Algorithm Theoretical Basis Document

for the GOME-2 NRT, Offline and Data Record Sun-Induced Fluorescence Products

Document No.	SAF/AC/KNMI/ATBD/004
Issue	0.03
Date	15-11-2019
Number of Pages	45
Distribution list	KNMI
Summary contents	ATBD of the Sun-Induced Fluorescence products
Period	N/A
Authors:	Maurits Kooreman
Contributors:	Olaf Tuinder (KNMI), Folkert Boersma (WUR), Erik van Schaik (WUR)

CHANGE RECORD SHEET

Date	Issue	Pages affected	Description
20190408	0.01	All	Draft version.
20190529	0.02	All	KNMI Internal Review.
20190606	1.01	All	Initial version.

RELATED PRODUCT LIST

ID number	Instrument	Product Descripion
O3M-403	GOME-2A	NRT SIF
O3M-404	GOME-2B	NRT SIF
O3M-405	GOME-2C	NRT SIF
O3M-406	GOME-2A	Offline SIF
O3M-407	GOME-2B	Offline SIF
O3M-408	GOME-2C	Offline SIF
O3M-411	GOME-2A/B/C	SIF Monthly Product
O3M-516	GOME-2A/B/C	SIF Data Record

ALGORITHM CHANGE LOG

Alg/Sw/Conf/OPF	Changes
1.02/0.01/0.01/1.0	Prototype algorithm.

Contents

Chapter 1

Introduction

1.1 Purpose of this document

This document is the Algorithm Theoretical Baseline Document of the GOME-2 Sun-Induced Fluorescence products, which is part of the AC SAF atmospheric composition product family. This document presents the scientific background of the algorithm, presents an outline of its implementation and assists the user of this product in its physical interpretation. Knowledge of the sensitivity of the SIF to its dependencies (such as cloud presence, surface pressure, and surface albedo scattering geometry, etc) is a vital ingredient for the interpretation of the SIF values. Therefore we also present a sensitivity study.

1.2 Scope

This ATBD provides information on the SIF retrieval background and product.

1.3 Heritage

This algorithm was developed at KNMI (the Royal Netherlands Meteorological Institute) in the MSc thesis work of Tommy van Leth [*van Leth et al.*, 2014], Maurits Kooreman [*Kooreman et al.*, 2015] and Erik van Schaik [*van Schaik et al.*, 2016].

This development was continued by the Satellite Application Facility on Ozone and Atmospheric Chemistry Monitoring (O3M SAF), renamed to Atmospheric Composition Monitoring SAF (AC SAF) as of March 2017 at the beginning of CDOP-3.

1.4 Glossary

1.4.1 Acronyms and abbreviations

Table 1.1: Acronyms and abbreviations.

ACSAF	Satellite Application Facility on Atmospheric Composition Monitoring
ATBD	Algorithm Theoretical Basis Document
CDOP	Continuous Development and Operations Phase
DFS	Degrees of Freedom for Signal
ECMWF	European Centre for Medium-range Weather Forecast
EOF	Empirical Orthogonal Function
EPS	EUMETSAT Polar System
ERS	European Remote Sensing Satellite
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FLEX	FLuorescence EXplorer
FRESCO	Fast Retrieval Scheme for Cloud Observables
FWHM	Full Width Half Maximum
GOME-2	Global Ozone Monitoring Instrument 2 (on METOP)
GOSAT	Greenhouse gasses Observing SATellite
HDF	Hierarchical Data Format
(N)IR	(Near-)Infrared
KNMI	Royal Netherlands Meteorological Institute
METOP	Meteorological Operational Satellite
MSC	Main Science Channels
NC	NetCDF Data Format
NRT	Near Real Time
OCO	Orbiting Carbon Observatory
OE	Optimal Estimation
OMI	Ozone Monitoring Instrument
OPF	Output Product Format
PC	Principal Component
PSC	Polar Stratospheric Clouds
PUM	Product User Manual
RMS	Root Mean Square
RTM	Radiative Transfer Model
SAA	Solar Azimuth Angle
SAF	Satellite Application Facility
SCIAMACHY	Scanning Imaging Absorption spectroMeter for Atmospheric ChartographY

Table 1.1 – *Continued from previous page*

SRD	Software Requirements Document
SUM	Software User Manual
SW	Software
SZA	Solar Zenith Angle
TOA	Top Of Atmosphere
UV	Ultra Violet
VAA	Viewing Azimuth Angle
VIS	Visible
VZA	Viewing Zenith Angle

Chapter 2

Introduction to EUMETSAT Satellite Application Facility on Atmospheric Composition Monitoring (AC SAF)

2.1 Background

The need for atmospheric chemistry monitoring was first realized when severe loss of stratospheric ozone was detected over the polar regions. At the same time, increased levels of ultraviolet radiation were observed.

Ultraviolet radiation (UV) is known to be dangerous to humans and animals (causing e.g. skin cancer, cataract, immune suppression) and having harmful effects on agriculture, forests and the oceanic food chain. In addition, the global warming – besides affecting the atmospheric chemistry – also enhances the ozone depletion by cooling the stratosphere. Combined, these phenomena have immense effects on the whole planet. Therefore, monitoring the chemical composition of the atmosphere is an important activity for EUMETSAT and the world-wide scientific community.

2.2 Objectives

The main objectives of the AC SAF are to process, archive, validate and disseminate atmospheric composition products (O_3, NO_2, SO_2, OClO, HCHO, BrO, H_2O), aerosols, surface ultraviolet radiation and auxiliary products utilising the satellites of EUMETSAT. The majority of the AC SAF products are based on data from the GOME-2 spectrometer and the IASI interferometer onboard the MetOp satellite series. Another important task of the AC SAF is the research and development in radiative transfer modeling and inversion methods for obtaining long-term, high-quality atmospheric composition products from the satellite measurements.

2.3 Product families

The AC SAF products are grouped into different families: total columns of trace gases, vertical profiles of trace gases, aerosol products and land surface products and UV dose products. An overview is given in Table 2.1.

Table 2.1: AC SAF Product families.

Near real-time	Total Columns: O_3, NO_2, O_3Tropo, NO_2Tropo, SO_2
	Vertical Profiles: Vertical Ozone Profile
	UV Index
	Absorbing Aerosol Index, ash
	Sun-Induced Fluorescence
Offline	Total Column: O_3, NO_2, O_3Tropo, NO_2Tropo, SO_2, BrO, H_2O, HCHO, OClO
	Vertical Profiles: Vertical Ozone Profile
	UV Index
	Absorbing Aerosol Index
	Sun-Induced Fluorescence
Data Record	Total Column: O_3, NO_2, O_3Tropo, NO_2Tropo, SO_2, BrO, H_2O, HCHO, OClO
	Vertical Profiles: Vertical Ozone Profile
	UV Index
	Absorbing Aerosol Index
	Surface Properties

2.4   Product timeliness and dissemination

Data products are divided in a few categories depending on how quickly they are available to users. See Table 2.1:

- Near real-time products: these are available in less than three hours after measurement. These products are disseminated via EUMETCast (such as NHP, NTO, NAP), the GTS (also NHP, NTO) or the internet (like NUV).

- Offline products: these are available within two weeks from the measurement and they are archived at the AC SAF archives at the Finnish Meteorological Institute (such as OHP, OUV, ARP) and the German Aerospace Center (such as OTO and related total columns). The products can be ordered via the EUMETSAT Data Centre (EDC).

- Data records: these products are 'static' in the sense that they cover a certain period and are produced once and are not updated as more data comes in. These products can be superseded by more recent versions. The Data Record products can also be ordered via the EUMETSAT Data Centre (EDC).

Products with "pre-operational" or "operational" status are disseminated to users. The up-to-date status of all the AC SAF products and ordering info is available on the AC SAF website.

2.5   Further information

Information about the AC SAF project in general, its NRT, Offline or Data Record products and its services can be found on the AC SAF website: http://acsaf.org/

The AC SAF Helpdesk can be contacted via: helpdesk@acsaf.org

Chapter 3

Sun-Induced Fluorescence: background and overview

Human activities have dramatically influenced the global carbon cycle since the industrial revolution in the second half of the 18th century. The global average concentration of atmospheric carbon dioxide (CO_2) reached 390.5 ppm in 2011 and exceeded the 400 ppm level since; 40% higher than in 1750 [*Hartmann et al.*, 2013]. Higher atmospheric CO_2 concentrations and the associated increase of global mean surface temperatures are expected to affect many aspects of the terrestrial system, possibly for centuries to come [*Solomon et al.*, 2009].

Emitted CO_2 is not permanently stored in the atmosphere. Marine and terrestrial sinks remove approximately 55% of the emitted CO_2 (e.g. *Le Quéré et al.* [2015]). The main player in this process is the terrestrial biosphere, which assimilates atmospheric CO_2 during photosynthesis. The amount of carbon assimilation, or gross primary productivity (GPP), is the most uncertain flux within the global carbon budget [*Ciais et al.*, 2013]. By means of autotrophic respiration, vegetation also puts carbon back into the atmosphere resulting in a net ecosystem exchange (NEE), which is defined as GPP minus TER (total ecosystem respiration). This net flux can be measured locally (e.g. by eddy-covariance flux towers). Computation of global carbon fluxes relies on input from a flux tower network, but there are only few such towers, especially in highly productive regions like the tropics. Necessary interpolation of the tower measurements results in uncertainties for global GPP estimates. Currently, model estimates range from 130 to 169 PgC yr^{-1} [*Anav et al.*, 2015]. The uncertainty is significant considering that we eventually want to estimate NEE and need to take into account errors in the respiratory processes as well. Recently, GPP was estimated from long-term atmospheric carbonyl sulfide (COS) observational records, suggesting substantial growth of global GPP in the 20th century [*Campbell et al.*, 2017].

Remotely sensed vegetation fluorescence is a promising global proxy for GPP. During photosynthesis, sunlight is absorbed by chlorophyll pigments in leaves and converted into chemical energy. Approximately 80% of the harvested energy is used for photosynthesis. Most of the rest is dissipated non-radiatively as heat and a small fraction ($\sim$1%) is emitted at longer wavelengths as fluorescence. This Sun-Induced Fluorescence (SIF) has a spectrally smooth signature with peaks at 683 (red fluorescence) and 737 nm (far-red fluorescence). Chlorophyll itself re-absorbs fluorescence within the canopy at wavelengths below 700 nm. The red SIF emission peak can be measured from space ([*Joiner et al.*, 2016; *Verrelst et al.*, 2015]), but this product focuses on retrieving the far-red SIF emission peak.

