Naethe, P.; Julitta, T.; Yao-Yung Chang, C.; Burkart, A.; Migliavacca, M.; Guanter-Palomar, LM.; Rascher, U. (2022). A precise method unaffected by atmospheric reabsorption for ground-based retrieval of red and far-red sun-induced chlorophyll fluorescence. Agricultural and Forest Meteorology. 325:1-19. https://doi.org/10.1016/j.agrformet.2022.109152

https://doi.org/10.1016/j.agrformet.2022.109152

A precise method unaffected by atmospheric reabsorption for ground-based retrieval of red and far-red sun-induced chlorophyll fluorescence

Paul Naethe [a,*], Tommaso Julitta [a], Christine Yao-Yun Chang [b], Andreas Burkart [a], Mirco Migliavacca [c, f], Luis Guanter [d], Uwe Rascher [e]

[a] *JB Hyperspectral Devices GmbH, Am Botanischen Garten 33, Düsseldorf 40225, Germany*

[b] *Adaptive Cropping Systems Laboratory, USDA, Agricultural Research Service, Beltsville Agricultural Research Center, 10300 Baltimore Ave., BLDG 001 BARC-WEST, Beltsville, MD 20705-2350, USA*

[c] *Max Planck Institute for Biogeochemistry, Hans Knöll Straße 10, Jena 07745, Germany*

[d] *Research Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, Camino de Vera, s/n., Valencia 46022, Spain*

[e] *IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences, Leo-Brandt-Str., Jülich 52425, Germany*

[f] *European Commission, Joint Research Centre, Via E. Fermi 2749, Ispra, VA 21027, Italy*

* Corresponding author.
E-mail address: paul@jb-hyperspectral.com (P. Naethe).

ABSTRACT

Remote sensing employs solar-induced chlorophyll fluorescence (SIF) as a proxy for photosynthesis from field to airborne and satellite sensors. The investigation of SIF offers a unique way of studying vegetation functioning from the local to the global scale. However, the passive, optical retrieval of the SIF signal is still challenging. Common retrieval approaches extract the SIF infilling directly from atmospheric oxygen bands in down-welling and up-welling radiance. They often involve a complex signal correction to compensate for atmospheric reabsorption and require long computing time. In contrast, the exploitation of solar Fraunhofer lines is devoid of atmospheric disturbances. We propose a new retrieval method for red and far-red SIF directly from up-welling radiance spectra in the spectral range between 650 nm and 810 nm by applying Partial Least Squares (PLS) regression machine learning. Solar Fraunhofer lines are exploited for SIF retrieval with the PLS approach by excluding telluric absorption features. The PLS models are trained and tested on synthetic reflectance and SIF data modeled with SCOPE. We identified a logarithmic relationship of the retrieval error with respect to signal- to-noise ratio of the instrument. The approach has been tested with real-world data measured by the Fluorescence Box (FloX) and evaluated against two well-established retrieval methods: the spectral fitting method (SFM) and the singular value decomposition (SVD). PLS models exploiting solar Fraunhofer lines retrieved meaningful SIF values with high precision and demonstrated robustness against atmospheric reabsorption, including from a 100m tall tower. In addition, PLS retrieval requires no complex correction for atmospheric reabsorption and computes 37 times faster than SFM. Hence, PLS retrieval allows fast and robust exploitation of SIF from solar Fraunhofer lines with high precision under conditions in which other retrieval approaches require complex atmospheric correction.

Keywords:
Solar induced chlorophyll fluorescence Field spectroscopy; Remote sensing SIF retrieval; Ground measurements FloX

1. Introduction

Remote sensing of solar-induced chlorophyll fluorescence (SIF) has been studied in the past decades as a non-invasive method to track photosynthesis from leaf to global scales (Mohammed et al., 2019). The signal is emitted as light in the red and near-infrared (NIR) wavelengths (Porcar-Castell et al., 2014). It is comprised of contributions from both photosystems II and I (Agati et al., 1995; Magney et al., 2019b). Thus, simultaneous monitoring of red and far-red fluorescence allows for obtaining direct insight into the light reactions and light-use efficiency of photosynthesis (Wieneke et al., 2018). At the leaf level, SIF changes as plants adjust photosynthesis and non-photochemical quenching in response to environmental conditions such as temperature, light, available water, and nutrients (Alonso et al., 2017; Camino Id et al., 2018; Campbell et al., 2019; Cendrero-Mateo et al., 2016; Martini et al., 2019, 2022; Sun et al., 2017; Zarco-Tejada et al., 2016). Additionally, canopy structure affects light absorption, scattering, and fluorescence emission, thereby also influencing the SIF signal (Dechant et al., 2020; Migliavacca et al., 2017; Van Wittenberghe et al., 2015). The true SIF signal is unknown and difficult to retrieve under natural conditions for its complex nature. The Soil Canopy Observation, Photochemistry and Energy fluxes (SCOPE) model fully integrates leaf-level radiative trans- fer and heat fluxes in the canopy and produces realistic fluorescence spectra for the validation of canopy SIF retrieval (van der Tol et al., 2009; Verhoef et al., 2018; Verrelst et al., 2016b; Yang et al., 2021). On the other hand, inversion of the SCOPE model allows the inference of parameters related to photosynthesis and energy balance from measured canopy hyperspectral reflectance (Guanter et al., 2012; Pacheco-Labrador et al., 2019b; van der Tol et al., 2016; Verrelst et al., 2015). Evaluation of SIF retrieval methods for the optical exploitation of the signal are performed with SCOPE simulated datasets, including known reflectance and SIF spectra

(Cogliati et al., 2015b).

The passive optical exploitation of SIF using hyperspectral sensors at the top-of-canopy (TOC) has great potential to enhance global monitoring of primary production across both natural ecosystems and agricultural landscapes (Damm et al., 2010a; Dechant et al., 2022; Goulas et al., 2017; Guanter et al., 2014; Magney et al., 2019a; Martini et al., 2019; Migliavacca et al., 2017; Nichol et al., 2019; Sun et al., 2018a; Tagliabue et al., 2019; Wohlfahrt et al., 2018). Towards this goal, automated SIF monitoring field spectrometer systems have been developed for deployment on stationary towers (Burkart et al., 2015; Cogliati et al., 2015a; Daumard et al., 2010; Drolet et al., 2014; Meroni et al., 2011; Rascher et al., 2009; Rossini et al., 2010), UAVs (Bendig et al., 2019; Chang et al., 2020b; Mac Arthur et al., 2014; Wang et al., 2021). Imaging sensors retrieve SIF on aircraft (Frankenberg et al., 2018; Wieneke et al., 2016). Current systems include, for example, TriFLEX (Daumard et al., 2010), Piccolo Doppio (Mac Arthur et al., 2014), PhotoSpec (Grossmann et al., 2018), FAME (Gu et al., 2019), FluoSpec2 (X. Yang et al., 2018), and Fluorescence Box (FloX) (Julitta et al., 2017). These field spectroscopy systems have been invaluable for calibration and validation of airborne and satellite missions (Porcar-Castell et al., 2015; Rossini et al., 2015) as well as investigating temporal, structural, and physiological dynamics of the monitored target (Fournier et al., 2012; Perez-Priego et al., 2015; Rossini et al., 2010). The potential of spectrometers was also further investigated in their own realm on the ground (Magney et al., 2017; X. Yang et al., 2018). Together, these ground-based measurements support the development of upcoming satellites such as the ESA Earth Explorer 8 candidate FLEX (Drusch et al., 2017; Middleton et al., 2017; Mohammed et al., 2019; Rascher et al., 2015), and enable calibration and validation of satellite SIF retrieval methods and data products (Hueni et al., 2017).

Different retrieval methods have been proposed to retrieve the SIF signal from continuous measurements of down-welling and up-welling radiance recorded by systems with different instrument configurations, as comprehensively reviewed by Meroni et al. (2009) and Cendrero-Mateo et al. (2019). The Fraunhofer Line Discrimination (FLD) method (Plascyk and Gabriel, 1975) is the basis for further, more advanced FLD retrieval algorithms, e.g. iFLD (Alonso et al., 2008). FLD estimates the emission of SIF through relative infilling of the telluric absorption bands, located around 687 nm (the Oxygen-B band, from which SIF-B is retrieved) and around 760 nm (the Oxygen-A band, from which SIF-A is retrieved). The Spectral Fitting Method (SFM) models the spectral shape of fluorescence and reflectance as mathematical functions to retrieve and correct signal intensities (Cogliati et al., 2015b; Meroni et al., 2010). SFM is a mechanistic approach based on physical principles and describes the spectral shape of the actual fluorescence signal but suffers from costly computation. Because of its reliability, we used it as the reference for field measurements in this study. The SFM is applied as the standard processing method for fluorescence retrieval with the FloX monitoring field spectrometer, optimizing first guessed SIF values from the iFLD. Both FLD and SFM approaches require high-resolution hyperspectral data around the telluric absorption features to exploit their infilling (Cendrero-Mateo et al., 2019; Cogliati et al., 2019; Julitta et al., 2016). Recently an improvement of the SFM method has been developed: The SpecFit method that allows for the retrieval of the full spectral SIF signal in ground-based field spectrometer measurements (Cogliati et al., 2019). Retrieval algorithms do not allow the exploitation of the spectral region of telluric water vapor (H2O) between 715 nm and 735 nm. Although it overlaps well with the SIF emission spectrum, rapid fluctuations impair signal retrieval (Ko¨hler et al., 2015).

However, current research highlights the necessity for correction of atmospheric influences due to varying air-column in the oxygen bands when retrieving SIF. In particular, reabsorption affects the accurate retrieval of SIF severely in the Oxygen-A absorption band around 760 nm and to a lesser extent in the Oxygen-B band around 687 nm (Sabater et al., 2018). Atmospheric reabsorption remains an issue with the computationally costly SFM and other retrieval approaches based on telluric absorption features. If the instrument is positioned in the midst of the atmospheric column more than a few meters above canopy, this effect is increasingly significant. (Aasen et al., 2019). An empirical correction approach partially addresses the issue but could not successfully demonstrate the retrieval of positive SIF-A values (Liu et al., 2019). An alternative is the use of solar Fraunhofer lines, which are in principle unaffected by changes in relative aerosol thickness of the atmospheric column and cloud cover. For example, the data driven Sin- gular Vector Decomposition (SVD) method can be configured to either exploit telluric oxygen absorption features or the solar Fraunhofer lines (Chang et al., 2020a; Guanter et al., 2013). SVD is capable of retrieving far-red SIF from satellite or ground-based field spectrometers with spectral resolution around 0.13 nm FWHM (Guanter et al., 2013, 2012; K. Yang et al., 2018). The SVD is limited to the accurate, SIF-free forward modeling of up-welling radiance and computes SIF as an inverse problem from the residuals. Furthermore, studies have shown reliable SIF retrieval solely by exploiting SIF infilling in solar Fraunhofer absorption lines between 757 nm and 771 nm from satellite data (Frankenberg and Berry, 2018; Sun et al., 2018b).

Here, we propose a novel retrieval of red SIF-B and far-red SIF-A based on the supervised machine learning algorithm Partial Least Squares (PLS) multivariate regression. The method developed aims at overcoming the shortcomings of the SVD and the SFM in terms of computing time, robustness, precisions, evading forward model inversion and without needing to compensate atmospheric reabsorption. PLS combines dimensionality reduction and statistical multivariate regression. The algorithm origins from quantitative analysis in chemo-metrics and was also applied to other fields of research (Biagioni et al., 2011; Biancolillo and

Marini, 2018; Dayal and MacGregor, 1997; Wagner et al., 2018). PLS is widely used for the retrieval of signals in mixture from a large number of variables (Jin and Wang, 2019; Schmidtlein et al., 2007; Wiklund, 2007). Further applications of the PLS algorithm include the analysis of vegetation with regards to leaf mass, canopy foliar chemicals, morphological and functional traits as well as the remote assessment of biodiversity with hyperspectral reflectance data (Cavender-Bares et al., 2016; Ma et al., 2019; Serbin et al., 2019; Serbin and Townsend, 2020; Singh et al., 2015). PLS regression decomposes the predictor data space into components (latent variables). As opposed to the SVD, latent variables are obtained with PLS in an iterative process, which optimizes the components to explain the most covariance be- tween predictor and response in the transferred orthogonal space of the singular vectors (Dayal and MacGregor, 1997). Component loading and score matrices are computed accordingly. Regression coefficients directly explain the variation in the response variable by using the transferred scores. In contrast to SVD-based SIF retrievals PLS does not compute the measured up-welling signal and residuals in a forward model. Instead, PLS is a supervised machine learning approach which predicts the response variable using a trained model to exploit the covariance structures between predicted variable (SIF) and predictor (preprocessed, up-welling radiance) in the transformed predictor space. In this study, we developed a new PLS retrieval method for the retrieval of SIF from up-welling radiance and evaluated its performance on a synthetic test dataset and on field measurements collected with the FloX instruments deployed over three field sites with differing vegetation targets. For this purpose, a dataset of synthetic up-welling radiance including SCOPE-modelled SIF was created. Two configurations of the PLS model were evaluated, the first covering the entire spectrum of the FloX sensor between 650 nm and 810 nm and the second excluding the telluric oxygen and water absorption bands. We further modelled a synthetic sensor with an adjustable Signal to Noise Ratio (SNR) to test the robustness of the approach across changing SNR with changing signal intensity. Finally, we compared the performance of PLS exploiting solar Fraunhofer lines against the two state-of-the-art retrieval methods (SFM and SVD) on both synthetic and measured data. The evaluation was focused on different aspects of the method: the performance, the speed, and the robustness against conditions where the SFM and SVD methods are not applicable.

2. Materials and methods

First, PLS models were tested for the prediction of red SIF-B around 687 nm and far-red SIF-A around 760 nm in different configurations using synthetic up-welling radiance data in which we could precisely control the spectral SIF contribution. Furthermore, the retrieval error of PLS was assessed against SCOPE simulated SIF in a synthetic test dataset in comparison to SFM and SVD. Secondly, we applied the optimized PLS models to retrieve SIF-B and SIF-A from up-welling radiance, measured at three different sites. The workflow of composing training data, tuning and evaluating the PLS models is described below.

2.1. Modelling synthetic data

We created a dataset of synthetic up-welling radiance spectra, which contained modelled fluorescence ranging between 0 and 5 mW m-2 sr-1 nm-11. To do so, we randomly selected 30 measurements of down-welling radiance from a clear day recorded by each of the three instruments. We used SCOPE to model 265 different SIF and reflectance spectra. The SCOPE dataset has been obtained through inversion against TOC reflectance collected over grassland (Pacheco-Labrador et al., 2019b). The down-welling radiance spectra were then combined with each of the SIF and reflectance spectra using a strong permutation, which randomly shuffled and 3-fold oversampled each SIF and reflectance spectra. Thus, a total synthetic dataset of 23,850 up-welling radiance spectra (L) at TOC was obtained according to Eq. (1). Here $\hat{T}$ is the total atmospheric transmittance (including both direct and diffuse radiation) of the entire path-length to the target and back to the sensor, R is the reflectance, E is the down-welling irradiance, F the fluorescence signal and $\breve{T}$ considers the reabsorption of fluorescence for the transmittance of the single path-length between target and sensor, specified for wavelength λ and view field direction Ω.

$$\mathbf{L}_{\lambda,\,\Omega} = \widehat{\mathbf{T}}_{\lambda,\Omega} \times \mathbf{R}_{\lambda,\,\Omega} \times \mathbf{E}_{\lambda,\,\Omega} + \mathbf{F}_{\lambda,\,\Omega} \times \breve{T}_{\lambda,\Omega} \qquad (1)$$

The relative fractional depth of solar Fraunhofer lines remains unaffected by atmospheric attenuation in spectral regions devoid of telluric absorption features (Frankenberg and Berry, 2018; Guanter et al., 2012). Thus, atmospheric modelling is greatly simplified if no telluric absorption bands are present. The influence of total atmospheric transmittance as well as reabsorption can be neglected if retrieving SIF from the fractional infilling of solar Fraunhofer lines in spectral regions with flat telluric absorption. In this study we retrieved SIF from both telluric oxygen absorption features (O2-A and O2-B) considering atmospheric attenuation and reabsorption, as well as the solar Fraunhofer lines devoid of telluric absorption features. We resampled the SCOPE-simulated SIF and reflectance spectra, which were initially generated at a resolution of 1 nm. Because they are relatively smooth, linear interpolation was used with respect to the FloX Instrument Spectral Response Function of the down-welling radiance measurements.

