Information Content in the Oxygen A-Band for the Retrieval of Macrophysical Cloud Parameters

Olena Schuessler, Diego Guillermo Loyola Rodriguez, *Senior Member, IEEE*, Adrian Doicu, and Robert Spurr

***Abstract*—Current and future satellite sensors provide measurements in and around the oxygen A-band on a global basis. These data are commonly used for the determination of cloud and aerosol properties. In this paper, we assess the information content in the oxygen A-band for the retrieval of macrophysical cloud parameters using precise radiative transfer simulations covering a wide range of geophysical conditions in conjunction with advance inversion techniques. The information content of the signal with respect to the retrieved parameters is analyzed in a stochastic framework using two common criteria: the degrees of freedom for a signal and the Shannon information content. It is found that oxygen A-band measurements with moderate spectral resolution (0.2 nm) provide two pieces of independent information that allow the accurate retrieval of cloud-top height together with either cloud optical thickness or cloud fraction. Additionally, our results confirm previous studies indicating that the retrieval of cloud geometrical thickness (CGT) from single-angle measurements is not reliable in this spectral region. Finally, a sensitivity study shows that the retrieval of macrophysical cloud parameters is slightly sensitive to the uncertainty in the CGT and very sensitive to the uncertainty in the surface albedo.**

***Index Terms*—Information content of hyperspectral measurements, oxygen A-band, retrieval of macrophysical cloud parameters.**

I. Introduction

CLOUDS are an important component of the global hydrological cycle and play a major role in the Earth's climate system through their strong impact on radiation processes [1].

The interplay of sunlight with clouds imposes major challenges for satellite remote sensing, both in terms of the spatial complexity of real clouds and the dominance of multiple scattering in radiation transport. The retrieval of trace gas products from UV/VIS spectrometers is strongly affected by the presence of clouds. The physics behind the influence of cloud on trace gas retrieval is well understood, and in general, there are three different contributions: 1) the albedo effect associated with the enhancement of reflectivity for cloudy scenes compared to cloud-free sky scenes; 2) the so-called shielding effect, for which that part of the trace gas column below the cloud is hidden by the clouds themselves; and 3) the increase in absorption, related to multiple scattering inside clouds which leads to enhancements of the optical path length. The albedo and in-cloud absorption effects increase the visibility of trace gases at and above the cloud-top, while the shielding effect normally results in an underestimation of the trace gas column. Several papers have quantified theoretically using radiative transfer modeling the influence of cloud parameters on the retrieval of trace gas columns (e.g., see [2]–[5]). These studies show that the cloud fraction (CF), cloud optical thickness (COT; albedo), and cloud-top pressure (height) are the most important quantities for cloud correction of satellite trace gas retrievals.

There have been many other attempts to retrieve cloud properties from a range of passive remote sensing instruments. Infrared measurements from instruments such as ATSR-2 (emission) and MODIS (backscatter and emission) yield brightness-derived cloud-top pressures at high spatial resolution. The OMI instrument uses UV and visible backscatter measurements to derive CF and representative cloud pressures. Other instruments providing cloud properties include AIRS (thermal emission), MISR (visible), and POLDER (A-band). In recent years, vertical profiles of cloud properties have become available from the active sensors CloudSat and CALIPSO on the NASA A-Train. There have been many intercomparisons and synergistic studies for various cloud products from active and passive spaceborne sensors (e.g., see [6] and [7]).

Cloud-property retrieval has been performed using a wide variety of inverse models—many of these are based on cost function minimization methods, with and without regularization. However, there has been limited work done on information content of such retrievals. Reference [8] made a detailed sensitivity and information-content study of visible and infrared measurements pertinent to cloud microphysical property retrievals from MODIS-type instruments; this work was aimed at optimizing spectral channel selection. In our paper, we present an information-theoretic study for the retrieval of cloud properties from moderate-resolution near-infrared spectrometers measuring in and adjacent to the O_2A band.

A continuous data record of oxygen A-band measurements is provided by European atmospheric composition sensors starting with GOME [9] launched in April 1995 onboard the ERS-2 satellite, SCIAMACHY [10] launched in March 2002 onboard the Envisat satellite, and GOME-2 [11] launched in October 2006 onboard MetOp-A platform. Two more GOME-2 sensors on the MetOp-B (launched in September 2012) and MetOp-C (to be launched in 2017) satellites will provide measurements until 2020. The series of atmospheric missions Sentinel-4/Sentinel-5 and Sentinel-5 Precursor (S5P) [12] will further extend these measurements beyond the next decade.

Manuscript received December 21, 2012; revised April 22, 2013; accepted May 20, 2013.

O. Schuessler, D. G. Loyola Rodriguez, and A. Doicu are with the Institut für Methodik der Fernerkundung (IMF), Deutsches Zentrum für Luft- und Raumfahrt (DLR), 82234 Weßling, Germany.

R. Spurr is with RT Solutions, Inc., Cambridge, MA 02138 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2013.2271986

The cloud parameters needed for the trace gas retrieval from GOME-type sensors are mainly obtained from the oxygen A-band measurements (see [13] and the references therein). The long-term data record of cloud properties derived from GOME compares well with other satellite data [14] as well as with the global and seasonal patterns from the multisatellite international satellite cloud climatology project. The use of the oxygen A-band generates complementary cloud information (especially for low clouds), as compared to traditional thermal infrared sensors (as used in most meteorological satellites) that are less sensitive to low clouds due to reduced thermal contrast.

In this paper, clouds are modeled as optically uniform layers of scattering water droplets; this is the clouds-as-layers (CAL) approach. This is a major point of departure from a treatment of clouds as Lambertian reflectors; the latter is the cloud-reflecting-boundaries (CRB) ansatz. CRB is commonly used in the operational retrieval of trace gases, and it is sometimes known as the mixed Lambert-equivalent reflectivity (MLER) method [15], [16]. The CRB assumption has been the default for GOME, SCIAMACHY, and GOME-2 operational processing [5], [17], [18].

The remainder of this paper is outlined as follows. In Section II, we first describe the CAL cloud model and discuss the limitations of the CRB approach; this is followed by a summary of the radiative transfer (RT) modeling. In Section III, the inversion scheme is described, including the determination of information content, and in Section IV, we present simulations covering a wide range of geophysical conditions and discuss the corresponding results. The conclusion is given in Section V.

II. Forward Modeling

The forward model in both CAL and CRB approaches involves multiple-scattering radiative transfer simulations of satellite intensities in and around the O_2A-band, but the CAL treatment is clearly more physically realistic, requiring a full consideration of scattering inside the cloud itself. Before we describe the CAL forward model setup in this paper, we will summarize some of the shortcomings of the CRB approach.

It is clear that the CRB approximation will be closest to reality for single-layer optically thick geometrically thin clouds. However, this type of cloud is the exception; the vast majority of clouds in the Earth atmosphere are not optically thick. For example, the area-mean COT over the ocean outside the polar regions is in the range of 3–5 [19]. Even then, the reflection from a semiinfinite cloud will not necessarily be close to the albedo from cloud-top. In the context of the OMI cloud pressure retrieval algorithm [16], the MLER model has been shown to produce significant errors in estimated cloud pressures for cloud of moderate or low optical thickness [15].

It is intrinsic to the CRB approach that it masks intracloud and below-cloud ozone in trace gas retrievals; in DOAS-style algorithms, it is necessary to use a climatology-derived ozone ghost column in air mass factor calculations [5], [18]. Satellite measurements are sensitive to intracloud ozone, and neglecting this effect in CRB-based retrieval may give rise to significant errors in the simulated backscatter signal, thereby inducing appreciable ozone total column errors [17]. To overcome this problem, a semiempirical method has been developed for deriving intracloud ozone and correcting the overestimation of the ozone ghost column induced by CRB [18].

It has also been observed that, in validation studies of satellite-derived total ozone compared with reference ground-based Brewer data, there is a notable underestimation of total ozone under (CRB-based) cloudy conditions, an effect that clearly decreases with increasing solar zenith angle [20]. This is likely to be partially remedied by a more physical approach to cloud modeling.

We describe now the CAL forward modeling used in this paper. The total (sun-normalized) radiance as seen at the satellite is taken to be a linear combination of independent radiances corresponding to the completely clear-sky atmosphere (I_{clear}) and the completely cloudy-sky atmosphere (I_{cloudy}), with the weighting expressed through the radiative CF. This is the independent pixel approximation (IPA).

