



**HAL**  
open science

## Geophysical validation of temperature retrieved by the ESA processor from MIPAS/ENVISAT atmospheric limb-emission measurements

M. Ridolfi, U. Blum, B. Carli, Valéry Catoire, S. Ceccherini, H. Claude, C. de Clercq, K. H. Fricke, F. Friedl-Vallon, M. Iarlori, et al.

► **To cite this version:**

M. Ridolfi, U. Blum, B. Carli, Valéry Catoire, S. Ceccherini, et al.. Geophysical validation of temperature retrieved by the ESA processor from MIPAS/ENVISAT atmospheric limb-emission measurements. *Atmospheric Chemistry and Physics*, 2007, 7 (16), pp.4459-4487. 10.5194/acp-7-4459-2007. hal-00328526

**HAL Id: hal-00328526**

**<https://hal.science/hal-00328526>**

Submitted on 18 Jun 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Geophysical validation of temperature retrieved by the ESA processor from MIPAS/ENVISAT atmospheric limb-emission measurements

M. Ridolfi<sup>1</sup>, U. Blum<sup>2</sup>, B. Carli<sup>3</sup>, V. Catoire<sup>4</sup>, S. Ceccherini<sup>3</sup>, H. Claude<sup>5</sup>, C. De Clercq<sup>6</sup>, K. H. Fricke<sup>7</sup>, F. Friedl-Vallon<sup>8</sup>, M. Iarlori<sup>9</sup>, P. Keckhut<sup>10</sup>, B. Kerridge<sup>11</sup>, J.-C. Lambert<sup>6</sup>, Y. J. Meijer<sup>12</sup>, L. Mona<sup>13</sup>, H. Oelhaf<sup>8</sup>, G. Pappalardo<sup>13</sup>, M. Pirre<sup>4</sup>, V. Rizi<sup>9</sup>, C. Robert<sup>4</sup>, D. Swart<sup>12</sup>, T. von Clarmann<sup>8</sup>, A. Waterfall<sup>11</sup>, and G. Wetzel<sup>8</sup>

<sup>1</sup>Dipartimento di Chimica Fisica e Inorganica, Università di Bologna, Italy

<sup>2</sup>Fraunhofer-Institut für Naturwissenschaftlich-Technische Trendanalysen, Euskirchen, Germany

<sup>3</sup>Istituto di Fisica Applicata “Nello Carrara”, Consiglio Nazionale delle Ricerche, Firenze, Italy

<sup>4</sup>Laboratoire de Physique et Chimie de l’Environnement, CNRS – Université d’Orleans, Orleans, France

<sup>5</sup>DWD, Observatory Hohenpeissenberg, Germany

<sup>6</sup>Institut d’Aéronomie Spatiale de Belgique, Bruxelles, Belgique

<sup>7</sup>Physikalisches Institut, Universität Bonn, Bonn, Germany

<sup>8</sup>Forschungszentrum Karlsruhe, IMK, Germany

<sup>9</sup>CETEMPS – Dipartimento di Fisica, Università de L’Aquila, L’Aquila, Italy

<sup>10</sup>Service d’Aéronomie, Institut Pierre Simon Laplace/UVSQ, Verrieres-Le-Buisson, France

<sup>11</sup>Earth Observation and Atmospheric Science, Space Science and Technology Department, Rutherford Appleton Laboratory, Oxfordshire, UK

<sup>12</sup>National Institute for Public Health and the Environment, RIVM – LVM, Bilthoven, The Netherlands

<sup>13</sup>Istituto di Metodologie per l’Analisi Ambientale, Consiglio Nazionale delle Ricerche, Tito Scalo, Potenza, Italy

Received: 5 April 2007 – Published in Atmos. Chem. Phys. Discuss.: 25 April 2007

Revised: 13 July 2007 – Accepted: 13 July 2007 – Published: 24 August 2007

**Abstract.** The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) has been operating since March 2002 onboard of the ENVIRONMENTAL SATellite of the European Space Agency (ESA). The high resolution (0.035 cm<sup>-1</sup> full width half maximum, unapodized) limb-emission measurements acquired by MIPAS in the first two years of operation have very good geographical and temporal coverage and have been re-processed by ESA with the most recent versions (4.61 and 4.62) of the inversion algorithms. The products of this processing chain are pressures at the tangent points and geolocated profiles of temperature and of the volume mixing ratios of six key atmospheric constituents: H<sub>2</sub>O, O<sub>3</sub>, HNO<sub>3</sub>, CH<sub>4</sub>, N<sub>2</sub>O and NO<sub>2</sub>. As for all the measurements made with innovative instruments and techniques, this data set requires a thorough validation. In this paper we present a geophysical validation of the temperature profiles derived from MIPAS measurements by the ESA retrieval algorithm. The validation is carried-out by comparing MIPAS temperature with

correlative measurements made by radiosondes, lidars, in-situ and remote sensors operated either from the ground or stratospheric balloons.

The results of the intercomparison indicate that the bias of the MIPAS profiles is generally smaller than 1 or 2 K depending on altitude. Furthermore we find that, especially at the edges of the altitude range covered by the MIPAS scan, the random error estimated from the intercomparison is larger (typically by a factor of two to three) than the corresponding estimate derived on the basis of error propagation.

In this work we also characterize the discrepancies between MIPAS temperature and the temperature fields resulting from the analyses of the European Centre for Medium-range Weather Forecasts (ECMWF). The bias and the standard deviation of these discrepancies are consistent with those obtained when comparing MIPAS to correlative measurements; however, in this case the detected bias has a peculiar behavior as a function of altitude. This behavior is very similar to that observed in previous studies and is suspected to be due to vertical oscillations in the ECMWF temperature.

Correspondence to: M.Ridolfi  
(Marco.Ridolfi@unibo.it)

The current understanding is that, at least in the upper stratosphere (above  $\approx 10$  hPa), these oscillations are caused by a discrepancy between model biases and biases of assimilated radiances from primarily nadir sounders.

## 1 Introduction

MIPAS (Michelson Interferometer for Passive Atmospheric Sounding, Fischer and Oelhaf 1996; Fischer et al. 2000) is a Fourier transform spectrometer, operating onboard of EN-VISAT, a satellite launched by the European Space Agency (ESA) on 1 March 2002 in a polar orbit with  $\approx 98$  deg. inclination. MIPAS measures the atmospheric limb-emission spectrum in the middle infrared (from 685 to 2410  $\text{cm}^{-1}$ ), a spectral region containing the signatures of the vibrational transitions of many atmospheric constituents. In the period from July 2002 to March 2004 the instrument operated mostly in the so called *nominal mode*. In this time frame the nominal operation mode consisted of measurements at high spectral resolution (0.035  $\text{cm}^{-1}$  full width half maximum, unapodized) of limb-scans in the altitude range from 6 to 68 km, with 3 km steps from 6 to 42 km and with 5 and 8 km steps above 42 km. These measurements have very good geographical and time coverage and have been re-processed by ESA with the most recent versions (4.61 and 4.62) of both Level 1b (Kleinert et al., 2007) and Level 2 (Ridolfi et al., 2000; Raspollini et al., 2006) algorithms. The products of the ESA Level 2 algorithm are pressures at the tangent points and geolocated profiles of temperature and of Volume Mixing Ratios (VMR) of six key atmospheric constituents:  $\text{H}_2\text{O}$ ,  $\text{O}_3$ ,  $\text{HNO}_3$ ,  $\text{CH}_4$ ,  $\text{N}_2\text{O}$  and  $\text{NO}_2$ . As for all the measurements made with innovative instruments and techniques, this data set requires a thorough validation. In this paper we focus on the validation of temperature. Besides its importance for atmospheric studies, in the case of MIPAS data processing, temperature also plays a key role in the determination of the accuracy of the atmospheric constituents that are retrieved in sequence after temperature, using this latter as an input (Raspollini and Ridolfi, 2000).

The validation has been carried out by comparing MIPAS retrieved temperature with correlative measurements made by radiosondes, lidars, in-situ and remote sensors operated either from the ground or stratospheric balloons.

As already pointed out by several authors (see e.g. Rodgers and Connor 2003; Ceccherini et al. 2003; Ridolfi et al. 2006), the intercomparison of measurements acquired by instruments that use different techniques, and therefore have different response functions to the real state of the atmosphere, is not a trivial task and requires specific methods to be used; therefore in Sect. 2 we provide an outline of the methods adopted for MIPAS temperature validation. In Sects. 3 to 7 we illustrate and discuss the results of the validation activities of the individual participating validation teams. In Sect. 8 we

summarize the results of the findings, and finally, in Sect. 9 we draw the conclusions.

## 2 General aspects of MIPAS validation

In this section we discuss a few general issues connected with the validation of MIPAS profiles. The specific validation approaches used by the individual validation teams are discussed later, in dedicated sections.

### 2.1 Spatial response function of the measurements

Remote emission measurements operated from satellites offer the great advantage of providing measurements with very good geographical and time coverage. Remote measurements, however, show characteristics that are intrinsically different from those of in-situ sounders that provide *local* and often *direct* measurements of the quantities of interest, as e.g. temperature or VMR of atmospheric constituents. Remote sensors do not measure directly the geophysical quantities of interest, but a complicated function of them (the limb-emission spectrum in the case of MIPAS); therefore these measurements must undergo an inversion process to extract the desired information. Both the characteristics of the inversion algorithm and instrument-specific features such as the instantaneous field of view, the sampling grid and the spectral resolution contribute to making the shape of the measurement spatial response a complex function.

Assuming the availability of well characterized correlative *reference* measurements, the intercomparison provides insight at two different levels: firstly, the statistical analysis of the discrepancies between MIPAS and reference measurements can be used to characterize both the bias and the precision of MIPAS (von Clarmann, 2006). The bias and precision estimates obtained from the intercomparison can be then compared with the available estimates based on error propagation analyses, hence corroborating them or raising question marks regarding their reliability. Secondly, if the intercomparison is carried out by avoiding or minimizing the known discrepancies due to the differences in the response functions of the intercompared measurements (Rodgers and Connor, 2003; Ceccherini et al., 2003; Ridolfi et al., 2006), the error budget of the profile differences is significantly reduced and it is then possible to investigate the remaining discrepancies with improved accuracy.

Unfortunately, in several cases the intercomparison cannot proceed up to this second level because the spatial response functions, or *averaging kernels* (AKs), which are the linear approximation (Rodgers, 2000) of the considered measurements, are not accurately known. In the MIPAS case, the correlative temperature measurements considered in this validation work are of two types:

1. Measurements with a vertical resolution better than that of MIPAS ( $\approx 3$  km in the range from 6 to 42 km).

Radiosonde and in-situ measured profiles have very high vertical resolutions and fine sampling steps (of the order of 100 m). Lidar measurements also have a good vertical resolution (ranging from 150 m up to 1–2 km).

2. Measurements with a vertical resolution comparable to that of MIPAS. Balloon limb-sounding FTS measurements and analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF, 1995) belong to this group.

Different spatial response functions of the measurements considered induce a significant *smoothing error* (Rodgers and Connor, 2003; Ridolfi et al., 2006) on the difference profile. In the intercomparison we check if this difference is statistically consistent with zero. The smoothing error of the difference and the other error components add up quadratically to form the total error of the difference. For comparison with measurements of type 1 we adopt different strategies depending on whether reliable MIPAS AKs are available or not for the atmospheric conditions in which the intercomparison takes place. Whenever the available MIPAS AKs, which were derived for standard atmospheric conditions, are considered representative of the actual atmospheric state, the smoothing error of the difference is reduced by comparing the MIPAS profile with the corresponding correlative measurement adjusted to the MIPAS spatial smoothing,  $\hat{\mathbf{x}}_{\text{ref,smooth}}$ , calculated as:

$$\hat{\mathbf{x}}_{\text{ref,smooth}} = \hat{\mathbf{x}}_0 + \mathbf{A}(\mathbf{x}_{\text{ref}} - \mathbf{x}_0) = \mathbf{A}\mathbf{x}_{\text{ref}} + (\hat{\mathbf{x}}_0 - \mathbf{A}\mathbf{x}_0) \quad (1)$$

where  $\mathbf{x}_{\text{ref}}$  is the correlative measurement in the original altitude resolution,  $\mathbf{x}_0$  is a pre-defined atmospheric state used as a linearization point for the calculation of the MIPAS averaging kernel  $\mathbf{A}$ , and  $\hat{\mathbf{x}}_0$  is the result of the MIPAS retrieval when the true state of the atmosphere is  $\mathbf{x}_{\text{ref}} = \mathbf{x}_0$ . The altitude grid of the profile  $\hat{\mathbf{x}}_{\text{ref,smooth}}$  obtained from Eq. (1) and the grid of the available correlative MIPAS profile are then matched using either the shrinking/stretching and interpolation technique described in Raspollini et al. (2006) or the regridding technique of Calisesi et al. (2005).

If the available MIPAS AKs are not considered representative of the actual atmosphere encountered in the intercomparison, the profiles are directly compared by matching the pressure or altitude grids using linear interpolation in altitude or in log pressure, and calculating the difference. Whenever significant, the smoothing error is evaluated as part of the random error budget of this difference. This latter approach is also used to intercompare MIPAS with correlative measurements of type 2.

## 2.2 MIPAS vertical grid

In the case of MIPAS, the interpretation of the retrieved profiles involves an additional complication linked with the vertical scale. The accuracy of the instrument elevation pointing was extensively tested during the commissioning phase

(Kiefer et al., 2007) and it was found that, although rather stable ( $\approx 300$  m) within the time interval required for the measurement of a limb-scan ( $\approx 70$  s), in absolute terms the elevation pointing may be affected by errors as large as 2 km in tangent altitude. This feature of the MIPAS pointing was also expected on the basis of the pre-launch requirements; therefore the retrieval algorithm was designed to be, as much as possible, independent from accurate pointing knowledge. Specifically, the ESA inversion algorithm (Ridolfi et al., 2000; Raspollini et al., 2006) retrieves temperature simultaneously with pressure at the tangent points of the limb-observations, taking into account the a-priori pointing knowledge supplied by the engineering pointing system. The altitude scale is then re-constructed using the barometric equation, the retrieved pressure and temperature at the tangent points and the altitude of a *reference* tangent point. Of course, if the altitude of the selected reference tangent point is affected by an error, the altitude scale obtained with this method is going to be shifted by an amount equal to this error. This intrinsic weakness of the re-constructed altitude scale has led to a recommendation by ESA to use, whenever possible, the pressure scale associated with MIPAS profiles rather than the altitude scale. A more reliable altitude scale can be obtained a-posteriori, if independent sources of information (such as ECMWF analyses) supply an accurate pressure profile as a function of altitude, for the same time and geolocation as the MIPAS profile under consideration.

Given this feature of the MIPAS profiles, whenever correlative measurements are supplied on a pressure grid, we carry out the intercomparisons in the pressure domain. This approach, however, is not applicable to the intercomparisons with lidar profiles, which are intrinsically represented as a function of altitude (see Sect. 6). In these cases we carry out the intercomparisons using the MIPAS altitude grid corrected with customized procedures, as described in Sects. 6.1 and 6.2.

## 2.3 Coincidence criteria

Ideally, validation requires comparison of measurements relating to the same air mass. In practice, however, perfect spatial and temporal coincidence of the measurements is very difficult to achieve, and therefore a compromise must be found. The specific compromise established for the intercomparison is usually referred as the *coincidence criterion*. As far as temperature is concerned, a spatial distance of less than 300 km and a time difference shorter than 3 h are generally considered reasonable coincidence criteria (Fischer et al., 2007). In several cases, however, in our intercomparison work these baseline criteria turned out to be too conservative, reducing dramatically the number of matching pairs of measurements to be compared. In these cases the baseline coincidence criteria have been revised by the individual validating teams on the basis of test attempts or model

corrections aimed at a reduction or a better estimation of the mismatch error.

### 3 Validation using radiosonde measurements

Radiosondes are lightweight boxes carried aloft by small balloons up to altitudes of almost 35 km. The instruments on board include temperature and pressure sensors characterized by high precision and accuracy. In particular, the largest majority of radiosonde temperature profiles used in this validation were measured with balloonborne Väisälä RS80 and RS90 sondes (Antikainen and Turtiainen, 1992; Antikainen and Jauheainen, 1995), equipped with pressure and temperature sensors characterized as follows (Luers, 1997; WMO, 2004):

- Pressure sensor. In the range from 3 to 1060 hPa, precision: 0.1 hPa; accuracy: 0.5 hPa.
- Temperature sensor. In the range from  $-90$  to  $60^{\circ}\text{C}$ , precision: 0.1 K; the accuracy is pressure-dependent: 0.2 K from ground to 50 hPa, 0.3 K from 50 to 15 hPa and 0.4 K for pressures below 15 hPa.

The radiosonde samples the atmosphere approximately every 10 seconds; therefore, since the average ascent rate of the balloon is about 4 m/s, the average vertical sampling step of the sonde profiles is about 40 m. Since the radiosonde performs *local* measurements, the vertical resolution of the measured profiles coincides with the sampling step. Although temperatures at different altitudes are measured at different times, for simplicity we interpret each individual radiosonde profile as *instantaneous* and refer (if not otherwise specified) to the mean time of the measurement interval. Considering that the duration of a profile measurement is about 2 h and that the allowed time mismatch between compared measurements is always equal or greater than 3 h (see Sect. 2.3), this is a reasonable approximation.

#### 3.1 Comparison with radiosonde measurements from Potenza and L'Aquila (Italy)

Both the Istituto di Metodologie per l'Analisi Ambientale del Consiglio Nazionale delle Ricerche (CNR-IMAA), Potenza (Italy) and the CETEMPS of the Department of Physics of the University of L'Aquila (Italy) are involved in the Ground Based Measurement Campaign for the validation of the MIPAS Level 2 temperature and water vapor. Validation measurements are acquired by water vapor lidars complemented with radiosonde launches in order to measure atmospheric pressure, temperature and relative humidity. Radiosondes are launched from Potenza (Lat.  $40.60^{\circ}\text{N}$ , Lon.  $15.72^{\circ}\text{E}$ , Altitude ASL 760 m) and L'Aquila (Lat.  $42.38^{\circ}\text{N}$ , Lon.  $13.31^{\circ}\text{E}$ , Altitude ASL 683 m) sites. In the time frame from July 2002 to December 2002, two radiosoundings and two lidar measurements per week coincident with ENVISAT over-

passes were carried out. From January 2003 to July 2003, one radiosounding and one lidar measurement per week were carried out. After the end of the validation campaign, measurements have still been acquired on a regular basis (Pappalardo et al., 2005).

For MIPAS temperature validation, radiosonde profiles were used. These profiles were measured with balloonborne Väisälä RS80 sondes with the features specified in Sect. 3.

