



# Long-term validation of ESA operational retrieval (version 6.0) of MIPAS Envisat vertical profiles of methane, nitrous oxide, CFC11, and CFC12 using balloon-borne observations and trajectory matching

Andreas Engel<sup>1</sup>, Harald Bönisch<sup>1,a</sup>, Tim Schwarzenberger<sup>1,b</sup>, Hans-Peter Haase<sup>1,c</sup>, Katja Grunow<sup>2,d</sup>, Jana Abalichin<sup>2</sup>, and Stephan Sala<sup>1,e</sup>

<sup>1</sup>Institute for Atmospheric and Environmental Sciences, Goethe Universität, Frankfurt, Germany

<sup>2</sup>Institut für Meteorologie, Freie Universität Berlin, Berlin, Germany

<sup>a</sup>now at: Institut für Meteorologie und Klimaforschung, KIT, Karlsruhe, Germany

<sup>b</sup>now at: Technologiezentrum Wasser, Karlsruhe, Germany

<sup>c</sup>now at: Deka Bank, Frankfurt, Germany

<sup>d</sup>now at: Senatsverwaltung für Stadtentwicklung und Umwelt, Berlin, Germany

<sup>e</sup>now at: Deutscher Wetterdienst, DWD, Offenbach, Germany

Correspondence to: Andreas Engel (an.engel@iau.uni-frankfurt.de)

Received: 18 May 2015 – Published in Atmos. Meas. Tech. Discuss.: 21 July 2015

Revised: 26 January 2016 – Accepted: 27 January 2016 – Published: 14 March 2016

**Abstract.** MIPAS-Envisat is a satellite-borne sensor which measured vertical profiles of a wide range of trace gases from 2002 to 2012 using IR emission spectroscopy. We present geophysical validation of the MIPAS-Envisat operational retrieval (version 6.0) of N<sub>2</sub>O, CH<sub>4</sub>, CFC-12, and CFC-11 by the European Space Agency (ESA). The geophysical validation data are derived from measurements of samples collected by a cryogenic whole air sampler flown to altitudes of up to 34 km by means of large scientific balloons. In order to increase the number of coincidences between the satellite and the balloon observations, we applied a trajectory matching technique. The results are presented for different time periods due to a change in the spectroscopic resolution of MIPAS in early 2005. Retrieval results for N<sub>2</sub>O, CH<sub>4</sub>, and CFC-12 show partly good agreement for some altitude regions, which differs for the periods with different spectroscopic resolution. The more recent low spectroscopic resolution data above 20 km altitude show agreement with the combined uncertainties, while there is a tendency of the earlier high spectral resolution data set to underestimate these species above 25 km. The earlier high spectral resolution data show a significant overestimation of the mixing ratios for N<sub>2</sub>O, CH<sub>4</sub>, and CFC-12 below 20 km. These differences need to be con-

sidered when using these data. The CFC-11 results from the operation retrieval version 6.0 cannot be recommended for scientific studies due to a systematic overestimation of the CFC-11 mixing ratios at all altitudes.

## 1 Introduction

Measurements of long-lived trace gases in the stratosphere serve a variety of purposes. They can be used to determine relative lifetimes (Plumb and Ko, 1992; Volk et al., 1997; Engel et al., 2013), to observe dynamical processes (Volk et al., 1996; Bauer et al., 1994; Plumb et al., 2002), to study transport timescales, the so-called mean age (Hall and Plumb, 1994; Engel et al., 2009), and also to derive chemical budgets (Engel et al., 1997; Laube et al., 2008). Such studies using tracers require high precision and accuracy in order to derive quantitative information. In situ observations in the stratosphere, which can be conducted from high flying aircrafts up to altitudes of 20 km and from large scientific balloons for altitudes up to 40 km, often provide information on tracers of interest at high spatial resolution and high ac-

curacy. The information obtained is however very sparse in space and time. On the other hand, satellite-borne sensors often provide very good temporal and spatial coverage, yet the vertical resolution is limited and the accuracy and precision of satellite observations are critical factors, which need to be evaluated carefully. A comparison of the high-accuracy in situ observations with satellite data retrievals is necessary to estimate the quality of satellite data and to identify possible biases.

As mentioned above, the density of satellite observations is often much higher than that of the respective in situ data to be used for the validation. During the acquisition of the in situ validation data, care should be taken to achieve a good match in time and space to the satellite observations. However, this typically can only be achieved for a single vertical profile from a given satellite data set. It is however desirable to use the valuable in situ data for validation of more than only one satellite profile. For this purpose either climatological approaches or tracer–tracer correlations can be used. Another approach is the use of backward and forward air mass trajectories to find more incidences during which the satellite measured the same air parcel as the in situ observations.

Here we present work on validation of long-lived trace gases  $\text{N}_2\text{O}$ ,  $\text{CH}_4$ , CFC-12, and CFC-11 derived from observations of the MIPAS (Michelson Interferometer for Passive Atmospheric Sounding) Fourier transform infrared (FTIR) spectrometer on board the Envisat satellite. Previous results on the validation of different data products for these gases have been published (Payan et al., 2009; Hoffmann et al., 2008; Kellmann et al., 2012; Vigouroux et al., 2007), and the data have been used, for example, to derive stratospheric lifetimes (Engel et al., 2013; Minschwaner et al., 2013) and climatologies and trends of CFC-11 and CFC-12 (Kellmann et al., 2012). While MIPAS is not operating anymore, it has provided a wealth of data that span nearly 10 years. A thorough validation of these data is the prerequisite for their scientific application. Here we present a comparison of measurements of long-lived trace gases obtained using a cryogenic whole air sampler and the MIPAS data set from the operational ESA retrieval version ML2PP/6.0. The validation of other retrieval algorithms of the MIPAS data for  $\text{CH}_4$  and CFCs 11 and 12 is subject to further studies (Laeng et al., 2015; Eckert et al., 2015). In order to enhance the number of coincidences between the in situ observations and the satellite data set, we apply trajectory mapping techniques. This technique and the different data sets are described in Sect. 2. In Sect. 3, we present the results of the comparison for the four trace gases  $\text{N}_2\text{O}$ ,  $\text{CH}_4$ , CFC-12, and CFC-11, which are then summarised and discussed in Sect. 4.

## 2 Data sets

### 2.1 MIPAS-E

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) is an emission spectrometer measuring atmospheric trace gases in the mid-infrared region using Fourier transform spectroscopy (Fischer et al., 2008). MIPAS was operated on different platforms. The satellite version on board the Envisat satellite operated by the European Space Agency (ESA) provided data for a large number of atmospheric trace gases from July 2002 until the end of operation in April 2012. The limb scanning geometry allows for both high vertical resolution of up to 3 km and good global coverage with about 1000 vertical profiles per day. Vertical profiles in so-called nominal mode are retrieved for about 16 tangent altitudes in a range between 8 and 68 km. The data are processed by ESA in a near-operational way for a limited number of trace gases. In addition the data are evaluated for a number of scientific data products (e.g. Stiller et al., 2008; Kellmann et al., 2012). The original spectral resolution was  $0.035\text{ cm}^{-1}$ , which had to be reduced to  $0.0625\text{ cm}^{-1}$  due to problems with the interferometer drive unit. The data set starting in January 2005 is based on these lower spectral resolution data. Due to this change in spectral resolution the validation results are presented separately for the time period prior to 2005 (high spectral resolution data in the following) and the data after January 2005 (low spectral resolution data in the following).

### 2.2 The cryogenic whole air sampler BONBON

The cryogenic whole air sampler (cryosampler) BONBON is an instrument designed for collection of large whole air samples during the flight of large stratospheric balloons. The instrument carries 15 internally electropolished stainless steel canisters which can be opened and closed at different altitudes. The sample canisters are cooled to 27 K using liquid neon. At this temperature the ambient air will condense on the inner surfaces of the sample canisters allowing the collection of very large air samples corresponding to up to 15 L at standard temperature and pressure, during rather short times. For example at 30 km the time to collect such a large sample is typically below 5 min. The samples are typically collected during the valve controlled slow descent of the balloon at vertical velocities of  $2\text{--}3\text{ m s}^{-1}$ . A 5 min sampling interval will thus result in 750 m vertical resolution. Depending on the external pressure, sampling time, and vertical velocities, the vertical resolution is usually between 100 m and 1 km. More details on the technique of cryogenic whole air sampling can be found in Lueb et al. (1975) and on the cryosampler used for this study in Engel et al. (1997).

