

*Course on "Inverse Methods in Atmospheric Science"  
1 - 12 October 2001*

301/1332-12

---

"Aerosol and Clouds Retrieved from GOME and SCIAMACHY  
Instruments. Methodological Approach and Results"

**R. GUZZI**  
ISAO-CNR  
Bologna, Italy

---

*Please note: These are preliminary notes intended for internal distribution only.*



**Aerosol and Clouds Retrieved from  
GOME and SCIAMACHY Instruments.  
Methodological Approach and Results**

**ICTP**

**Inverse Methods in Atmospheric Sciences  
1-12 October 2001**

**Rodolfo Guzzi**

**ISAO-CNR  
Via Gobetti 101  
I-40129 Bologna  
Italy**

---

## Chapter 1

# The aerosol retrieval method

### 1.1 Introduction

Current remote sensing retrieval of the airborne aerosol are based on complex methods and tools.

At the present we do not have an univocal mathematical theory that is able both to simulate the airborne aerosol optical properties and their radiances as seen by the satellite, and to retrieve the aerosol optical thickness by the measured radiances. That means that does not exist a mathematical theory that is able to treat together the direct and inverse problems related to optical models that are used in the radiative models of the atmospheric-surface system.

We have only fragmentary methods and approaches which are based on different techniques, among the other we can select the following:

- Tabular Look up

This involves pre-calculation of measurement properties for the range of conditions that the satellite experimented. Retrieval consists of looking up, in the LUT, the corresponding values closest to that one measured and then interpolating to produce the searched result.

- Linear and Nonlinear Least-Squares Fitting

The linear LSF is based on selected measurements data fitted using a linear model (in the parameters to which we are fitting), solving an equation of the type:

$$\sum_{i=1}^N \frac{1}{\sigma_i^2} [y_i - \sum_{j=1}^M a_j X_j(x_i)] X_k(x_i) = 0; k = 1, \dots, M \quad (1.1)$$

where the  $y_i$  are the measurements,  $\sigma_i^2$  are their standard deviations, the  $X$  are the model basis functions and 'a' are the linear parameters of the model.

The Nonlinear LSF is applied to models that are nonlinear both in the parameters obtained from measurement data both in the parameters 'a' of the model. Solutions are found by minimizing the  $\chi^2$  merit function:

$$\frac{\partial \chi^2}{\partial a_j} = \frac{\partial}{\partial a_j} \sum_{i=1}^N \frac{1}{\sigma_i^2} [y_i - y(x_i)]^2 = 0; j = 1, \dots, M \quad (1.2)$$

The Levenberg-Marquardt algorithm (1963) or steepest descent algorithm or a combination of both are the usual methods of solution for a set of nonlinear equation.

- Constrained Linear Inversion

This method is used when, within the measurement error, the solution  $f_j$  is not unique and the ambiguity can be removed by imposing an additional condition which enable one

---

of the possible sets  $f_j$  to be chosen. Thus the simplest criterion for a solution is that the measurement error is minimized while the solution is constrained to be closed to its mean value  $\bar{f}$ . This procedure is widely described by Twomey (1963) and successively developed or applied by various authors, among the others we cite King et Byrne (1976), Rizzi *et al.* (1982) for sun photometer data, Russell and McCormick (1996) for SAGE III data, Gabella *et al.* (1997) to evaluate aerosol profile on GOME simulated data.

- Optimal Estimation

Optimal estimation is considered the current state of the art for retrievals in underconstrained problems encountered in atmospheric measurements. The technique was introduced by Rodgers (1997) and includes the use of a *priori* data, that can be used to constrain the final solution, in combination with measurements.

The method involves the matrix of the derivatives of the model with respect the weighting functions  $K$ . The linear problem gives the solution vector

$$x = (C_a^{-1} + K^T C_y^{-1} K)^{-1} (C_a^{-1} + K^T C_y^{-1} y) \quad (1.3)$$

The non linear case is solved by a succession of linearisations at trial solutions to give

$$x = (C_a^{-1} + K_n^T C_y^{-1} K_n)^{-1} [K_n^T C_y^{-1} (y - y_n + K_n x_n) + C_a^{-1} x_a] \quad (1.4)$$

The iterations proceed until a convergence criterion is satisfied. Subscript 'a' indicate a priori information;  $C^{-1}$  is the covariance matrix; n indicates the  $n^{th}$  iteration.

The selection of the method to be applied is essentially driven by the level of error involved in the measurements, the number of constraints to be introduced, the mathematics technology (complexity of the models to retrieve, speed of the method).

In the case of data measured by a nadir sensor the number of constraints due to concurrent presence of different atmospheric constituents and the effects due to the surface reflectance prevent us to use inverse methods like constrained linear inversion, more suitable for limb measurements, or optimal estimation more suitable for retrieving, in the IR range and in the UV, temperature profile or ozone profile respectively (De Beek, 1997).

In order to retrieve the aerosol we adopted a pseudo-inversion method where the radiative transfer in atmosphere is firstly solved for many values of the parameters to be retrieved and afterward the results are compared with the measurements until the best fit is obtained by a nonlinear LSF method. This technique seems to be as the most promising to retrieve the cloud and aerosol present on the GOME scenarios.

In this chapter we deal with these techniques.

However, because scenarios experienced by GOME are far to be clear, we present also a study on aerosol retrieval, taking into account:

1. the effect of the cloud contamination and
2. the effect of mixed surface

In both cases we have simulated the cloud fraction and surface reflectance mixed scenarios, using MODTRAN3 with the extension to DISORT (with 8 streams), evaluating, by a sensitivity analysis, the effect of their presence in terms of error on aerosol retrieval.

---

## 1.2 The pseudo-inversion

The approach adopted to derive the aerosol characteristics from GOME reflectance spectra is based on the so called *pseudo-inversion* method. In this kind of method

- the radiative transfer is firstly **directly solved** for many values of the parameters to be retrieved and
- the results are compared with measurements until the **best fit** is obtained i.e. a minimum in the “figure of merit function” is reached (the merit function measures the agreement between the data and the model).
- Then, the **parameter values** that led to this minimum are considered the retrieved parameters.

In practice the GOME measured reflectance spectra are fitted to a theoretical reflectance function depending on parameters describing the aerosol loading:

- the aerosol optical thickness and
- the aerosol class, term with which we summarize the set of chemico-physical properties of the aerosol.

All the remaining parameters driving the atmospheric reflectance, as for instance the ground reflectivity or the ground level pressure, are considered fixed and have to be input to this method. The presented method is considered suitable for **cloud free** scenarios and for oceanic **low reflecting** areas: the effect of partial cloud contamination and of non-homogeneous surfaces over the aerosol retrieval results will be discussed in sec. 1.5.

The **outputs** of the fitting procedure are

- a value of the aerosol optical thickness ( $\tau_a$ ) at the reference wavelength of 500 nm for each selected class;
- the error estimate on the retrieved optical thickness and
- an index of the goodness of the fit.

The fitting is carried out by means of a standard method (the Levenberg-Marquardt method described in the next section) that is best suited for **nonlinear models** i.e. models that depend non-linearly on the set of  $M$  unknown parameters  $\vec{a} = (a_1, \dots, a_M)$ . Indeed the method does not require a real pre-computation of values of the model function for many values of the parameters: the function is on-line evaluated in a limited set of  $\vec{a}$  determined at each iteration by the method itself.

## 1.3 The Levenberg-Marquardt fitting method (LMFM)

The Levenberg-Marquardt is a standard and very robust fitting method. It is deeply described and discussed by Press *et al.* in the book “Numerical Recipes in FORTRAN” (1994).

Given a set of data with their own measurement errors, the aim of a procedure that fits these data to a “model” or, synonymously, to a “function” deriving from some underlying physical theory, is to **adjust the parameters of the model** so to achieve a minimum in the **merit function**: in fact the merit function is conventionally arranged so that small values represent close agreement between data and theory. The parameters that led to this minimum are called **best fit parameters**: in practice they are those values of  $a_1, \dots, a_M$  that maximize the probability that the measured data points  $(x_i, y_i)$   $i = 1, \dots, N$  are a statistical occurrence of the adopted theoretical function  $\vec{y}(\vec{x}) = \vec{y}(\vec{x}; a_1, \dots, a_M)$ . This form of parameter estimation is called **maximum likelihood estimation**

---

in the sense that the probability of the data set  $(x_i, y_i)$  given the parameters is identified with the likelihood of the parameters given the data (Press *et al.*, 1994). The particular form of the adopted merit function Chi-Square

$$\chi^2(\vec{a}) = \sum_{i=1}^N \left( \frac{y_i - y(x_i; a_1, \dots, a_M)}{\sigma_i} \right)^2 \quad (1.5)$$

directly derives from the a priori stringent assumption that the data  $y_i$  have measurements errors independently random distributed as a normal (Gaussian) distribution around the “true” model  $\bar{y}(\vec{x})$ . The standard deviations of these normal distributions are designed  $\sigma_i$ . Equation 1.5 is the well know expression of the “weighted least-square fitting” or “chi-square-fitting”.

In the LMF the minimization of the merit function  $\chi^2$  proceeds **iteratively**. Given trial values for the parameters, the trial solution is improved until  $\chi^2$  stops, or effectively stops, decreasing. To reach the minimum in the merit function the LMF makes use of two different approaches, the inverse-Hessian and the steepest descent method (Press *et al.*, 1994). These two methods are combined in such a way that the latter is used far from the minimum, switching continuously to the former as the minimum is approached. The method requires the evaluation of the first derivative of the function  $y(x_i)$  needed to compute the gradient of  $\chi^2$  with respect to  $\vec{a}$ . The derivative is numerically evaluated using a routine adapted from “Numerical Recipes”: in fact the analytical computation of the derivative is prevented when using a discrete ordinate-based RTM like FAST\_DISORT. Each step of the iterative procedure consists in solving the linear equations set giving the increments  $\delta a_k$  to be applied to the parameters; we used the standard routine `mrqmin` that at each call returns the new value of  $\chi^2(\vec{a} + \delta \vec{a})$ . The final call returns also the **covariance matrix** of the standard errors in the fitted parameters. To assess if the process reached the convergence, after each iteration step the difference between the actual value of  $\chi^2$  and the previous one is evaluated. We check if the relative and absolute differences between the two  $\chi^2$  fell below the corresponding pre-selected threshold:

$$\frac{|\chi_{(n)}^2 - \chi_{(n-1)}^2|}{\chi_{(n-1)}^2} < \varepsilon_{\text{rel}} \quad \text{and} \quad |\chi_{(n)}^2 - \chi_{(n-1)}^2| < \varepsilon_{\text{abs}}, \quad n = \text{number of iteration} \quad (1.6)$$

The procedure is stopped after  $\bar{n}$  consecutive steps giving rise to  $\chi^2$  values that fulfill these conditions.

## 1.4 Application of LMF to the retrieval of the aerosol optical thickness

In the application of the method to the aerosol retrieval task

- the independent variables  $x_i \equiv \lambda_i, i = 1, \dots, N = 28$  are the selected wavelengths;
- the data  $y_i, i = 1, \dots, N$  are the GOME measurements at  $\lambda_i$   $\text{REF}_{\text{GOME}}(\lambda_i)$  with their own measurement errors  $\sigma_i$ ;
- the vector of parameters to be retrieved reduces to a scalar quantity i.e.  $\vec{a} \equiv \tau_a$  and  $M = 1$ ;
- the theoretical function  $y(x_i; a_1, \dots, a_M)$  is the atmospheric reflectance  $\text{REF}_{\text{MODEL}}(\lambda_i; \tau_a, \text{ICLASS})$  computed by means of FAST\_DISORT.

So the merit function  $\chi^2$  becomes

$$\chi^2(\tau_a) = \sum_{i=1}^N \left( \frac{\text{REF}_{\text{GOME}}(\lambda_i) - \text{REF}_{\text{MODEL}}(\lambda_i; \tau_a, \text{ICLASS})}{\sigma_i} \right)^2 \quad (1.7)$$

In the test phase real GOME measured data have been substituted with **pseudo-measured** data obtained from the function

$$y_i(\lambda_i) \equiv \text{REF}_{\text{MODEL}}(\lambda_i; \tau_a, \text{ICLASS}) \quad (1.8)$$

plus a random gaussian error of given standard deviation, simulating the instrumental noise. The use of **pseudo-measured** reflectance spectra in the phase of analysis of the retrieval method has the twofold advantage of

- producing reference spectra having **controlled input parameters** and a well known error level;
- testing the method in the **best conditions** so to asses the performances of the method itself when the “measured” data are really a random sample of the parent distribution  $y(x)$ .

Lets take in consideration all the parameters and variables upon which the computed atmospheric reflectance depends one at time and describe how they are input to FAST\_DISORT.

- The criterion of selection of the set of **wavelengths**  $\lambda_i, i = 1, \dots, N = 28$  in the GOME spectral range suitable to the aerosol OT retrieval (see tab. 1.4) has been extensively described in chapters ?? and ??. The  $\lambda_i$  are all in the transparency windows i.e. in spectral ranges where gaseous absorption features are not present. With this choice the FAST\_DISORT computations required for the fitting are greatly simplified.

Wavelengths in nm					
361.0 ÷ 427.0 each 3.0 nm	752.9	755.9	777.0	780.0	783.0

Table 1.1: List of the 28 wavelengths selected for the aerosol retrieval

The use of  $N$  wavelengths implies that FAST\_DISORT has to be called  $N$  times each time  $\chi^2$  has to be recomputed according to eq. 1.5. Moreover in the minimization process the computation of the first derivative of  $\text{REF}_{\text{MODEL}}(\lambda_i; \tau_a, \text{ICLASS})$  is also required, and that means several more evaluations of the function itself corresponding to as many calls to FAST\_DISORT. The use of a fast RTM as FAST\_DISORT for the purpose of fitting GOME spectra appears then mandatory.

- The **15 aerosol classes** selected are collected in Tab. 1.4 (see also tab. ?? for the bibliographic references). The relevant optical properties to be input to FAST\_DISORT are: the extinction ( $\sigma_e$ ) and scattering ( $\sigma_s$ ) coefficients, the single scattering phase function ( $x(\gamma)$ ), and the single scattering albedo ( $\omega_0$ , see sec. ??). The optical properties are computed by means of the software tool described in chap. ??, app. ?? and in Levoni *et al.* (1997) in the suitable format to be input to FAST\_DISORT. The 15 reported classes have been used for testing the fitting procedure: the final choice of the set of classes to be considered when processing GOME measured spectra should maybe include a smaller number of classes, for the two reasons of saving computational time and of considering only those classes that, in the test phase, will result well distinguishable from other classes and will lead to acceptable errors in the retrieved optical thickness.
- The geometrical information for each pixel  $\mu_0, \mu, \varphi$  are taken from a file ad-hoc generated starting from GOME level 2 DLR products.
- The sea water spectral albedo  $A_p$  is taken from the corresponding file described in chap. ??.

The general structure of the fitting software is depicted in fig. 1.1: in practice, for **each GOME pixel**:

- i ) the GOME reflectances for the 28 selected wavelengths are read from the file corresponding to the current ground pixel.
  - ii ) The LMFM is independently applied for each of the 15 aerosol classes.
-

- iii ) These 15 fitting procedures produce, as output, 15 values of the best fit parameter  $\tau_a$ , the associated standard deviations  $\sigma_\tau$ , and the corresponding  $\chi^2$  values, called the **residual** of the fittings.
- iv ) The smaller fitting residual among the 15 available values is selected: the corresponding parameters  $\tau_a$ ,  $\sigma_\tau$  and ICLASS are the resulting aerosol characteristic of the pixel under examination.

Index ICLASS	Name of the aerosol class
1	Clean continental
2	Average continental
3	Urban
4	Clean maritime
5	Maritime polluted
6	Desert background
7	Desert wind-carry
8	Maritime
9	Continental
10	Urban industrial
11	Rural
12	Urban (Lowtran)
13	Maritime (Lowtran)
14	Volcanic 1
15	Volcanic 2

Table 1.2: List of aerosol classes used in the retrieval method tests

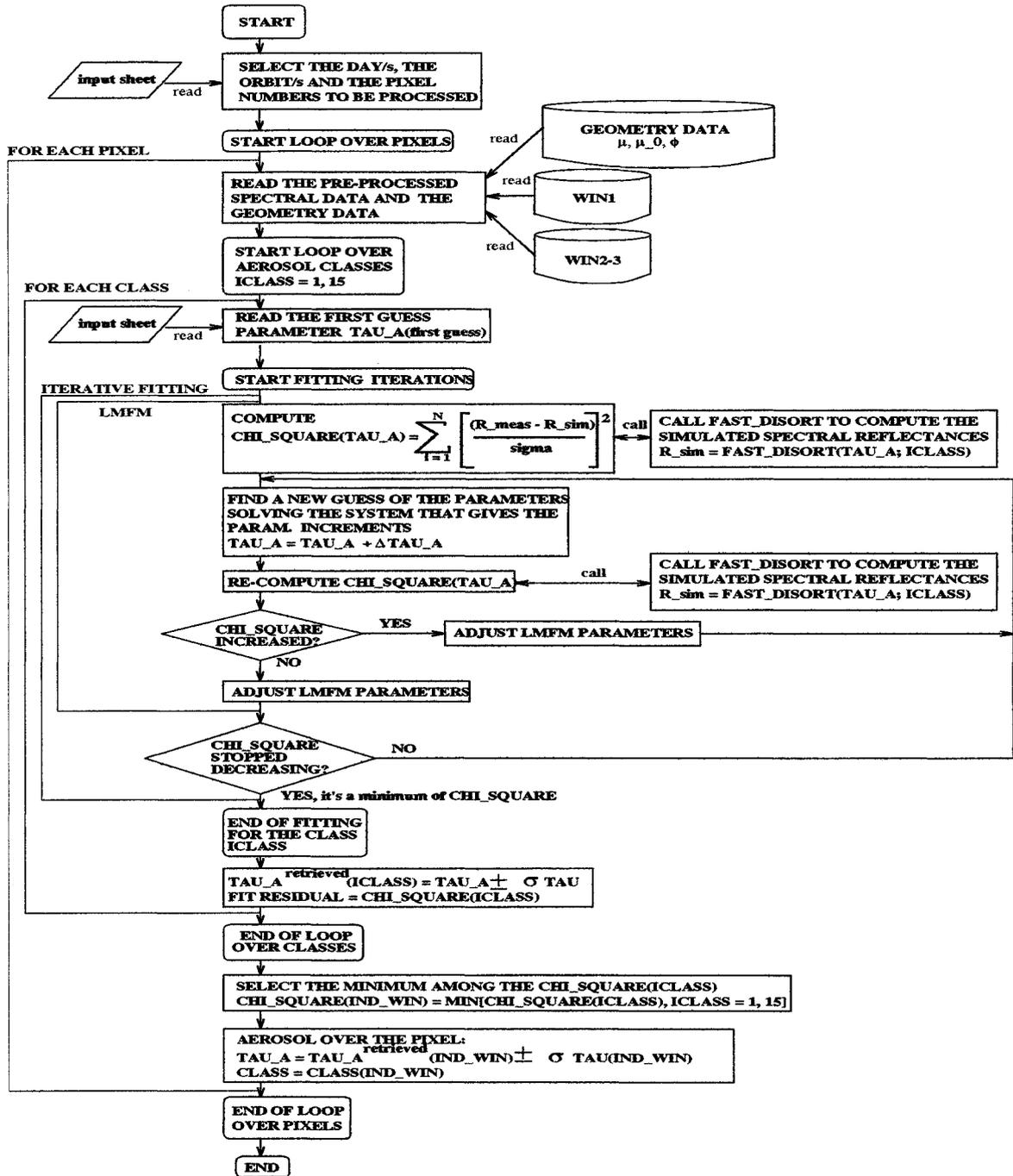


Figure 1.1: Aerosol fitting method: schematic flow diagram

## 1.5 Sensitivity analysis

It has to be stressed that several assumption have been made about the input of the forward model used to fit the measurements of the Earth surface plus the atmosphere. For instance, simulated spectra could not take into account the presence of clouds within the scene. Furthermore, the ground albedo is assumed to be equal to one of the five selected types within the dataset (see Chap.77).

Here the sensibility of the aerosol retrieval had been analysed, with respect to deviations of test scenarios as regards the previously mentioned two assumptions: the absence of clouds within the observational field; homogeneous Earth surface boundary.

Results will be discussed by focussing the attention on how much the retrieval deteriorates as the cloud fraction increases, on the one hand, and how the retrieval deteriorates as ground surface becomes less homogeneous due to the presence of small different areas on the other hand .

### 1.5.1 The influence of clouds over the aerosol retrieval

Even a low cloud fraction is reasonably able to determine the failure of aerosol retrieval, if the retrieval itself is performed under the hypothesis of a fully clear scene. For this reason a sensitivity analysis has to be carried out to evaluate the response of the retrieval as a function of the presence of cloud contamination over the scene (cloud fraction).

The following procedure needed to be built up:

- i simulation of cloud contaminated spectra have been performed by using reflectance data  $R(\lambda)$  obtained by combining cloud and clear sky reflectances,  $R_{cloud}(\lambda)$  and  $R_{clear}(\lambda)$  respectively. Reflectances of mixed scenarios with a cloud fraction  $f$  can be written as:

$$R(\lambda) = f \times R_{cloud}(\lambda) + (1 - f) \times R_{clear}(\lambda). \quad (1.9)$$

In this analysis the parameter  $f$  assumed the values shown in Tab.1.3:

Cloud fraction $f$ %	Step	no. of cases
0÷10	1	11
12÷20	2	5
25÷50	5	6

Table 1.3: Tested values of cloud fraction over the scene

The spectra  $R_{cloud}(\lambda)$  and  $R_{clear}(\lambda)$  have been produced with MODTRAN3.

