Technical Note on:

New functionalities implemented in ORM_ABC_1.2.3 and cloud detection algorithm

DRAFT

Delivery of the CCN7 of the study:

"Development of an Optimised Algorithm for Routine p, T and VMR Retrieval from MIPAS Limb Emission Spectra"

Contract No: **11717/95/NL/CN**

Prepared by:

Marco Ridolfi	University of Bologna

Checked by:

Piera Raspollini	IROE - CNR

Bologna, 16 November 2001

TABLE OF CONTENTS

Document Change Record

Issue	Revision	Date	Change description
Draft	-	16 Nov. 01	First draft of document, does not include cloud detection algorithm for which specifications have not been issued yet by the Leicester University team.

1. Reference Documents

#	Document	Issue	Title
RD1		1	M.Carlotti and L.Magnani, Tech. Note on "Results of WP 8300 and WP 8500 of the CCN5 of the contract 11717/95/NL/CN", (Issue 1 from 2 July 2001 and update from 14 August 2001)

2. Introduction

MIPAS (Michelson Interferometer for Passive Atmospheric Sounding) is an ESA developed instrument to be operated on board ENVISAT-1 as part of the first Polar Orbit Earth Observation Mission program (POEM-1). MIPAS will perform limb sounding observations of the atmospheric emission spectrum in the middle infrared region. Concentration profiles of numerous trace gases can be derived from MIPAS observed spectra.

According to the current baseline, from MIPAS measurements altitude profiles of atmospheric pressure and temperature (p,T), and of volume mixing ratio (VMR) of six species (O_3, H_2O, HNO_3, CH_4, N_2O and NO_2) will be routinely retrieved in near real time (NRT). The retrieval of these parameters from calibrated spectra (Level 1b data) is indicated as NRT Level 2 processing.

Level 2 processing is a critical part of the Payload Data Segment (PDS) because of both the long computing time that may be required and the need for a validated algorithm capable of producing accurate and reliable results.

In the frame of the on-going study "Development of an Optimized Algorithm for Routine p, T and VMR Retrievals from MIPAS Limb Emission Spectra" a scientific code has been developed for NRT Level 2 analysis, suitable for implementation in ENVISAT PDS and optimized for the requirements of speed and accuracy. The scientific code is called Optimized Retrieval Model (ORM) and is limited to the most critical elements of the MIPAS Level 2 algorithm, namely the p, T and the VMR retrieval components. It is the result of the primary study objective, to optimize the overall retrieval scheme with respect to the initial baseline, by introducing various mathematical and physical optimizations and by verifying the performance improvements on the basis of simulated test scenarii.

Also, a set of auxiliary data has been identified which allow to configure and operate the retrieval algorithm when integrated in the framework of the MIPAS Level 2 Processing Facility (MPF2).

An outcome of the previous study phase were a set of documents which define the functional elements of the p, T and VMR retrieval components, as well as the formats of the various input and output data. Based on this information a prototype code was developed in a parallel industrial study. This prototype is fully representative of the overall Level 2 algorithm as it is going to be implemented in MPF2. Furthermore it will incorporate a number of modifications with respect to the ORM s/w package. It will serve as a reference platform for use during the validation and verification of the MPF2 and shall allow the assessment of the overall ground processor performance under realistic conditions at an early stage.

Due to the postponement of ENVISAT launch date (launch currently scheduled for January 29th, 2002) it will be possible to implement in the ORM and / or in the MIPAS Level 2 prototype code a few new functionalities that are thought to improve the reliability and the robustness of the processor itself. Namely, the functionalities that are being considered are:

1. Introduction of a pT - H_2O loop in the retrieval chain in order to overcome the problem of lack of convergence in case of assumed or initial guess water profiles "too far" from truth.

2. Improvement of the algorithm robustness in case of large tangent altitude corrections.
3. Cloud-detection algorithm to filter-out measurements affected by clouds before Level 2 processing.

3. Objective

Objective of the present document is to provide a description of the new functionalities that have been implemented in the ORM_ABC or, as in the case of the cloud detection algorithm, that will be part of the framework of MIPAS Level 2 processor. The provided details should be sufficient for industry (ASTRIUM) to implement the considered functionalities in the Level 2 prototype code.

