ABSTRACT

A fast prototype processor for cloud parameter from MIPAS-Envisat (MIPclouds) measurements has been developed in the framework of an ESA funded study. Within the MIPclouds study we developed and improved retrievals of cloud parameters like cloud top height, temperature and extinction, as well as for microphysical parameter, e.g. effective radius and integrated quantities over the limb path like area density and volume density. We developed a classification algorithm for polar stratospheric clouds (PSC) types and the differentiation between liquid and ice cloud in the upper troposphere. The sensitivity in detection of different clouds is similar to space and ground based lidars, but the high cloud amount (<440 hPa) is on global scales significantly larger compared to passive nadir viewers like (A)ATSR, AIRS, SEVIRI or the ISCCP dataset.

1. INTRODUCTION

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on the ENVISAT satellite measures limb IR spectra in the 4 to 15 micron range [1]. The MIPAS cloud spectra contain a variety of crucial information about atmospheric processes such as cloud formation or chemical interaction of clouds and trace gases. The exploration of the scientific extremely valuable cloud spectra – measured globally with very high spectral resolution – has been just started. First publications already demonstrate the great scientific impact these data and specific algorithms can achieve [e.g. 2, 3, 4, and 5]. For example, IR limb measurements from space are extremely sensitive to the detection of optically thin clouds in the UTLS region [6], often denoted with subvisible or ultra-thin cirrus (SVC, UTTC). The physics behind these clouds, their impact on the radiation budget or the water entrance into the stratosphere are not well understood and quantified. In addition, MIPAS is the first instrument which allows the compilation of a pole covering climatology on the occurrence of polar stratospheric clouds (PSC) at daytime and nighttime conditions. On the other hand recent analyses show that the impact of cloud spectra on retrieval of p/T and trace gas profiles can be tremendously important and proper cloud-screening is necessary to avoid erroneous retrieval results [7,8]. The ESA operational level 1 and 2 products include no information on clouds and aerosols. This is not a surprising fact, because the analysis and retrieval of cloud parameter from limb IR spectra is still challenging task due to the complex radiative transport in the presence of clouds.

So far, no validated and consolidated MIPAS cloud product is available for the scientific community. Therefore the development of a cloud processor providing standardised and validated product parameters seems extremely valuable. The synergy between potential cloud parameter like cloud extinction, temperature, location, type classification, and microphysical information like effective radius (Reff) or ice water content (IWC) in combination with the operational level 1b and 2 products of MIPAS or other sensors would allow for example the analysis of processes in the Upper Troposphere and Lower Stratosphere (UTLS), e.g. with respect to cloud formation and water vapour transport, or the formation and chemical imprint of PSC in the polar regions in much more detail than currently feasible.

The principal objectives of the MIPclouds study are to:
1. explore the capabilities to retrieve cloud parameters from the MIPAS level 1b spectra
2. develop a time efficient cloud prototype processor (less than 1 hour processing time per orbit) for reuse in the scientific community and operational NRT applications
3. to perform the geophysical validation of the processor retrieval products

The instrument has measured nearly continuously from September 2002 to March 2004 in the so-called full resolution (FR) mode (0.025 cm⁻¹) of the spectrometer. These measurements are taken in the primary time period of interest for the application of the newly
developed prototype processor. However, special emphasis was also taken to keep the algorithms flexible in a way, that only minor modifications will be necessary in order to allow for the processing of ongoing measurements in the optimised resolution (OR) mode (0.0625 cm\(^{-1}\)).

2. STRUCTURE OF THE CLOUD PROCESSOR

The retrievability of potential cloud parameters has been investigated in a feasibility study of the MIPclouds project and is summarised in [13]. Figure 1 gives an overview of the processor and the corresponding scientific analysis. The outcome of the scientific analysis is closely linked with potential retrieval parameter of the processor. The Cloud Scenario Database (CSDB) was crucial for the algorithm development and the error assessment of the parameters and is described in more detail in section 3.

