Atmospheric
Measurement
Techniques

Improved SIFTER v2 algorithm for long-term GOME-2A satellite retrievals of fluorescence with a correction for instrument degradation

Erik van Schaik[1],[★], Maurits L. Kooreman[2],[★], Piet Stammes[2], L. Gijsbert Tilstra[2], Olaf N. E. Tuinder[2],
Abram F. J. Sanders[2,3,4], Willem W. Verstraeten[1,2,5], Rüdiger Lang[4], Alessandra Cacciari[4], Joanna Joiner[6],
Wouter Peters[1,7], and K. Folkert Boersma[1,2],[★]

[1]Meteorology and Air Quality group, Wageningen University, Wageningen, 6700 AA, the Netherlands
[2]Satellite Observations Department, Royal Netherlands Meteorological Institute, De Bilt, 3730 AE, the Netherlands
[3]Institute of Environmental Physics, University of Bremen, Otto-Hahn-Allee 1, 28359 Bremen, Germany
[4]EUMETSAT, EUMETSAT-Allee 1, 64295 Darmstadt, Germany
[5]Royal Meteorological Institute of Belgium (KMI), Ringlaan 3, 1180, Ukkel, Belgium
[6]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[7]University of Groningen, Centre for Isotope Research, Groningen, 9747 AG, the Netherlands
★These authors contributed equally to this work.

Correspondence: K. Folkert Boersma (folkert.boersma@wur.nl)

Abstract. Solar-induced fluorescence (SIF) data from satellites are increasingly used as a proxy for photosynthetic activity by vegetation and as a constraint on gross primary production. Here we report on improvements in the algorithm to retrieve mid-morning (09:30 LT) SIF estimates on the global scale from the GOME-2 sensor on the MetOp-A satellite (GOME-2A) for the period 2007–2019. Our new SIFTER (Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval) v2 algorithm improves over a previous version by using a narrower spectral window that avoids strong oxygen absorption and being less sensitive to water vapour absorption, by constructing stable reference spectra from a 6-year period (2007–2012) of atmospheric spectra over the Sahara and by applying a latitude-dependent zero-level adjustment that accounts for biases in the data product. We generated stable, good-quality SIF retrievals between January 2007 and June 2013, when GOME-2A degradation in the near infrared was still limited. After the narrowing of the GOME-2A swath in July 2013, we characterised the throughput degradation of the level-1 data in order to derive reflectance corrections and apply these for the SIF retrievals between July 2013 and December 2018. SIFTER v2 data compare well with the independent NASA v2.8 data product. Especially in the evergreen tropics, SIFTER v2 no longer shows the underestimates against other satellite products that were seen in SIFTER v1. The new data product includes uncertainty estimates for individual observations and is best used for mostly clear-sky scenes and when spectral residuals remain below a certain spectral autocorrelation threshold. Our results support the use of SIFTER v2 data being used as an independent constraint on photosynthetic activity on regional to global scales.

1 Introduction

Solar-induced fluorescence (SIF) by vegetation is directly related to light absorption by the chlorophyll complex during photosynthesis (Porcar-Castell et al., 2014; Mohammed et al., 2019). Most of the solar energy that a plant receives is used for photosynthesis, but part is released as heat and between 1 % and 2 % is re-emitted as fluorescence at longer wavelengths (Baker and Oxborough, 2004). The fluorescent light emitted by plants has a spectrally smooth signature with two peaks, one at 690 nm (red fluorescence) and one

at 740 nm (far-red fluorescence) (e.g. Daumard et al., 2012). Several studies have shown that the intensity of the SIF signal is correlated with the amount of absorbed photosynthetically active radiation (APAR) and light use efficiency (LUE) during the late morning and early afternoon (e.g. van der Tol et al., 2009; Zarco-Tejada et al., 2009). Other factors driving SIF are stress factors including diseases, canopy temperature, and nutrient and water availability. Therefore, measurements of SIF are promising indicators for photosynthetic activity of plants, and thus for monitoring carbon uptake in the Earth system, forest management, and agricultural applications.

SIF is detectable from space. Several studies have shown that SIF can be retrieved with satellite spectrometers such as MERIS (Guanter et al., 2007), GOSAT (e.g. Joiner et al., 2011; Frankenberg et al., 2011; Guanter et al., 2012), GOME-2 (e.g. Joiner et al., 2013; Köhler et al., 2015; Sanders et al., 2016), SCIAMACHY (Khosravi et al., 2015; Joiner et al., 2016), OCO-2 (Frankenberg et al., 2014; Sun et al., 2018), and recently TROPOMI (Köhler et al., 2018). Measurements by the GOME-2 sensors are particularly interesting because of the long mission lifetime of three successive MetOp satellites (2007 up to 2025–2030) with global coverage between 1 and 2 d depending on the number of instruments in orbit and their configuration. This long data record provides a good prospect of generating a climate data record of SIF (e.g. Parazoo et al., 2019). The spectral resolution of GOME-2 (~ 0.5 nm) allows a technique that matches the observed reflectances close to the wavelength of the far-red fluorescence peak to model reflectances. The modelled spectrum consists of contributions from surface reflectance, atmospheric transmittance, and fluorescence, where the latter fills in the Fraunhofer lines present in the incoming sunlight. This technique was pioneered by Joiner et al. (2013) and Köhler et al. (2015), and explored further by Sanders et al. (2016), who established the Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval (SIFTER) algorithm. Koren et al. (2018) used SIFTER fluorescence retrievals to study the effect of the 2015/2016 El Niño Amazon drought on the capacity of that tropical forest to store carbon relative to previous years. While the study suggested that SIFTER retrievals appropriately account for water vapour absorption signatures in the satellite spectra over the Amazon and provide a decadal data record, there were also clear indications for spurious negative trends in SIF that need further investigation.

Here we revisit KNMI's (Royal Netherlands Meteorological Institute) SIFTER retrieval approach for GOME-2. We propose a number of improvements to the SIFTER approach based on radiative transfer modelling tests. These improvements focus on optimising the spectral fitting window, and on calculating atmospheric transmittance terms from satellite spectra over a reference region without vegetation. The reference region is selected and sampled such that the transmittance terms are as representative as possible (in terms of surface reflectivity, viewing geometries, and atmospheric prop-

erties) for scenes with SIF. We then introduce two important quality control filters that test for the presence of unresolved spectral structures and for the presence of clouds and address the need for a latitudinal bias correction, as well as a correction for degradation in the level-1 reflectance input data. We compare the results of our new SIFTER v2 algorithm to SIFTER v1 and the data generated by NASA and discuss the uncertainty budget and limitations.

2   Data and methods

2.1   GOME-2 sensors

The GOME-2 spectrometers onboard EUMETSAT's MetOp-A (launched 19 October 2006), MetOp-B (launched 17 September 2012), and MetOp-C (launched 7 November 2018) satellites fly in a sun-synchronous polar orbit with a local Equator-crossing time of 09:30 LT in the descending node. For the SIF retrievals, we use (ir)radiances measured in the near-infrared channel (band 4, 593–790 nm), which has a spectral sampling of ~ 0.2 nm and a spectral resolution of ~ 0.5 nm. The scanning mirror of GOME-2 nominally covers a 1920 km swath back (1.5 s) and forth (4.5 s) in 6 s. The integration time of the detector is 0.1875 s, resulting in 24 forward pixels with a nominal spatial resolution of 80×40 km^2 and 8 backward pixels at 240×40 km^2. Only forward-scan pixels are processed in our data product. For the nominal swath the GOME-2 instrument achieves global coverage within a day beyond 40° latitude.

The launch of GOME-2B in 2012 motivated a reduction of the GOME-2A swath in order to achieve a better spatial resolution. The GOME-2A swath was reduced on 15 July 2013 to 960 km, resulting in a ground pixel size of 40×40 km^2 (Munro et al., 2016). Although the detectors in the GOME-2A sensors remain the same, the narrowing of the swath brings about changes in the viewing geometries and in the field-of-view. In what follows, we therefore consider that there are actually two GOME-2A sensors: one before 15 July 2013 (large pixels, wide range of viewing angles) and one after that date (small pixels, smaller range of viewing angles).

As with other space-borne sensors measuring reflected sunlight, GOME-2A is sensitive to degradation because of prolonged exposure to solar radiation and contamination of optical elements (Munro et al., 2016; Hassinen et al., 2016). The exact cause of the degradation is partially attributed to scan mirror surface contamination (also observed in GOME and SCIAMACHY) and partially due to contamination of optical surfaces in the rest of the optical light path by an unknown source of contaminants. The signal degradation is largely limited to the regions below 450 nm, but a change in reflectance values is observed throughout the spectral range, due to slightly different signal throughput variations in the solar (irradiance) optical path and the Earthshine path (because of additional degradation of the solar diffuser). This

also affects the absolute radiometric accuracy in the longer wavelength channels, like in band 4, but to a significantly lesser extent than in the shortwave channels. Since signal levels in band 4 remain very high at low noise levels, the signal-to-noise ratio is in contrast hardly affected in the regions around 740 nm that are used for fluorescence retrieval. Furthermore, the level-1b data available until now have been processed with different processor versions by EUMETSAT. Until May 2014, level-1b data were reprocessed with processor version 5.3, but since then a number of processor revisions have taken place (Table S1 in the Supplement).

To assess if signal throughput degradation or level-1 processing has influenced the stability and usefulness of the GOME-2A level-1b data, we track the variability of the top-of-atmosphere reflectances at 744 nm between 2007 and 2018. Figure 1 shows the time series of monthly mean reflectances for scenes with FRESCO+ (Wang et al., 2008) cloud fraction < 0.4 over the stable reference calibration site Libya4[1] and averaged over the globe (between 60° S and 60° N). Top-of-atmosphere reflectances are relatively stable for the early GOME-2A data with insignificant trends for the Libya4 and global mean reflectances for the period January 2007–December 2012. The late GOME-2A time series (from August 2013 onwards) however suggests a negative trend in top-of-atmosphere reflectance over Libya4 and in the global mean. The reflectance trends are consistent and amount to $-0.65\,\%\,\mathrm{yr}^{-1}$ over Libya4 and $-0.71\,\%\,\mathrm{yr}^{-1}$ averaged over the globe for August 2013–July 2018, and we will discuss this reduction later in the context of changes in the level-1 processor versions that occurred in June 2015 (version 6.1) and January 2018 (version 6.2). We conclude that GOME-2A can safely be used for fluorescence retrieval in the early period, but for the late GOME-2A data a degradation correction is warranted. We revisit the issue in Sect. 4.4.

The apparent "jump" between the baseline reflectance levels in the early (prior to July 2013) and late GOME-2A is due to the smaller range of viewing zenith angles encountered by the late GOME-2A. When we sample only the inner 12 scan positions (VZA < 35°) for the early GOME-2A period (grey line in Fig. 1b), we find a baseline reflectance level (0.196) that is consistent with the baseline reflectance level in August 2013, which is right after the detector settings change (0.195).

2.2 SIFTER retrieval algorithm

SIF can be retrieved with GOME-2 by matching a modelled spectrum that combines contributions from surface, atmosphere, and fluorescence characteristics to the measured top-of-atmosphere (TOA) reflectance spectrum in a spectral region where fluorescence is known to be substantial. Under the assumption that atmospheric scattering is negligible (for

---

[1]The Libya4 site is an often-used area for vicarious calibration because of its stable surface reflecting properties. It is centred at 28.55° N, 23.39° E; see Neigh et al. (2015).

a discussion, see Joiner et al., 2013), Eq. (1) describes the general model to simulate TOA reflectance (the dependency of reflectance R on wavelength, solar, and viewing angles is omitted here for brevity) in clear-sky situations:

$$R \approx a_\mathrm{s} T^{\uparrow} T^{\downarrow} + \frac{\pi I_\mathrm{SIF} T^{\uparrow}}{\mu_0 E_0}, \tag{1}$$

with a_s the surface albedo, $T^{\uparrow}$ and $T^{\downarrow}$ the upward and downward atmospheric transmission factors, I_SIF the fluorescence emission, μ_0 the cosine of the solar zenith angle, and E_0 the solar irradiance at the top of the atmosphere. The surface albedo and SIF emission are assumed to be spectrally smooth, but the transmission has distinct and variable spectral features from solar Fraunhofer lines, as well as atmospheric oxygen and water vapour absorption signatures. Disentangling the smooth albedo and fluorescence contributions to the TOA reflectance is possible because a number of Fraunhofer absorption lines near the fluorescence peak at 740 nm are dampened, or "filled in" by the additional fluorescence emission from the surface. In the absence of fluorescence, the Fraunhofer lines in the TOA reflectance spectrum are considerably deeper because in those circumstances the elastic scattering of sunlight at the surface maintains the relative spectral structure of the Fraunhofer lines.

The retrieval now proceeds in two distinct steps.

