

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Evaluating the predictive power of sun-induced chlorophyll fluorescence to estimate net photosynthesis of vegetation canopies: A SCOPE modeling study

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Verrelst, J., van der Tol, C., Magnani, F., Sabater, N., Rivera, J.P., Mohammed, G., et al. (2016). Evaluating the predictive power of sun-induced chlorophyll fluorescence to estimate net photosynthesis of vegetation canopies: A SCOPE modeling study. REMOTE SENSING OF ENVIRONMENT, 176(April 2016), 139-151 [10.1016/j.rse.2016.01.018].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/585571> since: 2017-05-10

*Published:*

DOI: <http://doi.org/10.1016/j.rse.2016.01.018>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

## Evaluating the predictive power of sun-induced chlorophyll fluorescence to estimate net photosynthesis of vegetation canopies: A SCOPE modeling study

Jochem Verrelst<sup>a,\*</sup>, Christiaan van der Tol<sup>b</sup>, Federico Magnani<sup>c</sup>, Neus Sabater<sup>a</sup>, Juan Pablo Rivera<sup>a</sup>, Gina Mohammed<sup>d</sup>, Jose Moreno<sup>a</sup>

<sup>a</sup> Image Processing Laboratory, Department of Earth Physics and Thermodynamics, University of Valencia, C/Catedrático José Beltrán 2, E-46980 Paterna, Valencia, Spain

<sup>b</sup> Department of Water Resources, Faculty ITC, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

<sup>c</sup> Department of Agricultural Sciences, University of Bologna, Bologna, Italy

<sup>d</sup> P&M Technologies, 66 Millwood Street, Sault Ste. Marie, Ontario, Canada

### A B S T R A C T

Progress in imaging spectroscopy technology and data processing can enable derivation of the complete sun-induced chlorophyll fluorescence (SIF) emission spectrum. This opens up opportunities to fully exploit the use of the SIF spectrum as an indicator of photosynthetic activity. Simulations performed with the coupled fluorescence–photosynthesis model SCOPE were used to determine how strongly canopy-leaving SIF can be re-lated to net photosynthesis of the canopy (NPC) for various canopy configurations. Regression analysis between SIF retrievals and NPC values produced the following general findings: (1) individual SIF bands that were most sensitive to NPC were located around the first emission peak (SIF<sub>red</sub>) for heterogeneous canopy configurations (i.e., varying biochemistry, leaf, canopy variables); (2) using two SIF retrieval bands, e.g. O<sub>2</sub>-B at 687 nm and O<sub>2</sub>-A at 760 nm, or the red and NIR emission peaks at 685 nm and 740 nm, led to stronger correlations than using only one band; (3) using the O<sub>2</sub>-B and the O<sub>2</sub>-A SIF retrieval bands was at least as effective as using the two emission peaks; (4) superior correlations were achieved by using the four main SIF retrieval bands (H<sub>α</sub>, O<sub>2</sub>-B, water vapor, O<sub>2</sub>-A); and (5) further improvements may be obtained by exploiting the full SIF profile and by using an adaptive, nonlinear regression algorithm such as Gaussian processes regression (GPR). Relationships can be due to variation in photosynthetic capacity (V<sub>cmax</sub>), but also from variation in leaf optical and canopy structural variables such as chlorophyll content and leaf area index. Overall, modeling results suggest that sampling the SIF profile in at least both O<sub>2</sub>-B and O<sub>2</sub>-A bands enables quantification photosynthetic activity of vegetation with high accuracy.

### Keywords:

Sun-induced fluorescence

Photosynthesis

SCOPE

Canopy

Regression

Band analysis

FLEX

### 1. Introduction

Sun-induced fluorescence (SIF) emitted by chlorophyll molecules is one of three main de-excitation mechanisms for energy captured by light harvesting pigments in plants. SIF emitted by vegetation is seen as a meaningful indicator of plant stress (Van Wittenberghe et al., 2013), instantaneous plant photosynthetic function (e.g., carbon fixation), and possibly gross primary productivity (GPP) at the ecosystem scale (Porcar-Castell et al., 2014).

Although the SIF flux emitted from the canopy is relatively small compared to reflected sunlight (about 1–5% in the near infrared; NIR), it is a broadband spectrum that typically spans about 650–800 nm

(but which can also extend somewhat below or above that range, e.g., 640–850 nm), and which contains useful information on the photosynthetic process (Franck, Juneau, & Popovic, 2002; Lichtenthaler & Rinderle, 1988). Even though the canopy will produce different SIF spectra under different environmental and structural conditions, the shape of SIF spectra preserves typical features. In general, the SIF spectrum is composed of two peaks, one located in the red (SIF<sub>red</sub>) spectral region with a maximum around 685 nm that is mainly attributed to the fluorescence emission of Photosystem II (PSII), and the other located in the NIR (SIF<sub>NIR</sub>) with a maximum around 740 nm that is attributable to both Photosystem I (PSI) and PSII (Baker, 2008; Papageorgiou & Govindjee, 2004).

With the advent of imaging spectrometers, the retrieval of SIF using remote sensing technologies has become a novel area of research (Alonso et al., 2007; Guanter et al., 2010; Meroni et al., 2009, 2010) aimed primarily at mapping SIF from the site-specific (Damm et al., 2014; Daumard et al., 2012; Moya, Daumard, Moise, Ounis, & Goulas,

\* Corresponding author at: Laboratory of Earth Observation (LEO), Image Processing Laboratory (IPL), Cientific Parc, BOX 22085, University of Valencia, 46071, Paterna (Valencia), Spain.

E-mail address: [jochem.verrelst@uv.es](mailto:jochem.verrelst@uv.es) (J. Verrelst).

2006; Perez-Priego, Zarco-Tejada, Miller, Sepulcre-Canto, & Fereres, 2005; Zarco-Tejada, Gonzalez-Dugo, & Berni, 2012; Zarco-Tejada, Morales, Testi, & Villalobos, 2013) to the global scale (Frankenberg et al., 2011; Joiner et al., 2011). In general, two strategies to extract SIF from passive detection methods have been pursued in recent years, exploiting either atmospheric (telluric) absorption features due to oxygen in the O<sub>2</sub>-A absorption region at 760 nm (Alonso et al., 2007; Guanter et al., 2010; Meroni et al., 2010), or solar Fraunhofer lines, which are narrow dark lines (absorption features) in the solar spectrum in which irradiance is strongly reduced (e.g. in the NIR 740–770 nm spectral window) (Frankenberg et al., 2011; Guanter et al., 2012). Regardless of the retrieval strategy, current remote sensing approaches have mostly emphasized the second emission peak region (SIF<sub>NIR</sub>) for photosynthetic quantification, especially from satellite-based atmospheric sensors. For example, Frankenberg et al. (2011) and Guanter et al. (2012) linked monthly-aggregated global SIF<sub>NIR</sub> observations with global GPP products using (biome-specific) linear relationships. More recently, at the local scale, Damm, Guanter, Paul-Limoges, et al. (2015) linked airborne SIF<sub>NIR</sub> measurements with eddy covariance flux tower GPP data and found that relationships were not linear but asymptotic when instantaneous rather than temporally aggregated measured data were used, and relationships were also ecosystem-specific.

One reason for the focus on the SIF<sub>NIR</sub> has been the absence of spaceborne sensors spectrally optimized to capture the full SIF emission spectrum. To fill this gap, the European Space Agency (ESA) has been conducting Phase A/B1 evaluations of a candidate Earth Explorer mission dedicated to measurement of SIF in terrestrial vegetation. The Fluorescence Explorer (FLEX) satellite, equipped with a Fluorescence Imaging Spectrometer (FLORIS) onboard, has recently been approved as ESA's Earth Explorer 8 mission (ESA, 2015). FLEX will operate in a tandem mission with ESA's Sentinel-3 satellite, the latter to provide atmospheric and land surface data needed for atmospheric corrections and accurate SIF characterizations. FLORIS will measure the radiance between 500 and 780 nm with a bandwidth between 0.3 nm and 2 nm (depending on wavelength), providing images with a 150 km swath and 300 m pixel size (Kraft et al., 2013; Moreno, Asner, Bach, et al., 2006). Such finely resolved spectral sampling will allow retrieval of the full broadband fluorescence emission spectrum and related products such as  $F_{\text{total}}$  (i.e., integral of the fluorescence broadband spectrum). In addition, a novel airborne imaging spectrometer *HyPlant* has become available recently which demonstrates the potential of a FLORIS-type sensor (Rascher et al., 2015). *HyPlant* has an ultra-high spectral resolution in the red and near-infrared spectral region (0.26 nm FWHM (Full Width at Half Maximum) in the 670–780 nm spectral range). This allows quantification of sun-induced fluorescence fluxes in physical units for SIF<sub>red</sub> and SIF<sub>NIR</sub> (Rossini et al., 2015), and eventually over the full SIF spectral region at a local scale. Another airborne experiment demonstrated that GPP is most strongly related to SIF<sub>red</sub> at the O<sub>2</sub>-B absorption band (Cheng et al., 2013), possibly due to the relevance of the red band to photosystem II processes (Baker, 2008).

Additional advances in signal retrieval and data processing are evident. Developments include the use of multiple absorption lines in both emission peak regions for SIF retrievals, and simulations of the influences of different atmospheric conditions (Liu & Liu, 2014). For example, Zhao et al. (2014) examined SIF retrieval in five absorption lines to allow reconstruction of the full SIF emission between 650 and 850 nm based on simulated data. Also within the FLEX scientific studies, a spectral fitting method has been developed that, when combined with an atmospheric correction algorithm, is able to reconstruct the full SIF spectrum directly from top of atmosphere (TOA) radiance data (Cogliati et al., 2015). Reconstruction of the full SIF spectrum will allow calculation of some other meaningful parameters relevant to detection of plant stress status, such as the spectral positions and FWHM of the SIF<sub>red</sub> and SIF<sub>NIR</sub> peaks, and the area under the SIF emission curve (Subhash & Mohanan, 1997; Zhao et al., 2014).

The FLEX scientific studies have also investigated radiative transfer modeling of the SIF signal through the leaf and canopy based on explicit leaf physiological descriptions. The SCOPE (Soil-Canopy Observation, Photosynthesis and Energy Balance) model (Van der Tol, Berry, Campbell, & Rascher, 2014; Van der Tol, Verhoef, Timmermans, Verhoef, & Su, 2009) has been coupled with new leaf fluorescence modules (resulting in Version 1.53). SCOPE combines the functionality of a Soil-Vegetation-Atmosphere-Transfer (SVAT) model with radiative transfer of reflected and emitted (thermal and fluorescent) radiation and enables the theoretical quantification of the canopy-leaving SIF broadband spectrum and canopy fluxes, such as the net photosynthesis of the canopy (NPC).

Now that derivation of the full SIF spectrum is possible, it creates opportunities to utilize more effectively the spectral information content related to NPC. An imminent requirement is to identify which SIF wavelengths are most sensitive to NPC. This leads to the main objective of this work: to analyze the sensitivity of single wavelengths, as well as combinations of SIF retrieval bands in estimation of NPC for various canopy configurations.

For this purpose, a SCOPE modeling study was applied. Simulations of canopy-leaving SIF and NPC outputs were conducted for different combinations of biochemical, leaf, canopy and micrometeorological variables. Simulated SIF spectra were subsequently analyzed with respect to their predictive power in estimating NPC. Specifically, the following aspects were investigated; (1) linear regression analysis between individual SIF bands and NPC outputs; (2) linear regression analysis between combined SIF bands and NPC outputs; and (3) adaptive, nonlinear machine learning regression between combined SIF bands and NPC outputs.

## 2. Materials and methods

### 2.1. SCOPE

The coupled fluorescence–photosynthesis model SCOPE simulates photosynthesis, radiative transfer in the leaf and canopy, and surface energy balance (Van der Tol et al., 2009, 2014). SCOPE recently became a virtual laboratory for studies on surface energy balance (Timmermans, Su, Van der Tol, Verhoef, & Verhoef, 2013), remote sensing thermal infrared measurements (Duffour, Olioso, Demarty, Van der Tol, & Lagouarde, 2015), and SIF-photosynthesis studies (Damm, Guanter, Paul-Limoges, et al., 2015; Verrelst et al., 2015; Zhang, Guanter, et al., 2014).