The clear-sky radiance depends on surface properties (the albedo s_a and surface height s_z) and on the viewing geometry (solar zenith angle, viewing zenith angle, and relative azimuth angle). The cloudy-sky radiance is additionally dependent on cloud parameters, which are the cloud-top height (CTH) c_z, cloud-bottom height c_b (or alternatively cloud geometrical thickness (CGT) $c_{gt} = c_z - c_b$), and COT τ_c. In the cloudy-sky simulation, the cloud is treated as a set of contiguous scattering layers; the entire cloud is assumed to be optically uniform, with scattering properties determined through Mie calculations for water droplet particles (Mie properties are discussed in Section IV-A). In the IPA, we may write

$$I_{\mathrm{sim}}^{\mathrm{CAL}} = c_f \cdot \langle I_{\mathrm{cloudy}}(\lambda, \Theta, \tau_c, c_z, c_{gt}, s_a, s_z)\rangle + (1 - c_f) \cdot \langle I_{\mathrm{clear}}(\lambda, \Theta, s_a, s_z)\rangle \,. \quad (1)$$

Here, c_f is the CF, Θ denotes path geometry (solar and line-of-sight angles), and λ is the wavelength. Radiances are typically calculated at high spectral resolution before convolution with the sensor slit function (denoted in (1) by the operation $\langle * \rangle$).

Radiances are calculated using the VLIDORT RT code [21], [22] in and adjacent to the O_2A-band. For full accuracy, it is necessary to calculate line-by-line (LBL) reflectances (typically at resolution 0.0025 wavenumber) using line-spectroscopic information for the O_2A-band (taken from the HITRAN database [23]) before convolution. For the range 758–771 nm which is commonly used for cloud-property retrieval in GOME-type sensors, there are more than 10 000 LBL calculations at high resolution.

In this paper, the clear-sky radiance is calculated using a 21-layer molecular-scattering atmosphere, with pressure and temperature profiles interpolated from the standard profile data set. Rayleigh scattering follows the specifications in [24]. O_2 absorption cross sections were computed using LBL calculations made with Generic Atmospheric Radiation LBL Infrared Code [25] using optimized rational approximations for the Voigt line profile [26]. For the cloudy-sky simulations, the placement of a cloud medium into this clear-sky atmosphere may introduce additional layers (up to two of them) in the

stratification. If the cloud-top level z_{upper} is between layer heights z_{n-1} and z_n, then layer n must be partitioned into two sublayers $[z_{n-1}, z_{\text{upper}}]$ (clear) and $[z_{\text{upper}}, z_n]$ (cloud filled), similarly for cloud-bottom level z_{lower}. If z_{upper} or z_{lower} (or both) coincides with one of the original stratification levels, no partitioning is required; conversely, if $z_{n-1} > z_{\text{upper}} > z_{\text{lower}} > z_n$, then layer n must be partitioned into three sublayers (two clear and the middle one cloud filled). All cloud sublayers are contiguous, and sublayer gas absorption and Rayleigh optical thickness values are adjusted accordingly to preserve the original values found for the clear-sky scenario.

VLIDORT is a multiple-scatter discrete-ordinate vector RT model that includes the effects of polarization and uses a precise calculation of the single-scattering radiation field in a spherically curved atmosphere. In general, the degree of linear polarization in the O_2A-band differs from that in the adjacent continuum, especially in the presence of polarizing aerosols [27]. Here, we do not consider aerosols; scattering by water-droplet clouds is depolarizing, and the Rayleigh polarization effect is not so high as in the UV. For these reasons, we have neglected polarization in this paper, although we will address polarization effects in a sequel paper. In the cloudy-sky simulations, effects from strong forward scattering by cloud particulates is dealt with using the delta-M approximation [28]; we have found that 16 discrete ordinates give sufficient accuracy for Fourier-azimuth convergence of the multiple-scattering diffuse field.

The forward model as outlined previously is run repeatedly to generate a complete data set of synthetic radiance templates for an appropriate range of viewing/solar geometries and surface geophysical scenarios, and for various combinations of cloud properties (this data set is summarized in Section IV-A).

III. Inverse Modeling

Denoting by $\mathbf{x}$ the n-dimensional state vector comprising the unknown cloud parameters, by $\mathbf{b}$ the known vector of auxiliary parameters, encapsulating the surface properties and the viewing geometry, and by $\mathbf{y}^\delta$ the m-dimensional measurement vector, we rewrite (1) in compact form as

$$\mathbf{y}^\delta = \mathbf{F}(\mathbf{x}, b) + \delta \tag{2}$$

where $\boldsymbol{F}$ is the forward model and δ is the error vector in the data.

Equation (2) is nonlinear and ill-posed, and some sort of regularization is required in order to obtain a solution $\mathbf{x}$ with physical meaning. Note that the ill-posedness of a discrete equation is reflected by a huge condition number $c(\boldsymbol{K})$ of the Jacobian matrix $K = d\mathbf{F}/d\mathbf{x}$, where $c(\mathbf{K}) = \gamma_{\max}/\gamma_{\min}$, with $\gamma_{\max}$ and $\gamma_{\min}$ as the largest and smallest singular values of $\mathbf{K}$, respectively.

The most widely used technique for regularizing discrete ill-posed problems is the method of Tikhonov regularization [29]. The key idea of this method is to include *a priori* assumptions about the smoothness and/or the size of the desired solution. Essentially, we define the regularized solution $\mathbf{x}_\alpha^\delta$ as the minimizer of the objective function, which is the weighted combination of the residual norm and the constraint norm

$$F_\alpha(\mathbf{x}, b) = \frac{1}{2}\left[\left\|\mathbf{F}(\mathbf{x}, b) - y^\delta\right\|^2 + \alpha\left\|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\right\|^2\right] \tag{3}$$

where α denotes the regularization parameter and $\mathbf{L}$ is the regularization matrix. The regularization parameter balances the information coming from the measurement $\|\mathbf{F}(\mathbf{x}, b) - y^\delta\|$ and the *a priori* information $\|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|$. The regularized solution can be computed with optimization methods for unconstrained minimization problems [30]. The minimizer of the Tikhonov function (3) can be computed by using the Gauss–Newton method, in which case the new iterate for $\mathbf{L} = \mathbf{I}$ is given by

$$\mathbf{x}_{i+1}^\delta = \arg\min_x \left\|\mathbf{K}_i\left(\mathbf{x} - \mathbf{x}_i^\delta\right) - \left[\mathbf{y}^\delta - \mathbf{F}\left(\mathbf{x}_i^\delta, b\right)\right]\right\|^2 + \alpha\left\|\mathbf{x}_i^\delta - \mathbf{x}_a\right\|^2. \tag{4}$$

A. Information Content

Statistical inversion methods are widely used in the field of atmospheric sounding for trace gas retrievals. Commonly, in the framework of statistical inversion, two criteria to analyze the information content of a signal are used: the degree of freedom for a signal [31] and the Shannon information content (SIC) [32]. In the statistical inversion theory, all parameters are treated as random variables, and it is assumed that they are continuous and that their probability distributions can be expressed in terms of probability densities and the *a posteriori* density represents the solution of the inverse problem. As shown in [30], the optimal estimation method [31], based on Bayes theorem [33], can be regarded as a stochastic version of Tikhonov regularization. More precisely, the maximum *a posteriori* solution coincides with the Tikhonov solution under the assumption that the state vector $\mathbf{x}$ and the noise vector $\boldsymbol{\delta}$ are Gaussian random vectors with covariance matrices, given by $C = \sigma_x^2\mathbf{I}_n$ and $C_\delta = \sigma^2\mathbf{I}_m$, respectively, where σ_x and σ are the standard deviations of the state vector and the noise vector, respectively. In this regard, the regularization parameter α is the ratio of the noise variance to the profile variance in our *a priori* knowledge, i.e., $\alpha = \sigma^2/\sigma_x^2$.

