Characterizing Sampling and Quality Screening Biases in Infrared and Microwave Limb Sounding

Luis F. Millán[1], Nathaniel J. Livesey[1], Michelle L. Santee[1], and Thomas von Clarmann[2]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
[2]Institut für Meteorologie und Klimaforschung, Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence to: L. Millán (luis.f.millan@jpl.nasa.gov)

Abstract. This study investigates orbital sampling biases and evaluates the additional impact caused by data quality screening for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) and the Aura Microwave Limb Sounder (MLS). MIPAS acts as a proxy for typical infrared limb emission sounders, while MLS acts as a proxy for microwave limb sounders. These biases were calculated for temperature and several trace gases by interpolating model fields to real sampling patterns and, additionally, screening those locations as directed by their corresponding quality criteria. Both instruments have dense uniform sampling patterns typical of limb emission sounders, producing almost identical sampling biases. However, there is a substantial difference between the number of locations discarded. MIPAS, as a mid-infrared instrument, is very sensitive to clouds, and measurements affected by them are thus rejected from the analysis. For example, in the tropics, the MIPAS yield is strongly affected by clouds, while MLS is mostly unaffected.

The results show that upper tropospheric sampling biases in zonally averaged data, for both instruments, can be up to 10% to 30%, depending on the species, and up to 3 K for temperature. For MIPAS, the sampling reduction due to quality screening worsens the biases, leading to values as large as 30% to 100% for the trace gases and expanding the 3 K bias region for temperature. This type of sampling bias is largely induced by the geophysical origins of the screening (e.g. clouds). Further, analysis of long-term time series reveals that these additional quality screening biases may affect the ability to accurately detect upper tropospheric long-term changes using such data. In contrast, MLS data quality screening removes sufficiently few points that no additional bias is introduced, although the vertical range of reliable measurements is slightly reduced. We emphasize that the results of this study refer only to the representativeness of the respective data, not to their intrinsic quality.

1 Introduction

Satellite limb sounders have provided a wealth of information for studies affecting climate, ozone layer stability, and air quality, as well as evaluation of reanalyses and chemistry climate models. Compared to ground-based instruments or aircraft field campaigns, satellite data provides continuous coverage over large areas (or even global scales, depending on their sampling), facilitating model evaluation on a large scale. Further, satellite missions such as the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) (Bernath et al., 2005) and the Aura Microwave Limb Sounder (MLS) (Waters et al., 2006)

have records that span more than a decade. In addition, data records constructed using several satellite instruments that span more than 3 decades (Froidevaux et al., 2015; Davis et al., 2016) provide the opportunity to study and evaluate long-term variability and trends. However, satellite observations sample the continuously changing atmosphere only at discrete locations and times, which can result in a biased depiction of the atmospheric state.

5   Several studies have evaluated the impact of orbital sampling by comparing raw model fields against satellite-sampled ones (e.g., McConnell and North, 1987; Bell and Kundu, 1996; Engelen et al., 2000; Luo et al., 2002; Brindley and Harries, 2003; Aghedo et al., 2011; Guan et al., 2013). For the limb sounding technique, Toohey et al. (2013) characterized the sampling bias for H_2O and O_3 for 16 satellite instruments, including limb scattering sounders, solar and stellar occultation instruments and limb emission sounders. They concluded that coarse non-uniform sampling leads to non-negligible biases, not only through
10   non-uniform spatial sampling but mostly through non-uniform temporal sampling, that is, producing means using measurements that span less than the full period in question. Millán et al. (2016) studied the sampling bias for temperature and several trace gas species for a subset of the instruments used by Toohey et al. (2013) and investigated the impact of such biases upon stratospheric trend detection. They found that coarse non-uniform sampling patterns can induce significant errors in the magnitudes of inferred trends, necessitating analysis of considerably more years of data to conclusively detect a trend. In contrast,
15   dense uniform sampling patterns accurately reproduce the magnitude of the trends, with the number of years of data required determined mostly by natural variability.

However, none of these studies have quantified the additional biases introduced through quality screening of the measurements. Many of the measurements discarded through quality screening have been affected by the presence clouds, which pose a substantial challenge to limb observations as the long limb path traverses hundreds of kilometers. The impact of such cloudy
20   scenes depends on the measurement technique used. For example, instruments measuring microwave emission are unaffected by all but the largest particles in the thickest clouds. Many other limb measurements are screened out because of temperature gradients near the poles, whose impact varies depending on the retrieval scheme, i.e., one dimensional versus tomographic, as well as how accurately the a priori or initial guess captures such gradients.

This study examines the sampling bias and quantifies the impact of quality screening upon two limb viewing instruments,
25   one using microwave emission (the Aura Microwave Limb Sounder - MLS) and the other one using infrared emission (the ENVISAT Michelson Interferometer for Passive Atmospheric Sounding - MIPAS). Both instruments have dense uniform sampling distributions, which should minimize the sampling biases; however, there is a substantial difference in the number of measurements rejected through quality screening for these techniques (see Figure 1). The discarded profiles tend to cluster geophysically, leading to biases in analyses that are based on the remaining measurements. We emphasize that the results of
30   this study refer only to the representativeness of the respective data, not to their intrinsic quality.

2   Data and Methodology

2.1   Model Fields

CMAM30-SD is a coupled chemistry climate model nudged to the winds and temperatures of the ERA-Interim reanalysis. This nudging exploits the much better dynamics of the reanalysis to reliably predict the chemical fields. More information can be found in Scinocca et al. (2008), de Grandpré et al. (2000) and McLandress et al. (2014). Extensive validation (de Grandpré et al., 2000; Hegglin and Shepherd, 2007; Melo et al., 2008; Jin et al., 2005, 2009) has shown that the free-running version of this model performs well against observations relevant to dynamics, transport, and chemistry. Comparisons against ACE-FTS and the Odin Optical Spectrograph and Infrared Imaging System (OSIRIS) have shown that CMAM30-SD has a good representation of stratospheric temperature, H_2O, O_3, and CH_4 in polar regions (Pendlebury et al., 2015). Further, CMAM30-SD has been used to construct a long-term H_2O record, acting as a transfer function between satellite observations (Hegglin et al., 2014), and it reproduces halogen-induced midlatitude O_3 depletion sufficiently well to be used in long-term O_3 trend studies (Shepherd et al., 2014).

The CMAM30-SD version used in this study has a horizontal resolution of approximately 3.75°, that is, approximately 400 km (similar to the ∼500 km limb viewing path length). It has a lid at 0.0007 hPa with 63 vertical levels that vary from ∼500 m in the lower troposphere to ∼3 km in the mesosphere. Here, we present results using the H_2O, O_3, CO, HNO_3, and temperature CMAM30-SD fields. Note that for this study it is not necessary for the model fields to be correct in absolute terms. CMAM30-SD is simply used as a representative evolving atmospheric state.

2.2   Satellite Instruments

We analyze the impact of sampling and quality screening of the limb emission sounders MIPAS and MLS. MIPAS (Fischer et al., 2000, 2008) was launched in March 2002 on the European Space Agency Environment Satellite. MIPAS was a Fourier transform spectrometer conceived to record limb emission spectra. It covered the mid-infrared region from 685 to 2410 cm^{-1} in five spectral bands, allowing retrievals of temperature, pressure and trace gases. MIPAS measured around 1350 vertical scans daily, providing global observations.

