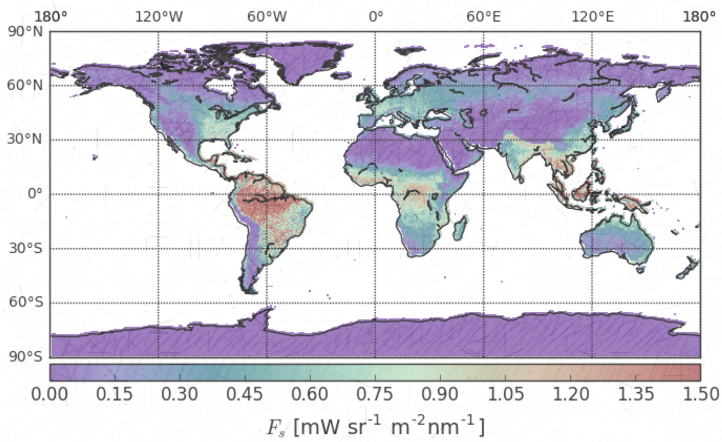
Retrieving Sun-Induced Fluorescence From the Global Ozone Monitoring Experiment 2

Author
Erik van Schaik

Supervisors
Folkert Boersma
Piet Stammes
Maurits Kooreman



An internship report
Royal Netherlands Meteorological Institute (KNMI)
Wageningen University
December 16, 2016

Acknowledgements

First and foremost I would like to thank the Royal Netherlands Meteorological Institute (KNMI) for facilitating this internship. This work would not have been possible without the guidance of my supervisors: Folkert Boersma, Piet Stammes and Maurits Kooreman. In addition, I would acknowledge Olaf Tuinder and Johan de Haan for their support on processing GOME-2 data and operating DISAMAR.

During this internship I received feedback and data from outside of KNMI as well. I would like to thank Bram Sanders and Willem Verstraeten for the fruitful discussions. Christian Frankenberg kindly supplied me with GOSAT and OCO-2 data and I would like to acknowledge Joanne Joiner for making their fluorescence product publicly available. In addition, Markus Reichstein gave me access to their GPP product. In addition I would like to acknowledge all people involved with GOME-2 project.

Abstract

Chlorophyll fluorescence is the re-emittance of radiation at a higher wavelength by vegetation and originates from the same internal apparatus as photosynthesis. Previous studies have shown that remotely sensed fluorescence contains information on actual photosynthesis rates and can thus provide constraints on gross primary production. This study expands upon a previous effort to retrieve chlorophyll fluorescence from the Global Ozone Monitoring Instrument 2 (GOME-2). As independent validation data is scarce simulated top-of-atmosphere spectra with a known fluorescence signal are used to assess retrieval performance. Based on this, four major changes are implemented in the Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval (SIFTER) algorithm: 1) The spectral fitting window is narrowed from 712-783 nm to 734-758 nm; 2) the number of principle components used to simulate atmospheric effects is reduced from 35 to 8; 3) Retrievals are rejected when autocorrelation in the fit residuals is larger than 0.2; 4) all retrievals within a 0.5-degree latitudinal bin are corrected with the daily mean fluorescence over all ocean areas within that bin to remove a zero-level offset.

The new retrieval (SIFTER v2) has correlates better with other space-born fluorescence products compared to SIFTER v1. Especially tropical forests show a large increase in estimated fluorescence for SIFTER v2. In contrast to other GOME-2 based fluorescence retrievals SIFTER v2 is able to pick up regional reductions in fluorescence that coincide (in time and space) with severe drought events.

Correlation between multi-year mean fluorescence and gross primary production (based on two independent datasets) is highest for SIFTER v2, compared to version 1 and several other space-borne fluorescence products. This highlights the potential of SIFTER v2 to be used as an independent new constraint on photosynthetic activity on regional to global scales.

Contents

Chapter 1

Introduction

1.1 Fluorescence and the carbon budget

During the last century atmospheric carbon dioxide (CO_2) concentrations have increased due to a combination of fossil fuel combustion, land-use change and biomass burning (Intergovernmental Panel on Climate Change, 2014). The increased CO_2 levels alter the radiative balance of the Earth and lead to increased surface temperatures (Friedlingstein et al., 2014; Jackson et al., 2015). The changes in both atmospheric CO_2 and surface temperatures are expected to cause profound changes in many components of the Earth system for possibly centuries to come (Solomon et al., 2009).

Not all CO_2 from anthropogenic origin ends up in the atmosphere. Approximately 56% of human-emitted CO_2 is removed from the atmosphere by marine and terrestrial sinks (Le Quéré et al., 2015). In absolute terms, uptake of CO_2 during photosynthesis is the single largest sink of atmospheric carbon. However, it is also the most uncertain flux within the global carbon budget (Ciais et al., 2013). Estimates of the gross carbon uptake (GPP) range from annual sums of 100 - 180 Pg C yr^{-1}, highlighting the need for new methods to better constrain this flux (Jung et al., 2011; Anav et al., 2015). A major obstacle in estimating GPP is that direct observation is difficult. Current measurement techniques (e.g. eddy-covariance flux towers) measure net carbon exchange (NEE), which is the combined flux of gross carbon uptake (GPP) and respiration. Estimates of the gross carbon fluxes for the biosphere on regional to continental scale therefore rely on model predictions (both statistical and mechanistic). However, the heterogeneity in vegetation, hydrology and other environmental conditions leads to large biases and uncertainties in one of the most important fluxes in the global carbon budget. In addition, a lack of independent global constraints on GPP severely hinders the development of terrestrial biosphere models. These models try to emulate the physical behaviour of vegetation under different environmental conditions and are currently the main method for predicting the future carbon cycle.

A promising new constraint on global GPP is remotely sensed fluorescence. Under sunlit conditions, plants use approximately 80% of absorbed photosynthet-

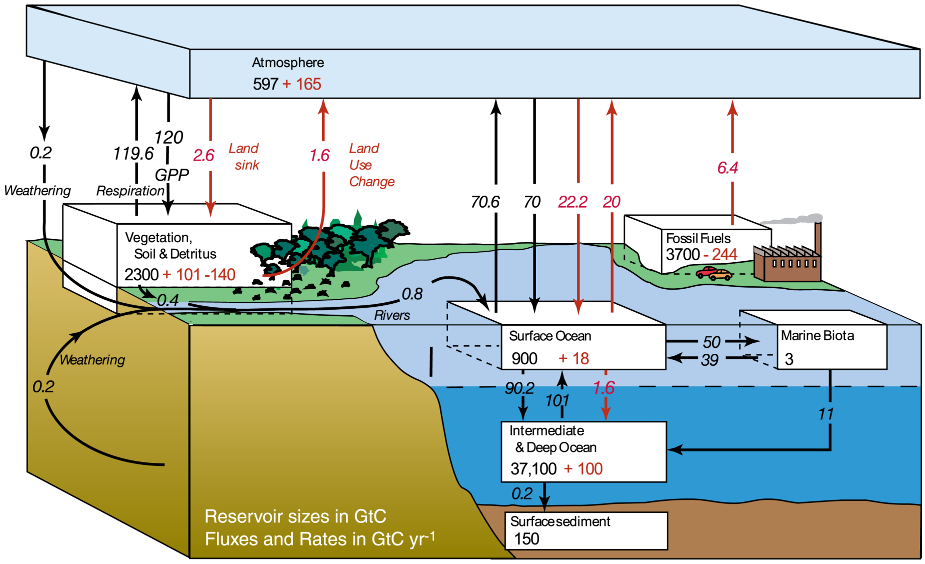


Figure 1.1: *Figure taken from Intergovernmental Panel on Climate Change (2014).* The global carbon budget. Black colour indicates the natural reservoir sizes (in Gt C) and fluxes (in Gt C yr^{-1}). Red indicates the changes in reservoirs and fluxes due to anthropogenic perturbation.

ically active radiation to drive photosynthesis, whilst the remainder is released as heat ($\sim 20\%$) or re-emitted at a higher wavelength ($\sim 1-2\%$) (Baker, 2008). The latter process is commonly known as sun-induced fluorescence (SIF), and will be referred to as fluorescence throughout the rest of this work. Chlorophyll fluorescence is directly related to the internal photosynthetic functioning of vegetation (i.e. photosystems I and II), and its emission has a spectrally smooth curve with peaks at 685 nm (red fluorescence) and 730 nm (far-red fluorescence, see Figure 1.2). Because carbon assimilation and fluorescence are both being derived through photosynthesis, their respective strengths are closely related. Previous studies found a linear correlation between GPP and fluorescence at leaf level, even under drought conditions (Schreiber and Bilger, 1987; Flexas et al., 2002).

Whilst fluorescence has long been used in plant sciences, only recently it was found to be detectable from space (Joiner et al., 2011; Frankenberg et al., 2011b). This lag is mainly caused by the weak signal of fluorescence, which only constitutes 1-2% of the total top-of-atmosphere (TOA) reflectance signal. The first global retrievals therefore relied on a method based on Fraunhofer line filling using observations from the Greenhouse gasses Observing SATtelite (GOSAT). Fraunhofer lines are narrow bands in the solar spectrum caused by absorption in the outer layers of the sun. Several of these lines overlap with the smooth emission spectrum of fluorescence. The observed top-of-atmosphere reflectance from the Earth above vegetated surfaces has a decreased depth in these lines (compared to the solar spectrum) due to fluorescence filling (Frankenberg et al., 2011a). This method has been successfully applied to retrieve fluorescence for the SCanning Imaging Absorption SpectroMeter for Atmospheric Chartogra-

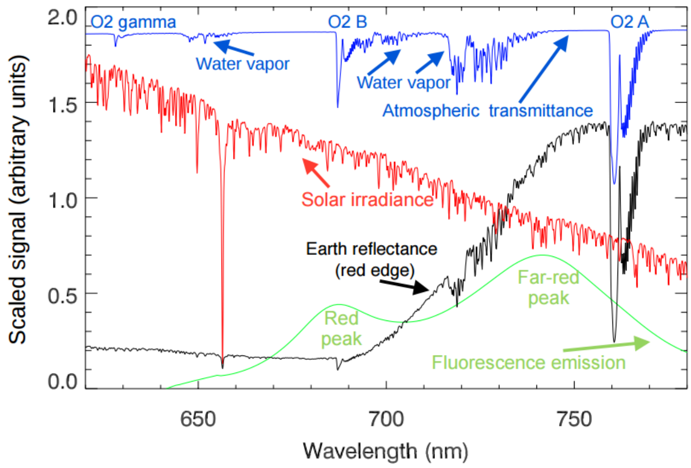


Figure 1.2: *Taken from Joiner et al. (2016).* Simulated spectra of different components required for fluorescence retrievals (solar irradiance, atmospheric transmittance, earth reflectance and fluorescence emission).

phY (SCIAMACHY) and the Orbiting Carbon Observatory-2 (OCO-2) as well (Joiner et al., 2012; Frankenberg et al., 2014). However, by design both GOSAT and OCO-2 only scan small areas of the globe on a daily basis as they take sampled measurements. SCIAMACHY has a lower spectral resolution and therefore relies on filling in the 866 nm Ca II solar Fraunhofer line. This line is far removed from the peak fluorescence (around 737 nm) and therefore suffers from low single measurement precision (Joiner et al., 2013).

Joiner et al. (2013) presented a method to derive fluorescence from observations made with the Global Ozone Monitoring Experiment (GOME-2) optical spectrometer on board of the MetOp-A satellite. GOME-2 has a lower spectral resolution of 0.5 nm and a nadir footprint of 80-by-40 km. Because of the spectral resolution it can not resolve individual Fraunhofer lines. Despite this, efforts to retrieve fluorescence from GOME-2 are undertaken because of the high spatial and temporal resolution of the sensor, compared to GOSAT and OCO-2. To retrieve fluorescence from GOME-2 observations Joiner et al. (2013) proposed to split the TOA signal into three different components: atmospheric absorption, surface reflectance and fluorescence (see Figure 1.2). This approach is followed by Sanders et al. (2016), who found the Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval (SIFTER) algorithm at the Royal Netherlands Meteorological Institute (KNMI). The concepts of retrieving fluorescence from GOME-2 will now be discussed in more detail.

1.2 Retrieving fluorescence from GOME-2

As mentioned in the previous section, fluorescence retrievals for the GOME-2 instrument are not able to utilize the Fraunhofer-line filling method. However, Joiner et al. (2013) found that fluorescence can be retrieved by optimizing the

TOA spectrum to a combination of surface, atmosphere and fluorescence characteristics. Below follows a short description of the method. However, more details and complete formulation can be found in Joiner et al. (2013, 2016); Sanders et al. (2016); Kooreman (2015); Leth (2014)

The idea behind the GOME-2 retrieval is that the observed TOA reflectance can be explained by atmospheric effects, surface reflectance and chlorophyll fluorescence. It uses principle component analysis (PCA) to account for the various atmospheric effects caused by aerosols and trace gasses. The contribution of surface reflectance to the TOA spectrum is approximated by a low-order polynomial. Fluorescence is assumed to have a Gaussian shape, with $\mu = 737$ nm and $\sigma = 33.9$ nm. Note that only far-red fluorescence is accounted for, though a recent study by Joiner et al. (2016) tried to retrieve red fluorescence using a similar approach. A linear combination of the various components (surface, atmosphere and fluorescence) is then minimized using the Levenberg-Marquardt algorithm (Sanders et al., 2016).

The various principle component are found by selecting so-called reference pixels, for which the TOA spectrum is assumed to be free of any contribution of vegetation (i.e. no fluorescence). Surface effects are removed from the TOA spectrum by fitting a low-order polynomial through wavelengths not severely affected by atmospheric effects (atmospheric windows). The remainder of the signals is aggregated for a large number of samples and subsequently transformed into principle components (PCs). These components can be interpreted as spectral features that explain the variance in the spectra devoid of surface effects (i.e. composed of only atmospheric effects). In other words, the first PC is similar to the mean, as this spectral shape explains most of the variance (see Figure 1.3). The second PC has its most pronounced features located near the absorption regions of water vapour (715-740nm), suggesting that variations in atmospheric water vapour content have a large impact in the TOA reflectance. The third PC has most of its features located near the area of oxygen absorption (760-770 nm). As each principle component aims to maximize the explained variance of the remainder of the signal, a linear combination of the first couple of PCs will usually explain most of the signal. Therefore, a model based on PCs can theoretically approximate the atmospheric effects in a statistical way. Field observations done by Guanter et al. (2013) support this. However, it should be noted that such an approach is linear and thus can lead to errors when extrapolating non-linear (e.g. saturation) effects in the atmosphere. In addition, the PCA approach does not try to physically explain the TOA spectrum, but rather builds a statistical model based on a large number of TOA samples.

The statistical atmospheric model constructed over non-vegetated surfaces can be used to retrieve fluorescence, under the assumption that the atmosphere has similar properties over vegetated areas. In the retrieval this set of atmospheric PCs together with a low-order polynomial for the surface and the Gaussian-shaped fluorescence signal are scaled linearly and optimized (towards a minimal root mean squared error between observed and modelled TOA spectrum).

Both the retrieval by (Joiner et al., 2013, 2016) (hereafter NASA SIF) and Sanders et al. (2016) (hereafter SIFTER v1) follow this approach. However,

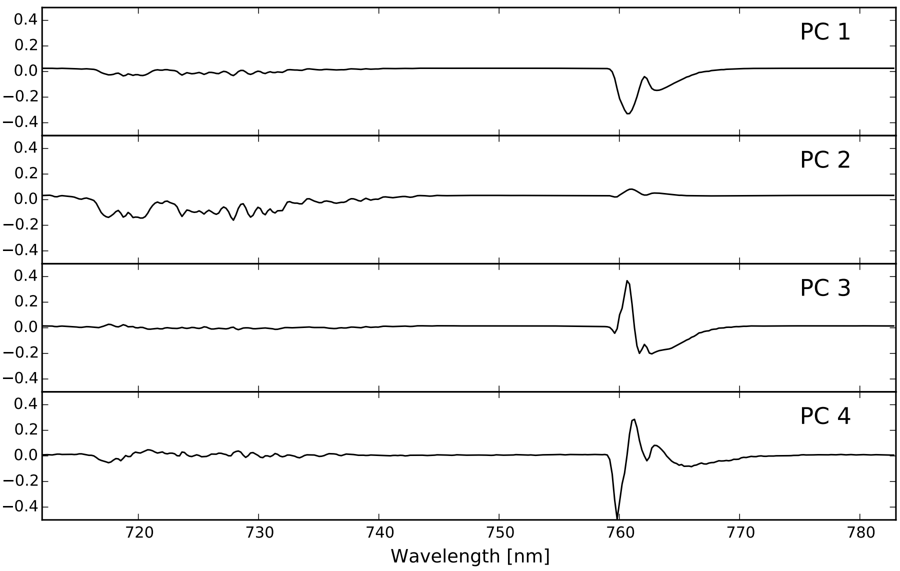


Figure 1.3: Principal components (PCs) as function of wavelength. The first PC is consistent with the mean atmospheric effect. The following PCs show the spectra with the highest explanatory fractions. For PC2, this corresponds to water vapour effects. For PC3 and 4 the major signals are found around the O2A band. The y-axis is dimensionless.