- ii These pseudo-measured spectra  $R(\lambda)$  can be obtained at selected known values of aerosol content (class and AOT); the retrieval of them has been tried by means of the PGADP.<sup>1</sup>

#### MODTRAN3 pseudo-measured spectra

To compute reflectances, radiance and irradiance are required, since

$$R_{cloud/clear}(\lambda) = \frac{\pi Rad(\lambda)}{\mu_0 Irrad(\lambda)} \quad (1.10)$$

where  $\mu_0$  is the cosine of the solar zenith angle. Furthermore, the computation of the AOT at 550 nm is needed, to know the “real” scenario under examination.

Some model settings are common for the whole set of spectral radiance simulation.

- The atmospheric layered model is the USSA76.
- The spectral range is within the so called “transparent windows” (Tab.1.4). Step and resolution have been chosen as to match those of GOME.

<sup>1</sup>Prototipal GOME Aerosol Data Processor: for details, see: Mochi, M., A. Bartoloni, C. Serafini, M. Cervino, C. Levoni, E. Cattani, 1997: GOME Data Processing at I-PAF: the Aerosol Optical Thickness Retrieval from GOME Spectra. Third ERS Symposium, ESA, Florence, Italy, 17-21 March 1997.

Spectral range $\text{cm}^{-1}$	Step $\text{cm}^{-1}$	Resolution $\text{cm}^{-1}$
27955÷25000	6	11
25000÷23240	6	15
13320÷13185	3	6
12910÷12705	3	5

Table 1.4: Wavenumber selection to simulate GOME response within the “transparent windows”

- The observing line of sight is nadir.
- The day of the year, to drive the Earth-Sun distance, is set to 180. Consequently, the season that drives the atmospheric vertical profiles has to be Spring-Summer.
- The surface at the bottom of the atmosphere is water.
- The stratospheric aerosol type was always background stratospheric.

In this exercise we have to solve 27 different problems, obtained by combining aerosol types, visibility range and solar zenith angles. For each boundary layer aerosol type we have 9 problems, being set the solar zenith angle and by changing the visibility (see Tab.1.5).

Fully cloudy reflectance spectra have been obtained by selecting the presence of a Stratus/Strato Cu. cloud type, with the optical thickness of 38.7 at 550 nm, encompassed between 0.66 and 2.00 Km of altitude. Aerosol vertical transmittances at 550 nm have been also computed to characterize

AEROSOL class	rural, maritime, desert
Visibility range (Km)	05, 23, 50
Solar Zenith Angle (deg.)	20, 40, 60

Table 1.5: Aerosol and Solar geometry for computations for fully clear and cloudy scenarios.

the tested scenarios. Since MODTRAN allows to describe four different aerosol layers vertically distributed among 0, 2, 10, 30 and 100 Km. levels, transmittance for each layer and for the total atmosphere have been computed and converted to AOT.

Finally, computation of the incident solar irradiance has been carried out, within the “transparent windows”, at the same day of the year.

### Pseudo-measured spectra for partially cloudy scene: results of fitting

Simulation of the previously described scenarios leads to 22 (one for each cloud fraction value) pseudo-measured spectra for each problem that consisted in one combination of aerosol class, visibility and solar zenith angle. Three of these spectra are shown as example in fig.1.2, for different cloud fraction values. Since clouds reflect a different amount of radiation at different angles, spectra computed at different solar zenith angles are shown in fig.1.3.

To depict as the growing cloud fraction drives the reflectance, in fig.1.4 a relative (percentage) root mean square difference among spectra as a function of the cloud fraction is shown.

Results analysis is based on the comparison of retrieved AOT in presence of a partially cloudy scenario with respect to that retrieved over the clear scenario.

An expected result is that as the cloud fraction increases, the retrieved AOT increases proportionally, as shown in fig.1.5. The obvious explanation is that the presence of unknown (partial) cloudiness is mixed up with AOT. In the figure, the absolute differences between retrieved AOT have been plotted, averaged over visibility and aerosol class.

In fig.1.6 averaged relative differences have been plotted, to show how the AOT is relatively over-estimated as the cloud fraction increases.

In fig.1.7 relative differences have been plotted, being averaged over solar zenith angle and aerosol class. The dependence on the turbidity of the cloudless scenario is evident: the relative difference trend is more steep for transparent atmospheres (dashed line) than for turbid ones (solid line, see the figure caption).

It could be said that if about 3% of the GOME pixel (that corresponds to half PMD ground pixel) is fully cloudy, the retrieved AOT results to be overestimated by a factor 1.1 to 1.8 (in average) of the value retrieved over the fully clear scenario, depending on the geometry and the atmospheric turbidity.

### 1.5.2 The effect of non-homogeneous surfaces over the aerosol retrieval

Since the effect of irregular surface boundary on the reflectance of the target (i.e. coastal areas, irregular boundaries among different surface types, areas with intensive agriculture use) is far to be known, almost on the GOME ground pixel scale, it is out of the aim of this work to be able to manage seriously mixed surface types; on the contrary it was possible to make some reasonable assumption about the surface spectral reflectance over homogeneous areas like oceans, deserts, large forests. The aim of this test is to analyze the effect of a possible contamination of the “a-priori” surface type by the presence of a small unknown fraction of a different type of surface. Many different combination of such situations may occur; we have selected four occurrences:

- i Ocean contaminated by Sand (unknown sand-banks, wrong coastlines).
- ii Ocean contaminated by Soil (unknown land, wrong coastlines).
- iii Sand contaminated by Vegetation (unknown desert morphology).
- iv Vegetation contaminated by Soil (unknown forest irregularities).

The following procedure needed to be built up:

- i Computation of contaminated surface albedos  $A(\lambda)$  as a function of a contamination linear parameter  $c$ , obtained by combining two spectral surface albedo types:

$$A(\lambda) = (1 - c) \times A_{main}(\lambda) + c \times A_{secondary}(\lambda). \quad (1.11)$$

$A_{main}$  being the surface albedo of the “a-priori” known surface type and  $A_{secondary}$  that of the unknown portion of surface type within the scene. In this analysis the  $c$  parameter assumed the values shown in tab.1.6

contamination $c$	Step	no. of cases
0÷0.1	0.01	11

Table 1.6: Contamination factor values used in this analysis.

Surface plus atmosphere spectra  $R(\lambda)$  have been produced with MODTRAN3 by inputting the overmentioned ad-hoc mixed surface albedo data.

- ii These pseudo-measured spectra  $R(\lambda)$  can be obtained at selected known values of aerosol content (class and AOT); the retrieval of them has been tried by means of the PGADP.

#### MODTRAN3 pseudo-measured spectra

Again, to compute reflectances, radiance and irradiance are required, as shown in eq.1.9. Irradiance and the computation of the AOT at 550 nm can be derived from the test on the cloud contamination exercise.

Several other model settings reflect the same choices made within the previous subsection, i.e the atmospheric layered model USSA76, the spectral range, the nadir observing line of sight, the day of the year and consequently the season of the year.

A more exhaustive presentation has to be made to explain the input for ground reflectances. MODTRAN3 allows users to define their own dataset for surface spectral albedos. The user-defined data have to be merged into the REFBKG. file which collect data from the official distribution of the code. IMGA has collected five spectral surface albedo for five surface types, 84 values in

the range 290-800 nm. They have been included within a new version of the input data file called REFBKG\_IMGA, each of them labelled with an index ranging from 21 to 25. Data are addressed by means of the SALB input flag, with negative values from -21 to -25.

These five dataset can be used for homogeneous surface boundaries. In carrying out the present sensitivity analysis, contaminated albedos have been computed as described in the previous section. Results have been included within the REFBKG\_IMGA. file as well. Indexes of dataset have been arranged as in tab.1.7.

Mixed surface types		Contamination par. ( $10^{-2}$ )
Main	Secondary	1,2,3,4,5,6,7,8,9,10
water	sand	31,32,33,34,35,36,37,38,39,40
water	soil	41,42,43,44,45,46,47,48,49,50
sand	veget.	51,52,53,54,55,56,57,58,59,60
veget.	soil	61,62,63,64,65,66,67,68,69,70

Table 1.7: Tabulation of reflectance data indexes used in MODTRAN3 to simulate contaminated surface types

Five tests on the selected surfaces have been carried out. Each of these five experiments is composed of one aerosol class at boundary layer, and of one (contaminated) surface type (see tab.1.8). Each

CASE	MAIN surface	SECOND. surface	AEROSOL class
1	water	sand	maritime
2	water	sand	desert
3	water	soil	maritime
4	sand	veget.	desert
5	veget	soil	rural

Table 1.8: Description of tested cases

test (one mixed surface, one boundary layer aerosol class) leads to 99 different problems: the first nine represent the input to compute radiances with reference to the main surface type without contamination, by changing solar zenith angles (20, 40, 60 deg.) and the visibility (5, 23, 50 Km) respectively. Whilst the others 90 input the surface contamination is modeled, only by changing the SALB flag as shown in tab.1.7

#### Pseudo-measured spectra for non-homogeneous surfaces: results of fitting

Simulation of the previously described scenarios leads to 11 (one for each cloud fraction value) pseudo-measured spectra for each problem that consisted in one combination of visibility and solar zenith angle. Two cases have been shown in fig.1.8 and fig.1.9. The first shows as the reflectance changes as the sand is corrupted by the presence of vegetation, whilst the second one shows the results of a partial presence of soil within a vegetation pixel.

To depict as the growing surface contamination drives the reflectance, in figs.1.10 a relative (percentage) root mean square difference among spectra as a function of the surface contamination is shown. It can be seen that the response to the contamination is low for bright surfaces as vegetation and sand (dotted and solid lines), while if the homogeneous surface is relatively dark, as in the water case (dashed line), even a small contamination results in a reflectance change. The "stepwise" trend of the dashed line is due to the use of a few significant digit in inputting the reflectance data in MODTRAN3.

Results analysis is based on the comparison of retrieved AOT in presence of a contaminated surface scenario with respect to that retrieved over the homogeneously bounded scenario.

As in the previous subsection, an expected result is that as the surface contamination increases, the retrieved AOT changes proportionally. By contrast with the previous test, the sign of the difference is now dependent on the difference between the main and the secondary surface albedo. It can be seen that the averaged absolute and relative differences (figs.1.11 and 1.12 increases with

the percentage of contamination in case of water as main surface, since the contamination by a brighter surface is mixing up with AOT. For instance, at 4% contamination level, AOT is overestimated by a factor 1.3 (4% contamination level corresponds to an area of 512 Km<sup>2</sup>).

Different results have been obtained if the main surface is brighter, as sand or vegetation. In these cases, both sign and trend of differences seems to depend on surface types and aerosol class. A better insight should be reached to evaluate the error. As a preliminary result, the case of vegetated target contaminated by soil seems not to be sensible to these small contamination.

---

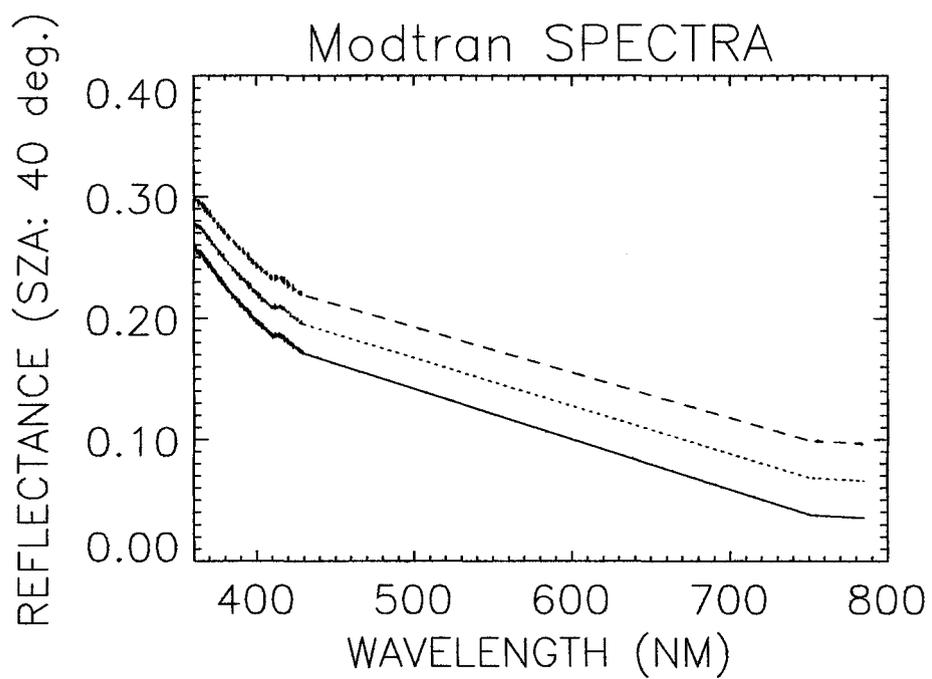


Figure 1.2: MODTRAN3 simulated spectra: aerosol class: Maritime; visibility: 23 Km; SZA: 40 deg. solid:  $f=0$ ; dotted:  $f=0.05$ ; dashed:  $f=0.1$ .

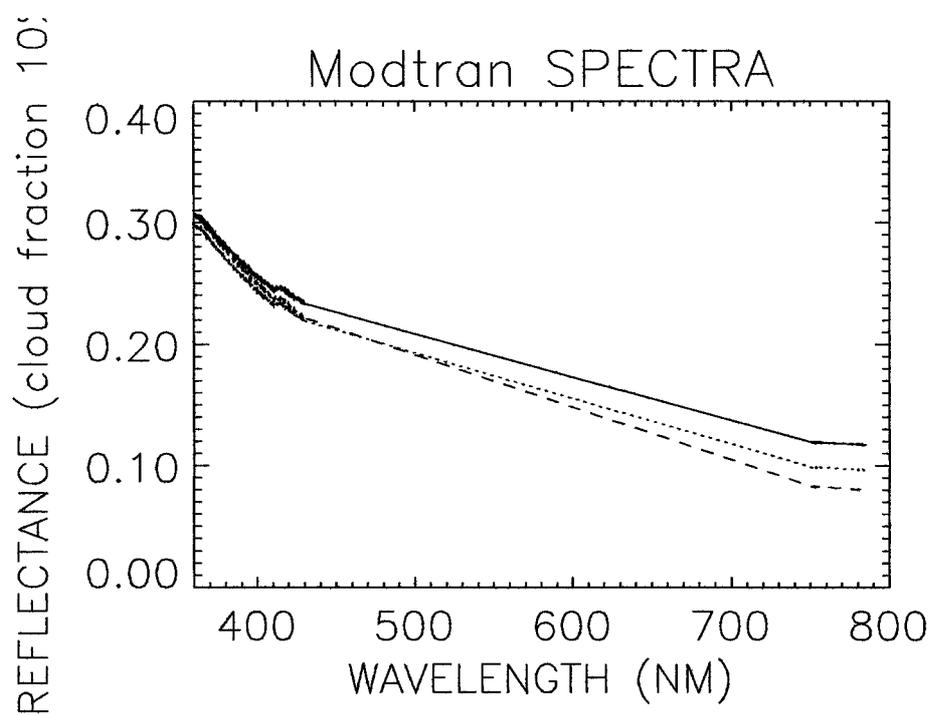


Figure 1.3: MODTRAN3 simulated spectra: aerosol class: Maritime; visibility: 23 Km; cloud fraction: 0.1. SZA: solid:20°; dotted:40°; dashed:60°.

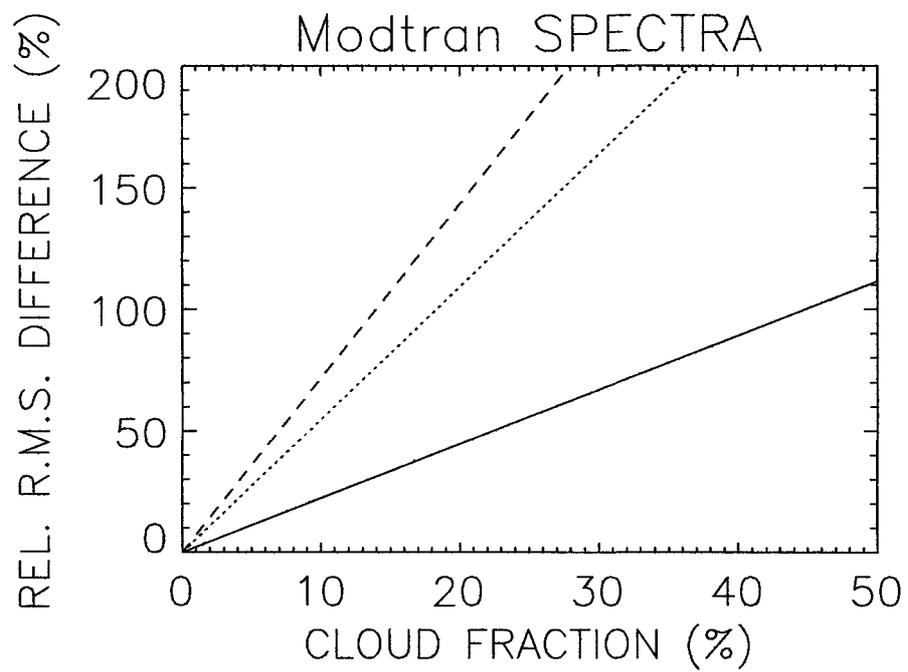


Figure 1.4: Root mean square differences among MODTRAN3 spectra as function of cloud fraction. The reference spectrum is for clear scenario. Aerosol class: Maritime. Visibility: solid:50km, dotted:23km, dashed:5km; SZA: 40 deg.

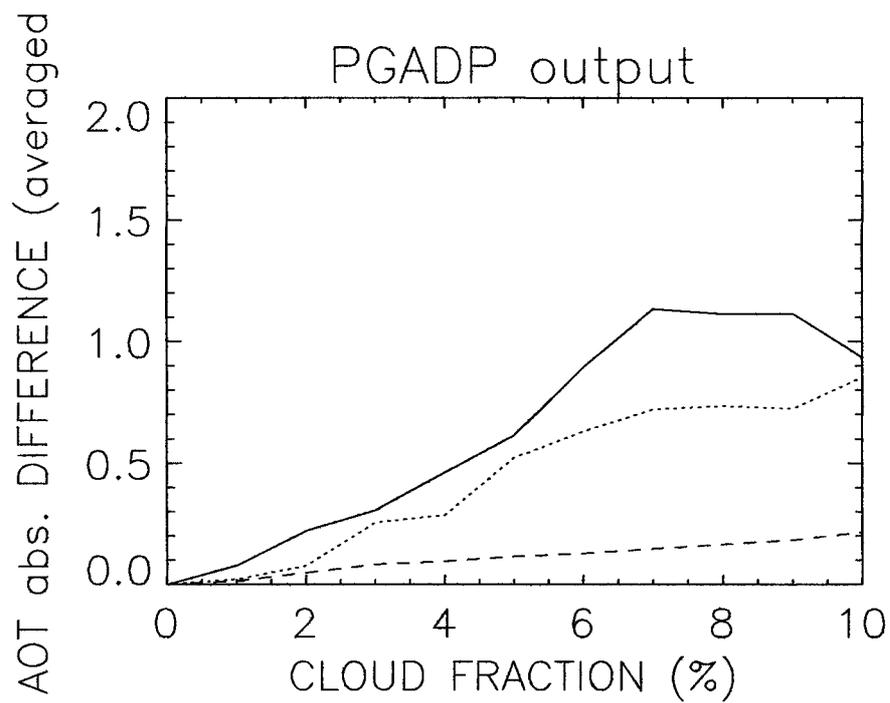


Figure 1.5: Averaged absolute differences between retrieved AOT as function of cloud fraction. The reference AOT is for clear scenario. Averages have been made over aerosol class and visibility. SZA(deg.): solid:20, dotted:40, dashed:60

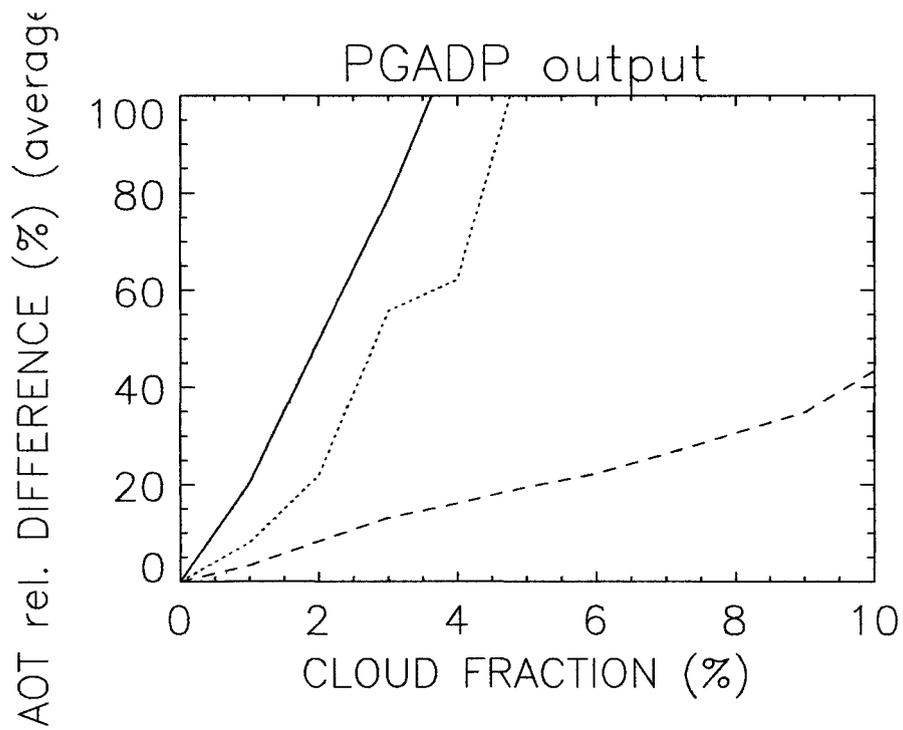


Figure 1.6: Averaged relative differences between retrieved AOT as function of cloud fraction. The reference AOT is for clear scenario. Averages have been made over aerosol class and visibility. SZA(deg.): solid:20, dotted:40, dashed:60