For this intercomparison we only considered radiosonde measurements within 300 km and 3 h of a MIPAS measured limb-scan. The errors due to the residual spatial and temporal mismatch between the measurements were accounted for by using the quantitative formal validation approach suggested by von Clarmann (2006). Here we use the terminology defined in von Clarmann (2006) and refer to the equations in that work as  $C_n$ , where  $n$  is the equation number.

Since for this set of data the intercomparisons take place at mid-latitudes, in atmospheric conditions not too far from those for which the standard MIPAS vertical averaging kernels were derived (Raspollini et al., 2006), as a first step we allowed for the smoothing error of MIPAS by downgrading all the radiosonde temperature profiles  $t_s$  to the MIPAS altitude resolution, using Eq. (1). As we also use the shrinking/stretching and interpolation technique described by Raspollini et al. (2006), each profile  $\hat{t}_s$  resulting from this operation is represented on the same vertical grid as the matching MIPAS profile  $t_M$ .

The error due to less than perfect coincidence (both in space and time) of each pair of measurements was then corrected for using the ECMWF temperature analyses. In particular, each profile  $\hat{t}_s$  was corrected using the following expression (see Eq. (C15) and (C24)):

$$t_c = \hat{t}_s + \text{diag}[\mathbf{A}_h \mathbf{T}_{\text{EC},M}] - t_{\text{EC},s}. \quad (2)$$

In this expression  $t_{\text{EC},s}$  is the ECMWF temperature interpolated to the radiosonde time and geolocation, while  $\mathbf{A}_h$  is the horizontal MIPAS averaging kernel matrix. The element  $i, j$  of  $\mathbf{A}_h$  represents the response of the  $i$ -th MIPAS retrieved temperature grid point to an infinitesimal variation of the temperature at the location  $j$  along the MIPAS line of sight. The matrix  $\mathbf{A}_h$  was derived in the following approximation: each entry of the row of the MIPAS vertical averaging kernel was assigned to the geolocation where, according to ray tracing when refraction is considered, the line of sight crosses the respective altitude. In the approximation of a quasi-transparent atmosphere the same weight (i.e. half the value of the respective entry of the vertical averaging kernel) was given to air parcels at the same altitude in front of and behind the tangent point. Averaging kernel elements below the tangent altitude were assigned to the tangent point geolocation. In Eq. (2)  $\mathbf{T}_{\text{EC},M}$  is a matrix constructed as follows: each column corresponds to a MIPAS limb-view, and the entries in each column are obtained by interpolating the ECMWF fields to the MIPAS time and to the geolocation of the points along the MIPAS line of sight considered

**Table 1.** Summary of validation against radiosondes operated from Potenza and L'Aquila.

Altitude [km]	$n_i$	$b_i$ [K]	$\sigma_{b_i}$ [K]	$\sigma_{b_i,\text{sys}}$ [K]	$p_i$ [K]	$\sigma_{d_i,\text{rnd}}$ [K]	$\chi_{\text{R},i}^2$	$L_i$
0.–7.5	6	0.35	2.27	1.25	5.54	2.17	6.27	1.000
7.5–10.5	9	−0.42	2.41	1.23	7.21	1.87	16.27	1.000
10.5–13.5	19	0.75	1.11	0.58	4.78	1.20	16.26	1.000
13.5–16.5	24	−0.96	0.55	0.59	2.53	1.20	4.51	1.000
16.5–19.5	24	−0.82	0.49	0.74	2.26	1.09	4.28	1.000
19.5–22.5	23	−0.45	0.43	0.95	2.03	1.48	1.88	0.992
22.5–25.5	23	−0.79	0.46	0.80	2.07	1.42	2.16	0.999
25.5–28.5	16	−0.38	0.63	1.04	2.49	1.74	2.03	0.990
28.5–31.5	15	−0.74	0.86	0.87	3.25	1.22	7.23	1.000
31.5–34.5	10	−1.02	1.31	0.92	4.01	1.53	6.52	1.000
34.5–37.5	2	5.35	5.49	0.82	1.78	1.63	1.20	0.727

for the construction of the corresponding row of  $\mathbf{A}_h$ . With this approach, Eq. (2) accounts for both the space and time mismatch between the measurements, and for the horizontal smoothing performed by MIPAS. Residual coincidence errors were neglected, firstly, because errors in ECMWF temperature fields are believed to have a correlation length considerably larger than the actual spatial and temporal mismatch, such that these errors to first order approximation cancel out in the difference, and secondly, because errors resulting from the interpolation of ECMWF data in space and time are hard to estimate, and thus the exercise would have been quite speculative.

Since MIPAS measurement errors are strongly correlated in altitude, while error correlations in the time or horizontal domain are negligibly small (apart from the bias which is explicitly corrected for), all available pairs of co-located measurements ( $t_{\text{M},i}$ ;  $t_{\text{c},i}$ , where  $t_{\text{M},i}$  are MIPAS temperatures and  $t_{\text{c},i}$  are radiosonde corrected temperatures at pressure  $i$ ), were sorted into pressure bins such that each bin contains only data relating to one nominal MIPAS limb viewing geometry, acquired at different times and locations. With this approach each bin contains up to 24 data pairs in total. Validation of MIPAS temperatures in each pressure bin allows one to neglect error correlations in the vertical (pressure) domain (see Sect. 7 of von Clarmann 2006).

For each pressure bin  $i$  we calculated an estimate of the bias  $b_i$  as (see Eq. (C30)):

$$b_i = \frac{1}{n_i} \sum_{k=1}^{n_i} (t_{\text{M},i}(k) - t_{\text{c},i}(k)) \quad (3)$$

where  $n_i$  is the number of data pairs in the  $i$ -th bin, and  $k$  numbers the pairs within this bin. The statistical uncertainty of this bias is estimated by the standard deviation (see Eq. C31):

$$\sigma_{b_i} = \sqrt{\frac{\sum_{k=1}^{n_i} (t_{\text{M},i}(k) - t_{\text{c},i}(k) - b_i)^2}{n_i(n_i - 1)}} \quad (4)$$

For consistent measurements, in each bin the bias  $b_i$  should be statistically consistent with zero, considering its random error bar  $\sigma_{b_i}$ . If this is not the case, the bias should at least be consistent with zero considering its total error  $\sigma_{b_i,\text{tot}}$ , which can be estimated as:

$$\sigma_{b_i,\text{tot}} = \sqrt{\sigma_{b_i}^2 + \sigma_{b_i,\text{sys}}^2} \quad (5)$$

where  $\sigma_{b_i,\text{sys}}$  is the systematic error of  $b_i$  which can be estimated as:

$$\sigma_{b_i,\text{sys}} = \sqrt{\sigma_{t_{\text{M},i,\text{sys}}}^2 + \sigma_{t_{\text{c},i,\text{sys}}}^2} \quad (6)$$

in which  $\sigma_{t_{\text{M},i,\text{sys}}}$  and  $\sigma_{t_{\text{c},i,\text{sys}}}$  are the estimates of the systematic errors of the MIPAS and sonde corrected profiles respectively, as determined from error propagation analysis. We call these errors *expected* errors. A deviation of  $b_i$  from zero beyond the error bar  $\sigma_{b_i,\text{tot}}$  indicates that the systematic error  $\sigma_{t_{\text{M},i,\text{sys}}}$  of the experiment to be validated has been underestimated significantly.

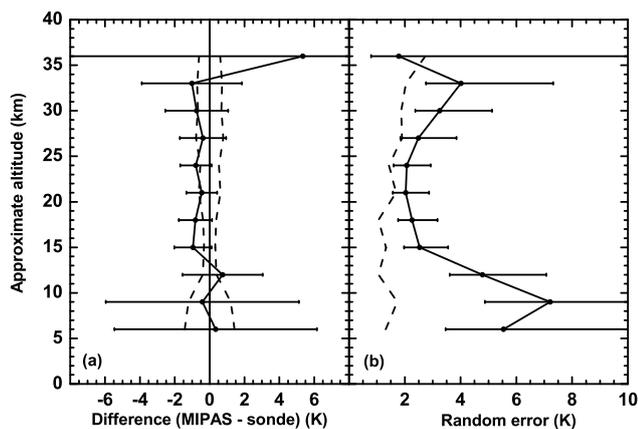
On the other hand, the random error  $\sigma_{d_i,\text{rnd}}$  of the differences  $d_i(k) = t_{\text{M},i}(k) - t_{\text{c},i}(k)$  can be estimated through the error propagation equation:

$$\sigma_{d_i,\text{rnd}} = \sqrt{\sigma_{t_{\text{M},i,\text{rnd}}}^2 + \sigma_{t_{\text{c},i,\text{rnd}}}^2} \quad (7)$$

in which  $\sigma_{t_{\text{M},i,\text{rnd}}}$  and  $\sigma_{t_{\text{c},i,\text{rnd}}}$  are the random errors of the MIPAS and sonde profiles respectively, and have been evaluated in the respective experiment characterization phases. This error estimate should be consistent with the precision estimate determined in the intercomparison experiment by (see Sect. 5 of von Clarmann 2006):

$$p_i = \sigma_{b_i} \sqrt{n_i}. \quad (8)$$

If we apply equations (3), (4), (6), (8) and (7) to our set of data we obtain the results reported in columns 3, 4, 5, 6 and 7 of Table 1 respectively. The first two columns of this table report the approximate altitude boundaries of the



**Fig. 1.** Summary of the intercomparison between MIPAS and radiosondes operated from Potenza and L'Aquila sites. Panel (a): bias  $b_i$  determined from the intercomparison (solid line) and expected systematic error  $\sigma_{b_i,sys}$  of the difference profile (dashed lines). Panel (b): precision  $p_i$  of the difference profile as estimated from the intercomparison (solid line), and expected random error  $\sigma_{d_i,rd}$  of the difference (dashed line). In both panels the error bars indicate the 95% confidence intervals determined on the basis of the  $t$ - (panel a) and on the chi-square (panel b) statistics, as explained in Sect. 3.1.

defined pressure bins and the number  $n_i$  of pairs per bin respectively. Here, the error  $\sigma_{t_{M,i},sys}^2$  (necessary for Eq. (6)) was obtained as the summation of the squared MIPAS error components (Dudhia, 2005) that are thought to be constant within our sample. Namely, these are errors affecting the forward model included in the MIPAS retrieval algorithm, and are due to: use of less than perfect spectroscopic line data; neglecting CO<sub>2</sub> line coupling; imperfect modeling of the instrument line-shape; neglecting non local thermodynamic equilibrium; and imperfect modeling of gaseous continua.

For the evaluation of  $\sigma_{d_i,rd}$  we calculated  $\sigma_{t_{M,i},rd}^2$  as the summation of the squared MIPAS error components that are thought to vary randomly within our sample. These error components affect both the MIPAS inversion model (such as errors due to: neglecting the horizontal variability of the atmosphere; the profile shape assumption outside the altitude range explored by the MIPAS scan; spectral interference of O<sub>3</sub>, CO<sub>2</sub>, ClONO<sub>2</sub>, CCl<sub>4</sub> and N<sub>2</sub>O<sub>5</sub>); and the measured spectra (measurement noise, frequency and intensity calibration errors).

For the calculation  $\sigma_{t_{c,i},rd}$  we considered both the radiosonde specified accuracy and the error on temperature arising from propagation of the pressure uncertainty through typical vertical temperature gradients.

The quantity  $\chi_{R,i}^2$  reported in column 8 of Table 1 is the reduced chi-square, testing the hypothesis that the differences  $d_i(k) = t_{M,i}(k) - t_{c,i}(k)$  are consistent with their expectation value  $b_i$  within their random error bars  $\sigma_{d_i,rd}$ .  $\chi_{R,i}^2$  is de-

defined as:

$$\chi_{R,i}^2 = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} \frac{(t_{M,i}(k) - t_{c,i}(k) - b_i)^2}{\sigma_{d_i,rd}^2} \quad (9)$$

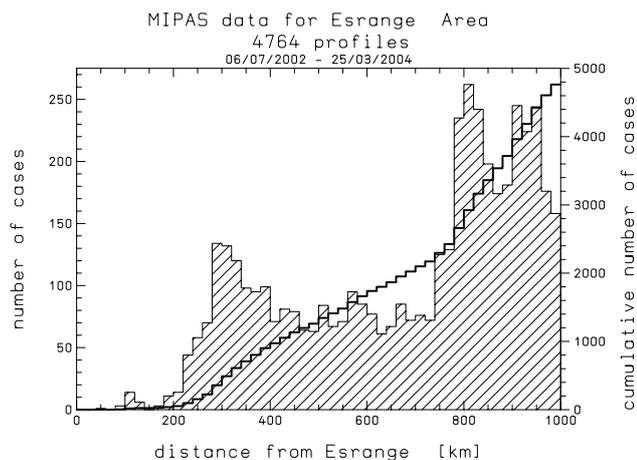
The expectation value of this quantity is unity; the probability  $L_i$  of getting a smaller value for this quantity in a new intercomparison, i.e. the probability of a substantial discrepancy between  $\sigma_{d_i,rd}$  and  $p_i$ , is reported in the rightmost column of Table 1.

For the sake of visual inspection, the results of Table 1 are also summarized in Fig. 1. In panel (a) of this figure we report, as a function of the approximate center altitude of each pressure bin, the bias  $b_i$  (solid line) and  $\pm\sigma_{b_i,sys}$ , the expected systematic error of the difference between the profiles (dashed lines). In panel (b) of Fig. 1 we report the precision  $p_i$  (solid line) as well as the expected random error  $\sigma_{d_i,rd}$  of the difference profile (dashed line). In both panels (a) and (b) the error bars of the estimated bias and standard deviation represent the 95% confidence interval calculated on the basis of the  $t$ - and chi-square statistics respectively (Gosset, 1908; Bevington and Robinson, 2003). Of course, since the chi-square probability distribution is not symmetric about its maximum, especially for a small number of degrees of freedom ( $=n_i - 1$ ), the confidence intervals of the standard deviation clearly reflect this property.

Below 10.5 km and above 34.5 km the size of the sample is too small for meaningful statistics, and therefore we do not discuss the results in those altitude ranges. In the altitude range above 10.5 km and below 34.5 km the detected bias is not statistically significant. However, there is strong indication that the estimated MIPAS precision is too optimistic by a factor of about 1.5 to 2. Discrepancies of this magnitude can hardly be explained by the neglect of higher order errors in the coincidence correction. Instead, they are attributed to occasional instabilities in the MIPAS retrievals causing spikes in the temperature profiles. These peculiar outliers cannot be assigned to individual limb scans for which the entire MIPAS profile is suspicious. At altitudes below 13.5 km the detected discrepancies are even larger. At these altitudes a candidate explanation is the deficiencies in the MIPAS cloud detection algorithm, which does not reliably reject all cloud-contaminated spectra from the Level 2 analysis.

### 3.2 Comparison with radiosonde measurements from the Esrange site

In this Section we compare MIPAS temperature with radiosonde measurements operated from the Arctic latitude site of Esrange, in northern Sweden. As an Arctic latitude site (67.9° N, 21.1° E) the atmosphere above Esrange undergoes extreme excursions with the seasons, as is manifested by clouds in the dry middle atmosphere, and the atmosphere is subject to heating and cooling during stratospheric warmings. This large geophysical variability provides an ideal



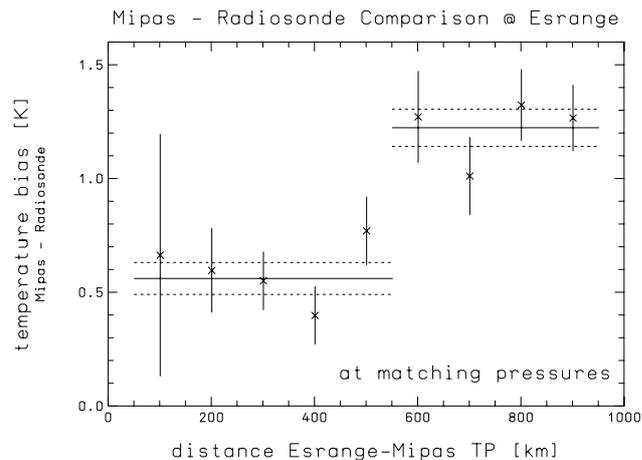
**Fig. 2.** Specific and cumulative distributions for the ground range from the MIPAS measurement locations to Esrange.

testing ground for the data derived from satellite borne instruments such as MIPAS.

The Esrange launches radiosondes to support campaigns conducted with rockets or balloons as the instrument carrier. As a launch site of opportunity the Esrange radiosonde data do not enter the ECMWF analyses. In the time period relating to the MIPAS data to be validated, (from 6 July 2002 to 25 March 2004) there were 146 flights with radiosondes. The campaigns took place in August, November, and December 2002, in January, February, March, June and July 2003, and in January, February and March 2004. Esrange uses the Väisälä Digicora sounding system with radiosondes of type RS80 characterized by the accuracy figures specified in Sect. 3. The Esrange radiosondes cover all pressures from the ground to 3 hPa. More than 75% of the sondes exceeded a top altitude of 10 hPa or about 30 km. Temperatures in the stratosphere varied in the range 185 to 270 K.

### 3.2.1 Testing the coincidence criteria

Application of the baseline coincidence criteria defined in Sect. 2.3 leads to very poor statistics, therefore we tune these criteria specifically for our intercomparison data set. For this purpose we define windows in time and horizontal range, which are believed to be acceptable for validation and which are to be met by the measurements. A vertical range window is not required for validation with radiosondes, since their altitude resolution far exceeds that of MIPAS, i.e. the MIPAS pressure can be matched exactly by the validation measurements. Using small windows reduces the effects of geophysical variability on the comparisons, however, it also reduces the number of available comparisons. The ultimate choice for time and ground range windows (i.e. the coincidence criterion) involves iterations resulting in a compromise among these competing requirements. To give an idea of the trade-off between the conservativeness of the adopted coincidence

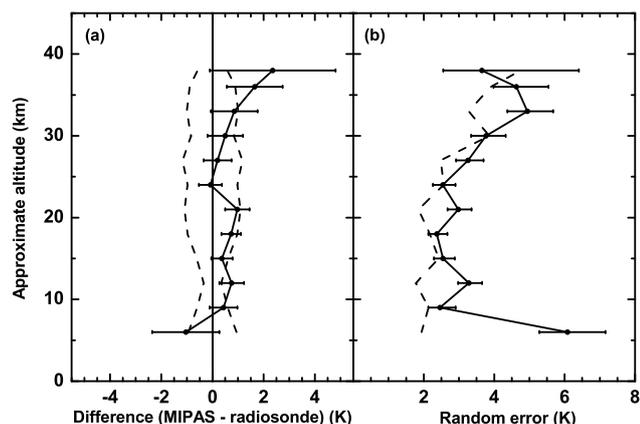


**Fig. 3.** Mean difference between MIPAS and radiosonde temperatures as a function of the ground distance from the MIPAS tangent point (TP) and Esrange in range windows of 200 km extent and centered every 100 km. The crosses denote the mean value of the temperature difference in each range window, vertical bars are the 1-sigma errors of these means. Solid horizontal bars mark the weighted means for range windows below and above 500 km; dashed lines are the 1-sigma errors of these weighted means.

criterion and the number of matching pairs of measurements, in Fig. 2 we report the specific and cumulative distributions of the available MIPAS measurements as a function of the ground spatial distance from Esrange.

