Remote Sensing of Sun-Induced Chlorophyll Fluorescence

MSc Thesis by Thomas C. van Leth

Supervisors: Dr. ir. Willem W. Verstraeten (KNMI, WUR),

Dr. Abram F.J. Sanders (KNMI), Dr. Geert-Jan H. Roelofs (UU)

September 2014

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

Universiteit Utrecht

Abstract

Terrestrial vegetation is a major sink of CO_2 through the process of photosynthesis. Assessing the photosynthesis rate is therefore of utmost importance for improved modelling and predicting the global carbon cycle. Simultaneous with photosynthesis, vegetation also emits a small fraction of its received solar energy as (sun-induced) fluorescent light at a longer wavelength. Consequently, sun-induced fluorescence (SiF) potentially can be used as a proxy for photosynthesis. Recently, novel techniques were developed to measure SiF from satellite sensors.

This research has focused on implementing an algorithm to retrieve SiF from the GOME-2 instrument using a statistical approach for accounting atmospheric effects on the measured spectrum. Global monthly averaged SiF maps for 2013 were obtained. To investigate the performance of the retrieval algorithm with respect to fitting parameters, several sensitivity test were performed.

The statistical approach for considering atmospheric effects on the measured spectrum is based on the selection of reference measurements of vegetation/fluorescence free areas. We show that the accuracy of the retrieval result depends on the selection of reference spectra. We also demonstrate that SiF retrievals perform well using a wider spectral window than in previous efforts.

The SiF results were evaluated using (i) gross primary productivity (GPP) derived from CO_2 measurements of a number of flux towers, (ii) the vegetation index NDVI (Normalized Difference Vegation Index), and (iii) other SiF retrieval efforts reported in the literature. The obtained spatio-temporal SiF data correspond well with these various evaluation sources. In addition to SiF, the fluorescence yield was computed. This is a proxy for the Light Use Efficiency, a common parameter in ecosystem carbon exchange models. The preliminary results are promising, pending further evaluation in future research.

Acknowledgements

I would like to thank the KNMI for providing me with the opportunity for this research internship and also for facilitating a trip to Vienna to present my findings there. I would like to thank especially my supervisors at KNMI, Willem Verstraeten and Bram Sanders, for their help throughout this endeavour. I would also like to thank Geert-Jan Roelofs, not just for being my supervisor at IMAU for this project, but also for helping me out as a mentor during my coursework.

Table of Contents

Abstract 3

Acknowledgements 5

1 Introduction and background 1

1.1 Global carbon cycle 2

1.2 Chlorophyll fluorescence 3

1.3 Measuring Methods 6

1.3.1 In Situ measurement 6

1.3.2 Ground-based remote sensing 7

1.3.3 Satellite-based remote sensing 9

1.3.4 Statistical retrieval methods 10

2 Retrieval Model 11

2.1 TOA reflectance model 11

2.2 Statistical Retrieval Approach 14

2.3 Spectral window 15

3 Retrieval Implementation 17

3.1 The GOME-2 Instrument 18

3.2 Minimization Scheme 20

3.3 Principal Component Algorithm 21

3.3.1 Principal component analysis 21

3.3.2 Iterative principal component scheme 21

3.4 Reference Spectra Selection 22

3.4.1 Desert 22

3.4.2 Cloudy Ocean 23

3.4.3 Ice/Snow 24

3.5 Target pixel filtering 25

3.6 Retrieval Settings 26

4 Baseline Retrieval Results 27

4.1 Spatial Patterns 27

4.2 Temporal Patterns 29

4.3 Error Analysis 32

5 Retrieval Experiments 35

5.1 Reference Datasets 35
5.1.1 Reference Areas 35
5.1.2 Reference Months 38
5.1.3 Seasonality 41
5.2 Number of Principal Components 44
5.3 Surface reflectance polynomial and window 48
6 Fluorescence Yield 53
6.1 Method 53
6.2 Results 54
7 Evaluation 59
7.1 Comparison with NDVI 59
7.2 Comparison with previous work 61
7.3 Australian Flux sites comparison 62
8 Discussion 69
8.1 Interpretation of baseline results 69
8.2 Interpretation of fluorescence yield results 70
8.3 Effect of retrieval parameters 71
8.4 Evaluation of obtained GPP estimate 72
8.5 Validation 72
9 Conclusions and recommendations 75
References 77
Appendix A: Additional algorithm aspects 83
A.1 High-sampling Interpolation 83
A.2 Pixel Selection Algorithm 83
A.3 Sun-glint rejection 84
Appendix B: Python Implementation Overview 87
Appendix C: Glossary 89

1 Introduction and background

Photosynthesis by terrestrial vegetation is a major sink of CO_2 from the atmosphere. An accurate quantification of this uptake by plants is therefore essential to global carbon cycle modeling and predicting future atmospheric CO_2 concentrations. Because of global land use change it is also important to quantify the differences in productivity between ecosystems. Plants that perform photosynthesis typically fluoresce in the red and near infrared. Measurements of this so-called Sun-induced chlorophyll fluorescence (SiF) from space have the potential to improve the accuracy of global photosynthesis maps.

In recent years new methods are being developed to retrieve sun-induced chlorophyll fluorescence using existing satellite instruments such as GOME2, GOSAT and SCIAMACHY (Frankenberg et al., 2011; Joiner et al., 2011, 2012, 2013; Guanter et al., 2012, 2013). A key feature of these methods is that extensive radiative transfer modelling is avoided by taking a statistical approach. These retrieval methods could possibly be applied to new instruments that are now becoming available such as OCO2 (Frankenberg et al. 2014) and TROPOMI.

This thesis project aims to develop an algorithm for the retrieval of sun-induced chlorophyll fluorescence from space based hyperspectral measurements using a statistical approach. The retrieval algorithm is applied to data from the GOME2 instrument aboard the MetOp-A platform. More specifically, we aim to use this retrieval product to obtain a better estimate of gross primary production (GPP) for terrestrial vegetation. To evaluate our SiF retrieval we compare with data from several tower based flux measurement sites.

The goals for this thesis work can be summarized as follows:

- Implementing an algorithm to retrieve sun-induced chlorophyll fluorescence from GOME-2 data;
- Investigating the sensitivities of the retrieved results to different algorithm settings and; developing a suitable baseline for operational use;
- Investigating ways to use SiF as a proxy for GPP;
- Validating the results using existing ground-based measurements.

This thesis is structured as follows. In **section 1** chlorophyll fluorescence is introduced and an overview of previous work is given. In s**ection 2**, the forward model used in the retrieval is described. In **section 3**, a detailed overview of the retrieval algorithm is given. In **section 4**, the results of our baseline fluorescence retrieval are discussed. In **section 5**, the sensitivity of the retrieval to different retrieval settings is analyzed. In **section 6**, a global fluorescence yield estimate is presented. In **section 7**, we evaluate our results by comparing them with flux tower measurements as well as other vegetation indices. The results are interpreted and discussed in **section 8** and finally we present our conclusions and recommendations for further research in **section 9**.

1.1 Global carbon cycle

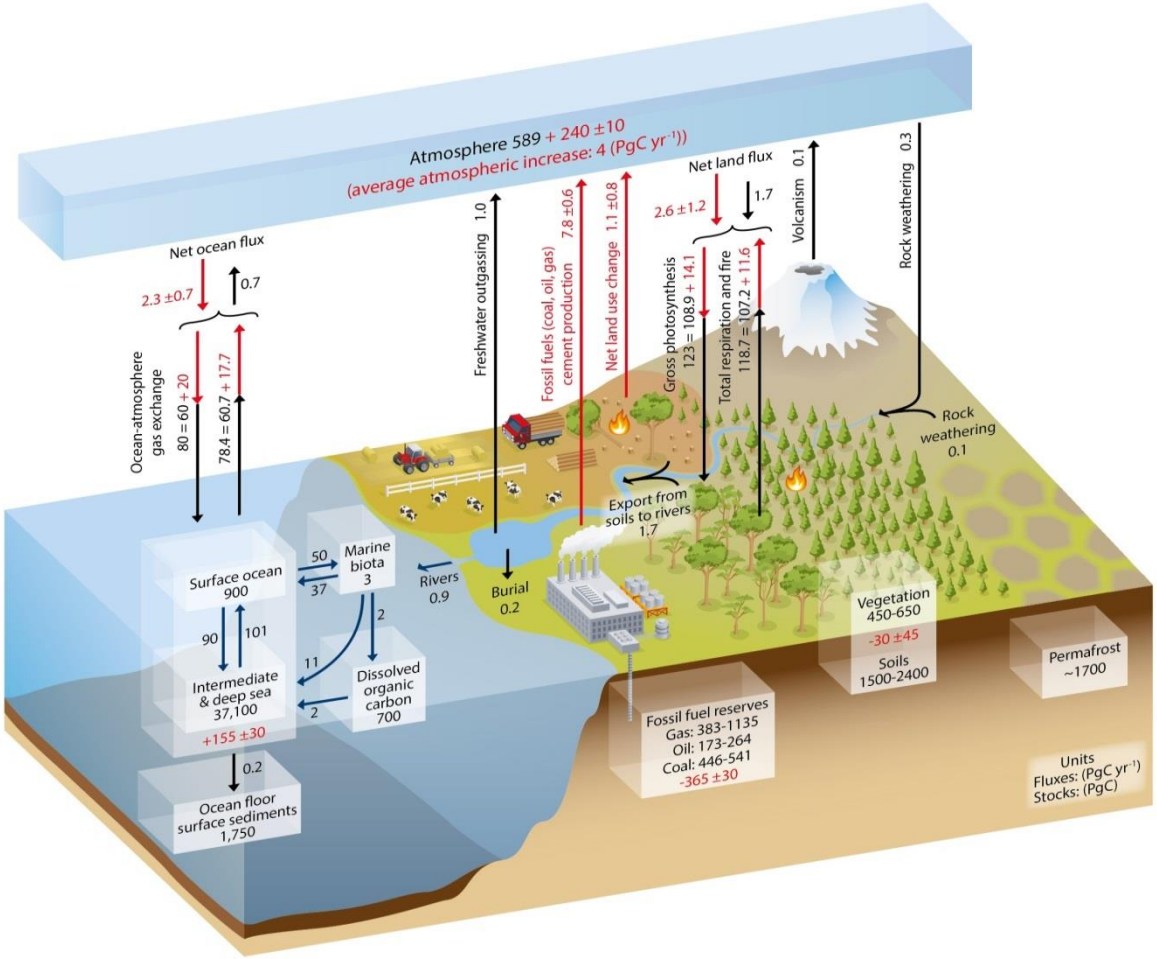


Figure 1.1: Estimation of global carbon flux and reservoir sizes from the IPCC fifth assessment report (Ciais et al. 2013)

Terrestrial vegetation can act as both a source and a sink of atmospheric carbon dioxide through respiration and photosynthesis respectively. As illustrated in Figure 1.1, both the respiration and the photosynthesis term are the largest source/sink terms for CO_2 in the atmosphere. The sink of CO_2 in the atmosphere due to photosynthesis can be expressed as gross primary productivity (GPP) which is usually measured in grams of carbon per square meter per day ($gC/m^2/d$). From Figure 1.1, it can also be observed that the difference between the sink of carbon through photosynthesis and the source of carbon from respiration is comparatively small. This quantity is called the net primary productivity (NPP). Because this net sink is so much smaller than the respiration and photosynthesis terms, a relatively small inaccuracy in either of those terms can lead to a relatively large uncertainty in the NPP. This is part of the reason why an accurate assessment of these terms is so important. CO_2 exchange of vegetation with the atmosphere can be measured in situ using a gas analyzer above the vegetated field under investigation. What is measured then is the net ecosystem production (NEP), which is the net conversion of CO_2 into organic material of the complete ecosystem including but not limited to the primary producers. Here, we focus just on the photosynthesis itself and thus the GPP is our quantity of interest. The net photosynthesis reaction is given in equation (1.1).

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2 \qquad (1.1)$$

GPP can be derived from flux tower based NEP measurements using an ecological model. However it cannot be directly measured through satellite remote sensing. In order to remotely sense GPP, one must look at the elements that drive GPP. GPP is mainly determined by the amount of sunlight that can be used for photosynthesis. GPP in a patch of vegetated land generally can be given by the following equation (Monteith 1972; Monteith 1977):

$$GPP = LUE \cdot fAPAR \cdot PAR \quad (1.2)$$

PAR is the photosynthetically active radiation, which is the incoming solar irradiance between 400 nm and 700 nm. fAPAR is the fraction of photosynthetically active radiation that is absorbed by vegetation. It is mainly determined by the fraction of the surface that is covered by vegetation and the density of the vegetation canopy. LUE is the light use efficiency. The LUE is a lumped factor that represents any physiological processes that might hamper photosynthesis on the leaf level. fAPAR can be remotely estimated to a reasonable amount of reliability from reflective greenness indices such as NDVI (e.g., (Olofsson & Eklundh 2007)). However, LUE is difficult to assess from satellite observations, while it is also highly variable between biomes (Turner et al. 2003). In models based on the approach by Monteith, the LUE is often assumed to be constant or split into factors that are empirically related to potential meteorological photosynthesis limiters (Schaefer et al. 2012).

Factors that play a role in limiting the efficiency of photosynthesis include: water availability, canopy temperature, atmospheric CO_2 concentration, nutrient availability, pollution, parasites, etc. (Schulze, Beck & Müller-Hohenstein 2005). Many of these are quite uncertain themselves and the physiological relations with light use efficiency are not completely understood. Therefore it would be beneficial to have a direct proxy for LUE that can be used without prior assumptions about causes for reduced LUE. In fact such a proxy could actually help to shed light on the relationships between those stress factors and the LUE.

1.2 Chlorophyll fluorescence

Inside the leaf of a typical terrestrial plant, sunlight is absorbed by so-called chloroplasts. The energy that is thus absorbed can be used to drive the photochemical reactions that produce glucose. However, since there may be other factors limiting the production capacity, it is usually not possible for the plant to fully utilize all of this energy for photochemical production. In this case, the plant needs to release this excess energy. There are generally two ways that a plant can do this: Through heat dissipation and through fluorescence. Of the sunlight that is absorbed, typically ~20% is released as heat; while only ~1% is emitted through means of fluorescence (see e.g. Baker (2008) for more detail). This is illustrated in Figure 1.2.

Chlorophyll-a is the main pigment employed in photosynthesis and often supplemented by other pigments such as chlorophyll-b and carotenes. The pigments in vivo are part of a larger light harvesting protein complex known as a photosystem. There are two photosystems in plants which work in tandem: PSI and PSII. The majority of fluorescence emissions originate from PSII.

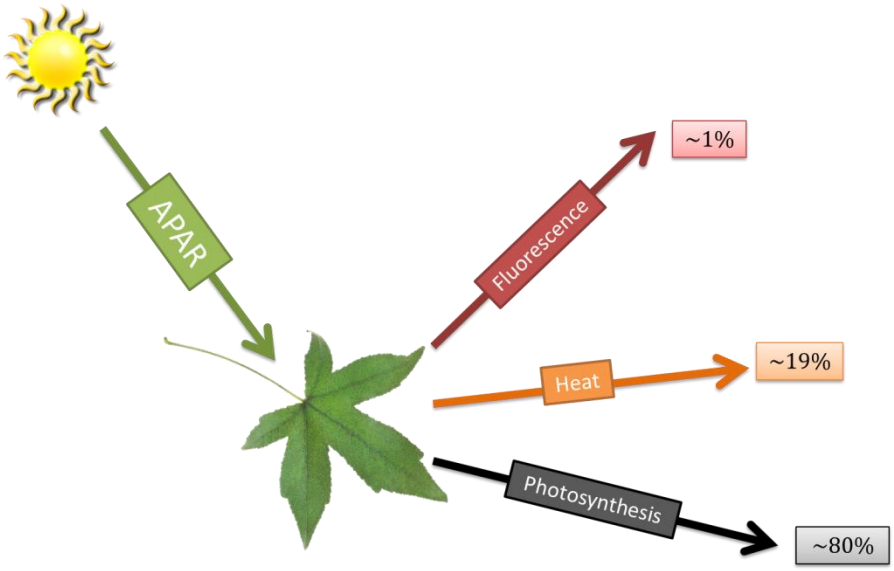


Figure 1.2: Basic overview of the different energy pathways of incoming absorbed sunlight. The percentages are a rough indication as they can be highly variable.

Fluorescence is the process by which absorbed radiation of a certain wavelength is emitted at a longer wavelength. For chlorophyll-a the absorbed radiation is the visible part of the radiation spectrum except for green (e.g. Gates et al. (1965)) and the radiation is emitted on the edge between the red part of the spectrum and the near infrared part. The fluorescence emitted under the influence of sunlight has a distinct spectral shape with one peak at 730-740 nm and another peak at 685-690 nm (Pedros et al. 2008). The exact spectral shape of chlorophyll-a absorption and fluorescence is different depending on whether one looks at chlorophyll in isolation, the emission from a single chloroplast, the emission from a leaf, the top of the canopy or the top of the atmosphere. Figure 1.3 illustrates the absorption and emission spectra of chlorophyll-a.

Since chlorophyll fluorescence is directly related to the internal photosynthesis mechanisms of a plant, the detection of chlorophyll fluorescence indicates photosynthetically active vegetation. This is different from other more traditional vegetation proxies such as NDVI since they only detect the presence of green leaves. This does not necessitate that those leaves are also optimally photosynthesizing. It does not take into account the actual PAR that is absorbed nor does it take into account short term stress factors that might limit the photosynthesis. Only long term stress factors have an effect on leaf greenness (e.g. Zhang et al. 2010), while chlorophyll fluorescence would only be present in leaves that are actually functioning.

Chlorophyll fluorescence is a relatively small signal. It is not visible with the naked eye and under daylight conditions the reflected sunlight is much stronger than the fluorescence signal. Nevertheless there is mounting evidence that it can be distinguished from the reflectance signal (Zarco-Tejada et al. 2000) and measured from the top of the atmosphere (Guanter et al. 2007). As such, it is a very promising proxy for space-born remote sensing of photosynthesis.

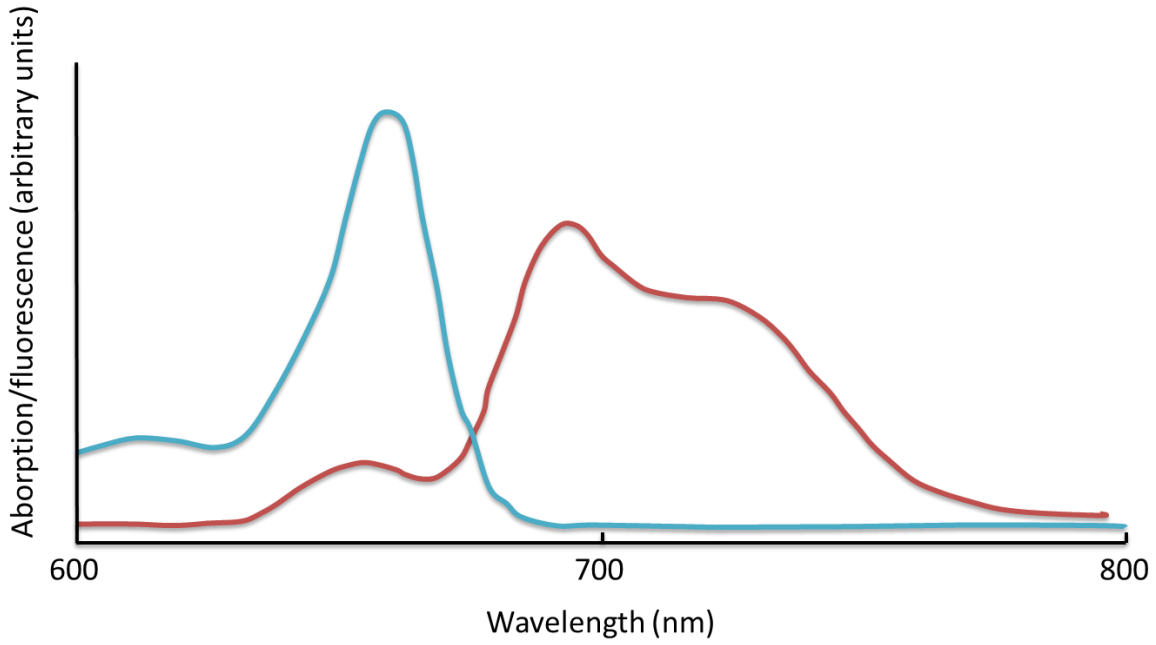


Figure 1.3: Example of typical Absorption and fluorescence spectra of green vegetation. spectra of a pigment extract from common hornbeam leaves. Adapted from Lichtenthaler et al. (1986). Note the depression in the fluorescence spectrum at around 670 nm which is caused by reabsorption in the extract.

The peak of fluorescence in the far-red spectral region is, on the chloroplast level, larger than the peak in the near-infrared spectral region. However this peak partially overlaps with a peak in the chlorophyll-a absorption spectrum. This means that light emitted by plants at this wavelength is partially reabsorbed within the leaf and within the larger canopy structure. The top of canopy spectrum and consequently the top of atmosphere spectrum have a much reduced far-red peak because of this. The amount of reabsorption is dependent on the chlorophyll content of the leaf and the geometry and density of the canopy structure (Buschmann 2007).

The reabsorption phenomenon makes the far-red peak on its own less useful as a proxy for photosynthesis compared with the near-infrared peak, although the ratio between the two could be used as a measure for reabsorption and as such provide extra information on chlorophyll content and canopy structure. In our study, we focus on retrieval of the near-infrared peak.

When the light use efficiency of the plant changes, the fraction of the incoming PAR that is emitted as fluorescence and the fraction of the incoming PAR that is emitted as heat changes as well. Here, it is useful to introduce the fluorescence yield which is defined as the fluorescence radiance as a fraction of the PAR absorbed by the plant. We can similarly define a heat dissipation yield as well.

$$\phi_f = \frac{I_F}{APAR} \qquad \phi_d = \frac{Q}{APAR} \qquad (1.3)$$

As the LUE (or equivalently photochemical yield) decreases, one or both of the fluorescence yield and/or the heat dissipation yield must increase. A high fluorescence yield value therefore indicates a low LUE. Although the heat dissipation yield is generally much larger than the fluorescence yield, for a large part of the range of LUE there exists a negative linear relationship between LUE and

fluorescence yield (van der Tol, Verhoef & Rosema 2009). This relationship switches sign at the low end of the LUE range and the point at which the relationship switches sign is dependent on the internal CO_2 concentration of the leaf.

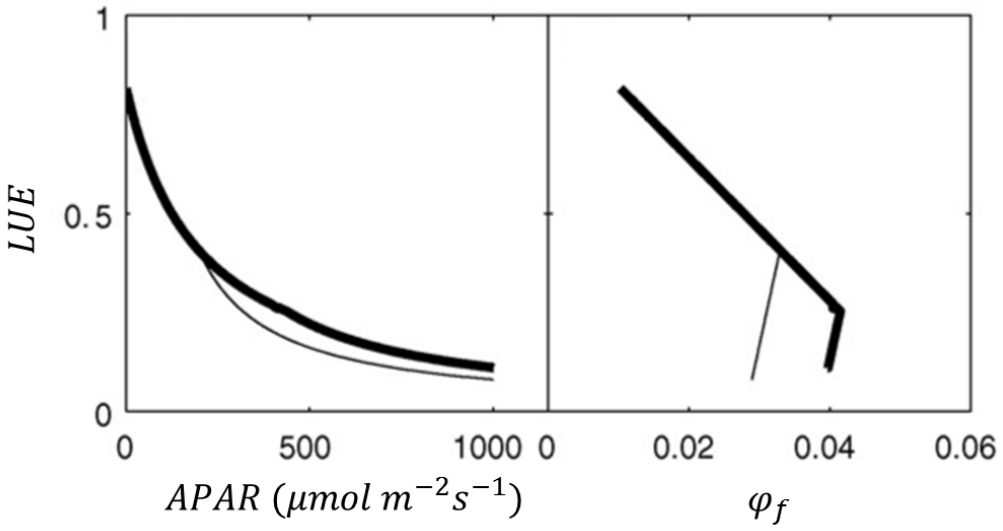


Figure 1.4: Theoretical relations between light use efficiency and APAR and between light use efficiency and fluorescence yield. The thin line illustrated the relation at a low CO_2 concentration while the thick line illustrated the relation at a high CO_2 concentration. Adapted from van der Tol et al. (2009).

This relationship could make the fluorescence yield also a useful proxy for LUE within the light-limited regime. Indeed, it was shown by Damm et al. (2010) that fluorescence yield can be used to track changes in LUE and thus GPP at least at the tower scale.

This complex relationship between fluorescence and photochemistry also affects the interpretation of fluorescence radiance. Van der Tol et al. (2009) conclude that in the mid-latitudes the relationship between fluorescence radiance and total photosynthesis rate is positive in the afternoon from around 11:00 onwards and more difficult to interpret in the morning. This can be a relevant concern given the overpass time of the instruments as we will see **in section 3.1**.

1.3 Measuring Methods

1.3.1 In Situ measurement

Laboratory and field measurement techniques of chlorophyll fluorescence have been used for several decades in botanical research. However, these measurement techniques are not suitable for measuring chlorophyll fluorescence at canopy level. These are active measurements which are performed close to the leaf itself and often rely on a controlled excitation radiation being applied to an otherwise light-isolated leaf surface. As such there is no need to distinguish the fluorescence signal from reflective signals from the leaf itself or its environment. By using a modulated excitation signal, the response of the leaf to the controlled signal can be distinguished from the response of the plan to ambient light. In this way fluorescence in leaves can also be measured in ambient light (Quick & Horton 1984), allowing measurements to be performed outside of the laboratory using a leaf clip. Because these measurements are performed close to the leaf itself, radiative extinction patterns are not relevant. For more general information on these types of chlorophyll fluorescence measurements, we refer to Maxwell and Johnson (Chlorophyll fluorescence - a practical guide 2000).

These leaf level in-situ measurements quantify the response of the plant to an isolated light signal and as such they can provide information on the health status of the plant and the functioning of its photochemistry mechanisms. This can give an indication of the capacity for production. However, to get information on the actual production of a vegetated field due to ambient conditions and especially for the production of a surface area of vegetation as opposed to an individual leaf, it is more useful to passively measure the chlorophyll fluorescence that is caused by the excitation by sunlight conditions itself at the top of a canopy.

1.3.2 Ground-based remote sensing

To measure sun-induced fluorescence, techniques are required that are able to distinguish the fluorescence from the ambient reflection. For ground-based remote sensing of fluorescence (for example using a spectrometer mounted on a tower structure) it is possible to actually make use of the telluric absorption spectrum. In particular, the O_2- A absorption band, which overlaps with the chlorophyll fluorescence emission peak, is often used for ground-based remote sensing. Figure 1.5 shows a typical transmission plot for the near-infrared part of the spectrum including the O_2-A band.

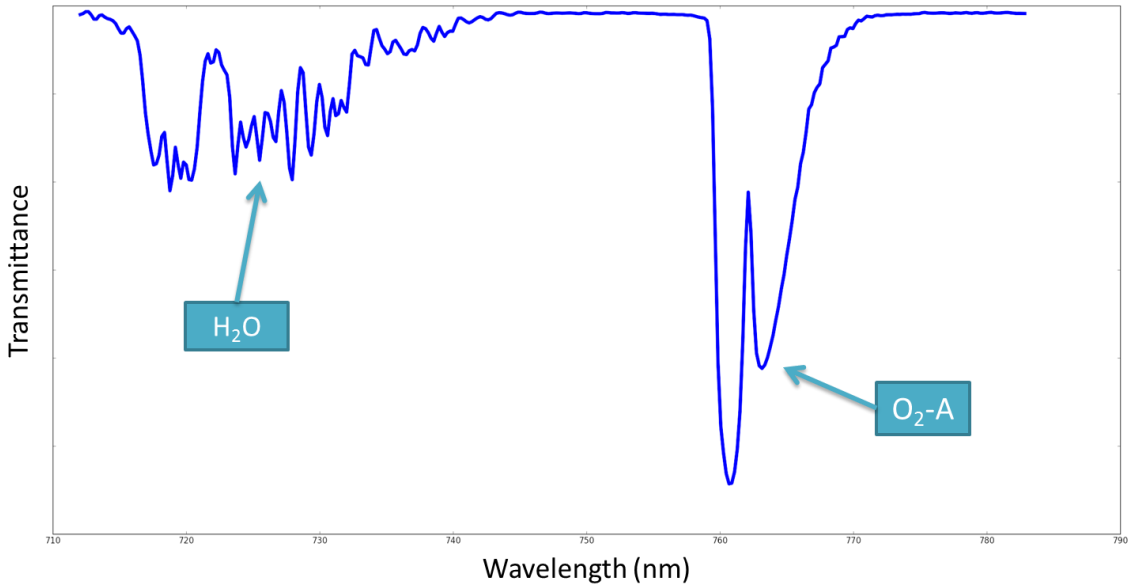


Figure 1.5: Typical atmospheric transmission spectrum, illustrating some prominent telluric absorption features within the spectral range employed in this work. The spectrum is derived from the data produced in this work.

The technique used in this case is based upon the in-filling of solar Fraunhofer lines (Plascyk & Gabriel 1975), but more often applied to the telluric O_2-A band instead (Moya et al., 2004). Fraunhofer lines are the spectral absorption features of gasses in the colder outer layer of the solar atmosphere. Therefore, any radiation originating from the sun contains this spectral signature. The telluric absorption lines, originating in the earth's atmosphere, are present in any signal that has travelled a significant distance through the atmosphere.

A spectrometer first measures a target patch of vegetation and then it measures a painted reference panel with known reflectivity held close to the target patch for which the same optical path can be assumed. From this reference measurement the solar irradiance incident upon the target vegetation can be determined.

$$I_r(\lambda) = \rho_r F_s(\lambda) \quad (1.4)$$

Here, $I_r(\lambda)$ is the measured radiance from the reference panel, ρ_r is the known reference reflectivity and $F_s(\lambda)$ is the incident solar irradiance spectrum. $F_s(\lambda)$ can then be used to calculate the contribution of fluorescence to the measured radiance from the target patch because it includes both the telluric and solar absorption patterns. Any radiative signal emitted from the targeted surface area (i.e. fluorescence) would not contain these patterns as it would have travelled a negligible distance trough the atmosphere and has no solar origin, while any radiative signal from sunlight reflected by the target area would contain these patterns. In this way, the amount of so-called infilling of absorption features can be used to quantify the contribution of fluorescence to the measured radiance.

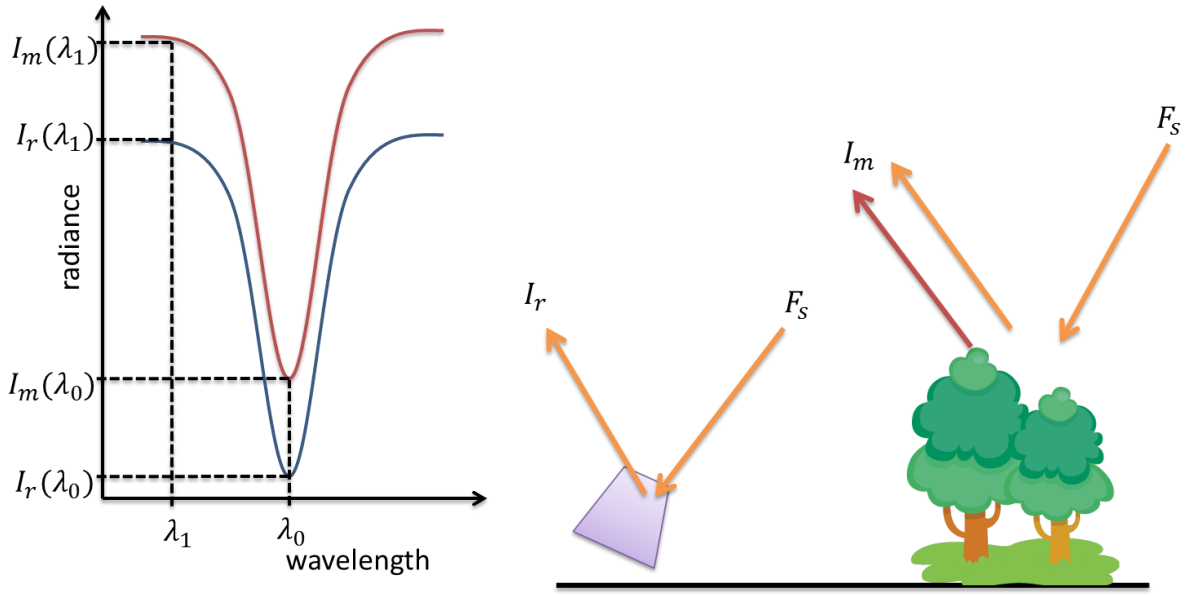


Figure 1.6: the Fraunhofer line in-filling principle. On the left, the spectral shapes of the radiance measured from the target vegetation and the radiance measured from the reference panel for a generic absorption feature are shown. On the right, the optical paths are illustrated.

