

Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative

M. Toohey,¹ M. I. Hegglin,² S. Tegtmeier,¹ J. Anderson,³ J. A. Añel,^{4,5} A. Bourassa,⁶ S. Brohede,^{7,8} D. Degenstein,⁶ L. Froidevaux,⁹ R. Fuller,⁹ B. Funke,¹⁰ J. Gille,^{11,12} A. Jones,¹³ Y. Kasai,¹⁴ K. Krüger,^{1,15} E. Kyrölä,¹⁶ J. L. Neu,⁸ A. Rozanov,¹⁷ L. Smith,¹¹ J. Urban,⁷ T. von Clarmann,¹⁸ K. A. Walker,¹³ and R. H. J. Wang¹⁹

Received 31 May 2013; revised 4 September 2013; accepted 28 September 2013; published 21 October 2013.

[1] Monthly zonal mean climatologies of atmospheric measurements from satellite instruments can have biases due to the nonuniform sampling of the atmosphere by the instruments. We characterize potential sampling biases in stratospheric trace gas climatologies of the Stratospheric Processes and Their Role in Climate (SPARC) Data Initiative using chemical fields from a chemistry climate model simulation and sampling patterns from 16 satellite-borne instruments. The exercise is performed for the long-lived stratospheric trace gases O₃ and H₂O. Monthly sampling biases for O₃ exceed 10% for many instruments in the high-latitude stratosphere and in the upper troposphere/lower stratosphere, while annual mean sampling biases reach values of up to 20% in the same regions for some instruments. Sampling biases for H₂O are generally smaller than for O₃, although still notable in the upper troposphere/lower stratosphere and Southern Hemisphere high latitudes. The most important mechanism leading to monthly sampling bias is nonuniform temporal sampling, i.e., the fact that for many instruments, monthly means are produced from measurements which span less than the full month in question. Similarly, annual mean sampling biases are well explained by nonuniformity in the month-to-month sampling by different instruments. Nonuniform sampling in latitude and longitude are shown to also lead to nonnegligible sampling biases, which are most relevant for climatologies which are otherwise free of biases due to nonuniform temporal sampling.

Citation: Toohey, M., et al. (2013), Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative, *J. Geophys. Res. Atmos.*, 118, 11,847–11,862, doi:10.1002/jgrd.50874.

1. Introduction

[2] Stratospheric trace gas observations are often used to produce monthly zonal mean data sets, or climatologies [e.g., von Clarmann et al., 2012; Groß and Russell, 2005; Hassler et al., 2009; Jones et al., 2012; Randel and Wu, 2007]. Monthly zonal mean data products are typically used as prescribed forcing for models [e.g., Cionni et al., 2011]

and are also useful for comparison with similarly averaged chemistry climate model output [e.g., Grewe et al., 2012; SPARC CCMVal, 2010].

[3] Monthly zonal mean data sets can be constructed through a variety of methods. A simple method involves binning observations into latitude and month bins and calculating average values for each bin. This is the method used

Additional supporting information may be found in the online version of this article.

¹GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany.

²Department of Meteorology, University of Reading, Reading, UK.

³Atmospheric Science, Hampton University, Hampton, Virginia, USA.

⁴Smith School of Enterprise and the Environment, University of Oxford, Oxford, UK.

⁵EPhysLab, Universidade de Vigo, Ourense, Spain.

⁶Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

⁷Department of Earth and Space Sciences, Chalmers University of

Technology, Göteborg, Sweden.

⁸Now at FluxSense AB, Göteborg, Sweden.

⁹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.

¹⁰Instituto de Astrofísica de Andalucía, Granada, Spain.

¹¹National Center for Atmospheric Research, Boulder, Colorado, USA.

¹²Center for Limb Atmospheric Sounding, University of Colorado Boulder, Boulder, Colorado, USA.

¹³Department of Physics, University of Toronto, Toronto, Ontario, Canada.

¹⁴National Institute of Information and Communications Technology, Koganei, Japan.

¹⁵Now at Department of Geosciences, University of Oslo, Oslo, Norway.

¹⁶Finnish Meteorological Institute, Helsinki, Finland.

¹⁷Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany.

¹⁸Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research, Karlsruhe, Germany.

¹⁹School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.

Corresponding author: M. Toohey, GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany. (mtoohey@geomar.de)

Table 1. Participating Instruments With the Time Period Used to Define the Sampling Pattern for Each Instrument

Instrument	Full Name	Mission Reference	Sample Reference Period
ACE-FTS	Atmospheric Chemistry Experiment–Fourier Transform Spectrometer	<i>Bernath et al.</i> [2005]	2005
Aura MLS	Aura Microwave Limb Sounder	<i>Waters et al.</i> [2006]	Jan 2005
GOMOS	Global Ozone Monitoring by Occultation of Stars	<i>Bertaux et al.</i> [2010]	2003
HALOE	Halogen Occultation Experiment	<i>Russell et al.</i> [1993]	1994
HIRDLS	High Resolution Dynamics Limb Sounder	<i>Gille et al.</i> [2008]	Sep 2006
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding	<i>Fischer et al.</i> [2008]	Jan 2009 ^a
OSIRIS	Optical Spectrograph and Infrared Imager System	<i>Murtagh et al.</i> [2002]	2009
POAM II	Polar Ozone and Aerosol Measurement II	<i>Glaccum et al.</i> [1996]	1995
POAM III	Polar Ozone and Aerosol Measurement III	<i>Lucke et al.</i> [1999]	2001
SAGE II	Stratospheric Aerosol and Gas Experiment II	<i>McCormick et al.</i> [1989]	1990
SAGE III	Stratospheric Aerosol and Gas Experiment III	<i>Thomason and Taha</i> [2003]	2003
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography	<i>Bovensmann et al.</i> [1999]	2010
SMILES	Superconducting Submillimeter-Wave Limb Emission Sounder	<i>Kikuchi et al.</i> [2010]	Oct 2009 – Apr 2010
SMR	Submillimetre Radiometer	<i>Murtagh et al.</i> [2002]	2010
TES	Tropospheric Emission Spectrometer	<i>Beer</i> [2006]	Jul 2007
UARS MLS	UARS Microwave Limb Sounder	<i>Waters et al.</i> [1999]	1992

^aThe MIPAS sampling pattern used here refers to the nominal reduced spectral resolution measurement mode, which was the regular MIPAS measurement mode from 2005 to 2012.

by the Stratospheric Processes and Their Role in Climate (SPARC) Data Initiative [Hegglin and Tegtmeier, 2011; M. I. Hegglin et al., SPARC Data Initiative: Comparison of trace gas and aerosol climatologies from international satellite limb sounder, manuscript in preparation, 2013], which aims to produce such climatologies for a number of stratospheric trace gas measuring limb sounding instruments to be used in model-observation comparisons and help guide future data merging activities.

[4] Each observation-based climatology is an estimate of the true atmospheric mean state; however, differences between the observation-based climatology and the truth may arise for a number of reasons. In many cases, the largest source of error in the climatology is due to errors in the measurements themselves. These errors are present in individual measured profiles and can be best estimated through careful comparison of coincident profiles. However, the construction of climatologies may itself introduce additional errors. The choice of averaging technique can lead to differences in the produced climatologies [Funke and von Clarmann, 2012]. The random error in a climatological mean due to the finite and potentially nonuniform sampling of an atmospheric field was investigated by Toohey and von Clarmann [2013]. Here we investigate how such sampling issues may lead to systematic errors, or biases, in atmospheric climatologies.

[5] Sampling bias is an error in a computed quantity which arises due to unrepresentative sampling of the population. In the estimation of an atmospheric mean value, sampling bias may occur when the atmospheric state within the time-space domain to be averaged over is not uniformly sampled. The magnitude of sampling bias will also be related to the degree and structure of the variability: If the atmospheric variability is weak, then nonuniformity of sampling will not strongly bias the sample mean compared to the population mean. Conversely, when variability is strong, then nonuniform sampling may lead to significant sampling bias.

[6] The impact of sampling and averaging on observation-based climatologies of tropospheric fields such as temperature, clouds, and chemical trace gases has been examined by a number of prior studies [e.g., Aghedo et al., 2011; Engelen et al., 2000; Guan et al., 2013]. The central technique of these studies is to subsample climate model or reanalysis fields based on the

sampling patterns of the instruments of interest and then to quantify differences between the fully resolved raw fields and the instrument-sampling-pattern-based sampled fields.

[7] In the present study, we use a similar technique as the above cited works, making use of chemistry climate model output, and sampling the model fields based on the sampling patterns of a suite of satellite instruments. However, we focus on instruments measuring stratospheric trace gas species, many of which have smaller sample density than the (often nadir viewing) instruments investigated in previous studies focusing on tropospheric measurements. Furthermore, we focus on the effect of averaging within 5° latitude bands, which is a coarser horizontal resolution than that typically used in prior studies. We use sampling patterns for 16 instruments, all of which are participants in the SPARC Data Initiative. We apply the sampling exercise to modeled O₃ and H₂O fields and characterize the structure and magnitude of the potential sampling biases associated with the sampling pattern of each instrument for the two chemical species. The results of this study should be used to help interpret instrumental climatology comparisons within the SPARC Data Initiative for O₃ and H₂O (Tegtmeier et al. [2013] and Hegglin et al. [2013], respectively) and inform future use of the climatologies for other purposes such as model comparisons.

2. Data and Methods

2.1. Instrumental Sampling Patterns

[8] Instruments participating in the SPARC Data Initiative and which are used in the present study are listed in Table 1, and include the limb emission sounders Aura MLS, HIRDLS, MIPAS, SMR, SMILES and UARS MLS, the limb scattering sounders SCIAMACHY and OSIRIS, the solar occultation instruments ACE-FTS, HALOE, POAM II, POAM III, SAGE II and SAGE III as well as the stellar occultation instrument GOMOS. We have also included the nadir-viewing TES emission sounder, which is used in the SPARC Data Initiative evaluation of upper troposphere/lower stratosphere (UTLS) ozone observations (J. L. Neu et al., The SPARC Data Initiative: Comparison of upper troposphere/lower stratosphere ozone climatologies from limb-viewing instruments and the nadir-viewing Tropospheric Emission Spectrometer (TES), submitted