Table 1.1: Most important differences between the NASA fluorescence retrieval and the SIFTER v1 algorithm.

	NASA SIF v26	SIFTER v1
Fitting window	734-758 nm	712-783 nm
Number of PCs	12	35
Reference area	Cloudy ocean	Sahara (desert)
Reference timespan	Daily	Yearly
Surface polynomial	Third order	Fourth order

differences do occur in several important areas. A selection of these differences is presented in Table 1.1. The larger fitting window chosen by SIFTER v1 includes both the water vapour absorption and oxygen absorption (see Figure 1.4). This is in contrast with the NASA SIF product, which excludes the oxygen A band and most of the water vapour absorption features. As a consequence, fewer PCs are required to model atmospheric effects for the NASA product.

A second major difference between NASA and SIFTER is the choice of reference area. SIFTER opts to select an ensemble of non-vegetated pixels over the Saharan desert and perform a PCA over all suitable TOA spectra within the preceding twelve months. In contrast, the NASA product constructs a reference set based on spectra for that day, taken mainly over cloudy ocean areas. This has the advantage of capturing daily instrument variations within the PCs, but

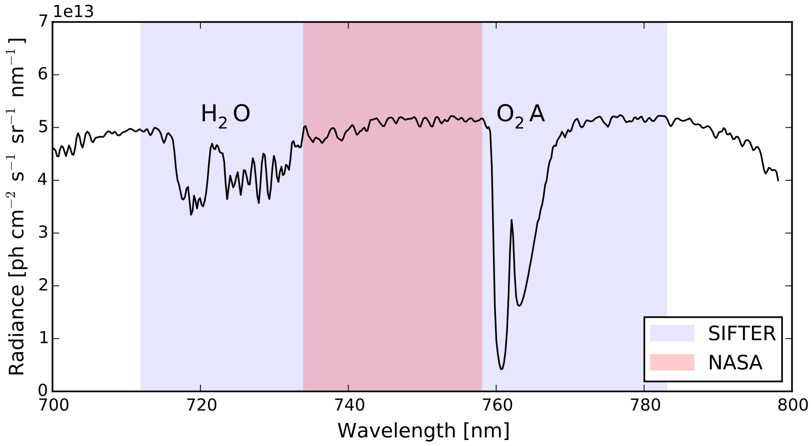


Figure 1.4: A typical GOME-2 observed TOA spectrum above the Saharan desert. The effect of water vapour absorption and oxygen are clearly visible. The blue shaded area shows the spectral fitting window of SIFTER (KNMI), whilst the red shade highlights the spectral fit window used by NASA.

could fail to capture atmospheric variability. However, for both products it can be debated whether the reference pixels contain similar atmospheric properties as the vegetated surfaces. If not, non-linear behaviour due to saturation and other phenomena can result in errors when retrieving fluorescence. For example, it can be debated whether air columns over the Sahara have similar properties (in terms of trace gas concentrations and aerosols) as air columns over the Amazon rain forest.

Currently it is not clear whether the retrieval parametrizations of NASA or SIFTER are better at retrieving chlorophyll fluorescence from GOME-2. A direct comparison between the two products show several differences, mainly in the tropics (Sanders et al., 2016). However, the lack of validation data on appropriate scales makes it difficult to state which set of parametrizations is performing better. Currently there are no observations (other than space-based) of fluorescence available at the scale of GOME-2 footprint (80-by-40 km). In addition, models that try to simulate top-of-canopy fluorescence are not yet mature enough to provide reliable estimates under real-world conditions (van der Tol et al., 2009; Verrelst et al., 2015). To overcome this problem, many studies make use of synthetic data derived from radiative transfer models (Joiner et al., 2011, 2013; Köhler et al., 2015; Frankenberg et al., 2014). The benefit of using a model to simulate TOA reflectance is that a known fluorescence signal can be added, with the downside that synthetic spectra can lack details present in real-world observations.

This study aims to provide insight in the choice of parametrization by assessing multiple combinations of fitting windows and PCs under different scenarios (e.g. changes in total column water vapour, viewing geometry, etc., see chapter 2). The results of these synthetic experiments are used to improve the algorithm to SIFTER v2 (chapter 3). The new fluorescence product is also compared to NASA SIF and other fluorescence products, looking at both spatial and temporal variation (chapter 4). At the end an overall conclusion and scientific outlook is presented.

Chapter 2

Synthetic Spectra

2.1 Introduction

As mentioned before, a major obstacle in validating fluorescence retrievals is a lack of independent reference measurements on appropriate spatial scales. Sanders et al. (2016) found that SIFTER and NASA SIF products compare well, but show no correlation in monthly mean fluorescence over the majority of tropical forests. Coincidentally, tropical forests are of crucial importance in regulating global atmospheric carbon dynamics due to their large aboveground biomass and carbon fluxes (Pan et al., 2011; Wang et al., 2013; van der Laan-Luijkx et al., 2015). Fluorescence can be an most useful tool to constrain photosynthesis rates in these regions, but this can only be done when the fluorescence in these regions is better known.

One way to gain insight in which of the retrieval settings perform better is to use model-generated TOA reflectances. Such an approach has the benefit of control over different important variables such as atmospheric water vapour content, viewing geometry and surface albedo. In addition, a top-of-canopy fluorescence signal can be added, thus allowing a direct comparison between true and retrieved fluorescence. This study uses the radiative transfer model Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval (DISAMAR) to determine retrieval performance under different settings for different scenarios in a so-called end-to-end test.

2.2 Methods

2.2.1 DISAMAR

DISAMAR is a radiative transfer model developed at KNMI by J. de Haan. DISAMAR can generate TOA reflectance spectra and has a built-in method for specifying surface fluorescence. It can be specified to convolve the reflectance signal with a pre-defined GOME-2 slit function.

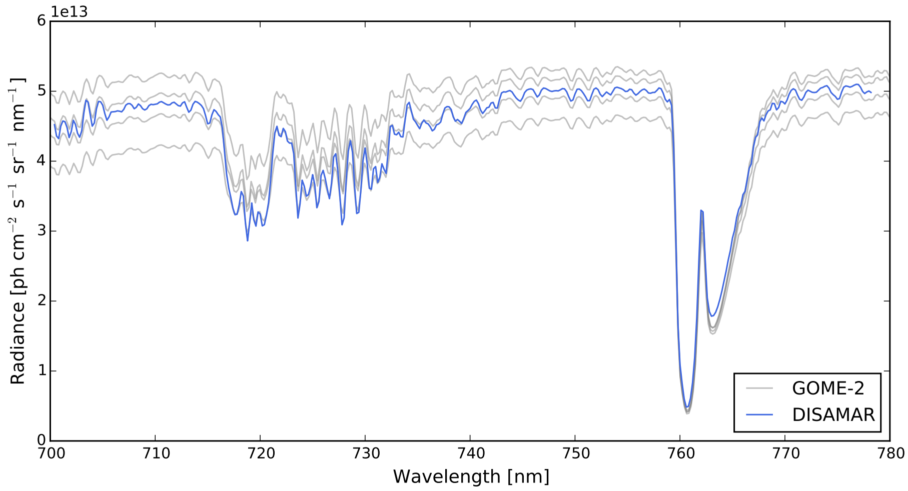


Figure 2.1: DISAMAR generated TOA spectrum (blue) and GOME-2 observed TOA spectra (grey) above the Saharan desert. DISAMAR is tuned to environmental conditions typically found over the Saharan desert.

2.2.2 Experimental set-up

DISAMAR was used to construct a reference set of spectra over non-fluorescent scenes located in the Saharan region. Some 2,000 spectra were generated with similar properties (albedo, surface pressure, viewing geometry) as found in the GOME-2-based reference set used by the SIFTER. The prescribed parameters were chosen randomly within the variation found in the SIFTER v1 reference set (see Table 2.1). A combination of observed solar zenith angle (SZA) and viewing zenith angle (VZA) was chosen randomly from a set of possible combinations to maintain realistic geometry. In contrast to the GOME-2 observations, which has effective cloud fractions up to 0.4 (see Koelemeijer et al. (2001)), DISAMAR generated spectra assume cloud-free conditions.

Five different datasets of each 1,000 spectra were generated to allow a quantitative assessment of the retrieval under different conditions. For each experiment fluorescence (F_s) was added by incorporating a Gaussian shape with $\mu = 737$ nm and $\sigma = 33.9$ nm. The peak value of F_s at 737 nm was set to five different strength (0.0, 0.5, 1.0, 2.0 and 4.0, all in mW m^{-2} sr^{-1} nm-1), such that each dataset contains 200 spectra per strength of F_s.

All five datasets share the majority of settings of the reference set. The first experiment (FLUOR) tries to quantify the uncertainty in retrieval of F_s under 'optimal' conditions, e.g. fluorescent scenes that are completely similar to the reference set but for an added source of F_s at the surface (see Table 2.1). The WATER experiment assesses the influence of water vapour concentrations on retrieval accuracy by replacing the Saharan water vapour columns by representative amounts over the Amazon rain forest (based on ERA-interim total column water vapour, Berrisford et al. (2011)). Water vapour is one of the main absorbing trace gasses within the spectral fitting window and non-linear effects

Table 2.1: Most important DISAMAR settings for the reference set and experiments. Solar Zenith Angle (SZA) and Viewing Zenith Angle (VZA) combinations are taken from GOME-2 observations. Total column water vapour is varied randomly between seasonal minimum and maximum values over the Sahara (taken from ERA interim, Berrisford et al. (2011)). Surface albedo is set to be representative of the Saharan desert. Signal-to-Noise Ratio (SNR) is taken to be representative of GOME-2.

Scenario	SZA [degrees]	VZA [degrees]	Column water [kg m-2]	Surface albedo [-]	SNR [-]
Reference	21.4 - 66.8	0.4 - 53.8	4.0 - 40.0	See Figure 2.2	1,000
FLUOR	ref	ref	ref	ref	ref
WATER	ref	ref	30.0 - 65.0	ref	ref
VEGFIX	ref	ref	ref	See Figure 2.2	ref
VEGDYN	ref	ref	ref	See Figure 2.2	ref
GEO	54.9 - 69.6	3.0 - 53.8	ref	ref	ref

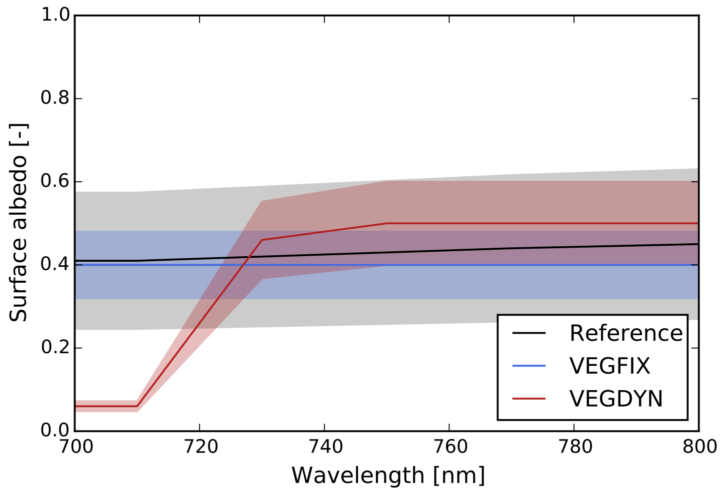


Figure 2.2: Surface albedo prescribed in DISAMAR for the default (reference) scenario (in black), fixed (VEGFIX, in blue) and with a red-edge (VEGDYN, in red). The shading denotes the minimum and maximum albedo within which the surface albedo can vary.

can not be captured effectively by PCs. Therefore, a correct representation of water vapour is crucial for modelling a TOA spectrum. VEGFIX and VEGDYN assess the sensitivity of the retrieval to different surface albedo's which are more consistent to that found over vegetated areas (with and without a red edge). These experiments should change the fitted surface polynomial and can be used to quantify uncertainty related to that part of the optimization. The final experiment, GEO, changes the solar and instrument angles to values found over a Russian boreal forest (latitude between 55-65 degrees) during an end-of-summer day (DOY: 253). Solar angles and therefore the distance travelled

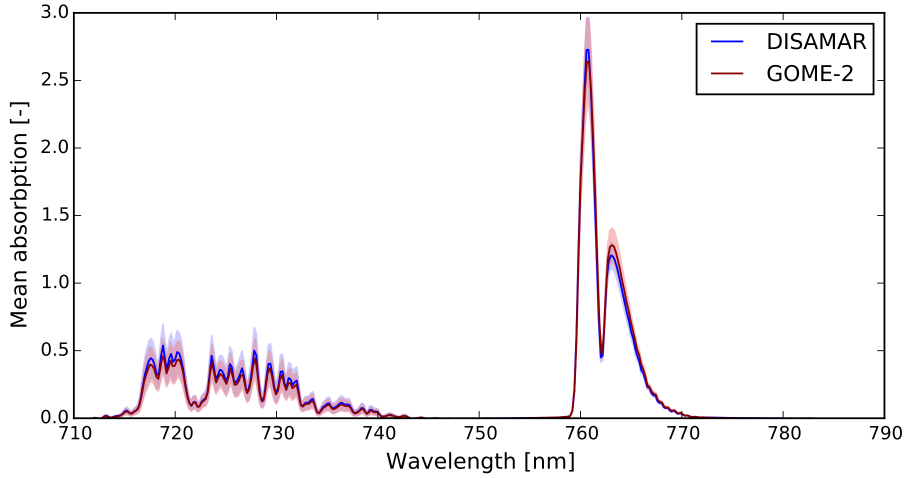


Figure 2.3: Mean wavelength-dependent absorption for real (GOME-2, red, n $\approx$ 30,000) and DISAMAR-generated (blue, n = 2,000) top-of-atmosphere spectra. The shading shows one standard deviation.

by photons through the atmosphere varies between areas near the equator and high-latitude regions. This could introduce non-linear effects that will decrease the goodness of fit (i.e. increase RMSE in fitting residuals). As the Saharan desert as reference sector has limited variability in SZA and VZA it is necessary to assess the impact of viewing geometry.

Subsequently, the generated spectra for each experiment are retrieved under different retrieval settings. Four different spectral fitting windows are evaluated, spanning the entire range (712-783 nm, SIFTER settings), excluding the water vapour and oxygen absorption bands (734-758 nm, NASA settings), or a combination (712-758 nm, H_2O window; 734-783 nm, O_2A window). In addition to a differences in spectral window, all four combinations are retrieved for three different counts of PCs, yielding a total of twelve combinations. After retrieval several statistical parameters are calculated for each scenario.

2.3 Results

2.3.1 Reference set

Reference TOA spectra generated by DISAMAR under Saharan conditions are very similar to observed GOME-2 TOA spectra (see Figure 2.3). This suggests that DISAMAR is able to accurately include the various atmospheric and surface effects that make up the signal. It also enhances the validity of results found in a purely synthetic simulation as differences are small.

GOME-2 observations over non-vegetated Saharan areas are also similar to DISAMAR generated spectra when looking at the PCs (Figure 2.4). The first two PCs, which represent the mean signal and the relative influence of water

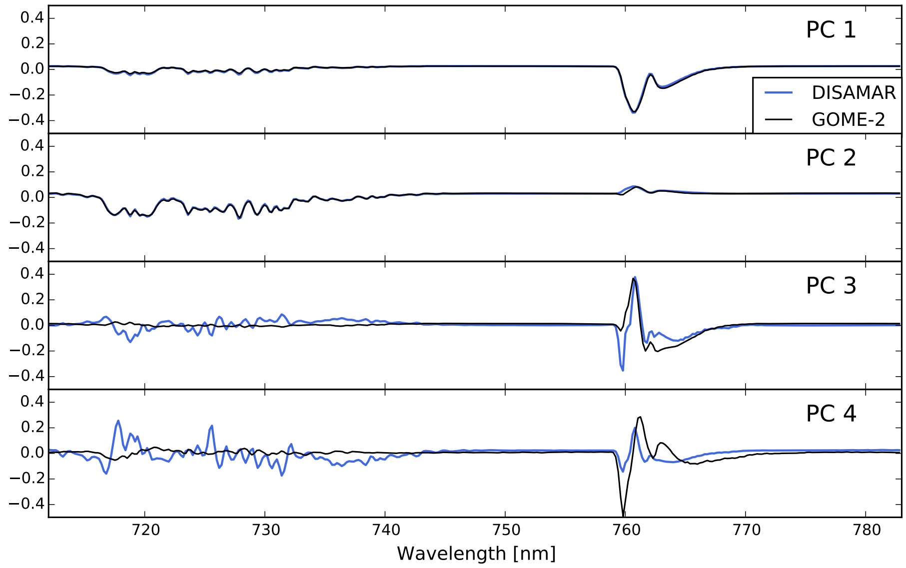


Figure 2.4: First four principle components (PCs) as calculated for one year of TOA spectra measurements above the Sahara from GOME-2 (n = ∼30000) and generated by DISAMAR (n = 2000).

vapour content, are extremely similar. The third and fourth PC show more variation between observed and modelled, which can be caused by only modelling cloud-free scenes, whilst the observations allow FRESCO-derived cloud coverages up to 0.4. In addition, the differences between GOME-2 and DISAMAR-based PCs can highlight the occurrence of other atmospheric compounds or small differences in spectral behaviour between DISAMAR and the real world (e.g. stray light, instrument calibration, etc.).