![Flowchart MIPclouds Processing](image)

**Figure 1. Overview of the MIPclouds processing scheme in interaction with the scientific analysis and validation activities**

In a first step various data sources (L1B, L2, ECMWF data) are merged based on the calibrated L1B spectra. Pre-processing takes place to create a consolidated profile based dataset of radiances for a number of specified micro windows. The cloud detection is then a crucial part of the processor where various methods like different colour ratios (CR) can be applied and add-on information like the influence of high water vapour continuum is retrieved for the best possible decision which spectra of a profile are influenced by cloud emissions. After this step the retrieval of cloud macroscopic cloud properties (cloud extinction, top height, and temperature) starts. Afterwards a cloud classification takes place based on the top two cloudy tangent heights. In the free troposphere (>5km) and UTLS region the classification works for liquid and cirrus clouds, and for the winter polar Stratosphere for the PSC types ice, nitric acid trihydrate (NAT), and liquid sulfuric ternary solutions (STS). Based on the classification, additional microphysical parameter can be estimated like effective radius (Reff), limb volume or area density path (ADP), volume density, or ice water content. A list of processor output parameters is summarised in Table 1.

At each processing step various input parameters and look-up tables are necessary. Various output parameters are created and stored in the data product files in netCDF format.

3. CLOUD SCENARIO DATABASE

A major part of the study was the compilation of a comprehensive database containing modeled MIPAS radiance measurements in the presence of various cloud types and related Jacobians with respect to cloud microphysical parameters and interfering variables. Currently, the database contains more than 70000 different cloud scenarios and about 600000 cloud spectra. To our knowledge, this is the first time that such extensive simulations have been performed for mid-IR limb-emission sounding of clouds.

Tests to limit the extent of spectral windows required for the cloud analyses have been performed by calculating broadband Jacobians with respect to aerosol extinction either with or without gas-contribution. The quotient of these Jacobians is to first order equal to the gas-transmission spectra at tangent altitude. The rational of these simulations is to exclude regions of the spectrum with strong interference of trace gases and of already opaque spectral intervals. In addition, constrains like already selected regions of existing algorithm have been taken into account and resulted in the following optimised list of window regions, in total a range of 137 cm\(^{-1}\), for the database: 782-841, 940-965, 1224-1235, 1246-1250, 1404-1412, 1929-1935, 1972-1985, 2001-2006, and 2140-2146 cm\(^{-1}\). An example spectrum for an optically thin cirrus cloud with a cloud top height at 17.5, cloud thickness of 4 km at 12 km tangent height is given in Fig. 2. Obviously line emissions are still visible in the spectrum overlaid with a moderate continuum offset due to the cloud.

All CSDB spectra have been generated with the KOPRA model [9], which takes single scattering into account [10]. KOPRA results have been compared with multiple scattering calculation with the FM2D/SHDOM [11]. The conclusions from these simulations are that (for details see [13]):

(a) differences between FM2D and KOPRA are as expected, with multiple scattering being important for intermediate cloud optical thickness, but good agreement for fully optically thin and fully optically thick clouds;
(b) radiance differences introduced by multiple scattering are generally within the range of other uncertainties affecting cloud radiative transfer (about 2-5% at 950 cm\(^{-1}\) but significant larger 20-50% differences at 2000 cm\(^{-1}\)).
Figure 2: Example spectrum of the CSDB for optically thin cirrus (black) observed at 12 km tangent height with a cloud top height at 17.5 km compared to a cloud-free spectrum in the selected wave number intervals. Radiance units in \( \text{nW/(cm}^2\text{sr cm}^{-1}\)).

(c) consideration should be given to the effect of solar scattering (for day-time retrievals) in the short-wave.

4. ALGORITHMS AND TECHNIQUES

Here we will give just a brief summary of the applied algorithms in the prototype processor, which are described in detail in the algorithm technical basis document (ATBD) [15] and the references cited below.

