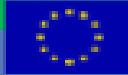




The European Commission

Community Research

**Energy, Environment  
and Sustainable Development**



Project title: **Energy- Specific Solar Radiation Data  
from Meteosat Second Generation (MSG):  
The Heliosat- 3 Project**

Project Acronym: **HELIOSAT- 3**

Contract No.: **NNK5- CT- 2000- 00322**

Project No.: **NNE5- 2000- 00413**

Deliverable Title: **Atmospheric parameter retrieval  
Including  
Cloud processing scheme  
Water vapour retrieval scheme  
Aerosol parameterisation scheme  
Ozone processing scheme**

Deliverable No.: **D9.1, D9.2, D9.3, D9.4**

Main Author: **DLR**

Contractual Date  
of Delivery: **31. 05. 2004**

Actual Date  
of Delivery: **31. 05. 2004 (electronic form by e- mail)  
31. 05. 2004 (hardcopy mailed)**



### Document Signature Table

Responsible authors:

M. Schroedter- Homscheidt, DLR

Additional authors:

L. Bugliaro, DLR

T. Erbertseder, DLR

G. Gesell, DLR

T. Holzer- Popp, DLR

### Document Change Record

<b>Issue Revision</b>	<b>/</b>	<b>Date</b>	<b>Changed Pages/Sections</b>	<b>Comment on changes</b>
Issue 1.0		31- May- 2004	all	Initial issue



**Contents**

1.Introduction .....7  
 1.1.Objectives of the HELIOSAT-3 project.....7  
 1.2.Aim and structure of the report .....7  
 2.Retrieval of cloud mask and cloud optical properties.....9  
 2.1.Introduction .....9  
 2.2.Cloud retrieval scheme.....10  
 2.2.1.Technical changes.....10  
 2.2.2.Physical changes.....11  
 2.3.Cloud parameter file format.....12  
 2.4.Example results.....16  
 2.5.Further modifications for operational processing.....23  
 2.6.References.....24  
 3.Retrieval of water vapour .....26  
 3.1.Introduction .....26  
 3.2.Water vapour retrieval development (level 2).....26  
 3.3.Realisation in the current HELIOSAT-3 software.....34  
 3.3.1.Input data .....34  
 3.3.2.Parameter settings in level 2 processing.....35  
 3.3.3.Generation of level 3 water vapour product.....36  
 3.4.Water vapour level 3 file format.....37  
 3.5.Example results.....39  
 3.6.Further modifications for operational processing.....40  
 3.7.References.....41  
 4.Parameterisation of aerosol amount and type.....41  
 4.1.Introduction .....41  
 4.2.Aerosol parameterisation scheme.....42  
 4.2.1.Current status of aerosol climatology data set.....42  
 4.2.2.Further development on the SYNAER method during HELIOSAT-3.....48  
 4.3.Aerosol file format.....50  
 4.4.References.....51  
 5.Retrieval of ozone.....52  
 5.1.Introduction .....52  
 5.2.Retrieval of Total Column Ozone.....52  
 5.2.1.GOME.....52  
 5.2.2.SCIAMACHY .....53

- 5.2.3.TOMS.....54
- 5.3.Level 3 products .....54
  - 5.3.1.GOME and TOMS.....55
  - 5.3.2.SCIAMACHY .....56
  - 5.3.3.Gridding to HELIOAT-3 format .....58
  - 5.3.4.Near- Real- Time Processing chains.....59
- 5.4.References.....61
- 6.Acronyms .....62

# 1.Introduction

## 1.1.Objectives of the HELIOSAT-3 project

A successful integration of solar energy technologies into the existing energy structure highly depends on a detailed knowledge of the solar resource. HELIOSAT-3 will supply high quality solar radiation data gained from the exploitation of existing Earth observation technologies and will take advantage of the enhanced capabilities of the new Meteosat Second Generation (MSG) meteorological satellites. The expected quality represents a substantial improvement with respect to the available methods and will better match the needs of the users of the resulting products. In particular, HELIOSAT-3 will provide:

Solar irradiance data with high accuracy and space-time resolution necessary for solar energy applications, plus a large geographical coverage

Additional solar energy specific data (direct and spectral irradiance, angular distribution of diffuse irradiance, spatial structure of irradiance) according to the needs of end- users

Information on HELIOSAT-3 products, its sustainability as a service and its potential benefits to end- users.

The objectives of HELIOSAT-3 will be achieved by the development and establishment of a new type of solar irradiance calculation schemes based on physical models of radiative transfer. This scheme will be based on satellite retrieved atmospheric data instead of ground based climatologies.

## 1.2.Aim and structure of the report

This report describes the retrieval of atmospheric data from satellites as applied in the current version of the new HELIOSAT-3 software.

The structure of this report is related to the description of project work which is part of the contract. As a consequence each work package has its own chapter as part of deliverable D9:

WP2020 Retrieval of cloud mask and cloud optical properties (deliverable D9.1)

WP2030 Retrieval of water vapour (deliverable D9.2)

WP2040 Parameterisation of aerosol amount and type (deliverable D9.3)

WP2050 Retrieval of ozone (deliverable D9.4)

Each section includes a description of the retrieval method, example results, and comments on requirements for an operational processing. Wherever refereed publications in easily available scientific journals exist, only a citation will be given. Heliosat-3 is mainly project which combines existing remote sensing technologies for a new application in solar energy. New development has only been made in order to fill gaps in existing knowledge. Therefore, we report mainly on these adaptations and changes instead of copying material that has been published partly years before.

An extended validation will only be given in the validation report (D16) at the end of the project duration.

Please note in general, that the work presented here has been strongly affected by the delay of MSG data availability. Instead of one and a half year of research only roughly half

a year of research could be performed with real MSG data. Therefore, methods presented in this report cannot be taken as 'final' versions. All methods will be further improved during the remaining time of HELIOSAT-3. Information on these modifications will be made available to the partners via refereed publications and a further update of this documentation. Nevertheless, the methods are described in their current status. This is especially relevant as input data for the validation group in HELIOSAT-3 will be produced with this version of the HELIOSAT-3 software.

## 2.Retrieval of cloud mask and cloud optical properties

### 2.1.Introduction

In WP 2020, DLR was mainly engaged in the adaptation of the cloud retrieval scheme APOLLO to MSG. APOLLO was originally designed for the NOAA AVHRR instrument and has now been extended to the MSG SEVIRI instrument. This APOLLO-version makes use of the AVHRR-like SEVIRI channels, i.e. SEVIRI 0.6, 0.8, 1.6, 10.8, and 12.0  $\mu\text{m}$ . In the following, we use "APOLLO/AVH" and "APOLLO/SEV" for the AVHRR- and SEVIRI-version of APOLLO respectively, but simply "APOLLO" for general descriptions.

The MSG-SCENES software for the retrieval of macro-physical cloud properties available from EUMETSAT was also implemented. MSG-SCENES is based on the AVHRR APOLLO version but has been further modified by EUMETSAT. At the start of the project it was expected that the cloud detection part of MSG-SCENES is nearly 100% comparable to APOLLO/AVH. Unfortunately, it turned out that the SCENES software has some disadvantages in comparison to the APOLLO package:

- SCENES needs additional input, e.g. EMCWF weather forecasting model data, and additional computing time to ingest and process these data.
- Within SCENES the setting of thresholds for individual cloud detection tests and the way to combine test results is still in a rough, preliminary stage, e.g. which test is better over land and which over sea and how can they be combined to a comprehensive mask etc..
- SCENES tries to use all spectral SEVIRI channels, but the experience with these channels, especially in an operational environment, is rather poor.

Much more development than estimated at the beginning of HELIOSAT-3 would have been necessary to get the SCENES software as suitable as the APOLLO package for the MSG-SEVIRI processing chain.

Therefore, the HELIOSAT-3 cloud processing software uses SCENES only for

Reading XPIF or HRIT formatted MSG SEVIRI Level 1.5 data

Calibration into radiances

Generation of brightness temperatures and reflectances

Generation of geolocation information

Generation of satellite and sun geometry information

Generation of suitable output as input for the APOLLO/SEV software

Calling the APOLLO/SEV package

Transforming APOLLO/SEV binary output into HDF4 format which is the interchange format the HELIOSAT-3 partners agreed on

The APOLLO/SEV software is called internally from the SCENES software. From the scientific point of view, all cloud products generated are a result of the APOLLO/SEV software. Only basic data handling is done in the SCENES package. Therefore, in the

following sections the main features of the APOLLO software are described including all necessary modifications due to the adaptation of APOLLO/AVH to APOLLO/SEV.

A short description is given how APOLLO works. For a more detailed information it is recommended to refer to the publications cited below. In this report we concentrate on necessary changes for SEVIRI which were implemented as part of the HELIOSAT-3 project.

## 2.2. Cloud retrieval scheme

The AVHRR Processing scheme Over cLOUDs Land and Ocean (APOLLO) was the first AVHRR data processing scheme to make use of all five spectral channels during daytime. It discretises all pixels into four different groups called cloud-free, fully cloudy, partially cloudy (i.e. neither cloud-free nor fully cloudy) and snow/ice-contaminated, before deriving physical properties (Saunders and Kriebel 1988, Kriebel et al. 1989, Gesell 1989, Kriebel et al, 2003).

Within APOLLO, clouds are categorised into three layers according to their top temperature. The layer boundaries are set to 700 hPa and 400 hPa. The associated temperatures are derived from standard atmospheres. Further, each fully cloudy pixel is checked to see whether it is thick or thin cloud, depending on its 11  $\mu\text{m}$  and 12  $\mu\text{m}$  brightness temperatures and, during daytime, its channel 0.6  $\mu\text{m}$  and 0.8  $\mu\text{m}$  reflectances. Thin clouds (with no thick clouds underneath) are taken as ice clouds, i.e. cirrus, whereas thick clouds are treated as water clouds.

Cloud cover is derived for each cloud type separately, which is trivial for the fully cloudy (100%) and the cloud-free pixels. To derive the fractional cloud cover of the partially cloudy pixels at daytime, the relationship of the measured 0.6  $\mu\text{m}$  and 0.8  $\mu\text{m}$  reflectances to the mean of the fully cloudy and cloud-free reflectances in a local neighbourhood (e.g. 50 by 50 pixels) is used. At night-time the relationship of measured to local mean cloud-free and fully cloudy 10.8  $\mu\text{m}$  radiances is used. More details of these methods can be found in Saunders and Kriebel (1988).

At daytime, for each fully cloudy pixel, clouds optical depth, liquid/ice water path and IR-emissivity are derived by means of parameterisation schemes using the reflectance at 0.6  $\mu\text{m}$ , i.e. which does not depend on and is not influenced by water vapour absorption above  $\sim 0.75 \mu\text{m}$ . The parameterisation scheme is based on the directional hemispherical cloud top reflectance, which is obtained from the (measured) bi-directional top of atmosphere reflectance by applying an anisotropy correction, correction due to ozone absorption and subtracting the surface part of the reflectance transmitted through the cloud. Details of the method can be found in Kriebel et al. (1989).

At daytime, as well as at night-time, for each fully cloudy pixel excluding thin clouds, the cloud top temperature is obtained by means of a correction for the water vapour absorption above the cloud and for the (attenuated) surface radiance through the cloud. The detailed description of the method can be found in Saunders (1988).

These four products can also be obtained as weighted averages for local (e.g. 10x10) pixel boxes compatible with model grid values.

The adaptation of APOLLO to MSG-SEVIRI comprises the following changes, developments and tests which are described in the following sections.

### 2.2.1. Technical changes

- Metadata and a metadata management are necessary within APOLLO to identify the data, to check the correct input quantities for the algorithms and schemes, to avoid combinations of mismatching subsets and to document the processing history. For

SEVIRI these metadata have been defined and their management has been implemented into the relevant APOLLO/SEV routines.

- The APOLLO software uses internal buffers to keep the input and output image data as well as intermediate results. The intermediate results are either single pixel values or mean values and histogram data for boxes of pixels representing local neighbourhoods. The size of the boxes can vary between 32 and 64 lines and elements and the dimensions of I/O-buffers must be 3712 x 3712 in APOLLO/SEV instead of 2048 x ~5500 in APOLLO/AVH. All these buffers have been adapted, i.e. extended in all subroutines.
- The access to geo-referencing and the sun/satellite viewing geometry is indispensable. APOLLO/AVH uses as interface files with gridded data with 32 x 32 pixels grid size and interpolates in-between. This interpolation is not reasonable for SEVIRI data because of large interpolation errors at the edge of the earth-disk. For APOLLO/SEV a new interface has been written for direct read of latitude, longitude, sun/satellite zenith and sun/satellite azimuth angles pixel by pixel from the corresponding SCENES output without any interpolation. Further, check for space-pixels has been implemented where-ever critical operations could happen, e.g. zero-divide etc.
- Some APOLLO algorithms use parameters depending on the relative air-mass pixels were recorded through. For AVHRR, the relative air-mass of each pixel depends only on the element number, but for SEVIRI it depends also on the line number. Therefore, the development of an algorithm to obtain the correct relative air-masses was necessary for APOLLO/SEV.
- The processing of a SEVIRI scene is carried out in three steps. The first is the ingestion of SCENES-output into the APOLLO/SEV environment. The second is cloud detection and the third step is product generation. Each step is implemented as a UNIX shell-script and each script is a combination of APOLLO/SEV commands and shell commands. These scripts have been developed taking into account special features of the SEVIRI sensor, i.e. the masking of the space-pixels and the use of HRIT-strips as subsets. The latter has made the use of 58x58 pixel neighbourhood-boxes necessary because the HRIT size of 3712 x 464 pixels is a multiple of 58. The choice of 58 x 58 pixels as box size should avoid tailoring effects while dealing with such subsets.
- The value ranges of most quantities of SEVIRI data or metadata are different to those of AVHRR data. Since APOLLO keeps all value ranges stored for comparison and enhancement purposes, an adaptation of all their checks and presets has been done. Therefore, APOLLO/SEV uses other limitations and defaults than APOLLO/AVH .
- The generation of quicklooks (QLs) for quality control is different in APOLLO/SEV compared to APOLLO/AVH. Mixed day/night- scenes are much more frequent in SEVIRI data than in AVHRR and the presence of space-pixels is quite common. Therefore, more care has been laid on the enhancement and layout of the QLs in APOLLO/SEV resulting in additional processing steps.

### 2.2.2. Physical changes

- The defaults of the thresholds for most of the APOLLO algorithms have to be set depending on the region covered by the data to be processed. The thresholds itself are either set explicitly or are dynamically derived from local histogram calculations in neighbourhood-boxes. Therefore, all explicitly set thresholds and all default settings have been changed, i.e. their values have been adapted to the regions visible by SEVIRI (especially necessary for non-European areas). These presets must be physically reasonable, i.e. are climatologically or empirically determined for the complete SEVIRI

view and usually depend on the region and also the season. Therefore, all tables with presets are different between APOLLO/SEV and APOLLO/AVH.

- Step 2 and 3 in the APOLLO processing contain corrections for regional peculiarities, e.g. cold currents and thermal fronts in the ocean, deserts with frequent sand storms etc. These local peculiarities can cause misclassifications and wrong values for the cloud parameters and the processing must correct for it. These corrections depend not just on the region with the certain peculiarity but also on the degree of experience with it. Therefore, each APOLLO processing scheme is specific for the region it is applied to and APOLLO/SEV and APOLLO/AVH differ here as much as the AVHRR-view (covering mainly Europe due to field-of-view of DLR's receiving system at Oberpfaffenhofen) and SEVIRI-view are different.
- A check if results of APOLLO/SEV are reasonable, has been done by means of comparisons of typical cloud parameter values derived from both, SEVIRI and AVHRR, and by means of a quality analysis based on many years of experiences with APOLLO/AVH.

## 2.3. Cloud parameter file format

APOLLO delivers cloud mask, cloud classification, cloud optical depth, liquid and ice water path, cloud top temperature and infrared emissivity as cloud parameter products.

It was agreed between the HELIOSAT-3 partners to store the following parameters in the intermediate cloud product for further use in the irradiance calculation part of HELIOSAT-3:

- Cloud mask
- Cloud coverage and type
- Cloud top temperature
- Cloud optical thickness

The storage of further APOLLO parameters in the HELIOSAT-3 cloud product is possible if requested by the developers of the surface irradiance scheme. Up to now the listed parameters seem to be sufficient and we restrict the HELIOSAT-3 cloud product to them in order to reduce transfer and storage amount.

Figures 1 to 5 show attributes as written into the HDF4 format and describing each data set. These figures were taken as screenshots from the freely available JAVA HDF Viewer software package. The example given is a data set from January 17<sup>th</sup>, 2004, 18 UTC for HRIT segment no. 6 which covers in this case all MSG lines no. 2500 to 2784 and MSG elements no. 600 to 3000. This data set is part of the first test data sent to HELIOSAT-3 partners for further development of the surface irradiance calculation scheme. It was agreed to use a Europe subset of the data – therefore, only MSG elements from no. 600 to 3000 are used.

Attributes are included in each cloud product. Each cloud product is therefore fully described by both its filename and its internal attributes.

Filename structure:

Clouds\_YYYYMMDD\_HHMM\_segmentid\_LAAAABBBBECCCCDDDD.hdf

- YYYY = year in 4 digits
- MM = month in 2 digits
- DD = day in 2 digits

HH = hour in 2 digits

MM = minute in 2 digits

Segmentid = HRIT segment number  
(1 .. 8, 1 = southernmost HRIT segment, 8 = northernmost HRIT segment)

AAAA = first line of data (1 .. 3712, 1 = southernmost SEVIRI line)

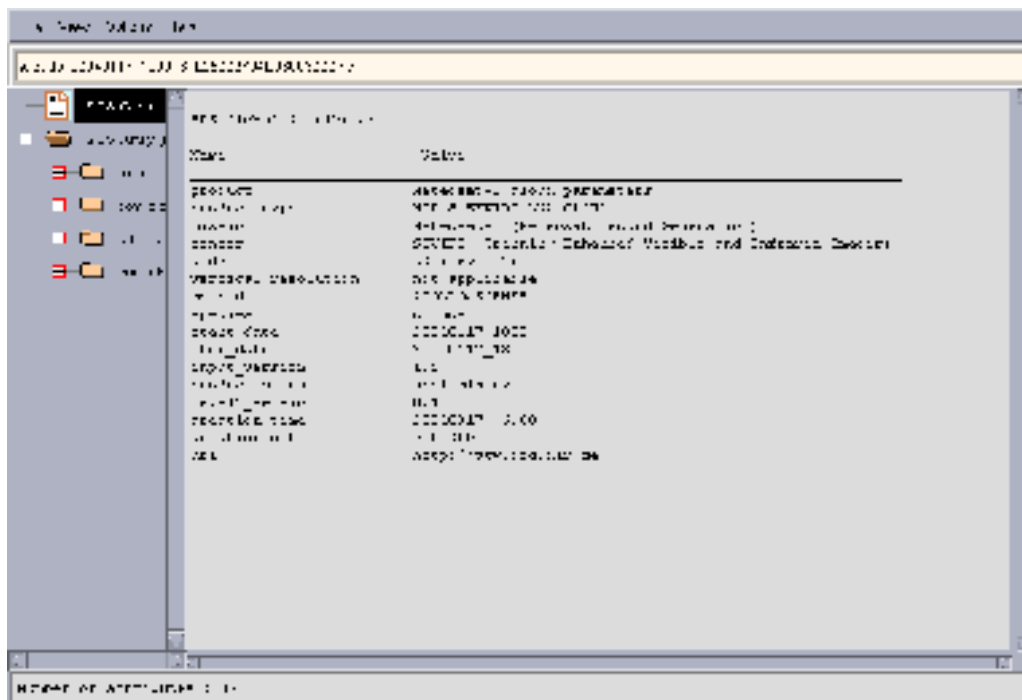
BBBB = last line of data (1 .. 3712)

CCCC = first element of data (1 .. 3712, 1 = easternmost SEVIRI sample)

DDDD = last element of data (1 .. 3712)

Example as shown in figures 1 to 5:

clouds\_20040117\_1800\_6\_L25002784E06003000.hdf is for January 17<sup>th</sup>, 2004, 18 UTC, HRIT segment no. 6, subset from line 2500 to 2784 and from element 600 to 3000.



*Figure 1: Cloud product - global parameters  
Basic information concerning the whole cloud product as satellite, instrument, level, date, input version, level2 version, generation data and place*

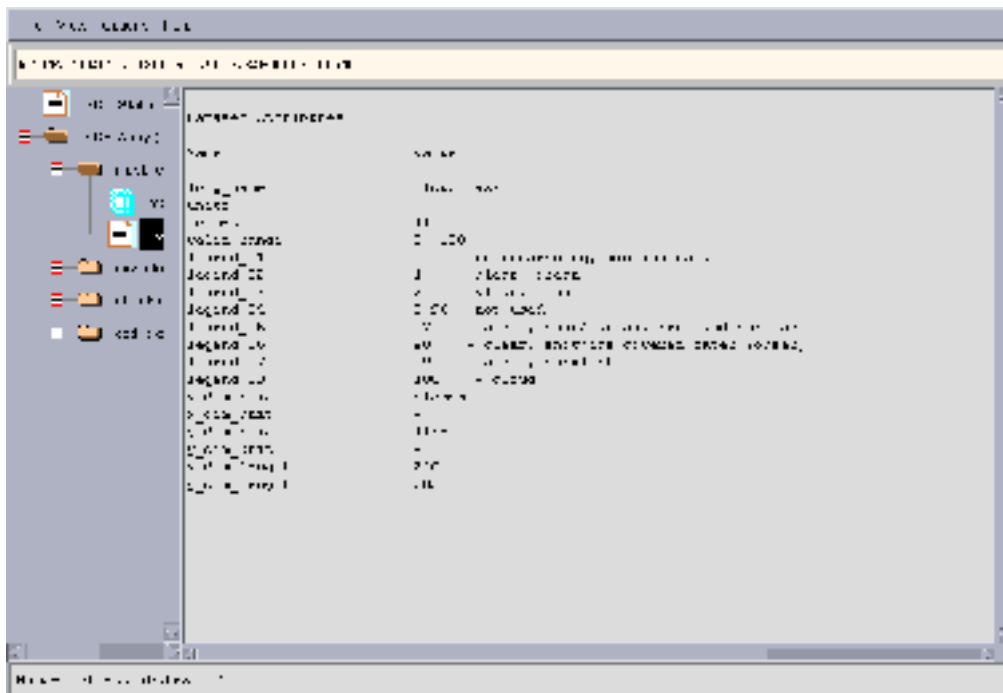


Figure 2: Cloud product – cloud mask  
Valid range is 0 to 100, several numbers give cloud mask classes.