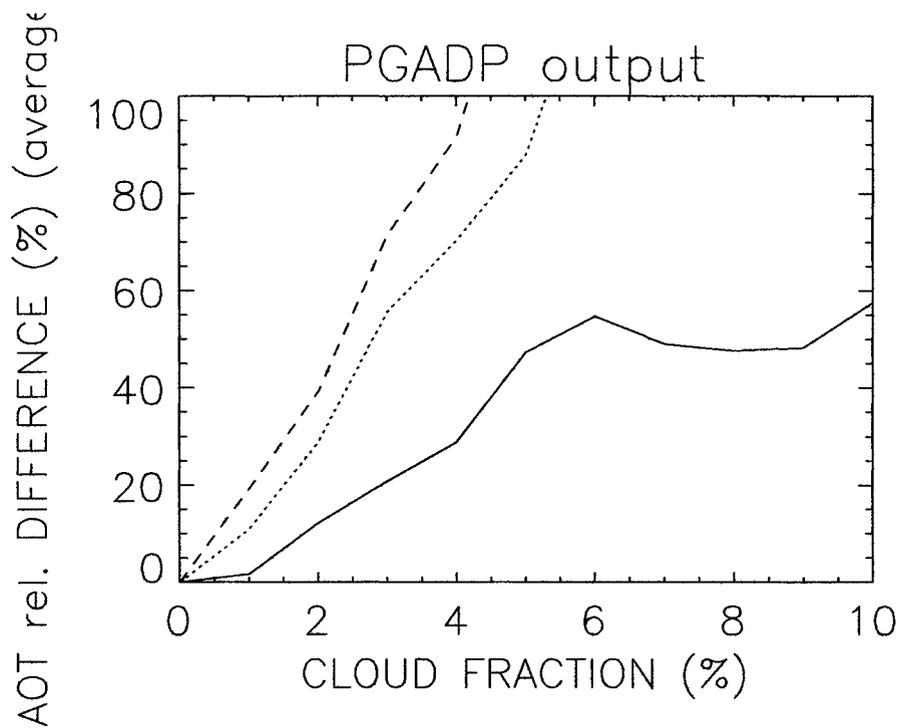


Figure 1.7: Averaged relative differences between retrieved AOT as function of cloud fraction. The reference AOT is for clear scenario. Averages have been made over aerosol class and Solar Zenith Angle. Visibility(Km): solid:5, dotted:23, dashed:50

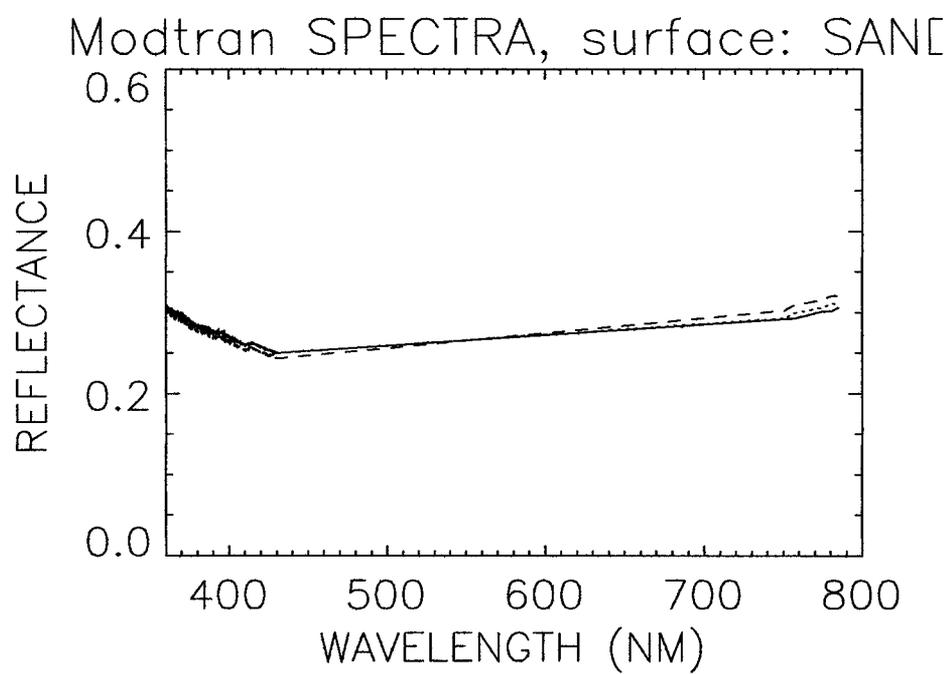


Figure 1.8: MODTRAN3 simulated spectra: aerosol class: Desert; visibility: 23 Km; SZA: 40 deg. Main Surface type: Sand. Contaminating surface: Vegetation: solid:  $c=0$ ; dotted:  $c=0.05$ ; dashed:  $c=0.1$ .

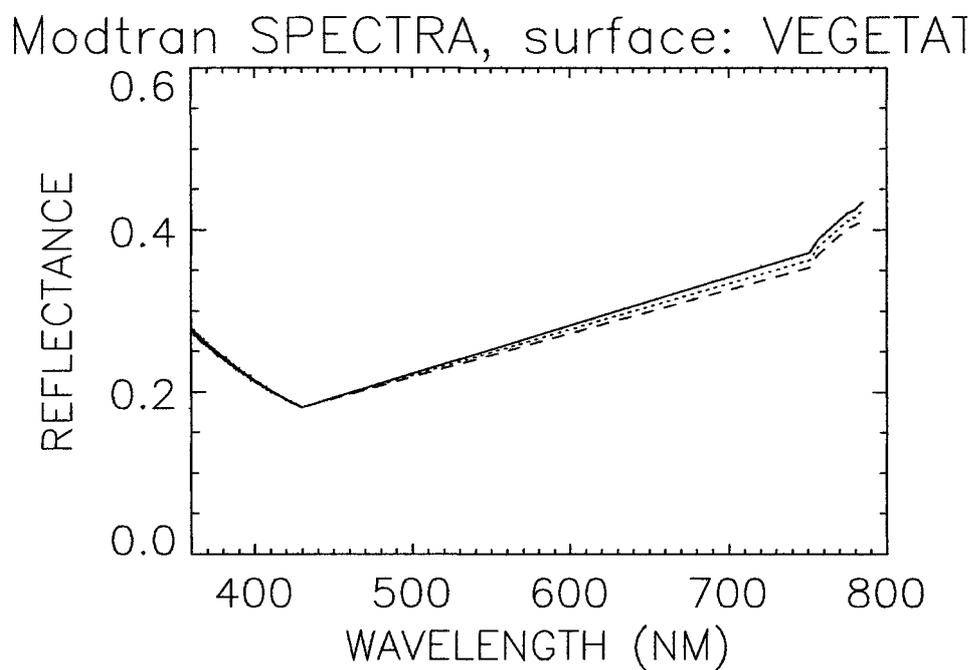


Figure 1.9: MODTRAN3 simulated spectra: aerosol class: Rural; visibility: 23 Km; SZA: 40 deg. Main Surface type: Vegetation. Contaminating surface: Soil: solid:  $c=0$ ; dotted:  $c=0.05$ ; dashed:  $c=0.1$ .

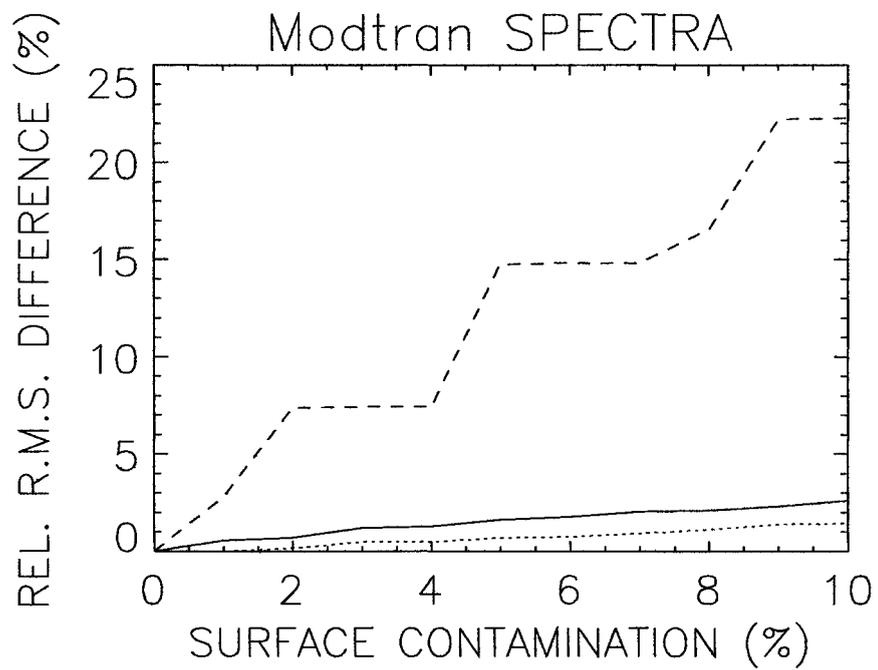


Figure 1.10: Root mean square differences among MODTRAN3 spectra as function of surface contamination. The reference spectrum is for homogeneous surface scenario. SZA: 40 deg. Visibility: 23 Km. Main surface (secondary) [aerosol]: solid:sand (veget.) [desert]; dotted:veget. (soil) [rural]; dashed:water (sand) [maritime].

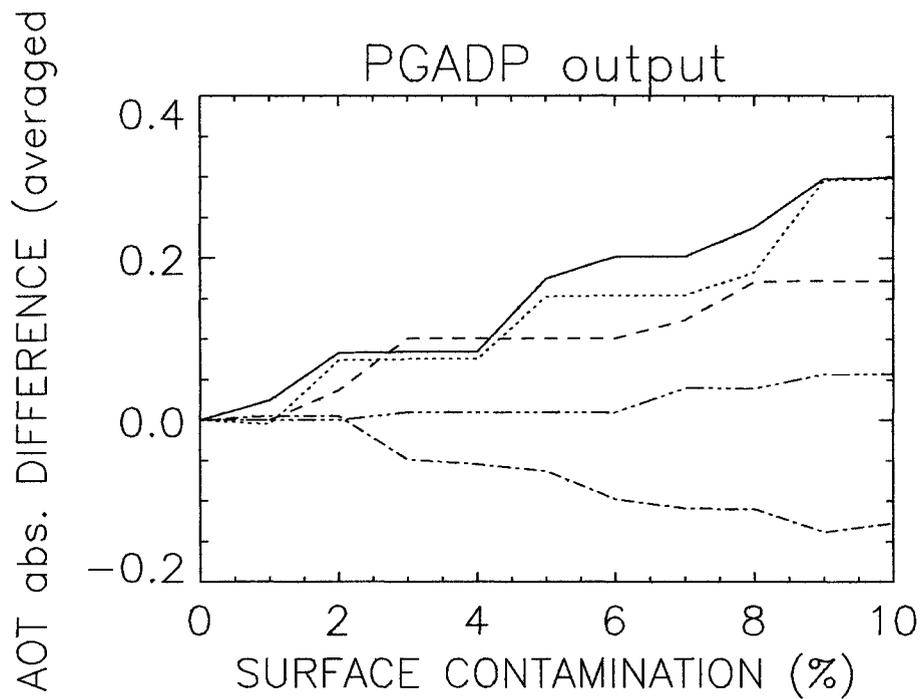


Figure 1.11: Averaged absolute differences between retrieved AOT as a function of surface contamination. The reference AOT is for homogeneous surface scenario. Averages have been made over SZA and visibility. Main surface (secondary) [aerosol]: solid:water (sand) [maritime]; dotted:water (sand) [desert]; dashed:water (soil) [maritime]; dash-dotted:sand (veget.) [desert]; dash-3-dotted:veget. (soil) [rural]

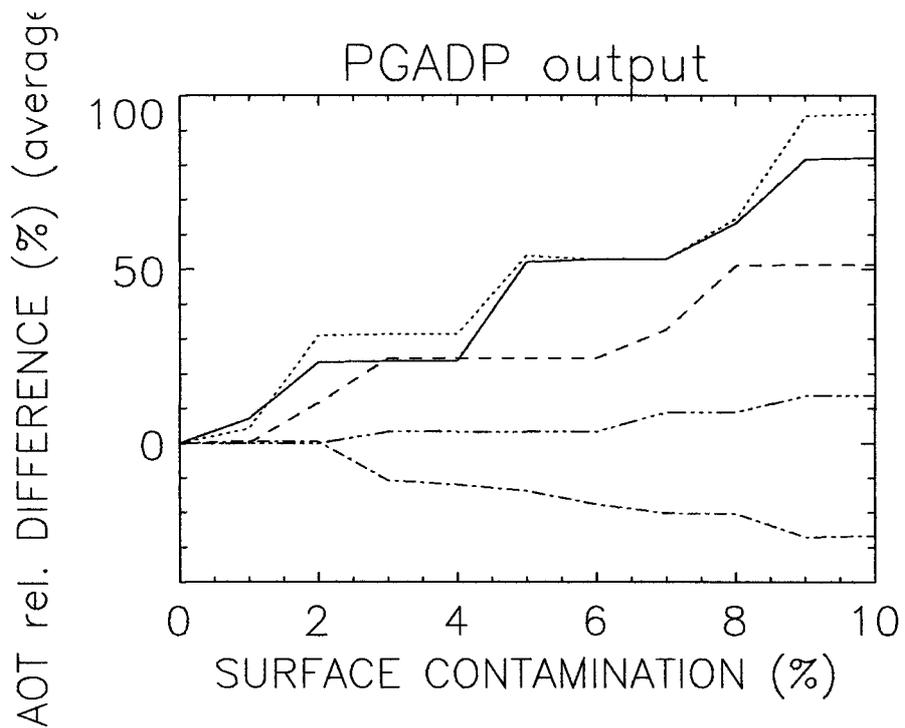


Figure 1.12: Averaged relative differences between retrieved AOT as a function of surface contamination. The reference AOT is for homogeneous surface scenario. Averages have been made over SZA and visibility. Main surface (secondary) [aerosol]: solid:water (sand) [maritime]; dotted:water (sand) [desert]; dashed:water (soil) [maritime]; dash-dotted:sand (veget.) [desert]; dash-3-dotted:veget. (soil) [rural]

## 1.6 Conclusions

Results of the aerosol retrieval with pseudo measured spectra indicate that:

- the presence of cloud fraction over the measured scene increases the retrieved AOT because the cloud fraction is mixed with the aerosol and it is indistinguishable from the aerosol. The overestimation is strongly dependent on the SZA and on the atmospheric turbidity. Higher and higher is the atmospheric turbidity lower and lower is the effect of the cloud contamination on the AOT retrieved as it is well shown in fig 11.6. On the contrary the rate of increasing of the AOT retrieved is higher for low SZA indicating that the cloud reflectance obtained by MODTRAN3 simulation is higher at low SZA than at high SZA. A visual inspection to MODTRAN3 cloud reflectance data shows that the cloud reflectance decreases with higher SZA (this behaviour is in agreement with the simulations carried out in Guzzi *et al.*, 1995) confirming the AOT results in fig 11.5.
- In the case of non-homogeneous surfaces, the result of the fitting shows that the AOT departure from the 'true value' is strongly dependent on the surface reflectance contrast. In fact when the contrast between the main and secondary surface is low (see fig 11.9), the AOT relative difference are very low, whereas increases when the contrast is higher (like in the water-sand, water-sand and water-soil mixed surfaces) (see fig 10.10 and 11.11)

Summarizing we can say that the presence of cloud fraction over the GOME scene produces an increase in AOT that can be relevant already when the GOME pixel is 3 % fully cloudy (corresponding to half PMD ground pixel) and the AOT retrieved is overestimated by a factor 1.1 to 1.8 (in average) depending on the SZA and turbidity.

In the case of non-homogeneous surface, the AOT retrieved strongly depends on the surface reflectance contrast. In case where the main surface is dark as like the water, the presence of a brighter surface that produce a contamination a level of 4 % (corresponding to an area of 512  $km^2$ ) produces an AOT overestimation by a factor 1.3 while if the main surface is bright as in the case of sand or vegetation the departure from the AOT 'true value' depends on surface type and aerosol class (as previously defined). Vegetated surface, contaminated by soil does not produce any aerosol variation.

In conclusion the accuracy of the AOT retrieved depends on:

- the cloud clearing accuracy.
- the information on the surface properties
- the accuracy of the radiative model and minimization techniques used in the pseudo-inversion

As far as regards GOME, the first point can be very accurate using the cloud retrieval approach outlined in previous work (Guzzi *et al.*, 1995) and improved and tested in the present study in combination with PMD cloud detection algorithm, to decouple the fraction of cloud from the cloud optical thickness (see section 4.1). The high accuracy presently reached, obtained making sensitivity experiments, however, should be validated over a large set of data measurements obtained by already validated airborne instrument or from ground based station. Only a large statistics can guarantee that the approach used and the model used are overall valid (we remind that the problem of 'missing physics' in the cloud sun absorption properties is an open problem yet).

The second point is strongly dependent on the GOME ground pixel from one side, and from ability to obtain the simultaneous retrieval of aerosol and surface reflectance. For non-dark homogeneous surface this task is an open question, despite several theoretical works (among the other we cite Kondratiev and Smokti, 1973, King and Herman, 1978, Pinty et al, 1990). From the simulation here presented it is evident that an accurate AOT retrieval can be obtained over large homogeneous dark surfaces like ocean in clear sky condition. AOT retrieval over non-dark, but homogeneous surfaces, like desert or vegetated areas can be recovered provided that 'a priori' information can

be supplied. Non-dark non-homogeneous surfaces, like water-soil and water-sand should be to flag out because critical.

The third point is more general and regards the pseudo-inversion approach we have, here, outlined. The level of confidence of the method here presented depends on the radiative model and the minimization techniques accuracy. Tests we carried out, presented in chapter 4 and appendix ??, indicate our approach is highly accurate and flexible because is able to retrieve, with the proper radiative model, not only the AOT but also the cloud optical thickness, fraction and altitude.

---

## 1.7 Bibliography

- De Beek R., R. Hoogen, V. Rozanov and J.P. Burrows: 1997 "Ozone Profile Retrieval from GOME satellite data. I: Algorithm Description" Proc Third ERS Symposium, Florence, 17-21 March 1997 ESA SP-414,
- Gabella M. , R. Guzzi, V. Kissilev, G. Perona: 1997, "Retrieval of the Aerosol Profile Variations in the Visible and Near Infrared: 1 Theory and Application of the single scattering approach", Applied Optics, 1328-1336
- Guzzi R., J. Burrows, M. Cervino, K. Chance, T. Kurosu, V. Rozanov, F. Torricella, P. Watts: 1996, "A study of cloud detection", ESA contract 10997/94/NL/CN
- King M. D. and D. M. Byrne: 1976, "A method for inferring total ozone content from spectral variations of optical depth obtained with a solar radiometer", J. Atmos. Sci. **35**, 2242-2251
- King M. D. and B. M. Herman: 1979 "Determination of the ground albedo and the index of absorption of atmospheric particulates by remote sensing. Part I: theory" J. Atmos Sci. **36**, 163-173
- Kondratiev K. Ya and O. I. Smokti: 1973, "Influence of aerosols on the spectral albedo of the atmosphere-underlying surface system", Izv. Atmos and Oceanic Physics, **9**, no 12, 1269-1282
- Levoni, C., M. Cervino, R. Guzzi, F. Torricella: 1997, Atmospheric aerosol optical properties: a Data Base of radiative characteristics for different components and classes, Applied Optics, **36**, no. 30, pp. 8031-8041
- Pinty B. , M. M. Verstraete, R. E. Dickinson: 1990, "A Physical model of the bidirectional reflectance of vegetation canopies. Inversion and validation ", J.G.R., **95**, no d8, 11767-11775
- Press, W.H., S.A. Teukolski and W.T. Vetterling, B.P. Flannery: 1994, "Numerical Recipes in FORTRAN. The Art of Scientific Computing Second Edition", Cambridge University Press, New York, USA, 963 pp.
- Rizzi R. , R. Guzzi, R. Legnani: 1982, "Aerosol Size spectra from Extinction Data: the Use of a Linear Inversion Method", Applied Optics, **21**, no. 9, pp. 1578-1587
- C. Rodgers: 1997, "Inverse methods for atmospheric sounding: Theory and practice", World Scientific, London
- Russel P. B., M. P. McCormick: 1996 "SAGE III algorithm theoretical basis document (ATBD) aerosol data products", NASA V1 Nov 1996, <http://eosdis.larc.nasa.gov>
- Taylor V. R. and L.L. Stowe: 1984, "Atlas of reflectance patterns for uniform Earth and Cloud surfaces (Nimbus-7 ERB-61 Days)", Us Dept of Commerce NOAA Tech Rep NESDIS 10
- Twomey S. : 1977, "Introduction to the mathematics of inversion in remote sensing and indirect measurements", Elsevier Sci. Pu Amsterdam 243 pp
-

## Chapter 2

# Application of the method to the GOME measured data: preliminary results

### 2.1 Introduction

The present chapter is based on the results of the retrieval method developed and explained in previous chapters. The elements we will analyze will be strictly connected with the GOME data and our aerosol retrieval results. Due to uneven spatial properties of the data and their complexity a visual inspection of GOME data cannot give information on aerosol event present over the measured scenarios. The presence of cloud, of any thickness, over the scenarios and unknown surface reflectance can produce misleading results. So we decided to use METEOSAT data to select those measurements which evidenced aerosol events spatially relevant and with optical thickness apparently large. Such modality did not question the validity of our results but served to drive our work. We found the METEOSAT data in the Project MEDUSE (The MEditerranean DUST Experiment) at:

(<http://halo.hi.is/meduse>)

MEDUSE is a two years project initiated on 1 March 1996, supported by the European Commission (DG XII), with partners-subcontractors in France, Germany, Greece, Iceland and Italy.

