Investigation of photon path length distributions derived from oxygen A-band measurements of the GOSAT satellite instrument

Beke Kremmling

Dissertation for Attaining the Academic Degree of
"Doktor rerum naturalium" (Dr. rer. nat.) of the Departments

- 08 – Physics, Mathematics, and Computer Science,
- 09 – Chemistry, Pharmaceutics, and Geoscience,
- 10 – Biology,
- University Medicine

of the Johannes Gutenberg University

MAX-PLANCK-INSTITUT
FÜR CHEMIE

Max Planck Graduate Center
mit der Johannes Gutenberg-Universität Mainz

Satellite Remote Sensing Group
Max Planck Institute for Chemistry, Mainz

Title: *Investigation of photon path length distributions derived from oxygen A-band measurements of the GOSAT satellite instrument*

Author: *Beke Kremmling, born 1985 in Wiesbaden*

Reviewers:

Date of examination: *03.05.2018*
Location: *Mainz*

Contents

1 Abstract **5**

2 Zusammenfassung **6**

3 Introduction **7**

4 Absorption and scattering in the atmosphere **11**
4.1 Absorption of electromagnetic radiation 11
4.1.1 Energy states and transitions 11
4.1.2 Natural lineshape 12
4.1.3 Collisional broadening and Doppler broadening 12
4.1.4 The Voigt lineshape 13
4.1.5 Further lineshape effects 13
4.2 The oxygen A-band 14
4.3 Scattering 15
4.3.1 Rayleigh scattering 15
4.3.2 Raman scattering 16
4.3.3 Particle scattering 16
4.4 Absorption spectroscopy 18
4.5 Photon path length distributions and radiative transport in clouds 18

5 GOSAT **20**
5.1 General remarks 20
5.2 TANSO-FTS 20
5.3 Selected TANSO-FTS products 22
5.3.1 Estimated Geolocation Data (EGD) 22
5.3.2 FTS L1B 23
5.3.3 L2 Input Data Set (IDS) 25
5.4 TANSO-FTS instrument line shape function (ILSF) 25
5.5 TANSO-CAI 26
5.5.1 L2 cloud property product 27

6 Radiative transfer simulations and photon path statistics **29**
6.1 McArtim 29
6.2 Input parameters 31
6.2.1 Atmosphere 31
6.2.2 Solar irradiance 32
6.2.3 O_2 absorption cross sections 32
6.2.4 Sun-satellite geometry 33
6.2.5 Cloud parametrization 35
6.2.6 Ancillary simulation options 35
6.3 Rotational Raman scattering 35
6.4 Photon event information 36
6.5 Photon penetration altitude 37
6.6 Photon path length distributions 38

7 Dataset and case studies **41**
7.1 CALIPSO 41
7.2 CloudSat 43
7.3 Collocated measurements and cloud characteristics from CALIOP 43
7.4 Cloud characteristics from TANSO-CAI and CPR 45
7.5 Dataset 45
7.6 Case studies 46

7.6.1 Case 1 . . . 47
7.6.2 Case 2 . . . 51
7.6.3 Case 3 . . . 55
7.7 Appendix . . . 59
7.7.1 Datasets used for the cloud characterization . . . 59

8 Spectrum fit **60**
8.1 Motivation . . . 60
8.2 Cloud parameters . . . 60
8.3 Fit function . . . 63
8.4 Test fits and sensitivity tests . . . 69
8.4.1 Synthetic reference spectrum . . . 69
8.4.2 Influence of surface albedo variations . . . 73
8.4.3 Accuracy of the applied ILSF . . . 74
8.4.4 Conclusions on test fits and sensitivity tests . . . 76

9 Cloud-free atmospheric scenarios **78**
9.1 Measurement over dark surface . . . 78
9.2 Measurement over bright surface . . . 78
9.3 Conclusions . . . 81

10 Case study 1 **83**
10.1 Simulated cloud scenarios . . . 83
10.2 Spectrum fit . . . 83
10.3 Fit results . . . 85
10.4 Fit results in sub regions of the O_2 A-band . . . 87
10.5 Influence of a different asymmetry parameter g . . . 88
10.5.1 Fit results . . . 88
10.5.2 Photon path length statistics . . . 89
10.6 Further investigations . . . 91
10.6.1 Horizontal heterogeneity of cloud case 1 . . . 91
10.6.2 Influence of atmospheric pressure-temperature profiles . . . 92
10.6.3 Influence of the spectral sampling of the O_2 absorption cross sections . . . 93
10.7 Appendix . . . 95
10.7.1 Cloud configurations for case 1 using an asymmetry parameter g=0.85 . . . 95
10.7.2 Cloud configurations for case 1 using an asymmetry parameter g=0.80 . . . 96
10.7.3 Fit results in sub regions of the O_2 A-band . . . 97
10.7.4 Influence of the simulation noise on the fit result . . . 98
10.7.5 O_2 weighted photon path length distributions . . . 99

11 Case study 2 **102**
11.1 Simulated cloud scenarios . . . 102
11.2 Spectrum fit and results . . . 102
11.3 Assumed simplifications and heterogeneity of cloud case 2 . . . 105
11.4 Appendix . . . 107
11.4.1 Cloud configurations for case 2 using an asymmetry parameter g=0.85 . . . 107
11.4.2 Horizontal heterogeneity of cloud case 2 . . . 108
11.4.3 O_2 weighted photon path length distributions . . . 109

12 Case study 3 **110**
12.1 Simulated cloud scenarios . . . 110
12.2 Spectrum fit and results . . . 111
12.2.1 Comparison of photon path length distributions ($B_{fit} \approx 1$ vs. $B_{fit} < 1$) . . . 114
12.2.2 Comparison of photon path length distributions ($B_{fit} \approx 1$ vs. $B_{fit} > 1$) . . . 116

12.2.3 Best results and photon penetration altitudes 117
12.3 Fit results in sub regions of the O_2 A-band 119
12.4 Altitude resolved photon path length distributions 119
12.5 Appendix 122
12.5.1 Fit results in sub regions of the O_2 A-band 122
12.5.2 O_2 weighted photon path length distributions 123

13 Discussion and Outlook **125**

List of Figures **133**

List of Tables **137**

References **139**

1 Abstract

Clouds in the atmosphere have different macroscopic shapes and microphysical characteristics, strongly influencing the photon trajectories of the incoming solar radiation. Reflection at the cloud top leads to a shortening of the photon path lengths while multiple scattering inside the cloud and reflections between different cloud clusters can significantly increase the path lengths. All these modifications affect the short-wave energy deposition of the atmosphere.

Within this thesis, photon path length distributions of reflected sunlight and other cloud properties are retrieved from high spectral resolution satellite measurements of the oxygen A-band by performing comparative radiative transfer simulations. The measured radiances originate from the Fourier Transform Spectrometer TANSO-FTS onboard the GOSAT satellite. Radiative transfer simulations of different cloud scenarios are performed with the Monte Carlo model McArtim and compared to the measurements. The simulation output provides direct access to the scattering events, allowing the calculation of photon path length distributions. The comparison between measurement and simulation is achieved by means of an optimization procedure, which was developed within this thesis.

The investigation is applied to selected measurements in presence of single layer clouds and one cloud system with multiple layers. For those cases, also collocated lidar measurements of CALIOP (CALIPSO) and radar measurements of CPR (CloudSat) are available. The cloud properties of the simulations, which agree best to the measurement, are compared to the collocated cloud profiles from CALIOP and CPR as well as to the cloud optical depths from TANSO-CAI (GOSAT).

In general, the results show a reasonable agreement with the independent collocated measurements and the compared cloud optical depths agree well. From the comparison with the collocated cloud profiles, a systematic overestimation of the retrieved cloud top heights was found for the single layer cloud cases and an indication for an overestimation of the simulated O_2 absorption in the order of 5-10%. In order to resolve this discrepancy, different sensitivity studies have been performed, evaluating the horizontal homogeneity of the cloud system and simulation input parameters like the O_2 absorption cross sections and the asymmetry parameter. So far, no explanation for the observed discrepancy has been found. However, it is unlikely that it is caused solely by the simulation uncertainties. Interestingly, a similar overestimation for the O_2 absorption has been observed for a clear sky case over a bright surface. A clear-sky case over a dark surface, however, shows the opposite effect, suggesting an underestimation of the O_2 absorption. It is interesting to note that the corresponding radiance spectrum shows many negative values, indicating possible calibration problems. Here, more clear-sky cases of different surface albedos should be investigated for a better understanding of the performance of the instrument. The case study of the multiple layer cloud system leads to a better comparison to the collocated cloud profiles. Besides the observed overestimation of the O_2 absorption, the developed investigation method works well and can be applied to satellite measurements on a global scale.

2 Zusammenfassung

Wolken in der Atmosphäre haben unterschiedliche makroskopische Formen und mikrophysikalische Eigenschaften, welche starken Einfluss auf die Photonenwege der einfallenden solaren Einstrahlung ausüben. Photonenwege werden durch Reflektion der Strahlung an der Wolkenoberkante verkürzt, können durch Vielfachstreuung innerhalb der Wolke und Reflektionen zwischen verschiedenen Wolkenclustern aber auch erheblich verlängert werden. Diese Modifikationen der Photonwege beeinflussen den Energieeintrag kurzwelliger Strahlung in der Atmosphäre.

In der vorliegenden Arbeit werden Photonenweglängenverteilungen von reflektiertem Sonnenlicht und weitere Wolkenparameter bestimmt. Dies wird durch den Vergleich spektral hochaufgelöster Satellitenmessungen der Sauerstoff A-Bande mit Strahlungstransportsimulationen erreicht. Die gemessenen Radianzen entstammen dem Fourier-Transform-Spektrometer TANSO-FTS des GOSAT-Satelliten. Strahlungstransportsimulationen verschiedener Wolkenszenarien werden mit dem Monte Carlo Modell McArtim durchgeführt und mit den Messspektren verglichen. Die Modellergebnisse bieten direkten Zugang zu den simulierten Streuereignissen, was die Berechnung von Photonenweglängenverteilungen ermöglicht. Der Vergleich von Messung und Simulation wird durch einen Optimierungsprozess erzielt, welcher im Rahmen dieser Arbeit entwickelt wurde.

Die Auswertung wird auf ausgewählte Fälle von Einschichtwolken angewandt, sowie auf ein Szenario mit unterschiedlichen Wolkenschichten. Für diese Fälle existieren Überschneidungen von LIDAR Messungen durch CALIOP (CALIPSO) und RADAR Messungen durch CPR (CloudSat). Die Wolkenszenarien, welche am besten mit der Messung übereinstimmen, werden mit den Wolkenprofilen von CALIOP and CPR verglichen, sowie mit den optischen Dichten der Wolke, gemessen durch TANSO-CAI (GOSAT).

Allgemein zeigen die Ergebnisse der Auswertung eine befriedigende Übereinstimmung mit den unabhängigen Vergleichsmessungen, und die ermittelten optischen Dichten der Wolken werden gut getroffen. Aus dem Vergleich mit den unabhängigen Wolkenprofilen ergibt sich jedoch eine systematische Überschätzung der ermittelten Wolkenoberkanten für die Einschichtwolken, was mit einer Überschätzung der simulierten Sauerstoffabsorption in der Größenordnung von 5-10% einher geht. Um diese Unstimmigkeit aufzulösen, wurden verschiedene Sensitivitätsstudien durchgeführt, welche die horizontale Homogenität der Wolken und Konfigurationen der Simulationen wie die Absorptionsquerschnitte von Sauerstoff und den gewählten Asymmetrieparameter genauer betrachten. Bisher konnte keine Erklärung für die festgestellte Abweichung gefunden werden, aber es ist unwahrscheinlich dass diese alleine durch Unsicherheiten im Rahmen der Simulationen verursacht wird. Interessanterweise wurde für eine Messung mit wolkenlosem Himmel und heller Oberfläche eine ähnliche Überschätzung der Sauerstoffabsorption festgestellt. Die Analyse einer Messung mit wolkenlosem Himmel über einer dunklen Oberfläche wiederum zeigt das Gegenteil, nämlich eine Unterschätzung der simulierten Sauerstoffabsorption. Hier ist die Feststellung interessant, dass das zugehörigen Messspektrum viele negative Radianzen aufweist, was ein Indiz für mögliche Probleme der Kalibrierung ist. In Zukunft sollten weitere Fälle mit klarem Himmel und unterschiedlicher Bodenalbedo betrachtet werden, um die Leistungsfähigkeit des Instrumentes besser einschätzen zu können. Die Auswertung eines mehrschichtigen Wolkensystems zeigt eine bessere Übereinstimmung mit den Vergleichsprofilen der Wolke. Abgesehen von der beobachteten Überschätzung der Sauerstoffabsorption funktioniert der entwickelte Auswertungsprozess gut und kann auf Satellitenmessungen in globalem Maßstab angewandt werden.

3 Introduction

The presented work focuses on the investigation of atmospheric radiative transport in the clear and cloudy atmosphere, involving the retrieval of photon path length distributions within selected altitude regions. They are derived from measurements of the oxygen A-band (O_2 A-band, 759-772 nm) with high spectral resolution together with radiative transfer simulations. The measurements originate from the TANSO-FTS instrument [45] onboard the GOSAT satellite and the Monte Carlo based radiative transfer simulations are performed with the radiative transfer model McArtim [25].

In cloud- and aerosol-free conditions, the radiative transport is well characterized as light path trajectories are mainly influenced by the wavelength dependent Rayleigh scattering as well as the geometrical configuration of the situation (e.g. sun position) and the surface albedo. Clouds occur in very heterogeneous macroscopic structures and morphologies [28] as well as very different microphysical characteristics. Their presence significantly influences the radiative transport and thus the energy deposition in the atmosphere. The top of atmosphere energy budget is quantified by the so-called *Cloud Radiative Forcing*, describing the budget change in presence of clouds compared to a clear sky situation. While clouds cool the surface by shielding incoming shortwave radiation (*albedo effect*), they warm the surface by trapping longwave radiation (*greenhouse effect*) (see Quante [79] and references therein). The effect of clouds on the atmospheric energy budget depends on their influence on the radiative transfer and thus on the distribution of photon path lengths. Also, absorbing aerosol particles within the cloud influence the radiative energy budget [8] and weak continuum absorbers (such as water dimers [77]), which gain importance in presence of very long photon paths.

Typical modifications of the photon trajectories and thus photon path length distributions due to clouds are visualized in fig. 1 for the case of a satellite measurement of reflected sunlight. If incoming solar radiation is reflected at the cloud top, shorter photon path lengths result compared to the corresponding clear sky situation. The photons which penetrate into the cloud, are subject to multiple scattering events, which can also enlarge the trajectories of the photons [69]. In very dense clouds, these multiple scattering events can eventually lead to diffusion-like behaviour of the photon transport [55]. Clouds can also occur in clustered structures, causing additional reflections between different cloud layers or patches [103].

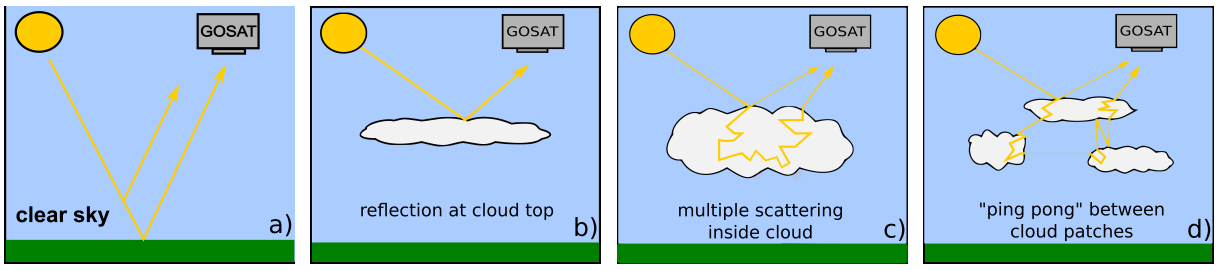

Figure 1: Examples of possible photon paths (from the Sun to the GOSAT detector) for different cloud scenarios (b-d) compared to a clear sky scenario (a). Example (b) shows a shortening of the light path by a cloud (shielding effect), while (c) also shows an elongation of the light path due to multiple scattering in a thick cloud and (d) between different cloud patches.

Photon path lengths can be described in terms of a probability density function (PDF) [87] and have been subject to several ground based studies and theoretical considerations, a selection is mentioned below. An access to this quantity is obtained through high spectral resolution measurements of the oxygen absorption. It is directly related to photon path lengths and results in the absorption structures of molecular oxygen which can be remotely measured by different devices, such as Fourier transform or grating spectrometers. A wavelength range well-suited for the retrieval of photon path lengths is the O_2 A-band in the near-infrared spectral region.

The O_2 A-band spectra show absorption structures of very different strengths, as illustrated in fig. 2 with two satellite measurements of the O_2 A-band, covering a sea (dark) and a land (bright) surface in clear-sky conditions. Atmospheric O_2 shows a well-known and constant mixing ratio in the atmosphere [52; 78]. The O_2 A-band lies in a spectral region which is nearly free of interference of other absorbers [52; 98]. The absorption line shapes of O_2 can be accessed via spectroscopic databases (e.g. HITRAN [83]) in combination with lineshape models. While deep absorption lines are mostly sensitive to short photon paths, weak absorption structures are mostly sensitive to long photon path lengths [85]. Many weak absorption lines do not saturate, even for very long photon path lengths [52].

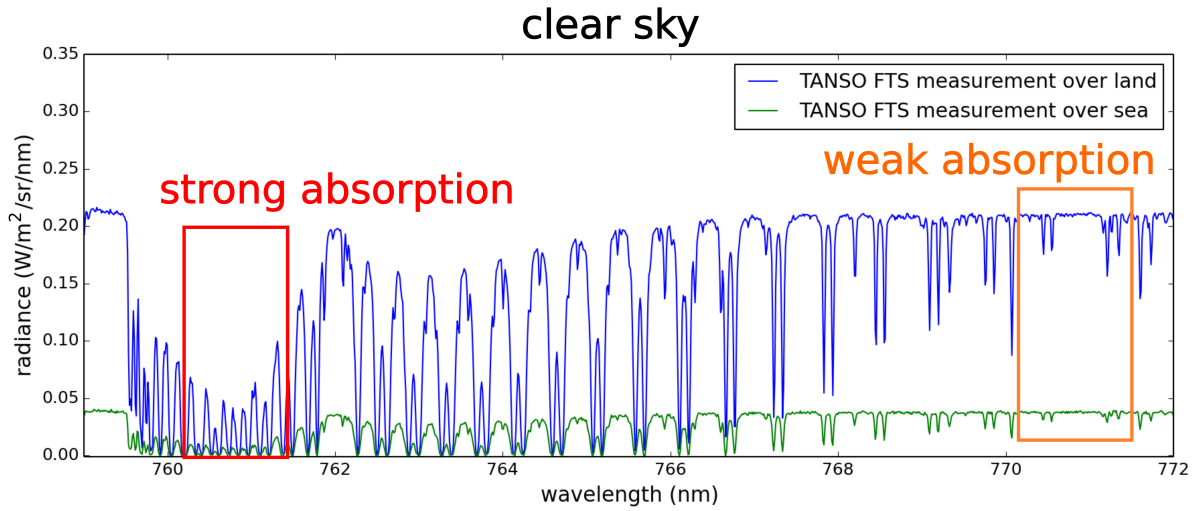

Figure 2: Space-based measurements of the oxygen A-band over land (blue) and sea surface (green). The measurement data originates from the TANSO-FTS spectrometer onboard GOSAT.

The O_2 A-band has been used within different remote sensing activities, like the retrieval of the surface pressure or the cloud top height (see e.g. references in Tran and Hartmann [98] and Long et al. [52]). Due to its importance for remote sensing purposes, scientific efforts have been carried out to characterize the spectroscopic models (e.g. [98; 99] or [51; 52]) to a great extent. Sensitivity analyses on the potential of the O_2 A-band spectra with different polarizations for cloud property retrievals, such as the cloud optical depth and cloud top height, from high resolution satellite spectrometers have been conducted by Sanghavi et al. [85] using model simulations.

In the past, several ground-based studies have been concerned with the retrieval of photon path length distributions from measurements of the O_2 A-band (see Pfeilsticker et al. [75; 76], Funk and Pfeilsticker [29] as well as Scholl et al. [87] and references therein). Among other results, it was shown that the assumption of plane parallel homogeneous cloud layers systematically underestimates photon path lengths [76]. Photon path length distributions can be obtained by assuming an analytical shape of the distribution and applying an inverse Laplace transformation on the measurement [34]. Most of these ground-based studies are based on the assumption of Γ-distributed photon path lengths and often, only low order moments of the photon path length distributions are retrieved. Besides a Γ-distribution, Funk and Pfeilsticker [29] also assumed log-normal distributed photon path lengths and validated the results with a Monte Carlo radiative transfer code. A number of studies were also concerned with the diffusion regimes of photons in different cloud scenarios, including classical Gaussian diffusion and Lévy transport (e.g. Pfeilsticker [75]); the latter leading to an increased probability of short and very long path lengths [34]. A more recent study investigated the use of photon path length distributions for the detection of multiple cloud layers (see Li and Min [49]). As shown in Oshchepkov et al. [72] (and references therein), parametrized photon path length statistics from O_2 A-band measurements can be used to identify measurements from the GOSAT satellite where

light paths are only little modified by cirrus clouds and aerosols (performed within the context of CO_2 retrievals). Simulation studies performed by Heidinger and Stephens [34] were concerned with the effect of cloud heterogeneities on the photon path lengths based on O_2 A-band spectra of reflected sunlight as given for spaceborne instrumentation.

The ground-based studies, together with theoretical and radiative transfer investigations, laid an important foundation for the retrieval of photon path length statistics in remote sensing. However, there is a difference between the photon path length statistics of transmitted and reflected photons [92]. Ground-based measurements allow the retrieval of information on the trajectory lengths of photons emerging downwards from the cloud, thus in the case of very dense clouds, only a small part of photons is considered. In order to obtain information on the reflected photons, space-based measurements are a suitable choice. Nowadays, spaceborne spectrometers, like the Fourier transform spectrometer onboard GOSAT (Greenhouse Gases Observing SATellite), are able to measure the O_2 A-band with very high spectral resolution, enabling the retrieval of photon path lengths. To our knowledge, the space-based retrieval of photon path length distributions, together with radiative transfer simulations, has never been investigated before and motivated the topic of this thesis. These measurements do not only allow a global investigation but a more general view with regard to the atmospheric energy deposition as a larger part of the incident solar radiation is considered compared to ground-based measurements below a cloud cover.

As mentioned before, photon path length statistics in cloudy skies are related to the atmospheric energy budget and information retrieved from the O_2 A-band can thus contribute to a more detailed knowledge on the deposition of short-wave energy in the atmosphere. They can also help to investigate the influence of weak continuum absorbers or weakly absorbing aerosols (single scattering albedo slightly below unity) embedded in the cloud structure or droplets. The effect of these absorbers for the short-wave energy budget becomes especially important for very long photon path lengths.

In this study, high resolution O_2 A-band measurements from the TANSO-FTS instrument on the GOSAT satellite are used. GOSAT was successfully launched in 2009 and continues operation well beyond the originally assumed time frame of 5 years [45]. The spectral resolution within the spectral range of the O_2 A-band is $\delta\lambda \approx 0.022$ nm [40] which almost allows to entirely resolve individual absorption lines. The measurements are compared to radiative transfer simulations using the Monte Carlo model McArtim which allows a flexible configuration of the model atmosphere and has the advantage of providing direct access to the simulated photon trajectories from the explicit modelling of absorption and scattering events. An example of a simulated photon path inside a cloud is shown in fig. 3. The comparison of the model to the measurement is achieved by simulating different cloud scenarios and evaluating the resemblance to the measurement spectrum using a fit routine which was developed within this thesis. The results are then compared to collocated measurements from lidar and radar observations.

The thesis is structured as follows: Section 4 starts with a short overview on atmospheric absorption and scattering processes and introduces photon path length distributions as well as some notations on cloud properties. Section 5 describes the GOSAT satellite mission, relevant instruments and related data products which are used within the retrieval. Information on the radiative transfer model McArtim and the chosen model setup is summarized in section 6, together with a description of the investigated photon path length statistics. The developed algorithm is applied on selected case studies of well-characterized cloud situations. These are selected by identifying collocated measurements of TANSO-FTS and the lidar and radar instruments on CALIPSO [105] and Cloudsat [91], respectively. The information on the cloud profile allows a comparison to the cloud properties retrieved. The selection procedure and a presentation of the investigated case studies is presented in section 7, followed by a description

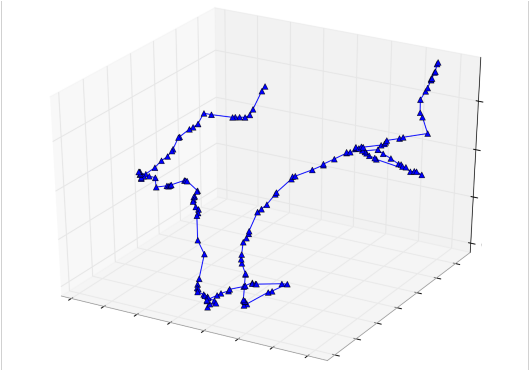

Figure 3: A possible photon trajectory entering and leaving a homogeneous cloud structure, simulated with the radiative transfer model McArtim. Only particle scattering events within the cloud are shown, the event positions are connected by lines.

of the fitting procedure in section 8, which evaluates the resemblance of the measurement to the simulations. The results are presented and discussed in sections 10, 11, 12 and 13, which includes a detailed outlook for future investigations.

4 Absorption and scattering in the atmosphere

The interaction of electromagnetic radiation with the constituents of the Earth's atmosphere is introduced in the following. Well-described overviews are to be found in e.g. Platt and Stutz [78], Burrows et al. [10] and Liou [50]. The following introduction covers the extinction of solar radiation by absorption and scattering processes and introduces the concept of absorption spectroscopy. The oxygen A-band (O_2 A-band) is presented as well as the notation of selected cloud properties and photon path length distributions.

Within this thesis, satellite measurements of O_2 A-band radiance spectra are compared to radiative transfer simulations. As input for the radiative transfer simulations, a solar irradiance spectrum is used. Here, a brief description on the terms radiant flux, irradiance and radiance is given, which include a wavelength dependency that is omitted in the following presentation for reasons of clarity.

The **radiant flux** Φ defines how much energy per unit time is emitted by a light source. It is independent of the emission direction and the corresponding unit is W. The term **irradiance** I_0 provides a measure, how much of the radiated energy arrives at the area A_i. If the area is perpendicular to the propagation direction of the radiation, it is expressed as:

$$I_0 = \frac{\Phi}{A_i} \frac{\mathrm{W}}{\mathrm{m}^2} \tag{1}$$

If the area A_i has a non-perpendicular orientation to the propagation direction of arriving radiation, an additional factor $\cos(\theta)$ has to be considered where θ is the angle between the surface normale and the arriving radiation.

The term **Radiance** I, on the other hand, defines the radiant flux per unit of area for a specific solid angle Ω:

$$I = \frac{\Phi}{\Omega \cdot A_r} \frac{\mathrm{W}}{\mathrm{sr} \cdot \mathrm{m}^2} \tag{2}$$

[78; 82].

4.1 Absorption of electromagnetic radiation

During the trajectory of light through the atmosphere, air molecules interact with electromagnetic radiation by scattering and absorption. When the arriving photon is absorbed, the participating molecule is transferred into a higher state of energy which exhibits discrete values due to the discrete photon energies [10]. There exist different kinds of transitions, namely electronic transitions ($\sim$ 1 eV), vibrational transitions ($\sim$ 0.1 eV) and rotational transitions ($\sim 10^{-3}$ - 10^{-2} eV) [78]. Vibrational transitions always include rotational transitions as well which leads to so-called "vibrational-rotational bands" in the infrared spectral region [50].

4.1.1 Energy states and transitions

The energy of a pure **rotational state** is

$$E_j = B \cdot J(J+1) \tag{3}$$

with the rotational quantum number J and the rotational constant $B = \frac{\hbar^2}{2\Theta}$ (Θ is the moment of inertia of the molecule). In order to have pure rotational transitions, the molecule must possess a dipole moment. The resulting transitions split into the P-branch (ΔJ = -1), the Q-branch (ΔJ = 0) and the R-branch (ΔJ = +1). It should be noted that the Q-branch is only possible in presence of a simultaneous electronic transition. The corresponding photon energy and thus the energy difference between two states is proportional to the rotational quantum number J:

$$\Delta E_j = E_{j+1} - E_j = 2B(J+1) \tag{4}$$

The resulting spectrum for purely rotational transitions consists of equally distanced lines with a spacing in the order of $10^{-3} - 10^{-2}$ eV.

The energy of a **vibrational state** can be approximated by the description of a quantum harmonic oscillator:

$$E_\nu = \left(\nu + \frac{1}{2}\right) \cdot \hbar\omega_0 \tag{5}$$

Here, the vibrational quantum number is $\nu = 0, 1, 2, \ldots$. According to the selection rules, $\Delta\nu = \pm 1$. The energy range of the vibrational states corresponds to the infrared spectral range. In order to be coupled to the electronic field, the molecule in question must possess a permanent dipole moment. This, for example, is not the case for O_2 [10; 78]. However, a transition is still possible for molecular oxygen because it has a weak magnetic dipole moment, leading to absorption bands in the red spectral range [36; 52] (e.g. O_2 A-band).

When rotational excitations take place alongside with vibrational transitions, this leads to a further splitting of the vibrational lines into finer rotational structures. For molecules, the rotational structure can be too fine to be resolved by the spectrometer or the overlap is too strong which leads to a continuous appearance of the absorption structures.

Electronic energy states are characterized by a combination of the vibrational and rotational states with differing moment of inertia Θ according to the prevailing configuration. Because of the dependence of the rotational constant B on Θ, the final spacing between the rotational lines is not equidistant. Electronic states are either allowed or forbidden without the existence of simple selection rules.

The relation between two **populations N(E) of states** with given energies $E_i > E_0$ (where E_0 is the ground state) is expressed using the Boltzmann distribution:

$$\frac{N(E_2)}{N(E_1)} = \frac{G_2}{G_1} \cdot \exp\left(-\frac{E_2 - E_1}{kT}\right) \tag{6}$$

k is the Boltzmann constant, G_1 and G_2 denote the degeneracy factors (number of different states with energy E_1 or E_2, respectively). While the vibrational states show no degeneracy, the degeneracy of the rotational states is given by $G_J = 2J + 1$ [10; 78].

4.1.2 Natural lineshape

When a molecule in an excited energy state decays to another state, e.g. the ground state, the result is a Lorentzian shaped transition line:

$$I_L(\omega)d\omega = I_0 \frac{\delta_L^2/4}{(\omega - \omega_0)^2 - \delta_L^2/4} d\omega \tag{7}$$

Here, $\delta_L = \sum \delta_i$ is the full width half maximum (FWHM) of the lineshape where δ_i are the transition frequencies of the different states. The value of δ_L mostly depends on the lifetime of the state under consideration (usually about 10^{-8} s) [10; 78].

4.1.3 Collisional broadening and Doppler broadening

There are different broadening effects which occur when the molecules are incorporated into a gas being subject to brownian motion, the most known are the pressure (collisional) broadening and the Doppler broadening.

Pressure (collisional) broadening results from collisions of molecules which reduce the lifetime of the exited states. The reduced lifetime results in a broadening of the natural lineshape leading to a FWHM of $\delta_p > \delta_L$. The FWHM δ_p is calculated by the product of the kinetic collision frequency of the considered gas and the deactivation probability of each collision (which depends on the interacting molecules). This includes collisions of different molecular species (air broadening) or species of the same kind (self-broadening). The air broadening is usually dominant in relation to the self-broadening. Relating the collision frequency to the gas density and the average molecular speed, the following relation of δ_p to pressure and temperature is derived (p_0 and T_0 denote reference values):

$$\delta_p(p,T) = \delta_P(p_0,T_0) \cdot \frac{p}{p_0}\sqrt{\frac{T_0}{T}} \tag{8}$$

As for the natural lineshape, the pressure broadened lineshape I_P is described by a Lorentzian shape and can thus be expressed with eq. 7, replacing δ_L with δ_p.

Doppler broadening is also referred to as Gaussian broadening because of the resulting spectral lineshape. The effect is caused the Doppler effect which modifies the emission frequency of the radiation. The Doppler broadened lineshape I_D is expressed as

$$I_D(\omega)d\omega = I_0 \cdot \exp\left(-\frac{\delta_D}{2} \cdot (\omega_0 - \omega)^2\right) d\omega \tag{9}$$

centered at ω_0 with the FWHM of

$$\delta_D = \omega\frac{2\sqrt{2R\ln 2}}{c} \cdot \sqrt{\frac{T}{m}} \tag{10}$$

[10; 78].

4.1.4 The Voigt lineshape

Except at very low pressures, the lineshape is governed by the pressure and Doppler broadening which have larger FWHM than the natural lineshape. The overall lineshape is frequency dependent and takes different values according to the prevailing temperature and pressure values of the environment. The so-called Voigt lineshape is the outcome of a convolution of the spectral lineshapes I_D and I_P, caused by Doppler and pressure broadening, respectively:

$$I_V(\omega,\delta_D,\delta_L) = \int_{-\infty}^{\infty} I_D(\omega',\delta_D) \cdot I_P(\omega-\omega',\delta_P)d\omega' \tag{11}$$

It should be noted that the above mentioned lineshapes have been described in terms of spontaneous emission. Because absorption is the inverse process of emission, these lineshapes also apply to absorption lines and thus lead to the wavelength (frequency) dependent absorption cross sections $\sigma(\lambda)$ ($\sigma(\omega)$) [10; 78].

4.1.5 Further lineshape effects

Nowadays, it is common knowledge that the Voigt lineshape alone does not accurately describe the shape of the absorption cross sections. Here, two other effects are shortly described which are incorporated in the model for the derivation of the O_2 absorption cross sections (Tran and Hartmann [98]; Tran et al. [99]) which is used within this study.

Line mixing (LM)

Line mixing (LM) can occur when the different transition frequencies are very near and the lineshapes cannot not be considered "collisionally isolated" [33]. The frequency separation should not be larger than the broadened linewidths. Inelastic collisions between molecules of the same

kind cause a population transfer between the energy states of the participating molecules. This is schematically illustrated in fig. 4. Here, two optical transitions of similar energy differences ΔE_{fi} and $\Delta E_{f'i'}$ are considered. For example, an inelastic collision between the molecules can cause a molecule to be transferred from energy level E_i to E_i'. The molecule can then be excited from energy level E_i' to E_f' through absorption of a photon (leading to a corresponding absorption line). Another inelastic collision between the molecules and the implied energy exchange can cause the molecule to relax to level E_f [33].

In comparison to the Voigt profile, the result is a change of the observed spectral lineshapes, where the absorption becomes weaker in the line wings but is shifted towards the line center. If only an isolated lineshape is modeled, this leads to a narrowing of the spectral lineshape [33; 74]. For a more detailed description on line mixing, please refer to e.g. Hartmann et al. [33].

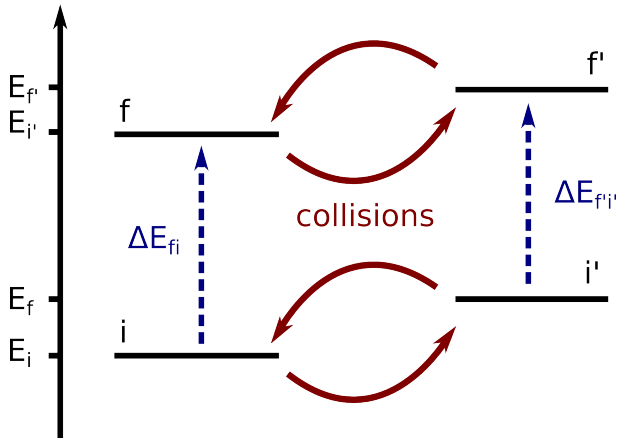

Figure 4: Schematical description of the line mixing process (slightly modified from [33]).

Collision Induced Absorption (CIA) for O_2

Although molecular oxygen does not possess an electric dipole moment, a short-lived electric dipole can be created by collisions between two oxygen molecules which leads to additional absorption. This is called the collision induced absorption (CIA) whose contribution is proportional to the squared density of O_2. Because the collisions are short-lived, the weak additional absorptions do not show distinct structures [80; 89]. Within the work of Tran et al. [99], the contribution of CIA was investigated for collisions within the oxygen A-band between O_2 and O_2 as well as O_2 and N_2 (where CIA is much more important for O_2-O_2 collisions than for O_2-N_2 collisions), having a weak temperature dependence between 200 K and 300 K. Together with the effect of line mixing, the overall absorption was enlarged in comparison to the traditional Voigt lineshapes.

4.2 The oxygen A-band

Molecular oxygen is an abundant (the volume mixing ratio is 20.948%) and well-characterized absorber in the atmosphere. In this study, the focus is laid on the O_2 A-band which defines the spectral region between 12800 and 13200 cm^{-1}. The band's center is at the vibrational transition of 762 nm ($[b^1\sum_g^+ \leftarrow X^3\sum_g^-(0,0)]$) [50; 52]. Although forbidden due to the missing electric dipole, this transition is possible due to the magnetic dipole moment [36; 52]. The band center is surrounded by two rotational transition branches: the P-branch (ΔJ = -1, centered at appr. 764 nm) and the R-branch (ΔJ = +1, centered at appr. 761 nm) [85].

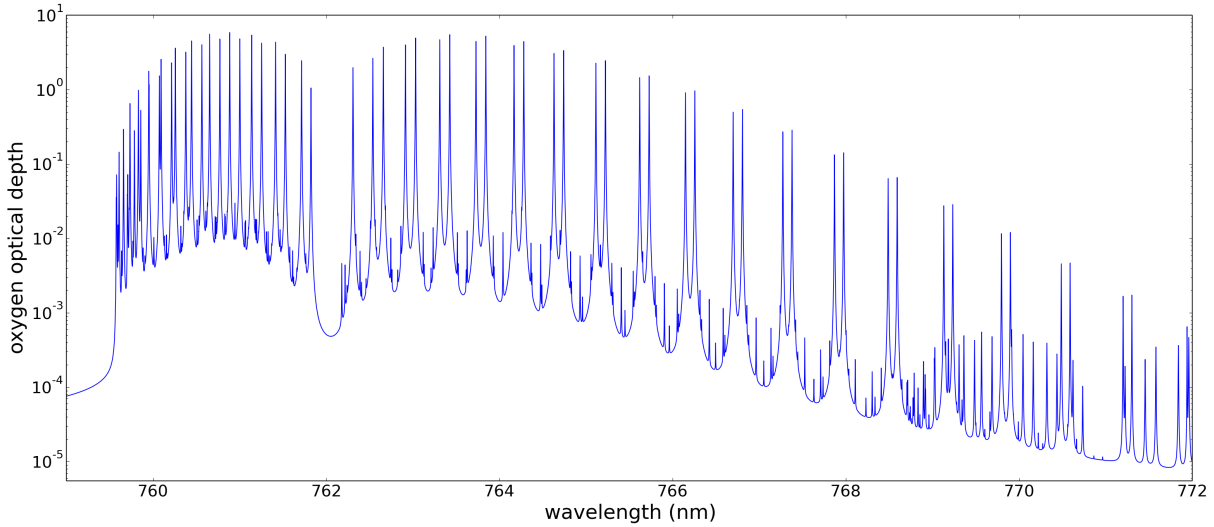

Figure 5: Vertical column density of O_2 in the model atmosphere within the O_2 A-band, showing lines of many different absorption strengths.

The O_2 A-band has a special status because it is situated in a spectral region which is nearly free from interferences of other absorbing gases [52; 98]. It also shows a large range of different line strengths, resulting in optical depths of seven orders of magnitude. However, the finite resolution of instruments reduces the variety of observed optical depths [90]. Also, many absorption lines are not saturated, even for very long photon path lengths [81]. The vertical column density of O_2 in the simulation atmosphere within the O_2 A-band is shown in fig. 5 as calculated with the absorption cross sections from the database and model described in section 6.2.3.

4.3 Scattering

Scattering can be described as the interaction of radiation with atmospheric constituents like air molecules, cloud particles (formed of liquid or solid water) or aerosol particles which induce a change of the propagation direction of the radiation. While elastic scattering does not include energy changes and thus only affects the propagation direction, inelastic scattering includes energy transfer between the photon and the scattering object [78]. The type of scattering depends on the size and the properties of the scattering object.

If the dimension of the scattering object is small compared to the wavelength of the participating photon, Rayleigh scattering takes place. For wavelengths in the visible and near-infrared spectral region, this concerns especially air molecules but it can also apply to small aerosol particles with linear dimensions below 0.1 μm. If the wavelength is not small to the linear dimension of the scattering object, particle scattering occurs. Possible scattering objects are cloud or aerosol particles [82]. Particle scattering is often referred to as Mie scattering. Because Mie theory in principle only applies to spherical particles [82], particle scattering seems a more accurate description at this place.

4.3.1 Rayleigh scattering

Rayleigh scattering describes the elastic scattering of photons on air molecules. Its scattering cross section is expressed as:

$$\sigma_R(\lambda) = \frac{24\pi^3}{\lambda^4 N_{air}^2} \cdot \frac{(n_0(\lambda)^2 - 1)^2}{(n_0(\lambda)^2 + 2)^2} \cdot F_K(\lambda) \approx \frac{8\pi^3}{3\lambda^4 N_{air}^2} \cdot (n_0(\lambda)^2 - 1)^2 \cdot F_K(\lambda) \qquad (12)$$

Here, $n_0(\lambda)$ is the index of refraction of air, N_{air} the number density of air and $F_K(\lambda)$ is a correction factor to account for the anisotropy of air molecules [78; 81]. The scattering cross

section is almost proportional to λ^{-4} [82], the probability of elastic scattering is thus low in the O_2 A-band around 760 nm. Because of the complexity of eq. 12, different approximations exist (e.g. Penndorf (1957, [73])).

The extinction coefficient for Rayleigh scattering is the product of the scattering cross section $\sigma_R(\lambda)$ and the air density N_{air}:

$$\epsilon_R(\lambda) = \sigma_R(\lambda) \cdot N_{air} \tag{13}$$

The corresponding phase function describes the probability of the resulting scattering direction θ. The phase function depends on the state of polarization of the incident light. For unpolarized light, it is given by:

$$\Phi(\cos(\theta)) = \frac{3}{4}\left(1 + \cos^2(\theta)\right) \tag{14}$$

[78; 81]

4.3.2 Raman scattering

Inelastic scattering between photons and scattering objects is referred to as Raman scattering. Here, energy transfer takes place and part of the photon's energy is transferred to the scatterer or vice versa. Rotational Raman scattering (RRS) describes the events where only rotational transitions are initiated [10]. In the atmosphere, rotational Raman scattering concerns only 2-4% of the scattering on air molecules, resulting in wavelength shifts of values up to appr. 2 nm [82]. In contrast to Rayleigh scattering, the scattering angle and the polarization of the incident light have only a small influence on the phase function. Rotational Raman scattering is the cause of the often observed Ring effect. Due to this effect, the absorption depths of the Fraunhofer lines are less deep than expected because the redistribution of absorptions results in a "filling-in" of the structures [81].

Due to the frequency changes, additional band structures appear. It is distinguished between Stokes scattering and Anti-Stokes scattering. The Stokes band consists of lower frequencies than the original photon energy while the Anti-Stokes band shows higher frequencies [78].

If vibrational excitation takes place as well, one refers to (rotational)-vibrational Raman scattering which is much weaker than the rotational Raman scattering. For a temperature of 273 K and a wavelength of 770 nm, the cross section of rotational Raman scattering is 6.1% and that of vibrational Raman scattering 0.1% in comparison to Rayleigh scattering [78]. Nevertheless, vibrational Raman scattering by water (e.g. ocean surfaces) can lead to significantly larger backscattered radiances [10]. Vibrational Raman scattering is usually strongest over ocean surface with a very low amount of suspended particles and thus a deep penetration of the incident radiation within the water [81].

For more details on Raman scattering, please refer to e.g. Platt and Stutz [78].

4.3.3 Particle scattering

In general, the scattering properties depend on the particle characteristics which include the size of the scattering object, the composition (real and imaginary index of refraction) and shape (e.g. spherical water droplets). For spherical scattering objects, the well-known Mie theory (Mie (1908, [57]), van de Hulst (1981, [100])) offers a good description of the scattering characteristics. Nevertheless, the theory can often be used as approximation for non-spherical particles (e.g. aspherical aerosol particles) in case that not an individual scattering particle, but an ensemble of scattering objects is considered [81]. The scattering is elastic and the wavelength dependency is weak (a typical value is $\lambda^{-1.3}$). In contrast to Rayleigh scattering, a dominant contribution

in the forward direction of the scattered radiation is observed [78]. Larger particle sizes increase the asymmetry between scattering in the forward and backward direction [81].

The phase functions are usually calculated by numerical models because they include summations over series that are only slowly converging. An example of the numerical calculation of the Mie phase functions $\Phi_{Mie}(\theta)$ for water droplets of different radius r at 750 nm is shown in fig. 6. The data was calculated with the software *MieTool* (2013) from T. Deutschmann.

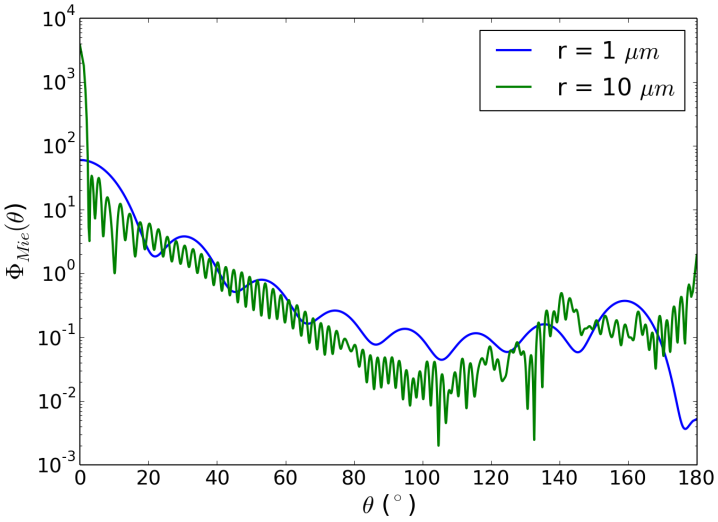

Figure 6: Examples of the Mie phase function $\Phi_{Mie}(\theta)$ for a water droplet with radius r=1, 10 μm at λ = 750 nm.

An alternative to the numerical calculation of the scattering phase functions is the use of an approximate formula, e.g. the Henyey-Greenstein phase function (Henyey and Greenstein [35], 1941) for particle ensembles, which is used in this study:

$$\Phi(\cos(\theta)) = \frac{(1-g^2)}{4\pi(1+g^2-2g\cos(\theta))^{3/2}} \tag{15}$$

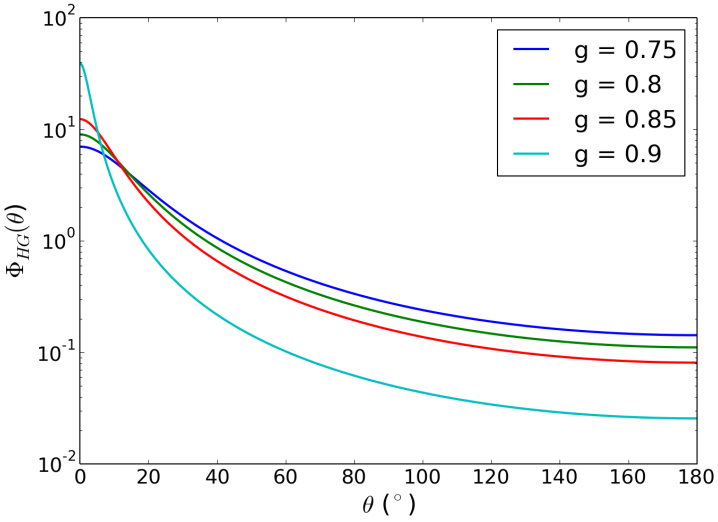

Figure 7: The Henyey-Greenstein phase function $\Phi_{HG}(\theta)$ for different asymmetry parameters g.

The above equation is characterized by the asymmetry parameter g, which describes the shape of the scattering phase function. If g=1, the scattering takes place entirely in the forward direction, while g=0 describes isotropic scattering [78; 81]:

$$g = \frac{1}{2} \int_{-1}^{1} P(\cos(\theta)) \cdot \cos(\theta) \cdot d\cos(\theta) \tag{16}$$

A typical value for liquid cloud droplets is g=0.85 [82], aerosol particles in the troposphere can usually be well described with g≈ 0.6-0.7 [78]. The Henyey-Greenstein phase function $\Phi_{HG}(\theta)$ calculated for different values of the asymmetry parameter g, is shown in fig. 7.

4.4 Absorption spectroscopy

The absorption of electromagnetic radiation is described by the Lambert-Beer law which is given by:

$$I(\lambda) = I_0(\lambda) \cdot \exp\left(-\sigma(\lambda) \cdot c \cdot L\right) \tag{17}$$

Here, the initial radiance I_0 of a light source is subject to absorption, characterized by the wavelength dependent absorption cross section $\sigma(\lambda)$ of the present species with homogeneous concentration c. L is the distance, the light travels within the medium. In the atmosphere, many absorbers with different absorption cross sections are present. Also, scattering and absorption events take place, e.g. on air molecules and particles which reduces I_0. Instrument effects will further modify the measured radiance I. A more complete description of I is thus

$$I(\lambda) = I_0(\lambda) \cdot \exp\left[-L \cdot \left(\sum_j (\sigma_j(\lambda) \cdot c_j) + \epsilon_R(\lambda) + \epsilon_M(\lambda)\right)\right] \cdot A(\lambda) \tag{18}$$

The extinction terms are given by the Rayleigh extinction ϵ_R and the Mie extinction ϵ_M, the instrument effects are summarized in $A(\lambda)$ [78]. So far, a well-defined light source has been considered. For spaceborne measurements of scattering sunlight, $A(\lambda)$ also includes a so-called "efficiency factor". It considers effects from the reflectance at the Earth's surface and atmospheric scattering of the solar irradiance into the instrument's field of view [81].

The optical depth is defined as [10]:

$$\tau = -\ln\left(\frac{I(\lambda)}{I_0(\lambda)}\right) \tag{19}$$

Consequently, the Lamber-Beer law can be re-written as:

$$I(\lambda) = I_0(\lambda) \cdot \exp\left(-\tau(\lambda)\right) \tag{20}$$

4.5 Photon path length distributions and radiative transport in clouds

In a multiple scattering atmosphere, the photons follow different trajectories leading to a multitude of geometric photon path lengths l. This can be expressed by the **photon path length distribution** [34; 92] which modifies eq. 17 to

$$I(\lambda) = I_0(\lambda) \int_0^\infty p(l) \cdot \exp(-\sigma(\lambda) \cdot c \cdot l) \cdot dl \tag{21}$$

Using this relation, the equivalence theorem (see e.g. Stephens et al. [92] for details) states that the radiance I in a certain direction within an absorption line can be determined if the radiance I_0 outside of the absorption line of the same direction and the photon path length distribution is known. Here, it should be noted that the photon path length distributions are different for reflected and transmitted photons [92] and thus for ground- and space-borne measurements.

Photon path length distributions can mathematically be obtained through an inverse Laplace transform (for a description, see e.g. Funk and Pfeilsticker [29]) of the optical depth but an

analytical shape has to be assumed due to the restricted measurement information which does not cover the imaginary part of the Laplace transformation [87]. For transmission spectra, Γ- as well as log-normal distributed photon path lengths have been assumed (e.g. Funk and Pfeilsticker [29]). An alternative method, which is used in this thesis, is to obtain photon path length statistics through Monte Carlo simulations of the atmospheric radiative transport when the trajectory information can be accessed.

Within this thesis, also the **cloud optical depth** is retrieved from measurements of the O_2 A-band, assuming homogeneous cloud layers. In case of a homogeneous cloud with constant cloud extinction coefficient $\epsilon_{ext,cloud}$, the cloud optical depth τ_{cloud} is defined as the geometrical thickness l_{cloud} of the cloud multiplied with the extinction coefficient [42; 43]:

$$\tau_{cloud}(\lambda) = \epsilon_{ext,cloud}(\lambda) \cdot l_{cloud} \tag{22}$$

In a more general case, e.g. if the cloud shows altitude dependent extinction coefficients due to internal inhomogeneities, the cloud optical depth is obtained by integration of these coefficients along the vertical extent of the cloud [43]:

$$\tau_{cloud}(\lambda) = \int_{z_0}^{z_N} \epsilon_{ext,cloud}(\lambda, z_i) \cdot dz_i \tag{23}$$

In the case of a homogeneous and plane-parallel cloud layer, the **average number of scattering events** N therein can be related to the cloud optical depth τ_{cloud} in the diffusion limit ($(1-g) \cdot \tau_{cloud} \gg 1$, thus valid for optically thick clouds). For reflected photons, the average number of scattering events is proportional to the cloud optical depth:

$$N_R(\tau_{cloud}, g) \sim \tau_{cloud} \tag{24}$$

For transmitted photons, the average number of scattering events also depends on the asymmetry parameter g:

$$N_T(\tau_{cloud}, g) \sim (1-g) \cdot \tau_{cloud}^2 \tag{25}$$

[54; 55]

5 GOSAT

5.1 General remarks

The Greenhouse Gases Observing SATellite (GOSAT) was launched in 2009 and designed to achieve high-accuracy measurements of tropospheric carbon dioxide (CO_2) and methane (CH_4) on a global scale. The mission is a joint project of the Japan Aerospace Exploration Agency (JAXA), the Ministry of the Environment (MOE) and the National Institute for Environmental Studies (NIES).

The nadir-pointing satellite is orbiting at a sun-synchronous orbit in an altitude of 666 km and has an observation recurrence cycle of three days. The satellite's main specifications are summarized in table 1. The high resolution Fourier transform spectrometer TANSO-FTS performs intensity measurements in different absorption bands in the near-infrared wavelength region as well as the thermal infrared. From these measurements, column densities of the Greenhouse Gases are derived. Measurements of the oxygen A-band are included to estimate the effective photon path lengths in the atmosphere. TANSO-CAI, a CCD camera, provides pixel radiances of the observation areas and allows to evaluate cloud and aerosol coverages [6; 45]. The instruments still continue data aquisition in 2017, GOSAT is thus well beyond the estimated lifetime of 5 years. The status on instruments and calibrations are published on a regular basis, e.g. in Kuze et al. [48].

Table 1: Main specifications and orbital configurations of the GOSAT satellite (slightly modified from [96]).

size	main body: 3.7 m x 1.8 m x 2.0 m wing span: 13.7 m
mass	1.75 kg
power	3.8 kW
lifetime	5 years
launch date	23.01.2009
orbit type	sun-synchronous, quasi-recurrent
altitude	666 m at Equator
inclination angle	98.06°
orbits per day	14 + 2/3 revolutions
recurrence	44 revolutions per 3 days
descending node time	13 h ± 15 min

5.2 TANSO-FTS

The TANSO-FTS (Thermal and Near infrared Sensor for Carbon Observation Fourier Transform Spectrometer) performs near-nadir measurements of the reflected radiance using different observation patterns with an instantaneous field of view (IFOV) of circular shape with a 10.5 km diameter. Reflected radiances are measured in different wavelength bands in the near infrared and short-wave infrared. Additionally, thermal infrared measurements are performed [45; 48]. The specifications of the individual wavelength bands are described in table 2.

The instrument itself consists of an optics unit, a control unit and a redundant electronics unit. All elements of the optics unit are mounted on a temperature-controlled optical bench. The unit (see fig. 8) can be separated into the pointing system, the Fourier-transform spectrometer and after-optics (red in fig. 8) and the detectors. A monitor camera is included to observe the

Table 2: Specifications of TANSO-FTS measurement bands 1-4. The spectral resolution is defined as $\frac{1}{2 \cdot \max(\text{path difference})}$*. All data within the table is copied or adapted from [96].*

Band name	**Band 1 (Visible/NIR[1])**
Polarimetric observation	Yes
Wavenumber region	12900-13200 cm^{-1}
Spectral resolution	0.2 cm^{-1}
FWHM of the instrument line shape	0.6 cm^{-1} or less
targeted gases and measurements	O_2 → information on air pressure and cirrus clouds
Band name	**Band 2 (SWIR)**
Polarimetric observation	Yes
Wavenumber region	5800-6400 cm^{-1}
Spectral resolution	0.2 cm^{-1}
FWHM of the instrument line shape	0.27 cm^{-1} or less
targeted gases and measurements	CO_2, CH_4
Band name	**Band 3 (SWIR)**
Polarimetric observation	Yes
Wavenumber region	4800-5200 cm^{-1}
Spectral resolution	0.2 cm^{-1}
FWHM of the instrument line shape	0.27 cm^{-1} or less
targeted gases and measurements	CO_2, H_2O → information on cirrus clouds
Band name	**Band 4 (TIR)**
Polarimetric observation	No
Wavenumber region	700-1800 cm^{-1}
Spectral resolution	0.2 cm^{-1}
FWHM of the instrument line shape	0.27 cm^{-1} or less
targeted gases and measurements	CO_2, CH_4, H_2O, O_3

TANSO-FTS viewing direction.

The pointing system is designed to have two kinds of angular movements: the cross track (CT) and the along track (AT) direction. By combination of these movements, different locations on the Earth's surface can be observed following defined observation patterns (optimized throughout the satellite's lifetime) or by pointing specific targets. Deep space, blackbody and diffuser measurements can also be performed by specific rotations of the pointing system, e.g. for calibration measurements [45; 46]. The CT movement includes angles of ± 35° from nadir, the AT movement performs physical angular movement of ± 10° plus ± 20° in the line of sight. A redundant secondary pointing mechanism is mounted on the same optical bench [45]. It replaced the primary pointing mechanism in 26 January, 2015 by inserting a fold mirror into the optical path [45; 48].

The TANSO-FTS is a double pendulum Fourier transform spectrometer using two corner reflectors to create path differences of ± 2.5 cm of the entering radiance. The passing light is collected at a common circular stop and passed through dichroic filters (DF1 - DF3), which separate it into narrow wavelength bands. Polarization beam splitters further divide the beams into two polarization directions P and S. The signal is recorded by a Si detector in band 1 and InGaAs detectors in band 2 and 3. The resulting real and imaginary spectra can be used to perform phase correction. The variable scan speed is set to a nominal value of 4 seconds. This leads to a small electrical bandwidth which ensures a signal-to-noise ratio of at least 300. For the calibration of the instrument functions, the TANSO-FTS has fully redundant 1,31 μm

[1]The O_2 A-band is situated in the near-infrared wavelength region, although the official description of the TANSO-FTS band 1 assigns it to the Visible.

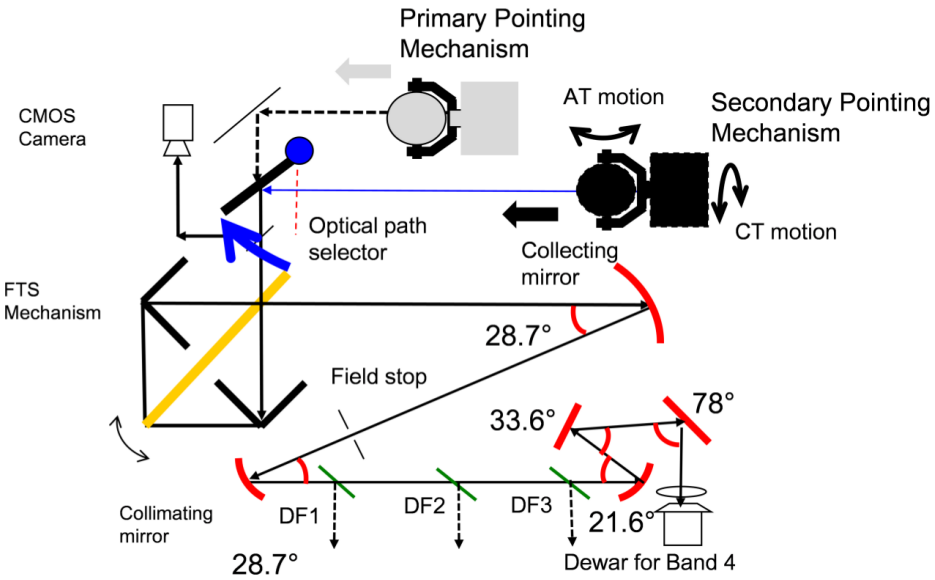

Figure 8: The design of the TANSO-FTS pointing system and optics after the switch to the secondary pointing mechanism. Figure from Kuze et al. [48].

distributed feedback lasers with a wavelength stability of $\frac{\Delta\lambda}{\lambda} < 10^{-7}$, which is secured by precise temperature control. The signal of band 4 is measured by a photoconductive HgCdTe detector [45].

5.3 Selected TANSO-FTS products

Different main data products based on GOSAT TANSO-FTS measurements are provided by JAXA/NIES/MOE. The products are subject to processing levels ranging from level L1A/B to level L4. The raw interferogram data is available in the L1A product, while the L1B product contains the measured radiance spectra. Further processing of the radiance spectra allows the determination of CO_2 and CH_4 concentrations (L2). The global concentrations of these greenhouse gases are found in the L3 data product. The corresponding sources and sinks are included in the L4 product [46].

5.3.1 Estimated Geolocation Data (EGD)

The primary pointing mechanism of TANSO-FTS has an offset, which cannot be detected with the angular resolver ("pointing offset" [48]). The measurement geolocation, provided in the FTS L1B radiance product, is calculated from the GOSAT viewing vector. It thus depends on both the AT and CT angles which show varying offset values ΔAT and ΔCT, respectively. To determine these offset values, the onboard CAM8 camera is used, which has the same line of sight as TANSO-FTS. Clear sky measurements of ground control points allow to estimate the latitude and longitude offsets, which can then be converted to ΔAT and ΔCT values [67]. The offset values are estimated two weeks after the measurement and used to recalculate the viewing vector for the L2 processing. The offset values are averaged to a certain degree, which leads to a remaining geolocation error of less than 1 km [48; 67].

Because the corrected geolocations are not included in the L1B product versions V160 and earlier, an additional product "Estimated Geolocation Data" is provided to users. It includes the uncorrected geolocations of the center of the TANSO-FTS IFOV together with the best estimated values including ortho-rectification and correction of the geometric error [68].

In January 2015, the switch to the secondary pointing mechanism took place, the optical path

selector now acts as a folding mirror. So far, no pointing offset has been detected since then by the angular resolver, but a CT angle dependent offset was created due to a slight misalignment of the optical path axis and the CT axis. However, the offset is well-characterized and the pointing accuracy is better than 0.5 km after correction, so that no CAM images are necessary for the correction [48].

5.3.2 FTS L1B

The measured intensities by TANSO-FTS are stored in the FTS L1B product whose newest version is named L1B V201.202. The intensity spectra cover the real and imaginary parts of the three SWIR bands (the O_2 A-band at 0.765 μm, weak CO_2 and CH_4 bands at 1.60 - 1.68 μm, the strong CO_2 band around 2.06 μm) of the two linear polarizations [47] and the TIR band (1.67 μm) [46]. Earlier versions provide the radiance in units of V/cm^{-1} which are converted to units of $W/cm^2/cm^{-1}/sr$ using provided radiance conversion factors. The new L1B dataset also includes the converted radiances [48].