As a threshold for the time mismatch between the measurements to be compared we arbitrarily choose a time-interval of 4 h between the radiosonde launch time and the MIPAS measurement time. MIPAS data located within 1000 km distance from Esrange separate time-wise into two groups: the morning window with southbound passes lasts from 07:20 UT to 10:30 UT and the evening window with northbound passes from 19:05 UT to 22:20 UT. Since many of the Esrange radiosondes were launched at 05:00 UT or 06:00 UT the related measurements are thus associated with the MIPAS southbound morning passes.

We divided the MIPAS datasets into range windows of 200 km extent centered every 100 km (which leads to partial overlap of adjacent range windows). For each MIPAS profile matching a radiosonde profile in the selected time-range window we searched for the pressure level measured by MIPAS in the radiosonde profile and extracted the difference between the MIPAS and radiosonde temperature. Figure 3 shows the global average of these differences as a function of the center of the chosen range-window. The mean temperature differences increase almost stepwise at a ground range of about 500 km. The specific cause for this behavior is not known, and we attribute it to a combination of geophysical variability and MIPAS sampling strategy (MIPAS measurements are mostly clustered around fixed latitude circles). The presence of this “step” allows us to choose as the



**Fig. 4.** Plot summarizing the agreement between temperatures from MIPAS and radiosonde measurements operated by the University of Bonn from Esrange. The format of the plot is the same as for Fig. 1.

spatial coincidence criterion the distance of 500 km: such a criterion permits us simultaneously to obtain good statistics (see Fig. 2) without significantly impacting on the overall size of the observed discrepancies between MIPAS and the validation measurements acquired at Esrange.

### 3.2.2 Results of comparison to MIPAS Temperature

For each MIPAS profile matching a radiosonde profile within the spatial and temporal margins discussed in Sect. 3.2.1, we searched for the pressure levels measured by MIPAS in the corresponding radiosonde profile and extracted the difference between the MIPAS and radiosonde temperatures. The obtained differences were then grouped in pressure bins centered around the pressures of the nominal MIPAS tangent points. Since the available MIPAS AKs, calculated for standard atmospheric conditions, can not be considered representative of the Arctic atmosphere of the Esrange area (Ceccherini, private communication, 2006<sup>1</sup>), the radiosonde profiles were not adapted to the MIPAS vertical resolution prior to the intercomparison. For this reason, the *smoothing* (see Sect. 2) performed by MIPAS acts as a random error (smoothing error), which was evaluated and included in the budget of the expected random error of the observed differences. The remaining expected error components considered in this intercomparison coincide with the ones described in Sect. 3.1.

The results of this intercomparison are summarized in Fig. 4, with the same format as Fig. 1. Panel (a) of Fig. 4 shows that the bias is only statistically significant (i.e. beyond the 95% confidence interval) for some bins between 12

<sup>1</sup>Ceccherini, S.: Test on the range of validity of the first order expansion of the retrieved profile as a function of the true profile and some considerations on the MIPAS averaging kernels, private communication, 2006.

and 21 km and at 36 km. However the bias is always consistent with the expected systematic error of the differences.

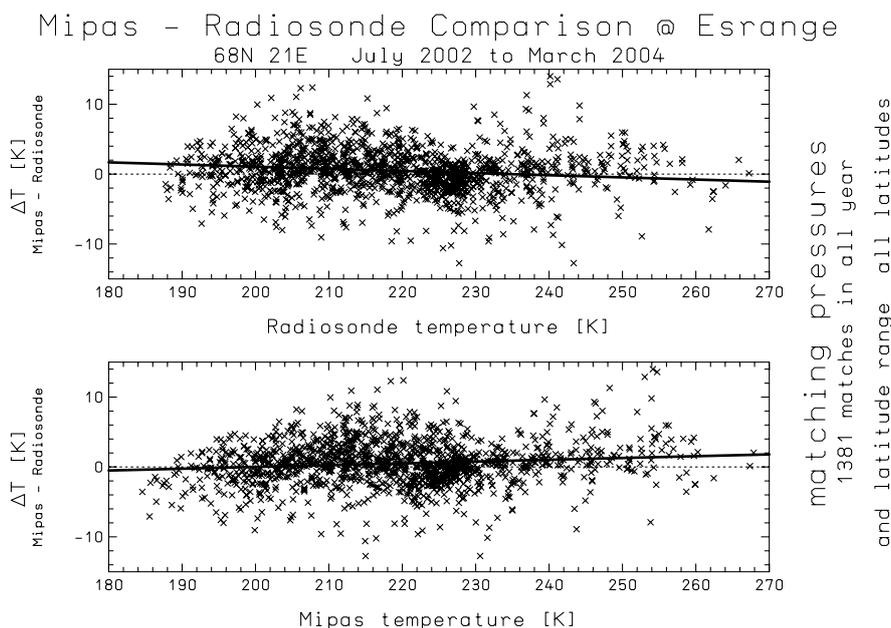
Panel (b) of Fig. 4 shows that the standard deviation of the differences between the MIPAS and correlative temperatures generally tends to be larger than the expected random error of the differences. This occurs especially for the bin around 6 km. Possibly at this altitude the observed effect could be caused by insufficient cloud filtering in the MIPAS data processing. Above 6 km the horizontal smoothing error (see Sect. 3.3), which was not evaluated in this analysis, may have played a role in the determination of standard deviations slightly exceeding the expected random error of the differences.

In order to check whether the MIPAS temperature error depends on the value of the temperature itself, in Fig. 5 we plotted the behavior of the temperature differences, MIPAS-radiosonde, as a function of both radiosonde (top panel) and MIPAS (bottom panel) temperatures. Figure 5 demonstrates that the temperature differences are independent of the temperature measured. The thick lines represent a linear fit to the data. The slopes deviate only marginally from zero, which confirms the lack of a systematic variation of the temperature differences with temperature. Partitioning of these data by season (winter: November to March, and summer: April to October) also does not suggest a systematic variation with season, even though there are only 142 data matches during summer out of the total of 1374 matches. Partitioning the data into latitudes north and south of the Esrange does not reveal any systematic variation either.

### 3.3 Comparison with NDACC/GAW radiosonde network data

In this Section, MIPAS temperature profiles are compared with radiosonde measurements acquired as part of WMO's Global Atmosphere Watch (GAW) and two of its main contributors, the Network for the Detection of Atmospheric Composition Change (NDACC, Kurylo and Zander 2001) and the Southern Hemisphere ADDitional OZonesondes program (SHADOZ, Thompson et al. 2003). Contributing instruments and stations are listed in Table 2. Note that the radiosonde data at Esrange considered for this validation are those associated with ozonesonde flights operated by NIES (National Institute for Environmental Studies, Tsukuba, Japan) and are independent from those of the University of Bonn discussed in Sect. 3.2. Due to the large amount of available data, the comparisons are limited here to the year 2003.

Before being used for validation, the quality of the collected correlative measurements was checked by comparison against climatological data, taking into account data remarks inferred from NDACC time series and ECMWF fields. Outliers and physically unacceptable data were excluded from the validation.



**Fig. 5.** Temperature differences MIPAS – radiosonde as a function of radiosonde temperature (top panel) and MIPAS temperature (bottom panel). Thick lines are linear fits to the data.

### 3.3.1 Discussion of coincidence criteria

For validation with this dataset, coincidence criteria were set to a maximum distance of 500 km between the MIPAS mean tangent point and the ground station, and a time mismatch of 6 h (12 h for the study of time series) between the measurements. As discussed in Sect. 3.2, using more conservative coincidence criteria leads to a drastic reduction of the number of usable coincidences, hence decreasing the significance of the statistics. Furthermore, for each intercompared set of profiles we also evaluate the error due to less than perfect spatial coincidence and, as shown later, usually this error component turns out to provide only a marginal contribution to the total error budget of the calculated profile differences.

### 3.3.2 Comparison error budget

The difference between the MIPAS and correlative measurement temperatures is affected by several error components. In general the total error of this difference can be expressed in terms of the following error covariance matrix  $\mathbf{S}$ :

$$\mathbf{S} = \mathbf{S}_M + \mathbf{S}_N + (\mathbf{A}_{Mv} - \mathbf{A}_{Nv}) \mathbf{S}_v (\mathbf{A}_{Mv} - \mathbf{A}_{Nv})^T + (\mathbf{A}_{Mh} - \mathbf{A}_{Nh}) \mathbf{S}_h (\mathbf{A}_{Mh} - \mathbf{A}_{Nh})^T + \mathbf{S}_{\Delta r}. \quad (10)$$

In this equation the subscripts “M” and “N” indicate quantities referring to MIPAS and to the validation measurement respectively.  $\mathbf{S}_M$  and  $\mathbf{S}_N$  are the total error covariances of the two compared measurements,  $\mathbf{A}_{Mv}$  and  $\mathbf{A}_{Nv}$  are the vertical AKs, while  $\mathbf{A}_{Mh}$  and  $\mathbf{A}_{Nh}$  are the horizontal AKs.  $\mathbf{S}_v$  and  $\mathbf{S}_h$  are covariances describing the vertical and the horizontal variability of the atmosphere respectively. The error terms

involving the AKs account for the smoothing error due to the different spatial (vertical and horizontal) response functions of the two measurements (Rodgers and Connor, 2003).  $\mathbf{S}_{\Delta r}$  is the covariance relating to the error due to the spatial and temporal mismatch between the measurements considered.

For the MIPAS and radiosonde total errors we assume the errors described in Dudhia (2005) and in Sect. 3 respectively. Ideally, these are the error bars that should be used to test the differences, MIPAS – ground measurement, against zero if the compared measuring systems and sounded air masses were identical. In practice, however, profile differences are larger due to smoothing and collocation errors. In this study we have estimated separately the errors due the difference in horizontal and vertical resolution and the errors due to differences in geolocation.

**Vertical smoothing error.** Starting from Rodgers’ theory and formalism (Rodgers, 2000), we estimate the effect of the difference in vertical resolution by means of the vertical AKs. First, the vertical AKs of MIPAS are used to map the high-resolution ground-based measurement to the MIPAS vertical resolution according to Eq. (1). Secondly, the smoothing error is estimated as the difference between the smoothed profile resulting from the previous step and the original high-resolution profile.

**Horizontal smoothing error.** The ECMWF temperature field can be used to estimate the derivative of temperature in the direction of the MIPAS line-of-sight. Multiplication of this derivative by a suitable fraction of the length of the optical path of the MIPAS observation estimates the horizontal smoothing.

**Table 2.** List of radiosonde stations contributing to the NDACC/GAW network.

Station	Location	Lat. (deg)	Long. (deg.)	Institute
Alert	Canada	82.5	−62.33	MSC
Eureka	Canada	80.05	−86.42	MSC
Ny-Alesund	Svalbard	78.91	11.88	AWI
Thule	Greenland	76.51	−68.76	DMI
Resolute	Canada	74.72	−94.98	MSC
Scoresbysund	Greenland	70.48	−21.97	DMI
Esrang	Sweden	67.88	21.06	NIES
Sodankylä	Finland	67.37	26.67	FMI
Keflavik	Iceland	63.97	−22.6	INTA
Orlandet	Norway	63.42	9.24	NILU
Jokioinen	Finland	60.82	23.48	FMI
Churchill	Canada	58.75	−94.07	MSC
Edmonton	Canada	53.55	−114.1	MSC
Goose Bay	Canada	53.32	−60.38	MSC
Legionowo	Poland	52.4	20.97	INWM
Debilt	Netherlands	52.1	5.18	KNMI
Valentia	Ireland	51.93	−10.25	ME
Uccle	Belgium	50.8	4.35	KMI
Praha	Czech Republic	50.02	14.45	CHMI
Hohenpeissenberg	Germany	47.8	11.02	DWD
Payerne	Swiss Alps	46.49	6.57	MCH
Tsukuba	Japan	36.05	140.13	JMA
Paramaribo	Surinam	5.81	−55.21	KNMI
San Cristobal	Galapagos	−0.92	−89.6	CMDL
Nairobi	Kenya	−1.27	36.8	MCH
Malindi	Kenya	−2.99	40.19	RPSM
Natal	Brazil	−5.42	−35.38	INPE
Watukosek	Java	−7.5	112.6	JAXA
Ascension Island	Congo	−7.98	−14.42	NASA
Tutuila	Samoa	−14.23	−170.56	CMDL
Fiji	Fiji	−18.13	178.42	CMDL
Saint-Denis	La Reunion	−21.06	55.47	CNRS
Irene	South Africa	−25.25	28.18	SAWS
Lauder	New Zealand	−45.03	169.68	NIWA
Marambio	Antarctica	−64.28	−56.72	INTA
Dumontd'Urville	Antarctica	−66.67	140.01	CNRS
Syowa	Antarctica	−69	39.58	JMA
Neumayer	Antarctica	−70.65	−8.25	AWI
Belgrano	Antarctica	−77.87	−34.63	INTA

Algebraically, we estimate the horizontal smoothing error  $\delta_{h,i}$  of a given MIPAS profile level  $i$  as:

$$\delta_{h,i} = |\nabla_{r,i} T_{\text{ECMWF}}| r_{90,i} \quad (11)$$

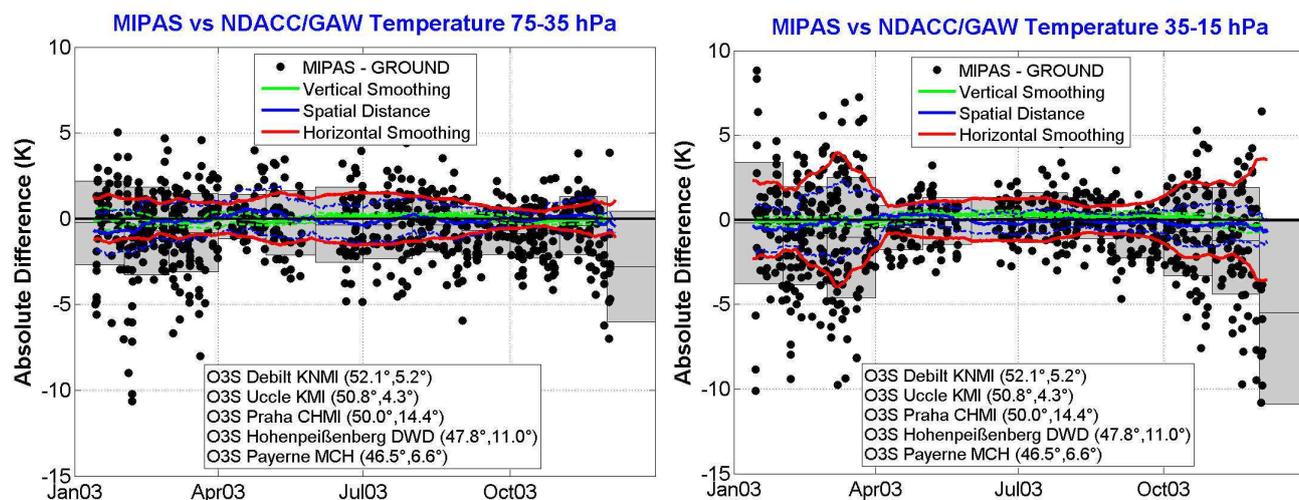
where  $|\nabla_{r,i} T_{\text{ECMWF}}|$  is the modulus of the component of the temperature gradient along the line-of-sight of the MIPAS limb-observation with its tangent point located at the retrieval level  $i$ ; this directional derivative is evaluated at the tangent point, assuming the ECMWF temperature fields.  $r_{90,i}$  is the length of the MIPAS optical path, at altitude  $i$ , corresponding to 90% of the measured CO<sub>2</sub> radiance from which temper-

ature is retrieved. This length was estimated using a simple radiative transfer model capable of calculating MIPAS limb radiance emission spectra in a two-dimensional atmosphere (De Clerq and Lambert, 2006).

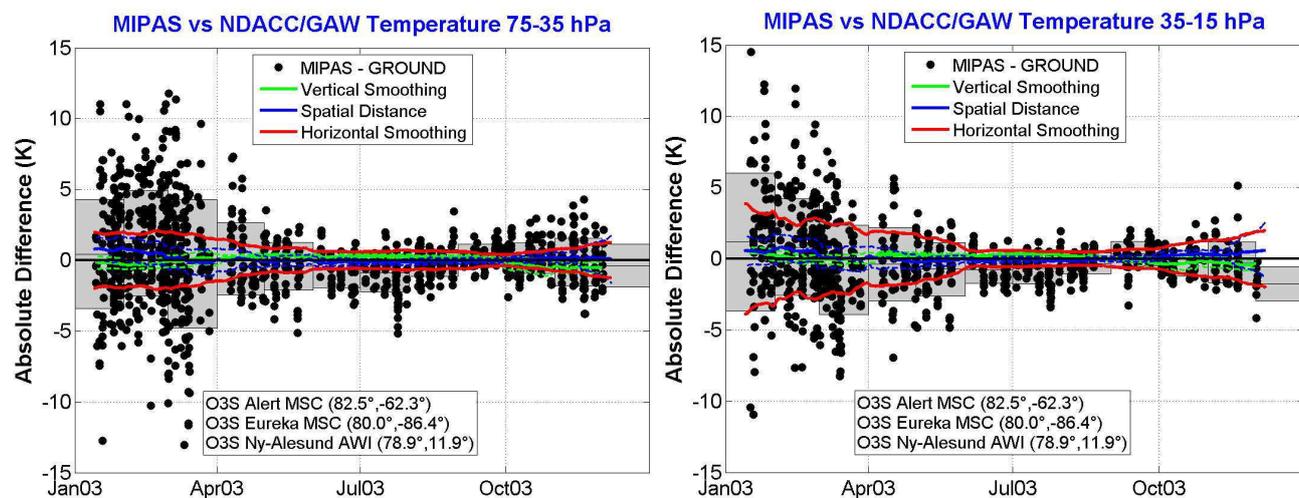
**Coincidence error.** The spatial distance between a generic pair of MIPAS and validation measurements induces a difference  $\delta_t$  that can be estimated as:

$$\delta_t = T_{\text{ECMWF}}(\mathbf{r}_{\text{MIPAS}}) - T_{\text{ECMWF}}(\mathbf{r}_{\text{VAL}}) \quad (12)$$

where  $T_{\text{ECMWF}}(\mathbf{r}_{\text{MIPAS}})$  and  $T_{\text{ECMWF}}(\mathbf{r}_{\text{VAL}})$  denote the ECMWF temperatures interpolated to the time and geoloca-



**Fig. 6.** European mid-latitude ( $46.5\text{--}52.1^\circ\text{ N}$ ) time series of the differences (full symbols) between MIPAS and NDACC/GAW radiosondes operated from the stations indicated in the plot key. Temperatures were averaged in the pressure layers from 75 to 35 hPa (left plot) and from 35 to 15 hPa (right plot). The grey areas indicate monthly averages (center) and  $\pm$ standard deviations (edges). The solid red (green) curves indicate the  $\pm$ average horizontal (vertical) smoothing error. The solid blue line indicates the error due to the spatial mismatch between the measurements. The dashed lines indicate the  $\pm$ standard deviation of the error estimates with the same color code of the error itself.



