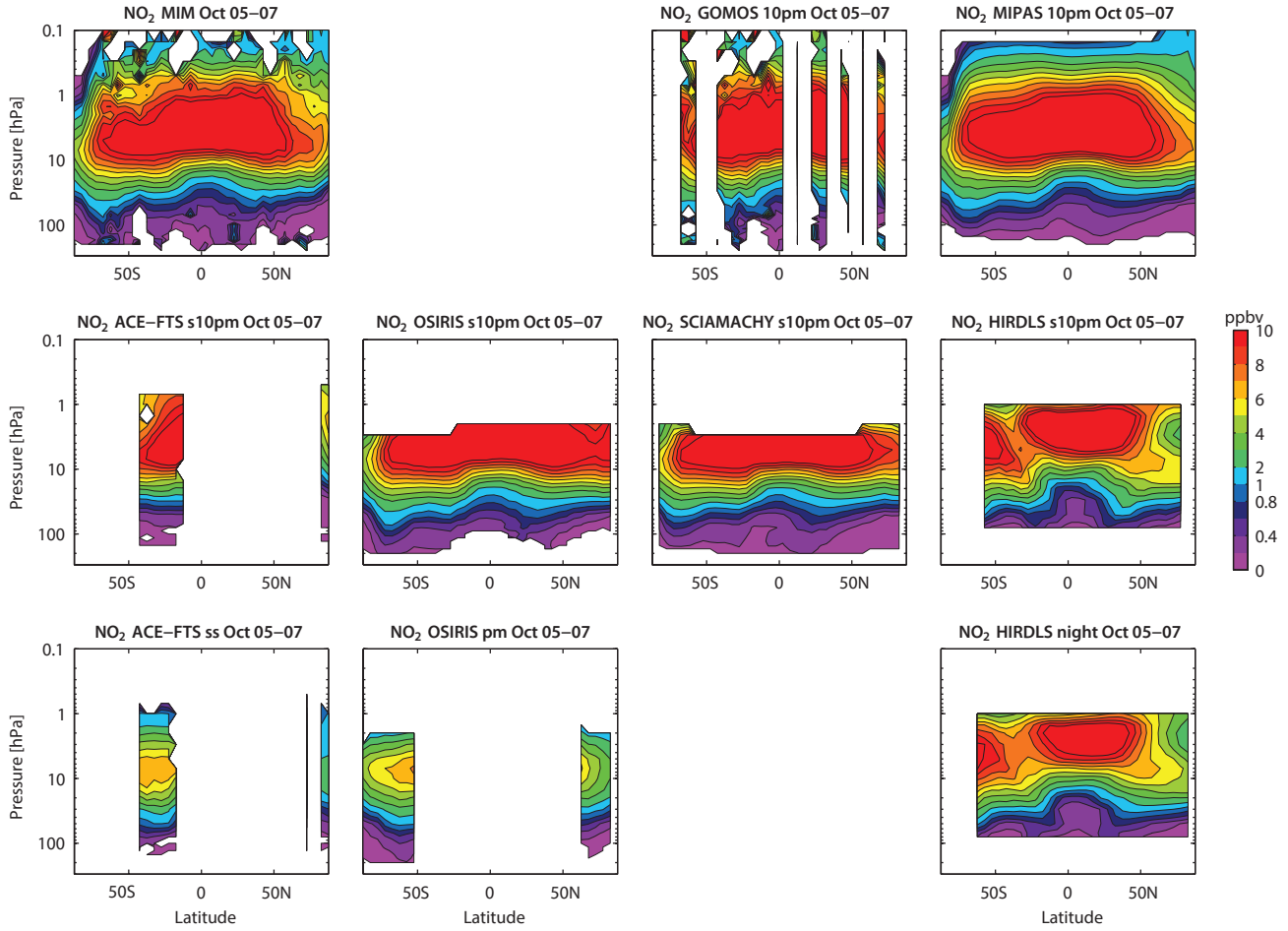


Figure 4.11.11: Cross sections of zonal mean night-time NO_2 for October 2005-2007. *Monthly zonal mean NO_2 cross sections for October 2005-2007 are shown for the MIM (upper panel) and the individual instruments. Measurements correspond to 10pm LST (MIPAS, and GOMOS) or are scaled to 10pm LST. Note that scaled HIRDLS data are only available for June 2005 – May 2006. In addition, unscaled data from ACE-FTS, OSIRIS, and HIRDLS are shown.*

dataset, unlike the results from OSIRIS. In these cases, it is possible that the errors introduced by the scaling based on look up tables outweigh the improvement achieved by the correction to a specific LST. For HIRDLS, large negative deviations are apparent in the scaled and unscaled datasets, indicating that the differences are related to measurement differences and cannot be corrected by accounting for the measurement LST.

NO_2 night-time climatologies for October 2005-2007 are shown in **Figure 4.11.11**. NO_2 abundances are considerably larger for the night-time climatologies than for daytime climatologies, as expected from the diurnal cycle. Maximum NO_2 values can be observed in the SH mid-latitudes and tropics between 1 and 10 hPa. As for the daytime climatologies, MIPAS data (corresponding to 10pm LST) as well as scaled ACE-FTS, OSIRIS, SCIAMACHY and HIRDLS data (all scaled to 10pm LST) are available and can be compared directly. Additionally, GOMOS measures night-time NO_2 at 10pm LST. Also shown in **Figure 4.11.11** are the unscaled OSIRIS night-time climatology (corresponding to a range of LSTs), the HIRDLS night-time climatology (corresponding to approximately 00:30am) and the ACE-FTS local sunset climatology. Note that for SCIAMACHY no measurements during the night exist and that the scaled 10pm SCIAMACHY climatology is based on daytime SCIAMACHY measurements.

Differences of the individual night-time climatologies from the MIM are displayed in **Figure 4.11.12**. The MIM is based on the climatologies corresponding directly to 10pm LST (MIPAS, GOMOS), and scaled to 10pm LST (ACE-FTS, OSIRIS, SCIAMACHY and HIRDLS). MIPAS and SCIAMACHY have positive deviations in the MS and negative deviations in the LS, and agree well with each other and with OSIRIS. The climatologies from HIRDLS and ACE-FTS show lower values for most latitude bands, with deviations from the MIM of up to -50%. The GOMOS dataset is somewhat noisier than the other climatologies but shows small differences from the MIM (up to ±10%) between 10 and 1 hPa. In the LS, MIPAS, SCIAMACHY and ACE-FTS observe the lowest values while GOMOS and OSIRIS show positive deviations from the MIM. Differences can become as large as ±100% locally. Note that SCIAMACHY data scaled to 10pm represent a scaling to completely different illumination conditions, namely from day to night, while the scaling of SCIAMACHY data to 10am is a daytime to daytime scaling, with only slightly different illumination conditions. The fact that the SCIAMACHY night-time climatology does not show larger differences from the MIM

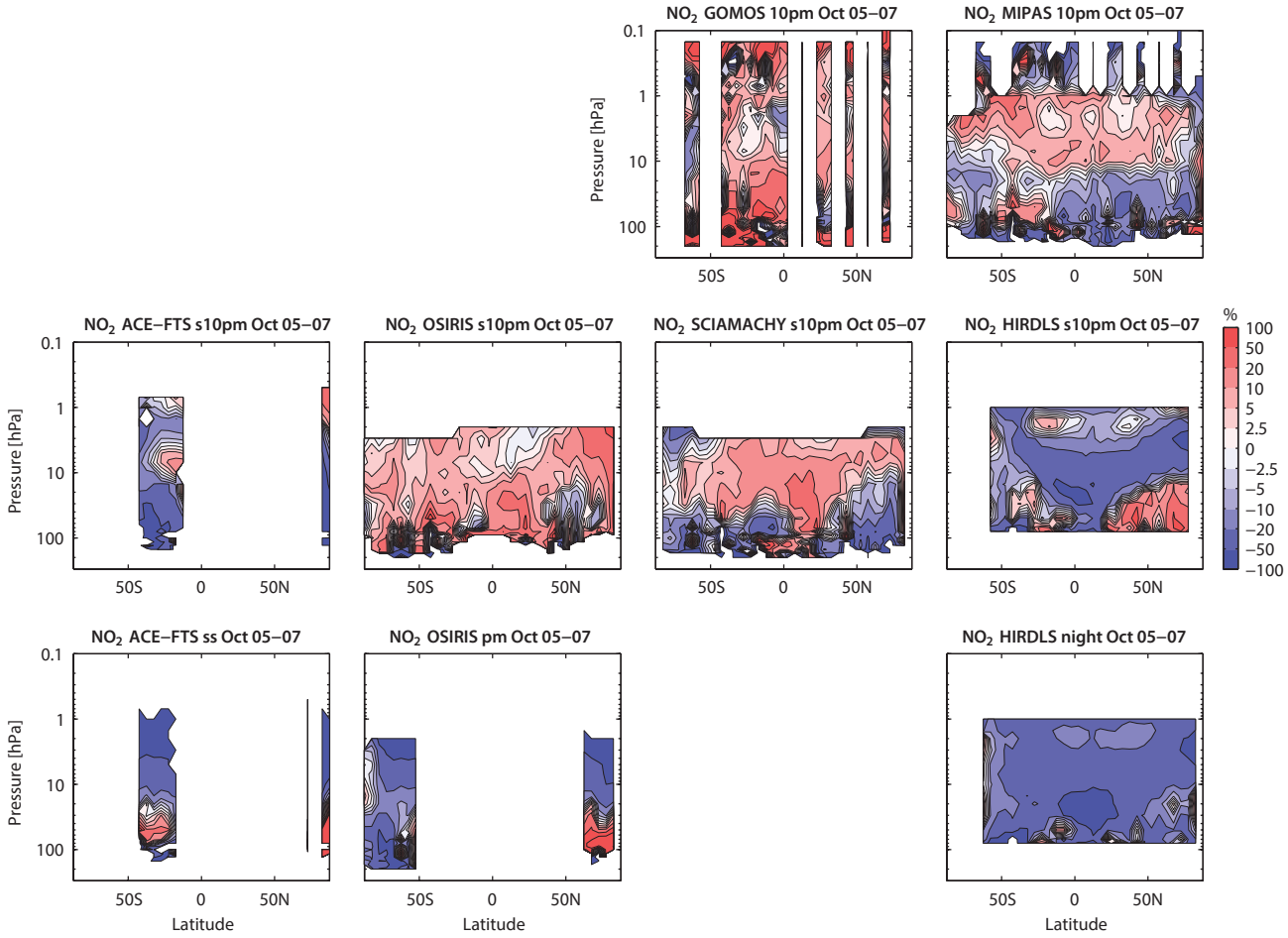


Figure 4.11.12: Cross sections of zonal mean night-time NO_2 differences for October 2005-2007. *Monthly zonal mean NO_2 differences from the MIM for October 2005-2007 are shown. Measurements correspond to 10pm LST (MIPAS, and GOMOS) or are scaled to 10pm LST. Note that scaled HIRDLS data are only available for June 2005 – May 2006. In addition, unscaled data from ACE-FTS, OSIRIS, and HIRDLS are shown.*

than the daytime climatology suggests that no large errors have been introduced by this scaling procedure.

A comparison of the night-time climatologies without GOMOS (see **Figure A4.11.9** in *Appendix A4*) shows that the observed differences are quite similar to the daytime evaluations, and that both sets of climatologies provide a consistent picture. Note that differences are slightly smaller between the night-time climatologies than between the daytime climatologies, which might be related to the larger NO_2 abundances during night-time.

In addition to the NO_2 datasets discussed above, which cover the time period after 2000, the very first satellite borne NO_2 profile measurements from the limb emission instrument LIMS are available for 1978/1979. The LIMS daytime climatology corresponds to 1pm at low and mid-latitudes and shifts to late afternoon at higher latitudes, while the LIMS night-time climatology corresponds approximately to 11pm LST. Since there are no daytime or night-time NO_2 measurements available before 2002, LIMS will be compared to the 2005-2007 climatologies presented above. Note that the stratospheric aerosol content is very low for both time periods, which facilitates the comparison of LIMS NO_2 with the other datasets.

Figure 4.11.13 shows the relative differences of LIMS day- and night-time climatologies to the respective 10am/pm 2005-2007 MIM from the evaluations above. LIMS shows good agreement in a narrow pressure range between 10 and 5 hPa, with differences mostly between ±10%. Below this level differences increase quickly, reaching +100% in the LS. Above this level, the daytime climatologies differ by up to -50% while the night-time climatologies show slightly better agreement, with differences of up to -20%. Note that LIMS measurements are not taken at 10am or 10pm LST, and it is therefore not possible to easily separate the differences attributable to a real measurement bias, and the effects of the diurnal NO_2 cycle.

4.11.4 NO_2 evaluations: Seasonal cycles

NO_2 exhibits a strong seasonal cycle in the extra-tropics due to the effects of sunlight on the partitioning of the NO_y family. Seasonal cycles of the NO_2 daytime and night-time climatologies will be evaluated below. Since the seasonal variations of the sunset and sunrise climatologies from the solar occultation instruments can be difficult to analyse due to sparse data coverage, they are not shown here.

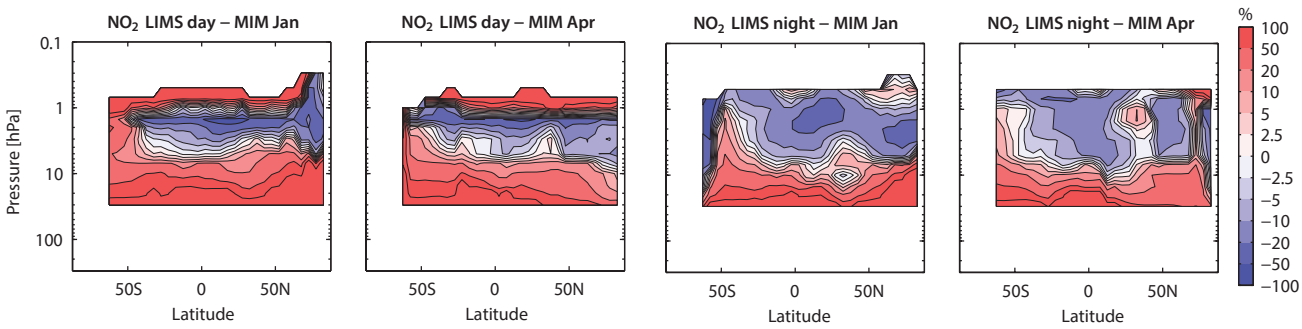


Figure 4.11.13: Cross sections of zonal mean daytime/night-time NO_2 differences between LIMS (1978-1979) and MIM (2005-2007). *Monthly zonal mean NO_2 differences between LIMS (1978-1979) and the MIM (2005-2007) are shown for January and April. Measurements correspond to 1pm for LIMS and 10am for the MIM (left panels) and the 11pm for LIMS and 10pm for MIM (right panels). The MIM is based on the 10am/pm climatologies for 2005-2007 presented above.*

Figure 4.11.14 shows the seasonal cycle of 2005-2007 daytime NO_2 climatologies for the NH and SH mid-latitudes at 10 hPa and the SH tropics at 3 hPa. The five datasets evaluated here (MIPAS and scaled ACE-FTS, OSIRIS, SCIAMACHY, and HIRDLS) correspond to 10am LST. The mid-latitude seasonal cycle with maximum values in summer and minimum values in winter is reproduced by all instruments. The datasets agree better during the summer, and show a larger spread during the hemispheric winter. The seasonal cycles from HIRDLS and ACE-FTS have larger amplitudes than the other datasets in both hemispheres, particularly in the NH mid-latitudes. In contrast, the seasonal cycle observed by MIPAS shows the smallest amplitude. Note that OSIRIS does not provide data for the SH winter months, leading to a flattening of the fitted seasonal cycle. In the tropics, seasonal variations are weak and the instruments have problems reproducing the signal. MIPAS and SCIAMACHY display the same seasonal cycle, with a maximum in April and a minimum in August/September. The seasonal cycle observed by OSIRIS agrees reasonably well but has a slightly larger amplitude and earlier minimum (July). The largest deviations are observed by HIRDLS, which shows a stronger amplitude and a weak second maximum in January/February, indicating a semiannual oscillation. Note that ACE-FTS in the SH tropics provides data only at the beginning and end of the year, and therefore does not offer sufficient information to fit the seasonal cycle. While ACE-FTS values for January and August are close to MIPAS and HIRDLS, the ACE-FTS values for February, April and October do not follow the seasonal signal suggested by the other instruments, indicating that the scaled ACE-FTS data observe at best a weak annual cycle.

Comparing the scaled climatologies in **Figure 4.11.14** with the seasonal cycle of the respective unscaled climatologies (**Figure A4.11.10** in *Appendix A4*) demonstrates a strong improvement in agreement of the seasonal cycle for OSIRIS, HIRDLS, and ACE-FTS for the SH and NH mid-latitudes, compared to each other and with MIPAS and SCIAMACHY. In the tropics, however, OSIRIS observes the same mean values and the same seasonal signal as MIPAS before the scaling, while the scaled dataset has lower mean values and a slightly shifted phase. Similarly, HIRDLS in

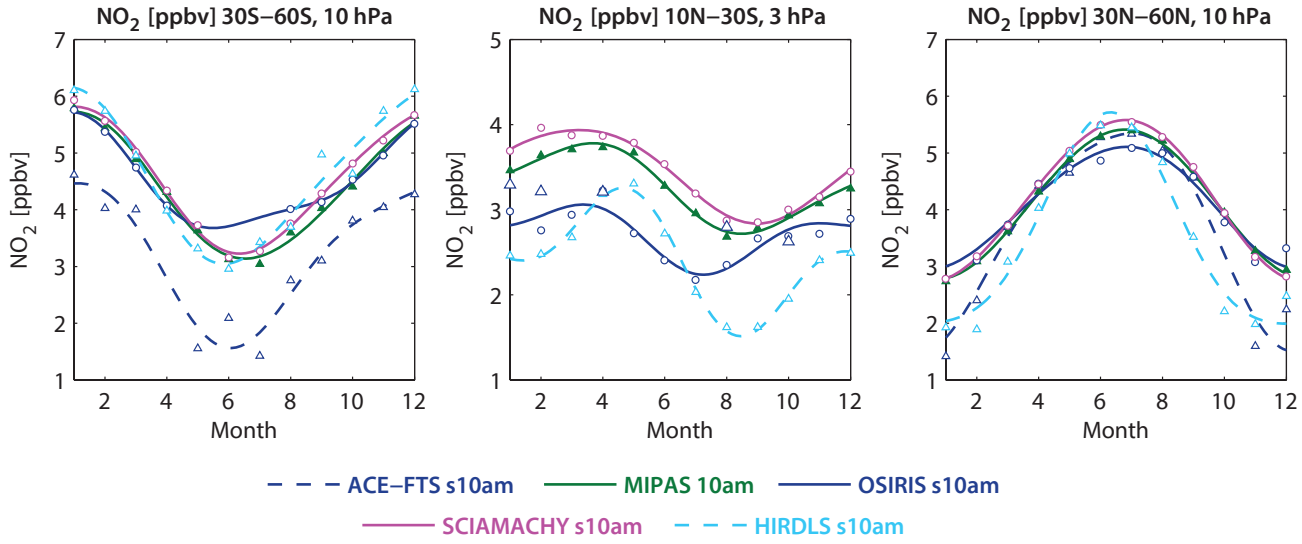


Figure 4.11.14: Seasonal cycle of daytime NO_2 for 2005-2007. *Seasonal cycle of monthly zonal mean NO_2 for 30°S-60°S at 10 hPa (left panel), 10°S-30°S at 3 hPa (middle panel) and 30°N-60°N at 10 hPa (right panel). Measurements correspond directly to 10am LST (filled symbols) or are scaled to 10am LST (open symbols). Note that scaled HIRDLS data are only available for June 2005 – May 2006. ACE-FTS in the SH tropics does not provide sufficient data coverage to estimate a fit of the seasonal cycle.*

the tropics has more similarities with the other instruments before the scaling in the sense that it shows no semiannual component. Finally, tropical unscaled ACE-FTS data have the same signal as MIPAS and SCIAMACHY as opposed to the scaled dataset.

Figure 4.11.15 displays the seasonal cycle of night-time NO_2 for 2005-2007. Although the general shape of the seasonal cycle is very similar to the one derived from daytime measurements, the instruments show a much larger spread in the signal. Excellent agreement is found between SCIAMACHY and GOMOS for the SH mid-latitudes. In this region, MIPAS and HIRDLS show the same phase but HIRDLS has a stronger amplitude than MIPAS. Similar to the daytime measurements, OSIRIS has little data available during SH winter, and the seasonal fit displays a flattened and shifted peak. In the NH mid-latitudes, the instruments disagree on the timing of the maximum, from June (HIRDLS), over July (MIPAS, SCIAMACHY) and August (OSIRIS), to September (GOMOS). Additionally, there is considerable spread in the amplitude of the seasonal signal, with HIRDLS displaying the largest amplitude and MIPAS and OSIRIS the smallest. The tropical seasonal cycles of MIPAS, GOMOS and SCIAMACHY agree well, while OSIRIS has a shifted phase and HIRDLS has an amplitude that is too strong.

4.11.5 NO_2 evaluations: Interannual variability

Apart from the climatological differences between the datasets it is of interest to evaluate how well the instruments detect signals of interannual variability. **Figure 4.11.16** shows the time series of daytime NO_2 mean values (upper panels) and deseasonalised anomalies (lower panels) for the tropical latitude band 20°S-20°N at 10 hPa. Datasets corresponding to 10am LST are displayed in the left panels and the original datasets 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, might also display 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_2 shows strong interannual variability dominated by an approximately two year long cycle, which is linked to the QBO-modulated transport of NO_y [*Zawodny and McCormick*, 1991]. MIPAS, OSIRIS and SCIAMACHY anomalies in the tropics agree extremely well and display the expected QBO cycle. Note that although unscaled OSIRIS mean values display strong deviations from MIPAS and SCIAMACHY, their multiplicative anomalies agree as well as the additive anomalies of the scaled datasets, creating confidence in the method applied. The scaled ACE-FTS climatology does not show a QBO signal and reveals only very little interannual variability. However, the unscaled ACE-FTS anomalies display a QBO signal very similar to the other instruments indicating that the missing interannual variability must be introduced by the scaling procedure. The scaled HIRDLS dataset is only available from June 2005 to May 2006 and does not provide any information on interannual variability. The unscaled HIRDLS climatology covers 3 years and shows similar interannual signals as the other datasets but with larger month-to-month variations. Evaluations of tropical nighttime climatologies (not shown) give very similar results.

Figure 4.11.17 shows the evaluation of the interannual anomalies of the longer time series from SAGE II and HALOE in comparison with the interannual variability of

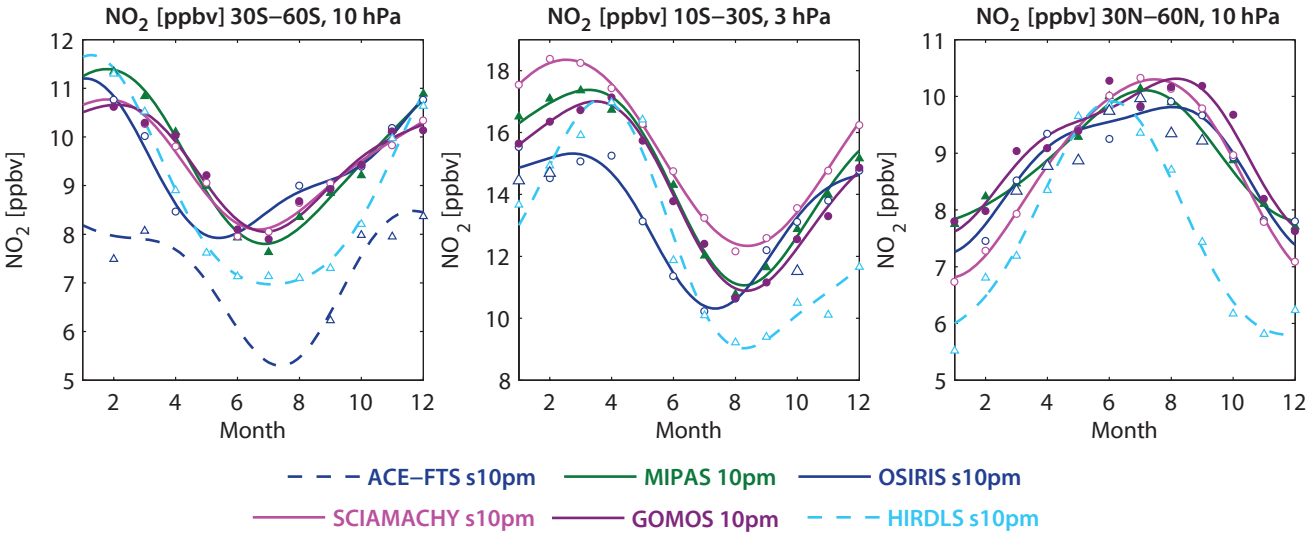


Figure 4.11.15: Seasonal cycle of night-time NO_2 for 2005-2007. *Seasonal cycle of monthly zonal mean NO_2 for 30°S-60°S at 10 hPa (left panel), 10°S-30°S at 3 hPa (middle panel) and 30°N-60°N at 10 hPa (right panel). Measurements correspond directly to 10pm LST (filled symbols) or are scaled to 10pm LST (open symbols). Note that scaled HIRDLS data are only available for June 2005 – May 2006. ACE-FTS in the SH tropics and NH mid-latitudes does not provide sufficient data coverage to estimate a fit of the seasonal cycle.*

ACE-FTS, MIPAS, OSIRIS and SCIAMACHY. Since no scaled versions of SAGE II and HALOE data are available, the comparison focuses on multiplicative anomalies of the sunrise/daytime NO_2 climatologies including the SAGE II, HALOE, and ACE-FTS local sunrise datasets as well as the MIPAS, OSIRIS, and SCIAMACHY 10am datasets. The comparison of the mean values (upper panel) shows very good agreement of MIPAS and scaled SCIAMACHY measurements, and good agreement of sunrise SAGE II, HALOE and ACE-FTS data for the overlap period 2004-2005. Scaled OSIRIS data lie between the two sets of climatologies and, surprisingly, are slightly closer to the local sunrise measurements rather than the 10am climatologies from MIPAS and SCIAMACHY, as one would expect. From 2003 onwards the multiplicative anomalies of all datasets display the expected QBO signal, with the best agreement between MIPAS, OSIRIS and SCIAMACHY. While HALOE agrees with the minimum values in 2003 and 2005 and maximum values in 2004, it shows much stronger month-to-month fluctuations. Also, while HALOE data indicate a QBO signal from 1993 to 1998, it has rather low variability and no clear QBO cycle for 1998 to 2002. SAGE II has similar month-to-month variability compared to HALOE, but shows no clear indication of a QBO signal, possibly as a result of its relatively sparse data coverage intensified by the separation into sunrise/sunset measurements. Note that the evaluation of the local sunset/night-time climatologies (**Figure A4.11.11** in *Appendix A4*) give similar results with excellent agreement of the interannual anomalies of MIPAS, SCIAMACHY, OSIRIS, and GOMOS. One difference to the local sunrise/daytime evaluations is that SAGE II and HALOE sunset climatologies agree better on their interannual variability and display the QBO signal over the whole time period.

In the mid-latitudes, NO_2 shows less interannual variability and the datasets show less good agreement (**Figure A4.11.12** in *Appendix A4*) when compared to the tropics. The largest deviations are found for GOMOS and HIRDLS, which display strong month-to-month fluctuations. The evaluation of interannual anomalies at SH polar latitudes (**Figure A4.11.13** in *Appendix A4*) is based on multiplicative anomalies calculated relative to the monthly mean values of the time period 2003-2005 (in order to include POAM III local sunrise) and on multiplicative anomalies calculated relative to the time period 2005-2007 (in order to include MIPAS). Interannual variations are most pronounced during the SH winter, but the datasets do not always agree on sign or magnitude of the anomalies. For the comparison relative to the time period 2003-2005, POAM III shows the largest anomalies, and for the comparison relative to the later time period 2005-2007, the MIPAS anomalies are strongest. Note that the evaluation of the polar anomalies might be impacted by sampling artefacts., *e.g.*, compared to SCIAMACHY, MIPAS observes higher latitudes of the SH in winter and might see less NO_2 due to polar vortex denitrification.

4.11.6 NO_2 evaluations: Downward transport of NO_2 during polar winter

In the polar mesosphere, NO_x is produced by ionizing Energetic Particle Precipitation (EPP) [*Barth*, 1992; *Solomon et al.*, 1982]. Observations have shown that inside the polar vortex NO_x is transported downwards into the stratosphere [*Funke et al.*, 2005b; *Seppälä et al.*, 2007] causing elevated NO_2 levels during polar winter with strong year-to-year variability. How well the limb-viewing satellite datasets

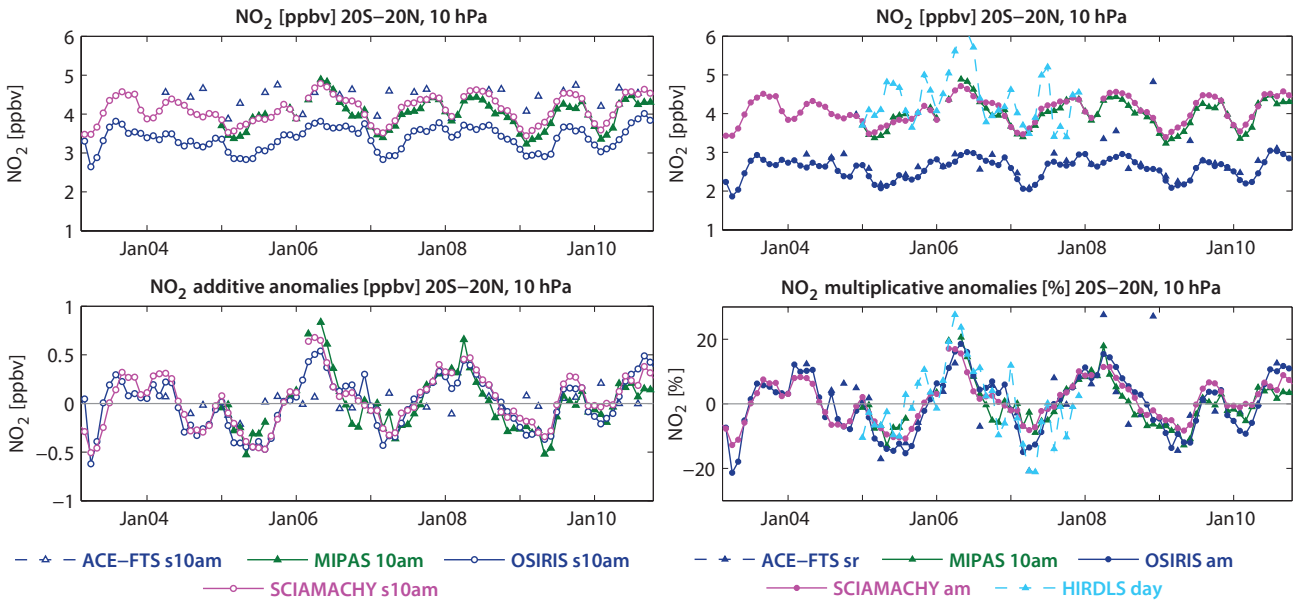


Figure 4.11.16: Time series of daytime tropical NO_2 mean values and anomalies for 2003-2010. *Monthly mean values (upper panels) and deseasonalised anomalies (lower panels) of NO_2 between 20°S – 20°N at 10 hPa. The 10am climatologies (left panel) correspond directly to 10am LST (filled symbols) or are scaled to 10am LST (open symbols). The daytime climatologies (right panel) correspond to a variety of LSTs as described in Table 4.11.3. The anomalies are calculated in an additive manner for the 10am and in a multiplicative manner for the daytime climatologies, as further explained in the text.*

agree on this phenomenon is evaluated in **Figure 4.11.18**, which shows NO_2 time series in the USLM (1 hPa) for nighttime climatologies at high NH and SH latitudes. Since at sunrise/sunset most of the atmospheric NO_x is available as NO and not NO_2, solar occultation measurements of NO_2 do not show very strong EPP signals. Therefore, the following comparison focuses on emission and star occultation measurements from HIRLDS, MIPAS and GOMOS. As the only exception, scaled ACE-FTS data are included allowing for the evaluation of the EPP signal in sunrise/sunset NO_2 by scaling it to 10pm NO_2 abundances.

In both hemispheres, measurements during the polar winter reveal very large NO_2 values (up to 10 times higher than the 10 hPa maximum in the summer hemisphere) related to the polar winter descent of NO_x. In the NH, MIPAS and GOMOS data show very good agreement on the timing and magnitude of the elevated NO_2 values for most years with stronger month-to-month fluctuations for GOMOS. In the SH, the two datasets also agree very well for the winter months, however, larger deviations are found before and after the winter, when MIPAS values decay but GOMOS values remain high or increase even further. Note that GOMOS does not provide any measurements during the SH polar summer and one can therefore not compare the winter NO_2 abundance to the annual mean state. HIRDLS measurements in both hemispheres show elevated values during the winter months very similar to MIPAS and GOMOS, except for the SH winter 2007 and the NH winter 2007/2008 when HIRDLS is low. During the rest of the year, HIRDLS NO_2 decays only slightly compared to the winter months, in contrast to MIPAS data, which is very low in the spring and summer. Since the HIRDLS nighttime climatology includes only measurements at SZA>90° with a measurement time varying around 0:30am, the absolute HIRDLS values cannot be compared directly with MIPAS or GOMOS. Very likely the larger NO_2 abundances observed by HIRDLS in summer are due to the fact that the HIRDLS climatology contains no measurements at illuminated conditions when NO_2 is rapidly transformed into NO. Based on the MIPAS NO_x climatologies in the same region we would expect the overall NO_x abundance to be ~7-8 ppbv during the NH summer 2005 (see **Figure 4.12.9** in *Section 4.12*), which is roughly consistent with the HIRDLS NO_2 values during this time. Scaled ACE-FTS data, although less frequent, shows the same signal as MIPAS data with elevated NO_2 in winter between 5 and 10 ppbv and very low NO_2 in summer with less than 1 ppbv. Note that unscaled ACE-FTS data do not reveal any elevated wintertime NO_2 levels, as expected due to the NO/NO_2 partitioning in the LM at sunrise/sunset.

The elevated NO_2 abundances caused by EPP in the mesosphere propagate downward into the upper and middle stratosphere during the polar winter as evident from MIPAS and GOMOS observations in **Figure 4.11.19**. The time-altitude cross sections show that elevated NO_2 exists for every winter but show a large interannual variability throughout the whole LM and US. Particularly strong events have been observed for 2004 and 2005 by both instruments. For the two last winters

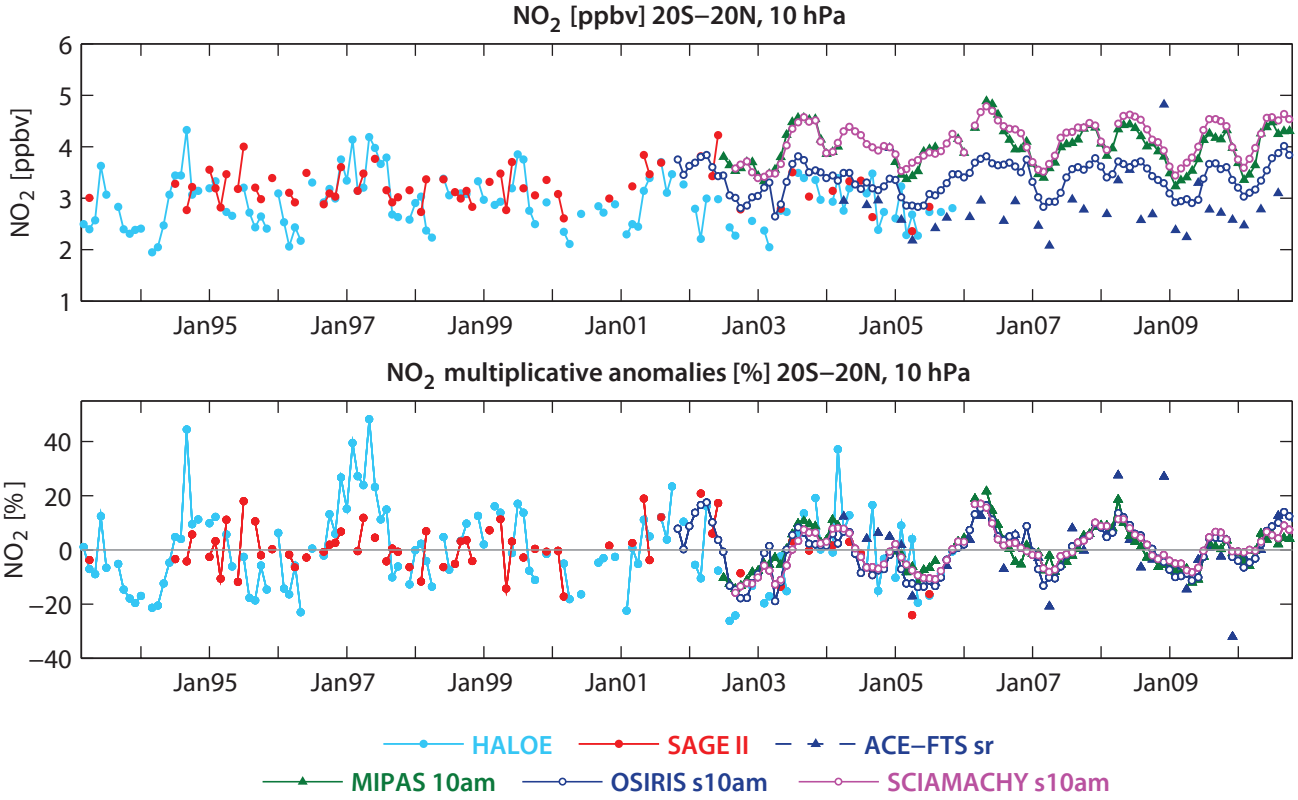


Figure 4.11.17: Time series of tropical local sunrise and daytime NO_2 mean values and anomalies for 1993-2010. *Monthly mean values (upper panels) and deseasonalised anomalies (lower panels) of NO_2 between 20°S – 20°N at 10 hPa. Datasets correspond to local sunrise or to 10am LSTs as described in the text.*

of the time series, MIPAS measurements show only slightly elevated NO_2 abundances, while the GOMOS time series includes some larger values that are, however, somewhat isolated in the noisy dataset. Differences of the two instruments to their MIM are of changing sign reaching ±50%. Overall, during periods of elevated abundances, MIPAS detects less NO_2 than GOMOS with the exception of early 2004.

Note that LIMS measurements in the NH winter 1978/1979 also reveal elevated NO_2 abundances from the LM down to 20 hPa. Due to the strong interannual variability of EPP indirect effects, LIMS observations belonging to a different period are not directly comparable to the other datasets. Additional to the elevated NO_2 in the US and LM there is an apparent excess of LIMS NO_2 at 10 hPa at high latitudes. This could partly be due to the fact that there were large wave-1 events in January 1979 that moved the vortex off the Pole and brought some NO_2 to high latitudes.

4.11.7 Summary and conclusions: NO_2

A comprehensive comparison of NO_2 profile climatologies from 12 satellite instruments (LIMS, SAGE II, HALOE, POAM II, POAM III, OSIRIS, SAGE III, MIPAS, GOMOS, SCIAMACHY, ACE-FTS, and HIRDLS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the NO_2 mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots. The first summary plot (**Figure 4.11.20**) provides information on the NO_2 mean state at local sunrise (am LSTs), local sunset (pm LSTs), 10am and 10pm. Additionally, the uncertainty derived from the spread between the datasets is given for all four illumination conditions. The second summary plot (**Figure 4.11.21**) shows specific inter-instrument differences in the form of the deviations of the instrument climatologies from 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 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

Middle stratosphere (30-5 hPa)

The uncertainty in our knowledge of the atmospheric NO_2 annual mean state is smallest in the tropical and

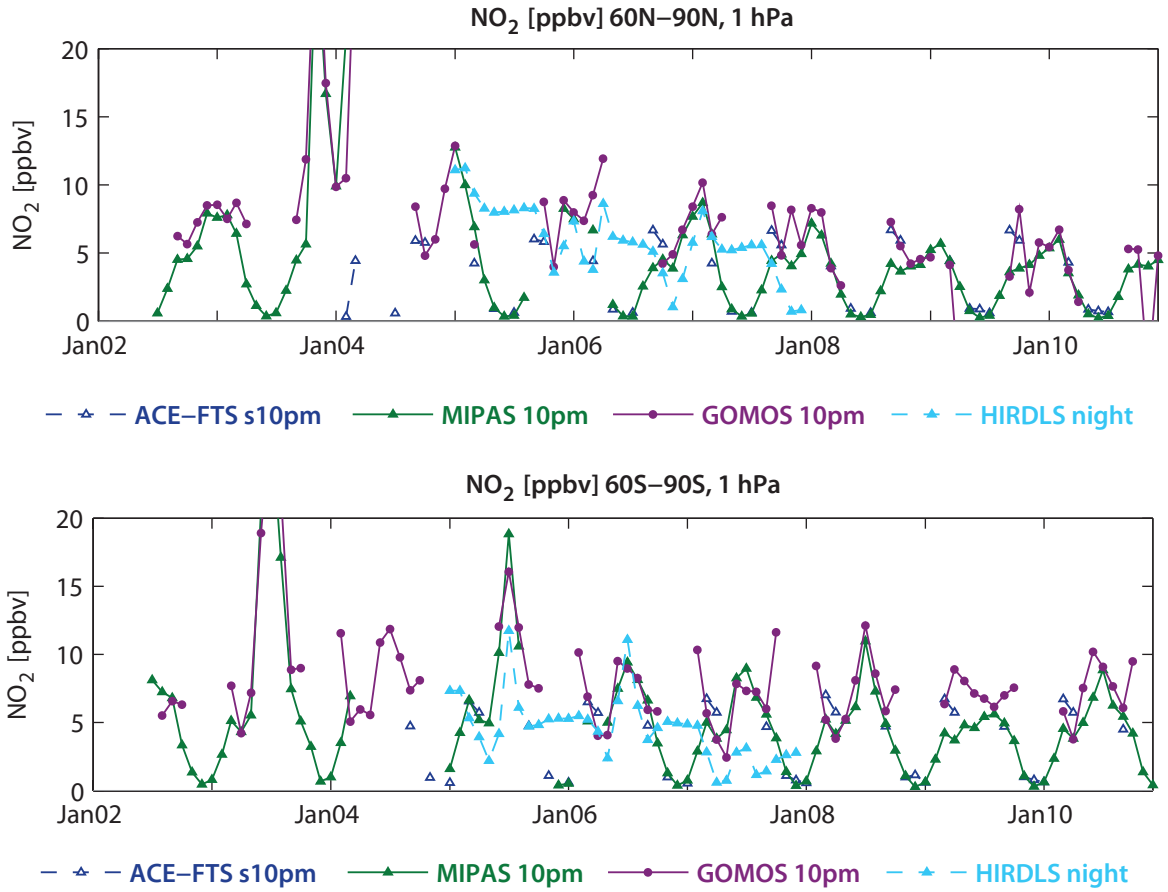


Figure 4.11.18: Time series of polar night-time NO_2 for 2002-2010. *Time series of night-time NO_2 at 1 hPa for 60°N-90°N (upper panel) and for 60°S-90°S (lower panel) from 2002 to 2010 are shown.*

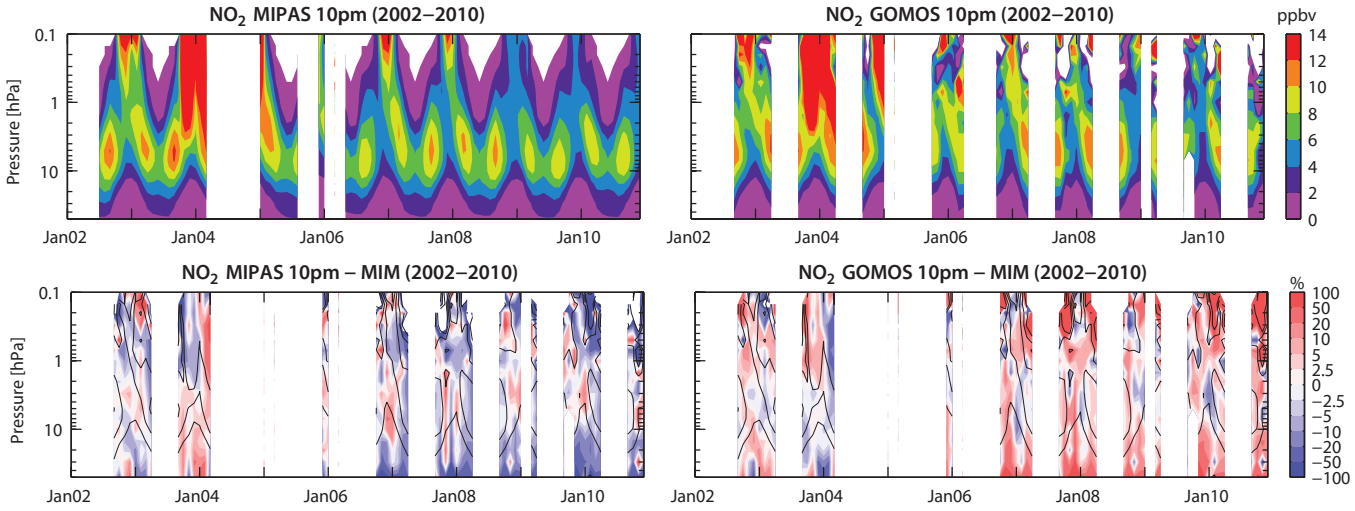


Figure 4.11.19: Altitude time evolution of NO_2 mean values and differences to the MIM in the Arctic. *Altitude-time sections of monthly zonal mean NO_2 (upper panels) and differences to the MIM (lower panels) corresponding to 10pm for 60°N-90°N from 2002 to 2010 are shown. The differences to the MIM are displayed by colour contours while the overlaid black contours show the MIM NO_2 field.*

mid-latitude MS. The evaluation of three solar occultation local sunrise/sunset datasets for 2004-2005 reveals a 1σ multi-instrument spread in this region of ±5% to ±10% (**Figure 4.11.20**, right panel). The datasets corresponding to 10am/pm LST give a slightly larger spread in the MS of ±10% to ±20%. Note that the latter comparison is based on climatologies derived by scaling measurements with a chemical box model to a common LST.

Lower stratosphere (100-30 hPa)

In the LS, the NO_2 abundances decrease quickly and the local sunrise/sunset climatologies from the solar occultation instruments show a large spread (1σ of ±50%), while a slightly better agreement is found for the daytime/nighttime climatologies (1σ of ±10% to ±50%). In particular, the 10am climatologies show a good agreement in the NH mid-latitudes down to 100 hPa.

Upper stratosphere (5-1 hPa)

In the US, the best agreement is found for the climatologies corresponding to 10pm LST, which give a multi-instrument spread in the tropics and mid-latitudes of ±5% to ±10%. Larger deviations are found for the other sets of climatologies, in particular above 2 hPa with a 1σ spread of up to ±30%.

High latitudes

At high latitudes, the instruments show larger deviations than at lower latitudes. In the MS, best agreement is found for the local sunrise datasets (1σ of ±10% to ±20%) while all other sets of climatologies give a larger uncertainty of the mean state (1σ of ±50%). In the USLM, the high latitude annual mean NO_2 abundance is dominated by the polar night NO_x descent, which is best reported by the limb emission instruments.

Instrument-specific conclusions

Local sunrise/sunset climatologies from solar occultation instruments

SAGE II, HALOE and ACE-FTS show very good to good agreement in the MS, with mean differences below ±5% for their local sunset climatologies and below ±10% for their local sunrise climatologies (**Figure 4.11.21**). Above and below this level, the relative differences increase steadily reaching mean values of up to ±20% in the US and up to ±50% in the LS. For most regions, the NO_2 local sunrise and sunset evaluations give a consistent picture, however, some differences exist (*e.g.*, for ACE-FTS and SAGE II in the tropical LS). Despite their sparser data coverage when compared to other limb sounders, all three solar occultation instruments observe the tropical QBO cycle except for the SAGE II sunrise climatologies, in which no QBO signal can be identified. One important characteristic of the solar occultation climatologies is the stronger month-to-month fluctuations. A comparison of the long-term climatologies from SAGE II and HALOE (1992-2005) leads to similar results as the comparison of their short-term climatologies (2004-2005) discussed above.

- When compared to the other solar occultation datasets, **SAGE II** detects mostly larger NO_2 abundances, with the exception of the MS local sunrise climatologies, for which SAGE II is lowest. In the US, a relatively large regional spread (over all US grid points) of the SAGE II differences is found, indicating individual monthly mean differences larger than +50%.
- **HALOE** detects less NO_2 than the other two instruments except around 10 hPa where, depending on the latitude bin, it can show larger values. Evaluations of the HALOE local sunrise and sunset climatologies give consistent results.

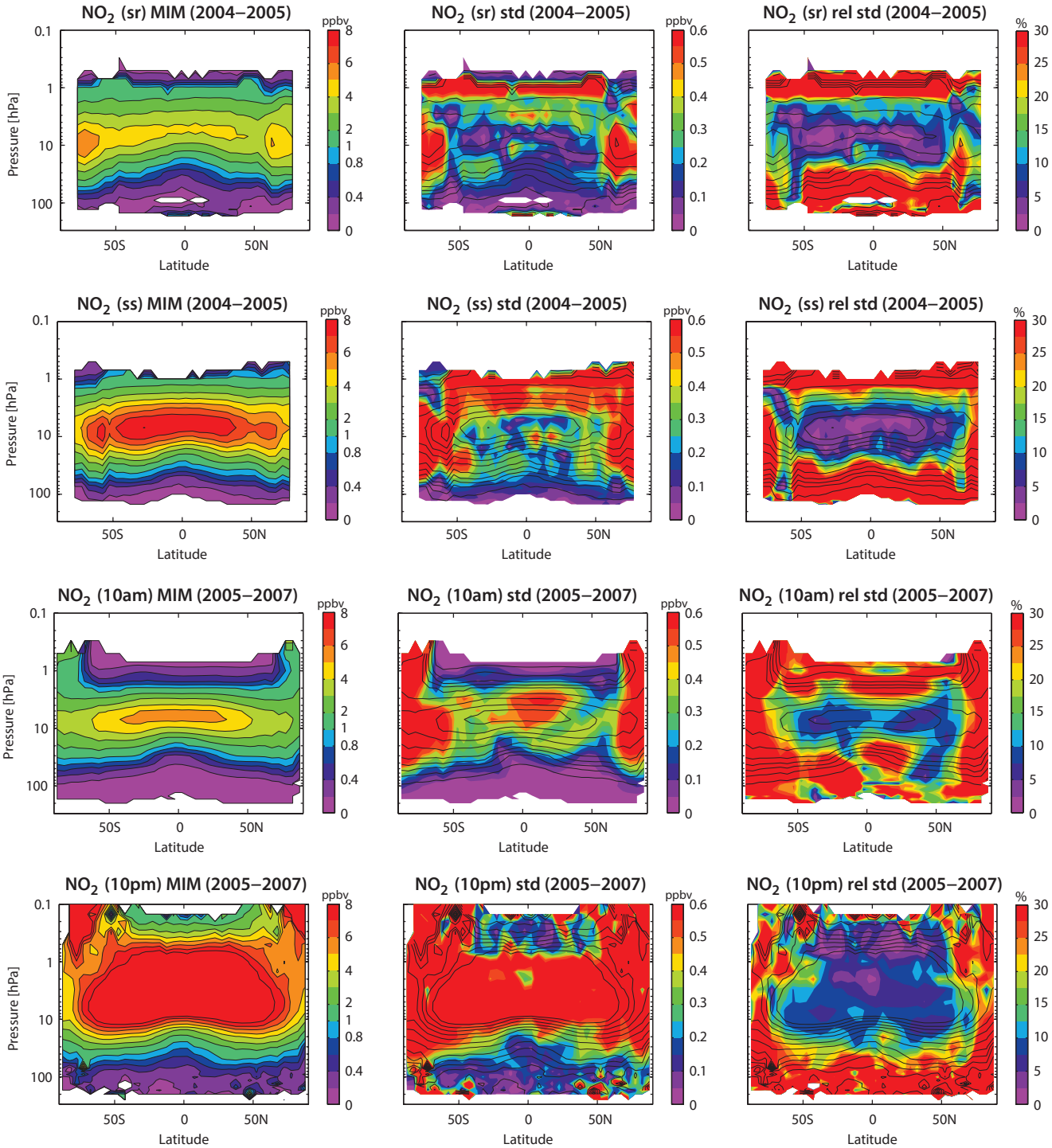


Figure 4.11.20: Summary of NO_2 annual zonal mean state for 2004-2005 and 2005-2007. *Annual zonal mean cross sections of the NO_2 MIM are shown in the left panels for local sunrise (upper row), local sunset (second row), 10am (third row) and 10pm (lower row) illumination conditions. The local sunrise and sunset mean state for 2004-2005 are based on SAGE II, HALOE and ACE-FTS. The mean state at 10am for 2005-2007 is based on MIPAS at 10am and OSIRIS, SCIAMACHY, ACE-FTS, and HIRDLS scaled to 10am. The mean state at 10pm for 2005-2007 is based on MIPAS and GOMOS at 10pm and OSIRIS, SCIAMACHY, ACE-FTS, and HIRDLS scaled to 10pm. Additionally, for all four illumination conditions, the standard deviation over all respective 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 MIM and standard deviation are only displayed for regions where at least two instruments provide measurements.*

- The **ACE-FTS** NO_2 local sunset climatologies are in the mid-range between the other two instruments. ACE-FTS local sunrise climatologies in the LSMS, however, display largest NO_2 abundances.
- **SAGE III**, **POAM II**, and **POAM III** provide measurements over a narrow range at high latitudes. The POAM II climatology for 1994-1996 reports smaller values than SAGE II and HALOE with differences from

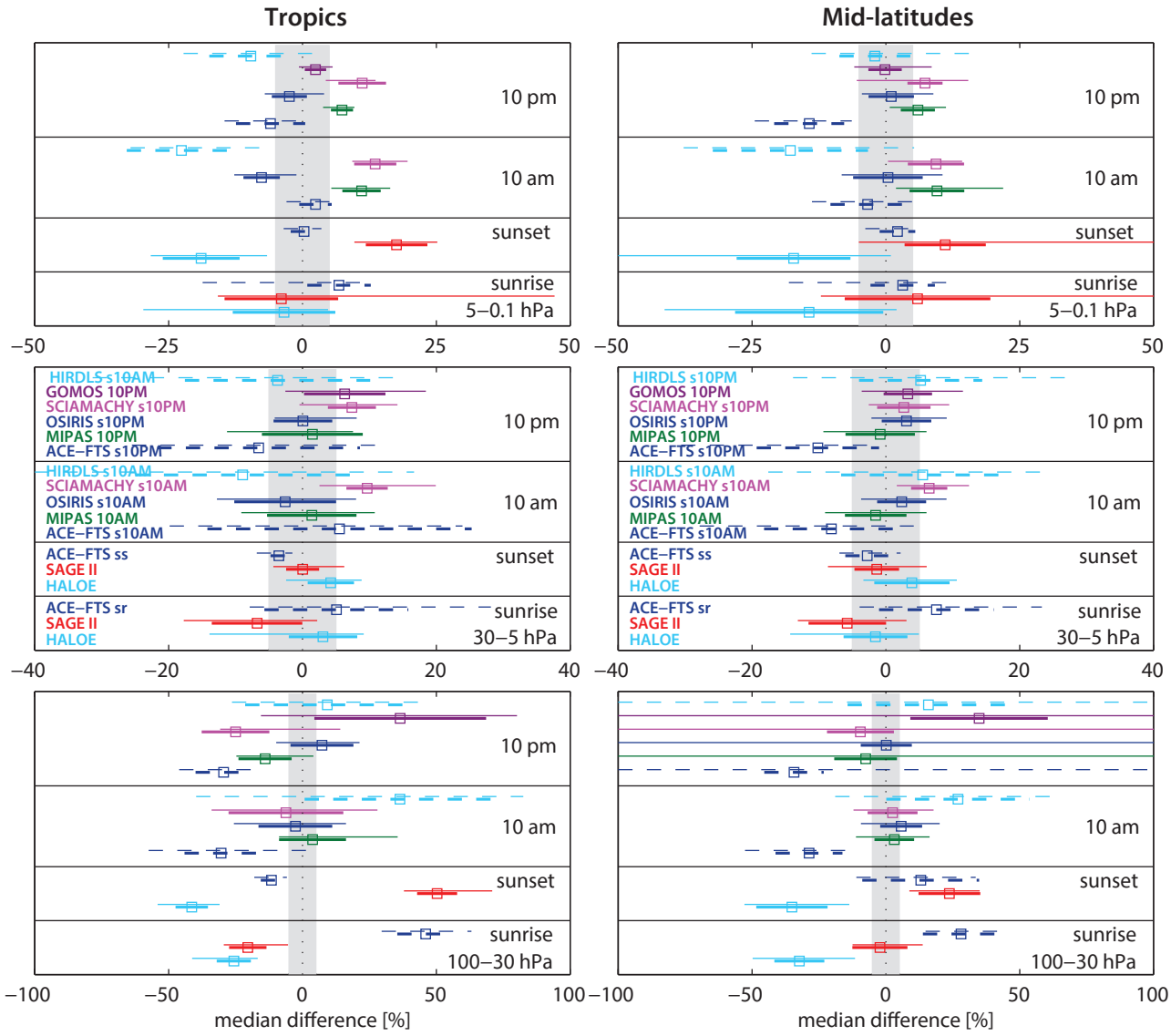


Figure 4.11.21: Summary NO_2 differences for 2004-2005. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the 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 three different altitude regions from the UT up to the LM between 300 and 0.1 hPa for the reference period 2004-2005.*

the MIM of around -10%. The comparison of POAM III and SAGE III to the other instruments for 2004-2005 yields a good agreement in the MS with differences often below ±5% except for a divergence between ACE-FTS and SAGE III in March (±10%). In general, SAGE III is similar to SAGE II and shows larger NO_2 values than the other datasets, while POAM III mostly resides in the mid-range.

10am/pm climatologies

The limb emission and scattering instruments MIPAS, OSIRIS, SCIAMACHY, and HIRDLS are evaluated based on their 10am/pm climatologies, with the latter three derived from scaling with a chemical box model. Additionally, a 10pm climatology from the stellar occultation instrument GOMOS and 10am/pm climatologies from the scaled local sunrise/sunset measurements of the solar occultation instrument ACE-FTS are included in the evaluation. All climatologies show a good agreement in the MS with mean differences of ±5% to ±10%. In particular MIPAS, GOMOS, OSIRIS and SCIAMACHY agree very well with differences below ±5% and in some cases even below ±2% (*e.g.*, for 10pm climatologies in mid-latitudes). In the LS, overall mean differences can be as large as ±40%, however, MIPAS, OSIRIS, and SCIAMACHY are very close to each other with differences among them in the range of ±5% for most cases (±20% for the 10pm climatologies in the tropics). In the US, the inter-instrument spread is about ±25%, with the lower end of the range given by HIRDLS or ACE-FTS depending on latitude and LST. OSIRIS and GOMOS in the middle range agree very well (±2%) and the same is true for MIPAS and SCIAMACHY (±2%) at the upper end of the measurement range. Monthly zonal mean cross sections (**Figures 4.11.9 – 4.11.12**) reveal that for most climatologies the deviations from the MIM can change sign

depending on the latitude band and month. All 10am/pm climatologies show the tropical QBO signal with the best agreement found between MIPAS, OSIRIS, SCIAMACHY and GOMOS.

- **MIPAS** measurements correspond directly to 10am/pm and have not been scaled for the evaluations presented in this chapter. The MIPAS climatologies, when compared to other datasets, are mostly in the middle range, with slightly larger NO_2 abundances and only small deviations from the MIM of up to +5% in MS and up to +15% in the US. Only in the LS, MIPAS shows negative deviations (-15%).
- The **OSIRIS** climatologies, based on measurements scaled to 10am/pm, agree very well with the MIM. Largest deviations are found in the tropical US where OSIRIS data corresponding to 10am show a mean difference of -7%. In the LS, a relatively large spread between all datasets is observed with OSIRIS in the middle range.
- The scaled **SCIAMACHY** climatologies agree overall very well with MIPAS and OSIRIS. Notable differences occur in the tropical US, where SCIAMACHY sets the upper end of the range of climatological values with mean differences from the MIM of +15%. In the tropical LS, SCIAMACHY values scaled to 10pm are on the low side with deviations from the MIM of up to -25%. Note that the SCIAMACHY daytime climatology does not show smaller differences to the MIM than the evaluation of the night-time climatologies (except for the tropical LS) although the latter is based on a scaling to completely different illumination conditions.
- **GOMOS** stellar occultation measurements are only available in the night. The GOMOS 10pm climatology is in the middle range of the other measurements in the US, where they compare very well with OSIRIS and also in the MS, where they are very close to MIPAS, OSIRIS and SCIAMACHY. In the LS, however, GOMOS observes more NO_2 than all other datasets resulting in large mean deviations from the MIM of up to +40%. Additionally, large deviations of GOMOS from the MIM in the mid-latitudes below 100 hPa cause a large inter-instrument spread whenever GOMOS measurements are present and therefore large variations of the differences of the other instruments from the MIM over the mid-latitude LS. GOMOS interannual anomalies in the mid-latitudes are characterised by strong month-to-month fluctuations.
- Scaled **HIRDLS** data agree well with the other datasets in the mid-latitude MS. Here, mean differences are less than 5%, however, a relatively wide regional spread (over all MS mid-latitude grid points) of the differences is found, indicating individual monthly mean differences larger than ±20%. In the US, there is a notable difference between HIRDLS 10am and HIRDLS 10pm climatologies with the latter agreeing quite well (up to -10%) while the first show mean differences to the MIM of up to -30%. The reverse is true for the LS, where the HIRDLS 10pm climatologies are in the middle range while the 10am climatologies show large positive differences of up to +40%. The NO_2 seasonal cycle derived from HIRDLS data often show stronger amplitude than suggested by the other instruments; the HIRDLS interannual anomalies also generally exhibit larger month-to-month fluctuations than the other observations.
- In the US and tropical MS, scaled 10am **ACE-FTS** data show a good agreement with the other datasets with small differences from the MIM (1% in the US, 5% in the MS). Scaled 10pm ACE-FTS data, on the other hand, are low and close to HIRDLS in the tropical MS/US (-5%) and well below all other datasets in the mid-latitude US. This inconsistency between ACE-FTS 10am and 10pm climatologies is not observed in the LS and mid-latitude MS, where both climatologies are always on the low side with differences of up to -40% in the LS.

Comparing local sunrise/sunset measurements and 10am/pm climatologies

The ACE-FTS climatologies are available in unscaled form where they can be compared to local sunrise/sunset climatologies from solar occultation instruments, and in scaled form where they can be compared to the 10am/pm climatologies from limb emission and scattering instruments. In the tropical MS, ACE-FTS agrees well with the 10am/pm climatologies and with the sunrise/sunset climatologies with differences up to ±5% (except for SAGE II local sunrise data). This agreement suggests that all available measurements at different LSTs are consistent with each other in this region. In the mid-latitude MS, ACE-FTS agrees quite well with SAGE II and HALOE but is on the lower side of the 10am/pm climatologies with differences of up to -10%. One needs to keep in mind that such differences could have been introduced by scaling the ACE-FTS data. However, if one assumes no errors from the scaling, then the solar occultation instruments would observe less NO_2 than the emission and scattering instruments in the mid-latitude MS. In the US, the same approach would place SAGE II measurements in the middle range with slightly positive differences and would give negative difference for HALOE versus most other datasets. In the LS, SAGE II and HALOE would both be on the low side when compared to the other instruments *via* ACE-FTS (except for SAGE II sunset data).

4.12 Nitrogen oxides – NO_x

Nitric oxide (NO) and nitrogen dioxide (NO_2) are together known as the nitrogen family NO_x. Sources of tropospheric NO_x include fossil fuel burning, lightning, chemical processes in soils, and biomass burning (see *Sections 4.10* and *4.11*). The primary source of NO_x in the stratosphere is the oxidation of N_2O also originating from soil emissions (see *Section 4.4*), which is transported from the troposphere into the stratosphere. NO_x is an efficient catalyst for the destruction of stratospheric ozone [*Crutzen*, 1970; *Johnston*, 1971]. While chlorine- and bromine-containing halocarbons have been successfully reduced by the Montreal Protocol and its Amendments and Adjustments,

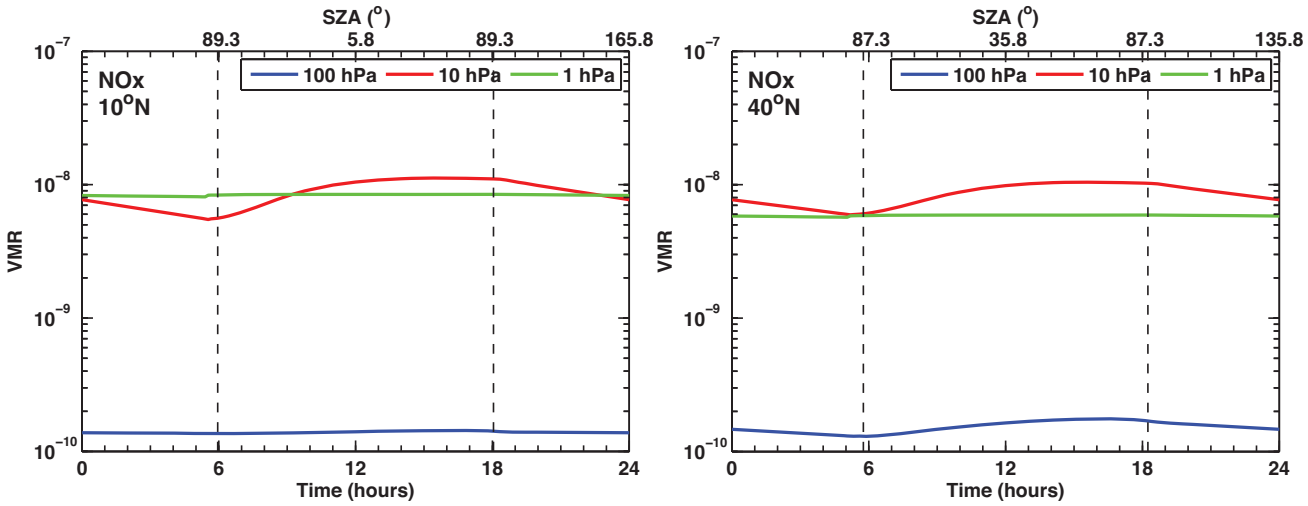


Figure 4.12.1: Diurnal NO_x cycle. *NO_x variations as function of LST are shown at 10°N and 40°N for at 1, 10 and 100 hPa for March 15.*

N_2O emissions are unregulated and are expected to become the most important ozone-depleting emission during the 21st century [*Ravishankara et al.*, 2009] on the basis of NO_x driven catalytic ozone loss. The second important source of stratospheric NO_x is the enhancement of upper atmospheric NO_x through ionizing Energetic Particle Precipitation [*Barth*, 1992; *Solomon et al.,* 1982] and the NO_x downward transport inside the polar vortex [*Funke et al.*, 2005b; *Seppälä et al.*, 2007]. This mechanism is enabled by the absence of solar radiation during the polar winter when chemically long-lived NO_x is dominated by atmospheric transport and has been suggested to impact global ozone variability [*e.g.*, *Vogel et al.*, 2008; *Reddmann et al.*, 2010].

Stratospheric NO_x, given as the sum of NO_2 and NO, displays a weak diurnal cycle in the LS and MS. **Figure 4.12.1** shows the diurnal NO_x cycle as a function of LST for three different pressure levels as derived from a chemical box model [*McLinden et al.*, 2010].

4.12.1 Availability of NO_x measurements

HALOE, ACE-FTS and MIPAS measure both NO and NO_2 and can therefore provide vertically resolved NO_x climatologies based on the sum of the two trace gases. OSIRIS and SCIAMACHY measure NO_2 but not NO. For these two instruments the NO_x climatologies are compiled based on their NO_2 measurements and on NO profiles derived from a chemical box model [*McLinden et al.*, 2010]. Both of these NO_x datasets as well as ACE-FTS data are afterwards scaled to 10am/pm in order to be comparable to MIPAS measurements. For OSIRIS, the scaling of NO_2 to NO_x at 10am/pm is done profile-by-profile with the photochemical box model initialised with measured trace gas abundances and temperature. The scaling for SCIAMACHY and ACE-FTS on the other hand is performed based on lookup tables calculated from the photochemical box model initialised with climatological inputs (see *Section 3.1.2*). In addition to the climatologies corresponding to 10am/pm, solar occultation measurements from HALOE and ACE-FTS can be compared directly if separated into local sunrise and sunset data. The two instruments overlap during 2004-2005.

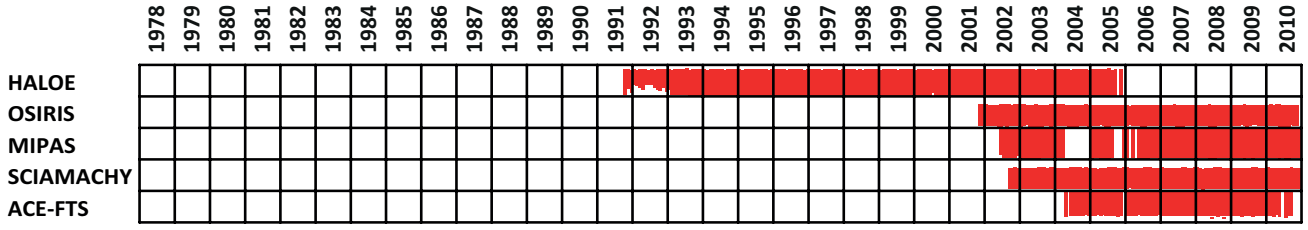
Table 4.12.1 summarises all available NO_x climatologies derived from satellite measurements including time period and vertical range. Information on NO and NO_2 measurements used for deriving the NO_x climatologies, including time period, vertical range and resolution, and relevant references are given in *Sections 4.10* and *4.11*.

4.12.2 NO_x evaluations: Zonal mean cross sections

HALOE and ACE-FTS (2004-2005)

For the time period 2004-2005, the long NO_x time series from HALOE overlaps with the more recent time series from ACE-FTS. **Figure 4.12.2** shows the HALOE, and

Table 4.12.1: Available NO_x datasets based on measurements 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.*



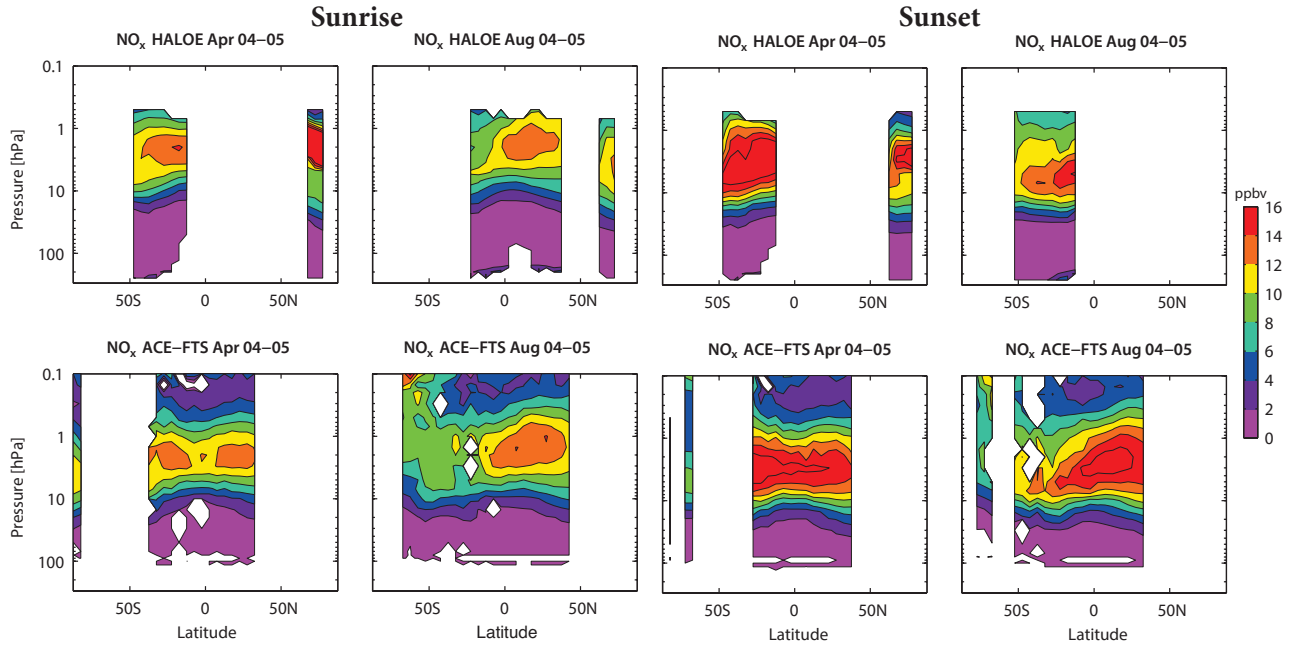


Figure 4.12.2: Cross sections of monthly zonal mean, local sunrise and sunset NO_x for 2004-2005. *Monthly zonal mean, local sunrise (column 1 and 2) and sunset (column 3 and 4) NO_x cross sections for April and August are shown for HALOE (upper row) and ACE-FTS (lower row).*

ACE-FTS local sunrise and sunset NO_x climatologies for April and August, while **Figure 4.12.3** shows the differences of both instruments from their MIM. Note that we use the comparison of both datasets to their MIM (and not a direct comparison) in order to stay consistent with other parts of the report. Departures from the mean in the MS and US are small and mostly below ±10%. HALOE shows larger values than ACE-FTS in the MS (and also in the US for April) and lower values otherwise. In the tropical LS, the relative differences increase to up to ±50%. Evaluations for the rest of the year agree in general with the April and August differences. However, the vertical extent of the region where HALOE shows larger NO_x values compared to ACE-FTS can differ with season and latitude and can sometimes extend through the whole MS and US. Note that there is less overlap between the two instruments for the other months of the year. While some month-to-month variations of the differences exist, the deviations between the local sunrise measurements are consistent with the comparison of the local sunset measurements over the same month and latitude.

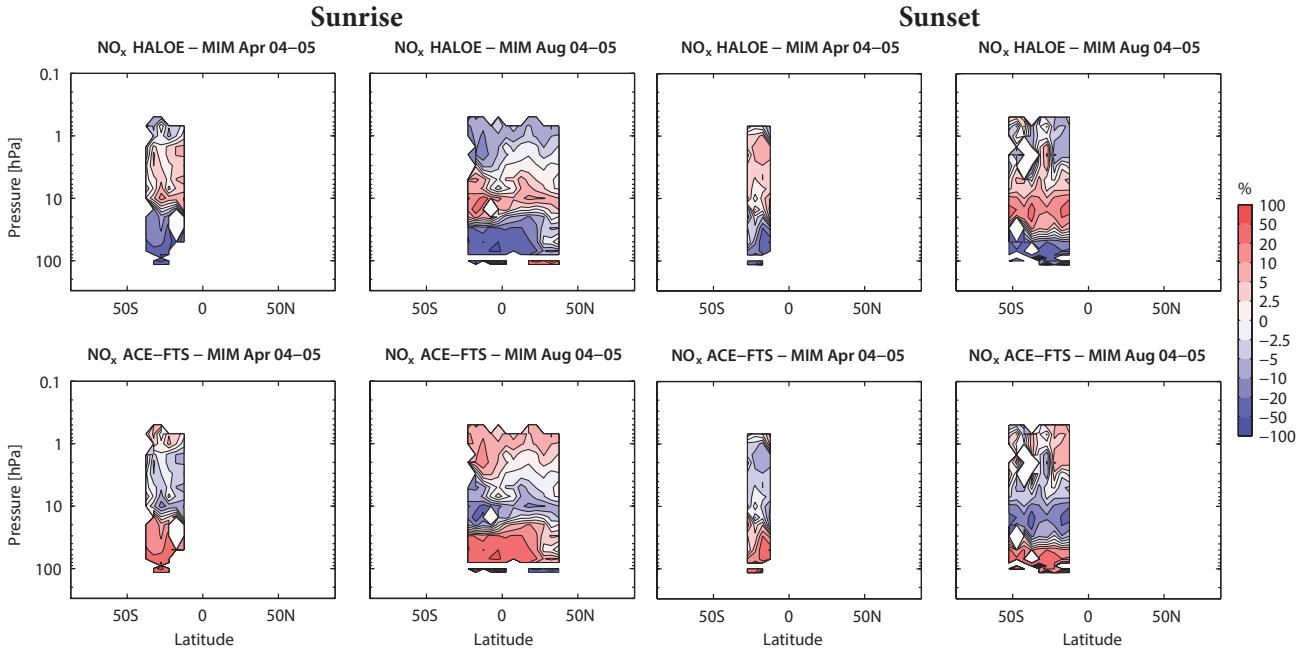


Figure 4.12.3: Cross sections of monthly zonal mean, local sunrise and sunset NO_x differences for 2004-2005. *Monthly zonal mean, local sunrise (column 1 and 2) and sunset (column 3 and 4) NO_x differences for April and August between the individual instruments (HALOE and ACE-FTS) and their MIM are shown.*

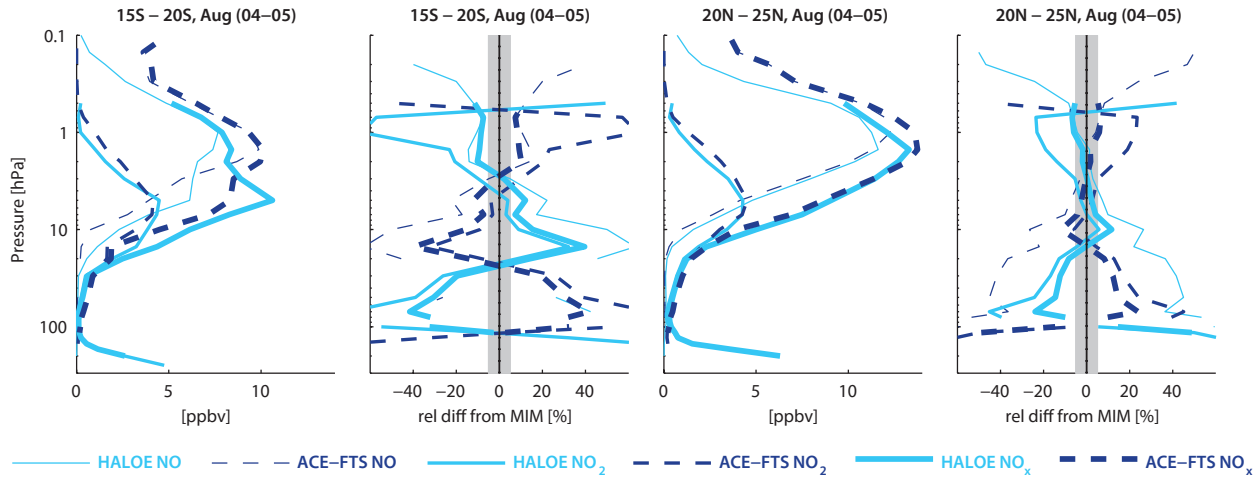


Figure 4.12.4: Profiles of monthly zonal mean, local sunrise NO, NO_2, and NO_x for 2004-2005. *Zonal mean NO, NO_2, and NO_x profiles for 15°S-20°S and 20°N-25°N for August are shown together with their differences from the MIM.*

For both local sunrise and sunset measurements, the evaluation of NO_2 from HALOE and ACE-FTS shows in general the same features as the evaluation of NO_x from the two instruments, as illustrated in **Figure 4.12.4** for individual profile comparisons. For NO_2, HALOE is larger than ACE-FTS only between 20 and 3 hPa but lower otherwise, in agreement with the results from the NO_x deviations. In addition to the NO_2 and NO_x profiles, the corresponding NO data are shown in **Figure 4.12.4**. NO differences are consistent with NO_x in the US and MS, but not in the LS where HALOE measures more NO, but less NO_2 and NO_x.

OSIRIS, SCIAMACHY, MIPAS, and ACE-FTS (2005-2010)

Figure 4.12.5 shows the NO_x 10am and 10pm climatologies for August 2005-2010. The datasets correspond to either 10am or 10pm LST, with scaled ACE-FTS, OSIRIS and SCIAMACHY (labelled as s10am and s10pm in the figure titles), and can be directly compared to each other. The four datasets corresponding to 10am or 10pm LST, respectively, show a similar NO_x distribution. However, some differences exist (*e.g.*, different meridional gradients around 10 hPa for SCIAMACHY). MIPAS 10am and 10pm measurements during the polar night above 10 hPa reveal very large NO_x values related to the polar winter descent of NO_x produced by energetic particle precipitation in the upper atmosphere.

Differences of the individual datasets from their respective MIM for August 2005-2010 are displayed in **Figure 4.12.6**. In general, the differences for the 10am climatologies in the MS are below ±10%, and reach values of ±10 to ±20% only in some regions, similar to the comparison of the solar occultation instruments. MIPAS is in the middle of the measurement range, with deviations in the MS of only up to ±5%. ACE-FTS and OSIRIS display opposite structures in their deviations from the MIM, with strong negative (positive) deviations in the SH extra-tropics. The structure of the differences is similar to that found for the comparison of the NO_2 climatologies (see **Figure A4.12.1** in *Appendix A4*). However, the NO_x climatologies, in particular the ones from MIPAS, agree better than the NO_2 climatologies and show smaller differences from their respective MIM.

NO_x 10pm climatologies show deviations of around ±10 to ±20% in the MS. Overall, the deviations are consistent with the ones derived for the 10am climatologies. While the results for SCIAMACHY and OSIRIS are very similar, some inconsistencies can be observed for MIPAS, for which negative deviations of -20(50)% are found in the MS(LS) that do not exist for the 10am climatologies. Since most of the 10pm NO_x is present as NO_2, the differences of the NO_x climatologies resemble the differences of the NO_2 climatologies (see **Figure A4.12.1** in *Appendix A4*). Not only do they display the same structure but they are also of similar or slightly larger magnitude, in contrast to the 10am climatologies where the relative differences for NO_x are smaller than those for NO_2. A comparison of climatologies for other months reveals similar differences (see **Figures A4.12.2** and **A4.12.3** in *Appendix A4*). The yearly mean climatologies and their relative differences with respect to the MIM (see **Figures A4.12.4** and **A4.12.5** in *Appendix A4*) show a similar structure but overall larger differences.

For both 10am and 10pm climatologies, the departure of unscaled datasets (*e.g.*, unscaled OSIRIS data corresponding to a variety of LSTs) from the MIM is quite different to that of the scaled data. This indicates that, although the diurnal cycle is weaker for NO_x than for NO or NO_2, the NO_x climatologies corresponding to the different LSTs cannot easily be compared unless the dependence on the LST is taken into account. A comparison of the two local sunrise/sunset datasets from HALOE and ACE-FTS to the four datasets corresponding to 10am/pm LST will be complicated by the diurnal cycle, and deviations cannot be attributed to actual measurement differences. Therefore, such a comparison is not provided here.

4.12.3 NO_x evaluations: Seasonal cycles

NO_x exhibits a strong seasonal cycle due to the effects of sunlight on the partitioning of the nitrogen family.

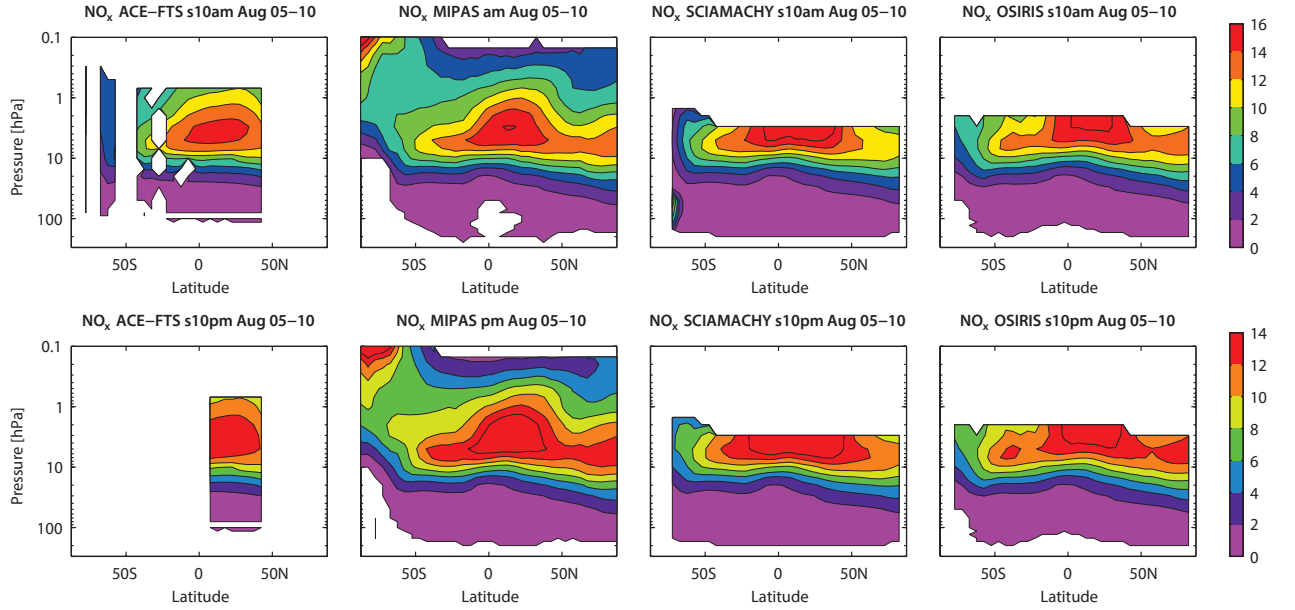


Figure 4.12.5: Cross sections of monthly zonal mean NO_x for August 2005-2010. *Monthly zonal mean NO_x cross sections of 10am (upper panels) and 10pm (lower panels) climatologies for August 2005-2010 are shown.*

Figure 4.12.7 displays the seasonal cycle of the 10am (upper panels) and 10pm (lower panels) NO_x climatologies for NH and SH mid-latitudes and tropics.

In the SH mid-latitudes, all four datasets agree very well on the seasonal cycle, with maximum values in summer and minimum values in winter. The best agreement is found for MIPAS and SCIAMACHY, which both show the same amplitude and phase, in particular for the 10am climatologies. For OSIRIS, the phase is shifted by one month (with a minimum in May) due to missing data coverage in June and July, as well as slightly elevated values in August compared to the other two datasets. The phase shift is more pronounced for the OSIRIS 10pm climatologies. ACE-FTS mean values are lower and show a larger spread around the fitted seasonal cycle, but still display very similar amplitude and phase. Note that for the ACE-FTS 10pm climatology there are no data during the SH autumn and winter, and only the September value constrains the seasonal cycle. The resulting amplitude agrees very well with that based on 10pm SCIAMACHY data, with both datasets producing an amplitude that is approximately 10% larger than that detected by MIPAS and OSIRIS.

In the NH mid-latitudes, the 10am MIPAS, OSIRIS and SCIAMACHY climatologies show the same annual cycle, with only slight differences in the amplitude of the seasonal signal. While OSIRIS has the smallest and SCIAMACHY the largest amplitude, MIPAS is in the middle range between the two instruments. ACE-FTS at 10am agrees on the

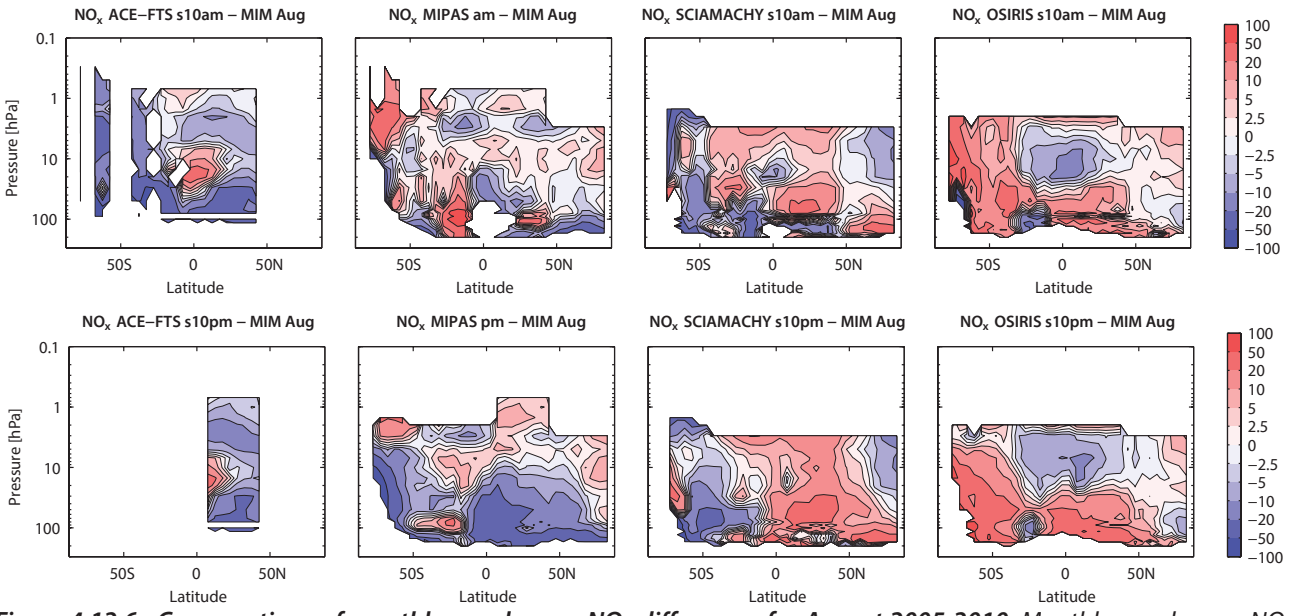


Figure 4.12.6: Cross sections of monthly zonal mean NO_x differences for August 2005-2010. *Monthly zonal mean NO_x differences from the MIM of 10am (upper panels) and 10pm (lower panels) climatologies for August 2005-2010 are shown. The MIM is based on all displayed climatologies corresponding to the respective LST.*

general structure of the seasonal signal but shows a larger amplitude due to low values in January and November. The 10pm datasets all show a very similar phase, including ACE-FTS. Here, the largest deviations to the evaluation of the 10am climatologies are the pronounced differences in the amplitude of the seasonal cycle. SCIAMACHY observes an amplitude four times larger than ACE-FTS, while MIPAS and OSIRIS agree very well in the middle range.

In the SH tropics, all four instruments detect a seasonal cycle with maximum values in the SH summer/early autumn, however they disagree on the details of the cycle. SCIAMACHY has larger mean values than the other datasets for most of the year (particularly pronounced for the 10pm climatologies) but a seasonal cycle very similar to the one detected by MIPAS. OSIRIS shows a slightly smaller amplitude than SCIAMACHY and MIPAS and a flattened maximum extending over 5 months. The largest deviations of the 10am seasonal cycle are found for ACE-FTS, which does not show the expected minimum in August but detects lowest NO_x in November. ACE-FTS at 10pm does not provide sufficient data to fit a seasonal cycle.

For SCIAMACHY, large deviations are found in the US at high latitudes (**Figure A4.12.6** in *Appendix A4*), where it does not produce the minimum in winter as observed by the other datasets. Note that this deficiency is not related to the choice of a wide latitude band (60°-90°) as similar results are obtained by analysing narrower latitude bands (60°-65°).

4.12.4 NO_x evaluations: Interannual variability

In addition to the absolute differences between the climatologies, it is of importance to evaluate how well the instruments detect signals of interannual variability. **Figure 4.12.8** shows the time series of 10am and local sunrise NO_x mean values (upper panels) and deseasonalised anomalies (lower panels) from 2003 to 2010. Datasets corresponding to 10am LST (MIPAS, scaled OSIRIS, scaled SCIAMACHY) and the two local sunrise datasets (ACE-FTS and HALOE) are compared in the tropics (20°S-20°N) at 10 hPa and NH mid-latitudes (40°N-50°N) at 7 hPa. The anomalies of the climatologies are calculated in a multiplicative sense as

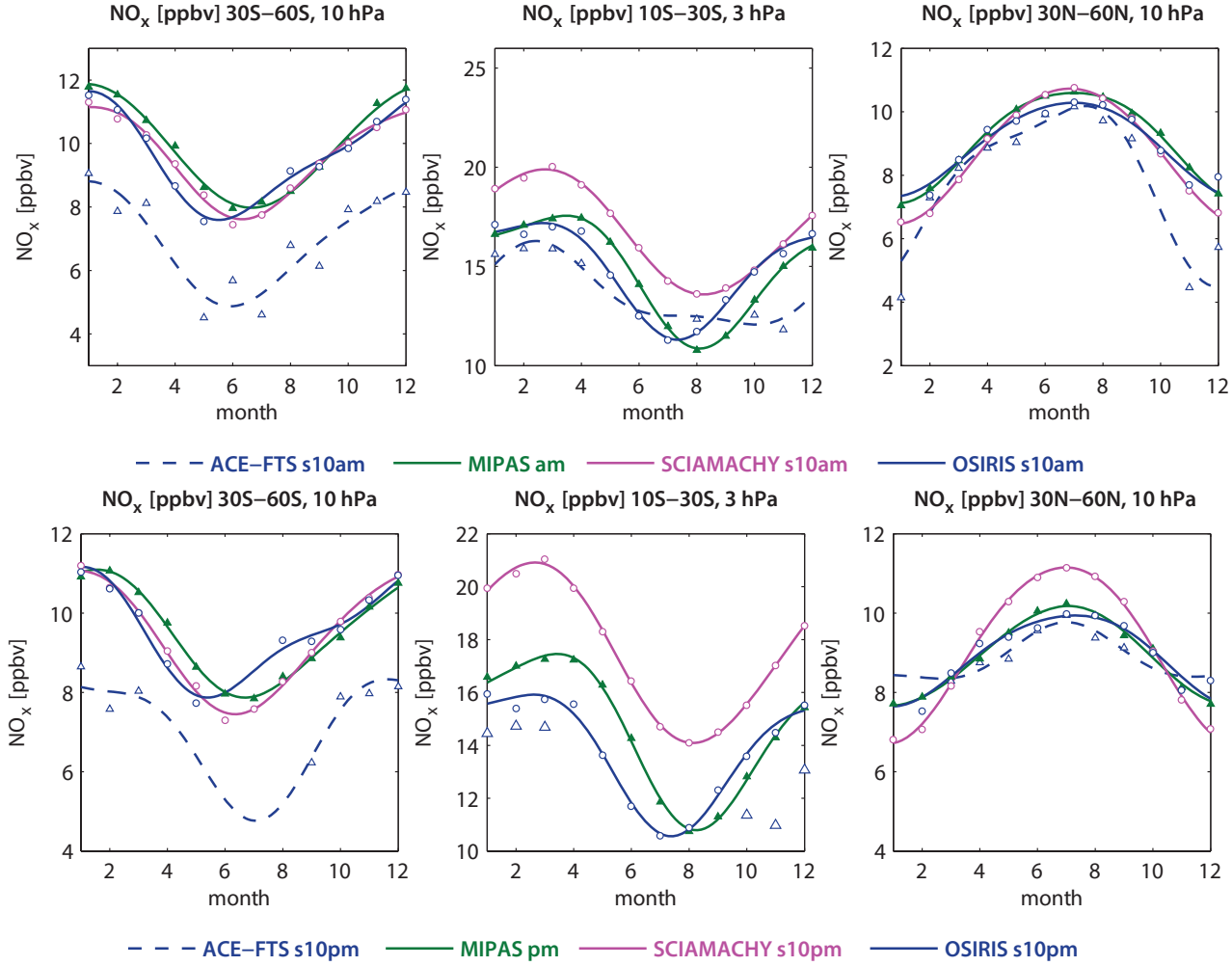


Figure 4.12.7: Seasonal cycle of 10am and 10pm NO_x for 2005-2010. *Seasonal cycle of monthly zonal mean NO_x for 30°S-60°S at 10 hPa (left column), 10°S-30°S at 3 hPa (middle column) and 30°N-60°N at 10 hPa (right column) for 10am (upper row) and 10pm (lower row) climatologies. Measurements correspond directly to 10am/pm LST (filled symbols) or are scaled to 10pm/am LST (open symbols). ACE-FTS scaled to 10pm in the SH tropics does not provide sufficient data coverage to estimate a fit of the seasonal cycle.*

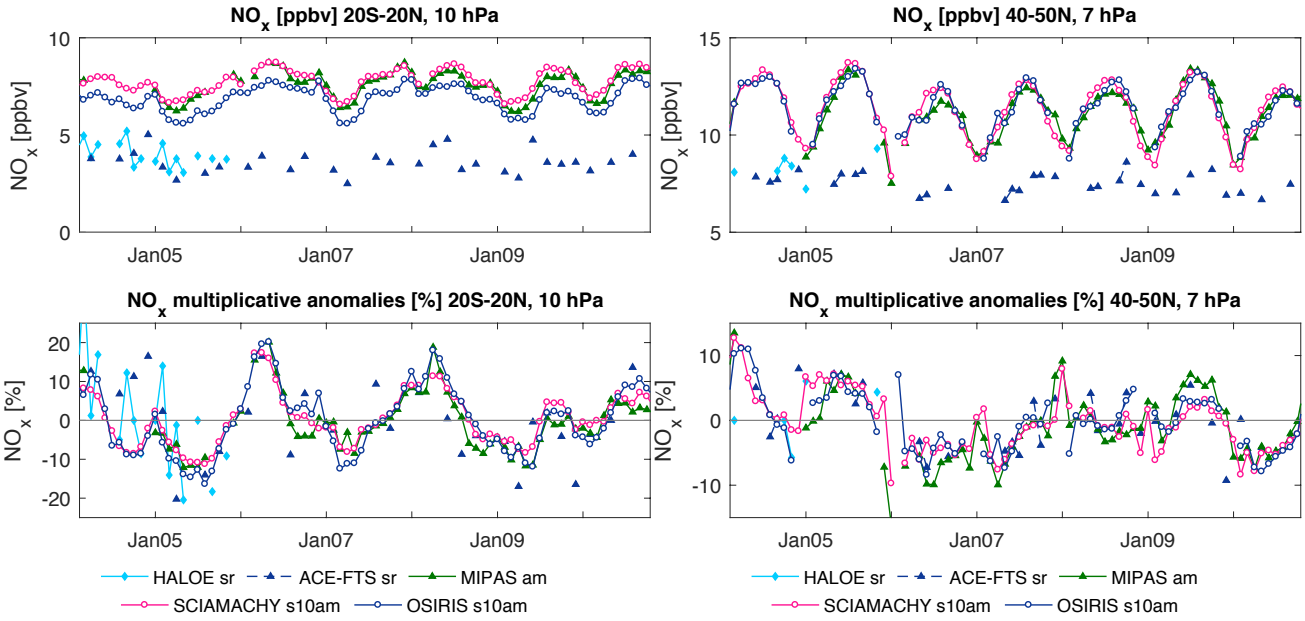


Figure 4.12.8: Time series of 10am and local sunrise NO_x mean values and anomalies for 2003-2010. *Monthly mean values (upper panels) and deseasonalized anomalies (lower panels) of NO_x at 10 hPa for 20°S – 20°N and at 7 hPa for 40°N-50°N. Measurements correspond to local sunrise conditions (ACE-FTS sr, HALOE sr), to 10am LST (MIPAS am) or are scaled to 10am LST (SCIAMACHY s10am, OSIRIS s10am).*

percent deviations from the monthly multi-year mean values, a quantity that is less affected by the diurnal variations than anomalies calculated in an additive sense.

In the tropics, NO_x is dominated by an approximately two year long cycle which is linked to the QBO (see *Section 4.11*). MIPAS, OSIRIS, and SCIAMACHY anomalies in the tropics agree very well and display the expected QBO cycle. Unscaled ACE-FTS and HALOE data show a QBO signal similar to the other datasets but also exhibit stronger month-to-month variability. The interannual variability of the scaled ACE-FTS data (not shown here) do not agree as well with the other datasets compared to the unscaled ACE-FTS data, similar to the results for NO_2.

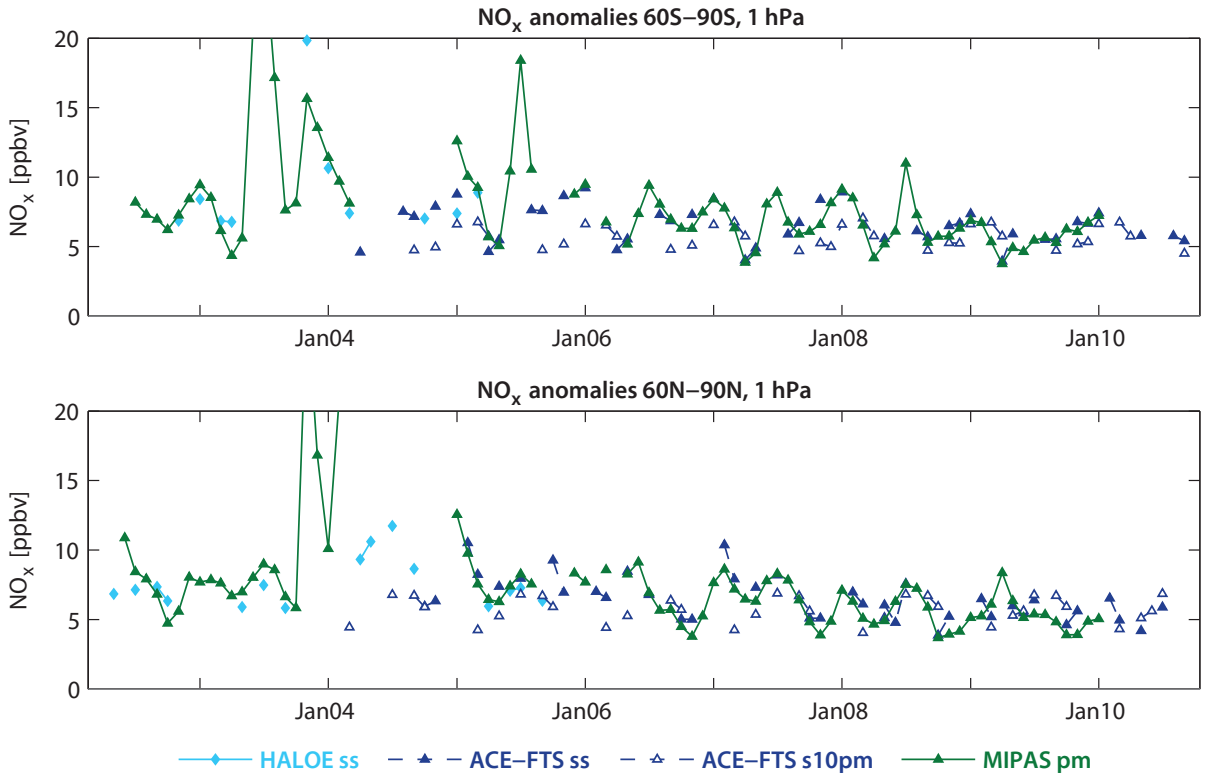


Figure 4.12.9: Time series of polar NO_x for 2003-2010. *Time series of polar 10pm and local sunset NO_x for 60°S-90°S (upper panel) and 60°N-90°N (lower panel) at 1 hPa from 2003 to 2010 are shown.*

In the mid-latitudes, the dataset anomalies show similar signals related to interannual variability, although the agreement is not as good compared to the tropics, consistent with results from the evaluations of the NO_2 interannual variability. Again, the scaled ACE-FTS data (not shown here) mostly fail to reproduce the interannual signal observed by the other instruments.