The equivalence between the Bayesian approach and the method of Tikhonov regularization enables us to analyze the information content of the signal with respect to the retrieved parameters in a stochastic framework. The degree of freedom for a signal (DFS) is a measure of independent pieces of information in the measurement and gives the minimum number of parameters, which can be used to define a state vector without loss of information. It is defined as the trace of the averaging kernel matrix, which represents the sensitivity of the retrieval to the changes in the *true* state. The DFS can be computed as

$$\text{DFS} = \sum_i^n \frac{\gamma_i^2}{\gamma_i^2 + \alpha} \tag{5}$$

where γ_i^2 represents the singular values of the matrix $\mathbf{K}$. The degree of freedom for the signal indicates how much information

TABLE I
O_2 A-BAND SIMULATION PARAMETERS AND RANGES

Parameter	Range of Values
Surface height	0, 1, 2 km
Surface albedo	0.07 to 1.0 (4 values)
Cloud-top height	0.5 to 5.0 km (intervals 0.5 km)
Cloud geometrical thickness	0.5 to 3.5 km (intervals 0.5 km)
Cloud optical thickness	0, 1.3, 3.6, 9, 10, 23, 125
Solar zenith angle	15, 30, 45, 55, 65, 70, 75, 80, 85, 87, 88 deg
Viewing zenith angle	0 to 70 deg (every 5 deg)
Relative azimuth angle	0, 45, 90, 135, 180 deg

is inferred from the measurement and how much from our *a priori* knowledge about the solution. For example, if $\mathrm{DFS} = 0$, all information comes from the *a priori*, while for $\mathrm{DFS} = n$, all information comes from the measurement. Thus, the DFS helps us to assess the overall performance of the retrieval algorithm for a variety of retrieval scenarios.

Another useful criterion for the estimation of the retrieval quality is the SIC, which is a measure of the incremental gain in information, defined as the entropy difference between the *a priori* and *a posteriori* states; the corresponding formula reads as

$$\mathrm{SIC} = \frac{1}{2}\sum_{i}^{n} \log\left(1 + \frac{\gamma_i^2}{\alpha}\right). \tag{6}$$

IV. SIMULATIONS AND RESULTS

We have set up a comprehensive data set of simulated reflectance for various ranges of cloud and surface parameters, and for a selection of solar and viewing geometries covering the full viewing geometry extent of current and future atmospheric composition missions. The template classification is summarized in Table I.

We have confined our attention to low-lying cumulus and stratiform clouds over land (class C1 according to the classification in [34]). Ten CTHs are specified from 0.5 to 5.0 km at 0.5-km intervals, with seven CGT values also set at 0.5-km intervals. We used seven values of COT from 0 (clear sky) to 125, three surface heights, and four surface albedos (0.07, 0.20, 0.45, and 1.0). For the geometrical angles, there are 11 solar zenith angles (15–88°), 11 viewing zenith angles (0–70°), and 5 azimuth angles. These geometrical angle ranges are representative for modern nadir-viewing satellite remote sensing instruments; in particular, the wide-swath viewing for GOME-2 and S5P is covered by the range of viewing zenith angles.

The optical properties for this C1 cloud type are determined through Mie calculations for purely scattering spherical water droplets with refractive index (1.33, 10^{-8}). Droplets are assumed to be polydispersed according to a modified-Gamma size distribution function parameterized by the mode radius r_m in millimeter and constants α and γ describing the shape of the distribution (the scheme follows that in [35]). The values of these parameters for the C1 cloud type are $r_m = 4.75$, $\alpha = 5.0$, and $\gamma = 1.61$. The cloud macrophysical properties (classifications of CTH and CGT) are based on the tables in [35].

A. Information Content

We selected a representative subset of 25 986 simulated measurements from the template data set and added measurement noise by specifying a signal-to-noise ratio (SNR) of 100 for all observations. For each retrieval scenario, the choice of best regularization parameter is performed based on the analysis of retrieval error for each of the cloud parameters and on the measurements of information content given by DFS and SIC. The retrieval errors are summarized in Table II.

In the ideal case, we want to retrieve all four macrophysical cloud parameters describing the cloudy atmosphere. The state vector is represented as $x = \{c_{th}, c_{gt}, c_\tau, c_f\}$, with c_{th} as the CTH, c_{gt} as the CGT, c_τ as the COT, and c_f as the CF.

The condition number values for this retrieval case are very high, with a median value of 2142.8 (see left panel of Fig. 1). Therefore, the Hadamard condition on stability [36] is violated, and this retrieval problem is ill-posed. The inverse problem is solved using (4); we find that the best results are obtained by applying Tikhonov regularization with regularization parameter $\alpha = 10^{-6}$. The averaged retrieval errors for c_{th}, c_{gt}, c_τ, and c_f, together with DFS and SIC, are depicted in the right panel of Fig. 1.

The errors for each retrieved cloud parameter are depicted in Fig. 2. All four retrieved parameters significantly deviate from the *true* values with the following error mean and standard deviation values: c_{th} is 0.243 ± 0.906 km, c_{gt} is 0.061 ± 1.01 km, c_τ is 15.1 ± 40.2, and c_f is 0.02 ± 0.22. The value of DFS for the optimal α is about 2.4. This indicates that the information contained in the oxygen A-band spectra is not enough to extract all four cloud parameters.

As a next step, we assume that the CF c_f is known and is delivered from independent measurements. In this case, our state vector is $x = \{c_{th}, c_{gt}, c_\tau\}$. For this problem, the values of the condition number (shown in Fig. 3) are lower than those for the retrieval of four parameters but still relatively high, with an average of 552.6. This illustrates that, the same way as in the first case, the Hadamard condition on stability [36] is violated and that regularization is required. We perform Tikhonov regularization, and the best results are obtained for $\alpha = 10^{-6}$ as on the previous case.

The retrieved values of COT are of good accuracy, with an average error of 0.19 ± 12.3, the errors of CTH are acceptable (0.092 ± 0.606 km), but the errors of the CGT are high (0.127 ± 0.96 km; see Fig. 4).

The DFS for this retrieval configuration has a value of about 2.1, i.e., the information in the oxygen A-band is also not enough to retrieve these three cloud parameters. In particular, variations in CTH and CGT cause similar changes in the simulated spectrum. In Fig. 5, we plot two spectra, corresponding to different combinations of c_{th} and c_{gt} values under the same atmospheric conditions (surface height of 0 km, surface albedo of 0.45, COT of 1.3, sun zenith angle of 65°, viewing zenith angle of 50°, and relative azimuth of 90°). The difference between these spectra is, in general, smaller than 1%, and therefore, it is below the instrumental and modeling errors. As a consequence, it is not possible to distinguish between these two spectra during the retrieval.

TABLE II
RESULTS FOR THE RETRIEVAL OF DIFFERENT COMBINATIONS OF MACROPHYSICAL CLOUD PARAMETERS

Nr. of retrieved parameters (N)	CTH [km]	CGT [km]	COT [-]	CF [-]	DFS/N	SIC/N
4	0.243±0.906	0.061±1.010	15.10±40.20	0.02±0.22	0.610	1.917
3	0.092±0.606	0.127±0.960	0.19±12.30		0.704	2.005
2	0.013±0.390		-0.56±14.92		0.994	5.036
2	0.039±0.587			0.008±0.06	0.996	5.443
2			12.47±37.39	0.015±0.167	0.894	4.332
1		0.442±0.465			0.913	2.651