1. First, a table with possible solutions describing the atmospheric transmittance for a wide variety of viewing conditions is constructed. A large ensemble of observed reflectance spectra is collected over a so-called reference region that is sufficiently large and known to be free of vegetation (i.e. no fluorescence). The contribution of the surface albedo to these reference sector TOA reflectance spectra is removed by subtracting a second-order polynomial function obtained by fitting to the reflectance in three spectral bands with virtually no atmospheric absorption (712–713, 748–757, and 775–783 nm). In these bands, the TOA reflectance may be assumed to be close to the surface albedo as long as aerosol loading is modest and clouds are absent. These conditions are fulfilled by accepting only scenes with effective cloud fractions (here from FRESCO+) below a certain threshold (effective cloud fraction < 0.4) in the ensemble. The albedo-corrected spectra (centred around the 740 nm far-red fluorescence peak) are then aggregated and transformed into principal components $f_k(\lambda)$, which are described further in Sect. 4. The main components (PCs) can be interpreted as spectral features that explain the variance in spectra devoid of surface effects, i.e. dominated by atmospheric absorption. $f_0(\lambda)$ describes the main principal component, i.e. the mean transmission spectrum, and $f_1(\lambda)$ represents the second most important source of variability from, e.g. water vapour absorption. Higher-order PCs represent variability caused by water vapour, oxygen, and from
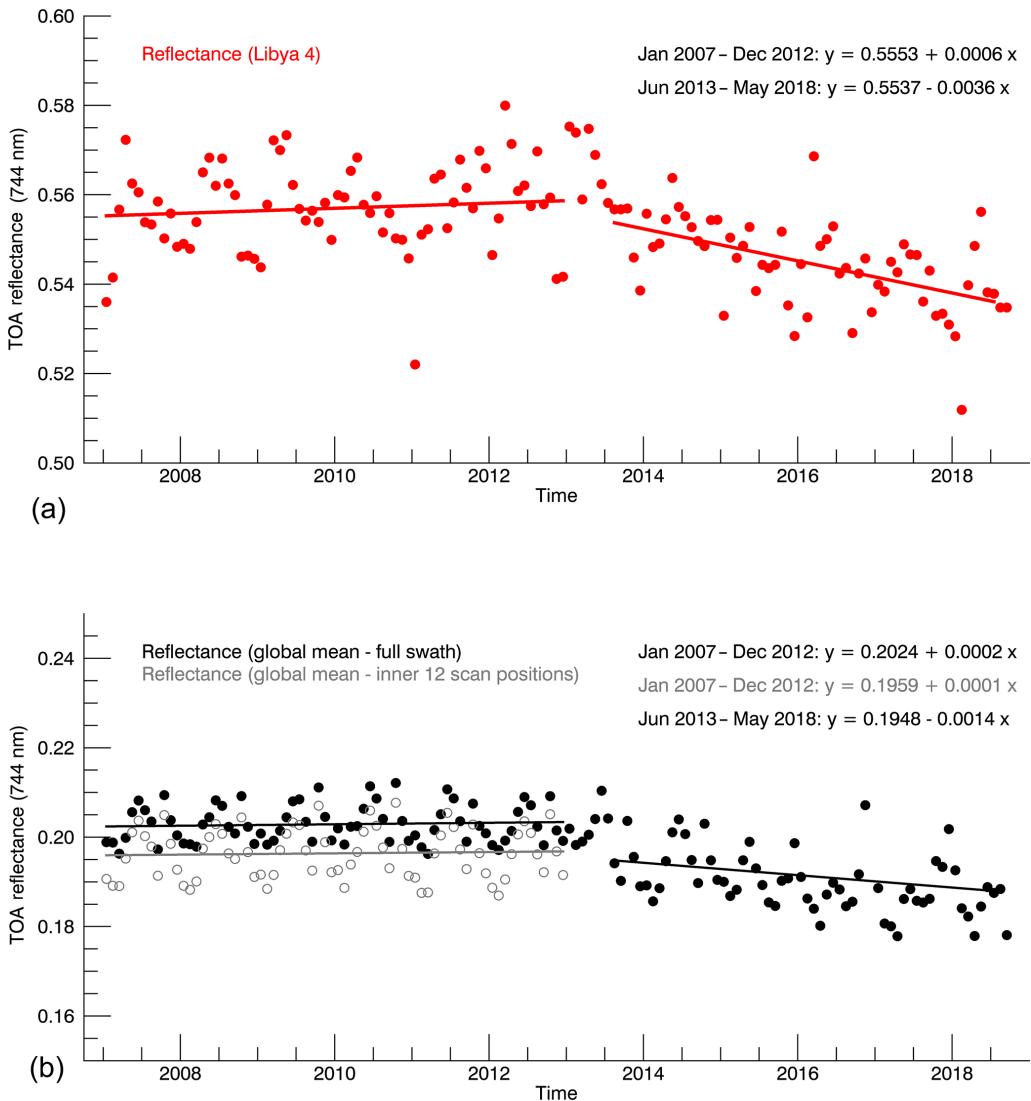
Figure 1. Top-of-atmosphere reflectance measured by GOME-2A at 744 nm as a function of time for Libya4 (**a**, red) and globally for each of the 24 scan mirror positions in the forward scan (**b**, black) and for the 12 inner scan mirror positions (**b**, grey). GOME-2A had $80 \times 40\,\text{km}^2$ pixels until 15 July 2013, and a narrower swath with $40 \times 40\,\text{km}^2$ pixels after that date. The solid lines indicate the linear trend in reflectance obtained in the early GOME-2A period for a period of exactly 6 years (January 2007–December 2012) and in the late GOME-2A period for exactly 5 years, when pixels were smaller (August 2013–July 2018). The x in the regression equation is the fractional number of years relative to January 2007 or June 2013.

other sources such as noise, unresolved surface, and instrumental effects. As each PC aims to maximise the explained variance of the transmission signal, a linear combination of the first couple of components usually explains most of the transmission signature.

2. Then, for each individual pixel, the differences between the observed ($R_\text{o}(\lambda)$) and modelled reflectance spectrum ($R_\text{m}(\lambda)$) are minimised with a Levenberg–Marquardt least-squares regression within an appropriate spectral window. The fit model is a representation of Eq. (1) and reads as follows:

$$R_\text{m}(\lambda) \approx \left(\sum_{j=1}^{n} a_j \lambda^j \right) e^{-\sum_{k=1}^{m} b_k f_k(\lambda)}$$
$$+ \frac{\pi c I_\text{SIF}(\lambda)}{\mu_0 E_0} e^{-\frac{\mu^{-1}}{\mu^{-1} + \mu_0^{-1}} \sum_{k=1}^{m} b_k f_k(\lambda)}, \quad (2)$$

where a_j is the fit coefficient best describing the surface reflectance contribution, b_k the coefficient that best matches the set of PCs to the contribution from the transmittance, and c the fitting coefficient that, multiplied with I_SIF, describes the intensity of the fluores-

cence signal. As there is a total of $n + m + 1$ fitting parameters (n for albedo, m for transmittance, and 1 for SIF) or unknowns, the spectral fitting window should be wide enough to establish an overdetermined system, i.e. include far more than $n + m + 1$ spectral samples. Leaf-level measurements over vegetated fields suggest that the SIF signal resembles a Gaussian with a width of 34 nm and a peak at 737 nm (Daumard et al., 2012), and recent work suggests that the shape of the fluorescence spectrum is stable, especially for wavelengths larger than 740 nm (Magney et al., 2019). A careful trade-off is required in the selection of the fitting window: it should preferably overlap with the stable part of the fluorescence reference spectrum (less uncertainty) and contain several Fraunhofer lines, but preferably avoid strong absorption features from oxygen and water vapour, as will be discussed in Sect 3.2.

In essence, the strength of the SIF signal is determined from the degree to which Fraunhofer lines are filled in. The retrieval method minimises the differences between modelled and observed reflectance by simultaneously fitting surface albedo, atmospheric transmittance (important for wide spectral windows such as used here), and SIF. The combination of fit coefficients that best reproduces the observed reflectance spectrum is considered to be the solution to the retrieval problem.

2.3 Retrieval improvements

We revisited a number of settings of the SIFTER v1 algorithm described in Sanders et al. (2016). A motivation for our work is that the SIFTER v1 and NASA SIF products compare reasonably well on the global scale but show little spatio-temporal correlation over the majority of the tropical forests (Fig. 9 in Sanders et al., 2016). Another incentive was the anomalous and persistent decrease in SIF over the Amazon in various studies, which required substantial detrending efforts (Koren et al., 2018). In order of relevance, these settings of interest include the following:

1. the selection of the spectral fitting window,

2. the number of principal components (PCs) used to calculate the table of transmittance,

3. the choice for the reference sector and calculation method for the PCs,

4. the application of a correction for the degradation of GOME-2A radiances.

In SIFTER v1, a wide fitting window was selected (712–783 nm), which includes spectral features from the oxygen-A band (759–769 nm) and water vapour absorption (714–734 nm). These features potentially complicate the calculation of the transmittance terms with the principal component method. Here we investigate the possibility to reduce the number of PCs $f_k(\lambda)$ by selecting a narrower window that includes the strong fluorescence signature but excludes the adjacent O_2-A and water vapour features.

In SIFTER v1, PCs were calculated from all top-of-atmosphere spectra taken over the non-vegetated parts of the Saharan desert in the 12 months preceding the GOME-2A measurement of interest. We test to what extent these Saharan-based PCs are representative, i.e. contain comparable atmospheric properties and viewing geometries for the retrieval conditions encountered for vegetated areas and locations.

3 Observation simulation experiments

3.1 Base test

In order to test our assumptions to model atmospheric transmittance, which is crucial for fluorescence retrieval, we simulated top-of-atmosphere spectra for a wide range of conditions with the Determining Instrument Specifications and Analysing Methods for Atmospheric Retrieval (DISAMAR, De Haan, 2011) radiative transfer model, developed at KNMI (based on Doubling–Adding KNMI, DAK). DISAMAR generates TOA reflectance spectra, accounts for absorption by H_2O and O_2, and describes the effects of multiple scattering and a spectrally varying surface albedo and allows convolution of a simulated reflectance signal with the predefined GOME-2 slit function (width ~ 0.5 nm). The simulations with DISAMAR allow us to control important retrieval variables such as atmospheric water vapour content, the effects of oxygen absorption, surface albedo, and viewing geometries. A top-of-canopy fluorescence signal can also be included to test which retrieval settings (spectral fitting window, number of PCs) best reproduce the magnitude of the fluorescence under different circumstances in a so-called end-to-end test.

In this first-base test, we generated 2000 DISAMAR TOA spectra in the range 712–783 nm for different prescribed parameters (albedo, surface pressure) and viewing geometries representative for those encountered by GOME-2. The prescribed parameters were obtained from a random draw of parameters found in the SIFTER v1 dataset. In the base test there was no fluorescence, as the intention was to simulate background reference spectra to be used for the determination of the PCs. We evaluated DISAMAR TOA reflectance spectra simulated for the "base" case against a large ensemble of observed GOME-2A spectra over the Sahara. Figure 2 compares the mean DISAMAR and GOME-2 atmospheric absorption signatures in the 712–783 nm range. The good agreement between the two spectra provides confidence in DISAMAR as a test tool.
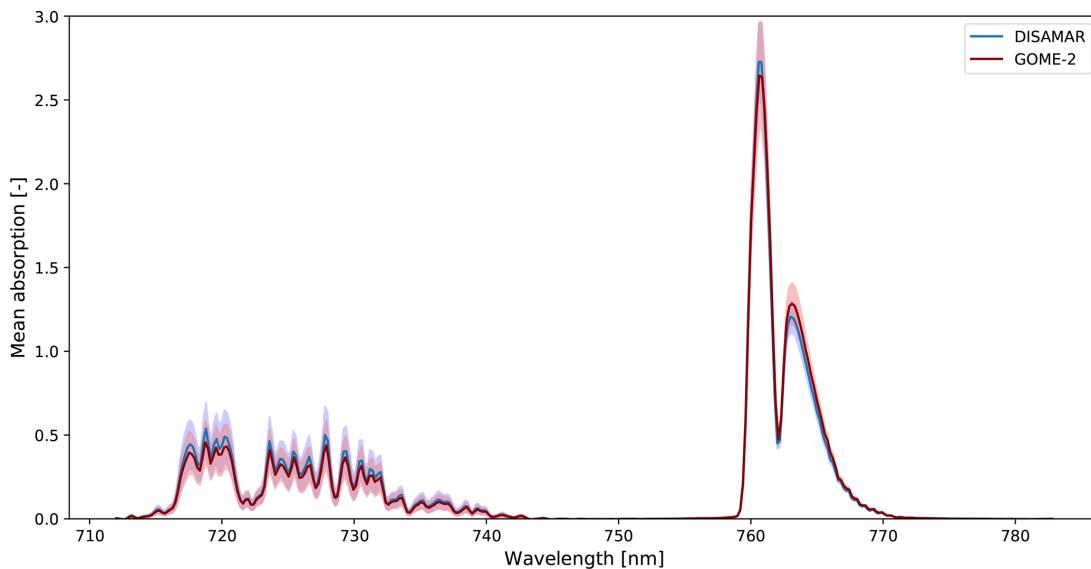
Figure 2. Mean wavelength-dependent absorption spectra for GOME-2A (red, $n \approx 30\,000$) and DISAMAR-generated (blue, $n = 2000$) top-of-atmosphere spectra, with n the number of spectra. The mean absorption corresponds to the two-way-slant optical thickness resulting from water vapour and oxygen absorption. The shading shows 1 SD.