SIF-B and SIF-A values were extracted from the SCOPE modelled SIF spectra at 687 nm and 760 nm, respectively, and used as response variables for the model training, testing and retrieval benchmark. The measured down-welling radiance data already contains atmospheric absorption at sensor height together with characteristic spectral response and noise from the respective instrument. This measured noise was statistically decreased as the same down-welling radiance measurements were repeatedly used to model multiple synthetic up-welling radiances in different combinations with SCOPE-modelled SIF and true reflectance. Thus, the following noise assumptions were added to account for the actual noise levels found in real-world measurements. First, Noise equivalent delta Radiance (NedL) was characterized according to Schaepman and Dangel (2000), using the FloX spectrometer configured for 0.3 nm FWHM spectral resolution and 0.1 nm spectral sampling rate around 820 SNR. As shown in Eq. (2), NedL is computed with respect to wavelength for each pixel with the radiometric calibration gain g. The detector was characterized with 110 measurements of a stable light source and dark current in digital numbers for an integration time (IT) optimized at 80% of the detector's dynamic range, with σ2(N) and σ2(NDC) being the standard deviation of the lamp and the dark-current, respectively,

$$NedL(\lambda) = \frac{\sqrt{\sigma^2(N_\lambda) + \sigma^2(N_{DC,\,\lambda})}}{IT} \cdot g(\lambda) \qquad (2)$$

In our case, the noise was considered consistent due to an automatic optimization of the integration time (IT) in the FloX, resulting in a steady signal-level around 80% of the detector's dynamic range. Note that this assumption does not hold for devices without automatic optimization of the signal level, as the noise scales with changing signal level at the detector.

To simulate random noise as described in Eq. (3), we then applied Monte Carlo simulation using a normal distribution N centered at mean zero and standard deviation equal to NedL to propagate instrument-characteristic uncertainty for ca. 80% optimized signal exposure at detector level to the synthetic dataset.

$$Noise_{\lambda,\,\Omega} = \mathrm{N}(0,\ c\cdot NedL_\lambda)_\Omega \qquad (3)$$

The artificial noise level was controlled with scaling factor c to replicate the actual SNR in real measured data. Thus, simulated noise was computed individually and added to each instance in the synthetic dataset. Signal to noise ratio (SNR) was calculated as denoted in Eq. (4):

$$SNR = \frac{\mu L_{\lambda,\,\Omega}}{\sigma Noise_{\lambda,\,\Omega}} \qquad (4)$$

with μL being the mean signal intensity with respect to wavelength and σNoise being the standard deviation of the simulated, random noise with respect to wavelength for this detector. Note that zero noise was not simulated as the SNR is not defined or infinite in this case.

Following the approach of Caporaso et al. (2018), we computed the first spectral derivative with respect to wavelength as a final pre-treatment. The first spectral derivative expresses the change in spectral shape rather than the actual intensities in the spectral continuum. Training the PLS on the first spectral derivative ensured that the algorithm exploits relative changes in the up-welling radiance instead of absolute intensities. Thus, offsets from absolute intensities were eliminated. We then split the synthetic dataset into 60% training and 40% testing datasets by randomized selection. The PLS was trained to exploit the covariance in the first spectral derivatives with respect to SIF-B values or SIF-A values given from the SCOPE simulation, decomposing the spectral-temporal domain into orthogonal scores and loadings. This testing dataset was further used for evaluation of the retrieval error against the known SCOPE SIF, where we compared the methods of PLS, SFM and SVD, as described below. We additionally compared the performance of PLS and five other machine learning algorithms against the baseline of a linear regression model and found that PLS was the most promising with respect to retrieval error and computing time over 30 repetitions (Appendix A1).

2.2. Evaluation of PLS retrieval performance

Two configurations of the synthetic data simulating the physical instrument were investigated in the spectral range of the detector be- tween 650 nm and 810 nm for model training, cross-validation and testing of PLS:

1) Including the full, usable spectral range of the FloX between 650 nm and 810 nm (hereafter referred to as PLSfull),
2) Excluding the spectral region of telluric absorption bands for O2- B between 684 nm and 700 nm, for H2O between 715 nm and 735 nm and for O2-A between 759 nm and 770 nm (hereafter referred to as PLS).

Each of the two response variables, SIF-B and SIF-A, were extracted from the SCOPE-simulated full SIF spectrum at 687 nm and

760 nm, respectively, and used in separate regression models. The models were trained using the “pls” package (Mevik and Wehrens, 2007) in R (Core Team, 2017) facilitating the orthogonal scores algorithm. Hereafter we refer to these as the SIF-B model and SIF-A model. The SIF-B model can predict only red-SIF values (SIF at 687 nm) and the SIF-A model can only predict far-red SIF values (SIF at 760nm) from the first spectral derivative of the up-welling radiance spectrum.

Random k-fold stratified cross-validation with four segments was applied during the training process for 360 latent vectors ranked ac- cording to the explained signal covariance. In a second cross-validation step, the models were evaluated in their ability to predict the response variable of the unknown synthetic testing dataset. The Root Mean Squared Error of Prediction (RMSEP) was calculated with respect to the actual known SCOPE-modelled value according to Eq. (5):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{n}} \qquad (5)$$

where ̂x is the actual known fluorescence value and x is the predicted value from the PLS regression model for n instances in the testing dataset. The process was repeated 100 times each for different configurations of the testing and training data to determine mean and standard deviation goodness-of-fit metrics. The average RMSEP and standard deviation were then calculated with respect to the number of components of prediction in the testing data. To investigate the algorithm’s limitations on scaling noise, we assumed an ideal, synthetic detector with controllable noise. Using a single down-welling measurement, a training and testing dataset was modelled which was practically free of random noise but still included the specific response characteristics of the sensor. Increasing amounts of Monte-Carlo simulated noise were added by changing the scaling factor c in Eq. (3) stepwise in 100 iterations to simulate detector SNR between 10,000 and 10 to account for scaling noise with changing signal intensity (see Eq. (4)). RMSEP was reported with respect to simulated SNR. The covariance scale was calculated for both of the abovementioned spectral configurations for the standardized covariance of the SIF signal with the first spectral derivative of synthetic up-welling radiance continuum, i.e., cov (Fλ, Ω, Lλ, ΩdL dλ). Each covariance value was scaled by the standard deviation of the total covariance in the respective waveband to normalize the covariance in the spectral continuum. Inversely, the covariance scale multiplied with the standardized latent variables results in the regular, temporal covariance of the computed components (scores) in the deflated spectral continuum.

2.3. Instruments for field measurements

The FloX monitoring field spectrometer(JB-Hyperspectral Devices, Düsseldorf, Germany) is a fully autonomous dual field of view system, which records long term time-series of down-welling and up-welling radiance computed towards hyperspectral reflectance and SIF (Acebron et al., 2021; Burkart et al., 2015; Dechant et al., 2022; Julitta et al., 2016; Kra¨mer et al., 2021).

We measured hyperspectral reflectance using three different FloX systems installed in France, Germany and Italy (Table 1). The instruments at both the Italian and German sites were installed at close distance to canopy. In contrast, the instrument in France was mounted at a 100 m distance from the canopy, to evaluate atmospheric reabsorption of the SIF signal with increasing air-column between sensor and canopy. Fig. 1 shows the setup conditions of the instruments in the field. Data processing and further use of the data is described in the next sections.

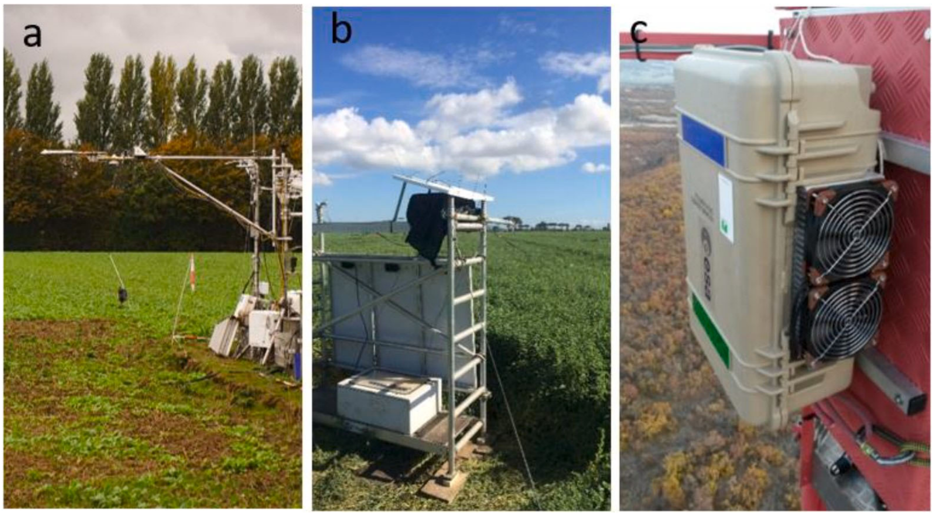
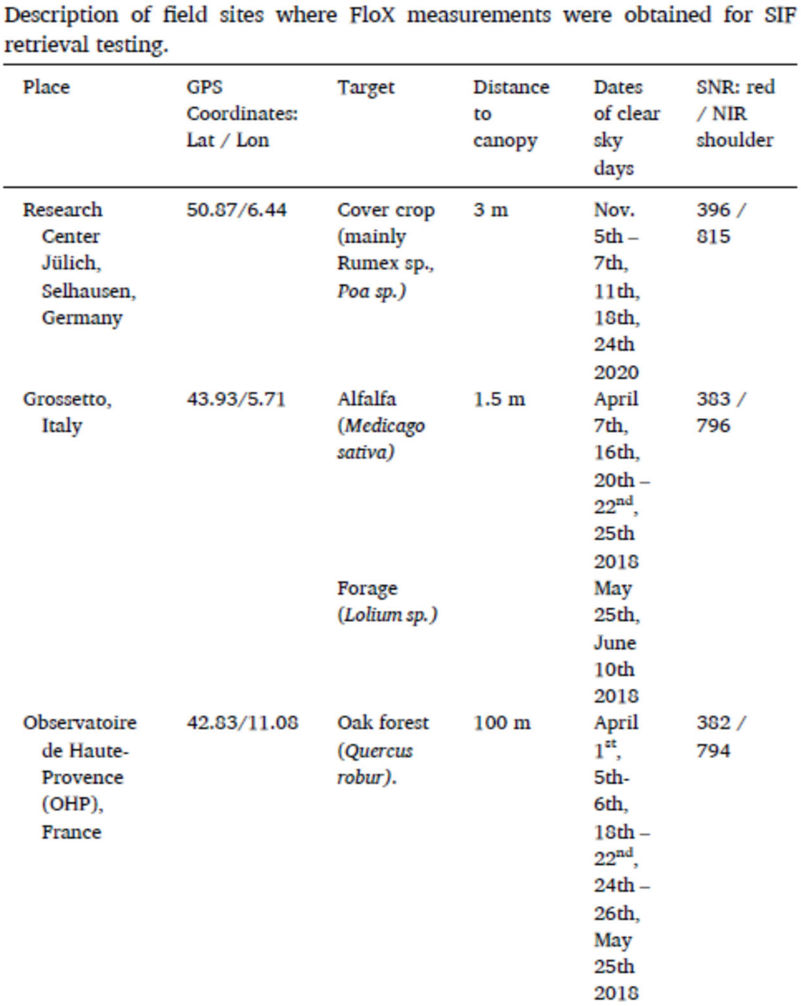


Fig. 1. From left to right, setup of the FloX systems (a) at TOC over Rumex in Germany, (b) at TOC over Alfalfa in Italy and (c) at 100 m distance to canopy over oak forest in France.

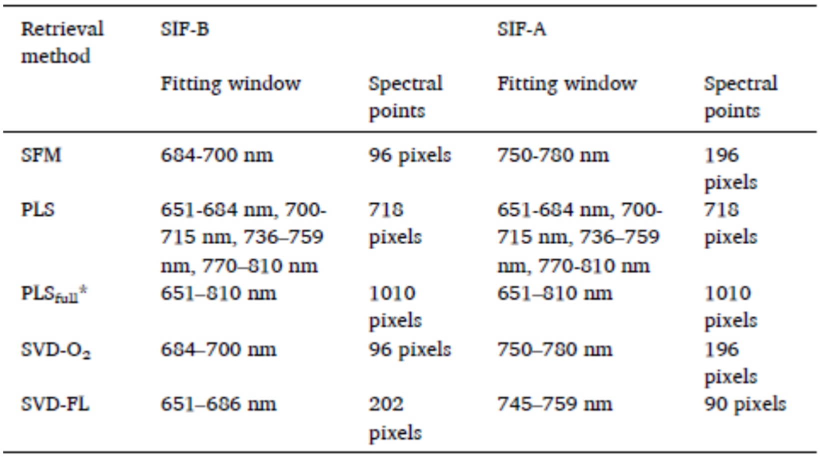
Table 1



Description of field sites where FloX measurements were obtained for SIF retrieval testing.

Place	GPS Coordinates: Lat / Lon	Target	Distance to canopy	Dates of clear sky days	SNR: red / NIR shoulder
Research Center Jülich, Selhausen, Germany	50.87/6.44	Cover crop (mainly Rumex sp., *Poa sp.)*	3 m	Nov. 5th – 7th, 11th, 18th, 24th 2020	396 / 815
Grossetto, Italy	43.93/5.71	Alfalfa (*Medicago sativa)*	1.5 m	April 7th, 16th, 20th – 22nd, 25th 2018	383 / 796
		Forage (*Lolium sp.)*		May 25th, June 10th 2018	
Observatoire de Haute-Provence (OHP), France	42.83/11.08	Oak forest (*Quercus robur)*.	100 m	April 1st, 5th-6th, 18th – 22nd, 24th – 26th, May 25th 2018	382 / 794

Table 2. Fitting windows and spectral points acquired by the FloX field spectrometer used for the retrieval of SIF-B and SIF-A with SFM, PLS, SVD-O2 and SVD-FL. * the full spectral configuration of PLS was tested in synthetic to estimate the loss of information when excluding the telluric absorption features under ideal conditions.



Retrieval method	SIF-B		SIF-A	
	Fitting window	Spectral points	Fitting window	Spectral points
SFM	684-700 nm	96 pixels	750-780 nm	196 pixels
PLS	651-684 nm, 700-715 nm, 736–759 nm, 770–810 nm	718 pixels	651-684 nm, 700-715 nm, 736–759 nm, 770-810 nm	718 pixels
PLS_{full}*	651–810 nm	1010 pixels	651–810 nm	1010 pixels
SVD-O$_2$	684–700 nm	96 pixels	750–780 nm	196 pixels
SVD-FL	651–686 nm	202 pixels	745–759 nm	90 pixels

2.4. Evaluation of retrieval methods using simulated and field measurements

Finally, the SFM and SVD retrievals were compared with PLS retrieval of SIF-B and SIF-A in synthetic testing data against known SCOPE modelled SIF values.

Fitting windows used for each retrieval method are shown in Table 2. We used the SFM algorithm implemented in the R packages

FieldSpectroscopyCC (https://github.com/tommasojulitta/FieldSpectroscopyCC) and FieldSpectroscopyDP (https://github.com/tommasojulitta/FieldSpectroscopyDP) as part of the standardized open source FloX processing. Here, SFM was applied as the standard processing method for fluorescence retrieval in the FloX monitoring field spectrometer, after optimizing first guessed SIF values based on iFLD retrieval. The SVD was implemented as reported in Guanter et al. (2013) and Chang et al. (2020a) to exploit the telluric oxygen absorption bands (SVD-O2) and solar Fraunhofer lines (SVD-FL). We calculated RMSEP with respect to the known SCOPE SIF values in each instance. The error was investigated with respect to changes in the SCOPE-modelled SIF, respectively. Following the approach of Cogliati et al. (2015b), SCOPE simulated SIF and reflectance spectra (Appendix A2) were used to assess the retrieval performance of the SFM, SVD-FL and SVD-O2 retrievals with a known "true SIF". We then used the three field datasets (Table 1) to retrieve SIF-B and SIF-A with SFM, SVD-FL and SVD-O2, respectively. Given that the true SIF signals in the measured data is unknown, the SFM retrieval was used as reference for later comparison.

We trained individual PLS models specifically for each instrument and response variable to retrieve red SIF-B at 687 nm and far-red SIF-A at 760 nm. The optimal number of components for prediction was selected with respect to the first minimum in average RMSEP with synthetic testing data simulating each instrument. The number of components for SIF-B and SIF-A retrieval were thus individually identified for the PLS models. SIF was then retrieved with PLS models trained with noise contribution according to in-field conditions, excluding major telluric absorption features, or with SFM, SVD-O2 or SVD-FL models, as described in Table 2. To compare the performance of PLS, SVD-O2, SVD-FL retrievals from measured observations, we calculated RMSEP as well as the coefficient of determination (R2), incident and slope from linear regression with respect to the SFM-retrieved SIF values.

3. Results

The results of the model training, testing and retrieval validation against SCOPE SIF with synthetic data are described in the following section, followed by the retrieval results from field measurements of up- welling radiance.

3.1. Model training and testing

The covariance of the PLS model across the spectrum describes the influence of spectral regions in the predictor space which covary with the response variable of the model. In this section we investigate the covariance of the two PLS Models for SIF-B and SIF-A response, respectively, and the impact of the exclusion of the telluric absorption features in the fitting window. When trained for the data configuration including telluric absorption features, the PLS model found high covariance predominantly at the O2-A band at 760 nm, with increased covariance also observed in the water vapor absorption band between 715 nm and 735 nm and in the O2-B band at 687 nm (Fig. 2b). In the model configuration excluding atmospheric absorption features, the covariance was predominantly observed around 660 nm and a large number of smaller yet pronounced peaks across the remaining spectrum (Fig. 2c). The fitting window excluding atmospheric absorption bands between 650 nm and 810 nm is an almost perfect overlay of the top of atmosphere solar irradiance spectrum (Fig. 2a) based on satellite data (Brault and Neckel, 1999; Thuillier et al., 2004). As a result, the latter model configuration is trained to exploit solar Fraunhofer lines as relevant spectral features for the fluorescence retrieval.

