**Atmospheric
Chemistry
and Physics**

Bias determination and precision validation of ozone profiles from MIPAS-Envisat retrieved with the IMK-IAA processor

T. Steck[1], T. von Clarmann[1], H. Fischer[1], B. Funke[2], N. Glatthor[1], U. Grabowski[1], M. Höpfner[1], S. Kellmann[1], M. Kiefer[1], A. Linden[1], M. Milz[1], G. P. Stiller[1], D. Y. Wang[3], M. Allaart[4], Th. Blumenstock[1], P. von der Gathen[5], G. Hansen[6], F. Hase[1], G. Hochschild[1], G. Kopp[1], E. Kyrö[7], H. Oelhaf[1], U. Raffalski[8], A. Redondas Marrero[9], E. Remsberg[10], J. Russell III[11], K. Stebel[6], W. Steinbrecht[12], G. Wetzel[1], M. Yela[13], and G. Zhang[1]

[1]Institut für Meteorologie und Klimaforschung, Forschungszentrum Karlsruhe, Germany
[2]Instituto de Astrofísica de Andalucía, Granada, Spain
[3]Physics Department, University of New Brunswick, Canada
[4]KNMI (Royal Netherlands Meteorological Institute), De Bilt, The Netherlands
[5]Alfred Wegener Institute for Polar and Marine Research, Potsdam, Germany
[6]Norwegian Institute for Air Research (NILU), Polar Environmental Centre, Tromsø, Norway
[7]FMI-Arctic Research Centre, Sodankylä, Finland
[8]Swedish Institute of Space Physics (IRF), Kiruna, Sweden
[9]Instituto Nacional de Meteorologia (INM), Santa Cruz de Tenerife, Spain
[10]Atmospheric Sciences Competency, NASA Langley Research Center, Hampton, USA
[11]Department of Physics, Hampton University, Hampton, USA
[12]Deutscher Wetterdienst (DWD), Hohenpeißenberg, Germany
[13]Instituto Nacional de Técnica Aeroespacial (INTA), Torrejón de Ardoz, Spain

Abstract. This paper characterizes vertical ozone profiles retrieved with the IMK-IAA (Institute for Meteorology and Climate Research, Karlsruhe – Instituto de Astrofisica de Andalucia) science-oriented processor from high spectral resolution data (until March 2004) measured by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) aboard the environmental satellite Envisat. Bias determination and precision validation is performed on the basis of correlative measurements by ground-based lidars, Fourier transform infrared spectrometers, and microwave radiometers as well as balloon-borne ozonesondes, the balloon-borne version of MIPAS, and two satellite instruments (Halogen Occultation Experiment and Polar Ozone and Aerosol Measurement III). Percentage mean differences between MIPAS and the comparison instruments for stratospheric ozone are generally within ±10%. The precision in this altitude region is estimated at values between 5 and 10% which gives an accuracy of 15 to 20%. Below 18 km, the spread of the percentage mean differences is larger and the precision degrades to values of more than 20% depending on altitude and latitude. The main reason for the degraded precision at low altitudes is attributed to undetected thin clouds which affect MIPAS retrievals, and to the influence of uncertainties in the water vapor concentration.

1 Introduction

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) is a Fourier transform infrared (FTIR) spectrometer measuring the emission of the Earth's atmosphere in limb-viewing mode (Fischer and Oelhaf, 1996; Fischer et al., 2007). MIPAS was launched on the environmental satellite Envisat on 1 March 2002 into a sun-synchronous polar orbit with equatorial local crossing times of 10:00 (descending node) and 22:00 (ascending node). MIPAS operated in its nominal mode from July 2002 to March 2004 in high spectral resolution of nominal $0.025 \, \text{cm}^{-1}$. Flight altitude of Envisat is $800 \, \text{km}$ and one orbit takes about $100 \, \text{min}$ allowing to measure 72 limb sequences per orbit with a sampling of about $500 \, \text{km}$ along track. With these characteristics, MIPAS is able to provide vertical profiles of temperature and a large number of trace gases globally during day and night. The nominal observation mode comprises 17 tangent altitudes

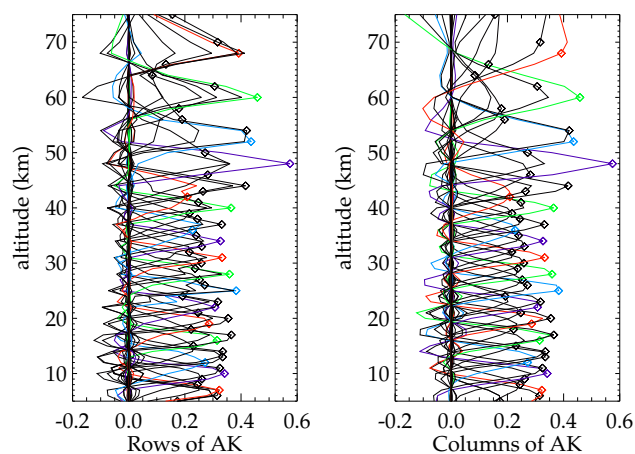*Correspondence to:* T. Steck
(tilman.steck@imk.fzk.de)

Fig. 1. Rows (left) and columns (right) of the averaging kernel matrix for a typical MIPAS ozone retrieval. Rows and colomns of the averaging kernel closest to a tangent altitude are highlighted in colors.

Fig. 2. Vertical resolution of MIPAS ozone derived from the FWHM of columns (solid) and rows (dashed) of the averaging kernel matrix shown in Fig. 1. Dotted lines are for guidance only.

per limb sequence from 6 to 68 km (3 km step-width between 6 and 42 km, and at 47, 52, 60, and 68 km). The 3 km step-width is chosen due to the vertical instantaneous field-of-view (FOV) of MIPAS, which is about 3 km, while the across track FOV is about 30 km. The generation of calibrated radiance spectra, the so-called level 1B data, is performed by the European Space Agency (ESA) (Nett et al., 2002). ESA additionally provides vertical profiles of temperature and 6 key species (H_2O, O_3, HNO_3, CH_4, N_2O, and NO_2), the so-called Level 2 product. Furthermore, there are several institutes which have developed retrieval codes to determine trace gases beyond the ESA key species. The ozone distributions used for validation in this paper were derived with the retrieval processor of the Institut für Meteorologie und Klimaforschung (IMK) and the Instituto de Astrofisica de Andalucia (IAA) (von Clarmann et al., 2003b) and are version V3O_O3_7 (publicly available at http://www-imk.fzk.de/asf/ame/envisat-data/).

Validation of a data product involves various self-consistency tests and comparisons (cf. Rodgers, 2000). The characterization of the calibrated radiance spectra has been extensively studied by e.g. Kleinert et al. (2006) and references therein. The performance of the retrieval processor has been studied in a pre-flight analysis by von Clarmann et al. (2003a), and the characteristics of the retrieved ozone profiles and self-consistency of results have been studied in a sensitivity study by Glatthor et al. (2006). Here we report the final step of validation, which is the comparison of MIPAS ozone profiles to those from other instruments in order to detect any potential bias and to verify the predicted precision.

Comparisons of MIPAS IMK-IAA ozone profiles with other satellite instruments have been performed in Wang et al. (2005) and Bracher et al. (2005). In these papers however, older data versions were used and no vertically high
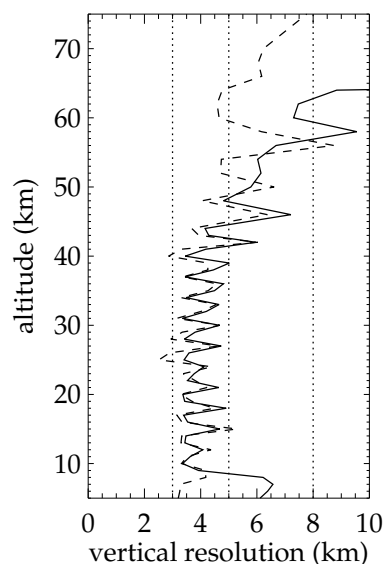
resolved measurements, like ozonesondes, were considered for comparison.

2 MIPAS IMK-IAA ozone data

The ozone data investigated in this paper is produced by the IMK-IAA retrieval processor (von Clarmann et al., 2003b). These data are complementary to the official ESA data product in a sense that the IMK-IAA retrieval aims at the best possible accuracy in a wide altitude region, while the advantage of the ESA data are their better temporal coverage.

2.1 Description

Since launch, MIPAS performed nearly continuous measurements until 26 March 2004 in high spectral resolution with a maximum optical path difference of 20 cm. Due to limitations in computational resources, at IMK only episodes of particular scientific interest are analyzed which is about 20% of all MIPAS measurements. MIPAS measurements which range from pole to pole (87° S to 89° N) are analyzed in quite regular temporal intervals of at least every tenth day. The selected spectral regions used for ozone retrieval, the so-called microwindows, are within the ranges 740–800 cm^{-1} and 1060–1110 cm^{-1}. For more details see Glatthor et al. (2006). Furthermore, spectra are excluded which are contaminated with cloud signal. For this, the method described by Spang et al. (2004) is used, however with a more restrictive cloud index of 4.0.

The retrievals are performed under the assumption of local thermodynamic equilibrium (LTE), which is valid in the
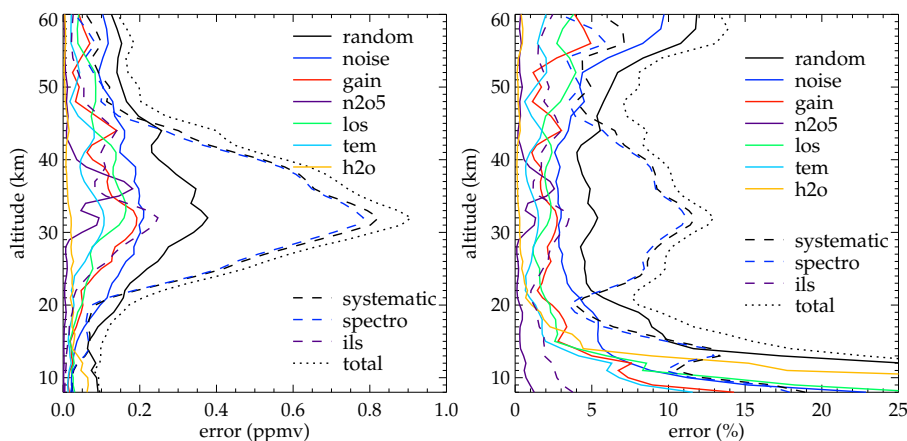
Fig. 3. Estimated ozone error budget of MIPAS averaged over selected locations (4 seasons, 5 latitude bands, night and day). Left: absolute errors, corresponding to values in volume mixing ratios (vmr). Right: percentage errors.

troposphere and most of the stratosphere in the selected spectral regions (Echle et al., 2000). The retrieval is performed on a fixed altitude grid with 1 km step-width up to 44 km, and 2 km step-width above with linear interpolation in altitude between the levels. Since the retrieval grid step-width is finer than the tangent altitude grid (about 3 km, see Sect. 1), regularization is necessary to stabilize the retrieval. Therefore, the iterative retrieval algorithm uses a Tikhonov's first order smoothing constraint, where the strength of the constraint is based on the number of degrees of freedom (Rodgers, 2000; Steck, 2002). More details on the retrieval setup can also be found in Glatthor et al. (2006).

2.2 MIPAS vertical resolution

The vertical resolution of MIPAS ozone is derived from the averaging kernel matrix (Rodgers, 2000; Steck, 2002). Figure 1 (left) shows the rows of the averaging kernel matrix which gives the contribution of the true ozone values to the retrieved values. The columns of the averaging kernel (Fig. 1, right) gives the response of delta-peak-like perturbations at each altitude. Full-width at half-maximum (FWHM) of both rows and columns of the averaging kernel gives a measure of the vertical resolution of MIPAS ozone retrieval. Figure 2 shows the vertical resolution derived from the averaging kernel matrix (Fig. 1). The values are increasing from 3.5 km at 10 km altitude to about 5 km at 42 km, and around 8 km at 60 km altitude. These values are a result of the combination of vertical sampling of measurements (between 3 and 8 km), width of FOV (about 3 km), and regularization.

The oscillating nature of the curve is the result of the following: The spacing of the retrieval grid is 1 km whereas the width of the FOV is 3 km, which means that only some retrieval gridpoints matches a tangent altitude. Generally, this results in a larger diagonal value of the averaging kernel matrix (Fig. 1, colored curves) and smaller values in the ver-

tical resolution. Retrieval gridpoints between two tangent alitudes result in smaller diagonal values of the averaging kernel (Fig. 1, black curves) and hence to larger values in vertical resolution.

2.3 MIPAS error budget

Figure 3 shows the MIPAS ozone error budget averaged over selected locations (4 seasons, 5 latitude bands, night and day), which are used in the precision validation (Sect. 5). The estimated total random error (Fig. 3 right, solid black) varies between 4 and 7% between 20 and 52 km corresponding to values in volume mixing ratios (Fig. 3 left, solid black) of 0.15 parts per million by volume (ppmv) and 0.35 ppmv in this altitude region. The percentage errors are slightly increasing towards 60 km (12%) and down to 15 km (10%). Below 15 km the percentage errors are rapidly increasing to values in the order of 25% for polar and midlatitude conditions or more than 50% for tropical conditions, where the vmr is small. The error in vmr remains below 0.1 ppmv. The estimated random error is dominated by the instrumental noise above 14 km (Fig. 3 left, solid blue). Below 14 km, the error due to uncertain water vapor concentration becomes dominant (Fig. 3 left, solid beige). The reason for that is the exponentially increasing water vapor with decreasing altitude which leads to a dominant water vapor signal in MIPAS spectra. These strong water vapor lines are slightly interfering with ozone lines leading to a dependence of the retrieved ozone on the pre-retrieved water vapor amount. Also errors due to uncertain gain calibration, N_2O_5, line-of-sight (LOS) pointing, and temperature contribute noticeably.

The estimated total systematic error (Fig. 3 right, dashed black) is mainly between 4 and 14% (corresponding to 0.1 and 0.8 ppmv, Fig. 3 left, dashed black) with maxima near the ozone maximum and in the lowermost stratosphere and below. The error is dominated by uncertainties in spectroscopic

data (Fig. 3 right, dashed blue). The altitude-dependence of errors due to spectroscopic data is due to the fact that the microwindows used in the retrieval are varying with altitude. Errors caused by uncertainties in the ILS (instrumental line shape) are in the order of 1 to 4% and thus nearly negligible compared to spectroscopic uncertainties.