For photosynthesis, SCOPE uses either the model of Von Caemmerer (2000, 2013) or Collatz, Ball, Grivet, and Berry (1991), Collatz, Ribas-Carbo, and Berry (1992). These physiological models originally were developed to interpret measurements of leaf gas exchange. The main boundary conditions for photosynthesis are energy supply (light) and carbon dioxide diffusion into the leaf. The models calculate photosynthesis under the condition that these two aspects, the energy supply and the carbon dioxide flux, are in equilibrium. Electron transport (i.e., transfer of the energy supply) is calculated classically from active fluorescence measurements (e.g. Genty, Wonders, & Baker, 1990; Maxwell & Johnson, 2000; Weis & Berry, 1987). However, active techniques typically are only feasible in a laboratory or small scale field study due to requirements for saturating light flashes, and for modulation of the measuring light beam when fluorescence is assessed in natural outdoor conditions (Maxwell & Johnson, 2000; Van der Tol et al., 2014). For this reason, the photosynthesis model in SCOPE is complemented with these alternative predictive models for fluorescence that serve to simulate the fluorescence leaf emission efficiency,  $\epsilon$ , as a function of weather conditions and photosynthesis parameters, normalized by the leaf fluorescence emission efficiency in (near) dark or pre-dawn conditions. The model of Van der Tol et al. (2014), is a semi-empirical model based upon field and laboratory experiments of unstressed

and drought-stressed vegetation and hereon referred to as TB12 and TB12-D models, respectively. The model of Magnani et al. (2009), Dayyoub, 2011), hereon referred to as MD12 model, has a more explicit parameterization of fluorescence quenching mechanisms. In its most recent development by Magnani (summarized in Mohammed et al., 2014), the module also incorporates effects on photosynthesis and fluorescence of seasonal changes in PSII photoinhibition and sustained, dark-adapted non-photochemical quenching (NPQ) (Porcar-Castell, 2011). The MD12 module is not limited to empirical calibration (compared to the TB12 models) and is therefore able to reproduce intermediate conditions using two additional variables, those being the rate constant of sustained thermal dissipation (kNPQs) and the fraction of functional reaction centers (qLs) (Porcar-Castell, 2011).

The radiative transfer of incident light and SIF emission in the leaf is handled with a separate sub-model, Fluspect. The Kubelka–Munk theory is applied to the mesophyll layer of the leaf to calculate diffuse scattering and absorption of both incident light, as in the model PROSPECT (Jacquemoud & Baret, 1990), and SIF. Two SIF spectra as published by Franck et al. (2002), one for PSI and another for PSII, are used to convert incident light at any depth in the leaf into fluorescence emission spectra for both photosystems. The output of Fluspect consists of reflectance and transmittance, and four matrixes  $\mathbf{M}$  that quantify the probability of incident light (between 400 and 750 nm) to return as SIF (between 640 and 850 nm) from PSI and PSII at the illuminated and shaded side of the leaf. In the calculation of  $\mathbf{M}$ , values for the emission efficiencies of 0.002 for PSI and 0.01 for PSII (the latter in dark adapted conditions) have been assumed.

The output of Fluspect is combined with the fluorescence emission efficiency scale factor  $\varepsilon$  to obtain the leaf SIF spectra:

$$\begin{aligned} F_f &= (\varepsilon \mathbf{M}_{f,PSII} + \mathbf{M}_{f,PSI})E \\ F_g &= (\varepsilon \mathbf{M}_{g,PSII} + \mathbf{M}_{g,PSI})E \end{aligned} \quad (1)$$

Where  $E$  is the incident irradiance vector (400–750 nm) on the leaf ( $\text{W m}^{-2} \text{sr}^{-1} \mu\text{m}^{-1}$ ). The variable  $\varepsilon$  scales the PSII SIF spectrum as a function of micrometeorological conditions and photosynthetic parameters. The separation between the biochemical model on the one hand (for  $\varepsilon$ ) and Fluspect on the other hand (for  $\mathbf{M}$ ) has the advantage that the effects of scattering and re-absorption due to leaf structure (the matrixes  $\mathbf{M}$ ) can be calculated in advance, because they are functions of leaf structure only and do not depend on micrometeorological conditions.

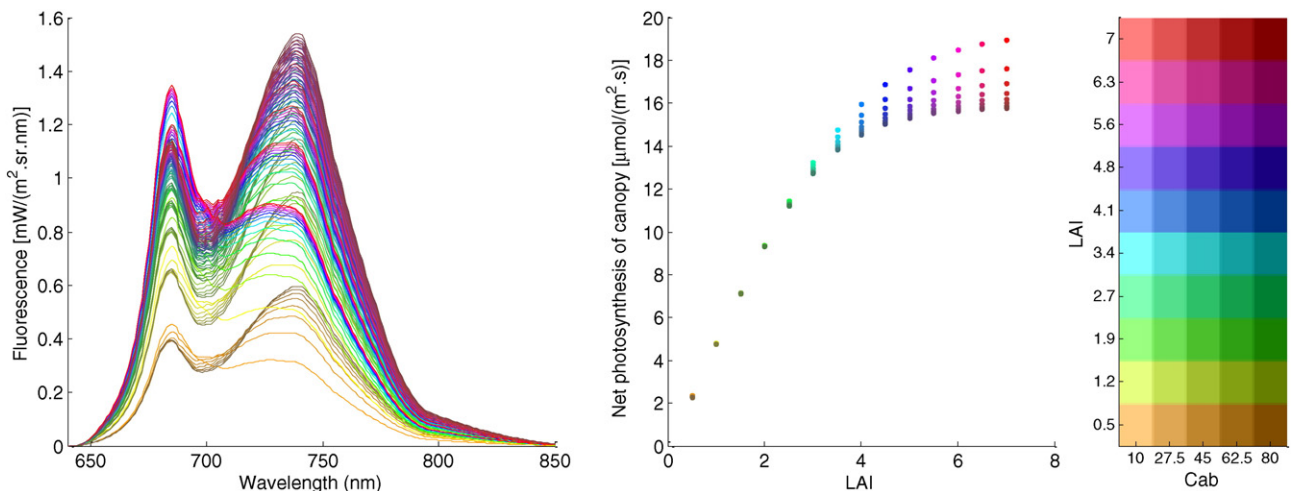
In Eq. (1) it is assumed that the micrometeorologically induced variations of the fluorescence emission ( $\varepsilon$ ) affect only PSII. This assumption

is crucial as it determines the wavelength dependence of the sensitivity of SIF to micrometeorological conditions and photosynthesis parameters. PSI emits SIF only in the near-infrared part of the spectrum, while PSII emits over a wide spectrum (640–800 nm) and peaks in the red SIF. Assuming a constant PSI fluorescence implies that  $\text{SIF}_{\text{red}}$  is more sensitive to photosynthetic parameters and weather conditions than is  $\text{SIF}_{\text{NIR}}$  (Porcar-Castell et al., 2014).

Both the incident light on the individual leaves ( $E$ ) and the propagation of SIF throughout the canopy are calculated with radiative transfer models based on the Scattering of Arbitrarily Inclined Leaves (SAIL) model (Verhoef, 1984). SAIL is a 1-D vertical model that distributes the canopy into 60 horizontal layers with an optical thickness of  $\text{LAI}/60$  (where LAI is leaf area index). Within each layer, leaves are described by their optical properties (from Fluspect) and their inclination. The zenith inclination distribution of the leaves is described stochastically with user defined parameter values, and distribution of leaf orientations in the azimuthal (horizontal) direction is uniform. SCOPE also simulates thermal radiation, net radiation, soil heat flux and the turbulent heat fluxes through an aerodynamic resistance scheme. Leaf temperature and leaf boundary layer gas concentrations are simulated as well.

SCOPE also simulates a diversity of fluxes, one of which is NPC. Net photosynthesis is the total gross photosynthesis minus the flux of  $\text{CO}_2$  associated with foliage respiration (photorespiration and ‘dark’ respiration), or gross primary productivity less the dark respiration of the foliage. NPC is calculated here by simply aggregating the photosynthesis over the leaf area of the canopy, because photosynthesis in SCOPE is the gas exchange flux of  $\text{CO}_2$  between atmosphere and leaf, which is a scalar quantity. NPC from SCOPE may be used to compute GPP for approximate comparisons with that derived from eddy covariance (EC) flux measurements over canopies, for example by setting the respiration parameter to zero, or by summing the net photosynthesis and leaf dark respiration; however, the GPP from both approaches will not be precisely identical because of differences in specific assumptions and methodological approaches used in EC-derived GPP.

An example of bidirectional, canopy-leaving SIF and NPC simulations for ranging leaf chlorophyll content (Cab) and LAI is shown in Fig. 1. Cab and LAI were earlier identified as key driving variables in governing canopy-leaving SIF spectra (Verrelst et al., 2015). That can also be observed in the plotted SIF spectra. SIF is especially governed by LAI, with a higher LAI (i.e., more leaves within a given area) leading to a more pronounced SIF signal. The same is true for NPC, where the occurrence of a higher LAI is linked with a higher NPC, i.e. greater photosynthetic capacity at the canopy scale.



**Fig. 1.** Example of full SIF profiles (left) and net photosynthesis of the canopy (NPC) (right) simulated by SCOPE as a function of leaf area index (LAI; 14 samples) and leaf chlorophyll content (Cab; 8 samples).



## 2.2. SIF retrieval bands

### 2.2.1. Single bands for SIF retrievals

Current SIF retrieval methods may be considered as single-line in-filling approaches that are applicable where absorption lines are present. Four commonly used absorption lines were chosen: the telluric atmospheric absorption lines of O<sub>2</sub>-A (centered at approximately 760 nm), O<sub>2</sub>-B (687 nm), and water vapor (719 nm), and the solar Fraunhofer line of H $\alpha$  (656 nm) (Zhao et al., 2014). Other absorption lines such as solar Fraunhofer lines around 755–759 nm were not selected because they were considered too narrow and of insufficient depth for capturing the subtle SIF signal with sufficient accuracy. Additionally, the peaks of the red emission (centered at 685 nm) and the near-infrared emission (centered at 740 nm), and the mid-valley (centered at 699 nm) between the peaks were also selected (Fig. 2, Table 1).

### 2.2.2. Combined bands

Assuming the successful retrieval of single bidirectional SIF bands, the following combinations were analyzed for their predictive value: (1) O<sub>2</sub>-B and O<sub>2</sub>-A absorption lines, (2) the four main absorption lines (H $\alpha$ , O<sub>2</sub>-B, water vapor, O<sub>2</sub>-A), (3) the two peaks (red, NIR), (4) the SIF peak-ratio (red:NIR), which is an indicator of chlorophyll content and plant status (Hak, Lichtenthaler, & Rinderle, 1990; Pedrós, Goulas, Jacquemoud, Louis, & Moya, 2010), and (5) the two peaks and mid-valley. Additionally, (6) given the possibility to reconstruct the full SIF profile, it is possible to calculate  $F_{\text{total}}$  (taken here as the integrated hemispherical SIF from 641 to 850 nm). Finally, (7) the reconstruction of the full SIF profile also enables inclusion of all single bands (from 650 to 790 nm was used here) in the regression analysis. An overview is provided in Table 2.

## 2.3. Canopy configurations

SIF–NPC relationships were assessed for various canopy configurations based on the variables that are known to be drivers of SIF emission characteristics. In order to identify driving variables a global sensitivity (GSA) analysis on SCOPE SIF simulations was conducted in a related study (Verrelst et al., 2015). Variance-based GSA explores the full input variable space and evaluates the relative importance of each input variable in a model (Saltelli, Tarantola, & Chan, 1999). The method can be used to identify the most influential variables affecting model outputs. In variance-based GSA the contribution of each input variable to the variation in outputs is averaged over the variation of all input variables, i.e., all input variables are changed together (Saltelli et al., 1999).