**Fig. 7.** Same as Fig. 6 but for Northern Polar latitudes ( $78.9\text{--}82.5^\circ\text{ N}$ ).

tion of the MIPAS and validation measurements respectively. The spatial coincidence error is considered with its sign. In fact, MIPAS is sampling the temperature field on fixed latitudinal circles, and a station occupies by definition a fixed geolocation. As the mean temperature field exhibits a meridional gradient, any permanent spatial distance (e.g. between MIPAS tangent points and a station) in the North-South direction will lead to a systematic difference in temperature. Hence it is convenient to use the spatial coincidence error with its own sign.

### 3.3.3 Intercomparison analyses

Given the large amount of correlative measurements available, in this case it was possible to carry out both time-series analyses and vertically resolved analyses of the discrepancies between the MIPAS and correlative measurements.

**Time series.** In order to monitor the quality of the MIPAS temperature as a function of time we have chosen to compare average temperature values in pre-defined layers rather than local profile values. In particular, average temperatures are calculated for two layers: from 75 to 35 hPa

( $\approx 18$ – $23$  km) and from 35 to 15 hPa ( $\approx 23$ – $28$  km), using alternatively MIPAS and correlative measurements. The approach of comparing layer mean temperatures offers several advantages over the usual comparison of local profile values; in particular both the random and the smoothing error components of the local profile differences are significantly reduced by the averaging process. The reduced size of these errors makes it possible to breakdown the intercomparison dataset on a monthly basis and to study the time behavior of the discrepancies for several geographical areas and synoptic systems.

Figures 6 and 7 are respectively European mid-latitude ( $46.5$ – $52.1^\circ$  N) and Arctic ( $78.9$ – $82.5^\circ$  N) time series of the differences (full symbols) between layer mean temperatures from MIPAS and radiosondes operated from the stations indicated in the plots' keys. Temperatures were averaged in the pressure layers from 75 to 35 hPa (left plots) and from 35 to 15 hPa (right plots). The grey areas denote monthly averages (center) and  $\pm$ standard deviations (edges). The solid red (green) curves indicate the  $\pm$ average horizontal (vertical) smoothing error. The solid blue lines indicate the error due to the spatial mismatch between the measurements. The dashed lines indicate the  $\pm$ standard deviation of the error estimates with the same color code used for the error itself. From these figures it is evident that the horizontal smoothing error dominates over the other error sources. Furthermore, the behavior of the standard deviation of the discrepancies follows very closely the estimate of the horizontal smoothing error, hence confirming the accuracy of the estimate of this error component derived with the procedure explained in Sect. 3.3.2. We also note that both the horizontal smoothing of MIPAS and the coincidence error are generally largest in wintertime (at the edges of the plots). This is because in this period the horizontal variability of the stratosphere is also largest, due to strong planetary waves that can propagate easily through the poleward meridian component of the wintertime circulation (the so-called Dobson-Brewer circulation). In summer time, the weaker and essentially zonal circulation prevents planetary waves from propagating deeply into the stratosphere; therefore the atmosphere is dynamically more quiet (Salby, 1984).

Figures 6 and 7 show also that the quality of MIPAS data is almost constant versus time, the bias is not statistically significant and the standard deviation is in agreement with the predicted random error of the differences. The agreement found in these figures is a typical case, however there are a few cases in which discrepancies larger than the predicted errors are found. These large discrepancies, however, are restricted to very limited data sets, with very poor statistics, which hence do not permit us to draw quantitative conclusions.

**Vertically resolved analyses.** In this type of analyses, for each matching pair of temperature profiles, we first adapted the high-resolution radiosonde profile to the MIPAS perception using Eq. (1), with a regridding technique inspired by

the work of Calisesi et al. (2005); secondly, we calculated the differences between profiles at corresponding pressure levels. Profile differences resulting from the whole set of intercompared pairs were finally grouped into pressure bins centered about the mean pressure of the MIPAS retrieval levels in the considered ensemble of measurements. The average and the standard deviation of the profile differences in the individual pressure bins were finally studied in relation to their systematic and random error components. Note that in this case, since the MIPAS vertical smoothing was applied to the correlative profiles using Eq. (1) (using the standard MIPAS AKs), the comparison results should be hardly affected by the vertical smoothing error.

Altitude-resolved analyses were carried out for the major geographical areas and synoptic systems identified in the time-series analyses.

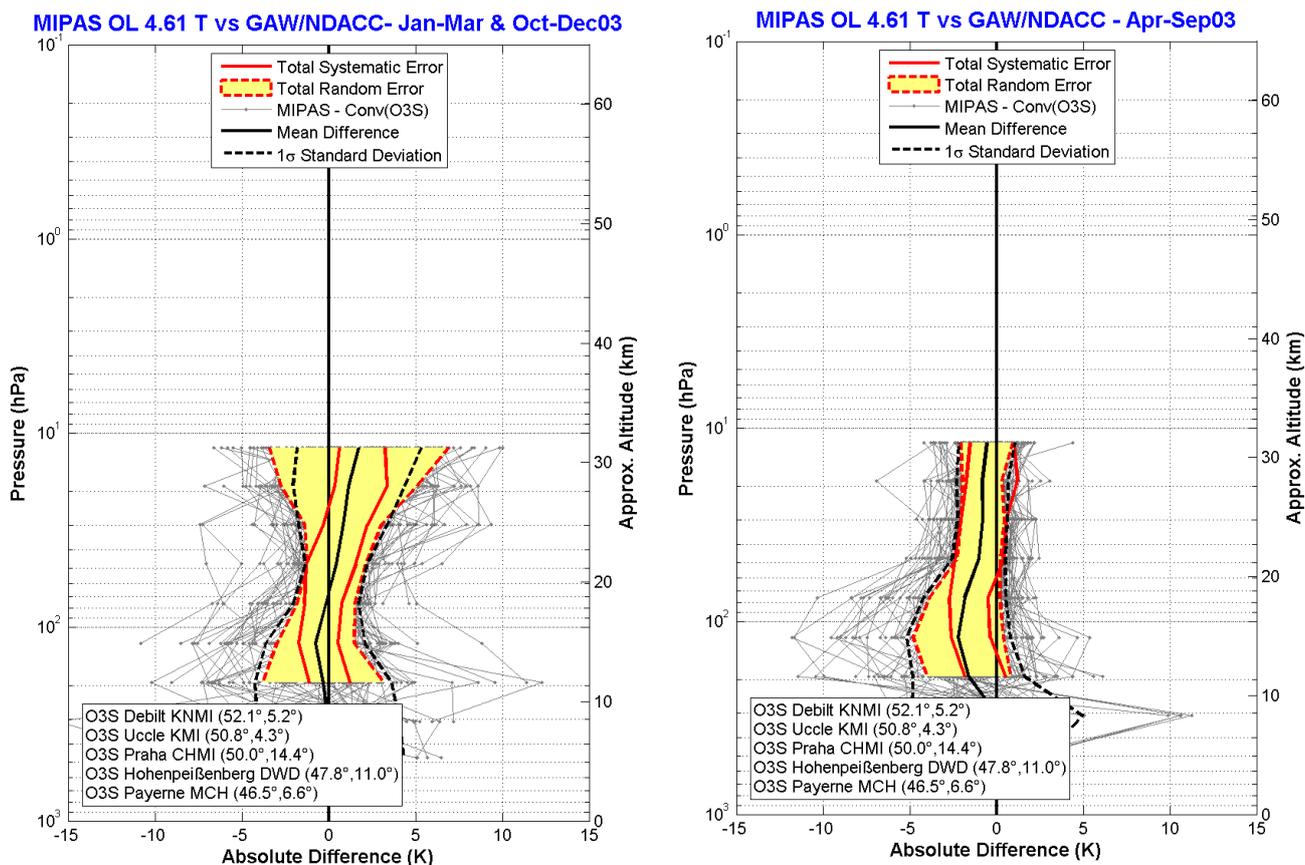
In Figs. 8 and 9 we report, for Mid- and Arctic- latitudes respectively, an example of the altitude-resolved analysis of the differences between MIPAS and NDACC/GAW radiosondes (gray lines with symbols) for the same stations considered in Figs. 6 and 7. The black lines indicate the average (solid) and the standard deviation (dash) of the differences. The solid red lines indicate the expected total systematic error of the differences, while the yellow-filled areas delimited by the dashed red lines indicate the expected random error of the differences. The left plots refer to measurements from January to March and from October to December 2003, while the right plots refer to the months from April to September 2003.

From Figs. 8 and 9 we see that the detected bias is always consistent with the expected systematic error. The amplitude and the altitude behavior of the bias depends both on the selected time period and on the latitude band. Moreover, the standard deviation is consistent with the expected random error. This latter error is dominated by the contribution due to the horizontal smoothing.

The altitude-resolved analyses presented here are typical results; in general we see that within the altitude range considered in this intercomparison (12–36 km) there are no specific intervals in which the mean or the standard deviation of the profile differences particularly exceed the corresponding expected errors.

#### 4 Comparison with SPIRALE measurements

SPIRALE (French acronym for infrared absorption spectroscopy by diode lasers) is a balloon-borne spectrometer with six tunable diode lasers dedicated to in-situ measurements of trace compounds in the upper troposphere and the stratosphere up to an altitude of 35 km. Its principle, operation and estimation of the measurement uncertainties have been detailed in a previous paper (Moreau et al., 2005). Briefly, absorption of mid-infrared laser beams takes place in an air-open Herriott cell, between two mirrors separated

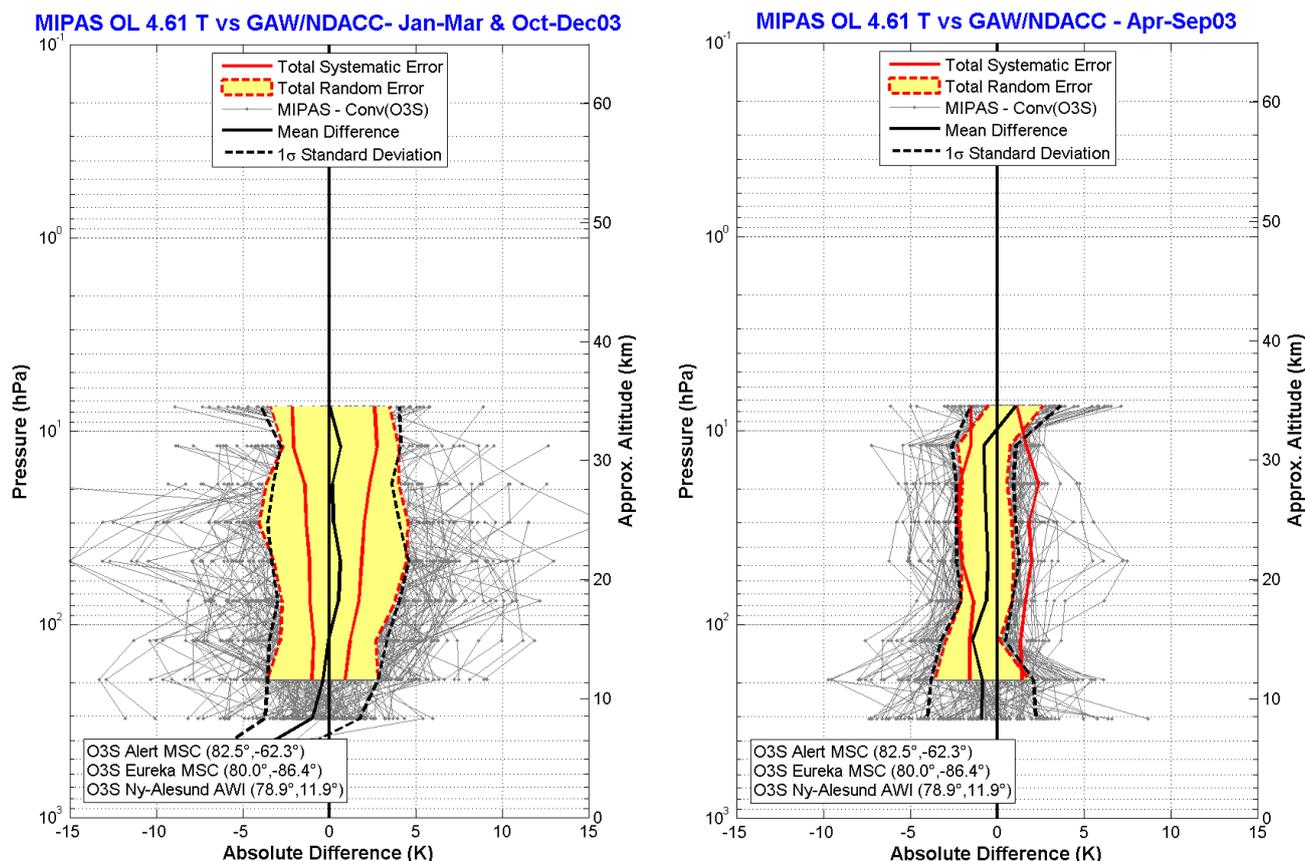


**Fig. 8.** Altitude resolved analysis of the differences (gray lines with symbols) between MIPAS and NDACC/GAW radiosondes for the same European mid-latitude stations considered for Fig. 6. The black lines indicate the average (solid) and the standard deviation (dashed) of the differences. The solid red line indicates the expected total systematic error of the differences, while the yellow-filled area delimited by the dashed red line indicates the expected random error of the differences. The left plot refers to measurements from January to March and from October to December 2003; the number of collocations contributing to this analysis is 177. The right plot refers to the months from April to September 2003; the number of contributing collocations is 160.

by 3.5 m, thus enabling a very long optical path (up to 544 m). Vertical profiles of concentrations of a great number of species, such as  $O_3$ ,  $CH_4$ ,  $CO$ ,  $N_2O$ ,  $HNO_3$ ,  $NO_2$ ,  $NO$ ,  $HCl$ ,  $HOCl$ ,  $H_2O_2$ ,  $COF_2$ , are measured with very high vertical resolution (a few meters), high sensitivity (down to volume mixing ratios of 20 pptv) and high accuracy (5 to 20%). Since altitude-resolved volume mixing ratio profiles are retrieved from SPIRALE measurements assuming known atmospheric temperature and pressure, very accurate in situ temperature measurements are required. For this purpose, two temperature probes made of resistive platinum wire are deployed during the flight, at the extremities of two horizontal masts of 2.5 m length. The two probes are located at the opposite sides of the main axis of the sampling cell, and therefore at least one probe is undisturbed by the thermal activity of the gondola. The bias of the air temperature is estimated to be smaller than 1 K, resulting in quite a poor accuracy compared to the intrinsic precision of the probe itself ( $\approx 0.05$  K). This is due to the difficulty of accounting for

the thermal influence of the wire holder and of radiative effects. Pressure is also measured aboard the gondola by two calibrated and temperature-regulated capacitance manometers of 0–1034 hPa and 0–100 hPa full scale ranges. This results in an accuracy of 0.5 hPa in the lower part of the profiles (200 hPa), improving to 0.1 hPa in the upper part (5 hPa). Assuming standard atmosphere temperature vertical gradients, this pressure uncertainty translates into an almost constant and negligible error (0.1 K) on the whole temperature profile, with respect to the accuracy of the temperature sensor itself.

Two flights were successfully completed in the frame of the ENVISAT validation campaign. The flights were carried out from different latitudes, namely from Kiruna, (Sweden, 68° N, 20° E) on 21 January 2003 and from Aire-sur-l'Adour, (France, 43.6° N, 0°) on 2 October 2002. During these validation flights, pressure and temperature measurements were acquired every 1.1 s, hence providing profiles with a vertical resolution of about 5 m. Among the two vertical temperature



**Fig. 9.** Same as Fig. 8 but for Arctic latitude stations. The number of collocations contributing to the analysis is 388 for the left plot and 243 for the right plot.

profiles measured in these campaigns, only the profile measured at Kiruna falls in the proximity of MIPAS measurements. Even for this profile, however, the obtained time and space coincidence does not meet the recommended criteria given in Sect. 2.3. Namely, we find that the two nearest MIPAS measurements belong to scan 20 of orbit 4677 (460 km and 8 min apart from SPIRALE) and to scan 6 of orbit 4678 (650 km and 1 h 50 min apart from SPIRALE).

Therefore we compared the SPIRALE profile with the aforementioned MIPAS profiles. In both cases, prior to the intercomparison the SPIRALE profile was corrected for the spatial and temporal mismatch between the compared measurements. The correction was once again done according to Eq. (C15) of von Clarmann (2006), and using temperature profiles obtained by interpolating in space and time ECMWF fields with a spatial resolution of  $0.5^\circ \times 0.5^\circ$  and a temporal resolution of 3 h.

Since the intercomparison takes place in polar atmospheres, once again in this case the available MIPAS AKs were not considered representative of the actual measured atmosphere and no correction was applied to the SPIRALE profile to match the MIPAS vertical resolution.