The fraction of explained variance per PC is dependent on the spectral window (see Table 2.2). Choosing a spectral window that includes oxygen absorption results in less explained variance per PC. In other words, more PCs are required to accurately model the behaviour of the atmosphere than when the O_2A band is excluded. In addition, the GOME-2 observed spectra have lower explained variance per PC than synthetic spectra. Again, this could be attributed to the 'clean' atmosphere (no clouds, no other atmospheric compounds, absence of instrument-specific features) in DISAMAR.

2.3.2 Experiment 1: Saharan fluorescence

Retrieving fluorescence under similar conditions as the reference set, but with the addition of a fluorescent signal at the surface, highlights that a small fitting window (NASA) with a low number of PCs (8) provides the best fit (Table 2.3). However, this is true only after faulty pixels, i.e. pixels that fail to reproduce the observed spectrum, are removed. Analysis indicates that whether a spectrum is successfully reproduced is not correlated with any of the variables (e.g. solar zenith angle, surface albedo, etc.). A common way to quantify goodness

Table 2.2: Fraction of explained variance per Principle Component (PC). In general, PCs derived from DISAMAR TOA spectra are better explained in the first few principle components. Including both oxygen and water absorption features (KNMI window) greatly reduces the explained variance in the first principle component. This suggests that a wider spectral fitting window required more PCs to accurately capture atmosphere effects.

	H_2O	NASA	KNMI	O_2A	H_2O	NASA	KNMI	O_2A
	Real (GOME-2)				Synthetic (DISAMAR)			
PC 1	0.9963	0.9968	0.5562	0.9159	0.9995	0.9995	0.7232	0.9201
PC 2	0.0028	0.0017	0.4263	0.0497	0.0005	0.0003	0.2762	0.0789
PC 3	0.0005	0.0004	0.0139	0.0291	0.0000	0.0000	0.0003	0.0005
PC 3	0.0001	0.0001	0.0024	0.0045	0.0000	0.0000	0.0001	0.0004

of fit is to look at the root mean square differences (RMSD) in the fit residuals. Fit residuals are retrieved by subtracting the modelled TOA spectrum from the observed. A high RMSD is considered an indicator for a failed fit. However, using RMSD as a filter is not sensitive enough to remove all unrealistic retrievals. Figure 2.5 shows the fit residuals for two different retrievals with an a priori fluorescence level of 4.0 mW m^{-2} sr^{-1} nm^{-1}. Whilst RMSD is similar, the retrieved fluorescence is not. However, the wavelength-dependent fitting residuals do contain information on the goodness of fit (Figure 2.5). If the model explains all the aspects of a spectrum the residuals will look like noise (i.e. an autocorrelation of 0). When the model is not able to resolve all components of the TOA reflectance the fitting residuals will contain structure, which can be expressed by autocorrelation.

In general, there is a strong correlation between the autocorrelation of the fitting residuals and the deviation of retrieved fluorescence from true fluorescence. As an example, Figure 2.6 shows increased deviations when the autocorrelation is larger than 0.2. Such relations are found for all combinations of fitting window are number of PCs, and also apply to retrieved fluorescence from GOME-2.

The statistical summary of this experiment (Table 2.3) suggests that a smaller fitting window is more likely to retrieve a fluorescence strength similar to the input fluorescence. This can be deduced from the relatively low root mean squared error (RMSE), a small bias and a linear fit that suggests similar values (i.e. slope of 1, intercept of 0, high correlation).
Inclusion of the O_2A band tends to lead to an underestimation of retrieved fluorescence (slope <1). This is visualized in Figure 2.7. Root mean squared error (RMSE) increases when the number of used PCs is increased (from eight) for the NASA and H_2O window. This is in agreement with the explained variance per PC, which shows that the atmospheric effects are captured by the first couple of PCs. Thus, additional PCs will likely lead to overfitting of the atmospheric effects. The fitting windows that include the O_2A band benefit from more than eight PCs, though RMSE increases when using 35 (instead of 20) PCs. The NASA window also results in the least amount of faulty pixels (i.e. pixels with an autocorrelation greater than 0.2), which indicates that more structure re-

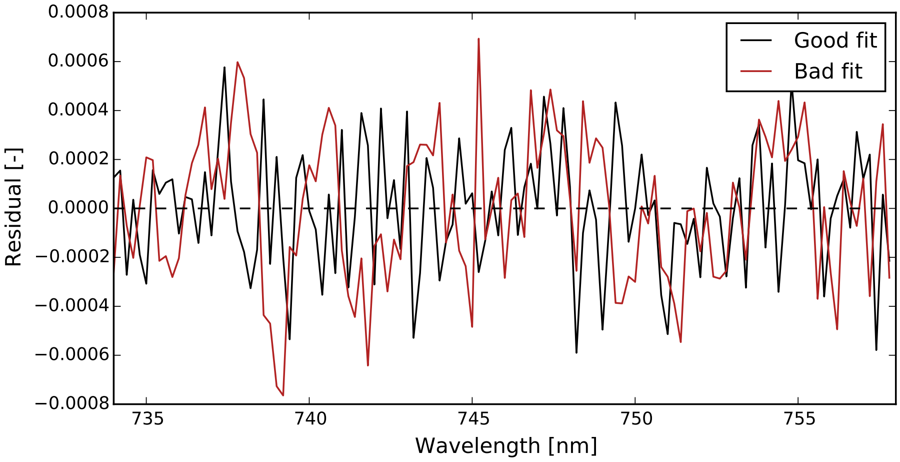
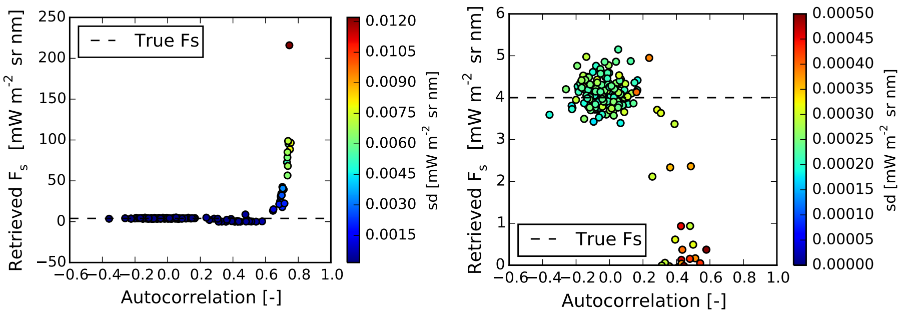


Figure 2.5: Wavelength-dependent fit residuals for a successful and unsuccessful retrieval. Input fluorescence as specified in DISAMAR is 4.0 mW m^{-2} sr^{-1} nm^{-1}. The successful fit (good fit, black) has a retrieved fluorescence strength of 3.9 mW m^{-2} sr^{-1} nm^{-1} and a RMSD of 0.24. The unsuccessful (bad fit, red) has a similar RMSD (0.29), but a different fluorescence strength of 0.9 mW m^{-2} sr^{-1} nm^{-1}. The autocorrelation for the successful fit is 0.04, whilst autocorrelation for the unsuccessful fit is 0.48.



(a) **(b)**

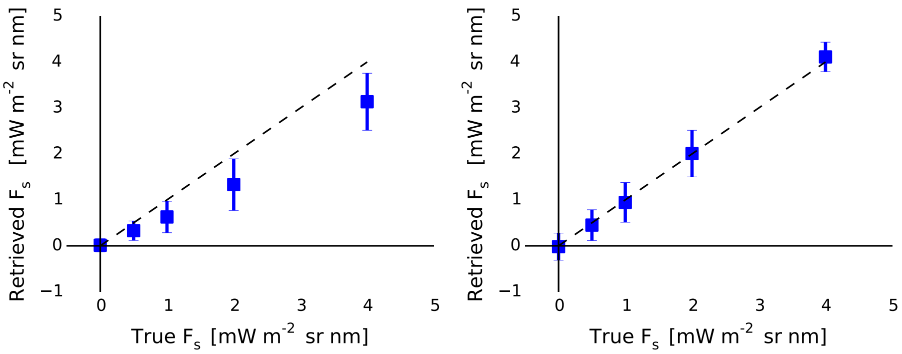
Figure 2.6: Scatterplots of autocorrelation against retrieved fluorescence. Input fluorescence is set at 4.0 mW m^{-2} sr^{-1} nm^{-1}. Colours show the standard deviation within the fit residuals for each retrieval. Figure b is zoomed in on the y-axis, but similar to a.

mains in the residuals when the water vapour and oxygen absorption bands are included.

In summary, the narrow spectral window between 734-758 nm in combination with eight principle components has the best agreement (lowest RMSE) between true and retrieved fluorescence. The other statistical parameters show similar results, with low bias and, a slope close to 1, and a relative low number of

Table 2.3: Statistical summary of Experiment 1 (FLUOR). Basis statistics are provided for all twelve combinations of fit window and number of principle components. A perfect retrieval would have a RMSE, bias, standard deviation and intercept of 0, whilst correlation (r) and slope would be 1. The absolute number of filtered pixels (autocorrelation >0.2) are given (n = 1,000 for unfiltered retrievals).

ID	λ_1 [nm]	λ_2 [nm]	# PC	RMSE [*]	r	bias [*]	σ [*]	slope	intercept [*]	faulty
	712	758	8	0.46	1	-0.2	0.41	0.94	-0.12	236
H_2O	712	758	20	0.5	1	-0.23	0.44	0.91	-0.11	201
	712	758	35	0.53	1	-0.26	0.47	0.89	-0.11	182
	734	758	8	0.39	1	0	0.39	1.04	-0.06	165
NASA	734	758	20	0.43	1	0.04	0.42	1.04	-0.02	162
	734	758	35	0.5	1	-0.03	0.5	1.01	-0.05	146
	712	783	8	0.71	0.99	-0.47	0.53	0.73	-0.16	347
KNMI	712	783	20	0.59	0.99	-0.37	0.45	0.82	-0.15	306
	712	783	35	0.62	1	-0.37	0.5	0.78	-0.09	288
	734	783	8	0.75	0.99	-0.49	0.57	0.69	-0.11	419
O_2A	734	783	20	6.41	0.96	-0.48	6.4	0.98	-0.43	379
	734	783	35	0.49	1	-0.25	0.43	0.83	-0.01	354



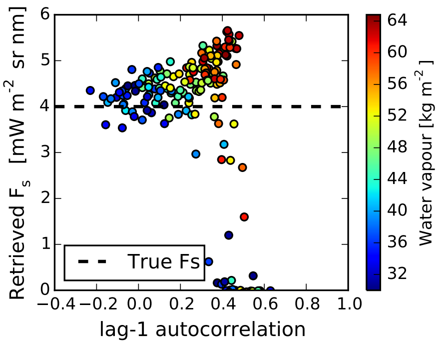
(a) (b)

Figure 2.7: True against retrieved fluorescence for (a) 712-783 nm fit window with 35 PCs and (b) 734-758 nm fit window with 8 PCs. Each square represents the mean and standard deviation of 200 retrievals. The dashed black line represent the 1:1 line.

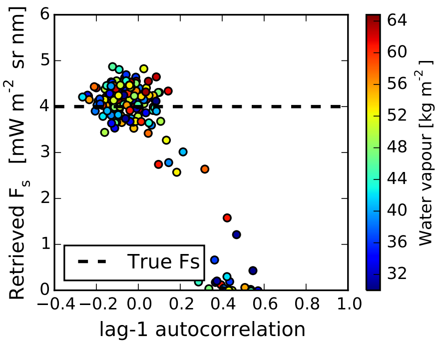
rejected (faulty) retrievals.

2.3.3 Experiment 2: The role of water vapour

Increasing water vapour content in the atmospheric column compared to the reference set leads to a strong increase in faulty retrievals, compared to the previous experiment (Table 5.1 in the Appendix). This suggests that the retrieval



(a)



(b)

Figure 2.8: True against retrieved fluorescence for a fit window between 734-758 nm and 8 PCs. Figure (a) has total column water vapour between 4.0 and 40 $kg\ m^{-2}$ in the reference set. Figure (b) is similar, but with 30-65 $kg\ m^{-2}$ in the reference (similar to experimental conditions). Colours indicate the total column water vapour for each individual retrieval. The dashed black line shows the true fluorescence.

is worse in explaining the spectrum and more structure remains in the residuals. Figure 2.8 shows that there is indeed a correlation between autocorrelation in the fitting residuals and exceeding total column water vapour content exceeding that of the reference set (40.0 $kg\ m^{-2}$). Non-linear behaviour created by water vapour content larger than present in the reference is the main suspect, as all other variables are unchanged (compared to the previous experiment). In addition, Figure 2.8 also shows that when a reference set with total column water vapour similar to tropical forest (and thus this experiment) is used, the mentioned correlation disappears.

The number of faulty retrievals increases most for the three configurations that include the water vapour absorption features around 730 nm due to more unexplained structure in the fit residuals. However as some water vapour features are still present at 734 nm the NASA / O_2A spectral windows are still affected. In general, the statistics suggest that a small fitting window with a low number of PCs (NASA 8) has the best match with prescribed fluorescence. Inclusion of the O_2A band still results in underestimation of fluorescence, which is to be expected as total column oxygen is unchanged.

2.3.4 Experiment 3: Vegetation - flat albedo

Statistics of the retrieved fluorescence for a spectrally flat surface albedo are very comparable to experiment 1 (sloping albedo, see Figure 2.2), except for the KNMI fitting window. This is the result of nearly all pixels being filtered out as their autocorrelation of the fitted residuals exceeds 0.2. Recreating the spectral shape of the surface as fitted by the retrieval shows that some curvature exists, where the albedo is expected to remain constant throughout the spectrum. Especially the large fitting window shows significant non-linearity near the edges of the window. This results in an incorrect representation of the surface and therefore an inability of the retrieval to resolve all structure in the retrieval.

The other, smaller fitting windows are not affected by the autocorrelation criteria and perform similar to experiment 1. Again, the small fitting window of NASA, combined with a limited number of PCs, gives the best results in terms of RMSE and other statistical components (Table 5.2 in the Appendix).

2.3.5 Experiment 4: Vegetation - red edge

Inclusion of a strong red edge between 700-730 nm increases the number faulty retrievals for all fitting windows, but mainly for the KNMI and H_2O windows. This is expected, as the NASA and O_2A fit window begin at 734 nm and are therefore not confronted with the red-edge as prescribed in this experiment. Large windows suffer from incorrect fits of the surface, which disable the retrieval from explaining all structure in the signal. The small fitting window of NASA seems to be able to keep the rejected pixels to a minimum, whilst retaining good skill in estimating fluorescence.

Loosening the autocorrelation filter will decrease the number of rejected retrievals, but also greatly reduces the retrieval accuracy and precision (results not shown). It should be noted that the prescribed surface albedo is not as smooth as seen in nature (see Figure 2.2).

2.3.6 Experiment 5: Geometry

Changing the viewing geometry to reflect the angles found during a late-summer day over boreal Russia does not severely affect the retrieval results (Table 5.4 in the Appendix). A slight increase in rejected pixels (autocorrelation > 0.2) is found, whilst the statistics are very compare between this and the benchmark of experiment 1. These results suggest that a reference set based on the limited combination of viewing geometries found over the Saharan desert can also be used for retrieving fluorescence over high-latitude regions.

Table 5.4 also shows a large increase in RMSE for the KNMI fitting window with 20 and 35 PCs, and the O_2A fitting window with 20 PCs. This is due to one erroneous pixel per retrieval which is not filtered out by the autocorrelation. These pixels can have fluorescence strengths over 200 mW m^{-2} sr^{-1} nm^{-1} (where a maximum strength of 4.0 mW m^{-2} sr^{-1} nm^{-1} is used as input), resulting in large increases in RMSE and standard deviation.