From July 2002 to March 2004, MIPAS operated in full resolution mode, with a spectral spacing of 0.025 cm^{-1}; however, following persistent malfunctions with the interferometer slide mechanism, instrument operations were temporarily suspended. In January 2005 operations were resumed with MIPAS operating at a spectral spacing of 0.0625 cm^{-1}. This mode of operation is known as optimum resolution, and it is characterized by finer vertical and horizontal sampling attained through the degraded spectral spacing. MIPAS took quasi-continuous measurements until April 2012, when the European Space Agency lost contact with ENVISAT.

In the optimum resolution operation, MIPAS has several measurement modes: the nominal mode, with 27 tangent heights from 6 to 70 km; the middle atmosphere mode, with 29 tangent heights from 18 to 102 km; and the upper atmosphere mode, with 35 tangent heights from 42 to 172 km. The nominal mode covers the entire stratosphere extending into both the upper

troposphere and the lower mesosphere to study linkages between these atmospheric layers. In this study we use the geolocations of this measurement mode because it covers around 80% of the measurement time (Fischer et al., 2008).

Several retrieval algorithms have been developed for the MIPAS spectra (e.g. Ridolfi et al., 2000; von Clarmann et al., 2003; Hoffmann et al., 2005; Carlotti et al., 2006; von Clarmann et al., 2009; Dudhia, 2017). Here we use the profiles generated by

5 the Institute of Meteorology and Climate Research (IMK) in cooperation with the Instituto de Astrofísica de Andalucía (von Clarmann et al., 2009). This retrieval algorithm uses a Tikhonov regularization; it is capable of handling deviations from local thermodynamic equilibrium, and it includes temperature horizontal gradients along the line of sight to prevent many retrievals from failing to converge, particularly near the boundary of the poles.

MLS (Waters et al., 2006) was launched in July 2004 on the Aura spacecraft. MLS measures limb millimetre and sub-

10 millimetre atmospheric thermal emission in spectral regions near 118, 191, 240, and 640 GHz, and 2.5 THz. These radiances are inverted using a tomographic optimal estimation algorithm (Livesey et al., 2006) that allows the retrieval of temperature, composition and ice cloud properties. MLS measures around 3500 vertical scans daily, providing near-global (82°S to 82°N) observations.

To investigate the impact of sampling and quality screening, daily CMAM30-SD model fields were linearly interpolated

15 to the actual latitude and longitude of the satellite measurements. For the MIPAS sampling pattern, for each calendar day of the year we identify the year with the most measurements obtained on that date. That is to say, for the 1st of January, we use the locations of the 1st of January for the year with the most measurement locations, etc. This allows us to have a complete year of MIPAS measurements without interruptions due to MIPAS changing measurement modes. For MLS, we use 2008 as a representative year. To avoid differences attributed to diurnal cycles, all satellite measurements were assumed to be made

20 at 12:00 UT on a given day, avoiding any interpolation in time. Further, we used the vertical grid of the CMAM30-SD fields; that is, we assume that MIPAS and MLS vertical resolution is good enough to resolve these model fields, at least in the upper troposphere / lower stratosphere (UTLS).

We constructed three time series: one using the raw CMAM30-SD fields; another using all the measurement locations available; and lastly one using only the measurement locations remaining after the quality screening recommended for each

25 instrument was applied, in other words, after those points flagged as bad values in the actual data were eliminated. The screening procedure applied to MIPAS data is as follows: We neglect profile points where the diagonal element of the averaging kernel is less than 0.03 — to avoid retrievals influenced by the a priori — and discard points where the visibility flag was set to zero — which indicates that MIPAS has not seen the atmosphere at those particular altitudes. The screening procedure applied to MLS data follows the guidelines detailed by Livesey et al. (2017), which vary product by product.

30 Figure 1 shows typical daily MIPAS and MLS geolocations overlaid on top of a modeled water vapor map. Both instruments have dense coverage that, as noted by Toohey et al. (2013), is relatively uniform with latitude and time. Figure 1 also displays those geolocations for which the retrieved values are not recommended for scientific studies, that is, they are screened out by the quality criteria. As shown, these failed or in many cases skipped retrievals cluster in the tropics or near the poles. Overall, in the tropics, these missing retrievals are due to clouds. Near the poles, the retrieval failures are presumably due to temperature

horizontal gradients and, in the case of MIPAS, also due to the presence of polar stratospheric clouds. The substantial difference between the number of failed/missing retrievals in the tropics for MIPAS and MLS is the main motivation for this study.

To quantify this further, Figure 2 displays the yield given by,

$$Y = \frac{N_{QS}}{N_A} \tag{1}$$

5 where N_A is the number of measurements available and N_{QS} is the number of measurements left after applying the quality screening criteria at each latitude and each pressure level. Again, MIPAS low yield values accumulate in the tropics and near the poles. Overall, MIPAS yields drop below 60% near the South Pole at pressures greater than $\sim 20\,\mathrm{hPa}$ and below 30% in the tropics at pressures greater than $100\,\mathrm{hPa}$. In contrast, in general MLS yield values are better than 90%. The two exceptions are the yield values for H_2O near the South Pole, which drop below 90%, and HNO_3 near the equator, which drop to 60%. Note

10 that MIPAS yield values drop below 10% well into the troposphere.

3   Induced Sampling and Quality Screening Biases

Following Millán et al. (2016), we evaluate the sampling biases as well as the quality screening biases associated with constructing monthly zonal means using the raw CMAM30-SD fields, Z_R, versus those using the satellite-sampled measurements, Z_A, or only those passing the quality screening criteria, Z_{QS}. The difference between Z_A or Z_{QS} and Z_R gives the sampling

15 or the quality screening induced bias, respectively. For each instrument and for each month throughout one year, we computed these biases as a function of latitude and pressure. Note that the quality screening bias is the sampling bias plus the additional impact of screening out more locations and, hence, reducing the sampling frequency.

Figure 3 shows examples of the sampling and screening biases for June 2005. Percentage biases are shown for the trace gases to cope with their large vertical variability. MIPAS and MLS sampling biases are practically identical. For the trace gases,

20 sampling biases are larger in the upper troposphere, where the variability is larger, while the temperature sampling biases are larger near the edges of the polar regions, where there are substantial temperature gradients. The impact of the MIPAS quality screening is evident in the tropics (in particular near $20°N$), where the yield values are expected to be greatly affected (see Figure 2) by clouds. In this region, on top of the sampling biases, all parameters studied display an underestimation, for example, up to -50% for H_2O. Although this resembles the expected dry bias in clear-sky tropospheric infrared measurements

25 (e.g., Sohn et al., 2006; Yue et al., 2013) — that is, the fact that infrared instruments cannot measure cloudy regions where H_2O is high, resulting in a dry bias — the biases shown here are mainly due to the reduced sampling frequencies rather than high H_2O values associated with deep convection. Note that this is also applicable to other parameters; that is, the quality screening biases shown here are not an indication of trace gas (or temperature) / deep convection relationships. In contrast to those of MIPAS, except for the reduced vertical ranges, MLS sampling biases are unaffected by data quality screening.

30 To summarize the potential sampling and quality screening biases, Figure 4 shows their root-mean-square (RMS) computed over one year's worth of data. Again, the MIPAS and MLS RMS sampling biases are almost identical: H_2O displays a bias of up to 30% at pressures greater than $\sim 150\,\mathrm{hPa}$; CO, O_3 and HNO_3 show biases (up to 30% for O_3 and HNO_3) near mid-latitudes

(around 40°S and 40°N), where there are sharp trace gas gradients and variability due to tropopause folding; and temperature displays a bias as large as 3 K near the polar edges.