The samples are subsequently analysed in the laboratory using a variety of gas chromatographic methods and reported on internationally accepted calibration scales. In this paper,

**Table 1.** Abbreviation, number of air samples, altitude range, location, and flight date of the cryosampler flights considered for the validation. The abbreviation BXX refers to flight no. XX of the balloon-borne whole air samplers. University of Frankfurt operates two nearly identical cryogenic whole air samplers allowing the performance of two flights during a campaign.

Abbreviation	Number of air samples	Altitude range	Location	Flight date
B39	14	12.2–32.4 km	44° N, Aire-sur-l'Adour, France	24 Sep 2002
B40	13	11.2–30.4 km	68° N, Kiruna, Sweden	6 Mar 2003
B41	10	13.3–28.9 km	68° N, Kiruna, Sweden	9 Jun 2003
B42	11	15.2–34.0 km	5° S, Teresina, Brazil	8 Jun 2005
B43	14	15.2–33.6 km	5° S, Teresina, Brazil	25 Jun 2005
B45	11	10.3–31.8 km	68° N, Kiruna, Sweden	10 Mar 2009
B46	8	12.5–23.3 km	68° N, Kiruna, Sweden	1 Apr 2011

data for the trace gases  $\text{N}_2\text{O}$ ,  $\text{CH}_4$ , CFC-12, and CFC-11 are compared.  $\text{N}_2\text{O}$  and  $\text{CH}_4$  were measured at the University of Heidelberg by gas chromatography with electron capture (Schmidt et al., 2001) and flame ionisation (Levin et al., 1999) detection respectively.  $\text{N}_2\text{O}$  was in addition also measured at University of Frankfurt. The data used for the intercomparison are those from University of Frankfurt except for flight B46 where  $\text{N}_2\text{O}$  data from University of Heidelberg are used. The chlorofluorocarbons CFC-11 and CFC-12 were measured at University of Frankfurt by gas chromatography coupled to mass spectrometry (Laube et al., 2008; Hoker et al., 2015), with the exception of the measurements prior to 2004, which were made using electron capture detectors (Engel et al., 1997).  $\text{N}_2\text{O}$  is reported on the NOAA 2006 scale for all observations (Hall et al., 2007),  $\text{CH}_4$  on the NOAA 2004 (Dlugokencky et al., 2005) scale and the CFC-11 and CFC-12 are reported relative to the NOAA 1993 and 2001 scales (Montzka et al., 2003), respectively. Typical reproducibilities for these four substances are 0.5–1 % (CFC-11), 0.5 % (CFC-12), 0.2 % ( $\text{CH}_4$ ), and 0.5 % ( $\text{N}_2\text{O}$ ). Absolute accuracies of the calibration scales are 1 % or less for all species. As these uncertainties are much smaller than those of the satellite sensors, they are neglected in the discussion of the validation results.

### 2.3 Available validation flights

A total of seven flights of our cryogenic whole air sampler were carried out during the time that MIPAS-E was operational. Not all flights were explicitly for the purpose of validation resulting in different quality of the direct coincidence with MIPAS-E observations. The flights took place at different latitudes ranging from the tropics (5° S) to the high northern latitudes (68° N) under different meteorological conditions and during different seasons. The dates, launch locations, latitudes, and altitude ranges covered by the flights are shown in Table 1.

In the following we will give some very brief information about the individual flights. Flight B39 was launched from southern France in autumn 2002 and represents typical undisturbed conditions during the turnaround from sum-

mer to winter circulation at the midlatitudes of the Northern Hemisphere. Flight B40 was launched from the high-latitude site of ESRANGE near Kiruna in northern Sweden during the polar winter. Some very low mixing ratios were observed inside the polar vortex due to the well-known effect of diabatic descent inside the vortex. This flight is discussed in detail in Engel et al. (2006), in particular with respect to the intrusion of mesospheric air into the stratosphere. A further flight (B41) was launched to probe the high-latitude stratosphere during summer. Flight B42 and B43 were launched from Teresina in Brazil (5° S) to probe the stratosphere in the inner tropics. The vertical profiles of many trace gases show much smaller decrease with altitude in the lower tropical stratosphere and then steep gradients with altitude inside the loss regions for the specific substances. These two flights are discussed in more detail in Laube et al. (2008). The two most recent flights are again from high northern latitudes during winter to early spring and were launched on 9 March 2009 (B45) and 1 April 2011 (B46) respectively. The flight B45 mainly sampled air from outside the polar vortex, with the exception of a tongue of vortex air sampled between about 20 and 25 km altitude. The flight B46 was launched very close to the edge of the polar vortex, which was unusually strong and showed very high wind speeds at the edge during this year (e.g. Manney et al., 2011). Due to these unusually strong horizontal winds only a very short flight could be carried out and the maximum altitude reached was only 23 km. The observed mixing ratios of the long-lived tracers are typical of conditions at the inner edge of the polar vortex with rather low mixing ratios close to those observed well inside the vortex.

### 2.4 Trajectory matching and data handling

As mentioned above the amount of satellite data which can be compared by a direct match is very limited and usually consists of only one or two scans of the satellite instrument. In order to enhance the number of satellite vertical profiles which can be used for the intercomparison, we have applied the technique of trajectory mapping using the University of Berlin trajectory model (Naujokat and Grunow, 2003;

Grunow, 2009). Trajectories are initiated about every 200 m in altitude (for flights prior to 2009) or every 10 s, i.e. about every 20–50 m (for flight in 2009 and 2011) along the flight path of the balloon and run forward and backward in time for 5 days (Grunow, 2009). MIPAS-E data are interpolated to the respective potential temperature of the trajectory at the point of the match. All MIPAS-E observations which are within a 500 km radius and a 1 h time mismatch with the trajectory are regarded as having sampled the same air mass. In addition only data have been used for validation where matches were obtained for at least four trajectories covering an altitude of at least 1.5 km. In this way we ensured that a match cannot be obtained from a single trajectory as single trajectories may show significant uncertainties (Schwarzenberger, 2014). These values were found to give the best compromise between finding a sufficient amount of matches and ensuring that differences in air mass were sufficiently small.

The trajectory calculations were performed on potential temperature surfaces, and vertical displacement in potential temperature space was taken into account using climatological heating rates (Grunow, 2009). The choice of using climatological heating rates or heating rates calculated specifically for the trajectories was found to give no significant differences, as tested for the validation of  $\text{N}_2\text{O}$  and  $\text{CH}_4$  (Grunow, 2009). When using isentropic trajectories vertical differences to trajectories using explicitly calculated heating rates were on the average 120 m in the tropics and 170 m in high latitudes during winter (Grunow, 2009). As trajectories are not calculated for more than 5 days in general, this only leads to very small deviations in the matches from trajectories using explicitly calculated heating rates (Grunow, 2009). For CFC-11, which shows the strongest vertical gradient of all gases investigated here, the associated uncertainty is about 3 ppt, which is about 1.5 % of the current tropospheric value. For clarity the results are presented on an altitude scale, which is that of the balloon profile along which the trajectories were initiated. This altitude is based on GPS measurements. For the flights prior to 2009 operational ECMWF data on  $2.5^\circ \times 2.5^\circ$  grid were used, while the calculations for the flights in 2009 and 2011 are based on  $1.25^\circ \times 1.25^\circ$  operational ECMWF data. The trajectory model is run using 25 vertical levels for the  $2.5^\circ \times 2.5^\circ$  grid data and 59 levels for the  $1.25^\circ \times 1.25^\circ$  grid data.

For a comparison with the balloon observations the matched satellite data points were averaged over 1 km intervals. For the quantitative intercomparison the balloon observations are then interpolated to the mean potential temperature of the matched satellite data. Again, for clarity the results are reported on an altitude scale based on the balloon trajectory, while the calculation of the matches is based on potential temperature.

