Interactive comment on “Envisat MIPAS measurements of CFC-11: retrieval, validation, and climatology” by L. Hoffmann et al.

L. Hoffmann et al.

Received and published: 27 May 2008

We thank the anonymous referee for the time and effort spend on reading and correcting the paper. We greatly acknowledge the helpful comments and suggestions. Below please find the reply to each comment (indicated by >>>>) and a description of the actions taken.

General Comments

The paper describes a new retrieval technique for analysing MIPAS spectra based on (a) using a spectrally-averaged radiance, effectively synthesising a filter radiometer, and (b) using a band-averaged forward model, both of which represent significant differences from the approaches used by other groups (ie using measurements at their full spectral sampling and monochromatic forward models). The advantage of such a
technique is speed, although it is limited to molecules whose emission features dominate particular regions of the spectrum (ie those which could also be retrieved using a filter radiometer). As presented here, it is applied to the retrieval of CFC-11 and relies on retrievals of pressure, temperature and the main absorbing molecules from the ESA operational retrievals. It is not clear whether the technique could be used as an alternative to the entire ESA retrieval, in which case the CPU saving could be significant, or whether it simply represents a faster technique for retrieving particular additional species, in which case the speed is less important.

To better put the work described in this paper in a context, we added these paragraphs to the introduction: "In this paper we discuss a retrieval of CFC-11 global distributions from measurements by the Envisat MIPAS satellite experiment utilising a new fast forward model and optimal estimation retrieval processor developed in Juelich. The forward model utilises the emissivity growth approximation (EGA) to significantly accelerate the radiative transfer calculations. Envisat MIPAS radiance measurements are spectrally averaged for the retrieval, effectively synthesising a filter radiometer. This represents a significant difference from the approach used by other groups, i.e. using measurements at their full spectral sampling and monochromatic forward models." and "The advantage of the EGA-technique is speed. Hence, it is especially suited for comprehensive analysis of infrared remote-sensing measurements from satellite missions. EGA-based forward models have been used for operational data processing for several satellite experiments (e.g. CLAES, HALOE, CRISTA, SABER, and HIRDLS). Since the European Space Agency (ESA) provides consolidated and validated Envisat MIPAS retrieval data for temperature and six major trace gases, we do not attempt to create a retrieval system that entirely replaces this work. We try to complement the ESA data with comprehensive datasets of minor trace species not covered by the operational analysis that will be of particular interest for scientific users."

While there is no reason to expect that this will be better than the conventional approach to retrieving CFC-11 from MIPAS spectra, the validation results presented show
that it appears to be consistent at the 10% level with other measurements and retrievals. This builds on earlier work by the same group (Hoffmann et al 2005), which describes a retrieval of both CFC-11 and CFC-12, but includes a more comprehensive error analysis and validation. However, by continuing to synthesise a simple broad-band filter from the MIPAS spectra using a box-car function, I feel that they have missed an opportunity in developing the theory further: better use could be made of the spectral resolution available from MIPAS by constructing a "customised" filter, e.g. increasing the relative response to regions of the spectrum where the target molecule emission dominates and reducing it where interfering species dominate. Such an approach would have allowed better precision and accuracy, and extend the number of potential target species.

>>> We studied the option of "customised" filters as part of a PhD thesis (Hoffmann, 2006). However, this approach does not lead to significant improvements of the retrieval results. Compared with the "boxcar" filters used here (note: the real filter functions are not boxcar, see comment 6), the increase in Shannon information content by "customising" is less than 0.5% for the CFC-11 band and less than 1.1% for the continuum band. This may be plausible considering the fact that CFC-11 has two rather broad spectral bands in the mid-IR and radiance contributions of interfering species are rare in these band. Using "customised" filters may introduce additional retrieval errors (e.g. wavelength calibration or uncertainties of ILS) that can safely be neglected with the current approach. Hence, for the CFC-11 retrieval the "boxcar" filters are a good choice.

Specific Comments

1) Abstract should give vertical range over which statements relating to retrieval accuracy and a priori contributions apply.