Most features observed for deseasonalised 10am and local sunrise NO_x also hold for the 10pm and local sunset NO_x time series (see **Figure A4.12.7** in *Appendix A4*). The tropical anomalies show a very similar QBO signal that agrees quite well between all instruments. Overall, the three 10pm datasets agree better with each other than with ACE-FTS and HALOE. In the SH mid-latitudes, SCIAMACHY shows smaller anomalies than the other datasets for some years (*e.g.*, 2007).

4.12.5 NO_x evaluations: Downward transport of NO_x during polar winter

In the polar mesosphere, NO_x is produced by EPP [*Barth*, 1992; *Solomon et al.*, 1982]. Observations have shown that NO_x in the polar mesosphere is transported downwards into the stratosphere inside the polar vortex [*Funke et al.*, 2005b; *Seppälä et al.*, 2007] causing elevated NO_x levels during polar winter. How well the limb-viewing satellite datasets agree on this phenomenon is evaluated in **Figure 4.12.9**, which shows NO_x time series in the USLM (1 hPa) for local sunrise and 10pm climatologies at high NH and SH latitudes.

MIPAS shows very high NO_x abundances for the SH winters 2003 and 2005, and for the NH winters 2003/2004 and 2004/2005. For most of these pronounced events HALOE and ACE-FTS do not provide the monthly zonal means, so that a direct comparison is not possible. The only exception is the NH winter 2004/2005, for which sunset ACE-FTS in February confirms high NO_x values exceeding 10 ppbv as observed by MIPAS. For the rest of the observation time period the polar NO_x time series shows a semi-annual oscillation with a maximum in summer due to a build up from reservoirs and (for most years) a maximum in winter due to EPP events. For the NH winters 2006/2007 and 2007/2008, MIPAS and sunset ACE-FTS datasets both confirm elevated NO_x values, while for the remaining winters a direct comparison is more complicated due to missing data coverage for ACE-FTS. Note that scaled ACE-FTS NO_x data show no clear signals of EPP events opposite to scaled ACE-FTS NO_2 data (*Section 4.11*).

4.12.6 Summary and conclusions: NO_x

A comprehensive comparison of NO_x profile climatologies from five satellite instruments (HALOE, OSIRIS, MIPAS, SCIAMACHY, and ACE-FTS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the NO_x mean state and important characteristics of the individual datasets are presented in the following summary, including three synopsis plots. The first summary plot (**Figure 4.12.10**) provides information on the NO_x mean state at 10am and 10pm. Additionally, the uncertainty

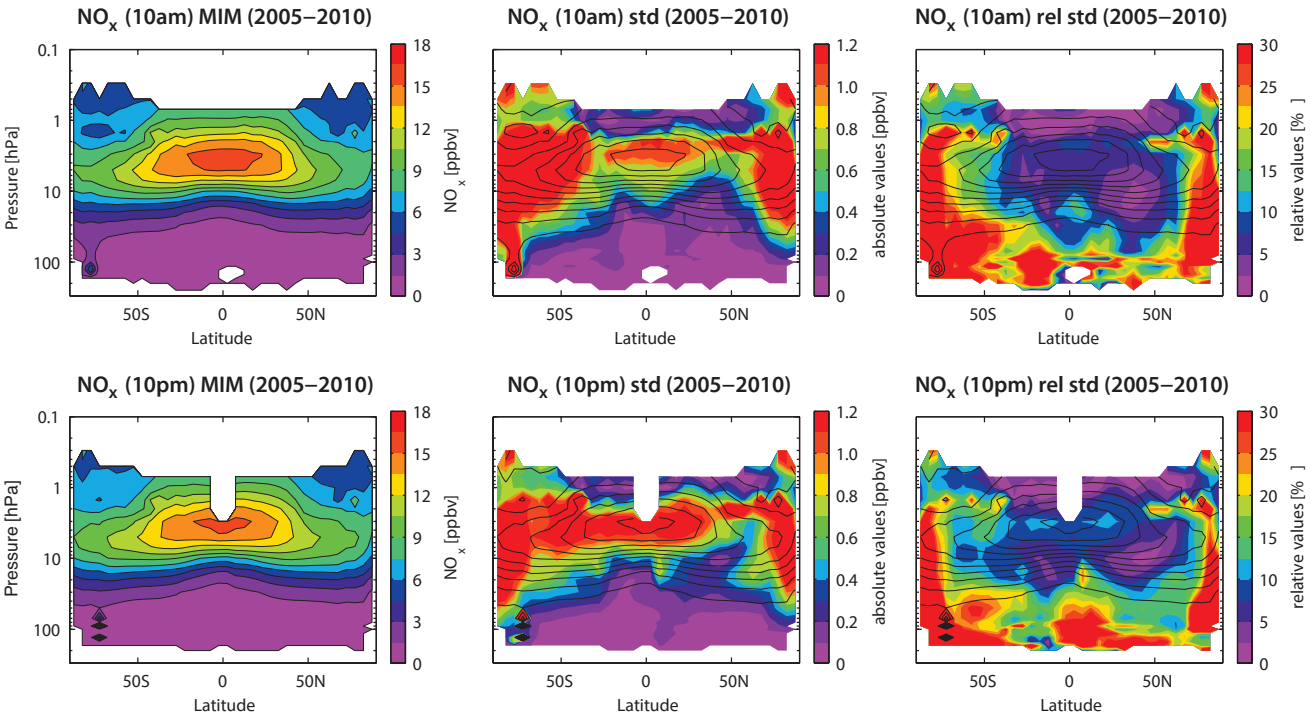


Figure 4.12.10: Summary of NO_x annual zonal mean state for 2005-2010. *Annual zonal mean cross sections of the NO_x MIM are shown in the left panels for10am (upper row) and 10pm (lower row) illumination conditions. The NO_x mean state at 10am (pm) is based on MIPAS at 10am (pm), ACE-FTS scaled to 10am (pm) and OSIRIS and SCIAMACHY derived from NO_2 with a chemical box model and scaled to 10am (pm). For both illumination conditions, the standard deviation over all respective instruments is presented in the middle panel, and the relative standard deviation in the right panel. Black contour lines give the MIM distribution. The MIM and standard deviation are only displayed for regions where at least two instruments provide measurements.*

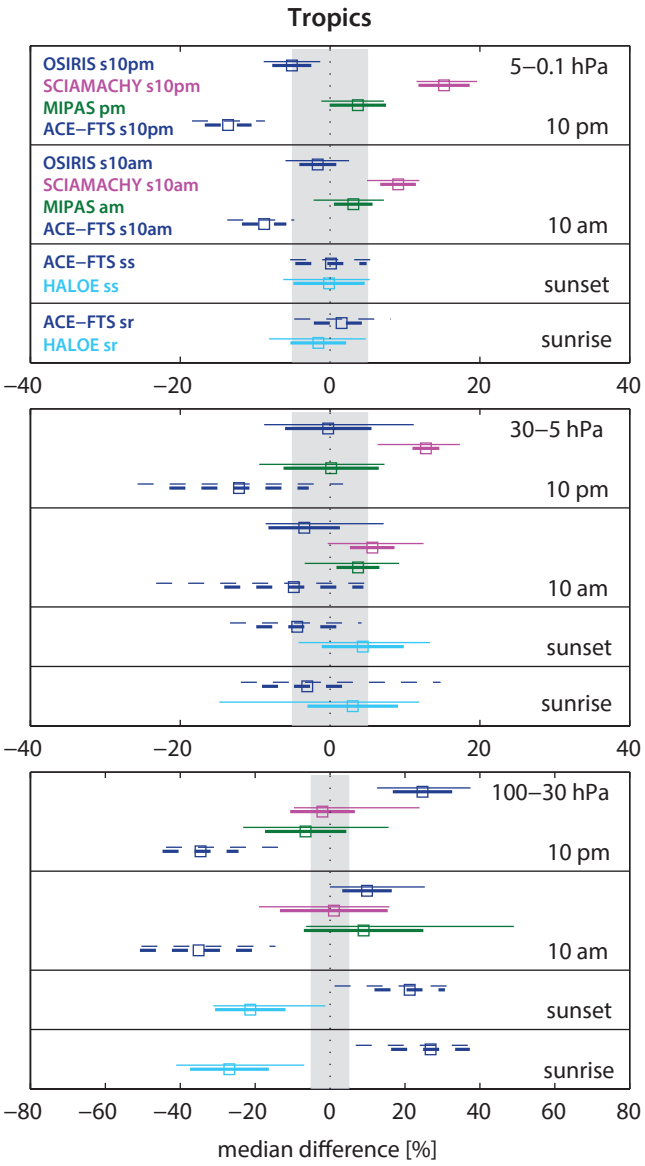


Figure 4.12.11: Summary NO_x differences in the tropics for 2004-2005. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (30°S-30°N) for three different altitude regions from the UT up to the LM between 100 and 0.1 hPa for the reference period 2004-2005.*

derived from the spread between the datasets is given for both illumination conditions. The second summary plot (**Figures 4.12.11** and **4.12.12**) shows specific inter-instrument differences in form of deviations of the instrument climatologies from the MIM climatology. For each region four separate evaluations for the four different illumination conditions (10am, 10pm, ss, sr) are included. For each LST, instrument and selected region, the deviation from 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 assessment of the atmospheric NO_x annual mean state is based on four climatologies corresponding to 10am and 10pm, respectively. Note that three out of four climatologies have been derived by scaling the individual measurements with a chemical box model to 10am/pm LST.

Middle and upper stratosphere (30-1 hPa)

The uncertainty in our knowledge of the atmospheric NO_x annual mean state is smallest in the tropical and NH mid-latitude MS/US (**Figure 4.12.10**, right panel), with a 1σ multi-instrument spread in this region of up to ±10%. In particular, in the NH mid-latitude MS, the instruments agree very well (1σ of ±5%). In the SH subtropics, the inter-instrument spread is comparable to the tropics (up to ±10%), but deviations increase in the SH mid-latitudes, in particular for the 10am climatologies (up to ±20%).

Lower stratosphere (100-30 hPa)

In the LS, the NO_x abundances decrease quickly with decreasing altitude and large deviations of up to ±30% are found in the SH mid-latitudes and for the 10pm climatologies also in the tropics. The instruments show a better agreement in the NH mid-latitudes and SH subtropics with a spread of up to ±20%.

High latitudes

At high latitudes, the instruments show larger deviations than at lower latitudes. In the MS, the best agreement is found for the 10pm climatologies in the NH where a spread larger than ±20% is found only north of 80°N. In the USLM, the high latitude annual mean NO_x abundance is dominated by the polar night NO_x descent, causing an increase of the inter-instrument spread for levels above 1 hPa.

Instrument-specific conclusions

Local sunrise/sunset climatologies from solar occultation instruments

HALOE and **ACE-FTS** show an excellent agreement in the US with mean differences below ±2.5% for their local sunset and sunrise climatologies (**Figures 4.12.11** and **4.12.12**). In the MS, HALOE detects slightly larger NO_x abundances than ACE-FTS resulting in differences with respect to their MIM of ±5%. The only exception to this is the comparison of the local sunset climatologies in the mid-latitude MS where both instruments show differences of up to ±10%. In the LS, the relative differences increase steadily, reaching mean values of up to ±30% with HALOE on the

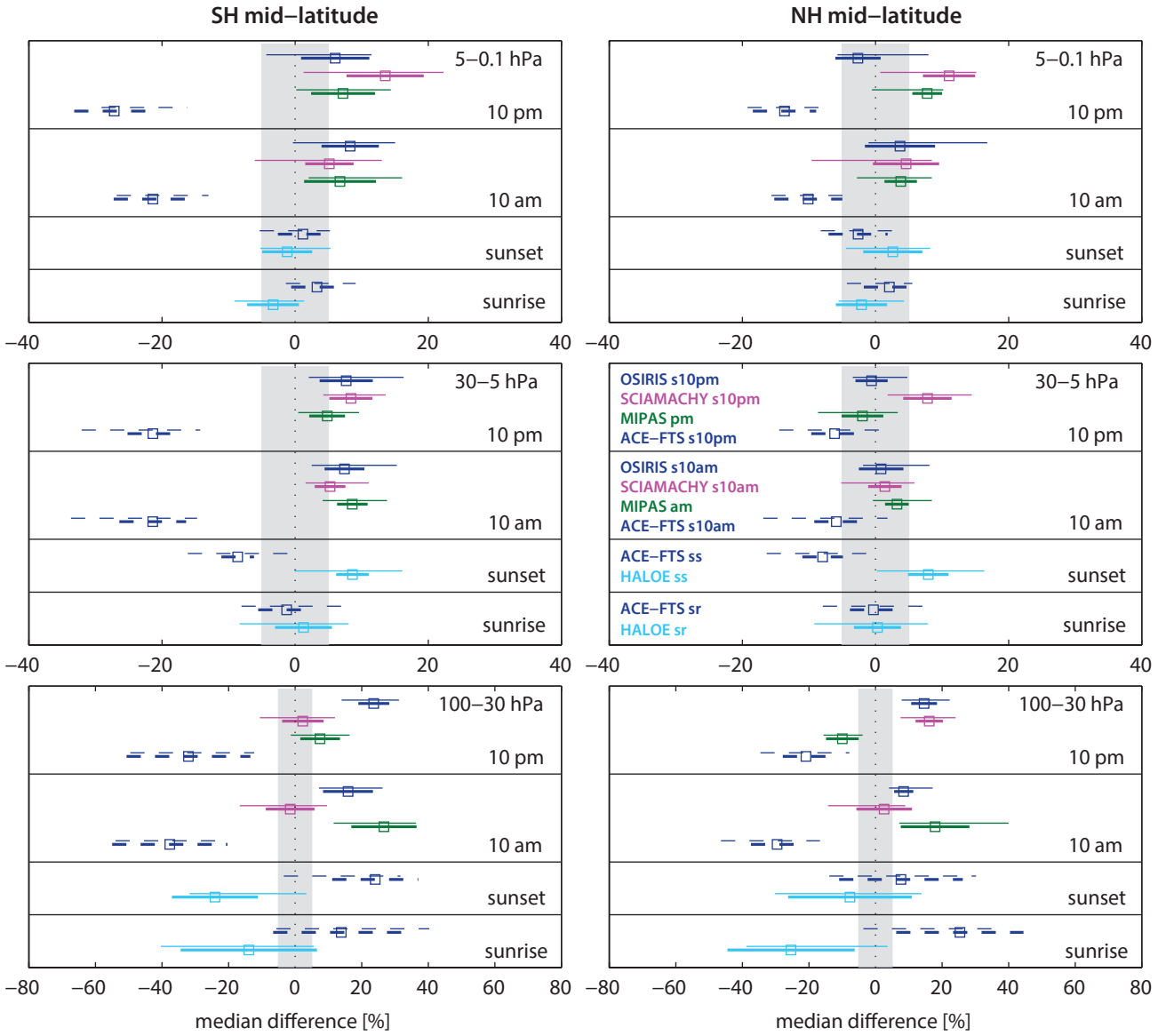


Figure 4.12.12: Summary NO_x differences in the mid-latitudes for 2004-2005. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the NH mid-latitudes (30°N-60°N) and for the SH mid-latitudes (30°S-60°S) for three different altitude regions from the UT up to the LM between 100 and 0.1 hPa for the reference period 2004-2005.*

low side and ACE-FTS on the high side. In particular in the mid-latitude LS, both datasets show a large regional spread (over all grid points in this region) indicating that the deviations are not well defined. Overall, the NO_x local sunrise and sunset evaluations give a consistent picture, with the exception of the mid-latitude MS. Despite their lower data coverage when compared to other limb sounders, both solar occultation instruments display important signals of interannual variability like the tropical QBO cycle. At the same time they show stronger month-to-month fluctuations probably related to sampling impacts.

10am/pm climatologies

The limb emission and scattering instruments MIPAS, OSIRIS, and SCIAMACHY, are evaluated based on their 10am/pm climatologies, with the latter two derived from scaling with a chemical box model. Additionally, 10am/pm climatologies from the scaled local sunrise/sunset measurements of the solar occultation instrument ACE-FTS are included in the evaluation. All climatologies show a good agreement in the tropical and NH mid-latitude MS with mean differences of ±5% to ±10%. In particular, the 10am climatologies from MIPAS, OSIRIS, and SCIAMACHY agree very well in the mid-latitude MS and US with differences of less than ±5%. Monthly zonal mean cross sections (**Figure 4.12.6**) reveal that for most climatologies the deviations from the MIM can change sign depending on the latitude band and month. All 10am/pm climatologies show the tropical QBO signal with the best agreement found between MIPAS, OSIRIS, and SCIAMACHY.

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 other datasets, is mostly in the middle range with relatively small deviations with respect to the MIM of up to ±10%. Only the 10am climatologies in the LS show negative deviations of up -20% in agreement with a similar finding for the MIPAS NO_2 climatologies.

The 10am **SCIAMACHY** climatology agrees very well with the MIM with differences of up to ±5% in most cases. The 10pm climatology, however, shows larger values than the other datasets with deviations of up to +10% in the MS and +15% in the US. This could be caused by the larger impact of the scaling procedure on the 10pm climatology for which SCIAMACHY measurements need to be scaled to completely different illumination conditions. Note, however, that for NO_2, which shows a much more pronounced diurnal cycle, no such severe differences in the performance of the 10am/pm climatologies exist.

OSIRIS is in the middle of the range in the MS and US, but shows larger deviations in the LS where it sets the upper measurement range and displays differences of up to +30% for the 10pm climatologies. The fact that in the tropical LS, OSIRIS 10am climatology agrees within ±5% with MIPAS and SCIAMACHY, while the OSIRIS 10pm climatology is 20-30% larger than the other two datasets, is consistent with the NO_2 evaluations.

Scaled ACE-FTS in the tropical and NH mid-latitude MS agrees well with the other datasets, with deviations of ±5% to ±10% in the MS and deviations of up to ±15% in the US. However, in the SH mid-latitudes, scaled ACE-FTS data are considerably lower than the other datasets, with differences of up to -30%. This inconsistency between NH and SH mid-latitudes causes the larger inter-instrument spread in the latter region, apparent also in **Figure 4.12.10**. In the LS, scaled ACE-FTS sets the lower end of the measurement range with differences of up to -40%.

4.13 Nitric acid – HNO_3

Nitric acid (HNO_3) is a member of the total reactive nitrogen family NO_y and has a large impact on stratospheric ozone destruction in the polar regions through its role in PSC formation. HNO_3 contributes to the composition of nitric acid trihydrate (NAT), forming Type Ia PSCs [*Toon et al.*, 1986; *Hanson and Mauersberger*, 1988], and to supercooled ternary solution particles ($H_2SO_4/H_2O/HNO_3$), forming Type Ib PSCs [*Carslaw et al.*, 1994; *Lowe and MacKenzie*, 2008]. Heterogeneous reactions occurring on PSC surfaces convert halogens from relatively inert reservoir species into their reactive forms, driving halogen-catalyzed ozone loss in polar spring [*e.g.*, *Peter*, 1997; *Solomon*, 1999]. HNO_3 is irreversibly removed if the solid PSC particles grow to large sizes and sediment out of the stratosphere [*Fahey et al.*, 2001], a process that is referred to as denitrification. If PSC particles do not sediment out of the stratosphere but eventually evaporate, the photolysis of HNO_3 will cause increasing NO_2 concentrations, thus enhancing the halogen deactivation process in polar springtime.

Stratospheric HNO_3 displays a weak diurnal cycle in the US that increases in the LM. **Figure 4.13.1** shows the diurnal HNO_3 cycle as a function of LST for three different pressure levels as derived from a chemical box model [*McLinden et al.*, 2010].

4.13.1 Availability of HNO_3 measurements

The first stratospheric HNO_3 measurements were made by LIMS, covering the end of 1978 and the first half of 1979. The next vertically resolved HNO_3 satellite measurements included in the SPARC Data Initiative came from UARS-MLS, covering the years 1991 to 1999. After 2001, several HNO_3 datasets became available from limb emission and solar occultation instruments. The time period 2005-2010 is covered by SMR, MIPAS, ACE-FTS, and Aura-MLS

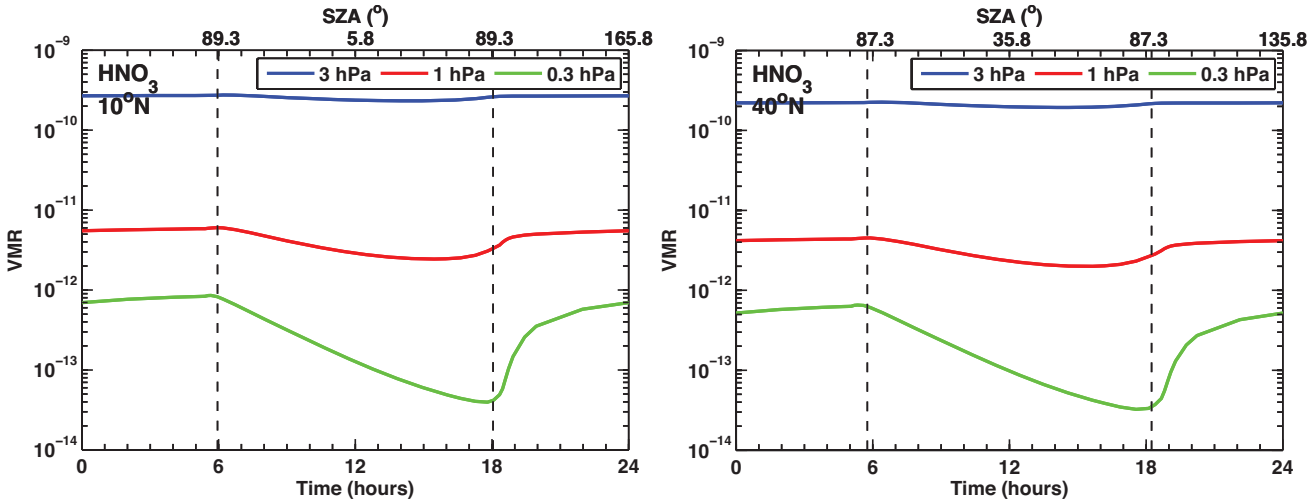


Figure 4.13.1: Diurnal HNO_3 cycle. *HNO_3 variations as function of LST are shown at 10°N and 40°N at 0.3, 1 and 3 hPa for March 15.*

allowing for an inter-instrument comparison maximising the number of instruments and number of years. HIRDLS measurements (2005-2007) are also included in the basic 2005-2010 comparison since sensitivity tests show that the results are not affected by HIRDLS covering a shorter time period. **Tables 4.13.1** and **4.13.2** compile information on the availability of HNO_3 measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report.

4.13.2 HNO_3 evaluations: Zonal mean cross sections and vertical profiles

SMR, MIPAS, ACE-FTS, Aura-MLS, and HIRDLS (2005-2010)

Zonal mean cross sections and vertical profiles are first compared for the overlap period 2005-2010. **Figure 4.13.2** shows all annual mean HNO_3 climatologies, which have maxima in the MS (around 20-30 hPa) at the mid- and high latitudes of both hemispheres. While the datasets agree very well on the overall distribution at low latitudes, they show large differences at polar latitudes, in particular in the SH. Here, ACE-FTS observes much higher and SMR much lower abundances than MIPAS in the annual zonal mean distribution. Examining monthly zonal mean cross sections (*e.g.*, **Figure A4.13.1** in *Appendix A4*) reveals that for ACE-FTS this feature is related to the sampling coverage of this region, which varies over the year resulting in an annual mean cross section that is not representative. In contrast, the zonal mean evaluations for SMR yield similar results to the annual mean evaluations.

Figure 4.13.3 shows the relative differences of the five instruments with respect to their MIM. For most instruments, the differences have a pronounced vertical gradient (negative and positive) and change sign between 30 and 10 hPa. At higher altitudes, *i.e.*, above the HNO_3 maximum, Aura-MLS and HIRDLS abundances are lower than the MIM, with differences that can reach up to -50%. ACE-FTS is larger, with differences of up to +50%, while MIPAS and SMR are also larger than the MIM but exhibit less extreme differences. In the LS, the situation is reversed

Table 4.13.1: Available HNO_3 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.*

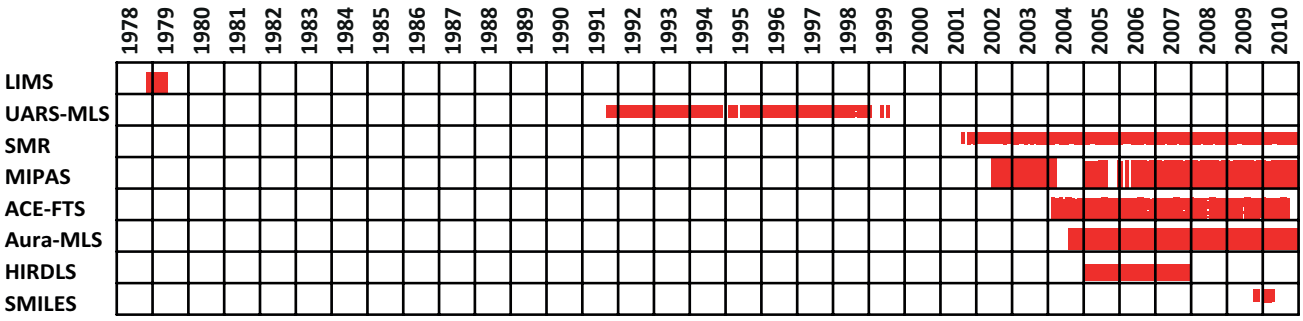


Table 4.13.2: Data version, time period, vertical range, vertical resolution, references and other comments for HNO_3 datasets participating in the SPARC Data Initiative.

Instrument and data version	Time period	Vertical range	Vertical resolution	References	Additional comments
LIMS V6.0	Nov 78 – May 79				
UARS-MLS V6	Oct 91 – Oct 99	100 – 4.6 hPa	5 – 10 km	*Livesey et al.*, 2003	Significant low bias (1-3 ppbv) exists for p<15 hPa. Some evidence for high bias below VMR peak.
SMR V2-0	Jul 01 –	18 – 45 km	1.5 – 2 km	*Urban et al.*, 2006 *Urban et al.*, 2009	Empirical scaling correction applied (see *Urban et al.*, 2009)
MIPAS **MIPAS(1)** V9 **MIPAS(2)** V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	6 km (cloud top altitude) – 70 km	 4 – 6 km 3 – 5 km	*Mengistu Tsidu et al.*, 2005 *Wang et al.*, 2007 *von Clarmann et al.*, 2009a	
ACE-FTS V2.2	Mar 04 –	5 – 37 km	3 – 4 km	*Wolff et al.*, 2008	
Aura-MLS V3-3	Aug 04 –	147 – 1 hPa	3 – 5 km	*Santee et al.*, 2007 *Livesey et al.*, 2013 *Fiorucci et al.*, 2013	
HIRDLS V6.0	Jan 05 – Mar 08	215 – 7.5 hPa	1 km	*Gille and Gray*, 2011	Latitude range 63°S-80°N
SMILES V2-0-1	Oct 09 – April 10	18 – 45 km	3 – 4 km	*Kreyling et al.*, 2013	Bias due to problems in spectroscopic parameter and altitude shift (corrected for in v3-0-0).

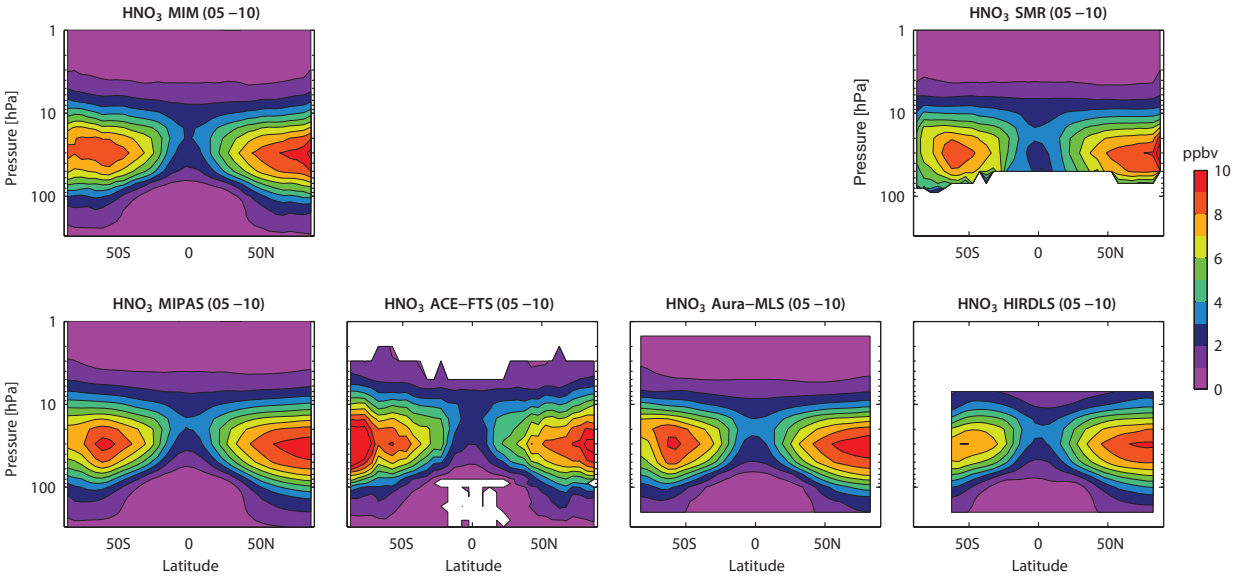


Figure 4.13.2: Cross sections of annual zonal mean HNO_3. *Annual zonal mean HNO_3 cross sections for 2005-2010 are shown for the MIM in the upper left, SMR in the upper right and MIPAS, ACE-FTS, Aura-MLS, and HIRDLS in the lower panels.*

with Aura-MLS values being high and ACE-FTS values low. Such a tendency between the Aura-MLS and ACE-FTS profiles was already noted in the Aura-MLS data quality documentation by *Livesey et al.* [2013], although the average coincident profile differences discussed there are somewhat smaller than the differences mentioned here. The smallest deviations are found for SMR, MIPAS, and HIRDLS in the mid-latitudes. At high latitudes, monthly mean cross sections need to be analysed instead of annual means, since the latter can be impacted strongly by the sampling patterns of the instruments. Overall, differences are largest in the SH polar winter, and spring at latitudes higher than 60°S (see **Figure A4.13.2** in *Appendix A4*), with MIPAS reporting more and Aura-MLS and SMR less HNO_3.

Detailed evaluations of monthly zonal mean differences for individual latitude bands are shown in **Figure 4.13.4**. At high SH latitudes (75°S-80°S) in July, all instruments clearly show the removal of HNO_3 from the gas phase. However, the signal is much stronger in Aura-MLS and SMR than in MIPAS, leading to differences of up to ±50% relative to the MIM at around 50 hPa. Very likely, this can be attributed to the higher sensitivity of infrared emission sounders like MIPAS to PSCs, leading to a more rigorous rejection of PSC-contaminated measurements and thus to higher HNO_3 mean values during the Antarctic winter conditions. In the 65°S-70°S latitude band, the HNO_3 contribution to PSC formation is smaller but MIPAS still clearly shows higher HNO_3 abundances than the other instruments (differences of up to +40%). Above 30 hPa, ACE-FTS shows better agreement with MIPAS. The Aura-MLS profile at high latitudes is characterised by small oscillations not found for any of the other instruments. In the tropics (0°N-5°N for January), the strongest disagreement is found for HIRDLS, which is larger than all other datasets in the LS and MS. The positive deviations of up to +80% (of small values) at

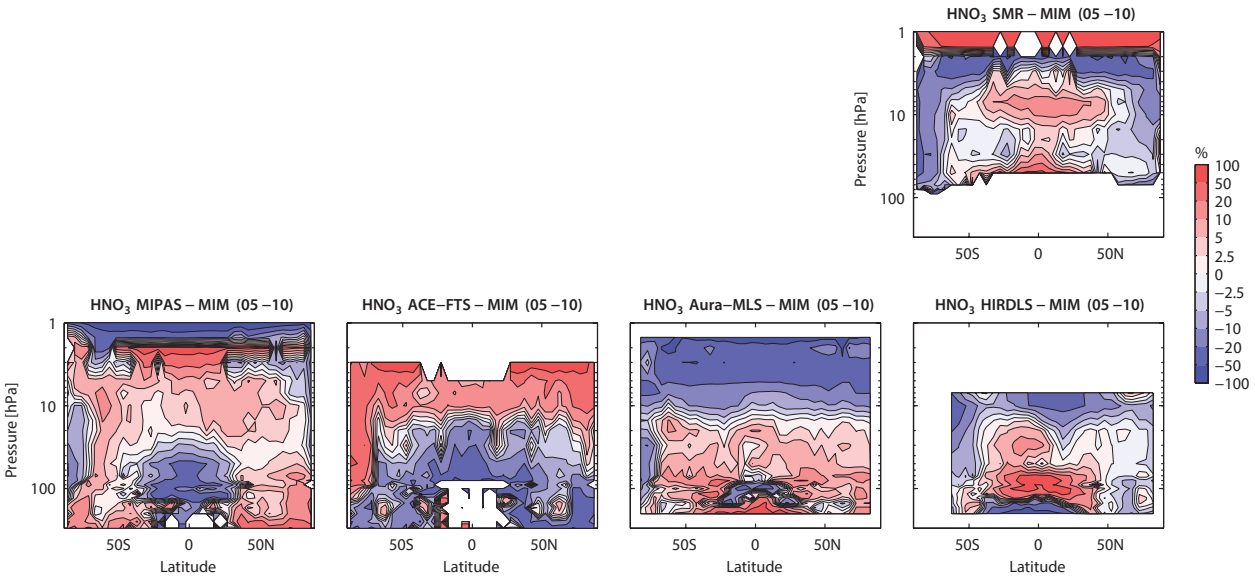


Figure 4.13.3: Cross sections of annual zonal mean HNO_3 differences. *Annual zonal mean HNO_3 differences for 2005-2010 between the individual instruments (SMR, MIPAS, ACE-FTS, Aura-MLS, and HIRDLS) and the MIM are shown.*

around 70 hPa are probably due to the lack of correction for aerosol emission in V6. The comparison of the NH mid-latitude profiles (50°N-55°N for May) confirms very good agreement in a wide altitude range from 200 to 15 hPa. Above this level, Aura-MLS falls off faster than HIRDLS and ACE-FTS, leading to differences of up to ±30% at 5 hPa. In the tropics and mid-latitudes, SMR has a positive vertical gradient between 3 and 1 hPa, where all other datasets approach zero, resulting in large deviations at these upper levels.

LIMS, UARS-MLS, and SMILES

For LIMS and UARS-MLS, no direct comparison for HNO_3 is possible, thus we include a comparison to the 2005-2010 MIM (restricted to January-April for the LIMS evaluation). The resulting differences can be caused by inter-instrument differences or by long-term changes in HNO_3, therefore clear attribution of the differences is not possible. The trend in N_2O is expected to lead to an upward trend in all nitrogen species, modulated by shifts in the total reactive nitrogen family partitioning due to changes in ozone, temperature, halogens, and possibly aerosol loading [*Fish et al.*, 2000; *McLinden et al.*, 2001].

Cross sections of LIMS (January-April 1979) show, as expected, considerably smaller HNO_3 abundances than the 2005-2010 MIM (**Figure 4.13.5**). The difference between the two (**Figure 4.13.6**) is mostly negative (up to -100% in the tropical LS), except for the mid-latitude LS and the US. At NH high latitudes, deviations are relatively small (up to ±5%), suggesting similar HNO_3 abundance in 1979 and 2005-2010 during Arctic winter and spring.

UARS-MLS detects less HNO_3 for the time period 1991-1998 compared to the annual mean 2005-2010 cross sections (**Figure 4.13.5**). Particularly evident is the reduced HNO_3 amount at high SH latitudes during the 1990s, consistent with strong PSC formation during this time period. The deviations between UARS-MLS and the 2005-2010 MIM are negative over a large range (**Figure 4.13.6**). A negative bias in UARS-MLS HNO_3 was also noted in comparsion with ATMOS and Improved Limb Atmospheric Spetrometer (ILAS) data for pressures less than 15 hPa [*Livesey et al.*, 2003]. In the mid-latitude LS and the NH

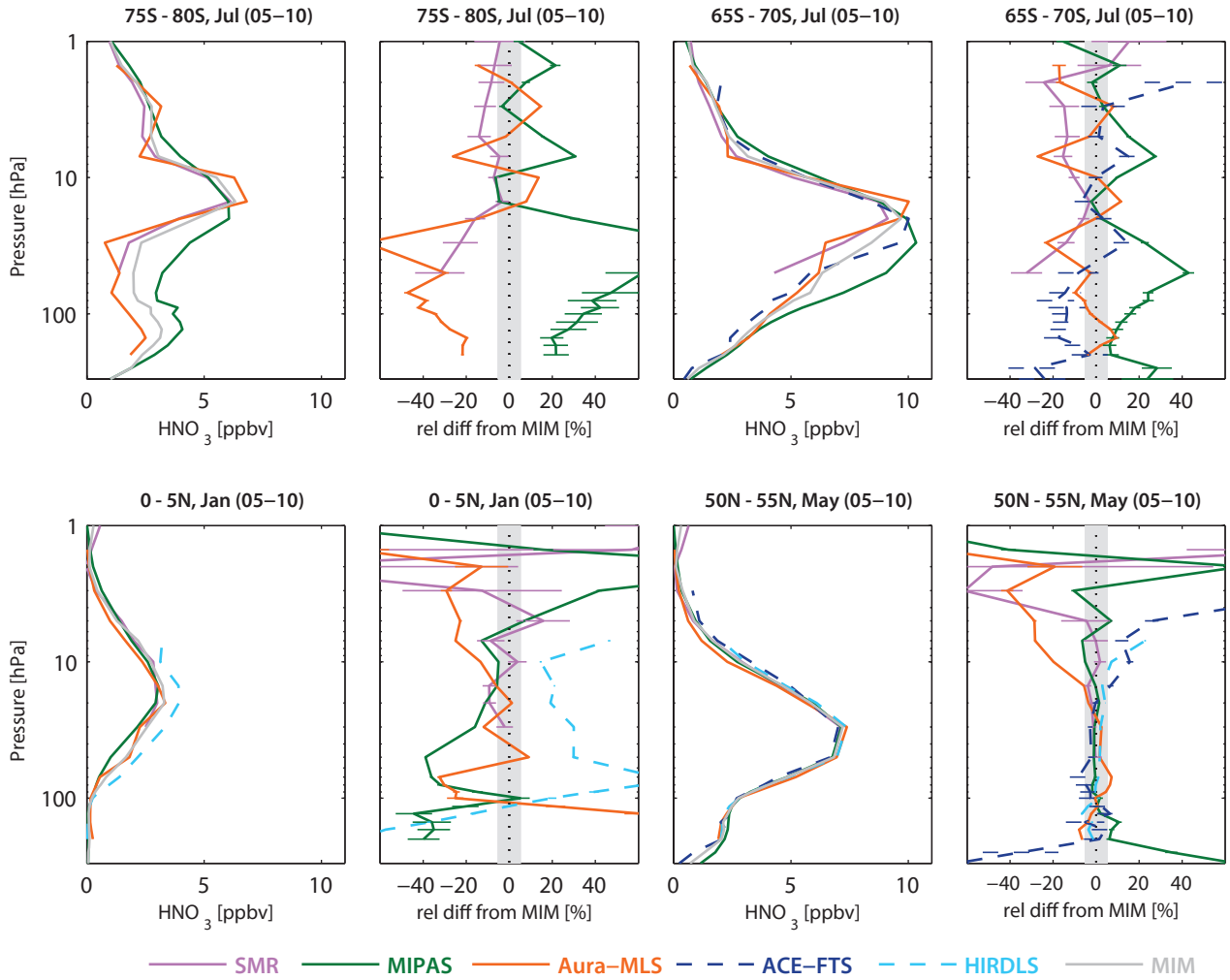


Figure 4.13.4: Vertical profiles of monthly zonal mean HNO_3 for 2005-2010. *Zonal mean HNO_3 profiles for 75°S-80°S and 65°S-70°S in July, 0°N-5°N in January, and 50°N-55°N in May are shown in columns 1 and 3. Differences between the individual instruments and the MIM profiles are shown in columns 2 and 4. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are within ±5%.*

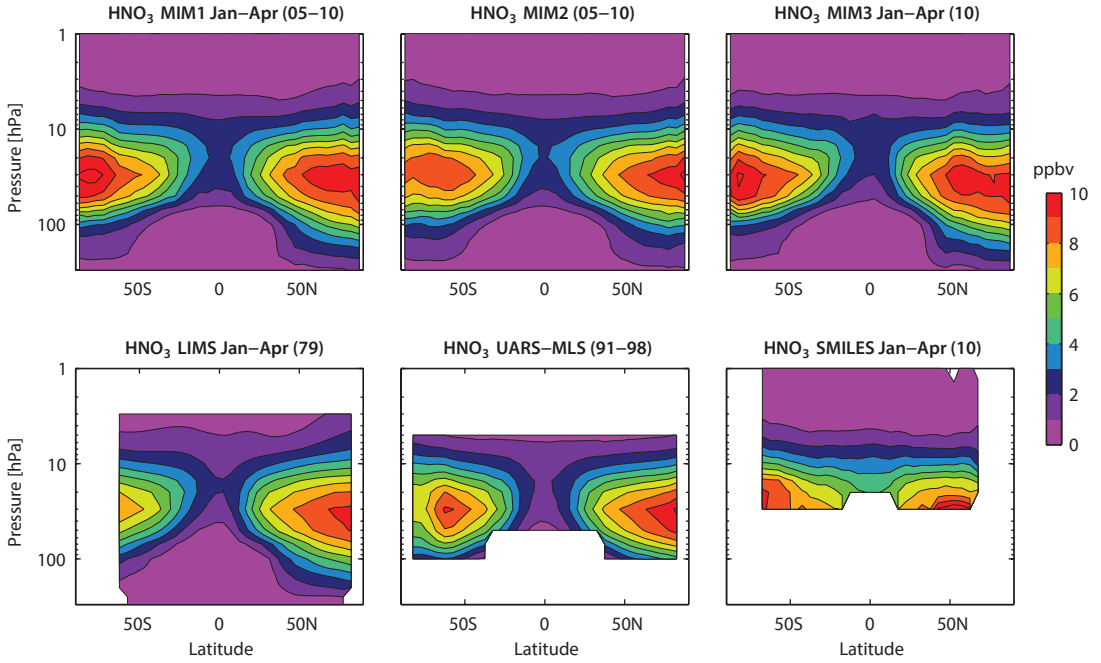


Figure 4.13.5: Cross sections of zonal mean HNO_3. *Zonal mean HNO_3 cross sections of the MIM1 (January-April 2005-2010), MIM2 (2005-2010) and MIM3 (January-April 2010), all based on SMR, MIPAS, ACE-FTS, Aura-MLS, and HIRDLS, are shown in the upper panels. Cross sections of LIMS (January-April 1979), UARS-MLS (1991-1998), and SMILES (January-April 2010) are displayed in the lower panels.*

high latitude LS, UARS-MLS is larger than the 2005-2010 MIM. Such positive deviations of an older dataset *versus* the newer datasets in the mid-latitude LS are also evident in the comparison of LIMS *versus* the 2005-2010 MIM but are not consistent with the expected positive NO_y trend. They could be related to inter-instrument differences caused by factors such as different altitude resolutions, to shifts in the nitrogen partitioning or possibly to changes in transport and mixing properties. More detailed comparisons *versus* models would be necessary to help investigate such issues.

The most recent satellite HNO_3 dataset is from SMILES and a comparison to the MIM of SMR, MIPAS, ACE-FTS, and Aura-MLS for the time period January-April 2010 is given in **Figures 4.13.5** and **4.13.6**. SMILES detects much higher abundances of up to ±50% in the MS, likely related to line mixing spectroscopic parameters and an altitude shift.

4.13.3 HNO_3 evaluations: Seasonal cycles

Figure 4.13.7 shows the HNO_3 seasonal cycles at high and tropical latitudes at 50 hPa. HNO_3 exhibits a strong seasonal cycle in the high latitude LS due to chemistry and transport effects. In winter, descending air masses transport HNO_3 downwards into the LS. This dynamical effect is counteracted by PSC formation, which removes HNO_3 from the gas phase and leads to a HNO_3 minimum at the end of winter/beginning of spring in the SH. During SH spring and summer, HNO_3 recovery through renitrification occurs, slowed down by the conversion between HNO_3 and the other nitrogen species. Comparison with the NO_y seasonal cycle (**Figure 4.17.5** in *Section 4.17*) further indicates that nearly the entire amplitude of the Antarctic HNO_3 seasonal cycle (around 8 ppbv between maximum and minimum) is related to the HNO_3 heterogeneous chemistry. Amplitudes derived from MIPAS and Aura-MLS seasonal cycles

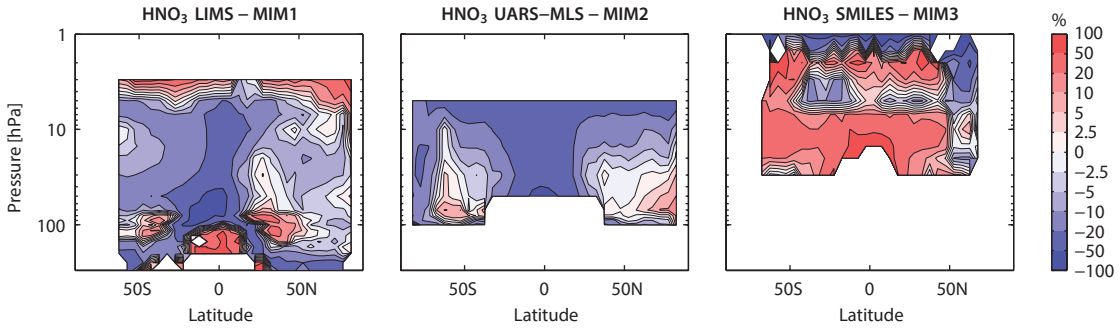


Figure 4.13.6: Cross sections of zonal mean HNO_3 differences. *Annual zonal mean HNO_3 differences between the individual instruments LIMS, UARS-MLS, SMILES and MIM1, MIM2, and MIM3, respectively, are shown. Time periods for the individual instruments and the MIMs are given in Figure 4.13.5.*

agree well and are relatively large. ACE-FTS and SMR have smaller amplitudes, but their total values show a considerable offset towards each other for most months, with smaller values from SMR. At these high southern latitudes, MIPAS shows minima and maxima of the HNO_3 seasonal cycle with a delay of about 1 and 2 months, respectively, compared to Aura-MLS and SMR.

In the NH high latitudes, the HNO_3 seasonal cycle is also impacted by the conversion between HNO_3 and NO_x as well as by heterogeneous chemistry, as evidenced by the much smaller chemistry-caused seasonal cycle in NO_y (**Figure 4.17.5** in *Section 4.17*). All HNO_3 datasets show a very similar phase but HIRDLS and SMR have smaller amplitudes than the other instruments. While the seasonal cycles in HNO_3 and NO_y are consistent for ACE-FTS and MIPAS, the same is not true for SMR, indicating that inconsistencies are introduced by the compilation of the Odin NO_y climatology, which is partially based on photochemically modelled member species (see *Section 4.17*).

In the tropics, transport variations are expected to cause a weak annual cycle as seen by MIPAS, Aura-MLS, and HIRDLS. Again, HIRDLS is affected by the uncorrected effects of tropical aerosols (see also *Section 4.5*). MIPAS and Aura-MLS agree very well on the amplitude and phase of the seasonal cycle but show a large offset. There is no agreement of the ACE-FTS and SMR seasonal signals with the other instruments. The ACE-FTS sampling is sparser in the tropics than at higher latitudes, which is probably largely responsible for apparent differences in this region.

4.13.4 HNO_3 evaluations: Interannual variability

Interannual variability of HNO_3 at polar latitudes (60°N-90°N and 60°S-90°S) at 50 hPa and in the tropics (10°S-10°N) at 30 hPa is shown in **Figure 4.13.8**. The 2005-2010 time series of deseasonalised anomalies in the tropics at 30 hPa shows an excellent agreement between MIPAS, Aura-MLS, and ACE-FTS. The pronounced QBO signal is recorded with the same amplitude and phase by the three satellite datasets. SMR and HIRDLS also detect the QBO signal but show some deviations compared to the other datasets. SMR has a lower QBO amplitude, particularly at the beginning of 2007 and the end of 2009/beginning of 2010. HIRDLS has unrealistic January anomalies due to a too high January value in 2005 at the beginning of the measurement period. Note that the 2005 January HIRDLS value is based on two days of measurements only and will be omitted in future HIRDLS data versions. Furthermore, HIRDLS shows positive anomalies at the end of 2007 while the other datasets have negative anomalies during this time period. Note that a comparison of the anomalies for the 2005-2007 time period leads to very similar results.

At polar latitudes, the datasets show good agreement, but some outliers in the anomaly time series can be found as well. For most years the largest anomalies occur during polar winter, related to the strong interannual variability of chemical and dynamical processes impacting the polar stratospheric nitrogen budget. In the Arctic, the largest differences coincide with the largest anomalies in January. The smallest differences are found during NH summer when the smallest anomalies occur. In general, the maximum anomalies are followed by a slow decay of the anomalies over the following months until next autumn. This seasonal persistence is also found in the Antarctic where the interannual anomalies are more pronounced. MIPAS shows largest deviations in June and SMR displays stronger month-to-month fluctuations.

4.13.5 Summary and conclusions: HNO_3

A comprehensive comparison of eight HNO_3 profile climatologies has been carried out. Overall findings on the systematic uncertainty in our knowledge of the HNO_3 mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots. The first summary plot (**Figure 4.13.9**) provides information on the HNO_3 mean state and the

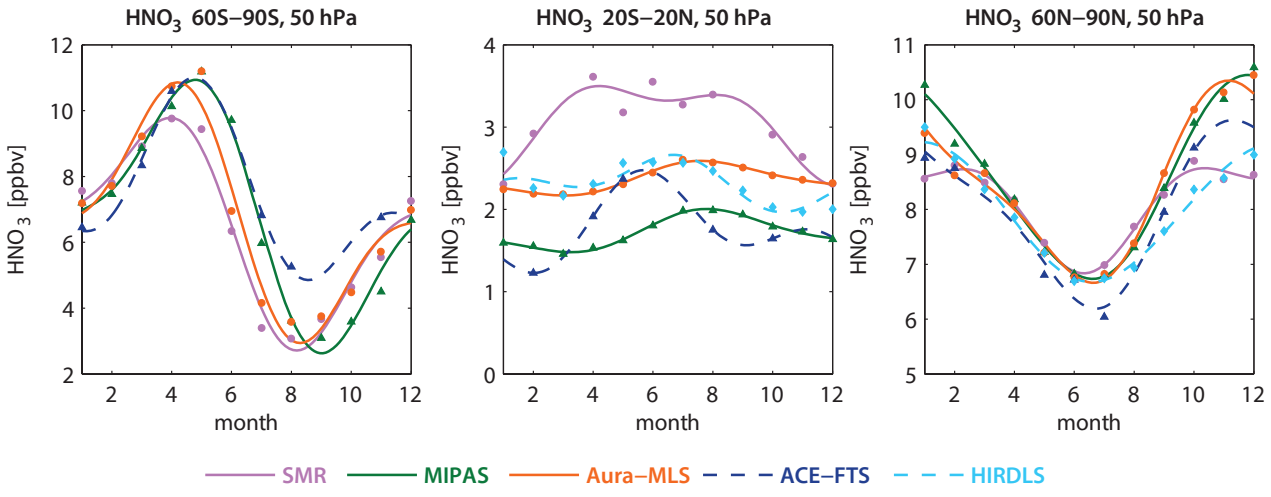


Figure 4.13.7: Seasonal cycle of HNO_3 for 2005-2010. *Seasonal cycle of monthly zonal mean HNO_3 for 60°S-90°S (left column), 20°S-20°N (middle column) and 60°N-90°N (right column) at 50 hPa.*

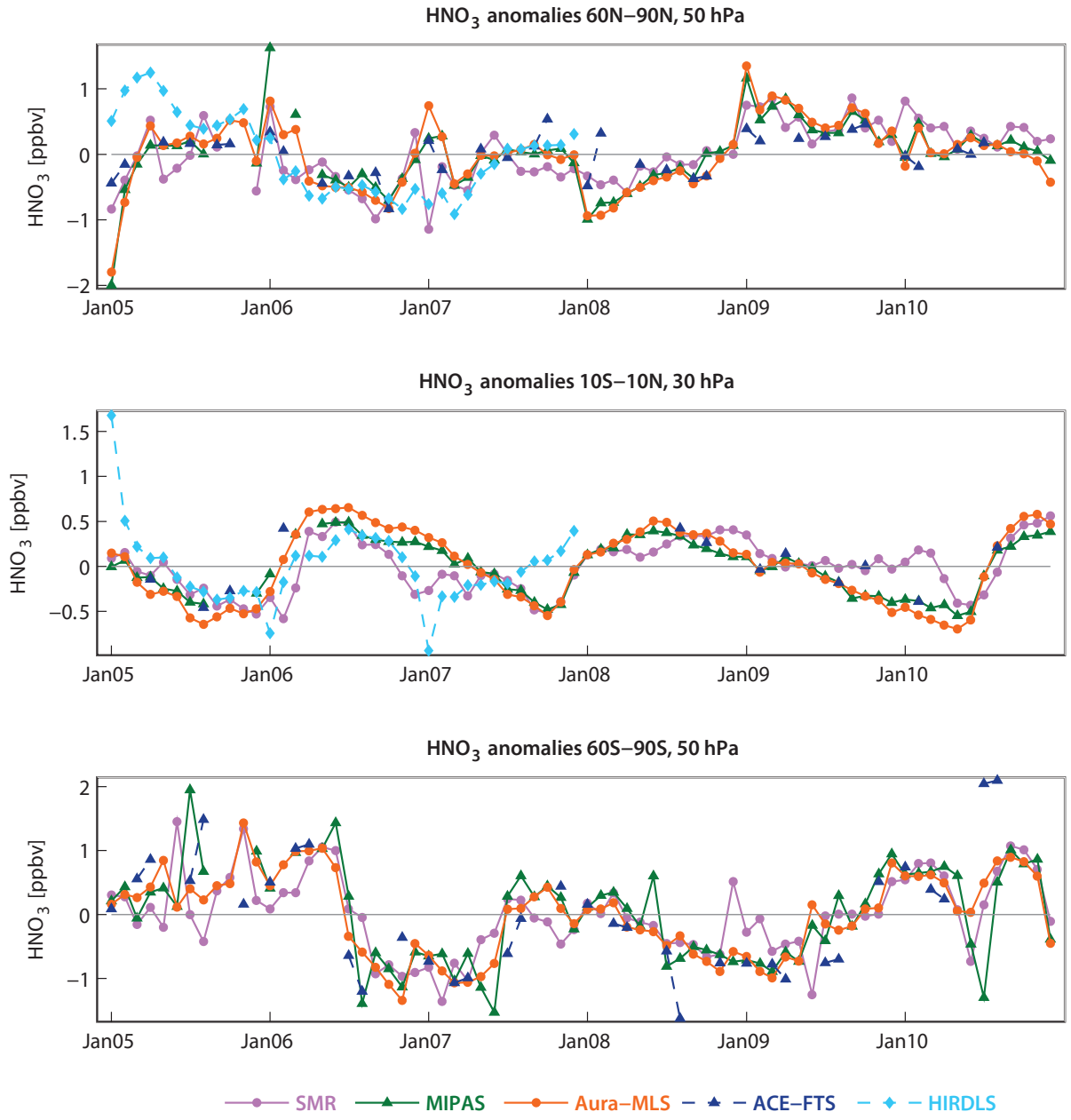


Figure 4.13.8: Time series of HNO_3 anomalies for 2005-2010. *Monthly zonal mean deseasonalised HNO_3 anomalies at 50 hPa for 60°N-90°N and 60°S-90°S, and at 30 hPa for 10°S-10°N.*

uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.13.10**) shows specific inter-instrument differences in form of deviations of the instrument climatologies relative 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. A detailed description of the rationale behind the summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The assessment of the atmospheric HNO_3 annual mean state is based on five climatologies overlapping for the time period 2005-2010.

Middle stratosphere (30-5 hPa)

The uncertainty in our knowledge of the atmospheric HNO_3 annual mean state is smallest in the MS (**Figure 4.13.9**, right panel) with a 1σ multi-instrument spread in this region between ±5% and ±15%. However, in the SH at the highest latitude bands (south of 80°S), the spread can reach values of ±30% partially related to larger deviations

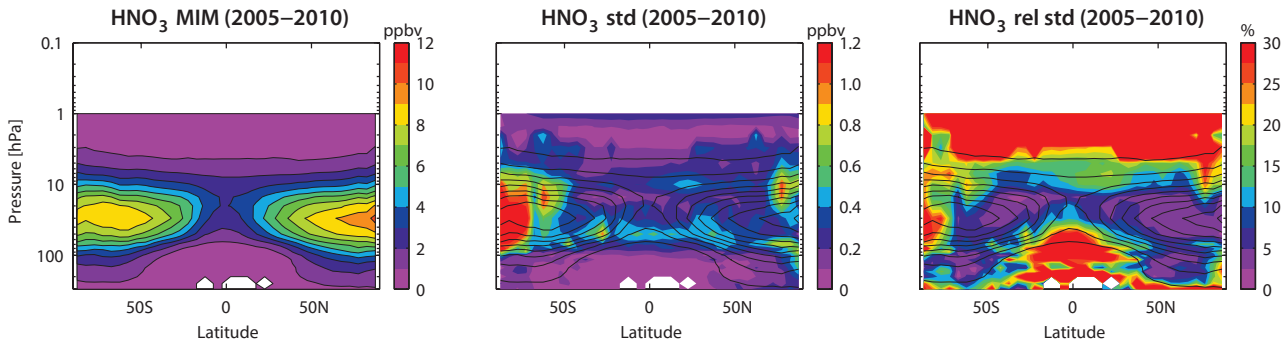


Figure 4.13.9: Summary of HNO_3 annual zonal mean state for 2005-2010. *Shown are the annual zonal mean cross section for the MIM of HNO_3 (left panel), the standard deviation over all instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black contour lines in the two rightmost panels give the MIM distribution. Instruments included are SMR, MIPAS, ACE-FTS, Aura-MLS, and HIRDLS. The MIM and standard deviation are only displayed for regions where at least two instruments provide measurements.*

during Antarctic winter. In addition, the larger deviations at the SH high latitudes are related to the fact that for some instruments (ACE-FTS) annual mean fields are not representative and evaluations in this region should be based on monthly mean fields.

Upper troposphere/lower stratosphere (300-30 hPa)

In the tropical UT and LS, the HNO_3 abundances decrease with decreasing altitude to very low values and the datasets show a large relative spread of up to ±50%. A much better agreement (±5% to ±10%) is found in the mid-latitude LS where the HNO_3 mixing ratios are larger compared to the tropics. In the mid-latitudes between 300-100 hPa, HNO_3 decreases and the values are comparable to the tropical LS; however, the deviations are considerably smaller giving an inter-instrument spread of around ±10%.

Upper stratosphere (5-1 hPa)

Above 3 hPa, the inter-instrument spread reaches very large values of ±50% to ±100% probably related to low HNO_3 mixing ratios close to the detection limit. In addition, diurnal variations become important above this level which can further impact the inter-instrument differences.

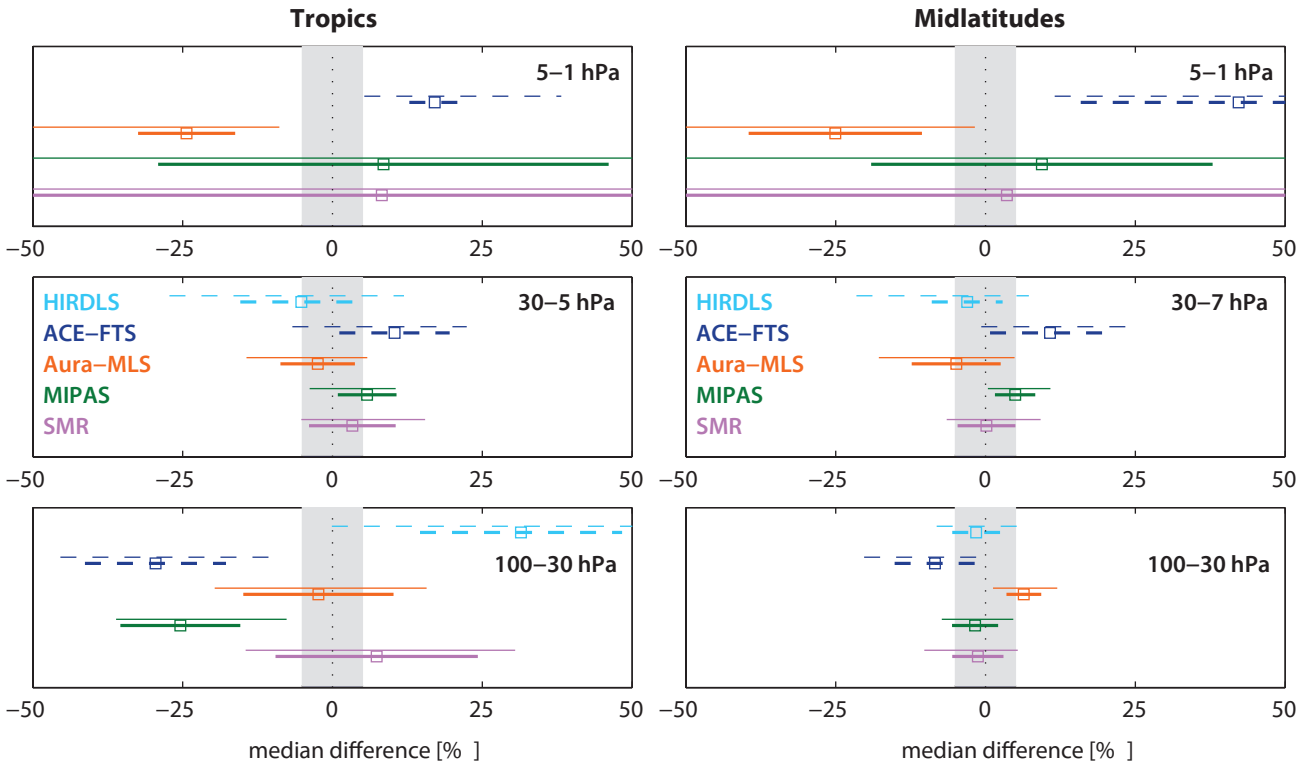


Figure 4.13.10: Summary of HNO_3 differences for 2005-2010. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (20°S-20°N) and midlatitudes (30°S-60°S and 30°N-60°N) and for three different altitude regions from the LS up to the US between 100 and 1 hPa for the reference period 2005-2010. The grey shaded area indicates where relative differences are within ±5%.*

Instrument-specific conclusions

LIMS and **UARS-MLS** provide the oldest HNO_3 satellite measurements available to the SPARC Data Initiative, and show negative deviations with respect to the later datasets covering 2005-2010. This difference is very likely attributable to the N_2O induced trend in HNO_3. At mid-latitudes and NH high latitudes in the LS, the deviations relative to the 2005-2010 data are positive, which is possibly related to changes in transport and mixing.

SMR shows excellent agreement with the MIM in the mid-latitude LS and MS, with differences smaller than ±2.5%. In the tropical MS, SMR is still in the middle range with differences smaller than +5%. In the tropical LS and in the US, SMR shows mostly positive deviations (smaller than +10%) and has a very large regional spread indicating a wide distribution of the relative differences around their mean.

The **MIPAS** climatology shows very similar behaviour to the SMR dataset, except for the tropical LS where MIPAS has a negative offset of -25%. MIPAS agrees well with the other datasets in terms of seasonal cycle and interannual variability, except for the Antarctic winter when it shows different anomalies. During Antarctic winter, MIPAS reports larger HNO_3 monthly mean values than the other instruments, very likely related to the rejection of PSC-impacted measurements.

Overall, **ACE-FTS** shows the largest deviations relative to the MIM that change from negative in the LS (up to -30%) to positive in the US (up to +40%). The mean and median of the deviations in the US are well defined with a low spread as opposed to SMR and MIPAS.

In the tropical LSMS, **Aura-MLS** is mostly in the middle of the range, but in the US it shows negative deviations of up to -25%. At mid-latitudes, Aura-MLS is towards the end of the range given by all measurements, with mean differences varying between +6% in the LS and -25% in the US.

HIRDLS is mostly in the middle of the range, except for the tropical LS where it shows large positive deviations. At high latitudes, the HIRDLS seasonal cycle has a much smaller amplitude than the other instruments.

4.14 Peroxynitric acid – HNO_4

Peroxynitric acid (HNO_4) acts as a reservoir of NO_x and has very low atmospheric mixing ratios. It is thought to be formed exclusively in the gas phase by the reaction of HO_2 with NO_2 and a collision partner [*DeMore et al.*, 1997], and thus has indirect tropospheric sources such as soil emissions and fossil fuel burning (see *Section 4.11*). HNO_4 easily decomposes back into its precursor species and has a local lifetime of minutes to days depending on temperature. In the stratosphere, the reaction of HNO_4 with OH can lead to NO_x-catalysed loss of odd hydrogen radicals, presenting a significant sink of the latter [*Brasseur and Solomon,* 2005]. HNO_4 has a strong diurnal cycle in the US but is relatively constant in the LS and MS as demonstrated by examples of the diurnal HNO_4 cycle for three different pressure levels (**Figure 4.14.1**) derived with a chemical box model [*McLinden et al.*, 2010].

4.14.1 Availability of HNO_4 measurements

The assessment of the atmospheric HNO_4 annual mean state is based on the climatologies from ACE-FTS and MIPAS. **Tables 4.14.1** and **4.14.2** compile information on the availability of HNO_4 measurements, including time period, altitude range, vertical resolution, and references relevant for the data products used in this report. ACE-FTS measurements are split into local sunrise and sunset data, and MIPAS measurements are split into 10am and 10pm data in order to identify differences between the datasets attributable to different LSTs.

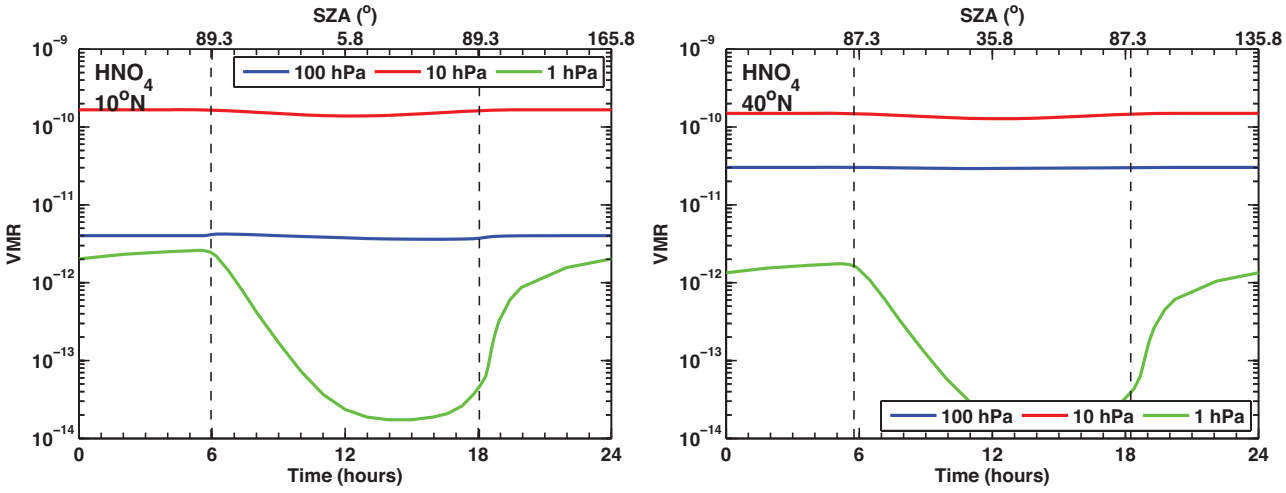


Figure 4.14.1: Diurnal HNO_4 cycle. *HNO_4 variations as function of LST are shown at 10°N and 40°N at 1, 10 and 100 hPa for March 15.*

Table 4.14.1: Available HNO_4 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.

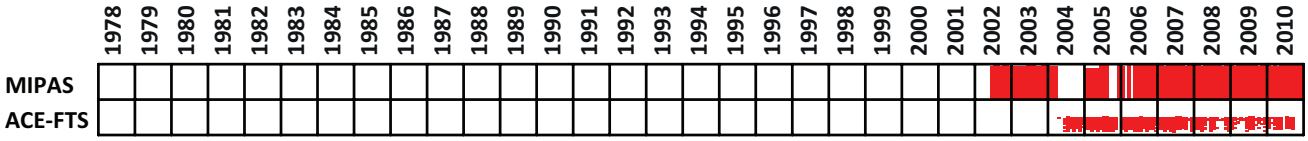


Table 4.14.2: Data version, time period, vertical range, vertical resolution, references, and other comments for HNO_4 datasets participating in the SPARC Data Initiative.

Instrument and data version	Time period	Vertical range	Vertical resolution	References	Additional comments
MIPAS **MIPAS(1)** V12 **MIPAS(2)** V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	Cloud top – 55 km	MIPAS(1): 4 – 6 km below 40 km; 10 – 15 km above MIPAS(2): 3 – 8 km below 40 km; 10 – 15 km above	*Stiller et al.*, 2007	
ACE-FTS V2.2	Mar 04 –	12 – 25 km	3 – 4 km	*Jones et al.*, 2011	Research product used here.

4.14.2 HNO_4 evaluations: Zonal mean cross sections and vertical profiles

Zonal mean cross sections are compared for the overlap period 2005-2010. As a first step, the annual mean MIPAS 10am and 10pm climatologies are compared to each other in order to identify the regions where varying LSTs may cause differences. **Figure 4.14.2** shows the cross sections of the two datasets, revealing very similar HNO_4 distributions below 10 hPa and significant differences above 10 hPa. For both LSTs the HNO_4 maximum is found in the tropical and subtropical LS to MS. While for the 10pm climatologies the maximum extends all the way through the MS with a steep gradient above, the 10am data decrease gradually above 10 hPa.

Differences of the MIPAS 10am and MIPAS 10pm datasets from their MIM are caused primarily by diurnal variations and are shown in **Figure 4.14.3**. Between 70 and 10 hPa, differences are very small (<±2.5%) suggesting only a very weak diurnal cycle in this region in accordance with the results from the chemical box model (**Figure 4.14.1**). At around 100 hPa in the tropics, slightly larger differences are found, increasing up to ±20%. The diurnal cycle has a large impact above 10 hPa causing differences of up to ±100% between the datasets corresponding to different LSTs. In the US, the 10pm measurements result in higher HNO_4 mixing ratios than the 10am measurements, while the situation is reversed in the LM. Differences around 100 hPa in the tropics and above 1 hPa are large in percentage units but should not be over interpreted since the mixing ratios here are very small as are the absolute diurnal variations.

A direct comparison between the MIPAS and ACE-FTS climatologies needs to be restricted to regions below 10 hPa and in the tropics to regions between 70 and 10 hPa. Outside of this range, different LSTs will cause differences in the HNO_4 mixing ratios. This conceals the true inter-instrument spread as the evaluations of MIPAS 10am and MIPAS 10pm climatologies have demonstrated. As the ACE-FTS datasets cover 130-20 hPa (70-20 hPa in the tropics), the comparison can be undertaken for the whole measurement range. **Figure 4.14.4** shows the annual zonal mean mixing ratios of the two MIPAS (10am and 10pm) and the two ACE-FTS (local sunrise and local sunset) datasets. While there is a good agreement on the magnitude of the HNO_4 mixing ratios and on the overall distribution, smaller differences exist leading to deviating tracer isopleths. While for MIPAS the tracer isopleths slope downwards from the tropics to the mid-latitudes, for ACE-FTS isopleths larger than 0.1 ppbv seem to slope upwards. Additionally, the ACE-FTS isopleths are noisier, which is very likely related to noisier profiles (compared to other ACE-FTS species) and to smaller sample sizes for the calculation of the monthly zonal mean values (compared to MIPAS mean values).

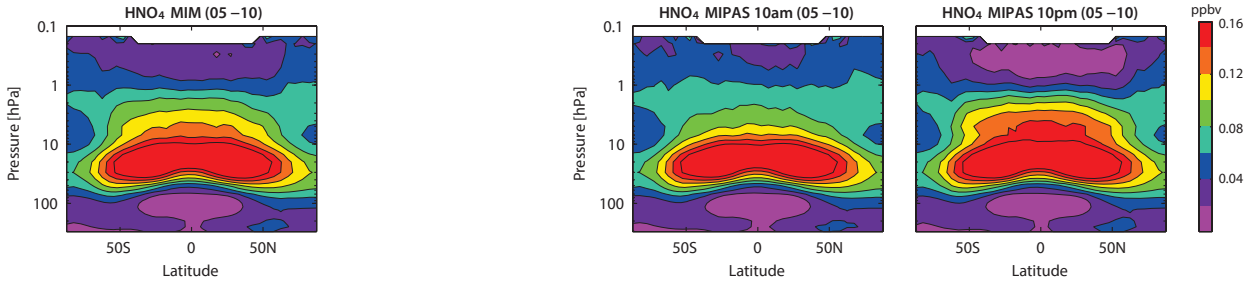


Figure 4.14.2: MIPAS cross sections of annual zonal mean HNO_4 for 2005-2010. Annual zonal mean HNO_4 cross sections are shown for MIPAS 10am and MIPAS 10pm in the right panels and for their MIM in the left panel.

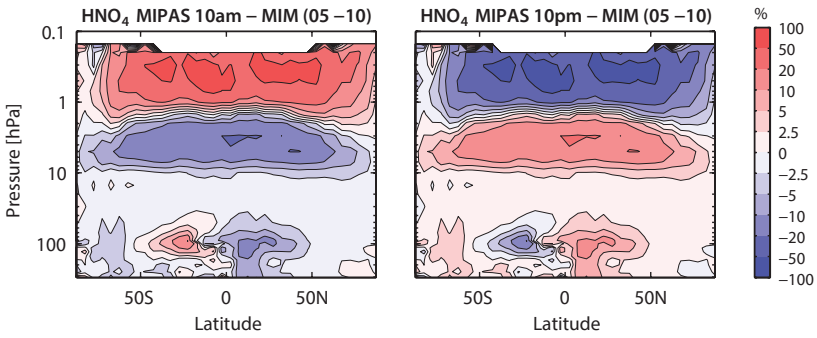


Figure 4.14.3: MIPAS cross sections of annual zonal mean HNO_4 differences for 2005-2010. *Annual zonal mean HNO_4 differences between MIPAS 10am, MIPAS 10pm and their MIM are shown.*

Differences between the four datasets and their MIM are shown in **Figure 4.14.5**. ACE-FTS detects smaller HNO_4 mixing ratios than MIPAS, leading to differences of up to ±50%. In all cases the differences between the two different instruments are larger than the differences caused by varying LSTs, as identified above. The only exception is at SH high latitudes where ACE-FTS local sunset data show much lower HNO_4 mixing ratios than ACE-FTS local sunrise data. This is likely due to the specific months that are sampled by sunrises and sunsets throughout the year.

4.14.3 Summary and conclusions: HNO_4

HNO_4 climatologies are available from two limb sounders, namely, ACE-FTS and MIPAS, with measurements at local sunrise/sunset and measurements at 10am/pm, respectively. The strong diurnal cycle above 10 hPa prevents a thorough comparison of the datasets in this region. Below 10 hPa, diurnal variations are weak allowing for a direct comparison of the datasets corresponding to different LSTs. For nearly all cases, ACE-FTS detects smaller and MIPAS larger HNO_4 mixing ratios leading to differences of up to ±50%. Comparisons of seasonal variations or interannual

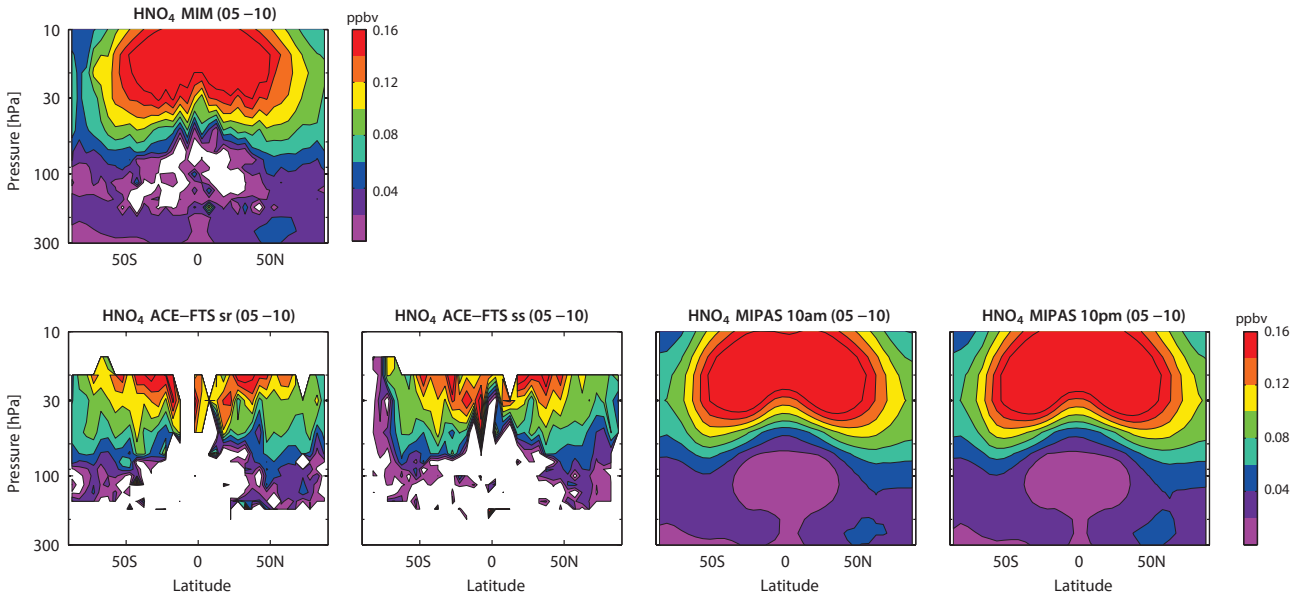


Figure 4.14.4: Cross sections of annual zonal mean HNO_4 for 2005-2010. *Annual zonal mean HNO_4 cross sections are shown for the MIM, ACE-FTS local sunrise (sr), ACE-FTS local sunset (ss), MIPAS 10am and MIPAS 10pm.*

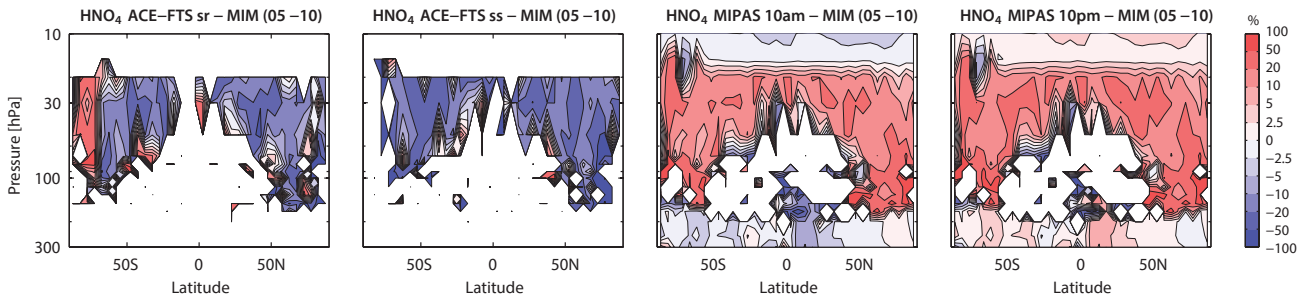


Figure 4.14.5: Cross sections of annual zonal mean HNO_4 differences for 2005-2010. *Annual zonal mean HNO_4 differences between the individual datasets and their MIM are shown.*

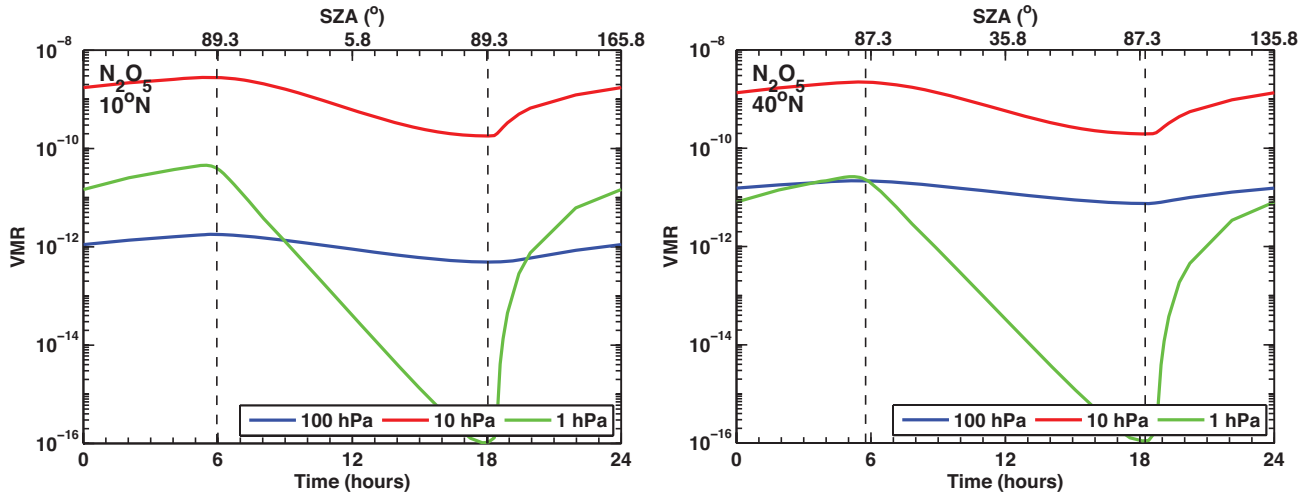


Figure 4.15.1: Diurnal N_2O_5 cycle. *N_2O_5 variations as function of LST are shown at 10°N and 40°N at 1, 10 and 100 hPa for March 15.*

variability are not possible due to the sparse ACE-FTS sampling pattern.

4.15 Dinitrogen pentoxide – N_2O_5

Dinitrogen pentoxide (N_2O_5) is a dominant night-time reservoir for the reactive nitrogen family. N_2O_5 is formed by the reaction of NO_2 with NO_3 and a collision partner, and thus has indirect tropospheric sources such as soil emissions and fossil fuel burning (see *Section 4.11*). Since NO_3 and N_2O_5 itself are rapidly photolysed by sunlight, N_2O_5 can only be formed and accumulated at night, reaching a maximum just before sunrise. During polar night, a significant fraction of NO_y in the lower to mid stratosphere will be available in the form of N_2O_5. In the presence of polar stratospheric clouds or volcanic aerosols, N_2O_5 can be converted to HNO_3, which has a large impact on stratospheric ozone destruction and can be irreversibly removed from the stratosphere by denitrification (see *Section 4.13*). The diurnal cycle of N_2O_5 for three different pressure levels derived with a chemical box model [*McLinden et al.*, 2010] is shown in **Figure 4.15.1**.

4.15.1 Availability of N_2O_5 measurements

The assessment of the atmospheric N_2O_5 annual mean state is based on the climatologies from ACE-FTS and MIPAS. **Tables 4.15.1** and **4.15.2** compile information on the availability of N_2O_5 measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report. ACE-FTS measurements are split into local sunrise and sunset data, and MIPAS measurements are split into 10am and 10pm data in order to identify differences between the datasets attributable to different LSTs.

4.15.2 N_2O_5 evaluations: Zonal mean cross sections and vertical profiles

Zonal mean cross sections are shown for the overlap period 2005-2010 for ACE-FTS local sunrise and sunset data and for MIPAS 10am and 10pm data (**Figure 4.15.2**). As expected from the diurnal cycle, ACE-FTS local sunrise data report overall the highest and ACE-FTS local sunset data the lowest N_2O_5 mixing ratios. MIPAS data at 10am

Table 4.15.1: Available N_2O_5 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.*

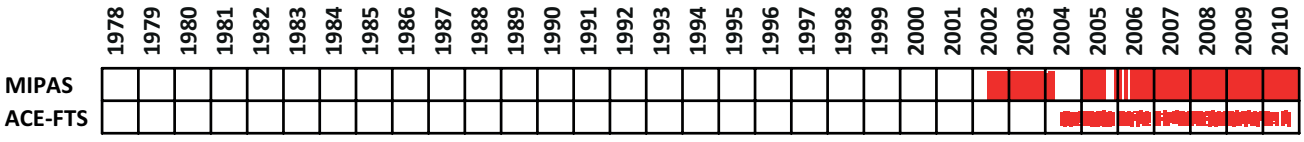


Table 4.15.2: Data version, time period, vertical range, vertical resolution, references and other comments for N_2O_5 datasets participating in the SPARC Data Initiative.

Instrument and data version	Time period	Vertical range	Vertical resolution	References	Additional comments
MIPAS **MIPAS(1)** V10 **MIPAS(2)** V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 cloud top altitude - 0.3 hPa (52 km)	 5 – 7 km below 2 hPa, 9 – 10 km above	 *Mengistu Tsidu et al.*, 2004	
ACE-FTS V2.2	Mar 04 –	15 – 40 km	3 – 4 km	*Wolff et al.*, 2008	

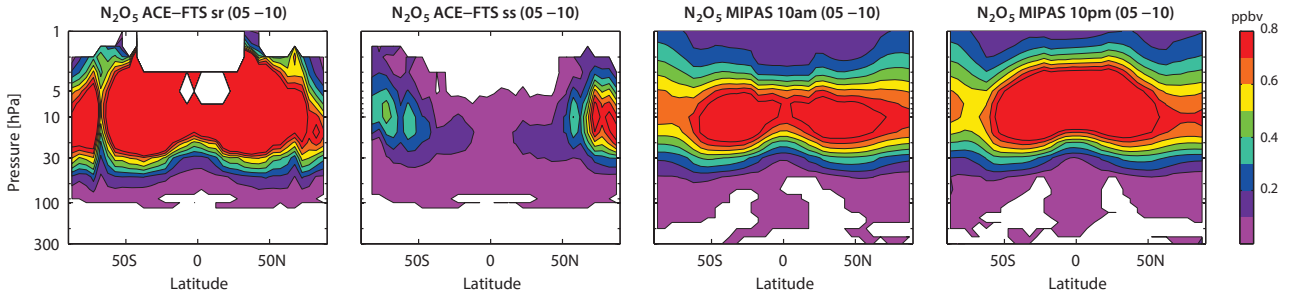


Figure 4.15.2: Cross sections of annual zonal mean N_2O_5 for 2005-2010. *Annual zonal mean N_2O_5 cross sections are shown for ACE-FTS local sunrise (sr), ACE-FTS local sunset (ss), MIPAS 10am, and MIPAS 10pm.*

and at 10pm are mostly in the middle of the range. An exception occurs at high latitudes. For example, in the NH at polar latitudes, ACE-FTS local sunset data are higher than both MIPAS climatologies, and in the US are also higher than ACE-FTS local sunrise data. These unexpectedly large values of the ACE-FTS local sunset dataset are caused by the fact that the annual mean includes more NH winter months, when N_2O_5 is higher, than the other three datasets. Similarly, at the SH polar latitudes, the annual mean ACE-FTS local sunset data are lower than expected related to the fact that not as much data from the SH winter is available and contributing to the annual mean. The comparison demonstrates that the annual mean ACE-FTS data at high latitudes, and in particular the local sunset data, are not representative and monthly means need to be evaluated instead.

Figure 4.15.3 shows monthly mean N_2O_5 profiles for NH high latitudes (65°N-70°N) during winter. The ACE-FTS local sunrise profiles show the largest values throughout most of the stratosphere, as expected from the diurnal cycle. Below 20 hPa, ACE-FTS local sunrise data are smaller than the two MIPAS profiles, contradicting the expectations based on N_2O_5 diurnal variations. Overall, differences in the MS and US are about ±20% between ACE-FTS local sunrise and MIPAS 10am/pm profiles and about ±50% between ACE-FTS local sunrise and local sunset profiles. In the latitude band 60°N-65°N, the instruments show a completely different picture, with MIPAS 10am profiles that are larger than ACE-FTS local sunrise data throughout the whole stratosphere (**Figure 4.15.4**). Similarly, at high SH latitudes in July and August, the ACE-FTS local sunrise data are lower than the MIPAS 10am and 10pm climatologies. Based on the current knowledge of the N_2O_5 diurnal cycle, the local sunrise data should give larger values than the other measurements. Therefore the comparisons in the 60°N/S-65°N/S latitude bands point to an inconsistency in either the MIPAS or the ACE-FTS data. Note, however, that local sunrise in this latitude band is between 8:30 and 9:30am reducing the diurnal variation-driven difference between ACE-FTS local sunrise and MIPAS 10am measurements. Furthermore, these latitude bands are characterised by large latitudinal gradients, and sparse sampling can have an impact on the representativeness of the monthly zonal mean values. As a result, sampling-driven differences can be larger than deviations caused by different LSTs causing the inconsistencies observed in the 60°N/S-65°N/S latitude bands. This assumption is further supported by the fact that the ACE-FTS sample size for the latitude bin 65°N-70°N is about 4 times larger than for the latitude bin 60°N-65°N.

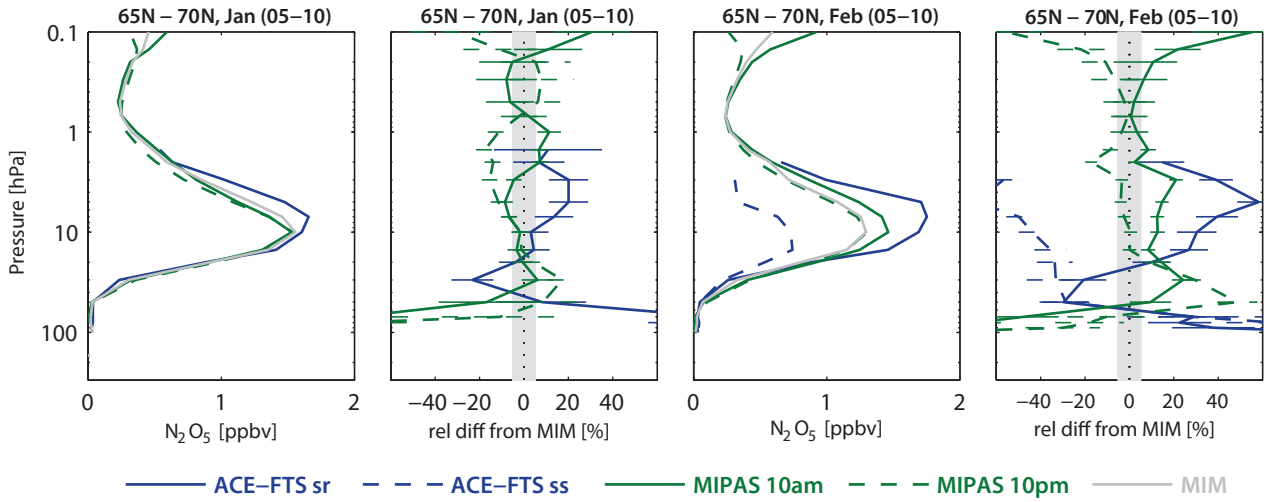


Figure 4.15.3: Vertical profiles of monthly zonal mean N_2O_5 for 2005-2010. *Zonal mean N_2O_5 profiles for 65°N-70°N in January and February are shown together with the differences between the individual instruments and the MIM profiles. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are within ±5%.*

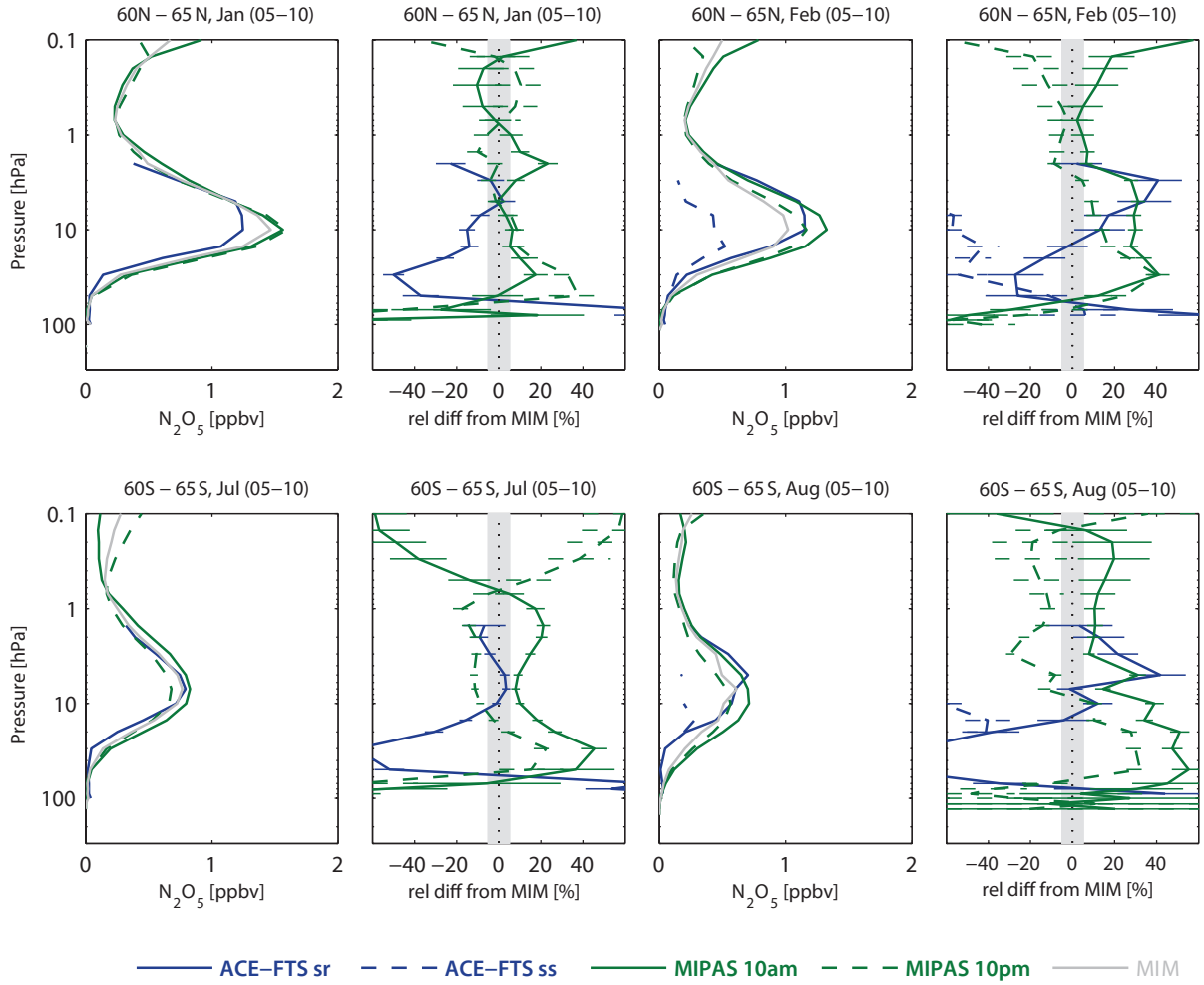


Figure 4.15.4: Vertical profiles of monthly zonal mean N_2O_5 for 2005-2010. *Zonal mean N_2O_5 profiles for 60°N-65°N in January and February and 60°S-65°S in July and August are shown in columns 1 and 3. Differences between the individual instruments and the MIM profiles are shown in columns 2 and 4. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are within ±5%.*

4.15.3 Summary and conclusions: N_2O_5

N_2O_5 climatologies are available from the limb sounders ACE-FTS and MIPAS, with measurements at local sunrise/sunset and measurements at 10am/pm, respectively. The strong diurnal cycle above 100 hPa prevents a thorough comparison of the datasets in this region. Below 100 hPa, diurnal variations are weak and the mixing ratios are low. For nearly all cases, ACE-FTS local sunrise detects the largest and ACE-FTS local sunset the lowest N_2O_5 mixing ratios, consistent with the diurnal cycle. ACE-FTS local sunrise is lower than the MIPAS 10am/pm datasets during winter at higher latitudes, in contradiction to the diurnal cycle. However, this issue may be attributable to sampling artefacts due to the latitudinal gradient, which is particularly pronounced in the latitude bins investigated. When using the datasets of trace gases with strong diurnal cycles, we recommend that additional information such as average latitude, day of month, sample size and LST (provided in the data files) should be taken into account.

4.16 Chlorine nitrate – $ClONO_2$

Chlorine nitrate ($ClONO_2$) is a reservoir species for stratospheric nitrogen and chlorine and therefore is important for ozone chemistry. $ClONO_2$ is produced from NO_x by reaction with ClO and a collision partner. During the daytime, $ClONO_2$ is photolyzed at ultraviolet wavelengths and reaches minimum abundances. During polar night, the presence of $ClONO_2$ reduces the amount of active chlorine and nitrogen and thus chemical ozone destruction. In the presence of polar stratospheric clouds, however, $ClONO_2$ can undergo heterogeneous reactions with H_2O and HCl to release chlorine into its chemically active form. The diurnal cycle of $ClONO_2$ for three different pressure levels derived with a chemical box model [*McLinden et al.*, 2010] is shown in **Figure 4.16.1**.

4.16.1 Availability of $ClONO_2$ measurements

The assessment of the atmospheric $ClONO_2$ annual mean state is based on the climatologies from ACE-FTS and MIPAS. **Tables 4.16.1** and **4.16.2** compile information on

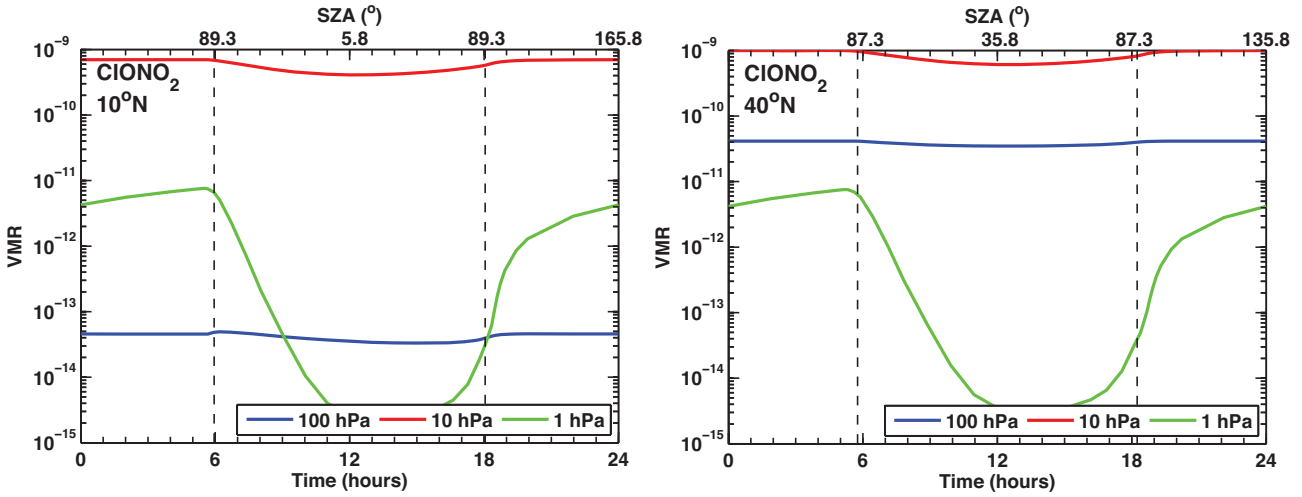


Figure 4.16.1: Diurnal $ClONO_2$ cycle. *$ClONO_2$ variations as function of LST are shown at 10°N and 40°N at 1, 10 and 100 hPa for March 15.*

the availability of $ClONO_2$ measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report. ACE-FTS measurements are split into local sunrise and sunset data, and MIPAS measurements are split into 10am and 10pm data in order to identify differences between the datasets attributable to different LSTs.

4.16.2 $ClONO_2$ evaluations: Zonal mean cross sections and vertical profiles

Zonal mean cross sections are shown for the overlap period 2005-2010 for ACE-FTS local sunrise and sunset data and for MIPAS 10am and 10pm data in **Figure 4.16.2**. The annual mean $ClONO_2$ distribution shows maxima at the mid-latitudes between 50 and 10 hPa. For the ACE-FTS local sunrise measurements, the maxima extend higher (up to 5 hPa). As expected from the diurnal cycle, the local sunrise data are overall the highest, whereas the ACE-FTS local sunset data and MIPAS 10am data show the lowest $ClONO_2$ mixing ratios in most regions. At high latitudes, the annual mean climatologies compare in a different way, and at the NH polar latitudes ACE-FTS local sunset data are higher than the other three datasets. Note that the same characteristic was found for the N_2O_5 climatologies (see *Section 4.15*). For both gases, the unexpectedly high values of the ACE-FTS local sunset datasets are caused by the fact that the annual mean includes more NH winter months, when $ClONO_2$ and N_2O_5 are higher than during the rest of the year. The comparison demonstrates that the annual mean ACE-FTS data at high latitudes, and in particular the local sunset data, are not representative and monthly means need to be evaluated instead.

Annual mean differences between the datasets and their MIM (**Figure 4.16.3**) in the low- and mid-latitudes show that in the LS, where the diurnal variations are expected to have only a small impact, the datasets agree quite well with differences of up to ±10%. Consistent with the diurnal cycle estimated from box model calculations, differences become significantly larger in the MS, in particular above 20 hPa for local sunrise and 10am measurements and above 10 hPa for local sunset and 10pm measurements. Differences below 100 hPa are also larger than in the LS. However, at these levels the impact of the diurnal cycle is supposed to be negligibly small and cannot explain the large deviations. Thus, the relative differences in the UT (or below 100 hPa) appear

Table 4.16.1: Available $ClONO_2$ 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.*

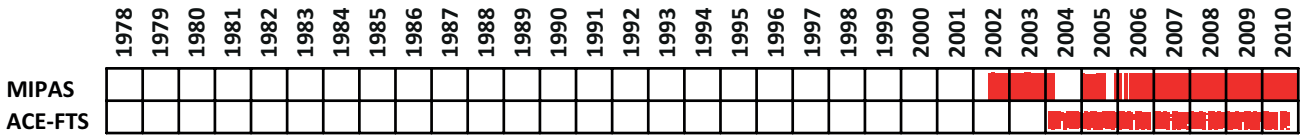


Table 4.16.2: Data version, time period, vertical range, vertical resolution, references and other comments for $ClONO_2$ datasets participating in the SPARC Data Initiative.

Instrument and data version	Time period	Vertical range	Vertical resolution	References	Additional comments
MIPAS					
MIPAS(1) V10	Mar 02 – Mar 04	8 – 65 km	3.2 – 8.5 km	*Höpfner et al.*, 2004, 2007	
MIPAS(2) V220	Jan 05 – Apr 12	8 – 65 km	2.5 – 9 km	*von Clarmann et al.*, 2009a	
ACE-FTS V2.2	Mar 04 –	12 – 35 km	3 – 4 km	*Wolff et al.*, 2008	

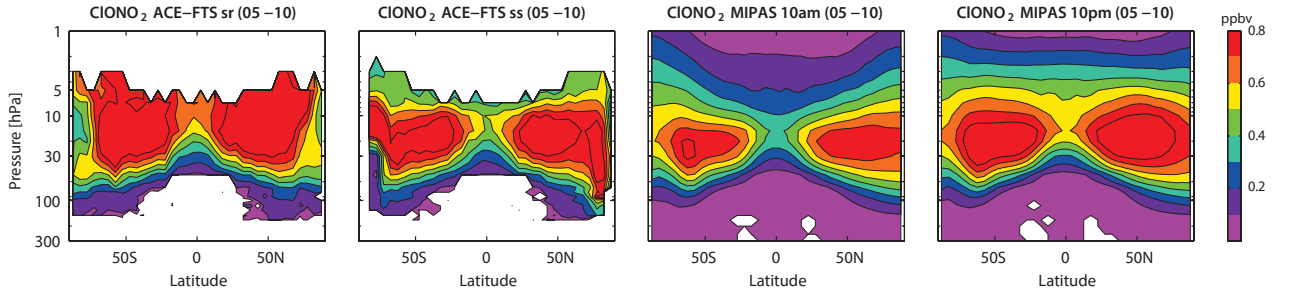


Figure 4.16.2: Cross sections of annual zonal mean $ClONO_2$ for 2005-2010. *Annual zonal mean $ClONO_2$ cross sections are shown for ACE-FTS local sunrise (sr), ACE-FTS local sunset (ss), MIPAS 10am and MIPAS 10pm.*

to be real inter-instrument differences further amplified by the small mixing ratios in this region.