The impact of MIPAS quality screening is especially evident in H_2O and HNO_3, which have potential biases as large as 100%, but quality screening also affects the rest of the parameters: CO and O_3 biases approach 30% in the tropics, while the region with 3 K temperature bias expands near the South Pole. As before, except for the reduced vertical ranges, the impact of the MLS quality screening is negligible; that is, the screening biases are almost identical to the sampling biases.

To exemplify the impact of these quality screening biases, Figure 5 (left) shows time series (1979-2012) of 20°S to 20°N H_2O at 200 hPa using the raw CMAM30-SD fields, the full satellite-sampled fields and only those points passing the screening criteria. All time series show the expected features, with an annual cycle related to the seasonality of the cold point tropopause temperature. The MIPAS time series constructed using the full satellite-sampled fields is almost identical to the one constructed using the raw CMAM30-SD field. However, as suggested by the screening bias shown in Figure 4, the MIPAS time series using the quality screening displays a substantial dry bias. In contrast, no evidence of such a bias is seen in the MLS time series; that is, both the time series constructed using the full satellite-sampled field and that based on only those points passing the screening criteria are almost identical to the CMAM30-SD one.

Figure 5 (right) shows the area-weighted scatter between these time series. MIPAS sampling scatter, that is, the scatter between MIPAS when using all available measurements and the raw CMAM30-SD fields, is small and their correlation tight, with a bias better than -1.5%, a slope of ∼1.05, and a coefficient of determination of 0.98. The contrast with the MIPAS screened scatter is dramatic in this particular latitude/pressure region; it displays considerably more scatter and, as in the time series (Figure 5-left), a discernible bias. Quantitatively, MIPAS screened data displays a bias of 16.13%, a slope of 1.32, and a coefficient of determination of ∼0.8 (which implies that 20% of the total variation cannot be explained). MLS sampling and screened scatterplots are almost the same.

To explore this further, Figure 6 shows these metrics versus pressure using different latitude bands for the MIPAS sampling scatter. As shown, the coefficients of determination as well as the slopes are close to one and the biases close to zero in most cases. The most notable exceptions are the biases between 20° and 45° (either north or south) for O_3 and HNO_3, which can be up to -10%. In these regions, Figure 3 and Figure 4 indicate biases due to the sharp trace gas gradients associated with tropopause folding. Note that both the MLS sampling and the MLS screened scatter are almost identical to the MIPAS sampling scatter and, hence, are not shown.

The MIPAS screened scatter results are shown in Figure 7. The largest impact can be found in the tropics (the 20°S-20°N latitude band) at pressures greater than 100 hPa. Here, the coefficients of determination, the biases and the slopes are severely degraded. The coefficients of determination rapidly decrease, especially for H_2O, O_3 and HNO_3, whose values are as low as 0.5 at 200 hPa and worsen further at lower altitudes. The biases for O_3 and HNO_3 oscillate between -10% and 10% and can be as large as 40% for H_2O. Lastly, all the slopes vary from 0.5 to 1.5, depending on pressure level. These poor metrics imply that any trends derived at these pressure levels will also be impacted by quality screening induced biases: the magnitude of the trends will be affected because of the change in the slope, and the number of years of observations required to conclusively

detect trends will considerably increase due to the noise associated with the worsening of the coefficients of determination (e.g., Millán et al., 2016).

4 Summary and Conclusions

This study explored the implications of sampling in the UTLS for two satellite instruments, MIPAS and MLS, for H_2O, O_3,
5   CO, HNO_3, and temperature. We quantify sampling biases by interpolating CMAM30-SD fields, used as a proxy for the atmospheric state, to the measurement locations and computing monthly means. Both of these instruments have dense uniform sampling, with around 1350 points spread globally for MIPAS and around 3500 spread from 82°S to 82°N for MLS, resulting in almost identical sampling biases for the two instruments. For the trace gases, the largest sampling biases are found in the upper troposphere, where there is more natural variability: H_2O displays a bias of up to 30%, while CO, O_3 and HNO_3 show
10   biases near mid-latitudes of up to 10% for CO or 30% for O_3 and HNO_3 due to sharp trace gas gradients and variability arising from tropopause folding. The temperature sampling bias is negligible (less than 1 K), except near the polar edges, where the bias can be as large as 3 K, presumably due to horizontal temperature gradients.

Besides the orbital sampling biases, this study also evaluated the impact of quality screening, which further reduces the sampling frequency. In the tropics (see Figure 2), MIPAS is substantially impacted by clouds, as they act as grey bodies with
15   high opacity, greatly altering the radiances below the cloud top. Cloud effects are evident, with H_2O and HNO_3 biases up to 100% and CO and O_3 biases up to 30%. In contrast, because of their longer wavelengths, MLS measurements are unaffected by all but the thickest clouds, negligibly impacting the sampling frequency. However, pressure broadening washes out the microwave signal in the troposphere, reducing the MLS vertical range.

Analysis of scatterplots of time series constructed using the raw model fields versus those using all the available measurement
20   locations (either for MIPAS or MLS) reveal that at most pressure levels and most latitude bands, the coefficient of determination and the slope of the fits are close to one, while the biases are close to zero. However, when only those measurements passing the screening criteria are used, MIPAS upper tropospheric measurements are severely impacted in some regions. In the tropics, the coefficients of determination rapidly decrease, especially for H_2O, O_3 and HNO_3, from ∼1 at 100 hPa to as low as 0.5 at 200 hPa, and they worsen further at lower altitudes. The biases for O_3 and HNO_3 oscillate between -10% and 10% and can be
25   as large as 40% for H_2O. Lastly, all the slopes vary from 0.5 to 1.5, depending on pressure level. These results imply that any trends derived from measurements made using techniques substantially affected by clouds will be biased and that the number of years required to conclusively detect trends from such data will be considerably larger than that for data records less sensitive to clouds. While this sampling bias can be mitigated for model validation studies by also rejecting cloud-affected data from the model sample, thus far no straightforward solution to this problem has been identified in the context of trend quantification.
30   Note that although these results were derived for MIPAS, they are applicable to other instruments with dense sampling but for which quality screening (e.g., for clouds) severely impacts their yield. For example, a UV-visible instrument such as the Ozone Mapping and Profiler Suite (Jaross et al., 2014) or a constellation of solar occultation instruments would suffer from similar substantial cloud-screening-induced data gaps.

5 Data availability

The datasets used in this study are publicly available: CMAM30-SD fields can be found in the Canadian Centre for Climate
Modeling and Analysis webpage (http://www.cccma.ec.gc.ca/data/cmam/output/CMAM/CMAM30-SD/index.shtml), MLS data
can be found in the NASA Goddard Space Flight Center Earth Sciences Data and Information Services Center (http://disc.sci.gsfc.nasa.gov/
5   holdings/MLS/index.shtml), and MIPAS data can be found in the Karlsruhe Institute of Technology webpage (https://www.imk-
asf.kit.edu/english/308.php).

Acknowledgements. Work at the Jet Propulsion Laboratory, California Institute of Technology, was done under contract with the National
Aeronautics and Space Administration. We thank David Plummer of Environment Canada for his assistance in obtaining the CMAM30-SD
dataset. Government sponsorship acknowledged.

