



Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements

Luis Guanter^{a,e,*}, Christian Frankenberg^b, Anu Dudhia^a, Philip E. Lewis^c, José Gómez-Dans^c, Akihiko Kuze^d, Hiroshi Suto^d, Roy G. Grainger^a

^a Atmospheric, Oceanic and Planetary Physics, University of Oxford, UK

^b Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

^c Department of Geography and National Centre for Earth Observation, University College London, UK

^d Japan Aerospace Exploration Agency, Tsukuba, Japan

^e Institute for Space Sciences, Free University of Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany

ARTICLE INFO

Article history:

Received 23 November 2011

Received in revised form 5 February 2012

Accepted 5 February 2012

Available online 7 March 2012

Keywords:

Fluorescence retrieval

GOSAT-FTS

Fraunhofer-line approach

Singular vector decomposition

Gross primary production

Vegetation indices

ABSTRACT

The recent advent of very high spectral resolution measurements by the Fourier Transform Spectrometer (FTS) on board the Greenhouse gases Observing SATellite (GOSAT) platform has made possible the retrieval of sun-induced terrestrial chlorophyll fluorescence (F_s) on a global scale. The basis for this retrieval is the modeling of the *in-filling* of solar Fraunhofer lines by fluorescence. This contribution to the field of space-based carbon cycle science presents an alternative method for the retrieval of F_s from the Fraunhofer lines resolved by GOSAT-FTS measurements. The method is based on a linear forward model derived by a singular vector decomposition technique, which enables a fast and robust inversion of top-of-atmosphere radiance spectra. Retrievals are performed in two spectral micro-windows (~2–3 nm width) containing several strong Fraunhofer lines. The statistical nature of this approach allows to avoid potential retrieval errors associated with the modeling of the instrument line shape or with a given extraterrestrial solar irradiance data set. The method has been tested on 22 consecutive months of global GOSAT-FTS measurements. The fundamental basis of this F_s retrieval approach and the results from the analysis of the global F_s data set produced with it are described in this work. Among other findings, the data analysis has shown (i) a very good comparison of F_s intensity levels and spatial patterns with the state-of-the-art *physically-based* F_s retrieval approach described in Frankenberg et al. (2011a), (ii) the overall good agreement between F_s annual and seasonal patterns and other space-based vegetation parameters, (iii) the need for a biome-dependent scaling from F_s to gross primary production, and (iv) the apparent existence of strong directional effects in the F_s emission from forest canopies. These results reinforce the confidence in the feasibility of F_s retrievals with GOSAT-FTS and open several points for future research in this emerging field.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

The terrestrial sun-induced chlorophyll fluorescence signal (F_s) is emitted by the photosystem II of the chlorophyll molecules of assimilating leaves: part of the energy absorbed by chlorophyll is not used for carbon fixation, but re-emitted at longer wavelengths in the 650–800 nm spectral region (Baker, 2008; Papageorgiou, 1975). The F_s signal originates at the cores of the photosynthetic machinery and responds instantaneously to perturbations in the environmental conditions such as light and water stress. This makes it a more direct proxy to the plant photosynthetic activity than the reflectance-based vegetation indices traditionally used in vegetation remote sensing applications (Coops et al., 2010). The very recent publication of the first

global maps of chlorophyll fluorescence (Frankenberg et al., 2011b; Joiner et al., 2011) opens the door to a new, unexplored field of research that couples remote sensing with vegetation and atmospheric sciences.

The F_s retrieval methods developed by Joiner et al. (2011) and Frankenberg, et al. (2011a) share a common fundamental basis: the fractional depth of solar Fraunhofer lines decreases due to *in-filling* by F_s . The feasibility of F_s retrieval using individual Fraunhofer lines had been proposed some years ago (Sioris et al., 2003). However, it has not been until the advent of the Greenhouse gases Observing SATellite “IBUKI” (GOSAT), carrying the Thermal And Near infrared Sensor for carbon Observation (TANSO) (Kuze et al., 2009), that the first global F_s maps have been produced. TANSO includes a Fourier Transform Spectrometer (FTS) measuring with high spectral resolution in the 755–775 nm range, which allows to resolve individual Fraunhofer lines overlapping the F_s emission.

* Corresponding author at: Institute for Space Sciences, Free University of Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany.

E-mail address: guanter@atm.ox.ac.uk (L. Guanter).

The research presented in this contribution covers a wide range of aspects related to both the retrieval of F_s from GOSAT-FTS measurements and the analysis of the F_s signal on a global basis. Key points addressed in this work are:

- (1) A new approach for F_s retrieval from GOSAT-FTS data is proposed. This method is intended to simplify the implementation and to minimize the computation time with respect to the one proposed by Frankenberg et al. (2011a), which is taken as a reference for the reasons described later in this paper.
- (2) The global maps of F_s produced with this method are compared to those by Frankenberg et al. (2011b) and Joiner et al. (2011) in order to establish an agreement in the range of F_s values to be expected for the temporal and spatial scales accessible from GOSAT measurements.
- (3) The F_s data set produced with the proposed method after the processing of 22 months of GOSAT-FTS data is used to investigate intrinsic properties of the F_s signal and the sensitivity of the retrievals against atmospheric and instrumental factors.
- (4) The potential and limitations of the retrieved F_s maps for the assessment of the vegetation condition is further studied by means of the comparison with other satellite-based vegetation parameters such as vegetation indices and gross primary production products.

The rest of this paper is organized as follows: an overview of the state-of-the-art in F_s retrieval from GOSAT data is provided in Section 2. The F_s retrieval approach and the data sets used in this work are described in Section 3. Results from F_s retrieval and the comparison with other products are presented in Section 4. Finally, a summary of the most relevant findings is given in Section 5.

2. Background

The main challenge in F_s retrieval from passive measurements is to decouple the F_s signal from the solar radiation reflected by vegetation and atmosphere, which can be about 100–150 times more intense than F_s at the top-of-atmosphere (TOA) level (Meroni et al., 2009). Spectrally-resolved measurements of back-scattered sunlight in solar Fraunhofer lines (Sioris et al., 2003) or in telluric absorption bands (Guanter et al., 2007) overlapping the F_s emission can be used to separate the F_s -emitted signal from the solar-reflected radiation. According to their spectral position, the atmospheric O₂-A and O₂-B absorption bands in 760.5 nm and 687.5 nm and the Fraunhofer lines around 757.5 and 770.1 nm have potential for F_s retrieval.

In the case of F_s retrieval in atmospheric absorption bands, the different atmospheric optical path crossed by the F_s -emitted and solar-reflected signals provides the information required to decouple those two components from the radiance measured at the TOA level. F_s retrieval requires then a very rigorous formulation of the surface-atmosphere radiative transfer which interposes a very complex inversion problem (Guanter et al., 2010). As a counterpart, the relatively coarser spectral resolution needed to resolve atmospheric telluric bands (~0.1 nm, as compared to >0.05 nm for solar Fraunhofer lines) enables instrument designs sampling a wider spectral window. The retrieval problem is based in this case on the multi-parameter inversion of F_s and correlated atmospheric and vegetation variables. This approach based on medium spectral resolution measurements (of the order of 0.1 nm) with a wide spectral window (around 200–300 nm) is the basis of the instrument concept of the FLEX mission currently under development phase A/B1 of the 8th ESA Earth Explorer Program. The main advantage of this retrieval scheme is that (a) the entire fluorescence broadband spectrum can be retrieved, and (b) F_s and other vegetation parameters, such as leaf area index and

chlorophyll content, can be derived consistently from the same TOA radiance spectrum. However, although feasible according to simulations (Guanter et al., 2010) and ground-based measurements (Meroni & Colombo, 2006), this type of wide-band F_s retrieval approach has not yet been demonstrated on real spaceborne data.

On the other hand, F_s retrieval in solar Fraunhofer lines relies on the evaluation of the *in-filling* of the Fraunhofer lines due to F_s . Since the fractional depth of Fraunhofer lines is not affected by atmospheric scattering and absorption in narrow spectral windows free from telluric absorption features, the atmospheric modeling required is much simpler than with atmospheric bands. The main shortcoming is that a spectral resolution better than 0.05 nm is needed to resolve the Fraunhofer lines for F_s retrieval (Frankenberg et al., 2011a). Taking advantage of GOSAT-FTS high resolution measurements, the existing F_s retrieval methods from Joiner et al. (2011) and Frankenberg et al. (2011a, b) exploit Fraunhofer lines as a basis for F_s retrieval. Both methods make use of an extraterrestrial solar irradiance spectrum as a reference to quantify the filling-in of the Fraunhofer lines and convert this into an F_s estimation. This set-up is illustrated in Fig. 1. A real top-of-atmosphere radiance spectrum measured by TANSO-FTS in band-1 is plotted together with an extraterrestrial irradiance spectrum from Chance and Kurucz (2010) and with a real F_s measurement. The F_s spectrum was measured with a FieldSpec FR spectroradiometer (ASD Inc., Boulder CO, US) (Amorós-López et al., 2008). Spectral shifts in FTS measurements with respect to the solar irradiance data set can be observed, which can be explained by both spectral calibration errors in the FTS and by Doppler shifts due to both the satellite-Earth motion for off-nadir observation geometries and the sun-Earth relative velocity, which depends on the sun position. The F_s curve in this figure corresponds to the right wing (when displayed in wavelengths) of the second peak of a typical F_s spectrum.

The method described in Joiner et al. (2011) makes use of one Fraunhofer line, which is the strong K I line around 770.1 nm. It employs real solar irradiance measurements from the TANSO-FTS in order to avoid the explicit modeling of the instrument line shape function (ILSF). In turn, the method proposed in Frankenberg et al. (2011a) extends this single-line approach to two broader spectral windows centered in 755 and 770 nm. This retrieval, making use of broader spectral windows containing several Fraunhofer lines, is expected to be less sensitive to instrumental noise than that based on one single line, while performing the retrievals for the two separate windows provides more independent measurements to be used to enhance the signal-to-noise ratio of the final F_s products. Critical steps of this method are the explicit modeling of the asymmetric ILSF of the FTS, the use of an *in-house* high accuracy solar transmission model, and the consideration of spectral shift and stretch as part of the state vector to be inverted from FTS measurements. These characteristics lead to a non-linear retrieval problem which has to be solved iteratively. The retrieval is performed independently for 4 fitting windows, which correspond to the two spectral windows in Fig. 1 and the two polarizations the FTS is measuring (S and P). Processing the two polarizations independently is useful to evaluate potential issues related to one or the other polarization. Retrievals from the two polarizations are combined in the final F_s product.

The fundamental assumption of the method proposed in this work is that FTS measurements in the 4 fitting windows can be expressed as a linear combination of a series of orthogonal spectral functions, which allows a linearization of the forward model. If valid, the forward model can be inverted by non-weighted linear least squares. This makes unnecessary (i) the modeling of the ILSF, (ii) the use of external extraterrestrial irradiance data sets, and (iii) the iterations in the minimization of the cost function required for spectral alignment as described by Frankenberg et al. (2011a).

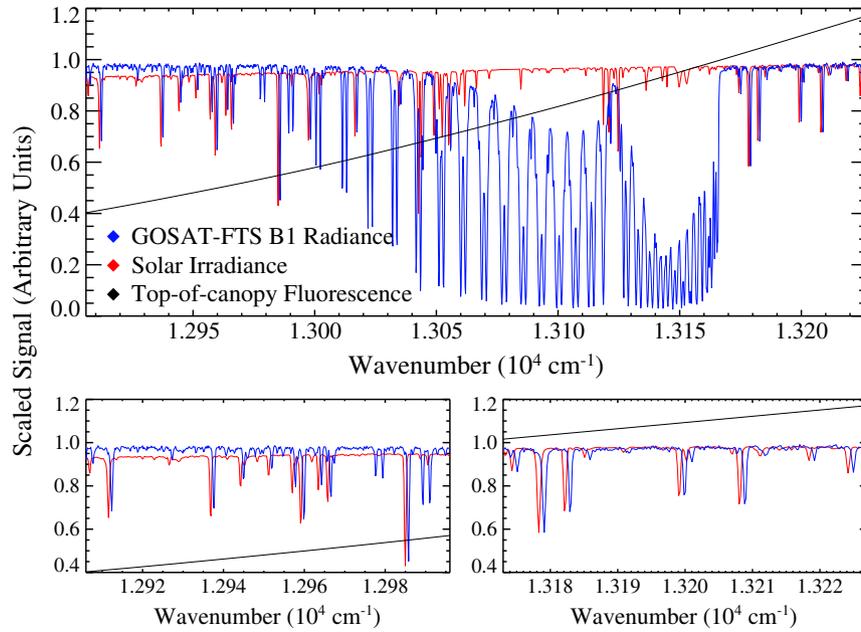


Fig. 1. Normalized radiance spectrum measured by GOSAT-FTS together with reference extraterrestrial solar irradiance and top-of-canopy fluorescence spectra. The two spectral micro-windows used for F_s retrieval in this work (770 nm on the bottom left, 755 nm on the bottom right) are magnified in the lower panels.