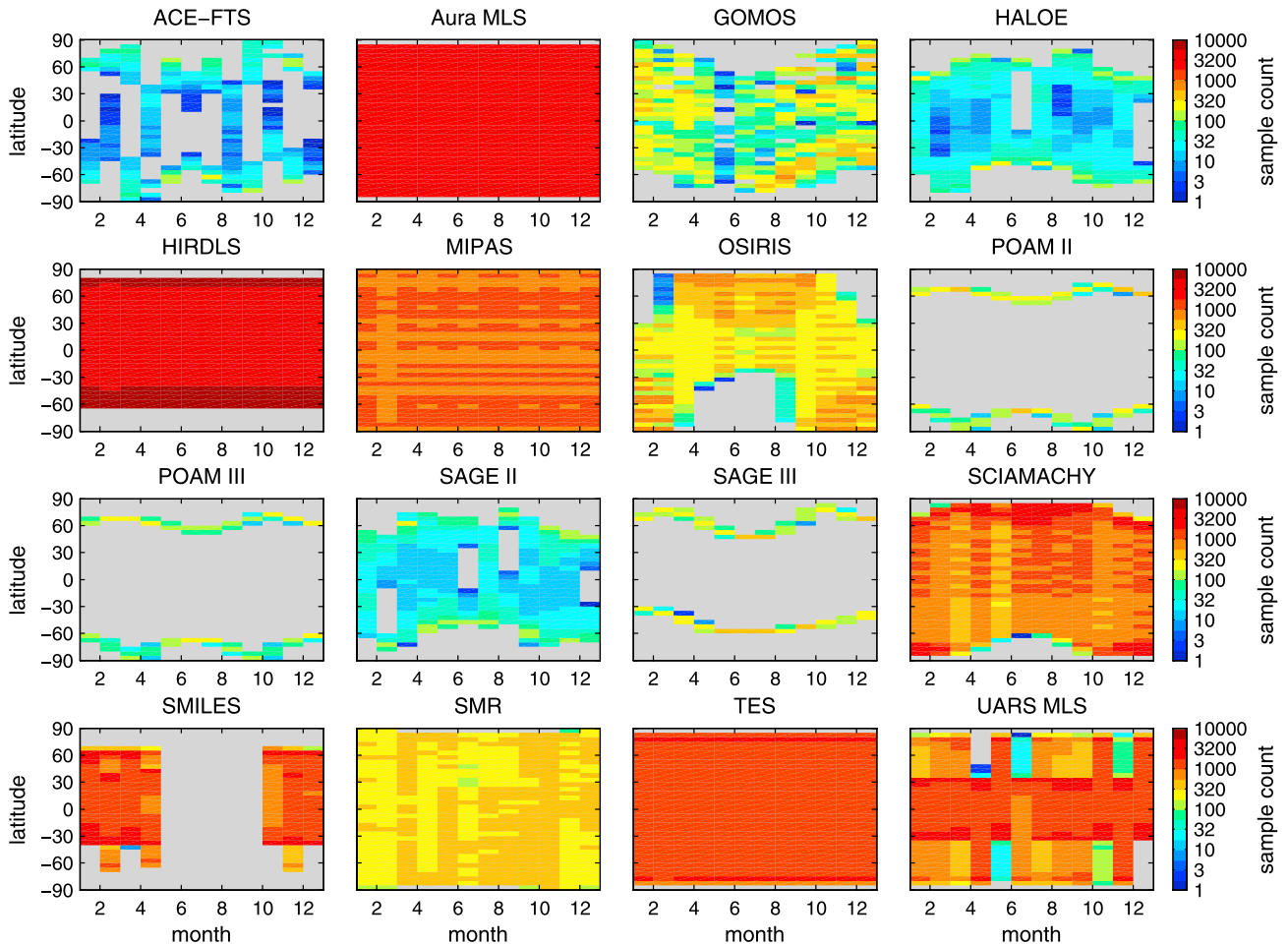


Figure 1. Monthly sample count in 5° latitude bins for 16 instrumental sampling patterns. Gray regions denote regions of no measurements.

to *Journal of Geophysical Research*, 2013). TES has much coarser vertical resolution ($\sim 6\text{--}7$ km versus $\sim 2\text{--}4$ km) but much higher horizontal resolution (~ 10 km versus ~ 200 km) than the other instruments in this study. Detailed information on the individual instruments including their sampling patterns and retrieval techniques can be found in the SPARC Data Initiative report (section 2) and Hegglin et al. (manuscript in preparation, 2013).

[9] Sampling patterns have been compiled for each instrument, defined as day, time, latitude, and longitude of measurement locations. For most 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 for reasons such as data download limitations). The particular years that were used to define each instrument's sampling pattern are included in Table 1. Year-to-year variations in sampling patterns due to such factors as changing orbital states, changing instrument capabilities, or irregular data gaps—which may be significant for a number of instruments—are not addressed by this study. For Aura MLS, HIRDLS, MIPAS, and TES, month-to-month variations in sampling are generally negligible, and we have therefore used a typical month from their sampling pattern and repeated this for all months of the year. For SMR, sampling patterns are defined separately for O_3 and

H_2O , based on locations of retrieved profiles from the 501.8 GHz and 488.9 GHz measurement bands, respectively. (In addition to the 488.9 GHz band H_2O retrievals, which extend from 20 to 70 km, SMR also produces climatologies of H_2O from ~ 16 to 20 km through measurements of the 544.6 GHz band [Hegglin et al., 2013], however these measurements are not considered here.) For SMILES, the sampling patterns correspond to the samples over the full lifetime of the mission, from October 2009 to April 2010.

[10] Monthly sample counts within the 5° bins of the SPARC Data Initiative grid are shown for the instrumental sampling patterns in Figure 1. The different sample patterns of remote sounders of the stratosphere lead to substantial differences in the sample density patterns shown in Figure 1. Solar occultation instruments typically have the sparsest sample density, and therefore the smallest monthly sample counts, because their number of measurements is limited to a maximum of two per orbit. The latitudinal coverage of solar occultation instruments varies with time, leading to samples concentrated within finite latitude ranges (usually one in the Northern Hemisphere (NH) and one in the Southern Hemisphere (SH)) each month and also leaving unsampled latitudes each month. Solar occultation instruments with high inclination orbits (~ 57 to 74° , e.g., ACE-FTS, HALOE, and SAGE II) have sampling which spans nearly the full globe over a year. Other

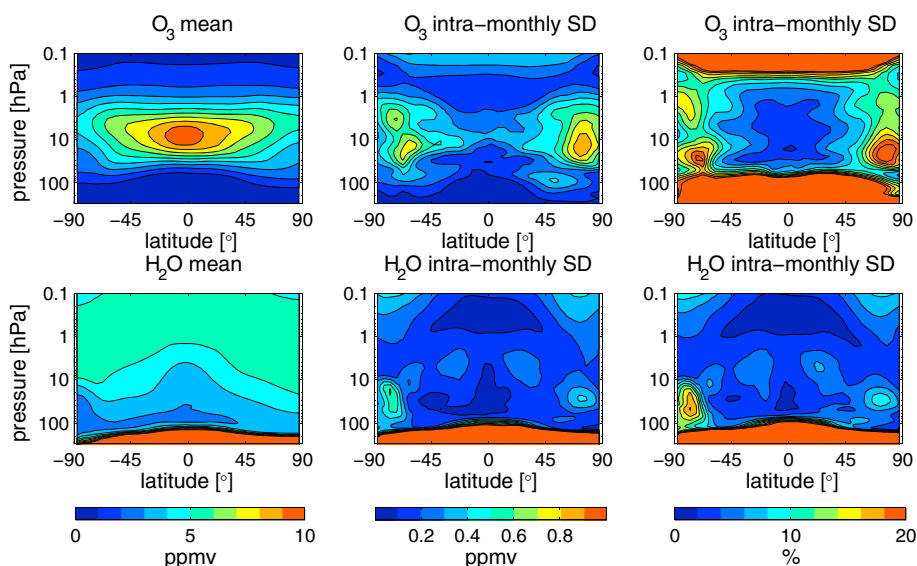


Figure 2. WACCM simulated O₃ and H₂O mean distribution and variability. (left) Annual mean O₃ and H₂O, average of monthly standard deviations over 12 months in (middle) absolute and (right) percent.

solar occultation instruments (POAM II, POAM III, and SAGE III) with Sun-synchronous orbits have sample ranges concentrated in the middle to high latitudes. Climatologies produced for the stellar occultation instrument GOMOS use only night measurements and therefore exclude measurements from the polar summer. Instruments which measure scattered sunlight (SCIAMACHY, OSIRIS) have denser sampling but are restricted to measuring within the sunlit portion of the atmosphere, leading to a smaller number of samples in the winter hemisphere, which may be just the polar latitudes, or a larger portion of the winter hemisphere depending on the pointing of the instrument. Instruments which measure atmospheric emission (Aura MLS, HIRDLS, MIPAS, SMILES, SMR, TES, and UARS MLS) do not have such constraints related to the radiation source, and thus typically have sampling coverage that is denser than solar occultation and limb scattering instruments and relatively uniform with latitude and time. An exception to the latter case is the sampling pattern of UARS MLS, which produced measurements spanning 34° on one side of the equator to 80° on the other side, with hemispheric coverage switching when the satellite performed a 180° yaw maneuver 10 times per year, at approximately 36 day intervals.

2.2. Model Fields

[11] Chemical fields are taken from an integration of the Whole Atmosphere Community Climate Model version 3 (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 essentially the same as that used for the second Chemistry-Climate Model Validation Activity [Morgenstern *et al.*, 2010; SPARC CCMVal, 2010]. The model’s horizontal resolution is 1.9° by 2.5° (latitude by longitude). WACCM has been extensively evaluated as part of the SPARC CCMVal report [2010]. WACCM has a good distribution and variability of O₃ and H₂O in the stratosphere [SPARC CCMVal, 2010, chapter 9] and UTLS [Gettelman *et al.*, 2010; Hegglin *et al.*, 2010]. We use here model output with daily

resolution at 0 UTC from 1 year of a transient simulation under modern conditions.

[12] Annual zonal mean mixing ratios for O₃ and H₂O are shown in Figure 2. Since sampling biases result from nonuniform sampling of a varying field, we also show for reference a measure of “average variability” for the WACCM fields, calculated by taking the standard deviation of the daily fields for each month and averaging over the 12 calendar months. The monthly standard deviations are calculated over all model values for each latitude and pressure level, i.e., over all times and longitudes, and as such represent the magnitude of the variability due to temporal progression of the seasonal cycle within any month and that due to dynamically induced disturbances to zonal symmetry. Maximum variability of O₃ is seen in the high latitudes, where the O₃ seasonal cycle is strongest, and where meridional gradients in the mean O₃ field mean that dynamical variability results in chemical variability. The variability of H₂O mixing ratios is notably weaker, a result of its weaker seasonal cycle and weaker meridional gradients. Strongest variability of H₂O is seen in the SH lower stratosphere high latitudes and is a result of the dehydration of the SH winter polar vortex.

2.3. Sampling Bias Calculation

[13] Starting with full resolution model fields, the goal of the exercise is to emulate as closely as possible the sampling of each instrument and the climatology procedure used to produce the SPARC Data Initiative climatologies and compare these climatologies to the “true” model climatology.

[14] First, the 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. Once the model data have been interpolated to each sample location, the subsampled fields are processed in the same manner as the real measurements in the production of climatologies, i.e., binned according to the SPARC Data Initiative latitude grid, and the mean calculated for each latitude bin and pressure surface. The

