f

MINUTES of MEETING

date de la réunion	18/19 Sep. 2007	ref./*réf.* SE-R&D-min-02-10-07\CZ	page/*page* 1/28

Meeting place *lieu de la réunion*	IMK - Karlsruhe	Chairman Secretary	C. Zehner C. Zehner
minute's date *dates de minute*	02 October 2007	participants *participants*	Enclosed Listing
subject/objet	**6th MIPAS Science Team Meeting**	copy/*copie*	G. Kohlhammer, S. Briggs, H. Laur, M. Doherty, E. Herland, YL. Desnos

description/description	**action/action**	**due date/date limite**
Open ACTIONS		
AI1 on A. Dudhia to provide to M. Laurentis a 1 page description on the measurement scenario for the proposed MIPAS operations observing parallel to the terminator.	Univ. Oxford	asap
AI2 on P. Raspollini to check in detail the computational requirements for the usage of a Variance Covariance Matrix ~~o~~for systematic errors in an operational environment**.**	IFAC	asap
AI3ST5 on the MIPAS Science Team to define 8 days during 2007 (equinox as starting point) of orbits sideward viewing over the full orbit by providing following input to M. D'Laurentis: • 20 Azimuth Angles (e.g. twice the same 10) • Optimised Tangent Point Altitudes	MST	asap
AI5ST5 on H. Oelhaf to update the current MIPAS mission planning based on the outcome of this meeting.	IMK	asap

a
ESRIN
Via Galileo Galilei - Casella Postale 64 - 00044 Frascati (Rome)- Italy
Tel. (39) 06 941801 - Fax (39) 06 94180 280

1. Welcome

H. Oelhaf as organizer of this meeting welcomes all participants.

2. Action Items Status of last Meetings

Zehner (ESA)

All open Action Items (AIs) of previous MIPAS Science Team Meetings have been reviewed and still open AIs were included into the listing of this meeting (see above).

3. MIPAS Instrument Status

Fehr (ESA)

The whole Envisat payload is still performing well after the target life-time (5 years anniversary was on 28 Feb. 2007). Different Envisat mission scenarios (changing its orbit after the year 2010) are being investigated right now to extend the mission lifetime beyond the year 2011.

MIPAS instrument errors have decreased by a factor of 10 during the last 12 months and based on this good instrument performance the duty cycle has been increased in April 2007 up to 60 % and since 8 June the MIPAS instrument has been operating with a duty cycle of 80%.

4. MIPAS Data Processing Status

Fehr (ESA)

The baseline for IPF Version 5.0 has been finalised and the start of off-line processing is foreseen by the end of 2007. The MIPAS user community will get access to MIPAS data products early 2008. Re-processing of MIPAS measurements over the full instrument lifetime might be finalised by mid 2008.

Based on the increased MIPAS duty cycle ESA will try to implement again MIPAS Near Real Time data delivery by latest mid of 2008.

MIPAS Science Team Recommendation to ESA:

Recognizing that the operational data processing of MIPAS RR data is now delayed by more than 2.5 years the MIPAS Science Team recommends vigorously that ESA should make any possible effort to provide validated MIPAS level 2 RR data as soon as possible. Many potential users of MIPAS data of the international community are waiting now for a long time and may believe that the MIPAS instrument does not yield valuable data although the duty cycle of MIPAS has been increased recently to 80 % (the quality of the MIPAS data is very good as shown by MIPAS Science Team members). After the great success of ENVISAT hardware this extremely slow progress in preparing operational data processing for MIPAS may damage the high reputation of ESA.

5. Overview on executed MIPAS Instrument Operations during the last half year

M. De'Laurentis (ESA)

Since the last MIPAS Science Team Meeting only 1 campaign (ECOWAR) has been supported by dedicated MIPAS operations during the time period March 5-17 2007. Beside baseline operations (with the emphasis on NOM mode measurements) NLC mode during July, AE mode during September, and 3 times UA mode measurements were executed. Beside nominal calibration measurements 4 times pointing tests with fixed azimuth angle were performed as well.

Detailed MIPAS mission planning has already been implemented by October 15 2007.

6. Review of new MIPAS Modes/Measurements/ and Future Mission Planning

No campaigns have to be supported by the end of March 2008. Based on the new instrument duty cycle and the need for more Upper Atmosphere measurements in order to study temperature and NO (e.g. transport down into the stratosphere from the thermosphere) at high altitudes the following new baseline measurement scenario is being proposed:

3 days NOM - 1 day MA – 1 day UA – 3 days NOM – 2 days off

AE mode measurements should be performed once around Christmas time and calibration measurements should be performed as during the last half year.

7. Proposed Future MIPAS Processor Upgrades

Based on discussions within the MIPAS Quality Working Group P. Raspollini has prepared a document (see Annex B) describing in detail proposed future MIPAS Procesor Upgrades which were reviewed and prioritised by the MIPAS Science Team:

A - Handling of Systematic Errors

A1 - Variance Covariance Matrix of systematic errors (error spectra instead masks):
In principal a very good idea but in case there are problems with the error spectra this might have significant negative impact on the retrieval and is therefore not recommended for immediate implementation. **AI2** on P. Raspollini to check in detail the computational requirements for the usage of Variance Covariance Matrix of systematic errors in an operational environment.

A2 - Diurnally varying input files:
Immediate implementation recommended as NO_2 retrieval (accuracy) will be improved as well as processing speed.

A3 - Correction of NLTE:
Better NLTE correction will certainly increase the retrieval altitude range (especially for T and H_2O) and lead to a better error characterization but the implementation might be complicated. More impact studies are required before an implementation is feasible.

A4- LUTs with CO_2 line mixing correction:
Should be implemented as it will facilitate the usage of more information from the MIPAS spectrum (Q-branch).

B - Change in the Retrieval Approach

B1 - Optimal estimation:

B2 - Regularisation inside iteration loop:

B3 - Extra points in retrieval grid outside the measurement range:
For all three items immediate implementation is not recommended as these are not necessary for the current target species being retrieved within the ESA operational MIPAS processor. More studies shall be performed in order to get more insight into the benefits/differences of these approaches.

C - Multi-target vs. single target retrieval

Should be implemented especially to handle the retrieval of species having reference spectra being affected by other species and in order to increase the flexibility of ESA MIPAS processor.

D - Handling of horizontal inhomogeneities

D1 - Two dimensional retrieval:

In principal the right approach which should be used in the future, but prior to operational implementation further studies should be performed to get more evidence on the improved retrieval in respect to the needed coding effort.

D2 - Modelling of horizontal gradient in one dimensional retrieval:

Should be implemented even it is only relevant for off-line processing as the temperature gradient has to be included using temperatures coming from previous (e.g. NRT) processing.

E - New target species

E1 - New target species retrievable with the current approach:

Supported for implementation especially for species that can already be retrieved with prototype processor (e.g. CFC-11, CFC-12, and $ClONO_2$~~N_2O_5~~)

E2 - Very weak targets:

Not recommended for immediate implementation as the retrieval of weaker species will be more difficult in low resolution mode and more studies will be needed to identify these.

F - Middle Atmosphere mode retrieval

Not recommended for immediate implementation as there are still too many questions open.

G - Cloud height in forward model

More studies recommended prior to operational implementation.

H - Horizontal Averaging Kernels:

Recommended for implementation

I - Upgrade of non-LTE error spectra:

Recommended for implementation

J – To use less apodized spectra:

Usage of less apodized spectra should ease the retrieval of weaker species and should be at least considered in a dedicated study.

K - To retrieve Isotopes:

Not recommended for implementation as this retrieval is still too complicated. Furthermore there should be a distinction between species which are not so weak (like HDO) and other isotopes that are weak (to be taken into account under E2).

8. MIPAS Data Exploitation

A. Dudhia (Univ. Oxford)

Long term trends: The Oxford MIPAS Level 2 Processor has been used to process 1 day per month (zonal means at different pressure levels) using full and reduced resolution MIPAS data covering the time-span July 2002 until July 2007. This allows intercomparing low resolution with high resolution measurement which shows in general quite good agreement (5-10%) for the ESA target species. For species not being retrieved by ESA like CFC-11 there seems to be a trend in the data (going up), which might be explained by using constant CO_2 in the retrieval.