>>> We added "in the altitude range 10 to 25 km" to the abstract according to the referees comment.
2) Section 2: discusses MIPAS in the present tense. While some statements are appropriate, eg MIPAS still measures over the range 685-2410cm⁻1, other statements belong in the past tense, such as spectral resolution and those relating to the limb scanning sequence 6-68km. To make sense of this, it should be stated here that MIPAS is currently operating at a reduced resolution with a different scanning sequence. p4565, line 8: MIPAS is now operating continuously again, and hasn’t been in "campaign-orientated" mode since Feb 2007.

>>> We rewrote the last paragraph of sec. 2 and added a new par. 2: "The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) (Fischer, 1996, Endemann, 1999, Fischer, 2007) is one of the atmospheric experiments aboard Envisat. Envisat MIPAS provided nearly continuous measurement coverage for the period July 2002 to March 2004. Due to an unexpected technical problem with the interferometer slides in March 2004, the instrument was operated at a reduced spectral resolution and in a campaign-orientated mode till February 2007. It is now operating continuously again with reduced spectral resolution and a new spatial scanning sequence. In this paper we analyse measurements obtained during the first two years of operation where consolidated Level-1B and Level-2 data products are available from ESA (processing software version 4.61 and 4.62)." We put the other text of the section in the past tense if appropriate.

3) Section 3 Retrieval of CFC abundance p4565, line 20: I wasn’t aware of a CFC-11 band from 910-960cm⁻1, and nothing seems to be listed in the HITRAN database. If there were such a feature it might be promising for a retrieval since this generally a cleaner part of the atmospheric window than 800-885cm⁻1.

>>> We rewrote the first paragraph of section 3: "In the mid infrared CFC-11 strongly radiates in the nu_4 band at 800 to 885 cm⁻1 with a band strength of about 6.5 x 10⁻17 cm²/(molec cm) at 293 K and the nu_1 band at 1045 to 1120 cm⁻1 with a band strength of about 2.5 x 10⁻17 cm²/(molec cm) at 293 K (e.g. McDaniel, 1991, Varanasi, 1991). The nu_2+nu_3 combination band weakly emits at 910 to 960 cm⁻1.
1 with a band strength of $0.17 \times 10^{-17} \text{ cm}^2/(\text{molec \ cm})$ at 293 K. The $\nu_2+\nu_3$ combination band is too weak for a retrieval analysis and spectroscopic data are not provided in the High Resolution Transmission (HITRAN) compilation (Rothman, 2003).

4) p4566, lines 10-14: The given percentages do not total 100% so what are they percentages of? There are several different ways in which the contributions of different absorbers can be defined (e.g., radiance of an atmosphere containing one absorber only, or change in radiance if one absorber is removed). What has been assumed here?

>>> We added the sentences "For this purpose the radiance of an atmosphere containing one absorber only is compared with the radiance of an atmosphere containing all absorbers." and "Since the maximum contributions are obtained at different heights and for different atmospheric conditions the percentages listed above do not total to 100%." to clarify.

I would have liked to have seen a spectral plot showing the contributions of the different absorbers at some representative tangent height.

>>> We added the spectra plot in Fig. 1b. This allows for a comparison with measured data (Fig. 1a). Please note that the radiance units in the old plot in Fig. 1a were wrong and had to be corrected.

5) p4566 lines 26: are these S/N values (3.1-190) for the CFC-11 retrieval or the continuum retrieval? Presumably CFC-11 is also retrieved from the "continuum" window but it’s not clear from what has been written.

>>> We rewrote: "In the 10 to 40 km altitude range the signal-to-noise ratio varies between 2.1 and 650 in the CFC-11 spectral window (844.275 to 850.575 cm⁻¹). It varies between 3.1 and 190 for the continuum spectral window (830.350 to 839.475 cm⁻¹)." It is pointed out later in the paper that aerosol and CFC-11 are retrieved in a multi-target approach, i.e. both targets are retrieved from both spectral windows.

6) p4567/68 Data pre-processing Apodisation and spectral averaging are both linear...
processes, ie the end product, spectrally averaged radiance over a microwindow, is a linear function of the set of (unapodised) L1B radiances within the microwindow. I don’t understand why each points has to be apodised, rather than simply applying a trapezoidal-like function to the raw spectra which is (a) much faster and (b) mathematically identical. See also comment (11).