3.2 Spectral fitting window experiments

The five other experiments focused on addressing a particular aspect of the SIF retrieval (influence of water vapour absorption, albedo characteristics, and viewing geometry) and included various strengths of fluorescence. For these experiments, 1000 TOA spectra were generated (with noise). The overall aim of the experiments was to establish the optimal spectral window and number of PCs with which to retrieve SIF. To do so, we evaluate four different spectral windows: a wide window stretching out from 712 to 783 nm, as was used in SIFTER v1, and three smaller windows that avoid the main absorption signatures from the O$_2$-A band (712–758 nm), from water vapour (734–783 nm), and a small window avoiding both bands (734–758 nm).

Table 1 summarises the settings used in the experiments and their purpose. The "fluorescence" experiment has settings that are identical to the "reference" case and allows us to investigate to what extent SIF can be reproduced in optimal conditions, as if fluorescence would be monitored over a Sahara-type surface. The "water" experiment replaces the water vapour columns over the Sahara by amounts that are more representative for tropical rain forests (ERA-Interim fields; Berrisford et al., 2011). This allows us to assess the influence of strong water vapour absorption not fully represented in the set of PCs on the SIF retrieval. The two "vegetation" experiments evaluate the sensitivity of SIF retrieval to different surface albedos that are more representative for vegetated areas than in the Sahara reference region. In the last experiment ("geometry"), the viewing geometry is changed to angles found over a Russian boreal forest (55–65° N) on a late summer day.

In the (Saharan) "fluorescence" experiment, we tested to what extent the fluorescence assumed in the generation of the TOA spectra could be reproduced with the retrieval approach described in Sect. 2. Table 2 summarises the results of this experiment in terms of retrieval bias and uncertainty. The results from this end-to-end test strongly suggest that limiting the spectral fitting window, and thereby minimising the number of PCs required, leads to the most accurate reproduction of the fluorescence signal. Clearly, the combination of a narrow fitting window (734–758 nm) and a low number of PCs (eight) gives an unbiased retrieval and the lowest uncertainties (approximately 25 %).

As a further test of the retrieval, we evaluated the goodness-of-fit, using the concept of spectral autocorrelation (Sect. S2 in the Supplement). For every spectrum, we analyse the residuals of the fit remaining after minimising the differences between the DISAMAR spectrum and the modelled spectrum from Eq. (1). In general, high fit residual values are considered to be indicative of a poor fit, but the degree of spectral lag-one autocorrelation in the residuals is another indicator as it tests the non-randomness of the residuals (NIST/SEMATECH, 2018). If the modelled spectrum perfectly explains the TOA spectrum, fit residuals appear as noise with zero autocorrelation. But if the model fails to resolve structural components of the TOA spectrum, then this will show up as structure in the fitting residuals, and, consequently, autocorrelation will be higher than zero. Our tests suggest a strong relationship between the bias and the value for χ^2_{red} in the retrieval and the degree of autocorrelation in the residual spectrum, when the autocorrelation exceeds 0.2 (See Fig. S2). This relationship is apparent for all selections

Table 1. Settings for the five experiments with DISAMAR. All experiments were done without clouds, and signal-to-noise ratio of the simulated reflectance spectra was 1000. SZA stands for solar zenith angle, and VZA stands for viewing zenith angle.

Experiment	SZA range	VZA range	H_2O column $(kg\,m^{-2})$	Surface albedo	Fluorescence $(mW\,m^{-2}\,sr^{-1}\,nm^{-1})$	Purpose
Fluorescence	21.4–66.8°	0.0–53.8°	4.0–40.0	0.41–0.45	0.0–4.0	Reproduce SIF in ideal conditions with noise
Water	21.4–66.8°	0.0–53.8°	30.0–65.0	0.41–0.45	0.0–4.0	Test influence of too low H_2O in PCs and SIF retrieval
Vegetation albedo	21.4–66.8°	0.0–53.8°	4.0–40.0	0.4	0.0–4.0	Test influence of constant spectral albedo in PCs and SIF retrieval
Vegetation red edge	21.4–66.8°	0.0–53.8°	4.0–40.0	0.06–0.5[a]	0.0–4.0	Test influence of realistic red edge spectral albedo in PCs and SIF retrieval
Viewing geometry	54.4–69.6°	0.0–53.8°	4.0–40.0	0.41–0.45	0.0–4.0	Test whether SIF can still be retrieved for viewing conditions at high latitudes

[a]The spectral albedo increases from 0.06 at 710 nm to 0.5 at 750 nm (see Fig. 2.2 in van Schaik, 2016).

Table 2. Results of the "fluorescence" experiment (mean of 1000 spectra) to reproduce fluorescence for different fitting windows and number of PCs used. The bias is defined here as the mean of the differences between assumed and retrieved fluorescence strength (on average 1.5 mW m^{-2} sr^{-1} nm^{-1}), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE. "Faulty" retrievals are characterised by high spectral autocorrelation (> 0.2) in their fit residuals. All experiments were done without clouds, and the signal-to-noise ratio of the simulated reflectance spectra was 1000.

	Bias $(mW\,m^{-2}\,sr^{-1}\,nm^{-1})$	RMSE $(mW\,m^{-2}\,sr^{-1}\,nm^{-1})$	Faulty
712–783 nm (SIFTER v1)			
8 PCs	−0.47 (−31 %)	0.71	34.7 %
20 PCs	−0.37 (−25 %)	0.59	30.6 %
35 PCs	−0.37 (−25 %)	0.62	28.8 %
712–758 nm (exclude O_2-A band, keep H_2O)			
8 PCs	−0.20 (−13 %)	0.46	23.6 %
20 PCs	−0.23 (−15 %)	0.50	20.1 %
35 PCs	−0.26 (−17 %)	0.53	18.2 %
734–783 nm (exclude H_2O band, keep O_2-A)			
8 PCs	−0.49 (−33 %)	0.57	41.9 %
20 PCs	−0.48 (−32 %)	6.41	37.9 %
35 PCs	−0.25 (−17 %)	0.49	35.4 %
734–758 nm (exclude both bands)			
8 PCs	0.0 (0 %)	0.39	16.5 %
20 PCs	0.04 (3 %)	0.43	16.2 %
35 PCs	−0.03 (−2 %)	0.50	14.6 %

of fitting window and number of PCs. We consider retrievals "faulty" when the spectral autocorrelation in the fit residuals exceeds 0.2 (corresponding with $\chi^2_{red} > 3$) for fluorescence retrievals from GOME-2 as well. Auto-correlation is strongly correlated with the fitting residuals (or rms error), and both metrics are included in the data file. For non-faulty retrievals in Table 2, the mean χ^2_{red} value was close to 1, e.g. 1.23 and 0.92 for the 734–758 nm window with 8 and 35 PCs, respectively, and 1.49 and 1.14 for the 734–783 nm window with 8 and 35 PCs, respectively.

As to why including the O_2-A band harms the retrieval of SIF from space, we performed an additional study into the sensitivity of top-of-atmosphere radiance to SIF at the surface. We simulated TOA radiances for two ensembles: one without SIF and one with a SIF strength of 4.0 mW m^{-2} nm^{-1} sr^{-1} (at 737 nm). The settings in DISAMAR were such that the ensemble average surface albedo, surface pressure, and viewing geometry were the same, so the essential difference between the two ensembles is in the presence of a SIF signal. No clouds or aerosols were included in the simulations.

The DISAMAR simulations show that the presence of a SIF signal leads to a small addition of radiance across the spectrum (Fig. 3a). The surplus radiance closely follows the magnitude and spectral shape of the fluorescence source spectrum between 740 and 758 nm but is weaker between 734 and 740 and 759 and 766 nm, where water vapour and oxygen partly absorb the SIF signal travelling from the Earth's surface towards the sensor. The sensitivity of the radiances to changes in the "state" thus shows a strong spectral dependence. Put simply, within the O_2-A band transmission is low and only half of the SIF signal makes it to the sensor. But between 740 and 758 nm (and also for 768–783 nm), transmission is full and all SIF photons reach the sensor.

To quantify this non-uniform sensitivity to SIF across the spectral range, we may use the concept of the air mass factor (AMF), which is widely used for the retrieval of weakly absorbing gases in the atmosphere. The AMF for an absorber at a vertical layer l can be calculated as $M_l(\lambda) = \frac{\partial I(\lambda)}{\partial \tau_l(\lambda)}$ (e.g. Eskes and Boersma, 2003), where $\partial I(\lambda)$ is the change in TOA
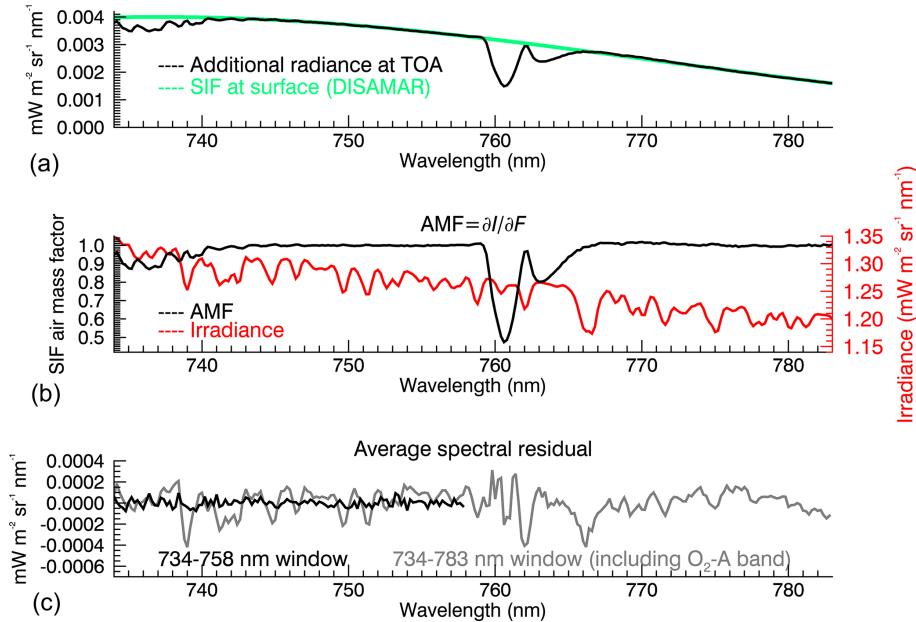
Figure 3. (a) Difference between ensemble-average simulations of top-of-atmosphere radiances with and without a SIF signal present at the surface (black line). The DISAMAR ensemble consists of 140 simulations with an average surface pressure of 977 hPa, surface albedo of 0.46 (at 750 nm), and geometrical AMF (defined as $(\cos\theta)^{-1}$) of 1.21. The fluorescence at the surface in green has a Gaussian shape with a peak at 737 nm of $4\,\mathrm{mW\,m^{-2}\,nm^{-1}\,sr^{-1}}$ and a width of 33.9 nm. (b) SIF AMF as a function of wavelength, and solar irradiance (red) based on the high-resolution spectrum from Chance and Kurucz (2010) convolved with the GOME-2A slit function. (c) Average spectral fitting residuals for end-to-end retrievals with the narrow (734–758 nm) and wide (734–783 nm) windows.

radiance caused by the addition of optical thickness $\partial\tau_l(\lambda)$ by an absorber in layer l. Scattering of light by air molecules and the generally low surface albedo results in small UV-Vis air mass factors, i.e. low vertical sensitivities for trace gases close to the Earth's surface.

In the near infrared, atmospheric scattering is an order of magnitude weaker than in the UV-Vis, thus transmission is generally close to 100 % (very long light paths being the exception). TOA sensitivity to SIF at the Earth's surface is anticipated to be optimal. To evaluate the spectral sensitivity to fluorescence, we define a spectrally resolved SIF AMF as $M_{\mathrm{SIF}}(\lambda) = \frac{\partial I(\lambda)}{\partial F(\lambda)}$, with $\partial I(\lambda)$ the change (here: addition) in the TOA radiance caused by the addition $\partial F(\lambda)$ of fluorescence at the Earth's surface. In the "ideal" retrieval situation, the AMF has a value of 1, i.e. full sensitivity to SIF at the surface. In suboptimal conditions, such as when optically thick clouds hide the SIF from detection, the AMF will be zero (e.g. Frankenberg and Berry, 2018). Within absorption bands, the AMF may be lower than 1, indicating reduced sensitivity to SIF.

Figure 3b shows that the SIF AMFs calculated from the DISAMAR between 742 and 758 nm are indeed close to 1, demonstrating the good sensitivity to fluorescence in this spectral range for our ensemble. Between 734 and 742 nm the AMF has values close to 0.9, and within the O_2-A absorption band the AMF drops to values of ~ 0.5. This explains why a spectral fit with a wide spectral window that includes the

O_2-A band will not reproduce but rather underestimate the SIF signal prescribed in the simulations. The spectral fitting procedure attempts to match all spectral features within the window. For the wide window this comprises the in-filling of the Fraunhofer lines in spectral regions where sensitivity to SIF is close to 1, but also where sensitivity to SIF drops to 0.5, i.e. the in-filling within the O_2-A band. The single, "window-average" retrieved SIF value then becomes a trade-off between partial SIF in-filling within the O_2-A band and complete in-filling outside the absorption bands. The result is a compromise: a structural underestimate of SIF. Figure 3c shows the diagnosis: we observe much larger average spectral residuals (model minus observation) for the 734–783 nm window in grey, especially around the Fraunhofer features, than for the 734–758 nm window (in black).