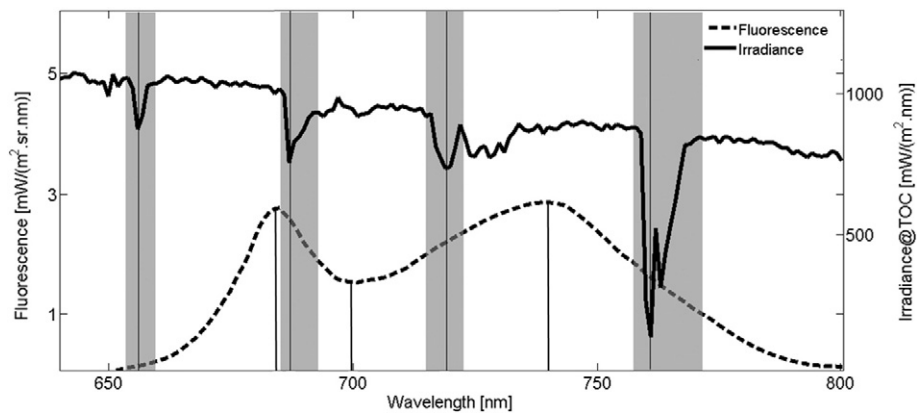
**Table 1**

Single absorption lines and features tested in regression analysis.

Index	Element	Central wavelength (nm)	Spectral range (nm) absorption lines
1	H $\alpha$ absorption line	656	653–662
2	Red peak (attributed to SIF emission of Photosystem II)	685	
3	O <sub>2</sub> -B absorption line	687	683–692
4	Mid-valley between red and NIR peaks	699	
5	Water vapor absorption line	719	714–722
6	Near-infrared peak (attributed to SIF emission of Photosystem I and to PSII)	740	
7	O <sub>2</sub> -A absorption line	760	757–771

In Verrelst et al. (2015), the method of Saltelli et al. (2010) was used to identify the driving variables that shape the variability of canopy-leaving SIF spectrum across its full spectral range. The method has been demonstrated to be effective in identifying both the main sensitivity effects (first-order effects, i.e., the contribution to the variance of the model output by each input variables,  $S_i$ ) and total sensitivity effects (the first-order effects plus interactions with other input variables,  $S_{Ti}$ ) of input variables. In this study, the GSA analysis was extended to include the integrated SIF ( $F_{\text{total}}$ ) from 641 to 850 nm. As such, the relative contribution of each input variable to SIF can be disentangled and quantified. The MD12 biochemical sub-model was used here, based on earlier performance comparisons to the TB12 modules, and keeping KNPQs and qLs constant (Mohammed et al., 2014). Further, the full variable space of SCOPE was analyzed for a spherical leaf angle distribution and without varying soil variables. See Table 3 and Verrelst et al. (2015) for details on the model variable boundaries. In order to catch the full model variability, it must be noted that some variable boundaries likely exceed normal real-world situations (for instance, the variation of CO<sub>2</sub> concentration in the air is considerably higher than would be expected under normal circumstances). Variables were sampled according to Sobol's quasi-random sequence generator. In total,  $(N(k + 2))$  model simulations were run, where  $N$  is the sample size and equals 2000, and  $k$  is the number of input variables and equals 25. This produced 54,000 simulations. Only total order sensitivity effects ( $S_{Ti}$ ) expressed as percentages were considered.

Fig. 3 provides the  $S_{Ti}$  results of the SCOPE v.1.53 input variables. Most variables exerted a negligible effect on  $F_{\text{total}}$ . The driving variables were: maximum carboxylation capacity at optimum temperature ( $V_{\text{cmo}}$ ) (sometimes referred to as  $V_{\text{cmax},25}$  in the literature, where 25 stands for an optimum temperature of 25 °C), dry matter content



**Fig. 2.** A typical incident solar irradiance spectrum at top of canopy simulated by MODTRAN-5 for a standard mid-latitude summer model atmosphere, and the default rural aerosol model. Both the spectral resolution and spectral sampling interval are 1 nm. Four main absorption lines used for SIF retrieval are shown in the figure. Their spectral wavelengths (spectral positions of the maximum absorption) and spectral ranges are marked with vertical lines and gray shade, respectively, which are further specified in Table 1. SIF radiance leaving the canopy and simulated by SCOPE has been added.

**Table 2**

Combined SIF absorption lines and features used in regression analysis.

Index	Combined wavelengths	Wavelengths (nm)
1	O <sub>2</sub> -B and O <sub>2</sub> -A absorption lines	687, 760
2	H $\alpha$ , O <sub>2</sub> -B, water vapor absorption lines, and O <sub>2</sub> -A	656, 687, 719, 760
3	Two SIF emission peaks	685, 740
4	Peak ratio	685/740
5	Two SIF emission peaks and mid-valley	685, 699, 740
6	F <sub>total</sub>	Hemispherically and spectrally integrated SIF at the TOC (from 641 to 850 nm)
7	F <sub>all</sub>	All individual bidirectional SIF wavelengths (from 650 to 790 nm)

(Cdm), Cab, LAI, canopy height (hc), within-canopy-layer resistance (rwc), air pressure (P), atmospheric vapor pressure (ea), atmospheric CO<sub>2</sub> concentration (Ca), air temperature (Ta), and broadband incoming shortwave radiation (Rin). Altogether these variables explained 97.5% of the total variance (taking interactions into account).

Because SCOPE is a SVAT model, analysis of input variables to the SIF–NPC relationships can be undertaken at various scales and contexts, e.g., biochemical (physiological), leaf, canopy, geometrical, and/or micrometeorological. A balance must be struck between including a sufficient number of ranging variables to achieve good representation of reality but not so many variables that the escalating heterogeneity becomes uninterpretable. Various canopy configurations with increasing heterogeneity were generated. First, only the main driving biochemical variable was ranged, then more variables were ranged at biochemical, leaf, and canopy scales, eventually

varying all variables at all scales. The driving variables taken as a starting point were V<sub>cmo</sub> at the biochemical scale, Cab at leaf scale, and LAI at canopy scale. The number of ranging variables was then increased to eventually produce a total of 12 canopy configurations, and then lastly all SCOPE variables were ranged (Table 4). Each variable was randomly sampled 2000 times within their minimum–maximum boundaries according to Table 3 (see also Verrelst et al., 2015). From the combined variable space, a uniform random set of 2000 simulations was selected. In the unlikely event that the radiative transfer equations were unsuccessfully resolved, e.g., due to unrealistic variable combinations, another random simulation was taken. Finally, from all SCOPE output variables the NPC output flux, the bidirectional SIF spectra and F<sub>total</sub> were collected.

## 2.4. Regression analysis

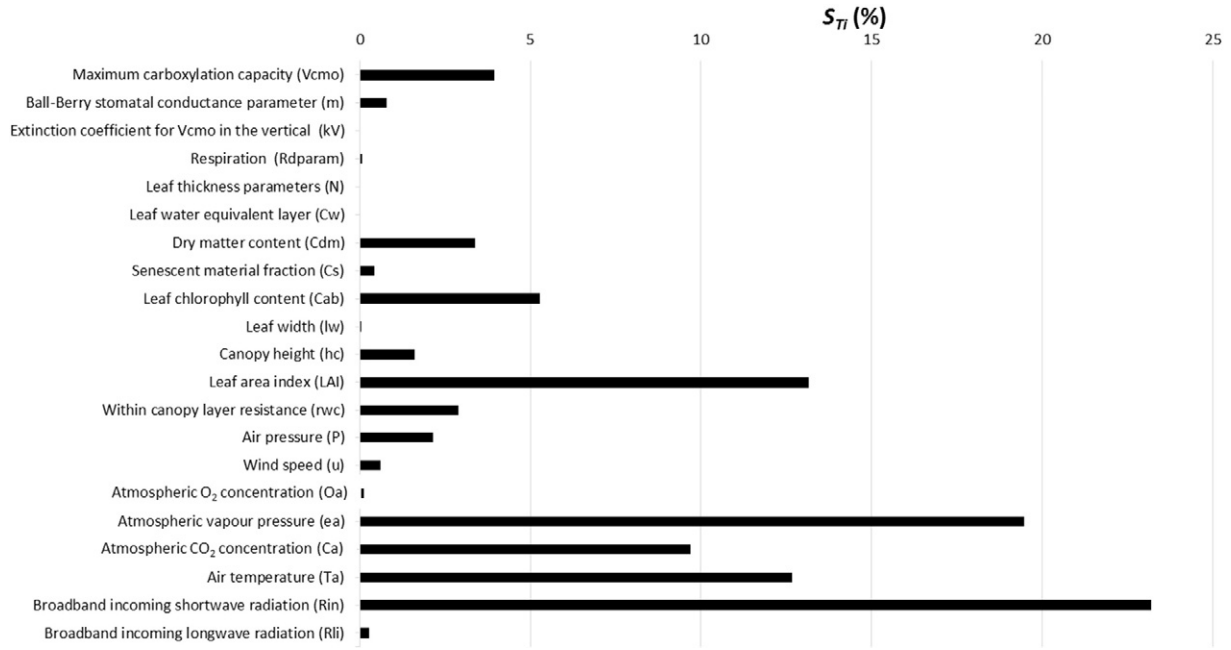
To enable estimation of NPC from SIF retrievals, a regression analysis was used. For each canopy configuration a random subset of 2000 simulations was selected and then split into 50% for regression model calibration and 50% for validation. An ordinary least squares linear regression (LR) and an adaptive, nonlinear regression algorithm regression called Gaussian processes regression (GPR) were applied. LR was used to analyze all bands individually and in combinations. GPR was applied to determine whether and how much improvement could be achieved for the combined band analyses.

GPR is a machine learning regression technique in a Bayesian framework equivalent to kernel ridge regression, least squares support vector machine and kriging (Rasmussen & Williams, 2006). In a previous study (Verrelst, Muñoz, et al., 2012), GPR outperformed neural networks, support vector regression and kernel ridge regression for the majority of biophysical parameter retrievals. The GPR model establishes a relation

**Table 3**

Input variables and their boundaries and default values of the SCOPE model. Soil and aerodynamic variables and kNPQs, qLs and leaf angle distribution have been kept to their default values. See also Verrelst et al., 2015.

Input	Definition	Unit	Min	Max	Default
<i>Leaf biochemistry</i>					
V <sub>cmo</sub>	Maximum carboxylation capacity (at optimum temperature)	μmol m <sup>-1</sup> s <sup>-1</sup>	0	200	30
m	Ball–Berry stomatal conductance parameter	[–]	2	20	8
Rdparam	Parameter for dark respiration (Rd = Rdparam * V <sub>cmo</sub> )	[–]	0.001	0.03	0.015
kV	Extinction coefficient for a vertical profile of V <sub>cmo</sub> (maximum value of V <sub>cmo</sub> occurs at the top of the canopy).	[–]	0	0.8	0.64
<i>Leaf optical</i>					
N	Mesophyll structural parameter in PROSPECT	[–]	1	2.5	1.4
Cw	Water content in PROSPECT	g cm <sup>-2</sup>	0	0.1	0.009
Cdm	Dry matter content in PROSPECT	g cm <sup>-2</sup>	0	0.05	0.012
Cs	Senescence factor in PROSPECT	[–]	0	0.9	0
Cab	Chlorophyll content in PROSPECT	μg cm <sup>-2</sup>	0	80	40
<i>Canopy</i>					
lw	Leaf width	m	0.01	0.1	0.1
LIDFa	LIDF parameter a, which controls the average leaf slope	[–]	–1	1	–0.35
LIDFb	LIDF parameter b, which controls the distribution's bimodality	[–]	–1	1	–0.15
LAI	Leaf area index	m <sup>2</sup> m <sup>-2</sup>	0	7	3
hc	Canopy height	m	0.1	2	1
<i>Micrometeorological</i>					
p	Air pressure	hPa	300	1090	970
u	Wind speed	m s <sup>-1</sup>	0	50	2
Oa	O <sub>2</sub> concentration in the air	ppm	0	220	209
ea	Atmospheric vapor pressure	hPa	0	150	15
Ca	CO <sub>2</sub> concentration in the air	ppm	50	1000	380
Ta	Air temperature	°C	–10	50	20
Rin	Incoming shortwave radiation	W m <sup>-2</sup>	0	1400	600
Rli	Incoming longwave radiation	W m <sup>-2</sup>	0	400	300
<i>Geometry</i>					
VZA	Viewing zenith angle	Degree	0	10	0
RAA	Relative azimuth angle	Degree	0	180	0
SZA	Sun zenith angle	Degree	0	60	30