3. Materials and methods

3.1. The forward model: singular vector decomposition

The underlying idea for the proposed F_s retrieval approach is that any radiance spectrum can be modeled as a linear combination of a F_s -free radiance spectrum plus the F_s component propagated to the top-of-the-atmosphere. The other assumption is that the F_s -free radiance spectrum can, in turn, be expressed as the linear combination of the singular vectors describing a F_s -free set of measurements. The singular vector decomposition (SVD) technique (Press et al., 2007; Rodgers, 2000) has been chosen to generate the basis of spectral functions used to model radiance spectra in the two spectral windows. The SVD technique, related to principal component analysis, is a powerful tool for the solution of linear systems, and is widely used in statistics to analyze and reduce the dimensionality of large data sets. This makes it also very suitable for atmospheric remote sensing applications (e.g. Hurley et al., 2009; Klüser et al., 2011).

SVD is based on the transformation of a large set of correlated variables into a smaller set of uncorrelated signals called singular vectors. The basis of singular vectors can be arranged so that the amount of signal variability each vector accounts for decreases monotonically. The singular vector decomposition of an $m \times n$ matrix \mathbf{M} (in our case, m spectra $\times n$ spectral points) is

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (1)$$

where \mathbf{U} is an $n \times m$ matrix of left singular vectors, $\mathbf{\Sigma}$ is defined to be an $n \times n$ diagonal matrix composed of the singular values, and \mathbf{V} is an $n \times n$ unitary matrix containing the right singular vectors in its n columns. T is the transpose operator.

Using the concept of singular vectors, the forward model \mathbf{F} consisting of the F_s -free plus the TOA F_s contribution can be written as

$$\mathbf{F}(\omega, F_s) = \sum_{i=1}^{n_v} \omega_i v_i + F_s^{\text{TOA}} I, \quad (2)$$

where ω_i is the weight of the singular vector v_i , F_s^{TOA} is the intensity in radiance units of the fluorescence signal at the top-of-atmosphere, I is an identity vector of size n , and n_v is the number of singular vectors

used to model the input radiance spectrum. The fit coefficients are then the n_v weights ω_i and F_s^{TOA} . Concerning n_v , in view of the percentage of captured variance by each singular vector, a number n_v of 4 and 5 singular vectors has been selected for the forward modeling of the 755 and 770 nm windows, respectively. This particular number of singular vectors has been chosen with an empirical threshold of 0.05% on the percentage of the variance explained by each singular vector (see Fig. 3). This number is a compromise between the degree to which the input radiance can be reconstructed and a safe limit to avoid data over-fitting.

It must be noted that the strongest atmospheric features in the 770 nm window are removed from the inversion due to the fact that no physical formulation of atmospheric absorption and scattering is included in the forward model. Including the explicit modeling of atmospheric absorptions would complicate the formulation without an expected improvement in the retrieval accuracy. Another aspect to be pointed out is that the input radiance spectra are normalized by their slope (divided by a straight line with the same slope) so that the spectra used in the inversion are as spectrally-flat as possible. The same is done with the training spectra used to calculate the singular vectors. The purpose of this normalization is to avoid retrieval errors from spectra with spectral slopes not represented by those of the surfaces covered by the training set. For example, snow surfaces may present very pronounced spectral slopes within the 760 nm spectral window. Due to this normalization, no spectral slope is assumed for the F_s signal. It must be acknowledged, though, that this normalization has generally a very small effect on the retrievals.

Eq. (2) is inverted by linear least squares without spectral weights, as white noise can be assumed for an FTS. Under this assumption, the retrieval error covariance S_e is given by

$$S_e = \sigma_m^2 (\mathbf{J}^T \mathbf{J})^{-1}, \quad (3)$$

where S_e is a matrix of $(n_v + 1) \times (n_v + 1)$, σ_m is the measurement error, and \mathbf{J} is the matrix consisting of the first n_v singular vectors and I . Following Frankenberg et al. (2011b), σ_m is estimated as the standard deviation of the first and last 500 spectral samples of each FTS band-1 spectrum. These samples are considered to be completely out of the band pass filter and can then provide a good estimation of

instrumental noise. It must be noted that Eq. (3) gives only an estimation of how instrumental noise propagates to F_s values. Potential error sources such as rotational Raman scattering, residual atmospheric contamination or radiometric offsets are not considered in S_e . In particular, rotational Raman scattering and the Ring effect (Vountas et al., 1998) are neglected following the analysis provided by Joiner et al. (2011). Concerning atmospheric scattering, since the two fitting micro-windows are selected so that strong atmospheric absorption features are avoided, atmospheric extinction for the surface-to-sensor path is expected to be small. This allows one to assume F_s^{TOA} equals top-of-canopy F_s , as it is discussed in (Joiner et al., 2011). This effect is further investigated in Section 4.2.

A training set \mathbf{M} of non-vegetated land targets has been compiled in order to generate the set of singular vectors \mathbf{V} . The goal is to produce a training data set which covers the widest span of radiance levels and atmospheric conditions. Vegetation-free spectra are selected according to the following criteria: (1) land surfaces, (2) acquired in the high-gain mode of the FTS, (3) radiance levels are in the range $[1.5 - 8] \cdot 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$, (4) spectra are over Antarctica or Greenland, or fulfill ($R_{760}^n < R_{1600}^n$ or $R_{1600}^n < R_{2100}^n$), with R_λ^n the TOA reflectance in the continuum of any of the three FTS bands. This latter condition is very effective to discriminate bare soils from any other surface (vegetation, water, cloud or snow). Snow surfaces in Antarctica and Greenland are added to increase the range of radiance levels in the data set. In order to capture the maximum variability in illumination and atmospheric conditions, spectra have been extracted from global data in 3 days (1, 15, 30) of 4 different months (July and December of 2009 and 2010). This has led to \mathbf{M} consisting of a total of 9513 spectra. The global distribution of these spectra is displayed in Fig. 2. Large fractions of desert areas are not included in the training set because the spectra are acquired under the medium-gain mode of the FTS (condition 2). On the other hand, Fig. 2 shows that spectra from areas which might contain vegetation are included in the training set. Assuming condition 4 is very strict in discriminating bare soil surfaces from the rest of land covers, spectra from those areas must come from a month (either in July or December) for which no vegetation is present.

3.2. Correction of systematic errors in F_s from FTS radiometric offsets

A systematic error in F_s associated to an apparent nonlinear performance of the TANSO-FTS band-1 is reported in Frankenberg et al. (2011b). This error is easily detected over Antarctica, where a positive bias in F_s reproduces spatial patterns of at-sensor radiance. Such an offset might be due to a detector signal non-linearity which can cause a zero level intensity offset in the spectrum (Abrams et al., 1994). In fact, many FTS's have detector non-linearity problems. Laboratory and simulation tests show that the TANSO-FTS band-1 non-

linear performance can be explained as an analog circuit non-linearity. Frankenberg et al. (2011b) propose an empirical correction of the resulting F_s offsets by calibrating them as a function of the average at-sensor radiance over Antarctica.

The same idea has been followed in this work for the quantification and correction of F_s biases correlated to the average radiance in the FTS band-1. The main difference with respect to the methodology in Frankenberg et al. (2011b) is that not only data over Antarctica have been used, but also any vegetation-free spectrum, which includes ocean and cloud-contaminated spectra. In this way, the amount of data available to generate the F_s -radiance correction curves is substantially increased, which enables the calculation of the correction curves on a monthly basis. This is in turn useful in the determination of possible temporal changes in the instrument response, such as the degradation of the radiometric performance. A similar approach has recently been proposed by Joiner et al. (2012). The criteria for the selection of vegetation-free spectra are in this case (1) acquired in high-gain mode, (2) $R_{760}^n > 0.5$ or $R_{760}^n < R_{1600}^n$ or $R_{1600}^n < R_{2100}^n$ or acquisition over ocean, (3) $|F_s| < 5 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$, and (4) at-sensor radiance levels in the range $[1.5 - 8] \cdot 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$. The average radiance in the FTS band-1 is calculated from the spectral window $[12936 - 13245] \text{ cm}^{-1}$ ($[755 - 773] \text{ nm}$), which is the spectral range assumed to be covered by the FTS band pass filter. The average radiance range $[1.5 - 8] \cdot 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$ is divided into 500 radiance levels into which the F_s offsets for a given radiance interval are binned. The resulting F_s -offset versus average radiance curves are finally low-pass filtered to reduce noise.

The filtering criteria for the selection of vegetation-free spectra are considered strict enough to prevent the reference spectra to be affected by residual F_s contamination. Having F_s -affected spectra in the generation of the F_s -radiance correction curves would lead to an over-correction of the F_s errors. Results from the correction of the errors in F_s due to radiometric offsets are described in an Appendix of this paper.

3.3. Production of global F_s maps

Global maps of F_s are generated from the offset-corrected F_s retrievals by spatial and temporal averaging of all successful retrievals. Following Frankenberg et al. (2011b) and Joiner et al. (2011), the F_s retrievals are binned as monthly averages in $2^\circ \times 2^\circ$ bins. The quality filtering criteria are as follows:

- (1) Average radiance $\in [1.5 - 8] \cdot 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$
- (2) $|F_s| < 5 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$
- (3) Out-of-band signal $< 1\%$
- (4) FTS high radiometric gain
- (5) 100% land pixels
- (6) Sun zenith angle $< 65^\circ$
- (7) $\chi_r^2 < \overline{\chi_r^2} + 0.15$

where χ_r^2 is the reduced χ^2 of the fit. These thresholds are expected to remove all inaccurate retrievals, including spectra affected by severe cloud contamination. The mean value $\overline{\chi_r^2}$ from all the retrievals is used instead of 1 to discriminate bad quality retrievals because it has been observed that fits in 755 nm have a $\overline{\chi_r^2}$ around 1.1 for both polarizations, which may be explained by a systematic underestimation of noise by the out-of-filter signal approach. $\overline{\chi_r^2}$ is around 1 for the 770 nm window.

Error-weighted mean F_s values are calculated for each cell following

$$\bar{F}_s = \frac{\sum_{i=1}^{n_f} (F_{s,i} / \sigma_{e,i}^2)}{\sum_{i=1}^{n_f} (1 / \sigma_{e,i}^2)} \quad (4)$$

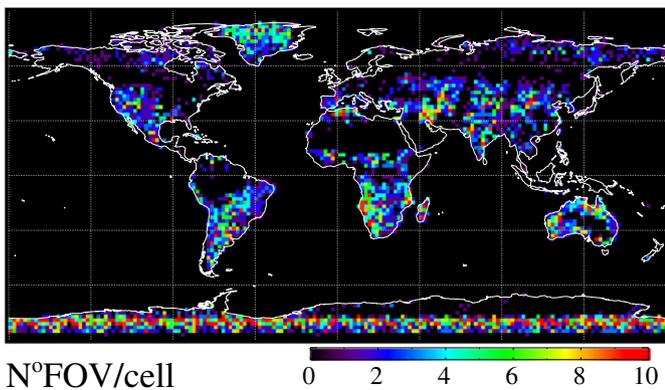


Fig. 2. Spatial distribution of the 9513 spectra in the training data set.

where σ_e is the single-retrieval error derived from Eq. (3) and n_f refers to the number of spectra that passed the quality assurance filtering criteria. The variance of the weighted mean is then given by

$$\sigma_{\bar{F}}^2 = \frac{1}{\sum_{i=1}^{n_f} (1/\sigma_{e,i}^2)}, \quad (5)$$

which is adopted as the square of the standard error to be associated to the mean value in each of the $2^\circ \times 2^\circ$ bins.

