Global retrievals of terrestrial chlorophyll fluorescence from spaceborne spectrometers

Philipp Köhler

27.05.2016

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
am Fachbereich Geowissenschaften
der Freien Universität Berlin

1. Gutachter: Prof. Dr. Jürgen Fischer

2. Gutachter: Prof. Dr. Luis Guanter

Tag der Disputation: 15.07.2016

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Arbeit wurde in keinem früheren Promotionsverfahren eingereicht.

Philipp Köhler

Abstract

This thesis is based on research articles (either published in or submitted to peer-reviewed journals), contributing to the assessment of global vegetation photosynthesis through the monitoring of sun-induced chlorophyll fluorescence (SIF) from space which is the aim of the GlobFluo project funded by the Emmy Noether Programme of the Deutsche Forschungsgemeinschaft (GU 1276/1-1). SIF is a weak electromagnetic signal emitted by the photosynthetic machinery of active vegetation in the visible and near-infrared spectral regions (650–800 nm). In this thesis, two fundamentally different algorithms were developed (physically- and statistically-based) to decouple the SIF signal from the about 100-times-more-intense reflected solar radiation in the measured top-of-atmosphere (TOA) radiance spectrum. The physically-based approach has been applied to data acquired from JAXA's Greenhouse Gases Observing Satellite (GOSAT) between June 2009–August 2011. The statistically-based retrieval has been employed to derive SIF from data of two different satellite instruments: (i) ESA's SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) between August 2002–March 2012, and (ii) EUMETSAT's Global Ozone Monitoring Experiment-2 (GOME-2) between January 2007–April 2015. Comparisons of these three new global SIF data sets with each other as well as with existing data sets reveal highly consistent spatiotemporal patterns. However, differences arise when comparing absolute values which is presently less critical since the absolute value is not applied in current research. Furthermore, the influence of potentially confounding factors on the retrieval has been analyzed. It could be proven that results from both retrieval approaches are relatively robust against cloud contamination. In contrast, the South Atlantic Anomaly (SAA), a region with an anomalously reduced strength in the Earth‘s magnetic field, leads to an increased retrieval noise in large parts of the South American continent. Therefore, a particularly stringent filtering ensured that the impact of the SAA on GOME-2 SIF measurements in a case study in the Amazon rainforests is as low as possible. Regarding the seasonality of SIF in the Amazon, two often disregarded aspects have been identified to be particularly important for the interpretation of SIF: (i) an essential influence of the satellite overpass time due to the response of SIF to instantaneous photosynthetically active radiation (PAR), and (ii) a significant impact of the sun-surface-sensor geometry due to anisotropic SIF emission characteristics. Given the intrinsic influence of canopy structure on SIF measurements, combining SIF with appropriate vegetation parameters (ideally retrievable from the same sensor) may be one key aspect to relate SIF to canopy photosynthesis. In particular, the canopy scattering coefficient (CSC), resulting from the concept of canopy spectral invariants, has been identified as a promising candidate for this purpose when observing densely vegetated areas. Therefore, the first global CSC data set has been derived from atmospherically corrected GOME-2 data and was examined within the case study. It has been found that both SIF and CSC may serve as complementary and valuable vegetation parameters under the conditions in the Amazon rainforest.

Kurzzusammenfassung

Diese Dissertation basiert auf Forschungsartikeln (veröffentlicht oder eingereicht in Peer-Review Journalen), die zur Analyse der terrestrischen Photosynthese durch die satellitengestützte Fernerkundung der von solarer Strahlung induzierten Chlorophyllfluoreszenz (SIF) beitragen. Damit wurde das Hauptziel des GlobFluo Projekts verfolgt, welches durch das Emmy Noether Programm der Deutschen Forschungsgemeinschaft (GU 1276/-1) gefördert wird. SIF ist ein schwaches elektromagnetisches Signal, welches vom Photosyntheseapparat aktiver Pflanzen im sichtbaren und nahen infraroten Spektralbereich (650–800 nm) emittiert wird. Es wurden zwei grundlegend verschiedene Algorithmen entwickelt (physikalisch und statistisch), die es ermöglichen das SIF-Signal von der in etwa 100 mal intensiveren reflektierten solaren Strahlung am Oberrand der Atmosphäre (TOA) zu separieren. Das physikalische Verfahren wurde auf Messungen von JAXAs Greenhouse Gases Observing Satellite (GOSAT) zwischen Juni 2009–August 2011 angewendet. Daten zweier verschiedener Satelliteninstrumente wurden für das statistische Verfahren verwendet: (i) ESAs SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) zwischen August 2002–März 2012 und (ii) EUMETSATs Global Ozone Monitoring Experiment-2 (GOME-2) zwischen Januar 2007–April 2015. Vergleiche dieser drei neuen globalen SIF-Datensätze untereinander sowie mit bereits bestehenden SIF-Datensätzen zeigen eine hohe räumliche und zeitliche Konsistenz. Unterschiede ergeben sich jedoch in Absolutwerten, die allerdings als weniger kritisch einzustufen sind, da Absolutwerte in der gegenwärtigen Forschung keine Anwendung finden. Desweiteren wurde der Einfluss möglicher Störfaktoren auf die SIF-Ableitungsverfahren analysiert. Es konnte gezeigt werden, dass die Ergebnisse beider Verfahren relativ robust gegen Wolkenkontamination sind. Eine bedeutsame Fehlerquelle stellt hingegen die Südatlantische Anomalie (SAA, eine regionale Anomalie im Magnetfeld der Erde) dar, welche für ein verstärktes Messrauschen in weiten Teilen Südamerikas verantwortlich ist. Deshalb wurde ein besonders restriktiver Filter angwendet um sicherzustellen, dass der Einfluss der SAA auf GOME-2 SIF für eine Fallstudie im Amazonischen Regenwald so gering wie möglich ist. Hinsichtlich der Saisonalität von SIF in der Amazonasregion wurden zwei relevante Aspekte identifiziert, die bisher kaum berücksichtigt wurden: (i) ein essentieller Einfluss der Überflugzeit des Satelliten aufgrund der Sensitivität von SIF auf photosynthetisch aktive Strahlung (PAR) und (ii) ein signifikanter Einfluss der Beobachtungsgeometrie aufgrund der Anisotropie der SIF-Emission. In Anbetracht des inhärenten Einflusses der Vegetationsstruktur auf SIF-Messungen ist eine Verknüpfung von SIF mit weiteren Vegetationsparametern (idealerweise vom selben Sensor ableitbar) erstrebenswert, um einen Zusammenhang zwischen SIF und der Photosyntheseaktivität herzustellen. Insbesondere der „Canopy Scattering Coefficient" (CSC), welcher aus der Theorie über spektrale Invariante folgt, wurde in diesem Sinne als vielversprechend befunden. Deshalb wurde der erste CSC Datensatz aus atmosphärenkorrigierten GOME-2 Daten abgeleitet und im Rahmen der Fallstudie analysiert. Dabei wurde festgestellt, dass unter den Bedingungen des Amazonischen Regenwaldes sowohl SIF als auch CSC als komplementäre und nützliche Vegetationsparameter dienen können.

Contents

Chapter 1

Introduction

1.1 Rationale & Structure

This thesis is based on research conducted within the framework of the GlobFluo project funded by the Emmy Noether Programme of the Deutsche Forschungsgemeinschaft (GU 1276/1-1) and is a compilation of scientific articles. The overall objective of the GlobFluo project is to assess global vegetation photosynthesis through the monitoring of sun-induced chlorophyll fluorescence (SIF) from space. Photosynthesis is the most essential process on Earth, considering that oxygen and the very foundation of the food chain (plant material) are formed through the conversion of sun-light to chemical energy. Gradually changing environmental conditions, e.g., increasing concentrations of atmospheric trace gases, raising temperatures or reduced water availability will most likely affect photosynthetic productivity. Therefore, consistent long-term time series of vegetation parameters, ideally describing the global photosynthetic performance, are required to detect trends and potential threats to biodiversity. A promising and innovative approach for the monitoring of terrestrial photosynthesis on the global scale is the estimation of SIF from spaceborne spectrometers. SIF is a unique spectral signal emitted by the photosynthetic machinery of green vegetation and is closely related to photosynthesis (e.g., Zarco-Tejada et al., 2003).

The advent of first global SIF retrieval approaches from Joiner et al. (2011) and Frankenberg et al. (2011a,b) marked the dawn of a new research field, aiming at the assessment of dynamics in global photosynthetic productivity by relating SIF to photosynthesis. Several recently published studies (e.g., Lee et al., 2013; Guanter et al., 2014; Joiner et al., 2014; Parazoo et al., 2014; Zhang et al., 2014; Yang et al., 2015; Walther et al., 2015) reveal that SIF data have the potential for providing complementary and in many cases enhanced information with respect to traditional reflectance-based vegetation indices (VIs), which evaluate the "greenness" (mixture of canopy structure, biomass, and pigment concentration) of vegetation. In this context, two aspects appear to be particularly relevant for current and future research. First, all studies relating SIF to photosynthesis rely on the quality of the underlying data sets which justifies a further

development and advancement of retrieval techniques. Second, the interpretation of SIF data is challenging due to several factors that influence the SIF signal such as vegetation status, environmental conditions, canopy structure, viewing and illumination geometry, and atmospheric conditions. This thesis seeks to address both points by (i) revisiting technical aspects related to the retrieval of SIF from spaceborne spectrometers, and (ii) examining SIF data with respect to relevant influencing factors.

In the following introductory Sect. 1.2, I will give a brief overview on the chlorophyll fluorescence signal and its link to photosynthesis, including current limitations for the full exploitation of the signal. Furthermore, I will address the estimation of SIF from spectroscopic measurements with emphasis on spaceborne instruments, and outline the frame conditions for a case study in the Amazon region. A detailed description of research objectives is given in Sect. 1.3. The core of this thesis is comprised of three chapters (2-4) reproducing research articles, either published in or submitted to peer-reviewed journals. Each chapter is therefore divided into distinct sections such as abstract, introduction, methods, results, summary/conclusions, and acknowledgments. In consequence, these chapters can be read individually, implying that recurring material occurs to a limited extent. Furthermore, it should be noted that the nomenclature was kept as originally published/submitted, leading to slight inconsistencies (e.g., SIF and F_s are used interchangeably) which is due to differing guidelines among journals and suggestions by reviewers. Contributions of author and co-authors are indicated at the end of each chapter. Chapter 2 and 3 primarily deal with technical aspects related to different retrieval techniques (physically-based, statistically-based), while chapter 4 assesses the potential of SIF and another promising vegetation parameter, namely the canopy scattering coefficient (CSC), for the monitoring of dense vegetation in the Amazon rainforest. Chapter 5 summarizes main findings, provides an overview on follow-up studies (which took advantage of the SIF data produced within the framework of this thesis), and discusses future prospects.

1.2 Research Background

This section provides an introductory overview on the chlorophyll fluorescence signal and its link to photosynthesis on different spatial scales (Sect. 1.2.1). Subsequently, Sect. 1.2.2 addresses the measurement of SIF and Sect. 1.2.3 outlines prerequisites for assessing the photosynthetic activity in the Amazon region, which has been selected as a study case because of the relevance of this area to the global carbon cycle.

1.2.1 Chlorophyll fluorescence and its link to photosynthesis

Fluorescence generally refers to the spontaneous emission of light by a substance that has absorbed light of another wavelength, while the emission occurs usually at longer wavelengths associated with lower energy (known as Stokes shift). The released photon (fluorescence) has typically a lower energy (longer wavelength) due to internal energy conversion. This process takes place on the scale of atoms/molecules, while a measure-

ment of plant chlorophyll fluorescence from space has a pixel size in the order of a few square kilometers. In the following, I will therefore describe the chlorophyll fluorescence signal and its link to photosynthesis according to the bottom-up principle.

Molecular & leaf level

During the initial stage of photosynthesis, light is absorbed to cover the energy demand for synthesizing carbon dioxide (CO_2) and water to glucose and oxygen. In case of higher plants, the light-harvesting takes place in the chloroplasts, a sub-unit of the leaf's cells. More specifically, two distinct types of photosystems, termed PSI and PSII, are embedded in the membrane of thylakoids within the chloroplasts (e.g., Lichtenthaler and Rinderle, 1988). Both photosystems consist of (i) an antennae with light-harvesting complexes (LHCs) and (ii) a reaction center (Heldt, 2004). The LHCs involve several pigment molecules (chlorophyll a, chlorophyll b, carotenoids, and anthocyanins) bounded by proteins (Liu et al., 2004), where photons of certain wavelengths are absorbed by the pigments. The resulting excitation energy is rapidly passed to the reaction center (resonance energy transfer), where a pigment-protein complex (containing chlorophyll a and β-carotene) performs the photosynthetic light reaction (Lichtenthaler and Rinderle, 1988). Due to differing pigment compositions of PSI and PSII, light of slightly different wavelengths is absorbed by each photosystem, while chlorophyll represents the main photosynthetic pigment in both PSI and PSII (Heldt, 2004). Both photosystems operate in series, in which electrons are being transferred from PSII to PSI (e.g., Porcar-Castell et al., 2014). Flexible mechanisms balance disturbances in the electron transport chain and adjust the energy partitioning between the photosystems. The light absorption involves the excitation of chlorophyll a from the ground state to a higher energetic level (Lichtenthaler and Rinderle, 1988), which essentially represents the origin of the fluorescence emission. So far, it has been described how the absorbed light reaches the chlorophyll a molecules. In addition, a wavelength specific review of this process is necessary for an understanding of the fluorescence emission. The paragraph below refers to Fig. 1.1, illustrating the emergence of chlorophyll a fluorescence.

Leaves absorb light mainly restricted to a small portion of the sunlight's spectrum between 400-700 nm, termed photosynthetically active radiation (PAR), where chlorophyll pigments absorb electromagnetic radiation preferentially in blue (∼450 nm) and red (∼650 nm) wavelengths. In the visible spectrum (400–700 nm), enhanced reflection takes place in the green (∼550 nm) which leads to the characteristic appearance of plants as seen with the human eye. The Planck-Einstein relation (Eq. 1.1) states that a photon's energy occurs only in discrete levels (quanta), separated by a fundamental constant:

$$E = h\nu = \frac{hc}{\lambda}, \tag{1.1}$$

where h is the Planck's constant (6.626 x 10^{-34} Js) and ν is the frequency of the electromagnetic wave. The frequency is in turn related to the wavelength λ through $\nu = c\lambda^{-1}$, where c is the speed of light (3 x 10^8 ms^{-1}). Light energy therefore can only be ab-

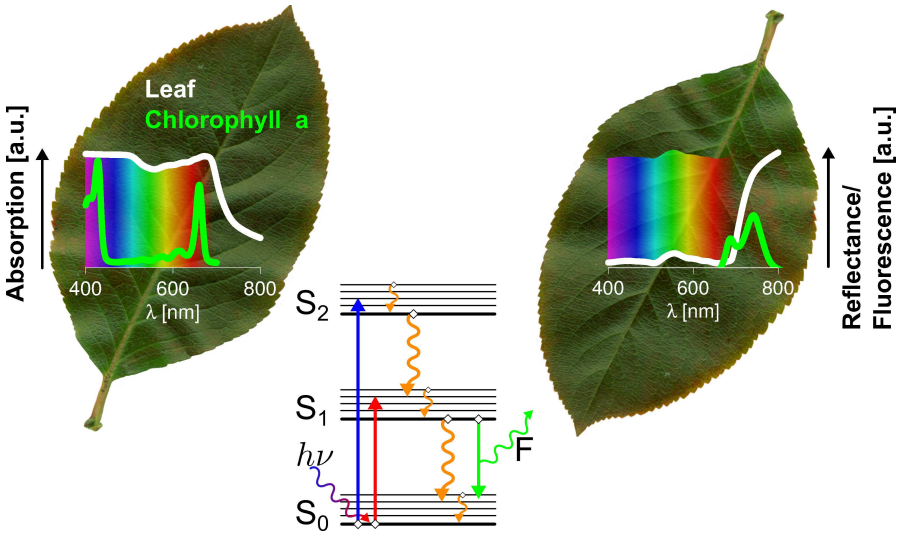


Figure 1.1: Origin of the chlorophyll *a* fluorescence emission. The left plot depicts the absorption spectrum of a healthy green leaf and of chlorophyll *a* (dissolved in diethyl ether, Dixon et al., 2005). The idealized Jablonski diagram (similar to Porcar-Castell et al., 2014) in the middle depicts the ground (S_0) and excited states (S_1 and S_2) of chlorophyll *a* molecules, including the excitation in case of blue and red photons ($h\nu$) with the corresponding thermal (orange) and fluorescent (green) pathways of de-excitation. Note, however, that most of the absorbed energy will typically be used for photosynthesis. The plot on the right depicts the spectral leaf reflectance and the chlorophyll *a* fluorescence emission spectrum, where the ratio of fluorescence peak values (F685, F740) depends on leaf structure and chlorophyll content. Absorption, reflectance, and emission spectra have been scaled for visualization purposes.

sorbed in discrete quanta, resulting also in discrete excitation states. Consequently, two main excitation states of chlorophyll *a* molecules occur due to the two distinct absorption maxima. Once a photon has been absorbed by a chlorophyll *a* molecule (or energy has been passed from other pigments in the antennae), one of its electrons is being excited to a singlet state (S_1 or S_2) with rotational and vibrational sublevels as indicated by the fine lines in Fig. 1.1 (Heldt, 2004). In case of blue light (containing more energy than red light), the second singlet state (S_2) is excited and decays rapidly (within $\sim 10^{-12}$ s) under heat emission (by rotations and vibrations) to the first singlet state (Heldt, 2004). The de-excitation from of the first singlet state can in turn proceed by thermal energy dissipation or through fluorescence, while this S_1 state can also be excited by red light directly (Heldt, 2004). However, the most important pathway of de-excitation from the S_1 state for photosynthesis consists of using the energy for chemical work (Heldt, 2004).

In addition, rare transitions from exited singlet states to potentially harmful triplet states can occur (Heldt, 2004). The probability for other de-excitation pathways than thermal energy dissipation increases at S_1 due to its much longer half-life of about 10^{-9} s. The resulting chlorophyll *a* fluorescence emission occurs roughly between 650–800 nm, building a double peak feature with maximum values at ∼685 nm (F685) and ∼740 nm (F740), respectively (Lichtenthaler and Rinderle, 1988; Buschmann, 2007). Note, however, that the major part of the absorbed light energy (under favorable conditions $> 80\%$) is used to generate reductants (ATP and NADPH) for the CO_2 assimilation in the dark reaction of photosynthesis (Calvin-Benson cycle), while oxygen is released as a by-product of the initial light reaction (Buschmann, 2007).

In summary, absorbed light can mainly undergo three competing pathways: (i) drive photosynthesis, (ii) dissipate as heat (also known as non-photochemical quenching) or (iii) be re-emitted as chlorophyll fluorescence. The probability of each pathway is subject to biochemical and environmental conditions (van der Tol et al., 2009a), and involves several regulatory mechanisms to dissipate excess energy (Porcar-Castell et al., 2014). That means the relationship between photosynthesis and fluorescence is in general non-linear, requiring large efforts in order to correctly interpret the fluorescence signal.

As a consequence of the interlinked pathways and regulatory mechanisms, fluorescence can be highly dynamic. This was first demonstrated by the observation from Kautsky and Hirsch (1931), revealing a rapid increase of fluorescence when dark-adapted leaves are exposed to light. The initially elevated intensity of fluorescence continuously decreases to a steady-state level within a few minutes ("Kautsky effect"). Simultaneously, the rate of CO_2 assimilation increases, suggesting an inverse relationship between fluorescence and photosynthesis during the period of light-acclimation (also known as induction kinetics). Their observation initiated a vast body of research based on such or similar kinetically processes, e.g., measurements of the temporal course of fluorescence induced by saturating light pulses under various conditions (see review articles by Lichtenthaler and Rinderle, 1988; Krause and Weis, 1991; Maxwell and Johnson, 2000; Govindjee, 2004; Schreiber, 2004; Baker, 2008). These active measurement techniques with external light sources, however, limit measurements to small spots in combination with a short distance to the sensor. Contrastingly, passive remote sensing enables the continuous monitoring of large areas, even from space, but it is generally not trivial to transfer findings obtained with active measurements to those of passive measurements. Since this thesis deals with passive remote sensing of SIF, I will hereinafter focus on the light-adapted steady-state fluorescence (or SIF).

Though SIF is emanating from both PSII and PSI, it is widely established to assume that the PSII contribution dominates, presumably due to a more efficient photosynthetic pathway in PSI (Porcar-Castell et al., 2014). In contrast to the more dynamic PSII fluorescence, the relatively low PSI contribution appears to be rather constant, even under stress conditions. Therefore, PSI fluorescence is often neglected or treated as an offset value (Porcar-Castell et al., 2014), while PSII fluorescence is more sensitive to the light use efficiency of photosynthesis and thus of primary interest. However, Porcar-Castell et al. (2014) emphasize that seasonal variations of the PSI fluorescence have not

yet been explored. Moreover, Franck et al. (2002) suggest that the relative contribution of PSI is indeed relevant, depending on wavelength and state (>40% at wavelength > 770 nm, dark-adapted state). There is, by contrast, a general agreement that the first peak, F685, can be attributed mostly to the PSII fluorescence emission, while the second peak, F740, is emanating from both PSI and PSII (Lichtenthaler and Rinderle, 1988; Gitelson et al., 1998; Buschmann, 2007; Porcar-Castell et al., 2014). Overall, the amount of fluorescence emitted by PSI and PSII is rather low and typically represents less than 5% of the total light absorbed (Zarco-Tejada et al., 2003).

The leaf spectral reflectance increases rapidly in the 680–730 nm wavelength region (see Fig. 1.1), termed red edge, mostly due to the spectral variation in chlorophyll absorption. In vivo (outside of a 'living' leaf) measurements of SIF are therefore inherently subject to re-absorption effects within the leaf, basically depending on leaf structure, thickness, and chlorophyll content (Gitelson et al., 1998). Since the red edge causes a selective absorption in red wavelengths, F685 is much more affected than F740. Consequently, the spectral shape of the SIF emission is considerably modified with respect to its origin inside the leaf. Nevertheless, the SIF spectrum in the leaf interior is important to fully exploit the information regarding photosynthesis. Thus, scaling approaches from the leaf to the molecular level are necessary to accurately model the fluorescence emission and its link to photosynthesis under various conditions. This includes the separation of the relative contributions of PSI and PSII fluorescence within a leaf, requiring a combination of measurements and modeling (e.g., Gitelson et al., 1998; Pedrós et al., 2008). Furthermore, the FlourMODleaf model (Pedrós et al., 2010) is capable of simulating the re-absorption of SIF within leaves. Hence, inversion schemes can be applied to relate leaf scale SIF measurements to photosynthetic functioning on the molecular scale. However, the link between SIF and photosynthesis (on the leaf level) can also be explored by using deterministic approaches, which incorporate the available knowledge of involved processes, but do not explicitly account for the radiative transfer within the leaf. More specifically, conventional photosynthesis models have been extended to simulate the response of SIF to changing biochemical and environmental conditions (van der Tol et al., 2009a,b, 2014).

Canopy & ecosystem level

When departing from a single leaf towards the canopy scale, it is highly important to additionally consider structural effects (e.g., Fournier et al., 2012); for instance, a complex canopy will:

1. lead to a further modification of the spectral shape of SIF due to the selective absorption of red wavelengths,

2. decrease the detectable amount of SIF with respect to the integrated SIF emission over all canopy layers due to an enhanced probability for photons to be absorbed or scattered in different directions,

3. have a directional component in the SIF emission (which also applies for less complex canopies) due to leaf angle distributions and the amount of sunlit/shaded leaves,

4. have an enriched contribution of SIF from the uppermost layer.

Considering these effects, it is crucial that recently developed/extended models allow to combine the SIF emission with canopy structure effects to simulate the impact on SIF at the top-of-canopy (TOC) level (e.g., Kobayashi and Iwabuchi, 2008; Zarco-Tejada et al., 2013). This eventually provides the opportunity to relate TOC SIF measurements to photosynthesis.

Besides structural effects of closed canopies, it should be considered that the canopy layer can be discontinuous (e.g., crowns and gaps). The first attempt to explore the influence of a heterogeneous canopy layer on SIF measurements (with a coarser resolution than the closed canopy components) has been realized by the FluorFLIM model. Zarco-Tejada et al. (2013) coupled the FluorMODleaf (Pedrós et al., 2010) and FluorSAIL (Scattering by Arbitrarily Inclined Leaves, Miller et al., 2005) with the geometric forest light interaction model (FLIM, Rosema et al., 1992), accounting for shadowing and crown transparency. FluorFLIM simulations proved that spatial resolution effects significantly influence the SIF signal from heterogeneous canopies, i.e., a weak relationship between SIF from pure, sunlit tree crowns and from aggregated pixel (including soil and shadows) was observed. Recently, the leaf scale photosynthesis-fluorescence model approach from van der Tol et al. (2014) has been combined with the Forest Light Environmental Simulator (FLiES Kobayashi and Iwabuchi, 2008), capable of simulating multiple scattering within the canopy. This integrated atmosphere-three dimensional canopy radiative transfer model can be used to evaluate the link between TOC SIF and canopy photosynthesis under various atmospheric conditions, also for heterogeneous landscapes. In summary, the complex response of SIF to canopy structure demonstrates that the domain of canopy–atmosphere radiative transfer is essential to interpret TOC SIF measurements, while progress in this field has been achieved recently.

The continuous development of methods for the (passive) remote sensing of SIF (see Meroni et al., 2009, and references therein for airborne measurements) enabled the acquisition of SIF data for larger scales. In particular, satellite-based SIF retrievals (Sect.1.2.2) opened a new opportunity to estimate photosynthetic rates for whole ecosystems. However, considering the lack of up-scaling approaches from the canopy level towards larger scales, the question arises whether a unique link between SIF and photosynthesis can be established. Faced with this question, recently conducted research has focused on comparisons between SIF data (predominantly retrieved from spaceborne spectrometers) and other data streams related to photosynthesis (Frankenberg et al., 2011b; Guanter et al., 2012; Lee et al., 2013; Parazoo et al., 2013; Guanter et al., 2014; Joiner et al., 2014; Zhang et al., 2014; Guan et al., 2015; Perez-Priego et al., 2015; Walther et al., 2015).

A common parameter, particularly suited for a comparison with SIF at the ecosystem level, is the estimated carbon uptake through photosynthesis, termed gross primary production (GPP) (e.g., Beer et al., 2010). Traditionally, GPP has been modeled based on

the concept of photosynthetic light use efficiency (Monteith, 1972)

$$\mathrm{GPP} = \mathrm{PAR} \cdot \mathrm{fPAR} \cdot \epsilon, \tag{1.2}$$

expressing GPP as product of photosynthetically active radiation (PAR, 400-700 nm), fraction of PAR (fPAR) that is absorbed, and the light use efficiency ϵ, at which the absorbed PAR (APAR) is used for photosynthesis. Remarkably, SIF similarly contains information on the light absorbed by vegetation and the efficiency at which the light is used for fluorescence emission, which in turn appears to be related to the photosynthetic light use efficiency under certain conditions (Guanter et al., 2014). In order to model GPP, remotely sensed data is typically used to infer APAR, while look-up tables provide values for the light use efficiency (see review article from Hilker et al., 2008). In the more elaborated approach from Jung et al. (2011), GPP is estimated in two steps: (i) General relationships between several explanatory environmental variables and GPP are inferred from measurements at specific sites, and (ii) these relationships are then globally up-scaled by using remote sensing, meteorological and land cover data. The first step comprises a training of model tree ensembles (a machine learning technique) by means of data sets obtained from a global network of eddy covariance flux towers (FLUXNET, Baldocchi et al., 2001). In the second step, the resulting statistical models are used to infer global GPP data.

Strong linear relationships between GPP and SIF were observed by Frankenberg et al. (2011b) and Guanter et al. (2012), though, depending on the biome (similar ecosystems on the global level) under consideration. This biome dependency can partly be attributed to canopy structure effects in different ecosystems such as forests and croplands. Their encouraging findings initiated more detailed studies at the ecosystem level, reflecting the beneficial value of SIF with respect to traditional reflectance-based products for several ecosystems:

1. Amazon rainforests (Lee et al., 2013; Parazoo et al., 2013) and global tropical forests (Guan et al., 2015),
2. Crop- and grasslands (Guanter et al., 2014; Zhang et al., 2014),
3. Boreal forests (Walther et al., 2015),
4. Temperate deciduous forest (Yang et al., 2015),
5. Mediterranean savanna (Perez-Priego et al., 2015).

In view of these findings, it seems reasonable to assume that SIF provides a new and potentially improved source of data to constrain GPP models or to infer GPP directly from SIF measurements. Thus, uncertainties in global GPP predictions could be reduced (Parazoo et al., 2014).

So far, biome dependent empirical relationships are necessary to relate SIF to photosynthesis, but a bottom-up approach is required to find an appropriate parameterization

for different canopy structures at the ecosystem level. Furthermore, polar orbiting satellites (there are currently no geostationary satellites capable of retrieving SIF) acquire measurements from the same area only every few days at a certain daytime. Therefore, information on important short term variations in photosynthetic activity is potentially missing. Last but not least, it must be emphasized that all of the aforementioned results rely on the quality of the underlying SIF data sets, which leads to the main subject of chapters 2 and 3, introduced in the following section.

1.2.2 Measurements of chlorophyll fluorecence

As outlined in Sect. 1.2.1, it is well known that the SIF signal is spectrally smooth and comprises two broad peaks (F685 and F740), while the amount of SIF represents only a few percent of the total light absorbed. The weak SIF emission is added to the reflected solar radiation, which poses the challenge of decoupling the SIF signal. In the following, I will first outline the basic concept enabling the measurement of SIF at the TOA level. Afterwards, the different contributions of the spectral radiance received by a spaceborne spectrometer will be discussed. Third, an overview on satellite-based observations of SIF will be given.

In-filling of absorption lines by chlorophyll fluorescence

A unique opportunity to decouple the SIF emission from reflected radiation arises from the existence of absorption lines in the spectral irradiance. These dark lines in the irradiance spectrum correspond to elements absorbing photons of certain wavelength (equivalent to the selective absorbtion by chlorophyll molecules in Fig 1.1), which can either originate from absorption in the sun's chromosphere, known as solar Fraunhofer

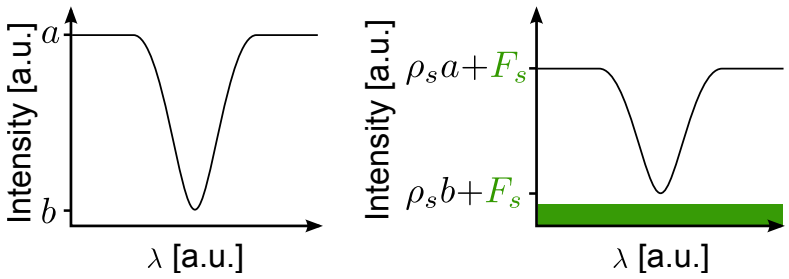


Figure 1.2: Principle of the in-filling of absorption lines by SIF. The left plot depicts an absorption line profile as it could be observed in a narrow wavelength range of the solar irradiance, where a is the intensity in the continuum and b is the intensity in the central wavelength of the absorption feature. The plot on the right illustrates the absorption line profile after being reflected by a fluorescent surface. For the sake of simplcity, surface reflectance (ρ_s) and fluorescence emission (green, F_s) are assumed to be wavelength independent (constant).

lines, or from atmospheric absorption. By performing measurements in spectral regions where reasonably strong solar Fraunhofer lines or atmospheric absorption features occur, the amount of SIF can be computed by exploiting the change in their fractional depth with respect to a non-fluorescent target. Specifically, a decrease in the fractional depth of absorption lines will occur in the presence of any SIF emission (assuming that there is no other source of radiation in the considered wavelength range), which is illustrated in Fig. 1.2. In the trivial case of no SIF emission ($F_s = 0$), the ratio between intensity in the continuum (a) and intensity in the central wavelength of the absorption feature (b) will be identical before and after being reflected ($a/b = (\rho_s a)/(\rho_s b)$). However, as soon as a fluorescence emission occurs, the central wavelength of the absorption feature will, in relative terms, be more affected than the continuum, i.e., the fractional depth decreases after being reflected ($a/b > (\rho_s a + F_s)/(\rho_s b + F_s)$, for $a > b$ and $F_s > 0$). The required comparative measurements (with/without fluorescence emission) can, for example, be realized by a reference panel in the field of view.

A method to exploit the in-filling of solar Fraunhofer lines by luminescent (fluorescent by unknown reasons) substances was firstly proposed by Plascyk (1975), termed Fraunhofer line discriminator (FLD). Based on this principle, the estimation of SIF from ground-based and airborne spectrometers has been addressed in several studies (see review article by Meroni et al., 2009), dating back to the 1980s. However, in contrast to near-surface retrievals, the estimation of SIF from spaceborne instruments involves additionally several atmospheric contributions which are addressed in the following.

Top-of-atmosphere radiance contributions

The complexity a satellite-based retrieval of SIF has to deal with is illustrated in Fig. 1.3, depicting the main contributions of the spectral radiance received by an instrument at the top-of-atmosphere (TOA). Starting with the TOA solar irradiance ($\mathbf{I}_{\text{sol}}$), the prominent hydrogen alpha line at 656.3 nm can be identified, which is the deepest Fraunhofer line in the solar irradiance spectrum (due to the fact that hydrogen is the most abundant element of the sun). The solar irradiance is partly attenuated when the light travels through the atmosphere, while the strength of the attenuation largely depends on the atmospheric state and wavelength. For the sake of simplicity, Fig. 1.3 only shows the total atmospheric transmittance in down- and upward direction ($\mathbf{T}_{\downarrow\uparrow} = \mathbf{T}_{\downarrow} \cdot \mathbf{T}_{\uparrow}$), where the atmospheric absorption in the considered wavelength range (600-800 nm) can mostly be attributed to oxygen (O_2 B-band:$\sim$687 nm, O_2 A-band:$\sim$760 nm) and water vapor. Reaching the surface, the spectral surface reflectance ($\boldsymbol{\rho}_{\text{s}}$), which includes the clearly observable red edge, determines the portion of reflected light in the direction of the sensor. At this stage, SIF ($\mathbf{F}_{\text{s}}$) is added to the signal, but consequently influenced by the atmospheric transmission in the upward direction ($\mathbf{T}_{\uparrow}$). Moreover, a small portion of the solar irradiance does not reach the surface, but is instead reflected by the atmosphere itself ($\mathbf{F}_0$) and then reaches the sensor. This contribution is known as atmospheric path radiance and, as can be seen, its magnitude can be larger than the fluorescence emission, even in the absence of clouds (due to aerosols). Nevertheless, it might be noted that this contribution does not affect the fractional depth of solar Fraunhofer lines. In total, the

spectral TOA radiance ($\mathbf{F}_{\mathrm{TOA}}$) can be formulated as

$$\mathbf{F}_{\mathrm{TOA}} = \mathbf{F}_0 + \frac{\mathbf{I}_{\mathrm{sol}} \cdot \frac{\mu_0}{\pi} \cdot \boldsymbol{\rho}_{\mathrm{s}} \cdot \mathbf{T}_{\downarrow\uparrow} + \mathbf{F}_{\mathrm{s}} \cdot \mathbf{T}_{\uparrow}}{1 - \mathbf{S}_{\mathrm{atm}} \cdot \boldsymbol{\rho}_{\mathrm{s}}}, \tag{1.3}$$

where μ_0 is the cosine of the solar zenith angle and $\mathbf{S}_{\mathrm{atm}}$ is the atmospheric spherical albedo (i.e., the reflectance of the atmosphere for isotropic light entering it from the surface). The product of atmospheric spherical albedo and surface reflectance, accounting for multiple scattering between atmosphere and surface, is usually very small ($\mathbf{S}_{\mathrm{atm}} \cdot \boldsymbol{\rho}_{\mathrm{s}} << 1$) and is therefore neglected in Fig. 1.3.

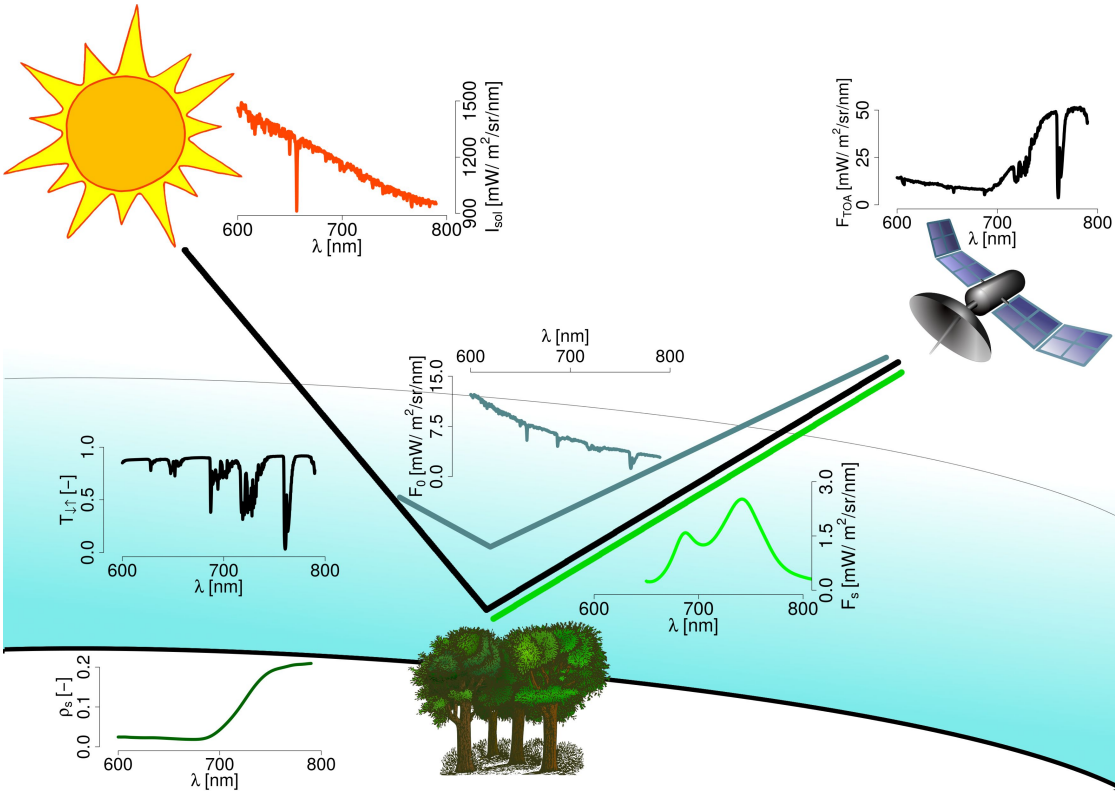


Figure 1.3: Main contributions of a spectroscopic measurement with a spectral resolution of 0.5 nm when photosynthetically active vegetation is observed (under cloud free conditions). Depicted are solar irradiance ($\mathbf{I}_{\mathrm{sol}}$), atmospheric path radiance ($\mathbf{F}_0$), total atmospheric transmission ($\mathbf{T}_{\downarrow\uparrow}$), surface reflectance ($\boldsymbol{\rho}_{\mathrm{s}}$), SIF ($\mathbf{F}_{\mathrm{s}}$), and radiance received by a satellite sensor ($\mathbf{F}_{\mathrm{TOA}}$). Please, see main text for a detailed description.

Satellite-based observations of terrestrial chlorophyll fluorecence

Near-surface SIF retrievals almost exclusively evaluate the in-filling of O_2 absorption bands by SIF (Meroni et al., 2009). These atmospheric absorption features are relatively strong, while absorption and scattering processes between target and sensor can be neglected (Frankenberg et al., 2011a). In contrast to near-surface measurements, spaceborne SIF retrievals using O_2 absorption bands are particularly challenging due to the extinction of the SIF emission through O_2, which is becoming important as the distance to the sensor increases. In addition, atmospheric scattering (e.g., by clouds and aerosols) in the O_2 A-band can affect the TOA radiance similar to the SIF emission itself (Frankenberg et al., 2011a). Nevertheless, the first SIF measurement from space was achieved by deriving the in-filling of the O_2 A-band by SIF (Guanter et al., 2007). Specifically, Guanter et al. (2007) used data acquired from MERIS (MEdium-spectral Resolution Imaging Spectrometer) on board ENVISAT (ENVIronmental SATellite) to retrieve SIF for a vegetated area surrounded by bare soil. In this case study, the conditions (flat area, clear-sky, large green fields surrounded by bare soils) allowed to correct for an offset in the retrieval due to uncertainties in the atmospheric state and the instrument performance. Nonetheless, a comparison of retrieval results from Guanter et al. (2007) to airborne and ground-based SIF estimates revealed reasonably good correlations, thus demonstrating the feasibility to retrieve SIF from space for the very first time. Only a few years later, studies from Joiner et al. (2011), Frankenberg et al. (2011b), and Guanter et al. (2012) presented the first global maps of SIF, using data acquired from the Japanese Greenhouse Gases Observing Satellite (GOSAT). In contrast to the first spaceborne (Guanter et al., 2007) and the most near-surface retrieval approaches, the new retrieval methods (Joiner et al., 2011; Frankenberg et al., 2011a; Guanter et al., 2012) focused on evaluating the in-filling of solar Fraunhofer lines by SIF in narrow spectral regions (fitting windows) mostly devoid of atmospheric absorption. This considerably simplifies the retrieval problem since the SIF emission can pass the atmosphere without being significantly attenuated under cloud free conditions (Frankenberg et al., 2011a). Consequently, explicit radiative transfer calculations are not required. Since then, the availability of global SIF data sets has progressively increased. Meanwhile, SIF retrieval approaches, relying on the in-filling of solar Fraunhofer lines, have been applied to data from several sensors, recently or currently operational in space:

1. Greenhouse Gases Observing Satellite-Fourier Transform Spectrometer (GOSAT-FTS) (Frankenberg et al., 2011a; Joiner et al., 2011, 2012; Guanter et al., 2012; Köhler et al., 2015a)

2. SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) on board ENVISAT (Joiner et al., 2012; Köhler et al., 2015b; Wolanin et al., 2015)

3. Global Ozone Monitoring Experiment-2 (GOME-2) on board the Meteorological Operational Satellites (MetOp-A, MetOp-B) (Joiner et al., 2013; Köhler et al., 2015b; Wolanin et al., 2015)

4. Orbiting Carbon Observatory-2 (OCO-2; Frankenberg et al., 2014)

In general, two fundamentally different retrieval approaches are currently pursued: (i) physically-based approaches (Guanter et al., 2007; Frankenberg et al., 2011a; Joiner et al., 2011, 2012; Frankenberg et al., 2014; Köhler et al., 2015a; Wolanin et al., 2015) and (ii) statistically-based/data-driven approaches (Guanter et al., 2012, 2013, 2015; Joiner et al., 2012; Köhler et al., 2015b). This classification refers to the underlying method to model the TOA radiance. In the first case, the modeling is based on a physical formulation of the radiative transfer between target and sensor. Such approaches are typically restricted to micro-fitting windows of a few nanometers since the description of surface reflectance, SIF, and atmospheric absorption becomes more complicated when extending the fitting window (Guanter et al., 2013). On the other hand, statistically-based retrieval methods rely on the assumption that any TOA radiance spectrum can be modeled by a linear combination of spectral functions, consisting of low and high frequency components. These spectral functions are usually derived from training data sets composed of vegetation-free spectra. An important advantage of statistically-based methods compared to physically-based approaches arises from avoiding the very challenging explicit formulation of atmospheric radiative transfer effects. Therefore, statistically-based methods can be relatively easily extended to broadband fitting windows (>10 nm), including atmospheric absorption features.