References

Aghedo, A. M., Bowman, K. W., Shindell, D. T., and Faluvegi, G.: The impact of orbital sampling, monthly averaging and vertical resolution on climate chemistry model evaluation with satellite observations, Atmospheric Chemistry and Physics, 11, 6493–6514, doi:10.5194/acp-11-6493-2011, 2011.

5   Bell, T. L. and Kundu, P. K.: A Study of the Sampling Error in Satellite Rainfall Estimates Using Optimal Averaging of Data and a Stochastic Model, Journal of Climate, 9, 1251–1268, doi:10.1175/1520-0442(1996)009<1251:asotse>2.0.co;2, 1996.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F.,

10   Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, Geophysical Research Letters, 32, doi:10.1029/2005gl022386, 2005.

Brindley, H. E. and Harries, J. E.: Observations of the Infrared Outgoing Spectrum of the Earth from Space: The Effects of Temporal and Spatial Sampling, Journal of Climate, 16, 3820–3833, doi:10.1175/1520-0442(2003)016<3820:ootios>2.0.co;2, 2003.

15   Carlotti, M., Brizzi, G., Papandrea, E., Prevedelli, M., Ridolfi, M., Dinelli, B. M., and Magnani, L.: GMTR: Two-dimensional geo-fit multitarget retrieval model for Michelson Interferometer for Passive Atmospheric Sounding/Environmental Satellite observations, Applied Optics, 45, 716, doi:10.1364/ao.45.000716, 2006.

Davis, S. M., Rosenlof, K. H.and Hassler, B., Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H., Fujiwara, M., and Damadeo, R.: The Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) database: a long-term database for climate studies, Earth System

20   Science Data, 8, 461–490, doi:10.5194/essd-8-461-2016, 2016.

de Grandpré, J., Beagley, S. R., Fomichev, V. I., Griffioen, E., McConnell, J. C., Medvedev, A. S., and Shepherd, T. G.: Ozone climatology using interactive chemistry: Results from the Canadian Middle Atmosphere Model, Journal of Geophysical Research: Atmospheres, 105, 26 475–26 491, doi:10.1029/2000JD900427, 2000.

Dudhia, A.: MIPAS Orbital Retrieval using Sequential Estimation, http://eodg.atm.ox.ac.uk/MORSE/, 2017.

25   Engelen, R. J., Fowler, L. D., Gleckler, P. J., and Wehner, M. F.: Sampling strategies for the comparison of climate model calculated and satellite observed brightness temperatures, Journal of Geophysical Research: Atmospheres, 105, 9393–9406, doi:10.1029/1999jd901182, 2000.

Fischer, H., Blom, C., Oelhaf, H., Carli, B., Carlotti, M., Delbouille, L., Ehhalt, D., Flaud, J.-M., Isaksen, I., López-Puertas, M., McElroy, C. T., and Zander, R.: ENVISAT-MIPAS, the Michelson Interferometer for Passive Atmospheric Sounding; An instrument for atmospheric

30   chemistry and Climate Research, ESA SP-1229 Readings and R. A. Harris eds European Space Agency, Noordwijk, The Netherlands, 2000.

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmospheric Chemistry and Physics,

35   8, 2151–2188, doi:10.5194/acp-8-2151-2008, http://www.atmos-chem-phys.net/8/2151/2008/, 2008.

Froidevaux, L., Anderson, J., Wang, H.-J., Fuller, R. A., Schwartz, M. J., Santee, M. L., Livesey, N. J., Pumphrey, H. C., Bernath, P. F., Russell, J. M., and McCormick, M. P.: Global OZone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS): method-

ology and sample results with a focus on HCl, H_2O, and O_3, Atmospheric Chemistry and Physics, 15, 10 471–10 507, doi:10.5194/acp-15-10471-2015, 2015.

Guan, B., Waliser, D. E., Li, J.-L. F., and da Silva, A.: Evaluating the impact of orbital sampling on satellite-climate model comparisons, Journal of Geophysical Research: Atmospheres, 118, 355–369, doi:10.1029/2012jd018590, 2013.

5   Hegglin, M. I. and Shepherd, T. G.: O3-N2O correlations from the Atmospheric Chemistry Experiment: Revisiting a diagnostic of transport and chemistry in the stratosphere, Journal of Geophysical Research: Atmospheres, 112, doi:10.1029/2006JD008281, d19301, 2007.

Hegglin, M. I., Plummer, D. A., Shepherd, T. G., Scinocca, J. F., Anderson, J., Froidevaux, L., Funke, B., Hurst, D. amd Rozanov, A., Urban, J., von Clarmann, T., Walker, K. A., Wang, H. J., Tegtmeier, S., and Weigel, K.: Vertical structure of stratospheric water vapour trends derived from merged satellite data, Nat. Geosci., 7, 768–776, doi:10.1038/ngeo2236, 2014.

10   Hoffmann, L., Spang, R., Kaufmann, M., and Riese, M.: Retrieval of CFC-11 and CFC-12 from Envisat MIPAS observations by means of rapid radiative transfer calculations, Advances in Space Research, 36, 915–921, doi:10.1016/j.asr.2005.03.112, 2005.

Jaross, G., Bhartia, P. K., Chen, G., Kowitt, M., Haken, M., Chen, Z., Xu, P., Warner, J., and Kelly, T.: OMPS Limb Profiler instrument performance assessment, Journal of Geophysical Research: Atmospheres, 119, 4399–4412, doi:10.1002/2013jd020482, https://doi.org/10.1002/2013jd020482, 2014.

15   Jin, J. J., Semeniuk, K., Jonsson, A. I., Beagley, S. R., McConnell, J. C., Boone, C. D., Walker, K. A., Bernath, P. F., Rinsland, C. P., Dupuy, E., Ricaud, P., De La Noë, J., Urban, J., and Murtagh, D.: Co-located ACE-FTS and Odin/SMR stratospheric-mesospheric CO 2004 measurements and comparison with a GCM, Geophysical Research Letters, 32, doi:10.1029/2005GL022433, l15S03, 2005.

Jin, J. J., Semeniuk, K., Beagley, S. R., Fomichev, V. I., Jonsson, A. I., McConnell, J. C., Urban, J., Murtagh, D., Manney, G. L., Boone, C. D., Bernath, P. F., Walker, K. A., Barret, B., Ricaud, P., and Dupuy, E.: Comparison of CMAM simulations of carbon monoxide (CO),

20   nitrous oxide (N_2O), and methane (CH_4) with observations from Odin/SMR, ACE-FTS, and Aura/MLS, Atmospheric Chemistry and Physics, 9, 3233–3252, doi:10.5194/acp-9-3233-2009, 2009.

Livesey, N. J., Snyder, W. V., Read, W. G., and Wagner, P. A.: Retrieval algorithms for the EOS Microwave limb sounder (MLS), IEEE Transactions on Geoscience and Remote Sensing, 44, 1144–1155, doi:10.1109/TGRS.2006.872327, 2006.

Livesey, N. J., Read, W., Wagner, P. A. Froidevaux, L., Lambert, A., Manney, G. L., Millán Valle, L., Pumphrey, H. C., Santee, M. L.,

25   Schwartz, M. J., Wang, S., Fuller, R. A., Jarnot, R. F., Knosp, B. W., and E., M.: Version 4.2x Level 2 data quality and description document, JPL D-33509 Rev. C, availablefromhttp://mls.jpl.nasa.gov, 2017.