In the "water vapour experiment", we investigate which fitting window is optimal when higher water vapour concentrations (30–$65\,\mathrm{g\,m^{-2}}$) than encountered over the Sahara reference region (4–$40\,\mathrm{g\,m^{-2}}$) are encountered. The detailed results are summarised in Table S2. Again, the retrieval in the narrow fitting window (734–758 nm) performs best in reproducing the prior fluorescence levels (bias $< 10\,\%$), albeit with a considerable fraction of retrievals with autocorrelation > 0.2 ($40\,\%$–$65\,\%$). Retrievals with other fitting windows all performed worse in reproducing the fluorescence signal, reflecting that the relatively strong water vapour signatures around 730 nm lead to more unexplained structures

Table 3. Overview of the five experiments (mean of 1000 spectra) to reproduce fluorescence for the narrow fitting window (734–758 nm) and eight PCs. Bias is the mean of the differences between assumed and retrieved fluorescence strength (on average $1.5\,\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$), and the RMSE stands for the root of the mean of the squared deviations. Faulty retrievals were not included in the calculation of the bias or the RMSE. "Faulty" retrievals are characterised by high spectral autocorrelation (> 0.2) in their fit residuals.

	Bias ($\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$)	RMSE ($\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$)	Faulty
Fluorescence	0.0 (0 %)	0.39	16.5 %
Water vapour	+0.12 (+8 %)	0.42	64.5 %
Vegetation albedo	−0.02 (−1.5 %)	0.44	13.6 %
Vegetation red edge	−0.01 (−1 %)	0.39	19.1 %
Viewing geometry	+0.02 (+1.5 %)	0.35	23.2 %

in the fit residuals. The other three experiments ("vegetation albedo", "vegetation red edge", and "viewing geometry") led to very similar conclusions: fluorescence is best reproduced with a 734–758 nm window and with eight PCs. Inclusion of a distinct "red edge" in the spectral surface albedo (700–730 nm) did not affect the small fitting window very much, whereas fits with the wider fitting windows suffered when reproducing the assumed fluorescence levels. One notable result is that changing the viewing geometries to those representative of a late-summer boreal morning over Russia did not strongly affect the retrieval results either. This suggests that a reference set based on the relatively small range of solar zenith angles encountered over the Saharan reference region can still be used for retrieving fluorescence over high-latitude regions. Table 3 presents the overview of the main results for the narrow fitting window.

To summarise, our tests show that fluorescence is consistently reproduced with a narrow fitting window (734–758 nm) and eight PCs under a wide range of different but realistic simulated retrieval conditions. This 734–758 nm window is similar to the fitting window selected by Joiner et al. (2016). The test results also suggest that it is important to filter out retrievals with persistent spectral structure in their residuals, and this can be done via the auto-correlation test proposed here. Higher amounts of atmospheric water vapour than present over the Sahara may lead to absorption imprints in the TOA spectra not captured by the PCs, and this leads to a relatively large fraction of faulty retrievals. This indicates that defining a reference sector to include scenes with similar water vapour amounts as encountered over vegetated areas has good potential to improve the retrieval. We will investigate this further in Sect. 4. Limiting the spectral range of the fitting window by excluding absorption bands is effective in reducing the number of PCs needed, both for GOME-2A and DISAMAR spectra. Including the bands with O_2-A and water vapour absorption signatures (as in SIFTER v1) strongly reduces the explained variance of principal components 1 to 4. Wider spectral fitting windows thus require more PCs to capture atmospheric effects, and selecting a narrow fitting window in the 734–758 nm range improves our

SIFTER retrievals. Our AMF analysis and Fig. 3 suggest that the influence of H_2O variability could be reduced further in future algorithm improvements by narrowing down the spectral window even further to 740–758 nm.

4 Retrieval method

4.1 Selection reference sector and period

Two questions for the calculation of the principal components need to be addressed.

- Is the pool of spectra large enough to generate a sufficiently accurate set of PCs (length of period and size of reference area)?

- Are the PCs based on atmospheric scenarios sufficiently representative for vegetated areas (spatial representativeness)?

In SIFTER v1, PCs were determined from spectra observed over the barren Sahara in the previous year. The drawback of taking an ensemble from the previous year is that atmospheric conditions over the reference area may differ from year-to-year, making it difficult to compare year-to-year changes in fluorescence strength. Experiments with shorter reference periods (not shown) showed much stronger inter-annual variability in retrieved SIF, suggesting sensitivity to the period chosen. A possible explanation for this sensitivity is that variability in water vapour over the Saharan reference sector translates into variability in SIF over, e.g. the Amazon region. We thus generate PCs for GOME-2A over the Saharan reference sector for the entire period 2007–2012 instead of the moving yearly window. PCs from a multi-year ensemble of Saharan measurements capture the wider dynamic range of water vapour (as shown in ECMWF ERA-Interim data in Dee et al., 2011, and evaluated in, e.g. Grossi et al., 2015) encountered over the 6-year period compared to water vapour variability occurring within 1 year. GOME-2A settings changed in July 2013, and thus for retrievals done after that date (with a smaller field-of-view and smaller range of viewing zenith angles) the set of PCs was based on

narrow-swath-only measurements over the reference sector (all PCs except the last (PC9) were spectrally similar). The Sahara reference sector here is defined as the region (16–30° N, 8° W–29° E), with scenes free of vegetation according to the USGS Global Land Cover Characterisation database version 2 (https://doi.org/10.5066/F7GB230D; Loveland et al., 2000) and effective cloud fraction (Wang et al., 2008) from FRESCO < 0.4.

We then evaluated whether the SIFTER v1 reference area over the Sahara was sufficiently representative in terms of total water vapour. Total water vapour is typically lower over deserts than over moist tropical forests. Our sensitivity tests indicate that non-linear effects due to water vapour saturation cause effects that are not captured by the PCs, and the autocorrelation filter then automatically rejects those cases (Table 3). We tested whether using spectra from a different reference sector, the moister tropical Atlantic Ocean (15–30° N, 30–45° W), was useful to generate a more representative set of PCs. Because the low ocean albedo leads to low TOA reflectance levels, we also included ocean scenes with effective cloud fractions larger than 0.4 in the ensemble. We found that the tropical Atlantic ensemble has very similar PCs (the first four were practically indistinguishable) to the Sahara set and that using Atlantic Ocean PCs instead of Sahara PCs leads to SIF values that are highly consistent with the ones retrieved with the Sahara set, consistent with Sanders et al. (2016). This might seem counter-intuitive, as higher water vapour concentrations are present over the Atlantic Ocean than over the Sahara. However, using mostly cloudy scenes implies that the lower atmosphere is partly obscured from view. Water vapour in the lower troposphere may therefore (still) be under-represented in the TOA reflectance and therefore in the PCs. Because using the Atlantic reference area (cloudy pixels) did not result in major differences in PCs or SIF, we keep the Sahara reference sector (for more details, see Sect. 3.3.2 in Van Schaik, 2016, and Sanders et al., 2016).

4.2 Retrieval settings and demonstration

Table 4 summarises the main changes between SIFTER v1 and SIFTER v2, as motivated by the observation simulation experiments (Sect. 3) and tests with real data (Sect. 4).

As a demonstration of the retrieval principle, here we show the operation of the fitting approach from the GOME-2A spectra measured on a clear-sky summer day (15 July 2007) over the eastern US. We performed two different retrievals on the GOME-2A spectra: one in which the fitting model (Eq. 2) accounted for a SIF term and one in which the SIF term was omitted (a reflectance model with only albedo and transmission contributions on the right-hand side of Eq. 2). Figure 4 shows the observed and modelled radiances for the two approaches over the eastern US on 15 July 2007. The differences between the two modelled and observed radiances are indistinguishable by eye, but the magnitude of the unexplained fitting residuals clearly reveals a better fit when in-

cluding SIF in the fitting model. In fact, the rms of the residuals is some 40 % smaller for the fit including the fluorescence term. The residuals improve in spectral regions with prominent solar Fraunhofer lines, highlighting that the broad fluorescence signal from vegetation leads to infilling of the solar Fraunhofer lines from below, and is well detectable even for an instrument with a relatively coarse spectral resolution of 0.5 nm such as GOME-2A.

4.3 Zero-level adjustment

Similar to other retrievals (Köhler et al., 2015; Joiner et al., 2016), global SIFTER v2.0 retrievals show indications for a bias in the retrieved fluorescence levels that systematically depend on latitude. The cause of the bias is likely related to changes in the instrument slit function driven by changing optical bench temperatures over the orbit (Munro et al., 2016), resulting in an increase in slit function width along the orbit (Azam and Richter, 2015; De Smedt et al., 2012; Miles et al., 2015; Beirle et al., 2017). Our PCs capture the slit function variability due to seasonal and long-term optical bench temperature changes only as far as they occur within the reference sector over the years 2007–2012, i.e. between 0 and 20° N. A widening of the slit function over the orbit would have implications for the observed depth of the Fraunhofer lines: the wider the slit, the less deep the Fraunhofer lines. Shallower Fraunhofer lines are interpreted by the fitting algorithm to have been caused by fluorescence, which should lead to a positive fluorescence bias in the Southern Hemisphere. Conversely, sharper and deeper Fraunhofer lines (relative to the width of the lines over the Sahara) are interpreted as being caused by negative fluorescence, and this is anticipated as a negative bias for latitudes north of the Sahara.

Here we make an estimate of the zero-level bias by using the fluorescence signal strength observed over the Pacific Ocean (130–150° W), where far-red fluorescence is supposed to be zero. We obtain a daily bias estimate in steps of 1° latitude, and then subtract this value for every fluorescence retrieval within the same 1° latitude band on that day, thus assuming that the bias does not depend on longitude. The bias is determined by a linear regression of the retrieved fluorescence to the reflectance levels for all GOME-2 retrievals with a cloud fraction < 0.4 and autocorrelation < 0.2 within the 1° by 20° Pacific box. The regression coefficients a (offset) and b (slope) then describe the bias B_i for a particular pixel i as $B_i = a + b \cdot R_i^{744}$ with as R_i^{744} the reflectance level at 744 nm for pixel i. Tests indicate that such bias correction results in a smoother bias estimate than simply prescribing the correction B_i determined from the mean reference sector reflectance (right panel of Fig. S3).

Daily bias corrections are similar to monthly mean corrections, as shown by the similarity between the daily zero-level estimate for 15 July and the average zero-level adjustment for the month of July in Fig. 4a. Note that there are very few

Table 4. Summary of individual changes in fluorescence retrievals from SIFTER v1 (Sanders et al., 2016) to SIFTER v2 (this work).

	SIFTER v1	SIFTER v2 (Jan 2007–June 2013)	SIFTER v2 (Aug 2013–Dec 2018)
Spectral fitting window	712–783 nm	734–758 nm	734–758 nm
Reference sector region	Sahara	Sahara	Sahara
Number of PCs	35	10	10
Reference period	Past 12 months	Jan 2007–Dec 2012	Jan 2007–Dec 2012, but only scenes with VZA < 35°.
Albedo retrieval	fourth-order polynomial (5 terms)	fourth-order polynomial (5 terms)	fourth-order polynomial (5 terms)
Solar irradiance spectrum	Chance and Kurucz (2010) spectrum, convolved with GOME-2 slit function and scaled with Earth–Sun distance.	Chance and Kurucz (2010) spectrum, convolved with GOME-2 slit function and scaled with Earth–Sun distance.	Chance and Kurucz (2010) spectrum, convolved with GOME-2 slit function and scaled with Earth–Sun distance.
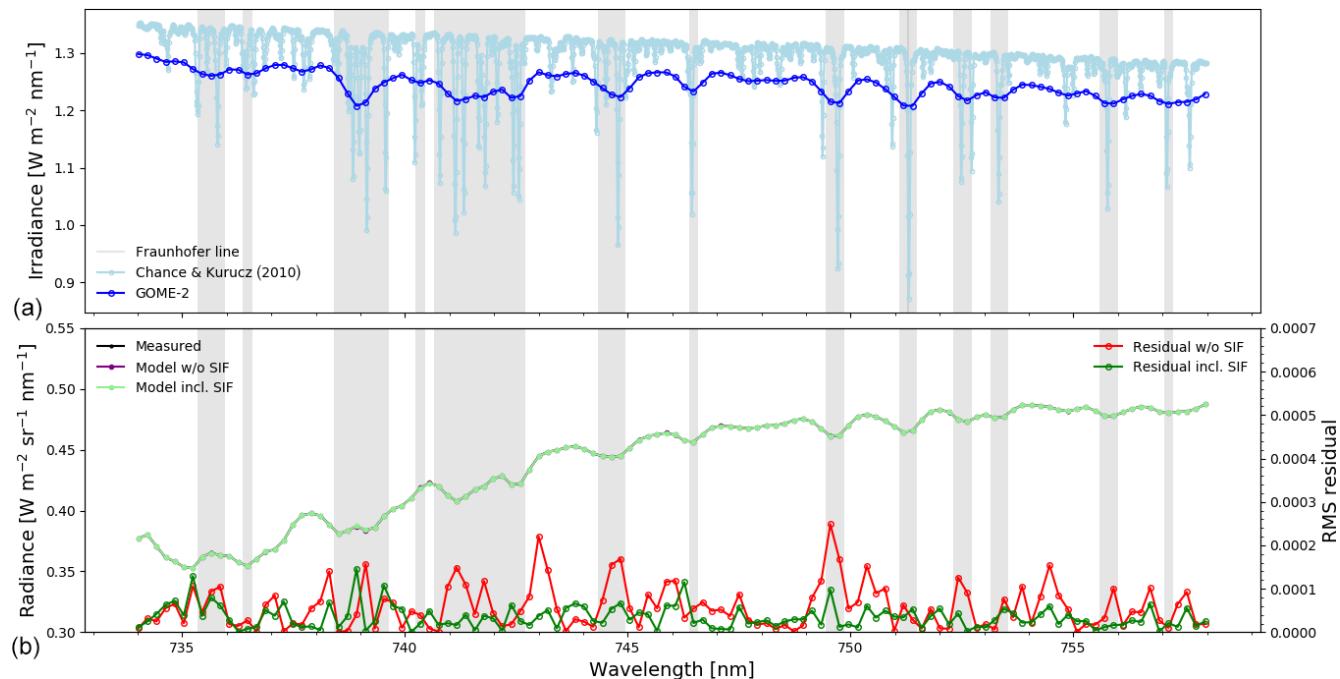
Figure 4. (b) The measured and modelled (with and without accounting for fluorescence) radiance spectrum in the SIFTER v2 fit window (734–758 nm) on 15 July 2007 over vegetated areas in the eastern US (orbit 3821). The spectra represent the average over 378 cloud-free pixels with a fluorescence signal strength $> 1.5\,\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$. The measured and modelled spectra are indistinguishable, and thus only the light green line is visible. The dark green and red lines show the absolute radiance residuals for a fit including and excluding fluorescence. Panel **(a)** shows the 0.01 nm resolution solar irradiance spectrum in light blue (Chance and Kurucz, 2010) and the irradiance spectrum convolved with the GOME-2A slit function in dark blue. The grey lines connecting panels **(a)** and **(b)** indicate the spectral regions where the solar Fraunhofer line infilling by fluorescence is relevant.