Table 1.1 summarizes relevant instrument specifications for satellite sensors which are currently used to retrieve SIF. In particular, the spectral resolution is essential to resolve less pronounced Fraunhofer lines and to accurately estimate the SIF emission in

Table 1.1: Relevant characteristics of satellite instruments currently used to retrieve SIF. Specifically, specifications for SCIAMACHY (Bovensmann et al., 1999), GOME-2 (Munro et al., 2006), GOSAT (Kuze et al., 2009), and OCO-2 (Frankenberg et al., 2014) have been compiled.

	SCIAMACHY	**GOME-2/MetOp-A**	**GOSAT-FTS**	**OCO-2**
Data availability	Aug 2002–Mar 2012	Aug 2007–present	Oct 2009–present	Sep 2014–present
Overpass time	10:00 LST	09:30 LST	13:00 LST	13:30 LST
Nadir pixel size	30 km x 60 km/ 30 km x 240 km	40 km x 80 km/ 40 km x 40 km	10.5 km diameter	1.3 km x 2.25 km
Swath width	960 km	1920 km 960 km	750 km (5 measurements)	10.3 km
Revisit time	6 days	1.5 days/3 days	3 days	16 days
Spectral coverage (with respect to SIF)	604–805 nm	590–790 nm	755–773 nm	757–775 nm
Spectral resolution	0.48 nm	0.5 nm	0.025 nm	0.042 nm
Signal-to-noise ratio (SNR)	< 3000	< 2000	> 300	< 1000

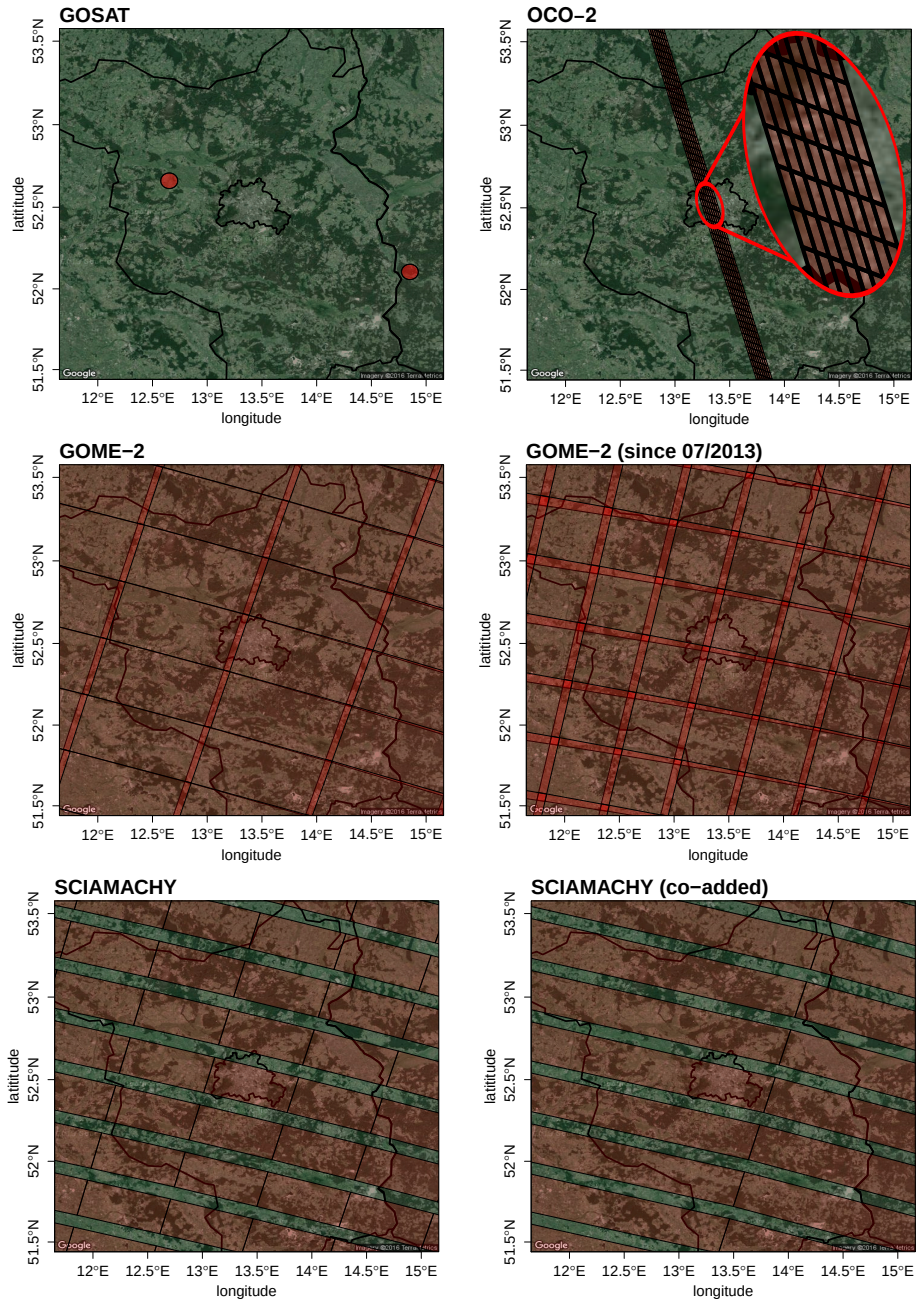


Figure 1.4: Individual spatial footprints (transparent reddish color) during one overpass for GOSAT (05.07.2009, computed from center latitude/longitude), OCO-2 (computed for orbit ground track 208, 16.07.2014), GOME-2 (03.07.2009 and 10.07.2014) on board MetOp-A, and SCIAMACHY (06.07.2009) based on the Brandenburg region with Berlin in the centered position. The corresponding nadir pixel size can be found in Table 1.1. The swath width of GOME-2 on board MetOp-A is reduced since 15th July 2013, resulting in a reduced ground pixel size. In case of SCIAMACHY, pixel were partly co-added due to data rate limitations.

narrow fitting windows. GOSAT-FTS and OCO-2 provide a very high spectral resolution (<0.05 nm), allowing to apply both physically-based and statistically-based approaches, but it should be mentioned that their spectral coverage is limited to the right shoulder of the second SIF emission peak (F740). In contrast, the spectral band setting of SCIAMACHY and GOME-2 essentially covers the full SIF emission spectrum, but their spectral resolution is much coarser (∼0.5 nm) compared to that from GOSAT and OCO-2. Therefore, the use of either deep Fraunhofer lines or broader fitting windows (statistically-based methods) is required when retrieving SIF from SCIAMACHY or GOME-2 data. Overall, the differences among the instrument specifications reflect the trade-off between spectral and spatial resolution, and signal-to-noise ratio (SNR).

Fig 1.4 illustrates the spatial sampling of the satellite instruments based on the Brandenburg region with Berlin in the centered position. The ascending orbit from OCO-2 results in a different orientation of the swath compared to GOSAT-FTS, SCIAMACHY, and GOME-2 which are flown on satellite platforms with descending orbits. In view of the sparse spatial sampling of GOSAT-FTS (two measurements in the considered area during one overpass), it might be questioned whether the SIF observations concerned are spatially representative. This applies also for OCO-2 measurements, where gaps in the order of 1.5° longitude arise between satellite footprints after a full repeat cycle (16 days), even though the single pixel size of 2 km x 1.3 km at nadir (along-track x across-track) outperforms those of the other instruments. Contrastingly, GOME-2 and SCIAMACHY in principle provide a spatially continuous sampling, but their nadir pixel size is much larger compared to that of GOSAT and OCO-2. The spatial resolution of GOME-2 (on board MetOp-A) has been slightly improved since 15th July 2013 by changing the swath width from 1920 km to 960 km, resulting in a reduced ground pixel size of about 40km x 40km (previously 80km x 40km). Unfortunately, SCIAMACHY's original ground pixel size of 30 x 60 km is only available in about 30% of cases. Due to data rate limitations, radiances in the spectral range of interest were partly downlinked at 240 km (four co-added pixel, ∼50%) or even at 480 km (eight co-added pixel, ∼20%) across track. In addition, SCIAMACHY alternately measures in nadir and limb mode, doubling the revisit time with respect to GOME-2 (at the same swath width, see Table 1.1). In summary, Fig 1.4 illustrates the intrinsic limitation of current satellite-based SIF data sets, arising from coarse spatial resolutions when the data are mapped. On the contrary, it must be considered that a high spectral and spatial resolution in combination with a spatially continuous sampling (neglecting technical limitations) would lead to an excessive amount of information to store, transmit, and process.

1.2.3 Prerequisites for a case study in Amazon forests

The Amazon region has been selected as a target area to assess the potential of spaceborne SIF observations regarding the monitoring of dense vegetation. In particular, the relevance of Amazon forests for the global carbon cycle and the challenging conditions for remote sensing in that area require continuing research and the development of innovative strategies. The following paragraphs provide a brief overview of the research background and aspects which need to be considered for tracking vegetation dynamics

in Amazon forests.

Global tropical rainforests account for more than one third of the terrestrial carbon uptake through photosynthesis (GPP), and build by far the most productive biome (Beer et al., 2010). The Amazon Basin spans an area as large as Western Europe, including the world's largest remaining rainforest. Thus, it is one of the most important areas to monitor vegetation status and photosynthetic activity to estimate the global carbon budget. Furthermore, understanding the metabolism of Amazon forests and identifying main drivers is essential to estimate the response to gradually changing climate conditions.

Even though Amazon forests generally retain green foliage all year round, evidence suggests that vegetation dynamics are subject to substantial seasonal swings (e.g., Myneni et al., 2007; Wu et al., 2016). For instance, many studies report a 'green-up', i.e., an enhanced leaf area along with an increase in photosynthetic activity during months with less intense precipitation, typically ranging from June-October (e.g., Huete et al., 2006; Samanta et al., 2012; Bi et al., 2015). However, there is a controversial debate on causes, characteristics, and magnitude of the seasonal swings in the Amazon region (e.g., Restrepo-Coupe et al., 2013; Morton et al., 2014; Guan et al., 2015). Besides the idea of water and/or light as key factors determining the photosynthetic activity, it has emerged that leaf phenology (leaf flushing, aging, and litterfall) could drive the seasonality of photosynthesis to a great extent (Samanta et al., 2012; Restrepo-Coupe et al., 2013; Wu et al., 2016). It appears that there is a large degree of interaction between leaf phenology and environmental conditions, while it remains challenging to identify main drivers valid for the entire Amazon region. Additionally, it is well known that reflectance-based VIs become weakly sensitive to vegetation changes under the conditions of the Amazon rainforest due to the saturation of reflected solar radiation in dense vegetation. Moreover, it has increasingly been recognized that view-illumination geometry and residual atmospheric effects must be carefully considered when evaluating satellite data with respect to seasonal vegetation dynamics (e.g., Bi et al., 2015; Maeda and Galvão, 2015; Hilker et al., 2015).

An alternative way for examining the photosynthetic seasonality in the Amazon Basin recently arose with the advent of SIF data sets, but only a few research articles using this new data stream for analyses addressing the Amazon region can be found in the literature so far (Lee et al., 2013; Parazoo et al., 2013; Guan et al., 2015). Lee et al. (2013) and Parazoo et al. (2013) examined GOSAT SIF data and reported that substantial seasonal variations in photosynthetic activity (proxied by SIF) occur predominantly in southern Amazon forests, associated with pronounced variations in precipitation. Guan et al. (2015) investigated the link between photosynthetic productivity and hydroclimate in global tropical evergreen forest regions by means of GOME-2 SIF and MODIS (Moderate Resolution Imaging Spectroradiometer) EVI (Enhanced Vegetation Index). They concluded that canopy photosynthesis can be maintained throughout the seasons if the annual water availability is sufficient. Nonetheless, the research with global SIF data is still in its infancy, compared with the numerous studies using reflectance-based VIs. In this context, it might be recalled that SIF is inevitably influenced by canopy structure effects as outlined in Sect. 1.2.1. Furthermore, directional effects could impact SIF mea-

surements similar to the hot spot effect in reflected solar radiation. Another point to be considered is the satellite overpass time because instantaneous illumination conditions may drive the observed seasonality in SIF, while daily averages of solar illumination/SIF would result in a different seasonal trajectory (e.g., different times for annual maximum and minimum values).

In view of this background, combining SIF data with other vegetation parameters, especially accounting for structural effects, might be necessary. One promising and innovative vegetation parameter for this purpose may be found in the canopy scattering coefficient (CSC). The underlying theory to compute CSC is based on decades of research on the concept of canopy spectral invariants (summarized by Huang et al., 2007), according to which the spectral reflectance can be expressed in terms of directional area scattering factor (DASF) and CSC (Knyazikhin et al., 2013). A detailed review of the methodology and related works can be found in Sect. 4.2.5. In essence, the spectral invariant DASF needs to be computed to relate spectroscopic measurements to CSC. For this purpose, the approach from Knyazikhin et al. (2013) has been adopted to derive DASF and CSC from atmospherically corrected GOME-2 data within the framework of this thesis. Thus, SIF and CSC are consistently derived from one sensor, providing the first data set of vegetation parameters based on spectroscopic measurements in red/NIR wavelengths which covers the entire Amazon Basin. Assumed to be particularly sensitive to leaf optical properties and less affected by view-illumination geometry as well as macro-structural properties (Samanta et al., 2012; Knyazikhin et al., 2013), the CSC might be related to leaf age effects when considering that matured tropical leaf spectra reveal an increased absorption in NIR wavelengths (Roberts et al., 1998). Hence, a basin-wide measure of leaf phenology (represented by CSC) could be established and related to photosynthetic activity (proxied by SIF). However, the experience with real-life CSC data is yet very limited and based on only two studies (Schull et al., 2011; Knyazikhin et al., 2013), thus more research is needed to evaluate the prospects and drawbacks of CSC. Chapter 4 therefore deals with both SIF and CSC, assessing the potential of these new vegetation parameters to track vegetation dynamics in Amazon forests with respect to the complex issues outlined above.

1.3 Objectives and research questions

The core objectives of this thesis are: (i) developing new methods for the global retrieval of SIF and (ii) exemplarily analyzing SIF data with respect to relevant influencing factors. Accordingly, the thesis aims to contribute to the development of state-of-the-art SIF retrieval techniques as well as to establish a basis for interpreting satellite-based SIF data. Both elements are crucial to assess global terrestrial photosynthesis through the monitoring of SIF from space, which is the overarching goal of the GlobFluo project. Based on the main objectives, chapters 2-4 seek to address the following research questions:

- How can SIF retrieval methods be streamlined in order to improve their efficiency and robustness?

- How do results from different retrieval approaches compare with each other and how consistent are different data sets?
- Which are the main confounding factors and pitfalls when analyzing SIF data?

Beyond SIF, chapter 4 additionally considers another promising vegetation parameter, namely the CSC derived along with the DASF, which is in simplified terms a measure of the canopy's complexity. The research question associated with the expanded core research topic is:

- Can the concept of canopy spectral invariants (underlying theory to compute DASF and consequently CSC) be utilized to add information regarding structural effects on SIF data?

In the following, I will briefly introduce the individual research articles (chapters 2-4) accompanied by their specific objectives, addressing the aforementioned overall research questions.

Specific Objectives

Chapter 2 describes a physically-based retrieval method comparable to that proposed by Frankenberg et al. (2011a,b), published as:

Philipp Köhler, Luis Guanter and Christian Frankenberg: "Simplified Physically Based Retrieval of Sun-Induced Chlorophyll Fluorescence From GOSAT Data." Geoscience and Remote Sensing Letters, IEEE 12.7 (2015): 1446-1450.

Though the basis for this research article was established in my Master's Thesis (Köhler, 2012), I continued working on this subject to refine the method, increase the computational speed, and produce a global data set of SIF using GOSAT data. Specific objectives are:

- Simplifying some of the assumptions in the state-of-the-art retrieval approach from Frankenberg et al. (2011a) without a loss in retrieval accuracy.
- Comparing results from existing retrieval approaches (Frankenberg et al., 2011b; Guanter et al., 2012) with results from the presented retrieval approach.
- Evaluating the impact of cloud contamination on SIF time series qualitatively by empirical radiance thresholds.

Chapter 3 proposes a new statistically-based approach to derive SIF from medium resolution spaceborne spectrometers, published as:

Philipp Köhler, Luis Guanter and Joanna Joiner: "A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data" Atmospheric Measurement Techniques, 8, 2589-2608, 2015.

The method described takes advantage of previous studies (Guanter et al., 2013; Joiner et al., 2013) and has been applied to data acquired from GOME-2 on board MetOp-A and SCIAMACHY on board ENVISAT to generate two new SIF data sets. These data sets span the August 2002–March 2012 (SCIAMACHY) and the January 2007–December 2011 (GOME-2) time period, respectively. It might be noted that the SCIAMACHY fluorescence retrievals in the 740 nm spectral region are the first ones in the literature. Specific objectives are:

- Providing a solution for the arbitrary selection of free model parameters in the retrieval approaches from Guanter et al. (2013) and Joiner et al. (2013).
- Evaluating the retrieval performance through a simulation-based retrieval test.
- Applying the same retrieval to data acquired from different sensors (GOME-2 and SCIAMACHY).
- Comparing SIF data from different sensors and retrieval approaches with respect to their consistency (GOSAT, GOME-2, and SCIAMACHY).
- Identifying possible limitations and shortcomings of the retrieval results.
- Evaluating the impact of cloud cover on SIF time series by using an additional data set of relative cloud fractions.

Chapter 4 comprises a case study in Amazon forests, submitted as:

Philipp Köhler, Luis Guanter, Hideki Kobayashi, Sophia Walther and Wei Yang: "Assessing the potential of Sun-Induced Fluorescence and the Canopy Scattering Coefficient to track large-scale vegetation dynamics in Amazon forests", Remote Sensing of Environment, submitted on April 23, 2016.

In essence, this article examines prospects and drawbacks of SIF for the monitoring of dense vegetation in the Amazon region, additionally involving the theory of canopy spectral invariants by examining CSC data, which have been derived along with the DASF. For this purpose, the GOME-2 SIF data set has been extended until April 2015 (previously January 2007–December 2011), while this study examines GOME-2 SIF data only from August 2013 onwards to take advantage of the reduced ground pixel size (see Sect. 1.2.2). Moreover, the first global DASF/CSC data set has been produced from atmospherically corrected GOME-2 data for the same time period (August 2013–April 2015). This part of the thesis also includes a series of supplementary data streams (simulated, satellite-based, and re-analysis data) necessary for the atmospheric correction and beyond that, for providing a basis to interpret the results. Specific objectives are:

- Analyzing the sensitivity of SIF and CSC to atmospheric, structural, and view-illumination (directional) effects under the conditions of Amazon forests with respect to the well established NDVI (Normalized Difference Vegetation Index).

- Comparing SIF data from different sensors (GOME-2, OCO-2) with particular emphasis on differences due to their overpass times (morning and noon, respectively).
- Examining prospects of CSC for the monitoring of dense vegetation in general and of leaf optical properties in particular.
- Investigating large-scale vegetation dynamics represented by SIF and CSC in Amazon forests with respect to environmental conditions.

Chapter 2

Simplified physically-based retrieval of sun-induced chlorophyll fluorescence from GOSAT data

Philipp Köhler[1,2], Luis Guanter[1,2], and Christian Frankenberg[3].

Journal: IEEE Geoscience and Remote Sensing Letters (Volume: 12, Issue: 7), pp. 1446–1450.

Status: Manuscript received September 22, 2014; revised January 15, 2015 and February 18, 2015; accepted February 19, 2015. Date of publication March 9, 2015.

Publisher version: http://dx.doi.org/10.1109/LGRS.2015.2407051

[1] Institute for Space Sciences, Freie Universität Berlin, Berlin, Germany.
[2] German Centre for Geosciences(GFZ), Remote Sensing Section, Potsdam, Germany.
[3] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA.

Abstract

First global retrievals of sun-induced chlorophyll fluorescence (F_s) have been achieved in the last years by using data from the Fourier Transform Spectrometer (FTS) on-board the Japanese Greenhouse Gases Observing Satellite (GOSAT). The high spectral resolution (approximately 0.025 nm) of the FTS enables measurements of F_s by the evaluation of the in-filling of solar Fraunhofer lines around 755 nm. This study presents a new F_s retrieval algorithm (GARLiC, for *GOSAT Retrieval of chlorophyll fluorescence*) and compares its results to F_s from two previously implemented approaches. GARLiC is intended to simplify some of the assumptions of existing retrieval approaches without a loss in retrieval accuracy. We show that GARLiC F_s retrievals are comparable to the previously used methods. We also assess the effect of clouds on F_s retrieval from GOSAT data through the analysis of the effect of different cloud filter thresholds on F_s time series. Our results show a low sensitivity of Fraunhofer line-based F_s retrievals to cloud contamination.

2.1 Introduction

Sun-induced chlorophyll fluorescence (F_s) is an electromagnetic signal emitted in the 650–800 nm spectral region by the chlorophyll-a of photosynthetically active vegetation (e.g. Zarco-Tejada et al., 2003). F_s represents a part of the excess energy during the process of photosynthesis and provides a measure of photosynthetic activity.

Although several studies have dealt with the estimation of F_s from airborne spectrometers in the last decade (see Meroni et al., 2009, and references therein), F_s maps remained only available for selected study areas and times. An extension to the global scale has been achieved in the last several years by the use of satellite sensors (Joiner et al., 2011; Frankenberg et al., 2011b; Guanter et al., 2012; Joiner et al., 2012, 2013). It can be stated that all methods make use of the in-filling of solar Fraunhofer lines and/or atmospheric absorption features by F_s. While all initial global F_s retrievals (Joiner et al., 2011; Frankenberg et al., 2011b; Guanter et al., 2012; Joiner et al., 2012) have used data from the Fourier Transform Spectrometer (FTS) on-board the Japanese Greenhouse Gases Observing Satellite (GOSAT, Kuze et al., 2009), recent studies (Joiner et al., 2013) explored also the use of medium spectral resolution spectrometers such as the Global Ozone Monitoring Experiment (GOME-2) on-board of the Meteorological Operational Satellites (MetOp-A and MetOp-B). The high spectral resolution of the GOSAT-FTS (approx. 0.025 nm) enables an accurate F_s retrieval, but only a limited spatial coverage is provided by the sensor. On the contrary, GOME-2 provides a continuous spatial sampling, but the pixel size (40 x 80 km^2) is much coarser than that from the GOSAT-FTS (10.5 km diameter). The first global F_s retrieval from GOSAT-FTS data (Joiner et al., 2011) was based on the potassium (K) I line at 770.1 nm (using a retrieval window between 769.9–770.25 nm). It turned out that the narrow spectral

range between 758.45–758.85 nm (covering another Fraunhofer line) used by Joiner et al. (2012) is more suitable to retrieve F_s due to a closer proximity to the second of two characteristic peaks of the fluorescence emission spectrum at 740 nm. Alongside this development, Frankenberg et al. (2011b) and Guanter et al. (2012) used two broader fitting windows covering the 756–759 nm and 770.5–774.5 nm spectral regions, which has led to an improved signal-to-noise ratio and thus to less noisier retrievals. Guanter et al. (2012) based their method on a singular value decomposition (SVD) technique with the basic assumption that any radiance spectrum in a narrow spectral window can be expressed as a linear combination of singular vectors plus fluorescence. This technique has the advantage of providing simpler and faster retrievals based on a linear forward model. A caveat of this method is that fine-tuning is needed in order to select the optimal number of singular vectors, which has an effect on the retrieval accuracy and precision. The most accurate results can therefore presumably be obtained using the physically-based method proposed by Frankenberg et al. (2011b). The spectral window between 756–759 nm is particularly suitable to retrieve F_s, because this wavelength range is devoid of atmospheric absorption features and highly insensitive to atmospheric scattering effects. However, applying the retrieval also to the 770 nm region (containing weak O_2 absorption lines) enables a supplementary retrieval of surface pressure, which is useful to filter cloudy scenes if ancillary surface pressure data show considerably higher values. Cloud contamination would potentially lead to an underestimation of F_s as indicated by simulation studies (e.g. Frankenberg et al., 2012).

This investigation proposes a simplified physically-based fluorescence retrieval method (similar to Frankenberg et al., 2011b) in the spectral range between 755–759 nm, which allows faster retrievals without a loss of performance with respect to currently available algorithms. We use only one fitting window in the region around 757 nm and linearize the estimation problem in order to solve the forward model without iterations. A secondary objective is the exploration of the impact of clouds on our F_s retrieval from GOSAT. We compare results of the presented retrieval with those obtained from existing methods and examine uncertainties caused by noise and cloud contamination.

2.2 Materials and Methods

2.2.1 GOSAT-FTS

GOSAT (Kuze et al., 2009) has a sun-synchronous orbit at an altitude of approximately 666 km with an overpass time around 13:00 local solar time (dayside) and a revisit time of three days. One cross track of the FTS has a swath of 750 km and contains from three to five measurements, each covering a circular area with a diameter of about 10.5 km at ground level. The distance between measurements along the satellite track ranges between 90 and 280 km. The GOSAT-FTS unit is composed of two identically constructed detectors measuring perpendicular (S) and parallel (P) polarized light. The FTS level 1B product (Kuze et al., 2009) consists of uncalibrated spectra in voltage per wavenumber [V cm] with two independent measurements per spectrum, which are converted to

spectral radiances through the application of radiometric calibration coefficients.

The primary purpose of the GOSAT mission is the estimation of CO_2 and CH_4 concentrations. F_s retrievals from GOSAT-FTS data are also possible because of the high spectral resolution of the FTS. The first band of the FTS samples the 755–773 nm spectral window (Fig. 2.1 depicts an example), which is nearby the second peak of the fluorescence emission spectrum (∼740 nm), with a spectral resolution of about 0.025 nm.

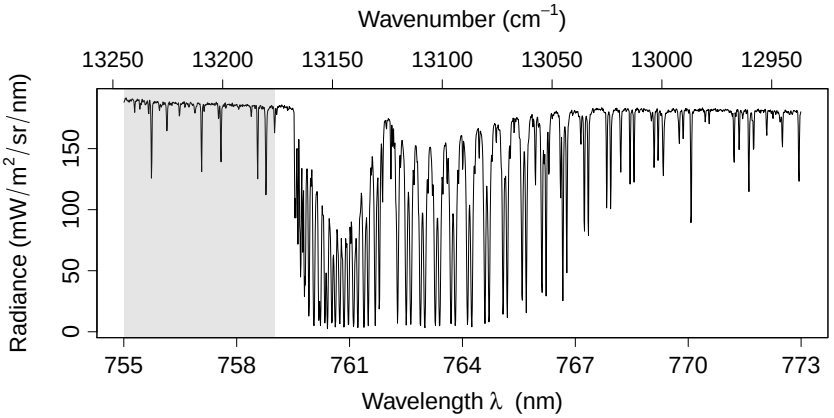


Figure 2.1: Sample GOSAT-FTS spectrum in band 1P (P-parallel polarized light). The spectral window to retrieve F_s (755–759 nm) is marked in gray.

2.2.2 GOSAT Retrieval of chlorphyll fluorescence (GARLiC)

The retrieval of F_s from high resolution radiance measurements at the top-of-atmosphere (TOA) relies on the idea that any additive signal fills in absorption features, i.e., makes their fractional depth smaller which was firstly mentioned by Plascyk (1975). F_s is such an additive signal, but it must be remarked that F_s is about 100 times smaller than the reflected solar radiation in the near-infrared (NIR) spectral region (e.g. Amorós-López et al., 2008).

The TOA radiance measured by the instrument (L_{TOA}) over a Lambertian reflecting surface can be formulated as

$$L_{TOA} = L_0 + \frac{\frac{1}{\pi} \cdot \rho_s \cdot T_{\downarrow} \cdot \mu_s \cdot I_{sc} T_{\uparrow} + F_s \cdot T_{\uparrow}}{1 - S \cdot \rho_s}, \tag{2.1}$$

where L_0 is the backscattered atmospheric radiation, ρ_s is the surface reflectance, I_{sc} is the solar irradiance at the TOA, μ_s is the cosine of the sun zenith angle, $T_{\uparrow}$ is the upward atmospheric transmittance, $T_{\downarrow}$ is the downward atmospheric transmittance, S is the atmospheric spherical albedo and F_s is the steady state sun-induced fluorescence. All variables in Eq. (2.1) have a spectral component. In the 755–759 nm NIR window where we perform the retrievals, the atmosphere is roughly transparent ($T_{\uparrow} \approx 1$) for near-nadir

observations under clear-sky conditions (Guanter et al., 2012), and $S \cdot \rho_s << 1$. Splitting the atmospheric path radiance L_0 into the product of solar irradiance and atmospheric reflectance ($\mu_s \cdot I_{sc} \cdot \rho_0$) allows us to express Eq. (2.1) in the form:

$$L_{TOA} = \frac{\mu_s \cdot I_{sc}}{\pi} \cdot (\rho_0 + \rho_s \cdot T_{\downarrow}) + F_s. \tag{2.2}$$

A spectrally-linear variation of atmospheric scattering and surface reflectance in the 755–759 nm range can be assumed, because spectral properties as well as atmospheric scattering indicate no discontinuity in this narrow spectral region. With this assumption, the forward model results in:

$$L_{TOA}(F_s, a_i) =< I_{sc} > \cdot(a_0 + a_1 \cdot \lambda) + F_s. \tag{2.3}$$

We take the widely used high-resolution solar irradiance spectrum (I_{sc}) from Kurucz (2005) and convolve it to the lower resolution spectrometer grid (from 0.001 nm to 0.025 nm), which is represented by the <> operation. There are three unknowns to generate a synthetic measurement at 350 spectral points with the forward model, a_0 and a_1 for the modeling of atmospheric scattering and surface reflectance, and the amount of F_s between 755–759 nm. We processed spectra acquired in each of the two polarizations independently.

Frankenberg et al. (2011b) observed an offset in the O_2-A band of GOSAT-FTS spectra, which was endorsed by GOSAT Project members. Since F_s and the radiance offset are both additive signals, which result in apparently weaker absorption lines, a correction is a necessary but critical implementation step. As found by previous studies (Frankenberg et al., 2011b; Guanter et al., 2012; Joiner et al., 2012) the magnitude of the offset is changing with the input radiance level, which introduces high biases especially for higher radiance levels (associated with bright surfaces and/or small solar zenith angles). Thus, retrieval results from regions where no vegetation fluorescence is expected (e.g. ice shield in Greenland, Antarctica) are recorded in dependency of the average TOA radiance in the FTS band 1. The resulting function is then smoothed by a boxcar average using a width of 100 points in order to obtain the final transfer function between input radiance level and F_s. Because this zero-level offset also has a slight temporal dependency and is different for the two detectors, the correction is applied on a monthly basis and independently of each polarization.

The GARLiC (*GOSAT Retrieval of chlorophyll fluorescence*) algorithm for the production of global composites of F_s from GOSAT-FTS spectra are based on the forward model in Eq. (2.3), which was implemented by following the steps below:

1. Perform a prefiltering to reduce the number of unnecessary retrievals by using only spectra acquired in the high radiometric gain mode of the FTS instrument and with 100% land surface in the field-of-view (determined with a land-sea mask).

2. Determine the exact spectral position of the measurement since the instrument is not perfectly calibrated and Doppler shifts (caused by relative velocity between

earth and sun) as well as a relative movement between satellite and earth (at off-nadir angles) may occur. The 1-D Brent's minimization method is used to calculate the wavelength shift leading to a minimal residual between the FTS measurement and the solar irradiance after convolution.

3. Calculate $< I_{sc} >$ through the spectral convolution of the high-resolution solar irradiance from Kurucz for the updated wavelength array.

4. Fit the TOA radiance measurements with the forward operator in Eq. (2.3), using an ordinary least squares fit in order to obtain F_s and a_i.

5. Correct the obtained F_s values from a zero-level offset caused by a nonlinearity of the detector response (Frankenberg et al., 2011b; Guanter et al., 2012; Joiner et al., 2012).

6. Perform a quality filtering to exclude cloudy cases and nonphysical results (e.g. extreme F_s values / outliers) by the following criteria:
   - average radiance in the FTS band 1 (755–773 nm) amounts $1.5 - 8 \cdot 10^{-7}$ W/cm^2/sr/cm (native unit of the FTS; according to Frankenberg et al., 2011b; Guanter et al., 2012)
   - $0.8 < \chi^2_r < 1.5$, χ^2_r is the reduced χ^2 (goodness of fit).
   - $|F_s| < 5$mW/m^2/sr/nm (according to Guanter et al., 2012)

7. Produce global composites by averaging quality-filtered retrievals from each polarization in 2° x 2° gridboxes. The coarse resolution is due to the relatively high retrieval noise and the poor spatial sampling of the measurements.

The 1-σ error of each retrieval is calculated by propagating the measurement noise, which is estimated from the first and last 500 channels in the band 1 spectrum (Frankenberg et al., 2011b; Guanter et al., 2012).

Fig. 2.2 shows a sample fit of the forward model Eq. (2.3) to a real GOSAT-FTS measurement in the 755–759 nm window. The residual consists only of white noise as expected for an FTS (Ljung-Box test indicates independently distributed data). Since the residual is free from spectral features, it can be assumed that a spectral stretch is negligible in the considered spectral range.

As mentioned in Sect. 2.1, the retrieval approach from Frankenberg et al. (2011b) is similar to the presented one. The main difference is the decoupling of the spectral shift estimation from the forward model and the neglection of spectral stretch effects in the GARLiC algorithm. As opposed to the non-linear forward operator proposed by Frankenberg et al. (2011b), the GARLiC forward model shown in Eq. (2.3) is linear and thus faster, since there is no iterative calculation needed. Furthermore, GARLiC uses only one fitting window under the presumption that the presented filtering criteria are sufficient and a supplementary retrieval of surface pressure is not required.

We evaluate the performance of the $8 \cdot 10^{-7}$ W/cm^2/sr/cm threshold for cloud screening (also used by Frankenberg et al., 2011b; Guanter et al., 2012) by using a more relaxed

threshold ($10 \cdot 10^{-7}$ W/cm^2/sr/cm) and a more restrictive one ($6 \cdot 10^{-7}$ W/cm^2/sr/cm) as it is shown in Sect. 2.3.3.

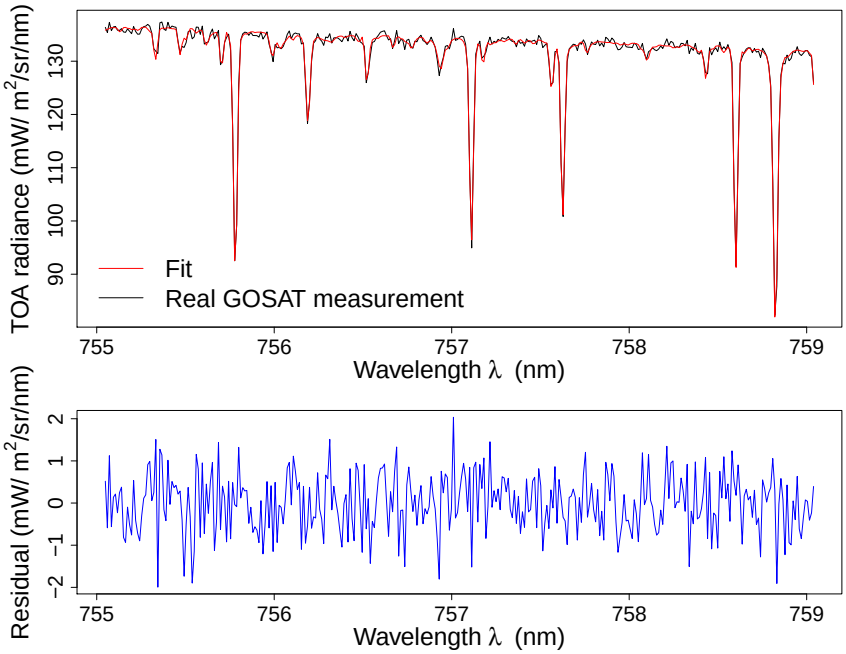


Figure 2.2: Fit of a GOSAT-FTS measurement in the 755–759 nm fitting window with the forward model in Eq. (2.3). The spectral residual of the fit is displayed in the bottom panel.

2.3 Results

2.3.1 Monthly composites

Although GOSAT has a revisit time of three days, global F_s composites are only meaningful when averaging retrieval results on larger time scales such as monthly means (because of the presence of clouds and the high retrieval noise). The error weighted monthly mean (according to Frankenberg et al., 2011b; Guanter et al., 2012) of F_s derived with GARLiC for July 2010 and its standard error of the weighted average is depicted in Fig. 2.3 as an example. The highest July F_s values were observed in the northern mid-latitudes associated with crops in the US Corn Belt, eastern Europe, and China. These pronounced F_s averages have also been observed in other studies (Frankenberg et al., 2011b; Guanter et al., 2012; Joiner et al., 2012). As expected, high F_s values occur also in tropical regions, whereas regions with none or sparse vegetation show values close to zero. There is no measurement available over the Sahara because the low radiometric gain mode of the FTS instrument is used to avoid a saturation in the detector over bright surfaces. As mentioned in Sect. 2.2.2, these measurements are not processed by GARLiC.

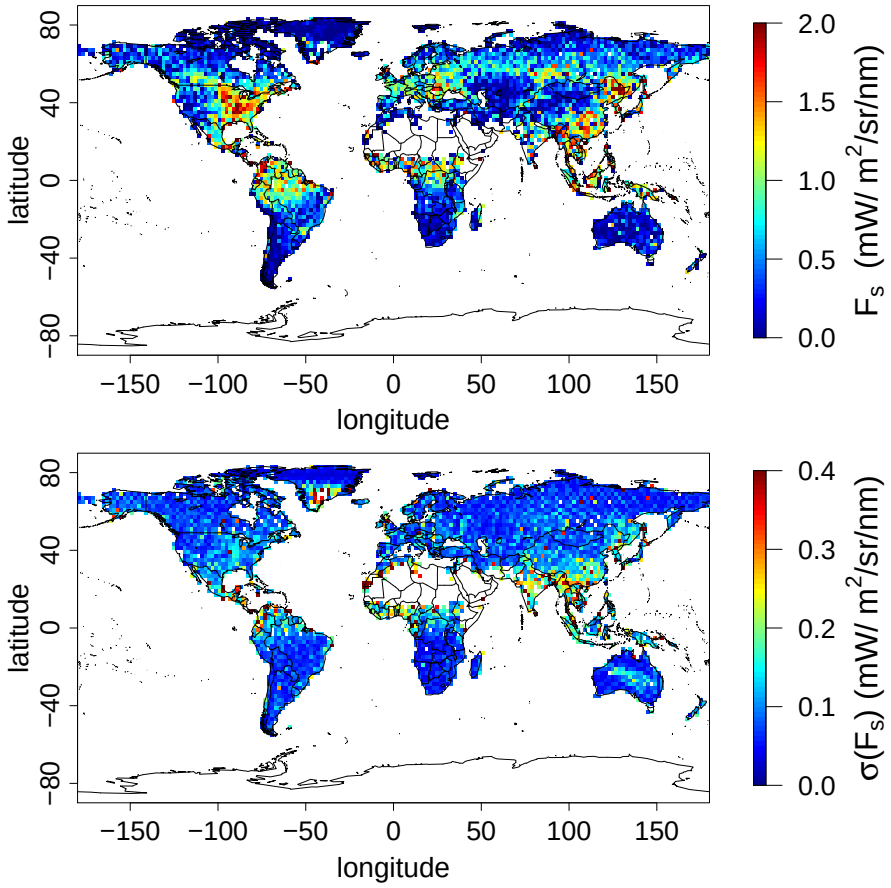


Figure 2.3: Error weighted monthly mean of F_s derived with GARLiC from GOSAT-FTS data for July 2010 (top) and standard error of the weighted average (bottom). No F_s retrievals were performed in white regions.

The standard error of the weighted mean is an estimate of the overall uncertainty of monthly composites. It represents the instrumental noise weighted by the number of observations per cell. In our July 2010 composites, the error usually amounts to one third of the derived F_s value and decreases with a higher number of measurements per cell. Therefore, more observations in northern latitudes resulting from the satellite orbit reduce the error, while more sparse overpasses produce a higher error in tropical regions.

2.3.2 Comparison with existing retrieval results

F_s annual averages for the 06/2009–05/2010 time period have been used to compare our retrieval results with those obtained by Frankenberg et al. (2011b) and Guanter et al. (2012). Fig. 2.4 shows that values from Frankenberg et al. (2011b) are very similar but

slightly higher than results from GARLiC. The main reason for the small bias is a slight polarization dependence in retrievals from Frankenberg et al. (2011b). Although we do not expect any sensitivity to polarization in the F_s retrievals, the data from Frankenberg et al. (2011b) show slightly higher values in the P- than in the S-polarization (on average 0.1 mW/m^2/sr/nm for retrievals included in Fig. 2.4). This polarization dependence is not observed in the GARLiC retrievals (mean difference is $9 \cdot 10^{-4}$ mW/m^2/sr/nm for retrievals included in Fig. 2.4), which provide more consistent values for either measurement polarization. The reason can be attributed to the post processing since there is no difference in the formulation of the two methods that could lead to more or less sensitivity to the polarization. Comparing the annual GARLiC results with those from Guanter et al. (2012) reveals that there is a clear linear relationship with an increasing difference at higher F_s values (with higher F_s values from GARLiC).