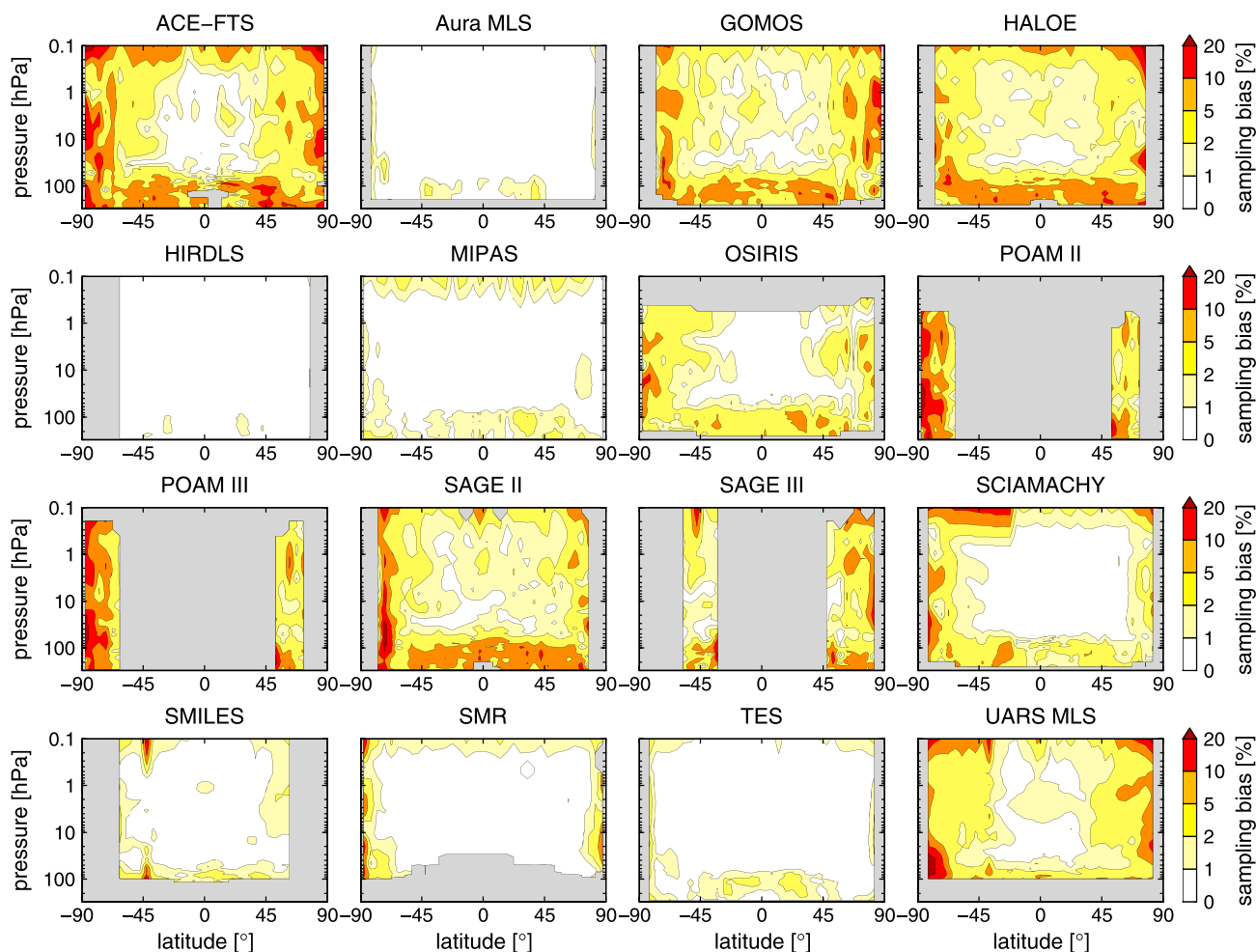


Figure 3. Latitude height sections of calculated RMS monthly percent sampling bias for O_3 , based on sampling patterns of instruments as labeled in each panel. Gray regions denote regions of no measurements.

only exception is that in the case of MIPAS SPARC Data Initiative climatologies, measured mixing ratios were interpolated to the central latitude of the bin, in order to avoid sampling biases. This operation has not been accounted for in this study; our analysis of the sampling error is based on the original geolocations of the measurements. 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 latitude bin. This interpolation is performed due to the fact that the two model latitudes (at 1.9° resolution) within each 5° latitude bin will not be perfectly centered with respect to the bin midpoint, therefore simply averaging the model data within a latitude bin could introduce a small sampling bias to the true model climatology. The difference between the instrument-sampling-pattern-based subsampled 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. The vertical range of presented results for each instrument is then “cropped” based on the range of measurements in the corresponding SPARC Data Initiative climatologies.

[15] This method produces an estimate of sampling bias due to the horizontal sampling of the instruments, and no effort is made here to understand differences between instrumental climatologies due to differences in the vertical resolution of the measurements. Furthermore, by using daily mean model data, we cannot address possible biases in instrumental climatologies due to different sampling of a chemical species' diurnal cycle. As such, we apply our study here only to chemical species with long stratospheric lifetimes.

[16] Ideally, in an exercise such as this, one would like the resolution of the sampled model fields to be similar to the horizontal resolution of the measurements they are meant to emulate. This is because the variability of a sampled field depends on the resolution of the measurement: measurements made at fine resolution will have larger variance than those made at coarser resolution. While such issues are important when comparing measured and modeled variability [e.g., Toohey *et al.*, 2010], when examining average values, the horizontal resolution should not have a large impact on the results, as long as sample sizes are large enough that the mean is well estimated (i.e., the standard error of the mean is small). Nevertheless, the resolution of the model fields used here is roughly consistent with that of the limb-sounding instruments—the horizontal resolution of limb

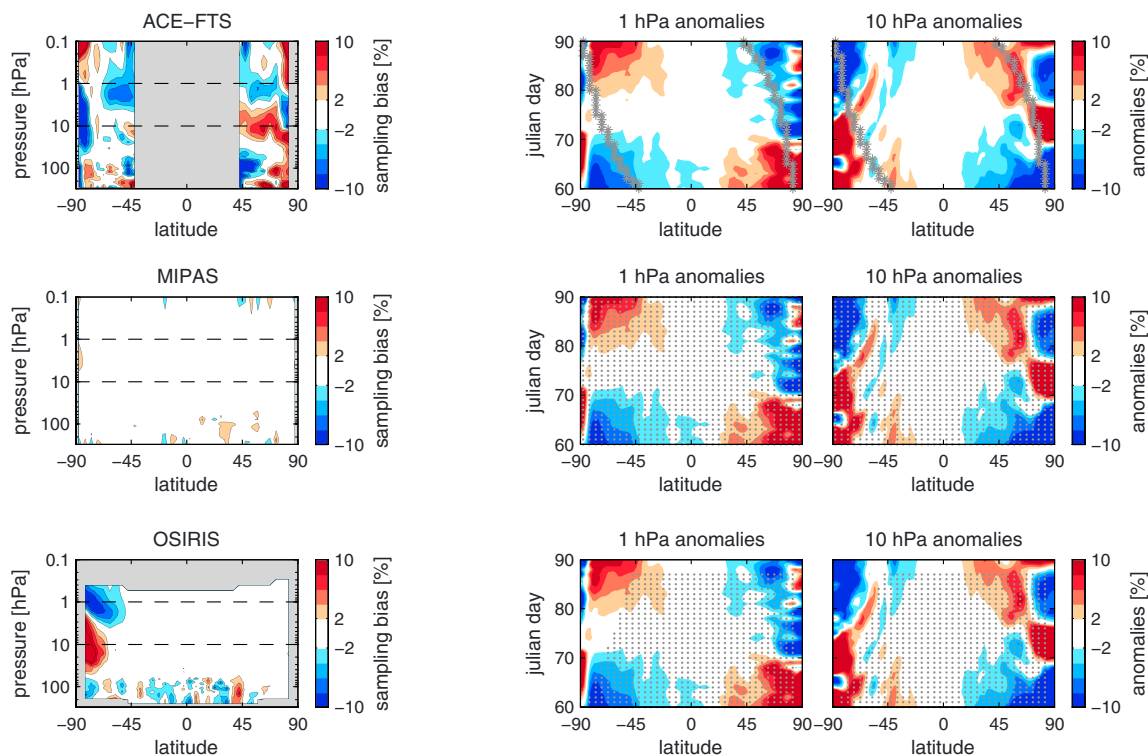


Figure 4. (left column) March O_3 percent sampling bias as function of latitude and pressure for ACE-FTS, MIPAS, and OSIRIS sampling patterns. March WACCM zonal mean O_3 intra-monthly percent anomalies from the monthly mean are shown as a function of latitude and Julian day for the (middle column) 1 hPa and (right column) 10 hPa surfaces, along with gray markers showing either the locations of all samples (ACE-FTS) or the locations of latitude bins which contain measurements (MIPAS and OSIRIS) based on each instrument's sampling pattern.

sounding retrievals is typically on the order of a few hundred kilometers [e.g., von Clarmann *et al.*, 2009]. The model fields used here have latitudinal resolution of about 200 km and longitudinal resolution that ranges from about 250 km at the equator to 200 km in the midlatitudes. At high latitudes, the rough agreement between model and observation resolution no longer holds, as the longitudinal resolution of the model fields is finer than that of the observations; however, we find (as shown in section 3) that estimated sampling biases are well explained by mechanisms for which the horizontal resolution of the model fields should not impact the results.

[17] Sampling biases calculated by the method described above are the biases introduced to the monthly zonal mean of model fields by the respective instrument sampling pattern. Whether the results represent reasonable estimates of the true sampling bias for any instrument's measurement of the real atmosphere depends on how similar the model fields are to reality. The model can be expected to reproduce in a fairly accurate way the seasonal cycle in chemical trace gases, and short-term (subseasonal) variability can be expected to be consistent with reality in a statistical sense. However, it should be clear that any sampling bias estimate resulting from short-term variability within this 1 year of model data is simply one realization of potential sampling bias. In this way, the sampling bias estimates should be considered example cases, and we focus on characterizing the general features of sampling bias and its dependence on latitude and height for the given sampling patterns.

3. Results

[18] Since sampling bias depends on the space-time gradients of the measured field, it can be different for different trace gases. In the following, we explore sampling bias for two trace gas species included in the SPARC Data Initiative climatologies: O_3 and H_2O .

3.1. O_3 Sampling Bias

3.1.1. Monthly Mean Sampling Bias

[19] Zonal mean sampling bias for O_3 is shown for each instrument and each calendar month as a function of latitude and height in the supporting information. To summarize the absolute magnitude and spatial structure of sampling bias, Figure 3 shows the root-mean-square (RMS) of the 12 monthly sampling bias estimates for each instrument.

[20] In general, at pressure levels above ~ 70 hPa, O_3 sampling bias for all instruments is weak in the tropics where intramonthly variability is weak (see Figure 2). In the extratropics and polar latitudes, where variability is stronger, sampling bias becomes much larger. Sampling bias is found to be weakest for the instruments with dense and uniform sampling density, Aura MLS, HIRDLS, MIPAS, and TES. RMS bias for these instruments is less than 1% over most of the stratosphere, with maximum values of 1–3% occurring in isolated regions at the high latitudes, and within the upper troposphere, where O_3 mixing ratios are quite small. SMR, SMILES, and SCIAMACHY have similarly

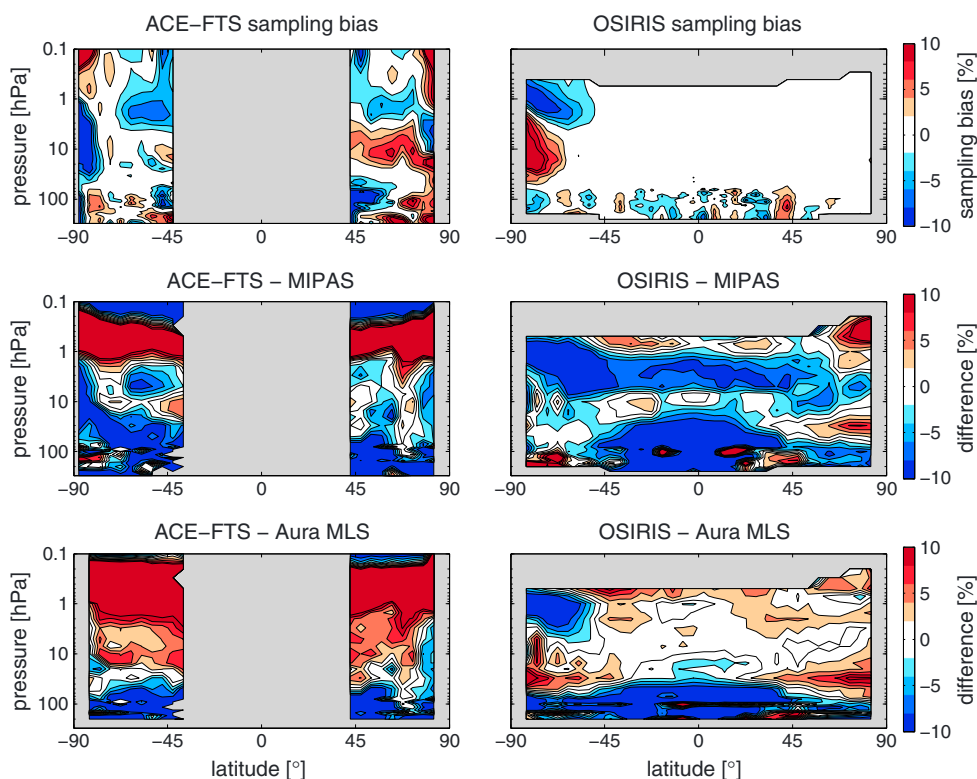


Figure 5. (top row) March O_3 percent sampling bias for ACE-FTS and OSIRIS, as in Figure 3. Differences between each instrument's March O_3 climatology and (middle row) MIPAS and (bottom row) Aura MLS.

small ($<1\%$) sampling bias through most of the tropical and subtropical stratosphere, but with larger RMS sampling biases in the middle to high latitudes. SMILES and SCIAMACHY show RMS sampling bias of $>1\%$ poleward of 45° in the NH and SH, respectively, while SMR and SCIAMACHY sampling bias reaches $>5\%$ in regions of the high latitudes of both hemispheres. RMS sampling bias of $>2\%$ covers most of the middle to high latitudes for OSIRIS and UARS MLS, and inspection of the monthly sampling bias plots reveals that the results are variable in time, with large sampling bias ($>10\%$) in the middle to high latitudes for certain months of the year. Finally, the occultation instruments ACE-FTS, GOMOS, HALOE, POAM II, POAM III, SAGE II, and SAGE III show strong sampling bias, with RMS values greater than 5% over much of the high latitudes, and reaching values of $>10\%$ in isolated locations. A number of these instruments also show RMS sampling biases of $1\text{--}2\%$ through much of the tropical stratosphere. The large sampling biases at high altitudes (above 0.4 hPa) calculated for many instruments should be considered to be an artifact of the exercise, due to the fact that O_3 has an appreciable diurnal cycle at these heights, and since the model fields are saved at 0 UTC, the diurnal variability of O_3 is expressed as a longitudinal structure in the O_3 field.