useful observations (due to clouds) between 40 and 55° N in the Pacific Ocean reference zone (130–150° W), and thus on 15 July the zero-level adjustment has quite pronounced values. Figure 5b and c display the unadjusted and adjusted fluorescence levels retrieved on 15 July 2007. In the operational retrieval, the bias correction is determined based on the most recent days as soon as a sufficient number of pixels has been collected in a latitude bin (at least 10 per bin). In prac-

tice, the bias correction is often based on the last day, but for some latitude bands, the correction can be based on pixels dating back at most 14 d. We see only small differences over land, and the most distinct difference is the removal of the positive values over the oceans in the Southern Hemisphere. Figure S3 shows the annual mean bias correction as a function of latitude for different years. The magnitude of the bias can exceed $0.2\,\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$, some 10 % of the high-

(a)

(b)

(c)
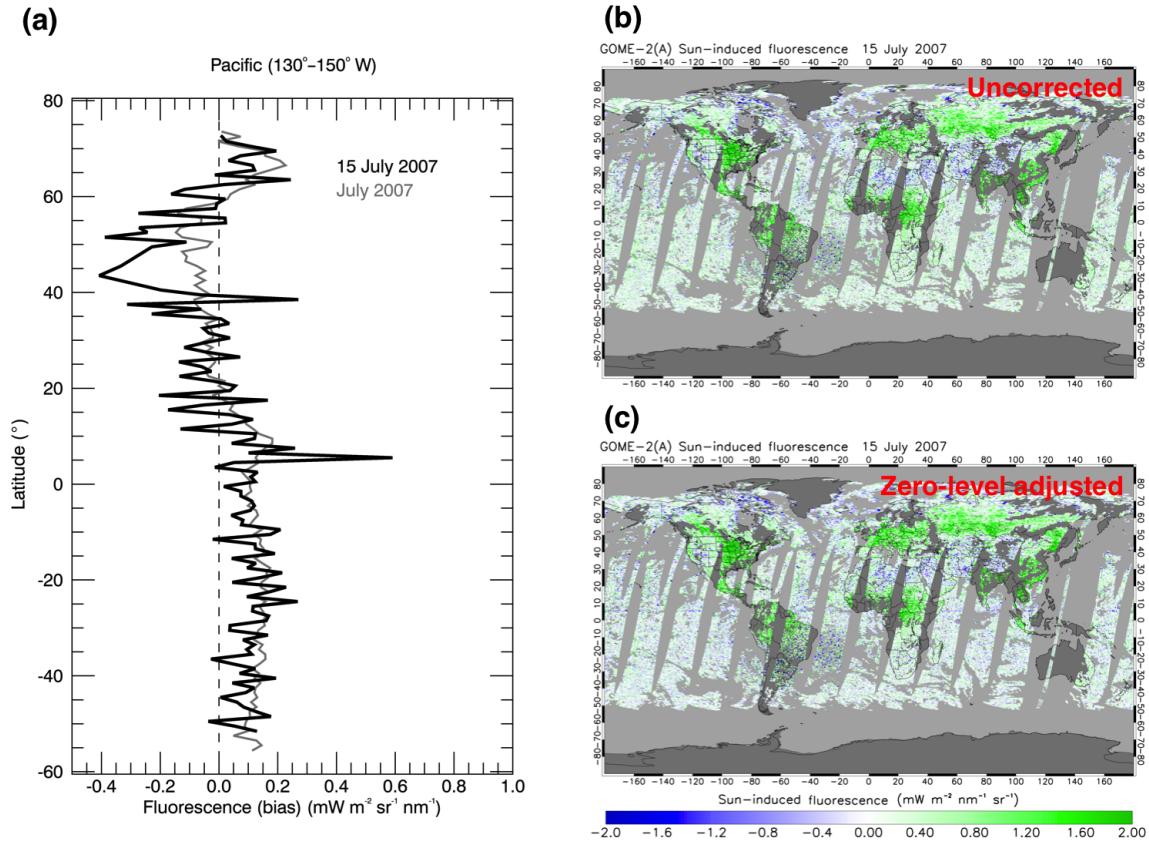
Figure 5. (a) Zero-level fluorescence found over the Pacific (130–150° W) on 15 July 2007 and averaged over July 2007, **(b)** level-2 SIF data on 15 July 2007 without zero-level adjustment, and **(c)** level-2 SIF data after the zero-level adjustment for 15 July 2007. The effect of the adjustment can be seen over the Southern Ocean, which after adjustment has SIF values close to $0\,\mathrm{mW\,m^{-2}\,nm^{-1}\,sr^{-1}}$.

est fluorescence signals. The bias correction follows a similar pattern from year to year (Fig. S3) and we do see indications for a weak long-term trend in the bias correction that is most likely caused by the gradual change in optical bench temperatures and hence in the slit function.

Over desert and high-altitude regions, negative fluorescence values are retrieved, and these may be the consequence of the Sahara reference set not being sufficiently representative for these areas. Radiative transfer tests described in Botia (2017) show that the TOA reflectance decreases with decreasing surface pressure. This suggests that observed reflectances are evaluated with too high reflectances derived from PCs over the Sahara (surface pressures 900–1000 hPa) and that these reflectance overestimations are then compensated for by negative fluorescence. This negative bias can be resolved by using PCs obtained from a reference set over barren high-altitude regions (Botia, 2017).

4.4 Correction for degradation in GOME-2A reflectances

Previous results with SIFTER v2, such as those discussed in Koren et al. (2018), showed a substantial reduction of the SIF signal in the later years of the GOME-2A period, which required detrending. The reduction in SIF is not real but the consequence of reductions in throughput levels ("degradation") or changes in calibration key data. Here we investigate the feasibility of a correction for the reduction of the reflectances shown in Fig. 1. We focus on changes in the post-2013 reflectances over the Sahara, since the principal components $f_k(\lambda)$ are derived from stable reflectances obtained over this region in the 2007–2012 period. Figure 6a shows the degradation of the reflectances averaged over the period that a level-1 processor was applied. The degradation is weakest in version 6.0 (June 2014–June 2015) and strongest in the v6.2 (January–December 2018) level-1 processor. The degradation changes with season (not shown). Based on these findings, we calculate degradation correction spectra[2] for all seasons since June 2014 to account for the fact that the principal components based on 2007–2012 spectra would lead

---

[2]The seasonal correction spectra for, e.g. version 6.0 (July 2014–June 2015) are defined as $c(\lambda) = \frac{R_{\mathrm{S,v5.3}}(\lambda)}{R_{\mathrm{S,v6.0}}(\lambda)}$, i.e. the ratio between the mean climatological (2007–2012) reflectance spectrum in a particular season over the Sahara in version 5.3 and that of the seasonal reflectance spectrum over the Sahara in version 6.0.

to too strong transmission terms in Eqs. (1) and (2). As an example we show correction spectra for December–January–February (DJF) in the years after the processors changed in Fig. 6b, which illustrates that the corrections generally have a spectral pattern with slightly stronger corrections required at the shortwave part of the spectrum where water vapour absorbs. The correction spectra, with values in the range from 1.00 to 1.05 are applied to the observed reflectance spectra $R_o(\lambda)$. The correction ensures that prior to fitting the observed reflectance spectrum is elevated to a level anticipated if no degradation had taken place after June 2014, so that the spectral fitting with the principal components obtained for the stable 2007–2012 period (when reflectances were higher) is still appropriate. The effect of correcting for throughput reductions on retrieved SIF levels is discussed in Sect. 5.3.

4.5 Uncertainty budget

Uncertainty in the retrieved fluorescence is driven by uncertainties in the measured reflectance (i.e. by measurement noise) and by uncertainties associated with solving the inverse non-linear problem. In the non-linear Levenberg–Marquardt regression, the differences between the measured and modelled reflectance are minimised by the parameters best describing the surface albedo (5 fit parameters), the atmospheric transmission (10 parameters for the 10 principal components), and the intensity of the fluorescence via Eq. (1). Correlations among the parameters may play an important role in the fitting, and contribute to (additional) uncertainty in the fluorescence. The diagonal element of the error covariance matrix provides the uncertainty in the fluorescence fit parameter. We can consider this a realistic estimate of the uncertainty for a single retrieval, as long as the description of surface albedo and transmission is appropriate and does not suffer from possible misrepresentations. Such misrepresentations could be too low extinction from atmospheric water vapour described in the set of PCs (as compared to the scene of interest) or too much extinction in the PCs for scenes with higher terrain height than the Sahara (Botia, 2017).

Figure 7 shows typical pixel uncertainties obtained from the Levenberg-Marquardt fit as a function of the fluorescence signal strength over land (Fig. 7a) and as a function of scene brightness (Fig. 7b). The median single-pixel uncertainty is $0.6\,\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$ and is independent of the fluorescence signal itself. This implies that the relative uncertainty for strongly fluorescent scenes is 25 %–50 %, but it can easily exceed 100 % for scenes with a moderate signal of $< 0.5\,\mathrm{mW\,m^{-2}\,sr^{-1}\,nm^{-1}}$. Figure 7b shows that fluorescence retrieval is most precise for scenes that are relatively dark. This illustrates that the additional top-of-atmosphere radiance signal from SIF is optimally retrieved against a dark background.

4.6 Data filtering and recommended usage

We recommend using the SIFTER v2 data after applying the autocorrelation filter. Both synthetic and real retrievals are of much better quality when spectral autocorrelation < 0.2. This filter rejects faulty retrievals, where too much (structured) spectral residual remains after fitting, as discussed in Sect. 3.2 and in Sect. S2 of the Supplement. The filter also greatly reduces the SD in the monthly mean retrieved fluorescence, indicating that it successfully removes unrealistically high or low fluorescence values. On average, the autocorrelation filter rejects 5 %–10 % of the retrievals. The majority of the rejected pixels overlap with the South Atlantic Anomaly and to lesser extent with very moist vegetated regions in the tropics (Fig. 3.12 in van Schaik, 2016).

The retrievals have been done for scenes with (FRESCO) cloud fractions < 0.4. With such a cloud fraction, there is sufficient light coming from the Earth's surface to effectively detect fluorescence (Frankenberg et al., 2014; Guanter et al., 2015). Users can expect better quality retrievals for scenes with the lowest cloud fractions, as the surface fluorescence signal then propagates unhindered to the sensor (e.g. Köhler et al., 2015).

5 Retrieval results

5.1 Comparison between SIFTER v2 and SIFTER v1

We first examine SIF from SIFTER v2 against our previous SIFTER v1 algorithm (Sanders et al., 2016) for January and July 2011. The spatial distribution of monthly mean SIF agrees reasonably well between SIFTER v1 and v2 (Fig. S4) with the highest levels of fluorescence in the vegetated regions of the world and low values over the oceans, deserts, and barren regions. The spatial correlation between the averaged and gridded SIFTER v2 and v1 data is good: $R^2 = 0.77$ ($n = 162\,169$) in January and $R^2 = 0.79$ ($n = 163\,553$) in July.