In the processing of the radiance spectra, the radiometric degradation (described below) has already been considered. Quality degradations still arise from pointing fluctuations and saturation at the detector level. A master quality flag identifies most of the poor interferograms affected by pointing fluctuations of the measurement system. Next to standard ancillary information, such as geolocation information and solar angles, additional information has been included in the product. Among other data, it includes the Mueller matrices used for the individual consideration of two linear polarizations and the CAM pictures. Error sources have been further reduced in the processing of the interferograms and resulting spectra. The remaining errors are within the random noise level [48].

Spectral calibration

By comparing the measured spectra to the Fraunhofer line positions, a spectral shift was observed. At the same time, a degradation of the laser detection level could be observed. They are most likely related and probably caused by an arising misalignment of the laser beam after launch [46]. An absolute spectral calibration can be achieved by using the time dependent correction factor, as described in Kuze et al. [46]:

$$\nu_{new} = \nu/(a_0 \exp(a_1 t) + a_2) \tag{26}$$

The wavenumbers before the correction are ν, those after the applied correction ν_{new}. The coefficients a_1 and a_2 result from the calibration and the passed time since launch of the satellite is denoted by t. The change of the correction factor since launch is less than 10 ppm. According to Kuze et al. [48], the spectral calibration as well as the primary and secondary pointing system are very stable.

Radiometric calibration and degradation

Before the launch of GOSAT, the radiometric response was calibrated by inserting unpolarized light in an integrating sphere. On this basis, radiometric conversion factors were obtained to convert the measured response of V/cm^{-1} in radiances ($W/cm^2/sr/cm^{-1}$) [45]. These radiometric conversion tables are provided for all bands.

To estimate the time-dependent radiometric degradation of the SWIR radiances, different calibration activities contributed to the actual radiometric degradation factors (RDF). The resulting RDF show a fast degradation within the first two years in orbit which then slowed down [48].

The first activity uses solar calibration data from measurements on orbit. Once a month, the backside of the spectralon diffuser is exposed to the solar radiation during the passage over the

northern polar region. The backside of the spectralon diffuser is assumed to have no significant degradation as the front side is usually exposed. From the calibration data, a time dependent model is derived for each band. In contrast to the work of Kuze et al. [46], Yoshida et al. [107] also considers spectral and polarization dependencies. The degradation shows a time-dependent exponential decrease before slowing down. The results indicate the largest degradation for band 1, which is $\leq 2\%$ until 120 days after launch (Yoshida et al. [107]).

Additionally, yearly vicarious calibration campaigns have been performed since 2009 in Railroad Valley, Nevada, USA. Measurements of ground-based radiances, as well as radiosonde and MODIS[2] measurements are used as input for radiative transport simulations to calculate the radiance at the top of the atmosphere. Comparison of the modeled radiances and space-based GOSAT measurements at approximately the same time lead to RDF factors for each band. These are combined with the smooth degradation model from onboard solar calibration data as described above [47]. The results are shown in figure 9.

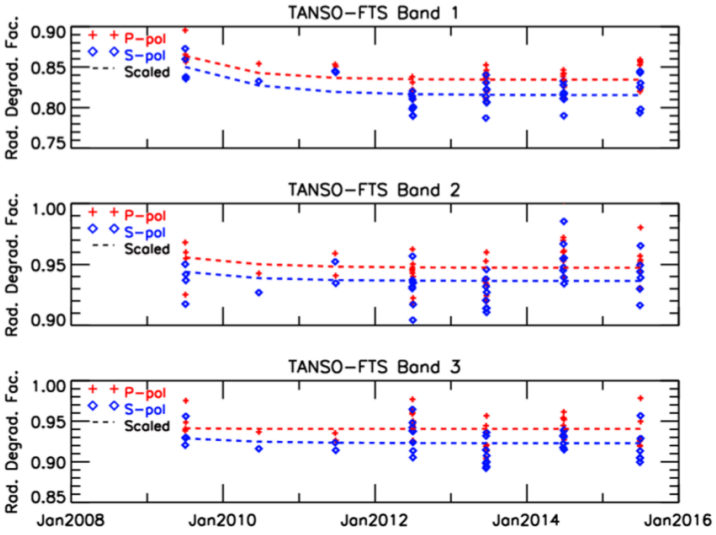

Figure 9: Radiometric degradation factors for TANSO-FTS for bands 1-3 and polarizations S and P. The symbols are the results from the annual vicarious calibration campaigns. The lines are the results of the model based on the solar calibration data. Figure from Kuze et al. [48].

According to Kuze et al. [48], the most likely source for the degradation lies in the FTS beam splitter efficiency. The evolution of recent radiance degradations is very slow. Yearly vicarious calibration campaigns after 2012 confirm the combined degradation model which can be used on the FTS dataset of the entire mission. Also the change to the secondary pointing system showed no significant difference to the previous results. The degradation model is applied to the radiance spectra in FTS L1B V201 [48].

Approximation for full radiance spectra

TANSO-FTS performs measurements of two linear polarizations I_P and I_S, which are defined in a specific reference system of GOSAT [70]. The polarized radiances can be compared to simulated values of the Stokes vector by applying a polarization model, which is described in detail in references [45] and [70]. If only the first element of the Stokes vector is of interest, the intensity can be approximated to [70]:

$$I = \frac{1}{2}(I_P + I_S) \tag{27}$$

[2]Moderate resolution imaging spectroradiometer on the Terra and Aqua satellites [106]

5.3.3 L2 Input Data Set (IDS)

All information needed for higher level processing of the TANSO-FTS data (L2) is summarized in the FTS L2 Input Data Set. Each measurement is defined by a unique scan ID, which includes the date and time of the measurement, the satellite pass and scene numbers as well as the performed observation mode. Besides the measured spectra, the dataset includes the (pointing error corrected) geolocation information on the measurement and satellite positions, solar angles, cloud cover of the FTS FOV and quality reports. Details can be found in the corresponding product format description [64].

5.4 TANSO-FTS instrument line shape function (ILSF)

To accurately compare simulated spectra to the TANSO-FTS measurements, the simulations have to be convolved with the instrument line shape function (ILSF) of the instrument. The TANSO-FTS ILSF has been numerically simulated for all bands at different wavenumbers to consider the wavelength dependency of the ILSF [46]. According to Kuze et al. [46], the shape of the ILSF has been very stable during the time in orbit. Here, the *ILSF2015* model for the TANSO-FTS L1B Version V201 is used, which is applicable to the entire dataset of observations[3]. The central parts of the ILSF models for band 1 are shown in figure 10:

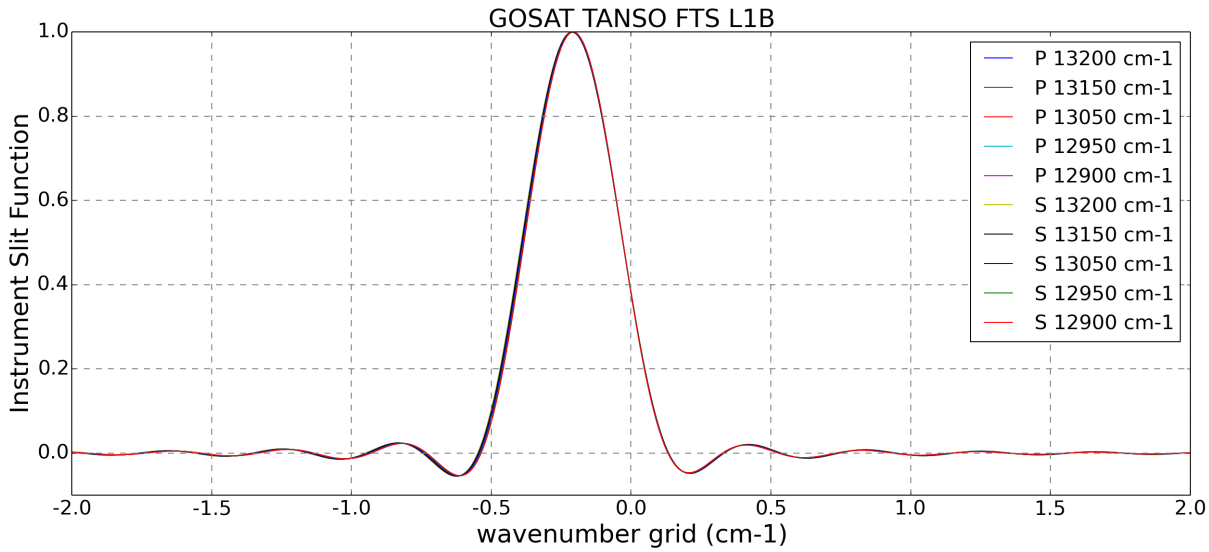

Figure 10: The central parts of the TANSO-FTS ILSF for band 1 at different wavenumbers and polarizations P and S. The wings extend to ± 50 cm^{-1}.

All simulated McArtim spectra are convolved with a mass-centered, averaged and approximated version of the above ILSF:

- The ILSF of the P- and S- polarization measurements are averaged because only the combined radiances (first element of the Stokes Vector) are used in this study.
- The wavelength dependence is neglected by averaging all provided ILSF within band 1.
- To reduce the number of necessary simulation wavelengths, the wings of the averaged ILSF are cropped at a distance of 1.5 cm^{-1} from the center.

To test the accuracy of the wavelength averaging, each radiance value of a simulated spectrum is convolved with a wavenumber dependent ILSF derived from a linear interpolation between the ILSFs at 12900 cm^{-1}, 12950 cm^{-1}, 13050 cm^{-1}, 13150 cm^{-1} and 13200 cm^{-1}. The result

[3]The 'ILSF2015' data was acquired from NIES (https://data.gosat.nies.go.jp) on July 20, 2016.

is compared to the same simulated spectrum convolved with the averaged ILSF. As can be seen in figure 11, the induced errors are very small.

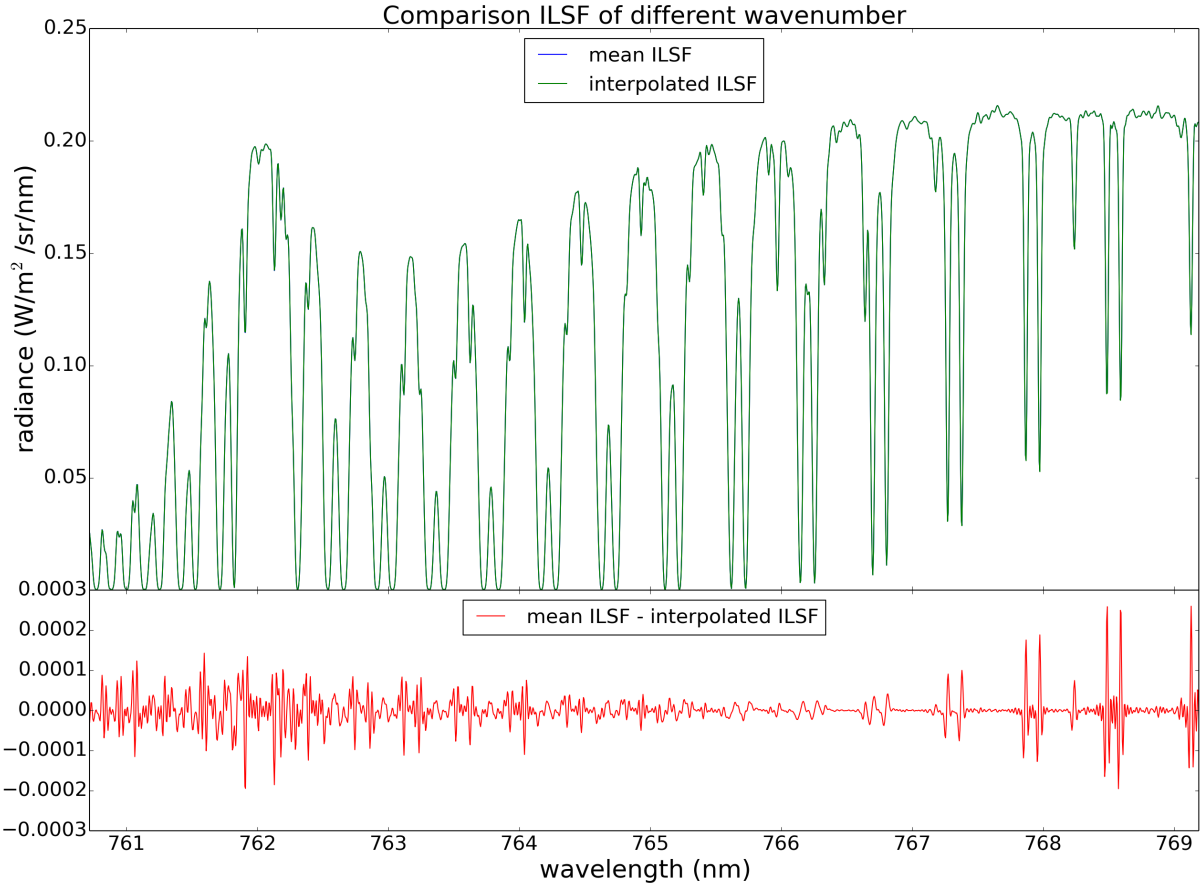

Figure 11: Blue: TANSO-FTS spectrum convolved with the polarization and wavenumber averaged ILSF. Green: Same TANSO-FTS spectrum convolved with the polarization averaged ILSF but considering the wavenumber dependency. Red: Absolute difference between the two convolved spectra.

To minimize the number of necessary radiance simulations, the wings were cropped to values ranging from 0.15 cm^{-1} to 5 cm^{-1} at each side of the ILSF center. A simulated spectrum was convolved with each shortened ILSF and compared to the same spectrum convolved with the ILSF with full wings. The final choice is a reduction of the wings to 1.5 cm^{-1}. The test results are shown in figure 12, again the induced errors are very small.

5.5 TANSO-CAI

The Cloud and Aerosol Imager TANSO-CAI serves to determine the cloud and aerosol cover within the TANSO-FTS field of view during the measurements. It consists of an optics unit with three telescopes, analogue signal-processing and a redundant electronics unit. The radiometer measures pixel radiances in four different wavelength bands that are centered at 380 nm, 674 nm, 870 nm with a bandwidth of 20 nm (band 1-3) and at 1600 nm with a bandwidth of 90 nm (band 4). Bands 1-3 have a projected field of view of 0.5 km (2048 pixel) and the intensities are measured by Si detectors. The instantaneous field of view of band 4 is 1.5 km (512 pixel) and uses a lnGaAs detector for the measurements. The signal-to-noise ratio for all bands is smaller than 200. Because horizontal formations of aerosol are usually larger than the TANSO-FTS field of view, TANSO-CAI has a wider swath than TANSO-FTS. The measurement swath in cross track direction is ± 36.1° for band 1-3 and ± 30° for band 4 [45].

Like TANSO-FTS, TANSO-CAI also features different classes of data products, provided by

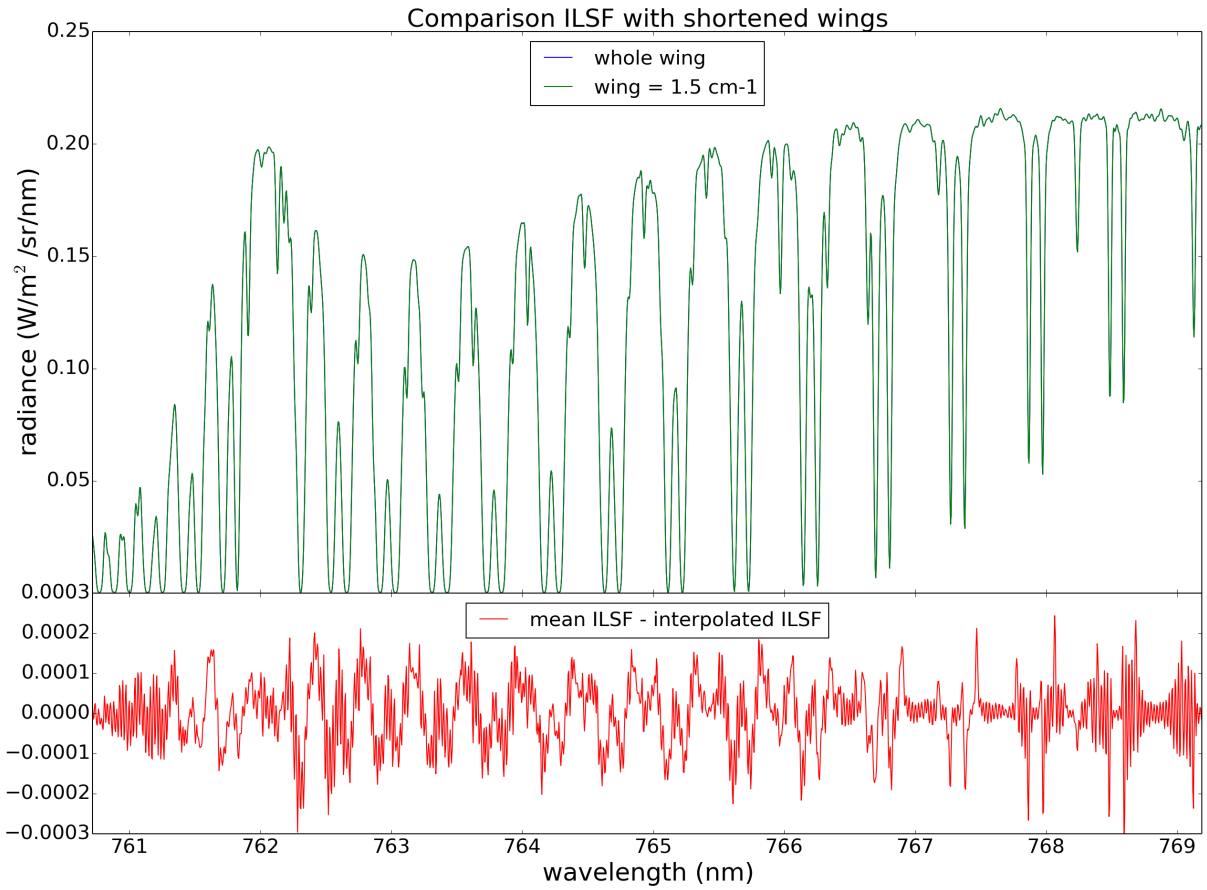

Figure 12: Blue: TANSO-FTS spectrum convolved with the polarization and wavenumber averaged ILSF with wings cropped to ± 1.5 cm^{-1}. Green: Same TANSO-FTS spectrum convolved with the polarization and wavenumber averaged ILSF with complete wings. Red: Absolute difference between the two convolved spectra.

JAXA/NIES/MOE. The unprocessed signal is available from level L1A data, while the level L1B product contains the corresponding radiances (spatially resampled in L1B+). Higher data products are the L2 datasets (including cloud flag, cloud properties and aerosol properties) and the L3 datasets with global distributions of radiances, reflectances and NDVI (normalized difference vegetation index) [46; 63].

5.5.1 L2 cloud property product

The TANSO-CAI L2 cloud property product is based on the TANSO-CAI L2 cloud flag and the TANSO-CAI L1B data. The calculated cloud properties include the dimensionless cloud optical depth and the effective cloud particle radius. The spatial resolution of the pixels is reduced compared to the TANSO-CAI L1B input data. The pixel grid is classified into 3x3 pixels and the central pixels are used for the determination of cloud properties.

For all pixels determined as cloudy in the TANSO-CAI L2 cloud flag, the cloud optical depth and effective particle radius are determined from the radiometric corrected TANSO-CAI radiances of band 2 and band 4, using look-up-tables and a radiative transfer code (Nakajima and Tanaka, 1988 [62]). Influences on the measured radiances like light reflected from the Earth's surface and transmitted by the cloud, are removed by generalized assumptions.

The cloud optical depth and the effective particle radius are determined as follows: The CAI band 2 measures radiances at a wavelength of 670 nm, where the imaginary index of refraction of water is very small ($\approx 10^{-8}$). Consequently, the amount of radiance absorbed inside the cloud

is negligible. The higher the cloud optical depth, the more radiance will be measured by the satellite due to an increased upward reflection of radiance. A saturation of the observed radiance is expected at values of the cloud optical depth of about 70.

The CAI band 4 radiance is measured at 1600 nm, where the imaginary part of the complex index of refraction of water is 10^{-5} - 10^{3} times larger than the one of visible light. The near-infrared light will thus encounter absorption inside the cloud. A higher cloud particle radius enhances absorption and thus leads to less detected radiance by the satellite.

The look-up table simulations show that the detected radiances of band 2 and 4 are higher for increasing cloud optical depth and that the radiance of band 4 decreases with increasing cloud particle radii.

The above information as well as further details on the product and the retrieval process can be found in the "Release Note on the TANSO-CAI Level 2 Cloud Property Product" [71]. The product information describes the CAI L2 cloud property version 1.00. In our retrieval, a newer version 2.00 is used alongside with version 1.00, depending on the availability.

6 Radiative transfer simulations and photon path statistics

6.1 McArtim

For the simulation of the O_2 A-band radiances, the radiative transport model McArtim (Version 3, Release from 28.06.2014) is used. McArtim (Monte Carlo Radiative transfer inversion model) is a backward Monte Carlo radiative transport code. It was designed to support remote-sensing applications in the spectral range between the ultra violet (UV) and the near-infrared (NIR). A detailed description of the mathematical methods and their implementation is given in Deutschmann [24] and [25]. Here, a short introduction is provided, using the information and notations from Deutschmann et al. [25].

A basic description of the radiative transfer in the atmosphere is provided by the scalar (unpolarized) radiative transfer equation (RTE), which describes the transport of non-interacting photons in a medium. Polarization effects of the electromagnetic field are considered in the vector RTE. In this thesis, polarization effects are taken into account, but only the first element of the Stokes vector is used from the simulated radiances.

Assuming monochromatic conditions, the RTE can be expressed as

$$\vec{\omega}\vec{\nabla}I(\vec{r},\vec{\omega}) = -\epsilon(\vec{r})I(\vec{r},\vec{\omega}) + \frac{\vec{\epsilon_S}(\vec{r})}{4\pi}\int_{4\pi} I(\vec{r'},\vec{\omega'})P(\vec{r},\vec{\omega}\cdot\vec{\omega'})d\vec{\omega'} + S(\vec{r}) \quad . \tag{28}$$

The gradient of the radiance I at position $\vec{r}$ and the direction $\vec{\omega}$ is described by the differential light attenuation by extinction (first term on the right-hand side of eq. 28), the integral over the scattered radiance to position $\vec{r}$ (second term) and a source density $S(\vec{r})$. $\epsilon(\vec{r})$ is the extinction coefficient, $\epsilon_S(\vec{r})$ the scattering coefficient and $P(\vec{r},\vec{\omega}\cdot\vec{\omega'})$ the (isotropic) scattering phase function. The source density $S(\vec{r})$ includes the Sun as photon source and the thermal radiation emitted from the Earth (which is negligible within the considered spectral range and thus not included in the further description).

The RTE can also be expressed in an integral form (IRTE) using the definition of the collision density $f(\vec{r},\vec{\omega}) = \epsilon(\vec{r})I(\vec{r},\vec{\omega})$:

$$f(\vec{r},\vec{\omega}) = \int_M\int_{4\pi}\delta_{\vec{r}}(\vec{r},\vec{\omega})\frac{k_p[(\vec{r'},\vec{\omega'})\rightarrow(\vec{r},\vec{\omega})]}{|\vec{r}-\vec{r'}|}f(\vec{r'},\vec{\omega'})d\vec{\omega'}d\vec{r'} + \Psi(\vec{r},\vec{\omega}) \tag{29}$$

or in operator form:

$$f = Kf + \Psi \tag{30}$$

The first term on the right-hand side of eq. 29 includes the transition density k_p which can be seen as the probability of the photon performing a transition from $(\vec{r'},\vec{\omega'})$ to $(\vec{r},\vec{\omega})$ through a scattering event and a free path length. The spatial region M includes both the Earth and the Sun. The initial transition density $\Psi(\vec{r},\vec{\omega})$ (second term) defines the initial density of radiation at $(\vec{r},\vec{\omega})$ after being subject to extinction within the Earth's atmosphere.

The solution of the IRTE leads to the Neumann-Series

$$f = \sum_{n=0}^{\infty} K^n\Psi \quad (\text{if } ||K|| < 1) \tag{31}$$

where the summation elements K show to have a direct relation to the order of the scattering. The principle of reprocity allows to reverse the time direction of the radiative transfer and results in the so-called adjoint IRTE (leading to an adjoint collision density $f^\dagger$):

$$f^\dagger = K^\dagger f^\dagger + \phi \tag{32}$$

with ϕ being a photon sink represented by the detector. The solution is found analogous to eq. 30:

$$f^{\dagger} = \sum_{n=0}^{\infty} K^{\dagger n} \phi \tag{33}$$

The solution can be numerically estimated by using a ray tracing algorithm which simulates the radiative transport process.

The backward Monte Carlo method, used in McArtim, is based on the reprocity theorem: Instead of a photon source term, a sink term (detector) is introduced. For the simulation of the radiative transport, it follows that the trajectory starts at the detector and is sampled until an absorption event of the photon occurs or the photon leaves the atmosphere.

The simulation environment in McArtim is built on a plane-parallel or spherical grid. In each bin, pressure and temperature are defined and used to calculate the air density assuming the ideal gas law. Optical properties for each bin are the total extinction coefficient ϵ, the scattering coefficients (σ_{Ray}: Rayleigh scattering, σ_{RRS}: Rotational Raman scattering, σ_p: particle scattering) and necessary parameters for the phase function calculation. Also used are the absorption coefficients σ_{a_i} of gases present in the atmosphere and the real part of the index of refraction of air $\vec{n}_{air}$.

The total extinction coefficient is given by

$$\epsilon = \vec{n}_{air}(\sigma_{Ray} + \sigma_{RRS}) + \sum_{i}^{particles} \epsilon_{p_i} + \sum_{i}^{absorbers} n_{a_i}\sigma_{a_i} \quad , \tag{34}$$

the total scattering coefficient is

$$\epsilon_S = \vec{n}_{air}(\sigma_{Ray} + \sigma_{RRS}) + \sum_{i}^{particles} \varpi_{0i}\epsilon_{p_i} \quad , \tag{35}$$

where ϖ_{oi} is the single scattering albedo for particle i.

The raytracing method in McArtim relies on a Markov process where the position $\vec{x}_{n+1}$ does only depend on the preceding position $\vec{x}_n$. The first event is placed at the detector position $\vec{r}^*$ and has an initial propagation direction $\vec{\omega}^*$. The transition from $\vec{x}_n$ to $\vec{x}_{n+1}$ includes the simulation of the single scattering albedo $\vec{\varpi_0}(\vec{r}) = \frac{\epsilon_S(\vec{r})}{\epsilon(\vec{r})}$, the selection of a scattering angle from an effective phase function and the mean free path traveled after the scatter event. The value of the single scattering albedo is determined by drawing a random number $\alpha_{\vec{\varpi_0}(\vec{r})}$. Absorption takes place, if $\alpha_{\vec{\omega_0}(\vec{r})} > \vec{\omega_0}(\vec{r})$, and defines the end of the trajectory. If the photon is subject to a scattering event, the scattering object is determined (e.g. air molecule or particle) and the scattering angle is selected from the effective phase function, both based on the use of random numbers. The free path length of the photon following the scatter event is determined by sampling an optical thickness $\tau_l(\alpha_l)$ and then adding the optical depths of small space segments along the chosen propagation direction such that τ_l is not exceeded. If τ_l is larger than the summed optical depth through the atmosphere, the photon escapes the atmosphere, which also ends the trajectory. Photons can also be absorbed or scattered at the Earth surface. The absorption probability relies on the surface albedo. If a scattering event is performed, the trajectory direction is determined by assuming a Lambertian reflector. A backward trajectory of a photon is schematically shown in figure 13.

The derived physical quantities include the radiance, light path integrals, heating rates and actinic fluxes as well as the Jacobians of these parameters with respect to optical parameters. The radiance I_* of a single photon trajectory can be expressed by

$$I_* \approx \sum_{n=0}^{N} c(\vec{r}_n) w_n \tag{36}$$

with

$$w_n = \exp(-\tau_{*\to n}) P(\vec{r}_n, \mu_n^*) \quad . \tag{37}$$

The element $c(\vec{r}_n)$ is approximately the area of the sun disc solid angle from the position $\vec{r}_n$. The weight w_n is composed of the probability that the light is transmitted between the Sun and the position $\vec{r}_n$ and the probability that a scattering event with the cosine representation of the scattering angle μ_n takes place. N is the number of scattering events during the trajectory.

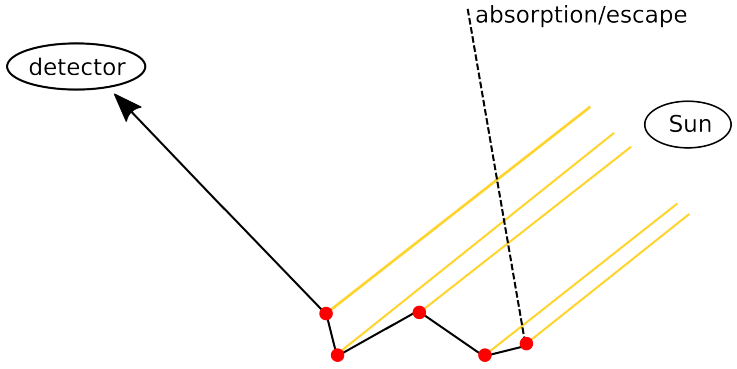

Figure 13: A backward trajectory of a photon. Originating from the detector position, the photon is subject to scattering events until it is absorbed or ejected into the atmosphere (adapted from Deutschmann et al. [25]).

6.2 Input parameters

6.2.1 Atmosphere

In order to characterize the model atmosphere, atmospheric profiles are included into the simulation environment. McArtim is configured to simulate in three dimensions within a one-dimensional spherical atmosphere. The planned simulations include tropospheric clouds below 8 km, which requires a fine altitude grid in the lower atmosphere. The altitude grid is chosen to include

- 0.2 km vertical altitude bins between 0 and 15 km,
- 2.5 km vertical altitude bins between 15 and 50 km,
- 5.0 km vertical altitude bins between 50 and 80 km as well as
- an additional vertical altitude bin between 80 and 1000 km.

The last altitude bin spans over 920 km and assures that the detector altitude ($\approx$ 666 km for GOSAT) is included in the model domain. Necessary atmospheric parameters are the pressure, the temperature, the O_2 number density and the O_2 absorption cross sections. The pressure and temperature values are extracted from an atmospheric profile of the US Standard Atmosphere[4] [22] and gridded to the chosen altitude bins. The number mixing ratios of O_2 are also extracted from the mentioned profile and converted to number densities for each altitude bin by using the altitude dependent air density n_{air}[5]:

$$n_{O_2}(cm^{-3}) = n_{air}(cm^{-3}) \cdot n_{O_2}(ppm) \cdot 10^{-6} \tag{38}$$

[4] datafile included in SCIATRAN [84]

[5] data from T. Wagner, personal communication

6.2.2 Solar irradiance

To characterize the incoming solar radiation, a solar irradiance spectrum is used. Sun reference spectra measured by GOSAT are not available, thus a solar reference spectrum from other measurements is used. The chosen reference spectrum is the disc-integrated solar spectrum (version 27.01.2016) from Toon [97]. The spectrum was created by extracting a list of solar line positions from different measurements (e.g. ATMOS, Kitt Peak) and fitting the Earth-atmosphere related absorptions using the HITRAN database. The resulting line list together with a solar line shape function led to the provided pseudo-transmittance spectrum. The available disc-integrated spectrum is suitable for measurements using reflected sunlight. In order to obtain the solar irradiance, the solar Planck function has to be multiplied to the pseudo transmittance. This, however, proves to be complicated because of the temperature dependence of the Planck function [97]. As only a small wavelength region is considered within this thesis, a fit on another solar irradiance spectrum is performed for the calibration. The pseudo-transmittance spectrum I_{PTS} is converted to the solar irradiance I_{irr} by fitting the wavelength (λ) region between 755 and 775 nm to the Kurucz spectrum I_K [44] using a linear function. The result is given below and illustrated in figure 14.

$$I_{irr} = (0.260 - 0.003 \cdot \lambda) \cdot I_{PTS} \tag{39}$$

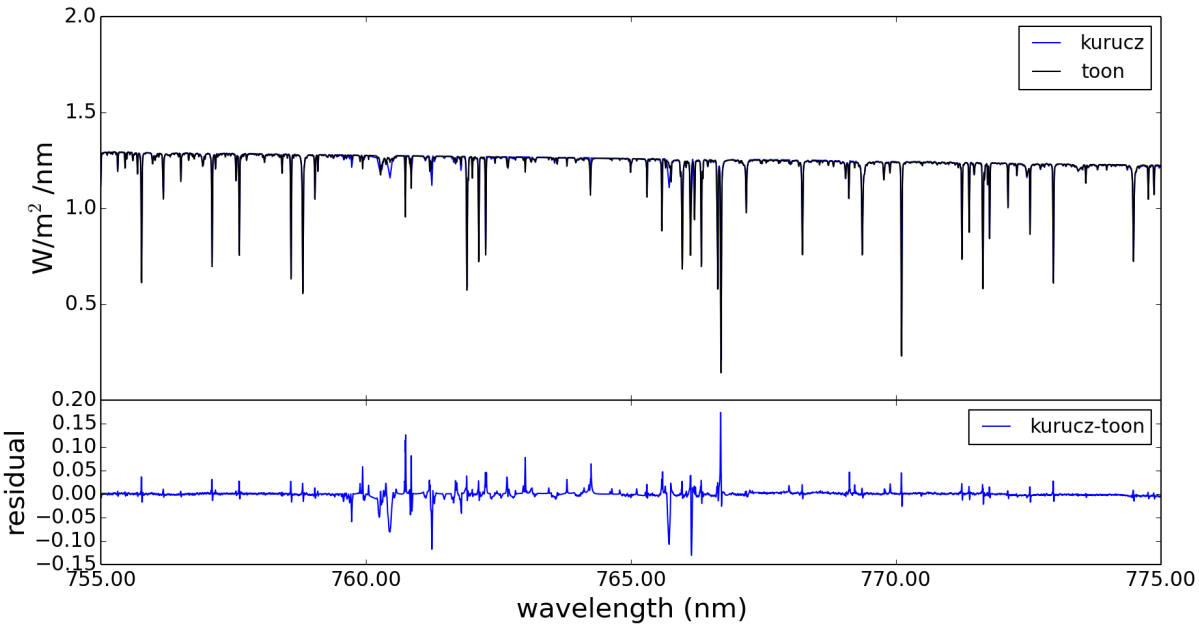

Figure 14: Top: Solar irradiance spectrum of Kurucz [44] (blue) and fitted irradiance spectrum using the pseudo-transmittance from Toon [97] (black). Bottom: Residual between both spectra.

6.2.3 O_2 absorption cross sections

The contributions of other trace gases besides atmospheric oxygen are negligible in the considered spectral range. Within this work, it it thus sufficient to accurately describe the concentration profile and absorption cross sections of O_2. The absorption lineshapes of molecular oxygen are subject to broadening effects from collisions and the brownian movement of the molecules. This is traditionally described by the Voigt Line shape [78]. Additionally, small effects as line mixing (LM) and collision induced absorption (CIA) arise (see section 4.1.5). For the oxygen A-band, the resulting line shape profiles have been modeled by Tran et al. [98; 99].

In this study, altitude and wavelength dependent O_2 absorption cross sections within the A-band are calculated from this model using a retrieval algorithm from J.-M. Hartmann and H. Tran, adapted by A. Butz (personal communication). Additionally, a scaling parameter of 1.03 is applied to the resulting absorption cross sections, as proposed by Butz et al. [11]. The cross sections are retrieved for wavenumbers between 12800 - 13400 cm^{-1} with a step size of 0.05 $cm^{-}1$. This corresponds to an approximate wavelength range of 746-781 nm and a wavelength step of $\delta\lambda \approx 0.003$ nm. The height grid was chosen such that the middle of each model altitude layer is used for the retrieval of the O_2 absorption cross sections.

The extracted lineshapes for different altitudes within a small region of the oxygen A-band are shown in figure 15 (for illustration purposes, the resolution is increased to 10^{-5} cm^{-1}).

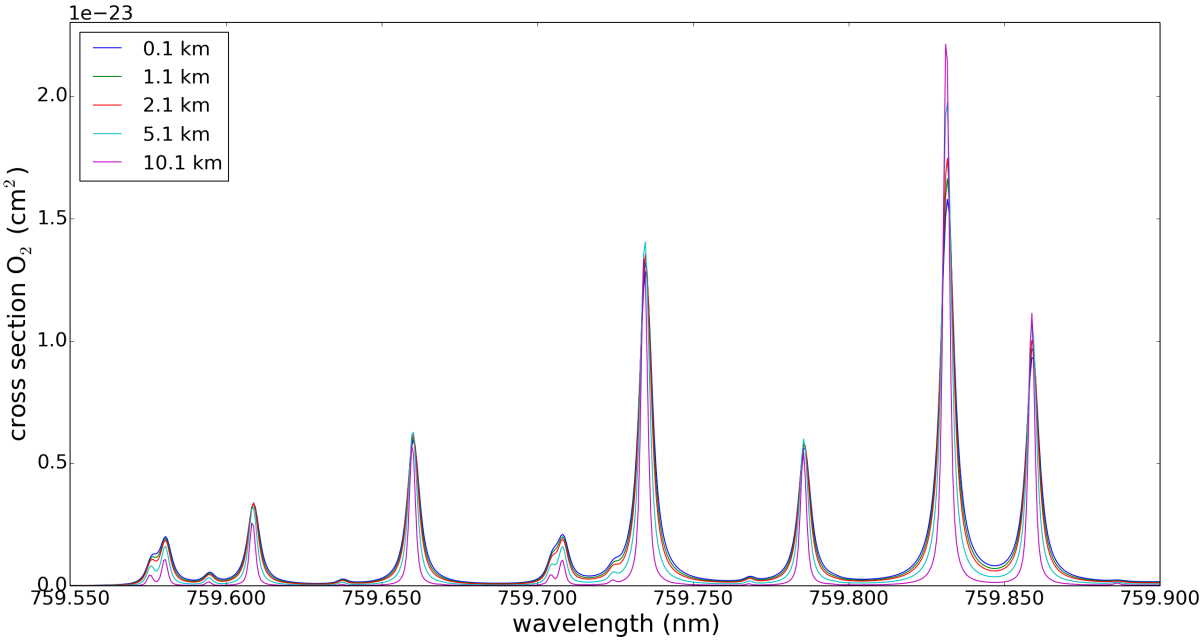

Figure 15: Absorption cross sections of O_2 for different altitudes using the retrieval algorithm mentioned in the text.

6.2.4 Sun-satellite geometry

In order to correctly set up the radiative transfer simulations, the model geometry has to be adapted to the measurement geometry of each TANSO-FTS measurement. Detailed information on solar and satellite angles are provided within the TANSO-FTS data products in relation to the center of field of view. The radiative transfer model McArtim assumes a spherical Earth and relates all input geometry to the satellite position, it follows that the measurement geometry has to be converted accordingly.

Coordinates (latitude and longitude)

All further angles rely on the latitude (lat) and longitude (lon) coordinates of both the center of field of view and the satellite position. The coordinates of the measurement location are provided within the GOSAT data product, as well as the cartesian coordinates of the satellite position. This information is converted [104] to sub satellite coordinates of GOSAT and the satellite altitude assuming a WGS-84 ellipsoid (e.g. [31]).

Solar angles

The solar zenith and solar azimuth angles at the satellite position are obtained from the *pyorbital* module [23] by providing the geographic coordinates and the measurement time.

Elevation angle

The elevation angle -α_{el} can be calculated using the sub satellite points SSP (lon_{SSP}, lat_{SSP}) and the coordinates of the field of view FOV (lon_{FOV}, lat_{FOV}), following [26; 88]. The relevant distances and angles are given in fig. 16.

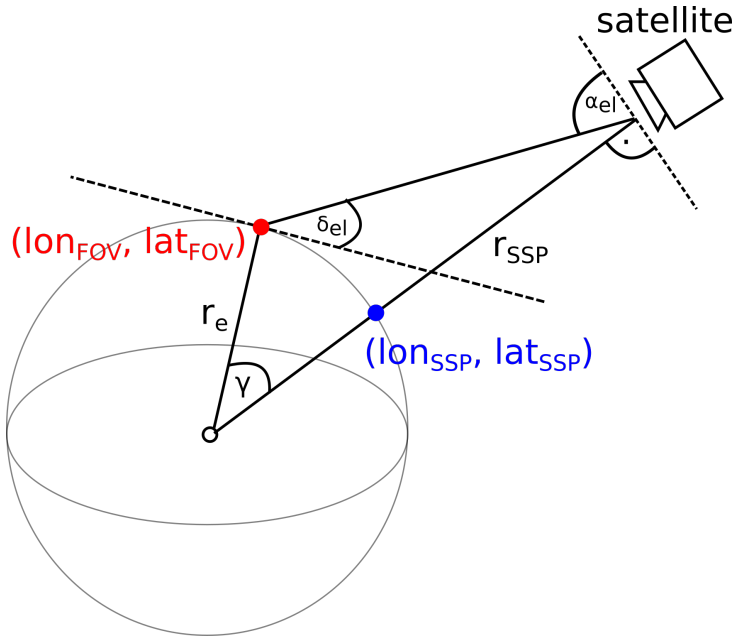

Figure 16: McArtim satellite elevation angle -α_{el} as well as further angles and distances for it's derivation.

The angle $\gamma \geq 0$, separating the vectors of the Earth center to the sub satellite point of GOSAT and the Earth center to the projected field of view, can be determined via the relation:

$$\cos(\gamma) = \cos(lat_{FOV})\cos(lat_{SSP})\cos(lon_{SSP} - lon_{FOV}) + \sin(lat_{SSP})\sin(lat_{FOV}) \tag{40}$$

Using the value of γ, δ_{el} is determined by:

$$\cos(\delta_{el}) = \frac{\sin(\gamma)}{\sqrt{1 + (\frac{r_e}{r_{SSP}})^2 - 2\frac{r_e}{r_{SSP}}\cos(\gamma))}} \tag{41}$$

The McArtim geometry description is relative to the detector position. The elevation angle is expressed by -α_{el}, where a value of -90° defines a nadir measurement.

$$\alpha_{el} = (\gamma + \delta_{el}) \tag{42}$$

Azimuthal angle

The azimuthal angle θ of the vector connecting the sub satellite position to the observation location on Earth is calculated via the relation [32]:

$$\tan(\theta) = \frac{\cos(lat_{SSP}) \cdot \sin(lon_{FOV} - lon_{SSP})}{\cos(lat_{SSP}) \cdot \sin(lat_{FOV}) - \sin(lat_{SSP}) \cdot \cos(lat_{FOV}) \cdot \cos(lon_{FOV} - lon_{SSP})} \tag{43}$$

6.2.5 Cloud parametrization

Within this study, only water clouds are considered. All modeled clouds have one or two layers with base and top altitudes fixed to the altitude grid and extend indefinitely in the horizontal direction. Within the layers, the cloud optical properties are assumed to be homogeneous and are defined by the Henyey-Greenstein phase function (see section 4.3.3). If not noted otherwise, a single scattering albedo of 1 and an asymmetry parameter of 0.85 are used. The cloud optical depth is individually determined for each simulation, which is described in section 8.2.

6.2.6 Ancillary simulation options

The simulations are performed on a wavelength grid of 0.005 nm and the first element of the Stokes vector is used for the simulated radiance. Within the output, these radiances are given twice - once considering polarization and once neglecting it. The precision of the simulation is 1% of the first element of the Stokes vector where polarization of the radiance is not considered. For the retrieval within this thesis, the output considering polarization effects is used.

The surface albedo of the model Earth depends on the surface type. In case of water surfaces, an albedo of 0.03 is assumed. In case of land surfaces, the Filled Land Surface Albedo Product from MODIS measurements (product from combined datasets during 2000-2004) [2; 61] serves to estimate a starting value for the ground reflectance at the measurement location.

6.3 Rotational Raman scattering

The simulations account for rotational Raman scattering (RRS, see section 4.3.2) in the O_2 A-band (details on RRS within McArtim are given in Deutschmann [24]). The filling-in effect of the RRS is one of the newer features of McArtim. A study of Vasilkov et al. [101] was used to test the performance of the RRS implementation within the O_2 A-band for an example clear sky case.

Vasilkov et al. [101] investigated the influence of Rotational Raman scattering in the oxygen A- and B-bands by simulating spectra for different atmospheric situations and instrument resolutions. The simulations were performed for different solar zenith angles and satellite measurement geometries. The resulting radiance spectra were convolved with a Gaussian function with FWHM of 0.03, 0.1, 0.5 and 1.0 nm to simulate instruments of different spectral resolution. In their work, it was found that in the case of clear sky radiances above a lambertian surface of 0.3 albedo, the filling-in factor (defined as the percentage-ratio between the inelastic and elastic radiance [39]) shows values up to 2% in the deep absorption lines of high spectrally resolved spectra. In presence of low clouds, the effect approximately doubles when the cloud optical depth is increased from 1 to 50 because the amount of RRS is dependent on the photon path. Although high clouds mostly reduce the filling-in by shielding part of the atmosphere, the effect of the RRS can be slightly increased compared to a clear sky scenario. In summary, the spectral effect of RRS is small but not negligible [101].

One of the results is presented in the left plots of figure 17 (as published by Vasilkov et al. [101]). Here, the top of atmosphere inelastic and elastic radiances were calculated using a surface albedo of 0.3, a solar zenith angle of 45° and a nadir direction of the satellite detector. The solar irradiance was taken from Chance and Kurucz [21]. In order to test the RRS simulations within McArtim, the simulation configuration was adapted as closely as possible to the study of Vasilkov et al. [101] [6] to reproduce the results (see right plots of figure 17). The comparison of both simulations shows that the results do not differ much and we can thus assume that the RRS effect in the McArtim simulations of the O_2 A-band is well considered.

[6] One difference is the source of the O_2 absorption cross sections.

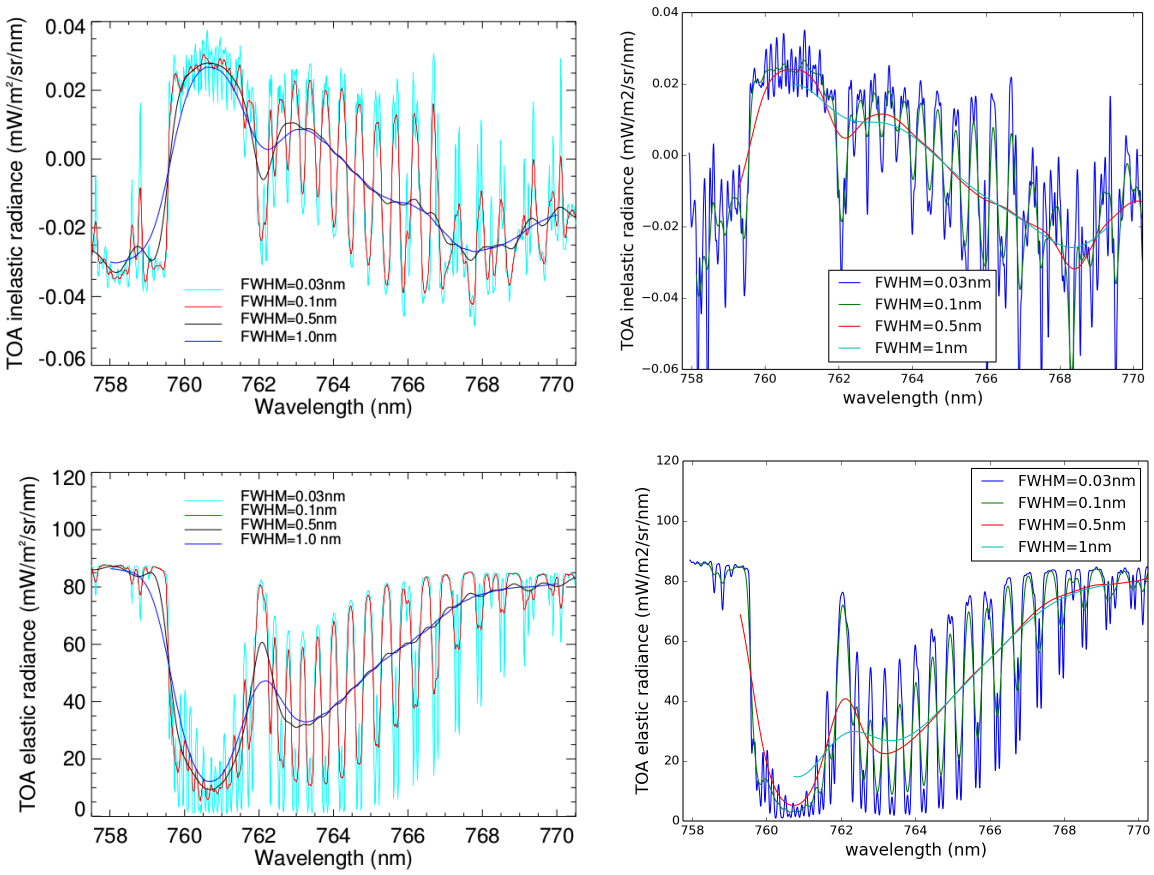

Figure 17: Left: Inelastic (top) and elastic (bottom) radiances as published in Vasilkov et al. [101]. Right: Inelastic (top) and elastic (bottom) radiances calculated with McArtim.

6.4 Photon event information

During the backward Monte Carlo process of the simulation, subsequently simulated photons ("photon ensembles" in the following) start at the position of the detector and pursuit a trajectory through the model atmosphere, therein being subject to scattering and absorption events. At each scatter event, the ensemble is forced into the direction of the model Sun with a calculated probability, which concerns thus part of the ensemble. The remaining part of the ensemble continues the trajectory through the atmosphere until it is terminated by e.g. an absorption event. Bearing the backward simulation routine in mind, it can be imagined that the photon ensemble "arrives" from the Sun direction with a given probability and pursuits the scattering trajectory to the detector position.

McArtim includes the option to create an additional simulation output ("trajectory output") that contains all information about the simulated scattering or absorption events, such as the event position in the model atmosphere or the type of events. The event information is saved in a consecutive way but can be separated to retrace individual photon trajectories[7] and combined to derive quantities like the air mass factor or photon path length statistics. The computational routine to effectively separate these absorption events was developed in cooperation with R. Sörensen from MPIC.

[7]Necessary information was obtained from the (former) web presentation of McArtim and personal communication with T. Deutschmann.

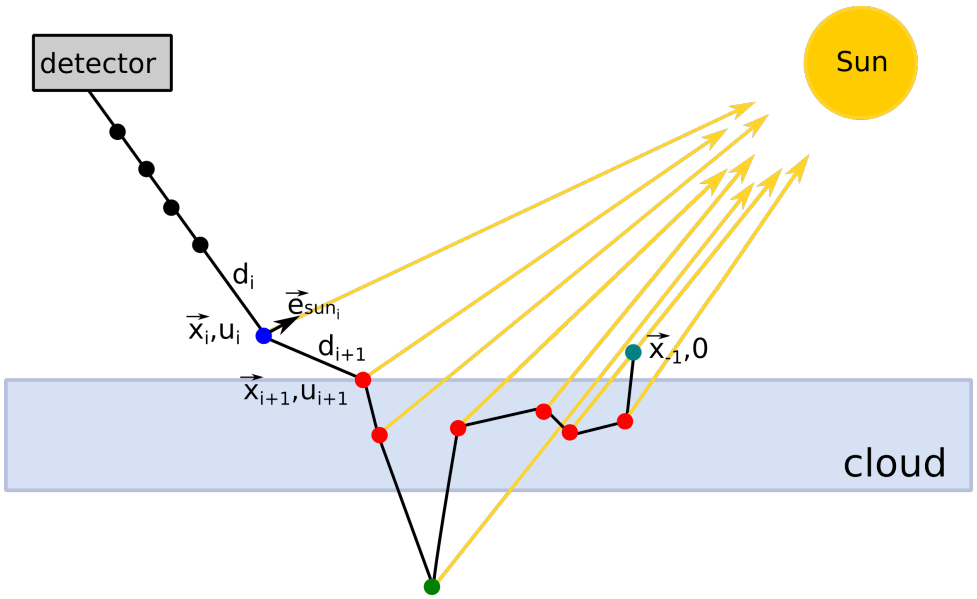

Figure 18: Schematic overview of a photon trajectory through the atmosphere. Rayleigh scattering events are shown in blue, particle scattering events in red, a ground scattering events in green and the final absorption event in cyan. In reality, the arrows pointing from the scattering events to the model Sun are nearly parallel due to the large distance between Sun and Earth.

The derived photon path length quantities are the photon penetration altitude, the photon path length distribution below a reference altitude (here: 10 km) and altitude resolved photon path length distributions. Furthermore, the information can be used to create visualizations of individual simulated photon paths within the model atmosphere. In principle, the output of the photon ensemble can be created for each simulated wavelength. Because of the resulting data volume, the information within this study is restricted to the very weak oxygen absorption at 772.145 nm. At this low absorbing wavelength, photons have the possibility of experiencing very long trajectories.

The event information of interest is as follows:

- the position of the event in three-dimensional space: $\vec{x}_i$
- the distance between two consecutive events: $d_i = |\vec{x}_i - \vec{x}_{i-1}|$
- the type of event: Rayleigh scattering, ground scattering, particle scattering, rotational Raman scattering or absorption
- the contribution of the scattering event to the radiance: $\vec{u}_i$
- the direction from the Sun to the scattering event: $\vec{e}_{sun,i} = (-1, \theta_{sun}, \phi_{sun})_i$

6.5 Photon penetration altitude

In a clear atmosphere, most photons in the near-infrared spectral range will travel unhindered to the Earth's surface where they get either absorbed or reflected back to the atmosphere. In presence of clouds, the cloud will usually act as a reflection layer as well as induce multiple scattering events in the vicinity of the cloud. Consequently, only part of the photons will enter the lower atmospheric layers below the cloud base. To create a measure of penetration of the detected photons into the atmosphere, the photon penetration altitude is defined as the lowest atmospheric layer reached by the modeled photons. For one photon ensemble, this is illustrated in fig. 19, where the magenta points indicate the different penetration altitudes. The three leftmost penetration altitudes correspond to the parts of the photon ensemble, which are

scattered into the sun direction at the event position. The penetration altitude at the ground layer, however, corresponds to all parts of the photon ensemble directed to the Sun at this and subsequent event positions because it is the lowest altitude of the remaining trajectory.

The penetration altitude statistics allow to visualize cloud effects, e.g. the radiance attenuation from different cloud optical depths or how much of the measured intensity was probing the bottom layer of a multi-layered cloud. In the later presented visualizations, the penetration altitudes are classified into the altitude layers defined by the altitude grid of the simulation environment and plotted against the corresponding fractions of the detected radiance.

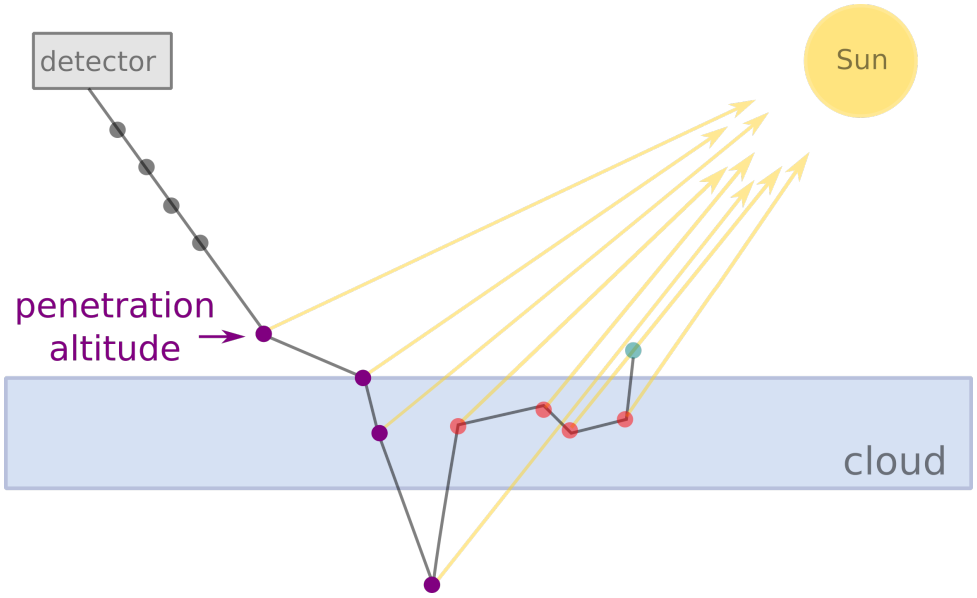

Figure 19: In the background, parts of picture 18 are shown, overlayed by the penetration altitude (magenta) for each fraction of the photon ensemble leaving the trajectory. The penetration altitude is defined as the lowest altitude of each individual photon trajectory.

6.6 Photon path length distributions

The photon path length distribution (PPLD) defines the distribution of all simulated photon path lengths below a reference altitude (here: 10 km). The following description gives a very simplified view on the photon trajectories:

In order to derive the path lengths of all detected photons, the trajectories of all simulated photon ensembles are considered. Because of the backward simulation routine, the trajectory of each photon ensemble starts at the satellite detector. The position of the detector marks the first event E_0 and allows to separate the individual ensembles. The following events $E_{i>0}$ describe the movement of the photon ensemble through the atmosphere. A schematic representation of a photon path below the altitude a_{ref} is shown in figure 20. At each scattering event E_i at position $\vec{x}_i$, the photon ensemble is directed into the Sun with a given probability ($\rightarrow$ contribution u_i). The path length for this part of the photon ensemble with contribution u_i then is the complete trajectory below the reference altitude a_{ref}. It is possible that the photon ensemble re-enters the atmosphere below a_{ref} which must also be taken into account. For an event E_n and thus the contribution u_n, it includes the added distances $\sum_0^n d_i(a_i < a_{ref})$ to the event E_n plus the trajectory s_n from the current position up to a_{ref} into the sun direction. The additional trajectory s_n is treated as a straight line with no additional scattering or refraction events. Following all scattering events of the trajectory, the different lengths of the photon trajectories are calculated and the resulting distribution is sorted into a pre-defined path length grid. After

the consideration of each simulated photon ensemble, the distribution is normalized to the entire detected radiance.

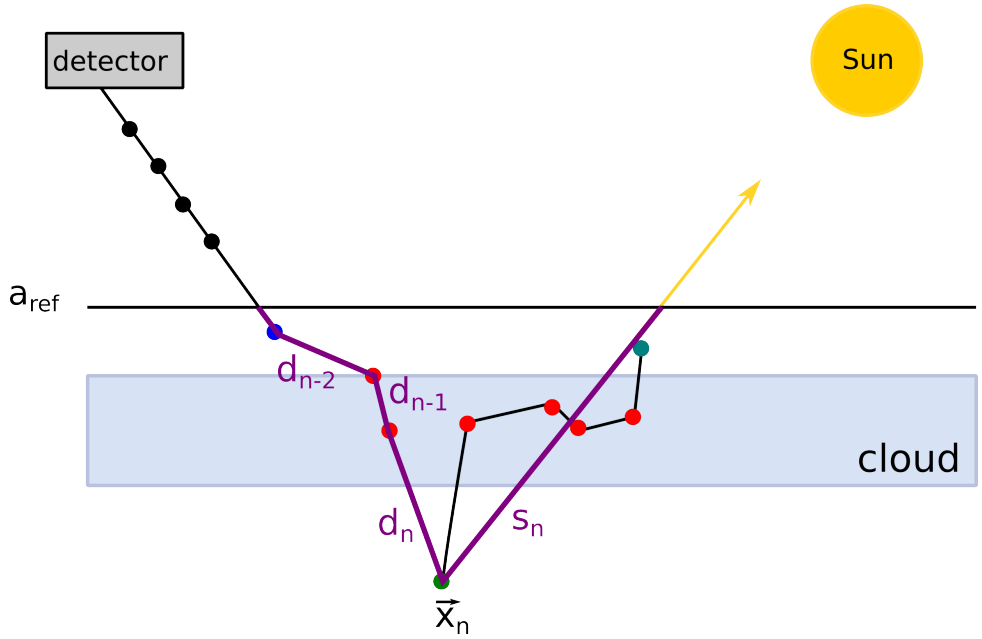

Figure 20: Schematic representation of the photon path length for the event at $\vec{x}_n$, considering the trajectory below the altitude a_{ref}. For this example the resulting path length is given by the highlighted path distances (magenta).

To assemble the geometric photon paths of all photons, different quantities from the trajectory information are used, namely the coordinates $\vec{x}_i$ and the altitude layer a_i of each event i, the geometric path $d_i = |\vec{x}_{i-1} - \vec{x}_i|$ and the contribution u_i.

The individual photon path lengths p_i are calculated as follows and saved together with the contribution u_i. If the altitude a_i of event E_i lies below the reference altitude a_{ref}, the geometric distance from $\vec{x}_i$ to the reference altitude into the solar direction s_i is added.

$$\text{if } \; a_i < a_{ref} : \quad p_i = s_i + \sum_{j=0}^{i} d_j (a_j < a_{ref}) \tag{44}$$

$$\text{if } \; a_i \geq a_{ref} : \quad p_i = \sum_{j=0}^{i} d_j (a_j < a_{ref}) \tag{45}$$

The path length s_i into the solar direction is calculated by using the direction angles from the sun position to the considered event E_i:

$$\vec{s}_i = s_i \cdot \vec{e}_{sun,i} = s_i \cdot (-1, \theta_{sun}, \phi_{sun})_i = s_i \cdot (e_{x,sun}, e_{y,sun}, e_{z,sun})_i \tag{46}$$

The value of s_i can be derived from the geometry of figure 21 (r_e is the radius of the Earth):

$$|\vec{x}_i + \vec{s}_i| = r_e + a_{ref} \tag{47}$$

$$s_i = -\vec{x}_i \cdot \vec{e}_{sun,i} + \sqrt{(\vec{x}_i \cdot \vec{e}_{sun,i})^2 - (r_e + a_i)^2 + (r_e + a_{ref})^2} \quad . \tag{48}$$

All individual photon path lengths p_i of an ensemble are classified into a grid together with their corresponding contributions u_i. After consideration of all simulated photon ensembles, the

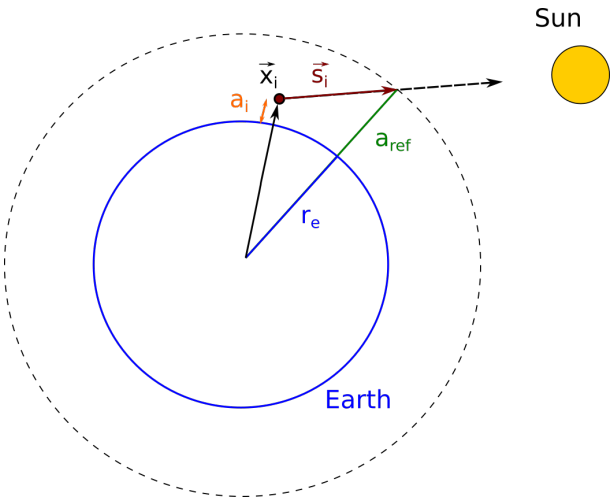

Figure 21: Geometric calculation of the path s_i from the scattering event to the reference altitude a_{ref} in direction of the Sun.

distribution is normalized to the detected radiance. From these photon path length statistics, quantities such as the mean can be calculated and the distributions visualized.