The overall objective of Project MEDUSE was to develop and implement a prototype system for routine monitoring and prediction of the atmospheric transport of desert dust in the Mediterranean region, based on a numerical weather prediction model enhanced by a module for the simulation of the dust uptake, transport, and wet and dry deposition. The dust simulations were being validated and further enhanced by means of an extensive measurement programme, including aerosol remote sensing by satellite, lidar and sunphotometer, chemical characterisation of aerosol particles, and deposition flux measurements.

The dust monitoring and prediction system is used both for the simulation of past events and for experimental forecasts of the atmospheric transport and deposition of desert dust in the Mediterranean region. These daily regional forecasts are normally produced for 60 hours ahead. In addition to forming the basis for desert dust alerts, the forecasts are being made available to interested agencies and scientists.

In the frame of this project we located two large aerosol events occurred at June 97 along the west and south west African coasts close to Sahara.

During the same period ground sunphotometry measurements were also carried out in the frame of AERONET (AErosol RObotic NETwork) project found at:

---

(<http://spamer.gsfc.nasa.gov/>)

AERONET is an optical ground based aerosol monitoring network and data archive supported by NASA's Earth Observing System and expanded by federation with many non-NASA institutions. The network hardware consists of identical automatic sun-sky scanning spectral radiometers owned by national agencies and universities. Data from this collaboration provides globally distributed near real time observations of aerosol spectral optical depths, aerosol size distributions, and precipitable water in different aerosol regimes. The data undergo preliminary processing (real time data), reprocessing (final calibration around 6 months after data collection), quality assurance, archiving and distribution from NASA's Goddard Space Flight Center master archive and several identical data bases maintained globally. The data provide algorithm validation of satellite aerosol retrievals and as well as characterization of aerosol properties that are unavailable from satellite sensors.

## 2.2 Meteosat and GOME data comparison

METEOSAT data are not wholly comparable with GOME data because METEOSAT is a geostationary satellite and pixels at equatorial locations are of the order of 5x5 Km against the GOME pixels that are of 320x40 Km. Data obtained from METEOSAT, however, permit a visual inspection more pregnant than that one provided by GOME and can guide the analysis we wanted to carry out. On the contrary GOME, having a medium spectral resolution in the Uv-Red wavelength range is able to extract spectral information useful to measure the gases content and the aerosol optical thickness present in the atmosphere with higher accuracy than METEOSAT. In such context we can say that METEOSAT and GOME can be considered synergics.

On the basis of such considerations we have analyzed the GOME data looking for the aerosol presence during three days of June 6,7,8, 1997 to verify if GOME was able to capture the aerosol events seen by METEOSAT (see Fig.2.1,2.2,2.3). Aerosol data have been coloured in green or red as a function of the lower or higher optical thickness respectively.

The application of our retrieval method guaranteed high accuracy and fast results. We focussed our attention on two events over the sea, one towards the west of Sahara, that we considered the main event, the other located in the south-west of Sahara that we considered as a secondary event.

Results obtained processing GOME measurements carried out over the aerosol events show that, not only, higher optical thicknesses due to the presence of desert wind carry aerosol over the sea were detected by GOME but also the smallest optical thicknesses, as for instance those close to the south African coasts. Data, however, were still affected by the presence of clouds and surface reflectance. This could produce large errors as has been evidenced in previous chapters.

## 2.3 Selection of data to be retrieved and results

One of the main requirements, mentioned in previous chapters is that the channels signal follow a well defined rule with no jumps. An inspection over data indicates that several jumps occurred, not allowing the employment of all data. Results, however, show that the main event can be still evidenced, while the secondary one is almost totally flagged. These results are shown in (see Fig.2.4,2.5,2.6). They indicate that the required information, on the main and secondary aerosol events we have selected are partially obtained. Then we decided to use all the measurements without any preselection on the data. Results presented in (see Fig.2.7,2.8,2.9) show that the aerosol events selected are well evidenced and are located in the same areas shown by METEOSAT. These results suggest also jumps can be retained as a physical effect of the target over the sensors (adjacency effect, effect of the roughness and inhomogeneity of the surface etc) that needs to be studied furtherly. While the range optical thicknesses between 0 - 3 was coloured with different shading of yellow and red, the AOT greater than 5 was coloured with green. These last set of data can be retained as due to different effects:

- cloud presence,

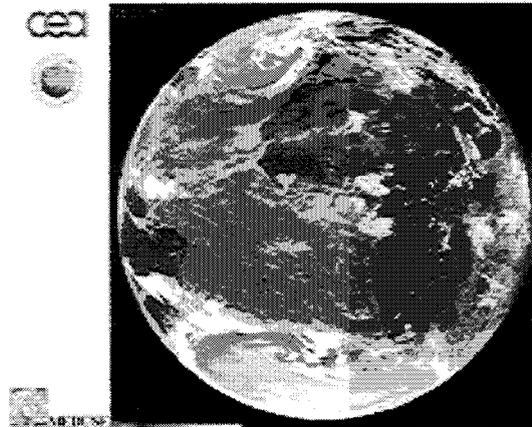


Figure 2.1: Meteosat picture obtained by MEDUSE project. Red and green colors show, respectively, the higher and the lower optical thickness of the desert aerosol. June 6,1997

- higher ground reflectance we expected ( probably due to sea roughness, substances suspended upon the sea),
- ground surface reflectance not corrected ( like in the case of desert reflectance)
- sunglint

However, in order to evaluate and define the level of the errors introduced in retrieval, due to also those effects, we had the opportunity to compare the AOT data obtained by GOME with ground data measured at Capo Verde.

## 2.4 Ground measurements

We used as ground validation the Aerosol Optical Thickness (AOT) data measured by sunphotometers located at Capo Verde. Atmospheric studies by sunphotometers are well known and were used to validate several satellite data. The initial remote sensing aerosol monitoring network initiated and funded by NASA's EOS under the name AERONET was expanded by national and international collaboration by using recent weather resistant automatic and sky scanning spectral radiometers to detect aerosol optical properties and precipitable water at remote sites.

Data are available by the net covering a large portion of the globe even though the operative frequency of collection is different from a site to another. Measurements are essentially based on three different measurements methodologies, the first based on sunphotometry introduced by Volz (1959) and successively implemented by several authors. Modern units can collect quickly and accurately data by a filtered detector measuring the spectral solar direct extinction according the Bouger-Lambert-Berr's law. The second is based on multifilter rotation shadowband radiometer measuring the spectral total and diffuse radiation to obtain the direct component from which is calculated the aerosol extinction ( Guzzi et al. 1985, Harrison et al. 1994). The third is a sky

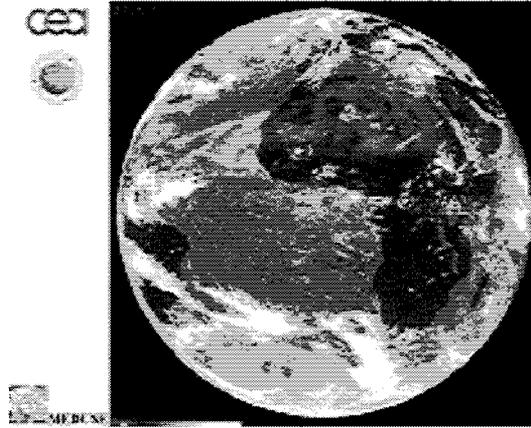


Figure 2.2: Meteosat picture obtained by MEDUSE project. Red and green colors show, respectively, the higher and the lower optical thickness of the desert aerosol. June 7, 1997

scanning spectral radiometers measuring the sky radiance at known angular distances from the sun. By such a method the knowledge of the aerosol properties such as size distribution and phase function can be efficiently derived (Nakajima et al 1983, 1996), Tanré et al (1988), Shiobara et al. (1991) and Kaufmann et al (1994).

Despite all those methodologies, the central problem remains the agreement on the accuracy by which the aerosol optical thickness is derived. Although Bouger-Lambert-Beer's law is very straightforward, however, the uncertainties in computation of the airmass ( $m$ ), the calculations for the Rayleigh and ozone optical depths and water vapor expressed as total column abundance or precipitable water ( $P_w$ ) as well as strategies for calibration of the instruments and monitoring the long term change in calibration all combine to preclude any globally accepted processing scheme.

In order to obtain comparable results the solutions that AERONET adopted make the raw data and calibration data available to the user and provide a basic processing package (of published, widely accepted algorithms).

We used the data as such as are given from AERONET and shown in Fig 2.10. Data show the AOT trend in three different wavelengths for the period in which aerosol event occurred. During the three selected days the AOT of all the three wavelengths presented in the figure ranges between 1.8 - 2.2. These findings together with the decreasing of the value around the 10th of June and the subsequent AOT increasing up to the final decreasing around the end of the month show that a big aerosol event was in act. Such event has been also recorded by GOME whereas our retrieval model evidences a relevant increasing of optical thickness respect to the surrounding areas. Since the GOME retrieved data, obtained at the 500 nm reference wavelength, were comparable with those given by AERONET we can consider our result very accurate, beyond the quality of raw data.

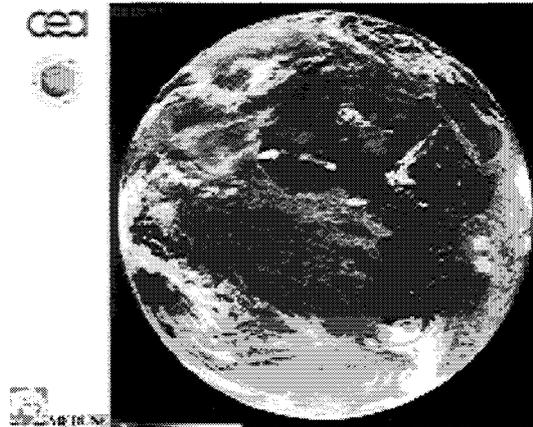


Figure 2.3: Meteosat picture obtained by MEDUSE project. Red and green colors show, respectively, the higher and the lower optical thickness of the desert aerosol. June 8,1997

## 2.5 Cloud and aerosol

The METEOSAT pictures, however, evidenced the presence of clouds in sites not too far from events where we retrieved the aerosol. While the main event is quite free of clouds, the secondary event shows mixed scenarios containing aerosol and clouds whose optical thickness could be also comparable with the aerosol one. In such a case we are not able to distinguish between aerosol and clouds unless we are undergoing GOME data to the cloud retrieval algorithm we developed and that will be presented in the next chapters. Such approach evidences ( see Fig 2.11 and Fig 2.12) the presence of clouds in the same areas shown where METEOSAT show clouds, while evidences, over the aerosol main event, no cloud.

## 2.6 Defining the Aerosol class

On the basis of the AOT obtained, we have also analyzed, for the selected days, the class of the aerosol by a ranking algorithm. The first results (june 6) are reported in Fig 2.13. They show that, in the area where METEOSAT pictures evidences the desert aerosol, our first ranking shows that volcanic or continental aerosol are predominant. Such behaviour contradicts findings expected but can be explained by analyzing reflectances of the four different aerosol classes employed, as shown in figure 2.14.

Comparing reflectances of the four different aerosol classes used, we can see that the wind carry desert aerosol class reflectance we used is far from GOME reflectance data, while the volcanic aerosol class matches them. Such behaviour, primary suggests that:

- the aerosol class found in the literature, mainly that proposed by D'Almeida et al., 1991, is not able to match the aerosol effectively seen by GOME;
- the absorption inherent the desert aerosol optical properties prevents the reflectance model to match GOME data unlike a non absorbing aerosol, like the volcanic one.

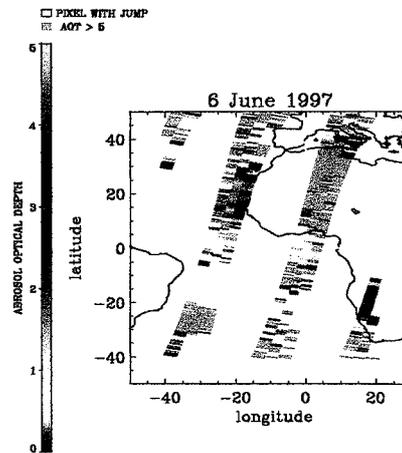


Figure 2.4: Aerosol retrieval obtained flagging data with jumps

Furthermore we can argue that also the following other effects, even though of lesser importance should be taken into account:

- the spectral sea reflectance properties;
- the number of the streams used in Legendre polynomial and the angular discretization of the phase function.

However we found that our radiative model with a different desert aerosol, as proposed by Shettle (1984), and a maritime polluted aerosol was able to properly fit the reflectance measured by GOME. Then, on the basis of the previous considerations, we re-analyzed GOME data using the radiative properties of those new aerosol classes. Final results of our algorithm with new classes are shown in figure 2.15, for June 6, and in figure 2.16, June 8, 1997. Comparing results shown in figure 2.13 with these one obtained using the new aerosol classes, as shown in figure 2.15 we can see that the desert aerosol is well evidenced and covers the areas previously classified volcanic. The same happens with marine polluted aerosol that replaces the continental aerosol. Same results are also obtained for June 8, 1997. These results show that the best fit must be obtained taking into account not only the most probable occurrence of the aerosol, but also its optical properties. In the same figures we have also drawn, by using the green color, the cloud occurrence detected by the aerosol retrieval algorithm. As one can see clouds lay in the same sites detected by cloud detection algorithm.

## 2.7 Conclusions

In this chapter we have shown some preliminary, but controlled examples of aerosol and cloud retrievals carried out using the retrieval method, presented in the previous chapters, applied on GOME data around the African coasts. In order to identify a large desert aerosol transport occurred in this area, we used as reference data those obtained from METEOSAT in the frame of MEDUSE project. Our retrieval method not only shows the presence of the aerosol in the same

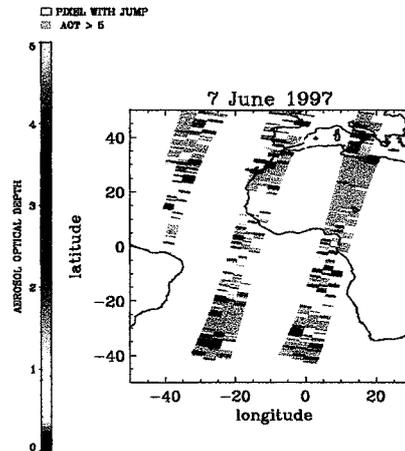


Figure 2.5: Aerosol retrieval obtained flagging data with jumps

location area shown by METEOSAT, but gives AOTs in good agreement with those extracted from sunphotometry data measured at the ground in the same area, carried out in selected days at Capo Verde in the frame of AERONET project. Results show the ability of GOME, not only to capture the presence of aerosol but also to give an accurate quantitative measurement of the AOT. The application in the same area of cloud retrieval algorithm evidences the presence of cloud in the same location shown by METEOSAT, while gives no presence of cloud in the locations we identified as covered with aerosol. Furthermore a properly tuned aerosol ranking algorithm is able to identify the desert aerosol in the same sites where METEOSAT evidences the desert aerosol.

In conclusion, a complete algorithm containing the sea-land mask and with a sea spectral reflectance a priori given, is able to retrieve the aerosol optical thickness with high accuracy and to classify the corresponding class correctly. The cloud retrieval algorithm is able to discriminate between cloudy scenarios and those are covered with aerosol, retrieving clouds in the same areas where were detected by METEOSAT.

## 2.8 Bibliography

- d'Almeida, G. A., P. Koepke, E.P. Shettle: Atmospheric Aerosols. Global Climatology and radiative characteristics. A. Deepak, Hampton Virginia, pp 561, 1991
- Guzzi, R., G. Maracci, R. Rizzi, A. Siccardi: A Spectroradiometer for Ground-Based Atmospheric Measurements Related to Remote Sensing in the Visible from Satellite, Applied Optics, 1985, vol. 5, pp. 2859-2864
- Harrison, L and J. Michalsky and J. Berndt, 1994. The automated multifilter rotating shadowband Radiometer: An instrument for Optical depth and radiation measurements, submitted to Appl. Optics.
- Kaufman, Y.J., A. Gitelson, A. Karnieli, E. Ganor, R.S. Fraser, T. Nakajima, S. Mattoo, B.N. Holben, 1994: 'Size Distribution and Phase Function of Aerosol Particles

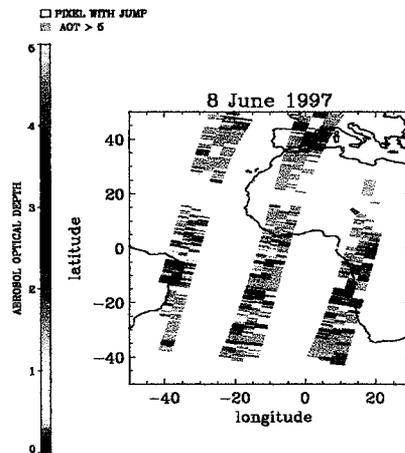


Figure 2.6: Aerosol retrieval obtained flagging data with jumps

- Retrieved from Sky Brightness Measurements', JGR-Atmospheres, 99, 10341-10356.
- Nakajima, T., M. Tanaka and T. Yamauchi, 1983. Retrieval of the optical properties of aerosols from aureole and extinction data. *Appl. Opt.*, 22, 2951-2959.
- Nakajima, T, G. Tonna, R. Rao, P. Boi, Y. Kaufman and B. Holben, 1996. Use of sky brightness measurements from Ground for Remote Sensing of Particulate Polydispersions, in press *Applied Optics*.
- Shiobara, M., Tadahiro Hayasaka, T. Nakajima and M. Tanaka, 1991. Aerosol monitoring using a scanning spectral radiometer in Sendai, Japan. *J. Meteorol. Soc. of Japan*, 60 (1), 57-70.
- Tanré, D., C. Deuaux, M. Herman and R. Santer, 1988. Radiative properties of desert aerosols by optical ground-based measurements at solar wavelengths. *J. Geophys. Res.*, 93, 14223-14231.

