

# Solar occultation with SCIAMACHY: algorithm description and first validation

J. Meyer, A. Bracher, A. Rozanov, A. C. Schlesier, H. Bovensmann, and J. P. Burrows

Institute of Environmental Physics, University of Bremen, Germany

Received: 20 September 2004 – Published in Atmos. Chem. Phys. Discuss.: 10 January 2005

Revised: 5 April 2005 – Accepted: 21 April 2005 – Published: 22 June 2005

**Abstract.** This presentation concentrates on solar occultation measurements with the spaceborne spectrometer SCIAMACHY in the UV-Vis wavelength range. Solar occultation measurements provide unique information about the vertical distribution of atmospheric constituents. For retrieval of vertical trace gas concentration profiles, an algorithm has been developed based on the optimal estimation method. The forward model is capable of simulating the extinction signals of different species as they occur in atmospheric transmission spectra obtained from occultation measurements. Furthermore, correction algorithms have been implemented to address shortcomings of the tangent height pre-processing and inhomogeneities of measured solar spectra. First results of O<sub>3</sub> and NO<sub>2</sub> vertical profile retrievals have been validated with data from ozone sondes and satellite based occultation instruments. The validation shows very promising results for SCIAMACHY O<sub>3</sub> and NO<sub>2</sub> values between 15 to 35 km with errors of the order of 10% and 15%, respectively.

## 1 Introduction

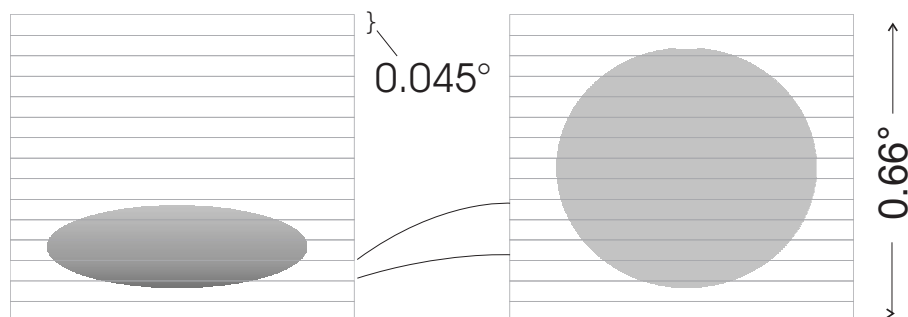
The concern about the anthropogenic impact on the ozone layer dramatically increased after the discovery of the Antarctic ozone hole in the early 1980s (Farman et al., 1985). Large efforts have been directed towards global ozone monitoring systems and the investigation of stratospheric chemistry. As a result, our knowledge of atmospheric dynamics, photochemistry, heterogeneous and homogeneous chemistry has increased significantly. However, the requirements of satellite remote sensors are still demanding as the complexity of scientific questions has increased correspondingly. Many other atmospheric constituents have been identified to play a key role in the ozone chemistry. Hence, there is a great

demand for height resolved concentration profiles of many different trace gases.

SCIAMACHY (SCanning Imaging Absorption spectrometer for Atmospheric CHartography) is a remote sensing instrument, which measures radiation in the UV-Vis-IR wavelength range in nadir, limb, and occultation geometry. It has successfully been launched onboard ENVISAT (ENVironmental SATellite) on 1 March in 2002. The instrument provides measurements of earthshine radiance and solar irradiance for the retrieval of height resolved concentration profiles of many different species in the Earth's atmosphere. Especially, the solar occultation measurements are expected to provide spectra with high signal-to-noise ratios, which allow for very accurate retrieval results. Details about scientific goals and instrument design can be found in Bovensmann et al. (1999). The high potential of spaceborne solar occultation measurements have already been demonstrated by various instruments such as SAGE I–III (Stratospheric Gas and Aerosol Experiment) (McCormick et al., 1979; Mauldin et al., 1985; Cunnold et al., 1989), POAM II/III (Polar Ozone and Aerosol Measurement instrument) (Glaccum et al., 1996; Lucke et al., 1999), and HALOE (HALogen Occultation Experiment) (Russel et al., 1993; Brühl et al., 1996).

ENVISAT's orbit is sun-synchronous and near-polar with an inclination of 98°. Nadir and limb measurements are performed alternately during day time. Solar occultation measurements take place whenever SCIAMACHY arises from the Earth's shadow occurring always adjacent to the North Pole. Depending on the season, air parcels investigated by solar occultation measurements are situated roughly between 50° and 70° North. The main scientific goals in this region are the Arctic ozone depletion and related studies of halogen compounds. In this study, the first validation results of O<sub>3</sub> and NO<sub>2</sub> vertical profile retrievals are presented. Since these species are more easily detected, they are appropriate for the first validation of the retrieval algorithms. Moreover, O<sub>3</sub> is of course the major target of most remote sensors and

Correspondence to: J. Meyer  
(jerome@iup.physik.uni-bremen.de)



**Fig. 1.** Occultation scan image. SCIAMACHY's fields of view are plotted onto the solar disk for a complete scan. One up or down scan over  $0.66^\circ$  takes two seconds. A typical scan through the atmosphere with refractive effects is shown on the left and a reference scan above the atmosphere is depicted on the right.

$\text{NO}_2$  is one of the key species required to understand atmospheric chemistry. SCIAMACHY is also designed for lunar occultation in the southern hemisphere, but this presentation is only focussed on the evaluation of solar occultation measurements.

The transmission through the atmosphere measured in occultation geometry can easily be modelled by the Lambert-Beer-Law:

$$\Upsilon = I_0 e^{-\int_0^{s_0} \alpha(s) ds}, \quad (1)$$

where  $\Upsilon$  is the radiation as measured by SCIAMACHY,  $I_0$  the extraterrestrial radiation from the Sun and  $\alpha$  the extinction coefficient, which depends on position  $s$  on the full path length  $s_0$  through the atmosphere. This approach assumes that the only radiation measured is directly transmitted through the atmosphere. The atmospheric transmission is calculated by dividing a measurement by an extraterrestrial spectrum. The latter can be taken from the same measurement sequence reaching tangent heights above 100 km. Atmospheric contributions are assumed to be negligible for this height range.

Modelled transmissions for each tangent height are calculated to retrieve height profiles using the optimal estimation method or the eigenvector decomposition method. Unfortunately, this procedure implies some difficulties concerning spatial inhomogeneities on the solar disk. A general limb darkening can be observed, i.e. measurements close to the solar limb will result in much smaller intensities than one would expect from pure geometric considerations. Furthermore, the limb darkening differs strongly for different Fraunhofer lines in the solar spectrum. This affects our retrieval algorithm as the differential structures of the transmissions are concerned.

Therefore, it is necessary to determine the exact pointing of SCIAMACHY's field of view on the solar disk. Figure 1 illustrates the problems of assigning a measurement to the corresponding reference measurement. Atmospheric

refraction makes it even more difficult as the image of the solar disk appears to be flattened. We developed an algorithm that eliminates mostly the Fraunhofer structure from the measured transmissions. The measurement scans above the atmosphere provide enough information about variations of the Fraunhofer structure.

Another important task is the line of sight geometry. In occultation geometry, the line of sight is uniquely defined by its tangent height. The exact knowledge of the line of sight geometry is important for the estimation of atmospheric refraction as well as for the radiative transfer modeling. The information on tangent heights can be improved by  $\text{O}_2$  retrievals. The  $\text{O}_2$  height profile is well known as it is directly related to pressure and temperature. So the tangent height grid can be adjusted until the  $\text{O}_2$  retrieval results match the "true" a-priori profile of  $\text{O}_2$ . A correction algorithm for tangent heights will also be presented in this study.

## 2 Occultation with SCIAMACHY

### 2.1 The spectrometer

SCIAMACHY is a passive remote sensing moderate-resolution imaging spectrometer. The spectral range is divided into eight channels each having 1024 wavelength pixels as shown in Table 1. In the solar occultation mode, the integration time is 62.5 ms. The signal-to-noise ratio is approximately 4000 in channel 6 and is well above 1000 throughout the visible and near infrared wavelength range. Details of SCIAMACHY's conceptual design can be found in Bovensmann et al. (1999).

### 2.2 Occultation scanning

SCIAMACHY's instantaneous field of view (IFOV) in solar occultation mode is  $0.045^\circ$  along the effective height and  $0.7^\circ$  along the width, which is a bit wider than the apparent diameter of the Sun. The relevant part of the atmosphere is

**Table 1.** Wavelengths and resolutions for the spectral channels of SCIAMACHY.

Channel	Wavelength [nm]	Resolution [nm]
1	240–314	0.24
2	309–405	0.26
3	394–620	0.44
4	604–805	0.48
5	785–1050	0.54
6	1000–1750	1.48
7	1940–2040	0.22
8	2265–2380	0.26

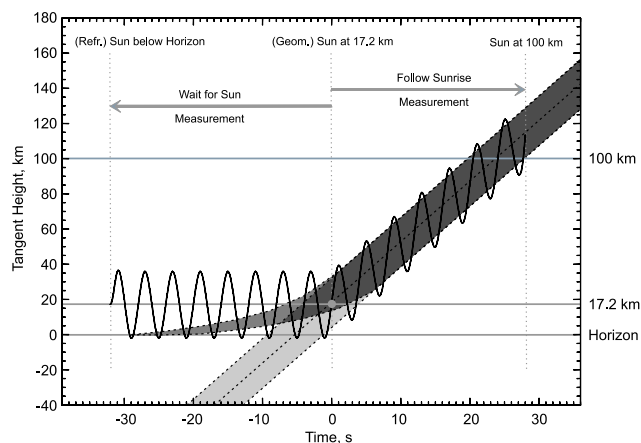
assumed to be 100 km high, which corresponds to an apparent height of approximately  $1.5^\circ$  seen from SCIAMACHY. The position of the IFOV is determined by two mirrors for azimuth and elevation movement. They correspond with the horizontal and vertical direction in the region of interest, respectively. The Sun is scanned up and down during sunrise. One complete scan (up and down) takes four seconds corresponding to a scanning speed of  $0.33^\circ/\text{s}$ . A complete scanning sequence is shown in Fig. 2. During a pixel exposure time of 62.5 ms, the FOV moves about  $0.02^\circ$  or nearly half of its own vertical dimension. The scanning speed is approximately six times higher than the elevation rate of the rising Sun.