In addition to the photon path length distribution below a reference altitude, altitude resolved distributions are considered. This is illustrated in figure 22 which shows a model atmosphere and a schematic trajectory of a photon ensemble. The considered altitude layer is filled in magenta. As can be seen from this overview, all trajectories ending after the ground scattering event (green) have passed the altitude layer. The photon paths within this layer are calculated from the event information, similar to the photon path length distribution below a reference altitude a_{ref}. In contrast to the entire photon path lengths below a_{ref}, each altitude layer is considered individually. For each event E_i, all paths $p_{i,alt}$ within the chosen altitude layer are added and - if necessary - completed by paths from the scattering event into the direction of the Sun which fall into the considered altitude bin. It must be taken into account that the altitude layer may be crossed multiple times by the same photon ensemble.

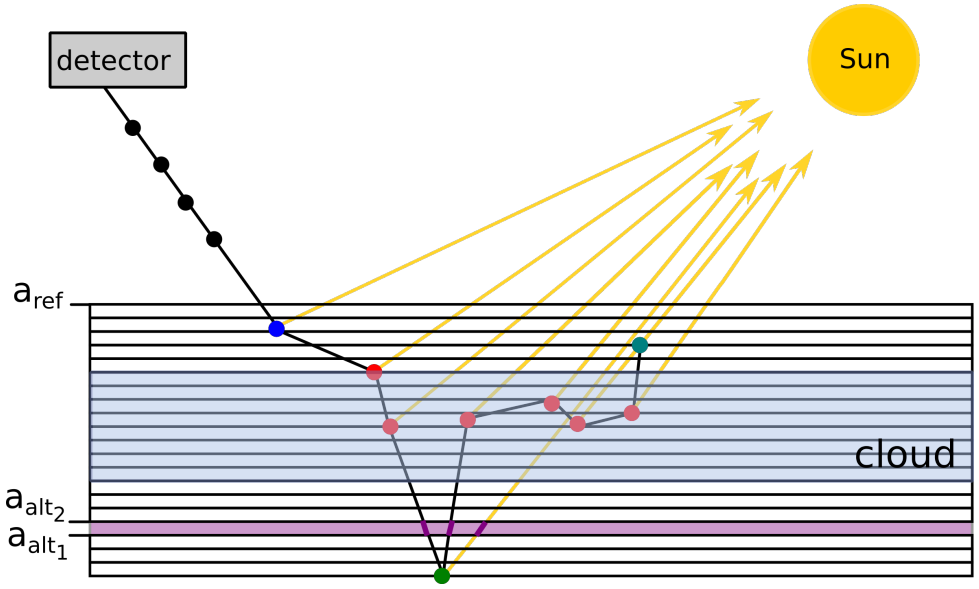

Figure 22: The altitude resolved photon path length distributions are derived from the photon trajectories within the individual altitude bins. An example altitude bin is shown between $a_{alt,1}$ and $a_{alt,2}$ and the considered photon path lengths are highlighted (magenta).

7 Dataset and case studies

During the retrieval process, the TANSO-FTS measurement spectrum is fitted to the results from radiative transfer simulations. From this fit, best estimated cloud parameters and photon path statistics for the considered cloud scenarios are obtained. In order to verify if the resulting cloud values are realistic, case studies with well characterized cloud situations are performed. These are selected by considering measurements of the TANSO instruments, as well as lidar and radar data from instruments onboard the CALIPSO and CloudSat satellites.

The section starts with a short description of the CALIOP lidar instrument on CALIPSO and the CPR radar instrument on CloudSat. From the lidar observations, altitude information of cloud and aerosol layers is obtained. Information on the altitude resolved cloud extinction and the cloud classification is provided by the radar measurements. Further information on the cloud scenario is obtained from TANSO-CAI, the second GOSAT instrument which performs measurements of pixel radiances that are collocated with the TANSO-FTS measurements. Here, the cloud property dataset gives information on the total cloud optical depth and the estimated particle radii (see section 5.5 for details).

In order to combine the cloud characterizations of these satellite instruments, collocated measurements have to be identified. The CloudSat footprint is very similar to the CALIPSO footprint and the measurements show only a small time difference. This reduces the task to obtain temporal and spatial collocations of CALIPSO and GOSAT. This procedure as well as the resulting dataset is explained in the second part of this section.

The section ends with a detailed description of the chosen case studies and the corresponding characterization of the cloud scenarios.

7.1 CALIPSO

The CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) satellite was launched in 2006 to be included in the NASA A-train and is part of the NASA ESSP (Earth System Science Pathfinder) mission in collaboration with CNES. The main instrument onboard the satellite is the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). The satellite flies in a sun-synchronous orbit at about 705 km altitude where it completes one measurement cycle in 16 days. The equator crossing time is approximately 13:30 (local time) [105].

Lidar instruments are active sensors which can be used to obtain atmospheric profile information, e.g. on clouds and aerosols. CALIOP sends laser pulses at 532 nm and 1064 nm of about 20 ns duration down to the Earth which result in footprints with a diameter of about 70 m at the surface. The distance between the subsequent footprints is 335 m. The elastic backscatter signals originate from receiver footprints of about 90 m diameter and are measured by a three-channel receiver in two orthogonal polarizations at 532 nm and as a single backscatter signal at 1064 nm. By analyzing the backscatter data, so-called features in the atmosphere can be detected. A feature consists of a particulate ensemble (e.g. clouds or aerosols) that is extended over a certain region and shows stronger backscatter signals than the molecular background [102; 105].

Lidars, such as CALIOP, are the only kind of satellite instrument to produce aerosol profiles of high resolution. Measurements of aerosol over bright surfaces and clouds can also be performed. Because of the polarization sensitivity of CALIOP, the vertically resolved cloud phase can be retrieved from the depolarization of the backscattered signal [105].

For the identification of cloud and aerosol layers, the attenuated backscatter profiles at 532 nm are used together with the selective iterative boundary locator (SYBIL) algorithms (see Vaughan

et al. [102] for further details). The type of the features is detected by the scene classification algorithms (SCA) which use all three receiver channels to identify clouds, aerosols, the type of aerosol and the cloud phase. If more than one layer is identified within a measurement column, the retrieval is first applied to the upper layer. Afterwards, the backscatter data is corrected for the attenuation by the upper layers. Profiles of the backscatter data and extinction coefficients are available from the hybrid extinction retrieval algorithm (HERA). The extinction products are not to be used for water clouds because multiple scattering effects within the large lidar footprints lead to significant errors, which are not accounted for in the retrieval [105].

The main data product from CALIOP measurements used here is the feature location in the atmosphere. The heights can be retrieved for clouds of optical depth larger than 0.01 and for aerosols with an optical depth larger than 0.005. The vertical extent can be determined for optically thin cloud layers as long as the cloud optical depth is below 5 (3 for aerosols). Backscatter data below 8.2 km is available with a resolution of 30 m in the vertical and 333 m in the horizontal direction. For higher altitudes, the signal data is averaged in both spatial directions onboard the satellite in order to reduce the telemetric bandwidth for the data transmission [105].

For the characterization of the cloud scenarios within this work, the Level 2 Cloud Layer and Aerosol Layer products are chosen. The top and base altitudes are retrieved for each layer but it has to be accounted for that the confidence is higher for the top height of a layer because of the signal attenuation within the layer and multiple scattering effects. This is also true for multi-layered features where the upper layers are easier to detect than lower features. The vertical resolution is 30 m below 8.2 km, 60 m below 20.2 km and 180 m for higher altitudes [12]. In the horizontal direction, data products with different resolutions are available. Because strongly scattering features are easier to identify, one laser shot is usually sufficient for the detection of clouds in the boundary layer [12; 105]. As space-based lidar has a low signal-to-noise ratio, horizontal averaging is needed to identify weakly scattering elements like thin cirrus clouds or aerosols [105]. Within this thesis, the dataset without horizontal averaging (resolution of 333 m) is used for the identification of cloud boundaries below 8 km. In order to detect the presence of cirrus clouds, the 5 km horizontal averaging is chosen which considers higher altitudes as well. Aerosol layer information is only available from 5 km horizontal averaging.

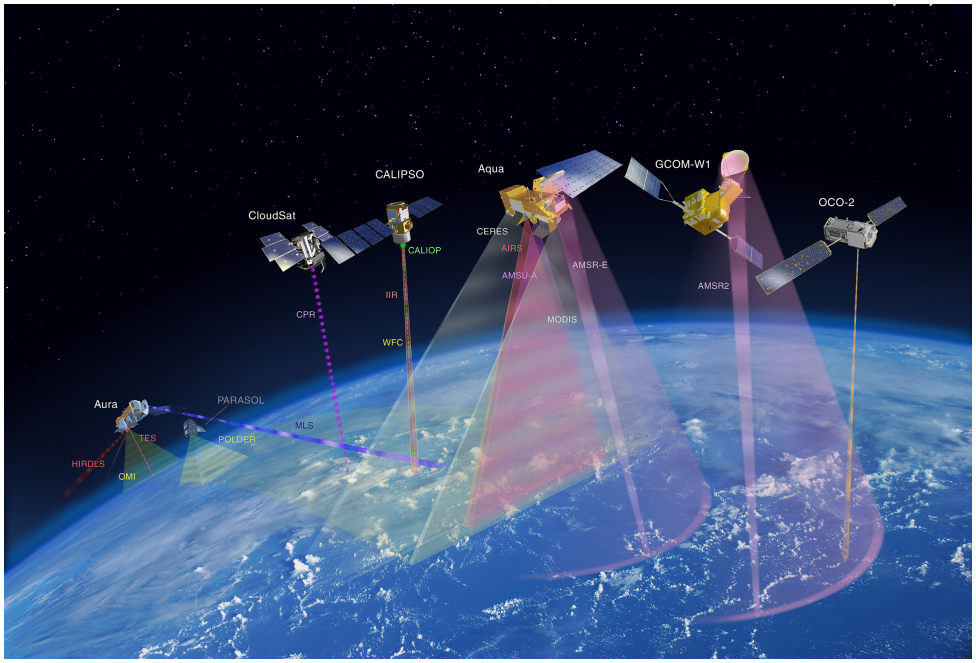

Figure 23: Impression of the NASA A-train (afternoon) formation in 2013 (from https://atrain.nasa.gov/images/A-Train2013_Web.jpeg, accessed on 11 Oct., 2017).

7.2 CloudSat

CloudSat is also part of the ESSP missions and was launched together with CALIPSO on April 28, 2006, carrying the first spaceborne 94-Ghz (3.2 mm) radar instrument, the Cloud Profiling Radar (CPR) [86; 93]. CloudSat flies in close formation with CALIPSO. The short time difference of about 15 seconds enables the synergetic use of the CPR and CALIOP sensors. Orbit manoevers are performed to restrict the distance of the footprint within 2 km along track to each other, which should be achieved in at least 50% of the time. According to Stephens et al. [93], the requirement was kept in more than 90% of the analyzed time.

CPR performs near-nadir measurements (0.16° since August 2006) by emitting pulses of 3.3 μs and measuring the backscattered power as a function of time. The signal is backscattered at atmospheric constituents, e.g. cloud droplets and ice crystals, the minimum measured detection sensitivity is -30 dBZ. The received power is then converted to distances from the detector. The resolution of the measurements is 1.3 - 1.4 km in cross track and 1.7 - 1.8 km in along track direction. Each recorded profile consists of 125 vertical samples of appr. 240 m. The data is oversampled as the range resolution is about 500 m [93; 95]. The backscatter profiles represent vertical slices of the atmosphere in along-track direction and lead to vertically resolved structures of underlying cloud scenarios [91]. Here, CPR has the ability to detect most tropospheric clouds while those with small particles (e.g. cold cirrus clouds) usually are invisible to the instrument because of the scattering power law dependence on the particle size. [86]. One of radar's extraordinarily strengths is the simultaneous detection of cloud droplets and precipitation [93].

From the backscatter profiles, a cloud mask is created by the CloudSat hydrometeor detection algorithm. The algorithm checks whether the backscattered signal results from scattering on hydrometeors or from noise. Details on the algorithm are found in Marchand et al. [53]. This results in a geometrical profile which is named 2B-GEOPROF and also includes the radar reflectivity, quality flags and an estimation of gas absorption due to O_2 and H_2O in the atmosphere. The range bin which is closest to the mean sea level is set to the same bin number in all profiles. The bins directly above the surface can be influenced by surface clutter which results from surface reflections 2-5 orders higher than the signal from hydrometeors. The contribution is most important in the first two bins above the surface, the original sensitivity is only regained in the 5^{th} bin above the surface which corresponds to approximately 1.2 km [53].

The main data products of interest for our study are the cloud classification 2B-CLDCLASS and the cloud optical depth 2B-TAU. The cloud classification algorithm (see Sassen and Wang [86]), assigns one of eight cloud types to each cloud cluster identified by the 2B-GEOPROF product. The algorithm also detects precipitation and the possibility of a mixed phases within the cloud. The 2B-TAU data product includes the retrieved (short-wave) cloud optical depth as well as the effective cloud particle radius for each radar bin. The retrieval is based on the retrieved cloud profile of the 2B-GEOPROF product together with MODIS (Aqua) reflectivities from different measurement channels. Further input parameters are the geolocation of the measurement, meteorological parameters and the surface albedo. For a detailed description, please refer to reference [60]. The total cloud optical depth for the atmospheric column is also provided and can be compared to the values measured by TANSO-CAI.

7.3 Collocated measurements and cloud characteristics from CALIOP

In order to obtain collocated measurements of TANSO-FTS and CALIOP, the instrument ground tracks have to be compared in time and space. As starting point, the Estimated Geolocation Data (see section 5.3.1) is used to identify the date, time and coordinates of all TANSO-FTS measurements between April 2009 and October 2015. The coordinates represent the center of the field of view for each measurement. The measurement time and location of the CALIOP instrument is approximated by the sub satellite points (SSP) of the CALIPSO satellite which is

the projection of the trajectory on the Earth's surface. The CALIOP lidar measures near-nadir, consequently the SSP should be sufficient for a preselection of the measurements.

The SSP are calculated in time steps of one minute from the NORAD (North American Aerospace Defense Command) Two Line Element (TLE) data, describing the orbital movement through space. For CALIPSO, the orbital parameters are typically recorded twice a day. TLE data contains information such as the Epoch date and time, inclination and eccentricity of the satellite. The used data was obtained from CelesTrak [41] which provides also detailed information on the content of the datasets. To calculate the satellite SSP at a given time, the freely available software package *pyorbital*, developed by Martin Raspaud, is used (for details, see [23]).

The main criteria for the preselection of collocations are a small time difference between the measurements of the different sensors and the necessity that the CALIOP measurement is located within the TANSO-FTS field of view. The TANSO-FTS field of view is spherical with a projected diameter of about 10.5 km. CALIOP performs measurements with a small footprint of less than 100 m diameter. Measurement pairs are identified as potentially collocated if the maximum distance between the center of the TANSO-FTS FOV and the CALIOP SSP is smaller than 5 km (see fig. 24) and the time difference between the overpasses does not exceed 20 minutes. The time frame is chosen as a compromise to ensure a high number of collocations while minimising the temporal change of the cloud properties at the same time.

All potential collocations are saved for future reference, including the geolocations, the time stamps and the solar zenith angle (provided by *pyorbital*) which should not exceed 50° in the final selection.

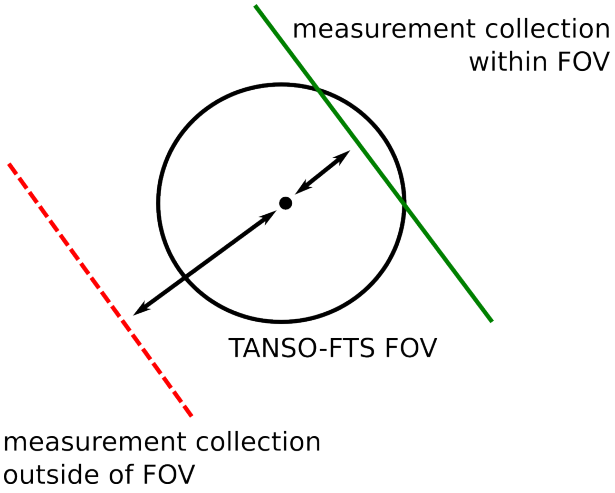

Figure 24: Spatial collocation criterium: If sub satellite points of CALIPSO coincide (green line) with the TANSO-FTS FOV (black circle), the measurement pair is identified as potential collocation. It is not, if the sub satellite points of CALIPSO (red dotted line) lie outside the TANSO-FTS FOV.

The resulting data is checked for collocations using the real measurement information (time, longitude, latitude) from CALIOP. The necessary measurement files are downloaded from the ICARE FTP Server of the University of Lille[8] by an automatic routine. If available for the given time and location, the geolocation as well as the measurement time are compared to the corresponding TANSO-FTS measurement. If the collocation conditions are met, the number of cloud layers is retrieved from the CALIOP L2 Cloud Layer product.

The number of cloud layers below 8 km is determined for all measurements within the TANSO-

[8]ICARE Data and Services Center: http://www.icare.univ-lille1.fr

FTS FOV and the average top and base altitudes are retrieved. The standard deviation of the cloud top height provides an estimate of the flatness of the cloud top. Here, cloud cases with a high degree of spatial homogeneity are searched, thus only cloud scenarios with a standard deviation of the cloud top heights below 100 m are considered in the final dataset. Nearby cloud layers are included in the visualization for the case study description to give an impression of the homogeneity of the cloud spatial structure and possible temporal changes. For altitudes above 8 km, only the existence of cloud layers is checked. In this altitude region, thin cirrus clouds may occur which can significantly change the atmospheric radiative transfer. The result is saved along with the information of cloud layers in the lower atmosphere.

In addition to the lidar information on cloud layers, the existence of aerosols in the TANSO-FTS FOV region and the closer environment is considered by using the CALIOP L2 Aerosol Layer product. As for high-altitude clouds, the information is saved for future reference and does not influence the final selection of collocated measurements.

7.4 Cloud characteristics from TANSO-CAI and CPR

TANSO-CAI measures pixel radiances of the TANSO-FTS observation location in four different wavelength bands at approximately the same time as TANSO-FTS (see section 5.5 for details). One of the derived datasets is the TANSO-CAI cloud property which includes pixel-resolved cloud optical depths and particle radii. By attributing the pixel information to the geolocation of the TANSO-FTS FOV, one can retrieve the geometrical cloud fraction and an average cloud optical depth for the corresponding TANSO-FTS measurement.

This procedure is applied to all collocated measurements between TANSO-FTS and CALIOP. In addition to the cloud cover and the cloud optical depth, different quality flags concerning the TANSO-FTS measurement are checked. In case that all given criteria are passed, the status of the cloud cover, the mean and standard deviation of the cloud optical depth as well as the geolocation information (surface type of the FOV region, presence of sunglint and rough terrain) are added to the dataset. Furthermore, it is verified if saturation in any of the four CAI bands occurs and if cirrus clouds or aerosols are detected within the TANSO-FTS field of view.

The vertical structure of the cloud can also be accessed using radar measurements from the Cloud Profiling Radar (CPR) on CloudSat, which precedes CALIPSO within the A-train formation by about 15 seconds. The satellite movements are monitored to cover the same ground track. Nevertheless, small distances between the measurements occur, consequently collocation criteria are also applied for CPR measurements. The data products are again acquired by an automatized download from the ICARE FTP server. The allowed time window of the measurement is $\pm$ 25 minutes. Because the footprints of CALIOP and CloudSat are not perfectly aligned, a slightly larger timeframe was assumed than for CALIOP measurements. If a measurement of the cloudy atmosphere is available within the time frame but not within the TANSO-FTS FOV, the nearest measurement is taken into account. The altitude dependent extinction and total cloud optical depth is retrieved from the CPR 2B-TAU data product. The cloud classification product 2B-CLDCLASS allows an identification of the cloud scenario including the occurrence of precipitation and the uncertainty of the measurement. The retrieved data is added to the cloud characterization data if the collocated CPR measurement is available.

7.5 Dataset

For each scenario of the resulting dataset, the TANSO-FTS L1B V201.202 products are downloaded, including the measured radiance spectra and the measurement geometry. All of the retrieved collocated data points are shown in figure 25.
The collocations are only found in the Northern Hemisphere between 25° N and 60° N. This is caused by the fact that daytime measurements are performed in the descending node by GOSAT

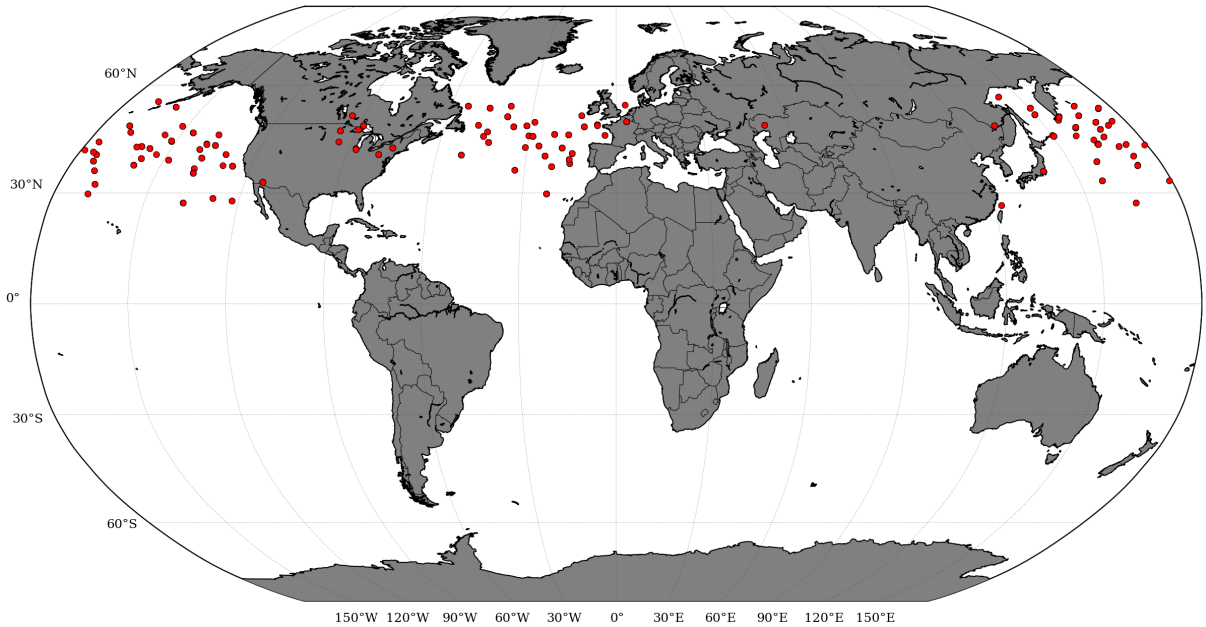

Figure 25: The TANSO-FTS measurement locations for the derived collocations (red), mapped onto the Earth's globe.

and the ascending node by CALIPSO. Most collocated measurements within cloudy atmospheres are found over ocean. The reason has not been further investigated but it is possible that the stringent condition on the spatial homogeneity of the cloud top height leads to a preferred selection of cloud scenarios over water.

The collocated dataset is sorted according to the TANSO-FTS based solar zenith angle, the cloud optical depth τ_{cloud} from the TANSO-CAI cloud property and the CALIOP-derived average cloud top height. Cloud top heights h_{cloud} are classified into $h_{cloud} \leq 0.8$ km, $0.8 < h_{cloud} \leq 1.5$ km (cloud ≈ 1 km), $1.5 < h_{cloud} \leq 2.5$ km (cloud ≈ 2 km) and $h_{cloud} > 2.5$ km (higher clouds). If more than one TANSO-FTS measurement observes approximately the same scene with only some seconds time difference (e.g. from special observation pointing modes), the cloud scenario is counted only once and sorted according to the mean cloud properties. The results with the number of found measurements are displayed in tables 3 to 6. For those presented cases, data from CloudSat, CALIPSO and GOSAT is available. Also, no high clouds in the TANSO-FTS FOV are detected by CALIOP, no aerosol by TANSO-CAI and none of the four TANSO-CAI bands shows saturation during the measurement. Possible detected aerosol by CALIOP does not influence the selection but is included in each dataset as additional information.

7.6 Case studies

The following section presents the case studies considered in section 8 and the result sections 10, 11 and 12. The cloud situations present during the TANSO-FTS measurements are described in detail using measurement data from multiple satellites. The individual scenarios are named after the TANSO-FTS L2 IDS scan ID and are attributed consecutive numbers for convenient identification.

- case 1: F090604215013382401
- case 2: F090828235909422301
- case 3: F150326232730412204

Table 3: Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 10° and 20°.

	$0 < \tau_{cloud} \leq 10$	$10 < \tau_{cloud} \leq 15$	$15 < \tau_{cloud} \leq 20$	$20 < \tau_{cloud} \leq 30$	$30 < \tau_{cloud}$
cloud $\leq$ 0.8 km	-	-	-	-	-
cloud $\approx$ 1 km	-	1	1	1	-
cloud $\approx$ 2 km	-	-	-	-	-
higher cloud	-	-	-	-	-

Table 4: Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 20° and 30°.

	$0 < \tau_{cloud} \leq 10$	$10 < \tau_{cloud} \leq 15$	$15 < \tau_{cloud} \leq 20$	$20 < \tau_{cloud} \leq 30$	$30 < \tau_{cloud}$
cloud $\leq$ 0.8 km	-	-	1	-	3
cloud $\approx$ 1 km	4	7	1	5	-
cloud $\approx$ 2 km	2	1	-	3	-
higher cloud	-	-	1	-	3

Table 5: Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 30° and 40°.

	$0 < \tau_{cloud} \leq 10$	$10 < \tau_{cloud} \leq 15$	$15 < \tau_{cloud} \leq 20$	$20 < \tau_{cloud} \leq 30$	$30 < \tau_{cloud}$
cloud $\leq$ 0.8 km	4	2	-	-	-
cloud $\approx$ 1 km	3	4	5	2	-
cloud $\approx$ 2 km	1	1	4	1	-
higher cloud	-	3	4	4	2

Table 6: Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 40° and 50°.

	$0 < \tau_{cloud} \leq 10$	$10 < \tau_{cloud} \leq 15$	$15 < \tau_{cloud} \leq 20$	$20 < \tau_{cloud} \leq 30$	$30 < \tau_{cloud}$
cloud $\leq$ 0.8 km	1	-	1	-	-
cloud $\approx$ 1 km	1	1	4	4	-
cloud $\approx$ 2 km	-	1	1	-	-
higher cloud	-	1	1	-	2

While cases 2 and 3 are exclusively used as case studies for the retrieval of cloud properties, case 1 is also used to test the performance of the implemented spectrum fit. Case 1 and 2 represent rather simple cloud scenarios featuring a cloud with top height of about 1 km or 2 km, respectively. Cases 1 and 2 have been chosen randomly under the condition that one cloud scenario of the collocated measurements with approx. 1 km as well as approx. 2 km top height is included. Case 3 was selected, because it shows a more complicated scenario, consisting of a highly structured cloud that spans a large altitude range.

As described before, the used cloud information originates from lidar measurements of the CALIOP instrument (cloud and aerosol layers (references in Appendix, section 7.7.1)), radar measurements from the CPR instrument (cloud optical depth and classification [59; 60]) as well as the TANSO-CAI cloud property [71].

7.6.1 Case 1

The considered dataset is one of the early GOSAT measurements, taken on 4 June, 2009 in the middle of the pacific ocean (longitude = -129.52°, latitude = 37.41°). The surface type

is water and the solar zenith angle at the Earth projected measurement point is 21.75°. The measurement is recorded in near-nadir direction with a satellite zenith angle of 1.31°. The FOV image is shown in fig. 26 and further information on the measurement in table 7. The FOV image, taken at the same time as the corresponding TANSO-FTS measurement, originates from the camera mounted on the same optical bench as the Fourier transform spectrometer [45].

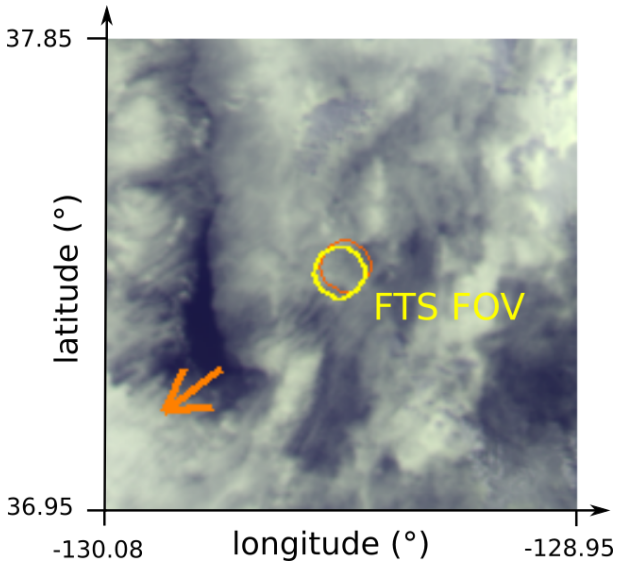

Figure 26: FOV image from the GOSAT monitor camera for case study 1 (modified from [3]). The yellow circle describes the FOV location of TANSO-FTS. The assumed FOV before the geolocation correction is shown in orange and the orange arrow denotes the direction of the Sun from the measurement location (for information on the geolocation correction, see section 5.3.1).

Table 7: Geometrical characteristics of case study 1 as well as selected surface properties at the measurement location as provided by the GOSAT FTS L1B and L2 IDS products.

longitude	-129.52°	surface of FOV	sea
latitude	37.41°	sunglint	no
timestamp of measurement	2009-06-04 21:50:13	rough terrain flag	no
solar zenith angle	21.75°	satellite zenith angle	1.31°
solar azimuth angle	232.02°	satellite azimuth angle	20.47°

The TANSO-FTS radiances are shown in fig. 27. The TANSO-FTS radiances are measured in the P- and S-polarization (see section 5) and are very similar, indicating that the radiance is largely non-polarized. Here, the approximated total radiance is calculated following O'Brien et al. [70]:

$$I = \frac{(I_P + I_S)}{2} \tag{49}$$

The main source of information for the cloud optical depth of the measurement situation is the TANSO-CAI cloud property product (the filenames of all used cloud products are given in section 7.7, table 10). The cloud optical depth together with the footprints of the TANSO-FTS, CALIOP and CPR is shown in fig. 28.

The average cloud optical depth within the TANSO-FTS FOV, as derived from TANSO-CAI

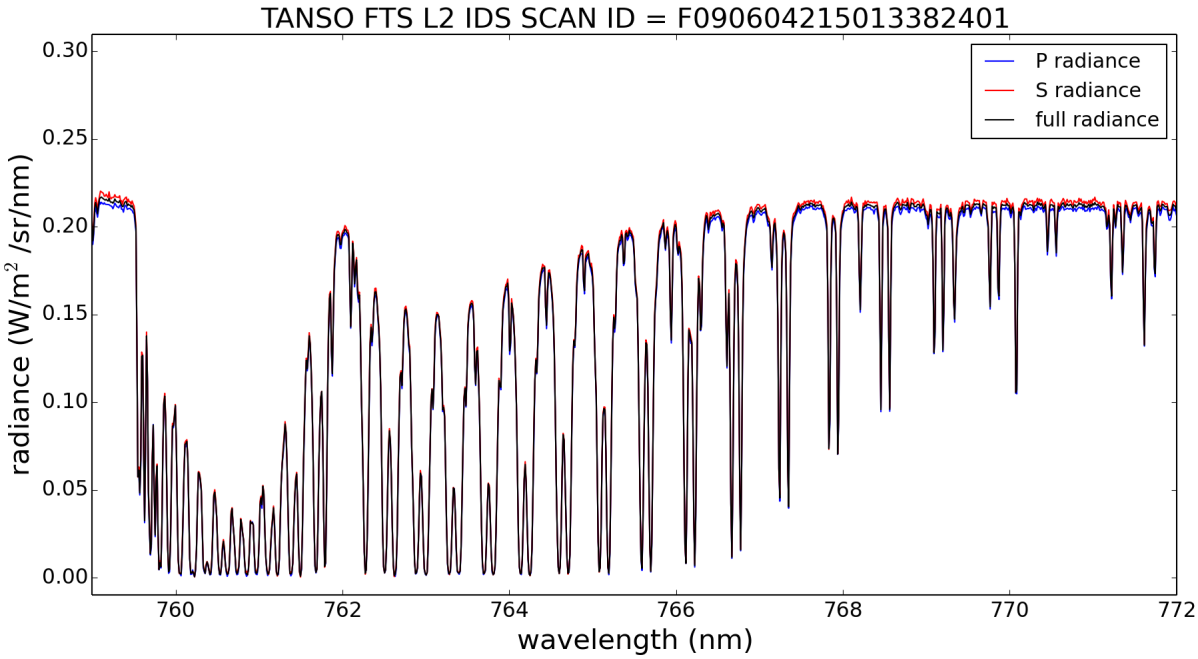

Figure 27: The TANSO-FTS radiances of case study 1 between 759 and 772 nm. The polarized radiances (P in blue and S in red) are shown together with the approximation of the full radiance spectrum (black).

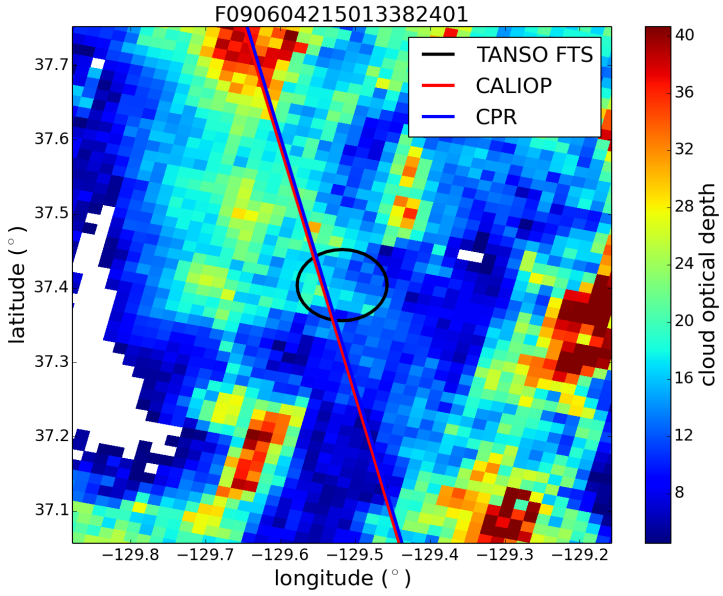

Figure 28: Cloud optical depth from TANSO-CAI together with the TANSO-FTS FOV (black circle) and the CALIOP (red) and CPR (blue) measurement ground tracks for case study 1. White pixels identify invalid data.

measurements, is 14.65 with a standard deviation (SD) of 2.95. The TANSO-CAI cloud property algorithm detects no significant aerosol contamination[9] of the measurement.

A graphical representation of the cloud layers from CALIOP measurements is shown in fig. 29. The used CALIOP data has a horizontal resolution of 333 m and a vertical resolution of 30 m. Atmospheric structures until an altitude of 8 km are considered. Higher altitudes are taken into account in the strongly horizontally averaged (5 km) data product which can be used for cirrus

[9]The term used in the GOSAT data product is *heavy aerosol load*.

cloud screening, as done by e.g. Min et al. [58]. The lidar measures a horizontally extended cloud structure with a rather homogeneous cloud top height at about 1.03 km (average value within the TANSO-FTS FOV with a standard deviation of 0.02 km for 24 measurement points). The approximate time difference between the measurements is 11-12 minutes. Aerosol features are not found for the collocated measurements but in the closer environment (not shown here). According to the TANSO-CAI cloud flag, the probability of high cirrus clouds[10] is given for about one third of the TANSO-FTS FOV, also at the coordinates of the CALIOP measurement. However, the stronger information source is the lidar data from CALIOP, which finds no high clouds in the collocated region with the TANSO-FTS FOV.

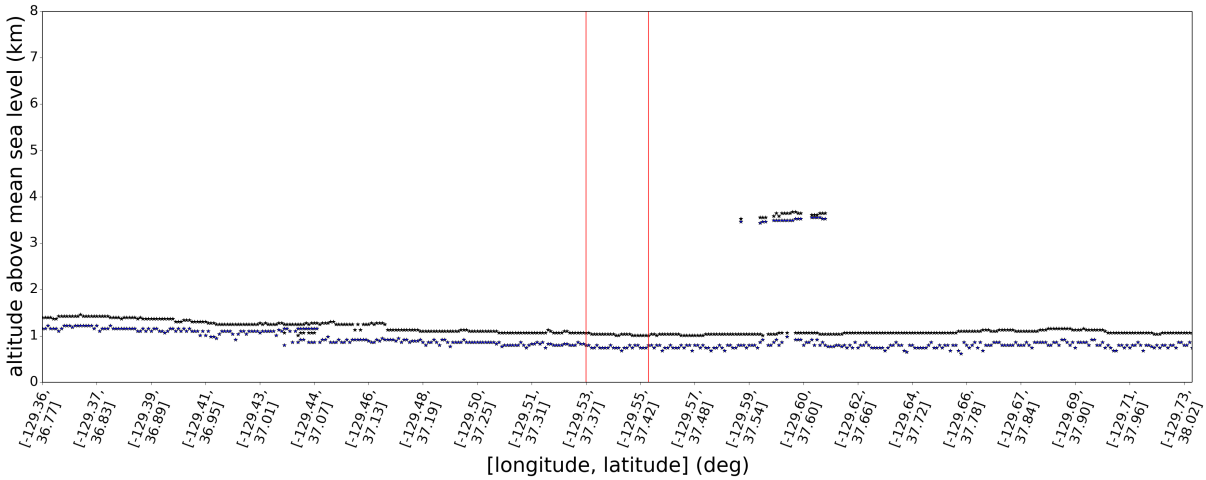

Figure 29: Cloud top heights (black) and cloud base heights (blue) for case study 1 retrieved from lidar measurements of CALIOP. All measurements that fall within the coordinates of the TANSO-FTS FOV are located between the red boundaries.

The vertical extent and profile of the cloud can best be determined from radar measurements. The cloud optical depths for each radar bin until an altitude of 8 km are shown in fig. 30. The strength of the extinction (layer optical depth) is expressed by the colorbar values, each vertical bin is 240 m. Invalid radar measurements are shown in white. The radar bins within the TANSO-FTS FOV are identified by the red box. Because the locations of the radar bins have slightly different altitude levels, elevation lines are drawn in white for each measurement. The horizontal white lines correspond to the altitudes above the mean sea level. Enhanced extinctions are found between 1 and 3 km with the highest values at about 1 km. It is possible that the data is influenced by the previously described surface clutter. Although not visualized, no higher cloud structures are detected.

The cloud classification product derived from CPR measurements presents a cloud in the 0.86-1.20 vertical bin and the 1.10-1.34 km vertical bin, for some measurements only in the lower bin. The cloud is classified as midlatitude stratocumulus. The time difference between the CPR and TANSO-FTS measurements is similar to the lidar measurement but the measurement swath is slightly nearer to the center of the TANSO-FTS FOV. The CPR derived total optical depth has an average value of 15.40 $\pm$ 1.24 (SD). The uncertainty of this total cloud optical depth, as given in the 2B-TAU product, is 0.46 in average. This value is in agreement with the total cloud optical depth from TANSO-CAI measurements.

[10]The *probability of high cirrus clouds* is a term used in the GOSAT data products. According to the definition, it is given if I(0.87 μm)/I(1.6 μm) $\leq$ 0.8 [65].

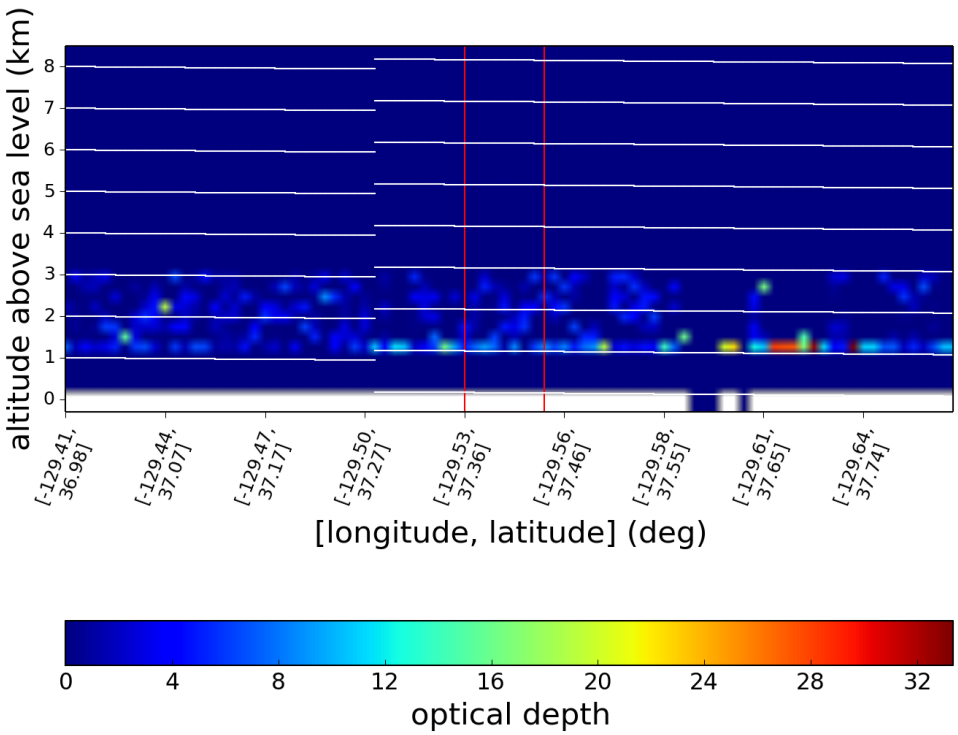

Figure 30: Altitude resolved cloud optical depth (of the 240 m bins) from CPR measurements and location of the collocated coordinates of the TANSO-FTS field of view (red box) for case study 1. The altitude levels of the horizontal bins are drawn in white. Cloud structures at 1-3 km are found.

7.6.2 Case 2

The TANSO-FTS measurement of case 2 is taken over a stratocumulus structure in the Northern Pacific Ocean on 28 August, 2009. Figure 31 shows the detector's field of view (yellow circle) of TANSO-FTS. The TANSO-FTS measurement is performed near nadir (satellite zenith angle of 1.36°) and the solar zenith angle has a value of 37.28°. The geometry of the measurement situation is summarized in table 8.

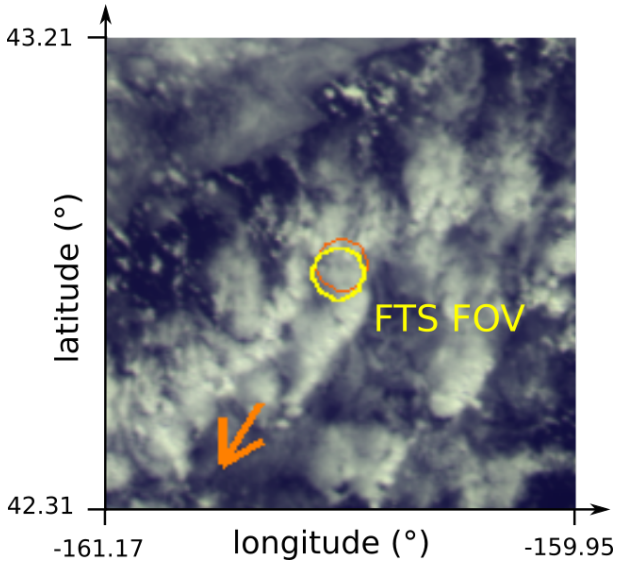

Figure 31: Image of the measurement situation of case study 2 (modified from [4]) including the geolocation-corrected TANSO-FTS field of view (yellow circle). The orange circle shows the non-corrected FOV, and the orange arrow denotes the Sun direction from the measurement location.

Table 8: Geometrical characteristics of case study 2 as well as selected surface properties at the measurement location as provided by the GOSAT FTS L1B and L2 IDS products.

longitude	-160.56°	surface of FOV	sea
latitude	42.76°	sunglint	no
timestamp of measurement	2009-08-28 23:59:09	rough terrain flag	no
solar zenith angle	37.28°	satellite zenith angle	1.36°
solar azimuth angle	211.97°	satellite azimuth angle	17.91°

The TANSO-FTS measurements are shown in fig. 32. It is interesting to note that the measured radiances of the two polarizations show large differences.

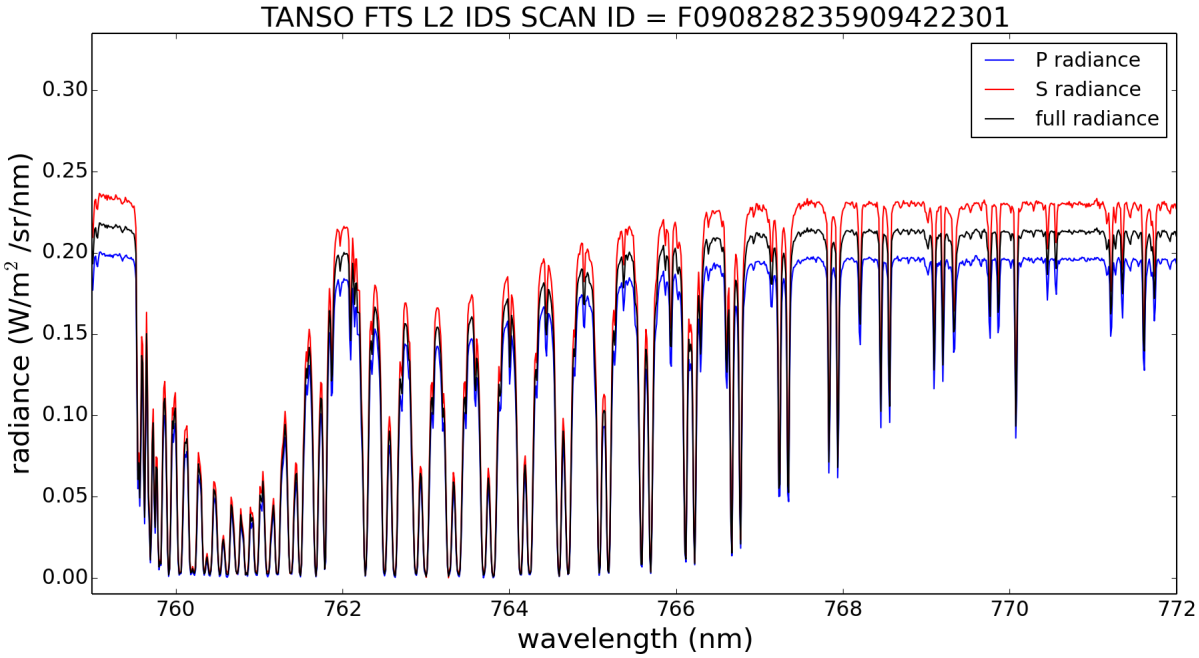

Figure 32: The TANSO-FTS radiances of case study 2 between 759 and 772 nm. The polarized radiances (P in blue and S in red) are shown together with the approximation of the full radiance spectrum (black).

Fig. 33 visualizes the collocation situation together with the TANSO-CAI cloud optical depth. The TANSO-FTS FOV lies within the black circle which is crossed by both the CALIOP measurement swath (red line) and the CPR measurement swath (blue line). Considering the structure of the cloud optical depth, the CALIOP and CPR results are focused on the brighter cloud structures within the TANSO-FTS measurement field. The datafiles used for the cloud properties are summarized in section 7.7, table 11.

The TANSO-CAI cloud property identifies 100% cloudy pixels within the TANSO-FTS FOV with a mean cloud optical depth of 19.72 ± 6.63 (SD). The data quality of the 39 pixels is judged positive and neither high aerosol loads nor saturation within any of the four CAI bands are detected. A probability of thin cirrus clouds is given for 87.18 percent of the pixels.

The lidar measurements from the CALIOP instrument take place near-simultaneously with the TANSO-FTS measurement having a mean time difference of about 16 seconds. The mean distance to the center of the TANSO-FTS FOV is 3.94 km. Figure 34 shows the measured cloud top and base altitudes, the collocated part of the TANSO-FTS FOV is located inside the red boundaries. CALIOP observes a single cloud layer in the collocated region with an average top altitude of 2.04 km. The measurement suggests a spatially homogeneous cloud top as the

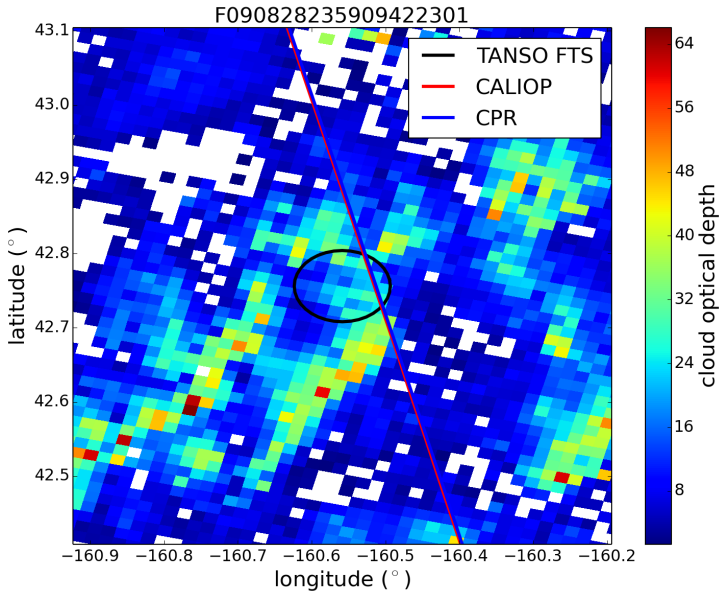

Figure 33: Cloud optical depth from TANSO-CAI together with the TANSO-FTS FOV (black circle) and the CALIOP (red) and CPR (blue) measurement ground tracks for case study 2. White pixels identify invalid data.

standard deviation is only 0.02 km for 23 measurement points. The cloud top altitude seems very stable, also in the larger environment of the considered measurement region. The cloud base height is estimated to an average height of 1.75 km but one has to keep in mind that these lidar measurements are more sensitive to cloud top heights and the sensitivity to lower cloud layers decreases depending on the attenuation and multiple scattering of the backscatter signal in the upper layers [12]. For higher optical depths like in this case, it can thus be assumed that the lidar is blind to the cloud bottom. Although not considered in the visualization (fig. 34), no cloud structures above 8 km are found for the measurement locations within the TANSO-FTS FOV.

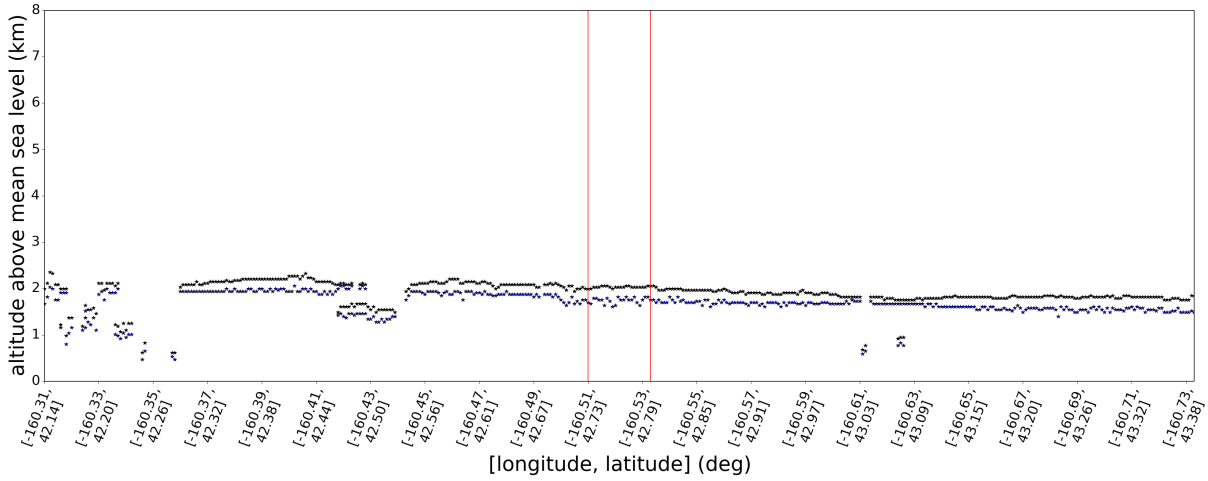

Figure 34: Cloud top heights (black) and cloud base heights (blue) for case study 2 retrieved from lidar measurements of CALIOP. All measurements that fall within the coordinates of the TANSO-FTS FOV are located between the red boundaries.

In addition to the cloud products, aerosol information is available from CALIOP measurements. In order to check the aerosol contamination of the TANSO-FTS measurement region, the aerosol

base and top altitudes are extracted from the aerosol product. Although significant amounts of aerosol can be observed in the larger environment of the measurement region, the region itself is estimated to be aerosol-free. Given the short measurement time difference between the TANSO instruments and CALIOP, these results can be applied to the collocated measurement region during the TANSO-FTS measurement time.

The CloudSat CPR measurements are even closer in time to the TANSO-FTS measurement with an average time difference of 3-4 seconds. From the Cloud Profiling Radar (CPR) datasets, the altitude dependent cloud optical depth, the total integrated cloud optical depth and the cloud classification are considered.

The total cloud optical depth within the TANSO-FTS FOV averages to 25.75 $\pm$ 5.93 (SD), the provided uncertainty is 1.03. Closer inspection of the spatial distribution of the cloud optical depths shows that this value is in agreement with the TANSO-CAI cloud optical depth. The average distance to the center of the TANSO-FTS FOV is 4.38 km. All 8 radar measurements which fall within the collocated region, are judged to have good data quality. The measured cloud optical depths for the individual radar bins are pictured in fig. 35.

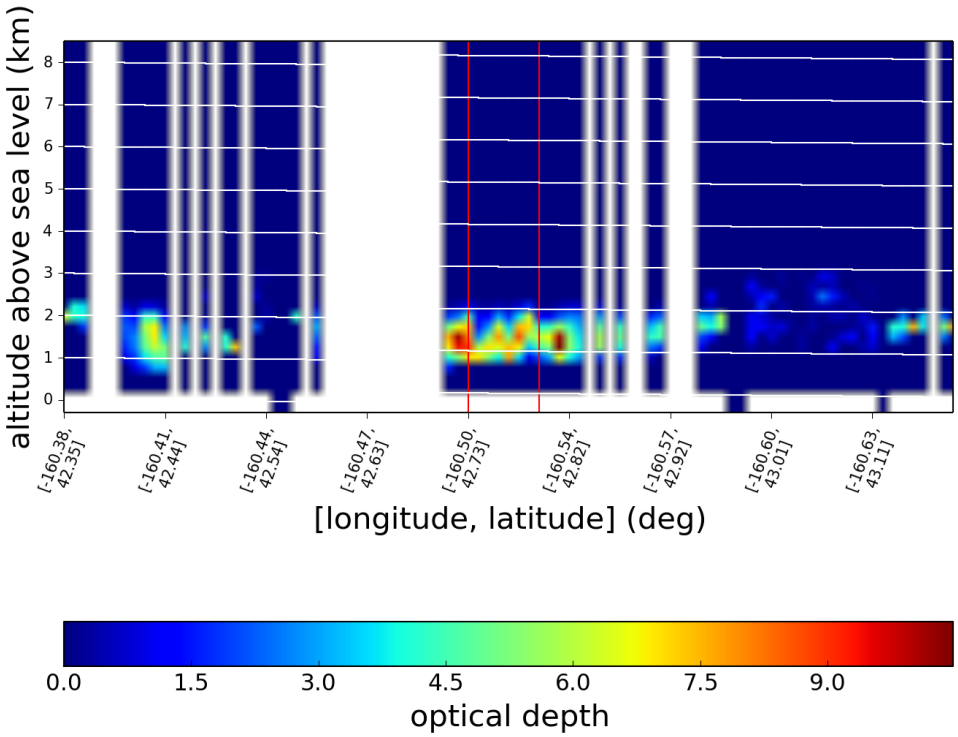

Figure 35: Altitude resolved cloud optical depth from CPR measurements and location of the collocated coordinates of the TANSO-FTS field of view (red box) for case study 2. The altitude levels of the horizontal bins are drawn in white.

The primary information source for the cloud base altitude are the CPR measurements. Dominant cloud structures are found between approximately 0.8 and 2.0 km with varying extinction. Here, the strength of radar measurements becomes obvious as the radar pulses are able to penetrate further into the cloud than the lidar pulses which leads to lower cloud base values from CPR measurements than those of CALIOP.

The last information used from CloudSat measurements is the cloud class classification. The algorithm detects a continuous cloud between the altitude bins 0.61-0.85 km and 2.3-2.54 km. The cloud within the TANSO-FTS measurement FOV is entirely classified as midlatitude stratocumulus cloud with possible drizzle and liquid precipitation in the inner cloud layers.

7.6.3 Case 3

Case 3 was chosen because it features a more inhomogeneous cloud scenario. Like both previous cases, the measurement was taken over the Pacific Ocean.

The measurement date and time is 26 March, 2015 at 23:27:29 UTC. The solar zenith angle at the location of the TANSO-FTS has a value of 41.05°, the corresponding satellite zenith angle is 22.35°. More geometrical and surface properties are summarized in table 9. The FOV picture (fig. 36) is recorded from the GOSAT CAM and shows a bright and extended cloud structure with regular horizontal features.

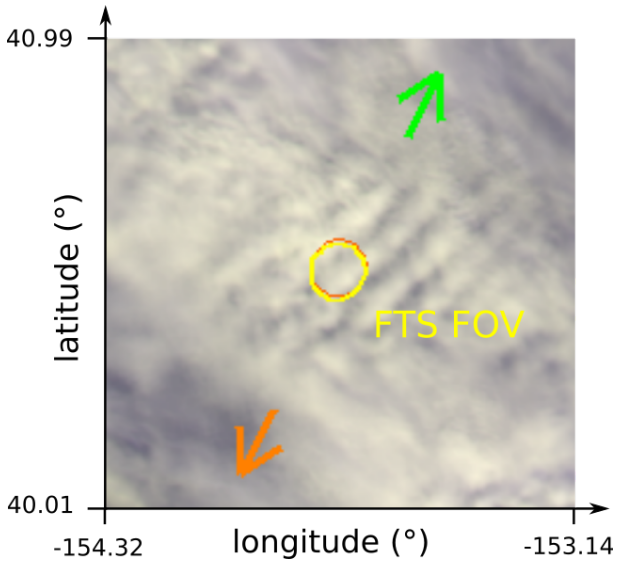

Figure 36: FOV image from the TANSO monitor camera for case study 3 (modified from [5]). The yellow circle shows the TANSO-FTS FOV of the measurement, while the orange circle is the originally assumed FOV before applying the geolocation correction. The directions to the Sun and the satellite from the FOV are drawn in orange and green arrows, respectively.

Table 9: Geometrical characteristics of case study 3 as well as selected surface properties at the measurement location as provided by the GOSAT FTS L1B and L2 IDS products.

longitude	-153.74°	surface of FOV	sea
latitude	40.54°	sunglint	no
timestamp of measurement	2015-03-26 23:27:29	rough terrain flag	no
solar zenith angle	41.05°	satellite zenith angle	22.35°
solar azimuth angle	205.97°	satellite azimuth angle	25.53°

The radiance spectra of the P- and S- polarizations as well as the calculated full radiances of the O_2 A-band are shown in fig. 37. The P- and S- polarized spectra are rather similar which shows that the emerging radiance from the cloud structure is mostly non-polarized.

The collocation situation together with the cloud optical depth from TANSO-CAI is pictured in fig. 38. The TANSO-FTS FOV is situated within the black circle which coincides with a small number of CALIOP measurement points at the eastern edge. The parallel cloud features from the FOV picture can also be identified from the optical depth values. The CPR measurement track does not cross the TANSO-FTS FOV, but the nearest measurement to the TANSO-FTS

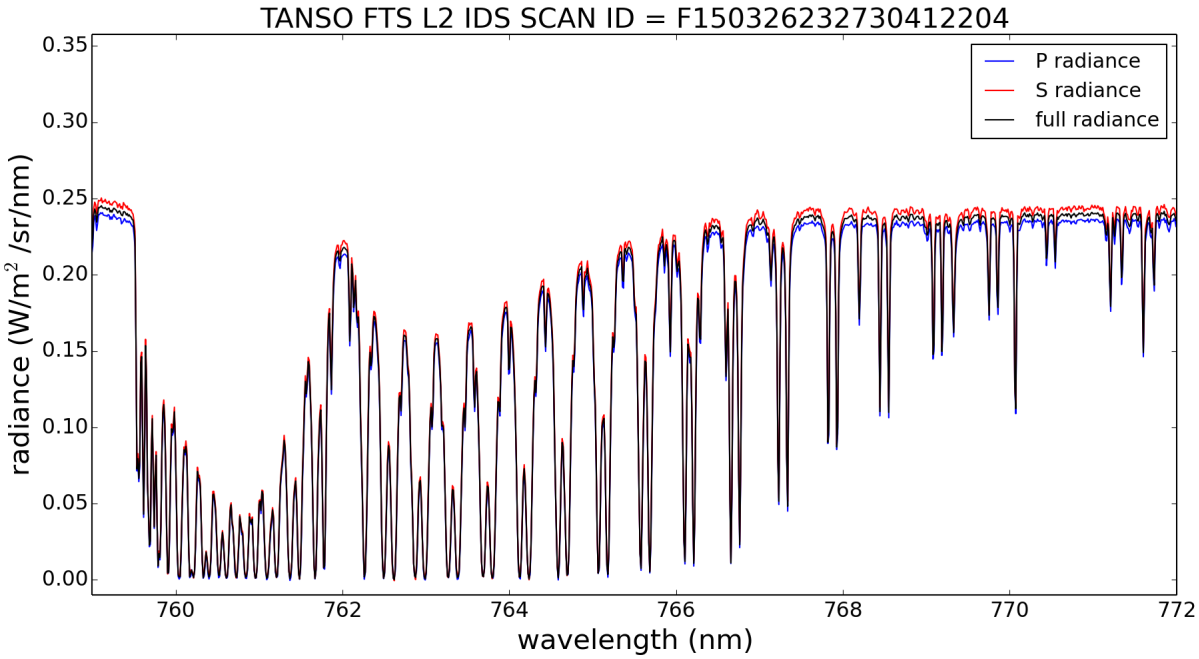

Figure 37: The TANSO-FTS radiances of case study 3 between 759 and 772 nm. The polarized radiances (P in blue and S in red) are shown together with the approximation of the full radiance spectrum (black).

center of FOV within a measurement time frame of ± 25 minutes is taken into account. The datafiles used for the cloud characterization are found in section 7.7, table 12.