The radiance measurements are done at different wavelengths, both inside and outside the absorption band. In theory, only two wavelengths are needed as is illustrated in Figure 1.6, although in practice three are often used. By solving the system of equations

$$\begin{aligned} I_m(\lambda_0) &= \rho_t F_s(\lambda_0) + I_F \\ I_m(\lambda_1) &= \rho_t F_s(\lambda_1) + I_F \end{aligned} \quad (1.5)$$

Where I_m is the radiance measured at the target patch, ρ_t is the reflectivity of the target, which is assumed independent of wavelength, the radiance contribution of fluorescence I_F, which is also assumed to be independent of wavelength, can be determined. Usually one wavelength is used in the center of the absorption feature and the other is taken at the edge of the absorption feature (or at both edges). Since the O_2-A band is a broader feature compared to the solar absorption features; a coarser spectral resolution can be employed than would be possible with solar absorption features.

1.3.3 Satellite-based remote sensing

When measuring sun-induced fluorescence from space, using the telluric line in-filling as in ground based remote sensing is much more complicated, because the fluorescence signal is also subject to atmospheric absorption as it travels towards the top of the atmosphere. Therefore, an extensive radiative transfer model that explicitly accounts for all optically active atmospheric constituents would be needed to be able to solve for the fluorescence component. This has been done (Guanter et al. 2007), but the drawback to this method is that it requires the use of derived atmospheric properties as inputs as well as a large computational expense.

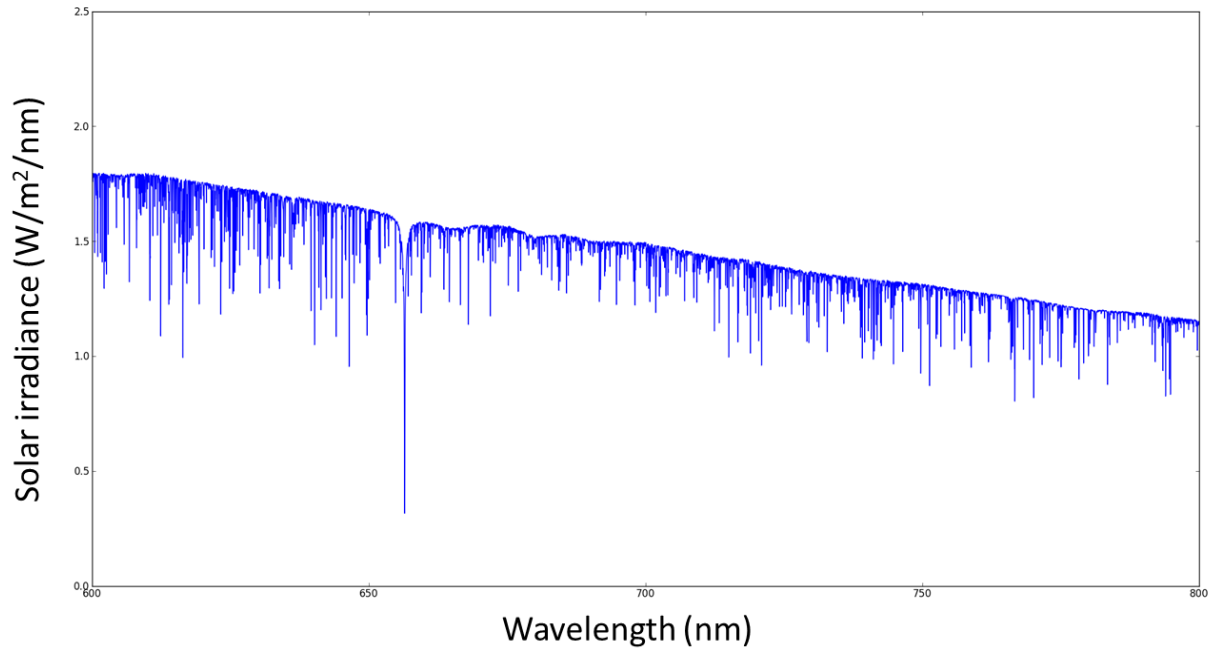


Figure 1.7: Part of the SAO2010 Reference solar spectrum (Chance & Kurucz 2010), illustrating the solar absorption features that are superimposed upon the blackbody spectral curve.

An alternative to a full radiative transfer model is to make use of a narrow window surrounding a solar Fraunhofer line outside of the atmospheric absorption bands (see Figure 1.7). Such a retrieval method has been performed by Joiner et al. (2011) using the K1 line, which lies just outside of the O_2-A absorption band. This method is similar to the ground-based retrieval method using Fraunhofer line infilling. By avoiding the telluric absorption bands, atmospheric absorption and scattering can be neglected, therefore avoiding the explicit radiative transfer modelling. In this way, this retrieval method relies solely on the solar absorption features for infilling. There is no reference reflective panel on the ground, but instead the solar spectrum is measured directly. A disadvantage is that a much finer spectral resolution is needed than for telluric line infilling. This limits the applicability of the method when considering existing or near future instruments. One example of an instrument that is capable of such spectral resolutions is the GOSAT satellite. However, this comes at a price since the spatial coverage of GOSAT is poor.

Retrieval based on only one line leads to a relatively high noise and a method was developed to use data from multiple Fraunhofer lines (Frankenberg, Butz & Toon 2011). This requires explicit modeling of the instrument line shape function and thus a non-linear fitting algorithm because a solar reference spectrum is used instead of a measured solar spectrum.

1.3.4 Statistical retrieval methods

Guanter et al. (2012) and Joiner et al. (2012) provided a statistical based approach that needs neither explicit modelling of the instrument line shape function nor explicit radiative transfer modelling in order to use the data of a large number of Fraunhofer lines within a spectral window. This approach assumes that the radiance spectrum can be expressed as a linear combination of several fluorescence-free principal components (or singular vectors) plus the fluorescence radiance. The singular vector decomposition is applied to a set of fluorescence free reference spectra. The spectral windows used in this approach are only a few nanometers wide and centered on Fraunhofer lines outside of the telluric absorption bands, similar to Joiner et al. (2011). On this scale the spectral dependence of sun-induced fluorescence is small and is therefore considered constant. Both the reference spectra and the target spectra are normalized by their slope. This statistical approach also allows retrievals to be performed with moderate spectral resolution hyperspectral imagers such as SCIAMACHY and GOME-2 (Joiner et al. 2012; Joiner et al. 2013).

To be able to perform retrievals within the O_2-A band using hyperspectral imagers, an extension of the statistical approach using principal components was developed (Guanter et al. 2013; Joiner et al. 2013). In this case, a broader spectral window is used that covers the entire O_2-A absorption band as well as some wavelengths outside of it instead of several micro windows centered on Fraunhofer lines. This method similarly applies a principal component analysis to certain fluorescence free reference spectra, but instead it makes use of both the telluric absorption spectrum and the Fraunhofer lines equally. This method requires identifying the spectral shape of the fluorescence signal and of the surface albedo. In this way a 'spectral signature' is defined for each of the three elements (surface albedo, fluorescence and telluric absorption) and this is used to disentangle these elements from each other.

In this research we have applied this same approach towards an even broader spectral window that covers both the O_2-A band and the nearby H_2O band. We will show that this yields comparable results and even improves upon the relative noise ratio. In the next chapter we will give a more extensive explanation of this method.

2 Retrieval Model

In this study we have adopted the retrieval approach developed recently by Joiner et al. (2013). We do not use micro-windows centered on isolated Fraunhofer lines, but instead we use a broad and continuous wavelength window that contains many fraunhofer lines as well atmospheric absorption features, surrounding the SiF peak in the near infrared. The main challenge is to characterize the atmospheric transmission spectrum within this window. In order to do so without explicit physical modelling, a statistical method is applied.

In this section we will discuss the forward model that is used in this approach. In the next section we will discuss the practical implementation of the retrieval method.

2.1 TOA reflectance model

We start with a radiative transfer model based upon the Lambertian approximation, which states that the ground surface can be assumed to reflect only isotropically (see e.g. Thomas & Stamnes, 2002). The reflectance measured by the instrument at the top of the atmosphere including the contribution from chlorophyll fluorescence can then be expressed as:

$$\rho_m(\lambda,\theta_v,\theta_s) \approx \rho_0(\lambda) + \frac{\rho_s(\lambda)\mathcal{T}(\lambda,\theta_s)\bar{\mathcal{T}}(\lambda,\theta_v)}{1-\rho_s(\lambda)\bar{\rho}(\lambda)} + \frac{\pi I_F(\lambda)\bar{\mathcal{T}}(\lambda,\theta_v)}{[1-\rho_s(\lambda)\bar{\rho}(\lambda)]F_s(\lambda)\mu_0(\theta_s)} \qquad (2.1)$$

Here $\rho_m(\lambda,\theta_v,\theta_s)$ is the reflectance measured by the instrument at the top of the atmosphere, $\rho_0(\lambda)$ is the atmospheric reflectance term (i.e. the reflectance in the case the ground surface is black), $\rho_s(\lambda)$ is the surface reflectance, $\bar{\rho}(\lambda)$ is the spherical albedo of the atmosphere for illumination from below. $\mathcal{T}(\lambda,\theta_s)$ is the atmospheric transmission along the optical path from the top of atmosphere (TOA) towards the reflecting surface from the angle of the incoming solar radiation. $\bar{\mathcal{T}}(\lambda,\theta_v)$ is the atmospheric transmission from the surface of the earth towards the top of the atmosphere in the direction of the satellite. Also, $F_s(\lambda)$ is the solar irradiance spectrum and $I_F(\lambda)$ is the sun-induced fluorescence emission. The wavelength is indicated with λ, while θ_v is the viewing zenith angle and θ_s is the solar zenith angle. In this work μ is used as shorthand for the cosine of the zenith angle:

$$\mu = \cos(\theta_v), \mu_0 = \cos(\theta_s) \qquad (2.2)$$

The first two terms on the right hand side of eq. (2.1) are actual reflectance terms. The last term on the right hand side is an emission term, which represents the fluorescence emission.

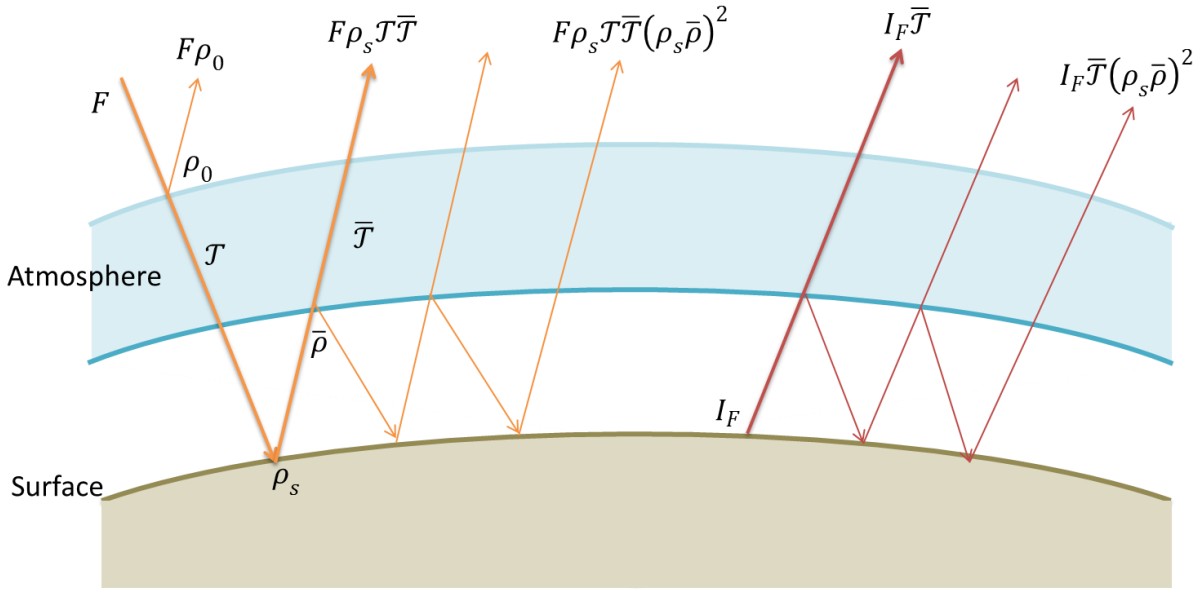


Figure 2.1: Radiative transfer model schematic. The orange arrows represent the reflection optical path, while the red arrows represent the fluorescence emission optical path. The bold arrows represent the optical paths in the case of negligible atmospheric scattering.

The possible optical paths within this model are indicated in Figure 2.1. The orange arrows indicate the optical path in the case of reflection of sunlight and the red arrows indicate the optical paths in the case of fluorescence. In both cases, the light can be reflected multiple times between the atmosphere and the surface. This is an abstract representation of the scattering processes that occur in the atmosphere. In fact, there are an infinite number of optical paths contributing to the signal received at the top of the atmosphere. These contributions get weaker by a constant factor for each iteration. This infinite series is a Taylor series and can be rewritten as such according to eq. (2.3). This factor is included in eq. (2.1).

$$\sum_{n=0}^{\infty}(\rho_s(\lambda)\bar{\rho}(\lambda))^n = \frac{1}{1-\rho_s(\lambda)\bar{\rho}(\lambda)} \qquad (2.3)$$

Since the reflectance signal that reaches the instrument actually includes the emission from sun-induced fluorescence, which is not an actual reflectance term, one can think of $\rho_m(\lambda)$ as an 'apparent' reflectance. However, from here on we will simply call this top-of-atmosphere (TOA) reflectance.

The main scattering constituents of the atmosphere are molecular nitrogen and molecular oxygen. These molecules are much smaller than the wavelengths in the optical and near infrared and thus the scattering regime in this case is Rayleigh scattering (e.g. Thomas & Stamnes, 2002). The Rayleigh scattering cross section exhibits a strong wavelength dependency:

$$\sigma_s \propto \frac{1}{\lambda^4}$$

Because of this, within the wavelength window of interest (700-800 nm), the atmospheric scattering cross section is small (in the order of $1 \cdot 10^{-27}$ to $2 \cdot 10^{-27}$ cm^2) and therefore $\rho_0(\lambda)$ is much smaller than the other terms in eq. (2.1) and $\bar{\rho}(\lambda) \ll 1$. This leads to a simplification of the model where

only two optical paths are now possible (illustrated in Figure 2.1 with bold arrows). The resulting simplification is as follows:

$$\rho_m(\lambda) \approx \rho_s(\lambda)\mathcal{T}(\lambda,\theta_s)\bar{\mathcal{T}}(\lambda,\theta_v) + \frac{\pi I_F(\lambda)\bar{\mathcal{T}}(\lambda,\theta_v)}{F_s(\lambda)\mu_0(\theta_s)} \quad (2.4)$$

Both transmission factors can be expressed in terms of the atmospheric optical depth:

$$\rho_m(\lambda) = \rho_s(\lambda)e^{-\tau(\lambda)} + \frac{\pi I_F(\lambda)}{\mu_0(\theta_s)F_s(\lambda)}e^{-\tau(\lambda)\frac{\mu(\theta_v)^{-1}}{\mu(\theta_v)^{-1}+\mu_0(\theta_s)^{-1}}} \quad (2.5)$$

Here, $\tau(\lambda)$ is the two-way atmospheric slant optical depth. The observation geometry is known and thus there are three unknown quantities and all three are wavelength dependent: Optical depth, surface albedo and the fluorescence emission. Note that the optical depth is also dependent on μ and μ_0, because this is a slant optical depth. Figure 2.2 illustrates the typical spectral shapes that are assumed for these three components as well as the measured TOA reflectance. We will elaborate on these spectral shapes in the following paragraphs.

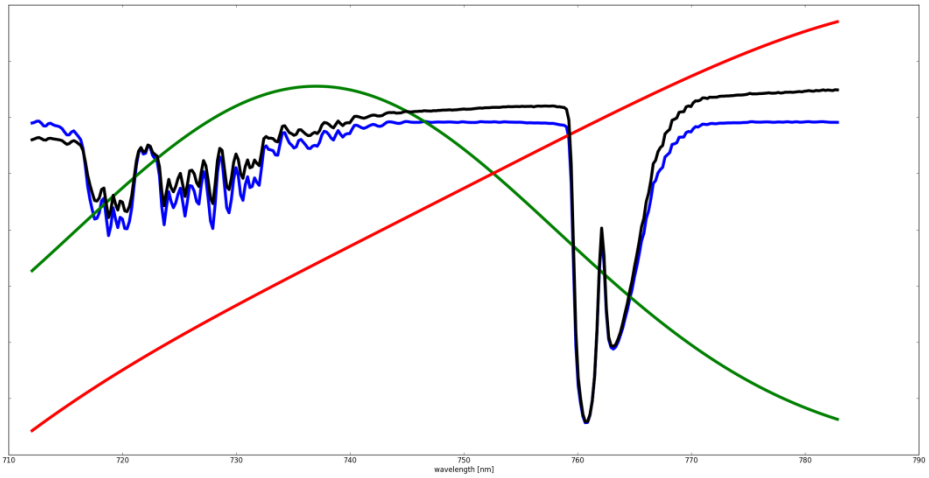


Figure 2.2: Illustration of the spectral model components within the wavelength window of interest. The black line is the TOA reflectance; the red line is a typical surface reflectance; the green line is the sun-induced chlorophyll fluorescence for vegetated land; and the blue line is the atmospheric transmission. These are all from the same typical retrieved scene. The components are shown in arbitrary units.

We assume that the fluorescence spectral shape can be adequately described as a superposition of two Gaussian shapes, representing the two peaks of chlorophyll fluorescence (Zarco-Tejada et al., 2000).

$$I_F(\lambda) = I_{F,0}\left(ce^{-\frac{1}{2}\left(\frac{\lambda-\mu_l}{\sigma_l}\right)^2} + e^{-\frac{1}{2}\left(\frac{\lambda-\mu_h}{\sigma_h}\right)^2}\right) \quad (2.6)$$

Here c is the ratio between the two peaks, σ_l and σ_h are the widths of the two peaks, μ_l and μ_h are the center wavelengths of the two peaks and $I_{F,0}$ is the amplitude of fluorescence. h and l stand for higher and lower representing the near-infrared and far-red peaks respectively.

For this retrieval attempt, the shape of the fluorescent signal is kept fixed. This means that both the center and width of the Gaussian are kept constant and only the amplitude is varied. Also, we use a

wavelength window covering only one of the peaks (the near-infrared peak), so a single Gaussian is used.

$$I_F(\lambda) = I_{F,0} e^{-\frac{1}{2}\left(\frac{\lambda-\mu_h}{\mu_h}\right)^2} \quad (2.7)$$

Here, μ is assigned the value 737 nm; σ is assigned the value 21 nm as in Joiner et al. (2013). Thus, $I_{F,0}$ is the radiance of the sun-induced fluorescence at 737 nm and this is the retrieved quantity of interest.

For the surface albedo, the shape shows more variability. However, the spectral dependence is assumed to be reasonably smooth (Joiner et al., 2013), so a low order polynomial could be used to capture most of the features of the shape.

$$\rho_s = \sum_{i=0}^{n} a_i \lambda^i \quad (2.8)$$

For most surface materials such as soil, a typical value of n would be two. However, due to the strong absorption characteristics of chlorophyll in the red, the surface albedo spectrum for vegetated land shows a strong increase on the edge between the red and near infrared wavelengths; the so-called red edge. Because our wavelength window is within this red edge, a higher order polynomial is needed for vegetated land than for non-vegetated land. Typically n is then in the order of four, which we will assume for most of this work (**section 4**), but we will also present the results when using a different number for n **(section 5.3**).

2.2 Statistical Retrieval Approach

The most complex to model is the optical depth. The fine spectral structure, which is caused particularly by the solar Fraunhofer lines as well as the H_2O and to a lesser extent the O_2-A absorption bands, cannot be parameterized by a simple analytical function of course. However, we want to avoid doing complicated explicit radiative transfer modelling calculations because these rely on externally derived atmospheric properties and also lead to a higher computational expense.

The alternative to radiative transfer modelling that we will use here is to assume that the optical depth for realistic atmospheric conditions can be written as a linear combination of a limited number of basis functions.

$$\tau(\lambda) = \sum_{j=1}^{m} b_j \cdot f_j(\lambda) \quad (2.9)$$

Here, $f_i(\lambda)$ are the basis functions and b_j are the unknown coefficients of these basis functions. The crux of the statistical method lies in how these basis functions are obtained: They are not analytical functions, but principal components.

To obtain these basis functions a critical assumption has to be made: The same set of basis functions that can be used to describe the optical depth of a column of atmosphere above non-fluorescent land can be used to describe a column of atmosphere above fluorescent land. With this assumption, the basis vectors can be obtained by aggregating a large number of non-fluorescent scenes with

diverse characteristics and applying a principal component analysis (PCA) to them. The m number of principal components with the highest explained variance are then the basis functions.

This assumption is not necessarily true, however: The characteristics of the ground surface may have an effect on the local atmospheric characteristics through, for example, different aerosol types and water vapor concentrations. Indeed, we have found that the choice of which non-fluorescent reference pixels to include in the PCA affects the retrieval results greatly (see **section 3.4** and **section 5.1**).

When no fluorescence is assumed to be present in a given scene, eq. (2.5) can be further simplified.

$$\rho_m(\lambda) = \rho_s(\lambda)e^{-\tau(\lambda)} \qquad (2.10)$$

Within the broad spectral window under investigation it is possible to identify certain sub-windows (λ_t) that are outside any atmospheric absorption bands and thus can be assumed to have an optical depth approaching zero.

$$\rho_m(\lambda_t) = \rho_s(\lambda_t) \qquad (2.11)$$

Within these narrow windows, the TOA reflectance is then approximately the same as the surface albedo. By fitting a second[1] order polynomial through these sub-windows and extrapolating to the entire window, one gets an approximation for $\rho_s(\lambda)$ in the entire window. The atmospheric optical depth for this particular scene is then given by the following equation:

$$\tau(\lambda) = -\ln\frac{\rho_m(\lambda)}{\rho_s(\lambda)} \qquad (2.12)$$

Optical depths obtained from a diverse set of vegetation-free scenes with diverse observation and solar angles and from a range of latitudes and longitudes are then used as the input for a principal component analysis algorithm.

2.3 Spectral window

The spectral window that is employed in this work is taken from 712 nm to 780 nm. This range covers the O_2-A band, the H_2O band and some atmospheric absorption free wavelengths between and around them. By taking such a broad window, the fitting parameters are more constrained and in this way we hope to achieve a more precise fitting than would be the case with a smaller window. By using a large window, the assumptions we made about the shape of surface albedo and the applicability of the reference optical depth to our retrieval target could become less obvious, however. We will present our findings on this aspect in **section 5.3** and **section 7.2**.

[1] Remember: this is the value for non-vegetated land or sea; we use a higher order for the surface albedo of the (vegetated) target pixel!

3 Retrieval Implementation

In this section we will first give an overview of the retrieval algorithm and then give a more detailed description of some key elements. A technical overview of the algorithm pipeline is given in appendix B.

The satellite instrument that we use data from is the GOME-2 imager aboard the MetOp-A platform (**section 3.1**). We use the operational level 1b data for 2013 and apply the retrieval method to data measured over the entire globe. First, level 1b data for 2013 is extracted from native level 1b files. The retrieval algorithm then consists of three distinct steps:

- **Determination of PCs**: From the data set, suitable fluorescence-free reference spectra are selected (**section 3.4**). Then a principal component analysis of these spectra is performed in an iterative scheme (**section 3.3**). The most significant principal components are selected to describe the optical depth (eq. (2.8)) and complete the forward model.
- **Actual retrieval:** Next, pixels for which a fluorescence retrieval is attempted are selected and filtered for e.g. clouds (**section 3.5**). For each pixel in the data set a cost-function is built from the TOA reflectance, the instrument error and the forward model and minimized using the Levenberg-Marquardt scheme **(section 3.2**). From this a number of optimized fit parameters are obtained: the amplitude of the fluorescence peak, the coefficients of the surface albedo polynomial and the coefficients of the optical depth basis functions.
- **Postprocessing:** post-retrieval filters are applied. The center coordinates of the GOME-2 footprints are then regridded to a regular 0.5^{o} x 0.5^{o} latitude longitude box grid and the fluorescence values are averaged per month and per grid-cell.

The process is summarized in Figure 3.1.

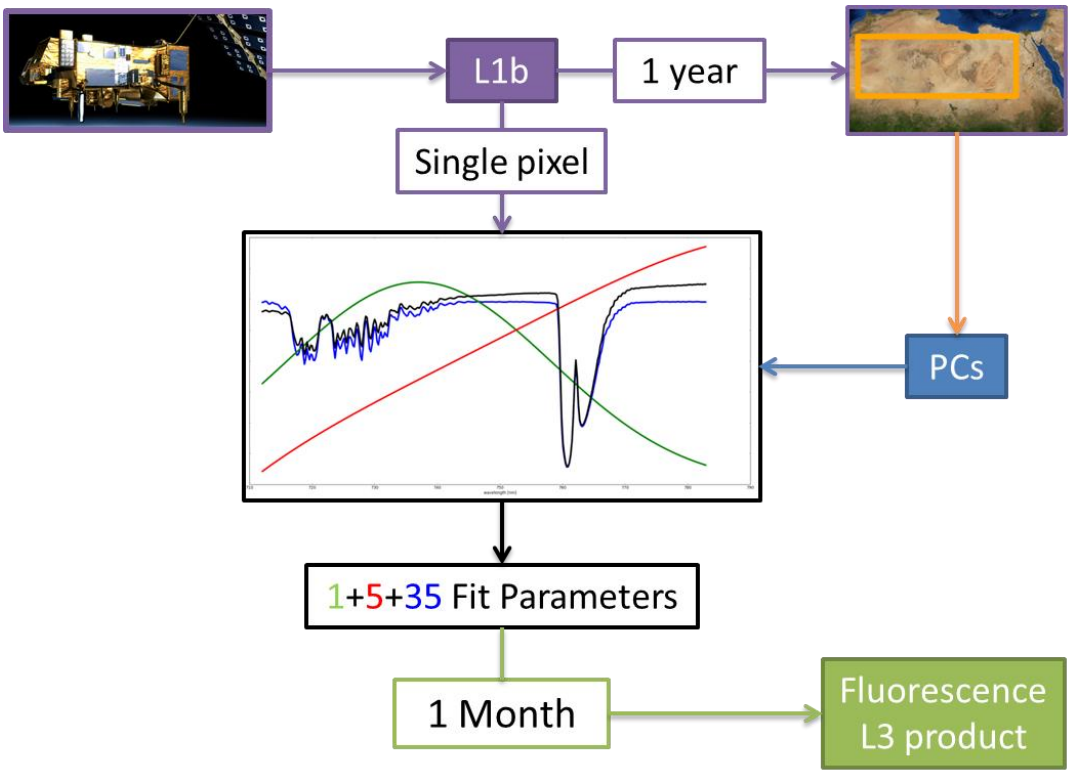


Figure 3.1: Simplified flow diagram of the algorithm. The input is GOME2 level 1b data. From this data, fluorescence-free reference pixels are selected; in the baseline retrieval these are all from the Sahara desert. One year (2013) of reference pixels are used in the principal component algorithm. The 35 most significant PCs are used to build a cost-function. For each pixel the minimization of the cost-function yields 41 fit-parameters (1 fluorescence parameter, 5 surface albedo polynomial coefficient and 35 PC coefficients). The fluorescence parameter is then averaged over a month per grid-cell resulting in the final level 3 product.

3.1 The GOME-2 Instrument

An instrument that has sufficient spectral resolution for the above method is the GOME-2 (Global Ozone Monitoring Experiment-2) imager aboard the MetOp-A and MetOp-B satellites (Callies et al. 2000). Each one of these has a global coverage in slightly more than a day. In this project MetOp-A data was used although the study could be expanded by including both satellites to get a better coverage. GOME-2 was designed to measure atmospheric ozone profiles but it also provides data on NO_2, SO_2, aerosol optical depth and water vapour among others (Callies et al. 2000). GOME-2 is normally nadir scanning, and it scans with a variable angular speed such that each pixel in the scan has an equally sized footprint.

Although an imager such as GOME-2 has a much better coverage compared to a sampling instrument such as GOSAT, its footprint is larger. The instrument aboard MetOp-A has a footprint of 40 km x 40 km from July 2013 onwards. Before that, its footprint was 40 km x 80 km. These large footprints imply that a single pixel in most cases represents a mixture of different surface types. This poses a challenge for evaluation of the results with independent in-situ measurements, since such reference measurements basically represent a point measurement. Only when the landscape is extremely homogeneous can it be expected that ground and satellite measurements actually measure the same properties.

MetOp-A operates in a sun-synchronous polar orbit and has an equatorial overpass time of around 9:30 local time. Its flight direction is from north to south during the dayside of the orbit (when measurements are done).

We make use of the operational level 1b product as provided by EUMETSAT (version 5.3). We have used only the 'nadir scanning' mode, and not any of the alternative modes. The instrument detects radiation between 240 nm and 790 nm. This range is divided into four channels. We have used only the channel 4 data, which has a spectral range from 593 nm to 790 nm. This range covers the red and near infrared parts of the radiative spectrum. Each channel is divided into 1024 spectral bins, each of which is a separate detector pixel. The nominal size of the bin is 0.21 nm; the full-width at half-maximum (FWHM) resolution within channel 4 is about 0.48 nm.

In addition GOME-2 also performs measurements along a narrow swath of 320 km for one day per month instead of the normal measurements with a swath of 1920 km or 960 km. We have used these measurements as well and they are indiscriminately combined with the regular measurements to obtain monthly averages. In each scan sweep of the instrument, measurements are conducted both forwards and backwards. We have only used the forwards scans, because the backwards scans overlap with the forward scans and the backwards pixels are three times as large.

Table 3.1: Summary of relevant GOME-2 statistics

Quantity	value
Spatial resolution (before July 15 2013)	40km x 80 km
Spatial resolution (after July 15 2013)	40km x 40 km
Spectral resolution (channel 4)	0.48 nm (FWHM)
Spectral bin size	0.21 nm
Swath width (before July 15 2013)	1920 km
Swath width (after July 15 2013)	960 km
Overpass at the equator	09:30
Years active	2006 - present
Operation mode used	Nadir scanning; forward scans

In addition to the earthshine measurements, GOME-2 performs measurements of the solar irradiance spectrum once per day using almost the same optical path within the instrument as the regular measurements. We used these to convert the TOA radiance into TOA reflectance using equation (3.1):

$$\rho_m(\lambda) = \frac{\pi I_m(\lambda)}{\mu_0 F_m(\lambda)} \qquad (3.1)$$

Where $I_m(\lambda)$ is the measured earthshine radiance and $F_m(\lambda)$ is the measured solar irradiance. By dividing the two and working with a TOA reflectance, any multiplicative instrument bias such as resulting from instrument degradation is to some extent compensated for.

Note that the solar spectrum and the earthshine spectrum are not measured in the same wavelength bins. Therefore we have to interpolate the solar irradiance values towards the wavelengths used in the earthshine. However, a simple linear interpolation will not suffice, because of the fine Fraunhofer

line structure. These lines are thinner than the spectral resolution of the instrument, and the interpolation could mean the difference between being in the middle of a Fraunhofer line and being between two Fraunhofer lines. In order to account for this, we used a so-called high-sampling interpolation technique, described in more detail in appendix A.1.