Detection: A number of complementary methods are implemented in the processor, the multi colour-ratio (Cloud Index: CI) approach in various wavelength regions with improved threshold definition similar to [12], a new singular value decomposition (SVD) approach for cloud detection [16], a new multi wavelength micro windows (10) at 930-960 cm\(^{-1}\) method, and finally a weighted combination of the cloud detection flag of each method for the determination of a detection confidence. For each of the detection methods a cloud top temperature (CTT) and cloud top pressure (CTP) is retrieved, where CTP and CTT are simply the corresponding ECMWF pressure and temperature at the CTH location.

Macro physical parameter: An optimal estimation retrieval for CTH, CTT and extinction (CEX) based on a simple continuum fit has been developed. The cloud effective fraction in the FOV (3km) is considered for the CTH.

Micro physical parameter: Fit coefficients of a CSDB based regression analysis are implemented in the processor for the parameter area density path (ADP) – using the operational CI of MIPAS band A as a proxy, and for an estimate of the effective radius (Reff) – using the brightness temperatures at three wavelengths (827, 941, and 1227 cm\(^{-1}\)).

Classification: Multi colour ratios [2,5,15], brightness temperature differences (BTD) and a naive Bayes classification [17] using the CSDB as a training dataset have been implemented in the processor. The latter method is selecting the best suitable micro window pairs (up to 10) by optimising the product probability of the probability density distribution of all potential BTDs for 1 cm\(^{-1}\) broad MWs. The following cloud types can be retrieved: (a) Polar Stratospheric Cloud types: ice (ICE), Nitric Acid Trihydrate (NAT) and sulphuric ternary solution droplets (STS), (b) cirrus (cir) and liquid (liq) clouds in the free troposphere and upper troposphere and lower stratosphere (ULTS).

5. VALIDATION ACTIVITIES

The validation of the retrievable products is a prime task for a consolidated data product. Various spaceborne, airborne (SAGE, GLAS lidar, SEVIRI, ATSR) and ground based measurements (several NH and SH lidar stations) have been investigated with respect to potential coincidences with the MIPAS instrument for the time period September 2002 to March 2004. Three types of validation methods can be applied to the processor output parameters, depending on the dataset or parameter of interest:

(1) The validation on the basis of coincident measurements in a certain miss-time and miss-distance window (abbreviated with CM),

(2) The statistical comparison of parameters based on temporal and spatial means (SM),

(3) Blind test retrievals with a sufficient number of modelled input radiance profiles with well characterised cloud parameters (BTR).
Table 1: MIPAS cloud processor output parameters with short cut comments for SM: statistical means (zonal or seasonal-means for various altitude bins), CM: coincidence method for the validation dataset, BTR: blind test retrievals. For details on validation method see the following sections. (*) For transformation from limb path or slant column quantities like IWP to nadir IWP or respectively IWC one has to assume the simplification of a homogeneous cloud layer filling the complete tangent height layer.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>comments on validation method, errors, etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>cloudiness flag</td>
<td>cf</td>
<td>status flag for each spectrum 1/0</td>
</tr>
<tr>
<td>Cloud occurrence frequencies</td>
<td>COF</td>
<td>statistical means: SM (e.g. zonally, seasonally)</td>
</tr>
<tr>
<td>Cloud Top Height</td>
<td>CTH</td>
<td>SM, BTR</td>
</tr>
<tr>
<td>Cloud Top Temperature</td>
<td>CTT</td>
<td>SM, BTR</td>
</tr>
<tr>
<td>Cloud Top Pressure</td>
<td>CTP</td>
<td>SM, BTR</td>
</tr>
<tr>
<td>Cloud Base Height</td>
<td>CBH</td>
<td>SM, BTR</td>
</tr>
<tr>
<td>Cloud Extinction</td>
<td>CEX</td>
<td>SM, CM</td>
</tr>
<tr>
<td>Cloud Classification: Stratospheric Cloud Types:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Nitric Acid Trihydrate</td>
<td>NAT</td>
<td>CM, BTR for all types</td>
</tr>
<tr>
<td>- Sulfuric Ternary Sol.</td>
<td>STS</td>
<td>CM, BTR for all types</td>
</tr>
<tr>
<td>- Ice</td>
<td>ICE</td>
<td></td>
</tr>
<tr>
<td>Cloud types in the free troposphere / UTLS:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Cirrus / Ice clouds</td>
<td>CirC</td>
<td></td>
</tr>
<tr>
<td>- Liquid</td>
<td>LiqC</td>
<td></td>
</tr>
<tr>
<td>Area Density Path</td>
<td>ADP</td>
<td>BTR; along the limb path; for threshold values and estimates only;</td>
</tr>
<tr>
<td>(*) limb Ice/Liquid Water Path</td>
<td>IWP /</td>
<td>BTR, (CM); quantities along the limb path</td>
</tr>
<tr>
<td>or Volume Density Path</td>
<td>LWP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDP</td>
<td></td>
</tr>
<tr>
<td>(*) limb Ice / Liquid Water Content or Volume Density</td>
<td>IWC /</td>
<td>BTR, SM, (CM); only estimates available</td>
</tr>
<tr>
<td></td>
<td>LWC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VD</td>
<td></td>
</tr>
<tr>
<td>Effective and/or mean radius</td>
<td>R_{eff} / R_{mean}</td>
<td>BTR, SM, (CM); only course size bin retrieval</td>
</tr>
</tbody>
</table>