Other measures of vegetation like the Normalized Difference Vegetation Index (NDVI) use the *greenness* of
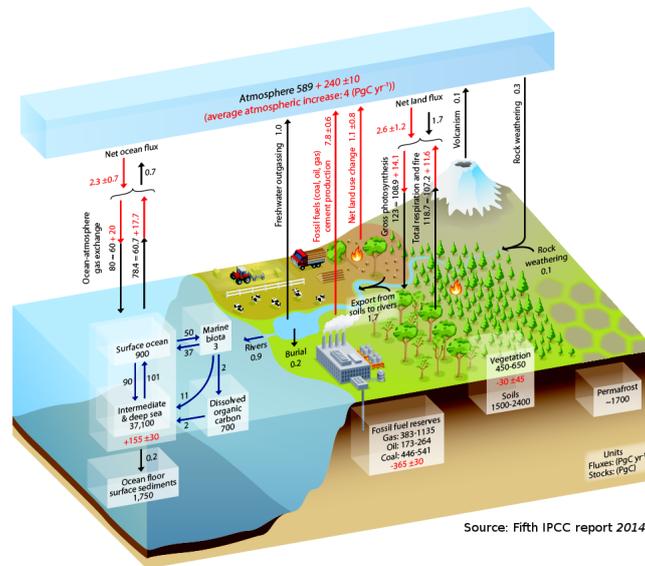
Figure 3.1: Illustration from the fifth IPCC report (2014) showing the different global carbon budgets. Gross photosynthesis is one of the major atmospheric carbon sinks.
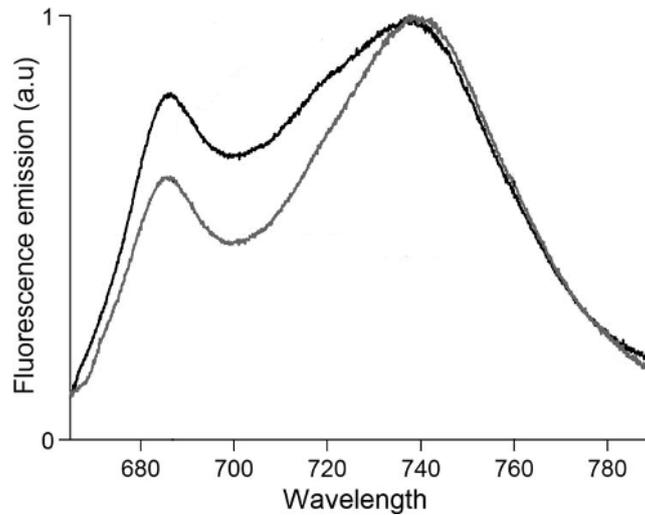

Figure 3.2: Chlorophyll fluorescence emission spectra at top of canopy for dense (grey) and sparse (black) leaf coverage. The discrepancy towards shorter wavelengths is caused by re-absorption of fluorescence by chlorophyll itself.

leaves as indicator of chlorophyll content. Unfortunately, chlorophyll content adapts only slowly to stress (it can take weeks for leaves to lose their green color) and is therefore not as useful as a proxy for direct GPP. As photosynthesis is the core mechanism that drives both carbon assimilation and fluorescence, fluorescence can be exploited to serve as a more direct proxy for GPP. Previous studies have shown that fluorescence and GPP correlate linearly at leaf level, even under drought conditions [*Schreiber and Bilger*, 1987; *Flexas et al.*, 2002].
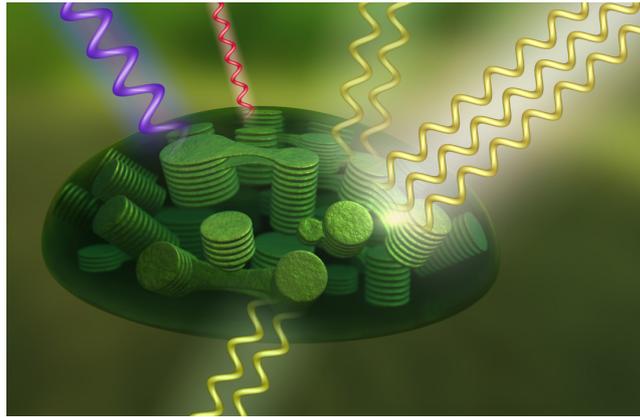
Figure 3.3: Schematic of chloroplast, the core of the photosynthetic mechanism. Approximately 80% of the harvested energy (bright minus dark yellow) is used for photosynthesis. Roughly 19% is dissipated non-radiatively as heat (purple) and a small fraction of is emitted at longer wavelengths as fluorescence (red). Credit: NASA

3.1 Measuring Fluorescence from Space

Space based spectrometers observing the near-infrared spectral region with sufficient spectral resolution and signal to noise ratio are able to detect the SIF signatures in top-of-atmosphere (TOA) reflectance. Following recent advancements in space borne observations, methods have been developed to disentangle the small SIF signal from earthshine spectra. Initially, the retrieval method relied on measuring the filling-in of individual solar Fraunhofer lines. This method requires a very high spectral resolution (typically < 0.05 nm) in order to resolve the fine individual Fraunhofer line structure and has been applied to space based instruments like the Greenhouse gasses Observing SATtelite (GOSAT) [*Joiner et al.*, 2011, 2012; *Frankenberg et al.*, 2011b] and the Orbiting Carbon Observatory 2 (OCO-2) [*Frankenberg et al.*, 2014]. However, it comes at the cost of sparse spatial sampling. To obtain accurate information on global and seasonal SIF, better spatial and temporal coverage is needed. This is achieved for instance by the SCanning Imaging Absorption spectroMeter for Atmospheric CHartography (SCIAMACHY) [*Joiner et al.*, 2013; *Khosravi et al.*, 2015] and the Global Ozone Monitoring Instrument 2 (GOME-2) [*Joiner et al.*, 2013; *Köhler et al.*, 2015; *Sanders et al.*, 2016]. Other instruments capable of retrieving SIF include the TROPOspheric Monitoring Instrument (TROPOMI) [*Guanter et al.*, 2015; *Köhler et al.*, 2018] and the dedicated FLuorescence EXplorer (FLEX) [*Kraft et al.*, 2012].

Chapter 4

Algorithm background

In this section, the theoretical derivation of the SIF retieval is given as well as the mathematical application to spectrometer data.

4.1 Retrieval Method

It is well possible to retrieve SIF with instruments originally designed to measure atmospheric composition. Instruments such GOME-2 and TROPOMI provide global coverage every day, but lack the spectral resolution to measure individual Fraunhofer lines explored from dedicated sensors such as OCO-2 [*Frankenberg et al.*, 2014; *Sun et al.*, 2018] and GOSAT [*Frankenberg et al.*, 2011a,b]. Studies showed that SIF can still be retrieved from moderate spectral resolution ($\sim$0.5 nm) spectrometers if a broad spectral fitting window is used, to observe multiple Fraunhofer lines. The approach relies on statistical methods such as Singular Value Decomposition (SVD) or Principal Component Analysis (PCA) to determine a representative background spectrum (e.g. *Guanter et al.* [2013]; *Joiner et al.* [2013]) based on an aggregate of reference spectra over non-fluorescent targets. Over vegetated areas, the method then assumes that the top-of-atmosphere radiance near 740 nm can be represented by a superposition of three radiative components; (effective) surface reflectance, atmospheric absorption and SIF. The Royal Netherlands Meteorological Institute (KNMI) recently developed a SIF retrieval approach for GOME-2 [*Sanders et al.*, 2016]. Continuing its development, sensitivity tests have been performed using the radiative transfer model DISAMAR. The tests revealed algorithm dependencies (e.g. on water vapor, spectral retrieval window, reference sector), which were subsequently analyzed to determine the optimal configuration for SIF retrieval. The development led to an improved retrieval algorithm [*van Schaik et al.*, 2016; *Botía et al.*, 2017; *Koren et al.*, 2018], that has been applied to GOME-2.

4.2 Forward model

The TOA reflectance as function of wavelength λ can be approximated using a Lambertian model which includes an atmosphere bounded from below by an isotropically reflecting surface. Superimposed on this model is the fluorescence radiance field emitted from the surface, which is also assumed to be isotropic.

$$R(\lambda,\mu,\mu_0) = \rho_0 + \frac{A_{\mathrm{s}}(\lambda)}{1-A_s(\lambda)\bar{\rho}}T^{\downarrow}(\lambda,\mu)T^{\uparrow}(\lambda,\mu_0) + \frac{\pi I_{\mathrm{F}}}{[1-A_s(\lambda)\bar{\rho}]E(\lambda)\mu_0}T^{\uparrow}(\lambda) \qquad (4.1)$$

Here, ρ_0 is the atmospheric reflectance contribution in absence of a boundary (i.e. the surface is black), A_{s} is the scene albedo, and $\bar{\rho}$ is the spherical albedo of the atmosphere for illumination from below. $T^{\uparrow}$ and $T^{\downarrow}$ are the sun-to-ground and ground-to-sensor atmospheric transmissions respectively. I_{F} is the SIF emission from the surface which is scaled to $E(\lambda)$, the solar irradiance, in order to comply as a reflectance term. μ and μ_0 are the cosines of respectively the viewing and solar zenith angles.

Atmospheric scattering is mostly caused by the presence of molecular nitrogen and oxygen. The Rayleigh scattering cross-section is a smooth function, strongly dependent on wavelength (λ^{-4}) and is as such, small in the NIR wavelength region. This implies that $A_s\bar{\rho} \ll 1$. Following this assumption and assuming that the remaining atmospheric reflectance can be described as part of the scene albedo, Equation 4.1 reduces to

$$R(\lambda,\mu,\mu_0) = A_{\mathrm{s}}(\lambda)T^{\downarrow}(\lambda,\mu)T^{\uparrow}(\lambda,\mu_0) + \frac{\pi I_{\mathrm{F}}(\lambda)}{E(\lambda)\mu_0}T^{\uparrow}(\lambda,\mu). \qquad (4.2)$$

For conversion to a numerical model, it is convenient to combine $T^{\downarrow}$ and $T^{\uparrow}$ and write them in terms of optical thickness τ.

$$T^{\downarrow\uparrow} = e^{-\tau_v\left(\frac{1}{\mu}+\frac{1}{\mu_0}\right)} = e^{-\tau^{\downarrow\uparrow}}$$

Here $\tau^{\downarrow\uparrow}$ is the two-way slant absorption optical thickness and τ_v is the vertical absorption optical thickness. When writing $\tau^{\uparrow}$ in terms of $\tau^{\downarrow\uparrow}$ an expression for $T^{\uparrow}$ follows[1].

$$T^{\uparrow} = e^{-\tau_v\frac{1}{\mu}} = e^{-\tau^{\downarrow\uparrow}\frac{\mu^{-1}}{\mu^{-1}+\mu_0^{-1}}}$$

Equation 4.2 can now be written in a form which allows conversion to a numerical model.

$$R(\lambda,\mu,\mu_0) = A_{\mathrm{s}}e^{-\tau^{\downarrow\uparrow}} + \frac{\pi I_{\mathrm{F}}}{E(\lambda)\mu_0}e^{-\tau^{\downarrow\uparrow}\frac{\mu^{-1}}{\mu^{-1}+\mu_0^{-1}}}. \qquad (4.3)$$

---

[1]This requires the step: $\tau^{\uparrow} = \tau_v\frac{1}{\mu} \cdot \frac{\frac{1}{\mu}+\frac{1}{\mu_0}}{\frac{1}{\mu}+\frac{1}{\mu_0}} = \tau^{\downarrow\uparrow}\frac{\frac{1}{\mu}}{\frac{1}{\mu}+\frac{1}{\mu_0}}$

4.3 Application to spectrometer data

This section describes the conversion of the analytic model to a numerical model for the application to spectrometer data.