3.2 Minimization Scheme

The statistical model yields a cost function that includes the principal components. In this case we have used a cost-function that is weighted by the measurement errors that are determined by the GOME-2 level 1b algorithm and reported in the level 1b product. The cost-function is described in eq. (3.2).

$$C = \sum_{i=0}^{n} \left(\frac{\rho_m(\lambda_i) - \rho_p(\lambda_i, \boldsymbol{\beta})}{\sigma_m(\lambda_i)} \right)^2 \qquad (3.2)$$

Where $\rho_m(\lambda_i)$ is the reflectance as measured by the instrument, $\rho_p(\lambda_i)$ is the reflectance obtained from the model, $\sigma_m(\lambda_i)$ is the instrument measurement error and **β** is the parameter vector. The modelled reflectance can be described by eq. (3.3):

$$\rho_p(\lambda_i, \boldsymbol{\beta}) = \left(\sum_{j=0}^{n} \beta_{j+1} \lambda_i^j \cdot exp\left[\sum_{k=0}^{m} \beta_{k+n+1} f_k(\lambda_i) \right] \right) + \left(\frac{\pi}{\mu_0 F_s(\lambda_i)} \cdot \beta_0 e^{-\frac{1}{2}\left(\frac{\lambda_i - \mu_h}{\sigma_h}\right)^2} \cdot exp\left[\sum_{k=0}^{m} \beta_{k+n+1} f_k(\lambda_i) \right]^{\frac{\mu^{-1}}{(\mu^{-1}+\mu_0{}^{-1})}} \right) \qquad (3.3)$$

This cost function is optimized using the Levenberg-Marquardt algorithm (Levenberg 1944; Marquardt 1963). This is an iterative method which starts from an initial guess for **β** and then adjusts this vector by a vector **δ** for each iteration. If there are local minima in the cost-function, then the choice of initial parameter values matters. In this case, all initial parameter values are set to zero. These parameters are then used to calculate the cost-function. From this, **δ** can be calculated by solving a linear set of equations which can be expressed in vector notation as follows as shown in eq. (3.4).

$$(\boldsymbol{J}^T \boldsymbol{J} + \alpha \boldsymbol{I})\boldsymbol{\delta} = \boldsymbol{J}^T \left(\frac{\boldsymbol{\rho_m} - \boldsymbol{\rho_p}(\boldsymbol{\beta})}{\boldsymbol{\sigma_m}} \right) \qquad (3.4)$$

Here, **J** is the Jacobian matrix of the cost-function; **I** is the identity matrix; and α is a relaxation factor. From this **δ**, a new **β** is calculated as shown in eq. (3.5).

$$\boldsymbol{\beta}_{new} = \boldsymbol{\beta}_{old} + \boldsymbol{\delta} \qquad (3.5)$$

This process is iterated until a prescribed minimum relative difference between β_{new} and β_{old} is reached (e.g. $1*10^{-8}$) or if the number of iterations exceeds a prescribed maximum (e.g. 4200). In the latter case, the retrieved pixel is flagged as non-converging and later filtered out.

The Jacobian of the cost function can be derived analytically and this speeds up the algorithm as otherwise the Jacobians would need to be numerically derived as well.

3.3 Principal Component Algorithm

3.3.1 Principal component analysis

We apply a so-called principal component analysis (PCA) to the set of vegetation free reference spectra. A PCA is a method to identify the orthogonal components from a multidimensional correlated dataset. It is often used to reduce a dataset with many dimensions to a smaller number of dimensions over which the variation is most significant.

The most straightforward way to calculate the principal components is as follows. First arrange all the data in a matrix X (e.g. a matrix with a number of rows equal to the number of spectra and a number of columns equal to the number of spectral bins). Then calculate the covariance of each dimension (e.g. a spectral bin) with every other dimension and arrange this in a matrix. Then calculate the eigenvectors and the corresponding eigenvalues of this covariance matrix. There will be a number of eigenvectors equal to the number of spectral bins and each eigenvector has a number of dimensions equal to the number of dimensions of the data (e.g. the number of spectral bins).

These eigenvectors are referred to as the basis vectors or the loading vectors of the principal components. The corresponding eigenvalues are a measure of the energy present in this basis vector, or in other words, the variance explained by this basis vector. Arranging the loadings from highest eigenvalue to lowest and projecting the original data onto the new basis yields the so called principal component score vectors. In meteorology these are also known as empirical orthogonal functions (EOFs).

The first n number of components are the principal components and by rejecting the lower components it is possible to reduce the dimensionality of a dataset with a minimal loss of information. Of course, which values of n give an acceptable loss of information depends on the dataset in question.

In our case, we are interested in the loading vectors, which are used in the minimization scheme (section 3.2) and applied to all target spectra.

3.3.2 Iterative principal component scheme

The amount of training data points is very large (e.g. in the order of 30000 spectra for a year worth of data from the Sahara) and thus calculating the principal components using the covariance method explained above is computationally problematic. Since we will only use a relatively few principal components (about 35) we use the so-called nonlinear iterative partial least squares (NIPALS) algorithm to solve for the PCs (Wold 1966; Wold 1973). This allows the principal components to be calculated sequentially in order of decreasing explained variance without having to calculate the entire covariance matrix. A short description of the NIPALS algorithm is given below.

We start by defining a matrix X as the array of reference spectra with each row a single measurement of the spectrum and each column a single wavelength bin. Here, the spectra are all centered and normalized first. We define the principal component loadings as vectors where $\boldsymbol{p}_j$ is the j^{th} principal component loading. The principal components are numbered in order of descending percentage of variance explained.

The algorithm is an iterative procedure, where the first principal component $\boldsymbol{p}_1$ is calculated first before the second can be calculated, etc. We initialize by defining:

$$j = 1;\ X_1 = X \qquad (3.6)$$

Then we take a vector $\boldsymbol{t}_1$ to be the first column of X_1. In fact, any column will do for the purposes of this algorithm, but in our implementation we simply chose the first. This vector is the first estimation for the first principal component score vector. Then we can assign an estimate of the first principal component loading as follows:

$$\boldsymbol{p}_j = \frac{X_j^T \boldsymbol{t}_j}{\left\| X_j^T \boldsymbol{t}_j \right\|} \qquad (3.7)$$

From that we can obtain a new estimate for the score vector t_1:

$$\boldsymbol{t}_{j,new} = X_j \boldsymbol{p}_j \qquad (3.8)$$

This procedure is then repeated until the difference between $\boldsymbol{t}_1$ and $\boldsymbol{t}_{1,new}$ is reasonably low. We quantify this as the dot-product of the difference vector with itself. In our case we used a threshold of 1E-8. If this point is reached, we consider p_1 to be a reasonable estimate of the true first principal component loading and $\boldsymbol{t}_1$ to be the first principal component score. Then we proceed to the next principal component.

$$X_{j+1} = X_j - \boldsymbol{t}_j \boldsymbol{p}_j^T \qquad (3.9)$$

From this we can repeat the procedure of equations (3.7) and (3.8) with j=j+1. We continue this until the required number of principal components is reached.

3.4 Reference Spectra Selection

To obtain a set of fluorescence free reference spectra, we must find scenes that can reasonably be considered to be free of vegetation. We have considered three different kinds of scenes that fulfill that requirement: desert scenes, cloudy ocean scenes and ice sheets. We will explain the selection procedure for each case in the following paragraphs. In all three cases, the selected scenes are aggregated over an extended period of measurements. Initially they were aggregated over one month, and in the more recent experiments over one year (2013). We have also considered combinations of the three primary types of reference scenes. The effect of the choice of reference set will be discussed in **section 5.1**.

3.4.1 Desert

In order to select suitable desert scenes from the global level 1b dataset we defined a latitude-longitude box that covers the inner part of the Sahara region. This is the largest area of near-continuous desert on the planet and, considering the coarse spatial resolution of the GOME2 instrument, using this region instead of smaller or more irregularly shaped desert areas reduces the risk of including pixels with residual vegetation. However, there are still some patches of desert within this area that contain some very-sparse desert vegetation as well as occasional oases. Ideally, only the completely barren dune seas are used. In order to do this, we have incorporated in the algorithm a selection procedure that compares the corner coordinates of the GOME-2 pixels with a 1

km x 1 km land cover climatology database. We used the USGS global land cover Characteristics database v2.0; specifically the 'seasonal land cover regions for Africa' database[2].

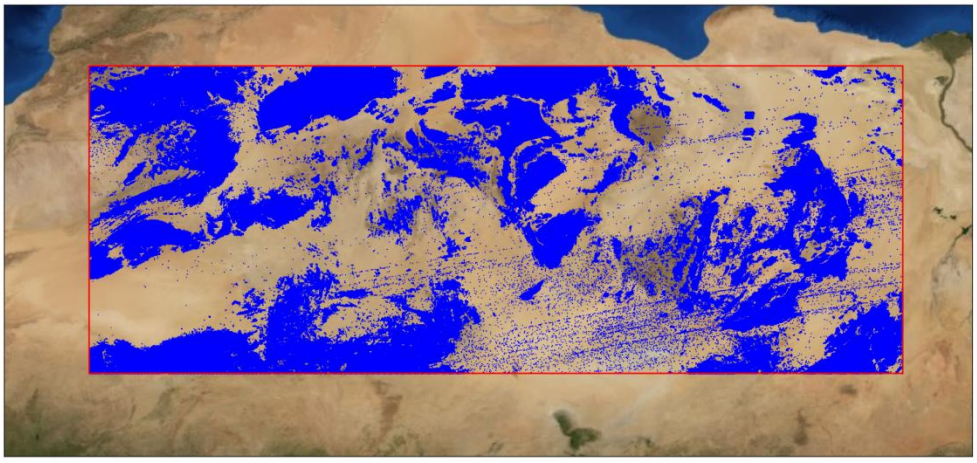


Figure 3.2: Area from which the desert reference dataset is acquired. The red boundary indicates the preselected area corresponding with the inner Sahara. The blue dots are the climatology grid cells that are not classified as completely barren by the USGS land cover database.

The selection algorithm is designed to be rather restrictive in its admission of reference pixels. Only GOME2 pixels which contain absolutely no vegetation according to the land cover climatology within their footprint area are allowed. Figure 3.2 shows the rejected climatlology gridcells together with the boundary of the reference area overlaid on a true color image of the Sahara. For a description of the selection algorithm, we refer to appendix A.2.

Furthermore the reference pixels are rejected when they are too cloudy based on data from the FRESCO algorithm that is included within the GOME-2 level1b product files. The FRESCO algorithm provides the effective cloud fraction and the cloud top height/pressure. The algorithm uses the O_2-A band in its calculations (see Koelemijer et al. (2001) for more details). An important concept to keep in mind here is that of effective cloud fraction. It is not possible to determine the actual cloud fraction and the cloud optical thickness separately and therefore the effective cloud fraction is a measure of both. We have chosen to reject pixels which have cloud cover fraction above a certain threshold (e.g. 0.4).

The restrictiveness of this selection limits the number of pixels that are available. However, this is balanced by aggregating pixels over a timespan of a month or even a year instead of per day. Aggregation over a year instead of over a month also takes into account the seasonal changes in atmospheric conditions over the region. One significant seasonal effect, for example, is the variable H_2O column. These changing conditions can have a significant effect on the representativeness of the reference scenes. This is further discussed in the results section.

3.4.2 Cloudy Ocean

Another possible reference category is that of cloudy ocean scenes. To obtain a diverse dataset of reference scenes, a random sampling of the global oceans was performed. To determine if a given pixel is in the ocean or not, the pixel coordinates where compared with the GSHHG coastline database (Wessel & Smith 1996) that is included in the python software package SciPy. The random

[2] Data made available by US Geological Survey. URL: http://edc2.usgs.gov/glcc/tablambert_af.php

sampling is done to reduce the total amount of pixels that are included in the principal component analysis when accumulating the reference spectra over a year. This serves to reduce unnecessary computational expense and also to avoid a biased reference set when combining with desert and/or ice pixels.

The inherent problem with using reflection spectra of oceans in the far red/ near infrared part of the radiation spectrum is that the ocean is almost completely absorbent. This would not be representative for the brighter vegetated ground surfaces. Therefore, instead of the ocean surface, the top of clouds are used, which are far more reflective. However, low hanging clouds are strongly preferred since otherwise the obtained optical depth is not actually representative of the entire atmospheric column. The algorithm therefore uses the FRESCO algorithm to select both a minimum cloud fraction (e.g. 0.9) and a minimum cloud top pressure (e.g. 900 hPa).

When looking at ocean pixels another phenomenon needs to be filtered for: the so-called sun-glint. This is the specular reflection of the sun on the earth surface and thus it violates the Lambertian assumption used in our radiative transfer model. This is further explained in appendix A.3.

3.4.3 Ice/Snow

Ice sheets are some of the most barren landscapes on the planet and would therefore appear to be a very good candidate for obtaining vegetation free pixels. There are some difficulties however. First, ice sheets can be up to several kilometers thick and thus have a high altitude. This could potentially make the references not representative for lower laying landscapes. Second, they are also some of the driest places on earth; the atmospheric column contains less moisture than for most deserts. Third, they are found only at high latitudes, which limits the availability of year round data and the data available is biased towards very high solar zenith angles which might not be properly representative of other latitudes.

The first part of the selection algorithm, similar to the other two categories, is to define a latitude-longitude box that includes a suitably large continuous area containing the landscape type we are looking for. We have chosen an area containing Greenland and the Canadian Arctic Archipelago.

Secondly, a land cover climatology database is used to select only the pixels that contain just ice, similar to what is done with the Sahara desert. In this case the land cover characteristics database for North America is used[3]. In Figure 3.3 the reference area and the rejected climatology grid cells within this area are shown.

[3] Data made available by US Geological Survey. URL: http://edc2.usgs.gov/glcc/tablambert_na.php

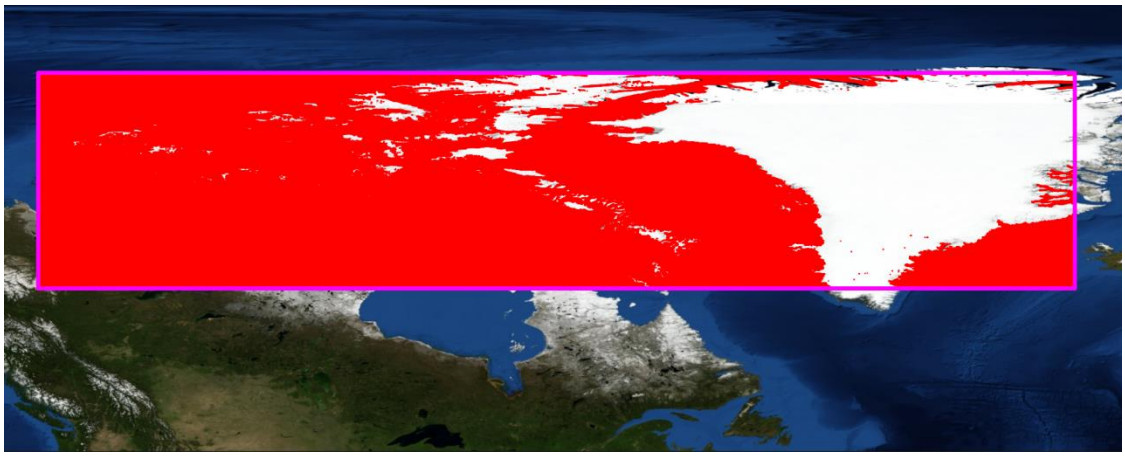


Figure 3.3: Reference area from which ice scenes for the training dataset are selected. The purple boundary indicates the preselected area corresponding to Greenland and the Canadian Arctic Archipelago. The red dots are the land cover grid points that are not classified as ice or completely barren.

Thirdly, the pixels have to be filtered for unsuitable cloud cover. This presents another problem, since the FRESCO algorithm used for determining cloud cover does not provide meaningful results over ice/snow scenes. Instead, the only parameters that are available are the surface pressure and the scene pressure. However, a scene pressure that is close to the surface pressure would indicate that cloud cover is only slight, or if there is more cloud cover, it consists of mostly very low altitude clouds. Therefore, we chose to define a maximum difference between scene pressure and surface pressure as selection criterion. For this criterion we initially took 100 hPa, but this value could possibly be optimized.

3.5 Target pixel filtering

The vegetated target pixels need to be filtered as well, for example for the presence of clouds. Some filters are applied to the data before and some after the retrieval process.

The filters that are applied before the retrieval:

- The target level 1b data is filtered for extremely high solar zenith angles. As a cutoff we have chosen 70^{o}.
- Furthermore the target pixels are rejected when they are too cloudy based on data from the FRESCO algorithm (effective cloud fraction threshold 0.4). Pixels where the FRESCO algorithm did not yield meaningful results (e.g. due to ice/snow) are already flagged in the level 1b data files. We exclude these flagged scenes from the retrieval as well.

After the retrieval, some more filters are applied:

- The residuals are checked, and those scenes which have a spectral RMS residual over the entire spectrum of over 1% are rejected.
- Also, any scene retrievals that take an anomalous number of iterations (> 4200) to converge are aborted and rejected. This is actually much larger than the typical number of iterations needed for convergence (roughly about five iterations).

All scenes that do pass these filters are then binned in a regular latitude-longitude grid of 0.5^{o} x 0.5^{o} and then the temporal average is taken per grid cell over a period of a month.

3.6 Retrieval Settings

To conclude the algorithm description we will list the typical values of the retrieval settings. Some of which are also varied in different retrieval sensitivity experiments.

Table 3.2: The algorithm settings and filters used in the baseline retrieval.

Setting	value
Sunglint threshold	> 18^{o}
cloud fraction threshold for target pixels	< 0.4
SZA threshold	< 70^{o}
NIPALS convergence threshold	< 1E-8
Levenberg-Marquardt convergence threshold	< 1E-8
Number of basis vectors	35
Order of surface albedo polynomial in SiFretrieval	4
Retrieval spectral window	712-780 nm
Absorption free spectral window	712-713,748-757,775-785 nm
Max. RMS residual	0.01
Peak of fluorescence curve	737 nm
Sigma width of fluorescence curve	21.2 nm

4 Baseline Retrieval Results

In this chapter we will discuss the results of the baseline retrieval algorithm. This baseline was defined after a number of initial retrieval experiments. We will present a few of the retrieval experiments in the next chapter to support the choices we have made.

The baseline algorithm that we have defined has been applied to all the GOME2 data of 2013. We have also applied the algorithm to some parts of earlier years. The baseline uses only the data from the Sahara for its set of reference spectra. For SiF retrievals from spectra observed in 2013 we use PCs that are determined from reference spectra collected over the entire year of 2013. For the surface reflectance in the retrieval a 4^{th} order polynomial is used. The retrieval window is from 712 nm to 780 nm. The 35 principal components with the highest explained variance are used. The retrieval is only applied to spectra from pixels with a FRESCO cloud fraction of 0.4 or lower. Reference spectra are subject to the same cloud fraction criterion.

4.1 Spatial Patterns

Retrieval results are obtained for both continents and oceans. We have included non-vegetated target pixels here to check whether the retrieved SiF is indeed zero and for further error analysis. The resulting sun-induced fluorescence values are ranging roughly from 0 to 7 mW/sr/m^2/nm at the peak of 737 nm over most of the continents. The global average of SiF including oceans is highest in July (μ_g = 0.688) and lowest in April (μ_g = 0.292). Over the oceans, the SiF values are roughly around zero and show no clear spatial pattern. There are clear and systematic spatial patterns visible on the continents. Below, we will discuss the spatial patterns that can be seen in July 2013 (see Figure 4.1) and what they signify.

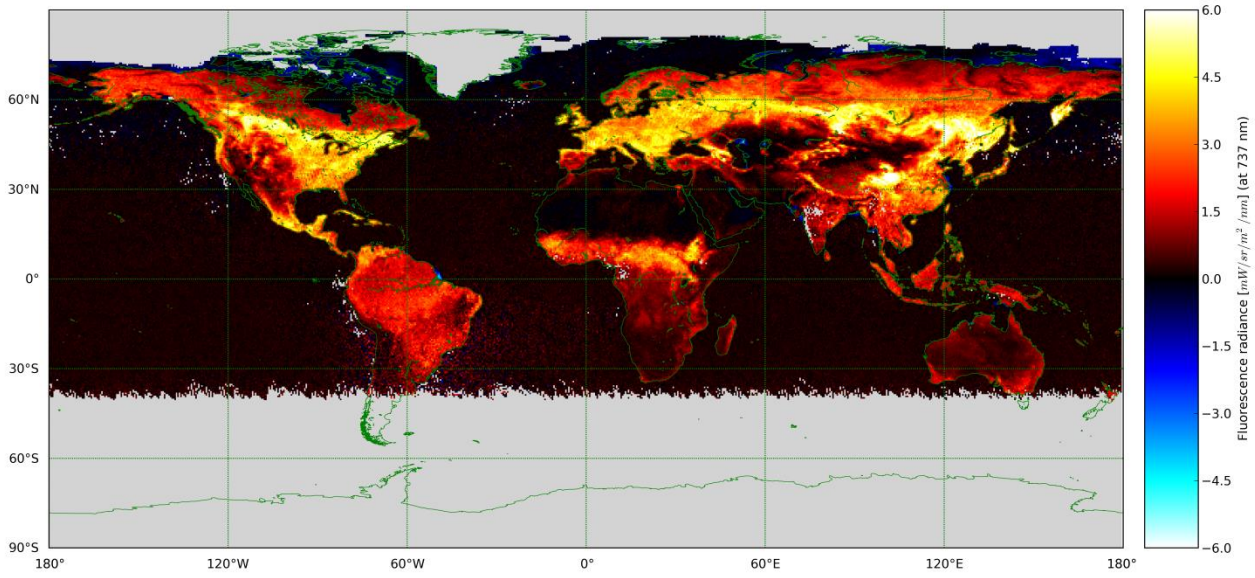


Figure 4.1: Retrieval of sun-induced chlorophyll fluorescence at 737 nm for July 2013 using the baseline retrieval setup (desert reference spectra only). Values are averaged over the 0.5° x 0.5° grid-cells. Grey pixels indicate no data.

Within the Sahara, Arabian Desert, Gobi desert, the Central Asian steppes, western United States, Iranian plateau, southwestern Andes region and the Australian Outback the SiF values are similar to

those values found in the ocean pixels. This is expected, given that these are very arid regions with little or no vegetation.

Within the other regions observed SiF values are generally positive and well above the suggested noise level (see below). The highest values can be found in July in the Siberian and Western Canadian Taigas as well as the Chinese Sichuan Basin and in June in Europe and the eastern United States. Mid-latitude and high latitude regions on the northern hemisphere in the months May, June, July and August generally show higher SiF values than equatorial regions throughout the year.

The retrieved values within these fluorescence emitting regions are spatially and temporally very heterogeneous. This is especially true in regions where narrow strips of fertile agricultural land are situated within a larger arid area, such as in North Africa and the Fertile Crescent (see Figure 4.2). Here, significantly higher retrieved values are found over the known agricultural land than the surroundings. These agricultural lands also show the highest values in March, April and May, while becoming significantly lower in subsequent months. During October, November and December the values in the Fertile Crescent no longer rise above those of the surrounding desert.

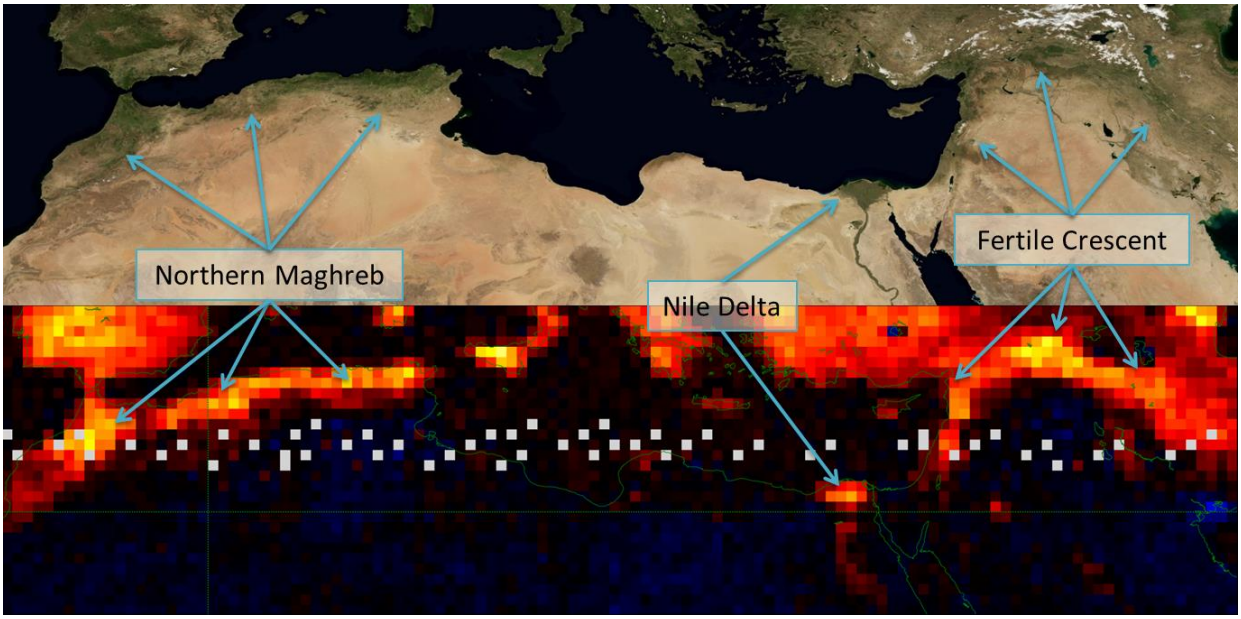


Figure 4.2: Retrieval over North Africa and the Middle-East. The lower half shows the SiF retrieval results (April 2013), while the upper half shows the same area in true color in April 2004[4]. Indicated are some significant agricultural areas bordered by desert and steppe. Colors in the SiF map are the same as in Figure 4.1. Grey pixels indicate no data.

These spatial patterns observed provide indications that we can indeed see a vegetation induced signal in these retrieval results as there is a clear contrast between fertile and infertile lands as has been shown in the above paragraph.

Some pixels show significantly negative SiF values. These have no physical interpretation; therefore they indicate a circumstance where the retrieval algorithm cannot fit the model towards the observations properly. Surprisingly, most of these negative values are found in areas of shallow and/or murky waters such as: The Northern Caspian Sea, the Aral Sea, the Balkhash Lake, the North-American Great Salt Lake, Lake Titicaca, the Bahamas and the mouths of the Amazon, Indus, Yellow

[4] Blue Marble: Next Generation by NASA Earth Observatory (Stöckli et al. 2005)

river and Mekong. Especially the Caspian Sea clearly illustrates how this phenomenon is depth related as can be seen in Figure 4.3.

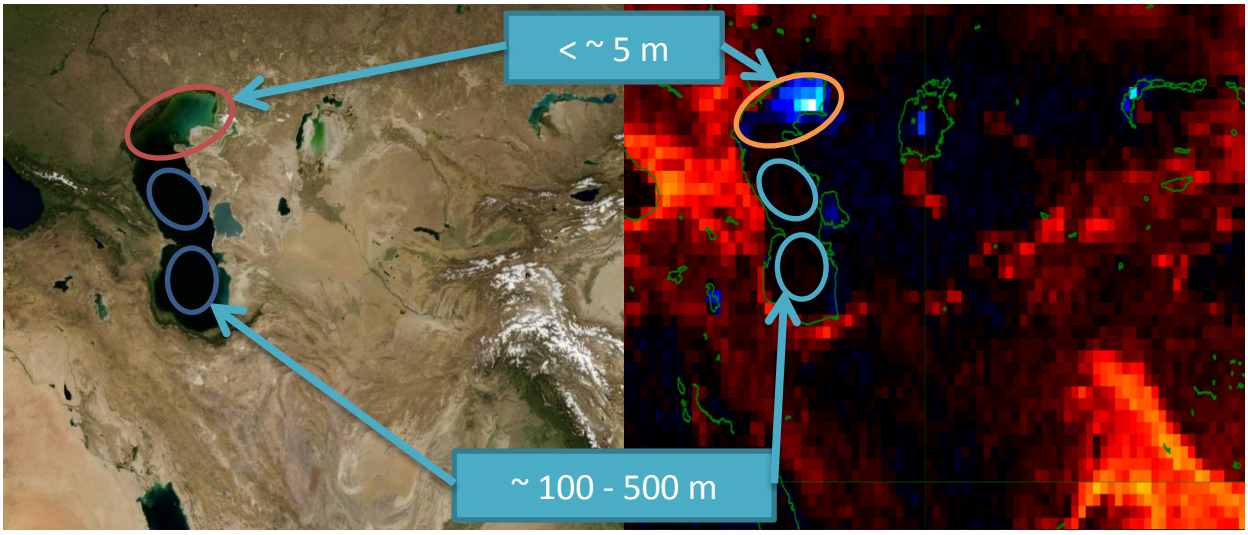


Figure 4.3: Retrieval over the Central Asian Lakes region. The right side shows the SiF retrieval results (September 2013), with red hues indicating positive values and blue hues indicating negative values. The left side shows the same area in true color taken from the NASA Blue Marble Next Generation project (September 2004). Note especially the Caspian Sea, which has a large depth difference between the center and South of the basin and the North of the basin.

4.2 Temporal Patterns

Strong seasonal variations in retrieved SiF values can be found in most of vegetated areas as illustrated in Figure 4.4.

In Western and Central Europe retrieved SiF values start increasing from March onwards and peak in June. Afterwards they decrease again until a minimum is reached in November. During the winter months (December, January, February) little data is available due to the lack of sunlight over most of these regions at the overpass time. However the available data suggest a constant value throughout this period. Similar temporal patterns exist in other temperate and subtropical regions on the northern hemisphere, notably the eastern US, south-eastern China and the northern Mediterranean coast.

In the continental climate zones of Eastern Europe/ Russia and southern Canada/north-eastern US the retrieved SiF values only rise above the background level from May onward. They also peak in June and are mostly back to background level by October.

In the Boreal climate of northernmost Eurasia and North-America SiF values above background level do not appear until June. They have a peak in July and are already back to background level by September.

In the lower latitudes, significant SiF values are retrieved in all months, but the location where the highest values are found in a given month shifts gradually throughout the year in a roughly North-South motion. In South America the SiF values within the Amazon basin are roughly constant throughout the year. However, the savanna in southern Brazil has a minimum of fluorescence in September and a maximum in January.

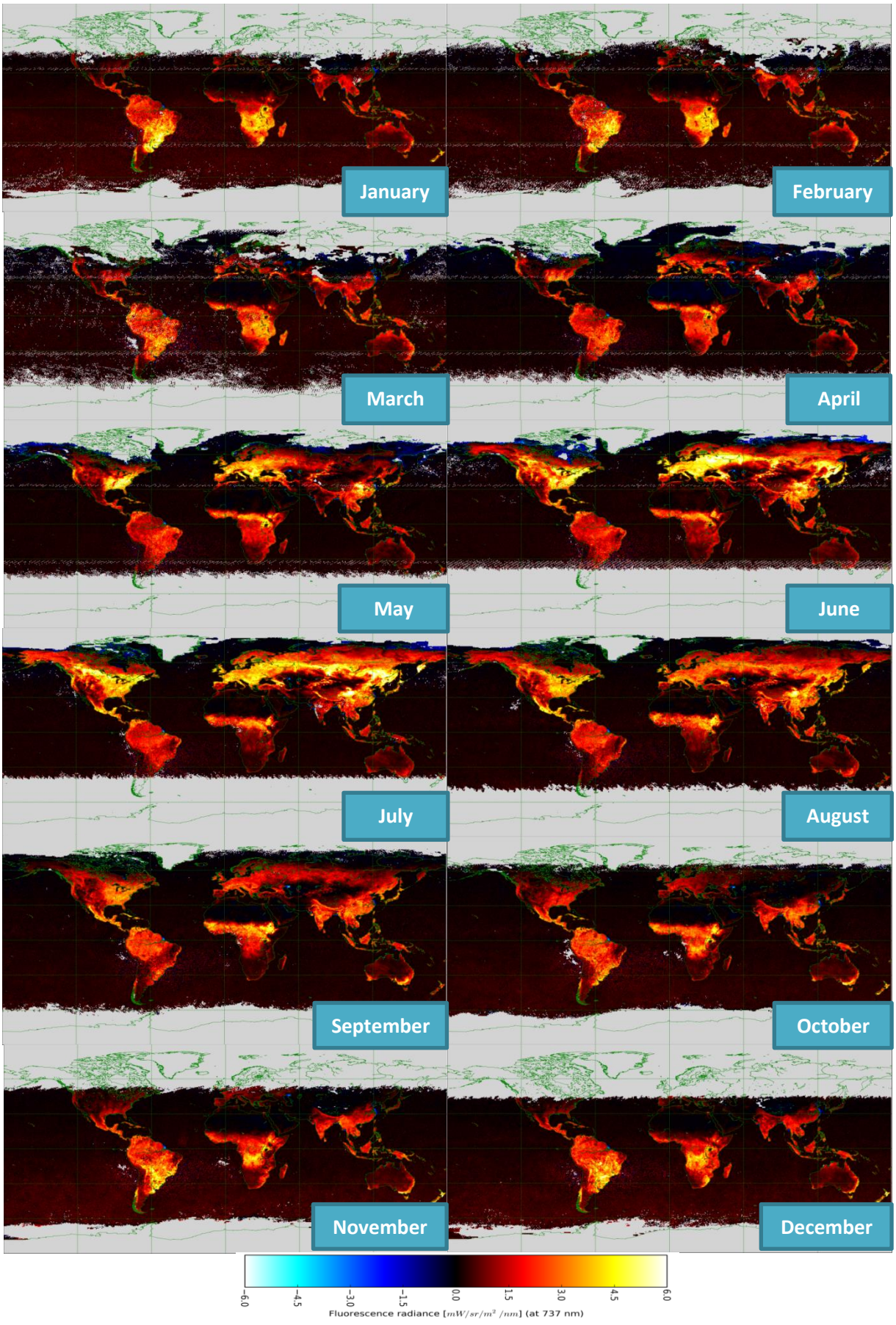


Figure 4.4: Retrieved monthly mean SiF values in 2013. All maps use the same color scale.

