

CLOUD RETRIEVAL FOR SCIAMACHY

Diego G. Loyola R.

*German Aerospace Center (DLR)
Postfach 11 16, D-82234 Weßling, Germany
Email: Diego.Loyola@dlr.de*

INTRODUCTION

The ENVISAT Satellite Payload includes the SCIAMACHY instrument designed to measure a number of chemically important atmospheric trace constituents on a global basis. Accurate determination of cloud information out of SCIAMACHY nadir measurements is needed for the retrieval of ozone and other trace gases and it is crucial for more demanding applications like aerosol retrieval and the retrieval of trace gas concentrations in the troposphere.

Determination of cloud coverage is done traditionally using threshold tests to distinguish between cloud-free and cloud-contaminated pixels. The probably most notable work using this approach is the one presented in [6], where different test are applied to examine variations in brightness temperatures or albedos measured by NOAA's Advanced Very High Resolution Radiometer (AVHRR).

The major problem with the threshold techniques is that the threshold settings, even dynamical thresholding or threshold functions, and the generation of global threshold maps are tedious and complicated tasks that are usually done with theoretical models or are performed empirically with the help of human experts.

The Optical Cloud Recognition Algorithm (OCRA) [5] is a novel algorithm for cloud recognition which uses the information provided by the GOME-2 and SCIAMACHY polarization measurement devices. OCRA overcomes the problems of the traditional threshold testing and it is based on the fusion of multitemporal images to determine a so called "cloud-free composite".

The paper is divided in two parts. The first section describes the cloud recognition algorithm OCRA, the second part presents the application of OCRA to the GOME/ERS-2 and SCIAMACHY/ENVISAT sensors.

CLOUD RECOGNITION ALGORITHM

The basic idea of OCRA is to decompose the image irradiance measured by an optical sensor into two components: the background and the clouds. Fig. 1 illustrates these decomposition. The reflectance ρ is defined as a linear combination of the surface reflectance ρ_{CF} of the background (cloud-free) and the reflectance of the clouds ρ_C

$$\rho = \rho_{CF} + \rho_C \quad (1)$$

The recognition of clouds from an optical image can be done using (1) and a cloud-free background. The background could be computed using models and optical properties of the sensor of interest, but a more natural and precise approach is to use historical multitemporal data measured by the same sensor.

Multitemporal data fusion for the generation of cloud-free composites is done in two steps as shown in Fig. 2. First a pre-processing is needed where radiometric and geometric correction are carried out. The fusion is then done by selecting the less cloudy pixels among the multitemporal data obtaining as result a global cloud-free composite.

The next sub-sections describe in more detail how the pre-processing and the multitemporal data fusion is done.

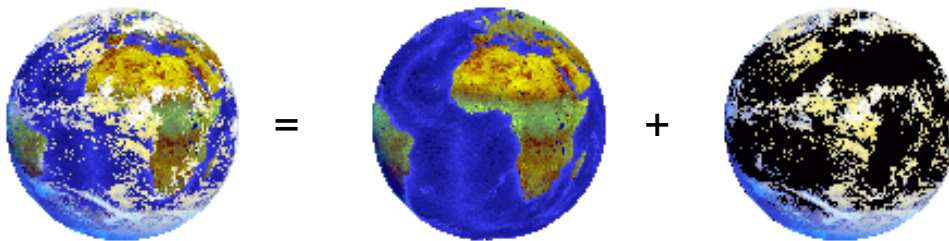


Fig. 1: Decomposition of image irradiance in background and clouds.