We simulated a synthetic FloX sensor with adjustable SNR between 10,000 and 10 to investigate the PLS algorithm limitations with respect to scalable noise. Both SIF-B and SIF-A PLS regression models are very sensitive to noise across all components (Fig. 3). The propagation of uncertainties increases RMSEP directly in proportion to the amount of SNR on a logarithmic scale and becomes increasingly unstable below 150 SNR (dotted line). The SIF-A model is slightly more susceptible to very low SNR at the NIR shoulder in comparison with the SIF-B model. Furthermore, the worst error is exceeding 0.5 mW m-2 sr-1 nm-1 in both models and varying largely due to the random noise contribution for very low SNR below 150. Thus, retrieval errors beyond SNR 150 cannot be reliably predetermined and is not suitable for the fluorescence retrieval. With the SNR of FloX instruments ranging around 390 in the red and 800 in the NIR shoulder wavelengths in operational scenarios due to automatic optimization of the signal level, we predicted a retrieval error of below 0.1 mW m-2 sr-1 nm-1 in both SIF-B and in SIF- A, using synthetic training and testing data.

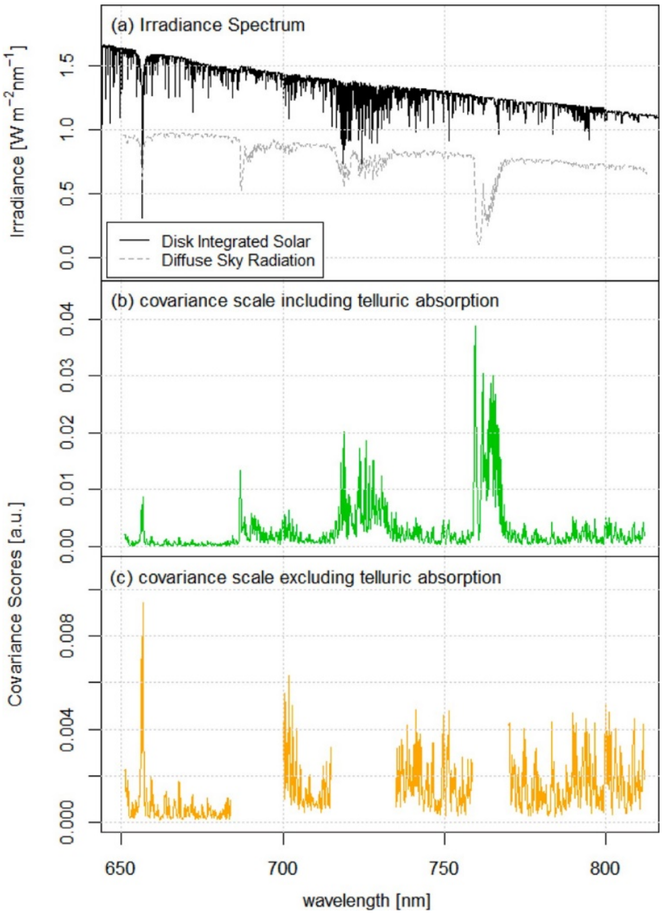


Fig. 2. Disk Integrated Solar Irradiance spectrum with Fraunhofer lines at 0.05 nm resolution after Thuillier et al. (2004) (black solid line) and measured Diffuse Sky Radiation Irradiance spectrum with telluric absorption features at0.3 nm resolution (grey dashed line) (a). Covariance between SIF and the first spectral derivative of up-welling radiance across the spectrum exploited by the PLS algorithm including telluric absorption features (b) and excluding telluric absorption features (c). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

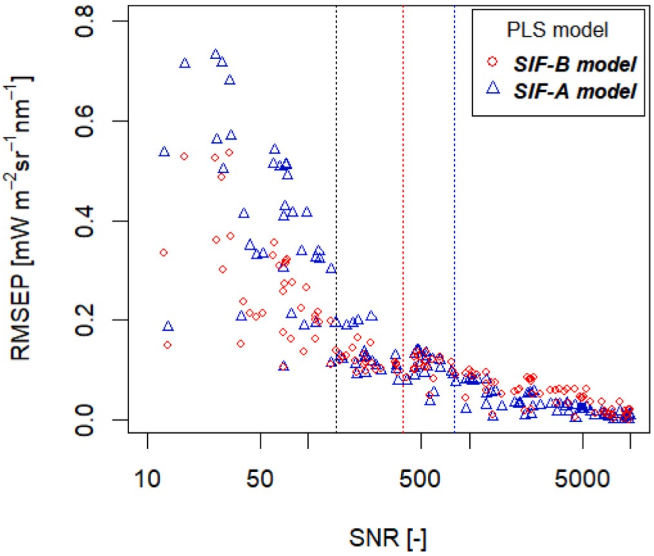


Fig. 3. Root Mean Square Error of Prediction (RMSEP) from testing data of PLS models with response variable SIF-B (red) and SIF-A (3) for simulated spectrometer SNR on logarithmic scale in 100 simulations with scaled Signal to Noise Ratio (SNR). SNR 150 is indicated by black dotted line. Operational SNR around 390 in the red and around 800 in the NIR shoulder wavelengths are marked with a red dotted line blue dotted line, respectively (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

When including telluric absorption and noise simulated according to field conditions, RMSEP significantly varied by model complexity over 100 randomized training and test datasets (Fig. 4a). Testing error significantly increased when model complexity deviated from the optimal number of components. We found an optimum of lowest RMSEP at 134 components for prediction of SIF-A with RMSEP around 0.07 mW m-2 sr-1 nm-1 and 245 components of SIF-B and RMSEP around 0.09 mW m-2 sr-1 nm-1. The standard deviation of the error also increased below and above the optimal number of components for prediction. When a low number of components was used, the error of the SIF-B model tended to be slightly lower than the error of the SIF-A model. This tendency was inverted as the number of components increased and resulted in an overall lower error of the PLS model for SIF-A prediction compared with SIF-B prediction.

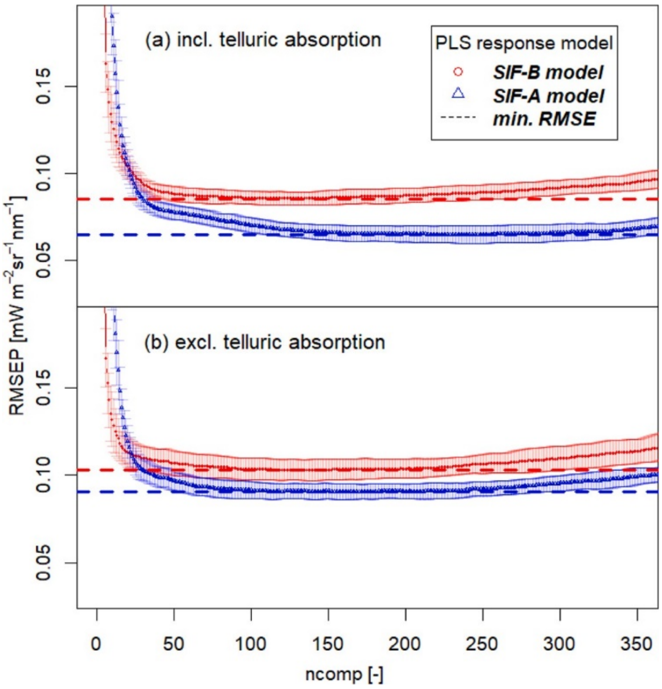


Fig. 4. Root Mean Square Error of Prediction (RMSEP) and its standard deviation across 100 PLS model training and validation cycles is shown for each of the separate PLS models in predicting the response variable SIF-B (red dots) and SIF-A (blue triangles), respectively, with increasing number of components (ncomp). Telluric absorption features are included in (a) and excluded in (b). The optimal number of components and associated lowest RMSEP is indicated with dashed lines for the SIF-B and the SIF-A model, respectively in red and blue (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

In contrast, when excluding telluric absorption and using noise simulated according to field conditions, very little increase in RMSEP was observed with respect to PLSfull (Fig. 4b). The error of SIF prediction was similar excluding telluric absorption features compared to the previous case which included those spectral regions (Fig. 4a). The mean RMSEP is minimal with 0.09 mW m-2 sr-1 nm-1 around 136 components for SIF-A model and 0.11 mW m-2 sr-1 nm-1 for the SIF-B model with 169 components, respectively. Furthermore, less complex PLS models performed with less error for SIF-B retrievals than for SIF-A retrievals. At higher complexity, the performance of the SIF-A models was again superior compared to the SIF-B model. The standard deviation increased slightly more compared to the former configuration with increasing number of components for prediction. While the PLS models for SIF retrieval showed a slightly enhanced performance when including telluric absorption features (i.e., RMSEP with respect to model complexity and standard deviation with respect to randomly permuting training and testing), the difference was not significant.

The retrieval performance of PLS, SFM, SVD-O2 and SVD-O2 was assessed with respect to known SCOPE-modelled SIF, extracted at 687 nm and 760 nm (Fig. 5). SVD-FL exhibited a strong dependence on the actual signal intensity, scaling the median RMSE almost proportional to the SCOPE-modelled SIF signal intensity. This behavior was not observed with the other retrieval algorithms, which exhibited a steady error across the investigated SCOPE SIF value range. In SIF-B exhibited PLS median RMSE very close or slightly below median RMSE of SFM (Fig. 5a). The median RMSE of SFM was lowest in SIF-A, while RMSE of SFM and SVD-O2 were very similar (Fig. 5b). PLS exhibited a slightly larger error with respect to SFM, even if the median value did not exceed 0.1 mW m-2 sr-1 nm-1.

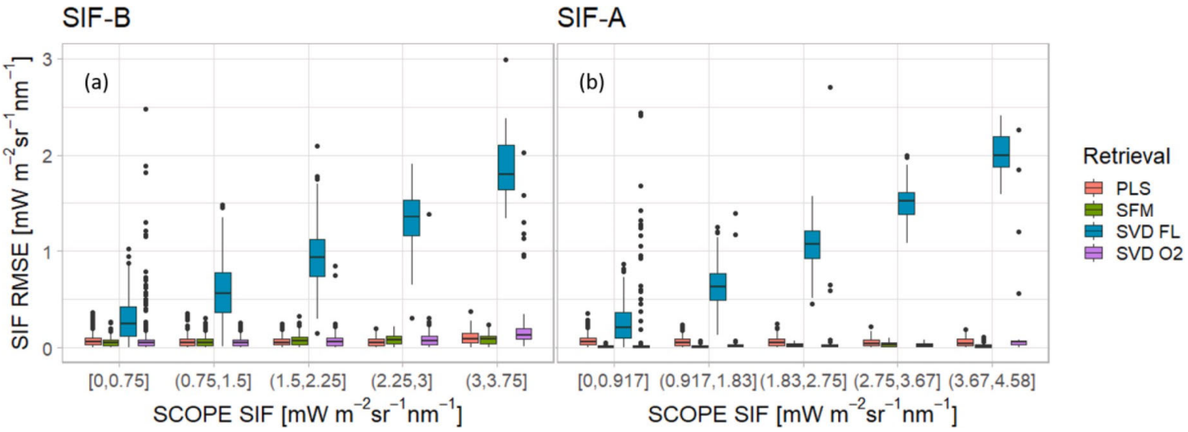


Fig. 5. RMSE of the PLS (red), SFM (green), SVD-FL (turquois) and SVD-O2 (purple) with respect to SCOPE simulated SIF for SIF-B extracted at 687 nm (a) and SIF-A extracted at 760 nm (b), with respect to changing SIF intensities in five equally large intervals given with [including boundary and (excluding boundary marks. Interquartile range is indicated by the expand of the boxes in y-direction, the median RMSE by a solid line. Whiskers expand to the last member inside 1.5 x interquartile range and outliers are marked with dots (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

3.2. Retrieving sun-induced fluorescence from field data

The previously tested PLS models excluding telluric absorption features were transferred to measured data. Typical diurnal cycles out of the entire measured data set are presented in detail for each instrument (Fig. 6). Retrieval results for the entire set are presented in Fig. 7. Finally, PLS, SVD-O2 and SVD-FL were compared with respect to SFM retrieved SIF across the entire dataset for each instrument and results are reported in Fig. 8. The differences of the PLS, SVD-O2 and SVD-FL retrieval with respect to the SFM were tested and are reported in Table 3.

3.2.1. Diurnal cycles and time series

Measured up-welling radiance data from three different instruments were processed with SFM, SVD-O2, SVD-FL and PLS. A representative diurnal cycle was selected for each site to illustrate a typical clear-sky day for a canopy in different stages of development under non- stressed conditions, and aggregated to half-hour intervals for presentation clarity. At the German site, the diurnal pattern of SIF exhibited an inverted smile shape for PLS, SVD-O2, SVD-FL and SFM (Fig. 6 a,b). This pattern is typical for non-stressed vegetation. The magnitude of the SIF- B and SIF-A signals were low (< 1 mW m-2 sr-1 nm-1) across all retrieval methods, possibly because measurements were collected towards the end of the growing season. SVD-FL overestimated SIF-B with respect to all other methods and underestimated SIF-A with respect to all other methods. The standard deviation for SVD-FL across the diurnal course was also noticeably higher, compared to all other presented retrieval methods. For SFM and SVD-O2, half-hourly means and standard deviations of retrieved SIF-B and SIF-A values were almost indistinguishable. SIF values retrieved with PLS also agreed well with both SFM and SVD-O2 in SIF-B and SIF-A.

At the Italian site, very high values were retrieved for SIF-B and SIF-A during the peak of the growing season (Cogliati et al., 2019). Again, the SVD-FL retrieval of SIF-B exhibited significantly higher values compared to all other methods (Fig. 6 c,d). Standard deviation, and therefore the uncertainties around the diurnal trend, were very high for SIF-B and SIF-A retrieved using the SVD-FL. SIF-B values retrieved using PLS were slightly above those from SVD-O2 and SFM retrieval. Differences be- tween the SFM, the SVD-O2 and PLS were almost indistinguishable in SIF-A, except for a slight overestimation of the PLS round midday after 10am UTC until 12am UTC with respect to the other methods.

Lastly, the French site exhibited very low to negative SIF from SFM and SVD-O2 retrievals, with the diurnal pattern inverted in SIF-A, compared to SVD-FL and PLS (Fig. 6 e,f). At this site, data collected by the FloX mounted at 100 m above the canopy were used without any atmospheric compensation for retrieving SIF-B and SIF-A with the SFM, SVD-O2, SVD-FL and PLS. This setup is known to cause errors in the SIF retrievals based on telluric absorption bands. In contrast, PLS and SVD- FL both exhibited positive, inverted smile patterns, typical for this kind of vegetation. Still, a difference between SVD-FL and PLS was noticeable in both SIF-B and SIF-A. SVD-FL predicted SIF-A values were lower and partially negative with a larger standard deviation around the half- hourly mean. Consequently, we retrieved only positive SIF-B and SIF- A values with lower noise using PLS. Differences between SVD-O2 and SFM were barely noticeable.

Across the entire time series, very similar values were retrieved using the three retrieval methods SVD-O2, SFM and PLS at the German site (Fig. 7 a,b). Especially SFM and SVD-O2 were always consistent. SVD-FL values for both SIF-A and SIF-B were noisier than telluric band retrievals or the PLS retrieval. Again, SVD-FL retrieval had a tendency to overestimate SIF-B. Towards the end of the growing season, magnitude of SIF was quite low, with midday peaks around 0.5 mw m-2 sr-2 nm-1 in SIF- A.