In Sub-Saharan Africa, the savanna regions northward of the equator and southward of the equator are showing seasonally varying levels of SiF intensity that are in anti-phase with each other. The northern savanna has a minimum in fluorescence during January/February and a maximum in August/September, and vice versa for the southern savanna.

Temporal patterns can also be observed in South Asia and Oceania. In south-western and south-eastern Australia the SiF values are high in September, October and November and are weakest in May. In the north of Australia the SiF values are highest in January until April and are very low for the rest of the year.

All these examples illustrate how the SiF values obtained over different regions follow the dominant seasonal climate induced vegetation patterns for that region. This could be an indicator that the retrieved SiF values are related to the actual vegetation activity.

The seasonality of the fluorescence retrievals can be further illustrated by categorizing pixels by land-use type. This is done by using a global land cover database (land cover 2000 version 1.1 published by the JRC). This database is then reprojected and interpolated to our SiF grid and is used to categorize each gridcell. Because the land cover database uses a much finer grid resolution than our grid, we have to scale the land cover database to our gridcells. We did this by choosing the most frequently occurring land type within each gridcell as the type for the entire gridcell. Then, the spatial averages are taken over all grid-cells that fit into a certain category within a single hemisphere and the time dependency of these averages is displayed in Figure 4.5.

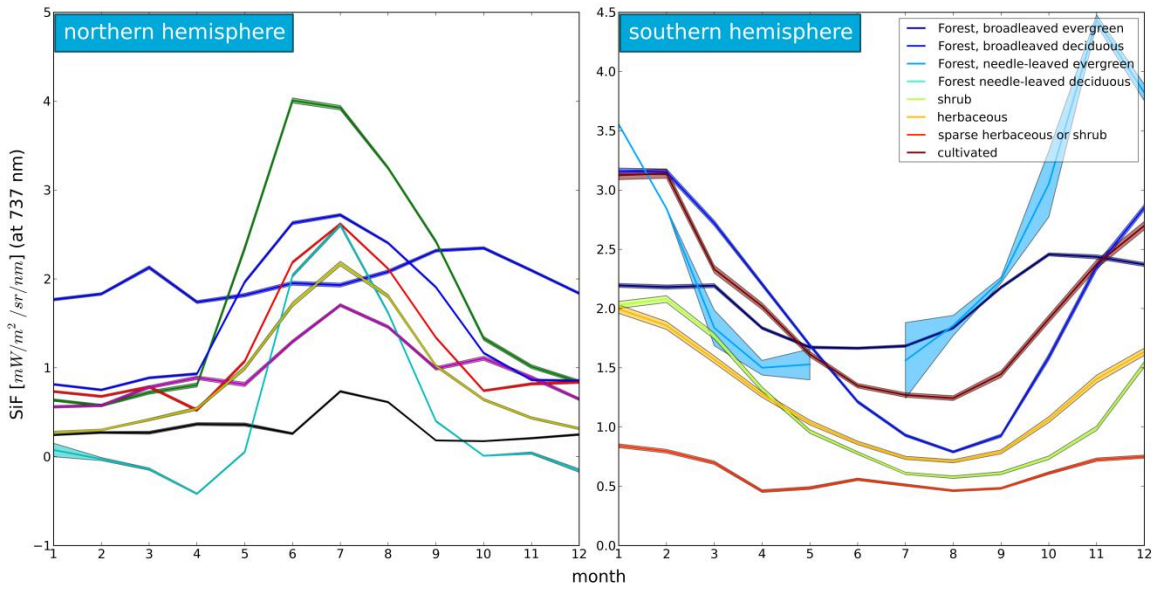


Figure 4.5: Time-series SiF for the year 2013. All values are monthly values. Each line represents the average of the fluorescence values found for a particular land-use type. On the left are the time-series for the northern hemisphere. To the right are the time-series for the southern hemisphere. The shaded areas represent the extent of the standard deviations taken over the samples.

Here it can be seen that tropical rainforests (evergreen broadleaved forest) have relatively constant SiF values in the northern hemisphere. Tropical forests in the southern hemisphere show a peak in February and in November. Evergreen needle leaved forests in the Southern hemisphere show a

peak in November. All other categories show some manner of seasonal dependence with a peak in July in the northern hemisphere and a peak in February in the southern hemisphere.

The lowest values throughout the year are in the sparsely vegetated areas, except for the deciduous needle leaved forests which reach zero in May and October. The highest values in the northern hemisphere are found with broadleaved deciduous trees (temperate forests) from May until September, while the rest of the year the tropical forests have higher SiF values.

4.3 Error Analysis

In order to get an idea of the amount of noise in the retrieval results and to better compare these results with the retrieval experiments with varying retrieval settings, we calculate some basic statistic values for this data. These values are computed using only data from September 2013 in order to better compare these baseline values to the results of the other retrieval experiments that are presented in section 5.

A tentative error analysis can be performed by assuming the true values for the oceans to be zero and then computing the mean and standard deviation from all the ocean pixels. Then, the standard deviation would be equal to the stochastic noise and the mean would be a measure for the offset of the retrieved value compared to the true value. In order to get a noise value that can be compared between retrieval experiment results with very different SiF magnitudes, we can also define a 'relative noise' that is normalized with respect to the mean SiF value over the continents:

$$RN = \frac{\sigma_o}{\mu_c} \cdot 100\% \qquad (4.1)$$

Here, σ_o is the standard deviation over all ocean pixels and μ_c is the mean over all land pixels. The average SiF value for September 2013 over all the continents is 1.070 mW/sr/m^2/nm and the standard deviation around the mean is 1.077 mW/sr/m^2/nm. The average for all ocean areas is 0.180 mW/sr/m^2/nm and the standard deviation around the mean is 0.222 mW/sr/m^2/nm (again for September 2013). This would then suggest that the stochastic noise is in the order of 0.222 mW/sr/m^2/nm. The relative error in SiF over the continents would then be, on average, 21 %.

The daily retrieved SiF values are averaged over a month which increases the reliability of the results at the cost of the temporal resolution. The number of data points is generally less than 30 because the filters (see **section 3.5** and **section 3.6**) remove a number of data points (e.g. because they are cloudy scenes). The number of samples per gridcell is shown in Figure 4.6 for two months. It can be seen here that there are a few areas with no samples at all surrounded by a broader area with very few samples. These are all areas with consistently high cloud cover fractions throughout the month. This means that even when averaging over a month it is not always possible to have complete global coverage with this level of filtering.

Furthermore, there are regularly spaced meridional bands with higher sample density than the immediate surroundings. These are caused by the narrow swath measurements that GOME2 makes once per month and that are included in the monthly averages.

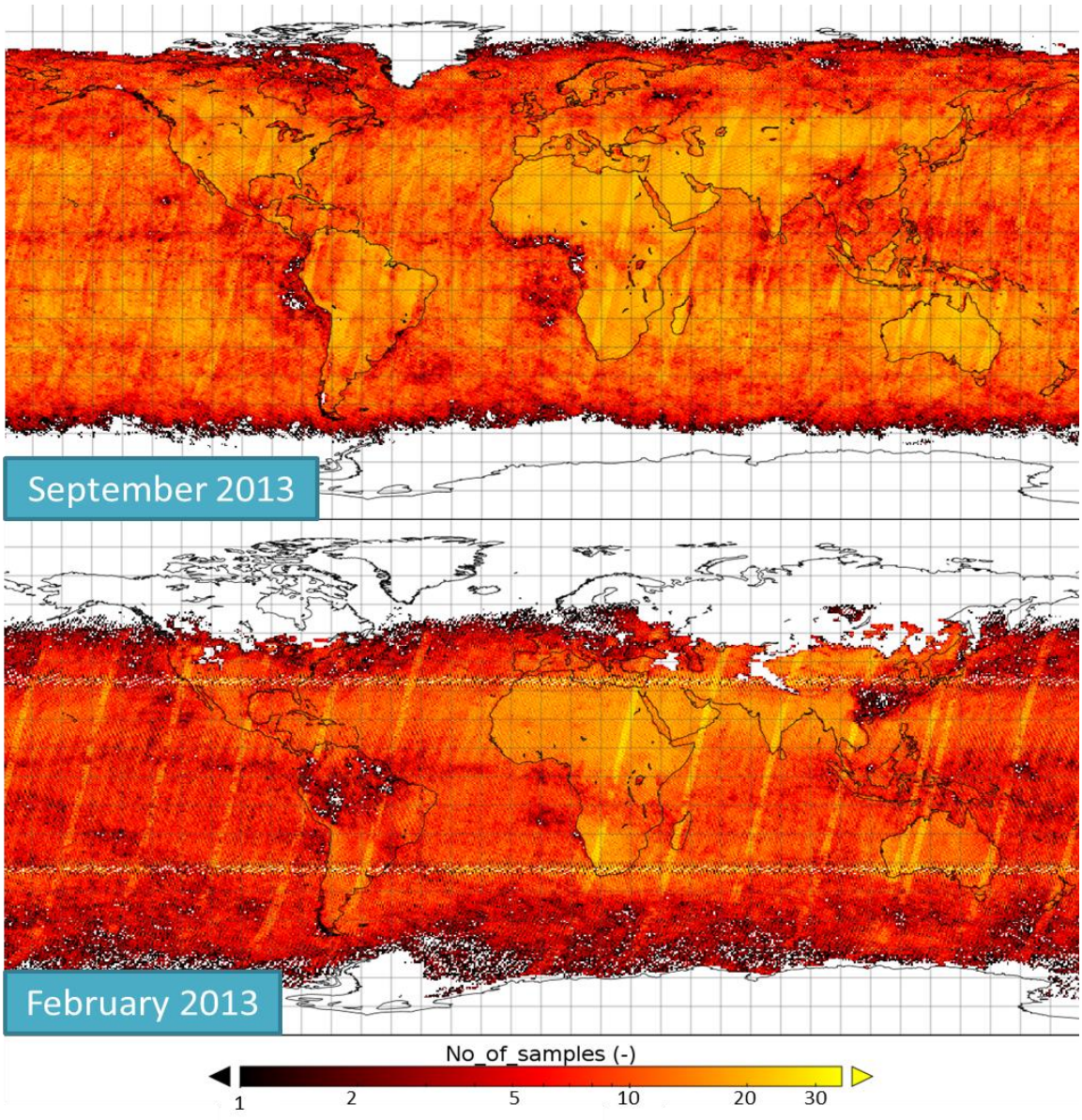


Figure 4.6: Sample density for the monthly averaged GOME2 SiF retrieval. The plot scale is logarithmic and white values mean zero samples. The largest number of sample in a single gridcell is 39 in February and 32 in September.

A third item that is striking, are the zonal oriented bands in the February sample density map that has gridcells with no samples occurring in a regular pattern. These are artifacts caused by the mapping of the pixel footprints to the regular lat-lon grid. This could perhaps be improved by using a more fitting interpolation scheme. These artifacts are visible for all months prior to July and they are not present for the months of July and beyond. This is because the nominal GOME-2 pixel footprint size has been altered from July 2013 onwards. Note also that sometimes the number of samples in a grid cell is actually higher than the number of days in the month because the center coordinates of multiple pixels may fall into the same lat-lon gridcell.

5 Retrieval Experiments

There are many different settings that can be used for the retrieval and it is not always obvious what values yield the most accurate results. We have performed several retrieval experiments spanning one month of data (September 2013) to define a reasonable baseline setup. The retrieval experiments illustrate some of the sensitivities of the retrieval: The area and month from which the reference spectra used for the PCA are taken (**section 5.1**), the number of PCs included in the forward model (**section 5.2**), the order of the surface albedo polynomial in the retrieval and the number of wavelength bins included in the retrieval (**section 5.3**).

5.1 Reference Datasets

5.1.1 Reference Areas

First we will discuss the effect of using different kinds of vegetation free areas to build the set of reference spectra from which to compute the principal components. We have compared the use of ice/snow, desert and ocean areas. All other retrieval settings are the same as in the baseline. Reference pixels are collected from the entire year.

When using only snow/ice reference pixels in the set of reference spectra, the results are quite different than when using desert reference pixels, as shown in Figure 5.1, where the global maps for September 2013 are compared. What can immediately be seen is that almost all areas that had high fluorescence emissions values in the baseline do now yield negative values. This includes tropical forests and savannas as well as temperate and boreal zones. The values in tropical areas are much more negative, however, than those in northern Eurasia and North America. Exceptions to this are the Mediterranean coastline and the temperate zones of Southern Oceania which mostly show slightly positive values. Most arid areas which yield fluorescence values around zero in the baseline retrieval actually yield significant positive values with this retrieval setup. These areas include the Sahara, Arabia, Iranian Plateau, Mongolia, southwestern United States and many parts of Australia. The average SiF value over all land areas combined is -0.838 mW/sr/m^2/nm and the standard deviation is 2.527 mW/sr/m^2/nm. The average over the oceans is 0.051 mW/sr/m^2/nm and the standard deviation is 0.443 mW/sr/m^2/nm. It seems then that there is little offset in the signal over the ocean as compared to the baseline, however it should be noted that there seems to be a spatial pattern present in the ocean for this retrieval that was not present in the baseline, which makes the situation less clear. There is a band of negative SiF values that covers the equatorial oceans. There are also patches of slightly positive values off the western coasts of South America and Africa. This spatial pattern also makes interpreting the standard deviation over the oceans as noise problematic.

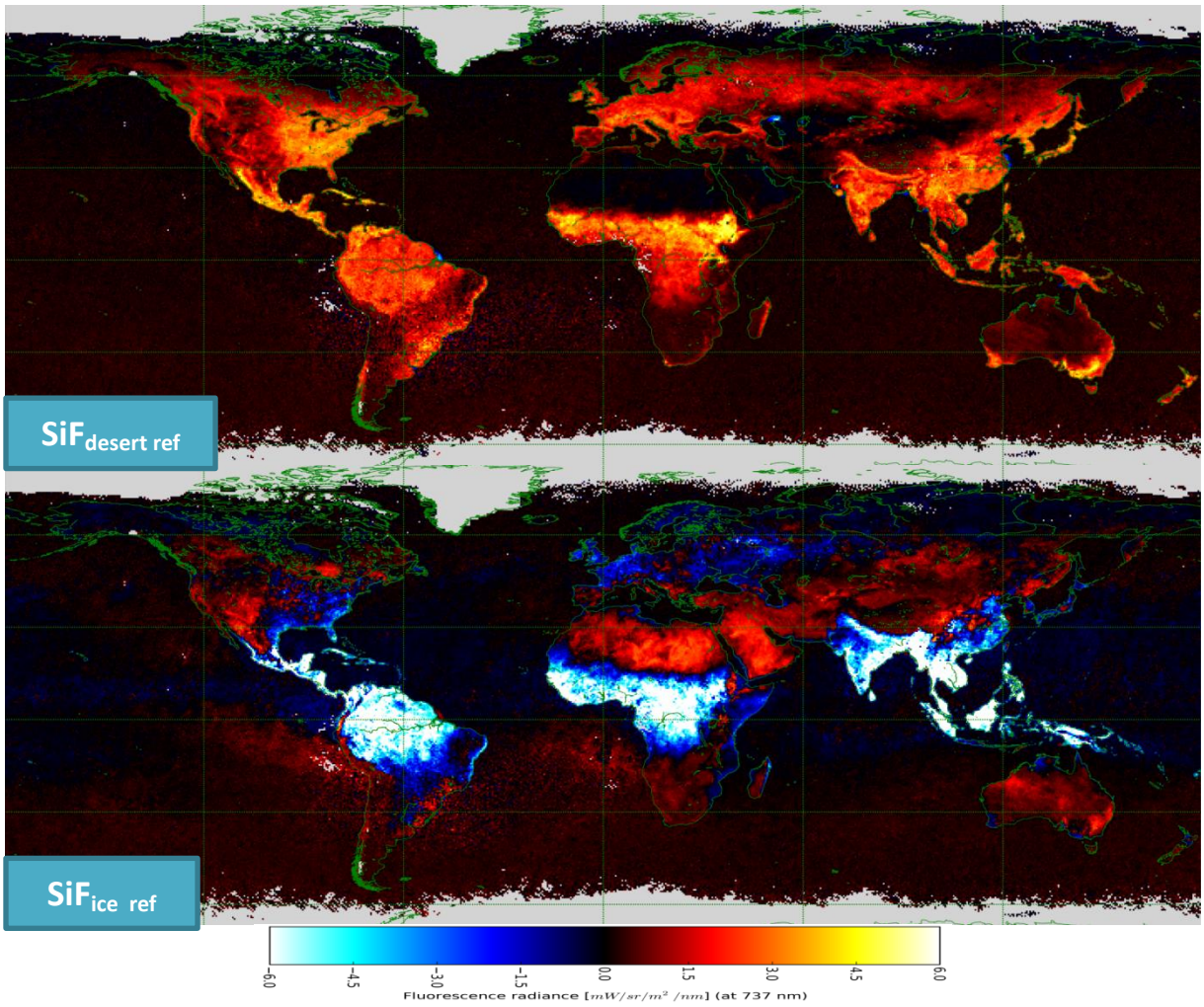


Figure 5.1: Retrieval of SiF using only snow/ice scenes for the reference dataset compared with retrieval using only desert in the reference dataset. Reference data is collected over 2013. Values are monthly means over the 0.5° x 0.5° grid-cells. The month retrieved is September 2013. Grey pixels indicate absence of data. The indicated fluorescence values are the values at 737 nm.

When using only cloudy ocean pixels in the reference set (as seen in Figure 5.2), results are more similar to the baseline results. There are substantial differences however. Just like with the sow/ice reference retrieval result, there are areas with negative SiF retrieval results where there are positive emissions found in the baseline retrieval. These areas are far more limited in extent though, and for most of the globe the SiF retrievals show a similar spatial pattern as for the baseline. The areas where negative values are found are mainly in the Tropical areas of South-East Asia, Especially in the Brahmaputra/Ganga delta and the Mekong delta. Negative values are also found to a lesser extent along the course of the Amazon River.

Slightly negative values are also observed in arid areas, especially the Sahara desert. Slightly negative values are also observed over the ocean surrounding Indonesia/Malaysia. Similarly to the baseline, there are also negative values over shallow lakes.

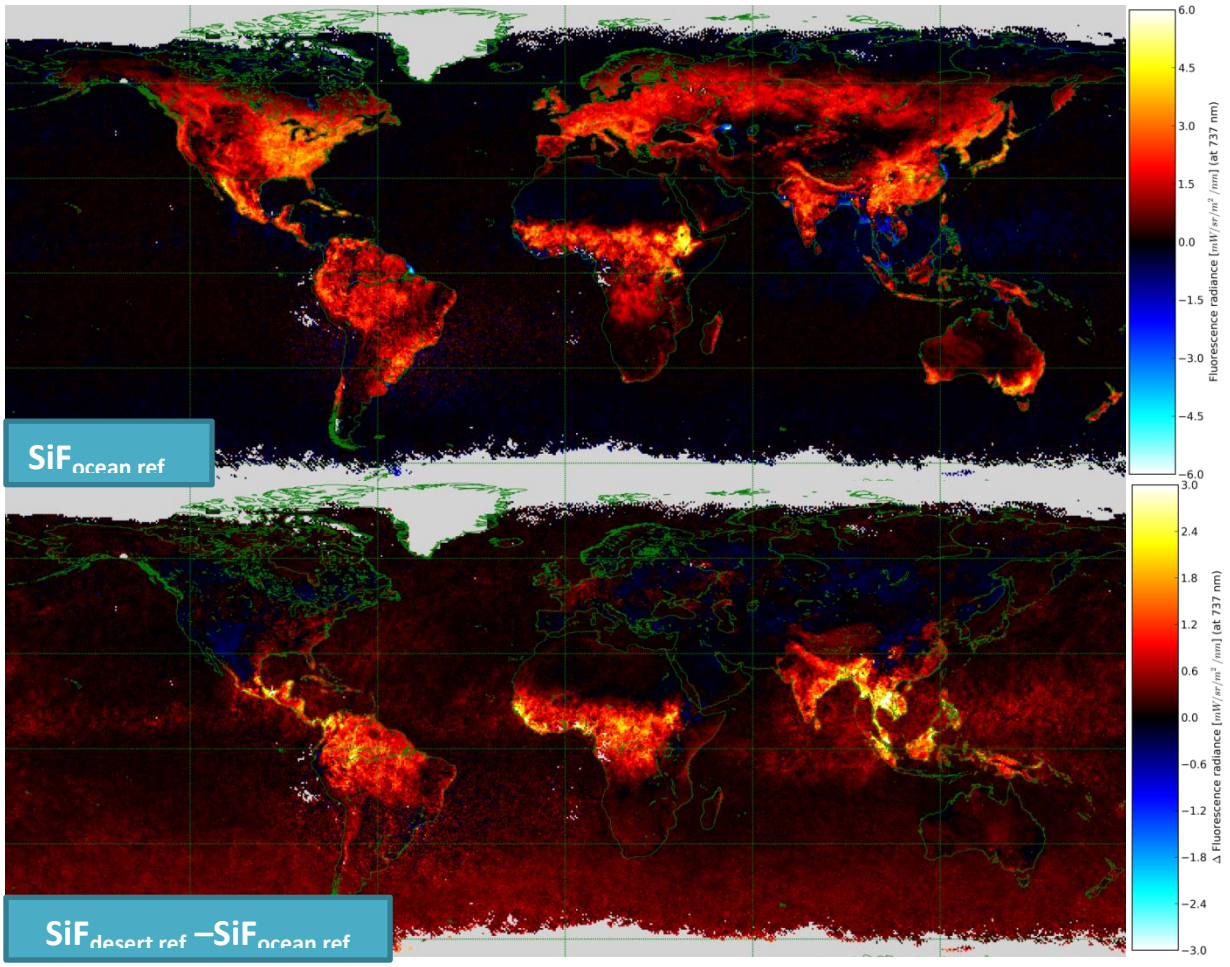


Figure 5.2: Retrieval of SiF using only ocean scenes for the reference dataset. Reference data is collected over a year. Values are monthly means over the 0.5^{o} x 0.5^{0} grid-cells. The month retrieved is September 2013. Grey pixels indicate missing data. The indicated fluorescence values are the values at the assumed peak at 737 nm. In the lower panel the difference between the baseline and the ocean based retrieval is given. Note the difference in the color scale between the panels.

The retrieved SiF values over the rest of the terrestrial areas are somewhat lower than for the baseline retrieval, but with similar spatial patterns. The average over all terrestrial areas is 0.842 mW/sr/m^2/nm and the standard deviation is 0.995 mW/sr/m^2/nm. The average SiF value retrieved over the marine areas is -0.021 mW/sr/m^2/nm, so there seems to be no offset, which is the case in the baseline. The standard deviation is similar to the baseline: 0.226 mW/sr/m^2/nm. However, it could be argued that the algorithm is here trained specifically for assuming oceans to be zero while in the baseline this is true for deserts. Also the negative patch surrounding Indonesia/Malaysia would tilt the average to lower values.

The retrieval results when using only the Sahara desert as a reference area (the baseline) has already been described in section 4. In Figure 5.1 the baseline retrieval for September 2013 is also shown for comparison.

At this point it should be stressed that without an in-depth and independent validation it is not possible to make conclusive statements about what retrieval results are more accurate or realistic. However, results that can be dismissed as unrealistic on the basis of these experiments are significantly negative SiF values as they can safely be regarded as unphysical, excluding those values

that are around zero. Since both the retrieval using ice/snow and the retrieval using cloudy ocean reference sets yields negative fluorescence radiance for large areas of interest and the retrieval using deserts does not (with the exception of the shallow/murky waters), we have decided to continue experimenting only with the desert-based reference spectra.

5.1.2 Reference Months

The results that are discussed above were all obtained using a set of reference spectra obtained from the entire year 2013 and applied to data from September 2013. Our initial retrievals where done using a reference set obtained only from September 2013 and applied to data from that same month. The spatial patterns obtained from that where very similar to that obtained from an entire year of reference data. But, the absolute values were quite different. The average SiF value over the continents when using a reference set obtained from one month only was 1.630 mW/sr/m^2/nm in September. The standard deviation from the mean was 1.646 mW/sr/m^2/nm. This is significantly higher than the values obtained from the baseline retrieval. The exceptions to this are the Australian outback, the Kalahari, the Tibetan Plateau and the western Andes where the retrieved SiF values are actually lower when using just September for training data. The results of both retrieval experiments and their difference are plotted in Figure 5.3.

Using this same algorithm (set for Deserts) to January 2013 yielded surprising results (see Figure 5.4): A retrieval of January, using a reference set obtained from January gives negative results in the equatorial areas, even though the retrieval for September did not. Specifically the areas that are negative are the tropical rainforests of Amazonia, Congo and Indonesia. Also, a band in the tropical oceans that coincides with the inter-tropical convergence zone has negative values.

When applying the reference set from September to retrieve the data for January, the results are all positive over the equatorial areas, just as when a yearly training data set had been used. However, the arid areas including the Sahara itself show negative values, where they were around zero in the baseline.

In both cases the SiF values northward of the equatorial region are mostly lower than the baseline and the SiF values southward of the tropics are higher than the baseline. This is opposite to what was found in September.

When using the set of reference spectra from January, the standard deviation over the ocean (which could roughly be interpreted as noise) is quite higher (0.351 mW/sr/m^2/nm) than in the baseline retrieval for January (0.224 mW/sr/m^2/nm), while it is more similar to the baseline when using reference spectra from September (0.266 mW/sr/m^2/nm). The average SiF over the oceans is negative when using reference spectra from September (-0.114 mW/sr/m^2/nm) which is similar to the values obtained over arid areas for this retrieval variant. The average SiF over the oceans is positive when using reference spectra from January (0.209 mW/sr/m^2/nm) and smaller than in the baseline retrieval (0.301 mW/sr/m^2/nm). It should be remarked that since there is a clear spatial pattern in this retrieval variant, that this average cannot be straightforwardly interpreted as a retrieval offset.

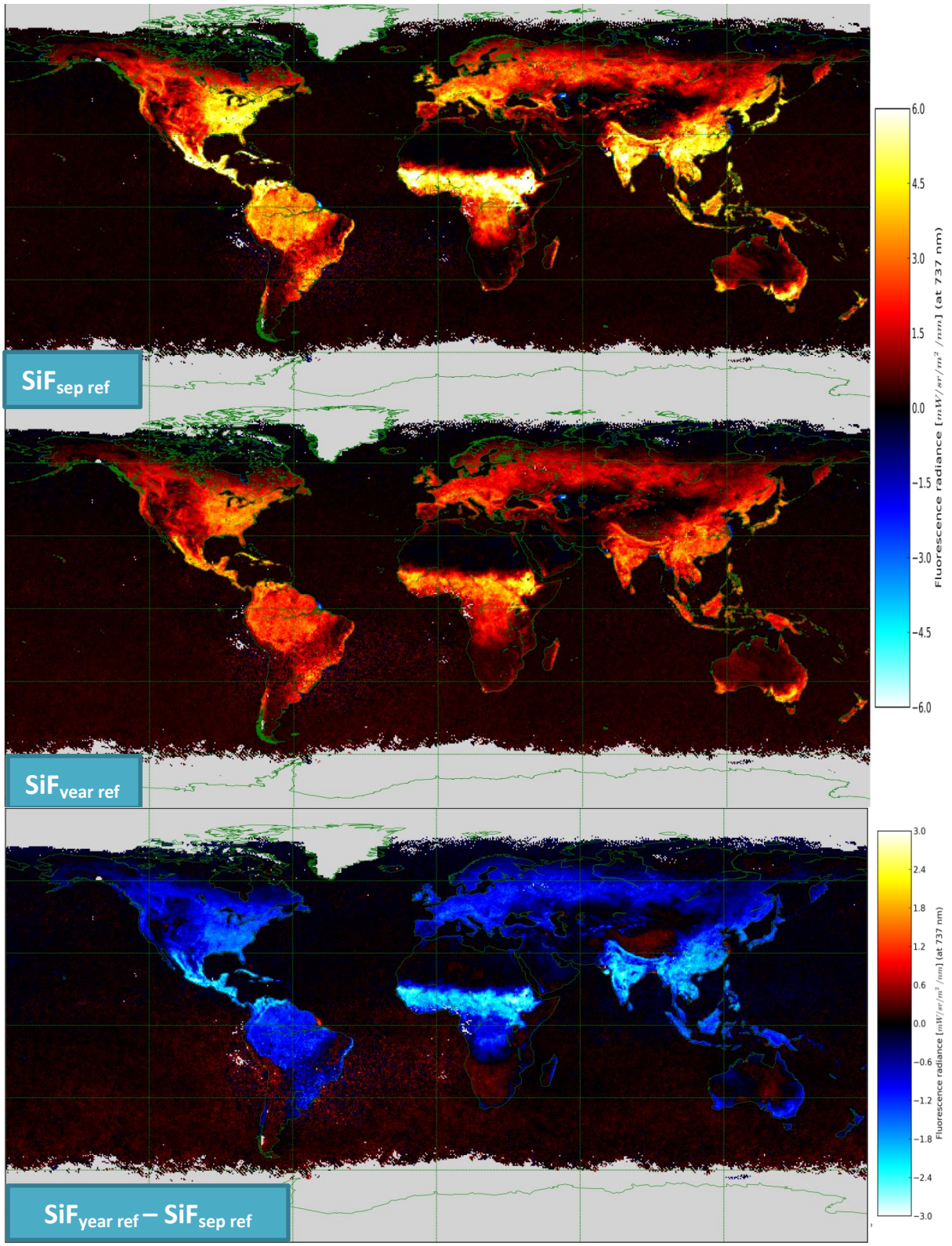


Figure 5.3: SiF retrieval using pixels from the Sahara desert as a training data set. Images show monthly mean values for September 2013. Upper frame: retrieval using pixels from September 2013 only. Middle frame: retrieval using pixels from the entirety of 2013. Lower frame: The radiance difference of sun-induced chlorophyll fluorescence between a retrieval using training data from a year and a retrieval using training data from September (2013).

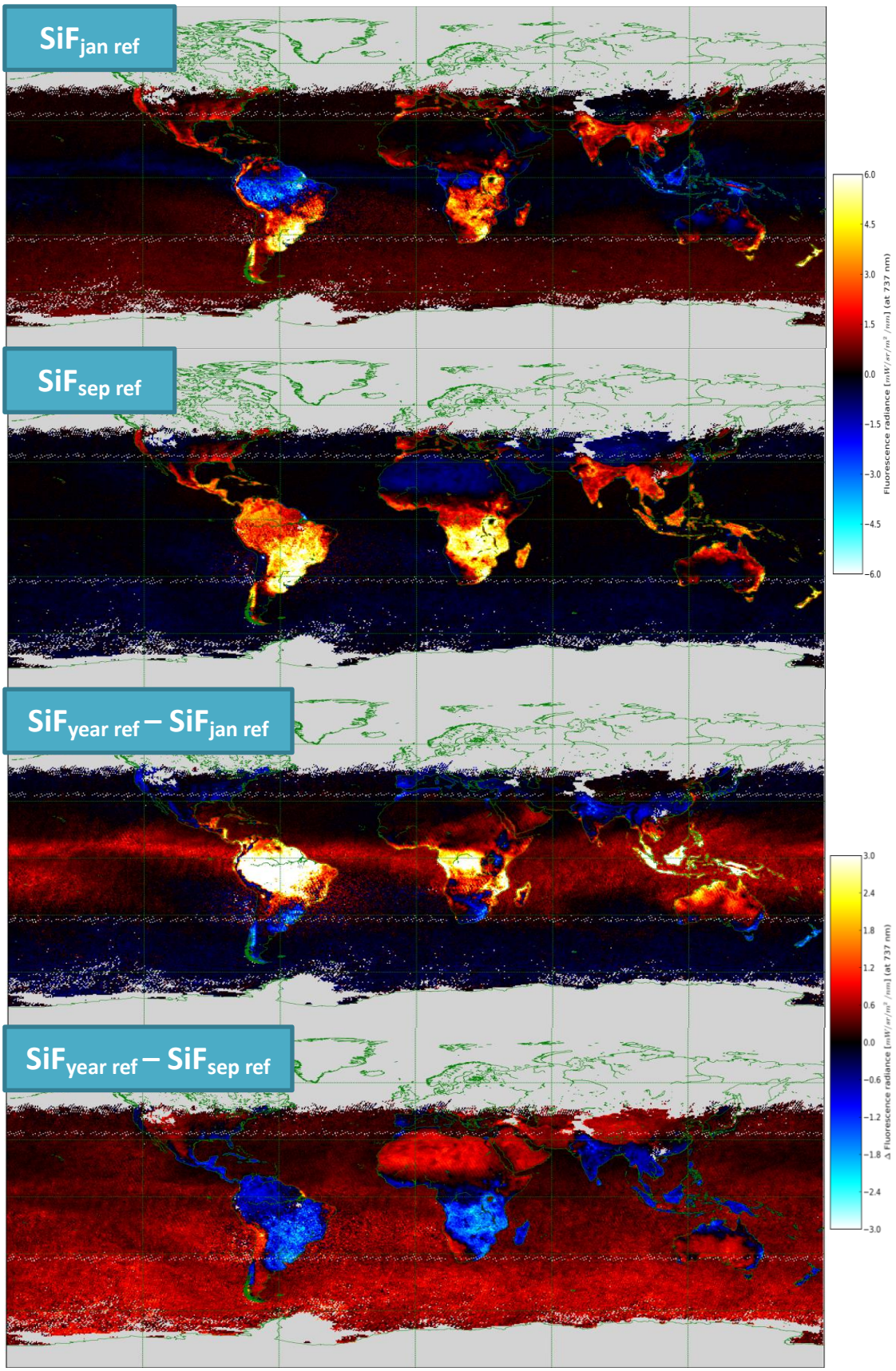


Figure 5.4: Retrieval results of January 2013 for a desert based training dataset taken from different months. Uppermost frame: retrieval from January 2013 training dataset. Second frame: retrieval from September 2013 training set. Third frame: difference between retrieval from all months 2013 training set and January training set. Lowest frame: difference between retrieval from all months training set and September training set.