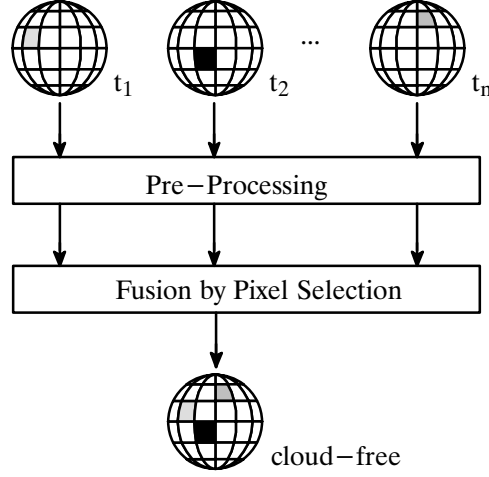


Fig. 2: Processing steps for the generation of the cloud-free composite fusing multitemporal data t_1 to t_n .

Pre-processing

The core of the pre-processing algorithm is based on the idealized physical model of image irradiance for remote sensing developed by [7]. The image irradiance $E(x, y)$ with explicitly dependence on wavelength λ is defined as

$$E(x, y, \lambda) = T_u(\lambda) \frac{\rho(x, y, \lambda)}{\pi} \times \left[E_0(\lambda) T_d(\lambda) \cos(i) + E_s(\lambda) \left(\frac{1 + \cos(e)}{2} \right) \right] + L_p(\lambda) \quad (2)$$

where $T_u(\lambda)$ is the upward transmission through the atmosphere, $\rho(x, y, \lambda)$ is the surface reflectance of the ground cover projecting to the image location (x, y) , $E_0(\lambda)$ is the solar irradiance at the top of the atmosphere, $T_d(\lambda)$ is the downward transmission through the atmosphere, i is the solar angle of incidence, $E_s(\lambda)$ is the integrated radiance over the hemisphere of the sky, e is the surface slope, and $L_p(\lambda)$ is the radiant energy that reaches the sensor due to backscatter from the direct solar beam (path radiance). Ground cover is assumed as Lambertian surface.

The reflectance factor $\rho(x, y, \lambda)$ is by definition invariant to conditions of illumination and viewing for Lambertian surfaces. The surface reflectance at location (x, y) can be computed from two measurements at different times

$$\rho(x, y, \lambda) = \frac{2\pi}{T_u(\lambda)T_d(\lambda)} \times \frac{(E_a(x, y, \lambda) - E_b(x, y, \lambda))}{E_0(\lambda)(\cos(i_a) - \cos(i_b)) E_s(\lambda)(\cos(e_a) - \cos(e_b))} \quad (3)$$

where $T_u(\lambda)$, $E_0(\lambda)$, $T_d(\lambda)$ and $E_s(\lambda)$ are assumed to remain constant at times t_a and t_b .

Equation (3) can be used to directly compare effects of the atmosphere in different data sets of the same location. The comparison would show changes in the atmosphere mainly influenced by cloud contamination. The most important variables, as expected from [7], are the cosine of the solar incident angle and cosine of the slope. A simplified model for the reflectance factor can be rewritten in

$$\rho(x, y, \lambda) = \frac{E(x, y, \lambda)}{E_0(\lambda) \cos(i) \cos(e)} \quad (4)$$

where no iteration between the wavelength and the geometric dependence of reflection is assumed. This very simple model for comparison of reflections conforms to reality with considerable accuracy for applications like the one described in this work. The model has also the additional advantage that a correction for instrument degradation and earth to sun distance is not needed when the solar irradiance $E_0(\lambda)$ can be directly measured by the same sensor.

Cloud-free Composite Generation

Red-Green-Blue (RGB) images can be obtained using the corrected reflectance from (4) in three visible bands: λ_R , λ_G , λ_B . A so called "cloud-free composite" is created by selecting the pixels with minimum brightness reflectance (less cloudy) out of the multitemporal data set.

The normalized RGB (rgb) representation of color has been used extensively in computer vision to separate color properties of the surface from the brightness component [3]. The rgb transformation is obtained normalizing the RGB colors by its intensity

$$r = \frac{R}{R + G + B}, g = \frac{G}{R + G + B}, b = \frac{B}{R + G + B} \quad (5)$$

It is sufficient to use the two components r and g , the third one is given by $b=1-r-g$. The color characteristics can be then represented in a rg -space. All colors that can be represented by the three primary colors R, G and B are confined within a triangle in the rg -space as shown in Fig. 3.