Figure 4.16.4 shows monthly zonal mean cross sections for April and October. At the SH high latitudes, the April climatologies show maxima at 50 and 5 hPa. After the polar summer, the $ClONO_2$ mixing ratios are low and only start to increase with decreasing solar radiation. This increase is faster at around 5 hPa than at 10 hPa, possibly related to more nitrogen being available at the higher level. Together with the persistence of the summer time maximum at 50 hPa, this results in a vertical structure with a local minimum at around 10 hPa. ACE-FTS local sunrise measurements, available in the latitude band between 80°S and 90°S, confirm the structure observed by MIPAS. At the NH polar latitudes, the October climatologies from MIPAS show a very broad maximum and no local minimum as their SH autumn counterpart. ACE-FTS local sunrise data, however, suggest a similar structure in the NH with a local minimum around 10 hPa.

Differences of the monthly mean datasets with respect to their MIM (**Figure A4.16.1** in *Appendix A4*) are similar to the differences of the annual mean climatologies. Only at high latitudes, ACE-FTS monthly mean differences are smaller than the annual mean differences demonstrating the impact of irregular sampling on the annual mean evaluations at these latitudes.

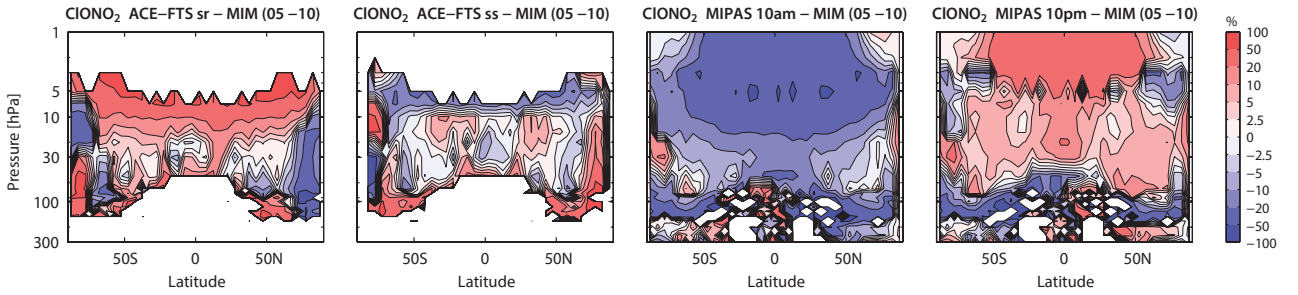


Figure 4.16.3: Cross sections of annual zonal mean $ClONO_2$ differences. *Annual zonal mean $ClONO_2$ differences for 2005-2010 between the datasets and the MIM are shown.*

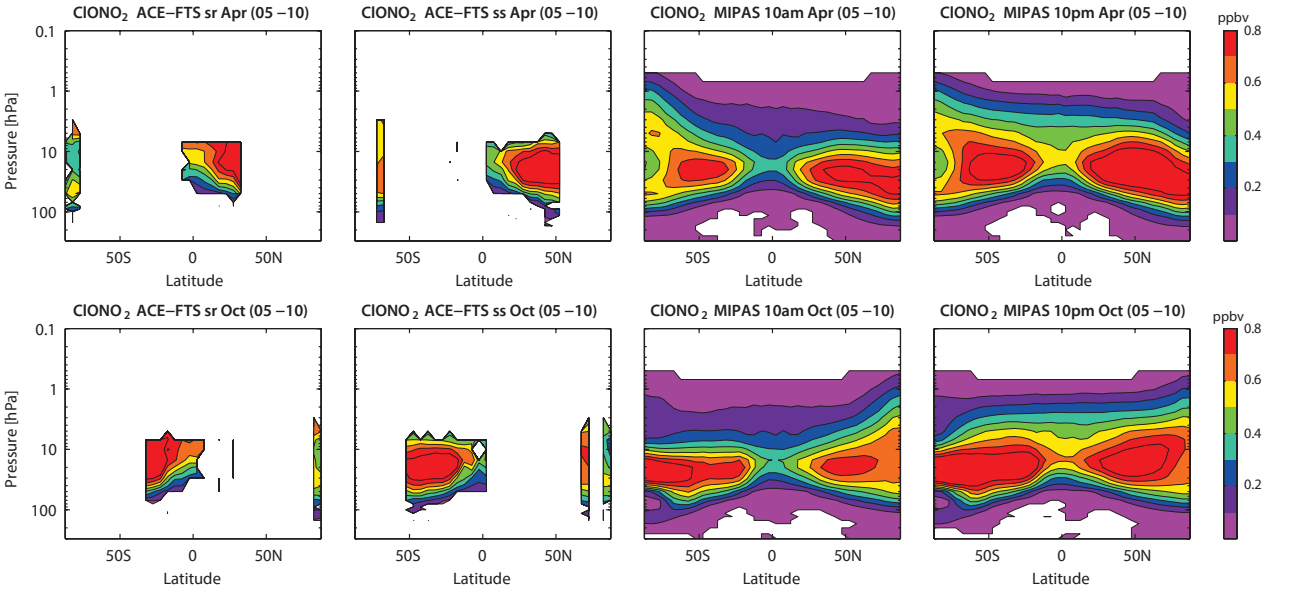


Figure 4.16.4: Cross sections of monthly zonal mean $ClONO_2$ for April and October. *Monthly zonal mean $ClONO_2$ cross sections are shown April (upper panels) and October (lower panels) 2005-2010 for ACE-FTS local sunrise (sr), ACE-FTS local sunset (ss), MIPAS 10am and MIPAS 10pm.*

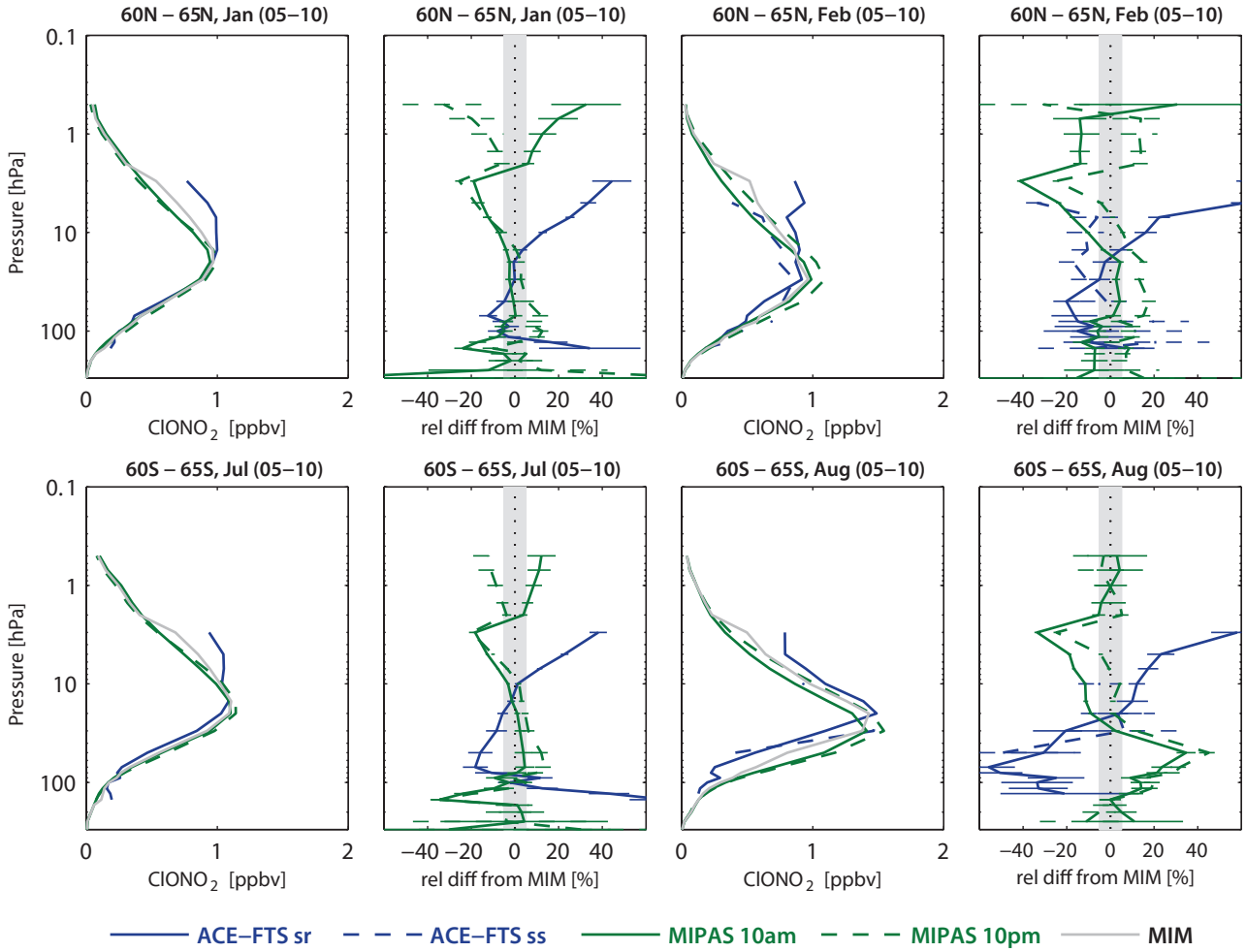


Figure 4.16.5: Vertical profiles of monthly zonal mean $ClONO_2$ for 2005-2010. *Zonal mean $ClONO_2$ profiles for 60°N-65°N in January and February and 60°S-65°S in July and August are shown in columns 1 and 3. Differences between the individual instruments and the MIM profiles are shown in columns 2 and 4. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are within ±5%.*

Figure 4.16.5 shows monthly mean $ClONO_2$ profiles at high latitudes (60°N-65°N and 60°S-65°S) during winter. The rapid increase of the deviations of ACE-FTS local sunrise data in the MS to US identify the level above which the diurnal cycle significantly impacts the climatology comparison. In the LS, where $ClONO_2$ has a longer lifetime, the ACE-FTS local sunrise profile seems to be shifted with regard to the MIPAS profiles, with the largest shift at 60°S-65°S in August. As a result the ACE-FTS local sunrise climatology is smaller than the MIPAS 10am and 10pm climatologies between 100 and 20 hPa. This larger difference, when compared to the July evaluations, could be related to the fact that considerably less data points are available for the construction of the ACE-FTS 60°S-65°S mean value in August than in July.

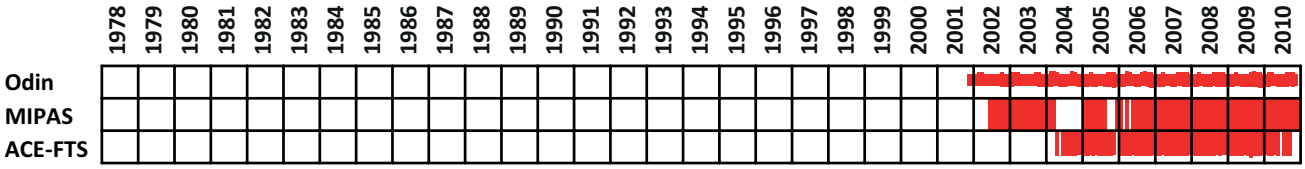
4.16.3 Summary and conclusions: $ClONO_2$

$ClONO_2$ climatologies are available from the limb sounders ACE-FTS and MIPAS with measurements at local sunrise/sunset and measurements at 10am/pm, respectively. The strong diurnal cycle in the MS and US prevents a thorough comparison of the datasets in this region. Below 10 hPa, diurnal variations are quite weak and the datasets agree well with differences from their MIM of up to ±10%. In the LS, however, relative differences are large partially related to the low mixing ratios. For nearly all cases, ACE-FTS local sunrise detects largest and MIPAS 10am lowest $ClONO_2$ mixing ratios consistent with the diurnal cycle. Only at higher latitudes during winter, ACE-FTS local sunrise is lower than the MIPAS 10am/pm datasets; a difference that cannot be explained by the diurnal cycle.

4.17 Total reactive nitrogen – NO_y

Total reactive nitrogen (NO_y) is the sum of all atmospheric reactive nitrogen species (NO_y = NO + NO_2 + NO_3 + HNO_3 + HNO_4 + $2N_2O_5$ + $ClONO_2$ + $BrONO_2$ + aerosol nitrate + …). Tropospheric NO_y originates largely from sources of NO and NO_2 released by fossil fuel burning, lightning, chemical processes in soils, and biomass burning (see *Sections 4.10* and *4.11*). The primary source of NO_y in the stratosphere is the oxidation of N_2O also originating from soil emissions (see *Section 4.4*). The dominant sink of stratospheric NO_y is through HNO_3 wash-out and sedimentation of HNO_3-containing PSCs. Reactive nitrogen species play an important role in stratospheric ozone

Table 4.17.1: Available NO_y 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.



chemistry through different mechanisms. First, NO_x is an ozone-depleting substance due to the catalytic NO_x cycle [*Crutzen*, 1970; *Johnston*, 1971]. Second, the presence of HNO_3 in an air parcel supports the formation of type I PSCs, which in turn enhance the release of chlorine from reservoir species into its ozone-destroying reactive form (see *Section 4.13*). Third, NO_2 reduces halogen-catalyzed ozone loss by converting reactive chlorine, bromine, and hydrogen compounds into more stable reservoir substances ($ClONO_2$, $BrONO_2$, HNO_3). The overall impact of reactive nitrogen on stratospheric ozone chemistry is determined by the balance of the different processes that are ultimately coupled since the amount of available NO_2 is, among other things, controlled by HNO_3 photolysis.

4.17.1 Availability of NO_y measurements

The assessment of the atmospheric NO_y annual mean state is based on the climatologies from ACE-FTS, MIPAS, and Odin. For ACE-FTS [*Jones et al.*, 2011] and MIPAS [*Funke et al.*, 2014] the climatologies are compiled from NO, NO_2, HNO_3, HNO_4, $2\times N_2O_5$, and $ClONO_2$ (six species climatologies) all directly measured by the instruments. The Odin climatology [*Brohede et al.*, 2008] is based on NO_2 from OSIRIS, HNO_3 from SMR and NO, $2\times N_2O_5$ and $ClONO_2$ taken from scan-based chemical box model simulations [*McLinden et al.*, 2010], while HNO_4 is not included (five species climatology). In addition, a MIPAS climatology not including HNO_4 is available (five species climatology) and analysed in the cross section evaluations in *Section 4.17.2*. For the remainder of *Section 4.17*, the MIPAS climatology based on six species is used. In all figures the instrument names will be augmented by lower indices indicating the number of species used to compile the climatology, *e.g.*, $MIPAS_5$ for the MIPAS five species climatology. Note that the ACE-FTS and Odin NO_y products are daytime climatologies and do not include polar night data, in contrast to MIPAS.

Table 4.17.1 summarises all available NO_y climatologies derived from satellite measurements, including time period and vertical range. Information on individual nitrogen species measurements used for deriving the NO_y climatologies, including time period, vertical range and resolution, and relevant references have been given in *Sections 4.10-4.16*.

4.17.2 NO_y evaluations: Zonal mean cross sections and vertical profiles

Zonal mean cross sections and vertical profiles are compared for the overlap period 2005-2010. **Figure 4.17.1** displays all annual mean NO_y climatologies. Note that the MIM consists of the six species climatologies from MIPAS and ACE-FTS and the five species climatology from Odin. The five species climatology from MIPAS is not included in order to prevent double counting of the MIPAS data. The two MIPAS climatologies are very similar and show differences from their MIM of less than 1% above 170 hPa and differences of less than 5% between 300 and 170 hPa. The small differences between the two climatologies suggest that the impact of HNO_4 is negligible in most regions, and that Odin can be compared directly to the six species climatologies from MIPAS and ACE-FTS.

Similar to the NO_2 distribution, the maximum values are found between 10 and 5 hPa with downward sloping isopleths towards higher latitudes due to isentropic mixing between the ascending and descending branches of the Brewer-Dobson circulation. Above 5 hPa, the distribution becomes more uniform with relatively flat isolines. Overall, the NO_y distributions of the three datasets agree very well with each other. The noisier isolines and the large values at the highest latitudes in ACE-FTS are related to sampling issues, *i.e.*, the region south of 85°S is only covered by ACE-FTS in March and April, resulting in anomalously high values when compiling annual means.

Figure 4.17.2 shows the differences of the three instruments from their MIM. Below 2 hPa, ACE-FTS is clearly lower than MIPAS with differences from the MIM ranging from -10% (ACE-FTS) to +10% (MIPAS), respectively. The Odin climatology changes from positive differences (below 10 hPa) to negative differences (above 10 hPa), mostly in the range of ±10%. Overall, there is good agreement between the three datasets, which is confirmed by the evaluation of the monthly mean climatologies (**Figure A4.17.1** and **A4.17.2** in *Appendix A4*) and consistent with evaluations [*Jones et al.*, 2011]. Above 2 hPa, only ACE-FTS and MIPAS data are available and differences can be larger reaching even ±50% at the high latitudes. The differences show a strong latitudinal gradient with ACE-FTS on the high side and MIPAS on the low side in the tropics and subtropics and *vice versa* poleward of about 50°.

Detailed evaluations of monthly zonal mean differences for individual latitude bands (0°N-5°N and 70°S-75°S for August) are shown in **Figure 4.17.3**. The comparison of the tropical profiles confirms the very good agreement in a wide altitude range from 40 to 0.3 hPa. While ACE-FTS and MIPAS agree also very well below 40 hPa, Odin shows clearly larger values with differences of up to +30%. The maximum at 5 hPa as seen by MIPAS is about 5 to 10% too large when compared to the other two climatologies.

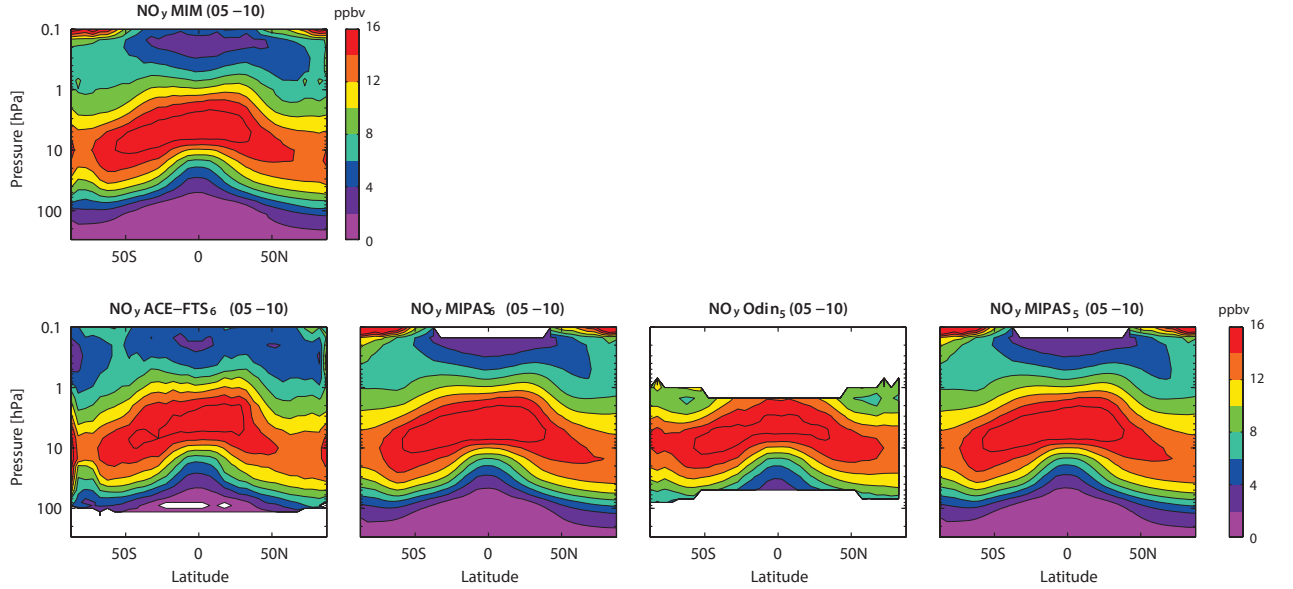


Figure 4.17.1: Cross sections of annual zonal mean NO_y for 2005-2010. *Annual zonal mean NO_y cross sections are shown for the MIM in the upper panel and for ACE-FTS_6, $MIPAS_6$, $Odin_5$ and $MIPAS_5$ in the lower panels. Note that $MIPAS_5$ is not included in the MIM.*

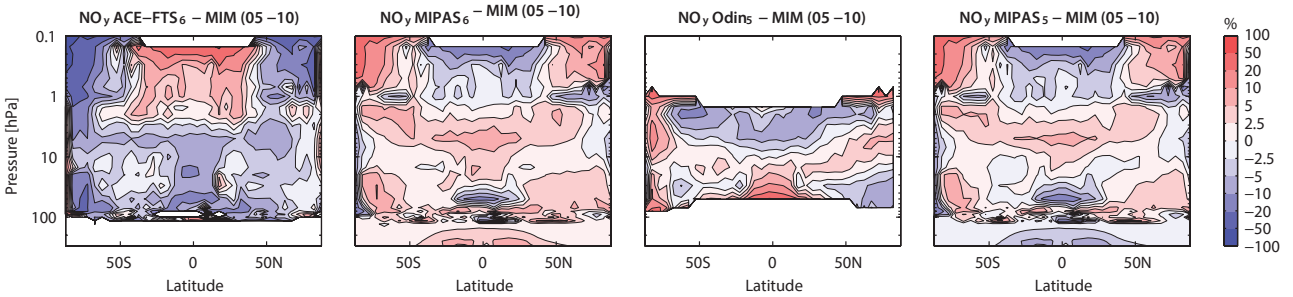


Figure 4.17.2: Cross sections of annual zonal mean NO_y differences for 2005-2010. *Annual zonal mean NO_y differences between the individual instruments and the MIM are shown. Note that $MIPAS_5$ is not included in the MIM.*

At high latitudes, the datasets agree less well than in the tropics. In the SH high latitude winter displayed here, NO_y has an S-shaped profile caused by Energetic Particle Precipitation NO_y intrusions, which is seen differently by the three instruments. In particular, the position of the MS maximum varies between 20 hPa (ACE-FTS and Odin) and 3 hPa (MIPAS). Furthermore, Odin shows the USLM minimum at a much lower level (2 hPa) than the other two instruments (0.5-0.2 hPa). Overall, the high latitude differences range between ±20% except for the LS where ACE-FTS has a large offset of around -60%. The similarity of the August monthly mean deviations and the annual mean deviations indicates that the negative offset found for annual mean ACE-FTS at the SH high latitudes between 100 and 50 hPa (**Figure 4.17.2**) is not related to annual sampling patterns.

At 1 hPa, the largest fraction of NO_y is located in the NO_x family while the other nitrogen species have smaller atmospheric abundances. Both nitrogen families, NO_y and NO_x, are available from MIPAS and ACE-FTS and are displayed in **Figure 4.17.4** as meridional profiles at 1 hPa for April 2005-2010. The NO_y profiles from the two instruments (upper row) agree very well, with differences of less than ±5% except for the high latitudes. The comparison of the NO_x profiles (middle row) also shows good agreement, but slightly larger relative differences of up to ±10% are found. The 1 hPa level has been chosen here because of the low diurnal variations NO_x displays at this level (±5%) in the tropics and mid-latitudes. The remaining fraction of the NO_y family (lower row) is calculated here as the difference between the NO_y and NO_x fields and is referred to as the residual in the following. The residual consists basically of the sum of HNO_3, HNO_4, $2N_2O_5$ and $ClONO_2$ and is expected to have similar (but inverted) diurnal variations as NO_x itself. The absolute differences of the MIPAS and ACE-FTS residuals are smaller than the absolute differences of the NO_y and NO_x families. As expected, the residuals act to remove the impact of the diurnal cycle and reduce the differences between the two instruments found in NO_x when added to give NO_y. Relative to NO_x the residual deviations are quite small (~5%) and are therefore consistent with deviations expected due to the diurnal cycle in the species and different LSTs of the measurements.

4.17.3 NO_y evaluations: Seasonal cycles

Figure 4.17.5 displays the NO_y seasonal cycles at high and tropical latitudes at 50 hPa. NO_y exhibits a strong seasonal cycle in the SH high latitude LS due to chemistry and transport effects. During polar night, descending air masses

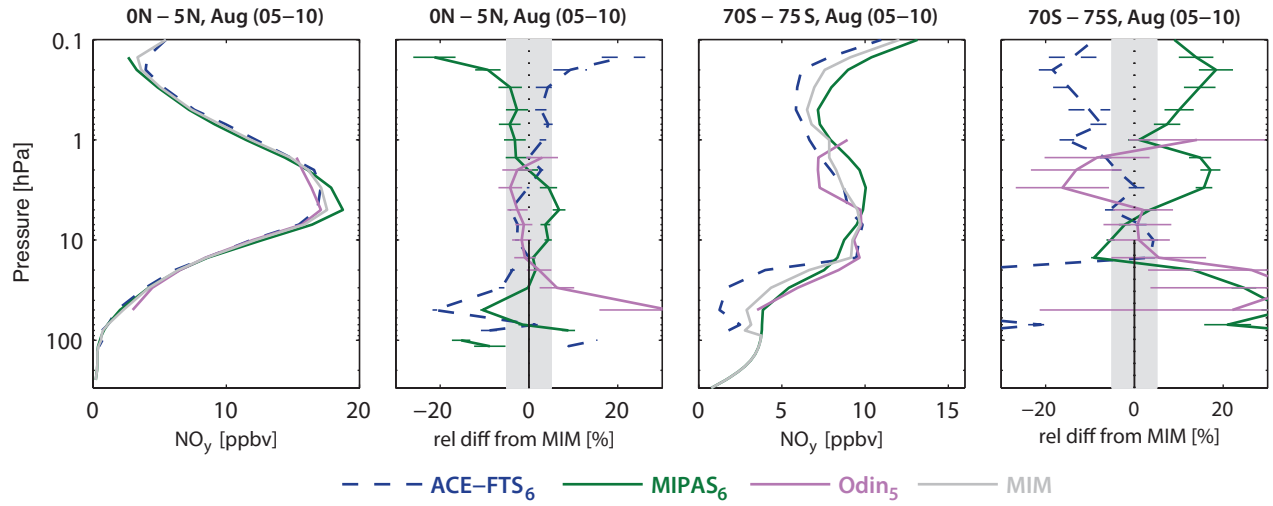


Figure 4.17.3: Vertical profiles of monthly zonal mean NO_y for 2005-2010. *Zonal mean NO_y profiles for 0°N-5°N and 70°S-75°S for August are shown in panels 1 and 3. Relative differences between the individual instruments and the MIM profiles are shown in panels 2 and 4. Bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are within ±5%.*

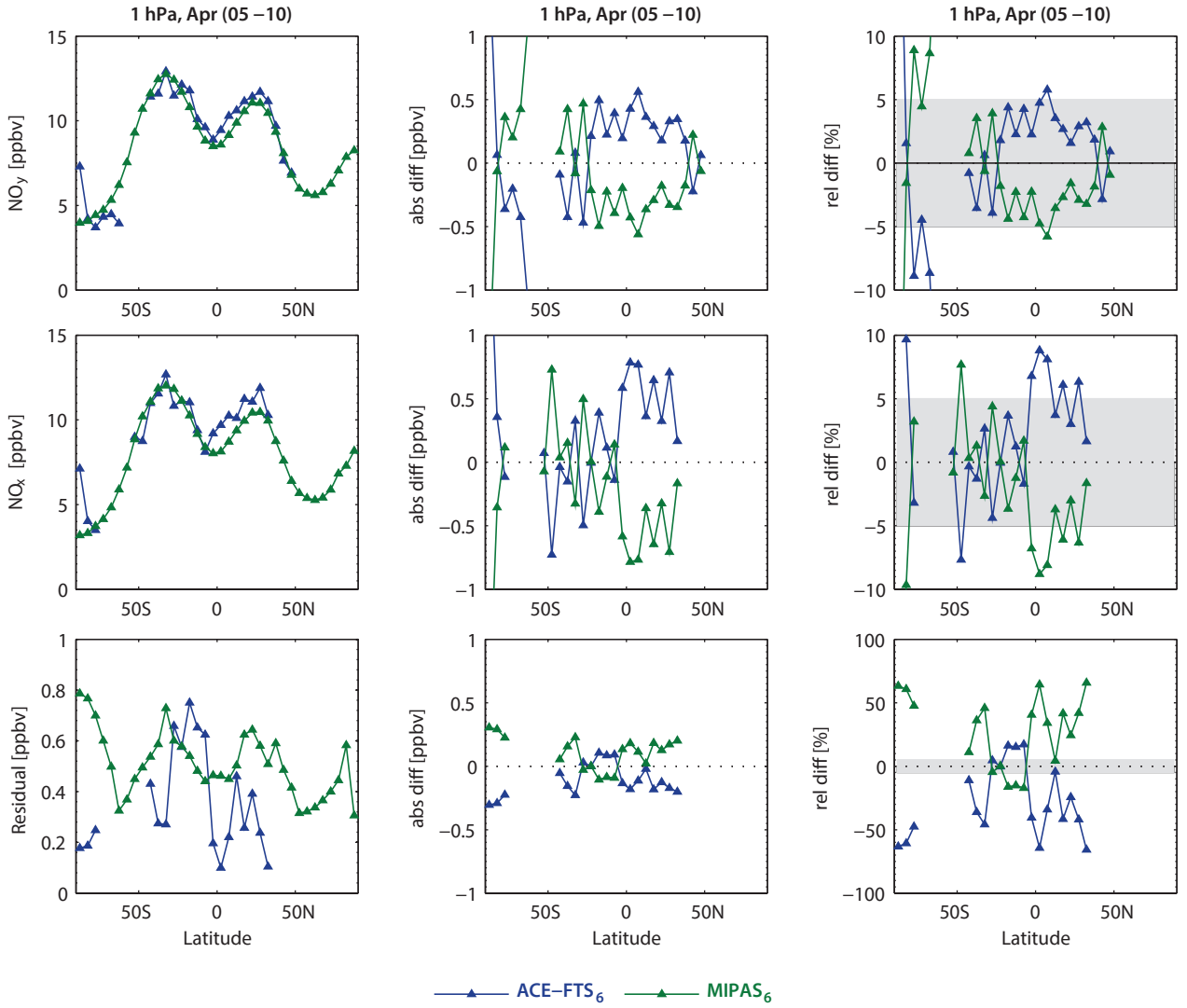


Figure 4.17.4: Meridional profiles of monthly zonal mean NO_y, NO_x, and their residual for 2005-2010. *Meridional profiles of monthly zonal mean NO_y, NO_x, and their residual at 1 hPa for April 2005-2010 are shown in the left column. Absolute (middle column) and relative (right column) differences between the two instruments (ACE-FTS and MIPAS) and their MIM are also shown for the two nitrogen families and the residual. The grey shaded area indicates where relative differences are smaller than ±5%.*

within the polar vortex bring higher NO_y concentrations into the LS, but this increase is counteracted by the formation of type I PSCs, which removes HNO_3 particles from the gas phase and causes a pronounced minimum in SH winter/spring. All three datasets show the same overall shape of the seasonal cycle. However, there are some differences in SH summer after renitrification when values stay approximately constant over three months; NO_y from ACE-FTS is smaller than from the other two datasets. There are also differences in SH winter when Odin shows larger values, and does not report the same level of HNO_3 removal as the other two datasets.

In the NH polar regions, the less frequent PSC type I formation leads to the opposite seasonal cycle. Here the downward transport dominates in autumn, producing a weak maximum at the beginning of winter. Later in the NH winter, removal of HNO_3 leads to a weak decline of the total nitrogen, which levels off in May. The main difference from the Antarctic counterpart is that there is no indication of strong renitrification in late spring and summer. MIPAS and ACE-FTS agree reasonably well in terms of amplitude but show some offset regarding phase and absolute values - MIPAS is larger and starts to decline earlier. Odin, on the other hand, displays an opposite seasonal cycle with a minimum in winter and spring that is not consistent with our understanding of polar processes determining the NO_y abundance. Note that the SMR HNO_3 seasonal cycle in this region is in general consistent with the other HNO_3 datasets with slightly smaller values during NH winter (see **Figure 4.13.7** in *Section 4.13*). Additionally, the OSIRIS NO_2 seasonal cycle in this region agrees very well with the other NO_2 datasets. Thus the deviations of the Odin NO_y seasonal cycle are not consistent with the underlying SMR HNO_3 or OSIRS NO_2 data, but are very likely introduced through the use of the photochemical model during the climatology compilation.

In the tropics, transport variations are expected to cause a weak annual cycle as seen by MIPAS. In addition to the relatively large absolute deviations between the datasets, there is no agreement on the seasonal signal, and both ACE-FTS and Odin show a semi-annual cycle with maxima in June/July and December/January, respectively. Note that 50 hPa is at the lower edge of the SMR measurement range for the HNO_3 product used to derive the Odin NO_y climatology. At higher levels (*e.g.*, 30 hPa) in the tropics, the Odin seasonal cycle agrees better with MIPAS, exhibiting the expected annual cycle. However, in the NH polar regions Odin agrees better with the other two datasets only above 10 hPa.

4.17.4 NO_y evaluations: Interannual variability

In addition to the absolute differences between the climatologies, it is important to evaluate how well the instruments detect signals of interannual variability. **Figure 4.17.6** shows the deseasonalised NO_y anomaly time series from 2005 to 2010 at polar and tropical latitudes at different pressure levels. The comparison in the tropics (10°S-10°N) at 10 hPa shows excellent agreement between the three datasets with only occasional deviations for ACE-FTS. The pronounced QBO signal is recorded with the same amplitude and phase by all satellite datasets.

At polar latitudes, the datasets also agree well, but stronger deviations between the three anomaly time series can be seen. For most years, the largest interannual anomalies occur during polar winter, related to the strong interannual variability of the chemical and dynamical processes that impact the polar stratospheric nitrogen budget. In both hemispheres, differences between ACE-FTS and MIPAS are also largest in polar winter months, a time period that is not covered by Odin. The maximum anomalies in winter are followed by a slow decay of the signal during spring and summer until the next autumn, which is a characteristic of HNO_3 (see *Section 4.13*). During spring and summer the three instruments agree better, and in most cases follow the winter anomaly signal given by MIPAS and not the one suggested by ACE-FTS.

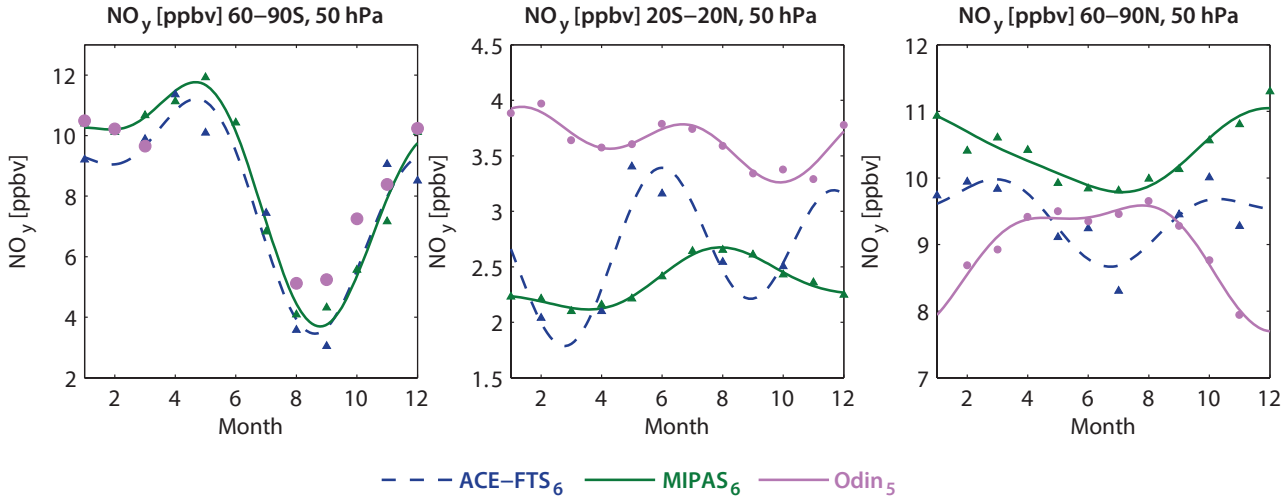


Figure 4.17.5: Seasonal cycle of NO_y for 2005-2010. *Seasonal cycle of monthly zonal mean NO_y for 60°S-90°S (left column), 20°S-20°N (middle column) and 60°N-90°N (right column). At SH high latitudes at 50 hPa, Odin does not provide sufficient coverage for fitting a seasonal cycle.*

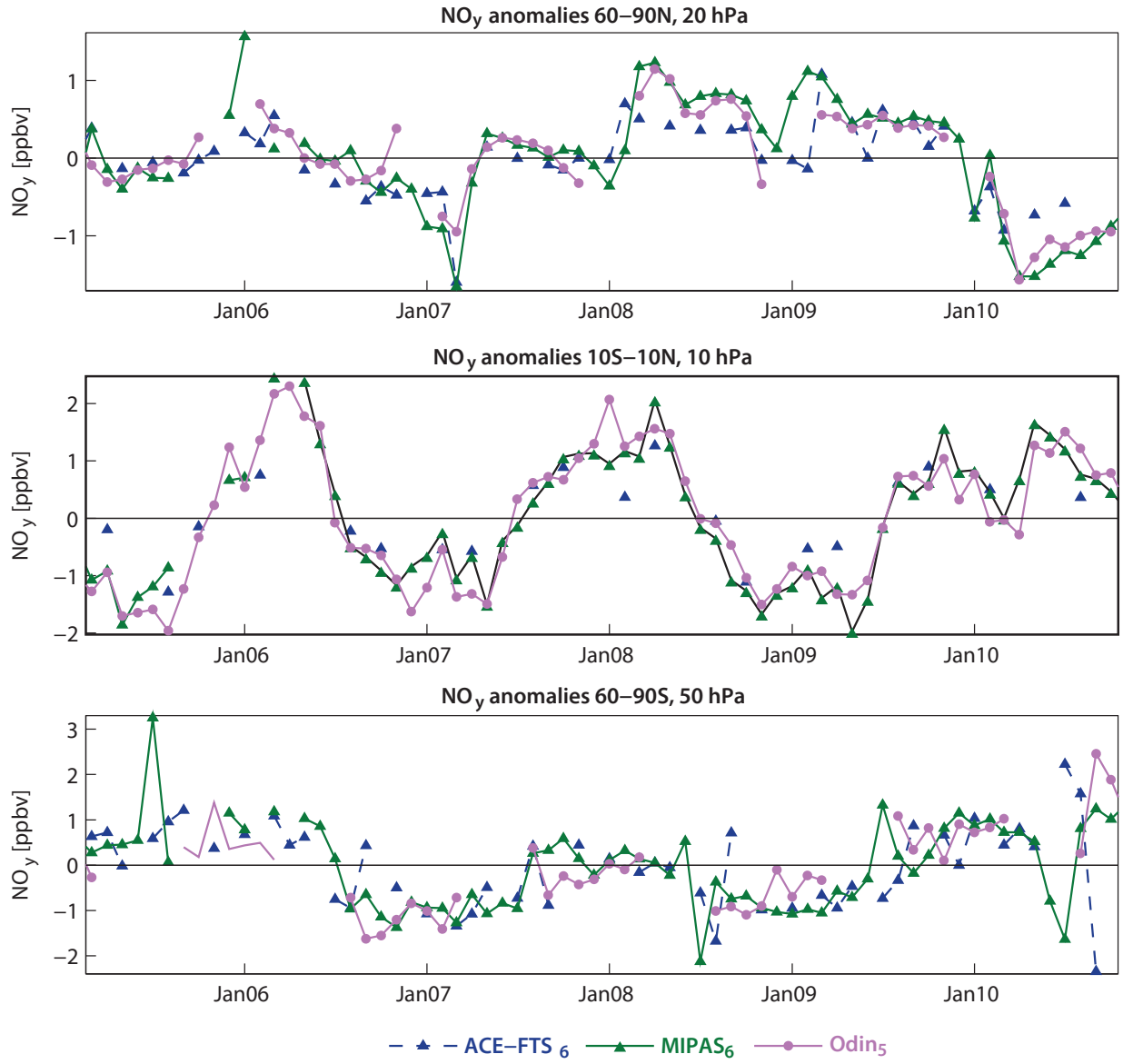


Figure 4.17.6: Time series of NO_y anomalies for 2005-2010. *Monthly zonal mean deseasonalised NO_y anomalies at 20 hPa for 60°N–90°N, at 10 hPa for 10°S-10°N, and at 50 hPa for 60°S-90°S.*

4.17.5 Summary and conclusions: NO_y

A comprehensive comparison of three NO_y profile climatologies (from ACE-FTS, MIPAS, and Odin) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the NO_y mean state and important characteristics of the individual datasets are presented in the following summary in the form of two synopsis plots. The first summary plot (**Figure 4.17.7**) provides information on the NO_y mean state and the uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.17.8**) shows specific inter-instrument differences in the form of the deviations of the instrument climatologies relative to the MIM climatology. For each instrument and selected region, the deviation relative 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 discussion of the rationale behind these summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The assessment of the atmospheric NO_y annual mean state is based on three climatologies with two datasets (ACE-FTS, MIPAS) constructed purely from measurements while the other dataset (Odin) is based upon NO_2 and HNO_3 measurements and chemical box model simulations of the remaining NO_y species.

Lower stratosphere (100-30 hPa)

In the LS, the NO_y abundance decreases with decreasing altitude, but the agreement in the mid-latitudes and NH polar latitudes is overall very good, with a spread of ±5%. The

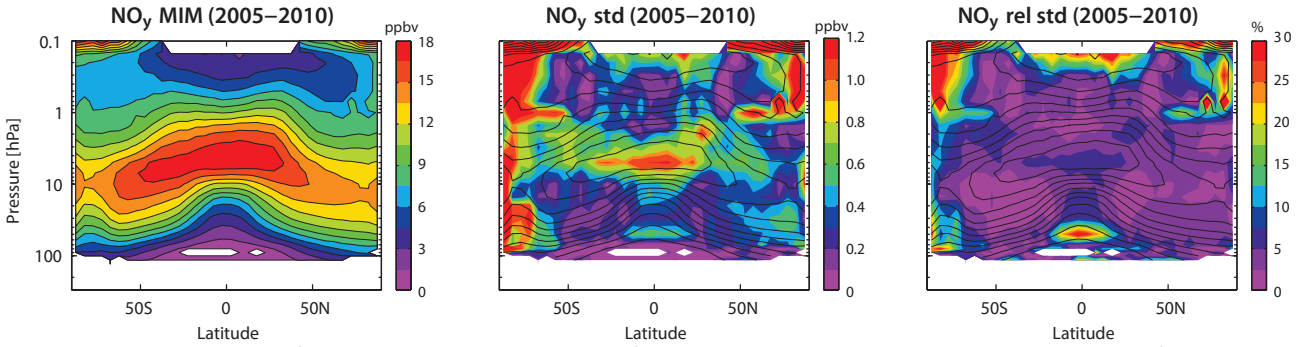


Figure 4.17.7: Summary of NO_y annual zonal mean state for 2005-2010. *Annual zonal mean cross section of the NO_y MIM for 2005-2010 is shown in the left panel. Additionally, the standard deviation over all three instruments is presented in the middle panel. Relative standard deviation (calculated by 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. Instruments included are ACE-FTS$_6$, MIPAS$_6$ and OSIRIS/SMR(Odin$_5$). The MIM and standard deviation are only displayed for regions where at least two instruments provide measurements.*

three datasets show a larger spread in the tropical LS with an inter-instrument spread of up to ±30% and at the SH high latitudes related to large deviations during polar winter.

Middle and upper stratosphere (30-1 hPa)

The uncertainty in our knowledge of the atmospheric NO_y annual mean state is smallest in the MS/US (**Figure 4.17.7**, right panel) with a 1σ multi-instrument spread in this region of mostly up to ±5%, in some regions up to ±7.5%. In the SH highest latitude bands (south of 80°S) the spread can reach values of ±12.5%.

Lower mesosphere (1-0.1 hPa)

At high latitudes, and in particular in the SH, the instruments show larger deviations with a spread of up to ±30%. In the mid-latitudes and tropics, on the other hand, the datasets agree very well and deviations are comparable to the MS/US.

Instrument-specific conclusions

ACE-FTS is generally lower compared to MIPAS and Odin, with mean deviations of around -5% except for the tropical and mid-latitude US and LM, where it shows positive deviations from MIPAS. The solar occultation instrument displays a similar seasonal cycle as MIPAS at high latitudes but an unrealistic semi-annual cycle in the tropical LS. Inter-annual anomalies from ACE-FTS differ from MIPAS in polar winter in a way that might be unrealistic considering the development of the anomaly in the following months.

MIPAS measurements are mostly on the high side (+5%) except for the tropical and mid-latitude US and LM. A

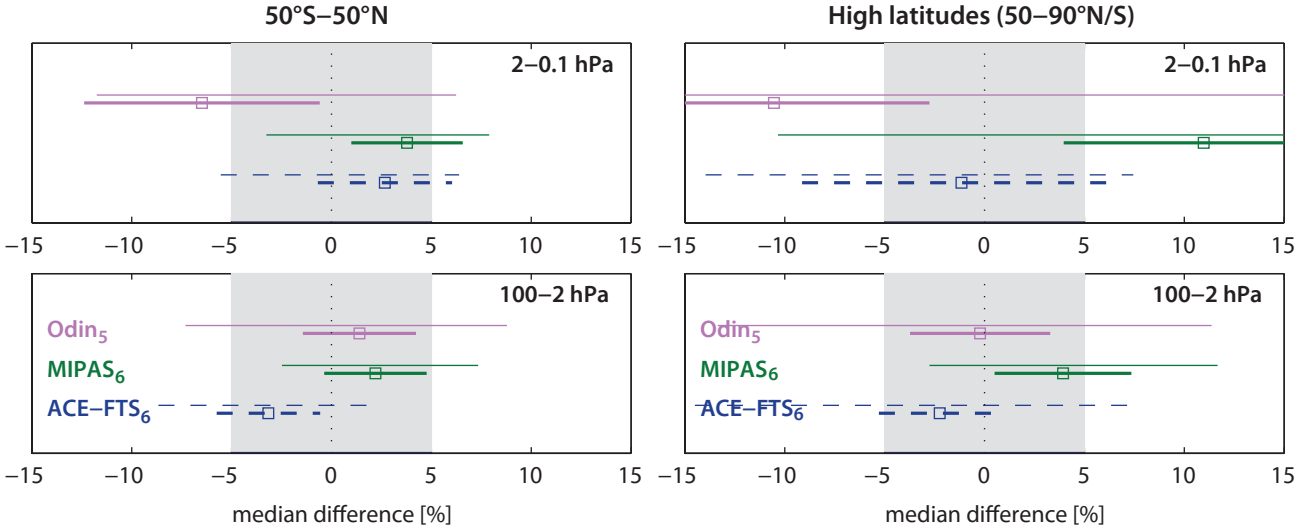


Figure 4.17.8: Summary NO_y differences for 2005-2010. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the lower (50°S-50°N) and the higher (50°S-90°S and 50°N-90°N) latitudes and for two different altitude regions from the LS up to the LM between 100 and 0.1 hPa for the reference period 2005-2010. The grey shaded area indicates where mean and median relative differences are within ±5%.*

comparison of the MIPAS and ACE-FTS nitrogen families (NO_y and NO_x) at 1 hPa suggests that the NO_y differences between the two instruments are consistent with the slightly larger NO_x differences, which are within the uncertainty given by the diurnal variations and differences in local times of measurements.

The **Odin** NO_y climatology covers the LS to US and is mostly in the mid-range between the other two datasets. In SH winter/spring, Odin reports higher values compared to ACE-FTS and MIPAS but shows similar signals of interannual variability. In the NH middle and upper stratosphere, the NO_y seasonal cycle from Odin shows severe differences compared to the other two datasets.

4.18 Hydrogen chloride – HCl

Hydrogen Chloride (HCl) is the main reservoir gas for reactive chlorine in the Earth's atmosphere. In the troposphere, the main sources of HCl are the ocean (acidification of salt spray) and volcanic activity; however, HCl is readily removed by wet scavenging, and hence it has a short lifetime and a relatively low concentration in the troposphere. In the stratosphere, HCl is the most abundant inorganic chlorine species with a lifetime of approximately one month. Its main source is the photolysis of anthropogenic chlorofluorocarbons (CFCs, see *Sections 4.5* and *4.6*) or hydrochlorofluorocarbons (HCFCs) and the subsequent reaction of the generated Cl-radical with methane (CH_4, see *Section 4.3*). HCl is processed on polar stratospheric clouds (PSCs) during winter, thereby releasing reactive chlorine atoms that can destroy ozone, and create the Antarctic ozone hole during SH springtime. Hence, the abundance of HCl (scaled to yield total reactive chlorine Cl_y) in the stratosphere is often used as a diagnostic for chemistry-climate models [*SPARC*, 2010].

4.18.1 Availability of HCl measurements

Measurements of HCl available to the SPARC Data Initiative start in 1991. The longest time series comes from HALOE, and the only instruments that overlap with this data record and have ongoing measurements are ACE-FTS and Aura-MLS. SMILES offers two HCl isotopologue products (from $H^{35}Cl$ and $H^{37}Cl$, named SMILES(1) and SMILES(2), respectively) for comparison during its mission time between October 2009 and April 2010.

Tables 4.18.1 and **4.18.2** compile information on the availability of HCl measurements including time period, height range, vertical resolution, and references relevant for the data product used in this report.

4.18.2 HCl evaluations: Zonal mean cross sections, vertical and meridional profiles

The HCl zonal mean cross section evaluations have been carried out using climatologies averaged over the years

Table 4.18.1: Available HCl 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.*

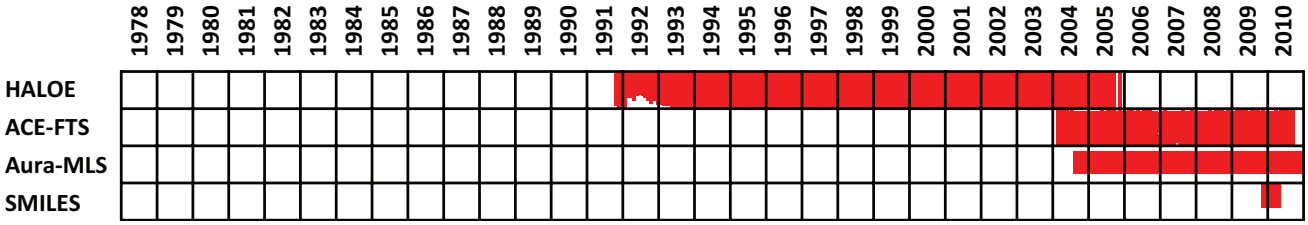


Table 4.18.2: Time period, vertical range, vertical resolution, references and other comments for HCl measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
HALOE V19	Oct 91 – Nov 05	up to 70 km	3.5 km	*Grooß and Russell*, 2005	
Aura-MLS V3.3	Jul 04 –	100 – 0.32 hPa (16 - 57 km)	3 km (in LS) 4 – 6 km (above 1 hPa)	*Froidevaux et al.*, 2008b (for V2.2) *Livesey et al.*, 2011 (for V3.3 update)	The MLS continuous band 14 HCl data used here are not suitable for upper stratospheric trend studies (see refs.). Values at 147 hPa are biased very high at low latitudes, but are probably more useful at high latitudes; all values below 100 hPa were flagged with bad value (-999.) for the current climatology.
ACE-FTS V2.2	Mar 04 –	8 – 57 km	3 – 4 km	*Mahieu et al.*, 2008	
SMILES HCl-35 V2.1.5 SMILES(1) HCl-37 V2.1.5 SMILES(2)	Oct 09 – Apr 10	100 hPa – 0.0001 hPa (16 – 110 km)	3 km	*Kreyling et al.*, 2013	Diurnal variation is observed.

2002-2008. HCl exhibits relatively small trends during this time period and averaging over a longer time period decreases the sampling bias. As shown in **Figure A4.18.1** in *Appendix A4*, the instruments show similar although somewhat noisier behaviour when only the year 2005 is considered. Additionally, vertical and meridional profiles are evaluated.

Note that SMILES measured over only a few months during 2009 and 2010 and hence was not included in the calculation of the MIM. SMILES(1) HCl is based on $H^{35}Cl$ isotopologue data and multiplied with a scaling factor of 1/0.7578, while SMILES(2) is based on $H^{37}Cl$ isotopologue data and multiplied with scaling factor of 1/0.2422. The scaling factors account for the natural isotopic abundance in the atmosphere.

HALOE, ACE-FTS, Aura-MLS, and SMILES (2002-2008)

Figure 4.18.1a shows the annual zonal mean HCl climatologies for 2002-2008. The relative differences between the instruments and the MIM are displayed in **Figure 4.18.1b**. **Figure 4.18.1a** reveals that HCl is a mostly stratospheric trace gas exhibiting very low mixing ratios in the troposphere. HCl increases with increasing altitude, opposite to its source gases that consist mostly of CFCs (see *Sections 4.5* and *4.6*). Below 5 hPa, the trace gas isopleths slope downwards between the tropics and the extra-tropics as expected for intermediate to long-lived trace gases. The instruments seem to agree well through most of the LS and MS, however show less agreement in the USLM. ACE-FTS, for example, shows much larger values at highest altitudes than the other instruments providing measurements in this region. This is a known feature reported in *Froidevaux et al.* [2008b] and this difference is reduced in the more recent ACE-FTS data version 3. The two SMILES datasets reveal very similar structures to the overall HCl distributions, however do not show increasing values in the LM as seen in the other instruments.

The difference plots in **Figure 4.18.1b** reveal upon close inspection that Aura-MLS and ACE-FTS show excellent agreement within ±2.5% through most of the stratosphere and into the LM up to the altitudes where Aura-MLS is available. Exceptions to this are the Southern Hemisphere polar region and the LS, where ACE-FTS (Aura-MLS) shows a slightly larger negative (positive) departure from the MIM of ±20%. The larger differences in these regions may be due to sampling bias (as suggested by smaller monthly values seen in **Figure A4.18.2** in *Appendix A4*). HALOE exhibits negative departures from the MIM with values below -10%. Validation studies of earlier versions of HALOE (v17) have already pointed toward a low bias in HALOE when compared to other correlative measurements such as balloon and ATMOS measurements, with differences between 10% and 20% [*Russell et al.*, 1996b]. The somewhat low HALOE (v19) HCl values have also been discussed in a number of previous studies, including those by *Froidevaux et al.* [2006; 2008b] for Aura-MLS versus HALOE, by *McHugh et al.* [2005] and *Mahieu et al.* [2008] for ACE-FTS versus HALOE, and by *Lary et al.* [2007], who used various space-based measurements in a neural network analysis. Both SMILES data products used in the SPARC Data Initiative show similar values as HALOE, with negative deviations from the MIM through most of the USLM. However, a newer version of SMILES with a better altitude registration tends to increase HCl, so that the differences versus Aura MLS are much smaller. Therefore, HALOE can be considered typically lower than all other three instrument retrievals. In the UTLS, SMILES shows structures in the deviations from the MIM that resemble the impact of different QBO phases on trace gas distributions [*cf., Randel et al.*, 1999]. Indeed a strong easterly phase was observed during 2009/2010. An evaluation of monthly cross sections

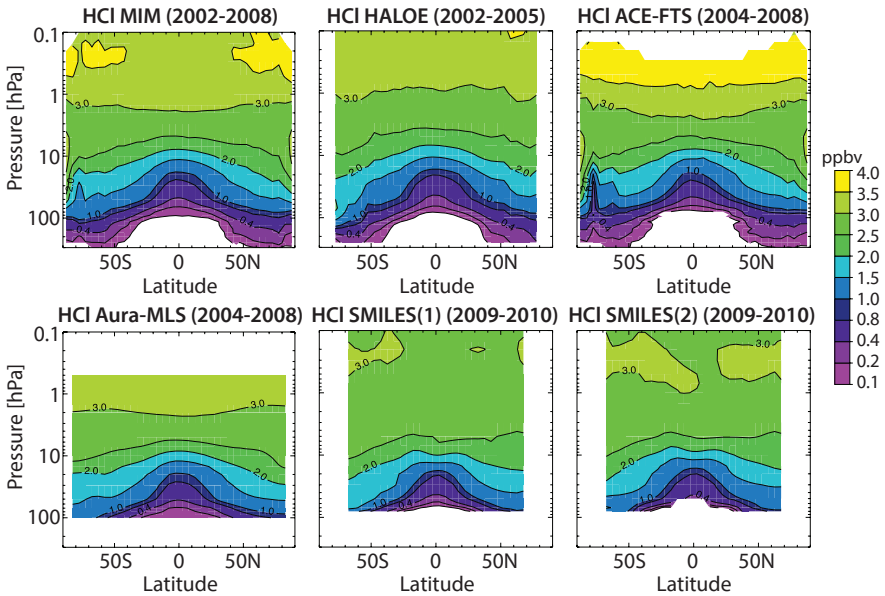


Figure 4.18.1a: Cross sections of annual zonal mean HCl for 2002-2008. *Annual zonal mean HCl cross sections are shown for 2002-2008 for the MIM, HALOE, ACE-FTS, Aura-MLS and SMILES. Note that the instruments provide data for different time periods as indicated in the panel titles. SMILES data have not been included in the MIM.*

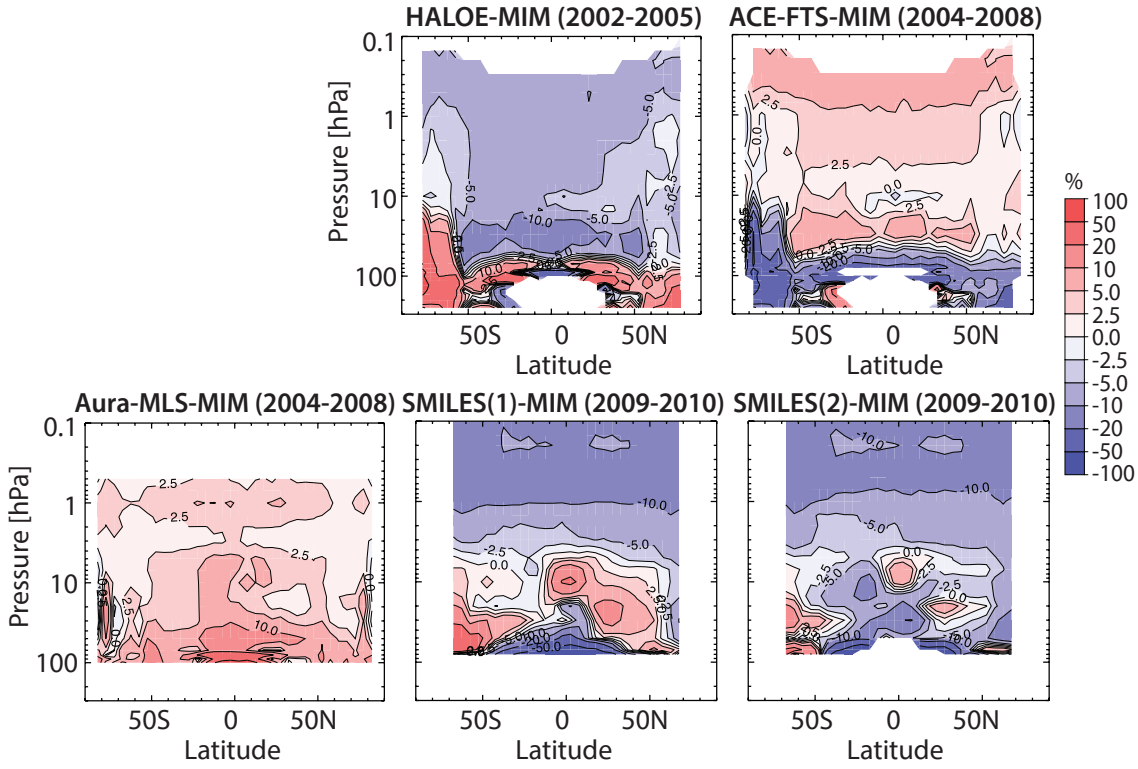
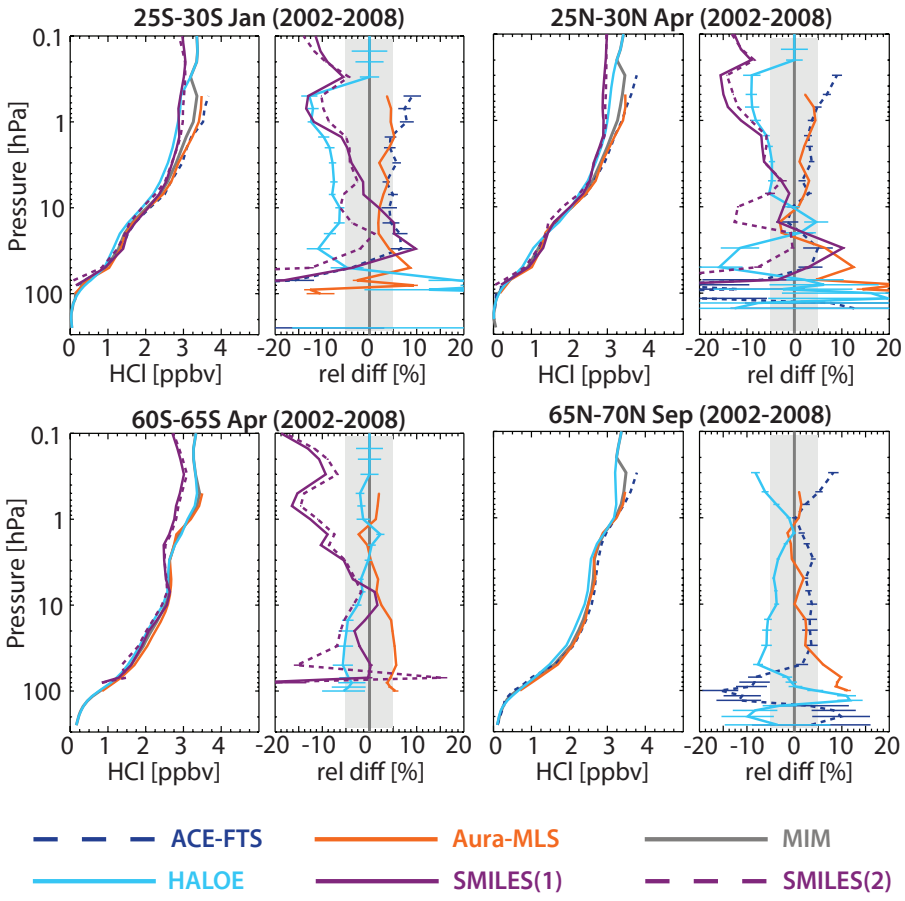


Figure 4.18.1b: Cross sections of annual zonal mean HCl differences for 2002-2008. *Annual zonal mean HCl differences between the individual instruments (HALOE, ACE-FTS, Aura-MLS, and SMILES) and the MIM are shown. SMILES data have not been included in the MIM.*

in the years 2009 and 2010 shows that these QBO-like structures disappear in the SMILES deviations from the MIM (see **Figure A4.18.2** in *Appendix A4*).

The altitude profiles shown in **Figure 4.18.2** support the above findings and reveal further structural details on shorter (monthly) time scales. ACE-FTS and Aura-MLS

Figure 4.18.2: Vertical profiles of zonal mean HCl for 2002-2008. *Zonal mean HCl profiles for January 25°S-30°S and April 25°N-30°N (upper panels), and for April 60°S-65°S and September 65°N-70°N (lower 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 (expressed here as ±1 SEM). SMILES is not included in the MIM since it measures during a limited time period in 2009-2010 only.*



agree well throughout the stratosphere and the LM for all months for which ACE-FTS is available, while HALOE is biased low by about 10%-15% except for the SH high latitudes. In the LS, however, ACE-FTS and Aura-MLS drift apart, with ACE-FTS (Aura-MLS) showing negative (positive) departures from the MIM as already seen in the annual mean cross sections. HALOE also shows a positive departure from the MIM in the UTLS. Both SMILES products tend to agree with ACE-FTS and Aura-MLS in the LS, but show negative deviations from the MIM above, with values more similar to those of HALOE. Monthly deviations can exceed annual mean deviations locally.

The meridional profiles in **Figure 4.18.3** illustrate how the relative differences from the MIM are decreasing with increasing altitude. At 100 hPa, HALOE and Aura-MLS show relative differences from the MIM in the extra-tropics and tropics of up to ±25% and ±80%, respectively. All instruments show relative differences around ±15% at 10 hPa, and around ±8% at 1 hPa. The relative differences increase again above 1 hPa. Note that the improved SMILES retrieval of HCl is expected to bring SMILES closer to Aura-MLS and ACE-FTS.

4.18.3 HCl evaluations: Latitude-time evolution

Figure 4.18.4 shows the latitude-time evolution of HCl at 50 hPa. Only Aura-MLS has frequent enough spatial and temporal sampling to provide year-around coverage at all latitudes. Nevertheless, the comparison with the interpolated fields from HALOE and ACE-FTS show consistent features. Generally, there is not much intra-annual variability seen in the field (except in the polar regions). As expected from the annual zonal mean cross sections, the isopleths are shaped by the Brewer-Dobson circulation, with lower values in the tropics than in the extra-tropics. In the tropics, HALOE shows somewhat lower values than Aura-MLS. However, a strong minimum in HCl is found in the polar vortex region during Southern Hemisphere winter, a feature consistently reproduced by Aura-MLS and ACE-FTS, and to some extent by HALOE. The physical explanation for the minimum in HCl is the build-up of polar stratospheric clouds (PSCs) during winter, on which HCl (together with HNO_3) reacts to release active chlorine [*e.g.*, *Santee et al.*, 2008]. The activation of chlorine atoms then leads to the severe catalytic depletion of ozone (and the Antarctic ozone hole) during polar spring.

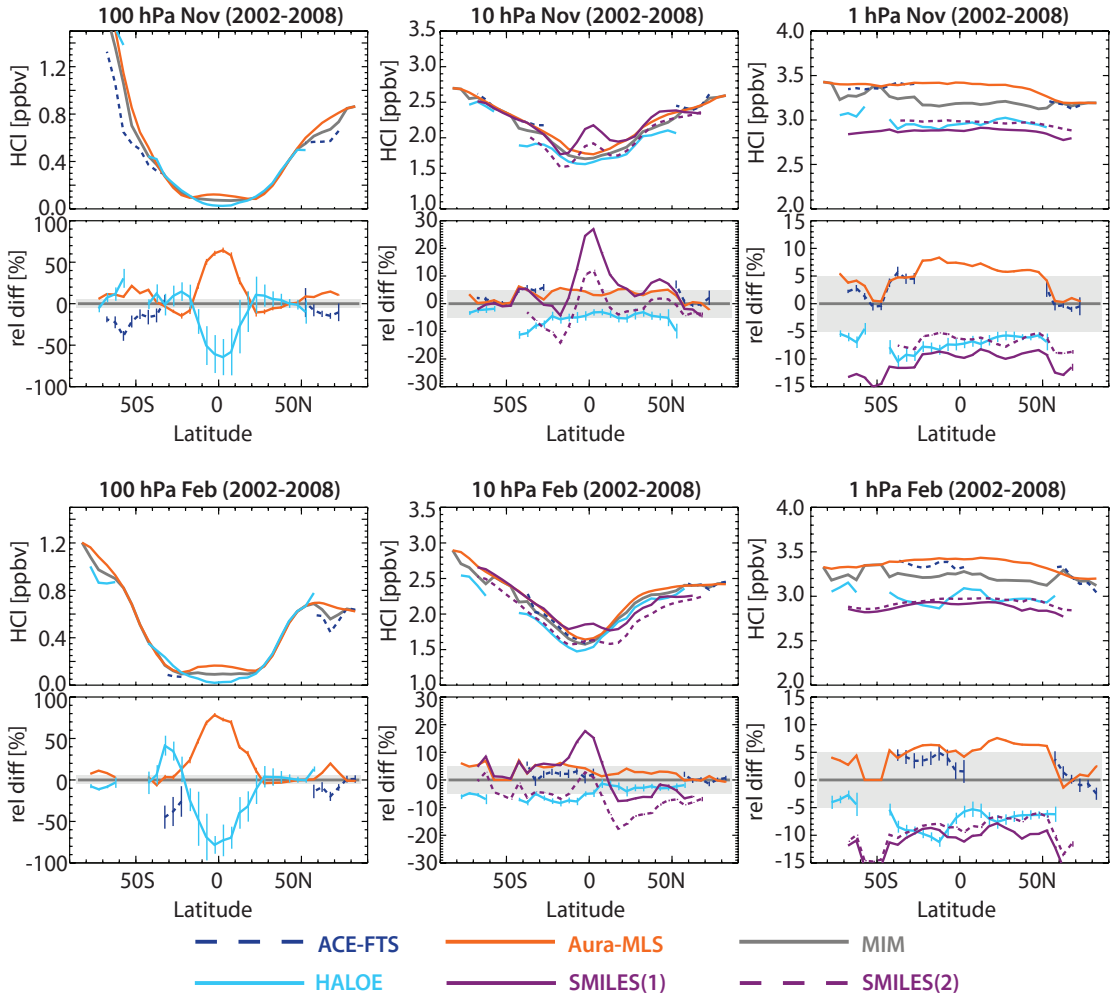


Figure 4.18.3: Meridional profiles of zonal mean HCl for 2002-2008. *Meridional zonal mean HCl profiles at 100, 10, and 1 hPa for November (upper row) and February (lower row). Differences between the individual instruments (HALOE, ACE-FTS, Aura-MLS, and SMILES) and the MIM are shown in the lower panels. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences (expressed here as ±1 SEM). Note, SMILES is not included in the MIM.*

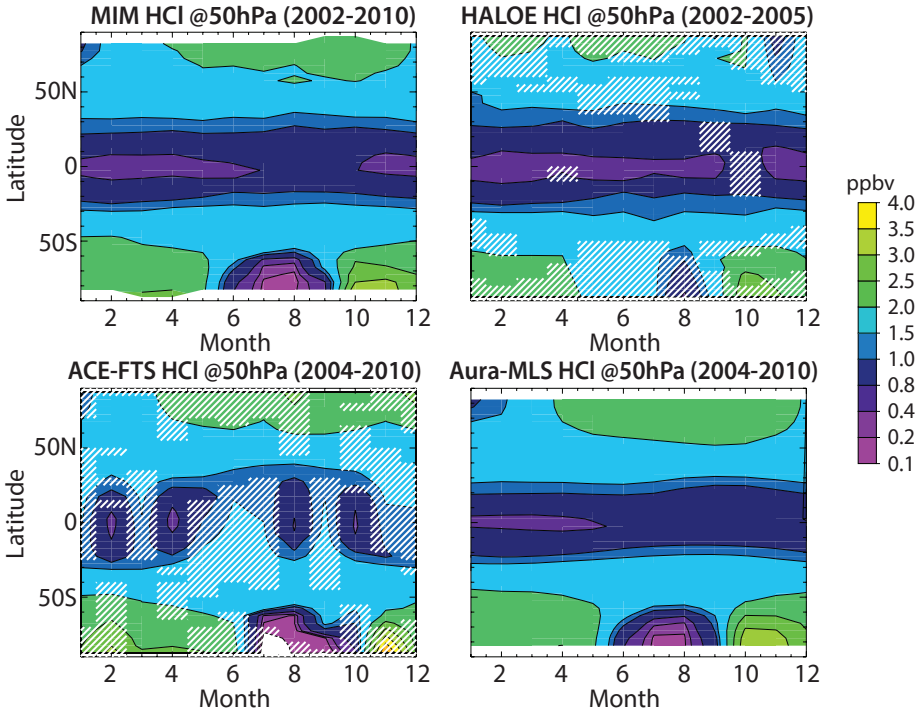


Figure 4.18.4: Latitude–time evolution of HCl at 50hPa. *The latitude-time evolution of HCl at 50 hPa is shown for the MIM (2002-2010 average) in the leftmost upper panel and the instruments HALOE, ACE-FTS, and Aura-MLS. HALOE and ACE-FTS show interpolated fields, with hatched regions indicating where no measurements are available.*

4.18.4 HCl evaluations: Interannual variability

In addition to climatological differences, the instrument performances in reproducing interannual variability is tested. While there is not enough overlap to thoroughly evaluate HALOE versus the later instruments and SMILES does not provide a long enough time series to be included in the comparison, one can deduce some information from the comparison on the behaviour of Aura-MLS and ACE-FTS. **Figure 4.18.5** shows anomalies for the different instruments at different pressure levels in the polar region of the Northern Hemisphere. Aura-MLS and ACE-FTS agree well, although ACE-FTS exhibits less frequent sampling and incomplete

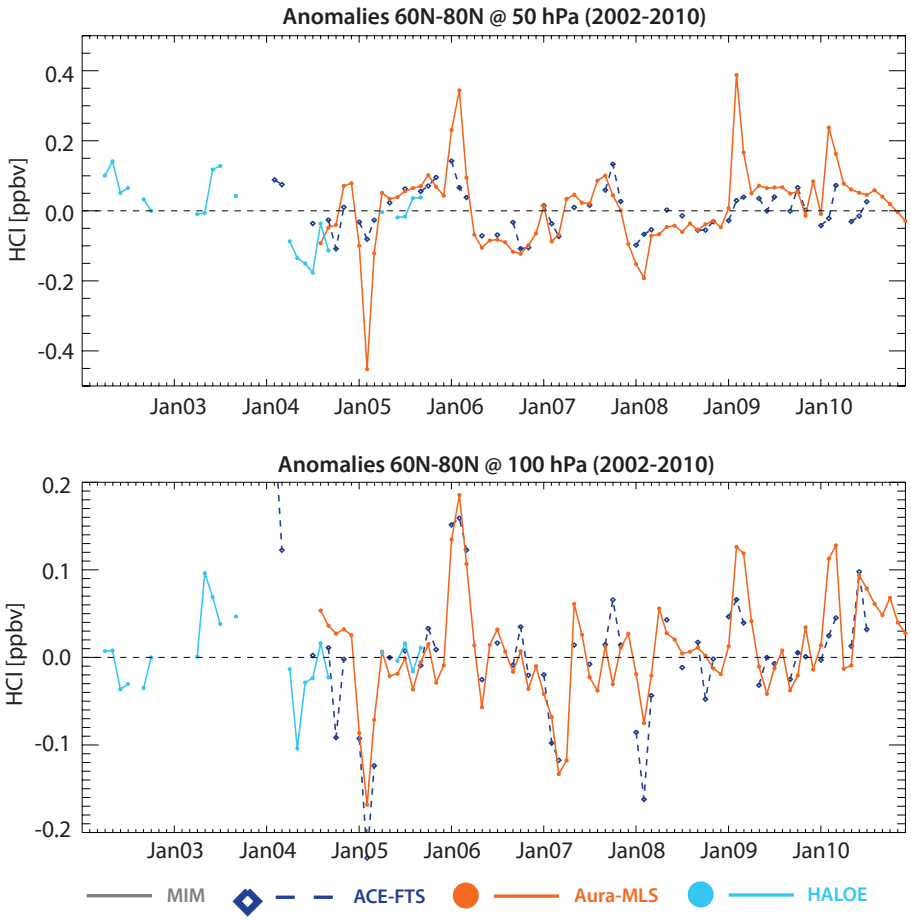


Figure 4.18.5: Time series of deseasonalised HCl anomalies in the Northern Hemisphere extra-tropics. *Deseasonalised HCl anomalies are shown for 60°N-80°N at 50 hPa (upper panel) and 100 hPa (lower panel).*

annual information. January maxima and minima in the data record are somewhat better resolved by ACE-FTS at 100 than at 50 hPa. The observed maxima stem from the descent of HCl-rich air during warmer winters, while the minima indicate strong processing on polar stratospheric clouds. The temperatures during the 2004/2005 Arctic winter for example were at a record low, with much higher PSC abundances than in other years. In the NH polar latitudes, the anomalies show the strongest disagreement during the winter months. This could be related to a different sampling of the polar vortex by the two instruments. Note that ACE-FTS is only of limited use when looking at interannual variability in the tropics, where the instrument's temporal coverage is more limited.

4.18.5 Summary and conclusions: HCl

HCl climatologies from four limb-sounders (HALOE, ACE-FTS, Aura-MLS, and SMILES) have been compared within the SPARC Data Initiative. SMILES measures two isotopologues ($H^{35}Cl$ and $H^{37}Cl$), which have been evaluated separately (using the SMILES(1) and SMILES(2) notation). Overall findings on the systematic uncertainty in our knowledge of the HCl 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 uncertainty in our knowledge of the atmospheric HCl annual mean state as derived from the four satellite instruments and measured by the multi-instrument spread is smallest in the MS and US, and smaller in the polar regions (±4%) than in the tropics (±8%) (see **Figure 4.18.6**). Good knowledge is obtained in the LM and tropical LS, where the uncertainty is about ±10-15%. The uncertainty is largest

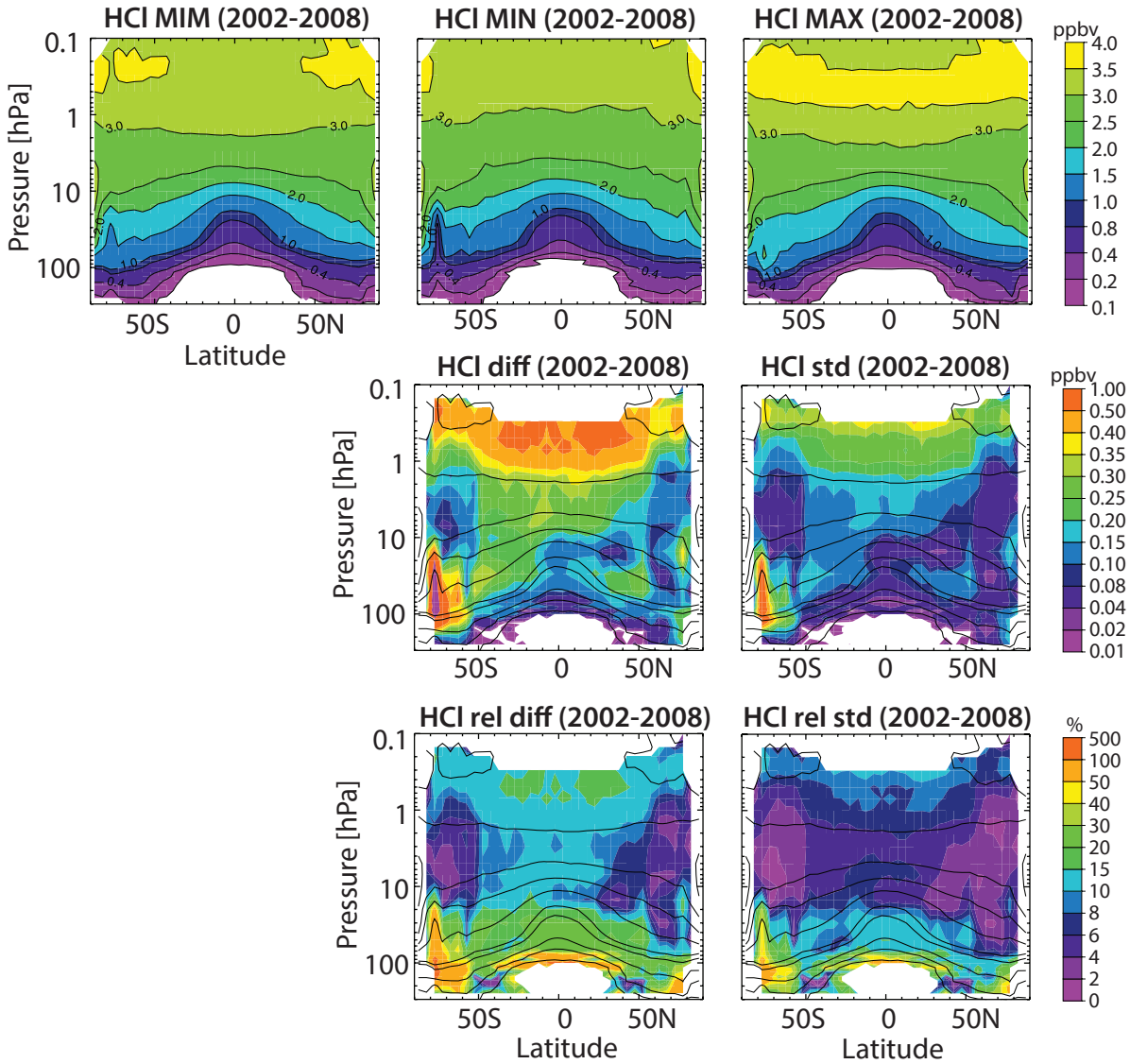


Figure 4.18.6: Summary of HCl annual zonal mean state for 2002-2008. *Annual zonal mean cross sections for 2002-2008 of the MIM, minimum (MIN), and maximum (MAX) HCl values are shown in upper row. The absolute and relative differences over all instruments (MAX-MIN) and the absolute and relative standard deviations over all instruments are presented in the middle and lower row, respectively. Black contours in lower panels repeat the MIM distribution. Instruments considered are ACE-FTS, Aura-MLS, and HALOE. Note SMILES(1) and SMILES(2) are not included in the MIM due to their limited sampling.*

in the Southern Hemisphere polar vortex region and the UTLS (reaching values higher than ±50%). The large uncertainty in the UTLS may be explained by the relatively small HCl mixing ratios that the instruments have to be able to detect.

Performance by region

In the USLM (0.1-5 hPa), median values are well defined as indicated by small MADs. The derived inter-instrument differences, which lie mostly within ±10%, are therefore well defined as well. SMILES shows somewhat larger deviations from the MIM, especially in the tropics.

In the MS (5-30 hPa), excellent agreement is found between the instruments in the extra-tropics. In the tropics, the uncertainty increases to ±8%. However, large MADs indicate that these inter-instrument differences are not well defined and local difference can be much larger.

In the LS (30-100 hPa), good agreement within ±10% is found in the extra-tropics, but considerable disagreement with inter-instrument differences of up to ±40% is found in the tropics.

In the UT (100-300 hPa), only HALOE and ACE-FTS provide measurements. HALOE (ACE-FTS) shows mostly positive (negative) differences in this region reaching mean values of up to ±20%. Note that the median values are not well defined as seen in relatively large MADs. Also, Aura-MLS HCl data at these lower altitudes may be useful at high latitudes, but they have a high bias at low latitudes [see *Froidevaux et al.*, 2008b] and hence were not included in the SPARC Data Initiative climatologies. This issue could also explain the tendency for Aura-MLS HCl values at 100 hPa to be on the high side in the tropics.

Instrument-specific conclusions

HALOE shows a negative bias throughout the stratosphere and the LM, with relative differences with respect to the MIM of between -5% and -10%. The negative bias in HALOE is known from previous studies and also agrees in magnitude [*Russell et al.*, 1996b; *Lary et al.*, 2007]. For pressures larger than 100 hPa, HALOE shows a positive deviation from the MIM (which is mostly determined by the comparison to ACE-FTS).

Below 30 hPa, **ACE-FTS** shows much lower values than the other instruments in the extra-tropical UTLS. Above 30 hPa, however, it shows excellent agreement with the Aura-MLS dataset.

As noted previously, **Aura-MLS** and ACE-FTS show excellent agreement with each other, except in the LS where Aura-MLS shows much higher values than ACE-FTS.

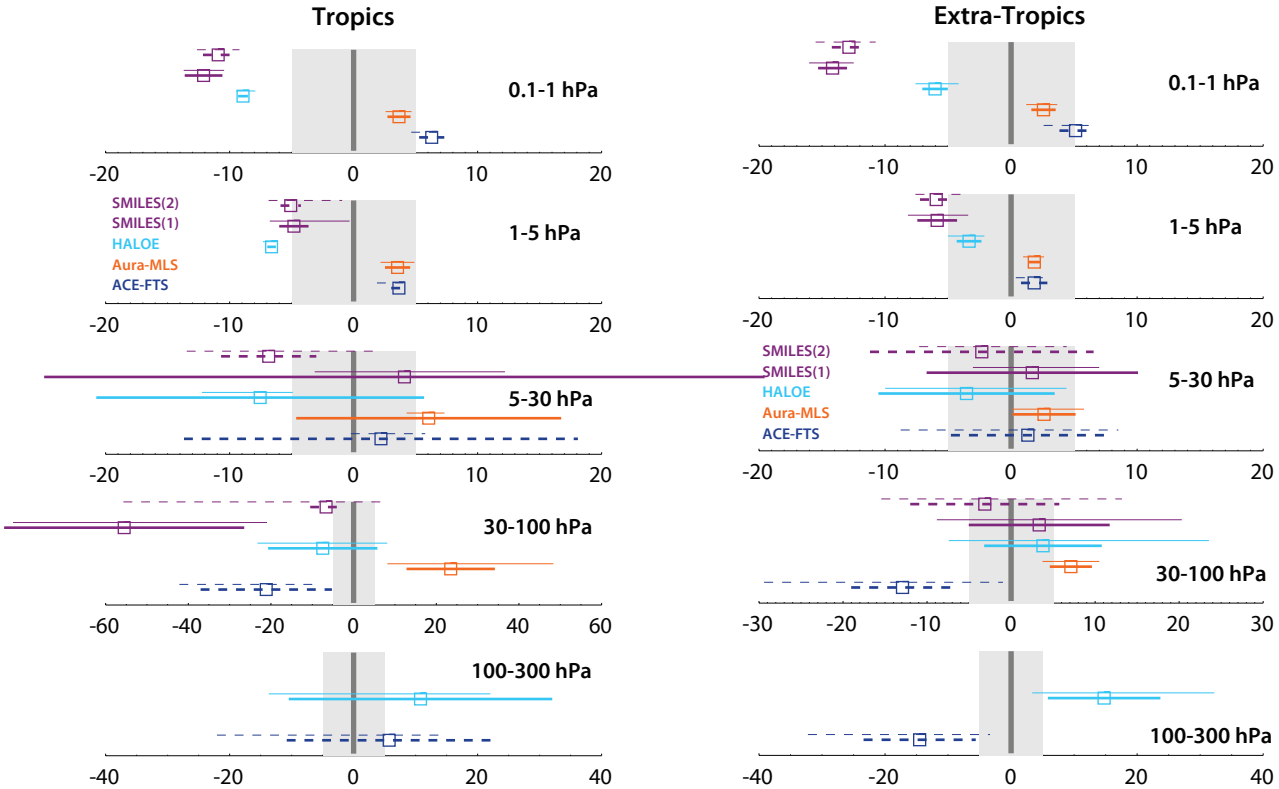


Figure 4.18.7: Inter-instrument differences in HCl calculated for the tropics (left) (20°S–20°N) and (right) extra-tropics (40°S–80°S and 40°N–80°N) and for five different altitude regions from the UT up to the LM. *Shown are the median (squares), median absolute deviations (MAD, thick lines), and the 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 2002-2008. Note SMILES(1) and SMILES(2) are not included in the MIM due to their limited sampling.*

SMILES(1) and **SMILES(2)** versions available to the SPARC Data Initiative are very similar in their overall structure, except in the tropical LS. Here, SMILES(1) shows a large negative bias with deviations from the MIM of around 50%. In the USLM, the two datasets generally are showing even larger negative deviations from the MIM than HALOE, which is known to exhibit a low bias. Note that an improved version of SMILES HCl has become available in the meantime, which largely removes the low bias and makes the retrieved values closer to those measured by Aura-MLS and ACE-FTS.

4.18.6 Recommendations: HCl

The long-term monitoring capability for stratospheric HCl hinges on two instruments (ACE-FTS and Aura-MLS), which are both past their expected lifetimes. New height-resolved measurements are needed to maintain this capability in order to be able to fulfill obligations to the Montreal Protocol.

4.19 Chlorine monoxide – ClO

Chlorine monoxide (ClO) is one of the most important reactive chlorine species involved in chlorine-catalyzed destruction of stratospheric ozone. The primary sources of ClO are CFCs (see *Sections 4.5* and *4.6*), but also hydrochlorofluorocarbons (HCFCs), methyl chloride (CH_3Cl), and carbon tetrachloride (CCl_4). These source gases reach the stratosphere, and are transported by the stratospheric circulation to higher altitudes, where they are photolyzed by UV radiation and release atomic chlorine (Cl). Cl then reacts with ozone to form ClO. In the upper stratosphere, ClO represents a substantial fraction of total available stratospheric chlorine. However, in the lower and middle stratosphere, reactive chlorine (ClO and Cl) is mostly tied up in the reservoir species HCl and $ClONO_2$ that do not directly destroy ozone. It is only during winter/spring in polar regions that HCl and $ClONO_2$ are converted back into ClO via heterogeneous chemical reactions on polar stratospheric clouds; these processes create enhanced levels of ClO that are a precursor to ozone destruction [*Solomon et al.*, 1986; *Molina and Molina*, 1987].

ClO exhibits a relatively strong diurnal cycle in the LS and MS, however, a less pronounced diurnal cycle in the US. **Figure 4.19.1** shows two examples of the diurnal ClO cycle as a function of local solar time or solar zenith angle for three different pressure levels as derived from a chemical box model [*McLinden et al.*, 2010]. A comparison of satellite-based ClO measurements corresponding to different local solar times would, ideally, account for this dependence on SZA.

4.19.1 Availability of ClO measurements

The assessment of ClO is based on the climatologies from SMR, Aura-MLS, MIPAS, and SMILES observations available to the SPARC Data Initiative. The climatologies start in 2001, with SMR currently providing the longest record. Both night-time and daytime ClO climatologies are available for the instruments used here, but only daytime measurements are presented and evaluated in this chapter. Measurements used from MIPAS correspond roughly to 10am, from Aura-MLS to about 1:30pm, from SMR to 6:30am (SMR(1))and also 6:30am measurements scaled to 1:30pm (SMR(2)), while SMILES measurements are taken throughout the day.

Older measurements are available from UARS-MLS, but are not included in the comparisons shown here, partly because significant trends are expected in atmospheric ClO concentrations between the 1990s and the 2000s as a result of reductions of chlorine source gas emissions following the Montreal Protocol regulations. Moreover, a careful selection of UARS-MLS data is needed as local time varies during each month complicating the creation of regular monthly datasets at all latitudes. This is due to periodic satellite yaws, which change the global viewing geometry of UARS-MLS roughly every 36 days. *Nedoluha et al.* [2011]

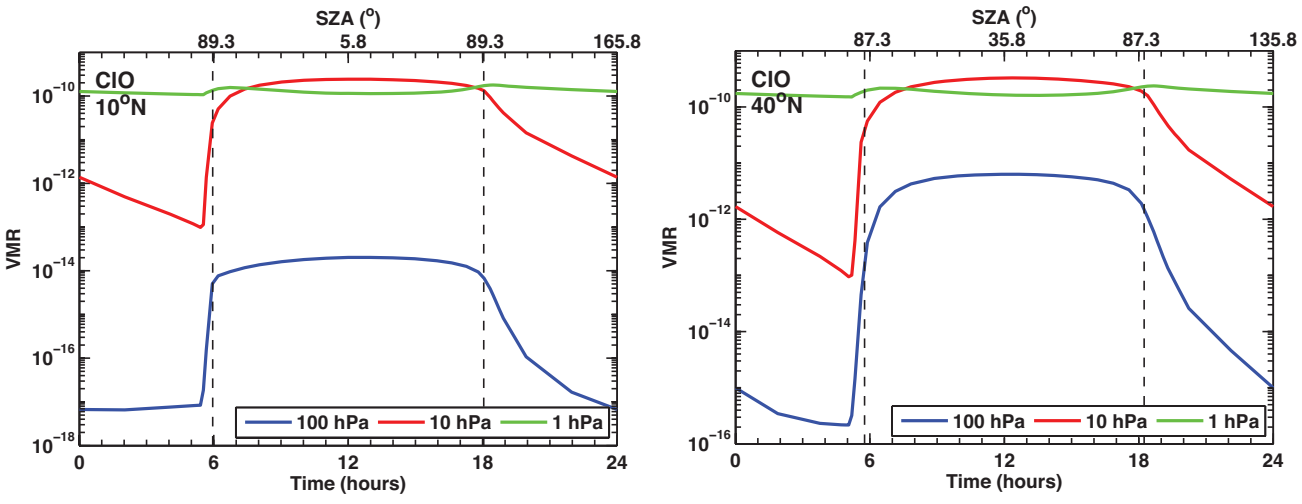


Figure 4.19.1: Diurnal cycle in ClO. *ClO diurnal variations are shown as a function of LST or SZA at 10°N (left panel) and 40°N (right panel) for 1, 10 and 100 hPa. The diurnal cycle is derived using a chemical box model [McLinden et al., 2010].*

Table 4.19.1: Available ClO 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.

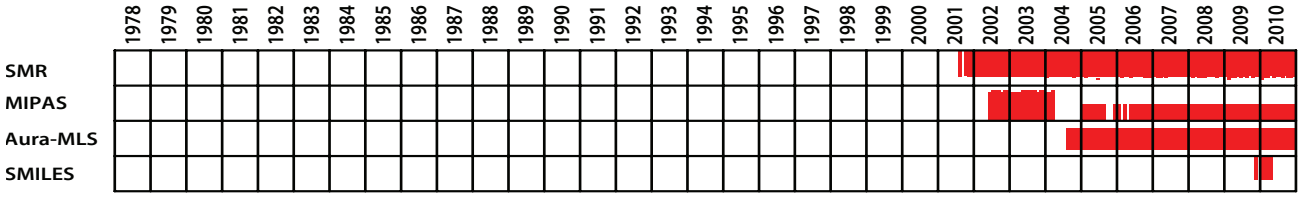


Table 4.19.2: Data version, time period, vertical range, vertical resolution, references and other comments for ClO measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
Aura-MLS V3.3	Aug 2004 –	130 – 1 hPa	3 – 4.5 km	*Santee et al.,* 2008 *Livesey et al.,* 2013	about 1:30am/pm (but changes towards high latitudes)
SMR V2.1	Jul 2001 –	20 – 50 km (variable range depending on minimum mixing ratios)	2.5 – 3 km	*Urban et al.,* 2005b *Urban et al.,* 2006 *Khosravi et al.,* 2013	6:30am/pm; am/pm climatologies scaled to 1:30am/pm are also available
MIPAS					10am/pm; close to detection limit; measurement mode switched in 2005 from high spectral to high vertical resolution; Note that for single measurement profiles, use of a restricted altitude range of 10 – 30 km is recommended for both products.
MIPAS(1) V11	Mar 02 – Apr 04	10 – 45 km	4 km (below 25 km) > 8 km (above 25 km)	*Glatthor et al.,* 2004	
MIPAS(2) V220	Jan 05 – Apr 12	10 – 30 km	3.3 – 7.5 km	*von Clarmann et al.,* 2009a	
SMILES V2.1	Nov 2009 – Apr 2010	100 – 0.01 hPa	3 – 5 km (stratosphere) 5 – 8 km (mesosphere)	*Sato et al.,* 2012 *Sagawa et al.,* 2013	C-band measurements

compared daytime portions of the UARS-MLS and Aura-MLS datasets versus time series of (daytime) ground-based profiles in the upper stratosphere and found no significant adjustment need between these two satellite datasets in this region, given their inferred average agreement of 1% ± 4%.

Tables 4.19.1 and **4.19.2** compile information on the availability of ClO measurements, including data version, time period, height range, vertical resolution, and references relevant for the data product used in this report.

4.19.2 ClO evaluations: Zonal mean cross sections

Monthly zonal mean cross sections will be analysed for daytime climatologies in order to investigate mean biases between the various datasets. Daytime climatologies are chosen, since the ClO diurnal cycle is smaller during the day than the night.

SMR(1), SMR(2), MIPAS(2), Aura-MLS, and SMILES (2009-2010), and MIPAS(1) (2002-2004)

Figure 4.19.2a shows the monthly zonal mean ClO climatologies for the reference period 2009-2010 chosen for all available measurements. MIPAS(1) measured during an earlier time period (2002-2004), but is here also compared to the 2009-2010 MIM reference. Note that in the following, SMR(1) is not included in the calculation of the MIM, but that the SMR(2) scaled product is considered in the limited region available.

The ClO distribution indicates a maximum in the upper stratosphere, representing a major contribution to total chlorine apart from HCl. The maximum values around 3-5 hPa reach around 800 pptv. ClO decreases towards the LM and also the LS to values smaller than 5 pptv, which is close to most instruments' detection limit. A localised maximum in the ClO distribution is found in the LS polar vortex region during winter/spring, indicating the enhancement of ClO through heterogeneous reactions on polar stratospheric clouds.

The different instruments' monthly climatologies look very similar, although MIPAS(1) shows somewhat larger values than the other instruments, especially across the US in the tropics. The higher values may be at least partially attributable to a negative trend in stratospheric ClO over the 2000s, the consequence of the successful implementation of the Montreal Protocol regulations on chlorine source gases. At the other end, SMR(1) shows lowest ClO mixing ratios, consistent with its local measurement time being slightly

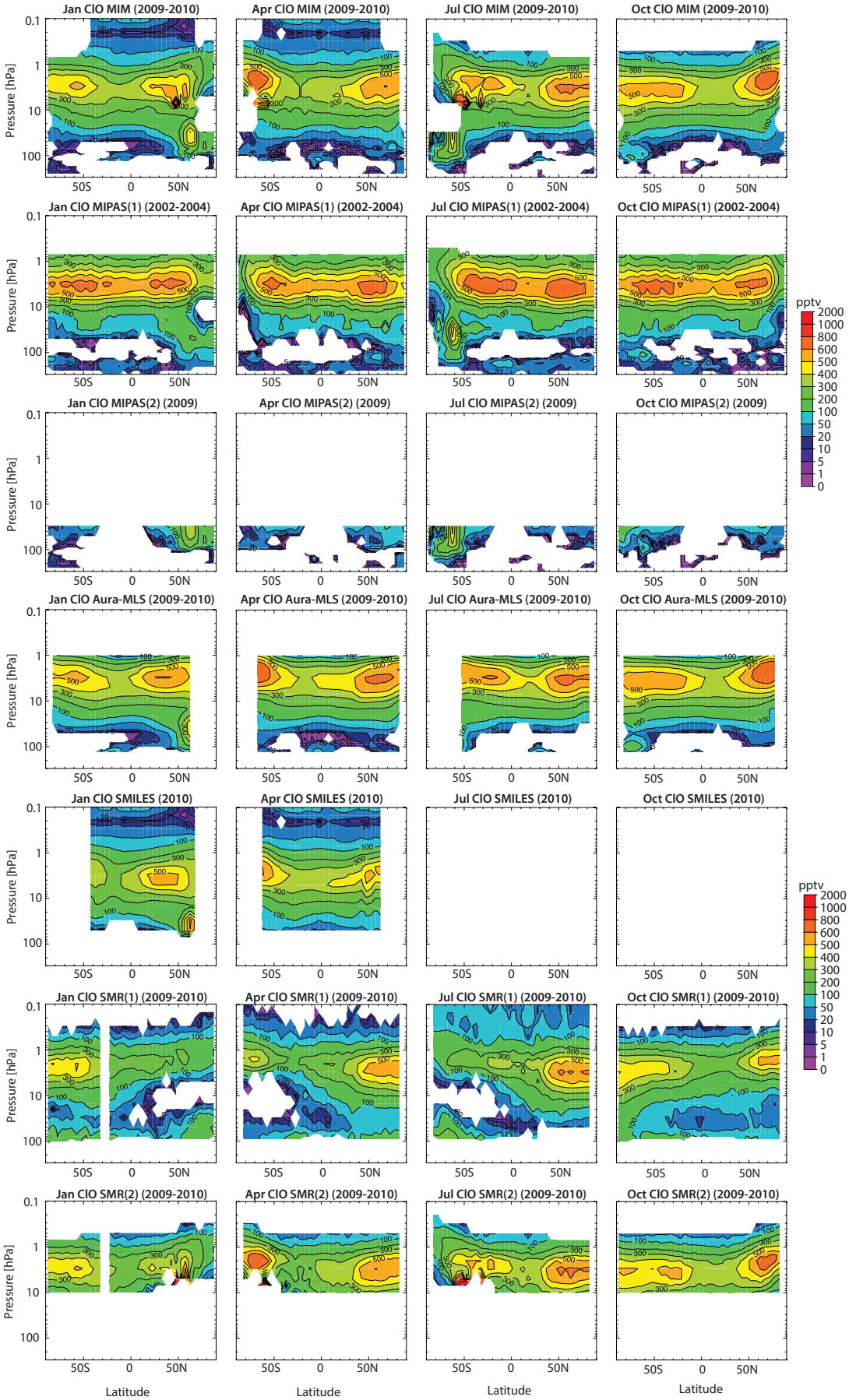


Figure 4.19.1a: Cross sections of monthly zonal mean ClO for 2009-2010. *ClO cross sections for 2009-2010 (or according to availability as indicated) are shown for the MIM, MIPAS(1) (2002-2004), MIPAS(2) (2009), Aura-MLS, SMILES (2010), SMR-am data unscaled (SMR(1)) and scaled to 1:30pm (SMR(2)). Note, MIPAS(1) and SMR(1) are not included in the MIM.*

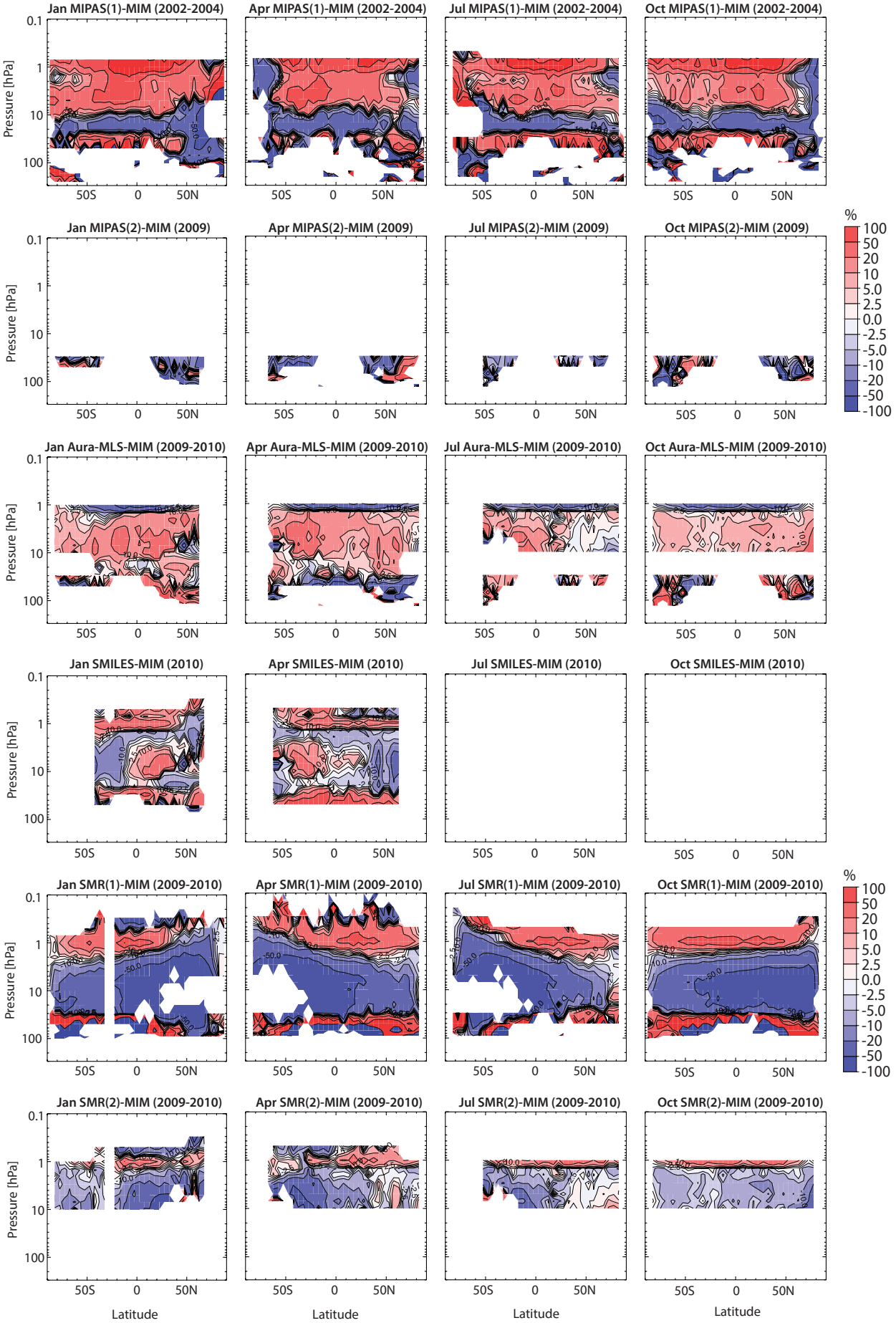


Figure 4.19.1b: Cross sections of monthly zonal mean ClO differences for 2009-2010. *ClO relative differences with respect to the MIM are shown for the individual instruments as in Figure 4.19.2a.*

before the maximum daily values are reached (which then are more or less constant through the day).