2.4 Discussion

DISAMAR-generated spectra are very comparable to GOME-2 observed TOA reflectances, suggesting that the results found in this study are representative for the GOME-2 based retrieval. It should be noted here that no clouds were added in DISAMAR. The lack of clouds can reduce the variance in the signal, leading to fewer PCs required to capture atmospheric effects as no scattering and absorption by clouds is present. Joiner et al. (2013) also used synthetic spectra to assess the retrieval errors, based on simulations with another radiative transfer modelling. Comparing the explained variance per PC (see Table 2.2) with the results of Joiner et al. (2013) shows that DISAMAR generated spectra are comparable in their level of variance.

Comparing the retrieval results of this study with that presented in Joiner et al. (2013) reveals that the statistics are very comparable. The RMSE found for the NASA window with 8 PCs is around 0.4 mW m^{-2} sr^{-1} nm^{-1} throughout all experiments, where Joiner et al. found a RMSE of 0.33 mW m^{-2} sr^{-1} nm^{-1} for their best fit. However, a direct comparison is unfortunately not possible as Joiner et al. used a spectral window of 712-747 nm, 25 PCs and a SNR of 2000. The other statistics (r, slope, intercept) are of similar order of magnitude though the circumstances are somewhat different.

The results of the different experiments show that a small fitting window (NASA) with a low number of PCs (8) consistently has the best statistics. However, in order to attain this it is crucial to filter out retrieved pixels that retain persistent structure in their residuals. A filter based on autocorrelation of the fit residuals is able to flag the vast majority of retrievals that contain unrealistically high fluorescence values. Prior analysis suggests that this approach is also suitable for GOME-2 observations and can both improve retrieval accuracy as well as reduce noise (more in the next chapter). A more in-depth study could focus on identifying repeating patterns in fit residuals to determine whether a consistent (and therefore correctable) error is present.

Viewing geometry and the spectral shape of the surface albedo have no strong impact on the retrieval results for the small NASA fitting window. However, this study does show that enhanced total column water vapour, as representative for tropical regions, does impact the reproducibility of the retrieval in a negative way. The possible non-linear effects due to saturation can induce atmospheric effects that can not be captured by the PCs. This leads to an increased number of rejected pixels (based on autocorrelation). The results clearly indicate that inclusion of reference areas with similar total column water vapour as vegetated areas improves the retrieval. The next chapter will assess how retrieved fluorescence changes when clouded ocean scenes, rather than the Saharan desert, are used as reference.

Small changes to the NASA fitting window do not improve the retrieval, suggesting that the window between 734-758 nm is appropriately chosen. In addition, no significant improvement was found when changing the number of PCs between 4-12. The experiment done with DISAMAR-generated spectra therefore suggests that four PCs are enough to capture atmospheric effects. However, GOME-2 derived PCs show that less variance is explained per PC, indicating that the real atmosphere as well as instrument effects result in more variance than the synthetic atmosphere. Therefore, 8 PCs seem a good compromise to avoid under- or over-fitting. However, Köhler et al. (2015) implemented a statistical method to derive the appropriate number of PCs per retrieved pixel. Such an approach can benefit this retrieval as well, but falls outside the scope of this study.

2.5 Conclusion

A small fitting window (734-758 nm) with a low number of PCs to capture atmospheric effects consistently produces the best results in terms of retrieval statistics. This confirms the choice of made by Joiner et al. (2016). In addition, it is essential that the reference areas contain similar total column water vapour content as vegetated areas over which fluorescence is derived. Even under idealized conditions, the retrieval is prone to produce erroneous fits, which can be a factor 100 larger than expected fluorescence. These faulty retrievals share that they contain more structure in the fit residuals than successful retrievals, and thus can be removed by filtering for autocorrelation. This novel method to assess the quality of a retrieval shows great promise and could potentially be applied outside of fluorescence retrievals.

The next chapter will continue by implementing the suggested changes and retrieve fluorescence from actual GOME-2 observations.

Chapter 3

Towards a new KNMI Retrieval

3.1 Introduction

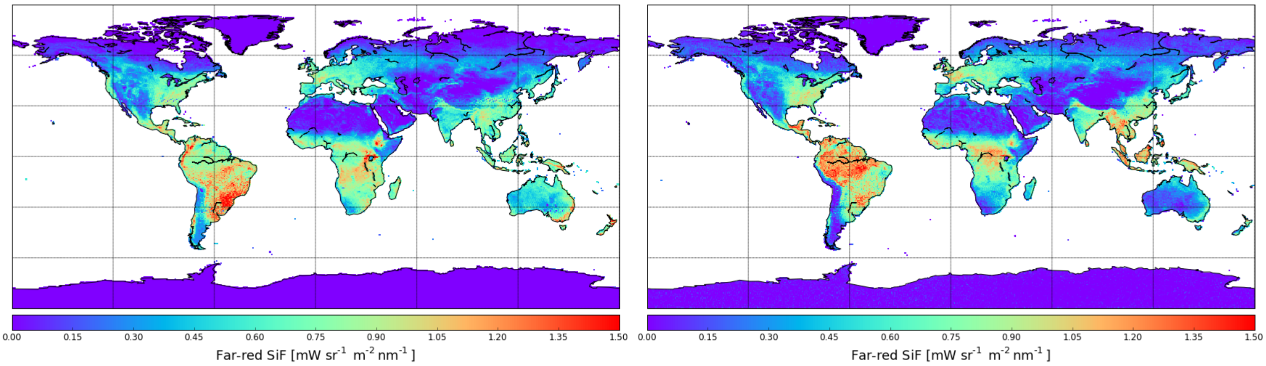
The previous chapter highlights a specific set of parameter-settings to use in the KNMI fluorescence retrieval. Specifically, it shows that a spectral window of 734-758 nm with eight principle components is preferred. Furthermore the experiments based on synthetic spectra show that the variation in atmospheric water vapour should be similar (when possible) between the reference locations and the actual location of retrieval. However, whilst the synthetic experiments aim to approximate real GOME-2 observations, they fail to capture phenomena such as the latitude-dependent bias described by Köhler et al. (2015).

The KNMI fluorescence product described by Sanders et al. shows little correlation with the NASA fluorescence product described by Joiner et al. (2013, 2016) in tropical regions (Sanders et al., 2016) (see also Figure 3.1). The synthetic experiments performed with DISAMAR suggest that the broad spectral fitting window used in KNMI v1 should be narrowed to the window used by NASA (in version 26) with eight principle components.

This chapter aims to investigate fluorescence retrievals under the new set of parameter settings. A correction for the latitude-dependent bias will be described and applied. Application of these changes yields a new fluorescence retrieval: the Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval (SIFTER). The final retrieval results of based on the new settings (SIFTER v2) will be compared with the original retrieval described by Sanders et al. (2016) (from here-on referred to as SIFTER v1).

3.2 Methods

The new retrieval (SIFTER v2), based on the small spectral window (734-758) and low number of PCs (8), is applied to retrieve fluorescence for the year 2011. This specific year is chosen as it is not affected by significant instrument events



(a) **(b)**

Figure 3.1: Annual mean fluorescence (2011) as retrieved by SIFTER v1 (a) and NASA v26 (b) on a 0.5-by-0.5 degree grid.

or global climate extremes (such as the El Nino in 2010) (Kooreman, 2015). GOME-2a level1b data is processed with version 5.3 of the processor.

A latitudinal bias exists in fluorescence retrievals (Köhler et al., 2015). Joiner et al. (2016) uses a polynomial fitted on latitude and irradiance levels to account for the bias. This study uses a different approach based on a reference sector. Over the ocean, far-red fluorescence can be assumed to be zero. Therefore, mean fluorescence levels over the ocean can indicate the offset for a specific latitude. Ocean-pixels are selected by applying the 0.5 degree land ocean mask of Jet Propulsion Lab (2013). Fluorescence estimated over all ocean pixels per 0.5 degree latitudinal band is averaged to derive the offset. Bias-correction is applied on a daily basis or over monthly aggregated fluorescence maps.

The synthetic experiments suggest that total column water vapour over reference areas should be as similar as possible to the levels encountered over vegetated surfaces. The current reference area over the Saharan desert could lead to errors as total column water vapour does not reach levels found over the tropics (see Figure 3.2). Atmospheric water vapour over tropical oceans does reach the desired levels. However, the low albedo of ocean diminishes the signal strength. Therefore, Joiner et al. (2013) opted for clouded ocean scenes, as clouds reflect and scatter more light back to the sensor. A similar approach is used in this study, where clouded ocean scenes over (sub-)tropical Atlantic with an effective cloud fraction larger than 0.4 (see Koelemeijer et al. 2001) are selected for reference. The results of a retrieval based on an ocean reference set are compared to the default retrieval that uses the Sahara as reference.

Similar to the synthetic retrievals the autocorrelation filter is assessed and applied to GOME-2A observations. The structure in residuals of the fit for each individual retrieval is quantified using autocorrelation. Retrievals that have an autocorrelation >0.2 are removed.

The other settings of the retrieval are similar to Kooreman (2015) and Sanders

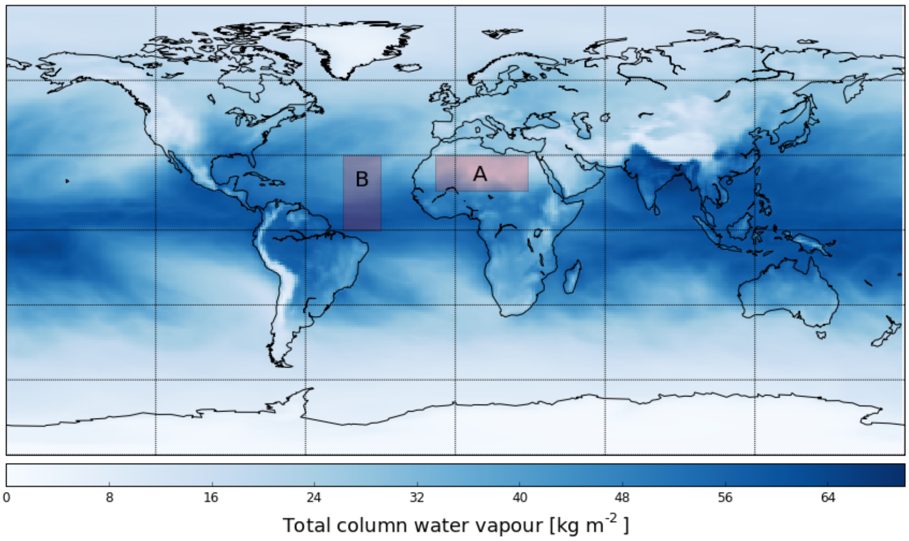


Figure 3.2: Maximum total column water vapour derived from ERA interim for the years 2010-2011. Red areas present the Saharan reference sector (A) and the Atlantic ocean reference sector (B) from which clouded ocean scenes are used.

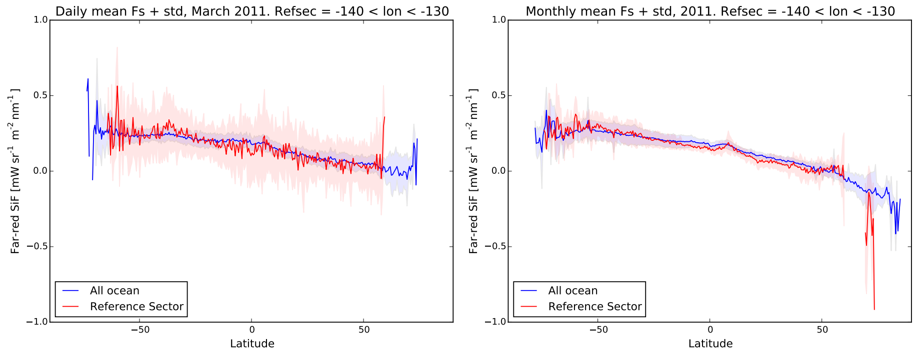
et al. (2016). Pixels with an effective cloud fraction >0.4 are not retrieved. The surface is fitted using a fourth-order polynomial (in contrast to a third-order polynomial used in NASA v26). Individual retrievals are aggregated and averaged on a 0.5-by-0.5 degree grid, both for individual days as well as on a monthly basis. PCs are constructed based on a full year of reference data (in contrast to the daily reference sets employed in the NASA product). Under normal operation, the year of reference data spans the day of the retrieval plus the preceding 364 days. Instrument functioning can warrant a change in reference period (see Kooreman 2015).

3.3 Results

3.3.1 Latitudinal bias

The latitudinal bias is present throughout all 2011 (see Figure 3.3). The magnitude of the bias is on the order of tenths of mW m^{-2} sr^{-1} nm^{-1} at 737 nm. The relative impact of the bias on the retrieved fluorescence signal can be over 100% for high-latitude regions. A similar bias (both in direction and magnitude) can be found for the SIFTER v1 product. In addition, the strength and direction of the bias are similar the zero-level offset for the NASA product Joiner et al. (2016).

To correct for this a reference sector approach is used, where oceans and seas are assumed to have no far-red fluorescence. However, such an approach assumes that there is no longitudinal bias. In addition, it assumes that latitudinal difference in land-sea distribution are not present. Using a latitudinal band with



(a) **(b)**

Figure 3.3: Latitudinal-dependent daily mean bias (a) and monthly mean bias (b) over the entire ocean (blue) or a latitudinal cross-section between -140 and -130 longitude (Pacific ocean). Shading presents one standard deviation of the day-to-day or month-to-month variation. Daily comparison is done for March 2011, and monthly comparison is done for all months in 2011.

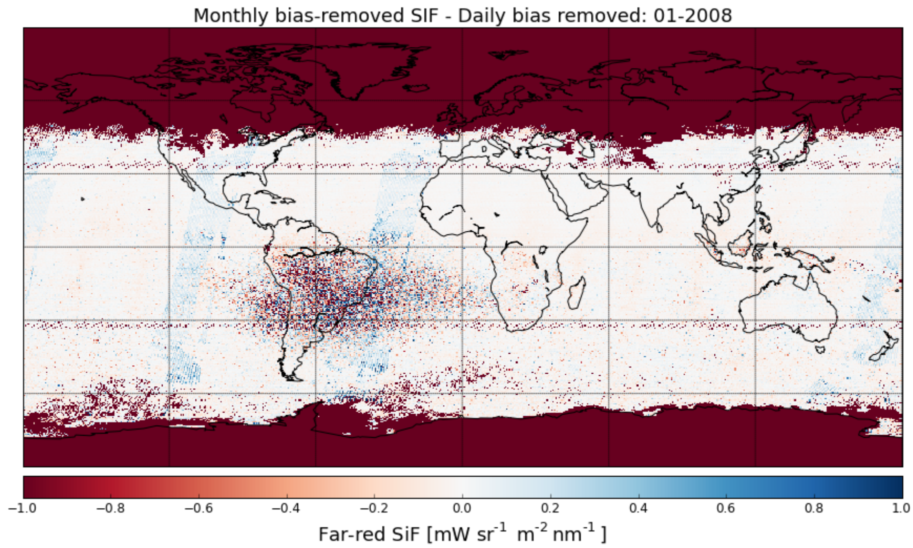
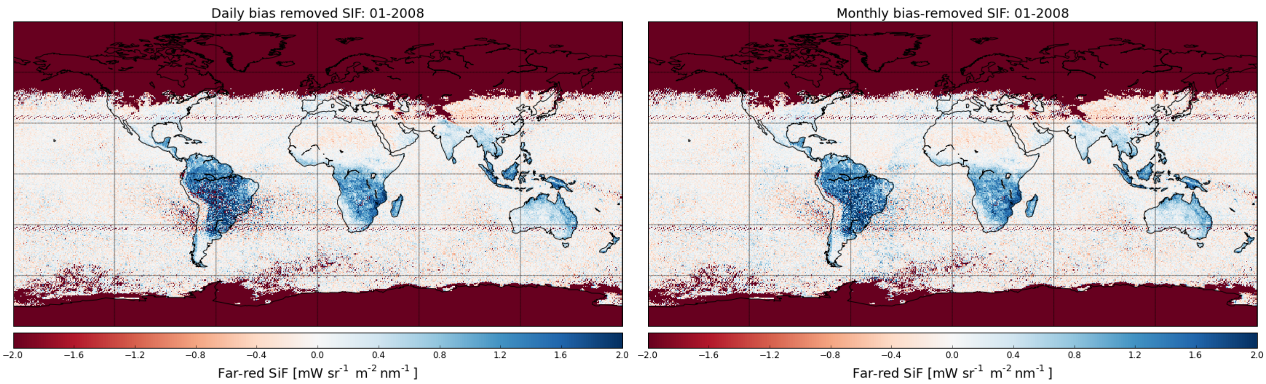


Figure 3.4: Difference between far-red fluorescence for January 2008 after application of a daily bias correction or monthly bias correction.

a longitudinal width of 10 degrees over the Pacific ocean (Figure 3.3 and the Atlantic (not shown) show very similar bias as when taking the entire ocean as reference. This indicates that the bias is not (severely) affected by land-ocean distribution or longitudinal effects.