Luo, M., Beer, R., Jacob, D. J., Logan, J. A., and Rodgers, C. D.: Simulated observation of tropospheric ozone and CO with the Tropospheric Emission Spectrometer (TES) satellite instrument, Journal of Geophysical Research, 107, doi:10.1029/2001jd000804, 2002.

McConnell, A. and North, G. R.: Sampling errors in satellite estimates of tropical rain, Journal of Geophysical Research, 92, 9567,

30   doi:10.1029/jd092id08p09567, 1987.

McLandress, C., Plummer, D. A., and Shepherd, T. G.: Technical Note: A simple procedure for removing temporal discontinuities in ERA-Interim upper stratospheric temperatures for use in nudged chemistry-climate model simulations, Atmospheric Chemistry and Physics, 14, 1547–1555, doi:10.5194/acp-14-1547-2014, 2014.

Melo, S. M. L., Blatherwick, R., Davies, J., Fogal, P., de Grandpré, J., McConnell, J., McElroy, C. T., McLandress, C., Murcray, F. J., Olson,

35   J. R., Semeniuk, K., Shepherd, T. G., Strong, K., Tarasick, D., and Williams-Rioux, B. J.: Summertime stratospheric processes at northern mid-latitudes: comparisons between MANTRA balloon measurements and the Canadian Middle Atmosphere Model, Atmospheric Chemistry and Physics, 8, 2057–2071, doi:10.5194/acp-8-2057-2008, 2008.

Millán, L. F., Livesey, N. J., Santee, M. L., Neu, J. L., Manney, G. L., and Fuller, R. A.: Case studies of the impact of orbital sampling on stratospheric trend detection and derivation of tropical vertical velocities: solar occultation vs. limb emission sounding, Atmospheric Chemistry and Physics, 16, 11 521–11 534, doi:10.5194/acp-16-11521-2016, 2016.

Pendlebury, D., Plummer, D., Scinocca, J., Sheese, P., Strong, K., Walker, K., and Degenstein, D.: Comparison of the CMAM30 data set with ACE-FTS and OSIRIS: polar regions, Atmospheric Chemistry and Physics, 15, 12 465–12 485, doi:10.5194/acp-15-12465-2015, 2015.

Ridolfi, M., Carli, B., Carlotti, M., von Clarmann, T., Dinelli, B. M., Dudhia, A., Flaud, J-M.and Höpfner, M., Morris, P. E., Raspollini, P., Stiller, G., and Wells, R. J.: Optimized forward model and retrieval scheme for MIPAS near-real-time data processing, Applied Optics, 39, 1323, doi:10.1364/ao.39.001323, 2000.

Scinocca, J. F., McFarlane, N. A., Lazare, M., Li, J., and Plummer, D.: Technical Note: The CCCma third generation AGCM and its extension into the middle atmosphere, Atmospheric Chemistry and Physics, 8, 7055–7074, doi:10.5194/acp-8-7055-2008, 2008.

Shepherd, T. G., Plummer, D. A., Scinocca, J. F., Hegglin, M. I., Fioletov, V. E., Reader, M. C., Remsberg, E., von Clarmann, T., and Wang, H. J.: Reconciliation of halogen-induced ozone loss with the total-column ozone record, Nat. Geosci., 7, 443–449, doi:10.1038/ngeo2155, 2014.

Sohn, B.-J., Schmetz, J., Stuhlmann, R., and Lee, J.-Y.: Dry Bias in Satellite-Derived Clear-Sky Water Vapor and Its Contribution to Long-wave Cloud Radiative Forcing, Journal of Climate, 19, 5570–5580, doi:10.1175/jcli3948.1, 2006.

Toohey, M., Hegglin, M. I., Tegtmeier, S., Anderson, J., Añel, J. A., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Neu, J. L., Rozanov, A., Smith, L., Urban, J., von Clarmann, T., Walker, K. A., and Wang, R. H. J.: Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative, Journal of Geophysical Research: Atmospheres, 118, 11,847–11,862, doi:10.1002/jgrd.50874, 2013JD020298, 2013.

von Clarmann, T., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Tsidu, G. M., Milz, M., Steck, T., Stiller, G. P., Wang, D. Y., Fischer, H., Funke, B., Gil-López, S., and López-Puertas, M.: Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2003JD003602, 4736, 2003.

von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: 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, Atmospheric Measurement Techniques, 2, 159–175, doi:10.5194/amt-2-159-2009, http://www.atmos-meas-tech.net/2/159/2009/, 2009.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Snyder, W. V., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1075–1092, doi:10.1109/TGRS.2006.873771, 2006.