The corresponding relative differences of the instruments with respect to the MIM are displayed in **Figure 4.19.2b** The figure quantifies the qualitative discussion on instrument differences above. Note that the white bar in the Aura-MLS difference climatology between 30 and 10 hPa during July and October is due to the fact that the MIM at these altitudes consists of Aura-MLS data only (MIPAS(1) and SMR(1) are not considered in the MIM). Aura-MLS, SMILES, and the scaled SMR(2) climatologies agree reasonably well, mostly within ±10-20%, confirming earlier validation studies [*Sagawa et al.*, 2013; *Kreyling et al.*, 2013]. Aura-MLS and SMR(2) agree even better than ±10% (as seen for July and October). This result is in good agreement with an earlier comparison by *Livesey et al.* [2013] who used tight coincidence criteria in terms of SZA for Aura-MLS and SMR profiles and who showed differences within 10-15% (except at 68 to 100 hPa where both datasets are subject to larger biases, and for which no comparison could be made here). The good quality of Aura-MLS ClO is supported by comparisons versus ground-based observations. Antarctic spring (year 2005) enhanced Aura-MLS ClO profiles, appropriately convolved to match the coarser vertical resolution of ground-based microwave ClO profiles, have been shown by *Connor et al.* [2007] to agree, on average, with the ground-based profiles (coincident in space and local time) near the lower stratospheric peak (within 11 ± 8%, and within the combined 2σ uncertainties); Aura-MLS values are slightly lower than the ground-based values.

The unscaled SMR climatology, SMR(1), shows large negative relative differences from the MIM throughout most of the LS and MS (reaching more than -50%), which is expected from its measurement time at 6:30am shortly before values reach their daily maximum. Note that around 1 hPa, the SMR(1) differences are positive, which again can be expected from the diurnal cycle, which is different than at lower altitudes and shows localised maxima just after 6am and 6pm (see **Figure 4.19.1**).

MIPAS(1) shows large positive differences reaching more than +20% in the US, with a band of large negative differences reaching -20% just below in the MS. Note that a comparison between MIPAS(1) and SMR(2) over the same time period (2002-2004; not shown) does not yield better agreement between the instruments, despite avoiding a potential temporal sampling issue.

4.19.3 ClO evaluations: Vertical and meridional profiles

The vertical and meridional profiles shown in **Figures 4.19.3** and **4.19.4** emphasise details in the structure of the differences in the monthly zonal mean cross sections.

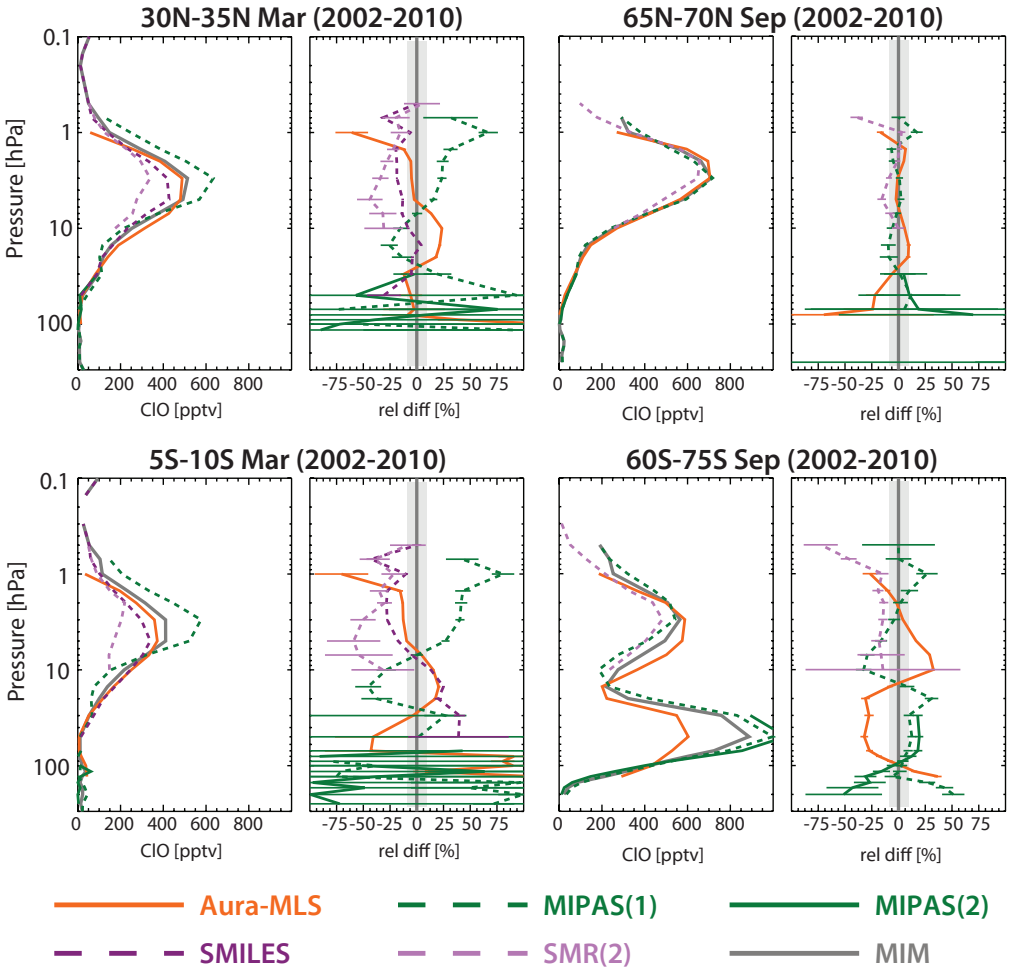


Figure 4.19.2: Vertical profiles of monthly zonal mean ClO for 2002-2010. *Vertical ClO profiles for March 30°N-35°N and 5°S-10°S (left panels), and for September 65°N-70°N and 60°S-75°S (right panels) are shown together with their differences from the MIM for both. Note, while MIPAS(1) and MIPAS(2) are included in the MIM, SMR(2) is not. The grey shading indicates where the relative differences are smaller than ±10%. Bars indicate the uncertainties in the relative differences based on the SEM of each instrument's climatology.*

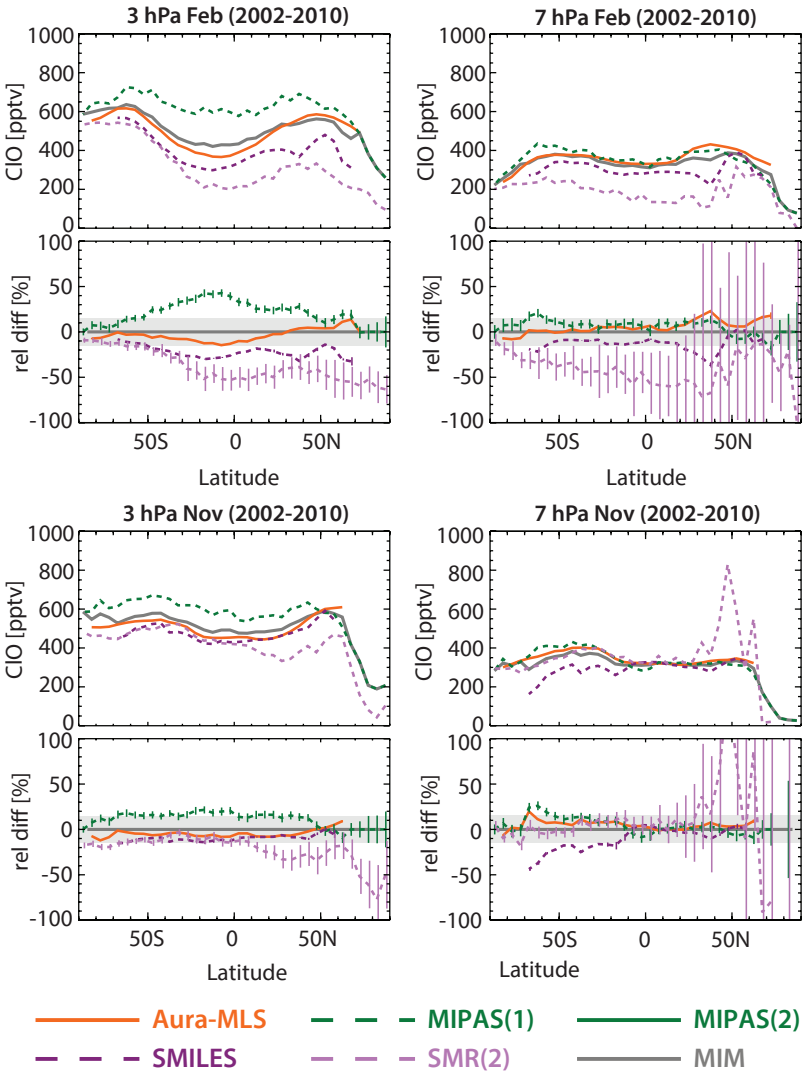


Figure 4.19.3: Meridional profiles of monthly zonal mean ClO for 2002-2010. *ClO profiles are shown at 3 and 7 hPa for February (upper row) and November (lower row) averaged over 2002-2010. Relative differences between the individual instruments (SMR(2), MIPAS(1), SMILES, and MLS) and the MIM profiles are shown in the lower panels. Note, SMR(2) is not included in the MIM. The grey shading indicates where the relative differences are smaller than ±15%. Bars indicate the uncertainties in the relative differences based on the SEM of each instrument's climatology.*

Good agreement and smallest relative differences of ±10% are found between the different instruments during September in the northern high-latitudes, when stratospheric dynamics is rather quiet and as a consequence geophysical variability is small. Small variability is also expected in the tropics throughout the year, however, the vertical profiles at 5°S-10°S show fairly large relative differences between the instruments. An exception is Aura-MLS and SMILES, which show very good agreement over most of the MS and US and similar maximum values around 3 hPa. The high bias in MIPAS(1) in the upper stratosphere may be partially explained by the expected trend in the chlorine loading between the early and late 2000s.

Only reasonably good agreement to considerable disagreement is found for March at northern mid-latitudes and September at southern high-latitudes. These are dynamically and chemically active seasons and sampling issues may play a stronger role in determining the relative differences between the instruments, especially for such a short-lived species as ClO that shows in addition strong interannual variability. In the face of the high natural variability it is hence difficult to draw firm conclusions about the instruments' retrievals during these dynamically active seasons. Note also that for Aura-MLS daytime data at higher latitudes, the local time is shifted away from 1:30pm to around 3:30pm, so that the diurnal change effect will tend to introduce larger biases than anticipated when comparing to the SMR product scaled to 1:30pm.

Meridional monthly zonal mean ClO profiles for 2002-2010 are shown in **Figure 4.19.4** for two levels in the MS (7 hPa) and US (3 hPa). At these levels, MIPAS(1), SMILES, and Aura-MLS show generally good to reasonable agreement (between ±10 and ±20%). The scaled SMR(2) product is, despite the diurnal adjustment, still on the low side of the other instruments' climatologies and also shows much larger SEM values, *i.e.,* its mean is not as clearly defined. MIPAS(1) is generally on the high side.

4.19.4 Summary and conclusions: ClO

ClO climatologies are available from four limb satellite instruments: MIPAS, SMR, Aura-MLS, and SMILES. While SMILES observes the full diurnal cycle, MIPAS measurements are made at about 10am/pm, Aura-MLS at about 1:30am/pm, and SMR at 6:30am/pm respectively. The observations allow therefore for the compilation of both daytime and night-time climatologies. Only daytime climatologies are evaluated here, since the diurnal cycle is deemed to cause

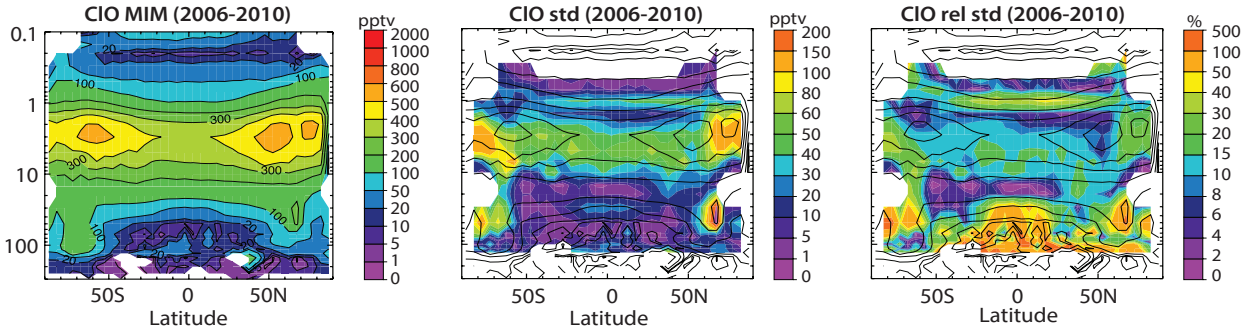


Figure 4.19.4: Summary of ClO annual zonal mean state for 2006-2010. *Annual zonal mean cross sections for 2006-2010 of the MIM, the absolute, and the relative standard deviations over all instruments are presented from left to right. Black contours in right-hand panels repeat the MIM distribution. Instruments considered are MIPAS(2), Aura-MLS, SMILES, and the scaled SMR product (SMR(2)). MIPAS(1) has not been included since its measurement period does not overlap with the rest of the instruments.*

smaller variations during the day than during night, and since lower stratospheric night-time values are much smaller (and harder to measure) than daytime values in this region.

Both the unscaled and scaled SMR climatologies are on the low side of the other instruments. It has been found that SMR am measurements scaled to 1:30pm improve the comparison with the other instruments, but generally still underestimate ClO mixing ratios through most of the stratosphere and LM. Nevertheless, the differences between the scaled SMR and Aura-MLS climatologies are generally well below 10-20% (or ±5-10% with respect to the MIM) in our zonal monthly mean comparisons between 10 and 1 hPa (the range the scaled SMR product is available for), indicating very good to good agreement. This result agrees with averaged coincident ClO profile differences between SMR and Aura-MLS (see *Livesey et al.*, 2013).

SMILES introduces somewhat larger differences compared to these two instruments, with values around ±10% with respect to the MIM depending on the region, with the results being largely consistent with *Sagawa et al.* [2013]. The high-spectral resolution MIPAS climatology, MIPAS(1), exhibits generally somewhat larger values (±20% with respect to the MIM) than those of the other instruments, with a contribution to the differences of only 3-5% likely attributable to the decreasing trend in stratospheric total chlorine following the Montreal Protocol regulations.

The overall agreement between the satellite ClO measurements excluding MIPAS(1) (see **Figure 4.19.5**) and as expressed by the relative standard deviation is very good to good throughout the MS and US (5-15%), and especially in regions where maximum values are found (US around 3 hPa at high latitudes). The uncertainty in the atmospheric mean state increases towards regions where ClO mixing ratios are close to the detection limit of the instrument (*e.g.*, tropical UTLS, or LM) with relative standard deviations larger than 50% (or absolute mixing ratios of only 20 pptv).

The above results and comparison with earlier validation literature highlight the general usefulness of the climatological validation approach even for shorter-lived species, yielding complementary information on latitude-height dependent measurement differences to that obtained from profile coincidences. For ClO, expected diurnal variations are relatively small during the day, so that most instruments can be compared directly with each other using daytime climatologies. An exception to this is SMR, which measures shortly before the 'stable' daytime ClO mixing ratios are reached, and for which scaling to mid-day values improves the comparisons significantly. Note that in regions where ClO variability is higher or daytime sampling differs from the stated equator crossing time (as in the case for Aura-MLS measurements at high latitudes), larger differences are found. To minimise such effects, validation and related inter-comparisons necessitate a lot of care and the use of selected coincident profiles (in terms of SZA and location) from either ground-based or satellite data.

4.20 Hypochlorous acid – HOCl

Hypochlorous acid (HOCl) in the stratosphere and LM is a short-lived reservoir gas for active chlorine (ClO_x) and hydrogen oxide (HO_x) species [*von Clarmann et al.*, 2012]. HOCl is photolyzed or reacts with atomic oxygen or hydroxyl radicals (OH) to form reactive chlorine, thereby contributing to ozone loss in the HOCl catalytic cycle [*Johnson et al.*, 1995]. HOCl is produced by the reaction between ClO and HO_2 (see *Sections 4.19* and *4.23*, respectively). The build-up of HOCl depends on the abundance of HO_2 relative to atomic oxygen (which competes for ClO to form the chlorine radical Cl), and available sunlight (which determines the rate of HOCl photolysis). HOCl can also be transformed into reactive chlorine species in heterogeneous processes on polar stratospheric clouds. Increases in upper stratospheric polar HOCl abundances following strong solar proton events have been observed in satellite data (MIPAS and Aura-MLS), as discussed by *von Clarmann et al.* [2005] and *Damiani et al.* [2012]. There is also some tropospheric HOCl chemistry in the marine boundary layer that produces HOCl, however, tropospheric concentrations remain very small because the HOCl lifetime (determined by wet scavenging there) is even shorter than in the stratosphere.

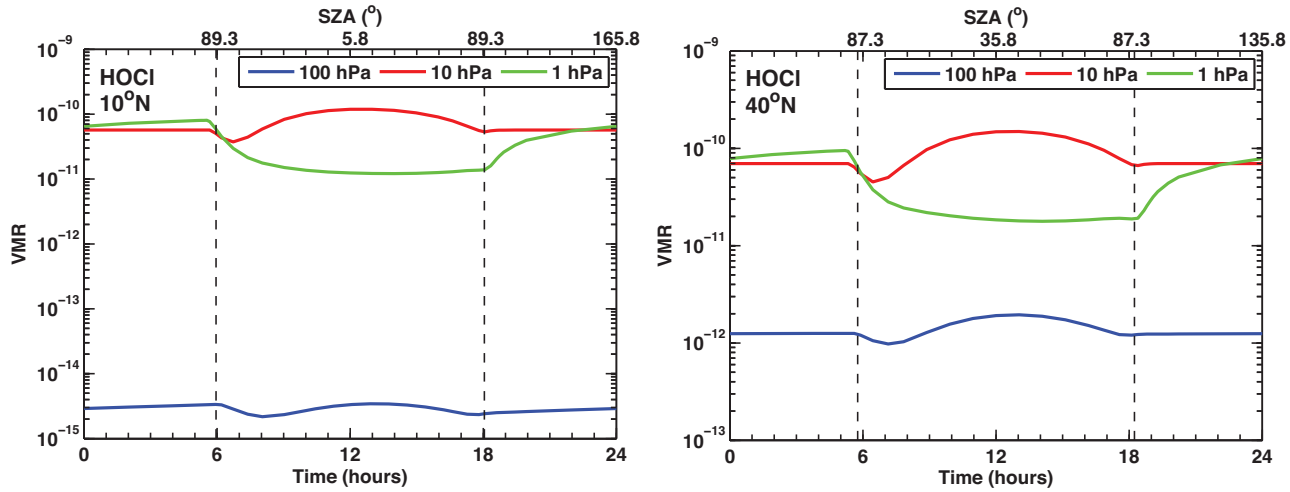


Figure 4.20.1: Diurnal cycle in HOCl. *HOCl diurnal variations are shown as function of LST or SZA at 10°N (left panel) and 40°N (right panel) for 1, 10 and 100 hPa. The diurnal cycle is derived using a chemical box model [McLinden et al., 2010].*

Figure 4.20.1 shows the diurnal cycle in HOCl at two different latitudes as a function of local solar time or solar zenith angle and for three different pressure levels as derived from a chemical box model [*McLinden et al.*, 2010]. The diurnal cycle in HOCl varies with altitude, depending on the relative strengths of formation and destruction reactions of HOCl and the abundances of species that lead to HOCl formation (in particular HO_2 and ClO). A quantification of the increases in night-time HOCl found in the US due to the reaction ClO + HO_2 -> HOCl is presented in *Kuribayashi et al.* [2014]. An evaluation of the diurnal cycle from SMILES measurements can be found in *Kreyling et al.* [2013]. A comparison of satellite-based HOCl measurements, which correspond to different local solar times would hence ideally account for the dependence of HOCl on the SZA.

4.20.1 Availability of HOCl measurements

The assessment of HOCl is based on the climatologies from Aura-MLS, MIPAS, and SMILES observations. **Tables 4.20.1** and **4.20.2** compile information on the availability of HOCl measurements, including data version, time period, altitude range, vertical resolution, and references relevant for the data product used in this report. Measurements of HOCl are available for the high-spectral MIPAS measurements between 2002-2004, Aura-MLS from 2004 onwards, and SMILES between October 2009 and April 2010. The Aura-MLS measurement range is limited to the US. Measurements are available for both night-time and daytime, but only night-time measurements will be presented and evaluated in this chapter because of a smaller diurnal cycle amplitude during the night. A data-comparison taking

Table 4.20.1: Available HOCl 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.*

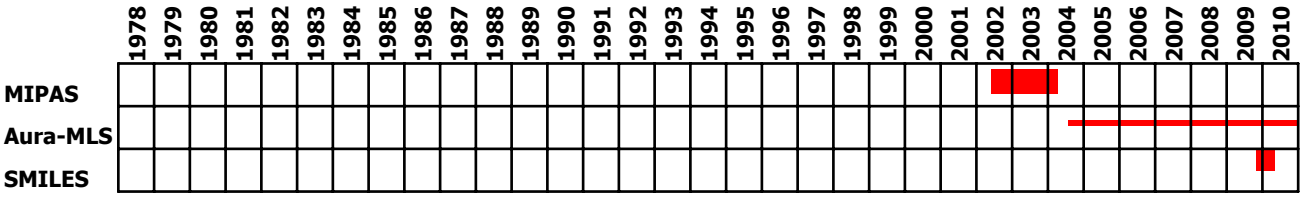


Table 4.20.2: Data version, time period, vertical range, vertical resolution, references and other comments for HOCl measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
Aura-MLS V3.3	Jul 2004 –	10 – 2.2 hPa	3 – 6 km	*Khosravi et al.*, 2013	2am/pm
MIPAS V4	Mar 2002 – Apr 2004	20 – 70 km (50 – 0.04 hPa)	9 km	*von Clarmann et al.*, 2006 *von Clarmann et al.*, 2009b	10am/pm
SMILES V2.1.5	Nov 2009 – Apr 2010	30 – 70 km	3 – 6 km	*Kreyling et al.*, 2013 *Khosravi et al.*, 2013 *Kuribayashi et al.*, 2014	day-/night-time Band A

into account the HOCl diurnal cycle has been presented by *Khosravi et al.* [2013], however with a focus on tropical data at three different altitude levels only.

4.20.2 HOCl evaluations: Zonal mean cross sections

Monthly mean zonal mean HOCl cross sections derived from night-time measurements for the time period 2002-2010 are compared in order to investigate mean biases between the various datasets. We choose to compare the night-time climatologies, since the diurnal cycle (at least below 10 hPa) shows weakest changes during the night. Measurements from MIPAS correspond roughly to 10pm, measurements from Aura-MLS correspond roughly to 2am and SMILES measurements take place over the whole nighttime.

MIPAS (2002-2004), Aura-MLS (2004-2010), and SMILES (2009-2010)

Figure 4.20.2a shows the monthly zonal mean HOCl night-time climatologies averaged over 2002-2010 for available measurements. The evaluation of the monthly mean instead of the annual mean datasets is chosen to avoid sampling biases for SMILES, which only covers a few months of the year. The MIM monthly mean climatologies for NH late autumn to early spring show highest values in the US at pressure levels between 5 and 2 hPa. Two distinct maxima are found at mid-latitudes of each hemisphere reaching up to about 200 pptv, and with somewhat higher values in the winter hemisphere than in the summer hemisphere. The maximum values increase towards the equinoxes. HOCl

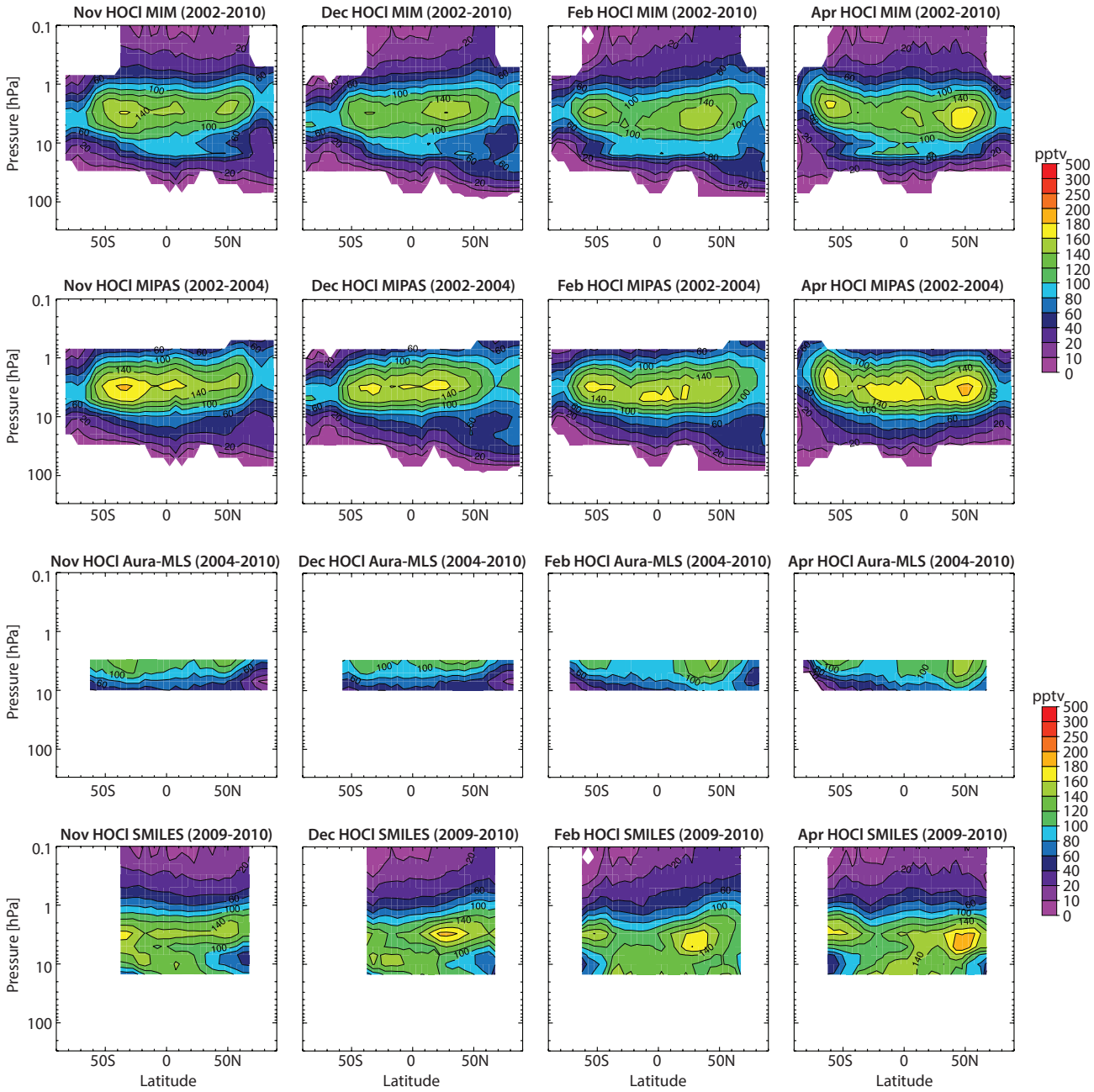


Figure 4.20.2a: Cross sections of monthly zonal mean HOCl for 2002-2010. *Monthly zonal mean HOCl cross sections for November, December, February and April over 2002-2010 are shown for the MIM, MIPAS, Aura-MLS, and SMILES.*

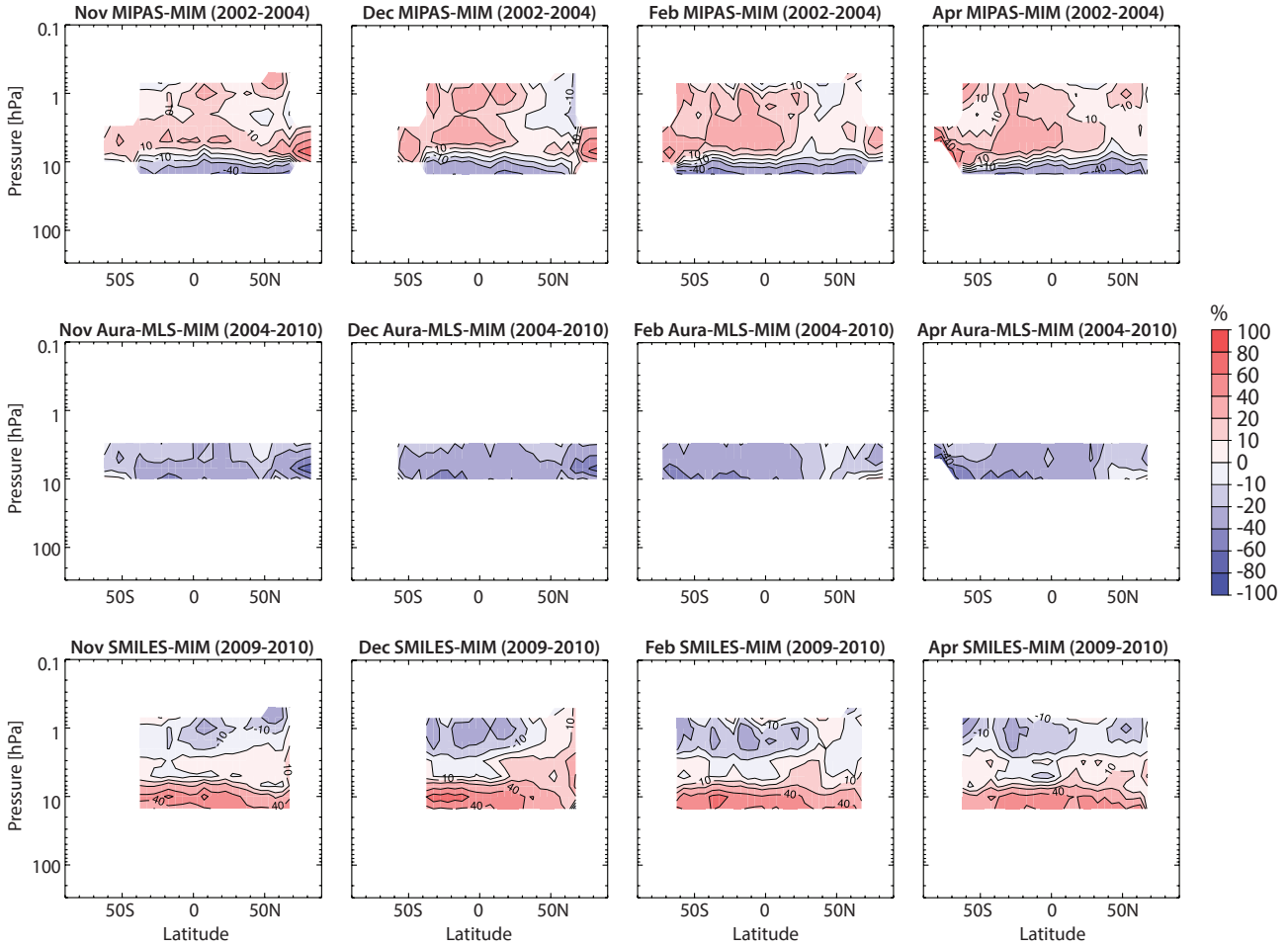


Figure 4.20.2b: Cross sections of monthly zonal mean HOCl differences for 2002-2010. *Monthly zonal mean HOCl differences for 2002-2010 between the instrument climatologies and the MIM are shown. From top to bottom: MIPAS, Aura-MLS, and SMILES.*

quickly drops to below 60 pptv in the LM above 1 hPa, and in the LS below 30 hPa. MIPAS indicates a third maximum in the tropical US region, which is not as pronounced in the SMILES observations. SMILES values, on the other hand, do not decrease as quickly below 10 hPa as seen in MIPAS (or Aura-MLS at its lower boundary). Differences are likely due to a decrease in the sensitivity of SMILES data below 10 hPa, which arises from a low signal-to-noise ratio attributable to contamination from pressure broadened ozone lines and air continuum absorption as described in *Kreyling et al.* [2013].

The differences in the instruments' HOCl climatologies relative to the MIM are displayed in **Figure 4.20.2b** and reflect the features seen in SMILES and MIPAS discussed above. The differences are strongly positive (and are opposite of what would be expected from the diurnal cycle) for MIPAS in the tropical US and for SMILES below about 6 hPa, consistent in sign with the results by *Khosravi et al.* [2013]. However, our evaluations indicate somewhat smaller differences than found in their study despite the fact that they attempted to take the diurnal cycle effects into account. Aura-MLS HOCl is slightly low throughout its limited measurements range when compared to the other two instruments' climatologies. Note the differences between Aura-MLS and SMILES are the same when looking only at the 2009/2010 months for which SMILES took measurements, so are not resulting from looking at different years. The negative trends in stratospheric chlorine since the late 1990s will contribute (by a few percent) to the positive differences seen in MIPAS data when compared to SMILES and Aura-MLS values for later years.

4.20.3 HOCl evaluations: Vertical and meridional profiles

The vertical profiles shown in **Figure 4.20.3** for subtropics (25°N-30°N) and mid-latitudes (50°N-55°N) further illustrate details in the vertical structure of the differences found in the monthly mean cross sections. In the US and LM, MIPAS and SMILES agree well or reasonably well within 10-20% except in April at mid-latitudes. This agreement is better than indicated in *Khosravi et al.* [2013], which may be partially due to their evaluation being performed in the deeper tropics. Aura-MLS shows decreasing negative differences with respect to the MIM with increasing altitude.

HOCl meridional mean profiles for 2002-2010 are shown in **Figure 4.20.4**. This evaluation highlights that MIPAS and SMILES generally agree reasonably well (within 20%) in the USLM, but that MIPAS agrees better with Aura-MLS in the MS with differences within 20%. Note that the meridional profiles also show somewhat better agreement at

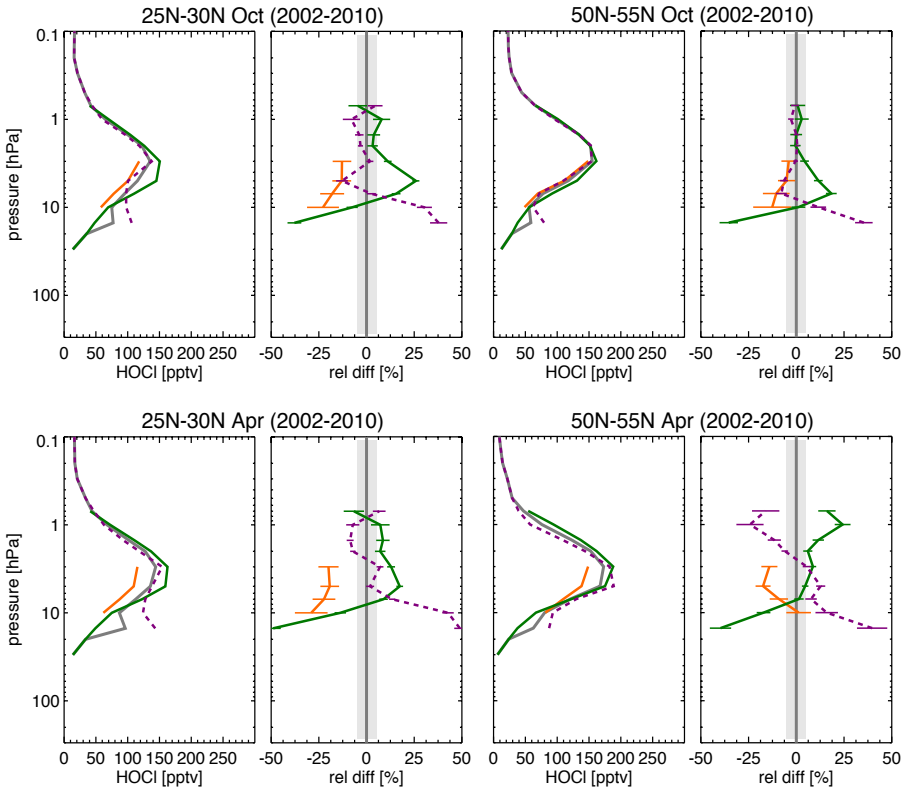


Figure 4.20.3: Vertical profiles of zonal mean HOCl for 2002-2010. *Vertical zonal mean HOCl profiles for October and April 25°N-30°N (left panels) and 55°N-60°N (right panels) are shown together with their differences from the MIM. MIPAS, Aura-MLS, and SMILES climatologies are averaged over the years 2002-2004, 2004-2010, and 2009-2010, respectively.*

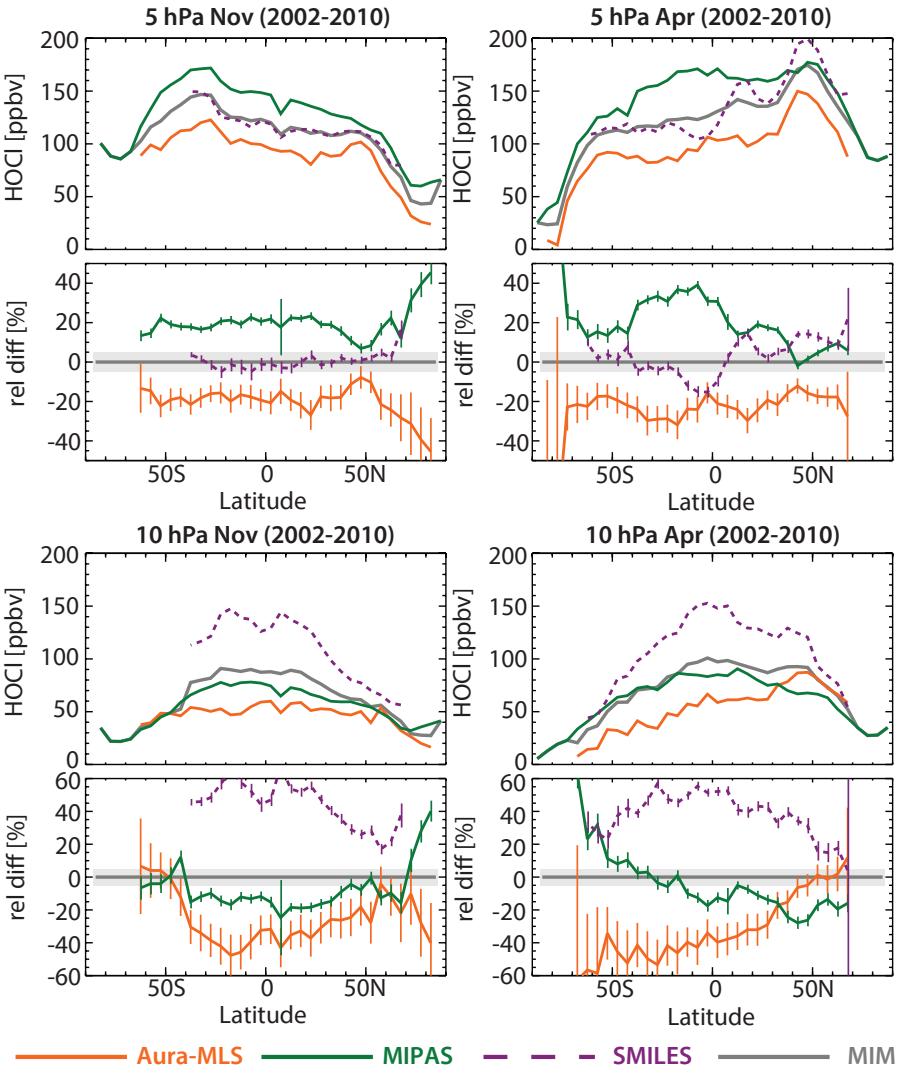


Figure 4.20.4: Meridional profiles of zonal mean HOCl for 2002-2010. *Meridional zonal mean HOCl profiles are shown for November (left) and April (right) at 5 hPa (upper row) and at 10 hPa (lower row). The climatologies are averaged over different time periods between 2002 and 2010 (see caption of Fig. 4.20.3). Differences between the individual instruments' (MIPAS, Aura-MLS, and SMILES) HOCl and the MIM profiles are shown in the lower panels.*

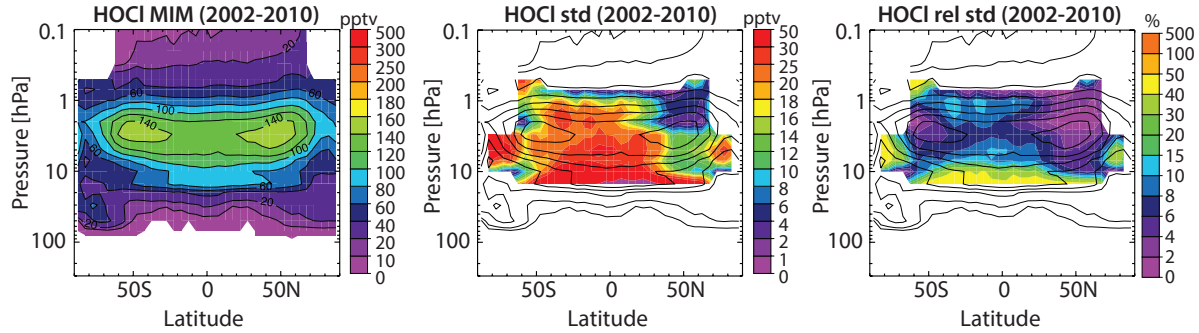


Figure 4.20.5: Summary of HOCl annual zonal mean state for 2002-2010. *Annual zonal mean cross sections for 2002-2010 of the MIM, the absolute, and the relative standard deviations over all instruments are presented from left to right, respectively. Black contours in right-hand panels repeat the MIM distribution. Instruments considered are MIPAS, Aura-MLS, and SMILES.*

mid-latitudes and the extra-tropics than in the inner tropics (20°S-20°N) where *Kohsravi et al.* [2013] chose to perform their comparison.

4.20.4 Summary and conclusions: HOCl

HOCl climatologies are available from three limb satellite instruments: MIPAS, Aura-MLS, and SMILES. While SMILES observes the full diurnal cycle, MIPAS measurements are made at about 10am/pm and Aura-MLS at about 2am/pm, respectively. The observations allow therefore for the compilation of both day- and night-time climatologies. Only the latter were evaluated here, since the impact of the diurnal cycle on night-time comparisons is deemed smaller (or even negligible below 10 hPa). The overall agreement between the instruments (see **Figure 4.20.5**) is better in the US than in the MS, despite the expectation that the diurnal cycle in HOCl may lead to a larger impact on the night-time comparison at higher altitudes.

In the US, MIPAS and SMILES agree reasonably well to within mostly 20%. In the MS, MIPAS agrees better with Aura-MLS (largely to within 30%). This can be explained by SMILES loosing sensitivity below about 10 hPa, which leads to a decrease in the quality of its HOCl product [*Kreyling et al.*, 2013]. Aura-MLS is generally on the low side of the other two instruments, with a few percent of the differences to MIPAS being explained by negative trends in stratospheric chlorine between the measurement periods of MIPAS (2002-2004) and Aura-MLS (2004-2010).

4.21 Bromine oxide – BrO

Bromine oxide (BrO) is another important reactive species involved in the catalytic destruction of stratospheric ozone, especially in the LS below about 20 km [*Brasseur and Solomon*, 1986]. While stratospheric abundances of inorganic bromine (Br_y) are much smaller than those of inorganic chlorine, bromine is around 60 times more effective in destroying ozone on a per-atom basis when compared to chlorine [*Sinnhuber et al.*, 2009]. The primary sources of stratospheric BrO are organic bromine-containing compounds including the longer-lived methyl-bromide (CH_3Br; from both anthropogenic and natural sources), halons (*e.g.*, CF_2ClBr, $CBrF_3$; from anthropogenic sources), and very short-lived bromine-containing source gases such as CH_2Br_2, $CHBr_3$, $CHBr_2Cl$ (mainly from natural sources) that are transported from Earth's surface into the stratosphere [*Carpenter et al.*, 2014]. In the stratosphere, these compounds are photolyzed by UV into bromine, which then reacts with ozone to form BrO. BrO is the most abundant inorganic bromine species in the LS and MS during daylight, making up to 50% of total inorganic bromine (Br_y; *Brasseur and Solomon*, 2005). Total inorganic bromine entering the stratosphere in 2012 as derived from tropospheric observations was 20.1 pptv, with a contribution from very-short lived substances of around 5 pptv [*WMO*, 2014].

BrO exhibits a relatively strong diurnal cycle in the LS and MS, and a somewhat smaller diurnal cycle in the US. **Figure 4.21.1** shows examples of the diurnal cycle of BrO at two different latitudes and as a function of local solar time or solar zenith angle for three different pressure levels as derived from a chemical box model [*McLinden et al.*, 2010]. A comparison of satellite-based BrO measurements corresponding to different local solar times would, ideally, account for this dependence on SZA.

4.21.1 Availability of BrO measurements

The assessment of BrO is based on the climatologies from OSIRIS, SCIAMACHY, and SMILES observations available to the SPARC Data Initiative. The climatologies start in 2001, with OSIRIS providing the longest record. Even where both night-time and daytime BrO climatologies are available for an instrument, only daytime measurements are evaluated here since diurnal variations are smaller during the day (see **Figure 4.21.1**). OSIRIS measurements available for 6:30am and scaled to 10am, SCIAMACHY measurements available for 10am (equator crossing time) and scaled to 10am (note that the local time is substantially different at higher latitudes), and SMILES daytime measurements are used.

Other BrO measurements are produced from Aura-MLS over a limited altitude range between 10 and 4.6 hPa [*Millán et al.*, 2012], although these were not provided as part of this comparison work; the latter authors note that for the 10 hPa level

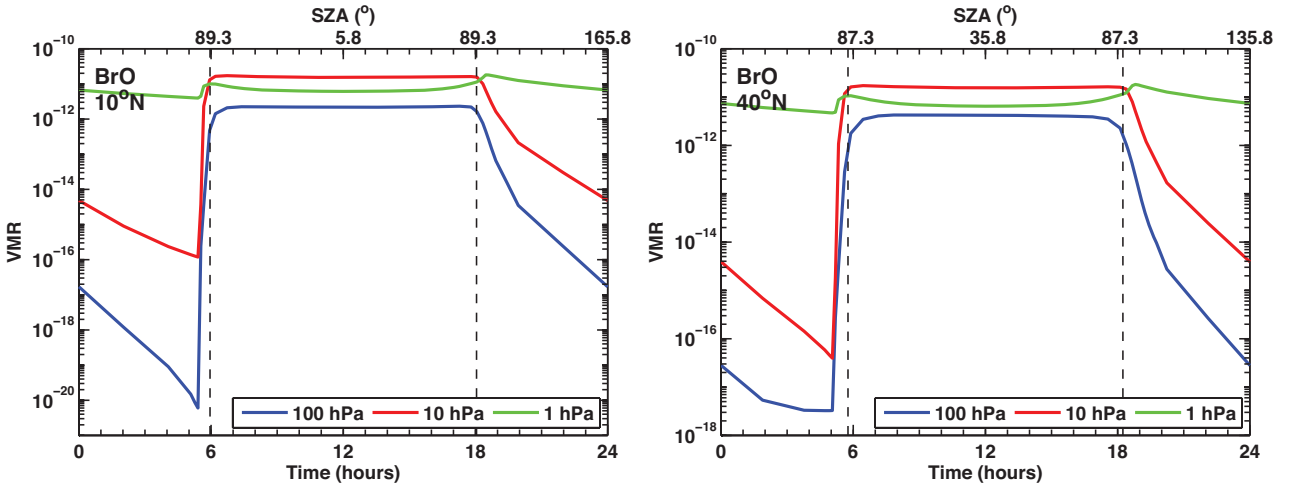


*Figure 4.21.1: **Diurnal cycle in BrO.** BrO diurnal variations are shown as function of LST or SZA at 10°N (left panel) and 40°N (right panel) for 1, 10 and 100 hPa. The diurnal cycle is derived using a chemical box model [McLinden et al., 2010].*

of overlap, MLS BrO generally agrees, within the respective uncertainties, with OSIRIS and SCIAMACHY BrO.

Tables 4.21.1 and **4.21.2** compile information on the availability of BrO measurements, including time period, height range, vertical resolution, and references relevant for the data product used in this report.

4.21.2 BrO evaluations: Monthly zonal mean cross sections

Monthly zonal mean cross sections are analysed for daytime climatologies in order to investigate mean biases between the various datasets. Daytime climatologies are chosen, since variations in the BrO diurnal cycle are smaller during the day than the night.

SMILES (2009-2010), OSIRIS, and SCIAMACHY (2003-2010)

Figure 4.21.2a shows the monthly zonal mean BrO climatologies for 2003-2010 (except for SMILES, 2009-2010) of all instruments available as derived from their daytime measurements and with the data for OSIRIS and SCIAMACHY scaled to 10am. The SMILES measurement range does not overlap substantially with OSIRIS or SCIAMACHY, so the comparison in the USLM is basically between the two SMILES products themselves as derived from the A-band (SMILES(1)) versus C-band (SMILES(2)) measurements. The comparison between OSIRIS and SCIAMACHY is further limited by the seasonally varying latitude coverage of OSIRIS.

The multi-instrument mean (MIM) reveals the structure of BrO throughout the domain, however it should be kept in

*Table 4.21.1: **Available BrO 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.*

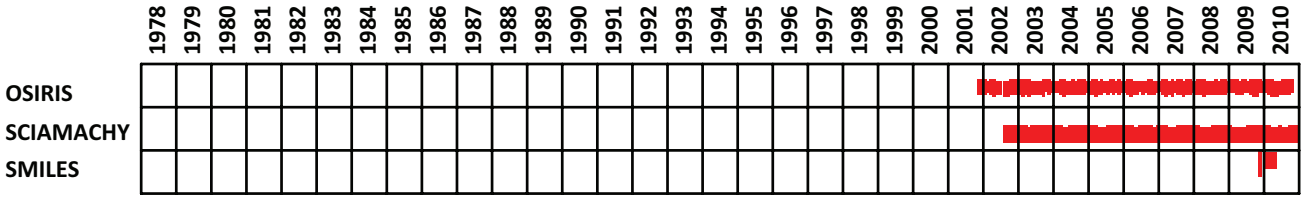


Table 4.21.2: Time period, vertical range, vertical resolution, references and other comments for BrO measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
OSIRIS V5.0	Oct 01 –	16 – 36 km (100 – 5 hPa)	3 – 5 km	*McLinden et al.*, 2010	Zonally averaged spectra with 10° lat bins; 6:30am/pm and scaled to 10am/pm
SCIAMACHY V4.1	Sep 2002 – Apr 2012	14 – 30 km (150 – 10 hPa)	3 – 5 km	*Rozanov et al.*, 2011	10am (LT of crossing at equator) and scaled to 10am/pm
SMILES V2.1.5	Oct 2009 – Apr 2010	30 – 60 km (10 – 0.2 hPa)	4 – 5 km (stratosphere) 5 – 8 km (mesosphere)	*Kreyling et al.*, 2013 *Kasai et al.*, 2013	Band A (SMILES(1)) and C (SMILES(2)) measurements day/night

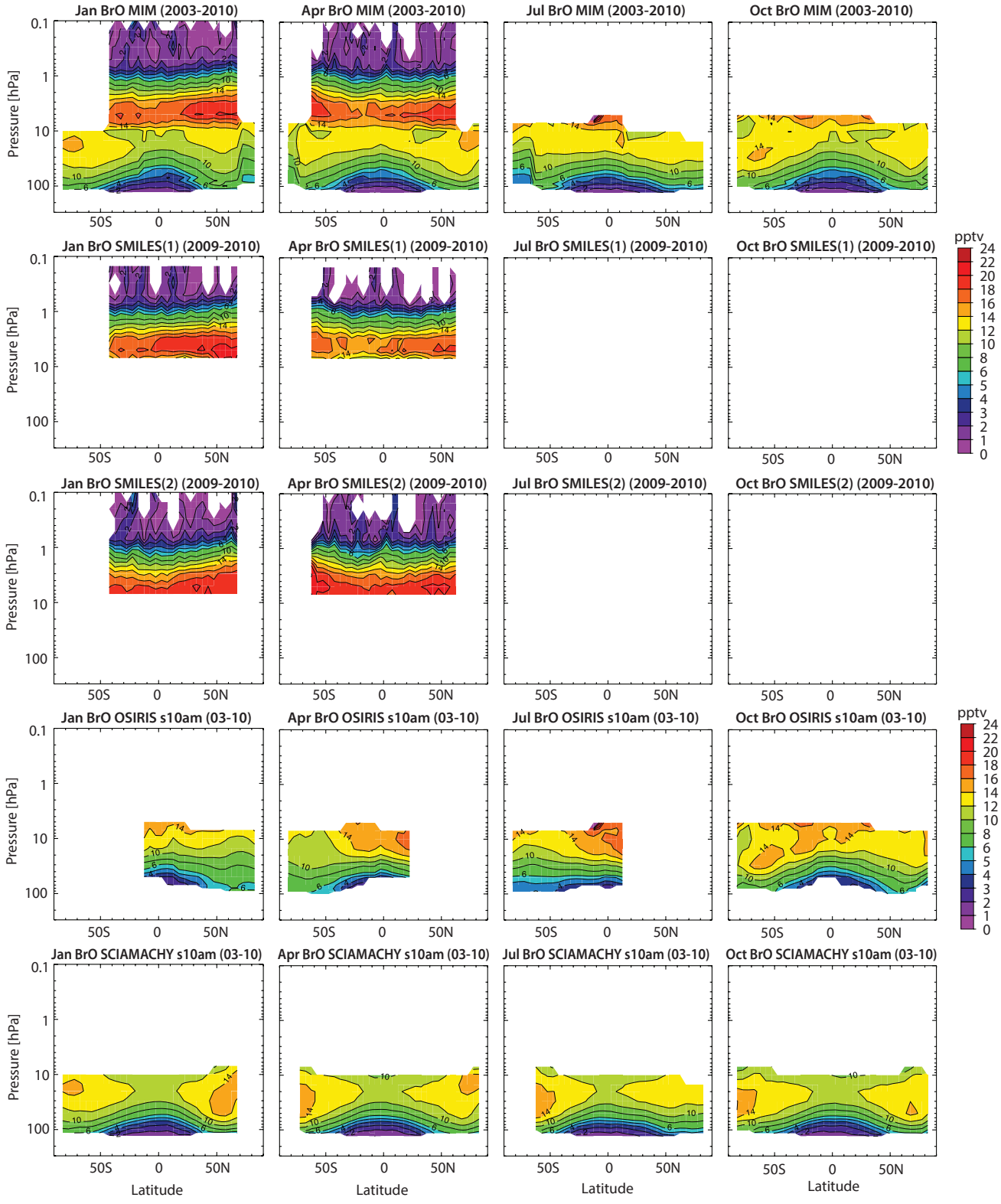


Figure 4.21.2a: Cross sections of monthly zonal mean BrO for 2003-2010. *Monthly zonal mean BrO cross sections are shown for the MIM, daytime SMILES(1) and SMILES(2) (2009-2010) measurements, OSIRIS 6:30am measurements scaled to 10am (2003-2010), and SCIAMACHY 10am measurements scaled to 10am (2003-2010).*

mind that this distribution exhibits considerable uncertainty because it is based on at most two instruments (just one in the USLM). Very low mixing ratios (< 1 pptv) are found in the tropical UT. Beyond the tropopause, and for the first few kilometers in the LS, mixing ratios are steadily increasing (with values up to around 8 pptv), and with isopleths that are seen to roughly follow the tropopause shape as expected from a trace gas with longer-lived tropospheric gases as its source. A band of high BrO values (with a maximum of about 18 pptv) is found between 30 and 5 hPa before mixing ratios decrease again towards the LM. This behaviour is similar to that seen in ClO (see *Section 4.19*).

Figure 4.21.2b shows the differences of the monthly zonal mean climatologies from the MIM. In the MS (except around 20 hPa where the instruments agree very well), SCIAMACHY (OSIRIS) shows negative (positive) differences from the MIM of around ±10%. On the other hand in the LS, SCIAMACHY (OSIRIS) shows positive (negative) differences from the MIM that are up to ±20% and larger. The structure and magnitude of the differences is observed for all months and is consistent with the study by *McLinden et al.* [2010]. In the US, the two A-band and C-band climatologies derived from SMILES agree very well to within ±2.5% at least during April. The differences between the two products are larger during January (up to ±10%) and also increase towards the LM. As can be seen in **Figure A4.21.1** in *Appendix A4*, the scaled climatologies when compared to those that are unscaled do not lead to a considerable improvement in the agreement between the instruments' monthly climatologies. This is likely because the diurnal variations during daytime are very small.

4.21.3 BrO evaluations: Vertical and meridional profiles

The vertical and meridional profiles shown in **Figures 4.21.3** and **4.21.4** provide more detailed information on inter-instrument differences, here considering only single-month evaluations in 2009 and 2010 during the time period when SMILES was operating.

Figure 4.21.3, which shows vertical profiles at different latitudes, indicates that the three instruments' climatologies seem not to agree on the altitude level where a maximum in BrO is found. SCIAMACHY indicates a maximum around 20 hPa, OSIRIS at around 10 hPa. SMILES, although it

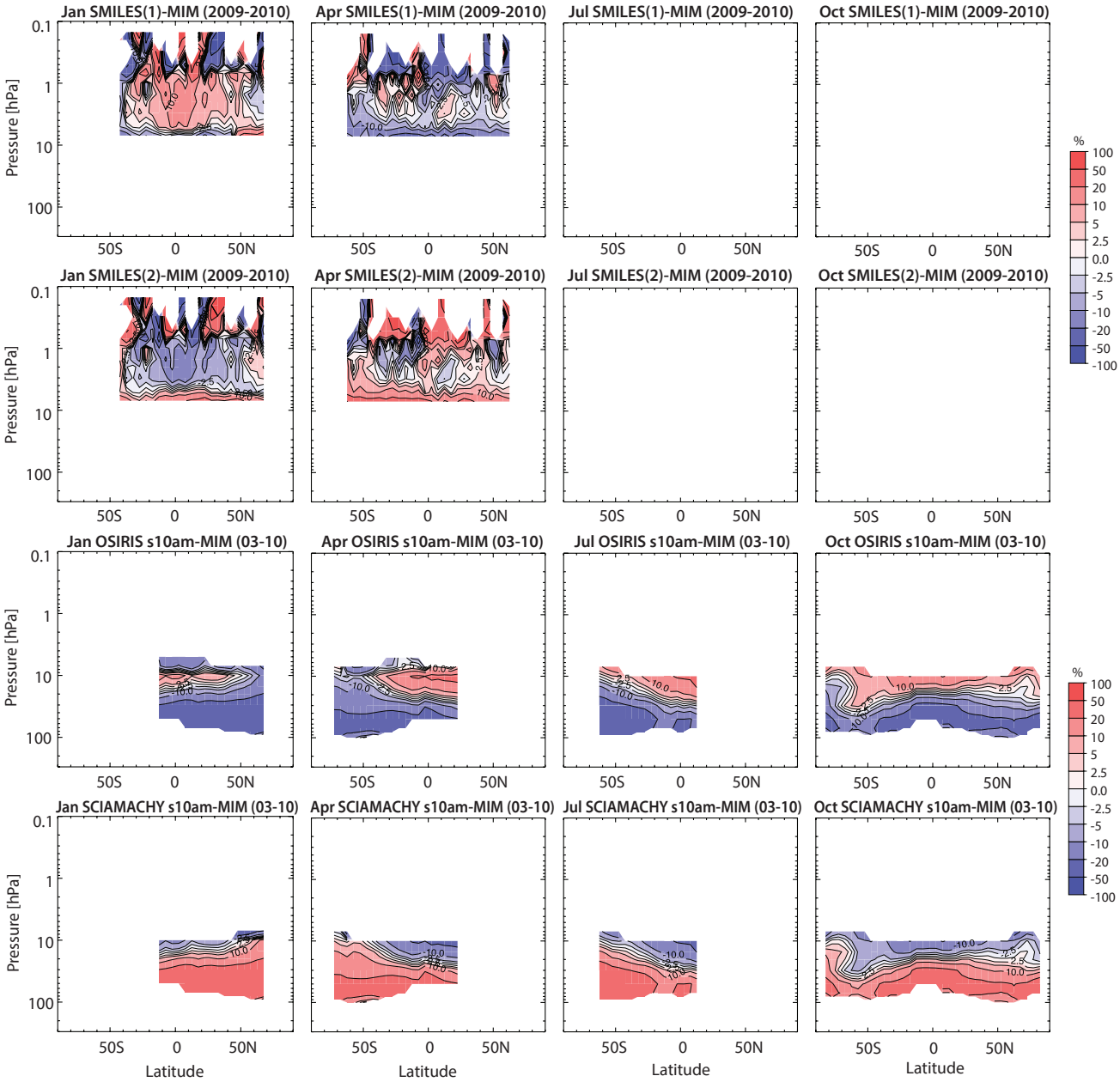


Figure 4.21.2b: Cross sections of monthly zonal mean BrO differences for 2003-2010. *Monthly zonal mean BrO differences between the individual instruments (SMILES, OSIRIS, and SCIAMACHY) and the MIM are shown.*

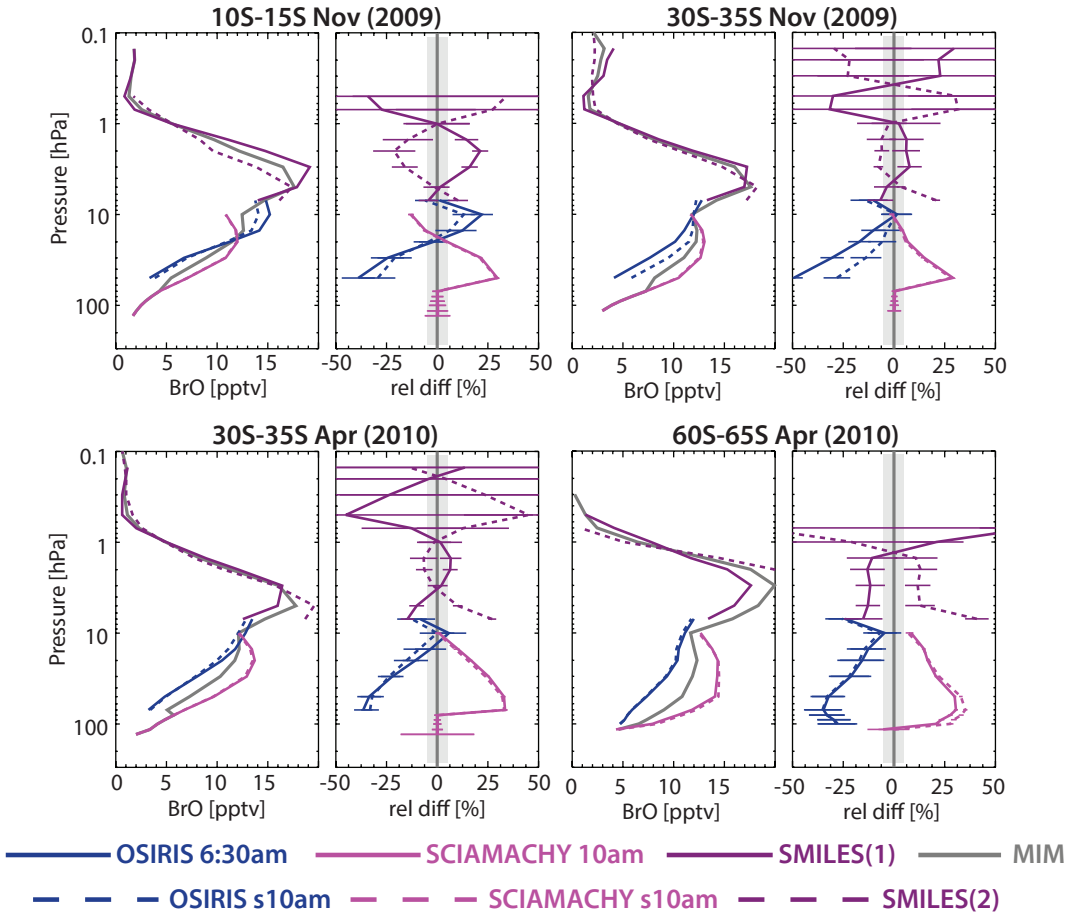


Figure 4.21.3: Vertical profiles of zonal mean BrO for 2009/2010. *Vertical zonal mean BrO profiles for November 2009 (upper panels) and April 2010 (lower panels) are shown for different latitude bands, together with their differences from the MIM. Note, the unscaled OSIRIS and SCIAMACHY products are excluded from the MIM.*

has no coverage below 10 hPa and hence cannot confirm or disprove the existence of a maximum at lower altitudes, shows a maximum at around 3 hPa (which is confirmed by Aura-MLS BrO data, although not shown here, see *Millán et al.*, 2012). The figure highlights again that SCIAMACHY and OSIRIS show only reasonably good agreement (±20% from the MIM) throughout most of the LS, but somewhat better agreement in the MS. The scaling of the climatologies in general leads to an improvement in the agreement of around 10%. The differences between these two instruments are mostly consistent during different months and at different latitudes.

Figure 4.21.4 finally shows meridional profiles during September 2009 and February 2010 at different altitudes. The figure indicates that the large differences in the LS (at 50 hPa) are statistically significant using the SEM as indicator for the uncertainty in the climatology, whereas the smaller differences in the MS (at 20 hPa) during both months lie within the range of the observational uncertainty across most latitudes except the polar region. In the limited measurement range where OSIRIS and SMILES show overlap, the scaled OSIRIS and the SMILES(1) agree mostly very well (within ±5%). Note that despite good agreement between the two SMILES products at 5 hPa, this result is not representative for higher altitudes (see **Figure 4.21.3**).

4.21.4 Summary and conclusions: BrO

BrO climatologies are available from three limb satellite instruments: OSIRIS, SCIAMACHY, and SMILES. While SMILES observes the full diurnal cycle, OSIRIS measurements are made at about 6:30am/pm and SCIAMACHY at about 10am local time at equator crossing, respectively. The observations allow therefore for the compilation of both day- and night-time climatologies. Only the daytime are evaluated here, since the impact of the diurnal cycle on daytime comparisons is deemed smaller (or even negligible). OSIRIS and SCIAMACHY cover the lower altitude range (LS and MS), while SMILES covers the higher altitude range (US and LM). SMILES offers two different BrO products (A-band and C-band retrievals), which are compared in this region.

The overall structure in BrO indicates a minimum in the tropical tropopause region, increasing mixing ratios with increasing altitude in the LS and MS, and a maximum in the US (around 5 hPa) with values around 18 pptv. The BrO distribution thereby resembles strongly the ClO distribution (see *Section 4.19*).

OSIRIS and SCIAMACHY show considerable disagreement in the LS with differences from the MIM of up to

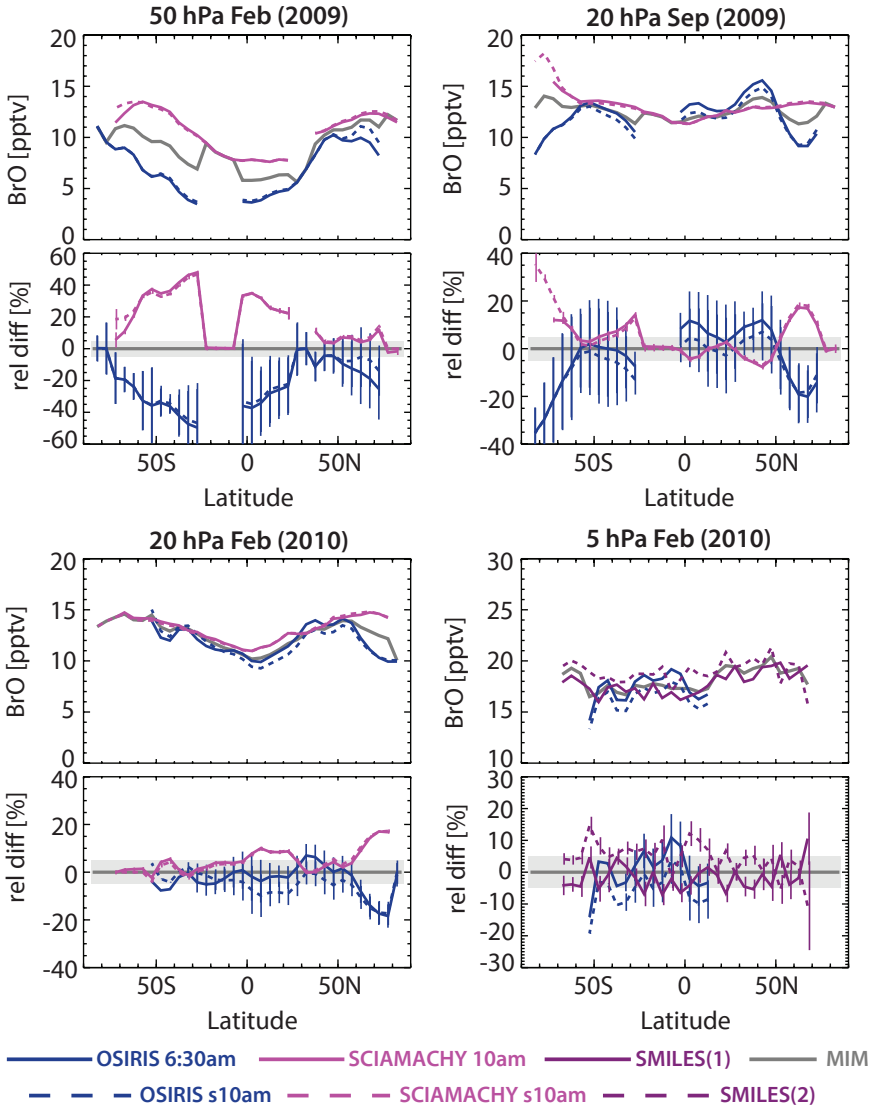


Figure 4.21.4: Meridional profiles of zonal mean BrO for 2009/2010. *Meridional zonal mean BrO profiles for September 2009 at 50 hPa and 20 hPa (upper panels), and for February 2010 at 20 hPa and 5 hPa (lower panels) are shown together with their differences from the MIM. Note, the unscaled OSIRIS and SCIAMACHY products are excluded from the MIM.*

±30%. Note, BrO mixing ratios are relatively low (around a few pptv) in the LS, hence the instruments measure near their detection limit in this region. On the other hand, good agreement (±10%) between the two instruments is found in the MS (except at high latitudes). Good to very good agreement is found between OSIRIS and SMILES(1) where their measurements overlap (around 5 hPa), whereas somewhat larger differences are found between OSIRIS and SMILES(2).

4.22 Hydroxyl radical – OH

The hydroxyl radical (OH) is one of the most reactive molecules in the atmosphere and of great importance to both tropospheric and stratospheric chemistry. OH is known as the cleansing agent of the atmosphere, since it helps in the destruction of many air pollutants and greenhouse gases. However, it is also involved in the catalytic reaction cycles of the HO_x family (OH, HO_2, H) that destroy stratospheric ozone. The HOx cycle was the first catalytic reaction cycle to be identified [*Bates and Nicolet*, 1950]. The primary source of OH in the stratosphere is the reaction of $O(^1D)$ (from the photolysis of O_3) with H_2O. In the mesosphere, another relevant source is the photolysis of water vapour. The OH abundance is generally not dominated by transport (except in polar night), and it depends mainly on local source gas abundances. OH responds very quickly to changes in available sunlight (UV radiation), which leads to OH variations on daily (see **Figure 4.22.1**), seasonal, and longer timescales (*e.g.*, following the 11-year solar cycle variations).

4.22.1 Availability of OH measurements

Measurements of OH are available to the SPARC Data Initiative only from the Aura-MLS instrument and currently only between 2004 and 2009. The climatologies discussed here are derived from daytime measurements (near 1:30pm local time at low- to mid-latitudes). Other OH measurements have been performed in the past on balloon soundings, such as from the Middle Atmosphere High-Resolution Spectograph Investigation (MAHRSI) in 1994 and 1997 [*Conway et al.*, 1999; 2000], the Far-Infrared Spectrometer (FIRS-2) [*Jucks et al.*, 1998], and the Balloon OH (BOH) instrument [*Pickett et al.*, 2006]. There have been other measurements of mesospheric OH from space, namely by the Spatial Heterodyne Imager for Mesospheric Radicals

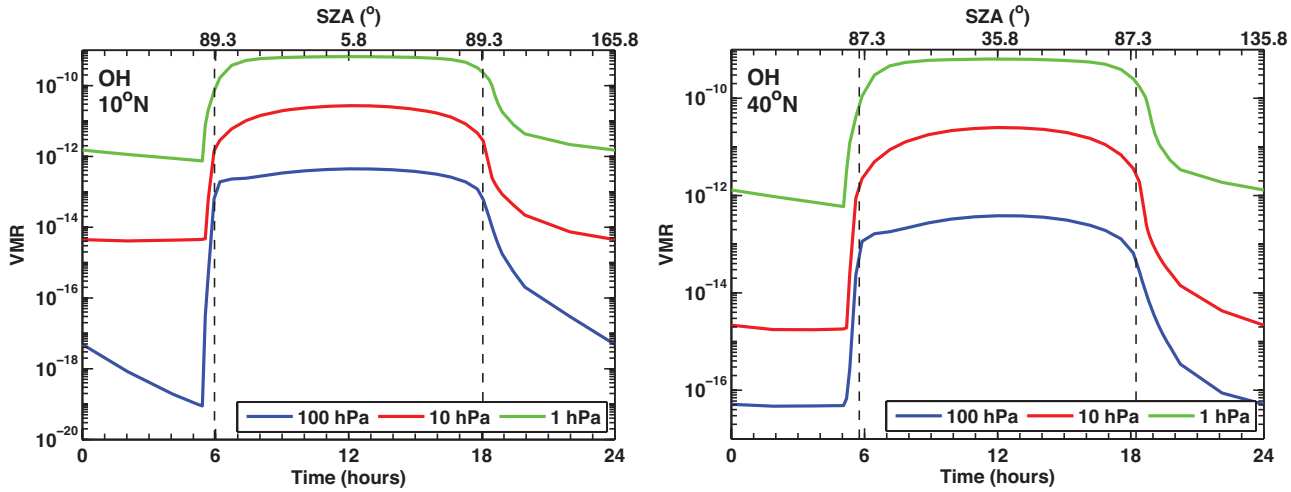


Figure 4.22.1: Diurnal cycle of OH. *OH variations as a function of LST are shown at 10°N and 40°N at 1, 10 and 100 hPa for March 15. The diurnal cycle is derived using a chemical box model [McLinden et al., 2010].*

(SHIMMER), between March 2007 and October 2009, and analyses of this dataset have been provided by *Englert et al.* [2008; 2010] and *Siskind et al.* [2013]. Good agreement (within 10-20%) between mesospheric OH profiles from SHIMMER and Aura-MLS was generally observed (at various latitudes and time periods), although the MLS values were found to be lower than SHIMMER values by more than 20% near 63-67 km [*Englert et al.*, 2010]. However, most of the altitude range for SHIMMER OH retrievals is above the range of the data files included in the SPARC Data Initiative (with a top pressure level at 0.1 hPa).

Validation of the Aura-MLS measurements has been described by *Pickett et al.* [2006] in terms of comparisons versus balloon-borne data, and by *Wang et al.* [2008] for column OH comparisons versus ground-based data. No major issues have been noted, and the MLS OH data are generally quite consistent with the few other correlative datasets. The 11-year solar cycle variation impacts on OH have also been described by *Wang et al.* [2013], using both ground-based data and the Aura-MLS dataset.

Tables 4.22.1 and **4.22.2** compile information on the availability of OH measurements, including time period, height range, vertical resolution, and references relevant for the data product used in this report.

4.22.2 OH zonal mean cross sections

Satellite OH measurements were only made available to the SPARC Data Initiative from the Aura-MLS instrument, hence there are no comparisons herein with other instruments. Instead, we show annual and monthly zonal mean cross sections from Aura-MLS daytime OH data with the aim to illustrate the general behaviour of and variability in OH in the atmosphere as known from the measurements from this instrument. Note that a correction was applied to the daytime zonal means, by subtracting long-term night-time averages, for the bottom few pressure levels (32 to 10 hPa), where this correction is non-negligible. This is done because the night-time values should be very small, and non-zero MLS night-time averages are more an indication of instrument systematics than real night-time values. The corrected daytime values are found in the SPARC Data Initiative (daytime) data files.

Notable in both the annual and monthly zonal mean OH cross sections are the rather flat distributions of the OH isopleths throughout the tropics and mid-latitudes. No strong seasonality is observed in these regions. At latitudes closer to the poles (higher than approx. 50°N or 50°S) a somewhat stronger seasonality in OH is observed, as further illustrated in the next section. In particular, OH mixing ratios

Table 4.22.1: Available OH 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.*

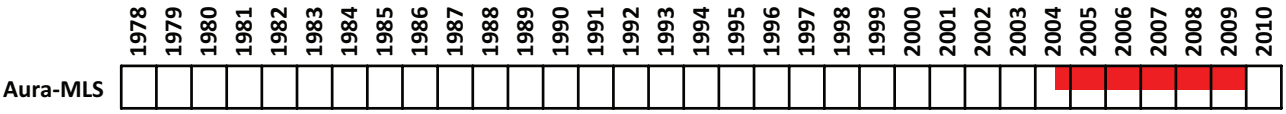


Table 4.22.2: Time period, vertical range, vertical resolution, references and other comments for OH measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
Aura-MLS V3.3	Jul 2004 – Dec 2009	32 – 0.0046 hPa	2.5 km at most pressure levels	*Pickett et al.*, 2008 *Wang et al.*, 2008 *Livesey et al.*, 2013	Only daytime measurements are considered here

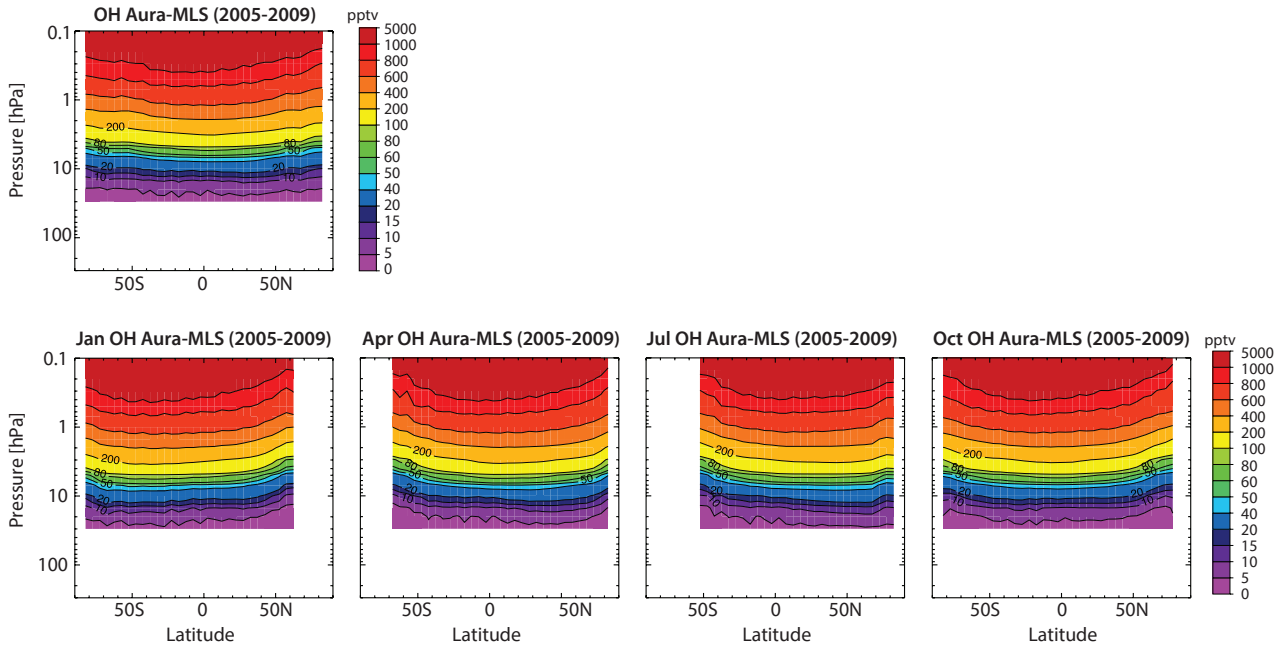


Figure 4.22.2: Zonal mean cross sections. *The annual zonal mean cross section from Aura-MLS averaged over measurements between 2005 and 2009 is shown in the upper left corner. Monthly zonal mean cross sections are shown for January, April, July, and October.*

decrease towards the poles during autumn/early winter. Changes in OH are largely driven by variations in sunlight (solar zenith angle).

4.22.3 OH vertical profiles from Aura-MLS

Aura-MLS zonal mean vertical profiles of OH illustrate the seasonality identified in the above cross sections more quantitatively. Again, in the tropics, there is basically no seasonality. The OH seasonal cycle becomes larger when moving towards higher latitudes, with decreases in the OH mixing ratios at mid-latitudes between July and January by approximately 30%, and in the NH polar region between July and October by up to 50%. The SH mid-latitude region exhibits a somewhat stronger OH decrease during winter than that in the NH.

4.22.4 Summary and conclusions: OH

Aura-MLS is the only instrument providing OH measurements within the SPARC Data Initiative. Aura-MLS OH climatologies are available between 2004 and 2009 and consist of daytime measurements (near 2pm local time, for low- to mid-latitudes). The seasonality in daytime OH is discussed. However, no comparisons with other OH measurements are provided here. An earlier version of the Aura-MLS OH data was validated by *Pickett et al.* [2006, 2008], including estimated uncertainties and comparisons with balloon-borne OH datasets. Based on this work, MLS stratospheric OH systematic uncertainties of about 10% are implied.

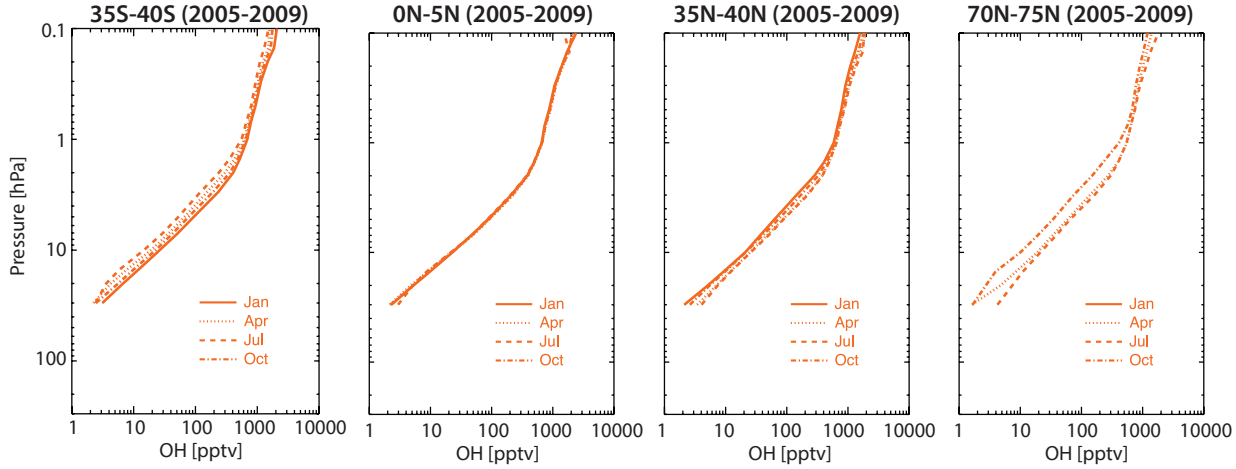


Figure 4.22.3: Zonal mean vertical profiles of daytime OH. *Zonal mean vertical profiles of OH averaged over 2005-2009 from Aura-MLS are shown for different months and latitude bands (SH mid-latitudes and NH equatorial, mid-latitude, and polar region). Note, there is no data during polar night.*

4.23 Hydroperoxy radical – HO_2

The hydroperoxy radical (HO_2) together with the hydrogen atom (H) and hydroxyl radical (OH; see *Section 4.22*) forms the HO_x family. HO_2 is formed in the reaction between a hydrogen atom and molecular oxygen (O_2), or between O_3 and an OH molecule. HO_2 is a highly reactive molecule and plays an important role in stratospheric ozone chemistry through its role in the HO_x catalytic reaction cycles that destroys ozone. The HO_x cycle was the first catalytic reaction cycle to be identified [*Bates and Nicolet*, 1950]. HO_x chemistry dominates ozone destruction above 40 km, while NO_x dominates ozone destruction in the lower stratosphere.

HO_2 exhibits relatively strong day-night differences (see **Figure 4.23.1**), in particular in the lower stratosphere. The relative differences decrease with increasing altitude. Maximum values are found during the day and are relatively constant, minimum values are found before sunrise with more or less steadily decreasing values during the night, when the main reservoir for HO_x is H_2O_2 (hydrogen peroxide), produced from the reaction between two HO_2 radicals. The daytime photolysis of H_2O_2 releases HO_x, and a rapid photochemical equilibrium is set up between OH and HO_2. Two-monthly climatologies of diurnal variations are provided from SMILES observations by *Kreyling et al.* [2013]. *Kuribayashi et al.* [2014] discussed in detail the night-time decrease in HO_2 due to the reaction HO_2 + ClO -> HOCl.

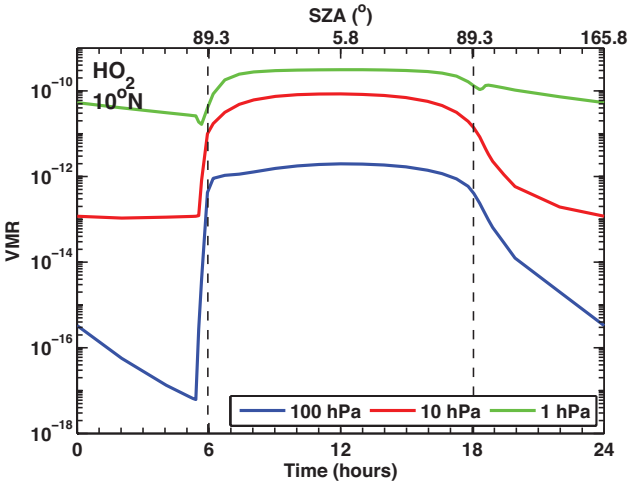


Figure 4.23.1: Diurnal cycle of HO_2. *HO_2 variations as a function of LST are shown at 10°N at 1, 10 and 100 hPa for March 15. The diurnal cycle is derived using a chemical box model [McLinden et al., 2010].*

4.23.1 Availability of HO_2 measurements

Measurements of HO_2 are available within the SPARC Data Initiative from three limb emission satellite instruments, SMILES, SMR, and Aura-MLS, which measure in the sub-mm/microwave wavelength bands. Other available stratospheric HO_2 datasets are restricted to balloon campaigns, such as from the Middle Atmosphere High-Resolution Spectograph Investigation (MAHRSI) in 1994 and 1997 [*Conway et al.*, 1999; 2000] and Far-Infrared Spectrometer (FIRS-2) [*Jucks et al.*, 1998]. There is no temporal overlap of the measurements from the three satellite instruments, since SMR currently provides HO_2 data only as a research product during 2003 and 2004, while the SMILES mission lasted from October 2009 to April 2010. Another difficulty in comparing the observations of these three instruments stems from the fact that they measure at different solar zenith angles, which poses a problem for all short-lived species with strong diurnal cycles. Aura-MLS and SMR are in sun-synchronous orbits and measure at about 1:30am/pm and 6:30am/pm, respectively. SMILES on the ISS was in a non-sun-synchronous orbit and hence measured the full diurnal cycle. For HO_2, no attempt is made here to correct for the differences in local measurement times. However, we focus on the daytime climatologies of SMILES and Aura-MLS, since the HO_2 abundances are larger and fairly constant during the day, and on SMR-am data, which are expected to be closer to daytime values than the SMR-pm data (compare **Figure 4.23.1**). For the MLS HO_2 daytime data, instrumental biases obtained from the night-time values have been subtracted, following the data usage recommendations from the MLS team [see *Pickett et al.*, 2008; *Livesey et al.*, 2013].

Tables 4.23.1 and **4.23.2** compile information on the availability of HO_2 measurements, including time period, height range, vertical resolution, and references relevant for the data product used in this report.

4.23.2 HO_2 evaluations: Zonal mean cross sections

Monthly zonal mean cross sections are analysed to investigate mean biases between the various datasets. Note that the climatologies of the different instruments are not representative of the same years or solar zenith angles, which likely explains some of the differences found between the datasets. SMILES daytime, Aura-MLS 1:30pm, and SMR 6:30am climatologies are used. A study by *Khosravi et al.* [2013] accounts for the diurnal cycle in the validation of the three instruments, however for limited pressure levels, seasons, and latitudes only.

SMR (2003-2004), Aura-MLS (2004-2010), and SMILES (2009-2010)

Figure 4.23.2 shows the comparison of the monthly zonal mean HO_2 climatologies from the different instruments in November and April, averaged over their respective observation periods. As seen in the MIM, mixing ratios are similar in both months in the tropics, indicating a weak seasonal cycle in the daily values in this region, where solar zenith angles do not vary very much with season. Lowest mixing ratios are found in the polar regions of the winter hemisphere (during high solar zenith angle conditions), indicating a more pronounced seasonal cycle in the daily values in these regions.

Table 4.23.1: Available HO_2 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.

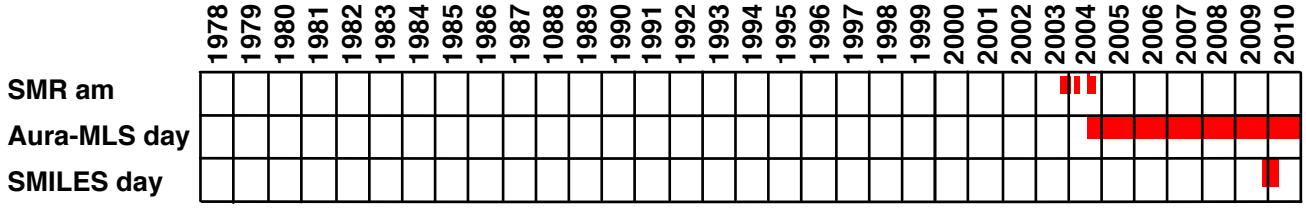


Table 4.23.2: Time period, vertical range, vertical resolution, references and other comments for HO_2 measurements.

Instrument	Time Period	Vertical range	Vertical resolution	References	Comments
MLS Aura V3.3	Jul 04 –	22 – 0.046 hPa	4 – 10 km	*Pickett et al.,* 2006 *Pickett et al.,* 2008 *Khosravi et al.,* 2013	SPARC Data Initiative files are daytime zonal monthly means (corrected by subtracting the small means from night climatology)
SMILES V2.0.1	Oct 09 – Apr 10	26 – 95 km (20 – 0.001 hPa)	4 – 5 km	*Kreyling et al.,* 2013 *Khosravi et al.,* 2013 *Kuribayashi et al.,* 2014 *Millán et al.,* 2015	
SMR V2.1	Oct 03 – Oct 04	30 – 60 km (10 – 0.3 hPa)	4 – 8 km	*Khosravi et al.,* 2013	

The differences between the SMILES and Aura-MLS instruments show distinct features. Aura-MLS exhibits lower values than SMILES by 10-20% through most of the stratosphere and LM, indicating a good to reasonably good agreement between the two instruments, similar to the differences found in *Khosravi et al.* [2013]. Differences are somewhat larger in the respective autumn hemisphere. Aura-MLS (SMILES) show positive (negative) differences at higher latitudes and the lowest/highest altitude regions covered by the SPARC Data Initiative climatologies. Note that using only years for which the instruments overlap yields somewhat better agreement between the datasets, but the overall structures in the differences remain similar. The SMR monthly zonal mean climatologies are based on a few days of data only for each of the chosen months and hence were not included in the calculation of the MIM. SMR exhibits lower values throughout the USLM by more than 50%. Since sunset observations can be more than 100% smaller than daytime measurements, scaling the SMR to daytime values for the comparison would be expected to yield better comparison results. *Khosravi et al.* [2013] have compared these three instruments' measurements in the tropics at three different pressure levels, scaling the data with a model (via local SZA), and find relatively good agreement (within about 10-20%, and also for SMR); this is better than their combined systematic uncertainty estimates (which are at least 15-20%).

4.23.3 HO_2 evaluations: Vertical profiles

The vertical profiles in **Figure 4.23.3** reveal further details of the structure in the differences of the monthly mean cross sections. In order to minimise sampling biases, we here compare the instruments during years of overlap only. Aura-MLS and SMR profiles are compared for August (65°N-70°N) and September (0°S-5°S) for the year 2004. Aura-MLS and SMILES are then compared for November (55°N-60°N) in 2009, and for April (10°S-15°S) in 2010.

For SMR, the time-consistent comparison yields better results than the multi-annual monthly mean comparison in *Section 4.23.2*. Due to its small number of profiles going into the monthly mean, SMR shows however large fluctuations in HO_2 mixing ratios with altitude, and a larger SEM in its differences from the MIM. SMR does not exhibit consistently lower values than Aura-MLS, as would be expected from its sampling times, although the overall behaviour *versus* pressure for this species with strong vertical gradients and small abundances is matched fairly well. For August at northern high latitudes, differences between SMR and Aura-MLS are mostly within 50%. For the September comparison in the Southern Hemisphere tropics, larger differences up to 100% are found, but SMR shows larger SEM values, leading to the differences not being statistically significant.

The comparisons between Aura-MLS and SMILES vertical profiles show consistent structures in their relative differences as in the zonal monthly mean cross section evaluation over the longer time period, with very good agreement in the tropical stratosphere during April. Relative inter-instrument differences lie mostly within 20% except for the lowest levels of the measurement range (with small mixing ratios) (around 10 hPa), where the differences increase to over 100%.

4.23.4 Summary and conclusions: HO_2

HO_2 climatologies are available from three limb-sounders: SMR, Aura-MLS, and SMILES. Here, we compare daytime climatologies from Aura-MLS and SMILES, since daytime HO_2 values are fairly constant allowing for a meaningful comparison despite different local sampling times. For SMR,

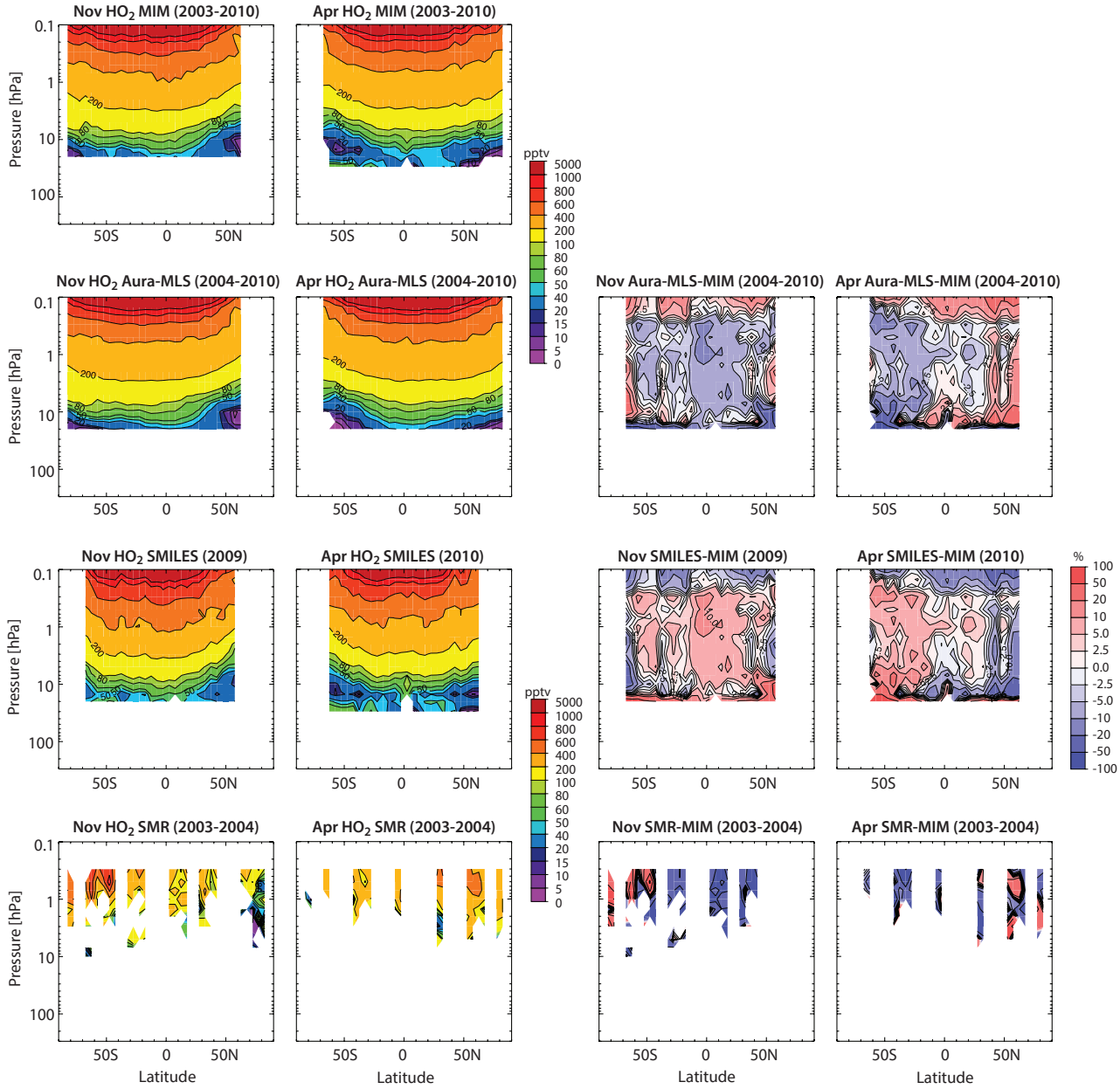


Figure 4.23.2: Cross sections of monthly zonal mean HO_2 and differences from the MIM. *Monthly zonal mean HO_2 cross sections (left two columns) and their differences from the MIM (right two columns) for the different instruments are shown for November and April as averaged over available years. Note SMR is excluded from the MIM due to its limited number of profiles in these monthly zonal mean climatologies.*

we use the climatologies derived from am-measurements, since they are expected to be closer to the daytime measurements than the pm-measurements, although on the low side of those due to the increase in HO_2 when sunlight starts appearing. Note that for a more detailed comparison of SMR with SMILES and Aura-MLS, scaling towards the daytime measurement would have to be applied [see *Khosravi et al.*, 2013]. Overall findings on the systematic uncertainty in our knowledge of the HO_2 mean state are presented in the following summary including the discussion of **Figure 4.23.4**.

Atmospheric mean state

Knowledge of the atmospheric HO_2 annual mean state presented here is based mainly on Aura-MLS, with the information being corroborated primarily by comparison with SMILES. The uncertainty in the annual mean state as derived from the two satellite instruments is smallest in the LM, and also in the tropical MS and US with values smaller than ±10%. The uncertainty increases towards the mid-latitudes with inter-instrument spreads of ±30%. Largest uncertainties are found at the lowest levels at which measurements are available (around 20 hPa). Here, HO_2 mixing ratios are very low (around 40 pptv) and differences increase to up to ±100%. The annual mean results are consistent with single-month comparisons.

Overall, the comparisons between Aura-MLS and SMILES daytime climatologies show promising results and good agreement over much of the altitude range the measurements cover (with inter-instrument differences mostly well

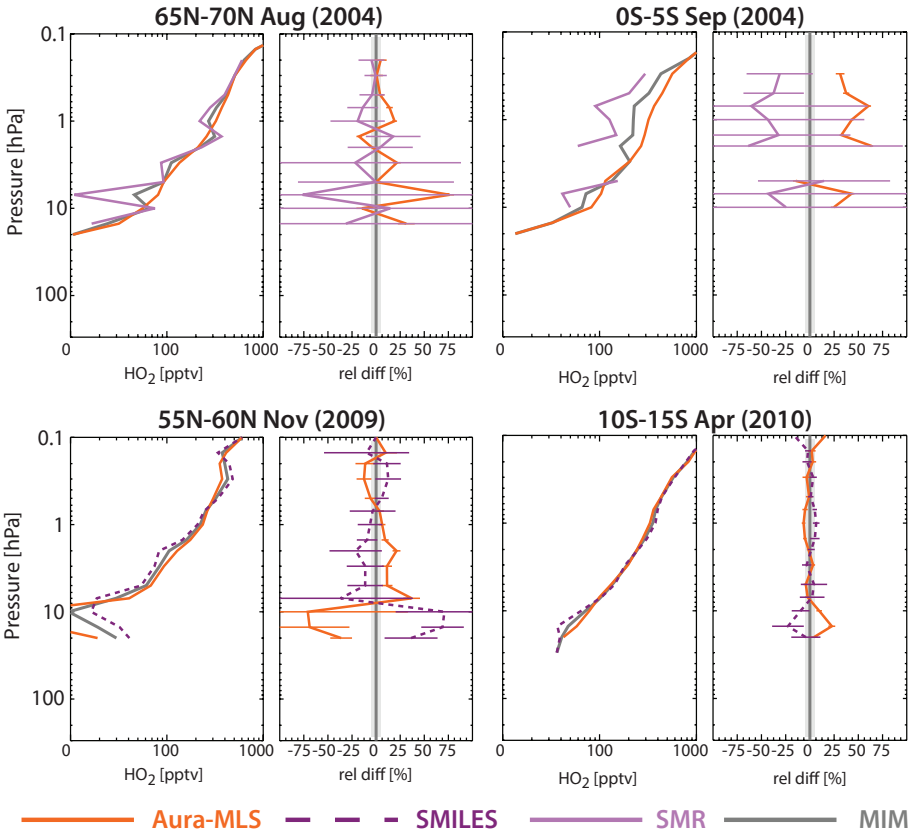


Figure 4.23.3: Altitudinal profiles of zonal mean HO_2. *Altitudinal zonal mean HO_2 profiles are shown along with their differences from the MIM for 65°N-70°N in August and 0°S-5°S in September 2004 for SMR and Aura-MLS (upper panels), and for 55°N-60°N in November 2009 and 10°S-15°S in April 2010 for SMILES and Aura-MLS (lower panels). Error bars in the difference plots indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates that there are few regions where relative differences are smaller than ±5%.*

within 20%, comparable to the *Khosravi et al.*, [2013] results for the tropical region). SMR measurements exhibit larger deviations from these other two instruments, most likely attributable to using unscaled climatologies that are derived from measurements at higher solar zenith angles, but maybe also because a smaller number of SMR profiles are used in the climatologies presented here.