In the tropics, SIFTER v2 shows a distinctly different SIF pattern from SIFTER v1, with up to 25 % higher fluorescence over the Amazon and Kalimantan in SIFTER v2 (Table S5). We attribute the low tropical fluorescence values to the PCs used in SIFTER v1, which have been derived from a 1-year ensemble of reference scenes only and likely do not capture the full range of water vapour concentrations encapsulated in the 6-year (2007–2012) PC set used in SIFTER v2. Moreover, the SIFTER v1 fit window is much wider, and presumably has stronger sensitivity to interference from water vapour and oxygen absorption, which leads to biases in the fluorescence (Table 2). Another reason for the differences between SIFTER v2 and v1 is the use of the autocorrelation filter in the former. On average, the autocorrelation filter rejects < 5 % of the retrieved pixels, but the vast majority of these rejections are located within the region
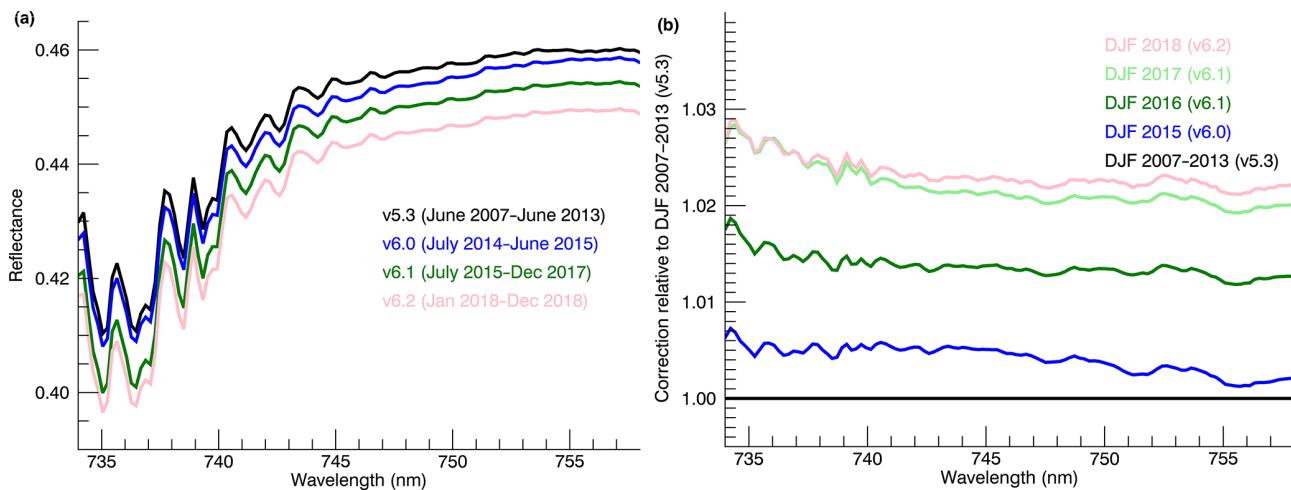
Figure 6. (a) Average reflectance spectra over the Sahara reference area for the different processor versions of GOME-2A. For version 5.3, only scenes with viewing zenith angles < 35° have been selected to stay consistent with the viewing geometries after the detector switch in June 2013. **(b)** Spectral corrections calculated as in footnote 2 for the winter seasons following the level-1 processor version changes.
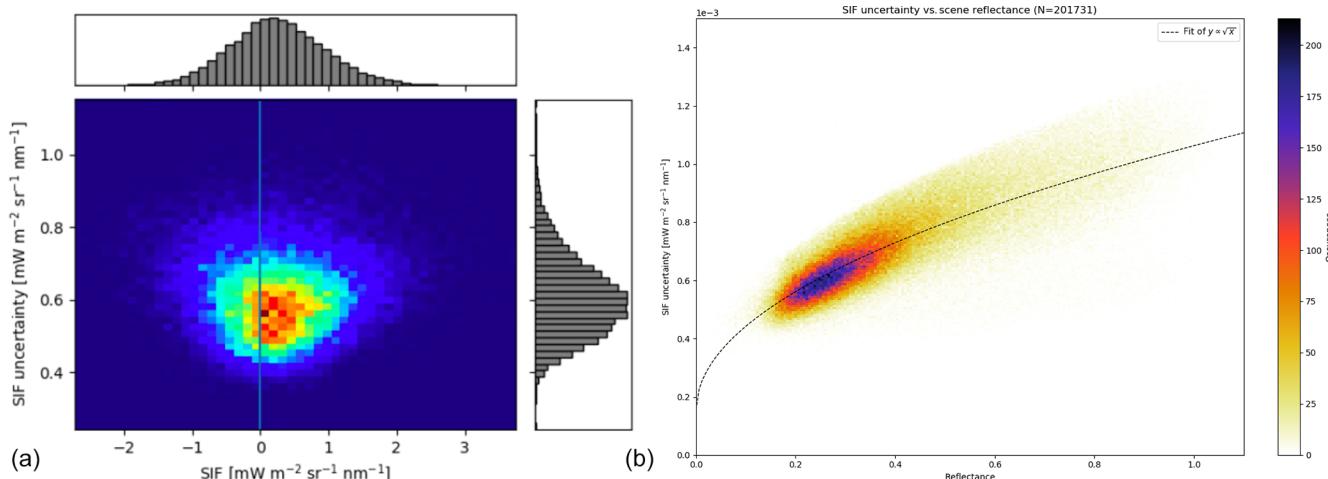
Figure 7. (a) Uncertainties in GOME-2A fluorescence retrievals as a function of signal strength on 1 August 2012. Only pixels over land (FRESCO+ albedo (at 758 nm) > 0.08) with cloud fractions < 0.4 have been selected. The colours indicate the number of occurrences of combinations of SIF and SIF uncertainty values. **(b)** SIF uncertainties as a function of scene reflectance for August 2012.

of the Southern Atlantic Anomaly. This stricter filtering of faulty retrievals also leads to higher SIFTER v2 fluorescence over the northwestern Amazon and lower values over eastern Brazil (Fig. 3.12 in van Schaik, 2016).

SIFTER v2 fluorescence shows more plausible values (close to zero) over the Northern Hemisphere landmasses in January and over the oceans in January and July. SIFTER v1 gives substantial positive fluorescence in absence of active photosynthesis there. The differences can be understood from the absence of a zero-level adjustment in SIFTER v1. The zero-level adjustment tends to reduce fluorescence at latitudes below 20° N (see Fig. 5). We also see substantial difference over the Himalayas and deserts. This is related to the negative fluorescence in SIFTER v2 in those regions.

5.2 Comparison between SIFTER v2 and NASA v2.8

We now evaluate the SIFTER v2 fluorescence retrievals against those from the most recent NASA suite (NASA v2.8). The NASA v2.8 retrieval algorithm uses the same 734–758 nm spectral fitting window as SIFTER v2, but the number of principal components (12 for NASA), the reference area (cloudy ocean and Sahara for NASA, Sahara for SIFTER v2), reference sector time span (daily for NASA, multi-year for SIFTER v2), and use of solar irradiance (daily for NASA, fixed Earth-Sun distance corrected for SIFTER v2) differ substantially between the two approaches. We first carried out a global pixel-to-pixel comparison of fluorescence values in the level-2 products for

15 January and 15 July 2011. Figure 8 shows scatter diagrams and histograms of the pixel-by-pixel differences between the SIFTER v2 and NASA v2.8. We find good correlation between the two retrievals ($R^2 = 0.45$ on 15 January and $R^2 = 0.61$ on 15 July 2011) and indications that SIFTER v2 is some 5 %–10 % lower than NASA v2.8 SIF.

Figure 8c and d show the probability distribution of the differences in fluorescence values between the SIFTER v2 and NASA v2.8 retrievals. The most frequently occurring differences are close to zero, and SIFTER v2 fluorescence is generally slightly lower than NASA v2.8 by $<$ 0.1 mW m^{-2} nm^{-1} sr^{-1}. On both days, the fluorescence differences resemble a Gaussian distribution function with a width of 0.4–0.5 mW m^{-2} nm^{-1} sr^{-1}. We interpret the width of the distribution as being caused by the combined independent SIFTER v2 (σ_S) and NASA v2.8 (σ_N) uncertainties, and if we assume that both retrievals are equally uncertain ($\sigma = \sqrt{\sigma_S^2 + \sigma_N^2}$), we arrive at estimates of $\sigma_S = 0.52$ mW m^{-2} nm^{-1} sr^{-1} on 15 January and $\sigma_S = 0.47$ mW m^{-2} nm^{-1} sr^{-1} on 15 July 2011. These estimates are consistent with the values discussed in Sect. 4.4 for SIFTER v2 and in Joiner et al. (2013) for NASA retrievals.

The spatial distribution of monthly mean SIF agrees well between SIFTER v2 and NASA v2.8, as shown in Fig. 9. Both retrievals show very similar patterns with the highest values over South America (Brazil, Bolivia, Argentina) in January, and over the US Corn Belt in July 2011. SIFTER v2 shows slightly negative SIF values over barren and elevated terrain (e.g. western China in January). These negative values can be explained by the choice of reference set as discussed in Botia (2017). SIFTER v2 has lower fluorescence signals than NASA v2.8 at high latitudes ($> 60°$ N) in July, which may also reflect the differences in reference dataset selected for the two retrievals.

We examined the agreement of the monthly mean SIF between our new SIFTER v2 products and the NASA v2.8 product for six vegetated regions across the globe. Within these regions, the two products agree to within 0.4 mW m^{-2} nm^{-1} sr^{-1} and often to within 0.2 mW m^{-2} nm^{-1} sr^{-1}. Both data products capture the seasonality of SIF. SIFTER v2 returns 15 %–35 % higher SIF values than NASA v2.8 in the evergreen equatorial regions in January, but 10 %–30 % lower SIF for all six regions in July (Table 5).

5.3 Impact of the degradation correction on SIF time series

Figure 10 shows time series of SIFTER v2 SIF values retrieved with and without the correction for throughput reductions for the six vegetated regions in Table 5. The corrections are applied on level-1 data retrieved from version 6.0 onwards, i.e. after June 2014. While the figure does not provide direct evidence that the SIFTER v2 data can be used

safely for long-term trend analysis, it does suggests that the level-1 degradation correction helps to stabilise our SIF retrievals in later years. A future long-term validation exercise could identify whether spurious trends exist in our SIFTER v2 retrievals.

6 Conclusions

We exploited end-to-end tests with the DISAMAR radiative transfer model to inform improvements for our new Sun-Induced Fluorescence of Terrestrial Ecosystems Retrieval version 2 (SIFTER v2) retrieval algorithm. The settings that optimise retrieval of accurate fluorescence strength are as follows.

– A smaller spectral fitting window (734–758 nm) compared to v1 (712–783 nm). The smaller window excludes the O_2-A band and most of the water vapour absorption features.

– A reduction in the number of principal components (PCs) used to describe atmospheric transmission effects to 10 (v2) from 35 (v1). Fewer PCs are needed to explain atmospheric transmittance since O_2-A and water vapour features no longer need to be explained in the fitting procedure.

– The calculation of the PCs reference set from all top-of-atmosphere spectra taken over the non-vegetated parts of the Saharan desert in the entire 2007–2012 period, when the quality of the GOME-2A level-1 data was stable, instead of from the 12 months preceding the GOME-2A measurement of interest.

– Application of a zero-level adjustment by correcting each retrieval within a 0.5° latitudinal bin with the mean retrieved fluorescence obtained over a Pacific Ocean reference sector within that bin. Such a zero-level adjustment was lacking in SIFTER v1.

The end-to-end tests indicated that the GOME-2A sensitivity to SIF (infilling) is much reduced within the O_2-A absorption bands compared to 740–758 nm, where sensitivity is close to 100 %. Spectral analysis of the infilling of the Fraunhofer lines over a wide spectral window with different sensitivities to SIF therefore becomes problematic. This can be diagnosed by spectral structure in the fit residuals being a strong indicator for unrealistic fluorescence estimates. Such faulty retrievals are effectively filtered out by requiring that spectral autocorrelation in the fitting residuals (or χ^2_{red}) stays below a certain threshold value.