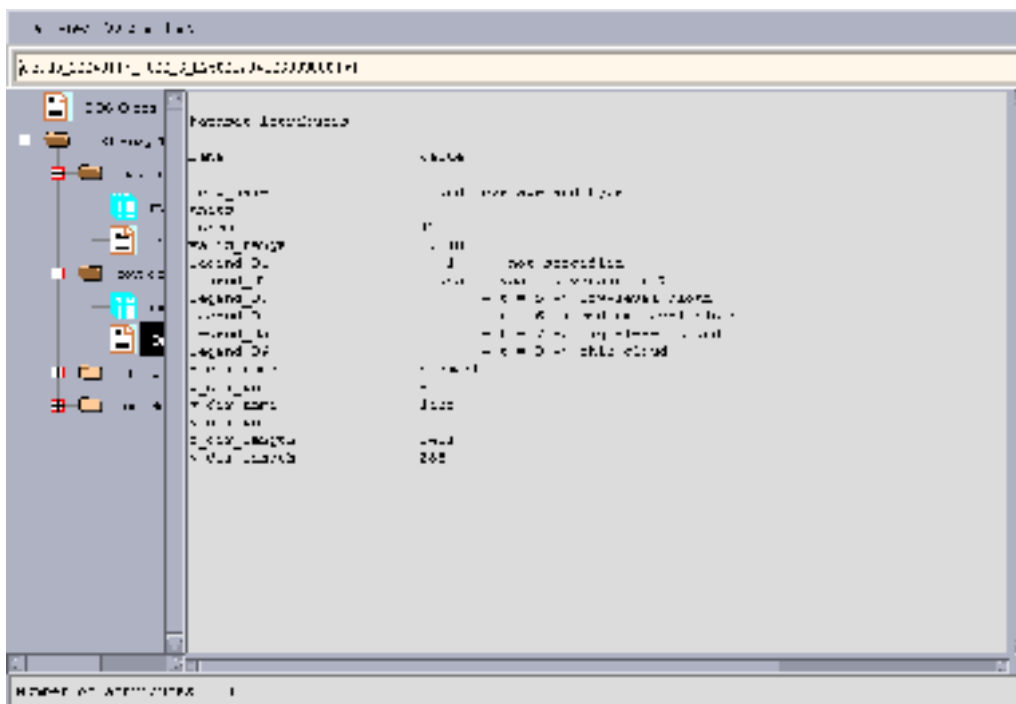


Figure 3: Cloud product - cloud coverage and type  
Valid range is 0 .. 1008. 4 digits are used to code both coverage in % and the type of cloud.  
First 3 digits is coverage in %, 4<sup>th</sup> digit is flag for low level cloud (value 5), medium level cloud (value 6), high level cloud (value 7), and thin cloud (value 8) . Example: 278 means 27% of thin cloud.



## 2.4. Example results

The following colour images (fig 6 to 12) show APOLLO products together with a composite of the SEVIRI channels at 0.6, 0.8 and 12  $\mu\text{m}$ . This is an example of a set of typical quicklook images of the APOLLO/SEV cloud products at daytime. The daytime scheme uses only pixels with enough sunlight to derive optical cloud parameters. In the quicklook images all pixels with not enough sunlight (dark pixels) or space pixels are masked. In the colour composite, both, space and too dark pixels are masked in marine blue. In the product images, the space pixels are masked in black and dark pixels of the earth in red. The example SEVIRI scene is from January 23<sup>rd</sup>, 2004, 11:00 UTC.

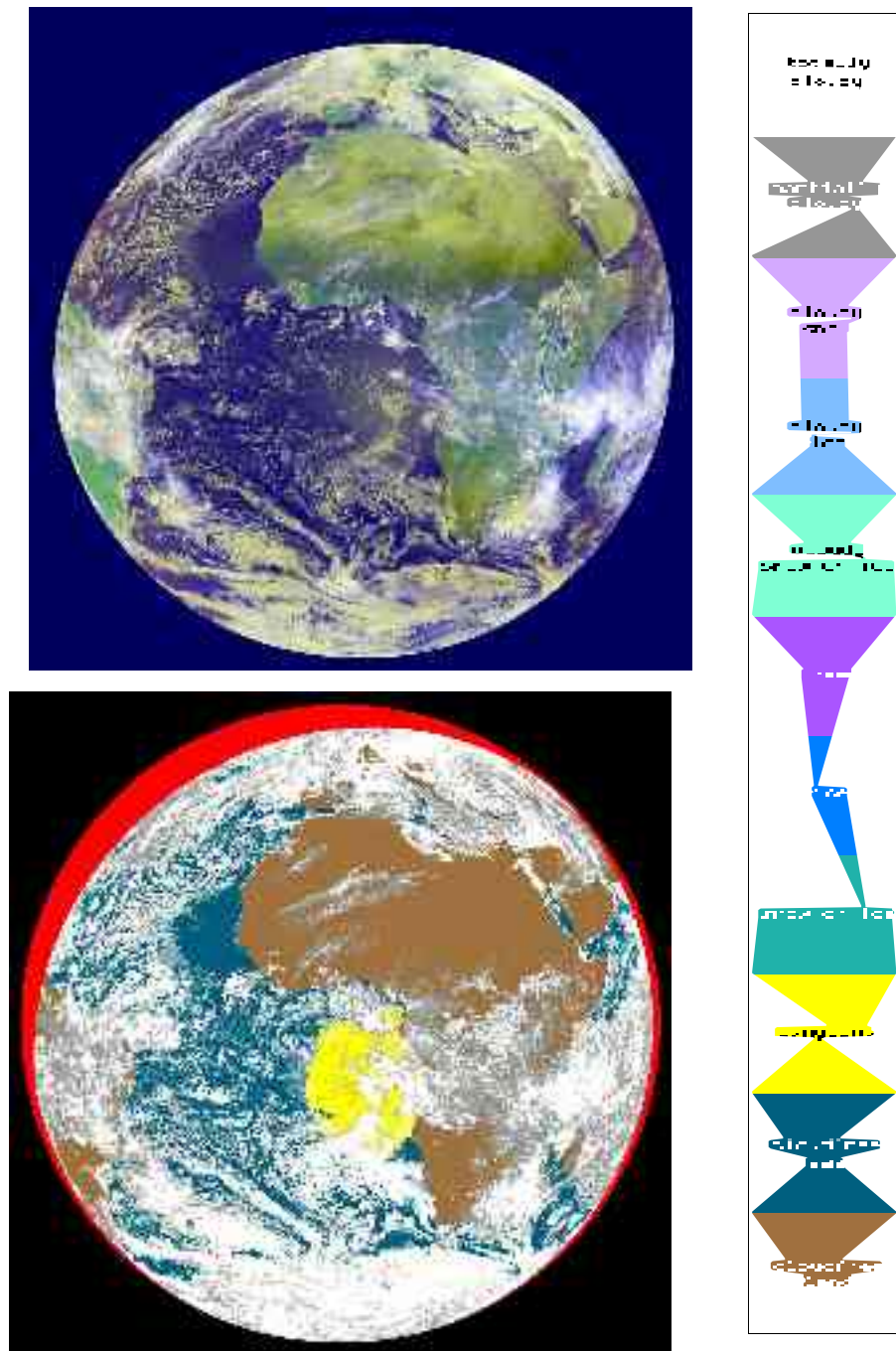
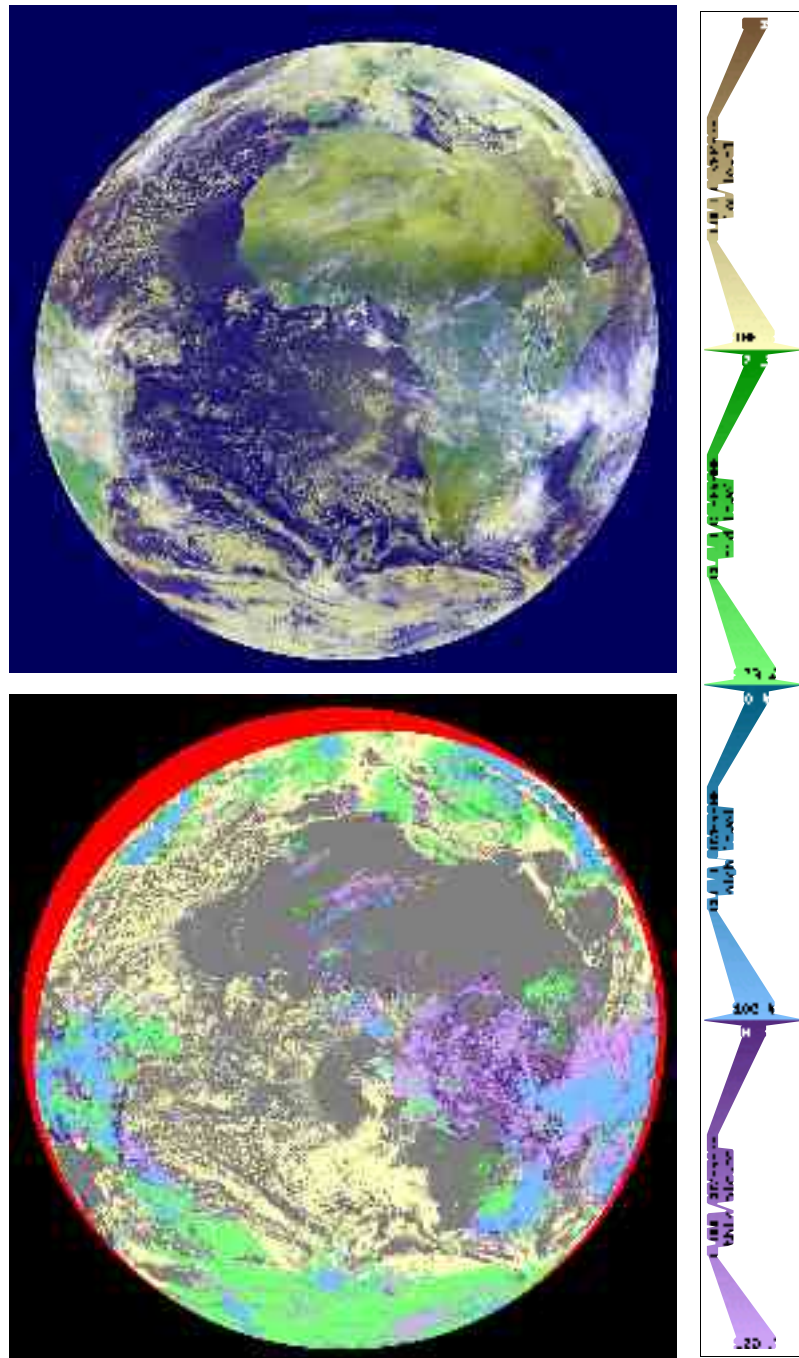


Figure 6: APOLLO colour composite and cloud mask for January 23<sup>rd</sup>, 2004, 11:00 UTC

The lower image is the typical quicklook of an extract of the APOLLO/SEV cloud-mask. The original APOLLO/SEV cloud-mask contains the information of all applied cloud detection algorithms. In this quicklook, the cloud-mask content is reduced to an overall cloud/snow/ice-information without details on the algorithms combined with a land/sea/sunglint-information from another APOLLO/SEV mask.



*Figure 7: APOLLO colour composite and cloud coverage for January 23<sup>rd</sup>, 2004, 11:00 UTC*

The cloud fraction, i.e. the pixel-coverage, is derived for four cloud classes, i.e. for three levels of thick clouds and for thin clouds. For the thin clouds no level can be given. In the lower image four colours indicate those four cloud classes: Yellow for low level, green for mid-level, blue for high level clouds and lilac for thin clouds. The brightness of the colours indicates the cloud-fraction from brightest=100% to darkest=0%.

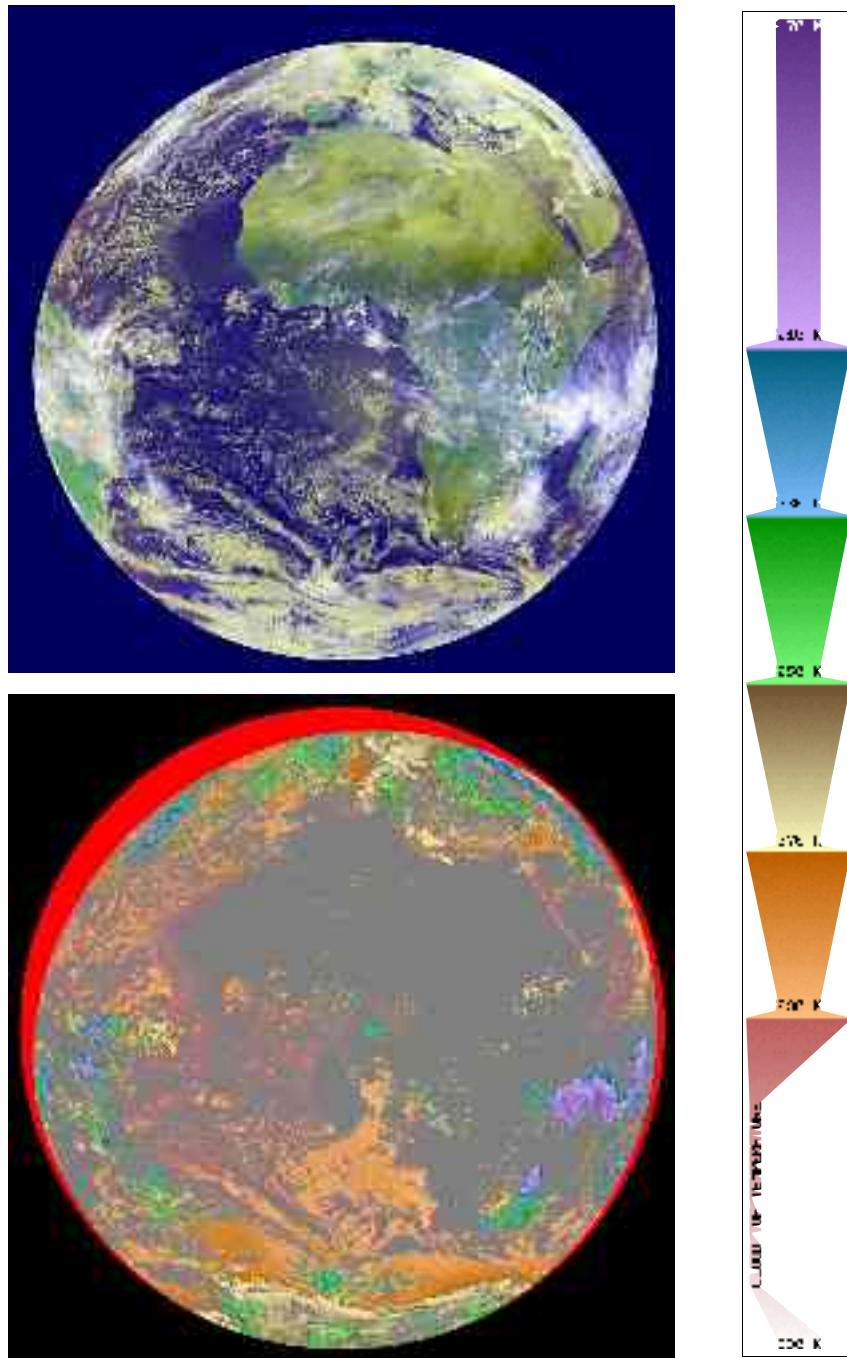


Figure 8: APOLLO colour composite and cloud top temperature for January 23<sup>rd</sup>, 2004, 11:00 UTC

APOLLO/SEV calculates the cloud top temperature for all totally cloudy pixels except thin clouds. The lower image displays the cloud top temperature using the colour scale above with changing colours at 290, 270, ...Kelvin. The lowest temperatures are found on top of the inner-tropical cumulonimbus clouds north of Madagascar at approximately 180 K.

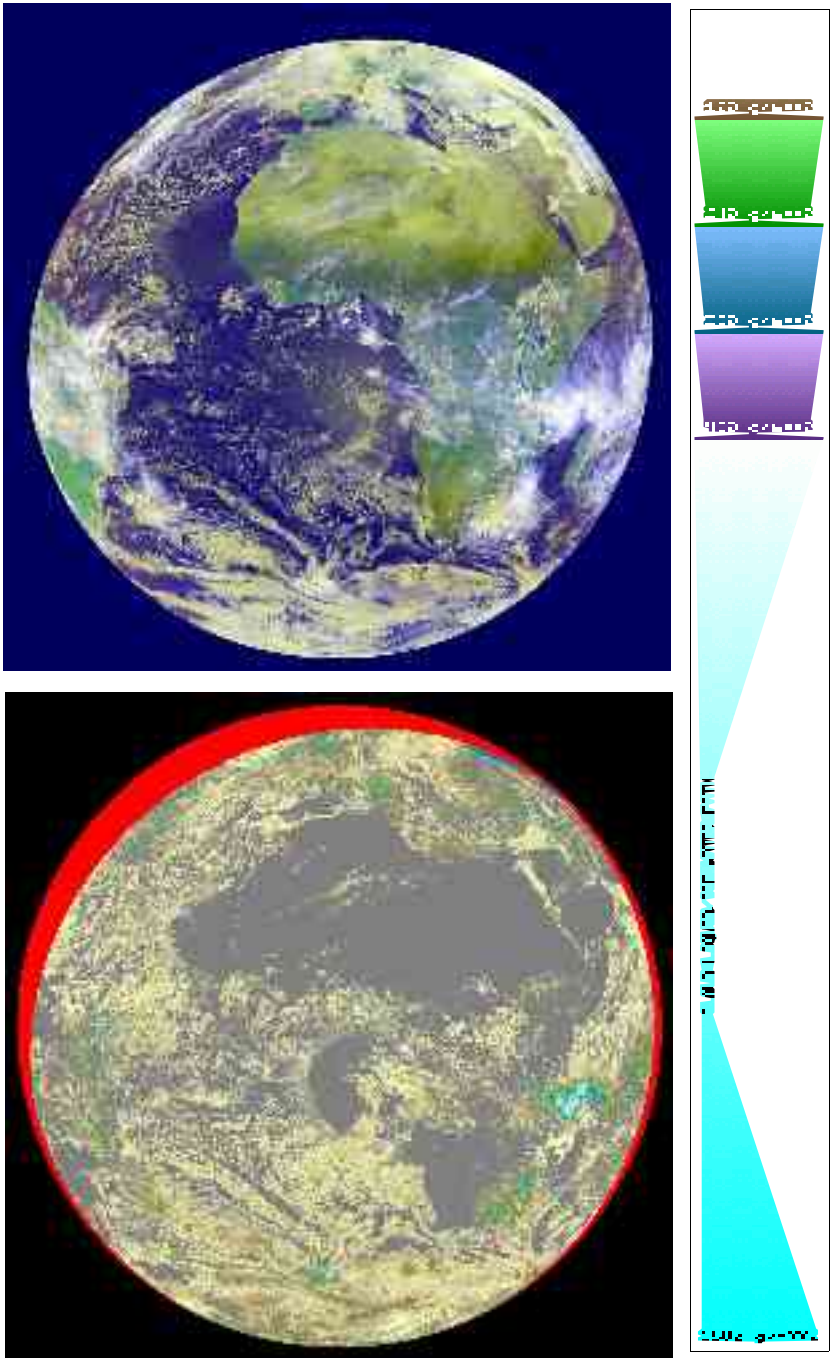
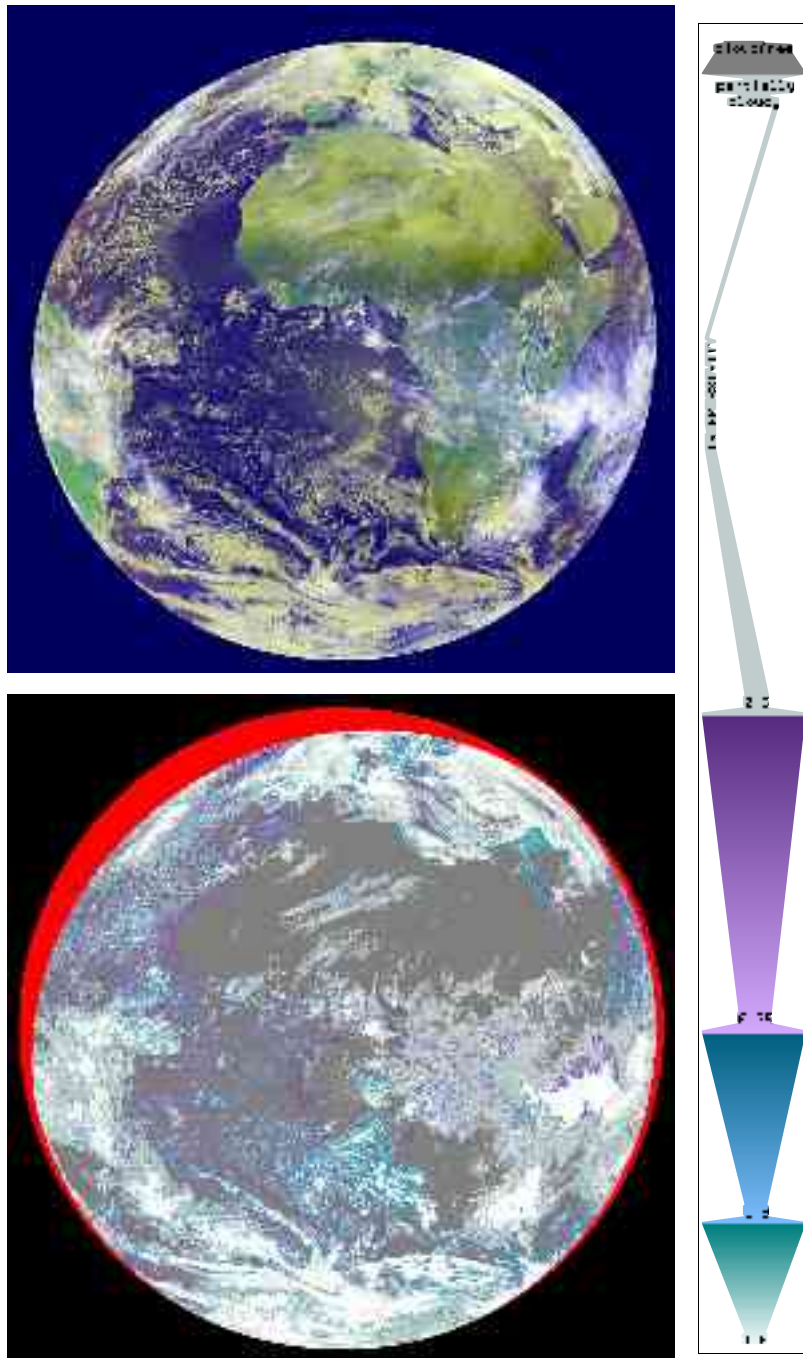


Figure 9: APOLLO colour composite and cloud liquid water path for January 23<sup>rd</sup>, 2004, 11:00 UTC

The lower colour image shows the columnar cloud water content in  $\text{g/m}^2$ , i.e. the cloud liquid/ice water path. The highest values of approximately  $1100 \text{ g/m}^2$  correspond with the lowest cloud top temperatures.