The reflectance defined in (4) can be translated to the rg -space using (5)

$$r = \frac{\rho(x, y, \lambda_R)}{\sum_{i=R,G,B} \rho(x, y, \lambda_i)}, \text{ and } g = \frac{\rho(x, y, \lambda_G)}{\sum_{i=R,G,B} \rho(x, y, \lambda_i)} \quad (6)$$

An interesting issue in the above definition of reflectance ratios is that the geometry component is eliminated from (4) since the factor $\cos(i) \cos(e)$ is the same for each of the RGB functions.

Let $M = \{rg_1, rg_2, \dots, rg_n\}$ be a set of n normalized multitemporal measurements over the same location (x, y) , then a cloud-free (less cloudy) pixel $rg_{CF} \in M$ is selected using a brightness comparison such that

$$\|rg_{CF} - w\| \geq \|rg_i - w\| \text{ for } i=1, \dots, n \quad (7)$$

where w is the *white-point* in the rg chromaticity diagram given by $w = [1/3, 1/3]$.

A global cloud-free composite is obtained merging the reflectance Q_{CF} corresponding to the cloud-free pixels rg_{CF} .

Cloud Recognition

The fractional cloud cover f , defined in the range $[0, 1]$, is determined as the distance between the measured reflectance Q and the cloud-free reflectance Q_{CF}

$$f = \sqrt{\sum_{\lambda_i} \alpha(\lambda_i) \times \max(0, [\rho(\lambda_i) - \rho_{CF}(\lambda_i)]^2 - \beta(\lambda_i))} \quad (8)$$

where $Q(\lambda_i)$ and $Q_{CF}(\lambda_i)$ are the reflectance at wavelength λ_i , $\alpha(\lambda_i)$ and $\beta(\lambda_i)$ are scaling and offset factors proportional to the position of the w point in the rg -space.

The scaling and offset factors can be determined using histogram analysis of selected scenarios, or can also be fitted from (8) using known values of the cloud fraction f .

APPLICATION TO GOME AND SCIAMACHY DATA

The OCRA algorithm presented in the previous section can be applied to the optical information provided by GOME and SCIAMACHY polarization measurement devices (PMDs). Table 1 shows the spatial and spectral resolution of the GOME and SCIAMACHY polarization measurement devices.

Cloud-free Composites

Single PMD data are converted to reflectances using (4) and the PMD solar irradiance measured by the same sensor. The R, G, and B components used in (5) correspond to the GOME PMD₃, PMD₂, and PMD₁, and the SCIAMACHY PMD₄+PMD₃, PMD₂, and PMD₁.

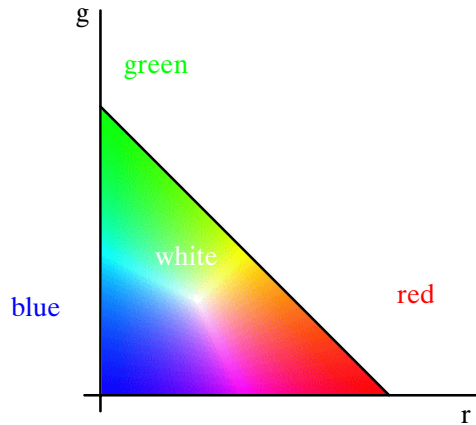


Fig. 3: Chromaticity diagram in the rg -color space.

Table 1. Spectral bands of the GOME and SCIAMACHY polarization measurement devices.