We note also that a recent offline retrieval dataset consisting of daily zonal averages for Aura-MLS HO_2 has been created [see *Millán et al.*, 2015], with some advantages over the MLS routine production data (in the spatial and vertical coverage as well as in the night-time values).

4.24 Formaldehyde – CH_2O

Formaldehyde (CH_2O) is the simplest aldehyde in the atmosphere. CH_2O is an important intermediate in the oxidation of methane and other hydrocarbons, involving sunlight and oxygen. In the stratosphere, its main source is the oxidation of methane. CH_2O is destroyed in the reaction with OH or via photolysis [*Brasseur and Solomon*, 1986]. CH_2O exhibits a small diurnal cycle at least in the middle and upper stratosphere, shown for three different pressure levels derived with a chemical box model [*McLinden et al.*, 2010] in **Figure 4.24.1**. In the troposphere, sources in addition to CH_4 oxidation include oxidation of non-methane hydrocarbons resulting from biomass and fossil

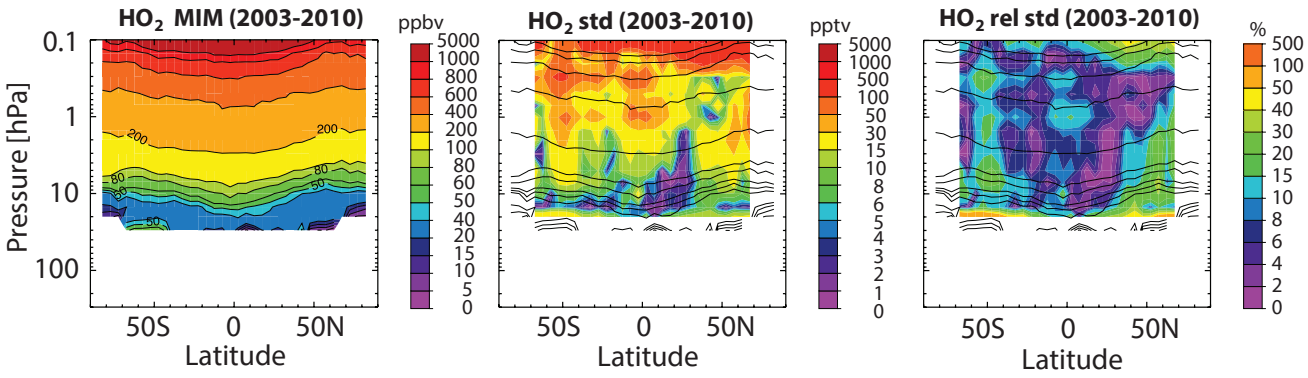


Figure 4.23.4: Summary of HO_2 annual zonal mean state for 2003-2010. *The figures show the annual zonal mean cross sections for 2003-2010 including the MIM, the absolute standard deviation, and the relative standard deviation over the HO_2 fields (from left to right, respectively). Black contours in the middle and right-hand panels repeat the MIM distribution. Instruments considered are Aura-MLS and SMILES. SMR is not included in the MIM due to its limited sampling.*

fuel burning. There are also industrial sources from manufacturing of building materials and household products. Smoking also produces CH_2O. CH_2O is highly toxic to humans and animals, causing allergic reactions, eye, nose, and lung irritations, and potentially cancer.

4.24.1 Availability of CH_2O measurements

CH_2O climatologies are available to the SPARC Data Initiative from the MIPAS and ACE-FTS instruments, which measure in the mid-IR (see **Table 2.2**). MIPAS provided measurements only during the first years of its mission, when the instrument was operating in the high spectral resolution mode [*Steck et al.*, 2008]. It obtains measurements from the upper troposphere to the upper stratosphere. ACE-FTS on the other hand extends into the mid troposphere while its upper bound is limited to the lower stratosphere [*Coheur et al.*, 2007; *Dufour et al.*, 2009]. SMR produced a CH_2O product for the upper stratosphere [*Ricaud et al.*, 2007], however, these observations are not available here.

Tables 4.24.1 and **4.24.2** compile information on the availability of CH_2O measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report.

4.24.2 CH_2O evaluations: Annual zonal mean cross sections

Annual mean zonal mean cross sections are here analysed to investigate mean biases between the two available datasets. **Figure 4.24.2** shows the annual zonal mean CH_2O climatologies for MIPAS (2002-2004) and ACE-FTS (2004-2009), along with the differences of ACE-FTS to the MIM. Note there is an overlap of only one month (March 2004) between the instruments.

The MIM in the upper stratosphere (here defined by MIPAS only) shows a clear maximum in the annual mean cross section with values that are higher than in the tropical upper troposphere. The strong maximum in CH_2O found in the tropics is present all year and is consistent with model calculations [see also *Ricaud et al.*, 2007]. In the lower

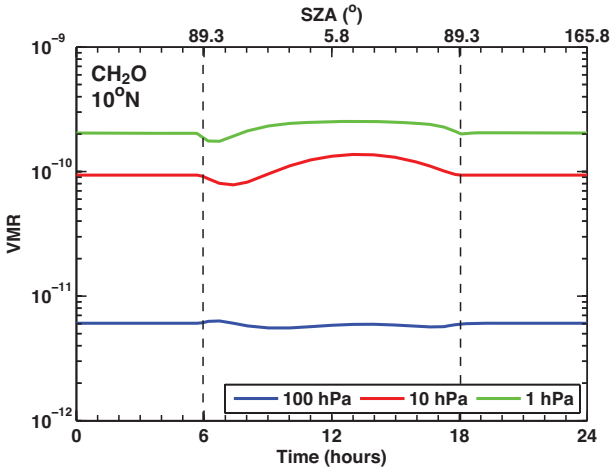


Figure 4.24.1: Diurnal cycle of CH_2O. *CH_2O variations as a function of LST or SZA are shown at 10°N at 1, 10 and 100 hPa for March 15. The diurnal cycle is derived using a chemical box model [McLinden et al., 2010].*

stratosphere, MIPAS shows minima in CH_2O located above the subtropical jets of each hemisphere. These minima are not seen in the ACE-FTS annual mean climatology, which instead shows tropopause-following isopleths. MIPAS averaging kernels [*Steck et al.*, 2008] indicate that CH_2O measurements below 18 km do not contain substantial altitude-resolved information in the UTLS. Thus the differences between MIPAS and ACE-FTS resulting from a comparison without application of averaging kernels should not be over-interpreted. In the tropical and mid-latitudinal lower stratosphere, ACE-FTS exhibits mostly larger mixing ratios (more than 100%) than MIPAS, except in the subtropical jet regions where the ACE-FTS measurements indicate lower values than MIPAS. At higher latitudes and at the upper end of the observational range, ACE-FTS measurements show also lower values than MIPAS. In the UTLS, inter-annually varying sources of CH_2O may contribute to the differences seen between the two instruments.

4.24.3 CH_2O evaluations: Meridional profiles

Meridional profile comparisons are shown in **Figure 4.24.3** and illustrate the details in the differences between the MIPAS

Table 4.24.1: Available CH_2O 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.*

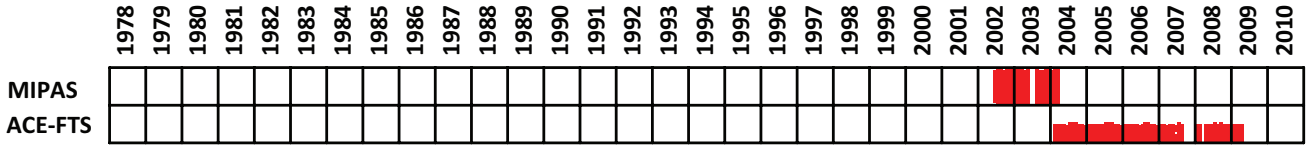


Table 4.24.2: Time period, vertical range, vertical resolution, references and other comments for CH_2O measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
ACE-FTS V2.2	Mar 04 –	5 – 25 km (500 – 25 hPa)	3 – 4 km	*Dufour et al.*, 2009	Research product from v2.2 used here
MIPAS V2	Jul 02 – Mar 04	10 – 60 km (250 – 0.2 hPa)	11 km	*Steck et al.*, 2008	Data only in high spectral resolution mode available

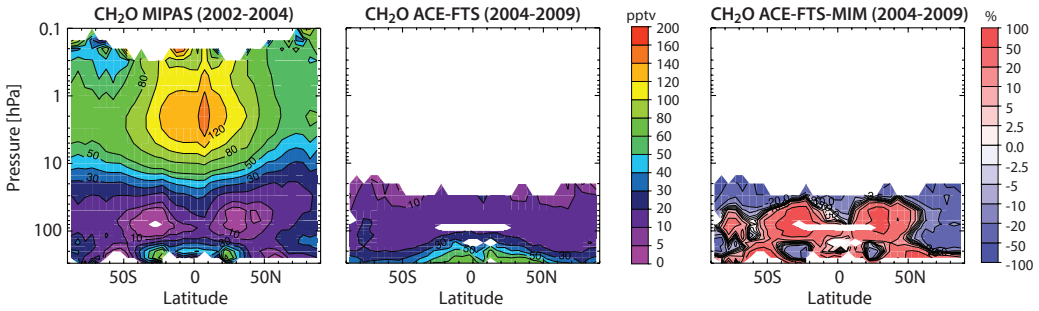


Figure 4.24.2: Cross sections of annual zonal mean CH_2O. *Annual zonal mean CH_2O cross sections are shown for MIPAS (2002-2004) and ACE-FTS (2004-2009), as well as the difference of ACE-FTS from the MIM.*

and ACE-FTS annual mean climatologies on a monthly basis. Note again that due to the low altitude resolution of MIPAS for this trace gas species and altitude range, the differences between MIPAS and ACE-FTS should be interpreted with care.

The measurements show smaller (or based on the SEM statistically less significant) differences in the SH than in the NH high latitudes, similar to the results found by *Steck et al.* [2008] for higher altitudes in comparisons with SMR. Relative differences in the NH between the two instruments are larger for March (up to 100%) than for September (mostly smaller than 50%). This result may be partially explained by higher interannual variability at polar latitudes during winter/spring than summer/autumn months together with the fact that ACE-FTS and MIPAS monthly means were not averaged over the same years. In contrast to the annual mean and March evaluations, ACE-FTS shows positive differences from the MIM at higher latitudes in September that are however statistically less significant.

4.24.4 Seasonality in CH_2O

Monthly zonal mean cross sections are compared for January, April, July, and October in order to illustrate the seasonal cycle found in CH_2O and to investigate the consistency of inter-instrument differences throughout the year in more detail (**Figure 4.24.4**). The monthly zonal mean differences are generally consistent with those derived from the annual mean evaluation, although the differences show somewhat more varying structures.

For a focus on seasonality in CH_2O in the MS and US, vertical profiles of MIPAS data are shown in **Figure 4.24.5** for

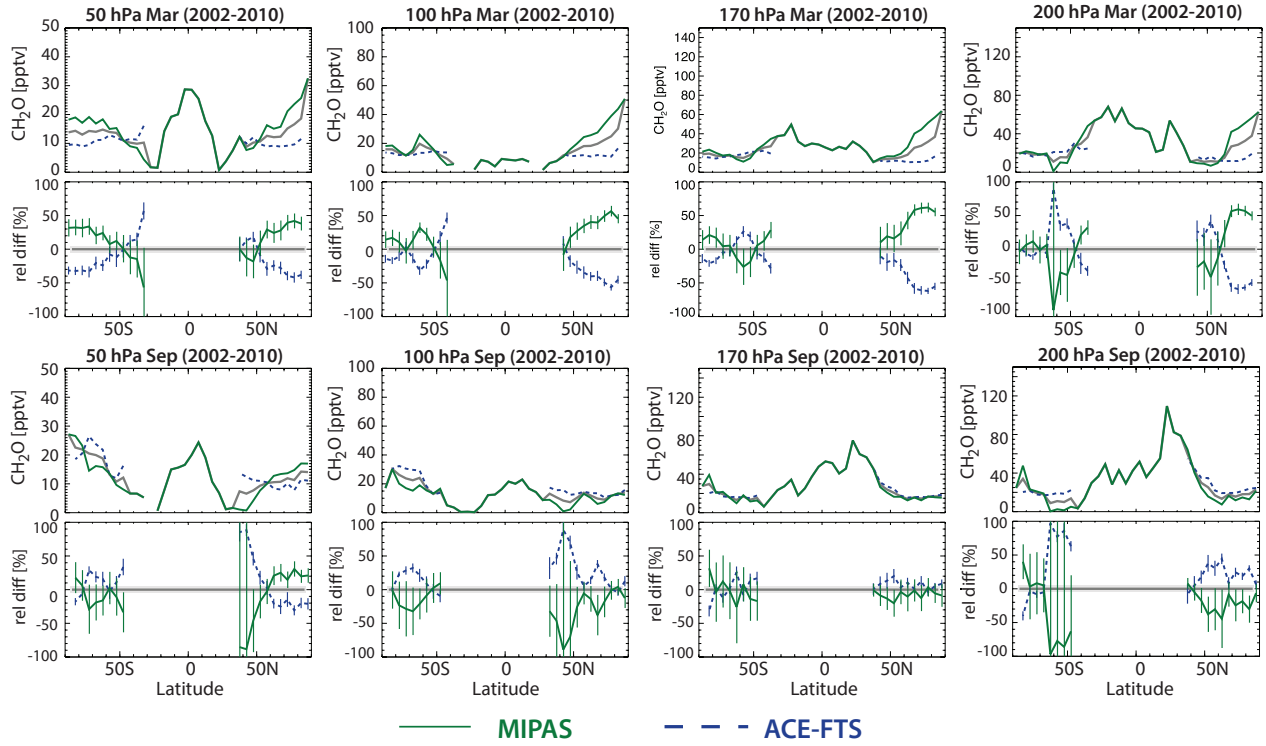


Figure 4.24.3: Meridional profiles of zonal monthly mean CH_2O. *Meridional zonal monthly mean CH_2O profiles for March (upper panels) and September (lower panels) are shown together with their differences from the MIM at 50, 100, 170, and 200 hPa. Error bars in the difference plots indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.*

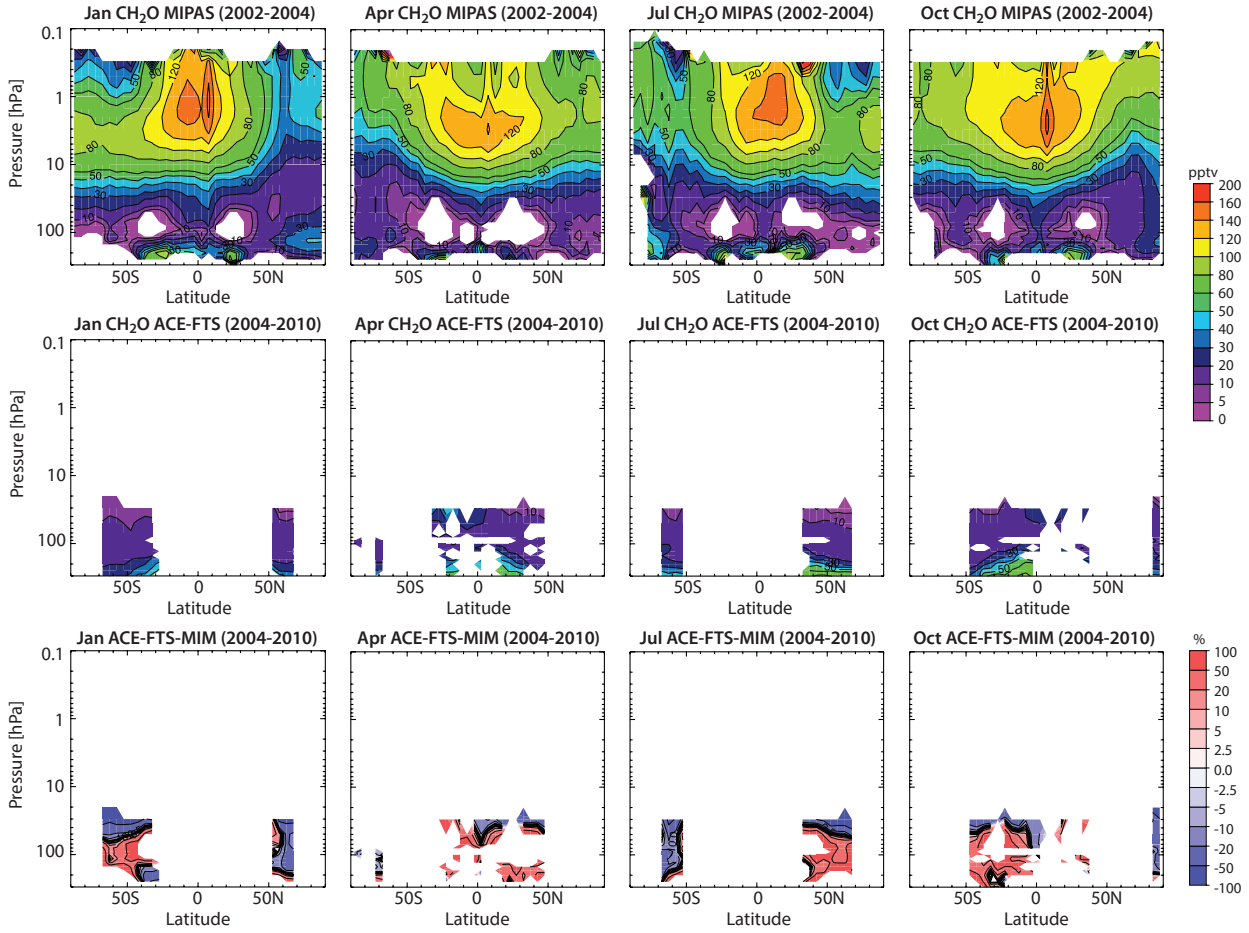


Figure 4.24.4: Monthly zonal mean cross sections of CH_2O. *Monthly zonal mean cross sections for MIPAS (2002-2004; upper row) and ACE-FTS (2004-2009; middle row) are shown for January, April, July, and October. Also shown (lower row) are the relative differences of ACE-FTS to the MIM.*

different latitudes. CH_2O in the extratropical LS and MS exhibits lowest values during autumn and winter. Maxima in extratropical CH_2O mixing ratios are found in the US during spring and summer when available sunlight triggers the production of CH_2O.

4.24.5 Summary and conclusions: CH_2O

CH_2O climatologies are available in the SPARC Data Initiative from MIPAS and ACE-FTS. The measurements were obtained in different years, with ACE-FTS focusing on

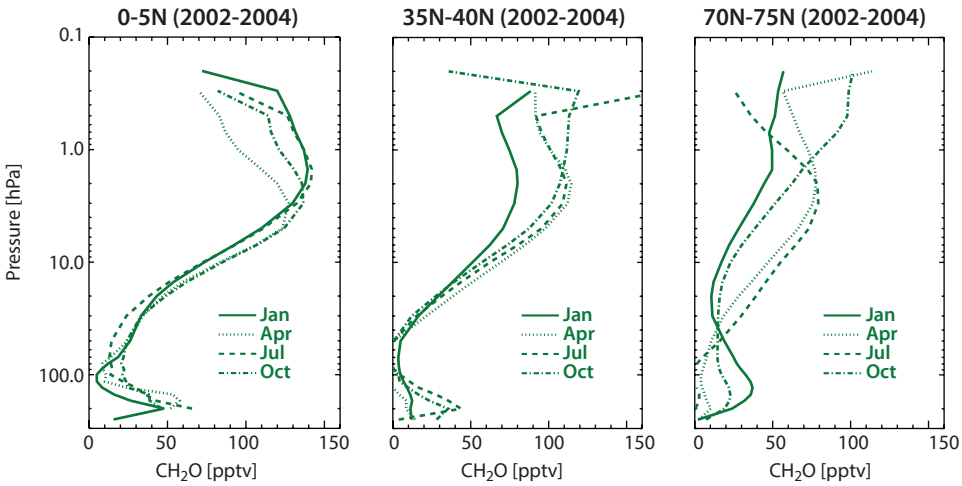


Figure 4.24.5: Altitude profiles of CH_2O. *Zonal monthly mean altitude profiles from MIPAS are shown for three different latitude bands in the tropics (0°N-5°N), at mid-latitudes (35°N-40°N) and in polar regions (70°N-75°N) for January, April, July and October.*

the troposphere/lower stratosphere, while MIPAS obtained measurements into the upper stratosphere. Differences in regions of overlap are relatively large and reach from 40% to over 100%. Largest differences are found in the subtropical LS and attributable to the relatively low vertical resolution of MIPAS, which is not adequate to resolve the distinct structures in CH_2O tracer fields around the tropopause and jet regions. In general, ACE-FTS shows mostly higher values than MIPAS with the exception of the subtropical jet regions and towards the upper limit of its measurement range.

4.25 Acetonitrile - CH_3CN

Acetonitrile (CH_3CN) is an organic trace gas in the atmosphere, which is relatively long-lived (lifetime of around 20 years at 20 km). Its main source is biomass burning [*Arijs and Brasseur*, 1986], with ocean uptake believed to be the dominant sink of the constituent [*de Gouw et al.*, 2003]. Chemical loss occurs with the hydroxyl radical (OH), however, this reaction is known to be rather slow. Tropospheric background levels are between 50 and 200 pptv [*de Gouw et al.*, 2003], while greatly enhanced values are found in biomass burning regions. Early measurements from UARS-MLS pointed to the possible existence of an additional low latitude stratospheric source of CH_3CN between 30-10 hPa (25-30 km) [*Livesey et al.*, 2001], which has not been fully explained.

4.25.1 Availability of CH_3CN measurements

Measurements of CH_3CN are available to the SPARC Data Initiative from one instrument only, SMILES, and hence only cover a few months. The first satellite-based global CH_3CN observations were made by UARS-MLS [*Livesey et al.*, 2001; 2004], and more CH_3CN products/analyses have now become available. ACE-FTS offers also measurements of global, lower stratospheric CH_3CN, which are described by *Harrison and Bernath* [2013]. While Aura-MLS measurements of CH_3CN exist, this noisy product has not been sufficiently characterised and some high biases appear to exist in the lower stratosphere [*Livesey et al.*, 2013]. Significant enhancements in lower stratospheric CH_3CN, as well as CO and HCN, were observed by Aura-MLS in the Southern Hemisphere, after the large Australian bush fires in February, 2009 [*Pumphrey et al.*, 2011].

Tables 4.25.1 and **4.25.2** compile information on the availability of CH_3CN measurements, including time period, height range, vertical resolution, and references relevant for the data product used in this report.

4.25.2 CH_3CN evaluations: Zonal mean cross sections

Zonal monthly mean cross sections are shown in **Figure 4.25.1** for selected months over the SMILES mission period (2009-2010). The cross sections reveal a strong maximum between 10 and 2 hPa, located in the SH tropics just off the equator. This maximum is at somewhat higher altitudes in the atmosphere than that identified in the early UARS-MLS measurements [*Livesey et al.*, 2001]. However, mixing ratios are similar to those found in these early measurements with maximum values between 50-60 pptv. ACE-FTS has an observational range that is limited to the lower stratosphere, and shows no indications of a tropical maximum at these altitudes [*Harrison and Bernath*, 2013]. Rather, it shows values decreasing from about 150 ppt at 11.5 km to less than 40 ppt at 25.5–29.5 km, the upper limit of its observational range.

No such maximum is seen in the extra-tropics as also illustrated in **Figure 4.25.2**. The vertical profiles reveal only a small seasonality in the tropics, increasing towards higher latitudes. At 45°N-50°N, lowest values are found in early spring, and highest in early winter at 10 hPa, although the full amplitude of the seasonal cycle cannot be determined due to the limited observation period.

4.25.3 Summary and conclusions: CH_3CN

CH_3CN climatologies are available to the SPARC Data Initiative from the SMILES instrument only. The monthly zonal mean CH_3CN climatologies from SMILES have limited temporal and vertical coverage, but reveal a maximum in the tropical upper stratosphere. Monthly profiles also indicate some seasonality that is increasing towards higher latitudes. No comparisons have been made with other instruments, however CH_3CN mixing ratios are comparable with those found in the literature on early UARS-MLS observations [*Livesey et al.*, 2001]. CH_3CN observations are now also available from Aura-MLS [*Livesey et al.*, 2013] and

Table 4.25.1: Available CH_3CN measurement record 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.*

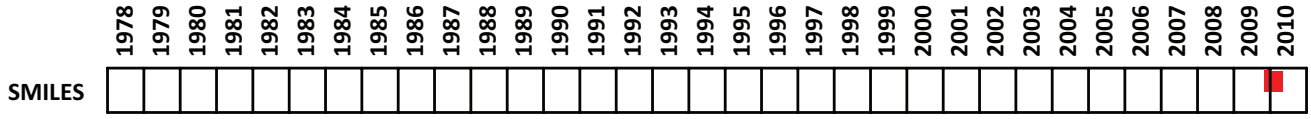


Table 4.25.2: Time period, vertical range, vertical resolution, references and other comments for CH_3CN measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
SMILES V2.1.5	Nov 2009 – Apr 2010	16 – 0.1hPa	3 – 5 km	See **http://smiles.nict.go.jp/index-e.html** for information	Here, the total CH_3CN product is used

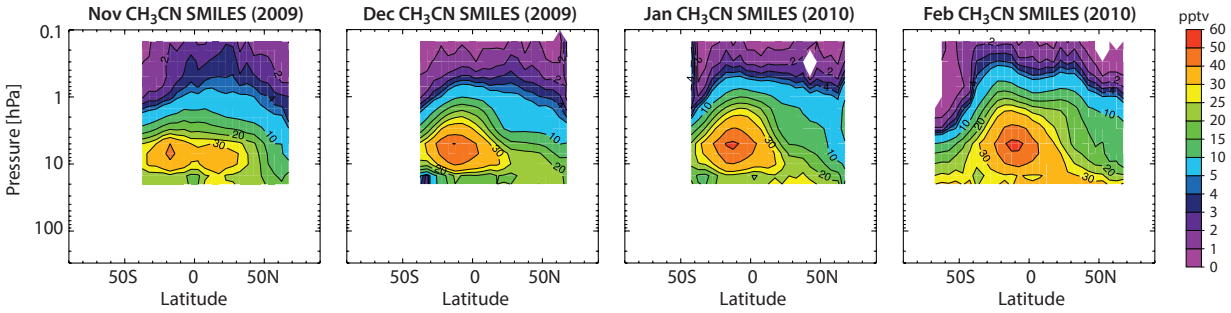


Figure 4.25.1: Cross sections of monthly zonal mean CH_3CN for 2009/2010. *Monthly zonal mean CH_3CN cross sections for November and December 2009, and January and February 2010 as obtained from SMILES observations.*

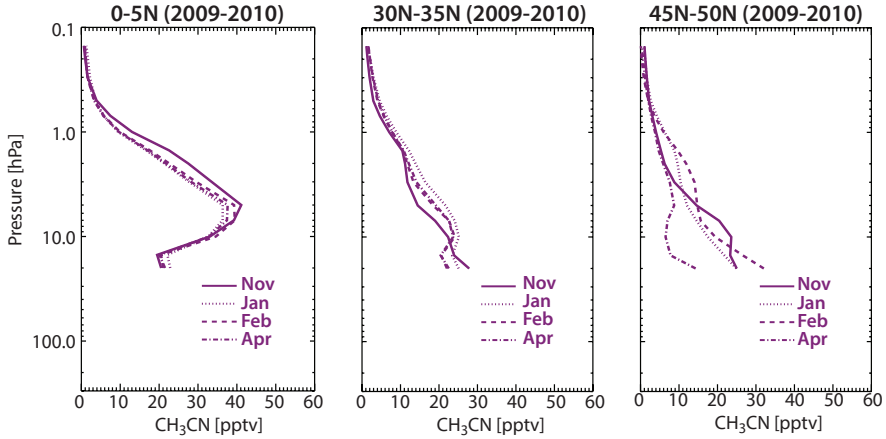


Figure 4.25.2: Vertical profiles of CH_3CN. *Zonal mean vertical profiles of CH_3CN from SMILES are shown for different months and latitude bands (equator, subtropics, and mid-latitudes).*

ACE-FTS (although these measurements are limited to the lower stratosphere)[*Harrison and Bernath*, 2013].

4.26 Aerosol

Aerosol has both natural and anthropogenic sources. Stratospheric aerosol consists primarily of liquid hydrated sulfuric acid droplets with an effective radius in the submicrometer range [*Carslaw et al.,* 1997]. The dominant source gases of stratospheric aerosol are (in order of importance) OCS, SO_2, CS_2 and Di-Methyl-Sulfide (DMS), entering the stratosphere primarily in the tropics [*SPARC*, 2006]. Natural sources of these precursors are ocean spray and desert dust, biological activity in oceans, meteoritic material, and volcanoes, the latter being capable of injecting material directly into the stratosphere in both the tropics and extra-tropics. The most important human source is fossil fuel combustion (in particular from air traffic) and biomass burning. Stratospheric aerosol has a key role in chemistry and the radiation budget of the atmosphere [*McCormick et al.*, 1995]. It offers a surface for heterogeneous reactions, controlling abundances of stratospheric NO_x [*Angell et al.*, 1985; *Hofmann and Solomon*, 1989]. NO_x in turn helps determining abundances of ClO_x and HO_x, species that are involved in stratospheric ozone depletion [*Wennberg et al.*, 1994; *Solomon et al.*, 1996]. Direct radiative forcing caused by increased aerosol loadings after volcanic eruptions leads to stratospheric warming [*Labitzke and McCormick*, 1992] and tropospheric cooling [*Manabe and Wetherald*, 1967]. Aerosol also serves as cloud condensation nuclei in the upper troposphere and the polar vortex regions [*Laaksonen et al.*, 2000], leading to an indirect radiative forcing effect. The stratospheric aerosol layer [*Junge et al.*, 1961], also called the Junge layer, has been shown to be highly variable resulting from both major and minor volcanic eruptions that reach the stratosphere [*Vernier et al.*, 2011].

4.26.1 Availability of aerosol measurements

Measurements of aerosol since 1984 are available to the SPARC Data Initiative, with the longest time series of 20 years from SAGE II ending in 2005. HALOE provides aerosol measurements from 1991 to 2005. Aerosol measurements over shorter time periods and with more limited latitude coverage are also available from SAGE III, POAM II, and POAM III. The newer generation of satellite instruments includes data from OSIRIS, GOMOS (two data products, AERGOM and v6.01, the latter hereafter simply referred to as GOMOS), and SCIAMACHY with measurements starting in the early 2000s. The data presented in this report are representative but not comprehensive. A number of valuable aerosol products can also be obtained from SAGE I [*Chu and McCormick*, 1979], SAM II [*McCormick et al.*, 1979; 1981], CLAES [*Roche et al.*, 1993], SME [*Eparvier et al.*, 1994; *Rusch et al.*, 1994], ORA [*Fussen et al.*, 2001], ILAS and ILAS II [*Sasano*, 2002], ISAMS [*Taylor et al.*, 1993; *Lambert et al.*, 1993], ACE imager [*Vanhellemont et al.*, 2008], HIRDLS [*Gille and Gray*, 2011] and CALIPSO [*Winker et al.*, 2003].

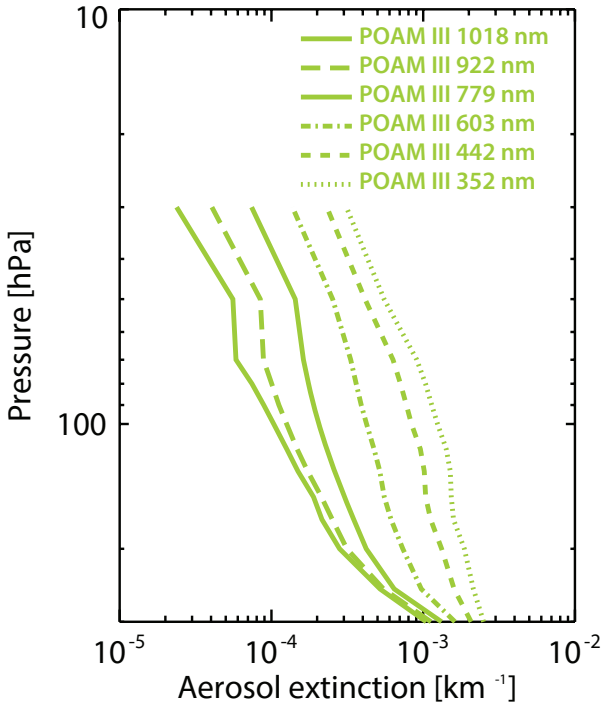


Figure 4.26.1: Aerosol extinction for April 2003. *Shown are the zonal monthly mean altitude profiles of aerosol extinction from POAM III at 60°N as derived from measurements at different wavelengths. The longer the wavelength, the smaller are the aerosol extinction values.*

The most frequently measured aerosol parameter from space that is evaluated in this report is aerosol extinction (with the unit km^{-1}). Full characterisation of aerosol on the other hand would rely on knowledge of size distribution and composition. The retrieved extinction coefficients are a measure of the attenuation of the light passing through the atmosphere due to scattering and absorption by aerosol particles. The extinction varies substantially with wavelength - longer wavelengths have smaller values as can be seen in **Figure 4.26.1**. This complicates multi-instrument comparisons, such as those included in this report, since the instruments measure over a wide range of wavelengths (see **Table 4.26.2**). HALOE products for example are retrieved at much longer wavelengths than products from other instruments, which does not allow for a direct comparison.

It should be noted that there exist fundamental differences between the aerosol extinction products from occultation and limb scattering instruments. Generally, solar, lunar, or stellar occultation sounders are able to provide a direct measurement of the atmosphere's optical depth, which can be translated into aerosol extinction without requiring assumptions on aerosol composition or size. Limb scattering instruments, on the other hand, only provide a derived quantity and require assumptions for aerosol properties in order to characterise the scattering properties of the measured light. A clear benefit of the limb scattering instruments is, however, much better geographical coverage. Note that the two SCIAMACHY products at 470 nm and 750 nm are not independent climatologies. The ratio between the extinction coefficients at these wavelengths is prescribed by the assumed particle size distribution and is not changed by the retrieval. The results at 750 nm are considered to be more reliable due to the much stronger sensitivity at this wavelength. Note also that this is essentially the same technique as used by OSIRIS.

In this report, the monthly zonal mean climatologies of aerosol extinction will be validated using two different approaches. The first approach is to compare aerosol climatologies that are derived from observations at similar wavelengths. These evaluations have to be interpreted with care, since disagreement between two instruments may be attributable to differences in the wavelengths used for the retrieval, while agreement may merely reflect the result of compensating errors. A second approach, which is introduced in more detail by *Hegglin et al.* (in preparation), applies a normalisation factor derived from each climatology during quiescent periods to the available time series. This scaling of the aerosol extinction (AE) to obtain relative anomalies follows the equation:

$$AE_{rel\_ano} = (AE/f - 1)*100$$

where *f* is a scaling factor and represents the mean aerosol loading in 2004.

This approach has the advantage that all climatologies (even when retrieved at very different wavelengths) can be compared to each other at the same time. The normalisation factor is in principle dependent on the particle size distribution of the aerosol, which is neglected here and may introduce some error in the comparison. However, this error is considered to be of second-order importance compared to other systematic inter-instrument differences including those arising from the wavelength dependency even at similar wavelengths. Only a few examples of this new validation approach will be shown here. An assessment using standard validation approaches between the earlier satellite and ground-based measurements has been provided in the *SPARC ASAP Report* [2006].

Tables **4.26.1** and **4.26.2** compile information on the availability of aerosol measurements used in this report, including time period, altitude range, vertical resolution, relevant references, and list of aerosol retrieval wavelengths.

4.26.2 Aerosol evaluations: Vertical and meridional profiles at similar wavelengths

The available aerosol climatologies are evaluated using zonal monthly means rather than annual means in order to prevent large sampling biases, which would be expected in annual mean comparisons due to the high variability found in aerosol distributions along with the limited geographical coverage solar occultation instruments offer. In the following, we compare aerosol climatologies retrieved at similar wavelengths. To this end the climatologies are classified into five categories according to the wavelength used for the retrievals:

Table 4.26.1: Available aerosol 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.*

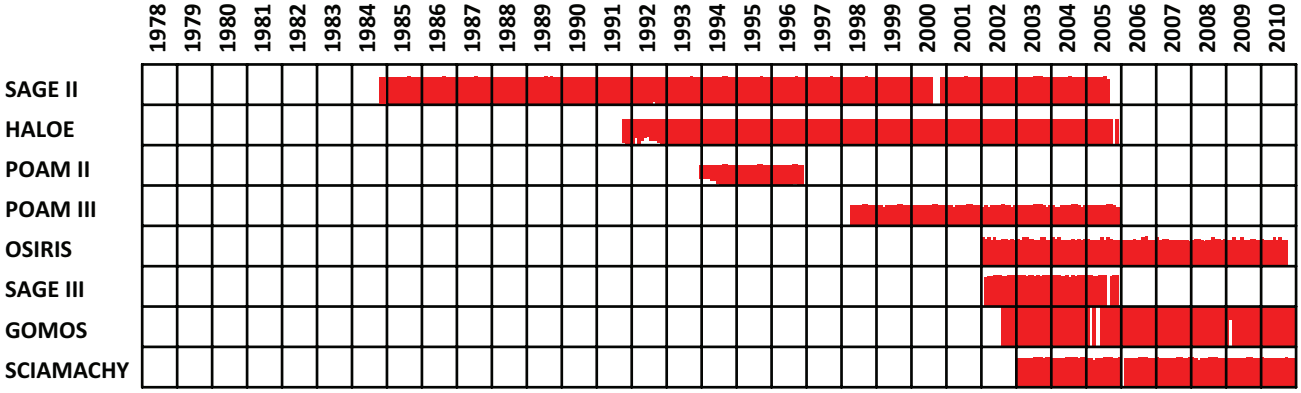


Table 4.26.2: Time period, vertical range, vertical resolution, references and other comments for aerosol extinction measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
POAM II V6.0	Oct 1993 – Nov 1996	10 – 30 km	1 – 1.5 km	*Lumpe et al.,* 1997 *Randall et al.,* 2000	Retrieved/available at 352, 442, 601, 781, 921, 1060 nm
POAM III V4.0	Apr 1998 – Dec 2005	5 – 25 km	1 – 1.5 km	*Lumpe et al.,* 2002 *Randall et al.,* 2001	Retrieved/available at 353, 442, 603, 779, 922, 1018 nm
GOMOS V6.01 AERGOM	Aug 2002 – Apr 2012	10 – 40 km	4 km	*Vanhellemont et al.,* 2010; 2016	Retrieved/available at 500 nm (v6.01), and 350, 450, 470, 500, 525, 600, 750 nm (AERGOM)
	Aug 2002 – Apr 2012	15 – 32 km	3 – 5 km	*von Savigny et al.,* 2015	Retrieved/available at 750 (470) nm
SAGE II V7.0	Jan 1985 – Aug 2005	~5 – 40 km (channel dependent)	1 km	*Damadeo et al.,* 2013	Retrieved/available at 386, 452, 525, 1020 nm
SAGE III V4.0	Feb 2002 – Dec 2005	~5 – 40 km (channel dependent)	1 km	*Thomason,* 2010	Retrieved/available at 520, 755, 1020 nm
HALOE V19	Oct 1991 – Nov 2005	*dependent on aerosol loading	~2 km	*Hervig et al., 1996b Hervig and Deshler, 2002*	Retrieved/available at 2450, 3400, 3460, 5260 nm
OSIRIS V5.07	Nov 2001 – present	15 – 30 km	2.2 km	*Bourassa et al.,* 2007; 2012	Retrieved/available at 750 nm

- **~350 nm:** 350-390 nm
- **~450 nm:** 440-470 nm
- **~550 nm:** 500-610 nm
- **~750 nm:** 750-780 nm
- **~1050 nm:** 1010-1060 nm

Note that no such comparisons could be made for HALOE, since it derives aerosol extinction only at longer wavelengths (2450-5260 nm).

SAGE II, POAM III, and AERGOM (2003, wavelenghs ~350 nm)

Figure 4.26.2 shows the zonal monthly mean cross sections for aerosol extinction climatologies are available from measurements retrieved at wavelengths around 350 nm. Aerosol extinction climatologies from SAGE II are retrieved at 386 nm, AERGOM at 350 nm, and POAM III at 353 nm. Note the different vertical range (300-10 hPa) used in the following figures when compared to previous, monthly zonal mean cross sections in this report.

Figure 4.26.2 reveals the general features of the stratospheric aerosol layer, with high aerosol extinction values in the tropical UT that leak into the tropical LS, and mostly tropopause following aerosol isolines similar as expected from a longer-lived tropospheric source gas. Due to higher sensitivity to temperature, chemistry, and transport and its effects on aerosol, variability however is much higher than for other long-lived chemical trace gases. Towards 10 hPa, the atmosphere is becoming much 'cleaner' and the instruments start to reach their detection limit.

AERGOM (SAGE II) shows higher (lower) values than the MIM by around 10% in the middle stratosphere, and by around 50% in the UT and LS. Note that in the following the MIM has been constructed using all products shown in

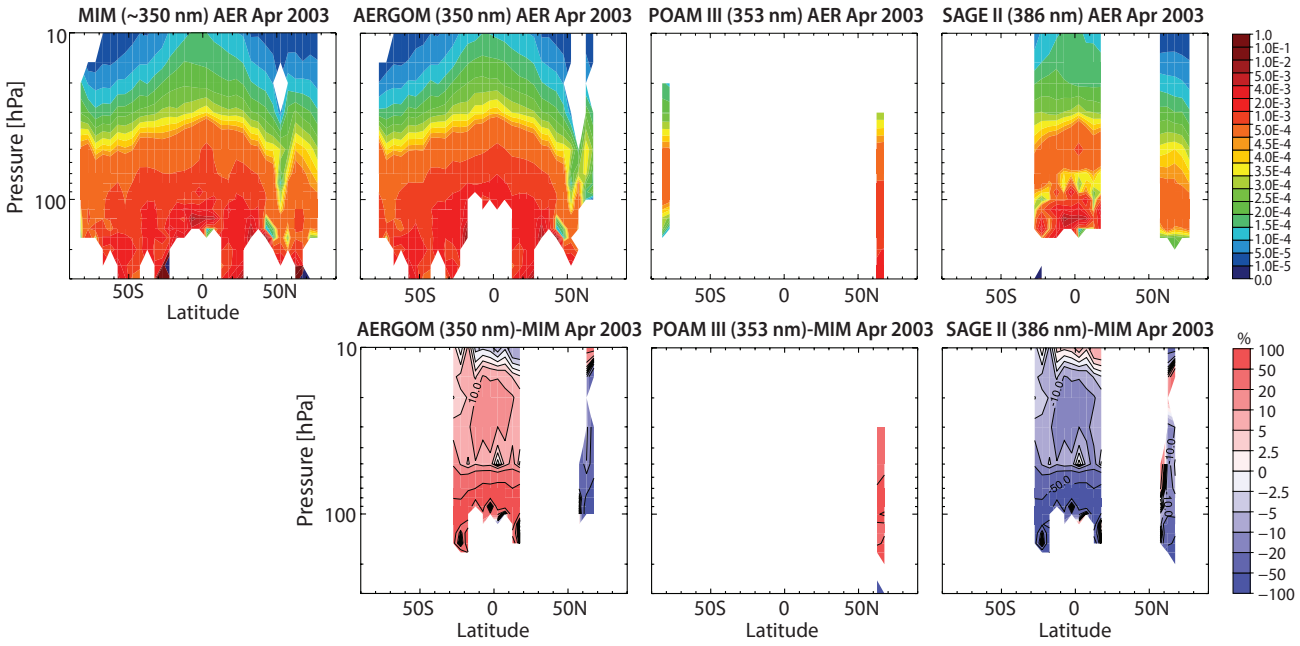


Figure 4.26.2: Cross sections of monthly zonal mean aerosol extinction (retrieved at around 350 nm) for April 2003. *The cross sections are shown for the MIM, AERGOM (at 350 nm), POAM III (at 353 nm), and SAGE II (at 386 nm)(upper row). Also shown are the relative differences of each instrument from the MIM (lower row). Note, the aerosol extinction climatologies are ordered according to rising retrieval wavelength.*

each figure, even though a particular instrument may offer more than one product in the considered wavelength range. The differences can be at least partially explained by the difference in wavelengths the two instruments use for the aerosol extinction retrieval (with SAGE II retrieving at longer wavelengths that yield lower extinction values). POAM III however, which retrieves the aerosol at a similar wavelength as AERGOM, indicates more positive values than both of these other instruments.

SAGE II, POAM III, AERGOM, and SCIAMACHY (2003, wavelengths ~450 nm)

Figure 4.26.3a and **b** show the zonal monthly mean cross sections for aerosol extinction climatologies and their differences from the MIM as deduced from measurements at wavelengths around 450 nm. Aerosol extinction climatologies are available from SAGE II retrieved at 452 nm, AERGOM at 450 and 470 nm, SCIAMACHY at 470 nm, and POAM III at 442 nm. Note that the SCIAMACHY climatologies at 470 nm have a very low information content and basically just represent a scaled 750 nm product.

AERGOM and SAGE II retrieve at very similar wavelengths in this category with their 450 and 452 nm products, respectively. Nevertheless, they show relatively large differences comparable to those seen in the previous comparison. Differences from the MIM are around ±10%, with AERGOM on the positive side and SAGE II on the negative side. Note that the AERGOM product retrieved at 470 nm still shows higher values than SAGE II (against expectations from wavelength considerations), indicating a real negative bias in SAGE II (or high bias in AERGOM). SCIAMACHY, which has a product at 470 nm as well, shows also quite large negative differences from the MIM through most of the tropical/subtropical MS. These are partially expected due to the wavelength-dependency. However, when compared to the AERGOM product at 470 nm, the differences are still much larger. In the UT and LS, SCIAMACHY shows strong positive differences from the MIM, attributable to the fact that here data were used to produce the zonal means that were not filtered for cloud occurrence. The only overlap with POAM III is seen in the Northern Hemisphere, where the POAM III retrieval shows positive deviations from the MIM, in best agreement with SAGE II and SCIAMACHY. The relatively strong negative differences from the MIM in AERGOM at these higher latitudes indicate that the products have likely a latitudinal structure in their biases, potentially arising from sampling issues that are more severe during the late winter/spring season where the polar vortex may have generated large and persisting horizontal gradients in the aerosol fields.

SAGE II, SAGE III, POAM III, GOMOS, and AERGOM (2003, wavelengths ~550 nm)

Figure 4.26.4a shows the zonal monthly mean cross sections for aerosol extinction climatologies for measurements retrieved at wavelengths around 550 nm. Available are GOMOS v6.01 aerosol extinction climatologies retrieved at 500 nm, and AERGOM at 500, 525, and 600 nm, SAGE II at 525 nm, SAGE III at 520 nm, and POAM III at 603 nm. Differences between single instruments and the MIM are shown in **Figure 4.26.4b**.

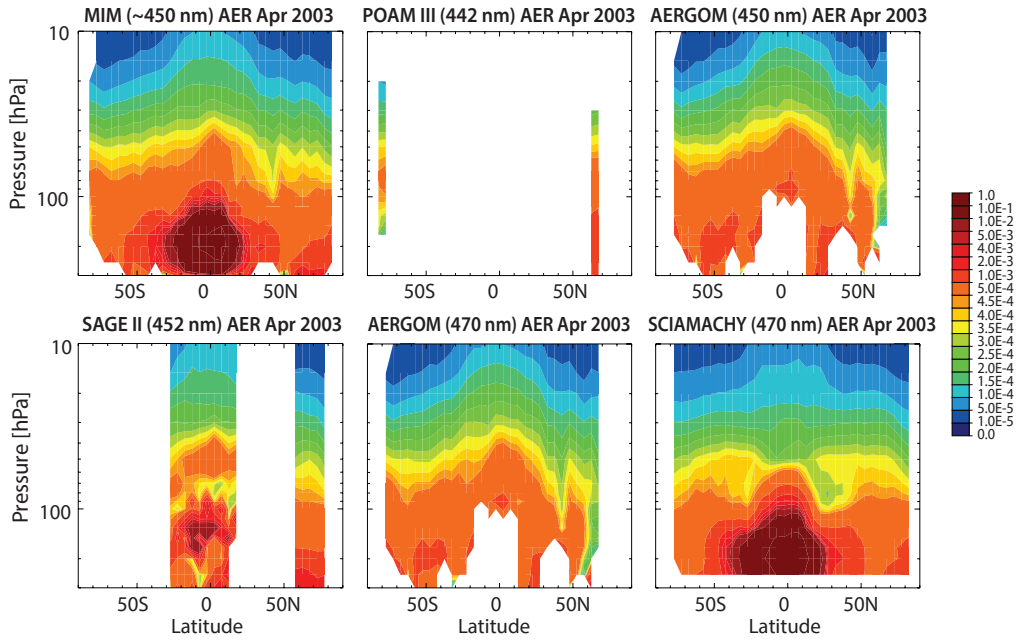


Figure 4.26.3a: Cross sections of monthly zonal mean aerosol extinction (retrieved at around 450 nm) for April 2003. *The cross sections are shown for the MIM, POAM III (442 nm), AERGOM (450 nm) (upper row), SAGE II (452 nm), AERGOM (470 nm) , and SCIAMACHY (470 nm) (lower row).*

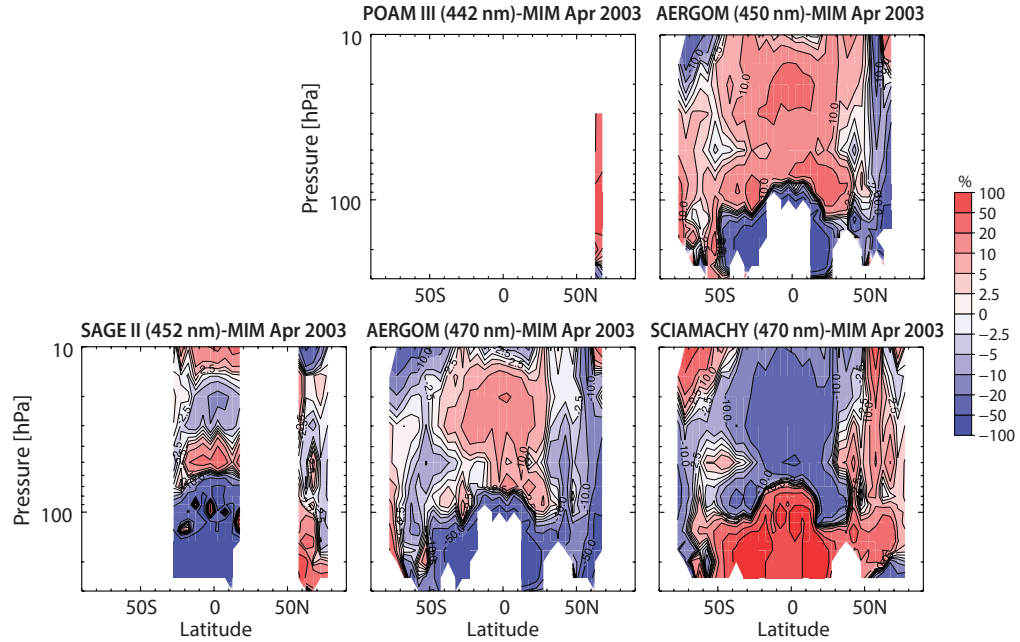


Figure 4.26.3b: Cross sections of monthly zonal mean differences in aerosol extinction (retrieved at around 450 nm) for April 2003*. Monthly zonal mean relative differences between the individual instruments and the MIM are shown.*

A strong wavelength-dependency is obvious in this comparison category with large positive deviations from the MIM for measurements obtained at 500 nm (GOMOS and AERGOM) and large negative deviations from the MIM at 600 nm (AERGOM). POAM III shows in contrast only moderate and mostly positive differences from the MIM, despite measuring at 600 nm as well. The other instruments retrieving at similar wavelengths show reasonably good agreement with differences of mostly below ±10% throughout the LS and MS (also SAGE III at 520 nm, and SAGE II and AERGOM at 525 nm). Only in the tropical upper troposphere, differences increase to above ±20%.

SAGE III, POAM III, OSIRIS, SCIAMACHY, and AERGOM (2003, wavelengths around 750)

Figure 4.26.5a shows the zonal monthly mean cross sections for aerosol extinction climatologies from measurements retrieved at wavelengths around 750 nm. OSIRIS retrieves its aerosol extinction at 750 nm, SCIAMACHY at 750 nm, AERGOM at 750 nm, SAGE III at 755 nm, and POAM II at 779 nm. OSIRIS, SCIAMACHY, and AERGOM fields are hence directly comparable, and even the wavelength used for SAGE III retrievals is close enough so that derived

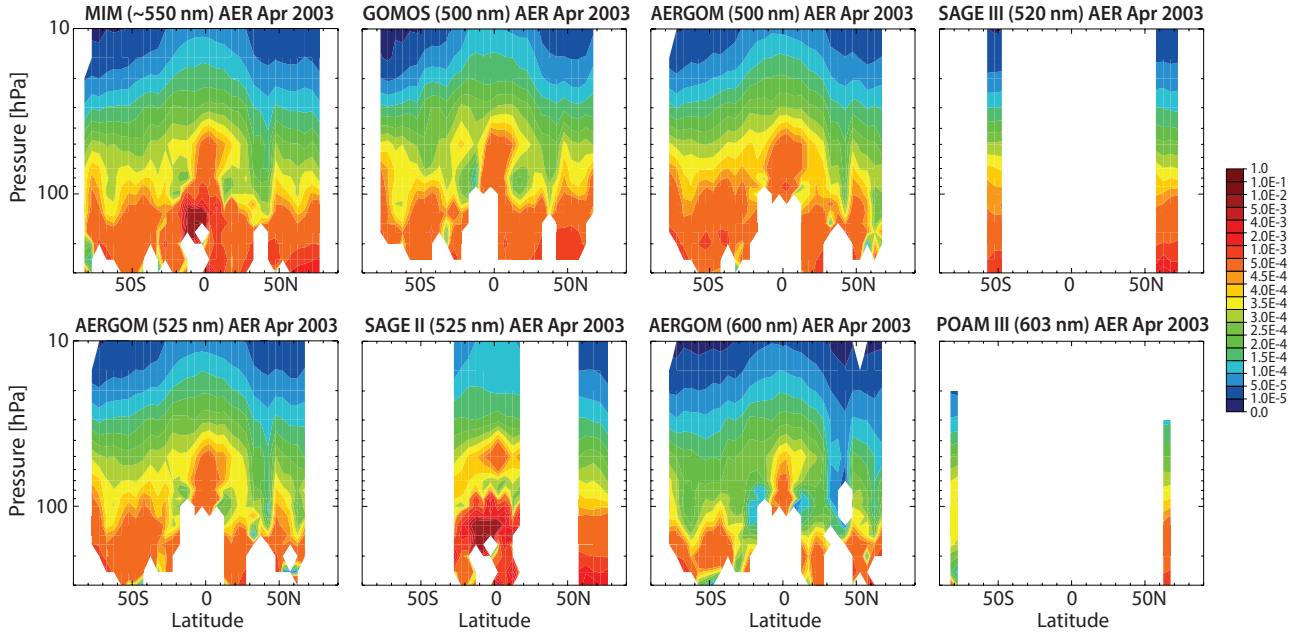


Figure 4.26.4a: Cross sections of monthly zonal mean aerosol extinction (retrieved at around 500 nm) for April 2003. *The cross sections are shown for the MIM (upper left panel), GOMOS (500 nm), AERGOM (500 nm), SAGE III (520 nm), AERGOM (525 nm), SAGE II (525 nm), AERGOM (600 nm), and POAM III (603 nm).*

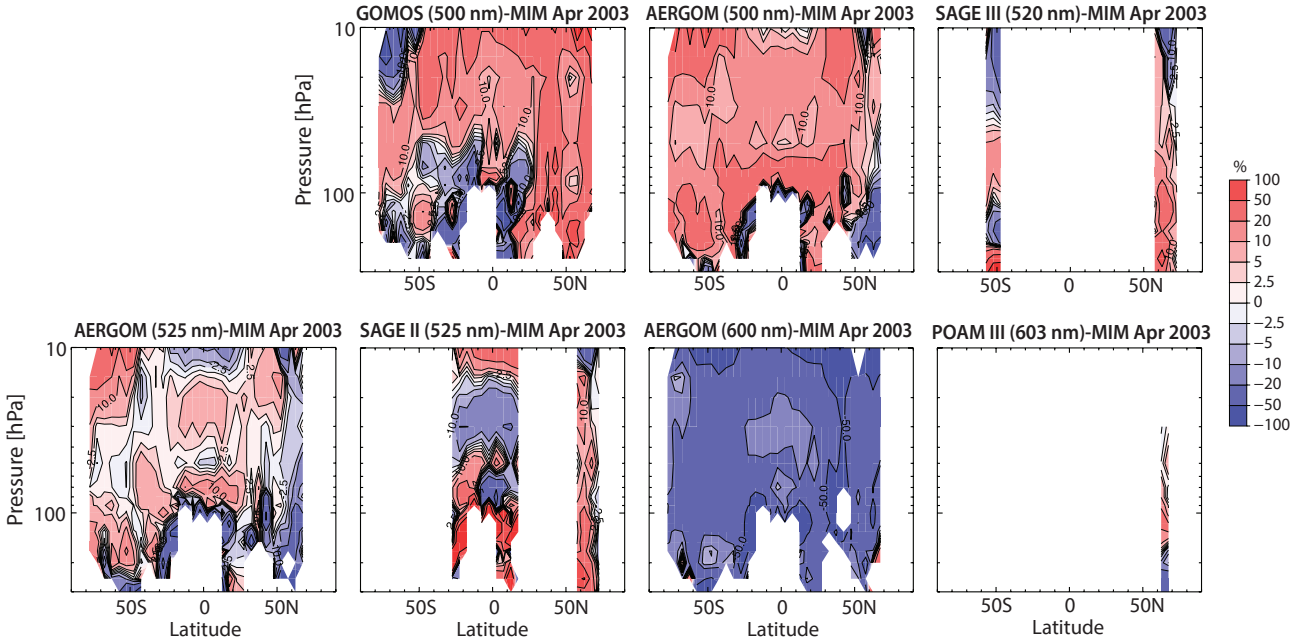


Figure 4.26.4b: Cross sections of monthly zonal mean differences in aerosol extinction (retrieved at around 550 nm) for April 2003. *Monthly zonal mean relative differences between the individual instruments and the MIM are shown.*

differences can be interpreted as real instrument differences.

A distinct feature of high aerosol extinction values is seen by all instruments in the Southern Hemisphere polar vortex (here early spring is shown), which is caused by the presence of polar stratospheric clouds [*e.g.*, *Benson et al.*, 2006].

Through most of the tropical and mid-latitude middle stratosphere, OSIRIS and SCIAMACHY show good to very good agreement (within ±5-10%). Relative differences increase towards the LS (±20%), the UT (±50%), and towards the polar region of the winter hemisphere (here the SH) (±50%) where dynamical variability and aerosol extinction values are larger. SAGE III compares very well with the MIM (mostly within ±5%), except below 100 hPa where differences from the MIM increase slightly. POAM III shows positive differences in the SH of 20% (however here agreeing better with SCIAMACHY), and a negative bias in the NH of up to 20% in the MS, which is partially expected given the higher wavelength its product is retrieved at. Differences from the MIM increase towards the LS where negative biases of up to 50% are found. The AERGOM product at 750 nm was expected to yield less accurate results and hence was excluded from the calculation of the MIM. Indeed, the differences from the MIM are as large as -50% (and worse) through most of the stratosphere, with equally large positive differences in the tropical lower stratosphere.

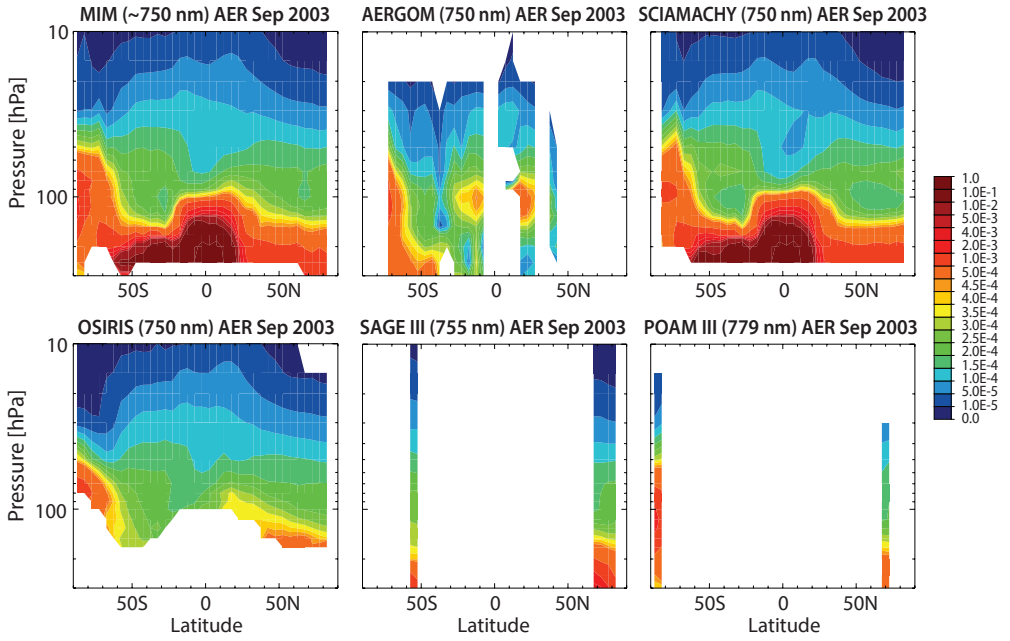


Figure 4.26.5a: Cross sections of monthly zonal mean aerosol extinction (retrieved at around 750 nm) for September 2003. *The cross sections are shown for the MIM (upper left panel), and for AERGOM (750 nm), SCIAMACHY (750 nm), OSIRIS (750 nm), SAGE III (755 nm), and POAM III (779 nm). Note that AERGOM is excluded from the MIM.*

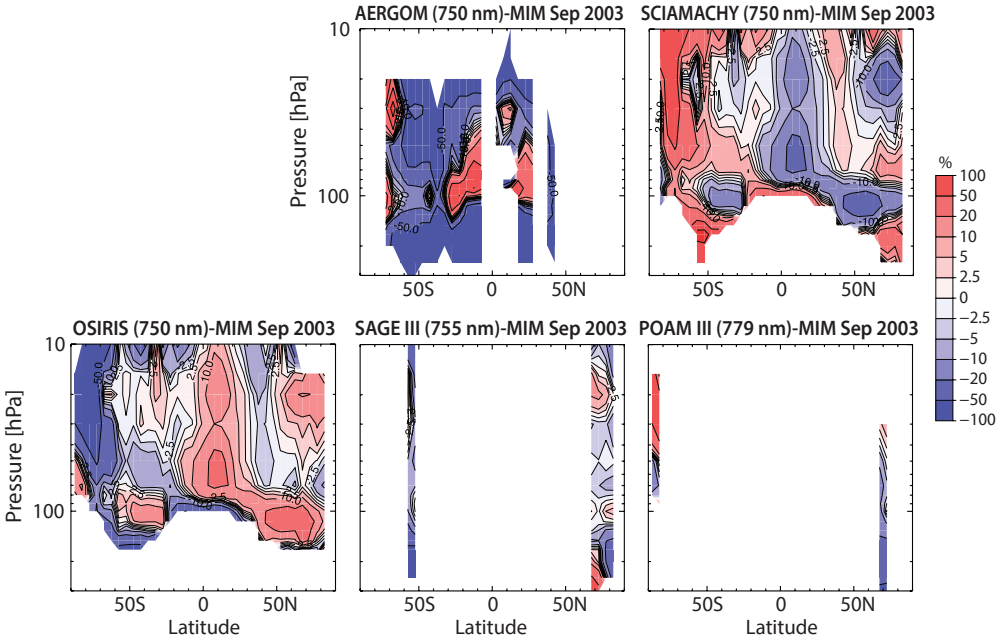


Figure 4.26.5b: Cross sections of monthly zonal mean differences in aerosol extinction (retrieved at around 750 nm) for September 2003. *Monthly zonal mean relative differences between the individual instruments and the MIM are shown.*

SAGE II, SAGE III, and POAM III (2003, wavelengths around 1050 nm)

Figure 4.26.6 finally shows the zonal monthly mean cross sections for aerosol extinction climatologies from measurements retrieved at wavelengths around 1050 nm. The instruments offering products retrieved in this wavelength region are SAGE II and SAGE III (with observations at 1020 nm) and POAM III (with observations at 1018 nm). All three aerosol extinction products are hence very well comparable to each other and show relative differences from the MIM of mostly less than ±10%, indicating good agreement between the climatologies.

4.26.3 Aerosol evaluations: Altitude profiles

Given the limitations of comparing aerosol extinction products retrieved at different wavelengths with each other, we here show only a few more evaluations of vertical profiles in addition to the above monthly zonal mean evaluations using this comparison approach. The examples are chosen to provide additional seasonal information and to perform comparisons between instruments that due to limited overlap with respect to geographical coverage were not considered above.

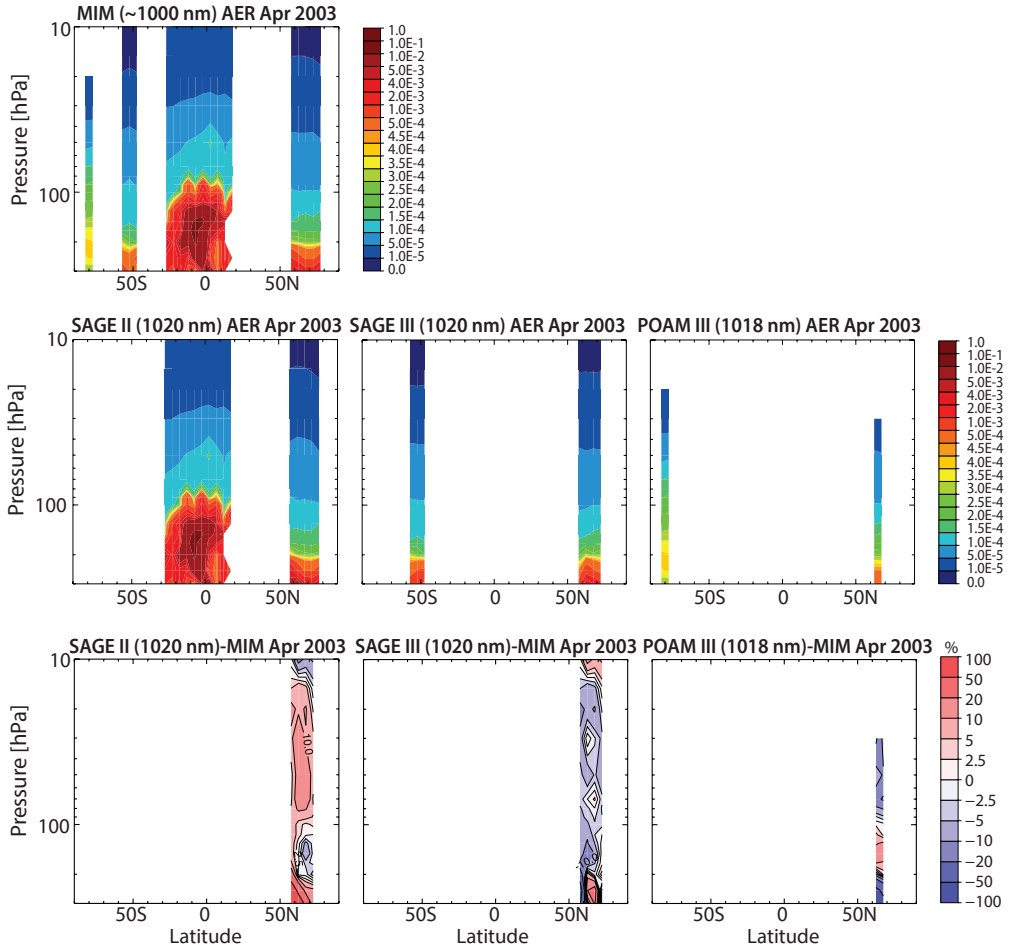


Figure 4.26.6: Cross sections of monthly zonal mean aerosol extinction (retrieved at around 1050 nm) for April 2003. *The cross sections are shown for the MIM (upper left panel), and for the instruments SAGE II, SAGE III, and POAM III (middle row). Also shown are the relative differences of each instrument to the MIM (lower row).*

Figure 4.26.7 shows a comparison between SAGE II and POAM II, with the latter not having been considered above. Two sets of products retrieved at different wavelengths are compared, SAGE II at 452 nm with POAM II at 442 nm, and SAGE II at 1020 nm with POAM II at 1060 nm. While for the first set, SAGE II should show slightly lower values than POAM II in the 450 nm comparison given above discussed wavelength dependency, SAGE II should show slightly higher values than POAM II in the 1050 nm comparison. However, this is not the case. The instruments show good agreement (between ±10%) in most examples, except for the 450 nm case in the Southern Hemisphere in February, where differences are up to ±25%. The fact that the instruments do not show the same sign in the difference from the MIM in the two hemispheres indicates a hemisphere-dependent instrument bias. These results are consistent with the POAM II/SAGE II validation analysis of *Randall et al.* [2000]. Differences were smaller in 1994, but by 1995-1996, with a cleaner atmosphere, hemispheric biases showed up that are thought to be caused by altitude registration errors in one or both datasets.

Figure 4.26.8 shows comparisons in the wavelength category of products retrieved at around 450 nm for April 2003. In the Northern Hemisphere high latitudes, the AERGOM products at 450 and 470 reveal extremely good agreement with the SAGE II product at 452 nm through most of the lower and middle stratosphere between 70 hPa and 15 hPa. While POAM III retrieved at 442 nm exhibits a high bias in the Northern Hemisphere, it shows excellent agreement with AERGOM in the Southern Hemisphere over the same altitude range. At altitudes below 70 hPa, however, AERGOM (SAGE II and POAM III) exhibit large negative (positive) differences from the MIM in the Northern Hemisphere and large positive (negative) differences from the MIM in the Southern Hemisphere, respectively.

Figure 4.26.9 shows comparisons in the wavelength category of products retrieved at around 550 nm. Good agreement between most instrument products is found throughout the LS above 100 hPa and MS below 15 hPa. Only the AERGOM product at 600 nm seems to be somewhat an outlier with much stronger negative differences from the MIM than what could be expected from the wavelength difference to other products. Even at altitudes below 100 hPa, there is generally good agreement between GOMOS, POAM, and AERGOM, while the SAGE II and III data products show rather large positive differences from the MIM. Note that the cross sections in **Figure 4.26.4b** yield a more complete picture on inter-instrument differences, given that these show varying behaviour with latitude.

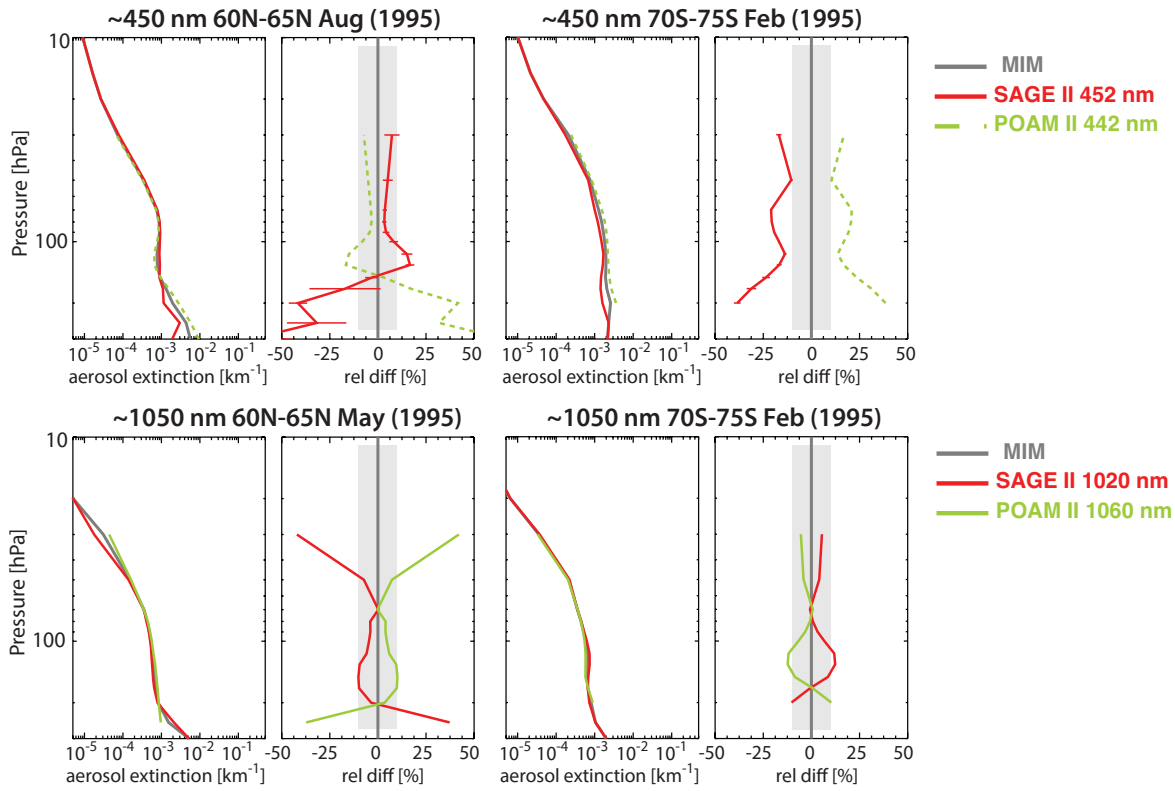


Figure 4.26.7: Vertical profiles of zonal monthly mean aerosol extinction in 1995. *The aerosol extinction profiles are shown for SAGE II and POAM II at 60°N-65°N and 70°S-75°S on the left and right, and for products retrieved at wavelengths around 450 nm and 1000 nm in upper and lower panels, respectively. The relative differences between the individual instruments and the MIM are shown in the right-hand panels. 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 10%.*

Finally, **Figure 4.26.10** shows comparisons in the wavelength category of products retrieved at around 750 nm. At these wavelengths, the instruments again show good to reasonably good agreement through most of the LS and MS above 100 hPa, except for AERGOM, which shows large negative deviations from the MIM and the other instruments. Note that it was anticipated that the 750 nm data product would yield poorer results, and AERGOM hence was excluded from the MIM calculation. SCIAMACHY shows the tendency to be higher than the other instruments below 100 hPa (possibly due to the fact that unfiltered climatologies were used in this comparison). Finally, note the remarkable agreement between SAGE III and OSIRIS.

4.26.4 Aerosol evaluations: Interannual variability

Another important aspect of instrument performance apart from the representation of the climatological mean structure is the instruments' capability to capture interannual variability. For this evaluation, we now turn to the second approach introduced above for the comparison of the aerosol extinction climatologies (see also *Hegglin et al.*, in preparation), which uses scaling factors applied to the different wavelength products to make them better comparable to each other and also to be able to include HALOE in the comparison.

The upper panel in **Figure 4.26.11** shows the original time series of aerosol extinction in the tropics at 50 hPa averaged over 20°S-20°N as derived from a set of instruments at different retrieval wavelengths. A large spread between the different time series is apparent, illustrating once more the wavelength-dependency of aerosol extinction retrievals described above (see **Figure 4.26.1**). Note that we included only three products from AERGOM (at 350, 500, and 600 nm) so to keep the figure better readable.

The lower panel in Figure **4.26.11** shows each instrument's time series now scaled following equation (1) and using the year 2004 as reference year. The applied scaling largely removes the differences between the time series and collapses the curves on top of each other. The time series derived from different wavelengths become thereby comparable to each other. At least after 1999 and during episodes of relatively low aerosol loading, the agreement between the instruments is mostly good and lies within ±10%. During episodes with higher aerosol loading, discrepancies between the datasets increase.

In particular towards the earlier years of the comparison, the differences between the time series increase to more than 300%. In the mid to late 1990s, SAGE II at 1020 nm is a clear outlier, showing a different relaxation timescale towards background aerosol extinctions than its other wavelength products and also HALOE. However, this SAGE II product shows good agreement with the other aerosol extinction products available from 2002 onwards, indicating that the enhanced aerosol loading and changing aerosol

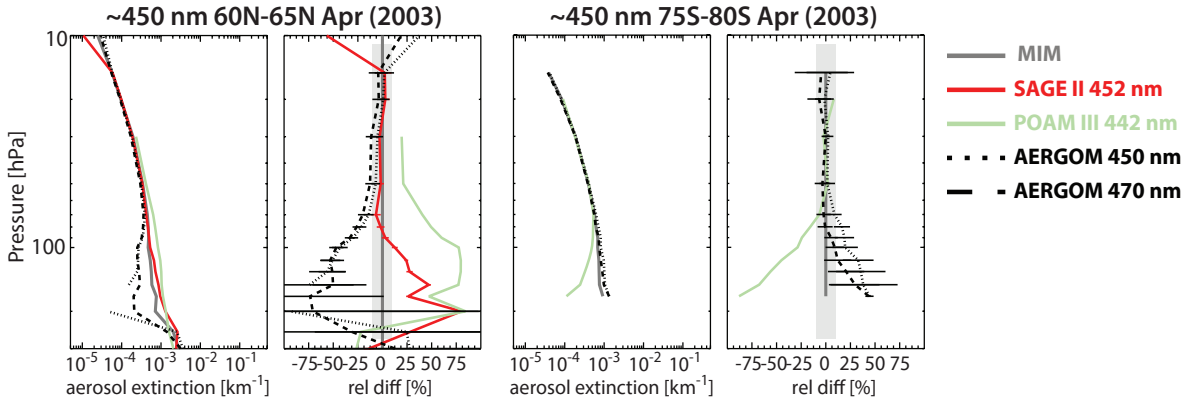


Figure 4.26.8: Vertical profiles of zonal monthly mean aerosol extinction in 2003 (retrieved at around 450 nm). *The aerosol extinction profiles are shown for different instruments (SAGE II, POAM III, and AERGOM) at 60°N-65°N and 75°S-80°S during April 2003 on the left and right, respectively. The relative differences between the individual instruments and the MIM are shown in the right-hand panels. 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 10%.*

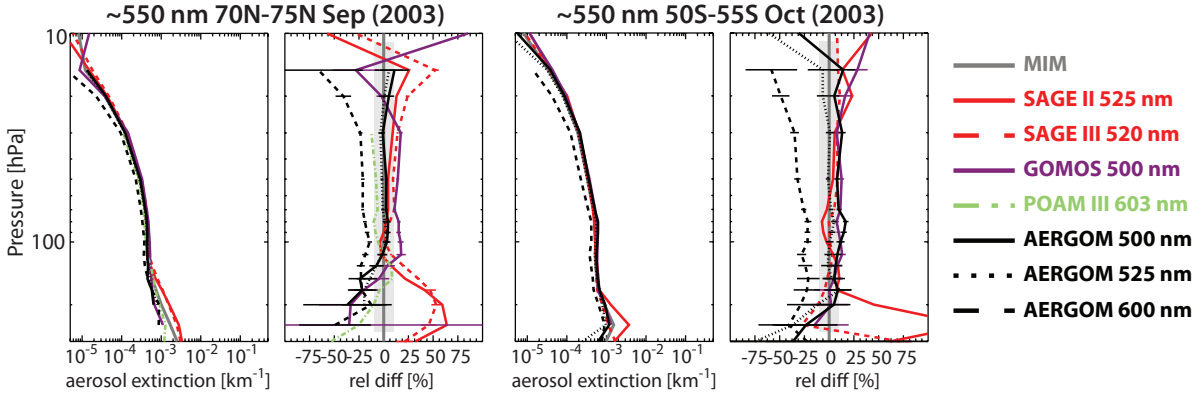


Figure 4.26.9: Vertical profiles of zonal monthly mean aerosol extinction in 2003 (retrieved at around 550 nm). *The aerosol extinction profiles are shown for different instruments and data products (SAGE II, SAGE III, POAM III, GOMOS, and AERGOM) at 70°N-75°N during September and 50°S-55°S during October on the left and right, respectively. The relative differences between the individual instruments and the MIM are shown in the right-hand panels. 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 10%.*

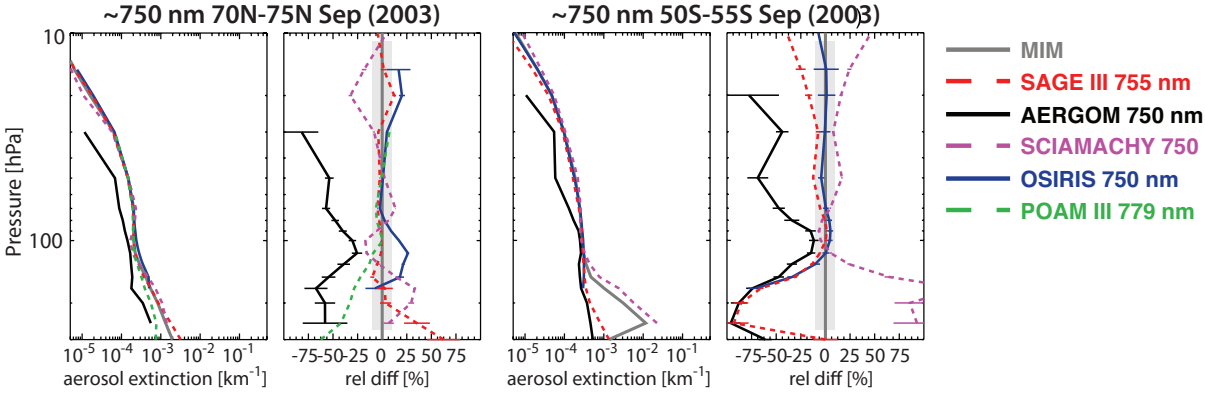


Figure 4.26.10: Vertical profiles of zonal monthly mean aerosol extinction in 2003 (retrieved at around 750 nm). *The aerosol extinction profiles are shown for different instruments (SAGE III, AERGOM, SCIAMACHY, OSIRIS, and POAM III) at 70°N-75°N and 50°S-55°S on the left and right, respectively. The relative differences between the individual instruments and the MIM are shown in the right-hand panels. 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 10%. Note AERGOM is not included in the MIM calculation.*

size distribution may affect the comparisons made here using scaled products adversely during the early years. Note that with increasing wavelength the time series of SAGE II (increasing relative anomalies) and HALOE (decreasing relative anomalies) show opposite behaviour. The dependency on aerosol size distribution may have been especially an issue within aged aerosol particles long after the Mount Pinatubo eruption. On the other hand, the aerosol

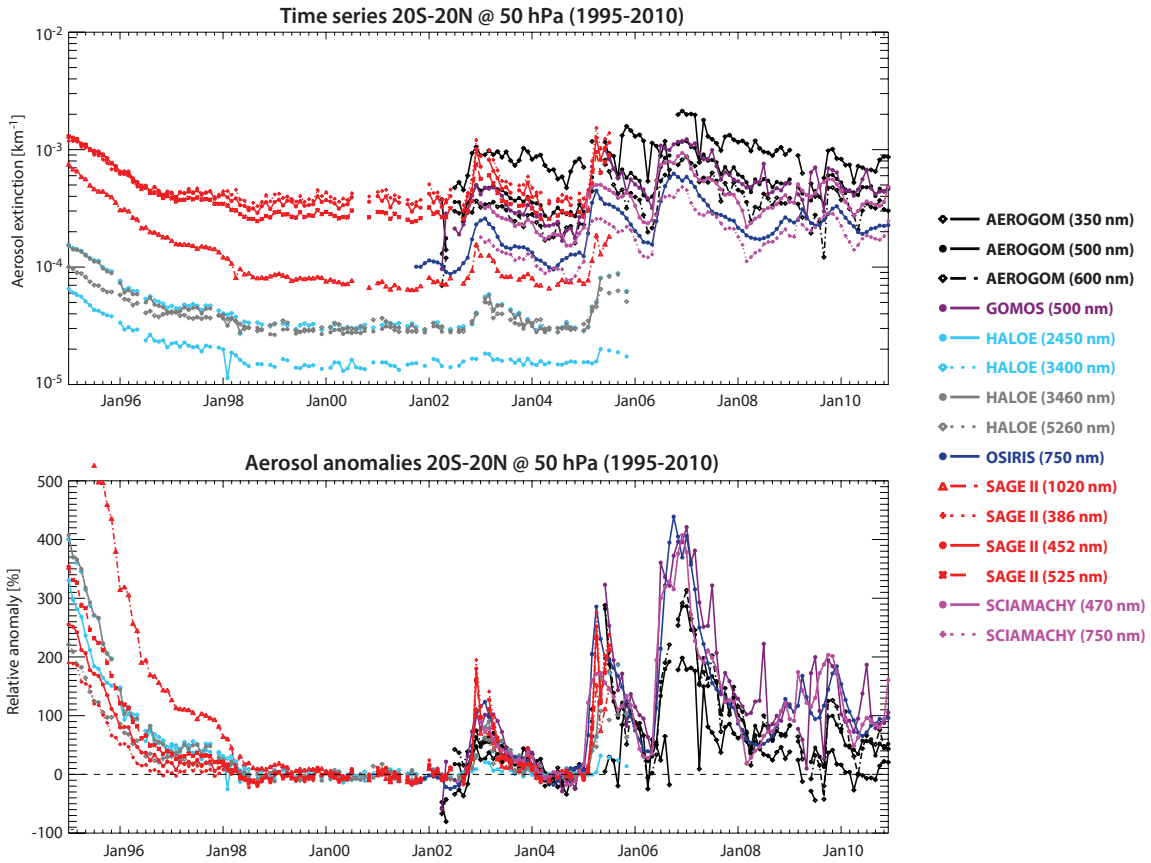


Figure 4.26.11: Tropical aerosol time series at 50 hPa. *Time series of aerosol extinction (upper panel) and deseasonalised and normalised aerosol extinction anomalies (lower panel) are shown for the latitude band 20°S-20°N at 50 hPa.*

extinction products of SAGE II at 452 and 525 nm and of HALOE at 3400 and at 3460 nm agree all fairly well with each other. HALOE at 2450 and 5260 nm seem to underestimate relative anomaly values during 2003 and 2005 when other products seem to agree well with each other. A problem in the HALOE 2450 and 5260 nm products has also been pointed out in a study by *Thomason* [2012], thereby lending support to our approach taken here.

In the later time period from 2002 onwards, GOMOS, while generally agreeing well with SCIAMACHY and OSIRIS, exhibits some strong spikes in its time series, which may be attributable to small sampling sizes. AERGOM disagrees with OSIRIS and SCIAMACHY (but not GOMOS) most strikingly during 2009, and all of its products tend towards lower aerosol extinctions than those from other instruments. The AERGOM product at 350 nm yields largest discrepancies.

Figure 4.26.12 shows the same as the previous figure, but for 80 hPa in the tropics. At this altitude, the scaled time series show somewhat larger disagreement (mostly within ±20-25%), although the main feature of a double-peaked structure with maxima during May and November is captured by most of them. Larger disagreement is expected since at these levels the retrievals are most influenced by high geophysical variability, inhomogeneous sampling, or cloud effects. SAGE II at 1020 nm shows somewhat higher variability than its other wavelength products. It remains to be investigated whether this indicates a potential retrieval problem or real natural variability that the longer wavelengths at these altitudes can capture better than the shorter wavelengths. Both GOMOS and AERGOM do not agree with the rest of the instruments during 2009. SCIAMACHY on the other hand shows too high values when compared to the other instruments in the beginning of its record.

Finally, **Figure 4.26.13** shows the same evaluation for the extra-tropics between 50°N-70°N. At these latitudes, POAM II and POAM III measurements can also be included in the comparison. Note that POAM II has been scaled using the scaling factor derived from POAM III measurements, which are measured at very similar wavelengths (see **Table 4.26.2**), since there are no POAM II measurements in the reference year 2004 available. POAM II exhibits the same wavelength dependency as SAGE II, with increasing relative anomalies with increasing wavelength, and the two instruments agree generally reasonably well with each other (mostly within 30% for POAM II at 921 and 1060 nm when compared to SAGE II at 1020 nm, and also for POAM II at 601 nm compared to SAGE II at 386, 452, and 525 nm). POAM III shows also a similar wavelength dependency during years with lower aerosol loading, and agrees well with HALOE and SAGE II overall. Finally after 2005, SCIAMACHY shows somewhat stronger maximum-to-minimum fluctuations than the other instruments. At least part of this behaviour may be attributed to the fact that the SCIAMACHY data have not undergone filtering to PSCs and clouds before inclusion in the monthly mean

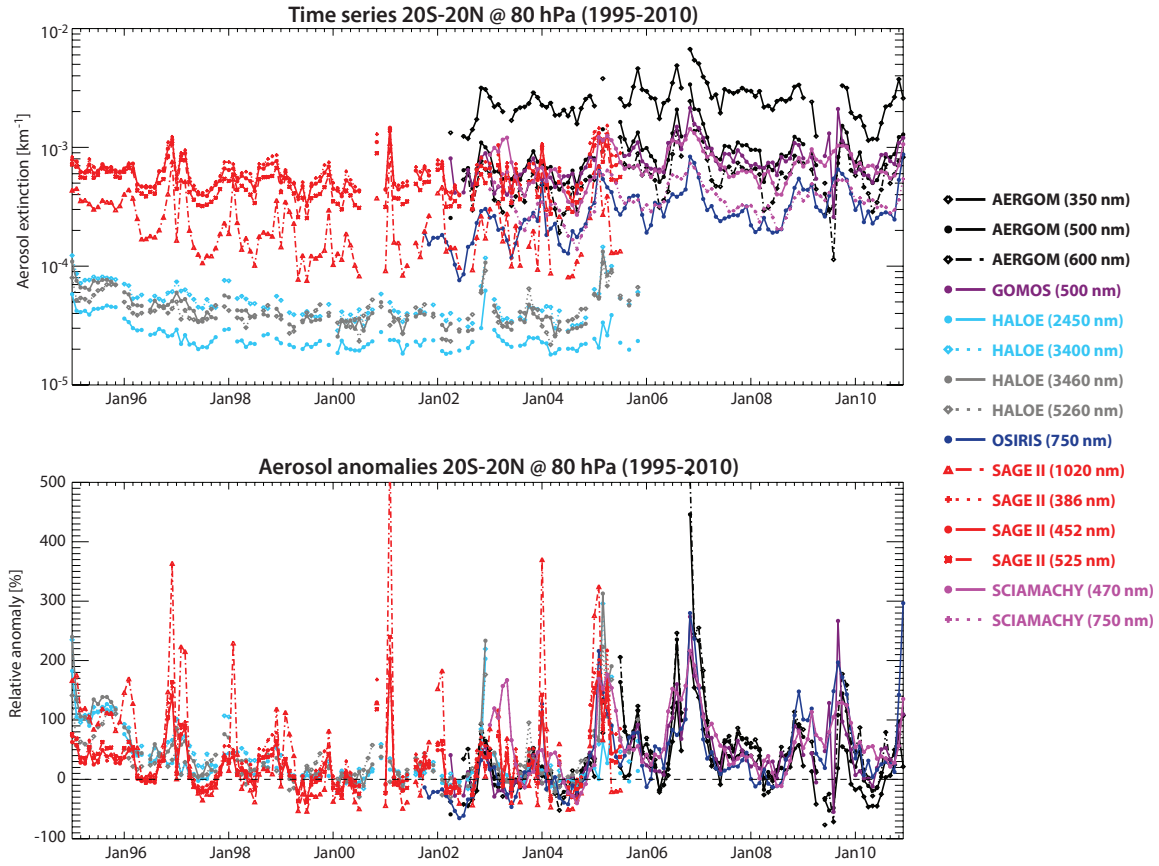


Figure 4.26.12: Tropical aerosol time series at 80 hPa. *Timeseries of aerosol extinction (upper panel) and deseasonalised and normalised aerosol extinction anomalies (lower panel) are shown for the latitude band 20°S-20°N at 80 hPa.*

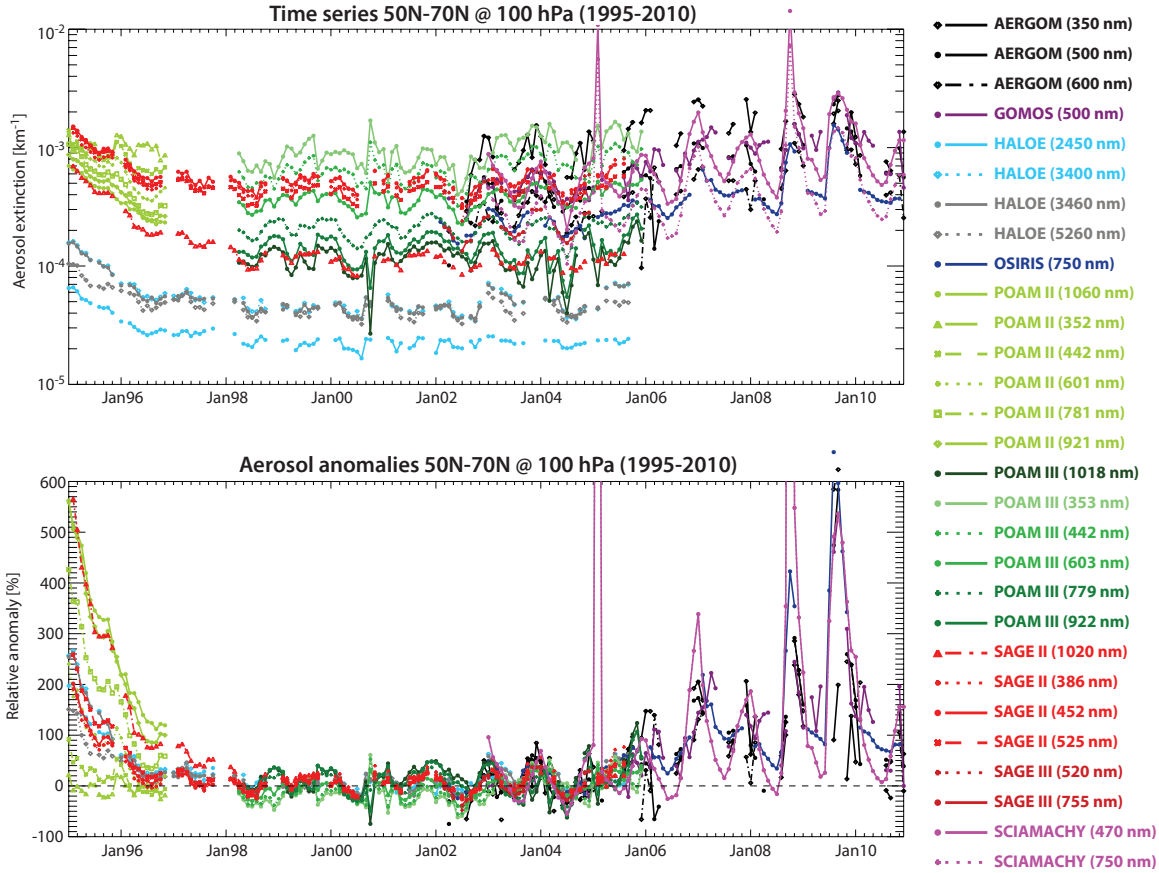


Figure 4.26.13: Extra-tropical aerosol time series at 100 hPa. *Timeseries of aerosol extinction (upper panel) and deseasonalised and normalised aerosol extinction anomalies (lower panel) are shown for the latitude band 50°N-70°N at 100 hPa.*

climatologies. GOMOS and AERGOM (except at 350 nm) indicate mostly good agreement with OSIRIS even during 2009, however show some spikes in its time series as noted for other levels/latitudes above.

Figure 4.26.14 shows the evolution of relative anomalies as derived from Equation 1 for the different instrument observations and for the MIM (upper left) now as a function of time and altitude. The original, unscaled data can be seen in **Figure A4.26.1** in *Appendix A4*. Note, the MIM is not to be mistaken for a climate data record, but merely represents a reference, since it includes all the available aerosol climatologies and not only those considered to be of high quality.

The MIM anomaly time series clearly reveals distinct pulses of enhanced stratospheric aerosol after the Nevado del Ruiz (1985), Kelut (1990), and Mount Pinatubo (1991) volcanic eruptions, a phase of relaxation towards background aerosol extinction values towards the late 1990s, and a relatively 'clean' background stratosphere between 1998 and 2004. After that a period marked by some intermediate aerosol influence following the eruptions of some smaller volcanoes such as Manam (2005), Soufrière Hills and Tavurvur (2006) occured [see *Vernier et al.*, 2011]. The source of aerosol in 2009 may indicate the impact of the Australian 'Black Saturday' bushfires on the stratospheric aerosol layer [*Siddaway et al.*, 2011].

The comparison of the different instrument products with the timeseries of the MIM reveals overall good agreement in terms of the structures seen. Note that the high values in the tropical upper troposphere in SCIAMACHY are due to the fact that the observations had not been filtered for clouds before they were included in the monthly zonal mean climatologies. Above around 100 hPa, the structures in the SCIAMACHY anomalies however resemble those found in the MIM well. The AERGOM time series from retrievals at 350, 600, and 750 nm seem somewhat noisier than the AERGOM products derived at other wavelengths and hence do not represent the structures as well as the other instruments' time series.

In order to be more quantitative about the inter-instrument differences, **Figure 4.26.15** shows the relative differences between each instrument's anomaly time series and the MIM. We first focus on the SAGE II and HALOE products during the rather clean phase of stratospheric aerosol loading after 1998 up to 2005. Smallest deviations from the MIM are seen in the SAGE II products at 452 and 525 nm, and the HALOE products at 3400 and 3460 nm with values mostly between ±5%. SAGE II at 1020 nm largely agrees with these products except for altitudes above 20 hPa, where it shows larger positive differences from the MIM of +10 to +20%. HALOE at 3400 nm shows increasing deviations below around 50 hPa. Both HALOE products at 2450 and 5260 nm show somewhat larger differences with respect to the MIM of ±10%. These findings are mostly consistent with the findings by *Thomason* [2012], who stated that the HALOE products at 2450 and 5260 should not be used, and the product at 3460 nm may suffer from a NO_2 contamination in the retrieval below 19 km.

During the years with high aerosol loading following the Mt. Pinatubo eruption, the different products however show large differences that exhibit in some cases even a vertical structure to them that complicates the interpretation. SAGE II at 1020 nm is consistently high when compared to the MIM, while SAGE II at 386 and 452 nm, and HALOE at 5260 nm are low. SAGE II at 525 nm shows positive/negative deviations above/below approximately 30 hPa, and HALOE at 2450, 3400, and 3460 nm approximately the opposite behaviour. The products during this time period seem to suffer from inconsistencies most likely attributable to the differences in aerosol size distributions assumed in the different retrievals.

The comparison with the newer generation of limb-scattering and limb occultation instruments show largely encouraging results, although with somewhat larger biases even during quiescent years. For OSIRIS, SCIAMACHY at 750 nm (or 470 nm), GOMOS, and AERGOM at 450, 470, 500, and 525 nm differences to the MIM are of around ±10%, while for AERGOM at 350, 600, and 750 nm differences increase to more than ±20%. During the time period with higher aerosol loading, the differences to the MIM increase, with mostly negative features in most AERGOM products and SCIAMACHY, and positive features in GOMOS and OSIRIS. Note that the alternating positive/negative departures from the MIM in GOMOS, OSIRIS, and SCIAMACHY may point towards sampling issues of these instruments, while the mostly negative departures in AERGOM seem to reflect a systematic low bias in the retrievals compared to the other instruments' aerosol extinction products.

The equivalent anomaly time series as in **Figure 4.26.14** are shown in **Figure 4.26.16** but for the northern mid- to high-latitudes (50°N-70°N). At these latitudes, we can also include POAM II and POAM III measurements into the comparison. In the early years of the comparison, high aerosol extinction values after the Mt. Pinatubo eruption relax to lower background values, with similar features as seen in the tropics. The MIM after 2000 shows two different layers of enhanced aerosol, one above 30 hPa, which most likely reflects aerosol from the Mt. Pinatubo eruption transported within the deep branch of Brewer-Dobson circulation to higher altitudes and latitudes, and one below 70 hPa, which rather stems from more recent volcanic injections of aerosol precursors into the LS at higher latitudes.