[21] The largest sampling biases in Figure 3 can be understood to be a product of nonuniform sampling through the days of a month, as can be seen when one examines variations in O_3 over a month, and the correlation of these variations with instrument sampling patterns. Sampling biases for the month of March are shown in Figure 4 for

ACE-FTS, MIPAS, and OSIRIS, as examples of sampling bias estimates that span the range of results discussed above. Sampling bias for this month is strong for ACE-FTS (reaching values of $>5\%$), weak for MIPAS, and strong for OSIRIS in the SH. In Figure 4 (right column), the model O_3 field in March is shown in terms of percent anomalies from the monthly zonal mean and plotted as a function of latitude and Julian day for the 1 and 10 hPa pressure surfaces. Included in these plots are gray markers indicating the latitude bins in each day which contain measurements based on each instrument's sampling pattern.

[22] The MIPAS sampling pattern contains measurements in all latitude bins for all days: There is no variation in its 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 midlatitudes (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, O_3 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 SH latitudes. The seasonal cycle of O_3 is comparatively reversed at 10 hPa, leading to slightly positive bias in the SH midlatitudes 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 its sampling of the temporal variations in O_3 , which depend strongly on height and latitude. At 100 hPa, intramonthly O_3 variations are relatively

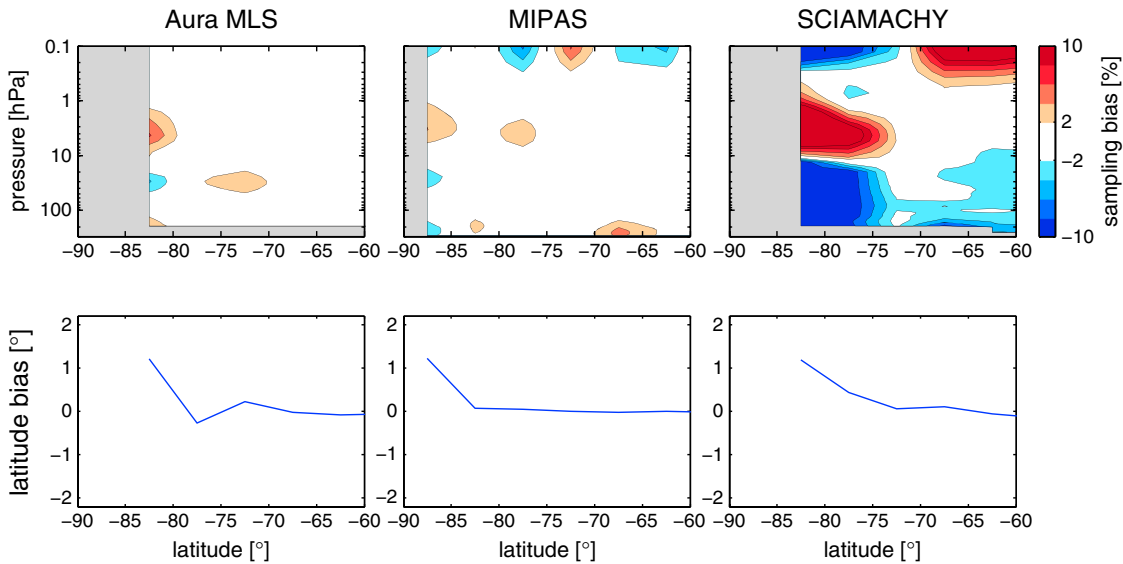


Figure 6. (top row) Sampling bias for September O_3 in SH high latitudes for selected instruments with (bottom row) differences between the latitude bin center and the mean sampled latitude.

noisy (not shown), and as a result, the sampling bias is dependent on the sampling of short-lived variability. We therefore can expect that in regions where the sampling bias is due to the nonuniform sampling of the slow seasonal variability through a month, the sign and approximate magnitude of the sampling bias calculated through our model exercise is a reasonably accurate estimate of the real sampling bias for each instrument. However, in regions where variability is dominated by short-term (time scale of days) 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 is only an example and should be used only to identify regions where sampling error may be important.

[23] The sampling pattern of OSIRIS in March is dense and uniform in the NH, but it samples SH latitudes only within the first half of the month. The OSIRIS sampling pattern, and thus its sampling bias, is intermediate to the two extreme cases explored above. As a result of the sampling pattern, OSIRIS means in the SH are biased low at 1 hPa and high at 10 hPa. This type of sampling bias occurs for OSIRIS in months where its sampling pattern latitudinal range shifts from one hemisphere to the other. Sampling biases for UARS MLS come from a similar source due to its latitudinal coverage shifting within months, leading to the mid to high latitudes being sampled for only a portion of each month.

[24] The applicability of the sampling bias estimates to real data is demonstrated through a brief case study. Assuming small O_3 sampling bias for MIPAS (as discussed above), one would expect that the difference of another instrument's climatology with respect to that of MIPAS, e.g., ACE-FTS minus MIPAS, should contain the effect of sampling bias in the ACE-FTS climatology, in addition to measurement errors. Figure 5 shows such differences for March climatologies of ACE-FTS and OSIRIS, with differences calculated with respect to both MIPAS and Aura MLS. Average values are shown over 2 years: 2005–2006 for ACE-FTS comparisons and 2008–2009 for OSIRIS comparisons. Differences

between ACE-FTS and the dense sampling instruments are dominated by positive values in the upper stratosphere and mesosphere (3–0.1 hPa). This feature is independent of latitude and is a common feature of comparisons between ACE-FTS and other instruments [Dupuy *et al.*, 2009]. Aside from this primary feature, the signature of the estimated sampling bias can be seen in the ACE-FTS comparisons: For example, the main features of the sampling bias in the SH, namely a negative bias at the highest SH latitudes between ~ 30 and 1 hPa and a negative bias in the midlatitude upper stratosphere (60 – 45° S, 3–1 hPa) can be detected in the ACE-FTS–MIPAS difference plot. In the NH, both ACE-FTS–MIPAS and ACE-FTS–Aura MLS show negative differences in the high latitudes between 30 and 1 hPa, much of which is consistent with the sampling bias estimate. For OSIRIS, the sampling bias signature of a strong negative bias in the SH upper stratosphere (3–0.8 hPa) is clearly evident in both OSIRIS–MIPAS and OSIRIS–Aura MLS difference plots. The predicted signature of positive bias in the high-latitude SH lower stratosphere is also seen in the OSIRIS–Aura MLS difference plots. The sampling bias estimate plot for OSIRIS suggests however that the positive difference between OSIRIS and both comparison instruments in the NH lower stratosphere high latitudes is not a result of sampling bias.

[25] Clearly, there are large differences between the climatologies shown in Figure 5 that are not explained by the sampling bias estimates, such as the positive difference between ACE-FTS and other instruments in the upper stratosphere and the negative bias of OSIRIS compared to MIPAS in the upper stratosphere. Such biases are almost certainly due to systematic biases in the measurements themselves. The comparisons highlight, however, that in certain regions and times, sampling biases make up a significant portion of the difference between monthly mean climatologies.

[26] In addition to temporal considerations discussed above, nonuniformity in latitude sampling can also lead to sampling bias, which can be most notable at the northern

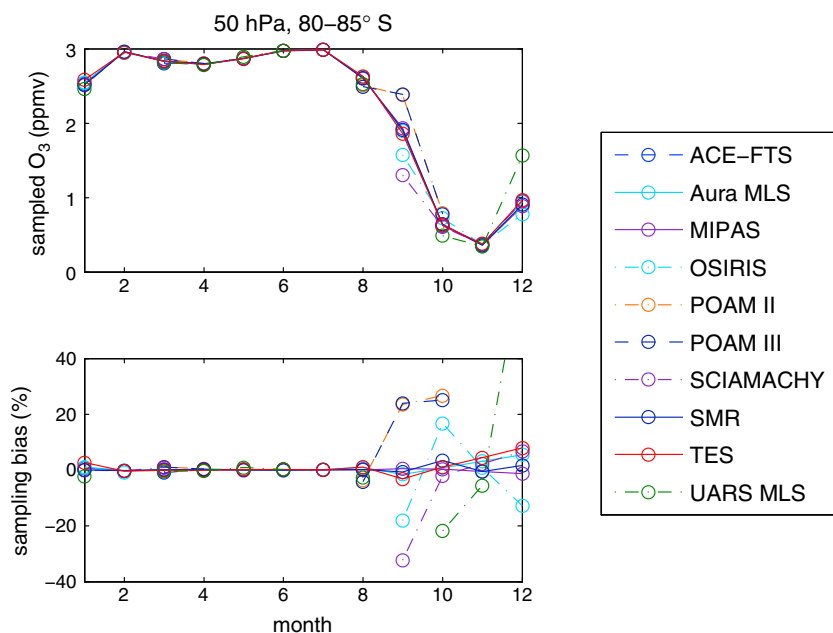


Figure 7. (top) Sample mean model O₃ and (bottom) relative sampling bias at 50 hPa in the 80–85°S latitude bin as a function of month of year, with sampling based on instrumental sampling patterns as defined in legend.

and southern limits of instruments' sampling patterns. In these cases, instrumental sampling patterns typically do not sample the full latitudinal extent of the latitude bin. If the measured species has a significant gradient from one side of the bin to the other, a sampling bias can occur. This sampling bias likely occurs for all instruments and all months to some degree but is most noticeable in cases where other sampling biases are not present, e.g., at the extreme southern latitudes for the dense samplers Aura MLS, HIRDLS, MIPAS and SMR.

[27] As an example, Figure 6 shows sampling biases from selected instruments in the SH high latitudes for the month of September, when Antarctic O₃ depletion leads to a strong gradient in O₃ mixing ratio across the polar vortex. Figure 6 (bottom row) shows the bias in sampled latitude (mean sampled latitude minus latitude midpoint), which is seen to be significant in the southernmost latitude bin of each instrument's sampling range, due to the incomplete coverage of this latitude band. Aura MLS, for example, samples only the northern half of the 80–85°S bin, which leads to a positive sampling bias (of around 4%) in the Aura MLS climatology in this latitude bin at around 3 hPa. In comparison, MIPAS samples the full 80–85°S bin, but only half of the 85–90°S bin. However, the latitudinal gradient in O₃ is weaker over this bin, so the MIPAS sampling bias at 3 hPa is smaller (~2%). All instruments display some degree of latitudinal sampling bias at the edges of their sampling domain; however, the bias in the measured field may be insignificant compared to other sources. For example, the bias in sampled latitude at high latitudes for SCIAMACHY is similar to the other instruments shown in Figure 6; however, the sampling bias for O₃ is dominated by temporal considerations since it measures only sunlit latitudes, in a way that varies with time during this time of year in the southern high latitudes. The bias seen for SCIAMACHY in the

mesosphere and in the UTLS at latitudes lower than 70°S is most probably due to an exclusion of measurements over the southern Atlantic (see below for details).