PMD Band	GOME	SCIAMACHY
1	300 to 400 nm	320 to 380 nm
2	400 to 600 nm	450 to 520 nm
3	600 to 800 nm	620 to 700 nm
4		800 to 900 nm
5		1500 to 1700 nm
6		2265 to 2380 nm

The global cloud-free composites for GOME are generated applying (7) to PMD reflectance data of one month. Almost 15 million single PMD data sets are fused in each monthly cloud-free composite (16x2200 PMD pixels per orbit x 14.3 orbits per day). Fig. 4 show the cloud-free composites for January 1997. The black regions indicate no data available: the poles are dark in winter and over the Himalayas there is a systematic dump problem of the ERS-2 tape recorder. It can be seen that some clouds remain mainly in the equatorial region. This is due to the fact that GOME provides global coverage at the equator within 3 days, i.e. only 10 different observations during a month are available for that region.

Finally the PMD monthly composites are merged to a yearly composite and in this way the temporal changing ice/snow regions are removed. Fig. 5 shows one year cloud-free composite using GOME data from July 1996, October 1996, January 1997 and April 1997. Ice/snow remains only in the Arctic, Antarctic, Alaska, Andean and Himalaya regions.

The GOME cloud-free composite will be initially used in the SCIAMACHY processors until enough historical SCIAMACHY data is available.

Cloud Recognition Results for GOME

The fractional cloud cover is determined with (8) using PMD measurement and the PMD cloud-free composite. The scaling and offset factors are determined using histogram analysis of selected GOME orbits.

Results show a good correlation between cloud coverage derived by OCRA, APOLLO and the Initial Cloud Fitting Algorithm (ICFA). The APOLLO algorithm [2] was applied to the ATSR-2/ERS-2 data. The ICFA algorithm, which is part of the GOME Data Processor system [4], computes fractional cloud cover using the relationship between the transmittance through the O₂ A-band (around 760 nm) and the cloud-top height given by the climatology. Correlation between OCRA, APOLLO and ICFA using ERS-2 data acquired on September 15th 1996 is shown in Fig. 6.

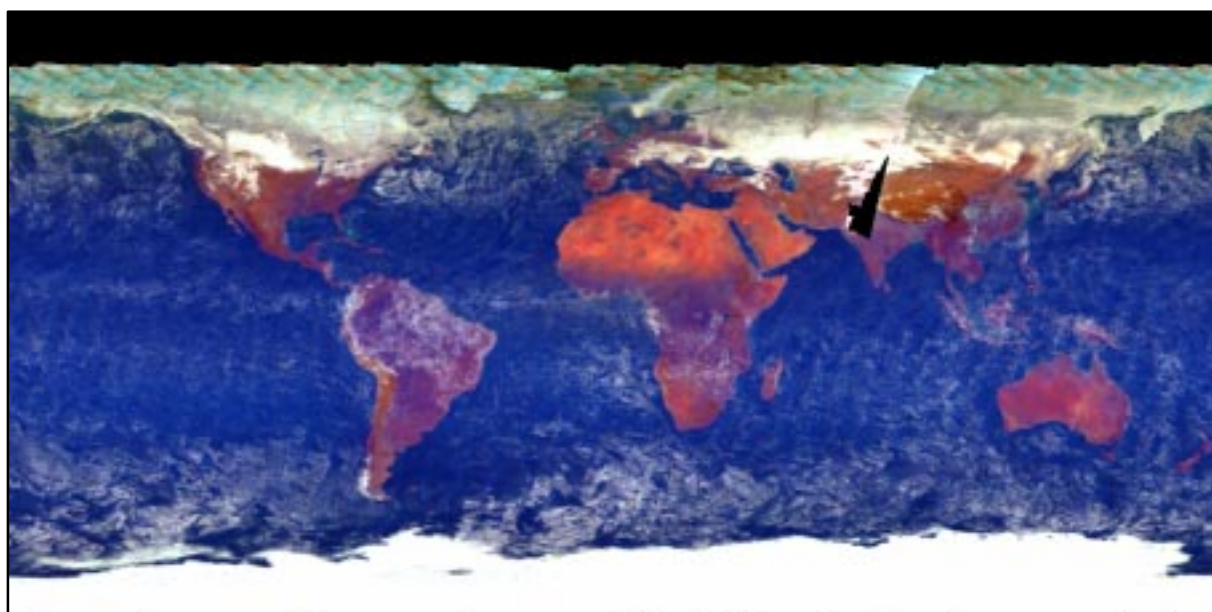


Fig. 4: GOME PMD cloud-free composite for January 1997.