*Figure 10: APOLLO colour composite and cloud infrared emissivity for 23. January 2004, 11:00 UTC*

In the lower image the cloud infrared emissivity is displayed for totally cloudy pixels. Additionally, partially cloudy pixels are displayed in bright grey. The higher values are associated with mid- and high- level clouds.

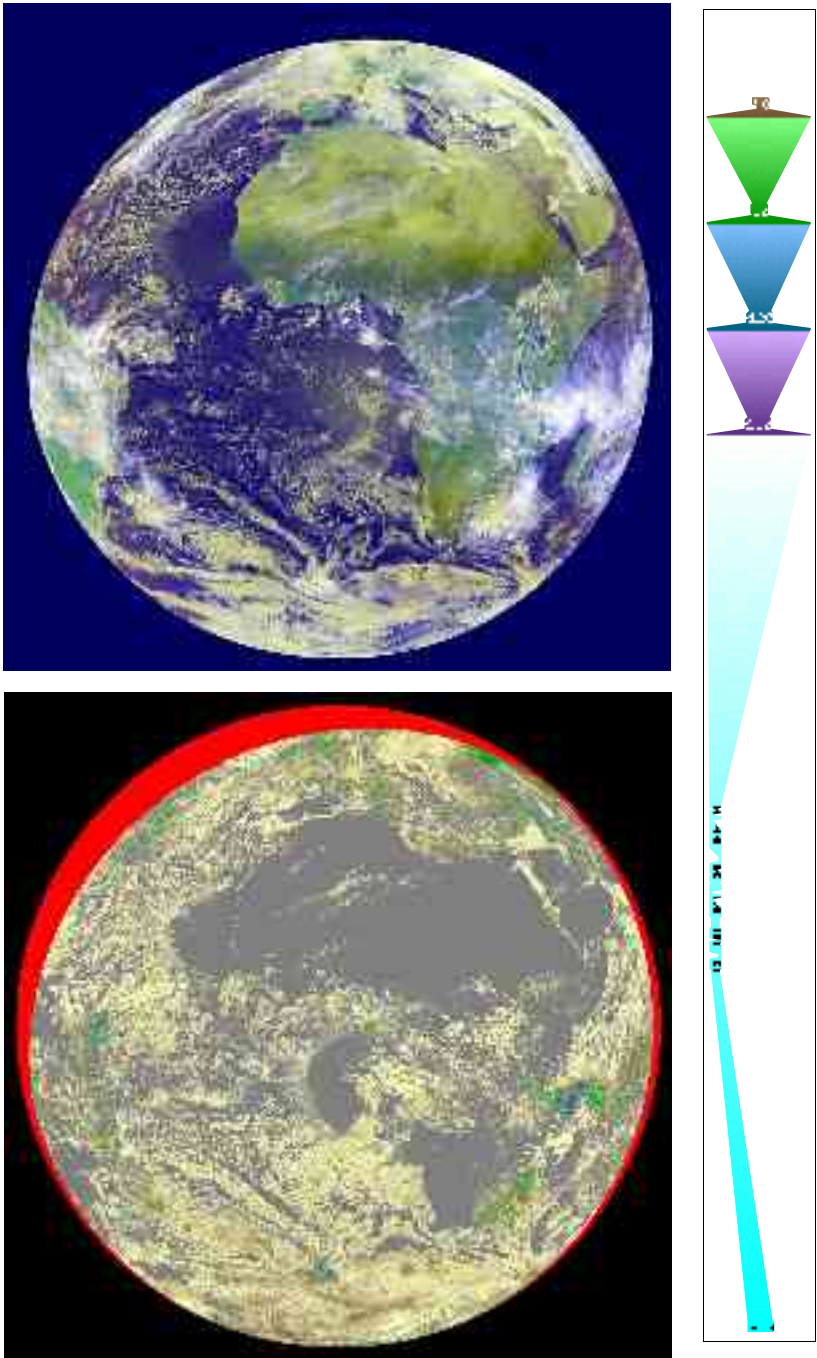
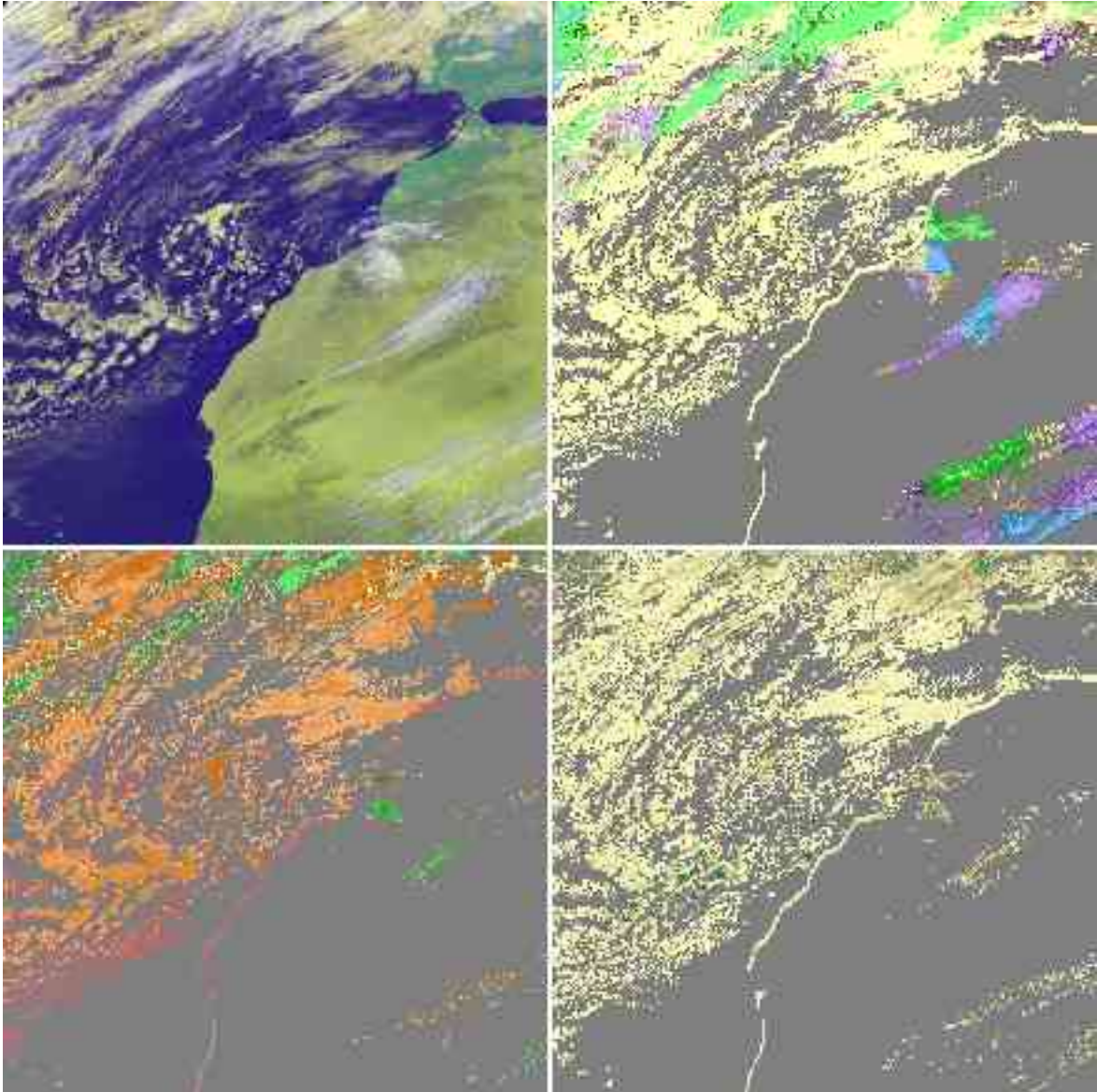


Figure 11: APOLLO colour composite and cloud optical thickness for 23. January 2004, 11:00 UTC

In APOLLO/SEV the values of the cloud optical depth can range from 0.5 to 500. Optical depths lower 0.5 are more typical for aerosols than for clouds and therefore not displayed in the lower image. The highest values are associated with the highest water path values and the lowest cloud top temperatures north of Madagaskar.

The four images below show a subset in higher resolution around the Canary Islands. Besides the composite (upper-left) the cloud coverage (upper-right), cloud top temperature (lower-left) and cloud optical depth (lower-right) are displayed using the same colour scale as in figures 6- 11. The misclassification along the African coast results from the still not completely solved problem with the accuracy of the geo-referencing. This is under investigation.



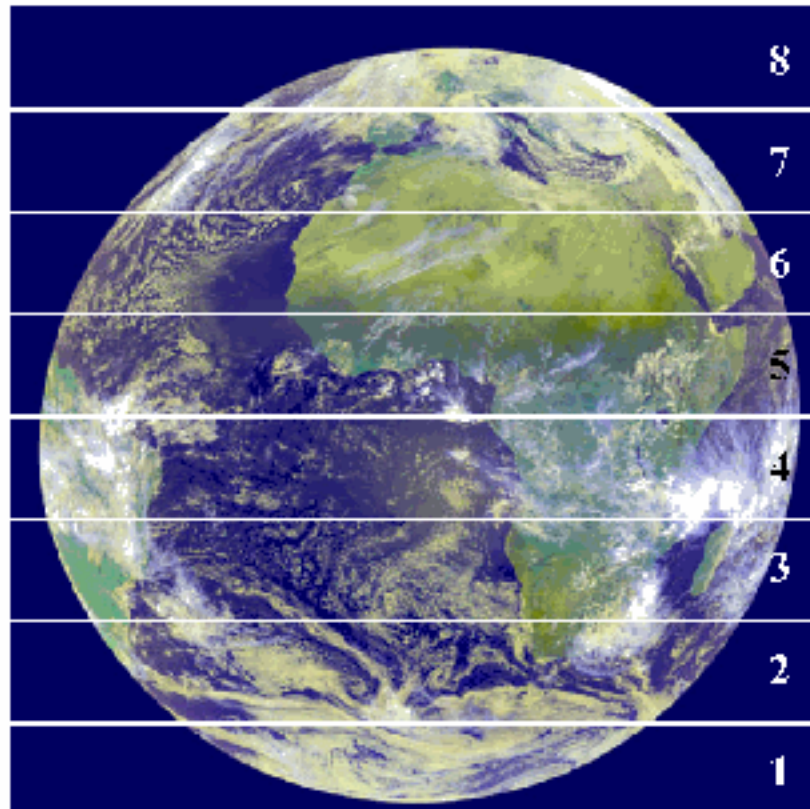
*Figure 12: Subset in higher resolution around the Canary Islands. (January 23<sup>rd</sup>, 2004) Besides the composite (upper- left) the cloud coverage (upper- right), cloud top temperature (lower- left) and cloud optical depth (lower- right) are displayed using the same colour scales as before.*

## 2.5. Further modifications for operational processing

MSG SEVIRI Level1.5 data is delivered from EUMETSAT via antenna in the so-called HRIT (High Rate Information Transmission) format. Due to an effective wavelet compression HRIT files are very small in comparison to e.g. uncompressed XPIF (Extended Processed

Image Format) files. Many users used PIF formatted data for the older Meteosat generation. As Meteosat Second Generation delivers much more Level1.5 data (1 Terabyte per month) the need for compression of archived data is obvious. Therefore, DLR-DFD archives HRIT files only.

In order to ensure an optimised transfer scheme EUMETSAT delivers HRIT files only for segments of MSG data. Each full disk of MSG is separated in 8 segments (figure 13). Basic idea is that using 8 small files instead of a single large file increases the success rate for data transmission between the EUMETSAT ground segment and user receiving stations.



*Figure 13: HRIT segmentation of MSG full disk (segment numbering starts in the South).*

This structure of HRIT segments is kept in the DLR processing chain. Therefore, all products derived directly from MSG data (e.g. cloud and also water vapour products) are also derived for this segmented structure. This means that a full disk cloud product is not processed in a single step. 8 segments of SEVIRI Level 1.5 HRIT data have to be processed and 8 cloud products (one for each segment) are the output.

We recommend to keep this structure also for irradiance products to be stored later at DLR-DFD. It is much easier for a potential end-user to get all SEVIRI products with the same temporal and spatial coverage independent of the physical parameter.

## 2.6. References

Gesell, G., 1989: An Algorithm for Snow and Ice Detection Using AVHRR Data: An Extension to the APOLLO Software Package. International Journal of Remote Sensing, Vol. 10, Nos. 4 and 5, pp. 897- 905

Kriebel, K.T., R.W. Saunders and G. Gesell, 1989: Optical Properties of Clouds Derived from Fully Cloudy AVHRR Pixels. Beiträge zur Physik der Atmosphäre, Vol. 62, No. 3, pp. 165-171, August 1989

Kriebel K. T., Gesell G., Kästner M., Mannstein H., The cloud analysis tool APOLLO: Improvements and Validation, Int. J. Rem. Sens., 24, 2389-2408, 2003

Saunders, R.W. and K.T. Kriebel, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. International Journal of Remote Sensing, 9, 123-150

Saunders, R.W., 1988: Cloud top temperature/height: A high resolution imagery product from AVHRR data, Meteorological Magazine, Vol 117, pp 211-221

## 3. Retrieval of water vapour

### 3.1. Introduction

Several retrieval methods to derive total water vapour content over land from MSG SEVIRI data have been tested at DLR. Comparisons indicate that a modified method based on the Kleespies and McMillan method (Kleespies and McMillin, 1990) is suited very well for MSG. The development of the retrieval method is based on the TIGR3 radiosonde data set. First results with MSG data will be shown. Validation will be given in deliverable D16 (validation report) at the end of the project duration.

### 3.2. Water vapour retrieval development (level 2)

The Thermodynamic Initial Guess Retrieval (TIGR, Chedin et al., 1985) data set is used to represent possible atmospheric states of temperature and humidity distribution. The TIGR data set was developed at the Laboratoire de Meteorologie Dynamique (LMD, Paris) especially for the development of retrieval methods. It includes more than 2300 radiosonde profiles from all over the world. The profiles are selected to represent the maximum variability of the atmosphere. This means, that especially extreme situations are collected in the TIGR data set and that it is not a climatological data set with representative mean values. For the development of robust and operational retrievals it is very important to have such a simulation data set which includes also extreme situations from all over the world. Temperature, humidity and ozone on 40 vertical levels between 1013 and 0.05 hPa, latitude, and longitude are given for each profile. For the development of the total water vapour content (TWC) algorithm temperature and humidity profiles were used. The global variability of total water vapour column in the TIGR data set is shown in figure 14 and 15.

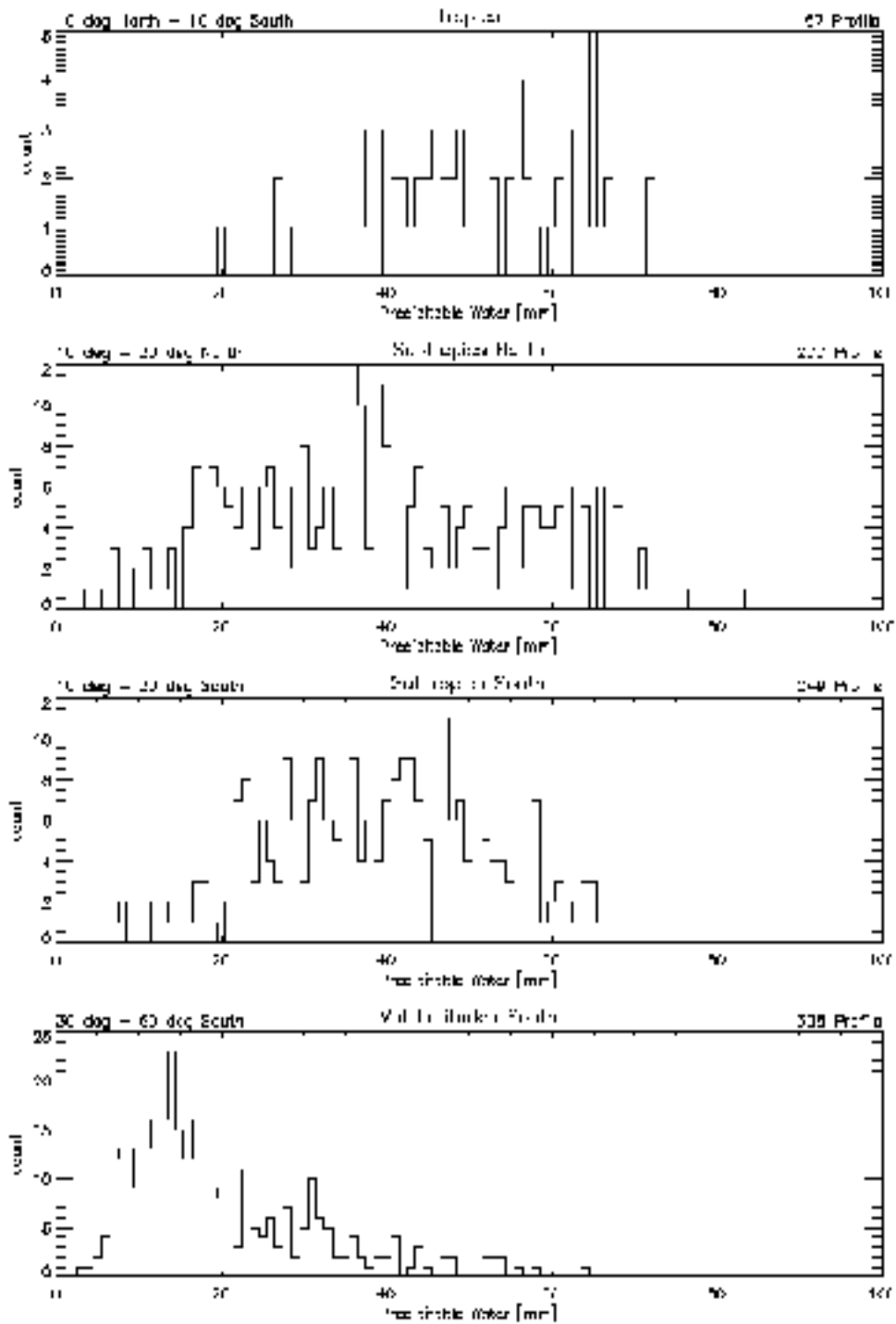


Figure 14: Global variability of total water vapour content in the TIGR data set (1/2)

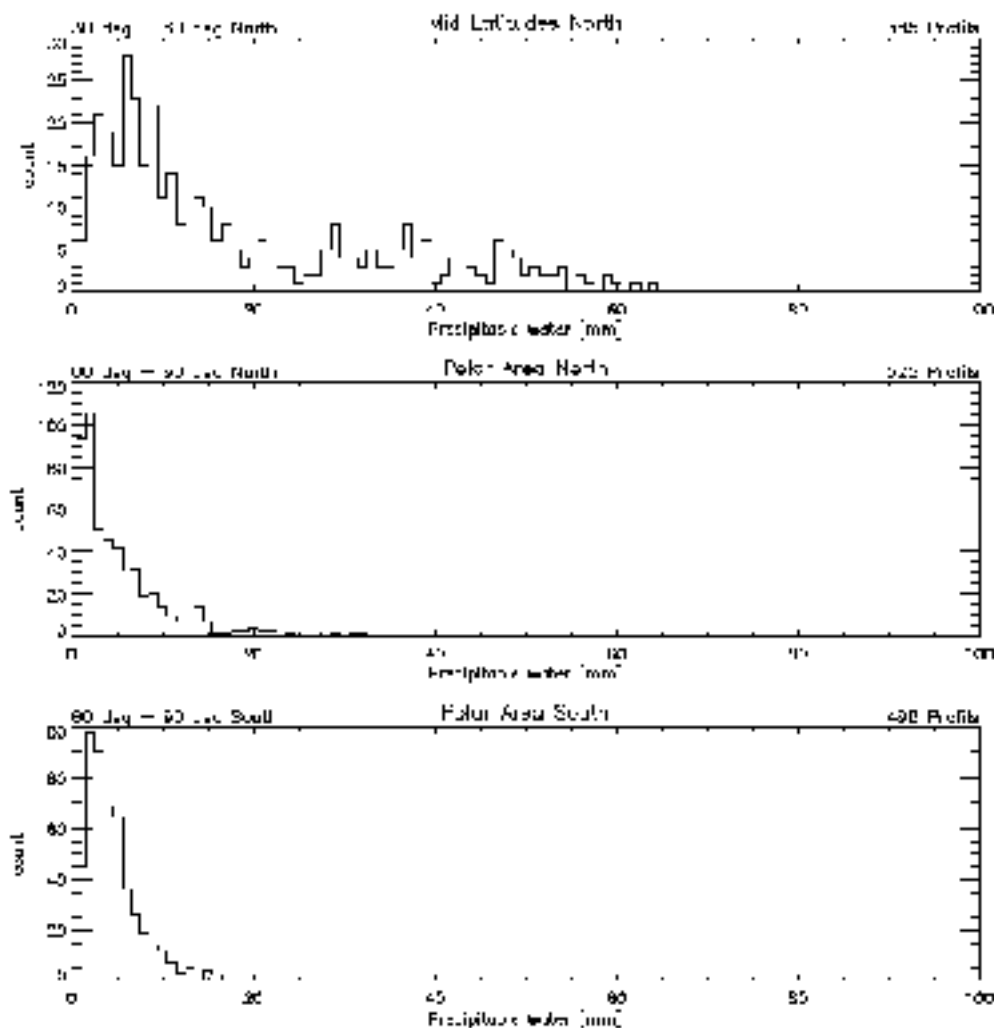


Figure 15: Global variability of total water vapour content in the TIGR data set (2/2)

The following steps were performed:

- a) all TIGR humidity profiles were integrated vertically to calculate 'true' TWC
- b) forward radiative transfer calculations (MODTRAN3.7, Abreu and Anderson, 1996; Berk et al., 1999) were performed with all TIGR temperature and humidity profiles to simulate MSG satellite measurements in the 10.8 and 12  $\mu\text{m}$  channels
- c) retrieval algorithms were applied to this simulated MSG radiances in order to derive TWC
- d) TWC column derived from these simulated MSG radiances were compared with the values calculated directly from the TIGR profiles

Both, MSG and MODIS were simulated and analysed during the algorithm development. MODIS (Moderate Resolution Imaging Spektoradiometer) is flown on the TERRA platform on a polar orbit. As the launch date for the MSG satellite was uncertain for a long time, we analysed the MODIS instrument as a possible backup solution.