The temporal variation in the bias is of similar magnitude across daily to yearly timescales. In other words, the day-to-day variation shows comparable magnitudes of deviation as month-to-month averages, or year-to-year averages. In general, a bias-correction applied to monthly aggregates of fluorescence is very similar as when a bias correction is applied daily (see Figure 3.4). Besides the expected noise around the Southern Atlantic Anomaly Figure 3.4 also shows 'striping'. This effect most likely originates from an instrument restart in 2008. Daily bias correction is able to resolve the instrument-induced errors, whilst monthly-averaged bias correction is not (Figure 3.5).



(a) **(b)**

Figure 3.5: Monthly mean fluorescence for January 2008 after application of a daily bias correction (a) or monthly bias correction (b). Striping effects due to a different retrieved fluorescence signal for a specific orbit occur in (b).

3.3.2 Reference area

Taking a reference area above the (sub-) tropical Atlantic ocean with a effective cloud fraction >0.4 does not lead to major changes in the first four PCs (Figure 3.6). This implies that within the spectral fit window atmospheric conditions between sub-tropical Atlantic ocean and the Saharan desert are similar. This might seem counter-intuitive, as Figure 3.2 clearly shows a higher maximum of total column water vapour content above the ocean reference sector. However, as only clouded scenes are used, the lower part of the atmosphere (below 850 hPa) is partly obscured from view. Water vapour present in the lower troposphere and planetary boundary layer may therefore be under-represented in the retrieved TOA reflectance.

A closer look at the PCs generated from TOA reflectance above the two reference sectors shows that not only the spectral shape, but also the explained variance per PC is comparable. A slightly higher amplitude is found in the second PC around 736 nm over the Saharan reference sector. This could be to caused by differences in atmospheric water vapour content. However, more research is required to investigate the cause and impact the difference. The current comparison removes all ocean pixels with effective cloud fractions <0.4 (based on the FRESCO product). Such a cloud fraction should return enough light to the sensor to effectively perform a sensitivity analysis. A higher cut-off results in less accepted measurements, which increase the noise in the PCs. A lower cut-off (0.2) does not lead to a significant change in the first four PCs (p

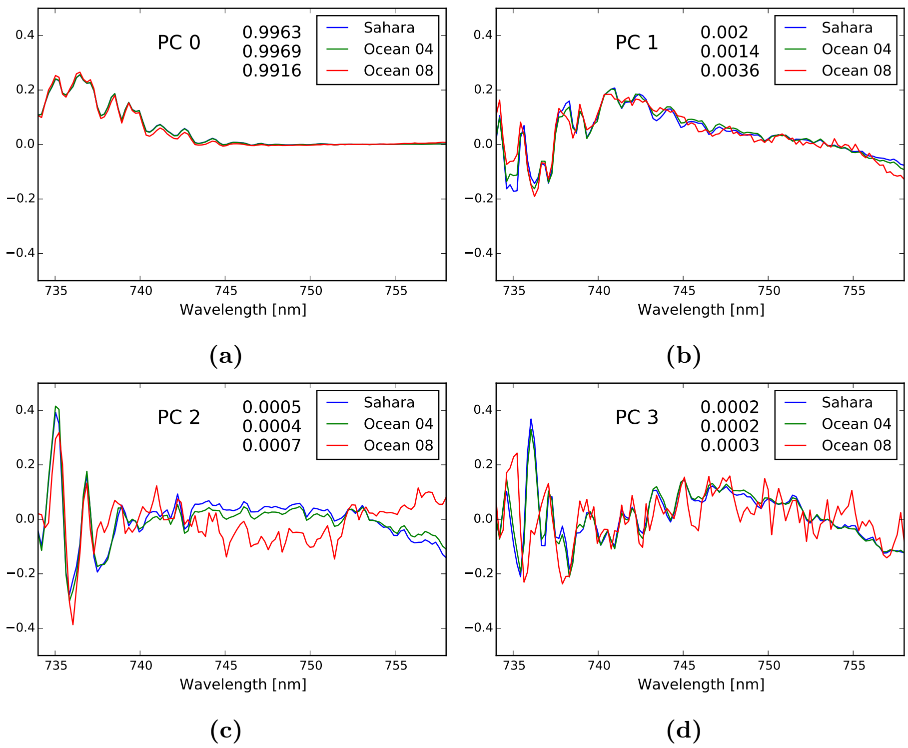


Figure 3.6: Spectral signature of the first four principle components derived over the Saharan reference area, the ocean reference sector with a minimal effective cloud cover fraction of 0.4 and with a minimal effective cloud cover fraction of 0.8. The numbers give the fraction of explained variance per principle component.

>0.05, n = 118 for each PC).

As expected from the PCs, the retrieved fluorescence from GOME-2A using both reference sets is highly correlated (r^2 of 0.99, see Figure 3.7). The choice of datasets introduces a small bias where fluorescence retrieved using the Atlantic ocean as reference is slightly lower than fluorescence retrieved using the Saharan desert as reference sector. Spatially, the difference in retrieved fluorescence is largest in the tropics and across dry areas (Figure 3.8). In general, using the Sahara as reference sector results in higher fluorescence over tropical areas compared to ocean-reference derived fluorescence. The reverse is true for dry areas, which tend to have lower retrieved fluorescence when the Sahara is used as reference sector. As everything except the PCs is similar between the two datasets the reason for the difference should reside in different atmospheric conditions. Unlike the first four PCs, which explain >99.9% of the variance, the last four PCs do diverge which could explain the differences. However, it is hard to find a physical explanation purely based on principle components.

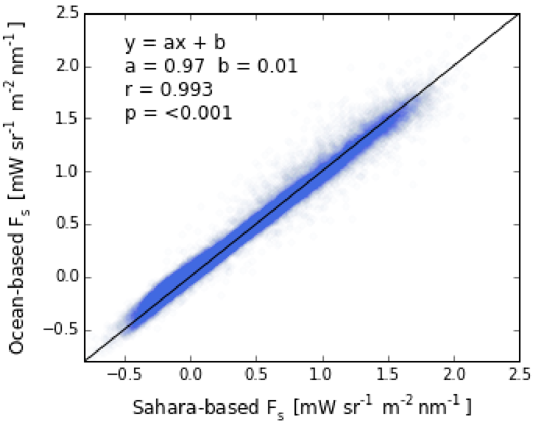


Figure 3.7: Scatterplot of 0.5-by-0.5 degree monthly mean fluorescence over land for 2011 based on a Saharan reference sector (x-axis) or ocean-based reference sector (y-axis). n = 88489.

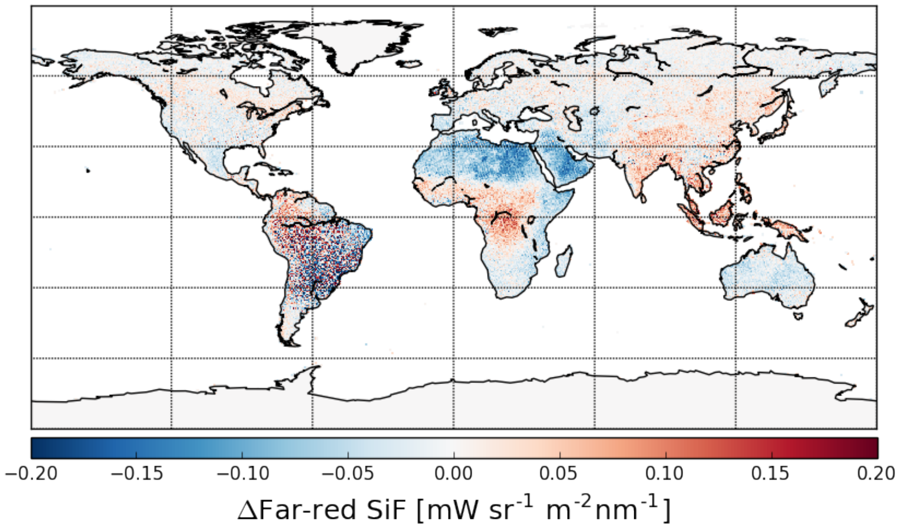


Figure 3.8: Annual mean 0.5-by-0.5 degree difference between a retrieval based on the Saharan reference sector or ocean reference sector. Negative (blue) values represent higher fluorescence in the ocean-based retrieval, whilst positive (red) values indicate higher fluorescence in the Saharan-based retrieval.

3.3.3 Comparison to SIFTER v1

The annual mean terrestrial fluorescence signal for 2011 as retrieved using the new algorithm (including bias correction and using the Sahara as reference sector) is presented in Figure 3.9. Tropical areas show the highest levels of fluorescence, whilst desert, high latitude and high altitude regions produce near-zero levels. Especially tropical regions show large differences between SIFTER v1 and v2 (Figure 3.10). SIFTER v2 estimates annual mean fluorescence values

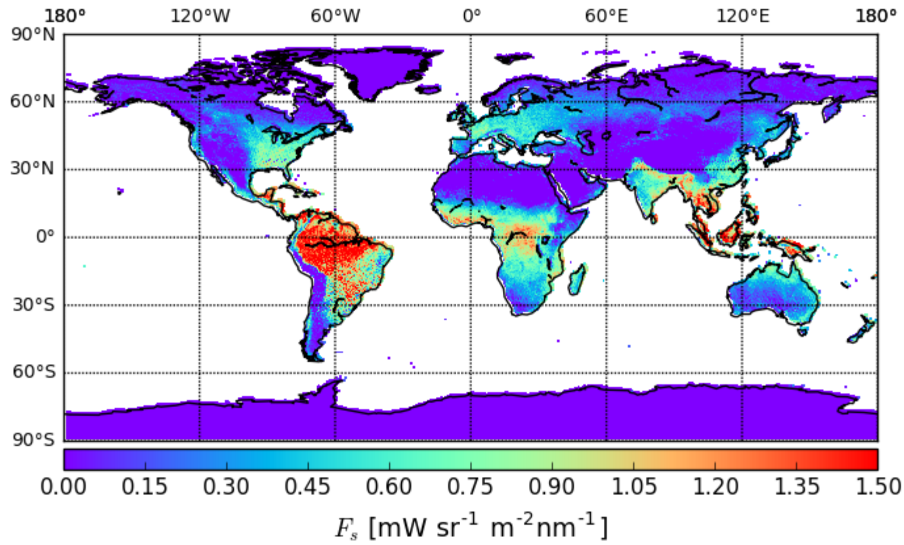


Figure 3.9: Annual mean fluorescence for 2011 as retrieved by SIFTER v2.

over tropical regions over 100% higher than v1. This is consistent for all tropical regions across the globe. Tropical regions aside, the estimated fluorescence of v2 is generally lower than v1. This is partly caused by the bias correction, which tends to reduce fluorescence levels south of 30 degree N (see also Figure 3.11). Large differences are also found in high altitude regions, specifically the Himalayas and the Andes. This is partly caused by negative fluorescence yields derived from v2, which will be addressed in more detail in the discussion section of this chapter.

The correlation between v1 and v2 for terrestrial global annual means on a 0.5 by 0.5 degree grid is 0.69 (r^2, n = 59077). In general, v1 produces higher values than v2 (slope of 0.92, intercept of -0.04, using y = ax+b and x = v1). The root mean squared error (RMSE) is 0.22 mW m^{-2} sr^{-1} nm^{-1}.

An additional change between SIFTER v1 and v2 is the inclusion of an autocorrelation filter. On average, the autocorrelation filter rejects $\sim$5% of the measurements. The majority of rejected pixels are located near the South-East of Southern America, overlapping with the Southern Atlantic Anomaly (Figure 3.12). For 2011, SIFTER v2 calculates autocorrelations larger than 0.2 for 4.7% (n = 24,361,289) of the pixels. Before the filter is applied, fluorescence retrieved over both land and ocean have μ = 0.59 mW m^{-2} sr^{-1} nm^{-1} and σ = 10.97 mW m^{-2} sr^{-1} nm^{-1}. Applying the filter changes μ to 0.11 mW m^{-2} sr^{-1} nm^{-1} and σ to 0.69 mW m^{-2} sr^{-1} nm^{-1}. The filtered values have a μ and σ of 10.25 and 49.47 (both in mW m^{-2} sr^{-1} nm^{-1}). The large reduction in standard deviation suggests that the new filter is able to successfully remove unrealistic high and low fluorescence values. Using autocorrelation as filter performs better than RMSE under synthetic experiments. Whether this is also the case for real-world observations is something that requires further

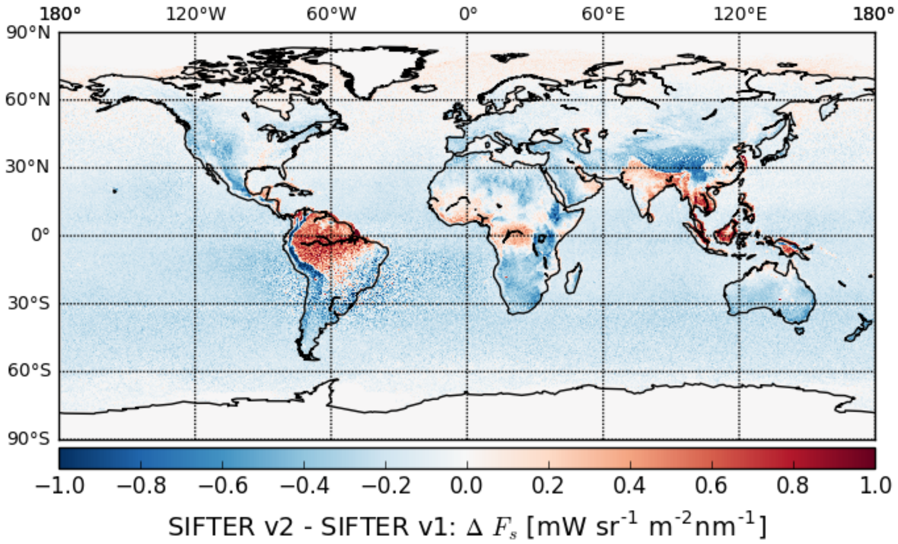


Figure 3.10: Annual mean difference in fluorescence (2011) between SIFTER v2 and SIFTER v2. Positive (red) values indicate areas where fluorescence retrieved by SIFTER v2 is higher than v1.

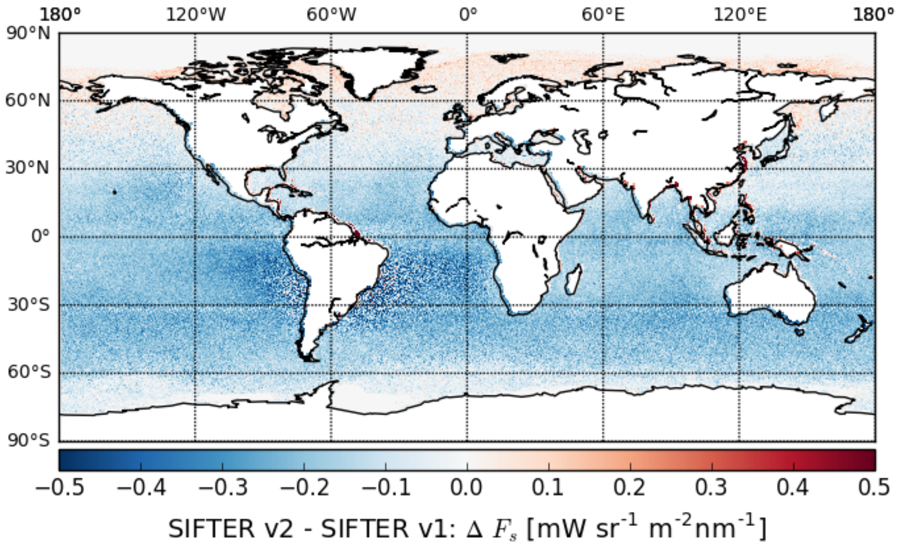


Figure 3.11: Annual mean difference in fluorescence (2011) between SIFTER v2 and SIFTER v1 over the ocean. Positive (red) values indicate areas where fluorescence retrieved by SIFTER v2 is higher than v1. The latitudinal bias is visible as it not removed in SIFTER v1.

investigation.