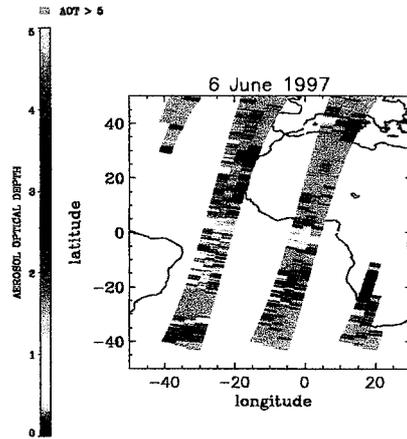


Figure 2.7: Aerosol retrieval obtained with all data

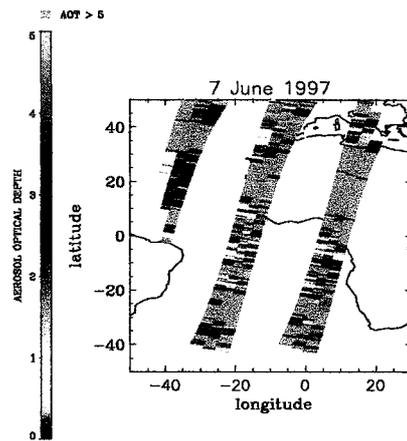


Figure 2.8: Aerosol retrieval obtained with all data

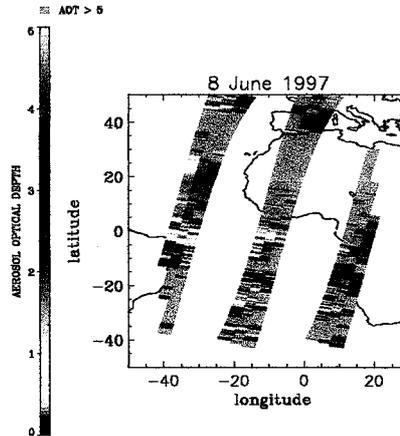


Figure 2.9: Aerosol retrieval obtained with all data

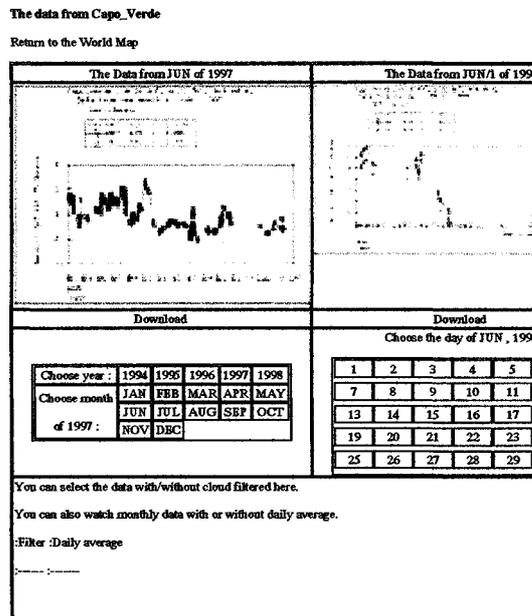


Figure 2.10: AOT retrieved by AERONET at Capo Verde

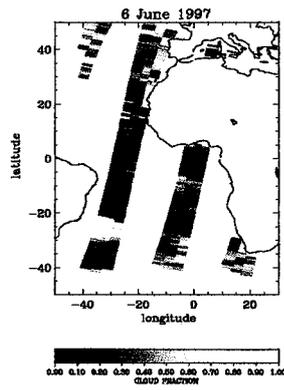


Figure 2.11: Cloud retrieval result

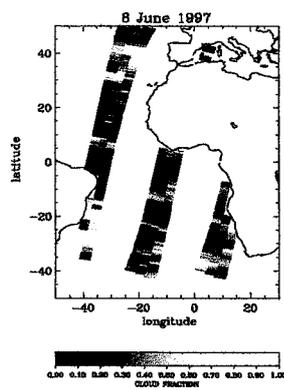


Figure 2.12: Cloud retrieval result

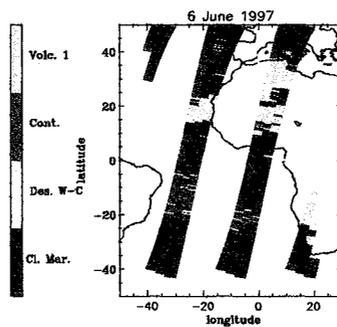


Figure 2.13: Aerosol ranking for June 6,1997

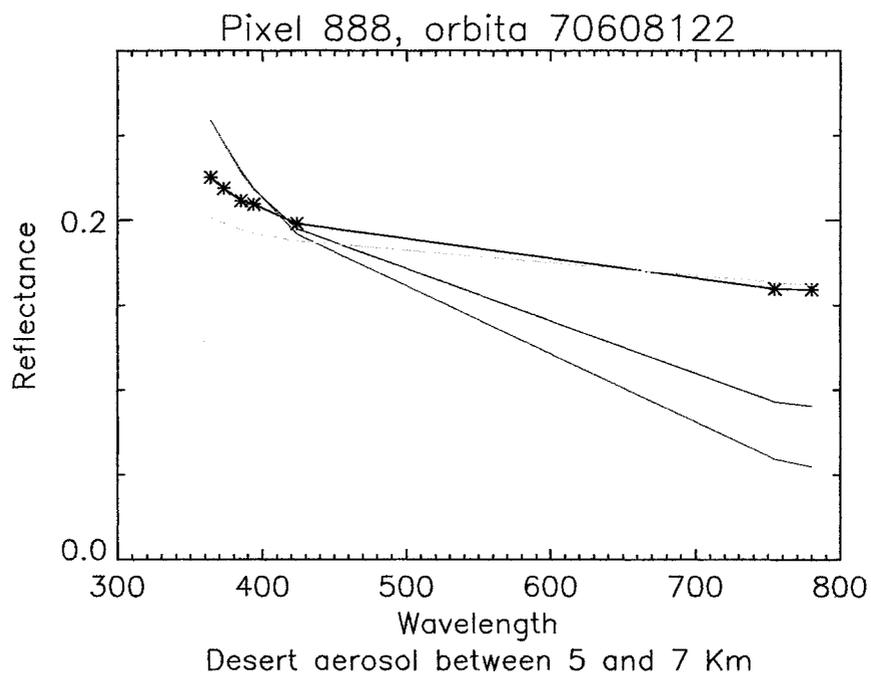


Figure 2.14: Comparison between the four aerosol classes reflectance and GOME data over the same pixels

---

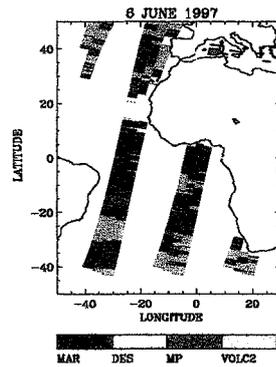


Figure 2.15: Aerosol ranking, June 6,1997

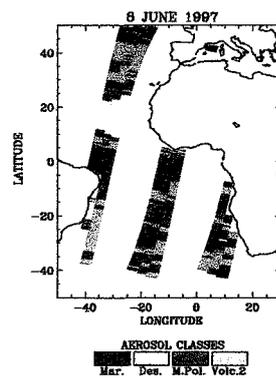


Figure 2.16: Aerosol ranking, June 8,1997

## Chapter 3

# Radiative transfer modeling for a cloudy atmosphere

### 3.1 History of atmospheric RTMs developed to simulate GOME sensed radiances or reflectances in presence of clouds

#### 3.1.1 The approximate model SIMPMOD

SIMPMOD is a RTM developed in IMGAs starting from 1993 for the aim of computing bidirectional reflectance functions around the molecular oxygen A band for a cloudy atmosphere. The evolution of the model can be found in a series of documents (Guzzi and Torricella, 1993; Canossi *et al.*, 1994; Guzzi *et al.*, 1996; Torricella and Cervino, 1996).

The radiative transfer in presence of both cloud droplet scattering and gaseous absorption is rendered especially complex, long and computer-time and memory consuming due to some specific characteristics of this problem, namely

- the strongly forward peaked phase function of the cloud droplets is cumbersome to treat, especially if the RT solution requires its expansion in Legendre coefficients;
- the great optical thicknesses typical of tropospheric water clouds require a RT solution able to accurately cope with high scattering orders;
- the presence of gaseous absorption features requires the RT equation be solved for a great number of spectral points, so to follow the rapid variations of the absorption cross section of the gas;
- the gas absorption strongly depending on temperature and pressure requires an appropriate stratification of the atmospheric model.

All these considerations joint with the need of having a RTM able to on-line compute reflectance spectra to be compared with GOME spectra in the fitting process, lead the authors to develop a model like SIMPMOD that, while able to describe the main physical and optical process driving the reflected radiance, performs drastic but controlled simplifications. The underlying simplifying approach of SIMPMOD is to decouple the two main process occurring at these wavelengths by considering no interaction between the heavy scattering processes that occurs inside the cloud and the oxygen absorption occurring above the cloud. So the cloud top reflectance (computed by means of the asymptotic theory) and the oxygen transmittance are independently computed and separately validated (see for instance: Guzzi *et al.*, 1996 and Torricella and Cervino, 1996, for an exhaustive description of the tests carried out to validate the cloud top reflectance; and Torricella (1996) and Canossi *et al.* for a description of the line-by-line transmittance code developed in IMGAs). The

---

reflectance of the cloud-free part of the scenario is separately computed by means of DISORT taking into account the whole complexity of the RT. In the end the cloudy and clear reflectance spectra are linearly combined using  $f$  to obtain the partially cloudy spectrum (as described in Guzzi *et al.*, 1996).

This choice of “taking to pieces” the reflectance function has the advantage of i) bringing explicit the dependence of the atmospheric reflection function upon  $\tau_c$  (the total optical thickness of the cloud); ii) assess more easily the accuracy with which each function concurring to the reflectance is evaluated and then the accuracy of the whole function; iii) to speed-up substantially the RT computations with respect to full RTMs like DISORT or GOMETRAN.

The basic simplification of neglecting the interaction between the cloud droplets scattering and the oxygen absorption obviously leads to approximate results for the wavelengths inside the absorption band (this phenomenon is also known as “extra-absorption” or photon penetration). An attempt to quantify the errors made in neglecting this phenomenon has been undertaken during the CIRS (1996) by comparing the reflection functions obtained using the two approaches devised for GOMETRAN, one including and the other neglecting the extra-absorption. What is planned as future work is to quantify the effects of such approximation directly on the results of the retrieval.

### 3.1.2 GOMETRAN and GOMETRAN++

GOMETRAN (respectively, GOMETRAN++) is a radiative transfer model specifically developed for the interpretation of GOME data. Its development started in 1991 under the scientific direction of Dr. Vladimir Rozanov [V. Rozanov, private communications].

The decision for the development of a GOME-specific model was driven by the necessity to have available an efficient method for the computation of weighting functions in addition to the radiation field<sup>1</sup>. Starting out from a basic GOMETRAN version, several projects for the extension of GOMETRAN were initiated during the last years. Most of these projects were realized as part of the activities on SCIAMACHY algorithm developments<sup>2</sup> and within the framework of ESA contract studies:

- During the *Tracegas Study* (Burrows *et al.*, 1994), the weighting function scheme for GOMETRAN was realized and a number of important precision assessments were performed.
- Parameterization schemes for terrestrial water clouds were implemented during the *Study of Cloud Detection* (Guzzi *et al.*, 1996).
- The inclusion of Ring spectra was realized in the framework of the *Study of the Ring Effect* (Burrows *et al.*, 1996).

Parallel to these studies, research activities at the University of Bremen lead to several improvements of the original GOMETRAN code. Major extensions consisted of

- the addition of all relevant trace gases for GOME,
- spherical geometry, including refraction,
- the computation of air mass factors,
- a new aerosol parameterization based on WMO recommendations, and also
- Line-by-line and ESFT band modelling for oxygen.

Most of these features are described in various publications on GOMETRAN (Kurosu, 1997; Kurosu *et al.*, 1997; Rozanov *et al.*, 1997a; Rozanov *et al.*, 1997b).

The result of all these research activities lead to a multiplicity of GOMETRAN versions, each of which contained some but not all of the above mentioned extensions. In order to bring together

<sup>1</sup>Weighting functions represent the response of the atmospheric radiation field to a change in atmospheric composition, and are required for the optimal estimation retrieval scheme, which has been proposed and developed for the interpretation of GOME data.

<sup>2</sup>With financial support from the *Deutsche Agentur für Raumfahrtangelegenheiten (DARA)*.

all features within a single model, the **GOMETRAN++ project** was initiated in 1993. Within this project, the development of GOMETRAN was placed under version control.

The first version of GOMETRAN++ (Version 1.0) was released on 1 June, 1995. It presented the common subset of all existent GOMETRAN codes, namely [M. Buchwitz, private communications]

- full multiple scattering,
- various single scattering options,
- plane-parallel geometry,
- satellite viewing geometry only,
- all major trace gases, and
- LOWTRAN aerosol parameterization.

However, many of the features already present in other GOMETRAN codes were not included.

Since the first release of GOMETRAN++, the work on unifying all GOMETRAN features within a single model has continued. The present version of GOMETRAN++ is Version 2.9, and the completion of Version 3.0 is expected before the end of 1997. The most important new features of this version include

- pseudo-spherical geometry, including refraction,
- ground-based viewing geometry,
- air mass factor computations,
- two parameterization schemes for terrestrial water clouds,
- an aerosol parameterization based on a WMO scheme,
- line-by-line and ESFT band modelling for oxygen and water,
- flux profile calculations,
- weighting functions for pressure and temperature, and
- speed-optimizations for matrix decompositions.

Currently, work is under way to improve the weighting function scheme, adding weighting functions for bi-directionally reflecting surfaces and cloud parameters.

## 3.2 Brief description of physical principles and main characteristics of SIMPMOD

### 3.2.1 The basic equation

As seen in section 3.1 the RTM SIMPMOD is based on the simplifying assumption that the bidirectional reflection function of a cloudy scenario computed at wavelengths around a gaseous absorption feature like the A band of O<sub>2</sub> can be expressed with good approximation as a simple combination of the gas transmittance function and of the reflectance of the surface bounding the atmosphere (cloud and/ or ground surface). In practice, if  $\lambda_i$  is one of the wavelengths in the GOME reflectance spectrum in the selected spectral range 750 – 780 nm the SIMPMOD computed reflectance is given by:

$$R_{\text{SIMPMOD}}(\lambda_i) = f \times R_{\text{cloud}}(\lambda_i; \tau_c; \mu, \mu_0, \varphi; A_g(\lambda_i); \text{CT}) \times \int_{\Delta\lambda} B(\lambda - \lambda_i) \times T(\lambda; h_p; \mu, \mu_0) d\lambda + (1 - f) \times R_{\text{clear}}(\lambda_i; \mu, \mu_0, \varphi; A_g(\lambda_i)) \quad (3.1)$$

where:

$\lambda_i$	is the measurement wavelength $750\text{nm} < \lambda_i < 780\text{nm}$ ;
$A_g(\lambda_i)$	is the ground albedo lambertian reflectance at $\lambda_i$ ;
$f$	is the cloud coverage or fraction of cloud present over the pixel ( $0 \leq f \leq 1$ , see fig.3.1);
$\tau_c$	is the optical thickness of the cloud;
$\mu$	is the cosine of the observing zenith angle $\theta$ ;
$\mu_0$	is the cosine of the sun zenith angle $\theta_0$ ;
$\varphi$	is the relative azimuth angle;
CT	is an index identifying the cloud type (see table 3.1);
$R_{\text{cloud}}(\lambda_i; \tau_c; \mu, \mu_0, \varphi; A_g(\lambda_i); \text{CT})$	is the cloud top bidirectional reflectance function;
$B(\lambda - \lambda_i)$	is the instrument response function or slit function;
$h_p$	is the cloud top pressure [hpa];
$T(\lambda; h_p; \mu, \mu_0)$	is the $\text{O}_2$ A band transmittance between the TOA and the pressure level $h_p$ ;
$R_{\text{clear}}(\lambda_i; \mu, \mu_0, \varphi; A_g(\lambda_i))$	is the BDRF of the clear part of the spectrum.

Lets analyze in more details the implications of such representation and the various functions that form the resulting BDRF of the scenario.

- The integral represents the process of convolution between the highly varying transmittance function and the GOME slit function. The convolution integral would extend from  $-\infty$  to  $\infty$ : practically, because the slit function rapidly vanishes when  $|\lambda - \lambda_i|$  increases, the interval  $\Delta\lambda$  has been selected in such a way that when  $B(\lambda - \lambda_i)$  falls below  $10^{-5}$  the integral is stopped (that leads to  $\Delta\lambda \simeq 7.5$  nm). The GOME slit function is the compound hyperbolic approximation with 6<sup>th</sup> order term.
- The cloud reflectance  $R_{\text{cloud}}(\lambda_i; \tau_c; \mu, \mu_0, \varphi; A_g(\lambda_i); \text{CT})$  is taken out the integral because it does not vary rapidly with  $\lambda$ ; rather it is considered constant over the entire fitting window i.e.  $R_{\text{cloud}}(\lambda_i) \equiv R_{\text{cloud}}(750)$  taking as reference wavelength  $\lambda = 750$  nm.
- The cloud BDRF is simply multiplied for the convoluted transmittance. That directly derives from the basic simplifying assumption of SIMPMOD:
  - the radiation from the TOA to the cloud top is considered simply attenuated by the oxygen absorption (no molecular scattering accounted for);
  - the cloud top BDRF includes the processes occurring from this level to the ground due to the cloud droplets and to the ground surface underlying the cloud (no molecular scattering or absorption are considered in this part of the atmosphere).

As far as concerns the scattering processes we are confident that neglecting the Rayleigh scattering does not introduce relevant errors in  $R_{\text{cloud}}$ . The only relevant source of error in this scheme is to neglect the absorption due to the  $\text{O}_2$  occurring *inside* the cloud and its interaction with the cloud droplets multiple scattering. On the other side this simple assumption and the adoption of the asymptotic theory to compute  $R_{\text{cloud}}$  allow to express the dependence upon  $\tau_c$  in a very useful and effective way, oriented to the fitting process.

- The clear part of the spectrum is computed by means of the full RTM DISORT for a stratified atmosphere including Rayleigh scattering, ground albedo and  $\text{O}_2$  absorption effects: this rigorous approach will allows to carefully fit GOME spectra when the cloud coverage  $f$  tends to zero.

### 3.2.2 The cloud BDRF

The cloud BDRF  $R_{\text{cloud}}(\lambda_i; \tau_c; \mu, \mu_0, \varphi; A_g(\lambda_i); \text{CT})$  is computed by means of the asymptotic theory. The AT applies to a plane-parallel, homogeneous, optically thick medium. This approximation holds for optically thick media being an approximate analytical solution of the RTE neglecting

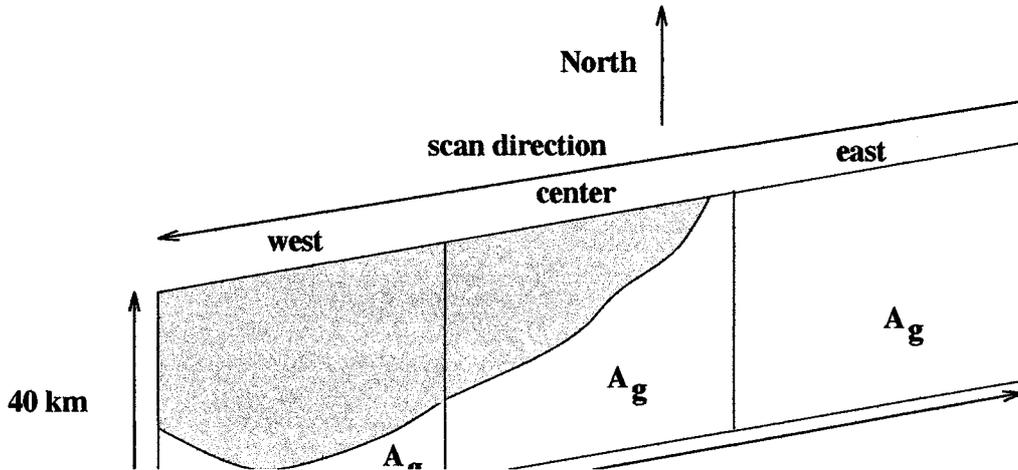


Figure 3.1: Schematic of cloudy GOME pixels: the shadowed area represents a cloud. The west pixel is almost overcast, the central pixel is partially covered and the east pixel is clear. The cloud coverage or cloud fraction  $f$  is defined for each pixel as the area of the part cloud covered normalized to the total area of the pixel.  $A_g$  is the ground lambertian albedo, in this case equal for the three GOME pixels.

CT	Cloud type
1	Stratus I
2	Stratus II
3	Stratocumulus I
4	Stratocumulus II
5	Nimbostratus
6	Altostratus
7	Fair Weather Cumulus
8	Cumulonimbus

Table 3.1: Stephens cloud types

terms that vanish when  $\tau_c \rightarrow \infty$ .

In the AT the reflectance of a cloud having an optical thickness  $\tau_c$  is computed by subtracting a correction factor depending on  $\tau_c$  from the reflectance of a theoretical “semi-infinite” cloud:

$$R_{\text{cloud}}(\lambda_i; \tau_c; \mu, \mu_0, \varphi; A_g(\lambda_i); \text{CT}) = R_{\infty}(\lambda_i; \mu, \mu_0, \varphi; \text{CT}) - \text{CORR}(\lambda_i; \mu, \mu_0; \text{CT}; \tau_c). \quad (3.2)$$

The computational methods needed to apply the asymptotic theory have been deeply analyzed as described in Torricella and Cervino (1996) and in Guzzi *et al.* (1996). The accuracy of the computed reflectances has been checked against the reference RTM DISORT and an agreement better than 1% has been found  $\forall \tau_c > 10$  and better than 2% when  $\tau_c > 5$ . Input to the computations are the optical properties of the clouds namely the single scattering albedo  $\omega_0$ , the phase function  $x(\gamma)$  and its Legendre coefficients  $\beta_i$ . These quantities characterizing the various cloud types are taken from a complete data-set supplied by Ife-Bremen University (T. Kurosu 1995, private communication). The data-base contains the optical properties of eight water clouds following Stephens’ classification (1979) (see tab. 3.1) given for 10 wavelengths in the range 225 and 800 nm. Because the OP of the various fitting cloud types are very similar in the narrow fitting window and, for each cloud, almost constant, we decided of taking as reference the Fair Weather Cumulus (FWC) cloud at the wavelength of 750 nm. Due the great difficulty in representing the strong forward peak of the cloud phase function with a low number of terms in its Legendre expansion,

the  $\delta_m$  approximation was used: this allows to use 299 terms or less maintaining a good accuracy in reflectance computations. A complete data-set of  $R_\infty$  Fourier components values has been precomputed for a set of  $125 \times 125 \mu$  and  $\mu_0$  values; this DS is accessed by SIMPMOD during  $R_{\text{cloud}}$  computation and the suitable value of  $R_\infty$  found by interpolation. In this way the slower part of the computation is avoided.

### 3.2.3 The O<sub>2</sub> transmittance

The high resolution O<sub>2</sub> transmittance has been computed for 21 pressure levels, ranging from 1000.0 hpa to 50 hpa (see table 3.2. The transmittance data are available for  $12800 < \tilde{\nu} < 13400$  (i.e.  $746 < \lambda < 781$ ) nm with a wavenumber step of  $0.002 \text{ cm}^{-1}$  that corresponds to 300,000 spectral points in the fitting window. The line shape is assumed to be described by the Voigt function. The input spectroscopic source was HITRAN86, but the IMGA O2TRANS code is already able to take in input the new Chance data (Guzzi *et al.*, 1996). The transmittances  $T_{\text{HR}}(\tilde{\nu}_i)$  have been computed for one air-mass and for an USSA76 atmosphere; because they are a very high resolution, transmittance values corresponding to different air masses i.e. to various values of  $\mu$ ,  $\mu_0$  are simply computed by

$$T(\tilde{\nu}_i; \mu, \mu_0) = \exp[(1/\mu + 1/\mu_0) \times \ln T_{\text{HR}}(\tilde{\nu}_i)]. \quad (3.3)$$

No	Top pressure [hpa]
0	50.0
1	58.1
2	67.5
3	78.4
4	91.0
5	105.7
6	122.8
7	142.7
8	165.7
9	192.5
10	223.6
11	259.7
12	301.7
13	350.5
14	407.1
15	472.9
16	549.3
17	638.0
18	741.1
19	860.9
20	1000.0

Table 3.2: Top height levels for which the high resolution transmittance files are available. The files contain the high resolution O<sub>2</sub> transmittances from the shown pressure level to the top of the atmosphere (one air mass). In the fitting process only tropospheric cloud top pressures are used (levels below the horizontal line)

### 3.2.4 The clear spectrum

The clear part of the spectrum is computed by means of DISORT for a 50 layers atmosphere. For each layer the Rayleigh scattering cross section is computed for a single particle according to Liou

(1980):

$$\sigma_s = \frac{\alpha^2 128 \pi^5}{3 \lambda^4} \quad (3.4)$$

Here  $\alpha$  is the polarizability given by

$$\alpha \simeq \frac{1}{4\pi N_s} (m_r^2 - 1) \quad (3.5)$$

where  $N_s$  is the numerical density of the particles and  $m_r$  is the real part of the refractive index of molecules given by:

$$(m_r - 1) \times 10^8 = 6432.8 + \frac{2,949810}{146 - \lambda^{-2}} + \frac{25,540}{41 - \lambda^{-2}} \quad (3.6)$$

The mass cross section is obtained multiplying by the Avogadro number and dividing for the molecular mean weight of dry air. The volume cross section is then derived for each layer ( $T_k, p_k$ ) by means of:

$$\sigma_s [\text{km}^{-1}] = \sigma_s [\text{cm}^2 \text{ g}^{-1}] \times \frac{T_0}{T_k} \times \frac{p_k}{p_0} \times \frac{W_{mol}}{V_{mol}} \times 10^5 \quad (3.7)$$

where now the scattering cross section is in the suitable unit of  $[\text{km}^{-1}]$ .  $W_{mol}$  is the molecular weight of the air and  $V_{mol}$  is the molar volume.