Yue, Q., Fetzer, E. J., Kahn, B. H., Wong, S., Manipon, G., Guillaume, A., and Wilson, B.: Cloud-State-Dependent Sampling in AIRS Observations Based on CloudSat Cloud Classification, Journal of Climate, 26, 8357–8377, doi:10.1175/jcli-d-13-00065.1, 2013.
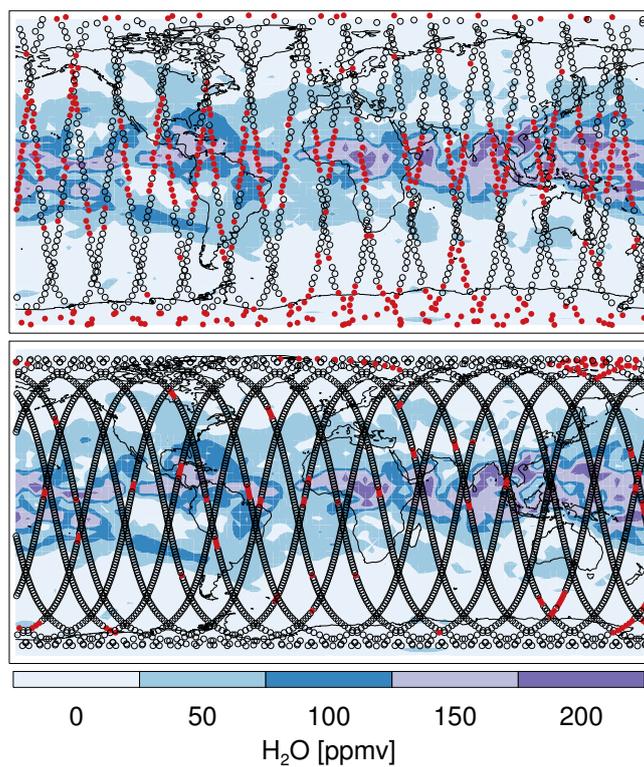
Figure 1. Typical MIPAS (top) and MLS (bottom) sampling overlaid on top of a modeled water vapor map (June 1st, 2005) at 200 hPa. Red dots show missed or failed retrievals: in the tropics, these are mostly due to clouds.
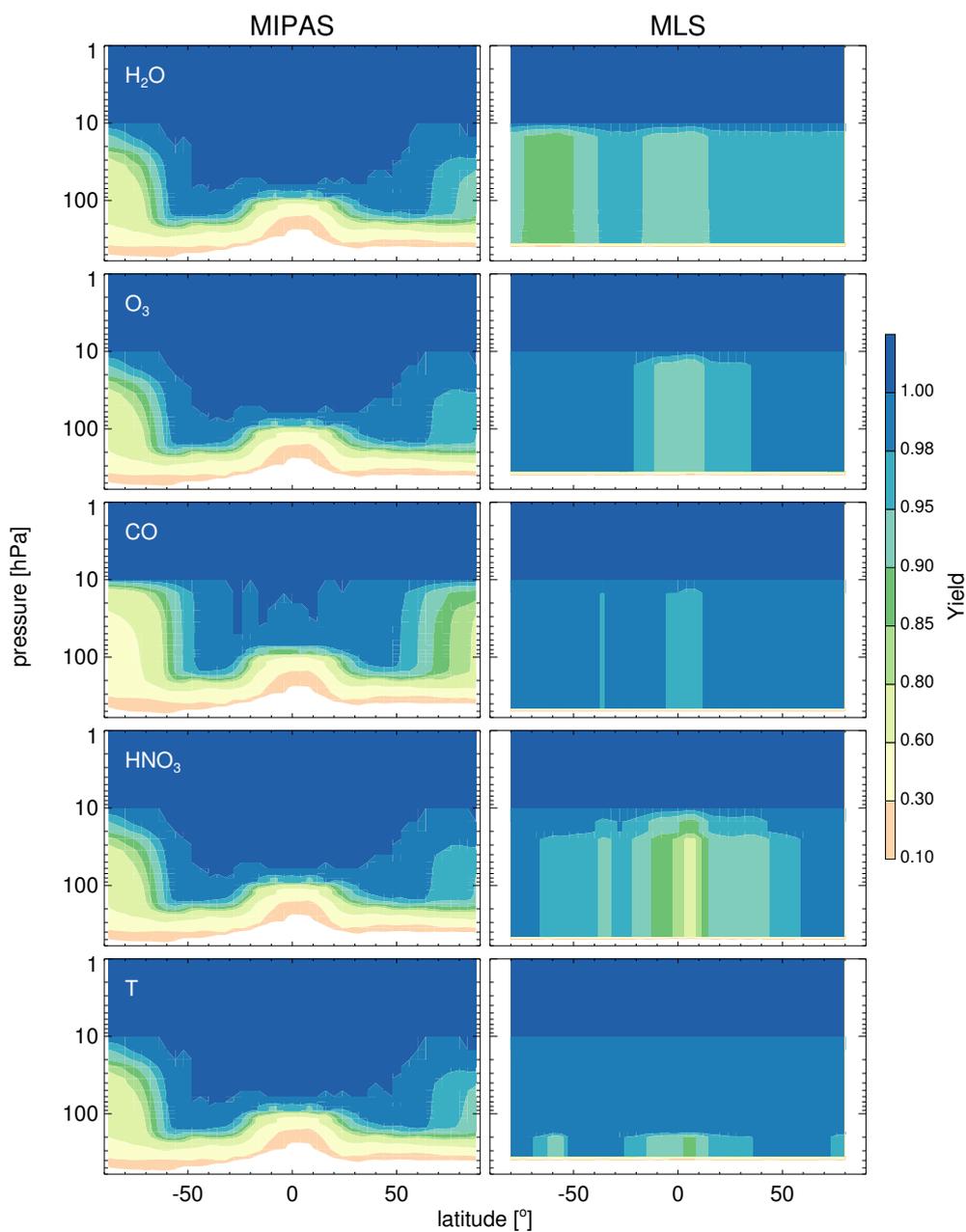
Figure 2. MIPAS and MLS zonal mean yield (see text) for H_2O, O_3, CO, HNO_3 and temperature for 2005, that is, sampling the modeled 2005 year with the sampling patterns as explained in the text. Note the non-linear color scale.