The scanning sequence can be divided into two parts: waiting for the Sun and following the Sun. On the basis of precalculated orbital parameters, SCIAMACHY begins to scan the estimated sunrise region shortly above the horizon. When the centre of the geometric Sun reaches a tangent height of 17.2 km (at time = 0 s in Fig. 2), a Sun tracking device adjusts the FOV to the apparent Sun in azimuthal direction. The FOV is then moved with a precalculated elevation rate additionally to the scanning movement. After a total of 16 up and down scans or 64 s, the tangent height is approximately 100 km. Here, the line of sight is adjusted also vertically to the brightest point of the solar disk. Usually, the scanning is then continued up to a tangent height of 300 km. Alternatively, a limb-staring mode is possible without any scanning movement.

### 3 Retrieval algorithms

#### 3.1 Forward model

Any approach to the retrieval of atmospheric parameters from spectroscopic measurements requires a fundamental knowledge about optical properties of the atmosphere. Even those constituents that are not interesting in the context of the evaluation have to be modelled quite accurately. Knowing exact concentrations and behaviour of the relevant



**Fig. 2.** Schematic view of a solar occultation measurement sequence. Tangent height in km is plotted vs. time in seconds. The solid line represents the movement of SCIAMACHY's FOV, the shaded areas illustrate the refracted and the imaginary true Sun, respectively.

atmospheric constituents, it is in fact possible to simulate all measurements.

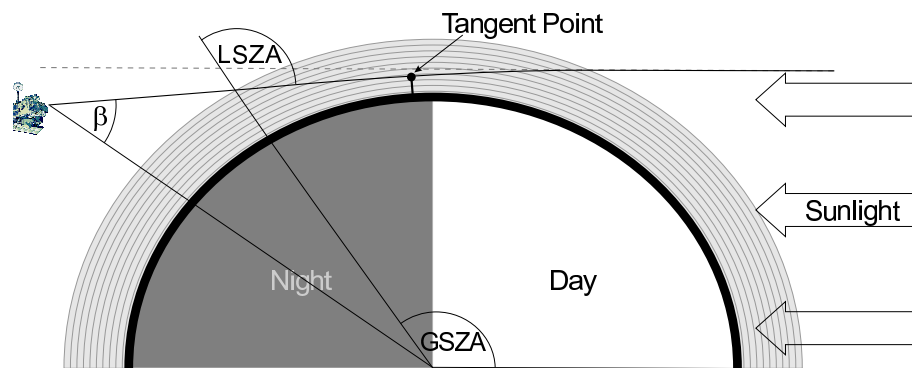
The occultation forward model is comparably simple as only directly transmitted radiation has to be taken into account. It has been developed based on the radiative transfer model GOMETRAN (Rozanov et al., 1997), which was originally dedicated to requirements of the GOME instrument (Burrows et al., 1999). The occultation geometry was implemented into the forward model by Rozanov (2001).

The major issue of an occultation forward model is the calculation of the full optical depth  $\tau$  along the line of sight as illustrated in Fig. 3.  $\tau$  is given by the integral in Eq. (1) and is a function of wavelength  $\lambda$  and tangent height. The tangent height in turn is uniquely defined by the solar zenith angle  $\psi$  (GSZA) at the top of atmosphere. Therefore,  $\tau$  can be written as

$$\tau(\psi, \lambda) = \int_0^{s_0} \alpha(\tilde{\psi}(s), \lambda) ds, \quad (2)$$

where  $\alpha$  is the overall extinction coefficient, which is a function of wavelength and local solar zenith angle  $\tilde{\psi}$  (LSZA) at the position  $s$  along the line of sight. In terms of computational considerations, the path length elements  $ds$  refer to an appropriate height layer grid. Horizontally, the atmosphere is assumed to be homogeneous. In the case of occultation geometry, it is furthermore important to take atmospheric refraction into account for calculation of  $ds$ . This is done by successive application of Snell's law.

Since real spectrometers cover a finite solid angle  $\Omega$  and single detector pixels correspond always to a finite



**Fig. 3.** Occultation geometry. For computational issues, the atmosphere is assumed to extend to 100 km height.

wavelength interval  $\Delta\lambda$ , the radiance  $\Upsilon$  measured by a real instrument can be written as

$$\Upsilon(\psi, \lambda, \beta) = \int_{\Omega} d\omega \int_{\Delta\lambda} d\lambda' I_0(\lambda', \beta) f(\omega) a(\lambda, \lambda') e^{-\tau(\psi, \lambda)}, \quad (3)$$

where  $\omega$  is the solid angle,  $a(\lambda, \lambda')$  is the slit function, and  $\beta$  is the elevation angle of the instrument's field of view (FOV). Again,  $I_0$  is the solar source function. A gaussian type slit function with wavelength dependent full-width half-maxima turned out to be sufficient for the requirements of the forward model (Ahlers and Dobber, 2000).  $f(\omega)$  describes the FOV characteristics. Theoretically, it must be integrated over the whole FOV. In practice, only those parts have to be considered that are illuminated by the solar disk, i.e.  $f(\omega)=1$  if  $d\omega$  is illuminated, otherwise  $f(\omega)=0$ . Note, that  $\Upsilon$  and  $I_0$  depend also on the scanning angle  $\beta$  of the instrument. In the context of forward modeling, it is useful to define  $\beta=0$  to be exactly centred on the solar disk with maximum intensity.  $I_0$  is reduced by the limb darkening effect described in Sect. 5 in combination with the geometrical effect of a decreasing solar disk area within the FOV as illustrated in Fig. 1. The impact of the limb darkening effect on occultation measurements with SCIAMACHY is described in detail by van Rie (1995) and is taken into account by the forward model. Additionally, the Sun flattening affects significantly the radiation  $\Upsilon$ , especially the dependence on  $\beta$ . This has not yet been implemented in the forward model as this effect is quite complex and hard to simulate. In Sect. 5, a correction algorithm is discussed which addresses this shortcoming of the forward model.

The extinction coefficient  $\alpha$  depends on wavelength  $\lambda$ , pressure  $p$ , and temperature  $T$ . Combining absorption cross-section  $\sigma$  and number density  $N$  of a species leads to the absorption coefficient  $\alpha$ . Furthermore, cross-sections and extinction/absorption coefficients are superposed to an overall extinction coefficient:

$$\alpha(\lambda, p, T) = \alpha_a(\lambda) + \alpha_r(\lambda) + \sum_{i=1}^M \sigma_i(\lambda, p, T) N_i. \quad (4)$$

Equation (4) is a sum of trace gas contributions, aerosol extinction coefficient  $\alpha_a$ , and Rayleigh scattering  $\alpha_r$ , where  $M$  is the number of relevant trace gases.  $\alpha_r$  is calculated as described by Bates (1984). Aerosol profiles and optical properties are taken from the LOWTRAN-7 aerosol model (Kneizys et al., 1988).

A forward model requires sufficient a-priori data with respect to the chosen height layer grid as well as molecule properties such as trace gas absorption cross sections. Cross sections of trace gases are either calculated with the so called line-by-line method or taken from measured moderate resolution cross section spectra such as (Bogumil et al., 1999) and (Burrows et al., 1998). Measured cross sections are commonly used in the UV-Vis wavelength range. The fine spectral structures of  $O_2$ ,  $H_2O$ , and typical infrared absorbers need to be simulated with the line-by-line method. Optical properties for the line-by-line calculations are taken from the HITRAN database (Rothman et al., 1998). A-priori vertical concentration profiles as well as information on pressure and temperature are taken from the US-standard climatology (NASA, 1976) and profiles provided by Anderson et al. (1986) as well as model calculations from the Max-Planck-Institut, Mainz (Brühl and Crutzen, personal communication, 1991).

### 3.2 Retrieval of trace gas profiles

The fundamental problem of spaceborne spectroscopy is to derive atmospheric parameters such as trace gas concentrations from spectral measurements, which are not related in a simple and obvious way. Mathematically, the retrieval problem is referred to as "inverse problem". The retrieval of atmospheric parameters is naturally an ill posed problem as real spectrometers have only limited spectral and spatial resolution and some components of interest do not contribute significantly to spectral signals.

It is the central task of the inverse problem to solve the following equation for vector  $\mathbf{x}$  denoting trace gas vertical

profiles or any other ensemble of parameters to be retrieved:

$$\mathbf{y} = \mathbf{A}(\mathbf{x}) + \epsilon, \quad (5)$$

where  $\mathbf{y}$  contains spectral intensities, e.g. measured by SCIAMACHY, and operator  $\mathbf{A}$  stands for the non-linear forward model.  $\epsilon$  comprises all possible errors of measurement and radiative transfer calculation. In this formulation, Eq. (5) describes the forward model introduced in Sect. 3.1. For numerical evaluation, the problem is approximated linearly by the first term of the Taylor series expansion:

$$\mathbf{y} = \mathbf{A}(\mathbf{x}) + \epsilon \approx \mathbf{y}_0 + \mathbf{K}_0(\mathbf{x} - \mathbf{x}_0). \quad (6)$$

$\mathbf{x}_0$  is an a-priori state vector as is already discussed in Sect. 3.1 and  $\mathbf{y}_0$  the corresponding measurement vector.  $\mathbf{K}_0$  is a linearised forward model operator. Just inverting the matrix  $\mathbf{K}_0$  is not useful as this problem is often ill-conditioned. Thus, there is a need of a-priori knowledge  $\mathbf{x}_0$  about the atmosphere which it is possible to find the most likely solution iteratively from.

A short introduction to the retrieval algorithms and their application to SCIAMACHY occultation measurements is given in the appendix.

#### 4 Correction of tangent height

All accuracies are limited at least by the knowledge about the line of sight geometry. Line of sight bending and tangent height are the most important quantities that are needed for radiative transfer calculations. Theoretically, it is possible to obtain tangent heights from pure orbital parameters of ENVISAT. The calculations can be implemented into evaluation algorithms using the CFI-Libraries (Customer Furnished Item) (Alvarez et al., 2000). The CFI routines are capable of calculating many useful geometric quantities, especially solar zenith angles (SZA) at the top of atmosphere and tangent heights with and without atmospheric refraction. They are also used by ESA in level-0 to level-1 data processing. Alternatively, refractive effects can be calculated by ray tracing routines as described by Kaiser (2001). Sometimes, it is even useful to use pure geometric tangent heights. In practice, it turned out that the calculations are not reliable and have to be improved.

##### 4.1 First approach to a tangent height retrieval

The idea of tangent height retrieval is based on the fact that there are gases whose volume mixing ratios are known quite accurately. This is true for  $\text{O}_2$ ,  $\text{O}_4$ , and – with some restrictions –  $\text{CO}_2$ . They are well mixed constituents of the atmosphere. Thus, their concentrations are directly related to pressure and temperature via the ideal gas law. Concentrations of  $\text{CO}_2$  have to be taken from appropriate climatologies. Since  $\text{CO}_2$  is chemically inert, it can be considered in a good approximation as well mixed gas for a wide height range in

the atmosphere. However, it is more and more photolysed within the stratosphere. The limiting factor is the knowledge of pressure and temperature profiles. For the following studies, globally averaged climatologies were used, such as the US-standard climatology (NASA, 1976). To some extent, they differ significantly from actual local profiles. However, it turned out that the differences can be neglected in the context of this presentation.