2.4 MIPAS data versions

The validation work presented in this paper is based on ozone data of version V3O_O3_7 which uses re-processed ESA spectra of version 4.61/4.62. Since there have been several publications with older ozone data (V2_O3_2 and V2_O3_3) based on spectra with versions 4.55 to 4.59, a comparison between the two data sets is performed in the following.

The main difference between version V3O_O3_7 and old versions (V2_O3_2 and V2_O3_3) is, besides the different spectra versions, the weaker regularization leading to more degrees of freedom of the retrieved ozone profile and better vertical resolution (Glatthor et al., 2006). Furthermore, the a priori profile is set zero in order to avoid artificial discontinuities in the retrieved ozone distributions.

Figure 4 shows the latitude-dependent comparison between the current MIPAS (V3O_O3_7) and old MIPAS (V2_O3_2 and V2_O3_3) IMK-IAA profiles. The mean profiles (Fig. 4, left) are in good agreement showing only little differences. The mean difference (for calculation see Eq. 3) is below ±0.2 ppmv for all altitudes and latitudes (Fig. 4, middle). The mean differences show some similar structures for all latitudes at the same altitudes (e.g., at 17 km, 28 km, and 45 km), which is not surprising, since the regularization is latitude-independent in both versions. The percentage mean differences (Fig. 4, right) are below ±5% above 19 km for all latitudes. Remarkable is the large percentage mean difference for tropical conditions (Fig. 4, right, middle row) around 17 km. The mean difference in vmr is similar to other latitudes, but due to the small absolute values, percentage differences reach 50%. This effect is discussed in more detail in Sect. 5.1 when dealing with ozonesonde comparison at the tropical station Paramaribo. At 14 km a compensation is seen for the differences at 17 km.

Beyond the differences in the retrieval setup, it was found that the different spectra versions (versions 4.55 to 4.59 compared to 4.61/62) also cause differences in the retrieval results. In general the mean differences are rather small between old and new MIPAS ozone data. Only at altitudes below 19 km differences can exceed 10%.

3 Validation

3.1 Comparison instruments

MIPAS measurements in high resolution are available from the period July 2002 to March 2004, which are used for comparison. An overview of geolocations of coincidences with

considered instruments, together with periods of comparison, mean spatial distances, mean temporal differences, and numbers of coincidences is summarized in Table 1. The different instruments are selected in such a way that all latitudes are covered.

To achieve a large number of correlative measurements, the coincidence criteria for non-satellite instruments are set to: maximum of 6 h in time, maximum of 800 km in distance with a maximum latitudinal distance of ±4°. The more stringent latitudinal criterion has proven to be a good choice for all comparisons, since latitudinal variations are in general more pronounced than longitudinal ones.

For the two satellite instruments under consideration a more stringent spatial criterion is used: maximum of 400 km in distance with a maximum latitudinal distance of ±2°. This still allows for plenty of correlative measurements.

3.2 Strategy and terminology

We use the terminology and formalism as summarized by von Clarmann (2006a,b). In particular, we understand bias is the mean deviation of the measurements from the truth, the relative bias is the mean deviation of profiles measured by two different instruments, and precision is the reproduceability of a measurement, i.e. the bias-corrected root mean squares difference between MIPAS profiles and the true ozone profiles, all under consideration of effects of finite resolution of the measurements. We further follow the strategy recommended by von Clarmann (2006a) to first determine the relative bias between two instruments and to use this for the subsequent precision validation.

In a first step, two co-incident data sets (profiles, averaging kernel matrices and covariance matrices), which are given on different altitude grids, need to be made comparable. The vertically higher sampled profiles are transformed to the lower sampled altitude grid by

$$x_h = W^* x_m, \tag{1}$$

where x_h is the high resolved profile from the original measurement grid x_m, and W^* is the pseudo-inverse of W (interpolation matrix from coarse to fine grid) with $W^*W=I$ (Rodgers, 2000).

There are two possibilities to handle the smoothing error problem. If the altitude resolutions of the co-incident measurements are in good approximation equal, the smoothing errors cancel out when the differences between co-incident measurements are calculated and consideration of smoothing errors is necessary neither in bias determination nor in precision validation. This applies to the comparison with MIPAS-B.

If the contrast in altitude resolution is so large that the finer resolved profile can be regarded as an ideal profile, the averaging kernel matrix of the coarser resolved profile A can be applied to the finer resolved profile x_h in order to adjust
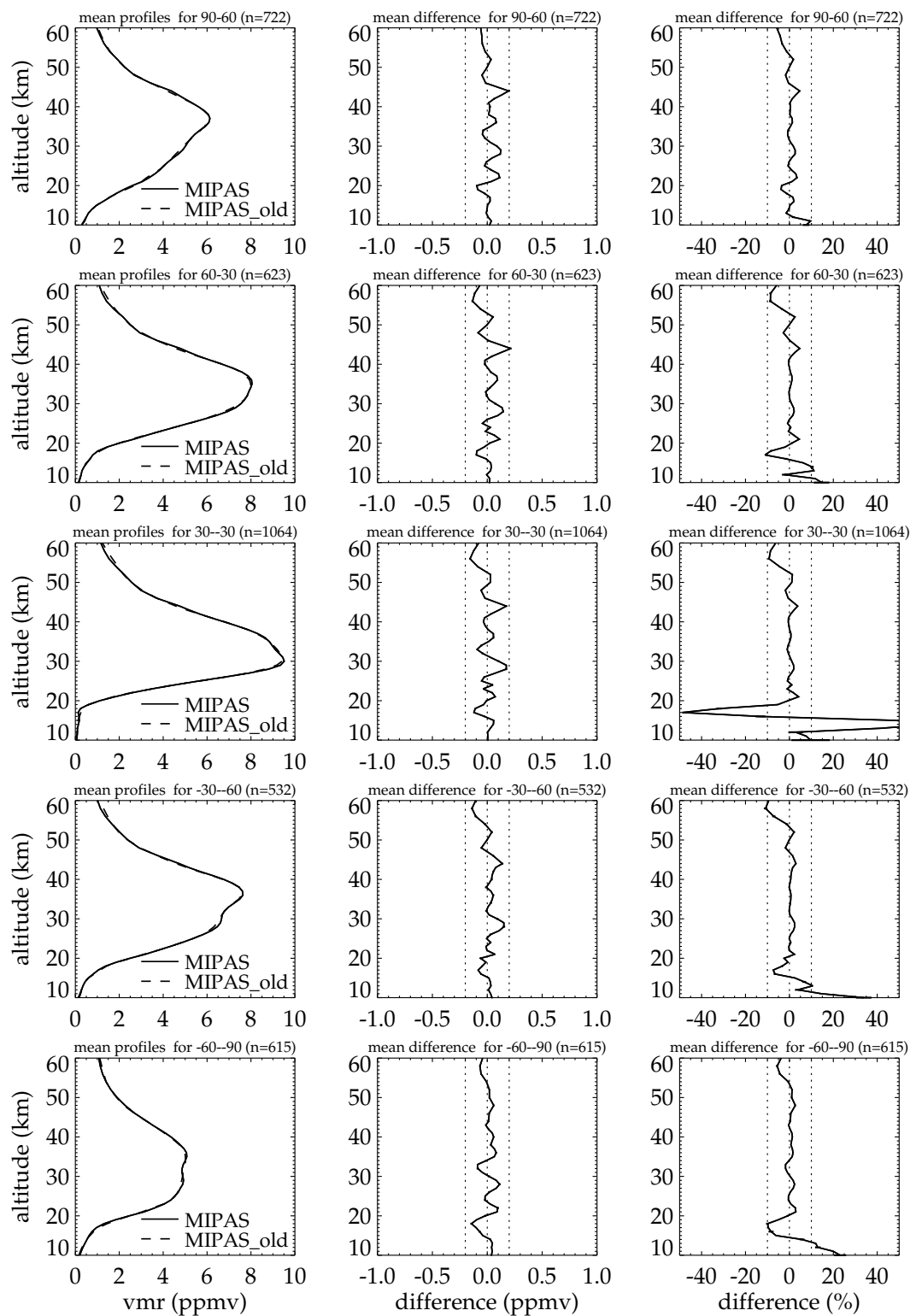
Fig. 4. Comparison of MIPAS version V3O_O3_7 and older versions (V2_O3_2 and V2_O3_3, MIPAS_old) ozone profiles for 5 latitude bins. From top to bottom (number of coincidences in brackets): 90° N to 60° N (722), 60° N to 30° N (623), 30° N to 30° S (1064), 30° S to 60° S (532), and 60° S to 90° S (615). From left to right: Mean ozone profiles, mean difference, and percentage mean difference. Dotted lines are for guidance only.

Table 1. Geolocations of considered instruments together with period of comparison, mean spatial distance $\overline{\Delta_\mathrm{d}}$, mean temporal difference $\overline{\Delta_\mathrm{t}}$, and number of comparison pairs.

Instrument location (type)	lat/lon [deg]	period	$\overline{\Delta_\mathrm{d}}$ [km]	$\overline{\Delta_\mathrm{t}}$ [h]	coinc. [#]
Ny-Ålesund (lidar)	78.9/11.9	10.2002–03.2003	460	2.6	362
Alomar (lidar)	69.3/16.0	09.2002–03.2004	456	2.6	108
Sodankyla (sonde)	67.8/26.6	08.2002–03.2004	456	2.8	163
Kiruna (FTIR)	67.8/20.4	09.2002–03.2004	498	1.9	498
Kiruna (microwave)	67.8/20.4	11.2002–12.2003	427	2.8	412
Hohenpeissenberg (lidar)	47.8/11.0	09.2002–03.2004	457	1.6	133
Hohenpeissenberg (sonde)	47.8/11.0	07.2002–03.2004	492	4.0	130
Zugspitze (microwave)	47.4/11.0	02.2003–07.2003	512	2.3	77
Izana (sonde)	28.5/–16.3	07.2002–03.2004	541	0.4	50
Izana (FTIR)	28.5/–16.3	09.2002–03.2004	519	2.8	189
Paramaribo (sonde)	5.8/–55.2	07.2002–03.2004	343	1.3	19
Belgrano (sonde)	–77.8/–34.6	07.2002–03.2004	393	2.8	127
MIPAS-B (balloon)	campaigns	09.2002–07.2003	274	2.2	9
HALOE (satellite)	globally	09.2002–02.2004	227	3.8	333
POAM (satellite)	N + S poles	09.2002–03.2004	214	3.3	674

it to the altitude resolution of the coarser resolved profile (Rodgers, 2000; Rodgers and Connor, 2003):

$$x = x_\mathrm{a} + \mathbf{A}(x_\mathrm{h} - x_\mathrm{a}), \qquad (2)$$

where x_a is the a priori profile of the coarser resolved profile. In case of MIPAS, where x_a is zero, Eq. (2) simplifies to $x = \mathbf{A}x_\mathrm{h}$. After this transformation, the smoothing characteristics have not to be considered any further. In case of ground-based sounding by microwave and FTIR instruments, their averaging kernels are used to smooth the MIPAS profiles, whereas the higher resolved lidar, ozonesonde, HALOE and POAM profiles are smoothed by the MIPAS averaging kernels. Residual smoothing error differences not accounted for correctly by the approaches above because the assumptions made may hold only in approximation, are discussed in von Clarmann and Grabowski (2007).

For bias determination and precision validation we follow the strategy proposed by von Clarmann (2006a). In particular we perform validation of ozone mixing ratios at discrete altitudes individually instead of profile validation (Rodgers, 2000; Migliorini et al., 2004), in order not to depend on covariance information between the altitudes. Since the true atmospheric state is not known, we can only estimate the relative bias $\check{b}_\mathrm{diff}$ (or mean difference) between two datasets from a sample of K co-incident pairs of measurements, which is

$$\check{b}_\mathrm{diff} = \frac{\sum_{k=1}^{K}(\hat{x}_{\mathrm{MIPAS};k} - \hat{x}_{\mathrm{ref};k})}{K}, \qquad (3)$$

where $\hat{x}_{\mathrm{MIPAS};k}$ are the ozone profiles retrieved by MIPAS, and $\hat{x}_{\mathrm{ref};k}$ are the co-incident profiles measured by the reference instrument. The statistical uncertainty of the bias $\check{\sigma}_{\mathrm{bias};n}$, at altitude gridpoint n is estimated as

$$\check{\sigma}_{\mathrm{bias};n} = \sqrt{\frac{\sum_{k=1}^{K}(\hat{x}_{\mathrm{MIPAS};n,k} - \hat{x}_{\mathrm{ref};n,k} - \check{b}_{\mathrm{diff};n})^2}{K(K-1)}}, \qquad (4)$$

for samples large enough to disregard t-statistics (Gosset, 1908). As pointed out by von Clarmann (2006a), this assessment does not need any error estimates of $\hat{x}_{\mathrm{MIPAS}}$ or $\hat{x}_{\mathrm{ref}}$. For percentage multiplicative bias estimates we use the percentage mean difference rather than the mean percentage difference.

With the relative bias between two instruments available, the precision at altitude gridpoint n is validated by altitude-wise testing of the de-biased mean squares difference of the co-incident measurements against the ex ante estimate of the variance of the difference $\sigma^2_{\mathrm{diff};n,k}$ in a χ^2 sense:

$$\chi^2 = \frac{\sum_{k=1}^{K}(\hat{x}_{\mathrm{MIPAS};n,k} - \hat{x}_{\mathrm{ref};n,k} - \check{b}_{\mathrm{diff};n})^2/K}{\sigma^2_{\mathrm{diff};n,k}}. \qquad (5)$$

The ex ante estimate of the variance of random error of the difference σ^2_diff (or in short form: estimated combined random error) includes the following additive components: the estimated random error variance of MIPAS ozone; the estimated random error variance of the ozone abundance measured with the reference instrument; the variance representing the expected difference due to less than perfect coincidence. In our application no smoothing error of the difference has to be considered here, because of application of Eq. (2) whenever relevant.

The error due to less than perfect coincidence (or spatial-temporal mismatch) is derived from ECMWF (European

Centre for Medium-Range Weather Forecasts) ozone distributions and is calculated as follows:

$$\sigma^2_{\text{mm};n,k}=(\overline{\delta x_{\text{t};n}}\Delta_{\text{t};k})^2+(\overline{\delta x_{\text{lat};n}}\Delta_{\text{lat};k})^2+(\overline{\delta x_{\text{lon};n}}\Delta_{\text{lon};k})^2. \quad (6)$$

$\delta x_{\text{lat};n}$ is the difference in ozone at altitude n between two different latitude points, divided by the distance in latitude direction, and averaged over a longitudinal bin of 10 degrees. $\delta x_{\text{lon};n}$ and $\delta x_{\text{t};n}$ are defined like $\delta x_{\text{lat};n}$ but for longitudinal and temporal differences. $\Delta_{\text{lat};k}$, $\Delta_{\text{lon};k}$, and $\Delta_{\text{t};k}$ are the actual differences in latitude, longitude, and time for the comparison pair k.