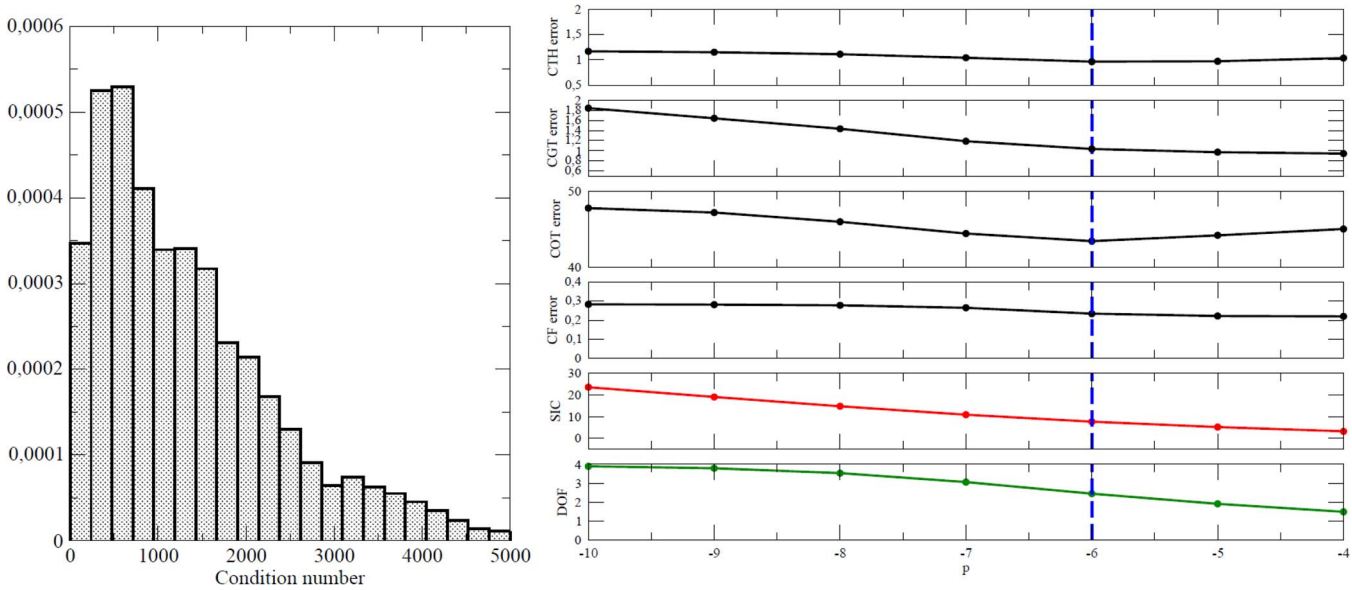

Fig. 1. (Left) Histogram of the condition number values for the problem of retrieving four cloud parameters $c_{\rm th}$, $c_{\rm gt}$, c_τ, and c_f. (Right) Corresponding averaged retrieval errors for $c_{\rm th}$, $c_{\rm gt}$, c_τ, and c_f, and the information content measurements SIC and DFS as a function of regularization parameter $\alpha = 10^p$. The dashed line (blue) shows the minimal retrieval error constellation.

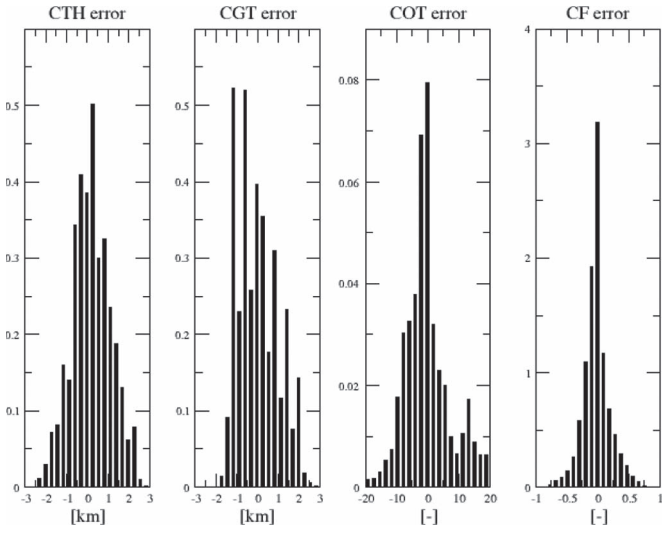

Fig. 2. Histograms of errors for the retrieval of four cloud parameters $c_{\rm th}$, $c_{\rm gt}$, c_τ, and c_f.

As a third case, we explore the possibility to retrieve two parameters c_{th} and c_τ, assuming known values of c_f and c_{gt}. The condition number for this problem is depicted in Fig. 6. It can be seen that the condition number values with an average of 18.5 are considerably lower than those of the case of the retrieval of four and three cloud parameters. However, the problem is still slightly ill posed and needs a small amount of regularization, and the best results are obtained for $\alpha = 10^{-8}$.

As seen from the histograms in Fig. 7 (left panel), the retrieval of these two cloud properties can be performed with a good accuracy of c_{th} of 0.013 $\pm$ 0.39 km and c_τ of -0.56 ± 14.92.

Similar good results are obtained for the retrieval of c_{th} and c_f, assuming known values of c_τ and c_{gt}. The values of the condition number are even slightly lower than those in the retrieval of c_{th} and c_τ with an average of 17.3. The accuracies of the retrieved c_{th} and c_f are good (c_{th} of 0.009 $\pm$ 0.587 km and c_f of 0.008 $\pm$ 0.06). The DOF for these last two cases (c_{th} and c_τ, and c_{th} and c_f) is almost 2.

We test the retrieval of two strongly correlated parameters c_τ and c_f, assuming known c_{gt} and c_{th}. Compared to the previous cases of two-parameter retrieval, the condition number values are larger, and we need more regularization. As expected, the retrieved values deviate significantly from the *true* values (c_τ of 12.47 $\pm$ 37.39 and c_f of 0.015 $\pm$ 0.167; see right panel of Fig. 7). This test confirms that it is not possible to retrieve simultaneously c_τ and c_f solely from the oxygen A-band spectrum.

Finally, we analyze the retrieval of a single cloud parameter. The retrieval works well for all cases, except for the CGT. In this case, a larger regularization is needed ($\alpha = 10^{-7}$), and the errors are relatively large (0.442 $\pm$ 0.465 km).

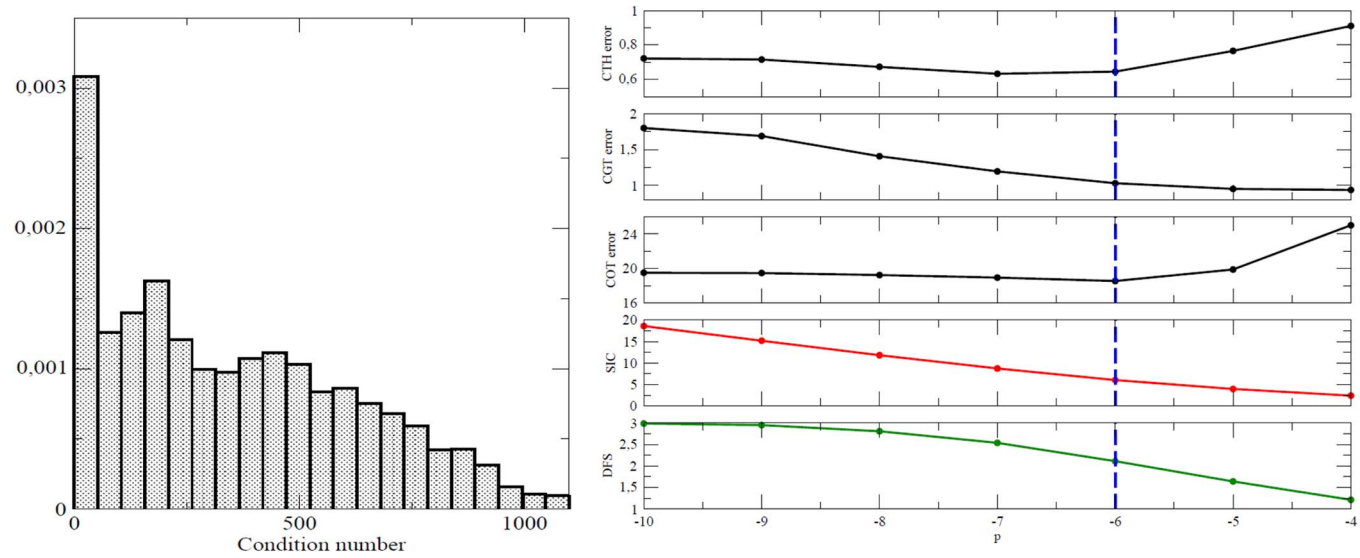

Fig. 3. Same as Fig. 1 but for the retrieval of three cloud parameters $c_{\rm th}$, $c_{\rm gt}$, and c_τ.

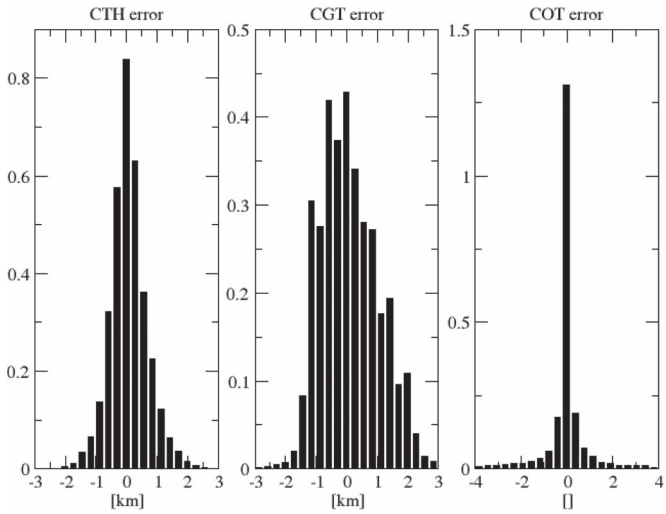

Fig. 4. Same as Fig. 2 but for the retrieval of three cloud parameters $c_{\rm th}$, $c_{\rm gt}$, and c_τ.