Vertical profiles of these species are regarded as true information for the tangent height retrieval. Corresponding profiles retrieved from SCIAMACHY measurements are divided by the true profiles bearing a measure of relative deviation from the true concentrations. An estimation for a better tangent height  $h$  at height layer  $i$  for iteration step  $k + 1$  is then calculated by

$$h_{i,k+1} = h_{i,k} - \frac{h_{i,k}^3}{f_{corr}} \ln \left( \text{abs} \left( \frac{C_{ox,i}^{ret}}{C_{ox,i}^{true}} \right) \right). \quad (7)$$

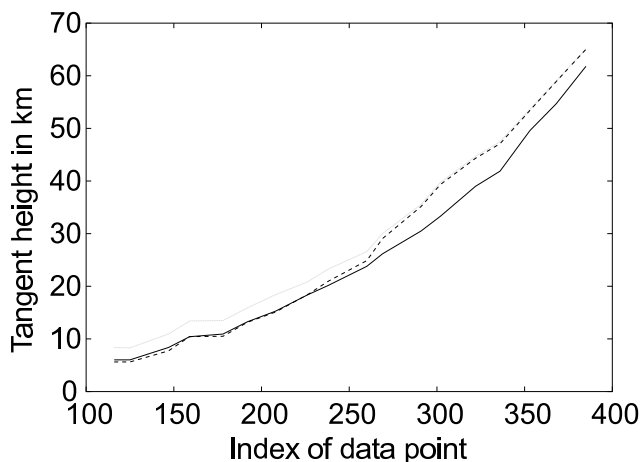
The superscripts *ret* and *true* indicate retrieval result and true information of trace gas concentration  $C_{ox,i}$ , respectively.  $f_{corr}$  is an empirical correction factor resulting from different parametrisation attempts. For  $\text{O}_2$  it is of the order of 6000. Smaller values force a faster but more instable retrieval, higher values lead to slower but stable calculations. The latter one is recommended if the initial values for the tangent heights provided by ESA differ strongly from the true values. For  $\text{CO}_2$ , Eq. (7) has to be modified slightly:

$$h_{i,k+1} = h_{i,k} - \frac{h_{i,k}^2}{f_{corr}} \ln \left( \text{abs} \left( \frac{C_{\text{CO}_2,i}^{ret}}{C_{\text{CO}_2,i}^{true}} \right) \right) \quad (8)$$

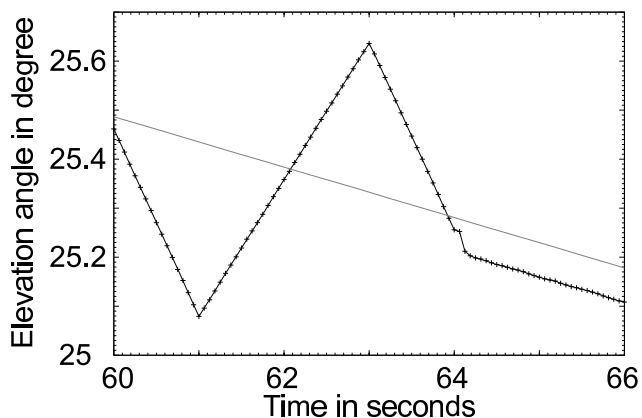
with a correction factor  $f_{corr}$  of the order of 100.  $\text{O}_4$  has not been tested so far. The mathematical formulation has to be seen as an empirical result. Equations (7) and (8) take advantage of the fact that each tangent height can be considered independently at a first approximation. It is a general feature of occultation geometry that absorption structures are mainly produced in an area around the tangent point. Therefore, it is legitimate that Eqs. (7) and (8) have no interdependence between the different tangent heights. In the case of exponentially decreasing profiles like that of the well mixed species, the independence of the heights is even clearer as the extinction shows also the exponential behaviour.

##### 4.2 Results

Tangent height retrievals for SCIAMACHY occultation measurements are performed for two reasons. First, they were needed for monitoring purposes, i.e. the Institute of Environmental Physics was involved in verifications of level-0 and level-1 data products. Secondly, tangent height retrievals might improve even geolocations that are calculated properly from orbital parameters. Figure 4 shows a comparison of tangent heights obtained from orbital parameters and  $\text{O}_2/\text{CO}_2$



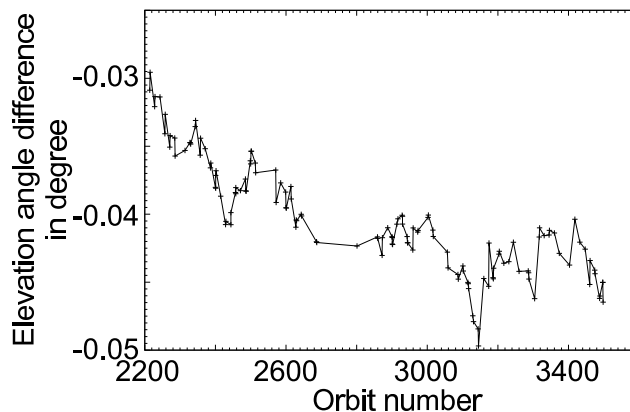
**Fig. 4.** Tangent heights vs. data index number of the occultation scan in orbit 2056 on July 22, 2002. Solid grey line: level-1 tangent heights. Solid black line: O<sub>2</sub> retrieval. Dotted line: CO<sub>2</sub> retrieval.



**Fig. 5.** Black line: Elevation angle of SCIAMACHY's line of sight vs. time in seconds. Grey line: Precalculated elevation angle of the Sun (Noel, personal communication, 2003).

retrievals in solid grey, solid black, and dotted lines, respectively. It can be seen that the algorithm shifted the tangent heights by 2–3 km throughout the measurement sequence. The CO<sub>2</sub> retrieval makes toward the initial tangent height information for higher altitudes. This can be explained by weak absorptions in this region leading to a lower information content. Up to now, similar shifts of the same order of magnitude were observed in all investigated data sets. Details about the underlying O<sub>2</sub>/CO<sub>2</sub> profile retrieval results can be found in Meyer et al. (2004). The tangent height retrieval turned out to be stable for measurements down to the troposphere.

Misalignments of precalculated detector pointing geometry account for the tangent height shift. This can be learned from Fig. 5. It shows a cutout from the movement of SCIAMACHY's elevation mirror during an occultation



**Fig. 6.** Elevation Misalignment vs. orbit number in the time from 1 August to 1 November 2002 (Noel, personal communication, 2003).

measurement sequence (see Sect. 2.2). The time interval was chosen such that the plot focusses on the Sun tracking adjustment in vertical direction. The zigzag line represents the motion of SCIAMACHY's FOV whereas the straight line shows the precalculated position of the Sun in terms of elevation angle. As can be seen, there is a difference of a bit less than 0.05°. The difference can be observed in all occultation measurements that have been investigated so far. Figure 6 shows the observed difference of Fig. 5 for successive orbits in the time between August and November 2002. Details of the time dependence are not understood. However, it can be seen that there is an average misalignment of approximately 0.04°. In terms of SCIAMACHY's instrument geometry, it corresponds to the height of the FOV. It can be transformed into an average height difference at the tangent point by simple geometric considerations. As a result, the observed tangent height shift can be identified with the differences in the elevation angle. Another evidence for a systematic error in the engineering data of ENVISAT is given by von Clarmann et al. (2003). Here, the data evaluation of the MIPAS instrument (Michelson Interferometer for Passive Atmospheric Sounding), which is looking into the opposite direction of SCIAMACHY, shows the same shift in tangent height with opposite sign.

## 5 Fraunhofer lines

### 5.1 Requirements

Since the Sun is not a homogeneous source of light, an algorithm has to be implemented that accounts for varying Fraunhofer structures. In order to obtain reasonable transmission spectra, every measurement has to be assigned to an appropriate reference measurement from above the atmosphere. However, it is almost impossible to address all effects properly that come along with atmospheric measurements. The

following issues have to be taken into account by an ideal matching algorithm:

- Consecutive measurement scans cannot be aligned exactly to the solar disk. Even measurement scans above the atmosphere and without refraction have different pointing characteristics as the elevation rate of the Sun is not known perfectly.
- The fraction of the solar disk covered by the FOV is significantly enlarged for tangent heights below 30 km. Although global intensities are leveled off due to a corresponding geometrical dilution (Dalaudier et al., 2001), it is more difficult to find an appropriate reference spectrum with respect to changing Fraunhofer lines.
- Real occultation measurements underlie spectral shifts as the spectral resolution of SCIAMACHY is too low with respect to narrow Fraunhofer lines.
- Measurements close to the solar limb encounter a different illumination of the FOV. In general, detector signals are a function of spatial illumination. It is questionable if this effect can be covered by a sophisticated assignment to the reference scan.
- The effect of flattening is a function of tangent height, even within the small height range of the FOV. It is still an open question whether it is negligible or not.
- Spectral inhomogeneities of the solar disk are also time dependent. Most effects should be minimised by the fact that there are reference measurements after each occultation event. Nevertheless, changes on a time scale of minutes are still possible as e.g. described by Neckel and Labs (1990).

An algorithm has been developed that covers at least the first two items. An optimal reference measurement is searched from a Sun scan above the atmosphere. Since interpolations between adjacent references are also taken into account, it is possible to find an optimal reference with respect to the exact pointing on the solar disk. Additionally, the algorithm allows for different seeming FOV extensions arising from the Sun flattening. The FOV seems to be extended in height as a larger fraction of the solar disk is covered by measurements at low tangent heights. So far, wavelength shifts are ignored in this context. As will be shown later, it might be useful to integrate the matching algorithm into a shift-and-squeeze algorithm, which is used anyway to overcome deficiencies in the wavelength calibration. The impact of different slit illuminations and Sun flattening within the FOV have not been investigated up to now. Variations of the solar spectrum on a time scale of less than minutes are expected to remain a minor but unsolved problem.