In order to investigate the temporal consistency between GARLiC and the other two data sets, we collocated the retrieval output (without averaging results in gridboxes) with the land cover classification of the International Geosphere and Biosphere Programme (IGBP, Friedl et al., 2002, derived from MODIS data) and produced biome specific time series. The three approaches compare very well for almost all areas, whereas one representative example is shown in Fig. 2.5. The 15-day average of F_s for evergreen needleleaf forest, located in the northern temperate zone (latitudes above 40° N), for the 06/2009–04/2011 period shows that an annual cycle with high F_s values during summertime is present in the three time series. Similarly to the annual comparison, F_s values from Frankenberg et al. (2011b) are slightly higher than the GARLiC values, also in winter when we expect very low F_s because of low illumination angles. Both the results

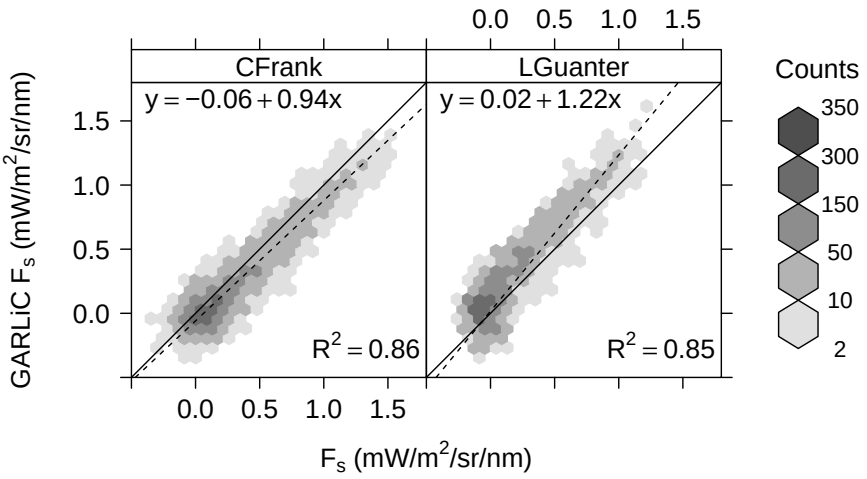


Figure 2.4: F_s retrieval results from GARLiC vs. results provided by Frankenberg et al. (2011b) (CFrank) and Guanter et al. (2012) (LGuanter) respectively. The F_s annual averages from 06/2009–05/2010 (rastered on 2° x 2° grid) have been used, whereas each evaluated gridbox contains at least 24 retrievals and is available in all data sets. The linear fits are depicted as dotted lines.

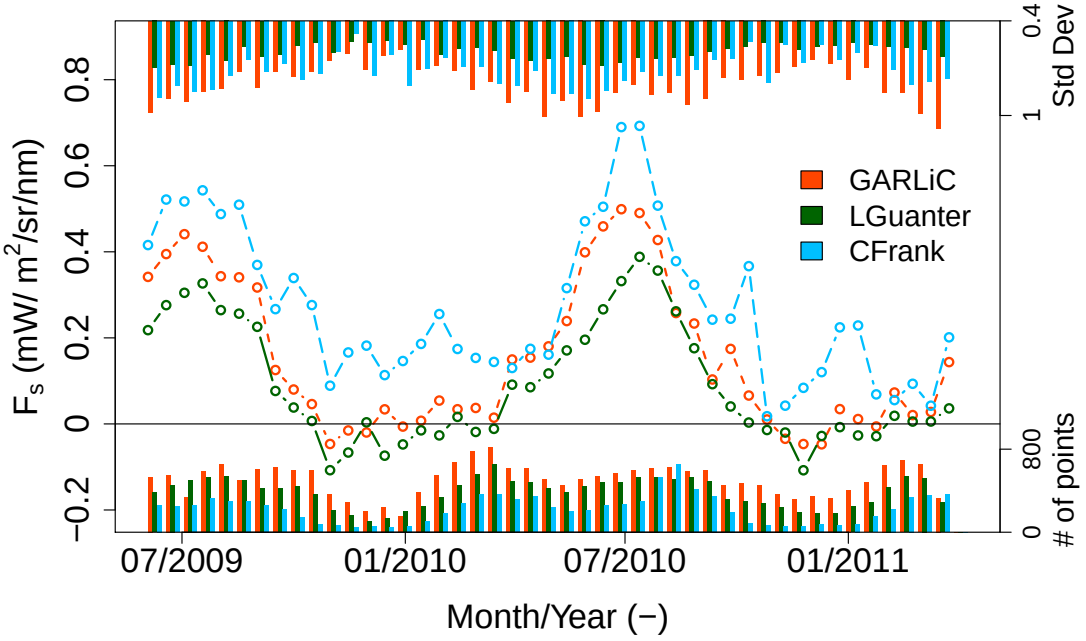


Figure 2.5: Comparison between GARLiC, Frankenberg et al. (2011b) (CFrank) and Guanter et al. (2012) (LGuanter) 15-day F_s means between 06/2009–04/2011 for evergreen needleleaf forest regions in latitudes above 40° N. The number of retrievals per time step and their standard deviation are depicted as barplots.

from the statistical approach by Guanter et al. (2012) and GARLiC show values falling slightly below zero during the wintertime. In theory the F_s value should be zero when the signal is totally absorbed or none is emitted. However, it should be emphasized that slight negative as well as slight positive results (up to ± 0.1 mW/m^2/sr/nm) can occur also in the absence of any F_s emission because of a relative high measurement noise. There are significant differences in 15-day F_s averages (especially between Frankenberg et al., 2011b; Guanter et al., 2012), but due to the lack of a ground truth it remains uncertain which values are the most realistic. Nevertheless, almost all GARLiC values lie in between those from the other two algorithms with F_s values close to zero during inactive periods.

Overall the GARLiC method provides the highest number of measurements for the evergreen needleleaf forest biome, which is due to a more relaxed quality filtering and leads to the highest standard deviation. Frankenberg's method shows the lowest number of retrievals for the 15-day mean and a standard deviation slightly lower than the GARLiC method. The standard deviation derived with the data from Guanter et al. (2012) is smaller than the other ones with a slightly lower number of retrievals than the GARLiC algorithm. It is important to note that the differences in the temporal series in Fig. 2.5 are not only due to the actual fundamental differences between the retrieval methods, but also to the subsequent quality filtering criteria applied by each algorithm. Thus, the

reduced number of retrievals from Frankenberg et al. (2011b) might be a result of the auxiliary surface pressure filter criterion. However, since spatial patterns and temporal course of F_s values retrieved with GARLiC and Frankenberg et al. (2011b) are similar, it might be assumed that the supplementary retrieved pressure is not essential for retrieving comparable F_s values.

2.3.3 Impact of clouds on the retrieval

We have used the temporal series of F_s between 06/2009–04/2011 to examine the impact of clouds on the retrieval. Although simulations indicate a low sensitivity to cloud coverage (Frankenberg et al., 2012), we still expect that clouds may affect the retrievals of F_s over tropical rainforest areas which present an almost persistent cloudy condition at the GOSAT overpass time (13:00 LT). Therefore, we have used GARLiC to produce three different time series of F_s over evergreen broadleaf forest regions (mostly tropical rainforest; identified by the IGBP land classification Friedl et al., 2002) using average radiance thresholds of $6/8/10 \cdot 10^{-7}$ W/cm^2/sr/cm (shown in Fig. 2.6) which corresponds to a strict/normal/relaxed cloud filtering.

Overall, the F_s values decrease slightly by relaxing the cloud filter while the standard deviation increases. Applying a more restrictive cloud filter implies a significant loss of

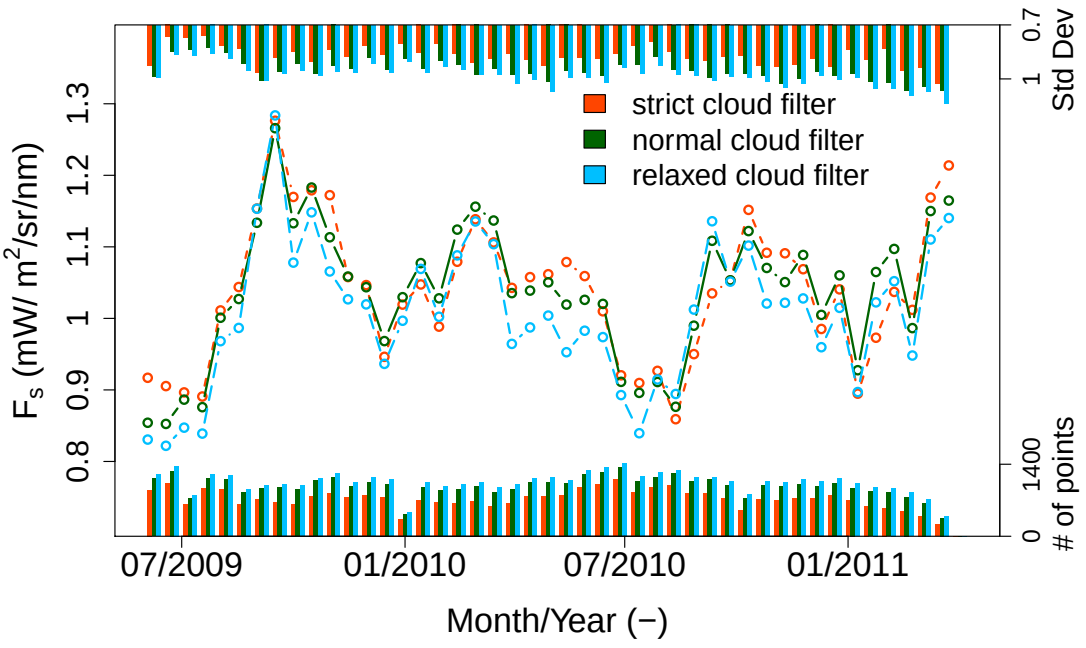


Figure 2.6: Temporal series of F_s retrievals from GARLiC over evergreen broadleaf forest regions using a strict/normal/relaxed cloud filter ($6/8/10 \cdot 10^{-7}$ W/cm^2/sr/cm). The time series corresponds to 15-day means between 06/2009–04/2011.

measurements without an appreciable change of the 15-day mean values. This confirms the low sensitivity of the retrieval to cloud coverage: strongly different levels of cloud contamination do not translate into similar changes in the retrieved F_s values. In contrast, a strict cloud filter leads to smaller fluctuations and is not affecting the reported values apparently, but it is associated with a loss of measurements. For this reason, a filter value of $8 \cdot 10^{-}7$ W/cm^2/sr/cm is a reasonable compromise between the loss of measurements and an acceptable influence on reported values.

In view of these results and of others based on simulations Frankenberg et al. (2012), it can be concluded that F_s retrievals are robust against cloud contamination.

2.4 Summary

A new approach to retrieve NIR F_s from GOSAT-FTS data was introduced in this contribution. It demonstrates that simplifying some assumptions of the existing algorithm by Frankenberg et al. (2011b) is still leading to comparable results. The simplifications consist of using only one fitting window, neglecting the spectral stretch and decoupling the spectral shift estimation from the forward operator so that the GARLiC forward model is linear and easy to invert with the ordinary least squares method. Thus, previous findings achieved by more stringent retrieval approaches could be confirmed. In addition, we have shown that F_s variations are increasing in presence of clouds using 15-day means but changes in reported values are only marginal. For this reason, it can be assumed that GARLiC and other F_s retrieval algorithms (based on Fraunhofer lines in spectral regions devoid of significant atmospheric absorption) are barely affected by cloud contamination. These findings impact future developments on F_s retrieval for the recently launched OCO-2 system (Frankenberg et al., 2014), which has a higher signal to noise ratio at a similar spectral resolution as compared to the GOSAT-FTS and provides about 100 times more measurements than GOSAT.

Author Contributions

Philipp Köhler and Luis Guanter designed the study. Philipp Köhler wrote code, processed the GOSAT data, analysed the output data, and wrote the manuscript. Luis Guanter gave conceptual advice, provided the LGuanter comparison data set, discussed the results and implications, and commented on the manuscript at all stages. Christian Frankenberg provided the CFrank comparison data set and commented on the manuscript.

Chapter 3

A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data

Philipp Köhler[1,2], Luis Guanter[1,2], and Joanna Joiner[3].

Journal: Atmospheric Measurement Techniques (Volume: 8), pp. 2589–2608.

Status: Manuscript received September 12, 2014. Published in Atmospheric Measurement Techniques Discussions December 04, 2014. Revised June 03, 2015; accepted June 04, 2015. Date of publication June 26, 2015.

Publisher version: http://dx.doi.org/10.5194/amt-8-2589-2015

[1] Institute for Space Sciences, Freie Universität Berlin, Berlin, Germany.
[2] German Centre for Geosciences(GFZ), Remote Sensing Section, Potsdam, Germany.
[3] NASA Goddard Space Flight Center, Greenbelt, MD, USA.

Abstract

Global retrievals of near-infrared sun-induced chlorophyll fluorescence (SIF) have been achieved in the last few years by means of a number of space-borne atmospheric spectrometers. Here, we present a new retrieval method for medium spectral resolution instruments such as the Global Ozone Monitoring Experiment-2 (GOME-2) and the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY). Building upon the previous work by Guanter et al. (2013) and Joiner et al. (2013), our approach provides a solution for the selection of the number of free parameters. In particular, a backward elimination algorithm is applied to optimize the number of coefficients to fit, which reduces also the retrieval noise and selects the number of state vector elements automatically. A sensitivity analysis with simulated spectra has been utilized to evaluate the performance of our retrieval approach. The method has also been applied to estimate SIF at 740 nm from real spectra from GOME-2 and for the first time, from SCIAMACHY. We find a good correspondence of the absolute SIF values and the spatial patterns from the two sensors, which suggests the robustness of the proposed retrieval method. In addition, we compare our results to existing SIF data sets, examine uncertainties and use our GOME-2 retrievals to show empirically the relatively low sensitivity of the SIF retrieval to cloud contamination.

3.1 Introduction

During the process of photosynthesis, the chlorophyll-*a* of photosynthetically active vegetation emits a small fraction of its absorbed energy as an electromagnetic signal (e.g., Zarco-Tejada et al., 2003). This signal, called sun-induced fluorescence (SIF), takes place in the 650–800 nm spectral region. Several studies have addressed the estimation of SIF from ground-based, airborne and spaceborne spectrometers in the last decade (see Meroni et al., 2009, and references therein). Here, we focus on SIF retrieval methods from space and their achievements.

The first global SIF observations have been achieved in the last 4 years by studies from Joiner et al. (2011), Frankenberg et al. (2011a) and Guanter et al. (2012) using data from the Fourier transform spectrometer (FTS) on board the Japanese Greenhouse Gases Observing Satellite (GOSAT). The first band of the GOSAT-FTS samples the 755–775 nm spectral window with a high spectral resolution of approximately 0.025 nm, which enabled an evaluation of the in-filling of solar Fraunhofer lines around the O_2 A absorption band by SIF. Joiner et al. (2011) based their retrieval on the strong K line around 770.1 nm on condition that a measured irradiance spectrum is available. In contrast, Frankenberg et al. (2011b) and Guanter et al. (2012) used two micro fitting windows around 757 and 770 nm by means of a reference solar irradiance data set. The 757 nm spectral region contains several solar Fraunhofer lines and is devoid of

significant atmospheric absorption, which minimizes the impact of atmospheric effects on the retrieval. Two micro fitting windows have also been used by Joiner et al. (2012) to reduce noise, whereas SIF-free Earth radiance spectra served as a reference. The method proposed by Frankenberg et al. (2011a) relies on the physical modeling of the in-filling of several solar Fraunhofer lines by SIF using the instrumental line shape function and a reference solar irradiance data set. The surface reflectance, atmospheric scattering as well as wavelength shifts have to be estimated for each measurement. Instead of explicitly modeling these parameters for each measurement, Guanter et al. (2012) proposed a data-driven approach, which is based on a singular value decomposition (SVD) technique. The basic assumption for this retrieval method is that any radiance spectrum can be expressed as a linear combination of singular vectors plus fluorescence. A caveat of this technique is an arbitrary selection of the optimum number of singular vectors, which has an effect on the retrieval accuracy and precision.

One intrinsic limitation of the GOSAT-FTS data arises from the coarse resolution of global maps ($2^\circ \times 2^\circ$), which is caused by a poor spatial sampling and a relatively high retrieval noise (Frankenberg et al., 2011b). Therefore, it was crucial that results from Joiner et al. (2012, 2013) indicated that instruments with lower spectral resolutions but better spatial coverage are also capable of estimating SIF from space. Joiner et al. (2012) examined the in-filling of the deep calcium (Ca) II solar Fraunhofer line at 866 nm by SIF using the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) satellite instruments, which has a spectral resolution of approximately 0.5 nm. Spatial and temporal variations of retrieved SIF were found to be consistent, although there is only a weak intensity of SIF in the examined spectral range. In this context, it should be considered that the representativeness of spatially mapped SIF is enhanced by using data from instruments which provide a continuous spatial sampling.

A further achievement was the data-driven study of Joiner et al. (2013). They proposed another SIF retrieval method for the Global Ozone Monitoring Experiment-2 (GOME-2), which enabled a significant increase of the spatiotemporal resolution with respect to previous works. In principle, this proposed approach extends the SVD (or principal component analysis, PCA) technique applied exclusively on solar Fraunhofer lines (Guanter et al., 2012) to a broader wavelength range making also use of atmospheric absorption bands (water vapor, O_2 A).

Here, we present a new SIF retrieval method using a comparable methodology to that developed for ground-based instrumentations by Guanter et al. (2013) and for space-based measurements by Joiner et al. (2013). A feature of the method is the automated determination of the principal components (PCs) needed. The proposed retrieval method is applied to simulated data (Sect. 3.4) as well as to data from GOME-2 and SCIAMACHY (Sect. 3.5). Besides a time series of GOME-2 SIF covering the 2007–2011 time period, we are able to present a SIF data set from SCIAMACHY data for the August 2002–March 2012 time span. We compare those data sets with each other (Sect. 3.5.3) and with existing SIF data sets from GOME-2 (Joiner et al., 2013) and GOSAT data (Köhler et al., 2015a) (Sect. 3.5.4). In addition, we examine uncertainties (Sect. 3.5.5, 3.5.6) and assess the effect of clouds on the retrieval through the analysis of different cloud filter

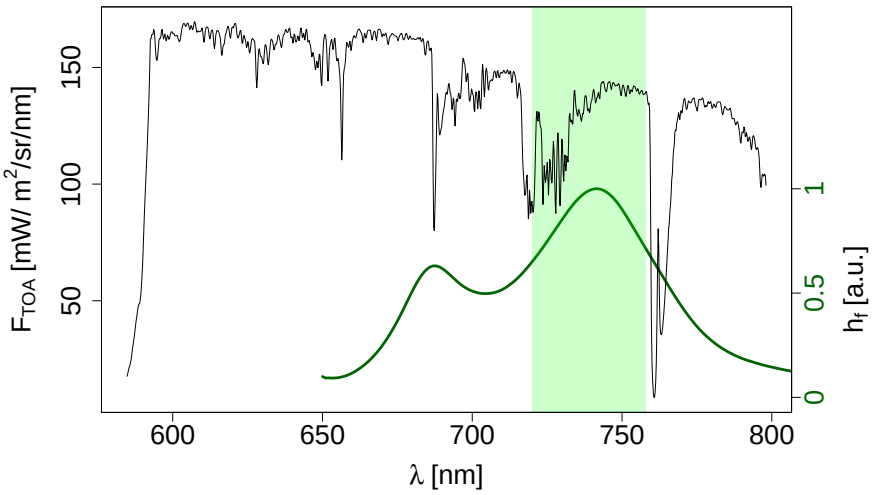


Figure 3.1: Sample GOME-2 spectrum in band 4. The spectral window that we use for SIF retrievals (720–758 nm) is marked in green. The reference fluorescence emission spectrum is depicted in green.

thresholds (Sect. 3.5.7).

3.2 Instruments

3.2.1 GOME-2

The Global Ozone Monitoring Experiment-2 (Munro et al., 2006) is a nadir-scanning medium-resolution UV/VIS spectrometer on board EUMETSAT's polar orbiting Meteorological Operational Satellites (MetOp-A and MetOp-B). MetOp-A (launched in October 2006) and MetOp-B (launched in September 2012) each carry 13 instruments performing operational measurements of atmosphere, land and sea surface. The polar orbit satellites reside in orbit at an altitude of approximately 820 km and have an equator crossing time near 09:30 local solar time. One revolution takes around 100 min. GOME-2 was designed to measure distributions of various chemical trace gases in the atmosphere. For this purpose, the spectral range between 240 and 790 nm is covered by four detector channels. The fourth channel (590–790 nm) encompasses the SIF wavelength region (Fig. 3.1). This channel has a spectral resolution of 0.5 nm and a signal-to-noise ratio (SNR) up to 2000. Here, we use a subchannel from GOME-2 on board MetOp-A, namely the spectral region between 720 and 758 nm covering 191 spectral points, to evaluate the SIF at 740 nm. The large default swath width of 1920 km with a footprint size of 80 km x 40 km enables a global coverage within 1.5 days. The GOME-2 level 1B product consists of radiance spectra and a daily solar irradiance measurement. Satellite data covering the 2007–2011 time period were used for this study.

3.2.2 SCIAMACHY

SCIAMACHY (Bovensmann et al., 1999) was 1 of 10 instruments on board European Space Agency's Environmental Satellite (ENVISAT). ENVISAT was launched in March 2002 and became non-operational in April 2012, when communications were lost. Similar to MetOp-A, the satellite has a sun-synchronous orbit at an altitude of approximately 800 km with an equator overpass time at 10:00 local solar time. Like GOME-2, the SCIAMACHY instrument was designed to measure distributions of various chemical trace gases in the atmosphere. Overall, a spectral range of 240–2400 nm is covered by eight detector channels, whereas a subset of the fourth channel (604–805 nm, spectral resolution of 0.48 nm, SNR up to 3000) is used to evaluate the amount of SIF at 740 nm.

SCIAMACHY measured alternately in nadir and limb mode, which leads to blockwise rather than continuous nadir measurements. Due to this default scan option, a global coverage is achieved within 6 days. The swath width of 960 km is only half as large as the swath width of GOME-2 in the nominal mode while the spatial resolution along track is 30 km and the nominal across track pixel size is 60 km. Due to data rate limitations, radiances in the spectral range of interest were downlinked at 240 km (four pixels binned) across track. Using the spectral interval between 720–758 nm to retrieve SIF leads therefore to a reduced spatial resolution by a factor of 2–3 compared to GOME-2.

As stated in Lichtenberg et al. (2006), two relevant corrections have to be applied in the processing chain of the level 1 data for the considered wavelength range: the memory effect correction and the dark signal correction. The first correction is probably more critical and a potential error source in the SIF retrieval because the memory effect modifies the absolute value of the signal and it involves the risk of artificial spectral features in the measurements. Both changes produce artifacts that are expected to influence SIF retrieval results since the level of SIF in comparison to the total signal at the sensor is low. Satellite data of SCIAMACHY for the August 2002–March 2012 time span have been used for this study.

3.3 Retrieval methodology

The main challenge when retrieving SIF from space-borne instruments is to isolate the SIF signal from the about 100 times more intense reflected solar radiation in the measured top-of-atmosphere (TOA) radiance spectrum. This section describes a strategy which is similar to the data-driven method proposed by Joiner et al. (2013).

3.3.1 Fundamental basis

Assuming a Lambertian reflecting surface in a plane-parallel atmosphere, the TOA radiance measured by a satellite sensor ($\boldsymbol{F}_{\mathrm{TOA}}$) can be described by

$$\boldsymbol{F}_{\mathrm{TOA}} = \frac{\boldsymbol{\rho}_{\mathrm{p}}}{\pi} \cdot \boldsymbol{I}_{\mathrm{sol}} \cdot \mu_0, \tag{3.1}$$

where $\boldsymbol{\rho}_\mathrm{p}$ is the planetary reflectance, $\boldsymbol{I}_\mathrm{sol}$ is the solar irradiance at the TOA and μ_0 is the cosine of the solar zenith angle (bold characters indicate variables with a spectral component). Verhoef and Bach (2003) split $\boldsymbol{\rho}_\mathrm{p}$ into contributions due to atmospheric path radiance, adjacency effect (path radiance from objects outside the field of view), reflected skylight by the target and reflected sunlight by the target. Assuming that a fluorescent target is observed, Eq. (3.1) can be modified as follows

$$\boldsymbol{F}_\mathrm{TOA} = \frac{\boldsymbol{\rho}_\mathrm{p}}{\pi} \cdot \boldsymbol{I}_\mathrm{sol} \cdot \mu_0 + F_\mathrm{s} \cdot \boldsymbol{h}_\mathrm{f} \cdot \boldsymbol{T}_\uparrow, \tag{3.2}$$

where F_s is the amount of sun-induced fluorescence at 740 nm (second peak of the emission spectrum), $\boldsymbol{h}_\mathrm{f}$ is a normalized reference fluorescence emission spectrum and $\boldsymbol{T}_\uparrow$ is the atmospheric transmittance in the upward direction.

As stated above, we use a statistical modeling approach similar to Joiner et al. (2013) in order to separate spectral features from planetary reflectance (including atmospheric absorption, atmospheric path radiance, surface reflectance) and SIF. The basic idea consists of modeling low and high frequency components of the planetary reflected radiance with a sufficient accuracy to be able to evaluate the changes in the fractional depth of solar Fraunhofer lines by SIF. It has to be emphasized that the main information for the retrieval is expected to originate from solar Fraunhofer lines, as simulations with a flat solar spectrum (Joiner et al., 2013) have shown that essentially no information about SIF can be retrieved. More specifically, it was not possible to disentangle atmospheric absorption from the SIF signal in the O_2 A-band with the statistical approach.

Here, $\boldsymbol{I}_\mathrm{sol}$ is known through satellite measurements (from GOME-2 and SCIAMACHY) or the external spectral library from Chance and Kurucz (2010), which has been used for the sensitivity analysis. Since $\boldsymbol{h}_\mathrm{f}$ is a prescribed spectral function, it is necessary to find appropriate estimates for $\boldsymbol{\rho}_\mathrm{p}$ and $\boldsymbol{T}_\uparrow$ to retrieve the amount of SIF at 740 nm (F_s). In order to model $\boldsymbol{\rho}_\mathrm{p}$ with a high accuracy, we use a combination of a third-order polynomial in wavelength to account for the spectrally smooth part and a set of atmospheric principal components (PCs) for the high frequency pattern of the spectrum. Thus, the preliminary forward model can be written as

$$\boldsymbol{F}_\mathrm{TOA} = \boldsymbol{I}_\mathrm{sol} \cdot \frac{\mu_0}{\pi} \cdot \sum_{i=0}^{3} (\alpha_i \cdot \boldsymbol{\lambda}^i) \cdot \sum_{j=1}^{n_\mathrm{PC}} (\beta_j \cdot \boldsymbol{PC}_j) + F_\mathrm{s} \cdot \boldsymbol{h}_\mathrm{f} \cdot \boldsymbol{T}_\uparrow, \tag{3.3}$$

where α_i, β_j, F_s and $\boldsymbol{T}_\uparrow$ are the unknowns which are necessary to generate a synthetic measurement. Note that Eq. (3.3) is equivalent to the forward model from Joiner et al. (2013) and Guanter et al. (2015). The difference lies in the derivation and interpretation, since so far no assumptions concerning effects of atmospheric scattering were made. The spectrally smooth contribution can mostly be attributed to surface reflectance and atmospheric scattering effects, which is henceforth referred to as apparent reflectance. The spectrally oscillating part is due to distinct atmospheric absorption features. Although the continuum absorption of the atmosphere is part of the apparent reflectance (low frequency component), we formally refer to the oscillating component of the spectrum as

effective atmospheric transmittance.

The statistically based approach assumes that spectra without SIF emission can reproduce the variance of the planetary reflectance also for spectra which contain SIF. For this purpose, the data are divided into a training set (measurements over non-fluorescent targets) and a test set (basically all measurements over land). According to Eq. (3.3), spectra of the training set need to be disaggregated into low and high frequency components in order to generate atmospheric PCs as described below.

3.3.2 Preparation of training set and generation of atmospheric principal components

The training set is exclusively composed of measurements over non-fluorescent targets and we assume that Eq. (3.2) is valid. As in Joiner et al. (2013), we use a principal component analysis (PCA) to convert these SIF-free spectra into a set of linearly uncorrelated components (PCs). Only a subset of all PCs (the same number as spectral points in the fitting window) will be needed to reconstruct each SIF-free spectrum with an appropriate precision.

First of all, $\boldsymbol{\rho}_{\mathrm{p}}$ is obtained through the normalization of the spectrum by the solar irradiance ($\boldsymbol{I}_{\mathrm{sol}}$) (which is also measured by GOME-2/SCIAMACHY), the cosine of the solar zenith angle (μ_0) and π. Figure 3.2 depicts such a normalized sample GOME-2 measurement in the 720–758 nm fitting window. The next step requires an estimate of the apparent reflectance (or the spectrally smooth component). Therefore, we assume two windows mostly devoid of atmospheric absorption within the retrieval window, namely the spectral regions between 721.5–722.5 and 743–758 nm, in order to estimate the ap-

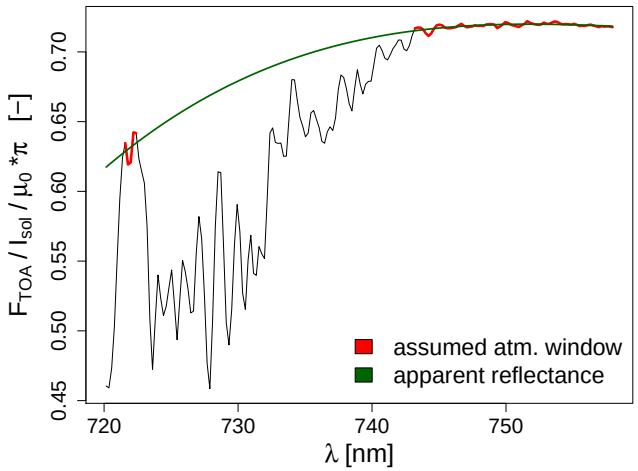


Figure 3.2: Estimation of apparent reflectance (green) from a sample GOME-2 measurement in the 720–758 nm fitting window by a third-order polynomial using atmospheric windows (red).

parent reflectance by a third-order polynomial. This is also shown in Fig. 3.2. In the next step, the apparent reflectance estimate is used to normalize $\boldsymbol{\rho}_\mathrm{p}$ in order to separate the effective atmospheric transmittance (high frequency component). If all spectra from the training set are processed in this way, a PCA is applied in order to obtain the atmospheric PCs. It is necessary to capture a sufficient number of atmospheric states within the training set so that the effective atmospheric transmittance can also be adequately constructed for the test data. Effects of correlation between the PCs and SIF will be investigated in detail below (Sect. 3.4.3).

An important advantage in using this method is that there is no need to perform explicit radiative transfer calculations to characterize atmospheric parameters, which affect the measurement (e.g., temperature and water vapor profiles).

3.3.3 Estimation of the ground to sensor transmittance

According to the preliminary forward model in Eq. (3.3), an estimate of $\boldsymbol{T}_\uparrow$ is the last necessary step to be able to retrieve SIF.

As it was shown by Joiner et al. (2013), $\boldsymbol{T}_\uparrow$ can be approximated in terms of the total atmospheric transmittance in downward and upward direction ($\boldsymbol{T}_{\downarrow\uparrow}$). Hence, $\boldsymbol{T}_{\downarrow\uparrow}$ and $\boldsymbol{T}_\uparrow$ are simultaneously computed from the state vector elements in the forward model from Joiner et al. (2013). Here, we estimate an effective ground to sensor transmittance ($\boldsymbol{T}^\mathrm{e}_\uparrow$) prior to solving the forward model in order to obtain an additional model parameter. The equation relating $\boldsymbol{T}^\mathrm{e}_\uparrow$ to $\boldsymbol{T}^\mathrm{e}_{\downarrow\uparrow}$ is given by

$$\boldsymbol{T}^\mathrm{e}_\uparrow = \exp\left[\ln(\boldsymbol{T}^\mathrm{e}_{\downarrow\uparrow}) \cdot \frac{\sec(\theta_\mathrm{v})}{\sec(\theta_\mathrm{v}) + \sec(\theta_0)}\right], \tag{3.4}$$

where θ_v is the viewing zenith angle and θ_0 is the solar zenith angle. Each measurement is normalized in the same way as for the training set to separate $\boldsymbol{T}^\mathrm{e}_{\downarrow\uparrow}$, that is in turn used to estimate $\boldsymbol{T}^\mathrm{e}_\uparrow$ through Eq. (3.4). Figure 3.3 depicts such an estimation for a sample GOME-2 measurement. However, a few implications arise through this approach, which are evaluated in the following.

Firstly, it is necessary to assume zero SIF, although these measurements potentially contain a SIF emission. Hence, an in-filling of atmospheric absorption lines might occur, which potentially affects the estimation of $\boldsymbol{T}^\mathrm{e}_\uparrow$ and in turn the retrieved SIF. By means of simulated TOA radiances (a detailed description can be found in Sect. 3.4), it becomes apparent that this effect is negligible. We have tested two scenarios without instrumental noise, a medium and a large difference in SIF emission (2.1 and 4.3 mW/m^2/sr/nm), and calculated $\boldsymbol{T}^\mathrm{e}_\uparrow$. Figure 3.4 depicts the resulting changes in $\boldsymbol{T}^\mathrm{e}_\uparrow$ due to the in-filling of atmospheric absorption lines, and as can be seen, differences are only marginal. Another consequence of the $\boldsymbol{T}^\mathrm{e}_\uparrow$ estimation procedure arises through the normalization of $\boldsymbol{T}^\mathrm{e}_\uparrow$ to one, although continuum absorption exists. This can lead to a bias in retrieved SIF values for inclined illumination/viewing angles and a high optical thickness (Guanter et al., 2015). It should also be noted, that the first assumed atmospheric micro-window (721.5–722.5 nm) might be influenced by water vapor absorption to a small extent, which

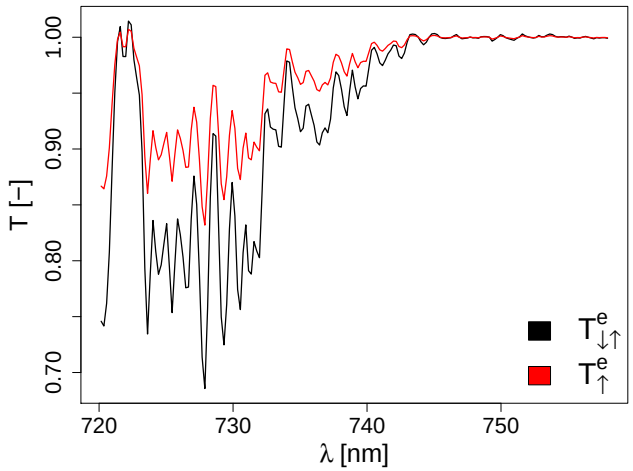


Figure 3.3: Estimation of the effective atmospheric transmittance in down- and upward direction (black) and in upward direction (red) using Eq. (3.4) from a sample GOME-2 measurement in the 720–758 nm fitting window.

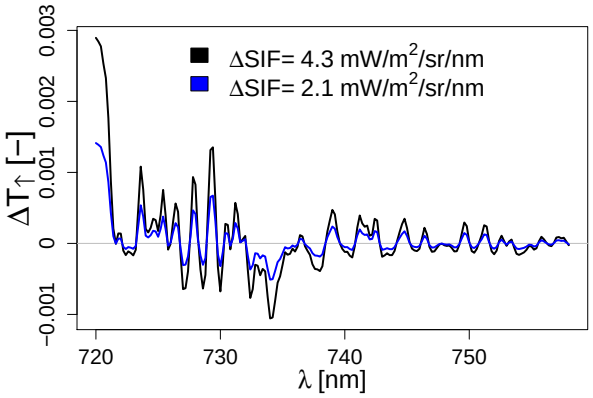


Figure 3.4: Changes in the estimated effective ground to sensor transmittance ($\boldsymbol{T}^{\mathrm{e}}_{\uparrow}$) for a medium (2.1 mW/m^2/sr/nm, blue) and a large (4.3 mW/m^2/sr/nm, black) SIF emission difference. Simulated TOA radiances (described in Sect. 3.4) without instrumental noise have been used for the calculation of $\boldsymbol{T}^{\mathrm{e}}_{\uparrow}$ in the 720–758 nm fitting window.

is due to the instrumental resolution of about 0.5 nm. In consequence, too high retrieval results may occur. However, the contribution of SIF in the lower wavelengths of the fitting window is decreased for two reasons: an increasing distance to the maximum of the prescribed spectral function at 740 nm and atmospheric absorption (except for the atmospheric micro-window). These theoretical considerations as well as our simulation results suggest that this effect is also negligible.

Our forward model is linearized as a result of the implementation from $\boldsymbol{T}_{\uparrow}^{\mathrm{e}}$ as a model parameter. That means the estimation problem could in principle be solved without iterations. However, it should be noted that the SIF estimation is inherently non-linear, and a linearisation is necessary, either at each step of an iterative algorithm (Joiner et al., 2013) or in advance.

3.3.4 Final Forward model

By using $\boldsymbol{I}_{\mathrm{sol}}, \boldsymbol{\lambda}, \boldsymbol{PC}$ and $\boldsymbol{T}_{\uparrow}^{\mathrm{e}}$ as model parameters and merging the high and low frequency components to model the planetary reflectance, the final forward model results in

$$\boldsymbol{F}_{\mathrm{TOA}} = \boldsymbol{I}_{\mathrm{sol}} \cdot \frac{\mu_0}{\pi} \cdot \sum_{i=0}^{3} \sum_{j=1}^{n_{\mathrm{PC}}} (\gamma_{i,j} \cdot \boldsymbol{\lambda}^i \cdot \boldsymbol{PC}_j) + F_{\mathrm{s}} \cdot \boldsymbol{h}_{\mathrm{f}} \cdot \boldsymbol{T}_{\uparrow}^{\mathrm{e}}, \tag{3.5}$$

where $\gamma_{i,j}$ and F_{s} are the state vector elements. Thus, there are $4 \cdot n_{\mathrm{PC}} + 1$ coefficients to be derived by an ordinary least squares fit. It might be noted that Eq. (3.5) is a simple algebraic transformation of Eq. (3.3).

A consequence of our approach is an increased number of state vector elements compared to Joiner et al. (2013), where the forward model is solved for $n_{\mathrm{i}} + n_{\mathrm{PC}} + 1$ elements. Note that a comparable forward model to our approach was proposed by Guanter et al. (2013), where the forward model is further simplified by convolving only the most significant PC with the low order polynomial in wavelength to avoid an overfitting of the measurement. Here, we omit an overfitting by applying a backward elimination algorithm to reduce the number of coefficients to fit. An important and unique consequence, which arises through this additional step, is an automated determination of the optimum number of PCs. The selection of the number of PCs has an effect on retrieval accuracy and precision as it is known from studies by Guanter et al. (2013) and Joiner et al. (2013). Hence, it is crucial to provide a solution for this issue, which is described in the following.

3.3.5 Backward elimination algorithm

The forward model in Eq. (3.5) contains $4 \cdot n_{\mathrm{PC}} + 1$ coefficients to fit, whereas not all of these coefficients are required. From a theoretical point of view, there is no reason to prefer an unnecessarily complex model rather than a simple one. To account for this fact, we use a backward elimination algorithm. Joiner et al. (2013) reported that only a few PCs (about 5, exact number depends on the fitting window) can explain already a very large amount of variance ($> 99.999\%$) of the normalized radiance. Using too many PCs

and thus coefficients to fit might therefore result in an overfitting of the measurement. That means an unnecessarily complex model loses its predictive performance by fitting noise. For this reason, we start with all candidate variables and remove each variable in order to test if any removal improves the model fit to the data. It should be noted that the coefficients of the first PC (carrying the most variance) and the targeted SIF are excepted from being removed. If any removal improves the fit, the removed variable, which improves the model fit the most, is abandoned in the state vector from the forward model in Eq. (3.5). This process is repeated until no further improvement occurs (for each single retrieval). The improvement is determined through the Bayesian information criterion (BIC, Wit et al., 2012), formally defined as

$$\boldsymbol{BIC} = -2l(\widehat{\Theta}) + p \cdot \log(n), \tag{3.6}$$

where $l(\widehat{\Theta})$ is the log-likelihood estimate of the model, p is the number of model parameters and n the number of spectral points. In simplified terms, it represents the goodness of fit weighted by the number of coefficients. The BIC is used as a statistical tool for model selection and balances the goodness of fit with the model complexity, where complexity refers to the number of parameters in the model. The BIC value itself is not interpretable but the model with the lowest BIC should be preferred because a lower BIC is associated with fewer model parameters and/or a better fit.

It should be mentioned that there are also other methods to compare and select models in order to avoid problems of overfitted measurements. Here, we decided to use the BIC because it penalizes the number of model parameters the most. It should further be noted that it would be necessary to test all possible combinations of model parameters to find the 'best' model, which is computationally too expensive. For this reason, a stepwise model selection (backward elimination) is performed.

Using the backward elimination algorithm has the consequence that the number of provided PCs is unimportant, as long as there are more PCs provided than actually necessary for an appropriate fit. In reverse it means that the optimum number of PCs is determined automatically. The detailed behavior of this supplementary step in comparison to a simple linear regression using all potential coefficients (all candidate variables) will be shown below in Sect. 3.4.3.

3.3.6 Uncertainty estimation

In order to assess the uncertainty of the SIF measurements, the 1σ retrieval error is calculated by propagating the measurement noise. We assume that measurement noise can be characterized by spectrally uncorrelated Gaussian noise. In this case, following Sanders and De Haan (2013), the signal-to-noise ratio (SNR) can be calculated for any radiance level if the SNR is known for a reference radiance level F_{ref} at a reference

wavelength λ_i^{ref} following

$$\mathrm{SNR}(\lambda_i) = \mathrm{SNR}_{\mathrm{ref}} \cdot \sqrt{\frac{F(\lambda_i)}{F_{\mathrm{ref}}}}. \tag{3.7}$$

The measurement noise scales with the square root of the signal level, which is an appropriate assumption for grating spectrometers such as GOME-2 and SCIAMACHY. Here, we determine $\mathrm{SNR}_{\mathrm{ref}}$ for each measurement (simulated and real) from the averaged radiance (F_{ref}) from 757.7–758 nm using calculations of the SNR vs. level 1B calibrated radiances (also averaged from 757.7–758 nm) under defined conditions (2% albedo, solar zenith angle of 0°, integration time of 1.5 s) performed by EUMETSAT. The error is then calculated through the evaluation of the retrieval error covariance matrix given by

$$\mathbf{S_e} = \left(\mathbf{K}^{\mathrm{T}}\mathbf{S_0}^{-1}\mathbf{K}\right)^{-1}, \tag{3.8}$$

where $\mathbf{K}$ is the Jacobian matrix formed by linear model parameters from Eq. (3.3) and $\mathbf{S_0}$ is the measurement error covariance matrix, which is a diagonal matrix (because of the assumption of spectrally uncorrelated noise) with the elements

$$\sigma_i^2 = \left(\frac{F(\lambda_i)}{\mathrm{SNR}(\lambda_i)}\right)^2. \tag{3.9}$$

In order to assess the restriction for the precision of spatiotemporal SIF composites due to instrumental noise only, the standard error of the weighted average can be calculated for each grid cell by

$$\sigma_{noise}(\overline{F_{\mathrm{s}}}) = \frac{1}{\sqrt{\sum_{i=1}^{n}(1/\sigma_i)^2}}, \tag{3.10}$$

where $\overline{F_{\mathrm{s}}}$ is the mean SIF value, and σ_i is the 1σ retrieval error given by Eq. (3.8).