H. Fischer (IMK)

Temperature and Trace gas retrieval using reduced resolution measurements at IMK and intercomparison to MLS data: Intercomparison of UTLS -1 mode with MLS V2.2, V1.5 and ECMWF (for temperature) give following results:

- **T:** good below 40 km w.r.t ECMWF
- MLS is a bit warmer in UTLS (15 km) w.r.t ECMWF and MIPAS
- MLS v2.2 show good agreement with MIPAS; v1.5 not so good
- **H2O:** Overall good agreement of MIPAS with MLS v2.2 and v1.5

~~MLS is higher in UTLS compared to MIPAS~~Difference plot (MIPAS-MLS) shows a layering in UTLS

- **RHI (Relative Humidity over Ice):** MIPAS shows slight super saturation but MLS shows a bit higher over the tropics.
- MIPAS < MLS above 17 km because MLS water vapour values are high w.r.t MIPAS
- MIPAS > MLS below 17 km because MLS temp. values are higher w.r.t MIPAS
- **O3:** Overall good agreement with MLS, more w.r.t new MLS v2.2 than v1.5
- Temperature and O3 plots over Darwin show features of equatorial Kelvin wave

In general there is good agreement between the two measurements sets (MIPAS showing more structures and more sensitivity). ~~Only for the trace gas HNO3 there are differences in the range of 30%.~~

C. Zehner (ESA)

Usage of MIPAS data within ESA projects: Within scientific projects (AO and Category I) projects MIPAS data are just partially being used since the instrument problem during 2004. For exploitation projects dealing with applications MIPAS profiles are only used within 1 project (www.gse-promote.org). Distribution to users of MIPAS low resolution Level 2 data and the restart of MIPAS NRT data delivery would clearly be beneficial do the MIPAS user community.

T. Fehr (ESA)

G-POD: There is an open Category I call on using the ESRIN GRID facilities for scientific purpose:

http://eopi.esa.int/esa/esa?aoid=400&ts=1191325038511&cmd=aodetail

9. Any Other Business

The publication of all MIPAS Validation Papers (under the lead of H. Fischer) is close to its finalisation

Next MIPAS Science Team Meeting: is planned at ESRIN or Salamanca – Mar. 17/18 2008

Agenda:

	Sep 18
09.00-09.10	Welcome/Logistics (H. Oelhaf)
09.10-10.00	Agenda/AIs from last Meetings (C. Zehner)
10.00-10.30	MIPAS Instrument Status (T. Fehr)
10.30-11.00	Coffee Break
11.00-11.30	Status on ESA MIPAS Processors (including reprocessing and data delivery to users so far) (T. Fehr)
11.30-12.00	Overview on executed MIPAS Instrument Operations during the last half year (M. De'Laurentis)
12.00-13.00	Lunch Break
13.00-15.00	Review of MIPAS (as performed so far) Operations and Definition of future Operations Scenario (including campaigns by end 2007) - Future planning – more UA measurements? (H. Oelhaf) - Scientific campaigns (All)
15.00-15.30	Coffee Break
15.30-18.00	Discussion on future Processor Upgrades (All)
	Sep 19
09.30-10.00	Any Scientific results or problems found using the new MIPAS Level 1 data sets (low spectral resolution – all modes)? (All)
10.00-10.30	Coffee Break
10.30-11.00	Ongoing/Planned Activities within ESA on MIPAS Data Exploitation (C. Zehner)

11.00-12.00	Ongoing/Planned Activities within EC and National projects on MIPAS Data Exploitation (All) AOB

List of Participants:

Raspollini, Piera	IFAC
Carlotti, Massimo	Universita Degli Studi di Bologna, Dip. di Chimica Fisica ed Inorganica
Dudhia, Anu	Oxford University, Atmospheric Oceanic & Planetary Physics Clarendon Laboratory
Fischer Herbert; Oelhaf Herman	Institut für Meteorologie und Klimaforschung, Forschungzentrum Karlsruhe
Flaud, Jean-Marie	LISA CNRS/Univ Paris
López-Puertas, Manuel	Instituto de Astrofisica de Andalucia
Claus Zehner, Fehr Thorsten; Marta de Laurentis;	ESA

Annex A:

Open ACTIONS from previous Meetings		status
AI9ST2 ESA to include a listing onto the Uranus ftp server providing an overview of re/processed MIPAS Level 1 and Level 2 data including the version number of the processing software being used.	ESA	closed
AI3ST3 to T. Fehr to check the available quality flags in the MIPAS L1 and L2 and report this back to IMK **changed to:** T. Fehr to make sure that key MIPAS Quality Flags are described in the MIPAS handbook.	ESA	closed
AI3ST4 on M. D'Laurentis on MIPAS emergency planning	ESA	Closed, execution at least within 50 hours after request
AI4ST3 to T. Fehr to check the feasibility to perform AE measurements over 1 full orbit.	ESA	Feasible and included into MIPAS operations planning
AI1ST5 on T. Fehr to provide to C. Zehner the new MIPAS IODD in order to upgrade the BEAT software tool for MIPAS data handling.	ESA	Closed and included into BEAT
AIST52 on M. D'Laurentis to plan during the first weekend in March 2007 sideward LOS calibration measurements.	ESA	closed
AI3ST5 changed to: plan alternating sidewards (fixed azimuth) and NOM mode measurements over a period of 24 hours	MST	closed
AI4ST5 on T. Fehr to check with Bomen if a schedule with either a 5 day or a 10 day period for the calibration measurements could be adopted instead of the present schedule with one week period.	ESA	closed
AI6ST5 on M. D'Laurentis to check the feasibility to calculate MIPAS LOS in respect to the terminator as a function of time (for 4 orbits for the 4 seasons) and to provide this input to the MIPAS Science Team.	ESA	ongoing
AI7ST5 on M. D'Laurentis to plan just after the ECOWAR campaign 1 day of MIPAS rearward operations with fixed azimuth (exact in the orbital plane) in order to characterise a possible roll offset between the MIPAS instrument and the Envisat platform.	ESA	closed
AI8ST5 on C. Zehner to provide the coordinates of an active volcano in Africa to M. D'Laurentis.	ESA	closed
AI9ST5 to M. D'Laurentis to plan a volcanic measurement scenario over 8 days for this case (in interaction with the Science Team).	ESA	ongoing

~~**Open ACTIONS from previous Meeting**~~		~~status~~

~~**Status of all Action Items as resulting from previous MIPAS Science Team Meetings**~~

Annex B:

Discussion document on possible changes in the MIPAS Level 2 operational processor

14/09/2007

Piera Raspollini, with contributions by Bruno Carli, Thorsten Fehr, Marco Ridolfi, Massimo Carlotti, Bianca M. Dinelli, Anu Dudhia, Manuel Lopez-Puertas and John Remedios

Introduction

ENVISAT mission, after the successful achievement of 5 years of measurements, corresponding to its planned originally lifetime, is continuing with all instruments of the payload working. Also MIPAS instrument, after encountering serious problems at the beginning of year 2004, seems now in good shape working at a reduced spectral resolution with respect to the nominal one, but with performances comparable to the ones of the first part of the mission (P. Raspollini & MIPAS QWG, 2007).

A proposal to extend the mission up to 2010-2014 is now under consideration by ESA. The extension of the mission is highly desirable in order to partially fill the gap of atmospheric chemistry observations foreseen before the launch of the GMES instruments..

Given that the present retrieval approach of the operation analysis is based on choices made more than 10 years ago, and considered the results of the MIPAS validation activity which highlighted the presence of instabilities in some retrieved profiles and some bias at low and high altitudes (near the borders of the retrieval range), it is time to consider upgrades in the operational processor and in the auxiliary data that also take advantage of the new computing capabilities.

During the 13th MIPAS Quality Working Group Meeting the following list of possible upgrades in retrieval products has been identified:

- A: Handling of systematic errors
  - A1: VCM of systematic errors (error spectra instead of masks)
  - A2: Diurnally varying input files
  - A3: Correction of NLTE
  - A4: LUTs with CO_2 line mixing correction
- B: Change in the retrieval approach:
  - B1: Optimal estimation
  - B2: Regularisation inside iteration loop
  - B3: Extra points in retrieval grid outside the measurement range
- C: Multi-target vs single target retrieval
- D: Handling of horizontal inhomogeneities
  - D1: Two dimensional retrieval
  - D2: Modelling of horizontal gradient in one dimensional retrieval
- E: New target species
  - E1: New targets retrievable with the current approach
  - E2: Very weak targets
- F: Middle Atm mode retrieval
- G: Cloud height in forward model
- H: Horizontal AK
- I: Upgrade of non-LTE error spectra

In this note, after a short review of the current retrieval approach, all the listed improvements are discussed: in particular, the advantages and disadvantages of each change are highlighted, and both the implementation effort and computing requirements are evaluated.