In principle, retrievals from all 4 fitting windows (2 spectral micro-windows, 2 polarizations) are to be treated as independent measures to calculate the temporal and spatial averages used to produce the global F_s maps. However, F_s in 755 nm is expected to be about two times higher than in 770 nm (see Fig. 1), and the retrievals in 770 nm might be affected by residual O_2 features not considered in the forward model. We can therefore expect that the SNR of the F_s composites would not be improved significantly by adding the lower F_s values from the 770 nm retrievals, which could in turn be affected by some systematic errors as those coming from residual atmospheric scattering. For these reasons, only retrievals in 755 nm are used to generate the final maps.

3.4. GOSAT-FTS data processed for this work

A data set consisting of level 1B GOSAT-FTS measurements in the time period between June 2009 and March 2011 (about 2.5 TB of data) has been processed to test the F_s retrieval algorithm described in this work. The band-1 of the GOSAT-FTS covers the 755–775 nm spectral window with a spectral resolution of around 0.025 nm. FTS level 1B data are delivered as unapodized calibrated spectra in two perpendicular polarizations, S and P. Each field-of-view covers a circular area with a diameter of about 10.5 km at ground. GOSAT-TANSO is a joint project of the Japan Aerospace Exploration Agency (JAXA), the Japanese National Institute for Environmental Studies (NIES), and the Japanese Ministry of the Environment (MOE).

The complete temporal series of FTS band-1 acquisitions consists of 4 different time periods, each one corresponding to a different version of the FTS level 1B product, labeled as 050050, 080080, 100100 and 110110. Differences between the 4 different sub-products, mostly linked to changes in the Level-0 processing chain, are considered negligible in terms of F_s retrieval, so the entire temporal series is treated as a single, continuous data set. FTS level 1B spectra are calibrated to radiance units using the standard FTS calibration factors and the detector degradation curves provided in the GOSAT User Interface Gateway.¹ Of the two FTS radiometric gain modes, only data acquired under the high gain mode have been analyzed in this work, as the medium gain mode is used only over bright desert areas.

3.5. Remote sensing-based vegetation products used in this work

The assessment of the F_s maps produced has been carried out by comparison with other satellite-based vegetation data sets. These comprise:

- (1) F_s scaled at 755 nm derived with the approach described in Frankenberg et al. (2011b).
- (2) Gross primary production (GPP) data sets from the MODIS MOD17A2 product (Running et al., 2004) and from the Biogeochemical Model-Data Integration Group of the Max Planck Institute for Biogeochemistry (MPI-BGC) (Jung et al., 2011).
- (3) Reflectance-based vegetation indices, including the normalized difference vegetation index and the enhanced vegetation index (NDVI and EVI, respectively) from the MODIS

MOD13C2 product (Huete et al., 2002), and the MERIS terrestrial chlorophyll index (MTCI) (Dash & Curran, 2004).

As stated earlier in this paper, the *physically-based* F_s retrieval approach described in Frankenberg et al. (2011a) is used as a reference F_s retrieval method for comparison in this work. Comparing the two F_s retrieval approaches is not only intended for validating the SVD-based approach presented here, but also for further consolidation of global F_s values and their spatial and temporal patterns.

Concerning the GPP products, vegetation GPP represents the conversion of chemical energy into biomass through photosynthesis, and is therefore the parameter which has a priori the most direct relationship to F_s . Both the MODIS and the MPI-BGC GPP data sets rely on satellite-based estimates of fraction of absorbed photosynthetically active radiation (FAPAR) for their modeling of GPP, but are fundamentally different. MODIS GPP is derived by the combination of tabulated light use efficiency with meteorological parameters and FAPAR, whereas the MPI-BGC GPP data set is produced by the global upscaling of site measurements of carbon dioxide, water and energy fluxes.

Reflectance-based vegetation indices, widely used by the remote sensing community, provide information about vegetation greenness (i.e. a combination of biomass, chlorophyll content and structural effects) which can be linked to canopy light interception. In particular, the MTCI has been designed to provide a high sensitivity to chlorophyll content, which should also be closely related to the chlorophyll fluorescence signal retrieved from GOSAT data. These vegetation indices can provide a direct indication of the link between F_s and vegetation phenology.

Finally, the temporal dimension of the F_s signal has been explored by means of the comparison of temporal series of F_s and GPP at given biomes with different vegetation cover types. For this analysis, the land cover classification scheme described in Frankenberg et al. (2011b) has been adopted. This is based on the IGBP land cover biome classification from the MODIS MCD12C1 product. Some classes are aggregated (e.g. *evergreen needleleaf forest* and *deciduous needleleaf forest* into *needleleaf forest*), and a separation between northern and southern hemisphere is done for some biomes. Monthly means of F_s , GPP and EVI have been calculated for each of the resulting classes.

4. Results

4.1. Assessment of the singular vectors

The first 8 v_i from the P polarization are plotted in Fig. 3. Very similar sets of vectors are obtained for the S polarization. The removal of the stronger O_2 features in the 770 nm window can be seen by comparison with Fig. 1. The two sets of O_2 features at both sides of the strongest Fraunhofer line in 770.1 nm do not appear in the first v_i s, although residuals can be seen in the higher order ones.

It can be noted that there is a comparable sequence of vectors for the two spectral windows. The first vector reproduces TOA radiance, determined mostly by surface albedo and illumination angle, whereas the second vector carries most of the information for the spectral alignment of FTS measurements. The third vector in 755 nm mostly accounts for spectral slopes in the radiance continuum, whereas the fourth one might reproduce a spectral band broadening. For the 770 nm window, the spectral slope information appears in both the fourth and fifth vectors, whereas the spectral broadening is in the third vector. Concerning the spectral slope piece of information, some residual spectral features with different sign are visible in the fourth and fifth vectors of the 770 nm window, which is not the case of the 755 nm window. The spectral slope information appears in spite of the slope normalization performed on the training set spectra, although it has been observed that it accounts for a smaller variance than when no slope normalization is performed. It can also be

¹ <https://data.gosat.nies.go.jp/gateway/gateway/MenuPage/open.do>.

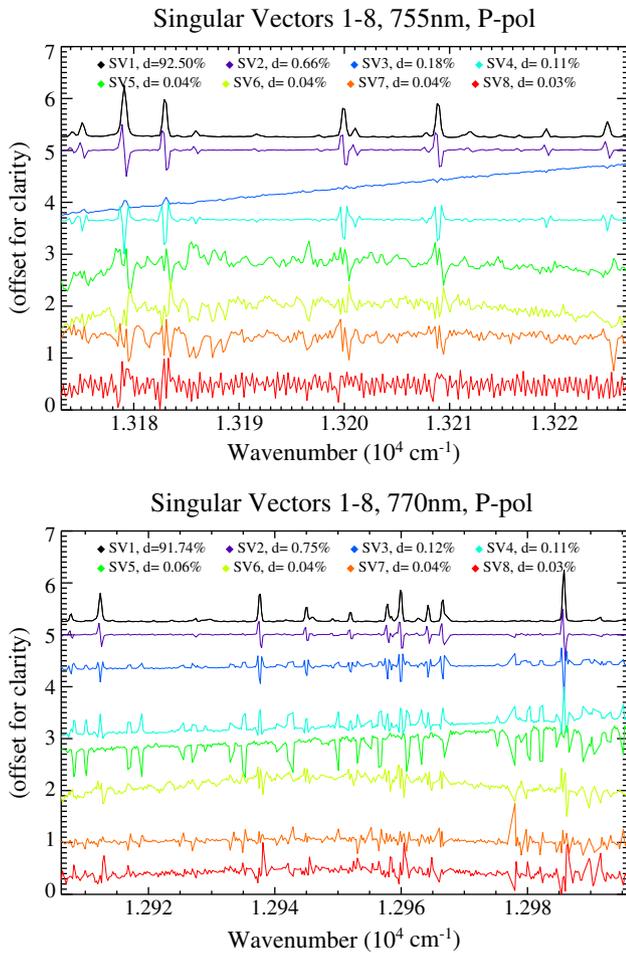


Fig. 3. First 8 singular vectors from the P polarization. Percentages in the legend refer to the relative contribution of each v to the total variance in \mathbf{M} . The first 4 (5) singular vectors have been selected for the forward modeling of the 755 nm (770 nm) window.

seen that radiance slopes are expected to have a higher variability in 755 than in 770 nm.

It could be argued that the spectral slope could be mistaken by F_s in the inversion. However, this is not observed in the results. This can be explained by the very different slope of the typical F_s emission and the third vector in 755 nm, and by residual features in the fourth vector of 770 nm. With respect to the potential spectral broadening detected, given that the solar nature of the Fraunhofer lines prevents them to be affected by pressure broadening, our current best understanding is that this is an instrumental effect: the digitization (16 bit analog-to-digital converting) causes the spectral broadening for low input signals. Only FTS band-1 is alternating-current sampled, and at the edge of the interferogram the level is almost zero so information with low number of digitization is lost. In the case of low input signals, the number of digitization bits is not enough and it behaves like apodization. It is then not a random but a systematic effect. It is observed that it has a very small contribution to the total variance.

4.2. Atmospheric effects on F_s retrievals

It has been discussed in Section 3.1 that the effect of atmospheric scattering on each single top-of-atmosphere F_s retrieval is expected to be small. The validity of this assumption has been further investigated. The 6SV1 atmospheric radiative transfer code (Kotchenova et al., 2006) has been used to perform a simulation of the atmospheric transmittance between the surface and the TOA for two 2.5 nm-wide spectral windows centered in 755 and 771 nm. This simulation

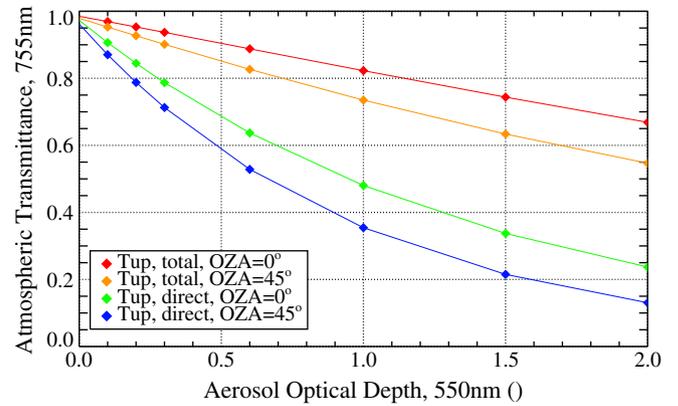


Fig. 4. Simulated total and direct upward transmittance T_{up} in 755 nm for two different observation zenith angles (OZAs). Simulations are for clear-sky conditions, a continental aerosol model and the surface located at the sea level. Very similar results are found for the 770 nm window.

assumes a mixed Rayleigh-aerosol atmosphere, clear-sky conditions, a continental aerosol model and the surface at the sea level. Total and direct transmittance factors for a range of aerosol optical depth (AOD) values and two different observation zenith angles are presented in Fig. 4. The total transmittance considering both direct and diffuse surface-to-sensor radiation applies to green canopies which are homogeneous within the atmosphere coupling scale, assumed to be of some kilometers. The direct transmittance would apply to a limiting case in which a fluorescent punctual source is surrounded by non-fluorescent surfaces. Due to the relatively coarse resolution of TANSO-FTS observations, it can be assumed that only large vegetated areas contribute to the binned fluorescence patterns so that the total transmittance applies. Adjacency effects and the separation between direct and diffuse radiation should be considered for an instrument with a higher spatial resolution.