Furthermore, the standard error of the mean from monthly mapped SIF has been computed for SCIAMACHY and GOME-2 as follows

$$\mathrm{SEM}_{\overline{F_{\mathrm{s}}}} = \frac{\sigma_{ret}}{\sqrt{n}}, \tag{3.11}$$

where σ_{ret} is the standard deviation and n is the number of observations per grid cell. In contrast to Eq. (3.10), this value is a measure of instrumental noise plus natural variability of SIF in the considered time range.

3.4 Sensitivity analysis

The retrieval approach has been tested for a wide range of conditions using simulated radiances in order to assess retrieval precision and accuracy, the effect of the backward elimination algorithm as well as the optimal retrieval window. This section describes

the underlying simulations briefly and examines retrieval properties and advantages with respect to a simple linear model without a backward elimination.

3.4.1 Simulated TOA radiances

As in Joiner et al. (2013), we use simulated sun-normalized TOA radiances from the Matrix Operator MOdel (MOMO) radiative transfer code (Fell and Fischer, 2001) with a spectral sampling of 0.005 nm. We take the widely used spectral library from Chance and Kurucz (2010) to simulate the solar irradiance. The simulations include two viewing zenith angles (0° and 16°), four solar zenith angles (15° , 30° , 45° and 70°), two atmospheric temperature profiles (middle latitude summer and winter), four surface pressures (955, 980, 1005 and 1030 hPa), four water vapor columns (0.5, 1.5, 2.5 and 4.0 g/cm^2), three aerosol layer heights (500–700, 600–800 and 700–900 hPa) using a continental aerosol model and five aerosol optical thicknesses at 550 nm (0.05, 0.12, 0.2, 0.3 and 0.4). Apart from the observation and illumination geometry, simulations for 480 different atmospheric states have been carried out in order to test the retrieval. In this case, the training set uses a spectral library of 10 different soil and snow surface reflectance spectra, which means that the training data contain 38 400 samples. A set of top-of-canopy (TOC) reflectances and SIF spectra derived with the FluorSAIL radiative transfer model has been utilized to produce the test data set. The surface reflectance is thereby a function of chlorophyll content and leaf area index (LAI), while SIF is a function of chlorophyll content (5, 10, 20 and 40 μg/cm^2), LAI (0.5, 1, 2, 3 and 4 m^2/m^2) and quantum efficiency (which affects the intensity of the SIF flux; 0.02, 0.05 and 0.08). It follows from these 60 diverse TOC fluorescence spectra and various simulations that the test data are composed out of 230 400 samples. We convolve the high spectral resolution TOA spectra (0.005 nm) to a lower resolution spectrometer grid with a 0.5 nm full-width at half-maximum and a spectral sampling interval of 0.2 nm which is similar to GOME-2 and SCIAMACHY. A realistic instrumental noise with respect to the calculated SNR from GOME-2 (in relation to the radiance level at a reference wavelength) provided by EUMETSAT is then added to the spectra using Eq. (3.7).

3.4.2 End-to-end simulation

Figure 3.5 depicts the result of the end-to-end simulation using the described retrieval approach with a fitting window ranging from 720–758 nm and providing eight PCs (derived from the synthetic training set). The mean and standard deviation of the 480 simulated atmospheric states as well as the illumination and observation angles were calculated for each of the 60 TOC SIF spectra. The good correlation between simulated and retrieved SIF in Fig. 3.5 suggests that the retrieval approach is basically appropriate for the separation of the SIF signal from the TOA radiance. However, to enhance confidence and to prove that several potential systematic effects can be excluded, we depict the retrieved minus simulated SIF in dependence on the simulated solar zenith angle, the water vapor column and aerosol optical thickness at 550 nm in Fig. 3.6. The only visible bias is with respect to the solar zenith angle. Low illumination angles cause a slightly higher variance,

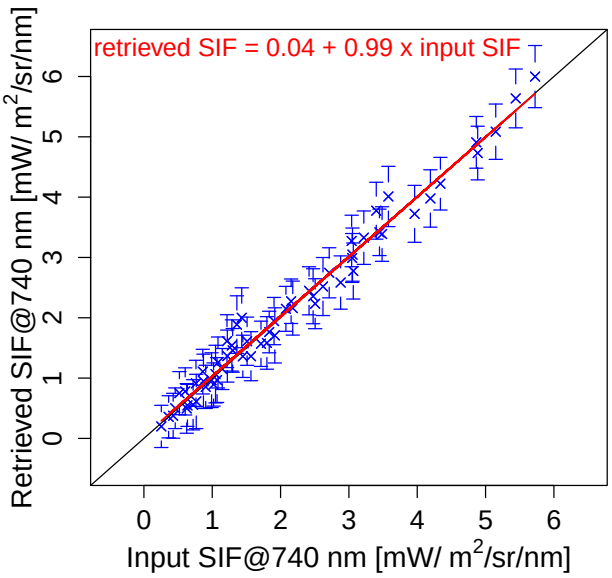


Figure 3.5: Input vs. retrieved SIF at 740 nm of the end-to-end simulation using the retrieval window from 720–758 nm and providing eight PCs. The average of the retrieved SIF for each top-of-canopy (TOC) measurement is shown together with its standard deviation (error bar). A linear fit of all TOC means is shown in red.

which can be expected since the noise level increases with a higher TOA radiance.

Fig. 3.7 illustrates the extent to which the theoretical random error, calculated through Eq. (3.8), matches the actual precision error. It can be seen that the root-mean-square error (denoted by crosses) is higher than the majority of our theoretical error estimates, but well within the statistical spread. This is due to the fact that the forward model is an approximation; therefore small systematic errors (as can be seen in Fig. 3.5) occur, which are not represented in the retrieval error covariance matrix $\mathbf{S_e}$. Furthermore, Fig. 3.7 reveals that the error (theoretical and actual) increases slightly with the signal level.

In view of the good correspondence between input and retrieved SIF in the end-to-end simulation, it can be stated that the retrieval method is appropriate. One limitation of this sensitivity study consists of the absence of clouds, which inevitably impact the retrieval of SIF. The result of other simulation studies by Frankenberg et al. (2012) and Guanter et al. (2015) suggests that an underestimation of SIF can be expected in the presence of clouds. The reason is a shielding effect, which is not captured by the forward model. Nevertheless, this should be of secondary importance when evaluating the pure retrieval methodology. Furthermore, we will assess the impact of clouds on the retrieval based on real satellite measurements in Sect. 3.5.7.

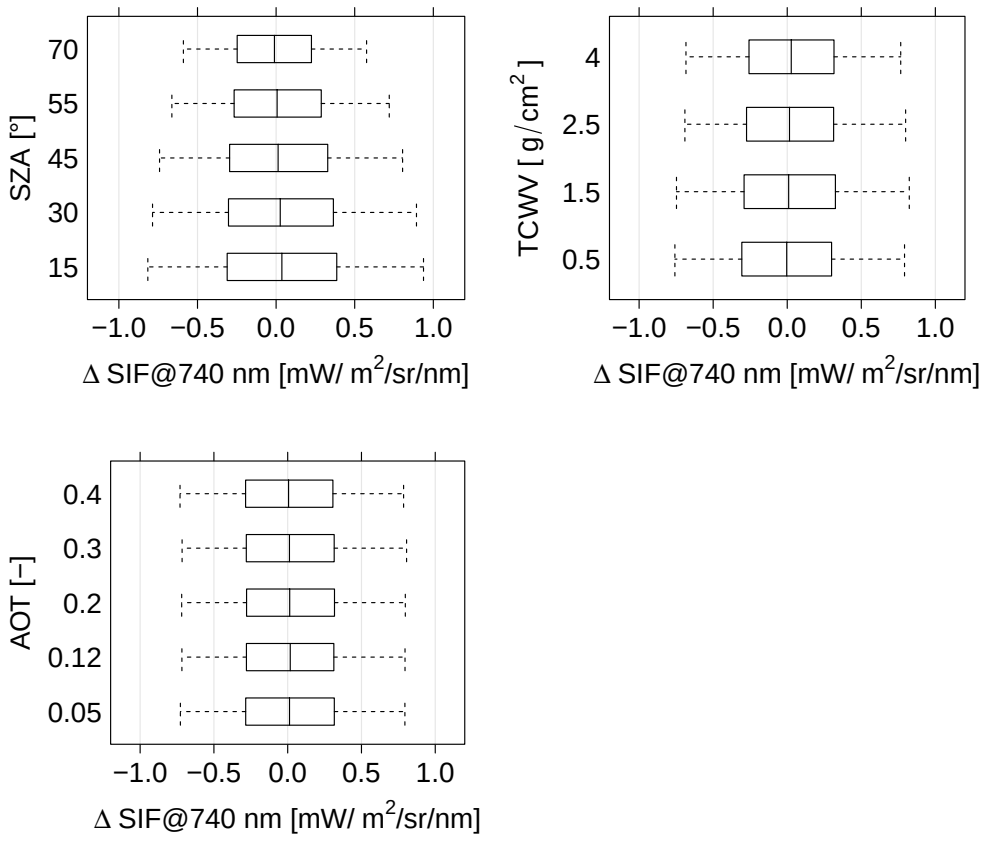


Figure 3.6: Retrieved minus simulated SIF in dependence on the simulated solar zenith angle (SZA), the water vapor column (TCWV) and aerosol optical thickness at 550 nm (AOT) for the same retrieval properties as in Fig. 3.5 (eight provided PCs, 720–758 nm retrieval window). The vertical bar represents the median, the box/error bar covers 50%/90% of differences.

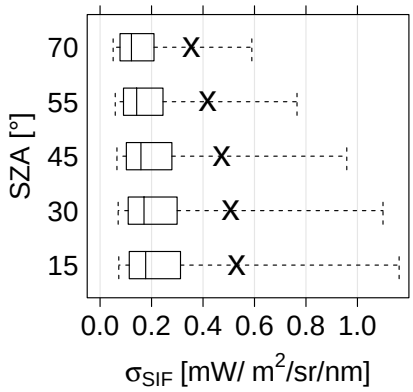


Figure 3.7: Estimated SIF error in dependence on the simulated solar zenith angle (SZA) for the same retrieval properties as in Fig. 3.5 (eight provided PCs, 720–758 nm retrieval window). The vertical bar represents the median, the box/error bar covers 50%/90% of the theoretical random errors as calculated from Eq. (3.8). The crosses denote the estimated root-mean-square errors of the simulation-based retrieval test.

3.4.3 Influence of number of PCs used and backward elimination

We performed the retrieval several times using 5–25 PCs in order to assess the sensitivity to the number of provided PCs. In addition, the backward elimination algorithm was disabled (all $4 \cdot n_{\mathrm{PC}} + 1$ coefficients from Eq. (3.5) are used) to evaluate whether the algorithm is capable of reducing the noise and avoid an overfitting. Results are presented in Fig. 3.8 in form of various statistical comparisons which are evaluated and described below.

The first plot in Fig. 3.8 shows the average number of selected atmospheric PCs and coefficients as a function of provided PCs. This plot is insufficient to decide how many atmospheric PCs should be provided to the algorithm, but it reveals that a saturation occurs at about seven selected PCs and 15 coefficients (combination of a third-order polynomial in wavelength with individual PCs), respectively. In the following, it should be taken into account that the linear model fit contains all $4 \cdot n_{\mathrm{PC}} + 1$ coefficients, while the backward elimination algorithm selects a maximum number of nine PCs/21 coefficients with a probability of not being exceeded in 90% of cases.

The bias which was calculated through the mean difference of retrieved minus input SIF represents the accuracy of the retrieval. It can be seen that the bias drops down to a value close to zero when providing more than seven PCs for both the linear model fit and the backward elimination fit. Using all coefficients leads to a slightly increasing positive bias associated with a larger number of PCs, which is not the case when the backward elimination is applied. Consequently, an unnecessary complex model would potentially result in overestimated SIF values.

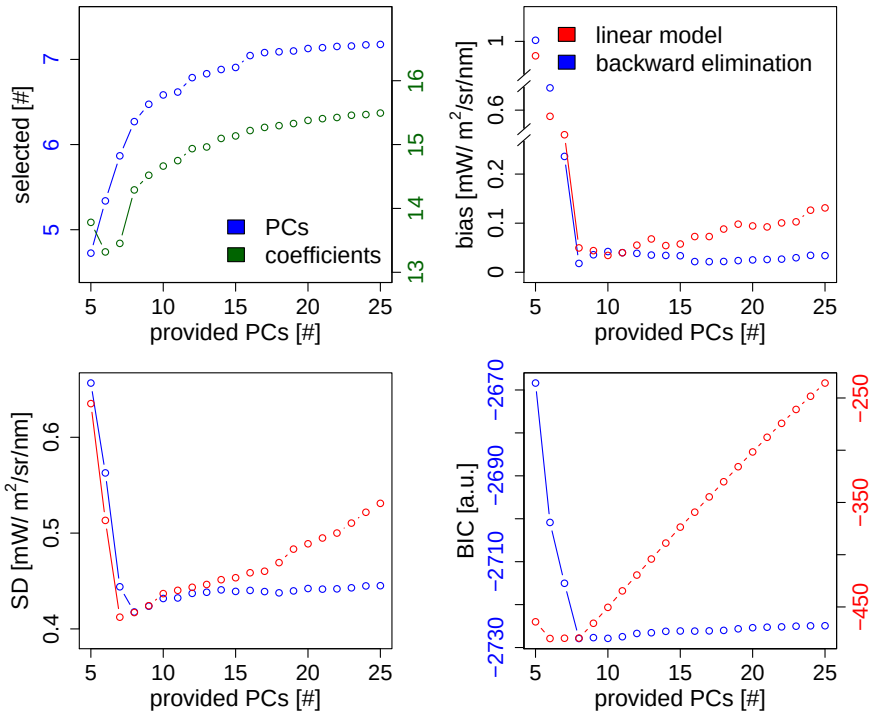


Figure 3.8: The plot on the upper left shows the average number of selected PCs (blue) and coefficients (green) as a function of PCs provided to the backward elimination algorithm. In the following, the retrieval results for the linear model fit using all potentials coefficients (red) are compared to the backward elimination fit (blue). Depicted are the bias (mean difference of retrieved minus input SIF), the average standard deviation of retrieved SIF for the 60 TOC SIF spectra (SD) and the mean Bayesian information criterion (BIC).

A difference between the disabled and enabled backward elimination algorithm can also be seen in the comparison of the mean standard deviation. Here, the resulting standard deviations of the retrieved SIF under differing atmospheric conditions for the 60 TOC SIF spectra were averaged to assess the retrieval precision. It can be noted that the precision of the backward elimination fit remains constant, while the linear model fit loses precision with a larger number of PCs. Also this comparison suggests providing at least seven PCs to the retrieval.

More pronounced differences arise in the comparison between the mean Bayesian information criterion (BIC) values. This fact is expected since the number of coefficients serves as weight for this criterion (as described in Sect. 3.3.5). For an interpretation of the BIC it has to be noted that the model with the lowest value is to be preferred. As a consequence, the BIC can be used to determine the appropriate number of PCs for the retrieval. It turns out that six to eight PCs should be used for the linear model fit using all coefficients, while at least eight PCs should be provided for the backward elimination

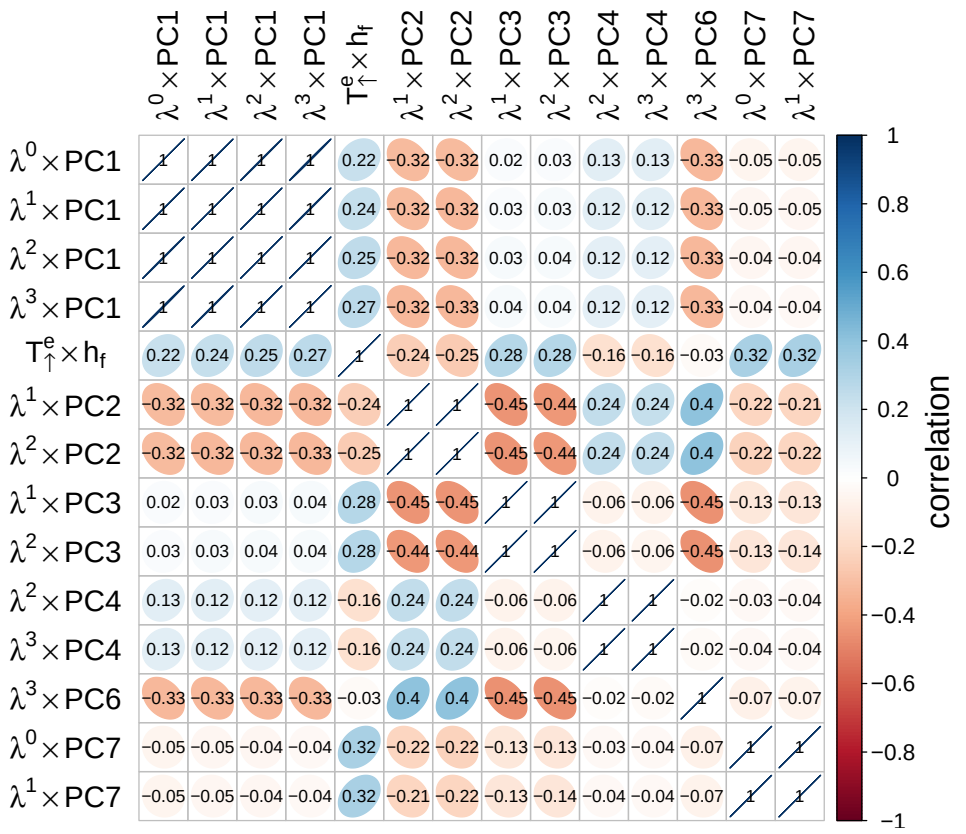


Figure 3.9: Correlation error matrix of a sample retrieval. A total of 10 PCs were supplied, atmospheric conditions were set to a middle latitude summer temperature profile, 955 hPa surface pressure, 700–500 hPa aerosol layer height, 0.05 aerosol optical thickness, 0.5 g/cm^2 water vapor column. The retrieved/input SIF at 740 nm amounts to 1.72/1.79 mW/m^2/sr/nm.

fit. This finding concurs with the results from the previous comparisons and supports the applicability of this criterion. It should be mentioned that the order of magnitude of BIC values clearly shows that the backward elimination fit should be preferred in any case. A potential overfitting of the measurement can be successfully prevented in this way.

Theoretically, there should be no more variability in data points for large numbers of initial PCs if the model parameter selection is enabled. Although Fig. 3.8 basically reflects this expectation, there is still a low variability. The reason can be found in the stepwise model selection, which is performed because a test of all possible model parameter combinations would be computationally too expensive (Sect. 3.3.5).

Even if differences are small, it can be concluded that the accuracy (no significant bias) and precision (decreased average standard deviation) are enhanced when the back-

ward elimination is enabled. Hence the noise is reduced by selecting only appropriate coefficients which is expected to be of particular importance for real satellite measurements. Unfortunately, there is no ground truth against which the retrieval could be adjusted or validated for real satellite measurements. As a consequence, it is not possible to determine the most appropriate number of PCs for real satellite data. Thus, it is advantageous that the backward elimination algorithm ensures stable results, regardless how many PCs are provided (with the restriction that there is a minimum number of required PCs). Furthermore, an overfitting of the measurement is avoided by using the discussed algorithm.

As a conclusion of these findings, we decided to provide initially 10 PCs for the retrieval applied to real GOME-2 and SCIAMACHY data. The backward elimination algorithm selects the required model parameters (a subset of candidate parameters from Eq. (3.5)) based on the BIC for each pixel automatically. In order to prove this assumption and to show that there are no significant correlations of the SIF spectrum to any of the selected model parameters, we supplied 10 PCs (as for the retrieval with satellite data) and calculated the correlation error matrix following Govaerts et al. (2010) for a sample fit (Figure 3.9). As can be seen, correlations of selected model parameters with $\boldsymbol{T}_{\uparrow}^{\mathrm{e}} \cdot \boldsymbol{h}_{\mathrm{f}}$ are not perfectly zero but absolute values do not exceed 0.3. Higher correlations between other model parameters occur, especially for those containing the same PC. However, this is unimportant since it is not intended to interpret the concerned coefficients. These model parameters are solely necessary to obtain an appropriate fit to the measurement. Nevertheless, it should be noted that higher correlations between $\boldsymbol{T}_{\uparrow}^{\mathrm{e}} \cdot \boldsymbol{h}_{\mathrm{f}}$ and other model parameters cannot per se be excluded because the algorithm selects the model parameters for each single measurement independently of correlations. A further test of the correlation of the PCs obtained by SCIAMACHY and the assumed SIF spectral shape ($\boldsymbol{h}_{\mathrm{f}}$) suggests that there are only low correlations (mean correlation coefficient is 0.2).

3.4.4 Selection of the retrieval window

In order to justify the 720–758 nm retrieval window, we performed retrievals in additional fitting windows (710–758, 715–758, 725–758, 730–758, 735–758 nm). Table 3.1 provides the linear fit results for supplying the most appropriate number of PCs to the retrieval algorithm in each tested fitting window. This number has been determined with the help of a statistical comparison as it is shown in Fig. 3.8. We found that both confined and extended retrieval windows lead to a slight bias. Furthermore, the extended retrieval windows require a larger number of PCs. The selected retrieval window from 720–758 nm is therefore reasonable in terms of retrieval accuracy and computation time. It might be worth noting that the described retrieval approach suggests using significantly less PCs as opposed to Joiner et al. (2014), where the latest algorithm (V25) uses 12 PCs in a retrieval window ranging from 734–758 nm. In the following, further theoretical aspects to select the retrieval window are discussed.

It is known that vegetation has a unique, spectrally smooth reflectance signature in the considered wavelength ranges, which means that there are no distinct absorption lines.

Table 3.1: Comparison of different retrieval windows. Shown is the linear fit (retrieved SIF = intercept + slope · input SIF) for supplying the most appropriate number of PCs (determined as in Fig. 3.8).

Exp.	**Fitting window**	**# of PCs**	**linear fit**
1	735–758	5	$y = 0.09 + 0.91 \cdot x$
2	730–758	6	$y = 0.10 + 0.89 \cdot x$
3	725–758	7	$y = 0.05 + 0.92 \cdot x$
4	720–758	8	$y = 0.04 + 0.99 \cdot x$
5	715–758	10	$y = 0.04 + 0.97 \cdot x$
6	710–758	11	$y = 0.26 + 0.86 \cdot x$

Nevertheless, the spectral reflectance of vegetated areas changes rapidly in the red edge region (680–730 nm), which is due to the spectral variation in chlorophyll absorption. Using the narrow atmospheric window at 722 nm assures that the apparent reflectance (which is in turn used to estimate $\boldsymbol{T}^{\mathrm{e}}_{\uparrow}$) can be modeled with sufficient precision for the part of the retrieval window which is affected by the red edge. Extending the retrieval window to lower wavelengths would lead to errors in the reflectance estimation caused by a lack of further atmospheric windows (the next spectral region with a high atmospheric transmittance is located at wavelengths below 710 nm). This error translates into the transmittance estimation, which propagates to errors in SIF. Extending the retrieval window to the O_2 A-band at 760 nm would be possible and has also been shown by Joiner et al. (2013), but a benefit cannot be expected. As already stated by Frankenberg et al. (2011a), the separation from SIF and atmospheric scattering properties is ambiguous when using only O_2 absorption lines. Also Joiner et al. (2013) reported that the removal of the O_2 A-band is not accompanied with a significant loss of information content. On the other hand, it is possible to confine the retrieval window, whereby a loss of information content regarding SIF is expected, which leads to a loss of retrieval precision and accuracy as can be seen in Table 3.1.

Advantages of the 720–758 nm retrieval window can be summarized as follows:

1. It covers the second peak of SIF emission at 740 nm.

2. It contains spectral regions with a high atmospheric transmittance (between 721.5–722.5 and 743–758 nm), which is necessary to characterize the apparent reflectance and in turn $\boldsymbol{T}^{\mathrm{e}}_{\uparrow}$.

3.5 Results

The presented SIF retrieval method has been used to produce a global SIF data set from GOME-2 data covering the 2007–2011 time period. In addition, the SIF retrieval has been implemented for SCIAMACHY data for the August 2002–March 2012 time span. This section describes the application of the algorithm to the satellite data, results

from spatiotemporal composites as well as a comparison to the results from Joiner et al. (2014). In addition, we compare the GOME-2/SCIAMACHY SIF results to a SIF data set derived from GOSAT data by Köhler et al. (2015a). We also discuss an important limitation, which arises due to the South Atlantic Anomaly (SAA). Furthermore, the impact of clouds on the retrieval will be assessed.

3.5.1 Application to GOME-2 and SCIAMACHY data

In contrast to the synthetic data set, where training and test sets are clearly separated, the real satellite data are first partitioned as described below.

The selection of the training set is relevant to obtain meaningful results because these data are used to model the planetary reflectance of the desired measurements over land (test set). It is therefore essential to capture as many atmospheric states and non-fluorescent surfaces as possible within the training set to ensure its representativeness. Random samples of measurements over areas where no SIF signal is expected (e.g., deserts, ice and sea regardless of the degree of cloudiness) are used for this purpose. The selection of such measurements is based on the determination of the land cover using the International Geosphere and Biosphere Programme (IGBP) classification (Friedl et al., 2002) derived from MODIS data. Care is taken to ensure that only homogeneous, non-vegetated land cover classes or measurements over sea and ice serve as a basis for the training set. Furthermore, it has to be noted that the training set is sampled on a daily basis for GOME-2 and a 3-day basis for SCIAMACHY data. The different period of time to sample the training data is due to the reduced number of soundings from SCIAMACHY. Based on the sampled training data, the atmospheric PCs are calculated as explained earlier. Instrumental artifacts and degradation are expected to be captured by this method and do not need to be considered.

In general, the test set is composed of all available land pixels, but cloud-contaminated measurements might make it problematic to retrieve SIF since it can be expected that the SIF signal is partly shielded in the presence of clouds, which potentially biases the retrieval. For this reason, the range of cloud fractions is limited to 0.5, which saves also computation time, whereas the restriction is explicitly not applied to the training set. We use the effective cloud fraction from the Fast Retrieval Scheme for Clouds from the O_2 A-band (FRESCO, Wang et al., 2008). In the case of GOME-2, the provided cloud fraction is already attached to the level 1B satellite data. The FRESCO cloud fraction data for SCIAMACHY were available separately and had to be collocated.

The FRESCO cloud fraction is not available for the presence of snow and ice, which is of particular relevance in the winter time at higher latitudes. In order to obtain a complete time series of SIF in affected regions, we evaluate also measurements with an unknown cloud fraction in the presence of snow. Hence, measurements over snow are determined using the ERA-Interim re-analysis data (Dee et al., 2011) on a 0.75° grid. The general cloud fraction threshold of 0.5 will be further discussed in Sect. 3.5.7.

Following the results of our sensitivity analysis, at least eight PCs should be provided for the SIF retrieval in the 720–758 nm spectral window. Furthermore, it has been evaluated that the number of provided PCs is not affecting retrieval results as long as

we provide more PCs than actually necessary. In contrast to the simulations, some PCs may also account for instrumental effects when applying the algorithm to real data. For this reason, we provided initially 10 PCs for the retrieval using real satellite data. It has emerged that the algorithm selects on average 6 PCs and decides in about 3% of cases that 10 PCs are actually needed when GOME-2 data are used. This finding corresponds to our sensitivity analysis, where on average seven PCs are selected. However, this changes when the retrieval is applied to SCIAMACHY data. In this case, the algorithm decides in about 40% of cases that 10 PCs are needed. Therefore, we increased the number of provided PCs to 20, whereas on average 14 PCs are selected. The probability that all 20 PCs are selected is below 1%. It is unclear why SCIAMACHY requires twice the number of PCs as GOME-2.

As a consequence of these findings, we provided 10 PCs to the retrieval for GOME-2 and 20 PCs when SCIAMACHY data are used. The processed data can be retrieved from ftp://ftp.gfz-potsdam.de/home/mefe/GlobFluo/.

3.5.2 Quality control

Since it cannot be excluded that the retrieval fails for single measurements, the retrieval results have to be checked. This is done by using the residual sum of squares (RSS) from each retrieval. The resulting coefficients are used to generate a synthetic measurement which is compared with the original measurement. The RSS value is then the discrepancy between the data and the model. One major issue, which causes high residuals, is the South Atlantic Anomaly (SAA), discussed below in Sect. 3.5.5. In general, it appears that the RSS value is around 0.5 $(\mathrm{mW/m^2/sr/nm})^2$. Single retrievals are removed if the RSS is above 2 $(\mathrm{mW/m^2/sr/nm})^2$ for both GOME-2 and SCIAMACHY. It turned out, that monthly composites from GOME-2 SIF retrieval results contain striping effects from single swaths in individual cases. This issue might be caused by special orbits (e.g., narrow swath) and is solved by removing distinct swaths with a high average of RSS values (above 1 $(\mathrm{mW/m^2/sr/nm})^2$). This striping effect was not observed for SIF results obtained from SCIAMACHY. Further filtering besides the residual check is not applied to our data set.

3.5.3 Spatiotemporal composites

In principle, it is possible to achieve a global coverage of SIF measurements within 1.5 days for GOME-2 and 6 days for SCIAMACHY, but the presence of clouds prevents such a high temporal resolution. Although it is important to consider different time scales, it is common to produce monthly means. Two monthly composites of SIF in January and July 2011 derived from SCIAMACHY and GOME-2 data are shown together with V25 GOME-2 SIF results provided by Joiner et al. (2014) in Fig. 3.10.

Overall all three results compare very well concerning spatial patterns, although the SIF composite derived with SCIAMACHY is provided in a spatial resolution of $1.5° \times 1.5°$, whereas SIF retrievals from GOME-2 are rastered in $0.5° \times 0.5°$ grid boxes. The lower spatial resolution for SCIAMACHY is due to the fact that the original pixels are co-

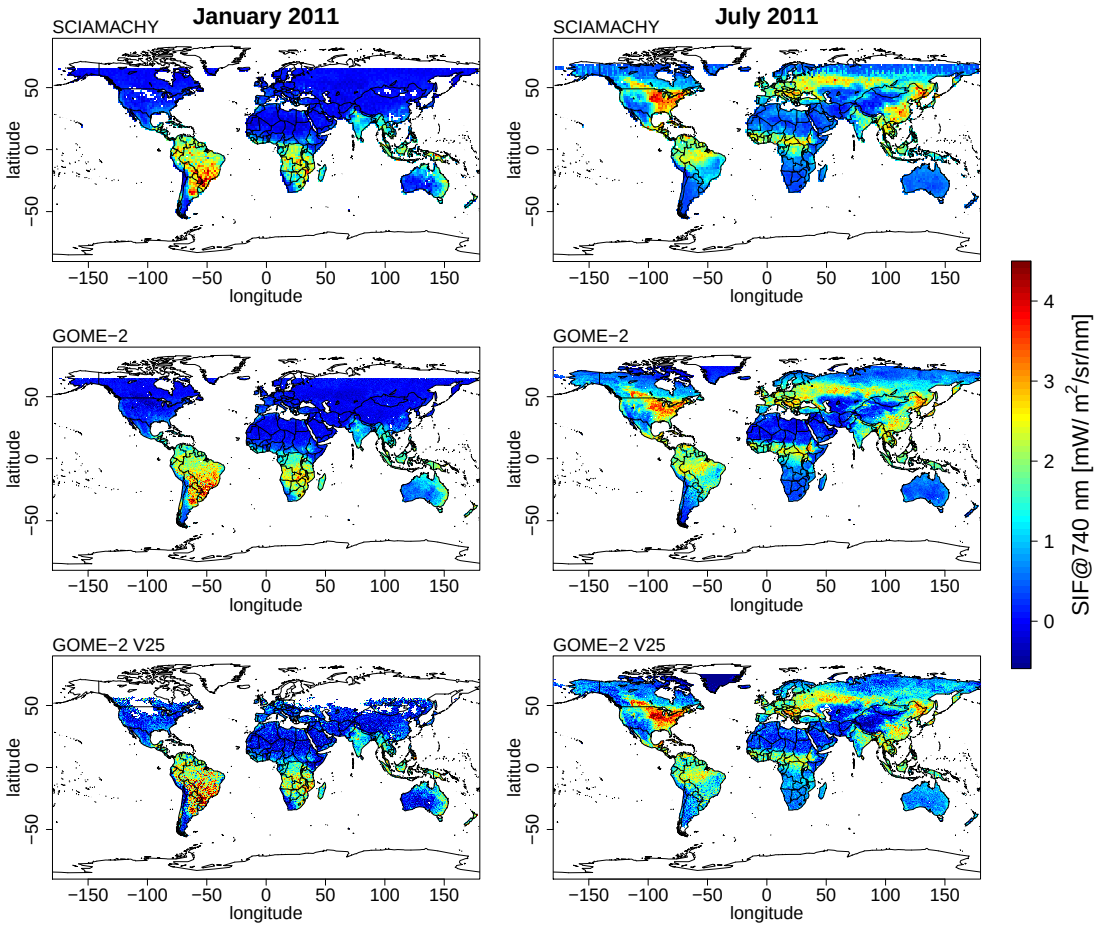


Figure 3.10: Monthly composites of SIF at 740 nm for January (left column) and July (right column) 2011. The upper row depicts SIF results using SCIAMACHY data and the rows below show SIF composites derived from GOME-2 data using our algorithm (middle row) and SIF results provided by Joiner et al. (2014) (bottom row). The SCIAMACHY composites and results from Joiner et al. (2014) are scaled by the relationships in Table 3.2 to ensure a good visual comparison.

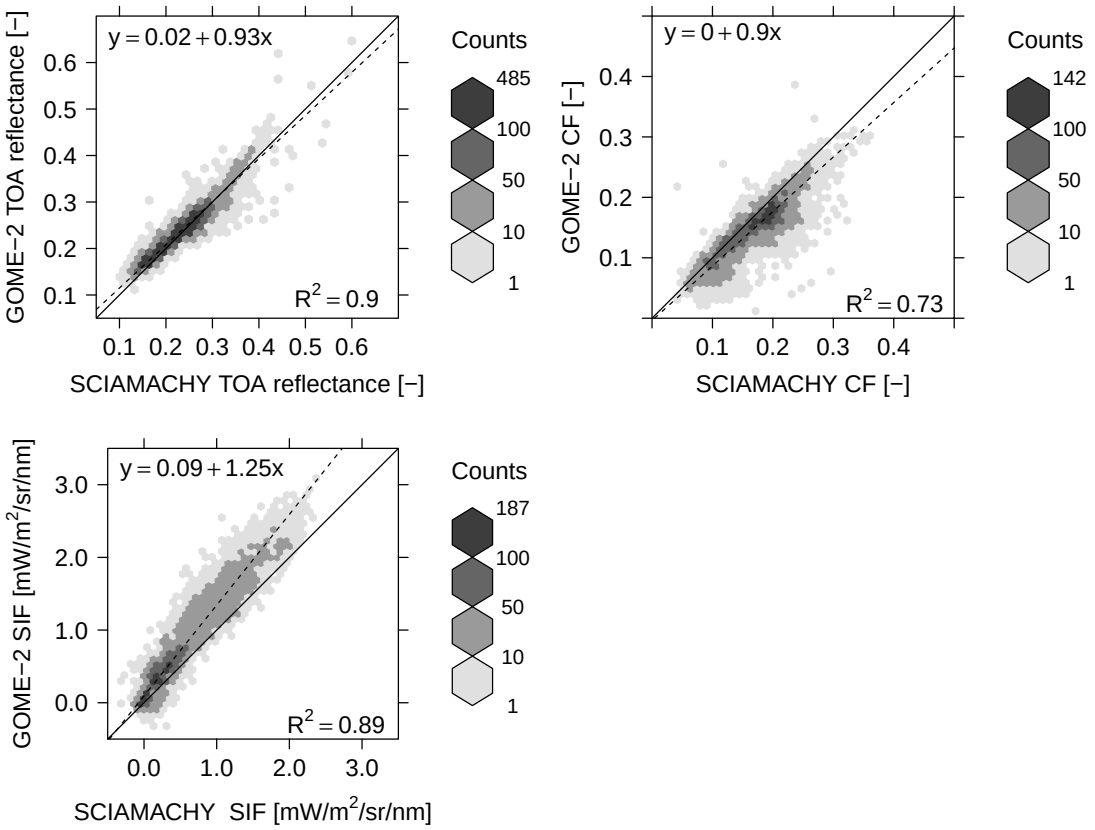


Figure 3.11: GOME-2 versus SCIAMACHY data of TOA reflectance at 720 nm, FRESCO (Wang et al., 2008) cloud fractions (CF) and SIF. Averages from February 2007 to December 2011 (rastered on $1.5° \times 1.5°$ grid) with a cloud fraction lower than 0.5 have been used. The linear fits are depicted as dotted lines.

added in certain wavelength regions in order to meet downlink limitations. However, it must be stated that absolute SIF values obtained from SCIAMACHY are slightly lower than from GOME-2, although such a difference is not expected since the overpass time differs only half an hour.

One reason might be a higher cloud contamination of the bigger footprints from SCIAMACHY. As already stated in Sect. 3.4.2, simulation studies by Frankenberg et al. (2012) and Guanter et al. (2015) have shown that an underestimation of SIF caused by a extinction through clouds can be expected. Another possible reason could be a relative radiometric offset between the two satellite sensors. Therefore, we assess the TOA reflectance at 720 nm (lower limit of the fitting window), the FRESCO cloud fraction and the retrieved SIF for the overlapping data sets in time from GOME-2 and SCIAMACHY in Fig. 3.11. As can be seen from the TOA reflectance comparison, there is most likely

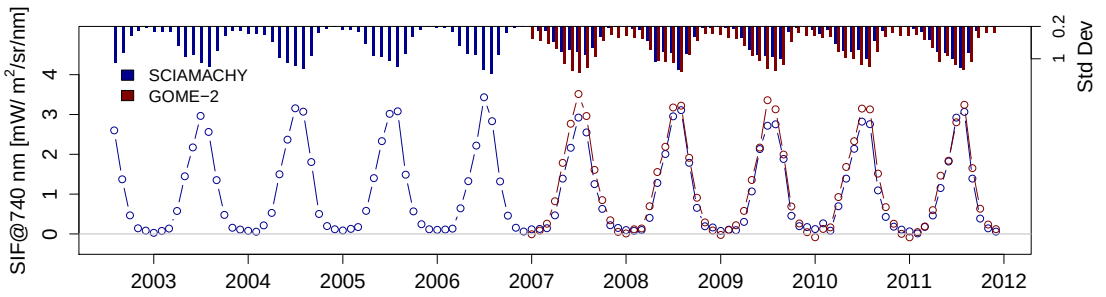


Figure 3.12: Comparison of monthly SIF averages derived from SCIAMACHY (blue) and GOME-2 (red) for croplands between 100° W–80° W, 35° N–50° N (shown in Fig. 3.16). The standard deviation per time step is depicted as a bar plot.

no offset. When considering the cloud fractions, the overall linear fit is comparable to that of the TOA reflectance but points scatter more widely ($R^2 = 0.73$) with a visible tendency to higher cloud fractions of the SCIAMACHY data. In contrast, the scattering of retrieved SIF values is not widespread ($R^2 = 0.89$) but GOME-2 SIF values are higher by a factor of 1.25. In view of these results, it is likely that lower SCIAMACHY SIF values are at least partly associated with a higher cloud contamination. Nevertheless, it needs to be mentioned that the relationship between SIF values from GOME-2 and SCIAMACHY is seasonally dependent. In order to ensure a good visual comparison, derived SIF values from SCIAMACHY data are therefore scaled in Fig. 3.10 by the relationship, which is obtained from monthly means (see Table 3.2).

The January 2011 SIF composites in Fig. 3.10 comprise a high SIF average in South America which must be treated with caution, because this region is particularly affected by the South Atlantic Anomaly. This limitation will be discussed in Sect. 3.5.5 but it should already be considered at this point to avoid a misinterpretation. Concurrently, SIF values in the Northern Hemisphere are close to 0, which is owing to the inactivity of vegetation during this month. The appearance of the map changes fundamentally for July. It is interesting to note the high SIF average in the US Corn Belt region with respect to other regions at the same latitude, which is produced by large crop- and grassland areas. This region with pronounced SIF averages was examined in detail by Guanter et al. (2014) with the finding that SIF retrievals provide a basis for a reliable estimation of gross primary production (GPP) in cropland and grassland ecosystems. Furthermore, it has been stated that current carbon cycle models on average underestimate the GPP in such regions.

Figure 3.12 depicts a time series of monthly averages over a sample cropland area between 100° W–80° W, 35° N–50° N (shown in Fig. 3.16) in order to investigate the temporal consistency of our SIF retrieval algorithm. Therefore, the retrieval output (without averaging the results in grid boxes) in that area has been collocated with the IGBP classification (Friedl et al., 2002). Those retrievals which were classified as cropland with cloud fractions lower than 0.5 or unknown cloud fractions in the presence of snow,

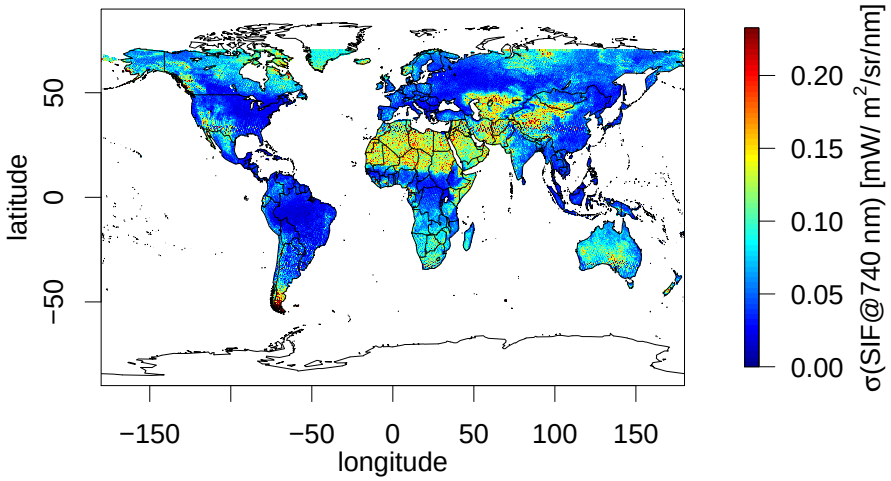


Figure 3.13: Standard error of the weighted average (using Eq. (3.10), measure of instrumental noise only) of GOME-2 SIF for July 2011.

serve as the basis for the monthly SIF averages in Fig. 3.12. Both time series show an aligned annual cycle with high SIF values during summertime. Beside the difference in absolute values (no scaling has been employed) it is noticeable that the interannual variability for the two time series may be slightly inconsistent, but the difference may also be due to noise or the different sampling. Considering that SIF values are close to zero during inactive periods, the standard deviation in wintertime may be used to estimate the precision for the SIF retrieval (specific for the considered area and monthly averages in Fig. 3.12), which is of the order of 0.2–0.4 mW/m^2/sr/nm. It seems reasonable that the GOME-2 SIF data exhibit a generally higher variability because of the fact that each SCIAMACHY pixel is similar to four averaged GOME-2 pixels.