With our improved SIFTER v2 algorithm we generated a multi-year global dataset of solar-induced fluorescence from GOME-2A. The fluorescence data are most reliable when retrieved under clear-sky conditions, when there is sufficient light coming from the Earth's surface, and for low spectral
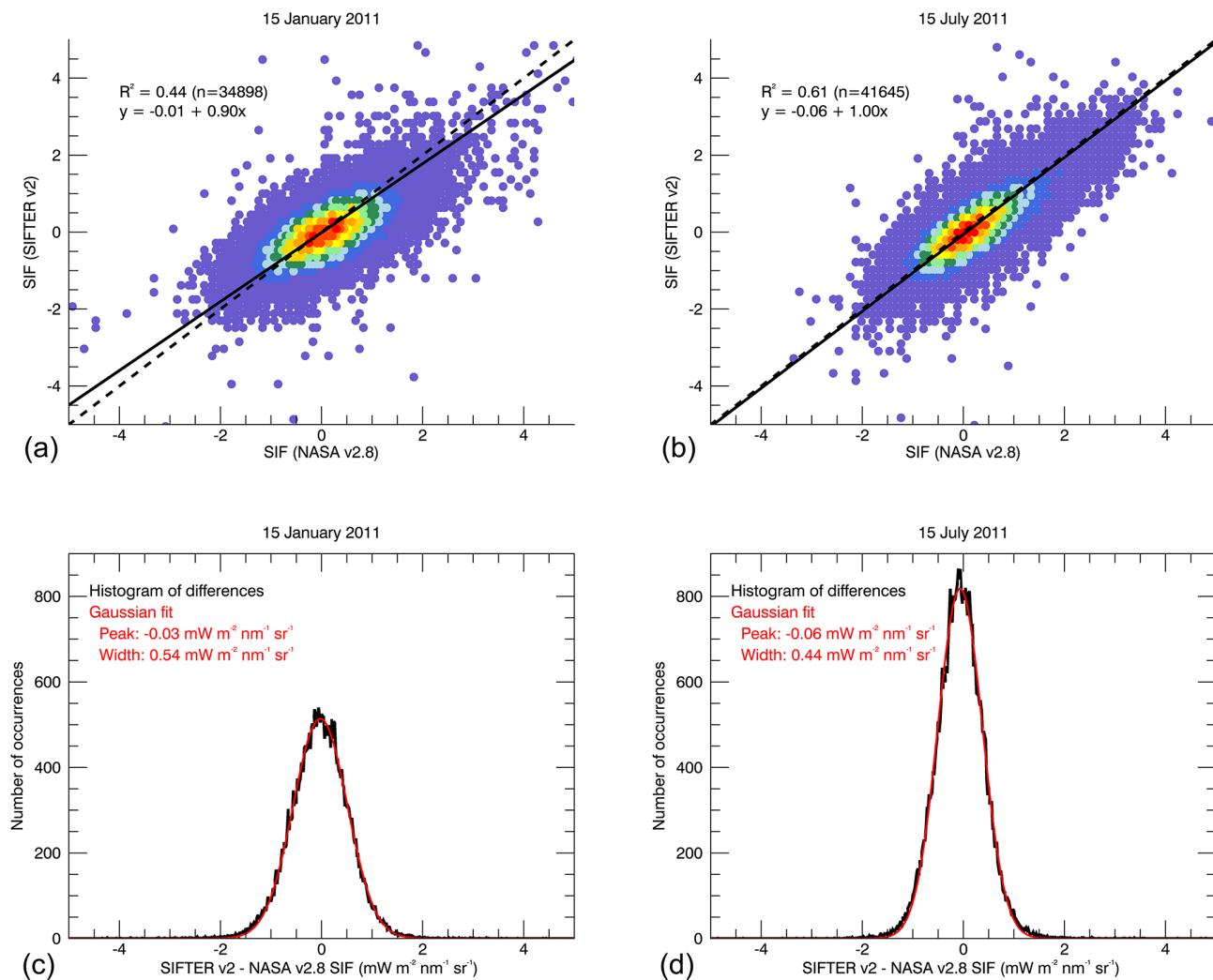
Figure 8. Scatterplots of SIFTER v2 vs. NASA v2.8 GOME-2A fluorescence retrievals (in $mW\,m^{-2}\,nm^{-1}\,sr^{-1}$) on 15 January and 15 July 2011. Only co-located pixels over land (albedo > 0.08) with (NASA) cloud fraction < 0.3 have been selected, and pixels were required to be valid in both retrievals. The colours indicate the density of occurrence of combinations of fluorescence values. The legend in **(a, b)** gives the result of a reduced major axis regression between the fluorescence retrievals.

autocorrelation in the fit residuals. We have the most confidence in the data quality for the period January 2007–July 2013, when the reflectance levels were stable and processed with one and the same level-1 processor version (5.3). After the detector switch in July 2013, we applied spectral correction factors to account for the degradation in the level-1 data that is apparent in later processor versions (v6.0 and beyond). Degradation correction has a stabilising effect on the SIF values retrieved since May 2014 with the introduction of the v6 processor versions. The SIFTER v2 (level-2) data product contains an uncertainty estimate associated with each valid retrieval. The uncertainty for individual retrievals is on the order of $0.6\,mW\,m^{-2}\,sr^{-1}\,nm^{-1}$ with fluorescence signal strengths typically between 0.0–$4.0\,mW\,m^{-2}\,sr^{-1}\,nm^{-1}$. The SIFTER v2 dataset shows the strongest fluorescence in tropical vegetated regions and in the agriculturally productive

regions in northern mid-latitudes. The SIFTER v2 product captures the seasonal variability in step with photosynthetic activity, with near-zero fluorescence in winter and high fluorescence in summer over northern mid-latitudes. The limitations of the GOME-2A fluorescence dataset when used for long-term trend analysis are in the quality of the level-1 data and the representativeness of the reference spectra used in the retrieval.

We evaluated SIFTER v2 fluorescence against our previous product and against the NASA v2.8 retrievals. Compared to SIFTER v1, SIFTER v2 returns higher fluorescence over strongly vegetated areas and shows a more pronounced seasonal cycle. The higher tropical fluorescence values in SIFTER v2 are the consequence of using a narrower fit window, a reference set based on a longer period (2007–2012) to capture a more realistic range of water vapour concentra-
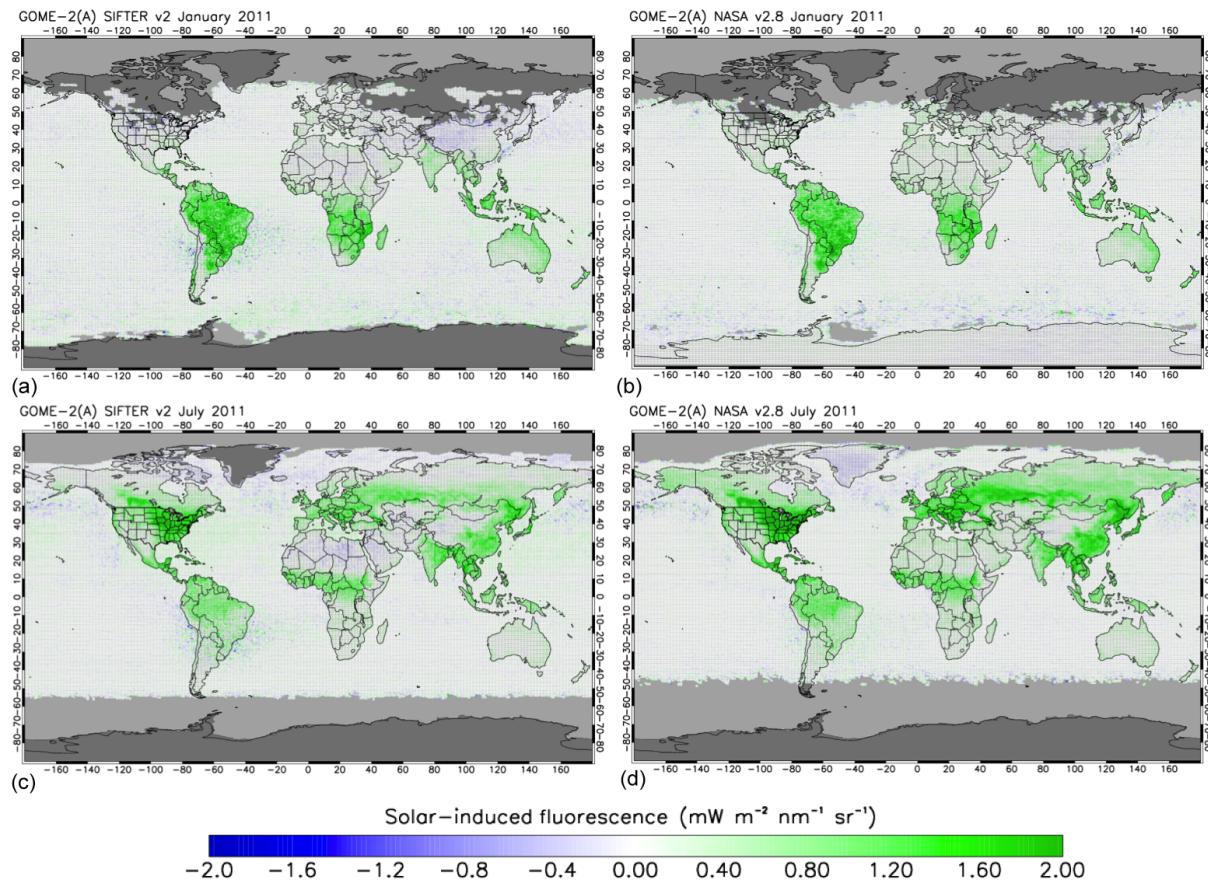
Figure 9. Gridded monthly mean SIF values retrieved from GOME-2A with SIFTER v2 (**a, c**) and NASA v2.8 (**b, d**) in January and July 2011. SIFTER data have been selected for autocorrelation < 0.2 and cloud fraction < 0.4, and NASA v2.8 data have been selected for cloud fraction < 0.3.

Table 5. Monthly mean gridded SIF values retrieved from GOME-2A with SIFTER v2 and NASA v2.8 for different vegetated regions throughout the world in January and July 2011.

	January 2011			July 2011		
	SIFTER v2 $(mW\ m^{-2}\ nm^{-1}\ sr^{-1})$	NASA v2.8 $(mW\ m^{-2}\ nm^{-1}\ sr^{-1})$	Relative difference	SIFTER v2 $(mW\ m^{-2}\ nm^{-1}\ sr^{-1})$	NASA v2.8 $(mW\ m^{-2}\ nm^{-1}\ sr^{-1})$	Relative difference
Amazon (0–15° S, 70–55° W)	1.47	1.27	+16 %	0.83	0.95	−13 %
Sub-Saharan Africa (5–10° N, 10° W–30° E)	0.21	0.31	N.A.	0.91	1.00	−9 %
Kalimantan (4° S–6° N, 110–115° E)	0.76	0.57	+34 %	0.49	0.58	−16 %
US Corn Belt (38–46° N, 96–81° W)	−0.09	0.07	N.A.	1.57	1.91	−18 %
Western Europe (44–52° N, 2° W–15° N)	0.07	0.19	N.A.	0.78	1.14	−32 %
Southeastern China (25–35° N, 100–120° E)	0.04	0.16	N.A.	1.10	1.37	−20 %

SIFTER v2 mean time series
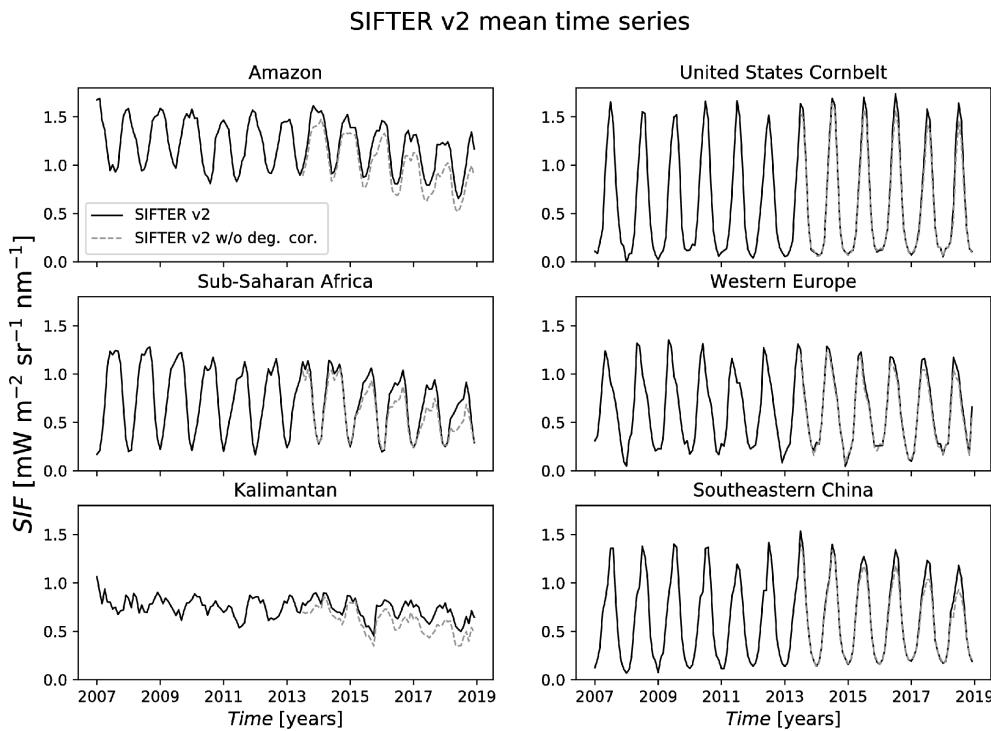
Figure 10. Time series of GOME-2A SIF retrieved with the SIFTER v2 algorithm with (solid black line) and without degradation correction (dashed grey line) over the six vegetated areas defined in Table 5. The degradation correction in our SIFTER v2 dataset is only applied for retrievals based on level-1 data retrieved from version 6.0, 6.1, and 6.2, i.e. from June 2014 onwards.

tions than in SIFTER v1, and application of a more strict spectral auto-correlation filter. The smaller window implies that SIFTER v2 is less prone to water vapour interference, while the 6-year reference set provides the retrieval with a more complete range of possible water vapour corrections. The more realistic near-zero fluorescence over the oceans and vegetated areas in wintertime are the result of the zero-level adjustment applied in SIFTER v2. NASA v2.8 and SIFTER v2 show a large degree of consistency in capturing spatial and temporal variability in fluorescence. Spatial correlation coefficients between the two level-2 products are high even on the day-to-day level (0.45–0.60). Both retrievals capture maximum fluorescence over the vegetated tropics and agricultural regions in the Northern Hemisphere in summer. In the evergreen tropics, SIFTER v2 fluorescence is 15 %–35 % higher than NASA v2.8, but over the summertime mid-latitudes SIFTER v2 is 10 %–20 % lower. Pixel-by-pixel differences between SIFTER v2 and NASA v2.8 retrievals allow evaluation of the combined uncertainties in the retrievals. These correspond to uncertainties of $0.5 \, \mathrm{mW \, m^{-2} \, sr^{-1} \, nm^{-1}}$ for both retrievals if we assume their uncertainties to be independent and contribute equally. This statistical estimate is slightly lower than the theoretical uncertainty estimate included in the data product.