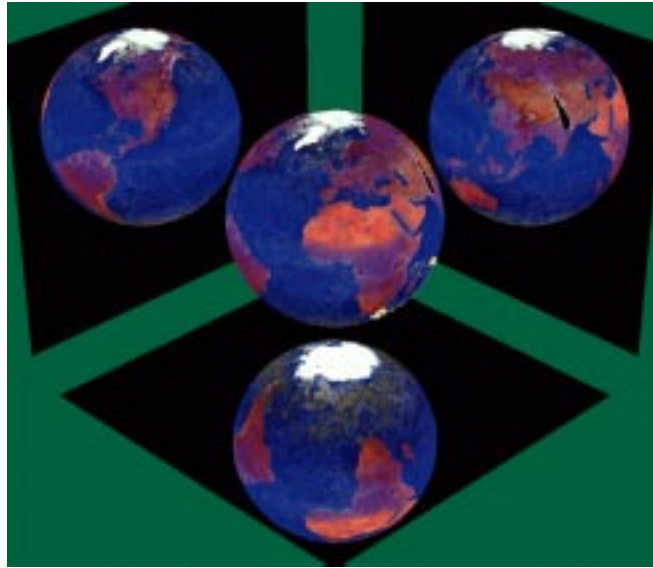


Fig. 5: Earthball projection of the GOME PMD cloud-free composite using data from July 1996 to April 1997.

A qualitative comparison of the cloud cover computed for the two GOME orbits from July 19th 1997 and a METEOSAT image for the same time period is shown in Fig. 7. Although there is more than one hour time difference from the begin of the GOME orbits to the METEOSAT data acquisition time, the cloud cover computed with OCRA matches like a puzzle in the METEOSAT image. A very good match between different type of clouds can be seen, not only for the clouds of a low pressure system covering the European continent and the convective clouds in the inner-tropical zone below the Sahara desert, but also for the much finer field of scattered cumuli clouds in the South Atlantic region.

Cloud coverage for a complete orbit can be computed in a few seconds, while the more computer demanding task of cloud-free composition determination is done off-line.

CONCLUSIONS

The cloud retrieval algorithm OCRA to be used in the operational near-real-time and off-line SCIAMACHY processors has been presented.

The key of the OCRA algorithm is the determination of a cloud-free composite using historical multitemporal data fusion techniques. A sophisticate pre-preprocessing step was developed based on a physical model of image formation invariant to topography, position of the sun, atmosphere, and position of the viewer.

The OCRA algorithm is robust and very fast, and the results using GOME data show a very good agreement in comparison with results of the ICFA and APOLLO algorithm, and with METEOSAT data.

Future work will concentrate on the distinction between clouds and snow/ice surfaces.