In order to provide an typical estimate of the uncertainty of monthly SIF composites due to instrumental noise, Fig. 3.13 shows the standard error of the weighted average (Eq. (3.10)) derived with GOME-2 for July 2011. Highest uncertainties occur over bright areas associated with a higher photon noise, e.g., deserts and regions with snow/ice. The resulting pattern can therefore be interpreted as an error increasing with TOA radiance.

3.5.4 Comparison with existing retrieval results

Since we present a similar SIF retrieval method to that proposed by Joiner et al. (2013), it is reasonable to compare both results. It should be noted that there are two released versions of GOME-2 SIF (V14 & V25), whereas this comparison focusses on the latest version (V25) described in Joiner et al. (2014). Main changes from V14 to V25 lie in a confined retrieval window (V14: 715–758 nm, V25: 734–758 nm) and subsequently in a reduced number of PCs used (V14: 25 PCs, V25: 12 PCs). Furthermore, the spectra are normalized with respect to earth spectra instead of the GOME-2 solar spectra in V25.

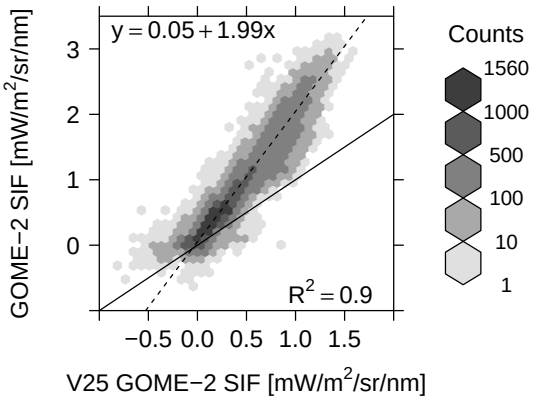


Figure 3.14: SIF retrieval results from the presented algorithm versus V25 results provided by Joiner et al. (2014). SIF averages from February 2007 to December 2011 (rastered on a $0.5° \times 0.5°$ grid) with cloud fractions lower than 0.5 have been used. The one-to-one line (solid line) is depicted as well as a linear fit (dotted line).

It should be considered that the change in retrieval window has lead to a significant decrease in absolute SIF values.

Similar to Fig. 3.11, a scatter plot between the V25 data set and SIF values derived with the presented algorithm is depicted in Fig. 3.14 for the overlapping period in time (February 2007 to December 2011). It is immediately noticeable that the absolute SIF values obtained from the presented retrieval are about 2 times higher. This is inconsistent with our simulation-based retrieval test with a similar fitting window to that of the V25 algorithm from Joiner et al. (2014), which only leads to a decrease of about 10% in absolute values (see Table 3.1, Exp. 1). In view of this discrepancy, it must be stated that there are still large uncertainties in absolute SIF values, while spatiotemporal patterns compare well. At the same time, it should be noted that a temporal consistency might be more relevant for potential users of SIF data sets at the moment, since the absolute value is not applied in current research. Although such a large difference is basically not expected, the data sets are highly correlated and there is a clear linear relationship. This relationship also exhibits seasonal variations as shown in Table 3.2, which has been used to scale the monthly composites in Fig. 3.10 in order to ensure a good visual comparison.

At this point, we are not able to judge which absolute values are closer to reality, since there is a lack of ground truth and validation. The only possibility to assess the validity of results from the data-driven approaches, besides the sensitivity analyses, is at present a comparison to physically based SIF retrieval results from GOSAT data. GOSAT and GOME-2/SCIAMACHY SIF values are expected to be different, which is due to different overpass times, evaluated wavelengths (GOSAT SIF is evaluated between 755–759 nm), illumination and especially observation geometries. Nevertheless, it can be presumed that there is a good correlation between those data sets.

Table 3.2: Relationship between SIF values derived from GOME-2 and SCIAMACHY on a monthly basis to scale SCIAMACHY and V25 GOME-2 SIF values in Fig. 3.10. Shown is the linear fit (GOME-2 SIF = intercept + slope · SCIAMACHY/V25 SIF) similar to those in Fig. 3.11 and Fig. 3.14.

... vs. GOME-2	**January 2011**	**July 2011**
SCIAMACHY	$y = 0.03 + 1.42 \cdot x$	$y = 0.35 + 0.95 \cdot x$
V25 GOME-2	$y = -0.09 + 2.18 \cdot x$	$y = 0.29 + 1.64 \cdot x$

In Fig. 3.15 we compare long-term SIF averages for the June 2009–August 2011 time period from GOME-2/SCIAMACHY SIF data to GOSAT SIF retrieval results obtained by Köhler et al. (2015a). The averaging interval has been selected according to the available overlapping periods for all four data sets. The underlying retrieval results are rastered on a $2° \times 2°$ grid, which is due to the poor spatial sampling of the GOSAT measurements. Furthermore, GOME-2/SCIAMACHY measurements with a cloud fraction lower than 0.5 serve as a basis for the long-term average while the cloud filter for GOSAT soundings consists of an empirical radiance threshold as reported in Köhler et al. (2015a). As can be seen in Fig. 3.15, there is an enhanced correlation to the GOSAT SIF data set for the presented retrieval algorithm. However, a small offset with respect to the GOSAT SIF data set can be observed in all comparisons. Our GOME-2 SIF data has the largest offset, which is probably introduced by a north–south bias in the training set (discussed in Sect. 3.5.6).

In view of these results, it can be assumed that the presented retrieval approach is not less valid than that of Joiner et al. (2013). In general, spatiotemporal patterns compare very well as can be seen in Fig. 3.10; nevertheless, it remains unclear which absolute values are more accurate.

3.5.5 South Atlantic Anomaly

Another point to be considered is that the error estimation might be too optimistic for large parts of the South American continent. The reason is the South Atlantic Anomaly (SAA), which is a region of reduced strength in the Earth's magnetic field. Hence, orbiting satellites are exposed to an increased flux of energetic particles, which leads to increased noise in the measurements. An impact of the SAA on the SIF retrieval using GOME-2 data has also been described in the study of Joiner et al. (2013). Here, we depict the standard error of the mean (SEM, Eq. (3.11)) of monthly SIF composites for January and July 2011 from our SIF retrieval using GOME-2 and SCIAMACHY data in Fig. 3.16 to illustrate the impact of the SAA. This comparison provides also an estimate of the uncertainty, whereas instrumental noise and natural variability are covered. The underlying SCIAMACHY SIF values to produce Fig. 3.16 are also scaled by the linear relationship in Table 3.2 and both composites are computed in a spatial resolution of $1.5° \times 1.5°$ in order ensure a comparison.

The center of the SAA is in close proximity to the coast of Brazil at about 40°W and

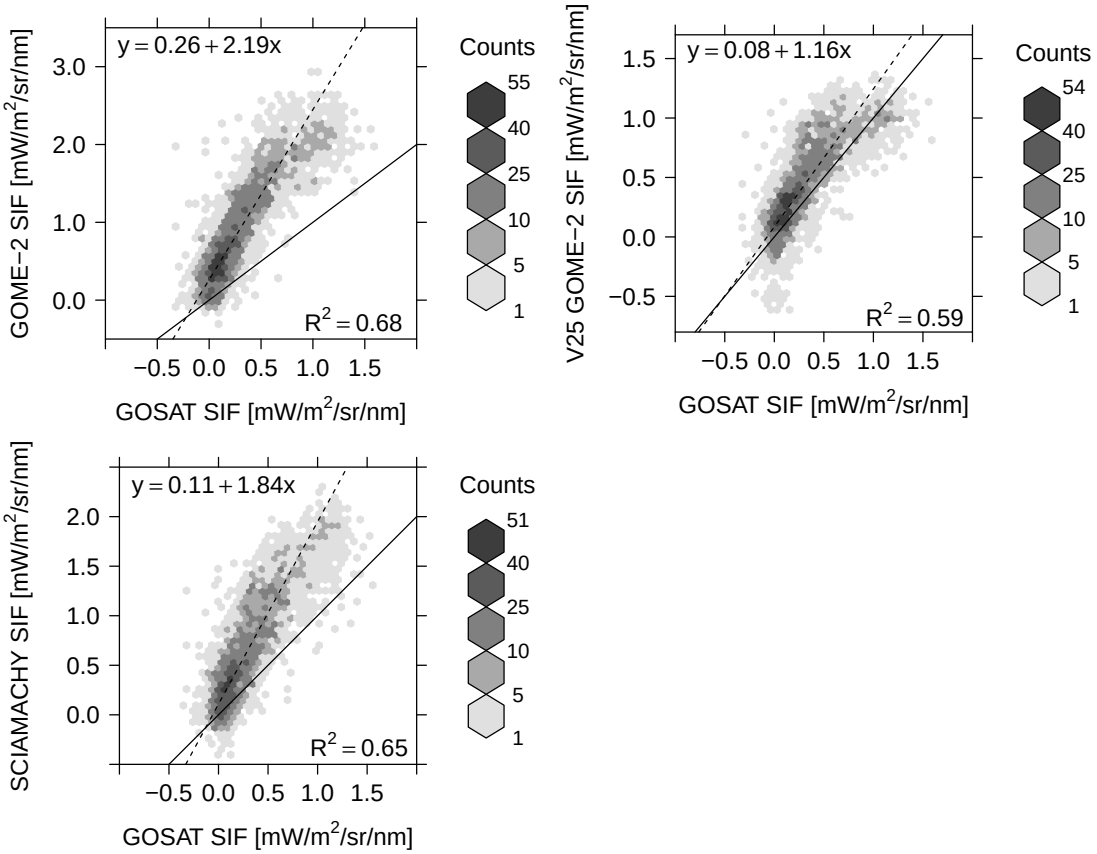


Figure 3.15: SIF retrieval results from GOSAT (evaluated at 755 nm, Köhler et al., 2015a) versus SIF retrieval results from the presented algorithm and V25 results provided by Joiner et al. (2014). SIF averages from June 2009 to August 2011 (rastered on a 2° × 2° grid) have been used. Each evaluated grid box contains at least 26 retrievals using GOSAT data, included GOME-2/SCIAMACHY SIF retrievals own a cloud fraction lower than 0.5. The one-to-one line (solid line) is depicted as well as a linear fit (dotted line).

30°S, which results in extraordinary high SEM values for the GOME-2 SIF data set in nearby regions. Overall, higher SEM values occur where a significant amount of SIF can be observed (see Fig. 3.10). It can also be seen that the SEM is generally lower for the GOME-2 SIF composites (except for the SAA affected regions). It is remarkable that the SAA seems to have a much lower impact on measurements performed by SCIAMACHY in particular for July 2011, although the difference in orbit height is only about 20 km. In view of Fig. 3.16 it is even questionable whether there is an impact of the SAA on SCIAMACHY data at all.

In contrast to the SCIAMACHY data set, the impact of the SAA on GOME-2 data translates also to the residual sum of squares (RSS) of the SIF retrieval (not shown). High

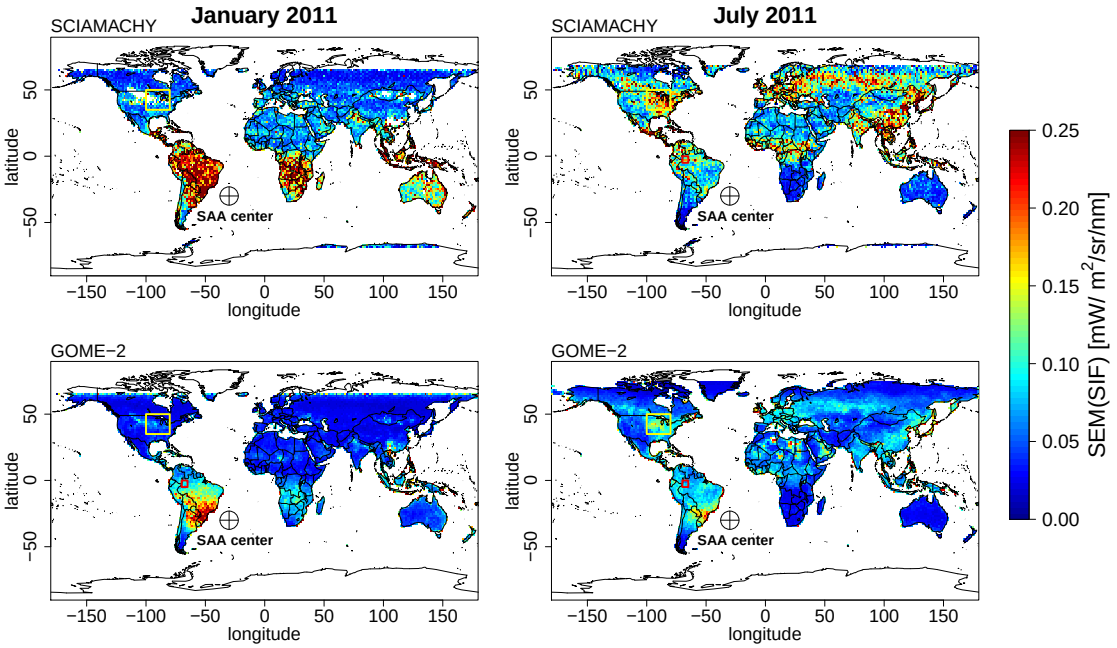


Figure 3.16: Monthly composites of the standard error of the mean (SEM, Eq. (3.11)) of SIF for January (left column) and July (right column) 2011 derived from SCIAMACHY (top) and GOME-2 data (bottom) in a spatial resolution of $1.5° \times 1.5°$. The underlying SCIAMACHY SIF values are again scaled by the relationship in Table 3.2 in order to ensure a good visual comparison. The center of the SAA at 40°W and 30°S, the analyzed region for Fig. 3.12 (yellow) and the analyzed region for Fig. 3.18 (red) are depicted in each map.

RSS values occur in particular in the region of large SEM values on the South American continent. This allows the conclusion that the noise and thus the measurement error is much higher than expected from our error propagation depicted in Fig. 3.13. Although all single retrievals with an RSS bigger than 2 $(\mathrm{mW/m^2/sr/nm})^2$ are filtered, care should be taken when using SIF retrieval results from affected areas, especially for GOME-2 SIF retrievals.

3.5.6 North–south bias in training set

Consistency checks and plausibility controls of the derived SIF values have lead to another source of error. We retrieved SIF values from the training set (sea, desert and ice) with the expectation that all SIF values are close to zero. It turned out that there is a slight time- and latitude-dependent offset in retrieved SIF values which amounts to up to 0.3 $\mathrm{mW/m^2/sr/nm}$ for GOME-2 SIF results. Two different ways to sample the training data have been tested for July 2011 in order to investigate if this bias is related

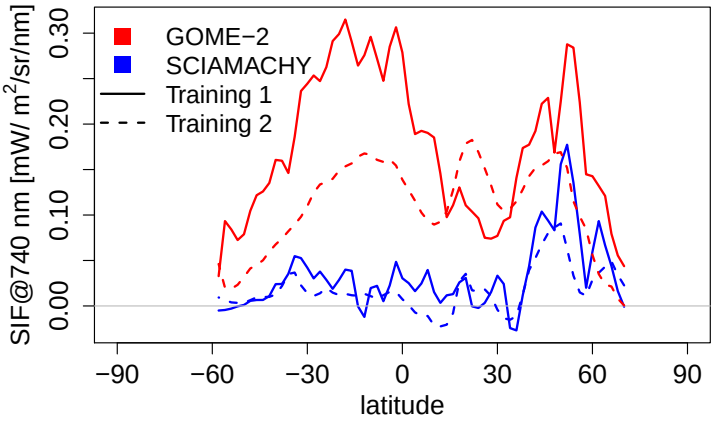


Figure 3.17: Mean SIF of two different training sets derived with data from GOME-2 and SCIAMACHY averaged over latitudes for July 2011. The first training set (Training 1, solid lines) corresponds to a daily sampling of the data while the second training set (Training 2, dotted lines) is sampled on a 3-day basis.

to the sampling method or if there is an instrumental issue. The first training set has been sampled on a daily basis as explained in Sect. 3.5.1 for both satellite instruments. In contrast, the second training set has been sampled on a 3-day basis, whereas the SAA affected region has been excluded (longitudes between 90°W–15°E and latitudes between 10°N–90°S) and spectra have been selected with respect to their latitudinal representativeness (the number of sampled spectra per latitude corresponds to the median number of spectra per latitude). Figure 3.17 depicts the monthly composite of SIF of the two different training sets averaged over latitudes derived from GOME-2 and SCIAMACHY data for July 2011. As can be seen, there is almost no difference between the different sampling methods of the SCIAMACHY training data. Nevertheless, a slight offset of about 0.1 mW/m^2/sr/nm is introduced in latitudes above 40°N. The offset in retrieved SIF values from the GOME-2 training set is decreased when the second sampling method is applied, even though the latitudinal dependency is conserved and the offset is more pronounced than in the two SCIAMACHY training sets. A slight zero level offset is also observable in the data provided by Joiner et al. (2014), which supports the thesis that this bias is related to an instrumental issue. One reason might be a temperature-dependent dark current change. This issue should be considered in further investigations.

3.5.7 Impact of clouds on the retrieval

In contrast to the usually applied pre-filtering of cloud contaminated measurements (discussed in Sect. 3.5.1), we run the retrieval using GOME-2 data in 2009 without restricted cloud fractions in order to examine the impact of clouds.

Tropical rainforest areas are frequently covered by clouds and therefore, it is expected that clouds may affect the retrieval of SIF in particular in such regions. Therefore, we

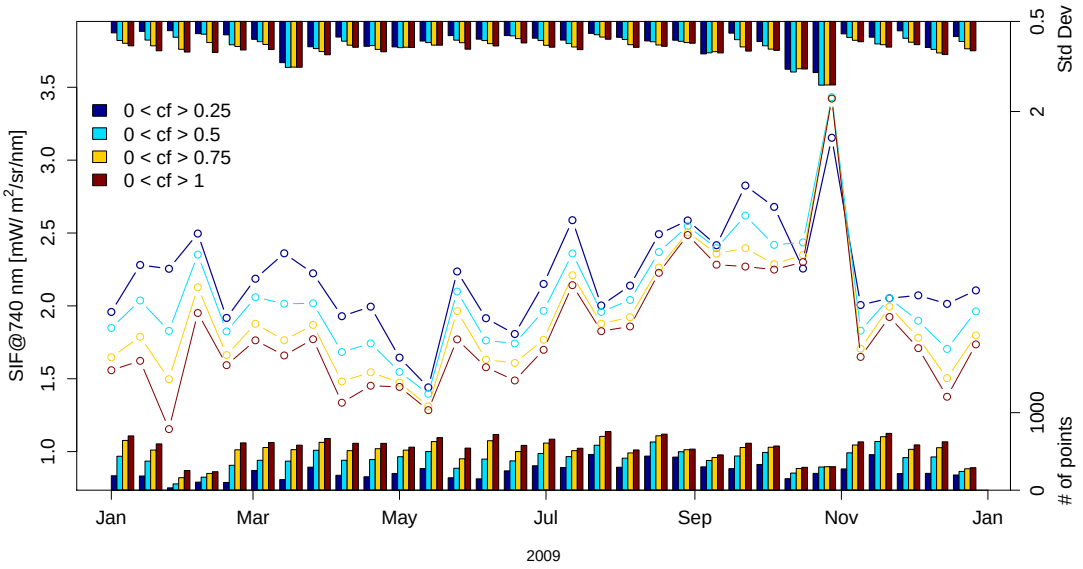


Figure 3.18: Time series of SIF retrievals from GOME-2 data over a box in the Amazon Basin (70°–65°W, 5°–0°S (shown in Fig. 3.16), covered with evergreen broadleaf forest) for different cloud fraction thresholds. The time series corresponds to 12 day means in 2009. The standard deviation (top) and the number of included retrievals (bottom) per time step are depicted as barplots.

produced a time series over a box in the Amazon Basin (70°–65°W, 5°–0°S (shown in Fig. 3.16), covered with evergreen broadleaf forest) with different cloud fraction thresholds, which is shown in Fig. 3.18. Care was taken when selecting the area to ensure that the impact of the SAA is as low as possible and the IGBP land cover (Friedl et al., 2002) (evergreen broadleaf forest) is mostly homogenous.

Overall it can be seen that SIF values are decreasing with an increasing cloud fraction threshold, whereby the temporal pattern remains almost unaffected. Using only retrievals with a low cloud fraction (< 0.25) implies a significant loss of measurements while the standard deviation (or fluctuation) is only slightly smaller and even equal for some time steps. SIF values successively decrease by relaxing the threshold, which is consistent with the expectation that the SIF signal is partly shielded by clouds (Frankenberg et al., 2012). An exception occurs in one time step where the reported standard deviation is significantly higher than in the other time steps. Simultaneously, the number of included retrievals is low and differences in cloud fraction values from FRESCO are only marginal. For these reasons, the time step at the end of October 2009 is not meaningful to evaluate the impact of clouds on the retrieval. Applying a cloud fraction threshold of 0.5 results in differences about 0.1 $\mathrm{mW/m^2/sr/nm}$ in comparison to the SIF averages of the lowest evaluated cloud fraction. The highest increase of measurements with respect

to a cloud fraction threshold of 0.25 can be achieved by using measurements with cloud fractions up to 0.5. A further increase of the cloud fraction has the consequence that more measurements can be used but the amount of SIF averages is distinctly decreasing. It must be noted that results remain consistent for different spatiotemporal scales and the presented region has been chosen as a representative example. On this basis, a cloud fraction threshold of 0.5 is a reasonable compromise between the loss of measurements and changes in the SIF average. As a consequence, we provide the global SIF data set for measurements with cloud fractions lower than 0.5 including also measurements with unknown cloud fractions in the presence of snow. Nevertheless, in view of these results it can be concluded that the presented SIF retrieval method is robust against moderate cloud contamination.

3.6 Conclusions

A new statistically based approach to retrieve SIF from GOME-2 and SCIAMACHY data has been introduced. Building upon previous works from Guanter et al. (2013) and Joiner et al. (2013), our approach solves the important issue of the arbitrary selection of the number of free model parameters.

The basic assumption to retrieve SIF from space is that the contribution of SIF can adequately be separated from low and high frequency components due to surface and atmospheric properties. In our forward model, the low and high frequency components are represented by a combination of a third-order polynomial in wavelength with atmospheric PCs. A backward elimination algorithm selects the required model parameters automatically with respect to the goodness of fit balanced by model complexity. Our sensitivity analysis reveals that the precision is enhanced, retrieval noise is reduced and the risk of an overfitting is minimized by applying the stepwise model selection to our forward model.

The retrieval approach has also been applied to spectra acquired by GOME-2 and SCIAMACHY. Thus, we were able to present a continuous SIF data set covering the 08/2002–03/2012 time period using SCIAMACHY data and the 2007–2011 time span using GOME-2 data. The number of selected PCs and total model parameters for the GOME-2 data is in line with the expectation from the sensitivity analysis while twice as many PCs (on average 14) are employed for SCIAMACHY data. Nevertheless, our approach suggests using a significantly smaller number of PCs as compared to Joiner et al. (2013, 2014).

Although the achievable spatial resolution of SIF maps from SCIAMACHY data is coarser than from GOME-2, there is now a decade long continuous SIF time series available. However, a significant discrepancy in absolute SIF values arises when comparing our retrieval results with the V25 results from Joiner et al. (2014), although spatiotemporal patterns are similar. A detailed comparison of the V25 algorithm with our algorithm using simulated data might explain the discrepancy, but that is beyond the scope of this work. It remains exceedingly difficult at present to judge which values are more accurate. Despite this issue, a comparison to GOSAT SIF data obtained by Köhler et al. (2015a)

showed an enhanced correlation for long-term averages between the different retrieval types, evaluated wavelength and satellite instruments for the presented approach. This allows us to conclude that progress has been achieved but also that there is still research to be done. In particular, there is a need for ground-based validation, ideally in areas with relatively homogeneous vegetation.

Furthermore, it must be considered that the SAA causes high uncertainties in large parts of the South American continent when using GOME-2 data. In contrast, a significant impact of the SAA on SCIAMACHY data has not been found. The analysis of different training sets revealed that there is a slight zero level offset in GOME-2 SIF retrievals which is probably related to an instrumental issue. On the basis of the GOME-2 SIF data set in 2009, we have shown that the retrieval is only moderately affected by cloud contamination. It has essentially been found that the SIF signal decreases with an increasing cloud fraction, but the seasonality is maintained.

Finally, it has to be noted that the flexibility of the retrieval method makes it also applicable to other instruments with a similar spectral and radiometric performance as GOME-2 and SCIAMACHY, such as the upcoming TROPOMI (Veefkind et al., 2012) on board the Sentinel-5 Precursor and the Earth Explorer 8 mission candidates FLEX (Rascher et al., 2008) and CarbonSat (Bovensmann et al., 2010).

Acknowledgements

The research is funded by the Emmy Noether Programme of the German Research Foundation. With thanks to EUMETSAT to make the GOME-2 data available and ESA for providing the SCIAMACHY data. Jochem Verrelst and Luis Alonso from the University of Valencia are gratefully thanked for the reflectance and fluorescence simulations produced in the framework of the ESA FLUSS project.

Author Contributions

Philipp Köhler designed the study, wrote code, processed the GOME-2 and SCIAMACHY data, conducted the simulation-based retrieval test, analysed the output data, and wrote the manuscript. Luis Guanter gave conceptual advice, discussed the results and implications, and commented on the manuscript at all stages. Joanna Joiner provided the V25 comparison data set, discussed the results and implications, and commented on the manuscript.

Chapter 4

Assessing the potential of Sun-Induced Fluorescence and the Canopy Scattering Coefficient to track large-scale vegetation dynamics in Amazon forests

Philipp Köhler[1], Luis Guanter[1], Hideki Kobayashi[2], Sophia Walther[1], and Wei Yang[3].

Journal: Remote Sensing of Environment.

Status: Manuscript received April 23, 2016.

[1] German Centre for Geosciences(GFZ), Remote Sensing Section, Potsdam, Germany.
[2] Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Department of Environmental Geochemical Cycle Research, Yokosuka, Japan.
[3] Center for Environmental Remote Sensing, Chiba University, Chiba, Japan.

Abstract

Two new remote sensing vegetation parameters derived from spaceborne spectrometers and simulated with a three dimensional radiative transfer model have been evaluated in terms of their prospects and drawbacks for the monitoring of dense vegetation canopies: (i) sun-induced chlorophyll fluorescence (SIF), a unique signal emitted by photosynthetically active vegetation and (ii) the canopy scattering coefficient (CSC), a vegetation parameter derived along with the directional area scattering factor (DASF) and expected to be particularly sensitive to leaf optical properties. Here, we present the first global data set of DASF/CSC and examine the potential of CSC and SIF for providing complementary information on the controversially discussed vegetation seasonality in Amazon forests. A comparison between near-infrared SIF derived from the Global Ozone Monitoring Experiment-2 (GOME-2) instrument and the Orbiting Carbon Observatory-2 (OCO-2) (overpass time in the morning and noon, respectively) reveals the response of SIF to instantaneous photosynthetically active radiation (PAR) and the response of SIF to changing pigment concentrations ('green-up'). Large-scale seasonal swings of GOME-2 SIF amount up to 21%, peaking in October and around February, while OCO-2 SIF peaks in February and November. We further examine anisotropic reflectance characteristics with the finding that the hot spot effect significantly impacts observed GOME-2 SIF values. On the contrary, our sensitivity analysis suggests that CSC is highly independent of sun-sensor geometry as well as atmospheric effects. The slight annual variability of CSC (4%) shows a seasonal cycle attributable to variations in leaf area ($R = -0.87$) and/or the amount of precipitation ($R = 0.87$) which rather supports the 'green-up' hypothesis for periods of less intense precipitation. Although there is still research to be done, SIF and CSC may serve as complementary and valuable vegetation parameters under the conditions in the Amazon rainforest.

4.1 Introduction

The importance of the Amazon rainforest regarding its role in the global carbon cycle (e.g., Tian et al., 2000) and its species diversity (e.g., Ter Steege et al., 2013) is represented by a tremendous number of studies. Simultaneously, the sheer number of studies also reveals the challenge posed by the complex ecosystem. Poorly accessible areas and the large extent of the Amazon Basin constrain ground-based observation capacities. Satellite-based remote sensing therefore provides a crucial opportunity to continually monitor these areas, but even recent findings of satellite-based analyses addressing vegetation dynamics, seasonality as well as its response to drought events remain contradictory.

The majority of studies report an increase ('green-up') in photosynthetic activity during months with less intense precipitation (usually June-October; e.g., Huete et al., 2006;

Samanta et al., 2012; Bi et al., 2015). In contrast, Morton et al. (2014) claim that structure and greenness and thus photosynthetic activity remain constant if the satellite data are accurately corrected for directional reflectance effects in the measurement. These controversial results might be explained due to the fact that the monitoring of dense vegetation represents a complicated case in optical remote sensing: Reflected solar radiation saturates and becomes weakly sensitive to vegetation changes, while there is a substantial influence of changing sun-sensor geometry and residual atmospheric effects (Verrelst et al., 2008; Hilker et al., 2015; Maeda and Galvão, 2015).

Thanks to recent advances in the analysis of spaceborne spectroscopic measurements, two new and promising vegetation parameters, sun-induced chlorophyll fluorescence (SIF) and the canopy scattering coefficient (CSC), have become available. While SIF is directly related to photosynthesis, CSC is a structural parameter assumed to be particularly sensitive to leaf properties (Samanta et al., 2012; Knyazikhin et al., 2013). The CSC may capture leaf age-effects when considering that Roberts et al. (1998) observed an 10% increase in the near-infrared (NIR) absorption in matured tropical leaf spectra. A strong argument to lay particular emphasis on such effects was recently raised by Wu et al. (2016). They specifically examined seasonal variations of several vegetation parameters from ground, tower and satellite measurements at four sites in the Amazon and concluded that the canopy phenology, i.e., the leaf age, drives the photosynthetic seasonality rather than seasonal variations of climate conditions. It would therefore be clearly beneficial if a basin-wide measure of leaf phenology could be established. On the other hand, three characteristics of SIF might be advantageous compared to traditional reflectance-based vegetation indices (VIs): (i) a saturation of SIF has not yet been observed, (ii) SIF appears to be less affected by sub-pixel clouds (Frankenberg et al., 2012; Guanter et al., 2015), (iii) SIF is expected to provide a direct link to actual photosynthetic rates, while greenness based indices indicate potential photosynthesis represented by their response to chlorophyll content. Both SIF and CSC may therefore contain relevant information to re-assess vegetation dynamics in the Amazon region.

Global SIF data sets are available from several spaceborne sensors, including:

1. Greenhouse Gases Observing Satellite-Fourier Transform Spectrometer (GOSAT-FTS; Frankenberg et al., 2011a; Joiner et al., 2011, 2012; Guanter et al., 2012; Köhler et al., 2015a),

2. SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY; Joiner et al., 2012; Köhler et al., 2015b; Wolanin et al., 2015),

3. Global Ozone Monitoring Experiment-2 (GOME-2; Joiner et al., 2013; Köhler et al., 2015b; Wolanin et al., 2015),

4. Orbiting Carbon Observatory-2 (OCO-2; Frankenberg et al., 2014).

However, only a few studies relating SIF to productivity in tropical forests can be found in the literature so far. Lee et al. (2013) and Parazoo et al. (2013) reported that wet season productivity (proxied by SIF) exceeds the 'dry' season productivity in southern Amazon forest regions, where pronounced precipitation variations occur. Guan et al. (2015)

examined global tropical evergreen forest regions and concluded that the photosynthetic activity can be maintained during the drier season if the annual water availability is sufficient.

Based on the concept of canopy spectral invariants (e.g., Knyazikhin et al., 2011), Samanta et al. (2012) expressed the spectral reflectance in terms of directional area scattering factor (DASF; Knyazikhin et al., 2013) and CSC (Smolander and Stenberg, 2005). By means of detailed theoretical considerations with respect to this expression, Samanta et al. (2012) argue that changes in NIR reflectance of Amazon forests could be explained by both changes in leaf area and leaf optical properties, but unlikely by changes in leaf area alone. Leaf optical property variations due to leaf aging effects translate into changes in CSC; however, Samanta et al. (2012) simultaneously acknowledge that leaf optical variations might also be caused by coating substances (e.g., water or dust).

GOME-2 is at present the only spaceborne spectrometer that provides time series of spectroscopic measurements in the 650–800 nm wavelength range, enabling simultaneous computations of DASF, CSC, SIF and NDVI. An encouraging aspect for retrieving several vegetation parameters from GOME-2 data lies in the wide range of covered observation geometries, allowing to address directional effects. Albeit the importance of directional effects has increasingly been recognized when analyzing VIs, less is known about the effects on satellite-based SIF data sets, though several studies suggest to similarly consider changing illumination and view conditions (van der Tol et al., 2009b; Guanter et al., 2012; Damm et al., 2015; Liu et al., 2016). Another benefit of simultaneous retrievals of different vegetation parameters from one sensor is the consistent spatial sampling, even though GOME-2 measurements have a relatively coarse spatial resolution as opposed to the more frequently used products obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS). The large footprint size of GOME-2 (40 km x 40 km) likely involves undetected sub-pixel clouds, representing a further limiting factor.

Here, we primarily aim to evaluate prospects and drawbacks of both SIF and CSC for tracking vegetation dynamics and productivity in Amazon forests by examining simulated and real satellite data. We present SIF time series derived from measurements of the GOME-2 instrument (Köhler et al., 2015b) and from OCO-2 (Frankenberg et al., 2014) for a region within the central Amazon Basin and we subsequently discuss limitations due to sun-sensor geometry and illumination conditions at the overpass time. Moreover, DASF and CSC have been computed from atmospherically corrected GOME-2 data which enables us to present the first global data set of these parameters. We have also computed the well-established Normalized Difference Vegetation Index (NDVI) and make use of supplementary data sets to provide a basis for interpreting the results obtained. Finally, we present long-term monthly averages to illustrate the large-scale seasonality of SIF, CSC and NDVI in the Amazon forest as seen from the GOME-2 instrument.

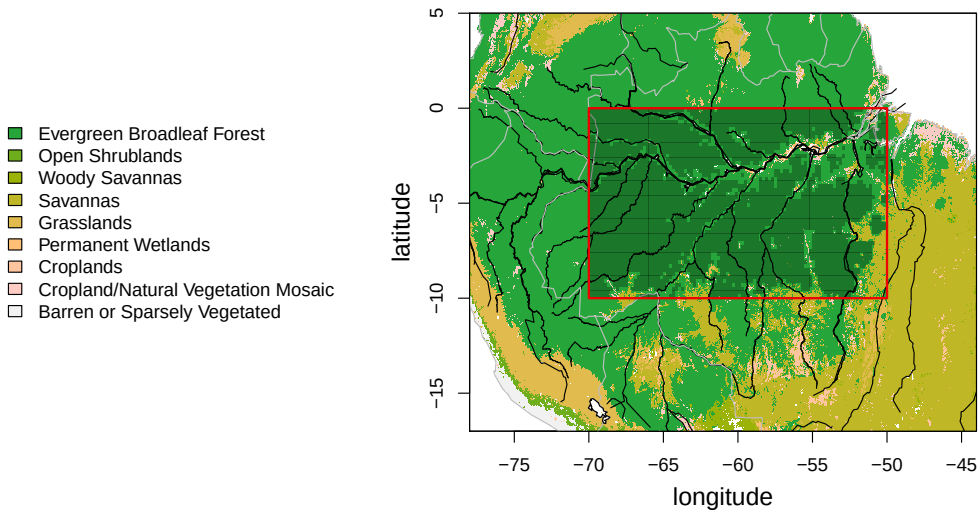


Figure 4.1: IGBP land classification (Friedl et al., 2002). The darkened areas within the red box define the study domain.

4.2 Data & Methods

4.2.1 GOME-2

We use spectral radiance measurements from the GOME-2 instrument onboard EUMETSAT's polar orbiting Meteorological Operational Satellite MetOp-A (Munro et al., 2006) acquired between August 2013 and April 2015. This time period has been selected because the swath width has changed since 15th July 2013 from 1920 km to 960 km, resulting in a reduced ground pixel size of about 40 km x 40 km (previously 80 km x 40 km). Spectra acquired by the fourth detector channel, covering the 600 to 790 nm spectral range with a resolution of 0.5 nm and a signal-to-noise ratio up to 2000, are relevant for our analysis. The effective cloud fraction from the Fast Retrieval Scheme for Clouds from the O2 A-band (FRESCO, Wang et al., 2008) is used for cloud screening.

4.2.2 Study Domain & Spatial Averaging

Our analysis of large-scale vegetation dynamics (Sect. 4.4.4) is focused on a domain in southern Amazonia, indicated by the red box in Fig. 4.1 (latitudes between 0° - 10° S and longitudes between 70°W–50°W). Special care has been taken to minimize the impact of structural composition effects on the resulting time series as it will be explained below.

At first, GOME-2 measurements are pre-selected based on the International Geosphere and Biosphere Program (IGBP) land classification (Friedl et al., 2002) shown in Fig. 4.1. The main land cover is determined under consideration that each GOME-2 footprint has a spatial extent represented by a polygon. Measurement polygons classified as evergreen broadleaf forest with a land cover fraction higher than 0.9 form the basis for the spatial

average. Pre-selected polygons are subsequently averaged on a monthly basis on a 0.2° x 0.2° grid (about 22 km x 22 km), while a value is transferred if the polygon covers the center of a grid cell.

The study domain is further confined by masking grid cells which potentially own differing structural properties compared to the majority of grid cells. In particular, a k-means approach has been applied to the supplementary LAI data set (see Sect. 4.2.8) to identify such grid cells. The remaining grid cells, shown as darkened areas in Fig. 4.1, define the underlying region for the actual analysis. It is remarkable that particularly grid cells close to the Amazon River and other water bodies were identified to have different structural properties which might be associated with large floodplain areas (Várzea forests).

4.2.3 Sun-induced chlorophyll fluorescence (SIF)

Chlorophyll pigments in terrestrial vegetation absorb sunlight to cover the energy demand during the process of photosynthesis. A part of the excess energy dissipates as sun-induced fluorescence (SIF) emitted by chlorophyll *a* in the red and NIR spectral region (650-800 nm). Though the amount of SIF represents only a few percent of the total light energy absorbed, SIF measurements can serve as a valuable tool to assess the photosynthetic performance of vegetation. The capability of spaceborne SIF measurements for providing a link to the carbon uptake through photosynthesis (gross primary production, GPP) has already been proven for Amazon regions (Lee et al., 2013; Parazoo et al., 2013) as well as for cropland and grassland ecosystems (Guanter et al., 2014). However, it is further important to consider the instantaneous nature of satellite-based SIF data when evaluating seasonal cycles. For example, SIF time series intrinsically depend on the illumination condition at the overpass time, while the temporal trajectory (e.g., maximum and minimum) changes with the considered daytime (discussed in detail in Sect. 4.4.2).

In contrast to reflectance based VIs, which are sensitive to canopy structure and pigment concentrations, it can be assumed that SIF also tracks rapid changes in the vegetation photosynthetic status. Drawbacks arise from the rather coarse spatial resolution of gridded SIF data, which is due either to a large footprint size (GOME-2, SCIAMACHY) or to spatially discrete measurements (GOSAT, OCO-2). Here, we analyze SIF estimates from GOME-2 (overpass time at ∼9:30 local solar time, LST) and OCO-2 (overpass time at ∼13:30 LST) data derived with fundamentally different retrieval approaches as discussed below.

GOME-2 SIF

The retrieval methodology of the GOME-2 SIF data set used in this study was described by Köhler et al. (2015b). This data set has been extended until April 2015 (previously 2007-2011) and can be retrieved from `ftp://ftp.gfz-potsdam.de/home/mefe/GlobFluo/`. GOME-2 SIF retrievals generally rely on a statistically based approach to separate the SIF emission from spectral features related to atmospheric absorption, scat-

tering and surface reflectance. This specific retrieval essentially reconstructs the radiance spectrum in the 720-758 nm interval and provides a value for the second of two characteristic peaks of the fluorescence emission spectrum at 740 nm. Modeled spectra are compared to the original measurements to exclude corrupted retrievals. This is done by a radiance residual check which has been reinforced for our investigation in the Amazon Basin to reduce the impact of potential error sources on resulting time series. In particular, Joiner et al. (2013) and Köhler et al. (2015b) reported that the South Atlantic Anomaly, a region with an anomalously reduced strength in the Earth's magnetic field, leads to an increased retrieval noise in large parts of the South American continent. The residual sum of squares (RSS) is therefore limited to 0.5 (mW m^{-2} sr^{-1} nm^{-1})^2 instead of using the 2 (mW m^{-2} sr^{-1} nm^{-1})^2 threshold documented by Köhler et al. (2015b). The number of evaluated retrievals within the study domain is consequently reduced by about two thirds.

OCO-2 SIF

The OCO-2 instrument was launched in July 2014 and provides spectrally high resolved measurements in the O_2 A-band (757-775 nm, FWHM=0.042 nm), which allows to evaluate the in-filling of solar Fraunhofer lines at 757 and 770.1 nm by SIF (Frankenberg et al., 2014). OCO-2 acquires 24 spectra per second with a much smaller ground-pixel size (1.3 km x 2.25 km) as compared to GOME-2 (40 km x 40 km). The high spectral resolution enables robust and accurate retrieval results (Frankenberg et al., 2014), but the fine spatial sampling with the narrow swath width of 10.3 km increases the revisit time up to 16 days (GOME-2: 3 days) and involves a lack of continuous global coverage (i.e., gaps between swaths for a full repeat cycle). Here, we use cloud free nadir observations of offset corrected OCO-2 SIF at 757 nm (version B7101) between September 2014 and November 2015.

4.2.4 Atmospheric correction of GOME-2 radiance spectra

Radiance received by GOME-2 at the top-of-atmosphere (TOA) is composed of direct and diffuse irradiance contributions from both atmosphere and surface. The magnitude of atmospheric effects can vary largely depending on the atmospheric state, surface properties and wavelength. Thus it is necessary to correct the TOA signal for atmospheric effects to estimate the actual surface reflectance (ρ_s), which is in turn used to derive other parameters (ρ_{NIR}, DASF, CSC and NDVI).

We based the atmospheric correction on simulations performed with the Matrix Operator MOdel (MOMO) radiative transer code (Fell and Fischer, 2001), mostly consistent to a subset of simulations used in Joiner et al. (2013); Guanter et al. (2015); Köhler et al. (2015b). Simulations were performed between 500 nm and 800 nm with a spectral sampling of 0.005 nm and include three viewing zenith angles (0°, 16°, 27°), four solar zenith angles (15°, 30°, 45°, 70°), a mid-latitude summer temperature profile, four surface pressures (955, 980, 1005, 1030 hPa), four water vapor columns (0.5, 1.5, 2.5,

4.0 g cm^{-2}), an aerosol layer height between 600–800 hPa using a continental aerosol model and five aerosol optical thicknesses at 550 nm (0.05, 0.12, 0.2, 0.3, 0.4).