The two data sets have different vertical resolution. The vertical resolution of the satellite is given by the averaging kernels. In the case of the cryosampler, the vertical resolution varies from sample to sample based on the vertical speed at

which the balloon was travelling and the sample integration time. As no information is available for the cryosampler measurements in between the samples (discontinuous sampling), a degradation of the vertical resolution of the cryosampler to those of MIPAS-Envisat was not possible. This paper thus only presents a validation of the geophysical results of the satellite retrieval.

### 3 Validation results

In the following we present the validation results for the four species mentioned above. Results are only shown for the version 6.0 of the operational ESA data retrieval. Measurements of the four trace gases are available for all samples collected with the cryogenic whole air sampler. We present the results for all species separately. The direct comparison of the vertical profile gives a qualitative impression of the quality of the intercomparison. For each species we present the direct comparison on an altitude vertical scale from flight B45 launched on 10 March 2009 from Esrange near Kiruna, as this flight covers a large altitude range and a wide range of mixing ratios. It also shows a sharp structure which is in particular challenging for satellite validation and shows the limitations of the method applied here.

For a more quantitative comparison the information from several flights is combined. For this purpose the in situ measurements are interpolated in altitude to give a value at the same altitude as the mean altitude of the MIPAS-E data in the respective bin. Results are separated for the period up to 2005 (high spectral resolution) and for the period with lower spectral resolution of MIPAS-E starting in January 2005. The flights were also conducted in different latitude regions. However, we found the variations with time (low spectral resolution vs. high spectral resolution) to be the dominant factor of variability. As the total number of flights was not sufficient to derive systematic differences also for different latitudes (and seasons), we have chosen to restrict the comparison to differentiate between different spectral resolution. We have chosen 1 km intervals for the binning of MIPAS-E data. As noted above the uncertainty in the in situ data is sufficiently small to be neglected in comparison to the uncertainty of the satellite data. As the number of satellite data points is averaged in each bin, there are two sources of uncertainty which have to be considered for this intercomparison. One is the uncertainty of an individual measurement point as specified from the retrieval algorithm, and the other is the variability of the individual data points in the bin around the mean value. For the direct comparison the error given for the MIPAS-E data is the standard deviation of the data points in the respective bin; it is thus the variability of the retrieved values. In case there are no systematic deviations it is expected that the MIPAS-E data should scatter around the in situ data. For the quantitative comparison two error bars are given. The ones in red are the same as those given in the direct comparison, while the blue lines show the mean of the uncertainty quoted

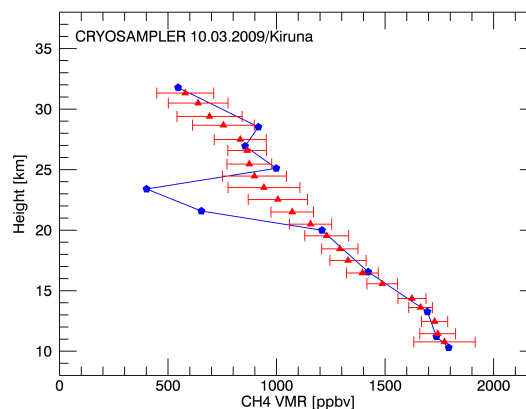
in the data files. This representation will thus show whether the uncertainty given in the data files is representative of the mean deviation between satellite data and in situ measurements.

### 3.1 MIPAS-E CH<sub>4</sub>

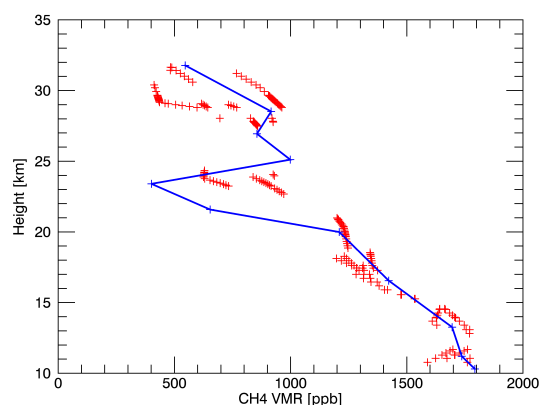
In Fig. 1 a direct comparison of the methane profile obtained from the cryosampler flight on 10 March 2009 (Flight B45) with the matched MIPAS-E values based on the version 6.0 ESA retrieval is shown. Below 20 km an excellent agreement is observed which is closer than the stated uncertainties. Above 25 km for the validation based on B45, good agreement of the MIPAS-E retrieval and balloon data is observed, which agrees within the scatter of the MIPAS-E data. The sharp structure observed between 20 and 25 km cannot be reproduced by the satellite observations. There is however an enhanced variability of the MIPAS-E values where the structure is observed in the in situ data. Some of the MIPAS-E data points actually show values which are very similar to the low values observed between 20 and 25 km in the in situ data. As this structure is caused by the presence of a filament of vortex air, it is possible that this structure is not captured or not reproduced by the trajectory matching technique. This rather sharp feature in flight B45 is a good example of the restriction of applying the match method to satellite validation, especially for a limited amount of validation data. In order to demonstrate this, we show the direct validation results (interpolated MIPAS data) for all trajectories which fulfill more stringent validation criteria (maximum difference 200 km and maximum trajectory travel time of 2 days). Figure 2 shows that in this case the meteorological feature is indicated also in the satellite data, but the resulting number of matches is rather low; therefore, a quantitative validation is not possible for these more stringent match criteria. Hence, the validation results for all species will also be given disregarding the altitude region between 20 and 25 km during flight B45.

Figure 3 shows the quantitative comparison for all cryosampler measurements available for this study: the mean differences (red lines in Fig. 3) and their scatter are compared with the mean error given from MIPAS-E database (blue line). Especially between 15 and 25 km the mixing ratio values of MIPAS-E are clearly overestimated. Above 25 km there is a good agreement according to the MIPAS-E error. In order to show the number of data which are available for the intercomparison, the right-hand side of the quantitative intercomparison plots shows the number of data available per bin and the number of independent scans for which a sufficiently good match has been obtained.

As noted before the spectral resolution of MIPAS-E was degraded after January 2005. Therefore the absolute differences are compared for the period before and after January 2005 separately. Figure 4 displays the mean deviation between MIPAS-E profiles and in situ profiles for both time

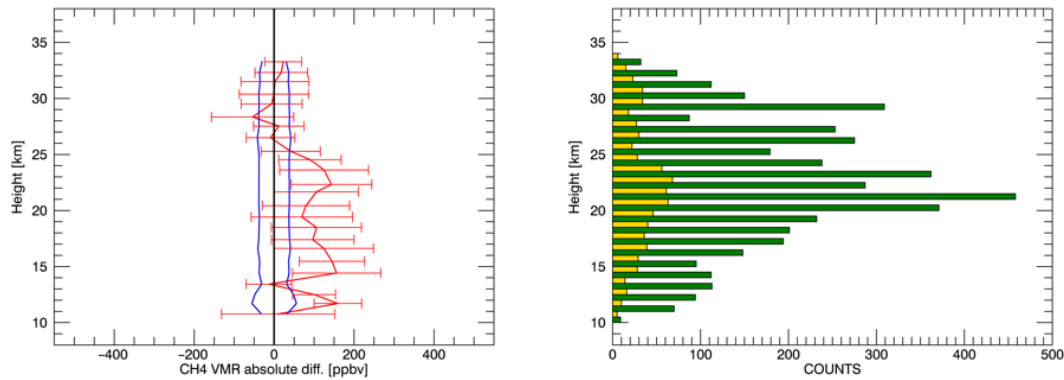


**Figure 1.** Comparison of CH<sub>4</sub> profile retrieved from B45 (1 April 2011, from Kiruna Sweden, 68° N) and the corresponding mean MIPAS E profile based on the trajectory matching results. Blue dots: cryosampler measurements. Blue line: interpolated CH<sub>4</sub> profile from cryosampler (BONBON). Red triangle: mean CH<sub>4</sub> mixing ratio values of MIPAS-E. Red error bars: standard deviation of matched Envisat data.