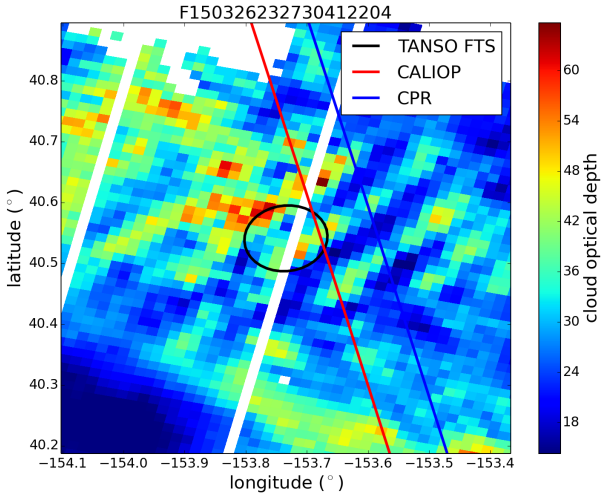

Figure 38: Cloud optical depth from TANSO-CAI together with the TANSO-FTS FOV (black circle) and the CALIOP (red) and CPR (blue) measurement ground tracks for case study 3. White pixels identify invalid data.

The corresponding TANSO-CAI cloud property product includes 37 good and 8 poor pixels (the poor pixels are marked white in fig. 38) within the considered TANSO-FTS FOV. All valid pixels are classified as cloudy with an average cloud optical depth of 34.78 ± 8.59 (SD). According to the TANSO-CAI measurements, the probability of high cirrus clouds is given for the entire FOV region of TANSO-FTS.

As mentioned before, the CALIOP measurements only cross the eastern part of the TANSO-FTS FOV (the number of coinciding measurements is 8). The mean distance of these measurement

points to the center of the TANSO-FTS FOV is 4.58 ± 0.25 km (SD). In average, approximately 4-5 minutes separate the measurements of TANSO-FTS and CALIOP. It should be kept in mind that the cloud structure might have slightly changed between both satellite overpasses.

The retrieved cloud top and base altitudes are shown in fig. 39. All measurements within the TANSO-FTS FOV (red box) show a one-layer cloud. The averaged cloud top altitude is 4.98 km and the averaged cloud base altitude is estimated to 4.74 km, but the latter information should be regarded as less confident because of the reduced sensitivity of lidar measurements for atmospheric regions below absorbing and scattering layers [12]. Cloud layers detected at altitudes above 8 km are not present.

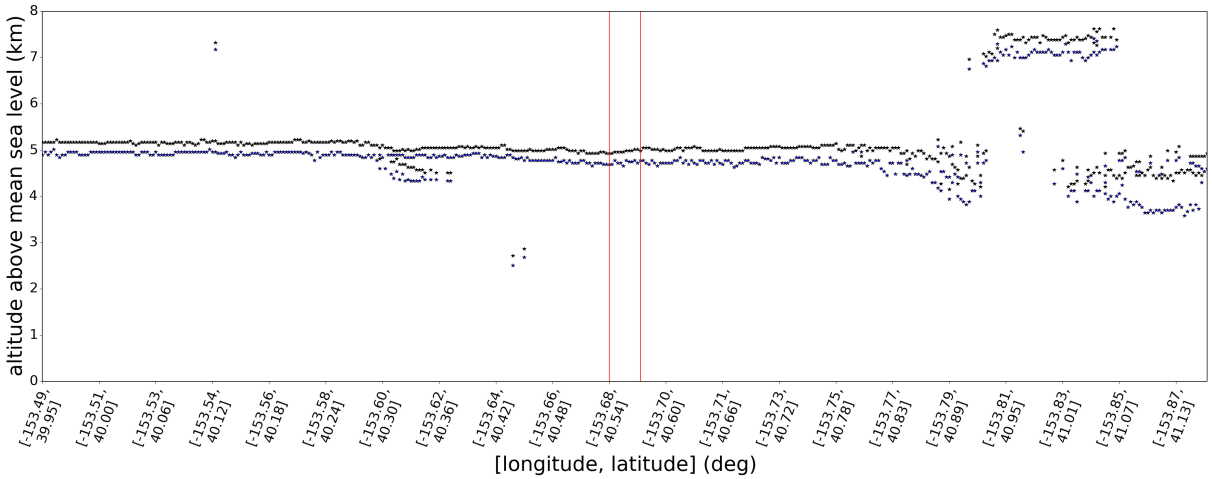

Figure 39: Cloud layers identified from CALIOP measurements (top layers in black, bottom layers in blue) and the location of the TANSO-FTS FOV (red boundaries) for case study 3. Near the TANSO-FTS FOV, the cloud top height appears spatially homogeneous.

The aerosol information shows no detected aerosol layers in the TANSO-FTS FOV region. Because of the large horizontal averaging of this data product, only one available data point corresponds to coordinates within the TANSO-FTS FOV.

As described before, the CloudSat satellite does not have a collocated overpass with GOSAT but a nearby measurement is available with a time difference of approximately 5-6 minutes. The distance to the TANSO-FTS center of FOV is 12.42 km. Considering that the TANSO-FTS FOV has an approximate diameter of 10.5 km, the CloudSat CPR measurement is located slightly more than 7 km outside of the field of view.

The retrieved total cloud optical depth at this data point is 23.18 with an estimated uncertainty of 0.76. The altitude resolved cloud optical depth until 8 km is shown in fig. 40. In contrast to the corresponding results from CALIOP measurements, the dominant cloud structures are found at lower altitudes, namely between approximately 0.8 and 2.0 km. The periodicity of the cloud can again be identified. The CloudSat cloud classification identifies a midlatitude Nimbostratus between the altitude bins of 0.45-0.68 km and 5.0-5.24 km. Characteristics of liquid precipitation are detected in all but the outer 240 m - cloud layers.

Case 3 represents a very interesting situation, considering that the CALIOP sensor only detects a thin cloud layer at the top of the cloud structure. The radar measurements from CPR on the other hand reveal a cloud of large vertical extent with a spatially structured cloud of high and varying optical depths in low altitudes bins and very small optical depths within the bins from 1-2 km to the cloud top.

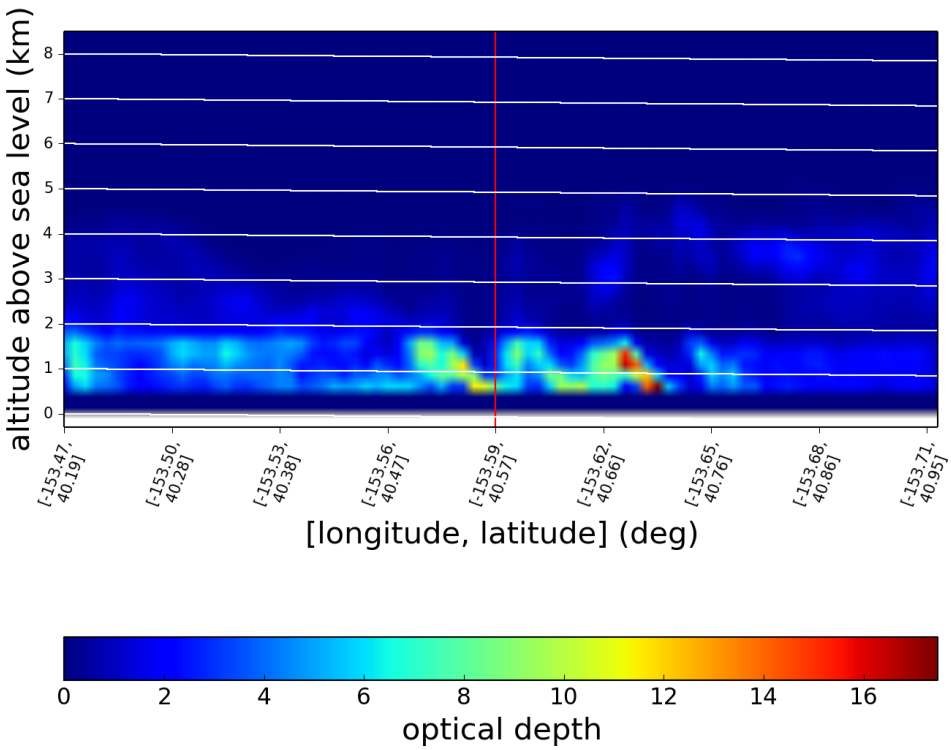

Figure 40: Cloud optical depth for each radar bin from the CPR derived cloud optical depth data product. The main extinction takes place below 2 km, less extinctive structures are found up to altitudes of approximately 5 km. The red line indicates the measurement nearest to the TANSO-FTS field of view.

7.7 Appendix

7.7.1 Datasets used for the cloud characterization

Table 10: Data products used for the cloud scenario characterization of case study 1.

data	filename
CAI cloud optical depth	GOSATTCAI2009060421500380250_02CLDPV02000.h5
CALIOP cloud layers [15; 19]	CAL_LID_L2_333mCLay-ValStage1-V3-01.2009-06-04T21-33-11ZD.hdf
	CAL_LID_L2_05kmCLay-Standard-V4-10.2009-06-04T21-33-11ZD.hdf
CALIOP aerosol layers [16]	CAL_LID_L2_05kmCLay-Standard-V4-10.2009-06-04T21-33-11ZD.hdf
CPR cloud optical depth	2009155210153_16507_CS_2B-TAU_GRANULE_P_R04_E02.hdf
CPR cloud classification	2009155210153_16507_CS_2B-CLDCLASS_GRANULE_P_R04_E02.hdf

Table 11: Data products used for the cloud scenario characterization of case study 2.

data	filename
CAI cloud optical depth	GOSATTCAI2009082823590420240_02CLDPV02000.h5
CALIOP cloud layers [13; 19]	CAL_LID_L2_333mCLay-ValStage1-V3-01.2009-08-28T23-25-17ZD.hdf
	CAL_LID_L2_05kmCLay-Standard-V4-10.2009-08-28T23-25-17ZD.hdf
CALIOP aerosol layers [14]	CAL_LID_L2_05kmALay-Standard-V4-10.2009-08-28T23-25-17ZD.hdf
CPR cloud optical depth	2009240225751_17746_CS_2B-TAU_GRANULE_P_R04_E02.hdf
CPR cloud classification	2009240225751_17746_CS_2B-CLDCLASS_GRANULE_P_R04_E02.hdf

Table 12: Data products used for the cloud scenario characterization of case study 3.

data	filename
CAI cloud optical depth	GOSATTCAI2015032623270410240_02CLDPV02000.h5
CALIOP cloud layers [17; 20]	CAL_LID_L2_333mCLay-ValStage1-V3-30.2015-03-26T22-56-14ZD.hdf
	CAL_LID_L2_05kmCLay-Standard-V4-10.2015-03-26T22-56-14ZD.hdf
CALIOP aerosol layers [18]	CAL_LID_L2_05kmALay-Standard-V4-10.2015-03-26T22-56-14ZD.hdf
CPR cloud optical depth	2015085223232_47400_CS_2B-TAU_GRANULE_P_R04_E06.hdf
CPR cloud classification	2015085223232_47400_CS_2B-CLDCLASS_GRANULE_P_R04_E06.hdf

8 Spectrum fit

8.1 Motivation

In the framework of this study, TANSO-FTS measurements of the oxygen A-band are compared to radiative transfer simulations with the McArtim model. In this context, a fit function was developed which is described in the following together with a number of tests and sensitivity studies.

The development of a fit routine is useful to test which simulated spectrum from a dataset of radiative transfer calculations fits best to the TANSO-FTS measurement. The model atmosphere for the simulations is adapted to the GOSAT measurement geometry (see section 6.2.4 for details). Additionally, atmospheric parameters like the O_2 absorption, the number density of oxygen molecules and pressure-temperature profiles are included. The simulated radiances used are the first elements of the Stokes vectors. During the simulation, polarization of the light is taken into account as well as rotational Raman scattering.

The dataset of simulated radiance spectra is created by assuming different cloud scenarios within the model atmosphere. In our case, the cloud scenarios range from one- to two-layered cloud structures of different cloud top heights h_{cloud}, cloud vertical extents e_{cloud} and cloud optical depths τ_{cloud}. The scattering characteristics of the clouds are represented by the Henyey-Greenstein phase function, adding the asymmetry parameter g to the parameter list. After completion of the radiative transfer calculations, each resulting radiance spectrum is convolved with the TANSO-FTS instrument slit function (see section 5.4).

This chapter is structured as follows: First, the free cloud parameters are reviewed, focussing on the possible reduction of the number of these free parameters. Then, a fit function is presented which takes small deviations from the measurement properties into account. For the fit, a wavelength shift of the simulations in comparison to the measurement has to be considered, which is mostly caused by the spectral calibration and degradation of the TANSO-FTS instrument (see section 5). After explaining further details and options of the fitting procedure, some tests are presented to verify the proper functioning of the fitting process.

8.2 Cloud parameters

The calculation of synthetic spectra is performed for different cloud scenarios which are based on the following macroscopic cloud characteristics:

- the cloud top height h_{cloud}
- the cloud vertical extent e_{cloud}
- the cloud optical depth τ_{cloud}

The microscopic characteristics of the cloud is described by the asymmetry parameter g.
In case of clouds composed of multiple vertical layers, different indices serve to distinguish the individual cloud parameters.

The first step in the retrieval is the reduction of the number of free cloud parameters, which is especially important in the case of multiple cloud layers. Assuming exclusively water cloud particles, an asymmetry parameter of 0.85, which is the typical value for water clouds [82], is used if not stated otherwise.

In the case of one-layered cloud scenarios, the remaining free parameters are h_{cloud}, e_{cloud} and τ_{cloud}. If more than one cloud layer is implemented into the simulation atmosphere, cloud optical depths are expressed as a total optical depth value τ_{all} of the cloud system and a fraction q_i

with $\tau_i = q_i \tau_{all}$ and $\sum_{i=1}^{N} q_i = 1$. A total cloud optical depth of 15, divided into 2 cloud layers with q_1=1/3 and q_2=2/3 thus leads to $\tau_1 = q_1 \tau_{all} = 5$ and $\tau_2 = q_2 \tau_{all} = 10$.

In order to save computational time for the spectrum simulations, the number of simulations is reduced by fixing the cloud optical depth τ_{cloud} through comparison of the continuum radiances of the measured radiances to the simulations. Here, the continuum (or "continuum region") is defined as the wavelength region of the highest radiances where (almost) no oxygen absorption occurs.

In our retrieval, convolved simulated radiances between 770.74 and 770.86 nm (continuum A, see red box in figure 41) are averaged to a mean radiance value I_{cont}. This spectral region was chosen because the absorption of molecular oxygen is negligible. One could also use non-convolved radiances because the convolution should not significantly change the values but it provides a smoothing of the radiances which are object to statistical noise from the Monte Carlo simulations. The noise effect is further reduced by the spectral averaging.

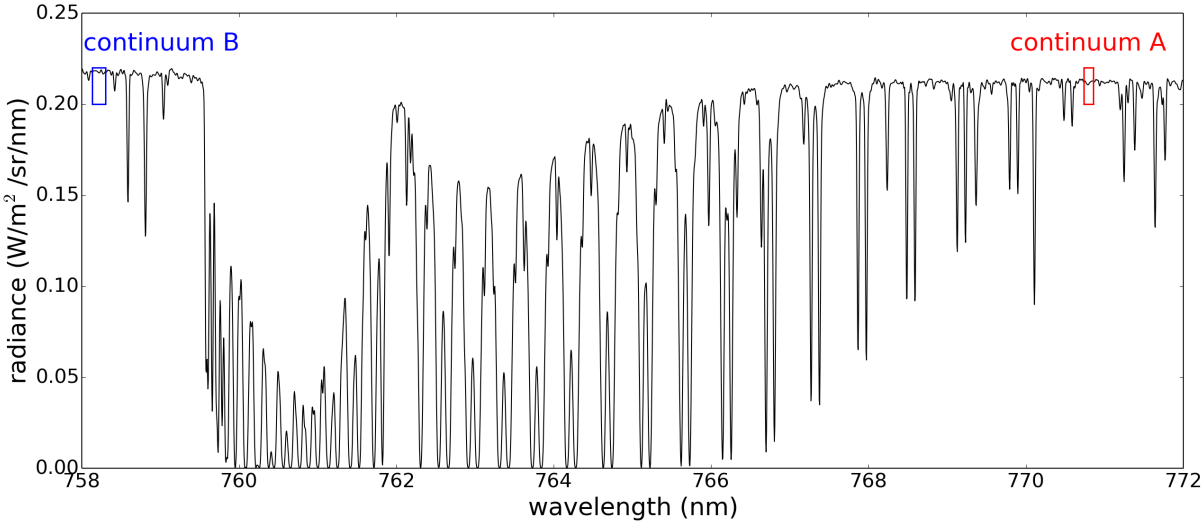

Figure 41: Two possible regions of the spectra outside of the absorption bands ("continuum radiances"). For the retrieval of the cloud optical depth, only continuum A is used as explained later in the text.

The reference continuum radiance I_{ref} of the TANSO-FTS measurement is calculated by averaging all radiances within the mentioned wavelength region. Because the spectral shift is not yet considered at this stage, the wavelength region is chosen in a way that both the simulated and the (slightly shifted) measured radiances are within the continuum region and thus do not contain any absorption bands.

The cloud optical depth τ_{cloud} can be determined by using the fact that the measured radiance increases approximately linear with the natural logarithm of τ. This is especially true if the optical depth takes values between 3 and 30 [38].

The retrieval procedure for τ_{cloud} is demonstrated in figure 42 and explained in the following:

In order to determine the cloud optical depth for a given cloud scenario, boundary values for τ_{cloud} are set, here $\tau_{b1} = 10$ and $\tau_{b2} = 25$. For both scenarios, the radiances in the continuum region of the spectrum are simulated with McArtim and the averaged continuum radiances I_{b1} and I_{b2} are calculated (corresponding to the cloud scenarios with τ_{b1} and τ_{b2} respectively). Linear interpolation between the natural logarithms of the cloud optical depths $[\ln(\tau_{b1}), \ln(\tau_{b2})]$ and the averaged radiances $[I_{b1}, I_{b2}]$ are performed (black interpolation curve in figure 42). The interpolated value at the reference radiance I_{ref} leads to an estimation of the cloud optical

depth $\tau_{b3} = 16.44$. After another radiative transfer simulation using the derived τ_{b3} for the cloud scenario, the resulting radiance I_{b3} is checked for the agreement to the reference radiance I_{ref}. If the agreement is better than 0.2 %, τ_{b3} is considered the final cloud optical depth τ_{cloud}. In this example, the condition is not met, consequently $\ln(\tau_{b3})$ and I_{b3} are included in the interpolation (blue interpolation curve), a new estimated cloud optical depth τ_{b4} is determined and radiative transfer simulations are repeated for this value. This process is continued until the agreement of the continuum radiances is better than 0.2% or the new continuum radiance fits worse to the measurement than the previous one (this is possible due to remaining noise effects). In the example, the deviation of I_{b4} to I_{ref} is smaller than 0.2%, thus $\tau_{cloud} = \tau_{b4} = 16.08$ (red point).

If the true cloud optical depth is outside the chosen boundaries, I_{b3} will be smaller than I_{b1} or larger than I_{b2}. The boundaries will then be extended and the process restarted.

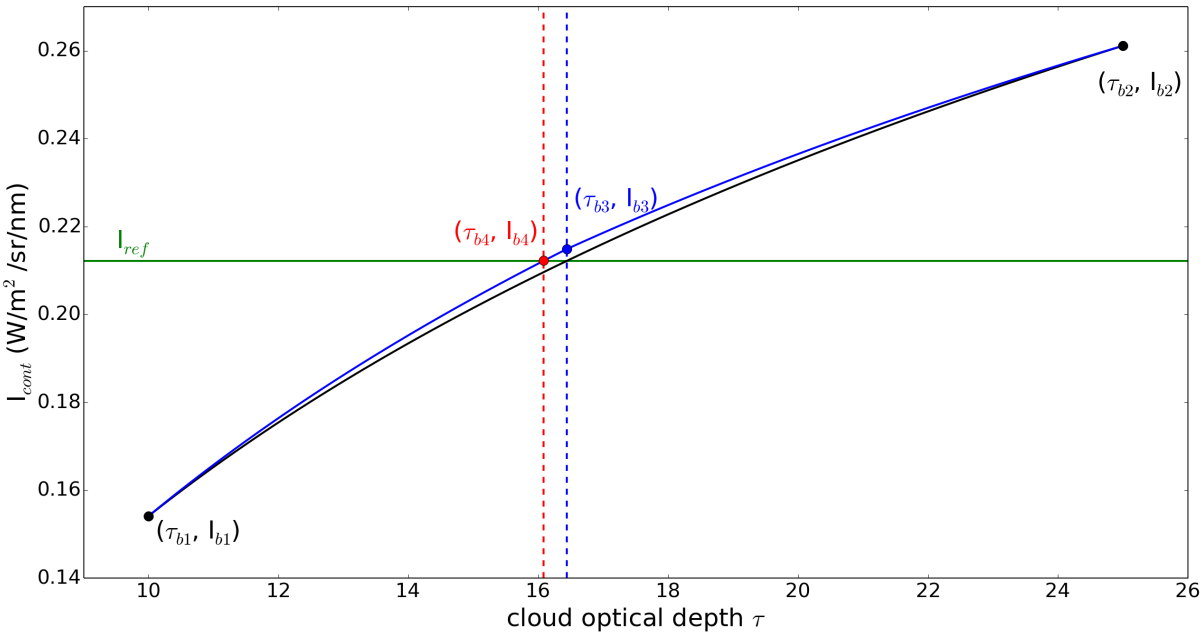

Figure 42: An example of how the cloud optical depth $\tau_{cloud} = \tau_{b4}$ for a given cloud scenario is determined. τ_{cloud} is approached in an iterative way, using interpolations between simulated continuum radiances and the corresponding cloud optical depths. The process is explained in the text.

In the cloud optical depth retrieval, only continuum region A is considered. In order to test if continuum radiances left of the deep absorption band should also be included, the differences between the continuum radiances A and the continuum radiances B (wavelengths between 758.134 and 758.305 nm, see figure 41) are considered. For the TANSO-FTS spectra, the retrieved dataset of collocated measurements (see section 7) is used to calculate average radiances of continuum A and B. The relative deviation of $I_{cont,B}$ to $I_{cont,A}$ is calculated, the mean and standard deviations are shown in table 13. As representative examples of the simulations, the three case studies (case study 1 with g=0.8 and g=0.85, case studies 2 and 3 with g=0.85) with the corresponding simulated cloud scenarios are chosen.

The mean relative deviation of $I_{cont,B}$ to $I_{cont,A}$ is -3.38% for the TANSO-FTS spectra with a standard deviation of 0.49%. The deviation is relatively small and similar for the considered spectra. The standard deviation is below the statistical simulation error of 1%. The simulations show deviations of approx. -2.8% to -2.9% which is almost equal for the three considered cases, the standard deviation is also approximately constant ($\approx 0.3\%$). $I_{cont,B}$ is always a little lower than $I_{cont,A}$ but the relative differences between continuum A and B of the different cases are very small and the differences between the simulated and measured relative differences is below

Table 13: Differences between the averaged continuum radiances of continuum A and B for measured TANSO-FTS spectra and McArtim simulations.

dataset	mean($\frac{I_{cont,A}-I_{cont,B}}{I_{cont,A}}$)	std($\frac{I_{cont,A}-I_{cont,B}}{I_{cont,A}}$)	number of spectra
TANSO-FTS spectra	-3.38%	0.49%	222
simulations case study 1	-2.90%	0.28%	240
simulations case study 2	-2.80%	0.26%	120
simulations case study 3	-2.88%	0.29%	21

1%. Consequently, it was decided that only one continuum should be used. In this retrieval, continuum A is chosen.

8.3 Fit function

To quantify the deviation of the measured spectra and the various model simulations, the simulated spectra are fitted to the measurement. In order to account for small deviations between the simulated radiances and the measurement, a fit function is proposed. The fit model is based on the Lambert-Beer law which describes the absorption of radiation in the atmosphere (see section 4.4):

$$I(\lambda) = I_0(\lambda) \exp\left(-L\left(\sum_j \sigma_j(\lambda) c_j\right) + \epsilon_R(\lambda) + \epsilon_M(\lambda)\right) A(\lambda) \tag{50}$$

I_0 denotes the solar irradiance, L the effective photon path length, σ_j the absorption cross sections and c_j the mean number densities in the atmosphere. $\epsilon_R(\lambda)$ describes the extinction due to Rayleigh scattering and $\epsilon_M(\lambda)$ due to Mie scattering. Effects from the instrument are described by $A(\lambda)$ [78]. Because the TANSO-FTS instrument does not use direct sunlight, $A(\lambda)$ also includes effects from scattering and reflection within the TANSO-FTS field of view, as explained in section 4.4.

Based on equation 50, a fit function (eq. 51) is proposed including a wavelength shift $\delta\lambda$, a multiplicative parameter A, a parameter B allowing for variations of the simulated O_2 absorption and a constant parameter C. In principle, the parameters A, B and C are designed to be polynomial functions of the wavelength λ. Within this thesis, they are considered to be constant, consequently the wavelength dependency is omitted in the following. The fit function is as follows:

$$I_{fit}(\lambda) = A \cdot I^B_{sim,ILSF}(\lambda + \delta_\lambda) + C \tag{51}$$

with

$$\delta_\lambda = \delta_{shift} + \delta_{squeeze} \cdot (\lambda - \overline{\lambda}) \tag{52}$$

The parameters are now individually described:

Wavelength alignment

In comparisons between measurement and simulations, a wavelength alignment should always be performed to account for small changes in the spectral positions. These can arise from temperature variations within the spectrometer and errors in the calibration of the measured wavelengths [78].

A wavelength shift between the TANSO-FTS spectra and the simulation spectra can indeed be observed. The TANSO-FTS spectra are subject to a spectral shift and additional spectral degradation over the lifetime (for details, see section 5.3.2). The discrepancy becomes obvious when comparing the positions of the solar Fraunhofer lines. The most probable reason for the shift is a change in the optical alignment of the laser beam. It can be corrected by applying a calibration factor which shows an exponential decrease between 2009 and 2011 with less than 10 ppm change [46]. The correction factor, as published by Kuze et al. [46], is represented in figure 43. Because the exact correction factors were not available to the public, the spectral calibration, applied to the TANSO-FTS spectra before the fit, is approximated by a correction factor of 0.99997. This significantly reduces the magnitude of the necessary wavelength correction within the fitting process.

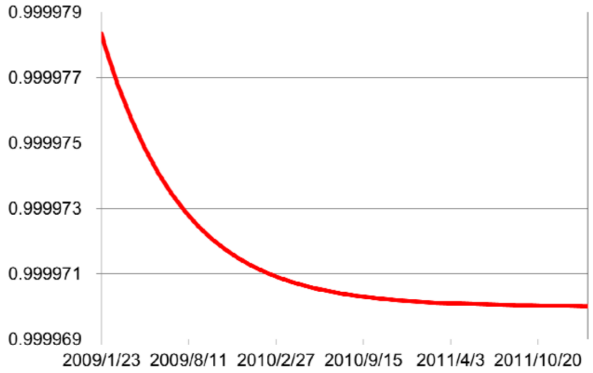

Figure 43: Spectral calibration factor of the TANSO-FTS instrument, which is a function of the days from launch (taken from Kuze et al. [46]). The plot shows the data for the first two years in orbit.

The wavelength alignment can be performed by applying a spectral shift within the fit routine. Following the description by Stutz and Platt [94], it is applied as follows:

$$\lambda_i' = \lambda_i + \delta_{\lambda_i} \tag{53}$$

$$\delta_{\lambda_i} = \sum_{j=0}^{J} \delta_j \left(\lambda_i - \overline{\lambda}\right)^j \tag{54}$$

$\overline{\lambda}$ is the average wavelength value within the considered boundaries, δ_j are called the alignment parameters and $\delta_j \left(\lambda_i - \overline{\lambda}\right)^j$ the alignment factors. In our case, the order of polynomial J=1 was chosen, but in principle higher orders are possible. The alignment factor of order 0 describes a constant spectral shift while the first order alignment factor performs a spectral stretch or squeeze [78], depending on the sign of δ_j. The above representation assures that the alignment parameters δ_j are zero if the wavelength positions are not shifted [94]. In the cases considered here, the spectral stretch δ_1 usually takes values near zero and has thus a small effect on the total spectral shift δ_λ.

Fit parameter A

The parameter A describes a multiplicative factor to the radiance. It can thus account for deviations of the input values in the simulations (e.g. surface albedo or cloud optical depth) as well as for instrument effects, or in the case of GOSAT, for uncertainties in the calibration/degradation of the measured radiances. For TANSO-FTS, both the radiance conversion as well as the application of the radiometric degradation is performed via a wavelength dependent multiplicative factor to the radiance. Here, it should be noted that in the case of cloud covered atmospheres,

the cloud optical depth within this retrieval is determined and applied to the radiative transfer simulations before the fitting process. Deviations of A from unity are thus expected to be very small.

Fit parameter B

The parameter B is an exponent to the radiance and the central parameter affecting the oxygen absorption. It acts as a multiplicative factor to the optical depth τ when considering $I \sim \exp(-\tau)$ (compare eq. 20) and accounts for deviations in the simulated O_2 absorption. Only in cases for which the O_2 absorption in the simulation shows large deviations from that in the measured spectrum (e.g. if the cloud top height is wrongly chosen), B values significantly below or above unity are expected. This is important for the interpretation of the fit results. During the determination of the parameter B within fit routine, the simulated radiances are temporarily expressed in a normalized way. This ensures that only the radiances within the absorption structures are affected by the parameter B, and that it does not change the value of the continuum radiances which is restricted by the chosen cloud optical depth τ_{cloud}.

Fit parameter C

The parameter C is an additive term to the radiance and includes all other (more or less) constant intensity contributions which may cause deviations of the simulated radiances from the measurement. The intensity contributions mainly include straylight from the spectrometer, remaining structures from rotational Raman scattering and effects from fluorescence [27; 101].

Choice of parameters

Except the wavelength alignment, all parameter are implemented as options that can be switched on or off as needed. In this study, only the parameter B is allowed as a free fit parameter and used as a polynomial of degree zero. We assume that the TANSO-FTS calibrations are correct and that straylight effects are negligible, which eliminates the use of parameters A and C. They are omitted in the further description of the fit routine. The results of the parameter B thus gives an evaluation of how well the simulated cloud scenario fits to the measurement. Together with the resulting fit residuals, the best-matching cloud scenario can be determined.

The fit results in residual structures where the residual is defined at the difference between the measured and fitted radiance value at each wavelength. In our retrieval, the residuals are quantified by the RMS (root mean square):

$$RMS = \sqrt{\frac{\sum_{i=1}^{N} (res_i)^2}{N - \nu}} \tag{55}$$

The variables are as follows: res_i are the resulting residuals, N the number of data point used within the fit and ν the number of free parameters.

Negligence of solar Fraunhofer lines during the fit evaluation

A spectrum of the solar irradiance is used for the simulation of the radiances. Since a direct solar reference spectrum from GOSAT is not available, the disc-integrated spectrum from Toon [97] is used and calibrated as described in section 6.2.2. The fit result should only depend on the O_2 absorption structures and not on potential errors of the solar reference spectrum in the measurement situation. This concerns especially mismatches within the Fraunhofer lines and they should consequently be excluded during the evaluation of the fit result. In this study, only deep Fraunhofer lines are not considered. Figure 44 shows the oxygen optical depths τ_{O_2} in the upper plot and the negative logarithm of the normalized calibrated solar reference spectrum in

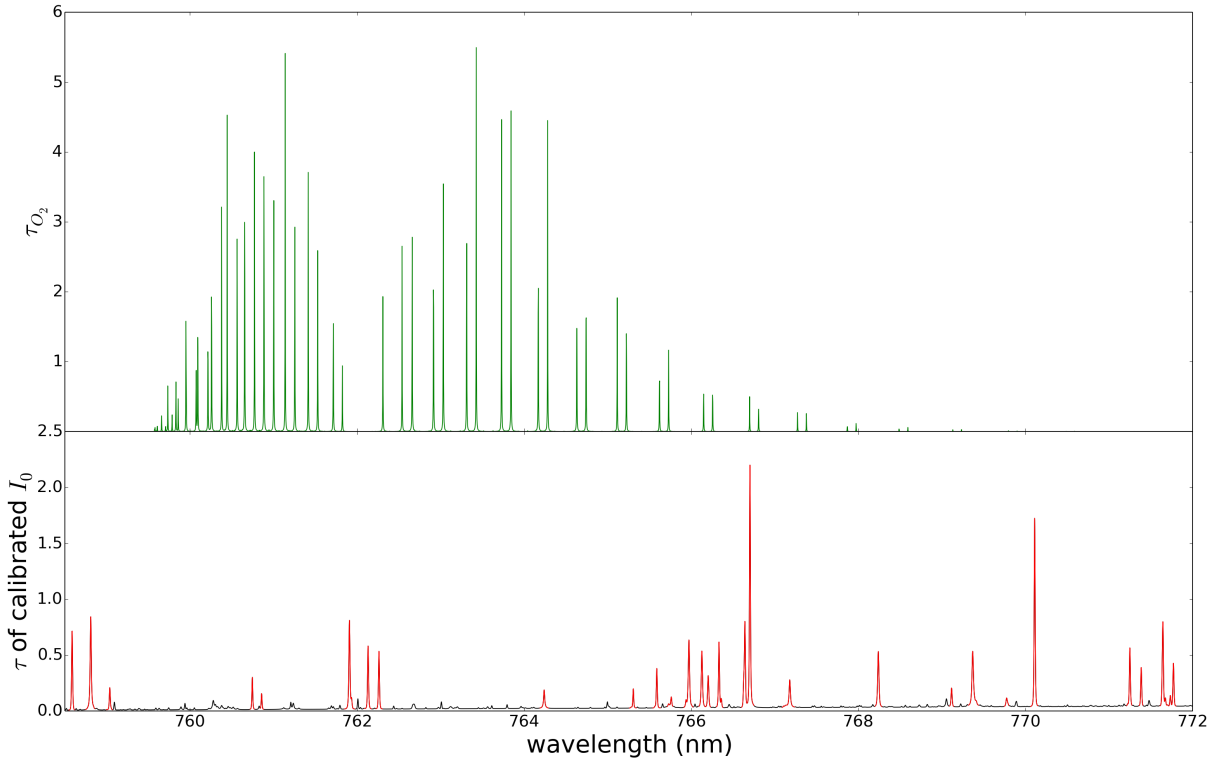

Figure 44: Oxygen optical depth τ_{O_2} (top) and negative logarithm of the normalized calibrated solar reference spectrum (Toon [97], bottom). Deep Fraunhofer lines are marked in red and are excluded from the fit evaluation.

the lower plot. All Fraunhofer structures containing radiances smaller than 1.15 W/m^2/sr/nm (marked red in figure 44) are excluded from the fit. Whenever possible, the boundaries of the excluded regions are chosen in a way that they are not situated in the vicinity of the O_2 absorptions.

Wavelength windows used for the fit

First, a fit including the entire O_2 A-band is performed. The O_2 A-band is composed of absorption structures of very different strength. It is also interesting to consider those separately within different subwindows. Figure 45 shows a simulated spectrum of the O_2 A-band and three selected subwindows. Subband 1 includes some of the strongest absorptions and is thus most sensitive to short photon path lengths because the absorption will be saturated for longer paths. Subband 2 considers intermediate oxygen absorption and subband 3 includes only weak absorption structures which are most sensitive to very long path lengths.

The fit results of the small spectral regions might allow further interpretation of the case study results:

- If all input parameters of the radiative transfer simulations are chosen in a "perfect" way, a good agreement between the simulated and measured spectra is expected.
- If the agreement of the spectral structures is good in the strongly absorbing subband 1 but a disagreement in the less absorbing subbands 2 and 3 is found, the distribution of short path lengths is probably close to the truth but the distribution of moderate and long path lengths might be wrong.
- Similar conclusions may be drawn from the agreement/disagreement of other subbands.

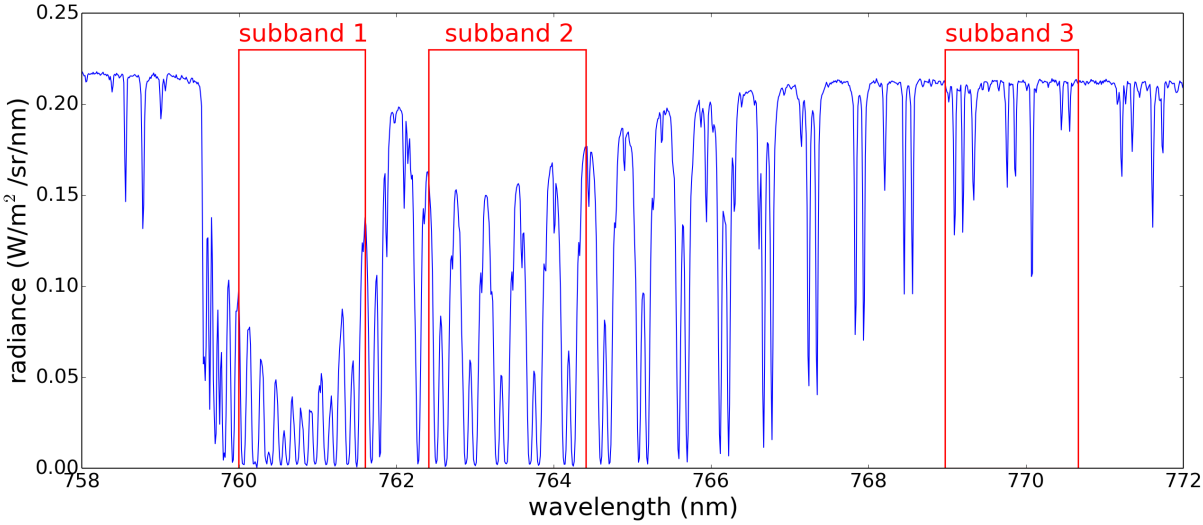

Figure 45: A simulated spectrum of the O_2 A-band with a cloud between 1.0 and 1.4 km and cloud optical depth of 16.08. Subbands 1-3 are indicated by red boxes and highlight strong, intermediate and weak oxygen absorptions, respectively.

Technical details and interpretation of results

The fitting procedure, which is applied after the determination of the cloud optical depths τ_{cloud}, is summarized in the following list and schematically shown in fig. 46:

- Simulated spectra with different cloud scenarios $I_{sim,ILSF}(\lambda_{sim})$: The radiative transfer model McArtim is used to simulate O_2 A-band spectra for different cloud situations on a wavelength grid spaced by $\Delta\lambda_{sim}$ =0.005 nm. The resulting spectra I_{sim} are convolved with the approximated GOSAT instrument slit function. The ILSF is given in wavenumber units on a grid of 0.01 cm^{-1}. For the convolution, the simulated spectrum is interpolated to the finer grid size, convolved, and transformed back to the original wavelength grid. The resulting radiances $\mathrm{I}_{sim,ILSF}(\lambda_{sim})$ are used as input for the fit routine.

- Measurement spectrum $I_{FTS}(\lambda_{FTS})$: The TANSO-FTS spectrum which is used in the fitting process, is not changed except the approximate calibration of the wavelengths described above.

- Determination of the wavelength boundaries for the fit evaluation: The wavelength boundaries are drawn from the TANSO-FTS spectrum $I_{FTS}(\lambda_{FTS})$ and $I_{sim,ILSF}(\lambda_{sim})$ and only include the common wavelength region. The chosen wavelength spacing is the one from TANSO-FTS which is about $\Delta\lambda_{FTS} \approx 0.012$ nm in the O_2 A - band. Here, also smaller wavelength regions can be selected, as in the case of the subband fits.

- Fit: The fit is based on the *minimize()* function of the *lmfit package* [66], and a Levenberg-Marquardt algorithm is used. The fit function varies and optimizes the parameter B and the wavelength alignment at the same time. Although the fit function is applied to the original wavelength grid from the simulated spectra ($\Delta\lambda$ =0.005 nm), the fitted radiances are interpolated to the TANSO-FTS wavelength grid before the residual is calculated, neglecting the radiances of deep Fraunhofer lines. The coarser TANSO-FTS wavelength grid ($\Delta\lambda \approx 0.012$ nm) is chosen to avoid undersampling effects through interpolation of the measured data.

- Fit result: From the outcome of the fitting process, the parameters δ_{shift}, $\delta_{squeeze}$ and B as well as the resulting RMS are derived and saved. Also, the fitted spectrum and the residuals are calculated on the GOSAT wavelength grid.

The fit result can be evaluated from the derived parameters B and the resulting RMS values. A small RMS value indicates a good agreement between the simulated and measured spectra. If the simulated cloud model meets the measurement situation, a value of B near unity is expected. If B is smaller (larger) than 1, the oxygen absorption is over- (under-) estimated in the simulation, for example the cloud top height is wrongly chosen. From the deviation of B from unity, one can draw conclusions on how the cloud model has to be modified to obtain a better agreement between simulation and measurement.

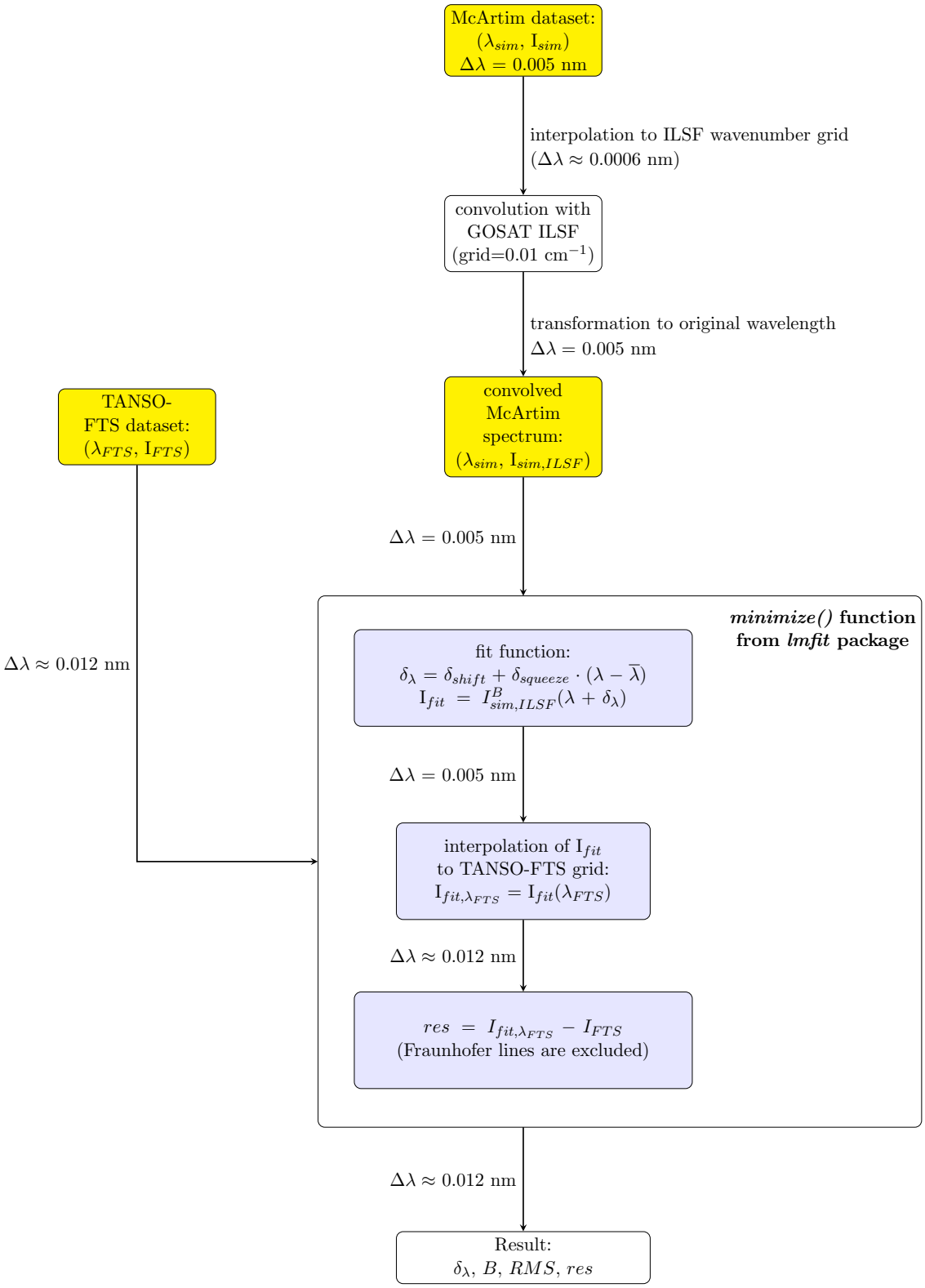

Figure 46: Schematic overview of the fitting procedure.

8.4 Test fits and sensitivity tests

The TANSO-FTS measurement from case 1 has been selected for testing purposes of the proposed fit. The cloud situation is described in detail in the section 7.6.1. Also, a synthetic spectrum calculated by radiative transfer simulations is used for a number of tests.

Using the radiative transfer model McArtim, 120 different cloud scenarios are implemented into the model atmosphere and radiance simulations are performed. The cloud parameter boundaries are loosely based on the information about the cloud scene derived from collocated satellite measurements. The cloud structure is defined as a homogeneous one-layer system with the cloud optical depth τ_{cloud} and different cloud top height h_{cloud} and vertical extent e_{cloud}. The microphysics are described by a Henyey-Greenstein phase function with asymmetry parameter g = 0.85. The cloud optical depth for each cloud scenario is determined as described in the beginning of the chapter. The boundary values as well as the step size for the cloud parametrization are summarized in table 14, all parameter combinations together with an identification number are shown in table 15.

Table 14: Parameter boundaries for the simulated model clouds (case study 1), characterized by the cloud top height h_{cloud}, the cloud vertical extent e_{cloud} and the cloud optical depth τ_{cloud}.

parameter	min. value	max. value	step size
h_{cloud}	0.2 km	3.0 km	0.2 km
e_{cloud}	0.2 km	h_{cloud}	0.2 km
τ_{cloud}	15.91	16.38	n/a

The full radiance spectrum as well as the performed radiative transfer simulations serve as input for the following tests: First, the results using a synthetic spectrum are discussed. Second, the influence of the surface albedo is tested. The next test probes the accuracy of the used instrument line shape function. The influence of the asymmetry parameter g is discussed in the results section, more explicitly in section 10.5.

8.4.1 Synthetic reference spectrum

In order to test the proper functioning of the fit routine, a synthetic spectrum acts as measurement spectrum ("reference spectrum" in the following). To create this test spectrum, a simulated radiance spectrum from the McArtim simulations is selected (h_{cloud} = 1.4 km, e_{cloud} = 0.4 km, τ_{cloud} = 16.08, number 22). The radiances are interpolated from the 0.005 nm grid of the McArtim output to the coarser GOSAT grid of about 0.012 nm. Additionally, a spectral wavelength shift of -0.009 nm (approximately the value from a typical TANSO-FTS wavelength alignment) is applied and random white noise is added. The noise level of the TANSO-FTS spectrum can be estimated from the (calibrated and radiometric degradation corrected) imaginary part of the radiance values of the L1B data product [46]. The mean of the imaginary radiances is approximately zero and the standard deviation is 0.0005 W/m^2/sr/nm. Here, larger noise is used, because it should be representative also for TANSO-FTS measurements being subject to higher noise levels. The standard deviation for the synthetic noise is 0.0022 W/m^2/sr/nm and thus is a factor of 4.37 higher than the TANSO-FTS noise for case 1. The resulting reference spectrum is shown in fig. 47, together with the original simulation and the added noise structure for comparison.

First, a simple test is performed to check the proper wavelength alignment. Here, the fit parameter B is set to unity. The resulting δ_{shift} and $\delta_{squeeze}$, summarized in figure 48, result from a wavelength alignment fit of all simulations described in table 15. As can be seen in the results,

Table 15: All performed simulations for case study 1 (g=0.85). Each parameter combination for the model cloud has an assigned number which is used in the result plots.

num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}
0	0.2	0.2	16.14	40	1.8	1.0	16.09	80	2.6	0.6	16.06
1	0.4	0.2	16.17	41	1.8	1.2	16.1	81	2.6	0.8	16.38
2	0.4	0.4	16.09	42	1.8	1.4	16.08	82	2.6	1.0	15.99
3	0.6	0.2	16.1	43	1.8	1.6	16.05	83	2.6	1.2	16.05
4	0.6	0.4	16.03	44	1.8	1.8	16.4	84	2.6	1.4	16.15
5	0.6	0.6	16.13	45	2.0	0.2	16.06	85	2.6	1.6	16.14
6	0.8	0.2	16.08	46	2.0	0.4	16.11	86	2.6	1.8	16.13
7	0.8	0.4	16.11	47	2.0	0.6	16.07	87	2.6	2.0	16.12
8	0.8	0.6	16.18	48	2.0	0.8	16.09	88	2.6	2.2	16.07
9	0.8	0.8	16.12	49	2.0	1.0	16.05	89	2.6	2.4	16.11
10	1.0	0.2	16.1	50	2.0	1.2	16.04	90	2.6	2.6	16.16
11	1.0	0.4	16.14	51	2.0	1.4	16.18	91	2.8	0.2	16.0
12	1.0	0.6	16.11	52	2.0	1.6	16.08	92	2.8	0.4	16.01
13	1.0	0.8	16.14	53	2.0	1.8	16.08	93	2.8	0.6	16.06
14	1.0	1.0	16.03	54	2.0	2.0	16.16	94	2.8	0.8	16.06
15	1.2	0.2	16.14	55	2.2	0.2	16.07	95	2.8	1.0	16.03
16	1.2	0.4	16.14	56	2.2	0.4	16.09	96	2.8	1.2	15.91
17	1.2	0.6	16.09	57	2.2	0.6	16.14	97	2.8	1.4	16.09
18	1.2	0.8	16.12	58	2.2	0.8	16.02	98	2.8	1.6	16.11
19	1.2	1.0	16.08	59	2.2	1.0	16.11	99	2.8	1.8	16.03
20	1.2	1.2	16.13	60	2.2	1.2	16.05	100	2.8	2.0	16.05
21	1.4	0.2	16.07	61	2.2	1.4	16.17	101	2.8	2.2	16.11
22	1.4	0.4	16.08	62	2.2	1.6	16.03	102	2.8	2.4	16.15
23	1.4	0.6	16.09	63	2.2	1.8	16.11	103	2.8	2.6	16.09
24	1.4	0.8	16.05	64	2.2	2.0	16.13	104	2.8	2.8	16.09
25	1.4	1.0	16.16	65	2.2	2.2	16.01	105	3.0	0.2	16.08
26	1.4	1.2	16.05	66	2.4	0.2	16.02	106	3.0	0.4	16.07
27	1.4	1.4	16.13	67	2.4	0.4	16.12	107	3.0	0.6	16.05
28	1.6	0.2	16.13	68	2.4	0.6	16.12	108	3.0	0.8	16.06
29	1.6	0.4	16.14	69	2.4	0.8	16.09	109	3.0	1.0	16.07
30	1.6	0.6	16.19	70	2.4	1.0	16.09	110	3.0	1.2	16.1
31	1.6	0.8	16.13	71	2.4	1.2	16.16	111	3.0	1.4	16.05
32	1.6	1.0	16.12	72	2.4	1.4	16.15	112	3.0	1.6	16.09
33	1.6	1.2	16.08	73	2.4	1.6	16.15	113	3.0	1.8	16.09
34	1.6	1.4	16.06	74	2.4	1.8	16.17	114	3.0	2.0	16.08
35	1.6	1.6	16.1	75	2.4	2.0	16.08	115	3.0	2.2	16.13
36	1.8	0.2	16.04	76	2.4	2.2	16.1	116	3.0	2.4	16.07
37	1.8	0.4	16.17	77	2.4	2.4	16.19	117	3.0	2.6	16.08
38	1.8	0.6	16.17	78	2.6	0.2	16.03	118	3.0	2.8	16.13
39	1.8	0.8	16.13	79	2.6	0.4	15.97	119	3.0	3.0	16.03

all shifts are near the applied shift of -0.009 nm and the squeeze values are very small. A slight variety of values still exists, probably due to the noise which was added to the original simulation. For the considered case, the wavelength alignment works well. Because the fit parameter B is set to unity, obvious differences can be observed in the deep absorptions (compare fig. 49). It is possible that for larger deviations within the absorption depths, the wavelength alignment without adjusting B leads to wrong alignment values. To avoid this, the fit parameter B will be adjusted together with the wavelength alignment from now on.

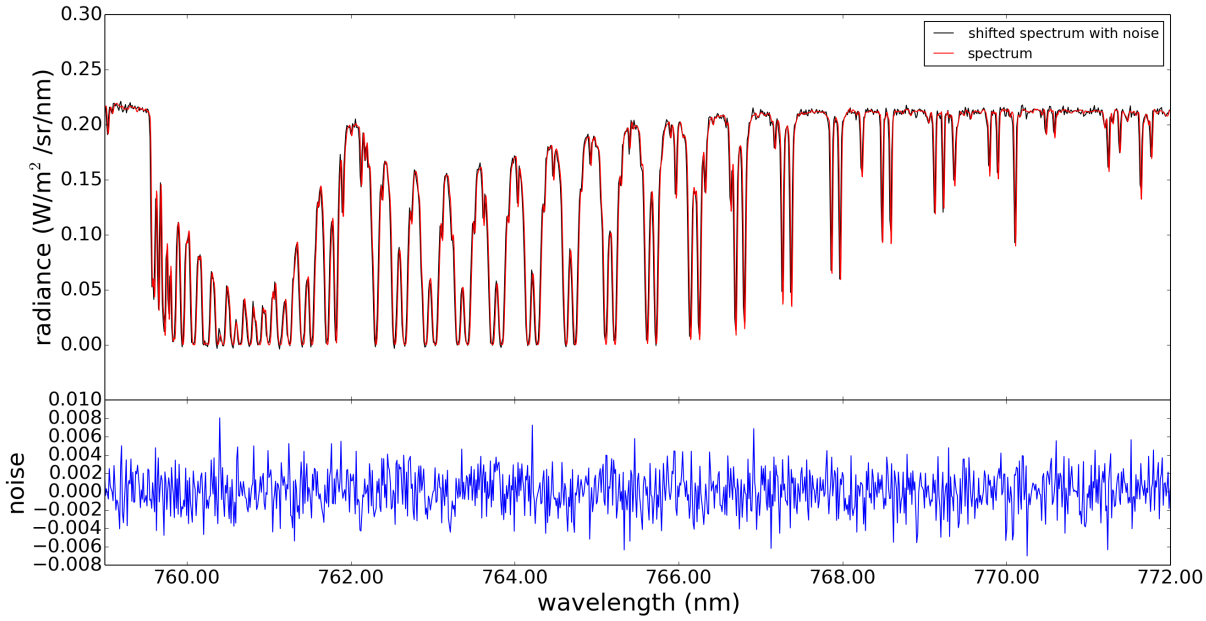

Figure 47: Top: The simulated spectrum of cloud scenario 22 before (red) and after modifications (black), which include a spectral shift and white noise ("reference spectrum"). Bottom: White noise structure added to the spectrum.

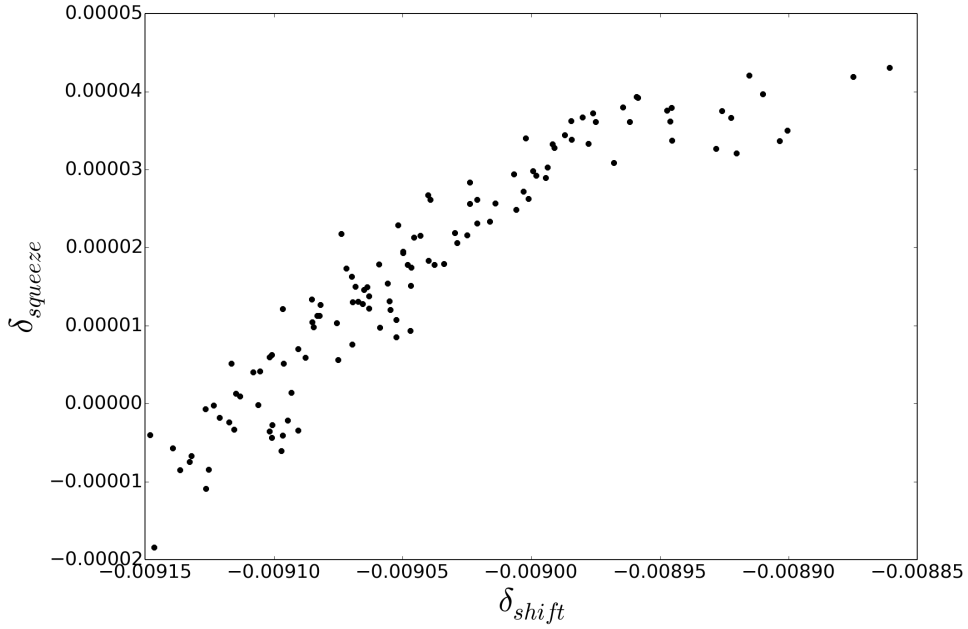

Figure 48: Resulting δ_{shift} and $\delta_{squeeze}$ from the pure wavelength alignment fit using the reference spectrum instead of the measurement. All δ_{shift} values are near the created shift of -0.009 nm.

The resulting spectra from the pure wavelength alignment fit are shown in figure 49. The aligned simulation spectra are drawn in grey, the reference spectrum in red and the simulation with the smallest RMS in green. As expected, the best comparison to the reference spectrum without applying any additional fit parameters, is number 22, thus the spectrum which the reference spectrum is based on. The residual structures, shown in the lower plot (grey for all residuals, the residual with the smallest RMS is highlighted in blue), point out the radiance differences of the different cloud scenarios, which are present especially within the strong and medium absorption structures. Large residuals in the continuum radiances are not to be expected because they are already optimized by the choice of τ_{cloud}.

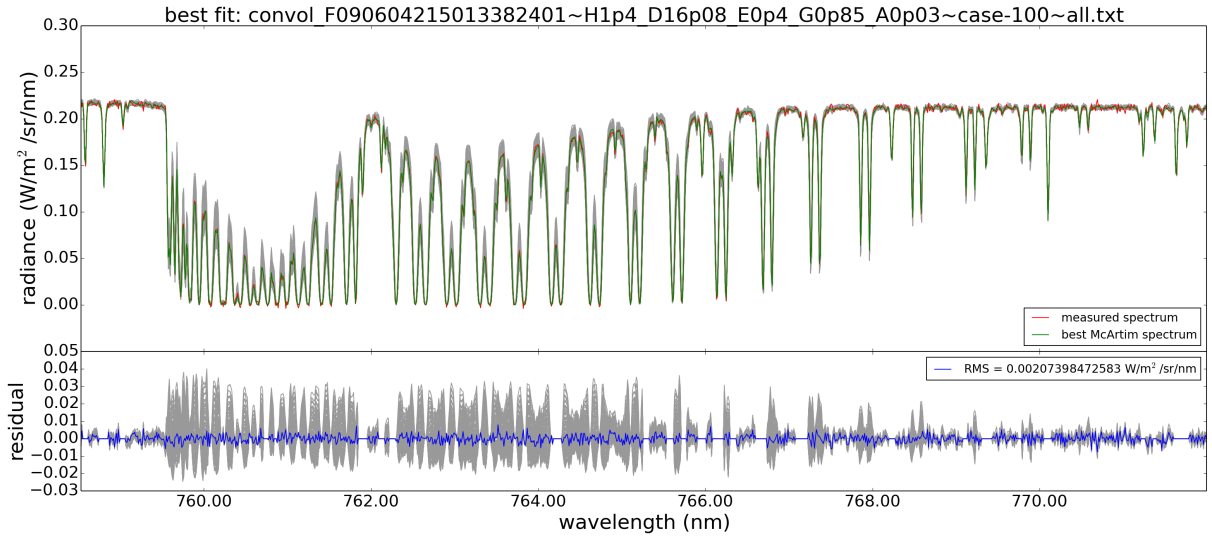

Figure 49: Spectrally aligned simulations and corresponding residuals (grey) to the reference spectrum (red). The residuals corresponding to simulated spectrum showing the smallest RMS from the fit, is drawn in blue.

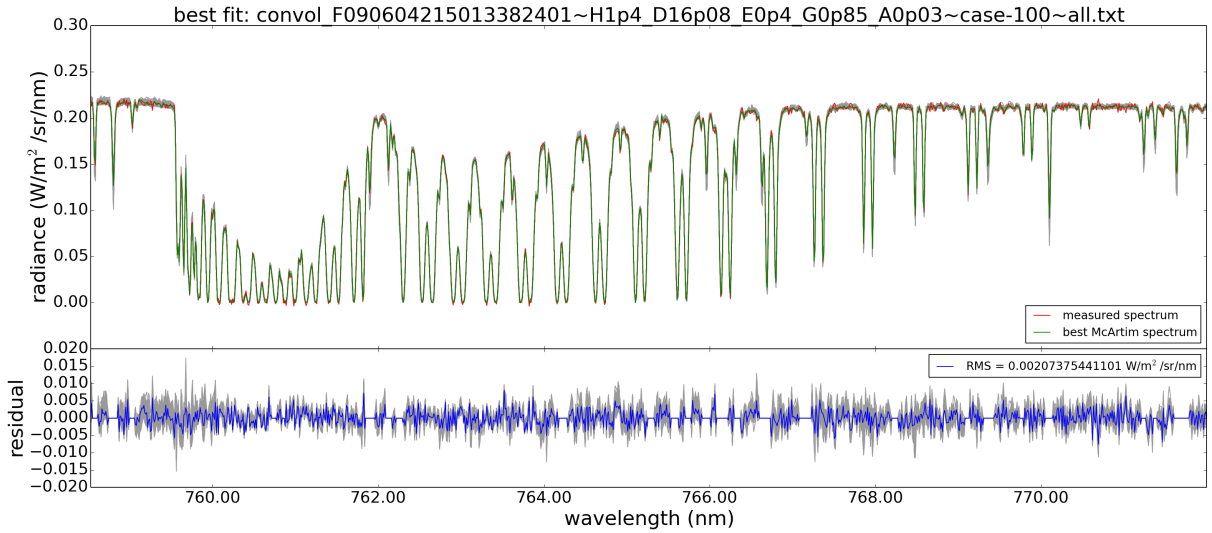

Figure 50: The fitted spectra (using the reference spectrum), also allowing a variation of the fit parameter B. The fit significantly reduces the residuals within the strong absorptions which can be seen from the comparison with figure 49.

The next step are the performance tests of the fit, including the wavelength alignment and the fit parameter B. The B parameter is assumed to have no wavelength dependency. The resulting spectra, together with the reference spectrum and the residuals, are shown in figure 50. The B parameter corrects the oxygen absorption in a way that the absorptions structures of the spectrum fit best to the reference spectrum - also in the deep absorptions. Consequently, the magnitude of the residuals is reduced. As before, simulation number 22 has the lowest RMS value.

Figure 51 includes the results of the fit parameter B and the corresponding RMS for each simulated spectrum. One can see that the lowest RMS values and thus the best agreement is found for cloud situations with B near 1. This is to be expected because the spatial extents of the model cloud apply to the simulations as well as the reference spectrum. It is interesting to note that the fit parameter B of case 22 is not exactly 1, but the deviation is negligibly small

(the resulting value is B≈1.00001). As the cloud scenario is not changed in comparison to the reference spectrum, the deviation is caused by the wavelength alignment and the added noise. Despite the small deviation of B from unity, the RMS of case 22 clearly has the lowest value. Other cases show B values different from unity. Higher B values indicate an underestimation of the simulated oxygen absorption, lower values an overestimation.