4 Bias determination

4.1 Comparisons with ground-based instruments

4.1.1 Lidar data

Comparison measurements are performed with stratospheric ozone lidar systems. The technique used is the so-called Differential Absorption Lidar (DIAL) (see, e.g. Godin et al., 1999). Ozone profiles are usually measured in reliable quality between 10 and 45 km altitude. The data was taken from the NILU (Norwegian Institute for Air Research) server for the stations at Hohenpeissenberg, Alomar, and Ny-Ålesund. Since the ozone concentration is given in number densities, the values are transformed into volume mixing ratios for proper comparison with MIPAS results.

Measurements at Hohenpeissenberg (47.8° N, 11.0° E) were available for the entire comparison period. The comparison is shown in Fig. 5 (top), where 133 comparison pairs were found. The mean profiles (Fig. 5, left) show good agreement with mean differences (Fig. 5, middle) below ±0.2 ppmv up to 30 km altitude, and reaching about 0.5 ppmv at 38 km. The percentage mean differences (Fig. 5, right) are within ±10% except at 16 km, where the percentage mean difference reaches 15%. The rather large positive difference around 38 km is likely due to a negative bias in the lidar data, which was found in comparison with SAGE II (W. Steinbrecht, personal communication, 2006).

Measurements at Alomar (69.3° N, 16.0° E) were available for the entire comparison period with 108 comparison pairs. The mean profiles (Fig. 5, middle left) agree very well up to 44 km with mean differences (Fig. 5, middle) below ±0.3 ppmv. The large differences at the upper end of the altitude range can be explained by the rather large uncertainty in lidar data there. The percentage mean difference (Fig. 5, right) is below ±10% between 13 and 44 km. At very low altitudes, mean differences become negative with values of up to –20%.

Measurements at Ny-Ålesund (78.9° N, 11.9° E) (Steinbrecht et al., 1999) were available between October 2002 and March 2003 resulting in 362 comparison pairs. The mean profiles (Fig. 5, bottom left) show good agreement up to

36 km with mean differences below about ±0.2 ppmv (Fig. 5, middle). Above 36 km mean differences between MIPAS and Ny-Ålesund lidar reach values of about −0.5 ppmv. Above 44 km the lidar values become again unreasonably large. The percentage mean differences (Fig. 5, right) are below ±10% between 17 and 43 km. At 13 km the difference reaches +15%.

4.1.2 FTIR data

Measurements are performed with ground-based Fourier transform spectrometers in the infrared region (FTIR). The spectrometers are directed to the sun measuring atmospheric absorption spectra. For ozone, a broad microwindow around 1000 cm^{-1} and two narrow ones around 780 cm^{-1} are used. The number of independent pieces of information (or degrees of freedom) is around 5, leading to a vertical resolution of about 6–7 km, depending on altitude and atmospheric situation.

Ozone profiles at Kiruna (67.8° N, 20.4° E) (Kopp et al., 2002.) were available for the entire comparison period. Figure 6 (top left) shows the comparison between MIPAS and FTIR mean ozone profiles. The agreement is reasonable except a pronounced discrepancy between 18 and 28 km where the mean difference exceeds the uncertainty of the mean difference by far. This is attributed to specific choice of the FTIR a priori profile, which is only corrected in part by application of Eq. (2) due to non-linearities in the ground-based retrieval. Percentage mean differences are below ±10% (Fig. 6, right) above 22 km.

In addition to profile comparison, partial zenith column densities of ozone between 10 and 40 km have been analyzed. The agreement between MIPAS and FTIR (Kiruna) is very good (see Table 2) with a percentage mean difference of 0.4% which is close to the estimated error of the mean difference (0.3%). This supports our explanation that differences in the profile comparison are caused by residual altitude resolution and a priori differences as discussed above.

Also Izana (28.2° N, 16.3° W) FTIR ozone profiles (Schneider et al., 2004) were available from September 2002 until March 2004. The mean profiles (Fig. 6, bottom left) show good agreement where the mean differences are less than ±0.4 ppmv (Fig. 6, middle). The percentage mean differences are below ±15% above 18 km altitude (Fig. 6, right). The large percentage mean difference around 17 km can partly be attributed to the spatial coincidence criteria. More restrictive criteria (600 km in distance and 3° in latitude) reduce the percentage mean difference from 40% to 20% (not shown).

Like in the Kiruna case, partial columns of ozone between 10 and 40 km have been compared in addition. The agreement between MIPAS and FTIR (Izana) is again very good (see Table 2) with a percentage mean difference of −0.4%.
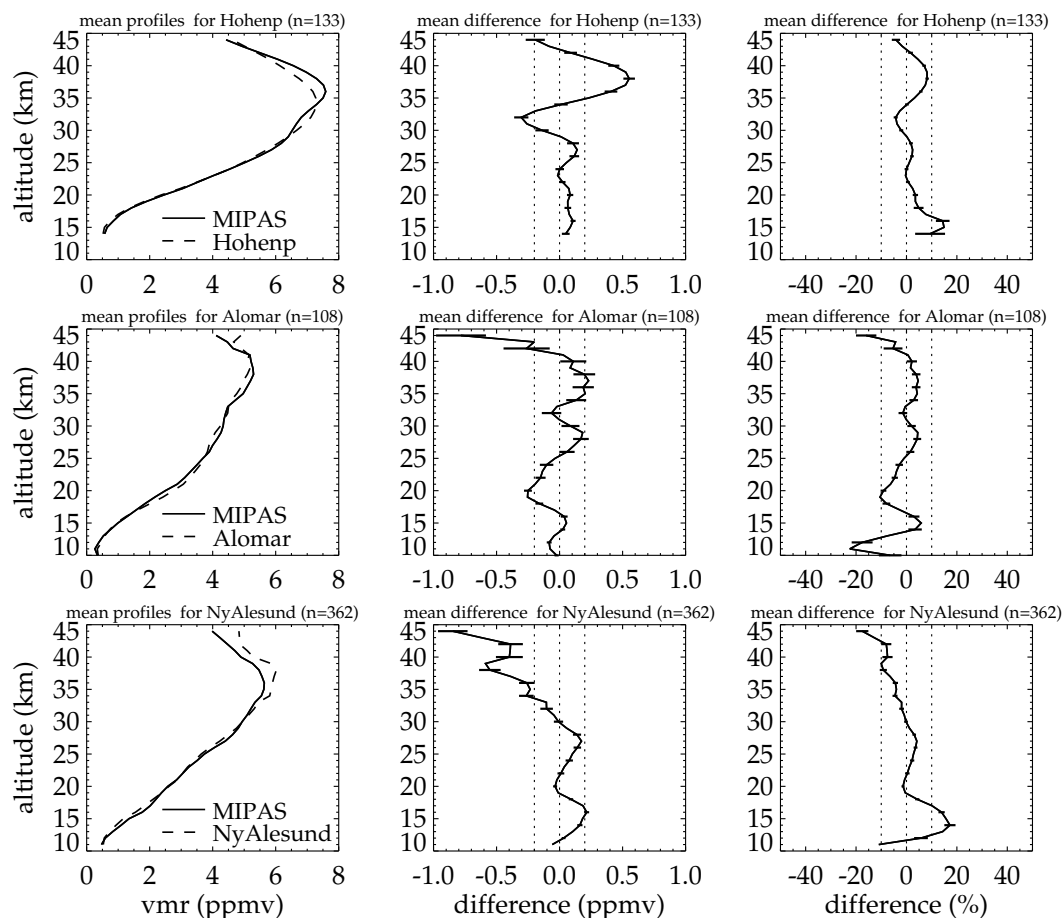
Fig. 5. Comparison of MIPAS and lidar ozone profiles. From top to bottom (number of coincidences in brackets): Hohenpeissenberg (133), Alomar (108), and Ny-Ålesund (362). From left to right: MIPAS (solid) and lidar (dashed) mean profiles, mean difference (solid) between MIPAS and lidar with uncertainty of the mean (error bars), and percentage mean difference (solid) between MIPAS and lidar with uncertainty of the mean (error bars). Dotted lines are for guidance only.

Table 2. Comparison of MIPAS and FTIR partial zenith column densities between 10 and 40 km altitude. Entries in the table from left to right are: station, number of comparison pairs, mean partial columns of MIPAS and FTIR, mean difference, and uncertainty of the mean. Values are given in units of $10^{20}\,\mathrm{m}^{-2}$.

Station	N	MIPAS	FTIR	$\breve{b}_{\mathrm{diff}}$	$\breve{\sigma}_{\mathrm{bias}}$
Kiruna	498	745.8	743.1	2.7 (0.4%)	2.1 (0.3%)
Izana	189	676.3	679.3	–3.0 (–0.4%)	2.0 (0.3%)

4.1.3 Microwave data

Two ground-based microwave radiometers were used for MIPAS validation, the 195–224 GHz Kiruna microwave radiometer KIMRA at IRF Kiruna (Raffalski et al., 2005) and the 268–280 GHz millimeter wave radiometer MIRA 2 of

IMK Karlsruhe (Kopp et al., 2002.). The retrieved ozone profiles of both instruments have a vertical resolution of about 6–8 km in the lower stratosphere degrading to more than 15 km in the lower mesosphere. The degrees of freedom amount to about 4 for measurements taken during good weather conditions.

Measurements from Kiruna (67.8° N, 20.4° E) were available from November 2002 until December 2003. Mean profiles of MIPAS and KIMRA are displayed in Fig. 7 (top left) showing very good agreement above 22 km with mean differences below ±0.2 ppmv (Fig. 7, middle). The nearly perfect agreement above 40 km is due to the strong influence of the microwave regularization at these altitudes since the measurement response is rather low. The smoothed MIPAS results are forced to the microwave a priori profile. Percentage mean differences (Fig. 7, right) are below ±5% above 22 km. In both, mean absolute and percentage mean difference, no indication of a systematic bias in MIPAS ozone is noticeable.
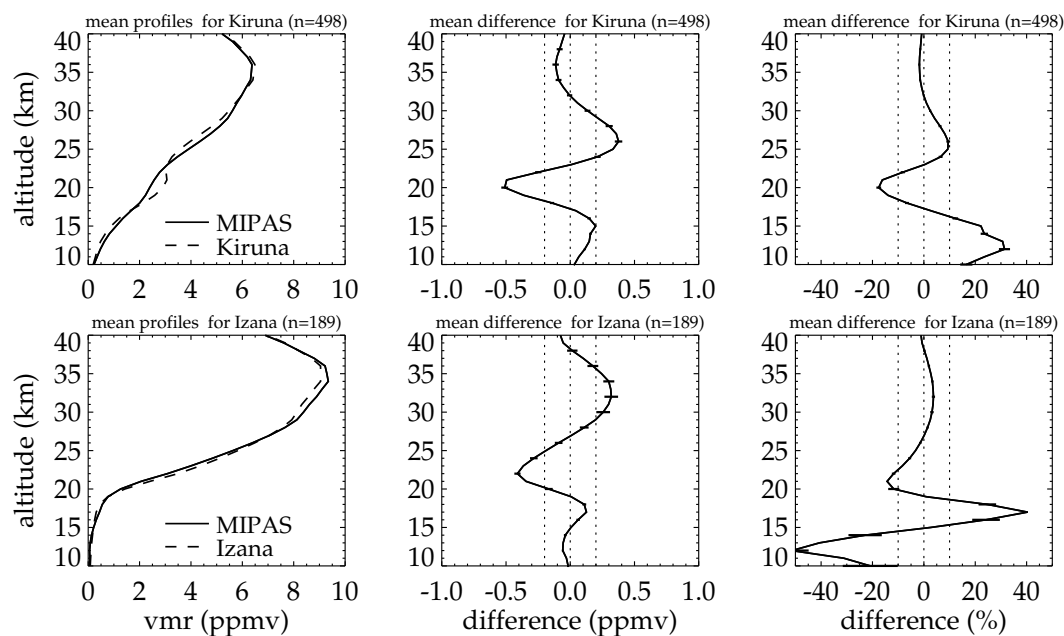
Fig. 6. Comparison of MIPAS and FTIR ozone profiles. From top to bottom (number of coincidences in brackets): Kiruna (498) and Izana (189). From left to right: see Fig. 5.

Table 3. Comparison of MIPAS and microwave partial zenith column densities between 20 and 60 km altitude. Entries in the table from left to right are: station, number of comparison pairs, mean partial columns of MIPAS and microwave (MW), mean difference, and uncertainty of the mean. Values are given in units of 10^{20} m^{-2}.

Station	N	MIPAS	MW	$\check{b}_{diff}$	$\check{\sigma}_{bias}$
Kiruna	412	404.0	403.7	0.3 (0.1%)	2.6 (0.6%)
Zugspitze	77	589.6	573.8	15.8 (2.7%)	3.6 (0.6%)

Table 4. Ozone sonde types and manufacturers for different stations. Abbreviations: ECC (Electrochemical Concentration Cell), BM (Brewer Mast).

Station	Type	Manufacturer
Sodankyla	ECC	both SPC-6A and ENSCI-Z
Hohenpeissenberg	BM	–
Izana	ECC	SPC-6A
Paramaribo	ECC	SPC-6A
Belgrano	ECC	both SPC-6A and ENSCI-Z

In addition to profile comparison, partial columns of ozone between 20 and 60 km have been analyzed. The agreement between MIPAS and microwave (Kiruna) is very good (see Table 3) with a percentage mean difference of 0.1% which is below the estimated error of the mean difference (0.6%).

Ozone profiles from Zugspitze (47.4° N, 11.0° E) were available for the MIRA campaign between February and July 2003. The comparison is shown in Fig. 7. There is good agreement below 25 km and above 38 km with mean differences (Fig. 7, middle) smaller than ±0.2 ppmv. In the ozone maximum, MIPAS mean values are larger than MIRA by up to 0.7 ppmv. The related percentage differences (Fig. 7, right) remain below ±10%. The larger differences at the upper end of the comparison region can be attributed to diurnal variations in ozone. Microwave measurements at Zugspitze have been performed mainly during daytime, whereas MIPAS measurements also have significant nighttime contributions with enhanced ozone values above about 50 km.