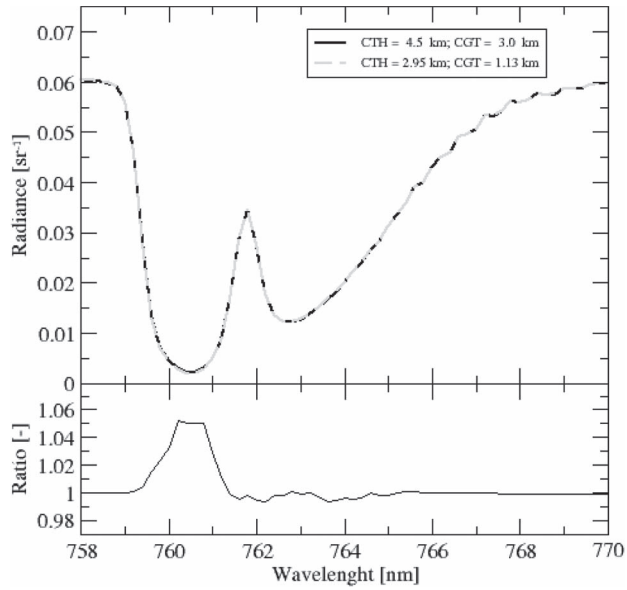

Fig. 5. Two simulated spectra under the same atmospheric conditions but different $c_{\rm th}$ and $c_{\rm gt}$ values (top) and corresponding ratio (bottom). The differences between the two spectra outside the 760–761-nm region are well below 1%.

B. Initial Sensitivity Studies

In inverse modeling, there are various sources of error in addition to measurement noise [31]. On the measurement side, there can be systematic errors (bias) due to calibration errors. On the modeling side, we distinguish between the following: 1) model parameter error, due to uncertainty in the values of assumed (i.e., not retrieved) parameters in the forward modeling, and 2) forward model error, due to uncertainties implicit in the mathematical and physical approximations used in the forward calculation itself. Most model parameter errors (and all forward model errors) are systematic.

For our inverse algorithm, there are a number of model parameters contributing to retrieval error. For the two-parameter retrieval of c_{th} and c_τ, the most significant parameters are geometrical cloud thickness, CF, and surface albedo, but also to be included in the list of such errors are quantities such as the droplet size distribution parameters assumed for cloud optical properties, the temperature and pressure values in the atmospheric stratification, and the *A*-band line spectroscopic parameters.

The list of forward model errors contains things such as the assumption of zero polarization in the radiative transfer, the choice of the number of computational layers and discrete ordinates in the radiative transfer, the assumption of cloud vertical homogeneity, the use of IPA, and the use, in general, of 1-D modeling at the expense of 3-D cloud effects. Some of these error sources are difficult, if not impossible, to quantify; in the sequel, we will focus on the first two of these (neglect of polarization and computational discretizations).

In this section, we give a flavor for error sensitivity studies in our algorithm by looking at two of the most significant sources of model parameter uncertainty: the parameter values for geometrical cloud thickness and the surface albedo. The inverse problem is restricted to the retrieval of the two macrophysical

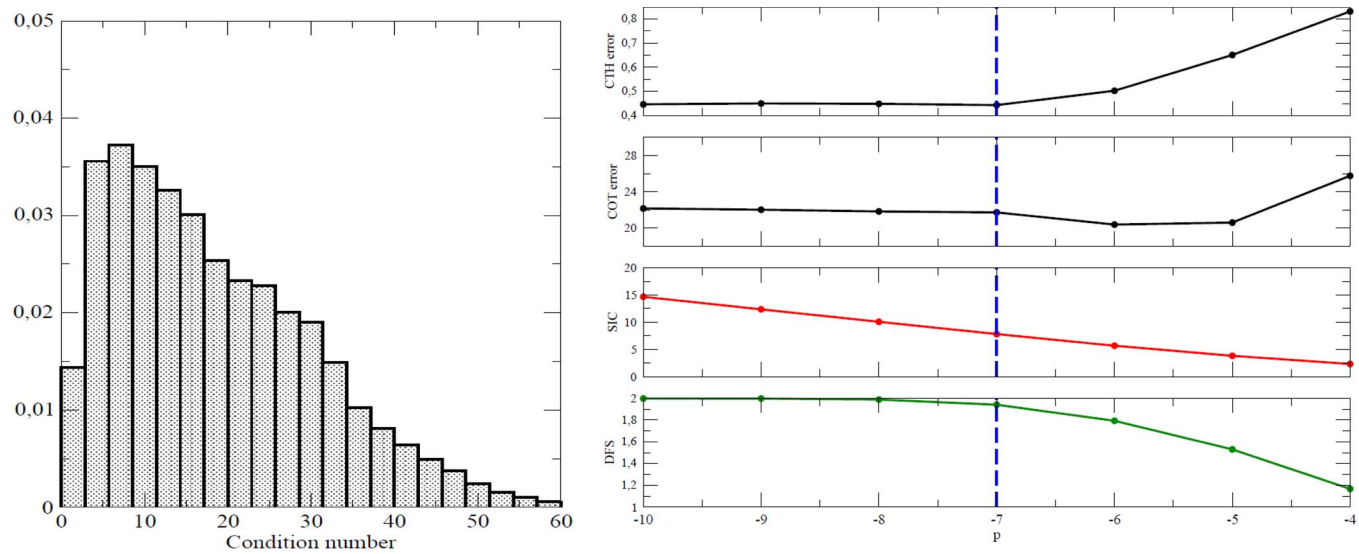

Fig. 6. Same as Fig. 1 but for the retrieval of two cloud parameters c_{th} and c_{τ}.

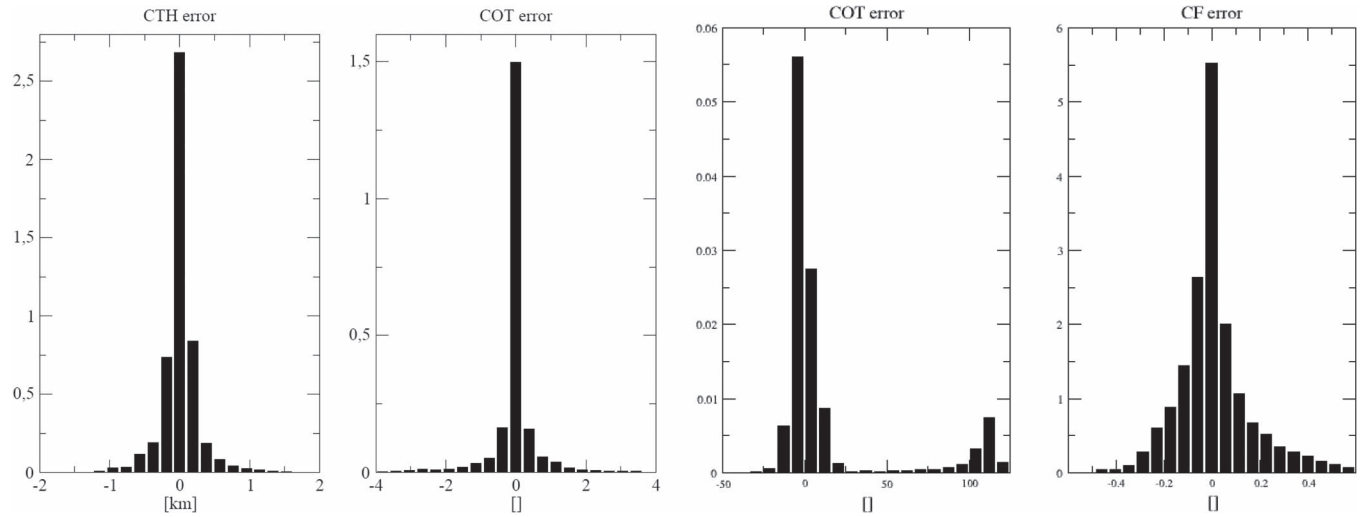

Fig. 7. Same as Fig. 2 but for the retrieval of two cloud parameters c_{th} and c_{τ} (left), and c_{τ} and c_f (right).