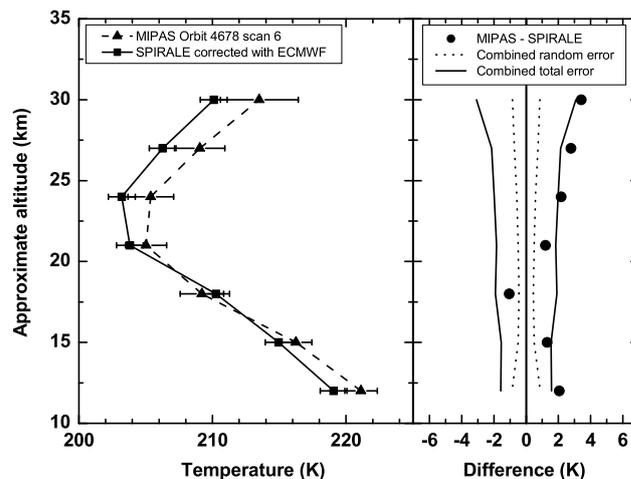
The intercomparison of SPIRALE with either of the two selected MIPAS measurements leads to very similar quantitative agreement. Figure 10 shows the results of the intercomparison between the ECMWF-corrected SPIRALE temperature and the MIPAS profile relating to scan 6 of orbit 4678. In particular, the left panel of Fig. 10 shows the two profiles (SPIRALE, full squares connected by solid line and MIPAS, full triangles connected by dashed line) with total error bars. Although the original SPIRALE profile has very high vertical resolution and a fine sampling step, in Fig. 10 we only plot the (corrected) SPIRALE profile points with pressures matching the ones of the MIPAS profile. The right panel of Fig. 10 shows the difference, MIPAS minus corrected SPIRALE (full symbols), as well as the expected values of both the random (dotted line) and the total (solid line) errors of the difference. From Fig. 10 it is clear that the compared profiles agree within the total error bars. Since only a single profile is available for this intercomparison nothing can be inferred regarding the individual (systematic and random) MIPAS error components.

## 5 Comparison with MIPAS-B measurements

MIPAS-B (Michelson Interferometer for Passive Atmospheric Sounding – Balloonborne) is an advanced cryogenic Fourier Transform Infra-Red (FTIR) spectrometer specially tailored to operation on a stratospheric balloon gondola (Fischer and Oelhaf, 1996). Equipped with suitable subsystems, MIPAS-B allows precise limb-emission sounding of chemical constituents related to the stratospheric ozone problem and to the greenhouse effect. This method is appropriate for obtaining vertical profiles of temperature, ozone and a considerable number of key radicals (e.g. NO, NO<sub>2</sub>) and reservoir species (HNO<sub>3</sub>, N<sub>2</sub>O<sub>5</sub>, ClONO<sub>2</sub>, and HO<sub>2</sub>NO<sub>2</sub>), as well as source gases (e.g. CH<sub>4</sub>, N<sub>2</sub>O, H<sub>2</sub>O, CFC-11, CFC-12, CFC-22, CCl<sub>4</sub>, CF<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, and SF<sub>6</sub>) simultaneously, with an altitude resolution of 2 to 3 km. Hence, MIPAS-B is able to measure the budget and partitioning of the complete NO<sub>y</sub> family (i.e. NO+NO<sub>2</sub>+HNO<sub>3</sub>+ClONO<sub>2</sub>+2 \* N<sub>2</sub>O<sub>5</sub>+HO<sub>2</sub>NO<sub>2</sub>), together with its source gas N<sub>2</sub>O, and to assess denitrification, dehydration, budgets and tracer correlations. As an emission sounder, MIPAS-B is independent of any extraterrestrial light source, which allows a high flexibility in terms of the time of the day of the launch and the observation geometry. Both azimuth and elevation angles can be commanded from the ground. This flexibility allows long integration times and permits the adjustment of the line of sight (LOS) to geophysical constraints in an optimal way (e.g. in the case of strong gradients in the constituents' fields, during sunrise/sunset, or for validation purposes). An innovative pointing system ensures high precision and stability of the acquired observation angles. The capability of MIPAS-B to measure day and night and to choose the observation angles independently of the Sun and Moon is very useful for the validation of satellite instruments.

MIPAS-B data processing, from the raw interferograms and the instrument housekeeping data to the calibrated spectra, is described in Friedl-Vallon et al. (2004) and references cited therein. Data processing includes instrument characterization, such as detector non-linearity, and a complete error budget of the calibrated spectra. Retrievals of temperature and trace gases profiles are performed with the KOPRAFIT code, an algorithm extension of KOPRA (Karlsruhe Optimized and Precise Radiative transfer Algorithm, Höpfner et al. 2002).

The validation measurements were performed with a vertical grid of 1.5 km while the retrieval grid was set to 1 km. Regularization was based on the Tikhonov-Phillips approach. The resulting vertical resolution typically lies between 1.5 and 3 km. Spectra were fitted in the six MIPAS-B proven microwindows included in the two spectral intervals 801–813 cm<sup>-1</sup> and 941–957 cm<sup>-1</sup> (Wetzel et al., 2002). The error estimation for the temperature retrieval includes random noise, calibration errors, errors in the CO<sub>2</sub> mixing ratio, line of sight inaccuracies, and spectroscopic data er-

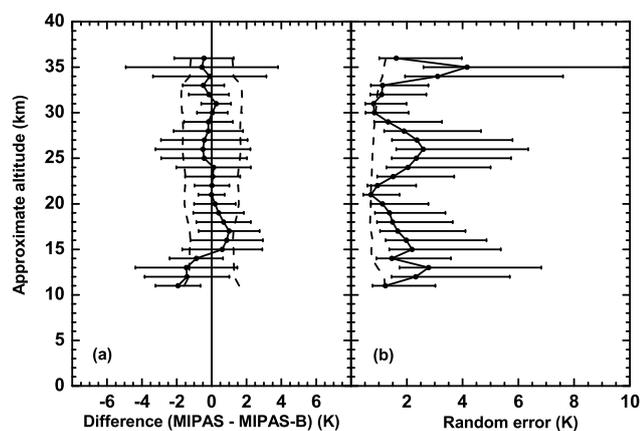


**Fig. 10.** Comparison of MIPAS with SPIRALE. Left panel: measured temperature profiles from MIPAS (triangles connected by dashed line) and SPIRALE (squares connected by solid line). The SPIRALE profile was corrected for the space and time mismatch using ECMWF fields. The right panel shows the profile differences (full circles) along with random (dotted line) and total (solid line) combined errors.

rors. Since the microwindows used for MIPAS-B retrievals are completely separate from the microwindows used by the ESA retrievals from MIPAS data (Raspollini et al., 2006), all these error components contribute to the error budget of the profile differences analyzed in this intercomparison. A detailed description of the Level 2 MIPAS-B data analysis is given in Wetzel et al. (2006) and references therein.

The first MIPAS-B flight within the ENVISAT validation activities took place from Aire sur l'Adour during the night of 24/25 September 2002. Launch was at 18:43 UT. The float was reached at a ceiling altitude of about 39 km at 21:47 UT i.e. 20 min before the ENVISAT overpass. The cut was at 23:50 UT. During the nominal measurement program (during ascent within the stratosphere and at ceiling) all systems worked nominally. The right launch time and the adjustable viewing direction of MIPAS-B have allowed a set of correlative data in excellent spatial and temporal coincidence with the satellite measurements to be obtained.

A second validation flight was performed from Erange, Kiruna (Sweden, 68° N, 21° E) on 20/21 March 2003. After a flight duration of more than 15 h from 18:22 UT (20 March) to 9:38 UT (21 March) touch down of the gondola was only about 50 km away from the launch site. The long duration of this flight was possible since the balloon was released right into the center of the polar vortex, where wind speeds are very low. Several limb sequences were measured matching the evening and morning overpass of ENVISAT (orbits 5508 and 5515, respectively). Most tangent points of MIPAS on ENVISAT have been matched within less than 100 km and 15 min respectively. All systems of the gondola worked stably and reliably.



**Fig. 11.** Summary of comparison between MIPAS/ENVISAT and MIPAS-B measurements. The reader should refer to Fig. 1 for the explanation of the various curves plotted.

The third ENVISAT validation flight was again carried out from the Esrange-Kiruna launch pad in the midnight sun on 2/3 July 2003. Two limb sequences of spectra were measured between 00:13 UT and 01:23 UT (3 July). Logistical problems prevented such perfect coincidences (in space and time) with ENVISAT as was the case for the other flights. In particular, the temporal offset was a couple of hours. However, taking into account the advection of the air masses, the coincidence between MIPAS on ENVISAT and MIPAS-B was finally better than expected from the unfavorable flight situation.

The set of MIPAS-B measurements available for validation of MIPAS is summarized in Table 3.

For intercomparison purposes the MIPAS temperature profiles matching the correlative measurements were interpolated (linear interpolation in log pressure) to the fixed vertical retrieval grid of MIPAS-B and the related error covariances transformed accordingly (von Clarmann, 2006). The optimal interpolation method suggested by Ridolfi et al. (2006) was not used here due to the low reliability of the currently available MIPAS AKs for Arctic latitudes. The matching pairs of measurements were then used to calculate, at each MIPAS-B retrieval pressure level, the statistics of the profile differences as described by the equations reported in Sect. 3.1.

For this intercomparison, only profile data points in the pressure range from 5 to 228 hPa (corresponding approximately to altitudes of 35 and 10 km respectively) were considered. This choice is motivated by the fact that: a) above 35 km the spatial response functions (namely the AKs) of the two considered instruments are too different for a meaningful intercomparison (see also Ridolfi et al. 2006) and b) below 10 km the MIPAS profile data are affected by unquantified systematic errors due to the assumption of profile shapes below the lowermost retrieved data point in each scan.

The results of this intercomparison are summarized in Fig. 11 with the same format as Fig. 1. From this figure (see panel (a) of Fig. 11) it is clear that the intercomparison does not highlight any significant bias in the MIPAS data except at the lowermost altitude around 11 km where differences of up to 2 K are visible, still consistent with the expected systematic error of the difference. Furthermore, panel (b) of Fig. 11 suggests once again that the expected random error of the profile differences could have been under-estimated by about a factor of 2.5 at most altitudes. This latter conclusion, however, is not corroborated by the statistics; i.e., due to the relatively small number of coincident measurements (only 6), the 95% confidence interval of the determined standard deviation of the profile differences is quite large.

In this intercomparison the standard deviation of the profile differences exceeding the expected value of the combined random error below 15 km could be partly due to the retrieval algorithms of the two experiments that operate with different cloud filtering criteria and may also react differently to the presence of a residual optically-thin cloud in the FOV. Furthermore, in the whole altitude range considered for this intercomparison horizontal and vertical smoothing errors (not included in the error estimates presented in Fig. 11) may also have played a role in the determination of both the large standard deviations and the bias detected at 11 km (see Ridolfi et al. 2006).

## 6 Validation using lidar measurements

Temperature profiles are mainly derived from Rayleigh lidar scattering (see e.g. Hauchecorne and Chanin 1980) by assuming that the atmosphere follows the ideal gas law and is in hydrostatic equilibrium. Temperature measurements are deduced from direct backscattering of photons on molecules in the aerosol-free part of the stratosphere and mesosphere. The precision of the derived temperatures is determined by Poisson counting statistics. Most of the profiles retrieved have close to a 1 h integration time and 1 km vertical resolution (which is of the same order as the MIPAS resolution). The temperature profile can be determined from the measurements using a seed value at the top of the profile, which is taken from statistical atmospheric models such as the MSIS90E model (Hedin, 1991). In an aerosol free atmosphere, temperatures are typically measured from 30 to around 60 km with an accuracy close to 1 K (Keckhut et al., 2004). Above this altitude the accuracy deteriorates rapidly. Several lidar systems involved in the MIPAS validation have Raman capabilities, i.e. they record both the Rayleigh and the Raman backscattered signals. Some of these systems (see Sect. 6.2) supply temperature profiles covering the range from 70 to 15 km. In these cases the profiles are derived using: only the Rayleigh signal in the altitude range from 70 to 30 km; only the Raman signal for altitudes from 25 to 15 km;

**Table 3.** Time and geolocation of MIPAS-B measurements available for validation of MIPAS on ENVISAT.

MIPAS-B	Flight #11		Flight #13		Flight #14
Location	Aire sur l'Adour (France)		Kiruna (Sweden)		Kiruna (Sweden)
Date	24 Sep. 2002	24 Sep. 2002	20 Mar. 2003	21 Mar. 2003	3 Jul. 2003
Sequence name	Seq. S	Seq. 3	Seq. N3a	Seq. D15c	Seq. 3
Mean time	21:50 UT	22:21 UT	20:56 UT	08:48 UT	01:12 UT
Mean latitude	39.9° N	47.0° N	65.8° N	65.6° N	69.3° N
Mean longitude	1.1° E	0.7° E	14.6° E	17.5° E	11.0° E
Altitude range	11.3–38.8 km	5.9–38.4 km	11.1–31.1 km	9.1–31.2 km	7.9–39.1 km
MIPAS (v4.61)	Orb. 2975	Orb. 2975	Orb. 5508	Orb. 5515	Orb. 7004
Mean time	22:05/22:06 UT	22:07 UT	21:10 UT	09:08 UT	09:39 UT
Mean latitude	36.6/41.7° N	46.4° N	65.7° N	65.6° N	70.5° N
Mean longitude	2.6/1.6° E	0.6° E	14.1° E	17.1° E	10.9° E
Distance (@20 km)	207.4/358.3 km	78.9 km	78.3 km	28.2 km	2.1 km

and a combination of the two signals in the interval from 25 to 30 km.

The other lidar systems derive the temperature using solely the Rayleigh signal and therefore provide profiles only above 25 or 30 km.

These routine lidar operations yield profiles of temperature as a function of altitude, with unique accuracy in the upper stratosphere and lower mesosphere. These profiles may be used for the validation of MIPAS temperatures as a function of altitude at some specific sites. Two similar approaches have been conducted during the MIPAS validation phase. The first one is performed by the University of Bonn and used the Esrange lidar. The second one was performed within the EQUAL project (ENVISAT Quality Assessment with Lidar) and used several different lidars located in very different places.

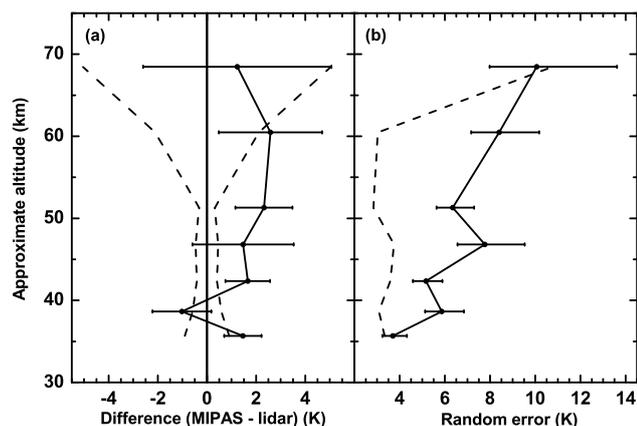
#### 6.1 Comparison with lidar measurements by the University of Bonn at Esrange

The University of Bonn lidar is located on Radarhill at Esrange near Kiruna, Sweden. Geographic coordinates are 67.88° N, 21.06° E and the altitude is 485 m ASL. The basic operation mode includes Rayleigh backscatter. The system operates at 532 nm with a vertical beam, range gates of 150 m width, and a power aperture product of 7 Wm<sup>-2</sup>. Details of the instrument are described by Blum and Fricke (2005). During the MIPAS measurement period the U. Bonn lidar was operated in 3 campaigns, one in summer 2002, one in winter 2002/03, and one in winter 2003/04. We accumulated about 687.5 h of integration time, which are recorded in 10 483 files with 5000 laser pulses each. The data quality is determined by the transmission of the troposphere. The data quality of a single file is measured by the top altitude which shows net signal after integrating those 5000 laser pulses (250 s). 97% of all recorded files reach at least 50 km altitude and 53% exceed 70 km. The precision of the derived temper-

atures is determined by Poisson counting statistics. Since the data are recorded as counts per range bin we improve the precision by adding several files, which of course reduces the temporal resolution.

Considering the results obtained from testing the coincidence criterion (see Sect. 3.2.1), to select lidar data suitable for validation we have searched our database for single lidar profiles that were measured within  $\pm 4$  h of a MIPAS measurement, within a range window of 500 km. We identified 70 such “coincidence windows”. For each coincidence we integrated the available lidar profiles to improve the measurement statistics and hence the precision of the lidar temperatures; the integration times varied from 26 to 131 min. The central times of the integrated lidar profiles were within  $-3.3$  to  $+2.2$  h of the MIPAS measurements.

The lidar temperature profiles are defined as a function of the altitude and typically span the range from 30 to 70 km. Pressure information relating to the altitude grid of the lidar profiles is not available, and therefore the comparison to MIPAS temperature was carried out by searching in the lidar profiles for the altitude levels of the corresponding MIPAS profiles, which is possible because the U. Bonn lidar profiles have a sampling step much finer than that of MIPAS. For this operation, the altitude grid of the MIPAS profiles was first corrected using the following approach exploiting the results of the intercomparison between MIPAS and radiosondes operated by the U. Bonn at Esrange (Sect. 3.2). For each MIPAS profile with a matching radiosonde profile, we searched the pressure measured by MIPAS in the radiosonde profile and extracted the difference between the MIPAS and radiosonde altitude. The global mean of this difference (averaged over all altitude grid points of all matching pairs of profiles) is 350 m, with the radiosonde altitude below the MIPAS altitude. This mean difference was interpreted as a *regional bias* of the MIPAS altitude scale and was therefore used to correct the altitude grid of the individual MIPAS profiles matching a lidar profile.



**Fig. 12.** Plot summarizing the agreement between temperatures of MIPAS and lidar measurements operated by the University of Bonn from Esrange. The format of the plot is the same as for Fig. 1.

The available pairs of measurements were restricted to a horizontal range of less than 500 km about Esrange for the MIPAS tangent location, and the differences, MIPAS – lidar temperatures, were computed in the altitude range from 35 to 68 km and grouped in altitude bins centered around the average MIPAS pointing altitudes. Again, in this case MIPAS AKs were not used to adjust the lidar profiles to the MIPAS vertical resolution, for the same reason reported in Sect. 3.2.

The results of the comparison with lidar measurements operated by the University of Bonn are summarized in Fig. 12, with the same format as Fig. 1. The expected systematic and random error budgets reported in this figure have been calculated using the same error components considered for the intercomparison reported in Sect. 3.2. Panel (a) of Fig. 12 shows that there seems to exist a positive bias of MIPAS of about 1.5 to 2 K; however this bias is only statistically significant (i.e. beyond the 95% confidence interval) for a subset of the considered altitude bins, and only for the bins at 42 and 52 km does it exceed (by a factor ranging from 3 to 4) the estimated systematic error of the differences.

Panel (b) of Fig. 12 shows that the standard deviation of the differences between the MIPAS and lidar temperatures generally tends to be larger than the expected random error of the differences. This occurs especially for the range from 42 to 60 km, where the standard deviation of the differences is up to a factor of 3 larger than the expected random error. In this range the component of the expected MIPAS random error due to the assumption of the temperature profile shape above the topmost retrieval altitude might have been underestimated. In fact, in the present calculation of the expected MIPAS random error, this error component is considered to be important only at 68 km and to vanish (rather unrealistically) very rapidly at lower altitudes (Dudhia, 2005). Another possible explanation for this finding could be that the combined random error in this intercomparison was evalu-

**Table 4.** Lidar stations involved in the EQUAL project and contributing to the temperature validation dataset used in this work.