4.3.1 The GOME-2 Spectrometer

The MetOp satellite series fly in a sun-synchronous polar orbit at a height of 817 km with a local equator overpass time of 09h30 in the descending node. The GOME-2 instrument has four main channels that measure radiance between 240 and 790 nm at a resolution of 0.2-0.4 nm. For the purpose of SIF retrieval, only the band-4 data is used, ranging from 593 nm to 790 nm.

A scan mirror enables across-track scanning in nadir with a swath width of 1920 km. This allows the instrument to achieve global coverage in 1.5 days. It takes 6 sec. for the mirror to go forwards (east to west, 4.5 sec.) and backwards (west to east, 1.5 sec.). As it goes back three times as fast, the ground pixel size is much larger and so we only use data from the forward scan. Pixel size depends on the satellite velocity and the scan integration time. With an integration time of 0.1875 sec., the along-track pixel size is 40 km and the cross-track pixel size is 80 km, the spatial resolution is thus 40 km x 80 km [*Callies et al.*, 2000].

This has been the nominal operation mode for GOME-2A until 15 July 2013. After that, the swath width has been reduced to 960 km resulting a pixel size of 40 km x 40 km. Global coverage is now achieved in 3 days. GOME-2B and -C still operate in nominal mode.

4.3.2 Reflectance measurement $R(\lambda)$

In the application of remote sensing retrieval, the monochromatic reflectance $R(\lambda)$ is replaced by the measured GOME-2 reflectance $R(\lambda_i)$.

4.3.3 Solar irradiance $E_0(\lambda)$

The monochromatic solar irradiance $E_0(\lambda)$ is replaced by the high-resolution solar irradiance reference spectrum from *Chance and Kurucz* [2010], convolved with the GOME-2 instrument response function $F_R(\lambda_i)$ and scaled to the actual sun-earth distance, denoted as $\hat{E}_0(\lambda_i)$. The scaling component is based on an elliptical orbit with an eccentricity e of 0.0167 and is calculated for each day of year d. The values 365, 89 and 3 correspond to the number of days in a year, the day of the irradiance reference intensity of the high-resolution spectrum (March 29, 1992) and the approximate day of perihelion (January 3) respectively. Hence,

$$\hat{E}_0(\lambda_i) \approx (E_0(\lambda) * F_R(\lambda_i)) \cdot \frac{\frac{2\pi}{365}\cos(d-3)(1+e)}{[\frac{2\pi}{365}\cos(d-89)(1+e)]^2}. \tag{4.4}$$

Because radiometric degradation would largely cancel out in the calculation of the reflectance, the actual measured irradiance spectrum could also be used to scale the modeled SIF term to reflectance. However, this would show up as a confounding trend in the SIF time series.

4.3.4 SIF emission I_F

The upwelling SIF radiance is parameterized as Gaussian curve with its peak located at 737.0 nm (μ_F) and a sigma-width of 33.9 nm (σ_F) following *Zarco-Tejada et al.* [2000]. Hence,

$$I_F(\lambda_i) \approx I_{F_0} e^{-\frac{1}{2}\left(\frac{\lambda_i - \mu_F}{\sigma_F}\right)^2}, \tag{4.5}$$

where I_{F_0} is the SIF strength at 737 nm in Watts per square meter per steradian per nanometer (W m^{-2} sr^{-1} nm^{-1}).

4.3.5 Scene albedo A_s

The scene albedo is calculated by fitting a low-order polynomial function to the top of atmosphere reflectance spectral micro-windows where the atmospheric absorption is close to zero (i.e. where all light comes from the surface): 712 - 713 nm, 748 - 757 nm and 775 - 785 nm.

$$A_s \approx \sum_{j=0}^{n} a_j \lambda_i^j \tag{4.6}$$

In the actual retrieval, $n=4$. This order is required to model the spectral variability of vegetated surfaces known as the red-edge (see e.g. *Tilstra et al.* [2017]).

4.3.6 Two-way slant absorption optical thickness $\tau_s^{\uparrow\downarrow}$

The two-way slant absorption optical thickness is determined in a statistical way. Reference scenes are selected that contain no vegetation (thus, I_F=0) and the last term in Equation 4.3 cancels. The two-way slant optical thickness $\tau_s^{\uparrow\downarrow}$ can now be written as

$$\tau_s^{\uparrow\downarrow}(\lambda) \approx -\ln\left(\frac{R_{\text{ref}}(\lambda; \mu, \mu_0)}{A_s(\lambda)}\right), \tag{4.7}$$

where R_{ref} is the reflectance of the reference scene. Next, a number of small spectral windows is selected for which the absorption optical thickness becomes very small (i.e. all radiation comes from the surface) to determine the scene albedo. Selected sub-windows are 712-713 nm, 748-757 nm and 775-783 nm. A second order polynomial (Equation 4.6; $n=2$) is fitted to the measured reflectance in these windows to calculate $A_s(\lambda)$.

GOME-2 spectra are selected over the Sahara region between latitudes 16N and 30N and longitudes 8W and 29E (see Figure 4.1) and included in a reference set only if they are absolutely free of vegetation according to the 1 km^2 resolution USGS Global Land Cover Characterization database (version 2; https://lta.cr.usgs.gov/GLCC) and if the scene effective cloud fraction in the infra-red is below 0.4. The cloud fraction is taken from the GOME-2 FRESCO cloud retrieval [*Koelemeijer et al.*, 2001; *Wang et al.*, 2008, 2010].
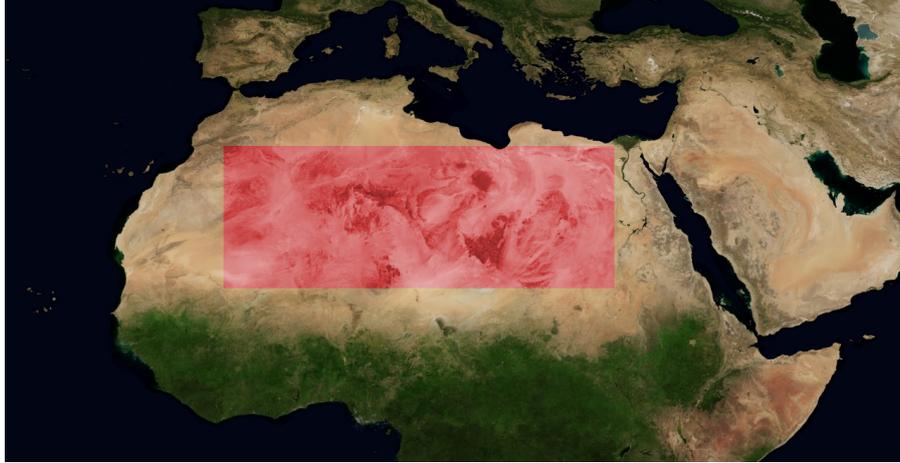
Figure 4.1: Measurements within the red section of the Sahara desert are checked for presence of vegetation using a land cover database. If they are completely devoid of vegetation, they are used to build a reference atmospheric slant optical thickness.

On an aggregate of reflectances from all scenes that match the criteria, principal component (PC) analysis is performed using the NIPALS algorithm [*Geladi and Kowalski*, 1986] resulting in a number of Empirical Orthogonal Functions (EOFs) $f_k(\lambda)$. These functions are the statistical representation of the two-way slant optical thickness. The EOFs describe the spectral variability due to absorption and light path geometries as well as instrumental features (stray light, dark current, etc.). It is assumed that this set of basis EOFs can be used to describe the atmospheric two-way slant absorption optical thickness over all target scenes, including scenes with vegetation as

$$\tau_s^{\uparrow\downarrow}(\lambda_i) \approx \sum_{k=0}^{m} b_k f_k(\lambda_i) \tag{4.8}$$

Where b_k is the principal component associated with its corresponding f_k.

The forward model for the fluorescence retrieval is now completed and written as

$$R(\lambda;\mu,\mu_0) \approx \left(\sum_{j=0}^{n} a_j \lambda_i^j \right) e^{-\sum_{k=0}^{m} b_k f_k(\lambda_i)} + \frac{\pi I_{F_0} e^{-\frac{1}{2}\left(\frac{\lambda_i-\mu_F}{\sigma_F}\right)^2}}{\mu_0 \hat{E}_0(\lambda)} e^{-\frac{\mu^{-1}}{\mu^{-1}+\mu_0^{-1}} \sum_{k=0}^{m} b_k f_k(\lambda_i)}. \tag{4.9}$$