This Technical Note is meant to provide a basis for a critical discussion aimed at identifying a sound strategy for a positive evolution of the Optimised Retrieval Model (ORM) code.

Summary of current operational MIPAS retrieval code

Before reviewing the possible upgrades that could be implemented in the operational code, a short summary of the retrieval approach currently adopted for the MIPAS operational retrieval is made. This introduction is also useful to define the formalism necessary for the discussion of the proposed new upgrades. A more detailed description of the algorithm implemented in the ORM, that is the scientific prototype of the L2 operational code, is contained in MIPAS ATBD (latest review on July 2007), Ridolfi et al. (2000) and Raspollini et al. (2006).

The current retrieval approach adopted in the ORM is based on the minimization of the χ^2- function.

The χ^2- function is given by the square residuals **n** (equal to the difference between the measurements **y** and the forward model simulations $F(x,b)$), weighted by the variance covariance matrix of the measurements $\boldsymbol{S}_y$:

$$\chi^2 = n^T S_y^{-1} n \quad (1)$$

with $\boldsymbol{n} = \boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x},\boldsymbol{b})$, $F(x,b)$ being the forward model function, that depends on the unknown vector $\boldsymbol{x}$ and a set of instrumental and geophysical parameters $\boldsymbol{b}$ assumed as known.

The solution is found with an iterative procedure that uses the Gauss-Newton method, modified according to the Levenberg-Marquardt criterion.
At each iteration the retrieved solution $\boldsymbol{x}_{iter}$ is given by:

$$x_{iter} = x_{iter-1} + (K_{iter-1}^T S_{y-1} K_{iter-1} + \lambda M)^{-1} K_{iter-1}^T S_{y-1} n_{iter-1}, \quad (2)$$

where $\boldsymbol{K}$ is the Jacobian of the problem at iteration (*iter-1*), i.e. the derivative of the forward model function with respect to the unknown quantities, λ is the Marquardt parameter that, during the retrieval iterations, is increased or decreased depending on whether the χ^2- function increases or decreases, $\boldsymbol{M}$ is a matrix, with the diagonal elements equal to those of the matrix $K^T S_y^{-1} K$ and the others equal to 0.

The solution matrix of the inverse problem or gain matrix is equal to $G = \left(K^T S_y^{-1} K + \lambda M\right)^{-1} K^T S_y^{-1}$, (3)

with $\boldsymbol{K}$ being the Jacobian at the last but one iteration, while the errors associated with the solution of the inversion procedure is characterised by the VCM of $\boldsymbol{x}$ given by:

$$S_x = \mathbf{G}\mathbf{S}_y G^T = \left(K^T S_y^{-1} K + \lambda M\right)^{-1} K^T S_Y^{-1} K \left(K^T S_y^{-1} K + \lambda M\right)^{-1}. \quad (4)$$

The diagonal elements of matrix $\boldsymbol{S}_x$ represent the mapping of measurement noise on the retrieved profile.
The total error affecting the retrieved profile is computed including the contribution of the random error provided by $\boldsymbol{S}_x$ and the forward model errors estimated a-priori. The forward model errors are due to uncertainties in instrument characterization and in input parameters of the forward model, as well as to approximations in the forward model itself.
Errors due to uncertainties in input parameters of the forward model are computed on the basis of the error spectra $\boldsymbol{K}_b(\boldsymbol{b}-\tilde{b})$, with $\boldsymbol{K}_b$ equal to the derivative of the synthetic spectra with respect to the forward model parameter $\boldsymbol{b}$, $\tilde{b}$ is the assumed value of $\boldsymbol{b}$, and are equal to $\boldsymbol{G}\ \boldsymbol{K}_b(\boldsymbol{b}-\tilde{b})$. Errors due to deficienties in the forward model are estimated on the basis of error spectra computed with sensitivity tests performed with reference models in which mathematical and physical approximations are removed at the best of our understanding [see http://www.atm.ox.ac.uk/group/mipas/err/].].
An a-posteriori regularization has been recently added to the retrieval approach, in order to process measurements in the new measurement scenario, characterised by limb-scanning steps smaller than the IFOV of the instrument.
The regularization is implemented a-posteriori, i.e. after convergence has been reached, and it consists in Tikhonov regularisation, (Tikhonov, 1963) with a constraint matrix obtained from the product $\boldsymbol{R}=\boldsymbol{L}_1^T\ \boldsymbol{L}_1$, with $\boldsymbol{L}_1$ equal to the first derivative matrix. The parameter driving the strength of the regularization is determined analytically using the criterion that the differences between the regularized and non-regularized profile should on average be equal to the errors of the regularized profile ('error consistency' method (Ceccherini, 2005; Ceccherini et al., 2007)).

A - Handling of systematic errors

As already said, in the current retrieval approach the χ^2- function is equal to the square residuals divided by the VCM of the measurements.

This implies that points with a small measurement error have a strong weight and hence a large effect in the retrieval, while points with a large measurement error have a small weight and hence a small effect in the retrieval.

The a-priori estimation of the systematic errors is used for selecting the points that are less affected by the forward model errors (selection of microwindow and of masks within each microwindow). The major advantage of this procedure is to limit as much as possible the computing time using only the points with the greatest information content, leading to a faster data processing.

A different approach for handling systematic errors is proposed in A1, consisting in the computation of the χ^2- function using as weight for the residuals not the VCM of the measurements but the VCM of the residuals.

This approach, consisting in representing all the systematic errors with a VCM (i.e. on the hypothesis that their error contribution can be combined quadratically) is valid for errors that have an average equal to 0. This is a reasonable assumption in the case of spectroscopic errors, p-T propagation errors, errors due to interfering species and calibration errors.
However, for errors that contribute with a constant and known bias, the bias introduced in the retrieved profiles should be removed. This is the case for errors due to interfering species that have a diurnal variation, see sub-Section A2, for Non-LTE errors, i.e. errors coming from the assumption of Local Thermodinamic Equilibrium (LTE), discussed in next sub-Section A3, for CO_2 line mixing approximation, that will be discussed in next sub-Section A4.

A1 - Variance Covariance Matrix (VCM) of systematic errors (error spectra instead of masks)
The proposed new approach for handling the forward model errors consists in using the VCM of the residuals $\boldsymbol{S}_T$, given by the combination of the VCM $\mathbf{S}_y$ of the measurements and the VCM $\mathbf{S}_{FM}$ of the forward model, as weight for the residuals for the computation of the χ^2- function:

$$S_T = S_y + S_{FM} \qquad (5)$$

With this approach it is no longer necessary to perform elaborate off-line selections of the points that must be used for the retrieval (masks), and the combination of the measurement noise with the systematic errors is performed by the retrieval, that directly provides the total error.

The calculation of $\mathbf{S}_{FM}$ requires the knowledge of the error spectra ε_k as described in the Section 'Summary of current operational MIPAS retrieval code' and it is given by:

$$S_{FM} = \sum_k \langle \varepsilon_k \varepsilon_k{}^T \rangle \qquad (6)$$

where the notation $\langle\rangle$ denotes the expectation value operator.
The individual sources of systematic errors can be considered fully uncorrelated among themselves. This assumption allows the combination of the different systematic errors with a simple summation of the VCM of their effects on the measurements.

The described approach has been successfully implemented in the retrieval code MARC (Carli et al., 2007) specifically designed to process MARSCHALS measurements.