Like in the Kiruna case, partial columns of ozone between 20 and 60 km have been compared in addition. The agreement between MIPAS and microwave (Zugspitze) is not as good as for the Kiruna station (see Table 3) with a percentage mean difference of 2.7%. However, this is not surprising considering the 10% difference in the mean ozone profiles near the ozone maximum.

4.2 Comparisons with balloon-borne instruments

4.2.1 Ozonesondes

Measurements are performed in situ with ozonesondes of the types electrochemical concentration cell (ECC) and Brewer-Mast (BM) on small balloons (see Table 4). Ozone profiles are usually measured with low random error and high vertical resolution (Smit and Straeter, 2004) from the Earth's surface
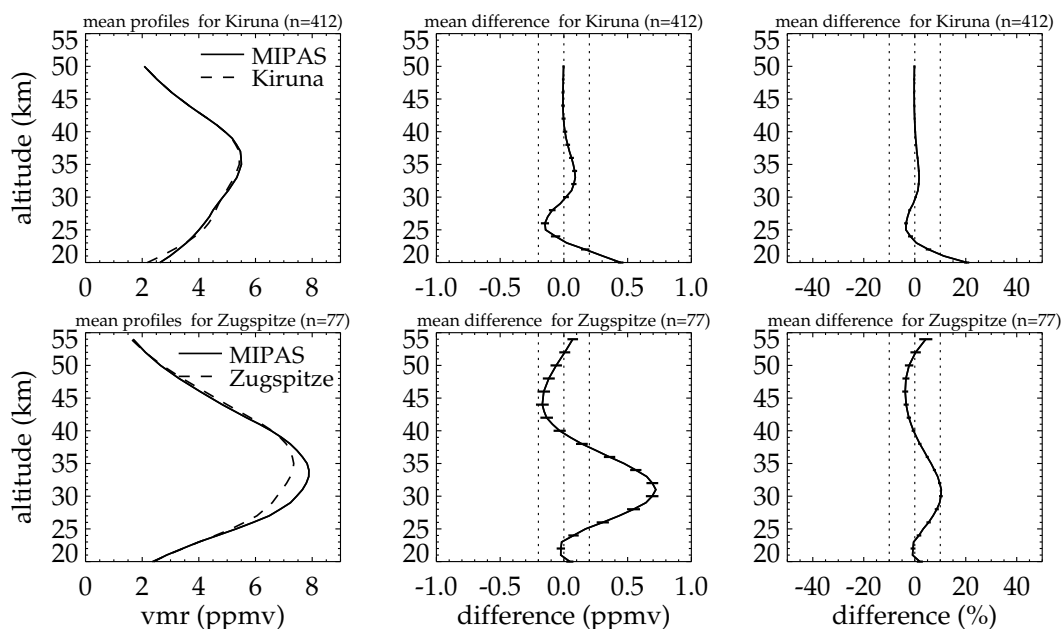
Fig. 7. Comparison of MIPAS and microwave ozone profiles. From top to bottom (number of coincidences in brackets): Kiruna (412) and Zugspitze (77). From left to right: see Fig. 5.

up to the mid stratosphere (about 32 km). Since ozone is measured in partial pressure, the values can be easily transformed into vmrs. The data was taken from the NILU (Norwegian Institute for Air Research) server except for Izana.

Measurements at Sodankyla (67.4° N, 26.6° E) were available for the entire measurement time of MIPAS. The comparison is shown in Fig. 8 (top), where the mean profiles (left) agree very well. The mean difference (middle) is below ±0.2 ppmv for the complete altitude range. Also the percentage mean difference (Fig. 8, right) is very small and exceeds ±6% difference only around 15 km and below 10 km altitude.

The comparison with measurements at Hohenpeissenberg (47.8° N, 11.0° E) is shown in Fig. 8 (second from top). The mean difference (middle) is increasing with altitude above 25 km and reaching values larger than +0.5 ppmv. The percentage mean difference reaches values of +20% below 16 km altitude. At higher altitudes the difference is continuously increasing and reaching +10% at 30 km. A positive mean difference is noticeable in both troposphere and stratosphere.

Izana (28.5° N, 16.3° W) measurements are compared with MIPAS and results are shown in Fig. 8 (third from top). The mean difference (Fig. 8, middle) is increasing with altitude above 21 km and reaches +0.5 ppmv at 28 km altitude. The percentage mean difference is below ±20% in the troposphere and below 10% in the lower stratosphere. Similar to the Hohenpeissenberg comparison, a positive mean difference between MIPAS and ozonesonde is noticeable.

Paramaribo (5.8° N, 55.2° W) measurements are also available for the entire comparison time interval. However, only 19 coincidences were found for this period. The mean profiles are displayed in Fig. 8 (second from bottom, left). The mean differences (middle) are below ±0.3 ppmv below 24 km altitude. Above 24 km altitude the differences are continuously increasing towards −0.7 ppmv. This is in contradiction to the higher MIPAS ozone mixing ratios at Hohenpeissenberg and Izana. A hint that these systematic differences are not a problem of MIPAS but of the different ozonesondes is given in Smit and Straeter (2004), who report biases between different ozonesonde systems of up to 10%.

At 17 km MIPAS is underestimating ozone values considerably by about 0.3 ppmv leading to even a negative mixing ratio in the mean. In the IMK-IAA MIPAS retrieval negative mixing ratios are not suppressed. While these are physically meaningless, a positivity constraint such as the retrieval of the logarithms of vmr instead of vmr adds complication to the statistical analysis of results and thus has not been applied to the current retrievals.

Investigation of the Paramaribo case has shown, that application of averaging kernels has basically no influence on the difference between MIPAS and sonde profiles, which means that the smoothing error of the MIPAS retrieval is not responsible for the negative values around 17 km. Dependence on temperature and water vapor was found negligibly small. However, in a revision of the MIPAS retrieval baseline, one suspicious microwindow around 742 cm^{-1} has been detected to be used around 17 km tangent altitude. Disregard of this microwindow decreases the difference between ozonesonde
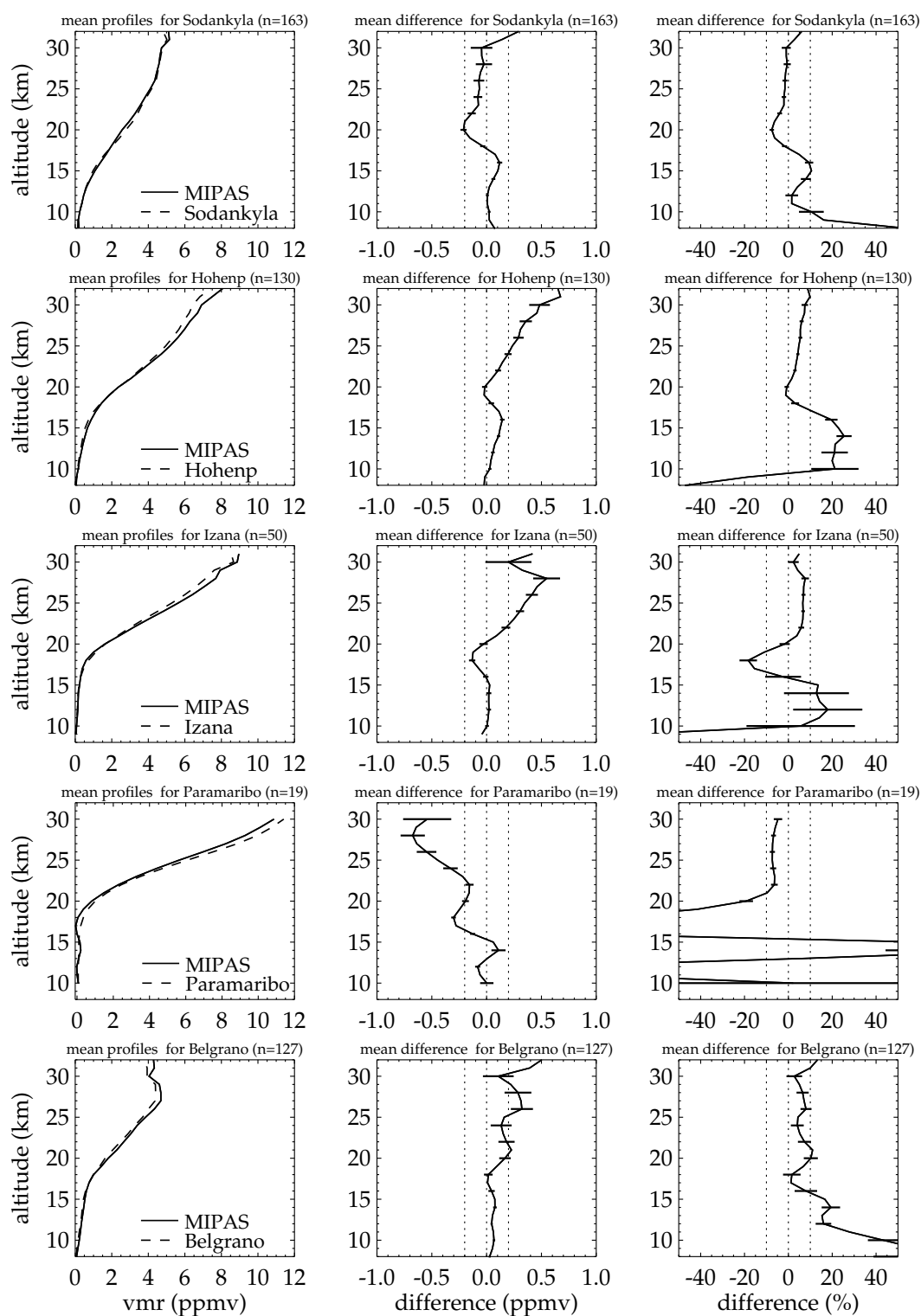
Fig. 8. Comparison of MIPAS and ozonesonde profiles. From top to bottom (number of coincidences in brackets): Sodankyla (163), Hohenpeissenberg (130), Izana (50), Paramaribo (19), and Belgrano (127). From left to right: see Fig. 5.
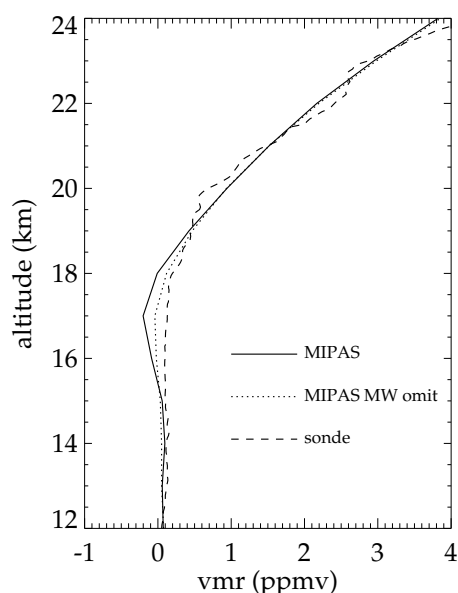
Fig. 9. Comparison of MIPAS and ozonesonde (Paramaribo) profiles for 6 February 2003. Omitting one microwindow at one altitude leads to better MIPAS result (dotted line).

and MIPAS (see Fig. 9) but does not fully remove it (dotted line in Fig. 9). This microwindow seems to be particularly sensitive to residual cloud signal in the spectra caused by e.g. very thin cirrus clouds undetected by the cloud filtering procedure. Negative ozone mixing values were retrieved by MIPAS particularly when the actual cloud index was low but still higher than the threshold used by the cloud detection algorithm. An estimation of the retrieval error due to cirrus clouds has been done in Glatthor et al. (2006), which shows a considerable increase of the ozone error. The negative tendency of the ozone values are not yet understood and further investigations are necessary.

The comparison with measurements at Belgrano (77.8° S, 34.6° W) is shown in Fig. 8 (bottom). The mean profiles (Fig. 8, left) agree well, with a rather positive mean difference. In absolute values this is mostly pronounced above 20 km (Fig. 8, middle). The percentage mean difference (Fig. 8, right) is below ±20% above 17 km altitude. Below 17 km the percentage mean difference reaches 30% but the mean absolute differences are below 0.2 ppmv.

4.2.2 MIPAS-B

MIPAS-B (Friedl-Vallon et al., 2004) is the balloon-borne version of MIPAS measuring the atmospheric emission in a limb viewing mode. The spectral resolution ($0.034\,\mathrm{cm}^{-1}$) and the vertical resolution (2–3 km) is similar to the space-borne MIPAS. MIPAS-B measurements performed near Aire sur l'Adour (24 September 2002, Envisat validation campaign) and Kiruna (20/21 March and 3 July 2003) are compared to those made by MIPAS-Envisat.

Figure 10 shows the comparison of MIPAS-Envisat and MIPAS-B ozone measurements. The mean profiles (Fig. 10, left) agree very well with mean differences below about ±0.3 ppmv for the complete altitude range. Percentage mean differences are within ±10% above 17 km. Around 15 km, the percentage mean difference is positive, which has already been detected in comparison with ozonesondes. The larger differences can partly be attributed to air parcels viewed by the instruments which are inside or at the edge of the vortex.

4.3 Comparisons with satellite instruments

4.3.1 HALOE

HALOE (Halogen Occultation Experiment) is a solar occultation instrument (Russell III et al., 1993) on board the Upper Atmosphere Research Satellite launched in September 1991. It measured ozone during sunrise and sunset atmospheric conditions. Due to the limb-viewing geometry a good vertical resolution (about 2.3 km) was achieved, which is slightly better than that of MIPAS. The measurements were made in the mid IR, at $9.6\,\mu\mathrm{m}$, which is partly the same band used in MIPAS retrievals. For the comparison, HALOE version 19 data is used.

Figure 11 shows the global comparison between MIPAS and HALOE. For nearly all altitudes, we see a positive mean difference (Fig. 11, middle) between MIPAS and HALOE with values up to 0.6 ppmv at 28 km and 44 km. The positive mean differences are attributed to the use of different spectroscopic databases. HALOE uses line strengths which are increased by 5% compared to HITRAN 92 leading to smaller mean ozone values (Brühl et al., 1996; Randall et al., 2003). This is confirmed by a bias in the percentage mean difference (Fig. 11, right) between roughly 0 and 10%. Since the microwindows used for MIPAS vary with altitude, it is expected that the bias also varies with altitude. Below 15 km, mean percentage differences become larger, which has already been detected in comparisons to ozonesondes (see Sect. 4.2.1).

The pronounced structure at around 18 km (see Fig. 11, middle) is attributed to the residual cloud signal problem in the suspicious microwindow used in the MIPAS retrieval, which has already been discussed in the context of the comparison with Paramaribo ozonesonde measurements (see Sect. 4.2.1). Too small MIPAS values do not only appear in the tropical region, but at all latitudes (Fig. 11). This indicates that not only cirrus, but also polar stratospheric clouds (PSCs) could be involved.