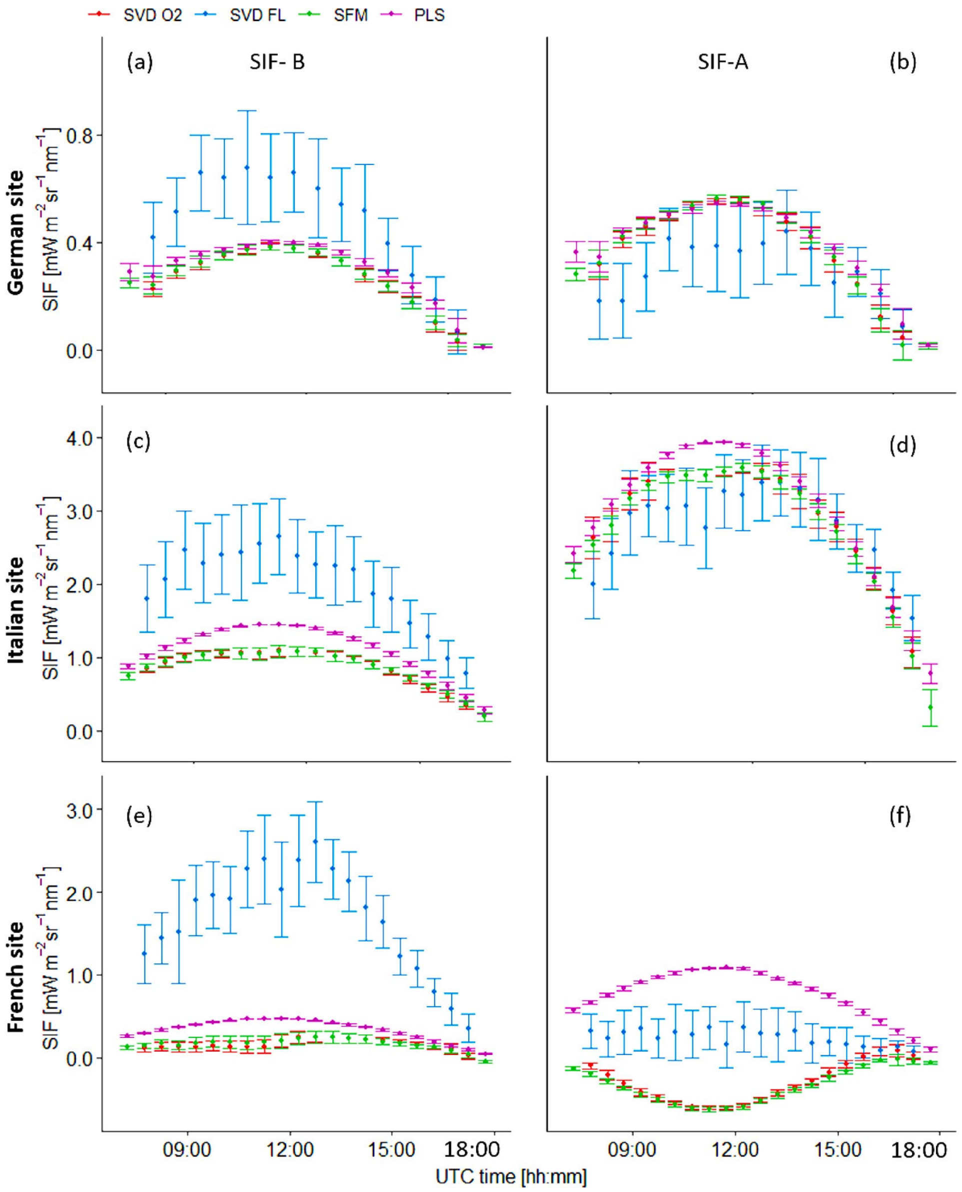


Fig. 6. Diurnal cycles of red SIF-B, resp. at 687 nm (left panel) and far-red SIF-A, resp. at 760 nm (right panel) retrieved with SVD-O2 (red), SVD-FL (blue), SFM (green) and PLS (purple). PLS models for retrieval have been trained on synthetic up-welling radiances data considering noise according to in-field conditions, excluding telluric absorption features. Data was measured at TOC in Germany over Rumex on November 6th 2020 (a, b), at TOC over Alfalfa in Italy on April 21st 2018 (c, d) and 100 m above oak forest canopy in France on April 25th 2018 (e, f). Measurements of one day were aggregated to half-hourly intervals; points indicate mean and error bars show standard deviation (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

At the Italian site, very high values were retrieved in SIF-B and SIF-A across the entire time series over both maturing alfalfa and forage canopies (Fig. 7 c,d). The error around the diurnal trend was slightly higher for SFM and SVD-O2 retrieval compared to PLS, while retrieved values were very similar in SIF-A. PLS SIF-B values were slightly above the values of SFM and SVD-O2 but remained well below the SVD-FL values. A difference in diurnal shape as well as in magnitude of SIF-A values was observed between the forage and alfalfa canopies before and after May 8th (Fig. 7 c,d). In SIF-B, the difference between the canopies was less pronounced.

At the French site, both, SIF-B and SIF-A showed mostly negative values across the entire time series when retrieved with SFM and SVD-O2 due to atmospheric distortion (Fig. 7 e,f). Notably, although PLS did not correct for atmospheric impacts, retrieval results were always positive. SVD-FL retrieved partially negative values with the seasonal pattern exhibiting a random jump in the second half of April. Furthermore, SVD- FL values in SIF-B were significantly higher in comparison with all other retrieval methods. The PLS results showed a gradual increase of SIF-B and SIF-A between April 18th and April 26th, which is in temporal agreement with the greening-up and increase of photosynthetic activity in the target canopy. This gradual increase was also present in the SVD- FL SIF-A values, while the diurnal pattern was barely distinguishable due to the high noise in the signal. Using the PLS retrieval, diurnal pattern in both SIF-B and SIF-A were exhibited with high precision.

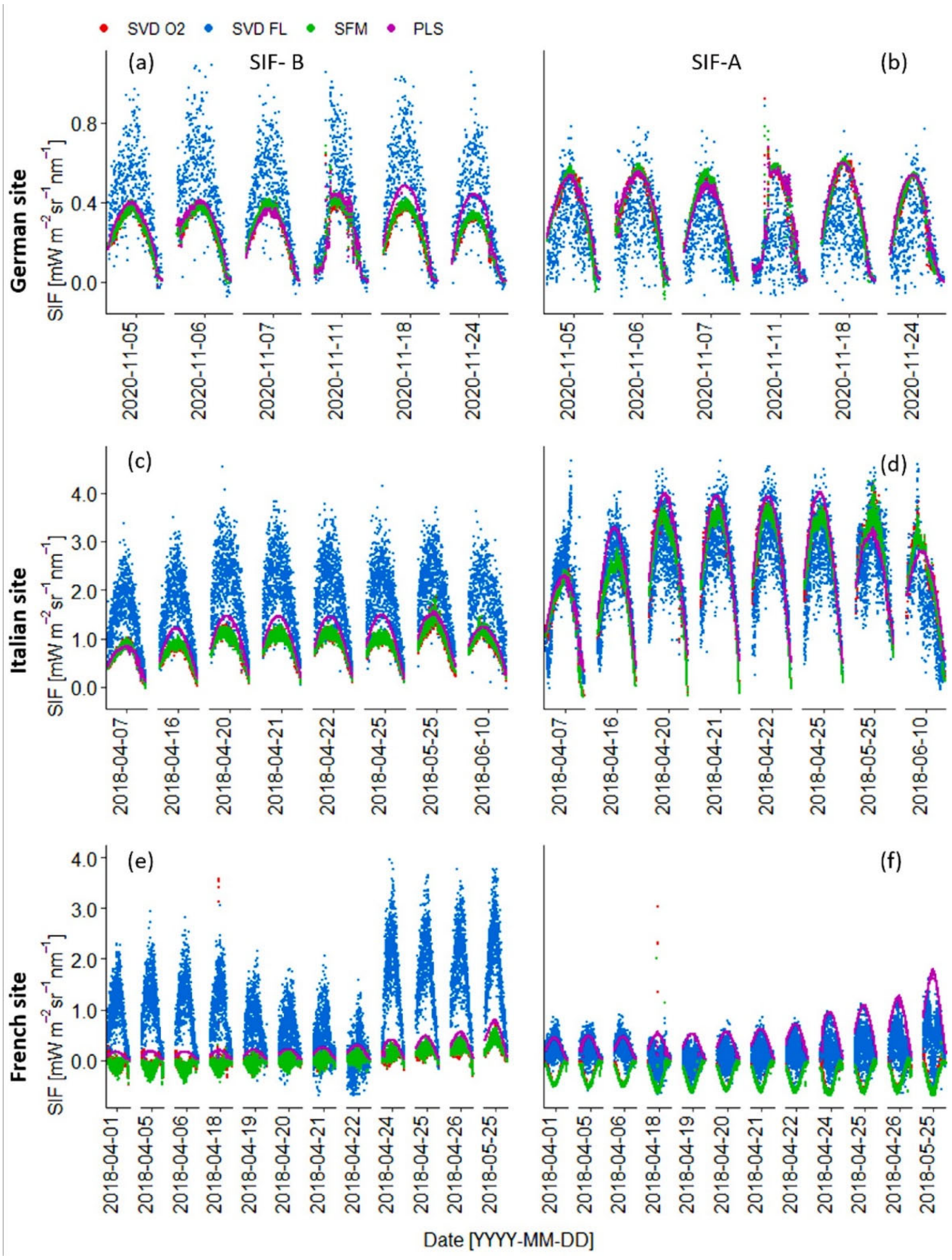


Fig. 7. Clear sky days selected for the retrieval of red SIF-B at 687 nm (left panel) and far-red SIF-A at 760 nm (right panel) with SVD-O2 (red), SVD- FL (blue), SFM (green) and PLS (purple). PLS models for retrieval have been trained on synthetic up-welling radiances data considering noise ac- cording to in-field conditions, excluding telluric absorption features to exploit solar Fraunhofer lines. Measurements were obtained at TOC over Rumex in Germany (a, b), at TOC over Alfalfa and Forage in Italy (c, d) and 100 m above an oak forest canopy in France (e, f) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Table 3. Root Mean Square Error (RMSE) of PLS retrieval excluding telluric absorption features, SVD O2 in the related atmospheric absorption bands and SVD FL in the solar Fraunhofer lines for SIF-B and SIF-A with respect to the SFM retrieved SIF, calculated in absolute values and percentage of the signal for the entire time-series.

Dataset	RMSEPLSSIF-B	RMSEPLSSIF-A	RMSESVD FLSIF-B	RMSESVD FLSIF-A	RMSESVD O2SIF-B	RMSESVD O2SIF-A	Unit
German site (TOC)	0.054	0.043	0.227	0.297	0.014	0.024	mW m^{-2} sr^{-1} nm^{-1}
	27.5	17.7	79.8	30.2	4.5	2.2	%
Italian site (TOC)	0.15	0.044	1.166	0.639	0.041	0.16	mW m^{-2} sr^{-1} nm^{-1}
	26.9	5.0	117.9	19.9	4.6	1.6	%
French site (100m above canopy)	0.724	1.572	1.443	0.718	0.067	0.050	mW m^{-2} sr^{-1} nm^{-1}
	55.3	366.1	187.7	178.1	4.3	3.4	%

3.2.2. Comparison of the result from the different retrieval methods

The performance of the PLS and SVD-FL for SIF retrieval based on solar Fraunhofer lines, and SVD-O2 based on telluric oxygen absorption bands, were evaluated against SFM retrieved SIF over 26 clear-sky days for three different instruments.

Similar RMSE values and, thus, difference with respect to SFM was found for PLS for the two instruments in Germany and Italy, which were both positioned relatively close to the canopy (Table 3). At the same time, SVD-O2 performed with very similar results to SFM in the two sites. This behavior was also reflected in the correlation of the SVD-O2 and PLS retrievals with respect to SFM, with very high R2 values (Fig. 8 a, b). The strongest correlation was observed between the telluric oxygen-line- based SVD-O2 and SFM with R2 0.99 and 0.98 in SIF-B and SIF-A, respectively. The slope and intercept described almost a one-to-one conversion from SFM to PLS with very low offset in SIF-B and a small multiplicative underestimation in SIF-A. The SVD-FL retrieval exhibited the highest disagreement with SFM (Table 3), poorest correlation and lowest R2 with SFM in the German site (Fig. 8 a, b). A multiplicative overestimation was recognized in SIF-B using SVD-FL retrieval with respect to SFM, with slope 1.7 and offset close to zero. On the contrary, a multiplicative underestimation with slope around 0.5 and very low offset was exhibited for SVD-FL with respect to SFM in SIF-A.

At the Italian site, PLS exhibited similar absolute but lower relative RMSE in SIF-A with respect to SFM, due to the higher absolute signal intensity compared to the German site (Table 3). The correlation with the SFM is high in SIF-B and SIF-A, with a small additive offset (Fig. 8 c, d). The multiplicative offset ranges around one. As with the German site, SVD-O2 exhibited strong agreement with SFM. SVD-O2 retrievals were highly correlated with SFM, with R2 around 0.98 in SIF-B and 0.97 in SIF-A, with slopes of 0.99 and offsets below 0.04. In contrast, SVD-FL retrieval in SIF-B and SIF-A exhibited the poorest correlation with SFM. The absolute and relative discrepancies were larger in SIF-B and SIF-A with SVD-FL in comparison to the other methods (Table 3), with intercept and slope deviating strongly from the one-to-one conversion (Fig. 8 c, d).

The instrument in France was positioned at 100 m distance to canopy without atmospheric correction applied to any of the SIF retrievals. High disagreement of PLS with respect to SFM and SVD-O2 were observed in both SIF-B and SIF-A (Table 3). Very high absolute and percentage RMSE was also found for SIF values retrieved using SVD-FL with respect to SFM. The performance of PLS and SVD-FL, which both exploited solar Fraunhofer lines and remained in principle unaffected by atmospheric distortion, differed from the telluric oxygen band based SFM and SVD- O2 in this site. SVD-O2 was again highly consistent with SFM with high R2 values in SIF-B and SIF-A (Fig. 8 d, e). In SIF-B, all retrieval methods exhibited regression lines with positive slope, with SVD-O2 and PLS below one and SVD-FL even above two. While the offset for SVD-O2 was negligible, SVD-FL exhibited a high offset in SIF-B. A smaller offset was recognized in SIF-B using PLS retrieval. Notably, PLS retrieved SIF-A was negatively correlated with SFM as a result of the atmospheric influence on the signal for this site. The SVD-FL retrieval exhibited a flat, negative slope around -0.07. The offsets in SIF-A for the two methods PLS and SVD-FL were both positive and similar (0.16 and 0.2, respectively).

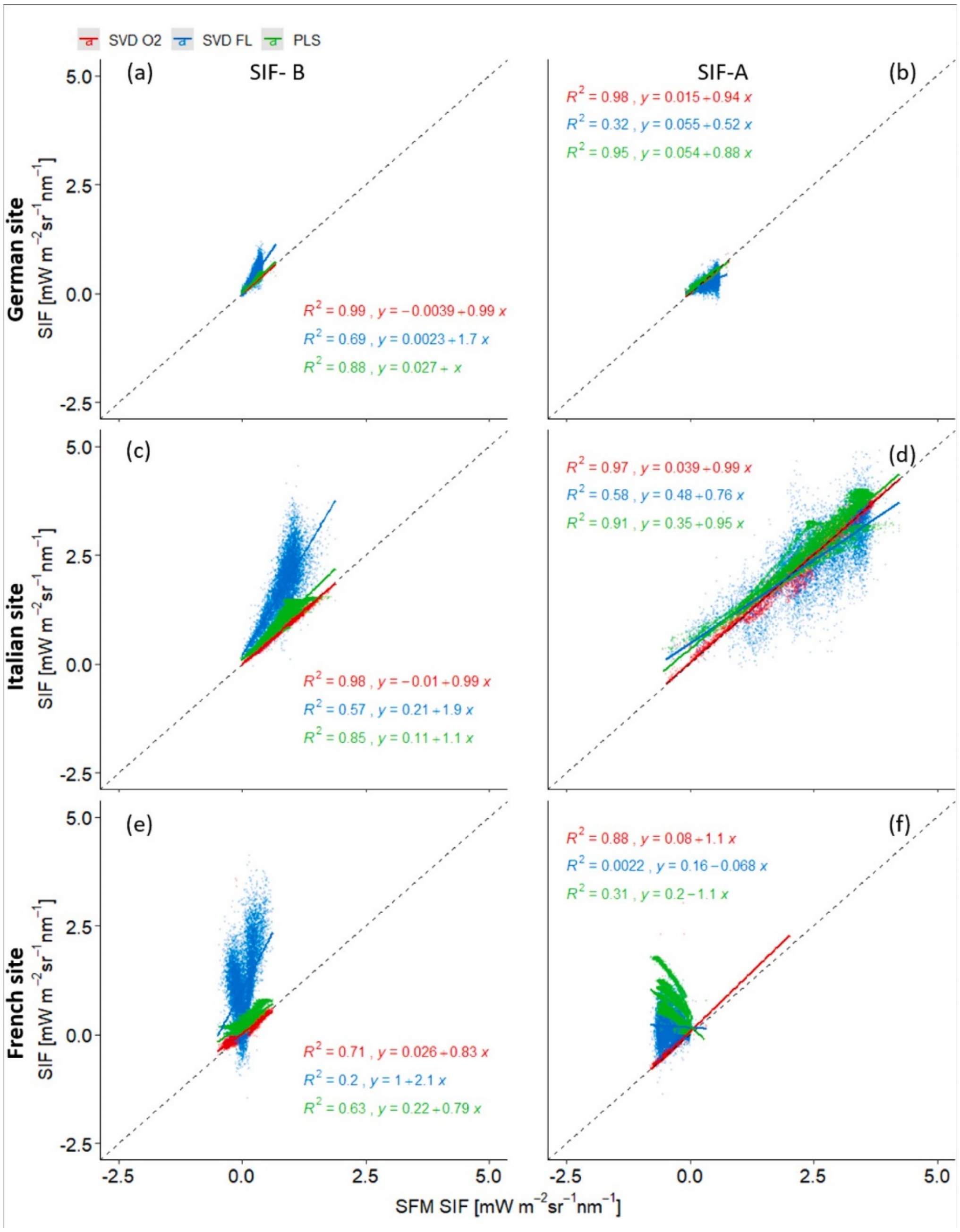


Fig. 8. Correlation of red SIF-B at 687 nm (left panel) and far-red SIF-A at 760 nm (right panel) with SVD-O2 (red), SVD-FL (blue) and PLS (green) with respect to SFM SIF from field measurements. PLS models were trained on synthetic up- welling radiances data considering noise according to in-field conditions, excluding telluric absorption features. Measurements were obtained from field sites located at TOC in Germany (a, b), at TOC Italy (c, d) and 100m above canopy France (e, f) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

4. Discussion

In the following two sections, we first discuss the results from modelling and testing with synthetic up-welling radiance data. We then discuss real world retrieval (i.e., observational data) results from up- welling radiance measured using FloX monitoring field spectrometers at three field sites.