Advantages

The approach of weighting the square residuals with the complete VCM of the residuals makes the selection of microwindows less critical, and unnecessary the selection of masks within each microwindow.
A second advantage is the full exploitation of the information content of the analysed spectral intervals, part of which is disregarded by the mask selection.
The third advantage is the direct calculation of the total retrieval error that takes into account both the measurements and the forward model errors.
Finally, the last, but not least, advantage is the reduction of the total retrieval error. The total error obtained when the complete VCM of the residuals is taken into account can never be larger than the error obtained when retrieval is performed on selected unmasked points. This results from the fact that the new approach allows to take into account the existing correlations between different spectral points throught the off-diagonal elements of the complete VCM of the residuals.
With this approach, the visual inspection of the residuals evaluated by themselves may be misleading, since different points have been differently weighted in the retrieval. They should instead be evaluated comparing their value with the total error.

Disadvantages

Due to forward model errors, correlations are present among spectral points of different MWs and of different geometries. As a consequence, a full VCM of the residuals S_T matrix must be considered.
The use of the complete VCM slows the retrieval since the inversion of large matrices can be a time consuming operation. Since the number of observations is larger than the number of unknowns, S_T is the largest matrix that must be inverted in the retrieval. However, matrix S_T must be inverted only once, and, in case that NESR do not change significantly from scan to scan, as it is normally the case, this inversion could be performed only for the scans for which NESR changed significantly.
It has to be noticed that currently, the product matrix $\boldsymbol{K}^T \boldsymbol{S}_y^{-1} \boldsymbol{K}$ is performed exploiting the fact that $\boldsymbol{S}_y$ is a block diagonal matrix.

Implementation effort

ADF: inclusion of all residual spectra.
Minor changes in the code needed: interface with error spectra, computation of the complete VCM of the residuals.

Computational requirements

The use of this approach has a strong impact on both memory requirements and computing time, since S_T is the largest complete matrix that must be inverted in the retrieval. However, memory is not a concern in modern computers and the extra time is not a significant fraction of total time.

Anu's comment: Theoretically the best approach but I'm not sure that benefits outweigh the extra effort and CPU time. Many of the more significant systematic errors (eg. spectroscopic database uncertainties, horizontal gradients) are poorly characterised and others, such as vertical smoothing error, are completely ignored so we might think that we have a more accurate retrieval at the end but I would not be convinced.
See also my comments on C.

A2 – Diurnally varying input files

In the current L2 processor the initial guess profiles and the assumed profiles of the interfering species are selected only according to the season and the latitude, and no distinction is performed between day and night conditions.

For example, at the moment we can only model N_2O_5 interference as a diurnal average since the same input file has to be used for both the ascending and descending parts of each orbit.

This source of bias can be removed adapting the processing to allow for diurnally varying input files. The modification consists in adding in the auxiliary files of the climatological profiles the profiles in both day and night conditions for the species that have a diurnal variation. Information on sun elevation contained in the level 1 file can be used by the L2 pre-processor to select the proper profile.

Advantages

The bias due to the assumption of a constant profile for diurnally varying profiles is removed. Furthermore, initial guess of NO2 retrieval could be selected more appropriately, and this could reduce the number of iterations needed to reach convergence.

Disadvantages

None

Implementation effort

The format of the auxiliary data of Initial Guess profiles has to be changed. The pre-processor has to be adapted for the selection of the proper profile making use of the sun elevation data.

Computational requirements

Negligible.

A3 - Correction of NLTE

The non-LTE modelling is not included in the operational retrieval algorithm since this has a high cost in computation time and requires a major effort for modelling and coding all non-LTE processes. The mw selection is made taking into account the non-LTE errors, that are computed by means of error spectra calculated from the difference between synthetic spectra generated with and without non-LTE. Several models are used for the computation of the Non-LTE populations of all the relevant species (including also gases not being retrieved operationally but potentially contaminant, such as CO and NO). A detailed description of the non-LTE models and the species and emitting levels for which non-LTE populations are computed can be found in Clarmann et al.(1998) and in a subsequent update by Lopez-Puertas et al.(2002).

A first order correction to reduce the bias due to the LTE assumption consists in adding the Non-LTE error spectra to the radiances simulated by the forward model in each retrieval iteration loop.

The complete VCM of the residuals should then include only the estimated residual error on error spectra.

This approach requires that different inputs files are read for day-time or night-time conditions.

Advantages

This type of approach is a simple first-order approximation to reduce the error due to the LTE assumption (more ambitious options are considered at point F).

Disadvantages

The bias due to Non-LTE is diurnally variable, and hence, as a further sofistication, it could be desirable to take into account this variability.

In particular, all the error spectra for different atmospheres, during both day and night, are necessary.

Implementation effort

Minor changes in the code are needed. Changes in the ADF2 consist in adding residual spectra that depends on diurnal conditions, latitude (?), season(?).

Computational requirements

Negligible.

Anu's comment: Since non-LTE effects have a strong diurnal signature, before we can implement a 1st order correction the processing will have to be adapted to allow for diurnally varying input files (Improvement A2).
Even if we do improve our handling of non-LTE, I don't think it will have a significant impact on the nominal altitude scanning but could be important if we try to retrieve to higher altitudes from the MA mode measurements.

Manuel's comment: The new proposed approach, would take some major efforts but could be tackle, at least for some gases. In order to customize the non-LTE error for each scan (which we think it is essentially to somehow perform an NLTE retrieval (mentioned here as first-order correction) we would need to parametrize the Tvibs as function of two major parameters: SZA and temperature profile. This should work for several species: CO2 (Tk), H2O, CH4, and NO2. O3 is somehow questionable because of parameterization of O (atomic oxygen). It could also be add the retrieval of CO. For N2O and HNO3 is not necessary and NO is definitely beyond this scope (photochemical processes, NLTE in rotational and spin, thermospheric contribution, etc.).
We have done some tests already for CO2 and it seems to work. The idea would be to develop Tvibs tables and a tool for computing Tvibs from them provided a given Tk profile and solar illumination conditions (sometimes guess profiles for other constituents).
We think this option is viable now. We should mentioned, however, that the workload required is much larger than that now assigned in the QWG contract.

In order to be able to include the NLTE approach we discussed, the ORM has to be modified to be able to compute NLTE radiances. In principle this is not difficult but I do not know how much effort will it need given the ORM structure. Of course it will also requires more CPU time, although I guess not a big deal since the Tvibs will be computed outside.
Another point in this case would be how do we treat the overlapping of bands. In LTE there is no problem since they all have the same excitation temperature, Tk, but in NLTE they do not. On the other hand overlapping is important in the LTE low region so we would then need to devise a transition from the LTE (overlapping) region to the NLTE (non-overlapping) one. Note that in NLTE the LUT of the transmissions (or optical depths?) that Anu calculates can't be used any more, or use a different one for each band.

John's comment: at this stage I am not in favour of correction for non-LTE.

A4 – LUTs with CO_2 line mixing correction.

Currently, CO_2 line mixing correction is not taken into account in the forward model, but the error due to this approximation is taken into account in the microwindow selection. Recently, a new and more accurate model of line-mixing has been finalised (Niro et al., 2006). With this model it is now possible to include effectively line mixing in the forward model computation, instead of considering it as a source of error.
Since cross-section computation is performed with the use of tabulated cross-section look-up tables (LUTs), line mixing modelling could be included in the LUTs. As a consequence, this upgrade does not require a code modification, but only involves the generation of the cross-section LUTs, that is performed off-line.

Advantages

A significant precision improvement for the retrieval of p and T is expected from the inclusion of line mixing in the forward model. Indeed, it has been proved (Niro et al., 2006) that, since the CO_2 Q branches that are more affected by line mixing are very sensitive to pressure and temperature, the capability to model accurately the line-mixing effects, and hence the possibility of including these points for the retrieval, opens the possibility to exploit at best these spectral regions to infer pressure and temperature distributions. The use of these spectral intervals provides a significant improvement (up to 70%) in the precision of the retrieved pressure and temperature profiles, while using a smaller number of observations with respect to the ones that are currently used.

Disadvantages

None.

Implementation effort

This improvement affects only the generation of the LUTs, that is performed off-line, and hence it has no impact on both the code and the format of the ADF2.

Computational requirements

None.

Anu's comment: This will in any case be implemented in the next version of the RFM used to generate the LUTs. However, since line-mixing generally only affects regions of intense absorption - which tend to be the same regions excluded from the retrieval due to their sensitivity to horizontal gradients - I do not expect a significant improvement in performance for a 1D retrieval.

John's comment: I would support use of CO_2 line mixing.