The final inversion applies a Levenberg-Marquardt least-squares regression to minimize the differences between the observed and modeled reflectance in the 734-758 nm window, by optimizing for 1+5+10=16 ($I_{F_0} + a_j + b_k$) free fitting parameters. A schematic overview of the retrieval is shown in Figure 4.2
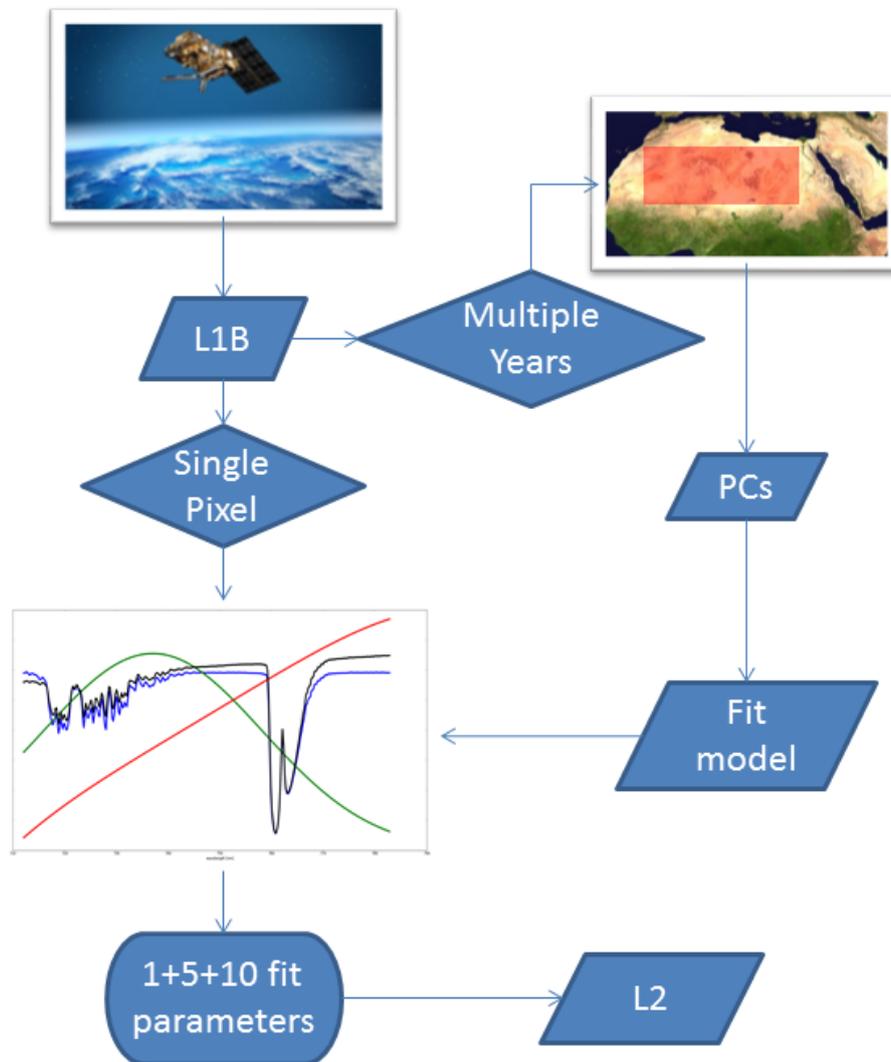
18

4.3.7 Algorithm flow diagram

Figure 4.2: Flowchart of the SIFTER retrieval algorithm.

4.4 Motivation for the selection of retrieval parameters

Understanding of the atmospheric transmittance term ($\tau^{\downarrow\uparrow}$, $\sum_{k=0}^{m} b_k f_k(\lambda_i)$), is crucial for the fluorescence retrieval. To test our understanding of this parameter, simulated top-of-atmosphere spectra for a wide range of conditions were generated with the DISAMAR radiative transfer model [*de Haan*, 2011], developed at KNMI (based on Doubling-Adding KNMI, DAK). DISAMAR generates TOA reflectance spectra, and allows convolution of a simulated reflectance spectrum with the pre-defined GOME-2 slit function. Simulation with DISAMAR allows control over important retrieval variables such as atmospheric water vapor content, the effects of oxygen absorption, surface albedo, and viewing geometries. Also, a top-of-canopy SIF signal can be included, to test which retrieval settings (spectral fitting window, number of PCs) best reproduce the magnitude of the SIF under different circumstances in a so-called end-to-end test.

In the baseline test, 2000 DISAMAR TOA spectra were generated[2] in the range 712 nm to 783 nm for different prescribed parameters (albedo, surface pressure) and viewing geometries representative for those encountered by GOME-2. The prescribed parameters were obtained from a random draw within the parameter value ranges. In the baseline test there was no SIF, as the intention was to simulate reference spectra to be used for the determination of the PCs.
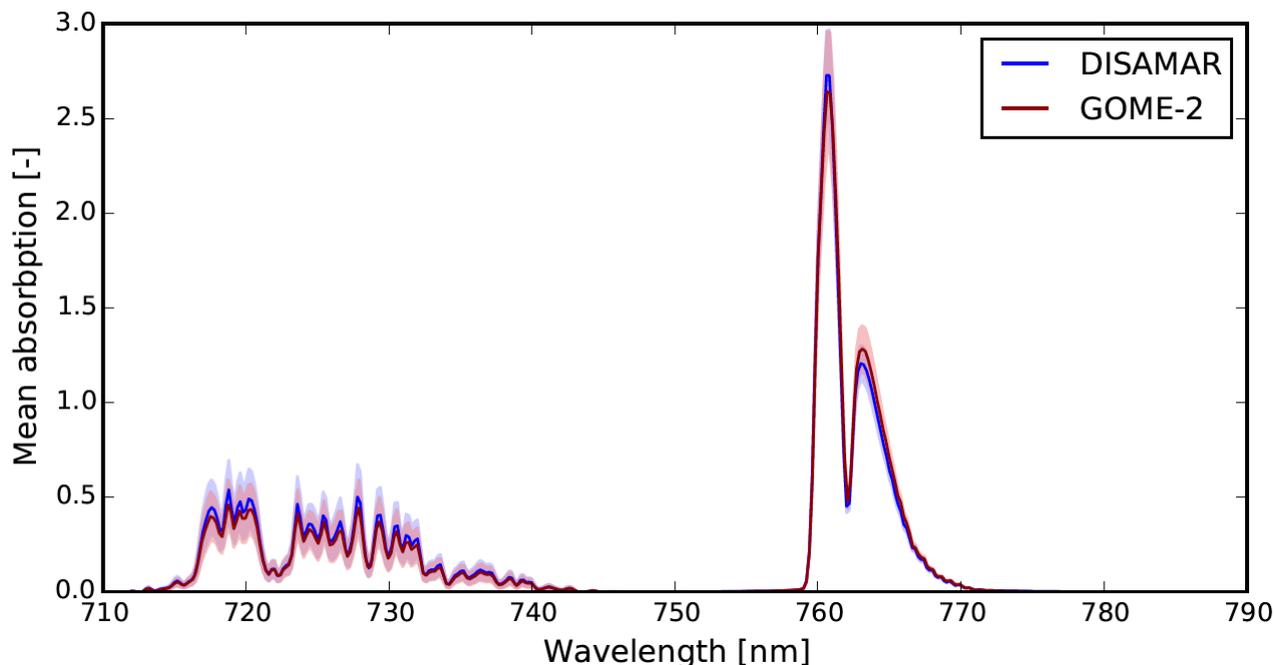
Figure 4.3: Mean top-of-atmosphere absorption for GOME-2 obervations (red, N≈30000) and DISAMAR generated spectra (blue, N≈2000). Shading shows the 1-σ deviation.

DISAMAR TOA reflectance spectra simulated for the baseline case were evaluated against an ensemble of observed GOME-2A spectra over the Sahara. Figure 4.3 compares the mean DISAMAR and GOME-2 atmospheric absorption signatures in the 712-783 nm range. The good agreement between the two spectra suggests that DISAMAR is capable of capturing the various relevant atmospheric and surface processes that drive the reflectance signal, providing confidence in DISAMAR as a test tool. The PCs are calculated in order of most explained variance. Comparing PCs obtained from the DISAMAR and GOME-2A spectra, we found that the

---

[2]A plane-parallel atmosphere was used here, justified by the moderate range of viewing geometries.

two most relevant PCs (1: the mean, 2: dominated by water vapor absorption), are highly similar (Figure 4.4). Higher-order PCs show stronger differences, probably related to fundamental differences between the simulations (cloud-free) and observations (cloud fractions up to 0.4), and to other small differences in spectral behavior such as unaccounted atmospheric absorption signatures, surface effects, or instrumental features such as stray light, wavelength calibration etc.
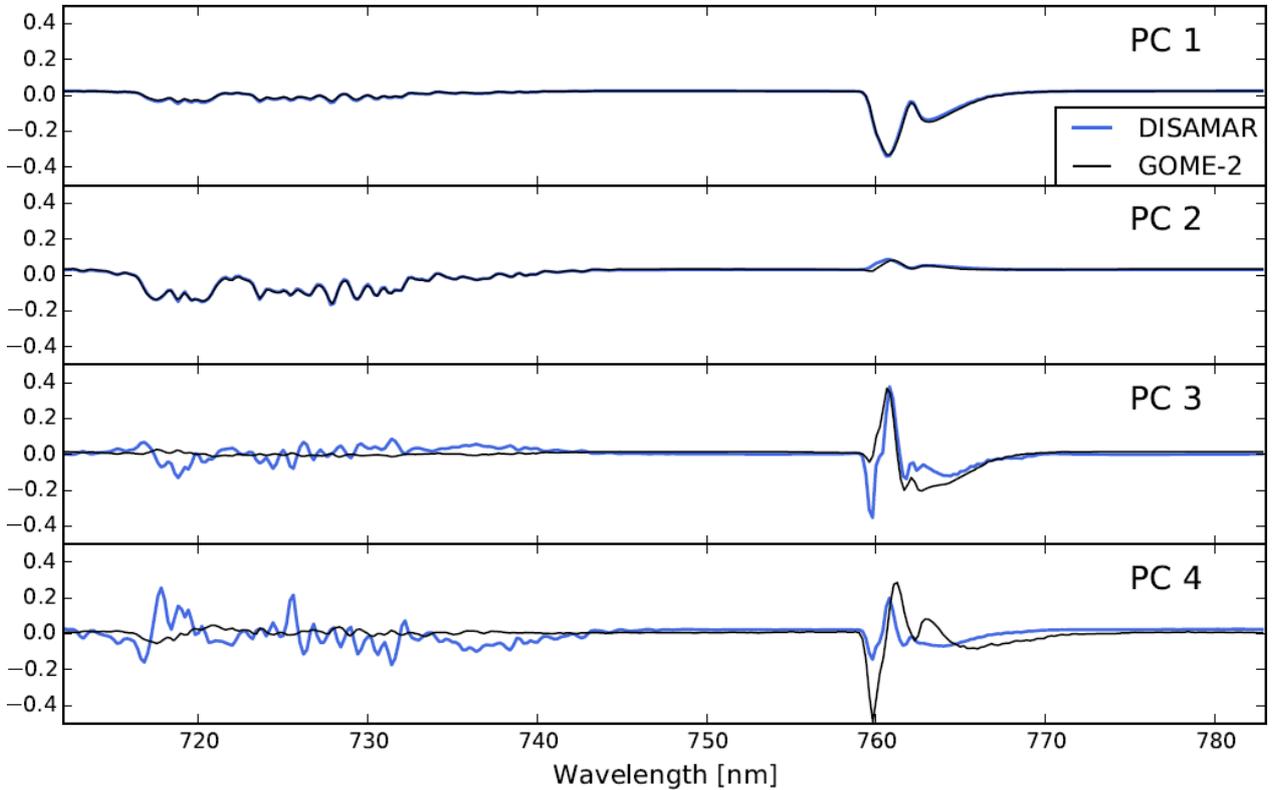

Figure 4.4: Mean of the first four PCs from the DISAMAR and GOME-2 reference sets. The first two components show very good agreement. Higher order components, which explain less and less of the spectral variability, deviate more significantly.