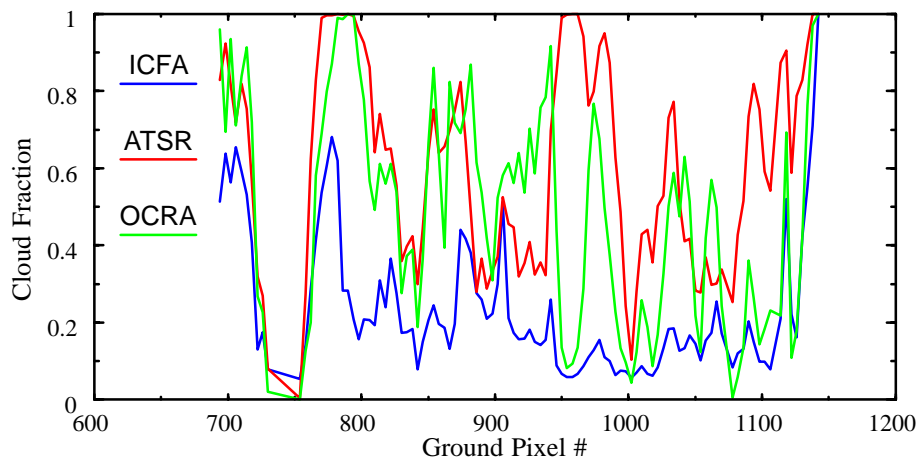


Fig. 6: Correlation of cloud cover derived by OCRA, ICFA (GOME data) and APOLLO (ATSR-2 data) algorithms.

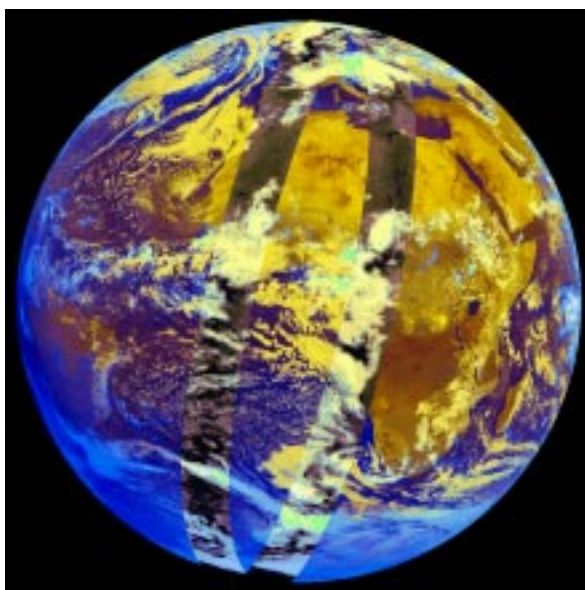


Fig. 7: OCRA cloud cover compared to METEOSAT data acquired on July 19th 1997, 11:00 UTC.

REFERENCES

- [1] M. Ehlers, "Multisensor image fusion techniques in remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 19, no. 5, pp. 823–854, 1991.
- [2] G. Gesell, "An algorithm for snow and ice detection using AVHRR data. An extension to the APOLLO software package," *International Journal of Remote Sensing*, vol. 10, no. 4–5, pp. 897–905, 1989.
- [3] P. Golland and A. M. Bruckstein, "Motion from Color," *Computer vision and Image Understanding*, vol. 68, no. 3, December 1997.
- [4] D. Loyola, W. Balzer, B. Aberle, M. Bittner, K. Kretschel, E. Mikusch, H. Muehle, T. Ruppert, C. Schmid, S. Slijkhuis, R. Spurr, W. Thomas, T. Wieland, M. Wolfmueller, "Ground Segment for ERS-2/GOME Sensor at the German D-PAF," *3rd ERS Symposium on Space at the service of our Environment*, Florence, Italy, vol. II, pp. 591–596, 1997.
- [5] D. Loyola, "A New Cloud Recognition Algorithm for Optical Sensors," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS'98 DIGEST VOLUME II*, pp. 572–574, Seattle 1998.
- [6] R.W. Saunders and K.T. Kriebel, "An improved method for detecting clear sky and cloudy radiances from AVHRR data," *International Journal of Remote Sensing*, vol. 9, pp. 123–150, 1988.
- [7] R. J. Woodham and M. H. Gray, "An Analytic Method for Radiometric Correction of Satellite Multispectral Scanner Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-25, no. 3, May 1987.

ACKNOWLEDGEMENTS

The author would like to thank the DLR colleagues T. Ruppert for his support during the generation of the GOME PMD cloud-free composites, A. von Bargaen and B. Aberle for the ATSR-2 data processing. The work was partially fund by BMBF and ESA.