Three different, so called 'split window' retrieval algorithms were tested. The methods were published by Eck and Holben (1994), Chesters et al. (1983, 1987) and Kleespies and McMillan (1984, 1990) and aim at the retrieval of water vapour column over land. Each of these methods was tested by the authors only with a few scenes in case studies. Objective of this study now was to test these methods for a global data set including as much as possible extreme atmospheric situations.

The method of Eck and Holben is not applicable as a global retrieval method. It turned out that it requires locally adapted empirically coefficients.

The method of Chesters et al. could be applied successfully only to TWC values below 40 mm. As the range of TWC in the TIGR data set is from 0 to 80 mm, the Chesters method seems inapplicable as an operational algorithm.

The method of Kleespies and McMillan is a modification of the Chesters method. It uses the split window channels at 10.8 and 12  $\mu\text{m}$ . In general, infrared measurements are affected by air temperature, surface temperature, surface emissivity, water vapour and absorption of other atmospheric gases. 'Split window' channels are selected close to each other, so that equal emissivity and absorption of other gases than water vapour can be assumed. Surface temperature is the same for both channels as this parameter is not dependent of the observer's wavelength. This reduces the dependence of measurements in these channels to air temperature and water vapour absorption. For the Chesters method different approaches to describe air temperature directly were tested, but all of them needed locally adapted coefficients. The method of Kleespies and McMillan avoids this. Two situations with varying surface temperatures are selected and therefore, two brightness measurements can be exploited. Having these two equations the air temperature dependence can be eliminated. MSG offers the possibility to use this approach as it measures in a 15 minutes temporal resolution and therefore, can deliver 'two situations with varying surface temperatures' during the daily temperature cycle.

The following equation describes the functional relationship between brightness temperatures and TWC.

$$\frac{\tau_{11}}{\tau_{12}} = \frac{(T_{11}^A - T_{11}^B)}{(T_{12}^A - T_{12}^B)} \quad (\text{Eq. 1})$$

$$TWC = fct \left\{ \frac{1}{\sec \theta} \ln \left( \frac{\tau_{11}}{\tau_{12}} \right) \right\}$$

$T$  is the brightness temperature in channel 11 or 12  $\mu\text{m}$ ,  $\theta$  is the satellite zenith angle,  $A$  and  $B$  are two temporal different situations and  $\tau$  stands for transmission.

Kleespies and McMillan proposed a linear relationship between transmission ratio term and TWC. Testing the whole TIGR data set it turned out that a third order polynomial describes the relationship better. Figure 16 shows the linear fit in green, a quadratic fit in blue and the third order polynomial in red. Assumably, the reasons for this non-linear behaviour are (a) non-neglectable absorption of other atmospheric gases if TWC is very small and (b) saturation of absorption lines for large TWC values.

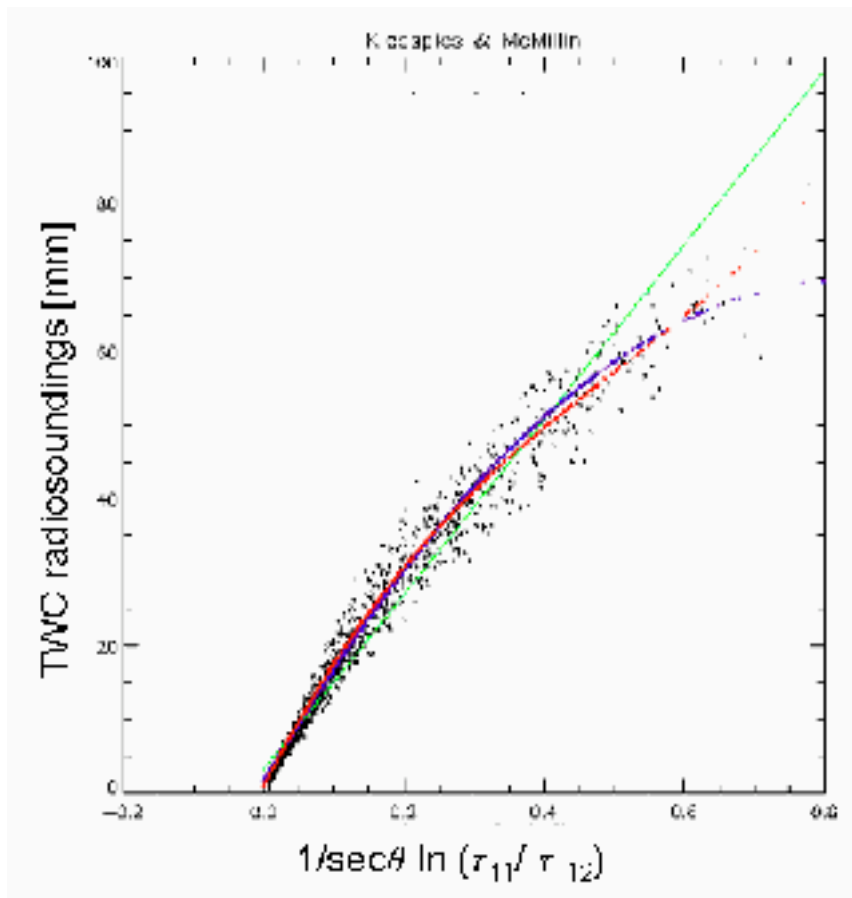


Figure 16: Different fit functions tested for the Kleespies and McMillan relationship. (fit function corresponds to eq. 1)

Using this approach TWC derived from simulated MSG radiances was compared with the 'true' TWC calculated directly from the radiosoundings. Figure 17 shows a good agreement with bias and standard deviation of  $1.16 \pm 1.61$  mm. The difference in surface temperature was assumed as 5K, surface emissivity was set to a global average value 0.975.

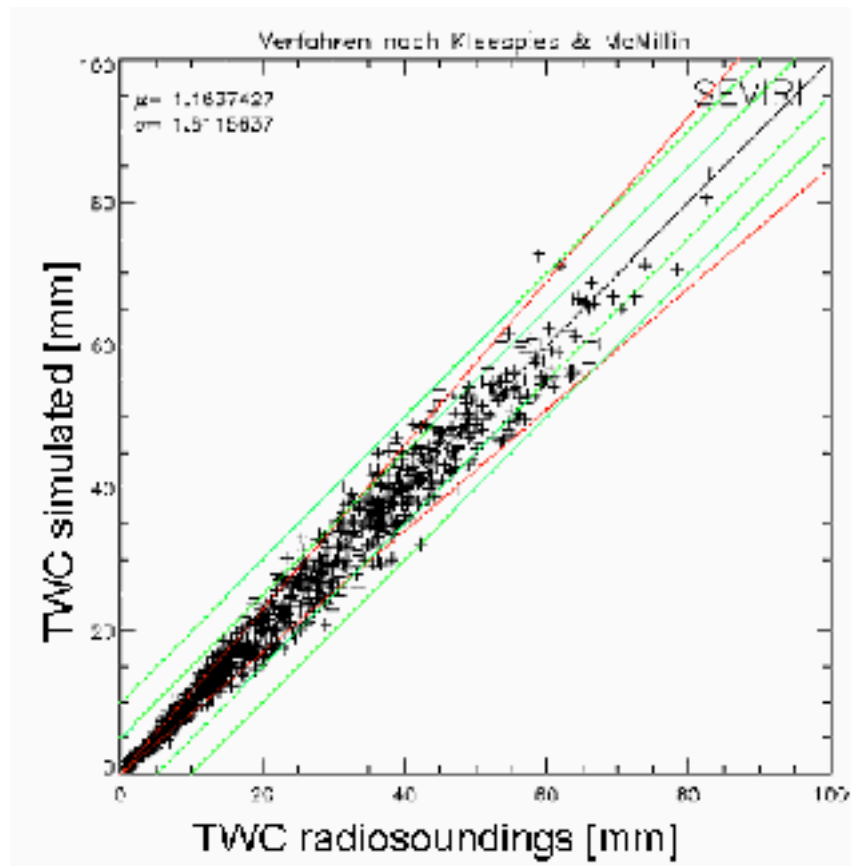


Figure 17: Comparison between measured TWC and TWC calculated with the modified Kleespies and McMillan method for all TIGR radiosonde cases

Up to now an ideal satellite without any measurement noise was assumed. MSG is specified to have a radiometric error of 0.25 K in the 10.8  $\mu\text{m}$  channel and 0.37 K in the 12  $\mu\text{m}$  channel. Figure 18 shows the results if noisy measurements are assumed. It turns out that 5K difference between the 'two situations with varying surface temperatures' is not sufficient if noisy input data is assumed. On the right part of figure 17 the same study with 10 K difference in surface temperature shows acceptable results with an agreement of  $4.5 \pm 3.7$  mm for a range of 0 – 80 mm. As a result of this study, a minimum difference of brightness temperatures in the 12  $\mu\text{m}$  channel of 8 K is asked for.

Of course, the assumption of equal surface emissivity in both channels is not perfectly true. The extreme case of sand was tested, but the influence of varying emissivity (0.970 in the 10.8  $\mu\text{m}$  channel and 0.9865 in the 12  $\mu\text{m}$  channel) on the retrieval accuracy was not significant.

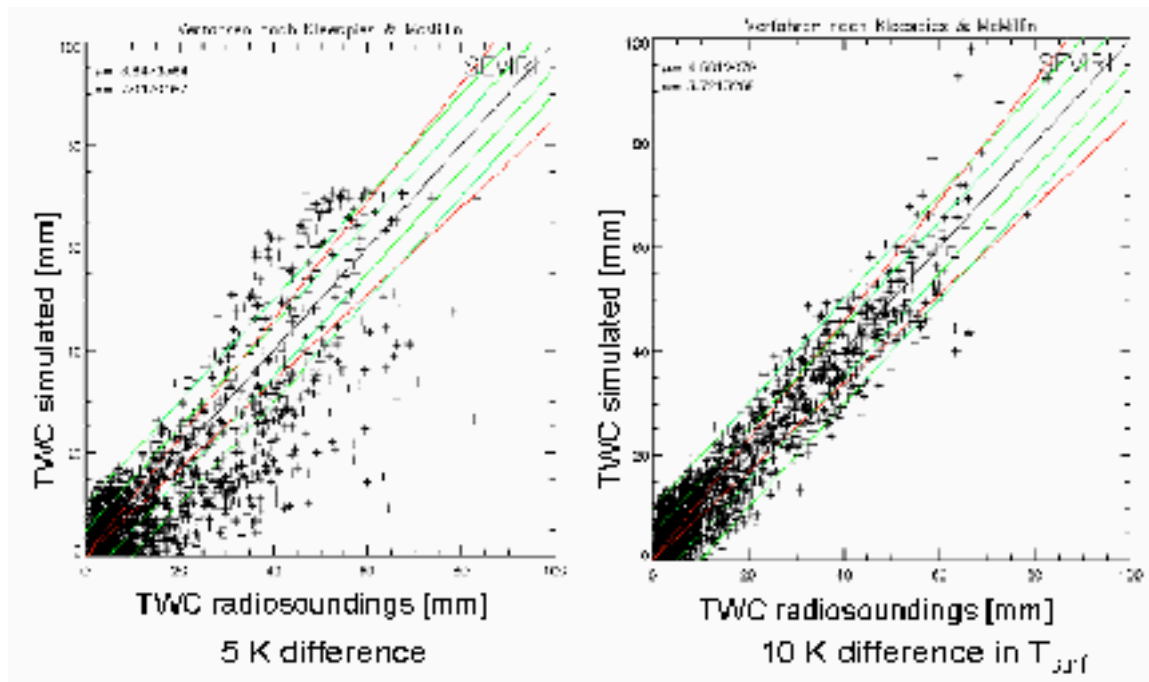


Figure 18: Simulation of effect of noisy signal on retrieved water vapour for a 5K (left) and a 10 K (right) difference in surface temperatures

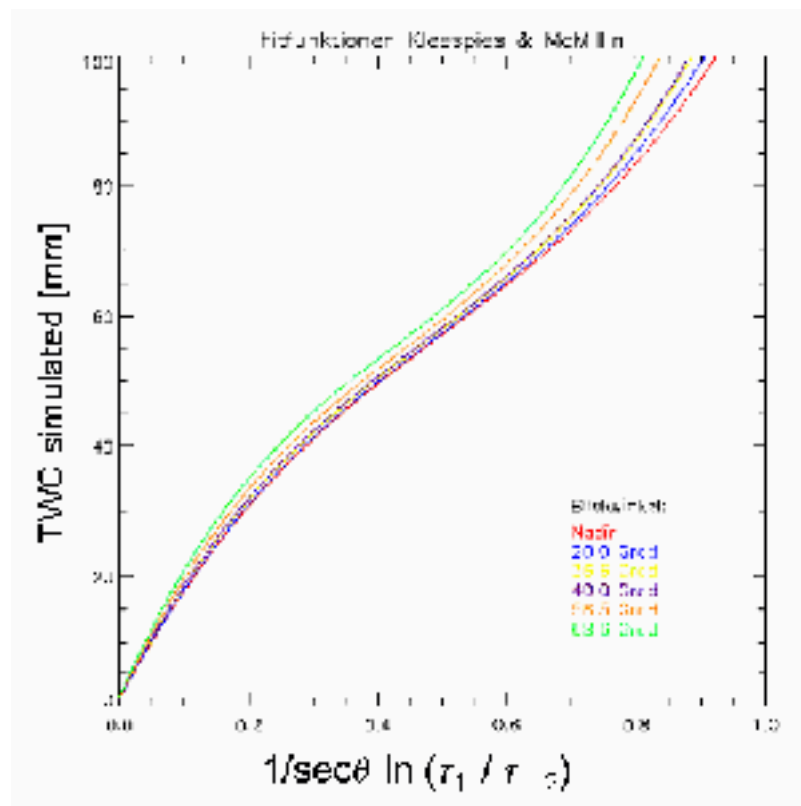


Figure 19: Scan angle dependency of fit function (fit function corresponds to eq. 1)

A significant dependence on satellite zenith angle was found. Figure 19 shows the third-order polynomial relationship mentioned above for several viewing angles. The relation includes only a scaling with the cosine of the zenith angle as airmass correction. Especially for larger TWC values this simple airmass correction is not suitable anymore. An explicit calculation of airmass correction in the infrared needs detailed information about the temperature profile of the atmosphere. Unfortunately, this will not be available later in the operational processing scheme. Therefore, we developed a parameterisation scheme. Figure 19 reveals that the change between the curves is rather smooth. Therefore, the coefficients of the third order polynomial a,b,c, and d can be parameterised with sufficient accuracy using a quadratic fit (see figure 20).

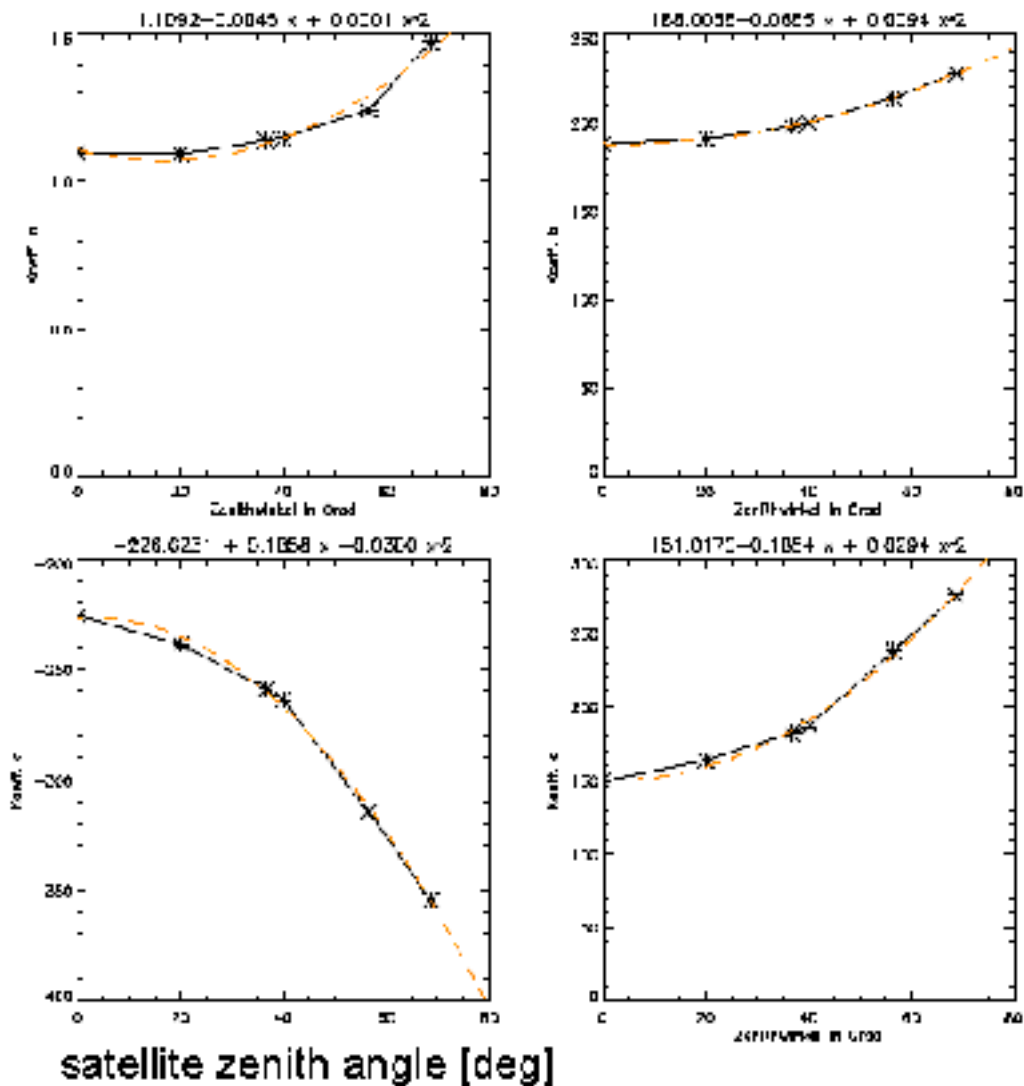


Figure 20: Parameterisation of fit function coefficients as function of solar zenith angle in degree. Coefficient a as function of solar zenith angle (upper left), coefficient b as function of solar zenith angle (upper right), coefficient c as function of solar zenith angle (lower left), coefficient d as function of solar zenith angle (lower right).

### 3.3. Realisation in the current HELIOSAT-3 software

To build up the current version of the water vapour algorithm the 'SCENES' software package was used as well. SCENES provides calibrated brightness temperature together with satellite geometry. Cloud detection is performed using the APOLLO module as described in chapter 2.

A separate program was written which implements the modified Kleespies & McMillan method as given above. In the following sections more details on this implementation are given.