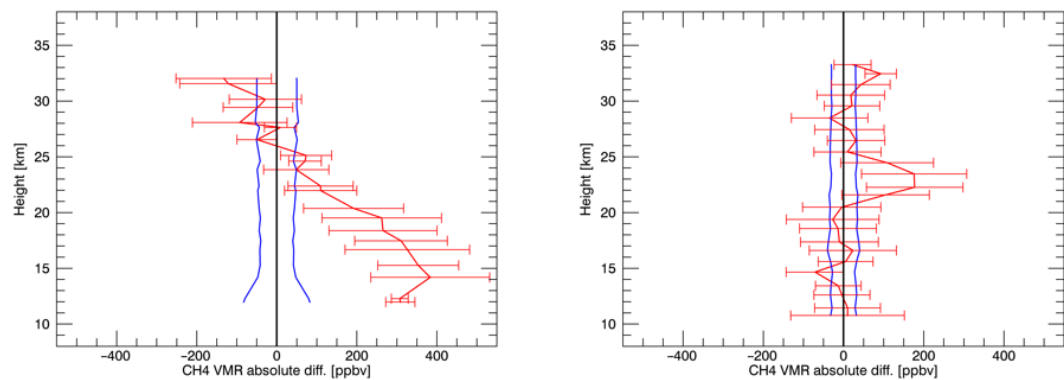


**Figure 2.** Comparison of the vertical profile obtained during flight B45 with the closest satellite measurements interpolated to the potential temperature of the trajectories at the point of the match. For this comparison only satellite data have been used which match the trajectory within 200 km and a maximum time delay of 2 days. While the low values observed between 20 and 25 km cannot be completely reproduced, these closest matches indicate a similar feature as in the balloon data (see text for discussion).

periods separately: the left panel presents the conclusions for earlier flights until 2003 (containing B39, B40, and B41) and the right panel for flights since 2005 (containing, B42, B43, B45, and B46). It is remarkable that the earlier high spectral resolution validation flights obviously show significantly worse correlations than the more recent low spectral resolution ones, especially below  $\sim 22$  km with differences up to 400 ppb (more than 20%). The deviations in this part of the profile cannot be explained by the measurement accuracy of MIPAS-E. The validation based on the more recent low spectral resolution validation flights shows agreement between



**Figure 3.** Left panel: CH<sub>4</sub> validation based on all cryosampler (BONBON) flights considered in this study. Red line: interpolated mean difference between MIPAS-E and cryosampler mixing ratio values. Red error bars: mean standard deviation of the difference between MIPAS-E and cryosampler mixing ratio value interpolated to the respective altitude. Blue line: interpolated mean error for CH<sub>4</sub> mixing ratio values from MIPAS-E. Right panel: corresponding database of MIPAS-E. Green bars: total no. of matches. Yellow bars: no. of independent scans which are matched.

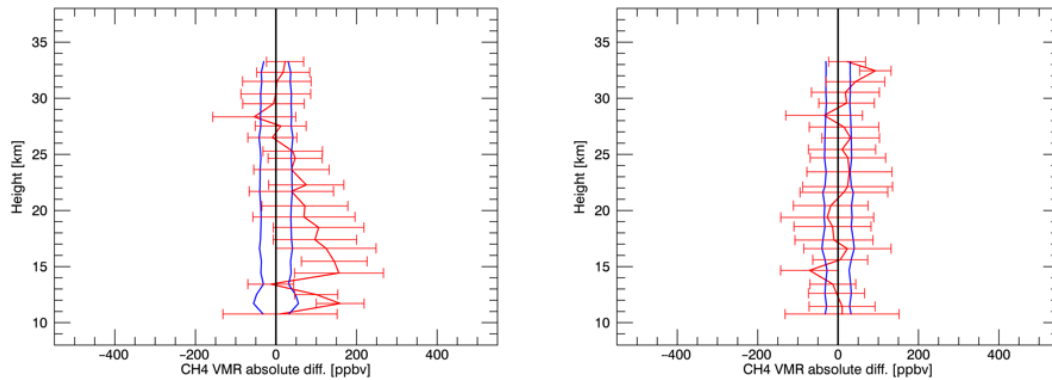


**Figure 4.** Left panel: CH<sub>4</sub> validation based on cryosampler (BONBON) flights B39, B40, and B41 (high spectral resolution). Red line: interpolated mean difference between MIPAS-E and cryosampler mixing ratios. Red error bars: mean standard deviation of the difference between MIPAS-E and cryosampler mixing ratios for the respective altitude range. Blue line: interpolated mean error for CH<sub>4</sub> mixing ratio values from MIPAS-E database. Right panel: CH<sub>4</sub> validation based on cryosampler flights B42, B43, B45, and B46 (low spectral resolution). Red line: interpolated mean difference between MIPAS-E and cryosampler mixing ratios. Red error bars: mean standard deviation of the difference between MIPAS-E and cryosampler mixing ratios for the respective altitude range. Blue line: interpolated mean error for CH<sub>4</sub> mixing ratios from MIPAS-E database.

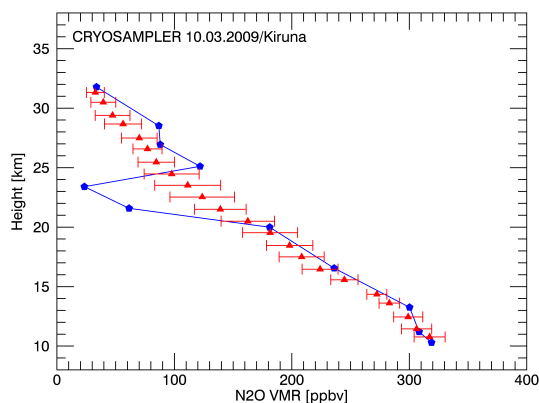
MIPAS-E and balloon measurements within the expected uncertainties. An exception is the altitude range of  $\sim 21\text{--}25$  km. This is mainly caused by the sharp structure observed during Flight B45 presented in Fig. 1. When disregarding the values between 20 and 25 km altitude from flight B45, the mean difference between MIPAS-E profiles and the in situ profiles becomes smoother (Fig. 5). In this case the deviation based on the CH<sub>4</sub> validation with the low spectral resolution validation flights can be explained almost for the full profile range by the quoted uncertainty in the MIPAS-E retrieval (Fig. 5, right panel).

### 3.2 Validation of MIPAS-E N<sub>2</sub>O

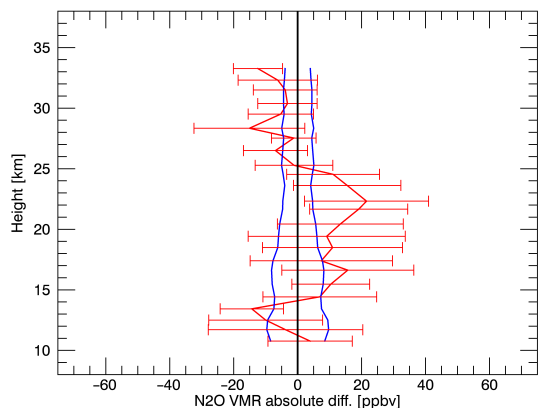
The N<sub>2</sub>O profile of flight B45 from the cryogenic whole air sampler is displayed in Fig. 6. It shows the same significant structure as observed for CH<sub>4</sub>. This leads to N<sub>2</sub>O mixing ratio values clearly overestimated by MIPAS-E in an altitude range between 20 and 25 km. Below 20 km the two N<sub>2</sub>O profiles fit together almost perfect. Above 25 km there is a tendency for MIPAS-E to underestimate the in situ data. As in the case of CH<sub>4</sub> enhanced variability is observed in the matched MIPAS-E data. Again, as in the case of CH<sub>4</sub>, the region with the strong vertical structure shows enhance variability in the MIPAS-E observations with some observations showing a similar structure as the in situ data. The discrepancy is thus most probably not representative of MIPAS-E



**Figure 5.** Deviations between MIPAS and cryosampler data after elimination of meteorological feature in profile B45 between 20 and 25 km altitude. Left panel: CH<sub>4</sub> validation based on all cryosampler (BONBON) flights considered in this study. Red line: interpolated mean difference between MIPAS-E and cryosampler mixing ratio values. Red error bars: mean standard deviation of the difference between MIPAS-E and cryosampler mixing ratio value for the respective altitude range. Blue line: interpolated mean error for CH<sub>4</sub> mixing ratio values from MIPAS-E database. Right panel: CH<sub>4</sub> validation based on cryosampler flights B42, B43, B45, and B46. Red line: interpolated mean difference between MIPAS-E and cryosampler mixing ratio values. Red error bars: mean standard deviation of the difference between MIPAS-E and cryosampler mixing ratio value for the respective altitude range. Blue line: interpolated mean error for CH<sub>4</sub> mixing ratio values from MIPAS-E database.