Station	Lat. (deg.)	Long. (deg.)	Exploit Raman
Eureka	80.05	−86.42	Yes
Esrange	67.88	21.10	No
Hohenpeissenberg	47.80	11.02	No
Obs. Haute Provence	43.94	5.71	No
Tsukuba	36.05	140.13	Yes
Table Mountain	34.40	−117.70	Yes
Mauna Loa	19.54	−155.58	Yes
La Reunion	−21.06	55.47	No

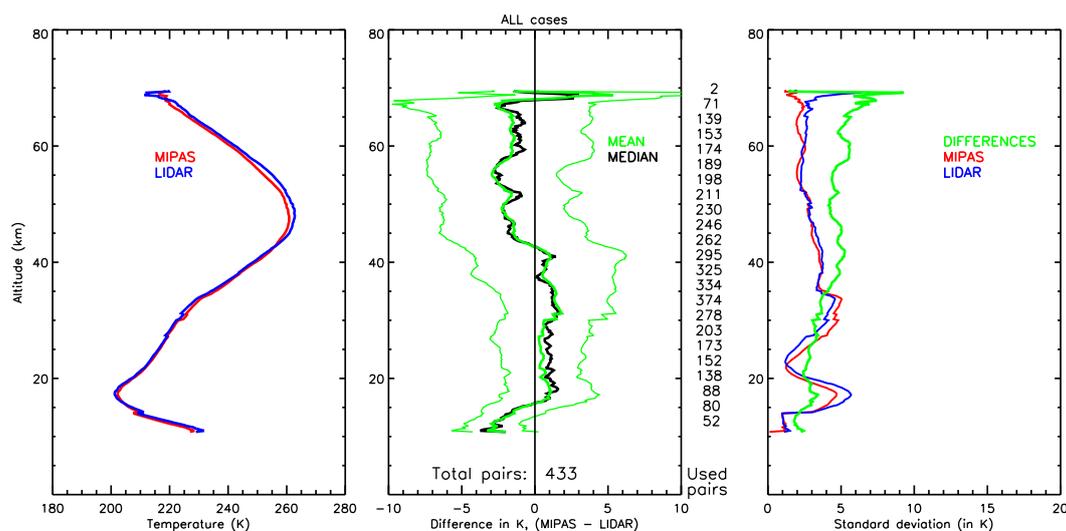
ated without consideration of the horizontal smoothing performed by MIPAS.

As a concluding remark, considering Figs. 4 and 12 (both concerning comparisons between MIPAS and U. Bonn measurements) we note that in the altitude range where the radiosonde and lidar measurements overlap (36–39 km), the predictions of the MIPAS data quality from the two types of intercomparison agree pretty well in terms of both bias and standard deviation.

## 6.2 Comparison with lidar measurements in the EQUAL project

The EQUAL project assesses the quality of ozone and temperature profiles derived from GOMOS, MIPAS and SCIAMACHY, the three chemistry instruments onboard ENVISAT, by comparison to lidar data from 13 stations spread worldwide. The large amount of validation data allows the analysis of the data quality for possible dependencies on several geophysical (e.g. latitude) and observational (e.g., solar illumination) parameters. Ozone and temperature profiles are measured with Differential Absorption Lidars (DIAL) and elastic backscatter lidars. Most of the lidars are also equipped with Raman channels. In the case of the DIAL lidars, the non-absorbed channel is pure elastic backscatter above 25–30 km and is used to derive temperature profiles. The lidar stations that contributed to building up the temperature validation data set processed in this work are listed in Table 4 (from north to south). Beyond the geolocation, Table 4 also indicates whether the individual contributing lidars have exploited the Raman signal to extend the temperature profiles below 25 or 30 km.

All lidar systems participating in the EQUAL project (except the one operating from Esrange) are part of the Network for the Detection of Atmospheric Composition Change (NDACC, [www.ndacc.org](http://www.ndacc.org)). Their measurements are regularly monitored for their quality via measurement and algorithm intercomparison campaigns performed under the NDACC protocol (Keckhut et al., 2004). Lidar profiles are



**Fig. 13.** Plot summarizing the agreement between temperatures from MIPAS and from lidar measurements operated in the frame of the EQUAL project. Left panel: average temperature profiles from MIPAS (red) and lidars (blue). Center panel: median (black), average and  $\pm 1\sigma$  standard deviation of the differences MIPAS – lidar (green). Right panel: standard deviations of the differences MIPAS – lidar (green), of MIPAS (red) and of lidar (blue) temperature profiles. The numbers reported on the right vertical axis of the center panel indicate the number of pairs of measurements involved in the intercomparison for some altitude bins.

routinely archived in the ENVISAT validation database at NILU ([www.nilu.no](http://www.nilu.no)).

For this intercomparison the selection of co-located pairs of MIPAS and lidar observations was based on matching criteria which were slightly relaxed with respect to the baseline established in Sect. 2.3. In order to get a sufficient number of matching profiles for a statistically meaningful comparison, the space and time coincidence criteria were set here to 400 km and 20 h respectively. This choice was made after verification that the overall conclusions of the intercomparison do not change if more stringent matching criteria, such as 200 km and 5 h, are used. In this latter case, however, only 39 matching profiles are found, leading to very poor statistics and significance of the intercomparison. With the adopted coincidence criteria, a total of 740 matching pairs of temperature profiles were identified and used to validate MIPAS temperature.

Since the altitude grid of the MIPAS Level 2 profiles is known to be affected by a large bias (Raspollini et al., 2006; Kiefer et al., 2007), prior to the intercomparison a corrected MIPAS altitude grid was set up, using the MIPAS retrieved pressures and ECMWF profiles of pressure versus Geopotential Height (GPH): the ECMWF pressure and GPH were interpolated to the MIPAS pressure grid and the resulting GPH values were then translated to geometric altitude. The comparison between MIPAS and lidar measurements was then carried out by merely calculating the statistics of the differences between co-located pairs of profiles interpolated to a common and fixed altitude grid with a step of 200 m. The intercomparison takes place in the altitude range from 10 to

68 km and is limited to the MIPAS profiles with a maximum expected random error smaller than 1 K.

Again, in this intercomparison MIPAS AKs were not used to adapt lidar profiles to the MIPAS vertical resolution. This choice is again motivated by the fact that the validation dataset is not restricted to the standard atmospheric conditions for which MIPAS AKs are available. Furthermore, the nominal vertical resolution of MIPAS ( $\approx 3$  km) is pretty similar to that of the lidar measurements used here ( $\approx 2$  km), and therefore we also expect that the smoothing error on the difference profiles will be small enough compared with the other error components affecting MIPAS measurements.

The results of the comparison for the whole set of co-located pairs are summarized in Fig. 13. The left panel of this figure shows the mean temperature profiles of MIPAS (red) and the lidars (blue). The center panel shows the median (black) of the differences, MIPAS – lidar, along with the mean and the standard deviation (green curves). On the right vertical axis of the center panel we also report, for some altitude bins, the population that contributed to the statistics. The right panel of Fig. 13 shows the standard deviations of MIPAS (red), of the lidar (blue) and of the difference, MIPAS – lidar (green).

From the center panel of Fig. 13 we see that:

- The mean and the median of the distribution of the differences nearly coincide in the whole altitude range considered. Therefore the number of asymmetric outliers is very small.
- In the altitude range from 12 to 15 km the MIPAS temperature has a negative bias with respect to the li-

dars. The bias amounts to about  $-2.5$  K and, even if it is slightly larger than the expected systematic error of MIPAS ( $\approx 1.2$  K, see Dudhia 2005), it is still within the combined systematic error of the difference profile. (As explained earlier, the systematic error of the lidar measurements could also be significant in this altitude range.)

- From 15 to 42 km MIPAS has a positive bias with respect to the lidars. This bias amounts to 1.5 or 2 K, and is consistent with the expected systematic error of MIPAS ( $\approx 1.5$  to 2 K in this range).
- From 42 to 68 km, the MIPAS bias again becomes negative with respect to the lidars and amounts to about 1.5–2 K. The value of the bias here is still consistent with the expected systematic error of MIPAS ( $\approx 2$  K in this range).

At a first glance this latter finding may seem to contradict the conclusions of the intercomparison of MIPAS against the lidar operated from Esrange by the University of Bonn (see Fig. 12 of Sect. 6.1). However, if we break down the intercompared dataset by latitude bands we find the results reported in Fig. 14. This figure consists of three frames, each with the same format as Fig. 13 and referring, from top to bottom, to the following latitude bands respectively: mid-latitude (from  $23.5$  to  $66.5^\circ$  latitude); tropical (from  $0$  to  $23.5^\circ$  latitude); and polar (from the latitude of  $66.5^\circ$  to the Pole).

From Fig. 14 we can see that the discrepancies between the MIPAS and lidar temperatures at mid- and tropical latitudes are pretty similar. Given the large size of the intercompared sample at these latitudes, the global results of the intercomparison shown in Fig. 13 are driven by the agreement found at these latitudes. The discrepancies observed in polar conditions (bottom frame of Fig. 14), however, are significantly different compared to the other latitude bands. In particular, for these latitudes the intercomparison suggests a positive bias of MIPAS with respect to the lidars from 42 to 68 km, in agreement with Fig. 12. The presence of a significant number of outliers (deduced from the difference between the mean and the median of the distribution) and the small size of the intercomparison sample does not permit, however, a more quantitative investigation of the nature of the discrepancies observed in this latitude band.

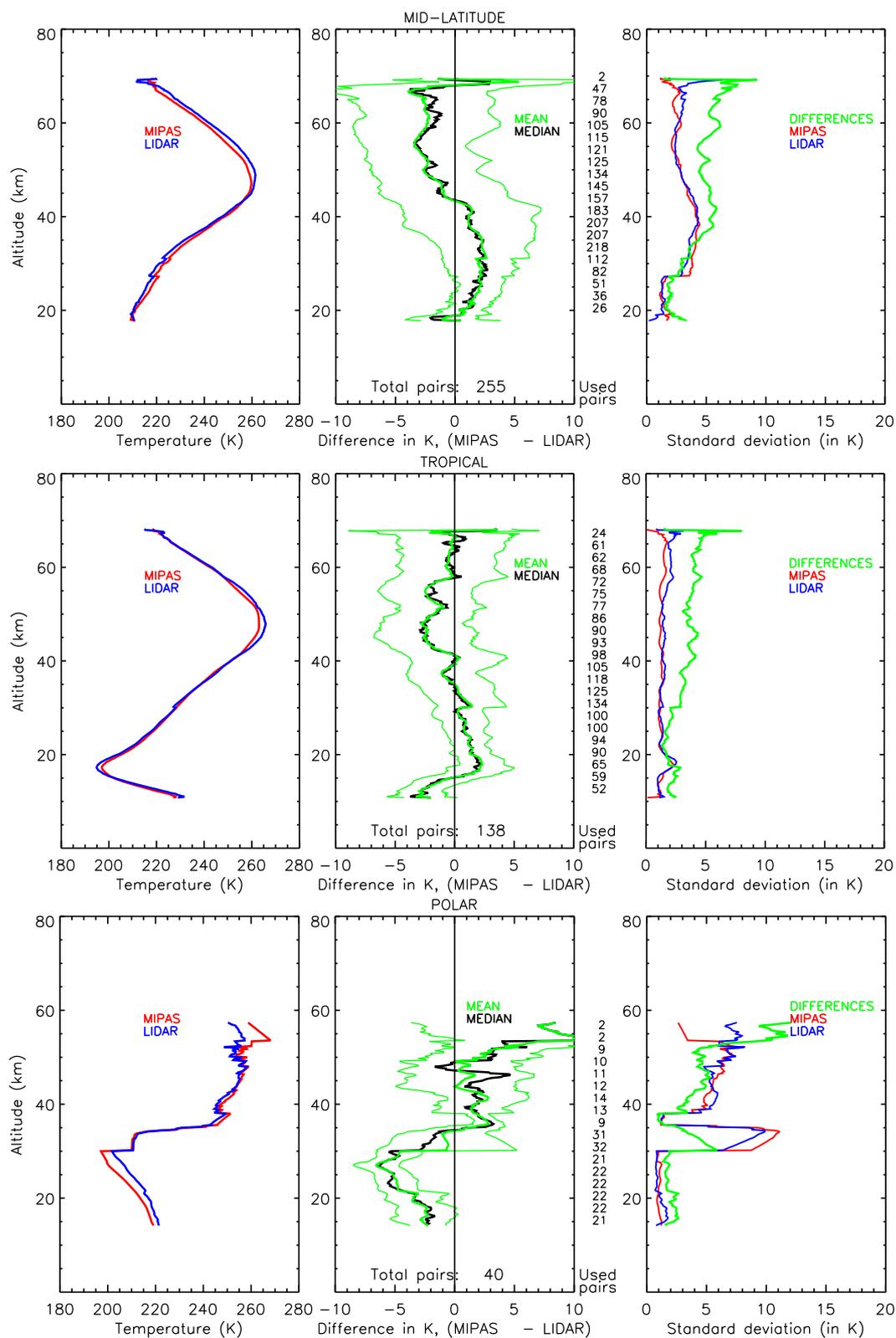
The global behavior of the bias versus altitude, as shown in the center panel of Fig. 13, also suggests that there could be a vertical shift between the altitude grids of MIPAS and the lidars. In fact, such an altitude shift would imply different positions for the temperature maximum at the stratopause and consequently a “first derivative” or a “wave like” shape of the profile differences. A similar behavior of the bias is found in the differences between MIPAS and ECMWF analyses (see Sect. 7). This hypothesis of a vertical shift will be further discussed in Sect. 8.

The right panels of Figs. 13 and 14 show that, in any case, in the altitude range from 40 to 60 km the observed standard deviation of the differences, MIPAS – lidar temperature, always exceeds its expected random error (Dudhia, 2005) by a factor of about two. This result again suggests that the MIPAS random error due to the assumption of the profile shape above the topmost retrieved profile point might have been underestimated while evaluating the expected MIPAS error budget. We also note however that both horizontal and vertical smoothing errors (not evaluated in this intercomparison) might have contributed to the observed effect.

## 7 Comparison with ECMWF data

The temperature comparisons reported here are based on meteorological analyses provided by the ECMWF operational system and archived at the British Atmospheric Data Centre (BADC). Analyzed temperature fields, provided every 6 h UT, had already been interpolated by the BADC onto a  $1.125 \times 1.125$  degree latitude/longitude grid. The data were then linearly interpolated in space and time to the location of the MIPAS data point (taken to be the average of the latitudes/longitudes of all the data points in the scan). ECMWF data have a vertical resolution of approximately 1.5 km in the middle stratosphere (between 70 and 3 hPa), decreasing with altitude above this, and increasing at lower levels.

Quantitative errors are not provided with the ECMWF data, and therefore no errors have been included for the ECMWF data in the validation work presented here. However, qualitatively ECMWF temperatures, particularly for altitudes below 30 km ( $\approx 10$  hPa), are generally believed to be of good quality. For example, validation against temperatures from CHAMP (Gobiet et al., 2005) indicate the seasonal zonal mean temperature biases between 10 and 30 km were generally smaller than 0.5 K. They suggested however that there is a cold bias at the low latitude tropopause (also observed against radiosondes, Simmons 2003), and a wave like bias structure in the Southern winter polar vortex in 2003. Simmons et al. (Simmons et al., 2005) compared the ECMWF temperature data against radiosondes for the southern hemisphere (SH) extratropics whilst studying the break up of the SH polar vortex in 2002, and concluded that mean errors in comparison to radiosondes were of the order of 0.5 K for temperature, with the suggestion that the analysis errors were much smaller than the errors of the radiosonde observations (including mismatch in location of radiosonde/ECMWF observations), and found that the implied random analysis errors were a few tenths of a Kelvin in temperature. At higher altitudes, there is known to be a cold bias at the model top over the winter pole, which can lead to some unrealistic structures in the ECMWF temperature profiles in the upper stratosphere and lower mesosphere in these polar regions (Dethof, 2003).



**Fig. 14.** Plot summarizing the agreement between temperatures from MIPAS and from lidar measurements operated in the frame of the EQUAL project. The validation dataset was broken down into three latitude bands. Top frame refers to mid-latitudes, center frame to tropics and bottom frame to polar latitudes. The format of each frame is the same as for Fig. 13.

## 7.1 Comparison to MIPAS

The MIPAS validation has been limited to version 4.61 data only, and has been performed for all unique and *good* profiles in the specified time periods. Profiles were not included in the comparison if the quality or convergence flag indicated a bad retrieval, or if there were (meaningless) negative variances in the profile. There were found to be very occasional cases where the pressure profile did not decrease monotonically with altitude, and these were rejected when they occurred for ease of analysis. In the case of more than one file containing the same profile, only one of the overlapping profiles were used.

The comparison has been performed for seasonal averages between July 2002 and March 2004. For each case the difference between each MIPAS profile and its corresponding ECMWF profile was calculated, first regridding the ECMWF profile to the pressure grid of the corresponding MIPAS profile by linear interpolation in the  $\log(\text{pressure})$  domain. Since the present intercomparison is not limited to mid-latitudes, and since the spatial response function of the ECMWF field is not punctual, the MIPAS averaging kernels, which are available only for standard atmospheric conditions, were not applied (Ceccherini, private communication, 2006<sup>1</sup>).

Seasonal averages were calculated by binning all the data into fixed latitude and pressure bins, with the mean value in each bin reported at the approximate altitude of the mean of all the pressures in that bin. The chosen pressure bins were defined about the seasonal mean value of pressure retrieved at each MIPAS level. Furthermore, in order to prevent correlations from more than one point from a single profile existing in any given pressure bin, only one point per bin, per profile was included. Although the top and bottom levels may have fewer points than others, there are no cases where they only have a few points.

For each pressure bin we calculated the seasonal (per year) mean difference between MIPAS and ECMWF (MIPAS - ECMWF) and the standard deviation. In Fig. 15 we show the results of a “global average” i.e. obtained by including in the statistical analysis pairs of MIPAS and ECMWF profiles relating to the whole globe. Figure 15 contains eight frames, each of which refers to a 2- or 3-monthly average as indicated by the key on top of the frame itself. In turn, each frame contains two plots with the same format as was adopted for Fig. 1. We must note however that, due to the very large number of matching profiles, in these plots the error bars reporting the 68% confidence intervals for the bias and the standard deviation are not visible. Furthermore, for this intercomparison, since ECMWF does not supply quantitative estimates for the error on their fields, the plotted errors refer only to the MIPAS profiles. Specifically, for the systematic error of MIPAS we assumed the estimates supplied by Dudhia (2005), and for the random component of the error due to measurement noise we assumed the estimate given in the MIPAS Level 2 files.