The latter validation approach was undertaken to produce a dataset of simulated cloudy radiances. These were designed as an input to the retrieval code, to allow the output parameters to be evaluated under a regime where the full characteristic of the cloud field is known. The radiative transport simulations are as such realistic as possible and are using 2-D radiative transport with multiple scattering calculations, and utilising cloud fields selected from ECWMF analysis data [11]. It is thus feasible to analyse the impact of the finite horizontal extent of clouds, their potential displacement far from the tangent point along the line of sight, the spatial variation of cloud parameters within the clouds, and multiple clouds within the atmosphere which all are scattering into the line of sight.

A more general note on the complexity of the validation of cloud parameter demands critical attention. The validation of cloud parameters is a more difficult task than for many other atmospheric parameters. The spatial and temporal scales where substantial changes in the parameter of interest can occur are very variable and can be quite small in respect to the characteristic sampling of the instrument. For these conditions the viewing geometry and field of view effects become very important when comparing measurements of two different instrument types. At the same time it is not a trivial task to consider the full complexity in a validation approach. Simplifications with respect to the viewing geometry and compromises in the selection criterions (miss distance and time) are necessary to achieve a certain number of potential coincidences.

6. FIRST VALIDATION RESULTS

The following examples are concentrating on the validation results with respect to detection sensitivity by the comparison with various sensors and are based on version 1.6 of the processor. More details on other processor parameters can be found in the product validation report [18].

6.1 Comparison of CTH with the GLAS lidar

One of the few techniques to obtain a continuous daily global coverage of subvisible/ultra-thin cirrus clouds and PSCs are space-borne lidars. During the first phase of MIPAS observations (June 2002 until March 2004) the Geoscience Laser Altimeter System (GLAS) instrument on ICESat [21] recorded continuous data between 25 September and 18 November 2003.

In Fig. 3 we show a summary of the mean differences between the CTHs of the various MIPclouds detection methods and the Lidar coincidences for 200 km and 3 h miss-distance and time respectively. The maximum GLAS CTH is used (GLAS_CTH_max). Numbers on the left inside the plots indicate the number of co-incident sample pairs which fall into the respective altitude. For better visibility, the differences are binned within altitude bands of 3 km height. Numbers on the left inside the plots indicate the number of co-incident sample pairs which fall into the respective altitude range. The following discussion will concentrate on those altitude bins with a sufficient number of samples (e.g. >20).
Figure 3: Altitude-binned differences of cloud top heights between the various MIPcloud database CTH entries and the lidar co-incidences (200 km/3 h).