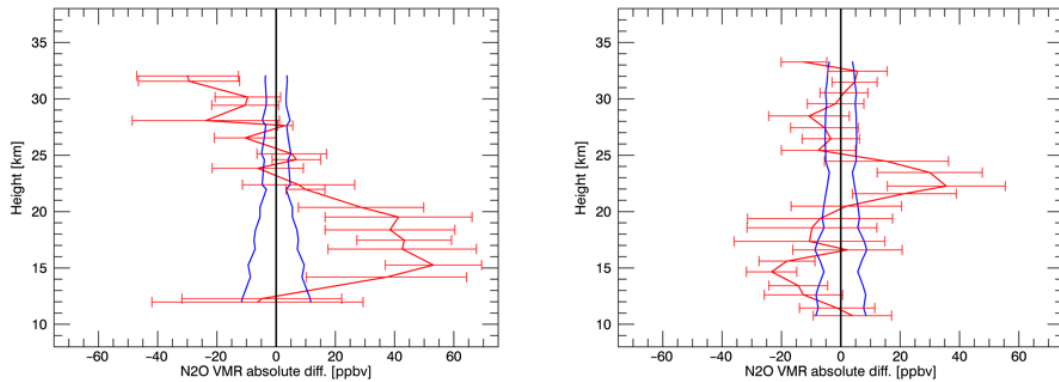
**Figure 6.** Same as Fig. 1, but instead of CH<sub>4</sub> now N<sub>2</sub>O.



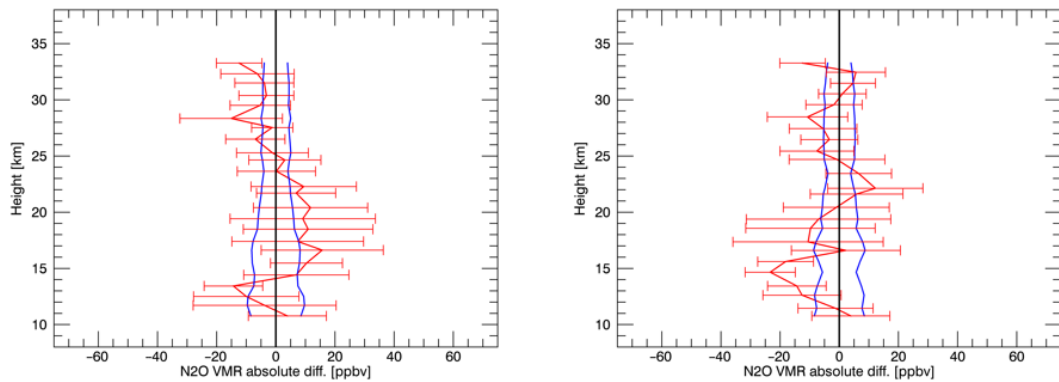
**Figure 7.** Same as left panel of Fig. 3, but instead of CH<sub>4</sub> now N<sub>2</sub>O.

observations in this region. The quantitative N<sub>2</sub>O validation based on all flights is displayed in Fig. 7. The most significant mean differences are located at an altitude of 20–25 km and are mainly caused by the sharp structure in the profile of B45. Above and below this altitude range the agreement is much better, so differences between MIPAS-E and in situ observations can be explained by the stated uncertainty of the MIPAS-E values.

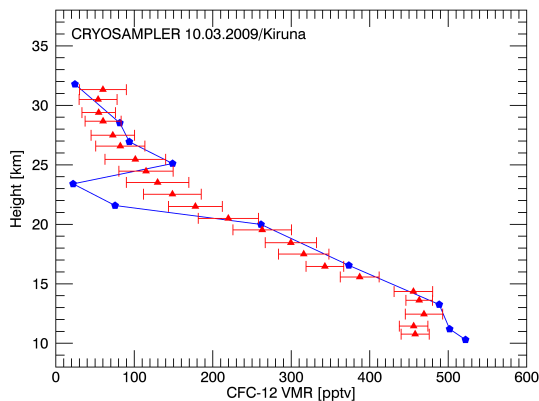
Figure 8 again distinguishes between high spectral resolution (flights B39, B40, and B41) and low spectral resolution validation measurements (flights B42, B43, B45, and B46) of MIPAS-E. For the validation based on the high spectral resolution validation flights between ~13 and 21 km, the deviations between N<sub>2</sub>O mixing ratio values measured by satellite and balloon respectively are significantly higher than the stated MIPAS-E N<sub>2</sub>O error, with MIPAS-E values consistently higher than the in situ data. The more recent low spectral resolution data tend to overestimate the mixing ratios below 20 km. The overall reasonable agreement below 20 km is thus in case of N<sub>2</sub>O in part due to an overestimation of N<sub>2</sub>O mixing ratios in the high spectral resolution and an underestimation in the low spectral resolution data. In addition, MIPAS-E consistently underestimates the mixing ratio values at altitudes above ~27 km for the high spectral resolution data set. In comparison the validation based on the low spectral resolution flights with lower spectral resolution generally shows agreement within the stated uncertainties (especially above 25 km), with the exception of the region between ~21 and 25 km which is again explained by the special feature in the B45 data (Fig. 6). Figure 9 shows the differences for the entire data set and the low spectral resolution data set when the region between 21 and 25 km from the flight B45



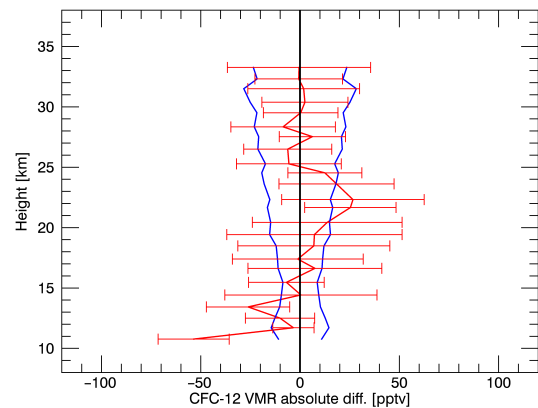
**Figure 8.** Same as Fig. 4, but instead of CH<sub>4</sub> now N<sub>2</sub>O.



**Figure 9.** Same as Fig. 5, but instead of CH<sub>4</sub> now N<sub>2</sub>O.



**Figure 10.** Same as Fig. 1, but instead of CH<sub>4</sub> now CFC-12.



**Figure 11.** Same as left panel of Fig. 3, but instead of CH<sub>4</sub> now CFC-12.

is eliminated from the data set. An overall display shows the mean deviation is affected by the sharp structure.

### 3.3 Validation of MIPAS-E CFC-12

The vertical structure of the CFC-12 profile for flight B45 (Fig. 10) looks similar to the profiles of CH<sub>4</sub> and N<sub>2</sub>O, which are displayed in Figs. 1 and 5 respectively. The altitude range

including the dynamical feature shows large differences between the matched MIPAS-E data and the in situ data. Again the variability in the matched MIPAS-E data is larger in the region of the dynamical feature indicating that the mismatch is most probably due to inaccuracies of the match technique applied here. In contrast to CH<sub>4</sub> and N<sub>2</sub>O there is also a significant discrepancy below about 13 km, where the CFC-



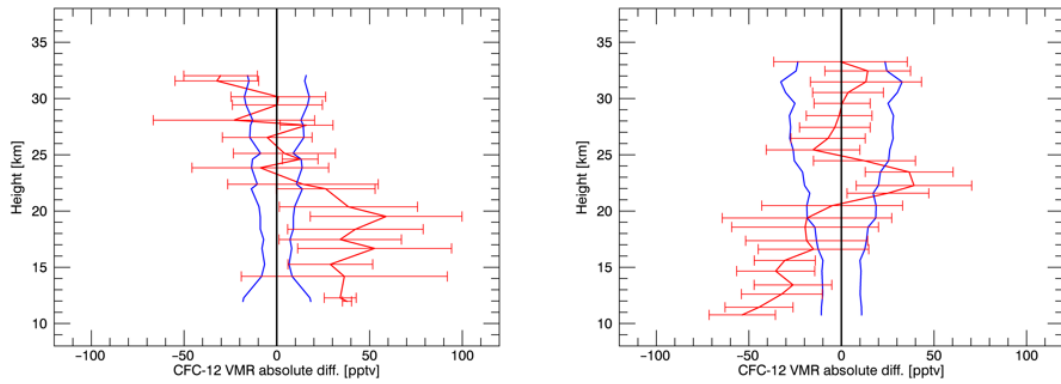


Figure 12. Same as Fig. 4, but instead of CH<sub>4</sub> now CFC-12.