In this intercomparison we also neglected the smoothing error component that originates from the differences in the spatial response functions of the MIPAS measurements and of the ECMWF model analysis. This choice is motivated by the contingent background in which the intercomparison is being operated: MIPAS AKs are available only for standard atmospheric conditions and ECMWF AKs are not available at all.

Since both the ECMWF and the smoothing errors have not been included in this analysis, it may be expected that the errors shown on the difference plots are an underestimate.

Given the very large amount of coincident MIPAS and ECMWF temperature estimates available for the validation, we also tried to break down the intercomparison dataset into the following latitude bands: 90 S–65 S, 65 S–20 S, 20 S–20 N, 20 N–65 N and 65 N–90 N. Figures equivalent to Fig. 15 for the individual latitude bands show behaviors of the agreement versus altitude generally similar to the ones reported in Fig. 15, therefore we do not report all of them. Here, we only report Fig. 16, relating to the comparison in the tropical region, because in this latitude band the agreement between MIPAS and ECMWF is slightly worse compared to the other bands.

In order to characterize, for each considered time period and latitude band, the consistency of the estimated bias between MIPAS and ECMWF with the systematic error of MIPAS, we also defined a quantifier  $\varepsilon$  as follows:

$$\varepsilon^2 = \frac{1}{N} \sum_{j=1}^N \frac{(X_{M_j} - X_{C_j})^2}{\sigma_{T_j}^2} \quad (13)$$

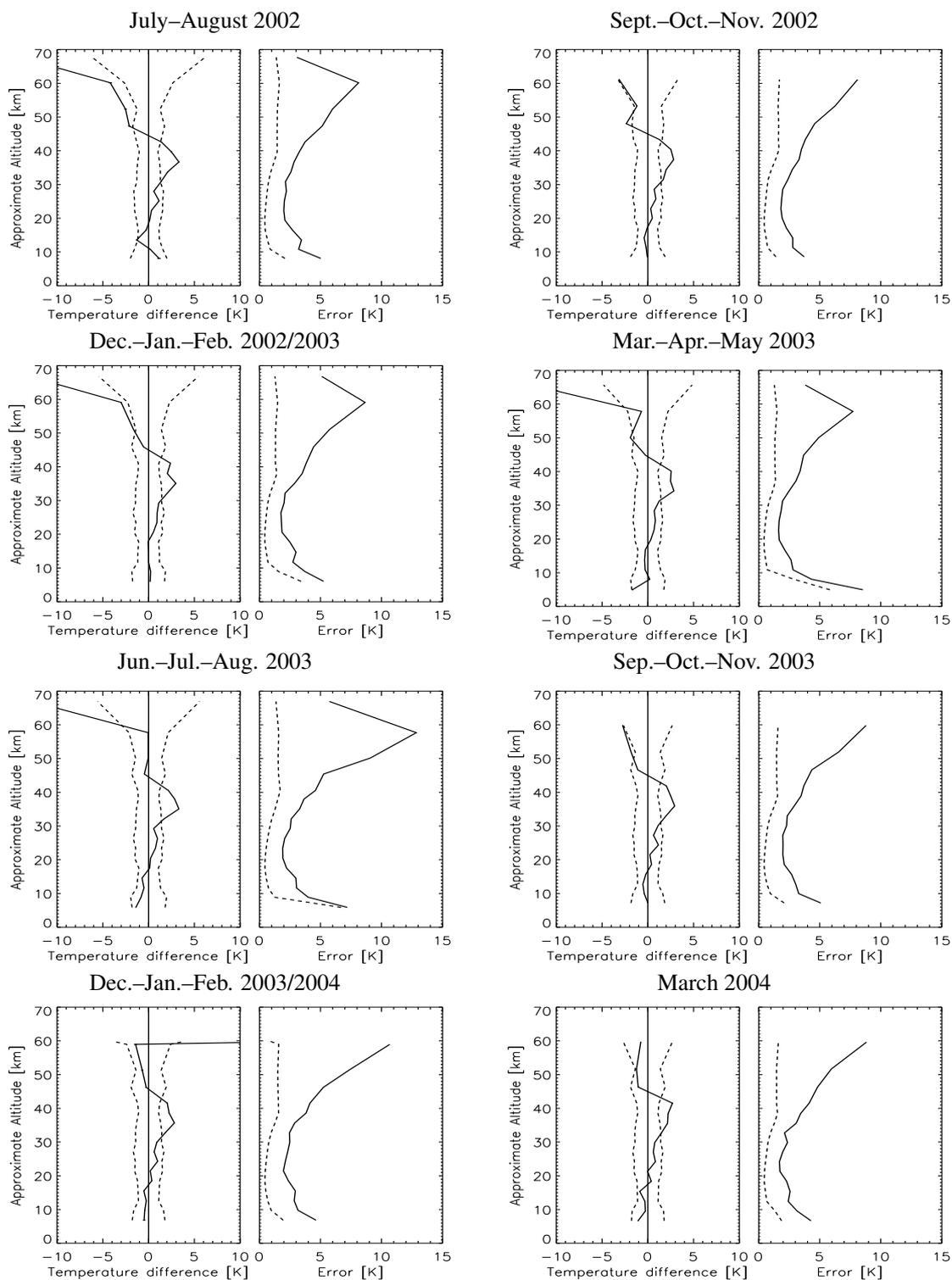
where  $N$  is the number of bins in the mean profile,  $X_{M_j}$  and  $X_{C_j}$  are the MIPAS and the ECMWF mean temperatures respectively in the  $j$ -th pressure bin and  $\sigma_{T_j}$  is the total error of the mean MIPAS profile which practically coincides with the MIPAS systematic error due to the typically large number of averaged temperature values. It is also worth noting that the topmost and the lowermost retrieved data points in each MIPAS temperature profile (corresponding to nominal altitudes of 68 and 6 km respectively) have been demonstrated to be especially affected by errors due to the profile shape assumptions outside the vertical range of the retrieval. This type of error is very difficult to quantify, therefore the ESA itself recommends not to make scientific use of these data points. For this reason for the calculation of  $\varepsilon$  we omitted the pressure bins corresponding to the extreme profile points of MIPAS.

The obtained values for the quantifier  $\varepsilon$  are reported in Table 5. Values exceeding unity indicate that the observed bias in the differences, MIPAS - ECMWF, is larger than the expected systematic error of MIPAS.

## 7.2 Discussion of results

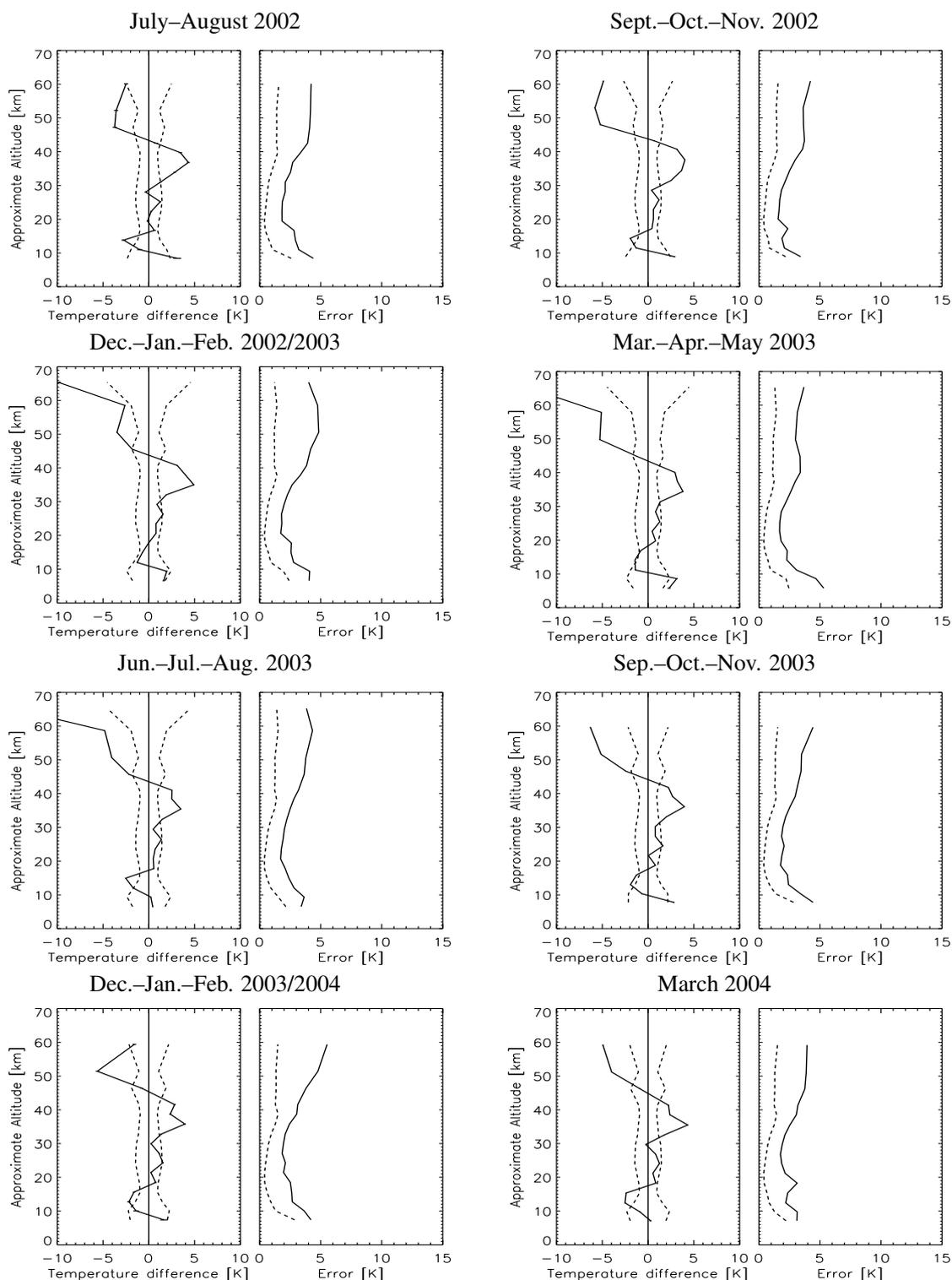
The results of the comparison of MIPAS against ECMWF temperatures can be summarized as follows:

## Global Mean



**Fig. 15.** Global mean seasonal differences (MIPAS-ECMWF). For each season, the left hand plot shows the mean temperature difference profile (solid line); error bars giving the 68% confidence intervals are not visible due to the large number of matching pairs. The dashed lines show  $\pm$  the estimated systematic error of MIPAS. The right hand plot shows the standard deviation (solid line) and the expected random error (dashed line) for MIPAS.

## Tropical Mean



**Fig. 16.** Tropical (20S–20N) mean seasonal differences (MIPAS-ECMWF). For each season, the left hand plot shows the mean temperature difference profile (solid line), error bars giving the 68% confidence intervals are not visible due to the large number of matching pairs. The dashed lines show  $\pm$  the estimated systematic error of MIPAS. The right hand plot shows the standard deviation (solid line) and the expected random error (dashed line) for MIPAS.

**Table 5.** Quantifier for the bias between MIPAS and ECMWF, calculated from Eq. 13.

	J-A 2002	S-O-N 2002	D-J-F 02/03	M-A-M 2003	J-J-A 2003	S-O-N 2003	D-J-F 03/04	March 2004
Global	1.39	1.12	1.13	1.07	1.13	1.16	1.01	1.01
90 S–65 S	1.66	0.93	1.10	1.09	2.34	0.90	1.03	0.91
65 S–20 S	1.56	1.39	1.68	1.15	1.39	1.53	1.59	1.40
20 S–20 N	2.02	2.37	2.15	2.08	1.93	2.15	2.07	2.02
20 N–65 N	1.76	1.39	1.37	1.42	1.68	1.44	1.27	1.46
65 N–90 N	0.93	1.07	0.87	0.83	1.01	1.21	1.79	1.95

- Occasionally very large (even greater than 10 K) discrepancies appear, corresponding to the extreme MIPAS profile points (6 and 68 km). As pointed out above, these discrepancies are not unexpected, therefore, in the subsequent discussion we ignore the pressure bins corresponding to the topmost and lowermost MIPAS profile points.
- Below approximately 30 km ( $\approx 10$  hPa), differences between MIPAS and ECMWF are generally much smaller than the assumed systematic errors of MIPAS, with differences typically less than 1 K. An exception occurs in the tropical region, where around 100 hPa, MIPAS temperatures are generally lower than ECMWF by around 2 K (approximately equal to the MIPAS expected systematic error, see Fig. 16).
- Between 10 hPa and 1 hPa (approximately between 30 and 48 km), MIPAS is generally a few Kelvin higher than ECMWF. This is particularly pronounced in the tropics (see Fig. 16), where differences of up to 5 K can be seen (larger than the expected systematic errors).
- Above 48 km (or pressures less than 1 hPa), MIPAS is generally lower than ECMWF, although a few exceptions exist. (e.g. over the SH winter pole, although it is known that ECMWF has problems in this region.)
- Values of the quantifier  $\varepsilon$  defined in Eq. (13) are normally between 1 and 2 (see Table 5), showing that on average the expected MIPAS systematic error has not been grossly underestimated. From Table 5 we also see that the defined quantifier reflects the poorer agreement between MIPAS and ECMWF at the tropics.
- At all altitudes there can be occasional outlying points where very large differences can be seen, particularly at the lowest MIPAS levels. The standard deviation can be seen to increase at the top and bottom of the comparison.
- In general the standard deviation of the differences between MIPAS and ECMWF exceeds the expected ran-

dom error of MIPAS by a factor of 2 or 3. This occurs in the whole altitude range of MIPAS, however the effect is largest at the edges of the profiles, where the MIPAS random error due to the profile shape assumption outside the vertical retrieval range might have been severely underestimated (by up to about a factor of 10 in extreme cases). At the lowest considered altitudes cloud detection scheme that is not conservative enough could also contribute to the observed inconsistency. We must stress however that no final conclusion on the MIPAS random error can be drawn from this intercomparison. One reason is, because of the unavailability of the AKs of the compared profiles, it was not possible to include in the analysis the effect of both vertical and horizontal smoothing errors.

- Test calculations (not reported in this paper) have shown that the agreement between MIPAS and ECMWF worsens slightly if MIPAS AKs are applied to the ECMWF profiles prior to the comparison. This result corroborates the assumption that the standard MIPAS AKs are not adequate to account for the smoothing error when data sets relating to the whole globe and all seasons are intercompared.

## 8 Summary of results

Different correlative measurements have provided a validation in different altitude ranges, therefore in this Section we discuss separately the results for the two altitude ranges from 6 to 36 km and from 30 to 70 km. The results of the comparison with ECMWF analyses is also discussed separately because, although ECMWF fields cover the whole altitude range of MIPAS, they are not the direct result of correlative measurements.

### 8.1 Altitude range from 6 to 36 km

The range from 6 to 36 km is covered by measurements from radiosondes, a temperature sensor installed onboard a stratospheric balloon (SPIRALE, 12–30 km), a stratospheric bal-

loon version of MIPAS (MIPAS-B, 11–36 km) and a few lidars with Raman capabilities (see Table 4). In this altitude range the conclusions on MIPAS bias and precision can be summarized as follows:

**Bias.** In several cases we find a bias statistically consistent with zero (see comparison with L'Aquila and Potenza radiosondes and MIPAS-B). For a limited number of altitude levels, a bias different from zero is found to be statistically significant in the comparisons against radiosoundings from Esrange and radiosoundings provided by GAW contributing networks. Whenever statistically significant, the absolute value of the bias ranges from 1 to 2 K, consistent with the expected systematic error budget of the compared measurements. Furthermore, whenever the bias is statistically significant, its altitude-dependent value seems to depend also on time and on the latitudinal band to which the intercomparison is restricted (see e.g. Figs. 8 and 9). This finding suggests that the detected biases are most likely to be attributable to error sources (such as MIPAS retrieval assumptions) that vary on temporal and latitudinal scales larger than those covered by the individual datasets used in the intercomparisons.

**Precision.** The observed standard deviation of the differences between the MIPAS and correlative measurements tends to be larger than the combined random error. This occurs especially at altitudes below 15 km, where the cloud filtering applied to the MIPAS observations may not have been effective enough. In some of the presented comparisons, however, differences due to unaccounted spatial smoothing may also have contributed significantly to the observed large standard deviations.

## 8.2 Altitude range from 30 to 70 km

The altitude range from 30 to 70 km is covered by lidar measurements. In this work we have considered measurements acquired by a lidar operated by the University of Bonn from Esrange and eight NDACC lidar stations (see Table 4) contributing to the EQUAL project. In this height range the conclusions can be summarized as follows:

**Bias.** The good statistics of correlative measurements available to the EQUAL project makes it possible to perform a validation with pseudo-global coverage. This pseudo-global validation indicates a statistically significant negative bias of MIPAS temperature of about 2 K in the altitude range from 45 to 65 km. This bias reverses its sign around 42 km and becomes positive, with an amplitude of about 1 K at lower altitudes. These “global” values of the bias, however, cannot be used to apply corrections to the MIPAS temperature because if we breakdown this intercomparison data set by latitude band we discover that the bias is latitude-dependent. In particular we find that for Northern Polar latitudes there seems to exist a positive bias of MIPAS of about 2 K in the range from 35 to 55 km, in agreement with the results of the comparison to lidar measurements done at Esrange by the University of Bonn. The detected biases are in

any case consistent with the expected combined systematic error of the compared measurements.

**Precision.** The standard deviation of the profile differences exceeds the combined random error estimate by a factor of 2 to 3 in the altitude range from 40 to 60 km. The intercomparisons against lidar measurements from the University of Bonn and from the EQUAL project agree fairly well with this conclusion.

The excess bias and standard deviation detected in the intercomparisons against lidar measurements could be due both to the neglect of horizontal and vertical smoothing effects in the intercomparisons, and/or to an underestimate of the expected MIPAS random error budget. We argue that probably the MIPAS error component due to the assumption of the profile shape above the topmost retrieved profile point could have been underestimated.

## 8.3 Results of comparison with ECMWF

In this work we also compared MIPAS temperature with the corresponding meteorological analyses of the ECMWF operational system. The conclusions of this type of intercomparison can be summarized as follows:

**Bias.** Below 30 km the detected bias is consistent with the MIPAS systematic error ( $\approx 1.5$  K). Above this altitude, the average difference between MIPAS and ECMWF may exceed the expected MIPAS systematic error. However at these altitudes the accuracy of ECMWF temperature is also controversial and it is not possible to assign the whole amount of bias detected exclusively to a MIPAS deficiency.