Results in Fig. 4 confirm a very small atmospheric effect on the F_s signal estimated at the TOA for low to medium aerosol loadings and near-nadir observations. However, assuming $F_s^{TOA} \approx F_s$ may lead to non-negligible errors for the largest AODs and observation angles (up to 30° in regular FTS observations). Since a random distribution of observation angles between 0 and 30° can be assumed for TANSO-FTS data, the dependence of the upward transmittance on the observation angle would lead to increasing the random error in the final F_s maps if this effect is not properly accounted for. In contrast, the existence of particularly high aerosol contents over some areas of the planet and at given times of the year may affect the spatial and temporal distribution of F_s in the final binned products. Using global AOD data from the reanalysis product of the MACC (Monitoring atmospheric composition and climate) project, it has been estimated that the monthly mean of the upward transmittance can drop up to 0.8 in for instance Equatorial regions of Africa and the Beijing area. This factor might lead to systematic errors in the final F_s maps over regions with persistently high aerosol contents. Moreover, additional errors can be introduced in single F_s retrievals by sudden aerosol loading outburst events such as wildfires or desert dust intrusions over vegetation areas, which would also lead to increasing the dispersion of the gridded mean values.

Only a simple approach has been adopted at this point for the correction of atmospheric extinction on F_s retrievals. This is based on an upward transmittance factor directly applied to the monthly $2^\circ \times 2^\circ$ F_s composites. To this end, a gap-free AOD climatology has been derived from the MACC project² aerosol reanalysis product. Such a correction has been shown to have a very small impact on the final F_s maps. It

² <http://www.gmes-atmosphere.eu/data/>.

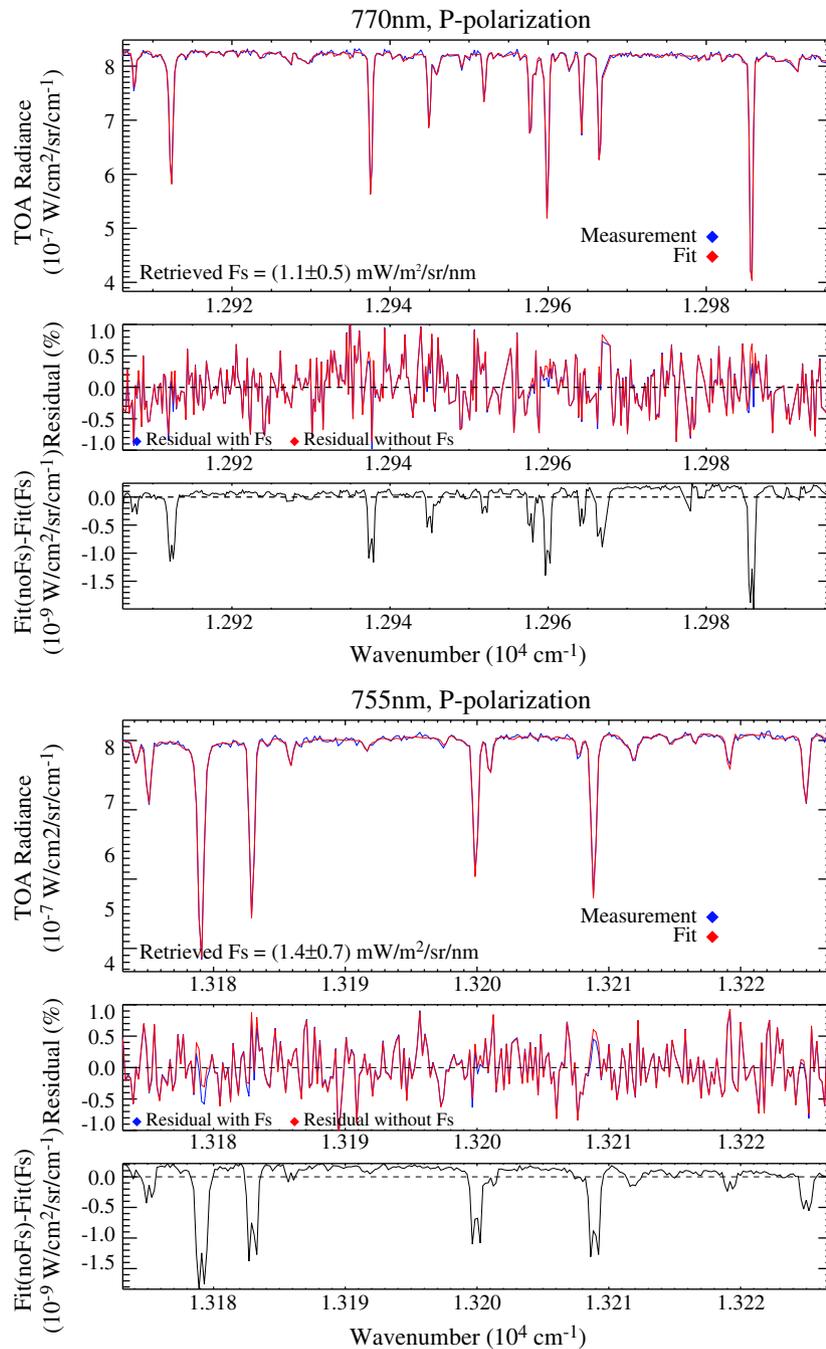


Fig. 5. Spectral fits and fit residuals for the two spectral micro-windows and P polarization. The retrieved F_s and its uncertainty is depicted. The impact of F_s on the fit is illustrated by the difference in the fit residuals for the forward model with and without F_s .

can only account for the systematic errors in F_s retrievals over areas where high aerosol loadings are expected on a monthly basis, so it cannot reduce the noise in F_s composites due to abnormally high AOD values or different observation angles. A per-spectrum atmospheric extinction correction is left as a matter of future research. The level-2 aerosol product from the TANSO-CAI (Cloud and Aerosol Imager) instrument could be used for this purpose.

4.3. Forward model fits

Exemplary spectral fits for the two spectral micro-windows in the P polarization are displayed in Fig. 5. Residuals with and without considering F_s in the forward model are plotted separately to illustrate the expected *in-filling* of Fraunhofer lines by F_s . A similar performance

has been found in S polarization fits, except for the higher amplitude of the residuals due to the lower SNR of the S polarization. It can be observed that no apparent spectral correlation is found between the input spectrum and the residuals, which suggests that spectral structures in the input spectra are properly modeled. The reason for the different radiance units used is that $\text{Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$ are the native units of the FTS, whereas $\text{mWm}^{-2}\text{sr}^{-1}\text{nm}$ are units normally used in the fluorescence literature.

4.4. Monthly F_s maps derived with the SVD-based retrieval approach

Global monthly composites of F_s are displayed in Fig. 6 for July and December 2009. The standard deviation of the F_s weighted mean σ_{F_s} as expressed in Eq. (5) is also depicted. Those two months have

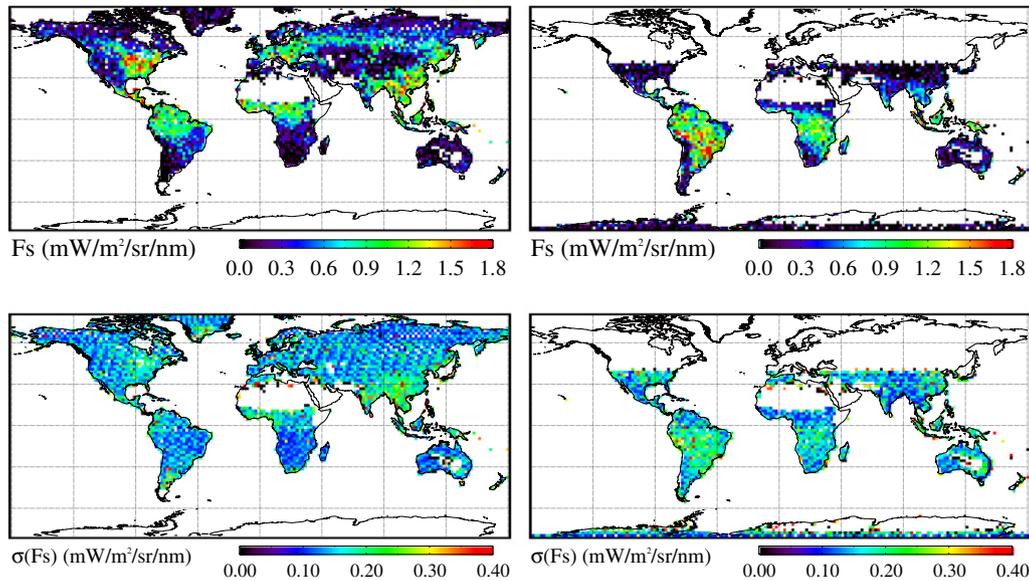


Fig. 6. Global F_s maps (composites from retrievals at 755 nm and S and P polarizations) for July 2009 (left column) and December 2009 (right column). The standard error of the weighted average (σ_{F_s}) is displayed in the second row.

been selected for the direct comparison with Joiner et al. (2011), who used the same two months as study cases. As expected, high F_s levels are detected over densely vegetated areas, such as the Amazon and the Central Africa rainforests in both July and December, the East coast of North America and the vegetated parts of the Far East in July, and the South of Brazil in December. These spatial patterns correspond well for both months with those shown in Joiner et al. (2011).

However, there is a clear mismatch in the range of F_s levels measured in this work and those shown in Joiner et al. (2011). On a global basis, this range is between 0 and $1.8 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ for the SVD-based approach and between -1 and $4 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ for Joiner et al. (2011). The lack of significant negative F_s values in the SVD-based maps is a good proof of consistency of the retrieval approach. It must be acknowledged that this is mostly due to the offset correction step performed after Frankenberg et al. (2011b), which compensates an F_s bias of about $0.5 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ for low signal levels (see Fig. 14). This correction was not considered in the pioneering work of Joiner et al. (2011), but has been accounted for in a very recent paper (Joiner et al., 2012). On the other hand, there is a large difference in the upper limit of the retrieved F_s . Since there is no reference F_s data set which can be used for the quantitative validation of GOSAT retrievals, it is difficult to say what this maximum F_s limit might actually be. Nevertheless, using the ground-based measurements reported in e.g. Meroni et al. (2009), Entcheva-Campbell et al. (2008), we can speculate that the $4 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ measured by Joiner et al. (2011) might be too high for $2^\circ \times 2^\circ$ cells, especially considering that the retrievals are performed in 770 nm where the signal is about half of that in 755 nm. On the other hand, as it will be discussed in Section 4.5, the range of F_s values derived with the SVD-based approach is very similar to those provided by the physically-based approach by Frankenberg et al. (2011b).

Concerning the retrieval uncertainty, σ_{F_s} values in Fig. 6(c)–(d) provide an estimation of the uncertainty to be associated to the average F_s in each cell. It can be seen that σ_{F_s} ranges between 0.1 and $0.4 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$, which give typical relative standard errors of 20–30% for these monthly composites. These values are directly related to the number of spectra which satisfy the quality filtering criteria. For example, σ_{F_s} values are particularly high for the Amazon basin in December, when the area is mostly cloudy during the wet season. A typical number of 25–35 spectra per cell passed the quality filtering in the monthly averaging, but this number may go down to

0–5 for the Amazon basin in December. It must be noted that this represents an optimistic uncertainty estimation, as only instrumental noise propagated to mean F_s values is considered by this error figure.

Information about the performance of the SVD-based F_s inversion approach can be obtained by analyzing the independent retrievals for each of the 4 fit combinations. Maps of F_s for July 2009 in the 4 fitting windows are presented in Fig. 7. By looking at the four maps it can be stated that no offset between the two polarizations in the 755 nm window is apparent, whereas a slight overestimation is found in 770 nm in the S polarization with respect to the P polarization. It is also clear that F_s in 755 nm is systematically higher than in 770 nm. This is a good proof of performance for the retrieval. The F_s ratio between 755 and 770 nm is roughly about 1.6 for most of the vegetated surfaces, which is somewhat lower than the 1.8–1.9 estimated from real measurements over crops (e.g. Amorós-López et al., 2008). A small retrieval offset in either 755 or 770 nm could account for this difference. A spectrally dependent re-absorption of fluorescence within the canopy up to the top-of-canopy level might also lead to changes of the slope of the F_s spectrum. This would depend on canopy structure effects. It must be underlined that available reference F_s spectra are usually acquired at the leaf-level and might therefore not represent complex canopies.