Author contributions. EvS and KFB designed and performed the sensitivity tests, improved the SIFTER v1 retrieval approach, and developed the zero-level adjustment and auto-correlation filter. MLK drafted the algorithm description, performed the theoretical uncertainty analysis, processed the retrieval algorithm to the entire GOME-2A dataset, and analysed the results of retrieval settings. KFB analysed the degradation of the GOME-2A level-1 data, analysed the zero-level adjustment, designed the spectral degradation correction, and compared SIFTER v2 to SIFTER v1 and NASA v28 retrievals. KFB wrote the manuscript. EvS, MLK, and KFB contributed equally to this work. PS, ONET, and LGT supported the processing and analysis of GOME-2A level-1 data at KNMI. AFJS and WWV designed and developed the code and processed the data for SIFTER v1, which was used as the basis for SIFTER v2. RL and AC helped interpret the degradation in GOME-2A level-1 data. JJ and WP provided insight in retrieval sensitivities and suggestions

References

Azam, F. and Richter, A.: GOME-2 on MetOp-B Follow-on analysis of GOME2 in orbit degradation, Final Report EUM/CO/09/4600000696/RM, EUMETSAT, available at: http://www.iup.uni-bremen.de/doas/reports/gome2_degradation_follow_up_final_report.pdf (last access: 8 June 2020), 2015.

Baker N. R. and Oxborough K.: Chlorophyll Fluorescence as a Probe of Photosynthetic Productivity, in: Chlorophyll a Fluorescence. Advances in Photosynthesis and Respiration, vol. 19, edited by: Papageorgiou G. C., Govindjee, Springer, Dordrecht, https://doi.org/10.1007/978-1-4020-3218-9_3, 2004.

Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, Atmos. Meas. Tech., 10, 581–598, https://doi.org/10.5194/amt-10-581-2017, 2017.

Berrisford, Dee, D. P., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kållberg, P. W., Kobayashi, S., Uppala, S., and Simmons, A.: The ERA-Interim archive Version 2.0. ERA Report, ECMWF, Shinfield Park, Reading, available at: https://www.ecmwf.int/node/8174 (last access: 11 August 2020), 2011.

Botia, S.: Importance of Surface Pressure and Albedo for retrieving Sun Induced Chlorophyll Fluorescence from GOME-2, Internship Report, Royal Netherlands Meteorological Institute and Wageningen University, Wageningen, the Netherlands, 39 pp., 2017.

Chance, K. and Kurucz, R. L.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared, J. Quant. Spectrosc. Ra.,

111, 1289–1295, https://doi.org/10.1016/j.jqsrt.2010.01.036, 2010.

Daumard, F., Goulas, Y., Champagne, S. Ounis, A., Olioso, A., and Moya, I.: Continuous monitoring of canopy level sun-induced chlorophyll fluorescence during the growth of a sorghum field, IEEE T. Geosci. Remote, 50, 4292–4300, https://doi.org/10.1109/TGRS.2010.2046420, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., Valks, P., Hao, N., and van der A, R.: Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmos. Meas. Tech., 5, 2933–2949, https://doi.org/10.5194/amt-5-2933-2012, 2012.

Frankenberg, C. and Berry, J.: 3.10 – Solar Induced Chlorophyll Fluorescence: Origins, Relation to Photosynthesis and Retrieval, Comprehensive Remote Sensing, Elsevier, 143–162, https://doi.org/10.1016/B978-0-12-409548-9.10632-3, 2018.

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.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, Geophys. Res. Lett., 38, L17706, https://doi.org/10.1029/2011GL048738, 2011.

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

Grossi, M., Valks, P., Loyola, D., Aberle, B., Slijkhuis, S., Wagner, T., Beirle, S., and Lang, R.: Total column water vapour measurements from GOME-2 MetOp-A and MetOp-B, Atmos. Meas. Tech., 8, 1111–1133, https://doi.org/10.5194/amt-8-1111-2015, 2015.

Guanter, L., Alonso, L., Goìmez-Chova, L., Amoroìs, J., Vila, J., and Moreno, J.: Estimation of solar-induced vegetation fluorescence from space measurements, Geophys. Res. Lett., 34, L08401, https://doi.org/10.1029/2007GL029289, 2007.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., and Grainger, R. G.: Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements, Remote Sens. Environ., 121, 236–251, https://doi.org/10.1016/j.rse.2012.02.006, 2012.

Guanter, L., Aben, I., Tol, P., Krijger, J. M., Hollstein, A., Köhler, P., Damm, A., Joiner, J., Frankenberg, C., and Landgraf, J.: Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence, Atmos. Meas. Tech., 8, 1337–1352, https://doi.org/10.5194/amt-8-1337-2015, 2015.

de Haan, J. F.: DISAMAR Algorithm Description and Background Information, Royal Netherlands Meteorological Institute, De Bilt, the Netherlands, 2011.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, https://doi.org/10.5194/acp-3-1285-2003, 2003.

Hassinen, S., Balis, D., Bauer, H., Begoin, M., Delcloo, A., Eleftheratos, K., Gimeno Garcia, S., Granville, J., Grossi, M., Hao, N., Hedelt, P., Hendrick, F., Hess, M., Heue, K.-P., Hovila, J., Jønch-Sørensen, H., Kalakoski, N., Kauppi, A., Kiemle, S., Kins, L., Koukouli, M. E., Kujanpää, J., Lambert, J.-C., Lang, R., Lerot, C., Loyola, D., Pedergnana, M., Pinardi, G., Romahn, F., van Roozendael, M., Lutz, R., De Smedt, I., Stammes, P., Steinbrecht, W., Tamminen, J., Theys, N., Tilstra, L. G., Tuinder, O. N. E., Valks, P., Zerefos, C., Zimmer, W., and Zyrichidou, I.: Overview of the O3M SAF GOME-2 operational atmospheric composition and UV radiation data products and data availability, Atmos. Meas. Tech., 9, 383–407, https://doi.org/10.5194/amt-9-383-2016, 2016.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., Corp, L. A., and Middleton, E. M.: First observations of global and seasonal terrestrial chlorophyll fluorescence from space, Biogeosciences, 8, 637–651, https://doi.org/10.5194/bg-8-637-2011, 2011.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, Atmos. Meas. Tech., 6, 2803–2823, https://doi.org/10.5194/amt-6-2803-2013, 2013.

Joiner, J., Yoshida, Y., Guanter, L., and Middleton, E. M.: New methods for the retrieval of chlorophyll red fluorescence from hyperspectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY, Atmos. Meas. Tech., 9, 3939–3967, https://doi.org/10.5194/amt-9-3939-2016, 2016.

Khosravi, N., Vountas, M., Rozanov, V. V., Bracher, A., Wolanin, A., and Burrows, J. P.: Retrieval of Terrestrial Plant Fluorescence based on the in-filling of far-red Fraunhofer lines using SCIAMACHY Observations, Front. Environ. Sci., 3, 78, https://doi.org/10.3389/fenvs.2015.00078, 2015.

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

Köhler, P., Frankenberg, C., Magney, T. S., Guanter, L., Joiner, J., and Landgraf, J.: Global Retrievals of Solar-Induced Chlorophyll Fluorescence With TROPOMI: First Results and Intersensor Comparison to OCO-2, Geophys. Res. Lett., 45, 10456–10463, https://doi.org/10.1029/2018GL079031, 2018.

Kooreman, M. L., Boersma, K. F., van Schaik, E., van Versendaal, R., Cacciari, A., and Tuinder, O. N. E.: SIFTER sun-induced vegetation fluorescence data from GOME-2A (version 2.0) (Data set), Royal Netherlands Meteorological Institute (KNMI), https://doi.org/10.21944/gome2a-sifter-v2-sun-induced-fluorescence, 2020.

Koren, G., van Schaik, E., Arau, A. C., Boersma, K. F., Gartner, A., Killaars, L., Kooreman, M. L., Kruijt, B., van der Laan-Luijkx, I. T., von Randow, C., Smith, N. E., and Peters, W.: Widespread reduction in sun-induced fluorescence from the Amazon during the 2015/2016 El Nino, Philos. T. Roy. Soc. B, 373, 20170408, https://doi.org/10.1098/rstb.2017.0408, 2018.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1 km AVHRR Data, Int. J. Remote Sens., 21, 1303–1330, https://doi.org/10.1080/014311600210191, 2000.

Magney, T. S., Frankenberg. C., Köhler, P., North, G., Davis, T. S., Dold, C., Dutta, D., Fisher, J. B., Grossmann, K., Harrington, A., Hatfield, J., Stutz, J., Sun, Y., and Porcar-Castell, A.: Disentangling changes in the spectral shape of chlorophyll fluorescence: Implications for remote sensing of photosynthesis, J. Geophys. Res.-Biogeo., 124, 1491–1507, https://doi.org/10.1029/2019JG005029, 2019.

Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. A. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, Atmos. Meas. Tech., 8, 385–398, https://doi.org/10.5194/amt-8-385-2015, 2015.

Mohammed, G. H., Colombo, R., Middleton, E. M., Rascher, U., van der Tol, C., Nedbal, L., Goulas, Y., Pérez-Priego, O., Damm, A., Meroni, M., Joiner, J., Cogliati, S., Verhoef, W., Malenovský, Z., Gastellu-Etchegorry, G.-E., Miller, J. R., Guanter, L., Moreno, J., Moya, I., Berry, J. A., Frankenberg, C., and Zarco-Tejada, P. J.: Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress, Remote Sens. Environ., 231, 111177, https://doi.org/10.1016/j.rse.2019.04.030, 2019.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, Atmos. Meas. Tech., 9, 1279–1301, https://doi.org/10.5194/amt-9-1279-2016, 2016.

Neigh, C. S. R., McCorkel, J., and Middleton, E. M.: Quantifying Libya-4 Surface Reflectance Heterogeneity With WorldView-1, 2 and EO-1 Hyperion, IEEE Geosci. Remote S., 12, 2277–2281, https://doi.org/10.1109/LGRS.2015.2468174, 2015.

NIST/SEMATECH: e-Handbook of Statistical Methods, available at: http://www.itl.nist.gov/div898/handbook/, last access: 6 July 2018.

Parazoo, N. C., Frankenberg, C., Köhler, P., Joiner, J., Yoshida, Y., Magney, T., Sun, Y., and Yadav, V.: Towards a Harmonized Long-Term Spaceborne Record of Far-Red Solar-Induced Fluorescence, J. Geophys. Res.-Biogeo., 124, 2518–2539, https://doi.org/10.1029/2019JG005289, 2019.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., and Berry, J. A.: Linking chlorophyll *a* fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges, J. Exp. Bot., 65, 4065–4095, https://doi.org/10.1093/jxb/eru191, 2014.

Sanders, A., Verstraeten, W., Kooreman, M., van Leth, T., Beringer, J., and Joiner, J.: Spaceborne Sun-Induced Vegetation Fluorescence Time Series from 2007 to 2015 Evaluated with Australian Flux Tower Measurements, Remote Sens.-Basel, 8, 895, https://doi.org/10.3390/rs8110895, 2016.

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., and Magney, T.: Overview of Solar-Induced

chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP, Remote Sens. Environ., 209, 808–823, https://doi.org/10.1016/j.rse.2018.02.016, 2018.

van der Tol, C., Verhoef, W., and Rosema, A.: A model for chlorophyll fluorescence and photosynthesis at leaf scale, Agr. Forest Meteorol., 149, 96–105, https://doi.org/10.1016/j.agrformet.2008.07.007, 2009.

van Schaik, E.: Retrieving sun-induced fluorescence from the Global Ozone Monitoring Experiment-2, Internship Report, Royal Netherlands Meteorological Institute and Wageningen University, Wageningen, the Netherlands, 49 pp., 2016.

Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M.: FRESCO+: an improved O_2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8, 6565–6576, https://doi.org/10.5194/acp-8-6565-2008, 2008.

Zarco-Tejada, P. J., Berni, J. A. J., Suárez, L., Sepulcre-Cantó, G.,Morales, F., and Miller, J. R.: Imaging chlorophyll fluorescencewith an airborne narrow-band multispectral camera for vegetation stress detection, Remote Sens. Environ., 113, 1262–1275, https://doi.org/10.1016/j.rse.2009.02.016, 2009.