The other sharp structure at around 32 km is caused by a discontinuity in the MIPAS background continuum emission, which is set zero above this altitude but fitted below (von Clarmann et al., 2003b). Glatthor et al. (2006) have found by sensitivity studies that this leads to compensation effects which introduce a structure in MIPAS ozone profiles
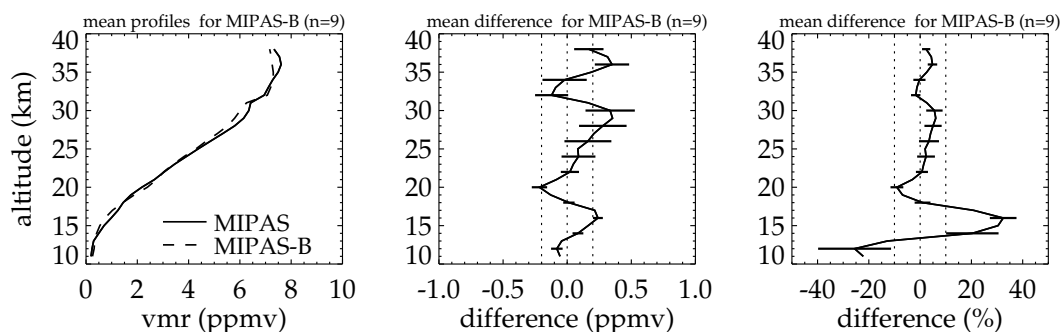
Fig. 10. Comparison of MIPAS-Envisat and MIPAS-B ozone profiles (9 coincidences). From left to right: see Fig. 5.

around 32 km. A revision of the retrieval strategy for the atmospheric continuum emission is under way.

4.3.2 POAM III

Polar Ozone and Aerosol Measurement III (POAM III, Lucke et al., 1999; Lumpe et al., 2002) on board the SPOT-4 satellite was launched in March 1998. POAM III uses, like HALOE, the solar occultation technique for measuring the vertical distribution of ozone during sunrise and sunset. The spectral region used for ozone measurements lies, unlike MIPAS and HALOE, in the visible region at 603 nm. The vertical resolution (Randall et al., 2003) in the stratosphere (15 to 50 km altitude) is 1 km, reaching 5 km at 5 km altitude, and 2.5 km at 60 km altitude. For the comparison, POAM version 4 data is used.

Due to the given orbit of SPOT-4, there are no measurements between about 50° N and 50° S available. Together with the coincidence criteria (Sect. 3.1) only comparisons polewards of 50° N and 60° S have been found leading to 674 comparison pairs.

Figure 12 shows the comparison between MIPAS and POAM correlative measurements divided into northern and southern polar regions. The mean profiles (Fig. 12, left) of MIPAS and POAM agree well. At 32 km a slight negative kink in the mean MIPAS results is visible (especially in the southern polar region) but not present in POAM. In the comparison with HALOE this feature has already been discussed and is due to the strategy of atmospheric continuum retrieval. Since this feature seems to be more pronounced in the southern polar region suggests that very thin PSCs may have an influence on the ozone retrieval due to upward error propagation. With a vertical FOV slightly larger than 3 km, the suggestion above becomes more likely.

When considering the entire altitude range, there is no obvious bias of any sign visible (Fig. 12, middle), leading to mean differences mainly below ±0.2 ppmv. The percentage mean differences (Fig. 12, right) are below ±10% for all altitudes and both polar regions. Below 20 km, a similar feature like in the HALOE comparison is visible, however less pronounced, which is due to the given latitude region.

4.4 Summary

In the previous sections, MIPAS ozone data have been compared to single instruments which give findings for specific locations and altitudes. In most altitudes and latitude bands, MIPAS ozone agrees well with the comparison instruments. Two problems have been identified: First, MIPAS ozone has a systematic negative kink around 32 km. This is explained by the treatment of atmospheric continuum emission in the retrieval (c.f. Sect. 4.3.1), which is forced to zero above this altitude. The transition will be moved to higher altitudes where the actual continuum signal can be expected to be zero. Second, around 18 km MIPAS ozone tends to low values. At least part of this problem has been solved and can be attributed to a particular microwindow used at these altitudes which will be removed in future versions.

In addition to the local and latitude band-wise investigation a comprehensive global comparison between MIPAS ozone and all other validation data from the different instrument types (lidar, FTIR, microwave, ozonesonde, MIPAS-B, HALOE, and POAM) is performed. The mean differences (Fig. 13, left) are within ±0.5 ppmv at all altitudes showing very good agreement between MIPAS and the comparison instruments. The mean differences have a slight positive tendency particularly in the lower stratosphere with values between −0.1 and 0.5 ppmv. FTIR values around 20 km seem to be too high compared to the other instruments. Mean differences below 18 km show values between −0.1 ppmv and 0.2 ppmv. Around 15 km altitude, MIPAS ozone is higher compared to most other instruments which is likely a compensation for the negative kink above.

The percentage mean differences (Fig. 13, right) are within ±10% above 18 km with very few exceptions. This is mainly within the expected systematic errors of MIPAS which are in the order of 5 to 10% (see Fig. 3). The large mean differences detected in specific latitude bands around 18 km are below −5% compared to most instrument types and are of less importance on a global scale. Below 17 km the percentage mean differences become larger with a positive tendency compared to most instrument types with values between −20
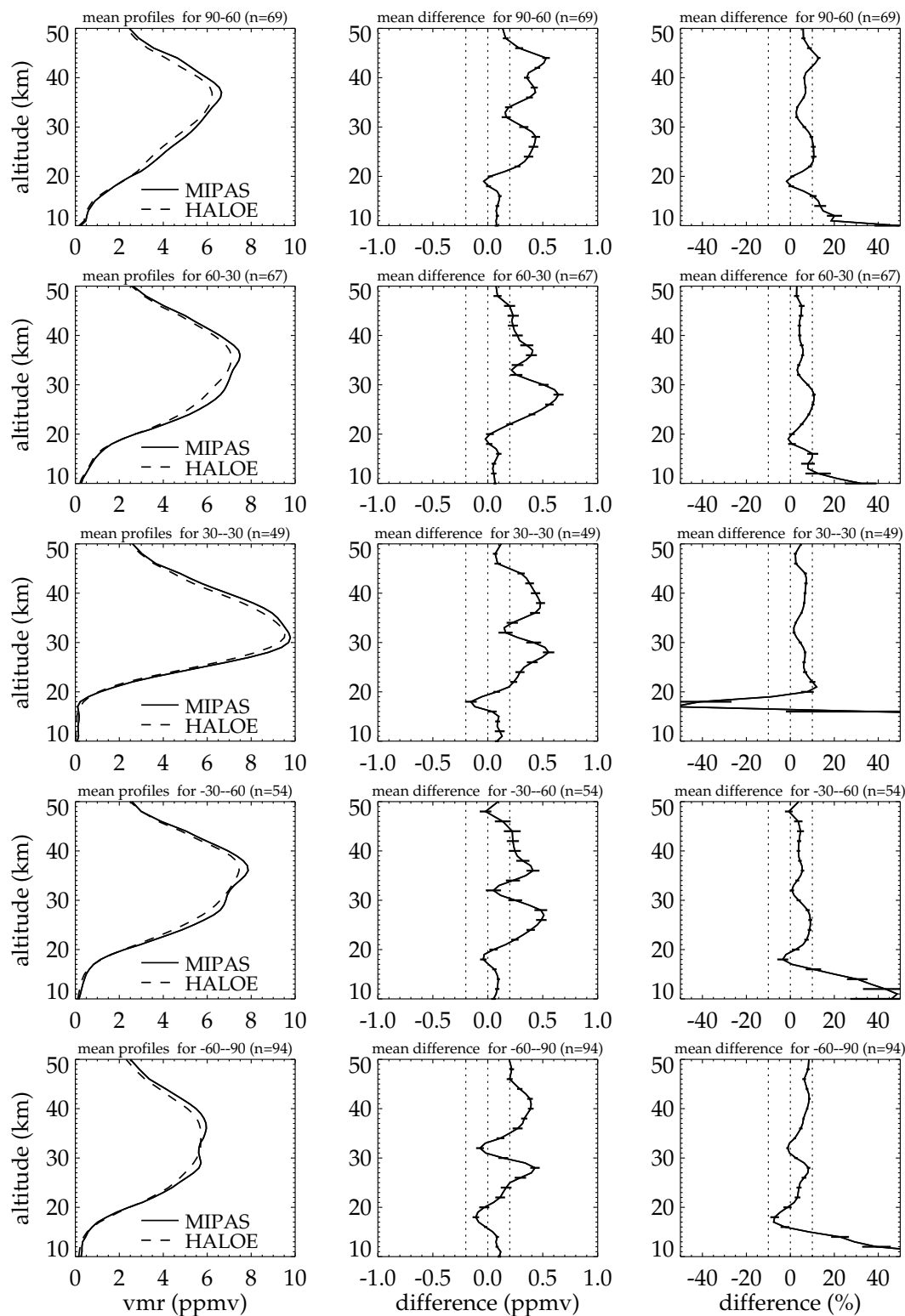
Fig. 11. Comparison of MIPAS and HALOE ozone profiles for 5 latitude bins. From top to bottom (number of coincidences in brackets): 90° N to 60° N (69), 60° N to 30° N (67), 30° N to 30° S (49), 30° S to 60° S (54), and 60° S to 90° S (94). From left to right: see Fig. 5.
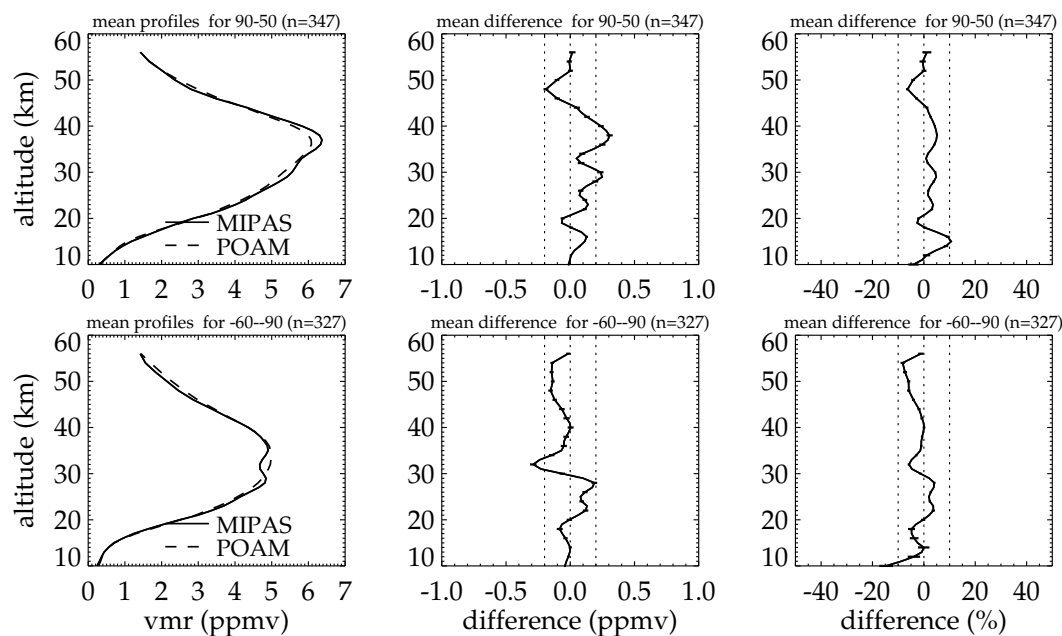
Fig. 12. Comparison of MIPAS and POAM ozone profiles for 2 latitude bins. From top to bottom (number of coincidences in brackets): 90° N to 60° N (347) and 60° S to 90° S (327). From left to right: see Fig. 5.
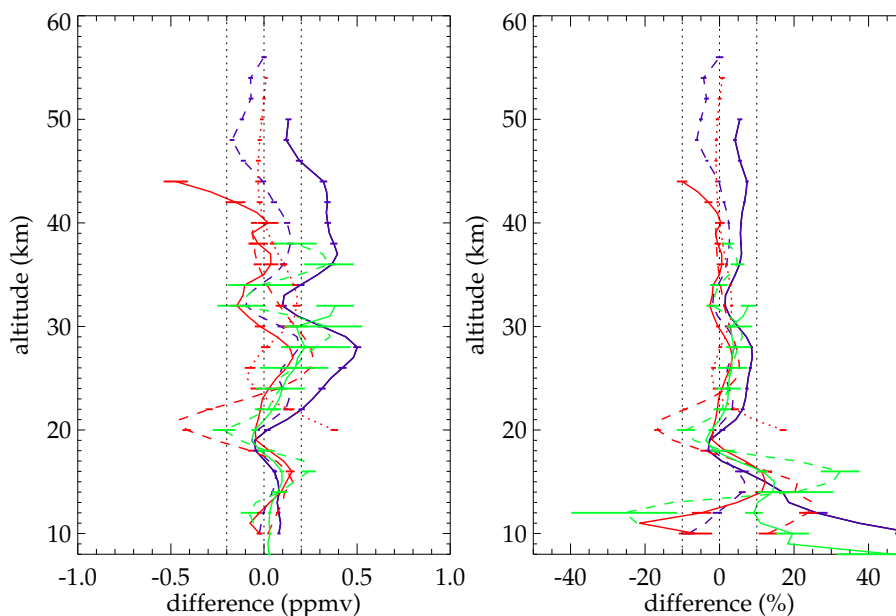

Fig. 13. Comprehensive global comparison of ozone from MIPAS with HALOE (333 comparison pairs, solid blue), POAM (674, dashed blue), lidar (603, solid red), FTIR (687, dashed red), microwave (489, dotted red), ozonesonde (489, solid green), and MIPAS-B (9, dashed green). Left: mean difference between MIPAS and others with uncertainty of the mean (error bars), right: percentage mean difference between MIPAS and others with uncertainty of the mean (error bars). Dotted lines are for guidance only.

and 30%. Here, the mean differences exceed the estimated systematic errors of MIPAS. But considering that the comparison instruments also have biases, the found mean differences are of reasonable size and should not be overstressed.