This surprising dependence of the retrieval on the month from which the set of reference spectra where obtained was a motivation to extend the timespan over which reference spectra are collected in the algorithm to an entire year, in order to ensure a more diverse set of circumstances in the set of reference spectra that can represent conditions throughout the year properly.

The differences between retrieval experiments using reference sets obtained from different months seem to indicate that atmospheric circumstances in the column above the Sahara pixels are quite different depending on the season. Apparently, the conditions in the Sahara in the winter alone cannot represent the conditions in the equatorial areas, while the conditions in the summer are not representative for the deserts in the winter months.

Indeed, the atmospheric circumstances in the Sahara region show a strong seasonal cycle. For instance, the total water vapor columns are much larger in summer than they are in winter in the Southern half of the Sahara region as can be seen in Figure 5.5. The water vapor data is taken from ERA-interim (Dee et al. 2011). Recall that the fit window covers a water vapor absorption band (Figure 2.2).

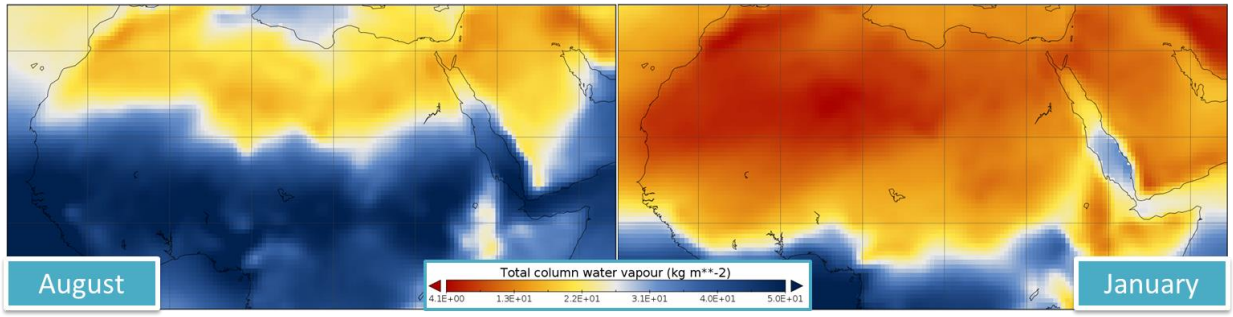


Figure 5.5: Total water vapor column over the Sahara region for the months of August and January in the year 2013. Both are monthly grid averages. The data is taken from the ERA-interim (ECMWF ReAnalysis interim) database (Dee et al. 2011).

5.1.3 Seasonality

For some smaller specific regions we have also run the retrieval algorithm for an entire year using several different reference sets for the principal components. These regions are: North-West Europe, the Arabian Peninsula, Indonesia and the western Mediterranean. These regions represent contrasting scenes: a temperate zone, a desert, a tropical rainforest and a Mediterranean climate respectively. The exact positions and extent of the regions under investigation are shown in Figure 5.6.

The averages over these regions (excluding sea pixels) are given in the timeseries of Figure 5.7. Here it can be seen that the extreme negative (unphysical) values of the ice/snow based retrieval in September are no exception and occur throughout the year in both the temperate region and the tropical region. In the desert region this retrieval variant produces SiF values higher than noise level through May until October while it is expected of the desert to yield zero SiF radiance throughout the year. In the Mediterranean region the ice-based retrieval does show a SiF temporal pattern that is quite similar to the other retrieval experiments.

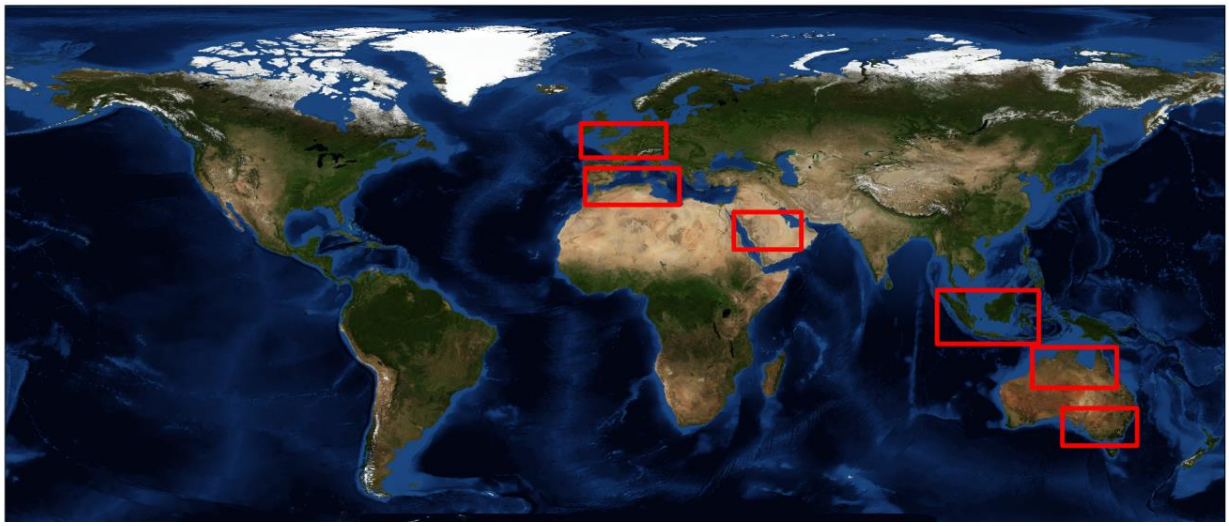


Figure 5.6: Location of test regions for year-long retrieval experiments.

The desert-based retrieval and the ocean-based retrieval produce results that are very similar to each other. In both Western Europe and the Mediterranean they display a seasonal cycle and in the Arabian Desert they remain below their respective noise estimates throughout the year. This is also in line with expectation. However, in the tropical test-region the ocean-based retrieval variant produces only values that are within the noise estimate, which is not likely to be true. The desert-based retrieval, on the other hand, produces a roughly constant positive SiF value within this region that is close to the global average of SiF for this retrieval variant.

We have also contrasted a desert-based retrieval using only data from August and a desert-based retrieval using only data from January. The reason to use August here and not September is because ERA-interim data seems to suggest that August is the peak of the seasonal water vapor cycle in this area. We see that in the temperate test-region they both show similar results, but in both cases retrieved values are higher than in the baseline, except for January until March where they are all very similar. In the Mediterranean region, only the January-based retrieval is higher. In Indonesia, the August-based retrieval is again more similar to the baseline, while the January based retrieval is negative during the first half of the year and positive during the second half of the year. In Arabia it can be seen that the August based retrieval gives negative results from January until August and gets slightly positive after that. The January based retrieval seems to follow a similar curve, but offset to higher values.

Thus, these temporal patterns seem to support our previous observations based on the spatial patterns found in **section 5.1.1** and **section 5.1.2**: The ice-based retrieval is unreliable in all instances; The retrieval using the all-year desert-based reference set is the only of the retrieval experiments that yields plausible results in all instances, while the retrieval using the only-August desert-based reference set is mostly plausible, except for vegetation poor areas.

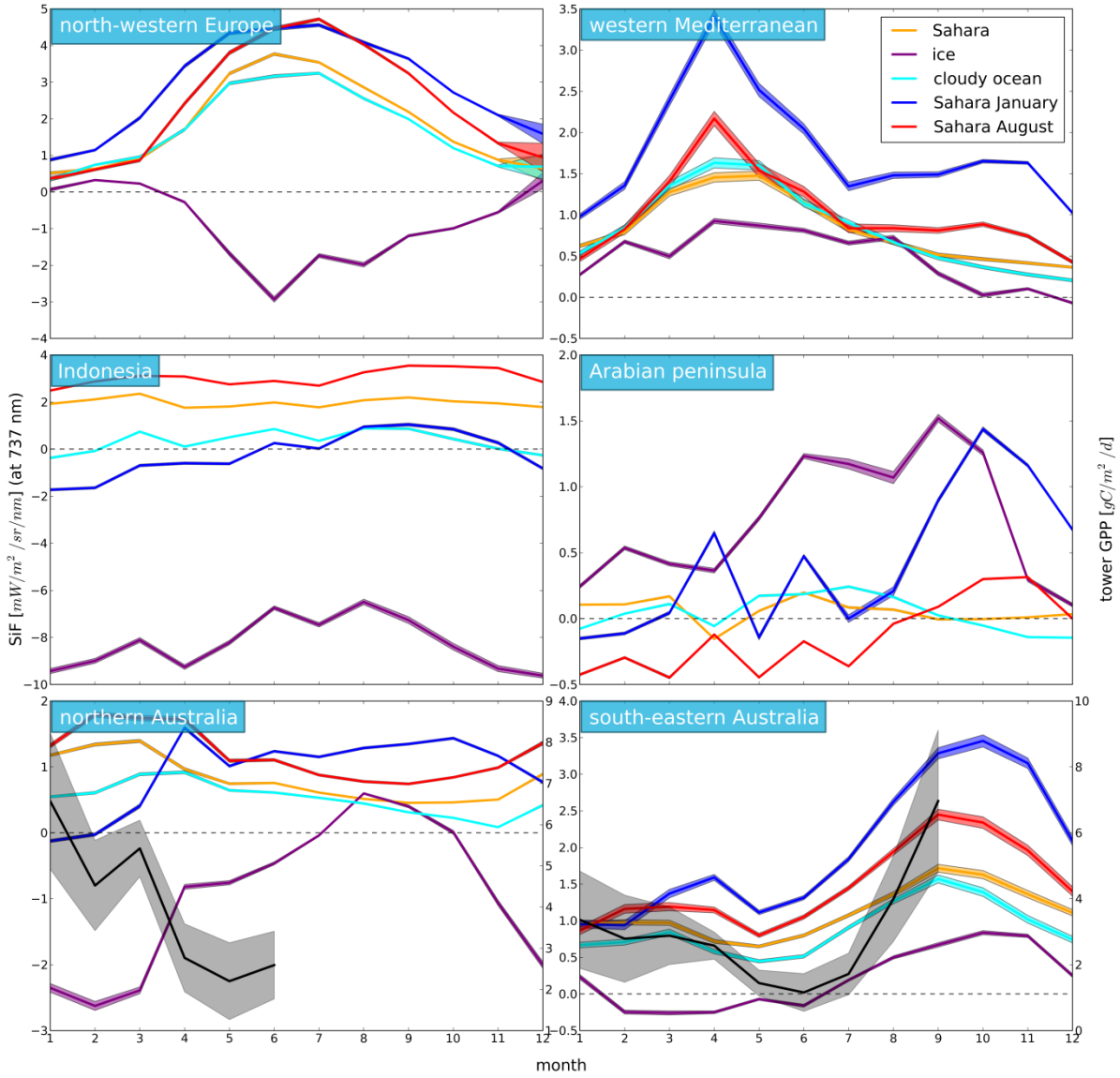


Figure 5.7: Timeseries for spatial average SiF values over six different areas. The bottom two frames also include the average GPP obtained from a number of flux towers in those areas as described in section 7.3. The shaded areas are the standard deviations over the sample pixels/tower sites.

We have also run the retrieval algorithm for an entire year in two distinct areas of the Australian mainland: The Tropical Savanna's in the north; and the south-eastern part, which has a Mediterranean climate. For these two areas we also plotted monthly GPP values obtained from several flux tower sites. The SiF in south-eastern Autralia shows a similar pattern as is observed in the Mediterranean region, but with the seasons shifted. The north shows a pattern that is more comparable with Indonesia. In the Australian south-east, the observed SiF seasonality in all the retrieval experiments was very similar to that observed in the flux tower GPP measurements. In the north of Australia this was more difficult to assess, because flux tower data was only available for the first half of 2013. However, it can be seen that the temporal pattern seen in the retrieval based on ice and the retrieval based on the Sahara in January does not match that seen in the tower GPP. A more in-depth evaluation of remotely sensed SiF in Australia using these flux tower measurements of GPP is given in **section 7.3**.

Table 5.1: statistics for the retrieval experiments using reference datasets obtained from different areas.

Reference dataset	μ_o	σ_o	μ_c	σ_c	RN
Ocean (entire year)	-0.021	0.226	0.842	0.995	27 %
Ice (entire year)	0.051	0.443	-0.838	2.527	-53 %
Desert (entire year)	0.180	0.222	1.070	1.077	21 %
Desert (September)	0.154	0.252	1.630	1.646	15 %

Table 5.2: statistics for retrieval experiments using reference datasets obtained from different periods. All experiments use the Sahara desert as their reference area.

Reference dataset	μ_o	σ_o	μ_c	σ_c	RN
Year	0.301	0.224	1.121	1.230	20 %
January	0.209	0.351	0.505	1.458	70 %
September	-0.114	0.266	1.167	1.884	19 %

5.2 Number of Principal Components

The principal components that are derived from the training data set are a crucial part of the retrieval. We have seen in the above section that the region and month chosen for the selection of reference spectra to be used in the principal component analysis can change the retrieval results appreciably and that we need to be careful to select the reference dataset that is most representative of the target scenes. But, how many principal components do we actually need for the retrieval algorithm to capture all of the variability related to the state of the surface-atmosphere system that is contained in the complete set of reference spectra?

First, we will look at the principal components themselves. Figure 5.8 shows the six principal components that have the highest fraction of explained variance for the baseline retrieval. The first principal component may be interpreted as a background state and the spectral dependence is clearly reminiscent of the O_2-A absorption band and the H_2O absorption band found in this spectral window. Features of both the O_2-A band and the H_2O band can be recognized in further principal components but they become increasingly harder to interpret. This is characteristic of a PCA based retrieval and it should always be kept in mind that PCs generally are not related to physical properties in the way that the parameters in a physical model would.

It can be seen that the first principal component already explains 99% of the variance in the reference spectra. This seems to suggest that a relatively low number of principal components should be sufficient to capture the atmospheric circumstances in the target pixels to a reasonable degree.

We have performed several different retrieval experiments with a varying number of principal components. The rest of the retrieval settings is kept the same as the baseline retrieval. The retrieved values are again for September 2013 and averaged over the grid-cells.

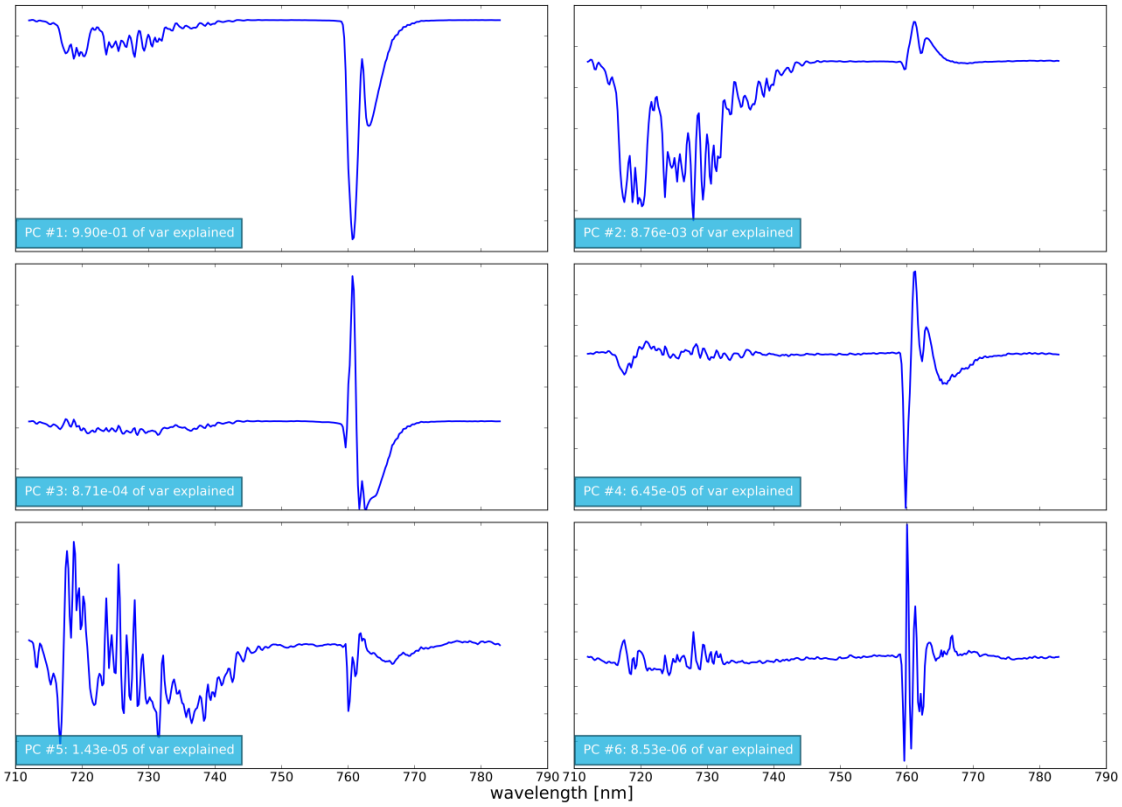


Figure 5.8: The first six principal components in the baseline retrieval.

It can immediately be seen from the results in Figure 5.9 that a number of principal components of 20 or lower does not suffice. What can be seen is that in South-East Asia and the Gangetic plain negative SiF values crop up. Also The Amazon area and the Congo area have unrealistically low or unphysically negative SiF values. When lowering the number of principal components even further, these negative values become more negative and all wet tropical areas become affected. The desert areas on the northern hemisphere, in particular the Sahara and the Arabian Desert also show negative SiF values.

When increasing the number of principal components to 25 or more there are no longer negative values over areas of interest or unrealistically low values and the retrieval result seems at first glance the same as the baseline, which uses 35 principal components. Thus, 25 principal components would be the minimum number of principal components that is needed to get plausible results (we have only varied the principal components in increments of 5).

Even though using 25 principal components yields plausible results, it is not true that using a higher number of principal components makes no difference for the result. As can be seen in Figure 5.10, there are still differences between the results using higher number of principal components. Increasing the number of principal components decreases the SiF values found in most areas of the globe, while the SiF values increase over South-East Asia, West Africa and parts of the Andes. The differences in results between retrieval experiments with a differing number of principal components become less as the number of principal components increases.

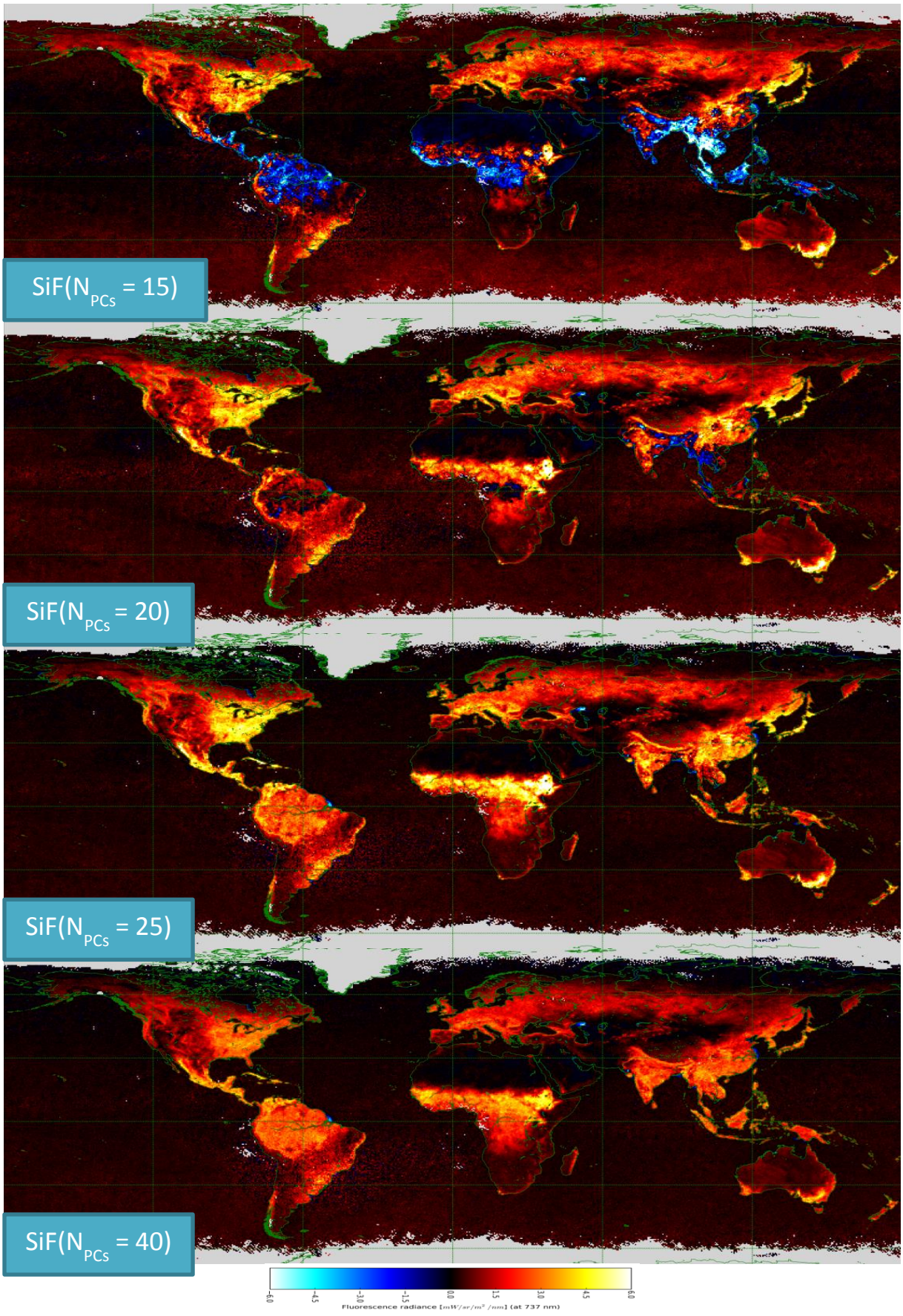


Figure 5.9: Retrieval results when using different numbers of principal components in the model. Uppermost frame: 15 PCs. Second frame: 20 PCs. Third frame: 25 Pcs. Lowest frame: 40 PCs. All four retrieval maps are for September 2013.

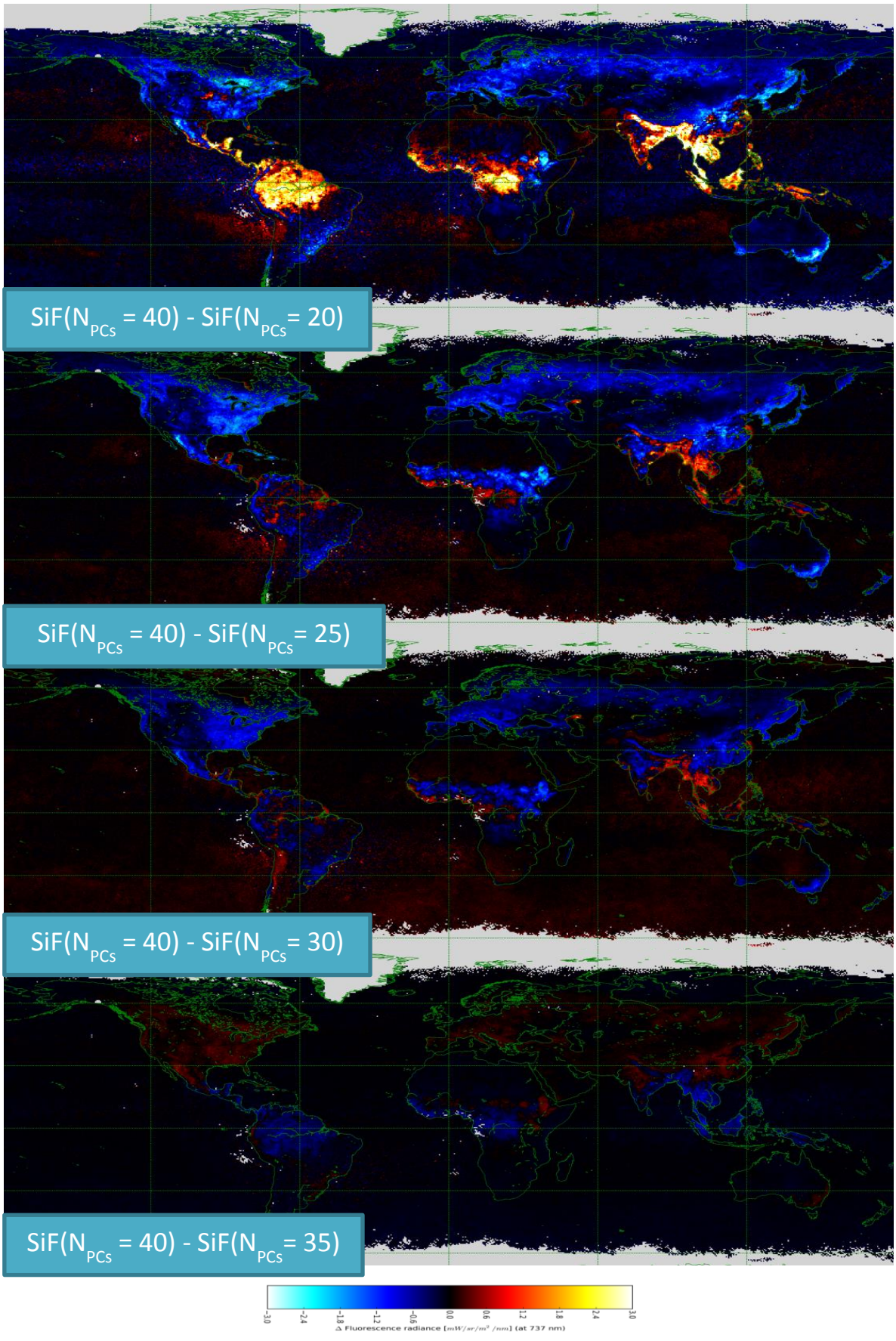


Figure 5.10: Maps of the difference in retrieved SiF values between a retrieval experiment using 40 principal components and retrieval experiments using several other numbers of principal components.

Increasing the number of principal components also increases the relative noise in the result, as can be seen in Table 5.3. This is related to the difference between the number of dimensions in the cost-function and the number of fit parameters. More fit parameters relative to the number of fit dimensions (in this case the dimensions are the number of discrete wavelength bins) increases the amount of stochastic noise that can be fit in the model. Because of this, one should try to limit the number of principal components used in the forward model. This should be traded off against the SiF biases found when using a too small number of principal components as seen in Figure 5.10.

Table 5.3: : Statistics for SiF retrievals using a decreasing number of principal components. For the last two entries no SNR value is given because these feature a non-negligible amount of negative valued pixels, for which the SNR definition does not hold.

N_{PCs}	μ_o	σ_o	μ_c	σ_c	RN
40	0.219	0.232	1.094	1.074	21 %
35	0.180	0.222	1.070	1.077	21 %
30	0.176	0.233	1.232	1.259	19 %
25	0.213	0.241	1.383	1.383	18 %
20	0.297	0.262	1.246	1.256	21 %
15	0.283	0.382	0.791	1.637	48 %

In the end we chose 35 principal components as a reasonable compromise for further retrieval experiments.

5.3 Surface reflectance polynomial and window

In the forward model (i.e. the forward model for the actual SiF retrieval), the surface albedo is characterized as a low order polynomial. Increasing the order of the polynomial in the retrieval model increases the number of fit parameters in the cost function. The exact order of the polynomial will therefore represent a balance between fitting noise (too high order) and making a model error for the surface albedo (too low order).

We have performed a number of retrieval experiments where we varied only the order of the surface reflectance polynomial. The rest of the retrieval settings are kept at the baseline values. In the baseline the order of the surface reflectance polynomial is four. Using a different polynomial order has substantial impact on the results. When the order is increased from four to five the spatial patterns of the SiF retrieval do not change significantly, as seen in Figure 5.11. On the other hand, the magnitude of the retrieval values is quite lower. At the same time, the relative noise level is higher. When increasing the polynomial to sixth order, a similar increase in noise level is found, but the difference is less. This is summarized in Table 5.4.

Table 5.4: Statistics for SiF retrievals using a wavelength window of 712 nm to 780 nm.

Polynomial order	μ_o	σ_o	μ_c	σ_c	RN
3rd	0.182	0.224	1.114	1.147	20 %
4th	0.180	0.222	1.070	1.077	21 %
5th	0.149	0.197	0.614	0.667	32 %
6th	0.152	0.192	0.569	0.554	33 %

As can be seen, increasing the order from four to five greatly increases the relative noise, while decreasing the order to three decreases it by a much smaller amount. From this limited number of retrieval experiments it seems that the order of the surface albedo polynomial has a much larger effect on the relative noise level than the number of principal components does. Because, of the small difference in relative noise between the orders three and four, we have chosen a fourth order as a reasonable compromise for the baseline.

All previous retrieval experiments have been performed using a wavelength window of 712 nm until 780 nm. With the spectral sampling of the GOME-2 instruments this means 340 discrete wavelength bins. We have also performed a few experiments with a smaller window of 712 nm to 758 nm, which means 230 discrete wavelength bins. This smaller wavelength window covers the H_2O absorption band, but not the O_2-A band covered by the broader window.

Narrowing the wavelength window does not really affect the spatial patterns observed, but it lowers the SiF signal strength considerably. The noise level is higher, even in the absolute sense; See also Table 5.5 and Figure 5.12. Reducing the number of degrees of freedom in accordance with the reduced number of equations in the forward model by lowering the order of the surface reflectance polynomial leads to a higher signal to noise ratio.

Table 5.5: Statistics for SiF retrievals using a wavelength window of 712 nm to 758 nm.

Polynomial order	μ_o	σ_o	μ_c	σ_c	RN
3rd	0.180	0.238	0.700	0.691	34 %
4th	0.194	0.234	0.421	0.464	56 %

In order to ensure a reasonably low relative noise we chose to keep the broader spectral window and to keep the surface reflectance polynomial at fourth order. Although a third order surface reflectance polynomial would yield an even smaller relative noise value, the difference between the noise at fourth order and third order is small and using a too small polynomial order would carry the risk that some surface could not be adequately modelled. Therefore fourth order seems to be a reasonable compromise.