>>> A trapezoidal-like function is not the correct filter function (though it may be a good approximation depending on the size of the spectral window). The correct filter function is obtained by averaging all ILS or AILS within the corresponding spectral window (based on the linearity argument pointed out by the referee). If apodization is neglected the filter function will have side-lobes around the boundaries of the spectral window. To better illustrate this we added the new Fig. 4 to the paper.

7) p4568/4569 Optimal Estimation Retrieval For those not familiar with OE, some equations would be clearer than attempting to explain in words. Similarly for the discussion of correlation lengths in covariance matrices (p4569 line 8).

>>> The corresponding equations have been added. The text was slightly modified to explain the equations.

8) p4569 lines 12-17: strictly speaking, most of these terms are errors in the forward model parameters rather than in the measurements, but are conveniently combined into the "measurement" covariance matrix.

>>> We agree that the term "measurement covariance" might not be best in this context. We use it to be compliant with the references.

9) p4569 line 18: "A priori atmospheric state". "A Priori" is conventionally used to refer to parameters which are then retrieved, and therefore also have a different "a posteriori" value. However in this case it seems to be used to refer to both parameters which are retrieved (CFC-11 and aerosol) and those which remain fixed (ESA products, other species profiles). Similarly p4575 line 26: reference to Temperature as "a priori".
We replaced "The a priori atmospheric state..." by "The a priori state vector and the parameter vector..." to be more precise. We also modified the wrong reference to temperature as "a priori" by "parameter". We checked all other references to "a priori" accordingly.

10) p4569 line 25: why is climatological data used from just this limited set of atmospheres when Remedios also provides the MIPAS "initial guess" atmospheres conveniently divided into season and latitude?

>>> When we started this work in 2003 the IG2 atmospheres were not commonly used and publicly available. A first reference to the IG2 atmospheres was published in 2007 in ACP. Originally, the IG2 atmospheres did not seem to be a good choice because standard deviation data, which are essential for the optimal estimation analysis, were not provided.

11) p4570 line 22: In order to justify the $1/\sqrt{n}$ scaling it is also assumed that the noise is uncorrelated between adjacent spectral points, from which it follows that it is also uncorrelated between different windows and different altitudes (as stated subsequently). This is a reasonable assumption for unapodised spectra and should also hold for simply-averaged apodised spectra.

>>> We agree with this statement and thank the referee for the comment.

12) p4570 line 28-29: Is there any justification for this 10 cm$^{-1}$ spectral correlation length? Since there are only two spectral windows it seems safer, and simpler, just to assume the worst case of no correlation. Vertical correlation is slightly more problematic, but if these really are calibration errors wouldn’t they apply to all tangent altitudes (ie infinite vertical correlation length)?

>>> Unfortunately, only standard deviations for offset and gain calibration errors are provided by ESA but correlation data is missing. Hence, the correlation lengths had to be chosen ad hoc, as pointed out in the paper. Colleagues with instrument expe-
perience pointed out that temperature drifts of the instrument or stray-light effects may reduce the spatial correlation lengths of the calibration errors compared with the fully correlated case (what one may expect from the routine calibration setup). To give the reader more guidance, we added in this paragraph: "Sensitivity studies indicate a weak dependence of the estimated retrieval error on the spectral correlation length (over the range from 1 to 100 cm^-1), but a strong dependence on the vertical correlation length (Hoffmann, 2006). The strongest dependence was found for the offset calibration error. The estimated retrieval error varies up to a factor 5 if the offset calibration errors are assumed to be either uncorrelated or fully correlated in altitude."

13) p4571 line 17: Spectroscopic data - why 10km vertical correlation length? If it’s basically a 3% uncertainty in the band strength shouldn’t it be fully correlated at all altitudes and between both spectral windows? Ie this should translate directly to a 3% uncertainty in the CFC-11 retrieval error.

>>> We agree with the referee that spectroscopic errors are fully correlated if the radiative transfer calculations are based on line-data. However, the calculations for CFC-11 are based on measured absorption cross-sections, tabulated for distinct pressure/temperature combinations in the HITRAN compilation. For an exact analysis we need to know how the errors of the database entries are correlated, but this information is not provided. In this case it seems safer to assume that the spectroscopic errors are not fully correlated. This will increase the retrieval error estimate. We added "for a conservative estimate" after the specified correlation length, to guide the reader.