B - Change in the retrieval approach

The current approach adopted in the operational Level 2 processor exploits the fact that the retrieval from MIPAS measurements of the vertical profile of the target species on the grid of the tangent altitudes is a well-posed problem and the use of a-priori information is not needed. The choice of avoiding any a-priori information is supported by the fact that it can introduce a bias on the averaged profile from different scans. Only the engineering information on pointing is used, but for this an independent measurement is available for each scan.
However, the problem can be ill-conditioned with the result that vertically contiguous values of the retrieved profiles often turn out to be anti-correlated (especially when the spacing between contiguous measurements is smaller than the IFOV, as it is the case for measurements after January 2005). As a consequence, the profile shows un-realistic oscillating values. The physical expectation of a smooth profile is injected in the retrieval through the regularization procedure described in the Section 'Summary of current operational MIPAS retrieval code'. Regularisation that is used in our code provides a constraint on the shape of the profile, but not on the value of the profile.
In order to improve the conditioning of the retrieval two different approaches can be used: Optimal Estimation and Tikhonov regularisation inside the iteration loop.
These improvements should, however, be evaluated in the light of the improvements already obtained with the use of a species dependent retrieval grid.

John's comment: I would support improvement of retrieval stability and sensible profiles – (removal of 10^{-10} problems).

B1 Optimal Estimation

Optimal Estimation (Rodgers, 1976) improves the conditioning of the retrieval combining the a-priori information on the target quantities with the information coming from the measurements.
If the a-priori information consists in an estimate $\mathbf{x}_a$ and its variance covariance matrix $\mathbf{S}_a$ of the state vector, our best estimate $\mathbf{x}_{oe}$ of the state vector is equal to the combination of the retrieved vector $\tilde{x}$ with the a-priori information $\mathbf{x}_a$. The combination of these two measurements can be obtained using the formula of the weighted average:

$$x_{oe}=\left[(S_x)^{-1}+(S_a)^{-1}\right]^{-1}\left\{(S_x)^{-1}\left[G\left(y-F(b,\tilde{x})\right)+\tilde{x}\right]+(S_a)^{-1}x_a\right\} \qquad (7)$$

Introducing the explicit expressions of **G** and $\mathbf{S}_x$, we obtain:

$$x_{oe}=\left[K^T\left(S_T\right)^{-1}K+\left(S_a\right)^{-1}\right]^{-1}\left\{K^T\left(S_T\right)^{-1}\left[\left(y-F\left(b,\tilde{x}\right)\right)+\left[K^T\left(S_T\right)^{-1}K\right]\tilde{x}\right]+\left(S_a\right)^{-1}x_a\right\} \qquad (8)$$

Eq. (8) can be used also at each retrieval iteration step, instead of Eq. (2), for deriving the new estimate of the state vector. In this case the retrieval process is faster and more stable (less affected by ill-conditioning).

Advantages

Optimal Estimation fills the gaps of information that may exist in the measurements. The use of Optimal Estimation allows to fit quantities for which measurements do not contain enough information (extra points in the retrieval grid (see Sect. B3), retrieval of species containing small information (see Sect. E2), retrieval of species in altitude ranges where concentration varies significantly from season to season). It could in principle allow to avoid the use of cloud filtering, since the reduced information content of the measurements in presence of a cloud could be handled with the a-priori information on the retrieved profile. However, in practice, also in case of Optimal Estimation cloud filtering is useful in order not to waste time in analysing measurements that do not contain information.

Disadvantages

We have to be very cautious when averaging profiles from different scans obtained with the Optimal Estimation approach, since the mean of many profiles is equal to the constant a-priori. People using retrieved profiles for performing data assimilation do not generally like to use profiles derived with Optimal Estimation.

Implementation effort

The matrix providing the solution must be changed and the a-priori profile and its error profile must be added in input. These modifications are, however, minor.

Computational requirements

In principle Optimal Estimation should make the convergence easier, as a consequence the expected number of iterations could be smaller than in the current approach with a consequent reduction in computing time.

Anu’s comment: This is what we use in Oxford for our retrievals. I don't think it is necessary for the current major species but probably more important for the minor species, especially where some interprofile averaging is required.
If it is simple to implement, we should incorporate it as a relatively simple
first step towards making B3, E, E' and G more feasible.

Marco’s comment: Optimal Estimation by itself does not provide a regularized profile. Tikhonov regularization has to be combined with Optimal Estimation to obtain a smooth profile.

B2 - Stronger Tikhonov regularisation inside iteration loop

An alternative approach to improve conditioning and stability of the problem is given by the use of Tikhonov regularisation inside the iteration loop.

Advantages

This approach provides a constraint on the shape of the profile instead of on the value of the profile. It is dedicated to reduce oscillations in the retrieved profile. It is a first step towards making B3 and G improvements more feasible

Disadvantages

The parameter determining the regularization strenght varies with the atmospheric status. An analytical method to derive it would be desirable.

Implementation effort
It depends on how the regularisation parameter is determined.

Computational requirements
It depends on how the regularisation parameter is determined.

B3 - Extra points in retrieval grid (if either optimal estimation or regularisation inside the iteration loop is used).

In the altitude domain, ORM determines discrete values of the VMR in correspondence of the measured tangent altitudes (or of a subset of them). In the nominal measurement scenario MIPAS measurements extend from 6 to 68 km, but the level 2 retrieval can be limited to a selected molecule dependent altitude range (retrieval range). Above the highest point and below the lowest point of the retrieval range the shape of the profile is assumed to be equal to that of the initial guess profile and the fit just applies a scaling factor that avoids discontinuities and preserves the shape of the initial guess profile. This may introduce an extra error in the retrieved profile which depends on the assumed shape. If the assumed shape of the profile above the highest tangent altitude is different from that in the real atmosphere, then the retrieval tries to compensate for the error in the estimated slant column by attributing a higher or lower value to the retrieved concentration at the highest tangent point. This error can propagate to next tangent points with an amplitude that quickly damps out as the distance from the highest tangent altitude increases.
The profile below the lowest tangent point is observed through the tail of the IFOV on the low altitude side. For this case, differences between the assumed and real shape of the profile lead to an incorrect computation of the IFOV convolution for the lowest tangent altitude and hence to an error in the retrieved concentration for the lowest retrieved point.
Tests performed after the ENVISAT launch proved the criticality of the adopted extrapolation strategy. Moreover the analysis of real data highlighted the need to extend the retrieval range of NO_2 at higher altitudes to account for its increase during the polar night.
In order to overcome the approximations introduced by the scaling of the initial guess profile outside the retrieval range, the strategy adopted for the definition of the retrieval range was changed to include all the points that are useful for improving the overall quality of the retrieval, independently of the quality attained at that specific altitude.
However, the fact that the standard deviation of the validation statistics close to the boundaries of the retrieval range is larger than the predicted random error (Ridolfi et al., 2007) indicates that the extension of the retrieval range to the whole measurement range is not sufficient to remove completely the problems in the retrieved profiles at both high and low altitudes (especially for temperature, H_2O, NO_2, N_2O and O_3).
A possible solution to reduce the error affecting the points near the boundaries of the retrieval range could be to fit additional points outside the altitude range of the measured tangent altitudes. In particular, an additional point could be fitted 1.5 km below the current lowest retrieved altitude, while some additional points could be fitted above the highest retrieved altitude up to 100 km. The number of additional points to be fitted should be species dependent.
In addition to the levels at the top and bottom of the profile, in case that improvement G (see description below) is adopted, it may also be necessary to add retrieved levels at the tropopause and cloud top.

However, this approach is feasible only with a different retrieval approach, in which either regularization with a stronger constraint or optimal estimation combined with regularisation are used inside the iteration loop. In case of regularization, the constraint should be altitude dependent and tunable from input.

Advantages
This approach would avoid the error due to a wrong shape in the initial guess profile outside the retrieval range by exploiting the information contained in the measurements to derive additional points outside the current retrieval range.

Disadvantages
It would require a change in the current retrieval approach, i.e. either Optimal Estimation or a different type of regularization inside the iteration loop should be implemented.

Implementation effort
New definition of the state vector is needed, with consequent increase of the number of derivatives to be performed. Change in the retrieval approach (use of either O.E. or regularization inside iteration loop is needed). Extensive tests are needed for the optimisation of the new approach

Computational requirements
No significant increase in either computing time or memory requirements is expected from the implementation of this option.