**Fig. 3.** Driving variables of  $F_{\text{total}}$  as identified by Global Sensitivity Analysis for SCOPE biochemistry (without KNPQs, qLs), leaf, canopy (without leaf angle distribution) and micrometeorology variables.

between the input (B-bands spectra)  $x \in \mathbb{R}^B$  and the output variable (leaf parameter)  $y \in \mathbb{R}$  of the form:

$$\hat{y} = f(x) = \sum_{i=1}^N \alpha_i K(x_i, x), \quad (2)$$

where  $\{x_i\}_{i=1}^N$  are the spectra used in the training phase,  $\alpha_i$  is the weight assigned to each one of them, and  $K$  is a sophisticated kernel function evaluating the similarity between the test spectrum and all  $N$  training spectra (Verrelst, Alonso, Camps-Valls, Delegido, & Moreno, 2012; Verrelst, Muñoz, et al., 2012). A scaled Gaussian kernel function was used,

$$K(x_i, x_j) = v \exp \left( - \sum_{b=1}^B \frac{(x_i^{(b)} - x_j^{(b)})^2}{2\sigma_b^2} \right), \quad (3)$$

where  $v$  is a scaling factor,  $B$  is the number of bands, and  $\sigma_b$  is a dedicated parameter controlling the spread of the relations for each particular spectral band  $b$ . Model parameters ( $v$ ,  $\sigma_b$ ) and model weights  $\alpha_i$  can

be automatically optimized by maximizing the marginal likelihood in the training set (Rasmussen & Williams, 2006).

The accuracies of LR and GPR models were validated using the validation dataset with coefficient of determination ( $R^2$ ) and the root-mean-square error (RMSE) calculated between retrieved and “true” (simulated) NPC values.

### 3. Results

#### 3.1. Single band analysis

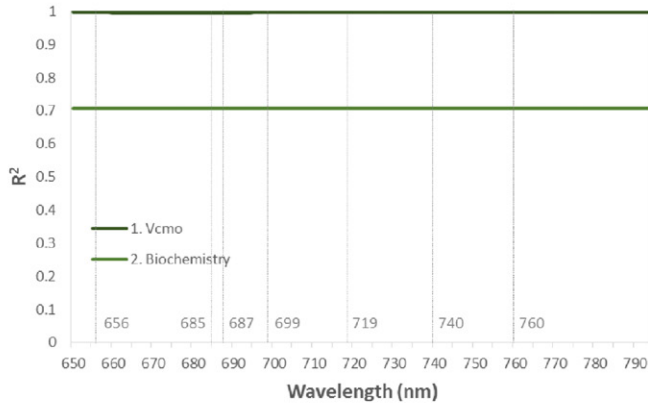
The sensitivity to NPC of each wavelength within the SIF emission spectrum was first assessed. LR was used and the  $R^2$  of validation data plotted in Fig. 4. Results are organized according to ranging variables at the scales of biochemistry, leaf, canopy, and at all SCOPE scales. The following trends were observed:

When ranging only variables at *biochemical scale* (Fig. 4a),  $V_{\text{cmo}}$  was the main variable driving NPC and produced a very strong relationship ( $R^2$  of 0.99). Relationships with NPC weakened when the other biochemical variables ( $m$ ,  $R_{\text{dparam}}$ ,  $kV$ ) were also varied, but impacts on

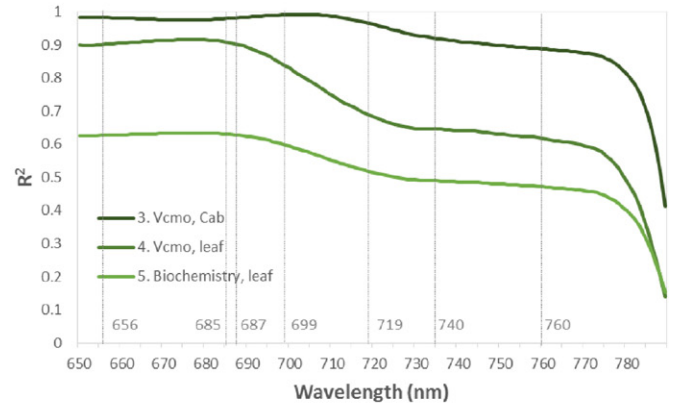
**Table 4**  
SCOPE canopy configurations with ranging variables. (See Table 3 for definitions of the variables.)

Index	Ranging variables	Justification
1	$V_{\text{cmo}}$	$V_{\text{cmo}}$ is the main biochemical driver of photosynthesis. Hence, this is the theoretical baseline when SIF is not influenced by any other variable.
2	Biochemistry	All biochemical variables ( $V_{\text{cmo}}$ , $m$ , $R_{\text{dparam}}$ , $kV$ ). Represents the most heterogeneous situation at the biochemical scale.
3	$V_{\text{cmo}}$ , Cab	Driving biochemical and leaf variables.
4	$V_{\text{cmo}}$ , leaf	Driving biochemical variable and all leaf variables ( $N$ , $Cw$ , $C_{\text{dm}}$ , $Cs$ , $Cab$ ).
5	Biochemistry, leaf	All biochemical and leaf variables. Represents the most heterogeneous situation at biochemical and leaf scales ( $V_{\text{cmo}}$ , $m$ , $R_{\text{dparam}}$ , $kV$ , $N$ , $Cw$ , $C_{\text{dm}}$ , $Cs$ , $Cab$ ).
6	Cab, LAI	Driving leaf and canopy variables.
7	$V_{\text{cmo}}$ , LAI	Driving biochemical variable ( $V_{\text{cmo}}$ ) with driving canopy variable (LAI)
8	$V_{\text{cmo}}$ , canopy	Driving biochemical variable ( $V_{\text{cmo}}$ ) with all varying canopy variables (LAI, $lw$ , $hc$ ).
9	$V_{\text{cmo}}$ , $N$ , $Cw$ , $C_{\text{dm}}$ , $Cs$ , $Cab$ , LAI, $hw$ , $hc$ (spherical LIDF)	Driving biochemical variable ( $V_{\text{cmo}}$ ) with all leaf and all canopy ( $N$ , $Cw$ , $C_{\text{dm}}$ , $Cs$ , $Cab$ , LAI, $lw$ , $hc$ ).
10	Biochemistry, leaf, canopy	All biochemical, leaf and canopy variables ( $V_{\text{cmo}}$ , $m$ , $R_{\text{dparam}}$ , $kV$ , $N$ , $Cw$ , $C_{\text{dm}}$ , $Cs$ , $Cab$ , LAI, $lw$ , $hc$ ). Represents the most heterogeneous situation at the canopy scale
11	Key SCOPE variables driving SIF	$V_{\text{cmo}}$ , $C_{\text{dm}}$ , $Cab$ , LAI, $hc$ , $rwc$ , $P$ , $ea$ , $Ca$ , $Ta$ , $R_{\text{in}}$ . These variables and their interactions explain 97.5% of the variability in $F_{\text{total}}$ .
12	All SCOPE variables	All SCOPE variables ( $V_{\text{cmo}}$ , $m$ , $R_{\text{dparam}}$ , $kV$ , $N$ , $Cw$ , $C_{\text{dm}}$ , $Cs$ , $Cab$ , LAI, $lw$ , $hc$ , $rwc$ , $P$ , $u$ , $Oa$ , $ea$ , $Ca$ , $Ta$ , $R_{\text{in}}$ , $R_{\text{li}}$ , $VZA$ , $RAA$ , $SZA$ ). Represents the most heterogeneous configuration.

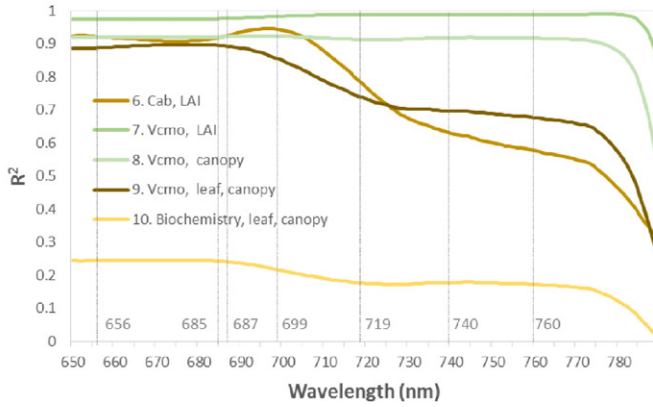
a: *biochemistry scale*



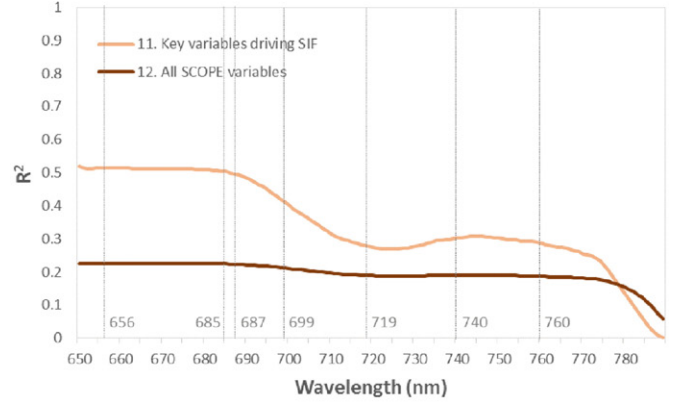
b: *leaf scale*



c: *canopy scale*



d: *all SCOPE*



**Fig. 4.** Strength of relationships between SIF and NPC for single wavelengths in the spectral emission profile. Regression analysis results ( $R^2$ ) of single wavelengths across the 650–790 nm spectral range were analyzed for different canopy configurations as outlined in Table 3.

$R^2$  were spectrally invariant. Although it should be noted that even though  $R^2$  was spectrally invariant, the slope of the relationship varied with wavelength (not shown) as  $SIF_{red}$  is affected more strongly by biochemical variables than is  $SIF_{NIR}$ .

When SCOPE variables were ranged at the *leaf scale* (Fig. 4b), varying the driving variables Cab and  $V_{cmo}$  produced strong relationships with NPC. However, relationships degraded at wavelengths beyond the red emission peak. The relationship degraded further when additionally varying other leaf variables, and a clear distinction between the first and second peak was observed. After 780 nm relationship degraded until zero approaching the end of PSII spectrum. It should be emphasized that the effect of the biochemical parameters ( $V_{cmo}$ ,  $m$ ,  $R_{param}$ ,  $kV$ ) on SIF is through parameter  $\epsilon$  in Eq. (1) (Van der Tol et al., 2014). Due to the fact that  $\epsilon$  applies only to PSII, which has its peak in the red fluorescence, the effect of biochemical parameters is the strongest in the red fluorescence. When also ranging all biochemical variables, correlations weakened further and the difference between the two SIF emission peaks was less prominent.

When SCOPE vegetation variables were subsequently ranged at the *canopy scale* (Fig. 4c), strong spectrally invariant relationships were obtained only in the case of  $V_{cmo}$  plus LAI and other canopy variables. The relationship broke down for the NIR emission peak when leaf variables were ranged. Also strong relationships could be obtained when varying only the driving leaf and canopy variables Cab and LAI, but these relationships also broke down for the second peak region. When simulating

fully heterogeneous canopies, and also ranging other biochemical variables then correlations degraded to such an extent that meaningful relationships could not be derived.

Combining input variables at *all SCOPE scales* (Fig. 4d), i.e., including also micrometeorological variables, caused relationships to deteriorate, especially for the second peak. When considering the driving SIF variables as identified by the GSA exercise (see Fig. 3), the first peak achieved a maximal  $R^2$  of about 0.5, whereas for the second peak it did not exceed about 0.3. Given that this dataset was mainly generated from varying the key micrometeorological variables ( $P$ ,  $ea$ ,  $Ca$ ,  $Ta$ ,  $Rin$ ), results suggest that these factors can weaken SIF–NPC relationships. Varying all SCOPE variables no longer produced meaningful relationships. This finding underlines the utility of applying a GSA study in order to constrain the number of input variables in an optimized way.