A solution for the radiative transfer in a plane-parallel atmosphere with a perfectly diffuse (Lambertian) surface can be formulated as

$$F_{TOA} = F_0 + \frac{E_g \cdot \rho_s}{\pi(1 - \rho_s S_{atm})}, \tag{4.1}$$

where F_{TOA} is the (measured) TOA radiance, F_0 is the atmospheric path radiance, E_g is the transmitted global irradiance (direct and diffuse), S_{atm} is the atmospheric spherical albedo (i.e., the reflectance of the atmosphere for isotropic light entering it from the surface) and ρ_s is the surface reflectance (all variables own a spectral component). F_{TOA}, F_0, E_g and S are assumed to be functions of solar zenith angle (SZA), viewing zenith angle (VZA), mean sea level pressure (MSL), water vapor column (WVC) and aerosol optical thickness at 550 nm (AOD). Azimuthal resolution of sun-sensor geometry and adjacency effect are not included. Although a natural surface reflects solar radiation anisotropically, we assume that Lambertian equivalent surface reflectance estimates through Eq. (4.1) satisfy the requirements for this analysis. Hu et al. (1999) evaluated relative errors in canopy reflectance estimates for MODIS data when using a Lambertian assumption. They reported that MODIS relative errors are about 5% in the red band and 2% in the NIR band (for hardwood forest and an AOD of 0.2; median AOD of observations used here is about 0.15).

A lookup table approach in conjunction with supplementary data sets (see Sect. 4.2.8) of MSL, WVC and AOD is applied to estimate the spectral bidirectional reflectance. Fig. 4.2 depicts a sample measurement spectrum of GOME-2 together with the atmospherically corrected surface reflectance estimate in the relevant wavelength range between 660–790 nm. It needs to be mentioned that the highly complex radiative transfer in the strong O_2 A-band (759-774 nm) causes a high frequency spectral pattern in ρ_s, although vegetation has a spectrally smooth reflectance signature in the considered wavelength range. We therefore fit a second order polynomial to the reflectance in two spectral windows around the O_2 A-band, assumed to be mostly devoid of atmospheric absorption (743-758 mn and 774-782 nm), and replace the unreasonable reflectance estimates.

4.2.5 Directional Area Scattering Factor (DASF) & Canopy Scattering Coefficient (CSC)

We adopted the approach of Knyazikhin et al. (2013) to retrieve CSC from atmospherically corrected NIR reflectances using GOME-2 data. This approach relies on the idea to scale scattering processes from the leaf to the canopy level. The theory of such a scaling has made significant progress since one of the initial attempts from Knyazikhin et al. (1998). In particular, the concept of canopy spectral invariants (summarized in Huang et al., 2007) provides the basis for a physically consistent and thus more direct approach to relate reflectance-based measurements to leaf structure and biochemistry.

First, it should be considered that the recollision probability (p) for scattered photons

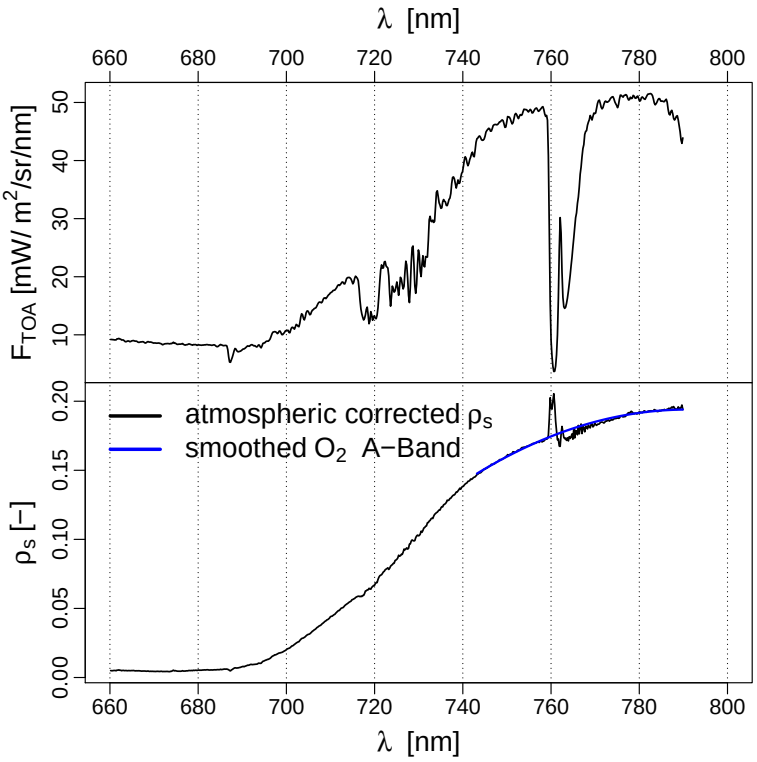


Figure 4.2: Sample measurement spectrum of GOME-2 and atmospherically corrected surface reflectance estimate between 660-790 nm (633 spectral points). The measurement within the study domain (center latitude 1.1° S, center longitude 57.3° W) was carried out at 07/04/2014, atmospheric correction was performed with following input parameters: SZA = 46.8°, VZA = 10.7°, AOD = 0.06, WVC = 3.87 g/cm^2, MSL = 1015.4 hPa. A second order polynomial fit using the 743-758 mn and 774-782 nm spectral intervals (blue) is applied to replace the unreasonable reflectance estimates inside the O_2 A-band (759-774 nm) by fitted values.

within a canopy is wavelength independent (spectral invariant) because scattering objects are large compared to the wavelength of solar radiation (Lewis and Disney, 2007). Smolander and Stenberg (2005) have shown that a combination of different recollision probabilities on the leaf level, namely for shoots and needles as the most important scattering elements within coniferous canopies, results in one equivalent recollision probability on the canopy scale. The scaling even applies from the leaf interior to the canopy level when neglecting leaf surface reflection (Lewis and Disney, 2007). Besides simulation based studies (Knyazikhin et al., 1998; Smolander and Stenberg, 2005; Huang et al., 2007; Lewis and Disney, 2007), the p theory has been succesfully applied to remote sensing applications (e.g, Schull et al., 2011; Knyazikhin et al., 2013). Particularly the

study from Knyazikhin et al. (2013) may have long-term consequences for the remote sensing of canopy chemistry. They introduced the following representation of the canopy bidirectional reflectance (BRF)

$$BRF_\lambda(\Omega) = DASF(\Omega) \cdot W_\lambda, \tag{4.2}$$

where Ω is the direction towards the sensor, DASF is the directional area scattering factor and W_λ the (wavelength dependent) canopy scattering coefficient. It might be noted that BRF_λ is equivalent to ρ_s in Eq. (4.1). The wording is changed in this section to remain consistent with the literature. Furthermore, it should be emphasized that W_λ refers to the spectral canopy scattering coefficient, while the abbreviation CSC refers to the canopy scattering coefficient at NIR wavelengths (discussed below in Sect. 4.2.5). The DASF is a spectral invariant estimate of the BRF_λ for a non-absorbing canopy and encapsulates macro-structural properties such as the tree spatial distribution. The key benefit of the approach from Knyazikhin et al. (2013) is that no prior knowledge is needed to estimate the DASF from reflectance spectra in the 710 to 790 nm spectral interval as described in the following.

DASF estimation

Assuming that the observed canopy is sufficiently dense (background/surface effects are negligible), the spectral bidirectional reflectance (BRF) can be approximated as (Knyazikhin et al., 2011)

$$BRF_\lambda(\Omega) = \frac{\rho(\Omega) i_0(\Omega_0) \omega_\lambda}{1 - \omega_\lambda p}, \tag{4.3}$$

where $\rho(\Omega)$ is the directional gap density giving the escape probability for scattered photons in direction Ω, $i_0(\Omega_0)$ is the initial collision probability or canopy interceptance for the direction of incoming radiation Ω_0 (fraction of photons that initially collide with foliage elements, close to unity in dense vegetation canopies), p is the mean probability that a scattered photon will interact within the canopy again (recollision probability) and ω_λ is the canopy albedo.

Eq. (4.3) is still formulated on the canopy level but p implicitly incorporates the scaling to the leaf level. Specifically, the scaling to the next smaller structural organization (i.e., from the canopy to the leaf level) can be expressed by accumulated scattering orders of local recollision probabilities (Smolander and Stenberg, 2005)

$$\omega_\lambda = \sum_{n=0}^{\infty} [p_L^n \omega_{L,\lambda}^n (1 - p_L)] = \frac{1 - p_L}{1 - p_L \omega_{L,\lambda}} \omega_{L,\lambda}, \tag{4.4}$$

where the leaf scale is indicated by the subscripted L. An empirical analysis of measured leaf albedo spectra suggests that one fixed reference leaf albedo ($\omega_{0,\lambda}$) can be used instead of the actual single leaf albedo ($\omega_{L,\lambda}$) without violating the spectral invariant relationship from Eq. (4.4) in the 710 to 790 nm spectral interval (Schull et al., 2011). It is further

necessary to decompose ω_λ into

$$\omega_\lambda = i_L \bar{\omega}_\lambda + s_L, \tag{4.5}$$

where i_L is the leaf interceptance, $\bar{\omega}_\lambda$ is the transformed leaf albedo giving the fraction of photons scattered by the interior of leaves and s_L is the portion of photons reflected from the leaf surface. Assuming that the scattering by the leaf interior dominates in the 710 to 790 nm spectral region, Eq. (4.4) becomes (Knyazikhin et al., 2013)

$$\omega_\lambda \approx i_L \bar{\omega}_\lambda = \frac{1 - p_L}{1 - p_L \omega_{0,\lambda}} i_L \omega_{0,\lambda}. \tag{4.6}$$

By substituting Eq. (4.6) into Eq. (4.3) and rearranging terms, Knyazikhin et al. (2013) arrived at

$$\frac{BRF_\lambda(\Omega)}{\omega_{0,\lambda}} = p_{eff} BRF_\lambda(\Omega) + i_L \rho(\Omega) i_0 (1 - p_L), \tag{4.7}$$

with $p_{eff} = p_L + i_L p(1 - p_L)$. It might be noted that p_{eff} is equivalent to the spectral invariant relationship between recollision probabilities of shoots and needles documented by Smolander and Stenberg (2005). Here, p_{eff} describes the scattering within leaves as seen from the canopy level, i.e., the probability of being scattered inside the leaf (p_L) is combined with the probability that a photon escapes the leaf ($1 - p_L$) and enters another leaf within the canopy ($i_L p$). A simple linear regression between $\mathrm{BRF}_\lambda/\omega_{0,\lambda}$ vs. BRF_λ (see Eq. 4.7) gives the slope $m = p_{eff}$ and intercept $n = i_L \rho(\Omega) i_0 (1 - p_L)$. Finally, the ratio $n/(1 - m)$ results in the following representation of the DASF (Knyazikhin et al., 2013)

$$DASF(\Omega) = \frac{i_L \rho(\Omega) i_0(\Omega_0)}{1 - i_L p}. \tag{4.8}$$

Note that the DASF becomes independent of p_L; and moreover, the DASF appears to be independent of the choice of $\omega_{0,\lambda}$ (Schull et al., 2011), which both seems reasonable considering that the DASF conveys macro-structural properties. A reference leaf albedo derived with the PROSPECT model (chlorophyll content of 16 μg cm^2, equivalent water thickness of 0.005 cm^{-1}, dry matter content of 0.002 g/cm^2 as suggested by Knyazikhin et al., 2013) is used for the linear regression in the 710-790 nm spectral region. The R^2 value of the linear fit indicates the retrieval quality and applicability of assumptions (e.g., sufficiently dense vegetation). Our analysis is therefore based on DASF retrievals with a R^2 value above 0.9 (this threshold will be further discussed in Sect. 4.4.1). It is further worth noting that NIR reflectance and DASF are closely related in terms of their sensitivity to view-illumination conditions.

CSC & NIR leaf albedo estimation

When expressing BRF_λ in terms of DASF and W_λ (Eq. 4.2) and considering that the DASF conveys macro-structural properties, leaf properties (represented by W_λ) can potentially be extracted from reflectance measurements through the ratio BRF_λ/DASF. In contrast to visible wavelengths where pigment absorption dominates, residual variations in the NIR are mostly due to canopy/leaf structure (Hikosaka et al., 2015). Here, the 780-785 nm NIR spectral interval serves as weakly absorbing wavelength range to compute CSC (or W_{NIR}).

An analytical representation of W_λ is given by (Knyazikhin et al., 2013)

$$W_\lambda = \frac{1 - pi_L}{1 - \hat{\omega}_\lambda pi_L}\hat{\omega}_\lambda, \tag{4.9}$$

with $\hat{\omega}_\lambda = \omega_\lambda / i_L$. In this case, p is the probability that a scattered photon will interact again within the canopy (recollision probability), ω_λ is the leaf albedo and i_L is the leaf interceptance (fraction of radiation that enters the leaf interior). Note that Eq. (4.9) includes no explicit dependency on illumination or observation geometry, although the leaf albedo is in principle subject to a certain bidirectional reflectance distribution function (BRDF). However, considering that the ground pixel size of GOME-2 (40 km x 40 km) involves a vast number of leaves, it can be anticipated that any angular preference will tend to average out. Conceptually, this vegetation parameter should therefore be independent of sun-sensor geometry, which is indeed supported by our results in Fig. 4.9 of Sect. 4.4.3.

W_λ increases as the albedo increases, but as can be seen from Eq. (4.9), W_λ is also subject to changes in canopy structure through the dependency on the p value. More specifically, p increases with LAI (Knyazikhin et al., 1998; Smolander and Stenberg, 2005) which can be explained by a higher chance for photons to be absorbed (Samanta et al., 2012). Consequently, the absolute value of W_λ decreases with increasing LAI. Smolander and Stenberg (2005) approximate the relationship between p and LAI by

$$p = p_{max}[1 - exp(-k\ LAI^b)], \tag{4.10}$$

with $p_{max} = 0.88$, $k = 0.7$ and $b = 0.75$. We use this relationship to extract large-scale variations in leaf properties (Sect. 4.4.4) by estimating p from the supplementary LAI data set and evaluating the rearranged Eq. (4.9)

$$\hat{\omega}_{NIR} = \frac{\omega_{NIR}}{i_L} = \frac{W_{NIR}}{i_L p(W_{NIR} - 1) + 1}, \tag{4.11}$$

under the assumption of $i_L = 1$, neglecting the sensitivity of CSC (W_{NIR}) to leaf surface properties (Smolander and Stenberg, 2005, have introduced W_λ under this assumption). However, it should be mentioned that the approximation in Eq. (4.11) might be critical as results rely on (i) the accuracy of both CSC and LAI estimates in addition to (ii) the validity and applicability of Eq. (4.11) under the conditions in the Amazon forests. We

will further evaluate this approach in Sect. 4.3.2, based on simulated data.

The formalism to compute DASF, CSC and NIR leaf albedo (ω_{NIR}) can be summarized as follows:

1. Perform atmospheric correction of measured BRF_λ (Sect. 4.2.4).
2. Perform a linear fit between $\text{BRF}_\lambda/\omega_{0,\lambda}$ vs. BRF_λ in the 710–790 nm spectral window (389 spectral points), where $\omega_{0,\lambda}$ is a reference leaf albedo (Sect. 4.2.5).
3. The ratio $n/(1-m)$ of slope, m, and intercept, n, from the linear fit results in DASF as represented by Eq. (4.8).
4. CSC (W_{NIR}) results from the ratio BRF_{NIR}/DASF, where BRF_{NIR} is the averaged atmospherically corrected reflectance between 780-785 nm.
5. Compute p from supplementary LAI data through Eq. (4.10) and evaluate Eq. (4.11) to estimate NIR leaf albedo (ω_{NIR}) from p and CSC.

4.2.6 Normalized Difference Vegetation Index

For the sake of comparison, the simple but effective Normalized Difference Vegetation Index (NDVI, Tucker, 1979) has been computed from GOME-2 data. The standard definition is given by the band ratio

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}, \tag{4.12}$$

where ρ_{NIR} and ρ_{RED} are computed from averaged atmospherically corrected reflectances between 780-785 nm and 660-665 nm, respectively. Joiner et al. (2013) reported that GOME-2 NDVI spatial patterns are similar to those of the MODIS NDVI product, even when atmospheric effects are not considered.

We acknowledge that the large-scale NDVI seasonality in the central Amazon Basin will potentially result from seasonal variations in cloud cover (Kobayashi and Dye, 2005; Hilker et al., 2015). Nevertheless, we consider that a spatially consistent comparison of GOME-2 SIF and CSC with a conventional VI provides an improved basis for the interpretation.

4.2.7 Characterization of Sun-Sensor Geometry

In general, four angles are needed to characterize the illumination and observation geometry, namely: solar zenith angle (θ_0), solar azimuth angle (ϕ_0), viewing zenith angle (θ_v) and viewing azimuth angle (ϕ_v). The position of the azimuth angles to each other can also be expressed by the relative azimuth angle (ϕ), where $\phi = 180°$ means that sun and satellite sensor are oppositely aligned. It is further appropriate to combine θ_0, θ_v and ϕ to the phase angle (γ) by applying the spherical law of cosines

$$\gamma = acos[cos(\theta_0)cos(\theta_v) + sin(\theta_0)sin(\theta_v)cos(\phi)], \tag{4.13}$$

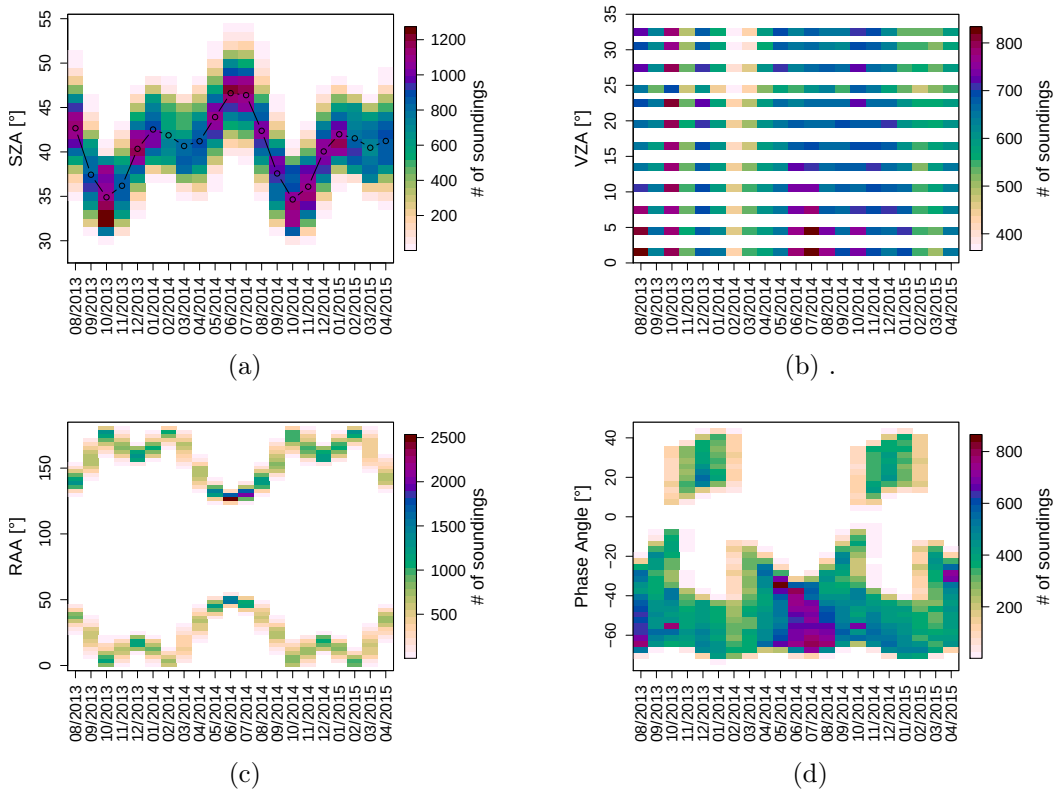


Figure 4.3: Viewing and illumination geometry of GOME-2 within our study domain. Shown are (a) solar zenith angle, (b) viewing zenith angle, (c) relative azimuth angle and (d) phase angle (Eq. 4.13).

which gives the angle between the directions to the sun and detector as seen from the surface (Hapke, 2012). Here, γ is defined positive if the solar azimuth angle is larger than viewing azimuth angle and negative otherwise to be able to separate those cases. The phase angle replaces three angles to characterize the sun-sensor geometry, while approaching $\gamma = 0°$ corresponds to the situation when sun and sensor are located along one axis.

Fig. 4.3 depicts the illumination and viewing conditions for GOME-2 measurements within our study domain. The seasonality of the sun-sensor geometry is comparable to MODIS observations shown by Bi et al. (2015), whereas several studies (Morton et al., 2014; Maeda and Galvão, 2015; Verrelst et al., 2008; Hilker et al., 2015) highlight the importance to take such a seasonality into account when evaluating VIs. In the case of MODIS, measurements are typically related to a fixed sun-surface-sensor geometry. We follow another strategy in Sect. 4.4.4: GOME-2 observations are limited to phase angles from $-65°$ to $-40°$ because this range of angles occurs in all considered months. Thus, a seasonality in sun-sensor geometry and a potential influence on time series can

be avoided.

4.2.8 Supplementary data

A series of data is required for our atmospheric correction and beyond that, for providing a basis to interpret the results obtained. The following list briefly summarizes all supplementary data sets used in this study.

(a) Cloud free nadir observations of offset corrected OCO-2 SIF at 757 nm (Frankenberg et al., 2014, version B7101) between September 2014–November 2015.

(b) IGBP land cover (Friedl et al., 2002) provided on a 0.05° x 0.05° (about 5 km x 5 km) grid.

(c) LAI (Baret et al., 2013) from the Copernicus Global Land Service. Specifically, the SPOT-VGT LAI (version 1.3) is used between 08/2013 and 05/2014 and the PROBA-V LAI (version 1.4) from 06/2014 to 04/2015. Both data sets are provided on a 10-day basis with a spatial resolution of 1 km and have been aggregated to monthly averages on a 0.2° x 0.2° grid.

(d) Collection 5 of the Aqua MODIS NDVI data provided on a 0.05° x 0.05° grid (MYD13C1, 2016, 16-day Vegetation Indices) quality flagged as good ('use with confidence'), aggregated to monthly averages on a 0.2° x 0.2° grid.

(e) Monthly precipitation data from the Tropical Rainfall Measuring Mission (TRMM, 2011, product 3B43 version 7) on a 0.25° x 0.25° grid.

(f) ERA-Interim re-analysis data (Dee et al., 2011) of photosynthetically active radiation (PAR, 400-700 nm, 3-hourly), WVC (6-hourly) and MSL (6-hourly). Here, WVC and MSL (PAR) data sets on a 0.75° x 0.75° (0.125° x 0.125°) grid have been used.

(g) AOD (6-hourly) from the Monitoring Atmospheric Composition and Climate (MACC, 2016) re-analysis data set on a 0.75° x 0.75° grid.

4.2.9 Simulated SIF and reflectance

Radiative transfer simulations of SIF and reflectance provide a possibility to review the basic validity of assumptions made. Furthermore, a departure of observations with respect to the simulated data may point to specific characteristics and problems.

Simulations of incoming PAR, top of canopy (TOC) reflectance, and TOC SIF between 650 and 850 nm have been computed by the spatially explicit three dimensional plant canopy radiative transfer model (FLiES, Kobayashi and Iwabuchi, 2008), capable of simulating the multiple scattering within the plant canopy. We adopted the boundary conditions from the sun-sensor geometry of GOME-2 (Sect. 4.2.7) in combination with monthly averages of the LAI data set for 2014. Specifically, monthly averages of the

solar zenith angle represent a simplified annual cycle of illumination, whereas all discrete viewing zenith angles as well as averages of the two separate regimes of the relative azimuth angle represent the observation geometry. The LAI ranges from 4.02 to 5.39 units, incorporating a changing canopy structure. Further, a realistic tropical forest landscape has been simulated by the empirical forest structure dynamics model (Yang et al., 2016, in preparation). The forest simulation includes a tree density of 2398 trees ha^{-1}, a canopy height of 5 to 50 m following a Weibull distribution, and a 99% crown cover. The parameters of the forest structure dynamics model were determined based on the allometric relationships obtained in the tropical forests. Optical parameters consist of a broadleaf type leaf spectral reflectance and transmittance, woody reflectance (medium reflectivity), and soil reflectance (medium reflectivity) data sets. These optical data were compiled from the existing literature and publicly available data sets (Kobayashi, 2015). The atmospheric condition is characterized by a tropical temperature profile and a continental aerosol model with an AOD of 0.2 for a cloud-free sky.

4.3 Sensitivity Analysis

4.3.1 Sensitivity to Cloud Effects

Even though the impact of clouds on computed GOME-2 vegetation parameters can be limited by means of the FRESCO cloud fraction data attached to GOME-2 measurements, residual cloud contamination (i.e., undetected sub-pixel clouds) might be unavoidable due to the frequent cloud cover in tropical rainforest areas in combination with the large GOME-2 footprint size. It is therefore crucial to examine potential cloud effects before a possible seasonality can be identified.

Here, simulated TOA radiance spectra as described in Sect. 4.2.4 have been used by incorporating different levels of cloud optical thickness (COT) values. It might be noted that the simulations are consistent with those used by Guanter et al. (2015). Fig. 4.4 illustrates the sensitivity of SIF, NDVI, ρ_{NIR}, DASF and CSC to COT values ranging from 0–3, corresponding to optically thin cirrus or scattered cumulus clouds. Surprisingly, CSC appears to be highly independent of atmospheric effects (+3% at a COT of 3) because the dependencies of DASF and ρ_{NIR} on cloud contamination tend to compensate by rationing. SIF appears to be only moderately affected by cloud contamination, which basically reflects previous findings of decreasing SIF values through the shielding effect by clouds (Frankenberg et al., 2012; Guanter et al., 2015). On the contrary, considerably decreasing NDVI values with an increasing COT may complicate the further evaluation. This strong effect can essentially be explained by the impact of clouds on both red and NIR wavelengths in addition to the shielding of the green vegetation.

Overall, it should be mentioned that the unknown degree of undetected cloud contamination of GOME-2 measurements represents the most limiting factor of this analysis.

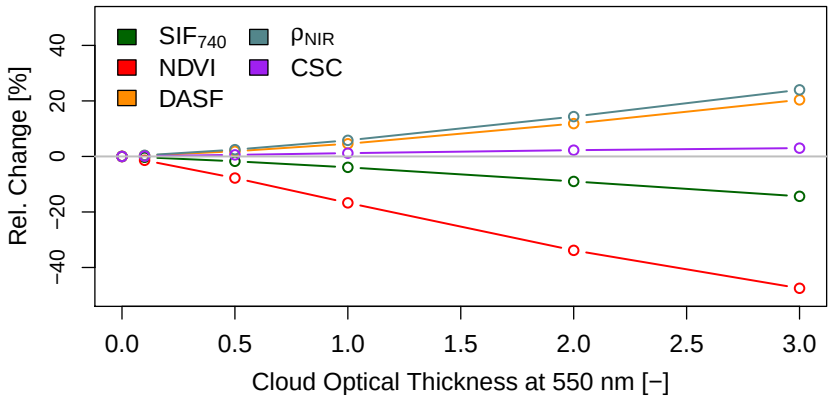


Figure 4.4: Simulated relative intensity changes of SIF, NDVI, ρ_{NIR}, DASF, and CSC as a function of cloud optical thickness (COT) for optically thin clouds.

4.3.2 Sensitivity of DASF, CSC & ω_{NIR} to LAI

The question may arise to which extent DASF and CSC are influenced by canopy structure (through the sensitivity of p to LAI, Eq. 4.10) under the conditions in the Amazon rainforest. Furthermore, the basic validity of our approach to extract leaf properties in terms of ω_{NIR} from CSC (Eq. 4.11) needs to be evaluated. It is therefore essential that the simulations (Sect. 4.2.9) allow us to review these aspects.

The simulations include soil absorption and are driven with constant leaf properties, while LAI and sun-sensor geometry vary similarly to real GOME-2 observations. Given that DASF is an estimate for the non-absorbing canopy's reflectance, the DASF should be consistently higher than the NIR reflectance (ρ_{NIR}) to support the basic applicability of the p theory on our simulations. A similar seasonality of CSC and LAI would then indicate the expectable magnitude of a response of CSC to LAI. Finally, the constant leaf properties in our simulation should lead to a seasonally flat ω_{NIR} to provide evidence for the basic capability to extract leaf properties through Eq. (4.11). Here, the LAI is exactly known as it is an input parameter, however, it should already be considered that this will not apply for real satellite data. Uncertainties in LAI will propagate into ω_{NIR} estimates, complicating the interpretation in case of real data. This approach might therefore be understood as proof of concept.

The first panel of Fig. 4.5 reveals that the DASF is indeed higher than ρ_{NIR}, consistent with the underlying theory. NIR reflectance/DASF intensity changes are highly correlated to the LAI input values ($R = 0.97/0.98$) and amount up to 8.3%/10.5% (from March to September). The second panel of Fig. 4.5 indicates that the CSC is sensitive to changes in LAI ($R = -0.88$), even though this sensitivity is rather low, i.e., a 25% increase in LAI from March to September translates into a 2.9% decrease in CSC. The low sensitivity can be explained by the asymptotic behavior of the recollision probability which reflects the saturation effect in dense vegetation canopies. A significant impact of

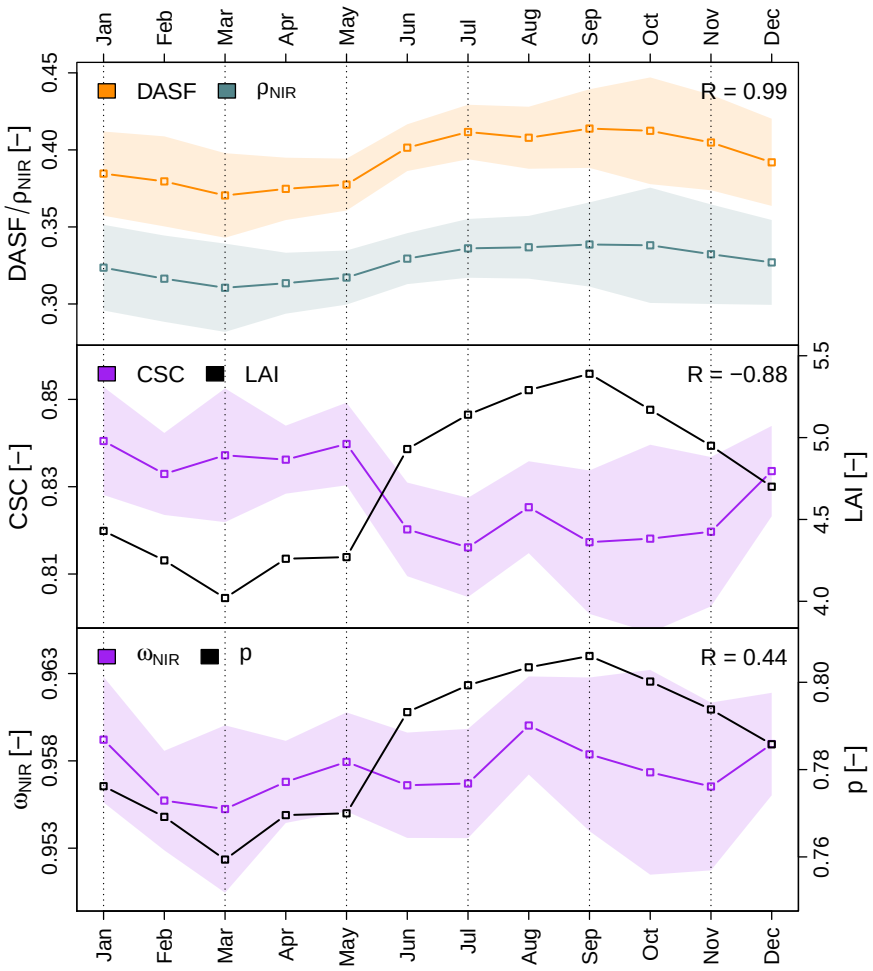


Figure 4.5: ρ_{NIR} (upper panel, blue), DASF (upper panel, orange), CSC (middle panel, purple) and ω_{NIR} (bottom panel, purple) computed from simulated data for a realistic tropical forest observed under a simplified GOME-2 sun-sensor geometry. The standard deviations (shaded areas) result from two relative azimuth angles and 14 viewing zenith angles (value depends on month, see Fig. 4.3). The LAI input value (middle panel, black) corresponds to the monthly average of the study domain for 2014. In the bottom panel, the canopy p value (black) is computed from the LAI input through Eq. (4.10), while ω_{NIR} (purple) is estimated from CSC and p through Eq. (4.11).

viewing zenith angle variations on absolute CSC values has not been found, but the standard deviation increases with viewing zenith angles. The third panel of Fig. 4.5 reveals a 5.8% increase of the canopy p value from March to September (computed by Eq. 4.10). Our strategy to extract leaf properties from CSC (Eq. 4.11) results in an average NIR leaf

albedo of 0.957 with a standard deviation of 0.001 albedo units, while intensity changes are below 0.5%. The basic validity of this approach might thus be confirmed because the simulations own constant leaf properties.

4.4 Results

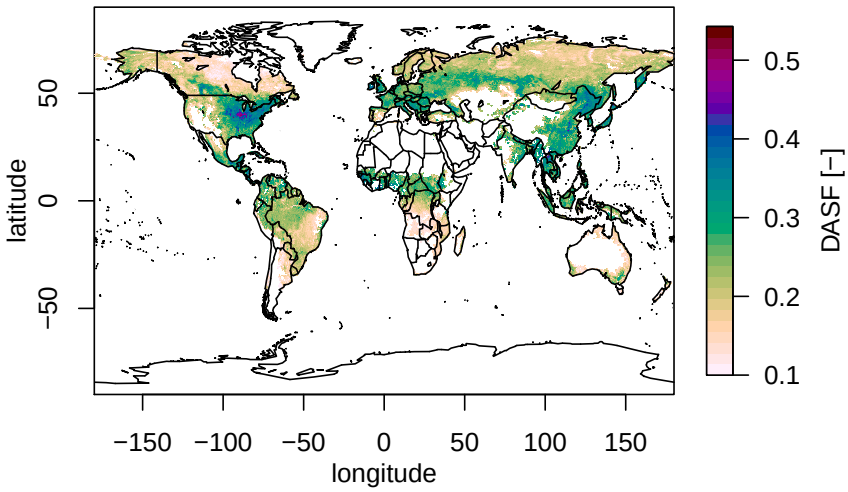
4.4.1 Spatiotemporal Composites

We performed global DASF retrievals from GOME-2 soundings with relative cloud fractions (from FRESCO, Wang et al., 2008) below 0.5 for the 08/2013-04/2015 time period. The processed data can be retrieved from `http://is.gd/GOME_2_DASF`. As described in Sect. 4.2.5, the R^2 value of the underlying linear fit Eq. (4.7) is a measure of retrieval quality and applicability of assumptions. We have therefore tested different R^2 value thresholds to examine the impact on resulting composites. In general, it appeared that non-vegetated areas and regions with sparse vegetation (e.g., deserts and ice sheets) are masked by a R^2 value above 0.7. Furthermore, it has emerged that a more stringent R^2 filter criterion also results in reduced relative cloud fraction averages:

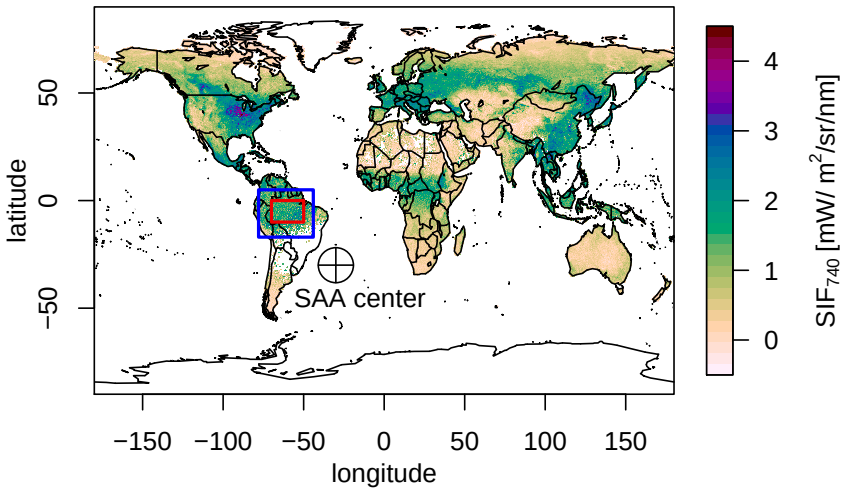
1. $R^2 > 0.7$: 7.4% of remaining soundings own a cloud fraction larger than 0.3
2. $R^2 > 0.9$: 3.5% of remaining soundings own a cloud fraction larger than 0.3

Global monthly composites of DASF(R^2 >= 0.9) and GOME-2 SIF for July 2014 on a 0.2° x 0.2° grid are shown in Fig. 4.6. It is interesting to note the high DASF and SIF average in northern mid-latitudes with peak values in the US Corn Belt region. This peak in SIF and DASF in large cropland and grassland areas might be explained by the rather less complex vegetation structure associated with an enhanced escape probability for scattered/emitted photons. Equally remarkable is that spatial DASF patterns apparently reproduce spatial SIF averages. This characteristic can be attributed to a similar sensitivity to canopy structure and illumination conditions. For example, a complex canopy structure leads to enhanced multiple scattering by which less photons escape regardless whether photons originate from the sun or from SIF.

Even though the ratio ρ_{NIR}/DASF (CSC) appears to be robust against cloud contamination (see Sect. 4.3.1), we based the further analysis on DASF retrievals with a R^2 value above 0.9 in combination with a strict cloud filter (cloud fraction equal zero) to ensure that the impact of clouds is as low as possible. Fig. 4.7 depicts composites of CSC, SIF and NDVI derived from GOME-2 data in the Amazon region for the July–September 2014 period. In addition, the MODIS NDVI (MYD13C1, 2016) composite for the same time period is shown, illustrating the largely consistent spatial patterns as well as consistent absolute values with respect to the GOME-2 NDVI composite. Both composites clearly demonstrate the saturation effect of the NDVI in dense vegetation canopies. The consistency is particularly remarkable against the background of a considerable difference in the spectral and spatial sampling of the satellite instruments. While spectral broadband measurements in channel 1 (red band: 620–670 nm) and channel 2 (NIR band: 841–876 nm) with a fine spatial sampling of 250 m form the basis to compute



(a)



(b)

Figure 4.6: Monthly (a) DASF(R^2 >= 0.9) and (b) SIF composites for July 2014 on a 0.2° x 0.2° grid derived from GOME-2 soundings with relative cloud fractions (from FRESCO, Wang et al., 2008) below 0.5. The center of the SAA at 40°W and 30°S, the extent of the region shown in Fig. 4.1, 4.7 (blue) and our study domain (red) are in addition depicted in map (b).

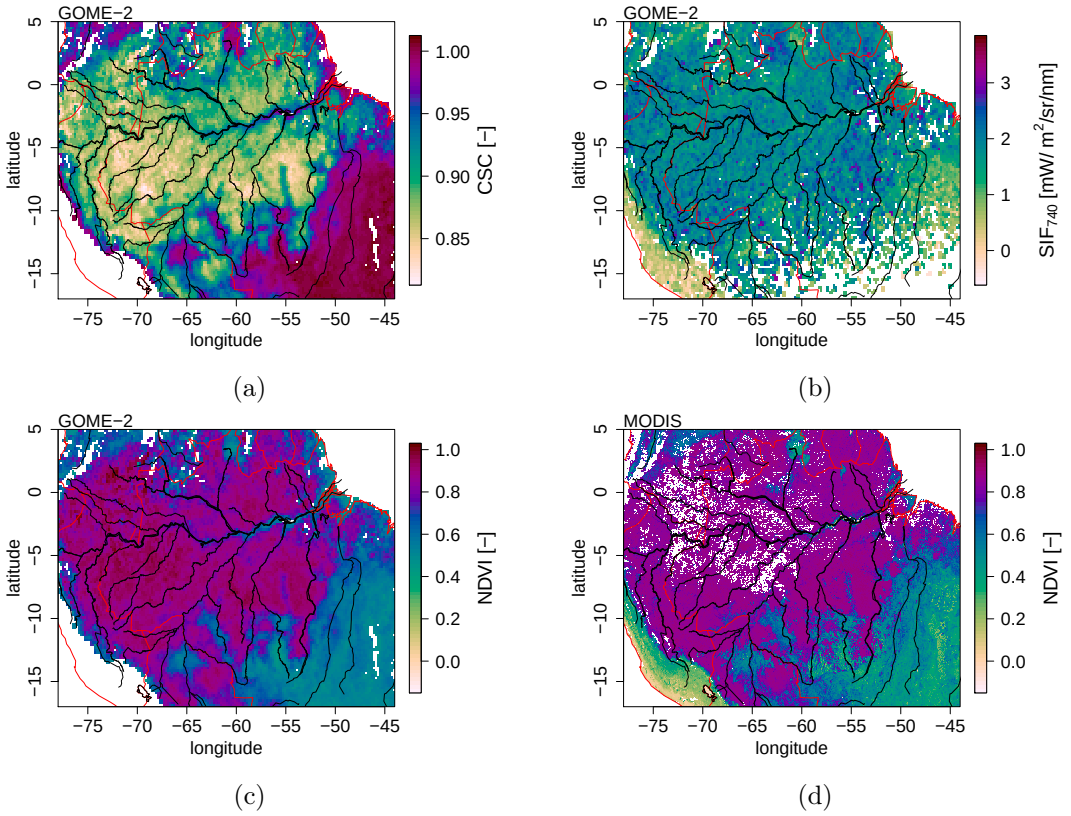


Figure 4.7: Composites of (a) CSC, (b) SIF and (c) NDVI derived from GOME-2 data with relative cloud fractions equal zero for the July–September 2014 period on a 0.2° x 0.2° grid in the Amazon region. Figure (d) shows a MODIS NDVI composite (MYD13C1, 2016, see Sect. 4.2.8) on a 0.05° x 0.05° grid for the same time period. White areas indicate missing values. The spatial sampling of CSC is consistent with that of GOME-2 NDVI, missing values are primarily due to underlying DASF retrievals with $R^2 < 0.9$ (see Sect. 4.2.5). Missing values in the GOME-2 SIF composite are mostly due to our restrictive RSS filter criterion in order to reduce the impact of the SAA (see Sect. 4.2.3).

the NDVI in case of MODIS, we use 2 subsets of the fourth channel from the GOME-2 instrument (660–665 nm, 780–785 nm) with a large ground pixel size of 40 km x 40 km for the same purpose.