5 Precision validation

In this section the estimated random error of the difference is compared with the bias-corrected root mean squares (rms) difference between MIPAS and the comparison instrument
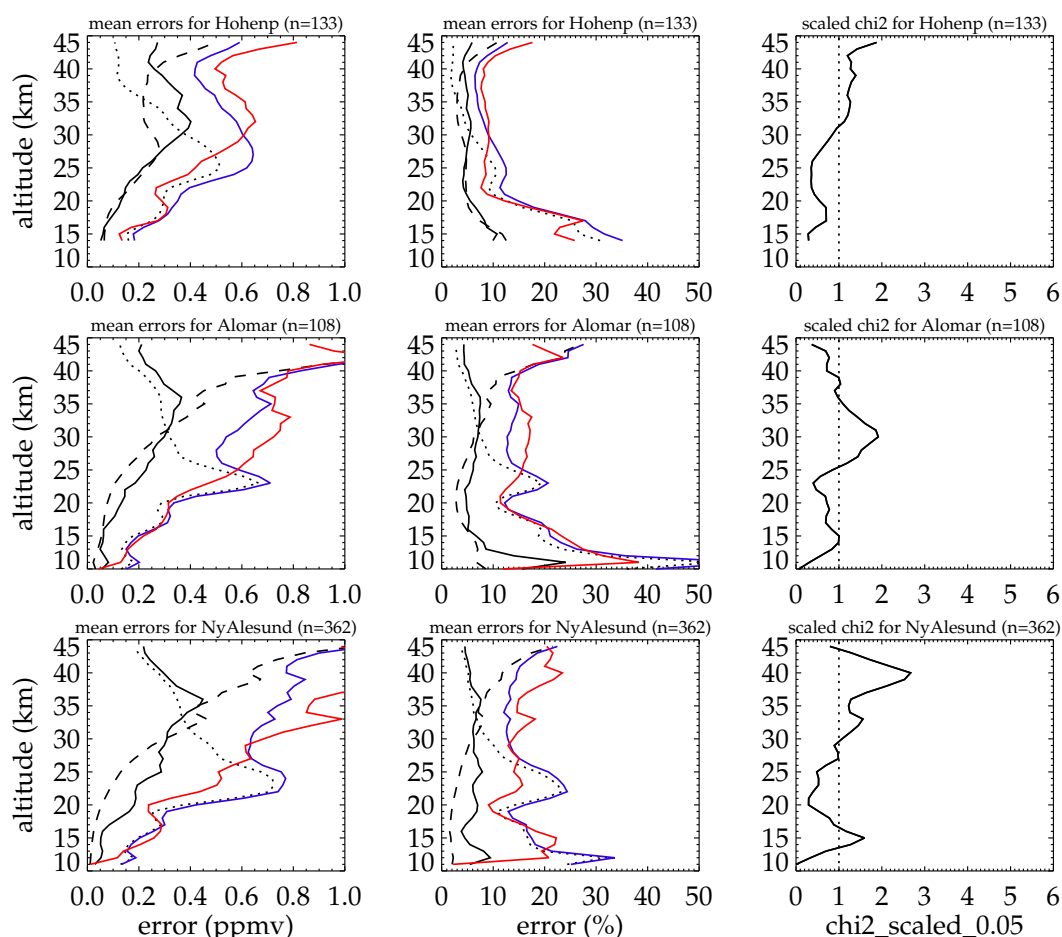
Fig. 14. Analysis of MIPAS and lidar ozone random errors. From top to bottom (number of coincidences in brackets): Hohenpeissenberg (133), Alomar (108), and Ny-Ålesund (362). From left to right: mean absolute errors (solid: MIPAS random, dashed: lidar random, dotted: mismatch, blue: combined random, red: de-biased rms difference), percentage mean errors (lines as for absolute errors), and scaled χ^2 values (dotted line marks the 95% limit).

(see Sect. 3.2). The instrument descriptions can be found in the previous section. The χ^2 value, introduced in Eq. (5), gives indication at which altitude two datasets agree or disagree for given random error estimation. Since the altitude range for comparison of two profiles can differ from location to location due to, e.g., cloud top height and top altitude of ozone sondes, the number of comparison pairs per altitude can differ with altitude. Therefore we have scaled each χ^2 by χ^2_f, corresponding to a significance level f for the appropriate number of pairs. Thus the ratio χ^2/χ^2_f should exceed unity only with probability f (Migliorini et al., 2004), where f is set to 5%.

5.1 Validation with ground-based instruments

5.1.1 Lidar

The analysis of MIPAS and lidar random errors is displayed in Fig. 14. The estimate of the total random error (solid blue)

for Hohenpeissenberg agrees very well with the de-biased rms difference (solid red, Fig. 14, top). The error due to less than perfect coincidence (Fig. 14, dotted line) plays the dominant role at altitudes below about 30 km where natural variability is very large. The scaled χ^2 values only slightly exceed the 95% confidence level.

For Alomar (Fig. 14, middle) and Ny-Ålesund (Fig. 14, bottom) the agreement between estimated random error and rms difference is similar to the case of Hohenpeissenberg but at a higher level of values. This can be attributed to the higher natural variability in the polar vortex region. Additionally, inside the polar vortex, very thin PSCs have affected MIPAS retrievals.

5.1.2 FTIR

The analysis of MIPAS and FTIR random errors is displayed in Fig. 15. For Kiruna (Fig. 15, top), the de-biased rms difference agrees well with the estimated random error for most
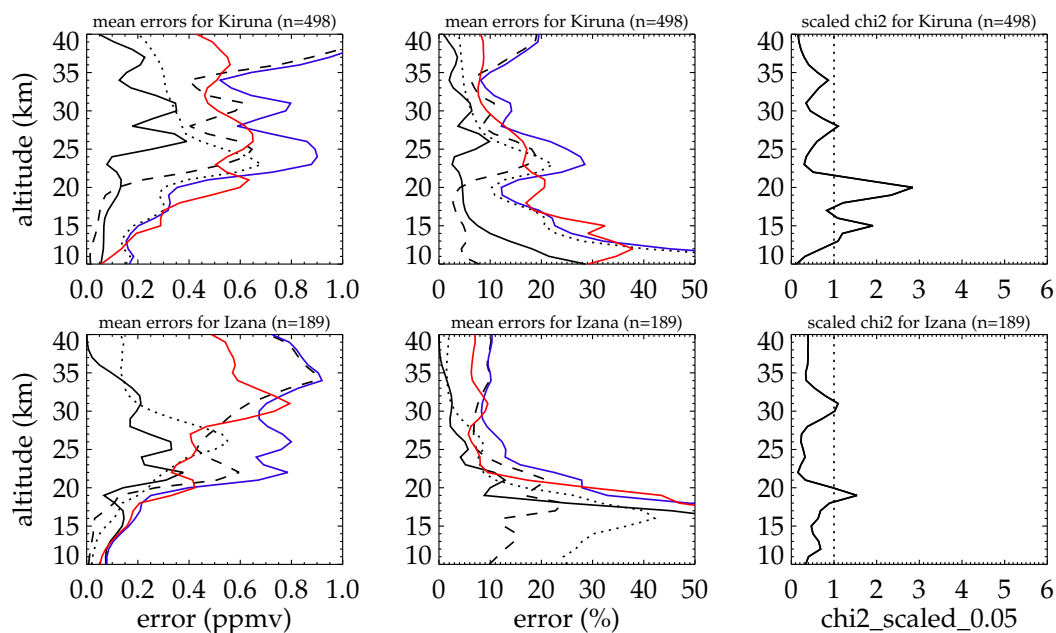
Fig. 15. Analysis of MIPAS and FTIR random errors. From top to bottom (number of coincidences in brackets): Kiruna (498) and Izana (189). From left to right: see Fig. 14.
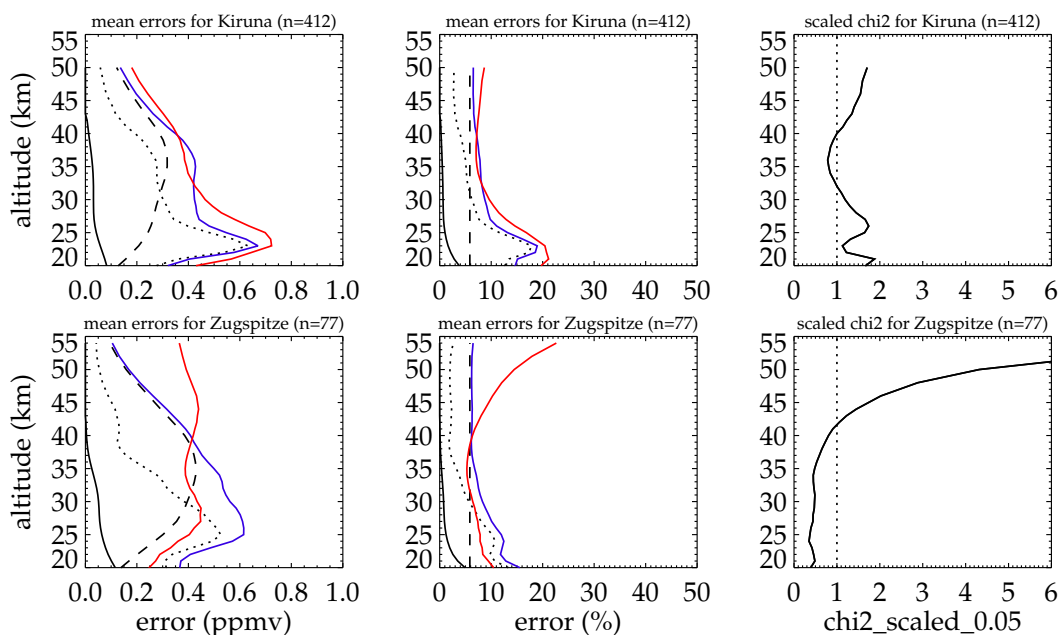

Fig. 16. Analysis of MIPAS and microwave random errors. From top to bottom (number of coincidences in brackets): Kiruna (412) and Zugspitze (77). From left to right: see Fig. 14.

altitudes. At 15 and 20 km the scaled χ^2 values are larger than 1. PSC affected MIPAS retrievals in the polar vortex region are the most probable reasons for this discrepancy. We also tested the inclusion of potential vorticity as additional coincidence criteria. However, the reduction of de-biased rms differences was only marginal (not shown).

For Izana (Fig. 15, bottom), the agreement between de-biased rms difference and estimated random error is good except for the tropopause region. The larger discrepancies there are probably due to the difficult retrieval situation to get the tropical tropopause correctly. In fact both random error components of MIPAS and FTIR are peaking 2 km higher
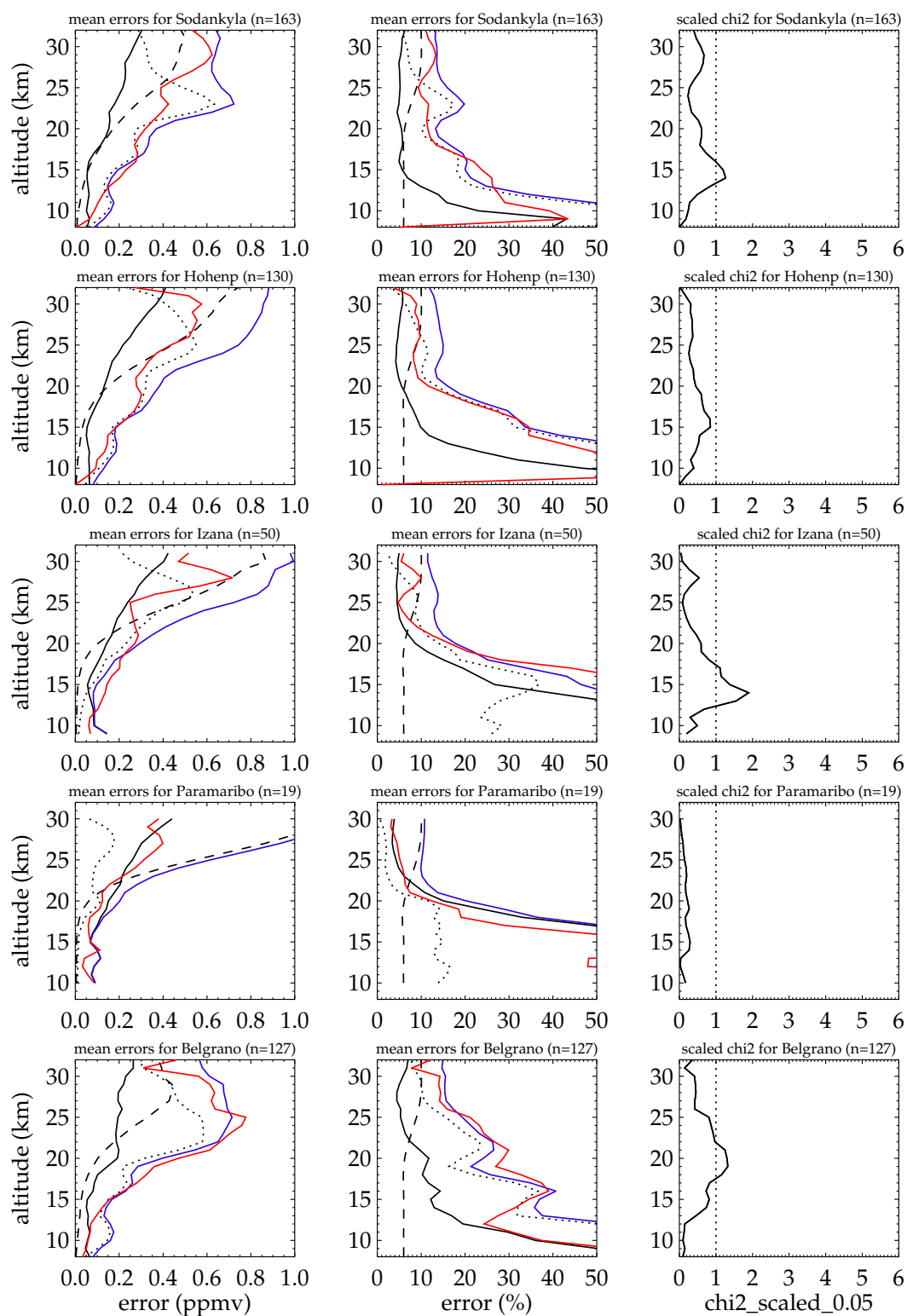
Fig. 17. Analysis of MIPAS and ozonesonde random errors. From top to bottom (number of coincidences in brackets): Sodankyla (163), Hohenpeissenberg (130), Izana (50), Paramaribo (19), and Belgrano (127). From left to right: see Fig. 14.
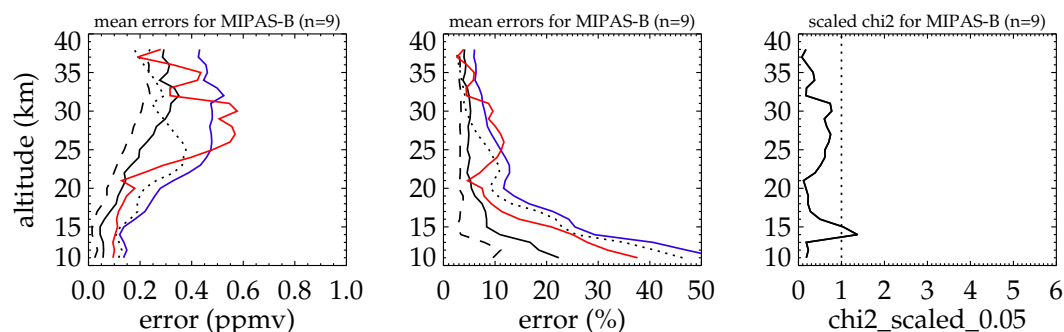
Fig. 18. Analysis of MIPAS-Envisat and MIPAS-B random ozone errors (9 coincidences). From left to right: see Fig. 14.

than the de-biased rms difference. Without this shift in altitude, the scaled χ^2 values would be almost always below 1. Above 35 km, the percentage random error of MIPAS is very small. This can be attributed to the strong influence of FTIR averaging kernels there.