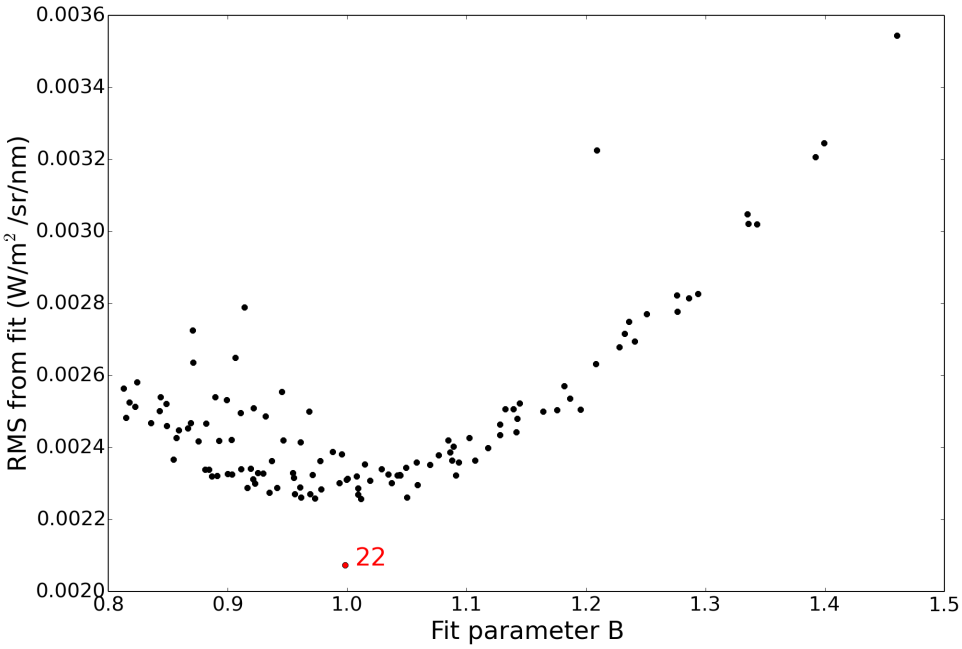

Figure 51: Dependence of RMS and the resulting parameter B from the fit for all simulations to the reference spectrum. In the case of a correct cloud model, the smallest values of RMS are found for B values near 1.

8.4.2 Influence of surface albedo variations

In all simulations, a surface albedo of 0.03 is assumed according to a water surface. This is a potential source for systematic errors. To test the influence of moderate albedo changes, the cloud characteristics of simulation 22 are kept constant but the surface albedo is varied from 0.01 to 0.1 in steps of 0.01. As before, the cloud optical depth is optimized for each simulation before the fit. The cloud characteristics are summarized in table 16. Comparing the surface albedo to the retrieved cloud optical depth τ_{cloud}, we see that τ_{cloud} is influenced by the albedo values as it slightly decreases with increasing albedo.

The convolved simulations are then fitted to the reference spectrum considering spectral alignment and a free parameter B. The fitted spectra are shown in grey in the upper plot in figure 52. The reference spectrum is shown in red and the fitted spectrum with the smallest RMS is highlighted in green. The residuals are shown in the lower plot, also in grey for all cases except the best-fitting case which is highlighted in blue. The magnitude of the residuals after fitting don't differ significantly between the different simulations.

An analysis of the resulting fit parameters and the fit quality shows that the B parameters are all in close proximity to B=1. This is to be expected because extinction within the cloud is dominant in this case and the cloud optical depth varies only slightly between the different simulations. Also as expected, the RMS value for a surface albedo of 0.03 is smallest. The comparison to the other situations shows no consistent dependency. The RMS shows only deviations in the 4^{th} decimal place of the values, which is very small. The RMS of the fitted spectra with a surface

Table 16: Properties of the model clouds for different values of the surface albedo after adapting the cloud optical depth τ_{cloud}.

albedo	h_{cloud}	e_{cloud}	τ_{cloud}
0.01	1.4	0.4	16.29
0.02	1.4	0.4	16.15
0.03	1.4	0.4	16.08
0.04	1.4	0.4	16.03
0.05	1.4	0.4	15.83
0.06	1.4	0.4	15.8
0.07	1.4	0.4	15.77
0.08	1.4	0.4	15.55
0.09	1.4	0.4	15.45
0.1	1.4	0.4	15.38

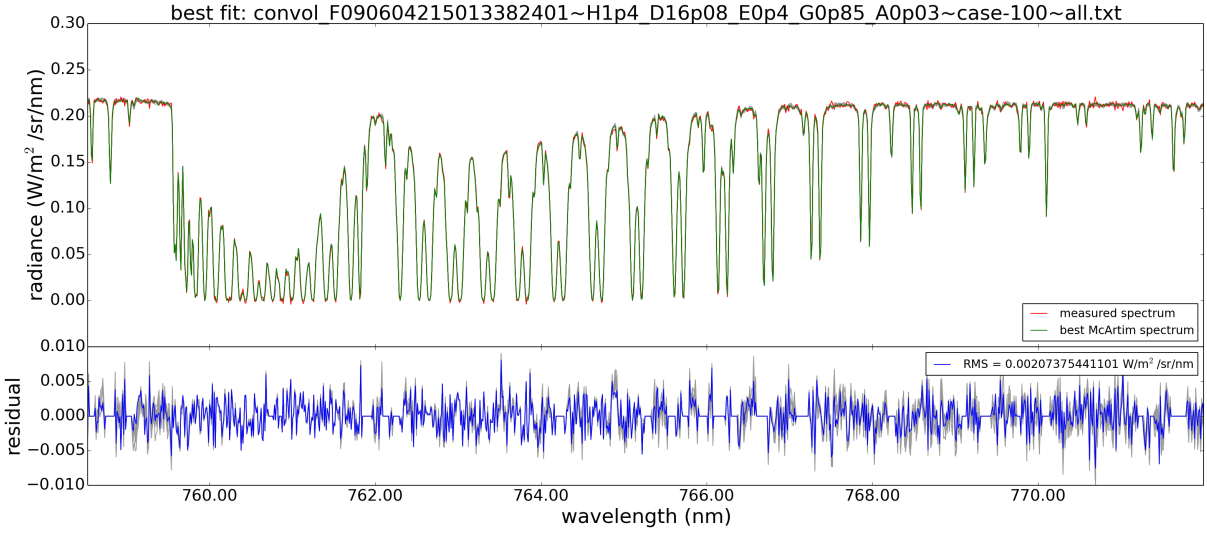

Figure 52: Fitted spectra and residuals for simulations using a constant cloud situation but different surface albedo and adapted cloud optical depths.

albedo other than 0.03 have approximately the same value as those in fig. 51 which are nearest to number 22 (RMS $\approx$ 0.00225). In principle, changes of the spatial cloud dimensions appear more important than changes of the surface albedo but it should be kept in mind that it is a source for systematic errors. Naturally, the influence of the surface albedo should become more important in case of clouds with very small optical depths or horizontal gaps within a cloud structure.

8.4.3 Accuracy of the applied ILSF

Here, the accuracy of the used instrument line shape function is tested and the real TANSO-FTS measurement is used instead of the created reference spectrum. First, the simulated spectra $I_{sim}(\lambda)$ are convolved with the TANSO-FTS instrument line shape function f_{ILSF} and are then subject to additional convolutions with a Gaussian function. Those convolutions following the first one, are applied to test whether the first convolution with the TANSO-FTS ILSF is sufficient or if further convolutions improve the agreement to the measured spectrum. Here, the discrete convolution operator is denoted by $*$.

$$I_{sim,0}(\lambda) = I_{sim}(\lambda) * f_{ILSF}(\lambda) \tag{56}$$

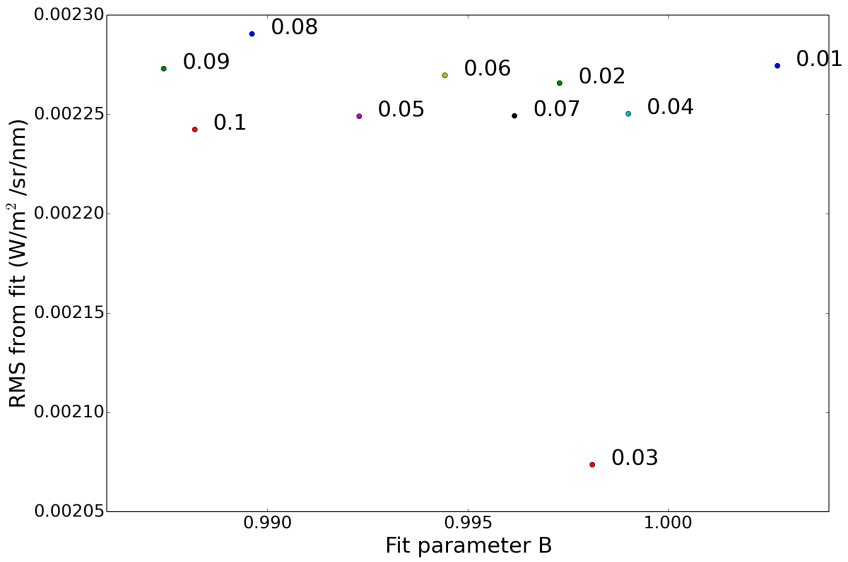

Figure 53: Fit RMS and B from fits considering spectral alignment and the variation of the B parameter. The cloud situation is constant, only the surface albedo is changed. The different albedos are given as annotation to the datapoints.

$$I_{sim,n}(\lambda) = I_{sim,n-1}(\lambda) * \frac{1}{2\pi\sigma} \exp^{-0.5\frac{\lambda^2}{\sigma^2}} \tag{57}$$

$$n \in \mathbb{N} \quad \text{in } [1, 9] \tag{58}$$

The Gauss convolution is characterized by the FWHM which corresponds to the standard deviation σ via the relation

$$\sigma = \frac{FWHM}{2\sqrt{2\ln(2)}} \tag{59}$$

In the case of N successive Gauss convolutions, the standard deviation σ calculates to (compare Bromiley [9]):

$$\sigma = \sqrt{\sum_{i=0}^{N} \sigma_i^2} \tag{60}$$

The Gaussian convolution provides an additional smoothing of the data and is applied one to nine times on each spectrum in order to describe convolutions of slowly increasing FWHM. After probing different values of the initial FWHM, it was set to 0.002 nm. The convolved spectra are then fitted to the TANSO-FTS measurement spectrum.

The performed fits use the standard approach including wavelength alignment and the B parameter. The RMS in relation to the resulting FWHM is shown in figure 54. To reduce the amount of data, the visual analysis is restricted to the 30 spectra which have the smallest RMS from fits without additional Gaussian smoothing. There is no consistent trend when examining the evolution of the RMS after additional Gaussian convolutions. For some cases, the RMS decreases before increasing again and for other cases, it steadily increases. The blue line (corresponding to the simulation with lowest RMS without Gaussian smoothing) shows a minimum of the fit RMS between a FWHM of 0.003 and 0.004 which is, however, not much pronounced.

A similar test is performed for the TANSO-FTS spectrum. The spectrum is modified by applying one to nine additional Gaussian convolutions (FWHM = 0.002 nm). Due to the implementation

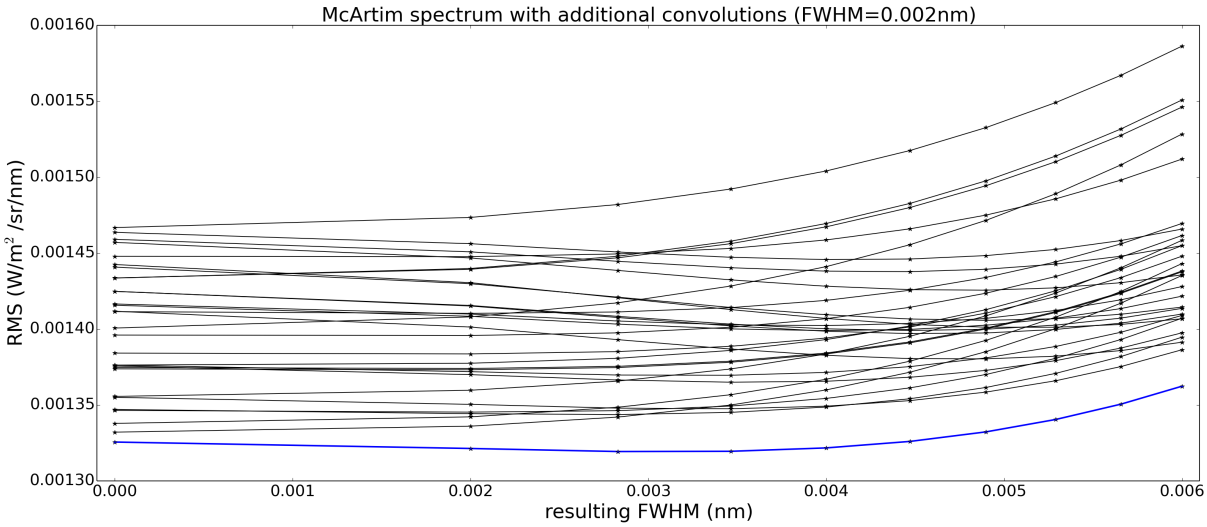

Figure 54: Resulting RMS for simulated spectra being modified through Gaussian convolutions with slowly increasing FWHM. Each simulation is indicated by a connecting line through the data points, the results for the spectrum with the lowest fit RMS (when no additional Gaussian smoothing is applied) are colored in blue.

of the fit algorithm, all simulated spectra are then fitted to each of the convolved TANSO-FTS spectra.

$$I_{sat,0}(\lambda) = I_{sat}(\lambda) \tag{61}$$

$$I_{sat,n}(\lambda) = I_{sat,n-1}(\lambda) * \frac{1}{2\pi\sigma} \exp^{-0.5\frac{\lambda^2}{\sigma^2}} \tag{62}$$

$$n \in \mathbb{N} \quad \text{in } [1, 9] \tag{63}$$

The result of the fits is shown in figure 55, using wavelength alignement and a parameter B. Like before, the visualization is restricted to the 30 best-fitting spectra. For some of the RMS evolutions, a minimum can be observed after a small number of additional Gaussian convolutions, but as in fig. 54, these are not very proncounced and present only for some of the shown curves. The blue line (again corresponding to the simulation with lowest RMS without Gaussian smoothing), for example, has the lowest RMS value when no further Gaussian smoothing is applied.

The results indicate that the TANSO-FTS spectra should not be subject to an additional Gaussian convolution before the fitting process is applied. Considering also the other shown evolutions of the RMS, it can be concluded that the applied convolution from the TANSO-FTS ILSF on the simulations is sufficient for this study.

8.4.4 Conclusions on test fits and sensitivity tests

In order to test the proper functioning of the fitting process and to probe the sensitivity of the fit results on selected influences, a number of tests has been performed. The functioning of the fitting routine is tested by fitting simulations with 120 different cloud scenarios to a synthetic (and thus well-characterized) radiance spectrum. As expected, a minimum value of the fit RMS and a parameter B near unity are found for the cloud scenario with the true spatial dimensions and τ_{cloud}. The wavelength alignment also works well. Varying the surface albedo within the simulations, shows that the spatial cloud dimensions are more important for the retrieved fit

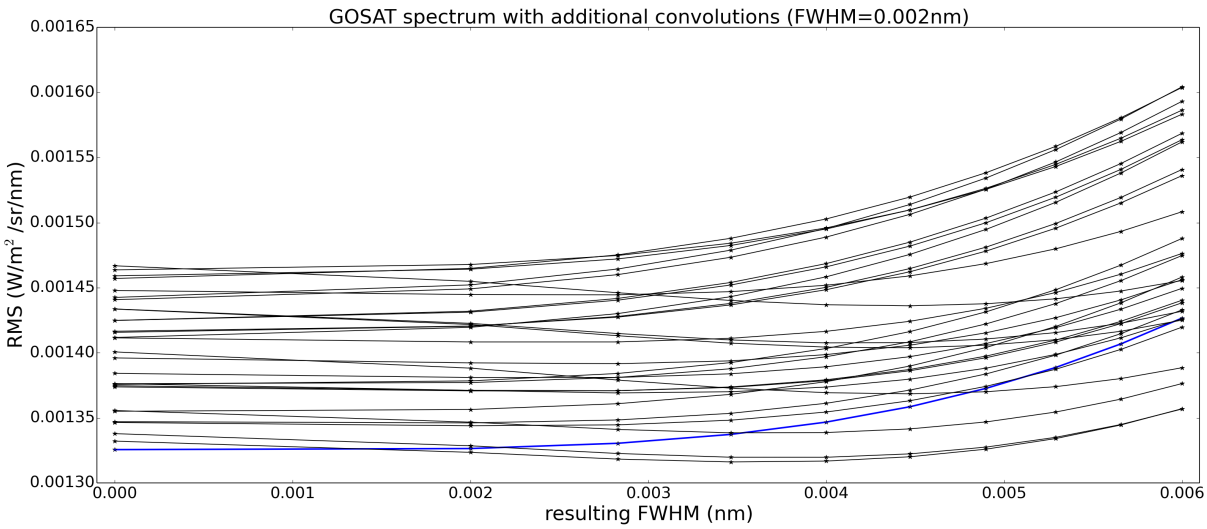

Figure 55: Resulting RMS for the simulated spectra (only the TANSO-FTS ILSF is applied beforehand) fitted to the TANSO-FTS spectrum which is smoothed through Gaussian convolutions with slowly increasing FWHM. The individual input spectra are identified by lines concatenating the data points, the results corresponding to the best-fitting spectrum (when no additional Gaussian smoothing is applied) are shown in blue.

result. This, however can change for other cloud scenarios than the one used (cloud between 0.8 and 1.4 km, $\tau_{cloud} = 16.08$), e.g. when the cloud optical depth is smaller.

Instead of a synthetic radiance spectrum, a real spectrum from TANSO-FTS measurements has been used to test the accuracy of the TANSO-FTS instrument slit function. The results show that a convolution of the simulations with the TANSO-FTS ILSF is sufficient and additional Gaussian convolutions on the simulations or the TANSO-FTS measurement do not systematically improve the fit result (RMS).

In section 10, further sensitivity tests will be shown which are also based on case study 1. These include the influence of the asymmetry parameter g on the fit results and the photon path length statistics, the effects of using different atmospheric profiles in the simulations and the importance of the wavelength sampling interval of the O_2 absorption cross sections for the simulation settings.

9 Cloud-free atmospheric scenarios

Before the retrieval presented in section 8 is applied to the cloudy sky cases studies, two cases with cloud- and almost aerosol-free atmospheres are considered, more explicitly a measurement over a water surface and one over desert. They were chosen in order to analyze at least one measurement over dark and one over bright surface. They are random choices from a manual search for clear sky measurements and not based on the previously presented collocation data. For the radiance spectra and ancillary information, the TANSO-FTS L1B dataset (version V201.202) is used which also includes a master quality flag which is set to good for both cases.

9.1 Measurement over dark surface

The measurement over water was performed on 8 July, 2014 over the Red Sea. The information on the cloud situation originates from the corresponding TANSO-CAI Cloud Property dataset, which defines all pixels within the TANSO-FTS FOV as clear without any heavy aerosol load. An indication for the possible contamination by high cirrus clouds is given for only about 14% of the pixels inside the TANSO-FTS FOV. Here, one should keep in mind that in general, the detection of cirrus clouds by TANSO-CAI is challenging because it does not include a thermal infrared channel [37].

For clear sky cases, the surface albedo is a critical parameter for the amount of the measured radiance. Here, the surface albedo is derived by comparing radiances within the wavelength continuum, which leads to an albedo $a_s = 0.078$ for the ocean measurement. This value is unusually high for water surfaces in this spectral region. Possible causes are remaining cloud structures or increased surface reflection due to strong wind (e.g. waves).

Table 17: Geometrical characteristics of the clear sky measurement over water.

longitude	36.02°	surface of FOV	sea
latitude	23.34°		
timestamp of measurement	2014-07-08 10:28:28		
solar zenith angle	10.96°	satellite zenith angle	26.72°
solar azimuth angle	267.70°	satellite azimuth angle	72.05°

The comparison of the measurement to the simulation after performing a wavelength alignment is shown in figure 56. In the upper plot, the measured TANSO-FTS spectrum (red) and the McArtim simulated radiances (green) are shown, the plot directly below this comparison includes the corresponding residual. The spectra appear very similar to each other. In order to allow a variation of the O_2 absorption of the simulation, a fit including the necessary wavelength alignment and also a parameter B_{fit} (B parameter, described in section 8.3) is performed for the wavelengths between 759.0 and 772.0 nm, leading to the fitted spectrum depicted in the lower plot of fig. 56. The residual between measurement and simulation is only slightly reduced to the comparison only applying the wavelength alignment. The resulting parameter $B_{fit} \approx 1.058$ suggests an underestimation of the O_2 absorption during the simulation by about 6%. Before this value is discussed, the measurement over desert is considered.

9.2 Measurement over bright surface

The second scenario is based on a measurement from 3 June, 2010 over a desert landscape on the African continent. The TANSO-CAI cloud property indicates that the entire FOV of the corresponding TANSO-FTS radiance measurement is free of clouds, thin cirrus and heavy aerosol load. The geometric properties are summarized in table 18. For the radiance simulations, the surface altitude is adapted to 400 m (the TANSO-CAI dataset indicates a ground elevation of $\approx$ 323 m but for the simulation, the chosen surface must match the altitude grid of the model

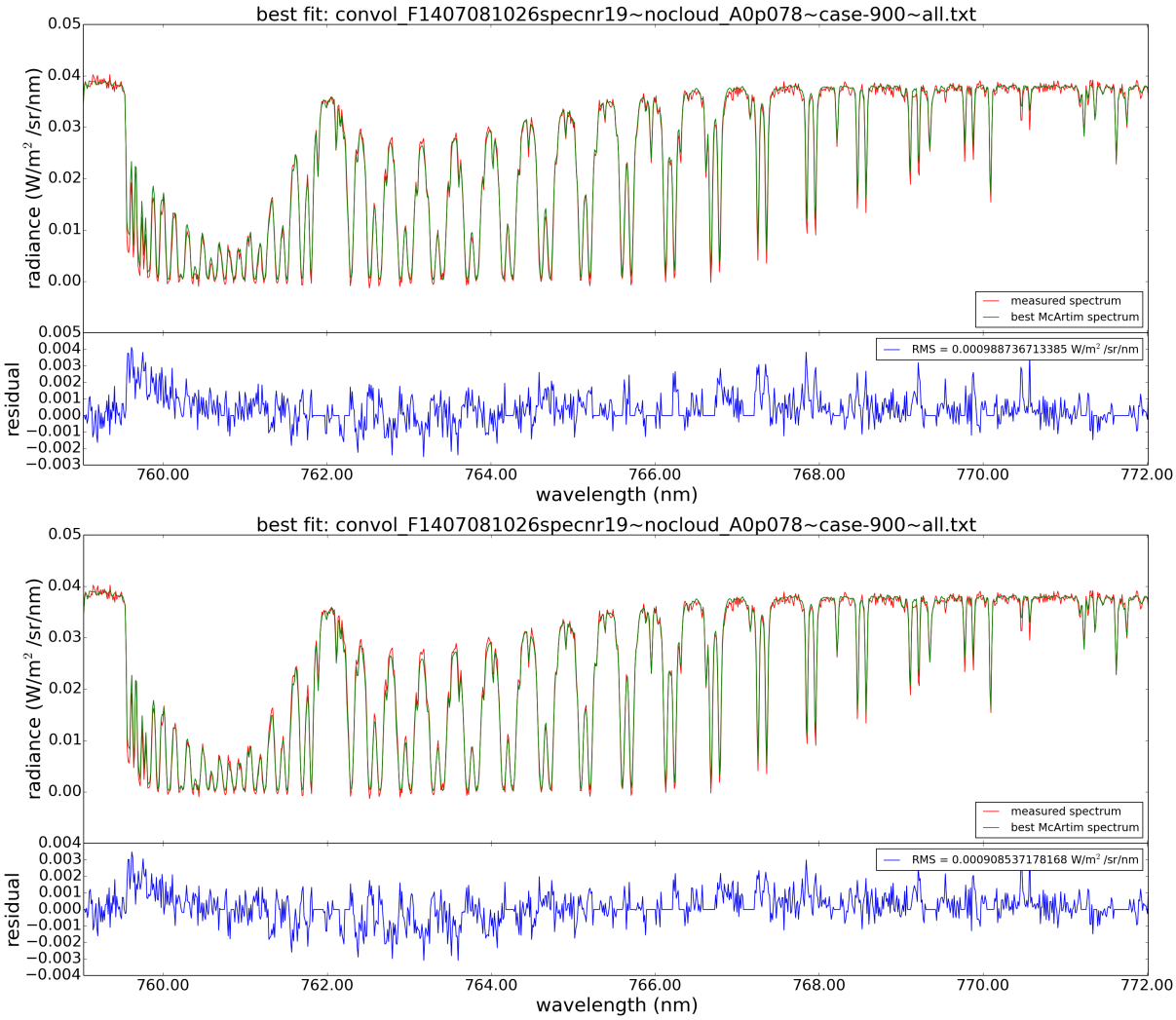

Figure 56: Top: Wavelength-aligned comparison of the simulation (green) and the measurement (red) with the residual directly below (blue) in the case of a clear-sky scenario over dark surface. Bottom: Fit of the simulation to the measurement also considering a parameter B_{fit} (same color scheme as above).

atmosphere) and the surface albedo is retrieved from the Filled Land Surface Albedo Product (Terra/MODIS) [61], which allows to define a starting value for the simulated ground albedo. The derived value for the white sky albedo is 0.605 and 0.594 for the black sky albedo[11]. From the comparison of measured and simulated radiances in the continuum wavelength region, an albedo of 0.523 is found, which defines a slightly less reflective surface.

Table 18: Geometrical characteristics of the clear sky measurement over land (desert).

longitude	28.52°	surface of FOV	land
latitude	21.68°		
timestamp of measurement	2010-06-03 10:59:22		
solar zenith angle	12.83°	satellite zenith angle	14.93°
solar azimuth angle	275.50°	satellite azimuth angle	93.98°

The comparison between the wavelength-corrected simulation and the measurement is shown in fig. 57 (top), the fit allowing a variation of the O_2 absorption in fig. 57 (bottom). The derived

[11]The white sky albedo is based on the assumption of diffuse, the black sky albedo on directional light [81].

parameter $B_{fit} \approx 0.934$ indicates an overestimation of the O_2 absorption by 6-7%. The fit leads to a better agreement of the radiances in wavelength regions of intermediate absorption but higher disagreement at the first strong absorption structure at about 759.95 nm. The RMS of the comparison is reduced from 0.0034 to 0.0024 by allowing a B_{fit} different from unity.

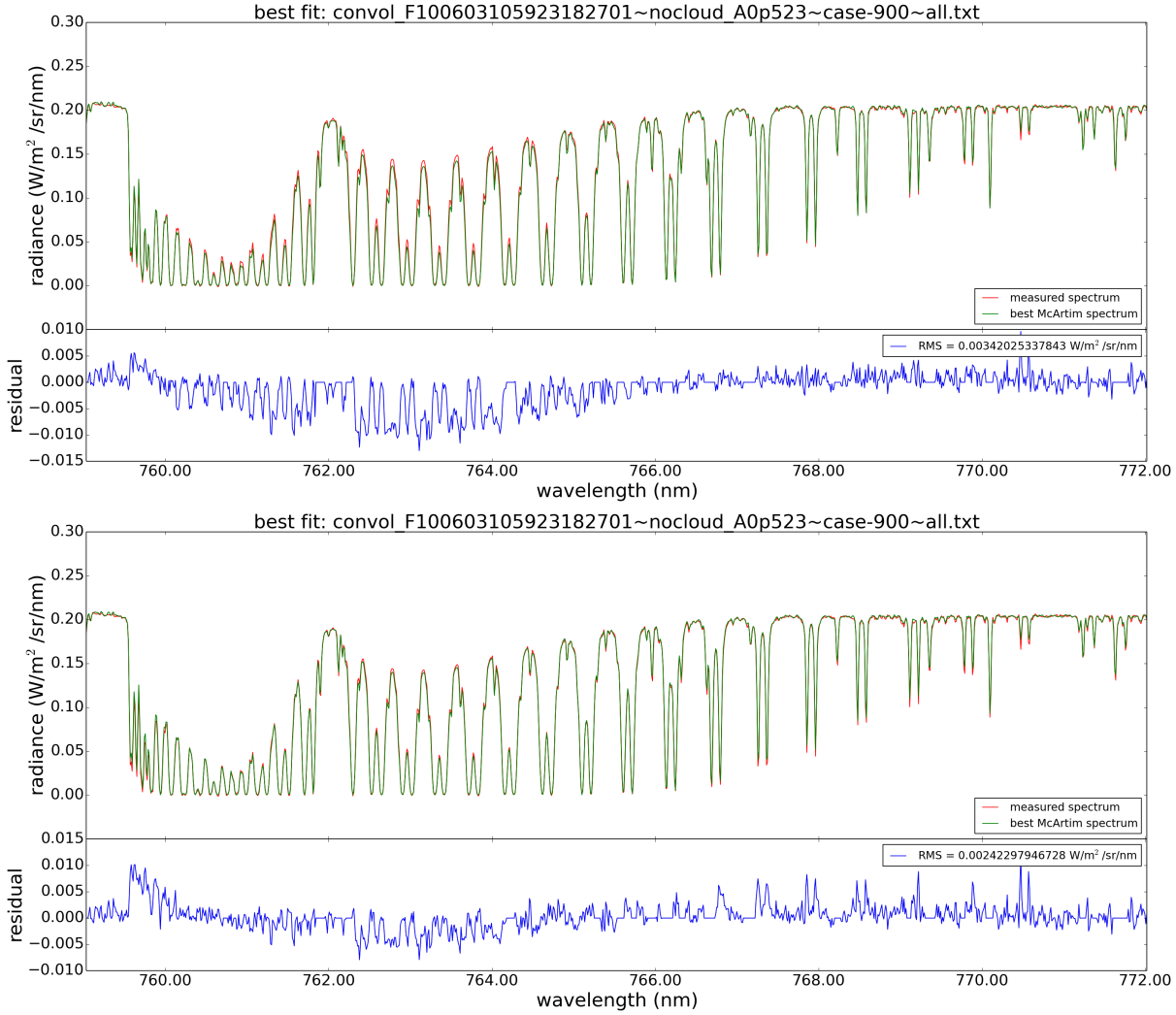

Figure 57: Top: Wavelength-aligned comparison of the simulation (green) and the measurement (red) with the residual directly below (blue) in the case of a clear-sky scenario over bright surface. Bottom: Fit of the simulation to the measurement also considering a parameter B_{fit} (same color scheme as above).

9.3 Conclusions

If the measurement as well as the simulation were perfect, B_{fit} should be (close to) unity and the RMS would consist of noise only. Although the residuals are small compared to the radiance values, we find deviations from $B_{fit} = 1$. In the case of the measurement over water (dark surface), an underestimation of the O_2 absorption is identified. The contrary results for the measurement over desert which is a highly reflective surface. There is thus a disagreement not only on the expected value of B_{fit}, but also on the direction of deviations from $B_{fit} = 1$.

The reasons for this discrepancy are currently not understood. As can be seen in figure 58 (water surface), many radiance values are negative, possibly due to an overcorrection of spectral straylight during the measurement or of an electronic offset of the detectors and/or electronics. This effect is smaller for bright surfaces, such as the measurement over desert (see fig. 59) or in presence of clouds. The fit results of the case study of the clear sky measurement over desert show an overestimation of the O_2 absorption by approx. 6%. As will be shown later, a similar overestimation is found for the considered measurements with one-layer cloud systems.

Another possibility for the discrepancies are the radiative transfer simulations which are subject to possible errors. Although no indications have been found in other studies, the used O_2 cross sections might have systematic errors. Concerning the effect of Raman scattering, it is interesting to note that it is less important over bright surfaces (desert, clouds) than dark ones like ocean. In general, the magnitude of the effect is rather small (see e.g. Vasilkov et al. [101]) and thus cannot explain the observed deviation of about 6% from $B_{fit} = 1$.

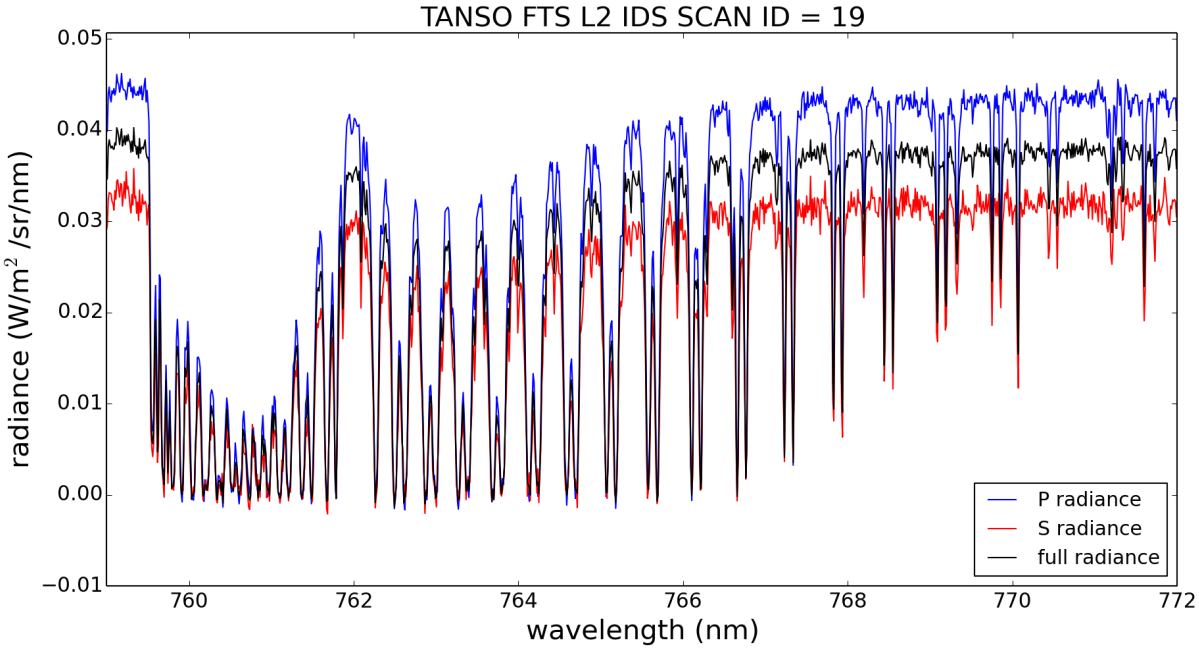

Figure 58: The TANSO-FTS radiances between 759 and 772 nm over water. The polarized radiances (P in blue and S in red) are shown together with the approximation of the full radiance spectrum (black).

In figure 58 (water) and 59 (desert), the measured polarized radiances (P- and S-radiance) are shown together with the approximated full radiance (see section 5.3.2 for details). For the water case, the radiances for the different polarizations are quite different compared to the magnitude of measured radiance, while they appear nearly identical for the desert case. This is to be expected because of the higher sensitivity of dark-surface measurements to Rayleigh scattering.

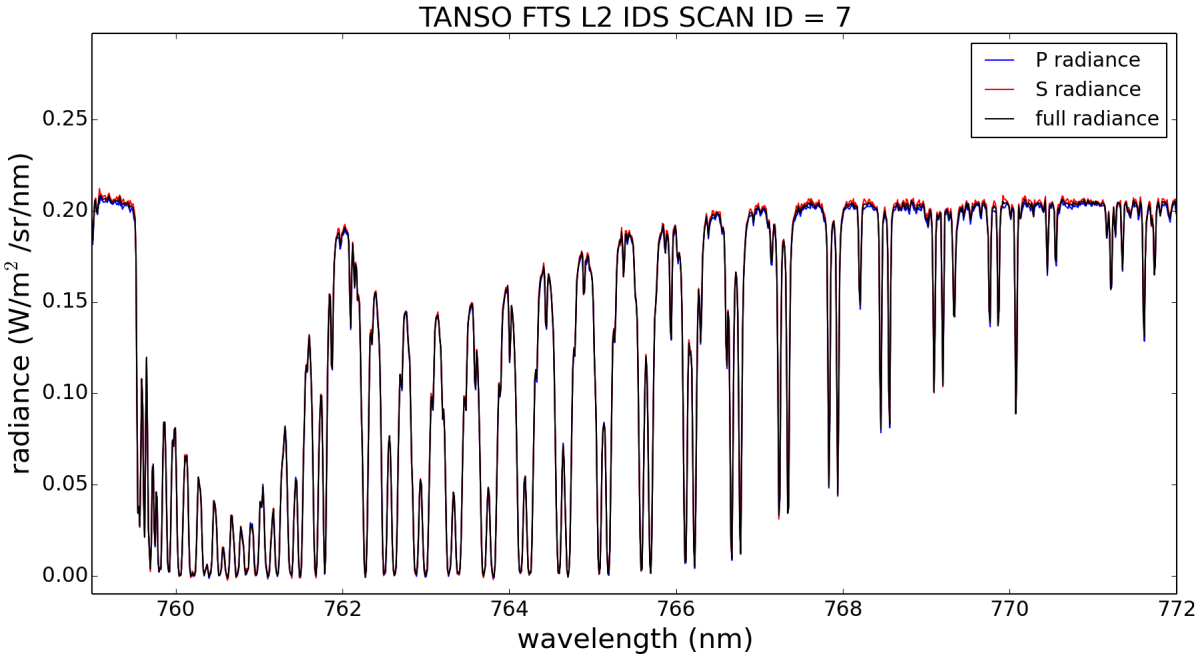

Figure 59: The TANSO-FTS radiances between 759 and 772 nm over desert. The polarized radiances (P in blue and S in red) are shown together with the approximation of the full radiance spectrum (black).

10 Case study 1

The complete fit algorithm is first applied to the case study investigated in section 8. Here, the TANSO-FTS radiances (instead of the artificial reference spectrum) are used as atmospheric measurement spectrum. According to the collocated measurements from CALIOP and CPR, the cloud top is situated at an altitude of approximately 1 km. The vertical extent of the cloud appears small, namely between 0.2 and 0.4 km. The cloud situation thus focuses on a low and vertically thin cloud. Mean cloud optical depths for the TANSO-FTS FOV region are available from measurements of TANSO-CAI (τ_{CAI}=14.65) and CloudSat CPR (τ_{CPR}=15.4), whose values are very similar.

10.1 Simulated cloud scenarios

Radiative transfer simulations with McArtim are performed for 120 different cloud scenarios and an assumed surface albedo a_s=0.03. The model clouds are designed to be homogeneous single-layer clouds and differ only in top height and vertical extent. The implemented cloud top heights range from $0.2 \leq h_{cloud} \leq 3.0$ km and all vertical extents e_{cloud} possible on a 0.2 km vertical grid are used. For each case, the cloud optical depth τ_{cloud} is optimized to find the best agreement in the continuum region, which leads to $\tau_{cloud} \approx 16.1$. Because of the low cloud top height, a water-phase cloud is assumed whose optical characteristics are described by a Henyey-Greenstein phase function with an asymmetry parameter of 0.85. A complete overview of all simulated cloud parameters is shown in table 21 (Appendix, section 10.7), together with the identification numbers for each cloud scenario.

10.2 Spectrum fit

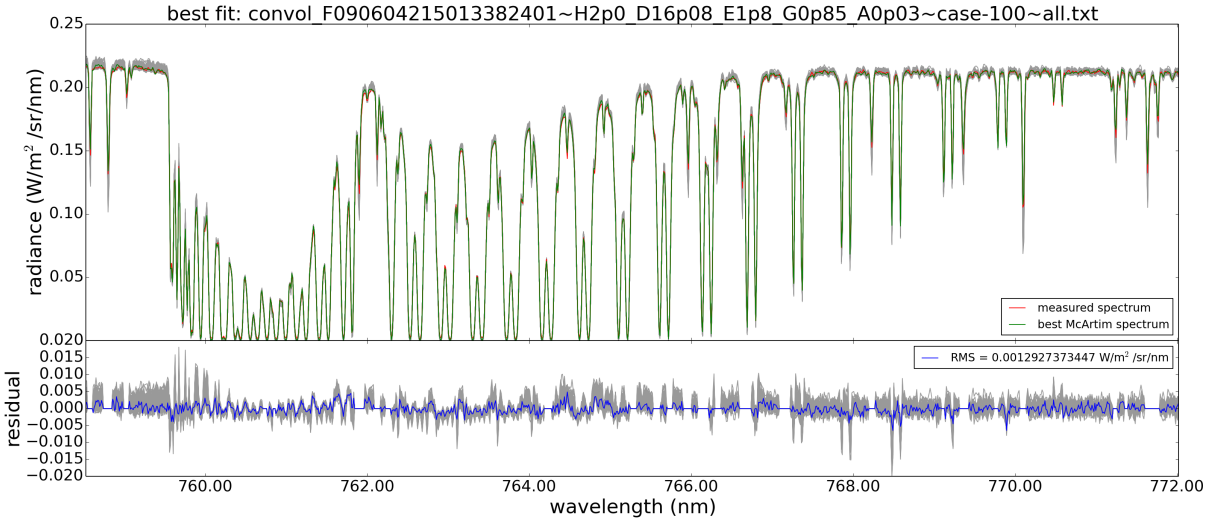

Figure 60: Top: Fitted spectra for all cloud scenarios (grey), the TANSO-FTS measurement spectrum (red) and the fitted spectrum with the lowest RMS (green). Bottom: The corresponding residuals (grey) and those with the smallest RMS (blue).

All simulations are fitted against the TANSO-FTS measurement using the entire wavelength region of the O_2 A-band. The fit parameters include the wavelength alignment parameters and B_{fit}, describing a correction of the O_2 absorption (see section 8 for details). The fitted spectra and their residuals to the measurement are visualized in fig. 60. The upper plot shows all fit results (grey) and the fit result with the lowest RMS (green), the TANSO-FTS spectrum is plotted for comparison (red). The residuals are shown in the lower plot (whole dataset in grey, the case resulting in the lowest RMS in blue). The fitting process adapts all simulations very

well to the measurement. To select the best matching simulation to the measurement, the fit RMS and/or the value of the parameter B_{fit} are considered.

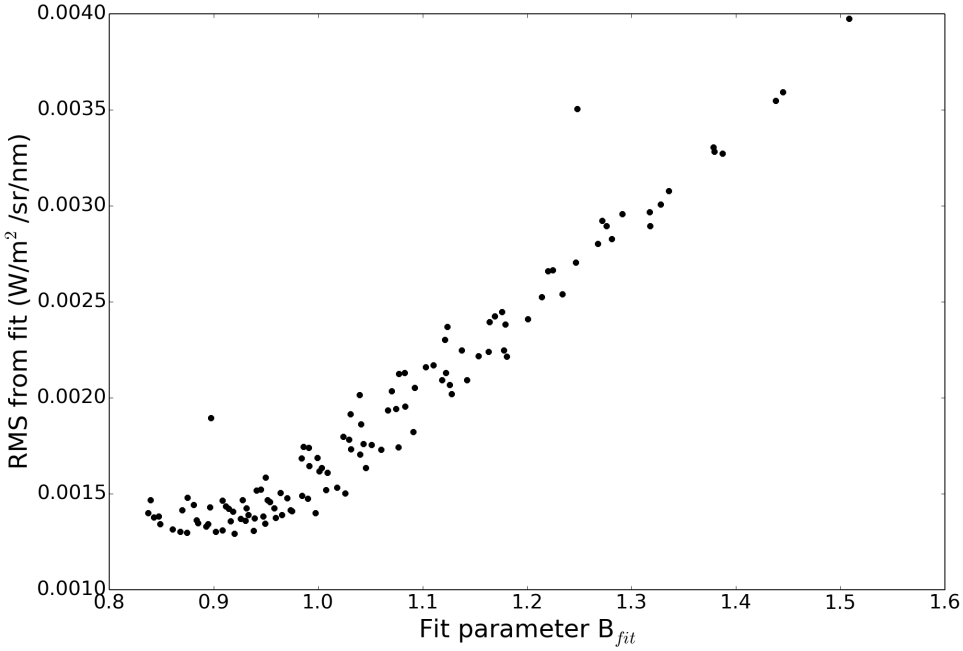

Figure 61: The resulting fit RMS in dependence of the fit parameter B_{fit} for all simulated cloud scenarios (g=0.85).

The dependence of the RMS on the derived fit parameter B_{fit} is shown in figure 61. In the hypothetical case that the cloud model and atmosphere perfectly match the reality, one would expect the lowest RMS for a case at $B_{fit} \approx 1$, similar to the fit against the synthetic reference spectrum (see section 8.4). Here, minimum values of the RMS are found at B_{fit} around 0.9 with RMS ≈ 0.0013 which indicates an overestimation of the simulated O_2 absorption by about 10%.

Figure 62 (top) shows a boxplot representation of the photon path length (PPL) distributions at 772.145 nm (weak O_2 absorption) in dependence of the fit parameter B_{fit}. All photon paths below an altitude of 10 km are considered. The median of the photon path length distribution is indicated in blue, while the 25^{th} and 75^{th} percentiles define the boundaries of the red box and the end bars are given by the 9^{th} and 91^{th} percentiles of the distribution. The lower plot of the same figure shows the median values (blue) together with the average photon path lengths $\overline{PPL}$ (green). Together with B_{fit}, the medians form a clear (almost linear) dependence with higher values for $B_{fit} < 1$ and lower values for $B_{fit} > 1$. $\overline{PPL}$ also anticorrelates with B_{fit} but there is a spread of $\overline{PPL}$ values for the same B_{fit}. Here, it follows that B_{fit} is not solely determined by the average photon path lengths but the distribution of photon path lengths within the O_2 A-band.

An alternative representation where a weight of the distances with relation to the O_2 absorption below 50 km is performed, can be found in Appendix (section 10.7), fig. 72.

It is noteworthy that only photon path length distributions with small differences between the 25^{th} and 75^{th} percentiles appear at the highest and smallest B_{fit} values. The cloud scenarios with the smallest B_{fit} parameters (and largest medians of the path length distributions) are those with the numbers 0, 2, 5, 9, 14 and 20 (above the magenta line on the left side in fig. 62). These are clouds with top altitudes until 1.2 km and they have in common that the vertical extent is equal to the cloud top height. All clouds extend down to the ground. From the fit

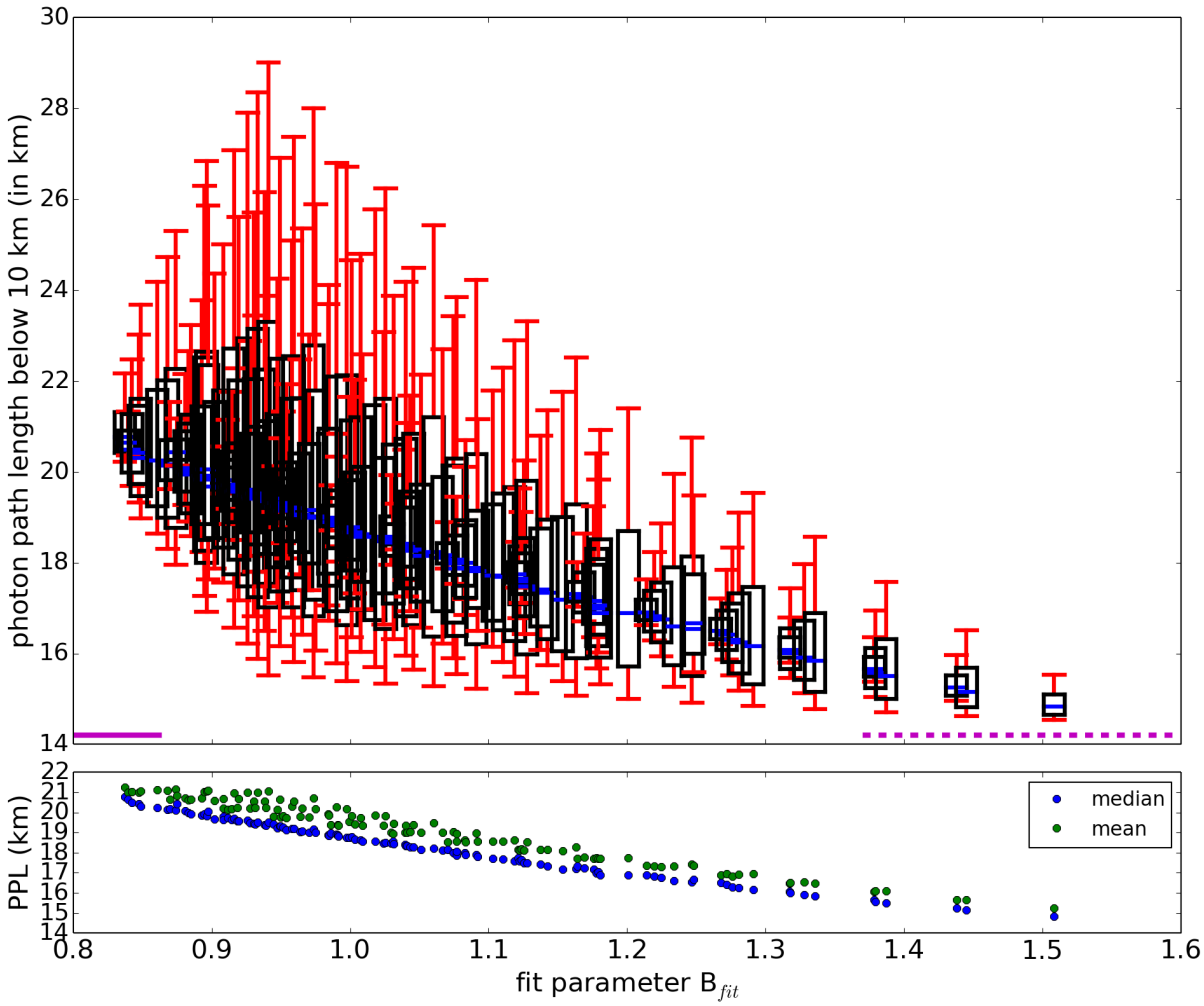

Figure 62: Top: Photon path length distributions for all assumed cloud scenarios (g=0.85) in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

results, values of $B_{fit} < 0.862$ are obtained, thus the simulated O_2 absorption is significantly overestimated. The widths of the distributions are rather small because of the low cloud top boundaries and thus the small vertical cloud extents.

The right side of the plot shows scenarios with low medians and large B_{fit} parameters. The corresponding case numbers are 105, 106, 91, 107, 92 and 78 (above the dotted magenta line on the right side in fig. 62). Here, the clouds are situated above 2.4 km and the cloud vertical extent is small with values from 0.2 and 0.6 km, depending on the cloud top height. Because of the high extinction coefficients of the clouds and the small vertical extents, most photons cannot enter the lower atmospheric layers. Due to the short photon path lengths, the absorption is significantly underestimated which is expressed by values of $B_{fit} > 1.37$.

10.3 Fit results

As mentioned before, the minimum of the RMS is expected for cloud scenarios with B_{fit} near unity. In this case, both the fit RMS and B_{fit} would be taken into account to select the best matching cloud situation. Here, the minimum RMS is found for cloud scenarios with B_{fit} smaller than unity. In principle, two interpretation strategies can be followed, either solely based on the lowest RMS or on the cloud scenarios with B_{fit} near unity from which the case with the lowest RMS is considered.

First, the scenarios with the lowest fit RMS are considered. The outcome of case study 1 shows minimum values of RMS ≈ 0.0013 W/m^2/sr/nm for cloud scenarios resulting in $B_{fit} \approx 0.9$. It follows that the simulation overestimates the O_2 absorption by about 10%. The cloud scenarios with the lowest RMS are 53, 35, 27, 34, 76 and 43 (characteristics and fit results are summarized in table 19). They show very different cloud top heights but have in common that the cloud bases are very low (at ground altitude or 200 m above). All scenarios show an overestimation of the simulated O_2 absorption and similar RMS (considering a minimum error of ΔRMS = $\pm$ 10^{-4} W/m^2/sr/nm, see Appendix (section 10.7.4) for details).

Table 19: Cloud parameters for the cloud scenarios with the lowest RMS values resulting from the fit together with the fit results RMS and B_{fit}. The fit RMS increases from the top of the table to the bottom.

num	h_{cloud}	e_{cloud}	τ_{cloud}	B$_{fit}$	RMS (W/m^2/sr/nm)
53	2.0	1.8	16.08	0.92	0.00129
35	1.6	1.6	16.1	0.874	0.0013
27	1.4	1.4	16.13	0.868	0.0013
34	1.6	1.4	16.06	0.902	0.0013
76	2.4	2.2	16.1	0.938	0.00131
43	1.8	1.6	16.05	0.909	0.00131

The spatial dimensions of the resulting cloud scenarios differ systematically from the collocated lidar and radar observations. The radar measurement detects a cloud bottom in higher altitudes and the lidar instrument observes a cloud top at about 1 km, while the above mentioned scenarios show higher cloud tops. The simulated O_2 absorption is overestimated by at least 5%, which is a possible reason for this discrepancy. Interestingly, a similar overestimation has been found for the clear sky measurement over a bright surface (desert, see section 9.2). As will be shown later, case study 2 (section 11) leads to similar results. So far, no explanation for this overestimation has been found, however, some sensitivity studies will be presented at a later point.

The corresponding photon path length distributions are shown in boxplot representation in fig. 64 together with the calculated mean and median values of the distributions (blue markers for g=0.85). The cloud scenarios resulting in the smallest fit RMS also show similarities of the photon path length distributions. The median values (19.3 km $< (P(PPL) \leq 0.5) <$ 20.2 km) are systematically lower than the mean photon path lengths (20.5 km $< \overline{PPL} <$ 21.2 km). This indicates that the distributions are centered towards shorter path lengths. The median values are reduced compared to the geometric clear-sky trajectory through the atmosphere. Half of the photon path lengths lie between the 25^{th} and 75^{th} percentiles of the distributions. The difference between the 25^{th} and 75^{th} percentiles is between 2.99 km and 4.71 km, thus a spread of only about 1.7 km is observed between the different cloud situations. The cloud scenario dependent difference between the 9^{th} and 91^{th} percentiles, thus comprising the path lengths of 82% of the photons, is only 3.3 km.

The second interpretation strategy is based on the assumption that the model and measurement are accurate and the final result must yield a B$_{fit}$ near unity. Considering the simulation noise error of $\Delta B = \pm\ 2 \cdot 10^{-3}$ (see section 10.7.4), the simulations with cloud scenarios 31, 74 and 117 fit best to the measurement. The cloud scenarios of this selection, together with the fit results, are summarized in table 20, including the cloud properties, B$_{fit}$ and the fit RMS. The spatial dimensions of the clouds show significant differences but again exhibit all higher cloud top heights as observed by CALIOP.

The scenario with the fit parameter B$_{fit}$ closest to 1 is number 31, describing a cloud between 0.8 and 1.6 km. It also compares best to the collocated profiles from the selection. However, also this interpretation strategy shows an ambiguity of possible cloud scenarios.

Table 20: Cloud parameters and fit results for the cloud scenarios with resulting B_{fit} close to unity and low RMS. The fit RMS increases from the top of the table to the bottom.

num	h_{cloud}	e_{cloud}	τ_{cloud}	B_{fit}	RMS ($W/m^2/sr/nm$)
117	3.0	2.6	16.08	0.998	0.0014
74	2.4	1.8	16.17	1.001	0.0016
31	1.6	0.8	16.13	0.999	0.0017

The photon path length distributions (fig. 64, blue markers) show averages between 19.5 and 20.2 km and slightly lower medians of appr. 18.8 km. The spread of the differences between the 25^{th} and 75^{th} percentiles as well as for the differences between the 9^{th} and 91^{th} percentiles are more than twice as large than for the cloud scenarios with the smallest RMS. While the medians of the distributions and $\overline{PPL}$ do not differ very much between the results of both interpretation strategies, the variations within the resulting groups of cloud scenarios is much larger for those scenarios with B_{fit} near unity. Moreover, and probably more important, the RMS values for these cloud scenarios show systematically higher values than those where only the smallest RMS is considered (see above).

10.4 Fit results in sub regions of the O_2 A-band

Fits were also performed using smaller wavelength regions of the O_2 A-band, as described in section 8.3. The selected wavelength ranges include regions of very strong O_2 absorption (subband 1), intermediate O_2 absorption (subband 2) and weak O_2 absorption (subband 3). In general, the deepest absorption lines are most sensitive to short photon path lengths and weak absorption structures are most sensitive to long photon path lengths. Similar to the fit of the entire O_2 A-band, two interpretation strategies can be followed by setting the focus either on the lowest fit residual (RMS) or the proximity of B_{fit} to unity.

The results of the different fitting regions are compared in fig. 70 (Appendix, section 10.7). The left plots show the 10 cloud scenarios leading to the smallest fit RMS while the right plots concentrate on the proximity of B_{fit}=1. The fitting region is the entire O_2 A-band in the upper row, followed by a wavelength region of strong, intermediate and weak O_2 absorption.

Regarding the fitting results of low RMS, the selection of cloud scenarios does not change much between using the entire wavelength range or strong and intermediate absorption regions. However, the values of RMS and B_{fit} show slight variations. Especially the magnitude of the RMS is lower for smaller fitting regions. This is not surprising, because the fit parameters must only adapt to a subregime of absorption depths. The different fitting regions (except very weak absorptions) also show smallest RMS for clouds extending down to the ground or ending shortly above. This shows that the fits in subregions of strong and intermediate absorption regimes of the O_2 A-band are in agreement with those performed on the complete wavelength range. The fit in the weakly absorbing subband 3 leads to very different results.

Concentrating on the scenarios leading to B_{fit} near unity, the following observations can be made: When the fit is performed in regions of strong O_2 absorption (subband 1), scenario 31 (cloud between 0.8 and 1.6 km) is still closest to $B_{fit} = 1$, although some cloud scenarios with lower RMS are found within an approximate 0.5% region of B_{fit}=1. For intermediate absorptions (subband 2), the scenarios with B_{fit} values nearest to 1, are 87 (cloud from 0.6-2.6 km), 40 (cloud from 0.8-1.8 km) and scenario 31. Scenarios 87 and 40 both show a slight underestimation of the O_2 absorption when the fit is performed within the entire O_2 A-band. Nevertheless, the result of this subband is similar to the entire band, because cloud scenarios 87 and 40 resemble 74 and 31 (present in the fit results of the entire band), both extending the cloud by 0.2 km in the upward direction. The fit results using only weakly absorbing wavelengths (subband 3) shows

very different findings. The scenario with B_{fit} closest to unity is number 115 (cloud from 0.8-3.0 km). As for the results using only the lowest RMS, there is a certain consistency of results from fits in all, strong and intermediate absorption regimes of the O_2 A-band.

10.5 Influence of a different asymmetry parameter g

10.5.1 Fit results

Under the assumption of a water cloud, an asymmetry parameter of g=0.85 for the Henyey-Greenstein phase function has been used so far. In order to probe the sensitivity for a slightly different phase function, the asymmetry parameter is reduced to a value of g=0.8, resulting in moderately less forward scattering. As before, the cloud optical depth is adapted for each cloud scenario to match the continuum region. While the retrieved cloud optical depths were close to 16.1 for g=0.85, they are reduced to values around 12.2 for g=0.8 (see table 22, Appendix section 10.7).

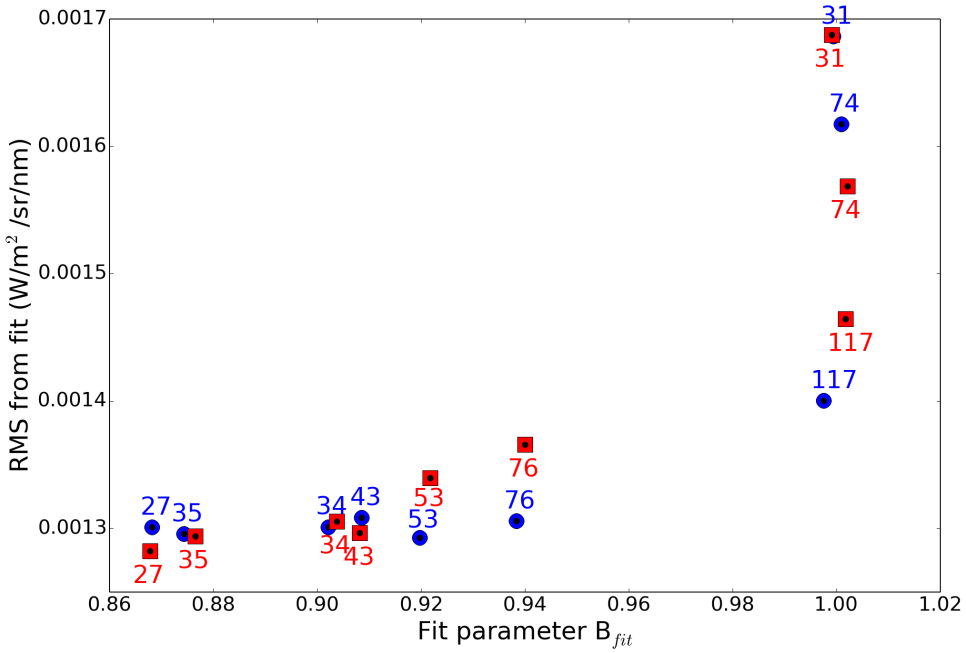

Figure 63: The resulting fit RMS and B_{fit} for selected cloud scenarios. Simulation results from using an asymmetry parameter g=0.85 are shown as blue circles, those using g=0.8 as red squares.

For the comparison, a selection of the performed simulation results is presented, which includes the cloud scenarios with the lowest RMS and those with B_{fit} closest to unity. They are shown in fig. 63, based on either g=0.85 (blue circles) or g=0.8 (red squares). The values of RMS and B_{fit} between both g differ only slightly. The boxplot representation of the corresponding photon path length distributions at 772.145 nm (weak O_2 absorption) in dependence of B_{fit} is shown in fig. 64 (top), together with the median and mean values (bottom). One can observe that also the percentile values and $\overline{PPL}$ of the photon path length distributions for the same spatial dimensions of the cloud have a high resemblance despite the different τ_{cloud} and g. The O_2 weighted distributions are shown in fig. 74 (see Appendix, section 10.7).

As mentioned before, the percentile representation of the photon path length distributions for the cloud scenarios with the lowest RMS are relatively similar while those with B_{fit} near unity are more different from each other. Taking all shown cloud scenarios into account, the median values of the distributions lie between 18.7 km and 20.2 km and the $\overline{PPL}$ are approximately 19.5-21.2 km. Despite the ambiguity of different results, some conclusions on these parameters

can be drawn from both interpretation strategies.

When considering the entire collection of simulations (instead of the above presented selection), the cloud scenarios with B_{fit} near unity are the same for both g. The lowest fit RMS for g=0.8 are obtained for the scenarios 27, 20, 35, 43 and 44 where the cloud top heights are equal or slightly lower than for the corresponding selection of g=0.85.