4. Lack of convergence in pT and / or H_2O retrievals, pT - H_2O loop

In some of the pT and water test retrievals (see RD1), the convergence criteria are not fulfilled after the maximum number of allowed (micro or macro) iterations. In the specific test cases analyzed, the lack of convergence is due to the very large differences between initial guess / assumed and real profiles (this occurs especially for the H_2O profile in the case of incorrect assumption of hygropause altitude).
In particular, in case of significantly incorrect initial guess H_2O profile it happens that:

- pT retrieval converges to pT values suffering of a large systematic error.
- The erroneous pT values prevent H_2O from converging

This behavior is due to the fact that pT microwindows have been selected assuming too small uncertainties for the H_2O profile. Namely, the assumed H_2O uncertainty does not allow for errors on the hygropause height which, on its own, may introduce errors as large as two orders of magnitude in the evaluation of the H_2O VMR profile.
Although this problem will be tackled by using more realistic estimates of the H_2O uncertainties for microwindow and occupation matrix selection, an option was implemented in the ORM in order to re-iterate pT and H_2O retrievals in case of lack of convergence in one of the two retrievals. In the following section we describe the details of this implementation.

4.1 ORM modifications for pT + H_2O loop

The following modifications have been applied to the ORM in order to implement the pT + H_2O loop:

1. Both retr_pt and retr_vmr write now the (pT / H_2O) profiles into the *_dump.dat files in all the following occasions:
   - when the convergence has been reached,
   - when the max. n. of macro-iterations has been reached,
   - when the max. n. of micro-iterations has been reached.

   See modified retr_pt and retr_vmr modules.
2. retr_pt (as retr_vmr is was already doing in the former ORM versions) can use as initial guess / assumed profiles the profiles recorded in pt_dump.dat and in gas_dump.dat. The usage of profiles stored in the "dump" files is commanded via switches that are recorded in the settings_pt.dat input file. In total, three switches are used: *lifptret* and *lifvmret* are equivalent to the corresponding switches used in VMR retrievals already in the former ORM releases. An additional switch (*lupdzt*) has been introduced in settings_pt.dat which allows to start pT retrieval from the engineering values of tangent altitudes while still using the last retrieved temperature. The interpolation / extrapolation operations of the "dump" profiles in pT retrieval

are done in a new FORTRAN module named "read_dump.inc" which is included (using the "include" statement) in the retr_pt module.

3. Both retr_pt and retr_vmr have now an argument that returns the convergence status at the end of the retrieval. Since now the retrieval is considered successful only if the convergence criteria are fulfilled, the retrieval convergence status is contained in the variable *lconverg* that is output of the convchk_* module (in the former ORM versions the retrieval was considered unsuccessful only when the max. number of micro-iterations was reached).

4. The main calling tree, in orm was modified as follows:
   a) Remove "dump" files (this step is necessary only in the ORM_ABC code that still handles a single scan).
   b) Read variable *orm_maxloop* from the environment (max. n. of pT + H_2O cycles)
   c) Begin loop *j* = 1, *maxloop*
      call retr_pt (*lconverg_pt*)(*)
      copy pt_dump.dat in the INP_FILES directory
      call retr_vmr(1, *lconverg_vmr*)(*)
      copy h2o_dump.dat in the INP_FILES directory
      if (*lconverg_pt* . AND . *lconverg_vmr*) EXIT from loop
    end loop
   d) If (.NOT. *lconverg_pt*) skip to the analysis of next scan, otherwise proceed with retrieval of O_3 from the current scan.