The global distribution of the singular vector weights ω_i obtained from the inversion of the forward model is also informative of the retrieval performance. The monthly composite of the weights of the first 4 singular vectors in 755 nm and the P polarization for July 2009 are shown in Fig. 8. Different scaling factors have been applied for visualization. These first 4 vectors are the ones in the state vector of the 755 nm micro-window, whereas a fifth one is used for 770 nm. As expected, the first weight ω_1 is directly correlated to the TOA radiance in 755 nm (not displayed). Vectors 2 and 4 roughly represent spectral alignment and band broadening, respectively. In both cases the dependence on the orbit and the scan angle can be observed (e.g. southern Africa for v_2 and northern Europe for v_4 , respectively). Spatial patterns and correlations with the TOA signal given by v_1 are also visible, which suggests that the two vectors are also affected by background changes and not only by the instrument spectral performance. This is expected for v_4 , as the band broadening is known to correlate with signal levels (see Section 4.1).

v_3 is the one with the most difficult interpretation: as shown in Fig. 4, it represents a spectral slope similar to that of F_s , and for this reason there might be some degree of linear dependency between

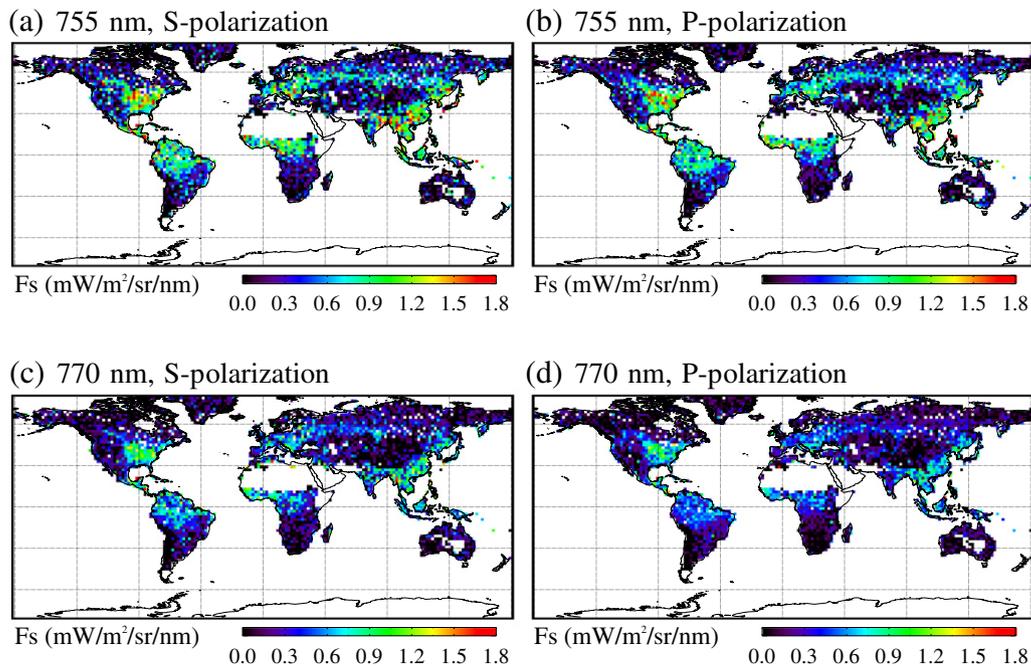


Fig. 7. Global F_s maps for July 2009 from GOSAT-FTS data in two spectral micro-windows and two polarizations.

the two. It has been discussed previously that v_3 in 755 nm can be associated to spectral radiance slopes, which are in turn strongly correlated with the particular background reflectance for a given area. For example, the reflectance of green vegetation canopies normally increases with wavelength around 755 nm due to chlorophyll absorption (Dash & Curran, 2004). Despite the spectral slope normalization performed on both the input spectra and the training set, clear background patterns are observed in v_3 , which demonstrates the spectral normalization by a straight line is not able to remove spectral slopes completely. From the F_s retrieval point of view, it is important to note the strong signal coming from vegetated areas such as the East coast of North America or China. As said, vegetation with high chlorophyll content (naturally correlated to F_s) is present in those areas, which can then explain high ω_3 values without leading to a conclusion of whether ω_3 and F_s are linearly dependent in the retrieval or not. However, there are other areas like the semi-arid regions in central Asia where ω_3 is also high without higher levels of F_s being measured. The same is true for Greenland, but the correction of F_s biases due to radiometric offsets, which has a strong effect over the brightest surfaces, causes that we cannot make a solid statement about the linear independence of F_s and v_3 .

4.5. Comparison of annual F_s retrievals with other remote sensing-based vegetation data sets

As described in Section 3.5, F_s maps derived with the SVD approach have been compared with the *physically-based* approach by Frankenberg et al. (2011a) and with two GPP products. Annual average maps (June 2009 to May 2010) for the two F_s data sets and the two GPP data sets are presented in Fig. 9. The F_s maps from the *physically-based* approach are generated by aggregating only S polarization retrievals from the two spectral micro-windows, with the 770 nm retrievals being scaled by 1.8 as described in Frankenberg et al. (2011b) in order to account for the decrease of F_s with wavelength in this part of the spectrum. Only 755 nm retrievals but in the two polarizations are used for the SVD approach for the reasons discussed in Section 3.3. A very good agreement between the two F_s data sets can be observed. The only noticeable difference between them seems to be in central Europe, where the *physically-based* approach gives

higher F_s values in the annual average. These higher F_s levels do not seem to appear in the GPP data sets. In any case, the fact that the range of F_s values is almost the same for these two relatively different approaches gives confidence on the F_s intensity levels found earlier by Frankenberg et al. (2011b), which are in turn very different from those reported by Joiner et al. (2011). F_s maps also compare well with the GPP products over most of the globe, as it was already reported in Frankenberg et al. (2011b).

A deeper look into these comparisons is provided by the scatter plots displayed in Fig. 10. The same F_s annual average from June 2009 to May 2010 gridded in $2^\circ \times 2^\circ$ cells is plotted against equivalent annual means of other parameters. Cells with $\sigma_{F_s} > 0.07 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ in either the *physically-based* or the SVD F_s products have been filtered out to remove outliers. The first conclusion to be extracted from this comparison is the very high agreement between the two F_s retrieval approaches, which show no bias and an almost identical range of values for these annual means. This very good comparison reinforces the confidence in the use of Fraunhofer lines to retrieve F_s as well as in the use of GOSAT data for this purpose, which is one of the most important findings of this work.

Concerning the comparison between F_s and the other vegetation products, the same qualitative conclusions reached in Frankenberg et al. (2011b) are found here: the best correlation with F_s is achieved by the MPI-BGC GPP product ($r^2 = 0.85$), followed by the MODIS GPP product ($r^2 = 0.82$). Worse correlations are found with the three vegetation indices, which seem to deviate from a linear relationship with F_s for high F_s levels (indicating high canopy chlorophyll content). This non-linearity is apparent for NDVI and MTCI, which seem to present some saturation. These effects are widely seen in vegetation index-GPP relationships (e.g. Phillips et al., 2008). Other relevant comparisons not displayed in Fig. 10 are the slightly worse correlations between the F_s *physically-based* approach and the MPI-BGC and MODIS GPP products (r^2 of 0.82 and 0.77, respectively), and the higher correlations of $r^2 = 0.92$ between the two GPP products and between each GPP product and EVI (0.89 and 0.88 for MPI-BGC and MODIS, respectively). In this latter case, nonlinear trends are again observed for high EVI values, which suggest that the relationship between GPP and vegetation indices is less linear than the one between F_s and GPP. The higher r^2 can then be partly

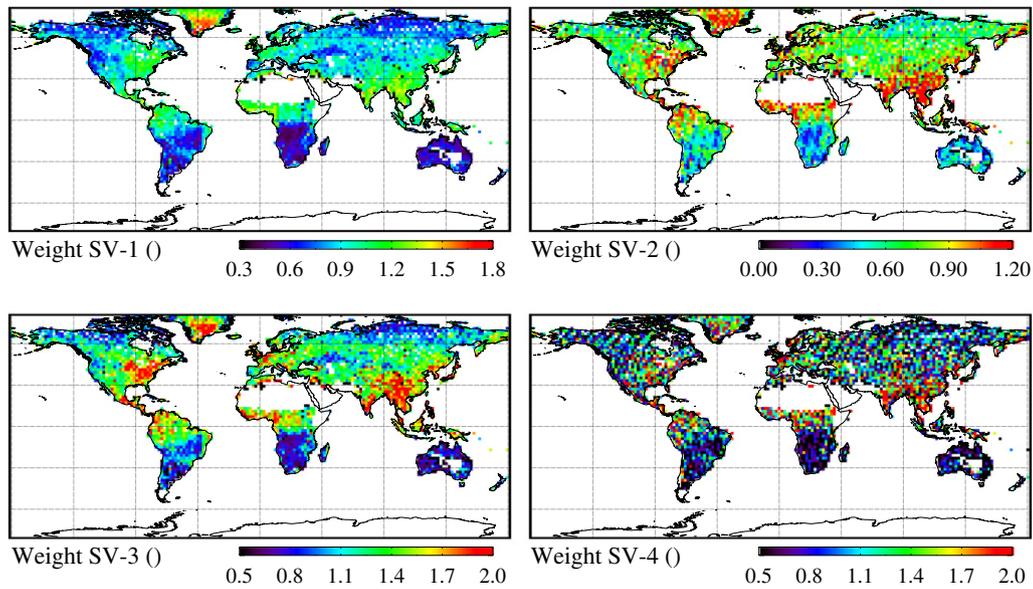


Fig. 8. Monthly composites (July 2009) of the weights ω_i of the first 4 singular vectors obtained after F_s retrieval in 755 nm and P-polarization. For visualization purposes, ω_1 is scaled by $-1 \cdot 10^5$, ω_2 by $1 \cdot 10^7$, ω_3 by $-1 \cdot 10^8$, and ω_4 by $1 \cdot 10^8$.

explained by the lower noise in EVI with respect to the gridded F_s retrievals.

It must be stated that this kind of comparison is not intended to represent any sort of validation of the F_s product. Fundamental differences can be expected between F_s and both GPP and vegetation indices. In the first case, GPP is a step further from electron transport and photosynthesis, which are the processes to which F_s is directly linked. Although the good correspondence between F_s and GPP can be predicted, it is also expected that the linear relationship between F_s and GPP breaks down under certain conditions such as strong water and temperature stresses. Besides, both the MPI-BGC and the MODIS GPP data sets are derived from a mixture of different data sources and modeling approaches with their own limitations. With respect to the vegetation indices, these are expected to provide a more static view of vegetation activity than F_s , as their sensitivity to environmental factors is just as much as that of top-of-canopy reflectance. For these reasons, no perfect comparison can a priori be expected from this comparison between F_s and the other vegetation products, but only an overall good correspondence. No reference data set is currently available for the actual validation of F_s retrievals at the spatial and temporal resolutions of the GOSAT product.

4.6. Temporal patterns in F_s and GPP

Scatter plots between F_s and MPI-BGC GPP, MODIS GPP and EVI are shown in Fig. 11. MPI-BGC GPP data cover the time period between 6/2009 and 9/2010, MODIS GPP the time period between 6/2009 and 12/2010, and MODIS EVI the time period between 6/2009 and 3/2011. For the comparison with EVI, F_s retrievals have been scaled by the cosine of the illumination angle in order to minimize the direct impact of changing illumination along the year. The first conclusion to be extracted is that F_s generally shows a high linear correlation with each of the data sets also in the temporal domain (the correlation in the spatial domain is already demonstrated by the scatter plots from the annual means in Fig. 10). This can in principle be explained by phenology and growth cycles being the first driver of F_s at a monthly scale. However, it is also observed that (a) the relationship between F_s and GPP has a very strong dependence on the biome, and (b) both the slope and the correlation between F_s and GPP are different for the two GPP data sets.

The fact that the relationship between GPP and F_s depends on the vegetation type could be expected a priori, but it is nonetheless very important to state that F_s alone cannot give a direct estimation of GPP, but some further modeling involving the particular photochemical and canopy structure properties of the different biomes is necessary for the most accurate use of F_s . Concerning the differences between F_s and the MPI-BGC and the MODIS GPP products, the most important discrepancies are in the cropland, the evergreen broadleaf forest and the savanna classes. For the first two of these biomes, the comparison between F_s and the MODIS GPP product shows low correlations, which is in principle surprising given that the MODIS EVI product shows good comparisons for the same data sets. It is difficult to determine where the main reason for these deviations might be due to the differences in the fundamental basis of each model (upscaling of flux tower data in the case of MPI-BGC, look-up table of light use efficiency driven by biome and meteorology for MODIS). It can be speculated that the different slopes in the F_s -GPP relationships for the two different GPP products may come from the particular implementation of the light use efficiency dependence on temperature and vegetation type in the MODIS product, which is not explicitly done for the MPI-BGC data set. Concerning the low correlation between F_s and GPP MODIS for some biomes, it can again be speculated that the MODIS product may suffer from some problem in highly cloudy areas, as reported in Zhao et al. (2005). An accurate analysis of these issues is considered to be out of the scope of this work.