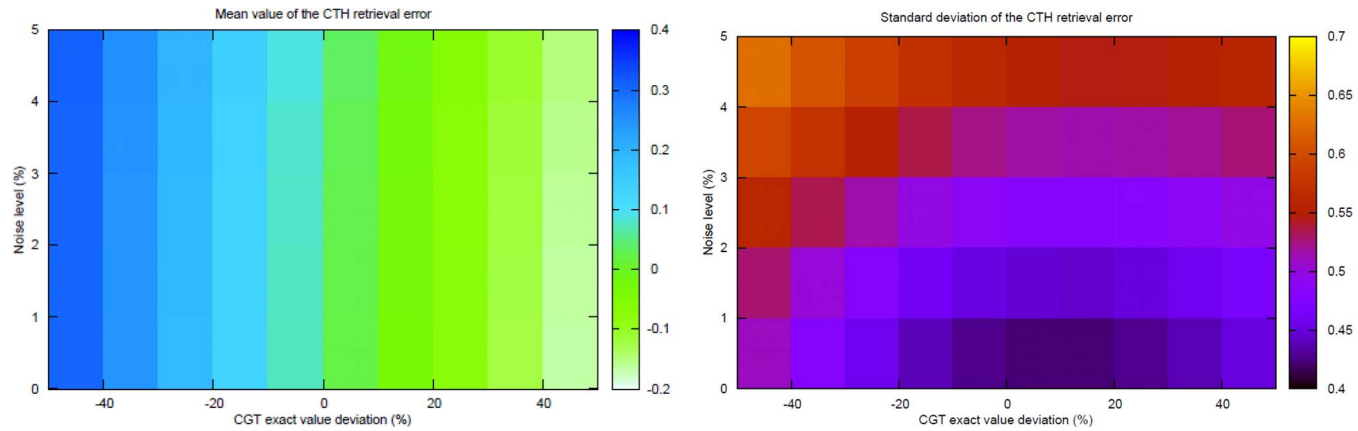

Fig. 8. (Left) Mean value and (right) standard deviation of the retrieval error for c_{th} as a function of noise level and c_{gt} exact value deviation.

parameters c_{th} and c_{τ} (similar results were obtained for the retrieval of c_{th} and c_f, not shown here).

First, we explore the influence of the uncertainty in the CGT on the retrieval. In case of c_{th}, the retrieval can be performed with sufficient accuracy even if the value of c_{gt} is underestimated by 40% or overestimated by 30% for the noise level under 1% (corresponding to an SNR of 100 as for the GOME instrument) or with over- and underestimation by 20% and -10% for the noise of 3% (corresponding to an SNR of 50). The mean value and standard deviation of the retrieval error for

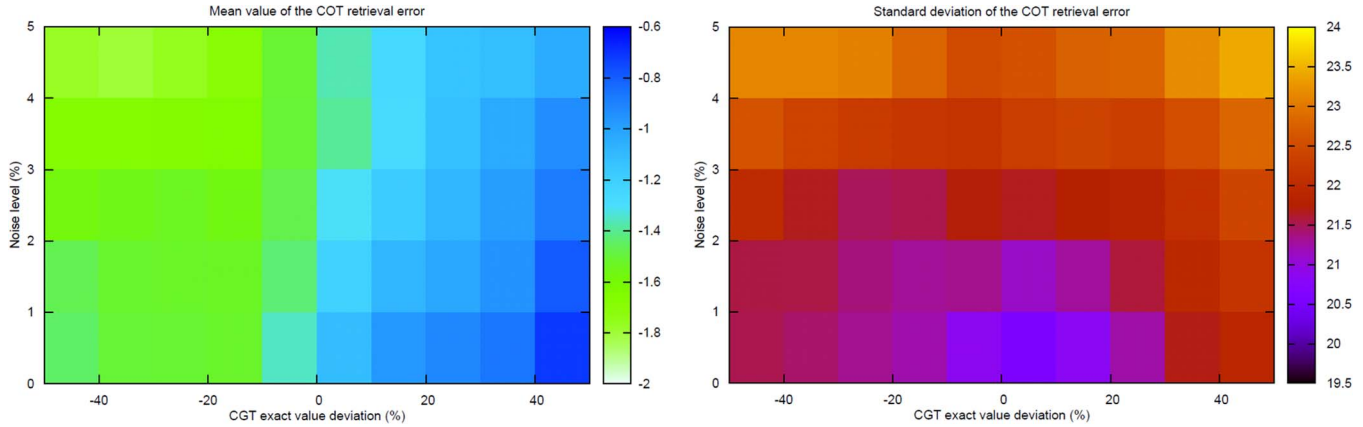

Fig. 9. (Left) Mean value and (right) standard deviation of the retrieval error for c_τ as a function of noise level and c_{gt} exact value deviation.

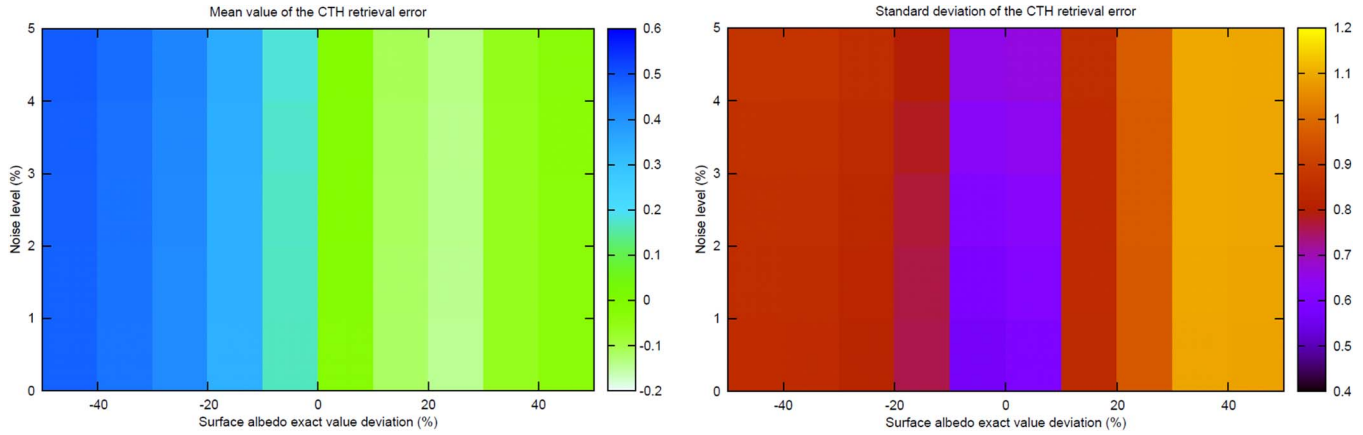

Fig. 10. (Left) Mean value and (right) standard deviation of the retrieval error for c_{th} as a function of noise level and surface albedo exact value deviation.

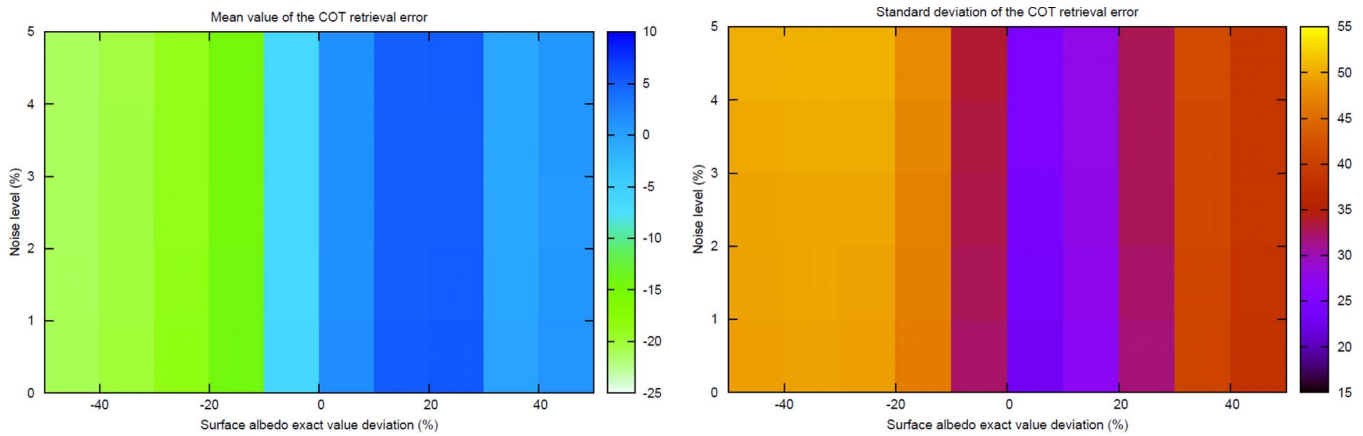

Fig. 11. (Left) Mean value and (right) standard deviation of the retrieval error for c_τ as a function of noise level and surface albedo exact value deviation.

c_{th} are depicted in Fig. 8. For the retrieval of c_τ, the retrieved values are slightly overestimated for all variations of c_{gt} (see Fig. 9). The retrieval of c_τ can be performed with satisfactory accuracy even in case of under- and overestimation of the c_{gt} value by 40% for the noise level under 2%.