4.1. Testing against modelled data

SIF was retrieved using a multivariate PLS model. We showed that the algorithm finds and exploits the covariance between SIF and specific spectral features in the first spectral derivative of up-welling radiance in the fitting window between 650 nm and 810 nm. For the spectral configuration including telluric absorption, covariance is predominantly observed around the O2-B band, centered around 687 nm, the water absorption feature between 715 nm and 735 nm, and the O2-A absorption feature around 760nm. Solar Fraunhofer lines are of minor contribution across the spectral range in this configuration, for they are very narrow. However, for training and testing data excluding telluric absorption features we identified significant covariance in the solar Fraunhofer lines across the spectral range of the detector. This means that the first derivative of up-welling radiance varies within the Fraunhofer lines due to the infilling of the varying SIF contribution.

Noise in the synthetic training and testing data significantly influenced the retrieval RMSEP in synthetic testing data. The retrieval error increased on a logarithmic scale by more than an order of magnitude when scaling noise was added to the ideal, noise-free, synthetic detector. In order to represent the expected characteristics of the FloX instruments in the field, noise must be included in the training data with respect to the automatic optimization at 80% signal of the total dynamic range of the detector. In operational scenarios were SNRs around 390 in the red and around 800 in the NIR shoulder region reported. SNR depends on the signal level at the detector and therefore changes with respect to wavelength depending on the spectral shape of the target. In effect, high noise and spectral drift affect the depicted shape and, thus, the accurate approximation of the infilling in the small solar Fraunhofer lines. Thus, instrumental noise is considered as an important limiting factor to the performance of the approach. This will be especially relevant when using the retrieval with different ground-based and air-borne SIF systems with varying SNR because of their various optical configurations. Therefore, we recommend a minimum operational SNR of around 150 in the NIR shoulder spectral region for the application of PLS retrieval based on the presented results.

In addition to the SNR, the spectral resolution (FWHM) of the spectrometer limits the maximum depth of the absorption bands which can be effectively exploited for SIF retrieval (Julitta et al., 2016; Pacheco-- Labrador et al., 2019a). The investigated FloX sensors featured a spectral resolution around 0.3 nm FWHM. Further research should investigate the PLS retrieval with other optical configurations, especially with different FWHM and SNR, to reproduce other instruments currently used in the community. The challenge of the PLS approach resides in the composition of the synthetic training and testing dataset with respect to the optimal level of added noise to model real conditions as closely as possible. Many other factors superimpose the fluorescence signal in remote measurements and require consideration to provide comparable readings (Damm et al., 2010b). One has to cover enough variance of these superimposed factors and instrumental noise for to identify these as components that are not correlated with the signal. At the same time, the synthetic data has to cover enough variance of fluorescence in the continuum so that the algorithm can correlate the components related to the actual signal. However, if the modelled data is very noisy, instances can potentially fall on multiple components at the same time and, thus introduce artifacts into the retrieved values. These artifacts are rather tricky to identify with common statistical error measures, as they often still represent the solution of best fit. In the retrieval of real measured SIF appear artifacts often as singular outliers, sudden jumps or shape inversions. However, since we cannot assume continuous trends in unknown data, contingency in real measurements must be respected. Thus, we strongly recommend to strive for an instrumental solution with optical specifications, accordingly. The performance of the algorithm is severely limited by the detector's capability to accurately measure the depth and infilling in relatively small spectral features. This was especially obvious when using the SVD-FL retrieval for SIF-B in this study. Other sensors, such as FLORIS and HyPlant, designed for SIF retrieval, exceed FloX-similar optical requirements (Drusch et al., 2017; Siegmann et al., 2019).

We investigated the model training and testing performance for two different configurations of the synthetic up-welling radiance data with noise to resemble in-field measurements: (1) including the full spectrum between 650 nm and 810 nm, (2) excluding spectral regions of major telluric absorption features. In all configurations, we observe for each model an increasing uncertainty of prediction with increasing number of components. This is known as overfitting, where the model increasingly tracks noise instead of the real signal. Conversely, the uncertainty of prediction also increases for a very low number of components. This is due to the model not tracking enough of the signal at low complexity (under-fitting). We identified the optimum where the mean error is lowest and the variation in the error is small. As demonstrated in Wagner et al. (2018), the first minimum of absolute RMSEP and smallest standard deviation indicate the optimal number of components for prediction for each model. No significant difference in mean RMSEP of the two configurations, i.e. either including or excluding the spectral regions of major telluric absorption features, was found in the PLS model testing with synthetic data. In particular, the variability attributed to changing detector noise or model complexity was significantly larger than the variability attributed to spectral configurations. The full spectral configuration performed slightly less retrieval error than the configuration excluding telluric absorption. Due to the higher complexity of spectral information available from the full spectrum, this difference in performance was anticipated. However, an RMSE around 0.1 mW m-2 sr-1 nm-1 for the PLS configuration excluding telluric absorption is still very similar to the SFM (Cogliati et al., 2015b). Atmospheric absorption is known to distort telluric retrievals of the SIF signal with increasing distance over the canopy, which requires correction using complex atmospheric compensation (Aasen et al., 2019). Therefore, we further investigated the PLS model configuration, which excludes the telluric absorption features and only exploits the solar Fraunhofer lines, thereby circumventing atmospheric reabsorption (Guanter et al., 2013).

We emphasize that the PLS regression model is a data driven approach similar in some aspects to SVD, nevertheless with some important differences. Both SVD and PLS employ dimensionality reduction. SVD uses a forward model and residuals to estimate the SIF signal, whereas PLS utilizes a direct regression between SIF and the deflated first spectral derivative of up-welling radiance spectra. SVD decomposes the variable space into principal components which are identified and ordered according to the amount of variance covered (Mardia et al., 1979). In contrast, PLS determines the response variable SIF through the projection of latent variables in such a way that scores and loadings explain as much covariance between predictor and response variable as possible (Mevik and Wehrens, 2007). Typically, the first latent variables represent highly correlated components of the predictor data space with respect to the response variable. However, it is possible that latent variables are equally weighted with respect to explained covariance between predictor and response variable (de Jong, 1993). Note that neither SVD nor PLS can model variations in the measurements which they were not trained for. The PLS algorithm has been specifically optimized for detection of signals in mixture in general and relies on a considerate construction of synthetic training and testing data to accurately resemble real measurements with modelled SIF contribution in this study.

We assumed that the modelled synthetic up-welling radiance data can resemble in-field conditions and, thus, can be transferred to a SIF retrievals applied to the real measurements (Pacheco-Labrador et al., 2019b). This assumption was further investigated with a comparison of the SFM, SVD-O2, SVD-FL and PLS retrieval against SCOPE simulated SIF as a reference. The results suggest a strong dependency of the SVD-FL retrieval error on the signal levels in both SIF-B and SIF-A. Thus, the spectral shape of the SIF signal is affecting this retrieval significantly. The spectral resolution of 0.3nm FWHM of the FloX, compared with 0.13nm in the literature is another limiting factor for the SVD-FL retrieval (Guanter et al., 2013). SFM was the most stable retrieval in this regard and therefore SFM was considered the reference to evaluate the other retrieval methods with field measurements. SVD-O2 and PLS showed similar errors compared with the SFM. Only very low SIF values were retrieved with an increased number of outliers. Note that SVD requires information about the fluorescence shape for precise retrieval, which can introduce error if the assumed fluorescence shape is incorrect (Chang et al., 2020a). PLS avoids this issue as no forward model is fitted to the up-welling radiance spectrum. Instead, the algorithm computes regression coefficients in each waveband across the fitting window by rotating the input matrix of spectral derivatives in such a way that covariance with the SIF variable is maximized.

4.2. Real world retrieval

SIF-B and SIF-A were retrieved with a PLS regression model approach configured to exploit solar Fraunhofer lines, considering instrumental noise according to in-field conditions from measured data from three different sites. Speeding up the measurement can be beneficial, especially for experiments aimed at investigating fast dynamics in fluorescence. However, simultaneous (or near-simultaneous) down-welling measurements are of great benefit for computing fluorescence yield and light use efficiency, even if more time consuming (Damm et al., 2010a; Rascher et al., 2010). The retrieval with SFM also required continuous down-welling light measurements, which are included together with automatic signal level optimization in each FloX measurement cycle.

We were able to retrieve positive and meaningful values both in SIF- B and SIF-A from FloX measurements with the PLS configuration in all sites. In close distance to canopy, PLS results agreed well with the oxygen-band-based retrieval methods SFM and SVD-O2. The most striking differences were found between retrievals based on telluric oxygen absorption lines and PLS retrieval at the 100 m tall tower in France. Complex atmospheric correction methods are required for SFM and SVD-O2 retrievals to achieve meaningful results with increasing distance to canopy (Sabater et al., 2018). With no atmospheric correc- tion in place, SFM and SVD-O2, both based on telluric oxygen absorption, retrieved negative SIF from the French site and were therefore unusable for further analysis. The German and Italian setups were both in relatively close proximity to their target canopies, and both sites exhibited similar agreement across PLS, SVD-O2 and SFM. In particular, SVD-O2 and SFM based on telluric absorption performed equally well with almost identical results. PLS slightly overestimated SIF with respect to SVD-O2 and SFM in the Italian site but exhibited very good agreement in the German site. Similar tendencies were observed comparing the retrieval methods against SCOPE simulated SIF in synthetic data. By exclusively exploiting solar Fraunhofer lines, the PLS retrieval is in principle unaffected by atmospheric distortion and clouds (Sun et al., 2018b). Furthermore, by using large fitting windows and exploiting multiple solar Fraunhofer lines between 650 nm and 810 nm, retrieval noise is kept very low and good robustness against atmospheric reabsorption is achieved with PLS. At the French site PLS also captured a gradual increase in SIF-B and SIF-A over time, which was related to the onset of photosynthesis during the greening-up of the forest canopy. This increase in fluorescence was also observed through SVD-FL but was not sufficiently resolved on a diurnal scale due to the high retrieval noise.

Overall, SVD-FL results were much noisier than PLS, SFM or SVD-O2. This could be due to instrumental limitations, mainly FWHM, and the small fitting window of the algorithm. The performance of the SVD algorithm also relies on accurate fits of a SIF-free model with measured up-welling radiance data (Du et al., 2018). Previous studies indicate that the exploitation of solar Fraunhofer lines using SVD was superior with a spectral resolution around 0.13 nm FWHM (Chang et al., 2020a; Guanter et al., 2013). FWHM and SNR affect the accurate exploitation of the SIF infilling and thus limiting retrieval accuracy further (Frankenberg and Berry, 2018; Liu et al., 2015). The FloX is commonly configured with 0.3 FWHM. Evidently, Frankenberg and Berry (2018) show the influence of spectral resolution affecting the apparent depth of the measured solar Fraunhofer lines in the same spectral window, which is also used for the SVD-FL retrieval in this work. As FWHM increases, the difficulty of accurately measuring the line depth increases, and the effect of atmospheric scattering on the infilling becomes more noticeable. When applied to the telluric oxygen absorption features, SVD exhibits similar retrieval noise as SFM, but loses the benefit of independence from atmospheric conditions. When exploiting narrow solar Fraunhofer lines, the SVD-FL is considered more sensitive to instrument noise, spectral resolution or drift due to the low SIF signal and narrow spectral features from field measurements (Chang et al., 2020a). Scattering and hotspot effects related to canopy structure and BRDF depending on the angle of incident light can substantially influence up-welling radiance and retrieved SIF (Pacheco-Labrador et al., 2016). Furthermore, SIF signals are known to be angle-dependent and directed in complex canopies (Rautiainen et al., 2018; Van Wittenberghe et al., 2015). Thus it is necessary to clarify that only at-sensor fluorescence can be directly measured without any inference of target canopy structure or scattering in the atmospheric path (Damm et al., 2014). Bearing this consideration in mind, a slight disagreement in absolute values between the different retrieval methods is expected due to the different physical principles on which they are based. Still, the diurnal shape of all retrievals should still produce similar patterns, unless the physical principle is obstructed by independent factors, e.g., atmospheric distortion. PLS outperformed SVD-FL in terms of retrieval noise and precision based on the presented results. Hence, the PLS algorithm appears more robust and suitable for SIF-retrieval based on solar Fraunhofer lines with a spectrometer of lower resolution (around 0.3 nm FWHM), as it exploits many more solar Fraunhofer lines in a wider fitting window.

Since PLS is a data driven approach, it requires training. A limiting factor to the retrieval of real-world data is the composition of the syn- thetic data for the PLS model training process. The employed SCOPE modelled vegetation data is based on actual in-field measurements over a grassland to allow for realistic variances of the signal covered in the synthetic dataset. We are aware that the vegetation spectra modelled with SCOPE do not represent all possible variations in canopy characteristics. Therefore, uncertainties in SIF retrieval can vary with other canopy characteristics, especially for complex structures. We assumed in this work that the modelled spectral fluorescence based on grassland can be extrapolated and transferred to other canopies. This assumption is based on the findings that the presented PLS algorithm exploits covariance between the pure SIF signal and first spectral derivative of up-welling radiance in the solar Fraunhofer lines. Furthermore, the assumption is supported by the results from comparing SFM, SVD-O2, SVD-FL and PLS against SCOPE simulated SIF, following the approach of Cogliati et al. (2015b) and achieving comparable retrieval errors. Finally, our retrieval results of PLS were similar to SVD-O2 and SFM in TOC settings. However, we anticipate increased uncertainty of prediction in structurally complex canopies. To address this limitation and account for structurally more complex canopies, a strong permutation of SCOPE modelled SIF and reflectance spectra is used in the training of the PLS models. Thus, each simulated SIF spectrum was combined with each simulated reflectance spectrum. This strong permutation by over- sampling increases the distinct instances in the training data and pro- vided larger variation of canopy characteristics. Thus, the PLS model becomes more capable of disentangling SIF from more diverse measured data, and improves the transferability of the PLS retrieval. At the same time, ambivalence and retrieval noise is reduced. However, this trick does not account for all variability in the spectral continuum of SIF and reflectance which can be affected by multiple factors in natural canopies in the field (Verrelst et al., 2016a, 2015). A thorough investigation optimizing a SCOPE dataset, specifically modelled for the training of statistical retrievals, could improve the robustness and transferability of machine-learning-algorithm-based retrieval approaches into specific canopy characteristics. However, this was beyond the scope of this study.

5. Conclusion

PLS models trained with SCOPE modelled SIF and reflectance transfer well to real measurements and retrieve values in agreement with other mechanistic (SFM) and statistical (SVD) telluric retrieval methods in TOC settings. As PLS is a supervised approach, its main limitations are identified in the model training and testing with SCOPE modelled up-welling radiance data of known SIF contribution. Further research is recommended to improve the performance of the PLS model in various and more complex canopies. We partially addressed this issue with a robust permutation approach, which oversamples the SCOPE modelled data to increase the number of possible SIF and canopy reflectance combinations. In addition, PLS models are sensitive to instrumental noise. The retrieval error increases with poorer SNR in a logarithmic relationship. For the retrieval of SIF in very noisy data, for example due to low-light conditions in which automatic signal optimization of the FloX is not possible, the PLS method is not very well suited. Furthermore, PLS and SVD retrieval methods require measurements of down-welling light for the model training process. Thus, it is recommended to acquire the measurements of down-welling light preferably with a white panel through the up-welling channel to accurately include the instrument's optical response. The instrumental configuration with a spectral resolution of 0.3nm FWHM provides a further limitation to the exploitation of the solar Fraunhofer lines with the FloX, especially recognized in the SVD-FL retrieval in combination with narrow fitting windows. In this context, the PLS approach was found more robust. For the first time, we have been able to retrieve positive and meaningful SIF- B and SIF-A values without any atmospheric correction from measured up-welling radiances on a 100m tower above ground, regardless of the reabsorption of signal in the atmospheric column, using PLS. The exploitation of solar Fraunhofer lines for SIF retrieval is very promising for being in principle independent of changing atmospheric disturbances. Since the solar Fraunhofer lines are distributed across the entire spectrum of SIF, more spectral information is available for the detection of SIF using the PLS regression. This additional spectral information reduces retrieval noise and could open a way to unveil the full spectral shape of fluorescence using PLS in the future. We consider the PLS regression model for SIF retrieval from solar Fraunhofer lines particularly promising under conditions for which retrieval methods based on telluric oxygen absorption require complex and computationally costly data correction. In consequence, PLS does not account for atmospheric correction but simply bypasses atmospheric reabsorption. Future research is encouraged to investigate also the performance of SIF retrievals under diffuse light conditions. On an ordinary consumer-grade laptop, PLS models were computed for an entire diurnal dataset within 6.62 s. Compared to the SFM method, this is 37.25 times faster. This is also superior to five other machine learning algorithms tested with and without dimensionality reduction in advance of this study.