4.5 Experiments

Five experiments focused on addressing a particular aspect of the SIF retrieval (influence of water vapor absorption, albedo characteristics, and viewing geometry), and included various strengths of fluorescence. For these experiments, 1000 TOA spectra were generated. The overall aim of the experiments is to establish the optimal spectral window and number of PCs with which to retrieve SIF. This is done by evaluating four different spectral windows: a wide window stretching out from 712-783 nm and three smaller sub-windows that (1) excludes the main absorption signatures from the O_2-A band (712-758 nm), (2) from water vapor (734-783 nm), and (3) avoids both bands (734-758 nm).

Table 4.1 summarizes the settings used in the experiments, and their purpose. The FLUOR experiment has settings that are identical to the baseline case, and allows us to investigate to what extent SIF can be reproduced in optimal conditions, as if fluorescence would be monitored over a Sahara-type surface. The WATER exper-

iment replaces the water vapor columns over the Sahara by amounts that are more representative for tropical rain forests (ERA-Interim fields, *Berrisford et al.* [2011]). This allows us to assess the influence of strong water vapor absorption not fully represented in the set of PCs on the SIF retrieval. The two vegetation experiments evaluate the sensitivity of SIF retrieval to different surface albedos that are more representative for vegetated areas than in the Sahara reference region. In the last experiment (LPATH), the viewing geometry is changed to angles found over a Russian boreal forest (55°-65° N) on a late summer day.

Table 4.1: Settings for the six experiments with DISAMAR. All experiments were done without clouds, and signal-to-noise ratio of the simulated reflectance spectra was 1000.

Experiment	SZA range	VZA range	H_2O column (kg m^{-2})	Surface albedo	Fluorescence (mW m^{-2} sr^{-1} nm^{-1})	Purpose
FLUOR	21.4°-66.8°	0.0°-53.8°	4.0-40.0	0.41-0.45	0.0-4.0	Reproduce SIF in ideal conditions
WATER	21.4°-66.8°	0.0°-53.8°	30.0-65.0	0.41-0.45	0.0-4.0	Test influence of too low H_2O in PCs and SIF retrieval
ALBEDO	21.4°-66.8°	0.0°-53.8°	4.0-40.0	0.4	0.0-4.0	Test influence of constant spectral albedo in PCs and SIF retrieval
RED EDGE	21.4°-66.8°	0.0°-53.8°	4.0-40.0	0.06-0.51	0.0-4.0	Test influence of realistic red edge spectral albedo in PCs and SIF retrieval
LPATH	54.4°-69.6°	0.0°-53.8°	4.0-40.0	0.41-0.45	0.0-4.0	Test whether SIF can still be retrieved for viewing conditions at high latitudes

In the (Saharan) FLUOR experiment, we tested to what extent the fluorescence assumed in the generation of the TOA spectra could be reproduced with the retrieval approach described in Section 4.2. Table 4.2 summarizes the results of this experiment in terms of retrieval bias and uncertainty. The results from this end-to-end test strongly suggest that limiting the spectral fitting window, and thereby minimizing the number of PCs required, leads to the most accurate reproduction of the fluorescence signal. Clearly, the combination of a narrow fitting window (734-758 nm) and a low number of PCs (8) gives an unbiased retrieval and lowest uncertainties (approximately 25%).

As a further retrieval test, the goodness-of-fit was evaluated, using the concept of spectral autocorrelation. For every spectrum, the residuals of the fit remaining after minimizing the differences between the DISAMAR spectrum and the modeled spectrum from Equation 4.3 were analyzed. In general, high fit residual values are considered to be indicative of a poor fit, but the degree of spectral lag-one autocorrelation in the residuals is another indicator as it tests the non-randomness of the residuals [NIST/SEMATECH, 2018]. If the modeled spectrum perfectly explains the TOA spectrum, fit residuals appear as noise with zero autocorrelation, but if the model fails to resolve structural components of the TOA spectrum, this will show up in the fitting residuals, and contributes to autocorrelation. The tests suggest a strong relationship between the bias in the retrieval and

the amount of autocorrelation in the residual spectrum. This relationship is apparent for all selections of fitting windows and number of PCs. For the sensitivity tests, retrievals are characterized as faulty when spectral autocorrelation in the fit residuals is larger than 0.2.

Table 4.2: Results of the FLUOR experiment (mean of 1000 spectra) to reproduce fluorescence for different fitting windows and number of PCs used. The bias is defined here as the mean of the differences between assumed and retrieved fluorescence strength (on average 1.5 mW m^{-2} sr^{-1} nm^{-1}), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE.

	Number of PCs	Bias (mW m-2 sr-1 nm-1)	RMSE (mW m-2 sr-1 nm-1)	Faulty
712-783 nm (full spectral window)	8	-0.47 (-31%)	0.71	34.7%
	20	-0.37 (-25%)	0.59	30.6%
	35	-0.37 (-25%)	0.62	28.8%
712-758 nm (excluding O$_2$-A band)	8	-0.20 (-13%)	0.46	23.6%
	20	-0.23 (-15%)	0.50	20.1%
	35	-0.26 (-17%)	0.53	18.2%
734-783 nm (exclude H$_2$O band)	8	-0.49 (-33%)	0.75	41.9%
	20	-0.48 (-32%)	6.41	37.9%
	35	-0.25 (-17%)	0.49	35.4%
734-758 nm (exclude both bands)	8	0.0 (0%)	0.39	16.5%
	20	0.04 (3%)	0.43	16.2%
	35	-0.03 (-2%)	0.5	14.6%

The WATER investigates which fitting window is optimal when higher water vapor concentrations are encountered than the concentrations present in the Sahara reference region. Again, the retrieval in the narrow fitting window (734-758 nm) performs best in reproducing the prior fluorescence levels (bias <10%), albeit with a considerable fraction of retrievals with autocorrelation > 0.2 (40-65%). Retrievals with other fitting windows all performed worse in reproducing the fluorescence signal, reflecting that the relatively strong water vapor signatures around 730 nm lead to more unexplained structure in the fit residuals. The other three experiments (ALBEDO, REDEDGE and LPATH) led to very similar conclusions: fluorescence is best reproduced with a 734-758 nm window, and with 8 PCs. Inclusion of a distinct red edge in the spectral surface albedo (700-730 nm) did not affect the small fitting window very much, whereas fits with the wider fitting windows suffered to reproduce the assumed fluorescence levels. One notable result is that changing the viewing geometries to those representative of a late-summer boreal morning over Russia, did not strongly affect the retrieval results either. This suggests that a reference set based on the relatively small range of solar zenith angles encountered over the Saharan reference region can still be used for retrieving fluorescence over high-latitude regions. Table 4.2 presents the overview of the main results for the narrow fitting window.

To summarize, the tests show that fluorescence is consistently reproduced with a narrow fitting window (734-

758 nm) and 8 PCs under a wide range of different, but realistic retrieval conditions. The 734-758 nm window is similar to the fitting window selected by *Joiner et al.* [2016]. The test results also suggest that it is important to filter out retrievals with persistent spectral structure in their residuals. This can be done via the auto-correlation test proposed here. Higher amounts of atmospheric water vapor than present over the Sahara may lead to absorption imprints in the TOA spectra which are not captured by the PCs, possibly leading to a relatively large fraction of faulty retrievals. This indicates that defining a reference sector to include scenes with similar water vapor amounts as encountered over vegetated areas has good potential to improve the retrieval. Limiting the spectral range of the fitting window by excluding absorption bands is effective in reducing the number of PCs needed, both for GOME-2A and DISAMAR spectra. Including the bands with O_2-A and water vapor absorption signatures strongly reduces the explained variance of principal components 1 to 4. Wider spectral fitting windows thus require more PCs to capture atmospheric effects.

The experiments presented here provide us with an estimate of the optimal retrieval settings. The settings used in the operational processing are provided in Table 4.3. Note that the number of PCs is increased from 8 to 10 to allow for slightly more variability due to an increase in the reference set period (from one to five years).

Table 4.3: Settings for the operational algorithm.

Model construction	
Solar model	*Chance and Kurucz* [2010]
Earth eccentricity	0.01671022
Day of Perihelion DOY	January 3
SIF Gaussian center	737 nm
SIF Gaussian one-sigma	33.9 nm
PC set area	Sahara
PC set temporal range	23.01.2007 - 31.12.2012 (GOME-2A)
	TBD (GOME-2B)
	TBD (GOME-2C)
PC set transparent windows	712-713 nm, 748-757 nm, 775-785 nm
Albedo polynomial order	2

Retrieval	
Spectral window	734-758 nm
Number of PCs	10
Albedo polynomial order	4

4.6 The origin of the SIF signal

How good the model can represent the observed spectral signal can be evaluated by looking at the RMS differences between the observation and the model fit. To analyze the importance of the addition of the SIF term, retrievals are performed with a model including and excluding the SIF term.
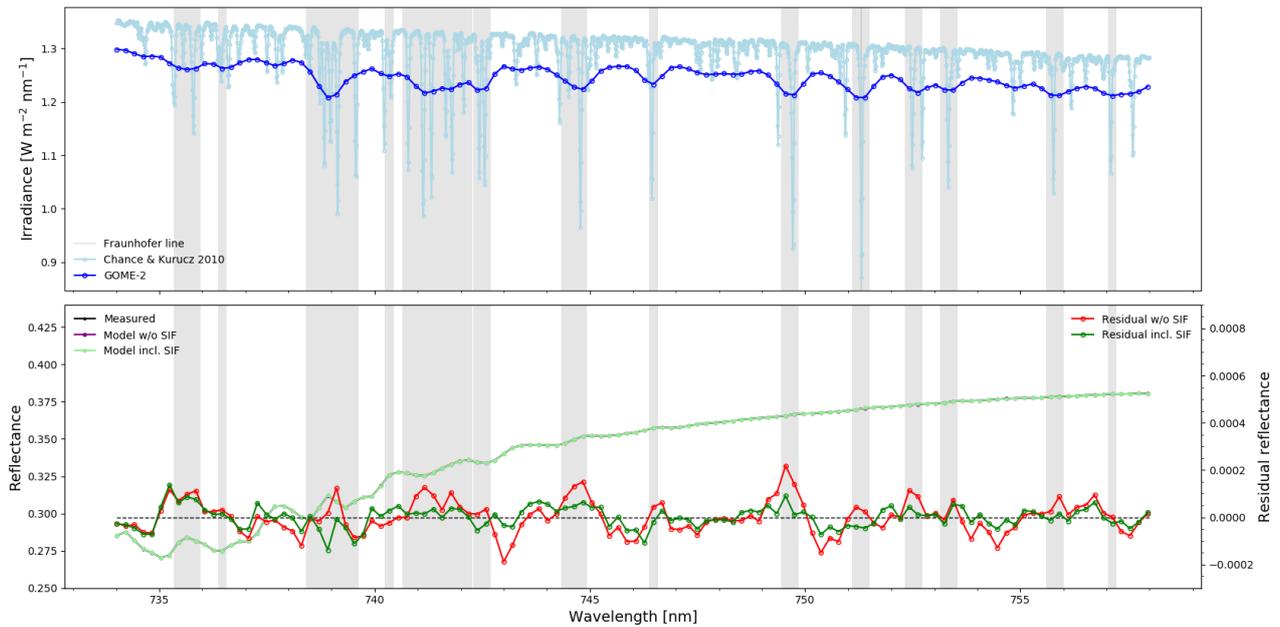
Figure 4.5: The lower panel shows the measured and modeled (with and without accounting for fluorescence) reflectance spectra in the fitting window (734-758 nm) on 15 July 2007 over vegetated areas in the eastern United States (orbit 3821). The spectra represent the average over 600 ground pixels with a fluorescence signal strength > 1.0 mW m^{-2} sr^{-1} nm^{-1}, a cloud fraction < 0.4 and a succesful fit ($\chi^2 < 2$). The measured and modeled spectra are indistinguishable, so that only the light green line is visible. The dark green and red lines show the raw reflectance residuals for a fit including and excluding fluorescence. The upper panel shows in light blue the 0.01 nm resolution solar irradiance spectrum and in dark blue the irradiance spectrum convolved with the GOME-2A slit function. The grey lines connecting the lower and upper panel indicate the spectral regions where solar Fraunhofer line infilling by fluorescence is relevant.