John's comment: I would support removal of retrieval instabilities - e.g. by addition of extra levels.

C - Multi-target vs single target
The ORM performs a sequential retrieval of all the target species profiles: first tangent pressure and temperature are retrieved simultaneously, and then the VMR profile of the target species are derived individually in sequence. The reason of this approach is that at the time of the code designing limited computer memory amount was available.

When the retrieval of many target quantities is performed sequentially, the result of each of them is used in the subsequent retrievals as a known input. With this approach the uncertainty of the already retrieved quantities acts as a systematic error source on the subsequent retrievals. The selection of the analysed spectral intervals is meant to minimise this error propagation; however, in many cases it remains signicant and may prevent the success of the retrieval. Moreover, a remarkable effort must be spent in the selection of optimal spectral intervals.
A more accurate approach is to retrieve simultaneously the target species that interfere in a significant way. In this case the cross-talk between different target quantities that contribute to the same spectral feature can be evaluated through the correlation coefficients of the VCM $\mathbf{S}_x$.

The performance of the proposed approach has been assessed comparing results obtained from a multi-target retrieval MTR (Dinelli et al., 2006) with those of a sequential analysis system (ORM) on the same target quantities. The comparison has shown that the simultaneous retrieval of p, T and water VMR does not lead, below the tropopause, to satisfactory results because of the high correlation between p and water VMR. This problem can be solved extending the MTR analysis to a further target whose spectral features decouple the retrieval of pressure and water VMR. It has been proved that ozone is a suitable target for this purpose.

The code has to be made sufficiently flexible to allow by input the choice of how many and which species have to be retrieved simultaneously.

Advantages
In the multi-target approach the uncertainty of the initial guess of the target quantities does not act as a source of systematic errors, the selection of spectral intervals to be used in the analysis is less critical than in the case of sequential retrievals, the information on pressure and temperature is gathered from the spectral features of all target species.
For weak species or species with limited spectroscopic bandwith (such as SF_6) there is a further advantage of a multitarget retrieval. If we retrieve these species in parallel with the interfering species (possibly using the previously retrieved concentrations of these interfering species as an a priori constraint with optimal estimation) this not only allows more spectral points to be used for the weak species (rather than excluding all points with a significant overlap from the interfering species) but also better quantifies the error due to the interference (rather than the current approach of treating these as a fixed systematic errors).

Disadvantages
None
Implementation effort
The number of elements of the state vector increases, as well as the derivatives to be computed, but the modifications to be performed in the code are minor. The modifications to be made are conceptually simple, a bit more elaborate from the practical point of view.

Computational requirements

Larger memory requirements are needed. Furthermore, since the number of elements of the state vector increases, the matrix G to be computed requires an increased computing time, but this is compensated by a reduction of total spectral points to be used compared to the case when the sequential retrieval of different species is performed.

John and Marco's comment: for multi-target retrievals - this sounds good and is correct in principle but we need to be sure that we are improving the data.

D Handling of horizontal inhomogeneities

The approach currently adopted in the ORM assumes that the atmospheric portion sounded by a limb scanning sequence is horizontally uniform. The vertical profile for each target species is retrieved analysing a single scan. The errors in the retrieved profiles caused by the approximation of horizontally uniform atmosphere have been estimated to be one of the most significant source of systematic errors at low altitudes, especially for p, T and O_3
[see http://www.atm.ox.ac.uk/group/mipas/err/ and Stiller et al., (2002)]
].

D1 - Two dimensional retrieval

An approach that allows to avoid the approximation of horizontal inhomogeneities is currently adopted in the GEOFIT program (Carlotti et al., 2001), that performs a simultaneous retrieval of all scans of one orbit and accounts for the horizontal variability of the atmosphere taking advantage of the partial superimposition of the portion of the atmosphere sounded by each scan.

Advantages

Two dimensional retrieval allows to avoid the errors due to the assumption of horizontal uniform atmosphere; this feature is mostly valuable in the UT/LS region.

The migration to a 2-D strategy would represent a desirable evolution towards a new generation of retrieval algorithms. The development of a 2-D algorithm may also include the most part of the upgrades considered in this document.

Disadvantages

It requires a major change in the structure of the retrieval code and also memory requirements increases significantly.

Implementation effort

An almost rewrite of the code is required

Computational requirements

Larger memory requirements are needed. From the computing time point of view, almost comparable run times are obtained with the GMTR and ORM on a whole orbit.

Anu's comment: probably too ambitious since a complete rewrite of the code is required.

John's comment: at this stage I am not in favour of 2-d retrievals operationally (Unless they can be shown to be superior and stable).

D2 - Modelling of horizontal gradient in one dimensional retrieval

Simulations have shown that horizontal gradients in temperature are the largest source of error among all gradients (Carli et al., 1998, Stiller et al., 2002). However, it has been proved (Carli et al., 2007) that the retrieval of the vertical profile of the gradients from a single limb sequence (one dimensional retrieval) is characterised by a very large error.

A simplified approach to handle horizontal inhomogeneities would be to simulate in the forward model a vertical profile of the horizontal gradient of temperature (and, possibly, of the other target species). For temperature, H_2O and O_3, the information on the vertical distribution of the gradient could be provided by ECMWF fields, for the other target species information could be taken from corresponding retrieved profiles of the previous orbits.
In this approach, the total error computation should then take into account the error on the assumed gradients.

Advantages

This type of approach could allow to reduce in a way easier than the one described in option D1 the error due to the assumption of horizontally uniform atmosphere.

Disadvantages

This approach is based on the external knowledge of the atmospheric gradient. The knowledge of T gradients from ECMWF data should be accurate enough.

Implementation effort

It is a major change and most of the optimisations that have been implemented in the forward model could no longer be applied (secant law approximation, equality of the paths along the line of sight simmetrical with respect to the tangent point). As a possible alternative, see point H.

Computational requirements

This would lead to an increase in computing time.

Anu's comment: I don't think it's as difficult task as suggested. The major impact of horizontal gradients is the variation of the Planck Function along the line-of-sight so it is just the variation of B along the path that needs to be considered as part of the radiance integration. All the main assumptions about spherical symmetry can probably still be used: ray-tracing, pressure, density, absorber amounts.

E - New target species

The list of molecules retrievable from MIPAS is significantly longer than the list of target species of the operational retrieval.We discuss the feasibility of processing other species with either current retrieval approach or with a different retrieval approach

E1 - New targets retrievable with the current approach

The species that could be retrieved sequentially without changing the retrieval approach are the following ones: CF11, CF12, N_2O_5, $ClONO_2$. The retrieval of these species with ORM has been already proven, at least for the full resolution measurements (Lahoz et al., 2007). Non-standard species produced by IFAC have been assimilated by BASCOE v3q33 together with MIPAS ESA off-line standard species (i.e., ten species assimilated simultaneously) for 15 days. From this analysis it results that including $ClONO_2$ and N_2O_5 as constraining species in the system adds information on the NO_2 content. Furthermore it is highlighted that the retrieval of these three species is consistent.

The retrieval of these species in sequence after the retrieval of the current target species requires only minor changes in the code but requires significant changes in the auxiliary data and in the output of the level 2 file. Of course this change has also an impact on computing time.

Advantages

Operational provision of additional species from MIPAS measurements.

Disadvantages

Increase in computing time.

Implementation effort
Duplication of the module that performs the single target retrieval for each new target species.
Duplication of the auxiliary data corresponding to the new target species.

Computational requirements
Any new target species requires additional computation resources.

Anu's comment: CFC-11 and CFC-12 are probably the highest priorities, especially in terms of a long-term mission and being relatively easy to retrieve with the current set-up. $ClONO_2$ is more variable so may pose problems if the concentrations tend to zero. N_2O_5 is the most difficult of those mentioned. For our own retrievals of this we've had to use a single continuum fit across the complete N2O5 feature rather than a separate continuum fit for each microwindow (I believe IMK do the same). I don't know how feasible this is with the current processor.
As well as new molecules we might also consider isotopomers and aerosol-type products.