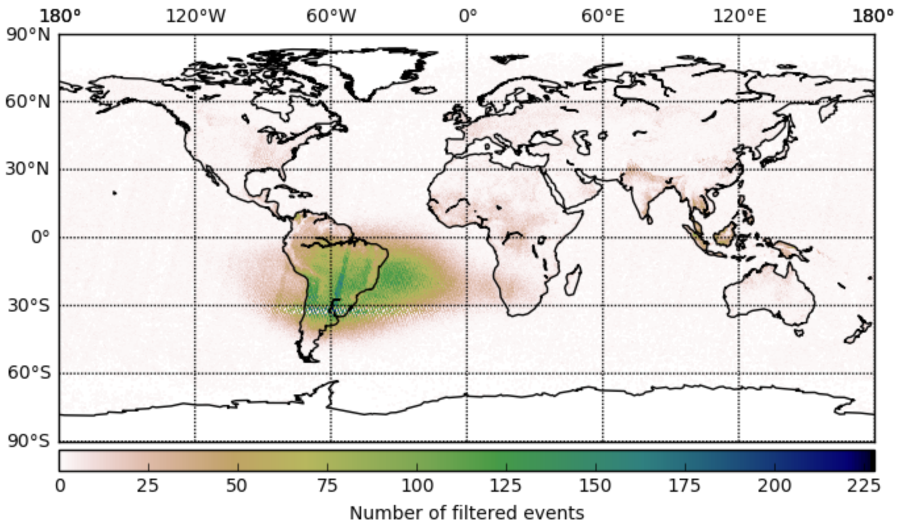


Figure 3.12: Spatial distribution of retrievals with an autocorrelation >0.2 in their fit residuals. The oval region with high numbers of filtered events in the South-East of South-America is located around the Southern Atlantic Anomaly.

3.4 Discussion

The new SIFTER v2 algorithm uses a smaller fitting window and fewer PCs than the previous version 1. This results in higher estimations of fluorescence over tropical regions, and lower fluorescence for most regions outside of the tropics. The synthetic experiments suggest that including the water vapour and oxygen absorption bands decreases the ability of the algorithm to fit a surface polynomial over vegetated areas.

As shown in the previous chapter, it is essential to account for variability in atmospheric water vapour content within the non-vegetated reference sets. This poses a problem, as the majority of non-vegetated areas with high water vapour content are located over the ocean, which has a low albedo in the red and near-infrared. Thus the TOA reflectances measured by GOME-2 have a weak signal strength which increases the signal-to-noise ratio and likely reduces the quality of the PCs. To circumvent the low signal strength we opted to filter for clouded ocean pixels, similar to Joiner et al. (2013). In theory, low clouds increase the signal strength due to their higher albedo, whilst retaining above-cloud atmospheric effects. The (sub-)tropical Atlantic ocean is rich in low clouds due to the high amount of aerosols (Kaufman et al., 2005). However, using this as reference area did not result in major differences in PCs or retrieved fluorescence, which suggests that the choice made by (Sanders et al., 2016) to choose the Sahara as reference area remains valid.

As noted by Köhler et al. (2015) and Joiner et al. (2016) there is strong latitudinal bias in retrieved fluorescence. This bias is also presented in spatial composites of SIFTER (both versions). As the magnitude of the bias can be as large as the annual mean of retrieved fluorescence it is vital to address. The cause of this bias remains unknown, though a recently discovered bias in instrument temperature could be driving this (personal communication with P. Stammes). Using daily bias correction removes orbit-specific artefacts and is therefore more suitable than a bias correction applied over monthly averages. SIFTER v2 addresses the latitudinal bias by assuming that it only depends on latitude (i.e. not affected by longitude, light intensity, surface effects, etc.). However, Joiner et al. (2016) suggests that absolute radiance levels also affect the bias. This is not account for in the current SIFTER v2 bias-correction scheme. As only (dark) ocean is used as reference, this could lead to an unsuccessful removal of bias over brighter vegetated surfaces.

Even in the idealized synthetic experiments large differences between true and retrieved fluorescence are found. To account for this filtering is required, as real-world retrievals also contain fluorescence values over 500 mW m^{-2} sr^{-1} nm^{-1} (where $<$2.5 is expected for monthly mean fluorescence over tropical regions). Based on the synthetic experiments an autocorrelation filter is implemented, which is able to greatly reduce standard deviation in retrieved fluorescence, suggesting that it is successful in removing unrealistic high or low fluorescence yields. The majority of retrievals with an autocorrelation larger than 0.2 are found near the Southern Atlantic Anomaly (SAA). Due to reduced strength of the magnetic field surrounding the Earth in this region satellites are confronted with an increased flux of energetic particles. Whilst Köhler et al. (2015) applies a filter based on RMSE in the fit residuals, the spatial distribution of rejected pixels based on RMSE is very similar to rejected pixels based on autocorrelation, with the vast majority clustering around the SAA. In addition to the SAA, tropical regions show an increased number of rejected retrievals. The synthetic experiments show that autocorrelation increases when total column water vapour over vegetated areas becomes higher than what is present in the atmosphere over which the reference set is constructed. Therefore, a higher number of rejected retrievals over such areas is expected. The apparent difficulty to explain observed TOA reflectance over areas with very moist atmosphere could also explain the fluorescence 'holes' found in SIFTER v1 over tropical regions.

Both versions of SIFTER produce negative values over desert and high-altitude regions. The negative values are larger for SIFTER v2. A possible explanation for the occurrence of such unexpected negative values is the difference in the surface polynomial during retrieval and before generating atmospheric principle components over the reference area. The latter uses atmospheric windows to fit a second-order polynomial, whilst the former has a free-varying fourth-order polynomial that is optimized together with the PCs and fluorescence. Negative values over the Saharan reference sector strongly suggest a closer inspection of the surface component as atmospheric PCs are similar to the reference set retrieval. However, other retrievals based on different instruments and different methods (i.e. Frankenberg et al. 2014) also give negative fluorescence over deserts and high-altitude areas. This can indicate that deepening of Fraunhofer lines can occur as well. However, more research is required to determine the root cause of negative fluorescence values in the retrieval.

3.5 Conclusion

A significant latitudinal bias exists within GOME-2 retrieved fluorescence. The applied reference sector approach seems able to correct for this, though it neglects possible biases in other parameters such as irradiance level. In addition, the majority of faulty retrievals can be removed by applying an autocorrelation-based filter, which rejects a retrieval if too much (unexplained) structure remains.

The choice for reference area (Sahara, tropical Atlantic ocean) does not lead to major changes in retrieved fluorescence for the year 2011. This is expected as PCs look similar, indicating similar atmospheric conditions above the two reference sectors.

SIFTER v2, which uses a narrower spectral fitting window, fewer PCs, a latitudinal bias correction and an autocorrelation filter shows higher fluorescence over tropical regions. This is more in line with expectations assuming fluorescence is positively correlation with photosynthetic activity. The next chapter will compare both version of SIFTER with other fluorescence products.

Chapter 4

Comparing Satellite-Based Fluorescence Products

4.1 Introduction

The lack of independent validation data on similar spatial scales as satellite retrievals greatly hampers an assessment of the ability of a retrieval to estimate sun-induced fluorescence. Previous studies try to circumvent this by comparing their product with other satellite-based fluorescence products (e.g. Joiner et al. 2013; Köhler et al. 2015; Sanders et al. 2016).

This chapter will compare multi-year retrieved fluorescence of SIFTER v1 (Sanders et al., 2016), SIFTER v2 (this study), NASA SIF (Joiner et al., 2013) and GOSAT SIF (Frankenberg et al., 2011a,b). Due to time limitations the main focus points are multi-year averaged spatial comparisons between the different fluorescence products.

In addition, two gross primary productivity (GPP) datasets are compared to all four fluorescence products to give a first insight in the skill of fluorescence to estimate GPP (i.e. photosynthetic activity).

4.2 Methods

Fluorescence products spanning at least the years 2010-2014 are used for comparison. The products are:

- SIFTER v1 (Sanders et al., 2016): Original KNMI retrieval that uses measurements of the GOME-2a instrument.
- SIFTER v2 (this study): Similar to SIFTER v1 but with several key changes to the algorithm (see previous chapters).
- NASA SIF (Joiner et al., 2013): Similar to SIFTER v1 but with several key differences. Also based on GOME-2a measurements.

- GOSAT SIF (Frankenberg et al., 2011b): Far-red fluorescence at 757 nm derived from GOSAT observations using deep-Fraunhofer line filling. Both the instrument and retrieval methodology are different from the other fluorescence retrievals.

For each dataset multi-year means on a coarse, 4-by-4 degree grid are calculated. The reason to choose such a coarse grid has to do with the poor spatial sampling and large single-measurement error associated with GOSAT SIF retrievals (Frankenberg et al., 2011b). It should be stressed that the GOSAT fluorescence product used here is an untested update over the published version. In addition, no level 3 data is available. As no filters are applied the results might be different and should be treated with caution. Furthermore GOSAT fluorescence used in this study is measured at 757 nm, rather than the peak at 737 nm. Therefore, GOSAT values should be increased by a factor 1.2 to be directly comparable to SIFTER and NASA products. The factor 1.2 is calculated by assuming fluorescence to be distributed according to a Gaussian function with mean 737 and standard deviation of 33.9 (in mW m^{-2} sr^{-1} nm^{-1}).

For the comparison only land pixels are used, which are selected by applying a 0.5-by-0.5 degree land use map (Jet Propulsion Lab, 2013). Gridcells that contain both land and water on a 4-degree grid are scaled by the fraction of land as removing these from the comparison eliminates import regions such as the Indonesian rainforests.

Similar to Frankenberg et al. (2011b) fluorescence is compared with GPP products (again on a 4-by-4 degree grid). Two different GPP datasets are used. The first is presented in Beer et al. (2010) and is a 0.5-by-0.5 degree multi-year mean based on extrapolation of flux tower observations using a machine learning algorithm. In contrast, the second set is derived from the terrestrial biosphere model Simple Biosphere/Carnegie-Ames-Stanford Approach (SiBCASA, see Schaefer et al. 2008). This model calculates the exchange of heat, water and carbon for all terrestrial surface on a 1-by-1 degree grid using three-hourly meteorological fields.

4.3 Results

4.3.1 Spatial comparison

SIFTER v2 has the highest absolute signal of all four products, which can be found in the tropics (Figure 4.1). The relative difference between tropical regions and agricultural areas in Western Europe and Eastern US is similar for SIFTER v2 and GOSAT. In contrast, NASA shows near similar values in France and the African tropical forests. SIFTER v2 furthermore has the lowest fluorescence signal of all four products at high (>40) northern latitudes. The high values found for SIFTER v1 in these regions are surprising as the latitude bias is not corrected for and tends to underestimate fluorescence above 30 degrees North. SIFTER v1 also has high fluorescence in Southern Brazil, close to the Southern Atlantic Anomaly. As these features are not present in the other three products it could very well be caused by the lack of filtering on fit residuals.

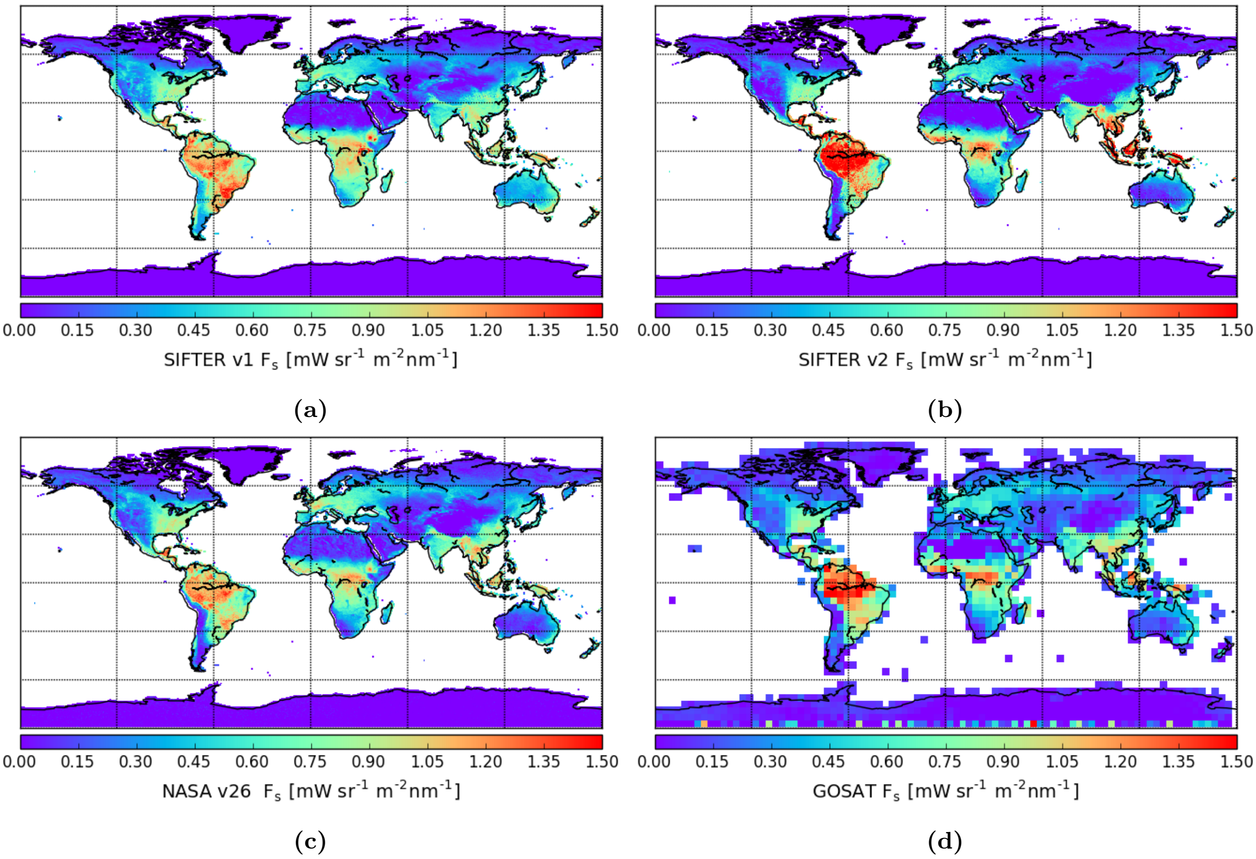


Figure 4.1: Annual mean fluorescence between 2010-2014 for SIFTER v1 (a), SIFTER v2 (b), NASA SIF (c) and GOSAT SIF (d). Note that GOSAT SIF has been multiplied by 1.2 to correct for the measurement at 757 nm, rather than at the peak value of 737 nm. Applying this correction allows direct comparison of the absolute values.

All products correlate well with SIFTER v2 (Figure 4.2). Increasing grid size from 0.5 degree to 4 degrees slightly increases the correlation between SIFTER v2 and SIFTER v1 ((a) en (b) in Figure 4.2). Whilst not shown, a similar slight increase in correlation is found when correlation SIFTER v2 with NASA at 0.5 degree and 4 degrees. Both the SIFTER v2 / SIFTER v1 and SIFTER v2 / NASA scatterplots show the effect of the tropical regions. At high fluorescence levels the linearity is reduced. This is a potential indication for saturating effects for SIFTER v1 and NASA in tropical regions as GOSAT fluorescence does retain a linear correlation with SIFTER v2.