Table 5 summarizes the predictive strength of the most important SIF spectral bands (as identified in Table 1) for estimating NPC. Additionally the wavelength that produced the strongest correlation is shown. The following observations can be made. Overall, the red peak,  $O_2$ -B, and  $H\alpha$  line showed similar predictive strength. The NIR peak and  $O_2$ -A were also similar in performance. In most instances the red peak or  $O_2$ -B band were better predictors than the NIR peak or  $O_2$ -A. The water vapor band was not as strong a predictor as the three red bands or the mid-valley, though it was superior to the NIR peak and  $O_2$ -A. The best performing wavelength differed by scenario and spanned



**Table 5**  
Sensitivity of single bands and spectral features under different canopy configurations using linear regression. Best performing  $R^2$  per canopy configuration is shown in bold-face text. The  $R^2$  cells that fall within 0.01 interval of the best performing  $R^2$  are highlighted in gray.  $R^2$  of the best performing wavelength is also indicated (far right column).

	Ranging SCOPE variables	H $\alpha$ (656 nm)		Redpeak (685 nm)		O2-B (687 nm)		Mid-valley (699 nm)		Water vapor (719 nm)		NIR peak (740 nm)		O2-A (760 nm)		Best wavelength (nm)	$R^2$
		$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE		
1	V <sub>cmo</sub>	0.9970	0.7484	0.9966	0.7948	0.9966	0.7890	0.9971	0.7361	0.9975	0.6760	0.9977	0.6491	0.9978	0.6426	790	0.9978
2	biochemistry	0.7070	8.5148	0.7065	8.5221	0.7066	8.5207	0.7072	8.5118	0.7079	8.5031	0.7082	8.4986	0.7082	8.4975	790	0.7087
3	V <sub>cmo</sub> , Cab	0.9830	1.7376	0.9801	1.8811	0.9819	1.7934	0.9911	1.2605	0.9652	2.4903	0.9110	3.9809	0.8883	4.4594	703	0.9922
4	V <sub>cmo</sub> , leaf	0.9026	4.3075	0.9092	4.1596	0.9040	4.2746	0.8371	5.5674	0.6887	7.6947	0.6415	8.2583	0.6175	8.5301	676	0.9159
5	biochemistry, leaf	0.6275	9.9114	0.6309	9.8653	0.6288	9.8938	0.5980	10.2940	0.5178	11.2759	0.4863	11.6382	0.4720	11.8004	677	0.6337
6	Cab, LAI	0.9208	1.4306	0.9197	1.4411	0.9257	1.3864	0.9438	1.2047	0.7789	2.3754	0.6297	3.0730	0.5772	3.2838	696	0.9459
7	V <sub>cmo</sub> , LAI	0.9744	2.3695	0.9760	2.2963	0.9766	2.2664	0.9829	1.9356	0.9880	1.6217	0.9869	1.6966	0.9875	1.6581	777	0.9895
8	V <sub>cmo</sub> , canopy	0.9199	3.4569	0.9211	3.4316	0.9215	3.4221	0.9227	3.4001	0.9132	3.6086	0.9179	3.5075	0.9166	3.5353	696	0.9232
9	V <sub>cmo</sub> , leaf, canopy	0.8879	4.2744	0.8947	4.1411	0.8925	4.1835	0.8540	4.8753	0.7377	6.5325	0.6985	7.0273	0.6768	7.2514	678	0.8974
10	Biochemistry, leaf, canopy	0.2453	29.3064	0.2429	29.3539	0.2411	29.3867	0.2168	29.8472	0.1773	30.5901	0.1782	30.5714	0.1727	30.6767	650	0.2462
11	Key variables driving SIF	0.5153	15.2639	0.5030	15.4557	0.4973	15.5454	0.4109	16.8229	0.2785	18.6249	0.3020	18.3188	0.2866	18.5130	650	0.5190
12	All SCOPE variables	0.2260	39.4620	0.2249	39.4886	0.2241	39.5090	0.2120	39.8133	0.1896	40.3767	0.1902	40.3663	0.1869	40.4500	650	0.2263

**Table 6**  
The sensitivity of combined bands and spectral features under different canopy configurations using linear regression. Best performing  $R^2$  per canopy configuration is shown in bold-faced text. The  $R^2$  cells that fall within 0.01 interval of the best performing  $R^2$  are highlighted in gray.

	Ranging SCOPE variables	O <sub>2</sub> -B, O <sub>2</sub> -A: 687, 760 nm		H $\alpha$ , O <sub>2</sub> -B, water vapor, O <sub>2</sub> -A: 656, 687, 719, 760 nm		Two peaks: 685, 740 nm		Peak ratio: 685/740		Two peaks and valley: 685, 699, 740 nm		F <sub>total</sub> : Integrated SIF (from 641 to 850 nm)		F <sub>all</sub> : All individual SIF wavelengths (from 650 to 790)	
		$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE
1	V <sub>cmo</sub>	0.9965	0.8255	0.9970	0.7412	0.9965	0.7869	0.8686	4.9370	0.9966	0.7831	0.9983	0.5665	1	0.0843
2	biochemistry	0.7234	8.0229	0.7111	8.1849	0.7195	8.1690	0.5874	9.8065	0.7141	8.3908	0.7201	8.2419	0.7011	8.4288
3	V <sub>cmo</sub> , Cab	0.9965	0.7792	0.9982	0.5657	0.9963	0.8123	0.4721	9.6652	0.9966	0.7816	0.9900	1.3493	0.9991	0.4056
4	V <sub>cmo</sub> , leaf	0.9333	3.5056	0.9462	3.0807	0.9305	3.6105	0.3062	11.1958	0.9573	2.8784	0.6467	8.0626	0.9783	2.0221
5	biochemistry, leaf	0.6356	9.7534	0.6205	10.2080	0.6337	9.8866	0.4056	12.3210	0.6423	9.3908	0.5176	10.9974	0.6828	8.9835
6	Cab, LAI	0.9529	1.1083	0.9354	1.2673	0.9547	1.0751	0.1955	4.5583	0.9434	1.2164	0.9426	2.0248	0.9728	0.8098
7	V <sub>cmo</sub> , LAI	0.9907	1.4723	0.9853	1.7952	0.9898	1.5035	0.7240	7.7957	0.9770	2.2615	0.9923	1.3046	0.9968	0.8415
8	V <sub>cmo</sub> , canopy	0.9315	3.2390	0.9333	3.2096	0.9321	3.2583	0.6404	7.3998	0.9244	3.4306	0.9105	3.7510	0.9490	2.8608
9	V <sub>cmo</sub> , leaf, canopy	0.8950	4.1334	0.9023	4.0438	0.9077	3.9263	0.3495	10.2994	0.8982	4.1313	0.7157	6.8358	0.9169	3.7433
10	Biochemistry, leaf, canopy	0.2356	31.8478	0.2388	30.8678	0.2348	30.8049	0.1152	33.2900	0.2144	30.2922	0.1292	32.3184	0.2805	30.4726
1	Key variables driving SIF	0.4581	13.7917	0.5068	13.6618	0.5342	12.5735	0.2484	15.8487	0.5407	12.6540	0.3078	16.1429	0.5693	12.4038
12	All SCOPE variables	0.2234	41.6271	0.2112	45.0060	0.2446	40.2953	0.1260	41.4454	0.2091	41.6138	0.2278	35.7368	0.2275	40.8113

the entire SIF emission spectrum. In realistic canopy scenarios (i.e., with ranging variables at scales of biochemistry, leaf and canopy; scenario 11) the best performing wavelength was situated on the slope before the first peak.

### 3.2. Combined bands analysis

Table 6 shows the predictive power of combined SIF wavelengths. Combining the O<sub>2</sub>-A and O<sub>2</sub>-B bands or the red and NIR peaks produced stronger relationships with NPC than those that were obtained when the O<sub>2</sub>-A band or the NIR peak was used individually (scenarios 3–6 and 8–12), but in contrast combinations produced small improvements over using only the O<sub>2</sub>-B band or the red peak. Hence, the O<sub>2</sub>-A and the NIR peak benefited the most from band combinations. The combination of O<sub>2</sub>-B and O<sub>2</sub>-A bands produced similar results as when combining the two peaks. Hence, these two combinations could be considered essentially equivalent from this analysis. In all cases, the peak ratio ( $F_{685}/F_{740}$ ) produced considerably poorer correlations than using the two bands individually. Combining the mid-valley with the two peaks produced only marginal improvements over the combined peaks. Also small improvements were obtained when combining SIF retrievals at the four absorption lines (H<sub>α</sub>, O<sub>2</sub>-B, water vapor, O<sub>2</sub>-A). The  $F_{\text{Total}}$  (integrated SIF) generally did not yield a predictive advantage and in several instances produced weaker correlations than other features. Conversely, further improvements were achieved for most of the scenarios when including all individual wavelengths in the regression analysis, but gains in explaining NPC variance were modest. Overall, selection of single SIF bands such as the red peak or the O<sub>2</sub>-B band, or a combination of the two peaks or the O<sub>2</sub>-A and O<sub>2</sub>-B bands appeared sufficient to estimate NPC.

### 3.3. Nonlinear Gaussian processes regression

The nonlinear GPR (Table 7) produced stronger relationships with NPC than did linear regression for the majority of cases, although improvements were generally modest. Considering the best two-band combinations from Table 6 (i.e., the two peaks or the O<sub>2</sub>-A and O<sub>2</sub>-B), R<sup>2</sup> values were higher in scenarios 2, 4–6, and 8–12. The strongest improvements were under conditions of increasing canopy and environmental heterogeneity. Again, including all individual wavelengths in the regression led to strongest relationships for the majority of scenarios, although improvements as compared to using the SIF absorption bands were generally modest. From a pragmatic perspective, using an adaptive, nonlinear regression method and retrieving SIF in the two deepest absorption lines could be sufficient to derive NPC with sufficient accuracy.

## 4. Discussion

### 4.1. Potentials and limitations of the applied approach

Progress in imaging spectroscopy technology and data processing will soon make it possible to derive and exploit the full SIF spectrum emitted from vegetation canopies. However, to date no imaging spectroscopy broadband SIF spectra are available for the canopy scale. Hence, simulation studies are required to predict the information content of this unique source of information. This study used SCOPE simulations to conduct a theoretical examination of SIF band sensitivity to net photosynthesis of the canopy (NPC). This type of modeling study can be useful in helping to disentangle the complex relationships between canopy-leaving SIF (as estimated by an imaging spectrometer) and vegetation photosynthetic activity. Importantly, it helps to establish

**Table 7**

The sensitivity of combined bands and spectral features under different canopy configurations using GPR. Best performing R<sup>2</sup> per canopy configuration is shown in bold-faced text. The R<sup>2</sup> cells that fall within 0.01 interval of the best performing R<sup>2</sup> are highlighted in gray.