5.1.3 Microwave

The analysis of MIPAS and microwave random errors and de-biased rms difference is displayed in Fig. 16. The microwave averaging kernel has a strong impact on MIPAS errors, which are basically zero above 40 km for both locations (solid black line in Fig. 16 left). The microwave random error is estimated to be 6% and constant for all altitudes.

For Kiruna (Fig. 16, top), the de-biased rms difference agrees well with the estimated random error in the region of the ozone maximum. Below about 30 km the de-biased rms difference is increasing slightly stronger than the estimated random error. For Zugspitze (Fig. 16, bottom) the agreement is very good below 42 km altitude with scaled χ^2 values smaller than 1. Above 42 km altitude, the percentage rms difference increases which can be attributed to diurnal variations in ozone. Microwave measurements have been performed mainly during daytime, whereas MIPAS measurements also have significant nighttime contributions with enhanced ozone values above about 50 km leading to a large variability.

5.2 Validation with balloon-borne instruments

5.2.1 Ozonesondes

The analysis of MIPAS and ozonesonde random errors is displayed in Fig. 17. For Sodankyla (Fig. 17, top), the de-biased rms difference exceeds the estimated random error only around 15 km altitude. For Belgrano (Fig. 17, bottom) the scaled χ^2 value exceeds 1 at 20 km which can be attributed again to the influence of PSCs on MIPAS ozone retrievals. For both polar conditions the error due to imperfect coincidence plays a very important role for the total estimated random error.

The results for Hohenpeissenberg (Fig. 17, second row) are even better than for the Sodankyla case. For Izana (Fig. 17, third row) there are discrepancies around 15 km altitude. Here very thin cirrus clouds not detected in the MIPAS cloud filter are at least partly responsible for the larger scaled χ^2 values. Since the MIPAS random error below 15 km is dominated by uncertainties in the water vapor amount (Fig. 3), scaled χ^2 values larger than 1 can also originate from underestimation of actual water vapor uncertainties.

For Paramaribo (Fig. 17, second from bottom) the estimated random error exceed the de-biased rms differences at all altitudes. The error estimates for the two instruments seem to be too high for this geolocation.

5.2.2 MIPAS-B

The analysis of MIPAS-Envisat and MIPAS-B ozone random errors is displayed in Fig. 18. The estimated random error and the de-biased rms differences agree very well. The scaled χ^2 for the confidence limit of 95% is exceeded at 15 km only. Occasional high χ^2 values, however, give no evidence of substantial disagreement but are explained by the χ^2 probability distribution.

5.3 Validation with satellite instruments

5.3.1 HALOE

The analysis of MIPAS-Envisat and HALOE ozone random errors is displayed in Fig. 19. The estimated random error and the de-biased rms differences agree well in the region of the ozone maximum for all latitudes. The largest rms differences are visible at 50 km altitude for southern midlatitude conditions. It is very likely that the diurnal variation of ozone is the reason for this discrepancy. Horizontal gradients can be very large at these altitudes for occultation measurements like HALOE. As described in Natarajan et al. (2005), twilight gradients have an impact in the order of 20% on ozone near 0.1 hPa. Furthermore, MIPAS usually measures during daytime or nighttime. This means that nighttime enhanced ozone
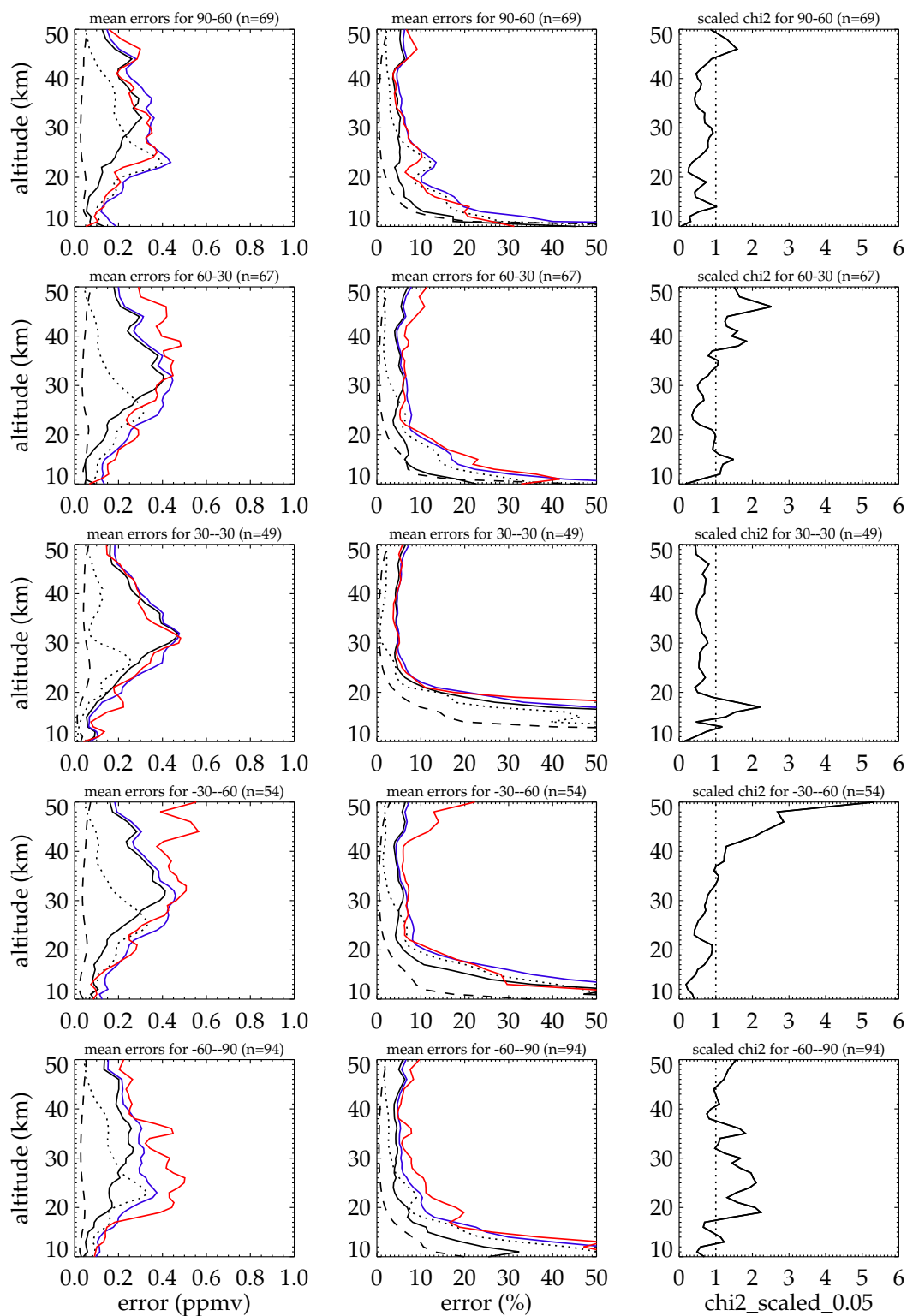
Fig. 19. Analysis of MIPAS and HALOE ozone random errors for 5 latitude bins. From top to bottom (number of coincidences in brackets): 90° N to 60° N (69), 60° N to 30° N (67), 30° N to 30° S (49), 30° S to 60° S (54), and 60° S to 90° S (94). From left to right: see Fig. 14.
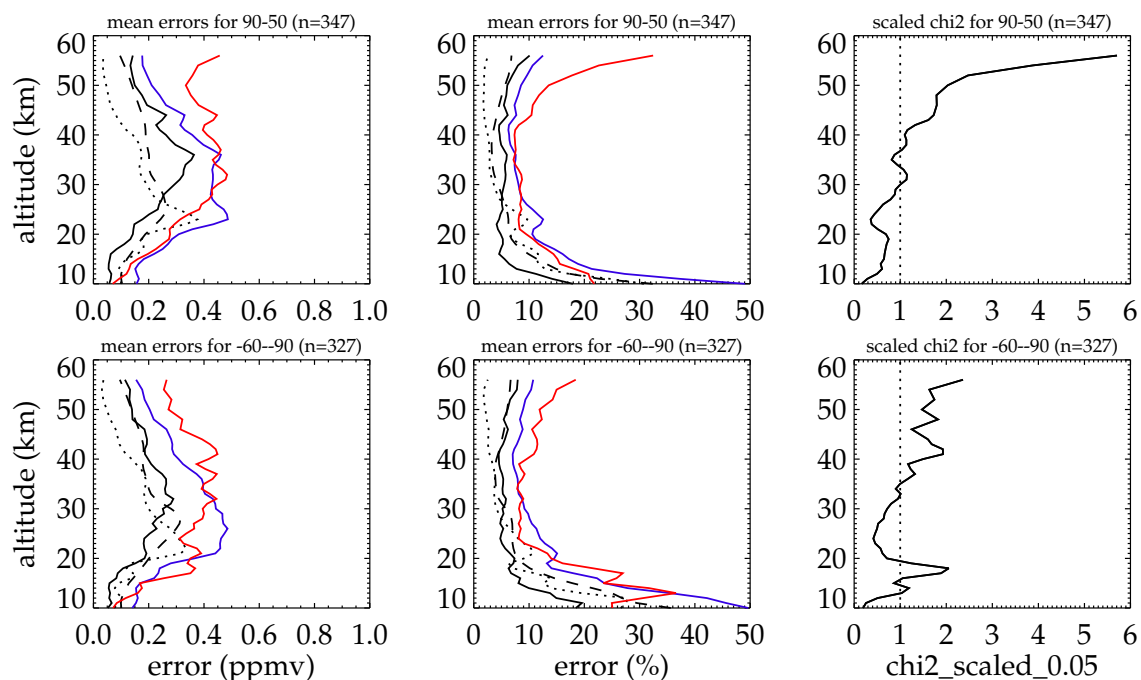
Fig. 20. Analysis of MIPAS and POAM ozone random errors for 2 latitude bins. From top to bottom (number of coincidences in brackets): 90° N to 60° N (347) and 60° S to 90° S (327). From left to right: see Fig. 14.

produces a large natural variability in the comparison not reflected by the error assessment. Division between day- and nighttime measurements of MIPAS demonstrates the impact on the de-biased rms differences (not shown). Especially MIPAS daytime measurements agree with HALOE well within the error estimates.

For polar conditions (Fig. 19, top and bottom) and lower altitudes, we find a similar behaviour of the scaled χ^2 values as for the ozonesondes Sodankyla and Belgrano (Fig. 17, top and bottom). Larger values can be found also above 20 km where PSCs can affect MIPAS ozone retrievals. Both, the criteria for cloud detection and the estimation of random errors may need adaption in particular for southern polar vortex conditions. Furthermore we find a dependence with seasons. The polar cases contain only few comparisons for winter conditions due to the occultation measurement technique of HALOE. For polar summer conditions, the estimated random error and the de-biased rms differences agree very well for all altitudes. The large values for the rms difference originates mainly from spring and autumn conditions where the natural variability is very high.

The inclusion of potential vorticity as additional coincidence criteria reduces the de-biased rms differences only slightly (not shown). This supports the statement that mainly undetected thin clouds in MIPAS retrievals are the reason for the discrepancy between estimated random error and de-biased rms difference.

For tropical conditions (Fig. 19, middle), we see the best agreement between estimated error and rms difference. For both midlatitude bands we also find good agreement with some larger scaled χ^2 values below 22 km. The larger discrepancies also have a seasonal component (not shown). The de-biased rms differences is largest for midlatitude winter conditions in both hemispheres where the variability of ozone is very large. As mentioned in Sect. 5.2.1, the influence of water vapor on the estimated ozone error may have been underestimated in the region around the tropopause and can contribute to χ^2 values larger than 1.

5.3.2 POAM III

The analysis of MIPAS-Envisat and POAM ozone random errors is displayed in Fig. 20. The estimated random error and the de-biased rms differences agree well for most of the altitudes. Above 50 km at northern latitudes (Fig. 20, top), the estimated errors are exceeding the rms difference considerably. As for HALOE, these discrepancies can mainly be attributed to diurnal variations in ozone. The effects which lead to larger discrepancies around 17 km for both latitude bands have been described for HALOE and are due to high undetected clouds in the MIPAS retrieval.

5.4 Summary

Total estimated random error, which contains MIPAS random error, random error of the correlative measurement, and

mismatch error, and de-biased rms difference show good agreement for all instruments and all latitudes particularly in the region of the ozone maximum. For tropical conditions we find good agreement at almost all altitudes. Some larger discrepancies occur below about 25 km for midlatitudinal and polar regions, where MIPAS ozone retrievals can be affected by undetected clouds. MIPAS random errors in these regions seem to be slightly underestimated caused by too small water vapor uncertainties. The error due to imperfect coincidence (mismatch error) plays an important role except for tropical conditions.

From the error analysis plots, MIPAS precision can be estimated to be between 5 and 10% for an altitude region of about 20 to 55 km. Above 55 km the precision is degrading towards values between 15 and 20%. Below 20 km the precision is in the order of 0.1 ppmv which gives percentage values between about 20 and 50% depending on the latitude.

6 Summary and conclusions

MIPAS ozone profiles have been compared to and validated with a large number of correlative measurements of different type and geolocation. Correlative instruments comprise ground-based lidars, Fourier transform infrared spectrometers, and microwave radiometers as well as balloon-borne ozonesondes, MIPAS-B, and satellite instruments (HALOE, POAM III) allowing comparisons with global coverage.