As can be seen from Figure 4.5, residuals are overall larger when the model does not include a SIF term. In this case, the largest difference in residuals coincide on many occasions with the presence of solar Fraunhofer lines in the spectrum. This tells us that the majority of the SIF signal that is retrieved originates from the in-filling of these absorption features.

Chapter 5

Data correction and filtering

As other spaceborne sensors measuring reflected sunlight, GOME-2A is sensitive to degradation because of prolonged exposure to solar radiation, contamination and icing. After the swath width reduction, it is unknown to what extent sensor degradation has affected the quality of the level-1 data in the NIR spectral channel. Furthermore, the level-1b data available until now has been processed with different processor versions by EUMETSAT. Until 17 June 2014, level-1b data was reprocessed with processor version 5.3, but since then a number of processor revisions have taken place (see Table 5.1).

Table 5.1: Major GOME-2 level-1b processor changes.

Period	L1B processor version	Main changes	Reference
23.01.2007 - 17.05.2014	5.3	All available level-1b data up to this date calibrated with this processor	Newsletter #29 [EU-METSAT, 2015]
18.05.2014 - 25.06.2015	6.0	Provision of additional cloud information from the AVHRR cloud mask on PMD read-out level.	Newsletter #35 [EU-METSAT, 2015]
26.05.2015 - 11.01.2018	6.1	Introduction of in-flight derived BSDF for solar radiometric calibration affecting the solar irradiance data. May have resulted in changes of the radiometric accuracy to unknown extent.	Newsletter #36 [EU-METSAT, 2015]
12.01.2018 - 17.12.2018	6.2	Implementation of the solar model for the solar visibility gaps. Visibility gaps are due to the instrument drift and will recur in the future.	
17.12.2018 - ongoing	6.3	A spike removal scheme was implemented to reduce spikes in the NO_2 fitting window. Negligible changes are expected in the NIR wavelength region.	Newsletter #38 [EU-METSAT, 2018]

5.1 Degradation Correction
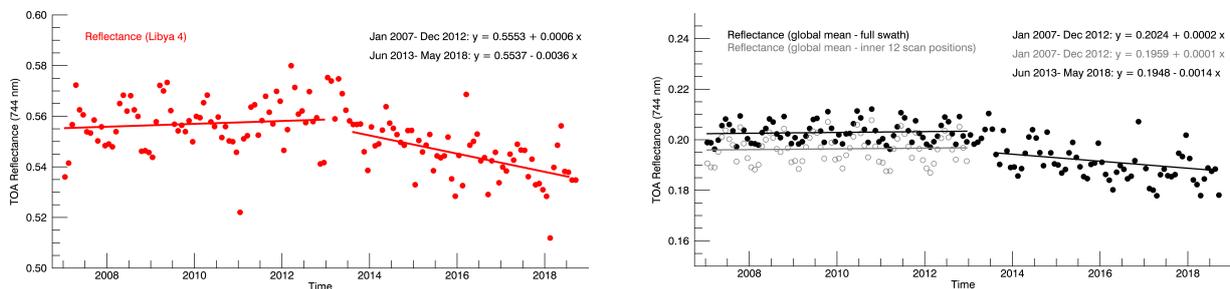
Figure 5.1: Left image: Monthly mean reflectance over the Libya4 calibration site measured by GOME-2A. Right image: Global monthly mean reflectance measured by GOME-2A. The grey dots indicate the reflectance level if only the inner 12 scan positions are taken into account. In this case the reflectance level is similar to the reflectance after the swath reduction in July 2013.

To assess if degradation or level-1 processing has influenced the stability and usefulness of the GOME-2A level-1b data, we track the variability of the top-of-atmosphere reflectances at 744 nm between 2007 and 2018. Figure 5.1 shows the time series of monthly mean reflectances over the stable reference calibration site Libya4[1] and averaged over the globe (between 60S and 60N). Top-of-atmosphere reflectances are relatively stable for the early GOME-2A with insignificant trends for the Libya4 and global mean reflectances for the period January 2007-December 2012. The late GOME-2A however shows an obvious negative trend in top-of-atmosphere reflectance over Libya4 and in the global mean. The reflectance trends are consistent and amount to -0.65%/yr over Libya4, and -0.71%/yr averaged over the globe for August 2013 - July 2018.

The apparent 'jump' between the baseline reflectance levels in the early and late GOME-2A can be understood from the smaller range of viewing zenith angles encountered by the late GOME-2A. When we sample only the inner 12 scan positions (VZA < 35°) for the early GOME-2A period, we find a baseline reflectance level (0.196) that is consistent with the baseline reflectance level in August 2013, right after the detector settings change (0.195). We conclude that GOME-2A can safely be used for fluorescence retrieval in the early period, but after the swath reduction, a degradation correction is warranted.

The degradation correction is based on an empirical determination of the spectral reflectance change compared to the July 2013 baseline. The following spectral reflectance correction is applied to observed reflectances after July 2013 prior to the retrieval:

$$R_c(\lambda) = R + R_0(\lambda) * t$$

Where R_c is the corrected reflectance, R is the original reflectance and t is the time in decimal years since the GOME-2A swath reduction. Factor $R_0(\lambda)$ is empirically determined for every season (i.e. DJF, MAM, JJA and SON).

---

[1] https://calval.cr.usgs.gov/apps/libya-4

5.2 Latitudinal bias

Similar to other retrievals [*Köhler et al.*, 2015; *Joiner et al.*, 2016], global SIF retrievals show indications for a bias in the retrieved fluorescence levels that systematically depend on latitude. This bias can be estimated by using the fluorescence signal strength observed over the Pacific Ocean (130°-150° W), where SIF is expected to be zero.

Figure 5.2 shows the daily mean bias as a function of latitude for 15 July 2007. The magnitude of the bias is 0.2 mW m^{-2} sr^{-1} nm^{-1} at most, some 10% of the highest fluorescence signals. The bias follows a similar pattern from year to year and we do not see any strong indications for a long-term trend.
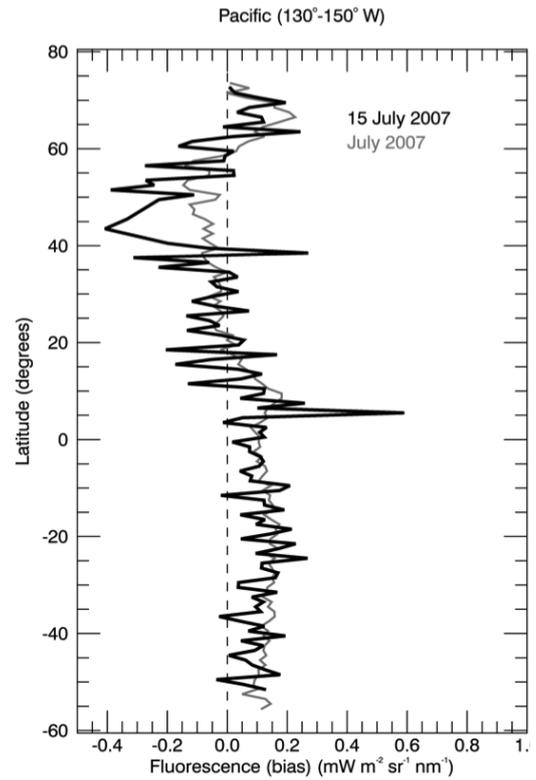
Figure 5.2: Zero-level bias determined over the Pacific Ocean for July 15 2007. Applied per latitude to correct for orbit phase bias. This bias is most likely due to heating of the instrument during its sunlit side passage.

5.3 Uncertainty Estimates

Uncertainty in the retrieved SIF is driven by uncertainties in the measured reflectance (i.e. by measurement noise) and by uncertainties associated with solving the inverse non-linear problem. In the non-linear Levenberg-Marquardt regression, the differences between the measured and modeled reflectance are minimized by the parameters best describing the scene albedo (5 fit parameters), the atmospheric transmission (10 parameters for the 10 principal components), and the intensity of SIF via Equation 4.3. Correlations among the parameters may play an important role in the fitting, and contribute to (additional) uncertainty in the fluorescence. The diagonal element of error covariance matrix provides the uncertainty in the fluorescence fit parameter. We can consider this a realistic estimate of the uncertainty for a single retrieval, as long as the description of scene albedo and transmission is appropriate, and does not suffer from possible misrepresentations. Such misrepresentations could be: too low extinction from atmospheric water vapor described in the set of PCs (as compared to the scene of interest), or too much extinction in the PCs for scenes with higher terrain height than the Sahara.
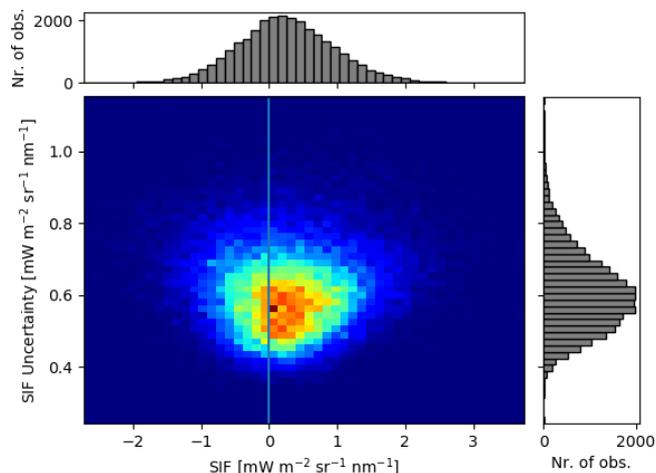
Figure 5.3: Uncertainties in GOME-2(A) SIF retrievals as a function of signal strength on 1 August 2012. Only pixels over land (albedo > 0.08) with cloud fractions < 0.4 have been selected. The colors are indicating the number of occurrences of combinations of SIF and SIF uncertainty values.