14) p4571 There are significant errors (~10%) associated with the ESA L2 estimates of H2O, HNO3, O3 etc - where are these? Given that O3 and HNO3, in particular, contribute ~10s% to the radiance I’m surprised if these are minor.

>>> We checked the error analysis and confirm that the retrieval errors due to 10% uncertainties (min) for HNO3 and O3 are less than 1ppt at all altitudes. The retrieval error for H2O is up to 6-7 ppt at 6 km altitude as shown in Fig. 5. This result may
not seem implausible considering that a 10% uncertainty in volume mixing ratio and radiances contributions less than 10% at relevant heights cause a radiances uncertainty smaller than 1% (for optically thin conditions) which is small compared to other error sources. We decided to not include the minor error sources to keep the plot simple.

15) A more sophisticated representation of the measurement-forward model covariance matrix is used in this retrieval than elsewhere: usually just the random noise term is used. However I don’t have any indication of whether the extra complication significant changes the mean value of the retrieved profiles. Was it necessary?

>>> We carried out test retrievals to assess this point. We conclude (sec. 3.3, par 4): "This is a more sophisticated representation of the measurement covariance matrix than used elsewhere. Compared to test retrievals which only consider noise in the measurement covariance, we find substantial differences in the retrieval results at lower altitudes at which noise is a minor source of error. Using the full error covariance increases the amount of a priori information by 5 to 10%, degrades the vertical resolution by about 10 to 20% and increases the estimated retrieval errors by about 20 to 30%. On the other hand, the negative impact of systematic errors on the retrieval results will be reduced significantly."

16) p4571 Smoothing Error: while it is no longer possible to use the a priori covariance to determine the smoothing error, it seems it should be possible to determine how the retrieval would smooth an atmosphere of the more appropriate CFC-11 a priori covariance matrix, hence determine the "correct" smoothing error, albeit not by the direct method.

>>> We carried out another test retrieval to investigate this. We conclude (sec. 3.4, last paragraph): "However, for a test retrieval we estimated the smoothing error based on the unscaled a priori covariance. The estimated smoothing error varies in between 6 to 12 ppt (3.7 to 5.4%)."

17) p4572/73 Vertical Resolution: I wasn’t aware of this definition of vertical resolution
(most people use width of the averaging kernel functions) but it seems plausible, and a more easily-defined quantity.

>>> We agree with this comment. The method is referenced in a comparison of resolution measures in (Rodgers, 2000).

18) p4573 Internal Quality Measures An equation defining chi2 would be clearer here. >>> We added a reference to Eq. (1). The chi2-test is applied to the final value of the objective function.

19) p4574 line 3: "significance at the 0.1% confidence level" sound like a low degree of confidence in the significance. I guess what is meant is that a high-ish value of chi2 has been used that would be expected to exclude only 0.1% of data generated by purely random statistics.

>>> To clarify we rewrote: "However, to prepare the CFC-11 data for subsequent scientific studies, we decided to use a weak filtering criterion and remove only the few extreme outliers. Hence, the individual retrievals are already accepted as successful if the value of the chi^2-statistics satisfies the 0.1% confidence level. This will exclude only 0.1% of data generated by purely random statistics."

20) p4574 lines 6-7: By the "chi2/m distribution for the initial guess" I assume that this means just the \((y-f(a))^T S_y^{-1} (y-f(a))\) part of the chi2 function since the \((x-a)^T S_a^{-1} (x-a)\) component is identically zero? (an equation would have been clearer).

>>> To clarify we rewrote in this paragraph: "For comparison the normalised chi^2/m-distribution for the initial guess is shown, too. Since the a priori state is used as initial guess, \(x_0=x_a\), this distribution measures only the difference \([y-F(x)]^T S_y^{-1} [y-F(x)]\) between the radiance measurements and the forward model fit while the difference \((x-x_a)^T S_a^{-1} (x-x_a)\) is identically zero. A peak value around 12 indicates that the initial fits are not consistent with the measurements. The final chi^2/m-distribution clearly indicates that the majority of retrieval results is consistent with the measure-
ments and a priori."