We have described the MIPAS ozone data version V3O_O3_7 with its errors and vertical resolution and performed a comparison with older data versions (V1_O3_2 and V1_O3_2) which use stronger regularization. Mean differences are below ± 0.2 ppmv between 10 and 56 km altitude and percentage mean differences are mainly below $\pm 3\%$ between 19 and 54 km.

The overall bias assessment of MIPAS (Fig. 13) shows good agreement between MIPAS and the other instrument types and good data quality. Mean differences are within ± 0.5 ppmv and mean percentage differences vary mainly between $\pm 10\%$ above 18 km altitude which is in agreement with the estimated MIPAS systematic error. For tropical conditions, percentage differences can be much larger. This has been attributed to a particular spectral analysis window used in the MIPAS retrieval which will no longer be used from version V3O_O3_9 on. The too small ozone values around 32 km, which have only little impact on a global scale, can be attributed to a problem in the strategy of atmospheric continuum emission retrieval, which is understood and will be solved (Glatthor et al., 2006). Below 18 km, mean differences are between -0.1 ppmv and 0.2 ppmv and percentage values are between -20% and 30%. Around 15 km altitude, MIPAS ozone is higher compared to most other instruments, which is likely a compensation effect for the negative kink around 18 km altitude.

The precision validation shows very good agreement between estimated random error and de-biased rms difference. Larger discrepancies at altitudes above 50 km are attributed to diurnal variations of ozone in MIPAS measurements which are only slightly present in the comparison instruments due to the solar occultation measurement technique of HALOE and POAM III. Discrepancies below 25 km can mainly be attributed to thin clouds which are not detected by the cloud filter and affect MIPAS ozone retrievals. Underestimation of water vapor induced error is also likely to contribute to these discrepancies. The error due to imperfect spatial-temporal coincidence plays an important role particularly at altitudes below 30 km for polar and midlatitude conditions. From the error analysis plots, MIPAS precision can be estimated to be between 5 and 10%, corresponding to vmr values of 0.15 ppmv and 0.4 ppmv, for an altitude region of about 20 to 55 km. Above 55 km and below 20 km the precision is degrading towards values of 20%. In tropical conditions the precision can reach percentage values in the order of 50%, whereas the absolute values stay around 0.1 ppmv.

Together, bias determination and precision validation suggests an accuracy of MIPAS ozone retrievals of 15 to 20% between 20 and 55 km altitude. Below 20 km MIPAS accuracy is degrading to values of 20 to 50% depending on altitude and latitude. The validation of the operational ozone product (Cortesi et al., 2007) generally shows degradation in the agreement with comparison data in both bias and precision at altitudes below 20 to 25 km. In comparison to that, the IMK-IAA ozone accuracy below 20 km altitude is also degrading but the data show better agreement with comparison instruments. This is in particular the case for precision validation.

References

Bracher, A., Bovensmann, H., Bramstedt, K., Burrows, J. P., von Clarmann, T., Eichmann, K.-U., Fischer, H., Funke, B., Gil-López, S., Glatthor, N., Grabowski, U., Höpfner, M., Kaufmann, M., Kellmann, S., Kiefer, M., Koukouli, M. E., Linden, A., López-Puertas, M., Mengistu Tsidu, G., Milz, M., Noël, S., Rohen, G., Rozanov, A., Rozanov, V. V., von Savigny, C., Sinnhuber, M., Skupin, J., Steck, T., Stiller, G. P., Wang, D.-Y., Weber, M., and Wuttke, M. W.: Cross comparisons of O$_3$ and NO$_2$ measured by the atmospheric ENVISAT instruments GOMOS, MIPAS, and SCIAMACHY, Adv. Space Res., 36, 855–867, doi: 10.1016/j.asr.2005.04.005, 2005.

Brühl, C., Drayson, S. R., Russell III, J. M., Crutzen, P. J., McInerney, J. M., Purcell, P. N., Claude, H., Gernandt, H., McGee, T. J., McDermid, I. S., and Gunson, M. R.: Halogen Occulta-

tion Experiment ozone channel validation, J. Geophys. Res., 101, 10 217–10 240, 1996.

Cortesi, U., Lambert, J. C., De Clercq, C., Bianchini, G., Blumenstock, T., Bracher, A., Castelli, E., Catoire, V., Chance, K. V., De Mazière, M., Demoulin, P., Godin-Beekmann, S., Jones, N., Jucks, K., Keim, C., Kerzenmacher, T., Kuellmann, H., Kuttippurath, J., Iarlori, M., Liu, G. Y., Liu, Y., McDermid, I. S., Meijer, Y. J., Mencaraglia, F., Mikuteit, S., Oelhaf, H., Piccolo, C., Pirre, M., Raspollini, P., Ravegnani, F., Reburn, W. J., Redaelli, G., Remedios, J. J., Sembhi, H., Smale, D., Steck, T., Taddei, A., Varotsos, C., Vigouroux, C., Waterfall, A., Wetzel, G., and Wood, W.: Geophysical validation of MIPAS-ENVISAT operational ozone data, Atmos. Chem. Phys. Discuss., 7, 5805–5939, 2007, http://www.atmos-chem-phys-discuss.net/7/5805/2007/.

Echle, G., von Clarmann, T., Dudhia, A., Flaud, J.-M., Funke, B., Glatthor, N., Kerridge, B., López-Puertas, M., Martín-Torres, F. J., and Stiller, G. P.: Optimized spectral microwindows for data analysis of the Michelson Interferometer for Passive Atmospheric Sounding on the Environmental Satellite, Appl. Opt., 39, 5531–5540, 2000.

Fischer, H. and Oelhaf, H.: Remote sensing of vertical profiles of atmospheric trace constituents with MIPAS limb-emission spectrometers, Appl. Opt., 35, 2787–2796, 1996.

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., Koopmann, 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, Atmos. Chem. Phys. Discuss., 7, 8795–8893, 2007, http://www.atmos-chem-phys-discuss.net/7/8795/2007/.

Friedl-Vallon, F., Maucher, G., Kleinert, A., Lengel, A., Keim, C., Oelhaf, H., Fischer, H., Seefeldner, M., and Trieschmann, O.: Design and characterisation of the ballon–borne Michelson Interferometer for Passive Atmospheric Sounding (MIPAS–B2), Appl. Opt., 43, 3335–3355, 2004.

Glatthor, N., von Clarmann, T., Fischer, H., Funke, B., Gil-López, S., Grabowski, U., Höpfner, M., Kellmann, S., Linden, A., López-Puertas, M., Mengistu Tsidu, G., Milz, M., Steck, T., Stiller, G. P., and Wang, D.-Y.: Retrieval of stratospheric ozone profiles from MIPAS/ENVISAT limb emission spectra: a sensitivity study, Atmos. Chem. Phys., 6, 2767–2781, 2006, http://www.atmos-chem-phys.net/6/2767/2006/.

Godin, S., Carswell, A. I., Donovan, D. P., Claude, H., Steinbrecht, W., McDermid, I. S., McGee, T. J., Gross, M. R., Nakane, H., Swart, D. P. J., Bergwerff, H. B., Uchino, O., von der Gathen, P., and Neuber, R.: Ozone differential absorption lidar algorithm intercomparison, Appl. Opt., 30, 6225–6236, 1999.

Gosset, W. S.: The probable error of a mean, Biometrika, 6, 1–25, 1908.

Kleinert, A., Aubertin, G., Perron, G., Birk, M., Wagner, G., Hase, F., Nett, H., and Poulin, R.: MIPAS Level 1B algorithms overview: operational processing and characterization, Atmos. Chem. Phys., 7, 1395–1406, 2007, http://www.atmos-chem-phys.net/7/1395/2007/.

Kopp, G., Berg, H., Blumenstock, T., Fischer, H., Hase, F., Hochschild, G., Höpfner, M., Kouker, W., Reddmann, T., Ruhnke, R., Raffalski, U., and Kondo, Y.: Evolution

of ozone and ozone-related species over Kiruna during the SOLVE/THESEO 2000 campaign retrieved from ground-based millimeter-wave and infrared observations, J. Geophys. Res., 107, 8308, doi:10.1029/2001JD001064, 2002.

Lucke, R. L., Korwan, D. R., Bevilacqua, R. M., Hornstein, J. S., Shettle, E. P., Chen, D. T., Daehler, M., Lumpe, J. D., Fromm, M. D., Debrestian, D., Neff, B., Squire, M., König-Langlo, G., and Davies, J.: The Polar Ozone and Aerosol Measurement (POAM) III instrument and early validation results, J. Geophys. Res., 104, 18 785–18 799, 1999.

Lumpe, J. D., Bevilacqua, R. M., Hoppel, K. W., and Randall, C. E.: POAM III retrieval algorithm and error analysis, J. Geophys. Res., 107, 4575, doi:10.1029/2002JD002137, 2002.

Migliorini, S., Piccolo, C., and Rodgers, C. D.: Intercomparison of direct and indirect measurements: Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) versus sonde ozone profiles, J. Geophys. Res., 109, D19316, doi:10.1029/2004JD004988, 2004.

Natarajan, M., Deaver, L. E., Thompson, E., and Magill, B.: Impact of twilight gradients on the retrieval of mesospheric ozone from HALOE, J. Geophys. Res., 110, D13305, doi:10.1029/2004JD005719, 2005.

Nett, H., Perron, G., Sanchez, M., Burgess, A., and Mossner, P.: MIPAS inflight calibration and processor validation, in: ENVISAT Calibration Review – Proc. of the European Workshop, 9–13 September 2002, ESTEC, Noordwijk, The Netherlands, CD-ROM, vol. SP-520, edited by: Sawaya-Lacoste, H., ESA Publications Division, ESTEC, P.O. Box 299, 2200 AG Noordwijk, The Netherlands, 2002.

Raffalski, U., Hochschild, G., Kopp, G., and Urban, J.: Evolution of stratospheric ozone during winter 2002/2003 as observed by a ground based millimetre wave radiometer at Kiruna, Sweden, Atmos. Chem. Phys., 5, 1399–1407, 2005, http://www.atmos-chem-phys.net/5/1399/2005/.

Randall, C. E., Rusch, D. W., Bevilacqua, R. M., Hoppel, K. W., Lumpe, J. D., Shettle, E., Thompson, E., Deaver, L., Zawodny, J., Kyrö, E., Johnson, B., Kelder, H., Dorokhov, V. M., König-Langlo, G., and Gil, M.: Validation of POAM III ozone: Comparisons with ozonesonde and satellite data, J. Geophys. Res., 108, 4367, doi:10.1029/2002JD002944, 2003.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics, edited by: Taylor, F. W., World Scientific, 2000.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, doi:10.1029/2002JD002299, 2003.

Russell III, J. M., Gordley, L. L., Park, J. H., Drayson, S. R., Hesketh, W. D., Cicerone, R. J., Tuck, A. F., Frederick, J. E., Harries, J. E., and Crutzen, P. J.: The Halogen Occultation Experiment, J. Geophys. Res., 98, 10 777–10 797, 1993.

Schneider, M., Blumenstock, T., Hase, F., Höpfner, M., Cuevas, E., Redondas, A., and Sancho, J.: Ozone profiles and total column amounts derived at Izaña, Tenerife Island, from FTIR solar absorption spectra, and its validation by an intercomparison to ECC-sonde and Brewer spectrometer measurements, J. Quant. Spectrosc. Radiat. Transfer, 91, 245–274, doi:10.1016/j.jqsrt.2004.05.067, 2004.

Smit, H. G. J. and Straeter, W.: JOSIE-2000, Jülich Ozone

Sonde Intercomparison Experiment 2000, The 2000 WMO international intercomparison of operating procedures for ECC-ozone sondes at the environmental simulation facility at Jülich, WMO Global Atmosphere Watch report series, No. 158 (Technical Document No. 1225), World Meteorological Organization, Geneva, 2004.

Spang, R., Remedios, J. J., and Barkley, M. P.: Colour indices for the detection and differentiation of cloud types in infra-red limb emission spectra, Adv. Space Res., 3, 1041–1047, doi:10.1016/S0273-1177(03)00585-4, 2004.

Steck, T.: Methods for determining regularization for atmospheric retrieval problems, Appl. Opt., 41, 1788–1797, 2002.

Steinbrecht, W., Neuber, R., von der Gathen, P., Wahl, P., McGee, T. J., Gross, M. R., Klein, U., and Langer, J.: Results of the Ny-Ålesund ozone monitoring intercomparison, J. Geophys. Res., 104, 30 515–30 523, 1999.

von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, Atmos. Chem. Phys., 6, 4311–4320, 2006a.

von Clarmann, T.: Addendum to "Validation of remotely sensed profiles of atmospheric state varaiables: strategies and terminology" published in Atmos. Chem. Phys., 6, 4311–4320, 2006, Atmos. Chem. Phys., 6, 5547–5547, 2006b.

von Clarmann, T. and Grabowski, U.: Elimination of hidden a priori information from remotely sensed profile data, Atmos. Chem. Phys., 7, 397–408, 2007,
http://www.atmos-chem-phys.net/7/397/2007/.

von Clarmann, T., Ceccherini, S., Doicu, A., Dudhia, A., Funke, B., Grabowski, U., Hilgers, S., Jay, V., Linden, A., López-Puertas, M., Martín-Torres, F.-J., Payne, V., Reburn, J., Ridolfi, M., Schreier, F., Schwarz, G., Siddans, R., and Steck, T.: A blind test retrieval experiment for infrared limb emission spectrometry, J. Geophys. Res., 108, 4746, doi:10.1029/2003JD003835, 2003a.

von Clarmann, T., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., 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), J. Geophys. Res., 108, 4736, doi:10.1029/2003JD003602, 2003b.

Wang, D.-Y., Stiller, G. P., von Clarmann, T., Fischer, H., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Milz, M., Steck, T., Wohnsiedler, S., López-Puertas, M., Funke, B., Gil-López, S., Kaufmann, M., Koukouli, M. E., Murtagh, D., Lautié, N., Jiménez, C., Jones, A., Eriksson, P., Urban, J., de La Noë, J., Le Flochmoën, É., Dupuy, É., Ricaud, P., Olberg, M., Frisk, U., Russell III, J., and Remsberg, E.: Comparisons of MIPAS/ENVISAT ozone profiles with SMR/ODIN and HALOE/UARS observations, Adv. Space Res., 36, 927–931, doi:10.1016/j.asr.2005.03.015, 2005.