Ranging SCOPE variables		O <sub>2</sub> -B, O <sub>2</sub> -A: 687, 760 nm		H <sub>α</sub> , O <sub>2</sub> -B, water vapor, O <sub>2</sub> -A: 656, 687, 719, 760 nm		Two peaks: 685, 740 nm		Peak ratio: 685/740		Two peaks and valley: 685, 699, 740 nm		$F_{\text{Total}}$ : Integrated SIF (from 641 to 850 nm)		$F_{\text{All}}$ : All individual SIF wavelengths (from 650 to 790)	
		R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
1	V <sub>cmo</sub>	<b>1</b>	0.0040	<b>1</b>	0.0053	<b>1</b>	0.0038	0.9900	1.4283	<b>1</b>	0.0039	<b>1</b>	0.0089	<b>1</b>	0.0028
2	biochemistry	0.7863	7.0925	<b>0.7901</b>	6.9813	0.7332	7.8987	0.7181	8.2289	0.7389	8.1721	0.7268	35.8424	<b>0.7996</b>	6.7456
3	V <sub>cmo</sub> , Cab	<b>0.9996</b>	0.2598	<b>0.9998</b>	0.1948	<b>0.9996</b>	0.2548	0.6668	7.6360	<b>0.9987</b>	0.4841	<b>0.9905</b>	1.3127	<b>1</b>	0.0811
4	V <sub>cmo</sub> , leaf	0.9695	2.3618	<b>0.9883</b>	1.4783	0.9739	2.1710	0.6382	8.0869	<b>0.9830</b>	1.7833	0.6934	7.5117	<b>0.9899</b>	1.3645
5	biochemistry, leaf	0.6797	9.2628	0.6640	9.2910	0.6833	9.0908	0.5779	10.5707	0.6777	9.1064	0.5459	10.6994	<b>0.7633</b>	7.9247
6	Cab, LAI	<b>1</b>	0.0101	<b>0.9919</b>	0.4640	<b>1</b>	0.0066	0.4044	3.8472	0.9879	0.5552	0.9589	1.7122	<b>1</b>	0.0040
7	V <sub>cmo</sub> , LAI	<b>0.9967</b>	0.8520	<b>0.9988</b>	0.5203	<b>0.9993</b>	0.3837	0.8926	4.8546	0.9791	2.1539	<b>0.9945</b>	1.1048	<b>1</b>	0.0070
8	V <sub>cmo</sub> , canopy	0.9456	2.9284	0.9431	2.9300	0.9447	2.9073	0.8274	5.2367	0.9397	3.0608	0.9416	3.0289	<b>0.9945</b>	0.9428
9	V <sub>cmo</sub> , leaf, canopy	0.9164	3.7109	0.9174	3.7290	0.9177	3.7213	0.5272	8.9396	0.8982	4.1313	0.7250	6.7224	<b>0.9401</b>	3.1440
10	Biochemistry, leaf, canopy	0.3180	27.7851	0.3573	27.6869	0.3392	30.4935	0.1938	33.3582	0.3280	27.6240	0.1935	31.0465	<b>0.3819</b>	28.0068
11	Key variables driving SIF	0.5881	12.1416	<b>0.6610</b>	10.9384	0.6000	11.2757	0.3184	15.4214	0.5374	13.0736	0.4053	14.7770	0.6411	11.7772
12	All SCOPE variables	<b>0.3131</b>	37.2426	0.2676	41.4844	<b>0.3204</b>	35.4161	0.1350	43.8470	0.2870	38.1309	0.2292	39.3470	0.2873	42.9957

a theoretical foundation upon which to build hypotheses and formulate future field studies, including identification of the sorts of ancillary information required to interpret and apply SIF findings.

Although the strengths of a theoretical study are well-known, i.e., full control of all variables, virtually unlimited capability to analyze canopy configurations and their interactions with incoming radiation (e.g., Verrelst et al., 2015), relying on simulated data also has its drawbacks. The model representativeness of actual field data may be limited by (1) random errors and bias in actual retrieved SIF that are not included in the simulations, (2) model representation errors and (3) representation errors in the selected input data. These errors are briefly discussed below.

First, we have assumed a perfectly known TOC signal, but it is recognized that signals acquired by airborne or space sensors must be atmospherically corrected. Currently, such correction can be achieved using sophisticated techniques such as inversion of the atmospheric radiative transfer codes from MODTRAN (Berk et al., 1999). However, atmospheric correction algorithms generally ignore the contribution of the atmospheric spherical albedo and assume a Lambertian surface behavior (Guanter, Gonzalez-Sanpedro, & Moreno, 2007), a known temperature vertical profile and a 'guesstimate' of the aerosols nature (Chavez, 1996). Even with sophisticated algorithms, propagated errors could potentially degrade the quality of SIF retrievals. Errors can arise due to e.g. instrumental noise, atmospheric interferences, and surface anisotropy effects (e.g., Damm, Guanter, Verhoef, et al., 2015). Thus, for a full sensitivity analysis of SIF retrievals from space, in principle all these aspects should be considered. Also, spectral shifts in band central wavelengths or a non-accurate characterization of the instrument spectral response function (ISRF) can become crucial when working at very high spectral resolutions ( $<1$  nm). Hence, for results to be valid for a specific sensor, sensor-specific signal-to-noise ratios plus sensor ISRF should be considered to determine if the technical capabilities are sufficient to successfully acquire SIF measurements. For instance, Zhao et al. (2014) showed that poor signal-to-noise ratio will degrade SIF retrievals. Consequently, we could expect concomitant degradation of relationships with photosynthesis. Another point is that not all SIF absorption bands equally enable accurate retrieval of SIF. SIF is typically more easily retrieved in the  $O_2$ -A region than in the other absorption lines due to a deeper and wider absorption at 760 nm (see also Fig. 2) and the TOC reflectance is a smooth profile in the NIR shoulder (Cogliati et al., 2015). In order to quantify and mitigate these limitations when moving towards spaceborne SIF retrievals, an end-to-end mission performance simulator (FLEX-E) has been developed within the framework of FLEX where most of these limitations are explicitly taken into account. FLEX-E combines the forward simulation of complex synthetic scenes using coupled SCOPE and MODTRAN-5 (Berk et al., 2006) with modeling of the satellite and instrument behavior and the full processing scheme up to the retrieval of the final SIF products (details in Vicent et al., submitted for publication).

Second, SCOPE, like any model, has representation errors. For instance, the radiative transfer in SCOPE is described with a 1D vertical model that assumes horizontal and vertical uniformity. The model is valid for vegetation in which scattering of radiation by adjacent objects of different leaf composition or canopy structure plays no significant role. Tree crowns surrounded by open space are not represented by SCOPE, because in that case illumination is mostly from the side of the crowns, and tree crowns cast shadows on other crowns. Also vertically varying leaf properties within the crown are not considered in SCOPE. Other representation errors are possible in the models for photosynthesis, stomatal regulation and turbulent atmospheric transport that contain semi-empirical relationships with coefficients that cannot be determined from physics, but require calibration to measurements (Van der Tol et al., 2014).

Third, the modeled NPC only approximates GPP as derived from eddy covariance (EC) flux measurements over canopies; in particular, SCOPE scales up net rather than gross photosynthesis. But EC-flux GPP

computations also have their limitations: for example, flux partitioning methods conventionally estimate GPP from net ecosystem exchange of  $CO_2$  by using night-time  $CO_2$  flux measurements to estimate day-time respiration (e.g. Reichstein et al., 2005). There are significant uncertainties with that approach and with other assumptions used in EC-derived GPP as discussed by various authors (e.g. Hilker et al., 2014; Wohlfahrt & Gu, 2015), hence, those computations should be also considered as estimates of GPP. In addition, there is a certain inconsistency over the usage of the term 'GPP' across disciplines (Wohlfahrt & Gu, 2015), which contributes to confusion when trying to compare outputs. Nonetheless, SCOPE NPC simulations (although directly called GPP in some publications) have demonstrated utility in inversion schemes (Zhang, Guanter, et al., 2014) or comparative studies against flux tower GPP data (Damm, Guanter, Paul-Limoges, et al., 2015).

Finally, representation in input variables affects results. In this study full ranges from minimum to maximum values of variables were used. It may be expected that in actual field situations not all of the variables might vary independently over the full ranges. Also, some variables may be readily quantifiable from actual measured data, such as short-wave incoming radiation ( $R_{in}$ ) and LAI (Baret et al., 2013; Zhang, Liang, Zhou, Wu, & Zhao, 2014), and thus would not need to be considered as unknown. Therefore, the actual set of ranging variables may differ in each field situation.

#### 4.2. Implications regarding exploitation of retrieved SIF data

This work underlines the complexity of SIF–NPC relationships in heterogeneous vegetation environments, simulated here through the ranging of SCOPE variables to represent varying levels of biological and environmental complexity. Results led to the following three key observations.

A first key observation is that with increasing heterogeneity, i.e. more ranging variables, poorer relationships were achieved, until a point is reached where no meaningful relationships with NPC could be derived. It indicates the difficulty of interpreting canopy-leaving SIF emitted by heterogeneous vegetation. However extreme situations with all SCOPE variables fully ranging seem unlikely in reality. In contrast, when varying only SCOPE's key variables or vegetation variables, both linear and advanced nonlinear regression methods were able to produce robust relationships between SIF and NPC. But varying only some variables also seems unlikely in reality. In this respect, combining modeling studies with experimental studies may be better suited for defining the type of relationships. For instance, Damm, Guanter, Paul-Limoges, et al. (2015) reported an asymptotic relationship using instantaneous experimental  $SIF_{760}$  and eddy covariance GPP flux tower data. In this respect, our modeling findings confirmed that nonlinear regression (i.e., GPR) is better able than linear regression to deal with heterogeneous situations. Moreover, GPR possesses several more interesting properties not exploited in this study. It provides (1) an indication of band relevancy for each variable; (2) a weight for the most relevant spectra contained in the training data set; and (3) probabilistic outputs, i.e. a mean estimate and an associated uncertainty interval (Verrelst, Alonso, et al., 2012; Verrelst, Muñoz, et al., 2012). Particularly the latter may be of interest in future analyses, as it is an elegant mathematical indicator of the reliability of the NPC prediction on a per-pixel basis. Typically, an increase in pixel heterogeneity provokes an increase in prediction uncertainty. Similarly, the associated uncertainties enable quantification of the portability of regression model in space and time (Verrelst, Rivera, Moreno, & Camps-Valls, 2013).

A second key observation is that SIF–NPC relationships likely are influenced, not only by variation in photosynthetic activity ( $V_{c_{mo}}$ ), but also of other key variables such as Cab and LAI. This is not surprising since  $V_{c_{mo}}$  is a rather modest key driver of SIF variability in natural field situations. For instance, LAI has a stronger influence. The relative influence of the various drivers can also be derived through a global sensitivity analysis, as shown in Fig. 3 or in Verrelst et al.,

2015). Consequently, this suggests that empirically obtained GPP maps, as has been assembled at the global scale (e.g. Frankenberg et al., 2011; Guanter et al., 2012; Joiner et al., 2011), might actually be more of a representation of spatial variations of chlorophyll content and canopy structural variables than of photosynthetic activity. Such aspects will require consideration in future SIF retrievals and photosynthesis mapping in order to account for possible confounding factors.

A third key observation is that a clear difference in performance of  $SIF_{red}$  and  $SIF_{NIR}$  can be observed for quantification of canopy photosynthesis.  $SIF_{NIR}$  was a significantly worse predictor than  $SIF_{red}$ , suggesting that exploitation of the second emission peak is not an optimal choice for the estimation of photosynthesis. Although current practices rely on  $SIF_{NIR}$  to quantify GPP, experimental data over a maize field (Cheng et al., 2013) confirmed that  $SIF_{red}$  led to stronger relationships than  $SIF_{NIR}$ . Two reasons could help explain the superior linkage of  $SIF_{red}$  to photosynthesis at the canopy scale. First, the red emission peak has a strong sensitivity to PSII processes such as NPQ (Porcar-Castell et al., 2014). Second, this peak, due to re-absorption in the red, is coming mostly from the upper leaves of the canopy and therefore the canopy-leaving  $SIF_{red}$  is less susceptible to multiple scattering effects. In turn, although the canopy-leaving  $SIF_{NIR}$  emission is typically more pronounced due to lower incidence of re-absorption as compared to  $SIF_{red}$ , the  $SIF_{NIR}$  is highly subject to scattering, implying that observations of spatial variations of the  $SIF_{NIR}$  flux may be more subject to misinterpretation due to the influence of leaf and canopy structure (Van Wittenberghe, Alonso, Verrelst, Moreno, & Samson, 2015; Van Wittenberghe et al., 2013). In order to develop links with photosynthetic activity, exploiting the first emission peak may therefore be more successful than focusing on the currently emphasized second emission peak. The best strategy, as results here suggest, would be to retrieve SIF from both emission peaks or the  $O_2-B$  and  $O_2-A$  bands, and combining them into a regression model. Further, if SIF can be retrieved in other absorption lines such as  $H\alpha$  and water vapor, theoretically that could lead to even stronger and stable correlations. But these regions are more difficult to retrieve with high accuracies. The  $H\alpha$  absorption line is at the edge of the SIF signal (656 nm) and may be too weak to retrieve meaningful values. Also SIF retrievals at the water vapor absorption line (719 nm) may be perturbed due to spatial variability of columnar atmospheric water vapor. Considering these various aspects, it is recommended to focus predominantly on SIF retrievals in the  $O_2-B$  and  $O_2-A$  regions or in the two peaks.