The different aerosol extinction products from the instruments show overall similar features, but the overall agreement is much worse than in the tropics. Much of these differences may be explainable by differences in sampling, which are considered to affect aerosol measurements more than long-lived trace gas species due to less homogeneous distributions in the stratosphere. However, differences between products of the same instrument reveal that major problems in certain wavelength products exist that must be real and

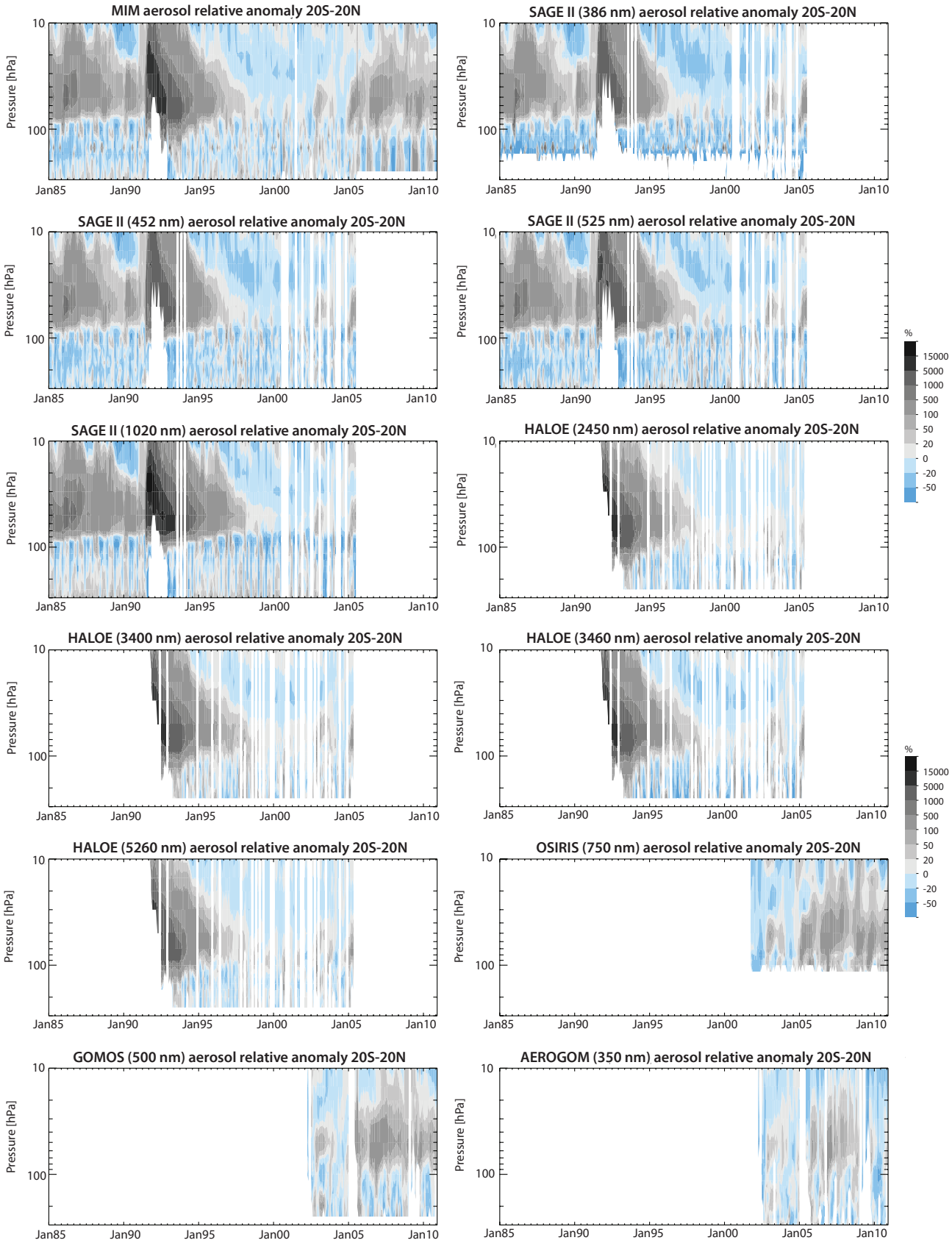


Figure 4.26.14: Time-altitude evolution of aerosol anomalies in the tropics. *The time-altitude evolution of normalised aerosol anomalies averaged over 20°S-20°N is shown for the MIM and all retrieval products of the different limb satellite sounders.*

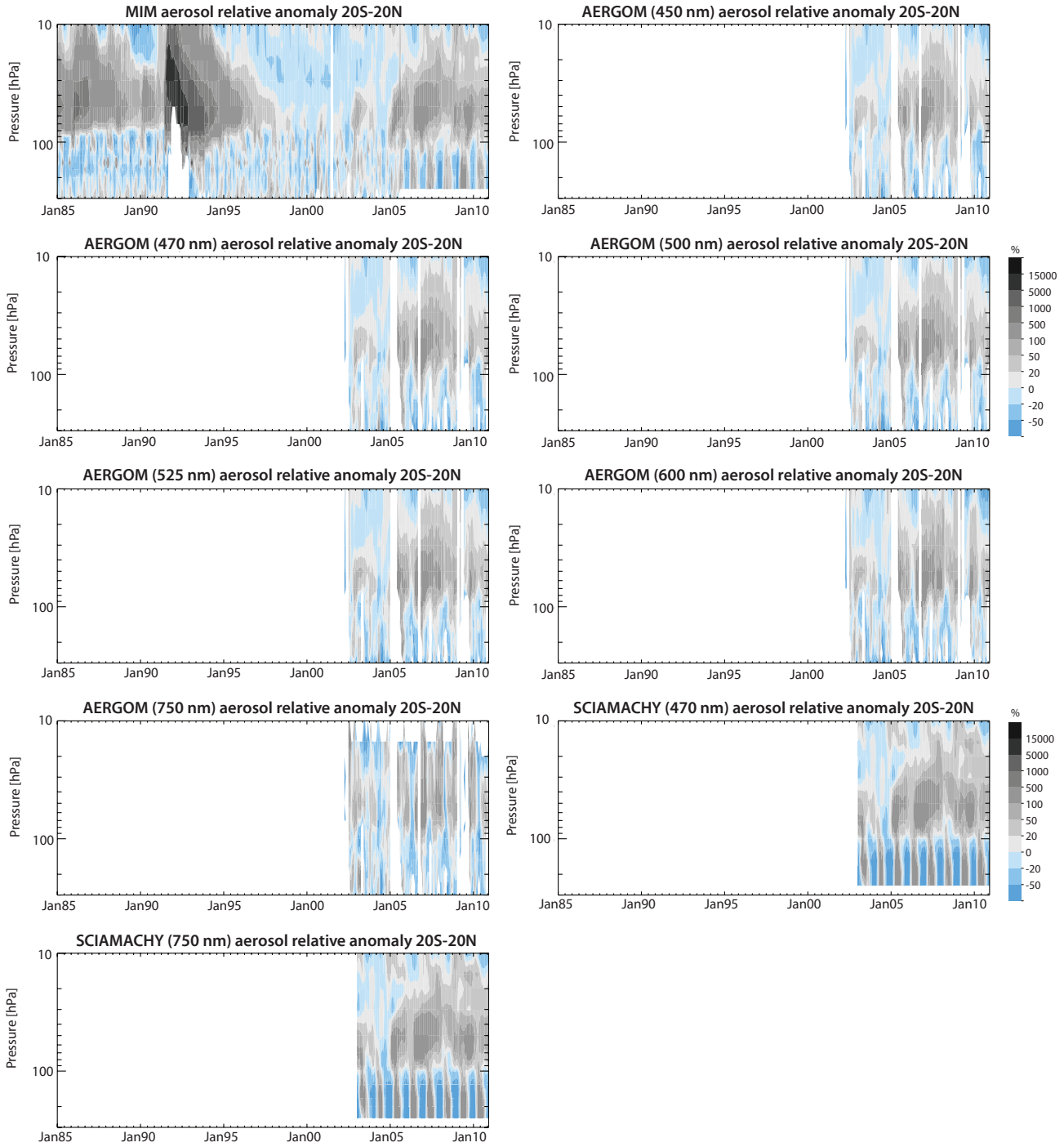


Figure 4.26.14 continued.

cannot be attributed to sampling differences. Most striking differences are found in HALOE at 2450 and 5260 nm, with both products not showing an enhanced aerosol extinction layer above 30 hPa. Also, POAM II at 352 nm seems not to show the same physical features as seen in the other time series. AERGOM time series at these latitudes have only intermittent coverage, but indicate also two layers of aerosol as seen in the other time series. Both POAM III products at 603 and 779 nm show very good to excellent agreement when compared to the MIM through most of the lower and middle stratosphere, along with slightly higher differences for the other products, which may be reflective of the limitation of the comparison methodology.

Relative differences from the MIM are shown for each anomaly time series in **Figure 4.26.17**. It is notable that these do not necessarily show the same behaviour as in the tropics. While SAGE II at 1020 nm shows positive deviations between 200 and 30 hPa in the beginning of the record as in the tropics, it here now shows negative differences above 30 hPa. All the other SAGE II products now show negative differences in the LS, and positive deviations in the MS. Also, the differences between the time series are here much larger with values around ±10% during 2000-2004, but with patches of larger and increasing differences that reach up to ±20 to ±50% towards the end of the time-series. The less reliable products from AERGOM at 600 and

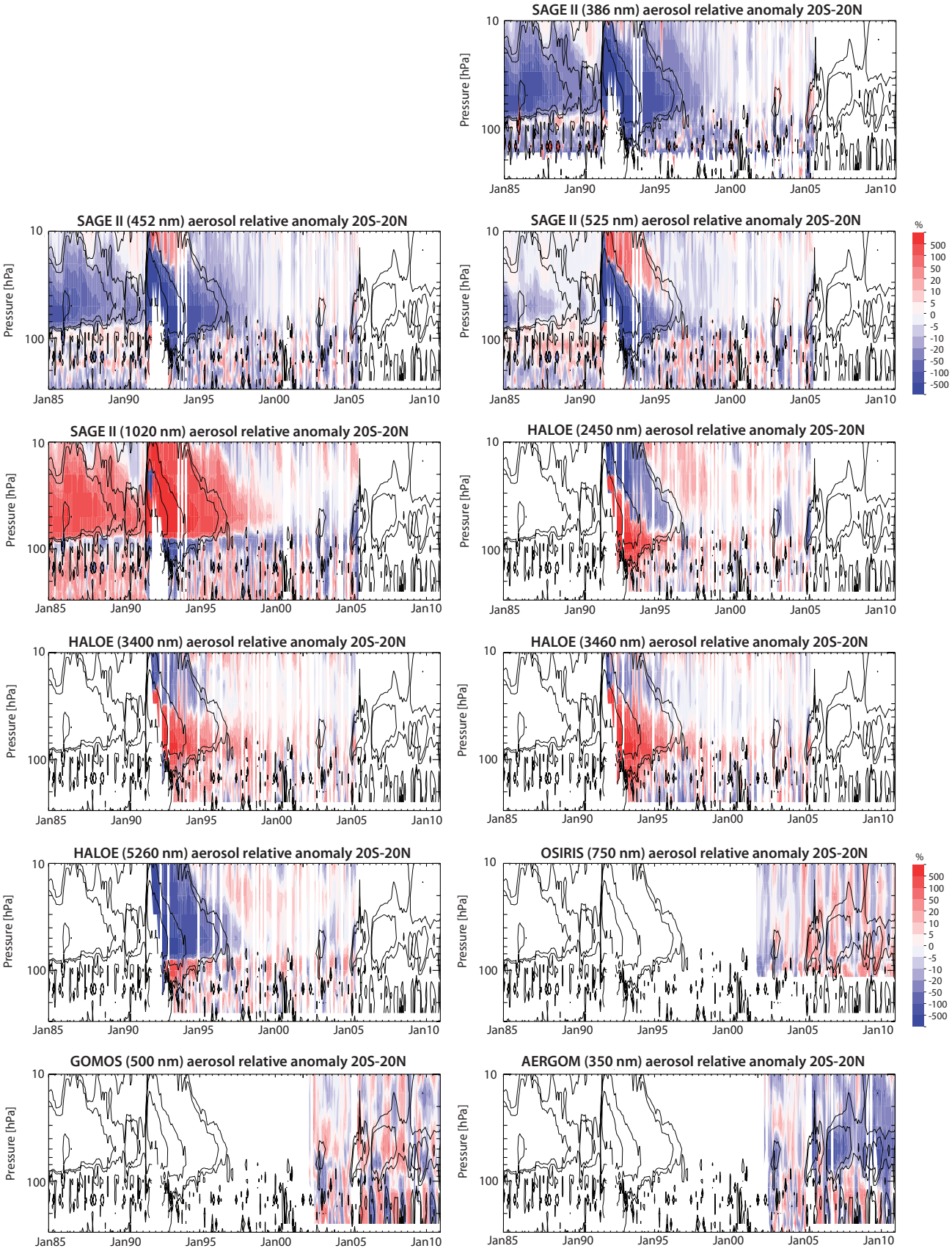


Figure 4.26.15: Time-altitude evolution of aerosol anomaly differences in the tropics. *The time-altitude evolution of normalised aerosol anomaly differences averaged over 20°S-20°N with respect to the MIM is shown for all retrieval products of the different limb satellite sounders.*

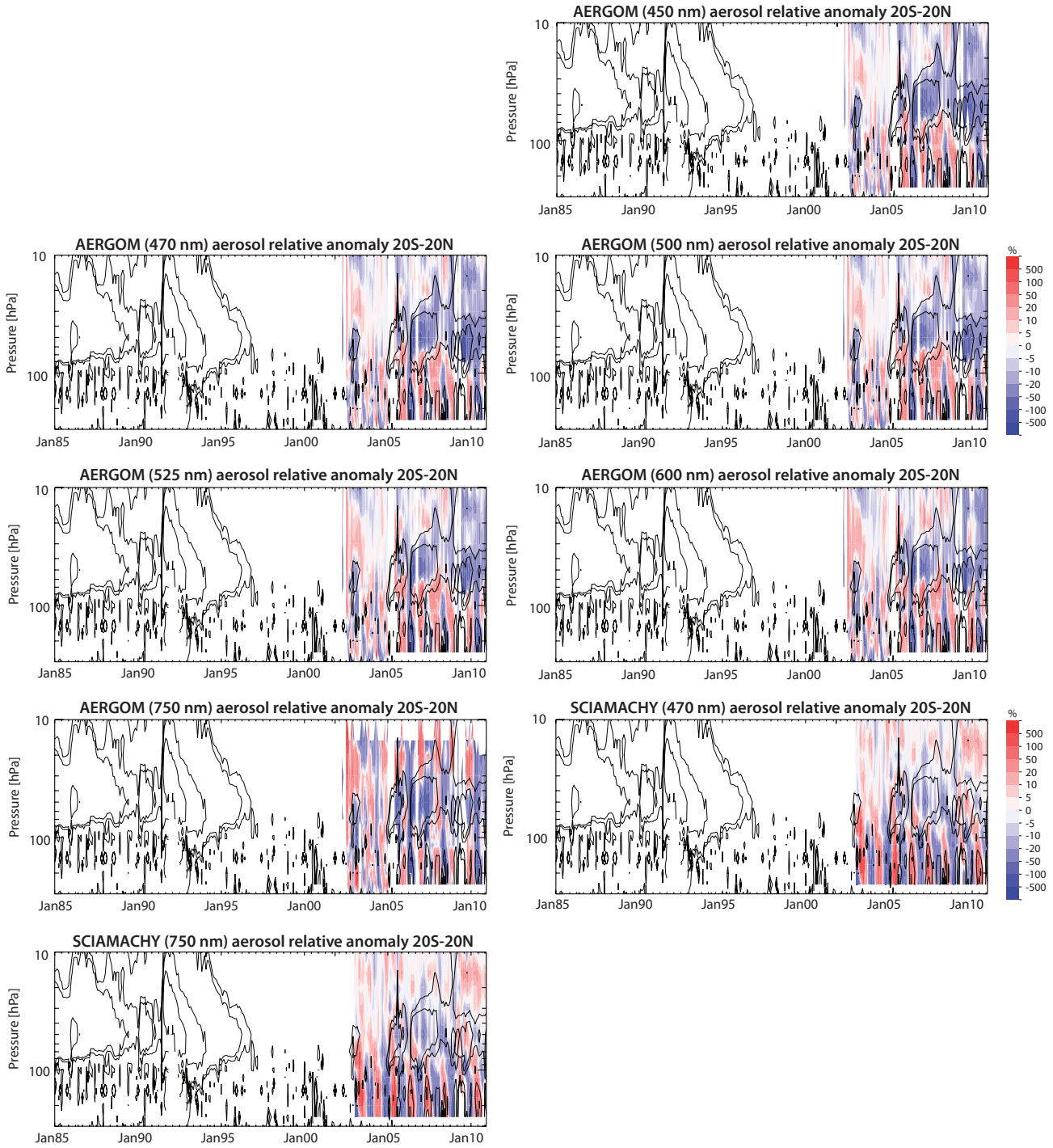


Figure 4.26.15 continued.

750 nm show mostly negative deviations as also shown in the evaluations considering similar wavelengths only (see **Figures 4.26.9** and **4.26.10**). Note, the original unscaled time series of aerosol extinction for the extra-tropics can be found in Figure **A4.26.2** in *Appendix A4.*

4.26.5 Summary and conclusions: Aerosol

Within the SPARC Data Initiative, a first overall comparison of available aerosol monthly zonal mean climatologies based on aerosol extinction profile measurements from 8 satellite instruments (SAGE II, HALOE, POAM II, POAM III, OSIRIS, SAGE III, SCIAMACHY, and GOMOS) has been carried out. From these instruments a total of 34 products are available, all retrieved at different wavelengths ranging from 350 to 5260 nm. Given the wavelength-dependency of aerosol extinction retrievals, the available products cannot all be directly compared to each other. Two different approaches have hence been chosen to compare the aerosol extinction products to each other to gain information on their quality and physical consistency. Note, interpretation of the findings from the two comparison approaches remains difficult and needs to be used with caution.

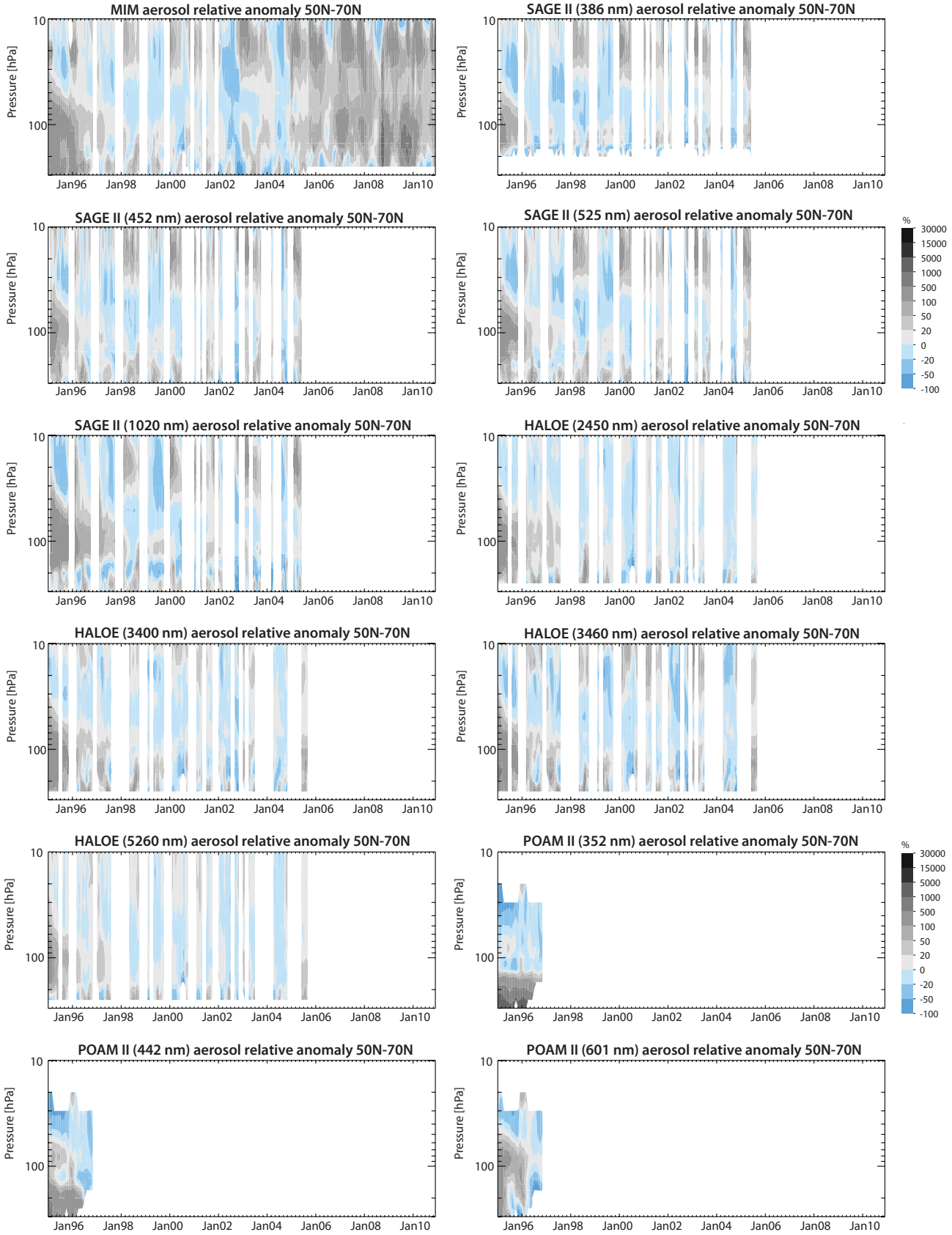


Figure 4.26.16: Time-altitude evolution of aerosol anomalies in the extra-tropics. *The time-altitude evolution of normalised aerosol anomalies averaged over 50°N-70°N with respect to the MIM is shown for all retrieval products of the different limb satellite sounders.*

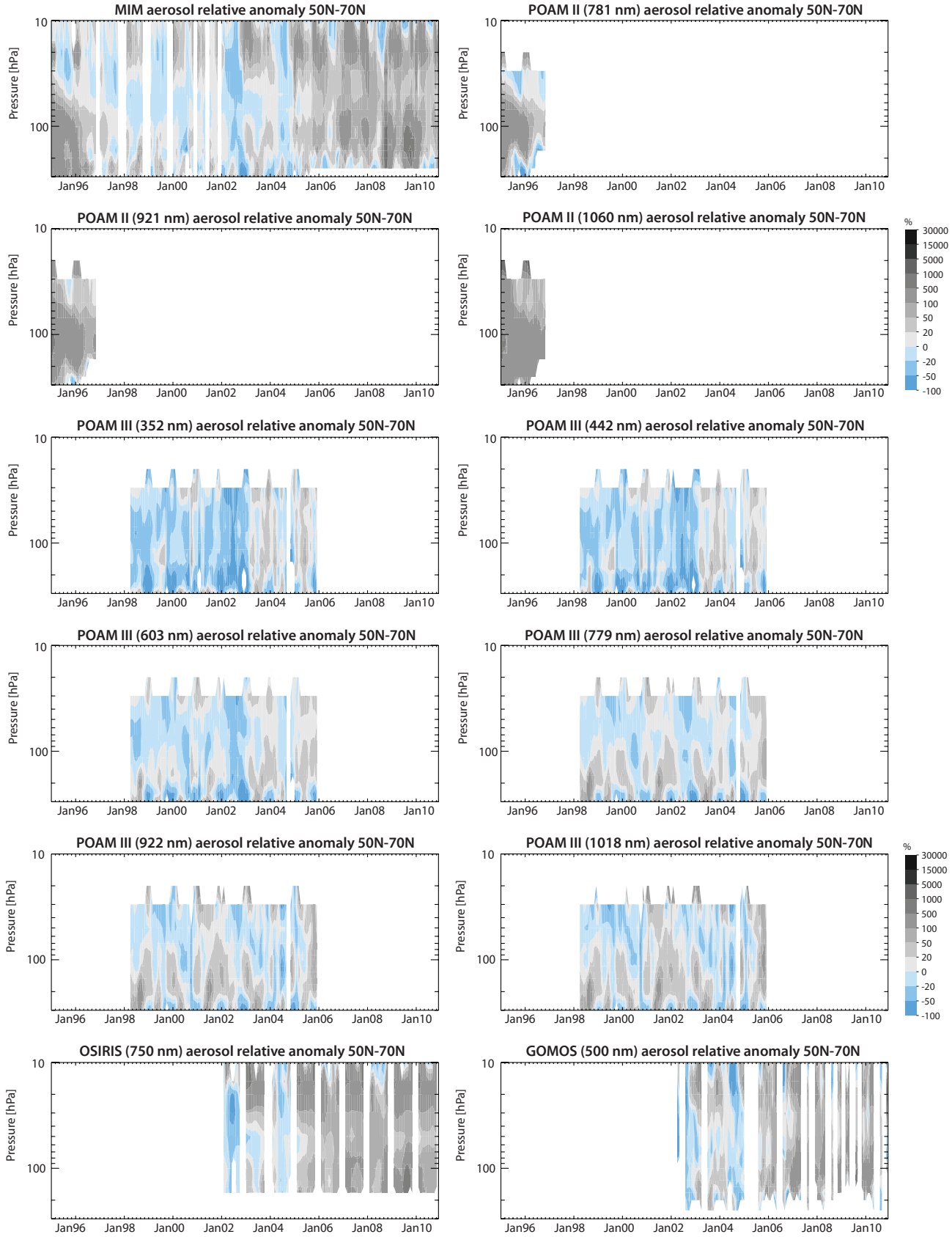


Figure 4.26.16 continued.

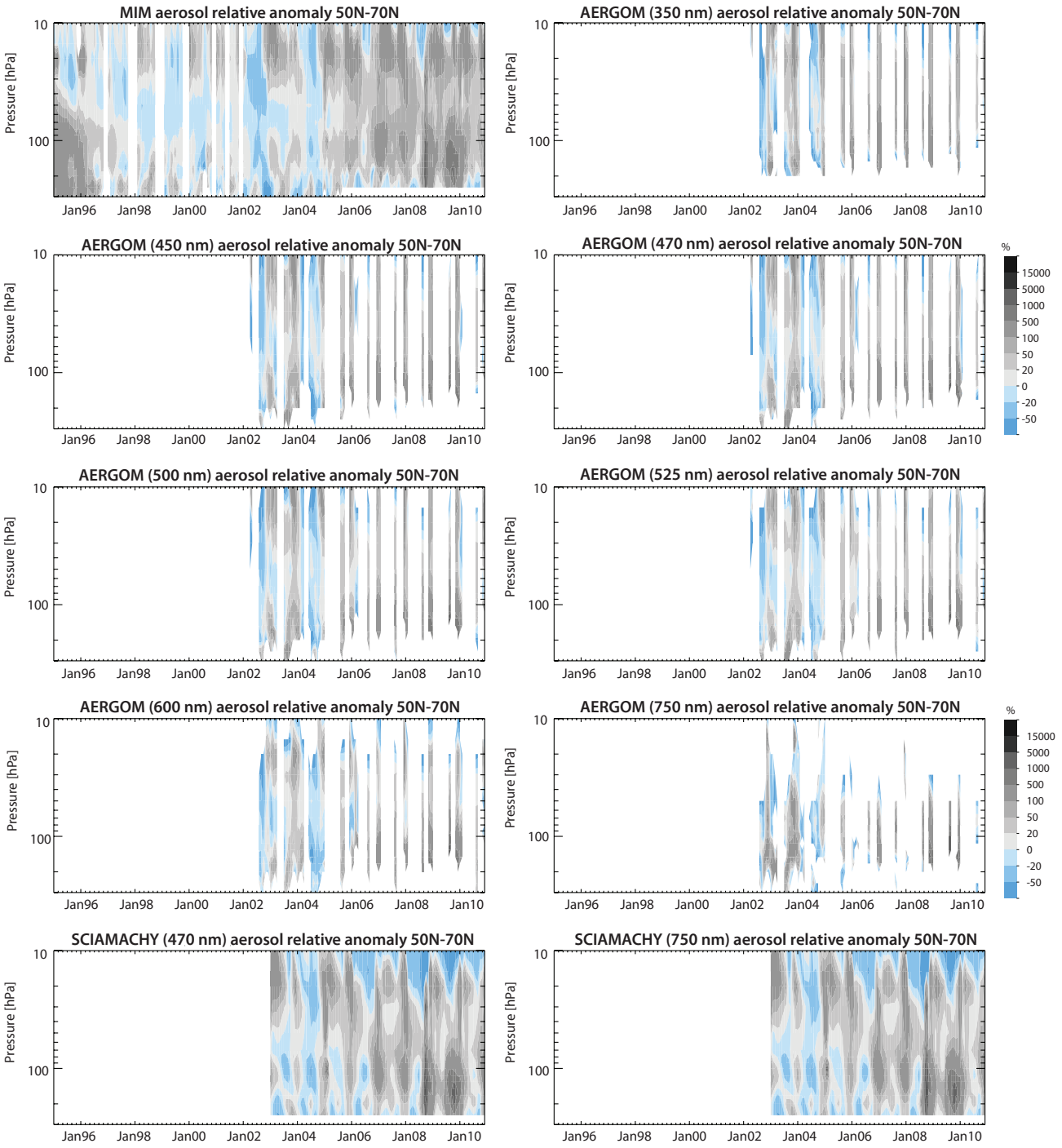


Figure 4.26.16 continued.

In the first approach, original aerosol extinction products retrieved at similar wavelengths are compared. Here, generally the aerosol extinction products show biases that at least partially reflect expected wavelength-dependencies. Some real inter-product differences however are revealed. For example, the AERGOM and SCIAMACHY products at 470 nm do show relatively large differences from the MIM in the tropics of around 30% (with AERGOM/SCIAMACHY on the high/low side) and *vice versa* in the mid-latitudes. The POAM III product at 603 nm shows mostly positive differences from the MIM, despite the fact that it was the product retrieved at the highest wavelength (and hence should show negative differences from the MIM). At 750 nm, AERGOM shows a clear negative bias when compared to the MIM (which consists mostly of products retrieved at the same wavelength). OSIRIS and SCIAMACHY at 750 nm on the other hand agree well with each other throughout the tropical and mid-latitude MS and LS, with differences largely within ±10% of the MIM. SAGE III at 755 nm compares even very well to OSIRIS (±5%).

In the second approach, a normalisation using the mean aerosol extinction value derived from each product's climatology during a quiescent period as scaling factor is applied to each product's time series. This approach neglects the spectral dependence of the normalisation factor on aerosol-size distributions, which we assume to

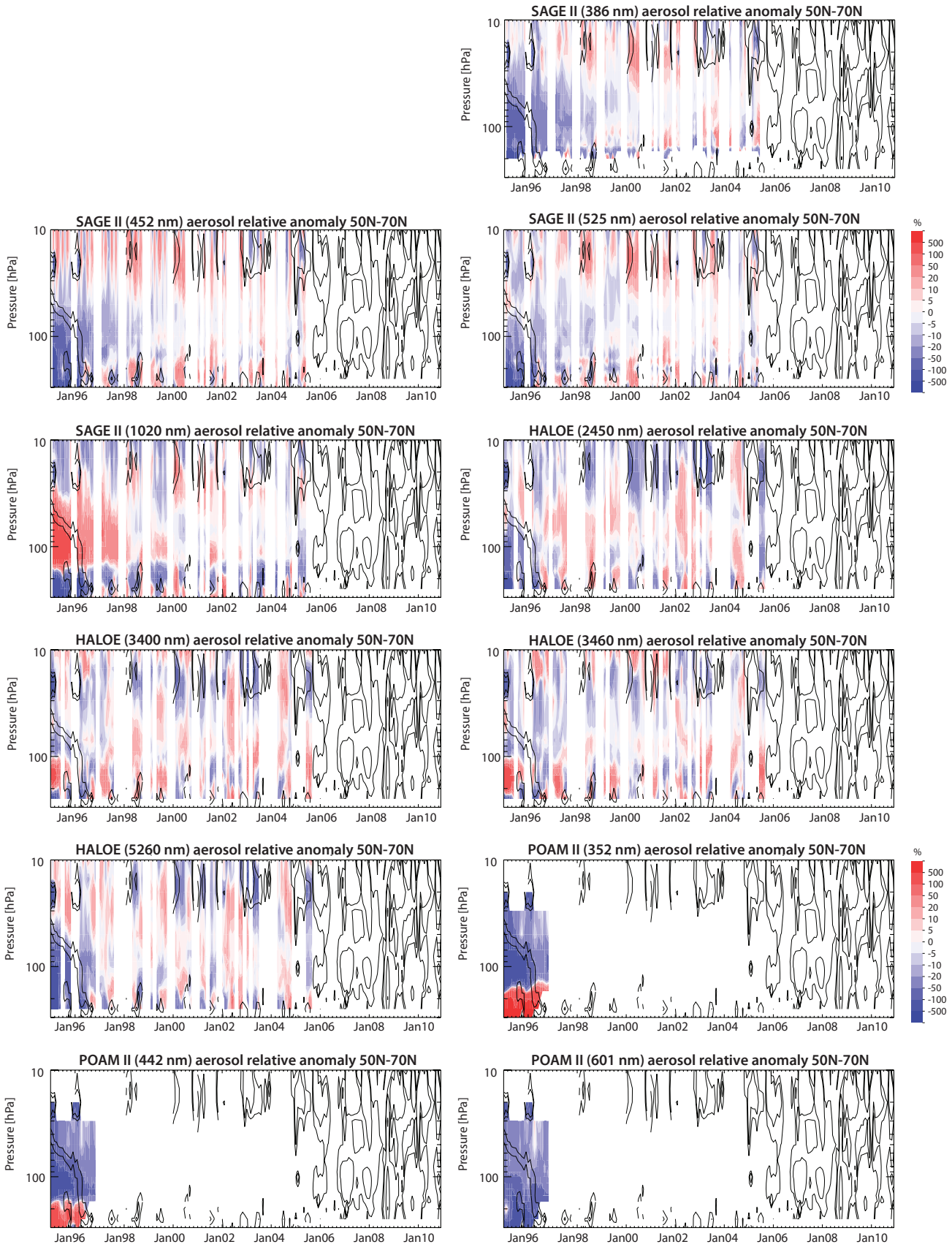


Figure 4.26.17: Time-altitude evolution of aerosol anomaly differences in the extra-tropics. *The time-altitude evolution of normalised aerosol anomaly differences averaged over 50°N-70°N with respect to the MIM is shown for all retrieval products of the different limb satellite sounders.*

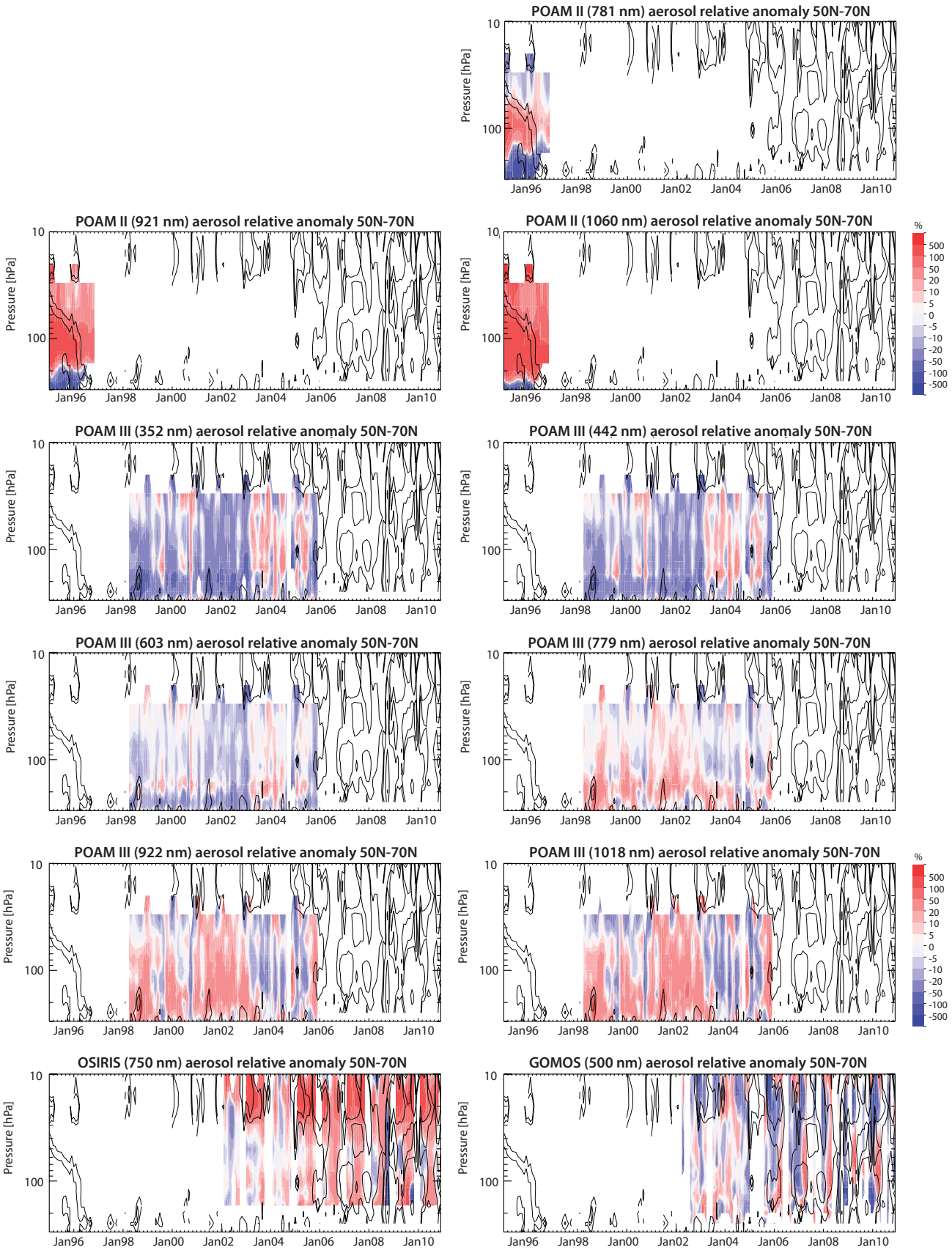


Figure 4.26.17 continued.

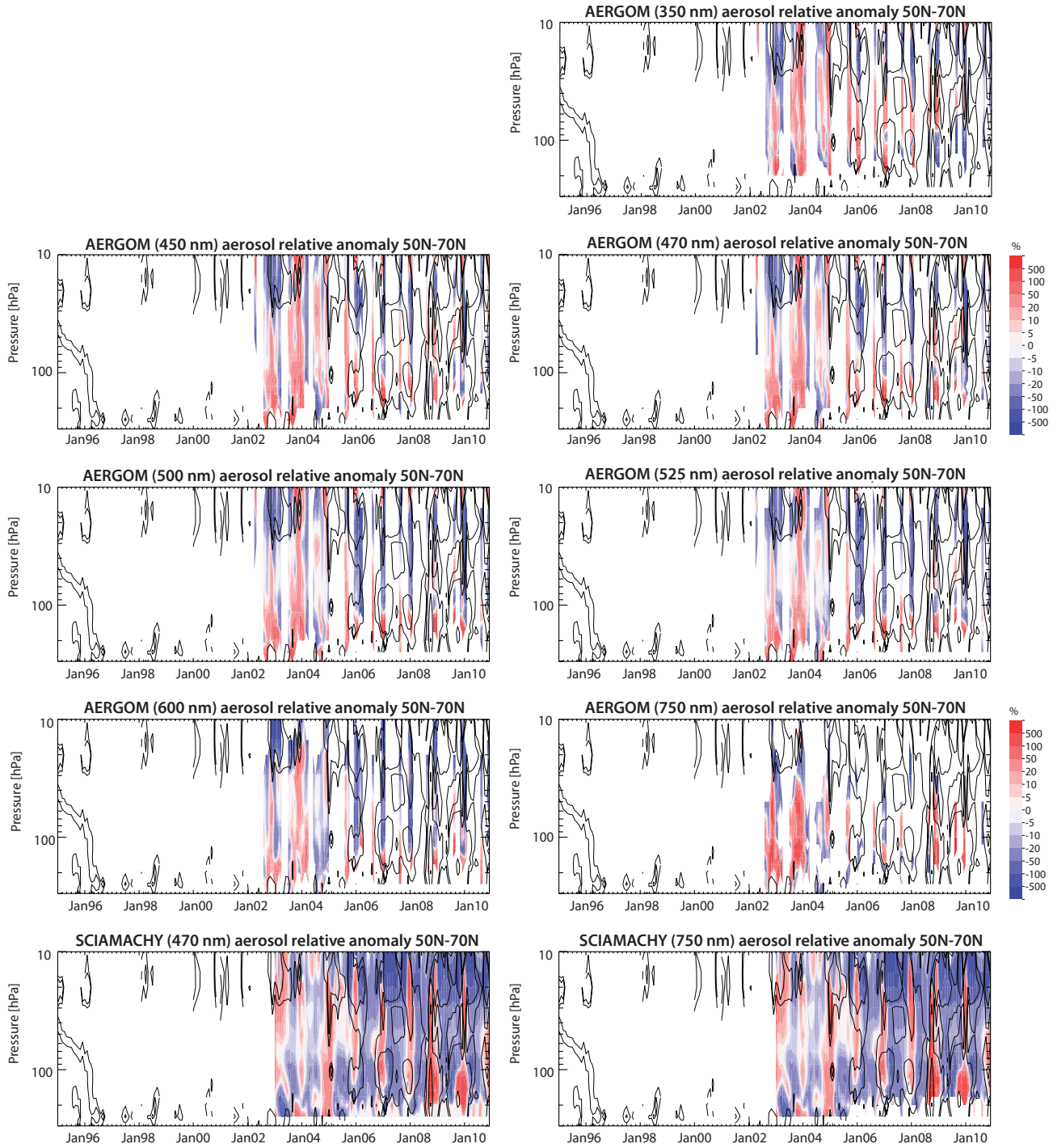


Figure 4.26.17 continued.

be of second-order importance (see also *Thomason*, 2012). Results from these comparisons can be summarised as follows.

Evaluation of anomaly time series in both the tropics and the extra-tropics reveal that most of the aerosol products capture the physical structures in the evolving aerosol layer of the stratosphere well. The comparison of the different instrument products with the time series of the MIM reveals overall good agreement in terms of the physical structures seen in the evolving stratospheric aerosol layer. Exceptions are HALOE at 2450 and 5260 nm, with both products missing to show an enhanced aerosol extinction above 30 hPa, and POAM II at 352 nm, which does not show the expected vertical structure. AERGOM time series at 350, 600, and 750 nm seem somewhat noisier than AERGOM products derived at other wavelengths and as a consequence capture the physical structure less well.

During the time period 1998-2004 with conditions of relatively low aerosol loading, most aerosol products agree very well to within ±5-10% from the MIM in the MS. At altitudes below around 70 hPa, the differences increase somewhat, but still remain largely within ±10-20%. Exceptions are the HALOE products at 2450 nm and 5260 nm, which show both somewhat larger (and also time-dependent) differences. These results confirm earlier findings

from *Thomason* [2012] who deemed these products as less reliable. The comparison with the newer generation limb-scattering and limb occultation instruments shows also encouraging results, although with somewhat larger biases, especially when moving into the extra-tropics. For OSIRIS, SCIAMACHY at 470 and 700 nm, GOMOS at 500 nm, and AERGOM at 450, 470, 500, and 525 nm differences to the MIM are of around ±10%, while for AERGOM at 350, 600, and 750 nm differences increase to up to ±20%. Note that the larger differences found in the extra-tropics between the instruments may at least partially be explained by larger sampling biases.

During the years with high aerosol loading following the Mt. Pinatubo eruption, the different products show increasing differences that exhibit often a vertical structure. This finding points towards problems in either the retrievals during periods with high aerosol loading, or problems with the comparison method that may over-simplify the wavelength-dependent sensitivity of the retrieval on aerosol size distributions.

4.27 Upper troposphere / lower stratosphere (UTLS) ozone evaluations based on TES averaging kernels

Section 4.1 provides a detailed description and comparison of the ozone climatologies from limb-viewing instruments, with a primary focus on the stratosphere. In this section we consider the distribution of ozone in the UTLS (300-70 hPa) and compare ozone measurements from six limb-viewing instruments (ACE-FTS, Aura-MLS, HIRDLS, MIPAS, OSIRIS, and SCIAMACHY) to those from the nadir-viewing Tropospheric Emission Spectrometer (TES) for 2005-2010. These results are also presented in *Neu et al.* [2014a]. TES is the only nadir-viewing instrument in this initiative, as well as the only instrument with a focus on tropospheric composition. Its ozone measurements have good sensitivity from the surface to 10 hPa and are well-validated against ozonesondes in the UTLS [*Nassar et al.*, 2008; *Boxe et al.*, 2010], as discussed in *Section 4.27.2*. Because TES is nadir viewing, it has relatively coarse vertical resolution (~6-7 km; **Figure 4.27.1a**) compared to the limb-viewing instruments discussed here, most of which have vertical resolutions of ~2-4 km. While TES has much finer horizontal resolution (<10 km) than the limb sounders (~200 km), the spacing between measurements is 182 km; thus its ability to resolve horizontal features is not much different than that of the limb sounders.

Given the strong gradients and small-scale structure of trace gas fields in the UTLS, differences in sampling and in vertical and horizontal resolution among instruments can lead to large differences that reflect sampling or smoothing error rather than systematic bias. *Toohey et al.* [2013] (see also *Section 3.2*) addresses the issue of sampling bias and shows, for example, that the construction of the zonal mean climatologies used here can lead to biases of a few percent in the subtropical jet regions (~30°N and 30°S) due to a combination of the sloping ozone surfaces in these regions and the increase in sampling density with latitude. Throughout the rest of this report, a simplified approach of directly comparing the limb-sounding climatologies without accounting for differences in vertical resolution has been used. However, because the vertical resolution of TES is so much

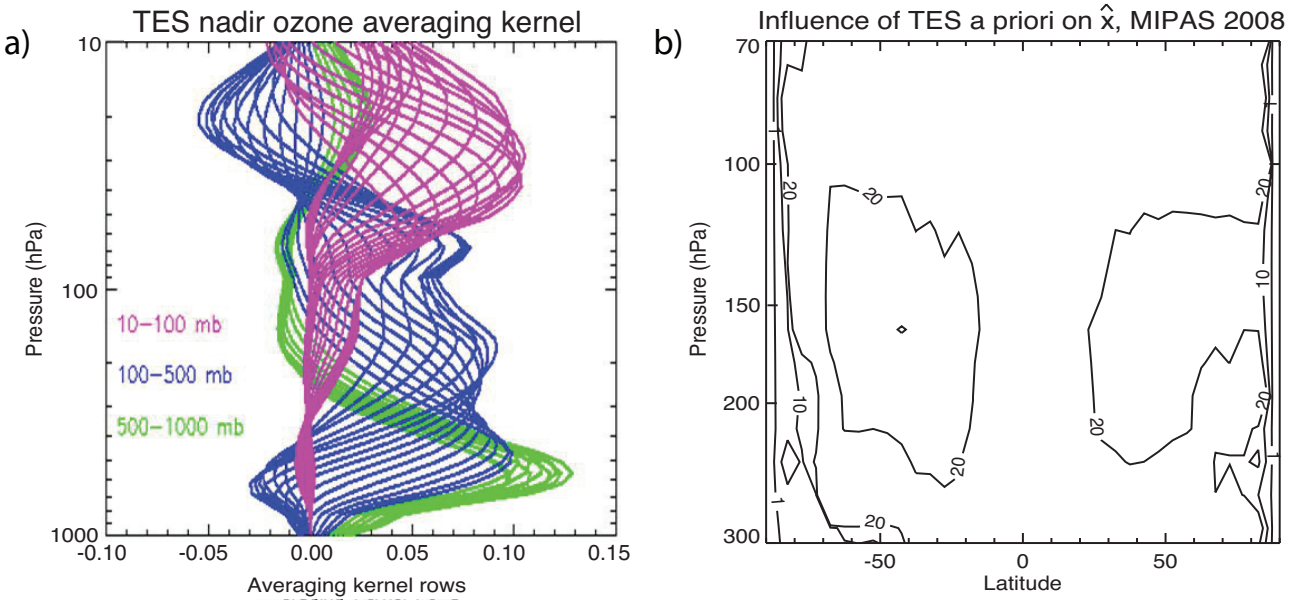


Figure 4.27.1: a) Sample TES averaging kernel. *The lines show the relative contribution of the "true" mixing ratio at each pressure level to the retrieved mixing ratio at 500-1000 hPa (green), 100-500 hPa (blue) and 10-100 hPa (purple). TES ozone averaging kernels vary with temperature, surface properties, clouds, and ozone.* ***b) Influence of the TES a priori on the virtual retrieval for MIPAS.*** *Annual mean value of the ratio of* Ax *(the contribution of the original climatology to the virtual retrieval) to* $x^a - Ax^a$ *(the contribution of the TES a priori to the virtual retrieval) for MIPAS for 2008. Results are very similar for other instruments and years. When the ratio is close to 1, the terms are of similar magnitude, so that the a priori and true profiles contribute equally to the retrieved ozone; values less than ~20 indicate a significant contribution of the a priori to the retrieval. Contour values of 1, 10, and 20 are shown.*

coarser than that of the limb viewing instruments and because vertical resolution is so critical in the UTLS, in this section we apply the TES observational operator (averaging kernel + constraint) to ozone climatologies from the limb-viewing instruments. This minimises the impact of vertical resolution disparities and allows identification of systematic differences in the large-scale structure and variability of UTLS ozone among the instruments. At the same time, the approach smooths the finer-scale vertical gradients present in the limb measurements and therefore represents a loss of this information. We thus include in this section an extensive analysis of the impact of the TES observational operator on the limb-viewing climatologies.

Ozone is the third largest component of radiative forcing [*Solomon et al.*, 2007], with maximum radiative effect in the UTLS [*Forster and Shine*, 2002]. Yet the processes that control the UTLS distribution of ozone and its trends and variability, including the exchange of air between the stratosphere and troposphere, are not well quantified [*WMO*, 2011]. The UTLS region is characterised by strong ozone gradients and complex and rapidly evolving small-scale features such as tropopause folds [*Gettelman et al.*, 2011 and references therein]. Aircraft measurements are well-suited for characterizing UTLS chemistry and dynamics because of their high spatial and temporal resolution. However, aircraft have only sparsely sampled the UTLS, raising questions about the representativeness of these measurements for applications such as evaluating free-running global chemistry-climate models [*SPARC*, 2010; *Hegglin et al.*, 2010]. Current satellite instruments lack the spatio-temporal resolution to resolve some UTLS features, such as thin, highly dynamic filaments. Furthermore, they can have low signal-to-noise in the UTLS because of the small ozone abundance there relative to the middle stratosphere, and clouds can interfere with trace gas retrievals. However, satellites provide much greater spatial and temporal coverage than aircraft, at a vertical resolution that is commensurate with that of most models [*SPARC*, 2010; *Hegglin et al.*, 2010], and their measurements have provided extensive improvements in our understanding of UTLS structure and processes [*e.g.*, *Hegglin et al.,* 2009; *Manney et al.,* 2011; *Peevey et al.*, 2012]. Assessing the differences between satellite measurements in the UTLS is critical to advancing our understanding of this region and evaluating UTLS processes in models because uncharacterised biases in satellite data can lead to incorrect conclusions about UTLS chemistry or radiative forcing.

We include TES data in this comparison and use it to evaluate measurements from the limb sounders in part because TES ozone data have been extensively validated against ozonesondes for a wide range of geophysical states and latitudes. Studies have shown that there are no observable changes in biases in the TES ozone data over time, and the bias is well characterised as a function of latitude [*Worden et al.*, 2007; *Verstraeten et al.*, 2013]. In addition, the sonde comparisons indicate that the calculated random errors are in agreement with actual errors. This means that evaluating other satellite measurements against TES provides an assessment of instrument bias rather than unquantified errors in the TES retrieval. TES is extensively used for the evaluation of upper tropospheric ozone and its precursors in chemistry transport models [*e.g.*, *Jones et al.*, 2009]. In addition, since TES measures over the entire wavelength range of ozone infrared absorption, it can provide the sensitivity of outgoing longwave radiation (OLR) to the vertical distribution of ozone [*Worden et al.*, 2008; *Worden et al.*, 2011]. As part of the Atmospheric Chemistry Climate Model Intercomparison Project (ACCMIP) TES tropospheric ozone and its effect on OLR have been compared to the same quantities derived from models and used to reduce uncertainties in ozone radiative forcing [*Bowman et al.*, 2013]. Assessing the differences between TES and other instruments measuring ozone in the UTLS region will provide a better understanding of the ozone gradients and variability that TES fails to capture due to its coarse resolution. Furthermore, improved characterisation of satellite measurements of ozone in this region will allow us to better quantify the significance of model-measurement differences in precursor emissions and radiative forcing in the UTLS.

4.27.1 Availability of UTLS ozone satellite datasets

The time period used for this analysis is determined by the availability of TES data and covers July 2005 to December 2010. TES provides global coverage from July 2005 through May 2008. To extend the life of the instrument, the latitudinal coverage was reduced in June 2008 to 60°S-82°N and in July 2008, to 50°S-70°N. From January to April 2010, the instrument went offline due to problems with the scanning mechanism. When operations resumed in May 2010, the latitude coverage was further reduced to 30°S-50°N. A second data gap of ~3 weeks occurred in October 2010, with only two Global Surveys conducted that month. **Tables 4.27.1** and **4.27.2** provide information of the availability of UTLS ozone measurements for 2005-2010 as well as the data version, vertical range, vertical resolution, and references for the instruments discussed in this section.

4.27.2 TES ozone and operational operator

TES ozone measurements have been extensively validated against ozonesondes [*Nassar et al.*, 2008; *Boxe et al.*, 2010; *Verstraeten et al.*, 2013]. In the 300-70 hPa region evaluated here, TES is positively biased with respect to sondes in all latitude regions except the southern low- and mid-latitudes (15°S-60°S), where it is negatively biased. The mean bias is smaller than 20% in all latitude regions. In the northern mid-latitudes, the bias is +15-20% for 100 hPa < p < 300 hPa and <+5% for 70 hPa < p < 100 hPa. The bias curve is "c-shaped" in the southern mid-latitude UTLS, with near-zero bias at 300 and 70 hPa and a maximum value of -20% at 150 hPa. In the tropical UTLS, TES shows a small positive bias (<10%) with respect to sondes. An analysis of seasonal variations in the northern mid-latitude (35°N-56°N) bias showed relatively small seasonal differences, except during summer when the bias decreases to <10% everywhere. Here, we include a comparison of the SPARC Data Initiative climatologies to a

Table 4.27.1: Available ozone measurements between 2005 and 2010 from limb-sounding satellite instruments and the nadir-viewing TES instrument. *The red filling of the grid boxes indicates the temporal (January to December) and vertical (300 to 70 hPa) coverage of the respective instruments.*

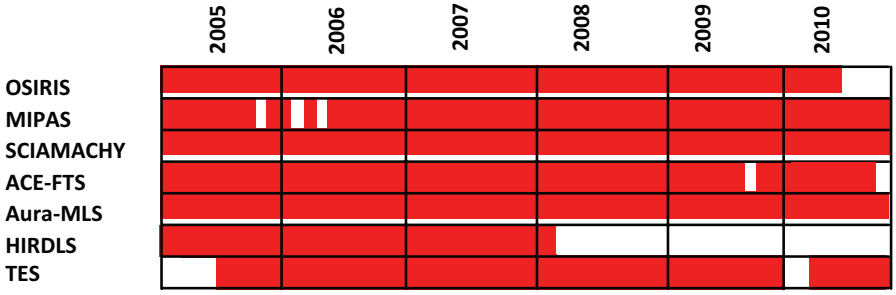


Table 4.27.2: Data version, vertical range, vertical resolution, and references for ozone datasets used for UTLS evaluations.

Instrument and data version	Vertical range	Vertical resolution	References
ACE-FTS V2.2update	5 – 95 km	3 – 4 km	*Dupuy et al.,* 2009
Aura-MLS V2.2	12 – 75 km	3 km	*Froidevaux et al.,* 2008; *Jiang et al.,* 2007
HIRDLS V6.0	10 – 55 km	1 km	*Nardi et al.,* 2008; *Gille et al.,* 2008
MIPAS V220	6 – 70 km	2.7 – 3.5 km	*von Clarmann et al.,* 2009a
OSIRIS V5-0	10 – 60 km	2 km	*Degenstein et al.,* 2009
SCIAMACHY V2.5	10 – 60 km	3 – 5 km	*Mieruch et al.,* 2012
TES V4	0 – 35 km	6 – 7 km	*Worden et al.,* 2004; *Boxe et al.,* 2010

"zonal mean" ozonesonde climatology (*Sections 4.27.3* and *4.27.4*) and find different biases for the TES climatology than those reported in the TES validation literature in some regions. These differences likely result from not accounting for 1) the sampling locations of the ozonesonde profiles and 2) the difference in vertical resolution between TES and the sondes in the climatological comparisons.

Use of zonal mean monthly mean averaging kernels

TES retrievals use the optimal estimation technique [*Rodgers*, 2000], with the retrieved profile, $\hat{x}$(ln(vmr)), given by:

$$\hat{x} = x^a + A^{xx}(x - x^a)$$

where x(ln(vmr)) is the true state, x^a(ln(vmr)) is the *a priori* profile, and A^{xx} is the averaging kernel matrix. For the comparisons shown here, the climatologies of the higher vertical resolution limb viewing instruments are taken to be the "true" state, x, and the TES observational operator (*a priori* and averaging kernel) are used to simulate a "virtual" TES retrieval, $\hat{x}$. Normally, this type of comparison is done on a profile-by-profile basis. However, due to the large number of instruments involved in this comparison and the focus on zonal mean climatologies, we apply the monthly mean zonal mean observational operator to the monthly mean zonal mean SPARC Data Initiative climatologies. The use of monthly mean zonal mean averaging kernels can be justified by the fact that the variations in TES averaging kernels are not highly correlated with variations in ozone. In the troposphere, ozone explains less than 25% of the variance in the averaging kernel diagonal at all latitudes due to the strong dependence of the averaging kernels on clouds, water vapor, and temperature, as discussed by *Aghedo et al.* [2011]. In the UTLS region, ozone explains up to 35% of the variance in the averaging kernel diagonal in mid-latitudes, with a minimum value at ~150-200 hPa where the sensitivity is relatively low and the *a priori* has a significant impact on the retrievals (see *Section 4.27.3*). In the tropical UTLS, ozone explains 20-60% of the variance in the averaging kernel diagonal, with maximum correlation at ~150 hPa. However, at all latitudes the dependence of the averaging kernel diagonal on ozone abundance is weak for ozone within ±40% of the mean value at each level; the correlations are primarily driven by ozone abundances more than 40% higher than the mean value. In the mid-latitudes, ozone abundances that are twice as large as the mean value at a given pressure level have a ~30% higher averaging kernel diagonal value. In the tropics, the slope of the relationship is somewhat higher, and a 100% increase in ozone over the mean value is associated with a 45% larger averaging kernel diagonal.

Aghedo et al. [2011] examined the error associated with using monthly mean averaging kernels in two climate models for p ≥ 100 hPa. They found differences in ozone of at most 3% when using monthly mean as compared to time-varying averaging kernels. To test the error involved in applying zonal mean averaging kernels to zonal mean data, we examined the difference between TES and Aura-MLS measurements for 2006. In the first case, we use 5°x10° gridded TES averaging kernels to smooth 5°x10° gridded Aura-MLS measurements and calculate zonal mean differences afterwards. In the second case, we use zonal mean TES averaging kernels to smooth zonal mean Aura-MLS data. The difference between the two cases is always smaller than 10%, and the difference in the zonal mean datasets is always smaller than the difference in gridded datasets (except in high southern latitudes during October). In addition, the difference between using zonal

mean averaging kernels aggregated from individual profiles and using the zonal mean of the gridded averaging kernels is negligible (< 2% everywhere). We therefore conclude that using zonal mean averaging kernels with zonal mean data provides a lower estimate that is within ~10% of the true difference between each instrument and TES. However, we note that because the averaging kernels are not fully independent of the ozone abundance, comparison using the TES observational operator may not accurately reflect the difference between TES and another instrument if there are large systematic differences between them. Given the fact that the averaging kernels depend only weakly on the ozone abundance for ozone within 40% of the mean value, we do not expect this to be an issue except where instruments differ from TES by more than 40%. In such cases, which are rare (see the discussion of **Figure 4.27.4** below), the error associated with using an averaging kernel that has sensitivity not appropriate for the ozone observed by the other instrument can only be quantified by recalculating the averaging kernel to "match" the instrument's ozone, which is beyond the scope of this report.

Applying the TES observational operator

For each instrument, we interpolate the monthly mean zonal mean climatologies from the SPARC pressure grid to the TES retrieval levels (67 levels between the surface and 0.1 hPa). We fill in the levels below the lowest measurement in each latitude bin (at pressure p_{max}) using the monthly mean, zonal mean TES *a priori* as a "fill profile". The "virtual" TES retrievals are calculated and then interpolated back to the SPARC pressure grid, and we average over all of the available data from 2005-2010 to create the climatologies shown here. We use the *a priori* as a fill profile because it makes $A(x - x^a) = 0$ in the "troposphere" (defined as $p \geq p_{max}$ for each instrument) since $x = x^a$ there, which is equivalent to applying the observational operator only to the levels where the limb-viewing instruments provide measurements. However, the fill profile can still impact the comparison to TES due to the vertical smearing of the averaging kernels. The difference between the virtual retrieval for a given instrument ($\hat{x}_{INST}$) and TES ($\hat{x}_{TES}$) can be written as

$$\hat{x}_{INST} - \hat{x}_{TES} = A^{SS}(x_{True}^{STRAT} - x_{INST}^{STRAT}) - A^{ST}(x_{True}^{TROP} - x_{INST}^{TROP})$$

where A^{SS} is the "stratospheric" component of the averaging kernel matrix ($p < p_{max}$), x_{INST}^{STRAT} the ozone profile measured by the limb-viewing instrument, A^{ST} represents the cross-terms of the averaging kernel that define the tropospheric influence on the stratosphere, and x_{INST}^{TROP} is the fill profile. To test the sensitivity of our results to our approach of using the TES *a priori* to fill in the profiles below the lowest measurement level, we have also calculated virtual retrievals in which we scale the TES *a priori*. We do so by multiplying the *a priori* profile by the percent difference between the individual instrument's ozone value and the TES *a priori* ozone at p_{max} for each latitude bin. Comparison of the virtual retrievals using the two different filling methods allows us to identify regions where our results are highly dependent on our assumptions for $p > p_{max}$, as discussed below.

4.27.3 UTLS ozone evaluations: Zonal mean cross sections, vertical and meridional profiles

As throughout *Chapter 4*, we use a series of diagnostics to evaluate differences in the vertical, latitudinal, and temporal structure of ozone as represented by the SPARC Data Initiative climatologies. However, rather than examining differences from the multi-instrument mean, we use the TES climatology as the standard to which the other climatologies are compared and, in some cases, include climatological ozonesonde measurements as an additional validation tool. We also analyse the impact of the TES observational operator on the climatologies from the limb-viewing instruments and assess how the use of the observational operator affects the ozone inter-comparison.

The second column of **Figure 4.27.2** shows the zonal mean ozone climatology for each instrument from 300 to 70 hPa averaged over 2005-2010 (2005-2007 for HIRDLS) using the data directly from the SPARC Data Initiative archive. All of the instruments show similar features, including the typical low tropical values, strong subtropical gradients, and relatively flat mid-latitude isopleths that reflect the competing effects of the stratospheric overturning circulation and mixing with the troposphere. The instruments also all show lower ozone values in the Southern Hemisphere than in the Northern Hemisphere in the annual mean due to the asymmetry in the overturning circulation. A few instruments show features not seen in the climatologies from any of the other instruments. The MIPAS climatology has an unusual contour shape in the tropics between ~200 and 100 hPa, with a slight "double ear" structure in the subtropics and a deep minimum near the equator, and Aura-MLS has very flat, tightly spaced contours near 100 hPa. It is likely that some of the differences in the climatologies in the upper tropical troposphere arise from differences in the impact of clouds on the retrievals and in criteria used for cloud screening, which can cause sampling artefacts. The OSIRIS climatology has an unusually strong zonal gradient at ~75°N below 250 hPa, which appears to reflect sampling bias in the climatology resulting from a lack of measurements in polar winter [*Toohey et al.*, 2013].

Impact of TES observational operator

The third column of **Figure 4.27.2** shows the virtual retrievals using the TES observational operator, and the fourth column shows the percent difference between the virtual TES retrieval (VTR) and the original climatology (OC) for each instrument (shown in the left column) (100*(VTR-OC)/OC). Hatched regions in the right column indicate where the choice of fill profile (TES *a priori* or scaled *a priori*) has a significant impact on the VTR, quantified (arbitrarily) as where the difference between the VTRs using the two fill profiles exceeds 10%. The HIRDLS climatology shows the

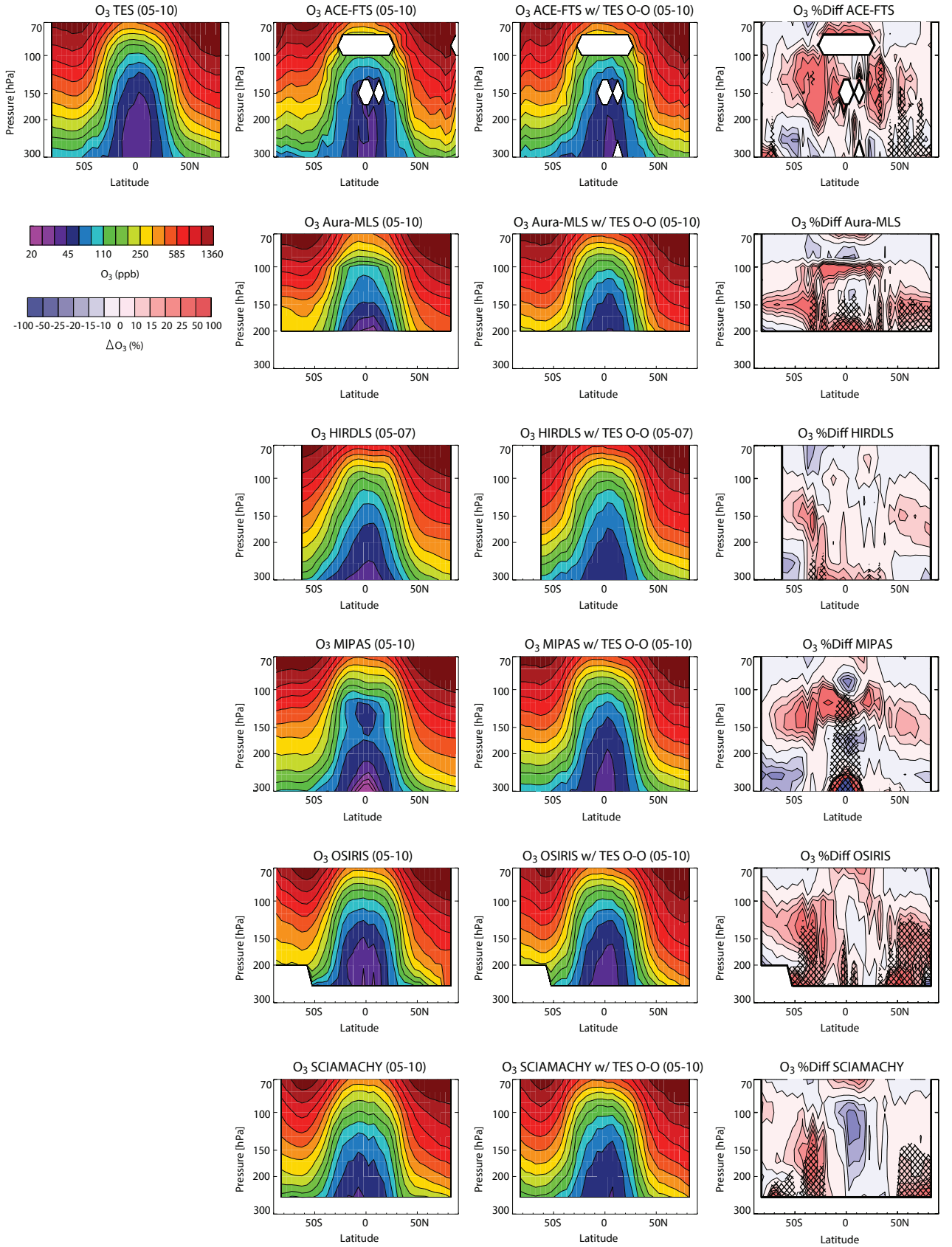


Figure 4.27.2: Cross sections of annual mean zonal mean ozone from 300 to 70 hPa for 2005-2010. *First column: Ozone cross section from TES. Second column: Ozone cross sections from ACE-FTS, Aura-MLS, HIRDLS (2005-2007), MIPAS, OSIRIS, SCIAMACHY. Third column: Cross sections from ACE-FTS, Aura-MLS, HIRDLS, MIPAS, OSIRIS, SCIAMACHY after application of the TES observational operator. Fourth Column: Percent change in annual mean zonal mean ozone introduced by the TES observational operator (100*(VTR – OC)/OC, where VTR=virtual TES retrival and OC=Original Climatology). Hatching indicates regions where the difference between the virtual retrievals using the TES a priori as the fill profile and those using the scaled a priori as the fill profile exceeds 10%. See text for details.*

most uniform and smallest changes in ozone after the application of the TES observational operator. The operator acts to smooth out the small-scale features seen in the MIPAS, Aura-MLS, and OSIRIS climatologies, as seen in the center column, due to the vertical smearing of the broad averaging kernels. In the tropics, the observational operator tends to increase ozone for $p \leq 80$ hPa and decrease it for $p > 80$ hPa relative to the original climatologies (with strong increases at 100 hPa associated with the unusual features in MIPAS and Aura-MLS). MIPAS and Aura-MLS are the only two instruments for which the choice of the fill profile has a significant impact in the tropics. It is unclear why this is the case for Aura-MLS, but the MIPAS tropical ozone values are very low at $p > 250$ hPa compared to the other instruments, and there is a very large difference between the TES *a priori* and the *a priori* scaled using the MIPAS measurements in this region.

In the extra-tropics, the observational operator tends to increase ozone at ~150 hPa and decrease it above and below, which increases the vertical and horizontal gradients of ozone in the virtual retrievals compared to the original climatologies. This increase in gradient clearly cannot result from the TES averaging kernels, which smooth and flatten vertical gradients. Rather the increase results from the influence of the *a priori*; comparison of the terms Ax, the contribution of the "true" profile from each climatology to the virtual retrieval, and, $x^a - Ax^a$, the contribution of the TES *a priori* to the virtual retrieval, (**Figure 4.27.1b**) shows that TES's sensitivity is lowest and the *a priori* profile makes the largest contribution to $\hat{x}$ in the mid-latitudes at ~150-200 hPa, as well as in the southern high latitudes at $p > 150$ hPa.

The sensitivity to the fill profile is largest in the extra-tropics, in particular for the climatologies whose range does not extend to 300 hPa (Aura-MLS, OSIRIS, and SCIAMACHY). Between ~200 and 300 hPa there are large vertical gradients in mid-latitude ozone that are not well-represented by the TES *a priori*, so that there is a large difference between the two fill profiles (the *a priori* and the scaled *a priori*). Furthermore, the averaging kernels spread the information from 300 hPa upward to ~100 hPa in the extra-tropics, so that changing ozone at 300 hPa has a significant influence over a large vertical range.

Figure A4.27.1 in *Appendix A4* shows the difference between the VTRs and OCs for January, April, July, and October. The results for the individual months are similar to the annual mean results shown here, with the exception of October. Because TES has relatively low sensitivity in southern high latitudes (see **Figure 4.27.1b** and discussion thereof) and the *a priori* does not adequately capture the characteristics of Antarctic ozone depletion, the application of the observational operator results in strong increases in ozone for $p <$~170 hPa and strong decreases in ozone for $p >$~170 hPa at latitudes poleward of 50°S in the virtual retrievals of all of the instruments. We also note a greater tendency for the difference between the VTRs using the two fill profiles to exceed 10% in the Southern Hemisphere during January-April and in the Northern Hemisphere during July-October.

Figure 4.27.3 shows a comparison of zonal mean vertical profiles in the northern mid-latitudes (April), tropics (annual mean), and southern mid-latitudes (October). The left column is once again the original SPARC Data Initiative climatology for each instrument, while the centre and right columns show the virtual retrievals using the TES observation operator and the percent difference between the virtual retrievals and the original data (100*(VTR-OC)/OC), respectively. Dashed lines in the right column indicate where the choice of fill profile affects the VTR by more than 10%. In the mid-latitudes, the observational operator smoothes the vertical profiles; it decreases ozone for $p >$150 hPa and increases it for $p <$ 150 hPa for all instruments for which the virtual retrievals do not depend strongly on the fill profile. In the tropics, as discussed above, the observational operator acts to slightly increase ozone at $p <$ 80 hPa and slightly decrease it at $p >$ 80 hPa, as well as to smooth small-scale vertical structures.

Percent difference from TES

Figure 4.27.4 shows the percent difference between the annual mean climatology for each instrument and TES (100*(OC-TES)/TES, left column) and the percent difference between the virtual retrievals for each instrument and TES (100*(VTR-TES)/TES, right column). The hatched regions in the right column indicate where differences in the VTR due to the choice of fill profile exceed 50% of the difference between the VTR and TES for each instrument. While 50% is an arbitrary choice, it highlights a combination of regions where the fill profile has a relatively large impact on the virtual retrievals (see **Figure 4.27.2** above) and regions where differences between the virtual retrievals and TES are small so that even small differences due to the fill profile are large relative to VTR-TES. The original climatologies from all instruments except ACE-FTS and OSIRIS show positive differences of more than 25% with respect to TES in the tropics. However, except for HIRDLS, which is affected by uncorrected emission from aerosol in the tropics (J. Gille, private communication, 2013), the biases are not uniform in pressure and there are some regions with negative biases, including for $p <$ 70 hPa (which can impact the region of interest when the TES observational operator is applied). The virtual retrievals represent the combined influence of the vertical smoothing of the TES averaging kernel and the *a priori*, whose influence is not negligible due to TES's imperfect sensitivity. Together these act to both vertically smooth the differences from TES and reduce them to ≤ 25% for the virtual retrievals from all instruments except HIRDLS. However, for the Aura-MLS and MIPAS virtual retrievals, the biases with respect to TES are robust (*i.e.,* not strongly dependent on the fill profile) only for $p <$ ~100 hPa. For ACE-FTS, which is a solar occultation instrument and has very sparse sampling in the tropics due to its orbit, the difference from TES may largely reflect a >5% tropical sampling bias in the climatology [*Toohey et al.*, 2013].

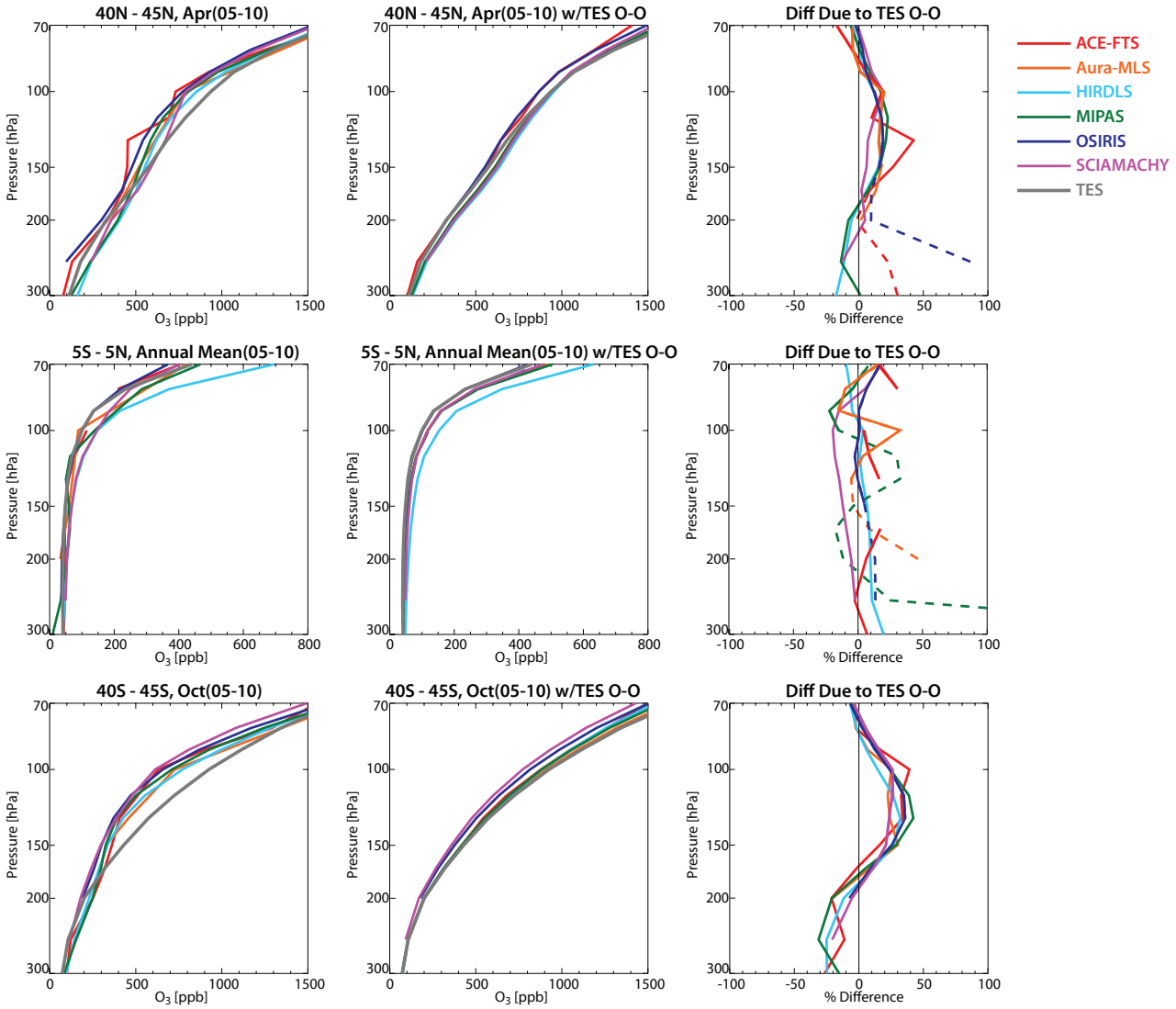


Figure 4.27.3: Vertical profiles of zonal mean ozone for 2005-2010. *Left column: Ozone profiles from the original climatology for each instrument. Centre column: Ozone profiles after application of the TES observational operator. TES measurements are the same as in the left column. Right column: The percent change in the ozone profiles introduced by the TES observational operator for all instruments except TES (100*(VTR-OC)/OC). Dashed lines indicate portions of the profile where the difference between the virtual retrievals using the TES a priori as the fill profile and those using the scaled a priori as the fill profile exceeds 10%. Top row: Zonal mean ozone profiles and differences for 40°N-45°N for April 2005-2010. Centre row: Annual mean zonal mean ozone profiles and differences for 5°S-5°N for 2005-2010. Bottom row: Zonal mean ozone profiles and differences for 40°S-45°S for October 2005-2010.*

HIRDLS and MIPAS also have annual mean positive differences of 10-30% with respect to TES at $p \geq 150$ hPa in the northern mid- and high latitudes. Again, the differences with respect to TES can be seen in the original climatologies, but the vertical extent of the positive biases is greater in the virtual retrievals. The climatologies from the other instruments also show positive differences from TES in the same region, but for the most part these are not seen in the original climatologies and are an artefact of the impact of the fill profile. In the Southern Hemisphere, the original climatologies are generally negatively biased with respect to TES, especially above 200 hPa. The TES observational operator strongly reduces the difference between the original datasets and TES in the Southern Hemisphere, such that the virtual retrievals agree with the TES climatology to within ~10%. This is likely because the differences between the original climatologies and TES occur largely in the region where TES has low sensitivity and the *a priori* plays an important role in the virtual retrievals (**Figure 4.27.1b**). OSIRIS and SCIAMACHY, which measure only in the sunlit portion of the atmosphere, have >5% negative sampling biases in their climatologies in the southern mid- and high-latitudes [*Toohey et al.*, 2013], which may at least partially explain their larger differences with respect to TES relative to the climatologies from other instruments.

Figure A4.27.2 in *Appendix A4* shows the percent difference between the virtual retrievals for each instrument and TES for the same months as shown in **Figure A4.27.1** in *Appendix A4*. The differences for individual months are generally similar to those for the annual mean (**Figure 4.27.4** right column). The positive difference between MIPAS and

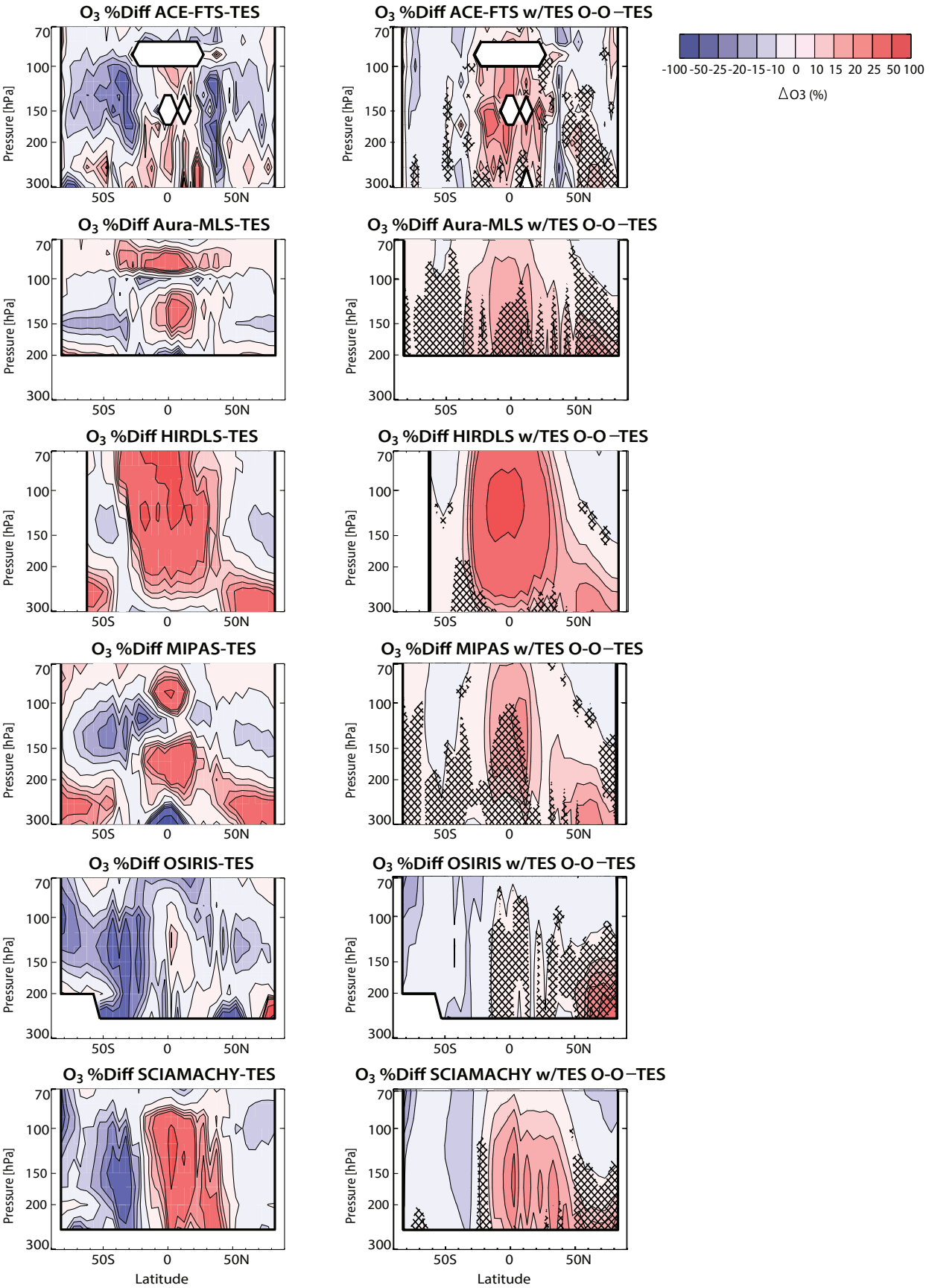


Figure 4.27.4: Cross sections of annual mean zonal mean ozone differences from 300 to 70 hPa for 2005-2010. *Left column: Annual mean zonal mean ozone percent differences between the climatology from each instrument and TES for 2005-2010 (HIRDLS: 2005-2007) (100*(OC-TES)/TES). Right column: Percent differences between the virtual retrieval from each instrument and TES after application of the TES observational operator (100*(VTR-TES)/TES). Hatched regions indicate where the difference in the virtual retrieval using the two different fill profiles exceeds 50% of the difference between the virtual retrieval and TES.*

TES in the tropics is larger in January than throughout the rest of the year, and the positive difference between both HIRDLS and MIPAS and TES in the northern mid- and high latitudes is largest in July. SCIAMACHY is the only instrument with a large (positive) difference with respect to TES in southern high latitudes during October, despite the fact that the virtual retrieval is strongly influenced by the *a priori* in that region and thus not expected to differ greatly from TES (see discussion above). The SCIAMACHY climatology has the highest October ozone values of any of the climatologies from ~150 hPa ≤ p ≤ ~100 hPa and ~70-80°S (see **Figure A4.27.5** in *Appendix A4*, discussed below), where there is a local maximum in TES Southern Hemisphere sensitivity (see **Figure 4.27.1b**). This maximum in sensitivity preserves the high ozone values in the SCIAMACHY virtual retrieval and spreads them vertically throughout the UTLS.

Latitudinal gradients on pressure surfaces

Figure 4.27.5 shows the 2005-2010 mean April ozone from each instrument as a function of latitude on four pressure surfaces. The original datasets (first column), virtual retrievals (second column), and percent differences between the original datasets and TES (100*(OC-TES)/TES, third column) and virtual retrievals and TES (100*(VTR-TES)/TES, fourth column) are all shown. Dashed lines in the fourth column indicate where differences in the VTR due to the choice of fill profile exceed 50% of the difference between the VTR and TES for an instrument. In addition to the satellite climatologies, the first column also includes a "zonal mean" ozone climatology from ozonesonde measurements at 48 stations from the datasets described by *Logan* [1999] (representative of 1980–1993) and *Thompson et al.* [2003] (representative of 1997–2011) (**Table 4.27.3**). We note that there are at most 4 ozonesonde stations in a given latitude band, and many latitude bands contain only one station, likely leading to large sampling biases. Furthermore, no attempt has been made to account for differences in vertical resolution between the satellites and the sondes, primarily because it is unclear whether the use of zonal mean averaging kernels would exacerbate the sampling bias. Nevertheless, we include the ozonesonde climatology to demonstrate the good agreement between the satellites and the sondes and to provide an additional tool to investigate biases in the satellite climatologies.

At p ≤ 200 hPa, the absolute differences between the climatologies are mostly small, except at high latitudes (>50°N/S, **Figure 4.27.5**, first column). Given the large ozone abundance at high latitudes, however, the absolute differences between the instruments translate to small relative differences; the limb-viewing instruments agree with each other and with TES (**Figure 4.27.5**, third column) to within ~30% at mid- and high latitudes for p ≤ 200. The ozonesonde measurements suggest that the TES climatology is positively biased by ~20% in the Northern Hemisphere extra-tropics for 80 < p < 200 hPa, in good agreement with the TES validation studies. In the Southern Hemisphere extra-tropics, the TES climatology shows positive biases of 20-30% with respect to the ozonesonde climatology for p > 100 hPa and negative biases of 20-30% for p < 100 hPa. The inconsistency between this comparison and the validation results discussed in *Section 4.27.2* likely arises from the sparse ozonesonde coverage in the Southern Hemisphere, the fact that we have not applied the TES operator to the ozonesonde measurements, and the comparison here being limited to a single month (results for the seasonal cycle are discussed in *Section 4.27.4*). For latitudes > ~40°N/S and p ≤ 200 hPa, it appears that the differences between the climatologies from the limb instruments and the TES climatology likely reflect biases in TES rather than any significant bias in the limb sounders' climatologies. The direct comparison of the satellite climatologies to TES, using the TES operational operator, results in agreement of all of the satellite climatologies at the 10% level for this region (**Figure 4.27.5**, fourth column). However, in reducing the difference between the limb sounders and TES, the use of the observational operator also introduces some of the TES bias into the virtual retrievals.

In the tropics, where ozone abundances are low, small absolute differences translate into large relative differences, as also seen in **Figure 4.27.4.** The limb-viewing instruments differ from one another and from TES by up to 90% in the tropics and subtropics (**Figure 4.27.5**, third column). The TES climatology does not differ systematically from the ozonesonde climatology in the tropics for p < 200 hPa, except at 150 hPa, where it is positively biased by ~25% in April. Thus, the greater ozone abundances in the limb sounder climatologies represent an over-estimate of tropical ozone throughout most of the UTLS for this month. However, although differences between the annual mean climatologies are generally similar to or smaller than the differences shown here for April, the annual mean TES climatology at 150 hPa is actually biased low by >20% relative to the annual mean ozonesonde climatology over much of the tropics (see also *Section 4.27.4*, discussion of **Figures A4.27.3-A4.27.5** in *Appendix A4* below). Again, this contradicts coincident validation studies showing a small positive bias for TES throughout the tropical UTLS. The annual mean positive differences between the limb sounder climatologies and TES seen in **Figure 4.27.4** thus likely reflect true positive biases for the limb climatologies only for p < 150 hPa. With the exception of HIRDLS, which has high ozone values over a deep vertical extent (thus limiting the impact of smoothing), the TES observational operator greatly reduces the differences between the climatologies, with agreement to within ~30% in the tropics and subtropics. The comparison to TES is most robust for p < 100 hPa, where the virtual retrievals are relatively free of influence from the fill profile. While the annual mean pattern of differences between HIRDLS and TES is more or less centred at the equator (**Figure 4.27.4**), the HIRDLS climatology shows the largest differences from TES and from the other climatologies in the southern subtropics in April, suggesting perhaps a seasonal variability in the aerosol effect on the ozone retrievals.

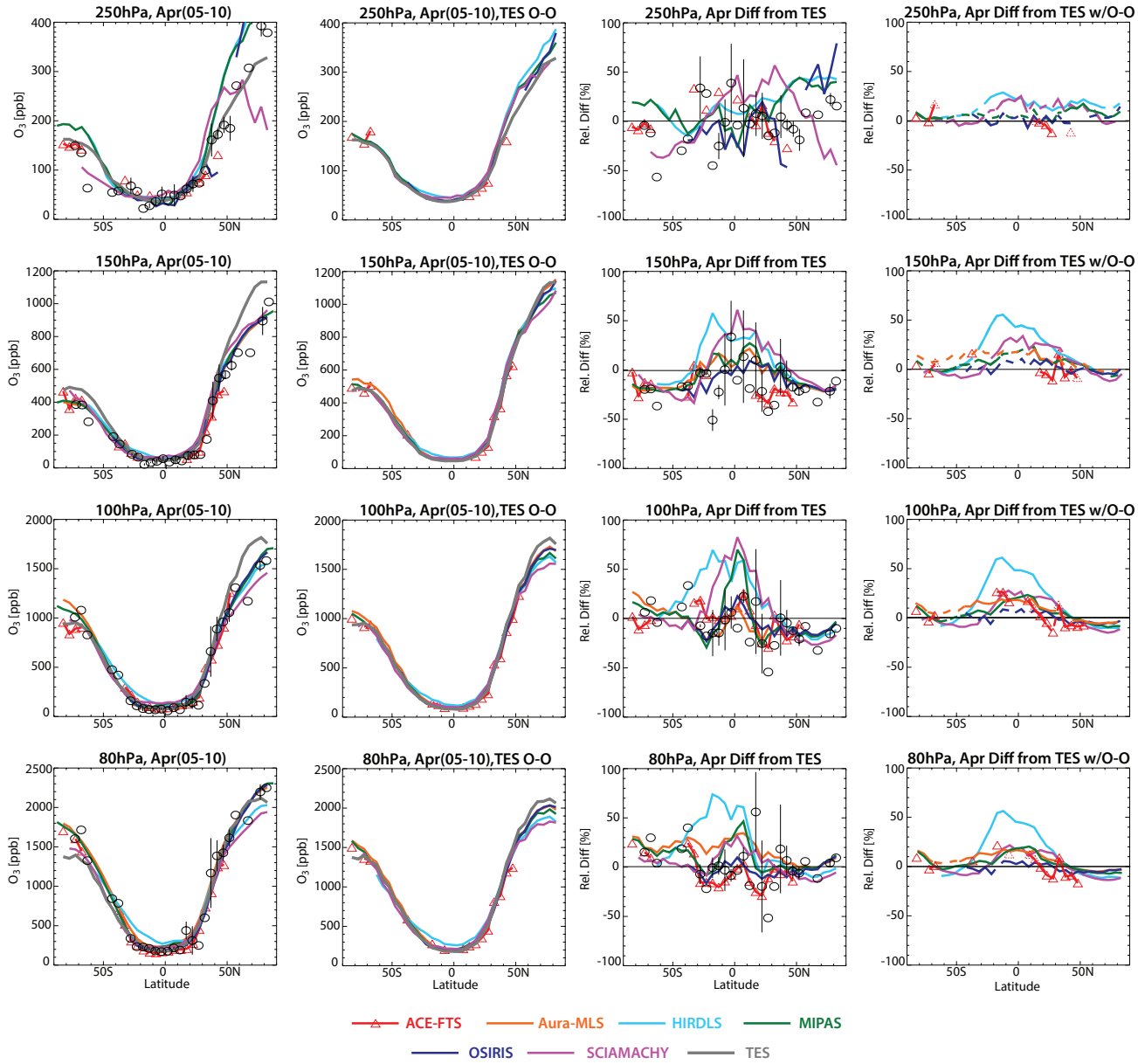


Figure 4.27.5: Meridional profiles of monthly mean zonal mean ozone for April 2005-2010. *First column: Meridional zonal mean ozone profiles from the climatology for each instrument at 250 hPa (top row), 150 hPa (second row), 100 hPa (third row), and 80 hPa (bottom row) for April 2005-2010. Black circles show the ozonesonde climatology; vertical bars represent the standard deviation of climatological mean values for latitude bands with more than one station. Second column: Meridional profiles after application of the TES observational operator to climatologies from each instrument. The TES measurements are the same as in the first column. Third column: Percent difference between each instrument and TES as a function of latitude on each pressure surface. (100*(OC-TES)/TES) Black circles show the ozonesonde climatology; vertical bars are as above. Fourth column: Same as third column, but for virtual retrievals with the TES observational operator applied (100*(VTR -TES)/TES). Dashed lines indicate portions of the virtual retrieval where the difference in the virtual retrieval using the two different fill profiles exceeds 50% of the difference between the virtual retrieval and TES.*

Figures A4.27.3-A4.27.5 in *Appendix A4* show results for January, July, and October. As discussed above, TES is biased low relative to the tropical sonde measurements at 150 hPa in each of these months, while the limb sounders are in good agreement with the sondes. In January, TES is also biased high relative to the sondes in the southern extra-tropics at this same pressure level; the limb sounders again agree well with the sonde measurements. July shows the best overall agreement among the climatologies in the tropics, both for the original climatologies and the virtual TES retrievals. During October the limb sounders are in much better agreement with the sondes in the southern extra-tropics than TES is, and the decrease in ozone for $p > 170$ hPa and increase in ozone for $p < 170$ hPa resulting from application of the TES observational operator at southern high latitudes (see discussion of **Figure A4.27.1** in *Appendix A4*) introduces negative biases and positive biases, respectively, in the VTRs.

Table 4.27.3: Station information for the ozonesonde climatology. *The latitude, longitude, average number of soundings per month and length of data record for each ozonesonde station used in the climatology is given, along with the zonal mean latitude bins (which are the same as those used for the satellite climatologies). The data are from Logan et al. [1999] (denoted by *), and Thompson et al. [2003] (denoted by †). Black rectangles show the stations that are averaged to calculate the seasonal cycle in the tropics (15°S-15°N) and northern and southern mid-latitudes (40°N-45°N and 40°S-45°S) in Figure 4.27.6. We include Asp. Laverton, located at 38°S, in the calculation of the southern mid-latitude seasonal cycle to avoid the use of a single station.*

Station name	Latitude	Longitude	Soundings/ month	Data record	Latitude bin
Forster*	71°S	12°E	28	1985-1991	70-75°S
Syowa*	69°S	39°E	18	1986-1993	65-70°S
Marambio*	64°S	57°W	20	1988-1995	60-65°S
Lauder*	45°S	170°E	24	1986-1990	45-50°S
Asp. Laverton*	38°S	145°E	24	1980-1995	35-40°S
Pretoria*	26°S	28°E	11	1990-1993	25-30°S
Irene†	26°S	28°E	19	1998-2012	25-30°S
La Reunion†	21°S	56°E	31	1998-2012	20-25°S
Suva†	18°S	178°E	22	1998-2011	15-20°S
Tahiti†	18°S	149°W	6	1998-1999	15-20°S
Am. Samoa†	14°S	170°W	37	1998-2012	10-15°S
Am. Samoa*	14°S	170°W	13	1986-1996	10-15°S
Ascension Island†	8°S	14°W	45	1998-2010	5-10°S
Watukosek†	8°S	113°E	24	1998-2012	5-10°S
Natal*	6°S	35°W	23	1978-1992	5-10°S
Natal†	5°S	35°W	39	1998-2011	5-10°S
Brazzaville*	4°S	14°E	7	1990-1992	0-5°S
Malindi†	3°S	40°E	8	1999-2006	0-5°S
Nairobi†	1°S	37°E	46	1998-2012	0-5°S
San Cristobal†	1°S	89.6°W	31	1998-2012	0-5°S
Kuala Lumpur†	3°N	102°E	23	1998-2012	0-5°N
Paramaribo†	6°N	55°W	33	1999-2012	5-10°N
Cotonou†	6°N	2°E	7	2005-2007	5-10°N
Panama*	9°N	80°W	4	1966-1969	5-10°N
Heredia†	10°N	84°W	6	2005-2012	10-15°N
Poona*	19°N	74°E	11	1966-1986	15-20°N
Hilo†	19°N	155°W	50	1998-2012	15-20°N
Hilo*	20°N	155°W	30	1985-1993	20-25°N
Ha Noi†	21°N	106°E	9	2004-2012	20-25°N
Naha*	26°N	128°E	15	1989-1995	25-30°N
Kagoshima*	32°N	131°E	19	1980-1995	30-35°N
Tateno*	36°N	140°E	37	1980-1995	35-40°N
Azores*	38°N	29°W	22	1983-1995	35-40°N
Cagliari*	39°N	9°E	25	1968-1980	35-40°N
Boulder*	40°N	105°W	27	1985-1993	40-45°N
Sapporo*	43°N	141°E	21	1980-1995	40-45°N
Sofia*	43°N	23°E	16	1982-1991	40-45°N
Biscarosse*	44°N	1°W	28	1976-1993	40-45°N
Payerne*	47°N	7°E	95	1980-1993	45-50°N
Hohenpeissenberg*	48°N	11°E	135	1980-1993	45-50°N

Table 4.27.3 continued.

Station name	Latitude	Longitude	Soundings/ month	Data record	Latitude bin
Lindenberg*	52°N	99°E	18	1980-1995	50-55°N
Edmonton*	53°N	114°W	41	1980-1993	50-55°N
Goose Bay*	53°N	60°W	45	1980-1993	50-55°N
Churchill*	59°N	147°W	43	1980-1993	55-60°N
Sodankyla*	67°N	27°E	20	1989-1992	65-70°N
Resolute*	75°N	95°W	45	1980-1993	75-80°N
Ny Alesund*	79°N	12°E	19	1990-1993	75-80°N
Alert*	83°N	62°W	29	1988-1993	80-85°N

4.27.4 UTLS ozone evaluations: Seasonal cycles

Figure 4.27.6 shows the seasonal cycle of ozone in the southern mid-latitudes, tropics, and northern mid-latitudes averaged over 2005-2010 for the original climatologies as well as for the virtual retrievals using the TES observational operator. Dashed lines in the right column of plots for each latitude region again indicate where differences in the VTR due to the choice of fill profile exceed 50% of the difference between the VTR and TES for an instrument. The ozonesonde climatology is included in the left column for each region. The seasonal variability of ozone is largely driven by seasonal changes in the Brewer-Dobson circulation [*Folkins et al.*, 2006; *Randel et al.*, 2007]. In the mid-latitudes, there is an annual cycle in ozone at all pressure levels, with maxima and minima in September and March, respectively, in the Southern Hemisphere and March and August, respectively, in the Northern Hemisphere [*Logan et al.*, 1999]. In the tropics, there is a weak semi-annual cycle driven primarily by mixing below ~150 hPa [*Konopka et al.,* 2010; *Ploeger et al.*, 2012], with maxima in June and September, transitioning to a strong single peak with a maximum in August at 100 hPa and above.

There are large differences among the climatologies in the timing and magnitude of the seasonal cycle in the tropical UT ($p \geq 100$ hPa). The OSIRIS climatology has the largest difference in peak-to-peak amplitude (>100%) relative to TES, while the SCIAMACHY climatology has the largest difference in timing, with a single broad peak from March to September. However, while the TES climatology seasonal cycle shows reasonable agreement with the sonde climatology at these levels (though with a general negative bias), the station-to-station variability in ozone from the sonde measurements is so large that the sonde climatology encompasses all of the satellite climatologies. The differences between the satellite climatologies are reduced in the comparison with the TES observational operator, so that the differences in seasonal cycle amplitude among all of the virtual retrievals are less than 50%, but are still much larger than in any other region. We note that the choice of fill profile significantly impacts most of the virtual retrievals in the tropical upper troposphere, so that, combined with the large variability in the sonde climatology, our conclusions are less robust here than anywhere else. We also note that, as discussed in *Section 4.27.3*, the difference between the TES and ozonesonde climatologies is smaller in April than any other time of the year at 150 hPa and that the TES climatology is negatively biased with respect to the sondes for $p \geq 150$ hPa.

In the tropics between 100 and 70 hPa, there are also large discrepancies between the satellite climatologies, with differences in the magnitude of the seasonal cycle of up to 85% relative to TES for the original climatologies. Here, however, the TES seasonal cycle agrees very well with that from the ozonesondes, and the variability in the ozonesonde climatology is largely diminished. The smoothing by the TES observational operator greatly improves the consistency in the seasonal cycle amplitude (to within 20% except for HIRDLS, which differs in peak-to-peak amplitude from TES by 25-45% at these pressure levels), with the largest impact at 100 hPa. In the original datasets, there are differences of up to 2 months in the timing of the ozone minimum and maximum; with the observational operator the consistency is improved to ±1 month. The changes in timing result from some combination of smoothing the seasonal cycle signal over a deep layer, with differences in phase throughout the layer, and the influence of the seasonal cycles in the TES a priori and averaging kernel.

There is excellent agreement among the satellite climatologies regarding the timing and magnitude of the seasonal cycle in northern mid-latitudes for $p < 200$ hPa; additionally, they all agree well with the ozonesonde climatology. For all instruments except ACE-FTS (which has limited sampling), the seasonal cycle peak-to-peak amplitude is consistent to within 25% for the original climatologies and for all instruments including ACE-FTS it is consistent within <5% using the TES observational operator. At $p \geq 200$ hPa, the MIPAS, SCIAMACHY, and HIRDLS climatologies have a 50-75% larger seasonal cycle than TES, whose climatology agrees well with the sondes. However, the magnitudes of the seasonal cycle in the satellite climatologies are consistent to within 5% when they are compared with the TES observational operator, with only

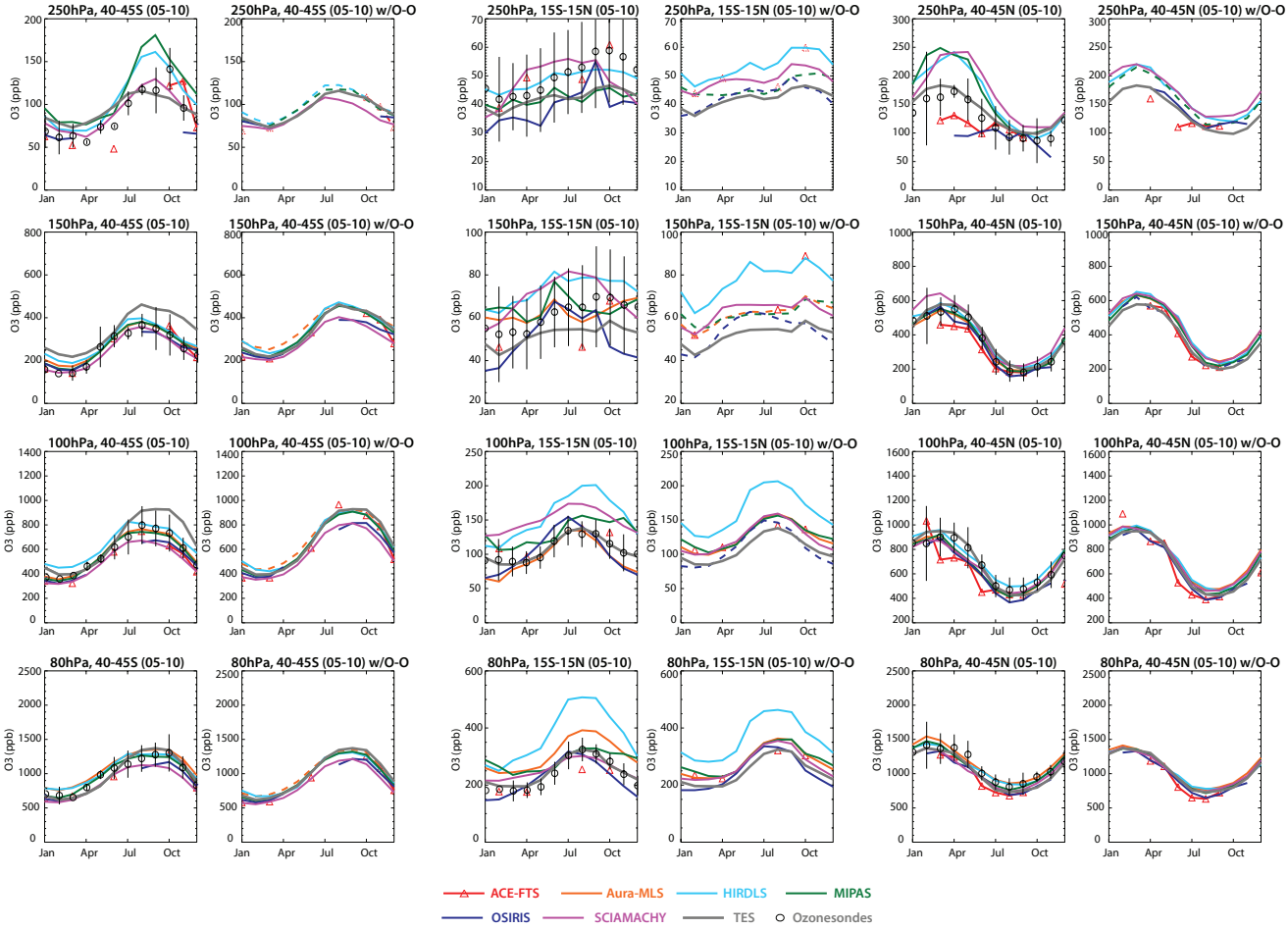


Figure 4.27.6: Seasonal cycle of ozone in the UTLS for 2005-2010. *Seasonal cycle of ozone from 40°S-45°S (two left columns), 15°S-15°N (two centre columns), and 40°N-45°N (two right columns) at 250 hPa (first row), 150 hPa (second row), 100 hPa (third row), and 80 hPa (fourth row). The left column in each grouping shows the seasonal cycle for each climatology, the right column in each grouping shows the seasonal cycle after application of the TES observational operator. The TES measurements are the same in both left and right columns of each group. Dashed lines in the figures in the right column of each group indicate portions of the virtual retrieval where the difference in the virtual retrieval using the two different fill profiles exceeds 50% of the difference between the virtual retrieval and TES in the annual mean. Black circles in the left columns of each grouping show the ozonesonde climatology; vertical bars represent the standard deviation of the climatological ozonesonde measurements from the stations in each latitude band.*

the MIPAS virtual retrievals being strongly dependent on the fill profile. In the southern mid-latitudes, the amplitude and timing of the TES seasonal cycle agree well with the ozonesonde climatology (though with a positive bias throughout the year at 150 hPa) at all levels except 100 hPa, where the ozone maximum in the TES climatology is almost 150 ppb larger than that seen in the ozonesondes and HIRDLS, MIPAS, and Aura-MLS climatologies. At $p \leq 100$ hPa, SCIAMACHY and OSIRIS have a 15% smaller seasonal cycle than those of the other instruments and sondes; the flatness results from an underestimate of the ozone maximum relative to the other climatologies and may be due to their limited sampling during winter. When the TES observational operator is applied to the climatologies, they agree to within 5% except for SCIAMACHY and OSIRIS. The vertical smoothing of the TES operator in fact spreads the low winter values in these climatologies to all of the pressure levels. We note that the increases in ozone at 150 and 100 hPa in the HIRDLS, MIPAS, and Aura-MLS virtual retrievals relative to the original climatologies represent another example of the TES operator introducing possible biases in the virtual retrievals.