21) p4574 line 7: states that the final chi2/m distribution shows that the majority of the retrieval results are consistent, but consistent with what?

>>> To clarify we rewrote "consistent with the measurements and a priori".

If the a priori and measurement covariance distributions were correct I would expect a peak value of around 1, yet the peak appears to be between 0.3 and 0.4, suggesting that either or both of the covariance matrices are over-pessimistic. It is earlier stated that the a priori covariance is multiplied by a factor 3, so is this sufficient to explain the position of the peak assuming that the measurement covariance is accurate? A better test would be to evaluate the chi2/m statistic assuming the more realistic a priori covariance and see if the peak is indeed near 1.

>>> To better put the result into a context we rewrote: "However, a peak value near 0.38 indicates that the covariances used are somewhat over-pessimistic. A more detailed analysis shows that both terms of the objective function contribute equally to the observed deviations (distributions peak near 0.17 for the measurements and near 0.15 for the state), i.e. both covariances are affected. Taking into account the more complicated nature of the covariances used in this study we consider the current agreement sufficient."

22) p4574 lines 8-11: the number of iterations required for convergence is mentioned, but there is no mention of what convergence criterion has been used. Obviously the tighter the criterion the more iterations will be required.

>>> In section 3.3 we now describe the convergence criterium used in this analysis: "Convergence is tested for by analysing the scaled step size in state space, 
\[ d_i^2 = (x_{i+1} - x_i)^T S^{-1} (x_{i+1} - x_i) \] 

where \( S^{-1} \) denotes the retrieval covariance matrix (i.e. its iterative estimate) and \( n \) denotes the number of state vector elements. The actual test is carried out by testing
\[ d_i^2/n < 0.1. \] The conclusions about linearity as presented in the paper remain valid.

23) p4574 OPERA: it is said that OPERA deals with average radiances, similar to this model, but doesn’t this just refer to the OPERA inverse model? Is the difference that OPERA performs forward model calculations using spectral integration whereas this is a simpler band model?

>>> We rewrote in section 4.2.1: "Similar to the approach presented in this paper, spectral mean radiances rather than detailed spectra are analysed in the Leicester OPERA scheme (Optimal Estimation Retrieval Algorithm) (Moore, 2006, 2007) in the inverse model. The MIPAS Reference Forward Model (RFM) is utilised as the forward model. In the OPERA scheme, spectrally averaged radiances are directly obtained from line-by-line calculations while the JURASSIC scheme relies on the emissivity growth approximation."

24) p4574/75: it is stated that OPERA results agree within 2-3% but no plots are shown. How many scans were compared? Given the only similarity between the two models is the treatment of the measurements as an average radiance I am surprised that such a degree of similarity is obtained - I would expect at least as much discrepancy between the forward model calculations alone.

>>> We estimated a 1-2% retrieval error (Fig. 5) due to 0.5% forward model errors (Fig. 3). The forward model errors were estimated based on comparisons between JURASSIC EGA calculations and RFM line-by-line calculations. Considering that the inverse model setups were rather similar, the good agreement does not seem implausible. To better explain we rewrote in section 4.2.1: "A direct comparison of about 400 individual scans reveals good agreement (i.e. better than 2 to 3%) between JURASSIC results and OPERA results. The differences found in this comparison exceed the estimated 1 to 2% retrieval errors due to forward model errors (due to the EGA method, compare Fig. 5). However, the rest is likely due to the minor differences in the retrieval setups of the schemes discussed above."
25) p4575 lines 22-24: In the text a systematic differences of +/- 6pptv is converted to a few percent "below 25km". However, with CFC-11 varying rapidly with altitude above 15km (from Fig.8) I would be more convinced if Fig 7 were plotted with a % scale on the x-axis instead of pptv. Similarly with the statement regarding statistical deviations (I assume this means "standard deviations"?) and p4576 line 9 "3.5% or less" systematic difference with the IMK data. For consistency, Figs 4 would also be better plotted with %error as the x-axis.