Remarks:
1. With the implemented approach the ORM writes profiles into "dump" files even in case of lack of convergence. These profiles are then used as initial guess / assumed profiles in the subsequent retrievals. This is consistent with the assumption that at each retrieval iteration the current estimate of the unknown profile represents always an improvement compared to the initial guess, even if the convergence has not been reached.
   In the ML2PP the retrieved profiles are weighted (optimal estimation) with the a-priori profiles before usage in the subsequent retrievals. In this optimal estimation procedure, the low reliability of the profiles retrieved without convergence should be reflected by a covariance matrix multiplied by the final the chi-test obtained in the retrieval.
   In general the covariance matrix scaled by the actual chi-test represents a better estimate of the error associated with the retrieved profiles independently from the convergence status. We therefore <u>recommend to use always the retrieval covariance scaled by the chi-test</u> for proper weighting of the retrieved profiles against a-priori profiles in the MIPAS Level 2 pre-processor.
2. We have tested the pT + H_2O loop against the "difficult cases" encountered in the full orbit analysis made in the frame of WP 8300. From these tests it turns-out that the pT + H_2O loop strategy is effective even if pT retrieval is repeated using as initial guess the engineering tangent altitudes and not the "last retrieved" tangent altitudes (i.e. the in the tested cases the approach is valid also if *lupdzt* = FALSE). Therefore the switch *lupdzt* could be permanently maintained FALSE (or not implemented at all in the ML2PP). This should overcome the inconsistency that may be encountered while reporting in Level 2 products tangent altitude corrections obtained by repeating pT retrieval with the "last retrieved" tangent altitudes.

(*) pT and H_2O retrievals use as IG / assumed profiles the pT and H_2O profiles contained in the most recent *_dump files whenever available. If no *_dump.dat files are available climatological estimates are used instead.

5. Improvement of ORM robustness in case of "large" tangent altitude corrections

5.1 Improvement of the grid used for profiles representation

A recognized weakness of the ORM up to version ABC_1.22.1 is embedded in the approach used for constructing the pressure and altitude grids in which the profiles are represented. In the ORM the tangent height (Z_1) of the lowest sweep of the scan is assumed as fixed, while pressure (P_1) and temperature (T_1) relating to the lowest tangent point are retrieved quantities. In the ORM versions up to ABC_1.22.1, the lowest point of profiles (*rvmrbase(ibase,igas)*, *rtbase(ibase)*, etc.) was set in correspondence of some pressure level (P_0) included in the so called *prof* profiles (initial guess or assumed profiles, this point is set in the routine chbase_pt). At each pT retrieval iteration the altitude Z_0 relating to P_0 is obtained (in updprof_pt) by means of hydrostatic equilibrium, assuming the altitude Z_1. Now, if P_1 approaches "too much" P_0 the profiles extension below the lowest tangent point (the difference (Z_1- Z_0)) could not be sufficient to simulate the FOV convolution and cause a crash (being the FOV pattern a geometrical property of the instrument, its width does not depend neither on pressure nor on temperature).

The ORM_ABC_1.2.3 checks the difference (Z_1- Z_0) and, if it is less than half of the FOV width, the point Z_0 is substituted by a new point $Z_0' = Z_1$ - *rintup* (*rintup* being the width of the FOV). The value of pressure corresponding to Z_0' is calculated using hydrostatic equilibrium, while temperature, VMR and continuum are calculated by extrapolation from the original *'base'* profiles. This operation is done in the modified module "updprof_pt.f" (see source code).

5.2 Constraining tangent pressures

Assuming the engineering LOS data to be a reliable source of information on the pointing, the retrieved tangent altitudes can not differ from their engineering estimate by an amount greater than the maximum estimated error on engineering data. Since the pointing-related parameters in pT retrieval are pressures, we applied a constraint to tangent pressures. In particular, in the routine retr_pt, after the call to newparest_pt (that generates a new estimate of the vector of the unknowns) we check the differences between the predicted and the initial guess (engineering) tangent pressures. If (for an individual sweep) this difference exceeds a user-defined threshold (variable *rmxpvar* read from settings_pt.dat), the predicted tangent pressure is set equal to its engineering value.

Modified ORM modules:

- retr_pt. Added check of tangent pressures as explained above, after call to newparest_pt routine.
- finput_pt. Added reading of new variable *rmxpvar*.

Remark

In order to choose an appropriate value for the threshold *rmxpvar* we have argued as follows:

- the max. absolute pointing error (on engineering data) can be assumed to be of the order of 1.5 km
- pressure roughly changes by 14 % every km of altitude (in a standard, average atmosphere)
- therefore initial guess and retrieved pressures should not differ by more than 21 %.

6. Cloud detection algorithm

7. Correction of recently discovered bugs

Routines "tcgeo_pt" and "gcgeo_vmr": while calculating igeocder(j,3) the loop on simulated geometries must be stopped at *jgeo=igeo-1* and not at *jgeo=igeo*, as in the former ORM versions. Furthermore it ha been set *igeocder(igeo,3)=igeocder(igeo,2)*.