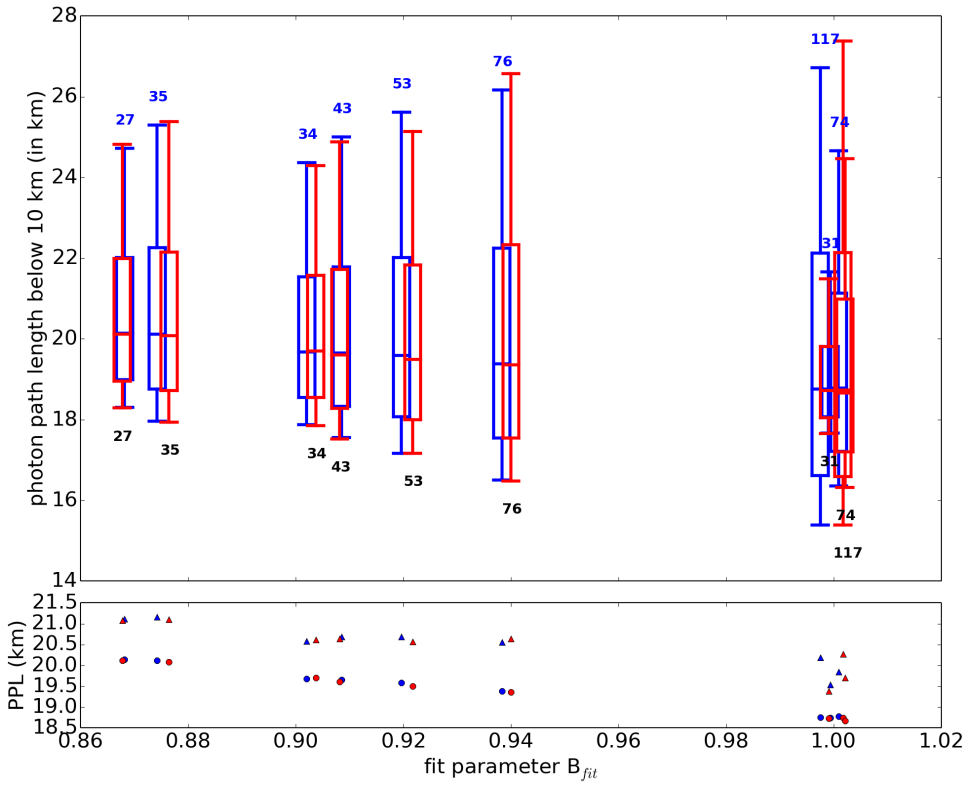

Figure 64: Top: Photon path length distributions for a selection of cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter for two values of g (g=0.85 in blue, g=0.8 in red). Bottom: Medians (blue dots for g=0.85, red dots for g=0.8) and means (blue triangles for g=0.85 and red triangles for g=0.8) of the photon path length distributions in dependence of B_{fit}.

For the presented case study, one can draw the conclusion that - as long as the cloud optical depth is optimized in the continuum region - the fit results appear very similar for g=0.85 and g=0.8. Here, the macroscopic parameters are more important for the derived photon path distributions than the choice of the microphysical parameter g.

10.5.2 Photon path length statistics

In the following, photon path length statistics are shown for the cloud scenarios 117, 74 and 31, considering both investigated values of g. Only a very small set of scenarios has been chosen for the visualization for purposes of clarity. For the following overview, the selected cases with B_{fit} near unity are chosen because both the spatial dimensions of the cloud scenarios as well as the photon path length distributions are quite different from each other. The cloud scenarios leading to the smallest overall RMS, on the other hand, show more similar photon path length distributions.

First the penetration altitudes of the detected photons are shown for both g=0.8 and g=0.85 in fig. 65. The penetration altitude defines the lowest altitude within the model atmosphere which

is traveled by the simulated photons. They are nearly identical for both asymmetry parameters and will be shortly discussed in the following:

Almost no photons with penetration altitudes above the cloud top are detected. The Rayleigh scattering cross section is small in the considered wavelength regime which leads to a low probability of these events. The fraction of photons touching the surface of the modeled Earth is around 1% and nearly identical for all clouds. This is to be expected because the optical depths of the clouds for the same g are approximately equal.

Most photons have penetration altitudes within the cloud layers due to the high probability of scattering on the cloud particles. The most probable penetration altitudes are situated around the upper third of the cloud vertical extent. Because of the different vertical dimensions of the cloud scenarios but similar cloud optical depths, the distribution of cloud droplets is denser for vertically less extended clouds. This increases the probability that the photons do not penetrate into lower regions of the cloud, leading to larger detected photon fractions for those penetration altitudes.

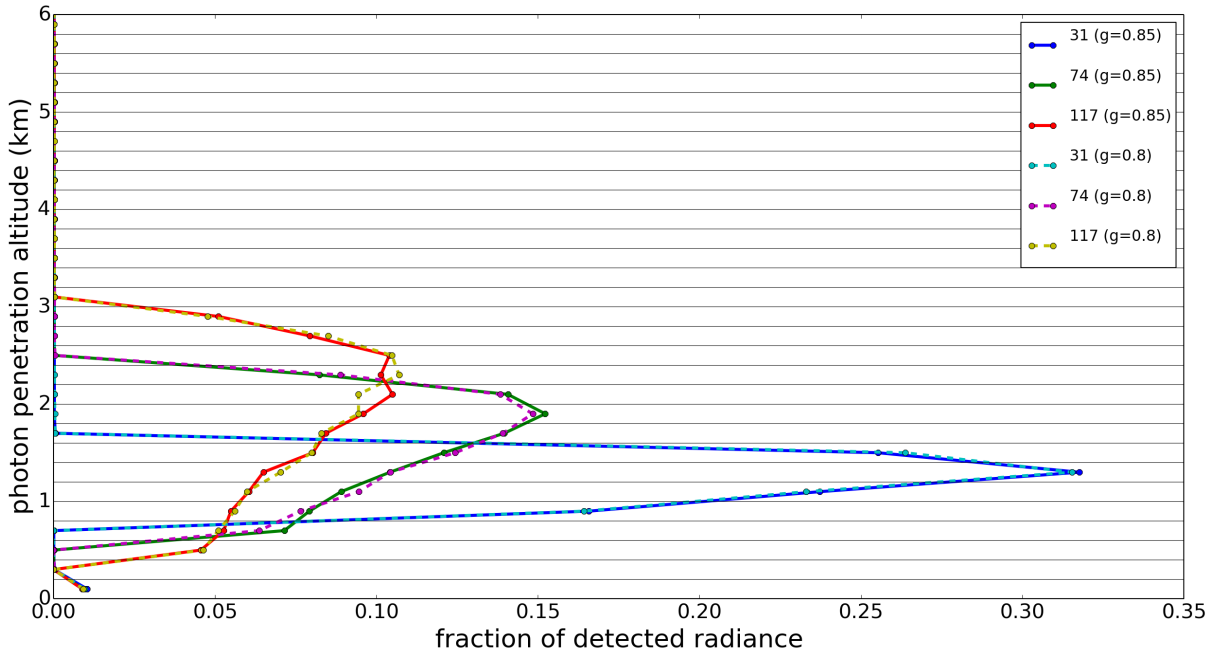

Figure 65: Penetration altitudes for the cloud scenarios 31, 74 and 117 with g=0.85 and g=0.8.

Another property of the calculated photon path length statistics are the path length distributions of the photons below altitudes of 10 km. They are visualized in fig. 66. The O_2 weighted distributions are shown in fig. 75 (Appendix, section 10.7). Comparing the photon path length distributions with clouds described by an asymmetry parameter of 0.8 and 0.85, again no significant differences are observed.

The distributions show that significant fractions of the detected radiance are contributed by photons being backscattered at the cloud top. As for the penetration altitudes, the distributions are less wide for clouds of smaller vertical extent but have higher peaks (probably resulting in similar median values of the photon path lengths) as shown in fig. 66. The slopes of the distributions for larger photon path lengths are caused by multiple scattering which can significantly enhance the trajectories. Cloud scenario 117 shows the largest fractions of longer photon path lengths, it also has the largest vertical extent. By having a similar cloud optical depth to the other two scenarios (of the same g), the mean distance between the cloud droplets and thus the distance

between scattering events, increases. The photons consequently travel longer paths inside the cloud system before leaving in the direction of the detector position.

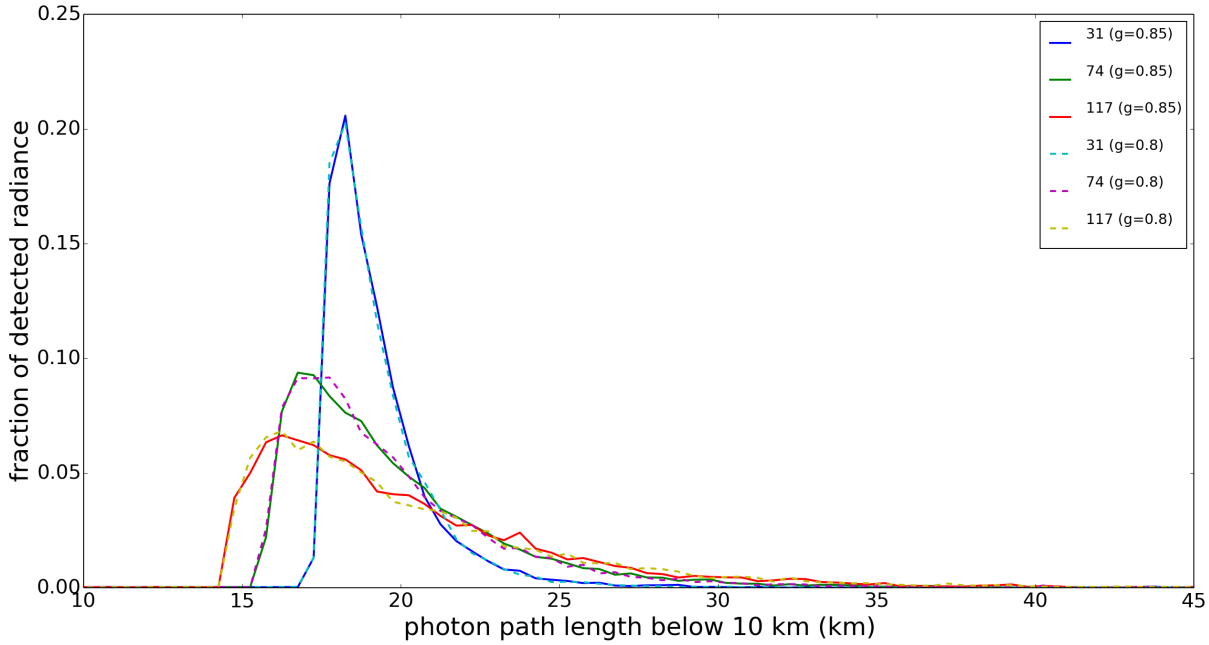

Figure 66: Comparison of the geometric photon path length distributions below 10 km for the cloud scenarios 31, 74 and 117 with g=0.85 and g=0.8.

Altitude resolved distributions of the photon path lengths can also be calculated but are not shown at this point. An example of this representation can be found for a cloud scenario within case study 3 (section 12.4).

10.6 Further investigations

In the following, further investigations are summarized which serve to identify potential uncertainty/error sources during the retrieval process. These include a measure of the cloud top height heterogeneity within the TANSO-FTS FOV, the consideration of the variability of atmospheric temperature and pressure profiles as well as the spectral sampling accuracy of the used O_2 absorption cross sections.

10.6.1 Horizontal heterogeneity of cloud case 1

Besides the cloud optical depth derived from TANSO-CAI measurements and the TANSO-FTS radiances, collocated radar and lidar data from NASA A-train satellite measurements are used to infer information on the cloud scenario under consideration. As can be seen in fig. 67, the collocated swaths only cover a narrow region of the TANSO-FTS FOV. Detailed information on the cloud profile within the total region is thus mostly unknown. One possible influence on disagreements between the retrieval results and the collocated cloud profile data is the possible horizontal heterogeneity of the cloud top heights (and also cloud optical depths). To gain further information on this, cloud top pressures with a spatial resolution of 1 km from measurements of MODIS (Moderate Resolution Imaging Spectrometer) on Aqua [7; 56], which also flies in the NASA afternoon constellation, are mapped onto the collocation plot (locations 0,1,2 in fig. 67).

The corresponding cloud top pressures are:

- location 0: p ≈ 855 hPa
- location 1: p ≈ 825 hPa
- location 2: p ≈ 825 hPa

The difference in the corresponding cloud top heights is approx. 300 m, which is between one and two altitude grids in our simulation. It can be concluded that the spatial inhomogeneity of the cloud top heights is small compared to the deviations between the lidar- and radar-derived cloud profiles and the vertical cloud dimensions obtained from the O_2 fit of the O_2 A-band.

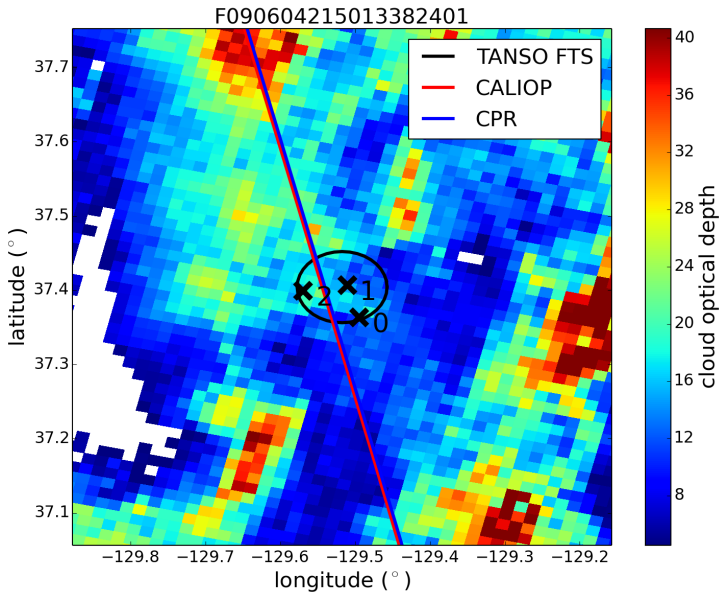

Figure 67: Collocation plot of the TANSO-FTS FOV, CALIOP and CPR measurements together with the MODIS Aqua measurement locations where the cloud top pressure is retrieved (datafile: MYD06_L2.A2009155.2200.051.2009157225111.hdf).

10.6.2 Influence of atmospheric pressure-temperature profiles

As mentioned in section 6.2.1, a common atmospheric pressure-temperature profile of the atmosphere is used for all simulations, regardless of the real atmospheric conditions during the measurement. In order to evaluate the influence of changes in the pressure-temperature profiles, profiles from the measurement position itself are used. Altitude resolved pressures and temperatures from the CALIOP 2B-GEOPROF data product of the collocated measurement locations are used as input for the simulations. These profiles are provided until an altitude of 40 km and are external information of a MERRA-2 product [1; 30].

The new model atmosphere below 40 km is derived by linear interpolation of these profiles to the McArtim altitude grid. For higher altitudes, the profiles from the US standard atmosphere are used. As a first test, only the temperature values are replaced in comparison to the standard grid, as a second test, both temperature and pressure values are replaced. The absorption cross sections of O_2 are calculated based on the updated atmosphere and the number density of O_2 is updated using the ideal gas law.

As selection, the cloud cases 31, 74 and 117 are considered. After setup of the new simulation environment, the cloud optical depth is derived as before and the fit is performed on the new radiance calculations. The results are visualized in figure 68 showing the parameters using

the standard profile (numbers with blue circles), the parameters considering solely an adapted temperature profile (numbers + T with green triangles) and the parameters using profiles with adapted pressure and temperature (numbers + A with red squares).

The fit residual is still smallest for the original profile but for most of adapted profiles slightly higher B_{fit} values are found (only exception: 74T). The comparison shows that an overestimation of the simulated O_2 absorption up to 1% can be explained. However, a systematic overestimation of the O_2 absorption in the order of 5-10% cannot be explained at this stage.

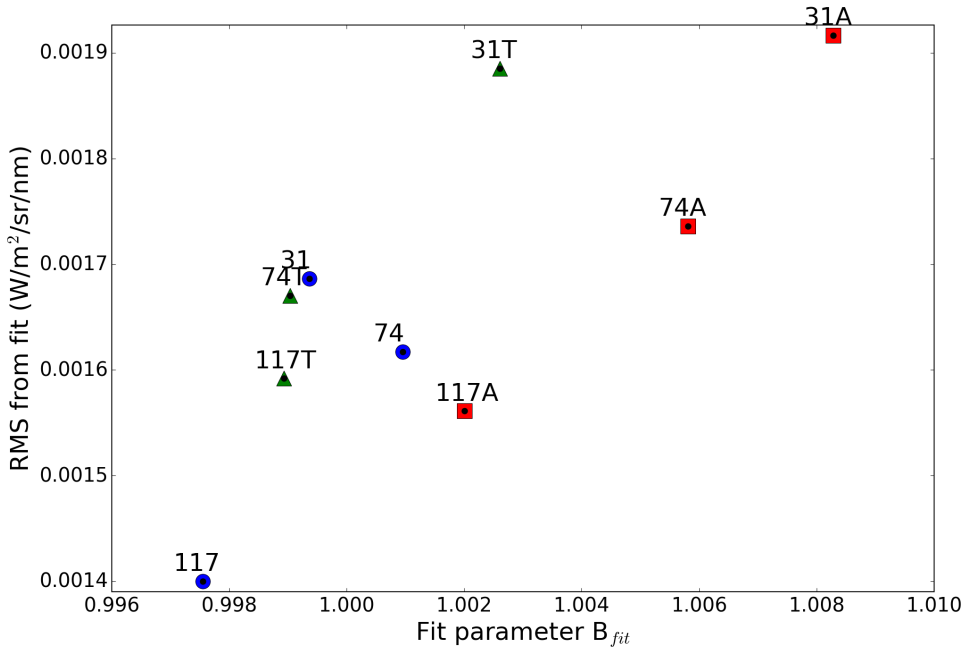

Figure 68: Fit results RMS and B_{fit} for selected cloud scenarios from simulated radiance spectra based on different atmospheric profiles (original profile in blue circles, temperature adapted profiles in green triangles and pressure-temperature adapted profiles in red squares).

10.6.3 Influence of the spectral sampling of the O_2 absorption cross sections

For all simulations, the absorption cross sections of O_2 are expressed in a look-up table on a wavenumber grid of 0.05 cm^{-1} ($\approx$ 0.003 nm). During the simulation, the values are interpolated to the wavelength values of the selected spectrum. The spacing of this wavelength grid is a possible source for systematic errors. To investigate the effect on the fit results, a finer grid size with a factor 5 is assumed (0.01 cm^{-1}).

The simulations for selected cloud scenarios have been repeated for the finer resolved look-up table for cases 6-27 (clouds with top heights until 1.4 km), cases 34 and 76 (low RMS) as well as cases 31, 74, 117 (B_{fit} close to unity). The cloud optical depths have not been adapted because only the (very low absorbing) continuum region is considered for the determination and a statistical error of the cloud optical depth retrieval on the results should not influence the outcome of the investigation. The resulting fit RMS and B_{fit} are shown in fig. 69 for all considered cloud scenarios. The additional F of the annotations indicates the usage of the finer resolved O_2 absorption cross section data (datapoints as green triangles, the original fit results are shown as blue circles).

Indeed, a systematic shift of B_{fit} to higher values is found, explaining a systematic overestimation of up to 1%. Case 16F, for example, which would agree well with the collocated lidar and radar measurements, now shows B_{fit} values closer to unity. Interestingly, cloud scenarios with

$B_{fit} < 0.97$ lead to a higher RMS when the finer spectral sampling is used while those with $B_{fit} > 0.97$ show smaller and higher RMS, depending on the cloud scenario. This provides an indication that the shape of the B_{fit} - RMS dependence (as seen in fig. 61) slightly changes for a finer resolved O_2 absorption cross section. Although the influence on B_{fit} is rather small, the updated look-up table of O_2 absorption cross sections should be used in future retrievals, also considering the high spectral resolution of the TANSO-FTS spectra. Because this sensitivity study was performed after consideration of all performed case studies, the results of case study 2 and 3 are still based on the coarser wavelength grid.

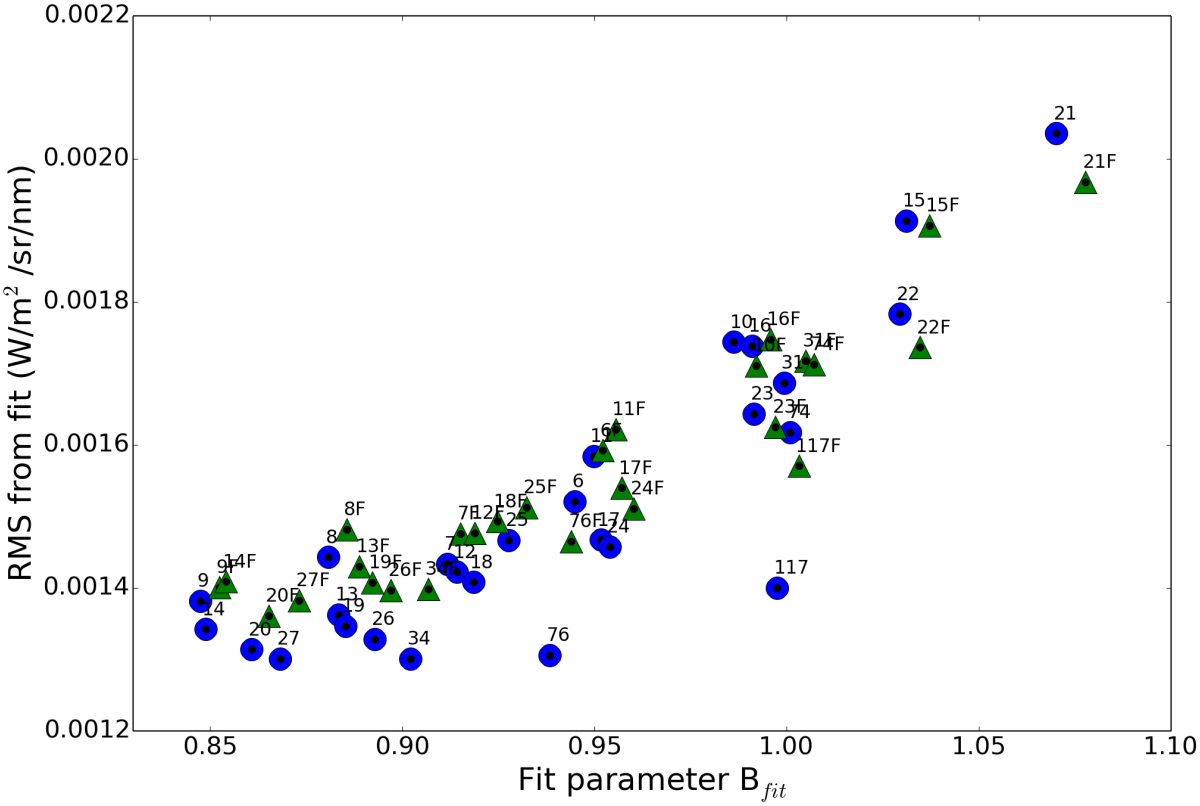

Figure 69: Fit results RMS and B_{fit} for radiance calculations based on different wavelength resolutions of the O_2 absorption cross sections (F defines a finer resolution).

10.7 Appendix

10.7.1 Cloud configurations for case 1 using an asymmetry parameter g=0.85

Table 21: All performed simulations for case study 1 using g=0.85. Each parameter combination for the model cloud has an assigned number which is used in the result plots.

num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}
0	0.2	0.2	16.14	40	1.8	1.0	16.09	80	2.6	0.6	16.06
1	0.4	0.2	16.17	41	1.8	1.2	16.1	81	2.6	0.8	16.38
2	0.4	0.4	16.09	42	1.8	1.4	16.08	82	2.6	1.0	15.99
3	0.6	0.2	16.1	43	1.8	1.6	16.05	83	2.6	1.2	16.05
4	0.6	0.4	16.03	44	1.8	1.8	16.4	84	2.6	1.4	16.15
5	0.6	0.6	16.13	45	2.0	0.2	16.06	85	2.6	1.6	16.14
6	0.8	0.2	16.08	46	2.0	0.4	16.11	86	2.6	1.8	16.13
7	0.8	0.4	16.11	47	2.0	0.6	16.07	87	2.6	2.0	16.12
8	0.8	0.6	16.18	48	2.0	0.8	16.09	88	2.6	2.2	16.07
9	0.8	0.8	16.12	49	2.0	1.0	16.05	89	2.6	2.4	16.11
10	1.0	0.2	16.1	50	2.0	1.2	16.04	90	2.6	2.6	16.16
11	1.0	0.4	16.14	51	2.0	1.4	16.18	91	2.8	0.2	16.0
12	1.0	0.6	16.11	52	2.0	1.6	16.08	92	2.8	0.4	16.01
13	1.0	0.8	16.14	53	2.0	1.8	16.08	93	2.8	0.6	16.06
14	1.0	1.0	16.03	54	2.0	2.0	16.16	94	2.8	0.8	16.06
15	1.2	0.2	16.14	55	2.2	0.2	16.07	95	2.8	1.0	16.03
16	1.2	0.4	16.14	56	2.2	0.4	16.09	96	2.8	1.2	15.91
17	1.2	0.6	16.09	57	2.2	0.6	16.14	97	2.8	1.4	16.09
18	1.2	0.8	16.12	58	2.2	0.8	16.02	98	2.8	1.6	16.11
19	1.2	1.0	16.08	59	2.2	1.0	16.11	99	2.8	1.8	16.03
20	1.2	1.2	16.13	60	2.2	1.2	16.05	100	2.8	2.0	16.05
21	1.4	0.2	16.07	61	2.2	1.4	16.17	101	2.8	2.2	16.11
22	1.4	0.4	16.08	62	2.2	1.6	16.03	102	2.8	2.4	16.15
23	1.4	0.6	16.09	63	2.2	1.8	16.11	103	2.8	2.6	16.09
24	1.4	0.8	16.05	64	2.2	2.0	16.13	104	2.8	2.8	16.09
25	1.4	1.0	16.16	65	2.2	2.2	16.01	105	3.0	0.2	16.08
26	1.4	1.2	16.05	66	2.4	0.2	16.02	106	3.0	0.4	16.07
27	1.4	1.4	16.13	67	2.4	0.4	16.12	107	3.0	0.6	16.05
28	1.6	0.2	16.13	68	2.4	0.6	16.12	108	3.0	0.8	16.06
29	1.6	0.4	16.14	69	2.4	0.8	16.09	109	3.0	1.0	16.07
30	1.6	0.6	16.19	70	2.4	1.0	16.09	110	3.0	1.2	16.1
31	1.6	0.8	16.13	71	2.4	1.2	16.16	111	3.0	1.4	16.05
32	1.6	1.0	16.12	72	2.4	1.4	16.15	112	3.0	1.6	16.09
33	1.6	1.2	16.08	73	2.4	1.6	16.15	113	3.0	1.8	16.09
34	1.6	1.4	16.06	74	2.4	1.8	16.17	114	3.0	2.0	16.08
35	1.6	1.6	16.1	75	2.4	2.0	16.08	115	3.0	2.2	16.13
36	1.8	0.2	16.04	76	2.4	2.2	16.1	116	3.0	2.4	16.07
37	1.8	0.4	16.17	77	2.4	2.4	16.19	117	3.0	2.6	16.08
38	1.8	0.6	16.17	78	2.6	0.2	16.03	118	3.0	2.8	16.13
39	1.8	0.8	16.13	79	2.6	0.4	15.97	119	3.0	3.0	16.03

10.7.2 Cloud configurations for case 1 using an asymmetry parameter g=0.80

Table 22: All performed simulations for case study 1 using g=0.8. Each parameter combination for the model cloud has an assigned number which is used in the result plots.

num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}
0	0.2	0.2	12.21	40	1.8	1.0	12.17	80	2.6	0.6	12.22
1	0.4	0.2	12.24	41	1.8	1.2	12.18	81	2.6	0.8	12.16
2	0.4	0.4	12.25	42	1.8	1.4	12.15	82	2.6	1.0	12.21
3	0.6	0.2	12.27	43	1.8	1.6	12.15	83	2.6	1.2	12.16
4	0.6	0.4	12.21	44	1.8	1.8	12.22	84	2.6	1.4	12.17
5	0.6	0.6	12.25	45	2.0	0.2	12.16	85	2.6	1.6	12.16
6	0.8	0.2	12.23	46	2.0	0.4	12.19	86	2.6	1.8	12.19
7	0.8	0.4	12.22	47	2.0	0.6	12.22	87	2.6	2.0	12.19
8	0.8	0.6	12.22	48	2.0	0.8	12.18	88	2.6	2.2	12.23
9	0.8	0.8	12.25	49	2.0	1.0	12.23	89	2.6	2.4	12.17
10	1.0	0.2	12.18	50	2.0	1.2	12.26	90	2.6	2.6	12.23
11	1.0	0.4	12.19	51	2.0	1.4	12.16	91	2.8	0.2	12.15
12	1.0	0.6	12.19	52	2.0	1.6	12.23	92	2.8	0.4	12.19
13	1.0	0.8	12.22	53	2.0	1.8	12.22	93	2.8	0.6	12.2
14	1.0	1.0	12.22	54	2.0	2.0	12.23	94	2.8	0.8	12.16
15	1.2	0.2	12.22	55	2.2	0.2	12.44	95	2.8	1.0	12.17
16	1.2	0.4	12.19	56	2.2	0.4	12.22	96	2.8	1.2	12.21
17	1.2	0.6	12.28	57	2.2	0.6	12.16	97	2.8	1.4	12.16
18	1.2	0.8	12.22	58	2.2	0.8	12.19	98	2.8	1.6	12.16
19	1.2	1.0	12.16	59	2.2	1.0	12.19	99	2.8	1.8	12.24
20	1.2	1.2	12.22	60	2.2	1.2	12.2	100	2.8	2.0	12.22
21	1.4	0.2	12.24	61	2.2	1.4	12.21	101	2.8	2.2	12.2
22	1.4	0.4	12.12	62	2.2	1.6	12.16	102	2.8	2.4	12.22
23	1.4	0.6	12.2	63	2.2	1.8	12.28	103	2.8	2.6	12.21
24	1.4	0.8	12.16	64	2.2	2.0	12.18	104	2.8	2.8	12.19
25	1.4	1.0	12.22	65	2.2	2.2	12.18	105	3.0	0.2	12.14
26	1.4	1.2	12.18	66	2.4	0.2	12.17	106	3.0	0.4	12.23
27	1.4	1.4	12.21	67	2.4	0.4	12.19	107	3.0	0.6	12.19
28	1.6	0.2	12.16	68	2.4	0.6	12.18	108	3.0	0.8	12.15
29	1.6	0.4	12.2	69	2.4	0.8	12.24	109	3.0	1.0	12.16
30	1.6	0.6	12.23	70	2.4	1.0	12.22	110	3.0	1.2	12.13
31	1.6	0.8	12.22	71	2.4	1.2	12.13	111	3.0	1.4	12.19
32	1.6	1.0	12.23	72	2.4	1.4	12.2	112	3.0	1.6	12.18
33	1.6	1.2	12.17	73	2.4	1.6	12.21	113	3.0	1.8	12.22
34	1.6	1.4	12.2	74	2.4	1.8	12.25	114	3.0	2.0	12.2
35	1.6	1.6	12.2	75	2.4	2.0	12.24	115	3.0	2.2	12.18
36	1.8	0.2	12.22	76	2.4	2.2	12.21	116	3.0	2.4	12.18
37	1.8	0.4	12.21	77	2.4	2.4	12.2	117	3.0	2.6	12.21
38	1.8	0.6	12.24	78	2.6	0.2	12.18	118	3.0	2.8	12.17
39	1.8	0.8	12.24	79	2.6	0.4	12.17	119	3.0	3.0	12.18

10.7.3 Fit results in sub regions of the O_2 A-band

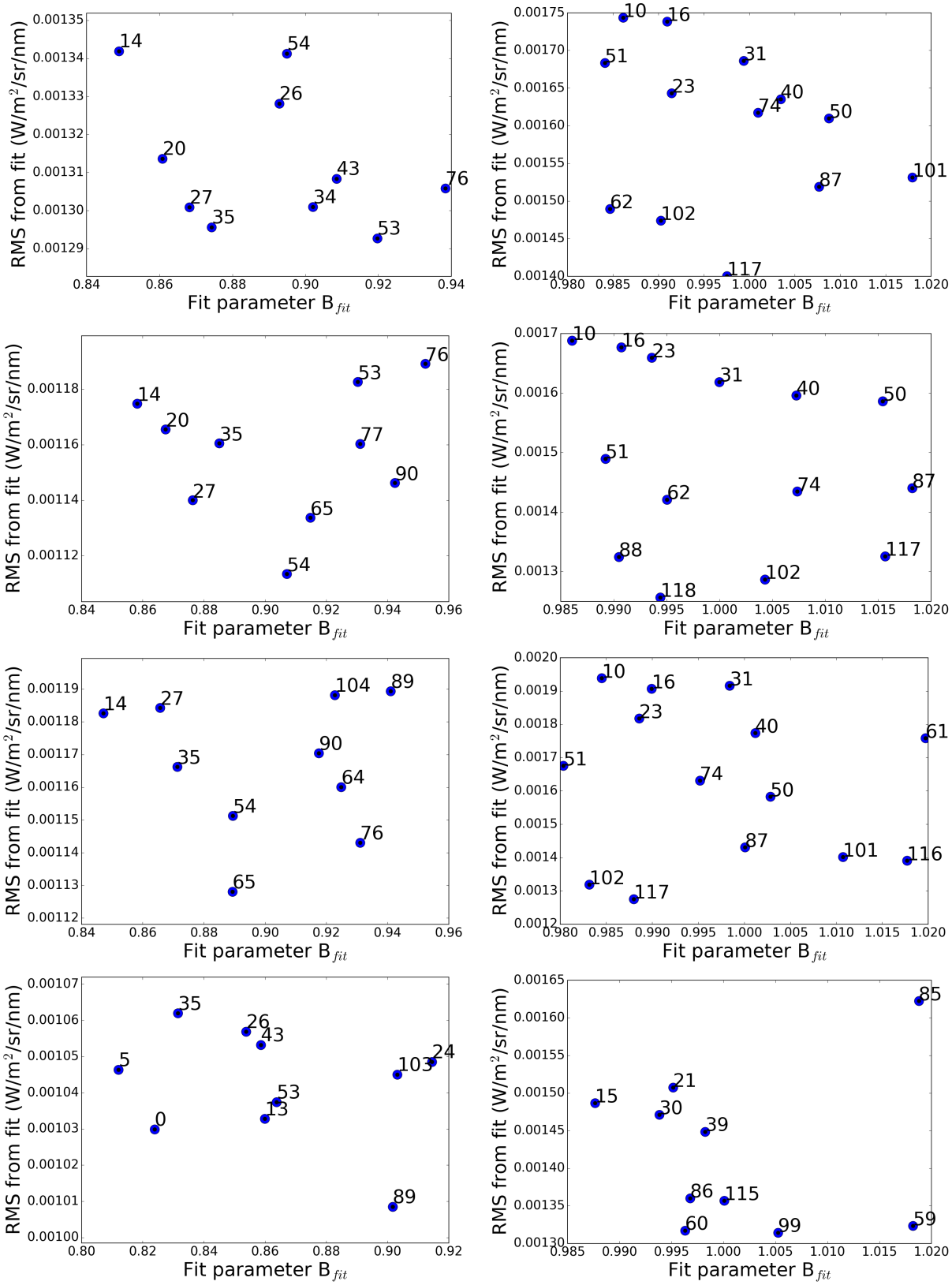

Figure 70: Fit results (RMS, B_{fit}) for different wavelength regions of the O_2 A-band. The fitting parameters of the 10 cloud scenarios with the smallest RMS are shown on the left, the right plots show those scenarios near B_{fit}=1. The wavelength regions are as follows: entire O_2 A-band (top), followed by strong O_2 absorption, intermediate O_2 absorption and weak O_2 absorption (bottom).

10.7.4 Influence of the simulation noise on the fit result

The simulation environment is configured to simulate as many photons as necessary to achieve a 1% precision on the unpolarized radiance. Furthermore, the radiance calculation for a given wavelength is repeated if the error of the entire polarized radiance is above 2%. The resulting noise of the Monte Carlo simulations can affect the fit result. In order to estimate this influence, the radiance simulations for three cloud scenarios have been performed 10 times without changing the simulation environment or the cloud characteristics. The cloud scenarios are as follows:

- cloud between 0.4 and 3.0 km, $\tau_{cloud} = 16.09$ (green squares in fig. 71)
- cloud between 0.6 and 2.4 km, $\tau_{cloud} = 16.17$ (red triangles in fig. 71)
- cloud between 0.8 and 1.6 km, $\tau_{cloud} = 16.09$ (black circles in fig. 71)

The resulting fit RMS and B_{fit} are shown in figure 71. From this small set of simulations, $\Delta RMS \approx \pm\ 10^{-4}$ W/m^2/sr/nm and $\Delta B_{fit} \approx \pm 2 \cdot 10^{-3}$ can be inferred.

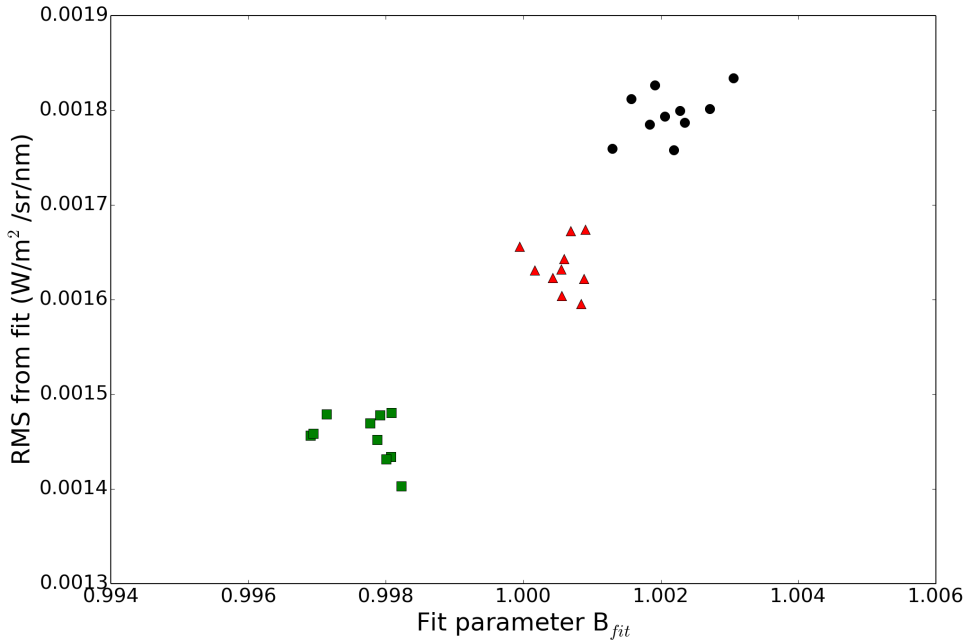

Figure 71: Fit results (RMS, B_{fit}) of three cloud scenarios (marked by green squares, red triangles and black circles) for 10 radiance calculations with the same set-up.

10.7.5 O_2 weighted photon path length distributions

Within the main description of the findings, the photon path length distributions are calculated using the geometric distances between the photon events, which allows a direct interpretation of the results with regard to the traveled distances. In principle, the measured or simulated spectra are not determined by the geometric distances but by the geometric distances weighted with the O_2 concentration. Thus an alternative representation of the photon path length distributions is included where a weighting of the individual distances between the events by the O_2 concentration is performed. The weighted photon path lengths are derived as follows: The O_2 concentration within each altitude layer is expressed in relation to the average O_2 concentration below 50 km altitude. Instead of using the geometrical lengths, the path lengths are multiplied by the derived O_2 ratios. Consequently, the photon path length distributions are not calculated below 10 km (fig. 66), but below 50 km (fig. 75). The O_2 weighted photon path length distributions for all cloud scenarios with g=0.85 are shown in figure 72. As shown in figure 73, there is almost a linear relationship between the mean and median values of the geometric and the O_2 weighted photon path length distributions. Although the alternative representation leads to the same conclusions as already discussed for the geometric photon path lengths, it is included for reasons of completeness.

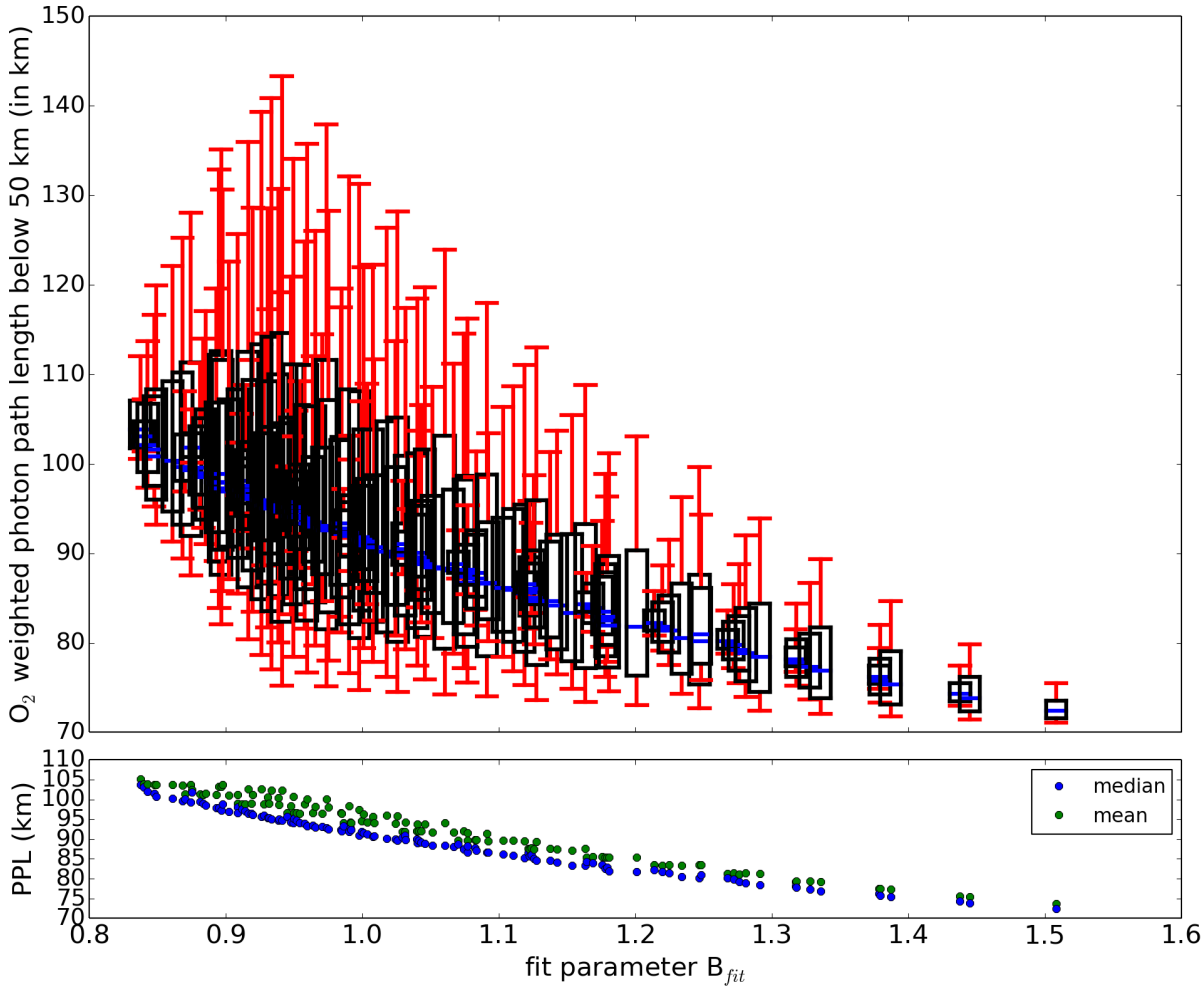

Figure 72: Top: O_2 weighted photon path length distributions for all assumed cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

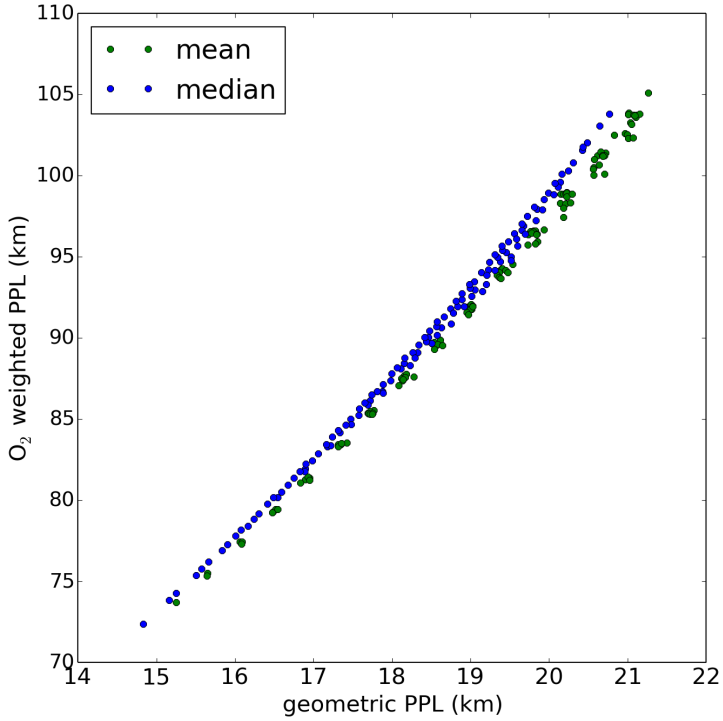

Figure 73: Mean (green) and median (blue) values of the photon path length distributions (PPL), using the geometric or the O_2 weighted results.

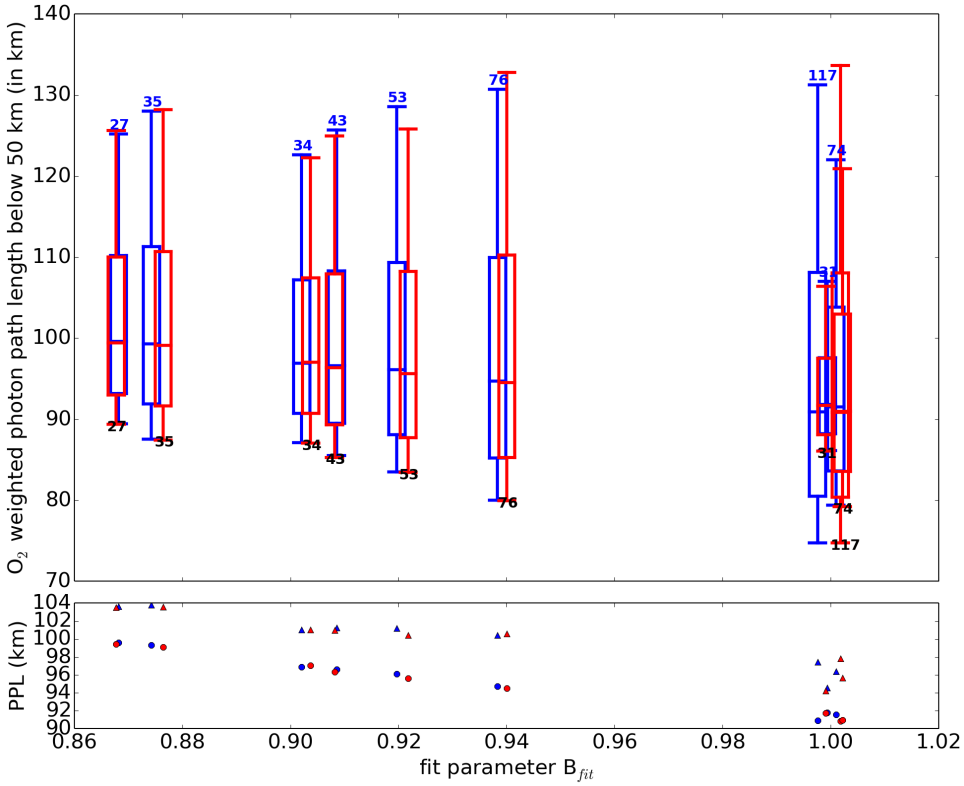

Figure 74: Top: O_2 weighted photon path length distributions for a selection of cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter for two values of g (g=0.85 in blue, g=0.8 in red). Bottom: Medians (blue dots for g=0.85, red dots for g=0.8) and means (blue triangles for g=0.85 and red triangles for g=0.8) of the photon path length distributions in dependence of B_{fit}.

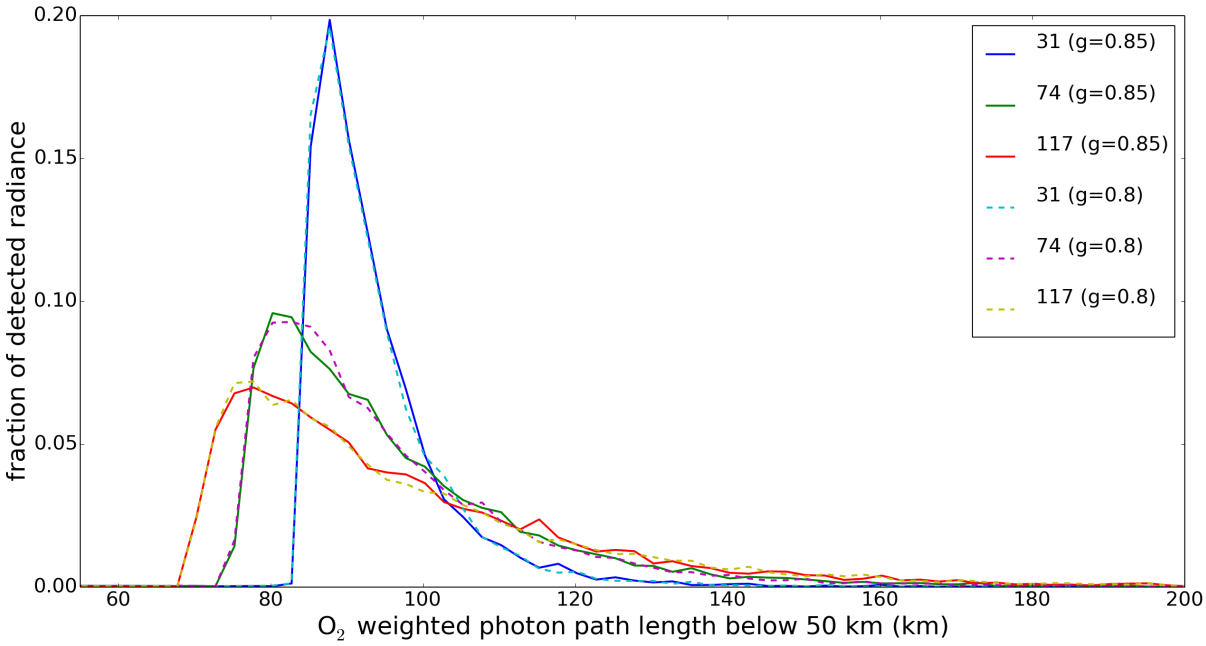

Figure 75: Comparison of the O_2 weighted photon path length distributions for the cloud cases 31, 74 and 117 with g=0.85 and g=0.8.

11 Case study 2

The scenario of case study 2 includes a one-layered cloud system which covers the entire TANSO-FTS field of view. The averaged cloud optical depth inside the TANSO-FTS FOV from the TANSO-CAI pixel radiances is $\tau_{CAI} = 19.72$ and the collocated radar measurements indicate $\tau_{CPR} = 25.75$ in average at those pixels inside the FOV where TANSO-CAI detects higher cloud optical depths (see fig. 80 in Appendix, section 11.4). The average cloud top height estimated from collocated lidar measurements is $h_{CALIOP} = 2.04$ km. For higher cloud optical depths, lidar signals do not penetrate deeply into the cloud, thus radar measurements from the CloudSat CPR are used for an estimation of the cloud vertical extent. Here, cloud optical depths of different strengths are found between altitudes of approx. 0.8 and 2.0 km.

11.1 Simulated cloud scenarios

A vertical (and horizontal) inhomogeneity of the cloud extinction is measured (see radar-based cloud profile in section 7.6.2, fig. 35). Nevertheless, a homogeneous cloud structure is assumed in the radiative transfer simulations of the radiances in a first attempt.

The clouds are modeled as one-dimensional entities of different cloud top heights (0.2 km $\leq h_{cloud} \leq$ 3.0 km) and cloud vertical extents ($0.2 \leq e_{cloud} \leq h_{cloud}$). The optical properties are described by a Henyey-Greenstein phase function using an asymmetry parameter of g=0.85. The reflectance of the water surface of the measurement location is set to 0.03. All cloud configurations used are listed in table 25 (Appendix, section 11.4). The spatial cloud configurations are identical to the cloud models of case 1, they only differ in the cloud optical depth τ_{cloud}. However, the measurement situation of the TANSO-FTS radiance spectrum features a higher cloud with a larger vertical extent compared to case 1 and the radar measurements show that a higher probability of vertical inhomogeneity within the cloud structure is given. Nevertheless, the results allow common conclusions. As a consequence and because this is the second case for a one-layer cloud system, the results will be only shortly discussed.

11.2 Spectrum fit and results

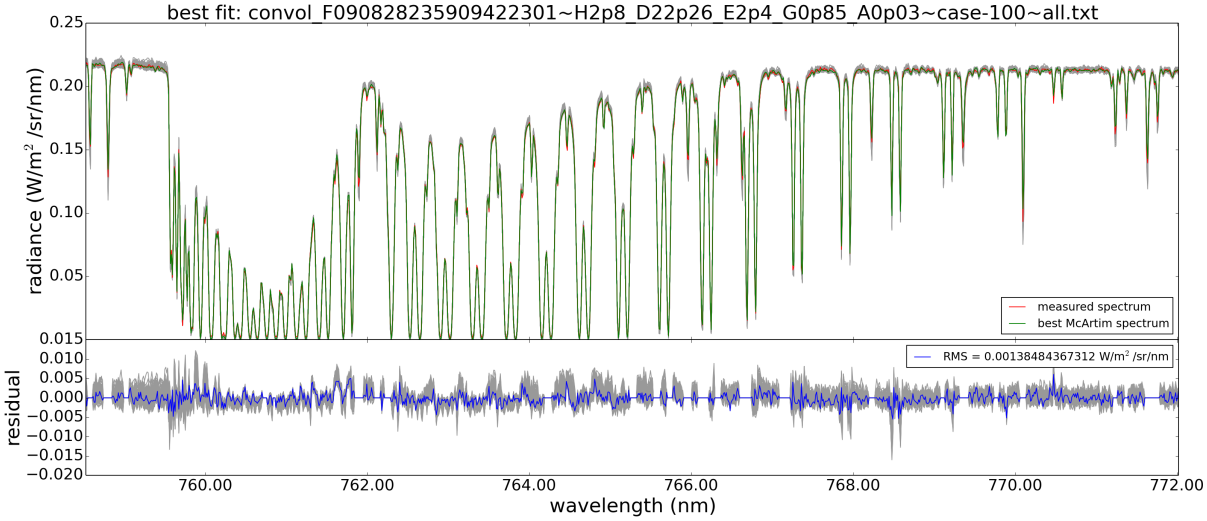

Figure 76: Top: Fitted spectra (grey), the TANSO-FTS measurement spectrum (red) and the fitted spectrum with the lowest RMS (green). Bottom: The corresponding residuals (grey) and those with the smallest RMS (blue).

The fitted spectra are shown in fig. 76. As before, the cloud optical depth for each cloud case

is determined in the continuum region. The fitting process results in a good alignment of the radiance values between the fitted simulations and the measurement.

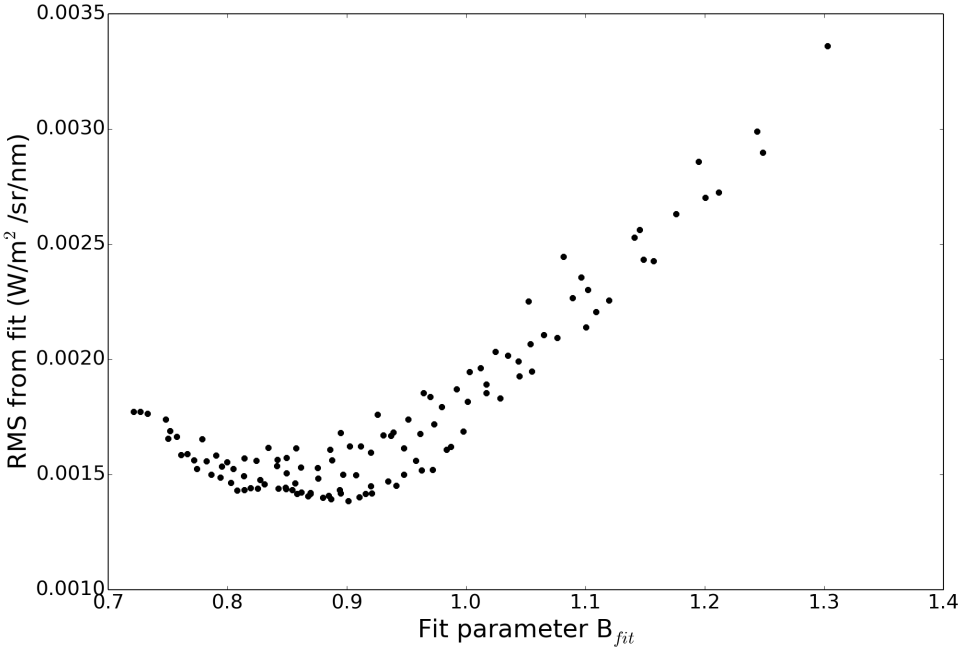

Figure 77: The resulting fit RMS in dependence of the fit parameter B_{fit} for all simulated cloud scenarios.

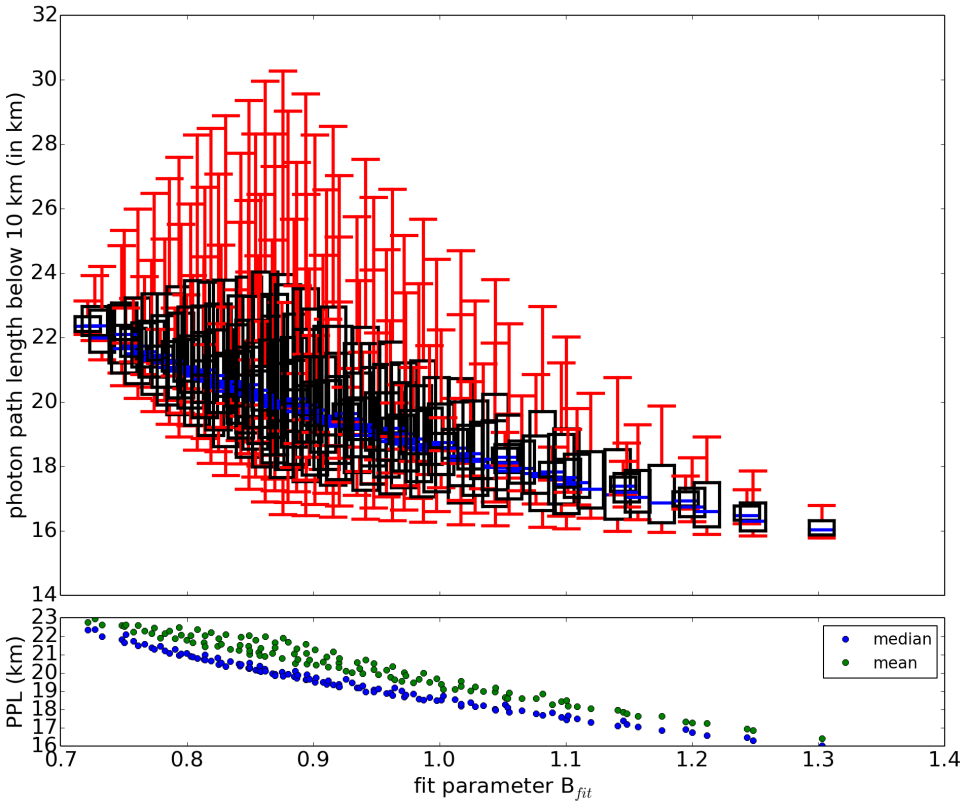

Figure 78: Top: Photon path length distributions for all assumed cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

The resulting fit RMS and B_{fit} for all performed simulations are presented in figure 77. Similar to case 1, the minimum RMS is found for values of B_{fit} around 0.9 which suggests a systematic overestimation of the O_2 absorption during the simulations. The photon path length distributions below 10 km (for λ=772.145 nm) are visualized as boxplots in the top plot of fig. 78. The average photon path lengths $\overline{PPL}$ (green) together with the medians of the distributions (blue) are shown in the lower plot. The results show that B_{fit} is anticorrelated to the median of the photon path length distribution. Different $\overline{PPL}$ lead to similar values of B_{fit} indicating that the complete photon path length distribution influences the results. The corresponding O_2 weighted distributions are shown in Appendix (section 11.4), fig. 81.

Again, two strategies are followed for the interpretation of the fit results. In a first attempt, only the cloud scenarios with the smallest resulting RMS are selected while in a second attempt, the residual of the fit is considered together with the value of B_{fit}. Both selections of cloud scenarios are presented in fig. 79 with the photon path length distributions, averages and medians from fig. 78. The left side scenarios are those with the smallest RMS and the right side shows those with the lowest RMS values with B_{fit} near unity. The O_2 weighted distributions are shown in Appendix (section 11.4), fig. 82.

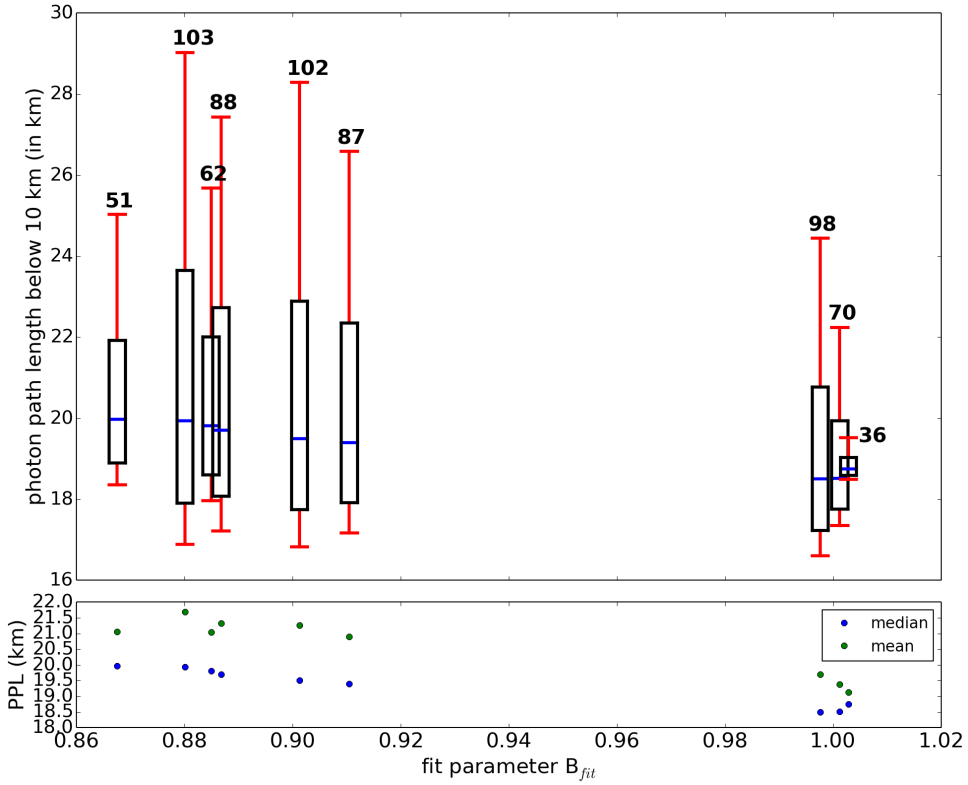

Figure 79: Top: Photon path length distributions for a selection (lowest RMS, $B_{fit} \approx 1$) of cloud scenarios in boxplot representation (9^{st}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

For the lowest RMS, B_{fit} values between $0.86 < B_{fit} < 0.92$ are found (see table 23), all showing an overestimation of the simulated O_2 absorption. The average photon path lengths $\overline{PPL}$ lie between 20.9 and 21.7 km, the median of the distributions takes lower values between 19.4 and 20.0 km. These values do not differ much, thus an approximate $\overline{PPL}$ and median can be given despite the ambiguity of resulting cloud scenarios. The average photon path lengths are reduced in comparison to the geometrical path through a clear atmosphere which agrees to the assumption that a large fraction of the photons is backscattered at the cloud top. The observation of lower medians than $\overline{PPL}$ is in agreement with the shape of the distributions

which has a larger spread of values for long photon paths. Many of the cloud scenarios from the selection have similar distribution widths. While scenario 51 (cloud between 0.6 - 2.0 km) and scenario 62 (cloud between 0.6 - 2.2 km) are close to the measured cloud profiles (cloud between approx. 0.8 and 2.0 km), the cloud top heights from the other scenarios are higher than observed by CALIOP.