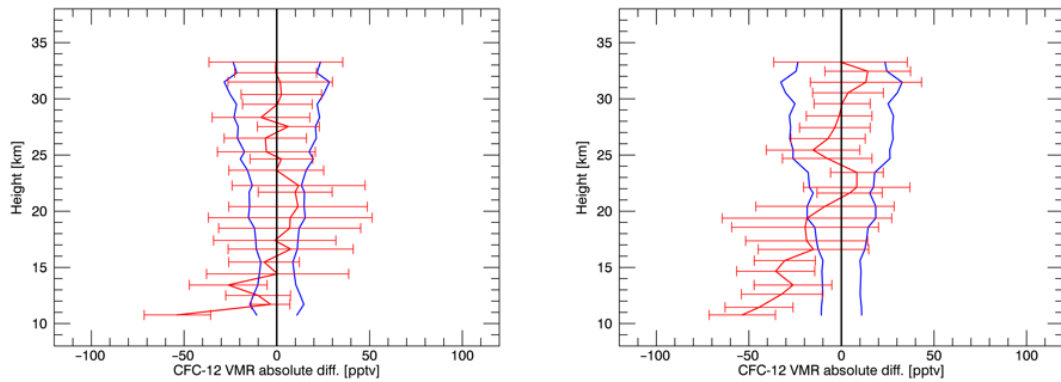


Figure 13. Same as Fig. 5, but instead of CH<sub>4</sub> now CFC-12.

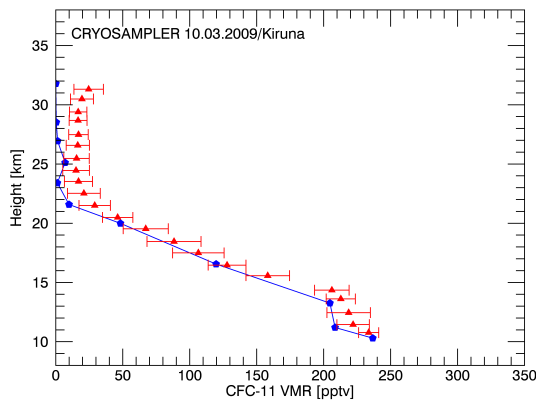


Figure 14. Same as Fig. 1, but instead of CH<sub>4</sub> now CFC-11.

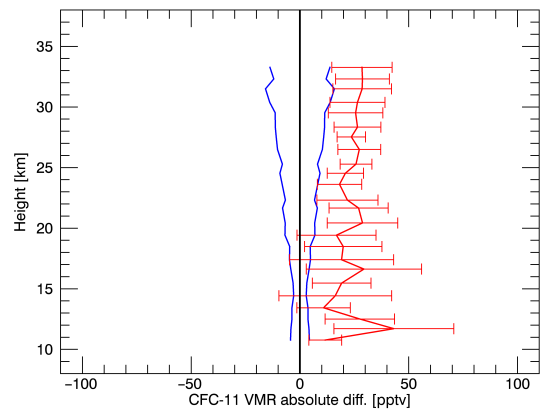
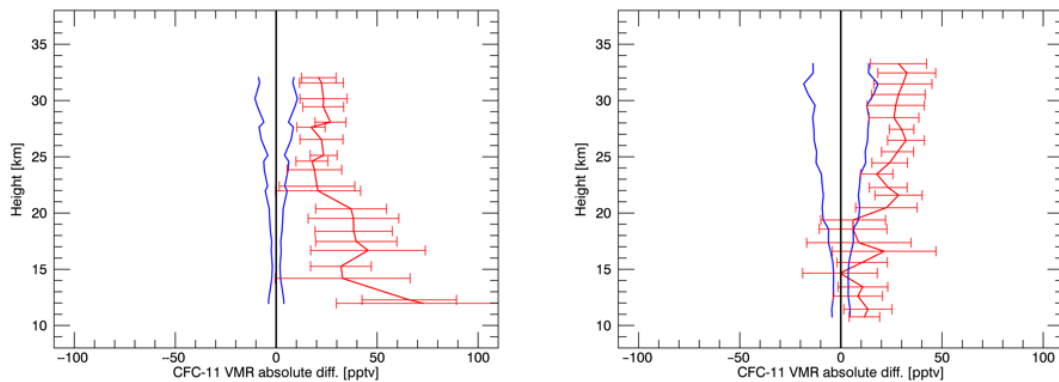


Figure 15. Same as left panel of Fig. 3, but instead of CH<sub>4</sub> now CFC-11.

12 mixing ratios measured by MIPAS-E underestimate the in situ data. Using all available flights the differences between MIPAS-E and in situ data are mostly within the measurement error specified for MIPAS-E (Fig. 11). Especially above 25 km the agreement is within the stated uncertainties. Below ~13 km the mixing ratio values are underestimated, most significantly in the lowest bin between 10 and 11 km. For the region below 20 km a very high variability is ob-

served in the validation results as indicated by the large error bars, which are substantially larger than the quoted uncertainty. As in the case of CH<sub>4</sub> and N<sub>2</sub>O the combination of all data may not reveal all aspects of the validation. Figure 11 shows the difference between the validation based on the high spectral resolution validation flights and the low spectral resolution ones respectively. For the validation based on



**Figure 16.** Same as Fig. 4, but instead of CH<sub>4</sub> now CFC-11.

the high spectral resolution validation flights (left panel), the mixing ratio values measured by MIPAS-E are overestimated below 20 km, whereas based on the low spectral resolution validation flights (right panel) there is an underestimation. So for the combined validation displayed in Fig. 10, these two discrepancies nearly compensate each other. This is reflected in the large variability below 20 km shown in Fig. 11. Figure 12 (right panel) shows again a significant discrepancy in the 21–25 km altitude region where the dynamical feature is observed in flight B45. Figure 13 again shows the result of eliminating this feature from the validation of the entire period (left panel) and the low spectral resolution data after January 2005 (right panel). In both cases the results from the validation above 20 km are within the given uncertainty estimate. Below 20 km the mean good agreement in the overall validation is again an effect of a cancellation of the overestimation of the high spectral resolution validation flights and the underestimation of the low spectral resolution validation flights.

### 3.4 Validation of MIPAS-E CFC-11

CFC-11 has the shortest stratospheric lifetime of all trace gases discussed here (SPARC, 2013). The vertical profile observed on 10 March 2009 (flight B45) is displayed in Fig. 14 in comparison to the matched satellite data. The strong dynamical feature observed in the profiles of N<sub>2</sub>O, CH<sub>4</sub>, and CFC-12 is much less obvious in the CFC-11 data of the cryogenic whole sampler. Upon close inspection higher mixing ratios of CFC-11 are indeed observed at 25 km (thus just above the dynamical feature) in comparison to the samples collected around 21 and 23 km altitude. The reason for this behaviour is that CFC-11 is nearly completely photolysed in this region of the stratosphere already, making it less sensitive to dynamical variability. The comparison in Fig. 14 does however reveal that the mixing ratios of CFC-11 are overestimated by the retrieval algorithm over the entire range of the profile. Even at altitudes where CFC-11 is no longer measured by the cryogenic whole sampler (i.e. the mixing ratio is

below the detection limit of about 0.5 ppt), the retrieval yields several tens of ppt of CFC-11, which is clearly unrealistic. As can be seen in Fig. 15 the combined validation of CFC-11 shows the lowest quality in comparison to the validation of CH<sub>4</sub>, N<sub>2</sub>O, and CFC-12. For the full altitude range the CFC-11 mixing ratio values are overestimated and not explainable with the uncertainty stated in the MIPAS-E data used here. The overestimation of the CFC-11 mixing ratios is obvious in both the high spectral resolution data and the low spectral resolution data, although the differences are largest in the upper part of the profiles for the later data and in the lowest part of the earlier data (see Fig. 16). The differences between satellite and balloon mixing ratio cannot be explained within the measurement uncertainty for the data with high spectral resolution. While the CFC-11 mixing ratios of the MIPAS derived with this retrieval for the low spectral resolution data still overestimate the balloon data, the agreement is within the uncertainties below 20 km altitude.