4.3.2 Temporal comparison

A massive drought ravaged Indian crops and vegetation in the (Northern hemispheric) summer of 2009 (Neena et al., 2011). As vegetation withered, photosyn-

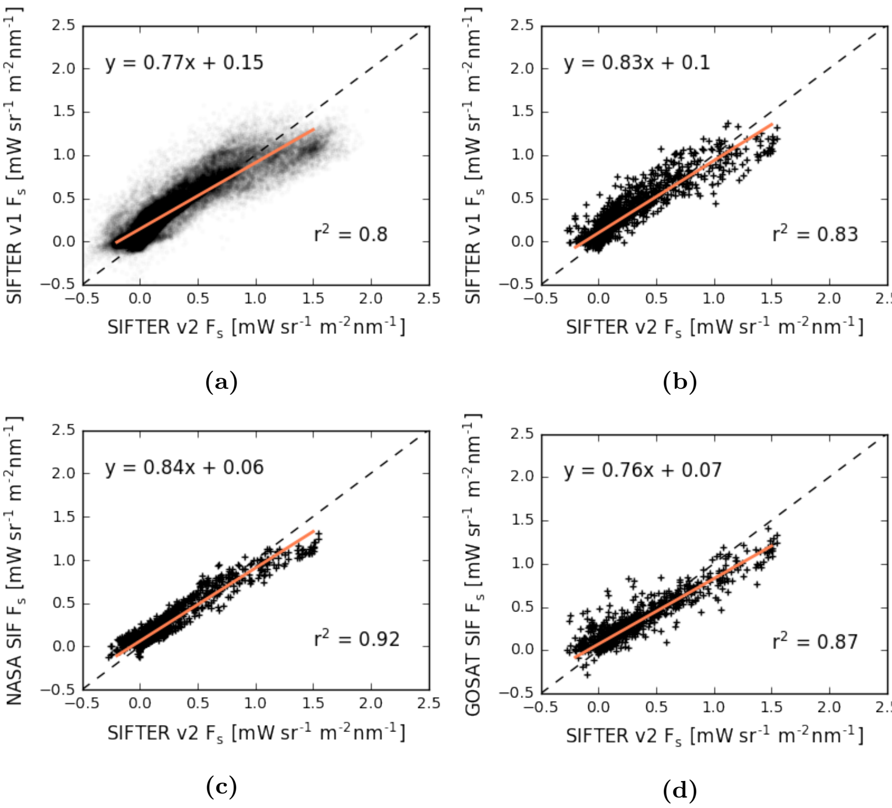
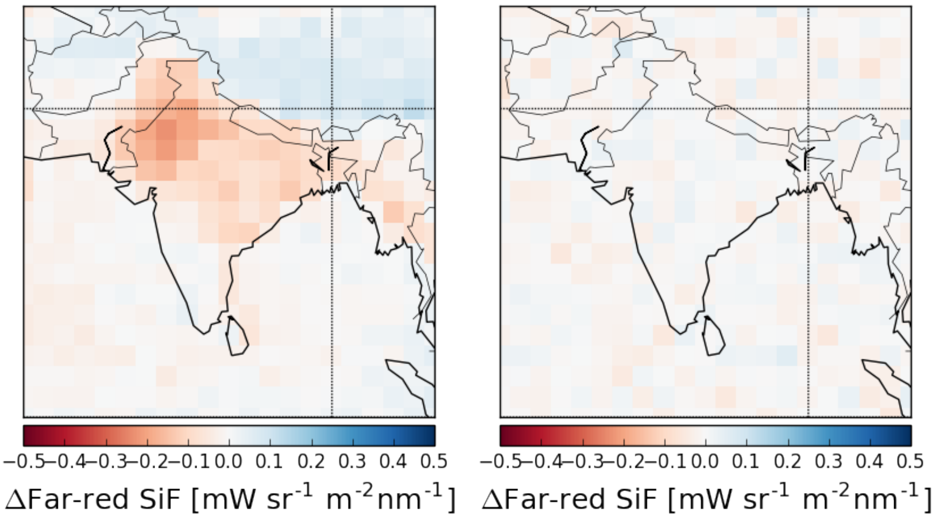


Figure 4.2: Scatterplots that correlate annual mean terrestrial fluorescence retrieved by SIFTER v2 with SIFTER v1 (a,b), NASA SIF (c) and GOSAT SIF (d). (a) shows the correlation at 0.5 degree spatial resolution, whilst figures (b,c,d) are derived from 4-by-4 degree grids. The dashed black line shows the 1:1 line. Red lines show the least of squares linear fit (parameter values also shown). It should be noted that GOSAT SIF is not scaled, thus represents far-red fluorescence at 757 nm rather than 737 nm.

thesis rates likely plummeted as well. Figure 4.3 shows the fluorescence anomaly in India for 2009, compared to a baseline spanning 2008-2014. A clear reduction in fluorescence is visible for SIFTER v2, whilst NASA shows no anomaly whatsoever. A closer inspection reveals that SIFTER v2 projects the main anomalous events to occur in the months July - September, which is consist with Neena et al. (2011). In addition, the location of the anomaly is also consistent with reports of the strongest affected regions. The annual average anomaly is $\sim$0.25 mW m^{-2} sr^{-1} nm^{-1} in North-Western India, which coincides with a 40% reduction in fluorescence. Whilst not shown here, no significant anomalies can be found in SIFTER v1. GOSAT has no complete record for 2009 and is not used.

SIFTER v2 is able to detect large-scale droughts such as the 2010 Amazon drought, the 2010 Russian drought, as well as others. More often than not these extremes do not show up in the SIFTER v1 and NASA product, whilst a



(a) **(b)**

Figure 4.3: Annual mean anomaly over for 2009 (compared to a baseline spanning 2008-2014) over India. Red areas denote a decrease in fluorescence during 2009. Shown are SIFTER v2 (a) and NASA (b).

strong impact on plant photosynthesis rates is expected. However, more analysis is required to substantiate these claims.

4.3.3 Fluorescence and Gross Primary Production

As stated in the introduction, the main scientific relevancy of fluorescence is the direct physiological link with photosynthetic activity in vegetation. Previous studies have found strong correlation between satellite-derived fluorescence and gross primary production (e.g. Guanter et al. 2014; Parazoo et al. 2013; Frankenberg et al. 2011b). 4-by-4 degree aggregated fluorescence, averaged over 2010-2014, are compared with the Beer et al. (Beer et al., 2010) GPP product and the terrestrial biosphere model SiBCASA for both version of SIFTER (Figure 4.4) and NASA / GOSAT (Figure 4.5).

Correlations are higher for SIFTER v2 than SIFTER v1 for both GPP products (Figure 4.4). SIFTER v2 shows strong linear behaviour throughout the entire range of fluorescence values with the Beer et al. GPP product. Compared to SiBCASA, there are several areas that have high GPP which don't directly correlate with fluorescence. These areas are all located in the African tropical forest regions, where SiBCASA estimates high GPP. In contrast, the Beer et al. estimation of GPP in tropical Africa are much lower, and correlate better with SIFTER v2 fluorescence. The linearity between fluorescence and GPP breaks down for high fluorescent scenes for SIFTER v1, suggesting an underestimation. This is consistent with previous findings (this study).

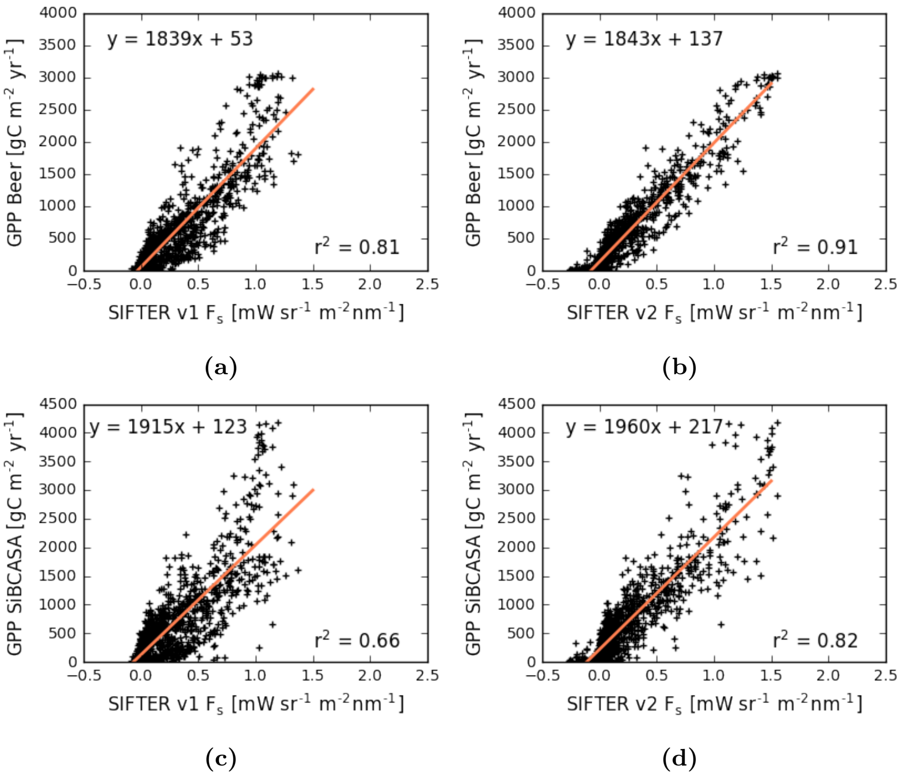


Figure 4.4: Scatterplots correlating multi-year averages of terrestrial fluorescence (SIFTER v1, v2) and GPP aggregated on a 4-by-4 degree grid. The red line visualizes the least of squares linear fit (values are presented in top-left corner).

Correlations between fluorescence retrieved by the NASA algorithm and GPP are higher than SIFTER v1 but lower than SIFTER v2 (Figure 4.5). Similar to SIFTER v1 linearity breaks down at high fluorescence / GPP. Similar correlations as NASA are found for GOSAT (Figure 4.5). In addition, the break-down of linearity is present for GOSAT, though slightly less so than NASA when correlated against SiBCASA GPP.

In general, correlations between GPP and fluorescence are high, suggesting that fluorescence can be a good proxy for photosynthetic activity. However, it should be mentioned that this analysis uses long-term, coarsely gridded averages. Therefore it remains to be seen whether these correlations hold when spatial and temporal resolution is increased.

4.4 Discussion

Multi-year spatial averages show a good visual agreement between SIFTER v2 and fluorescence derived from GOSAT. Especially the good agreement in absolute fluorescence in the Amazon rain forest and the relative reduction from

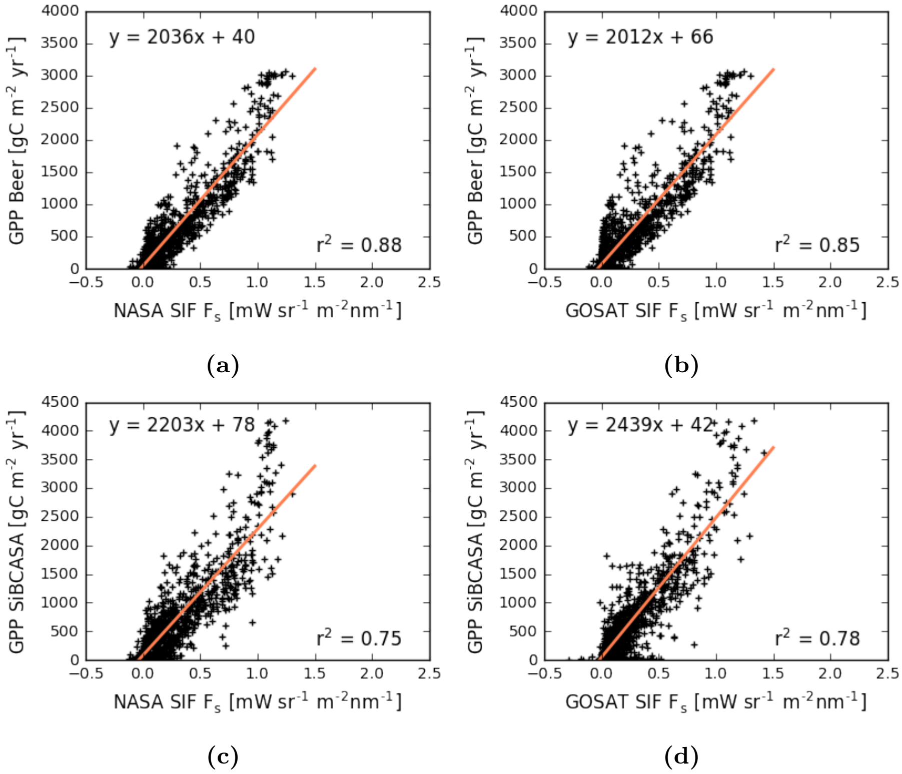


Figure 4.5: Scatterplots correlating multi-year averages of terrestrial fluorescence (NASA, GOSAT) and GPP aggregated on a 4-by-4 degree grid. The red line visualizes the least of squares linear fit (values are presented in top-left corner).

tropics to higher latitude agricultural regions (e.g. Europe, Eastern US) is promising. Whilst not shown in the results, similar results are found when comparing SIFTER v2 with OCO-2 data. As the latter has only data available from mid-2014 onwards it was not used in the comparison. The good agreement with GOSAT (and OCO-2) fluorescence gives confidence that actual fluorescence is observed, as similar patterns are found using different techniques and different sensors on different satellites. However, especially the difference in overpass time could influence the fluorescence retrieval, as GOME-2 passes in the morning (10:00 local time) and GOSAT in the early afternoon (13:00 local time, Hamazaki et al. 2005).

The difference in overpass could not only influence the absolute fluorescence signal, but also the correlation with gross primary production. van der Tol et al. (2009) found that GPP and fluorescence correlate well in afternoon conditions, but less so in the early morning. They further state that additional information on carboxylation rates is required to couple fluorescence to actual carbon assimilation. As the coupling between GPP and remotely sensed sun-induced, top-of-canopy fluorescence is yet poorly understood, one should take caution in translating one to another (Maxwell, 2000; van der Tol et al., 2014) .

That being said, correlations for a long-term annual mean between SIFTER v2 and two different GPP products (Beer et al., SiBCASA) are high, suggesting that the spatial distribution is similar. Compared to SIFTER v1 and other fluorescence products, the SIFTER v2 correlations are highest. However, as uncertainty is large both in fluorescence products as well as GPP estimates one should be careful with further conclusions.

A very promising result is the apparent ability of SIFTER v2 to detect regional anomalies in retrieved fluorescence that coincide (both spatial and temporal) with strong drought events. Several strong drought events (India 2009, Amazon 2010, Russia 2010) were successfully detected using SIFTER v2, whilst both SIFTER v1 and NASA did not show any anomalies. The droughts were of such magnitude that vegetation was severely impact. As fluorescence has been shown to decline under water stress (Flexas, 2002) it is reasonable to assume reductions in fluorescent signal during these extreme events. However, more research is required to validate such claims. In addition, no analysis was performed for GOSAT as the temporal resolution (monthly) falls outside of GOSAT range.

A strong negative anomaly over the entire tropical region was found for 2016 (SIFTER v2). Though some negative anomalies are expected due a strong El Nino, these anomalies are likely not to be realistic. As they occur mainly over tropical regions (i.e. high total column water vapour) it could be caused by a shortage of atmospheric water vapour in the reference area. Preliminary results indeed indicate that using the ocean as a reference area removes the strong anomalous events. This can be explained by the occurrence of moist air over the Saharan desert, as the moist conditions are only achieved 1-2 months a year. A dry year could therefore remove the moist conditions from your reference area. Ocean can be assumed to have more consistency in atmospheric water vapour from year to year. To resolve these anomalous events the Saharan and ocean-reference sectors can be merged. However, it is interesting to see how principle components of the atmosphere derived over the Saharan reference sector vary from year to year. If no significant variance occurs, one could opt to use the entire length of the timeseries (i.e. 2007-now) for the atmospheric model. Such a multi-year composite would include a wider range of atmospheric water vapour concentrations than currently found. However, it should be investigated how this affects the retrieval as instrument-specific degradation might no longer be captured within these PCs.

4.5 Conclusion

Application of a narrower spectral fitting window, fewer principle components, an autocorrelation filter and a latitudinal bias correction as done for SIFTER v2 increase correlation with other fluorescence products (NASA, GOSAT) compared to SIFTER v1. Especially the good correlations found compared with GOSAT are encouraging as this fluorescence product is derived using a different methodology and a different instrument.

SIFTER v2 finds local and regional reductions in fluorescence coinciding with severe drought events. SIFTER v1 and NASA are not able to pick up these signals, which really shows the potential of SIFTER v2 to be used as a new tool to assess vegetation stress during drought conditions. However, one must be wary of the uncertainty currently surrounding coupling between fluorescence emission and carbon assimilation rates.

Spatial correlation of multi-year mean averages of fluorescence and GPP correlate well, suggesting that the fluorescent signal contains information on photosynthetic activity. However, more research is required to validate SIFTER v2 against flux-tower-derived GPP (local scale) and terrestrial biosphere models. A promising candidate of such a model is the Simple Biosphere model 4 (SiB4), which calculates top-of-canopy fluorescence and can thus be compared with satellite observations. However, more independent validation data (e.g. aircraft measurements) is required to evaluate space-borne fluorescence products.

Chapter 5

Conclusions and Outlook

This study introduces the Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval (SIFTER), which estimates far-red chlorophyll fluorescence from GOME-2a observed top-of-atmosphere spectra. Important changes are made to SIFTER (version 1) based on an end-to-end test using DISAMAR-generated top-of-atmosphere spectra with a known fluorescent component. Main changes from v1 (based on Sanders et al. 2016) to v2 are:

- A smaller spectral fitting window (734-758 nm) in v2 compared to v1 (712-783 nm). This window excludes the O2A band and most of the water-vapour absorption features.
- A decrease in the number of principle components to estimate atmospheric effects from 35 (v1) to 8 (v2). This coincides with the smaller spectral fit window as atmospheric complexity is reduced due to exclusion of important absorption features.
- Filtering of retrievals based on autocorrelation within the fit residuals. This study shows that structure in the fit residuals is a strong indicator for unrealistic fluorescence estimates.
- Removal of the latitudinal bias by correcting each retrieval within a 0.5-degree latitudinal bin with the mean retrieved fluorescence over all ocean within that bin.