#### 3.3.1. Input data

The modified Kleespies and McMillan method needs basically brightness temperatures in 10.8 and 12  $\mu\text{m}$  channels. Water body pixels and cloudy pixels have to be excluded. Two measurements of brightness temperatures are needed for the early morning and around noon.

Therefore, a specific input data format is generated during the cloud processing as described in section 2. For each time slot (every 15 minutes) two files are produced:

ch108\_YYYYDDMM\_HHMM\_segmentid\_LAAAABBBBECCCCDDDD.dat

ch120\_YYYYDDMM\_HHMM\_segmentid\_LAAAABBBBECCCCDDDD.dat

with the following file name structure (same structure as cloud product described in section 2.3):

YYYY = year in 4 digits

MM = month in 2 digits

DD = day in 2 digits

HH = hour in 2 digits

MM = minute in 2 digits

segmentid = HRIT segment number

(1 .. 8, 1 = southernmost HRIT segment, 8 = northernmost HRIT segment)

AAAA = first line of data (1 .. 3712, 1 = southernmost SEVIRI line)

BBBB = last line of data (1 .. 3712)

CCCC = first element of data (1 .. 3712, 1 = easternmost SEVIRI sample)

DDDD = last element of data (1 .. 3712)

For each pixel the following values are possible:

- 999           space

- 17            water body

- 100           cloud

positive values       brightness temperature in K \* 100

This is the basic input for the modified Kleespies and McMillan algorithm. After processing of a full day of MSG SEVIRI data these two files exist for each 15 minute time interval.

In a second step one has to choose automatically the 'early morning' and 'noon' time in UTC for each pixel. In general pixels in the East of the SEVIRI field of view will have a smaller UTC time as e.g. 'early morning time' as pixels in the West. This dynamic change

during the daily cycle of the sun is modeled in a hourly resolution. Furthermore, a minimum time distance between ,early morning‘ and ,noon‘ of 6 hours is required for HRIT segments 2 to 7. For the most northern and southern HRIT segments no. 1 and 8 a minimum time distance of only 4 hours is required – otherwise there would be no data available in the winter months for some areas. Last but not least, a missing data handling is included in the selection of suitable input files. The following steps are required to fulfill all requirements:

A hourly resolution in modeling the daily solar cycle corresponds to a 15 degree longitude interval. For each longitude interval an optimum UTC time corresponding to 6:00 mean local time (MLT) and 12:00 MLT is defined as parameter.

For each longitude interval it is checked if daylight has already reached the whole interval area at this optimum UTC time. If not, up to 3 further hours of SEVIRI data are checked for daylight availability. This step covers the yearly cycle of the sun and results in a list of ,early morning UTC times‘ suitable for the specific day.

Next, a check of availability of these ,early morning UTC times‘ in the SEVIRI data set is performed. If a slot is missing, an available slot is searched for up to 2 hours later. At the end of this step a list of ,existing early morning UTC times‘ is available.

As a first guess for the list of ,noon UTC times‘ the ,existing early morning UTC times‘ + 6 hours are selected. For HRIT segments 1 and 8 only 4 hours are required in order to get input data at all during winter time.

The availability of these ,noon UTC times‘ in the SEVIRI data set is checked. If a slot is missing, an available slot is searched for up to 2 hours later.

Finally, a list of ,existing early morning UTC times‘ and of ,existing noon UTC times‘ is created as input for the water vapour software.

### 3.3.2. Parameter settings in level 2 processing

Level 2 processing starts with the list of ,existing early morning and noon UTC times‘ and reads the corresponding ch108\_YYYYDDMM\_HHMM\_segmentid\_LAAAABBBBECCCCDDDD.dat and ch120\_YYYYDDMM\_HHMM\_segmentid\_LAAAABBBBECCCCDDDD.dat files (see section 3.3.1).

The relationship described in section 3.2 was implemented as follows:

$$ratio = \frac{\tau_{11}}{\tau_{12}} = \frac{(T_{11}^A - T_{11}^B)}{(T_{12}^A - T_{12}^B)}$$

with :  $T_{12}^A - T_{12}^B > 8K$  as minimum

$$arg = \left\{ \frac{1}{\sec\theta} \ln\left(\frac{\tau_{11}}{\tau_{12}}\right) \right\}$$

with : arg in  $[0., 0.8]$  as requirement (see figure 16)

*sun angle dependency of coefficients :*

$$a = 1.11 - 0.0045 * \theta + 0.0001 * \theta^2$$

$$b = 188.01 - 0.0685 * \theta + 0.0094 * \theta^2$$

$$c = -226.62 + 0.1858 * \theta - 0.03 * \theta^2$$

$$d = 151.02 - 0.1854 * \theta + 0.0294 * \theta^2$$

*calculation of total water vapour column :*

$$TWC = a + b * arg + c * arg^2 + d * arg^3$$

$T$  is the brightness temperature in Kelvin in channel 11 or 12  $\mu\text{m}$ ,  $\theta$  is the satellite zenith angle in radiant,  $A$  and  $B$  are two temporal different situations,  $\tau$  stands for transmission and TWC is total water vapour column in mm.

Values lower than 0 and above 80 mm are set to the missing value -99.

These level 2 results for total water vapour column are stored pixel- wise in  $\text{mm} * 100$  in a HDF formatted file.

### 3.3.3. Generation of level 3 water vapour product

It was agreed with the HELIOSAT-3 partners to deliver a gridded level 3 water vapour column product on a half degree latitude/longitude grid instead of a pixel- wise level 2 product. The following steps are performed to generate such an level 3 product:

Step 1: Average all level 2 TWC values inside a 0.5 x 0.5 degree latitude and longitude box. A flag is set to 1.

Step 2: Fill missing boxes with a local 3 x 3 boxes average to interpolate from the direct neighbours. The flag is set to 2.

Step 3: Fill still missing boxes with the background climatology. The flag is set to 3.

Boxes containing only ocean areas are left empty as the modified Kleespies and McMillan method works only over land.

The background climatology is derived from the NASA Water Vapor Project (NVAP) data set. This data set consists of merged data from NOAA- TOVS, SSM/I and radiosonde water vapour column measurements. Daily merged data sets from January 1<sup>st</sup>, 1988 to December 31<sup>st</sup>, 1997 were averaged to monthly mean values for a 1 degree grid from -89.5 to 89.5 degree North and 0.5 to 179.5 degree East. The NVAP web site ([http://eosweb.larc.nasa.gov/PRODOCS/nvap/table\\_nvap.html](http://eosweb.larc.nasa.gov/PRODOCS/nvap/table_nvap.html)) gives further information on the underlying data sets.

Experience with first MSG data sets show that the NVAP climatology is used only in large cloudy areas and in areas with missing input data. Both cases are acceptable because (a) a TWC climatology is sufficient for the cloudy irradiance retrieval as the cloud is the mainly influencing parameter in such a case and (b) missing input data also results in missing irradiance data anyway.

For a detailed format description of level 3 water vapour column data see the following section.

### 3.4. Water vapour level 3 file format

Data set specific attributes are included in each water vapour product. Each water vapour product is therefore fully described by both its filename and its internal attributes.

The filename structure used is the same as for the cloud product (please see section 2.3).

Figures 21 and 22 show attributes as written into the HDF4 format. These figures were taken as screenshots from the freely available JAVA HDF Viewer software package. The example given is a level 3 data set from January 17<sup>th</sup>, 2004 for HRIT segment no. 6 which covers in this case all MSG lines no. 2500 to 2784 and MSG elements no. 600 to 3000. This data set is part of the first test data sent to HELIOSAT-3 partners for further development of the surface irradiance calculation scheme. It was agreed to use a Europe subset of the data – therefore, only MSG elements from no. 600 to 3000 are used.



### 3.5. Example results

Figure 23 gives an example for MSG derived total water vapour column (level 2). Shown is the water vapour column distribution for March 17<sup>th</sup>, 2004 for the HRIT segment no. 7 and the subset from line no. 2783 to 3248 and element no. 600 to 3000 in pixel resolution (test subset as used so far in HELIOSAT-3).

Please note, that the modified Kleespies and McMillan method derives water vapour information only over land. This is in accordance with the data requirements defined in WP 2010.

Further data gaps are due to cloud cover at this specific day. Please refer to figure 24 for a colour composite of MSG input data for the water vapour algorithm. Cloudy areas in figure 24 can easily be found as data gaps in figure 23.

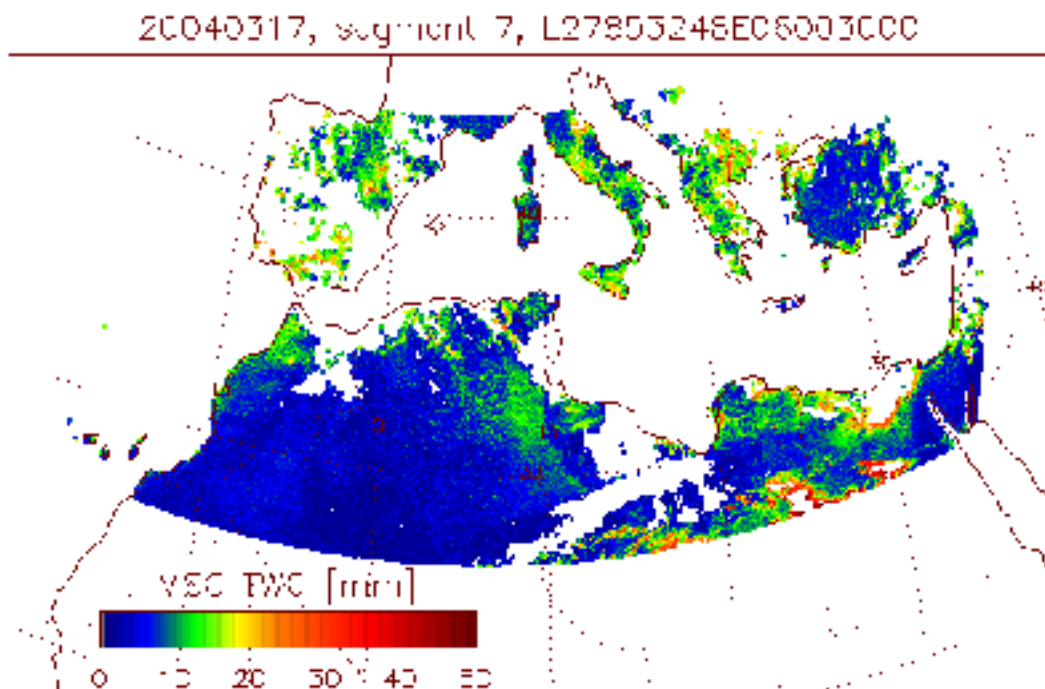


Figure 23: Example for MSG derived total water vapour column  
 March 17<sup>th</sup>, 2004, HRIT segment no. 7  
 (lines 2785 to 3248 and elements 600 to 3000)

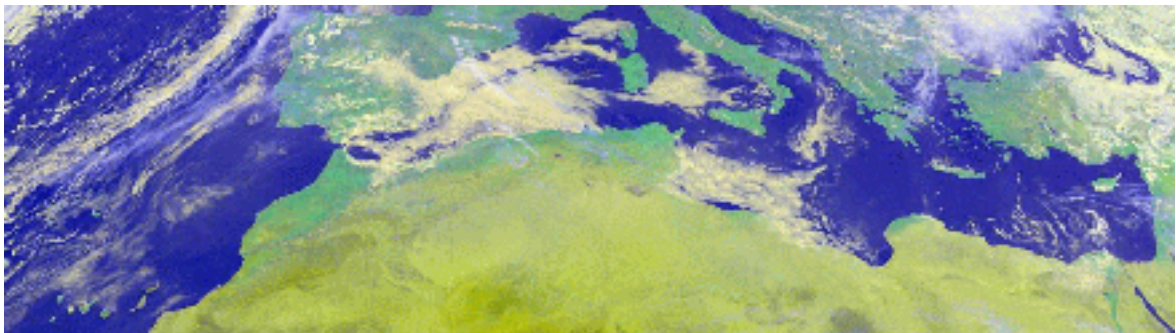


Figure 24: Colour composite of input data for example water vapour column product as shown in figure 23.

Figure 25 shows an example of the level 3 total water vapour column product. As described in section 3.3.3 this product is generated in a 3-step-approach: (1) average all level 2 data available in a 0.5x0.5° latitude/longitude grid box, (2) interpolate empty boxes locally, and (3) fill still empty boxes with a background climatology. The lower part of figure xxx shows the quality flag which is given for each grid box. Green values show where level 2 data is available and level 3 data was calculated through an average. Orange values show the local interpolation and red values show where a background climatology was used to complement the data set.

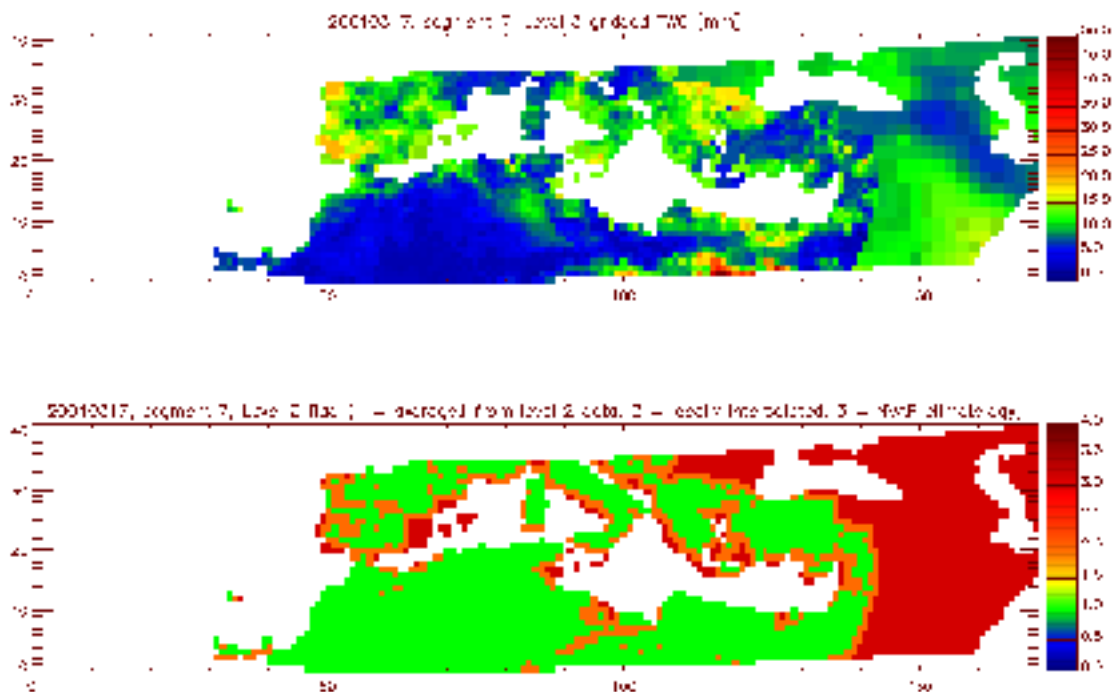


Figure 25: Example for level 3 water vapour column product for March 17<sup>th</sup>, 2004 for 26.75 ° to 47.25° North and from - 33.75° to 51.75° East (42 grid lines, 171 grid elements, 0.5° grid spacing)  
 upper figure: gridded total water vapour column in mm  
 lower figure: quality flag, green values = averaged from level 2 data,  
 orange values = interpolated from local neighbours,  
 red values = taken from background NVAP climatology

### 3.6. Further modifications for operational processing

As described already in section 2.5 (cloud products), MSG SEVIRI Level1.5 data is delivered from EUMETSAT via antenna in HRIT format separating the full disk data in 8 segments. This structure of HRIT segments is also kept in the DLR water vapour processing chain. Therefore, each cloud product for a single segment finds its corresponding water vapour level 2 and level 3 product for this segment.

## 3.7. References

- ABREU, L.W. & G.P. ANDERSON (Hrsg.) (1996): The MODTRAN 2/3 Report and LOWTRAN 7 MODEL.- Philips Laboratory, Hanscom.
- BERK, A., ANDERSON, G.P., ACHARYA, P.K., CHETWYND, J.H., BERNSTEIN, L.S., SHETTLE, E.P., MATTHEW, M.W. & S.M. ADLER-GOLDEN (1999): MODTRAN4 User's Manual.- Airforce Research Laboratory, Hanscom., URL: <http://www.cis.rit.edu/~dirsig/doc>.
- CHÉDIN, A., SCOTT, N.A., WAHICHE, C. & P. MOULINIER (1985): The improved initialisation method: a high resolution physical methods for temperature retrievals from satellites of the TIROS-N series.- *Journal of Applied Meteorology* 24:128- 143.
- CHESTERS, D., UCELLINI, L.W. & W.D. ROBINSON (1983): Low-level water vapour fields from VISSR Atmospheric Sounder (VAS) "Split-Window" channels.- *Journal of Climate and Applied Meteorology* 22:725- 743.
- CHESTERS, D., ROBINSON, W.D. & L.W. UCELLINI (1987): Optimized retrievals of precipitable water from the VAS "Split Window".- *Journal of Climate and Applied Meteorology* 26:10059- 1066.
- ECK, T.F. & B.N. HOLBEN (1994): AVHRR Split Window temperature differences and total precipitable water over land surfaces.- *International Journal of Remote Sensing* 15(3):567- 582.
- KLEESPIES, J.T. & L.M. MCMILLIN (1984): Physical retrieval of precipitable water using the Split Window technique.- *Conference on Satellite Meteorology, Remote Sensing and Applications (AMS)*, S. 55- 57.
- KLEESPIES, J.T. & L.M. MCMILLIN (1990): Retrieval of precipitable water from observations in the Split Window over varying surface temperatures.- *Journal of Applied Meteorology* 29:851- 862.

## 4. Parameterisation of aerosol amount and type

### 4.1. Introduction

A synergetic aerosol retrieval method (SYNAER), which exploits a spectrometer radiometer combination to retrieve aerosol optical thickness and type, has been developed and validated at DLR-DFD with GOME and ATSR-2, both onboard ERS-2 (Holzer-Popp et al., 2002a and 2002b). The methodology selects the most plausible type of aerosols from the remote sensing observations in each pixel without relying on any additional background data set.

Based on 14 months of data in 1997/1998 from GOME and ATSR-2 over the MSG observation area (due to GOME operations modes only for 3 days each month) a first climatological 5x5 degree data set has been derived, which describes annual average optical thickness values of major components (sulfate/nitrate, soot, dust, sea salt) of the atmospheric aerosol load. This data set is available at DLR for the partners to test the irradiance scheme.

A first validation of the method and recent investigations in the community have led to the definition of promising improvements in the aerosol model: Absorption features of soot and mineral components as well as the vertical profile of dust outbreak events require updating. A reprocessing of the climatology data set is under way. An improved climatology version will be provided to the partners soon.

The adaptation of SYNAER towards ENVISAT data is still an open task. Due to several delays in the ENVISAT ground segment it took some time to get SCIAMACHY and AATSR data from ESA. We succeeded in getting access to both NRT and offline delivery only during the last few weeks. Therefore, adaptation of the APOLLO scheme for AATSR cloud detection and adaptation of SYNAER towards SCIAMACHY and AATSR is now on DLR's agenda for summer 2004.

## 4.2. Aerosol parameterisation scheme

### 4.2.1. Current status of aerosol climatology data set

SYNAER (SYNergetic AERosol Retrieval, Holzer-Popp et al. 2002a) delivers boundary layer aerosol optical thickness (BLAOT) and type over both land and ocean, the latter as percentage contribution of 6 representative components from the OPAC (Optical Parameters of Aerosols and Clouds, Hess et al., 1998) data set to BLAOT. The high spatial resolution of the ATSR-2 instrument (Along Track Scanning Radiometer) permits accurate cloud detection, BLAOT calculation over automatically selected and characterised dark pixels and surface albedo correction for a set of 40 different pre-defined boundary layer aerosol mixtures. After spatial integration to the larger pixels of the spectrometer GOME (Global Ozone Monitoring Experiment) these parameters are used to simulate GOME spectra for the same set of different aerosol mixtures. A least square fit of these spectra to the measured spectrum delivers the correct BLAOT value and - if a uniqueness test is passed - the aerosol mixture. For humidity dependent components two models with 50% and 80% relative humidity have been included. Figure 26 gives an overview on the SYNAER method data flow.