E2 - Retrieval of very weak targets

The retrieval of species with a weak spectroscopic signal, (the indication on what species would be of interest should come from the Science team) requires a special strategy to be studied and hence a change in the retrieval approach. Indeed, when very little information is contained in the observations, the retrieval errors can be so large that the linear approximation may be not valid any longer. In order to reduce the retrieval error two approaches are possible: either to constrain the inversion in order to stay in the linear regime (by means of Optimal Estimation) or to perform the analysis on a redundant number of spectral points.

If Optimal Estimation is used, the sequential retrieval of standard and weak targets can be performed. For standard targets usual retrieval is performed. For weak targets the Optimal Estimation is used and the climatological profile with its VCM is used as a-priori information at the first analysed limb sequence, while the profile retrieved at the previous sequence with its VCM is used subsequently [Kalman Filter]
The final output contains the average profile of the weak species for the full set of limb sequences.
Using Optimal Estimation mode, the VCM associated with the retrieval is averaged with the VCM of previous retrieval with the effect of reducing the error about fitting parameters. Also the values of fitting parameters are averaged with values of parameters computed in previous retrieval (this average is weighted taking into account the VCM). This approach requires a different calculation of the AK.

If Optimal Estimation is not used, the retrieval of weak species can be done using a redundant number of spectral points. Assuming that the VMR of standard and weak species, as well as temperature, do not vary significantly along a set of sequences, all the spectra measured in N consecutive sequences can be used simultaneously to retrieve a single profile. (Marco's comment: this approach is equivalent to perform a retrieval as described above and then remove the a-priori contribution.)

Three possibilities exist for the retrieval of a single profile from spectra corresponding to the N considered sequences:

a) All limb scanning spectra of N sequences are included among the observations of the single retrieval. This seems possible without any change in the code, but we have to consider that the forward model inside ORM assumes that there is an altitude grid point for each simulated tangent altitude. The simultaneous inclusion of all spectra of N sequences has as a consequence that the atmospheric layering is made of numerous and almost overlapping grid points.
b) For a suitable subset of the grid points (reduced altitude grid) determined in the pressure domain, the mean of the spectra are derived. If the reduced altitude grid has a sufficiently fine spacing, the FOV is broadened by the averaging process in a negligible way. The computation of the mean spectra at the reduced altitude grid is not an easy operation, because different scans can observe a different atmospheric continuum.
c) A third possibility (proposed by Anu) could be to model everything in the target molecule microwindows apart from the target molecule itself, average these residual spectra (ie measured-modelled) and then perform a retrieval on the mean residual spectrum assuming average profiles of p,T and other interfering species.

For all the approaches, we assume that single scan retrieval of the target species is performed for all scans, and the average of the retrieved profiles is used as representative of the mean state of the atmosphere assumed in the retrieval of the weak species.

Extensive tests are needed for the definition of the most appropriate approach to retrieve weak species.

Advantages
Operational retrieval of additional species.

Disadvantages

A change in the retrieval approach is needed.

Implementation effort
Three different approaches have been proposed. Kalman filter is the approach that should require less substantial changes.

Computational requirements

Any new retrieved target species requires additional computation resources.

Anu's comment: The proposed retrieval (b) from averaged spectra assumes a linear relationship between radiance and all the variables. The temperature and tangent point pressure variations, in particular, may cause problems.
The optimal estimation approach avoids this by effectively averaging in the state space of the retrieved product rather than in radiance space, but does require the retrieval to be recast as an optimal estimation problem, and preferably retrieving on fixed pressure levels rather than tangent points.
The proposed c) option represents an intermediate solution.

John's comment: retrieval of more species would be very good - I think the original list of extra species should be revised to include species which are most interesting for people. I think CFCs will be proposed throughout the mission by Juelich so perhaps there is no need to do it operationally. ClONO2, N2O5, and weak species such as ClO, C2H6 might be most interesting. It is important for some species that we have the full processing of MIPAS data.

Bruno's comment: the variable atmospheric continuum is a problem for all three proposed possibilities a), b) and c). However, in the case of c) it is possible to consider the retrieval of the continuum (instead of modelling the other atmospheric contribution) making sure that this operation is made in a spectral interval where the weak species signal has been masked out.

F - Middle Atmosphere mode retrieval
The Middle Atmosphere (MA) observation mode covers most part of the stratosphere, the mesosphere and the lower thermosphere (from 18 to 102 km) with 29 steps at a constant 3 km step size. The along-track sampling is about 430 km. This observation mode is dedicated to study linkages between the upper atmosphere and the stratosphere, i.e. the global circulation and transport from CO and NOx from the mesosphere down to the stratosphere in polar winter hemispheres, as well as Solar Proton Events affecting both the upper atmosphere and the stratosphere. The mode is also used to monitor the quality of operational retrievals that are neglecting non-LTE effects.
The feasibility of using the ORM for analysing the MA measurements is discussed here.
As a first step the ORM could be used for retrieveing profiles in the range 18-70 km from these measurements. Above 70 km the combination of decreasing density and temperature leads to a rapid fall-off in S/N hence a much larger number of microwindows are required. Non-LTE effects also become more widespread, limiting the lines which can be used, although we could restrict ourselves to nighttime conditions when these are generally smaller.
Oxford team is currently running its own retrievals of MA mode data up to 90km using the standard microwindow approach, but these microwindows were only selected for polar-winter conditions so there may be strong non-LTE effects in other parts of the orbit. This approach is also very slow given the large number (~10) microwindows required for each molecule.

Other possible approaches to process MA measurements with ORM could be:

a) option described in A3;
b) Non-LTE calculations are included in the forward model using calculated non-LTE vibrational temperatures;
c) retrieval of some Non-LTE parameters is performed.

Implementation effort
It depends on the selected choice. See also discussion of improvement A3.

Computational requirements
It depends on the selected choice. See also discussion of improvement A3.

Anus' comment:
A PhD student, Luis Millan-Valle, is investigating a completely different approach based on a linear (as opposed to iterative non-linear) retrieval using the complete spectrum rather than microwindows but it is too early to say whether this will work.

Manuel's comment: we think options b) and c) in F are now far beyond of what we could do because it would requires major changes to the code as well as a much larger CPU cost.

G - Cloud height in FM

. The current approach for handling clouds consists in identifying and flagging clouds with high opacity so that retrieval of trace gas concentration is performed above the cloud top height.
Residuals clouds in the FOV of the instrument are suspected to be cause of some bias at low altitudes, especially for HNO_3, O_3 and pT retrieval (it has to be noticed that the threshold used for cloud filtering is significantly less conservative than the one used in IMK code.

The main difficulty that prevents the proper simulation of clouds in our forward model comes from the fact that even if the current continuum retrieval should allow for small amounts of cloud contamination, there is the fundamental problem that this can only represent an extinction profile on a 1.5 or 3 km grid, whereas most types of cloud would be better represented in the forward model as having a step-like function where the extinction changes from zero to infinity at a particular height.
As part of a separate study Oxford team has been investigating the possibility of retrieving cloud top height (as well as extinction) from MIPAS spectra and this could provide a better representation of cloud in the forward model used for the L2 retrievals. However it does require the forward model to allow an extra level (or closely-spaced pair of levels) to be inserted at an arbitrary altitude within the retrieval grid to represent the cloud top.

Advantages
The approach of changing the forward model to better represent non completely opaque clouds would allow to process a larger number of measurements and would reduce the bias in the retrieved profiles.

Disadvantages
The external information on the cloud top height (and the corresponding value of extinction) is necessary (or the algorithm for determining these quantities should be incorporated in the L2 pre-processor) for each scan. A change in the retrieval approach is needed (O.E. and/or stronger regularization inside the iteration loop).

Implementation effort
Change in the ADF2 needed. Changes in the algorithm consists in adding 1 or 2 points in the retrieval grid in correspondence of the cloud top height, where both continuum and VMR are fitted.

Computational requirements
Minor

John's comment: We can establish new cloud filtering threshold, that is not a problem. The REAL problem is to answer the question "why do retrievals not work better in the presence of real cloud?" Would this be a good study? We can simulate spectra of clouds reasonably for this purpose.
Then we can try to improve the operational retrieval to get the right answer.