Table 23: Cloud parameters for the cloud scenarios with the lowest RMS values resulting from the fit together with the fit results RMS and B_{fit}. The fit RMS increases from the top of the table to the bottom.

num	h_{cloud}	e_{cloud}	τ_{cloud}	B_{fit}	RMS (W/m^2/sr/nm)
102	2.8	2.4	22.26	0.901	0.0038
88	2.6	2.2	22.31	0.887	0.00139
103	2.8	2.6	22.28	0.88	0.0014
87	2.6	2.0	22.35	0.91	0.0014
51	2.0	1.4	22.15	0.868	0.0014
62	2.2	1.6	22.33	0.885	0.00141

Concentrating on the second attempt, which focuses on the smallest deviation of B_{fit} from unity, the cloud cases 36, 70 and 98 are found (see table 24). The spatial dimensions of these three cloud cases are quite different. Case 36 describes a cloud layer between 1.6 and 1.8 km which agrees quite well to the lidar retrieved cloud top height but the vertical extent is much smaller than measured by the radar. Case 70 is nearest to the spatial information from the collocated measurements, featuring a cloud layer between 1.4 and 2.4 km. Case 98 describes a cloud between 1.2 and 2.8 km, where both the cloud top height and vertical extent are significantly larger than measured by CALIOP and CPR.

Table 24: Cloud parameters and fit results for the cloud scenarios with resulting B_{fit} close to unity and low RMS. The fit RMS increases from the top of the table to the bottom.

num	h_{cloud}	e_{cloud}	τ_{cloud}	B_{fit}	RMS (W/m^2/sr/nm)
98	2.8	1.6	22.24	0.998	0.0017
70	2.4	1.0	22.25	1.001	0.0018
36	1.8	0.2	22.18	1.003	0.0019

All together, these three cloud scenarios for $B_{fit} \approx 1$ and low RMS have spatial dimensions with moderate deviations from the collocated profile data. Although the cloud scenarios differ in the vertical profile, they lead to similar fit results. Nevertheless, they have similar average photon path lengths. The widths of the photon path length distributions differ substantially (see fig. 79) while the differences between the median and $\overline{PPL}$ are low. They have values of 18.5 - 18.75 km and 19.1 - 19.7 km, respectively.

11.3 Assumed simplifications and heterogeneity of cloud case 2

Within the retrieval, different simplifications were made, including the negligence of vertical and horizontal inhomogeneities of the cloud. A certain vertical inhomogeneity is indicated in the collocated CPR profiles of the cloud optical depth. The cloud top pressures from collocated MODIS Aqua measurements for different pixels within the TANSO-FTS FOV (see fig. 80 in Appendix, section 11.4) are similar (730 hPa and 755 hPa), indicating similar cloud top heights as well. The pixel-resolved TANSO-CAI cloud optical depth, also plotted in fig. 80, shows higher cloud optical depths within the easterly part of TANSO-FTS measurement field. The FTS instrument naturally measures an average of the cloud extinction. The retrieved cloud optical depth in this study shows a satisfying agreement with this value. Using a slightly

smaller asymmetry parameter would most probably reduce the value of τ_{cloud} because of a reduced tendency for forward scattering, as seen for case study 1.

11.4 Appendix

11.4.1 Cloud configurations for case 2 using an asymmetry parameter g=0.85

Table 25: All performed simulations for case study 2. Each parameter combination for the model cloud has an assigned number which is used in the result plots.

num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}	num	h_{cloud}	e_{cloud}	τ_{cloud}
0	0.2	0.2	22.31	40	1.8	1.0	22.44	80	2.6	0.6	22.16
1	0.4	0.2	22.32	41	1.8	1.2	22.22	81	2.6	0.8	22.27
2	0.4	0.4	22.3	42	1.8	1.4	22.54	82	2.6	1.0	22.28
3	0.6	0.2	22.17	43	1.8	1.6	22.25	83	2.6	1.2	22.37
4	0.6	0.4	22.31	44	1.8	1.8	22.31	84	2.6	1.4	22.22
5	0.6	0.6	22.23	45	2.0	0.2	22.39	85	2.6	1.6	22.25
6	0.8	0.2	22.32	46	2.0	0.4	22.26	86	2.6	1.8	22.33
7	0.8	0.4	22.4	47	2.0	0.6	22.28	87	2.6	2.0	22.35
8	0.8	0.6	22.42	48	2.0	0.8	22.39	88	2.6	2.2	22.31
9	0.8	0.8	22.66	49	2.0	1.0	22.3	89	2.6	2.4	22.42
10	1.0	0.2	22.27	50	2.0	1.2	22.39	90	2.6	2.6	22.42
11	1.0	0.4	22.42	51	2.0	1.4	22.15	91	2.8	0.2	22.24
12	1.0	0.6	22.48	52	2.0	1.6	22.34	92	2.8	0.4	22.21
13	1.0	0.8	22.31	53	2.0	1.8	22.27	93	2.8	0.6	22.13
14	1.0	1.0	22.28	54	2.0	2.0	22.37	94	2.8	0.8	22.17
15	1.2	0.2	22.3	55	2.2	0.2	22.12	95	2.8	1.0	22.5
16	1.2	0.4	22.38	56	2.2	0.4	22.2	96	2.8	1.2	22.2
17	1.2	0.6	22.5	57	2.2	0.6	22.37	97	2.8	1.4	22.31
18	1.2	0.8	22.32	58	2.2	0.8	22.35	98	2.8	1.6	22.24
19	1.2	1.0	22.33	59	2.2	1.0	22.34	99	2.8	1.8	22.23
20	1.2	1.2	22.35	60	2.2	1.2	22.42	100	2.8	2.0	22.29
21	1.4	0.2	22.36	61	2.2	1.4	22.31	101	2.8	2.2	22.15
22	1.4	0.4	22.38	62	2.2	1.6	22.33	102	2.8	2.4	22.26
23	1.4	0.6	22.19	63	2.2	1.8	22.23	103	2.8	2.6	22.28
24	1.4	0.8	22.43	64	2.2	2.0	22.41	104	2.8	2.8	22.38
25	1.4	1.0	22.25	65	2.2	2.2	22.31	105	3.0	0.2	22.4
26	1.4	1.2	22.48	66	2.4	0.2	22.34	106	3.0	0.4	22.14
27	1.4	1.4	22.4	67	2.4	0.4	22.23	107	3.0	0.6	22.27
28	1.6	0.2	22.35	68	2.4	0.6	22.25	108	3.0	0.8	22.38
29	1.6	0.4	22.31	69	2.4	0.8	22.38	109	3.0	1.0	22.48
30	1.6	0.6	22.34	70	2.4	1.0	22.25	110	3.0	1.2	22.21
31	1.6	0.8	22.39	71	2.4	1.2	22.29	111	3.0	1.4	22.63
32	1.6	1.0	22.34	72	2.4	1.4	22.34	112	3.0	1.6	22.36
33	1.6	1.2	22.38	73	2.4	1.6	22.25	113	3.0	1.8	22.42
34	1.6	1.4	22.33	74	2.4	1.8	22.22	114	3.0	2.0	22.27
35	1.6	1.6	22.42	75	2.4	2.0	22.13	115	3.0	2.2	22.31
36	1.8	0.2	22.18	76	2.4	2.2	22.35	116	3.0	2.4	22.37
37	1.8	0.4	22.26	77	2.4	2.4	22.2	117	3.0	2.6	22.31
38	1.8	0.6	22.3	78	2.6	0.2	22.39	118	3.0	2.8	22.27
39	1.8	0.8	22.35	79	2.6	0.4	22.23	119	3.0	3.0	22.34

11.4.2 Horizontal heterogeneity of cloud case 2

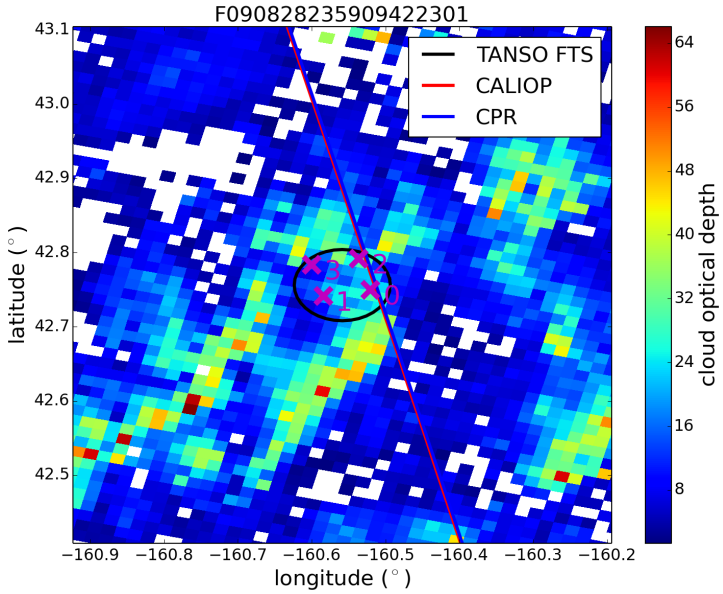

Figure 80: Collocation plot of the TANSO-FTS FOV, CALIOP and CPR measurements together with the MODIS Aqua measurement locations where the cloud top pressure is retrieved: $p \approx 755$ hPa for 1,2 and 4, $p \approx 730$ hPa for 3 (datafile: MYD06_L2.A2009240.2355.051.2009241174621.hdf)

11.4.3 O_2 weighted photon path length distributions

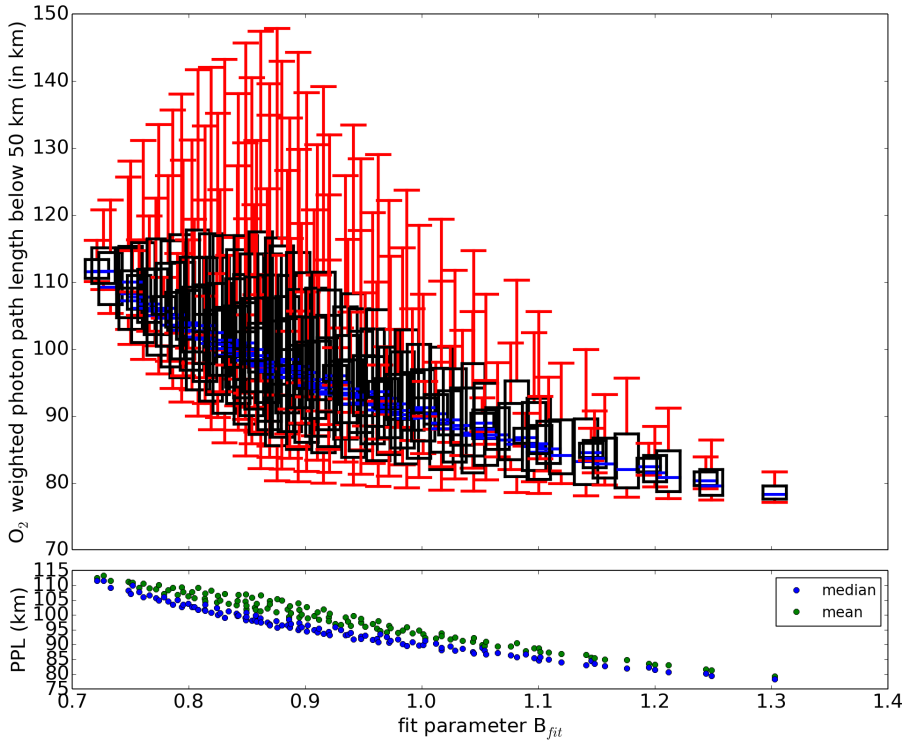

Figure 81: Top: O_2 weighted photon path length distributions for all assumed cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

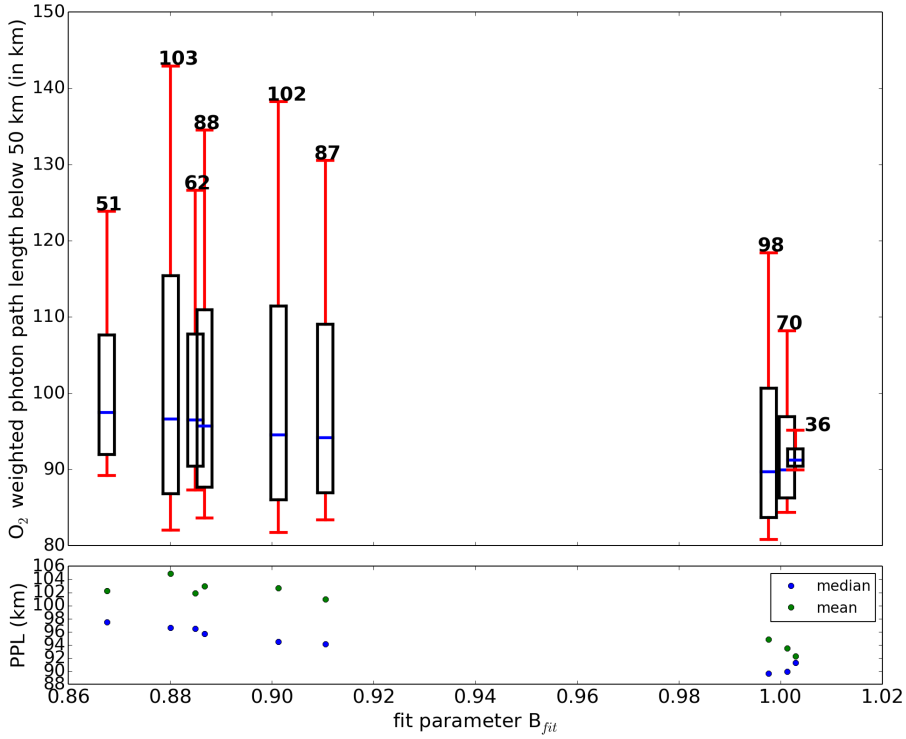

Figure 82: Top: O_2 weighted photon path length distributions for a selection (lowest RMS, $B_{fit} \approx 1$) of cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

12 Case study 3

After considering two single-layer cloud cases, now a measurement with a more complicated cloud scenario is investigated, which is described in detail in section 7.6.3. The CALIOP lidar detects cloud top heights at approximately 5 km and small vertical extents of the cloud. The radar signal of the CPR instrument allows to view the complete cloud profile and finds the cloud base height at about 0.48-0.68 km above ground. The profile structure indicates that high extinction takes place in the lower cloud layers, while the upper layers have smaller extinctions.

12.1 Simulated cloud scenarios

Different cloud scenarios are implemented in the McArtim model atmosphere to create a dataset of simulated radiance spectra. Because of the complexity of the considered cloud scenario, only a limited number of special scenarios is considered which are classified into three categories:

- Two-layer cloud structures with low and high fractions of cloud optical depths τ_{cloud} attributed to the layers having the same vertical extent
- Two-layer cloud structures adapted from lidar and radar measurements with low and high τ_{cloud} attributed to the layers
- One-layer cloud structures

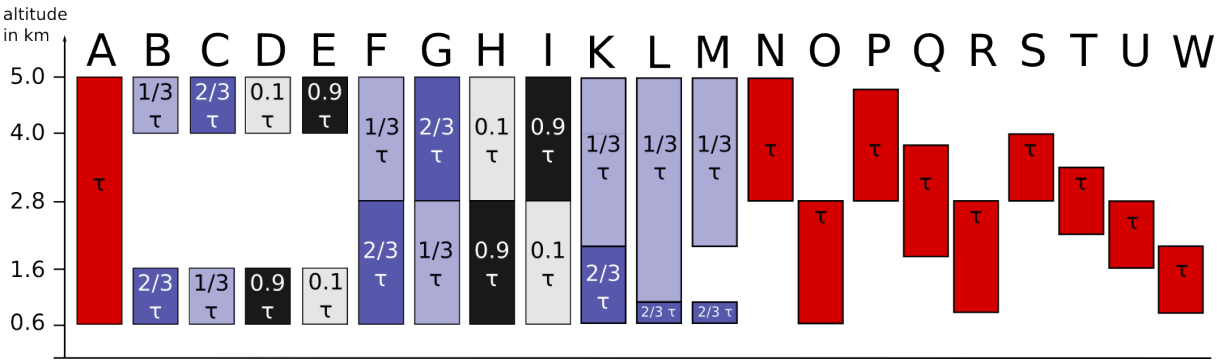

Figure 83: Visualization of the simulated cloud scenarios. The fraction of τ shows which percentage of the total cloud optical depth is attributed to the cloud layers.

All model clouds are schematically presented in figure 83 with the corresponding cloud parameters summarized in table 26. The common upper and lower cloud system boundaries for cases A-M were chosen by simulating a small number of one-layer clouds with a top altitude between 4.8 and 5.2 km, a vertical extent between 4.4 and 5.2 km and individually adapted cloud optical depths. This altitude range of the cloud bottom and top was estimated by taking the lidar and radar data into account. The best comparison to the TANSO-FTS measurement, which allows to create symmetric cloud structures on a 0.2 km model altitude grid, leads to the cloud boundary altitudes of cloud scenario A. Scenario A extends from 0.6 to 5.0 km with a constant cloud optical depth $\tau_{cloud} = 31.97$, which is close to the cloud optical depth from the TANSO-CAI cloud property ($\tau_{CAI} = 34.78$).

The first set of simulations (cloud scenarios B-I) includes spatially symmetric two-layer clouds, that are either composed of two layers of 1 km extent with a cloud-free gap of 2.4 km in between or of two layers of 2.2 km extent which are located on top of each other. The fractional distribution of τ_{cloud} is $\frac{1}{3} : \frac{2}{3}$ and $\frac{1}{10} : \frac{9}{10}$ and vice versa. This choice is based on the knowledge that the collocated radar data identifies the more extinctive structures in the bottom of the cloud. Contrasting structures with higher extinction in the upper cloud layer are also used

Table 26: All model clouds implemented for this case study including the fit results. The clouds have 1-2 homogeneous layers of different fractions of the total cloud optical depth τ_{cloud}.

ID	h_1 (km)	e_1 (km)	τ_1	h_2 (km)	e_2 (km)	τ_2	B_{fit}	RMS ($W/m^2/sr/nm$)
A	5.0	4.4	31.97	-	-	-	1.305	0.0043
B	1.6	1.0	21.0	5.0	1.0	10.5	1.215	0.0028
C	1.6	1.0	10.53	5.0	1.0	21.07	1.611	0.0056
D	1.6	1.0	28.8	5.0	1.0	3.2	0.896	0.0021
E	1.6	1.0	3.16	5.0	1.0	28.44	1.825	0.0075
F	2.8	2.2	21.07	5.0	2.2	10.53	1.161	0.0035
G	2.8	2.2	10.57	5.0	2.2	21.13	1.425	0.0051
H	2.8	2.2	28.53	5.0	2.2	3.17	0.98	0.0033
I	2.8	2.2	3.17	5.0	2.2	28.53	1.576	0.0063
K	2.0	1.4	21.07	5.0	3.0	10.53	1.042	0.0028
L	1.0	0.4	21.13	5.0	4.0	10.55	0.923	0.0021
M	1.0	0.4	21.0	5.0	3.0	10.5	0.991	0.0021
N	5.0	2.2	32.52	-	-	-	1.654	0.0075
O	2.8	2.2	32.0	-	-	-	1.028	0.0044
P	4.8	2.0	31.7	-	-	-	1.604	0.0069
Q	3.8	2.0	31.67	-	-	-	1.291	0.0053
R	2.8	2.0	31.95	-	-	-	1.05	0.0046
S	4.0	1.2	31.85	-	-	-	1.494	0.0067
T	3.4	1.2	31.94	-	-	-	1.311	0.0061
U	2.8	1.2	32.7	-	-	-	1.162	0.0057
W	1.8	1.2	32.28	-	-	-	0.933	0.0046

to probe the importance of the cloud optical depth distribution within these simple two-layer systems.

The next set of cloud configurations is composed of scenarios K-M. The chosen geometry is more closely inspired by the altitude resolved cloud optical depth profile from CPR, which was measured in a small distance from the TANSO-FTS FOV. The dense regions of the cloud are located at an altitude of around 1 km while the upper layers show significantly less extinction. The periodic cloud optical depth structures around 1 km (see fig. 40 in section 7.6.3) are ignored at this stage because of the simplified model and because the detailed cloud structures are largely unknown inside the TANSO-FTS measurement region.

The last set of cloud scenarios includes different one layer systems. The cloud elements have vertical extents of 1.2, 2.0 and 2.2 km which are located at different altitudes for each scenario. This set of simulations serves to estimate if the measured spectrum can be reproduced by an even simpler one-layer cloud configuration although the measurement indicates cloud regions of different extinction strengths.

Before the radiative transfer simulations of the O_2 A-band spectra are started, the cloud optical depth τ_{cloud} is individually determined for each model cloud scenario. All resulting values are near 32. The microphysics of the model clouds is further described by a Henyey-Greenstein phase function with an asymmetry parameter of 0.85.

12.2 Spectrum fit and results

The fits of the simulated spectra to the TANSO-FTS measurement spectrum are performed by considering the necessary wavelength alignement and the free parameter B_{fit} which allows

corrections of the O_2 absorption from the model simulation. The fits are performed for the entire O_2 A-band as well as for the three subbands of different absorption regions.

The fitted spectra, using the entire wavelength range of the O_2 A-band as input, are shown in the upper plot of figure 84 (grey), together with the TANSO-FTS measurement spectrum (red) and the fitted spectrum with the smallest resulting RMS (green). The residuals are also grey for all fits except the one with the smallest fit RMS (blue). The spectrum with the lowest RMS value is scenario L, showing a 0.4 km - layer of higher extinction at the cloud bottom and a 4 km - layer of lower extinction on top. The vertical profile of the cloud scenario corresponds well to a simplified description of the CPR radar measurement.

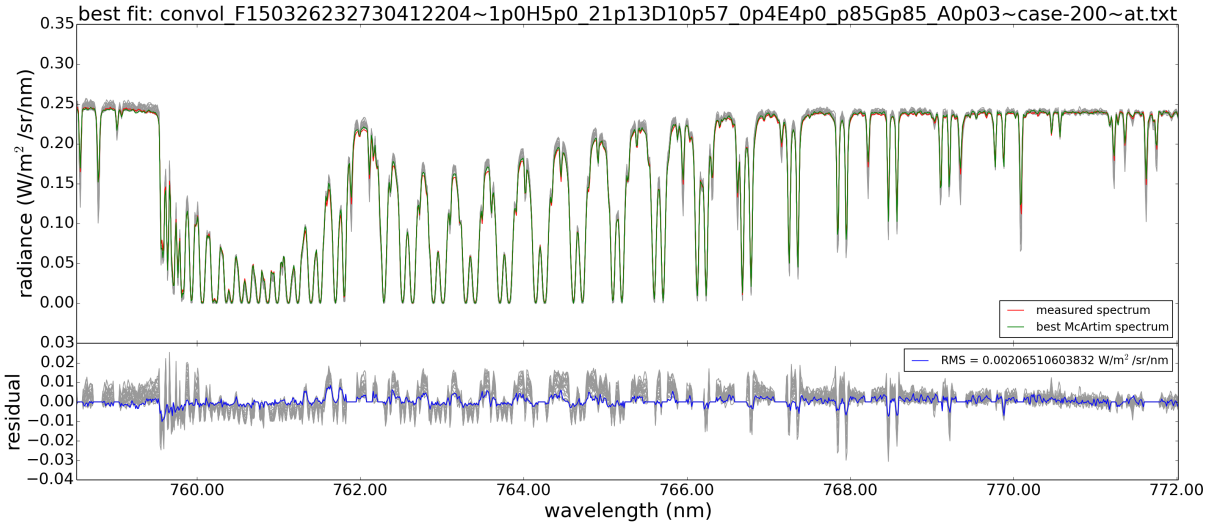

Figure 84: Top: Fitted spectra (grey), the TANSO-FTS measurement spectrum (red) and the fitted spectrum with the lowest RMS (green). Bottom: The corresponding residuals (grey) and those with the smallest RMS (blue).

The resulting RMS values and B_{fit} parameters for each cloud scenario are shown in figure 85. The cloud scenarios resulting in the smallest RMS are L, M and D (see table 26 for the fit results).

Besides the alphabetic labeling of the cloud scenarios, the following markers are attributed to the main categories of cloud optical depth distributions: The red markers indicate homogeneous one-layer clouds with τ_{cloud}, cloud models composed of two layers with the higher extinction in the lower layer are marked by ($\frac{2}{3}\tau_{cloud}$) or by ($\frac{9}{10}\tau_{cloud}$). In scenarios where the higher extinction is attributed to the upper cloud layer, the markers are ($\frac{2}{3}\tau_{cloud}$) or ($\frac{9}{10}\tau_{cloud}$).

In general, the scenarios with the higher extinction in the lower cloud layer, have smaller fit RMS and B_{fit} values closer to unity. For model clouds where the higher extinction is mainly attributed to the upper layers, high RMS and large deviations from B_{fit}=1 result. While the one-layer clouds show B_{fit} values near and far from unity, the RMS is generally larger than for the two layer scenarios with the larger extinction in the bottom cloud layer. This shows the importance of a multi-layer description of the cloud model for the fit outcome. The number of possible best matches can thus be reduced by only considering low RMS values. For this case study, one-layer clouds and two-layer clouds with high extinction in the upper layer could already be discarded at this stage.

The above-mentioned scenario L shows the smallest RMS, followed by the scenarios M, D, K and B. All of these scenarios have the largest extinction in the bottom cloud layer and a layer with less extinction on top, some of them with a cloud-free gap in between. The B_{fit} value

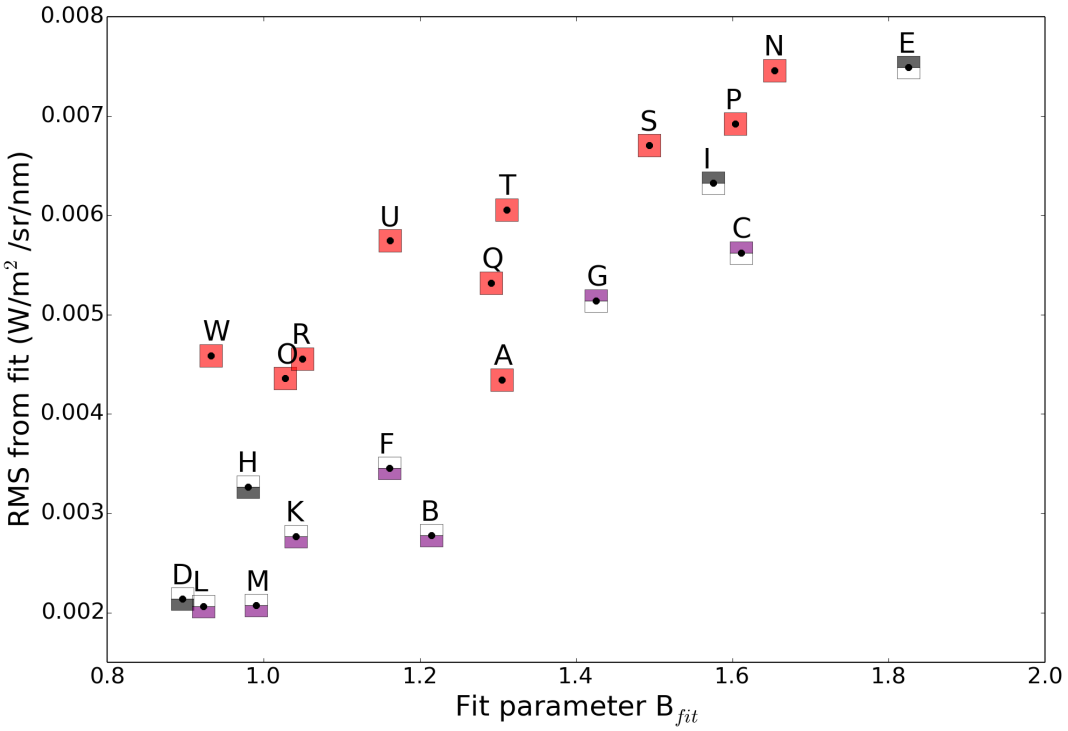

Figure 85: The resulting fit RMS in dependence of the fit parameter B_{fit} for all simulated cloud scenarios. $B_{fit} < 1$ indicates an overestimation of the O_2 absorption in the model simulation, $B_{fit} > 1$ indicates an underestimation. The meaning of the different markers is described in the text.

corresponding to scenario L (lowest RMS) is close to 0.9, which means that the absorption of O_2 in the simulation is overestimated. This is similar to the case studies with one-layer clouds. While this applies to cloud model D as well, the simulation with cloud model M results in B_{fit} near unity, also with a low RMS. Here, it should be noted that the difference between the RMS of these three scenarios are all within ΔRMS. It is also interesting to note that the RMS differences of all simulated cloud scenarios is about one magnitude higher than for the one-layer cloud case studies. This facilitates the interpretation of the fit results. The reason is most probably the choice of simulated cloud models, which are very similar to each other for the one-layer cases and rather different for the two-layer case.

A boxplot representation of the photon path length distributions for each fitted radiance spectrum is shown in fig. 86 (top), together with the mean of the distributions $\overline{PPL}$ and the median (bottom). The O_2 weighted distributions are shown in Appendix (section 12.5), fig. 92. The median is systematically lower than $\overline{PPL}$, indicating that an increased number of short photon paths exists besides fewer long path lengths.

The cloud scenarios with low fit RMS can be divided into three groups: The cloud scenarios M, K and H have B_{fit} values near unity. Scenarios D and L show values of B_{fit} significantly lower and scenarios B and F significantly higher than unity. In the following, the groups will be referred to as "$B_{fit} \approx 1$", "$B_{fit} < 1$" and "$B_{fit} > 1$", respectively. The average path lengths $\overline{PPL}$ of the simulated photons are similar within each group.

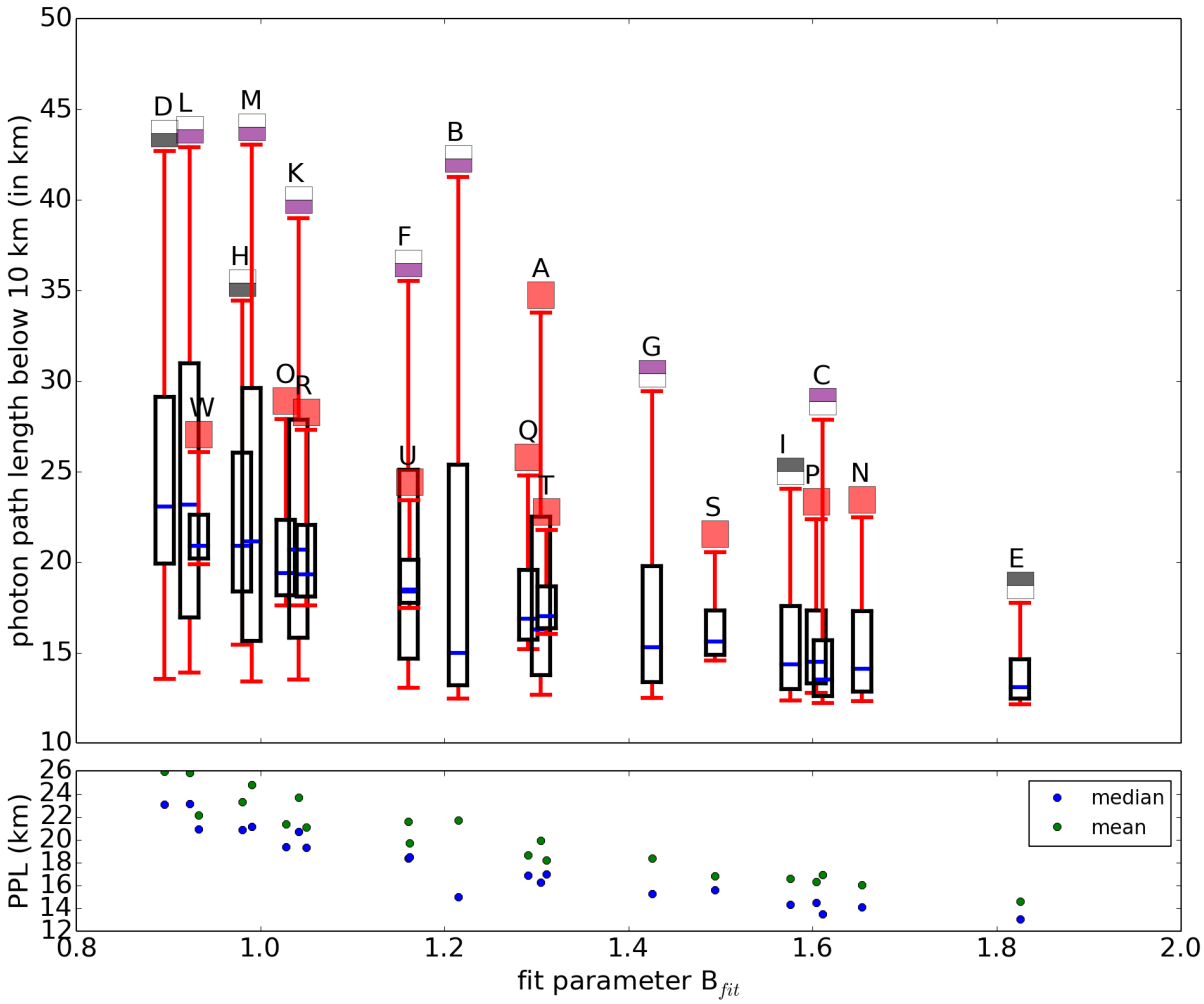

Figure 86: Top: Photon path length distributions for all cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

12.2.1 Comparison of photon path length distributions ($B_{fit} \approx 1$ vs. $B_{fit} < 1$)

The geometric photon path length distributions below 10 km altitude for the mentioned groups are shown in fig. 87, focusing on path lengths with values below 45 km. The plot shows the comparison of the cloud models where the O_2 absorption is clearly overestimated (D, L) and those models, where the absorption is correctly estimated (M, K, H).

First, those cloud scenarios with B_{fit} near unity are considered. Cloud scenarios M (blue) and K (green) show very similar photon path length distributions with a maximum near 15 km but for scenario M, the fit RMS and the deviation of B_{fit} from unity is lower than for scenario K. Both cloud scenarios only differ in the vertical extent of the lower layer, which is smaller for scenario M but leads to a cloud-free gap in between both cloud layers. Case H (red) also shows a B_{fit} value near unity but the fit RMS is higher than for the other two scenarios. The photon path length distribution of scenario H, where the extinction coefficient of the lower layer is much higher than for scenarios M and K, differs significantly from M and K. Here, the largest peak is found near a higher photon path length because a larger quantity of the photons arrives at lower altitudes.

All plots show a superposition of two distributions which results from the two distinct cloud layers for each scenario. For scenarios M and K, a superposition of distributions can be identified

with maximums near 15 km and 23 km. Case H shows are more pronounced superposition of distributions, including a smoother distribution starting near 12 km path length and thus corresponding to the upper cloud layer with low extinction and a sharper distribution emerging near 17 km, corresponding to the lower cloud layer with strong extinction. Although scenario H has a very different photon path length distribution compared to scenarios M and K, the average photon path length $\overline{PPL}$ is similar.

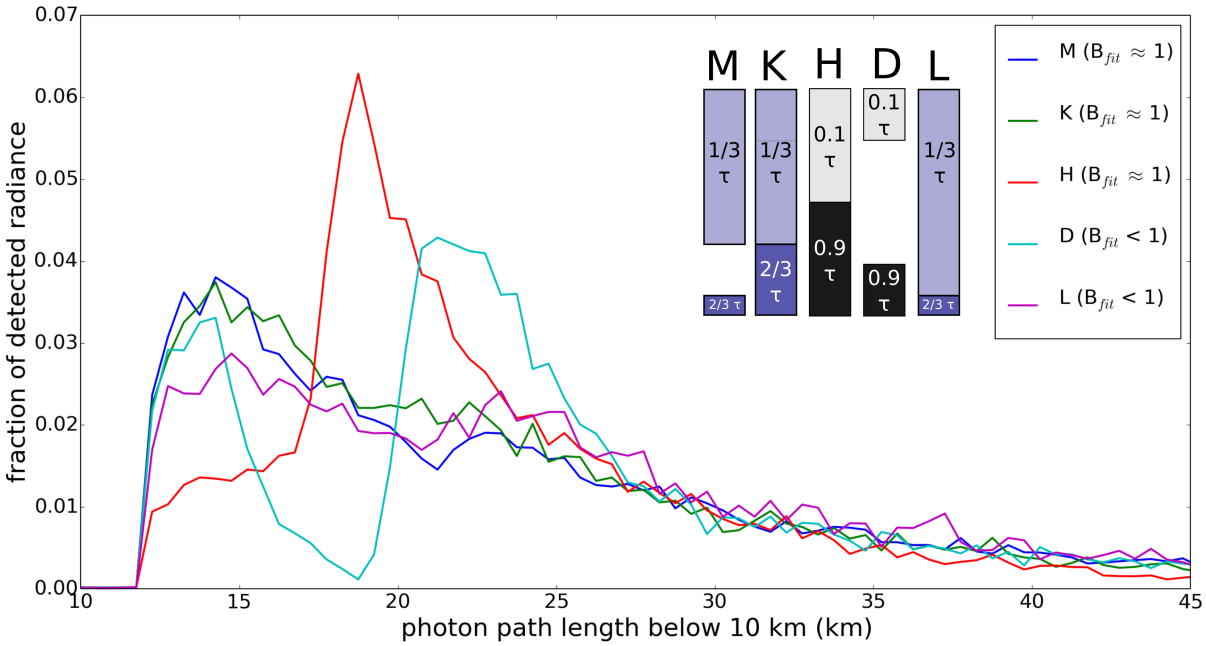

Figure 87: Comparison of the photon path length distributions below 10 km for simulations resulting in a B_{fit} near unity (well-assumed O_2 absorption) and B_{fit} value < 1 (overestimated O_2 absorption). The O_2 weighted distributions are shown in Appendix (section 12.5), fig. 93.

Now, simulations for cloud models that overestimate the O_2 absorption during the simulation ("$B_{fit} < 1$") are considered. At the first glance, the results of cloud scenario L (magenta) - which has an excellent fit RMS but shows a clear overestimation of the simulated O_2 absorption - has a very similar photon path length distribution to the cloud scenarios M and K where the O_2 absorption is well met. The spatial cloud configuration is indeed very similar to scenario M, with a larger vertical extent of the upper cloud but no gap between both cloud layers. Case L also shows a superposition of distributions, but the maximum of the shorter photon path lengths is lower and the maximum for the longer photon path lengths is a bit higher than for the scenarios M and K. This contributes to a slightly higher $\overline{PPL}$, as indicated in fig. 86.

Cloud scenario D (cyan) consists of two cloud layers of one kilometer vertical extent each. It has a low fit RMS and the photon path length distributions are slightly overlapping. The pronounced peaks correspond to the upper cloud layer of low extinction (left distribution) and the lower cloud layer of high extinction (right distribution). Due to the cloud-free gap, almost all photons emerging from the upper layer will also arrive at the lower layer. The result is a collection of very different photon path length values, which leads to the almost complete separation of the underlying distributions. In the model, $\frac{9}{10}$ of the cloud's extinction is attributed to the lower layer. B_{fit} is significantly lower than unity. To overcome the overestimation of the O_2 absorption, photons would need to travel shorter paths or probe altitudes with lower O_2 absorption more extensively. This, for example, could be achieved by attributing a larger fraction of the cloud optical depth to the upper layer or moving the lower cloud layer upwards.

Regarding once again cloud scenario H (red), it can be seen that the fit RMS is rather high but

the O_2 absorption seems to fit the measurement. In comparison to scenario D (cyan), which leads to an overestimation of the O_2 absorption, the spatial cloud dimensions are extended such that the cloud layers touch at an intermediate altitude. Here, the spatial distribution of the cloud optical depth causes the photons to be backscattered at higher altitudes than in scenario D, which leads to shorter photon paths in average and less absorption of O_2.

12.2.2 Comparison of photon path length distributions ($B_{fit} \approx 1$ vs. $B_{fit} > 1$)

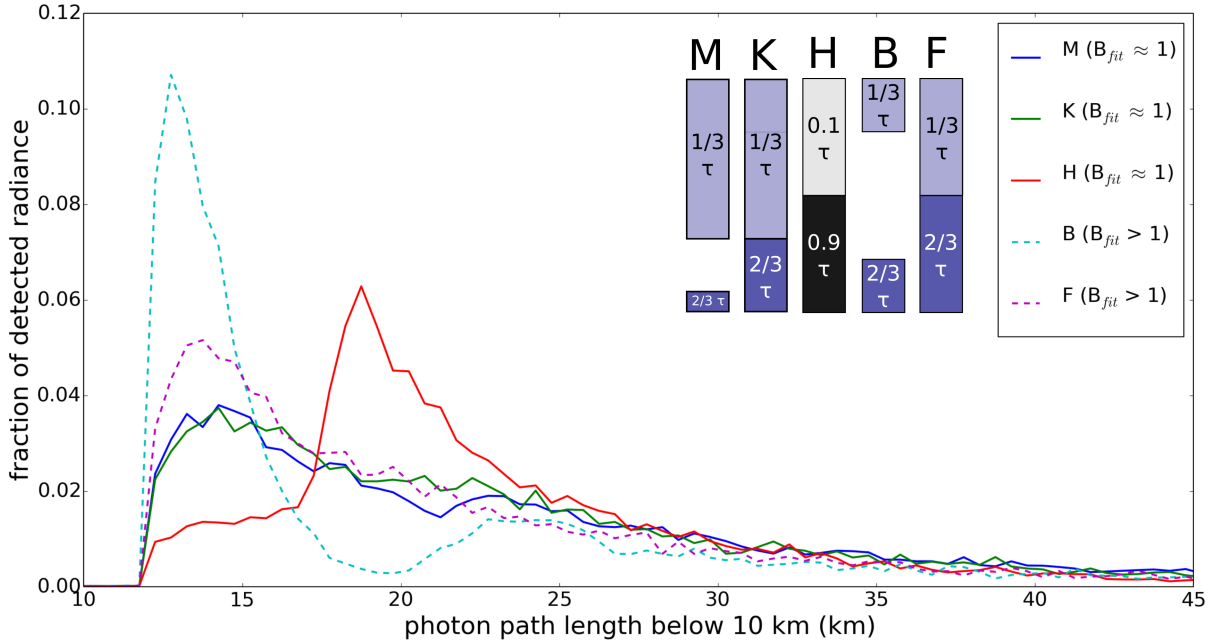

Figure 88: Comparison of photon path length distributions below 10 km altitude for simulations resulting in B_{fit} near unity and spectra of $B_{fit} > 1$ (underestimation of the O_2 absorption). The O_2 weighted distributions are shown in Appendix (section 12.5), fig. 94.

Figure 88 shows a comparison of the photon path length distributions between those simulations with resulting $B_{fit} \approx 1$ (cloud scenarios M, K, H) and the two simulations with the smallest fit RMS of $B_{fit} > 1$ (cloud scenarios B, F).

First, the two cloud scenarios are considered, where the O_2 absorption in the simulation is underestimated but the fit RMS is still moderate (clouds B and F). They are both constructed by two spatially symmetric layers with attributed $\frac{1}{3}\tau_{cloud}$ in the upper, and $\frac{2}{3}\tau_{cloud}$ in the lower cloud layer.

The average photon path length distribution $\overline{PPL}$ is nearly identical for both cloud scenarios. For cloud scenario B, a clear dominance of shorter photon paths can seen in figure 88, which is also indicated by the different median values (compare fig. 86). Similar to the comparison of cloud scenarios D and H (see fig. 87), the only difference between clouds B (cyan in fig. 88) and F (magenta) lies in the vertical extent of the two layers. It is 1 km for scenario B with a 2.4 km gap in between and 2.2 km for scenario F with no cloud-free gap. The photon path length distribution for scenario B shows that the photons travel extensively within the upper cloud layer, leading to a maximum near 14 km. A superpositioned distribution with a peak around 24 km shows the path lengths of those photons also much probing the lower cloud layer.

The photon path length distribution of cloud F (magenta) is similar to cloud M (blue) and K

(green). These three clouds mainly differ in the vertical extent of the lower cloud layer. The main peak near 15 km is larger for scenario F. This difference can be attributed to the smaller vertical extent of the upper cloud layer where the smaller fraction of the cloud extinction is defined. Here, the cloud is denser than in the other two scenarios which increases the probability for shorter photon path lengths.

When examining the boxplots of the photon path length distributions in fig. 86, it is interesting to note that the cloud scenarios F and U have very similar median values of the photon path length distributions although the spatial configuration of the cloud models is very different. Case F features a cloud of 4.4 km vertical extent, separated in a lower cloud layer of high extinction and an upper cloud layer of low extinction, both having the same vertical dimensions. Case U, on the other hand, is a one-layer cloud situated at an intermediate altitude. Although the medians of the photon path length distributions are similar, the fit RMS of scenario U is much worse. The main difference is the width of the photon path length distribution which is very small compared to scenario F. Here, the multiple scattering region of the photons is restricted to a 1.2 km altitude range where the main enlargement of the photon path lengths takes place. This leads to a smaller variety of different photon path lengths. Also, the cloud droplet concentration is relatively high, improving the probability that the photons are backscattered after shorter path lengths inside the cloud. Implementing cloud F, on the other hand, allows multiple scattering over a 4.4 km vertical extent where not only regions of different O_2 absorption exist but a smaller density of the cloud droplets is given.

12.2.3 Best results and photon penetration altitudes

As already discussed in the context of the previous case studies, two strategies for the interpretation of the fit results can be applied. The first attempt concentrates solely on the lowest fit RMS while the second one takes the value of B_{fit} into account.

If only the fit RMS is considered, cloud scenarios L, M and D show the best results. The difference of the RMS is below 10^{-4} W/m^2/sr/nm (the derived uncertainty from the simulation noise is 10^{-4} W/m^2/sr/nm, see also section 10.7.4). The construction of the cloud scenarios is as follows:

- scenario L: $\frac{2}{3}\tau_{cloud}$ between 0.6 and 1.0 km, $\frac{1}{3}\tau_{cloud}$ between 1.0 and 5.0 km
- scenario M: $\frac{2}{3}\tau_{cloud}$ between 0.6 and 1.0 km, $\frac{1}{3}\tau_{cloud}$ between 2.0 and 5.0 km
- scenario D: $\frac{9}{10}\tau_{cloud}$ between 0.6 and 1.6 km, $\frac{1}{10}\tau_{cloud}$ between 4.0 and 5.0 km

The above scenarios have in common that the lower cloud layer is relatively thin (0.4 or 1.0 km) and that the main part of the cloud optical depth is attributed to this layer ($\frac{2}{3}\tau_{cloud}$ or $\frac{9}{10}\tau_{cloud}$). The corresponding photon path length distributions are all shown within fig. 87. The distributions of scenarios L and M are not very different as is the spatial construction of the cloud scenarios. The main difference between the cloud scenarios is a cloud-free gap of 1 km between the layers for cloud scenario M.

Figure 89 shows how deep the simulated photons penetrate into the atmosphere. For each penetration altitude, the fraction of measured radiance is given. The two cloud layers of scenario M can be clearly identified as many photons are backscattered into the detector direction due to the multiple scattering within the layers. In cloud- and surface-free penetration altitudes, the photon fractions are negligible because only few scattering events on air molecules take place. Although not easy to identify in the visualization, the photon fractions in these regions show small positive values. The penetration altitudes within the lower cloud layers are nearly identical for scenarios L and M. This is to be expected because of the same cloud optical depth above and within the lower layer.

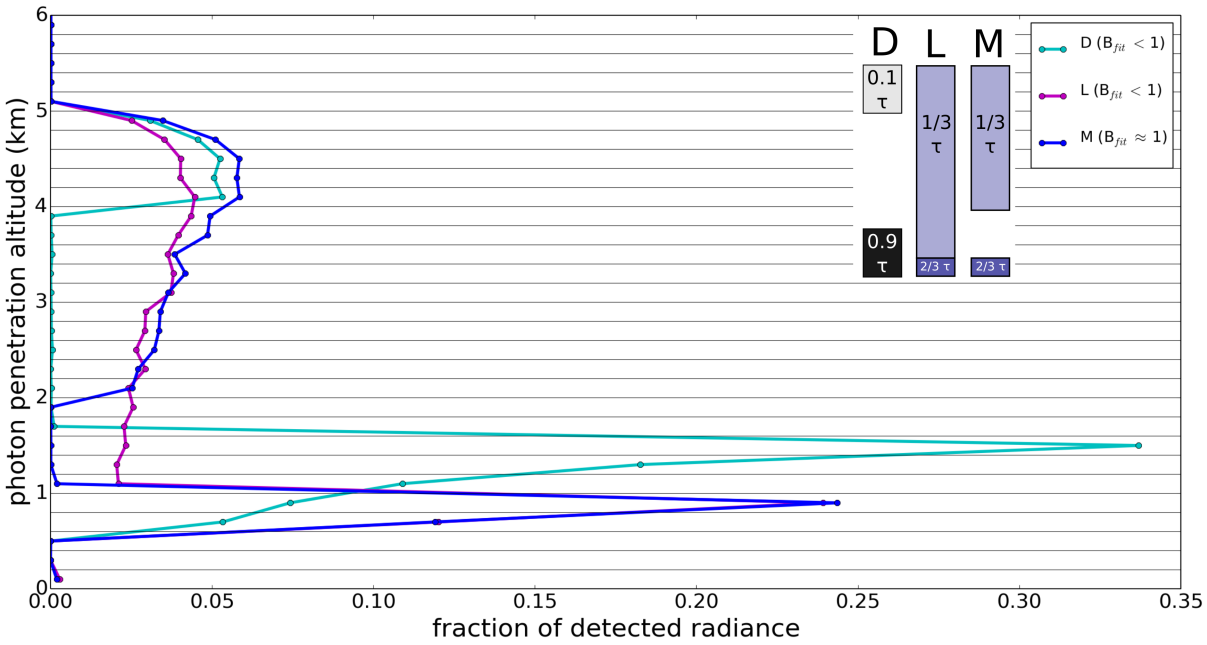

Figure 89: Penetration altitudes reached by the detected photons (scenarios D, L, M).

For scenario D, a significant amount of photons has penetration altitudes in the upper cloud layer and the main contribution in the lower cloud layer. This agrees to the distribution of τ_{cloud} within the cloud system. The amount of photons which reaches the ground is nearly identical for all three cloud scenarios because the similar total optical depth of the entire cloud system.

The results are now compared to the collocated cloud profiles from CALIOP and CPR. The average cloud top height, derived from CALIOP lidar measurements, is 4.98 km and the radar measurement from CPR finds cloud structures in radar bins of 0.45-0.68 km and above. For the model clouds, the cloud top and bottom altitudes were chosen based on the results of the collocated measurements. The findings confirm this choice. In future studies, slight variations of these values should be investigated as well.

The total cloud optical depth of the model cloud is $31.5 \leq \tau_{cloud} \leq 32.0$. The measured mean cloud optical depth by TANSO-CAI within the TANSO-FTS FOV is 34.78, which is very close to our retrieved τ_{cloud} for the model cloud. The CloudSat CPR finds a smaller τ_{cloud} of 23.18 but it has to be considered that this value is retrieved from only one measurement which is located some kilometers besides the TANSO-FTS FOV. Comparison with the altitude resolved cloud optical depth from the CloudSat CPR (see section 7.6.3 fig. 40) also shows that the cloud optical depth is measured at coordinates where the cloud structures with high extinction in the lower altitudes are smaller than in the close environment.

Using the second interpretation attempt, the result depends on both the RMS and B_{fit}. While scenario M results in B_{fit} near unity, scenario L ($B_{fit} \approx 0.92$) and scenario D ($B_{fit} \approx 0.9$) both show an overestimation of the simulated O_2 absorption similar to the case studies with one-layer clouds. Here, it would be concluded that scenario M is the best fit result which shows many similarities with the collocated data of the cloud characteristics. However, taking the previous case studies into account, both scenarios L and D would be possible choices as well.

In summary, the chosen cloud scenarios are all in the possible range of cloud properties when the collocated cloud profile measurements are taken into account. Depending on the focus of the fit parameters on the interpretation, either a collection of cloud scenarios (L, M, D) results when only using the RMS or one cloud scenario (M) when also the parameter B_{fit} is considered.

12.3 Fit results in sub regions of the O_2 A-band

So far, only fit results of the entire O_2 A-band have been considered. Here, the fit is applied to the radiances of smaller wavelength regions within the O_2 A-band. The dependence of RMS and B_{fit} is shown in the plots of figure 91 in the Appendix (top: strong O_2 absorption, middle: medium O_2 absorption, bottom: weak O_2 absorption). The cloud scenarios with low RMS are the same as before, except for the wavelength region of very low O_2 absorption where scenario H would need to be included as well.

Regarding mainly B_{fit}, the results for the strong and medium O_2 absorption also indicate that scenario M is the best match to our measurement. For the weak absorption, scenario M has the lowest RMS but a value of $B_{fit} < 1$. It follows that the O_2 absorption is overestimated for long photon path lengths. This can concern photons undergoing much multiple scattering also within in the lower cloud layer. It gives an indication, that a modification of the lower cloud layer could lead to an optimization of the cloud model.

12.4 Altitude resolved photon path length distributions

In the following, altitude resolved photon path length distributions are shown for cloud scenario M. This scenario was chosen as representative example because it has both a B_{fit} near unity and a small RMS. The photon path length distributions within individual altitude layers have the potential to provide additional information about the influence of the cloud structure on the photon trajectories.

The photon path length distributions within selected altitude layers (with a vertical extent of 200 m each) of the model atmosphere are shown in fig. 90 (black: outside cloud layers, red: within the upper cloud, blue: within the lower cloud) and include the following atmospheric heights:

- 5.0 - 5.2 km : first altitude layer above the top cloud
- 4.8 - 5.0 km : highest altitude layer of the top cloud
- 4.0 - 4.2 km : intermediate altitude layer of the top cloud
- 2.0 - 2.2 km : lowest altitude layer of the top cloud
- 1.8 - 2.0 km : first altitude layer below the top cloud
- 0.8 - 1.0 km : highest altitude layer of the bottom cloud
- 0.6 - 0.8 km : lowest altitude layer of the bottom cloud
- 0.4 - 0.6 km : first altitude layer below the bottom cloud

Above the upper cloud layer, almost 100% of the detected photons have simulated path lengths of almost 0.5 km. This corresponds to the geometric passage through the 0.2 km - altitude layer, considering the direction of the solar source and the satellite detector. Fractions of the detected radiance which are larger than zero are also present for smaller and larger path lengths due to scattering events on air molecules (to small to be visible in this visualization).

The upper cloud layer extends from 2.0 to 5.0 km altitude and has an optical depth of 10.5. Given the large vertical extent, the cloud optical depth within each altitude layer is small. When entering the upper cloud layer, the peak at about 5.0 km is reduced to $\approx 60\%$ of the measured photons but the slopes of the distribution become more pronounced. This indicates that many photons do not travel the entire altitude range twice but are deviated and backscattered due to multiple scattering events inside the cloud. When they are scattered into the detector direction

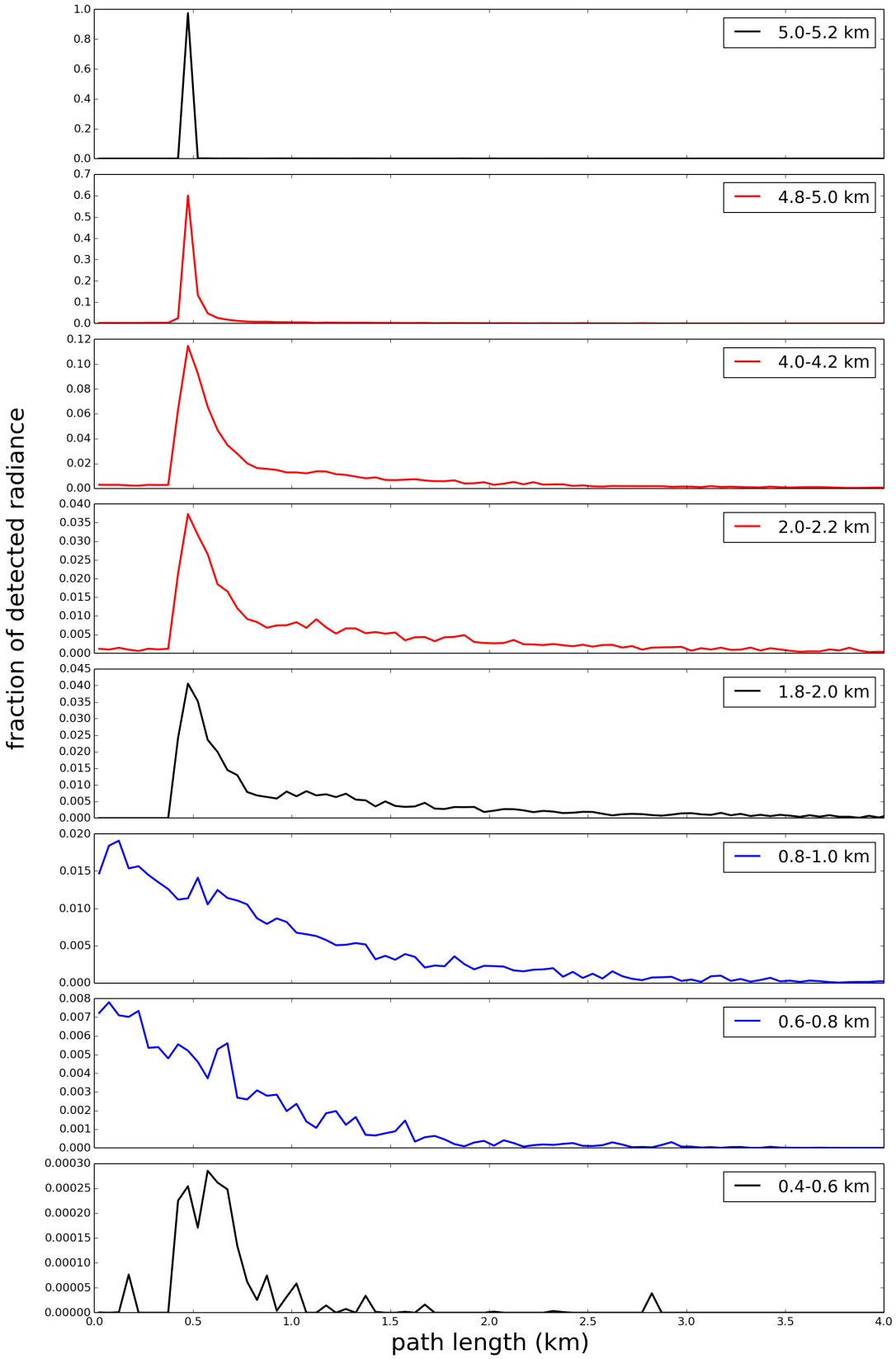

Figure 90: Photon path length distributions of scenario M for selected altitude layers. The distributions within cloud-free regions are shown in black. All distributions in the upper cloud layer are plotted in red and those in the lower cloud layer in blue.

shortly after entering the upper layer, it increases the detected radiance fraction for shorter paths. Some path lengths are also enlarged due to multiple scattering processes within the

cloud. Considering altitude layers deeper inside upper cloud, the fraction of measured photons associated with the major peak is further reduced (to 11%-12% within 4.0 - 4.2 km and below 4% within the lowest layer of the top cloud) but longer path lengths become increasingly important.

The photon path length distributions of the lowest altitude of the top cloud and the layer beneath (cloud gap) are very similar but significant fractions of very short photon paths are only present within the top cloud. This is to be expected because no cloud droplets allowing for multiple scattering processes are present in the cloud gap. Here, only scattering on air molecules and possibly multiple reflections between the two cloud layers occur. The similar slope expressing longer path lengths can be explained as follows: Those photons which emerge from the top cloud, will (approximately) follow the current direction in space and thus have a similar path length than in the superpositioned altitude layer. The added path which describes the way from the scattering position to the solar source is similar as well because no additional multiple scattering enlargement of the photon paths is considered for the calculated distance to the Sun (see section 6.6).

The following subplots (0.8 - 1.0 km and 0.6 - 0.8 km) show the photon path length distributions within the bottom cloud. The vertical extent is very small (0.4 km) but $\tau_1 = 21.0$, which leads to a large amount of multiple scattering. The density of the cloud manifests itself in the significant decrease of measured photons between the upper and lower layer of the bottom cloud and the shape of the photon path distribution. The distributions show very short photon paths, leading to the conclusion that the penetration of the photons inside the layers can be very small. Also, longer path lengths are observed more frequently. Overall, a smooth distribution of photon path lengths is obtained.

The lowest subplot shows the photon path length distribution in the layer beneath the lower cloud. Only few photons emerge downwards from the dense cloud structure above. A small peak can be observed around 0.6 km, corresponding to path lengths a little larger than those above the top cloud. A possible explanation is based on the direction of the emerging photons which can lead to different path lengths than the geometric assumption based on the sun-satellite geometry. This can also cause the longer photon paths, like the ones inside the cloud gap. Reflections between the surface and the lower cloud layer are not very probable due to the low albedo (0.03) of the water surface. Very short photon path lengths are caused by the rare scattering events on air molecules.