Figure A4.27.6 in *Appendix A4* shows the seasonal cycles for high latitudes. In the northern high latitudes, results are generally similar to those for mid-latitudes; there is relatively good agreement between the climatologies for $p < 200$ hPa, particularly for the VTRs. However, TES is biased high relative to the limb sounders and to the sondes during the first half of the year for $p \leq 150$ hPa, and the observational operator transfers this bias to the limb sounder VTRs. In the southern high latitudes, the seasonal cycles at 250 hPa and $p \leq 150$ hPa are very different, with maximum values in January/February at 250 hPa and in July/August for $p \leq 150$ hPa. TES does a very poor job of reproducing the seasonal cycle from the sondes and the other instruments at 150 hPa. At $p \leq 100$ hPa, all of the instruments except ACE-FTS agree very well. ACE-FTS is much lower than the other instruments for August and September, but is in very good agreement with the sonde

climatology for these months. However, given the fact that only one sonde station is included here and that ACE-FTS samples this region less frequently than many of the other instruments, it is difficult to say anything conclusive about sampling differences versus biases in this case.

4.27.5 UTLS ozone evaluations: Interannual variability

Figure 4.27.7 shows the deviations from the 2005-2010 climatological monthly mean ozone for each instrument in the southern mid-latitudes, tropics, and northern mid-latitudes using the original climatologies as well as the virtual retrievals with the TES observational operator. As for the seasonal cycle, the differences between the climatologies are largest in the tropics and the TES observational operator damps out much of the smaller scale variability particularly for $p \geq 100$ hPa, and greatly improves the consistency in this region. Overall, the interannual variability in ozone is relatively low in the tropics, as expected, and the only signal that is observed by all of the instruments is a pronounced minimum in early 2010 throughout the UTLS region. This minimum can be seen in the original climatologies and is not an artefact of the TES observational operator. The low ozone values result from changes in convection and an increase in the Brewer-Dobson circulation associated with the 2009-2010 El Niño and coincident strong Easterly shear phase of the stratospheric quasi-biennial oscillation (QBO) [*Neu et al.*, 2014b]. While a tropical QBO signal is not particularly apparent for other years due to the noise in individual climatologies and the spread between them at these levels, we note that the timing of the minima and maxima in the HIRDLS climatology at $p \leq 100$ hPa is consistent with the QBO phase for 2005-2008 and that because the HIRDLS measurements are averaged over a shorter time period than the other instruments, the QBO signal is amplified in the HIRDLS record. However, it is not clear whether this fully explains the very large differences between HIRDLS and the other instruments for $p \leq 100$ hPa, or whether there might be an additional signal from QBO-related variability of the aerosol that affects the HIRDLS ozone retrievals. Regardless of the explanation, the TES observational operator spreads the information from $p \leq 100$ hPa downward so that it increases the apparent differences between HIRDLS and the other instruments for $p > 100$ hPa in the virtual retrievals. ACE-FTS also shows large differences in interannual variability from the other instruments at $p \leq 100$ hPa, likely due to its sparse sampling of the tropics. The interannual variability in OSIRIS ozone is somewhat noisier than that of the other instruments, even with the TES observational operator, but it is generally consistent with the other climatologies for $p < 100$ hPa, where the fill profile has little influence on the virtual retrieval.

The interannual variability is more consistent between the various climatologies in the northern and southern mid-latitudes than in the tropics, both in the original datasets and in the vertically smoothed virtual retrievals. As was the case for the seasonal cycle, the HIRDLS climatology and virtual retrievals agree very well with those from the other instruments in mid-latitudes, despite their large differences from the other instruments in the tropics. The largest discrepancies in mid-latitude interannual variability can be seen in the climatologies from ACE-FTS and OSIRIS in the Southern Hemisphere. In the case of ACE-FTS, the sampling is likely to blame, though we note that there may be a contribution from the fact that the lowest retrieval level (and thus the influence of the fill profile) varies throughout the year and between years more for ACE-FTS than for any other instrument. OSIRIS does not continuously sample the 40°S-45°S latitude band, so that the climatological monthly mean and deviations from the mean are not well-defined in Southern Hemisphere winter. The TES observational operator reduces the ozone variability somewhat in mid-latitudes, but the major deviations in northern mid-latitude ozone in 2008 and 2010 are well-preserved, except during the Jan-April 2010 TES data gap. The northern mid-latitude ozone minimum in 2008 and maximum in 2010 result from changes in the Brewer-Dobson circulation associated with La Niña / Westerly shear QBO and El Niño / Easterly shear QBO, respectively [*Neu et al.*, 2014b]. In the southern mid-latitudes, the climatologies all show maxima in 2005 and 2007 and minima in 2006 at $p \leq 100$ hPa. The TES observational operator reduces the maxima in the virtual retrievals due to the vertical smearing. TES stopped sampling south of 30°S in January 2010.

Figure A4.27.7 in *Appendix A4* shows the interannual variability at high latitudes. northern high latitudes have similar interannual variability as that seen in mid-latitudes, with the exception of noticeably low ozone values during 2005. Southern Hemisphere high latitude ozone variability is considerably larger than that in mid-latitudes, with pronounced minima in late 2006 and 2008 and maxima in late 2007 and 2010. While the TES southern high latitude record is quite short, the main features of the limb sounder climatologies are preserved in the virtual retrievals.

4.27.6 Summary and conclusions: UTLS ozone

While the use of zonal mean climatologies for detailed UTLS process studies is obviously limited, the SPARC Data Initiative climatologies nevertheless represent our best knowledge of the abundance and temporal variability of ozone in the UTLS and the characterisation of the datasets presented here will provide valuable information for model evaluation. Overall findings are presented in the following summary, and **Figure 4.27.8** is a synopsis plot showing the median, median absolute deviation, and standard deviation of monthly mean relative differences between measurements from the limb viewing instruments and TES in the tropics and mid-latitudes, both for the original climatologies and the virtual TES retrievals. **Figure 4.27.8** clearly shows the reduction in the difference between the limb sounders and TES as well as the reduction in variance when using the TES observational operator. It also highlights the better agreement among the instruments in mid-latitudes relative to the tropics.

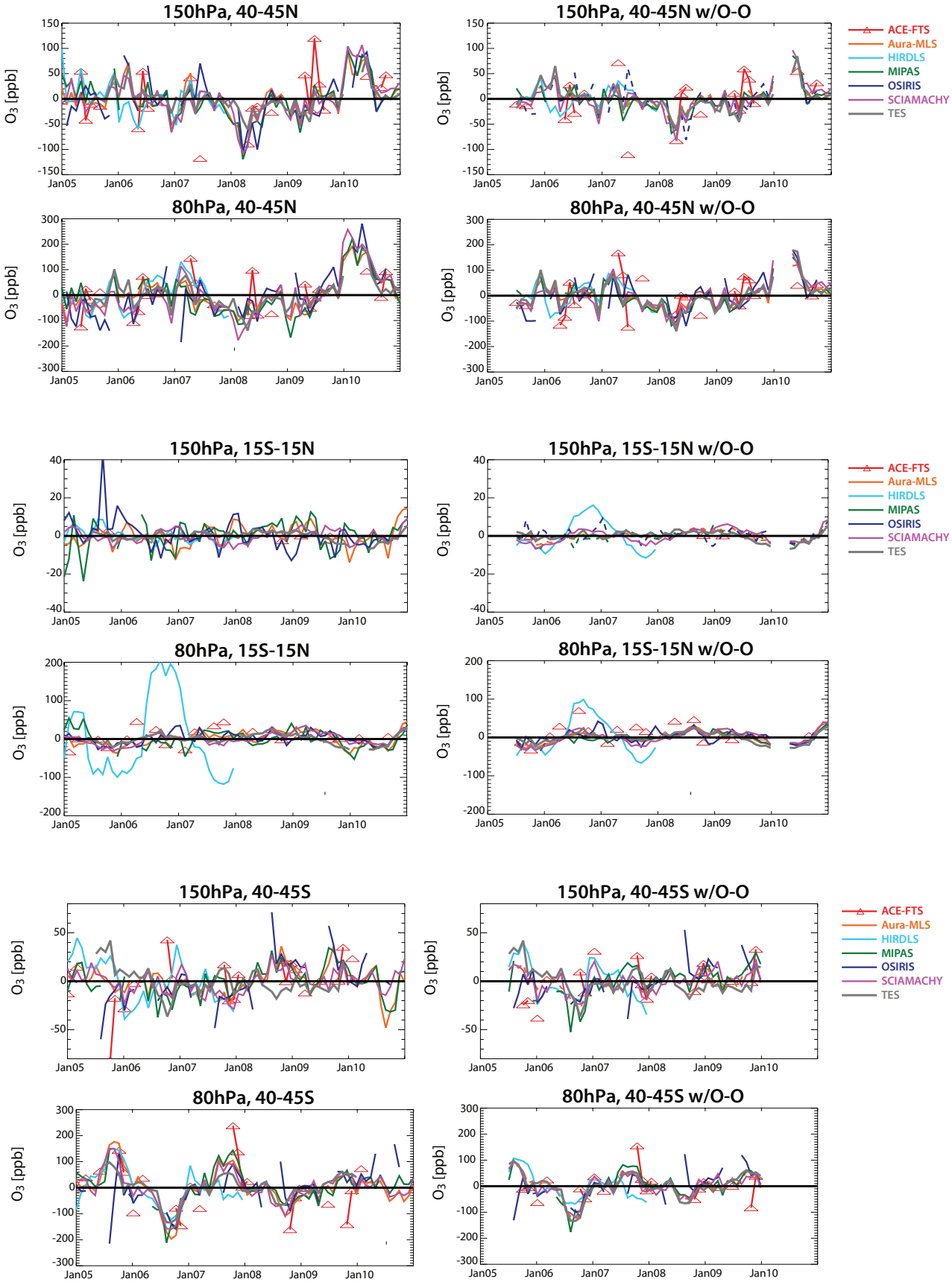


Figure 4.27.7: Interannual variability of ozone in the UTLS for 2005-2010. *Deseasonalised ozone anomalies at 150 hPa and 80 hPa for 40°N-45°N (top 4 panels), 15°S-15°N (middle 4 panels), and 40°S-45°S (bottom 4 panels). The left column shows the original climatologies, the right column shows the climatologies after application of the TES observational operator. Dashed lines in the figures in the right column indicate portions of the virtual retrieval where the difference in the virtual retrieval using the two different fill profiles exceeds 50% of the difference between the virtual retrieval and TES in the annual mean.*

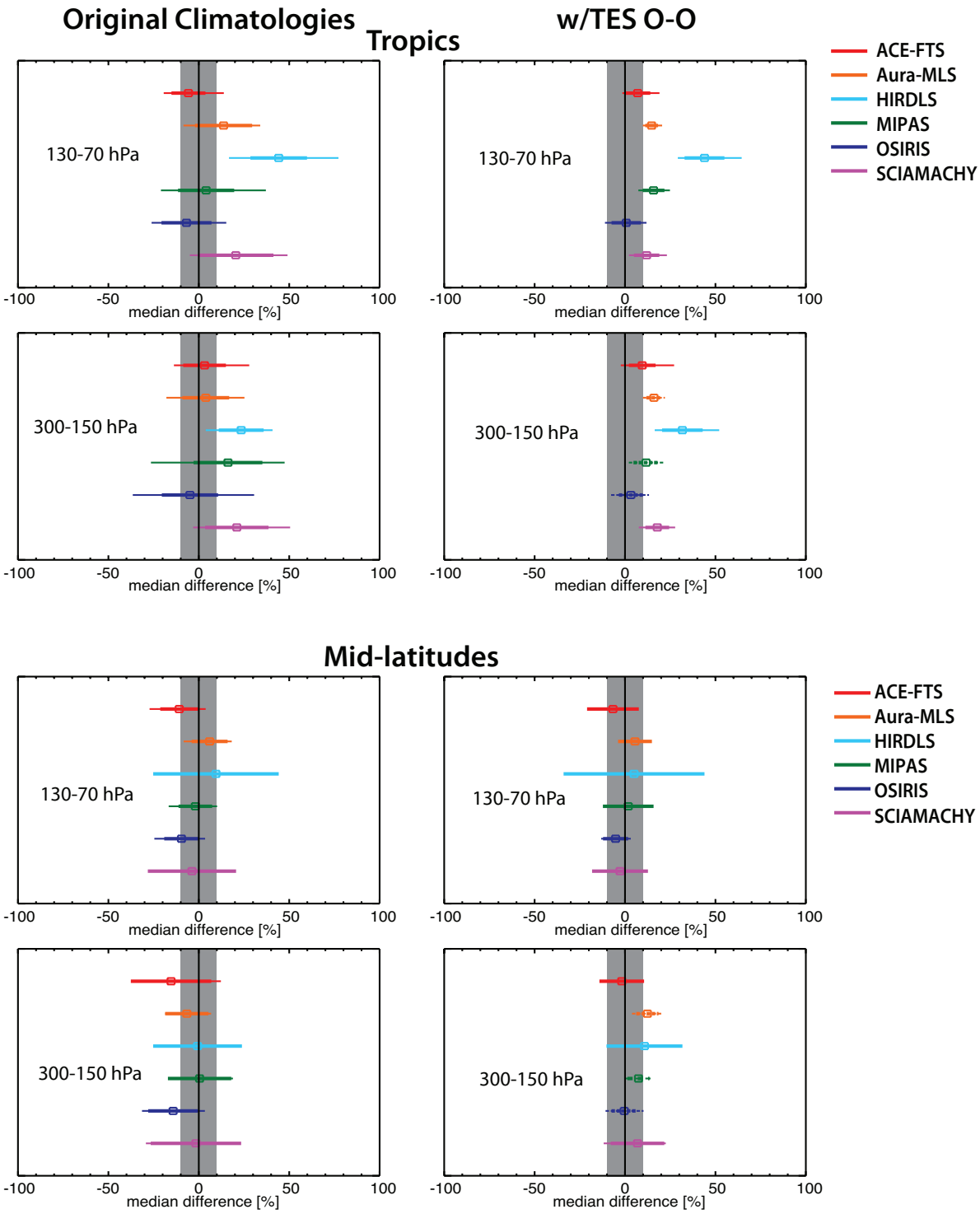


Figure 4.27.8: Summary plot of UTLS ozone. *Summary plot showing the median (square), median absolute deviation (thick line), and standard deviation (thin line) of monthly mean relative differences from TES. Here, the tropics are defined as 20°S-20°N, and mid-latitudes include 30°S-50°S and 30°N-50°N. Shaded grey regions show 10% difference from TES for reference. Left column shows results for the original climatologies, while the right column shows the results for the virtual TES retrievals.*

Key findings: nadir- and limb-viewing instruments comparisons

- Using the TES observational operator to vertically smooth the climatologies from the higher resolution limb-viewing instruments provides a common basis for comparison of the large-scale ozone morphology as well as the seasonal and interannual variability of ozone within the UTLS.
- However, this approach has several limitations, including:
  1) that the virtual retrievals can be sensitive to how one chooses to "fill in" the profiles below the lowest measurement level of the limb sounders,
  2) that the TES sensitivity varies in the UTLS such that the *a priori* profile has a significant influence near 150 hPa in the extra-tropics,

3) that the averaging kernels are not fully independent of the ozone abundance, resulting in errors in the virtual retrievals that are difficult to quantify.

We have tried to account for these factors when possible, and to focus on robust differences in the UTLS climatologies.

- The TES observational operator smoothes small-scale ozone structures and due to the influence of the *a priori* tends to increase tropical-extra-tropical ozone gradients as well as mid-latitude vertical ozone gradients in the climatologies from the limb-viewing instruments.
- The TES observational operator also reduces and vertically smoothes the differences between the limb climatologies and TES.
- The TES observational operator reduces the temporal variability of the ozone climatologies from the high-resolution instruments but also greatly improves the consistency between them. This indicates that the differences in vertical resolution among the limb-viewing instruments make a substantial contribution to differences in their retrieved ozone distributions both relative to TES and relative to one other.

Atmospheric mean state and variability

- Most of the limb-viewing instruments have climatological-mean positive differences (ranging from 5-75%) relative to TES ozone in the tropics, though for several instruments the differences depend strongly on the fill profile below ~100 hPa.
- For $p \leq 100$ hPa, the positive difference from TES likely reflect true positive biases for the climatologies given TES's lack of bias with respect to the ozonesonde climatology.
- In the Northern Hemisphere extra-tropics, only the HIRDLS and MIPAS climatologies have differences with respect to TES that are >15% and are also independent of the fill profile.
- In the southern extra-tropics, the TES observational operator greatly reduces differences between the limb-sounder climatologies and TES due to TES's low sensitivity at the pressure levels where the differences between the original climatologies and TES are largest.
- There are large differences in the timing and magnitude of the seasonal cycle in the tropical upper troposphere.
- At $p \leq 100$ hPa, the climatologies show a more consistent tropical seasonal cycle, particularly when smoothed to the TES vertical resolution. The TES observational operator reduces the differences in seasonal cycle amplitude to within 20% of TES for all instruments except HIRDLS.
- In general, there is very good agreement among the climatologies regarding both the timing and magnitude of the seasonal cycle in mid- and high latitudes, except that ozone from the OSIRIS and SCIAMACHY climatologies is ~15% low relative to the other instruments during the southern mid-latitude maximum, likely due to their limited sampling of this region. In southern high latitudes, all of the instruments except ACE-FTS overestimate the October ozone minimum relative to the ozonesonde climatology, but a definitive assessment of biases is precluded by sampling differences.
- All of the climatologies show low interannual variability in the tropics (except for HIRDLS) and higher variability in mid- and high latitudes.
- Northern mid-latitude interannual variability greatly exceeds that in southern mid-latitudes for $p > 80$ hPa, but variability in southern high latitudes is larger than that in northern high latitudes. The sampling of the ACE-FTS instrument is insufficient to capture interannual variability on monthly time scales in the tropics.
- The TES observational operator greatly reduces the interannual variability in ozone from the limb-sounder climatologies for $p > 100$ hPa in the tropics and $p > 200$ hPa in mid-latitudes. This improves the consistency between the datasets but may limit the usefulness of the virtual retrievals for quantifying UTLS variability.

Instrument-specific conclusions

- While the ACE-FTS climatology agrees well with the climatologies from TES and the other limb-viewing instruments for the annual mean ozone distribution, the instrument's sampling pattern impacts its ability to fully capture seasonal and interannual variability, particularly in the tropics.
- The HIRDLS climatology shows large positive differences from TES and the other limb sounders in the tropics, as well as a smaller positive difference in the NH extra-tropics for $p \geq 150$ hPa. The seasonal cycle amplitude is larger and the interannual variability is quite different in the tropics when compared to the other climatologies. However, both the seasonal and interannual variability are in very good agreement with the other climatologies in mid-latitudes. The large tropical ozone values are likely due to uncorrected emission from aerosol, and the differences in temporal variability relative to the other climatologies suggest that this aerosol effect may vary with time.
- The MIPAS climatology has a different morphology than is seen in the climatologies from the other instruments in the tropics between 200 and 100 hPa and much lower ozone values than any other climatology in the tropical upper troposphere. It is positively biased with respect to TES for $p < 100$ hPa in the tropics and for $p \geq 150$ hPa in the northern extra-tropics but agrees relatively well elsewhere.
- The MLS climatology has unrealistically "flat" contours in the tropics near 100 hPa, but otherwise agrees quite well with the climatologies from TES and the other limb sounders.
- The OSIRIS climatology shows the best overall agreement with TES. The limited sampling in polar winter results in unrealistic horizontal ozone gradients in the

northern high latitude lowermost stratosphere and in the southern mid-latitudes results in a slightly underestimated seasonal cycle and problems capturing the interannual variability of UTLS ozone in some months.

- The SCIAMACHY climatology is positively biased with respect to TES in the tropics. As for OSIRIS, the limited sampling during polar night leads to a smaller amplitude southern mid-latitude seasonal cycle than is seen in the other climatologies.

4.27.7 Recommendations: UTLS ozone

- A much more detailed UTLS inter-comparison using high spatial and temporal resolution measurements of multiple species is needed to fully characterise differences between instruments in this region, and has been proposed as a follow-on to the SPARC Data Initiative.
- Such an inter-comparison will require diagnostic tools that minimise geophysical variability and differences in sampling and resolution such as tracer-tracer correlations, probability distribution functions, tropopause-relative vertical coordinates, and jet-based coordinates [*e.g.*, *Hegglin et al.*, 2008; *SPARC*, 2010; *Manney et al.*, 2011].
- The proposed follow-on analysis promises to not only provide a detailed assessment of the quality of the satellite data, but also to improve our understanding of UTLS structure and processes.

Chapter 5: Implications of results

5.1 Implications for model-measurement intercomparison

Satellite trace gas datasets are crucial for the evaluation of transport and chemistry in numerical models. Datasets available from different satellite instruments vary in terms of measurement method, geographical coverage, spatial and temporal sampling and resolution, time period, and retrieval algorithm and thus have different strengths and shortcomings. Comparing numerical model output to different chemical datasets can lead to conflicting results depending on the particular application. Issues arising from the use of different observational datasets for model evaluations have been identified in the CCMVal report [*SPARC*, 2010]. It became clear that the characteristics of the satellite datasets, including quality, resolution, and representativeness, need to be known prior to their use and prior to the interpretation of model evaluation results. The CCMVal report's recommendations that "*A systematic comparison of existing observations is required in order to underpin future model evaluation efforts, by providing more accurate assessments of measurement uncertainty*" directly motivated the work for the SPARC Data Initiative presented in this report. While *Chapter 4* provides basic information on quality and consistency of the various data products, the following *Chapter 5* focuses on summarizing some implications of the results for model-measurement inter-comparisons. Examples of how knowledge of uncertainty and inter-instrument differences can be used to improve comparisons are given and particular diagnostics appropriate for model evaluations are recommended.

For the CCMVal report, the observational mean values and uncertainty range served as input for the performance metrics. Such metrics are used to quantify the ability of models to reproduce key stratospheric processes. One widely applied metric:

$$g = 1 - \frac{1}{n_g} \frac{|\mu_{mod} - \mu_{obs}|}{\sigma_{obs}} \qquad (5.1)$$

uses a scaling factor n_g as well as the observational uncertainty σ_{obs} and climatological mean μ_{obs} for the evaluation of the model climatological mean μ_{mod} [*Douglass et al.*, 1999; *Waugh and Eyring*, 2008; *SPARC*, 2010]. In the past, the observational uncertainty has most often been derived using the interannual variability of a single instrument only.

Our approach is to provide an alternative, more comprehensive uncertainty range derived from all available datasets, instead of recommending one particular satellite dataset for the model-measurement comparison. The selection of the data points suitable for the construction of the new climatological mean values and uncertainty range is based on their agreement with the mean state of the atmosphere as given by all instruments and on the specific satellite characteristics such as sampling patterns. The following general guidelines are applied for the selection process.

- The agreement of each individual dataset with the mean state of the atmosphere is determined based on the 1σ standard deviation over all instruments. For trace gases observed by more than five instruments, individual data points will be removed if they are outside of the $\pm 1\sigma$ standard deviation range. For trace gases observed by five or less instruments, the data points will be removed if they are outside of the $\pm 2\sigma$ standard deviation.
- Further specific criteria used to calculate the mean state and uncertainty range are chosen based on the instrument/retrieval performance identified in the different chapters of this report, and will change depending on the diagnostic and the trace gas. Detailed information on the evaluations is provided in the following paragraphs, structured according to evaluation diagnostic.
- For each diagnostic and even within one diagnostic, the datasets selected for the construction of the uncertainty range can be different depending on latitude, altitude, or time period considered. One example of this approach can be given for the evaluation of the ozone seasonal cycle: if one instrument presents a clear outlier for *e.g.*, March, then only the March value of this instrument is removed while the values for all other months stay included in the uncertainty range.
- The climatological mean μ_{obs} is defined as the multi-annual, multi-instrument mean value of all data points selected as suitable for the construction of the uncertainty range. Note that the climatological mean is different from the MIM used in the previous chapters which was based on all available datasets.
- The uncertainty range σ_{obs} is defined as the spread over all selected datasets and years. In general, the interannual spread needs to be accounted for when producing the uncertainty range, so that the free-running models can be compared against the observational mean state. Note however, that for model simulations nudged to meteorological reanalysis, the comparisons focus on the same years and the uncertainty range can be solely based on the spread over all selected datasets and not include interannual variations.

In summary, we derive an observational mean state and uncertainty range from multiple datasets from the SPARC Data Initiative for selected examples of evaluation diagnostics. For all evaluations listed in the following *Section 5.1*, the uncertainty range and climatological mean will be made publicly available through the SPARC Data Center.

5.1.1 Seasonal cycles

The seasonal cycle of long-lived atmospheric trace gases such as water vapor and ozone is often used as a diagnostic of transport processes in the stratosphere and in particular in the UTLS.

Ozone – O_3

The ozone seasonal cycle in the UTLS in mid-latitudes is determined by the seasonality of two processes: air mass transport with the Brewer-Dobson circulation and mixing with tropical air masses. In order to evaluate the model's representation of these large-scale transport and mixing processes, a comparison of the ozone seasonal cycle for the latitude bands 40°S/N-60°S/N at 100 and 200 hPa has been used [*SPARC*, 2010; *Hegglin et al.*, 2010]. While the calculation of the quantitative performance metric in the CCMVal report was based on MIPAS data alone, we will provide a new climatological mean state and uncertainty range derived from multiple datasets. The method (illustrated in **Figure 5.1.1**) is explained below for 40°N-60°N, 200 hPa.

Step 1: The ozone seasonal cycles for satellite datasets are derived from 2005-2010 multi-annual mean values. The time period has been chosen based on a maximum number of active satellite limb instruments. The uncertainty range (grey shading in **Figure 5.1.1**) is calculated as the $\pm 1\sigma$ standard deviation over all instruments' multi-annual mean values.

Step 2: All data points outside of the $\pm 1\sigma$ standard deviation from step 1 are removed. Additionally, data points from instruments with a very large interannual spread need to be excluded. Therefore, all multi-annual mean values with an interannual variability (vertical bars in the uppermost left panel of **Figure 5.1.1**) larger than the $\pm 2\sigma$ standard deviations from step 1 are removed. The new mean values and uncertainty range are calculated.

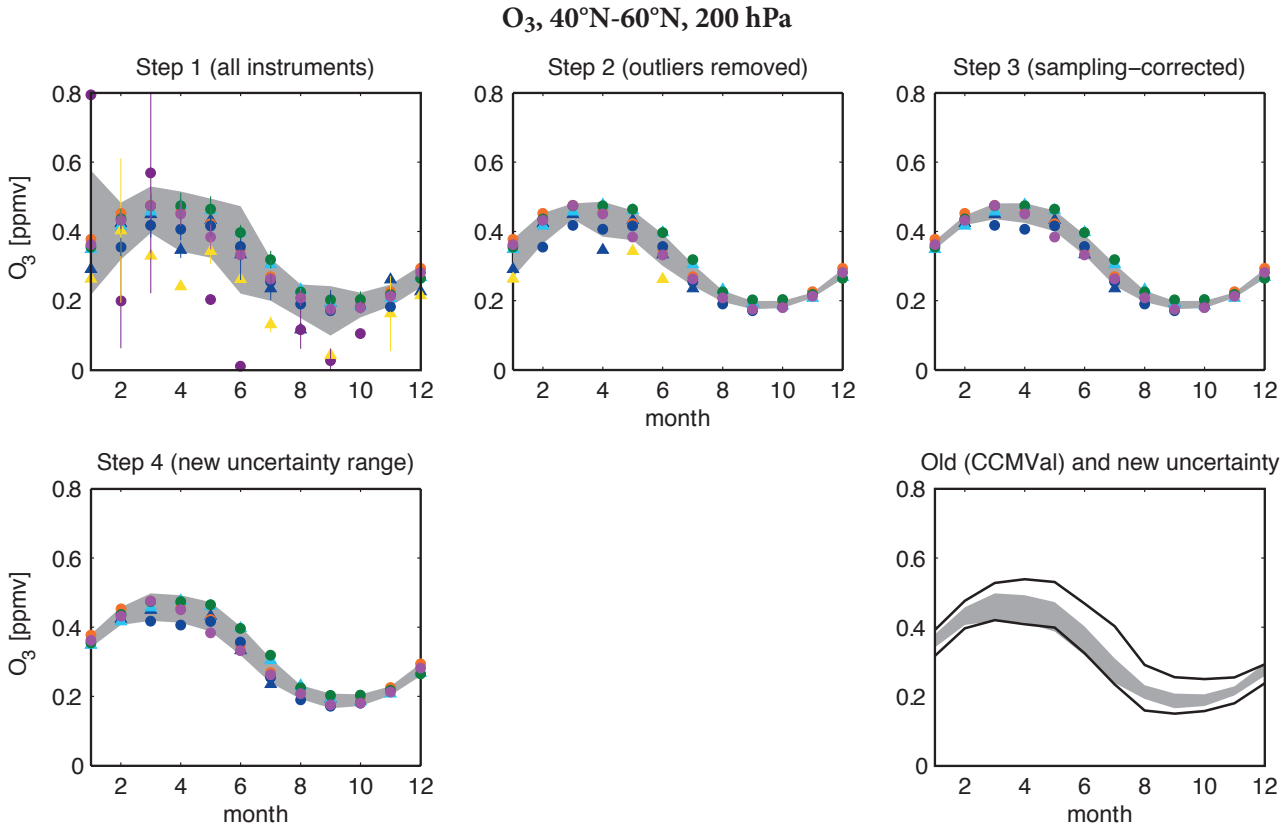


Figure 5.1.1: Ozone seasonal cycle diagnostic for 40°N-60°N at 200 hPa. *The individual steps of deriving the ozone seasonal cycle diagnostic are shown. The uncertainty range (grey shading) is given for each month by the standard deviation over all multi-annual means of the selected datasets. In the uppermost left panel the vertical bars indicate the interannual spread of each instrument calculated as the standard deviations over all years. For the selection of the datasets, outliers and data points strongly impacted by sampling are removed as illustrated in steps 1 to 3 and explained in detail in the text. In step 4 the uncertainty due to interannual variations is added to the uncertainty range. In the lower rightmost panel the old uncertainty range given in the CCMVal report and the new uncertainty range are compared.*

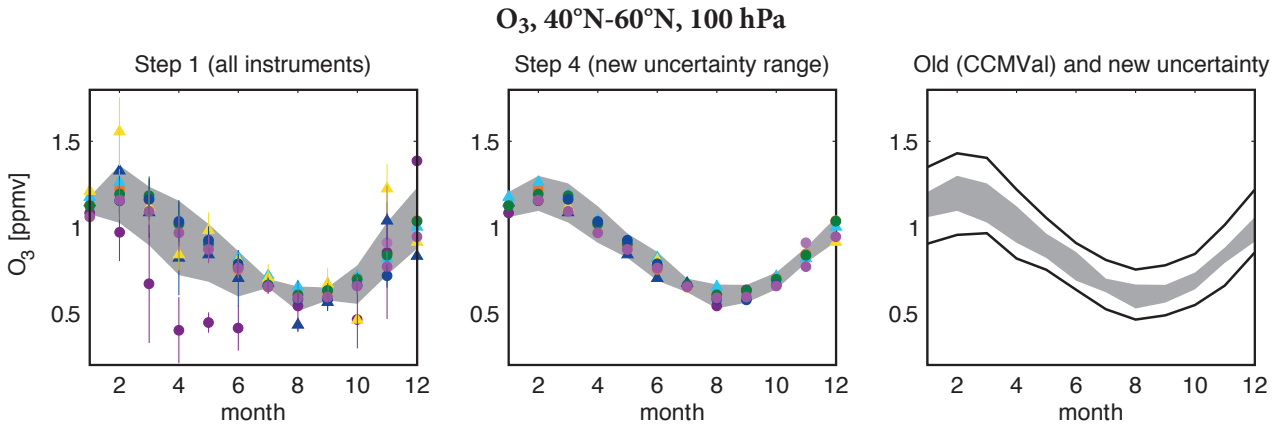


Figure 5.1.2: Ozone seasonal cycle diagnostic for 40°N-60°N at 100 hPa. *Steps 1 and 4 of deriving the ozone seasonal cycle diagnostic are shown. The uncertainty range (grey shading) is given for each month by the standard deviation over all selected datasets. In the rightmost panels the old uncertainty range given in the CCMVal report and the new uncertainty range are compared.*

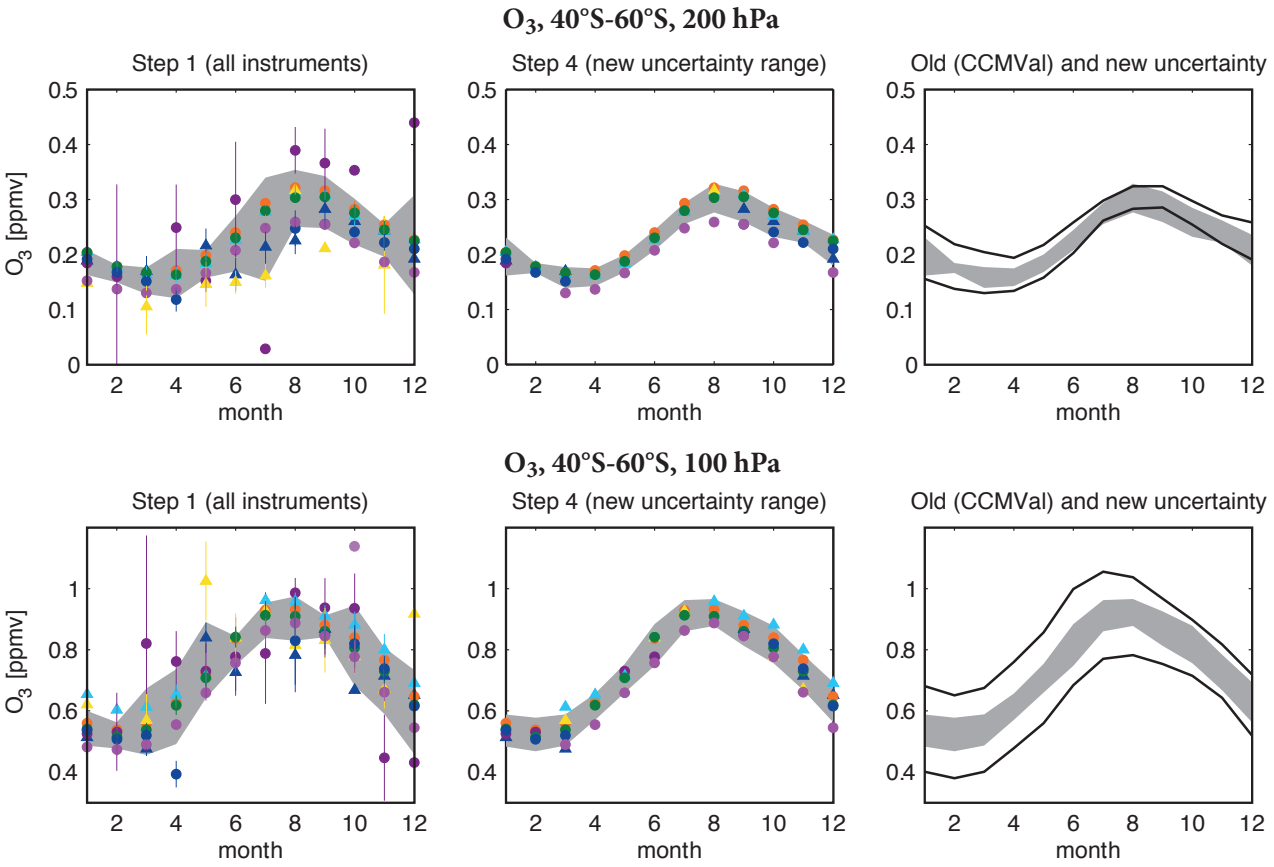


Figure 5.1.3: Ozone seasonal cycle diagnostic for 40°S-60°S at 200 and 100 hPa. *Same as Figure 5.1.2 but for 200 and 100 hPa at 40°S-60°S.*

Step 3: All data points impacted by a sampling bias estimated to be larger than 10% are removed. Such sampling bias can arise when averaging binned atmospheric measurements due to non-uniform sampling in time or space. These sampling biases have been identified by applying the sampling patterns of the satellite instruments to O_3 fields from coupled chemistry climate models (see *Chapter 3*; *Toohey et al.* [2013]). In the tropics, comparisons to ozonesondes are used to remove data points that show large deviations. The new mean values and uncertainty range are calculated.

Step 4: The uncertainty range is recalculated as the $\pm 1\sigma$ standard deviation over all remaining instruments and years, now taking not only the inter-instrument but also the inter-annual spread into account. Including the latter in this final step increases the uncertainty range for most cases, but is nevertheless important in order to produce an uncertainty that free-running models can be compared against.

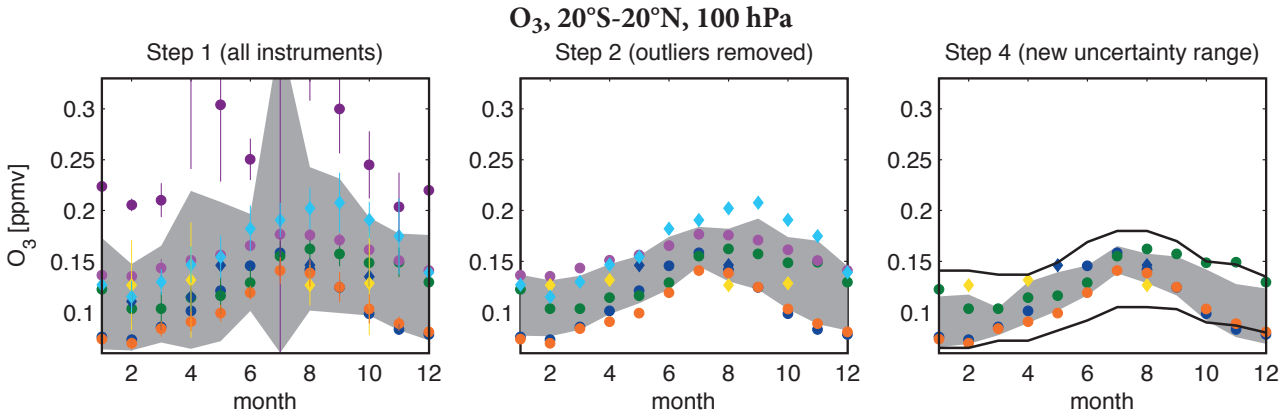


Figure 5.1.4: Ozone seasonal cycle diagnostic for 20°S-20°N at 100 hPa. *Steps 1, 2 and 4 of deriving the ozone seasonal cycle diagnostic are shown. The uncertainty range (grey shading) is given for each month by the standard deviation over all selected data points.*

To summarise, the final ozone seasonal cycle for the 2005-2010 period is calculated from the instruments' multi-annual mean values remaining after the removal of outliers and data points impacted by sampling bias (steps 1-3). The uncertainty range is calculated accordingly as the ±1σ standard deviation over all those instruments and over all years. The new uncertainty range is generally smaller than the old uncertainty range used in the CCMVal report (lower right panel in **Figure 5.1.1**). For some months, the uncertainty has been reduced by more than 50%. This reduced uncertainty range applied to the quantitative performance metric (**Eq. 5.1**) will provide a powerful constraint on the model results and thus the model representations of the ozone seasonal cycle can be differentiated more clearly than before. Additionally, the climatological mean is now shifted to lower values. The new lower mean values agree better with the CCMVal models whose multi-model mean values were found to be too low compared to the old climatological mean values (see Figure 7.22 in the *CCMVal report*; also Figure 11 in *Hegglin et al.* [2010]). The improved agreement suggests that most of the CCMVal models perform better than previously thought with regard to the ozone seasonal cycle in the UTLS.

For the presentation of the improved seasonal cycle diagnostic for other regions and trace gases, only step 1 and 4 and for some regions also step 2 as well as the comparison with the old CCMVal uncertainty range will be displayed. Figures containing each step of the derivation of the new uncertainty range are provided in *Appendix A5*. At 100 hPa in the Northern Hemisphere (NH) mid-latitudes (40°N-60°N), the ozone seasonal cycle is derived from the 2005-2010 multi-annual mean of 9 satellite instruments (**Figure 5.1.2**). Reducing the satellite datasets according to their agreement with the multi-instrument mean value and their sampling biases results in a much reduced uncertainty range in particular during NH winter and spring. For these months, most of the new uncertainty is caused by inter-annual variations and not by inter-instrument variations as becomes clear from the multi-annual mean values clustering in the center of the new uncertainty range. Similar to our results for 200 hPa, the new uncertainty range is much reduced when compared to the one used in the CCMVal report, and hence will be much better suited to identify badly performing models. For the NH summer and autumn, the reduction is about 2/3 of the old range. The climatological mean value, however, did not change systematically.

The evaluation of the Southern Hemisphere (SH) mid-latitude ozone seasonal cycle (**Figure 5.1.3**) follows the same steps as described above for the NH based on 2005-2010 multi-annual mean datasets from 9 satellite instruments. At 200 hPa, the new uncertainty range is very similar to the old one based on MIPAS observations only. Except for January, this uncertainty results mostly from the inter-instrument spread (and not from the inter-annual spread) with one instrument having particularly lower values than all other datasets. Despite this instrument seeming to be an outlier for some months, it agrees very well with SAGE II and HALOE ozone for the overlap time period 2003 (not included here) confirming this as the lower end of our uncertainty range. At 100 hPa, the new uncertainty range is reduced over the whole year compared to the old one based on MIPAS observations, with strongest improvement for SH summer, autumn and winter. The climatological mean values increase slightly for August-October, but do not change systematically for the rest of the year.

Tropical ozone exhibits a large annual cycle near and above the tropopause which is related to seasonal changes in vertical transport acting on the strong vertical ozone gradient in this region [*Randel et al.*, 2007] and in quasi-horizontal mixing [*Ploeger et al.*, 2012]. 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 has been used to analyse transport and mixing processes. The SPARC CCMVal evaluation of the seasonal cycle in tropical ozone is based on a comparison to the observational NIWA dataset [*Hassler et al.*, 2008] at 100 hPa for 20°S-20°N. A new uncertainty range based on the SPARC Data Initiative datasets is presented in **Figure 5.1.4.** After the removal of the outliers the uncertainty range (middle panel of **Figure 5.1.4**) is still relatively large and comparable to the

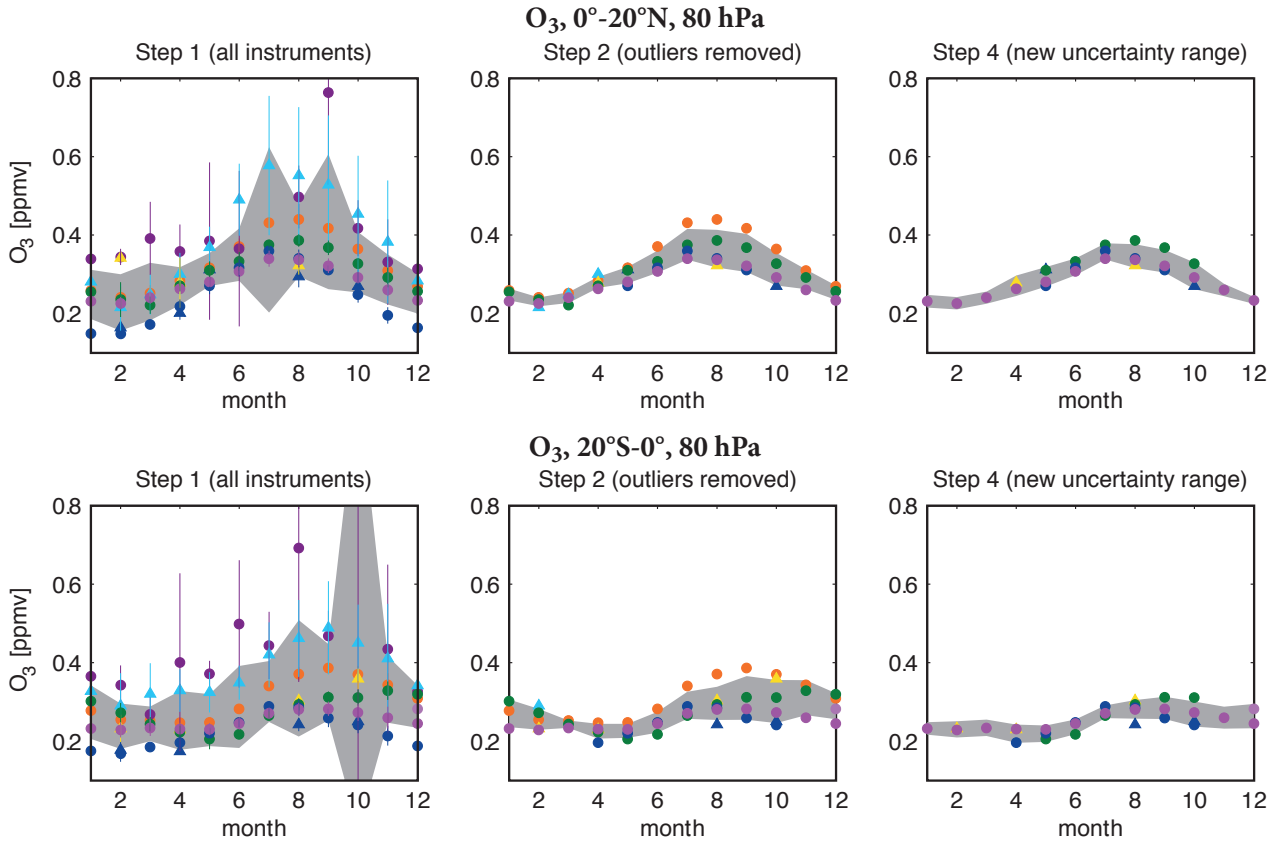


Figure 5.1.5: Ozone seasonal cycle diagnostic for 0°-20°N and 20°S-0° at 80 hPa. *Same as Figure 5.1.4 but for 80 hPa at 0°-20°N and 20°S-0°.*

NIWA-based uncertainty range (black lines in right panel of **Figure 5.1.4**). Evaluations of UTLS ozone after the application of the TES observational operator [*Section 4.27*; *Neu et al.*, 2014a] including a comparison to a "zonal mean" ozonesonde climatology indicate that in the tropics below 100 hPa most instruments have a positive bias. Removing all datasets that are outside of the ±1σ standard deviation of the climatological ozonesonde measurements (see **Figure 4.27.6** for details) results in a lower mean and also a reduced uncertainty range (right panel of **Figure 5.1.4**). Note that for November, the criteria has not been applied in order to avoid inconsistencies with the October and December uncertainty ranges. In particular for the time period from March to October the uncertainty range has been substantially reduced and is now smaller than the NIWA-based one.

Most studies analyzing the seasonal cycle of long-lived trace gases treat the tropics as a horizontally homogeneous region without differentiating between NH and SH. Very recently differences between the ozone seasonal cycles in the NH and SH tropics, related to hemispheric differences in the seasonal strength of vertical transport and horizontal mixing, have been pointed out by *Stolarski et al.* [2014]. Here, we follow their approach and derive uncertainty ranges for the ozone seasonal cycle at 80 hPa for the NH tropics (0°-20°N) and the SH tropics (20°S-0°) as illustrated in **Figure 5.1.5**. For both regions, the number of overall applicable datasets decreases substantially when identifying outliers and comparing to ozonesondes resulting in a new, narrow uncertainty range. Confirming the results from *Stolarski et al.* [2014] the seasonal cycle of tropical ozone is substantially different in the two hemispheres with a less pronounced and later occurring maximum in the SH. Model-evaluations of the ozone seasonal cycle should thus be based on two diagnostics differentiating between the NH and SH tropics.

Nitric acid – HNO_3

The HNO_3 seasonal cycle in the UTLS is used to evaluate transport and mixing processes in the models on typical time scales of weeks to months. Like ozone, HNO_3 is mostly produced in the stratosphere and thus has a similar seasonal cycle with some differences caused by chemistry and microphysics. **Figure 5.1.6** shows the evaluation diagnostics of the HNO_3 seasonal cycle for 40°N-60°N at 100 and 200 hPa. The seasonal cycles of five satellite instruments are derived from multi-annual mean (2005-2010) values. At both levels, but in particular at 100 hPa, the instruments are clustered together with only little inter-instrument spread (left panels in **Figure 5.1.6**). Compared to the ozone seasonal cycle, we have fewer instruments available (the ozone evaluations are based on 9 instruments) and thus choose a different criterion to identify outliers. Only data points outside of the ±2σ standard deviation calculated in step 1 will be removed. Note that the grey shading in **Figure 5.1.6** corresponds to the ±1σ standard deviation. At both levels, the agreement between the multi-annual mean states of the five instruments is very

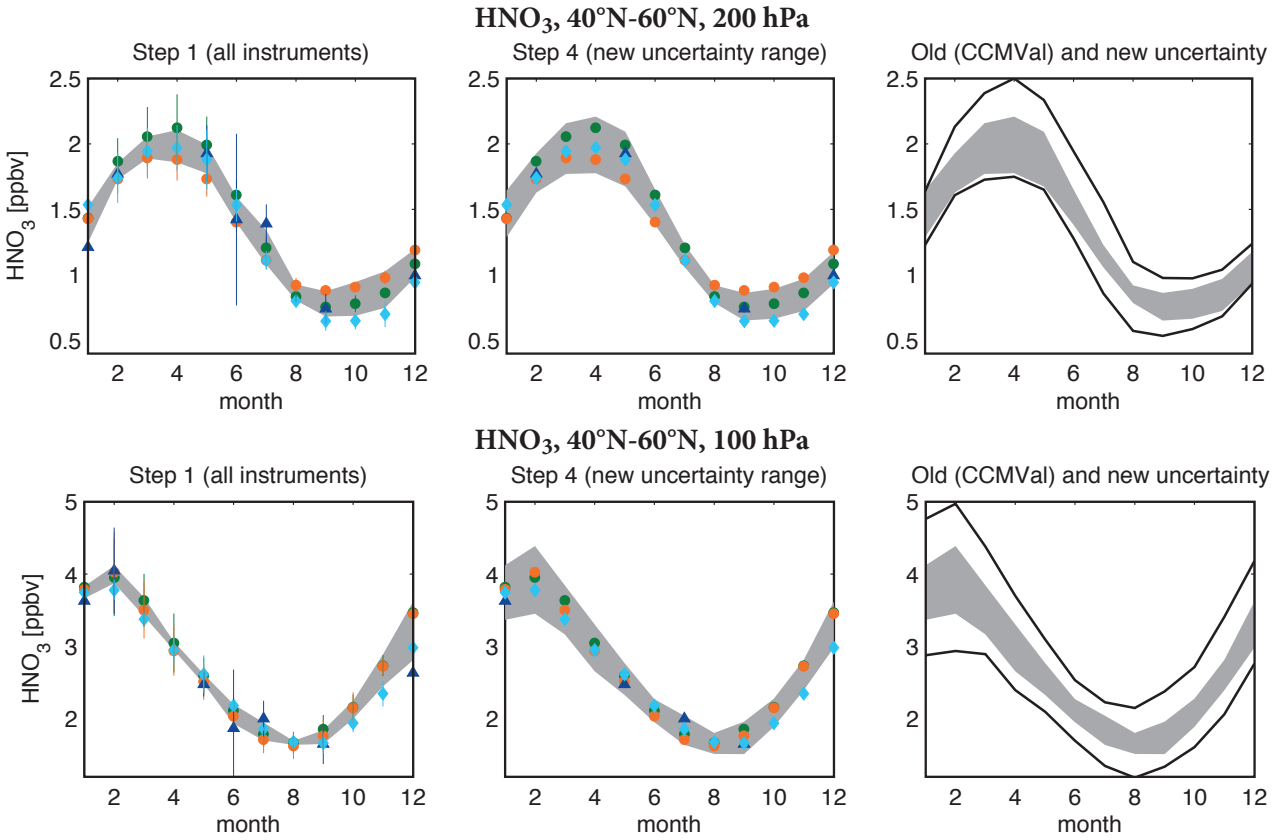


Figure 5.1.6: HNO_3 seasonal cycle diagnostic for 40°N-60°N at 200 and 100 hPa. *Steps 1 and 4 of deriving the HNO_3 seasonal cycle diagnostic are shown. The uncertainty range (grey shading) is given for each month by the standard deviation over all selected data points. In the rightmost panels the old uncertainty range given in SPARC (2010) and the new uncertainty range are compared.*

good and thus no data points are identified as outliers and excluded from the calculation of the uncertainty range. However, some instruments are removed due to their large interannual variability (illustrated by the vertical bars in the left panels). At 200 hPa, the uncertainty is more driven by the instrument spread than by the interannual variability and is smallest during NH summer. Compared to the CCMVal report the new climatological mean values during NH winter and spring are lower. The new lower mean values agree also better with most of the CCMVal models which were found to be too low when compared to the old climatological mean (see Figure 7.22 in *SPARC*, 2010). At 100 hPa, the new uncertainty range is largest during the NH winter as a result of the inter-annual variability of the remaining data. Comparisons to the uncertainty used in the CCMVal report and in *Hegglin et al.* [2010] show that the new uncertainty range is much reduced.

Similarly to the HNO_3 seasonal cycle in the NH UTLS, we derive a new uncertainty range and climatological mean (see **Figure A5.1.4** in *Appendix A5*) for the HNO_3 seasonal cycle in the SH UTLS (30°S-60°S, 100 and 200 hPa) as applied in the UTLS chapter of the CCMVal report. The strongest reduction of the uncertainty range with respect to the one used in the SPARC report is found at 100 hPa in the form of an 80% decrease.

Water vapour – H_2O

The H_2O seasonal cycle in the tropical tropopause region (at 80 hPa) is a key diagnostic to evaluate the amount of water vapour entering the stratosphere [*Gettelman et al.*, 2010]. Water vapour affects stratospheric ozone through HO_x-chemistry as well as the formation of polar stratospheric clouds, and also the radiative budget of the UTLS [*SPARC*, 2000]. The seasonal cycle in water vapour is closely related to the seasonal cycle in tropical coldpoint tropopause temperature, which in turn is dominated by the seasonally varying strength of the stratospheric Brewer-Dobson circulation.

Figure 5.1.7 shows the evaluation diagnostics of the H_2O seasonal cycle for 20°S-20°N at 80 hPa. The seasonal cycles of seven satellite instruments are derived from multi-annual means averaged over the time period 1996-2010. Choosing a shorter time period for which the instruments would show exact overlap does not improve the comparison between the instruments (see *Chapter 4*; *Hegglin et al.* [2013]), but would limit the number of instruments and information on interannual variability needed in step 4 to calculate an improved uncertainty range.

The instruments do not agree well on the mean values and hence the uncertainty range is relatively large.

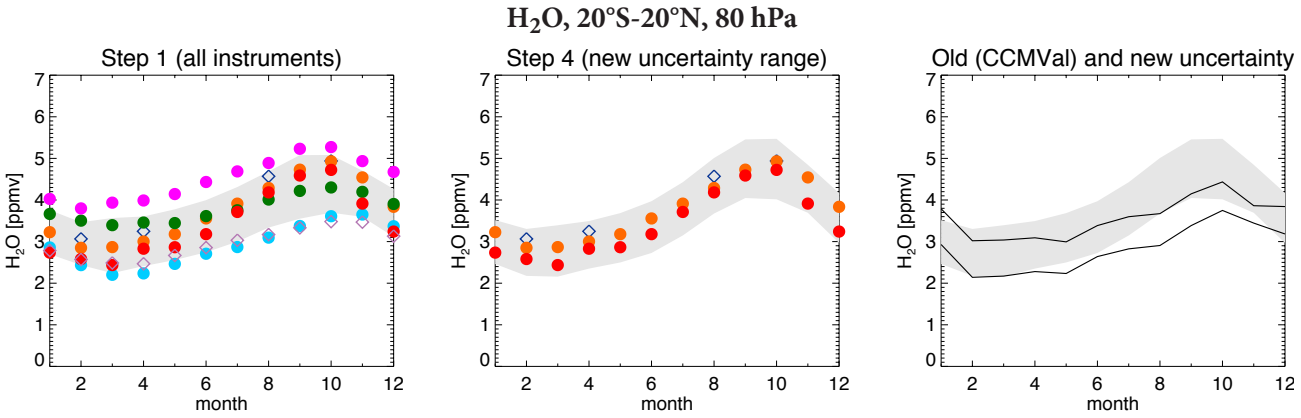


Figure 5.1.7: H_2O seasonal cycle diagnostic for 20°S-20°N at 80 hPa. *Steps 1 and 4 of deriving the H_2O seasonal cycle diagnostic are shown. The uncertainty range (grey shading) is given for each month by the 1σ standard deviation over all selected datasets (left panel). The middle panel shows the sub-selected datasets, but with the new uncertainty range accounting for interannual variability. In the right panel the new uncertainty range is compared to the uncertainty range given in the CCMVal report.*

Following the approach introduced in this chapter, we exclude instruments that lie outside the ±1σ uncertainty range given by the seven available satellite datasets. The instruments excluded were already identified in *Chapter 4.2* to have weaknesses, with SMR and HALOE showing a distinct low bias and SCIAMACHY showing a high bias due to too low resolution in the altitude range considered. In addition, we also remove MIPAS, since its averaging kernels are state-dependent (measuring with better altitude resolution in a more humid atmosphere), which leads to a seasonal cycle that exhibits too small an amplitude. The remaining instruments agree very well with each other, even though ACE-FTS has very limited sampling in the tropics. Thus no further data points are removed from the calculation and the new uncertainty range is calculated including interannual variability. The new mean seasonal cycle and its uncertainty imply that the models have been evaluated in the CCMVal report against a too low water vapour reference in terms of both mean values and seasonal cycle amplitude, while the old uncertainty range may have underestimated the impact of interannual variability.

5.1.2 Vertical and meridional profiles

Ozone – O_3

Another important aspect of CCM validation is the evaluation of polar spring time ozone profiles. Climatological mean vertical profiles in March at 75°N-85°N and in October at 75°S-85°S are compared between models and observations in order to test the models' representation of transport and chemistry in the polar regions. In contrast to the strong ozone decline driven by anthropogenically emitted ozone depleting substances until the mid-1990s, the Antarctic ozone hole has been controlled primarily by variations in stratospheric temperature and dynamical processes since 1997 [*WMO*, 2011]. In order to avoid the impact of the strong trend before the mid-1990s, we choose the time period 1997-2010 for the ozone profile evaluation in the polar regions. Over this long time period eight ozone datasets provide profile information for the Antarctic spring (**Figure 5.1.8**, left panel). Although some of the datasets cover only part of the time period, most of the profiles cluster together closely. We find one clear outlier with large deviations on the positive side, which is removed in step 2 (**Figure 5.1.8**, middle panel). In the last step, interannual variations are included in the construction of the uncertainty range resulting in slightly larger uncertainties (**Figure 5.1.8**, right panel). Overall in the MS and US, a well-defined mean ozone profile with a relatively narrow uncertainty range is derived for the Antarctic spring. In the LS, however, the spread is quite large which given the overall very small ozone abundances during this time of the Antarctic ozone hole, results in very large relative differences (see also *Chapter 4.1.6*; *Tegtmeier et al.* [2013]). The ozone hole with near-zero ozone values extends from 300 to nearly 50 hPa. Particularly between 100 and 50 hPa, the uncertainty is much higher than in other altitude ranges with similarly low abundance (above 0.3 hPa) or during other times of the year (not shown here). Such differences might be related to the different sampling patterns of the individual instruments and for detailed evaluations of high-latitude ozone in the LS we recommend the use of coincident measurement comparisons, polar vortex coordinates and the use of *in-situ* measurements.

Ozone evaluations can depend on the time period chosen. If we limit the ozone profile comparisons to shorter time periods such as 2000-2010 or 2005-2010 we get very similar mean profiles but a somewhat smaller uncertainty range. In particular, for the latter time period, the uncertainty range in the lower and middle stratosphere can be substantially reduced (see **Figure A5.1.5** in *Appendix A5*). While this suggests a better agreement of the instruments covering the latter time period, one needs to keep in mind that fewer instruments go into this evaluation (five instead of eight) which have at the same time a denser sampling pattern. The evaluation of the earlier time period 1991-2000

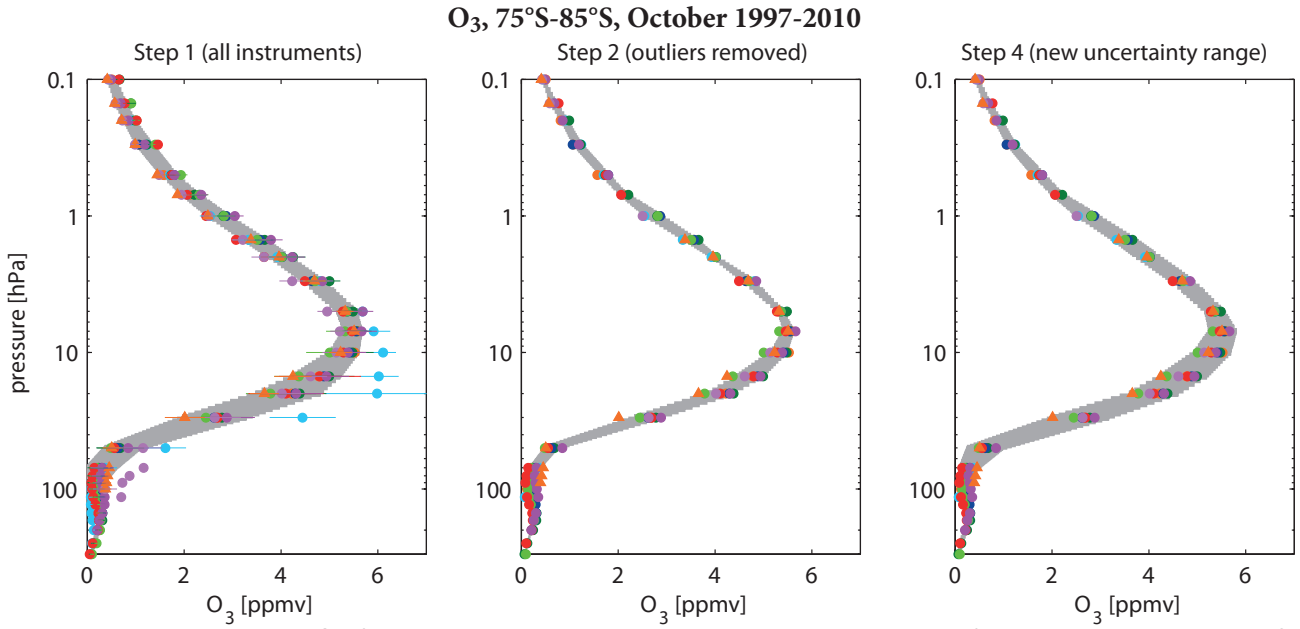


Figure 5.1.8: O_3 vertical profile for 75°S - 85°S in October 1997-2010. *Steps 1, 2 and 4 of deriving the O_3 vertical profile diagnostic are shown. The uncertainty range (grey shading) is given for each level by the standard deviation over all selected datasets.*

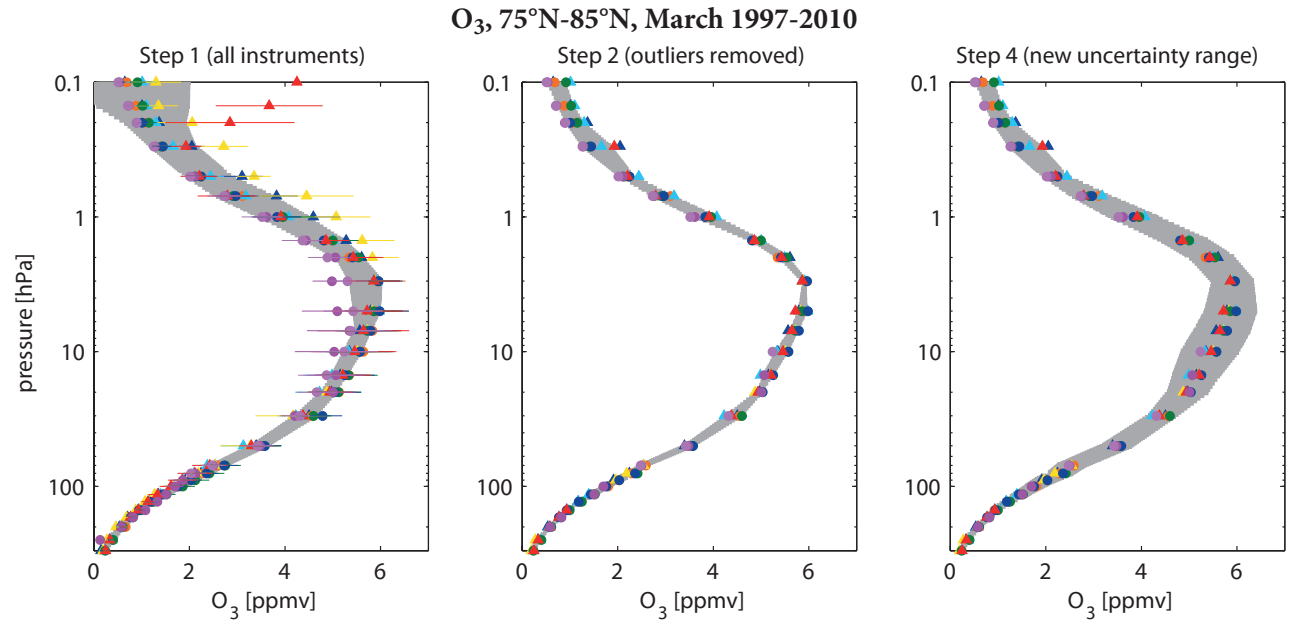


Figure 5.1.9: O_3 vertical profile for 75°N-85°N in March 1997-2010. *Same as Figure 5.1.8 but for 75°N-85°N in March.*

(**Figure A5.1.6** in *Appendix A5*), on the other hand, gives a different mean profile and a slightly larger uncertainty range due to larger interannual variability and, in comparison to 2005-2010, larger instrument-spread. In previous model evaluations focusing also on the 1990s [*SPARC*, 2010; *Eyring et al.*, 2006] the uncertainty range, based on the HALOE climatology and interannual standard deviations, was much larger than the new, multi-instrument uncertainty range introduced above.

Evaluation of the Arctic spring time ozone (here 75°N-85°N in March) shows a large inter-instrument spread, in particular in the MS/US (**Figure 5.1.9**, left panel). The spread is in most cases based on 1-2 outliers which are removed in step 2 resulting in a very narrow uncertainty range (**Figure 5.1.9**, middle panel). Due to the larger dynamical variability at the NH high latitudes, including the interannual standard deviation in the construction of the uncertainty range leads to much larger uncertainties, in particular in the MS. In contrast to the Antarctic, the inter-instrument spread in the LS is quite small leading to a well-defined profile with low uncertainties in this region.

Methane – CH_4

Methane (CH_4) meridional profiles are similarly used in model evaluation to study stratospheric transport characteristics (see *Eyring et al.* [2006]). As mentioned above, transport in the stratosphere involves both the residual mean circulation and isentropic mixing, with the

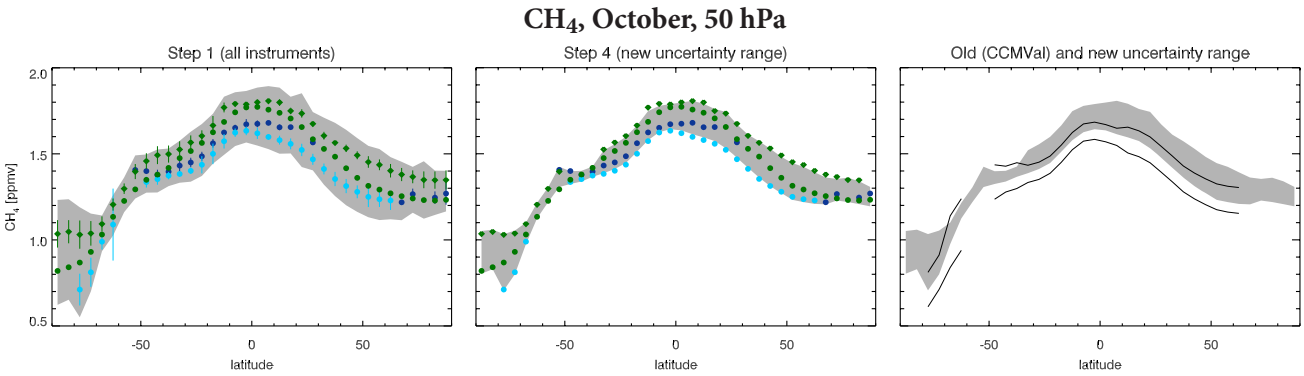


Figure 5.1.10: Climatological CH_4 meridional profile at 50 hPa in October over the time period 1998-2010. *Steps 1 and 4 of deriving the meridional profile of CH_4 are shown in the upper two panels. The uncertainty range (grey shading) is given for each month by the ±2σ standard deviation over all selected datasets (left panel). In the middle panel the newly derived uncertainty range (accounting for interannual variability) is shown, and in the right panel it is compared to the old uncertainty range given in the CCMVal report.*

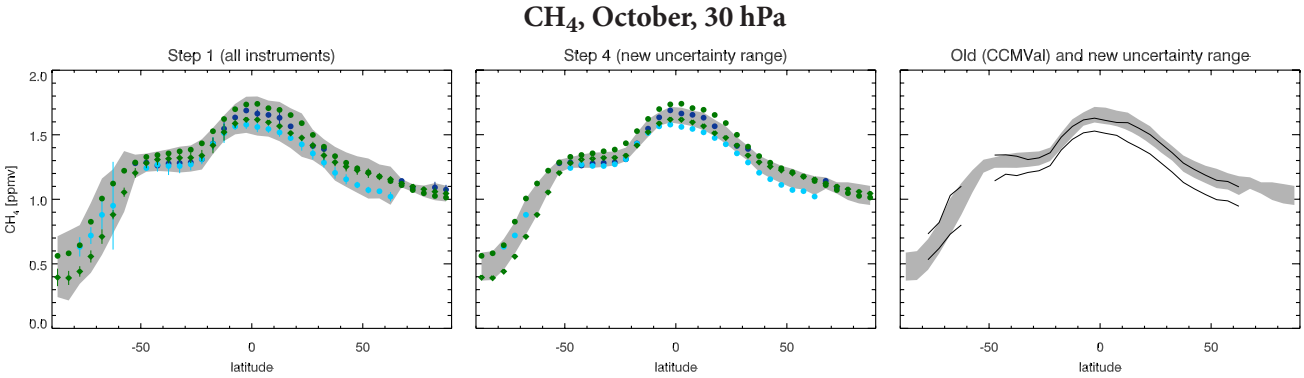


Figure 5.1.11: Climatological CH_4 meridional profile at 30 hPa in October over the time period 1998-2010. *Same as previous Figure, but for the 30 hPa level.*

latter being highly inhomogeneous in space and time. The winter hemisphere surf zone thereby constitutes a region of strong stirring and mixing, whereas the subtropical edges and the polar vortex are barriers to transport and mixing processes. Failing to reproduce the strength of these mixing barriers can lead to wrong distributions of long-lived and reactive trace gas species with potentially significant impacts on the ozone chemistry. The meridional profile of methane (or any other long-lived trace gas) reveals the existence of transport and mixing barriers in regions where tracer gradients are large. On the other hand, small tracer gradients indicate regions of strong mixing.

Only three instruments participating in the SPARC Data Initiative measured CH_4. The conclusions of *Chapter 4*, supported by other validation studies from the literature, suggest to treat the two MIPAS retrievals (high-spectral and high-spatial resolution) as two different instruments, hence **Figures 5.1.10** and **5.1.11** include four datasets each.

Figure 5.1.10 shows the meridional profile of methane at 50 hPa. The uncertainty range in step 1 is relatively large, especially in the SH polar vortex region, where the diagnostic is used to test the relative strengths of mixing across the polar vortex edge versus descent within the polar vortex. Removing multi-annual mean values with an interannual variability larger than the ±2σ standard deviations from step 1 and accounting for interannual variability yields a much smaller uncertainty range. This uncertainty range compared to the one used in *Eyring et al.* [2006] is shown to have improved in two aspects. First, the strong gradient across the polar vortex edge is much better defined than by using HALOE measurements alone. Second, HALOE mean values are much lower than the new multi-instrument mean values, in particular within the polar vortex region. The models (from Figure 5 in *Eyring et al.* [2006]) would hence compare much more favorably to the new instrument mean than to the old measurement diagnostic derived from HALOE. Note that the HALOE reference does not improve using a more limited range of years (*e.g.*, 2003-2005), but loses latitudinal coverage due to increasing sampling limitations towards the end of the mission.

Figure 5.1.11 shows the meridional profile of methane at 30 hPa. This level is chosen in order to illustrate that the comparison between the HALOE reference (as calculated in an equivalent way to that used in the CCMVal report at 50 hPa) and the multi-instrument mean and standard deviation from the SPARC Data Initiative datasets is altitude dependent. The comparison has much improved in terms of latitudinal structure, although the HALOE mean

values are still generally somewhat lower than those of the other instruments.

5.1.3 Recommendations for short-lived species

Short-lived species are characterised by chemically driven variations linked to the local solar time (LST). Limb-viewing instruments measure at LSTs that can differ from instrument to instrument, and between seasons and latitudes for the same instrument. Most of the instruments measure two distinct LSTs per latitude. These instruments are in polar sun-synchronous orbits, with one LST for the ascending portion of the orbit and one forthe descending portion. In the case of solar occultation sounders, measurements correspond to sunrise and sunset as seen from the satellite and the LSTs shift with the day of year.

The SPARC Data Initiative produced two types of climatologies for the diurnally varying species; climatologies from observations binned by LST (unscaled), and climatologies from observations scaled to a common LST. The climatologies from instruments in a sun-synchronous orbit are generally based on measurements separated into am and pm data. Climatologies from instruments that observe from non sun-synchronous orbits are generally separated into daytime and night-time measurements. Exceptions are the climatologies from solar occultation measurements which are based on data separated into local sunrise and sunset measurements. Additional climatologies are compiled using a photochemical box model to scale the measurements to a common LST, as explained in detail in *Section 3.1.2*.

When evaluating short-lived species from chemistry-climate models with the SPARC Data Initiative climatologies, the comparisons will be meaningless in most cases, if the monthly zonal mean model output is constructed in the traditional way by averaging over all longitudes at each output time step. Since most of the SPARC Data Initiative climatologies correspond to specific LSTs or times of day, the model output needs to be sampled in a similar manner. Even for instruments like SMILES, that observe species at varying LST because of their non sun-synchronous orbit, the constructed zonal mean climatologies are biased towards particular LSTs as a result of the non-homogeneous sampling patterns [*Kreyling et al.*, 2013]. Ideally, model data should be sampled with the satellite sampling patterns including the position and LST of each measurement. Trace gas climatologies derived from thus sampled model fields can be directly compared to the trace gas climatologies from the respective satellite instrument. While this approach is well suited for the comparison of short-lived species, it also means a lot of effort given that each satellite instrument has a different sampling pattern. Alternatively, the model output could be filtered according to LST in a manner similar to the SPARC Data Initiative climatologies in order to construct datasets corresponding to a particular LST, am/pm, day/night, or local sunrise/local sunset conditions. Another possibility is to restrict comparisons between model and satellite climatologies of short-lived species to latitude and altitude regions where the diurnal variations are small. Guidelines for appropriate comparisons of the individual short-lived species are given below.

- NO measurements show strong gradients at sunrise and sunset and model output should be filtered to construct sunrise and sunset (comparable to ACE-FTS or HALOE) or 10am LST (comparable to MIPAS or scaled ACE-FTS) climatologies.
- NO_2 diurnal variations are also most pronounced during sunset/sunrise. Model data should be filtered in order to construct sunrise/sunset (comparable to HALOE, SAGE II, POAM III, SAGE III and ACE-FTS) or 10am/10pm LST (comparable to MIPAS, SCIAMACHY, GOMOS or scaled OSIRIS, HIRDLS, and ACE-FTS) climatologies. If the model output is binned into daytime or night-time data instead (comparable to MIPAS, SCIAMACHY, GOMOS, OSIRIS, HIRDLS am/pm) differences of up to 20-30% can arise from the diurnal variations.
- NO_x is longer lived and has small diurnal variations in the MS. Data should be filtered to construct sunrise/sunset (comparable to HALOE and ACE-FTS) or 10am/10pm (comparable to MIPAS, SCIAMACHY, or scaled OSIRIS, and ACE-FTS) climatologies. Comparison of unfiltered monthly zonal mean climatologies can result in differences of around 20%. Binning the model output into daytime/night-time will not improve the comparison since there are no pronounced gradients at sunrise/sunset.
- HNO_3 is fairly long-lived in the UT to MS and shows a weak diurnal cycle in the US which increases further in the LM. Zonal mean climatologies can be compared directly at altitudes below 3 hPa.
- HNO_4, $ClONO_2$ and N_2O_5 climatologies show strong diurnal cycles above 10 hPa (100 hPa for N_2O_5) where model data needs to be binned according to sunrise/sunset (comparable to ACE-FTS) or 10am/10pm data (comparable to MIPAS). Below 10 hPa (100 hPa for N_2O_5), diurnal variations are weak allowing for a direct comparison of the datasets corresponding to different LSTs.
- ClO and BrO exhibit strong diurnal variations most pronounced during sunset/sunrise and with decreasing amplitude towards the USLM. Daytime variations are much smaller than night-time variations. For ClO, model data should be filtered in order to construct sunrise/sunset (comparable to SMR) or daytime climatologies (comparable to Aura-MLS pm, SMILES daytime, MIPAS am, or scaled daytime SMR climatologies). Comparisons should focus on the tropical/mid-latitude US. For BrO, model output should be filtered to construct daytime climatologies (comparable to scaled OSIRIS, scaled SCIAMACHY, or daytime SMILES climatologies). Comparisons should focus on altitude levels above 20 hPa. HOCl shows in contrast strong diurnal variations and model data need to be compiled according to instrument measurement times for a more meaningful comparison.

- HO_2 shows a strong diurnal cycle with smaller variations during daytime than during night-time. Model data can be binned into daytime climatologies (comparable to SMILES and Aura-MLS daytime) and compared in the altitude region between 10 and 0.5 hPa. OH has a strong diurnal cycle and model output should be filtered in order to construct daytime 2pm climatologies (comparable to Aura-MLS). CH_2O and CH_3CN show small diurnal variations, thus allowing for a direct comparison of datasets even if they apply to different LSTs.

5.1.4 Suggestions for new diagnostics

The monthly zonal mean SPARC Data Initiative datasets provide a unique source of observational data for model evaluation diagnostics. Here, we present suggestions for new diagnostics covering different aspects of model validation. The new diagnostics use, in addition to the monthly zonal mean climatologies, parameters from the SPARC Data Initiative datasets that describe variability, location and timing of the underlying measurements.

CFC-11 mean profiles and standard deviations

Profiles of long-lived tracers (as also shown in *Section 5.1.2* for CH_4) have been used extensively over the past to analyse the effects of diabatic descent and mixing in the polar vortex [*SPARC,* 2010]. Here, we show CFC-11 profiles at the high SH latitudes (80°S-85°S) at the beginning (June) and end (September) of the Antarctic winter for MIPAS and the Whole Atmosphere Community Climate Model (WACCM) (**Figure 5.1.12**). The comparison of June and September CFC-11 profiles provides information on the combined effects of vortex descent, bringing lower CFC-11 mixing ratios downward, and of transport from lower latitudes, bringing higher CFC-11 mixing ratios towards the pole. Between 100 and 50 hPa, WACCM shows lower mixing ratios at the beginning of the austral winter but higher mixing ratios at the end of the winter when compared to MIPAS. Thus the CFC-11 decrease is not strong enough in the model suggesting that there is too little descent and/or too much mixing across the vortex edge.

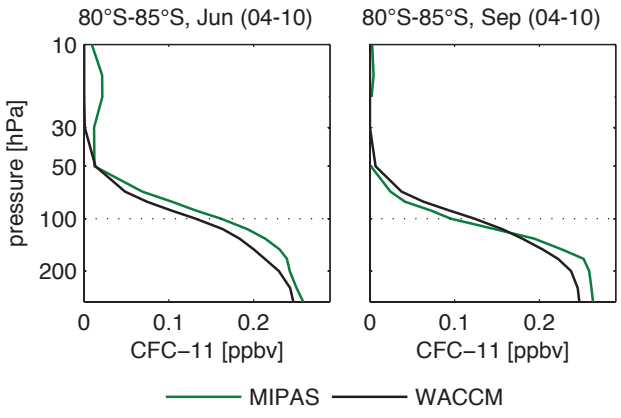


Figure 5.1.12: Vertical monthly zonal mean CFC-11 profiles for 80°S-85°S in June and September for MIPAS and WACCM.

Besides the monthly mean values, the SPARC Data Initiative datasets provide the standard deviations for each month, latitude bin and pressure level. **Figure 5.1.13** shows the standard deviation fields which describe the variability within each latitude band and month and are calculated over all given measurements in the respective bin. At 30 hPa (upper panels), elevated standard deviations at around 20°S/N indicate stronger variability in the trace gas field caused by breaking of planetary scale waves at the tropical pipe edge. The temporal extent (in the NH from December to March and in the SH from April to November) and magnitude of this event agree quite well between model and observations. Note that, at altitudes below 70 hPa, breaking synoptic scale waves cause more stirring and therefore prevent strong tracer gradients or any maxima in the standard deviation field. At 100 hPa in the tropics (~20°S-20°N), WACCM shows similar mean values but much lower standard deviations than MIPAS, which is very likely caused by the natural variability in this region being smaller than the MIPAS measurement error [*Toohey et al.*, 2010]. Most of the MIPAS variability is indeed explained by the MIPAS random error estimated to be around 17 pptv. Consequently, the standard deviation from observational fields should only be used for model evaluations in regions where the natural variability is larger than the measurement error. However, at 30 hPa the comparison reveals a striking absence of variability in the model in the SH high latitudes throughout the year, but in particular during SH winter, when the observations show high variability. This result implies a too low dynamical activity in the model, which may be related to the SH cold bias chemistry-climate models exhibit in this region [*Austin et al.*, 2003].

The comparison of the standard deviation fields from MIPAS and WACCM at 100 hPa (**Figure 5.1.13**, lower panels) reveals the absence of a mixing minimum during summer in the SH mid-latitudes in the model. The SH vortex edge region shows comparable variability during SH late winter, but higher variability in the model in early winter. The situation reverses at the very high SH latitudes, where the model has much lower variability over most of the year. In particular, the low standard deviations in the model during the winter from June to September suggest that the inner vortex south of 70°S in WACCM is less disturbed than implied by the MIPAS observations. Thus the missing decrease of the WACCM CFC-11 profiles in the vortex during winter (seen from the profile comparisons in **Figure 5.1.12**) is probably caused by too weak diabatic descent and not by too strong in-mixing. MIPAS on the other hand, shows elevated standard deviations during the winter related to zonal asymmetries in the CFC-11 field which can be either caused by asymmetric descent or by in-mixing. One to two months after the vortex breakdown the standard deviation of the CFC-11 field increases due to longitudinal asymmetries. This phenomenon can be observed earlier in MIPAS (December) than in WACCM

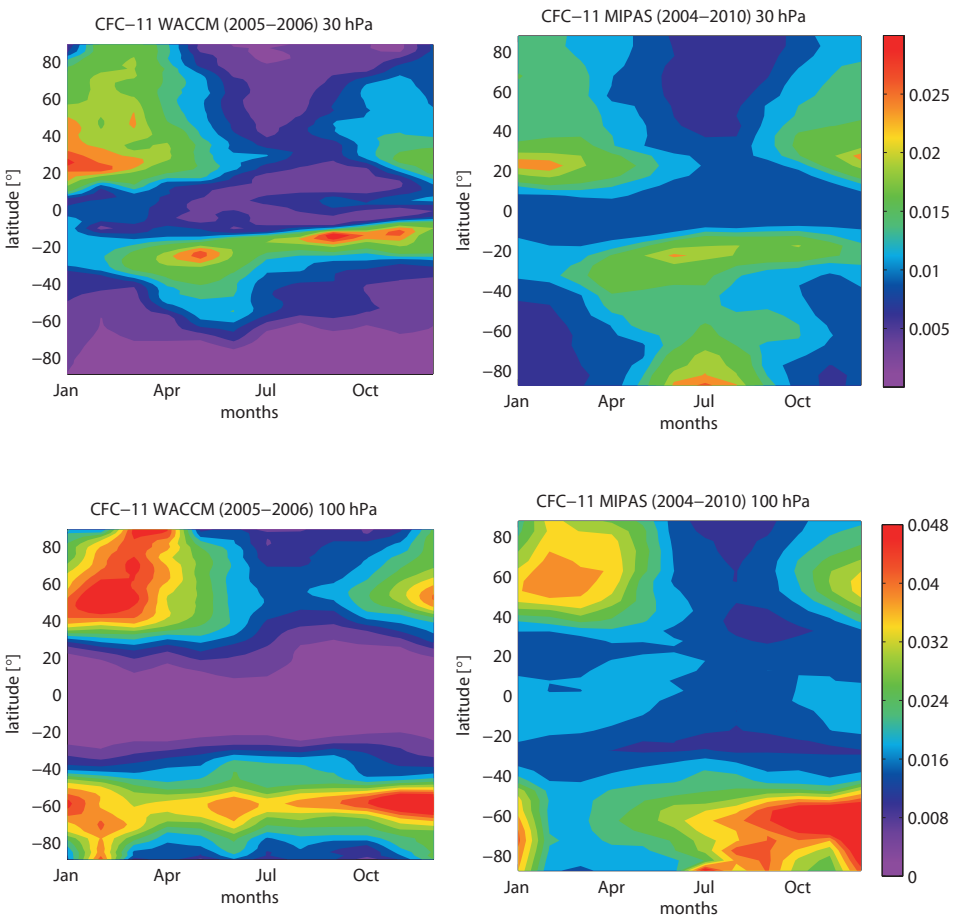


Figure 5.1.13: *Time-latitude cross-sections of CFC-11 standard deviation fields for MIPAS and WACCM at 30 hPa (upper panels) and at 100 hPa (lower panels). The standard deviation describes the variability within each latitude band and month and has been calculated over all given data points in the respective bin and month.*

(February) due to a late breakdown of the vortex in the model [*de laTorre et al.*, 2012]. At the NH high latitudes, the standard deviations show better agreement between observations and model suggesting more similarities in the dynamical situation.

CH_4 time-latitude evolution

While meridional and altitude profiles of CH_4 and N_2O have been extensively used in the past to test stratospheric transport in chemistry-climate models [*Eyring et al.*, 2006; *SPARC*, 2010; *Strahan et al.*, 2011], the SPARC Data Initiative monthly zonal mean climatologies lend themselves to also study the time evolution of these profiles. **Figure 5.1.14** shows to this end a comparison of the time-latitude evolution of CH_4 at two different pressure levels (2 and 10 hPa) between the multi-instrument mean derived from the HALOE, MIPAS, and ACE-FTS instruments, and the Canadian Middle Atmosphere Model using a simulation nudged to observed meteorology (CMAM30). As explained in more detail in *Chapter 4.3*, the feature at 2 hPa has been attributed to the equatorial Semi-Annual Oscillation (SAO) [*Choi and Holton*, 1991], with the maxima in tropical CH_4 coinciding with maxima in upwelling. The 2 hPa and 10 hPa levels are furthermore 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]. At 2 hPa, on the other hand, the minima show up in summer/autumn as the result of photochemistry, with CH_4 lifetimes decreasing to four months at these altitudes [*Randel et al.*, 1998; *Solomon*, 1986].

Comparison of CMAM30 with the observations yields overall encouraging results, with CMAM30 clearly indicating a SAO. Furthermore, the timing and extent of the low CH_4 in polar regions correspond well between observations and model at both levels. However, some differences can also be identified. For example in both hemispheres at 2 hPa, the photochemically induced minima during autumn are not quite as pronounced as in the observations. This could be due to a problem in the chemistry, but more likely results from too strong mixing between the tropics and the higher latitudes (partially due to numerical diffusion in the rather low model resolution). Likewise, the maxima seen in the tropics are not quite as pronounced as in the observations, along with the minima in polar regions at 10 hPa, indicating that CMAM30 exhibits too weak upwelling/downwelling or again too strong horizontal mixing. The overall good agreement between CMAM30 and the observations is partially due to using a model version that is driven by the observed meteorology. Note however that the influence of the nudging to the meteorological fields weakens towards higher altitudes above 10 hPa, so that the model seems to at least partially represent the right dynamical mechanisms that produce the SAO.

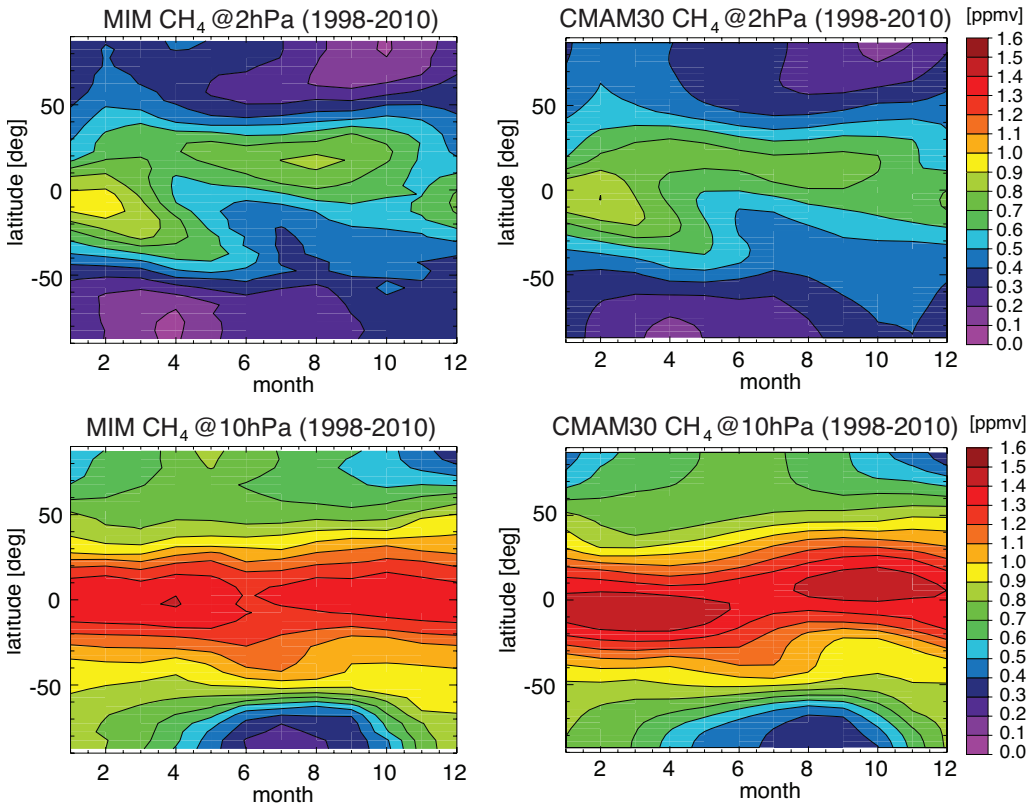


Figure 5.1.14: *Time-latitude cross-sections of CH_4 mixing ratios at 2 hPa (upper panels) and 10 hPa (lower panels) from the multi-instrument mean (left) and CMAM30 (right).*

5.2 Implications for merging activities

With monthly zonal mean time series of stratospheric constituents available from all the SPARC Data Initiative instruments, the obvious question is why these have not been merged into one homogeneous data product which globally covers multiple decades. The reason is that such a project is a challenge in itself which requires solving a number of technical and methodological problems. One needs to try to eliminate outliers or even whole datasets if many problems are discovered (*e.g.*, after a careful multi-instrument comparison). Currently, there is not even full agreement about what the most appropriate merging techniques are. Techniques range from a simple merge of two single datasets by accounting for an inter-instrument bias that is calculated over some overlap time period [*Bourrassa et al.*, 2014] to merging of multiple datasets including detailed calculations of uncertainties [*Froidevaux et al.*, 2015], statistical methods to fill in observational gaps [*Bodeker et al.*, 2013], or the use of a nudged chemistry-climate model as transfer function between the instruments [*Hegglin et al.*, 2014].

Some of the problems arising in data merging can be solved by directly using the parent datasets instead of the merged dataset and using an analysis tool that is immune against one or the other of these problems. One example is the trend estimator by *von Clarmann et al.* [2010] which is immune against biases between subsets of data. An ideal solution for the general data merging problem, however, does not yet exist. The first important step towards optimal data merging is to develop a common language and to develop schemes to evaluate and report retrieval errors, altitude resolution and content of prior information in the data in an inter-comparable manner. Given that all different merging techniques have their weaknesses and strengths, it remains important that independent research teams approach data merging so that their results can be compared and used to identify not only instrument errors but also uncertainties in the merging techniques themselves.

In the following sections we discuss the most prominent problem areas that arise in data merging.

5.2.1 Error characterisation of instruments

In the most straightforward scenario, multiple datasets are available for the same latitude bins and certain overlap time periods. In this case merging reduces to a weighted or unweighted mean of the data. The obvious advantage of weighting the data by their inverse estimated error, usually in terms of variance, is that reliable data dominate the merged product. Drawbacks and pitfalls, however, are:

- The error estimation schemes used for the different datasets may differ and different error types may be included. Thus, a better instrument can have larger error bars.
- For some instruments error covariances are reported, while for others only error bars are available. Optimal averaging, however, requires the covariance matrices.

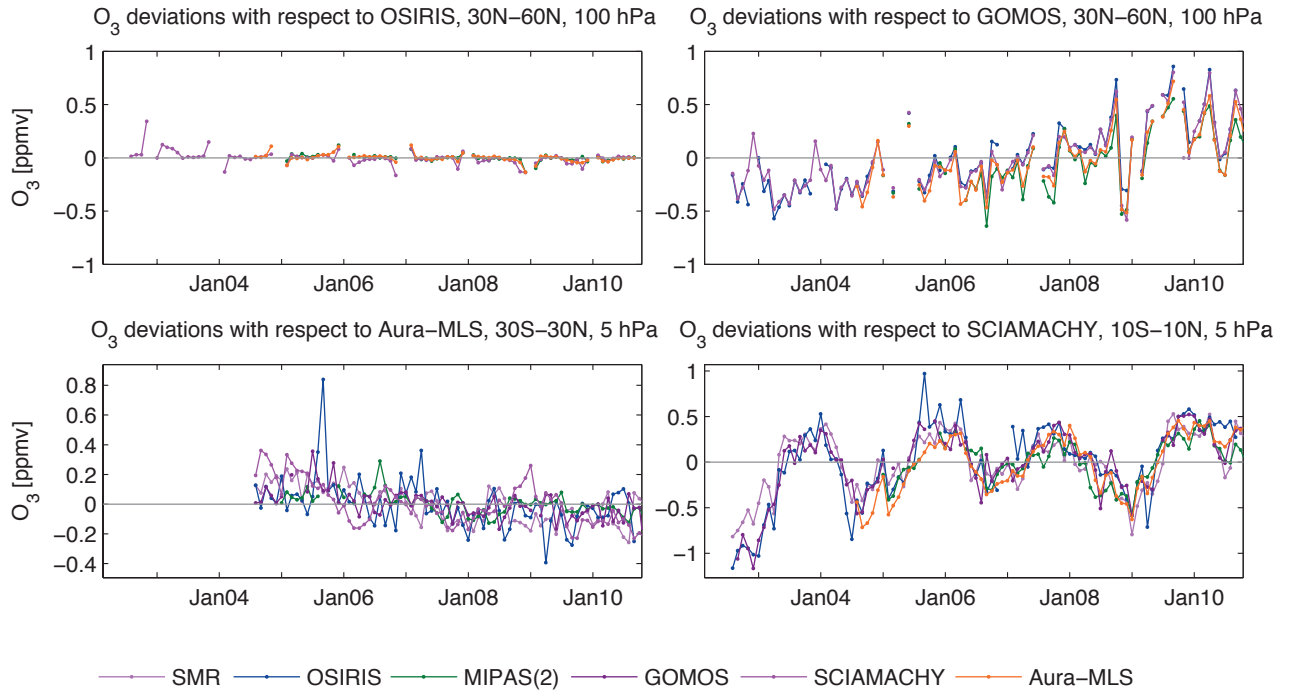


Figure 5.2.1: Time series of ozone deviations for 2002-2010. *Deviations of all instruments with respect to OSIRIS and GOMOS for 30°N-60°N at 100 hPa (upper panels), with respect to Aura-MLS for 30°S-30°N at 5 hPa (lower left panel) and with respect to SCIAMACHY for 10°S-10°N at 5 hPa (lower right panel) are shown.*

- For some instruments the error estimate includes by default the so-called smoothing error [*Rodgers*, 2000]. This quantity, however, does not follow the generalised Gaussian error propagation law [*von Clarmann,* 2014] and thus is not applicable to regridded data.

Trying to avoid these problems by using the sample standard error of the zonal mean instead of the error estimates is not as simple as one might think, because (a) in the case of regular sampling patterns measurements cannot be regarded as independent random samples and thus the standard error is not the sample standard deviation divided by the square root of the sample size [*Toohey and von Clarmann*, 2013]; and (b) sophisticated schemes are needed to distinguish the two components of the standard error of zonal means, namely measurement errors and natural variability [*Laeng et al.,* 2014; *Sofieva et al.*, 2014].

A particular problem is that the quality of the measurement can depend on the atmospheric state itself, *e.g.,* in infrared emission spectroscopy the signal is larger and thus the precision is better when it is warmer. Weighting by the inverse error variance in such a case would introduce a representativeness bias towards warmer parts of the atmosphere.

Another problem arises from denotation ambiguities. Many terms used for error characterisation are not clearly defined, used in different contexts, and have ambiguous meanings. Accuracy characterises in some cases the total error, in other cases only the systematic part, precision excluded. The systematic error in some documents includes all error sources except noise, in other cases only error terms which are - in amount and sign - time-independent. Noise often is referred to as the random part of the error while equally often it is used for the pure measurement noise only. Some total errors are more comprehensive than others. Some error budgets refer to an ideal point measurement and include the so-called smoothing error which characterises the expected difference between the atmospheric state at one idealised atmospheric point and in a finite air volume. Other error budgets refer to the atmospheric state at finite resolution and do not include the so-called smoothing error.

5.2.2 Drifts and jumps between datasets

Drifts within datasets are often unknown because, contrary to the usual validation measurements, drift estimation requires availability of long-term datasets. Even if these are available, it is not always clear which of the instruments compared to each other has a drift. For trace gases where a large number of instruments are available, such as ozone, the long-term changes of the differences can provide information on possible drifts [*Tegtmeier et al.*, 2013]. Therefore, for each instrument an analysis of the temporal variations of the differences with respect to each of the other instruments has been performed. Such time series are characterised by seasonal patterns and month-to-month variability. After removing the seasonal cycle, longer-term changes can be the dominant signal. However, for nearly all ozone datasets and regions included in this study the differences display no apparent long-term changes. One example for this consistency is shown in **Figure 5.2.1** (upper left panel) in the form of the instrument difference

with respect to OSIRIS in the NH mid-latitude LS. A few exceptions exist where clear changes of the differences over time can be identified (**Figure 5.2.1**). First, differences of all instruments with respect to GOMOS in the NH mid-latitude LS are mostly negative before 2008 and mostly positive afterwards indicating a change of GOMOS over time that is not seen by the other instruments. Note that GOMOS is excluded from the comparison to OSIRIS discussed above in order to present one example where the differences display no apparent long-term changes. For Aura-MLS, some discrepancies can be observed for the tropical US, with positive differences at the beginning and negative differences at the end of the time period, although not all instruments agree on this. SCIAMACHY differences in the tropics are dominated by the quasi-biennial oscillation (QBO) signal, while SMR (not shown here) displays larger values compared to the other datasets in 2003 but differences around zero after 2006. Note that here only drifts of a magnitude comparable to the deviations themselves have been identified; while for trend studies a more thorough analysis including possibly quite small long-term drifts is necessary.

Another option is the comparison of the instruments' time series with a model (used as a transfer function in the merging) which can yield additional evidence for which instrument is more likely to show a drift or a jump [*Hegglin et al.*, 2014]. An example for this is shown in **Figure 5.2.2**, where two data versions of SAGE II (v6.2 and v7.0) are compared to each other and a distinct difference in the beginning of the data record is revealed. The very good agreement between the maroon-coloured data version (v7.0) and the model (as seen in the bias-corrected differences fluctuating randomly around zero) provides the user with confidence that the red data version (v6.2) suffers from an inhomogeneity at the beginning of its record and therefore should not be used for merging during this time period.

Finally, the choice of well-established *in-situ* measurements as reference instruments, which usually are trusted more than remote sensing instruments, leads to the problem of often lacking statistical significance and representativeness due to low data amounts. Despite this shortcoming, ground-based measurements of ozone from sonde and lidar networks have been shown to allow for comprehensive analysis of the long-term stability of satellite ozone datasets [*Hubert et al.*, 2015]. A complication of all these types of validation studies is that the reference instrument (or model) itself needs thorough validation.