A different view of those temporal series is provided in Fig. 12. Temporal series of F_s and GPP are plotted for four selected biomes. Some of the trends observed in Fig. 11 are confirmed. The bad comparison between the MODIS GPP product and F_s is observed in the evergreen broadleaf forest biome, that in this case covers mostly the tropical rainforest in the Amazon basin and Central Africa. Both GPP MPI-BGC and F_s show the same seasonal patterns, which are very different to those shown by MODIS GPP. A very good comparison is found for the deciduous broadleaf forest biome, but the comparison between the two GPP products is again different for the cropland and savanna-north biomes. In the first case, the two GPP products give similar seasonal cycles but a very different amplitude, whereas the MODIS GPP product shows an unexpected temporal pattern for the savanna-north biome, which covers most of the vegetated areas between 0° and 30° latitude in Africa. It must be

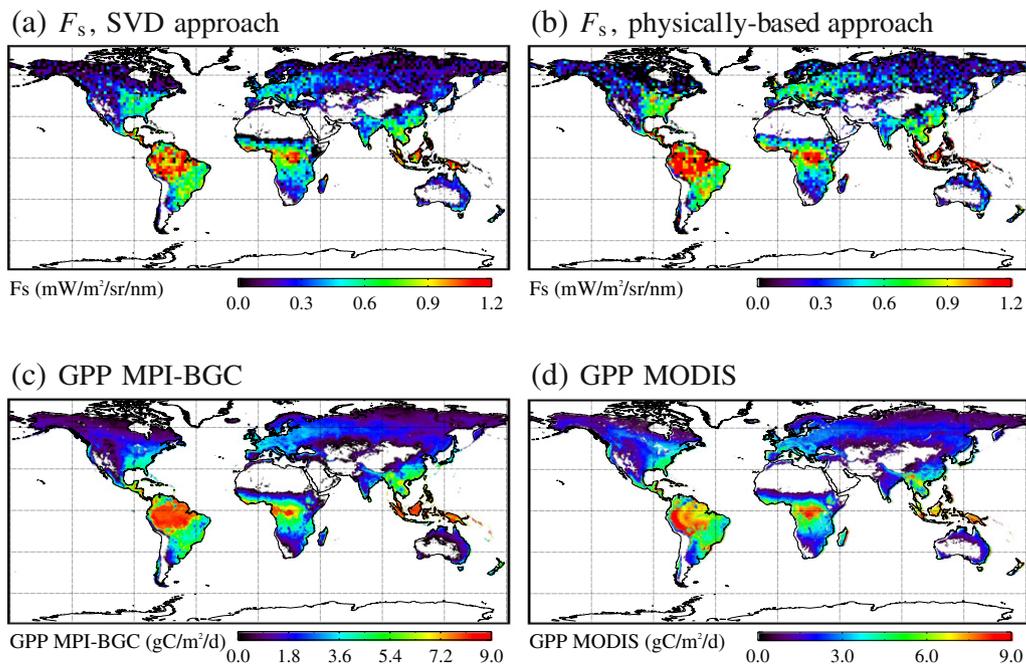


Fig. 9. Global maps of F_s and gross primary production (GPP) for the time period 6/2009–5/2010. F_s maps from both the SVD and the Frankenberg et al. (2011) (physically-based) retrieval approaches are displayed. GPP data are from the MPI-BGC and MODIS MOD17A2 data sets. The non-data mask from the MPI-BGC GPP product has been used for all 4 data sets.

noted that the MODIS EVI product (not shown) shows the same seasonality as F_s and the GPP MPI-BGC data sets, so that the unexpected seasonality may again be given to some particularity of the MODIS GPP model.

4.7. Directional effects in F_s retrievals

The last experiment that has been performed with the F_s data set produced with the SVD approach has been the assessment of directional effect on the F_s emission. Despite the fact that the F_s emission has traditionally been considered isotropic, it can be argued that some directionality can be expected by analogy with the thermal emission by vegetation covers (e.g. Timmermans et al., 2009). Indeed, Van der Tol et al. (2009) explored directional effects in vegetation fluorescence (685 nm) using the SCOPE model and found them to be quite pronounced. Their simulations suggest a broadly similar form to near infrared bidirectional reflectance distribution function (BRDF), particularly in the backscatter direction, with a pronounced retro-reflectance (hot spot) and angular variations of greater than a factor of 2 in F_s for a mature canopy.

For this experiment with GOSAT-derived F_s , all F_s retrievals over two densely vegetated areas around the Equator, one in America and one in Africa, have been grouped in five different scan angle bins of 8° width and centered in -30° , -15° , -2° , $+15^\circ$ and $+30^\circ$, which roughly reproduce the distribution of scan angles in the observations. The sign of these angles refers to which side of the orbit the instrument is pointing at. This sign can be associated to azimuth angles with respect to North by considering the inclination of the orbit. The two areas have been selected to be at the Equator in order to minimize changes in the illumination zenith angle, which is in both cases around $(20 \pm 5)^\circ$ during the entire year. Only observations between June 2009 and May 2010 have been used in order to avoid the switch from 5 to 3 scan angles in TANSO-FTS observations which occurred in August 2010.

Results are plotted in Fig. 13. F_s and TOA reflectance are plotted as a function of scan angle for those two sites. Even though this can only be considered a preliminary assessment of the directionality of the F_s emission from space, very clear directional trends

are visible. The signal is reduced by up to a 20% for the largest observation angles around 30° , and a maximum F_s is measured for the $+15^\circ$ observation angle, which seems to be the closest one to the hot-spot configuration according to the equivalent TOA reflectance patterns.

The primary physical mechanism for this is likely to be the varying amounts of sunlit leaf that are viewed at different angles. In the simulations of Van der Tol et al. (2009), F_s appears to show a broad “dark spot” in the anti-solar direction, and generally less pronounced forward emission than seen in near infrared BRDF which is consistent with this idea. Multiple scattering effects, that are generally large for vegetation reflectance in the near infrared, are likely small for F_s in the Fraunhofer lines. From such arguments, we might suppose directional fluorescence would broadly follow the form of visible reflectance (high leaf absorptance, so low multiple scattering) when bounded by an absorbing lower boundary (i.e. no fluorescence from the soil). This suggests the possibility that relatively simple BRDF models, such as those used for describing angular variations in reflectance for albedo estimation (e.g. Wanner et al., 1997) might be suitable for describing and normalizing such effects, as has already shown to be the case for describing directional emissivity at thermal wavelengths (Snyder & Wan, 1998). A more detailed analysis of this effect and its implications in the modeling of the signal will be carried out in future activities.

5. Summary and conclusions

A new method for the retrieval of chlorophyll fluorescence from GOSAT-FTS spaceborne measurements has been presented in this work. This method makes use of the Fraunhofer lines located in two spectral windows of 2–3 nm centered in 755 and 770 nm to disentangle F_s from the solar radiation reflected by the surface-atmosphere system. The inversion of TOA radiances is achieved by means of a linear forward model consisting of a basis of orthogonal vectors to reproduce F_s -free spectral patterns plus the fluorescence contribution. The basis of singular vectors has been derived with the singular vector decomposition of a training set of vegetation-free spectra for the two spectral micro-windows. Whereas this

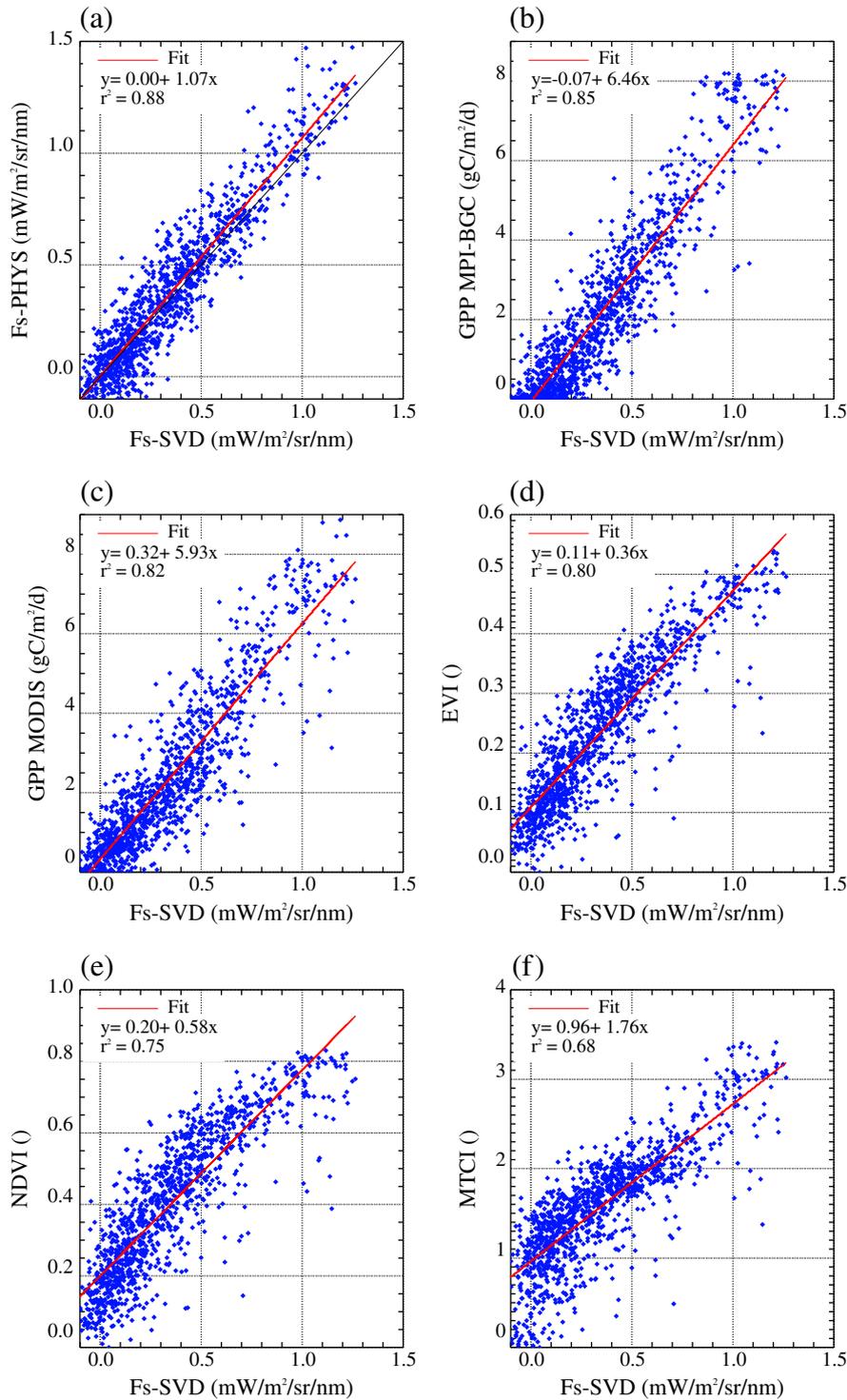


Fig. 10. Scatter plots of annual mean F_s against other vegetation remote sensing products for the time period 6/2009–5/2010 and 2° cells. In addition to the GPP data sets in Fig. 9, comparisons with MODIS EVI and NDVI and with the MERIS MTCI products are shown.

SVD-based method has some fundamental similarities with the state-of-the-art approach described in Frankenberg et al. (2011a, b) (same two spectral micro-windows, two polarizations processed separately, empirical correction of radiometric offsets), this method may have some advantageous features such as a simple implementation and a higher efficiency.