12.5 Appendix

12.5.1 Fit results in sub regions of the O_2 A-band

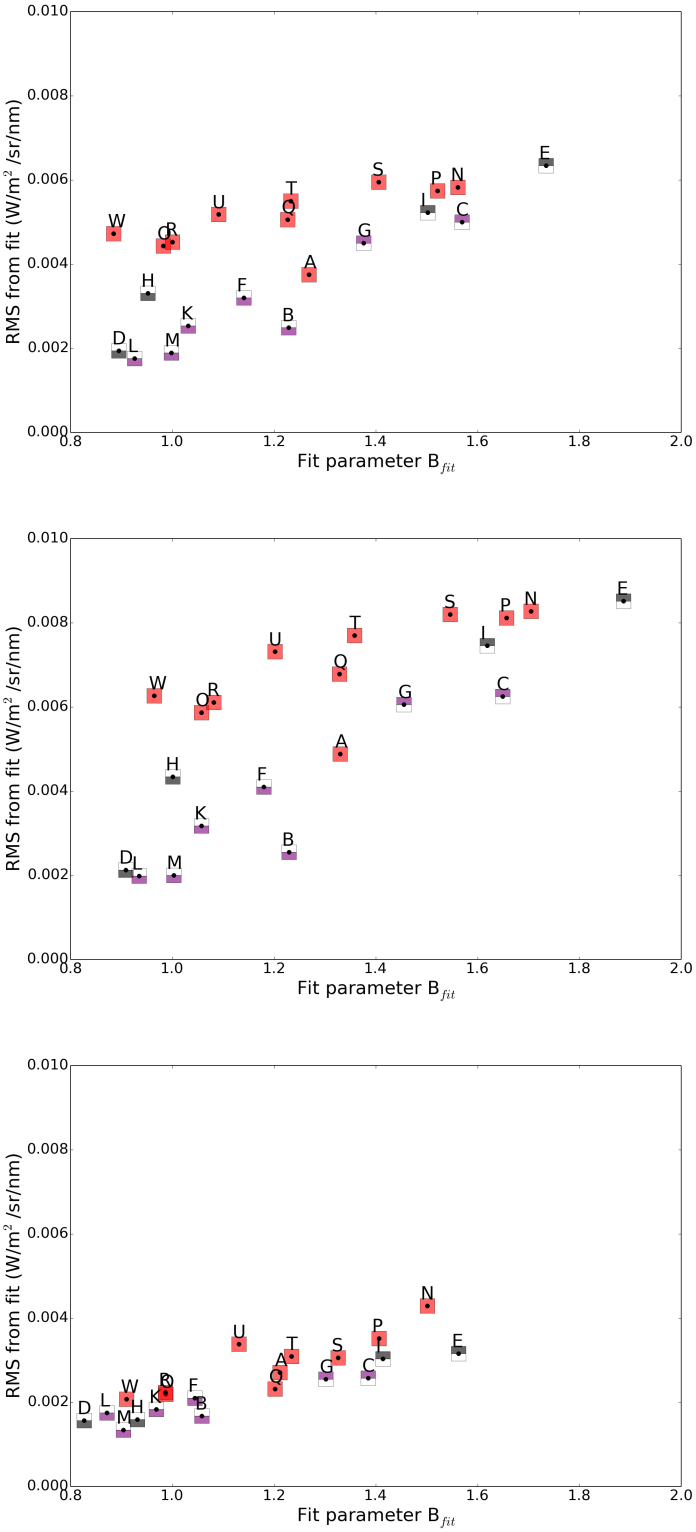

Figure 91: The resulting fit RMS in dependence of the optimized fit parameter B_{fit} for the three subbands of the O_2 A-band (top: strong absorption of O_2, middle: intermediate absorption of O_2, bottom: low absorption of O_2). $B_{fit} < 1$ indicates an overestimation of the O_2 absorption in the model simulation, $B_{fit} > 1$ indicates an underestimation.

12.5.2 O_2 weighted photon path length distributions

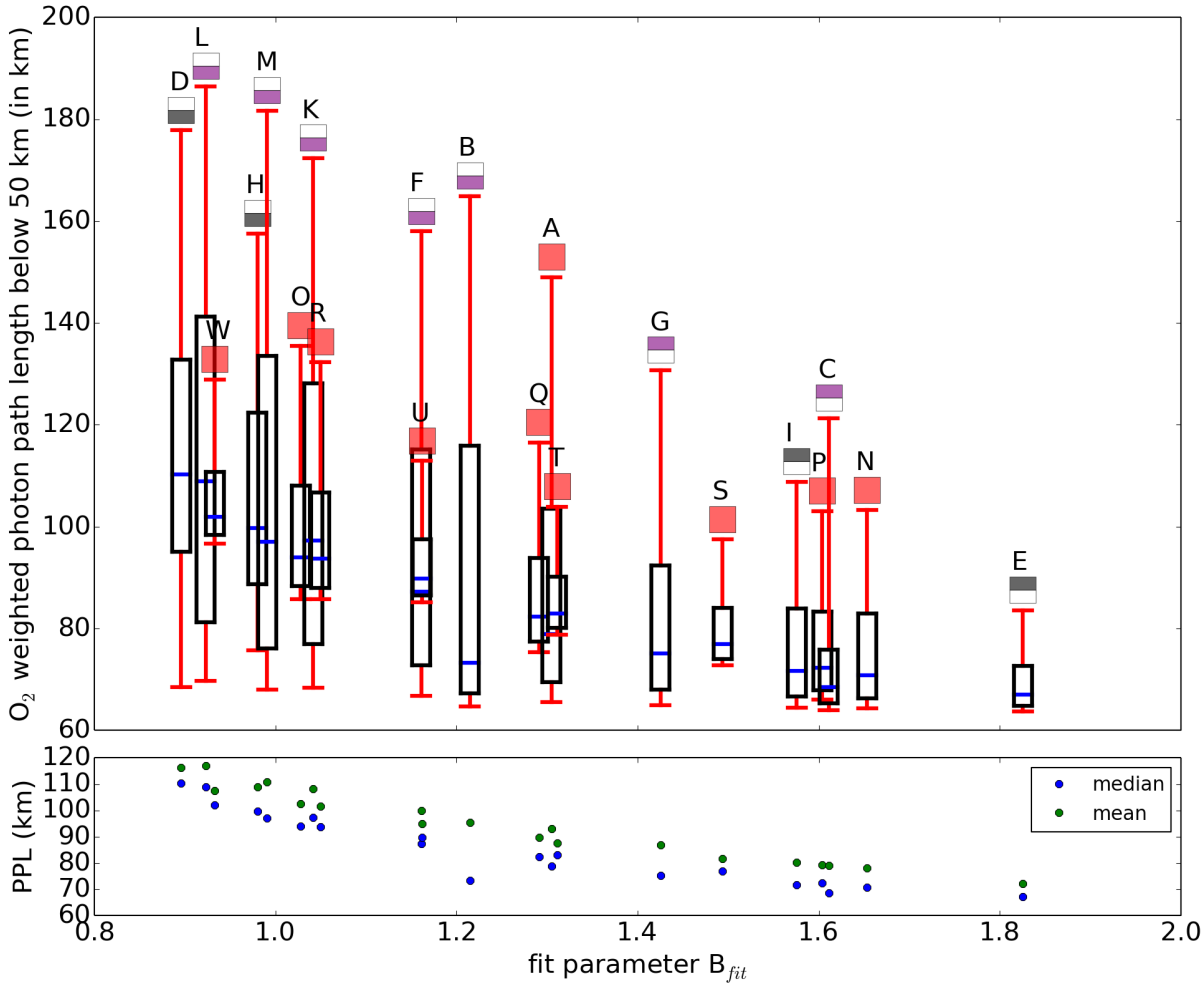

Figure 92: Top: O_2 weighted photon path length distributions for all cloud scenarios in boxplot representation (9^{th}, 25^{th}, 50^{th}, 75^{th} and 91^{th} percentiles) in dependence of the B_{fit} parameter. Bottom: Medians (blue) and means (green) of the photon path length distributions in dependence of B_{fit}.

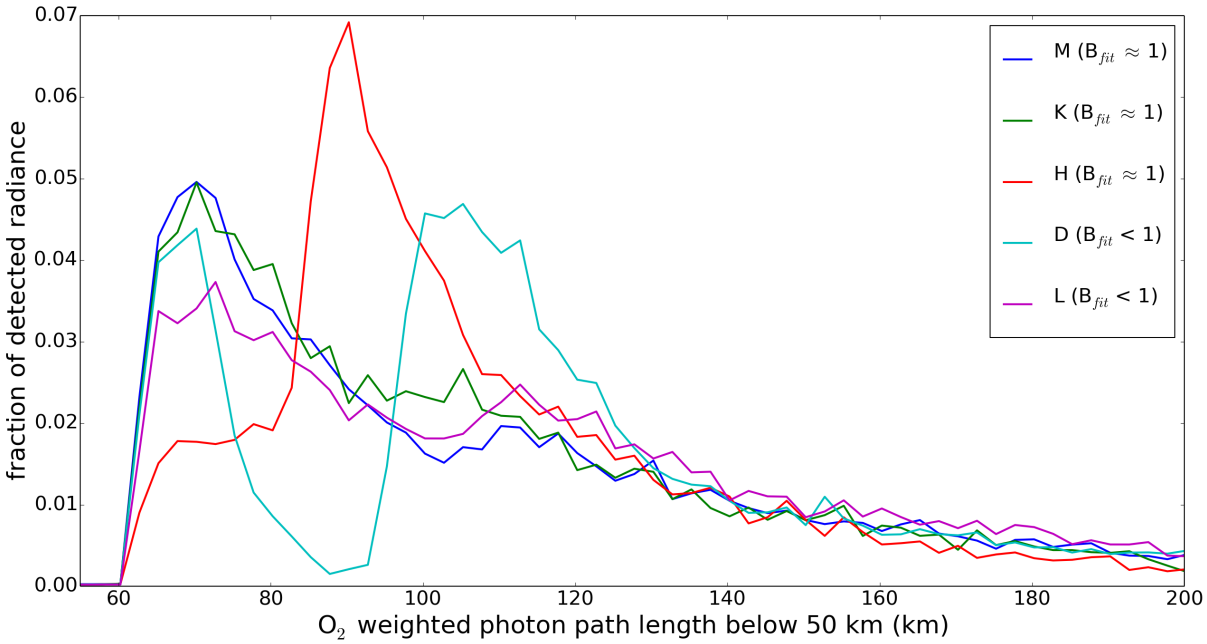

Figure 93: Comparison of the O_2 weighted photon path length distributions below 50 km for simulations resulting in a B_{fit} value near unity (well-assumed O_2 absorption) and $B_{fit} < 1$ (overestimated O_2 absorption).

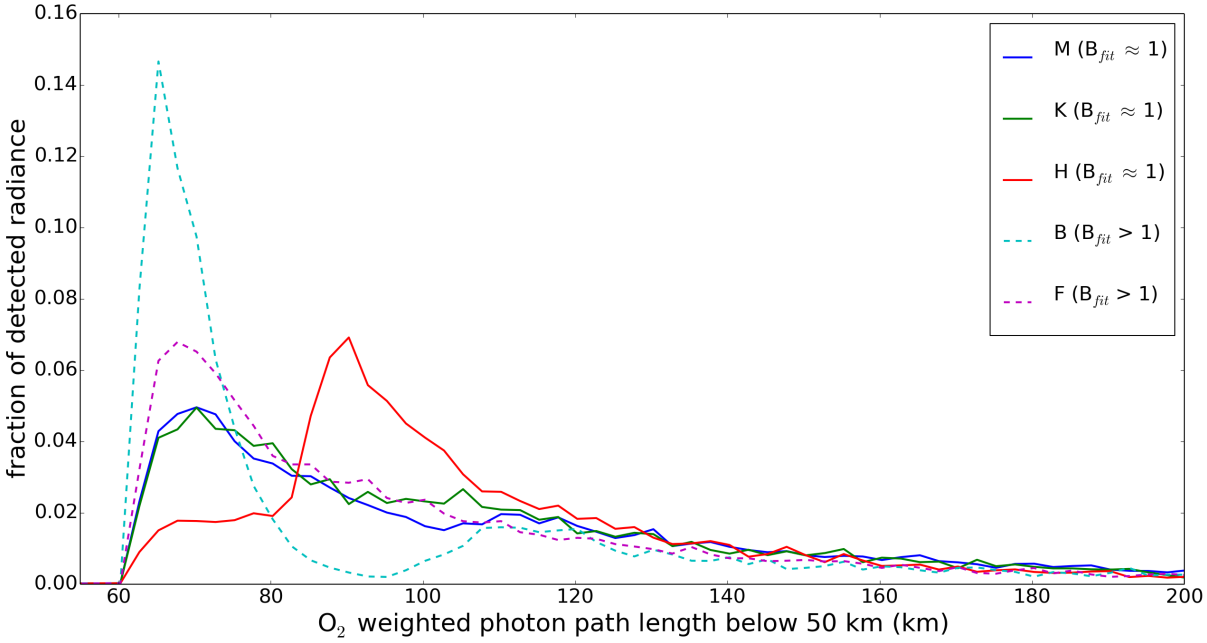

Figure 94: Comparison of O_2 weighted photon path length distributions below 50 km altitude for simulations resulting in B_{fit} near unity and spectra of $B_{fit} > 1$ (underestimation of the O_2 absorption).

13 Discussion and Outlook

Within this thesis, the basis of a forward retrieval for the derivation of cloud properties and photon path length statistics is developed, based on O_2 A-band measurements of high spectral resolution and radiative transfer simulations. It is applied on two clear-sky measurements and selected case studies which originate from a dataset of measurements with well-characterized cloud situations that is created by finding collocated measurements of GOSAT, CALIPSO and CloudSat instruments.

At the current stage, the water cloud scenarios, which are incorporated into the model atmosphere, are based on assumptions of the vertical cloud profile (horizontal cloud properties are assumed to be homogeneous). The number of simulated cloud scenarios depends on the considered cloud case and the assumed altitude region of the cloud top and the vertical extent. The microphysics of the clouds are described by a Henyey-Greenstein phase function with an asymmetry parameter of 0.85. This choice was found to be uncritical since very similar results were obtained by varying the asymmetry parameter. The cloud optical depth is derived by comparison of the measured and simulated radiances in the continuum region of the spectrum (very low O_2 absorption). The similarity of the simulations to the measurements is evaluated by a fit function which includes a wavelength alignment of the radiances (spectral shift and stretch) as well as a correction factor B_{fit} for the simulated O_2 absorption (multiplicative to the O_2 optical depth). For perfect measurement conditions and model descriptions, the smallest residual between the two spectra is expected at values of $B_{fit} \approx 1$. This was verified by using a synthetic measurement spectrum originating from radiative transfer simulations with added white noise.

In order to test the present stage of the retrieval, two TANSO-FTS measurements under cloud-free conditions are considered first. Then, selected case studies with well characterized cloud situations are investigated. Measurements with various well-defined cloud situations are identified by screening 5 years of TANSO-FTS measurement data for collocated measurements from lidar and radar observations of CALIOP and CPR, respectively. Additional information on the cloud scenario originates from the TANSO-CAI sensors and MODIS-Aqua. A complete characterization of a cloud scenario includes

- the pixel-resolved cloud optical depth from TANSO-CAI (τ_{CAI}), available for the entire FOV region of TANSO-FTS,
- the altitude resolved cloud optical depth on a 240 m vertical grid from CloudSat CPR (twice oversampled) as well as the total cloud optical depth τ_{CPR},
- information on the vertical extent of the cloud and the cloud and precipitation type from CloudSat CPR,
- the cloud top height h_{CALIOP} from CALIPSO CALIOP and
- the cloud top pressures inside the TANSO-FTS FOV from MODIS-Aqua.

Prior to the investigation of selected cloud case studies, two clear sky situations are considered. One measurement is taken over a dark surface (water), the second over a very bright surface (desert). Instead of τ_{cloud}, the surface albedo a_s is varied to adapt the radiance intensity in the continuum region of the O_2 A-band. For the dark-surface measurement, an unusually high albedo for water surfaces is found. Possible reasons include surface changes (due to e.g. strong winds) and remaining cloud structures. Since the clear sky case studies are not part of the collocated dataset, only information from the GOSAT measurements can be used to estimate the clear-sky status during the measurement (otherwise, lidar measurements from CALIOP could provide additional information). The subsequent fit on the complete radiance spectrum

shows an underestimation of the simulated O_2 absorption of about 6%. Possible reasons are also remaining cloud structures within the TANSO-FTS FOV, errors in the radiative transfer simulations and/or calibration effects of the measurement spectrum because many values below zero are found within the deep absorption structures.

The measurement of the second clear-sky case is taken over a desert region. Here, the derived surface albedo a_s is a bit lower than the value retrieved from the MODIS Filled Land Surface Albedo Product for the corresponding time of year. In contrast to the measurement over dark surface, an overestimation of the simulated O_2 absorption results from the fit, leading to a parameter B_{fit} around 0.94. Here, it should be noted that the P- and S-radiances of the TANSO-FTS measurement are almost equal which can be expected for measurements dominated by surface reflection. The discrepancy between both fit results and the deviations of B_{fit} from unity cannot be explained based on only two case studies. In order to systematically analyze the deviations from the expected simulated O_2 absorption, a larger number of clear sky cases would need to be considered. Ideally, collocated measurements from other satellite instruments would provide further information on possible cloud remains during the measurement. Also, fits to the measurement of simulated spectra with different surface albedos could provide further insights when the resulting B_{fit} is compared. As an example, the surface albedo for $B_{fit} \approx 1$ could be derived.

Following the clear-sky radiances, two one-layer cloud scenarios are chosen from the dataset of collocated (cloudy-sky) measurement situations. The choices were performed randomly on the basis that one measurement should be included in the rather large dataset group of low clouds and a measurement with a higher cloud is considered as well. Case study 1 features a vertically thin cloud (cloud extent approx. 0.2 - 0.4 km) with a cloud top height of approximately 1 km, case study 2 shows a more vertically extended cloud (cloud extent approx. 1.2 km) with a cloud top height of about 2 km. The performed radiative transfer simulations for both cloud cases include variations of the cloud top height between $0.2 \leq h_{cloud} \leq 3.0$ km and all possible vertical extents e_{cloud} on a 0.2 km altitude grid (thus including $e_{cloud} = 0.2$ km and $e_{cloud} = h_{cloud}$).

For case study 1, the use of an asymmetry parameter of 0.85 leads to $\tau_{cloud} \approx 16.1$, which agrees well with τ_{CAI} and τ_{CPR}. The fit results show a minimum root mean square (RMS) at $B_{fit} \approx 0.9$ and the corresponding photon path length distributions reveal an almost linear dependence of B_{fit} with the medians of the distributions. The average photon path lengths $\overline{PPL}$ also decrease with increasing B_{fit}, but show a spread of values for the same B_{fit}. This observation indicates that the fit results do not solely rely on $\overline{PPL}$, but on the complete photon path length distribution, leading to the observed spectrum.

In general, two interpretation strategies for the fit results are followed: In a first attempt, only the value of the RMS is taken into account. As a second strategy, ideal conditions for both the measurement and the simulation are assumed and the cloud scenarios with $B_{fit} \approx 1$ and low RMS are considered. For case study 1, the focus on the lowest RMS values reveals a systematic overestimation of the simulated O_2 absorption of 5-10%, which is similar to the clear-sky case study over a bright surface. The selection of a best corresponding cloud scenario is difficult because the minimum of the RMS - B_{fit} dependence covers many simulations whose RMS values differ only slightly (the minimum error - only due to the simulation noise - is $\Delta\text{RMS} \approx \pm 10^{-4}$ W/m^2/sr/nm). However, the group of cloud scenarios with the lowest RMS shows some common characteristics: The cloud top height is systematically higher than the one measured by CALIOP and the clouds all extend down to the ground surface or shortly above. Thus, also the vertical extent is systematically larger than measured by the CPR instrument. Although the disagreement is obvious, one should keep in mind that low clouds are more difficult to accurately be measured by the Cloud Profiling Radar because of the surface return signals (see section 7.2). However, the percentile values of the photon path length distributions at very

low O_2 absorption ($\lambda = 772.145$ nm) are rather similar.

For the second interpretation strategy, only those cloud situations with resulting B_{fit} near unity are considered. The lowest RMS in this region applies to three cloud scenarios (all within the minimum error range $\Delta B_{fit} \approx \pm\ 2 \cdot 10^{-3}$) which have very different vertical dimensions. As for the other interpretation possibility, the cloud top heights are higher than measured by the lidar instrument. The cloud scenario closest to the collocated profile data features a cloud between 0.8 and 1.6 km.

Fits were also performed in small subwindows of the O_2 A-band, covering wavelength regions of strong, intermediate and weak absorption regimes. In general, this method has the potential to provide further insight to the fit results, especially considering the sensitivity of the different regions to very long and very short photon paths. However, for the investigated examples, no additional information could be retrieved from these fits except a similarity of the results between the use of the entire O_2 A-band as well as strong and intermediate absorption regimes (indicating a certain degree of consistency).

From case study 2, featuring a higher one-layered cloud, similar conclusions can be drawn as from case study 1. The derived cloud optical depth is $\tau_{cloud} \approx 22.2$ and thus close to τ_{CAI}. Considering the lowest values of the fit RMS, an overestimation of the O_2 absorption of 5-10% is derived. The derived cloud top heights are between 2.0 and 2.8 km, the base heights between 0.2 and 0.6 km (where lower base heights belong to higher top heights), thus all derived cloud scenarios show a vertical extent of at least 1.4 km. Many of the included scenarios show higher cloud tops than detected by the lidar. However, two cloud scenarios (cloud layer from 0.6 - 2.0 km and 0.6 - 2.2 km) are close to the lidar-and radar based measurements (cloud between approx. 0.8 and 2.0 km). The photon path length distributions are similar for those two cloud scenarios. The remaining cloud scenarios show similar photon path length distributions as well.

Concentrating on the cloud scenarios where the simulation corresponds best to the O_2 absorption (B_{fit} close to unity), cloud scenarios of very different vertical dimensions result. However, the included cloud between 1.4 and 2.4 km is close to the collocated lidar and radar observations.

If the focus is set on the smallest fit RMS, both one-layer cloud cases as well as the clear sky measurement over a bright surface show a systematic overestimation of the O_2 absorption within the radiative transfer simulations to best fit the measurements. Consequences are increased cloud top heights and possibly a reduced surface albedo. The fine changes of the vertical cloud dimensions between the different scenarios result in fit RMS only differing by small values. A general uncertainty of the method must be assumed which makes the interpretation in this case challenging.

In order to exclude possible effects on the simulations and to inquire measurement related influences, several sensitivity tests were performed. The individual studies could not be explored in every detail but allow to exclude certain effects and/or provide a first measure of uncertainty due to others. The investigated topics are as follows:

- The uncertainties of the fit results solely due to the statistical noise of the Monte Carlo simulations are $\Delta\text{RMS} \approx \pm 10^{-4}$ W/m^2/sr/nm and $\Delta\ B_{fit} \approx \pm 2{\cdot}10^{-3}$. Indeed, these values must be considered to being a minimum uncertainty for RMS and B_{fit}. No uncertainties of the τ_{cloud} retrieval are considered so far, but the possible effect is expected to be small because the averaged continuum radiance of the simulation is required to be as close as 2% to the corresponding value of the measurement.
- The comparison of different cloud top pressures within the TANSO-FTS FOV from collocated MODIS-Aqua measurements shows that the cloud top heights of the cloud systems from case study 1 and 2 are mostly homogeneous.

- The use of a different asymmetry parameter within case study 1 (g=0.8) shows similar fit results to the original settings (g=0.85) as long as τ_{cloud} is adapted accordingly. Also, the resulting photon path length statistics are almost identical (verified for selected cloud scenarios).

- Adapting the model atmosphere to a location-dependent temperature and pressure profile leads to slightly higher B_{fit} values (verified for selected cloud scenarios), but the effect is below 1% and cannot explain a systematic shift of B_{fit} of 5-10% to lower values.

- An important influence might be the spectral sampling of the O_2 absorption cross sections. In the performed simulations, a spacing of $\delta\lambda = 0.05$ cm^{-1} was assumed which corresponds to approximately 0.003 nm for the O_2 A-band. For comparison, the spectral spacing of the simulations is $\delta\lambda_{sim} = 0.005$ nm and the spectral spacing of the TANSO-FTS spectra is $\delta\lambda_{FTS} \approx 0.012$ nm. McArtim interpolates between the values of the look-up table to derive the absorption cross sections which are used in the radiative transfer simulation. This can lead to a systematic error if the wavelength spacing is too large. Another look-up table of absorption cross sections was created using $\delta\lambda = 0.01$cm^{-1} (which corresponds to about $6 \cdot 10^{-4}$ nm in the O_2 A-band) and the simulations repeated for a number of cloud scenarios of case 1. In comparison to the original settings, B_{fit} is systematically shifted to higher values, but not exceeding shifts of 0.01. Although the finer wavelength grid of the O_2 absorption cross sections should definitely be used in future studies to obtain a higher degree of accuracy, a deviation of 5-10% of B_{fit} from unity can also not be explained. Interestingly, the RMS values for lower B_{fit} systematically slightly increase while those for higher B_{fit} slightly decrease or increase. This effect has to be investigated in more detail. Although the absorption line shapes of O_2 are subject of continuous improvement, no indication has been found in other studies that the used values are inaccurate.

- Depending on the case study, the polarized radiances (P and S) from the TANSO-FTS measurement are very different. The full radiance, however, is derived from the values of both polarization spectra. Comparing the case studies with one-layer clouds, case 1 shows similar polarized spectra while those of case 2 are very different. This is surprising and currently cannot be explained. In particular, it should be investigated if this is an indication for possible measurement problems. The radiance difference should be regarded for each future case study to detect possible influences on the fit results.

Although reasons for a small increase of B_{fit} closer to unity can be identified, none of them, also in combination, allows to justify a deviation of as much as 10%. However, only a small selection of possible uncertainties has been investigated and only for one cloud case study or even selected cloud scenarios. Besides a more thorough examination of these effects, possible additional influences on the fit results should be regarded. These include the description of the cloud model (homogeneous, Henyey-Greenstein phase function) which should be replaced by a more accurate Mie phase function and the applied modification of the TANSO-FTS instrument line shape function (see section 5.4) by mainly cropping the wings to a certain degree. Also the measurement uncertainties of the TANSO-FTS radiances should be taken into account. For the measurement spectra, the P- and S-radiances I_P and I_S provided by the GOSAT teams (including degradation corrections), are used. Subsequently, the entire radiance I is approximated by calculating the average of both spectra. Although the data is subject to continued investigation by the responsible science teams, the effect of these approximations should be investigated in more detail.

The initial objective of this study was the investigation of radiative transfer in presence of complex cloud systems, including vertical and horizontal heterogeneities. Case study 3 shows a vertically inhomogeneous cloud which is formed by at least two layers of different extinction strengths. While no special selection was performed for choosing the first two case studies, this cloud scenario was chosen for its more complex vertical structure. The entire cloud system

extends from approx. 0.6 to 5.0 km, and the largest extinction takes place in the bottom layers (measured by the CPR in some distance to the TANSO-FTS FOV). The measurement situation is especially interesting as CALIOP is blind to the underlying cloud layers.

Because of the complexity of the cloud situation, only selected cloud scenarios were simulated, including various two-layer clouds with identical cloud bases and tops of the complete system and one-layer clouds. In case of the two-layer systems, the cloud optical depth $\tau_{cloud} = \tau_1 + \tau_2$ was attributed to the layers by fulfilling the ratios $\tau_1 : \tau_2 = \frac{1}{3} : \frac{2}{3}$ or $\tau_1 : \tau_2 = \frac{1}{10} : \frac{9}{10}$ and vice versa. The derived total cloud optical depth is $\tau_{cloud} \approx 32$, which is only slightly lower than τ_{CAI}. From the fit results (RMS and B_{fit}), a preselection of the scenarios was performed: The cloud scenarios with higher extinction in the upper cloud layer show large underestimations of the O_2 absorption and those clouds consisting of a single layer systematically have higher RMS values. Because of these large deviations, these two groups of cloud scenarios were discarded from further consideration which reduces the variety of different cloud scenarios before possible refinements.

Further reducing the selection to the smallest RMS in and outside the proximity of $B_{fit} = 1$, three cloud scenarios remain. Two scenarios show $B_{fit} < 0.93$, and one cloud scenario results in the expected O_2 absorption. It should be noted that the assumed cloud scenarios are more divers from each other than those for case studies 1 and 2. This probably leads to larger differences between the RMS values (and less influence of small uncertainties) and thus an easier identification of well suiting scenarios.

Except the scenario with the smallest B_{fit}, all of them are similar to an approximated version of the cloud profile from the lidar and radar measurements. Both scenarios are in agreement with the two interpretation strategies mentioned for case studies 1 and 2. Naturally, the truth of the cloud structure inside the TANSO-FTS FOV remains mostly unknown, except the cloud top height h_{cloud} for selected datapoints. Case study 3 shows that despite the discussed discrepancies during the retrieval, basic properties of the cloud system can be distinguished with the current implementation of the retrieval comparing the O_2 A-band absorption structures to radiative transfer simulations. Case study 3 was also used to demonstrate the information content of the derived photon path length statistics at a low absorbing wavelength.

In principle, for the next steps of the project, a refinement of the case 3 cloud scenario would be planned together with the consideration of more complex cloud case studies. However, a discrepancy between the derived O_2 absorptions for the cloud scenarios with the lowest fit RMS and the expected O_2 absorption during the measurement was found (between 5-10%), which needs to be resolved first. Within this study, no explanation for this discrepancy was found. In order to examine this effect in detail, the following steps are proposed:

- Small adjustments of the simulation environment should be implemented, especially the O_2 absorption cross sections should be implemented on a finer wavelength grid.
- A larger number of clear sky cases over surfaces with different albedos a_s should be compared to radiative transfer simulations to examine the discrepancy between the so-far detected under- and overestimation of the O_2 absorption for a dark and a bright surface, respectively. Here, also fits of different a_s should be used to learn more about the effect of a_s on B_{fit}.
- The collocated dataset, which was created from 5 years of measurements from instruments onboard GOSAT, CALIPSO and CloudSat, should be used to investigate a larger number (if possible) of case studies having common characteristics (h_{cloud}, τ_{cloud}, SZA). The comparison of the B_{fit} values at minimum RMS can show if the observed discrepancy to $B_{fit} \approx 1$ and the systematic overestimation of the cloud top heights depend on some characteristics of the cloud system or applies to all retrievals. Here, a special focus should be

laid also on the comparison of the P- and S-radiances of the measurements, more explicitly the radiance differences.

Further objectives are the implementation of Mie phase functions in order to more accurately describe the cloud's microphysics and the inclusion of horizontal heterogeneities of the cloud (McArtim offers both methods), especially when further complex cloud systems are regarded.

List of Figures

1 Examples of possible photon paths for different cloud scenarios. 7
2 Space-based measurements of the O_2 A-band over a land and sea surface from TANSO-FTS. 8
3 A possible photon trajectory, simulated with McArtim. 10
4 Schematical description of the line mixing process. 14
5 Vertical column density of O_2 in the model atmosphere within the O_2 A-band. . 15
6 Examples of the Mie phase function $\Phi_{Mie}(\theta)$. 17
7 Examples of the The Henyey-Greenstein phase function $\Phi_{HG}(\theta)$. 17
8 Design of the TANSO-FTS pointing system and optics. 22
9 Radiometric degradation factors for TANSO-FTS for bands 1-3. 24
10 Central parts of the TANSO-FTS ILSF for band 1. 25
11 Comparison of the TANSO-FTS spectrum convolved with the polarization and wavenumber averaged ILSF and the TANSO-FTS spectrum convolved with the polarization averaged ILSF considering the wavenumber dependency. 26
12 Comparison of the TANSO-FTS spectrum convolved with the polarization and wavenumber averaged ILSF with wings cropped to $\pm$ 1.5 cm^{-1} and the TANSO-FTS spectrum convolved with the polarization and wavenumber averaged ILSF with complete wings. 27
13 A backward trajectory of a photon. 31
14 Solar irradiance spectrum (Kurucz) and fitted irradiance spectrum using the pseudo-transmittance (Toon). 32
15 Absorption cross sections of O_2 for different altitudes. 33
16 Derivation of the McArtim satellite elevation angle. 34
17 Comparison of inelastic and elastic radiances as published by Vasilkov et al. and as calculated with McArtim. 36
18 Schematic overview of a photon trajectory through the atmosphere. 37
19 Schematic representation of the photon penetration altitude. 38
20 Schematic representation of a photon path length below the altitude a_{ref}. 39
21 Geometric calculation of the path from the scattering event to a_{ref} in direction of the Sun. 40
22 Schematic representation of a photon path length within an individual altitude bin. 40
23 Impression of the NASA A-train (afternoon) formation in 2013. 42
24 Spatial collocation criterium between SSP of CALIPSO and the TANSO-FTS FOV. 44
25 TANSO-FTS measurement locations for the derived collocations. 46
26 Case 1: FOV image from the GOSAT monitor camera. 48
27 Case 1: TANSO-FTS radiances between 759 and 772 nm. 49
28 Case 1: TANSO-CAI cloud optical depth, TANSO-FTS FOV and CALIOP and CPR measurement ground tracks. 49
29 Case 1: CALIOP cloud top and cloud base heights. 50
30 Case 1: Altitude resolved CPR cloud optical depth. 51
31 Case 2: FOV image from the GOSAT monitor camera. 51
32 Case 2: TANSO-FTS radiances between 759 and 772 nm. 52
33 Case 2: TANSO-CAI cloud optical depth, TANSO-FTS FOV and CALIOP and CPR measurement ground tracks. 53
34 Case 2: CALIOP cloud top and cloud base heights. 53
35 Case 2: Altitude resolved CPR cloud optical depth. 54
36 Case 3: FOV image from the GOSAT monitor camera. 55
37 Case 3: TANSO-FTS radiances between 759 and 772 nm. 56
38 Case 3: TANSO-CAI cloud optical depth, TANSO-FTS FOV and CALIOP and CPR measurement ground tracks. 56
39 Case 3: CALIOP cloud top and cloud base heights. 57
40 Case 3: Altitude resolved CPR cloud optical depth. 58

41 Possible regions of continuum radiances in the O_2 A-band. 61
42 Example of the cloud optical depth determination for a given cloud scenario. . . 62
43 Spectral calibration factor of the TANSO-FTS instrument. 64
44 O_2 optical depth and location of deep Fraunhofer lines in the O_2 A-band. 66
45 Subbands of strong, intermediate and weak O_2 absorptions in the O_2 A-band. . . 67
46 Schematic overview of the fitting procedure. 68
47 Synthetic reference spectrum, input spectrum and added white noise. 71
48 δ_{shift} and $\delta_{squeeze}$ from the pure wavelength alignment fit using the reference spectrum. 71
49 Spectrally aligned simulations and residuals to the reference spectrum. 72
50 Fitted spectra and residuals (using the reference spectrum), also allowing a variation of the fit parameter B. 72
51 Dependence of RMS and B from the fit for all simulations to the reference spectrum. 73
52 Fitted spectra and residuals for simulated spectra with a constant cloud situation and different surface albedos. 74
53 Fit RMS and B for simulations with a constant cloud situation and different surface albedos. 75
54 Resulting RMS for simulated spectra being modified through Gaussian convolutions with slowly increasing FWHM. 76
55 Resulting RMS for the simulated spectra fitted to the TANSO-FTS spectrum which is smoothed by Gaussian convolutions with slowly increasing FWHM. . . . 77
56 Wavelength-aligned comparison of the simulation and the measurement in the case of a clear-sky scenario over dark surface. Fit of the simulation to the measurement also considering a parameter B_{fit}. 79
57 Wavelength-aligned comparison of the simulation and the measurement in the case of a clear-sky scenario over bright surface. Fit of the simulation to the measurement also considering a parameter B_{fit}. 80
58 TANSO-FTS radiances between 759 and 772 nm over water. 81
59 TANSO-FTS radiances between 759 and 772 nm over desert. 82
60 Case study 1: Fitted spectra for all cloud scenarios, the TANSO-FTS measurement spectrum and the corresponding residuals. 83
61 Case study 1: The resulting fit RMS in dependence of the fit parameter B_{fit} for all simulated cloud scenarios (g=0.85). 84
62 Case study 1: Photon path length distributions for all assumed cloud scenarios (g=0.85) in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. 85
63 Case study 1: The resulting fit RMS and B_{fit} for selected cloud scenarios for g=0.85 and g=0.8. 88
64 Case study 1: Photon path length distributions for selected cloud scenarios (g=0.85, g=0.8) in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. 89
65 Case study 1: Penetration altitudes for the cloud scenarios 31, 74 and 117 with g=0.85 and g=0.8. 90
66 Case study 1: Comparison of the geometric photon path length distributions below 10 km for the cloud scenarios 31, 74 and 117 with g=0.85 and g=0.8. . . . 91
67 Case study 1: Collocation plot of the TANSO-FTS FOV, CALIOP and CPR measurements together with the MODIS Aqua measurement locations where the cloud top pressure is retrieved. 92
68 Case study 1: Fit results (RMS, B_{fit}) for selected cloud scenarios from simulated radiance spectra based on different atmospheric profiles. 93
69 Case study 1: Fit results (RMS, B_{fit}) for radiance calculations based on different wavelength resolutions of the O_2 absorption cross sections. 94

70 Case study 1: Fit results (RMS, B_{fit}) for different wavelength regions of the O_2 A-band. . . . 97
71 Case study 1: Fit results (RMS, B_{fit}) of three cloud scenarios for 10 radiance calculations with the same set-up. . . . 98
72 Case study 1: O_2 weighted photon path length distributions for all assumed cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 99
73 Case study 1: Mean and median values of the photon path length distributions, using the geometric or the O_2 weighted results. . . . 100
74 Case study 1: O_2 weighted photon path length distributions for selected cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 100
75 Case study 1: Comparison of the O_2 weighted photon path length distributions for the cloud cases 31, 74 and 117 with g=0.85 and g=0.8. . . . 101
76 Case study 2: Fitted spectra for all cloud scenarios, the TANSO-FTS measurement spectrum and the corresponding residuals. . . . 102
77 Case study 2: The resulting fit RMS in dependence of the fit parameter B_{fit} for all simulated cloud scenarios. . . . 103
78 Case study 2: Photon path length distributions for all assumed cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 103
79 Case study 2: Photon path length distributions for selected cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 104
80 Case study 2: Collocation plot of the TANSO-FTS FOV, CALIOP and CPR measurements together with the MODIS Aqua measurement locations where the cloud top pressure is retrieved. . . . 108
81 Case study 2: O_2 weighted photon path length distributions for all assumed cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 109
82 Case study 2: O_2 weighted photon path length distributions for selected cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 109
83 Case study 3: Visualization of the simulated cloud scenarios. . . . 110
84 Case study 3: Fitted spectra for all cloud scenarios, the TANSO-FTS measurement spectrum and the corresponding residuals. . . . 112
85 Case study 3: The resulting fit RMS in dependence of the fit parameter B_{fit} for all simulated cloud scenarios. . . . 113
86 Case study 3: Photon path length distributions for all assumed cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. . . . 114
87 Case study 3: Comparison of the photon path length distributions below 10 km for simulations resulting in a B_{fit} near unity and B_{fit} value < 1. . . . 115
88 Case study 3: Comparison of photon path length distributions below 10 km altitude for simulations resulting in B_{fit} near unity and spectra of $B_{fit} > 1$. . . . 116
89 Case study 3: Penetration altitudes reached by the detected photons (scenarios D, L, M). . . . 118
90 Case study 3: Photon path length distributions of scenario M for selected altitude layers. . . . 120
91 Case study 3: The resulting fit RMS in dependence of the optimized fit parameter B_{fit} for the three subbands of the O_2 A-band. . . . 122

92 Case study 3: O_2 weighted photon path length distributions for all assumed cloud scenarios in boxplot representation in dependence of the B_{fit} parameter. Medians and means of the photon path length distributions in dependence of B_{fit}. 123
93 Case study 3: Comparison of the O_2 weighted photon path length distributions below 10 km for simulations resulting in a B_{fit} near unity and B_{fit} value < 1. . . 124
94 Case study 3: Comparison of O_2 weighted photon path length distributions below 10 km altitude for simulations resulting in B_{fit} near unity and spectra of $B_{fit} > 1$.124

List of Tables

1 Main specifications and orbital configurations of the GOSAT satellite. 20
2 Specifications of TANSO-FTS measurement bands 1-4. 21
3 Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 10° and 20°. 47
4 Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 20° and 30°. 47
5 Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 30° and 40°. 47
6 Number of collocated measurement situations for solar zenith angles at the TANSO-FTS center of FOV between 40° and 50°. 47
7 Geometrical characteristics of case study 1 as well as selected surface properties at the measurement location. 48
8 Geometrical characteristics of case study 2 as well as selected surface properties at the measurement location. 52
9 Geometrical characteristics of case study 3 as well as selected surface properties at the measurement location. 55
10 Data products used for the cloud scenario characterization of case study 1. . . . 59
11 Data products used for the cloud scenario characterization of case study 2. . . . 59
12 Data products used for the cloud scenario characterization of case study 3. . . . 59
13 Differences between the averaged continuum radiances of continuum A and B for measured TANSO-FTS spectra and McArtim simulations. 63
14 Parameter boundaries for the simulated model clouds (case study 1), characterized by h_{cloud}, e_{cloud} and τ_{cloud}. 69
15 All performed simulations for case study 1 (g=0.85). 70
16 Properties of the model clouds for different values of the surface albedo after adapting the cloud optical depth τ_{cloud}. 74
17 Geometrical characteristics of the clear sky measurement over water. 78
18 Geometrical characteristics of the clear sky measurement over land (desert). . . . 79
19 Case study 1: Cloud parameters and fit results for the cloud scenarios with the lowest RMS values. 86
20 Case study 1: Cloud parameters and fit results for the cloud scenarios with resulting B_{fit} close to unity and low RMS. 87
21 All performed simulations for case study 1 using g=0.85. 95
22 All performed simulations for case study 1 using g=0.8. 96
23 Case study 2: Cloud parameters and fit results for the cloud scenarios with the lowest RMS values. 105
24 Case study 2: Cloud parameters and fit results for the cloud scenarios with resulting B_{fit} close to unity and low RMS. 105
25 All performed simulations for case study 2. 107
26 All model clouds implemented for case study 3 including the fit results. 111

References

[1] CALIOP V4.10 L1 & L2 Release Announcement (10 November, 2016). https://eosweb.larc.nasa.gov/news/caliop-v410-l1-l2-release-announcement, accessed on 29.10.2017.

[2] Terra/MODIS-derived Land Surface Albedo Product. *http://modis-atmos.gsfc.nasa.gov/ALBEDO/*, accessed on 9.1.2012 and 4.10.2012.

[3] *FTS FOV image for scan ID F090604215013382401*, https://data.gosat.nies.go.jp/, accessed on 01.06.2016.

[4] *FTS FOV image for scan ID F090828235909422301*, https://data.gosat.nies.go.jp/, accessed on 01.06.2016.

[5] *FTS FOV image for scan ID F150326232730412204*, https://data.gosat.nies.go.jp/, accessed on 01.06.2016.

[6] Japan Aerospace Exploration Agency, National Institute for Environmental Studies, and Ministry of the Environment. *GOSAT/IBUKI Data Users Handbook*. 1st edition, March 2011.

[7] B. A. Baum, W. P. Menzel, R. A. Frey, D. C. Tobin, R. E. Holz, S. A. Ackerman, A. K. Heidinger, and P. Yang. MODIS cloud-top property refinements for Collection 6. *Journal of applied meteorology and climatology*, 51(6):1145–1163, 2012.

[8] O. Boucher, D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S.K. Satheesh, S. Sherwood, B. Stevens, and X.Y. Zhang. Clouds and Aerosols. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]*. United Kingdom and New York, NY, USA, 2013.

[9] P. Bromiley. Products and convolutions of gaussian probability density functions. *Tina-Vision Memo*, 3(4), 2003.

[10] J. P. Burrows, U. Platt, and P. Borrell. Tropospheric Remote Sensing from Space. In *The Remote Sensing of Tropospheric Composition from Space*, pages 1–65. Springer, 2011.

[11] A. Butz, S. Guerlet, O. Hasekamp, D. Schepers, A. Galli, I. Aben, C. Frankenberg, J.-M. Hartmann, H. Tran, A. Kuze, et al. Toward accurate CO_2 and CH_4 observations from GOSAT. *Geophysical Research Letters*, 38(14):L14812, 2011.

[12] Atmospheric Science Data Center. *CALIPSO Quality Statements Lidar Level 2 Cloud and Aerosol Layer Products Version Releases: 3.01, 3.02*.

[13] ICARE Data and Services Center. CAL_LID_L2_05kmCLay-Standard-V4-10 data product. http://www.icare.univ-lille1.fr, accessed on 17.04.2017.

[14] ICARE Data and Services Center. CAL LID L2 05kmALay-Standard-V4-10 data product. http://www.icare.univ-lille1.fr, accessed on 17.04.2017.

[15] ICARE Data and Services Center. CAL_LID_L2_05kmCLay-Standard-V4-10 data product. http://www.icare.univ-lille1.fr, accessed on 18.03.2017.

[16] ICARE Data and Services Center. CAL LID L2 05kmALay-Standard-V4-10 data product. http://www.icare.univ-lille1.fr, accessed on 18.03.2017.

[17] ICARE Data and Services Center. CAL_LID_L2_05kmCLay-Standard-V4-10 data product. http://www.icare.univ-lille1.fr, accessed on 21.03.2017.

[18] ICARE Data and Services Center. CAL LID L2 05kmALay-Standard-V4-10 data product. http://www.icare.univ-lille1.fr, accessed on 22.03.2017.

[19] ICARE Data and Services Center. CAL_LID_L2_333mCLay-ValStage1-V3-01 data product. http://www.icare.univ-lille1.fr, accessed on 27.04.2016.

[20] ICARE Data and Services Center. CAL_LID_L2_333mCLay-ValStage1-V3-01 data product. http://www.icare.univ-lille1.fr, accessed on 28.04.2016.

[21] K. Chance and R. L. Kurucz. An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(9):1289–1295, 2010.

[22] US Coesa. Standard atmosphere, 1976. *US Government Printing Office, Washington, DC*, 1976.

[23] The Pytroll crew. *pyorbital Documentation, Release v1.1.1* , 2017.

[24] T. Deutschmann. *On Modeling Elastic and Inelastic Polarized Radiation Transport in the Earth Atmosphere with Monte Carlo Methods*. PhD thesis, 2014.

[25] T. Deutschmann et al. The Monte Carlo atmospheric radiative transfer model McArtim: Introduction and validation of Jacobians and 3D features. *JQSRT (2011), doi:10.1016/j.jqsrt.2010.12.009*.

[26] G. D. Durgin. ASD5: Earth Station Look Angles, 2009.

[27] C. Frankenberg, C. O'Dell, L. Guanter, and J. McDuffie. Remote sensing of near-infrared chlorophyll fluorescence from space in scattering atmospheres: implications for its retrieval and interferences with atmospheric CO_2 retrievals. *Atmospheric Measurement Techniques*, 5(8):2081–2094, 2012.

[28] O. Funk. *Photonen Weglängenverteilungen bei bewölktem Himmel - Sauerstoff A-Banden Messungen und Modellrechnungen zum Strahlungstransport*. PhD thesis, Universität Heidelberg, 2000.

[29] O. Funk and K. Pfeilsticker. Photon path length distributions for cloudy skies? oxygen A-Band measurements and model calculations. *Annales Geophysicae*, 21(3):615–626, 2003.

[30] R. Gelaro, W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454, 2017.

[31] Militärisches Geowesen. World Geodetic System 1984 (WGS84). *Amt für militärisches Geowesen*, 1988.

[32] M. Geyer. Earth-referenced Aircraft Navigation and Surveillance Analysis. Technical report, John A. Volpe National Transportation System Center, 2016.

[33] J.-M. Hartmann, C. Boulet, and D. Robert. *Collisional effects on molecular spectra: laboratory experiments and models, consequences for applications*. Elsevier, 2008.

[34] A. K. Heidinger and G. L. Stephens. Molecular line absorption in a scattering atmosphere. Part III: Pathlength characteristics and effects of spatially heterogeneous clouds. *Journal of the atmospheric sciences*, 59(10):1641–1654, 2002.

[35] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *The Astrophysical Journal*, 93:70–83, 1941.

[36] G. Herzberg. Forbidden Transitions in Diatomic Molecules. *Astrophysical Journal*, 89:288, March 1939.

[37] H. Ishida, T. Nakajima, and N. Kikuchi. *Algorithm Theoretical Basis Document for GOSAT TANSO-CAI L2 cloud flag*, 1.00 edition, February 2011.

[38] H. Iwabuchi. Retrieval of cloud optical thickness and effective radius using multispectral remote sensing and accounting for 3D effects. In *Light Scattering Reviews 2*, pages 97–124. Springer, 2007.

[39] J. Joiner, P. K. Bhartia, R. P. Cebula, E. Hilsenrath, R. D. McPeters, and H. Park. Rotational Raman scattering (Ring effect) in satellite backscatter ultraviolet measurements. *Applied Optics*, 34(21):4513–4525, 1995.

[40] J. Joiner, Y. Yoshida, A. P. Vasilkov, Y. Yoshida, L. A. Corp, and E. M. Middleton. First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences*, 8(3):637–651, 2011.

[41] T. S. Kelso. CelesTrak. Public Domain Satellite Tracking Data, URL: http://celestrak.com/, accessed in 2016.

[42] A. Kokhanovsky. *Cloud Optics*. Atmospheric and Oceanographic Sciences Library. Springer Netherlands, 2006.

[43] A. Kokhanovsky, S. Platnick, and M. D. King. Remote sensing of terrestrial clouds from space using backscattering and thermal emission techniques. In *The Remote Sensing of Tropospheric Composition from Space*, pages 231–257. Springer, 2011.

[44] R. L. Kurucz. High resolution irradiance spectrum from 300 to 1000 nm. URL = Kurucz.harvard.edu/sun/irradiance2005/irradthu.dat, 2005, revised 2.1.2010.

[45] A. Kuze, H. Suto, M. Nakajima, and T. Hamazaki. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Applied Optics*, 48(35):6716–6733, 2009.

[46] A. Kuze, H. Suto, K. Shiomi, T. Urabe, M. Nakajima, J. Yoshida, T. Kawashima, Y. Yamamoto, and F. Kataoka. Level 1 algorithms for TANSO on GOSAT: processing and on-orbit calibrations. *Atmospheric Measurement Techniques Discussions*, 5:2959–3018, 2012.

[47] A. Kuze, T. E. Taylor, F. Kataoka, C. J. Bruegge, D. Crisp, M. Harada, M. Helmlinger, M. Inoue, S. Kawakami, N. Kikuchi, et al. Long-term vicarious calibration of GOSAT short-wave sensors: techniques for error reduction and new estimates of radiometric degradation factors. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7):3991–4004, 2014.

[48] A. Kuze, H. Suto, K. Shiomi, S. Kawakami, M. Tanaka, Y. Ueda, A. Deguchi, J. Yoshida, Y. Yamamoto, F. Kataoka, T. E. Taylor, and H. L. Buijs. Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space. *Atmospheric Measurement Techniques*, 9(6):2445–2461, 2016.

[49] S. Li and Q. Min. Diagnosis of multilayer clouds using photon path length distributions. *Journal of Geophysical Research: Atmospheres*, 115(D20), 2010.

[50] K.-N. Liou. *An introduction to atmospheric radiation*, volume 84. Academic press, 2002.

[51] D. A. Long and J. T. Hodges. On spectroscopic models of the O_2 A-band and their impact upon atmospheric retrievals. *Journal of Geophysical Research: Atmospheres*, 117(D12), 2012.

[52] D.A. Long, D.K. Havey, M. Okumura, C.E. Miller, and J.T. Hodges. O_2 A-band line parameters to support atmospheric remote sensing. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(14):2021–2036, 2010.

[53] R. Marchand, G. G. Mace, T. Ackerman, and G. Stephens. Hydrometeor detection using CloudSat-An Earth-orbiting 94-GHz cloud radar. *Journal of Atmospheric and Oceanic Technology*, 25(4):519–533, 2008.

[54] A. Marshak and A. B. Davis. Horizontal fluxes and radiative smoothing. In *3D Radiative Transfer in Cloudy Atmospheres*, pages 543–586. Springer, 2005.

[55] A. Marshak, A. Davis, W. Wiscombe, and R. Cahalan. Radiative smoothing in fractal clouds. *JOURNAL OF GEOPHYSICAL RESEARCH-ALL SERIES*, 100:26–26, 1995.

[56] W. P. Menzel, R. A. Frey, H. Zhang, D. P. Wylie, C. C. Moeller, R. E. Holz, B. Maddux, B. A. Baum, K. I. Strabala, and L. E. Gumley. MODIS global cloud-top pressure and amount estimation: Algorithm description and results. *Journal of Applied Meteorology and Climatology*, 47(4):1175–1198, 2008.

[57] G. Mie. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annalen der physik*, 330(3):377–445, 1908.

[58] M. Min, P. Wang, J. R. Campbell, X. Zong, and J. Xia. Cirrus cloud macrophysical and optical properties over North China from CALIOP measurements. *Advances in Atmospheric Sciences*, 28(3):653–664, 2011.

[59] CloudSat Project A NASA Earth System Science Pathfinder Mission. *Level 2 Cloud Scenario Classification Product Process Description and Interface Control Document*, 5.0 edition, 2007.

[60] CloudSat Project A NASA Earth System Science Pathfinder Mission. *Level 2 Cloud Optical Depth Product Process Description and Interface Control Document*, 5.0 edition, February 2008.

[61] Eric G. Moody, Michael D. King, Crystal B. Schaaf, and Steven Platnick. MODIS-Derived Spatially Complete Surface Albedo Products: Spatial and Temporal Pixel Distribution and Zonal Averages. *Journal of Applied Meteorology and Climatology*, 47(11):2879–2894, 2008.

[62] T. Nakajima and M. Tanaka. Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 40(1):51–69, 1988.

[63] GOSAT Project National Institute for Environmental Studies. *NIES GOSAT TANSO -CAI Level 2 Data Product Format Description, Version 3.00* , November 2016.

[64] GOSAT Project Office National Institute for Environmental Studies. *NIES GOSAT TANSO-FTS Level 2 Input Data Set Format Description*, 1.00 edition, April 2012.

[65] GOSAT Project National lnstitute for Environmental Studies. *Algorithm Theoretical Basis Document for GOSAT TANSO-CAI L2 cloud flag*, 1.00 edition, February 2011.

[66] M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, September 2014.

[67] JAXA EORC NIES GOSAT Project. *Geolocation Error in TANSO-FTS L1B and its countermeasure*, December 2016.

[68] JAXA EORC NIES GOSAT Project. *Notice to GOSAT General Users: Estimated Geolocation Data (EGD) is Now Available*, December 13 2016.

[69] D. M. O'Brien, I. Polonsky, P. Stephens, and T. E. Taylor. Feasibility of cloud screening using proxy photon pathlength distributions derived from high-resolution spectra in the near infrared. *Journal of Atmospheric and Oceanic Technology*, 27(1):135–146, 2010.

[70] D. M. O'Brien, I. Polonsky, C. O'Dell, A. Kuze, N. Kikuchi, Y. Yoshida, and V. Natraj. Testing the Polarization Model for TANSO-FTS on GOSAT Against Clear-Sky Observations of Sun Glint Over the Ocean. *IEEE Transactions on Geoscience and Remote Sensing*, 51(12):5199–5209, 2013.

[71] NIES GOSAT Project Office. *Release Note on the TANSO-CAI Level 2 Cloud Property Product* , 2.00 edition, February 2014.

[72] S. Oshchepkov, A. Bril, T. Yokota, I. Morino, Y. Yoshida, T. Matsunaga, D. Belikov, D. Wunch, P. Wennberg, G. Toon, et al. Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space: Validation of PPDF-based CO2 retrievals from GOSAT. *Journal of Geophysical Research: Atmospheres*, 117(D12), 2012.

[73] R. Penndorf. Tables of the refractive index for standard air and the Rayleigh scattering coefficient for the spectral region between 0.2 and 20.0 μ and their application to atmospheric optics. *Josa*, 47(2):176–182, 1957.

[74] A. Perrin. Review on the existing spectroscopic databases for atmospheric applications. In *Spectroscopy from Space*, pages 235–258. Springer, 2001.

[75] K. Pfeilsticker. First geometrical path length probability density function derivation of the skylight from high-resolution oxygen A-band spectroscopy, 2. Derivation of the Lévy index for the skylight transmitted by midlatitude clouds. *Journal of geophysical research*, 104(D43):4101–4116, 1999.

[76] K. Pfeilsticker, F. Erie, O. Funk, H. Veitel, and U. Platt. First geometrical pathlengths probability density function derivation of the skylight from spectroscopically highly resolving oxygen A-band observations 1. Measurement technique, atmospheric observations and model calculations. *Journal of geophysical research*, 103(D10):11483–11, 1998.

[77] K. Pfeilsticker, A. Lotter, C. Peters, and H. Bösch. Atmospheric detection of water dimers via near-infrared absorption. *Science*, 300(5628):2078–2080, 2003.

[78] U. Platt and J. Stutz. *Differential optical absorption spectroscopy: Principles and applications*. Springer Verlag, 2008.

[79] M. Quante. The role of clouds in the climate system. *J. Phys. IV France*, 121:61–86, 2004.

[80] C. Richard, I. E. Gordon, L. S. Rothman, M. Abel, L. Frommhold, M. Gustafsson, J.-M. Hartmann, C. Hermans, W. J. Lafferty, G. S. Orton, et al. New section of the HITRAN database: Collision-induced absorption (CIA). *Journal of Quantitative Spectroscopy and Radiative Transfer*, 113(11):1276–1285, 2012.

[81] A. Richter and T. Wagner. The use of UV, Visible and Near IR solar back scattered radiation to determine trace gases. In *The Remote Sensing of Tropospheric Composition from Space*, pages 67–121. Springer, 2011.

[82] W. Rödel and T. Wagner. *Physik unserer Umwelt: Die Atmosphäre*. Springer, 4^{th} edition, 2011.

[83] L.S. Rothman, I.E. Gordon, A. Barbe, D.C. Benner, P.F. Bernath, M. Birk, V. Boudon, L.R. Brown, A. Campargue, J.P. Champion, et al. The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(9-10): 533–572, 2009.

[84] A. Rozanov, V. Rozanov, M. Buchwitz, A. Kokhanovsky, and J.P. Burrows. SCIATRAN 2.0 - A new radiative transfer model for geophysical applications in the 175-2400 nm spectral region. *Advances in Space Research*, 36(5):1015–1019, 2005.

[85] S. Sanghavi, M. Lebsock, and G. Stephens. Sensitivity analysis of polarimetric O2 A-band spectra for potential cloud retrievals using OCO-2/GOSAT measurements. *Atmospheric Measurement Techniques*, 8(9):3601–3616, 2015.

[86] K. Sassen and Z. Wang. Classifying clouds around the globe with the CloudSat radar: 1-year of results. *Geophysical research letters*, 35(4), 2008.

[87] T. Scholl, K. Pfeilsticker, A. B. Davis, H.K. Baltink, S. Crewell, U. Löhnert, C. Simmer, J. Meywerk, and M. Quante. Path length distributions for solar photons under cloudy skies: Comparison of measured first and second moments with predictions from classical and anomalous diffusion theories. *J. Geophys. Res*, 111:D12211–12226, 2006.

[88] T. Soler and D. W. Eisemann. Determination of look angles to geostationary communication satellites. *Journal of surveying engineering*, 120(3):115–127, 1994.

[89] F. R. Spiering, M. B. Kiseleva, N. N. Filippov, H. Naus, B. Van Lieshout, C. Weijenborg, and W. J. Van Der Zande. Line mixing and collision induced absorption in the oxygen A-band using cavity ring-down spectroscopy. *The Journal of chemical physics*, 133:114305, 2010.

[90] G. L. Stephens and A. Heidinger. Molecular line absorption in a scattering atmosphere. Part I: Theory. *Journal of the atmospheric sciences*, 57(10):1599–1614, 2000.

[91] G. L. Stephens, D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, et al. The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, 83(12):1771–1790, 2002.

[92] G. L. Stephens, A. K. Heidinger, and P. M. Gabriel. Photon paths and cloud heterogeneity: An observational strategy to assess effects of 3D geometry on radiative transfer. In *3D Radiative Transfer in Cloudy Atmospheres*, pages 587–616. Springer, 2005.

[93] G. L. Stephens, D. G. Vane, S. Tanelli, E. Im, S. Durden, M. Rokey, D. Reinke, P. Partain, G. G. Mace, R. Austin, et al. CloudSat mission: Performance and early science after the first year of operation. *Journal of Geophysical Research: Atmospheres*, 113(D8), 2008.

[94] J. Stutz and U. Platt. Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurements with least-squares methods. *Applied Optics*, 35(30):6041–6053, 1996.

[95] S. Tanelli, S. L. Durden, E. Im, K. S. Pak, D. G. Reinke, P. Partain, J. M. Haynes, and R. T. Marchand. CloudSat's cloud profiling radar after two years in orbit: Performance, calibration, and processing. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (11):3560–3573, 2008.

[96] IBUKI Project Team, Space Application Mission Directorate, and Japan Aerospace Explaration Agency. IBUKI (GOSAT: Greenhouse Gases Observing Satellite), 2009/12F.

[97] G. C. Toon. Solar line list for GGG2014, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A., 2014.

[98] H. Tran and J.-M. Hartmann. An improved O_2 A band absorption model and its consequences for retrievals of photon paths and surface pressures. *Journal of Geophysical Research*, 113(D18):D18104, 2008.

[99] H. Tran, C. Boulet, and J.-M. Hartmann. Line mixing and collision-induced absorption by oxygen in the A band: Laboratory measurements, model, and tools for atmospheric spectra computations. *JOURNAL OF GEOPHYSICAL RESEARCH-ALL SERIES-*, 111 (D15):15210, 2006.

[100] H. C. van de Hulst. *Light scattering by small particles*. Courier Corporation, 1957.

[101] A. Vasilkov, J. Joiner, and R. Spurr. Note on rotational-Raman scattering in the O_2 A-and B-bands. *Atmospheric Measurement Techniques*, 6(4):981–990, 2013.

[102] M. A. Vaughan, K. A. Powell, D. M. Winker, C. A. Hostetler, R. E. Kuehn, W. H. Hunt, B. J. Getzewich, S. A. Young, Z. Liu, and M. J. McGill. Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements. *Journal of Atmospheric and Oceanic Technology*, 26(10):2034–2050, 2009.

[103] T. Wagner, T. Deutschmann, M. Grzegorski, B. Jäker, and M. Penning de Vries. Determination of atmospheric light path distributions. Presentation at the 3^{rd} GOSAT RA PI Meeting in Edinburgh, Scotland, UK, May 2011.

[104] J. Whitaker. pyproj 1.8.9. http://github.com/jswhit/pyproj.

[105] D. M. Winker, M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young. Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, 26(11):2310–2323, 2009.

[106] X. Xiong, B. N. Wenny, and W. L. Barnes. Overview of NASA Earth Observing Systems Terra and Aqua moderate resolution imaging spectroradiometer instrument calibration algorithms and on-orbit performance. *Journal of Applied Remote Sensing*, 3(1):032501–032501, 2009.

[107] Y. Yoshida, N. Kikuchi, and T. Yokota. On-orbit radiometric calibration of SWIR bands of TANSO-FTS onboard GOSAT. *Atmospheric Measurement Techniques Discussions*, 5: 4711–4734, 2012.

Acknowledgements

Declaration

I hereby declare that I wrote the dissertation submitted without any unauthorized external assistance and used only sources acknowledged in the work. All textual passages which are appropriated verbatim or paraphrased from published and unpublished texts as well as all information obtained from oral sources are duly indicated and listed in accordance with bibliographical rules. In carrying out this research, I complied with the rules of standard scientific practice as formulated in the statutes of Johannes Gutenberg-University Mainz to insure standard scientific practice.

Mainz, October 2017

Beke Kremmling