In summary, we investigated the potential of PLS as an approach for SIF retrieval in modelled data and real measurements from autonomous field-spectrometers with promising results. Our results suggest that PLS is superior, compared with other machine learning algorithms, in retrieving SIF signals in the mixture of hyperspectral, up-welling radiance. The fast computation time makes the approach especially appealing for fast processing to overview large datasets for which complex atmospheric correction was required otherwise and for future application of PLS with imaging high-resolution hyperspectral data. Furthermore, PLS exhibits very high precision compared with solar Fraunhofer line-based SVD. At the same time, PLS based on exploiting the infilling of solar Fraunhofer lines is highly robust against atmospheric reabsorption, compared with telluric oxygen-based methods.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data will be published separately and the DOI referenced in manuscript upon acceptance.

Acknowledgment

Field measurements were carried out under lead management and funding of European Space Agency in the frame of AtmoFLEX (4000122454/17/NL/FF/mg). This research was supported by the Action CA17134 SENSECO (Optical synergies for spatiotemporal sensing of scalable ecophysiological traits) funded by COST (European Cooperation in Science and Technology, www.cost.eu). The SCOPE modelled dataset and field-measured dataset from the instruments used in this study are available athttps://doi.org/10.5281/zenodo.7040578.

Appendix A1. Assessment of machine learning for SIF retrieval

Machine learning algorithms (MLA) are tools from data science, which provide methods to extract useful information from a number of input variables. Emulators and toolboxes have been developed to aid model inversion and retrieval of biophysical parameters from SCOPE using a variety of MLA (Berger et al., 2020; Rivera-Caicedo et al., 2014; Rivera et al., 2015; Verrelst et al., 2012). Typically, supervised MLA models are trained to predict the variable Y in the predictor space of X. Redundant information in a high-dimensional, hyperspectral predictor spaces are identified and eliminated using dimensionality reduction in a pre-processing step (Verrelst et al., 2017). Principal component analysis (PCA) decomposes the hyperspectral predictor space into components in such a way, that most of the variance in the data is explained by the first few components.

We tested six MLA and a linear regression model with and without applying PCA during the preprocessing. Neural Network (NN), Gaussian Process Regression (GPR), Elastic Net regression (ENet), Conditional Inference Random forest (CForest), Support Vector Machine (SVM) and a Linear Model (LM) were assessed in addition to the Partial Least Squares regression (PLS). Training and testing data were composed in the same fashion as presented above and the first spectral derivative was computed. Likewise, 60% training data and 40% testing data were split with a total number of 718 variables in the predictor space. A PCA was performed in the preprocessing and the principal components were tested for significance with the randomization procedure (Dray, 2008). For that purpose, axis in the training data were randomized in 900 repetitions. According to Eq. (A1), the cumulative proportion of variance explained σ2 for n components was computed from the standard deviation σ2 of each component i, with respect to the total variance of all m=718 components, expressed by the Eigenvalues of the rotated matrix using PCA:

$$\sigma^2_{\mathrm{cum}[\%],\ \mathrm{n}} = 100 \cdot \sum_{i=1}^{n} \sigma_i^2 \Big/ \sum_{i=1}^{m} \sigma_i^2 \tag{A1}$$

We identified no components, which were significant on a 95% confidence interval. Thus, 95% explained variance is statistically not missing any important information from the original dataset. In consequence, the number of reduced dimensions from the PCA was selected to explain 95% of the variance in the total predictor space, which resulted in using the first five components (see Fig. A1).

A k-fold, random, segmented (k=4) cross-validation approach with four repeats was used during the training. Training and testing procedures were repeated 30 times for each algorithm in SIF-A and SIF-B, respectively. The RMSE was computed in each run using k-fold cross-validation for the training fraction of the dataset during the training and for the predicted SIF values using the testing fraction of the dataset during the testing with respect the actual SCOPE SIF.

The mean errors and standard deviation are reported for each MLA with and without dimensionality reduction in Fig. A2. The associated mean computing time and standard deviation for one training are reported in Fig. A3. Our results suggest that PLS exhibited similar prediction errors to GPR, ENet or SVM without applying PCA during the preprocessing. The MLAs exhibited a significantly increased training and testing error in SIF retrieval when combined with PCA. Only in terms of computing time benefit CForest, Enet and SVM from dimensionality reduction slightly, while GPR, NN and LM need even more time to converge the internal optimization. However, PLS outperformed all other MLA significantly in terms of computing time. The results are explained since the SIF signal contributes only about 2% to the variance of the total spectral continuum of up-welling light. PCA compresses the components in such a way, to maximize the total variance of the predictor space within the first components. However, due to this compression is the information lost, which has only a small contribution to the continuum. Thus, PCA is limiting the MLAs in performing the internal optimization to find and exploit the fluorescence contribution. Note that differences in implementation of the MLAs between different environments (e.g. MATLAB, R, Python.) should also be considered when comparing the results with different studies. Given the high dimensionality of the hyperspectral data with 718 bands and the computed PCA explaining 95% of its variance within the first five components, we consider the results meaningful. In contrast, PLS exploits covariance between the predicted parameter (SIF in this case) and the predictor space (entity of first spectral derivatives of up-welling radiance in this case) by reducing the number of variables into correlated orthogonal scores. The PLS algorithm has been optimized to find and to obtain quantitative information, which contribute only small covariance to the spectra in hyperspectral near-infrared spectroscopy (Biancolillo and Marini, 2018; Jiang et al., 2020; Jin and Wang, 2019). PLS outperformed all other tested MLAs with and without PCA preprocessing in our study. The drawback of PCA, losing small covariance when compressing the data into principal components, was critical. We showed that PLS, by design, is superior in finding and exploiting small signals in mixture and, thus, is considered superior for the retrieval of SIF.

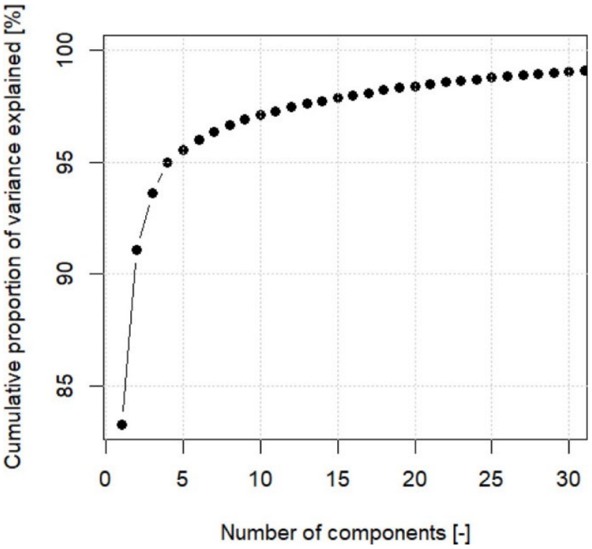


Fig. A1. Cumulative proportion of variance of the total predictor space explained in the training data by the first 30 principal components, computed by PCA in the preprocessing of training the different MLAs.

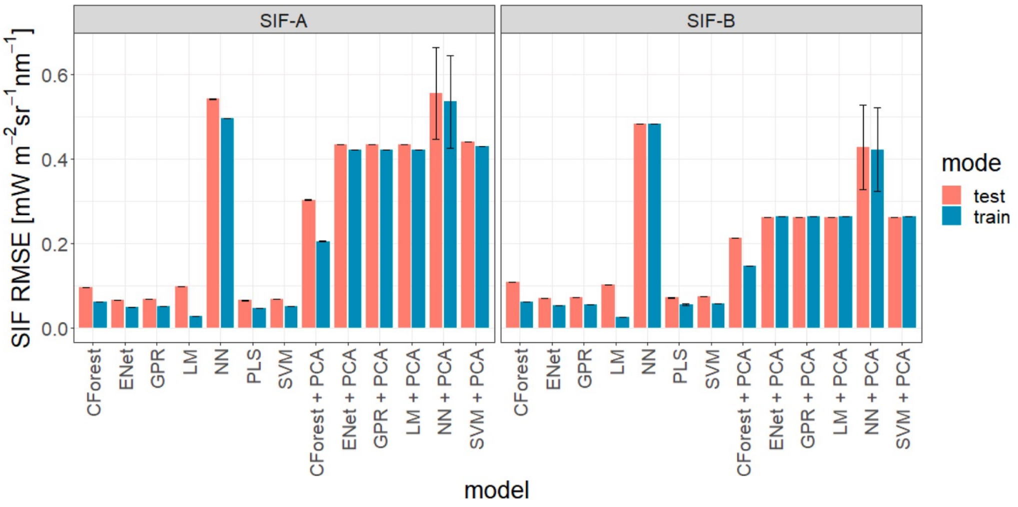


Fig. A2. Error of SIF prediction from six different MLA and linear regression with and without using PCA in the preprocessing with respect to SCOPE simulated SIF. RMSE was calculated separately for training and testing in 30 repetitions, with SIF-A or SIF-B as response variable, respectively. Whiskers show standard deviation.

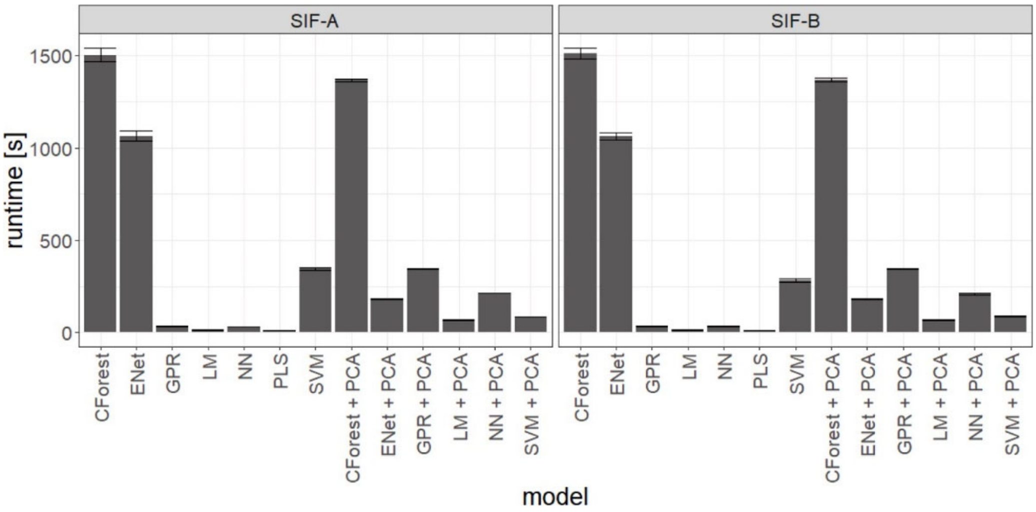


Fig. A3. Computing time in seconds from six different MLA and linear regression with and without using PCA in the preprocessing. Mean runtime was calculated separately for training with SIF-A or SIF-B as response variable, respectively. Whiskers show standard deviation.

Appendix A2. SCOPE simulated data

Following the approach of Cogliati et al. (2015b), SCOPE simulated SIF and reflectance spectra were used to assess the retrieval performance of the SFM, SVD-FL and SVD-O2 retrievals. The simulated spectra were used to compute up-welling radiance which resembles the optical configuration of the FloX. The range of SIF and reflectance is given for 212 different instances in Fig. A4.

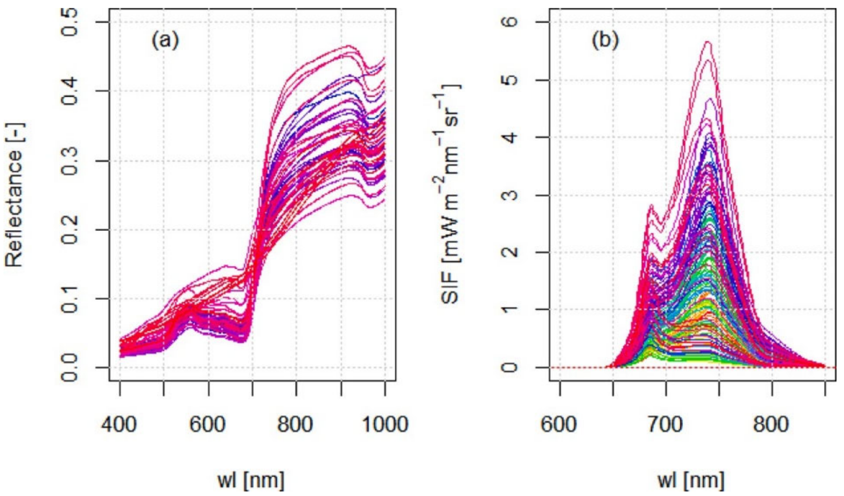


Fig. A4. Range of SCOPE simulated reflectance (a) and SIF (b) of the 212 different cases used for the composition of the combined training and testing datasets.

References

Aasen, H., Van Wittenberghe, S., Medina, N.S., Damm, A., Goulas, Y., Wieneke, S., Hueni, A., Malenovský, Z., Alonso, L., Pacheco-Labrador, J., Cendrero-Mateo, M.P., Tomelleri, E., Burkart, A., Cogliati, S., Rascher, U., Arthur, A.Mac, 2019. Sun-induced chlorophyll fluorescence II: Review of passive measurement setups, protocols, and their application at the leaf to canopy level. Remote Sens. https://doi. org/10.3390/rs11080956.

Acebron, K., Matsubara, S., Jedmowski, C., Emin, D., Muller, O., Rascher, U., 2021. Diurnal dynamics of nonphotochemical quenching in Arabidopsis npq mutants assessed by solar-induced fluorescence and reflectance measurements in the field. New Phytol. 229, 2104–2119. https://doi.org/10.1111/nph.16984.

Agati, G., Mazzinghi, P., Fusi, F., Ambrosini, I., 1995. The F685/F730 chlorophyll fluorescence ratio as a tool in plant physiology: response to physiological and environmental factors*. J. Plant Physiol. https://doi.org/10.1016/S0176-1617(11) 81882-1.

Alonso, L., Go´mez-Chova, L., Vila-Franc´es, J., Amoro´s-Lo´pez, J., Guanter, L., Calpe, J., Moreno, J., 2008. Improved fraunhofer line discrimination method for vegetation fluorescence quantification. IEEE Geosci. Remote Sens. Lett. 5, 620–624. https://doi. org/10.1109/LGRS.2008.2001180.

Alonso, L., Van Wittenberghe, S., Amoro´s-Lo´pez, J., Vila-Franc´es, J., Go´mez-Chova, L., Moreno, J., 2017. Diurnal cycle relationships between passive fluorescence, PRI and NPQ of vegetation in a controlled stress experiment. Remote Sens. https://doi.org/ 10.3390/rs9080770.

Bendig, J., Malenovsky, Z., Gautam, D., Lucieer, A., 2019. Solar-induced chlorophyll fluorescence measured from an unmanned aircraft system: sensor etaloning and platform motion correction. IEEE Trans. Geosci. Remote Sens. 1–8. https://doi.org/ 10.1109/TGRS.2019.2956194.

Berger, K., Verrelst, J., F´eret, J.B., Hank, T., Wocher, M., Mauser, W., Camps-Valls, G., 2020. Retrieval of aboveground crop nitrogen content with a hybrid machine learning method. Int. J. Appl. Earth Obs. Geoinf. 92, 102174 https://doi.org/ 10.1016/j.jag.2020.102174.

Biagioni, D.J., Astling, D.P., Graf, P., Davis, M.F., 2011. Orthogonal projection to latent structures solution properties for chemometrics and systems biology data. J. Chemom. 25, 514–525. https://doi.org/10.1002/cem.1398.

Biancolillo, A., Marini, F., 2018. Chemometric methods for spectroscopy-based pharmaceutical analysis. Front. Chem. 6, 576. https://doi.org/10.3389/ fchem.2018.00576.

Brault, J., Neckel, H., 1999. Announcement spectral atlas of solar absolute disk-averaged and disk-center intensity from 3290 to 12510 Å (Brault and Neckel, 1987) now available from Hamburg observatory anonymous ftp site. Sol. Phys. 184, 421. https://doi.org/10.1023/A:1017165208013. –421.

Burkart, A., Schickling, A., Mateo, M.P.C., Wrobel, T.J., Rossini, M., Cogliati, S., Julitta, T., Rascher, U., 2015. A method for uncertainty assessment of passive sun- induced chlorophyll fluorescence retrieval using an infrared reference light. IEEE Sens. J. 15 https://doi.org/10.1109/JSEN.2015.2422894.

Camino Id, C., Zarco-Tejada, P.J., Gonzalez-Dugo, V., 2018. Effects of heterogeneity within tree crowns on airborne - quantified SIF and the CWSI as indicators of water stress in the context of precision agriculture. Remote Sens. 10 https://doi.org/10.3390/rs10040604.