## 4 Discussion and conclusion

The comparison of the results of in situ measurements and satellite instruments is important in order to assess the quality of the satellite data. The database available in this study for the validation of MIPAS-E operational retrievals (version 6.0) of CH<sub>4</sub>, N<sub>2</sub>O, CFC-12, and CFC-11 is rather small with only seven flights, i.e. seven vertical profiles from the in situ observations. In addition the validation took place under different geophysical and meteorological conditions, and the spectral instrumental resolution of MIPAS-E was changed as of January 2005 due to technical problems, which requires additional attention as the spectral resolution will affect the retrieval results. The validation results are not always unambiguous; in particular significant differences are observed in the comparison with the high spectral resolution and the low spectral resolution data. Surprisingly the data after January 2005 (low spectral resolution) seem to show a better agreement with the in situ data. Above 20 km altitude there is generally a good agreement between both data sets

within the errors given in the MIPAS-E retrieval for CH<sub>4</sub>, N<sub>2</sub>O, and CFC-12. For the high spectral resolution data a significant underestimation of both CH<sub>4</sub> and N<sub>2</sub>O in this retrieval version is observed above 25 km which can be on the order of 100 ppb of CH<sub>4</sub> (up to about 20 %) and 30 ppb N<sub>2</sub>O (up to about 30 %) around 30 km altitude. An underestimation of the CFC-12 mixing ratios is also observed, but this feature is closer to the uncertainty range given in the CFC-12 retrieval. In contrast to this we find a significant overestimation of mixing ratios of CH<sub>4</sub>, N<sub>2</sub>O, and CFC-12 below 20 km, which may be as large as 300 ppb of CH<sub>4</sub> (up to 25 %), 40 ppb of N<sub>2</sub>O (up to 20 %), and 50 ppt of CFC-12 (up to 20 %). In combination this should result in an overestimation of the vertical gradient of these three trace gases in the high spectral resolution data.

After elimination of a dynamical feature, which is probably not well caught in the trajectory matching, the low spectral resolution data generally agree better with our in situ data. Above 20 km there is generally a good agreement which is within the estimated uncertainty limits for CH<sub>4</sub>, N<sub>2</sub>O, and CFC-12. For CH<sub>4</sub> a good agreement within the stated uncertainty is also observed below 20 km, while there seems to be a tendency for an underestimation of the mixing ratios of N<sub>2</sub>O (20 ppb or around 15 % at around 15 km) and also of CFC-12 (50 ppt or about 10 % at 10 km).

For CFC-11 the validation shows a very consistent picture. The ESA version 6.0 retrieval of MIPAS-E clearly and significantly overestimates the CFC-11 mixing over the entire stratosphere both in the high spectral resolution and the low spectral resolution data. This version of the CFC-11 retrieval cannot be recommended for scientific studies.

*Acknowledgements.* We would like to acknowledge Ingeborg Levin from University of Heidelberg for providing measurements of CH<sub>4</sub> and N<sub>2</sub>O for the samples of the cryosampler. The validation flights were part of the Envisat validation programme and were funded through projects from ESA (ESRIN contract no. 20752/08/I-OL) and from the German Federal Ministry of Education and Research (grant no. 50EE0016 and 50EE0842). Operational Envisat data were provided by ESA.

Edited by: C. von Savigny

## References

- Bauer, R., Engel, A., Franken, H., Klein, E., Kulesa, G., Schiller, C., Schmidt, U., Borchers, R., and Lee, J.: Monitoring the Vertical Structure of the Arctic Polar Vortex over Northern Scandinavia during Easoe – Regular N<sub>2</sub>O Profile Observations, *Geophys. Res. Lett.*, 21, 1211–1214, 1994.
- Dlugokencky, E. J., Myers, R. C., Lang, P. M., Masarie, K. A., Crotwell, A. M., Thoning, K. W., Hall, B. D., Elkins, J. W., and Steele, L. P.: Conversion of NOAA atmospheric dry air CH<sub>4</sub> mole fractions to a gravimetrically prepared standard scale, *J. Geophys. Res.*, 110, D18306, doi:10.1029/2005jd006035, 2005.
- Eckert, E., Laeng, A., Lossow, S., Kellmann, S., Stiller, G., von Clarmann, T., Glatthor, N., Höpfner, M., Kiefer, M., Oelhaf, H., Orphal, J., Funke, B., Grabowski, U., Haenel, F., Linden, A., Wetzel, G., Woiwode, W., Bernath, P. F., Boone, C., Dutton, G. S., Elkins, J. W., Engel, A., Gille, J. C., Kolonjari, F., Sugita, T., Toon, G. C., and Walker, K. A.: MIPAS IMK/IAA CFC-11 (CCl<sub>3</sub>F) and CFC-12 (CCl<sub>2</sub>F<sub>2</sub>) measurements: accuracy, precision and long-term stability, *Atmos. Meas. Tech. Discuss.*, 8, 7573–7662, doi:10.5194/amtd-8-7573-2015, 2015.
- Engel, A., Schmidt, U., and Stachnik, R. A.: Partitioning between chlorine reservoir species deduced from observations in the arctic winter stratosphere, *J. Atmos. Chem.*, 27, 107–126, 1997.
- Engel, A., Möbius, T., Haase, H.-P., Bönisch, H., Wetter, T., Schmidt, U., Levin, I., Reddmann, T., Oelhaf, H., Wetzel, G., Grunow, K., Huret, N., and Pirre, M.: Observation of mesospheric air inside the arctic stratospheric polar vortex in early 2003, *Atmos. Chem. Phys.*, 6, 267–282, doi:10.5194/acp-6-267-2006, 2006.
- Engel, A., Möbius, T., Bönisch, H., Schmidt, U., Heinz, R., Levin, I., Atlas, E., Aoki, S., Nakazawa, T., Sugawara, S., Moore, F., Hurst, D., Elkins, J., Schauffler, S., Andrews, A., and Boering, K.: Age of stratospheric air unchanged within uncertainties over the past 30 years, *Nat. Geosci.*, 2, 28–31, doi:10.1038/NGEO388, 2009.
- Engel, A., Atlas, E. L., Bernath, P. F., Bönisch, H., Brown, A., Laube, J., Minschwaner, K. R., Montzka, S. A., O'Doherty, S., Prinn, R. G., Rigby, M., Schauffler, S. M., Volk, C. M., and Yvon-Lewis, S.: Inferred Lifetimes from Observed Trace-Gas Distributions, in: SPARC Report No. 6 (2013) Lifetimes of Stratospheric Ozone-Depleting Substances, Their Replacements, and Related Species; edited by: Ko, M. K. W., Newman, P. A., Reimann, S., and Strahan, S. E., 2013.
- Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, *Atmos. Chem. Phys.*, 8, 2151–2188, doi:10.5194/acp-8-2151-2008, 2008.
- Grunow, K.: Anwendung von Trajektorien zur ENVISAT-Validierung und zur Untersuchung der Luftmassenherkunft in der Stratosphäre, PhD, Free University Berlin, Berlin, Germany, 2009.
- Hall, B. D., Dutton, G. S., and Elkins, J. W.: The NOAA nitrous oxide standard scale for atmospheric observations, *J. Geophys. Res.*, 112, D09305, doi:10.1029/2006jd007954, 2007.
- Hall, T. M. and Plumb, R. A.: Age as a diagnostic of stratospheric transport, *J. Geophys. Res.*, 99, 1059–1070, doi:10.1029/93jd03192, 1994.
- Hoffmann, L., Kaufmann, M., Spang, R., Müller, R., Remedios, J. J., Moore, D. P., Volk, C. M., von Clarmann, T., and Riese, M.: Envisat MIPAS measurements of CFC-11: retrieval, validation, and climatology, *Atmos. Chem. Phys.*, 8, 3671–3688, doi:10.5194/acp-8-3671-2008, 2008.
- Hoker, J., Obersteiner, F., Bönisch, H., and Engel, A.: Comparison of GC/time-of-flight MS with GC/quadrupole MS for halocarbon trace gas analysis, *Atmos. Meas. Tech.*, 8, 2195–2206, doi:10.5194/amt-8-2195-2015, 2015.