[28] The potential for sampling bias is greatest when temporal and spatial gradients of the measured field are largest. The Antarctic O₃ hole leads to extreme temporal change in O₃ mixing ratios and creates a strong gradient across the vortex edge in SH spring. The temporal evolution of sampling bias in the 80–85°S latitude bin in the lower stratosphere (50 hPa) is examined in Figure 7. Sampling bias is quite small through most of the year but reaches values of ±30% or more for some instruments in the SH spring months. Most of the behavior shown in Figure 7 can be understood in terms of previously discussed mechanisms. The low bias of SCIAMACHY and OSIRIS in September is a result of the fact that their measurements at high latitudes cover only the latter half of the month, when O₃ depletion has begun. A similar situation occurs for UARS MLS sampling in October and December, leading to a low bias in October and a high bias in December (for the sample year of UARS MLS sampling used here). Slight biases in the 80–85°S bin can be seen for Aura MLS and TES in November and December due to latitudinal bias in this latitude bin, as discussed above.

[29] Nonuniformity in sampled longitude can also lead to sampling bias if the measured field is not zonally symmetric. This is, for example, the case for the SCIAMACHY climatologies compiled for the SPARC Data Initiative. Due to data quality issues, measurements over the South Atlantic were not included in these climatologies, which results in a systematically nonuniform sampling in longitude for latitudes between 20°S and 70°S. The effect of this longitudinal sampling irregularity can be seen in the monthly mean sampling bias plots for SCIAMACHY (see the supporting information), which show systematically larger absolute magnitude sampling

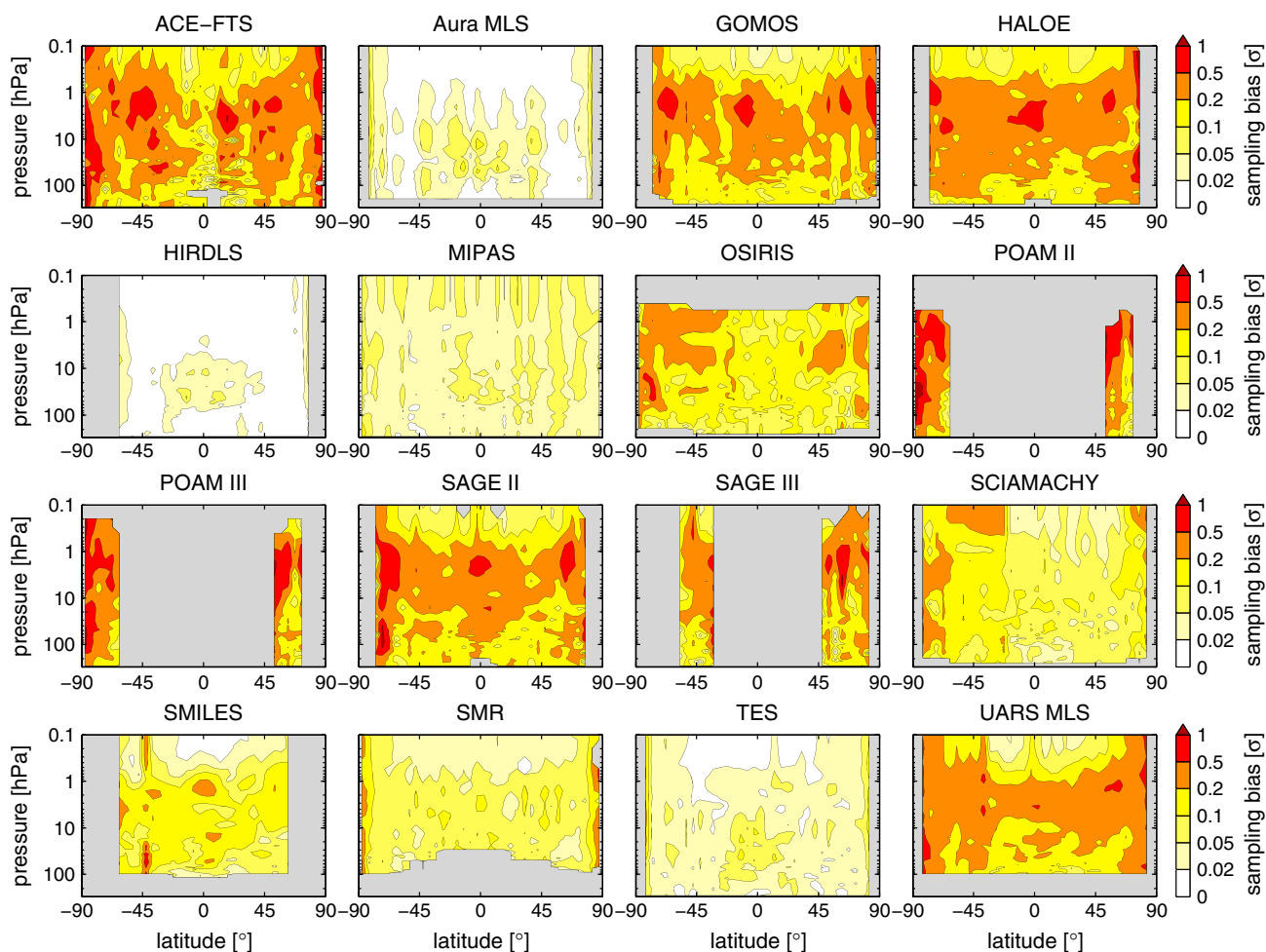


Figure 8. RMS O_3 sampling biases for each instrument, as in Figure 3, but with monthly sampling biases first normalized (by the standard deviation of O_3), and therefore shown in units of σ rather than percent.

biases (with RMS values of $>5\%$, see Figure 3) between 20 and 70°S latitude in the upper troposphere/lower stratosphere (UTLS, approximately 300 to 70 hPa) and lower mesosphere (approximately 1 to 0.1 hPa). The sampling biases at high altitudes should be considered to be an artifact, as discussed above. However, the sampling biases shown for SCIAMACHY climatologies in the UTLS are likely due to real longitudinal dependence of UTLS O_3 due to variations in tropospheric chemistry and dynamics and should be considered when making use of the SCIAMACHY climatologies in the UTLS.

[30] It is clear from the discussion above that sampling bias results from a combination of nonuniform sampling and variability of the sampled field. With this in mind, it is interesting to consider sampling biases normalized by the variability of the chemical field. Figure 8 shows RMS O_3 sampling biases, as in Figure 3, but normalized by the standard deviation of the O_3 field. RMS normalized O_3 sampling bias shown in Figure 8 is more uniform with respect to latitude and height compared to the percent values shown in Figure 3. Maximum RMS normalized sampling biases, as seen for the solar occultation instruments, are on the order of approximately $0.5\text{--}1\sigma$. For the dense sampling instruments, RMS normalized sampling biases are typically $0\text{--}0.1\sigma$. Normalizing by the standard deviation acts to better

isolate the impact of nonuniform instrumental sampling on the sampling bias, making it independent of the magnitude of the chemical field variability. In principle, the normalized O_3 sampling bias estimates could be used to calculate rough sampling bias values for other long-lived trace gases which show variability coherent with that of O_3 , if the variability of the chemical field is known.

3.1.2. Annual Mean Sampling Bias

[31] To assess how sampling bias can affect annual mean climatologies, we produce annual mean sampling bias estimates for each instrument by calculating annual averages of the sampled monthly mean fields for each instrument, then calculating the percent difference of this quantity compared to the model annual mean. The annual mean bias therefore includes not only the sampling biases present in the monthly mean biases discussed above but also any bias due to an instrument's incomplete sampling over the year for any latitude bin. Annual sampling biases for O_3 are shown for each instrument in Figure 9.

[32] The instruments with uniform sampling throughout the year (Aura MLS, HIRDLS, MIPAS, SMR, and TES) show very weak annual mean sampling biases of only a few percent. In contrast, the annual mean sampling biases for the other instruments, all of which have latitudinal sampling which

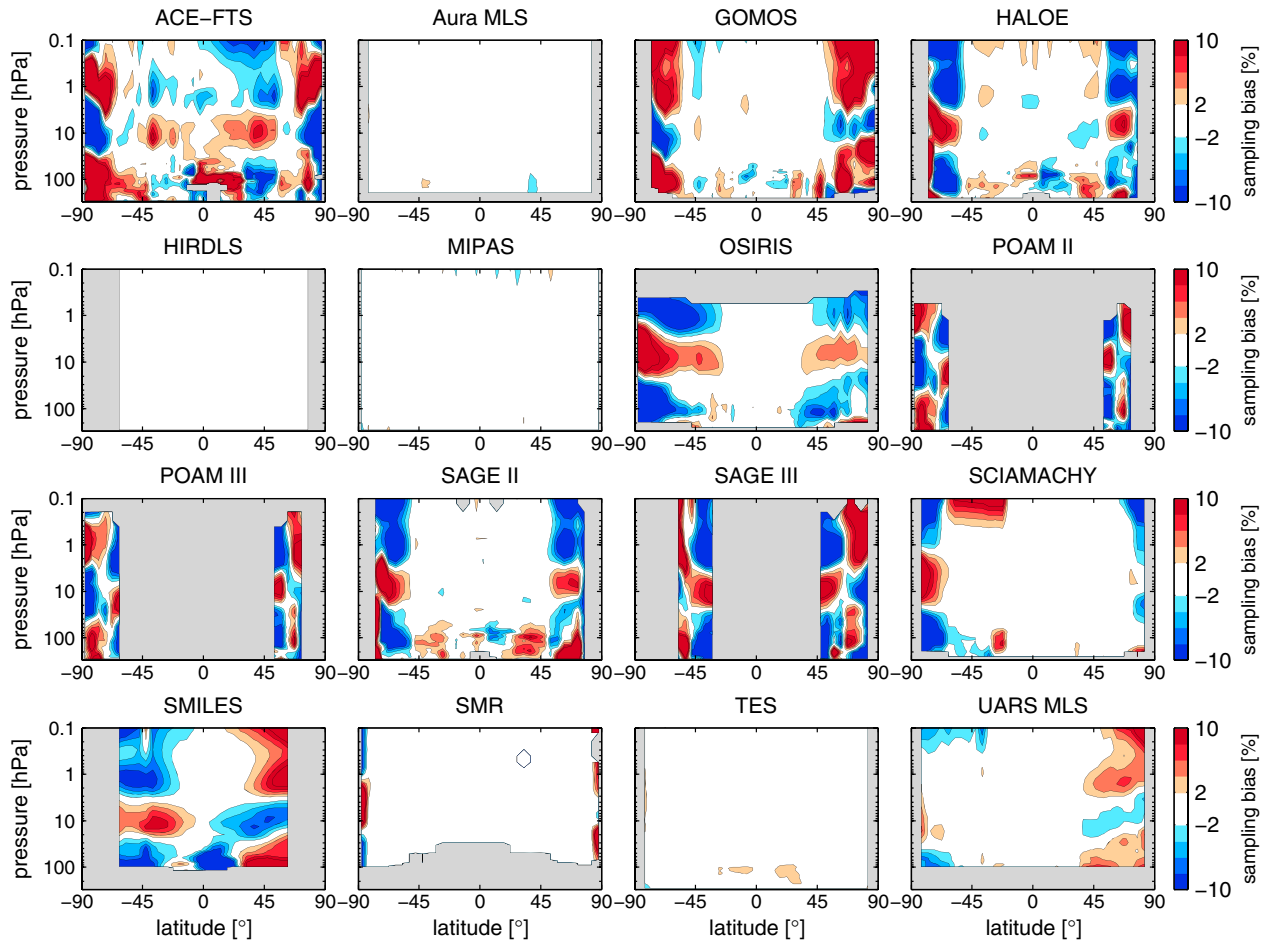


Figure 9. Latitude height sections of calculated annual mean percent sampling bias for O_3 , based on sampling patterns of instruments as labeled in each panel. Gray regions denote regions of no measurements.