## 5.2 The algorithm

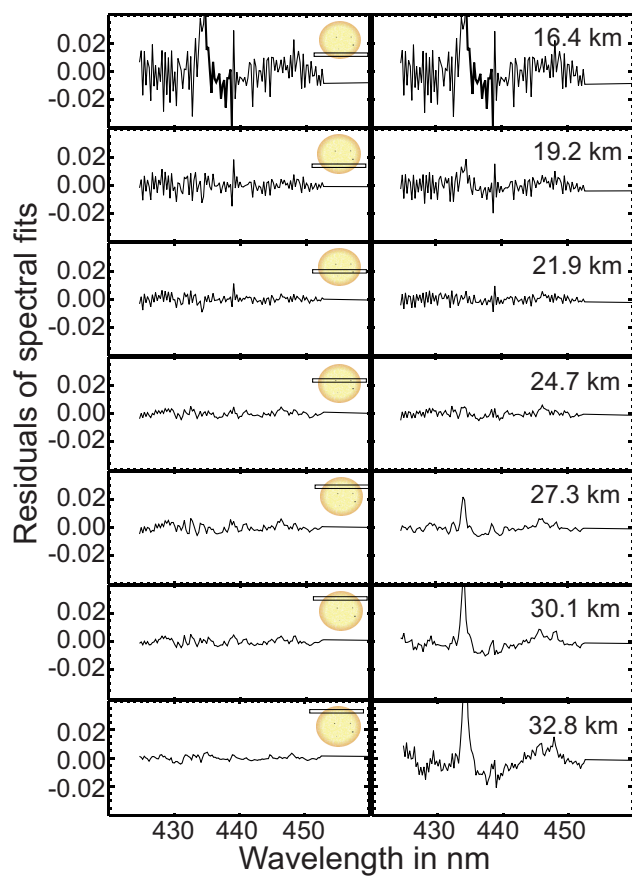
A typical occultation measurement scan consists of 16 or 32 single measurements depending on integration time and data transmission rate between SCIAMACHY and the ground segments (see Fig. 2). First of all, all atmospheric measurements are sorted out which do not match the Sun or have too weak intensities at the solar limb. The remaining data covers a more or less extended region around the centre of the solar disk. As a second step, a range of measurements above 100 km is identified as reference scan. Since all scans have slightly different pointing characteristics, it would be meaningful to use all the information that is available from above the atmosphere. However in reality, the reference scan is reduced to one up or down scan due to computational costs. The following procedure is then repeated for every atmospheric measurement:

1. Spectral windows with strong Fraunhofer lines are chosen.
2. The reference scan is divided into small steps with respect to their elevation angle. Elevation steps between adjacent reference measurements are interpolated.
3. The following procedure is repeated for every elevation step within the reference scan:
  - (a) The spectra within the Fraunhofer windows  $\Upsilon$  are divided by corresponding spectra of the reference scan  $\Upsilon_0$ .
  - (b) A polynomial  $p_n$  of the order 3 is subtracted to reduce the transmission to its differential structure. Fraunhofer lines can now be clearly identified.
  - (c) The absolute values of the differential transmission intensities are integrated over wavelength  $\lambda$ :

$$S = \int_{\Lambda} d\lambda \left| \frac{\Upsilon}{\Upsilon_0} - p_n \right|. \quad (9)$$

The integration sum  $S$  is now considered to be a measure for the strength of the Fraunhofer structure.

- (d) Steps a to c are repeated with different effective heights for the FOV in order to account for the Sun flattening.
4. The elevation step with the smallest integration sum  $S$  over the Fraunhofer windows is now considered to be the optimal reference for this measurement.
5. The whole spectrum needed by the retrieval is divided by the optimal reference, which is often an interpolation between two measurements of the reference scan.



**Fig. 7.** Residuals of spectral retrieval fits for consecutive measurements of an occultation scan. Left: With solar matching algorithm. Right: Without solar matching routine.

### 5.3 Results

Although transmissions derived from appropriate measurements are not expected to contain any structure except for atmospheric extinction signals, Fraunhofer lines are the main contribution to residuals of spectral fits. This section focusses on Fraunhofer structures in the wavelength range between 420 nm and 460 nm, a typical fitting window for  $\text{NO}_2$ . It is particularly interesting due to a strongly decreasing sensitivity towards shorter wavelengths in this region. This can be seen from the signal-to-noise investigations (Bovensmann et al., 1999). Many interesting weak absorbers such as BrO and OClO have a maximum absorption cross section towards the UV wavelength region. The ability to retrieve concentrations of these species depends on the reduction of disturbing spectral structures such as the Fraunhofer lines in the UV-Vis wavelength range. The residual of a spectral retrieval fit is a good indicator for unexpected spectral signals. It reflects all spectral structures that have not been modelled in the retrieval process.

Figure 7 shows residuals of different  $\text{O}_3$  and  $\text{NO}_2$  retrieval fits. They are obtained from a complete up scan over the solar disk. The estimated pointing on the solar disk is indicated by small icons. Two cases are distinguished for each pointing and corresponding tangent height. On the right side, no efforts were made to remove Fraunhofer lines. One extraterrestrial measurement pointing to the solar centre was taken as a reference for all atmospheric measurements. On the left, the matching algorithm described above was applied in advance of the trace gas retrieval. Here, every measurement was assigned to an “optimal” reference from above the atmosphere. At least one strong Fraunhofer line can be identified around 434 nm in this fitting window. It is the  $\text{H}_\gamma$ -line. It can clearly be seen that it has been removed successfully by the matching algorithm on the left side. Looking to smaller tangent heights and the solar centre, no differences can be seen between the residuals.

### 5.4 Other methods

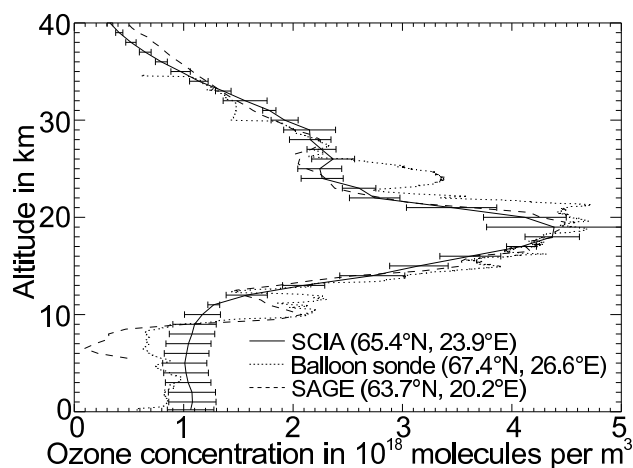
#### 5.4.1 The geometric approach

The matching algorithm introduced above is purely based on spectral properties of the measurements. Of course, a more obvious approach would be the geometric calculation of the pointing. Geometric calculations are performed e.g. in the case of POAM, as described in Glaccum et al. (1996) and in more detail in Lumpe et al. (1997). However, this approach needs reliable information on some geometric quantities such as the elevation angle of the instrument. Geometric calculations have limited accuracy and do not account for spectral issues, which are significant to the quality of the spectral fittings. Nevertheless, it seems to be a very promising task to combine geometric calculations with the matching algorithm.

#### 5.4.2 Spectral fitting of Fraunhofer lines

A third way of addressing the Fraunhofer line problem was developed at the Institute of Environmental Physics in Heidelberg, as described by Bösch (2002). The basic idea is to perform a spectral fit of changing Fraunhofer lines. In Bösch (2002), two high resolution spectra were taken to define the behaviour of Fraunhofer lines towards the solar limb. The first one is an average spectrum of the whole solar disk  $I_0(\lambda)$ , the second one is a measurement of the solar centre  $I_C(\lambda)$ . The true solar spectrum  $I_{true}(\lambda)$  is then replaced by an optimal estimation of both spectra. The spectral fitting is certainly the most elegant way of addressing the Fraunhofer problem. However, spatially and spectrally highly resolved reference spectra are needed in order to obtain significantly better results by this method. Furthermore, the reference spectra should be taken from space without any impact of atmospheric extinction. Otherwise, the problem turns into the opposite, i.e. atmospheric absorption features must





**Fig. 8.** Vertical concentration profiles of ozone as measured by SCIAMACHY (solid line), SAGE II (dashed line), and ozone sonde (dotted line).

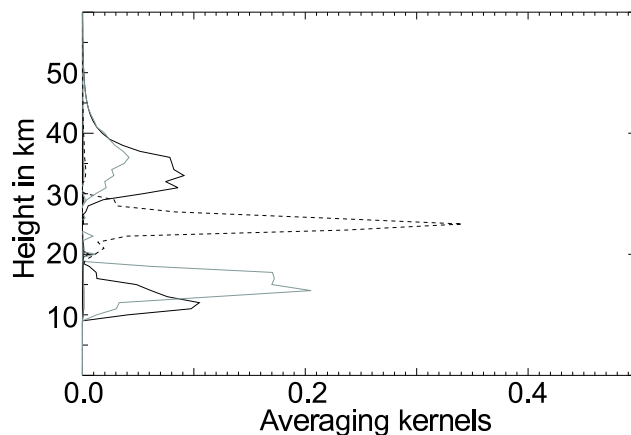
be removed elaborately. Finally, the reference spectra need to be updated regularly due to the time dependent changes of the Fraunhofer lines. So far, such extraterrestrial high resolution spectra do not exist at all.

## 6 Validation of O<sub>3</sub> and NO<sub>2</sub> retrievals

The retrieval algorithm discussed in the previous sections has successfully been applied to O<sub>3</sub>, NO<sub>2</sub>, O<sub>2</sub>, and CO<sub>2</sub> (Meyer et al., 2004). In this section, first validation results for O<sub>3</sub> and NO<sub>2</sub> are presented. Vertical profiles have been compared to corresponding results of the Stratospheric Aerosol and Gas Experiment II (SAGE II) and balloon-borne in-situ measurements. The SAGE II instrument has been extensively validated for ozone and NO<sub>2</sub> profiles (Cunnold et al., 1989, 1991; Wang et al., 2002). SAGE II provides the longest record of satellite high-resolution profile measurements operating since October 1984 onboard the Earth Radiation Budget Satellite (ERBS) measuring solar occultation in the UV-Vis-NIR wavelength range. The recently released SAGE II data version 6.2 (<http://www-sage2.larc.nasa.gov/data/>) is used in this study. The considered measurements were performed between 17 and 22 September 2002. 25 collocated solar occultation measurements of SCIAMACHY and SAGE II were found with a spatial distance between 40 to 250 km. The balloon sonde measurements were performed by Kyrö (2002) over Sodankylä, Finland. Corresponding measurements were carried out on the same day.