Second, the influence of the uncertainty in the surface albedo value is shown in Fig. 10 for c_{th} and in Fig. 11 for c_τ. In both cases, the retrieval can be performed accurately if the surface albedo value is under- or overestimated by not more than 10%, for noise values up to 5% (which corresponds to an SNR of 20).

V. Conclusion

In this paper, we have analyzed the information content in the O_2A-band for the retrieval of cloud macrophysical properties. Clouds are modeled as optically uniform layers of scattering water droplets, and the VLIDORT radiative transfer model is used in generating a comprehensive set of simulations. The method used for the retrieval is Tikhonov regularization with an optimal chosen regularization parameter. The choice of regularization parameter was performed individually for each retrieval

scenario by analyzing the averaged errors of the retrieval for each cloud property, the degree of freedom for the signal, and the Shannon information criteria.

For all retrieval problems discussed in this paper, the degree of freedom for the signal is a decreasing function of the regularization parameter. Thus, the less regularization is applied, the higher is the degree of freedom, approaching its maximum when no regularization is applied. With a high degree of regularization, the DFS becomes small, and the solution is dominated by the *a priori* contribution.

As with the DFS, the SIC is higher for small amounts of regularization and lower when more regularization is introduced.

It was shown that it is not possible to retrieve simultaneously four cloud properties such as CF, CTH, CGT, and COT with satisfactory accuracy as the measurements in the O_2A-band do not provide enough information for the retrieval of these four parameters.

For the retrieval of three parameters, assuming a known CF, the errors of c_{th} and c_{th} were acceptable, but it was not possible to retrieve c_{gt} accurately enough.

The smallest errors are obtained for the retrievals of the following:

1) CTH and COT;
2) CTH and CF.

The simultaneous retrieval of COT and CF is also not possible using solely the O_2 A-band measurements with the moderate spectral resolution of 0.2 nm as the information on these two parameters contained in the measurements is strongly correlated. This remains true, even for higher SNR values.

It is obvious that the retrieval of a single parameter (c_f, c_{th}, or c_τ), assuming known values for the rest of the parameters, will work at least as well as that for the coupled retrieval. To figure out the information content on c_{gt} in the O_2A-band, we have performed the retrieval of only this parameter. The retrieval errors tell us that the information in the measurements on c_{gt} is not strong enough to retrieve it with satisfactory accuracy, even assuming that all other parameters are known. This result confirms the work from [37] that shows the need of multidirectional O_2A-band measurements for the retrieval of cloud geometric thickness, taking advantage of the tight correlation between the angular standard deviation of the cloud oxygen pressure and the CGT.

Finally, the sensitivity study for the case of two-parameter retrieval $\{c_{th}, c_\tau\}$ shows that the results are slightly sensitive to the uncertainty in the c_{gt} input value and permits c_{gt} deviations from the exact value of up to $\pm 40\%$ for an SNR of 100. On the other hand, the retrieval is very sensitive to the uncertainty in the surface albedo; acceptable retrievals are obtained only up to a $\pm 10\%$ deviation of the surface albedo.

REFERENCES

[1] M. Zelinka, S. Klein, and D. Hartmann, "Computing and partitioning cloud feedbacks using cloud property histograms. Part II: Attribution to changes in cloud amount, altitude, and optical depth," *J. Climate*, vol. 25, no. 11, pp. 3736–3754, Jun. 2012.

[2] X. Liu, M. Newchurch, R. Loughman, and P. K. Bhartia, "Errors resulting from assuming opaque Lambertian clouds in TOMS ozone retrieval," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 85, no. 3, pp. 337–365, May 2004.

[3] Z. Ahmad, P. K. Bhartia, and N. Krotkov, "Spectral properties of backscattered UV radiation in cloudy atmospheres," *J. Geophys. Res.*, vol. 109, no. D1, pp. D01201-1–D01201-11, Jan. 2004.

[4] K. Boersma, H. Eskes, and E. Brinksma, "Error analysis for tropospheric NO_2 retrieval from space," *J. Geophys. Res.*, vol. 109, no. D4, pp. D44311-1–D44311-20, Feb. 2004.

[5] M. Van Roozendael, D. Loyola, R. Spurr, D. Balis, J.-C. Lambert, Y. Livschitz, P. Valks, T. Ruppert, P. Kenter, C. Fayt, and C. Zehner, "Ten years of GOME/ERS2 total ozone data: The new GOME Data Processor (GDP) Version 4: I. Algorithm description," *J. Geophys Res.*, vol. 111, no. D14, pp. D14311-1–D14311-21, Jul. 2006.

[6] M. Sneep, J. F. de Haan, P. Stammes, P. Wang, C. Vanbauce, J. Joiner, A. P. Vasilkov, and P. F. Levelt, "Three-way comparison between OMI and PARASOL cloud pressure products," *J. Geophys Res.*, vol. 113, no. D15, pp. D15S23-1–D15S23-11, Aug. 2008.

[7] D. L. Wu, S. A. Ackerman, R. Davies, D. J. Diner, M. J. Garay, B. H. Kahn, B. C. Maddux, C. M. Moroney, G. L. Stephens, J. P. Veefkind, and M. A. Vaughan, "Vertical distributions and relationships of cloud occurrence frequency as observed by MISR, AIRS, MODIS, OMI, CALIPSO, and CloudSat," *Geophys. Res. Lett.*, vol. 36, no. 9, pp. L09821-1–L09821-5, May 2009.

[8] T. S. L'Ecuyer, P. Gabriel, K. Lessman, S. J. Cooper, and G. L. Stephens, "Objective assessment of the information content of visible and infrared radiance measurements for cloud microphysical property retrievals over the global oceans. Part I: Liquid clouds," *J. Appl. Meteorol. Climatol.*, vol. 45, no. 1, pp. 20–41, Jan. 2006.

[9] J. P. Burrows, M. Weber, M. Buchwitz, V. V. Rozanov, A. Ladstädter-Weissenmayer, A. Richter, R. de Beek, R. Hoogen, K. Bramstedt, K.-U. Eichmann, M. Eisinger, and D. Perner, "The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results," *J. Atmos. Sci.*, vol. 56, no. 2, pp. 151–175, Jan. 1999.

[10] H. Bovensmann, J. P. Burrows, M. Buchwitz, J. Frerick, S. Noel, V. V. Rozanov, K. V. Chance, and A. P. H. Goede, "SCIAMACHY: Mission objectives and measurement modes," *J. Atmos. Sci.*, vol. 56, no. 2, pp. 127–150, Jan. 1999.

[11] J. Callies, E. Corpaccioli, M. Eisinger, A. Hahne, and A. Lefebvre, "GOME-2—MetOp's second generation sensor for operational ozone monitoring," *ESA Bull.*, vol. 102, pp. 28–36, May 2000.

[12] P. Ingmann, B. Veihelmann, J. Langen, D. Lamarre, H. Stark, and G. Bazalgette Courrèges-Lacoste, "Requirements for the GMES atmosphere service and ESA's implementation concept: Sentinels-4/-5 and -5P," *Remote Sens. Environ.*, vol. 120, pp. 58–69, May 2012.

[13] D. Loyola, W. Thomas, Y. Livschitz, T. Ruppert, P. Albert, and R. Hollmann, "Cloud properties derived from GOME/ERS-2 backscatter data for trace gas retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2747–2758, Sep. 2007.

[14] D. Loyola, W. Thomas, R. Spurr, and B. Mayer, "Global patterns in daytime cloud properties derived from GOME backscatter UV–VIS measurements," *Int. J. Remote Sens.*, vol. 31, no. 16, pp. 4295–4318, May 2010.

[15] A. Vasilkov, J. Joiner, R. Spurr, P. K. Bhartia, P. Levelt, and G. Stephens, "Evaluation of the OMI cloud pressures derived from rotational Raman scattering by comparisons with other satellite data and radiative transfer simulations," *J. Geophys. Res.*, vol. 113, no. D15, pp. D15S19-1–D15S19-12, Aug. 2008.

[16] J. Joiner and A. P. Vassilkov, "First results from the OMI rotational Raman scattering cloud pressure algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1272–1282, May 2006.