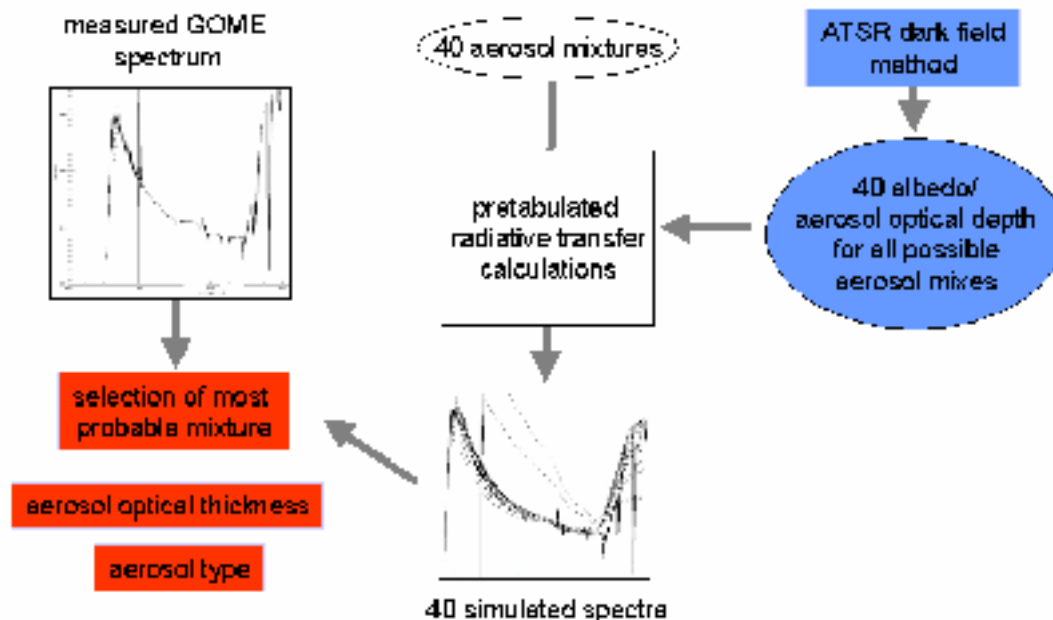


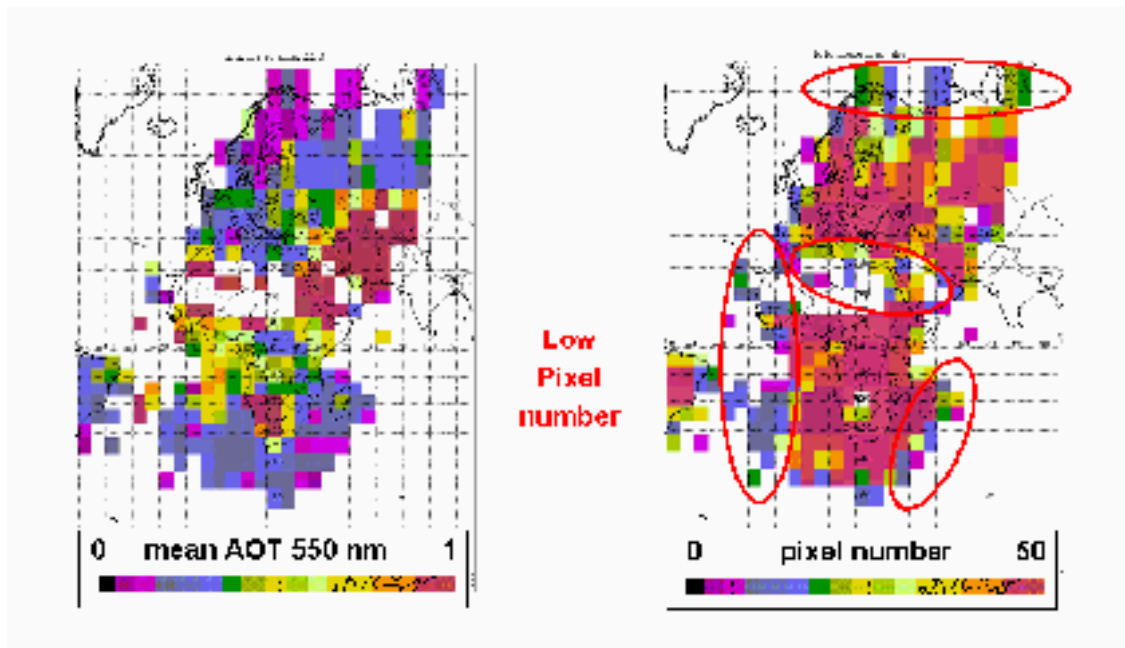
Figure 26: SYNAER method data flow as a combination of ATSR and GOME.  
 ATSR delivers aerosol optical depth and surface albedo but needs aerosol type.  
 GOME delivers aerosol type, but needs aerosol optical depth and surface albedo first.

Accurate cloud detection is an important prerequisite for each aerosol retrieval. The well established APOLLO (AVHRR Processing Scheme Over Cloud Land and Ocean; Kriebel et al. 2003) software was adapted to ATSR-2 data. The capability of retrieving cloud cover in

boxes of 1 km<sup>2</sup> means a significant strength of SYNAER because it reduces the erroneous aerosol detection due to the presence of sub pixel clouds significantly. It even allows the correction of partly cloudy GOME pixels. First case studies distributed over the globe using ground based sun-photometer measurements of the spectral aerosol optical thickness from NASA's Aerosol Robotic Network (AERONET) at 14 locations and one airborne lidar measurement show a good agreement with errors less than 0.1 in six wavelengths (340 – 870 nm) which indicates correct assessment of the amount and type (namely the spectral dependence of extinction) of aerosol (Holzer-Popp et al. 2002b). This first ground based validation comprised data from 4 continents in several climate zones (latitudes 17 South to 56 North) distributed over 2 years except the winter season with solar elevations ranging from 25 to 60 degrees. The cloud fraction inside the GOME pixels ranged up to 35%. Furthermore, a comparison of monthly mean results from SYNAER and other satellite aerosol retrievals as well as AERONET stations over ocean (Myhre et al., 2004) showed a satisfactory agreement with the other data sets.

It is the ultimate goal of this work to produce (and in future update) a satellite based aerosol climatology for the observation area of the European geostationary Meteosat Second Generation satellite. For this purpose the SYNAER method will be implemented and work operationally with the sensors SCIAMACHY and AATSR onboard ENVISAT. A backup climatology production based on GOME and ATSR-2 products (both onboard ERS-2) of the year 1997/98 is shown here. The final aim is to deliver 4 seasonal climatology data sets with a 1 degree horizontal grid. Due to cloud coverage and method inherent limitations a one year data set was obtained as first product for a 5 degree grid.

For this first application of the SYNAER method data products of the period July 1997 through August 98 covering Europe/Africa were received through the ESA AO project SENECA (AO ID 106). Unfortunately, GOME measures "small" pixels of 80x40 km<sup>2</sup> only for 3 days every month, and 320x40 km<sup>2</sup> pixels throughout the rest of the month. For producing the aerosol climatology only "small" pixels are meaningful in correspondance with horizontal variability scales of the aerosol loading and are thus exploited. With this small pixel mode GOME covers a swath width of 240 km with only 3 pixels in one scan line (SCIAMACHY will deliver pixels of 60x30 km<sup>2</sup> for a swath of 960 km, i. e. 16 pixels in one line). Although this means a severe limitation to the available data base, it opens the opportunity to test the methodology with a one year data set. Detailed handling of quality information from the retrieval process such as fit error, GOME-ATSR-2 cross-calibration deviation, spectral noise, surface elevation, solar elevation angle, etc. was optimized. Depending on the cloud coverage and method inherent limitations (surface brightness must not exceed 8% over land and 1.5 % over ocean, GOME pixel cloud fraction must not exceed 35%, differentiation of aerosol types is only reliable for optical thickness at 550 nm larger than 0.1, the ambiguity test rejects pixels with a fit error less than 0.005) extracting the seasonality was hardly possible with this first data set. Given the limited temporal coverage of 3 observations per month (which is further reduced due to cloudiness and technical errors) it was assured during averaging of the climatology values that a minimum of 2 orbits (i.e. 2 different times of observations) and of 5 pixels contributed to the box average (mean values for grid boxes with result values are 6 orbits and 41 pixels). To obtain a significant result a 14 month data set (average and standard deviation optical thickness, number of pixels contributing to one grid box, differentiation of optical thickness into basic components) using a total number of 18914 GOME pixels was produced as first product for a 5 degree grid.



*Figure 27: 14 months GOME/ATSR-2 aerosol climatology  
 Mean aerosol optical thickness is given on the left, pixel number used for  
 the climatology is given on the left. Areas with a small pixel number are marked in red.*

Figure 27 shows the climatology data set based on the 14 months of GOME/ATSR-2 data on a 5 degree grid. The low number of observations in some boxes (in particular in the South-West Atlantic and at the edges of the data set) leads to some unrealistic mean concentrations values which are determined by single episodes. This will be enhanced by using the daily ENVISAT observations.

However, despite of the small data base the aerosol distribution is rather smooth with a mean aerosol optical thickness of 0.26 and some interesting features can already be observed: Largest optical thickness values above 1 occur over / near the desert areas (fig. 29), whereas the Scandinavian area or oceanic zones far off from the continents show lowest values (fig. 28). Also biomass burning plumes from South America and Central Southern Africa are indicated over the Atlantic (fig. 30).

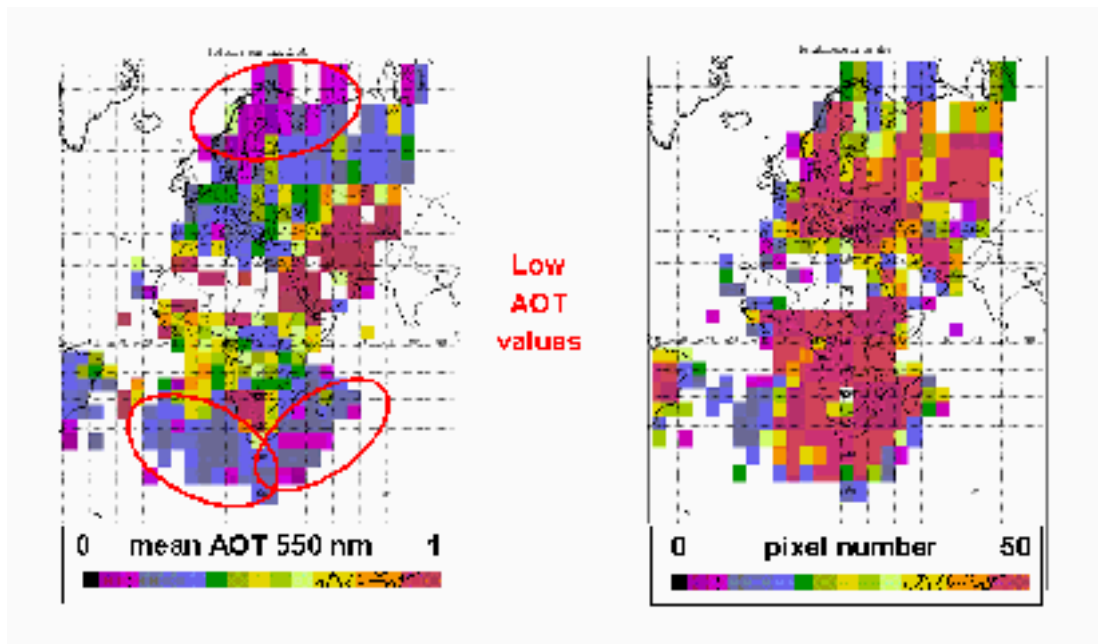


Figure 28: GOME/ATSR-2 aerosol climatology showing lowest values in the Scandinavian area or in oceanic zones. Available pixel numbers are given to show the significance of the result.

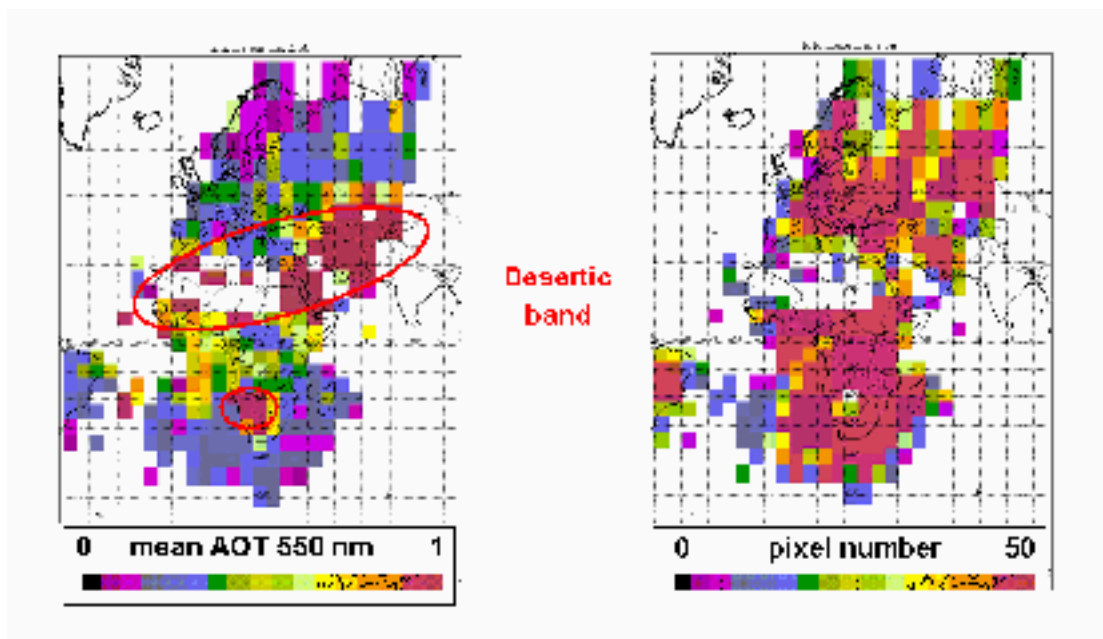


Figure 29: GOME/ATSR-2 aerosol climatology showing high optical thickness values in the desertic band. Available pixel numbers are given to allow the reader to judge on the significance of the result.

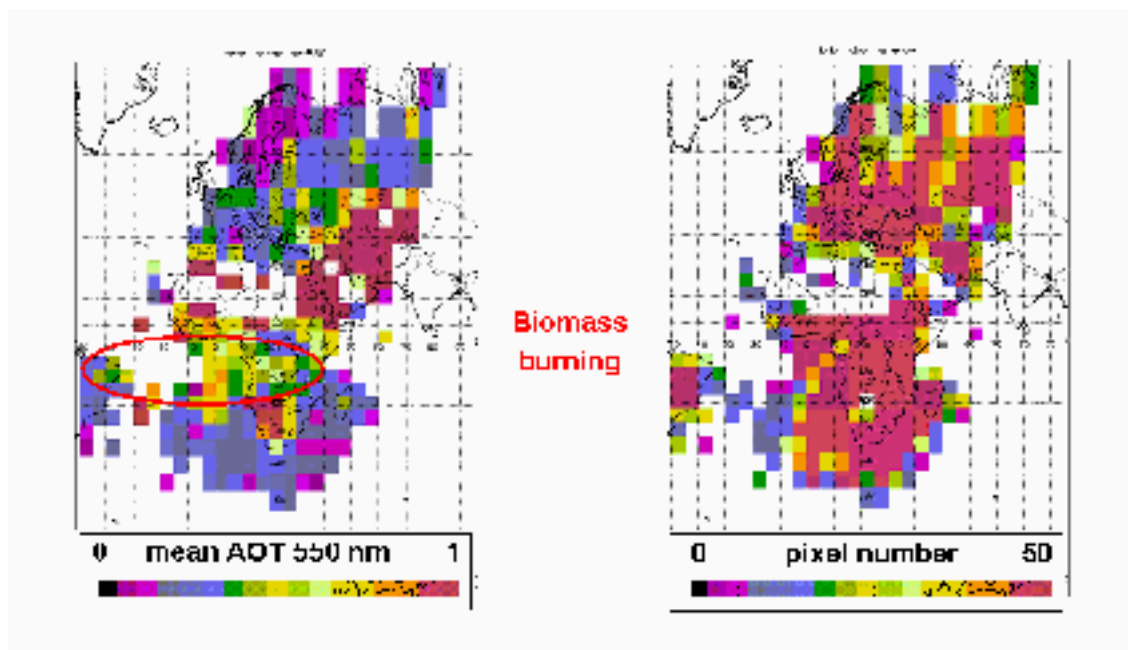


Figure 30: GOME/ATSR-2 aerosol climatology showing biomass burning plumes. Available pixel numbers are given as well.

Figure xxx shows also the component-wise mean aerosol optical thickness maps: Sulfate/nitrate aerosols which are included in all modelled aerosol types as background contribution show even clearer the unpolluted oceanic and Northern areas. Soot occurs most prominently over industrialised / densely populated areas in Central/Eastern Europe as well as over biomass burning source areas in Central/Southern Africa and in their plumes over the ocean (figure 31).

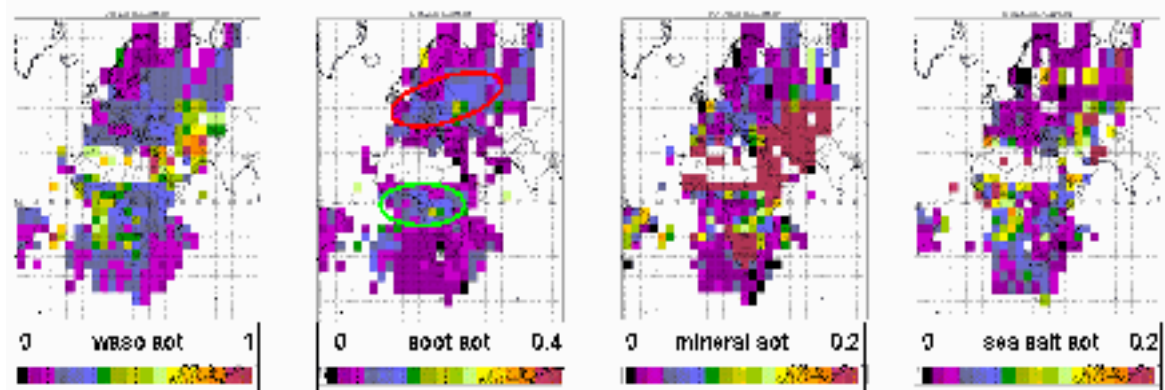


Figure 31: GOME/ATSR-2 aerosol climatology separated for watersoluble aerosols, soot, mineral dust and sea salt. Industrial areas and biomass burning areas can clearly be identified in the soot component.

Maximum 14 month 5 degree box average values are near 0.4. Few grid boxes show exceptionally high values which are due to a low number of contributing pixels. Dust is dominant in and around desertic areas (Sahara, Namib, Near East) and occurs with smaller amount in continental and oceanic aerosol. Sea salt occurs with low values also at inland locations which indicates to the limits of separating this component with low impact on the total optical thickness, but it should be noted that its maximum occurs over the

Southern Atlantic. Some areas are not covered by the basic data set because of either too high cloudiness or too bright surface. Due to miss-interpretation because of the simplified dust vertical structure high soot and sea salt values occur over deserts to some extent. Furthermore, an error in the land sea mask was found in the South-East-Atlantic which leads to some increased concentration levels over the ocean. Finally, new ground-based measurements revealed limitations in the soot and dust representation in the SYNAER aerosol model (see section 4.2.2), which are also a potential reason of wrong aerosol type detection.

As an outlook showing the usability of SYNAER to monitor the yearly and even daily variations, figure 32 shows a first attempt to evaluate the climatology for each season (3 months). Despite of the small number of available pixels so far, typical seasonal features can be detected. Aerosol optical thickness has a clear peak in Central Africa during the biomass burning season from August to November. This area is progressing South over the year. Another peak can be detected in the European aerosol optical thickness in summer. This corresponds to the expected longer lifetime of aerosols in the summer (due to decreased washing out by rain and stronger sun-light dependent sources as biomass and atmospheric organic chemistry). The winter behaviour of aerosols in Europe remains unclear, probably due to a larger cloud coverage and a lower sun in the winter time which further reduce the available pixel number. Especially for the seasonal description of aerosols a major step forward is expected from ENVISAT data.

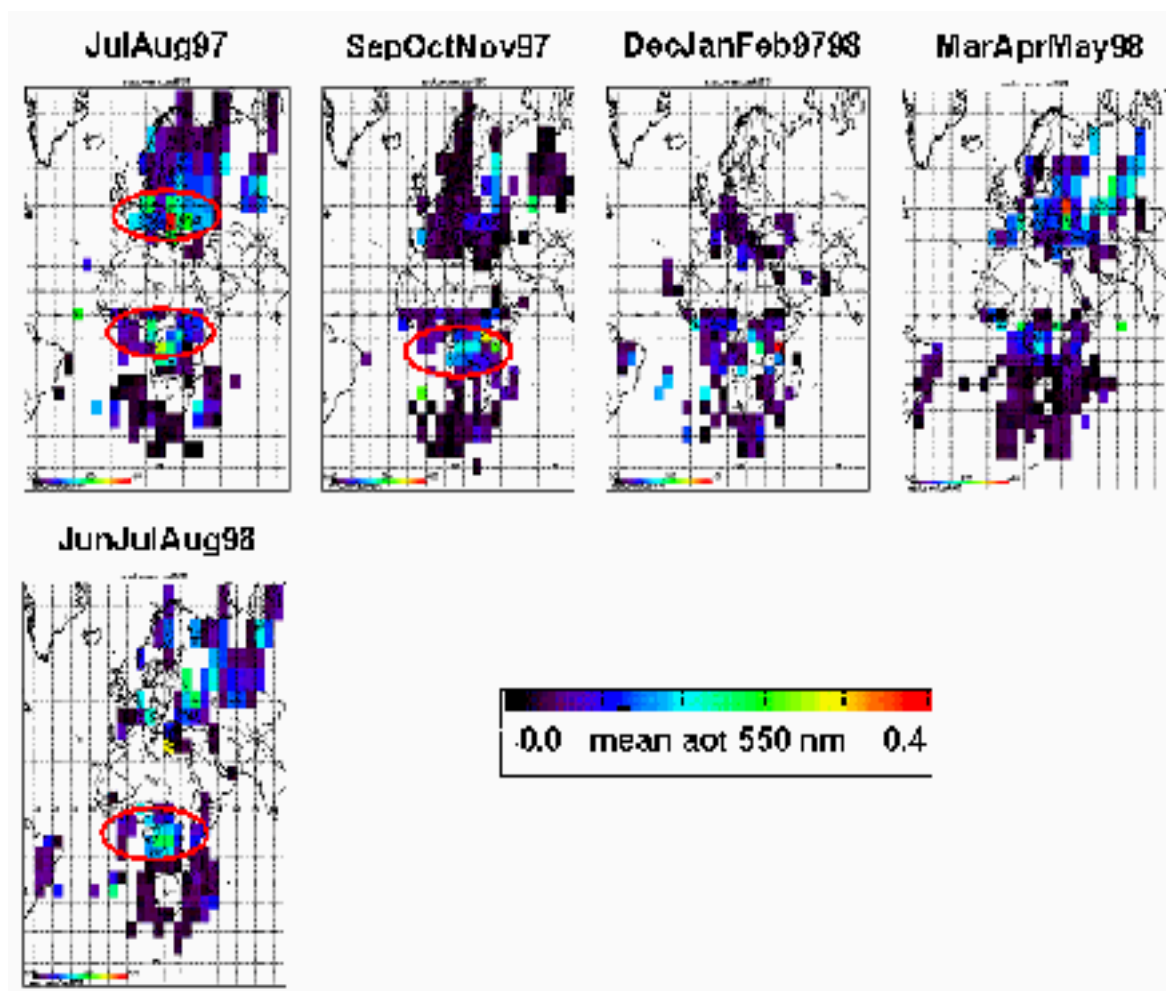


Figure 32: First seasonal evaluation of the aerosol climatology data set.