Figure 5.3 shows single pixel uncertainties obtained from the Levenberg-Marquardt fit as a function of the fluorescence signal strength over land. The median single-pixel uncertainty amounts to 0.6 mW m-2 sr-1 nm-1, and appears to be independent of the fluorescence signal itself. This implies that the relative uncertainty for strongly fluorescent scenes is 25%-50%, but can easily exceed 100% for scenes with a moderate signal of <0.5 mW m-2 sr-1 nm-1. The requirement for the single pixel uncertainty is 0.1-1.0 mW m^{-2} sr^{-1} nm^{-1}. For monthly mean gridded data, errors are estimated to be a factor 3-5 smaller.

To test for any problems in the retrieval we show additional metrics. Figure 5.4 shows SIF error estimate which is expected to relate to the square root of the signal level. Another fitness test is to check if the residual scales with the expected SNR ratio. To test this Figure 5.5 displays the spectral mean reflectance error and the spectral mean residual both normalized to the reflectance. This reveals that the residuals are indeed of the same order as the instrument SNR and also shows that unsuccessful retrievals (high χ^2_{red}) do not uphold the same relationship.

5.4 QA Value

A qa-value is provided in the daily L2 SIF files to provide users a quick data quality filter. This value is calculated from two parameters with a large influence on SIF retrieval precision: cloud fraction and χ^2_{red}. Although the above-mentioned autocorrelation is a useful measure for investigating residual structures in the fit residual, a more established parameter for determining retrieval success is the χ^2_{red}. Figure 5.6 shows the occurrence of χ^2_{red} of the fit residual. Optimally this number should be 1, values are well within an acceptable
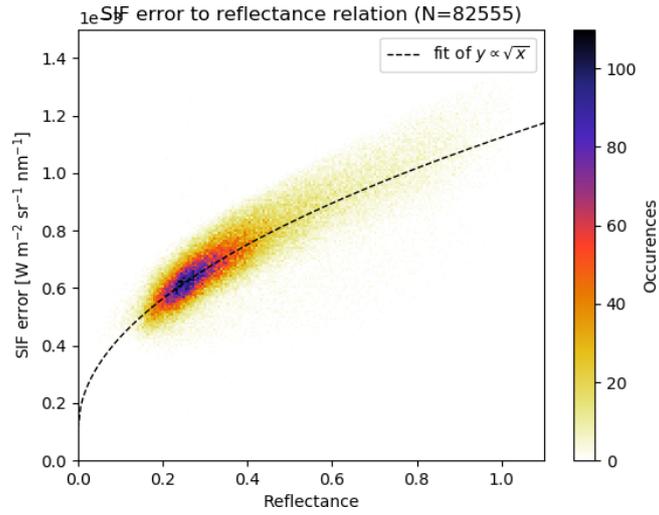
Figure 5.4: SIF error related to reflectance. The expected square root dependency is illustrated by fitting $y = a\sqrt{x}$ to the data (dashed line). Colors indicate the occurrence density.

range. Retrievals that show $\chi^2_{\text{red}} > 5$ can be considered underfitting.

The qa-value is calculated from these two parameters as follows:

$$qa = 1 - (w_\chi * \chi^2_{\text{red}}) * 0.01 - w_c * f_c \tag{5.1}$$

With $w_\chi = 3$ the weight of the χ^2_{red} and $W_c = 1$ the weight of the cloud fraction contribution.

IMPORTANT NOTE: Users are strongly advised to disregard data flagged with a qa_value < 0.6 as data quality can no longer be warranted.
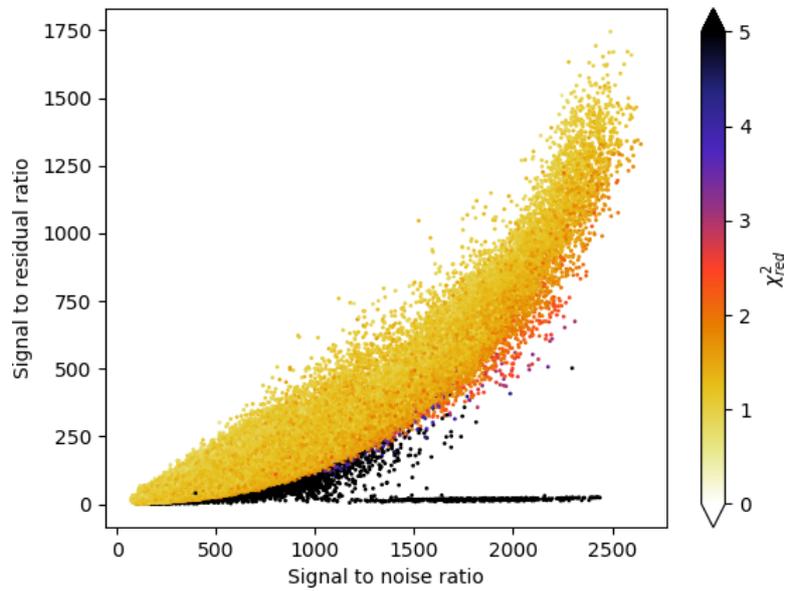
Figure 5.5: Relation between the residual scaled to the reflectance and the SNR. Colors indicate retrieval success measured as χ^2_{red}. Successful retrievals yield residuals which are proportional to the SNR. Fit failures occupy a distinct domain in this space.
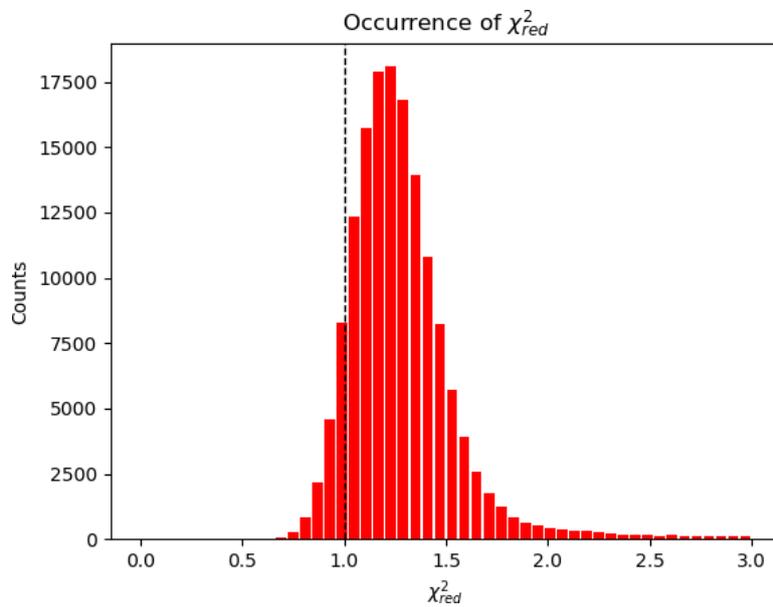
Figure 5.6: Histogram showing the occurence of χ^2_{red} for a single day. The median of ~ 1.2 is very close to 1, indicating a good fit.

Chapter 6

Appendix

6.1 Statistical summaries

Table 6.1: Results of the WATER experiment (mean of 1000 spectra) to reproduce fluorescence for different fitting windows and number of PCs used. The bias is defined here as the mean of the differences between assumed and retrieved fluorescence strength (on average 1.5 mW m^{-2} sr^{-1} nm^{-1}), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE.

	Number of PCs	Bias (mW m-2 sr-1 nm-1)	RMSE (mW m-2 sr-1 nm-1)	Faulty
712-783 nm (full spectral window)	8	-0.44 (-29%)	0.66	68.3%
	20	-0.53 (-35%)	1.59	38.3%
	35	-0.52 (-35%)	0.75	33.7%
712-758 nm (excluding O$_2$-A band)	8	-0.41 (-29%)	0.67	47.9%
	20	-0.28 (-19%)	1.73	47.7%
	35	-0.19 (-13%)	0.52	39.4%
734-783 nm (exclude H$_2$O band)	8	-0.34 (-23%)	0.57	67.1%
	20	-0.42 (-28%)	0.65	48.5%
	35	-0.52 (-35%)	0.79	29.6%
734-758 nm (exclude both bands)	8	0.12 (8%)	0.42	64.5%
	20	0.37 (24%)	0.62	58.7%

		35	0.07 (5%)	0.52	43.4%

Table 6.2: Results of the ALBEDO experiment (mean of 1000 spectra) to reproduce fluorescence for different fitting windows and number of PCs used. The bias is defined here as the mean of the differences between assumed and retrieved fluorescence strength (on average 1.5 mW m^{-2} sr^{-1} nm^{-1}), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE.

	Number of PCs	Bias (mW m-2 sr-1 nm-1)	RMSE (mW m-2 sr-1 nm-1)	Faulty
712-783 nm (full spectral window)	8			97.1%
	20			98.2%
	35			98.4%
712-758 nm (excluding O_2-A band)	8	-0.57 (-38%)	0.77	44.1%
	20	-0.17 (-11%)	0.44	38.4%
	35	-0.24 (-16%)	0.48	36.0%
734-783 nm (exclude H_2O band)	8	-0.26 (-17%)	0.58	16.7%
	20	-0.28 (-19%)	0.63	13.7%
	35	-0.29 (-19%)	0.65	12.7%
734-758 nm (exclude both bands)	8	-0.02 (-1%)	0.44	13.6%
	20	0.01 (1%)	0.48	13.1%
	35	-0.05 (-3%)	0.59	12.1%

Table 6.3: Results of the RED EDEGE experiment (mean of 1000 spectra) to reproduce fluorescence for different fitting windows and number of PCs used. The bias is defined here as the mean of the differences between assumed and retrieved fluorescence strength (on average 1.5 mW m^{-2} sr^{-1} nm^{-1}), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE.

	Number of PCs	Bias (mW m-2 sr-1 nm-1)	RMSE (mW m-2 sr-1 nm-1)	Faulty

712-783 nm (full spectral window)	8			100.0%
	20			99.9%
	35			99.9%
712-758 nm (excluding O_2-A band)	8	-0.55 (-37%)	0.72	50.5%
	20	-0.17 (-11%)	0.38	46.1%
	35	-0.24 (-16%)	0.42	42.1%
734-783 nm (exclude H_2O band)	8	0.22 (15%)	0.35	76.7%
	20	0.14 (9%)	0.30	66.7%
	35	0.04 (3%)	0.32	57.0%
734-758 nm (exclude both bands)	8	-0.01 (-1%)	0.38	19.1%
	20	0.03 (2%)	0.41	20.7%
	35	0.01 (1%)	0.44	23.2%

Table 6.4: Results of the LPATH experiment (mean of 1000 spectra) to reproduce fluorescence for different fitting windows and number of PCs used. The bias is defined here as the mean of the differences between assumed and retrieved fluorescence strength (on average 1.5 mW m^{-2} sr^{-1} nm^{-1}), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE.