These changes increase the correlation with other fluorescence products, as well as two independent gross primary production estimates (the Beer et al. product and the SiBCASA terrestrial biosphere model). In addition, fluorescence retrieved using the SIFTER v2 algorithm is reduced during regional drought events. Such anomalies were not found for the other GOME-2a based fluorescence products (SIFTER v1, NASA) and highlight the potential of SIFTER v2 to be used for detecting and quantifying plant stress under prolonged periods of extreme weather.

Two main research questions are still unanswered. First, why are desert and high-altitude regions estimated to have negative fluorescence? As this trend is visible not only in SIFTER v2 but also in GOSAT, OCO-2 and NASA retrieved fluorescence it could have a physical explanation. However, it could also be caused by the different fitting procedure of surface effects during the construction of the reference set and the actual retrieval. Secondly, why does fluorescence retrieved for 2016 show large (unexpected) negative anomalies across all tropical regions? Preliminary results suggest that this can be caused by a lack of atmospheric water vapour above the Saharan reference area for that specific year, which can be resolved by including clouded ocean scenes. However, this is far from certain and more research is required to validate such a claim. In addition, SIFTER v2 can be applied to GOME-2b measurements. This will allow a direct comparison between two similar instruments and can give new insight in the robustness of the retrieval.

As no independent validation data of top-of-canopy fluorescence on appropriate spatial scales is available it remains uncertain how accurate and precise the current retrievals are. Models predicting top-of-canopy emission of fluorescence, such as constructed by van der Tol et al. (2014) can improve understanding. In addition, aircraft measurements of fluorescence are vital for validation purposes as they can sample on the appropriate scales using robust techniques (e.g. filling in of the O2A band. Meroni et al. 2009). Retrieved fluorescence can also be correlated against local estimations of GPP, such as done in Frankenberg et al. (2011b); Guanter et al. (2014); Parazoo et al. (2013); Sanders et al. (2016). The high temporal variability under extreme conditions encountered in SIFTER v2 retrieved fluorescence could be a sign that this fluorescence product can enhance the correlation between fluorescence and GPP at local levels.

Future mission such as TROPOMI (Veefkind et al., 2012) and FLEX (Rasch et al., 2008) have the potential to provide fluorescence products on an unprecedented scale, both in terms of accuracy and spatial resolution (Guanter et al., 2015).

Level2 data of SIFTER v2 fluorescence is projected to be released on the Tropospheric Emission Monitoring Internet Service (TEMIS) data-portal in 2017.

Bibliography

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review: GPP Spatiotemporal Patterns. *Reviews of Geophysics*, 53(3):785–818.

Baker, N. R. (2008). Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo. *Annual Review of Plant Biology*, 59(1):89–113.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D. (2010). Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. *Science*, 329(5993):834–838.

Berrisford, Dee, D. P., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kållberg, P. W., Kobayashi, S., Uppala, S., and Simmons, A. (2011). The ERA-Interim archive Version 2.0.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., and others (2013). Carbon and other biogeochemical cycles. In *Chapter 6: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 465–570. Cambridge University Press.

Flexas, J. (2002). Drought-inhibition of Photosynthesis in C3 Plants: Stomatal and Non-stomatal Limitations Revisited. *Annals of Botany*, 89(2):183–189.

Flexas, J., Escalona, J. M., Evain, S., Gulias, J., Moya, I., Osmond, C. B., and Medrano, H. (2002). Steady-state chlorophyll fluorescence (Fs) measurements as a tool to follow variations of net CO2 assimilation and stomatal conductance during water-stress in C3 plants. *Physiologia Plantarum*, 114(2):231–240.

Frankenberg, C., Butz, A., and Toon, G. C. (2011a). Disentangling chlorophyll fluorescence from atmospheric scattering effects in O $_2$ A-band spectra of reflected sun-light: REMOTE SENSING OF CHLOROPHYLL FLUORESCENCE. *Geophysical Research Letters*, 38(3):n/a–n/a.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T. (2011b). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity: CHLOROPHYLL FLUORESCENCE FROM SPACE. *Geophysical Research Letters*, 38(17):n/a–n/a.

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

Friedlingstein, P., Andrew, R. M., Rogelj, J., Peters, G. P., Canadell, J. G., Knutti, R., Luderer, G., Raupach, M. R., Schaeffer, M., van Vuuren, D. P., and Le Quéré, C. (2014). Persistent growth of CO2 emissions and implications for reaching climate targets. *Nature Geoscience*, 7(10):709–715.

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

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

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J.-E., Moran, M. S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J. M., and Griffis, T. J. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences*, 111(14):E1327–E1333.

Hamazaki, T., Kaneko, Y., Kuze, A., and Kondo, K. (2005). Fourier transform spectrometer for Greenhouse Gases Observing Satellite (GOSAT). page 73.

Intergovernmental Panel on Climate Change, editor (2014). *Climate Change 2013 - The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge.

Jackson, R. B., Canadell, J. G., Le Quéré, C., Andrew, R. M., Korsbakken, J. I., Peters, G. P., and Nakicenovic, N. (2015). Reaching peak emissions. *Nature Climate Change*, 6(1):7–10.

Jet Propulsion Lab (2013). ISLSCP II Land and Water Masks with Ancillary Data.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C. (2013).

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

Joiner, J., Yoshida, Y., Guanter, L., and Middleton, E. M. (2016). New methods for the retrieval of chlorophyll red fluorescence from hyperspectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY. *Atmospheric Measurement Techniques*, 9(8):3939–3967.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Middleton, E. M., Campbell, P. K. E., Yoshida, Y., Kuze, A., and Corp, L. A. (2012). Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects: simulations and space-based observations from SCIAMACHY and GOSAT. *Atmospheric Measurement Techniques*, 5(4):809–829.

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

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, 116.

Kaufman, Y. J., Koren, I., Remer, L. A., Rosenfeld, D., and Rudich, Y. (2005). The effect of smoke, dust, and pollution aerosol on shallow cloud development over the Atlantic Ocean. *Proceedings of the National Academy of Sciences*, 102(32):11207–11212.

Koelemeijer, R. B. A., Stammes, P., Hovenier, J. W., and de Haan, J. F. (2001). A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment. *Journal of Geophysical Research: Atmospheres*, 106(D4):3475–3490.

Kooreman, M. L. (2015). Spaceborne Observations of Sun-Induced Fluorescence: A 2007-2015 global time series from GOME-2a.

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

Le Quéré, C., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Friedlingstein, P., Peters, G. P., Andres, R. J., Boden, T. A., Houghton, R. A., House, J. I., Keeling, R. F., Tans, P., Arneth, A., Bakker, D. C. E., Barbero, L., Bopp, L., Chang, J., Chevallier, F., Chini, L. P., Ciais, P., Fader, M., Feely, R. A., Gkritzalis, T., Harris, I., Hauck, J., Ilyina, T., Jain, A. K., Kato, E., Kitidis, V., Klein Goldewijk, K., Koven, C., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lima, I. D., Metzl, N., Millero, F., Munro, D. R., Murata, A., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., O'Brien, K., Olsen, A., Ono, T., Pérez, F. F., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Rödenbeck, C., Saito, S., Schuster, U., Schwinger, J., Séférian, R., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Vandemark, D., Viovy, N., Wiltshire, A., Zaehle, S., and Zeng, N. (2015). Global Carbon Budget 2015. *Earth System Science Data*, 7(2):349–396.

Leth, T. C. v. (2014). Remote Sensing of Sun-Induced Chlorophyll Fluorescence.

Maxwell, K. (2000). Chlorophyll fluorescence–a practical guide. *Journal of Experimental Botany*, 51(345):659–668.

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

Neena, J. M., Suhas, E., and Goswami, B. N. (2011). Leading role of internal dynamics in the 2009 Indian summer monsoon drought. *Journal of Geophysical Research*, 116(D13).

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., and Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. *Science*, 333(6045):988–993.

Parazoo, N. C., Bowman, K., Frankenberg, C., Lee, J.-E., Fisher, J. B., Worden, J., Jones, D. B. A., Berry, J., Collatz, G. J., Baker, I. T., Jung, M., Liu, J., Osterman, G., O'Dell, C., Sparks, A., Butz, A., Guerlet, S., Yoshida, Y., Chen, H., and Gerbig, C. (2013). Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of XCO $_2$ and chlorophyll fluorescence from GOSAT: CARBON CYCLE OF SOUTHERN AMAZONIA. *Geophysical Research Letters*, 40(11):2829–2833.

Rasch, P. J., Crutzen, P. J., and Coleman, D. B. (2008). Exploring the geoengineering of climate using stratospheric sulfate aerosols: The role of particle size. *Geophysical Research Letters*, 35(2).

Sanders, A., Verstraeten, W., Kooreman, M., van Leth, T., Beringer, J., and Joiner, J. (2016). Spaceborne Sun-Induced Vegetation Fluorescence Time Series from 2007 to 2015 Evaluated with Australian Flux Tower Measurements. *Remote Sensing*, 8(12):895.

Schaefer, K., Collatz, G. J., Tans, P., Denning, A. S., Baker, I., Berry, J., Prihodko, L., Suits, N., and Philpott, A. (2008). Combined Simple Biosphere/Carnegie-Ames-Stanford Approach terrestrial carbon cycle model. *Journal of Geophysical Research*, 113(G3).

Schreiber, U. and Bilger, W. (1987). Rapid assessment of stress effects on plant leaves by chlorophyll fluorescence measurements. In Tenhunen, J. D., Catarino, F. M., Lange, O. L., and Oechel, W. C., editors, *Plant Response to Stress*, pages 27–53. Springer Berlin Heidelberg, Berlin, Heidelberg.

Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6):1704–1709.

van der Laan-Luijkx, I. T., van der Velde, I. R., Krol, M. C., Gatti, L. V., Domingues, L. G., Correia, C. S. C., Miller, J. B., Gloor, M., van Leeuwen, T. T., Kaiser, J. W., Wiedinmyer, C., Basu, S., Clerbaux, C., and Peters, W. (2015). Response of the Amazon carbon balance to the 2010 drought derived with CarbonTracker South America: AMAZON CARBON CYCLE DURING 2010 DROUGHT. *Global Biogeochemical Cycles*, 29(7):1092–1108.

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

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

Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120:70–83.

Verrelst, J., Rivera, J. P., van der Tol, C., Magnani, F., Mohammed, G., and Moreno, J. (2015). Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence? *Remote Sensing of Environment*, 166:8–21.

Wang, W., Ciais, P., Nemani, R. R., Canadell, J. G., Piao, S., Sitch, S., White, M. A., Hashimoto, H., Milesi, C., and Myneni, R. B. (2013). Variations in atmospheric CO2 growth rates coupled with tropical temperature. *Proceedings of the National Academy of Sciences*, 110(32):13061–13066.

Appendix

Statistical Summaries

Table 5.1: Statistical summary of Experiment 2 (WATER). Basis statistics are provided for all twelve combinations of fit window and number of principle components. A perfect retrieval would have a RMSE, bias, standard deviation and intercept of 0, whilst correlation (r) and slope would be 1. The absolute number of filtered pixels (autocorrelation >0.2) are given (n = 1,000 for unfiltered retrievals).

ID	λ_1 [nm]	λ_2 [nm]	# PC	RMSE [*]	r	bias [*]	σ [*]	slope	intercept [*]	faulty
	712	758	8	0.57	0.99	-0.34	0.45	0.89	-0.22	671
H_2O	712	758	20	0.65	0.99	-0.42	0.49	0.87	-0.26	485
	712	758	35	0.79	0.99	-0.52	0.59	0.79	-0.26	296
	734	758	8	0.42	1	0.12	0.4	1.06	0.04	645
NASA	734	758	20	0.62	1	0.37	0.5	1.07	0.27	587
	734	758	35	0.52	1	0.07	0.52	1.04	0.01	434
	712	783	8	0.66	0.98	-0.44	0.5	0.77	-0.21	683
KNMI	712	783	20	1.59	0.98	-0.53	1.5	0.79	-0.28	383
	712	783	35	0.75	0.99	-0.52	0.55	0.75	-0.2	337
	734	783	8	0.67	0.99	-0.41	0.53	0.75	-0.14	479
O_2A	734	783	20	1.73	0.99	-0.28	1.71	0.86	-0.11	477
	734	783	35	0.52	1	-0.19	0.49	0.85	0.01	394

Table 5.2: Statistical summary of Experiment 3 (VEGFIX). Basis statistics are provided for all twelve combinations of fit window and number of principle components. A perfect retrieval would have a RMSE, bias, standard deviation and intercept of 0, whilst correlation (r) and slope would be 1. The absolute number of filtered pixels (autocorrelation >0.2) are given (n = 1,000 for unfiltered retrievals).

ID	λ_1 [nm]	λ_2 [nm]	# PC	RMSE [*]	r	bias [*]	σ [*]	slope	intercept [*]	faulty
1	712	758	8	0.58	0.99	-0.26	0.52	0.89	-0.11	167
2	712	758	20	0.63	0.99	-0.28	0.56	0.86	-0.09	137
3	712	758	35	0.65	0.99	-0.29	0.58	0.85	-0.08	127
4	734	758	8	0.44	1	-0.02	0.44	1.04	-0.08	136
5	734	758	20	0.48	1	0.01	0.48	1.04	-0.05	131
6	734	758	35	0.59	1	-0.05	0.59	1.02	-0.08	121
7	712	783	8							971
8	712	783	20							982
9	712	783	35							984
10	734	783	8	0.77	0.98	-0.57	0.52	0.79	-0.34	441
11	734	783	20	0.44	1	-0.17	0.41	0.89	-0.03	384
12	734	783	35	0.48	1	-0.24	0.42	0.86	-0.05	360

Table 5.3: Statistical summary of Experiment 4 (VEGDYN). Basis statistics are provided for all twelve combinations of fit window and number of principle components. A perfect retrieval would have a RMSE, bias, standard deviation and intercept of 0, whilst correlation (r) and slope would be 1. The absolute number of filtered pixels (autocorrelation >0.2) are given (n = 1,000 for unfiltered retrievals).

ID	λ_1 [nm]	λ_2 [nm]	# PC	RMSE [*]	r	bias [*]	σ [*]	slope	intercept [*]	faulty
1	712	758	8	0.35	1	0.22	0.27	1	0.23	767
2	712	758	20	0.3	1	0.14	0.27	1	0.14	667
3	712	758	35	0.32	1	0.04	0.31	0.98	0.07	570
4	734	758	8	0.38	1	-0.01	0.37	1.06	-0.09	191
5	734	758	20	0.41	1	0.03	0.41	1.04	-0.03	207
6	734	758	35	0.44	1	0.01	0.44	1.04	-0.05	232
7	712	783	8							1000
8	712	783	20							999
9	712	783	35							999
10	734	783	8	0.72	0.98	-0.55	0.47	0.87	-0.42	505
11	734	783	20	0.38	1	-0.17	0.34	0.92	-0.07	461
12	734	783	35	0.42	1	-0.24	0.35	0.9	-0.1	421

Table 5.4: Statistical summary of Experiment 5 (GEO). Basis statistics are provided for all twelve combinations of fit window and number of principle components. A perfect retrieval would have a RMSE, bias, standard deviation and intercept of 0, whilst correlation (r) and slope would be 1. The absolute number of filtered pixels (autocorrelation >0.2) are given (n = 1,000 for unfiltered retrievals).

ID	λ_1 [nm]	λ_2 [nm]	# PC	RMSE [*]	r	bias [*]	sigma [*]	slope	intercept [*]	faulty
1	712	758	8	0.41	1	-0.19	0.36	0.96	-0.13	350
2	712	758	20	0.43	1	-0.2	0.38	0.94	-0.12	268
3	712	758	35	0.47	1	-0.23	0.41	0.92	-0.12	226
4	734	758	8	0.35	1	0.02	0.35	1.07	-0.09	232
5	734	758	20	0.41	1	0.08	0.4	1.09	-0.04	198
6	734	758	35	0.45	1	-0.02	0.45	1.07	-0.12	159
7	712	783	8	0.59	0.99	-0.4	0.43	0.79	-0.18	471
8	712	783	20	3.68	0.94	-0.05	3.68	0.86	0.11	387
9	712	783	35	1.98	0.99	-0.31	1.95	0.78	-0.04	348
10	734	783	8	0.76	0.99	-0.51	0.57	0.69	-0.12	446
11	734	783	20	1.29	1	-0.32	1.25	0.79	-0.05	421
12	734	783	35	0.49	1	-0.31	0.39	0.85	-0.1	387