### 6.1 Ozone

O<sub>3</sub> is one of the strongest absorbers in the atmosphere. It is even so strong that the Hartley-Huggins absorption bands make the atmosphere opaque for UV radiation. Since the

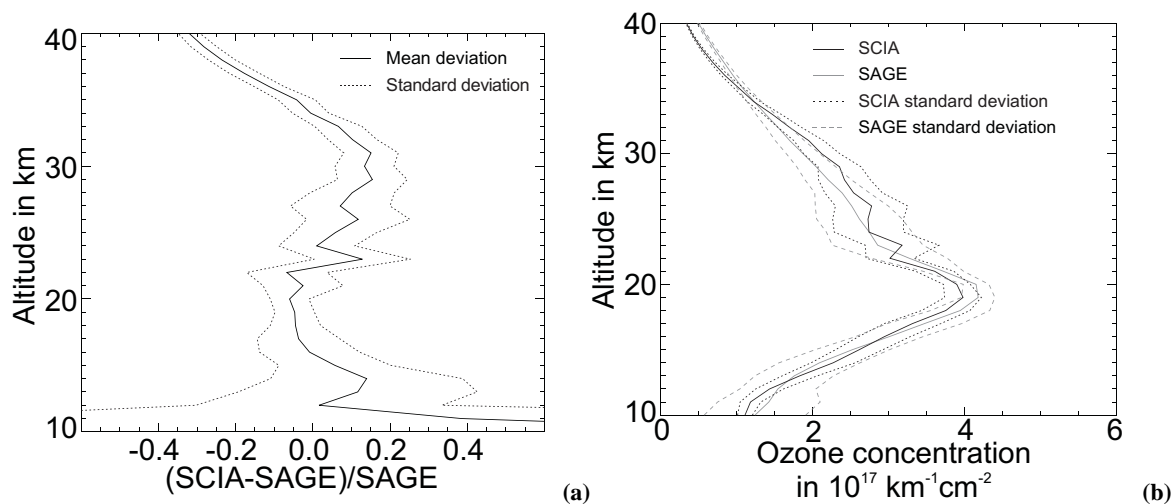


**Fig. 9.** Averaging kernels for ozone retrieval at five selected heights (10, 15, 25, 35, and 40 km) covering the region of interest.

occultation measurement strategy is based on transmitted radiation, it is impossible to choose a spectral fitting window in the UV wavelength range for altitudes up to the middle stratosphere. Only at high altitudes above the ozone layer, it becomes interesting to consider the Hartley-Huggins absorption bands. However, this was beyond the scope of first validation activities. Throughout the studies of this section, the Chappuis bands in the visible wavelength region have been used in the retrieval process. The fitting window ranged from 520 to 595 nm. On the one hand, it includes the strongest regions of the Chappuis bands. On the other hand, it matches also the sensitive range of SCIAMACHY channel 3, which ends at 620 nm.

There is also a clear differential structure in the cross sections, which is very important due to the subtraction of a low order polynomial. In the case of ozone retrievals, a polynomial of the order 3 is fitted to the transmission data. In order to obtain smooth vertical concentration profiles, the regularisation described in Sect. 7 is usually applied. Reasonable regularisation parameters are  $T_1=10^3$  and  $T_2=10^4$ .

In Fig. 8, a retrieved ozone profile is plotted in comparison with retrieval results from SAGE II occultation measurements (dashed line) and balloonborne measurements over Sodankylä (dotted line). Statistical errors obtained by Eq. (14) are indicated by the error bars. The overall agreement is quite good though some of the profile features are not matched by the retrieval results. Since ozone sondes perform direct in-situ measurements, their vertical resolution is quite high. Furthermore, balloons encounter all local variabilities, whereas remote measurements have to be considered as an average over a relatively large airspace. The small ozone peak between 10 and 12 km has not been retrieved from the SCIAMACHY measurements as the considered height range ends here. Therefore the information content was too small to have a significant influence on the result. All values below



**Fig. 10.** Top: Mean deviation between SCIAMACHY and SAGE II ozone concentration profiles. Bottom: Average profiles of collocated measurements.

10 km that are shown here are only based on the a-priori knowledge of the applied climatologies. Similarly, the values above 40 km make towards the a-priori ozone profile as the information content of the corresponding measurements decreases rapidly with height. A reasonable height range for trace gas retrievals can be estimated from the averaging kernels, which are introduced in Sect. 7. In Fig. 9, five selected averaging kernels are plotted as a function of height. The kernels are automatically calculated within the retrieval process. Of course, there are values in the kernel matrix for each height layer. Here, the focus is on the layers at 10, 15, 25, 35, and 40 km as it gives a clear overview of the retrieval characteristics. They are indicated by solid grey and black lines as well as by a dashed black line, respectively. The peak values of the averaging kernels indicate that height layer of the true state with the most significant impact on the retrieved concentration. The peak width is a measure for the vertical resolution of the retrieval result. It can be seen that there are clear peaks throughout the height range between 10 and 40 km. Around the ozone maximum, the peaks of the averaging kernels become more significant, i.e. the theoretical height resolution increases to approximately 3 km and the retrieval results are more strongly related to the true concentration values. Furthermore, it means that the influence of the a-priori profile has a minimum in this region. Towards 40 km, the impact of the measurement on the result begins to decrease again with a corresponding decrease of the height resolution.

In Fig. 10, two plots are shown concerning the statistical evaluation of comparisons between SCIAMACHY and SAGE II ozone retrieval results. The validation is based on 25 collocated measurements. On the top, a height profile of the mean relative deviations of retrieved ozone profiles is

shown. The relative deviation  $RD$  is calculated for each altitude level  $h_i$  by

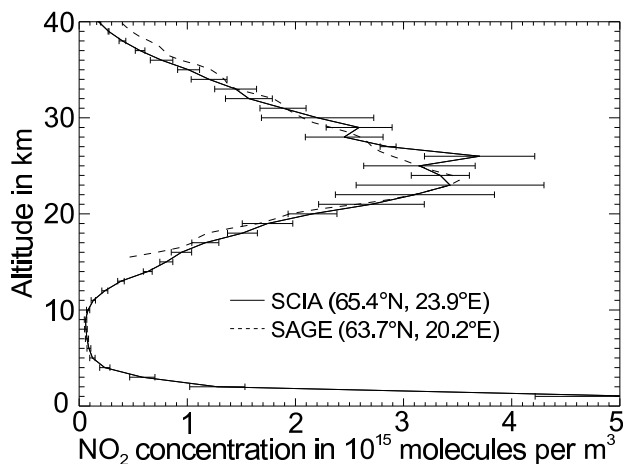
$$RD(h_i) = \frac{SCIA(h_i) - SAGE(h_i)}{SAGE(h_i)}. \quad (10)$$

Trace gas concentrations are indicated by the abbreviated instrument names. After that, the mean relative deviation and the corresponding root mean square (RMS) are calculated for each height layer. On the bottom, the average profiles of both instruments are compared directly. The mean relative deviation is between  $-2$  to  $15\%$  between 15 and 35 km altitude, which is a quite promising result. There is a trend towards positive differences with increasing height that cannot be explained at the moment.

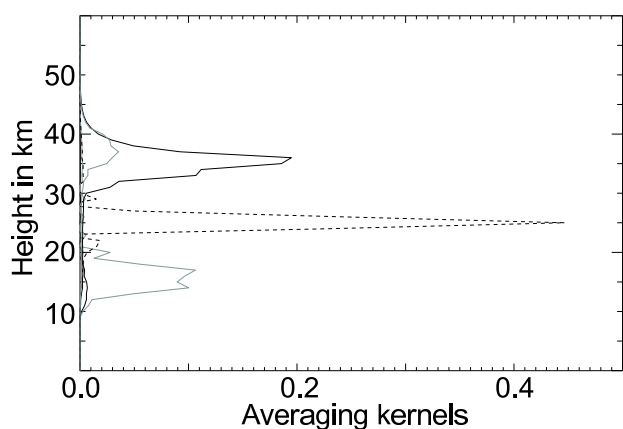
## 6.2 Nitrogen dioxide

Like ozone,  $\text{NO}_2$  has good absorption properties in the UV and visible wavelength range. However due to the much smaller abundance of  $\text{NO}_2$  compared to ozone, it is useful to focus strictly on the region of maximum absorption cross sections. The spectral fitting window used in the  $\text{NO}_2$  studies ranges from 420 to 460 nm. It matches also the sensitive detector region of channel 3, which starts at 394 nm. Furthermore, the spectra are well supplied with differential structures of  $\text{NO}_2$ . A limiting factor is a rapidly decreasing signal-to-noise ratio towards the UV range due to the strong ozone absorption.

It turned out that a simultaneous fit of ozone and  $\text{NO}_2$  is quite suitable for the spectral properties of both species. On the one hand of course, it is comfortable to retrieve both at the same time for standard data products. On the other hand, the spectral signals of both gases interfere significantly in the  $\text{NO}_2$  fitting window. The information content of  $\text{NO}_2$  in the



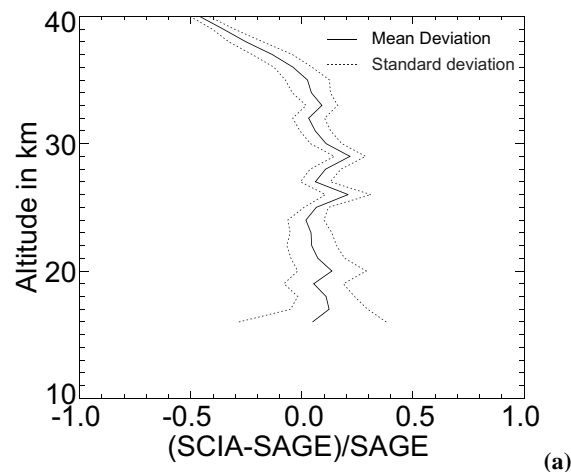
**Fig. 11.** Vertical concentration profiles of  $\text{NO}_2$  as measured by SCIAMACHY and SAGE II.



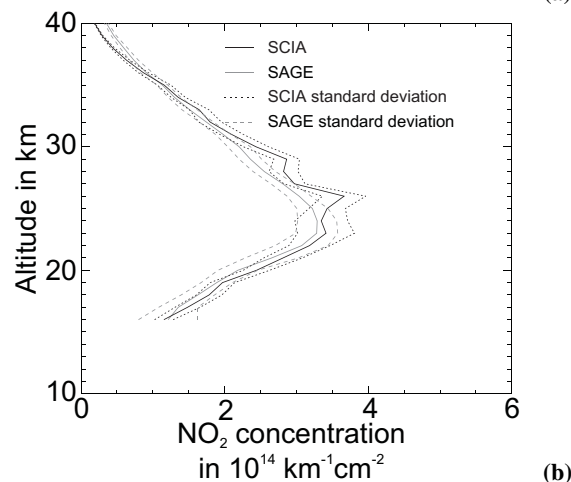
**Fig. 12.** Averaging kernels for  $\text{NO}_2$  at five selected heights (10, 15, 25, 35, and 40 km) covering the region of interest.

measurements is smaller than that of ozone. Therefore, it is necessary to apply different regularisation parameters. So far, best results were obtained with  $T_1=50$  and  $T_2=50$  (see Sect. 7). The subtraction polynomial is the same with an order of 3.