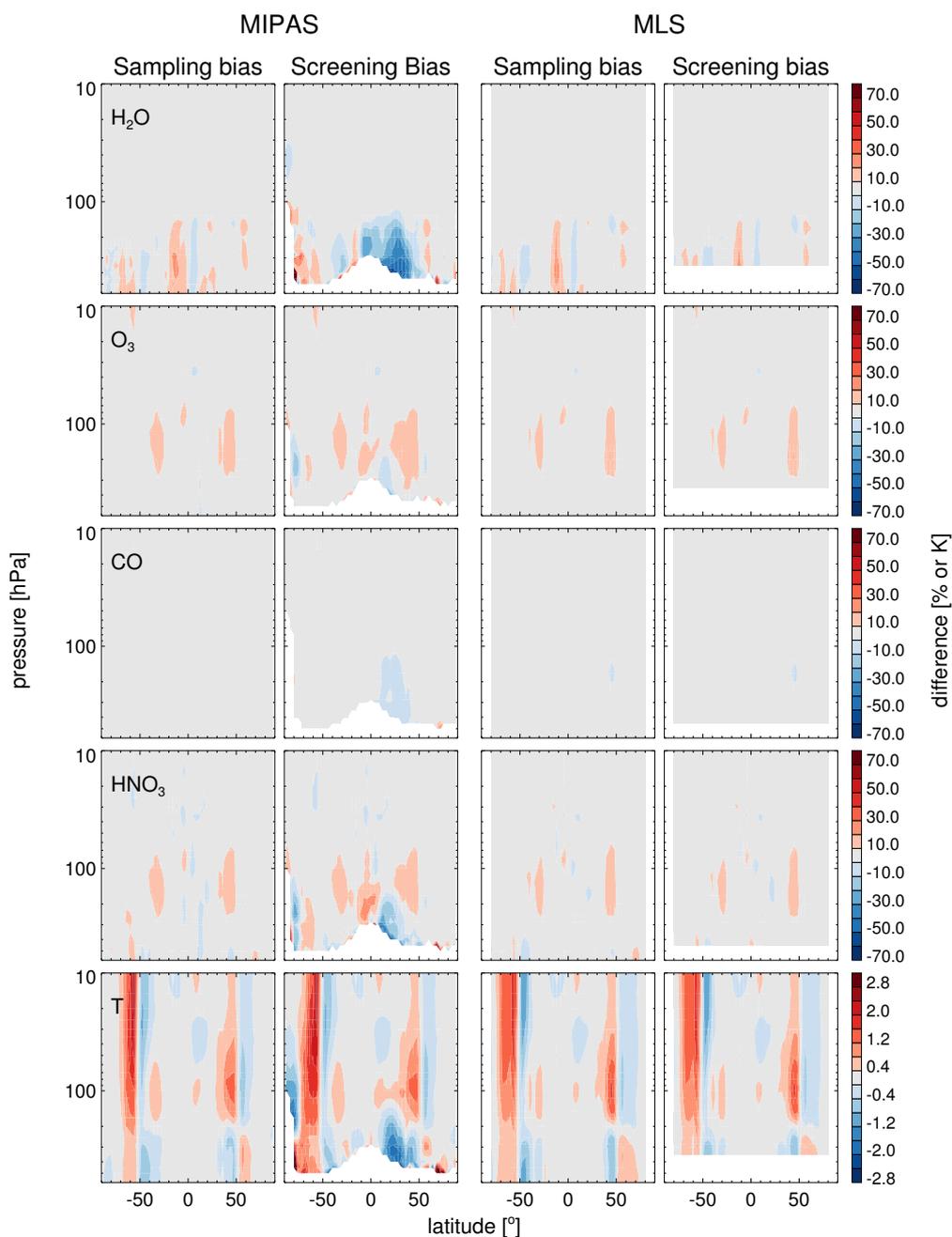
Figure 3. June 2005 sampling and quality screening biases as a function of latitude and pressure for H_2O, O_3, CO, HNO_3, and temperature as measured using MIPAS and MLS. White regions denote a lack of measurements.
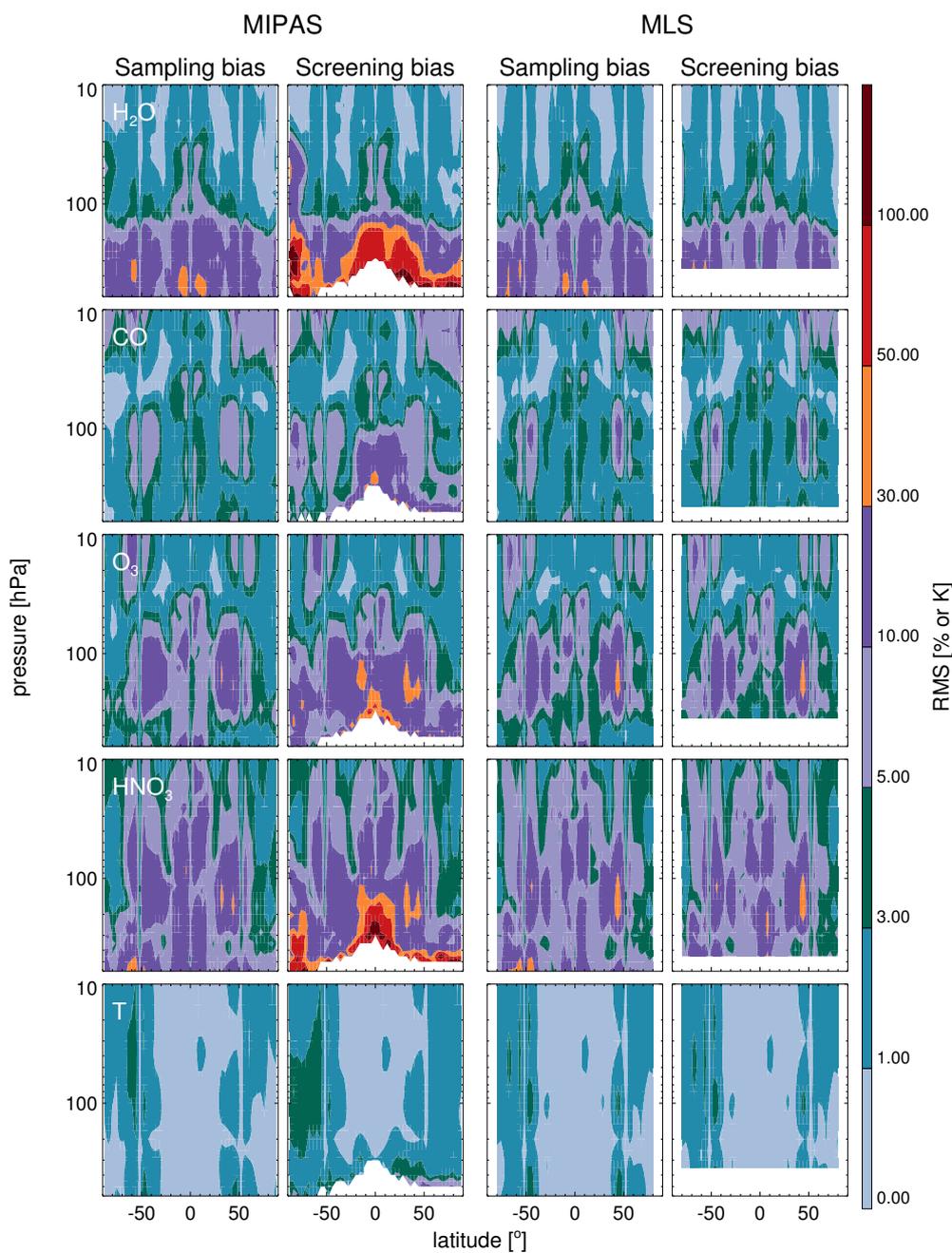
Figure 4. Root-mean-square sampling and quality screening biases for 2005 as a function of latitude and pressure for H_2O, O_3, CO, HNO_3, and temperature as measured using typical MIPAS and MLS data coverage.
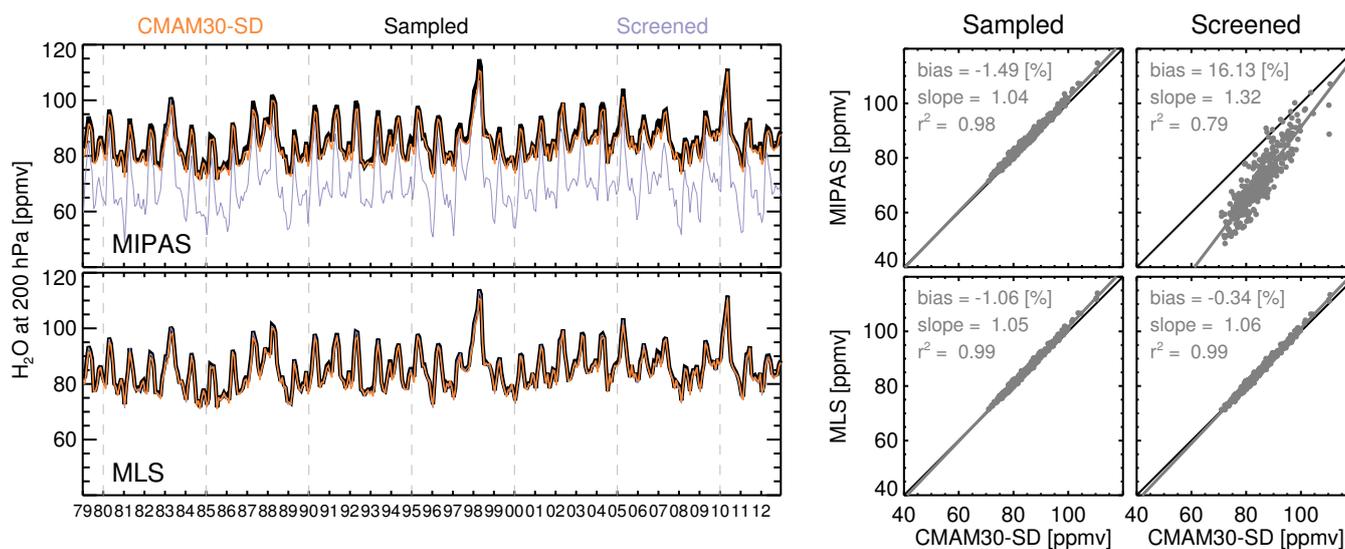
Figure 5. (left) Time series of $20°$S-$20°$N H_2O at 200 hPa for the raw CMAM30-SD fields (orange lines), the full satellited-sampled fields (black lines) and only those points passing