Campbell, P., Huemmrich, K., Middleton, E., Ward, L., Julitta, T., Daughtry, C., Burkart, A., Russ, A., Kustas, W., 2019. Diurnal and seasonal variations in chlorophyll fluorescence associated with photosynthesis at leaf and canopy scales. Remote Sens. 11, 488. https://doi.org/10.3390/rs11050488.

Caporaso, N., Whitworth, M.B., Grebby, S., Fisk, I.D., 2018. Rapid prediction of single green coffee bean moisture and lipid content by hyperspectral imaging. J. Food Eng. 227, 18–29. https://doi.org/10.1016/j.jfoodeng.2018.01.009.

Cavender-Bares, J., Meireles, J., Couture, J., Kaproth, M., Kingdon, C., Singh, A., Serbin, S., Center, A., Zuniga, E., Pilz, G., Townsend, P., 2016. Associations of leaf spectra with genetic and phylogenetic variation in oaks: prospects for remote detection of biodiversity. Remote Sens. 8, 221. https://doi.org/10.3390/rs8030221.

Cendrero-Mateo, M.P., Moran, M.S., Papuga, S.A., Thorp, K.R., Alonso, L., Moreno, J., Ponce-Campos, G., Rascher, U., Wang, G., 2016. Plant chlorophyll fluorescence: active and passive measurements at canopy and leaf scales with different nitrogen treatments. J. Exp. Bot. 67, 275–286. https://doi.org/10.1093/jxb/erv456.

Cendrero-Mateo, M.P., Wieneke, S., Damm, A., Alonso, L., Pinto, F., Moreno, J., Guanter, L., Celesti, M., Rossini, M., Sabater, N., Cogliati, S., Julitta, T., Rascher, U., Goulas, Y., Aasen, H., Pacheco-Labrador, J., Arthur, A.M., 2019. Sun-induced chlorophyll fluorescence III: benchmarking retrieval methods and sensor characteristics for proximal sensing. Remote Sens. 11 https://doi.org/10.3390/ rs11080921.

Chang, C.Y.Y., Guanter, L., Frankenberg, C., Ko¨hler, P., Gu, L., Magney, T.S., Grossmann, K., Sun, Y., 2020a. Systematic assessment of retrieval methods for canopy far-red solar-induced chlorophyll fluorescence (SIF) using high-frequency automated field spectroscopy. J. Geophys. Res. Biogeosci. https://doi.org/10.1029/ 2019jg005533.

Chang, C.Y.Y., Zhou, R., Kira, O., Marri, S., Skovira, J., Gu, L., Sun, Y., 2020b. An unmanned aerial system (UAS) for concurrent measurements of solar-induced chlorophyll fluorescence and hyperspectral reflectance toward improving crop monitoring. Agric. For. Meteorol. 294, 108145 https://doi.org/10.1016/j. agrformet.2020.108145.

Cogliati, S., Celesti, M., Cesana, I., Miglietta, F., Genesio, L., Julitta, T., Schuettemeyer, D., Drusch, M., Rascher, U., Jurado, P., Colombo, R., 2019. A spectral fitting algorithm to retrieve the fluorescence spectrum from canopy radiance. Remote Sens. 11, 1840. https://doi.org/10.3390/rs11161840.

Cogliati, S., Rossini, M., Julitta, T., Meroni, M., Schickling, A., Burkart, A., Pinto, F., Rascher, U., Colombo, R., 2015a. Continuous and long-term measurements of reflectance and sun-induced chlorophyll fluorescence by using novel automated field spectroscopy systems. Remote Sens. Environ. 164, 270–281. https://doi.org/ 10.1016/j.rse.2015.03.027.

Cogliati, S., Verhoef, W., Kraft, S., Sabater, N., Alonso, L., Vicent, J., Moreno, J., Drusch, M., Colombo, R., 2015b. Retrieval of sun-induced fluorescence using advanced spectral fitting methods. Remote Sens. Environ. 169, 344–357. https://doi.org/10.1016/J.RSE.2015.08.022.

Damm, A., Elber, J., Erler, A., Gioli, B., Hamdi, K., Hutjes, R., Kosvancova, M., Meroni, M., Miglietta, F., Moersch, A., Moreno, J., Schickling, A., Sonnenschein, R., Udelhoven, T., van der Linden, S., Hostert, P., Rascher, U., 2010a. Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP). Glob. Chang. Biol. 16, 171–186. https://doi.org/10.1111/j.1365- 2486.2009.01908.x.

Damm, A., Guanter, L., Laurent, V.C.E., Schaepman, M.E., Schickling, A., Rascher, U., 2014. FLD-based retrieval of sun-induced chlorophyll fluorescence from medium spectral resolution airborne spectroscopy data. Remote Sens. Environ. 147, 256–266. https://doi.org/10.1016/j.rse.2014.03.009.

Damm, A., Schickling, A., Schaepman, M., Rascher, U., 2010b. Deriving Sun-Induced Chlorophyll Fluorescence from Airborne

Based Spectrometer Data, 2010. Hyperspectral Work, Frascati, IT, pp. 17–19, 2010.

Daumard, F., Champagne, S., Fournier, A., Goulas, Y., Ounis, A., Hanocq, J.F., Moya, I., 2010. A field platform for continuous measurement of canopy fluorescence. IEEE Trans. Geosci. Remote Sens. 48, 3358–3368. https://doi.org/10.1109/TGRS.2010.2046420.

Dayal, B.S., MacGregor, J.F., 1997. Improved PLS algorithms. J. Chemom. 11, 73–85. https://doi.org/10.1002/(SICI)1099-128X(199701)11:1<73::AID-CEM435>3.0.CO; 2-#.

de Jong, S., 1993. SIMPLS: an alternative approach to partial least squares regression. Chemom. Intell. Lab. Syst. 18, 251–263. https://doi.org/10.1016/0169-7439(93) 85002-X.

Dechant, B., Ryu, Y., Badgley, G., Ko¨hler, P., Rascher, U., Migliavacca, M., Zhang, Y., Tagliabue, G., Guan, K., Rossini, M., Goulas, Y., Zeng, Y., Frankenberg, C., Berry, J. A., 2022. NIRVP: a robust structural proxy for sun-induced chlorophyll fluorescence and photosynthesis across scales. Remote Sens. Environ. 268, 112763 https://doi. org/10.1016/j.rse.2021.112763.

Dechant, B., Ryu, Y., Badgley, G., Zeng, Y., Berry, J.A., Zhang, Y., Goulas, Y., Li, Z., Zhang, Q., Kang, M., Li, J., Moya, I., 2020. Canopy structure explains the relationship between photosynthesis and sun-induced chlorophyll fluorescence in crops. Remote Sens. Environ. 241 https://doi.org/10.1016/j.rse.2020.111733.

Dray, S., 2008. On the number of principal components: A test of dimensionality based on measurements of similarity between matrices. Comput. Stat. Data Anal. 52, 2228–2237. https://doi.org/10.1016/j.csda.2007.07.015.

Drolet, G., Wade, T., Nichol, C.J., MacLellan, C., Levula, J., Porcar-Castell, A., Nikinmaa, E., Vesala, T., 2014. A temperature-controlled spectrometer system for continuous and unattended measurements of canopy spectral radiance and reflectance. Int. J. Remote Sens. 35, 1769–1785. https://doi.org/10.1080/ 01431161.2014.882035.

Drusch, M., Moreno, J., Del Bello, U., Franco, R., Goulas, Y., Huth, A., Kraft, S., Middleton, E.M., Miglietta, F., Mohammed, G., Nedbal, L., Rascher, U., Schuttemeyer, D., Verhoef, W., 2017. The FLuorescence EXplorer mission concept- ESA's earth explorer 8. IEEE Trans. Geosci. Remote Sens. 55, 1273–1284. https:// doi.org/10.1109/TGRS.2016.2621820.

Du, S., Liu, L., Liu, X., Zhang, Xiao, Zhang, X., Bi, Y., Zhang, L., 2018. Retrieval of global terrestrial solar-induced chlorophyll fluorescence from TanSat satellite. Sci. Bull. 63, 1502–1512. https://doi.org/10.1016/J.SCIB.2018.10.003.

Fournier, A., Daumard, F., Champagne, S., Ounis, A., Goulas, Y., Moya, I., 2012. Effect of canopy structure on sun-induced chlorophyll fluorescence. ISPRS J. Photogramm. Remote Sens. 68, 112–120. https://doi.org/10.1016/j.isprsjprs.2012.01.003.

Frankenberg, C., Berry, J., 2018. Solar induced chlorophyll fluorescence: origins, relation to photosynthesis and retrieval, in: Comprehensive Remote Sensing. 10.1016/B978- 0-12-409548-9.10632-3.

Frankenberg, C., Ko¨hler, P., Magney, T.S., Geier, S., Lawson, P., Schwochert, M., McDuffie, J., Drewry, D.T., Pavlick, R., Kuhnert, A., 2018. The chlorophyll fluorescence imaging spectrometer (CFIS), mapping far red fluorescence from aircraft. Remote Sens. Environ. 217, 523–536. https://doi.org/10.1016/j. rse.2018.08.032.

Goulas, Y., Fournier, A., Daumard, F., Champagne, S., Ounis, A., Marloie, O., Moya, I., 2017. Gross primary production of a wheat canopy relates stronger to far red than to red solar-induced chlorophyll fluorescence. Remote Sens. 9, 97. https://doi.org/10.3390/rs9010097.

Grossmann, K., Frankenberg, C., Magney, T.S., Hurlock, S.C., Seibt, U., Stutz, J., 2018. PhotoSpec: a new instrument to measure spatially distributed red and far-red solar- induced chlorophyll fluorescence. Remote Sens. Environ. 216, 311–327. https://doi.org/10.1016/j.rse.2018.07.002.

Gu, L., Wood, J.D., Chang, C.Y.Y., Sun, Y., Riggs, J.S., 2019. Advancing terrestrial ecosystem science with a novel automated measurement system for sun-induced chlorophyll fluorescence for integration with eddy covariance flux networks. J. Geophys. Res. Biogeosci. 124, 127–146. https://doi.org/10.1029/2018JG004742.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P.E., Go´mez-Dans, J., Kuze, A., Suto, H., Grainger, R.G., 2012. Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. Remote Sens. Environ. 121, 236–251.

https://doi.org/10.1016/j.rse.2012.02.006.

Guanter, L., Rossini, M., Colombo, R., Meroni, M., Frankenberg, C., Lee, J.E., Joiner, J., 2013. Using field spectroscopy to assess the potential of statistical approaches for the retrieval of sun-induced chlorophyll fluorescence from ground and space. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2013.01.017.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J.A., Frankenberg, C., Huete, A.R., Zarco-Tejada, P., Lee, J.-E., Moran, M.S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J.M., Griffis, T.J., 2014. Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proc. Natl. Acad. Sci. 111 https://doi.org/10.1073/ pnas.1320008111.

Hueni, A., Damm, A., Kneubuehler, M., Schlapfer, D., Schaepman, M.E., 2017. Field and airborne spectroscopy cross validation - some considerations. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 1117–1135. https://doi.org/10.1109/ JSTARS.2016.2593984.

Jiang, G., Zhou, S., Cui, S., Chen, T., Wang, J., Chen, X., Liao, S., Zhou, K., 2020. Exploring the potential of HySpex hyperspectral imagery for extraction of copper content. Sensors 20, 6325. https://doi.org/10.3390/s20216325.

Jin, J., Wang, Q., 2019. Evaluation of informative bands used in different PLS regressions for estimating leaf biochemical contents from hyperspectral reflectance. Remote Sens. 11 https://doi.org/10.3390/rs11020197.

Julitta, T., Burkart, A., Rossini, M., Schickling, A., Colombo, R., Rascher, U., Cogliati, S. M., Migliavacca, M., 2017. FLoX: A system for automatic long term measurements of top of canopy sun induced chlorophyll fluorescence. Remote Sensing of Fluorescence, Photosynthesis and Vegetation Status. ESA-ESRIN, Frascati RM, Italy.

Julitta, T., Corp, L.A., Rossini, M., Burkart, A., Cogliati, S., Davies, N., Hom, M., Mac Arthur, A., Middleton, E.M., Rascher, U., Schickling, A., Colombo, R., 2016. Comparison of sun-induced chlorophyll fluorescence estimates obtained from four portable field spectroradiometers. Remote Sens. 8 https://doi.org/10.3390/ rs8020122.

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

Kr¨amer, J., Siegmann, B., Kraska, T., Muller, O., Rascher, U., 2021. The potential of spatial aggregation to extract remotely sensed sun-induced fluorescence (SIF) of small-sized experimental plots for applications in crop phenotyping. Int. J. Appl. Earth Obs. Geoinf. 104, 102565 https://doi.org/10.1016/j.jag.2021.102565.

Liu, L., Liu, X., Hu, J., 2015. Effects of spectral resolution and SNR on the vegetation solar-induced fluorescence retrieval using FLD-based methods at canopy level. Eur. J. Remote Sens. 48, 743–762. https://doi.org/10.5721/EuJRS20154841.

Liu, X., Guo, J., Hu, J., Liu, L., 2019. Atmospheric correction for tower-based solar- induced chlorophyll fluorescence observations at O2-A band. Remote Sens. 11, 355. https://doi.org/10.3390/rs11030355.

Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., Kattge, J., Richter, R., Musavi, T., Baeten, L., Barnoaiea, I., Bohn, F.J., Bouriaud, O., Bussotti, F., Coppi, A., Domisch, T., Huth, A., Jaroszewicz, B., Joswig, J., Pabon-Moreno, D.E., Papale, D., Selvi, F., Laurin, G.V., Valladares, F., Reichstein, M., Wirth, C., 2019. Inferring plant functional diversity from space: the potential of sentinel-2. Remote Sens. Environ. 233, 111368 https://doi.org/10.1016/j. rse.2019.111368.

Mac Arthur, A., Robinson, I., Rossini, M., Davis, N., MacDonald, K., 2014. A dual-field-of- view spectrometer system for reflectance and fluorescence measurements (Piccolo Doppio) and correction of etaloning. In: Proceedings of the 5th International Workshop on Remote Sensing of Vegetation Fluorescence. EuropeanSpace Agenccy, Fifth International Workshop on Remote Sensing of Vegetation Fluorescence. Paris, 22–24 April.

Magney, T.S., Bowling, D.R., Logan, B.A., Grossmann, K., Stutz, J., Blanken, P.D., Burns, S.P., Cheng, R., Garcia, M.A., Köhler, P., Lopez, S., Parazoo, N.C., Raczka, B., Schimel, D., Frankenberg, C., 2019a. Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence. Proc. Natl. Acad. Sci. USA. https://doi.org/10.1073/pnas.1900278116.

Magney, T.S., Frankenberg, C., Fisher, J.B., Sun, Y., North, G.B., Davis, T.S., Kornfeld, A., Siebke, K., 2017. Connecting active to passive fluorescence with photosynthesis: a method for evaluating remote sensing measurements of Chl fluorescence. New

Phytol. https://doi.org/10.1111/nph.14662.

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

Mardia, K.V., Kent, J.T., John, T., Bibby, J.M., John, M., Birnbaum, Z.W., Lukacs, E., 1979. Multivariate analysis. eds Probability and Mathematical Statistics. Academic Press, London, GB, p. 521.

Martini, D., Sakowska, K., Wohlfahrt, G., Pacheco-Labrador, J., van der Tol, C., Porcar- Castell, A., Magney, T.S., Carrara, A., Colombo, R., El-Madany, T.S., Gonzalez- Cascon, R., Martín, M.P., Julitta, T., Moreno, G., Rascher, U., Reichstein, M., Rossini, M., Migliavacca, M., 2022. Heatwave breaks down the linearity between sun-induced fluorescence and gross primary production. New Phytol. https://doi. org/10.1111/nph.17920.

Martini, D., Pacheco-Labrador, J., Perez-Priego, O., et al., 2019. Nitrogen and phosphorus effect on sun-induced fluorescence and gross primary productivity in mediterranean grassland. Remote Sens. 11, 2562. https://doi.org/10.3390/ rs11212562.

Meroni, M., Barducci, A., Cogliati, S., Castagnoli, F., Rossini, M., Busetto, L., Migliavacca, M., Cremonese, E., Galvagno, M., Colombo, R., Di Cella, U.M., 2011. The hyperspectral irradiometer, a new instrument for long-term and unattended field spectroscopy measurements. Rev. Sci. Instrum. 82 https://doi.org/10.1063/ 1.3574360.

Meroni, M., Busetto, L., Colombo, R., Guanter, L., Moreno, J., Verhoef, W., 2010. Performance of spectral fitting methods for vegetation fluorescence quantification. Remote Sens. Environ. 114, 363–374. https://doi.org/10.1016/j.rse.2009.09.010.

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 Sens. Environ. 113, 2037–2051. https://doi.org/10.1016/ j.rse.2009.05.003.

Mevik, B.H., Wehrens, R., 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18, 1–23. https://doi.org/10.18637/jss.v018. i02.

Middleton, E.M., Rascher, U., Corp, L.A., Huemmrich, K.F., Cook, B.D., Noormets, A., Schickling, A., Pinto, F., Alonso, L., Damm, A., Guanter, L., Colombo, R., Campbell, P.K.E., Landis, D.R., Zhang, Q., Rossini, M., Schuettemeyer, D.,

Bianchi, R., 2017. The 2013 FLEX-US airborne campaign at the parker tract loblolly pine plantation in North Carolina, USA. Remote Sens. https://doi.org/10.3390/ rs9060612.