Figure 11 shows the results of the  $\text{NO}_2$  vertical profile retrieval from the same retrieval run as in Sect. 6.1. Again, it is compared to corresponding results from the SAGE II instrument. The agreement is fairly good. Figure 12 shows the  $\text{NO}_2$  averaging kernels obtained within the retrieval. Like in the case of the  $\text{O}_3$  kernels, the plot is reduced to the height layers at 10, 15, 25, 35, and 40 km. The results are quite similar to the  $\text{O}_3$  averaging kernels. An important difference can be observed at 10 km, where the impact of the true profile has almost vanished. Standard retrievals of  $\text{NO}_2$  are possible in the range between 13 and 40 km. The maximum information content of the measurements with respect to  $\text{NO}_2$  is between 20 and 30 km. However, there have already been successful



(a)



(b)

**Fig. 13.** Top: Mean deviation between SCIAMACHY and SAGE II  $\text{NO}_2$  concentration profiles. Bottom: Average profiles of considered collocated measurements.

retrieval fittings at tangent heights below 13 km. It is an indication that  $\text{NO}_2$  measurements are possible even down to the troposphere.

Figure 13 illustrates the statistical results of all comparisons with the results from SAGE II according to Eq. (10) and the calculation of the mean relative deviation. Between 16 to 35 km, SCIAMACHY shows a slight positive bias with deviations between 2 and 20%. Errors are of the order of 15% with a relatively large bias. Nevertheless, the results are quite promising for a first validation.

### 6.3 Error budget

Table 2 gives an overview of different error sources. The errors are supposed to be the most important ones with respect to retrieval precisions and quality. In some cases, the impact on the retrieval results can only be estimated qualitatively as detailed studies are missing. Some errors have been

**Table 2.** Error budget.

Error description	Value
Deviations obtained from validation results	
O <sub>3</sub>	~10%
NO <sub>2</sub>	~15%
Errors due to instrument noise	0.025–1%
Statistical errors (optimal estimation)	
O <sub>3</sub>	~5%
NO <sub>2</sub>	≤10%
Horizontal Inhomogeneities due to	
photochemical dependencies (NO <sub>2</sub> )	10–40%
dynamic processes	0–10%
Uncertainties of cross-sections	≤3%
Diffuse radiation	≪0.1%
Increased aerosol loading (enhanced diffuse light)	≪1%
Neglect of FOV height extension	<2%
Inhomogeneous illumination of detector	~1%
Wavelength shifts	~1%
Undersampling	~1%
Model errors obtained from spectral residuals	
with matching algorithm	~1%
without matching algorithm	1–6%
Sun spots	0–10%
Global limb darkening (due to pointing errors)	~7%
Geometry errors	
Tangent height uncertainty	
with height retrieval	<1 km
without height retrieval	1–3 km
Height resolution	3–5 km
Maximum pointing error	0.012°

estimated in various studies that have not been published so far. They are discussed in Meyer (2004). All estimations refer to the standard retrievals of O<sub>3</sub> and NO<sub>2</sub> described in this presentation.

Concerning the uncertainties of the retrieved profiles, the deviations obtained from the validation results are the most meaningful. The deviations show a small statistical behaviour and some bias, which is height dependent in the case of ozone between –5% to 15% and around 10% for NO<sub>2</sub>. The statistical deviations of the fast varying structures are less than 10%.

The statistical errors obtained by the optimal estimation are of the same order but significantly smaller as the theoretical precisions only account for the signal-to-noise ratio, the information content of the measurements, and the uncertainty of the a-priori. Another important error source with mainly statistical behaviour are dynamically induced inhomogeneities of the atmosphere. At the worst, they are around 10%, which is of the same order as the statistical errors from the optimal estimation. The error propagation according to gaussian statistics predicts an overall statistical error of around 14% in this case.

Most of the errors given here are bias errors. As can be seen in Table 2, they are negligible with the exception of the photochemically induced inhomogeneities. Furthermore, many errors will not be relevant if a polynomial is subtracted from the measurement data. Although many errors cannot be compared directly, they seem to describe consistently the limitations of the standard O<sub>3</sub> and NO<sub>2</sub> profile retrievals.

## 7 Conclusions

An algorithm for the evaluation of SCIAMACHY solar occultation measurements has been presented. It is mainly dedicated to the retrieval of height resolved trace gas concentration profiles such as ozone and NO<sub>2</sub>. The retrieval is based on the optimal estimation method. The forward model is a further development of the radiative transfer model GOME-TRAN.

Pre-calculation of tangent heights turned out to be not sufficient for good retrieval results. An algorithm has been implemented to correct the tangent heights. It is based on the retrieval of oxygen, which is a well mixed gas in the atmosphere. The results are used to adjust the measurement heights.

Another important task for solar occultation measurements is the exact pointing on the solar disk. Transmissions, which are obtained from divisions by extraterrestrial reference measurements, are expected to contain only atmospheric extinction signals. Nevertheless, spatially varying Fraunhofer lines remain part of the differential transmission spectra as the shape and the depth of Fraunhofer lines are a function of pointing on the solar disk. Retrieval of weak absorbers can be seriously affected due to their relatively small spectral signals. An algorithm has been presented that reduces these structures. It is shown that the impact of Fraunhofer lines will be minimised if atmospheric measurements are assigned to appropriate extraterrestrial reference measurements.

First validation results of ozone and NO<sub>2</sub> profiles have been presented. The agreement with results from SAGE II and POAM III is already fairly good. Height resolved concentration profiles of these species can be considered as standard product, which can now be provided for further use in the community. A more detailed error analysis will certainly help to optimise the retrieval of trace gas profiles as the theoretical potentials of SCIAMACHY are still not tapped. The scientific evaluation of solar occultation data with respect to weak absorbers such as BrO and OClO, infrared absorbing gases such as CH<sub>4</sub> and CO, and the evaluation of tropospheric as well as mesospheric measurements is still in the beginning.

## Appendix

### The optimal estimation method

The optimal estimation method was first described in the context of remote sensing by Rodgers (1976, 1990). The important results are presented here. Starting point is the linearised forward model in Eq. (6). Accuracies of a-priori knowledge and measurement errors have to be taken into account as well, to solve the inverse problem. They are introduced in form of covariance matrices  $\mathbf{S}_a$  and  $\mathbf{S}_y$ , respectively. Minimisation of

$$\|(\mathbf{y} - \mathbf{y}_0) - \mathbf{K}_0(\mathbf{x} - \mathbf{x}_0)\|_{\mathbf{S}_y^{-1}}^2 + \|\mathbf{x} - \mathbf{x}_0\|_{\mathbf{S}_a^{-1}}^2 \quad (11)$$

leads then to an optimal estimation for the state vector  $\mathbf{x}$ :

$$\mathbf{x} = \mathbf{x}_0 + (\mathbf{K}_0^T \mathbf{S}_y^{-1} \mathbf{K}_0 + \mathbf{S}_a^{-1})^{-1} \mathbf{K}_0^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{y}_0). \quad (12)$$

Since the solution is based on a linear approximation, it might be necessary to repeat the calculation of  $\mathbf{x}$  iteratively. Iteration step  $i + 1$  is given by

$$\mathbf{x}_{i+1} = \mathbf{x}_0 + (\mathbf{K}_i^T \mathbf{S}_y^{-1} \mathbf{K}_i + \mathbf{S}_a^{-1})^{-1} \mathbf{K}_i^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{y}_i + \mathbf{K}_i(\mathbf{x}_i - \mathbf{x}_0)). \quad (13)$$

After convergence,  $\hat{\mathbf{x}}$  is the final retrieval result. The theoretical precisions  $s_j$  of the result are represented by the diagonal elements of covariance matrix  $\hat{\mathbf{S}}$ :

$$\hat{\mathbf{S}} = (\hat{\mathbf{K}}^T \mathbf{S}_y^{-1} \hat{\mathbf{K}} + \mathbf{S}_a^{-1})^{-1}. \quad (14)$$

With  $\sigma_j$  denoting the square root of the  $j$ -th diagonal element of  $\hat{\mathbf{S}}$  and  $x_0^j$  as the a-priori value of the  $j$ -th parameter to be retrieved, precisions are obtained by

$$s_j = \frac{\sigma_j}{x_0^j}. \quad (15)$$

Another frequently used quantity is the so called averaging kernel matrix  $\hat{\mathbf{A}}_k$ . It is defined as

$$\hat{\mathbf{A}}_k = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}_t}, \quad (16)$$

where  $\mathbf{x}_t$  is the unknown true state vector. The averaging kernels are a measure of the impact of  $\mathbf{x}_t$  on the retrieved state  $\hat{\mathbf{x}}$ . Each row of the kernel matrix represents one height layer of the retrieved state, the columns indicate the influence of each height layer of the true state.

### Application to occultation measurements

For spectroscopic measurements of atmospheric constituents in occultation geometry, the retrieval is based on the linearised iterative approach of Eq. (13). Elements of the state

vectors  $\mathbf{x}_0$  and  $\mathbf{x}$  are defined as concentrations  $N_j$  for each height layer  $h_i$  and each considered trace gas  $j$ :

$$\mathbf{x}_0 = \begin{pmatrix} N_1(h_1) \\ \vdots \\ N_1(h_L) \\ \vdots \\ N_J(h_1) \\ \vdots \\ N_J(h_L) \end{pmatrix} \quad \text{and} \quad (17)$$

$$\mathbf{x} = \begin{pmatrix} N_1(h_1) + \delta N_1(h_1) \\ \vdots \\ N_1(h_L) + \delta N_1(h_L) \\ \vdots \\ N_J(h_1) + \delta N_J(h_1) \\ \vdots \\ N_J(h_L) + \delta N_J(h_L) \end{pmatrix} \quad (18)$$

with  $J$  as the total number of considered trace gases and  $L$  the total number of height layers. For computational reasons, a parameter vector  $\mathbf{p}$  is defined containing the relative deviations from the a-priori concentrations  $N_j(h_i)$ :

$$\mathbf{p} = \begin{pmatrix} \delta N_1(h_1)/N_1(h_1) \\ \vdots \\ \delta N_1(h_L)/N_1(h_L) \\ \vdots \\ \delta N_J(h_1)/N_J(h_1) \\ \vdots \\ \delta N_J(h_L)/N_J(h_L) \end{pmatrix}. \quad (19)$$