For certain regions and/or time periods, available datasets do not overlap in time. In this case, it is not clear if any jumps in the data reflect natural variability or instrument biases. A model as a transfer standard again can help here [*Hegglin et al.*, 2014]. While this approach may be seen as contaminating an otherwise purely empirical product with model information, it capitalises on our physical knowledge of the atmosphere and provides at least a best estimate of what happened during a time period when observations were not available.

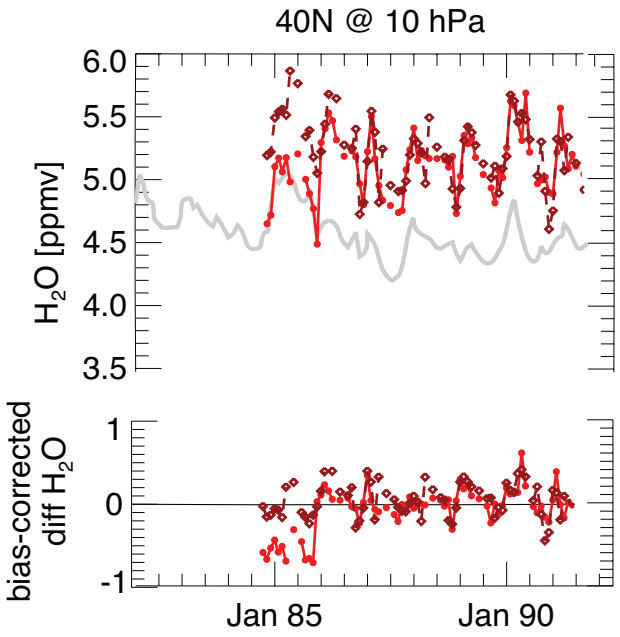


Figure 5.2.2: Model-based bias and drift estimates of observational data. *Time series of water vapour (upper panel) and bias-corrected difference to the model (lower panel) at 10 hPa and 40°N from two data versions of SAGE II water vapour (with red indicating v6.2 and maroon dots v7.0) and the Canadian Middle Atmosphere Model CMAM nudged to observed meteorology (grey line). Irregular behaviour in the bias-corrected differences reveals a problem in the red data version.*

5.2.3 Altitude resolution and a priori information

Key problems in any application where remote measurements of multiple instruments are considered are different altitude resolutions and different content of a priori information in the datasets. Some of the related problems can be solved by application of the averaging kernel matrix [*Rodgers,* 2000]; *e.g.*, the averaging kernel matrix can be used to degrade the altitude resolution of a high-resolution profile to make it comparable to a lower-resolution profile [*Rodgers and Connor*, 2003]. Such an approach has been applied in the SPARC Data Initiative when comparing the limb-viewing instruments with the nadir sounder TES in order to cross-validate ozone distributions in the UTLS with an independent dataset [*Section 4.27*; *Neu et al.*, 2014a]. TES measurements have been well-validated against ozonesondes in the UTLS and the dataset is frequently used for the evaluation of tropospheric ozone in chemistry-climate models. Observations of the higher vertical resolution limb sounders have been smoothed using the observational operator of TES. In the tropical UTLS, large positive biases of up to 50% have been identified for the limb-sounders with respect to the TES. While this study successfully provides a common basis for comparison of the large-scale ozone morphology in the UTLS, a couple of general problems remain unresolved for the general application of such comparisons:

- There exists a large number of datasets for which no averaging kernels are available, and can - due to the particular retrieval scheme used - not easily be produced.
- The application of the averaging kernel fails if the better resolved profile does not have sufficient altitude coverage to allow this operation for all relevant altitudes. There exist ad hoc solutions to this problem but these are not exact (see *Section 4.27* or *Neu et al.* [2014a]). In the SPARC Data Initiative evaluations, the TES a priori has been used to fill in the profiles below the lowest measurement level. To identify regions where the results are highly sensitive to this approach, virtual retrievals using two different filling methods have been calculated and compared.
- The situation is even worse if the altitude resolution of a measurement depends on the atmospheric state. This causes artefacts in estimated trends [*Yoon et al.*, 2013] or amplitudes of annual cycles (see *Section 4.2* or *Hegglin et al.* [2013]).

5.3 Implications for future planning of satellite limb-sounders

Past observations from limb satellite sounders have provided us with invaluable information on the chemistry (*e.g.*, *Waters et al.* [1993]; *Santee et al.* [1998]), transport (*e.g.*, *Park et al.* [2007]; *Stiller et al.* [2008]; *Hegglin et al.* [2009]; *Gille et al.* [2014]), and dynamics of the stratosphere (*e.g.*, *Randel et al.* [1993]; *Manney et al.* [2009]). This information has helped us understand many key aspects of the processes involved in stratospheric ozone depletion, the Antarctic ozone hole, and climate change. While we had a wealth of stratospheric limb observations during the past 30 years, it now has to be expected that there will be a lack of adequate limb measurements in the near future. This looming problem is due to an ageing fleet of currently still flying limb sounders (Aura-MLS, ACE-FTS, ACE-MAESTRO, OSIRIS and SMR) along with the lack of any concrete plans to launch new instruments except for SAGE III on the International Space Station (ISS) (which offers only limited spatial coverage) and the OMPS instruments (which only measure O_3, NO_2, and aerosol). These instruments may not be able to provide continuous temporal coverage, due to a nominal mission duration of Suomi NPP until 2016 and a replacement of the limb-viewing OMPS capacity on JPSS-2 in 2022 only.

The evaluations within the SPARC Data Initiative illustrate that there is no single best instrument that potentially covers all measurement needs, because instruments differ greatly in their measurement characteristics such as spatial and temporal sampling, viewing geometry, accuracy and precision, and measurement stability (*Chapters 2* and *3*). It is only through careful comparison between the instruments as done in this report that outliers can be detected, and weaknesses and strengths of instruments in measuring different species can be identified. An example is the realisation that SAGE II offers a valuable water vapour product that helps to extend the water vapour record from satellite observations (in particular HALOE) back to the late 1980s and also to improve this climate data record more generally [*Chapter 4.2*; *Hegglin et al.*, 2013; 2014].

Our evaluations also demonstrate clearly that there is no single instrument that can provide measurements of the full suite of atmospheric trace gas species with a high vertical and horizontal resolution, high accuracy and precision, and dense data coverage. Only a comprehensive set of high quality instruments that are complementary with respect to data coverage and target species allows development of a global picture of stratospheric composition. Such datasets enable among other things the analysis of temporal variations on different time scales and the quantification of important chemical budgets *e.g.*, of the chlorine family. As discussed in the previous *Section (5.2)*, data merging, even in the case of multiple overlapping instruments, poses a real challenge and complicates our understanding of long-term changes of the stratosphere in a changing climate. The future scenario we are currently facing with no overlap between instruments will render it impossible to derive reliable long-term changes of atmospheric trace constituents such as water vapor, ozone, and aerosol, and other important transport tracers.

Not only water vapour, but also other chemical trace gas species can be difficult to measure, especially when their atmospheric mixing ratios are close to the instruments' detection limits. Where agreement between instruments is found, the atmospheric mean state distributions and variability of trace gas species can be considered well-known (ozone [*Tegtmeier et al.*, 2013], water vapour [*Hegglin et al.*, 2013], N_2O, and CH_4 [*Hegglin et al.*, in prep.]). However, for other species that are measured by a few instruments only and for which not many ground-based validation measurements are available, our knowledge is still limited (many short-lived species such as HO_2, OH, BrO, ClO, *etc.*). It is key for the future planning of satellite limb sounders to design measurement systems that not only fit the purpose of covering specific measurement needs (in terms of scientific research question, region of interest, resolution, accuracy and precision, species list required), but also that offer redundancy between measurements, so that problems can be identified and adequately investigated.

References

Aghedo, A. M., *et al.*, 2011: The impact of orbital sampling, monthly averaging and vertical resolution on climate chemistry model evaluation with satellite observations. *Atmos. Chem. Phys.*, **11**, 6493-6514, doi:10.5194/acp-11-6493-2011.

Allen, D. R., *et al.*, 2000: Antarctic polar descent and planetary wave activity observed in ISAMS CO from April to July 1992. *Geophys. Res. Lett.*, **27**, 665-668, doi:10.1029/1999GL010888.

Angell, J. K., *et al.*, 1985: Ground-based and satellite evidence for a pronounced total-ozone minimum in early 1983 and responsible atmospheric layers. *Mon. Weather Rev.*, **113**, 641-646, doi:10.1175/1520-0493(1985)113<0641:GBASEF>2.0.CO;2.

Arijs, E., and G. Brasseur, 1986: Acetonitrile in the stratosphere and implications for positive ion composition. *J. Geophys. Res.*, **91**, 4003-4016, doi: 10.1029/JD091iD03p04003.

Austin, J. and N. Butchart, 2003: Coupled chemistry-climate model simulations for the period 1980 to 2020: Ozone depletion and the start of ozone recovery. *Q. J. Roy. Meteor. Soc.*, **129**, 3225-3249, doi:10.1256/qj.02.203.

Austin, J., *et al.*, 2003: Uncertainties and assessments of chemistry-climate models of the stratosphere. *Atmos. Chem. Phys.*, **3**, 1-27, doi:10.5194/acp-3-1-2003.

Austin, J., and F. Li, 2006: On the relationship between the strength of the Brewer-Dobson circulation and the age of stratospheric air. *Geophys. Res. Lett.*, **33**, L17807, doi:10.1029/2006GL026867.

Baehr, J., *et al.*, 2005: Validation of MIPAS-ENVISAT CH_4, N_2O, CFC-11 and CFC-12 by airborne in situ observations. Proceedings of the 2004 Envisat & ERS Symposium, SP-572, 6-10 September 2004, Salzburg, Austria, published by: ESA Publications Division, ESTEC, Noordwijk, The Netherlands, ISBN: 92-9092-883-2.

Baldwin, M. P., *et al.*, 2001: The quasi-biennial oscillation. *Rev. Geophys.*, **39**, 179-229, doi:10.1029/1999RG000073.

Barath, F. T., *et al.*, 1993: The Upper Atmosphere Research Satellite Microwave Limb Sounder instrument. *J. Geophys. Res.*, **98**, 10751-10762, doi:10.1029/93JD00798.

Baron, P., *et al.*, 2009: HO_2 measurements in the stratosphere and the mesosphere from the sub-millimetre limb sounder Odin/SMR. *Int. J. of Remote Sens.*, **30**, 4195-4208, http://dx.doi.org/10.1080/01431160902822831.

Baron, P., *et al.*, 2011: The Level 2 research product algorithms for the Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES). *Atmos. Meas. Tech.*, **4**, 2105-2124, doi:10.5194/amt-4-2105-2011.

Barth, C. A., 1992: Nitric oxide in the lower thermosphere. *Planet. Space Sci.,* **40**, 315-336, doi:10.1016/0032-0633(92)90067-X.

Bates, D. R., and M. Nicolet, 1950: The photochemistry of atmospheric water vapor. *J. Geophys. Res.*, **55**, 301-327, doi:10.1029/JZ055i003p00301.

Bauer, R., *et al.*, 2012: Validation of SCIAMACHY limb NO_2 profiles using solar occultation measurements. *Atmos. Meas. Tech.*, **5**, 1059-1084, doi:10.5194/amt-5-1059-2012.

Beer, R., *et al.*, 2001: Tropospheric Emission Spectrometer for the Earth Observing System's Aura satellite. *Appl. Opt.*, **40**, 2356-2367, doi:10.1364/AO.40.002356.

Beer, R., 2006: TES on the Aura mission: Scientific objectives, measurements, and analysis overview. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1102-1105, doi:10.1109/TGRS.2005.863716.

Benson, C. M., *et al.*, 2006: Polar stratospheric clouds in the 1998-2003 Antarctic vortex: Microphysical modeling and Polar Ozone and Aerosol Measurement (POAM) III observations. *J. Geophys. Res.*, **111**, D18206, doi:10.1029/2005JD006948.

Bernath, P. F., *et al.*, 2005: Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Lett.*, **32**, L15S01, doi:10.1029/2005GL022386.

Bernath, P., 2006: Atmospheric Chemistry Experiment (ACE): Analytical Chemistry from Orbit. *Trends Anal. Chem.*, **25**, 647-654, doi:10.1016/j.trac.2006.05.001.

Bertaux, J. L., *et al.*, 2010: Global ozone monitoring by occultation of stars: An overview of GOMOS measurements on Envisat. *Atmos. Chem. Phys.*, **10**, 12091-12148, doi:10.5194/acp-10-12091-2010.

Bhartia, P. K., *et al.*, 2004: Solar Backscatter Ultraviolet (SBUV) version 8 profile algorithm. Proceedings of the Quadrennial Ozone Symposium 2004, edited by: C. Zerefos, *Int. Ozone Comm.*, Athens, Greece, pp295-296.

Bodeker, G. E., *et al.*, 2013: A vertically resolved, global, gap-free ozone database for assessing or constraining global climate model simulations. *Earth Syst. Sci. Data*, **5**, 31-43, doi:10.5194/essd-5-31-2013.

Boone, C. D., *et al.*, 2005: Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer. *Appl. Opt.*, **44**, 7218-7231, doi:10.1364/AO.44.007218.

Bourassa, A. E., *et al.*, 2007: Stratospheric aerosol retrieval with optical spectrograph and infrared imaging system limb scatter measurements. *J. Geophys. Res.*, **112**, D10217, doi:10.1029/2006JD008079.

Bourassa, A. E., *et al.*, 2012: Odin-OSIRIS stratospheric aerosol data product and SAGE III intercomparison. *Atmos. Chem. Phys.*, **12**, 605-614, doi:10.5194/acp-12-605-2012.

Bourassa, A. E., *et al.*, 2014: Trends in stratospheric ozone derived from merged SAGE II and Odin-OSIRIS satellite observations. *Atmos. Chem. Phys.*, **14**, 6983-6994, doi:10.5194/acp-14-6983-2014.

Bovensmann, H., *et al.*, 1999: SCIAMACHY – Mission objectives and measurement modes. *J. Atmos. Sci.*, **56**, 127-150, doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2.

Bowman, K. W., *et al.*, 2002: Capturing time and vertical variability of tropospheric ozone: A study using TES nadir retrievals. *J. Geophys. Res.*, **107**, 4723, doi:10.1029/2002JD002150.

Bowman, K. W., *et al.*, 2006: Tropospheric Emission Spectrometer: Retrieval method and error analysis. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1297-1307, doi:10.1109/TGRS.2006.871234.

Bowman, K. W., *et al.*, 2013: Evaluation of ACCMIP outgoing longwave radiation from tropospheric ozone using TES satellite observations. *Atmos. Chem. Phys.*, **13**, 4057-4072, doi:10.5194/acp-13-4057-2013.

Boxe, C. S., *et al.*, 2010: Validation of northern latitude Tropospheric Emission Spectrometer stare ozone profiles with ARC-IONS sondes during ARCTAS: Sensitivity, bias and error analysis. *Atmos. Chem. Phys.,* **10**, 9901-9914, doi:10.5194/acp-10-9901-2010.

Bracher, A., *et al.*, 2005: Cross comparisons of O_3 and NO_2 measured by the atmospheric ENVISAT instruments GOMOS, MIPAS, and SCIAMACHY. *Adv. Space Res.*, **36**, 855-867, doi:10.1016/j.asr.2005.04.005.

Brasseur, G. and S. Solomon, 1984: Aeronomy of the middle atmosphere: chemistry and physics of the stratosphere and mesosphere. D. Reidel Publishing Company: Hingham, MA.

Brasseur, G., *et al.*, 1998: MOZART, A global chemical transport model for ozone and related chemical tracers 1. Model description. *J. Geophys. Res.*, **103**, 28265-28289, doi:10.1029/98JD02397.

Brohede, S., *et al.*, 2007a: A Stratospheric NO_2 Climatology from Odin/OSIRIS Limb-Scatter Measurement. *Can. J. Phys.*, **85**, 1253-1274, doi:10.1139/P07-141.

Brohede, S., *et al.*, 2007b: Validation of Odin/OSIRIS Stratospheric NO_2 Profiles. *J. Geophys. Res.*, **112**, D07310, doi:10.1029/2006JD007586.

Brohede, S., *et al.*, 2008: Odin stratospheric proxy NO_y measurements and climatology. *Atmos. Chem. Phys.*, **8**, 5731-5754.

Brown, A. T., *et al.*, 2011: Trends in atmospheric halogen containing gases since 2004. *J. Quant. Spectrosc. Ra.*, **112**, 2552-2566, doi:10.1016/j.jqsrt.2011.07.005.

Bruehl, C., *et al.*, 1996: Halogen Occultation Experiment ozone channel validation. *J. Geophys. Res.*, **101**, 10217, doi:10.1029/95JD02031.

Buehler, S. A., *et al.*, 2005: ARTS, the atmospheric radiative transfer simulator. *J. Quant. Spectros. Rad. Transf.*, **91**, 65-93, http://dx.doi.org/10.1016/j.jqsrt.2004.05.051.

Burrows, J. P., *et al.*, 1995: SCIAMACHY – Scanning Imaging Absorption Spectrometer for Atmospheric Chartography. *Acta Astronautica,* **35**, 445, doi:10.1016/0094-5765(94)00278-T.

Carleer, M. R., *et al.*, 2008: Validation of water vapour profiles from the Atmospheric Chemistry Experiment (ACE). *Atmos. Chem. Phys. Discuss.*, **8**, 4499-4559, doi:10.5194/acpd-8-4499-2008.

Carslaw, K. S., *et al.*, 1994: Stratospheric aerosol growth and HNO_3 gas phase depletion from coupled HNO_3 and water uptake by liquid particles. *Geophys. Res. Lett.*, **21**, 2479-2482, doi:10.1029/94GL02799.

Carslaw, K.S. *et al.*, 1997: Modeling the composition of liquid stratospheric aerosols. *Rev. Geophys.*, **35**, 125–154, doi:10.1029/97RG00078.

Chauhan, S., *et al.*, 2009: MIPAS reduced spectral resolution UTLS-1 mode measurements of temperature, O_3, HNO_3, N_2O, H_2O and relative humidity over ice: retrievals and comparison to MLS. *Atmos. Meas. Tech.*, **2**, 337-353, doi:10.5194/amt-2-337-2009.

Chipperfield, M. P., *et al.*, 1994: A two-dimensional model study of the QBO signal in SAGE II NO_2 and O_3. *Geophys. Res. Lett.*, **21**, 589-592, doi: 10.1029/94GL00211.

Choi, W. K, and J. R. Holton, 1991: Transport of N_2O in the stratosphere related to the equatorial semiannual oscillation. *J. Geophys. Res.*, **96**, 22543-22557, doi:10.1029/91JD02263.

Chu, W. P., and M. P. McCormick, 1979: Inversion of stratospheric aerosol and gaseous constituents from spacecraft solar extinction data in the 0.38-1.0 μm wavelength region. *Appl. Opt.*, **18**, 1404-1414, https://doi.org/10.1364/AO.18.001404.

Chu, W. P., *et al.*, 1989: SAGE II Inversion Algorithm. *J. Geophys. Res.*, **94**, 8339-8351, doi:10.1029/JD094iD06p08339.

Clerbaux, C., *et al.*, 2008: CO measurements from the ACE-FTS satellite instrument: data analysis and validation using groundbased, airborne and spaceborne observations. *Atmos. Chem. Phys.*, **8**, 2569-2594, doi:10.5194/acp-8-2569-2008.

Coheur, P.-F., *et al.*, 2007: ACE-FTS observation of a young biomass burning plume: first reported measurements of C_2H_4, C_3H_6O, H_2CO and PAN by infrared occultation from space. *Atmos. Chem. Phys.*, **7**, 5437-5446, doi:10.5194/acp-7-5437-2007.

Collins, W. D., *et al.*, 2004: Description of the NCAR Community Atmosphere Model (CAM3). NCAR Technical Note, NCAR/TN-464+STR, 226 pp.

Connor, B. J., *et al.*, 2007: Comparison of ClO measurements from the Aura Microwave Limb Sounder to ground-based microwave measurements at Scott Base, Antarctica, in spring 2005. *J. Geophys. Res.*, **112**, D24S42, doi:10.1029/2007JD008792.

Conway, R. R., *et al.*, 1999: Middle atmosphere high resolution spectrograph investigation. *J. Geophys. Res.*, **104**, 16327-16348, doi:10.1029/1998JD100036.

Conway, R. R., *et al.*, 2000: Satellite observations of upper stratospheric and mesospheric OH: The HO_x dilemma. *Geophys. Res. Lett.*, **27**, 2613-2616, doi:10.1029/2000GL011698.

Crutzen, P. J., 1970: The influence of nitrogen oxides on the atmospheric ozone content. *Q. J. R. Meteorol. Soc.*, **96**, 320-325, doi:10.1002/qj.49709640815.

Cunnold, D. M., *et al.*, 1991: Validation of SAGE II NO_2 Measurements. *J. Geophys. Res.*, **96**, 12913-12925, doi:10.1029/91JD01344.

Damadeo, R. P., *et al.*, 2013: SAGE version 7.0 algorithm: Application to SAGE II. *Atmos. Meas. Tech.,* **6**, 3539–3561, doi:10.5194/amt-6-3539-2013.

Damiani, A., *et al.*, 2012: Impact of January 2005 solar proton events on chlorine species. *Atmos. Chem. Phys*., **12**, 4159-4179, doi:10.5194/acp-12-4159-2012.

Daniel, J. S., and S. Solomon, 1998: On the climate forcing of carbon monoxide. *J. Geophys. Res.*, **103**, 13249-13260, doi:10.1029/98JD00822.

Degenstein, D., *et al.*, 2002: Volume emission rate tomography from a satellite platform. *Appl. Opt.*, **42**, 1441-1450, https://doi.org/10.1364/AO.42.001441.

Degenstein, D., *et al.*, 2009: Limb scatter ozone retrieval from 10 to 60 km using a multiplicative reconstruction technique. *Atmos. Chem. Phys.*, **9**, 6521-6529, doi:10.5194/acp-9-6521-2009.

de Gouw, J. A., *et al.*, 2003: Emission sources and ocean uptake of aceotnitrile (CH_3CN) in the atmosphere. *J. Geophys. Res.*, **108**(D11), doi:10.1029/2002JD002897.

de la Torre, L., *et al.*, 2012: Climatology and characteristics of stratospheric sudden warmings in the Whole Atmosphere Community Climate Model. *J. Geophys. Res.*, **117**, D04110, doi:10.1029/ 2011JD016840.

De Mazière, M., *et al.*, 2008: Validation of ACE-FTS v2.2 methane profiles from the upper troposphere to the lower mesosphere. *Atmos. Chem. Phys.*, **8**, 2421-2435, doi:10.5194/acp-8-2421-2008.

DeMore, W. B., *et al.*, 1997: Chemical kinetics and photochemical data for use in stratospheric modeling. JPL publication 92-94, Jet Propulsion Laboratory, Pasadena, California.

Douglass, A. R., *et al.*, 1999: Choosing meteorological input for the global modeling initiative assessment of high-speed aircraft. *J. Geophys. Res.*, **104**(D22), 27,545-27,564, doi:10.1029/1999JD900827.

Drummond, J. R., *et al.*, 1980: The stratospheric and mesospheric sounder on NIMBUS-7. *Philos. T. Roy. Soc.*, **A296**, 219-241, doi:10.1098/rsta.1980.0166.

Dufour, G., *et al.*, 2009: Global upper-tropospheric formaldehyde: seasonal cycles observed by the ACE-FTS satellite instrument. *Atmos. Chem. Phys.*, **9**, 3893-3910, doi:10.5194/acp-9-3893-2009.

Dupuy, E., *et al.,* 2004: Strato-mesospheric measurements of carbon monoxide with the Odin Sub-millimetre Radiometer: Retrieval and First Results. *Geophys. Res. Lett.*, **31**, L20101, doi:10.1029/2004GL020558.

Dupuy, E., *et al.*, 2009: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE). *Atmos. Chem. Phys.*, **9**, 287-343, doi:10.5194/acp-9-287-2009.

Ejiri, M. K., *et al.*, 2006: Validation of the improved limb atmospheric spectrometer-II (ILAS-II) Version 1.4 nitrous oxide and methane profiles. *J. Geophys. Res.*, **111**, D22S90, doi:10.1029/2005JD006449.

Ekström, M., *et al.*, 2008: Comparison of satellite limb-sounding humidity climatologies of the uppermost tropical troposphere. *Atmos. Chem. Phys.*, **8**, 309-320, doi:10.5194/acp-8-309-2008.

Ekström, M., and Eriksson, P., 2008: Altitude resolved ice-fraction in the uppermost tropical troposphere. *Geophys. Res. Lett.*, **35**, L13822, doi:10.1029/2008GL034305.

Englert, C. R., *et al.*, 2008: First results from the Spatial Heterodyne Imager for Mesospheric Radicals (SHIMMER), Diurnal variation of mesospheric hydroxyl. *Geophys. Res. Lett.*, **35**, L19813, doi:10.1029/2008GL035420.

Englert, C. R., *et al.*, 2010: Spatial Heterodyne Imager for Mesospheric Radicals on STPSat-1. *J. Geophys. Res.*, **115**, D20306, doi:10.1029/2010JD014398.

Eparvier, F. G., *et al.*, 1994: Solar Mesosphere Explorer satellite measurements of El Chichon stratospheric aerosols, 2. Aerosol mass and size parameters. *J. Geophys. Res.*, **99**, 20533-20544, doi:10.1029/94JD01841.

Eriksson, P., C. Jiménez and S. A. Buehler, 2005: Qpack, a tool for instrument simulation and retrieval work. *J. Quant. Spectrosc. Rad. Transf.*, **91**, 47-64, http://dx.doi.org/10.1016/j.jqsrt.2004.05.050.

Ernst, F., *et al.*, 2009: Retrieval of stratospheric aerosol distributions from SCIAMACHY limb measurements: first steps and methodology. Proceedings Atmospheric Science Conference, Barcelona, Spain, 7-11 Sept 2009, ESA Special Publication SP-676.

Eyring, V., *et al.*, 2006: Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past. *J. Geophys. Res.*, **111**(D22), D22308.

Fahey, D. W., *et al.*, 2001: The detection of large HNO_3-containing particles in the winter Arctic stratosphere. *Science*, **291**, 1026-1031, doi:10.1126/science.1057265.

Fiorucci, I., *et al.*, 2013: Ground-based stratospheric O_3 and HNO_3 measurements at Thule, Greenland: an intercomparison with Aura MLS observations. *Atm. Meas. Tech.*, **6**, 2441-2453, doi:10.5194/amt-6-2441-2013.

Fischer, H., *et al.*, 2008: MIPAS: An instrument for atmospheric and climate research. *Atmos. Chem. Phys.,* **8**, 2151-2188, doi:10.5194/acp-8-2151-200.

Fish, D. J., *et al.*, 2000: Possible causes of stratospheric NO_2 trends observed at Lauder. *Geophys. Res. Lett.*, **20**, 3313-3316, doi:10.1029/2000GL011700.

Flocke, F, *et al.*, 1999: An examination of chemistry and transport processes in the tropical lower stratosphere using observations of long-lived and short-lived compounds obtained during STRAT and POLARIS. *J. Geophys. Res.*, **104**, 26625-26642, doi:10.1029/1999JD900504.

Folkins, I., *et al.*, 2006: Seasonal cycles of O_3, CO, and convective outflow at the tropical tropopause. *Geophys. Res. Lett.*, **33**, L16802, doi:10.1029/2006GL026602.

Forkman, P., *et al.*, 2012: Six years of mesospheric CO estimated from ground-based frequency-switched microwave radiometry at 57°N compared with satellite instruments. *Atmos. Meas. Tech.*, **5**, 2827-2841, doi:10.5194/amt-5-2827-2012.

Forster, P. M. and K. P. Shine, 2002: Assessing the climate impacts of trends in stratospheric water vapour. *Geophys. Res. Lett.,* **29**, 1086-1089, doi:10.1029/2001GL013909.

Frisk, U., *et al.*, 2003: The Odin satellite: I. Radiometer design and test. *Astron. Astrophys.*, **402**, L27(34), doi:10.1051/0004-6361:20030335.

Froidevaux, L., *et al.*, 2006: Early validation analyses of the atmospheric profiles from EOS MLS on the Aura satellite. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1106-1121, doi:10.1109/TGRS.2006.864366.

Froidevaux, L., *et al.*, 2008a: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements. *J. Geophys. Res.*, **113**, D15S20, doi:10.1029/2007JD008771.

Froidevaux, L., *et al.*, 2008b: Validation of Aura Microwave Limb Sounder HCl measurements. *J. Geophys. Res.*, **113**, D15S25, doi:10.1029/2007JD009025.

Froidevaux, L., *et al.*, 2015: Global Ozone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS): methodology and sample results with a focus on HCl, H_2O, and O_3. *Atmos. Chem. Phys.*, **15**, 10471-10507, doi:10.5194/acp-15-10471-2015.

Fueglistaler, S., *et al.*, 2009: The tropical tropopause layer. *Rev. Geophys.*, **47**, RG1004, doi:10.1029/2008RG000267.

Fujiwara, M., *et al.*, 2010: Seasonal to decadal variations of water vapour in the tropical lower stratosphere observed with balloon-borne cryogenic frost point hygrometers. *J. Geophys. Res.*, **115**, D18304, doi:10.1029/2010JD014179.

Funke, B., *et al.*, 2001: A new non-LTE retrieval method for atmospheric parameters from MIPAS-Envisat emission spectra. *Adv. Space Res.*, **27**, 1099-1104, doi:10.1016/S0273-1177(01)00169-7.

Funke, B., *et al.*, 2005a: Retrieval of stratospheric NO_x from 5.3 and 6.2 μm nonlocal thermodynamic equilibrium emissions measured by Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on Envisat. *J. Geophys. Res.*, **110**, D09302, doi:10.1029/2004JD005225.

Funke, B., *et al.*, 2005b: Downward transport of upper atmospheric NO_x into the polar stratosphere and lower mesosphere during the Antarctic 2003 and Arctic 2002/2003 winters. *J. Geophys. Res.,* **110**, D24308, doi:10.1029/2005JD006463.

Funke, B., *et al.*, 2009: Carbon monoxide distributions from the upper troposphere to the mesosphere inferred from 4.7 μm nonlocal thermal equilibrium emissions measured by MIPAS on Envisat. *Atmos. Chem. Phys.*, **9**, 2387-2411, doi:10.5194/acp-9-2387-2009.

Funke, B. and von Clarmann, T., 2012: How to average logarithmic retrievals? *Atmos. Meas. Tech.*, **5**, 831-841, doi:10.5194/amt-5-831-2012.

Funke, B., *et al.*, 2014: Mesospheric and stratospheric NO_Y produced by energetic particle precipitation during 2002-2012. *J. Geophys. Res. Atmos.*, **119**, 4429-4446, doi:10.1002/2013JD021404.

Fussen, D., *et al.,* 2001: Evolution of stratospheric aerosols in the post-Pinatubo period measured by the occultation radiometer experiment ORA. *Atmos. Env.*, **35**, 5067-5078.

Garcia, R. R., *et al.*, 2007: Simulation of secular trends in the middle atmosphere, 1950-2003. *J. Geophys. Res.*, **112**, D09301, doi:10.1029/2006JD007485.

Garcia, R. R., *et al.*, 2011: On the Determination of Age of Air Trends from Atmospheric Trace Species. *J. Atmos. Sci.*, **68**, 139-154, doi:10.1175/2010JAS3527.1.

Geller, L.S., *et al.,* 1997: Tropospheric SF_6: Observed latitudinal distribution and trends, derived emissions, and interhemispheric exchange time. *Geophys. Res.*, **24**, 675-678, doi:10.1029/97GL00523.

Gettelman, A., *et al.*, 2010: Multimodel assessment of the upper troposphere and lower stratosphere: Tropics and global trends. *J. Geophys. Res.*, **115**, D00M08, doi:10.1029/2009JD013638.

Gettelman, A., *et al.*, 2011: The extra tropical upper troposphere and lower stratosphere. *Rev. Geophys.*, **49**, RG3003, doi:10.1029/2011RG000355.

Gille, J., and J. M. Russell III, 1984: The Limb Infrared Monitor of the Stratosphere (LIMS) experiment: Experiment description, performance, and results. *J. Geophys. Res.*, **88**, 5125-5140.

Gille, J., and J. Barnett, 1992: The High-Resolution Dynamics Limb Sounder (HIRDLS). An instrument for the study of global change. The Use of EOS for Studies of Atmospheric Physics, J. Gille and G. Visconti (Eds.), North-Holland, pp433-450.

Gille, J., *et al.*, 2008: The High Resolution Dynamics Limb Sounder (HIRDLS): Experiment overview, results and validation of initial temperature data. *J. Geophys. Res.*, **113**, D16S43, doi:10.1029/2007JD008824.

Gille, J., and L. Gray, 2011: High Resolution Dynamics Limb Sounder, Earth Observing System (EOS), Data Description and Quality. Version 6 (V6), available from **www.eos.ucar.edu/hirdls/**, **http://disc.sci.gsfc.nasa.gov/data-holdings**, or **http://badc.nerc.ac.uk/browse/badc/hirdls**.

Gille, J., *et al.*, 2014: The role of midlatitude mixing barriers in creating the annual variation of total ozone in high northern latitudes. *J. Geophys. Res. Atmos.*, **119**, 9578-9595, doi:10.1002/2013JD021416.

Glaccum, W., *et al.*, 1996: The Polar Ozone and Aerosol Measurement (POAM II) Instrument. *J. Geophys. Res.*, **101**, 14479-14487, doi:10.1029/96JD00576.

Glatthor, N., *et al.*, 1998: Airborne remote sensing of NO_2 in the Arctic winter of 1994-1995 and comparison with a three-dimensional chemical transport model. *J. Geophys. Res.*, **103**, 13315-13326., doi:10.1029/98JD00521.

Glatthor, N., *et al.,* 1999: Intercomparison of the KOPRA and the RFM radiative transfer codes. Proc. European Symposium on Atmospheric Measurements from Space, ESAMS'99, 18-22 Jan 1999, Noordwijk, European Space Agency, ESTEC, Noordwijk, The Netherlands, pp757-764.

Glatthor, N., *et al.*, 2004: Spaceborne ClO observations by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) before and during the Antarctic major warming in September/October 2002. *J. Geophys. Res.*, **109**, D11307, doi:10.1029/2003JD004440.

Glatthor, N., *et al.*, 2005: Mixing processes during the Antarctic vortex split in September/October 2002 as inferred from source gas and ozone distributions from ENVISAT-MIPAS. *J. Atmos. Sci.*, **62**, 787-800, doi:10.1175/JAS-3332.1.

Glatthor, N., *et al.*, 2006: Retrieval of stratospheric ozone profiles from MIPAS/ENVISAT limb emission spectra: A sensitivity study. *Atmos. Chem. Phys.*, **6**, 2767-2781, doi:10.5194/acp-6-2767-2006.

Gordley, L. L., *et al.*, 1996: Validation of nitric oxide and nitrogen dioxide measurements made by the Halogen Occultation Experiment for UARS platform. *J. Geophys. Res.*, **101**, 10241-10266, doi:10.1029/95JD02143.

Griesfeller, A., *et al.*, 2008: Intercomparison of ILAS-II Version 1.4 and Version 2 target parameters with MIPAS-Envisat measurements. *Atmos. Chem. Phys.*, **8**, 825-843, doi:10.5194/acp-8-825-2008.

Grooß, J.-U., and J. M. Russell III, 2005: Technical note: A stratospheric climatology for O_3, H_2O, CH_4, NO_x, HCl and HF derived from HALOE measurements. *Atmos. Chem. Phys.*, **5**, 2797-2807, doi:10.5194/acp-5-2797-2005.

Gunson, M., *et al.*, 1996: The Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment: Deployment on the ATLAS Space Shuttle missions. *Geophys. Res. Lett.*, **23**, 2333-2336, doi:10.1029/96GL01569.

Haley, C., *et al.*, 2004: Retrievals of stratospheric O_3 and NO_2 profiles from Odin Optical Spectrograph and InfraRed Imager System (OSIRIS) limb-scattered sunlight measurements. *J. Geophys. Res.*, **109**, D16303, doi:10.1029/2004JD004588.

Haley, C. and S. Brohede, 2007: Status of the Odin/OSIRIS Stratospheric O_3 and NO_2 Data Products. *Can. J. Phys.*, **85**, 1177-1194, doi:10.1139/P07-114.

Hanson, D. R. and K. Mauersberger, 1988: Laboratory studies of the nitric acid trihydrate: implications for the south polar stratosphere. *Geophys. Res. Lett.*, **15**, 855-858, doi:10.1029/GL015i008p00855.

Harries, J. E., *et al.*, 1996: Validation of water vapor measurements from the Halogen Occultation Experiment. *J. Geophys. Res.*, **101**, 10205, doi:10.1029/95JD02933.

Harrison, J. J. and P. F. Bernath, 2013: ACE-FTS observations of acetonitrile in the lower stratosphere. *Atmos. Chem. Phys.*, **13**, 7405-7413, doi:10.5194/acp-13-7405-2013.

Hartmann, G. K., *et al.*, 1996: Measurements of O_3, H_2O and ClO in the middle atmosphere using the Millimeter-Wave Atmospheric Sounder (MAS). *Geophys. Res. Lett.*, **23**, 2313-2316, doi:10.1029/96GL01475.

Hasebe, F., 1994: Quasi-biennial oscillations of ozone and diabatic circulation in the equatorial stratosphere. *J. Atmos. Sci.*, **51**, 729-745, doi:10.1175/1520-0469(1994)051<0729:QBOOOA>2.0.CO;2.

Hassler, B., G.E. Bodeker, and M. Dameris, 2008: Technical Note: A new global database of trace gases and aerosols from multiple sources of high vertical resolution measurements. *Atmos. Chem. Phys.*, **8**, 5403-5421, doi:10.5194/acp-8-5403-2008.

Hauchecorne, A., *et al.*, 2005: First simultaneous global measurements of nighttime stratospheric NO_2 and NO_3 observed by Global Ozone Monitoring by Occultation of Stars (GOMOS)/ENVISAT in 2003. *J. Geophys. Res.*, **110**, D18301, doi:10.1029/2004JD005711.

Hauchecorne, A., *et al.*, 2010: Response of tropical stratospheric O_3, NO_2 and NO_3 to the equatorial Quasi-Biennial Oscillation and to temperature as seen from GOMOS/Envisat. *Atmos. Chem. Phys.*, **10**, 8873-8879, doi:10.5194/acp-10-8873-2010.

Hedin, A. E., 1991: Extension of the MSIS thermosphere model into the middle and lower atmosphere. *J. Geophys. Res.*, **96**(A2), 1159, Document ID: 19910040911.

Hegglin, M. I., *et al.*, 2008: Validation of ACE-FTS satellite data in the upper troposphere/lower stratosphere (UTLS) using non-coincident measurements. *Atmos. Chem. Phys.*, **8**, 1483-1499, doi:10.5194/acp-8-1483-2008.

Hegglin, M. I., *et al.*, 2009: A global view of the extratropical tropopause transition layer from Atmospheric Chemistry Experiment Fourier Transform Spectrometer O_3, H_2O, and CO. *J. Geophys. Res.*, **114**, D00B11, doi:10.1029/2008JD009984.

Hegglin, M. I., *et al.,* 2010: Multimodel assessment of the upper troposphere and lower stratosphere: Extratropics. *J. Geophys. Res.*, **115**, D00M09, doi:10.1029/2010JD013884.

Hegglin, M. I., *et al.*, 2013: SPARC Data Initiative: Comparison of water vapor climatologies from international limb satellite sounders. *J. Geophys. Res. Atmos.*, **118**, doi:10.1002/jgrd.50752.

Hegglin, M. I., *et al.*, 2014: Vertical structure of stratospheric water vapour trends derived from merged satellite data. *Nature Geoscience*, **7** (10), 768-776, ISSN 1752-0894 doi:10.1038/ngeo2236.

Hegglin, M. I., *et al.*, in prep.: SPARC Data Initiative: Comparison of aerosol climatologies from international limb satellite sounders.

Hervig, M. E., *et al.*, 1996a: Validation of temperature measurements from the Halogen Occultation Experiment. *J. Geophys. Res.*, **101**, 10277-10285, doi:10.1029/95JD01713.

Hervig, M. E., *et al.*, 1996b: Validation of aerosol measurements from the Halogen Occultation Experiment. J. *Geophys. Res.*, **101**, 10267-10275, doi:10.1029/95JD02464.

Hervig, M. E., and T. Deshler, 2002: Evaluation of aerosol measurements from SAGE II, HALOE, and balloonborne optical particle counters. *J. Geophys. Res.*, **107**, 4031, doi:10.1029/2001JD 000703.

Hocke, K., *et al.*, 2007: Comparison and synergy of stratospheric ozone measurements by satellite limb sounders and the ground-based microwave radiometer SOMORA. *Atmos. Chem. Phys.*, **7**, 4117-4131, doi:10.5194/acp-7-4117-2007.

Hofmann, D. J., and S. Solomon,1989: Ozone destruction through heterogeneous chemistry following the eruption of El Chichón. *J. Geophys. Res.*, **94**, 5029-5041, Paper No.: 88JD04231.

Holton, J. R., and W.-K. Choi, 1988: Transport circulation deduced from SAMS trace species data. *J. Atmos. Sci.*, **45**, 1929–1939.

Höpfner, M., *et al.*, 2004: First spaceborne observations of Antarctic stratospheric $ClONO_2$ recovery: Austral spring 2002. *J. Geophys. Res.*, **109**, doi:10.1029/2004JD004609.

Höpfner, M., *et al.*, 2007: Validation of MIPAS $ClONO_2$ measurements. *Atmos. Chem. Phys.*, **7**, 257-281, doi:10.5194/acp-7-257-2007.

Hoor, P., *et al.*, 2010: Transport timescales and tracer properties in the extratropical UTLS. *Atmos. Chem. Phys.*, **10**, 7929-7944, doi:10.5194/acp-10-7929-2010.

Hubert, D., *et al.*, 2016: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records. *Atmos. Meas. Tech.*, **9**, 2497-2534, doi:10.5194/amt-9-2497-2016.

Hurst, D. F., *et al.*, 2011: Stratospheric water vapour trends over Boulder, Colorado: Analysis of the 30 year Boulder record. *J. Geophys. Res.*, **116**, D02306, doi:10.1029/2010JD015065.

Isaksen, I. S. A., *et al.*, 2009: Atmospheric composition change: Chemistry-climate interactions. *Atmos. Env.,* **43**, 5138-5192, doi:10.1016/j.atmosenv.2009.08.003.

Jackman, C. H., *et al.*, 2008: Short and medium-term atmospheric constituent effects of very large solar proton events. *Atmos. Chem. Phys.*, **8**, 765-785, doi:10.5194/acp-8-765-2008.

Jiang, J. H., *et al.*, 2005: Comparison of GPS/SAC-C and MIPAS/ENVISAT temperature profiles and its possible implementation for EOS MLS observations. Earth Observation with CHAMP: Results from Three Years in Orbit, C. Reigber, H. Lühr, P. Schwintzer, and J. Wickert (Eds.), Springer-Verlag Berlin Heidelberg New York, pp573-578.

Jiang, Y. B., *et al.*, 2007: Validation of aura microwave limb sounder ozone by ozonesonde and lidar measurements. *J. Geophys. Res.*, **112**, D24S34, doi:10.1029/2007JD008776.

Jin, J. J., *et al.*, 2009: Comparison of CMAM simulations of carbon monoxide (CO), nitrous oxide (N_2O), and methane (CH_4) with observations from Odin/SMR, ACE-FTS, and Aura/MLS. *Atmos. Chem. Phys.*, **9**, 3233-3252, doi:10.5194/acp-9-3233-2009.

Johnson, D. G., *et al.*, 1995: Estimating the abundance of ClO from simultaneous remote sensing measurements of HO_2, OH, and HOCl.*Geophys. Res. Lett.*, **22**, 1869-1871, doi:10.1029/95GL01249.

Johnston, H., 1971: Reduction of stratospheric ozone by nitrogen oxide catalysts from supersonic transport exhaust. *Science*, **173**, 517-522, doi:10.1126/science.173.3996.517.

Jones, R. L., and J. A. Pyle, 1984: Observations of CH_4 and N_2O by the NIMBUS 7 SAMS: A comparison with *in situ* data and two-dimensional numerical model calculations. *J. Geophys. Res.*, **89**, 5263-5279, doi:10.1029/JD089iD04p05263.

Jones, R. L., *et al.*, 1986: The water vapour budget of the stratosphere studied using LIMS and SAMS satellite data. *Quarterly J. Royal Meteorol. Soc.*, **112**, doi:10.1256/smsqj.47411.

Jones, A., *et al.*, 2009: Evolution of stratospheric ozone and water vapor time series studied with satellite measurements. *Atmos. Chem. Phys.*, **9**, 6055-6075, doi:10.5194/acp-9-6055-2009.

Jones, D. B. A., *et al.*, 2009: The zonal structure of tropical O_3 and CO as observed by the Tropospheric Emission Spectrometer in November 2004. Part I. Inverse modeling of CO emissions. *Atmos. Chem. Phys.*, **9**, 3547-3562, doi:10.5194/acp-9-3547-2009.

Jones, A., *et al.*, 2011: A global inventory of stratospheric NO_y from ACE-FTS. *J. Geophys. Res.*, **116**, D17304, doi:10.1029/2010JD015465.

Jones, A., *et al.*, 2012: Technical Note: A trace gas climatology derived from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) dataset. *Atmos. Chem. Phys.*, **12**, 5207-5220, doi:10.5194/acp-12-5207-2012.

Jucks, K. W., *et al.,* 1998: Observations of OH, HO_2, H_2O, and O_3, in the upper stratosphere: Implications for HO_x photochemistry. *Geophys. Res. Lett.*, **25**, 3935-3938, doi:10.1029/1998GL900009.

Junge, C. E., C. W. Chagnon, and J. E. Manson, 1961: Stratospheric aerosols. *J. Meteorol.*, **18**, 81-108., http://dx.doi.org/10.1175/1520-0469(1961)018<0081:SA>2.0.CO;2

Kanzawa, H., *et al.*, 2003: Validation and data characteristics of nitrous oxide and methane profiles observed by the Improved Limb Atmospheric Spectrometer (ILAS) and processed with the Version 5.20 algorithm. *J. Geophys. Res.*, **108**, 8003, doi:10.1029/2002JD002458.

Kar, J., *et al.*, 2007: Initial comparison of ozone and NO_2 profiles from ACE-MAESTRO with balloon and satellite data. *J. Geophys. Res.*, **112**, D16301, doi:10.1029/2006JD008242.

Kasai, Y., *et al.*, 2013: Validation of stratospheric and mesospheric ozone observed by SMILES from International Space Station. *Atmos. Meas. Tech.*, **6**, 2311-2338, doi:10.5194/amt-6-2311-2013.

Kellmann, S., *et al.*, 2012: Global CFC-11 (CCl_3F) and CFC-12 (CCl_2F_2) measurements with the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS): Retrieval, climatologies and trends. *Atmos. Chem. Phys.*, **12**, 11857-11875, doi:10.5194/acp-12-11857-2012.

Kerzenmacher, T., *et al.*, 2008: Validation of NO_2 and NO from the Atmospheric Chemistry Experiment (ACE). *Atmos. Chem. Phys.*, **8**, 5801-5841, doi:10.5194/acp-8-5801-2008.

Khosravi, R., *et al.*, 2009: Overview and characterization of retrievals of temperature, pressure, and atmospheric constituents from the High Resolution Dynamics Limb Sounder (HIRDLS) measurements. *J. Geophys. Res.*, **114**, D20304, doi:10.1029/2009JD011937.

Khosravi, M., *et al.*, 2013: Diurnal variation of stratospheric and lower mesospheric HOCl, ClO and HO_2 at the equator: comparison of 1-D model calculations with measurements by satellite instruments. *Atmos. Chem. Phys.*, **13**, doi:10.5194/acp-13-7587-2013.

Kiefer, M., *et al.,* 2010: Impact of temperature field inhomogeneities on the retrieval of atmospheric species from MIPAS IR limb emission spectra. *Atmos. Meas. Tech.*, **3**, 1487-1507, doi:10.5194/amt-3-1487-2010.

Kikuchi, K., *et al.*, 2010: Overview and early results of the Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES). *J. Geophys. Res.*, **115**, D23306, doi:10.1029/2010JD014379.

Kinnison, D.E., *et al.*, 2007: Sensitivity of chemical tracers to meteorological parameters in the MOZART-3 chemical transport model. J. *Geophys. Res.*, **112**, D20302, doi:10.1029/2006JD007879.

Ko, M. K. W., *et al.*, 1993: Atmospheric sulfur hexafluoride: Sources, sinks and greenhouse warming. *J. Geophys. Res.*, **98**, 10499-10507, doi:10.1029/93JD00228.

Kohri, W.J., 1981: LRIR Observations of the Structure and Propagation of the Stationary Planetary Waves in the Northern Hemisphere during December 1975. Cooperative Thesis No. 63, Drexel University and National Center for Atmospheric Research.

Konopka, P., *et al.*, 2010: Annual cycle of ozone at and above the tropical tropopause: observations versus simulations with the Chemical Lagrangian Model of the Stratosphere (CLaMS). *Atmos. Chem. Phys.*, **10**, 121-132, doi:10.5194/acp-10-121-2010.

Kreyling, D., *et al.*, 2013: SMILES zonal and diurnal variation climatology of stratospheric and mesospheric trace gasses: O_3, HCl, HNO_3, ClO, BrO, HOCl, HO_2, and temperature. *J. Geophys. Res. Atmos.*, **118**, 11888-11903, doi:10.1002/2012JD019420.

Krüger, K., B. Naujokat, and K. Labitzke, 2005: The unusual midwinter warming in the Southern Hemisphere stratosphere of 2002: A comparison to Northern Hemisphere phenomena. *J. Atmos. Sci*., **62**, 603-613, doi:10.1175/JAS-3316.1.

Kulawik, S. S., *et al.*, 2006a: TES atmospheric profile retrieval characterization: An orbit of simulated observations. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1324– 1333, doi:10.1109/TGRS.2006.871207.

Kulawik, S. S., *et al.*, 2006b: Calculation of altitude-dependent Tikhonov constraints for TES nadir retrievals. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1334–1342, doi:10.1109/TGRS.2006.871206.

Kuribayashi, K., *et al.,* 2014: Direct estimation of the rate constant of the reaction ClO + HO_2 -> HOCl +O_2 from SMILES atmospheric observations. *Atmos. Chem. Phys.*, **14**, 255-266, doi:10.5194/acp-14-255-2014.

Kyrölä, E., *et al.*, 2006: Night-time ozone profiles in the stratosphere and mesosphere by the Global Ozone Monitoring by Occultation of Stars on Envisat. *J. Geophys. Res.*, **111**, D24306, doi:10.1029/2006JD007193.

Kyrölä, E., *et al.*, 2010a: GOMOS O_3, NO_2, and NO_3 observations in 2002-2008. *Atmos. Chem. Phys.*, **10**, 7723-7738, doi:10.5194/acp-10-7723-2010.

Kyrölä, E., *et al.*, 2010b: Retrieval of atmospheric parameters from GOMOS data. *Atmos. Chem. Phys.*, **10**, 11881-11903, doi:10.5194/acp-10-11881-2010.

Laaksonen, A., *et al.*, 2000: Upper tropospheric SO_2 conversion into sulfuric acid aerosols and cloud condensation nuclei. *J. Geophys. Res.*, **105**, 1459-1469, doi:10.1029/1999JD900933.

Labitzke, K., and M. P. McCormick, 1992: Stratospheric temperature increases due to Pinatubo aerosols. *Geophys. Res. Lett*., **19**, 207-210, doi:10.1029/91GL02940.

Laeng, A., *et al.*, 2014: Validation of MIPAS IMK/IAA V5R_O3_224 ozone profiles. *Atmos. Meas. Tech.*, **7**, 3971-3987, doi:10.5194/amt-7-3971-2014.

Lambert, A., *et al.*, 1993: Measurements of the evolution of the Mount Pinatubo aerosol clouds by ISAMS. *Geophys. Res. Lett*., **20**, 1287-1290, doi:10.1029/93GL00827.

Lambert, A., *et al.*, 2007: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements. *J. Geophys. Res.*, **112**(D24S36), doi:10.1029/2007JD008724.

Lary, D. J., *et al.*, 2007: Variations in stratospheric inorganic chlorine between 1991 and 2006. *Geophys. Res. Lett.*, **34**, L21811, doi:10.1029/2007GL030053.

Leblanc, T., *et al.*, 2011: Measurements of Humidity in the Atmosphere and Validation Experiments (MOHAVE)-2009: Review of campaign operations and results. *Atmos. Meas. Tech.*, **4**, 2579-2605, doi:10.5194/amt-4-2579-2011.

Ling, X.-D., and J. London, 1986: The quasi-biennial oscillation of ozone in the tropical middle stratosphere: A one-dimensional model. *J. Atmos. Sci*., **43**, 3122-3137, doi:10.1175/1520-0469(1986)043<3122:TQBOOO>2.0.CO;2.

Livesey, N. J., and W. G. Read, 2000: Direct retrieval of line-of-sight atmospheric structure from limb sounding observations. *Geophys. Res. Lett.*, **27**, 891-894, doi:10.1029/1999GL010964.

Livesey, N. J., et al., 2001: Stratospheric CH_3CN from the UARS Microwave Limb Sounder. *Geophys. Res. Letts.*, **28**, 779-782, doi:10.1029/2000GL012144.

Livesey, N. J., *et al.*, 2003: The UARS Microwave Limb Sounder version 5 dataset: Theory, characterization and validation. *J. Geophys. Res.,* **108**, 4378, doi:10.1029/2002JD002273.

Livesey, N. J., *et al.*, 2004: Enhancements in lower stratospheric CH_3CN observed by UARS MLS following boreal forest fires. *J. Geophys. Res.*, **109**, D06308, doi:10.1029/2003JD004055.

Livesey, N. J., *et al.*, 2006: Retrieval algorithms for the EOS Microwave Limb Sounder (MLS) instrument. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1144-1155, doi:10.1109/TGRS.2006.872327.

Livesey, N. J., *et al.*, 2008: Validation of Aura Microwave Limb Sounder O_3 and CO observations in the upper troposphere and lower stratosphere. *J. Geophys. Res.*, **113**, D15S02, doi:10.1029/2007JD008805.

Livesey, N. J., *et al.*, 2011: EOS MLS Version 3.3 Level 2 data quality and description document. Tech. rep., D-33509, Jet Propulsion Laboratory.

Livesey, N. J., *et al.,* 2013: EOS MLS Version 3.3/3.4 Level 2 data quality and description document. Tech. rep., Jet Propulsion Laboratory, available from **http://mls.jpl.nasa.gov/**.

Llewellyn, E., *et al.*, 2004: The OSIRIS Instrument on the Odin Spacecraft. *Can. J. Phys.*, **82**, 411-422, doi:10.1139/p04-005.

Logan, J.A., 1999: An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models, and development of a gridded climatology for tropospheric ozone. *J. Geophys. Res.*, **104**, 16115-16149, doi:10.1029/1998JD100096.

Lossow, S., *et al.*, 2009: Wintertime water vapor in the polar upper mesosphere and lower thermosphere – first satellite observations by Odin/SMR. *J. Geophys. Res.*, **114**, D10304, doi:10.1029/2008JD011462.

Lowe, D., and R. MacKenzie, 2008: Review: Polar stratospheric cloud microphysics and chemistry. *J. Atmos. Sol.-Terr. Phys.*, **70**, 13-40, http://dx.doi.org/10.1016/j.jastp.2007.09.011.

Lucke, R. L., *et al.*, 1999: The Polar Ozone and Aerosol Measurement (POAM III) Instrument and Early Validation Results. *J. Geophys. Res.*, **104**, 18785-18799, doi:10.1029/1999JD900235.

Lumpe, J. D., *et al.*, 1997: POAM II Retrieval Algorithm and Error Analysis. *J. Geophys. Res.*, **102**, 23593-23614, doi:10.1029/97JD00906.

Lumpe, J. D., *et al.*, 2002: POAM III retrieval algorithm and error analysis. *J. Geophys. Res.*, **107**, 4575, doi:10.1029/2002JD002137.

Lumpe, J., *et al.*, 2006: Validation of Polar Ozone and Aerosol Measurement (POAM) III version 4 stratospheric water vapour. *J. Geophys. Res.*, **111**, D11301, doi:10.1029/2005JD006763.

Mahieu, E., *et al.*, 2008: Validation of ACE-FTS v2.2 measurements of HCl, HF, CCl_3F and CCl_2F_2 using space-, balloon- and ground-based instrument observations. *Atmos. Chem. Phys.*, **8**, 6199-6221, doi:10.5194/acp-8-6199-2008.

Manabe, S., and R. T. Wetherald, 1967: Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.*, **24**, 241 pp., http://dx.doi.org/10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.

Manney, G. L., *et al.*, 2009: Satellite observations and modelling of transport during the 2006 major stratospheric sudden warming. *Atmos. Chem. Phys.*, **9**, 4775-4795.

Manney, G. L., *et al.*, 2011: Jet characterization in the upper troposphere/lower stratosphere (UTLS): Applications to climatology and transport studies. *Atmos. Chem. Phys.*, **11**, 6115-6137, doi:10.5194/acp-11-6115-2011.

McCormick, M. P., *et al.*, 1979: Satellite studies of the stratospheric aerosol. *Bull. Am. Meteor. Soc.*, **60**, 1038-1046, http://dx.doi.org/10.1175/1520-0477(1979)060<1038:SSOTSA>2.0.CO;2.

McCormick, M. P., *et al.*, 1981: High-latitude stratospheric aerosols measured by the SAM II satellite system in 1978 and 1979. *Science*, **214**, 328-331, doi:10.1126/science.214.4518.328.

McCormick, M. P., *et al.*, 1989: An overview of SAGE-I and II ozone measurements. *Planetary and Space Science*, **37**, 1567-86, doi:10.1016/0032-0633(89)90146-3.

McCormick, M. P., *et al.*, 1995: Atmospheric effects of the Mt. Pinatubo eruption. *Nature*, **373**, 399-404, doi:10.1038/373399a0.

McElroy, C. T., *et al.*, 2007: The ACE-MAESTRO instrument on SCISAT: description, performance, and preliminary results. *Appl. Opt.*, **46**, 4341-4356, doi:10.1364/AO.46.004341.

McHugh, M. J., *et al.*, 2003: Improved mesospheric temperature, water vapor and polar mesospheric cloud extinctions from HALOE. *Geophys. Res. Lett.*, **30**, 1440, doi:10.1029/2002GL016859.

McHugh, M., *et al.*, 2005: Comparison of atmospheric retrievals from ACE and HALOE. *Geophys. Res. Lett.*, **32**, L15S10, doi:10.1029/2005GL022403.

McLinden, C. A., *et al.*, 2000: Stratospheric ozone in 3-D models: A simple chemistry and the cross-tropopause flux. *J. Geophys. Res.*, **105**, 14653-14665, doi:10.1029/2000JD900124.

McLinden, C. A., *et al.*, 2001: Understanding trends in stratospheric NO_y and NO_2. *J. Geophys. Res.*, **106**, 27787-27793, doi:10.1029/2000JD000100.

McLinden, C. A., S. Tegtmeier, and V. Fioletov, 2009: Technical Note: A SAGE-corrected SBUV zonal-mean ozone data set. *Atmos. Chem. Phys.*, **9**, 7963-7972, doi:10.5194/acp-9-7963-2009.

McLinden, C. A., *et al.*, 2010: Odin/OSIRIS observations of stratospheric BrO: Retrieval methodology, climatology, and inferred Br_y. *J. Geophys. Res.*, **115**, D15308, doi:10.1029/2009JD012488.

Mengistu Tsidu, G., *et al.*, 2004: Stratospheric N_2O_5 in the austral spring 2002 as rerieved from limb emission spectra recorded by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS). *J. Geophys. Res.*, **109**, D18301, doi:10.1029/2004JD004856.

Mengistu Tsidu, G., *et al.*, 2005: NO_y from Michelson Interferometer for Passive Atmospheric Sounding on Environmental Satellite during the southern hemisphere polar vortex split in September/October 2002. *J. Geophys. Res.,* **110**, D11301, doi:10.1029/2004JD005322.

Mieruch, S., *et al.,* 2012: Global and long-term comparison of SCIAMACHY limb ozone profiles with correlative satellite data (2002-2008). *Atmos. Meas. Tech.*, **5**, 771-788, doi:10.5194/amt-5-771-2012.

Millán, L., 2012: New Aura Microwave Limb Sounder observations of BrO and implications for Bry. *Atmos. Meas. Tech.*, **5**, 1741–1751, doi:10.5194/amt-5-1741-2012.

Milz, M., *et al.*, 2005: Water vapor distributions measured with the Michelson Interferometer for Passive Atmospheric Sounding on board Envisat (MIPAS/Envisat). *J. Geophys. Res*., **110**, D24307, doi:10.1029/2005JD005973.

Milz, M., *et al.*, 2009: Validation of water vapour profiles (version 13) retrieved by the IMK/IAA scientific retrieval processor based on full resolution spectra measured by MIPAS on board Envisat. *Atmos. Meas. Tech*., **2**, 379-399, doi:10.5194/amt-2-379-2009.

Millán, L., *et al.*, 2015: Stratospheric and mesospheric HO_2 observations from the Aura Microwave Limb Sounder. *Atmos. Chem. Phys.*, **15**, 2889-2902, doi:10.5194/acp-15-2889-2015.

Molina, L. T. and M.J. Molina, 1987: Production of Cl_2O_2 from the self-reaction of the ClO radical. *J. Phys. Chem*. **91**, 433-436, doi: 10.1021/j100286a035.

Moré, J. J., 1977: The Levenberg-Marquardt algorithm: Implementation and theory. *Proc. Biennial Conf. Numerical Analysis*, pp105-116, doi:10.1007/BFb0067700.

Morris, R. A., *et al.*, 1995: Effects of electron and ion reactions on atmospheric lifetimes of fully fluorinated compounds. *J. Geophys. Res.,* **100**, 1287-1294, doi:10.1029/94JD02399.

Mote, P.W., *et al.*, 1996: An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapour. *J Geophys. Res.*, **101**, 3989-4006, doi:10.1029/95JD03422.

Murtagh, D., *et al.*, 2002: An overview of the Odin atmospheric mission. *Can. J. Phys.*, **80**, 309-319, doi:10.1139/p01-157.

Nardi, B., *et al.*, 2008: Initial validation of ozone measurements from the High Resolution Dynamics Limb Sounder. *J. Geophys. Res.*, **113**, D16S36, doi:10.1029/2007JD008837.

Nassar, R., *et al.*, 2008: Validation of Tropospheric Emission Spectrometer (TES) nadir ozone profiles using ozonesonde measurements. *J. Geophys. Res.*, **113**, D15S17, doi:10.1029/2007JD008819.

Nedoluha, G. E., *et al.*, 2007: A comparison of middle atmospheric water vapour as measured by WVMS, E_S-MLS, and HALOE. *J. Geophys. Res.*, **112**, D24S39, doi:10.1029/2007JD008757.

Nedoluha, G. E., *et al.*, 2009: Water vapour measurements in the mesosphere from Mauna Loa over solar cycle 23. *J. Geophys. Res.*, **114**, D23303, doi:10.1029/2009JD012504.

Nedoluha, G. E., *et al.*, 2011: Ground-based measurements of ClO from Mauna Kea and intercomparisons with Aura and Uars MLS. *J. Geophys. Res.*, **116**, D02307, doi:10.1029/2010JD014732.

Neu, J. L., *et al.*, 2014a: The SPARC Data Initiative: Comparison of upper troposphere / lower stratosphere ozone climatologies from limb-viewing instruments and the nadir-viewing Tropospheric Emission Spectrometer (TES). *J. Geophys. Res.*, **119**, 6971-6990, doi:10.1002/2013JD020822.

Neu, J. L., *et al.*, 2014b: Tropospheric ozone variations governed by changes in stratospheric circulation. *Nature Geosci.*, **7**, 340-344, doi:10.1038/ngeo2138.

Newman, P. A., *et al.*, 2006: When will the Antarctic ozone hole recover? *Geophys. Res. Lett*., **33**, L12814, doi:10.1029/2005GL025232.

Niwano, M., *et al.*, 2003: Seasonal and QBO variations of ascent rate in the tropical lower stratosphere as inferred from UARS HALOE trace gas data, *J. Geophys. Res.*, **108**, 4794, doi:10.1029/2003JD003871.

Olberg, M., *et al.*, 2003: The Odin satellite: II. Radiometer data processing and calibration. *Astron. Astrophys.,* **402**, L35, http://dx.doi.org/10.1051/0004-6361:20030336.

Orsolini, Y., *et al.*, 2010: Descent from the polar mesosphere and anomalously high stratopause observed in 8 years of water vapor and temperature satellite observations by the Odin Sub-Millimeter Radiometer. *J. Geophys. Res.*, **115**, doi:10.1029/2009JD013501.

Pardo, J. R., *et al.*, 2001: Submillimeter atmospheric transmission measurements on Mauna Kea during extremely dry El Niño conditions: Implications for broadband opacity contributions. *J. Quant. Spectrosc. Rad. Transf.*, **68**, 419-433, doi:10.1016/S0022-4073(00)00034-0.

Park, J. H., *et al.*, 1996: Validation of halogen occultation experiment CH_4 measurements from the UARS. *J. Geophys. Res.*, **101**, 10183-10204, doi:10.1029/95JD02736.

Park, M., *et al.*, 2004: Seasonal variations of methane, water vapor, ozone, and nitrogen dioxide near the tropopause: Satellite observations and model simulations. *J. Geophys. Res.*, **109**, D03302, doi:10.1029/2003JD003706.

Park, M., *et al.*, 2007: Transport above the Asian summer monsoon anticyclone inferred from Aura MLS tracers. *J. Geophys. Res.*, **112**, D16309, doi:10.1029/2006JD008294.

Peevey, T. R., *et al.*, 2012: Investigation of double tropopause spatial and temporal global variability utilizing High Resolution Dynamics Limb Sounder temperature observations. *J. Geophys. Res.*, **117**, D01105, doi:10.1029/2011JD016443.

Peter, T., 1997: Microphysics and heterogeneous chemistry of polar stratospheric clouds. *Annu. Rev. Phys. Chem.*, **48**, 785-822, doi:10.1146/annurev.physchem.48.1.785.

Pickett, H. M., *et al.*, 2006: Validation of Aura MLS HO_x measurements with remote-sensing balloon instruments. *Geophys. Res. Lett.*, **33**, doi:10.1029/2005GL024048.

Pickett, H. M., *et al.*, 2008: Validation of Aura Microwave limb sounder OH and HO_2 measurements. *J. Geophys. Res.*, **113**, doi:10.1029/2007JD008775.

Ploeger, F., *et al.*, 2012: Horizontal transport affecting trace gas seasonality in the Tropical Tropopause Layer (TTL). *J. Geophys. Res.*, **117**, 1-16, doi:10.1029/2011JD017267.

Plummer, D. A., *et al.*, 2010: Quantifying the contributions to stratospheric ozone changes from ozone depleting substances and greenhouse gases. *Atmos. Chem. Phys.*, **10**, 8803-8820, doi:10.5194/acp-10-8803-2010.

Prather, M. J., 1992: Catastrophic loss of stratospheric ozone in dense volcanic clouds. *J. Geophys. Res.*, **97**, 10187-10191, doi:10.1029/92JD00845.

Pumphrey, H. C., 1999: Validation of a new prototype water vapor retrieval for the UARS Microwave Limb Sounder. *J. Geophys. Res.*, **104**, 9399-9412, doi:10.1029/1998JD200113.

Pumphrey, H. C., *et al.*, 2000: Lower stratospheric water vapor measured by UARS MLS. *Geophys. Res. Lett.*, **27**, 1691-1694, doi:10.1029/1999GL011339.

Pumphrey, H. C., *et al.*, 2007: Validation of middle-atmosphere carbon monoxide retrievals from the Microwave Limb Sounder on Aura. *J. Geophys. Res.*, **112**, doi:10.1029/2007JD008723.

Pumphrey, H. C., *et al.*, 2011: Microwave Limb Sounder observations of biomass-burning products from the Australian bush fires of February 2009. *Atmos. Chem. Phys.*, **11**, 6285-6296, doi:10.5194/acp-11-6285-2011.

Pumphrey, H. C., *et al.*, 2013: EOS MLS Version 3.3/3.4 Level 2 data quality and description document. Tech. Rep., Jet Propulsion Laboratory, available from **http://mls.jpl.nasa.gov/**.

Randall, C. E., *et al.*, 1998: POAM II measurements of stratospheric NO_2, 1993-1996. *J. Geophys. Res.*,**103**, 28361-28371, doi:10.1029/98JD02092.

Randall, C. E., *et al.*, 2000: Comparison of Polar Ozone and Aerosol measurement (POAM) II and Stratospheric Aerosol and Gas Experiment (SAGE) II aerosol extinction measurements from 1994 to 1996. *J. Geophys. Res.*, **105**, 3929-3942, doi: 10.1029/1999JD901024.

Randall, C. E., *et al.*, 2001: Validation of POAM III Aerosols: Comparison to SAGE II and HALOE. *J. Geophys. Res.*, **106**, 27525-27536, doi:10.1029/2001JD000528.

Randall, C. E., *et al.*, 2002: Validation of POAM III NO_2 measurements. *J. Geophys. Res.*, **107**, 4432, doi:10.1029/2001JD001520.

Randall, C. E., *et al.*, 2003: Validation of POAM III ozone: Comparisons with ozonesonde and satellite data. *J. Geophys. Res.*, **108**, 4367, doi:10.1029/2002JD002944.

Randel, W. J., 1990: Kelvin wave induced trace constituent oscillations in the equatorial stratosphere. *J. Geophys. Res.*, **95**, 18641-18652, doi:10.1029/JD095iD11p18641.

Randel, W. J., 1993: Global variations of zonal mean ozone during stratospheric warming events. *J. Atmos. Sci.*, **50**, 3308-3321.

Randel, W. J., *et al.*, 1998: Seasonal cycles and QBO variations in stratospheric CH_4 and H_2O observed in UARS HALOE data. *J. Atmos. Sci.,* **55**, 163-185, doi:10.1175/1520-0469(1998)055<0163:SCAQVI>2.0.CO;2.

Randel, W. J., *et al.*, 2004: Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures. *J. Atmos. Sci.*, **61**, 2133-2148, http://dx.doi.org/10.1175/1520-0469(2004)061<2133:ICOSWV>2.0.CO;2.

Randel, W. J., *et al.*, 2006: Decreases in stratospheric water vapor since 2001: Links to changes in the tropical tropopause and the Brewer-Dobson circulation. *J. Geophys. Res.*, **111**, D12312, doi:10.1029/2005JD006744.

Randel,W. J., *et al.*, 2007: A large annual cycle in ozone above the tropical tropopause linked to the Brewer-Dobson circulation. *J. Atmos. Sci.*, **64**, 4479-4488, doi:10.1175/2007JAS2409.1.

Ravishankara, A. R., *et al.*, 1993: Atmospheric lifetimes of long-lived halogenated species. *Science*, **259**, 194-199, doi:10.1126/science.259.5092.194.

Ravishankara, A. R., *et al.*, 2009: Nitrous Oxide (N_2O): The dominant ozone-depleting substance emitted in the 21st century. *Science*, **326**, 123-125, doi:10.1126/science.1176985.

Ray, E. A., *et al.*, 1994: The tropical semiannual oscillation in temperature and ozone as observed by MLS. *J. Atmos. Sci.*, **51**, 3045-3052, doi:10.1175/1520-0469(1994)051<3045:TTSOIT>2.0.CO;2.

Read, W. G., *et al.*, 2006: The clear-sky unpolarized forward model for the EOS Microwave Limb Sounder (MLS). *IEEE Trans. Geosci. Remote Sens.*, **44**, 1367-1379.

Read, W. G., *et al.*, 2007: Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H_2O and relative humidity with respect to ice validation. *J. Geophys. Res.*, **112**, D24S35, doi:10.1029/2007JD008752.

Reber, C. A., *et al.*, 1993: The Upper Atmosphere Research Satellite (UARS) mission. *J. Geophys. Res.*, **98**, 10643-10647, doi:10.1029/92JD02828.

Reddmann, T., *et al.*, 2001: Three-dimensional model simulations of SF_6 with mesospheric chemistry. *J. Geophys. Res.*, **106**, 14525-14537, doi:10.1029/2000JD900700.

Reddmann, T., *et al.*, 2010: Modeling disturbed stratospheric chemistry during solar-induced NO_x enhancements observed with MIPAS/ENVISAT. *J. Geophys. Res.*, **115**, D00I11, doi:10.1029/2009JD012569.

Remsberg E. E., et. al., 1990: Estimation of synopic fields of middle atmosphere parameters from Nimbus-7 LIMS profile data. *J. Atmos. Oceanic Tech.*, **7**, 689-705, doi:10.1175/1520-0426(1990)007<0689:EOSFOM>2.0.CO;2.

Remsberg, E. E., *et al.*, 2004: The Nimbus 7 LIMS Version 6 radiance conditioning and temperature retrieval methods and results. *J. Quant. Spectrosc. Rad. Transf.*, **86**, 395-424, doi:10.1016/j.jqsrt.2003.12.007.

Remsberg, E. E., *et al.*, 2007: On the quality of the Nimbus 7 LIMS version 6 ozone for studies of the middle atmosphere. *J. Quant. Spectrosc. Rad. Transf.*, **105**, 492-518, doi:10.1016/j.jqsrt.2006.12.005.

Remsberg, E. E., *et al.*, 2009: On the quality of the Nimbus 7 LIMS Version 6 water vapor profiles and distributions. *Atmos. Chem. Phys.*, **9**, 9155-9167, doi:10.5194/acp-9-9155-2009.

Remsberg, E. E., *et al.*, 2010: Improvements in the profiles and distributions of nitric acid and nitrogen dioxide with the LIMS version 6 dataset. *Atmos. Chem. Phys.*, **10**, 4741-4756, doi:10.5194/acp-10-4741-2010.

Remsberg, E. E., 2015: Methane as a diagnostic tracer of changes in the Brewer–Dobson circulation of the stratosphere. *Atmos. Chem. Phys.*, **15**, 3739-3754, doi:10.5194/acp-15-3739-2015.

Ricaud, P., *et al.*, 2007: Measurements of mid-stratospheric formaldehyde from the Odin/SMR instrument. *J. Quant. Spectroscop. Radiat. Transfer*, **107**, 91-104, http://dx.doi.org/10.1016/j.jqsrt.2007.01.058.

Riese, M., *et al.*, 1999: Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) data processing and atmospheric temperature and trace gas retrieval. *J. Geophys. Res.*, **104**, 16349-16368, doi:10.1029/1998JD100057.

Roche, A. E., *et al.*, 1993: The Cryogenic Limb Array Etalon Spectrometer (CLAES) on UARS: Experiment description and performance. *J. Geophys. Res.,* **98**, 10763-10775, doi:10.1029/93JD00800.

Rodgers, C. D., 1976: Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys.*, **14**, 609-624, doi:10.1029/RG014i004p00609.

Rodgers, C., 2000: Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific, London.

Rodgers, C. D., and B. J. Connor, 2003: Intercomparison of remote sounding instruments. *J. Geophys. Res.*, **108**, 4116, doi:10.1029/2002JD002299.

Rothman, L. S., *et al.*, 2005: The HITRAN 2004 molecular spectroscopic database. *J. Quant. Spectrosc. Rad. Transf.*, **96**, 139-204, doi:10.1016/j.jqsrt.2004.10.008.

Rousseeuw, P. J., and Croux, C., 1993: Alternatives to the median absolute deviation. *J. Amer. Statist. Assoc.,* **88**, 1273-1283, http://dx.doi.org/10.1080/01621459.1993.10476408.

Rozanov, A., *et al.*, 2005: NO_2 and BrO vertical profile retrieval from SCIAMACHY limb measurements: Sensitivity studies. *Advances in Space Research*, **36**, 846-854, doi:10.1016/j.asr.2005.03.013.

Rozanov, A., *et al.*, 2011a: BrO vertical distributions from SCIAMACHY limb measurements: comparison of algorithms and retrieval results. *Atmos. Meas. Tech.*, **4**, 1319-1359, doi:10.5194/amt-4-1319-2011.

Rozanov, A., *et al.*, 2011b: Retrieval of water vapor vertical distributions in the upper troposphere and the lower stratosphere from SCIAMACHY limb measurements. *Atmos. Meas. Tech.*, **4**, 933-954, doi:10.5194/amt-4-933-2011.

Rusch, D. W., *et al.*, 1994: Solar Mesosphere Explorer satellite measurements of El Chichon stratospheric aerosols 1. Cloud morphology. *J. Geophys. Res.*, **99**, 20525-20532, doi:10.1029/94JD01842.

Rusch, D.W., *et al.*, 1997: Validation of POAM II Ozone Measurements with Coincident MLS, HALOE, and SAGE II Observations. *J. Geophys. Res.*, **102**, 23615-23627, doi:10.1029/97JD00458.

Russell III, J. M., *et al.*, 1984: The variability of stratospheric and mesospheric NO_2 in the polar winter night observed by LIMS. *J. Geophys. Res.*, **89**, 7267-7275, doi:10.1029/JD089iD05p07267.

Russell III, J. M., *et al.*, 1993: The Halogen Occultation Experiment. *J. Geophys. Res.*, **98**, 10777-10797, doi:10.1029/93JD00799.

Russell III, J. M., *et al.*, 1996a: Validation of hydrogen fluoride measurements made by the Halogen Occultation Experiment from the UARS platform. *J. Geophys. Res.*, **101**, 10163-10174, doi:10.1029/95JD01705..

Russell III, J. M., *et al.*, 1996b: Validation of hydrogen chloride measurements made by the Halogen Occultation Experiment from the UARS platform. *J. Geophys. Res.*, **101**, 10151-10162, doi:10.1029/95JD01696..

Ruth, S., *et al.*, 1997: Seasonal, semiannual, and interannual variability seen in measurements of methane made by the UARS Halogen Occultation Experiment. *J. Geophys. Res.,* **102**, 16189-16199, doi:10.1029/97JD00868.

Sagawa, H., *et al.*, 2013: Comparison of SMILES ClO profiles with satellite, balloon-borne and ground-based measurements. *Atmos. Meas. Tech.*, **6**, 3325-3347, doi:10.5194/amt-6-3325-2013.

SAGE-III, 2002: SAGE III Algorithm Theoretical Base Document, Solar and Lunar Algorithm, Earth Observing System Project science Office web site, **http://eospso.gsfc.nasa.gov**.

Salby, M. L., *et al.*, 1990: Chemical fluctuations associated with vertically propagating equatorial Kelvin waves. *J. Geophys. Res.*, **95**, 20491-20505, doi:10.1029/90JD01371.

Santee, M. L., *et al.*, 1998: UARS Microwave Limb Sounder HNO_3 observations: Implications for Antarctic polar stratospheric clouds. *J. Geophys. Res.*, **103**(D11), 13285-13313, doi:10.1029/98JD00365.

Santee, M. L., *et al.*, 2003: Variations and climatology of ClO in the polar lower stratosphere from UARS Microwave Limb Sounder measurements. *J. Geophys. Res.*, **108**, D15, 4454 pp., doi:10.1029/2002JD003335.

Santee, M. L., *et al.*, 2007: Validation of Aura Microwave Limb Sounder HNO_3 Measurements. *J. Geophys. Res.*, **112**, D24S40, doi:10.1029/2007JD008721.

Santee, M.L., *et al.*, 2008: A study of stratospheric chlorine partitioning based on new satellite measurements and modeling. *J. Geophys. Res.*, **113**, D12307, doi:10.1029/2007JD009057.

Sasano, Y., *et al.*, 1999: Improved Limb Atmospheric Spectrometer (ILAS) for stratospheric ozone layer measurements by solar occultation technique. *Geophys. Res. Lett.*, **26**, 197-200, doi:10.1029/1998GL900276.

Sasano, Y., 2002: Preface. *J. Geophys. Res.*, **107**, 8204, doi:10.1029/2002JD002155.

Sato, T. O., *et al.*, 2012: Strato-mesospheric ClO observations by SMILES: Error analysis and diurnal variation. *Atmos. Meas. Tech.*, **5**, 2809-2825, doi:10.5194/amt-5-2809-2012.

Schneider, N., *et al.*, 2005: Seasonal and diurnal ozone variations: observations and modeling. *J. Atmos. Chem.*, **50**, 25-47, doi:10.1007/s10874-005-1172-z.

Schwartz, M. J., *et al.*, 2006: EOS MLS forward model polarized radiative transfer for Zeeman-split oxygen lines. *IEEE Trans. Geosci. Remote Sens.,* **44**, 1182-1191, doi:10.1109/TGRS.2005.862267.

Seinfeld, J. H., and S. N. Pandis, 2006: Atmospheric chemistry and physics: From air pollution to climate change. 2nd ed., John Wiley&Sons, New York, NY.

Seppälä, A., *et al.*, 2007: Arctic and Antarctic polar winter NO_x and energetic particle precipitation in 2002-2006. *Geophys. Res. Lett.*, **34**, L12810, doi:10.1029/2007GL029733.

Sheese, *et al.*, 2013: Odin observations of Antarctic nighttime NO densities in the mesosphere-lower thermosphere and observations of a lower NO layer, *J. Geophys. Res.*, **118**, 7414–7425, doi:10.1002/jgrd.50563.

Siddaway, J. M., and S. V. Petelina, 2011: Transport and evolution of the 2009 Australian Black Saturday bushfire smoke in the lower stratosphere observed by OSIRIS on Odin. *J. Geophys. Res.*, **116,** D06203, doi:10.1029/2010JD015162.

Sinnhuber, B.-M., *et al.*, 2009: The contribution of anthropogenic bromine emissions to past stratospheric ozone trends: a modelling study. *Atmos. Chem. Phys.*, **9**, 2863–2871, doi:10.5194/acp-9-2863-2009.

Siskind, D. E., *et al.*, 2013: Comparison of a photochemical model with observations of mesospheric hydroxyl and ozone. *J. Geophys. Res. Atmos.*, **118**, 195-207, doi:10.1029/2012JD017971.

Sofieva, V. F., *et al.*, 2004: Ozone profile smoothness as a priori information in the inversion of limb measurements. *Ann. Geophys.*, **22**, 3411-3420, doi:10.5194/angeo-22-3411-2004.

Sofieva, V., *et al.*, 2009: Influence of scintillation on quality of ozone monitoring by GOMOS. *Atmos. Chem. Phys.*, **9,** 9197-9207, doi:10.5194/acp-9-9197-2009.

Sofieva, V. F., *et al.*, 2010: Retrievals from GOMOS stellar occultation measurements using characterization of modeling errors. *Atmos. Meas. Tech.*, **3**, 1019-1027, doi:10.5194/amt-3-1019-2010.

Sofieva, V. F., *et al.*, 2014: Validation of GOMOS ozone precision estimates in the stratosphere. *Atmos. Meas. Tech.*, **7**, 2147-2158, doi:10.5194/amt-7-2147-2014.

Solomon, S., *et al.*, 1982: Photochemical coupling between the thermosphere and the lower atmosphere, I. Odd nitrogen from 50 to 120 km. *J. Geophys. Res.*, **87**, 7206-7220, doi:10.1029/JC087iC09p07206.

Solomon, S., *et al.*, 1986: On the depletion of Antarctic ozone. *Nature*, **321**, 755-758, doi:10.1038/321755a0.

Solomon, S., *et al.*, 1996: The role of aerosol variability in anthropogenic ozone depletion at northern midlatitudes. *J. Geophys. Res.*, **101**, 6713-6727, doi:10.1029/95JD03353.

Solomon, S., 1999: Stratospheric ozone depletion: A review of concepts and history. *Rev. Geophys.*, **37**, 275-316, doi:10.1029/1999RG900008.

Solomon, S., *et al.* (eds), 2007: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge Univ. Press, Cambridge.

Sonkaew, T., *et al.*, 2009: Cloud sensitivity studies for stratospheric and lower mesospheric ozone profile retrievals from measurements of limb-scattered solar radiation. *Atmos. Meas. Tech.*, **2**, 653-678, doi:10.5194/amt-2-653-2009.

SPARC, 2000: Upper Tropospheric and Stratospheric Water Vapour (WAVAS). D. Kley, J. M. Russell III, and C. Phillips (eds.), SPARC Report No. 2, WMO/TD-No. 1043, **www.sparc-climate.org/publications/sparc-reports/sparc-report-no2**.

SPARC, 2006: Assessment of stratospheric aerosol properties (ASAP). L. Thomason and T. Peter (eds.), SPARC Report No. 4, WCRP-124, WMO/TD-No. 1295., **www.sparc-climate.org/publications/sparc-reports/sparc-report-no4**.

SPARC, 2010: SPARC Report on the Evaluation of Chemistry-Climate Models. V. Eyring, T. G. Shepherd, D. W. Waugh (eds.), SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526, **www.sparc-climate.org/publications/sparc-reports/sparc-report-no5**.

Steck, T., and T. von Clarmann, 2001: Constrained profile retrieval applied to the observation mode of the Michelson Interferometer for Passive Atmospheric Sounding. *Appl. Opt.,* **40**, 3559-3571, https://doi.org/10.1364/AO.40.003559.

Steck, T., *et al.*, 2007: Bias determination and precision validation of ozone profiles from MIPAS-Envisat retrieved with the IMK-IAA processor. *Atmos. Chem. Phys.*, **7**, 3639-3662, doi:10.5194/acp-7-3639-2007.

Steck, T., *et al.*, 2008: Retrieval of global upper tropospheric and stratospheric formaldehyde (H_2CO) distributions from high-resolution MIPAS-Envisat spectra. *Atmos. Chem. Phys.*, **8**, 463-470, doi:10.5194/acp-8-463-2008.

Stephens, G. L., *et al.*, 2002: The CloudSat mission and the A-train: A new dimension of space-based observations of clouds and precipitation. *Bull. Amer. Meteor. Soc.*, **83**, 1771-1790, http://dx.doi.org/10.1175/BAMS-83-12-1771.

Stiller, G. P., *et al.*, 2002: Sensitivity of trace gas abundances retrievals from infrared limb emission spectra to simplifying approximations in radiative transfer modelling. *J. Quant. Spectrosc. Radiat. Transfer*, **72**, 249-280, doi:10.1016/S0022-4073(01)00123-6.

Stiller, G. P., *et al.*, 2007: Global distributions of HO_2NO_2 as observed by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS). J. *Geophys. Res.*, **112**, D09314, doi:10.1029/2006JD007212.

Stiller, G. P., *et al.*, 2008: Global distribution of mean age of stratospheric air from MIPAS SF_6 measurements. *Atmos. Chem. Phys.*, **8**, 677-695, doi:10.5194/acp-8-677-2008.

Stiller G. P., *et al.*, 2012: Validation of MIPAS IMK/IAA temperature, water vapour, and ozone profiles with MOHAVE-2009 campaign measurements. *Atmos. Meas. Tech.*, **5**, 289-320, doi:10.5194/amt-5-289-2012.

Stolarski, R. S., *et al.*, 2014: Seasonal variation of ozone in the tropical lower stratosphere: Southern tropics are different from northern tropics. *J. Geophys. Res. Atmos.*, **119**, 6196-6206, doi:10.1002/2013JD021294.

Strong, K., *et al.*, 2008: Validation of ACE-FTS N2O measurements. *Atmos. Chem. Phys.*, **8**, 4759-4786, doi:10.5194/acp-8-4759-2008.

Taha, G., *et al.*, 2004: Comparison of Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 water vapor with ballon-born and space-based instruments. *J. Geophys. Res.*, **109**, D18313, doi:10/1029/2004/2004JD004859.

Tamminen, J., *et al.*, 2010: GOMOS data characterisation and error estimation. *Atmos. Chem. Phys.*, **10**, 9505-9519, doi:10.5194/acp-10-9505-2010.

Taylor, F. W., 1987: Infrared remote sensing of the middle atmosphere from satellites: The stratospheric and mesospheric sounder experiment 1978-1983. *Surveys in Geophysics*, **9**, 123-148, doi:10.1007/BF01904119.

Taylor, F. W., *et al.*, 1993: Remote sensing of atmospheric structure and composition by pressure modulator radiometry from space: The ISAMS experiment on UARS. *J. Geophys. Res.*, **98**, 10799-10814, doi:10.1029/92JD03029.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183-7192, doi:10.1029/2000JD900719.

Tegtmeier, S., *et al.*, 2013: SPARC Data Initiative: A comparison of ozone climatologies from international satellite limb sounders. *J. Geophys. Res. Atmos.*, **118**, 12229-12247, doi:10.1002/2013JD019877.

Tikhonov, A. N., 1963: О решении некорректно поставленных задач и методе регуляризации. *Doklady Akademii Nauk SSSR*, **151**, 501-504. Translated in "Solution of incorrectly formulated problems and the regularization method". *Soviet Mathematics*, **4**, 1035-1038.

Thomason, L. W., and G. Taha, 2003: SAGE III aerosol extinction measurements: Initial results. *Geophys. Res. Lett.*, **30**, 1631, doi:10.1029/ 2003GL017317.

Thomason, L. W., *et al.*, 2004: A revised water vapor product for the Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 data set. *J. Geophys. Res.*, **109**, D06312, doi:10.1029/2003JD004465.

Thomason, L. W., *et al.*, 2010: An evaluation of the SAGE III version 4 aerosol extinction coefficient and water vapour data products. *Atmos. Chem. Phys.*, **10**, 2159-2173, doi:10.5194/acp-10-2159-2010.

Thompson, A. M., *et al.*, 2003: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements. *J. Geophys. Res.*, **108**, 8238, doi:10.1029/2001JD000967.

Toohey, M., *et al.*, 2010: Validating the reported random errors of ACE-FTS measurements. *J. Geophys. Res.*, **115**, D20304, doi:10.1029/2010JD014185.

Toohey, M., and von Clarmann,T., 2013: Climatologies from satellite measurements: The impact of orbital sampling on the standard error of the mean. *Atmos. Meas. Tech.*, **6**, 937-948, doi:10/5194/amt-6-937-2013.

Toohey, M., *et al.*, 2013: Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative. *J. Geophys. Res.*, **118**, 11847-11862, doi:10.1002/jgrd.50874.

Toon, O. B., 1986: Condensation of HNO_3 and HCl in winter polar stratospheres. *Geophys. Res. Lett.*, **13**, 1284-1287, doi:10.1029/GL013i012p01284.

Toon, G. C., 1991: The JPL MkIV Interferometer. *Opt. Photonics News*, **2**, 19-21, https://doi.org/10.1364/OPN.2.10.000019.

Twomey, S., 1975: Comparison of constrained linear iversion and an iterative non-linear algorithm applied to the indirect estimation of particle size distribution. *J. Comput. Phys.*, **18**, 188-198, doi:10.1016/0021-9991(75)90028-5.

Urban, J., *et al.*, 2004: MOLIERE (v5): A versatile forward- and inversion model for the millimeter and sub-millimeter wavelength range. *J. Quant. Spectrosc. Rad. Transf.*, **83**, 529-554, http://dx.doi.org/10.1016/S0022-4073(03)00104-3.

Urban, J., *et al.*, 2005a: Odin/SMR limb observations of stratospheric trace gases: Level 2 Processing of ClO, N_2O, O_3, and HNO_3. *J. Geophys. Res.*, **110**, D14307, doi:10.1029/2004JD005741.

Urban, J., *et al.*, 2005b: Odin/SMR limb observations of stratospheric trace gases: Validation of N_2O. *J. Geophys. Res.*, **110**, D09301, doi:10.1029/2004JD005394.

Urban, J., *et al.*, 2006: Odin/SMR Limb Observations of Trace Gases in the Polar Lower Stratosphere during 2004-2005. In: Proc. ESA First Atmospheric Science Conference, 8-12 May 2006, Frascati, Italy. Lacoste, H. (ed.), ESA-SP-628 Noordwijk: European Space Agency. ISBN-92-9092-939-1.

Urban, J., *et al.*, 2007: Global observations of middle atmospheric water vapour by the Odin satellite: An overview. *Planet. Space Sci.*, **55**, 1093-1102, http://dx.doi.org/10.1016/j.pss.2006.11.021.

Urban, J., 2008: Tropical ascent of lower stratospheric air analysed using measurements of the Odin Sub-Millimetre Radiometer. Proc. Reunion Island Int. Symp. Tropical Stratosphere – Upper Troposphere, 5-9 November 2007, St. Gilles, Reunion Island, France. Bencherif, H. (ed.), publisher: Université de la Réunion, pp29-34.

Urban, J., *et al.*, 2009: Nitric acid in the stratosphere based on Odin observations from 2001 to 2009 – Part 1: A global climatology. *Atmos. Chem. Phys.*, **9**, 7031-7044, doi:10.5194/acp-9-7031-2009.

Urban, J., *et al.*, 2012: Evolution and variability of water vapour in the tropical tropopause and lower stratosphere region derived from satellite measurements. In: Proc. ATMOS 2012, Advances in Atmospheric Science and Applications, ESA SP-708 (CD-ROM), ESA Communications, European Space Agency, Noordwijk, The Netherlands.

Vanhellemont, F., *et al.*, 2010: Optical extinction by upper tropospheric/stratospheric aerosols and clouds: GOMOS observations for the period 2002-2008. *Atmos. Chem. Phys.*, **10**, 7997-8009, doi:10.5194/acp-10-7997-2010.

Vanhellemont, F., *et al.*, 2016: AerGOM, an improved algorithm for stratospheric aerosol extinction retrieval from GOMOS observations – Part 1: Algorithm description. *Atmos. Meas. Tech.*, **9**, 4687-4700, doi:10.5194/amt-9-4687-2016.

Vernier, J.-P., *et al.*, 2011: Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade. *Geophys. Res. Lett.*, **38**, L12807, doi:10.1029/2011GL047563.

Verronen, P. T., *et al.*, 2005: A comparison of night-time GOMOS and MIPAS ozone profiles in the stratosphere and mesosphere. *Adv. Space Res.,* **36**, 958-966, doi:10.1016/j.asr.2005.04.073.

Verstraeten, W. W., *et al.*, 2013: Validation of six years of TES tropospheric ozone retrievals with ozonesonde measurements: implications for spatial patterns and temporal stability in the bias. *Atmos. Meas. Tech.*, **6**, 1413-1423, doi:10.5194/amt-6-1413-2013.

Voemel, H., *et al.*, 2007: Validation of Aura Microwave Limb Sounder water vapour by balloon-borne Cryogenic Frost Point Hygrometer measurements. *J. Geophys. Res.*, **112**, D24S37, doi:10.1029/2007JD008698.

Vogel, B., *et al.*, 2008: Model simulations of stratospheric ozone loss caused by enhanced mesospheric NO_x during Arctic Winter 2003/2004. *Atmos. Chem. Phys.*, **8**, 5279-5293, doi:10.5194/acp-8-5279-2008.

von Clarmann, T., *et al.*, 2002: Intercomparison of radiative transfer codes under nonlocal thermodynamic equilibrium conditions. *J. Geophys. Res.*, **107**, 4631, doi:10.1029/2001JD001551.

von Clarmann, T., *et al.*, 2003a: A blind test retrieval experiment for infrared limb emission spectrometry. *J. Geophys. Res.*, **108**, 4746, doi:10.1029/2003JD003835.

von Clarmann, T., *et al.*, 2003b: Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS). *J. Geophys. Res.*, **108**, 4736, doi:10.1029/2003JD003602.

von Clarmann, T., *et al.*, 2003c: Modelling of atmospheric mid-infrared radiative transfer: the AMIL2DA algorithm intercomparison experiment. *J. Quant. Spectrosc. Radiat. Transfer*, **78**, 381-407, doi:10.1016/S0022-4073(02)00262-5.

von Clarmann, T., *et al.*, 2005: Experimantal evidence of perturbed odd hydrogen and chlorine chemistry after the October 2003 solar proton events. *J. Geophys. Res.*, **110**, A09S45, doi:10.1029/2005JA011053.

von Clarmann, T., *et al.*, 2006: Global stratospheric HOCl distributions retrieved from infrared limb emission spectra recorded by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS). *J. Geophys. Res.*, **111**, D05311, doi:10.1029/2005JD005939.

von Clarmann, T., *et al.*, 2009a: Retrieval of temperature, H_2O, O_3, HNO_3, CH_4, N_2O, $ClONO_2$ and ClO from MIPAS reduced resolution nominal mode limb emission measurements. *Atmos. Meas. Tech.*, **2**, 159-175, doi:10.5194/amt-2-159-2009.

von Clarmann, T., *et al.*, 2009b: HOCl chemistry in the Antarctic Stratospheric Vortex 2002, as observed with the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS). *Atmos. Chem. Phys.*, **9**, 1817-1829, doi:10.5194/acp-9-1817-2009.

von Clarmann, T., *et al.*, 2010: Technical Note: Trend estimation from irregularly sampled, correlated data. *Atmos. Chem. Phys.*, **10**, 6737-6747, doi:10.5194/acp-10-6737-2010.

von Clarmann, T., *et al.*, 2012: The MIPAS HOCl climatology. *Atmos. Chem. Phys.*, **12**, 1965-1977, doi:10.5194/acp-12-1965-2012.

von Clarmann, T., 2014: Smoothing error pitfalls. *Atmos. Meas. Tech.*, **7,** 3023-3034, doi:10.5194/amt-7-3023-2014.

von Savigny, C., *et al.*, 2015: Improved stratospheric aerosol extinction profiles from SCIAMACHY: Validation and Sample results. *Atmos. Meas. Tech.*, **8**, 5223-5235, doi:10.5194/amt-8-5223-2015.

Wang, H. J., *et al.*, 1996: A critical analysis of SAGE ozone trends. *J. Geophys. Res.*, **101**, 12495-12514, Paper No. 96JD00581.

Wang, H. J., *et al.*, 2002: Assessment of SAGE version 6.1 ozone data quality. *J. Geophys. Res.*, **107**, 4691, doi:10.1029/2002JD002418.

Wang, D.-Y., *et al.*, 2004: Cross-validation of MIPAS/ENVISAT and GPS-RO/CHAMP temperature profiles. *J. Geophys. Res.,* **109**, D19311, doi:10.1029/2004JD004963.

Wang, D.-Y., *et al.*, 2005a: Comparisons of MIPAS/ENVISAT ozone profiles with SMR/ODIN and HALOE/UARS observations. *Adv. Space Res.*, **36**, 927-931, doi:10.1016/j.asr.2005.03.015.

Wang, D. Y., *et al.*, 2005b: Validation of stratospheric temperatures measured by Michelson Interferometer for Passive Atmospheric Sounding MIPAS on Envisat. *J. Geophys. Res.*, **110**, D08301, doi:10.1029/2004JD005342.

Wang, H.-J., *et al.*, 2006: SAGE III solar ozone measurements: Initial results. *Geophys. Res. Lett.*, **33**, L03805, doi:10.1029/2005GL025099.

Wang, D. Y., *et al.*, 2007: Validation of nitric acid retrieved by the IMK/IAA processor from MIPAS/ENVISAT measurements. *Atmos. Chem. Phys.*, **7**, 721-738, doi:10.5194/acp-7-721-2007.

Wang, S., *et al.*, 2008: Validation of Aura Microwave Limb Sounder OH measurements with Fourier Transform Ultra-Violet Spectrometer total OH column measurements at Table Mountain, California. *J. Geophys. Res.*, **113**, D22301, doi:10.1029/2008JD009883.

Wang, S., *et al.*, 2013: Midlatitude atmospheric OH response to the most recent 11-y solar cycle. *Proc. Nat. Acad. Sci.*, **110**, 2023-2028, doi:10.1073/pnas.1117790110.

Waters, J. W., *et al.*, 1993: Stratospheric ClO and ozone from the Microwave Limb Sounder on the Upper Atmosphere Research Satellite. *Nature*, **362**, 597-602, doi:10.1038/362597a0.

Waters, J. W., *et al.*, 1999: The UARS and EOS Microwave Limb Sounder (MLS) Experiments. *J. Atmos. Sci.*, **56**, 194-218, http://dx.doi.org/10.1175/1520-0469(1999)056<0194:TUAEML>2.0.CO;2.

Waters, J. W., *et al.*, 2006: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite. *IEEE Trans. Geosci. Remote Sens.,* **44**, 1075-1092, doi:10.1109/TGRS.2006.873771.

Waugh, D. W., and V. Eyring, 2008: Quantitative performance metrics for stratospheric-resolving chemistry-climate models. *Atmos. Chem. Phys.*, **8**, 5699-5713, doi:10.5194/acp-8-5699-2008.

Waugh, D. W., *et al.*, 2009: Impacts of climate change on stratospheric ozone recovery. *Geophys. Res. Lett.*, **36**, L03805, doi:10.1029/2008GL036223.

Weinstock, E. M., *et al.*, 2009: Validation of the Harvard Lyman-a in situ water vapour instrument: Implications for the mechanisms that control stratospheric water vapour. *J. Geophys. Res.*, **114**, D23301, doi:10.1029/2009JD012427.

Wennberg, P. O., *et al.*, 1994: Removal of stratospheric O_3 by radicals: In situ measurements, of OH, HO_2, NO, NO_2, ClO, and BrO. *Science*, **266**, 398-404, doi:10.1126/science.266.5184.398.

Winker, D. M., *et al.*, 2003: The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds. *Proc. SPIE Int. Soc. Opt. Eng.*, **4893**, 1-11, doi:10.1117/12.466539.

WMO, 2011: World Meteorological Organization Scientific Assessment of Ozone Depletion: 2010. Report 52, Global Ozone Research and Monitoring Project.

WMO, 2014: World Meteorological Organization Scientific Ozone Assessment of Ozone Depletion: 2014, Report 55, Global Ozone Research and Monitoring Project..

Wolff, M. A., *et al.*, 2008: Validation of HNO_3, $ClONO_2$ and N_2O_5 from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS). *Atmos. Chem. Phys.*, **8**, 3529-3562, doi:10.5194/acp-8-3529-2008.

Worden, J. S., *et al.*, 2004: Predicted errors of Tropospheric Emission Spectrometer nadir retrievals from spectral window selection. *J. Geophys. Res.*, **109**, D09308, doi:10.1029/2004JD004522.

Worden, H. M., *et al.*, 2007: Comparisons of Tropospheric Emission Spectrometer (TES) ozone profiles to ozonesodes: methods and initial results. *J. Geophys. Res.*, **112**, D03309, doi:10.1029/2006JD007258.

Worden, H. M., *et al.*, 2008: Satellite measurements of the clear-sky greenhouse effect from tropospheric ozone. *Nature Geosci.*, **1**, 305-308, doi:10.1038/ngeo182.

Worden, H. M., *et al.*, 2011: Sensitivity of outgoing longwave radiative flux to the global vertical distribution of ozone characterized by instantaneous radiative kernels from Aura TES. *J. Geophys. Res.*, **116**, D14115, doi:10.1029/2010JD015101.

Worden, H. M., *et al.*, 2013: Decadal record of satellite carbon monoxide observations. *Atmos. Chem. Phys.*, **13**, 837-850, doi:10.5194/acp-13-837-2013.

Yoon, J., *et al.*, 2013: Technical Note: Temporal change in averaging kernels as a source of uncertainty in trend estimates of carbon monoxide retrieved from MOPITT. *Atmos. Chem. Phys.*, **13**, 11307-11316, doi:10.5194/acp-13-11307-2013.

Zawodny, J. M., and M. P. McCormick, 1991: Stratospheric aerosol and gas experiment II measurements of the quasi-biennial oscillation of ozone and nitrogen dioxide. *J. Geophys. Res.*, **96**, 9371-9377.