varies throughout the year, are considerable, often exceeding 10%, and reaching values of >20% in some cases. The annual mean sampling biases can be qualitatively explained by the seasonality of sampling. For example, the SMILES

sampling pattern covers only approximately half the year, during NH winter and spring. Clearly, some sampling bias is to be expected when calculating an annual mean with a half year of SMILES data, and while this example is perhaps not

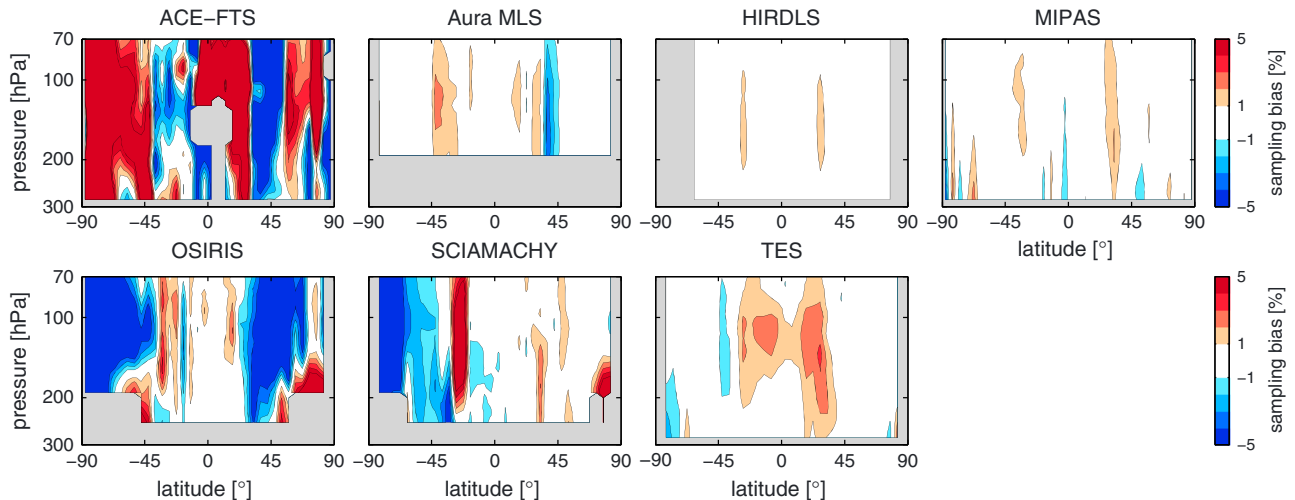


Figure 10. Annual mean O_3 sampling biases in the UTLS for instruments with applicable vertical coverage.

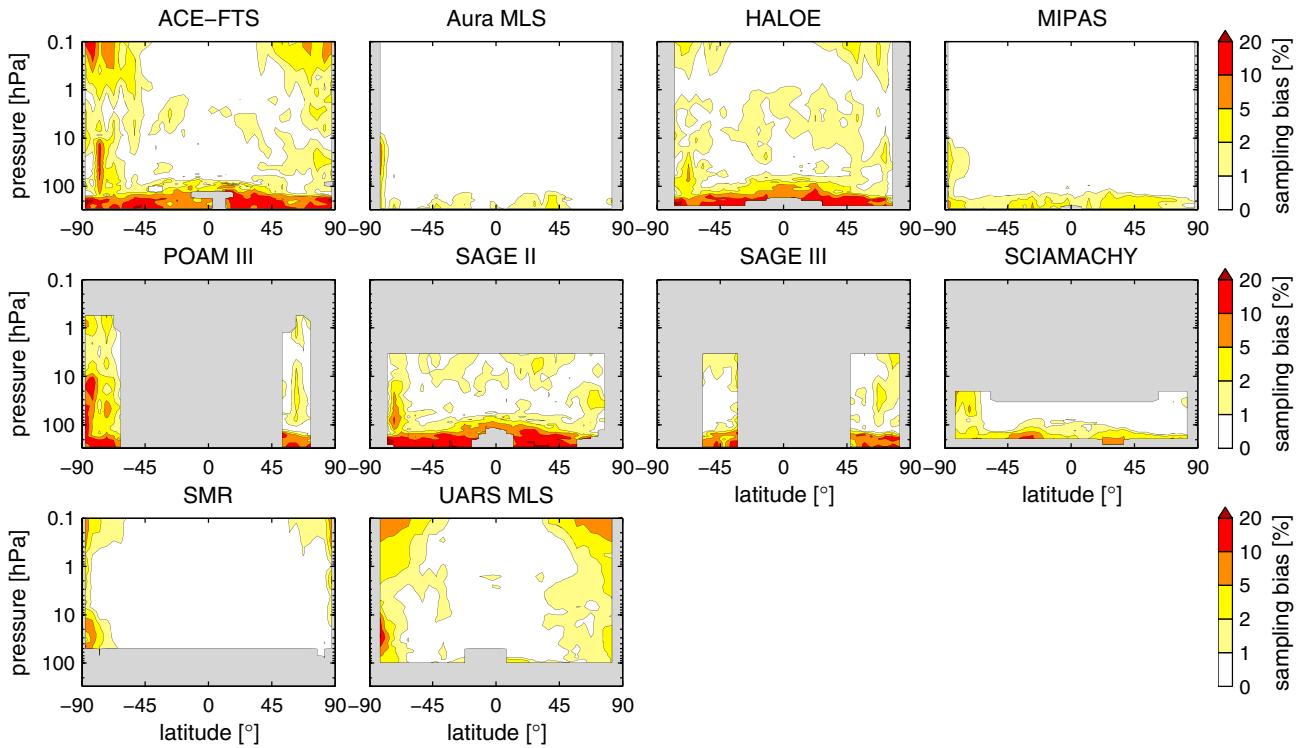


Figure 11. Latitude height sections of calculated RMS monthly percent sampling error for H_2O , based on sampling patterns of instruments as labeled in each panel. Gray regions denote regions of no measurements.

realistic, it helps to illustrate the mechanism behind the sampling biases of other instruments with less dramatic sampling nonuniformity throughout the year. The annual mean sampling bias for the SMILES sampling pattern in each hemisphere is characterized by a “tripole” pattern, with alternating sign with height, and is opposite in sign in each hemisphere. OSIRIS does not sample the high-latitude winters (Figure 1), and as a result, the annual mean sampling bias pattern shows the characteristic tripole pattern in the high latitudes, which is similar in morphology with the SMILES sampling bias in the “summer” SH. HALOE, SAGE II, and SCIAMACHY show similar biases toward summer sampling (see again Figure 1), leading to similar sampling bias morphology for these instruments. GOMOS sampling is biased toward winter in the high latitudes; as a result, its annual mean sampling pattern has a similar structure to that of OSIRIS, HALOE, SAGE II, and SCIAMACHY but with opposite sign. The more complicated annual mean sampling bias morphologies of ACE-FTS, POAM II, POAM III, and SAGE III result from the more complicated patterns of their sampling. For example, ACE-FTS samples the highest latitudes during autumn and early spring, with a lack of samples at these latitudes through the summer months. As a result, the annual mean sampling for the highest latitudes resembles that of GOMOS, while in the midlatitudes, ACE-FTS sampling covers most of the summer months but less of the winter, leading to a sign change in the sampling bias for any pressure level moving equatorward. A similar mechanism explains the complicated structures of the POAM II, POAM III, and SAGE III annual mean sampling biases.

[33] Figure 10 details the calculated annual mean sampling biases within the UTLS for instruments with applicable measurement ranges. A common feature of many of the

instrumental sampling biases is positive biases around 30°N and S between 200 and 100 hPa. In these regions, the tropopause slopes downward with latitude, creating a strong horizontal O_3 gradient on pressure surfaces. Instrumental sampling densities tend to increase modestly with latitude. As a result, within the latitude band that straddles the tropopause for a given pressure surface, instruments tend to sample the poleward side of the latitude band slightly more often than the equatorward side, leading to positive bias in the average O_3 mixing ratio. This feature is apparent to different degrees in the UTLS sampling bias estimates for Aura MLS, HIRDLS, MIPAS, and TES. The sampling biases for ACE-FTS and OSIRIS are dominated by nonuniform month of year sampling issues as discussed above, and for SCIAMACHY, the nonuniformity of longitudinal sampling is seen to have an influence between 20 and 70°S , seemingly amplifying the magnitude of the tropopause-related bias at 30°S .

3.2. H_2O Sampling Bias

3.2.1. Monthly Mean Sampling Bias

[34] Monthly mean sampling bias results for H_2O are shown in the supporting information plots. To summarize the absolute magnitude and spatial structure of H_2O sampling bias for each instrument, Figure 11 shows the RMS of the 12 monthly sampling bias estimates for each instrument.

[35] Sampling bias for H_2O is generally smaller than for O_3 . For example, solar occultation instruments which had sampling biases of greater than 5% over large portions of the high latitudes (poleward of $\pm 50^\circ$) for O_3 have H_2O sampling biases of 2% or less over most of the NH. This is related to the weaker temporal and spatial variability of stratospheric

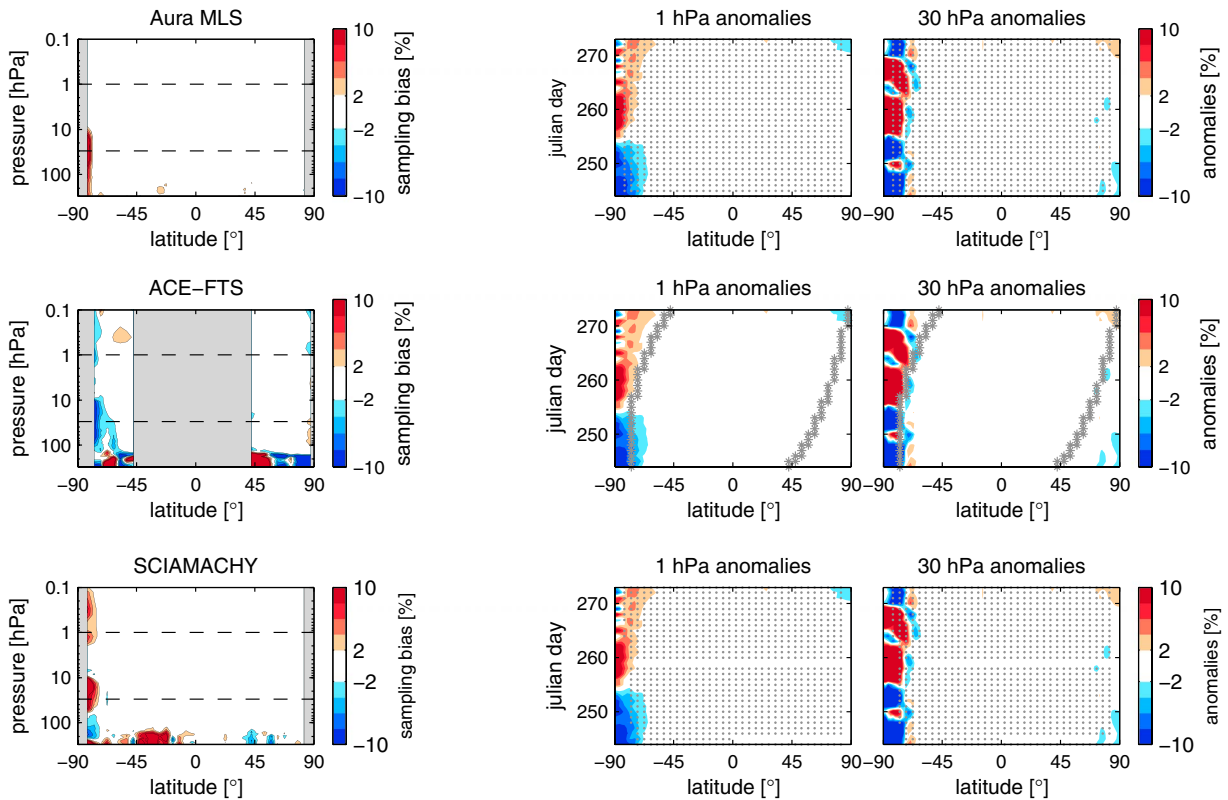


Figure 12. (left column) September H₂O percent sampling bias as function of latitude and pressure for Aura MLS, ACE-FTS, and SCIAMACHY sampling patterns. September WACCM zonal mean intramonthly percent anomalies from the monthly mean are shown as a function of latitude and Julian day for the (middle column) 1 hPa and (right column) 30 hPa surfaces, along with gray markers showing either the locations of all samples (ACE-FTS) or the locations of latitude bins which contain measurements (Aura MLS and SCIAMACHY) based on each instrument's sampling pattern.