The measurement vectors  $\mathbf{y}$  and  $\mathbf{y}_0$  are conceived in the same way:

$$\mathbf{y}_0 = \begin{pmatrix} \Upsilon(\psi_1, \lambda_1) \\ \vdots \\ \Upsilon(\psi_1, \lambda_\Lambda) \\ \vdots \\ \Upsilon(\psi_K, \lambda_1) \\ \vdots \\ \Upsilon(\psi_K, \lambda_\Lambda) \end{pmatrix}, \quad (20)$$

$$\mathbf{y} = \begin{pmatrix} \Upsilon(\psi_1, \lambda_1) + \delta \Upsilon(\psi_1, \lambda_1) \\ \vdots \\ \Upsilon(\psi_1, \lambda_\Lambda) + \delta \Upsilon(\psi_1, \lambda_\Lambda) \\ \vdots \\ \Upsilon(\psi_K, \lambda_1) + \delta \Upsilon(\psi_K, \lambda_1) \\ \vdots \\ \Upsilon(\psi_K, \lambda_\Lambda) + \delta \Upsilon(\psi_K, \lambda_\Lambda) \end{pmatrix}. \quad (21)$$

$\Lambda$  is the total number of wavelengths  $\lambda$  and  $K$  is the total number of tangent heights or solar zenith angles  $\psi$ . Intensity  $\Upsilon$  is given by Eq. (3). Throughout this presentation, only transmissions are used in the retrieval algorithm. In most cases, weakly wavelength dependent features such as Rayleigh scattering are removed by subtraction of a low order polynomial. Additionally, global variations of the transmission due to a non-perfect pointing on the solar disk are removed by the polynomial subtraction.

Formally, the intensity variation  $\delta\Upsilon$  can be related to the variation of atmospheric parameters  $\delta p_j$  by

$$\delta\Upsilon(\psi, \lambda_i, \mathbf{p}) = \sum_{j=1}^J \int_{h_b}^{h_t} W_j(\psi, \lambda, \bar{\mathbf{p}}) \delta p_j(h') dh', \quad (22)$$

where  $\bar{\mathbf{p}}$  denotes the parameter vector under a-priori conditions. The integration over height  $h$  is performed from the bottom  $h_b$  to the top  $h_t$  of atmosphere.  $W_j$  are the so called weighting functions. They correspond to the elements of matrix  $\mathbf{K}_0$  in Eq. (6) describing the intensity variation  $\delta\Upsilon$  with respect to variations of atmospheric parameters  $\delta p_j$ :

$$W_j(\psi, \lambda, \bar{\mathbf{p}}) = \left. \frac{\delta\Upsilon(\mathbf{p})}{\delta p_j} \right|_{\bar{\mathbf{p}}}. \quad (23)$$

For trace gas retrieval, the weighting functions  $W_j$  are the derivatives of measured intensity  $\Upsilon$  with respect to relative trace gas concentrations  $p_j$ . From Eqs. (2), (3), and (4) it can be seen how  $\delta\Upsilon$  is obtained. The derivation of Eq. (3) with respect to trace gas number densities  $N_j$  leads to

$$\delta\Upsilon(\psi, \lambda, \beta) = - \int_{h_b}^{h_t} dh' \delta N_j \int_{\Omega} d\omega \frac{f(\omega)}{\cos \tilde{\psi}(h', \omega)} \chi(\psi, h', \omega) \int_{\Delta\lambda} d\lambda' a(\lambda, \lambda') I_0(\lambda', \beta) e^{-\tau_j(\psi, \omega, \lambda')} \sigma_j(\lambda'), \quad (24)$$

changing the order of integration and pulling down the integration over the path length. Comparing Eqs. (22) and (24), the following expression is obtained for the weighting function of trace gas  $j$ :

$$W_j(\psi, \lambda, \mathbf{p}) = - \int_{\Omega} d\omega \frac{f(\omega) \chi(\psi, h', \omega)}{\cos \psi(h', \omega)} \int_{\Delta\lambda} d\lambda' a(\lambda, \lambda') I_0(\lambda') e^{-\tau_j(\psi, \omega, \lambda')} \sigma_j(\lambda', h'). \quad (25)$$

For numerical reasons, it is necessary to discretise variable  $h'$  in the Equations above. Hence, Eq. (22) becomes

$$\delta\Upsilon(\psi, \lambda, \bar{\mathbf{p}}) = \sum_{j=1}^J \sum_{i=1}^L \tilde{W}_{ji}(\psi, \lambda, \bar{\mathbf{p}}) \frac{\delta N_j(h_i)}{N_j(h_i)} q_i \quad (26)$$

with  $q_i$  as discretisation of  $\delta h'$ , which depends on the integration rule. Note, that  $\tilde{W}_{ji}$  is now appropriate to relative variations of parameters  $\mathbf{p}$ .

The linearised forward model operator  $\mathbf{K}_0$  from Eq. (6) can now be written for occultation trace gas retrieval:

$$\mathbf{K}_0 = \begin{pmatrix} \tilde{W}_{11}(\psi_1, \lambda_1) q_1 & \dots & \tilde{W}_{JL}(\psi_1, \lambda_1) q_L \\ \vdots & & \vdots \\ \tilde{W}_{11}(\psi_1, \lambda_\Lambda) q_1 & \dots & \tilde{W}_{JL}(\psi_1, \lambda_\Lambda) q_L \\ \vdots & & \vdots \\ \tilde{W}_{11}(\psi_K, \lambda_1) q_1 & \dots & \tilde{W}_{JL}(\psi_K, \lambda_1) q_L \\ \vdots & & \vdots \\ \tilde{W}_{11}(\psi_K, \lambda_\Lambda) q_1 & \dots & \tilde{W}_{JL}(\psi_K, \lambda_\Lambda) q_L \end{pmatrix} \quad (27)$$

Measurement errors are taken into account by covariance matrix  $\mathbf{S}_y$  representing the measurement noise for each measurement and wavelength. Usually, it is a diagonal matrix, i.e. the errors are assumed to be uncorrelated. The a-priori covariance matrix  $\mathbf{S}_a$  reflects the errors on the a-priori height profiles. It is a block diagonal matrix, where correlations are assumed to be found only between adjacent height layers of the same species. The off-diagonal elements within each block are approximated by

$$\mathbf{S}_{a,j} = \sigma_{j,i'} \sigma_{j,i''} e^{-\frac{|h_{i'} - h_{i''}|}{h_c}}. \quad (28)$$

Index  $j$  denotes trace gas number or number of diagonal block,  $h_c$  is the correlation length, and  $\sigma_{j,i'}/i''$  are the variances of trace gas concentrations at height levels  $h_{i'}$  and  $h_{i''}$ , respectively. A reasonable value for  $h_c$  is for example 5 km in the case of ozone (Hoogen et al., 1999).

### Regularisation

Real retrieval solutions of atmospheric parameters are often characterised by strong variabilities, which cannot be explained by real atmospheric features. A ‘first order’ regularisation is automatically involved by including the a-priori information, as described in Sect. 7. The relative strength of the regularisation is weighted by the covariance matrices  $\mathbf{S}_a$  and  $\mathbf{S}_y$ . However, even if the matrices contain reasonable estimations of errors and uncertainties, instable solutions will not be avoided in any event.

The retrieval algorithms provide more sophisticated constraints on the solution. They are based on considerations that were first published by Tikhonov (1963) and Twomey (1963). The method is known as Twomey-Tikhonov regularisation. The aim is to modify the inverse of the a-priori covariance matrix  $\mathbf{S}_a$  such that the estimated solution is forced into a more ‘realistic’ shape. Of course, the constraints should be as weak as possible and as strong as necessary. Otherwise, the solution will suffer from unrealistic features or will lose measurement information, respectively.

Looking at Eqs. (11) and (12), it can be shown that the diagonal elements of  $\mathbf{S}_a^{-1}$  are a direct weighting of the a-priori parameter values. The off-diagonal elements can be used to take also first and higher order derivatives of the a-priori values into account. A nice introduction to this idea is also given by Rodgers (1999). In the case of the occultation retrieval algorithms, two ‘‘Tikhonov parameters’’,  $T_1^j$  and  $T_2^j$ , are used for each trace gas  $j$ . The first one enhances the direct influence of the a-priori parameters replacing the diagonal elements of  $\mathbf{S}_a^{-1}$  by the chosen parameter  $T_1^j$ . The second parameter is intended for the first derivatives of the a-priori parameter values, i.e. the smoothness of the estimated profiles. A matrix  $\mathbf{R}_j$ , which accounts for the deviation of the a-priori state vector  $\mathbf{x}_0$  with respect to height  $h_i$ , can be written in the following form:

$$\mathbf{R}_j = \begin{pmatrix} \frac{\sqrt{T_2^j}}{h_1-h_2} & \frac{\sqrt{T_2^j}}{h_2-h_1} & 0 & \dots & \dots & \dots \\ \frac{\sqrt{T_2^j}}{h_1-h_3} & 0 & \frac{\sqrt{T_2^j}}{h_3-h_1} & \dots & \dots & \dots \\ 0 & \frac{\sqrt{T_2^j}}{h_2-h_4} & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & \frac{\sqrt{T_2^j}}{h_L-h_{L-2}} & \dots \\ \dots & \dots & \dots & \dots & \frac{\sqrt{T_2^j}}{h_L-h_{L-1}} & \frac{\sqrt{T_2^j}}{h_L-h_{L-1}} \end{pmatrix}. \quad (29)$$

$\mathbf{R}_j$  has to be defined for each trace gas  $j$ . The matrices  $\mathbf{R}_j$  are comprised in an overall block-diagonal  $\mathbf{R}$ -matrix:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & & \mathbf{0} \\ & \dots & \\ \mathbf{0} & & \mathbf{R}_J \end{pmatrix}. \quad (30)$$

Equation (11) can now be rewritten with a third term that has to be minimised:

$$\|(\mathbf{y}-\mathbf{y}_0)-\mathbf{K}_0(\mathbf{x}-\mathbf{x}_0)\|_{\mathbf{S}_a^{-1}}^2 + \|\mathbf{x}-\mathbf{x}_0\|_{\mathbf{S}_a^{-1}}^2 + \|\mathbf{x}-\mathbf{x}_0\|_{\mathbf{R}^T\mathbf{R}}^2. \quad (31)$$

Minimisation of Eq. (31) leads to a new optimal estimation of state vector  $\mathbf{x}$ . It can be written in the same formulation as Eq. (12), but with a sophisticated revision of the inverse a-priori covariance matrix  $\mathbf{S}_a^{-1}$ . The impact of the mean squared difference and the mean squared second difference on the result state vector is now completely ruled by the two Tikhonov parameters,  $T_1^j$  and  $T_2^j$ .