In view of the sensitivity of CSC to the canopy p value (Eq. 4.9) and subsequently to the LAI (Sect. 4.3.2), it seems reasonable that grasslands and savannas (see Fig. 4.1 for the regional land cover distribution) reach high, partly saturated CSC levels, while evergreen broadleaf forest regions are clearly separated through lower CSC values. In contrast to the NDVI, the saturation effect appears to be reversed; the CSC saturates in regions covered by less dense vegetation. This supports the CSC as valuable vegetation parameter particularly for dense vegetation canopies such as the Amazon forests. The GOME-2 SIF composite in Fig. 4.7b also reveals a few interesting aspects. First, a spatially highly variable structure can be observed which might either indicate an elevated noise level or may point to a high degree of spatial heterogeneity in the SIF emission. As mentioned in Sect. 4.2.3, we reduce the impact of the SAA on GOME-2 SIF by a restrictive filtering of potentially corrupted retrievals, leading to a concentrated appearance of missing values in south-eastern regions of the shown region. The large extent of the affected region on the South American continent can be seen from the global composite in Fig. 4.6b. Although the SAA represents a limiting factor, remaining grid cells in Fig. 4.7b reveal reasonable SIF averages with respect to the land cover distribution. For example, sparsely vegetated areas at Peru's largely arid Pacific coast and the Andes mountains (south-western sector of the map) show consistently SIF values close to zero.

4.4.2 Comparison between SIF from OCO-2 & GOME-2

This section addresses the question whether SIF from GOME-2 data is consistent with SIF from OCO-2 data. Fig. 4.8 depicts monthly SIF averages from both instruments for our study domain. A few properties and assumptions, which complicate a comparison, are discussed in the following.

First, it should be noted that the spatial coverage is substantially limited since only grid cells which contain SIF retrievals from both OCO-2 and GOME-2 measurements are compared. The cloud filter threshold of GOME-2 SIF retrievals had to be relaxed (maximum cloud fraction is 0.5, mean cloud fraction is 0.11) to ensure that there are overlapping regions in all evaluated months. Köhler et al. (2015b) have shown (for a similar area) that absolute SIF values decrease by relaxing the cloud filter threshold, while the temporal pattern remains almost unaffected. This can also be concluded from Fig. 4.10 in Sect. 4.4.4. Nevertheless, the number of overlapping grid cells per month is rather low and ranges from 36-523, which corresponds to 1%-14% of total grid cells. It might also be considered that absolute SIF values are expected to be different, which is primarily due to different overpass times (GOME-2: ∼9:30 LST, OCO-2: ∼13:30 LST) and evaluated wavelengths (GOME-2: 740 nm, OCO-2: 757 nm). Finally, the GOME-2 SIF data have been extended from May 2015 onwards by monthly averaged data between August 2013 and May 2015, which is due to the lack of complete overlapping data sets in time. This implies the assumption of a repeating cycle every year.

The most obvious difference in the time series lies in the seasonal cycle, which highlights

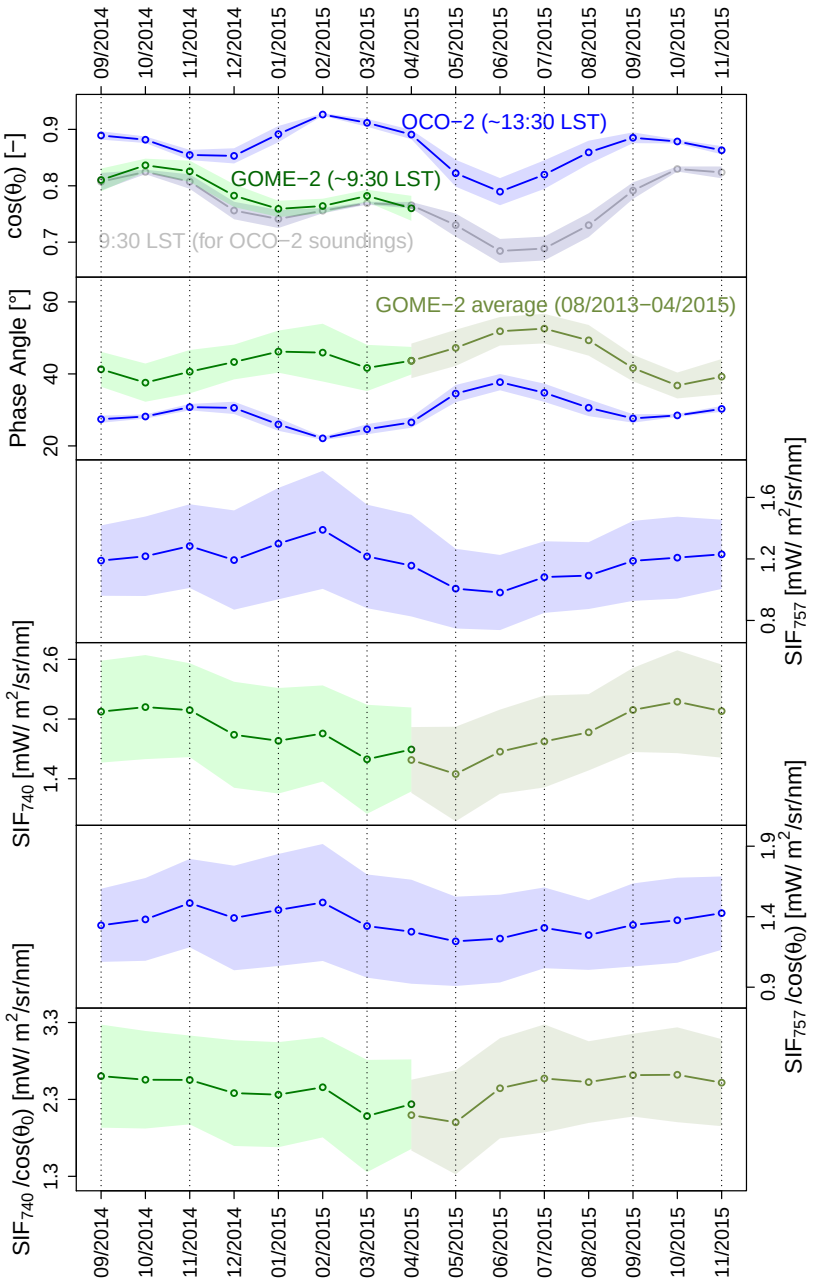


Figure 4.8: Comparison of the cosine of the solar zenith angle (θ_0), phase angle, SIF and SIF/cos(θ_0) from OCO-2 (blue) and GOME-2 (green). Only study domain grid cells which contain SIF retrievals from both instruments form the basis of the shown monthly averages. Monthly long-term averages of GOME-2 are displayed in darkgreen and have been computed due to the lack of complete overlapping data sets in time. Shaded areas indicate the spatial standard deviation.

the instantaneous nature of the SIF signal, i.e., the annual maximum of SIF retrieved from OCO-2 occurs in February, while the maximum SIF obtained from GOME-2 data is observed in October. Both annual peak values coincide with the associated maximum of the cosine of the solar zenith angle, which can be regarded as a proxy for instantaneous PAR (under cloud free conditions), and the minimum of the average phase angle. We included a time series of the cosine of the solar zenith angle for OCO-2 if the overpass time would be the same as for GOME-2 to emphasize the changing annual cycle of illumination with daytime. Considering the relaxed cloud filter in case of SIF from GOME-2 and the maximum of OCO-2 SIF during the wet season, it is remarkable that the large agreement of SIF with instantaneous PAR seems to discard a seasonality being driven by clouds.

Interestingly, a close examination of the GOME-2 time series shows that SIF roughly follows the course of illumination with two exceptions: (i) the slight increase in February 2015 followed by a drop in March 2015 and (ii) the increase during the June to August time period. The latter feature coincides with the period of less intense precipitation (dry season), which ranges typically from June to October. This feature might therefore indicate an increase in photosynthetic activity during the dry season, however, the September increase could again be explained by illumination. On the contrary, the OCO-2 SIF time series seems to be out of the illumination phase only in October and November 2014/2015 where SIF values slightly increase, while instantaneous PAR declines (proxied by the cosine of the solar zenith angle), resulting in a side peak in November.

In order to remove the influence of the instantaneous illumination angle, it might be appropriate to normalize the retrieved SIF values by a correction factor, e.g., the cosine of the solar zenith angle (Joiner et al., 2013). This approach is assumed to extract a possible underlying seasonality (besides the instantaneous PAR driven seasonality) but might only be applicable in areas where the SIF level is relatively high throughout the year. Noise in SIF would otherwise be amplified, in particular for low SIF values in combination with inclined illumination angles. Furthermore, it should be considered that second order characteristics such as the amount of sunlit/shaded leaves are likely to become more pronounced.

The bottom panels of Fig. 4.8 show the normalized time series and reveal that the SIF level is almost maintained during the time period covered, i.e., even the lowest value in May 2015 (for GOME-2 and OCO-2) barely exceeds the spatial standard deviation of the remaining months. A significant seasonality is not visible, but the identified features of both time series are preserved. Note the local maximum in February 2015 which could be linked to reduced precipitation and increased PAR during that particular month. Furthermore, the relatively low GOME-2 SIF/$\cos(\theta_0)$ level from March to May corresponds to months with the most intensive precipitation, while the increase in June is associated with the start of the dry season. This provides evidence for a response of SIF to changing pigment concentrations (green-up). On the contrary, the overall less characteristic seasonality in OCO-2 SIF is apparently not affected by the amount of precipitation. In addition, there is an opposing trend between GOME-2 and OCO-2 SIF during October and November when the difference in illumination and observation geometry is minimized. In view of the above, it might be concluded that the sensitivity

of SIF to environmental conditions varies with the time of day.

Overall, OCO-2 and GOME-2 SIF show consistent seasonal cycles driven by instantaneous PAR, clearly demonstrating the need to consider the time of measurement in addition to the observation geometry.

4.4.3 Directional Effects

The necessity to address the sun-surface-sensor geometry arises from anisotropic reflectance characteristics of natural surfaces, also known as directional effects. It has already been proven that a significant impact on reflectance data, which consequently translates into VIs, complicates the interpretation of satellite remote sensing products, particularly in the Amazon rainforest (Maeda and Galvão, 2015; Hilker et al., 2015). Simulations (van der Tol et al., 2009b) and measurements (Liu et al., 2016) at the TOC level reveal that SIF is also subject to the hot spot effect. Guanter et al. (2012) have shown that even satellite-based SIF measurements from GOSAT data are strongly affected by directional effects. Hence it is reasonable to examine the sensitivity of GOME-2 SIF to view-illumination effects, which has not been considered so far.

A wide range of phase angles from GOME-2 observations within the study domain occurs in October (see Fig. 4.9) and provides the opportunity to estimate and compare intensities and characteristics of directional effects on all retrieved vegetation parameters simultaneously. Although there is a lack of a controlled environment and a certain level of change may be anticipated, we assume that the canopy remains sufficiently invariant during one month and that the study domain is fairly homogeneous to provisionally estimate the directional effects.

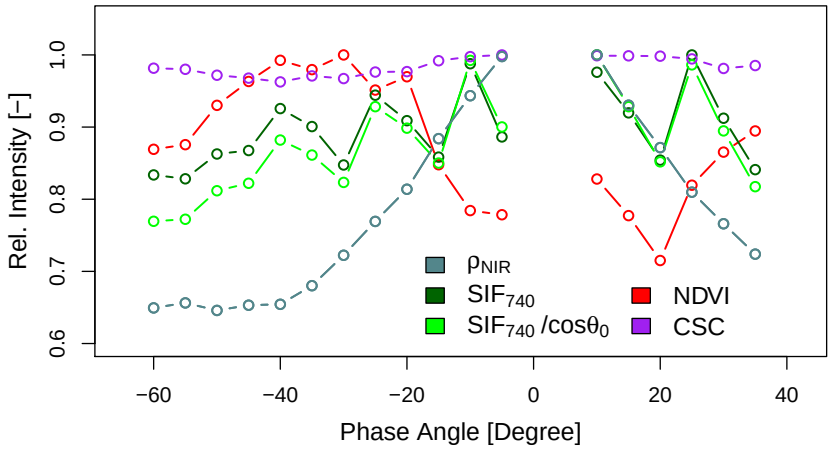


Figure 4.9: Relative intensity of ρ_{NIR}, NDVI, CSC, SIF and normalized SIF as a function of phase angle. Underlying absolute values are averages from GOME-2 in the study domain for October 2014, sampled in 5° phase angle steps.

Fig. 4.9 depicts relative intensities of ρ_{NIR}, NDVI, CSC, SIF and SIF/$\cos(\theta_0)$ as a function of phase angle in October 2014. It has emerged that the FRESCO algorithm preferentially detects higher effective cloud fractions when the phase angle approaches 0°. We therefore relaxed the cloud filter threshold similar to Sect. 4.4.2 (average cloud fraction is 0.11) in order to cover the largest possible phase angle range.

Overall, our results confirm that NIR reflectance is significantly affected by phase angle variations (e.g., Maeda and Galvão, 2015). In particular, the NIR reflectance shape illustrates the hot spot effect, which exhibits the most pronounced directional effect among the tested parameters (intensity varies within 35%). A less clear directionality is visible in the NDVI-phase angle dependency, where values tend to decrease with lower absolute phase angles, while the substantial variation amounts up to 30%. Compared to NIR reflectance and NDVI, SIF seems to be significantly less affected by changing phase angles (intensity changes are well below 20%, change in terms of absolute values amounts up to 0.4 mW m^{-2} sr^{-1} nm^{-1}). Nevertheless, there appears to be a tendency for higher SIF values towards 0° phase angle, which is somewhat similar to the hot spot effect in NIR reflectance, indicating a similar directionality as observed in simulations (van der Tol et al., 2009b) and GOSAT SIF data (Guanter et al., 2012). The actual magnitude of this effect might even be larger because the solar zenith angle increases (from $\sim 31°$ to $\sim 39°$) with decreasing absolute phase angle values (not shown). Accordingly, a decline of SIF values could be expected because of the decrease in instantaneous PAR with inclined illumination angles. However, the opposite behavior is observed, suggesting that the hot spot effect overcompensates the slight decrease in illumination. It should be noted that this effect is even amplified by normalizing SIF with the cosine of the solar zenith angle, resulting in intensity changes slightly above 20% (0.7 mW m^{-2} sr^{-1} nm^{-1}). Seasonal changes in average phase angles will therefore likely affect the SIF time series for the study domain if not accounted for. This finding suggests to carefully address seasonal changes in view-illumination geometry also in further investigations. In contrast to the previously evaluated parameters, it is remarkable that CSC is almost unaffected by phase angle variations as predicted by theory (Sect. 4.2.5). This characteristic combined with the benefit of being highly independent of atmospheric effects (Sect. 4.3.1) makes CSC initially a promising tool for the monitoring of leaf properties in dense canopies. However, the remaining sensitivity to canopy structure as suggested by our sensitivity analysis is a limiting factor regarding the assessment of actual leaf properties.

4.4.4 Large-scale vegetation dynamics

The persistence of recurring phase angles in GOME-2 observations (see Fig. 4.3) allows us to suppress seasonal changes in sun-sensor geometry by limiting the phase angle range. OCO-2 observations, by contrast, are subject to an inseparable seasonality in average phase angle values (not explicitly shown, but essentially consistent with the course shown in the second panel of Fig. 4.8). For this reason, we focus on retrieved vegetation parameters from phase angle limited (between $-65°$ and $-40°$) GOME-2 measurements in this section.

Fig. 4.10 depicts long-term monthly averages from August 2013 to May 2015 of com-

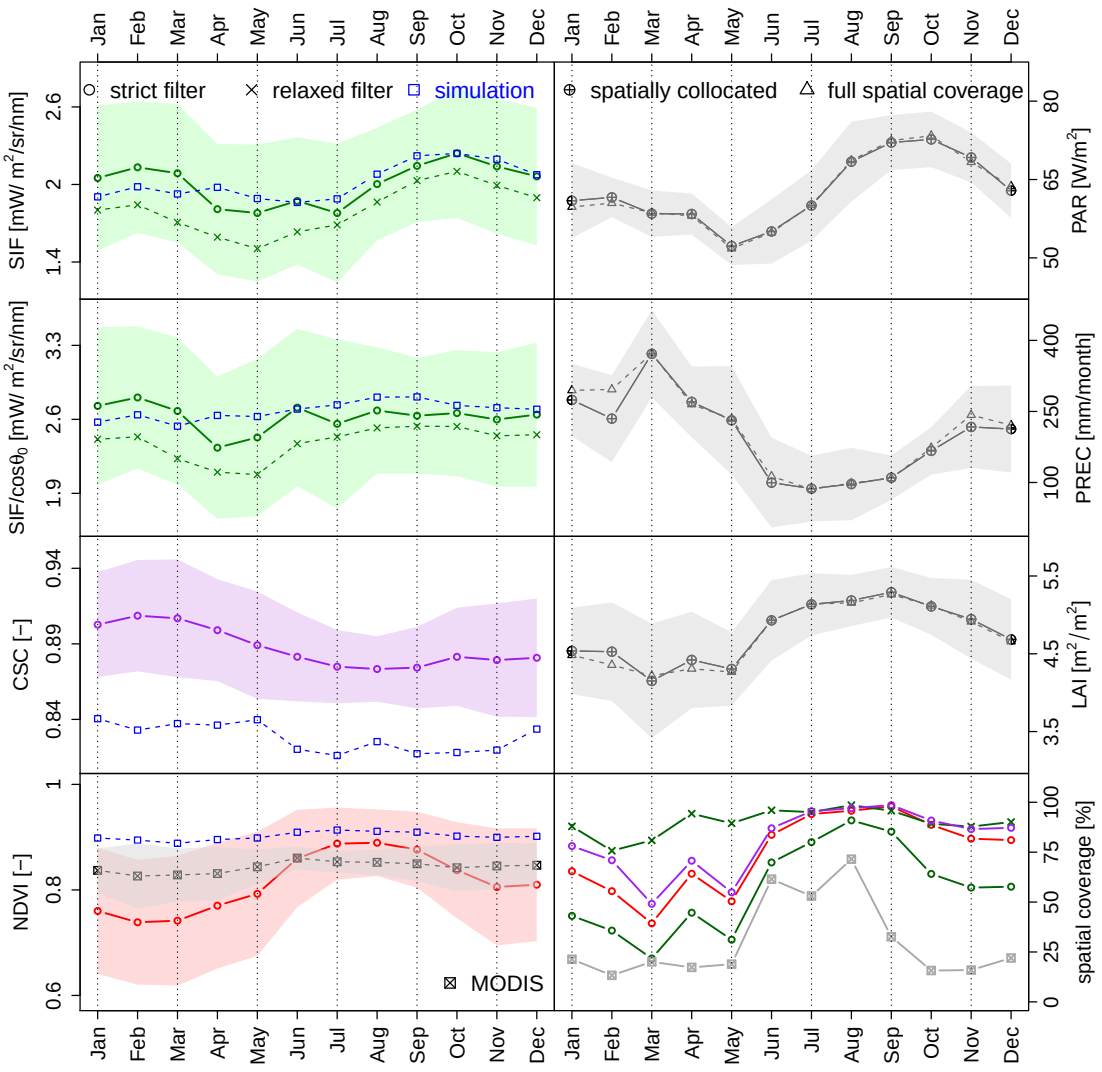


Figure 4.10: Long-term monthly averages (August 2013 – May 2015) of SIF (green), normalized SIF (green), CSC (purple) and NDVI (red) retrieved from GOME-2 data together with simulation results (blue) and MODIS NDVI (MYD13C1, 2016, black). The strict (relaxed) filter indicates that underlying measurements own a cloud fraction equal 0 (below 0.5) and phase angles between $-65°$ and $-40°$ (except for CSC). The simulations of a realistic tropical forest landscape are subject to following boundary conditions: simplified GOME-2 like sun-sensor geometry, cloud free sky, monthly LAI averages for 2014. In addition, averages of PAR (based on 3-hourly data), precipitation and LAI from supplementary data sets are depicted in black, where circles (triangles) indicate a spatial collocation (full spatial coverage) with respect to the spatial coverage of SIF. Spatial standard deviations are depicted by shaded areas.

puted vegetation parameters (SIF, SIF/cos(θ_0), CSC and NDVI) and of environmental conditions (PAR, precipitation, LAI) within our study domain (Sect. 4.2.2). The spatial coverage of each month is also depicted in Fig. 4.10 where differences for the retrieved vegetation parameters arise from individual filter criteria as described in the following.

The strict/relaxed filter (denoted by circles/crosses) constrains all underlying measurements to own a relative cloud fraction equal 0/below 0.5. In addition, SIF (strict filter) and NDVI retrievals are only included if the phase angle ranges between $-65°$ and $-40°$. The phase angle restriction is not necessary for CSC as it can be concluded from Fig. 4.9, however, the number of selected CSC retrievals might be reduced through the quality filter of the underlying DASF computation ($R^2 > 0.9$, see 4.2.5). The remaining SIF retrievals (strict filter) result in the lowest spatial coverage which is mainly due to the restrictive RSS threshold (see Sect. 4.2.3).

Fig. 4.10 reveals that the spatial coverage is substantially reduced during months with intense precipitation. The supplementary data sets have therefore been collocated with valid SIF grid cells (strict filter) to avoid a spatial mismatch, while monthly averages for the entire study domain are nevertheless denoted by triangles. The minor differences between the sparsest (SIF) and the full spatial coverage for the supplementary data sets suggest that the spatial sampling is sufficiently representative for the entire study domain. Nevertheless, it should be noted that the coverage is reduced to about 25% for averages in March and May.

Similar to Fig. 4.8, two annual peaks can be observed in the SIF time series. The maximum in October can be linked to the maximum in PAR, while the lower peak around February is not explained by the shown environmental variables. It seems that the SIF time series is a superposition of monthly averaged PAR (based on 3-hourly data, R=0.75) and an additional seasonality. In view of the comparison between OCO-2 and GOME-2 SIF (Sect. 4.4.2), it is remarkable that the illumination conditions at the overpass time of GOME-2 (9:30 LST) track seasonal variations of instantaneous as well as monthly PAR. A slight discrepancy between simulated and retrieved SIF occurs in the January to May period. At the same time, the spatial coverage is reduced to less than 50% and the spatial standard deviation of the LAI data set is increased. This may point to a departure of simulated conditions with respect to those actually observed: A higher true LAI, for example, could explain the difference. Large-scale leaf age effects (e.g., Wu et al., 2016) could possibly offer an alternative explanation for the observed SIF increase around February. The magnitude of seasonal swings in SIF amounts up to 21% (28% for the relaxed filtering), which is close to the 25% change in green leaf area reported by Myneni et al. (2007).

The normalization of SIF by the cosine of the solar zenith angle emphasizes the peak around February which is also observed in OCO-2 SIF. Similarly, relatively low averages during April and May, which can also be observed in OCO-2 SIF in Fig 4.8, become more pronounced. The seasonal swings are decreased to 17% (18% for the relaxed filter) when removing the influence from instantaneous PAR.

The large-scale seasonality of CSC is less variable (intensity change amounts up to 4%), revealing a peak in February and minimum in August. The CSC time series is

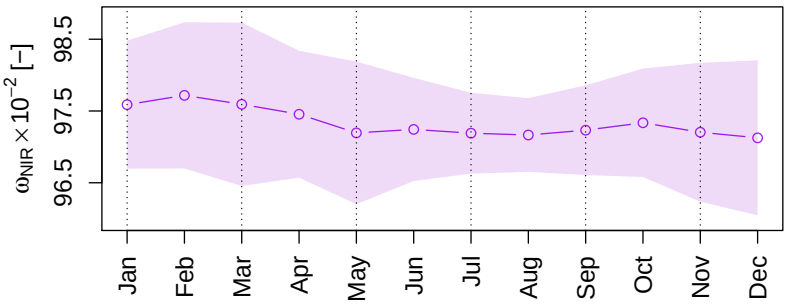


Figure 4.11: Long-term monthly averages (August 2013 – May 2015) of NIR leaf albedo (ω_{NIR}) estimated from monthly GOME-2 CSC averages in combination with monthly aggregated LAI values using Eq. (4.11). The spatial standard deviation is depicted by the shaded area.

highly anti-correlated with LAI (R=-0.87) and equally high positively correlated with the amount of precipitation (R=0.87). Although the sensitivity to LAI and cloud contamination appeared to be small, the low variability of CSC could entirely be controlled by these effects (LAI: 3%, cloud contamination:0-3%). However, particularly the negative correlation with LAI could indicate a green-up during the dry season.

The NDVI simulation results (with constant leaf, wood and soil properties; LAI > 4) and the MODIS NDVI indicate that the sensitivity of NDVI to LAI is almost saturated. At the same time results from our sensitivity analysis (Sect. 4.3.1) and previous investigations (e.g., Kobayashi and Dye, 2005; Hilker et al., 2015) demonstrate the strong sensitivity to atmospheric conditions. It seems therefore likely that the GOME-2 NDVI time series is dominated by undetected sub-pixel clouds, which in turn may point to an insufficient cloud filtering.

We have used monthly aggregated LAI values (per grid cell) to estimate the canopy p value (Eq.4.10) and subsequently the NIR leaf albedo (Eq. 4.11). The resulting large-scale seasonality is depicted in Fig. 4.11, showing only a fractional variability in monthly ω_{NIR} averages. It should be considered that changes are in the same order of magnitude as our simulation results with a constant leaf albedo (0.5%). At the same time, the spatial standard deviation is more pronounced which could indicate a high degree of inhomogeneity. However, our analysis reaches its limits in this case: The residual sensitivity of CSC to canopy structure is removed by LAI estimates with an accuracy of 0.5–0.8 units at a maximum change of one unit. This attempt might therefore be considered as non-conclusive.

4.5 Summary

We have examined two new vegetation parameters, namely SIF and CSC, with respect to their prospects and drawbacks for the monitoring of vegetation dynamics and productivity in the central Amazon Basin. For this purpose, we have used data derived from currently operating satellite sensors as well as simulated data from the MOMO and the FLiES radiative transfer models.

The comparison between OCO-2 and GOME-2 SIF emphasizes the instantaneous nature of SIF measurements, i.e., the annual cycle is closely related to illumination conditions at the overpass time (instantaneous PAR). To this end, OCO-2 and GOME-2 SIF show consistent seasonal cycles. However, we have illustrated that the seasonality of daily averaged illumination is in general not represented by the seasonality of instantaneous illumination. The GOME-2 overpass time around 9:30 LST appears to be an exception, its seasonality is indeed representative for variations in PAR (based on 3-hourly data) on the monthly time scale. A drawback of using SIF in the Amazon region arises from the South Atlantic Anomaly: Increased noise requires a strict filtering and leads to a considerably reduced number of reliable SIF retrievals. Nevertheless, large-scale averages of GOME-2 SIF are largely consistent with our simulations and reveal two distinct contributions to the seasonality of photosynthetic activity. First, the dominant seasonality is associated with variations in PAR with a peak in September/October. Thus, the repeatedly reported dry-season green-up (e.g., Huete et al., 2006; Samanta et al., 2012; Bi et al., 2015) is not rejected by our analysis. Second, a peak around February indicates an increase in photosynthetic activity during parts of the wet season and, consequently, similar wet and dry season SIF averages are observed. This is in agreement with the reported maintained photosynthetic activity if water availability is sufficient (Lee et al., 2013; Guan et al., 2015). Our analysis of phase angle effects provides evidence for a significant impact of the hot spot effect on GOME-2 SIF data.

Contrastingly, CSC appears to be highly independent of sun-sensor geometry as well as atmospheric effects. A major benefit of CSC for the monitoring of dense vegetation lies in the reversed saturation compared to NDVI. However, the slight response to LAI variations limits the informative value regarding the assessment of actual leaf properties. We have found an overall low variability of CSC from GOME-2 data within the central Amazon Basin, showing a consistent seasonality with respect to variations in LAI ('green-up'). Nevertheless, it should be noted that the seasonality of CSC could also be explained by cloud contamination. We have further tested if seasonal changes in NIR leaf albedo can be extracted from CSC by estimating the canopy p value through supplementary LAI data. Although our simulations suggested that this should in principle be possible, uncertainties in LAI weaken the applicability to real satellite data. The resulting minor large-scale variations in NIR leaf albedo might therefore be classified as non-conclusive.

The frequent cloud cover in tropical rainforests in combination with the relatively large pixel size of GOME-2 complicates the analysis because undetected sub-pixel clouds and areas of less dense vegetation are likely to occur. Indeed, the comparison between GOME-2 NDVI and simulated/MODIS NDVI suggests that cloud contamination represents the

most limiting factor of our study, although both SIF and CSC are significantly less affected by fractional cloud-cover than traditional VIs. However, further studies are needed to fully exploit the potential of space-based SIF and CSC observations in the Amazon region. Even though OCO-2 already provides a fine spatial sampling in case of SIF, it is nevertheless crucial that simultaneous and spatially continuous retrievals of both SIF and CSC will also be possible for the next generation of satellite instruments with a similar spectral resolution/coverage as GOME-2, such as the TROPOspheric Monitoring Instrument (TROPOMI; 7 km x 7 km, Veefkind et al., 2012) and the FLuorescence EXplorer (FLEX, 0.3 km x 0.3 km; Rascher et al., 2008).

Acknowledgement

The research is funded by the Emmy Noether Programme of the German Research Foundation (GU 1276/1-1). With thanks to EUMETSAT to make the GOME-2 data available. Jochem Verrelst and Luis Alonso from the University of Valencia are gratefully thanked for the reflectance simulations produced in the framework of the ESA FLUSS project. The OCO-2 SIF data were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center.

Author Contributions

Philipp Köhler designed the study, wrote code, processed the GOME-2 data (including atmospheric correction), analysed the output data, and wrote the manuscript. Luis Guanter gave conceptual advice, provided simulations for Sect. 4.3.1, discussed the results and implications, and commented on the manuscript at all stages. Hideki Kobayashi performed the simulations with the spatially explicit three dimensional plant canopy radiative transfer model (FLiES). Sophia Walther commented on the manuscript at all stages. Wei Yang provided the empirical forest structure dynamics model.

Chapter 5

Conclusions and Outlook

The objectives of this thesis are embedded in the overarching goal of the GlobFluo project of the Deutsche Forschungsgemeinschaft (GU 1276/1-1), which is to assess global terrestrial photosynthesis through the monitoring of SIF from space. More specifically, the presented thesis aims to contribute to the development of state-of-the-art SIF retrieval techniques and to establish a basis for interpreting satellite-based SIF data.

In this chapter, I will first summarize the main outcomes of this thesis with respect to the specific objectives, research questions, and main objectives outlined in Sect. 1.3. A brief review of published research articles, taking advantage of SIF data sets produced within the framework of this thesis, is given in Sect. 5.2. The outlook in Sect. 5.3 comprises future prospects and research recommendations.

5.1 Conclusions

The purpose of this section is to present key results according to the specific research objectives. Subsequently, the overall research questions and main objectives will be addressed.

Specific objectives

Chapter 2 describes a physically-based SIF retrieval method which has been applied to GOSAT-FTS data acquired from June 2009–August 2011.

- Simplifying some of the assumptions in the state-of-the-art retrieval approach from Frankenberg et al. (2011a) without a loss in retrieval accuracy.

Instead of using two micro-fitting windows, only one fitting window (755–759 nm) has been taken into account in the GARLiC algorithm. Furthermore, spectral stretch effects within this narrow fitting window have been neglected. Moreover, the spectral shift between solar reference and measurement spectrum has been estimated prior to solving the forward model. Consequently, the forward model is linear and easy to invert, resulting

in a more efficient algorithm in terms of computation time. A comparison to results from more stringent retrieval approaches (see next point) suggests that the retrieval accuracy is not affected by the simplifications made.

- Comparing results from existing retrieval approaches (Frankenberg et al., 2011b; Guanter et al., 2012) with results from the presented retrieval approach (GARLiC).

A comparison of global annual averages and biome-specific 15-day means revealed an overall good correlation in terms of absolute values as well as temporal trajectories. It has essentially been found that almost all GARLiC values lie in between those from Frankenberg et al. (2011b) and Guanter et al. (2012). This implies that the GARLiC algorithm works well, simultaneously reinforcing confidence on the other data sets derived from GOSAT.

- Evaluating the impact of cloud contamination on SIF time series qualitatively by empirical radiance thresholds.

Based on a strict, normal, and relaxed cloud filter threshold, it has been shown that strongly different levels of cloud contamination do not translate into similar changes in retrieved SIF. This means, cloud contamination barely affects SIF retrievals using Fraunhofer lines in spectral regions devoid of significant atmospheric absorption, confirming the finding from Frankenberg et al. (2012) which was based on simulated data.

Chapter 3 proposes a statistically-based method to retrieve SIF from medium spectral resolution spaceborne spectrometers.

- Providing a solution for the arbitrary selection of free model parameters in the retrieval approaches from Guanter et al. (2013) and Joiner et al. (2013).

The presented approach combines a third-order polynomial in wavelength with atmospheric PCs to model low and high frequency components of the TOA radiance spectrum in a linear way. Consequently, the linear forward model allows to apply a backward elimination algorithm, selecting the required model parameters automatically with respect to the goodness of fit balanced by model complexity.

- Evaluating the retrieval performance through a simulation-based retrieval test.

The retrieval algorithm was extensively tested to (i) assess the retrieval precision and accuracy, (ii) determine the optimum number of provided PCs, (iii) evaluate the influence of the backward elimination algorithm, and (iv) optimize the fitting window. In essence, it has been found that (i) input and retrieved SIF values agree very well, (ii) the described retrieval approach suggests using significantly less PCs as opposed to Joiner et al. (2013, 2014), (iii) noise is reduced by selecting only appropriate model parameters (overfitting is avoided), positively affecting retrieval accuracy and precision, and (iv) a fitting window ranging from 720–758 nm provides the most accurate results (regarding the tested scenarios).

- Applying the same retrieval to data acquired from different sensors (GOME-2 and SCIAMACHY).

In principle, the same retrieval method can be employed to retrieve SIF from both GOME-2 and SCIAMACHY data. However, it has emerged that the retrieval algorithm selects on average twice as many atmospheric PCs when using SCIAMACHY data (6 and 14 PCs, respectively). It is therefore crucial that the backward elimination algorithm ensures stable results, regardless how many PCs are initially provided to the retrieval. The entire time series of SCIAMACHY L1B data from August 2002–March 2012 has been processed, whereas the retrieved GOME-2 SIF covers the period from January 2007–April 2015 (until December 2011 at the time of publishing the research article). It might further be noted that the SCIAMACHY fluorescence retrievals in the 740 nm spectral region are the first ones in the literature.

- Comparing SIF data from different sensors and retrieval approaches with respect to their consistency (GOSAT, GOME-2, and SCIAMACHY).

First, a very good agreement has been found between SCIAMACHY and GOME-2 SIF derived with the method described. Absolute SIF values obtained from SCIAMACHY are slightly lower than those from GOME-2, which may be due to a higher cloud contamination of the larger SCIAMACHY footprints. Second, a comparison with NASA GOME-2 SIF data provided by Joiner et al. (2014) reveals a substantial difference in absolute values between the retrieval approaches, while spatial patterns compare well. Specifically, SIF values from the presented retrieval are about twice as large as those reported by Joiner et al. (2014). In this context, it should be noted that a significant decrease in absolute SIF values in successive versions of the NASA SIF data sets (Joiner et al., 2013, 2014) occurred. Furthermore, a comparison of the NASA GOME-2 SIF (Joiner et al., 2014) and the produced GOME-2/SCIAMACHY SIF data with the GOSAT SIF data set from chapter 2 shows an enhanced correlation for SIF data derived with the presented approach. In summary, SIF data sets are highly consistent concerning spatiotemporal patterns, while uncertainties reside in absolute values.

- Identifying possible limitations and shortcomings of the retrieval results.

Besides the uncertainty about absolute SIF values, an increased retrieval noise in large parts of the South American continent due to the SAA is one of the most noteworthy issues to consider. Moreover, a slight time- and latitude-dependent bias has been observed in the training data set (where SIF values should be zero) when using GOME-2 data. It has been hypothesized that this bias is related to an instrumental issue (e.g., a temperature-dependent dark current change).

- Evaluating the impact of cloud cover on SIF time series by using an additional data set of relative cloud fractions.

Based upon a time series of GOME-2 SIF retrievals in the Amazon Basin, the retrieval appears to be only moderately affected by cloud contamination. SIF values successively

decrease by relaxing the cloud filter threshold, but the seasonality is maintained. This is in agreement with the finding of Frankenberg et al. (2012), Guanter et al. (2015), and Köhler et al. (2015a) (here chapter 2).

Chapter 4 assesses the potential of SIF and CSC for tracking vegetation dynamics in Amazon forests. It might be noted that the first global data set of CSC (derived along with DASF) is presented within this chapter. Time series of GOME-2 SIF and CSC are examined for the August 2013–April 2015 period to take advantage of the reduced ground pixel size (see Sect. 1.2.2).

- Analyzing the sensitivity of SIF and CSC to atmospheric, structural and view-illumination (directional) effects under the conditions of Amazon forests with respect to the well established NDVI.

Simulations consistent with those used by Guanter et al. (2015) and GOME-2 SIF time series reveal that SIF is only moderately affected by cloud contamination, thus confirming results presented by Frankenberg et al. (2012), Köhler et al. (2015a) (here chapter 2), and Köhler et al. (2015b) (here chapter 3). A close examination of the GOME-2 SIF time series might indicate a response of SIF to changing pigment concentrations during periods with less intense precipitation (June–October 'green-up'). Moreover, the hot spot effect is found to significantly impact GOME-2 SIF data. CSC appears to be highly independent of atmospheric effects. The reason lies in similar sensitivities of NIR reflectance (ρ_{NIR}) and DASF to cloud contamination which tend to compensate by rationing (ρ_{NIR}/DASF=CSC). This principle also leads to a compensation of the hot spot effect, making CSC less prone to interference from seasonal changes in sun-sensor geometry. Furthermore, CSC exhibits a low sensitivity to structural effects, i.e., changes in leaf area translate into changes in CSC with about one-tenth of their intensity. In summary, both SIF and CSC are substantially less affected by cloud contamination and sun-sensor geometry than NDVI, which is essentially saturated under the conditions of the Amazon forests.

- Comparing SIF data from different sensors (GOME-2, OCO-2) with particular emphasis on differences due to their overpass times (morning and noon, respectively).

SIF time series derived with data from GOME-2 (Köhler et al., 2015b, here chapter 3) and OCO-2 (Frankenberg et al., 2014) show consistent seasonal cycles with respect to the response of SIF to instantaneous PAR. However, temporal trajectories of instantaneous PAR in the morning and at midday are different, i.e., annual peak values occur in September/October in case of GOME-2 (overpass time in the morning) and in February in case of OCO-2 (overpass time at noon). Even though second order characteristics such as the amount of sunlit/shaded leaves are likely to become more pronounced when SIF is normalized by the cosine of the solar zenith angle (as proxy for instantaneous PAR), a peak in February occurs in normalized GOME-2 and OCO-2 SIF time series.

- Examining prospects of CSC for the monitoring of dense vegetation in general and of leaf optical properties in particular.

Three properties emphasize the beneficial value of CSC for the monitoring of dense vegetation: (i) A low sensitivity to cloud contamination, (ii) insensitivity to sun-sensor geometry (hot spot effect), and (iii) a reversed saturation effect with respect to the NDVI. The residual sensitivity to canopy structure effects (changes in leaf area) is a limiting factor regarding the assessment of leaf optical properties. Therefore, a strategy to estimate the NIR leaf albedo from CSC has been proposed and tested. Although results were encouraging for simulated data, the application to real satellite data is weakened due to uncertainties in LAI estimates.

- Investigating large-scale vegetation dynamics represented by SIF and CSC in Amazon forests with respect to environmental conditions.

Large seasonal swings of GOME-2 SIF (changes amount up to 21%), with peak values in October and February, indicate two distinct contributions to the seasonality of photosynthetic activity in Amazon forests. Specifically, the dominant peak in October can be linked to the maximum in PAR, while the lower peak around February is not clearly attributable to environmental drivers (such as PAR, precipitation, and LAI). Therefore, leaf-age effects might explain this peak. CSC may capture changes in leaf optical properties and thus leaf-age effects, assuming that an increased absorption in NIR wavelength occurs in matured leaves (according to Roberts et al., 1998). However, the derived seasonality of CSC is only slightly variable (changes amount up to 4%) and might be attributed to variations in leaf area and/or the amount of precipitation. That means the residual sensitivity of CSC to canopy structure (see previous point) might control the observed seasonality. Consequently, the 'green-up' hypothesis would be supported, but it has to be mentioned that cloud contamination could also explain the observed seasonality. In particular, the comparison between simulated/MODIS NDVI and GOME-2 NDVI points to an undetected cloud contamination of GOME-2 measurements, representing a limiting factor for the analysis in chapter 4.

Research questions

- How can SIF retrieval methods be streamlined in order to improve their efficiency and robustness?

Chapter 2 shows that simplifying some of the assumptions of an existing physically-based retrieval approach is not accompanied by a loss in retrieval accuracy. Therefore, the efficiency in terms of computation time could be improved. A solution for the arbitrary selection of free model parameters in statistically-based retrieval methods is provided in chapter 3. Consequently, a simulation-based retrieval test revealed an enhanced precision, reduced retrieval noise, and an minimized risk of overfitting. These aspects are expected to be of particular relevance when using real satellite measurements. Thus, the proposed method improves the robustness of statistically-based SIF retrievals.

- How do results from different retrieval approaches compare with each other and how consistent are different data sets?

The different satellite data sets provide a valuable tool for consistency checks, even though a global validation of spaceborne SIF observations with a ground truth still remains to be done. In general, all SIF data sets (GOSAT, GOME-2, SCIAMACHY, and OCO-2) compare well in terms of their spatiotemporal patterns. However, discrepancies arise when considering absolute values. In this context, it should be emphasized that presently, a spatiotemporal consistency is of much more interest since absolute values are not applied in current research. When comparing seasonal cycles, special attention should be paid to the satellite overpass time. This is due to the response of SIF to instantaneous PAR. For instance, the comparison between GOME-2 and OCO-2 in chapter 4 clearly demonstrates that the seasonality of SIF in Amazon forests is primarily driven by instantaneous illumination conditions, which in turn differ with the considered daytime.

- Which are the main confounding factors and pitfalls when analyzing SIF data?

First and foremost, one needs to bear in mind that canopy structure inevitably impacts SIF measurements. For instance, a complex vegetation structure (e.g., rainforests) leads to an enhanced re-absorption of SIF within the canopy, decreasing the detectable amount of SIF. Structural composition effects can be taken into account through a partitioning of retrieval results with respect to the land cover classification and LAI as done in chapter 4. Second, the instantaneous nature of the SIF signal (response to PAR, see previous research question) needs to be considered. This includes to account for possible variations in the satellite overpass time due to orbit drifts/maneuvers when analyzing trends. One possibility to remove the influence of instantaneous PAR consists of normalizing SIF by a correction factor such as the cosine of the solar zenith angle. However, it should be noted that retrieval noise will be amplified in case of low SIF values and inclined illumination angles. Simultaneously, second order characteristics (e.g., the amount of sunlit/shaded leaves and the leaf angle distribution) become increasingly important. Third, directional effects can alter the observed seasonality in SIF if not accounted for. Therefore, the sun-sensor geometry should be carefully addressed; for instance, by evaluating the seasonality with respect to the observational phase angle. Furthermore, a significant issue related to the retrieval of SIF from space is an increased noise in large parts of the South American continent due to the SAA, addressed in chapter 3. Hence, care should be taken when evaluating SIF retrievals from affected areas. This issue can be addressed by employing a restrictive residual check to remove potentially corrupted SIF retrievals (as in chapter 4). Lastly, it is worth noting that results presented by Frankenberg et al. (2012), Guanter et al. (2015), and chapters 2–4 suggest that atmospheric effects due to cloud contamination appear to be, unlike traditional reflectance-based products, of minor importance for the retrieval of SIF from spaceborne instruments.