Interestingly, using the absorption lines at  $O_2-B$  and  $O_2-A$ , which are located on the slopes just beyond the red and NIR peaks respectively, was as successful as the actual emission peaks. Similarly, but at the leaf scale, Van Wittenberghe et al. (2014) analyzed the sensitivity of the SIF spectral bands using GPR by correlating it to chlorophyll content. GPR ranking of most relevant bands revealed that the most sensitive SIF bands were found on the slopes. By subsequently plotting the first derivative of the SIF profiles of the leaf dataset it was illustrated that largest range of variation is effectively to be found on the slopes.

While the use of a combination of SIF retrieval bands improved relationships, using the ratio of red to NIR peaks in the regression analysis caused a degradation of accuracies in comparison to using the peak values independently. We suggest the reason might be that the (nonlinear) regression algorithm can exploit spectral data more efficiently when provided with multiple data layers than when that data is transformed with a simple function to one data layer. The same phenomenon was also reported when comparing the performance of vegetation indices as opposed to individual reflectance data into GPR to estimate biophysical variables (Verrelst, Alonso, et al., 2012).

We are not aware of imaging spectroscopy studies that exploited the information content of the full SIF broadband spectrum, but this should soon become possible, e.g. with the HyPlant airborne sensor (Rascher et al., 2015). Based on findings presented here, retrieving and utilizing multiple SIF observations is strongly encouraged. Upon reconstruction

of the full broadband SIF signal, all individual bands could be directly inserted into a full-spectrum (nonlinear) regression analysis, as demonstrated successfully here.

#### 4.3. Towards unbiased global estimation of photosynthesis

Given the biochemical, leaf, structural and micrometeorological influences on the canopy-leaving SIF signal, it will be necessary to disentangle the information content related to photosynthetic activity from that due to vegetation structure in order to achieve unbiased estimations of photosynthetic carbon uptake. Two complementary strategies to address these issues may be envisaged: (1) exploitation of the full SIF emission, i.e., including both  $SIF_{red}$  and  $SIF_{NIR}$  emission features; and (2) use of jointly derived biophysical variables in order to account for structural effects. Firstly,  $SIF_{red}$  and  $SIF_{NIR}$  have different behavioral features:  $SIF_{red}$  is less scattered and more sensitive to PSII photochemistry, whereas  $SIF_{NIR}$  is less re-absorbed. Detection of these two signals in combination with reflectance based estimates of pigments could reveal phenological and physiological changes in canopies. Secondly, through simultaneous retrieval of key biophysical variables (e.g. LAI, Cab), such a priori information could be used to constrain the regression models or be incorporated into assimilation schemes, potentially leading to stronger SIF-photosynthesis relationships. Similarly, exploiting covariance linkages among key retrievable vegetation attributes (e.g. Cab, LAI) and biochemistry ( $V_{c_{\max}}$ ) in space and time could lead to improved understanding of plant physiological responses to environmental drivers and stresses.

Atmospheric platforms recently or currently operational in space (e.g. GOSAT, SCIAMACHY, GOME-2, OCO-2) or anticipated in the near future (e.g. GOSAT-2, Sentinels 4, 5 & 5P) are not optimized to realize multiple SIF retrievals along with estimations of biophysical variables. Such platforms either are too restricted spectrally to permit derivation of both  $SIF_{red}$  and  $SIF_{NIR}$  or their spatial resolution is too coarse to allow characterization of heterogeneity at canopy or stand scales. For example, the TROPOMI sensor, which will be on-board ESA's Sentinel-5 Precursor satellite, starts sampling at 675 nm, but there are outstanding issues to be resolved for reliable discrimination of the red peak, and the spatial footprint at  $7\text{ km} \times 7\text{ km}$  is well beyond the canopy scale (Guanter et al., 2014). Similarly, GOME-2, used in recent retrievals of red and NIR fluorescence, samples at an even coarser spatial resolution of  $40\text{ km} \times 40\text{ km}$  or  $40\text{ km} \times 80\text{ km}$  (Joiner et al., 2013; Wolanin et al., 2015). At the present time, the strongest prospect for acquisitions is the Fluorescence Explorer (FLEX) mission, now approved as ESA's Earth Explorer 8. FLEX, a small satellite flying in tandem with ESA's Sentinel-3, would be unique in the sense that it would be equipped with an imaging spectrometer (FLORIS) specifically designed to capture the full broadband SIF signal, including both emission peaks. The full spectral range would be 500 to 780 nm at high spectral resolution (up to 0.3 nm) in the regions of the fluorescence peaks and the PRI region. With a continuous spatial resolution of 300 m, FLEX would provide both high resolution and bi-weekly global coverage (Kraft et al., 2013). The visible part of the spectrum (500–677 nm) can allow for the concurrent estimation of chlorophyll absorption and xanthophyll-related NPQ, using information from, for example, the PRI (also noting precautions discussed by Porcar-Castell et al., 2012; Wong & Gamon, 2015a, 2015b) while the complete FLORIS spectrum allows for the estimation of LAI. These sources of information will facilitate interpretations of SIF data. Relevant meteorological variables would be available from Sentinel-3. Apart from supporting the interpretation of SIF, these meteorological variables would also serve for signal correction (Kraft et al., 2013). Such a platform should improve the capacity to monitor terrestrial vegetation vitality and to link SIF with photosynthetic carbon uptake.

## 5. Conclusions

A SCOPE modeling simulation study was conducted to examine how successfully top-of-canopy sun-induced fluorescence (SIF) can be



related to net photosynthesis of the canopy (NPC). To analyze the influence of biochemistry, leaf and canopy mechanisms impacting the SIF signal, multiple canopy configurations were simulated. Regression analyses between simulated SIF retrievals and NPC outputs indicated that robust relationships can be achieved. However, it should be emphasized that in heterogeneous conditions, the main biochemical variable regulating photosynthetic capacity ( $V_{\text{cmo}}$ ) is only one of the key variables affecting SIF–NPC relationships. Other key vegetation variables that drive SIF–NPC relationships include chlorophyll content (Cab) and canopy structural variables (e.g., LAI). Regarding the predictive strengths of SIF features the following conclusions may be drawn:

- (1) The most sensitive SIF bands to NPC were located around the first emission peak for heterogeneous canopy configurations.
- (2) Combining two SIF retrieval bands (e.g.,  $O_2\text{-B}$  and  $O_2\text{-A}$ ) led to stronger correlations than using only one SIF band, especially if that single band was the  $O_2\text{-A}$ .
- (3) Using a combination of  $O_2\text{-B}$  and  $O_2\text{-A}$  bands produced similar performances as using the two emission peaks. Using the peak ratio produced poorer relationships than when both bands were individually entered into the regression model.
- (4) Even stronger correlations were achieved using four main SIF retrieval bands ( $H\alpha$ ,  $O_2\text{-B}$ , water vapor,  $O_2\text{-A}$ ).
- (5) Superior correlations were obtained when all the individual bands of the full SIF emission spectrum (650–790 nm) were included in the analysis.
- (6) Using an adaptive, nonlinear regression algorithm (GPR) further improved the correlations for the great majority of cases. This suggests that SIF–NPC correlations are typically nonlinear.

Overall, it is recommended to sample the SIF profile in at least both  $O_2\text{-B}$  and  $O_2\text{-A}$  regions to enable accurate quantification of vegetation photosynthetic activity from SIF retrievals.

When advancing from a purely statistical approach towards a more process-based approach, it is also recommended to retrieve in parallel with SIF retrievals key biophysical and meteorological variables, in order to disentangle their influence on the SIF signal at the pixel scale. Such an assimilation approach would provide unique opportunities to derive unbiased global estimates of photosynthetic carbon uptake.

## Acknowledgment

Funding for this work was provided by the European Space Agency (ESA), ‘2012 FLEX/Sentinel-3 Tandem Mission Photosynthesis Study’ (ESA/ESTEC Contract No. 4000106396/12/NL/AF). Three anonymous reviewers are thanked for providing comments that helped to improve the quality of the original manuscript.

## References

Alonso, L., Gómez-Chova, L., Vila-Francés, J., Amorós-López, J., Guanter, L., Calpe, J., & Moreno, J. (2007). Sensitivity analysis of the Fraunhofer Line Discrimination method for the measurement of chlorophyll fluorescence using a field spectroradiometer. *Geoscience and Remote Sensing Symposium* 3756–3759. IGARSS 2007.

Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annual Review of Plant Biology*, 59, 89–113.

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P., & Smets, B. (2013). GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part I: principles of development and production. *Remote Sensing of Environment*, 137, 299–309.

Berk, A., Anderson, G. P., Acharya, P. K., Bernstein, L. S., Muratov, L., Lee, J., ... Shettle, E. P. (2006). MODTRAN™ 5: 2006 update. *Proceedings of SPIE - The International Society for Optical Engineering*, 6233 (II).

Berk, A., Anderson, G. P., Bernstein, L. S., Acharya, P. K., Dothe, H., Matthew, M. W., ... Hoke, M. L. (1999). MODTRAN4 radiative transfer modeling for atmospheric correction. *Proceedings of SPIE - The International Society for Optical Engineering*, 3756, 348–353.

Chavez, P. S., Jr. (1996). Image-based atmospheric corrections—revisited and improved. *Photogrammetric Engineering and Remote Sensing*, 62, 1025–1036.

Cheng, Y., Middleton, E., Zhang, Q., Huemmrich, K., Campbell, P., Corp, L., ... Daughtry, C. (2013). Integrating solar induced fluorescence and the photochemical reflectance index for estimating gross primary production in a cornfield. *Remote Sensing*, 5(12), 6857–6879.

Cogliati, S., Verhoef, W., Kraft, S., Sabater, N., Alonso, L., Vicent, J., ... Colombo, R. (2015). Retrieval of sun-induced fluorescence by advanced spectral fitting methods. *Remote Sensing of Environment*, 169, 344–357.

Collatz, G. J., Ball, J. T., Grivet, C., & Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, 54, 107–136.

Collatz, G. J., Ribas-Carbo, M., & Berry, J. A. (1992). Coupled photosynthesis–stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology*, 19, 519–538.

Damm, A., Guanter, L., Paul-Limoges, E., Van der Tol, C., Hueni, A., Buchmann, N., ... Schaepman, M. E. (2015a). Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches. *Remote Sensing of Environment*, 166, 91–105.

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

Damm, A., Guanter, L., Laurent, V. C. E., Schaepman, M. E., Schickling, U., & Rascher, U. (2014). FLD-based retrieval of sun-induced chlorophyll fluorescence from medium spectral resolution airborne spectroscopy data. *Remote Sensing of Environment*, 147, 256–266.

Daumard, F., Goulas, Y., Champagne, S., Fournier, A., Ounis, A., Olioso, A., & Moya, I. (2012). Continuous monitoring of canopy level sun-induced chlorophyll fluorescence during the growth of a sorghum field. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 4292–4300.

Dayyoub, A. (2011). *Novel Techniques for the Remote Sensing of Photosynthetic Processes*. (PhD Thesis) Italy: University of Bologna.

Duffour, C., Olioso, A., Demarty, J., Van der Tol, C., & Lagouarde, J. P. (2015). An evaluation of SCOPE: a tool to simulate the directional anisotropy of satellite-measured surface temperatures. *Remote Sensing of Environment*, 158, 362–375.

ESA (2015). [http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/New\\_satellite\\_to\\_measure\\_plant\\_health](http://www.esa.int/Our_Activities/Observing_the_Earth/New_satellite_to_measure_plant_health)

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

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., ... Yokota, T. (2011). New global observations of the terrestrial carbon cycle from GOSAT: patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, 38. <http://dx.doi.org/10.1029/2011GL048738>.