The proposed F_s retrieval method has been tested on 22 months of GOSAT TANSO-FTS level 1B data. Some of the most relevant findings from the analysis of the resulting F_s data set are:

- (1) A simple linear model can be used to estimate F_s from GOSAT data on a global basis. Thanks to the statistical approach chosen there is no need for an accurate solar irradiance data set and for the modeling of the instrument ILSF, while the linear nature of the forward model makes it more efficient in terms of computation time.
- (2) The very good comparison with the retrievals from Frankenberg et al. (2011b) in both F_s intensity levels and geographical patterns gives confidence in the consistent performance of both

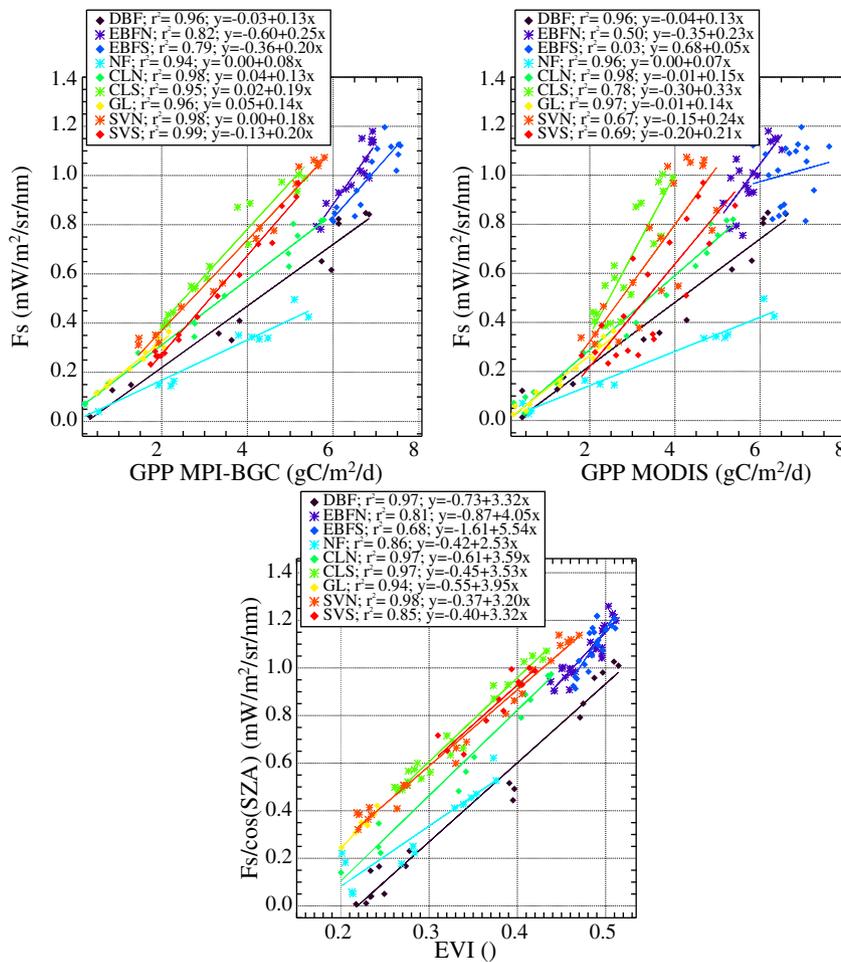


Fig. 11. Scatter plots between F_s and MPI-BGC GPP, MODIS GPP and MODIS EVI. Each symbol represents one month. Biomes follow the IGBP-based land cover classes detailed in Frankenberg et al. (2011). [DBF]=Deciduous Broadleaf Forest, [EBLN(S)]=Evergreen Broadleaf Forest in the northern (southern) hemisphere, [NF]=Needleleaf Forest, [CLN(S)]=Cropland in the northern (southern) hemisphere, [GL]=Grasslands, [SVN(S)]=Savannas in the northern (southern) hemisphere.

methods. It also enables the consolidation of the range of F_s values to be expected at the temporal and spatial scales sampled by GOSAT-based retrievals.

- (3) Also similar results to Frankenberg et al. (2011b) have been reached in terms of the comparison of annual means of F_s with satellite-based GPP and vegetation indices: the best correlation is found between F_s and the MPI-BGC GPP dataset, followed by the MODIS GPP data set. The comparison with vegetation indices has shown the relationship with F_s is not linear for the higher F_s levels.
- (4) Concerning F_s temporal patterns for different biomes, a very high linear correlation between F_s and MPI-BGC GPP has been found, although the slope of this linear correlation shows a strong dependence on the biome under consideration. This confirms that, even though F_s is expected to be a direct proxy to electron transport and photosynthesis, some modeling accounting for biome-related aspects such as the photosynthetic pathways or the canopy structure must be taken into account for the proper upscaling of F_s to GPP.
- (5) The same F_s -GPP comparison is worse for the GPP MODIS product, which has shown lower linear correlations than the MPI-BGC data set, and also substantially different slopes between GPP and F_s for some of the biomes. The comparison of the MODIS GPP data set with both F_s and the MPI-BGC GPP is particularly bad for some of the biomes investigated.
- (6) A strong directionality in the F_s emission has been observed for two densely vegetated Equatorial forest areas. This is opposed

to the traditional isotropic assumption for fluorescence. These directional effects might require modeling for the most accurate exploitation of the F_s signal.

Near-future research will be devoted to the systematic analysis of this F_s data set, putting the emphasis on the investigation of the fluorescence signal itself, the impact of different plant functional types in the F_s -GPP relationship, and the environmental controls driving the spatial and temporal patterns. Both statistical and theoretical approaches are expected to help us to understand the potential and limitations of this new parameter for the remote observation of vegetation functioning.

Acknowledgment

We would like to acknowledge JAXA, NIES and MOE for making GOSAT data available to the scientific community. Dr. H. Watanabe and NIES are especially thanked for opening the large volume GOSAT data distribution server to PIs. The MPI-BGC GPP was kindly provided by Martin Jung from the Max Planck Institute for Biogeochemistry, MODIS MOD17 GPP data were downloaded from the server of the Numerical Terradynamic Simulation Group at the University of Montana, MODIS MOD13 EVI/NDVI data were obtained from the MODIS LP DAAC archive, and MERIS-MTCI from the Infoterra Ltd server. H. Boesch and R. Parker from the University of Leicester are thanked for initial help with GOSAT data handling. C. Van der Tol from ITC is also thanked for useful comments on the manuscript.

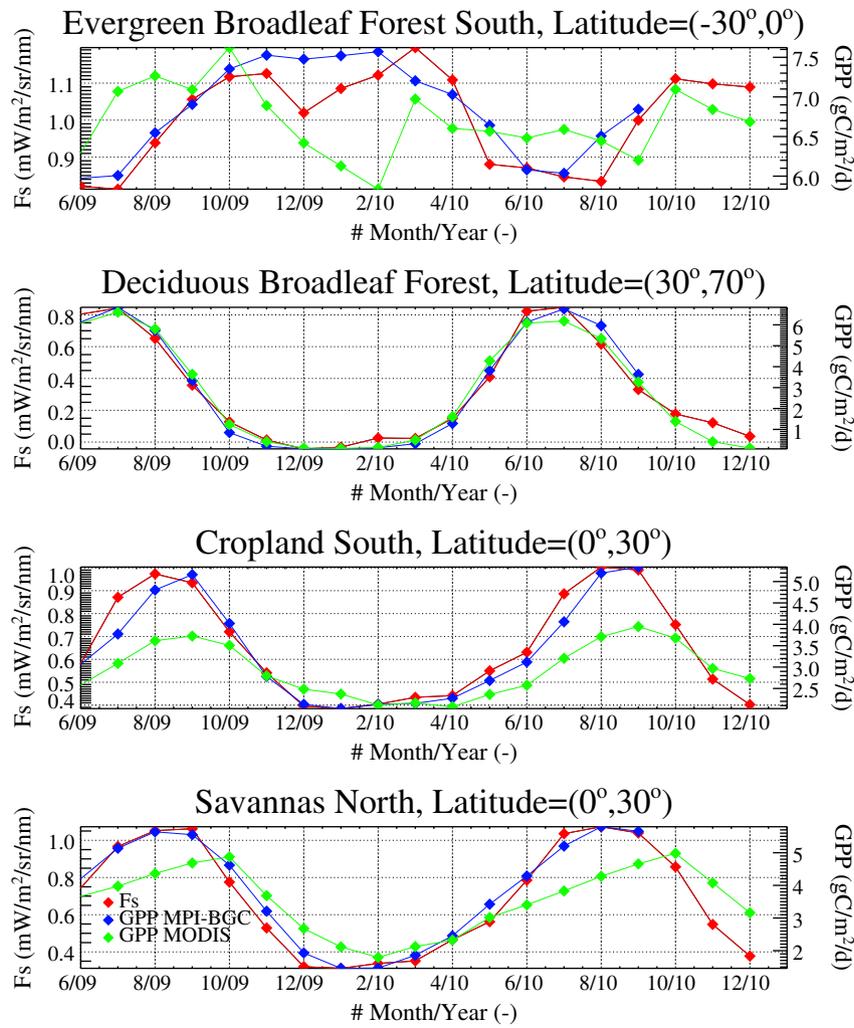


Fig. 12. Temporal series of F_s and GPP for 4 different biomes. The scaling of the vertical axes of all four plots differs.

The research of LG has been funded by the European Commission through the 7th Framework Marie Curie Actions Programme. We thank three anonymous reviewers for their constructive comments to improve the quality of this manuscript.

Appendix A. Assessment of time-dependent F_s biases from radiometric offsets

Time-dependent F_s offsets have been estimated over vegetation-free surfaces following the method described in Section 3.2. Fig. 14 displays F_s offsets for the 2 polarizations and the 2 micro-windows as a function of the average radiance in the 755–773 nm spectral window and for July 2009 and July 2010. These curves compare well to those shown in Frankenberg et al. (2011b), which were generated from relative F_s instead of from absolute F_s . It can be stated that the two polarizations follow different patterns, and that the curves are highly nonlinear in medium signal levels. These offsets correct for F_s biases of around $0.5 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ at low radiance levels, whereas they have a relatively low impact for the medium radiance levels around $5.5 \cdot 10^{-7} - 6.5 \cdot 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$ which typically correspond to tropical and mid-latitude vegetation covers.

On the other hand, a strong temporal dependence in the radiometric offsets is detected from the comparison of July 2009 and July 2010. In principle, these changes can only be explained by instrument-related causes, as the F_s -free data set used to generate the curves is expected to be statistically invariant for the same

month of the year. An insight into this effect is given in Fig. 15. The complete temporal series of F_s offsets and radiance spectral slopes for a radiance level of $6 \cdot 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$ are plotted. It can be seen that F_s offsets change with time in a similar fashion for the two spectral windows irrespective of the polarization, which does

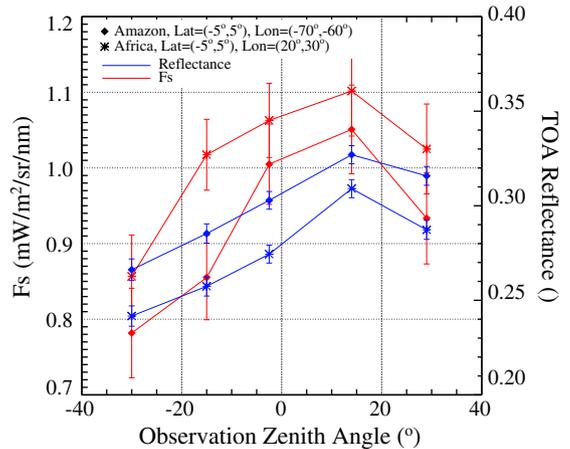


Fig. 13. F_s and top-of-atmosphere reflectance in 755 nm as a function of the observation zenith angle for two densely vegetated forest areas. Error bars (scaled by 0.1 for visibility) depict the standard deviation of all averaged retrievals. Sun zenith angle is around $(20 \pm 5)^\circ$ along the year for the two sites.