- Can the concept of canopy spectral invariants (underlying theory to compute DASF and consequently CSC) be utilized to add information regarding structural effects on SIF data?

Given the intrinsic influence of canopy structure, combining SIF with appropriate vegetation parameters, ideally retrievable from the same sensor and explicitly accounting for

structural effects, may be one key aspect to relate SIF to canopy photosynthesis without having to rely on biome dependent empirical relationships. Here, the concept of canopy spectral invariants, involving DASF and CSC, has been identified as a promising candidate for this purpose when observing densely vegetated areas by means of data from medium spectral resolution spectrometers such as GOME-2. Yet, the experience with satellite-based DASF/CSC data is very limited. This applies also for SIF time series concerning the impact of satellite overpass times and directional effects. Therefore, chapter 4 initially assessed prospects and drawbacks of SIF and DASF/CSC for the monitoring of dense vegetation in Amazon forests, explicitly addressing relevant influencing factors. However, a definitive answer to the research question is difficult, mostly due to the frequent cloud cover in tropical regions in combination with the relatively large footprint size of GOME-2 soundings. Hence, it must be stated that more research is needed to fully exploit the potential of SIF and DASF/CSC with respect to the research question posed.

Main objectives

- Developing new methods for the global retrieval of SIF.

This first core objective has been pursued in Köhler et al. (2015a) (here chapter 2) and Köhler et al. (2015b) (here chapter 3). In particular, a physically- and a statistically-based approach to retrieve SIF from spaceborne spectrometers were proposed. As a key result, three new SIF data sets have been produced:

1. GOSAT SIF (June 2009–August 2011),
2. SCIAMACHY SIF (August 2002–March 2012),
3. GOME-2 SIF (January 2007–April 2015).

Fig. 5.1 depicts monthly composites of the different SIF retrievals for July 2010, showing the overall good correspondence of spatial patterns.

- Exemplarily analyzing SIF data with respect to relevant influencing factors in order to establish a basis for interpreting satellite-based SIF data.

Based upon biome specific SIF time series and intercomparisons between the different SIF data sets, this second core objective is essentially addressed throughout the main part of this thesis (chapters 2–4). The results provide evidence for the robustness of satellite-based SIF retrievals against cloud contamination. On the other hand, two aspects have been identified to be of particular relevance which are often disregarded when evaluating the seasonality of photosynthetic activity proxied by SIF:

1. The need to carefully address sun-surface-sensor geometry due to anisotropic SIF emission characteristics.
2. The importance to consider the time of measurement due to the sensitivity of SIF to instantaneous PAR (instantaneous nature of the SIF signal).

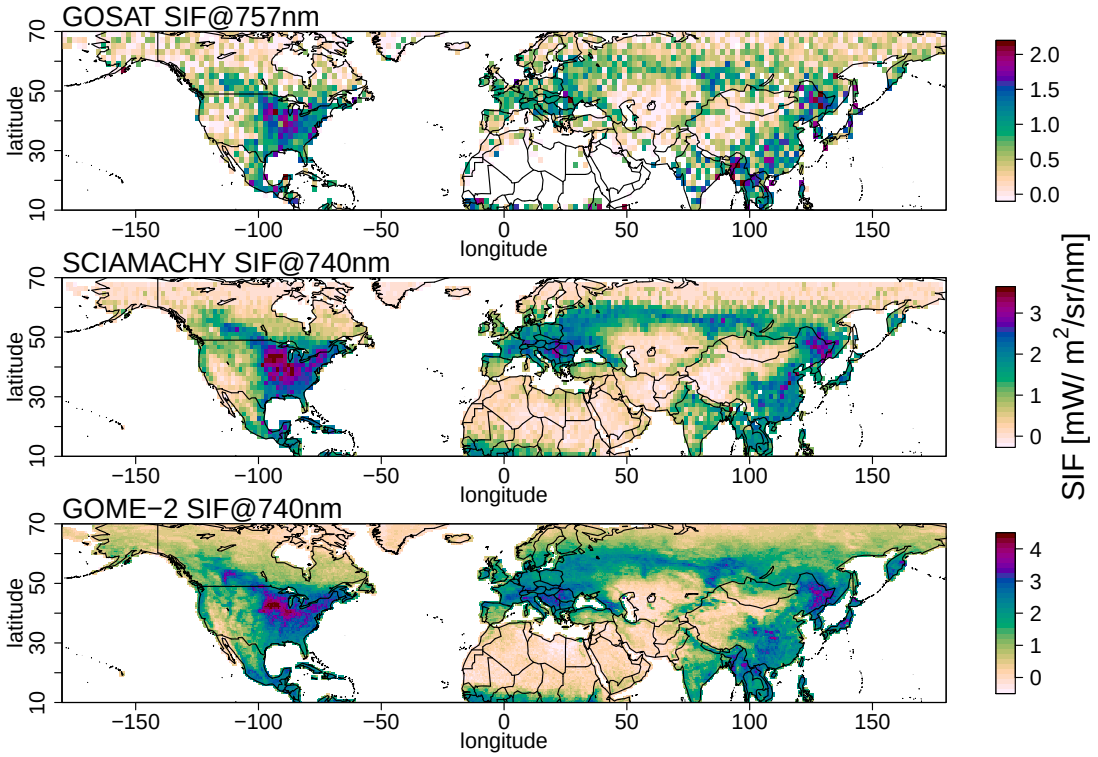


Figure 5.1: Monthly composites of SIF derived from GOSAT (2° x 2° grid, chapter 2), SCIAMACHY (1.5° x 1.5° grid, chapter 3), and GOME-2 (0.5° x 0.5° grid, chapter 3) for July 2010.

5.2 Overview on studies derived from data sets produced in this thesis

The global SIF data sets derived from GOSAT, SCIAMACHY, and GOME-2 in this thesis were made freely available to the user community and are frequently requested by research groups worldwide. Thus, several studies and follow-up publications take advantage of those data sets. Accordingly, a brief review of the published scientific literature is given below.

State-of-the-art XCO_2 (column-averaged dry-air mole fraction of atmospheric carbon dioxide) retrievals typically rely on multiple spectral bands, including the O_2 A-band, to account for atmospheric effects due to scattering in optically thin cirrus clouds and/or aerosol layers as well as surface pressure. Neglecting SIF may lead to biases in retrieved atmospheric parameters when the O_2 A-band is employed, potentially propagating into XCO_2 estimates (Frankenberg et al., 2012). Consequently, SIF may systematically interfere with XCO_2 retrievals using the O_2 A-band (Frankenberg et al., 2012). Reuter et al. (2015) explicitly account for SIF in their SCIAMACHY XCO_2 retrieval approach.

Specifically, the SCIAMACHY SIF product has been implemented in the post-processing of the Bremen Optimal Estimation DOAS[1] (BESD) algorithm for the retrieval of XCO_2 (version v02.01.01, Reuter et al., 2015).

Detmers et al. (2015) quantified the overall increase in carbon uptake in Australia during the La Niña phase from the end of 2010 to early 2012 by means of GOSAT XCO_2 retrievals. The variation of XCO_2 shows a consistent anti-correlation with GOME-2 SIF and soil moisture, justifying a high level of confidence in the results obtained.

Walther et al. (2015) analyzed the start and end of the photosynthetically active season of boreal forests by means of the GOME-2 SIF product. It was essentially found that the active period indicated by SIF is up to six weeks longer than indicated by the EVI (Enhanced Vegetation Index). Their finding supports model results and studies which evaluated in-situ observations.

5.3 Outlook

Given the short period of time since first satellite-based SIF retrievals have been developed, the wealth of outcomes is quite remarkable. However, several limitations need to be addressed, paving the way for future research. One intrinsic limitation of current SIF data sets arises from the rather coarse spatial resolution of global maps (0.5°-2°) which can either be attributed to a large footprint size in case of GOME-2 and SCIAMACHY or to spatially discrete measurements when using GOSAT and OCO-2 SIF products. It is therefore crucial that future satellites missions will contribute to overcome this issue. First, the upcoming TROPOspheric Monitoring Instrument (TROPOMI, Veefkind et al., 2012), payload of the Sentinel-5 Precursor mission, will provide spatially continuous spectroscopic measurements in the 675–775 nm spectral window with a much finer spatial resolution (7 km x 7 km) compared to GOME-2 (40 km x 40 km) and SCIAMACHY (240 km x 30 km). TROPOMI will enable a substantial improvement in SIF monitoring and can reduce uncertainties in mapping SIF by more than a factor of two with respect to GOME-2 (Guanter et al., 2015). Moreover, the FLuorescence EXplorer (FLEX, Rascher et al., 2008) is the first space mission specifically designed to measure the broadband SIF signal. FLEX was recently selected as ESA's Earth Explorer 8 mission, planned to be launched in 2022, and will carry the Fluorescence Imaging Spectrometer (FLORIS). FLORIS will measure spatially continuous with a small footprint of 300 m, outperforming the spatial resolution of GOSAT (10.5 km diameter) and OCO-2 (1.3 km x 2.25 km). However, it should be remarked that the narrow swath width of 150 km increases the revisit time up to 27 days. In view of the upcoming satellite instruments, it can be anticipated that the availability and quality of global SIF data will continue to increase over the next few years. However, it is widely acknowledged that further research is needed to fully exploit the potential of space-based SIF observations as an indicator of photosynthetic activity (e.g., Grace et al., 2007; Baker, 2008; Porcar-Castell et al., 2014). In this context, the following aspects might be particularly relevant:

[1] differential optical absorption spectroscopy (DOAS)

1. Future retrieval approaches should capture the red and NIR peak values of the SIF emission spectrum.

2. Illumination and directional effects (e.g., hot spot effect) must be carefully considered in further investigations (e.g., Damm et al., 2015, chapter 4).

3. Further simulation-based studies are needed to predict the influence of biochemistry, leaf and canopy structure on SIF (scaling from leaf to canopy level).

4. Ground-based validation is needed for a proper characterization of uncertainties and to obtain certainty about absolute SIF values obtained by satellite instruments.

5. Up-scaling approaches are required to fill the gap between canopy and ecosystem levels (Schaepman et al., 2008; Porcar-Castell et al., 2014).

6. An increased temporal resolution would be necessary to assess diurnal cycles of SIF.

While this thesis focuses on the retrieval of the NIR fluorescence peak, recent studies demonstrated the feasibility to also retrieve red fluorescence from space using GOME-2 and SCIAMACHY data (Wolanin et al., 2015; Joiner et al., 2016). A simultaneous retrieval of both red and NIR SIF or a joint data set is therefore the next step towards a better understanding of the link between SIF and terrestrial photosynthesis (point 1). Simultaneously, it is important to increase the efforts in incorporating the canopy structure and view-illumination effects (points 2 and 3). The comparison of ground-based SIF with satellite-based SIF measurements (point 4) will remain complicated due to the discrepancy between the observed scales. However, future validation campaigns will benefit from approved strategies of the Committee on Earth Observing Satellites' Working Group on Calibration and Validation (CEOS WGCV, e.g., Morisette et al., 2006) and the existing infrastructure of established validation networks such as FLUXNET (Baldocchi et al., 2001). Besides the ground-based validation, it will also be worth to compare different satellite data sets as it is done by Joiner et al. (2013), Köhler et al. (2015a,b) (here chapters 2 and 3) as well as chapter 4. In this context, an inter-sensor comparison between TROPOMI and OCO-2 is particularly promising. A possible pathway to close the gap between canopy and ecosystem levels (point 5) would be to up-scale SIF derived from airborne spectrometers to OCO-2 SIF and subsequently to TROPOMI SIF. Near-real time observations of global photosynthetic activity (point 6) would require a constellation of geostationary satellites capable of retrieving SIF. Unfortunately, there are currently no plans for such satellite missions, although a suitable instrumentation was already developed more than one decade ago, namely the Geostationary Imaging Fourier Transform Spectrometer (GIFTS, Smith et al., 2002). It may nevertheless be remarked that the Sentinel-4/UVN (Ultraviolet Visible Near-infrared; Bazalgette Courrèges-Lacoste et al., 2011) instrument on board the geostationary MTG-S (Meteosat Third Generation-Sounder) will enable a continuous monitoring of SIF in a spectral region close to the NIR fluorescence peak (750-775 nm) above Europe and

northern Africa. Summarising the above, it can be stated that there is a lot of research to be done with present and near-future resources.

List of Abbreviations

AOT (or AOD)	aerosol optical thickness
APAR	absorbed photosynthetically active radiation
ATP	adenosine triphosphate
BESD	Bremen Optimal Estimation DOAS algorithm for the retrieval of XCO_2
BIC	Bayesian information criterion
BRF	bidirectional reflectance
CarbonSat	Carbon Monitoring Satellite
CEOS WGCV	Committee on Earth Observing Satellites' Working Group on Calibration and Validation
CF	cloud fraction
CFrank	GOSAT SIF provided by Frankenberg et al. (2011b)
COT	cloud optical thickness
CSC (or W_{NIR})	canopy scattering coefficient
DASF	directional area scattering factor
DOAS	differential optical absorption spectroscopy
ECMWF	European Centre for Medium-Range Weather Forecasts
ENVISAT	ENVIronmental SATellite
ERA	global atmospheric reanalysis produced by the ECMWF
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
EVI	enhanced vegetation index
F685	red fluorescence peak at 685 nm

F740	NIR fluorescence peak at 740 nm
FLD	Fraunhofer line discriminator
FLEX	FLuorescence EXplorer
FLiES	forest light environmental simulator
FLIM	geometric forest light interaction model
FLORIS	Fluorescence Imaging Spectrometer
FlourMODleaf	model of chlorophyll *a* fluorescence emission by plant leaves (extension of the PROSPECT model)
FluorFLIM	combined radiative transfer model (FlourMODleaf, FluorSAIL, and FLIM)
FluorMOD	Project to develop an integrated leaf-canopy model to simulate solar-induced chlorophyll fluorescence in vegetation
FluorSAIL	Scattering by Arbitrarily Inclined Leaves (canopy fluorescence radiative transfer model)
FLUSS	atmospheric corrections for fluorescence signal and surface pressure retrieval over land (ESA project)
FLUXNET	network of regional flux tower sites
fPAR	fractional absorptance of photosynthetically active radiation
FRESCO	Fast Retrieval Scheme for Clouds from the O_2 A-band
FTS	fourier transform spectrometer
FWHM	full width at half maximum
GARLiC	GOSAT retrieval of chlorophyll fluorescence
GIFTS	Geostationary Imaging Fourier Transform Spectrometer
GlobFluo	Project for the global assessment of vegetation photosynthesis through the monitoring of chlorophyll fluorescence from space
GOME-2	Global Ozone Monitoring Experiment-2
GOSAT	Greenhouse Gases Observing Satellite
GPP	gross primary production
IGBP	International Geosphere and Biosphere Programme
JAXA	Japan Aerospace Exploration Agency
LAI	leaf area index
LGuanter	GOSAT SIF provided by Guanter et al. (2012)
LHC	light-harvesting complex
LST	local solar time

MACC	Monitoring Atmospheric Composition and Climate re-analysis data
MERIS	MEdium-spectral Resolution Imaging Spectrometer
MetOp	Meteorological Operational Satellite
MODIS	Moderate Resolution Imaging Spectroradiometer
MOMO	Matrix Operator MOdel
MSL	mean sea level pressure
MTG-S	Meteosat Third Generation-Sounder
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NASA	National Aeronautics and Space Administration
NDVI	normalized difference vegetation index
NIR	near infra-red
OCO-2	Orbiting Carbon Observatory-2
P	parallel polarized light
PAR	photosynthetically active radiation
PC	principal component
PCA	principal component analysis
PROSPECT	a radiative transfer model describing the optical properties of plant leaves from 400 nm to 2500 nm
PSI	photosystem I
PSII	photosystem II
RAA	relative azimuth angle
RSS	residual sum of squares
S	perpendicular polarized light
S_0	ground state of chlorophyy a molecules
S_1/S_2	excited states of chlorophyy a molecules
SAA	South Atlantic Anomaly
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY
SEM	standard error of the mean
SIF (or F_s)	sun-induced chlorophyll fluorescence
SNR	signal-to-noise ratio
SVD	singular value decomposition
SZA	solar zenith angle
TCWV (or WVC)	total water vapor column

TOA	top-of-atmosphere
TOC	top-of-canopy
TRMM	Tropical Rainfall Measuring Mission
TROPOMI	TROPOspheric Monitoring Instrument
UVN instrument	Ultraviolet Visible Near-infrared spectrometer on board MTG-S
V25 GOME-2	GOME-2 SIF provided by Joiner et al. (2014)
VI	vegetation index
VZA	viewing zenith angle
W_λ	spectral canopy scattering coefficient
XCO_2	column-average dry-air mole fraction of atmospheric carbon dioxide

List of Tables

List of Figures

Bibliography

Amorós-López, J., Gomez-Chova, L., Vila-Frances, J., Alonso, L., Calpe, J., Moreno, J., del Valle-Tascon, S., 2008. Evaluation of remote sensing of vegetation fluorescence by the analysis of diurnal cycles. International Journal of Remote Sensing 29 (17-18), 5423–5436.

Baker, N. R., 2008. Chlorophyll fluorescence: a probe of photosynthesis in vivo. Annu. Rev. Plant Biol. 59, 89–113.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., et al., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society 82 (11), 2415–2434.

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., Smets, B., 2013. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. Remote Sensing of Environment 137, 299–309.

Bazalgette Courrèges-Lacoste, G., Ahlers, B., Guldimann, B., Short, A., Veihelmann, B., Stark, H., 2011. The Sentinel-4/UVN instrument on-board MTG-S. EUMETSAT Meteorological Satellite Conference, Oslo, Norway.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., et al., 2010. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. Science 329 (5993), 834–838.

Bi, J., Knyazikhin, Y., Choi, S., Park, T., Barichivich, J., Ciais, P., Fu, R., Ganguly, S., Hall, F., Hilker, T., et al., 2015. Sunlight mediated seasonality in canopy structure and photosynthetic activity of amazonian rainforests. Environmental Research Letters 10 (6), 064014.

Bovensmann, H., Buchwitz, M., Burrows, J., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., Erzinger, J., 2010. A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications. Atmospheric Measurement Techniques 3 (4), 781–811.

Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V., Chance, K., Goede, A., 1999. SCIAMACHY: Mission objectives and measurement modes. Journal of the Atmospheric Sciences 56 (2), 127–150.

Buschmann, C., 2007. Variability and application of the chlorophyll fluorescence emission ratio red/far-red of leaves. Photosynthesis Research 92 (2), 261–271.

Chance, K., Kurucz, R., 2010. 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.

Damm, A., Guanter, L., Verhoef, W., Schläpfer, D., Garbari, S., Schaepman, M., 2015. Impact of varying irradiance on vegetation indices and chlorophyll fluorescence derived from spectroscopy data. Remote Sensing of Environment 156, 202–215.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137 (656), 553–597.

Detmers, R., Hasekamp, O., Aben, I., Houweling, S., Leeuwen, T., Butz, A., Landgraf, J., Köhler, P., Guanter, L., Poulter, B., 2015. Anomalous carbon uptake in Australia as seen by GOSAT. Geophysical Research Letters 42 (19), 8177–8184.

Dixon, J. M., Taniguchi, M., Lindsey, J. S., 2005. PhotochemCAD 2: a refined program with accompanying spectral databases for photochemical calculations. Photochemistry and photobiology 81 (1), 212–213.

Fell, F., Fischer, J., 2001. Numerical simulation of the light field in the atmosphere–ocean system using the matrix-operator method. Journal of Quantitative Spectroscopy and Radiative Transfer 69 (3), 351–388.

Fournier, A., Daumard, F., Champagne, S., Ounis, A., Goulas, Y., Moya, I., 2012. Effect of canopy structure on sun-induced chlorophyll fluorescence. ISPRS Journal of Photogrammetry and Remote Sensing 68, 112–120.

Franck, F., Juneau, P., Popovic, R., 2002. Resolution of the photosystem I and photosystem II contributions to chlorophyll fluorescence of intact leaves at room temperature. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1556 (2), 239–246.

Frankenberg, C., Butz, A., Toon, G., 2011a. Disentangling chlorophyll fluorescence from atmospheric scattering effects in o2 a-band spectra of reflected sun-light. Geophysical Research Letters 38 (3).

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., et al., 2011b. New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. Geophysical Research Letters 38 (17).

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

Frankenberg, C., O'Dell, C., Guanter, L., McDuffie, J., 2012. 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.

Friedl, M. A., McIver, D. K., Hodges, J. C., Zhang, X., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., et al., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83 (1), 287–302.

Gitelson, A. A., Buschmann, C., Lichtenthaler, H. K., 1998. Leaf chlorophyll fluorescence corrected for re-absorption by means of absorption and reflectance measurements. Journal of plant physiology 152 (2), 283–296.

Govaerts, Y., Wagner, S., Lattanzio, A., Watts, P., 2010. Joint retrieval of surface reflectance and aerosol optical depth from MSG/SEVIRI observations with an optimal estimation approach: 1. Theory. Journal of Geophysical Research: Atmospheres 115 (D2).

Govindjee, G., 2004. Chlorophyll a fluorescence: a bit of basics and history. Chlorophyll a fluorescence: a signature of photosynthesis. Springer, Dordrecht, 1–42.

Grace, J., Nichol, C., Disney, M., Lewis, P., Quaife, T., Bowyer, P., 2007. Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence? Global Change Biology 13 (7), 1484–1497.

Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., Caylor, K. K., Sheffield, J., Wood, E. F., Malhi, Y., et al., 2015. Photosynthetic seasonality of global tropical forests constrained by hydroclimate. Nature Geoscience 8 (4), 284–289.

Guanter, L., Aben, I., Tol, P., Krijger, J., Hollstein, A., Köhler, P., Damm, A., Joiner, J., Frankenberg, C., Landgraf, J., 2015. Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence. Atmospheric Measurement Techniques 8 (3), 1337–1352.

Guanter, L., Alonso, L., Gómez-Chova, L., Amorós-López, J., Vila, J., Moreno, J., 2007. Estimation of solar-induced vegetation fluorescence from space measurements. Geophysical Research Letters 34 (8).

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., Grainger, R. G., 2012. Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. Remote Sensing of Environment 121, 236–251.

Guanter, L., Rossini, M., Colombo, R., Meroni, M., Frankenberg, C., Lee, J.-E., Joiner, J., 2013. 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.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J.-E., et al., 2014. Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proceedings of the National Academy of Sciences 111 (14), E1327–E1333.

Hapke, B., 2012. Theory of reflectance and emittance spectroscopy. Cambridge University Press.

Heldt, H.-W., 2004. Plant biochemistry. Academic Press.

Hikosaka, K., Niinemets, Ü., Anten, N. P., 2015. Canopy Photosynthesis: From Basics to Applications.

Hilker, T., Coops, N. C., Wulder, M. A., Black, T. A., Guy, R. D., 2008. The use of remote sensing in light use efficiency based models of gross primary production: A review of current status and future requirements. Science of the Total Environment 404 (2), 411–423.

Hilker, T., Lyapustin, A. I., Hall, F. G., Myneni, R., Knyazikhin, Y., Wang, Y., Tucker, C. J., Sellers, P. J., 2015. On the measurability of change in Amazon vegetation from MODIS. Remote Sensing of Environment 166, 233–242.

Hu, B., Lucht, W., Strahler, A. H., 1999. The interrelationship of atmospheric correction of reflectances and surface BRDF retrieval: A sensitivity study. Geoscience and Remote Sensing, IEEE Transactions on 37 (2), 724–738.

Huang, D., Knyazikhin, Y., Dickinson, R. E., Rautiainen, M., Stenberg, P., Disney, M., Lewis, P., Cescatti, A., Tian, Y., Verhoef, W., et al., 2007. Canopy spectral invariants for remote sensing and model applications. Remote Sensing of Environment 106 (1), 106–122.

Huete, A. R., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutyra, L. R., Yang, W., Nemani, R. R., Myneni, R., 2006. Amazon rainforests green-up with sunlight in dry season. Geophysical Research Letters 33 (6).

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A., Middleton, E., Huemmrich, K., Yoshida, Y., Frankenberg, C., 2013. Global monitoring of terrestrial chlorophyll

fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2. Atmospheric Measurement Techniques 6 (10), 2803–2823.

Joiner, J., Yoshida, Y., Guanter, L., Middleton, E., 2016. New methods for retrieval of chlorophyll red fluorescence from hyper-spectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY. Atmospheric Measurement Techniques Discussions, 1–41.

Joiner, J., Yoshida, Y., Vasilkov, A., Middleton, E., Campbell, P., Kuze, A., Corp, L., 2012. 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 (4), 809–829.

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

Joiner, J., Yoshida, Y., Vasilkov, A., Schaefer, K., Jung, M., Guanter, L., Zhang, Y., Garrity, S., Middleton, E., Huemmrich, K., et al., 2014. The seasonal cycle of satellite chlorophyll fluorescence observations and its relationship to vegetation phenology and ecosystem atmosphere carbon exchange. Remote Sensing of Environment 152, 375–391.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., et al., 2011. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. Journal of Geophysical Research: Biogeosciences 116 (G3).

Kautsky, H., Hirsch, A., 1931. Neue Versuche zur Kohlensäureassimilation. Naturwissenschaften 19 (48), 964–964.

Knyazikhin, Y., Kranigk, J., Myneni, R. B., Panfyorov, O., Gravenhorst, G., 1998. Influence of small-scale structure on radiative transfer and photosynthesis in vegetation canopies. Journal of Geophysical Research: Atmospheres (1984–2012) 103 (D6), 6133–6144.

Knyazikhin, Y., Schull, M. A., Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., Latorre Carmona, P., Kaufmann, R. K., Lewis, P., Disney, M. I., Vanderbilt, V., Davis, A. B., Baret, F., Jacquemoud, S., Lyapustin, A., Myneni, R. B., 2013. Hyperspectral remote sensing of foliar nitrogen content. Proceedings of the National Academy of Sciences 110 (3), E185–E192.

Knyazikhin, Y., Schull, M. A., Xu, L., Myneni, R. B., Samanta, A., 2011. Canopy spectral invariants. Part 1: A new concept in remote sensing of vegetation. Journal of Quantitative Spectroscopy and Radiative Transfer 112 (4), 727–735.

Kobayashi, H., 2015. Leaf, woody and background optical data for GCOM-C LAI/FAPAR retrieval. JAXA GCOM-C RA4 #102 project documents, available from the author upon the request.

Kobayashi, H., Dye, D. G., 2005. Atmospheric conditions for monitoring the long-term vegetation dynamics in the Amazon using normalized difference vegetation index. Remote Sensing of Environment 97 (4), 519–525.

Kobayashi, H., Iwabuchi, H., 2008. A coupled 1-D atmosphere and 3-D canopy radiative transfer model for canopy reflectance, light environment, and photosynthesis simulation in a heterogeneous landscape. Remote Sensing of Environment 112 (1), 173–185.

Köhler, P., 2012. Satellitengestützte Fernerkundung von Chlorophyllfluoreszenz mithilfe von GOSAT-Daten. Master's thesis, Freie Universität Berlin.

Köhler, P., Guanter, L., Frankenberg, C., July 2015a. Simplified Physically Based Retrieval of Sun-Induced Chlorophyll Fluorescence From GOSAT Data. Geoscience and Remote Sensing Letters, IEEE 12 (7), 1446–1450.

Köhler, P., Guanter, L., Joiner, J., 2015b. A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data. Atmospheric Measurement Techniques 8 (6), 2589–2608.

Krause, G., Weis, E., 1991. Chlorophyll fluorescence and photosynthesis: the basics. Annual review of plant biology 42 (1), 313–349.

Kurucz, R. L., 2005. High Resolution Irradiance Spectrum from 300 to 1000 nm. Tech. rep., Harvard-Smithsonian Center for Astrophysics.

Kuze, A., Suto, H., Nakajima, M., Hamazaki, T., 2009. 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.

Lee, J.-E., Frankenberg, C., van der Tol, C., Berry, J. A., Guanter, L., Boyce, C. K., Fisher, J. B., Morrow, E., Worden, J. R., Asefi, S., Badgley, G., Saatchi, S., 2013. Forest productivity and water stress in Amazonia: observations from GOSAT chlorophyll fluorescence. Proceedings of the Royal Society of London B: Biological Sciences 280 (1761).

Lewis, P., Disney, M., 2007. Spectral invariants and scattering across multiple scales from within-leaf to canopy. Remote Sensing of Environment 109 (2), 196–206.

Lichtenberg, G., Kleipool, Q., Krijger, J., Soest, G. v., Hees, R. v., Tilstra, L., Acarreta, J., Aben, I., Ahlers, B., Bovensmann, H., Chance, K., Gloudemans, A. M. S., Hoogeveen, R. W. M., Jongma, R. T. N., Noël, S., Piters, A., Schrijver, H., Schrijvers, C., Sioris, C. E., Skupin, J., Slijkhuis, S., Stammes, P., Wuttke, M., 2006. SCIAMACHY Level 1 data: calibration concept and in-flight calibration. Atmospheric Chemistry and Physics 6 (12), 5347–5367.

Lichtenthaler, H. K., Rinderle, U., 1988. The role of chlorophyll fluorescence in the detection of stress conditions in plants. CRC Critical Reviews in Analytical Chemistry 19 (sup1), S29–S85.

Liu, L., Liu, X., Wang, Z., Zhang, B., 2016. Measurement and Analysis of Bidirectional SIF Emissions in Wheat Canopies. IEEE Transactions on Geoscience and Remote Sensing 54 (5), 2640–2651.

Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., Chang, W., 2004. Crystal structure of spinach major light-harvesting complex at 2.72 å resolution. Nature 428 (6980), 287–292.

MACC, 2016. Monitoring Atmospheric Composition and Climate. Available online: http://www.gmes-atmosphere.eu.

Maeda, E. E., Galvão, L. S., 2015. Sun-sensor geometry effects on vegetation index anomalies in the Amazon rainforest. GIScience & Remote Sensing 52 (3), 332–343.

Maxwell, K., Johnson, G. N., 2000. Chlorophyll fluorescence—a practical guide. Journal of experimental botany 51 (345), 659–668.

Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R., Moreno, J., 2009. Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. Remote Sensing of Environment 113 (10), 2037–2051.

Miller, J., Berger, M., Goulas, Y., Jacquemoud, S., Louis, J., Moise, N., et al., 2005. Development of a vegetation fluorescence canopy model, Final Report, May 2005. European Space Research and Technology Centre (ESTEC).

Monteith, J., 1972. Solar radiation and productivity in tropical ecosystems. Journal of applied ecology 9 (3), 747–766.

Morisette, J. T., Baret, F., Privette, J. L., Myneni, R. B., Nickeson, J. E., Garrigues, S., Shabanov, N. V., Weiss, M., Fernandes, R., Leblanc, S. G., et al., 2006. Validation of global moderate-resolution LAI products: A framework proposed within the CEOS land product validation subgroup. IEEE Transactions on Geoscience and Remote Sensing 44 (7), 1804.

Morton, D. C., Nagol, J., Carabajal, C. C., Rosette, J., Palace, M., Cook, B. D., Vermote, E. F., Harding, D. J., North, P. R., 2014. Amazon forests maintain consistent canopy structure and greenness during the dry season. Nature 506 (7487), 221–224.

Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., Albiñana, A. P., 2006. GOME-2 on MetOp. In: Proc. of The 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland. pp. 12–16.

MYD13C1, 2016. Vegetation Indices 16-Day L3 Global 0.05Deg CMG. Available online: https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13c1.

Myneni, R. B., Yang, W., Nemani, R. R., Huete, A. R., Dickinson, R. E., Knyazikhin, Y., Didan, K., Fu, R., Juárez, R. I. N., Saatchi, S. S., et al., 2007. Large seasonal swings in leaf area of Amazon rainforests. Proceedings of the National Academy of Sciences 104 (12), 4820–4823.

Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D., Cescatti, A., Pérez-Priego, Ó., Wohlfahrt, G., Montagnani, L., 2014. Terrestrial gross primary production inferred from satellite fluorescence and vegetation models. Global change biology 20 (10), 3103–3121.

Parazoo, N. C., Bowman, K., Frankenberg, C., Lee, J.-E., Fisher, J. B., Worden, J., Jones, D., Berry, J., Collatz, G. J., Baker, I. T., et al., 2013. Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of XCO_2 and chlorophyll fluorescence from GOSAT. Geophysical Research Letters 40 (11), 2829–2833.

Pedrós, R., Goulas, Y., Jacquemoud, S., Louis, J., Moya, I., 2010. FluorMODleaf: A new leaf fluorescence emission model based on the PROSPECT model. Remote Sensing of Environment 114 (1), 155–167.

Pedrós, R., Moya, I., Goulas, Y., Jacquemoud, S., 2008. Chlorophyll fluorescence emission spectrum inside a leaf. Photochemical & photobiological sciences 7 (4), 498–502.

Perez-Priego, O., Guan, J., Rossini, M., Fava, F., Wutzler, T., Moreno, G., Carvalhais, N., Carrara, A., Kolle, O., Julitta, T., Schrumpf, M., Reichstein, M., Migliavacca, M., 2015. Sun-induced chlorophyll fluorescence and photochemical reflectance index improve remote-sensing gross primary production estimates under varying nutrient availability in a typical Mediterranean savanna ecosystem. Biogeosciences 12 (21), 6351–6367.

Plascyk, J. A., 1975. The MK II Fraunhofer line discriminator (FLD-II) for airborne and orbital remote sensing of solar-stimulated luminescence. Optical Engineering 14 (4), 339–0.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., Berry, J. A., 2014. Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. Journal of Experimental Botany.

Rascher, U., Gioli, B., Miglietta, F., 2008. FLEX—fluorescence explorer: a remote sensing approach to quantify spatio-temporal variations of photosynthetic efficiency from space. In: Photosynthesis. Energy from the Sun. Springer, pp. 1387–1390.

Restrepo-Coupe, N., da Rocha, H. R., Hutyra, L. R., da Araujo, A. C., Borma, L. S., Christoffersen, B., Cabral, O. M., de Camargo, P. B., Cardoso, F. L., da Costa, A. C. L., Fitzjarrald, D. R., Goulden, M. L., Kruijt, B., Maia, J. M., Malhi, Y. S., Manzi, A. O., Miller, S. D., Nobre, A. D., von Randow, C., Sá, L. D. A., Sakai, R. K., Tota,

J., Wofsy, S. C., Zanchi, F. B., Saleska, S. R., 2013. What drives the seasonality of photosynthesis across the Amazon basin? A cross-site analysis of eddy flux tower measurements from the Brasil flux network. Agricultural and Forest Meteorology 182–183, 128 – 144.

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P. Heymann, J., Hilker, M., Schneising, O., 2015. Algorithm Theoretical Basis Document Version 4 (ATBDv4) - The Bremen Optimal Estimation DOAS (BESD) algorithm for the retrieval of XCO_2.

Roberts, D., Nelson, B., Adams, J., Palmer, F., 1998. Spectral changes with leaf aging in Amazon caatinga. Trees 12 (6), 315–325.

Rosema, A., Verhoef, W., Noorbergen, H., Borgesius, J., 1992. A new forest light interaction model in support of forest monitoring. Remote Sensing of Environment 42 (1), 23–41.

Samanta, A., Knyazikhin, Y., Xu, L., Dickinson, R. E., Fu, R., Costa, M. H., Saatchi, S. S., Nemani, R. R., Myneni, R. B., 2012. Seasonal changes in leaf area of Amazon forests from leaf flushing and abscission. Journal of Geophysical Research: Biogeosciences 117 (G1), n/a–n/a, g01015.

Sanders, A., De Haan, J., 2013. Retrieval of aerosol parameters from the oxygen A band in the presence of chlorophyll fluorescence. Atmospheric Measurement Techniques 6 (10), 2725–2740.

Schaepman, M. E., Kneubuhler, M., Bartholomeus, H., Malenovky, Z., Damm, A., Schaepman-Strub, G., Hueni, A., 2008. Scaling spectroscopic approaches-from leaf albedo to ecosystems mapping. In: Proceedings of ASD and IEEE GRS; Art, Science and Applications of Reflectance Spectroscopy Symposium, Vol. II. Boulder, Colorado: ASD Inc. pp. 1–19.

Schreiber, U., 2004. Pulse-amplitude-modulation (PAM) fluorometry and saturation pulse method: an overview. In: Chlorophyll a Fluorescence. Springer, pp. 279–319.

Schull, M. A., Knyazikhin, Y., Xu, L., Samanta, A., Carmona, P. L., Lepine, L., Jenkins, J., Ganguly, S., Myneni, R. B., 2011. Canopy spectral invariants, Part 2: Application to classification of forest types from hyperspectral data. Journal of Quantitative Spectroscopy and Radiative Transfer 112 (4), 736–750.

Smith, W., Harrison, F., Hinton, D., Revercomb, H., Bingham, G., Petersen, R., Dodge, J., 2002. GIFTS-the precursor geostationary satellite component of the future Earth Observing System. In: Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International. Vol. 1. IEEE, pp. 357–361.

Smolander, S., Stenberg, P., 2005. Simple parameterizations of the radiation budget of uniform broadleaved and coniferous canopies. Remote Sensing of Environment 94 (3), 355–363.

Ter Steege, H., Pitman, N. C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., Phillips, O. L., Castilho, C. V., Magnusson, W. E., Molino, J.-F., et al., 2013. Hyperdominance in the Amazonian tree flora. Science 342 (6156), 1243092.

Tian, H., Melillo, J., Kicklighter, D., McGuire, A., Helfrich Iii, J., Moore Iii, B., Vörösmarty, C., 2000. Climatic and biotic controls on annual carbon storage in Amazonian ecosystems. Global Ecology and Biogeography 9 (4), 315–335.

TRMM, 2011. Tropical Rainfall Measuring Mission Project (2011),TRMM/TMPA 3B43 TRMM and Other Sources Monthly Rainfall Product V7,version 7, Greenbelt, MD:Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC). Available online: http://disc.sci.gsfc.nasa.gov/datacollection/TRMM_3B43_V7.html.

Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote sensing of Environment 8 (2), 127–150.

van der Tol, C., Berry, J., Campbell, P., Rascher, U., 2014. Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence. Journal of Geophysical Research: Biogeosciences 119 (12), 2312–2327.

van der Tol, C., Verhoef, W., Rosema, A., 2009a. A model for chlorophyll fluorescence and photosynthesis at leaf scale. Agricultural and forest meteorology 149 (1), 96–105.

van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., Su, Z., 2009b. An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. Biogeosciences 6 (12), 3109–3129.

Veefkind, J., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H., De Haan, J., Kleipool, Q., et al., 2012. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sensing of Environment 120, 70–83.

Verhoef, W., Bach, H., 2003. Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models. Remote Sensing of Environment 87 (1), 23–41.

Verrelst, J., Schaepman, M. E., Koetz, B., Kneubühler, M., 2008. Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. Remote Sensing of Environment 112 (5), 2341–2353.

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Köhler, P., Jung, M., Varlagin, A., Guanter, L., 2015. Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. Global change biology.

Wang, P., Stammes, P., Pinardi, G., Roozendael, M. v., et al., 2008. FRESCO+: an improved O 2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals. Atmospheric Chemistry and Physics 8 (21), 6565–6576.

Wit, E., Heuvel, E. v. d., Romeijn, J.-W., 2012. 'all models are wrong...': an introduction to model uncertainty. Statistica Neerlandica 66 (3), 217–236.

Wolanin, A., Rozanov, V., Dinter, T., Noël, S., Vountas, M., Burrows, J., Bracher, A., 2015. Global retrieval of marine and terrestrial chlorophyll fluorescence at its red peak using hyperspectral top of atmosphere radiance measurements: Feasibility study and first results. Remote Sensing of Environment 166, 243–261.

Wu, J., Albert, L. P., Lopes, A. P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K. T., Guan, K., Stark, S. C., Christoffersen, B., Prohaska, N., et al., 2016. Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests. Science 351 (6276), 972–976.

Yang, W., Kobayashi, H., Chen, X., Kondoh, A., 2016. An empirical forest structure dinamics model for driving 3-D canopy radiative transfer simulations. In preparation.

Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., Munger, J. W., Kornfeld, A., Richardson, A. D., 2015. Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. Geophysical Research Letters 42 (8), 2977–2987.

Zarco-Tejada, P. J., Pushnik, J., Dobrowski, S., Ustin, S., 2003. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. Remote Sensing of Environment 84 (2), 283–294.

Zarco-Tejada, P. J., Suarez, L., González-Dugo, V., 2013. Spatial resolution effects on chlorophyll fluorescence retrieval in a heterogeneous canopy using hyperspectral imagery and radiative transfer simulation. Geoscience and Remote Sensing Letters, IEEE 10 (4), 937–941.

Zhang, Y., Guanter, L., Berry, J. A., Joiner, J., Tol, C., Huete, A., Gitelson, A., Voigt, M., Köhler, P., 2014. Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. Global change biology 20 (12), 3727–3742.

The CV is not published online for reasons of data protection.

The CV is not published online for reasons of data protection.

Danksagung

Nach einer intensiven Abschlussphase meiner Promotion möchte ich die Gelegenheit nutzen um mich bei allen Personen zu bedanken, die mich in den Jahren als Doktorand und insbesondere während der Erstellung dieser Arbeit unterstützt haben.

Ein besonderer Dank gilt meinem Betreuer und Zweitgutachter Prof. Dr. Luis Guanter, der durch seine fachliche Kompetenz, konzeptionelle Beratung und wertvolle Unterstützung einen großen Anteil am Gelingen dieser Arbeit hat.

Bei Herrn Prof. Dr. Fischer möchte ich mich für die Übernahme des Erstgutachtens sowie für seine engagierte Hilfe bei den Formalitäten bedanken.

Allen Mitarbeitern des Instituts für Weltraumwissenschaften an der Freien Universität Berlin und meinen Kollegen der Sektion Fernerkundung des GeoForschungsZentrums in Potsdam danke ich für die Inspiration, Hilfsbereitschaft und die angenehme Arbeitsatmosphäre.

Anna Grimm danke ich dafür, dass sie mehr als nur einmal ihre eigenen Interessen hintangestellt hat, um mich zu unterstützen und zu motivieren.

Meinen Eltern gebührt ein großer Dank für die finanzielle sowie moralische Unterstützung während meines Studiums und darüber hinaus. Insbesondere die Freiheit wegweisende Entscheidungen selbst zu treffen hat nicht zuletzt zu meiner erfolgreichen Promotion beigetragen.

Darüber hinaus gilt mein Dank allen Verwandten, Freunden und Bekannten, die mir eine willkommene Abwechslung zum akademischen Alltag geboten haben.