	Number of PCs	Bias (mW m-2 sr-1 nm-1)	RMSE (mW m-2 sr-1 nm-1)	Faulty
712-783 nm (full spectral window)	8	-0.40 (-26%)	0.59	47.1%
	20	-0.05 (-3%)	3.68	38.7%
	35	-0.31 (-21%)	1.98	34.8%
712-758 nm (excluding O_2-A band)	8	-0.51 (-34%)	0.76	44.6%
	20	-0.32 (-21%)	1.29	42.1%
	35	-0.31 (-21%)	0.49	38.7%
734-783 nm (exclude H_2O band)	8	-0.19 (-13%)	0.41	35.0%
	20	-0.20 (-13%)	0.43	26.8%

	35	-0.23 (-15%)	0.47	22.6%
734-758 nm (exclude both bands)	8	0.02 (1%)	0.35	23.2%
	20	0.08 (5%)	0.41	19.8%
	35	-0.02 (-1%)	0.45	15.9%

List of Figures

List of Tables

Bibliography

Anav, A., et al., Spatiotemporal patterns of terrestrial gross primary production: A review, *Reviews of Geophysics*, *53*(3), 785–818, doi:10.1002/2015RG000483, 2015.

Berrisford, P., P. Kållberg, S. Kobayashi, D. Dee, S. Uppala, A. Simmons, P. Poli, and H. Sato, Atmospheric conservation propoerties in ERA-Interim, *Quarterly Journal of the Royal Meteorological Society*, *137*, doi:10.1002/qj.864, 2011.

Botía, S., K. Boersma, P. Stammes, E. van Schaik, and M. Kooreman, Importance of Surface Pressure and Albedo for retrieving Sun Induced Chlorophyll Fluorescence from GOME-2, Master's thesis, Wageningen University, Wageningen, The Netherlands, 2017.

Callies, J., E. Corpaccioli, M. Eisinger, A. Hahne, and A. Lefebvre, GOME-2 - MetOp's Second-Generation Sensor for Operational Ozone Monitoring, *Bulletin*, ESA, 2000.

Campbell, J. E., J. A. Berry, U. Seibt, S. J. Smith, S. A. Montzka, T. Launois, S. Belviso, L. Bopp, and M. Laine, Large historical growth in global terrestrial gross primary production, *Nature*, *544*, 84–87, doi:10.1038/nature22030, 2017.

Chance, K., 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 & Radiative Transfer*, *111*, 1289–1295, doi:10.1016/j.jqrst.2010.01.036, 2010.

Ciais, P., et al., *Carbon and Other Biochemical Cycles. 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*, Cambridge Univsersity Press, Cambridge, UK and New York, NY, USA, 2013.

de Haan, J. F., DISAMAR: Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval, Algorithm description and background information, *Tech. rep.*, Royal Netherlands Meteorological Institute, 2011.

Flexas, J., J. Mariano Escolona, S. Evain, J. Gulías, I. Moya, C. B. Osmond, and H. Medrano, Steady-state chlorophyll fluorescence (Fs) measurements as a tool to follow variations of net CO_2 assimilation and stomatal conductance during water-stress in C_3 plants, *Physiologia Plantarum*, *114*, 231–240, 2002.

Frankenberg, C., A. Butz, and G. Toon, Disentangling chlorophyll fluorescence from atmospheric scattering effects in O_2 A-band spectra of reflected sun-light, *Geophysical Research Letters*, *38*, L03,801, 2011a.

Frankenberg, C., C. O'Dell, J. Berry, L. Guanter, J. Joiner, P. Köhler, R. Pollock, and T. E. Taylor, Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2, *Remote Sensing of Environment*, *147*, 1–12, doi:10.1016/j.rse.2014.02.007, 2014.

Frankenberg, C., et al., New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, *Geophysical Research Letters*, *38*, L17,706, 2011b.

Geladi, P., and B. Kowalski, Partial least-squares regression: A tutorial, *Analytica Chimica Acta*, *185*, 1–17, doi:10.1016/0003-2670(86)80028-9, 1986.

Guanter, L., M. Rossini, R. Colombo, M. Meroni, C. Frankenberg, J.-E. Lee, and J. Joiner, Using field spectroscopy to assess the potential of statistical approaches for the retrieval of sun-induced chlorophyll fluorescence from ground and space, *Remote Sensing of Environment*, *133*, 52–61, doi:10.1016/f.rse.2013.01.017, 2013.

Guanter, L., et al., Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence, *Atmospheric Measurement Techniques*, *8*, 1337–1352, doi:10.5194/amt-8-1337-2015, 2015.

Hartmann, D. L., et al., Observations: Atmosphere and Surface, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 159–251, 2013.

Joiner, J., Y. Yoshida, A. P. Vasilkov, E. M. Middleton, P. K. E. Campbell, Y. Yoshida, A. Kuze, and L. A. Corp, Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects: simulations and space-based observations from SCIAMACHY and GOSAT, *Atmospheric Measurement Techniques*, *5*, 809–829, doi:10.5194/amt-5-809-2012, 2012.

Joiner, J., L. Guanter, R. Lindstrot, M. Voigt, A. P. Vasilkov, E. M. Middleton, K. F. Huemmrich, Y. Yoshida, and C. Frankenberg, Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, an application to GOME-2, *Atmos. Meas. Tech.*, *6*, 2803–2823, 2013.

Joiner, J., Y. Yoshida, L. Guanter, and E. M. Middleton, New methods for the retrieval of chlorophyll red fluorescence from hyperspectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY, *Atmos. Meas. Tech.*, *9*, 3939–3967, 2016.

Joiner, J., et al., First observations of global and seasonal terrestrial chlorophyll fluorescence from space, *Biogeosciences*, *8*, 637–651, doi:10.5194/bg-8-637-2011, 2011.

Khosravi, N., M. Vountas, V. V. Rozanov, A. Bracher, A. Wolanin, and J. P. Burrows, Retrieval of Terrestrial Plant Fluorescence Based on the In-Filling of Far-Red Fraunhofer Lines Using SCIAMACHY Observations, *Frontiers in Environmental Science*, *3, 78*, doi:10.3389/fenvs.2015.00078, 2015.

Koelemeijer, R. B. A., P. Stammes, J. W. Hovenier, and J. F. de Haan, A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment, *Journal of Geophysical Research*, *106*, 3475–3490, 2001.

Köhler, P., L. Guanter, and J. Joiner, A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data, *Atmospheric Measurement Techniques*, *8*, 2589–2608, doi:10.5194/amt-8-2589-2015, 2015.

Köhler, P., C. Frankenberg, T. Magney, L. Guanter, J. Joiner, and J. Landgraf, Global Retrievals of Solar-Induced Chlorophyll Fluorescence with TROPOMI: First Results and Intersensor Comparison to OCO-2, *Geophysical Research Letters*, *45*, 10,456–10,463, doi:10.1029/2018GL079031, 2018.

Kooreman, M., A. Sanders, and W. Verstraeten, Spaceborne Observations of Sun-Induced Vegetation Fluorescence: A 2007-2015 global time series from GOME-2A, Master's thesis, Utrecht University, the Netherlands, 2015.

Koren, G., et al., Widespread reduction in sun-induced fluorescence from the Amazon during the 2015/2016 El-Niño, *Philosophical Transactions of the Royal Society B. Biological sciences*, *373*, doi:10.1098/rstb.2017.0408, 2018.

Kraft, S., U. Del Bello, M. Bouvet, M. Drusch, and J. Moreno, FLEX: ESA's Earth Explorer 8 candidate mission, *Proceedings of 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22-27 July.*, pp. 7125–7128, 2012.

Le Quéré, C., et al., Global Carbon Budget 2015, *Earth System Science Data*, *7*, 349–396, doi:10.5194/essd-7-349-2015, 2015.

Sanders, A. F. J., W. W. Verstraeten, M. L. Kooreman, T. C. van Leth, J. Beringer, and J. Joiner, Spaceborne Sun-Induced Fluorescence Time Series from 2007 to 2015 Evaluated with Australian Flux Tower Measurements, *Remote Sensing*, *8*, doi:10.3390/rs8110895, 2016.

Schreiber, U., and W. Bilger, *Rapid assessment of stress effects on plant leaves by chlorophyll fluorescence measurements.*, Plant Response to Stress, Edited by J. D. Tenhunen et al. Springer-Verlag, Heidelberg, Berlin, 1987.

Solomon, S., G.-K. Plattner, R. Knutti, and P. Friedlingstein, Irreversible climate change due to carbon dioxide emissions, *Proceedings of the Natural Acadamy of Sciences*, *106*, 1704–1709, doi:10.1073/pnas.0812721106, 2009.

Sun, Y., C. Frankenberg, M. Jung, J. Joiner, L. Guanter, P. Köhler, and T. Magney, Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP, *Remote Sensing of Environment*, *209*, 808–823, doi:10.1016/j.rse.2018.02.016, 2018.

Tilstra, L. G., O. N. E. Tuinder, P. Wang, and P. Stammes, Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY, *Journal of Geophysical Research: Atmospheres*, *122*, doi:10.1002/2016JD025940, 2017.

van Leth, T. C., A. Sanders, and W. Verstraeten, Remote Sensing of Sun-Induced Chlorophyll Fluorescence, Master's thesis, Utrecht University, the Netherlands, 2014.

van Schaik, E., K. Boersma, P. Stammes, and M. Kooreman, Retrieving Sun-Induced Fluorescence from the Global Ozone Monitoring Experiment 2, Master's thesis, Wageningen University, Wageningen, The Netherlands, 2016.

Verrelst, J., G. Camps-Valls, J. Muñoz Marí, J. Pablo Rivera, F. Veroustraete, J. Clevers, and J. Moreno, Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - a review, *Journal of Photogrammetry and Remote Sensing*, *108*, 273–290, 2015.

Wang, P., P. Stammes, R. van der A, G. Pinardi, and M. van Roozendael, FRESCO+: an improved O_2 a-band cloud retrieval algorithm for tropospheric trace gas retrievals, *Atmospheric Chemistry and Physics Discussions*, *3*, 2008.

Wang, P., O. Tuinder, and P. Stammes, Cloud retrieval algorithm for GOME-2: FRESCO+, *Tech. rep.*, Royal Netherlands Meteorological Institute, 2010.

Zarco-Tejada, P. J., J. R. Miller, G. H. Mohammed, and T. L. Noland, Chlorophyll Fluorescence Effects on Vegetation Apparent Reflectance: I. Leaf-Level Measurements and Model Simulation, *Remote Sensing of Environment*, *74*, 582–595, 2000.