#### 4.2.2. Further development on the SYNAER method during HELIOSAT-3

Basically the components (which summarise the optical features of all particles with similar optical behaviour) including their log-normal size distribution are taken from the OPAC database (Hess et al., 1998). Table 1 summarises their relevant micro-physical properties and derived (Mie-calculated) optical characteristics. However, more recent campaigns and AERONET data exploitations have been used to improve some specific items:

- the original soot component was split in 2 components for weakly absorbing biomass burning (Dubovik, 2002) and very strongly absorbing diesel (industrial) carbon (Schnaiter, et al., 2003). Furthermore, the real part of the refractive index was adjusted to Dubovik, 2002 representing a mixture of Amazonian, South American cerrado, African savannah and boreal forest fires.
- in the absorption of mineral dust (transported and insoluble) 2 components were introduced to take the dust origin with different hematite content and consecutively absorption into account (Dubovik, 2002 representing an average of Cape Verde, Saudi Arabia and Bahrain observations and Sinyuk et al., 2003 below 400 nm). In the case of desert outbreak (transported minerals) the lowest aerosol layer of 4-6 km was modeled as two distinct sub-layers (dust above background), as they occur in nature.

Component	Species	Complex Refract. Index at 550 nm	Mode radius [µm]	Stand. Dev. of size dist.	Particle density [g/cm <sup>3</sup> ]	Extinction coefficient for 1 particle per cm <sup>3</sup> at 550 nm [km <sup>-1</sup> ]	Single scattering albedo at 550 nm	Literature source
WASO, rH = 70%	Sulfate/nitrate	1.53 – 0.0055 i	0.028	2.24	1.33	7.9 e-6	0.981	Hess, et al. 1998
INSO	Mineral dust, high hematite content	1.53 – 0.008 i	0.471	2.51	2.0	8.5 e-3	0.73	Hess, et al. 1998
INSL	Mineral dust, low hematite content	1.53 – 0.0019 i	0.471	2.51	2.0	8.5 e-3	0.891	Dubovik, et al. 2002
SSAM, rH = 70%	Sea salt, accumulation mode	1.49 – 0 i	0.378	2.03	1.2	3.14 e-3	1.0	Hess, et al. 1998
SSCM, rH = 70%	Sea salt, coarse mode	1.49 – 0 i	3.17	2.03	1.2	1.8 e-1	1.0	Hess, et al. 1998
BISO	Biomass burning soot	1.63 - 0.036 i	0.0118	2.0	1.0	1.5 e-7	0.698	Dubovik, et al. 2002
DISO	Diesel soot	1.49 – 0.67 i	0.0118	2.0	1.0	7.8 e-7	0.125	Schnaiter, et al. 2003
MITR	Transported minerals, high hematite content	1.53 – 0.0055 i	0.5	2.2	2.6	5.86 e-3	0.837	Hess, et al. 1998

MILO	Transported minerals, low hematite content	1.53 – 0.0019 i	0.5	2.2	2.6	5.86 e-3	0.93	Dubovik, et al. 2002
------	--	-----------------	-----	-----	-----	----------	------	----------------------

Table 1: updated aerosol components (*new components are highlighted*)

WASO = watersoluble, INSO = insoluble, INSL = insoluble / low hematite,

SSAM = sea salt accumul. mode, SSCM = sea salt coarse mode, BISO = biomass burning soot,

DISO = diesel soot, MITR = mineral transported, MILO = mineral transported / low hematite

Table 2 shows the (updated) definition of the 40 mixtures used in the SYNAER retrieval method. Values in the table show the vertical profile, relative humidity in the boundary layer and the percentage contribution to the optical thickness at 550 nm of the respective components. Two groups of 20 mixtures, each are applied where either relative humidity or the hematite content of the mineral component (and consecutively the absorption) are altered. Differences between these two groups are marked with grey boxes in the table. E.g. mixture no. 1 has 50% relative humidity while mixture no. 21 has a relative humidity of 80%. An example for the low/high hematite content can be found e.g. in mixture no. 2 (5% insoluble with high hematite content) and in mixture no. 22 (5% insoluble with low hematite content). The set of 40 mixtures is meant to model all principally existing aerosol types and allow for some variability of the composition of each type. These set of mixtures has proven to provide a fit in the GOME spectra retrieval which is in many cases better than a 1% noise level. Work to determine the success rate of the retrieval approach using this set of mixtures with a 14 month data set is currently going on. This improved version of the aerosol climatology will be provided to the partners as soon as possible.

				Component contributions to AOT550 [%]										
No.		Name	Rel. Hum [%]	Vert. Prof. [km]	WASO	INSO	INSL	SSAM	SSCM	BISO	DISO	MITR	MILO	
1	21	Pure watersoluble	50/80	2	100									
2	22	Continental	50	2	95	5	5							
3	23				90	10	10							
4	24				85	15	15							
5	25	Maritime	50/80	2	30			70						
6	26				30			65	5					
7	27				15			85						
8	28				15			75	10					
9	29	Polluted watersoluble	50/80	2	90						10			
10	30				80					20				
11	31	Polluted Continental	50	2	80	10	10				10			
12	32				70	10	10			20				
13	33	Polluted Maritime	50/80	2	40			45	5		10			
14	34				30			40	10		20			
15	35	Desert Outbreak	50	2 - 4	25							75	75	
16	36			3 - 5	25								75	75
17	37			4 - 6	25								75	75
18	38	Biomass Burning	50/80	3	85					15				
19	39				70					30				
20	40				55					45				

WASO = watersoluble, INSO = insoluble, INSL = insoluble / low hematite, SSAM = sea salt accumul. Mode, SSCM = sea salt coarse mode, BISO = biomass burning soot, DISO = diesel soot, MITR = mineral transported, MILO = mineral transported / low hematite; Mixture number N and mixture number N + 20: alternative humidity or mineral composition, respectively

Table 2: updated aerosol mixtures

### 4.3. Aerosol file format

Aerosol parameters are delivered in HDF4 format as well.

For the first HELIOSAT-3 version the aerosol climatology is used. It is envisaged to use a daily updated aerosol product if ENVISAT aerosol products are available in future (approx. end of 2004). Therefore, the final aerosol file format cannot be described here. A further update of this section is foreseen until the end of the project in order to keep this deliverable as an up-to-date software documentation.

Nevertheless, it has already been tested and confirmed that the following aerosol parameters will be used as interface to the irradiance calculations part of HELIOSAT-3:

- Aerosol optical thickness at 550 nm to describe the aerosol amount
- Angstrom exponents  $\alpha$  and  $\beta$  to describe the aerosol type

These parameters can be used directly as input for the libradtran software package included in the irradiance calculations part. The libradtran input parameters needed are called 'aerosol\_angstrom' and 'aerosol\_set\_tau550'.

## 4.4. References

Dubovik, O., B. Holben, T.F. Eck, A. Smirnov, Y.J. Kaufman, M.D. King, D. Tanre, I. Slutsker, Variability of Absorption and optical Properties of Key Aerosol Types Observed in Worldwide Locations, *J. Atm. Sciences*, Vol 59, 590 – 608, 2002

Hess, M., Köpke, P., Schult, I., Optical Properties of Aerosols and Clouds: The Software package OPAC, *Bulletin of the American Meteorological Society*, 79, pp. 831- 844, 1998

Holzer-Popp, T., M. Schroedter, and G., Gesell, Retrieving aerosol optical depth and type in the boundary layer over land and ocean from simultaneous GOME spectrometer and ATSR-2 radiometer measurements, 1, Method description *J. Geophys. Res.*, 107, D21, pp. AAC16- 1 – AAC16- 17, 2002a

Holzer-Popp, T., M. Schroedter, and G., Gesell, Retrieving aerosol optical depth and type in the boundary layer over land and ocean from simultaneous GOME spectrometer and ATSR-2 radiometer measurements, 2, Case study application and validation, *J. Geophys. Res.*, 107, D24, pp. AAC10- 1 – AAC10- 8, 2002b

Kriebel K. T., Gesell G., Kästner M., Mannstein H., The cloud analysis tool APOLLO: Improvements and Validation, *Int. J. Rem. Sens.*, 24, 2389- 2408, 2003

Myhre, G., Stordal, F., Johnsrud, M., Diner, D. J., Geogdzhayev, I. V., Haywood, J. M., Holben, B., Holzer-Popp, T., Ignatov, A., Kahn, R., Kaufman, Y. J., Loeb, N., Martonchik, J., Mishchenko, M. I., Nalli, N. R., Remer, L. A., Schroedter-Homscheidt, M., Tanre, D., Torres, O., Wang, M., Intercomparison of satellite retrieved aerosol optical depth over ocean during the period September 1997 to December 2000, submitted, *JGR*, 2004

Schnaiter, M., H.Horvath, O. Möhler, K.-H. Naumann, H. Saathoff, O.W. Schöck, UV-VIS-NIR spectral optical properties of soot and soot-containing aerosols, *J. Aerosol Science*, 34, 1421- 1444, 2003

Sinyuk, A., O. Torres, O. Dubovik, Combined use of satellite and surface observations to infer the imaginary part of refractive index of Saharan dust, *GRL*, Vol 30, No 2, 1081, doi 10.1029/2002GL016189, 2003

Stedman J. R., The predicted number of air pollution related deaths in the UK during the August 2003 heat wave, *Atmos. Env.*, 38, 1087 – 1090, 2004

## 5. Retrieval of ozone

### 5.1. Introduction

Ozone total column data is provided by the instruments ERS-2 GOME, EP-TOMS and ENVISAT-SCIAMACHY. This degree of redundancy ensures a robust backup system and further allows for cross validation.

The Global Ozone Monitoring Experiment (GOME) has been operating successfully aboard ESAs ERS-2 satellite since 1995. Due to the failure of the tape recorder the data coverage has been limited since July 2003. A second GOME instrument (GOME-2) will be launched on the polar orbiter METOP in 2005.

In March 2002 an enhanced GOME was successfully brought into its near-polar synchronous orbit as payload of ENVISAT: the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) is currently delivering ozone vertical columns and vertical profiles. The expected lifetime of SCIAMACHY is 5 years.

The Total Ozone Monitoring Spectrometer (TOMS) has been provided daily global information on the total ozone distribution since 1978 on several platforms. Currently, a TOMS instrument is operated on NASAs Earth Probe platform. It was launched in 1996. The mission of the successor QuikTOMS failed during launch in September 2001. The launch of the Ozone Monitoring Instrument (OMI) on NASAs EOS-AURA satellite is currently foreseen for 2004.

Measurements of these three instruments can be used within HELIOSAT-3 to derive synoptic distributions and analyses of total ozone in near-real time.

### 5.2. Retrieval of Total Column Ozone

The above mentioned instruments GOME, SCIAMACHY and TOMS perform backscatter measurements in the ultraviolet and visible portion of the solar spectrum. The retrieval of the total column ozone content from these measured spectra is described in this section for each instrument.

#### 5.2.1. GOME

GOME is a nadir-looking across-track scanning spectrometer with a typical footprint size of about 320 x 40 km<sup>2</sup>. It is able to measure the atmospheric column content of a number of minor trace constituents with the focus on ozone between 240 nm and 793 nm. The spectral resolutions varies from 0.2 nm to 0.4 nm (ESA, 1996).

The core element of the retrieval is a DOAS (Differential Optical Absorption Spectroscopy) fitting technique that involves a multi-linear regression of GOME-measured optical densities against a number of reference spectra. It provides trace gas column amounts along the viewing path of the instrument (Burrows et al., 1998a). The differential absorption of trace species along the absorption path is modelled on Beer's law

$$dI(\lambda) = -I(\lambda) \cdot \sigma(\lambda) \cdot C(s) \cdot ds$$

where the incremental increase of intensity  $dI(\lambda)$  at wavelength  $\lambda$  through a slant path distance  $ds$  is proportional to the absorption coefficient  $\sigma(\lambda)$  times the intensity  $I(\lambda)$  and the absorber column amount  $C(s)$ . When there are several absorbers, the contributions are

additive. The ozone retrieval was performed between 325 nm and 335 nm. For ozone fitting, the set of reference spectra is used including a Ring spectrum and an undersampling spectrum (Thomas et al, 2003 and references therein). The resulting trace gas slant columns are then converted to geometry- independent vertical column amounts through division by appropriate air mass factors (AMF). Air mass factors describe the enhancement of the absorption of a given trace gas due to slant paths of incident light in the atmosphere.

$$AMF = \frac{\tau_{slant}}{\tau_{vert}}$$

where  $\tau_{slant}$ ,  $\tau_{vert}$  are the slant path and the vertical optical density, respectively. The slant optical densities, i.e. the AMFs are derived from radiative transfer calculations and do not need input from GOME measurements other than viewing geometry and geolocation information.

Ozone AMFs were calculated on the basis of the TOMS V7 ozone profile climatology (McPeters et al., 1996) using LIDORT (Spurr et al., 2001) and are restored from look-up tables by neural network techniques (Loyola, 1999). The total ozone content is derived from an iterative procedure (Spurr, 1999), that searches for the best suited ozone profile, i.e. the AMF as function of the integrated profile content (and a number of other geophysical parameters). It is followed up to a given threshold of  $10^{-4}$ .

Since GOME pixels cover a relatively large area, their footprints will often be partially or totally covered by clouds. Clouds are opaque in the UV/VIS spectral range (except optically thin cirrus clouds), which needs to be taken into account if trace gas total columns are retrieved. The cloud coverage is determined from GOME measurements and simulations in and around the oxygen A- band between 758 nm and 778 nm (Kuze and Chance, 1994).

A common way of calculating the vertical column density (VCD) of a species is to take the weighted average of clear sky and cloudy contributions, assuming a so-called ghost vertical column (GVC) below clouds that is derived from climatological trace gas profiles.

$$VCD = \frac{SC + f_c \cdot GVC \cdot AMF_c}{f_c \cdot AMF_c + (1 - f_c) \cdot AMF_g}$$

where  $f_c$  is the cloud coverage between 0 and 1, SC the fitted slant column, GVC the ghost vertical column, and  $AMF_c$ ,  $AMF_g$  the air mass factors down to ground and to cloud-top. A recent validation campaign (Lambert et al., 2002) of GOME data using TOMS and ground-based measurements indicates that the agreement of total ozone column is within 2-4% for solar zenith angles below  $70^\circ$ .

GOME level-1 (radiances / reflectances) and level-2 (trace gas amounts) products are computed from level-0 (raw) data using the GOME Data Processor (GDP) system which was designed and developed by the German Remote Sensing Data Center (DFD) of DLR on behalf of ESA in co-operation with other institutions involved in the GOME Science Advisory Group (DLR, 1996).

For this project GOME GDP level 2 data Version 3.0 from ESA/DLR are used, derived by the procedure outlined above.

### 5.2.2. SCIAMACHY

SCIAMACHY is an imaging spectrometer performing measurements in [nadir](#), [limb](#), and [solar/lunar occultation geometry](#). It measures atmospheric absorption in spectral bands from the ultraviolet to the near infrared (240 nm - 2380 nm). From these observations total column amounts

and stratospheric profiles of a multitude of atmospheric constituents are retrieved on a global scale (DLR, 2000).

In the nadir scanning mode, which is relevant for the HELIOSAT-3 project, SCIAMACHY has a typical footprint size of about 60 x 30 km<sup>2</sup>. Since SCIAMACHY has a similar nadir-scan measurement capability to GOME, the experience gained with GOME in the implementation of operational nadir UV-visible total column retrieval algorithms has proved essential for the development of corresponding algorithms in the SCIAMACHY level 0 to 2 processing elements. Thus, the main algorithm used within the GOME Data Processor, the DOAS fitting approach for the retrieval of O<sub>3</sub> columns is used directly for SCIAMACHY NRT applications. This approach is described in the section above. Main improvements, however, concern the implementation of an enhanced cloud algorithm. For further details see DLR (2000).

Currently, SCIAMACHY data of Version 5.01 is available, which is still subject of improvement and ongoing validation by the SCIAMACHY Validation Group (SCIAVALIG).

### 5.2.3. TOMS

The Total Ozone Mapping Spectrometer (TOMS) samples backscatter UV at six wavelengths and provides a contiguous mapping of total column ozone. Its nadir angle field of view is about 39 km on a side (McPeters et al. 1996, NASA, 2000).

The ozone retrieval is based on a normalised radiance, the ratio of the radiance backscattered by the Earth and atmosphere to the incident solar irradiance. Retrieval of total ozone comprises of a comparison between the measured normalised radiances and radiances derived by radiative transfer calculations for different ozone amounts and the conditions of the measurement. It is implemented by using radiative transfer calculations to generate a table of backscattered radiance as a function of total ozone, viewing geometry, surface pressure, surface reflectivity, and latitude.

Since TOMS data retrieval is not performed at DLR please see for further details NASA (1998).

For Earthprobe TOMS total ozone, the absolute error is  $\pm 3$  percent, the random error is  $\pm 2$  percent (though somewhat higher at high latitudes). The drift after 1.5 years of operation was less than  $\pm 0.6$  percent, but currently the instrument is experiencing scan bias problems. Within HELIOSAT-3 TOMS Data of Version 8.0 is used.

## 5.3. Level 3 products

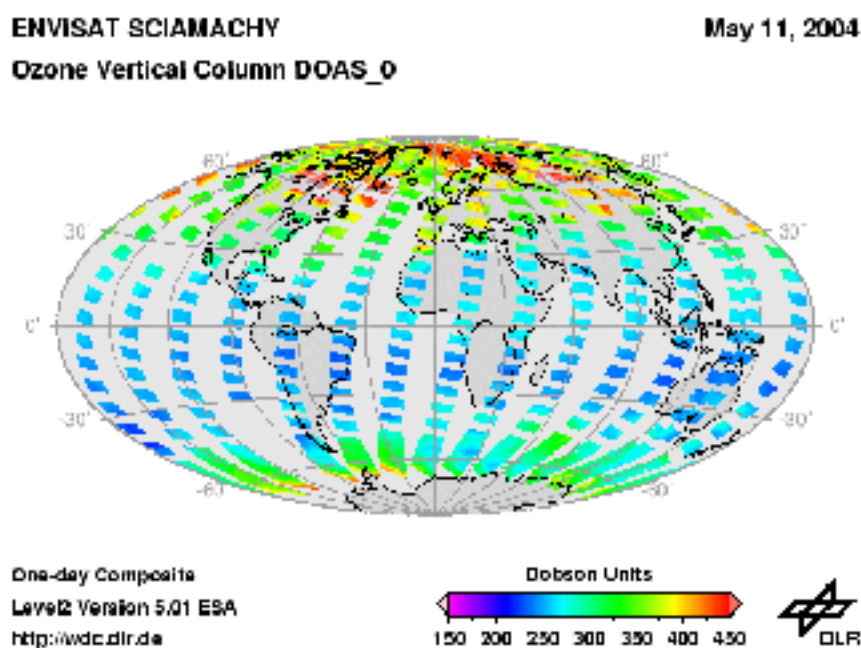
Due to the scanning geometry of the above mentioned spectrometers and the orbital parameters of the platforms they are mounted, the total column ozone observations are heterogeneously distributed in time and space. The sequential scanning along the orbit track implies that all observations are recorded at a different time.

Additionally, TOMS achieves global coverage after 2 days, GOME after 3 days and SCIAMACHY after 6 days only (depending on the latitude). It is quite obvious that the information content from averaged synoptic maps is rather limited and would result in artifacts. This is because the distribution of total column ozone is primarily controlled by dynamic processes in the atmosphere, which can significantly modulate the distribution on short time scales. Already Reed (1950) showed the high correlation between total ozone and high and low pressure systems or fronts in the troposphere.

To gain synoptic distributions of total column ozone considering atmospheric variability two different data assimilation techniques are applied depending on the footprint pattern and accordingly the temporal and spatial coverage of the observations.

For TOMS and GOME a simple spectral statistical approach is chosen, which proved to be well suited, since these instruments provide sufficient daily coverage.

Unlike TOMS and GOME, SCIAMACHY's daily coverage is sparse, since the instrument switches permanently from the nadir to the limb mode. This unique combination of measurement modes results in the atypical footprint pattern shown in figure 33. To countervail this peculiarity the SCIAMACHY observations are sequentially assimilated into a 3D Chemistry Transport Model, describing all relevant chemical and physical processes.



*Figure 33: Daily composite of SCIAMACHY total ozone observations for May 11, 2004. Note the unique footprint pattern due to the permanent switch between the nadir and limb mode.*

### 5.3.1. GOME and TOMS

In order to derive synoptic total column ozone fields from TOMS and GOME observations a spectral statistical approach is applied (Bittner and Erbertseder, 2002). It is based on the assumption that the total column ozone variability is mainly governed by planetary waves and thus can be described by a linear combination of harmonic functions. To consider the time and space dependency of the planetary waves the Fourier coefficients are estimated by a Kalman-Filter (Daily, 1999).