We should really put some more effort into testing the retrieval in cloudy and near cloudy situations. For gases like H2O and O3, the only way to do this really is by simulation.
Harjinder's work shows that there are problems with the O3 retrieval in clear-sky near the tropopause in any case. Clouds then cause further problems.
So I think a 3-step approach:

1) Test the performance of the current retrieval in cloudy situations by running a statistically significant number of retrieval simulations (against simulated spectra including clouds)
2) Revise the retrieval approach to improve retrieval stability (remove 1e-10 tendencies) and performance near the tropopause or PSCs.
3) Re-test the new retrieval approach on the same simulated cloudy spectra

We can then implement the new improved processor with a better cloud specification.
An alternative is to try to also retrieve some cloud properties to improve the performance.
Finally it is worth noting that this work could benefit from work on clouds going in the MIPClouds project.

H - Horizontal Averaging Kernel (AK)

During the MIPAS validation activity it emerged that the horizontal smoothing error can be a non-negligible source of discrepancy between MIPAS and measurements from other instruments (see Ridolfi et al., 2007, Cortesi et al., 2007).
As well as the vertical AK, that is currently computed on line by IPF V5.0 for each scan, also the horizontal AK could be very useful to characterise MIPAS products and to compare them with other measurements characterised by different horizontal resolution.
However, the horizontal AK has never been computed by ORM team. Two approaches have been followed so far to estimate the error due to the horizontal smoothing performed by MIPAS.
The first one has been proposed by T. von Clarmann (von Clarmann, 2006). It indicates an approximated method to compute the horizontal AK $\mathbf{A}_h$ from the vertical AK.
For each target species the element *i, j* of $\mathbf{A}_h$ represents the response of the i^{th} grid point to an infinitesimal variation of the target species at the location *j* along the MIPAS line of sight. The matrix $\mathbf{A}_h$ can be derived using the following approximation: each entry of the row of the MIPAS vertical averaging kernel is assigned to the geolocation where, according to ray tracing, 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) can be given to air parcels at the same altitude in front of and behind the tangent point. Averaging kernel elements below the tangent altitude are assigned to the tangent point geolocation.

The second approach to compute the horizontal smoothing error, for instance for temperature, has been proposed by De Clerq (De Clerq and Lambert, 2006).
The ECMWF temperature field is 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.
Algebraically, the horizontal smoothing error $\delta_{h,i}$ of a given MIPAS profile level *i* has been estimated as:

$$\delta_{h,i} = |\nabla_{r,i} T_{ECMWF}| r_{90,i}$$

where $|\nabla_{r,i} T_{ECMWF}|$ 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_2 radiance from which temperature 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).

A more precise approach to compute the horizontal AK is proposed here. This approach consists in computing numerically for some reference standard atmospheres the two-dimensional AK (i.e. horizontal and vertical at the same time), i.e. the derivative of the retrieved profile with respect to the true profile as a function of height and latitude.
The calculation of the two-dimensional AK, that should be performed off-line, needs a forward model that is able to handle horizontal inhomogeneities, like the one included in the GMTR algorithm (Carlotti et al., 2005). Starting from a non-uniform atmosphere, it simulates a sequence of spectra at the tangent altitudes corresponding to the nominal limb scanning sequence. Then spectra relative to the complete limb scanning sequence are simulated as many times as the number of points of the considered bi-dimensional atmospheric field spanning the area sounded by a limb scanning measurement with only one point perturbed each time.
A vertical profile must be retrieved by ORM for each of these perturbed limb scanning sequences and be compared with the retrieved profile obtained in the case of no perturbation. The ratio between the difference of the 2 retrieved profiles and the value of the perturbation provides the value of the bi-dimensional AK for all the points of the altitude grid in correspondence of the given altitude and latitude that has been perturbed.

Advantages

The avaibility of the horizontal AK is very important to characterise the retrieved profiles.

Disadvantages

The computation of the two-dimensional AKs requires a major work, but can be done off-line.
The proposed method allows to compute numerically the two-dimensional AK for some defined standard atmospheres. From the experience gained during the validation of MIPAS measurements, we know that the AKs computed for standard atmosphere may be not representative of all the studied cases. However, it is certainly a useful tool to estimate the size of the horizontal smoothing error. For this reason, the suggestion is not to change the current approach of computing on-line the vertical AK for each scan, but to compute numerically also the bi-dimensional AKs for some defined standard atmospheres.

Implementation effort

This computation is performed off-line, therefore it has no impact on both the code and the ADF2.

Computational requirements

None

I - Upgrades of non-LTE error spectra

Manuel's contribution: It is an appropriate time to upgrade the non-LTE vibrational temperatures. About 5 years have been elapsed and we could update the Tvib with all the studies we have done. To upgrade them, first we would have to compute the new Tvibs, then Anu would have to compute the forward NLTE spectra and, I guess, to redo the MW selections based on the new NLTE errors. So, even in the first case of not changing the NLTE strategy, this will take some time.

This implies a complete re-generation of the auxiliary data, but it has impact neither on the format of the ADFs nor on the retrieval algorithm.

REFERENCES

Carli, Dudhia et al., Final Report of the Study of the retrieval of atmospheric trace gas profiles from infrared spectra, ESA Contract 12055-96-NL-CN, 1998

Carli et al., High level algorithm definition and physical and mathematical Optimisations (MIPAS Level 2 Algorithm Theoretical Baseline Document, IFAC_GA_2007_12_SC, latest review: Issue 4.1, July 2007.

Carli et al., 'MARC: a code for the retrieval of atmospheric parameters from millimeter-wave limb measurements, Journal of Quantitative Spectroscopy and Radiative Transfer,105, 476-491, 2007.

Carlotti et al., 'Geo-fit approach to the analysis of limb-scannning satellite measurements', Appl.Opt., 40, 1872-1885, 2001.

Ceccherini, 'Analytical determination of the regularization parameter in the retrieval of atmospheric vertical profiles', Opt. Lett., 30, 2554-2556, 2005.

Ceccherini et al., 'Technical Note: Regularization performances with the error consistency method in the case of retrieved atmospheric profiles', ACP, 7, 1435-1440, 2007

Clarmann, v. T., Dudhia, A., Echle, G., Flaud, J.-M., Harrold, C., Kerridge, B., Koutoulaki, K., Linden, A., Lopez-Puertas, M., Lopez-Valverde, M. A., Martin-Torres, F. J., Reburn, J., Remedios, J., Rodgers, C. D., Siddans, R., Wells, R. J., and Zaragoza, G.: Study on the simulation of atmospheric infrared spectra, Tech. Rep., Final Report of ESA Contract 12054/96/NL/CN, European Space Agency, ESTEC, Noorwijk, The Netherlands, 1998.

Clarmann, Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, ACP, 6, 4311-4320, 2006.

Cortesi et al., Geophysical validation of MIPAS-ENVISAT operational ozone data, ACPD, 7, 5805-5939, 2007.

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.

Dinelli et al., 'Multi-target retrieval (MTR):the simultaneous retrieval of pressure,temperature and volume mixing ratio profiles from limb-scanning atmospheric measurements', Journal of Quantitative Spectroscopy & Radiative Transfer, 84, 141-157, 2004

Lahoz et al., 'The assimilation of ENVISAT data (ASSET) project',ACP, 7, 1773-1796, 2007.

López-Puertas, M., Funke, B., García-Comas, M., López-Valverde, M. A., Martín-Torres, J., Advanced MIPAS Level 2 Data Analysis (AMIL2DA), Project EVG1-CT-1999-00015, Report on the Climatology of Vibrational Temperatures, 2002.

F. Niro, G. Brizzi, M. Carlotti, E. Papandrea, M.Ridolfi, 'Precision improvements in the geo-fitt retrieval of pressure and temperature from MIPAS limb observations by modeling CO2 line-mixing, JQRST 103 (2007), 14-26.

Raspollini et al., 'MIPAS level 2 operational analysis', ACP, 6, 5605-5630, 2006

Raspollini for MIPAS QWG, Overview of MIPAS operational products', ENVISAT symposium, Montreux, April 2007.

Ridolfi et al., 'Optimized forward model and retrieval scheme for MIPAS near-real-time data processing, Appl. Opt., 39, 1323-1340, 2000

Ridolfi et al., Geophysical validation of temperature retrieved by the ESA processor from MIPAS/ENVISAT atmospheric limb-emission measurements, ACPD,7, 5439-5513, 2007.

Stiller et al., "Sensitivity of trace gas abundances retrievals from infrared limb emission spectra to simplifying approximations in radiative transfer modelling", JQSRT, 72, 249-280 (2002).