H₂O compared to that of O₃ (see Figure 2), and in fact RMS normalized H₂O sampling biases (normalized by the H₂O standard deviation, not shown) are quite similar to those of O₃ shown in Figure 8. One location of strong H₂O sampling bias is the high SH latitudes, where dehydration of the SH winter vortex leads to strong gradients in space across the vortex edge and temporal gradients during the period of dehydration and recovery. Most solar occultation instruments (all but SAGE III, which does not measure in the high SH latitudes) show signs of the impact of vortex dehydration in their H₂O sampling biases. Larger biases are seen for the instruments with irregular temporal sampling, e.g., ACE-FTS, HALOE, POAM II, POAM III, SAGE II, SCIAMACHY, SMR, and UARS MLS.

[36] Figure 12 illustrates the impact of irregular temporal sampling in the month of September, when SH vortex H₂O values increase over the month, with short-term variability superimposed on the steady recovery in the lower stratosphere. Aura MLS sampling is uniform, and as a result, the calculated sampling bias is small (<2%, except for the highest SH latitudes, as discussed below). ACE-FTS samples the highest SH latitudes only at the beginning of the month, leading to a low bias in its monthly mean estimate, which is strongest in the lowest stratosphere (>10%) but also present in the upper stratosphere (>4%). In contrast, SCIAMACHY sampling follows the return of sunlight to the highest latitudes, so its

sampling of the highest SH latitudes occurs in the latter half of the month, leading to a positive bias (reaching values >10%) compared to the true monthly mean.

[37] Aura MLS and MIPAS, which have temporally invariant sampling patterns, also show sampling bias in the high SH latitudes. As was shown for O₃, the strong gradient in H₂O across the vortex edge and the nonuniform sampling of the southernmost latitude bin lead to bias at that latitude. Figure 13 shows SH high-latitude H₂O sampling biases for Aura MLS, MIPAS, and SCIAMACHY in September and the bias in sampled latitude. The incomplete sampling of the latitude band can lead to biases of more than 10%, due to the strong gradient in H₂O and its very low mixing ratios.

3.2.2. Annual Mean Sampling Bias

[38] Annual mean sampling biases are shown in Figure 14 for H₂O. As was the case for O₃, the annual mean sampling bias for Aura MLS and MIPAS—instruments with relatively uniform seasonal sampling—is small throughout all of the stratosphere, while sizeable sampling biases are computed for instruments with seasonally varying coverage. Many instruments (HALOE, POAM II, SAGE II, SCIAMACHY, and SMR) show a low bias in the SH high-latitude lower stratosphere (reaching values of at least -8%), resulting from the fact that these instruments sample the H₂O mixing ratio minimum in SH spring following the dehydration of

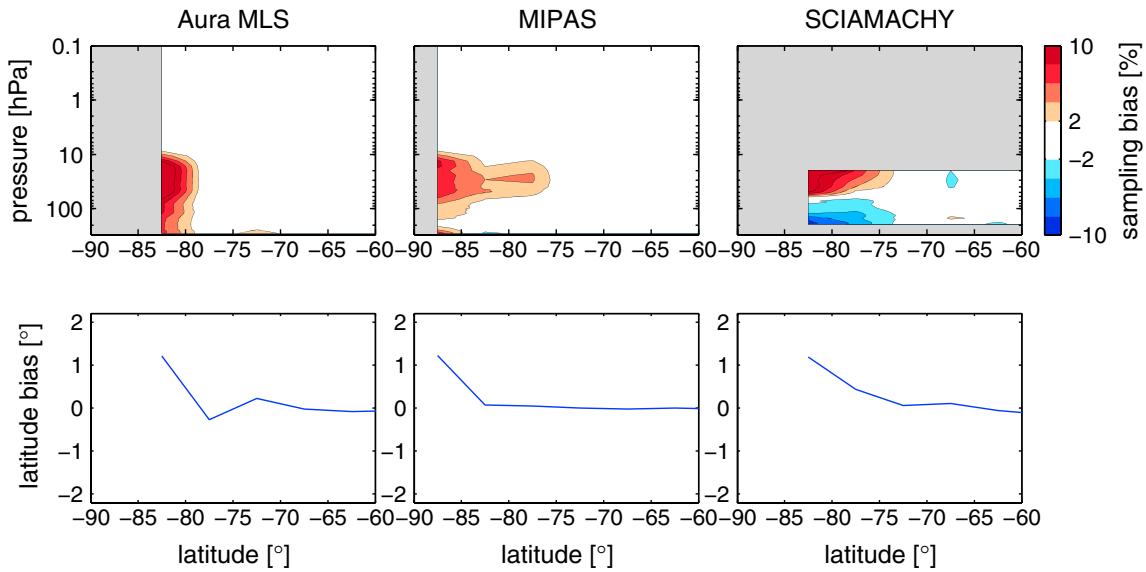


Figure 13. (top row) Sampling bias for September H₂O in SH high latitudes for selected instruments with (bottom row) differences between the latitude bin center and the mean sampled latitude.

the polar vortex but tend to undersample the maximum in H₂O mixing ratios during SH fall and winter. ACE-FTS, on the other hand, samples the highest SH latitudes only during March and April, leading to a >10% positive sampling bias there. For occultation instruments whose vertical range extends into the UTLS, sampling bias is large in the UTLS, with values >10%. It is clear that due to the significant variability of H₂O in the UTLS region, an accurate estimation of the annual mean H₂O mixing ratio requires the high

density and uniform sampling of instruments like Aura MLS and MIPAS.

4. Conclusions

[39] Monthly zonal mean climatologies produced through the method of binning and averaging measurements into months and 5° latitude bands may contain nonnegligible biases of >10% in some cases due to the nonuniform

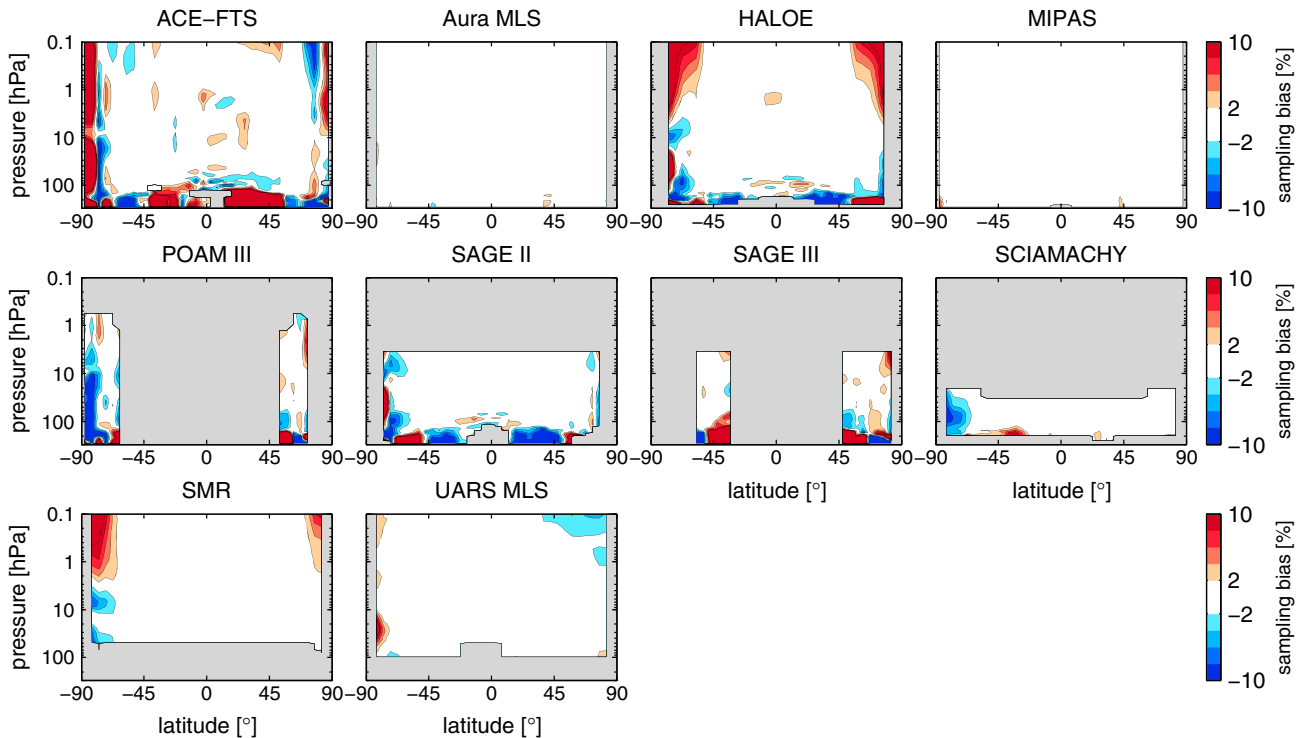


Figure 14. Latitude height sections of calculated annual mean percent sampling bias for H₂O, based on sampling patterns of instruments as labeled in each panel. Gray regions denote regions of no measurements.

temporal and spatial sampling of the atmosphere performed by the instruments. We have examined sampling biases produced by the sampling patterns of 16 instruments participating in the SPARC Data Initiative when applied to coupled chemistry-climate model output of O₃ and H₂O fields. Keeping in mind that our results are based on a single year of simulated fields from a single model, and the sampling patterns are based on single years of satellite measurements, a number of general statements regarding the impact of sampling bias on trace gas climatologies can be made. Specifically, we find that:

[40] 1. Climatologies built from measurements from instruments with regular and uniform sampling patterns have generally small sampling bias. Sampling bias may still exist, however, when the sampled latitudes within latitude bins are not uniformly distributed. For example, incomplete coverage of the northernmost and southernmost latitude bins can lead to sampling bias in those bins. This type of sampling bias appears most significantly in the high southern latitudes during SH winter, when the Antarctic vortex produces particularly strong latitudinal gradients.

[41] 2. Climatologies built from measurements from instruments whose latitudinal coverage varies with time can have strong sampling biases for certain months and locations. Sampling biases for O₃ were found in some instances to be as large as 10–40%. This is primarily due to nonuniformity in day-of-month sampling and occurs whenever an instrument provides measurements over only a portion of the month. In cases for which 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 short-term random fluctuations 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 is also important for any instrument whose sampling of certain latitudes is not uniform in time.

[42] 3. Sampling bias is a function not only of the sampling pattern but also the time-space variability of the field being averaged. Throughout most of the stratosphere, sampling bias is much more important for O₃ than for H₂O, since the variability of O₃ is stronger. A normalized sampling bias, shown in Figure 8, isolates the impact of an instrument's nonuniform sampling and can in principle be used to produce rough estimates of potential sampling bias for other long-lived trace gases if the variability of the trace gas is known.