[17] M. Van Roozendael, R. Spurr, D. Loyola, C. Lerot, D. Balis, J.-C. Lambert, W. Zimmer, J. van Gent, J. van Geffen, M. Koukouli, J. Granville, A. Doicu, C. Fayt, and C. Zehner, "Sixteen years of GOME/ERS2 total ozone data: The new direct-fitting GOME Data Processor (GDP) Version 5: I. Algorithm description," *J. Geophys. Res.*, vol. 117, no. D3, pp. D03305-1–D03305-18, Feb. 2012.

[18] D. Loyola, M. Koukouli, P. Valks, D. Balis, N. Hao, M. Van Roozendael, R. Spurr, W. Zimmer, S. Kiemle, C. Lerot, and J.-C. Lambert, "The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation," *J. Geophys. Res.*, vol. 116, no. D7, pp. D07302-1–D07302-11, Apr. 2011.

[19] C. J. Hahn, W. B. Rossow, and S. G. Warren, "ISCCP cloud properties associated with standard cloud types identified in individual surface observations," *J. Climate*, vol. 14, no. 1, pp. 11–28, Jan. 2001.

[20] M. Antón and D. Loyola, "Influence of cloud properties on satellite total ozone observations," *J. Geophys. Res.*, vol. 116, no. D3, pp. D03208-1–D03208-11, Feb. 2012.

[21] R. J. D. Spurr, "VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies

in multilayer multiple scattering media," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 102, no. 2, pp. 316–342, Nov. 2006.
[22] R. Spurr, "LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems," in *Light Scattering Reviews*, vol. 3, A. Kokhanovsky, Ed. Berlin, Germany: Springer-Verlag, 2008.
[23] L. Rothman, I. E. Gordon, A. Barbe, D. C. Benner, P. F. Bernath, M. Birk, V. Boudon, L. R. Brown, A. Campargue, J.-P. Champion, K. Chance, L. H. Coudert, V. Dana, V. M. Devi, S. Fally, J.-M. Flaud, R. R. Gamache, A. Goldman, D. Jacquemart, I. Kleiner, N. Lacome, W. J. Lafferty, J.-Y. Mandin, S. T. Massie, S. N. Mikhailenko, C. E. Miller, N. Moazzen-Ahmadi, O. V. Naumenko, A. V. Nikitin, J. Orphal, V. I. Perevalov, A. Perrin, A. Predoi-Cross, C. P. Rinsland, M. Rotger, M. Šimečková, M. A. H. Smith, K. Sung, S. A. Tashkun, J. Tennyson, R. A. Toth, A. C. Vandaele, and J. Vander Auwera, "The HITRAN 2008 molecular spectroscopic database," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 110, no. 9/10, pp. 533–572, Jun. 2009.
[24] B. Bodhaine, N. Wood, E. Dutton, and J. Slusser, "On Rayleigh optical depth calculations," *J. Atmos. Ocean. Technol.*, vol. 16, no. 11, pp. 1854–1861, Nov. 1999.
[25] F. Schreier and B. Schimpf, "A new efficient line-by-line code for high resolution atmospheric radiation computations incl. derivatives," in *IRS 2000: Current Problems in Atmospheric Radiation*, W. L. Smith and Y. Timofeyev, Eds. Hampton, VA, USA: A. Deepak Publishing, 2001, pp. 381–384.
[26] F. Schreier, "Optimized implementations of rational approximations for the Voigt and complex error function," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 112, no. 6, pp. 1010–1025, Apr. 2011.
[27] E. Boesche, P. Stammes, R. Preusker, R. Bennartz, W. Knap, and J. Fischer, "Polarization of skylight in the O_2 *A* band: Effects of aerosol properties," *Appl. Opt.*, vol. 49, no. 19, pp. 3467–3480, Jul. 2008.
[28] W. Wiscombe, "The delta-M method: Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions," *J. Atmos. Sci.*, vol. 34, no. 9, pp. 1408–1422, Sep. 1977.
[29] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*. Washington, DC, USA: Winston & Sons, 1977.
[30] A. Doicu, T. Trautmann, and F. Schreier, *Numerical Regularization for Atmospheric Inverse Problems*. Berlin, Germany: Springer-Verlag, 2010.
[31] C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Singapore: World Scientific, 2000.
[32] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana, IL, USA: Univ. of Illinois Press, 1949.
[33] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*., 2nd ed. Boca Raton, FL, USA: CRC Press, 2003.
[34] M. Hess, P. Koepke, and I. Schult, "Optical properties of aerosols and clouds: The software package OPAC," *Bull. Amer. Meteorol. Soc.*, vol. 79, no. 5, pp. 831–844, May 1998.
[35] J. Wang, W. B. Rossow, and Y. Zhang, "Cloud vertical structure and its variations from a 20-yr global rawinsonde dataset," *J. Climate*, vol. 13, no. 17, pp. 3041–3056, Sep. 2000.
[36] J. Hadamard, *Lectures on Cauchy's Problems in Linear Partial Differential Equations*., Dover Phoenix Editions ed. New York, NY, USA: Dover, 2003, [1923].
[37] N. Ferlay, F. Thieuleux, and C. Cornet, "Toward new inferences about cloud structures from multidirectional measurements in the oxygen a band: Middle-of-cloud pressure and cloud geometrical thickness from POLDER-3/PARASOL," *J. Appl. Meteorol. Climatol.*, vol. 49, no. 12, pp. 2492–2507, Dec. 2010.

Olena Schuessler received the B.Sc. degree in computer sciences from the National Technical University of Ukraine, Kiew, Ukraine, and the M.Sc. degree in earth-oriented space science and technology from the Technical University of Munich, Munich, Germany. She is currently working toward the Ph.D. degree in the Institut für Methodik der Fernerkundung (IMF), Deutsches Zentrum für Luft- und Raumfahrt (DLR), Weßling, Germany.

Her work is focused on the development of inversion algorithms for the retrieval of tropospheric ozone from UV/VIS nadir measurements and of cloud parameters from NIR measurements in the group of D. Loyola.

Diego Guillermo Loyola Rodriguez (SM'00) received the Lic.Inf. degree in computer science from ESLAI, Universidad de Lujan, Lujan, Argentina, and the Dr.Ing. degree in remote sensing from the Technische Universität München, München, Germany.

Since 1992, he has been a Scientist and a Project Manager with Deutsches Zentrum für Luft- und Raumfahrt (DLR), Oberpfaffenhofen, Germany. He was a Visiting Scientist with the NASA Goddard Space Flight Center in 2000. Since 2010, he has been with DLR's Institut für Methodik der Fernerkundung (IMF), where he leads the teams responsible for the operational atmospheric composition and cloud products from the European missions GOME/ERS-2, GOME-2/MetOp-A, GOME-2/MetOp-B, and the future Sentinel 5 Precursor. His research interests are remote sensing of atmospheric trace species and clouds, as well as computational intelligence techniques.

Dr. Loyola Rodriguez is a Senior Member of the IEEE Geoscience and Remote Sensing Society and IEEE Computational Intelligence Society and a member of the International Neural Network Society. He is an Associate editor of the EGU *Atmospheric Measurement Techniques* and IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Adrian Doicu received the Eng. and Ph.D. degrees from the Technical University of Bucharest, Bucharest, Romania, in 1986 and 1996, respectively.

He is currently a Senior Scientist with the Institut für Methodik der Fernerkundung (IMF), Deutsches Zentrum für Luft- und Raumfahrt (DLR), Weßling, Germany. His research interests include electromagnetic scattering theory, inverse problems, and radiative transfer.

Robert Spurr received the M.A. degree from Cambridge University, Cambridge, U.K., in 1979 and the Ph.D. degree from the Technical University of Eindhoven, The Netherlands, in 2001.

He was a Professional Meteorologist in the 1980s. He joined the remote sensing community in 1991, and he has 22 years of experience with radiative transfer in the Earth's atmosphere and ocean and with the retrieval of ozone and other atmospheric constituents. He spent 10 years working at the Harvard-Smithsonian Center for Astrophysics before leaving in January 2005 to set up RT Solutions, Inc., Cambridge, MA, USA, where he is the Director, offering consultancy services in radiative transfer and remote sensing applications. He is the main author of the LIDORT family of discrete ordinate radiative transfer codes.

Dr. Spurr is a member of the American Geophysical Union, AAAS, and Optical Society of America and is a Fellow of the Royal Meteorological Society.