The approach makes use of the fact that about 90-95% of the total ozone column stems from altitude regions from 10 km up to about 40 km. In this region, ozone can be regarded as being photochemically stable for time scales of a few days and weeks. Therefore, the distribution of total column ozone is mainly due to transport processes, that is dynamics. Several wave phenomenos on different time and spatial scales are governing the stratospheric dynamics. Examples are gravity waves, shorter-period

planetary waves, Kelvin and mixed Rossby-gravity waves near the equator, large-scale circulation, and surges above thunderstorms, cyclones or over sea surface temperature anomalies such as El Niño, to mention only a few.

Since the total ozone distribution is governed by waves, knowing the waves means knowing the ozone distribution. Therefore, the waves are modelled by means of a linear combination of harmonic functions.

The procedure starts in binning GOME or TOMS level-2 data into latitude segments. To obtain information on the temporal variability of total ozone the satellite observations from previous days are taken into account. This results in an ensemble of zonal data/time series of total column ozone. The Kalman filter now estimates the fourier coefficients for the trigonometric functions describing the planetary waves at each latitude bin a given time  $t$ . Hence, a synoptic ozone distribution can be derived for any given time ("snapshot") (Figure 34).

The state space approach Kalman-Filtering is capable to take into account even rapid changes in the atmospheric system and it is independent on the measurement geometry, that means the observations need not to be equidistant. Additionally, the presented approach is tolerant in treating data gaps. Generally, it has to be emphasised that no auxiliary data like global wind fields are needed. As validation with ground based measurements has shown the approach doesn't add significant noise to the data at all, while gaining a much better data availability compared to the pure level-2 data (Bittner and Erbertseder, 2002).

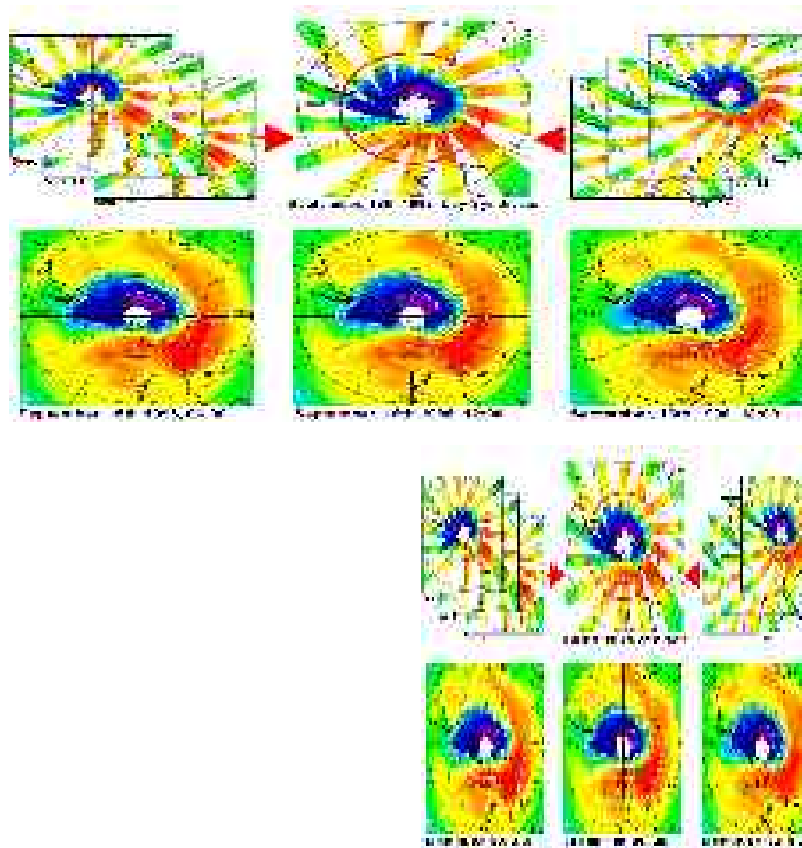
*Figure 34: Rough outline of the data assimilation approach as applied for TOMS and GOME data (here: GOME). Note the spectral statistical analysis of the ozone distribution for each latitude. To consider the temporal and spatial variation of the ozone field data of the previous days is used. The Fourier coefficients for each latitude ensemble are estimated by a Kalman Filter.*

### 5.3.2. SCIAMACHY

To account for the scan properties of SCIAMACHY its total column observations are sequentially assimilated into a 3D Chemical Transport Model (Erbertseder, 2004, Erbertseder et al., 2003b).

We use the 3D global chemical-transport-model DLR-ROSE 3.0. It is based on the ROSE model described in detail in Rose and Brasseur (1989) and Riese et al. (1999), but was significantly improved. The model covers all relevant gas-phase stratospheric chemical processes.

Heterogeneous processes are also included in the model. It accounts for about 100 reactions, including oxygen, hydrogen, carbon, nitrogen chlorine, and bromine species. The basic chemical time-step is 15 min. All species are transported every 90 minutes using a Lin and Rood scheme. The model is driven by wind- and temperature fields of the European Center for Medium Range Weather Forecast (ECMWF). The spatial discretisation of the DLR-ROSE 3.0 uses a  $2.8^\circ \times 2.5^\circ$  lon.- lat. spherical grid and 43 log-pressure levels between 0 and 56 km altitude (316. to 0.316 hPa) resulting in a vertical step size of 1.3 km. Further details are given in Table 3.



<b>Version</b>	<b>DLR 3.0</b>
Met. analysis	ECMWF
Resolution	43 layer, 0- 56km 2.8°•2.5° lon- lat grid
Gas chemistry	non- family concept
solver	Gauss- Seidel, Newton- Raphson
reaction rates	JPL2002
photolysis rates	Stamnes, 2003
Source gases	CFC11, CFC12, CH3Br
Heterogeneous chemistry	ICE, NAT, STS
Advection	Lin and Rood
Troposphere	Convection O3 relax. to climat. mean
Data assimilation	OI sequential
Data formats	HDF, NetCDF
Language spec.	F90, openMP, CRAY

--	--

Table 3: Model specifications of DLR-ROSE 3.0

Given the CTM as described in table 3, the SCIAMACHY data is sequentially assimilated (fig. 35) using an optimal interpolation scheme. It considers the time of observation, a spatial weighting function between observation and the grid and the errors of the model and the observation. Details on the assimilation procedure with GOME observations are given in Bittner et al (2004).

Globally gridded 3D fields will be available every 6 hours. A detailed report on the errors will be given in the HELIOSAT-3 validation report.

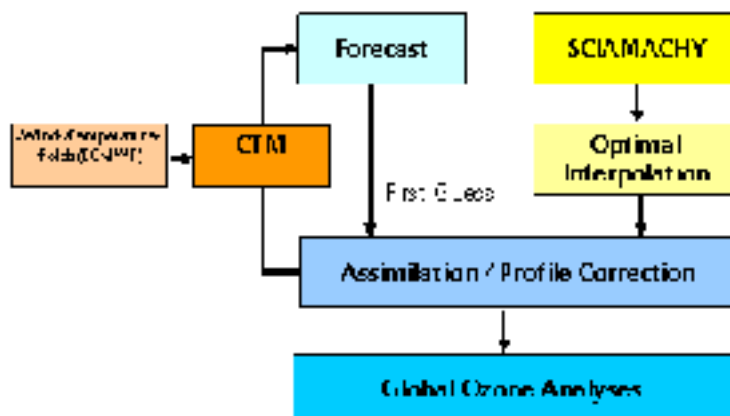


Figure 35: Data assimilation scheme of SCIAMACHY total ozone data.

### 5.3.3. Gridding to HELIOSAT-3 format

In a further step the total ozone analyses are regridded to the resolution as defined within the HELIOSAT-3 project. This guarantees simple interfacing between the workpackages.

In order to remap the data to the specified 0.5° x 0.5° latitude-longitude grid a nearest neighbour algorithm is used. It assigns an average value to each node that have one or more points within a radius centered on the node. The average value is computed as a weighted mean of the nearest point from each sector inside the search radius. The weighting function used is

$$w(r) = 1.0 / (1 + d^2)$$

where  $d = 3 * r / \text{search\_radius}$  and  $r$  is distance from the node. This weight is modulated by the original data points.

We use grid line registration where the nodes are centered on the grid line intersections and the data points represent the average value in a cell centered on the nodes (figure 36). For remapping the data we use routines provided by the Generic Mapping Tools (GMT, 2004).

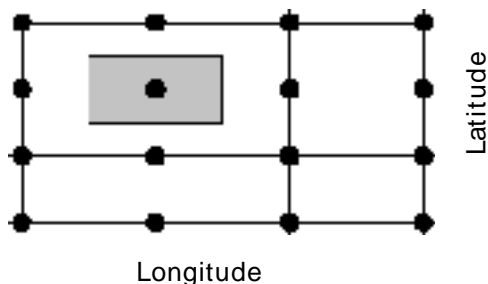


Figure 36: Grid line registration of data nodes (GMT, 2004)

### 5.3.4. Near- Real-Time Processing chains

#### 5.3.4.1. GOME processing chain

The GOME processing chain consists of data reception, geolocation and calibration of spectra, retrieval of trace gases and finally data assimilation to derive synoptic global total column ozone distributions (figure 37).

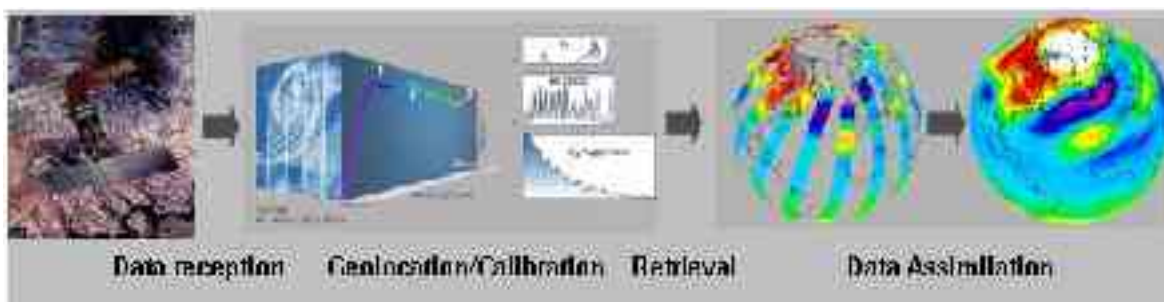


Figure 37. The GOME near- real- time processing chain at DFD.

The near-real-time access to GOME observation is achieved by the Global Ozone Network from GOME (N-GONG). It is a national financed joint collaboration between the German Aerospace Center (DLR/IMF and DLR/DFD) and the Institute of Environmental Physics (IUP) at the University of Bremen in co-operation with the European Space Agency (ESA). The N-GONG project provides an extreme cost effective near-real-time service utilising all the information within the GOME data. The GOME near-real-time service was initiated in January 1997 with the installation of GDP at the Kiruna station (10 out of 14 daily orbits) and completed in February 2002 with the GDP installation at the Maspalomas station. The GOME Level 2 data are available 3 hours after acquisition.

#### 5.3.4.2. TOMS processing chain

TOMS level 2 data is provided directly from NASA in near-real-time on a ftp server. After complete acquisition a synoptic ozone map is derived applying the spectral statistic data assimilation described above. This results in an operationally generated synoptic total ozone analysis.

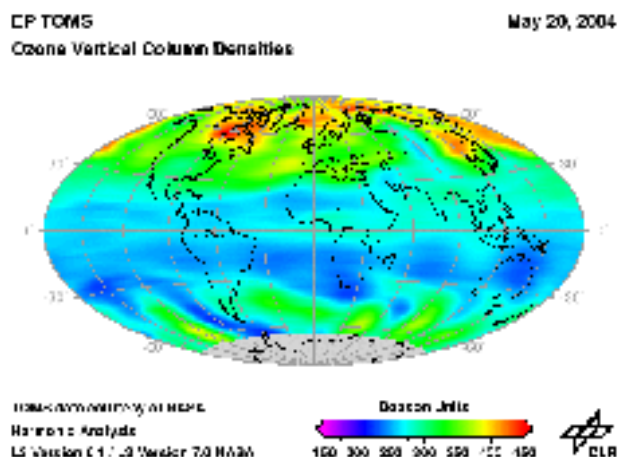


Figure 38: Example of an operationally generated TOMS total ozone analysis product for May 20, 2004 at 12 UT.

### 5.3.4.3. SCIAMACHY processing chain

The SCIAMACHY processor up to level 2 was developed by DLR on behalf of ESA. For NRT applications it is operated in Kiruna, Sweden and at ESRIN, Frascati Italy. NRT access to the SCIAMACHY level 2 data is guaranteed by ftp access to the processor outputs (Erbertseder, 2003a).

The ECMWF wind and temperature analyses are kindly provided by the European Center for Medium Range Weather Forecast.

The level 2 data (SCIAMACHY observations) are then operationally blended by a data assimilation approach into the 3D CTM DLR-ROSE 3.0 which is driven by the wind and temperature fields from ECMWF (figure 38) (Erbertseder, 2003b).

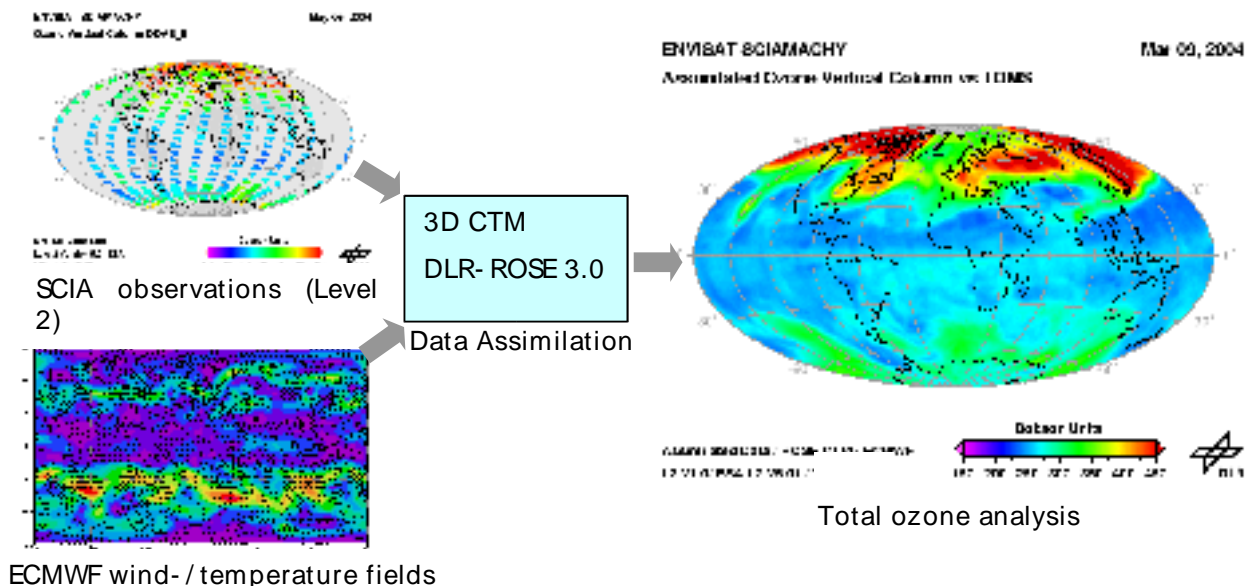


Figure 39. The SCIAMACHY near-real-time processing chain at DFD.

## 5.4. References

- Bittner M. & Erbertseder T (eds.), 2002, STREAMER, final report, Environmental and Climate Research Programme, Research area: 3.3 Earth Observation, No. ENV4-CT98-0756.
- Bittner, M, Baier, F, Brasseur, G, Elbern, H, Erbertseder, T, Hildenbrand, B, Morgenstern, O: INVERT – Final Report, BMBF, AFO 2000, 2004
- Burrows J.P., Weber M., Buchwitz M., Rozanov V., Ladstätter-Weißenmayer A., Richter A., de Beek R., Hoogen R., Bramstedt K., Eichmann K.-U., Eisinger M., and Perner D., 1998a: The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results, *J. Atm. Sci.*, 56, 151- 175.
- Daley R., 1991: Atmospheric Data Analysis, Cambridge University Press.
- DLR: GOME Level 1 to 2 Algorithm Description, ER-TN-DLR-GO-0025, Issue 2, August, 1996.
- DLR: ENVISAT-1 - SCIAMACHY Level 1c to 2 NRT and Off-line Processing Algorithm Description, ENV-ATB-SAO-SCI-2200- 0003, Issue 2, December, 2000
- Erbertseder, T: EVIVA – Value Adding of ENVISAT products for continuous monitoring of atmospheric trace gases and aerosols – ESA AO, 2003a
- Erbertseder T., F. Baier, M. Bittner, Hildenbrand, B: GLOBAL OZONE ANALYSES BY ASSIMILATING SCIAMACHY AND GOME OBSERVATIONS INTO A 3D CTM, EGS-AGU-EGU Assembly, Nice, 2003b
- Erbertseder, T., Baier, F., Bittner, M., Elbern, H., Schwinger, J.: Assimilation of ENVISAT products for continuous monitoring of atmospheric trace gases - First results from EVIVA, submitted to ESA ERS/ENVISAT Symposium 2004
- ESA: GOME Users Manual, ESA SP-1182, ESA/ESTEC, Noordwijk, The Netherlands, 1996
- Generic Mapping Tools, <http://gmt.soest.hawaii.edu>, 2004
- Kuze A. and Chance K.V., 1994: Analysis of Cloud- Top Height and Cloud Coverage from Satellites Using the O2 A and B Bands, *J. Geophys. Res.*, 99, 14481- 14491.
- Lambert J.-C., et al., 2002: ERS-2 GOME Data Products Delta Characterisation Report 2002. ESA Technical Report, 2002.
- Loyola D., Using Artificial Neural Networks for the Calculation of Air Mass Factors, 1999, European Symposium on Atmospheric Measurements from Space, Noordwijk, The Netherlands, ESA WP-161, 573- 575.
- McPeters R.D., Bhartia P.K., Krueger A.J., Herman J.R., Schlesinger B.M., Wellenmeyer C.G., Seftor C.J., Jaross G., Taylor S.L., Swissler T., Torres O., Labow G., Byerly W., and Cebula R.P., 1996: Nimbus-7 Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide, NASA Reference Publication.
- NASA, Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide, NASA Technical Publication 1998- 206895, 1998
- Reed R.J., 1950: The role of vertical motions in ozone weather relationships, *J. Meteorol.*, 7, 263- 267.
- Riese M., Tie X., Brasseur G., and Offermann D., 1999: Three-dimensional simulation of stratospheric trace gas distributions measured by CRISTA, *J. Geophys. Res.*, 104, 16419- 16435.

Rose K. and Brasseur G., 1989: A three-dimensional model of chemically active trace species in the middle atmosphere during disturbed winter conditions, *J. Geophys. Res.*, 94, 16387- 16403.

Spurr R.J.D., 1999: Improved climatologies and new air mass factor look-up tables for O3 and NO2 column retrievals from GOME and SCIAMACHY backscatter measurements, *European Symposium on Atmospheric Measurements from Space*, Noordwijk, The Netherlands, ESA WPP-161, 277- 284.

Spurr R.J.D., Kurosu T.P., and Chance K.V., 2001: A Linearized Discrete Ordinate Radiative Transfer Model for Atmospheric Remote Sensing Retrieval, *J. Quant. Spec. Radiat. Trans.*, 68, 689- 735.

Thomas, W; Baier, F; Erbertseder, T; Kästner, M: The Algeria severe weather event of November 1999 and its impact on ozone and NO2 distributions, *Tellus B*, November, Vol. 55, No. 5, pp. 993- 1006, November 2003

## 6. Acronyms

AATSR	Advanced Along Track Scanning Radiometer
AERONET	NASA's Aerosol Robotic Network
AMF	Airmass factor
AO	Announcement of Opportunity
AOT	Aerosol optical thickness
APOLLO/AVH	AVHRR Processing scheme Over cLouds Land and Ocean for AVHRR
APOLLO/SEV	AVHRR Processing scheme Over cLouds Land and Ocean for SEVIRI
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BLAOT	Boundary layer aerosol optical thickness
DFD	Deutsches Fernerkundungsdatenzentrum
DLR	Deutsches Zentrum für Luft- und Raumfahrt e.V.
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Center for Middle- Range Weather Forecasting
ERS	European Radar Satellite
ESA	European Space Agency
EUMETSAT Satellites	European Organisation for the Exploitation of Meteorological Satellites
GDP	GOME data processor
GMT	Generic mapping tools
GOME	Global Ozone Monitoring Experiment
GVC	Ghost Vertical Column

HDF	Hierarchical Data Format
HRIT	High Resolution Image Transmission
JPL	Jet Propulsion Laboratory
MLT	mean local time
MODIS	Moderate Resolution Imaging Spektroradiometer
MSG	Meteosat Second Generation
NASA	National Aeronautics and Space Administration
NAT	nitric acid trihydrate
N-GONG	Global Ozone Network for GOME
NOAA	National Oceanic and Atmospheric Administration
NRT	near real time
NVAP	NASA's water vapour project
OI	Optimal Interpolation
PIF	Processed Image File
SCIAMACHY Cartography	Scanning Imaging Absorption Spectrometer for Atmospheric
SEVIRI	Spinning Enhanced Visible and Infra Red Imager
STS	Super cool tenary solutions
SYNAER	SYNergetic AErosol Retrieval
TIGR	Thermodynamic Initial Guess Retrieval
TOMS	Total Ozone Monitoring Spectrometer
TWC	total water vapour column
UTC	Universal Time Coordinated
XPIF	Extended Processed File Format
VCD	vertical column density