[43] 4. In the UTLS region, space-time gradients in O₃ and H₂O (and in fact most trace gas species) are strong, and sampling bias is important. Instruments with uniform temporal sampling have monthly mean sampling biases of a few percent in the UTLS, while instruments with nonuniform temporal sampling have sampling biases of up to 10%. Due to the random nature of sampling bias in the UTLS, sampling biases are somewhat reduced in annual means. However, even for instruments with dense and uniform sampling, the strong horizontal gradients in chemical mixing ratios across the subtropical tropopause can lead to annual mean sampling biases of 1–2%.

[44] Users of monthly zonal mean climatologies of trace gases, including those of the SPARC Data Initiative, are encouraged to keep these results in mind when using such data products for chemistry-climate model validation, assessments

of observed variability, or other applications. It should be understood that in some cases, the inter-instrument spread of the climatologies is larger than the uncertainty in the true fields due to measurement errors. In such cases, the climatologies of instruments with highest sampling density should be considered more representative estimates of the truth, in the sense that differences between the climatologies and the true atmospheric mean state should be dominated by measurement error rather than sampling biases. It should also be noted that no instrumental climatology produced by binning of measurements into latitude bins is completely free of the influence of sampling bias, as even modest irregularities in sampling over latitude bins can lead to sampling biases of a few percent in regions of strong meridional gradients, such as across the polar vortices and the tropopause.

[45] **Acknowledgments.** The SPARC Data Initiative thanks the International Space Science Institute in Bern (ISSI) who supported the activity through their ISSI International Team activity program, and the World Climate Research Programme and the Toronto SPARC office for generous travel funds. The work by Matthew Toohey is supported by the BMBF MIKLIP project ALARM through the grant 01LP1130B. Michaela Hegglin thanks the Canadian Space Agency (CSA) and the European Space Agency for supporting her work for the SPARC Data Initiative. The work from Susann Tegtmeier was funded from the WGL project TransBrom and the EU project SHIVA (FP7-ENV-2007-1-226224). The work from Hampton University was partially funded under the National Oceanic and Atmospheric Administration's Educational Partnership Program Cooperative Remote Sensing Science and Technology Center (NOAA EPP CREST). Work at the Jet Propulsion Laboratory, California Institute of Technology, was performed under contract with the National Aeronautics and Space Administration. ACE is a Canadian-led mission mainly supported by the CSA. Development of the ACE-FTS climatologies was supported by grants from the Canadian Foundation for Climate and Atmospheric Sciences and the CSA. Odin is a Swedish-led satellite project funded jointly by the Swedish National Space Board (SNSB), the CSA, the National Technology Agency of Finland (Tekes), the Centre National d'Etudes Spatiales (CNES) in France, and the Third Party Mission program of ESA. The work of E. Kyrölä was supported by the Academy of Finland through the project MIDAT (134325). Work on HIRDLS was supported by NASA's EOS Program in the U.S. and in the UK by NERC. IAA was supported by the Spanish MINECO under grant AYA2011-23552 and EC FEDER funds. The work of the University of Bremen team on SCIAMACHY climatologies was funded in part by the German Aerospace Agency (DLR) within the project SADOS (50EE1105), by the DFG Research Unit FOR 1095 "Stratospheric Change and its role for Climate Prediction (SHARP)" (project GZ WE 3647/3-1), and by the State and University of Bremen. WACCM simulations were performed in the Centro de Supercomputacion de Galicia under a Reto2009 Project, and the authors thank Rolando García, who developed the WACCM model version used here, and Andrew Gettelman for guidance with the model. Finally, the authors thank Ted Shepherd, Peter Braesicke, and Karen Rosenlof for helpful feedback on early results from this work, and two anonymous reviewers who provided valuable comments on the submitted manuscript.

References

- Aghedo, A. M., K. W. Bowman, D. T. Shindell, and G. Faluvegi (2011), The impact of orbital sampling, monthly averaging and vertical resolution on climate chemistry model evaluation with satellite observations, *Atmos. Chem. Phys.*, *11*(13), 6493–6514, doi:10.5194/acp-11-6493-2011.
- Beer, R. (2006), TES on the Aura mission: Scientific objectives, measurements, and analysis overview, *IEEE Trans. Geosci. Remote Sens.*, *44*(5), 1102–1105, doi:10.1109/TGRS.2005.863716.
- Bernath, P. F., et al. (2005), Atmospheric Chemistry Experiment (ACE): Mission overview, *Geophys. Res. Lett.*, *32*, L15S01, doi:10.1029/2005GL022386.
- Bertaux, J. L., et al. (2010), Global ozone monitoring by occultation of stars: An overview of GOMOS measurements on ENVISAT, *Atmos. Chem. Phys.*, *10*(24), 12,091–12,148, doi:10.5194/acp-10-12091-2010.
- Bovensmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, S. Noël, V. V. Rozanov, K. V. Chance, and A. P. H. Goede (1999), SCIAMACHY: Mission objectives and measurement modes, *J. Atmos. Sci.*, *56*(2), 127–150, doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2.
- Cionni, I., V. Eyring, J. F. Lamarque, W. J. Randel, D. S. Stevenson, F. Wu, G. E. Bodeker, T. G. Shepherd, D. T. Shindell, and D. W. Waugh (2011), Ozone database in support of CMIP5 simulations: Results and corresponding

- radiative forcing, *Atmos. Chem. Phys.*, 11(21), 11,267–11,292, doi:10.5194/acp-11-11267-2011.
- Dupuy, E., et al. (2009), Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), *Atmos. Chem. Phys.*, 9(2), 287–343, doi:10.5194/acp-9-287-2009.
- Engelen, R. J., L. D. Fowler, P. J. Gleckler, and M. F. Wehner (2000), Sampling strategies for the comparison of climate model calculated and satellite observed brightness temperatures, *J. Geophys. Res.*, 105(D7), 9393–9406, doi:10.1029/1999JD901182.
- Fischer, H., et al. (2008), MIPAS: An instrument for atmospheric and climate research, *Atmos. Chem. Phys.*, 8(8), 2151–2188.
- Funke, B., and T. von Clarmann (2012), How to average logarithmic retrievals?, *Atmos. Meas. Tech.*, 5(4), 831–841, doi:10.5194/amt-5-831-2012.
- Garcia, R. R., D. R. Marsh, D. E. Kinnison, B. A. Boville, and F. Sassi (2007), Simulation of secular trends in the middle atmosphere, 1950–2003, *J. Geophys. Res.*, 112, D09301, doi:10.1029/2006JD007485.
- Gottelman, 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.
- Gille, J., et al. (2008), High Resolution Dynamics Limb Sounder: Experiment overview, recovery, and validation of initial temperature data, *J. Geophys. Res.*, 113, D16S43, doi:10.1029/2007JD008824.
- Glaccum, W., et al. (1996), The Polar Ozone and Aerosol Measurement instrument, *J. Geophys. Res.*, 101(D9), 14,479–14,487, doi:10.1029/96JD00576.
- Grewe, V., N. Moussiopoulos, P. Buitjes, C. Borrego, I. S. A. Isaksen, and A. Volz-Thomas (2012), The ACCENT-protocol: A framework for benchmarking and model evaluation, *Geosci. Model Dev.*, 5(3), 611–618, doi:10.5194/gmd-5-611-2012.
- Groß, J.-U., and J. M. Russell (2005), Technical note: A stratospheric climatology for O₃, H₂O, CH₄, NO_x, HCl and HF derived from HALOE measurements, *Atmos. Chem. Phys.*, 5, 2797–2807.
- Guan, B., D. E. Waliser, J.-L. F. Li, and A. da Silva (2013), Evaluating the impact of orbital sampling on satellite-climate model comparisons, *J. Geophys. Res. Atmos.*, 118, 355–369, doi:10.1029/2012JD018590.
- Hassler, B., G. E. Bodeker, I. Cionni, and M. Dameris (2009), A vertically resolved, monthly mean, ozone database from 1979 to 2100 for constraining global climate model simulations, *Int. J. Remote Sens.*, 30(15–16), 4009–4018, doi:10.1080/01431160902821874.
- Hegglin, M., and S. Tegtmeier (2011), The SPARC Data Initiative, *SPARC Newsletter*, 36, 22–23.
- 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 satellite limb sounders, *J. Geophys. Res. Atmos.*, 118, doi:10.1002/jgrd.50752.
- Jones, A., et al. (2012), Technical note: A trace gas climatology derived from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) data set, *Atmos. Chem. Phys.*, 12(11), 5207–5220, doi:10.5194/acp-12-5207-2012.
- 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.
- Lucke, R. L., et al. (1999), The Polar Ozone and Aerosol Measurement (POAM) III instrument and early validation results, *J. Geophys. Res.*, 104, 18,785–18,799, doi:10.1029/1999JD900235.
- McCormick, M. P., J. M. Zawodny, R. E. Veiga, J. C. Larsen, and P. H. Wang (1989), An overview of SAGE I and II ozone measurements, *Planet. Space Sci.*, 37(12), 1567–1586, doi:10.1016/0032-0633(89)90146-3.
- Morgenstem, O., et al. (2010), Review of the formulation of present-generation stratospheric chemistry-climate models and associated external forcings, *J. Geophys. Res.*, 115, D00M02, doi:10.1029/2009JD013728.
- Murtagh, D., et al. (2002), An overview of the Odin atmospheric mission, *Can. J. Phys.*, 80(4), 309–319, doi:10.1139/p01-157.
- Randel, W. J., and F. Wu (2007), A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data, *J. Geophys. Res.*, 112, D06313, doi:10.1029/2006JD007339.
- Russell, J. M., L. L. Gordley, J. H. Park, S. R. Drayson, W. D. Hesketh, R. J. Cicerone, A. F. Tuck, J. E. Frederick, J. E. Harries, and P. J. Crutzen (1993), The Halogen Occultation Experiment, *J. Geophys. Res.*, 98(D6), 10,777–10,797, doi:10.1029/93JD00799.
- SPARC CCMVal (2010), SPARC CCMVal Report on the evaluation of chemistry-climate models, edited by D. W. W. V. Eyring T. G. Shepherd, SPARC Report 5, WCRP-132, WMO/TD-No. 1526.
- Tegtmeier, S., et al. (2013), The SPARC Data Initiative: A comparison of ozone climatologies from international satellite limb sounders, *J. Geophys. Res. Atmos.*, doi:10.1002/2013JD019877.
- Thomason, L. W., and G. Taha (2003), SAGE III aerosol extinction measurements: Initial results, *Geophys. Res. Lett.*, 30(12), 1631, doi:10.1029/2003GL017317.
- Toohey, M., K. Strong, P. F. Bernath, C. D. Boone, K. A. Walker, A. I. Jonsson, and T. G. Shepherd (2010), Validating the reported random errors of ACE-FTS measurements, *J. Geophys. Res.*, 115(D20), D20304, doi:10.1029/2010JD014185.
- Toohey, M., and T. von Clarmann (2013), Climatologies from satellite measurements: The impact of orbital sampling on the standard error of the mean, *Atmos. Meas. Tech.*, 6(4), 937–948, doi:10.5194/amt-6-937-2013.
- von Clarmann, T., C. de Clercq, M. Ridolfi, M. Höpfner, and J.-C. Lambert (2009), The horizontal resolution of MIPAS, *Atmos. Meas. Tech.*, 2(1), 47–54.
- von Clarmann, T., B. Funke, N. Glatthor, S. Kellmann, M. Kiefer, O. Kimer, B.-M. Sinnhuber, and G. P. Stiller (2012), The MIPAS HOCl climatology, *Atmos. Chem. Phys.*, 12(4), 1965–1977, doi:10.5194/acp-12-1965-2012.
- Waters, J. W., et al. (1999), The UARS and EOS Microwave Limb Sounder (MLS) Experiments, *J. Atmos. Sci.*, 56(2), 194–218, doi: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(5), 1075–1092, doi:10.1109/TGRS.2006.873771.