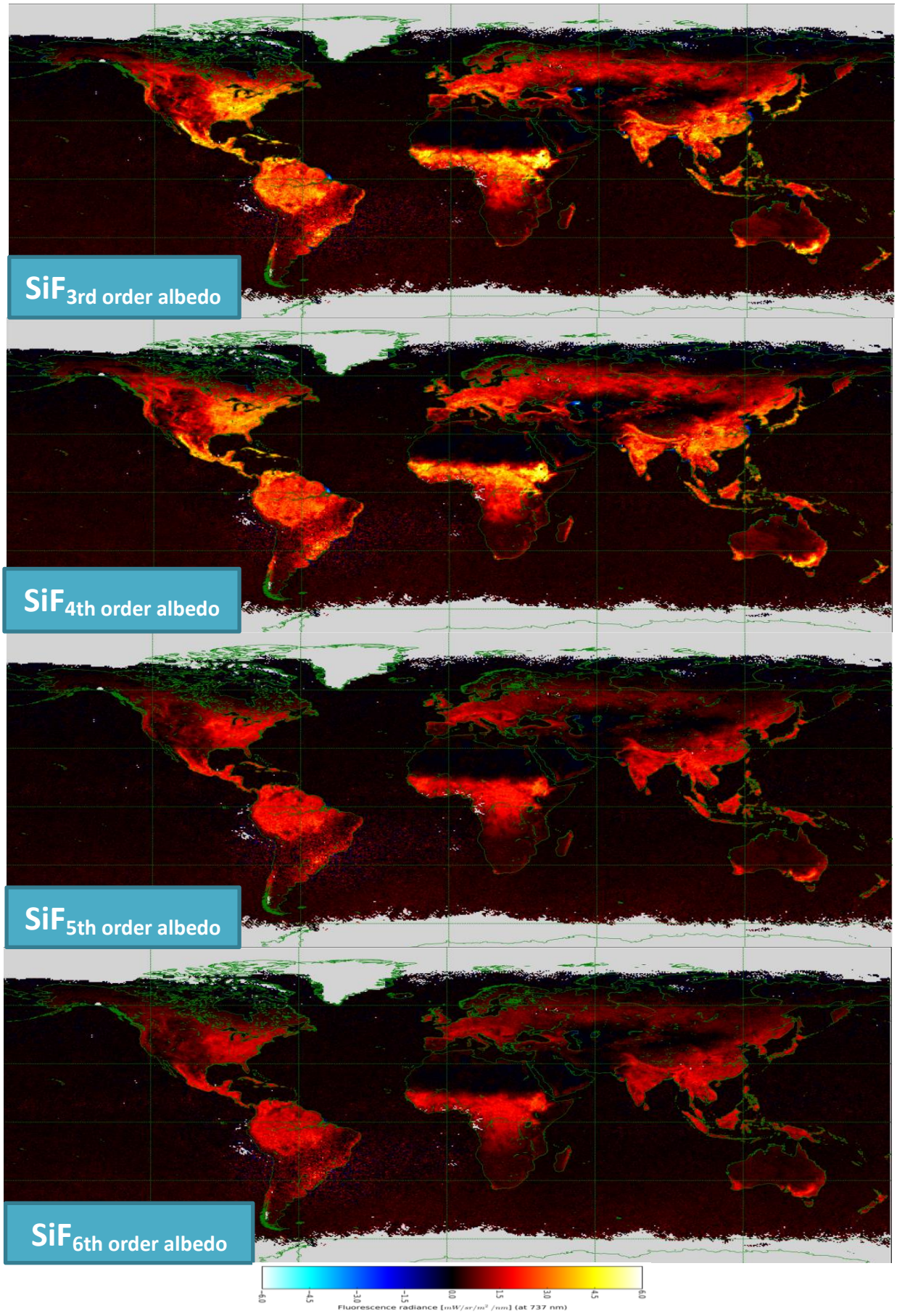


Figure 5.11: Retrieval results using different orders for the surface albedo polynomial.

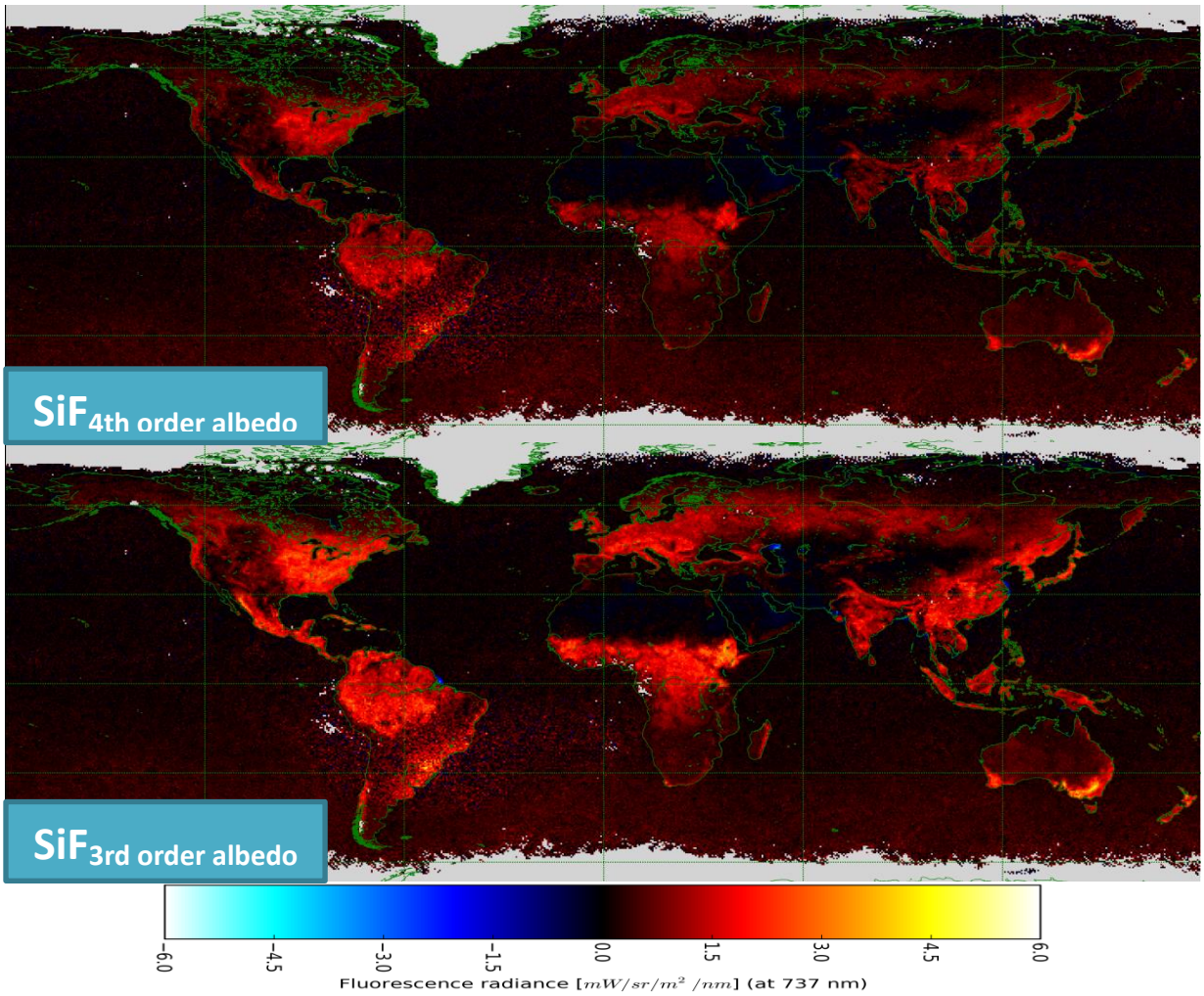


Figure 5.12: Retrieval results using a spectral window of 712-758 nm, using different orders for the surface albedo polynomial.

6 Fluorescence Yield

As a next step towards improving estimations of GPP, using fluorescence yield as a proxy for LUE might complement existing proxies better than using fluorescence radiance as a direct proxy for GPP. We have calculated the fluorescence yield by combining the retrieved fluorescence with data from other sources as follows:

$$\phi_F = \frac{I_F}{fAPAR \cdot PAR} \qquad (6.1)$$

Here, I_F Is the wavelength integrated fluorescence radiance, $fAPAR$ is the fraction of absorbed photosynthetically active radiation and PAR is the incident photosynthetically active radiation at the surface. For all these quantities daily values are used and the resulting daily fluorescence yields are averaged over the period of a month. We will first describe how the fAPAR and PAR datasets were obtained (**section 6.1**) and then we will discuss the results (**section 6.2**).

6.1 Method

The fAPAR data that is needed in equation (6.1) was obtained from MODIS aboard the TERRA satellite[5]. These values are derived from atmospherically corrected surface reflectances (Knyazikhin et al. 1999). Since MODIS has a much higher temporal resolution than GOME-2 (1 km x 1 km), we have reprojected and interpolated the fAPAR data towards the 0.5^{o} x 0.5^{o} geographic projection grid of the fluorescence data.

We have used only the data from the TERRA satellite and not the combined product of the TERRA and AQUA satellites. This is because TERRA has an overpass time of roughly 10:30 in the morning, while AQUA has an overpass time of roughly 13:30 in the afternoon. Since MetOp-A has an overpass time of 9:30 in the morning, the TERRA data would be more representative of the conditions during a GOME-2 fluorescence retrieval than the AQUA data. It should be noted that this data is only available as 8 day composites and therefore the same fAPAR values are used for 8 consecutive days. This could potentially be improved by using interpolated values instead.

The PAR at the surface dataset is taken from the ERA-interim reanalysis. It is available in a geographic projection at the same grid resolution as the retrieval data, 0.5^{o} x 0.5^{o}. We used the daily integrated clear sky PAR. However, the fluorescence that results from the retrieval is an instantaneous value at the time of overpass and not an integrated value. Therefore, the integrated PAR needs to be converted to an instantaneous PAR, iPAR. To do this the solar elevation at the time of overpass is needed, which can be obtained from the solar zenith angle (which is included in the level 1b product files) using equation (6.3).

$$\alpha_s = \cos\theta_s \qquad (6.2)$$

[5] MOD15 data obtained from the online Data Pool at NASA Land Processes Distributed Active Archive Center (LP DAAC)

We will also need to know the integral of all the solar elevations during each day for which PAR was integrated. To calculate this, first the solar declination must be known for each of these days. This is achieved via the following equation:

$$\delta_s = \varepsilon \cos\left(2\pi \frac{DOY + 10}{365}\right) \quad (6.3)$$

Here, δ_s is the solar declination angle, ε is the obliquity of the earth, DOY is the day of the year numbered from 1 for January 1st to 365. The northern winter solstice is assumed to be on December 21st. When the solar declination is known, the solar hour angle can be calculated for each grid point using that grid point's latitude.

$$\cos \psi_s = -\tan \varphi \tan \delta_s \quad (6.4)$$

Here, φ is the latitude and ψ_s is the solar hour angle at the time of sunset. From here, the daily integral of the solar elevation angle can be calculated.

$$\alpha_{s,int} = 2\psi_s \sin \varphi \sin \delta_s + \cos \varphi \cos \delta_s (\sin \psi_s - \sin(-\psi_s)) \quad (6.5)$$

Here, $\alpha_{s,int}$ is the daily integrated solar elevation. Now, the instantaneous PAR at the time of overpass can be calculated.

$$iPAR = PAR \frac{\alpha_s}{\alpha_{s,int}} \quad (6.6)$$

This value is then used to calculate the fluorescence yield.

In order to calculate the total fluorescence emission from the 737 nm emission, the integral over the complete spectral shape of the fluorescence must be taken. However, this must include the second peak of fluorescence at 690 nm. Since we do not know this, we cannot get an absolute value for the yield. The values obtained can therefore only be compared relatively.

The pixels which have a SiF value below zero are filtered out of the fluorescence yield results since the yield of a negative emission is a meaningless quantity. We have also filtered out pixels which have a fAPAR value below 0.1. This is done to prevent fluorescence yield values to 'blow up' in some barren areas due to random noise. The remaining values are averaged per gridcell per month.

6.2 Results

The fluorescence yields are all obtained from our baseline fluorescence radiance results. The values obtained here are of arbitrary magnitude and should not be construed to be of comparable magnitude to actual yield fractions. They can for now only be compared relative to each other. No data over the oceans is available, because the MODIS fAPAR data that was used in deriving the fluorescence yield did not include data over the oceans. The spatial patterns in fluorescence yield are illustrated using the results obtained from July 2013 as presented in Figure 6.1.

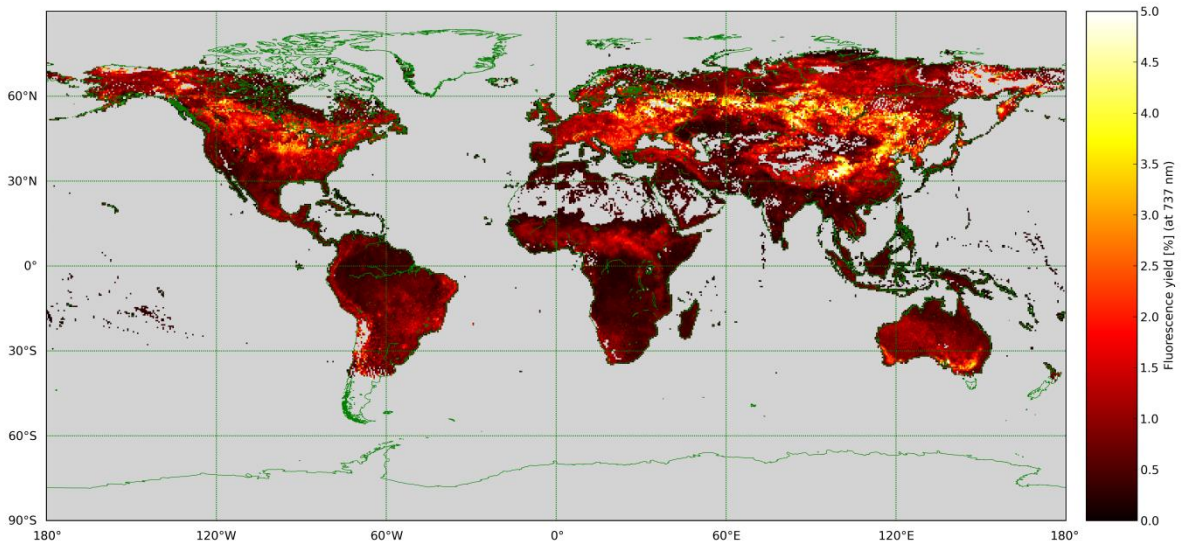


Figure 6.1: Fluorescence yield averaged over July 2013 in a 0.5° x 0.5° grid.

Consistent spatial and seasonal patterns are seen in the fluorescence yield results. These patterns are mostly very similar to those found in the fluorescence radiance, with some striking exceptions. The tropical rainforest areas of the Amazon, central Africa and Indonesia have very low values throughout the year. This is in contrast with the fluorescence radiance, which shows values close to global land average throughout the year. The arid western Andes region which has very low fluorescence radiance does show a reasonably strong fluorescence yield. Similarly the Australian outback lights up in June and July. Other similarly arid areas do not show much fluorescence yield. The Corn Belt in the Midwestern United States shows much higher fluorescence yield values than its surroundings during the months of June, July and August. This contrast is not seen in the fluorescence radiance.

Fluorescence yield maps for each month in 2013 are given in Figure 6.2 to illustrate the seasonality patterns. The seasonality observed in the Sahel is much more pronounced in the fluorescence yield values than in the fluorescence radiance values. This is also true for the seasonality in the Boreal regions. In the South-west and South-East of Australia it can also be seen that the seasonal cycle in fluorescence yield is shifted compared to the seasonal cycle in fluorescence radiance. The peak for the fluorescence radiance lies in September and October while the peak for fluorescence yield lies in the months of July and August.

We have also categorized the fluorescence yield results by land use type following the same procedure as we did with the fluorescence radiance, as illustrated in Figure 6.3. Here it can be seen that also in terms of yield the tropical rainforest areas remain constant values throughout the year, while the other land use types show a seasonal pattern. The peak of fluorescence yield is in July, just as with fluorescence radiance. A significant difference between the pattern of fluorescence yield and the pattern of fluorescence radiance is that the yield values for tropical rainforests are around the minimal values for deciduous forests, while the radiance values for tropical rainforests are closer to the yearly average value of deciduous forests.

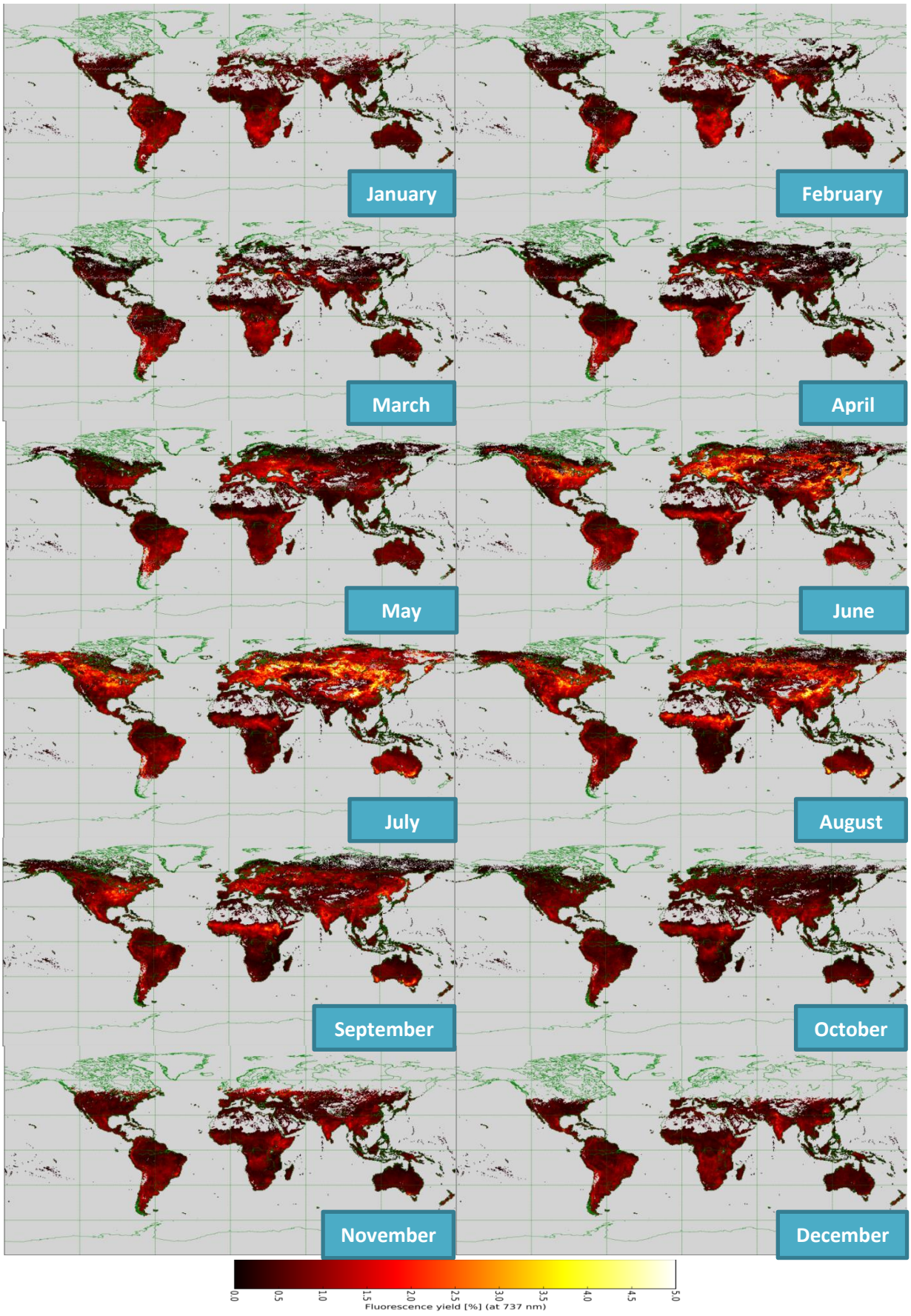


Figure 6.2: Monthly mean fluorescence yield retrieval for 2013.

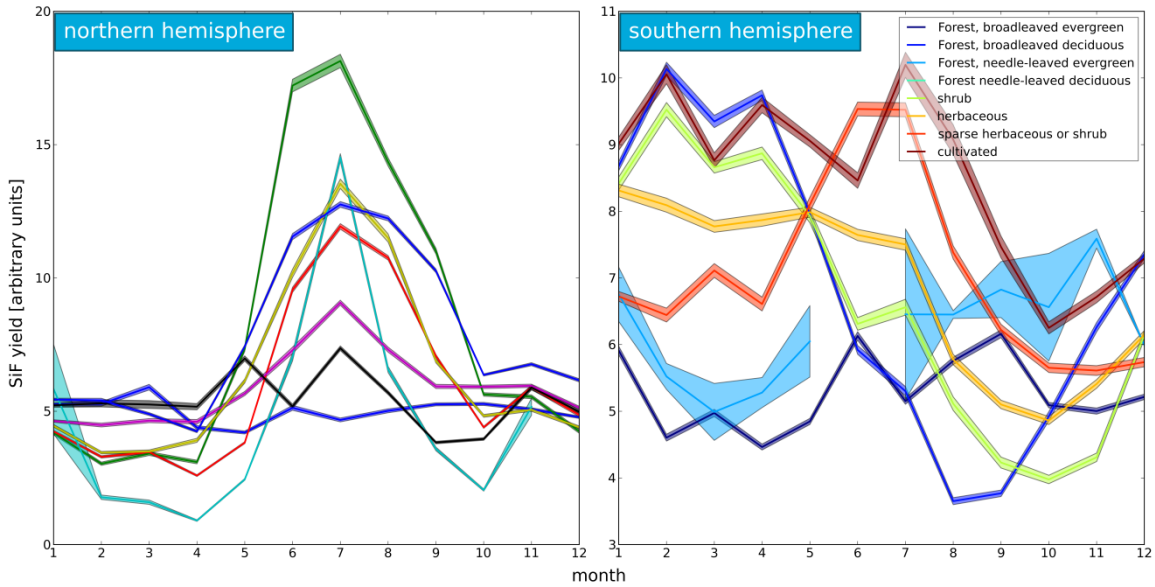


Figure 6.3: Timeseries of SiF yield for 2013. Each line represents the average SiF yield values for a particular land use type. Left side shows averages over the northern hemisphere; right side shows averages over the southern hemisphere. The shaded areas indicate the extent of the standard deviations over the sample pixels.

7 Evaluation

After having obtained a suitable baseline that can be considered plausible on a priori grounds, we need to subject our results to an independent validation. Unfortunately, at this time it was not yet possible to compare to suitable, reliable and independent ground measurements. The size of the GOME2 footprint plays a large role in this. Instead, we made several different comparative evaluations. First we compare with a completely different widely used vegetation index, the NDVI (**section 7.1**). Second, we evaluate the differences between our retrieval results and those of Joiner et al. (2013) (**section 7.2**). Thirdly we compare with a limited set of ground based flux tower measurements from Australia (**section 7.3**).

7.1 Comparison with NDVI

As a first step to evaluating the results obtained from the SiF retrieval algorithm we compared the obtained results with another satellite based vegetation proxy, the normalized difference vegetation index (NDVI). The NDVI is calculated as follows:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \qquad (7.1)$$

Where ρ_{NIR} is the TOA reflectance in the near-infrared wavelength band and ρ_{red} is the TOA reflectance in the red wavelength band. A high NDVI indicates a high amount of surface area that absorbs red light but not infrared light. This is characteristic of green leaves.

A high NDVI indicates more green-leaved vegetation. Therefore, signals coming mainly from those areas indicate signals likely coming from vegetation. We have compared the NDVI as obtained from the MODIS instrument with the SiF retrieval results. This is illustrated in Figure 7.1.

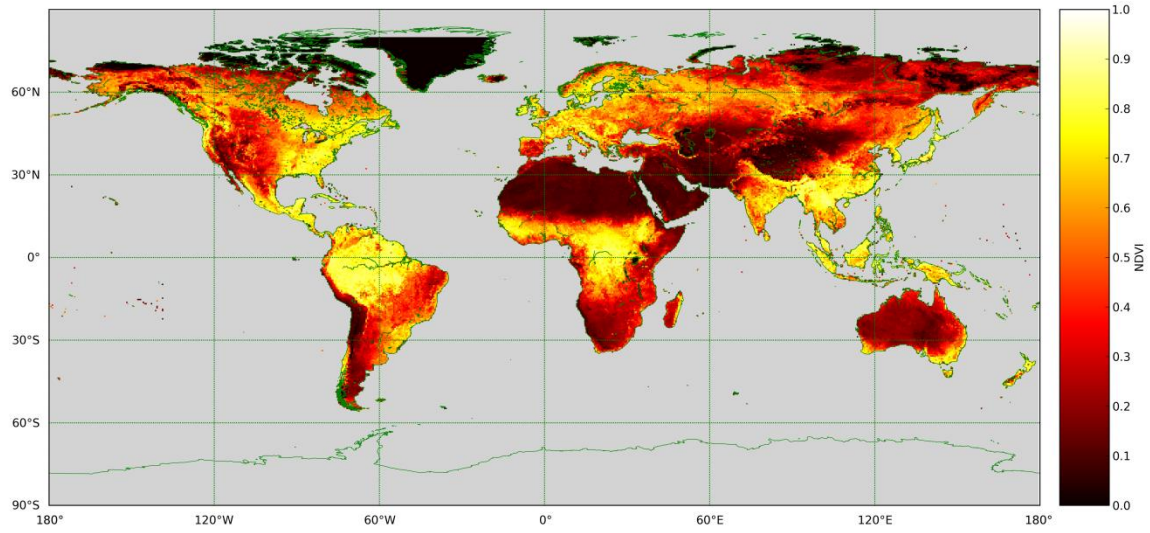


Figure 7.1: Monthly mean NDVI values for July 2013 from MODIS. Gridcells are 0.5° x 0.5°.

It can be seen here that in general the spatial patterns in the NDVI results are very similar to the spatial patterns in the SiF results. We can see that in both cases the arid regions have very low values, while temperate, continental and tropical savanna zones have higher values.

As an example we can look at the continental United States and Mexico (Figure 7.2) to see a striking spatial pattern that occurs both in the NDVI and the SiF results. We can see how the Eastern US has a high vegetation index, while the western US and northern Mexico have a low vegetation index. The northwestern coastal area of the US and the southern part of Mexico have a again a higher vegetation index. The same pattern is visible in in the SiF retrieval values.

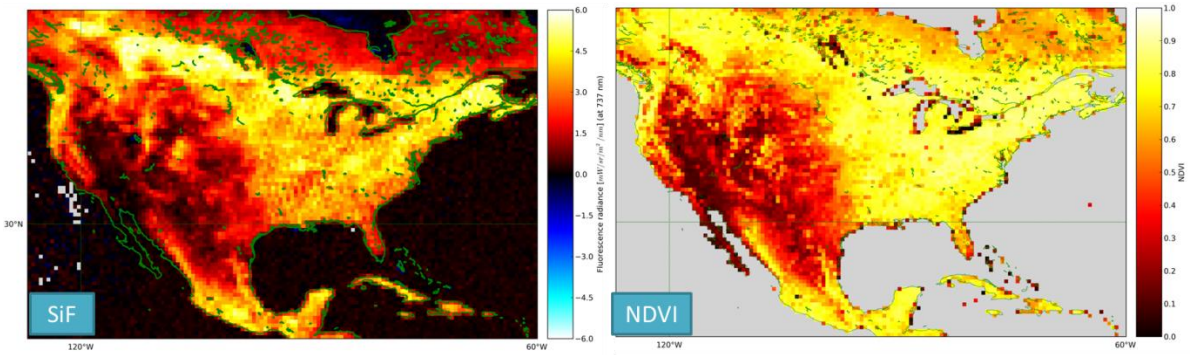


Figure 7.2: GOME-2 SiF emissions (left) and MODIS NDVI (right) over North America for July 2013.

These and other patterns found across the globe provide a clue that the results of the retrieval algorithm are indeed indicative of a signal originating with vegetation. However, there are also clear differences between the spatial patterns observed in the NDVI and the spatial pattern observed in the SIF. Two environments in particular stand out: Tundra areas and tropical rainforests.

Rainforest areas show slightly lower SiF radiance values compared to other areas with similar NDVI values. The boreal tundra and taiga areas in September still show high NDVI values while, there is virtually no SiF observed in the tundra in those months and the SiF is already significantly lower in the taiga. This indicates that we are still measuring a different signal than NDVI measurements. It also provides a clue as to where SiF measurements might provide information that is not covered by other vegetation indices.

We might also perform a linear regression of NDVI versus SIF radiance. In Figure 7.3 this is shown for all monthly grid averages of September 2013. There is a clear correlation between NDVI and SiF with an r^2 of 0.61. However, there is also quite some deviation around the regression line which indicates that the NDVI alone cannot predict the SiF value. It can also be observed that the deviation increases with higher NDVI values with SiF values around zero still possible at quite high NDVI values while higher SiF values are not occurring at low NDVI values. This confirms the suspicion that NDVI is an indicator of the potential for photosynthesis while SiF is a more direct indicator of photosynthesis occurring.

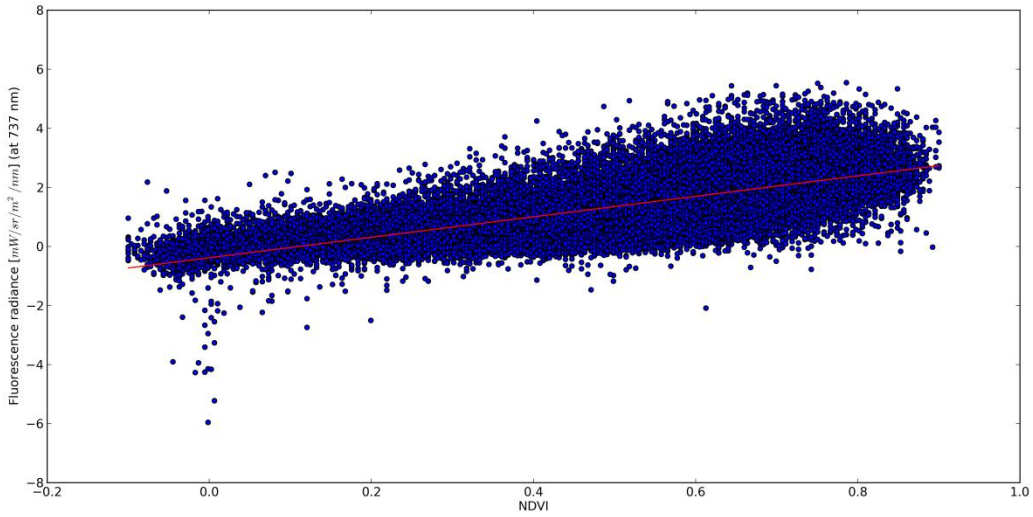


Figure 7.3: SiF versus NDVI for gridcell averages over September 2013.

7.2 Comparison with previous work

Since our work is based on Joiner et al. (2013), it may also be useful to compare our results with the results put forward in that article. For that, we have also applied our retrieval algorithm to September 2012 and compared it to the results of Joiner et al. for that same month. It should be noted that in this case we still used the set of reference spectra obtained from 2013 for logistical reasons.

The results obtained in both attempts are quite similar, which is as expected. Note that the results by Joiner et al. (2013) do not include oceans. One difference that is immediately obvious is the generally lower SIF values obtained by Joiner et al (2013). These lower values are also consistent with earlier SIF retrieval attempts using GOSAT (Frankenberg et al., 2011; Guanter et al., 2012). However, the SiF retrieval by Joiner et al. uses a smaller wavelength window than our retrieval and uses the same order of the surface reflectance polynomial. As we found in our retrieval experiments, narrowing the wavelength window does indeed lower the magnitude of the retrieved SIF values, but it also increases the relative noise in the retrieval.

Another difference that is immediately observed is the grainy pattern over southern South America retrieved by Joiner et al. which obscures any other spatial pattern that might be seen over this area. Its position is indicative of a noise pattern related to the so-called South Atlantic anomaly. This grainy pattern is severely reduced in our retrieval, although it is not entirely absent.