**PSCs:** A typical feature at high southern latitudes (-90° -70° and -70° - -50°) is a relatively high positive bias of many MIPAS CTH indices (exceptions are CTH_OPER_CI, CTH_CIOPT_CI, and CTH_SVD). Such high clouds above 13 km in the south are very probably PSCs. Inspecting corresponding correlation plots shows that only few Lidar samples indicate higher CTHs than MIPAS, but most are lower. This might be explained by (1) possibly a smaller sensitivity to optically thin polar stratospheric clouds of the lidar compared to MIPAS or (2) by a large inhomogeneity of the PSC field at the end of the polar winter. At that time, due to the preceding denitrification of the stratosphere, PSCs are optically thinner and are not as homogeneous as during the period from June-August.

**Tropical cirrus:** Compared to the PSCs, the CTHs of high cirrus clouds in the tropics at 11.5-17.5 km altitude compare better between the lidar and MIPAS. At these altitudes, in the latitude range -10° - 10° the mean CTH differences are around 1 km.

**Tropospheric mid- and high latitude clouds:** MIPAS and Lidar CTHs of tropospheric clouds at mid and high latitudes at around 6-11 km agree well, mostly within differences of 1-2 km. This agreement is slightly better for northern than for southern latitudes.

### 6.2 Comparison cloud occurrence with SAGE II

The analysis for subvisible cirrus (SVC) cloud occurrence in the SAGE data is described in detail in [19] and references therein. Here we have applied the algorithm to the time period 12/1998 to 11/2004 on a monthly and seasonally mean basis. SVC clouds are detected by a two wavelength extinction technique at 0.5 and 1 micron. Typically, for a solar occultation instrument with only ~30 limb profiles a day, it is necessary to compute a long time mean for an adequate count statistic.

The MIPAS prototype processor results have been analysed in respect of cloud occurrence in a similar way like the SAGE data. The CTHs were used for the computation of cloud occurrence frequencies \( f_c \) (COF) by the following expression:

\[
f_c(i) = \frac{N_c(i)}{N(i) - \sum_{i=1}^{n} N_c(i)} \times 100 \ \text{[\%]}
\]

where \( N(i) \) indicates the total number of MIPAS overpasses in a defined latitude–longitude grid box at altitude level \( z(i) \), with \( i = 1 \ldots n \) from the bottom to the top level of the analysis. \( N_c(i) \) is defined by the number of cloud events \( N_c(i) \) (detected CTHs) in the layer \( i \). By the definition of the CTH it is not unambiguous to detect clouds underneath the first detected cloud layer in a height profile. Therefore, in the dominator \( N(i) \) is diminished by all cloud observations in the layers above. A 1 km \( \times 10° \times 20° \) grid for altitude, latitude and longitude was defined and the parameter MIPAS SUM_CLOUD is usually used for the COF analyses. The parameter is a combination of various detection methods but in the current version of the processor (V1.6) dominated by the macro retrieval CTH information (RTV_CTH_MARCO). If the macro result is not available then weighted information of a selected number of detection methods is applied.
The comparison of the resulting cloud occurrences are illustrated for 15 km altitude in Fig. 4 for the Dec/Jan/Feb winter season (DJF). The peak COF-values of MIPAS are significantly larger than the SAGE ones. Differences up to 20-40% are obvious.

In general, MIPAS showed slightly higher cloud occurrence rates around the tropopause than SAGE II. This is partially an effect of the coarse field of view and sampling of MIPAS (3 km for both parameters) compared to the better resolved SAGE measurements (1 and 0.5 km respectively), which creates a slight high bias in the detected CTHs and systematic differences when comparing occurrence frequencies at a fixed altitude but the result is also an indicator for a slightly higher cloud sensitivity of MIPAS than SAGE, an effect by the new implementation of a combination of multiple detection method in the processor.