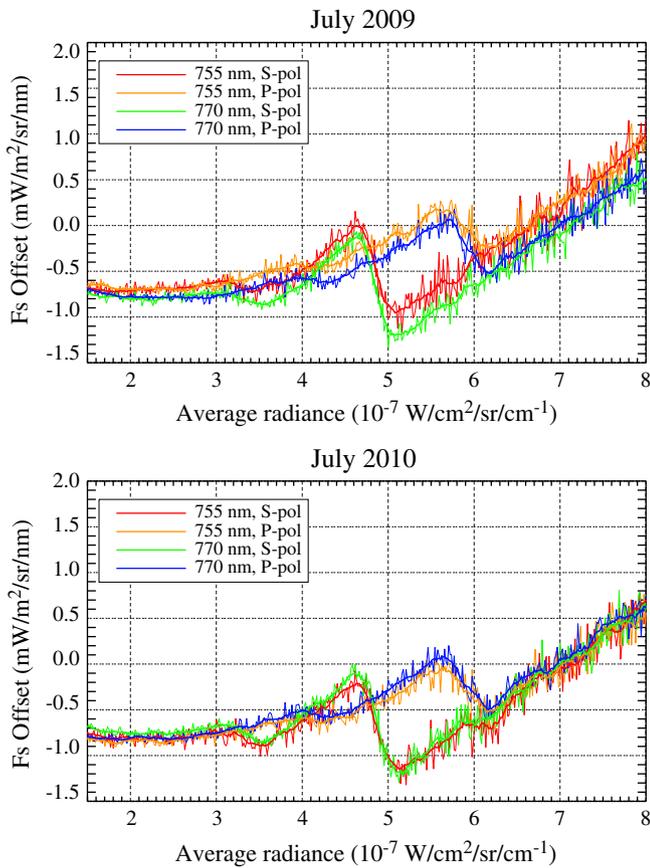


Fig. 14. F_s offsets in the 4 fitting windows as a function of the average radiance in the FTS band-1 for July 2009 and July 2010.

not agree with the apparent stability of the S polarization reported by Frankenberg et al. (2011b). It must also be observed that there is a clear seasonal trend in the estimated offsets, especially in 770 nm. The reason for such a seasonal pattern might be in either a seasonal-dependence in the instrument temperature (confirmed by in-flight calibration devices), which would affect the zero-level offsets, or in the type of surfaces dominating the monthly data sets used to estimate the offsets. A similar temporal dependence of radiometric offsets in the FTS band-1 have been recently reported by Joiner et al. (2012).

Concerning the spectral component of the temporal evolution of the F_s offsets, it is unexpected that the temporal changes in the 755 nm F_s offsets are different from those in 770 nm, as shown in Fig. 15. A temporal degradation of the FTS detector response reported in Kuze et al. (2011) could partially explain the decrease of the 755 nm offsets. However, no spectral component is in principle to be associated to this degradation, so it must be the same for 755 and 770 nm. Moreover, the magnitude of the detector degradation rate reported in Kuze et al. (2011) seems to be insufficient to account for the strong offset change in 755 nm along the entire time period.

In order to gain more information on this issue, it has been investigated whether these temporal patterns may come from any effect affecting the two spectral micro-windows in 755 and 770 nm differently. Fig. 15 also displays the temporal evolution of the spectral slope of the TOA radiance spectra for the 4 fitting windows. It can be observed that there is a noticeable temporal pattern in the radiance slopes in the 755 nm window, which suggests that there might be a spectral component associated to the evolution of the instrument performance. On-going laboratory analyses have shown that the possibility of the detector degradation is very low, as the FTS band-1 detector is a very simple Si diode whose characteristics are not expected to change. Optical efficiency degradation (FTS modulation efficiency or contamination on the optics) and an analog circuit seasonal change are more likely explanations.

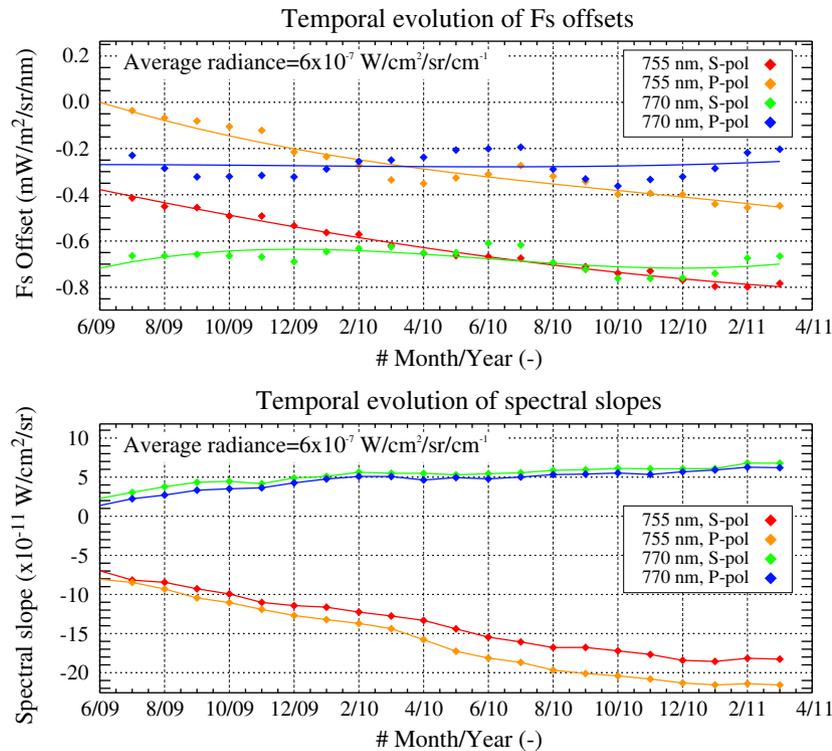


Fig. 15. Temporal evolution of F_s offsets and the spectral slope of TOA radiance for the 4 fitting windows and a radiance level of $6 \times 10^{-7} \text{ Wcm}^{-2}\text{sr}^{-1}\text{cm}^{-1}$. Diamonds depict the estimated offsets, whereas the solid lines correspond to a 2nd order polynomial fitted to the offsets for each average radiance level. F_s offsets in 770 nm have been shifted by $0.3 \text{ mWm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$ for visualization purposes.

In any case, the similarity between the temporal trends in both F_5 -offsets and the radiance spectral slopes enable the hypothesis that either the F_5 retrieval algorithm is sensitive to radiance spectral slopes, or the same instrumental issue making the spectral slopes to change is somehow affecting the fractional depth of the Fraunhofer lines with a different sign for the two spectral windows.

References

- Abrams, M. C., Toon, G. C., & Schindler, R. A. (1994). Practical example of the correction of Fourier-transform spectra for detector nonlinearity. *Applied Optics*, 33, 6307–6314.
- Amorós-López, J., Gómez-Chova, L., Vila-Francés, J., Alonso, L., Calpe, J., Moreno, J., et al. (2008). Evaluation of remote sensing of vegetation fluorescence by the analysis of diurnal cycles. *International Journal of Remote Sensing*, 29, 5423–5436.
- Baker, N. R. (2008). Chlorophyll fluorescence: A probe of photosynthesis in vivo. *Annual Reviews – Plant Biology*, 59, 89–113.
- Chance, K., & Kurucz, R. (2010). An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(9), 1289–1295.
- Coops, N. C., Hilker, T., Hall, F. G., Nichol, C. J., & Drolet, G. G. (2010). Estimation of light-use efficiency of terrestrial ecosystems from space: A status report. *BioScience*, 60(10), 788–797.
- Dash, J., & Curran, P. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25, 5003–5013.
- Entcheva Campbell, P. K., Middleton, E. M., Corp, L. A., & Kim, M. S. (2008). Contribution of chlorophyll fluorescence to the apparent vegetation reflectance. *Science of the Total Environment*, 404, 433–439.
- Frankenberg, C., Butz, A., & Toon, G. C. (2011). Disentangling chlorophyll fluorescence from atmospheric scattering effects in O₂A-band spectra of reflected sun-light. *Geophysical Research Letters*, 38, L03801, doi:10.1029/2010GL045896.
- Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. -E., et al. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, 38, L17706, doi:10.1029/2011GL048738.
- Guanter, L., Alonso, L., Gómez-Chova, L., Amorós, J., Vila, J., & Moreno, J. (2007). Estimation of solar-induced vegetation fluorescence from space measurements. *Geophysical Research Letters*, 34(L08401), doi:10.1029/2007GL029289.
- Guanter, L., Alonso, L., Gómez-Chova, L., Meroni, M., Preusker, R., Fischer, J., et al. (2010). Developments for vegetation fluorescence retrieval from spaceborne high-resolution spectrometry in the O₂-A and O₂-B absorption bands. *Journal of Geophysical Research-Atmospheres*, 115(D19303), doi:10.1029/2009JD013716.
- Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213.
- Hurley, J., Dudhia, A., & Grainger, R. G. (2009). Cloud detection for MIPAS using singular vector decomposition. *Atmospheric Measurement Techniques*, 2(2), 533–547.
- Joiner, J., Yoshida, Y., Vasilkov, A. P., Middleton, E. M., Campbell, P. K. E., Yoshida, Y., et al. (2012). Filling-in of far-red and near-infrared solar lines by terrestrial and atmospheric effects: simulations and space-based observations from SCIAMACHY and GOSAT. *Atmospheric Measurement Techniques Discussions*, 5(1), 163–210.
- Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., Corp, L. A., & Middleton, E. M. (2011). First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences*, 8(3), 637–651.
- Jung, M., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research—Biogeosciences*, 116(G00J07), doi:10.1029/2010JG001566.
- Klüser, L., Martynenko, D., & Holzer-Popp, T. (2011). Thermal infrared remote sensing of mineral dust over land and ocean: a spectral SVD based retrieval approach for IASI. *Atmospheric Measurement Techniques*, 4(5), 757–773.
- Kotchenova, S. Y., Vermote, E. F., Matarrese, R., & Klemm, F. J., Jr. (2006). Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance. *Applied Optics*, 45, 6762–6774.
- Kuze, A., O'Brien, D. M., Taylor, T. E., Day, J. O., O'Dell, C. W., Kataoka, F., et al. (2011). Vicarious calibration of the GOSAT sensors using the railroad valley desert playa. *IEEE Transactions on Geoscience and Remote Sensing*, 49(5), 1781–1795.
- Kuze, A., Suto, H., Nakajima, M., & Hamazaki, T. (2009). Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Applied Optics*, 48, 6716–6733.
- Meroni, M., & Colombo, R. (2006). Leaf level detection of solar induced chlorophyll fluorescence by means of a subnanometer resolution spectroradiometer. *Remote Sensing of Environment*, 103, 438–448.
- Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R., et al. (2009). Remote sensing of solar induced chlorophyll fluorescence: review of methods and applications. *Remote Sensing of Environment*, 113, 2037–2051.
- Papageorgiou, G. (1975). Chlorophyll fluorescence: An intrinsic probe of photosynthesis. *Bioenergetics of Photosynthesis* (pp. 319–371). New York, USA: Academic.
- Phillips, L., Hansen, A., & Flather, C. (2008). Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production. *Remote Sensing of Environment*, 112, 3538–3549.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing*, Vol. 3. : Cambridge University Press.
- Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: Theory and practice*. : World Scientific Publishing Co. Ltd.
- Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *BioScience*, 54(6), 547–560.
- Sioris, C. E., Courrèges-Lacoste, G. B., & Stoll, M. P. (2003). Filling in of fraunhofer lines by plant fluorescence: Simulations for a nadir-viewing satellite-borne instrument. *Journal of Geophysical Research-Atmospheres*, 108, L4133, doi:10.1029/2001JD001321.
- Snyder, C., & Wan, Z. (1998). BRDF models to predict spectral reflectance and emissivity in the thermal infrared. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 214–225.
- Timmermans, J., Gieske, A., Van der Tol, C., Verhoef, W., & Su, Z. (2009). Automated directional measurement system for the acquisition of thermal radiative measurements of vegetative canopies. *Sensors*, 9(3), 1409–1422.
- Van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., & Su, Z. (2009). An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. *Biogeosciences*, 6, 3109–3129.
- Vountas, M., Rozanov, V., & Burrows, J. (1998). Ring effect: Impact of rotational Raman scattering on radiative transfer in earth's atmosphere. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 60(6), 943–961.
- Wanner, W., Strahler, A., Hu, B., Lewis, P., Muller, J. P., Li, X., et al. (1997). Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR Data: Theory and algorithms. *Journal of Geophysical Research*, 102, 17143–17162.
- Zhao, M., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, 95, 164–176.