- Kellmann, S., von Clarmann, T., Stiller, G. P., Eckert, E., Glatthor, N., Höpfner, M., Kiefer, M., Orphal, J., Funke, B., Grabowski, U., Linden, A., Dutton, G. S., and Elkins, J. W.: Global CFC-11 ( $\text{CCl}_3\text{F}$ ) and CFC-12 ( $\text{CCl}_2\text{F}_2$ ) measurements with the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS): retrieval, climatologies and trends, *Atmos. Chem. Phys.*, 12, 11857–11875, doi:10.5194/acp-12-11857-2012, 2012.
- Laeng, A., Plieninger, J., von Clarmann, T., Grabowski, U., Stiller, G., Eckert, E., Glatthor, N., Haenel, F., Kellmann, S., Kiefer, M., Linden, A., Lossow, S., Deaver, L., Engel, A., Hervig, M., Levin, I., McHugh, M., Noël, S., Toon, G., and Walker, K.: Validation of MIPAS IMK/IAA methane profiles, *Atmos. Meas. Tech.*, 8, 5251–5261, doi:10.5194/amt-8-5251-2015, 2015.
- Laube, J. C., Engel, A., Bönisch, H., Möbius, T., Worton, D. R., Sturges, W. T., Grunow, K., and Schmidt, U.: Contribution of very short-lived organic substances to stratospheric chlorine and bromine in the tropics – a case study, *Atmos. Chem. Phys.*, 8, 7325–7334, doi:10.5194/acp-8-7325-2008, 2008.
- Levin, I., Glatzel-Mattheier, H., Marik, T., Cuntz, M., Schmidt, M., and Worthy, D. E.: Verification of German methane emission inventories and their recent changes based on atmospheric observations, *J. Geophys. Res.*, 104, 3447–3456, doi:10.1029/1998JD100064, 1999.
- Lueb, R. A., Ehhalt, D. H., and Heidt, L. E.: Balloon-Borne low-Temperature air sampler, *Rev. Sci. Instrum.*, 46, 702–705, doi:10.1063/1.1134292, 1975.
- Manney, G. L., Santee, M. L., Rex, M., Livesey, N. J., Pitts, M. C., Veefkind, P., Nash, E. R., Wohltmann, I., Lehmann, R., Froidevaux, L., Poole, L. R., Schoeberl, M. R., Haffner, D. P., Davies, J., Dorokhov, V., Gernandt, H., Johnson, B., Kivi, R., Kyro, E., Larsen, N., Levelt, P. F., Makshtas, A., McElroy, C. T., Nakajima, H., Parrondo, M. C., Tarasick, D. W., von der Gathen, P., Walker, K. A., and Zinoviev, N. S.: Unprecedented Arctic ozone loss in 2011, *Nature*, 478, 469–475, 2011.
- Minschwaner, K., Hoffmann, L., Brown, A., Riese, M., Müller, R., and Bernath, P. F.: Stratospheric loss and atmospheric lifetimes of CFC-11 and CFC-12 derived from satellite observations, *Atmos. Chem. Phys.*, 13, 4253–4263, doi:10.5194/acp-13-4253-2013, 2013.
- Montzka, S. A., Butler, J. H., Hall, B. D., Mondeel, D. J., and Elkins, J. W.: A decline in tropospheric organic bromine, *Geophys. Res. Lett.*, 30, 1826, doi:10.1029/2003gl017745, 2003.
- Naujokat, B. and Grunow, K.: The stratospheric arctic winter 2002/03: Balloon flight planning by trajectory calculations, 16th Esa Symposium on European Rocket and Balloon Programmes and Related Research, Proceedings, 530, 421–425, 2003.
- Payan, S., Camy-Peyret, C., Oelhaf, H., Wetzell, G., Maucher, G., Keim, C., Pirre, M., Huret, N., Engel, A., Volk, M. C., Kuellmann, H., Kuttippurath, J., Cortesi, U., Bianchini, G., Mencaraglia, F., Raspollini, P., Redaelli, G., Vigouroux, C., De Mazière, M., Mikuteit, S., Blumenstock, T., Velasco, V., Notholt, J., Mahieu, E., Duchatelet, P., Smale, D., Wood, S., Jones, N., Piccolo, C., Payne, V., Bracher, A., Glatthor, N., Stiller, G., Grunow, K., Jeseck, P., Te, Y., and Butz, A.: Validation of version-4.61 methane and nitrous oxide observed by MIPAS, *Atmos. Chem. Phys.*, 9, 413–442, doi:10.5194/acp-9-413-2009, 2009.
- Plumb, R. A. and Ko, M. K. W.: Interrelationships between Mixing Ratios of Long Lived Stratospheric Constituents, *J. Geophys. Res.*, 97, 10145–10156, 1992.
- Plumb, R. A., Heres, W., Neu, J. L., Mahowald, N. M., del Corral, J., Toon, G. C., Ray, E., Moore, F., and Andrews, A. E.: Global tracer modeling during SOLVE: High-latitude descent and mixing, *J. Geophys. Res.*, 108, 8309, doi:10.1029/2001JD001023, 2002.
- Schmidt, M., Glatzel-Mattheier, H., Sartorius, H., Worthy, D. E., and Levin, I.: Western European  $\text{N}_2\text{O}$  emissions: A top-down approach based on atmospheric observations, *J. Geophys. Res.*, 106, 5507–5516, doi:10.1029/2000jd900701, 2001.
- Schwarzenberger, T.: Validierung von MIPAS-ENVISAT-Messungen mithilfe von in situ Ballonmessungen, Goethe University Frankfurt, Frankfurt am Main, Germany, 2014.
- SPARC: SPARC Report on the Lifetimes of Stratospheric Ozone-Depleting Substances, Their Replacements, and Related Species, SPARC Report No. 6, WCRP-15/2013, 2013.
- Stiller, G. P., von Clarmann, T., Höpfner, M., Glatthor, N., Grabowski, U., Kellmann, S., Kleinert, A., Linden, A., Milz, M., Reddmann, T., Steck, T., Fischer, H., Funke, B., López-Puertas, M., and Engel, A.: Global distribution of mean age of stratospheric air from MIPAS SF<sub>6</sub> measurements, *Atmos. Chem. Phys.*, 8, 677–695, doi:10.5194/acp-8-677-2008, 2008.
- Vigouroux, C., De Mazière, M., Errera, Q., Chabrillat, S., Mahieu, E., Duchatelet, P., Wood, S., Smale, D., Mikuteit, S., Blumenstock, T., Hase, F., and Jones, N.: Comparisons between ground-based FTIR and MIPAS  $\text{N}_2\text{O}$  and  $\text{HNO}_3$  profiles before and after assimilation in BASCOE, *Atmos. Chem. Phys.*, 7, 377–396, doi:10.5194/acp-7-377-2007, 2007.
- Volk, C. M., Elkins, J. W., Fahey, D. W., Salawitch, R. J., Dutton, G. S., Gilligan, J. M., Proffitt, M. H., Loewenstein, M., Podolske, J. R., Minschwaner, K., Margitan, J. J., and Chan, K. R.: Quantifying transport between the tropical and mid-latitude lower stratosphere, *Science*, 272, 1763–1768, 1996.
- Volk, C. M., Elkins, J. W., Fahey, D. W., Dutton, G. S., Gilligan, J. M., Loewenstein, M., Podolske, J. R., Chan, K. R., and Gunson, M. R.: Evaluation of source gas lifetimes from stratospheric observations, *J. Geophys. Res.*, 102, 25543–25564, doi:10.1029/97jd02215, 1997.