### 6.3 Comparisons of high cloud amount with GEWEX

The MIPAS cloud products have been compared with a variety of cloud climatologies, e.g. the data sets prepared for the GEWEX cloud assessment. The GEWEX cloud assessment group (http://climserv.ipsl.polytechnique.fr/gewexca/presentation.html) was initiated by the GEWEX Radiation Panel (GRP) in 2005 to evaluate the reliability of available, global, long-term cloud data products, with a special emphasis on the International Satellite Cloud Climatology Project (ISCCP) [20].

Fig. 5 shows global maps of the high cloud amounts (altitudes above a pressure level of 440 hPa) for the different nadir viewing instruments at the resolution of the MIPAS level 3 product. The ATSR instruments show marginally more sensitivity to high cloud than the ISCCP data set, while the AIRS data set is the most sensitive. In addition the ATSR level 2 cloud data set is collocated in time with MIPAS and was used to generate an ‘MIPAS like’ product. This new product is more representative of a limb viewing instrument that will assign cloud over a large horizontal footprint. The ‘MIPAS like’ product was generated by the following steps: (1) Averaging the ATSR orbit information onto a 1° longitude by 2° latitude grid. (2) If the high cloud amount (cloud > 440hPa) was greater than 0 then the cloud fraction of the grid box was set to one. If a single observation is detected in the larger footprint it is assumed that the MIPAS instrument would have detected it. (3) The data was then projected onto the MIPAS monthly grid. The resulting product has been labelled ‘ATSRMIPAS’ and is also shown in Fig. 5 together with the MIPAS cloud occurrence information for March 2004.
Clearly the limb viewing instrument is more sensitive to thin cloud because of the longer path through the atmosphere than of nadir instrument. The high cloud amount (hca) for this ‘MIPAS like’ product is still less than the MIPAS product as would be expected due to differences in sensitivity and the limb view. However, this product shows strong spatial correlations especially in the tropics with the MIPAS product.

7. SUMMARY AND OUTLOOK

A new cloud parameter processor for MIPAS with near-real-time capability has been developed. The processor implements various cloud detection and classification methods as well as micro and macro-physical parameter retrievals. The validation of selected cloud parameters shows partially excellent results (e.g. for cloud detection, CTH and the climatology of cloud occurrence). Some parameter like the area density path (ADP), cloud top temperature (CTT) and the classification polar stratospheric clouds (PSC) types (not presented) show the capability for innovative new research objectives like the understanding of formation processes and the spatial and temporal distribution of optically thin cirrus and polar stratospheric clouds.

Various validation techniques have been applied. The analyses showed that it is sometimes difficult to compare parameters of complementary measurement techniques (e.g. nadir and limb). Validation comparisons need to consider differences in sensitivity, viewing geometries, FOV or the vertically and horizontally sampling of two sensors otherwise results may be misleading. A blind test retrieval approach has been used to bypass this problem.

The overall detection sensitivity of the processor is similar to space and ground based lidars. The MIPAS results for high cloud amount (<440 hPa) show on global scales significantly higher values compared to passive nadir viewers in the GEWEX dataset.

The processing of the full MIAPS time series (>8 years) would create a unique and complementary data series of cloud parameters (e.g. compared to products of nadir viewer) for climate related studies in respect to cloud processes. Comparisons with models with incorporated cloud physics – like the ECMWF, chemical transport and general circulation models – are a major issue for future applications. The MIPclouds data can be used to validate the model capabilities to predict correctly the cirrus distribution and coverage as well as the water transport in the UTLS region. In addition, new topics will come into the scope of the MIPclouds processor, for example, further improvements and developments are desirable for the classification of various particle types in the troposphere and lower stratosphere like the severe differentiation of various aerosol types (e.g. volcanic ash) from liquid and ice water clouds.

Acknowledgment:
This work was supported by the European Space Agency under the contract no. 20601/07/I-OL. R.S. likes to thank P.H. Wang, NASA, LaRC, for the
preparation of the SAGE II data and the algorithm for SVC detection.

8. REFERENCES