Genty, B., Wonders, J., & Baker, N. R. (1990). Non-photochemical quenching of Fo in leaves is emission wavelength dependent: consequences for quenching analysis and its interpretation. *Photosynthesis Research*, 26, 133–139.

Guanter, L., Aben, I., Tol, P., Krijger, J. M., Hollstein, A., Köhler, P., ... Landgraf, J. (2014). Potential of the TROPospheric monitoring instrument (TROPOMI) onboard the sentinel-5 precursor for the monitoring of terrestrial chlorophyll fluorescence. *Atmospheric Measurement Techniques Discussions*, 7, 12545–12588.

Guanter, L., Alonso, L., Gomez-Chova, L., Meroni, M., Preusker, R., Fischer, J., & Moreno, J. (2010). Developments for vegetation fluorescence retrieval from spaceborne high-resolution spectrometry in the  $O(2)\text{-A}$  and  $O(2)\text{-B}$  absorption bands. *Journal of Geophysical Research-Atmospheres*, 115. <http://dx.doi.org/10.1029/2009JD013716>.

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

Guanter, L., Gonzalez-Sanpedro, M. C., & Moreno, J. (2007). A method for atmospheric correction of ENVISAT/MERIS data over land targets. *International Journal of Remote Sensing*, 28, 709–728.

Hak, R., Lichtenthaler, H. K., & Rinderle, U. (1990). Decrease of the fluorescence ratio F690/F730 during greening and development of leaves. *Radiation and Environmental Biophysics*, 29, 329–336.

Hilker, T., Hall, F. G., Coops, N. C., Black, A. T., Jassal, R., Mathys, A., & Grant, N. (2014). Potentials and limitations for estimating daytime ecosystem respiration by combining tower-based remote sensing and carbon flux measurements. *Remote Sensing of Environment*, 150, 44–52.

Jacquemoud, S., & Baret, F. (1990). PROSPECT: a model of leaf optical properties spectra. *Remote Sensing of Environment*, 34, 75–91.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., ... Frankenberg, C. (2013). Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2. *Atmospheric Measurement Techniques*, 6, 2803–2823.

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, 637–651.

Kraft, S., Del Bello, U., Drusch, M., Gabriele, A., Harnisch, B., & Moreno, J. (2013). On the demands on imaging spectrometry for the monitoring of global vegetation fluorescence from space. *SPIE Proceedings*, 8870, 88700N–887012.

Lichtenthaler, H. K., & Rinderle, U. (1988). The role of chlorophyll fluorescence in the detection of stress conditions in plants. *CRC Critical Reviews in Analytical Chemistry*, 19(Suppl. 1), S29–S85.

Liu, X., & Liu, L. (2014). Assessing band sensitivity to atmospheric radiation transfer for space-based retrieval of solar-induced chlorophyll fluorescence. *Remote Sensing*, 6(11), 10656–10675.

- Magnani, F., Olioso, A., Demarty, J., Germain, V., Verhoef, W., Moya, I., ... Van der Tol, C. (2009). Assessment of vegetation photosynthesis through observation of solar induced fluorescence from space. *ESTEC Contract No. 20678/07/NL/HE, Final Report*.
- Maxwell, K., & Johnson, G. N. (2000). Chlorophyll fluorescence: a practical guide. *Journal of Experimental Botany*, 51, 659–668.
- Meroni, M., Busetto, L., Colombo, R., Guanter, L., Moreno, J., & Verhoef, W. (2010). Performance of spectral fitting methods for vegetation fluorescence quantification. *Remote Sensing of Environment*, 114, 363–374.
- Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R., & Moreno, J. (2009). Remote sensing of solar-induced chlorophyll fluorescence: review of methods and applications. *Remote Sensing of Environment*, 113, 2037–2051.
- Mohammed, G. H., Goulas, Y., Magnani, F., Moreno, J., Olejníčková, J., Rascher, U., ... Drusch, M. (2014). 2012 FLEX/Sentinel-3 Tandem Mission Photosynthesis Study. *Final report. ESA/ESTEC Contract No. 4000106396/12/NL/AF*.
- Moreno, J., Asner, G. P., Bach, H., et al. (2006). Fluorescence explorer (FLEX): an optimized payload to map vegetation photosynthesis from space. *AIAA 57th International Astronautical Congress, Vol. 3*. (pp. 2065–2074). Valencia: IAC.
- Moya, I., Daumard, F., Moise, N., Ounis, A., & Goulas, Y. (2006). First airborne multiwavelength passive chlorophyll fluorescence measurements over La Mancha (Spain) fields. *2nd International Symposium on Recent Advances in Quantitative Remote Sensing* (pp. 820–825).
- Papageorgiou, G. C., & Govindjee (2004). *Chlorophyll a fluorescence - a signature of photosynthesis*. Springer, 818.
- Pedrés, R., Goulas, Y., Jacquemoud, S., Louis, J., & Moya, I. (2010). FluorMODleaf: A new leaf fluorescence emission model based on the PROSPECT model. *Remote Sensing of Environment*, 114, 155–167.
- Perez-Priego, O., Zarco-Tejada, P. J., Miller, J. R., Sepulcre-Canto, G., & Fereres, E. (2005). Detection of water stress in orchard trees with a high-resolution spectrometer through chlorophyll fluorescence in-filling of the O-2-A band. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 2860–2869.
- Porcar-Castell, A. (2011). A high-resolution portrait of the annual dynamics of photochemical and non-photochemical quenching in needles of *Pinus sylvestris*. *Physiologia Plantarum*, 143(2), 139–153.
- Porcar-Castell, A., Garcia-Plazaola, J. I., Nichol, C. J., Kolari, P., Olascoaga, B., Kuusinen, N., ... Nikinmaa, E. (2012). Physiology of the seasonal relationship between the photochemical reflectance index and photosynthetic light use efficiency. *Oecologia*, 170(2), 313–323.
- Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Van der Tol, C., Flexas, J., Pfündel, E. E., ... Berry, J. A. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. *Journal of Experimental Botany*, 65(15), 4065–4095.
- Rascher, U., Alonso, L., Burkart, A., Cilia, C., Cogliati, S., et al. (2015). Sun-induced fluorescence – a new probe of photosynthesis: first maps from the imaging spectrometer HyPlant. *Global Change Biology*. <http://dx.doi.org/10.1111/gcb.13017>.
- Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.
- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., et al. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, 11(9), 1424–1439.
- Rossini, M., Nedbal, L., Guanter, L., Ać, A., Alonso, L., Burkart, A., ... Rascher, U. (2015). Red and far red Sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophysical Research Letters*, 42, 1632–1639.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181, 259–270.
- Saltelli, A., Tarantola, S., & Chan, K. P. S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41, 39–56.
- Subhash, N., & Mohanan, C. N. (1997). Curve-fit analysis of chlorophyll fluorescence spectra: application to nutrient stress detection in sunflower. *Remote Sensing of Environment*, 60, 347–356.
- Timmermans, J., Su, Z., Van der Tol, C., Verhoef, A., & Verhoef, W. (2013). Quantifying the uncertainty in estimates of surface-atmosphere fluxes through joint evaluation of the SEBS and SCOPE models. *Hydrology and Earth System Sciences*, 17, 1561–1573.
- Van der Tol, C., Berry, J. A., Campbell, P. K. E., & Rascher, U. (2014). Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence. *Journal of Geophysical Research – Biogeosciences*, 119, 2312–2327.
- 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.
- Van Wittenberghe, S., Alonso, L., Verrelst, J., Hermans, I., Delegido, J., Veroustraete, F., ... Samson, R. (2013). Upward and downward solar-induced chlorophyll fluorescence yield indices of four tree species as indicators of traffic pollution in Valencia. *Environmental Pollution*, 173, 29–37.
- Van Wittenberghe, S., Alonso, L., Verrelst, J., Moreno, J., & Samson, R. (2015). Bidirectional sun-induced chlorophyll fluorescence emission is influenced by leaf structure and light scattering properties – a bottom-up approach. *Remote Sensing of Environment*, 158, 169–179.
- Van Wittenberghe, S., Verrelst, J., Rivera, J. P., Alonso, L., Moreno, J., & Samson, R. (2014). Gaussian processes retrieval of leaf parameters from a multi-species reflectance, absorbance and fluorescence dataset. *Journal of Photochemistry and Photobiology B: Biology*, 134, 37–48.
- Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sensing of Environment*, 16, 125–141.
- Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., & Moreno, J. (2012a). Retrieval of vegetation biophysical parameters using Gaussian process techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5 PART 2), 1832–1843 (6071004).
- Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J. P., Camps-Valls, G., & Moreno, J. (2012b). Machine learning regression algorithms for biophysical parameter retrieval: opportunities for sentinel-2 and – 3. *Remote Sensing of Environment*, 118, 127–139.
- Verrelst, J., Rivera, J. P., Moreno, J., & Camps-Valls, G. (2013). Gaussian processes uncertainty estimates in experimental sentinel-2 LAI and leaf chlorophyll content retrieval. *ISPRS Journal of Photogrammetry and Remote Sensing*, 86, 157–167.
- Verrelst, J., Rivera, J. P., Van der Tol, C., Magnani, F., Mohammed, G., & Moreno, J. (2015). Global sensitivity analysis of the SCOPE model: what drives simulated canopy-leaving sun-induced fluorescence? *Remote Sensing of Environment*, 166, 8–21.
- Vicent, J., Sabater, N., Tenjo, C., Acarreta, J. R., Manzano, M., Rivera, J. P., ... Moreno, J. (2015). FLEX end-to-end mission performance simulator. *IEEE Transactions on Geoscience and Remote Sensing* (submitted for publication).
- Von Caemmerer, S. (2000). Biochemical models of Leaf Photosynthesis. *Techniques in plant science no 2*. Collingwood: CSIRO Publishing.
- Von Caemmerer, S. (2013). Steady-state models of photosynthesis. *Plant, Cell & Environment*, 36(9), 1617–1630.
- Weis, E., & Berry, J. A. (1987). Quantum efficiency of photosystem II in relation to energy dependent quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta*, 894, 198–208.
- Wohlfahrt, G., & Gu, L. (2015). Opinion: the many meanings of gross photosynthesis and their implication for photosynthesis research from leaf to globe. *Plant, Cell & Environment*. <http://dx.doi.org/10.1111/pce.12569>.
- Wolanin, A., Rozanov, V. V., Dinter, T., Noël, S., Vountas, M., Burrows, J. P., & Bracher, A. (2015). Global retrieval of marine and terrestrial chlorophyll fluorescence at its red peak using hyperspectral top of atmosphere radiance measurements: feasibility study and first results. *Remote Sensing of Environment*, 166, 243–261.
- Wong, C. Y. S., & Gamon, J. A. (2015a). Three causes of variation in the photochemical reflectance index (PRI) in evergreen conifers. *New Phytologist*, 206, 187–195.
- Wong, C. Y. S., & Gamon, J. A. (2015b). The photochemical reflectance index provides an optical indicator of spring photosynthetic activation in evergreen conifers. *New Phytologist*, 206, 196–208.
- Zarco-Tejada, P. J., Gonzalez-Dugo, V., & Berni, J. A. J. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment*, 117, 322–337.
- Zarco-Tejada, P. J., Morales, A., Testi, L., & Villalobos, F. J. (2013). Spatio-temporal patterns of chlorophyll fluorescence and physiological and structural indices acquired from hyperspectral imagery as compared with carbon fluxes measured with eddy covariance. *Remote Sensing of Environment*, 133, 102–115.
- Zhang, X., Liang, S., Zhou, G., Wu, H., & Zhao, X. (2014a). Generating global LAnd surface satellite incident shortwave radiation and photosynthetically active radiation products from multiple satellite data. *Remote Sensing of Environment*, 152, 318–332.
- Zhang, Y., Guanter, L., Berry, J. A., Joiner, J., Van der Tol, C., Huete, A., ... Köhler, P. (2014b). Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. *Global Change Biology*, 20, 3727–3742.
- Zhao, F., Guo, Y., Verhoef, W., Gu, X., Liu, L., & Yang, G. (2014). A method to reconstruct the solar-induced canopy fluorescence spectrum from hyperspectral measurements. *Remote Sensing*, 6(10), 10171–10192.