The absorption cross section of the  $\text{O}_2$  is rapidly varying in the considered spectral interval. We already mentioned that several thousands of spectral points are needed to follow properly these rapid variations. Because the duration of a single run of DISORT cannot be rendered faster than few seconds, this implies that the  $\text{O}_2$  absorption has to be managed not line-by-line but making use of some approximation. Indeed the k-correlated approach has been adopted to describe wavelength and pressure-temperature dependence of absorption cross sections. The suitable data-set of ESFT coefficients has been supplied by ife (M. Buchwitz, private communication). The ESFT data are given each about 0.2 nm. DISORT spectra so obtained are convoluted with the GOME slit function before be input to eq. 3.1. This convolution is not strictly legitimate, because the obtained clear spectra are not at high spectral resolution. Nevertheless tests carried out in IMGAs showed that

$$R_{\text{HR}} * \text{SLIT\_GOME} \simeq R_{\text{ESFT}} * \text{SLIT\_GOME} \quad (3.8)$$

where  $*$  is the symbol of convolution.

### 3.3 GOMETRAN and GOMETRAN++: relevant aspects for the cloud retrieval task

#### 3.3.1 GOMETRAN++ Cloud Parameterizations

The parameterization schemes of terrestrial water clouds for the radiative transfer model GOMETRAN++ have been developed in the framework of another ESA contract study and are described in detail in the respective study report (Guzzi et al., 1996). Additional information can be found in related publications (Kurosu, 1997; Kurosu et al, 1997). This section therefore gives only a short summary of the various aspects of cloud parameterization in GOMETRAN++, mainly excerpted from the final report of the ESA *Cloud Detection Study* (Guzzi et al., 1996).

#### General Cloud Properties

The scattering and absorption properties of terrestrial water clouds can be described with Mie theory (van de Hulst, 1981), which calculates the scattering of light on spherical particles of a certain refractive index. A particular cloud can then be described as an assembly of spherical

water droplets, distributed according to a  $\Gamma$ - size distribution with a particular Mode Radius  $r_{mod}$  (Stephens, 1979). The optical properties of the cloud, i.e., the scattering phase function and the scattering- and absorption coefficient, are as a result given by the optical properties of the water droplets integrated over the size distribution.

The following 10 cloud types are implemented into GOMETRAN++:

- Two types of stratus
- Two types of stratocumulus
- Altostratus
- Nimbostratus
- Two types of fair weather cumulus
- Two types of cumulonimbus.

For each cloud type, the quantities of scattering and extinction coefficients, single scattering albedo, and Legendre expansion coefficients for the phase function are available. All of these cloud parameters are given as a function of wavelength, covering the spectral range of 200 nm – 2600 nm.

#### Summary of GOMETRAN++ Cloud Parameterizations

The schemes for the parameterization of light scattering in clouds can be divided into two basic types:

- i *Clouds as Diffuse Reflecting Boundaries*, an approximate solution for optically thick clouds, and
- ii *Clouds as Layers*, a quasi-exact parameterization for clouds of arbitrary optical thickness.

The reasons for the choice of an approximative strategy for the case of optically thick clouds is the following:

- a. GOMETRAN++ employs the finite differencing method for solving the radiative transfer equation (Rozanov et al., 1997). In order to avoid numerical instabilities, the optical properties of adjacent atmospheric layers must not differ too drastically. A cloud therefore has to be divided into multiple sub layers such that optical thicknesses of all sub-layers not only follow a smooth gradient between adjacent layers, but also that the total optical thickness for each individual layer does not become too large. Consequently, thick clouds require a large number of sub-layers, which, in turn, results in a penalty in computational speed.
- b. An optically thick cloud can be well approximated by a semi-infinite atmosphere, i.e., a scattering medium which possesses a finite boundary of the depth variable only on one side and extends to infinity at the other. The scattering of light in such atmospheres is invariant to the addition or subtraction of additional (cloud-) layers. It is therefore unnecessary to treat thick clouds in full detail, once the reflection pattern has been established as diffuse reflecting boundaries, scattering the light effectively only at the cloud top. Finite clouds of large optical thickness can then be treated as a combination of an infinitely thick reflecting cloud together with a correction factor for finite optical thickness. Since no additional layers have to be added to the atmosphere, computational speed will not be influenced greatly by this kind of cloud parameterization.

For a detailed description of the mathematical concepts and the computational realization of these cloud parameterizations, the reader is referred to (Guzzi et al., 1996), (Kurosu, 1997), and (Kurosu et al, 1997).

### 3.3.2 GOMETRAN++ Spherical Geometry

For low solar elevations, atmospheric light paths are enhanced due to refraction effects in the (spherical-shell) Earth's atmosphere. In order to account for this effect, an approximate method for the computation of light path lengths in a spherical-symmetric, refractive atmosphere has been developed for GOMETRAN++. With this so-called *pseudo-spherical approach* it is possible to extend the validity regime of GOMETRAN++ to solar zenith angles of up to  $92^\circ$ .

At present no officially published documentation on the pseudo-spherical approach in GOMETRAN++ is available. However, two internal documents of the Institute of Remote Sensing are available: a summary report (Rozanov and Eichmann, 1997) and the report of a diploma thesis (Eichmann, 1995). Both documents can be obtained directly from the authors<sup>3</sup>.

**The design of the pseudo spherical approach** in GOMETRAN++ uses Snell's law of refraction and ray-tracing, together with an iterative scheme, to compute the *tangent height* of the incoming solar beam<sup>4</sup>. Once this quantity has been determined, the optical path length of an arbitrary solar beam from top of atmosphere to any given point inside the refractive atmosphere can be calculated.

It has to be noted that in GOMETRAN++ the *pseudo-spherical correction* is *only applied to single-scattered light* (V. Rozanov, private communications). The quasi-exact cloud parameterization scheme, on the other hand, treats clouds as multiply scattering, plane-parallel layers<sup>5</sup>. Since the pseudo-spheric correction is not applied to the multiple-scattered contribution of the atmospheric radiation field, it is *irrelevant for cloud scattering processes*.

---

<sup>3</sup>rozanov@gome5.physik.uni-bremen.de, eichmann@gome5.physik.uni-bremen.de

<sup>4</sup>The tangent height is defined as the minimum distance between a solar beam and the Earth's surface.

<sup>5</sup>The approximative cloud parameterization scheme need not be investigated here, since it treats the cloud only as a reflecting surface which bounds the atmosphere from below. In particular, atmospheric processes below the cloud top are not considered in this parameterization.

---

### 3.4 Bibliography

- Burrows, J.P. and D. Diebel and B. Kerridge and R. Munro and D. Pemberton and U. Platt and H. Frank: 1994 "A Study of Methods for Retrieval of Atmospheric Constituents, European Space Agency, Final Report, ESA Contract 9687/91/NL/BI, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands
- Burrows, J.P. and U. Platt and K. Chance and V. Rozanov and M. Vountas and A. Richter and H. Haug and L. Marquard: 1996 "Study of the Ring Effect", European Space Agency, Final Report, ESA Contract 10996/94/NL/CN, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands
- Dlugach, J. M., and E. G. Yanovitskij: 1974, "The optical Properties of Venus and the Jovian Planets. II. Methods and Results of Calculations of the Intensity of Radiation Diffusely Reflected from Semi-infinite Homogeneous Atmospheres", *Icarus*, **22**, pp. 66-81
- Eichmann, K.-U.: 1995 "Optimierung und Validierung des pseudo-sphärischen Strahlungstransportmodells GOMETRAN", Diplom Thesis, Institute of Environmental Physics, University of Bremen, Germany (in German)
- Guzzi R., F. Torricella: 1993, "Scientific background for the generation of level 2 data products concerning cloud coverage and classification, aerosol type and load determination from GOME Level 1 data", ESA Report, ESA Ref. LRX/0719/AH-nk.
- Canossi, I., M. Cervino, M. Gabella, R. Guzzi, F. Torricella: 1994, "Aerosol sensitivity and cloud reflectance studies," ESA Report, ESA ref. LRX/0853/AH-nk.
- Guzzi, R and J.P. Burrows and V.V. Rozanov and K. Chance and M. Cervino and T. Kurosu and F. Torricella and P. Watts: 1996 "A Study of Cloud Detection", European Space Agency, Final Report, ESA Contract 10997/94/NL/CN, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands
- Hanel, R. A.: 1961, "Determination of cloud altitude from a satellite", *Journal of Geophysical Research*, **66**, 1300
- Hansen, J. E.: 1971, "Multiple scattering of polarized light in planetary atmospheres, Part II. Sunlight reflected by terrestrial water clouds", *Journal of the Atmospheric Sciences*, **28**, pp. 1400-1426
- King, M. D.: 1983, "Number of terms required in the Fourier expansion of the reflection function for optically thick atmospheres", *Journal of Quantitative Spectroscopy and Radiative Transfer*, **2**, pp. 143, 161
- King, M. D.: 1987, "Determination of the Scaled Optical Thickness of Clouds from Reflected Solar Radiation Measurements", *Journal of the Atmospheric Sciences*, **44**, pp. 1734-1751
- King, M. D., and Harshvardhan: 1986, "Comparative Accuracy of Selected Multiple Scattering Approximations", *Journal of the Atmospheric Sciences*, **43**, pp. 784-801
- King, M. D., L. F. Radke and P. V. Hobbs: 1990, "Determination of the Spectral Absorption of the Solar Radiation by Marine Stratocumulus Clouds from Airborne Measurements within Clouds", *Journal of the Atmospheric Sciences*, **47**, pp. 894-907
- Kurosu, T.: 1997 "Die Modellierung des Strahlungstransports in Wolken für atmosphärische Fernerkundung im ultravioletten und sichtbaren Spektralbereich", PhD Thesis (in German), Mainz Verlag, Aachen; ISBN 3-89653-185-9
- Kurosu, T. and V.V. Rozanov, and J.P. Burrows: 1997 "Parameterization Schemes for Terrestrial Water Clouds in the Radiative Transfer Model GOMETRAN", *Journal of Geophysical Research* (102), No. D18, 21,809-21,823
- Kuze, A., and K.V. Chance: 1994, "Analysis of cloud top height and cloud coverage from satellites using the O<sub>2</sub> A and B band", *Journal of Geophysical Research*, **99**, D7, pp.14481-14491.
-

- Lenoble, J.: 1985, "Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures", Ed. J. Lenoble, A. Deepak Publishing, Hampton
- Rozanov, V.V. and D. Diebel and R.J.D. Spurr and J.P. Burrows: 1997 "GOMETRAN: A radiative transfer model for the satellite project GOME, the plane-parallel version", *Journal of Geophysical Research* (102), No. D14, 16,683-16,695
- Rozanov, V.V. and K.-U. Eichmann: 1997 "Spherical Geometry and Refraction Implemented in GOMETRAN", Internal Report, Institute of Remote Sensing, University of Bremen, Germany
- Sobolev V. V.: 1975, "Light Scattering in Planetary Atmospheres", Pergamon Press, Oxford
- Stamnes K., S-Chee Tsay, Wiscombe W. and Kolf Jayaweera: 1988, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media", *Applied Optics*, 27, 12, pp 2502-2509
- Stephens, G. L.: 1979a, "Optical Properties of Eight Water Cloud Types", Technical paper no. 36, Commonwealth Scientific and Industrial Research Organization, Division of Atmospheric Physics (CSIRO), Melbourne, Australia.
- Stephens, G. L.: 1979b, "Radiative profiles in extended water clouds I: Theory", *Journal of Atmospheric Sciences*, 35, 2111-2122.
- Torricella, F.: 1996, "Simulazione di spettri di radianza nell'intorno della banda di assorbimento A dell O<sub>2</sub>. Relazione conclusiva", IMGa report of professional collaboration, Modena
- Torricella, F., and M. Cervino: 1996, "The reflection of media of large optical thickness," Technical Report no.2/1996, July 1996, Consiglio Nazionale delle Ricerche - Istituto per lo studio delle Metodologie Geofisiche Ambientali, Pisa, Italy, pp.20.
- van de Hulst, H.C., 1981: "Light Scattering by Small Particles", Dover
-

## Chapter 4

# Cloud detection and cloud parameters retrieval: towards the development of ACFA

### 4.1 Introduction

This chapter describes the cloud detection algorithm for GOME which has been developed in the framework of this study. This algorithm consists of two parts:

- i A “cloud pre-processor”, utilizing GOME PMD signals only. In this first step of cloud detection, the information content of the high spatial resolution PMD signals is used to obtain the cloud fraction for each PMD sub-pixel. In addition, a priori information on the most likely cloud type, cloud top pressure, and cloud optical thickness is obtained from an ISCCP data set.
- ii The second step consists of the fitting of simulated spectra to the high spectral resolution array data in regions of oxygen absorption ( $O_2$ -A and B bands). Together with the a priori information from the PMD cloud pre-processor, information on cloud optical thickness and cloud top pressure is obtained.

In the following sections, a short introduction to the algorithm and its history is given: its principle design is outlined, and the flow of logic for the actual cloud detection process is presented. A detailed description of the PMD cloud detection algorithm can be found in a separate document (Kurosu, 1997a), annexed to the final report of the CADAPA study.

#### 4.1.1 The ACFA (Advanced Cloud Fitting Algorithm)

The measurements of satellite instruments such as GOME and SCIAMACHY, which observe Solar light reflected at the Earth’s atmosphere in nadir viewing geometry, are strongly influenced by the presence of clouds. The standard ground pixel of GOME for a scan angle of  $\pm 30^\circ$  is of size  $40 \times 320$  km<sup>2</sup>. Results from a study at *Meteo France* (Derrien, 1992) extrapolated to GOME ground pixel size, show that the probability for cloud contamination of such a pixel lies above 99.8% (Kurosu, 1997b). Neglecting cloud effects will lead to large errors in the data products of these instruments.

Methods for cloud detection and the simulation of cloud effects are clearly required. While suitable algorithms to simulate cloud radiative properties already exist (Kurosu, 1997b; Kurosu et al., 1997), the status of cloud detection for the GOME instrument is still not satisfactory. Currently, cloud detection for GOME is performed using the so-called **Initial Cloud Fitting Algorithm (ICFA)**, developed by *Kuze&Chance* (1994). It is based on a non-linear least squares fit of GOME radiances

---

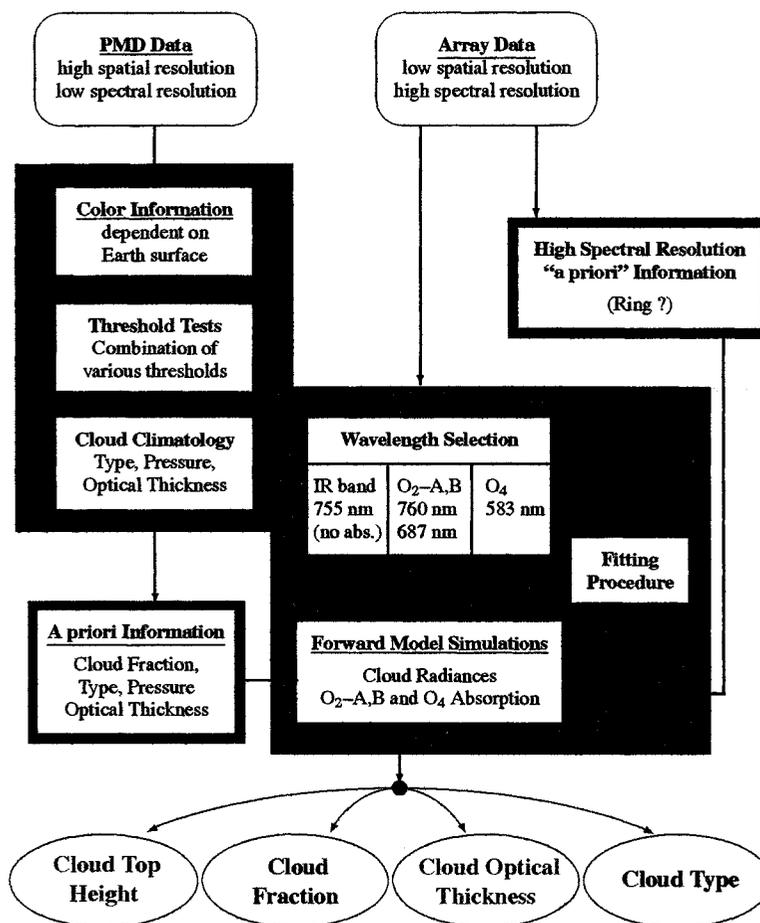


Figure 4.1: ACFA, the Advanced Cloud Fitting Algorithm. A proposed cloud detection scheme for GOME.

inside the A- and B-bands of molecular oxygen, and it returns the cloud fractional cover for a standard GOME ground pixel.

Efforts to improve cloud detection for GOME have been made in the framework of two previous ESA studies: *Fouquart et al.* (1994) proposed a scheme for the utilization of PMD signals for cloud detection, *Guzzi et al.* (1996) extended the ICFA scheme, including cloud top height, cloud optical thickness, and cloud type into the fitting. However, both algorithms suffered from shortcomings which made them unsuitable for an immediate application to GOME data.

The cloud detection algorithm described here focuses on the utilization of PMD signals for cloud detection. It is based that proposed by *Fouquart* (1994), but extends it in many ways. Its aim is to provide information on the cloud contamination of a PMD sub-pixel, taking into account fractional cloud cover and a priori information from cloud climatologies on the most probable cloud type and cloud top pressure. The algorithm is designed such that it can be used either in a stand-alone manner or as cloud pre-processing within a more extensive cloud detection algorithm. It is therefore a major step of the **Advanced Cloud Fitting Algorithm (ACFA)**, a cloud detection algorithm, utilizing both, PMD detection and fitting methods. A conceptual algorithm of this type was proposed by IFE&IMGA at the end of the last ESA study *A Study of Cloud Detection* (*Guzzi et al.*, 1996). Figure 4.1 shows the layout of the ACFA algorithm.

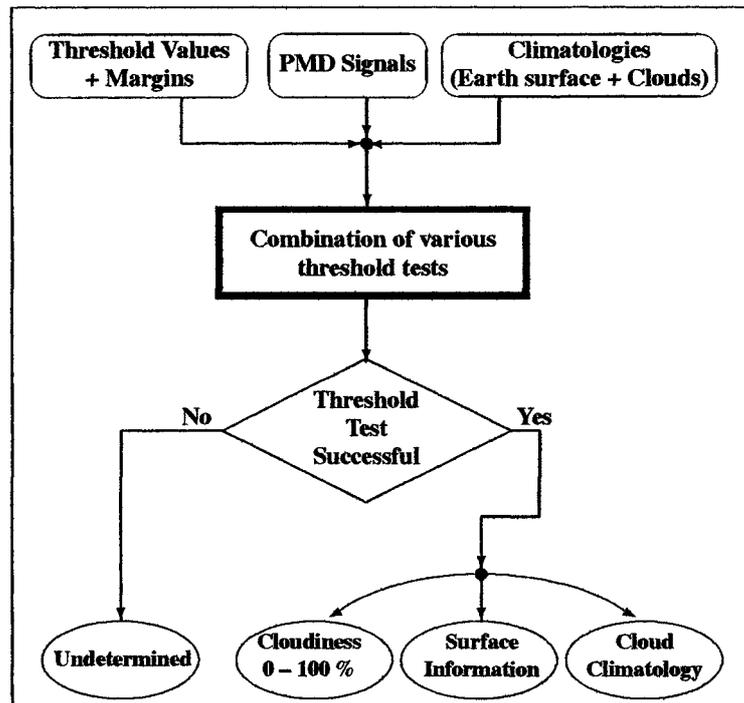


Figure 4.2: General outline of the PMD cloud detection algorithm.

## 4.2 The PMD Cloud Detection Algorithm

### 4.2.1 Design of the PMD Algorithm

The PMD cloud detection algorithm relies on the utilization of PMD radiances alone. The PMD signals<sup>1</sup> are compared with user-defined **threshold values**  $P_{\min, \max}$  having appropriate margins  $\delta_{\min, \max}$ , for clear and cloudy conditions. A PMD sub-pixel is then declared as **clear** respectively **cloudy**<sup>2</sup>, if the value of the PMD signal falls within  $P_{\min, \max} \pm \delta_{\min, \max}$ , and otherwise **undetermined**. For the latter, an approximation is made which enables the percentage of clouds in the pixel to be determined.