*Acknowledgements.* We would like to thank ESA/ESRIN for providing level-0 and level-1 spectral data from SCIAMACHY. We thank the SAGE II group (at NASA LaRC, especially L. Thomason and the NASA Langley Radiation and Aerosols Branch) and E. Kyrö (FMI-Arctic Research Centre, Sodankylä, Finland) for providing us with vertical trace gas profiles of  $\text{O}_3$  and  $\text{NO}_2$ . This work has been funded in part by the German Federal Ministry of Education and Research (BMBF) via the AFO2000 project (no. 07ATF52), the University of Bremen, and the state of Bremen.

Edited by: U. Pöschl

## References

- Ahlers, B. and Dobber, M.: OPTEC5: Slit function calibration of the SCIAMACHY PFM, Tech. rep., TNO TPD, Netherlands, 2000.
- Alvarez, L. J., Lopez, E., Sanchez, M., and Gonzales, J. A.: ENVISAT-1 Mission CFI Software – General Software User Manual, vers. 4.6, provided by GMV/ESA, 2000.
- Anderson, G. P., Clough, S. A., Kneizys, F. X., Chetwynd, J. H., and Shettle, E. P.: AFGL Atmospheric Constituent Profiles (0–120 km), Tech. Rep. AFGL-TR-86-0110, Phillips Laboratory, AD A 175173, 1986.
- Bates, D. R.: Rayleigh scattering by air, *Planet. Space Sci.*, 32, 785–790, 1984.
- Bogumil, K., Orphal, J., and Burrows, J. P.: Reference Spectra of Atmospheric relevant trace gases measured with the SCIAMACHY PFM satellite spectrometer, in *URF Sciences Exactes et Naturelle*, BP 1039, 51687 Reims Cedex, France, 1999.
- Bösch, H.: Studies of the Stratospheric Nitrogen and Iodine Chemistry by Balloon-Borne DOAS Measurements and Model Calculations, Ph.D. thesis, Universität Heidelberg, 2002.
- Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, *J. Atmos. Sci.*, 56, 127–150, 1999.
- Brühl, C., Drayson, S. R., Russell, J. M. I., Crutzen, P. J., McInerney, J. M., Purcell, P. N., Claude, H., Gernandt, H., McGee, T. J., McDermid, T. S., and Gunson, M. R.: Halogen Occultation Experiment Ozone Channel Validation, *J. Geophys. Res.*, 101, 10 217–10 240, 1996.
- Burrows, J. P., Dehn, A., Deters, B., Himmelmann, S., Richter, A., Voigt, S., and Orphal, J.: Atmospheric Remote-Sensing Reference Data from GOME: 1. Temperature-Dependent Absorption Cross Sections of  $\text{NO}_2$  in the 231–794 nm Range, *J. Quant. Spectrosc. Radiat. Transfer*, 60, 1025–1031, 1998.
- Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weissenmayer, A., Richter, A., deBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., and Eisinger, M.: The Global Ozone Monitoring Experiment (GOME): Mission Concept and First Scientific Results, *J. Atmos. Sci.*, 56, 151–175, 1999.
- Cunnold, D. M., Chu, W. P., Barnes, R. A., McCormick, M. P., and Veiga, R. E.: Validation of SAGE II Ozone Measurements, *J. Geophys. Res.*, 94, 8447–8460, 1989.
- Cunnold, D. M., Zawodny, J. M., Chu, W. P., Pommereau, J. P., Goutail, F., Lenoble, J., McCormick, M. P., Veiga, R. E., Murracay, D., Iwagami, N., Shibasaki, K., Simon, P. C., and Peetermans, W.: Validation of SAGE II  $\text{NO}_2$  Measurements, *J. Geophys. Res.*, 96, 12 913–12 925, 1991.
- Dalaudier, F., Kan, V., and Gurvich, A. S.: Chromatic refraction with global ozone monitoring by occultation of stars, I. Description and scintillation correction, *Appl. Opt.*, 40, 866–877, 2001.
- Farman, J. C., Gardner, B. G., and Franklin, J. D.: Large losses of total ozone in Antarctica reveal seasonal  $\text{ClO}_x$  and  $\text{NO}_x$  interaction, *Nature*, 315, 207–210, 1985.
- Glaccum, W., Lucke, R. L., Bevilacqua, R. M., Shettle, E. P., Hornstein, J. S., Chen, D. T., Lumpe, J. D., Krigman, S. S., Debrestian, D. J., Fromm, M. D., Dalaudier, F., Chassefiere, E., Deniel, C., Randall, C. E., Rusch, D. W., Olivero, J. J., Brogniez, C., Lenoble, J., and Kremer, R.: The Polar Ozone and Aerosol Measurement instrument, *J. Geophys. Res.*, 101, 14 479–14 487,

- 1996.
- Hoogen, R., Rozanov, V., and Burrows, J. P.: Ozone profiles from GOME satellite data: Algorithm description and first validation, *J. Geophys. Res.*, 104, 8263–8280, 1999.
- Kaiser, J.: Atmospheric Parameter Retrieval from UV-Visible-NIR Limb Scattering Measurements, Ph.D. thesis, Universität Bremen, ISBN 3-89722-830-3, Logos Verlag Berlin, 2001.
- Kneizys, F. X., Shettle, E. P., Abreu, L. W., Chetwynd, J. J., Anderson, G. P., Gallery, W. O., Selby, J. E. A., and Clough, S. A.: Users guide to LOWTRAN 7, Tech. Rep. AFGL-TR-88-0177, Air Force Geophysical Laboratory, Bedford, Massachusetts, 1988.
- Kyrö, E.: Balloon sonde measurements over Sodankylä, published in the NADIR database at the Norwegian Institute for Air Research (NILU), 2002.
- Lucke, R. L., Korwan, D. R., Bevilacqua, R. M., Hornstein, J. S., Shettle, E. P., Chen, D. T., Daehler, M., Lumpe, J. D., Fromm, M. D., Debrestian, D., Neff, B., Squire, M., König-Langlo, G., and Davies, J.: The Polar Ozone and Aerosol Measurement (POAM) III instrument and early validation results, *J. Geophys. Res.*, 104, 18 785–18 799, 1999.
- Lumpe, J. D., Bevilacqua, R. M., Hoppel, K. W., Krigman, S., Kriebel, D., Debrestian, D. J., Randall, C. E., Rusch, D., Brogniez, C., Ramanaherisoa, R., Shettle, E. P., Olivero, J. J., Lenoble, J., and Pruvost, P.: POAM II retrieval algorithm and error analysis, *J. Geophys. Res.*, 102, 23 593–23 614, 1997.
- Mauldin, L. E., Zaun, N. H., McCormick, M. P., Guy, J. J., and Vaughn, W. R.: Stratospheric Aerosol and Gas Experiment II Instrument: A Functional Description, *Opt. Eng.*, 24–307, 1985.
- McCormick, M. P., Hamill, P., Pepin, T. J., Chu, W. P., Swissler, T. J., and McMaster, L. R.: Satellite studies of the stratospheric aerosol, *B. Am. Meteor.*, 60, 1038–1046, 1979.
- Meyer, J.: Solar Occultation Measurements with SCIAMACHY in the UV-Vis-IR Wavelength Region, Ph.D. thesis, Universität Bremen, 2004.
- Meyer, J., Schlesier, A. C., Rozanov, A., Bovensmann, H., and Burrows, J. P.: Towards O<sub>3</sub> and NO<sub>2</sub> vertical profile retrieval from SCIAMACHY solar occultation measurements: First results, *Adv. Space Res.*, 34, 744–748, 2004.
- NASA: U.S. Standard Atmosphere Supplements, Tech. rep., U.S. Government Printing, Washington, D.C., 1976.
- Neckel, H. and Labs, D.: Variations of ‘wavelengths’ and ‘bisector indices’ of 70 solar spectral lines between 33 000 and 3960 Å in Kitt Peak FTS spectra, *Solar Phys.*, 126, 207–266, 1990.
- Rodgers, C. D.: Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation, *Rev. Geophys. Space Phys.*, 4, 609–624, 1976.
- Rodgers, C. D.: Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, 5587–5595, 1990.
- Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practise, World Scientific, Singapore, 1999.
- Rothman, L. S., Barbe, A., Chris Benner, D., et al.: The HITRAN molecular spectroscopic database and HAWKS (HITRAN atmospheric workstation), *J. Quant. Spectrosc. Radiat. Transfer*, 60, 665–710, 1998.
- Rozanov, A.: Modeling of the radiative transfer through a spherical planetary atmosphere: Application to the atmospheric trace gases retrieval from occultation- and limb-measurements in UV-Vis-NIR, Ph.D. thesis, Universität Bremen, ISBN 3-8325-0138-X, Logos Verlag Berlin, 2001.
- Rozanov, V. V., Diebel, D., Spurr, R. J. D., and Burrows, J. P.: GOMETRAN: A radiative transfer model for the satellite project GOME, the plane-parallel version, *J. Geophys. Res.*, 102, 16 683–16 695, 1997.
- Russel, J. M., Gordley, L. L., Park, J. H., Drayson, S. R., Hesketh, W. D., Cicerone, R. J., Tuck, A. F., Frederick, J. E., Harries, J. E., and Crutzen, P. J.: The Halogen Occultation Experiment, *J. Geophys. Res.*, 98, 10 777–10 797, 1993.
- Tikhonov, A.-N.: On the solution of incorrectly stated problems and a method of regularization, *Dokl. Acad. N.*, 151, 1963.
- Twomey, S.: On the numerical solution of Fredholm integral equation of the first kind by the inversion of the linear system produced by quadrature, *J. Ass. Comput. Mach.*, 10, 1963.
- van Rie, G. J. R. M. J.: Solar calibration of Earth Observation instruments in general and applied to SCIAMACHY, Tech. rep., Delft University of Technology, Netherlands, 1995.
- von Clarmann, T., Ceccherini, S., Doicu, A., Dudhia, A., Funke, B., Grabrowski, U., Hilgers, S., Jay, V., Linden, A., Lopez-Puertas, M., Martin-Torres, F., Payne, V., Reburn, J., Ridolfi, M., Schreier, F., Schwarz, G., Siddans, R., and Steck, T.: A blind test retrieval experiment for infrared limb emission spectrometry, *J. Geophys. Res.*, 108, doi: 10.1029/2003JD003835, 2003.
- Wang, H. J., Cunnold, D. M., Thomason, L. W., Zawodny, J. M., and Bodeker, G. E.: Assessment of SAGE version 6.1 ozone data quality, *J. Geophys. Res.*, 107, 4691–4700, 2002.