the quality screening criteria (thin purple lines) for MIPAS and MLS. (right) Scatterplots between these time series. The black lines are the 1:1 line, and the gray lines are the linear best fits, whose slopes are given. Also, the coefficient of determination, r^2, and the bias are shown.
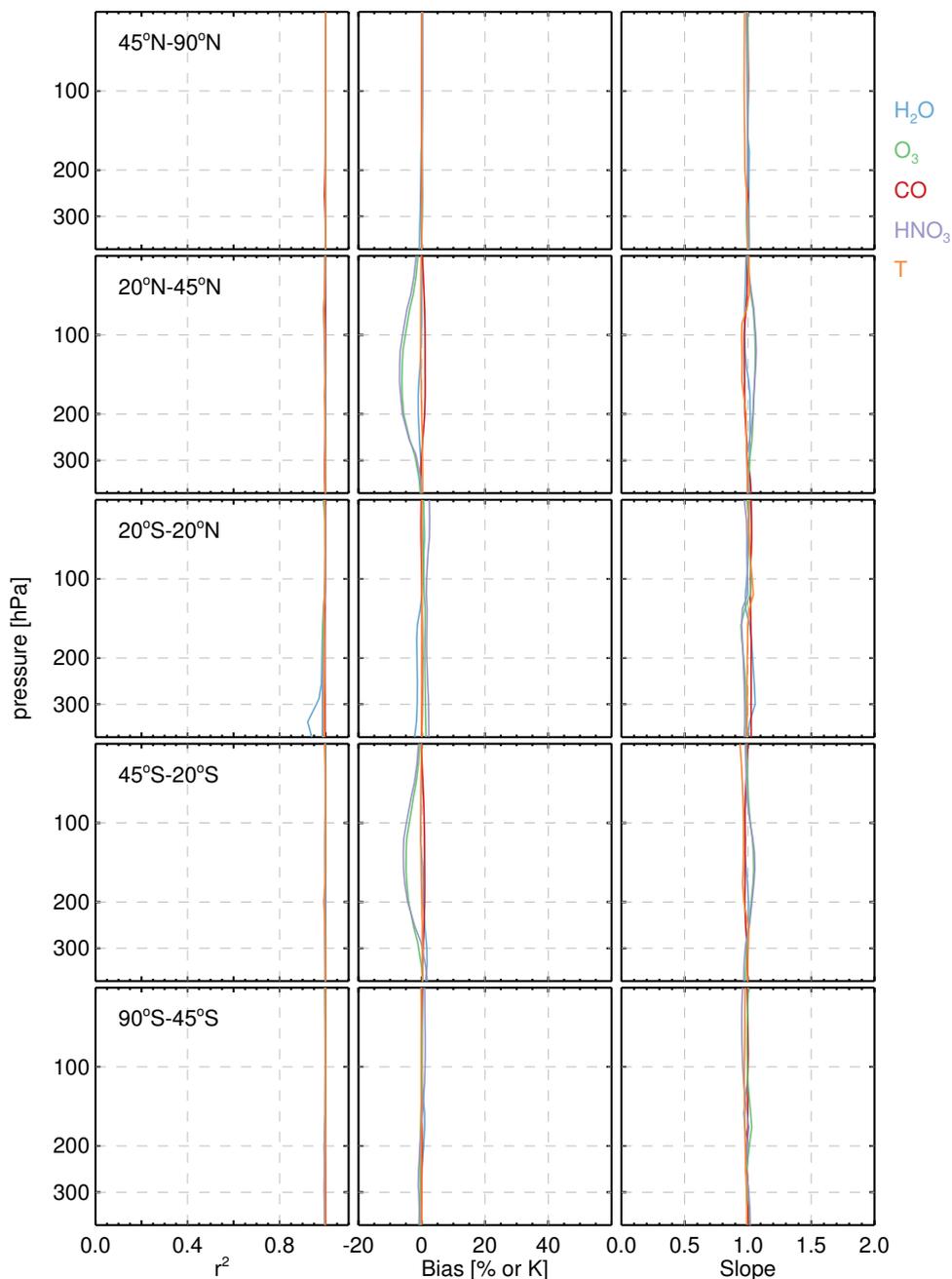
Figure 6. Vertical profiles of the coefficient of determination, the bias, and the linear fit slope for different latitude bands for the MIPAS vs CMAM30-SD scatter using the full satellite-sampled fields. Blue, green, red, purple, and orange lines represent H_2O, O_3, CO, HNO_3, and temperature metrics.
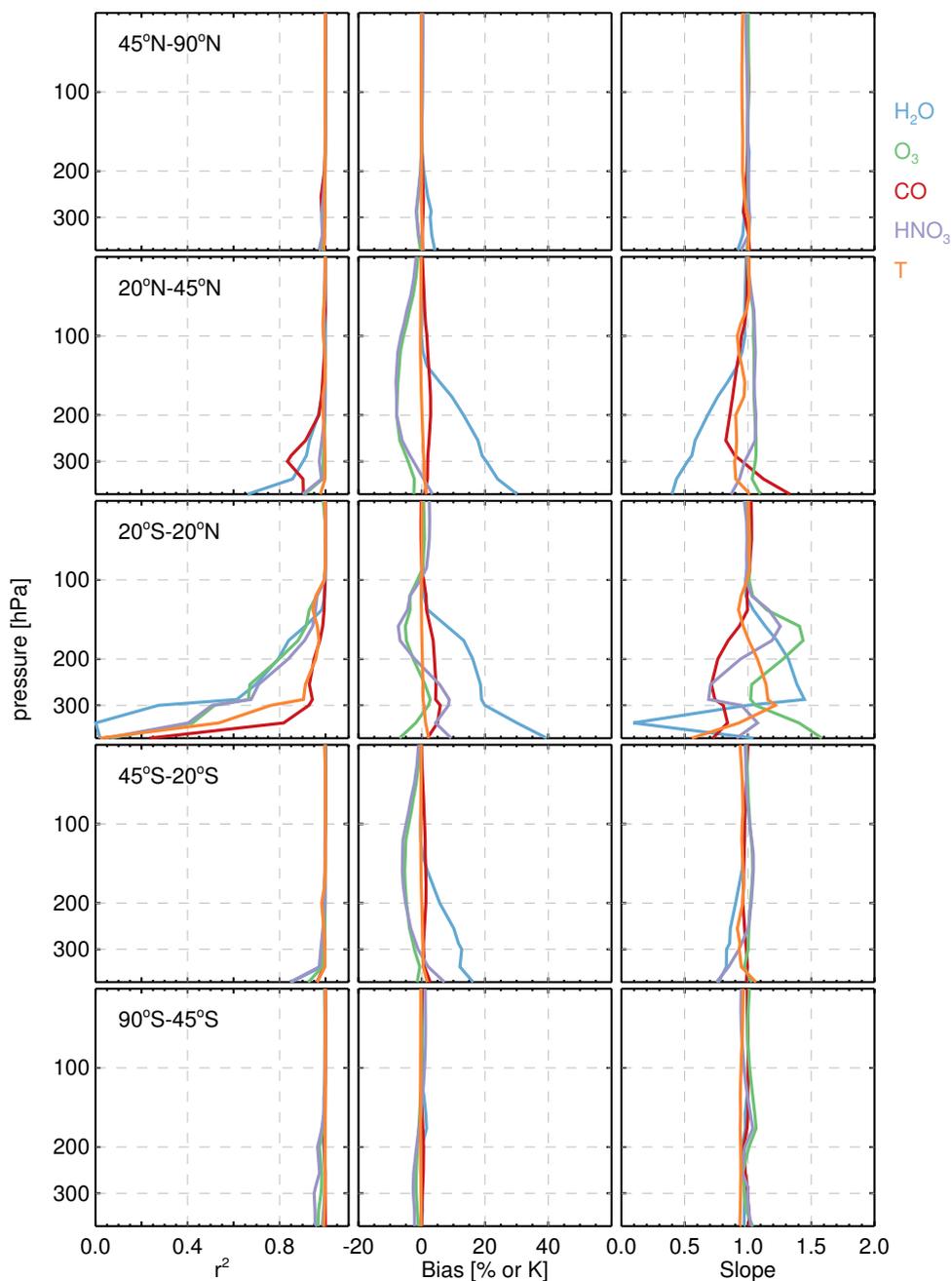
Figure 7. As in Figure 6 but using only the profiles that passed the quality screening.