A combination of various thresholds is used in order to maximize the retrieved information: tests on **absolute** PMD values are combined with **ratios** between two PMD signals. For each sub-pixel, a variety of different tests is performed before the decision between **clear**, **cloudy**, or **undetermined** is made. Since reflection from the Earth's surface strongly influences the signals, different strategies are employed for the various surface types.

For every PMD pixel, the algorithm returns an integer number between 0 and 100, representing the **cloudiness** of the pixel, with 0 (100) corresponding to **clear** (**cloudy**) condition. In cases where the algorithm concludes **undetermined**, the value of the PMD signals within the interval  $[P_{\min}, P_{\max}]$ , and this value is translated to the cloud cover fraction, represented by an integer number between 1 and 99. Where this fails or the algorithm cannot reach a decision due to, for example, sun-glint conditions, the PMD pixel is flagged with  $-1$ . In such case, a post-processing is invoked after all PMD sub-pixel have been classified, which reduces the number of sub-pixel that failed to be identified.

<sup>1</sup>The PMD signals are not used "as is" in the GOME level 1 product: they are first converted to spectral albedos by division by  $\mu_0$ , the Cosine of the Solar zenith angle. However, here and throughout the remaining part of this document, the terms *PMD signals*, *radiances*, or *measurements* will always be understood to include the division by  $\mu_0$ .

<sup>2</sup>Here and throughout the chapter, **cloudy** refers to completely overcast scenes.

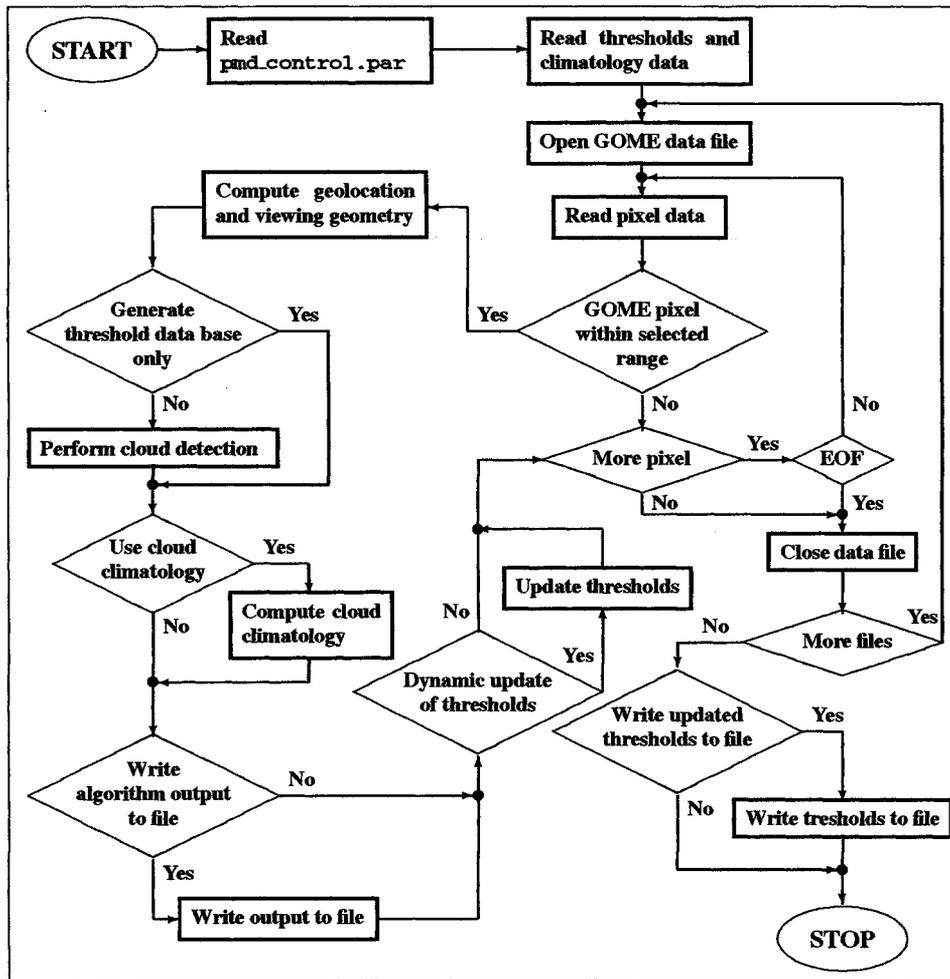


Figure 4.3: Flow-diagram for the PMD cloud detection algorithm.

The general layout of the PMD cloud detection algorithm is displayed in figure 4.2. At its center, represented by the square box in the middle of the diagram, lies a variety of threshold tests, i.e., comparisons of the PMD signals to be evaluated with predefined thresholds. As input, these test require, of course, the PMD signals themselves, and also a set of (either dynamic or static) threshold values with appropriate margins, and certain climatological data:

- The *PMD data* are read from an extracted standard GOME level-1 file. At this time, all necessary additional information such as the viewing geometry and geolocation are read from the level-1 file as well.
- *Thresholds* corresponding to clear (minima) and cloudy (maxima) conditions have to be provided in the form of precomputed input files, and the respective margins, i.e.,  $\delta$ -regions around the threshold values, have to be specified by the user.
- *Climatological data* is required for the processing of cloud detection (e.g., Earth surface types), and in addition it provides information such as cloud top height which cannot be determined from the PMD approach (e.g., ISCCP cloud climatology).

Provided the various threshold tests for a specific PMD pixel have been successful in some way, it is possible to determine the cloudiness of the pixel. For this case a number between 0 and 100, representing the percentage of cloud cover of the PMD pixel, is returned, and, if requested,

information on surface type, ground height, geolocation, and cloud climatology is returned by the algorithm. If, on the other hand, the algorithm fails to classify the PMD sub-pixel, the pixel is flagged with -1. Figure 4.3 shows the flow of logic for the PMD cloud detection algorithm in more detail. Various aspects like the algorithm I/O and the logic for cloud detection are assessed below.

The crucial element in a cloud detection algorithm based on spectral reflectances is the appropriate choice of thresholds. The center of the subroutine containing for the cloud detection process contains a number of straightforward If-Then-Else structures. Extreme care has to be taken when selecting the threshold values user to derive cloud cover fraction of a pixel. It has already been shown by *Fouquart* (1994) that an inappropriate choice of thresholds will lead to large uncertainties in the algorithm results. A substantial part of the development of the PMD cloud detection algorithm for GOME was therefore dedicated to the compilation of a set of GOME-PMD threshold data, which provide an optimum input for any cloud detection processing of GOME data.

The rationale for cloud detection (i.e., the specific type of threshold comparison) differs between different surface types: the spectral dependence of the green and red PMD reflectances, for example, are different over oceanic or continental areas. Such differences can be utilized for the actual cloud detection. Also, extreme care has to be taken over areas with a high spectral reflectance, such as ice and snow. Therefore the PMD algorithm enables thresholds and threshold margins for various surface types.

The values of the threshold margins are strongly dependent on the type of threshold data base used. While the GOME-PMD threshold set performs well as is, i.e., with margins equal to zero<sup>3</sup>, instrument non-specific thresholds might require a substantial fine tuning using these margins.

#### 4.2.2 PMD Reflectances and Geolocation Computation

Before cloud detection for a PMD sub-pixel can be performed, several tasks have to be carried out:

- i The PMD signals have to be converted to a spectral albedo, in order to assure that threshold tests are performed for quantities independent of viewing geometry. This is achieved by dividing all PMD signals by the cosine  $\mu_0$  of the solar zenith angle.
- ii All comparative tests require that the exact geolocation of the PMD sub-pixel is known. The following strategy has been adopted: starting from the geolocation information in the GOME level 1 data file<sup>4</sup>, the latitudes and longitudes of all PMD sub-pixel centers are computed from spherical trigonometry. These center coordinates are then located in the Earth surface grid on which a data set (Earth surface types, threshold sets, cloud climatologies, etc.) is given. The value determined in this manner is used for evaluating (classifying) the PMD signal.

It should be noted that, in the current version of the algorithm, no correction is made for the different viewing geometries at the satellite and at the top of the atmosphere (TOA), i.e., directly above the point on the Earth's surface where the incident direct solar radiation reaches the surface. While in the GOME level 1 data all zenith angles are given at the point of the instrument, the appropriate angles for a conversion of GOME signals to spectral albedos are those at TOA. Such a correction is performed only in the GOME level 1 to 2 processing (DLR-DFD, 1996).

Taking this correction into account requires either a patch for the viewing geometry from the GOME level 1 to 2 data processor (DLR-DFD, 1996) or the use of GOME level 2 data. The former approach could not be taken due to the unavailability of the sources, and the latter was rejected for the following reasons:

- Any processing of an additional level 2 data file will increase the complexity of the algorithm, and will make it more unwieldy to use.

---

<sup>3</sup>Such a performance is not too surprising, since the GOME-PMD thresholds are compiled from minima and maxima of PMD signals.

<sup>4</sup>For each standard GOME pixel, the coordinates of the four pixel corners and the pixel center are given. Out of these, only the corner coordinates are used in the PMD cloud detection algorithm.

---

- Most important, GOME level 2 data are not available for all of the corresponding level 1 data; a test for the availability of auxiliary level 2 data files would be required, and parts of the GOME level 1 data would have to be processed with an uncorrected viewing geometry. Such mixture is clearly undesirable for the cloud detection processing.

It is not to be expected that the use of the uncorrected viewing geometries will lead to large errors, since the difference between zenith angles at TOA and at the location of the satellite should not exceed  $4^\circ$  for a standard GOME pixel<sup>5</sup>. Nevertheless, the implementation of the correction patch should be considered for future versions of the PMD cloud detection algorithm.

### 4.2.3 The Threshold Margins

For each PMD signal  $P$  a set of two thresholds  $P_{\min, \max}$  can be specified, against which they are tested. The additional specification of margins  $\delta_{\min, \max}$  around the thresholds leads to a more flexible comparison scheme, as these margins can be selected arbitrarily small or large.

Any comparison of PMD signals  $P$  with threshold values is of one of the following two types:

$$P < P_{\min} + \delta_{\min} \quad \text{or} \quad P > P_{\max} - \delta_{\max}. \quad (4.1)$$

In the PMD cloud detection algorithm the  $\delta_{\min, \max}$  are specified as a **fraction of the threshold values**  $P_{\min, \max}$ , i.e., in the form

$$\delta_{\min, \max} = \varepsilon_{\min, \max} P_{\min, \max}. \quad (4.2)$$

The quantities  $\varepsilon_{\min, \max}$  are those which are in fact specified in the control file. The exact form of the comparisons (refeq:comp-1) as performed by the algorithm is thus

$$P < P_{\min}(1 + \varepsilon_{\min}) \quad , \quad P < P_{\max}(1 - \varepsilon_{\max}). \quad (4.3)$$

The  $\varepsilon_{\min, \max}$  can be used to selectively switch off the use of PMD signals for cloud classification (Kurosu, 1997a).

### 4.2.4 PMD Cloud Detection: Flow of Logic

The base of the cloud detection process is given by a set of **If-Then-Else** statements, which compare the PMD signals with a set of user-defined thresholds. The general flow of logic for this comparison is identical for all Earth surface types, as can be seen from figures 4.4 and 4.5, which display the flow diagrams of the cloud detection process over oceanic and continental areas:

- after an initial check for sun glint, various minimum threshold tests for cloud free conditions are performed; if this fails,
- the PMD signals are tested against threshold maxima for 100% cloudy pixel; should this also fail,
- the computation of the pixel cloud cover fraction is performed; only
- if all tests fail, and if no linear interpolation or constant extrapolation from successfully classified pixels is possible, the pixel remains **undetermined**.

While the *structure* of the threshold tests is identical for all surface types, the *actual tests* vary due to the different spectral reflection properties of the surfaces. This is most easily understood when considering the ratio  $P_3/P_2$ , i.e., the ratio of the red and green PMD signals, over **ocean** and **land**:

---

<sup>5</sup>This can be easily verified considering the fact that GOME orbits the Earth in approximately 800 km altitude, and its maximum scan angle is  $\pm 30^\circ$ . For an Earth radius of 6370 km, the maximum difference is less than  $30^\circ \times (800/6370) \approx 3.77^\circ$ .

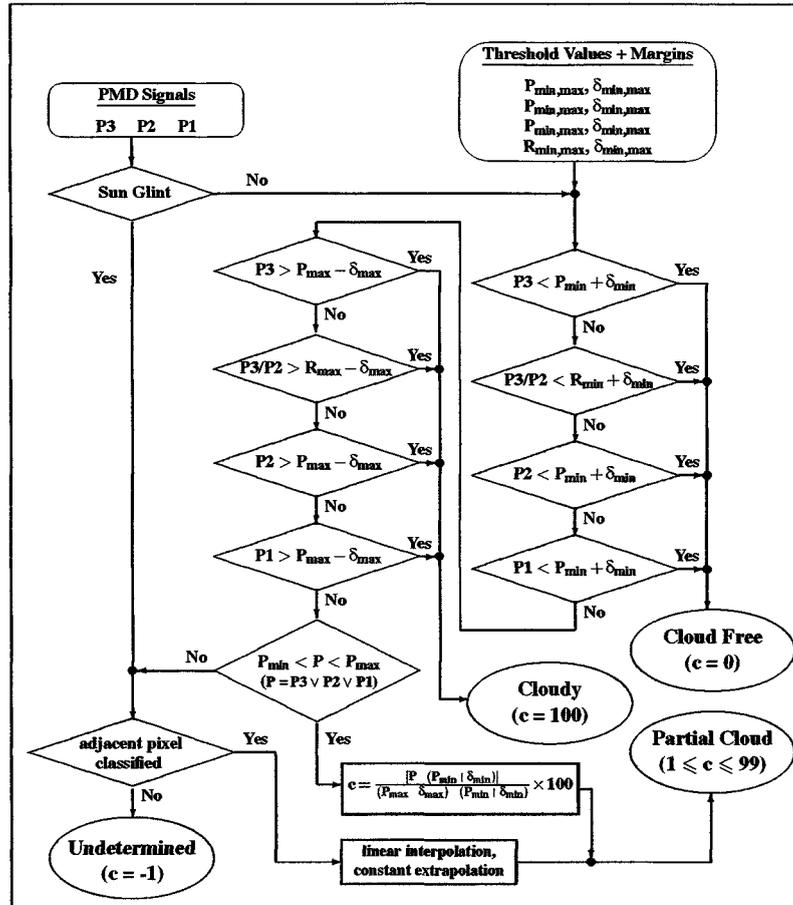


Figure 4.4: Threshold test over oceanic areas (snow and water type surfaces).

- For **cloud** conditions this ratio is approximately 1, independent of the Earth surface type. This is due to the fact that Mie scattering, which describes the scattering of light in clouds, has only a small dependence on wavelength.
- For **cloud free** conditions however, the ratio of the PMD signals strongly depends on the underlying surface type:
  - **Over Oceans**, very little red light is reflected. It follows that  $P_3 < P_2$ , leading to  $P_3/P_2 < 1$ .
  - **Over Land**, this situation is reversed: green light is absorbed more strongly than red over continental areas, and thus  $P_3 > P_2$ , from which follows  $P_3/P_2 > 1$ .

This spectral dependence has to be taken into account when performing threshold tests for various surfaces.

The PMD cloud detection algorithm in its current version discriminates between five different surface types, as classified by *R. Guzzi* (*R. Guzzi*, private communications). These types are sand, snow, soil, vegetation, and water. The same type of tests are performed for snow and water surfaces on the one hand, and sand, soil, and vegetation surfaces on the other. The difference concerns the use of the red and green PMD signals and their ratio, taking into account their spectral dependence as mentioned above: given thresholds  $R_{\min, \max}$  with margins  $\delta_{\min, \max} > 0$  for the signal ratio, the following rationale for cloud classification applies:

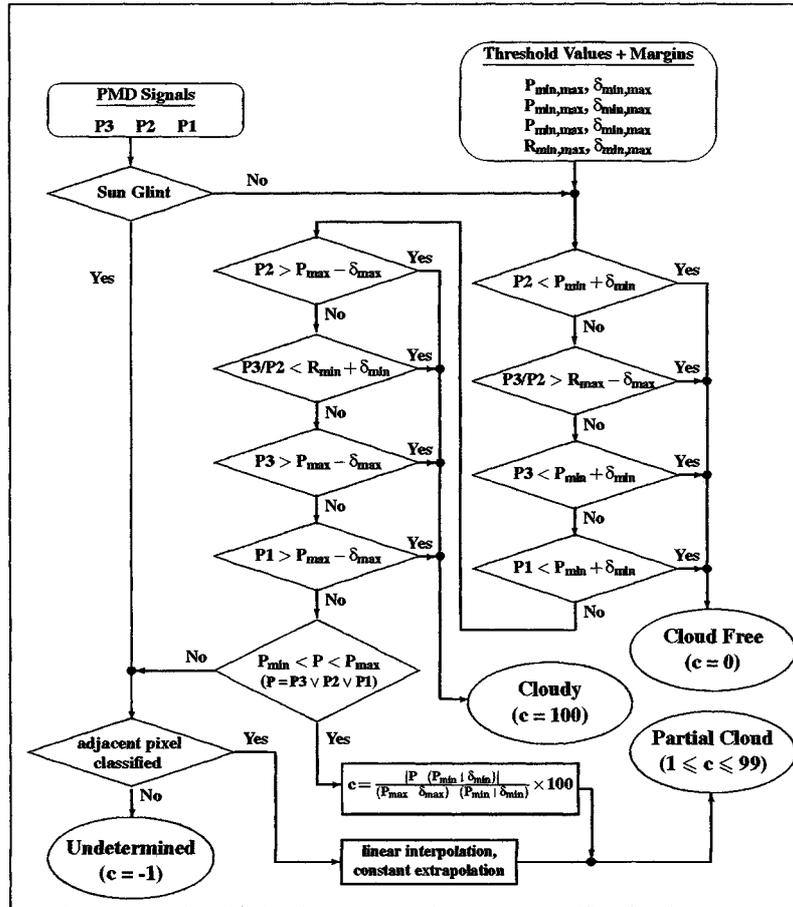


Figure 4.5: Threshold test over continental areas (sand, soil and vegetation type surfaces).

	Land	Ocean
Cloud Free	$P_3/P_2 > R_{max} - \delta_{max}$	$P_3/P_2 < R_{min} + \delta_{min}$
Cloudy	$P_3/P_2 < R_{min} + \delta_{min}$	$P_3/P_2 > R_{max} + \delta_{max}$

**A note on the use of PMD1 (blue):** In the original proposal for an algorithm using PMD reflectances for cloud detection (Fouquart, 1994), the use of the blue PMD was discouraged due to its corruption by Rayleigh scattering. As can be seen from the diagrams 4.4 and 4.5, the  $P_1$  signal has been included with the threshold tests. The rationale behind this is to provide the user with the possibility to utilize these data if she/he prefers to do so. Since it is always possible to switch off threshold tests with an appropriate choice of threshold margins the user can decide whether or not to include this PMD (Kurosue, 1997a).

#### 4.2.5 Determination of Pixel Cloudiness

In those cases where a PMD sub-pixel cannot be classified as neither completely clear or cloudy, the PMD algorithm attempts to compute the cloud fraction of the pixel. In order for this classification to be successful, at least one of the three PMD signals  $P_i$  ( $i = 1, 2, 3$ ) has to lie within the interval of minimum and maximum thresholds  $P_{i,min}$  and  $P_{i,max}$  (see figures 4.4 and 4.5), i.e.,

$$P_{i,min} < P_i < P_{i,max}. \tag{4.4}$$

If this is the case, then the computation of the pixel cloudiness  $C$  follows from linear interpolation:

$$C = \frac{|P_i - P_{i,\min}|}{P_{i,\max} - P_{i,\min}}. \quad (4.5)$$

The selection of the  $P_i$  and  $P_{i,\min,\max}$  for cloudiness determination follows a certain ranking<sup>6</sup>: over oceanic areas, the ranking is (1)  $P_3$ , (2)  $P_3/P_2$ , (3)  $P_2$ , (4)  $P_1$ , while over land the tests are performed in the order (1)  $P_2$ , (2)  $P_3/P_2$ , (3)  $P_3$ , (4)  $P_1$ . The first combination of thresholds and PMD signals which fulfills the condition (refeq:ival) is used for the cloudiness computation (refeq:cldness.)

#### 4.2.6 Sun-Glint

Sun-glint conditions pose a problem for the reflectance test, since the enhanced signals might lead to an erroneously high cloud cover. The following two criteria have to be fulfilled for a sun-glint scenario:

- A reflecting surface like snow or water;
- light from the sun reaches the instrument in forward scattering, i.e.,
  - the relative azimuth angle  $\phi$  between sun and instrument lies within the interval  $[-\pi/2 < \phi < \pi/2]$ , and
  - the difference  $\vartheta$  between the solar and line-of-sight zenith angles is less than  $0.5^\circ$ .

If both conditions are fulfilled, no threshold tests are performed, and in the first instance the pixel is classified as undetermined.

However, under most circumstances a post-processing of such pixel will allow the determination of their cloudiness: after all 16 PMD sub-pixel have been processed, any sun-glint contaminated pixel are revisited, and their classification is attempted based on the following approach:

- If sun-glint was present for a start or end PMD pixel (numbers 1 or 16), the nearest PMD pixel is sought which did not contain sun-glint and also could be classified successfully, and its cloudiness is assigned to the sun-glint pixel; this corresponds to a “constant extrapolation” procedure.
- If sun-glint conditions are present for non-start or end sub-pixels, the nearest successfully classified PMD pixel to their left and right are sought, and these values are linearly interpolated to the sun-glint contaminated pixel.