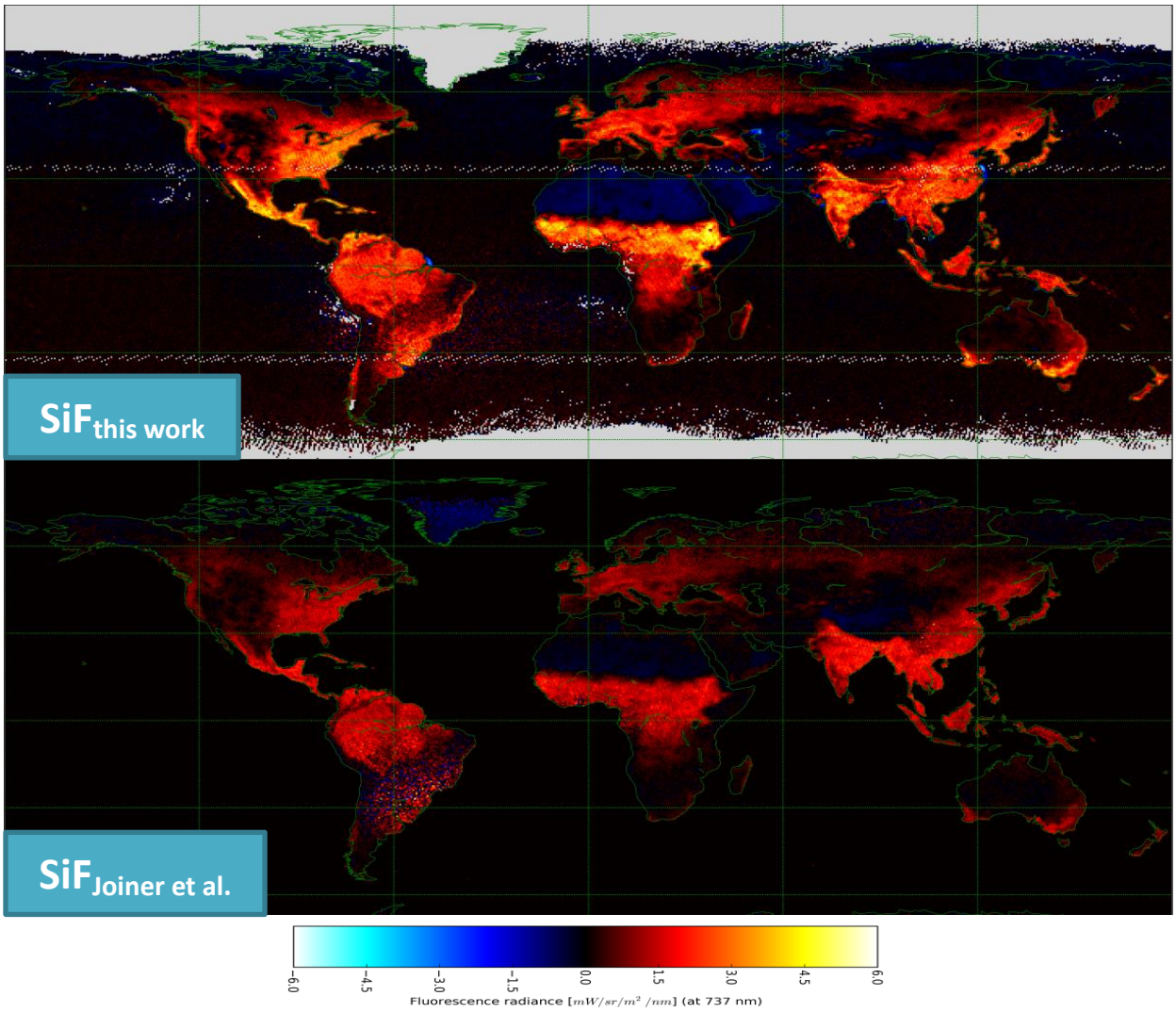


Figure 7.4: Comparison of fluorescence retrieval results for September 2012 means. Upper frame is the baseline retrieval result from this thesis. Bottom frame is the retrieval results published by Joiner et al. (2013).

7.3 Australian Flux sites comparison

In order to properly validate the fluorescence results obtained, the results must be compared with reliable ground based reference data. Simultaneous observations of the same test areas must be made using the investigated method and an already calibrated and tested method.

One way to do this is using carbon flux measurements from flux towers situated in homogeneous vegetated areas. A challenge in this case is the coarse footprint of GOME-2. The fluorescence values obtained from GOME-2 are representative of the average over the entire 40 km by 40 km or 80 km by 40 km area, while the tower based measurements are only representative of their immediate surroundings. Only when the ecosystem is extremely homogenous over a very wide area would this provide a useful validation. Despite this, we can still get a reasonable evaluation from this comparison.

In our case, we used data from nine Australian flux towers[6]; All part of the Australian Ozflux network, which is itself part of Fluxnet (Beringer et al. 2011). The sites include tropical savanna, pasture, sparse woodlands and plains; See Figure 7.5 and Table 7.1 for an overview of all sites.

Figure 7.5: locations of the flux tower sites that were used to evaluate the retrieved SiF data.

Table 7.1: overview of the carbon flux tower test sites.

Site name	Location	Landscape
Yanco (JAXA test site)	34°59'16.1376"S 146°17'26.7252"E	Grass/plains
Whroo	36°40'22.9"S 145°1'34.3"E	Woodland
Sturt Plains	17°9'2.48"S 133°21'0.79"E	Grass/plains
Riggs Creek	36°38'59.59"S 145°34'33.56"E	Grass/pasture
Dry river	15°15'31.62"S 132°22'14.04"E	Open forest savanna
Daly river uncleared	14.1592°S 131.3881°E	Woodland Savanna
Calperum	34°00.163'S 140°35.261'E	Semi-arid sparse shrub/woodland
Alice Springs	22.283°S 133.249°E	Semi-arid woodland
Howard Springs	12°49.428'S 131°15.231'E	Open woodland savanna (wet)

At the sites carbon fluxes are measured along with several meteorological parameters. GPP values are derived from these measurements. We used the provided GPP values for comparison with our SiF measurements in order to evaluate their use as a proxy for GPP. We used available tower data for 2013 to compare with the SiF data for 2013 that we had processed. For each flux tower we selected the SiF grid-cell with center coordinates closest to the tower site coordinates. The tower data was provided on a half hourly basis. In order to compare this to the monthly SiF values, we selected the

[6] Data provided by Jason Beringer, Monash University, Australia

tower data for the MetOp-A overpass time on each day and then averaged those values over each month. We then used a basic linear regression to compare site and satellite data.

In Figure 7.6, the results of a linear regression over all the available data are shown. A statistically significant correlation between tower-derived GPP and GOME-2 derived fluorescence radiance can be observed, although the r^2 value is only 0.48 indicating that while the correlation is significant ($p << 0.05$) it only accounts for half of the variance in the data.

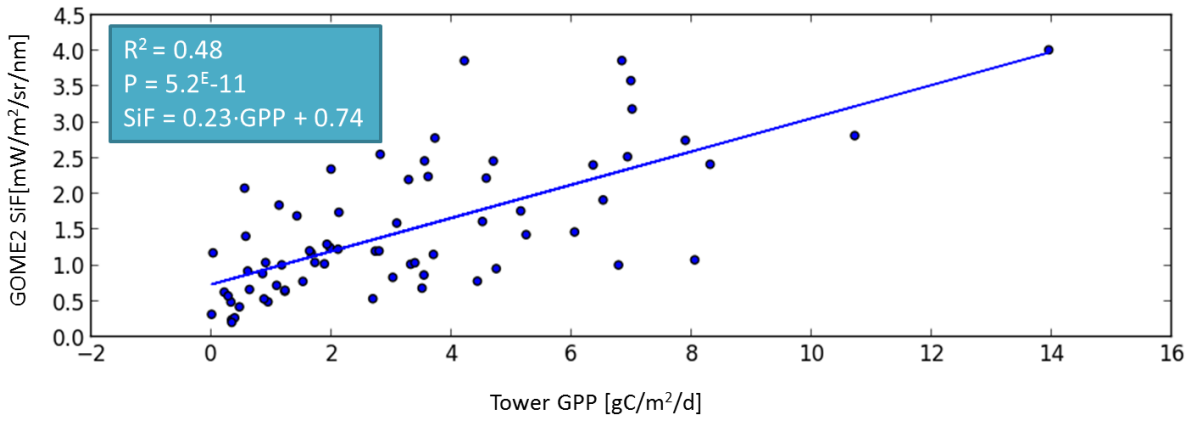


Figure 7.6: Comparison of flux tower derived GPP and GOME2 fluorescence radiance. Data shown here is the combined dataset for all flux towers. The continuous line is the linear regression line.

The correlations between tower GPP and GOME-2 SiF radiance for individual tower sites were also analyzed. Interestingly, the results for different sites differ enormously. Some of the sites have a very good correlation with GOME-2 data, while other sites have no correlation at all. In Figure 7.7 the regression plot is shown for both the best and the worst fit. Based on our limited data, homogeneous pastures and plains show a very good fit, while the more sparse shrub and woodland sites have a rather bad fit.

Leaving out the sites with p-values higher than 0.1, the regression plot changes significantly (as shown in Figure 7.8) resulting in a much tighter correlation. It can also be noticed that even among the sites that show a significant correlation, the empirical relation function obtained can be quite different per site. In particular, we can see two regimes in the sites. Although it is quite likely that different ecosystem types have different relations between GPP and observed SiF, we could not find a common characteristic between the different sites that follow a particular regime. In Table 7.2 the results for all sites with p-values lower than 0.2 are shown.

Table 7.2: Linear regression results obtained from several individual Australian carbon flux tower test sites when compared against the GOME2 SiF results corresponding to that location. Data used in the regression where 2013 monthly means. The slope and intercept are for the equation: GPP=SiF*slope+intercept.

Site	r value	p value	Slope	intercept
Sturt plains	0.898	0.015	4.588	-2.66
Riggs creek	0.947	0.0001	3.958	-2.57
Yanco	0.957	5.1E-5	1.947	-0.61
Alice Springs	0.732	0.061	1.752	-0.21
Daly river uncleared	0.811	0.027	3.287	-2.59
Howard Springs	0.616	0.19	2.034	-0.25

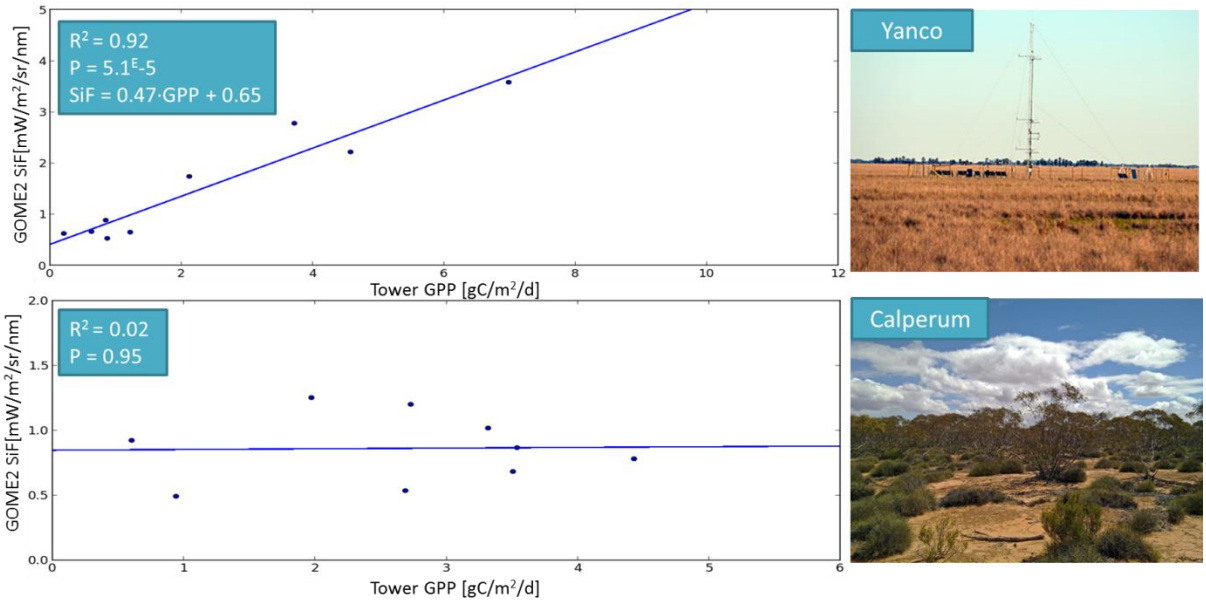


Figure 7.7: Comparison of flux tower derived GPP and GOME2 sun-indiced fluorescence radiance. Data shown here is for two individual flux tower test sites along with a photograph of the site in question. The Calperum site had the highest p-value of all the individual site regressions, while the Yanco JAXA site had the lowest.

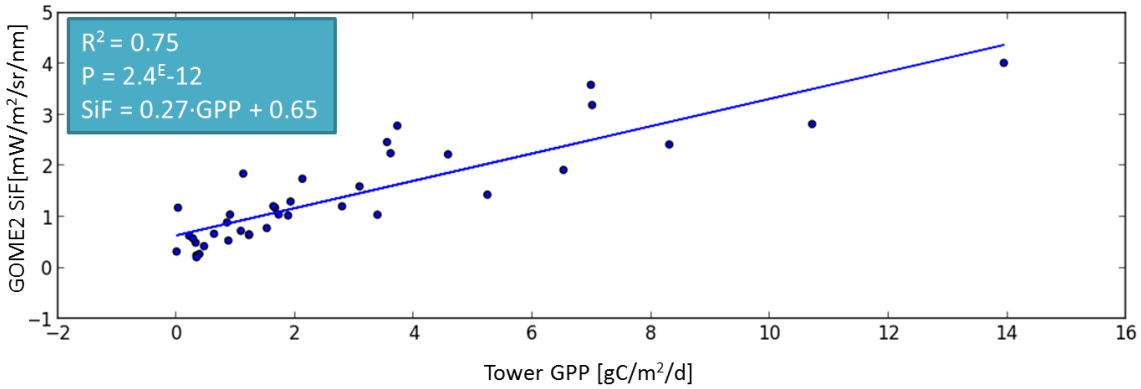


Figure 7.8: Comparison of flux tower derived GPP and GOME2 sun-induced fluorescence radiance. Data shown here is the combined dataset for all flux towers which had a p-value smaller then 0.1 for their individual regressions.

As a first attempt to obtain GPP values from our SiF retrieval we can then use the inverse of the empirical relation obtained from the regression to calculate GPP values for all our retrieved SiF radiance values. It must be noted that this empirical relation is only valid for this retrieval and not for spacebased SiF in general. We can also compare these results with the relations found by Guanter et al. (2014), as displayed in Table 7.3.

Table 7.3: Empirical relation found from linear regression of GPP from flux tower sites and remotely sensed SiF. The slope and intercept are for the equaton: GPP = SiF*slope+intercept.

	slope	intercept
This study (all sites)	4.35	-3.22
This study (6 most significant sites)	3.70	-2.41
Guanter et al. (Europe grassland)	3.71	0.35
Guanter et al. (US cropland)	3.55	-0.88

The GPP values obtained by the relation of Figure 7.8 from SiF are shown in figure Figure 7.9 for September 2013. To give an indication of the plausibility of these values we can compare them with global GPP values obtained from other sources. For example, we show the GPP data obtained from the MODIS GPP product[7] for the same month. This product is derived from other MODIS products, especially the fAPAR product (Running et al. 2004; Zhao et al. 2005)

We can see here, that the GPP values are very similar in the Amazon basin, the Congo basin and Indonesia between the MODIS product and this work. However, outside of these areas we see that in general the GPP from this work is higher than the MODIS GPP (including Autralia), but spatial patterns found herein are similar.

[7] MOD17A2 v05 data obtained from the online Data Pool at NASA Land Processes Distributed Active Archive Center (LP DAAC)

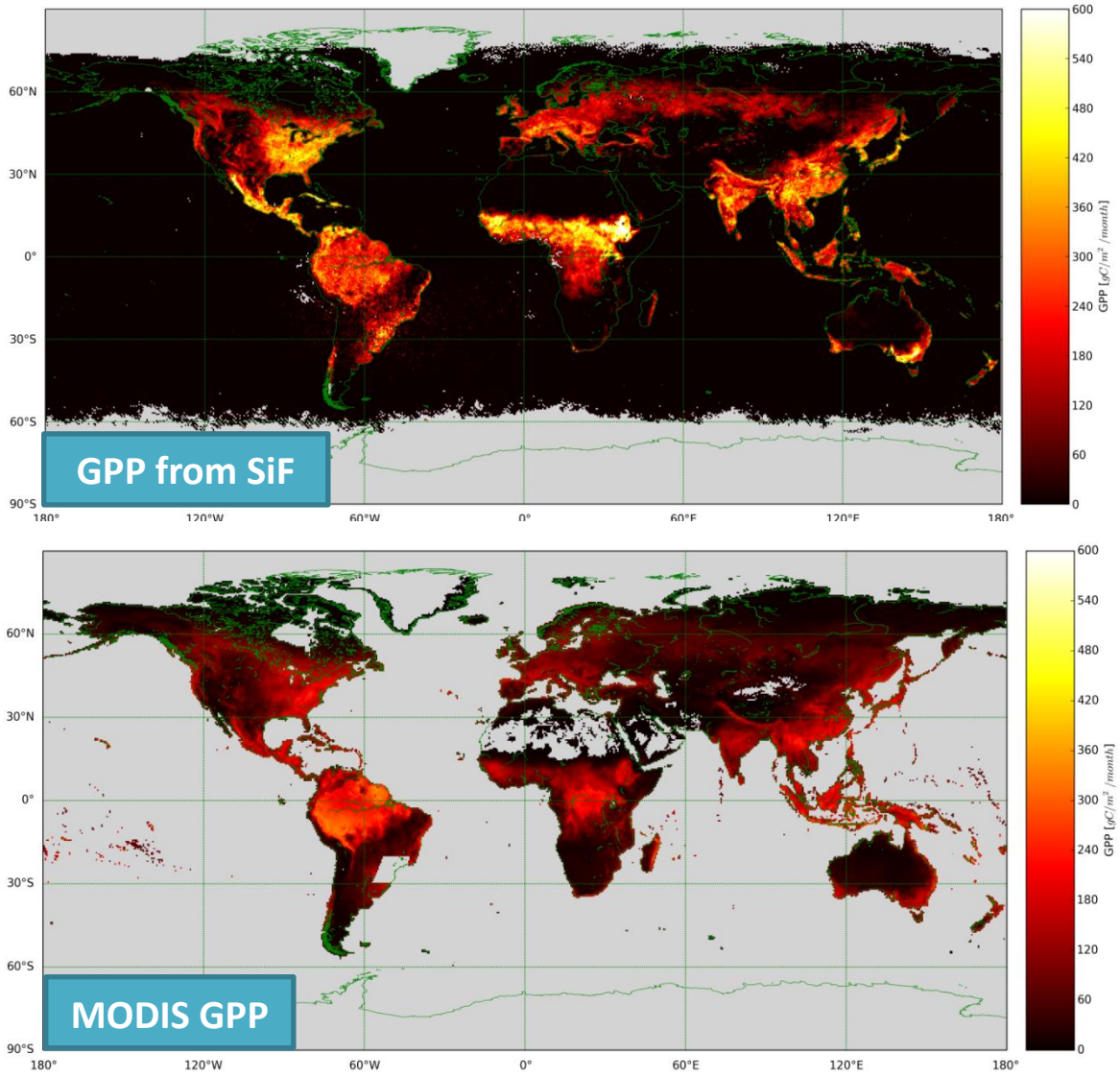


Figure 7.9: Global GPP values for September 2013. Upper frame: GPP obtained by inserting the SiF values measured by GOME2 into the empirical relation obtained in the regression with Australian flux tower data. Lower frame: GPP values obtained from MODIS.

8 Discussion

In this section we will interpret the results and discuss them in the context of other research. Based on this, recommendations for follow-up research are formulated. In **section 8.1** a general analysis of our results as compared to other SiF retrievals and NDVI is given. In **section 8.2** the preliminary fluorescence yield retrieval is discussed. In **section 8.3** the importance of different aspects of the retrieval algorithm are highlighted. In **section 8.4** we evaluate the obtained GPP estimate in the context of existing GPP estimates. In **section 8.5** we assess some of the challenges in validation of spaceborn SiF retrievals.

8.1 Interpretation of baseline results

The spatial and temporal patterns found in the global SiF retrieval are highly similar to patterns found in NDVI (Figure 7.2), as well as vegetation patterns found by qualitative observation of true color imagery (Figure 4.2). This provides a first indication that the retrieved SiF values are likely originating from vegetation.

At the same time though, those spatial patterns found in our SiF retrieval are not completely similar to the patterns found in vegetation indices. There are several striking differences between the patterns in NDVI and in SiF. The tundra areas still show relatively high NDVI values at the end of the summer, while the SiF values are comparatively low already. Similarly the tropical areas show very high NDVI values throughout the year, while the retrieved SiF values are relatively low throughout the year (**section 7.1**). This may suggest that GOME-2 is indeed measuring a vegetation parameter that is not directly related to the color (greenness).

The reason for low SiF values over the tropical rainforests could be due to a high average LUE in these areas, using all absorbed radiance for photosynthesis, leaving little energy for fluorescence. In tropical rainforests neither temperature nor moisture is generally limited. Due to heavy competition between vegetation in tropical rainforests, the availability of direct sunlight is often the most limiting factor (Nemani et al. 2003; Running et al. 2004) which is illustrated in Figure 8.1. As such, tropical plants would use a very high fraction of absorbed sunlight. Therefore, consistent with the supposed negative linear relation between fluorescence yield and LUE in light-limited conditions (van der Tol, Verhoef & Rosema 2009; Damm et al. 2010) (see also Figure 1.4), SiF would be relatively low in the tropical rainforest regions despite their high productivity.

The above example suggests that the SiF we measure can be related to photosynthetic activity and efficiency. This illustrates that SiF measurements can provide complementary information about the state of photosynthesis in ecosystems with respect to vegetation indices such as NDVI, yet at the same time they also illustrate how difficult it can be to properly interpret these results.

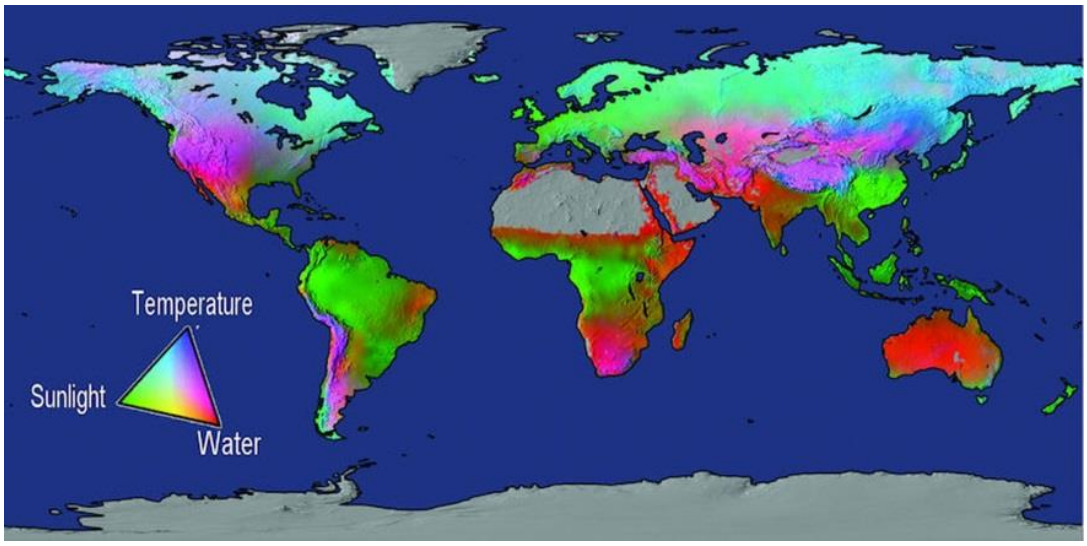


Figure 8.1: Potential climate restraints to plant growth (Running et al. 2004).

Compared with results of Joiner et al. (2013) and Guanter et al. (2014) we obtained very similar SiF patterns. This is as expected since a similar approach was used as in Joiner et al. (2013). While in general our retrieval performs rather well, some negative values are left. These negative values can be observed above shallow and murky surface water. This is indicative that the assumed model cannot fit the observed radiative spectrum in these specific circumstances. It is likely that such scenes contain characteristics which mimic the inverse spectral shape of the assumed shape of fluorescence in the model. It is unclear what these characteristics are, but it follows from the areas where this is observed that they are probably related either to the spectral characteristics of liquid surface water or the suspended sediments therein. Above deeper water, where almost all the light is absorbed, these negative values are not observed.

In the timeseries of fluorescence split by land use type, we can see that the fluorescence patterns obtained from different ecosystem types differ in magnitude and sometimes in shape. It is unclear whether this is due to the different atmospheric circumstances above these areas or due to a different fluorescence response to stress for typical vegetation in these ecosystems. Presumably, both play a role. The empirical relation between GPP and SiF might therefore depend on the ecosystem. A larger more diverse set of flux tower sites than has been used in this study would be needed to test this hypothesis.

8.2 Interpretation of fluorescence yield results

We have performed a preliminary test, combining our SiF product with several other datasets in order to obtain a fluorescence yield product. The plausibility of the results we obtained for fluorescence yield is rather less straightforward to assess than for the fluorescence radiance. We found a pattern where the fluorescence yield is exceptionally low in the tropical areas which could be interpreted as a very high LUE and very high values in arid and semi-arid areas which could be interpreted as a very low LUE considering the theoretical relationship between LUE and SiF (van der Tol, Verhoef & Rosema 2009) and the dominant constraints to plant growth in these areas (Figure 8.1). Very high fluorescence yield values were also found in boreal and temperate areas during summer. We also see strong seasonal variations in the monsoon regions. For example, in the Sahel in Africa we observe that the fluorescence yield values are highest in August coinciding with the peak of

the wet season (e.g. Sultan & Janicot, 2000 and references therein). It is not clear why the fluorescence yield would be as high in these instances.

Because of the potential of using fluorescence yield as a direct proxy for LUE, it is important to be able to unambiguously interpret these results. A regression analysis with LUE derived from independent ground measurements could shed light on this. Unfortunately, LUE data from flux towers was not available during the time of this study for the test sites. Therefore, we would recommend for a future follow-up on this work to concentrate towards obtaining suitable ground measurements for LUE.

8.3 Effect of retrieval parameters

The different retrieval experiments have shown in particular the importance of the selection of the set of reference spectra for the PCA. Since we have found that oceans and particularly ice/snow pixels in the reference dataset yield significantly negative SiF values for substantial areas of the globe (**section 5.1.1**), we reject these retrievals as SiF values must always be positive and thus negative values indicate that the model cannot fit the data properly. The reason for the failure of retrievals based on these reference spectra remains unclear, but we hypothesize that this may be due to low water vapor concentrations in these reference atmospheric columns, suggested by the spatial pattern observed in the oceans (Figure 5.1) which coincides with the inter-tropical convergence zone. Note that our spectral window features a major H_2O absorption band (Figure 1.5). It can also be observed from Figure 5.5 that the atmospheric column above the Sahara desert does not have a low water vapor column all year round, which would suggest that a set of PCs based upon the Sahara would include the effect of this water vapor as long as these parts of the year are included. Indeed, taking a PCA over a full year gives a significant improvement compared with taking a PCA over one-month dataset. This is actually surprising, given how other works use only a single day of data in the reference dataset (Joiner et al., 2013; Guanter et al., 2014). This may also be an interesting direction for further investigation. In the current retrieved data (for 2013), the reference dataset includes data from 2013. When the algorithm is applied to the entire history of GOME-2 measurements, it may be interesting to investigate whether this reference dataset can be applied to other years or whether reference data could better be taken from a moving year period.

The number of principal components that is needed to fully characterize the atmospheric absorption characteristics over all different areas of the globe is surprisingly high given the fact that the highest PCs already explain such a large part of the observed variance in the reference dataset (Figure 5.8). However, when the number of PCs in the retrieval is increased enough, the retrieval values eventually do start to converge. This could suggest that the atmospheric absorption characteristics over the reference areas are not typical for all retrieved areas per se, but those reference absorption spectra do contain similar components as in other areas but with a different weight. This simultaneously illustrates the greatest weakness and the greatest strength of the PCA approach to modelling the atmospheric absorption. It is not straightforward to assign a specific characteristic of the atmospheric column to a specific PC which makes interpretation of the different PCs and their effect on the final retrieval difficult. At the same time, this could be regarded as an advantage since the retrieval is not dependent on a particular model interpretation. As such, this is a very empirically oriented method.

We have found that increasing the order of the surface albedo polynomial increases the relative noise in the retrieval results and therefore a smaller order of the polynomial is preferred. Luckily, the actual spectral shape of the surface albedo is quite simple and thus a low polynomial order seems to suffice, as the spatial patterns do not significantly change between the different retrieval attempts with different polynomial orders; only the order of magnitude of the results and the relative noise changes. We have chosen a fourth order as a reasonable compromise since there is little decrease of the relative noise going from fourth order to third order (Table 5.4).

In our retrieval we have chosen a spectral window for the retrieval spanning from 712 nm to 780 nm, which is roughly double of that suggested by Joiner et al. (2013) which is 712 nm to 758 nm or 747 nm to 780 nm. As such it covers both the O_2-A absorption band and the nearby H_2O band at the same time. We have found that choosing a larger spectral window increases the magnitude of the fluorescence values retrieved while the relative noise in the global image is lower. This work shows that a larger spectral window can be used for retrieving fluorescence and it seems that a larger spectral window could possibly be preferable for retrieving fluorescence.

8.4 Evaluation of obtained GPP estimate

A preliminary global GPP estimate from only the average relation over the sites already gave promising results. Generally our GPP values where similar to those obtained in the model MODIS GPP model, despite having to rely on data from a limited set of ecosystems and a limited geographic extend.

However, the tropical areas show too low GPP values compared to these independent models, similar to how our SiF values also seem to be low compared to the NDVI in the tropics. This may indicate that our GPP estimates may be improved by obtaining an empirical relationship between fluorescence yield and light use efficiency and calculating GPP using this LUE and the APAR, instead of using a direct empirical relationship between fluorescence radiance and GPP.

Guanter et al. (2014) also performed an evaluation of GOME-2 derived SiF with several flux tower sites. They used only grassland and cropland sites in their regression and found a very good correlation (see Table 7.3). The empirical relation we found between flux tower GPP and GOME-2 derived SiF for our quality controlled set of Australian sites is very similar to what Guanter et al. (2014) found for grasslands, except that we have a very different intercept value. This might correspond to the fact that we find consistently higher SiF values globally than in Guanter et al. (2014) and Joiner et al. (2013).

8.5 Validation

In this work we mainly used a limited set of flux tower sites in order to evaluate our results. However, a thorough validation is still pending which is an essential next step for continued development of space-born SiF retrieval. One of the difficulties with this is the large footprint area of GOME-2. With this footprint size it will remain difficult to assess any correlation with small footprint area ground measurements. Luckily, with newer instruments under way such as OCO-2 (Launched in July 2014) (Crisp et al. 2004) and TROPOMI (Veefkind et al. 2012) a properly validated SiF product becomes more feasible. The reduced footprint area combined with a similar spectral resolution as compared with GOME-2, provides for a more reasonable comparison with ground measurements, although care still must be taken to use reasonable homogenously vegetated areas.

One interesting challenge in the interpretation of the fluorescence results is the fact that what we measured fluorescence at the top of the vegetation canopy. The wavelength dependent attenuation of the signal between the top of the canopy and the top of the atmosphere we assert is accounted for with the principal components. However, the quantity that is actually directly related to the photosynthesis mechanism and that would provide a measure for the GPP in an area would be the ground area average chlorophyll level fluorescence. In this work there has been no modelling of the radiative transfer from chlorophyll level to leaf level to canopy level. This is something that could be improved in further research. Leaf geometry within a patch of vegetated land might modulate the measured top of canopy signal (Fournier et al. 2012). An important contribution to attenuation of the chlorophyll fluorescence signal between the chlorophyll level and the canopy level is the overlap between the chlorophyll-a absorption spectrum and the chlorophyll-a emission spectrum which leads to reabsorption within the canopy (Buschmann 2007). In this work as well as preceding work the near-infrared peak of chlorophyll fluorescence is measured which is subject to very little reabsorption, while the far-red peak is not retrieved. An interesting future extension of this work could be to perform a combined retrieval of the far-red and the near-infrared peak and use both to obtain a measure for attenuation within the canopy.

9 Conclusions and recommendations

In this work we implemented an algorithm to retrieve SiF from GOME-2 radiance data. From studying the spatial and temporal patterns therein and comparing this with NDVI, other SiF retrieval attempts, global GPP models and a few ground based GPP measurement sites, we can conclude that the obtained results are actual fluorescence emission patterns.

We have tested several different retrieval settings and selected a reasonable baseline that can be developed towards operational use. Data has been processed for all of 2013 and for September 2012. The algorithm will be further applied to the entire catalog of GOME-2 data in the near future.

We found that selection of reference areas for use in the PCA is the most critical variable for a successful retrieval. References taken from only deserts turned out to be the most reliable, but care must be taken to include reference data from the entire seasonal cycle. Further fine-tuning the retrieval algorithm with particular attention paid towards the reference scenes, and a better understanding towards the underlying causes would be helpful. On the other hand, while retrievals based upon ice or ocean reference pixels could perhaps be made feasible, we recommend this to not be a first priority in further research. We would rather recommend fine-tuning and validating the retrieval based on desert pixels as we have already shown these to be feasible.

We also found that it is feasible to use a larger spectral window than has been used in previous studies, which covers both the O_2-A band and the nearby H_2O band. Retrievals with a broader spectral window have reduced noise and a higher SiF magnitude. An interesting direction for further research would be to explore the feasibility of even larger spectral windows in order to retrieve the complete spectrum of SiF including the far-red peak. In this way, it might be possible to use multiple fit-parameters for the fluorescence contribution to the TOA reflectance, and as such also map variations in the shape of the SiF emissions. We would recommend starting with a two-parameter approach where the SiF spectrum is modelled as a double Gaussian with distinct amplitudes.