>>> We follow the advise and add percentage plots to Fig. 4 and Fig. 8. As pointed out by the referee the relative differences must be analysed carefully at upper altitudes where the CFC-11 volume mixing ratio is practically zero. For the comparison with IMK data we restrict the dataset to measurements where the CFC-11 volume mixing ratio exceeds 10ppt (this is less stringent than previously used criteria). We rewrote this paragraph and corrected the percentage values: "Figure 8 shows a global comparison between our CFC-11 retrieval results and the IMK data (version V3O_F-11_8) for several days. The comparison reveals small systematic differences of about -6 to 6 ppt (-9 to 5% below 25 km). Standard deviations are in the range of 8 to 24 ppt (9 to 23%)."

26) p4575 lines 25-29: A simple model of temperature error just assumes that the radiance error is proportional to the change in Planck function, ie of the order of 3%/K for these wavelengths and temperatures. So it is unlikely that a discrepancy larger than 3% can be explained by the difference in temperatures and, in any case, the shape of the temperature differences does not really resemble the mirror image of the shape of the CFC-11 differences. The temperature differences will obviously contribute, but I doubt if it is the major contribution here (again, having Fig 7(a) on a % x-axis would help).

>>> We agree with the referee and removed Fig. 7(a) and the discussion from the text. The paragraph now reads: "A direct comparison indicates good to perfect agreement for many CFC-11 profiles. However, some profiles show large differences. Possible reasons to explain these differences are uncertainties in temperature and pres-
sure, tangent altitudes, abundances of interfering species (ESA operational data versus IMK data), as well as different approaches for the retrieval of radiometric background caused by aerosols and for the regularisation of the retrieval problem. These aspects have to be addressed in future work, but are out of the scope of this paper.

27) p4578 line 10: the 6% difference (15pptv) compared with tropospheric values is within the systematic error budget of a single profile, but this error is largely composed of gain, temperature and pressure errors all of which would be expected to vary pseudo-randomly over the entire MIPAS dataset - these terms would be much reduced when discussing the systematic error budget of the dataset as a whole.

>>> This is certainly true to some extent, but a detailed analysis would require more detailed data about the spatial and spectral correlation lengths of the individual errors source, which are not available (see e.g. comment 12 and 13). One may argue contrariwise that based on the result that the retrieved tropospheric mean and the ground-based value slightly disagree it can be concluded that the systematic errors do not vary quasi-randomly as suggested. We are not aware of a study where this problem was examined for Envisat MIPAS. To be more precise we replaced "systematic errors of the Envisat MIPAS measurements" by "systematic errors of the _individual_ Envisat MIPAS measurements" in this paragraph.

28) p4578 4.5 Comparison with ground measurements: There seems to be a single global tropospheric mean value being used for the comparison. Since one would expect the tropospheric concentrations to vary as a function of latitude, rather than compare the tropospheric values retrieved from individual latitude bands with this global average, it would make more sense to construct the equivalent global mean value from the retrieved dataset as a whole and just compare this number.

>>> We agree and carried out the analysis as suggested. We rewrote in this paragraph: "The global mean tropospheric volume mixing ratio derived from the Envisat MIPAS measurements is (243 +/- 17) ppt, which is about 12 to 14 ppt (5%) lower than the
global mean from the ground based measurements. However, this difference remains within the estimated systematic errors of the individual Envisat MIPAS measurements (compare Fig. 5). Please note that this comparison might be biased because the polar troposphere is not regularly covered by our analysis."

29) p4578 The cited Remedios climatology consists of 6 latitude bands and 4 seasons. If the intention is to provide an update for this, why isn’t the data classified in the same way? The 6 month averages for polar conditions seem particularly coarse.

>>> Please see reply to comment 10. However, we agree that an IG2-type climatology will be useful and will prepare such a file for the electronic supplement of the paper.

Technical Corrections

p4564 line 4: French GUIANA

>>> Corrected.

P4564 line 5: Envisat web-site (http://envisat.esa.int/category/index.cfm?fcategoryid=61) gives inclination as 98.55 deg, which is closer to 99 deg than the 98 deg quoted here.

>>> Corrected to 98.55 deg.

p4565, line 12: by "permanent" I think you mean "routine".

>>> Corrected as suggested.

p4568 line 6: "distance of the spectral windows" Since there are only two windows used, it might be clearer if you said "spectral separation between the two windows".

>>> Corrected as suggested.

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 4561, 2008.