In this way, most undetermined pixels can be classified. Only if all pixel contain sun-glint or none of them could be classified successfully for a different reason does the algorithm return the value -1.

#### 4.2.7 Dynamic Update of Thresholds

The method of dynamically updating thresholds provides the only reliable way to a reliable cloud classification and is strongly preferable over the use of static thresholds (Fouquart, 1994). The idea is that, after a sufficient number of cloud detections, the resulting threshold set will provide the optimum basis for cloud classification, as it has been derived directly from the instrument’s signals.

The PMD cloud detection algorithm allows a dynamic update procedure for the PMD thresholds. This updating is done **only** in those instances where a PMD pixel has been classified as 100% cloudy or completely cloud free. In the latter case, the minimum threshold value will be updated,

<sup>6</sup>This ranking can be influenced by de-selecting the use of certain PMD signals for cloud detection (Kurosu, 1997a).

in the former case it is the value for the maximum threshold. If the corresponding PMD signal is found to be either larger than the maximum or smaller than the minimum, the respective threshold value is replaced with the PMD signal. After all data files have been processed, the newly derived threshold data base can be written to a file (see figure 4.3) and thus is available as input for further calls of the PMD algorithm.

#### 4.2.8 Tests and Validation

At the CADAPA progress meeting held on 9 September, 1997 in Roma, Italy, it was originally decided to validate the PMD cloud detection algorithm against ATSR data; these data should be provided by ESA for use by *ife* (Levoni, 1997). However, this validation scheme could not be carried out since ATSR data were not available to ESA in the form as required by *ife* (*J. Callies*, private communication).

#### Comparison with AVHRR Images

A thorough validation of the PMD cloud detection algorithm using data from other satellite instruments such as ATSR and AVHRR is a time-consuming task. Given the limited time until the end of the study it was therefore decided to limit the validation of the algorithm to a first qualitative comparison with AVHRR images. These images are readily available from Worldwide-Web servers (NERC, 1997) for every day of the year over the period from January 1993 until December 1997.

As recommended by the WWW-server for the AVHRR data, Channel 2 images were taken for the comparison. This AVHRR channel measures in the visible to near-infrared region from 0.725–1.10  $\mu\text{m}$ .

The following four days were randomly selected for a qualitative comparison: 8 October 1996, 6 March 1997, 4 April 1997, and 7 July 1997. The PMD algorithm and user's manual document contains the AVHRR for these days, and also shows the results from the PMD cloud detection algorithm (Kurosui, 1997a). Since the GOME–AVHRR comparison has to be performed by eye and the images are not included here, the reader is referred to the PME algorithm description document. Only the conclusions are repeated here.

#### 4.2.9 Conclusion

The conclusion from the above validation is that the PMD algorithm developed here achieves aim of a cloud detection approach for GOME based on PMD reflectances. In particular, the following points have to be noted:

- i The PMD cloud detection is successful independent of the Earth surface type (snow and ice surfaces being currently the only notable exception).
- ii Not only can cloud presence be detected sufficiently well, but also the cloudiness of a PMD sub-pixel can be retrieved with a certain accuracy: the correlation between the AVHRR grey-scale and the GOME–PMD color scale, most obvious in the center regions of the large cloud fields, indicates that computation of the PMD sub-pixel cloudiness in fact leads to sensible results for the cloud cover fraction of the PMD pixel.

The next step in the validation procedure should consist of a quantitative analysis of the retrieved cloud cover fractions. Data from high spatial resolution instruments like AVHRR or ATSR should be used for this approach.

### 4.3 Cloud Fitting Algorithm

The Cloud Fitting Algorithm developed in the frame of this study is based on the principles firstly described by Kuze and Chance (1994) and developed later on by Guzzi *et al.* in the frame of the

---

ESA contract "A study of cloud detection" (1996). In the original method the cloud height and a cloud coverage parameter <sup>7</sup> are determined by least square fitting GOME measured reflectances to computed reflectance spectra at wavelengths inside and near the molecular oxygen A band. The knowledge of both cloud coverage, or sinonimously cloud fraction, and cloud top height are very important in the retrieval of total ozone retrieval from GOME data and in the distinction between the stratospheric and tropospheric O<sub>3</sub> burdens.

A direct application of the method presented by Kuze and Chance is realized in the operative cloud fitting algorithm (ICFA) by DLR-DFD.

In this section, after briefly revising the previous cloud fitting algorithms developed for the GOME instrument, the main novelties and improvements of the proposed fitting (called hereinafter Bologna Cloud Fitting Algorithm, BOCFA) are presented along with its schematic structure, the main required inputs and, finally, the results of the first applications of the method to selected GOME scenarios.

### 4.3.1 The use of the O<sub>2</sub> A band to derive cloud top height

The use of the O<sub>2</sub> A absorption band around 760 nm to remotely sense the cloud-top pressure is know from 1961, when the first paper by R. A. Hanel on this topic appeared. This technique attracted considerable attention both in United States and the Soviet Union in the 1970s and early 1980s. Exhaustive reviews of the topic are given by O'Brien and Mitchell (1992) and by Wu (1985). In practice the rationale of the method is to derive the transmittance (and then the optical depth) of the molecular oxygen between the instrument and the reflecting surface (cloud or ground) by comparing the intensities or reflectances at two wavelengths one inside and one near but outside the absorption band. Since the oxygen is an uniformly mixed gas, its vertical distribution reflects the vertical stratification of atmospheric pressure and temperature (i.e. basically depends only upon the atmospheric model) slightly depending on local varying atmospheric conditions. Then the derived oxygen transmittance gives the pressure of the surface that reflected the solar light when compared to pre-computed transmittance templates. The method is particularly suitable for passive remote sensing of the pressure because the O<sub>2</sub> A band is narrow and not overlapped by other absorption bands.

### 4.3.2 The Kuze and Chance method for GOME

The application of the O<sub>2</sub> A band method to the GOME instrument has been firstly proposed by Kuze and Chance in 1994. In general the two authors investigated the applicability of the principle to satellite-based spectrometer having large ground pixels. They recognized the statistical relevance of partially cloud-covered fields of view for instruments having IFOV of several thousands km<sup>2</sup> as GOME. In fact, it is rare that such large IFOVs are covered evenly by a cloud layer or are completely clear i.e. cloud free: partially cloud covered scenarios are very likely to occur and then a method able to cope with such common scenarios has to be devised. For this reason Kuze and Chance developed a method able to determine simultaneously the cloud top height and a coverage parameter by least square fitting to calculated reflectance in the oxygen bands, exploiting the moderately high spectral resolution of GOME and allowing for partially cloudy scenarios to be detected. A grid search method is used to search the parameter space to minimize an appropriate sum of squares of differences between measured (or pseudo-measured) and computed spectra.

### 4.3.3 The ICFA

The ICFA algorithm (Spurr, 1994) is a first practical application of the method presented in the previous sub-section. It is routinely used to produce operational cloud level 2 products in the GDP. Being devoted to an operative aim the algorithm makes use of several *a priori* assumptions and simplifications that tend to reduce the computational time necessary to process each GOME pixel.

---

<sup>7</sup>The coverage parameter is a combination of cloud top reflectivity, Earth ground albedo, scene averaged albedo and fractional coverage.

---

As already discussed in section 3.2 the computation of atmospheric spectra in presence of both multiple scattering processes and gas absorption represents one of the most complex and time consuming task in the RT. Most of the simplifying assumptions made in the ICFA algorithm can be summarized as follows:

- the clouds are approximated as bi-directionally reflecting lower boundaries of the atmosphere;
- the cloud top reflectance is computed by assuming an infinite cloud optical thickness resulting in a systematic underestimation of the cloud fraction;
- no scattering phenomena are considered: all the physical process other than O<sub>2</sub> absorption and cloud top and ground reflection are approximated by a closure term almost constant in the fitting window;
- the cloud top pressure is considered fixed to a canonical value depending only on latitude;
- the azimuth dependence of computed cloud reflectance is neglected;
- the non-linear fitting problem is transformed in a linear regression making use of the simplifying assumptions adopted;
- the viewing geometry used to compute the transmittance set is fixed ( $\theta_0 = 60^\circ$ ,  $\theta = 22^\circ$ ); the 7 lowest pressure levels in the model atmosphere are considered, corresponding to tropospheric cloud tops and are stored in a data-base.
- If the slit function and the calibration wavelengths are known, pre-convoluted transmittance data-bases are used, then introducing the need of a selection of non-saturated transmittances when scaling to other geometries.
- For a given cloud-top pressure and ground-level pressure the convoluted transmittance values needed are found by interpolation of the convoluted data-base.
- The Earth ground albedo is assumed to be Lambertian.

#### 4.3.4 The cloud fitting algorithm from CIRS

A step forward respect to the ICFA can be considered the cloud fitting proposed in 1996 as output of the ESA contract "A study of cloud detection" (Guzzi *et al.*, 1996, hereinafter CIRS). In this method several shortcomings of ICFA were overcome, especially the adopted RT used to simulate GOME reflection spectra was improved. The principal new features were:

- the cloud optical thickness is a free parameter that can assume all the values typical of thick tropospheric water clouds ( $\tau_c > 5$ ); the effect of the reflecting ground surface underlying the cloud is accounted for.
- Scattering processes are considered: the cloud free part of the spectrum is computed using a full RT like GOMETRAN, accounting for all the relevant scattering and absorption phenomena occurring;
- the cloud top pressure becomes a free parameter and can assume any value between the ground level and the TOA pressure;
- the method used to compute the cloud-top BRDF is substantially improved as far as concerns the accuracy (the computational procedure is deeply described in Torricella and Cervino, 1996).
- The data-base containing O<sub>2</sub> transmittances is obtained at an even higher resolution (the data are each 0.002 cm<sup>-1</sup>) and is available for all the atmospheric levels.

It should be noted that the CIRS fitting is still based on a grid search rather than a continuous non-linear least square technique, since the  $\chi^2$  is evaluated over a space of variables with discrete values. The RTFM adopted is the SIMPMOD version described in Guzzi *et al.*, 1996.

---

### 4.3.5 The BoCFA

The natural development of the CF produced by CIRS is realized by the BoCFA. This method retains the main features concerning the RTFM i.e. still uses a simplified FM instead of a full RTM as for instance GOMETRAN or DISORT but adopt a continuous non-linear least square technique to fit the GOME data to the model function instead of the grid search. The fitting method is the Levenberg-Marquardt described in some details in section 1.3. The parameters to be fitted for are

- the cloud optical thickness  $\tau_c$ ,
- the cloud-top pressure  $p_c$  and
- the cloud coverage or cloud fraction  $f$ .

The structure of the BoCFA is presented in fig. 4.6 and can be summarized as follows:

- the list of days, orbits and pixels to be processed is arranged;
  - for each pixel to be processed the software open the corresponding GOME reflectance file (channel 4) and the corresponding geometry information file containing the  $\theta_0$ ,  $\theta$  and  $\varphi$  values for each pixel. The reflectance file contains for each wavelength in the channel 4  $\lambda_i$  the reflectance values  $R_{\text{MEAS}}(\lambda_i)$  and the associated instrumental uncertainty ( $\sigma_i$ ).
  - The wavelengths inside the fitting window are selected. The fitting window can be varied and is specified in the input sheet. Typically the range 746 – 781 is used as fitting window.
  - The first guess values of the three parameters specified above are read from the input sheet. In the test phase in which pseudo-measured spectra were used instead of GOME data, the method as proved to be very robust to the selection of first guess values very different from the true values. So, for the time being, the first guess values are fixed for all observational conditions.
  - The fitting iterative process is started: the  $\chi^2$  value is computed by means of eq. 1.5. Whenever the  $\chi^2(R_{\text{MEAS}}; \theta_0, \theta, \varphi; \tau_c, p_c, f)$  has to be computed the RTM SIMPMOD is called.
  - (†) A new guess of the parameters is found by solving the set of equations that gives the increments  $\delta$  to be applied to the parameters to find the minimum in the  $\chi^2$  function.
  - A new value of  $\chi^2$  is computed using  $\tau_c + \delta\tau_c$ ,  $p_c + \delta p_c$  and  $f + \delta f$ .
  - The current  $\chi^2(R_{\text{MEAS}}; \theta_0, \theta, \varphi; \tau_c + \delta\tau_c, p_c + \delta p_c, f + \delta f)$  and the previous  $\chi^2$  value are compared:
    - if the  $\chi^2$  function increased or substantially decreased the Levenberg-Marquardt internal parameters are accordingly arranged and the iterative process is restarted from (†);
    - if the  $\chi^2$  stopped decreasing the process is terminated and the parameters values designed as retrieved parameters for the current pixel. The criterion for assessing if the  $\chi^2$  stopped decreasing is described in eq.(1.6). The two thresholds  $\varepsilon_{\text{rel}}$  and  $\varepsilon_{\text{abs}}$  have been both fixed = 0.01.
  - The final call to the Levenberg-Marquardt routine returns also the so called covariance matrix. The diagonal elements of such  $3 \times 3$  matrix are the variances of the retrieved parameters.
-

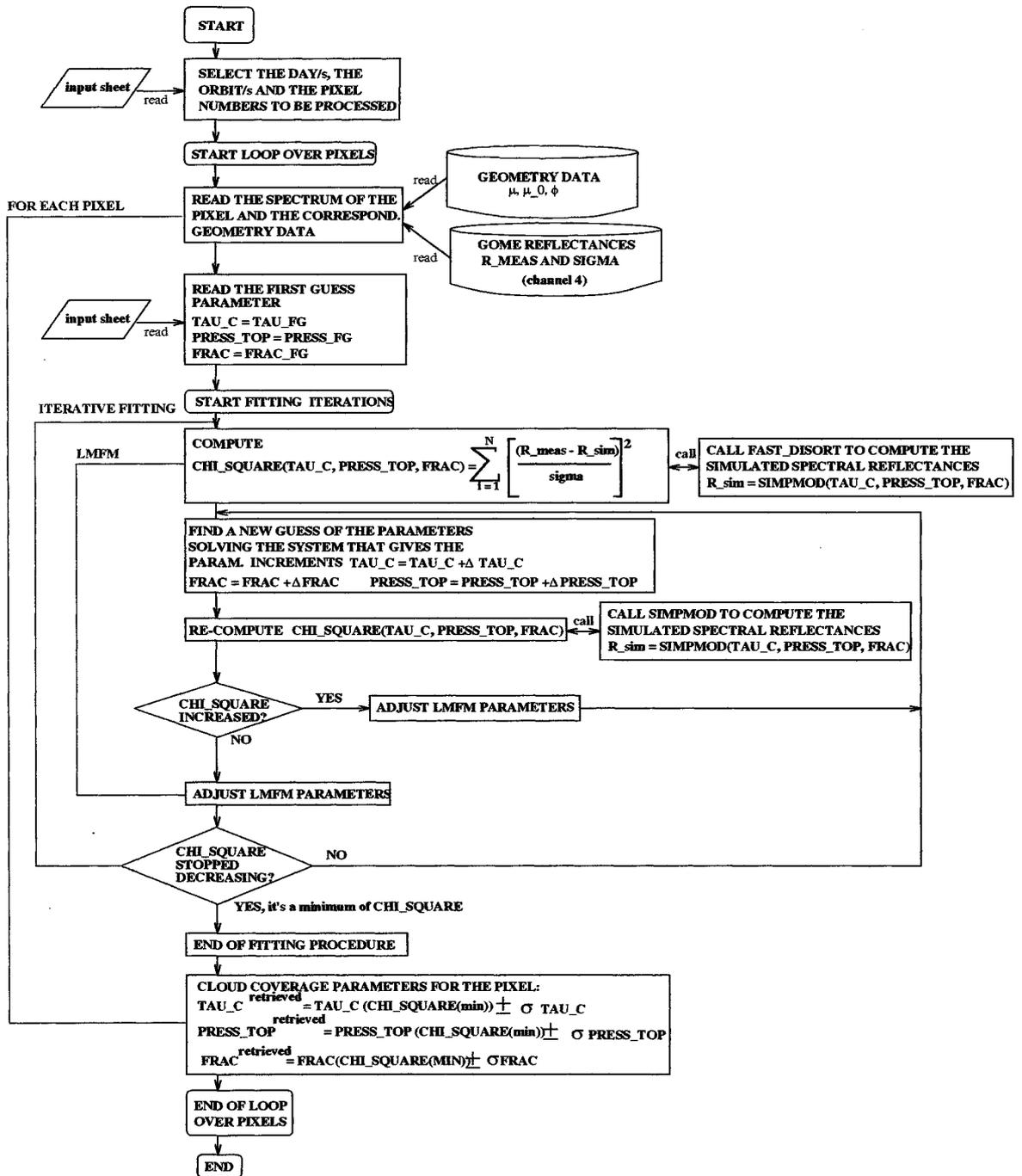


Figure 4.6: BoCFA schematic flow diagram

## 4.4 Bibliography

- M. Derrien, M.: 1992 "Influence of the Size of Field of View on the Contamination by Clouds", Meteo-France
- DLR-DFD: 1996 "Product Specification Document of the GOME Data Processor", Deutsche Forschungsanstalt für Luft- und Raumfahrt – Deutsches Fernerkundungsdatenzentrum, Technical Report No. ER-PS-DLR-GO-0016, Iss./Rev.3/C
- Fouquart, Y. and J.C. Buriez and P. Dubuisson and F. Parol and C. Vanbauce and H. Grassl and M. Kollewe" 1994 "Study on the Effects of Scattering on the Monitoring of Atmospheric Constituents, Volume 2:Scattering by Clouds", ESA-ESTEC AO/1-2542/91/NL/SF, Final report, 1994
- Guzzi, R. and J.P. Burrows and V.V. Rozanov and K. Chance and M. Cervino and T. Kurosu and P. Watts and F. Torricella and K. Muirhead: 1996 "A Study of Cloud Detection", European Space Agency, Final Report, ESA Contract 10997/94/NL/CN, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands
- Hanel, R. A.: 1961, "Determination of cloud altitude from a satellite", *Journal of Geophysical Research*, **66**, 1300
- Kurosu, T.: 1997a "PMD Cloud Detection Algorithm for the GOME Instrument; Algorithm Description and User's Manual", European Space Agency, Draft Final Report, ESA Contract 11572/95/NL/CN, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands
- Kurosu, T.: 1997b "Die Modellierung des Strahlungstransports in Wolken für atmosphärische Fernerkundung im ultravioletten und sichtbaren Spektralbereich", PhD Thesis (in German), Mainz Verlag, Aachen; ISBN 3-89653-185-9
- Kurosu, T. and V.V. Rozanov, and J.P. Burrows: 1997 "Parameterization Schemes for Terrestrial Water Clouds in the Radiative Transfer Model GOMETRAN", *Journal of Geophysical Research* (102), No. D18, 21,809–21,823
- Kuze, A. and K.V. Chance: 1994 "Analysis of Cloud Top Height and Cloud Coverage from Satellites Using the O<sub>2</sub> A and B Bands", *Journal of Geophysical Research* (99), No. D7, 14481–14491
- Levoni, C.: 1997 "Minutes of the Cloud and Aerosol DATA Product Algorithms (CADAPA) Progress Meeting", held in Roma (Nuova Telespazio), Italy, 09 September 1997
- NERC (Natural Environment Research Council): 1997 WWW-Page of the Dundee Satellite Receiving Station, Dundee University, UK, <http://www.sat.dundee.ac.uk/>
- O'Brien, D. M., and R. M. Mitchell: 1992, "Error Estimates for Retrieval of Cloud-Top Pressure Using Absorption in the A Band of Oxygen", *Journal of Applied Meteorology*, **31**, pp. 1179-1192
- Spurr, R.: 1994, "GOME level 1 to 2 algorithms description", Technical Note ER-TN-DLR-GO-0025, Iss./Rev. 1/A, DLR, Oberpfaffenhofen (DE)
- Wu, M. L. C: 1985, "Remote Sensing of the Cloud-Top Pressure Using Reflected solar Radiation in the Oxygen A-Band", *Journal of Climate and Applied Meteorology*, **24**, pp 539-546
-

## List of Abbreviations and Acronyms used in the report

ACFA	Advanced Cloud Fitting Algorithm
AERONET	AERosol Robotic NETwork
AOT	Aerosol Optical Thickness
AT	Asymptotic Theory
BDRF	Bi-Directional Reflection Function
BoCFA	Bologna Cloud Fitting Algorithm
CCD	Charge Coupled Device
CF	Cloud Fitting
CIRS	Cloud Impact on Remote Sensing
COLA	Center for Ocean-Land-Atmosphere Studies
DLR	Deutsche Forschungsanstalt für Luft- und Raumfahrt
DFD	Deutsches Fernerkundungsdatenzentrum
DB	Data Base
DS	Data Set
ERS-2	European Remote Sensing Satellite
ESA	European Space Agency
ESFT	Exponential Sum Fitting
FWC	Fair Weather Cumulus
GDP	GOME Data Processor
GOME	Global Ozone Monitoring Experiment
Grs	GOME reflectance spectrum
ICFA	Initial Cloud Fitting Algorithm
LMFM	Levenberg-Marquardt Fitting Method
LUT	Look Up Table
MEDUSE	MEDiterranean DUSt Experiment
METEOSAT	METEORologic SATellite
NCAR	National Center for Atmospheric Research
NIR	Near Infrared
OP	Optical Properties
OT	Optical Thickness
PGADP	Prototypal GOME Aerosol Data Processor
PMD	Polarization Measurement Device
SCHIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartography
RT	Radiative Transfer
RT(F)M	Radiative Transfer (Forward) Model
SW	Short Wave
SZA	Sun Zenith Angle
TOA	Top Of the Atmosphere
USSA76	United States Standard Atmosphere, 1976
UV	Ultra Violet