**Precision.** In general the standard deviation of the differences between MIPAS and ECMWF exceeds the expected random error of MIPAS by a factor of 2 or 3. This occurs in the whole altitude range of MIPAS, however the effect is largest at edges of the profiles, where the MIPAS random error due to the profile shape assumption outside the vertical retrieval range might have been severely underestimated (by up to about a factor of 10 in extreme cases). At the lowest considered altitudes a cloud detection scheme which is not conservative enough could also contribute to the observed inconsistency. We must stress however that in this case no final conclusion can be drawn regarding the MIPAS random error. One reason is, because of the unavailability of reliable AKs of the compared profiles, it was not possible to take into account in our analysis the effect of both vertical and horizontal smoothing.

As a further consideration it is worth pointing out that the observed biases of MIPAS with respect to ECMWF analyses have a wave like behavior versus altitude. A very similar behavior of the bias between ECMWF and external, independent, non-MIPAS measurements has recently also been observed by Niels Bormann (Bormann and Thépaut, 2006; Bormann et al., 2006). Antje Dethof also found a similar behavior of the bias while comparing MIPAS near real-time temperature retrievals to ECMWF analyses (De-

thof, 2003). The current understanding is that, at least in the upper stratosphere (above  $\approx 10$  hPa), the observed oscillations belong to the ECMWF temperature and are caused by a discrepancy between model biases and radiance biases from primarily nadir sounders (like the Advanced Microwave Sounding Unit, AMSU-A, integrated with the Atmospheric Infrared Sounder, AIRS on the Earth Observing System polar-orbiting platform, EOS Aqua). Such a problem in the ECMWF temperature could then also translate into an altitude offset when GPH is first derived from ECMWF pressure and temperature (using the barometric equation), and subsequently used to reconstruct the altitude scale of MIPAS profiles to be compared with lidars (see Sect. 6.2). Indeed, the wave like behavior of the bias observed in the comparison with the EQUAL project lidars can be mostly explained by an altitude shift between the compared profiles of about 800 m (MIPAS being shifted toward lower altitudes).

Although the aforementioned problem in the ECMWF temperature can certainly provide a significant contribution to the observed biases, at present we cannot also exclude a small bias in the MIPAS retrieved pressures. Theoretically, from the expected error budget of MIPAS pressure (Dudhia, 2005) we see that spectroscopic uncertainties are expected to be the main systematic error source of pressure above 30 km.

## 9 Conclusions

In this validation work we compared MIPAS/ENVISAT temperature profiles retrieved by the ESA processor versions 4.61 and 4.62 with correlative measurements acquired by independent instruments adopting different measurement techniques. A large number of teams have participated in this work, each of them carrying out the intercomparisons independently, and sometimes also using different methods; however all the teams reached similar conclusions.

Globally we find that the absolute value of the bias of MIPAS temperature profiles is generally smaller than 1 or 2 K depending on altitude and latitude. This estimate agrees pretty well with the MIPAS error predictions based on error propagation analyses.

Regarding the MIPAS random error estimated from the intercomparison, we find that it is larger (typically by a factor of two to three) than the corresponding expected value (Dudhia, 2005) derived on the basis of error propagation. This occurs especially at the edges of the altitude range covered by the MIPAS limb scan. Insufficient cloud-flagging at low altitudes and underestimation of the error due to the profile shape assumption outside the vertical retrieval range of MIPAS are possible candidates responsible for the excessive random error observed in this intercomparison. We must state however that our analysis cannot be conclusive on this point because in many cases it was not possible to account properly for the spatial smoothing operated by MIPAS.

For a single intercomparison data set (see Fig. 5) we also verified that the differences between the MIPAS and correlative measurement temperatures do not depend on the value of the temperature itself.

Finally, from the intercomparisons carried out in this work there is no evidence of unpredicted significant variations with time of the quality of the MIPAS temperature.

*Acknowledgements.* The analysis presented in Sect. 3.1 was carried-out during a co-location in which some of us (M.R., S.C. and T.v.C.) were hosted by the International Space Science Institute of Bern (Switzerland). The authors are grateful to C. Bellotti, N. Bormann, M. Carlotti, C. Cornacchia, V. Cuomo, and P. Raspollini for fruitful discussions and technical support.

The work reported in Sects. 3 and 6 was partly supported by the ESA-funded projects EQUAL and TASTE, by ProDEX and the Belgian Science Policy Office (CINAMON project), by the French CNES (VETO project) and INSU/CNRS. All these projects rely on nationally funded lidar and radiosonde station activities. In particular, lidar activities of the University of Bonn were supported by grant 50EE0009 from DLR Erdbeobachtung. Most of the radiosonde and lidar data used in this publication were obtained as part of WMO's GAW contributing networks including the Network for the Detection of Atmospheric Composition Change (NDACC, formerly NDSC) and the Southern Hemisphere Additional OZonesondes programme (SHADOZ). They are now publicly available through the NDACC (see <http://www.ndacc.org>), SHADOZ (<http://croc.gsfc.nasa.gov/shadoz>) and WOUDC (<http://www.woudc.org>) data archives.

The authors are grateful for contributions from the following institutes, and thank their co-workers who contributed to the generation of these data and to fruitful discussions: University of Bonn (Esrance station, lidar PI K.H. Fricke), MSC (Eureka station, lidar PI K. Strawbridge), DWD (Hohenpeissenberg Observatory, PI H. Claude), Université de la Reunion (Saint-Denis de La Reunion, PI H. Bencherif), NASA/JPL (Mauna Loa Observatory and Table Mountain Facility, PI I.S. McDermid), CNRS and CNES (Observatoire de Haute Provence, PI P. Keckhut), NIES (Tsukuba station, PI H. Nakane), and all other institutes listed in Table 2 and corresponding co-workers. Authors are also grateful to T. Leblanc from JPL/NASA and H. Bencherif from LACY.

The SPIRALE technical team (M. Chartier, L. Pomathiod, T. Lemaire, G. Chalumeau, J.-C. Rimbault, T. Vincent) is acknowledged for making the flight possible and successful. The campaigns were funded by the European Space Agency (ESA), the French Space Agency (CNES). The MIPAS and SPIRALE balloon teams thank the CNES launching team for the success of the validation campaigns.

Financial support by the DLR (Project 50EE0020) and ESA for the MIPAS-B balloon flights is gratefully acknowledged. We thank the Centre National d'Etudes Spatiales (CNES) balloon launching team and the Swedish Space Corporation (SSC) team at Esrange for excellent balloon operations and for their always quick and uncomplicated support during measurement campaigns. We also acknowledge the Free University of Berlin (K. Grunow and B. Naujokat) for meteorological support.

The authors acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing data which were used

in this paper. They are grateful to the British Atmospheric Data Centre (BADC) for providing access to these ECMWF data.

Finally, a special acknowledgement goes to the work performed by the Quality Working Group established by ESA for verification and monitoring of MIPAS products.

Edited by: P. Espy

## References

- Antikainen, V. and Turtiainen, H.: Solar and infrared temperature correction studies with a special Vaisala RS80 radiosonde. Preprints, WMO Tech. Conf. on Instruments and Methods of Observation, Vienna, Austria, WMO, 149–153, 1992.
- Antikainen, V. and Jauheainen, H.: The New Vaisala RS90 radiosonde. Preprints, Ninth Symp. on Meteorological Observation and Instrumentation, Charlotte NC, Amer. Meteor. Soc., 170–175, 1995.
- Bevington, P. R. and Robinson, D. K.: *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed. McGraw-Hill, New York, 2003.
- Blum, U. and Fricke, K. H.: The Bonn University Lidar at the Esrange: Technical Description and Capabilities for Atmospheric Research, *Ann. Geophys.*, 23, 1645–1658, 2005, <http://www.ann-geophys.net/23/1645/2005/>.
- Bormann, N. and Thépaut, J.-N.: Assimilation of MIPAS limb radiances in the ECMWF system. Part I: Experiments with a 1-dimensional observation operator, Technical Memorandum n. 495 of the European Centre for Medium-Range Weather Forecasts, available from [www.ecmwf.int](http://www.ecmwf.int), 2006.
- Bormann, N., Healy, S. B. and Hamrud, M.: Assimilation of MIPAS limb radiances in the ECMWF system. Part II: Experiments with a 2-dimensional observation operator and comparison to retrieval assimilation, Technical Memorandum n. 496 of the European Centre for Medium-Range Weather Forecasts, available from [www.ecmwf.int](http://www.ecmwf.int), 2006.
- Calisesi, Y., Soebijanta, V. T., and van Oss, R.: Regridding of remote soundings: Formulation and application to ozone profile comparison, *J. Geophys. Res.*, 110, D23306, doi:10.1029/2005JD006122, 2005.
- Ceccherini, S., Carli, B., Pascale, E., Prosperi, M., and Dinelli, B. M.: Comparison of measurements made with two different instruments of the same atmospheric vertical profile, *Appl. Opt.*, 42, 6465–6473, 2003.
- De Clercq, C. and Lambert, J.-C.: A forward model of limb infrared emission spectra in a two-dimensional atmosphere, Proc. ESA Atmospheric Science Conference, Frascati, Italy, 8–12 May 2006, ESA SP-628, 2006.
- Dethof, A.: Monitoring of retrievals from the MIPAS and SCIAMACHY instruments on board ENVISAT, Contract report to the European Space Agency, available from [www.ecmwf.int](http://www.ecmwf.int), December 2003.
- Dudhia, A.: MIPAS-related section of the web-site of the Oxford University: [www.atm.ox.ac.uk/group/mipas/err/](http://www.atm.ox.ac.uk/group/mipas/err/), 2005.
- ECMWF: The Description of the ECMWF/WCRP Level III-A Global Atmospheric Data Archive, 1995.
- 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., Blom, C., Oelhaf, H., Carli, B., Carlotti, M., Delbouille, L., Ehhalt, D., Flaud, J.-M., Isaksen, I., Lopez-Puertas, M., McElroy, C. T., and Zander, R.: *Envisat-MIPAS an instrument for atmospheric chemistry and climate research*, edited by: C. Readings and R. A. Harris, ESA Publication SP-1229, European Space Agency, 2000.
- Fischer, H., Birk, M., Blom, C., et al.: 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 characterization of the balloon-borne Michelson Interferometer for Passive Atmospheric Sounding (MIPAS-B2), *Appl. Opt.*, 43, 3335–3355, 2004.
- Gobiet, A., Foelsche, U., Steiner, A.K., Borsche, M., Kirchengast, G., and Wickert J.: Climatological validation of stratospheric temperatures in ECMWF operational analyses with CHAMP radio occultation data, *Geophys. Res. Lett.*, 32, L12806, doi:10.1029/2005GL022617, 2005.
- Gosset, W. S. (pen name: Student): The probable error of a mean, *Biometrika* 6 (1908) pp. 1–25, reprinted on pp. 11–34 in “Students” Collected Papers, edited by: Pearson, E. S. and Wishart, J., Cambridge University Press for the Biometrika Trustees, 1942.
- Hauchecorne, A. and Chanin, M.-L.: Density and temperature profiles obtained by lidar between 35 and 70 km, *Geophys. Res. Lett.*, 7, 565–568, 1980.
- Hedin, A. E.: Extension of the MSIS Thermospheric Model into the Middle and Lower Atmosphere, *J. Geophys. Res.*, 96, 1159–1172, 1991.
- Höpfner, M., Oelhaf, H., Wetzell, G., Friedl-Vallon, F., Kleinert, A., Lengel, A., Maucher, G., Nordmeyer, H., Glatthor, N., Stiller, G., von Clarmann, T., Fischer, H., Kröger, C., and Deshler, T.: Evidence of scattering of tropospheric radiation by PSCs in mid-IR limb emission spectra: MIPAS-B observations and KOPRA simulations, *Geophys. Res. Lett.*, 29(8), 1278, doi:10.1029/2001GL014443, 2002.
- Keckhut, P., Chanin, M.-L., and Hauchecorne, A.: Stratosphere temperature measurement using Raman lidar, *Appl. Opt.*, 29, 5182–5186, 1990.
- Keckhut, P., Hauchecorne, A., and Chanin, M., L.: A critical review on the data base acquired for the long term surveillance of the middle atmosphere by french Rayleigh lidars, *J. Atmos. Oceanic Technol.*, 10, 850–867, 1993.
- Keckhut, P., McDermid, I.S., Swart, D.P.J., McGee, T., Godin-Beekmann, T., Adriani, A., Barnes, J., Baray, J.-L., Bencherif, H., Claude, H., Fiocco, G., Hansen, G., Hauchecorne, A., Leblanc, T., Lee, C.H., Pal, S., Megie, G., Nakane, H., Neuber, R., Steinbrecht, W., and Thayer, J.: Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change, *J. Environ. Monitor.*, 6, 721–733, 2004.
- Kiefer, M., von Clarmann, T., Grabowski, U., De Laurentis, M., Mantovani, Milz, M., and Ridolfi, M.: Characterization of MIPAS elevation pointing, *Atmos. Chem. Phys.*, 7, 1615–1628, 2007, <http://www.atmos-chem-phys.net/7/1615/2007/>.

- 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/>.
- Kurylo, M. J. and Zander, R. J.: The NDSC – Its status after ten years of operation, in: *Proceedings of the Quadrennial Ozone Symposium 2000*, Hokkaido Univ., Sapporo, Japan, edited by: NASDA, 167–168, 2001.
- Leblanc, T., McDerimid, I. S., Hauchecorne, A., and Keckhut, P.: Evaluation of optimization of lidar temperature analysis algorithms using simulated data, *J. Geophys. Res.*, 103, 6177–6187, 1998.
- Luers, J. K.: Temperature Error of the Vaisala RS90 Radiosonde, *J. Atmos. Oceanic Technol.*, 14(6), 1520–1532, 1997.
- Moreau, G., Robert, C., Catoire, V., Chartier, M., Camy-Peillet, C., Huret, N., Pirre, M., Pomathiod, L., and Chalumeau, G.: SPIRALE: A multispecies in situ balloon-borne instrument with six tunable diode laser spectrometers, *Appl. Opt.*, 44(28), 5972–5989, 2005.
- Oelhaf, H., von Clarmann, T., Fergg, F., Fischer, H., Friedl-Vallon, F., Fritzsche, C., Piesch, C., Rabus, D., Seefeldner, M., and Volker, W.: Remote sensing of trace gases with a balloon borne version of the Michelson interferometer for passive atmospheric sounding MIPAS, in *Proceedings of the Tenth ESA Symposium on European Rocket and balloon Programmes and Related Research*, ESA-SP-317 European Space Agency, Paris, 207–213, 1991.
- Pappalardo, G., Amodeo, A., Boselli, A., Cornacchia, C., D'Amico, G., Madonna, F., Mona, L., and Pandolfi, M.: IMAA-CNR lidar systems for aerosol, clouds and water vapor study: Lidar Technologies, Techniques, and Measurements for Atmospheric Remote Sensing, *Proceedings of SPIE*, Vol. 5984, No. 12, Bruges, Belgium, 2005.
- Raspollini, P. and Ridolfi, M.: Mapping of temperature and line-of-sight errors in constituent retrievals for MIPAS/ENVISAT measurements, *Atmos. Environ.*, 34, 5329–5336, 2000.
- Raspollini, P., Belotti, C., Burgess, A., Carli, B., Carlotti, M., Ceccherini, S., Dinelli, B. M., Dudhia, A., Flaud, J.-M., Funke, B., Hoepfner, M., Lopez-Puertas, M., Payne, V., Piccolo, C., Remedios, J. J., Ridolfi, M., Spang, R.: MIPAS Level 2 operational analysis, *Atmos. Chem. Phys.*, 6, 5605–5630, 2006, <http://www.atmos-chem-phys.net/6/5605/2006/>.
- Ridolfi, M., Carli, B., Carlotti, M., von Clarmann, T., Dinelli, B. M., Dudhia, A., Flaud, J.-M., Hoepfner, 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, *Appl. Opt.*, 39(8), 1323–1340, 2000.
- Ridolfi, M., Ceccherini, S., and Carli, B.: Optimal interpolation method for intercomparison of atmospheric measurements, *Opt. Lett.*, 31, 855–857, 2006.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, Vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics (World Scientific, Singapore), 2000.
- Rodgers, C. D. and Connor, B., J.: Intercomparison of remote sounding instruments, *J. Geophys. Res.*, 108, 4116–4130, 2003.
- Salby, M. L.: Survey of planetary scale travelling waves, the state of theory and observations, *Rev. Geophys.* 22, 209–236, 1984.
- Simmons, A.: Representation of the stratosphere in ECMWF operations and ERA-40, Workshop on modelling and assimilation for the stratosphere and troposphere, Reading, 23–26 June, 2003.
- Simmons, A., Hortal, M., Kelly, G., McNally, A., Untch, A., and Uppala, S.: ECMWF Analyses and Forecasts of Stratospheric Winter Polar Vortex Breakup: September 2002 in the Southern Hemisphere and Related Events, *J. Atmos. Sci.*, 62(3), 668–689, 2005.
- Thompson, A., M., Witte, J., C., McPeters, R., D., Oltmans, S., J., Schmidlin, F., J., Logan, J., A., Fujiwara, M., Kirchhoff, V. W. J. H., Posny, F., Coetzee, G. J. R., Hoegger, B., Kawakami, S., Ogawa, T., Johnson, B., J., Vömel, H., and Labow, G.: Southern Hemisphere Additional Ozoneondes (SHADOZ) 1998–2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements, *J. Geophys. Res.*, 108(D2), 8238, doi:10.1029/2001JD000967, 2003.
- von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: Strategies and Terminology, *Atmos. Chem. Phys.*, 6, 4311–4320, 2006, <http://www.atmos-chem-phys.net/6/4311/2006/>.
- Wetzel, G., Oelhaf, H., Ruhnke, R., Friedl-Vallon, F., Kleinert, A., Kouker, W., Maucher, G., Reddmann, Th., Seefeldner, M., Stowasser, M., Trieschmann, O., von Clarmann, T., and Fischer, H.: NO<sub>y</sub> partitioning and budget and its correlation with N<sub>2</sub>O in the Arctic vortex and in summer mid-latitudes in 1997, *J. Geophys. Res.*, 107(D16), 4280, doi:10.1029/2001JD000916, 2002.
- Wetzel, G., Oelhaf, H., Friedl-Vallon, F., Kleinert, A., Lengel, A., Maucher, G., Nordmeyer, H., Ruhnke, R., Nakajima, H., Sasano, Y., Sugita, T., and Yokota T.: Intercomparison and validation of ILAS-II version 1.4 target parameters with MIPAS-B measurements, *J. Geophys. Res.*, 111, D11S06, doi:10.1029/2005JD006287, 2006.
- WMO: CIMO, Presentation on recent national tests/comparisons, Joint meeting Geneva, 17–20 March 2004, document CIMO/OPAG-Upper-AIR/ET-UASI-1/IOC-1/Doc, 3.2(5), 2004.