In this work, it has been shown that the SiF product obtained here can be related to GPP, at least for grassland and pasture areas. The number of evaluation sites was too little to make very definite conclusions as to the applicability to particular land types. More research is needed towards validating the results and empirically quantifying the link between SiF and GPP.

A promising alternate way to link SiF and GPP is via fluorescence yield and LUE. This could potentially give a more complete GPP estimate by combining the strengths of different types of vegetation proxies. We have obtained a fluorescence yield global dataset with an arbitrary magnitude. The next step would be to evaluate this product and obtaining an empirical relationship between SiF and LUE by comparing with independent LUE derived from ground-based measurements. A validated empirical relation between LUE and fluorescence yield would be useful to integrate in a global carbon cycle model to better estimate the plant photosynthesis.

References

Baker, NR 2008, 'Chlorophyll Fluorescence: A Probe of Photosynthesis in Vivo', *Annual Review of Plant Biology*, vol 59, pp. 89-113.

Beringer, J, Hutley, LB, Hacker, JM, Neininger, B & Tha Paw U, K 2011, 'Patterns and processes of carbon, water and energy cycles across northern Australian landscapes: From point to region', *Agricultural and Forest Meteorology*, vol 151, pp. 1409-1416.

Buschmann, C 2007, 'Variability and application of the chlorophyll fluorescence emission ratio red/far-red of leaves', *Photosynthesis Research*, vol 92, no. 118, pp. 261 - 271.

Callies, J, Corpaccioli, E, Eisinger, M, Hahne, A & Lefebvre, A 2000, 'GOME-2 - MetOp's Second Generation Sensor for Operational Ozone Monitoring', *ESA Bulletin*, vol 102.

Chance & Kurucz 2010, 'An improved high-resolution solar reference spectrum for Earth's atmosphere measurements in the ultraviolet, visible, and near infrared', *Journal of Quantitative Spectroscopy & Radiative Transfer*, vol 111, pp. 1289-1295.

Ciais, P, Sabine, C, Bala, G, Bopp, L, Brovkin, V, Canadell, J, Chlahra, A, DeFries, R, Galloway, J, Heimann, M, Jones, C, Le Quéré, C, Myneni, RB, Piao, S & Thornton, P 2013, 'Carbon and Other Biogeochemical Cycles', in TF Stocker, D Qin, G-K Plattner, M Tignor, SK Allen, J Boschung, A Nauels, Y Xia, V Bex, PM Midgley (eds.), *Climate CHange 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate CHange*, Cambridge University Press, Cambridge, United Kingdom.

Crisp, D, Atlas, RM, Breon, F-M, Brown, LR, Burrows, JP, Ciais, P, Connor, BJ, Doney, SC, Fung, IY, Jacob, DJ, Miller, CE, O'Brien, D, Pawson, S, Randerson, JT, Rayner, P, Salawitch, RJ, Sander, SP, Sen, B, Stephens, GL, Tans, PP, et al. 2004, 'The Orbiting Carbon Observatory (OCO) mission', *Advances in Space Research*, vol 34, no. 4, pp. 700-709.

Damm, A, Elbers, J, Erler, A, Gioli, B, Hamdi, K, Hutjes, R, Kosvancova, M, Meroni, M, Miglietta, F, Moersch, A, Moreno, J, Schickling, A, Sonnenschein, R, Udelhoven, T, Van der Linden, S, Hostert, P & Rascher, U 2010, 'Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP)', *Global Change Biology*, vol 16, no. 1, pp. 171-186.

Dee, DP, Uppala, SM, Simmons, AJ, Berrisford, P, Poli, P, Kobayashi, S, Andrae, U, Balmaseda, MA, Balsamo, G, Bauer, P, Bechtold, P, Beljaars, M, van de Berg, L, Bidlot, J, Bormann, N, Delsol, C, Dragani, R, Fuentes, M, Geer, AJ, Haimberger, L, et al. 2011, 'The ERA-Interum reanalysis: configuration and performance of the darta assimilation system', *Quarterly of the Royal Meteorological Society*, vol 137, no. 656, pp. 553-597.

European Comission Joint Research Centre 2003, *The Global Land Cover Map for the Year 2000*, <http://www-gem.jrc.it/glc2000>.

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*, vol 68, pp. 112-120.

Frankenberg, C, Butz, A & Toon, GC 2011, 'Disentangling chlorophyll fluorescence from atmospheric scattering effects in O2A-band spectra of reflected sunlight', *Geophysical Research Letters*, vol 38, no. 3.

Frankenberg, C, O'Dell, C, Berry, J, Guanter, L, Joiner, J, Köhler, P, Pollock, R & Tayler, TE 2014, 'Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2', *emote Sensing of Environment*, vol 147, pp. 1-12.

Gates, DM, Keegan, HJ, Schleter, JC & Weidner, VR 1965, 'Spectral Properties of Plants', *Applied Optics*, vol 4, no. 1, pp. 11-20.

Guanter, L, Alonso, L, Gómez-Chova, L, Amorós, J, Vila, J & Moreno, J 2007, 'Estimation of solar-induced vegetation fluorescence from space measurements', *Geophysical Research Letters*, vol 34, no. 8.

Guanter, L, Frankenberg, C, Dudhia, A, Lewis, PE, Gómez-Dans, J, Kuze, A, Suto, H & Grainger, RG 2012, 'Retrieval and global assesment of terrestrial chlorophyll fluorescence from GOSAT', *Remote Sensing of Environment*, vol 121, pp. 236-251.

Guanter, L, Rossini, M, Colombo, R, Meroni, M, Frankenberg, C, Lee, J-E & Joiner, J 2013, 'Using field spectroscopy to asses the potential of statistical approaches for the retrieval of sun-induced chlorophyll fluorescence from ground and space', *Remote Sensing of Environment*, vol 133, pp. 52-61.

Guanter, L, Zhang, Y, Jung, M, Joiner, J, Voigt, M, Berry, JA, Frankenberg, C, Huete, AR, Zarco-Tejada, P, Lee, J-E, Moran, MS, Ponce-Campos, G, Beer, C, Camps-Valls, G, Buchmann, N, Gianelle, D, Klumpp, K, Cescatti, A, Baker, JM & Griffis, TJ 2014, 'Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence', *PNAS*, vol 111, no. 14, pp. 1327 - 1333.

Joiner, J, Guanter, L, Lindstrot, R, Voigt, M, Vasilkov, AP, Middleton, EM, Huemmrich, KF, Yoshida, Y & Frankenberg, C 2013, 'Global monitoring of terrestrial chlorophyll fluorescence from moderate spectral resolution near-infrared satellite measurements: methodology, simulations, and applications to GOME-2', *Atmospheric Measurement Techniques*, vol 6, pp. 2803-2823.

Joiner, J, Yoshida, Y, Vasilikov, AP, Middleton, EM, Campbell, PKE, Yoshida, Y, Kuze, A & Corp, LA 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*, vol 5, pp. 809-829.

Joiner, J, Yoshida, Y, Vasilkov, AP, Corp, LA & Middleton, EM 2011, 'First observations of global and seasonal terrestrial chlorophyll fluorescence from space', *Biogeosciences*, vol 8, pp. 637-651.

Knyazikhin, Y, Glassy, J, Privette, JL, Tian, Y, Lotsch, A, Zhang, Y, Wang, Y, Morisette, JT, Votava, P, Myeni, RB, Nemani, RR & Running, SW 1999, *MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document*, <http://eospso.gsfc.nasa.gov/atbd/modistables.html>.

Koelemeijer, RBA, Stammes, P, Hovenier, JW & De Haan, JF 2001, 'A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment', *Journal of Geophysical Research*, vol 106, pp. 3475-3490.

Levenberg, K 1944, 'A Method for the Solution of Certain Non-Linear Problems in Least Squares', *Quarterly of Applied Mathematics*, vol 2, pp. 164-168.

Lichtenthaler, HK, Buschmann, C, Rinderle, U & Schmuck, G 1986, 'Application of chlorophyll fluorescence in ecophysiology', *Radiation and Environmental Biophysics*, vol 25, pp. 297-308.

Marquardt, D 1963, 'An Algorithm for Least-Squares Estimation of Nonlinear Parameters', *SIAM Journal on Applied Mathematics*, vol 11, no. 2, pp. 431-441.

Maxwell, K & Johnson, GN 2000, 'Chlorophyll fluorescence - a practical guide', *Journal of Experimental Botany*, vol 51, no. 345, pp. 659-668.

Monteith, J 1972, 'Solar-radiation and productivity in tropical ecosystems', *Journal of Applied Ecology*, vol 9, pp. 747-766.

Monteith, J 1977, 'Climate and efficiency of crop production in Britain', *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, vol 281, pp. 277-294.

Moya, I, Camenen, L, Evain, S, Goulas, Y, Cerovic, ZG, Latouche, G, Flexas, J & Ounis, A 2004, 'A new instrument for passive remote sensing 1. Measurements of sunlight-induced chlorophyll fluorescence', *Remote sensing of environment*, vol 91, pp. 186-197.

Nemani, RR, Keeling, CD, Hashimoto, H, Jolly, WM, Piper, SC, Tucker, CJ, Myeni, RB & Running, SW 2003, 'Climate-Driven Increases in GLobal Terrestrial Net Primary Production from 1982 to 1999', *Science*, vol 300, pp. 1560 - 1563.

Olofsson, P & Eklundh, L 2007, 'Estimation of absorbed PAR across Scandinavia from satellite measurements. Part II: modeling and evaluating the fractional absorption', *Remote Sensing of Environment*, no. 110, pp. 240-251.

Pedros, R, Moya, I, Goulas, Y & Jacquemoud, S 2008, 'Chlorophyll fluorescence emission spectrum inside a leaf', *Photochemical & Photobiological Sciences*, vol 7, no. 4, pp. 498-502.

Plascyk, JA & Gabriel, FC 1975, 'The Fraunhofer line discriminator MKII - an airborne instrument for precise and standardized ecological limuniscence measurements', *IEEE Transactions on Instrumentation and Measurement*, vol 24, pp. 306 - 313.

Quick, WP & Horton, P 1984, 'Studies on the induction of chlorophyll fluorescence in barley protoplasts. II. Resolution of fluorescence quenching by redox state and the transthylakoid pH gradient', *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol 220, no. 1220, pp. 371-382.

Running, SW, Ramakrishna, NR, Heinsch, FA, Zhao, M, Reeves, M & Hashimoto, H 2004, 'A continuous Satellite-Derived Measure of Global Terrestrial Primary Production', *BioScience*, vol 54, no. 6, pp. 547 - 560.

Schaefer, K, Schwalm, CR, Williams, C, Arain, MA, Barr, A, Chen, JM, Davis, KJ, Dimitrov, D, Hilton, TW, Hollinger, DY, Humphreys, E, Poulter, B, Raczka, BM, Richardson, AD, Sahoo, A, Thornton, P, Vargas, R, Verbeeck, H, Anderson, R, Baker, I, et al. 2012, 'A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis', *Journal of Geophysical Research*, vol 117, no. G3.

Schulze, E-D, Beck, E & Müller-Hohenstein, K 2005, *Plant Ecology*, Springer, Berlin Heidelberg.

Stöckli, R, Vermote, E, Saleous, N, Simmon, R & Herring, D 2005, *The Blue Marble Next Generation - A true color earth dataset including seasonal dynamics from MODIS*, <http://earthobservatory.nasa.gov/Features/BlueMarble/bmng.pdf>.

Sultan, B & Janicot, S 2000, 'Abrupt shift of the ITCZ over West Africa and the intra-seasonal variability', *Geophysical Research Letters*, vol 27, no. 20, pp. 3353-3356.

Thomas, GE & Stamnes, K 2002, *Radiative transfer in the atmosphere and ocean*, 2nd edn, Cambridge university press, Cambridge.

Turner, DP, Ritts, WD, Cohen, WB, Gower, ST, Zhao, M, Running, SW, Wofsy, SC, Urbanski, S, Dunn, AL & Munger, JW 2003, 'Scaling Gross Primary Production (GPP) over boreal and deciduous forest landscapes in support of MODIS GPP product validation', *Remote Sensing of Environment*, vol 88, no. 3, pp. 256-270.

van der Tol, C, Verhoef, W & Rosema, A 2009, 'A model for chlorophyll fluorescence and photosynthesis at leaf scale', *Agricultural and forest meteorology*, vol 149, no. 1, pp. 96-105.

Veefkind, JP, Aben, I, McMullan, K, Förster, H, de Vries, J, Otter, G, Claas, J, Eskes, HJ, de Haan, JF, Kleipool, Q, van Weele, M, Hasekamp, O, Hoogeveen, R, Landgraf, J, Snel, R, Tol, P, Ingmann, P, Voors, R, Kruizinga, B, Vink, R, 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*, vol 120, pp. 70-83.

Wessel, P & Smith, WHF 1996, 'A Global, Self-consistent, Hierarchical, High-Resolution Shoreline Database', *Journal of Geophysical Research: Solid Earth*, vol 101, no. B4, pp. 8741-8743.

Wold, H 1966, 'Estimation of principal components and related models by iterative least squares', in PR Krishnaiah (ed.), *Multivariate Analysis*, Academic Press, New York.

Wold, H 1973, 'Nonlinear Iterative Patial Least Squares (NIPALS) Modeling: Some Current Developmensts', in PR Krishnaiah (ed.), *Multivariate Analysis II, proceedings of an International Symposium on Multivariate Analysis held at Wright State University, Dayton, Ohio*, Academic Press, New York.

Zarco-Tejada, PJ, Miller, JR, Mohammed, GH & Noland, TL 2000, 'Chlorophyll fluorescence effects on vegetation apparent reflectance: I. Leaf-level measurements and model simulation', *Remote Sensing of Environment*, vol 74, no. 3, pp. 582-595.

Zhang, X, Goldberg, M, Tarpley, D, Friedl, MA, Morisette, J, Kogan, F & Yu, Y 2010, 'Drought-induced vegetation stress in southwestern North America', *Environmental Research Letters*, vol 5, no. 2.

Zhao, M, Heinsch, FA, Nemani, RR & Running, SW 2005, 'Imporovements of the MODIS terrestrial gross and net primary production global data set', *Remote Sensing of Environment*, vol 95, pp. 164 - 176.

Appendix A: Additional algorithm aspects

A.1 High-sampling Interpolation

The measured solar irradiance spectrum and the earthshine spectrum measured by GOME-2 do not have exactly the same wavelength bins. In order to use them together we need to interpolate the solar spectrum to the binning wavelengths of the earthshine spectrum. However, the spectra contain fine Fraunhofer structures that are narrower than the resolution of the measurements, which are important for interpolation. Therefore, we need to apply a so-called high-sampling interpolation technique. For this we need a high-resolution reference solar spectrum which contains information about the solar Fraunhofer lines. In this case, we have used the SAO 2010 reference solar spectrum (K. Chance and R.L. Kurucz, 2010). In order to make this comparable to the instrumental measurements, the reference spectrum needs to be convolved with the GOME-2 slit function. This slit function has been determined through lab measurements on the instrument before launch. From this convolved function, the spline interpolation function S(λ) is then determined. This spline function is then used to interpolate from the original solar irradiance wavelength bins λ_s to the earthshine radiance wavelength bins λ_e.

$$F_m(\lambda_e) = F_m(\lambda_s)\frac{S(\lambda_e)}{S(\lambda_s)}$$

These interpolated values are then used to calculate the TOA reflectance.

The SAO 2010 reference solar spectrum convolved with the GOME-2 slit function is also used as the solar irradiance in equation (2.1).

A.2 Pixel Selection Algorithm

The pixels of the land classification map are much smaller than the footprint of the GOME-2 pixels. Therefore, we regard the former as infinitesimal points and the latter as finite surfaces for the purposes of the selection algorithm. Since the footprint is an irregular quadrangle but the corner-coordinates of the GOME-2 footprint are explicitly known, I make use the following mathematical consideration to determine which of the points are within the footprint:

Given a line segment AB and a point P on a two-dimensional plane, one can define a vector **a** pointing from A to B and a vector **p** pointing from A to P (see figure). If the cross product is greater than zero:

$$\boldsymbol{a} \times \boldsymbol{p} > 0$$

then the point P lies on the right-hand side of the vector **a**.

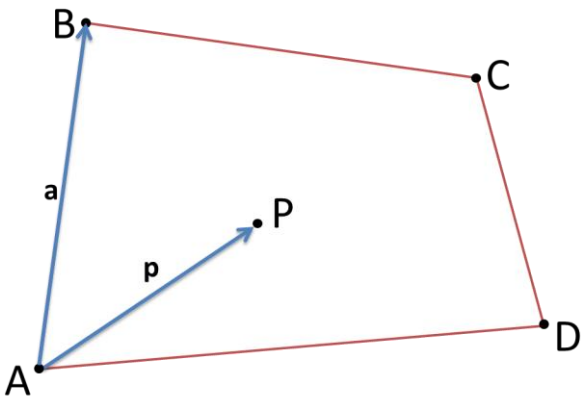


Figure 0.1: Schematic illustration of an arbitrary quadrangular footprint. The perimeter of the footprint is indicated in red, while the set of vectors corresponding to the first line-segment are indicated in blue.

Considering this, given a quadrangle ABCD, a point P lies within the boundaries of the quadrangle if it lies on the right-hand side of each of the four vectors **a**, **b**, **c** and **d** that correspond to the line segments AB,BC,CD and DA. This can then be expressed as the following set of conditions:

$$\boldsymbol{a} \times \boldsymbol{p} > 0$$

$$\boldsymbol{b} \times \boldsymbol{p} > 0$$

$$\boldsymbol{c} \times \boldsymbol{p} > 0$$

$$\boldsymbol{d} \times \boldsymbol{p} > 0$$

If these conditions are met regarding any point that is not classified as totally barren, the GOME-2 pixel is rejected.

A.3 Sun-glint rejection

The first filter that is applied across both the training scenes as well as the target scenes is the sun-glint rejection filter. Sun-glint is the spot on the earth surface that contains the specular reflection of the sun when seen from the position of the satellite. Specular reflection poses a problem for our retrieval since it strongly violates the Lambertian assumption that we have used in our radiative transfer model. Therefore, only diffuse reflectance is used. Only a suitably smooth surface can contain a specular reflection, which in practice means surface water. Therefore, the filter is applied only towards ocean scenes.

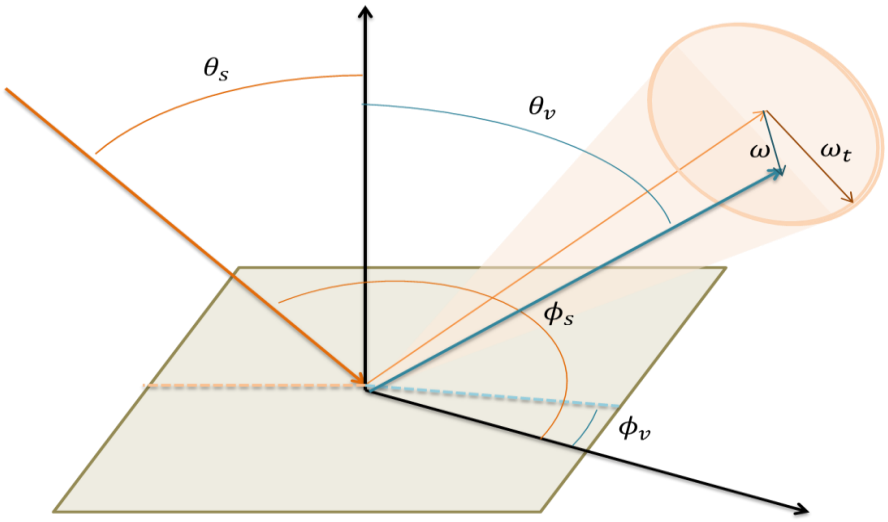


Figure 0.2: Clarifying illustration for the sun-glint angles.

For the purposes of our rejection scheme we neglected the specifics of the specular reflection and used the simplest possible approximation, which is a step-function approach. Here, we calculate the angle of the sun-glint cone ω, which is the angle between the view angle of the instrument and the angle of specular reflection:

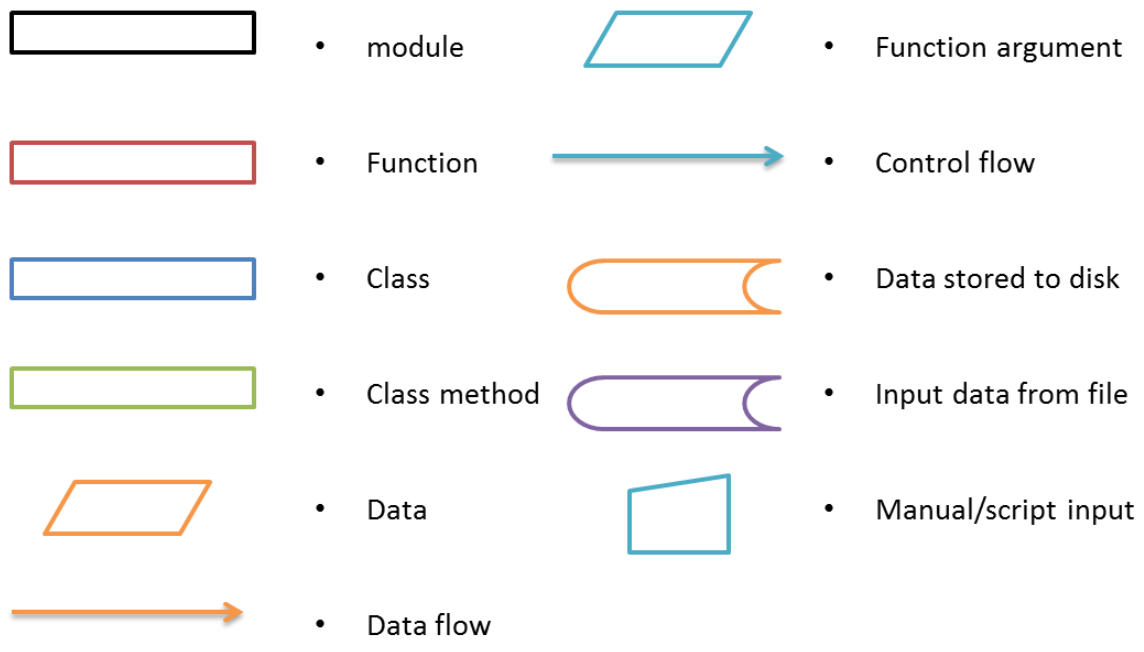
$$\cos\omega = \cos\theta_v \cos\theta_s + \sin\theta_s \cos|\phi_v - \phi_s - \pi|$$

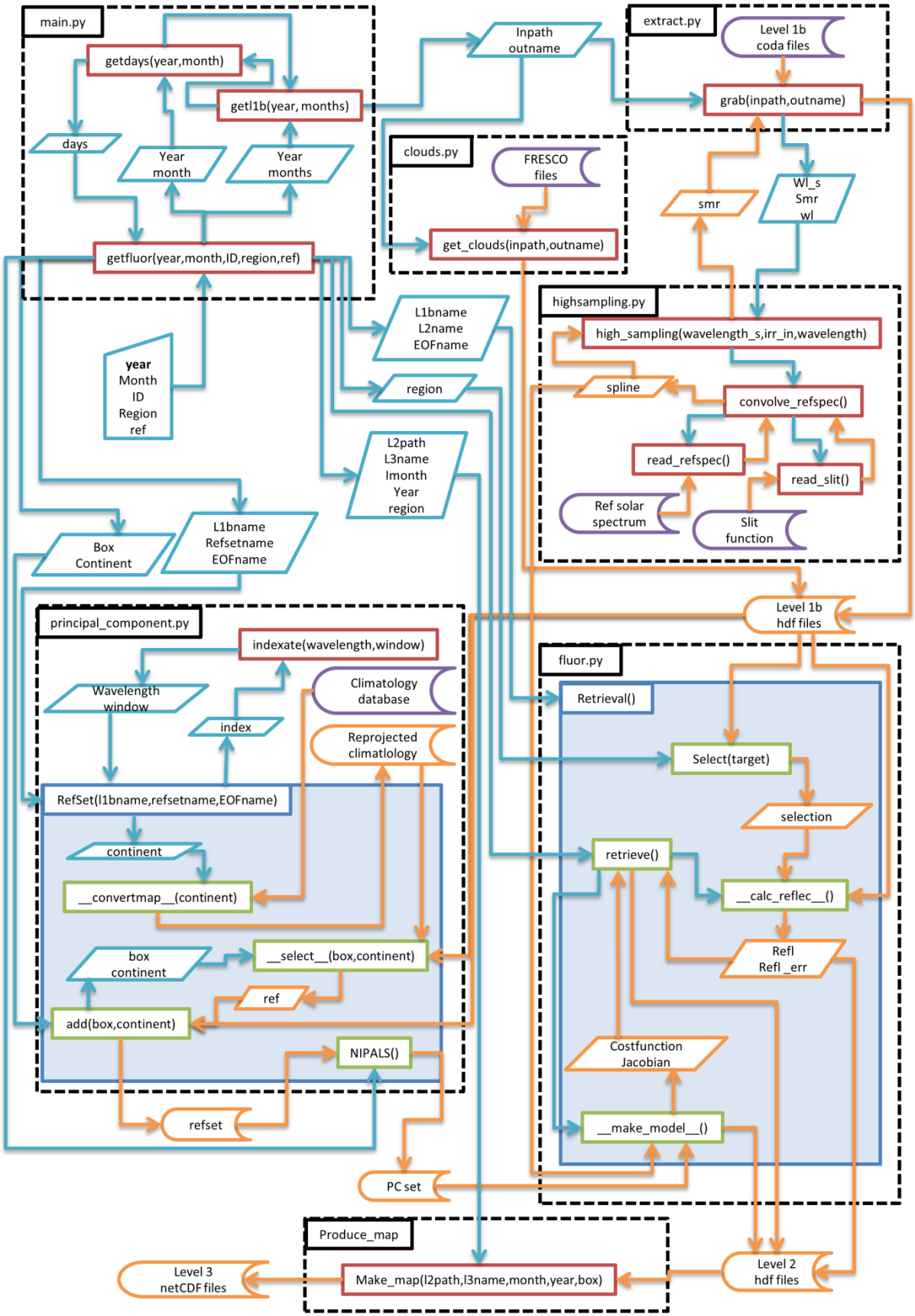
Where θ_s is the solar zenith angle, θ_v is the viewing zenith angle, ϕ_s is the solar azimuth angle and ϕ_v is the viewing azimuth angle. When the specular reflection angle is lower than a prescribed threshold $\omega < \omega_t$, then it is assumed to be within the sun-glint cone and rejected. We have chosen the threshold as 18°.

Appendix B: Python Implementation Overview

The complete retrieval algorithm from the fetching of operational level 1b data to the writing of gridded level 3 data to disk is implemented in Python using the SciPy package and several ancillary packages. The code is aimed to be able to be further developed into an operational retrieval code, as well as to facilitate further experimentation and to be easily expanded and adapted by future collaborators. Therefore the code is mostly modular.

A flowchart of the code is provided on the next page. Not shown in the flowchart is the config.py module which contains all retrieval settings. A legend is given below:




main.py
getdays(year,month)
getl1b(year, months)
days
Year month
Year months
getfluor(year,month,ID,region,ref)
year
Month
ID
Region
ref
Inpath
outname
extract.py
Level 1b
coda files
grab(inpath,outname)
clouds.py
FRESCO
files
get_clouds(inpath,outname)
smr
Wl_s
Smr
wl
highsampling.py
high_sampling(wavelength_s,irr_in,wavelength)
spline
convolve_refspec()
read_refspec()
read_slit()
Ref solar
spectrum
Slit
function
L1bname
L2name
EOFname
region
L2path
L3name
lmonth
Year
region
Box
Continent
L1bname
Refsetname
EOFname
principal_component.py
indexate(wavelength,window)
Wavelength
window
index
Climatology
database
Reprojected
climatlology
RefSet(l1bname,refsetname,EOFname)
continent
__convertmap__(continent)
box
continent
__select__(box,continent)
ref
add(box,continent)
NIPALS()
refset
Level 1b
hdf files
fluor.py
Retrieval()
Select(target)
selection
retrieve()
__calc_reflec__()
Refl
Refl _err
Costfunction
Jacobian
__make_model__()
PC set
Produce_map
Level 3
netCDF files
Make_map(l2path,l3name,month,year,box)
Level 2
hdf files

Appendix C: Glossary

Table 0.1: used acronyms and abbreviations.

Acronym	Full name
DOY	Day of the year
ECMWF	European Centre for Medium-range Weather Forecasts
ERA-interim	ECMWF reanalysis interim
EUMETSAT	EUropean organization for the exploitation of METeorological SATellites
EVI	Enhanced vegetation index
fAPAR	Fraction of absorbed photosynthetically active radiation
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
FWHM	Full-width half maximum
GOME-2	Global Ozone Monitoring Experiment 2
GOSAT	Greenhouse gasses Observing SATellite
GPP	Gross primary productivity
iPAR	Instantaneous Photosynthetically Active Radiation
ITCZ	Inter-Tropical Convergence Zone
JRC	Joint Research Centre
KNMI	Koninklijk Nederlands Meteorologisch Instituut
Lat.	Latitude
Lon.	Longitude
LUE	Light use efficiency
MODIS	MODerate resolution Imaging Spectroradiometer
NDVI	Normalized difference vegetation index
NEP	Net ecosystem productivity
NIPALS	Non-linear iterative partial least squares
NIR	Near infrared
NPP	Net primary productivity
OCO-2	Orbiting Carbon Observatory 2
PAR	Photosynthetically active radiation
PC	Principal component
PCA	Principal component analysis
RMS	Root mean square
RN	Relative Noise
SCIAMACHY	Scanning Imaging Absorption spectrometer for Atmospheric ChartographY
SiF	Sun induced fluorescence
SNR	Signal to noise ratio
SZA	Solar zenith angle
TOA	Top of atmosphere
TROPOMI	TROPOspheric Monitoring Instrument
USGS	United States Geological Survey
VZA	View zenith angle

Table 0.2: Used symbols

symbol	name
F_m	Measured solar irradiance
F_s	Solar irradiance
$I_{F,0}$	Amplitude of fluorescence
I_F	Sun-induced Fluorescence radiance
I_m	Measured earthshine radiance
a_i	Coefficient of i^{th} polynomial term of surface reflectance
b_j	Coefficient of j^{th} basis function of atmospheric slant optical depth
f_j	J^{th} basis function of atmospheric slant optical depth
$\overline{T}$	Atmospheric transmission from surface to TOA along angle of the instrument
$\alpha_{s,int}$	Daily integrated solar elevation
α_s	Solar elevation
δ_s	Solar declination angle
θ_s	Solar zenith angle
θ_v	View zenith angle
μ_0	Cosine of solar zenith angle
μ_H	Center wavelength of NIR peak of SiF
μ_L	Center wavelength of red peak of SiF
μ_c	Mean SiF over all continents
μ_g	Mean SiF over the entire globe
μ_o	Mean SiF over all oceans
$\overline{\rho}$	Spherical albedo of the atmosphere
ρ_0	Atmospheric reflectance
ρ_{NIR}	TOA reflectance in the near infrared
ρ_m	TOA measured reflectance
ρ_p	Model TOA reflectance
ρ_r	reference reflectivity
ρ_{red}	TOA reflectance in the red
ρ_s	Surface reflectance
ρ_t	Target reflectivity
σ_H	width of the NIR peak of SiF
σ_L	Width of the red peak of SiF
σ_c	Standard deviation in SiF over all continents
σ_g	Standard deviation in SiF over the entire globe
σ_m	Instrument measurement error
σ_o	Standard deviation in SiF over all oceans
σ_s	Scattering cross section
ψ_s	Solar hour angle
ϕ_d	Heat dissipation yield
ϕ_f	Fluorescence yield
ϕ_s	Solar azimuth angle
ϕ_v	View azimuth angle
Q	Heat
X	Array of reference spectra
c	Ratio between NIR and red peaks of SiF
I	Identity matrix

$\boldsymbol{J}$	Jacobian matrix
$\boldsymbol{p}$	Principal component loading vector
$\boldsymbol{t}$	Principal component score vector
$\mathcal{T}$	Atmospheric transmission from TOA to surface along the angle of incoming solar radiation
$\boldsymbol{\alpha}$	Relaxation factor in minimization scheme
$\boldsymbol{\varepsilon}$	Obliquity of the earth
$\boldsymbol{\lambda}$	wavelength
$\boldsymbol{\mu}$	Cosine of view zenith angle
$\boldsymbol{\tau}$	Two-way atmospheric slant optical depth
$\boldsymbol{\varphi}$	latitude
$\boldsymbol{\beta}$	parameter vector
$\boldsymbol{\delta}$	Iterative parameter vector adjustment term