Migliavacca, M., Perez-Priego, O., Rossini, M., El-Madany, T.S., Moreno, G., van der Tol, C., Rascher, U., Berninger, A., Bessenbacher, V., Burkart, A., Carrara, A., Fava, F., Guan, J.H., Hammer, T.W., Henkel, K., Juarez-Alcalde, E., Julitta, T., Kolle, O., Martín, M.P., Musavi, T., Pacheco-Labrador, J., P´erez-Burguen˜o, A., Wutzler, T., Zaehle, S., Reichstein, M., 2017. Plant functional traits and canopy structure control the relationship between photosynthetic CO2 uptake and far-red sun-induced fluorescence in a Mediterranean grassland under different nutrient availability. New Phytol. 214, 1078–1091. https://doi.org/10.1111/nph.14437.

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

Nichol, C.J., Drolet, G., Porcar-Castell, A., Wade, T., Sabater, N., Middleton, E.M., MacLellan, C., Levula, J., Mammarella, I., Vesala, T., Atherton, J., 2019. Diurnal and seasonal solar induced chlorophyll fluorescence and photosynthesis in a boreal scots pine canopy. Remote Sens. 11 https://doi.org/10.3390/rs11030273.

Pacheco-Labrador, J., Hueni, A., Mihai, L., Sakowska, K., Julitta, T., Kuusk, J., Sporea, D., Alonso, L., Burkart, A., Cendrero-Mateo, M.P., Aasen, H., Goulas, Y., Arthur, Mac, Mac, A., Pacheco-Labrador, J., Hueni, A., Mihai, L., Sakowska, K., Julitta, T., Kuusk, J., Sporea, D., Alonso, L., Burkart, A., Cendrero-Mateo, M.P., Aasen, H., Goulas, Y., Mac Arthur, A., 2019a. Sun-Induced chlorophyll fluorescence i: instrumental considerations for proximal spectroradiometers. Remote Sens. 11, 960.

https://doi.org/10.3390/rs11080960.

Pacheco-Labrador, J., Martïn, M.P., Riaïo, D., Hilker, T., Carrara, A., 2016. New approaches in multi-angular proximal sensing of vegetation: Accounting for spatial heterogeneity and diffuse radiation in directional reflectance distribution models. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2016.10.051.

Pacheco-Labrador, J., Perez-Priego, O., El-Madany, T.S., Julitta, T., Rossini, M., Guan, J., Moreno, G., Carvalhais, N., Martín, M.P., Gonzalez-Cascon, R., Kolle, O., Reischtein, M., van der Tol, C., Carrara, A., Martini, D., Hammer, T.W., Moossen, H., Migliavacca, M., 2019b. Multiple-constraint inversion of SCOPE. Evaluating the potential of GPP and SIF for the retrieval of plant functional traits. Remote Sens. Environ. 234 https://doi.org/10.1016/j.rse.2019.111362.

Perez-Priego, O., Guan, J., Rossini, M., Fava, F., Wutzler, T., Moreno, G., Carvalhais, N., Carrara, A., Kolle, O., Julitta, T., Schrumpf, M., Reichstein, M., Migliavacca, M., 2015. Sun-induced chlorophyll fluorescence and photochemical reflectance index improve remote-sensing gross primary production estimates under varying nutrient availability in a typical Mediterranean savanna ecosystem. Biogeosciences. https:// doi.org/10.5194/bg-12-6351-2015.

Plascyk, J.A., Gabriel, F.C., 1975. The fraunhofer line discriminator MKII-an airborne instrument for precise and standardized ecological luminescence measurement. IEEE Trans. Instrum. Meas. 24, 306–313. https://doi.org/10.1109/TIM.1975.4314448.

Porcar-Castell, A., Mac Arthur, A., Rossini, M., Eklundh, L., Pacheco-Labrador, J., Anderson, K., Balzarolo, M., Martín, M.P., Jin, H., Tomelleri, E., Cerasoli, S., Sakowska, K., Hueni, A., Julitta, T., Nichol, C.J., Vescovo, L., 2015. EUROSPEC: at the interface between remote-sensing and ecosystem CO2 flux measurements in Europe. Biogeosciences 12, 6103–6124. https://doi.org/10.5194/bg-12-6103-2015.

Porcar-Castell, A., Tyystja¨rvi, E., Atherton, J., Van Der Tol, C., Flexas, J., Pfündel, E.E., Moreno, J., Frankenberg, C., Berry, J.A., 2014. Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. J. Exp. Bot. https://doi.org/10.1093/jxb/eru191.

Core Team, R, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rascher, U., Agati, G., Alonso, L., Cecchi, G., Champagne, S., Colombo, R., Damm, A., Daumard, F., de Miguel, E., Fernandez, G., Franch, B., Franke, J., Gerbig, C., Gioli, B., Go´mez, J.A., Goulas, Y., Guanter, L., Guti´errez-de-la-C´amara, O´., Hamdi, K., Hostert, P., Jim´enez, M., Kosvancova, M., Lognoli, D., Meroni, M., Miglietta, F., Moersch, A., Moreno, J., Moya, I., Neininger, B., Okujeni, A., Ounis, A., Palombi, L., Raimondi, V., Schickling, A., Sobrino, J.A., Stellmes, M., Toci, G., Toscano, P., Udelhoven, T., van der Linden, S., Zaldei, A., 2009. CEFLES2: the remote sensing component to quantify photosynthetic efficiency from the leaf to the region by measuring sun-induced fluorescence in the oxygen absorption bands. Biogeosciences 6, 1181–1198. https://doi.org/10.5194/bg-6-1181-2009.

Rascher, U., Alonso, L., Burkart, A., Cilia, C., Cogliati, S., Colombo, R., Damm, A., Drusch, M., Guanter, L., Hanus, J., Hyv??rinen, T., Julitta, T., Jussila, J., Kataja, K., Kokkalis, P., Kraft, S., Kraska, T., Matveeva, M., Moreno, J., Muller, O., Panigada, C., Pikl, M., Pinto, F., Prey, L., Pude, R., Rossini, M., Schickling, A., Schurr, U., Sch?? ttemeyer, D., Verrelst, J., Zemek, F., 2015. Sun-induced fluorescence - a new probe of photosynthesis: first maps from the imaging spectrometer HyPlant. Glob. Chang. Biol. 21, 4673–4684. https://doi.org/10.1111/gcb.13017.

Rascher, U., Damm, A., van der Linden, S., Okujeni, A., Pieruschka, R., Schickling, A., Hostert, P., 2010. Sensing of photosynthetic activity of crops. Precision Crop Protection - the Challenge and Use of Heterogeneity. Springer, Netherlands, Dordrecht, pp. 87–99. https://doi.org/10.1007/978-90-481-9277-9_6.

Rautiainen, M., Lukeˇs, P., Homolova´, L., Hovi, A., Pisek, J., Mo˜ttus, M., 2018. Spectral properties of coniferous forests: a review of in situ and laboratory measurements. Remote Sens. https://doi.org/10.3390/rs10020207.

Rivera-Caicedo, J.P., Verrelst, J., Munoz-Mari, J., Moreno, J., Camps-Valls, G., 2014. Toward a semiautomatic machine learning retrieval of biophysical parameters. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 1249–1259. https://doi.org/10.1109/ JSTARS.2014.2298752.

Rivera, J., Verrelst, J., Go´mez-Dans, J., Mun˜oz-Marí, J., Moreno, J., Camps-Valls, G., 2015. An emulator toolbox to approximate radiative transfer models with statistical learning. Remote Sens. 7, 9347–9370. https://doi.org/10.3390/rs70709347.

Rossini, M., Meroni, M., Migliavacca, M., Manca, G., Cogliati, S., Busetto, L., Picchi, V., Cescatti, A., Seufert, G., Colombo, R., 2010. High resolution field spectroscopy measurements for estimating gross ecosystem production in a rice field. Agric. For. Meteorol. 150, 1283–1296. https://doi.org/10.1016/j.agrformet.2010.05.011.

Rossini, M., Nedbal, L., Guanter, L., Ac, A., Alonso, L., Burkart, A., Cogliati, S., Colombo, R., Damm, A., Drusch, M., Hanus, J., Janoutova, R., Julitta, T., Kokkalis, P., Moreno, J., Novotny, J., Panigada, C., Pinto, F., Schickling, A., Schttemeyer, D., Zemek, F., Rascher, U., 2015. Red and far red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. Geophys. Res. Lett. 42, 1632–1639. https://doi.org/10.1002/2014GL062943.

Sabater, N., Vicent, J., Alonso, L., Verrelst, J., Middleton, E., Porcar-Castell, A., Moreno, J., 2018. Compensation of oxygen transmittance effects for proximal sensing retrieval of canopy–leaving sun–induced chlorophyll fluorescence. Remote Sens. 10, 1551. https://doi.org/10.3390/rs10101551.

Schaepman, M.E., Dangel, S., 2000. Solid laboratory calibration of a nonimaging spectroradiometer. Appl. Opt. 39, 3754–3764. https://doi.org/10.1364/ AO.39.003754.

Schmidtlein, S., Zimmermann, P., Schüpferling, R., Weiß, C, 2007. Mapping the floristic continuum: ordination space position estimated from imaging spectroscopy. J. Veg. Sci. 18, 131–140. https://doi.org/10.1111/j.1654-1103.2007.tb02523.x.

Serbin, S.P., Townsend, P.A., 2020. Scaling functional traits from leaves to canopies. Remote Sensing of Plant Biodiversity. Springer International Publishing, pp. 43–82. https://doi.org/10.1007/978-3-030-33157-3_3.

Serbin, S.P., Wu, J., Ely, K.S., Kruger, E.L., Townsend, P.A., Meng, R., Wolfe, B.T., Chlus, A., Wang, Z., Rogers, A., 2019. From the arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance. New Phytol. 224, 1557–1568. https://doi.org/10.1111/nph.16123.

Siegmann, B., Alonso, L., Celesti, M., Cogliati, S., Colombo, R., Damm, A., Douglas, S., Guanter, L., Hanuˇs, J., Kataja, K., Kraska, T., Matveeva, M., Moreno, J., Muller, O., Pikl, M., Pinto, F., Quiro´s Vargas, J., Rademske, P., Rodriguez-Morene, F., Sabater, N., Schickling, A., Schüttemeyer, D., Zemek, F., Rascher, U., 2019. The high- performance airborne imaging spectrometer HyPlant—from raw images to top-of- canopy reflectance and fluorescence products: introduction of an automatized processing chain. Remote Sens. 11, 2760. https://doi.org/10.3390/rs11232760.

Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C., Townsend, P.A., 2015. Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecol. Appl. 25, 2180–2197. https://doi.org/10.1890/ 14-2098.1.

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Ko¨hler, P., Magney, T., 2018a. Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the orbiting carbon observatory-2: retrieval, cross-mission comparison, and global monitoring for GPP. Remote Sens. Environ. 209, 808–823. https://doi.org/10.1016/j.rse.2018.02.016.

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Ko¨hler, P., Magney, T., 2018b. Overview of solar-induced chlorophyll fluorescence (SIF) from the orbiting carbon observatory-2: retrieval, cross-mission comparison, and global monitoring for GPP. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2018.02.016.

Sun, Y., Geng, Q., Du, Y., Yang, X., Zhai, H., 2017. Induction of cyclic electron flow around photosystem I during heat stress in grape leaves. Plant Sci. https://doi.org/ 10.1016/j.plantsci.2016.12.004.

Tagliabue, G., Panigada, C., Dechant, B., Baret, F., Cogliati, S., Colombo, R., Migliavacca, M., Rademske, P., Schickling, A., Schüttemeyer, D., Verrelst, J., Rascher, U., Ryu, Y., Rossini, M., 2019. Exploring the spatial relationship between airborne-derived red and far-red sun-induced fluorescence and process-based GPP estimates in a forest ecosystem. Remote Sens. Environ. 231, 111272 https://doi.org/ 10.1016/j.rse.2019.111272.

Thuillier, G., Floyd, L., Woods, T.N., Cebula, R., Hilsenrath, E., Hers´e, M., Labs, D., 2004. Solar irradiance reference spectra for two solar active levels. Adv. Sp. Res. 34, 256–261. https://doi.org/10.1016/j.asr.2002.12.004.

van der Tol, C., Rossini, M., Cogliati, S., Verhoef, W., Colombo, R., Rascher, U., Mohammed, G., 2016. A model and measurement comparison of diurnal cycles of sun-induced chlorophyll fluorescence of crops. Remote Sens. Environ. 186 https:// doi.org/10.1016/j.rse.2016.09.021.

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., Moreno, J., Samson, R., 2015. Bidirectional sun-induced chlorophyll fluorescence emission is influenced by leaf structure and light scattering properties-bottom-up approach. Remote Sens. Environ. 158, 169–179. https://doi.org/10.1016/j.rse.2014.11.012.

Verhoef, W., van der Tol, C., Middleton, E.M., 2018. Hyperspectral radiative transfer modeling to explore the combined retrieval of biophysical parameters and canopy fluorescence from FLEX – sentinel-3 tandem mission multi-sensor data. Remote Sens. Environ. 204, 942–963. https://doi.org/10.1016/J.RSE.2017.08.006.

Verrelst, J., Mun˜oz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., Moreno, J., 2012. Machine learning regression algorithms for biophysical parameter retrieval: opportunities for sentinel-2 and -3. Remote Sens. Environ. 118, 127–139. https://doi.org/10.1016/j.rse.2011.11.002.

Verrelst, J., Rivera Caicedo, J., Mun˜oz-Marí, J., Camps-Valls, G., Moreno, J., 2017. SCOPE-based emulators for fast generation of synthetic canopy reflectance and sun- induced fluorescence spectra. Remote Sens. 9, 927. https://doi.org/10.3390/ rs9090927.

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 Sens. Environ. 166, 8–21. https://doi.org/10.1016/j.rse.2015.06.002.

Verrelst, J., Sabater, N., Rivera, J.P., Mun˜oz-Marí, J., Vicent, J., Camps-Valls, G., Moreno, J., 2016a. Emulation of leaf, canopy and atmosphere radiative transfer models for fast global sensitivity analysis. Remote Sens 8, 1–27. https://doi.org/10.3390/rs8080673.

Verrelst, J., van der Tol, C., Magnani, F., Sabater, N., Rivera, J.P., Mohammed, G., Moreno, J., 2016b. Evaluating the predictive power of sun-induced chlorophyll fluorescence to estimate net photosynthesis of vegetation canopies: a SCOPE modeling study. Remote Sens. Environ. 176, 139–151. https://doi.org/10.1016/J. RSE.2016.01.018.

Wagner, A., Hilgert, S., Kattenborn, T., Fuchs, S., 2018. Proximal VIS-NIR spectrometry to retrieve substance concentrations in surface waters using partial least squares modelling. Water Sci. Technol. Water Supply. https://doi.org/10.2166/ ws.2018.177.

Wang, N., Suomalainen, J., Bartholomeus, H., Kooistra, L., Masiliu¯nas, D., Clevers, J.G.P. W., 2021. Diurnal variation of sun-induced chlorophyll fluorescence of agricultural crops observed from a point-based spectrometer on a UAV. Int. J. Appl. Earth Obs. Geoinf. 96, 102276 https://doi.org/10.1016/j.jag.2020.102276.

Wieneke, S., Ahrends, H., Damm, A., Pinto, F., Stadler, A., Rossini, M., Rascher, U., 2016. Airborne based spectroscopy of red and far-red sun-induced chlorophyll fluorescence: Implications for improved estimates of gross primary productivity. Remote Sens. Environ. 184, 654–667. https://doi.org/10.1016/j.rse.2016.07.025.

Wieneke, S., Burkart, A., Cendrero-Mateo, M.P., Julitta, T., Rossini, M., Schickling, A., Schmidt, M., Rascher, U., 2018. Linking photosynthesis and sun-induced fluorescence at sub-daily to seasonal scales. Remote Sens. Environ. 219, 247–258. https://doi.org/10.1016/j.rse.2018.10.019.

Wiklund, S.S., 2007. Spectroscopic Data and Multivariate Analysis : Tools to Study Genetic Perturbations in Poplar trees. Department of Chemistry Research Group for Chemometrics, Umeå.

Wohlfahrt, G., Gerdel, K., Migliavacca, M., Rotenberg, E., Tatarinov, F., Müller, J., Hammerle, A., Julitta, T., Spielmann, F.M., Yakir, D., 2018. Sun-induced fluorescence and gross primary productivity during a heat wave. Sci. Rep. 8, 14169. https://doi.org/10.1038/s41598-018-32602-z.

Yang, K., Ryu, Y., Dechant, B., Berry, J.A., Hwang, Y., Jiang, C., Kang, M., Kim, J., Kimm, H., Kornfeld, A., Yang, X., 2018. Sun-induced chlorophyll fluorescence is more strongly related to absorbed light than to photosynthesis at half-hourly resolution in a rice paddy. Remote Sens. Environ. 216, 658–673. https://doi.org/ 10.1016/j.rse.2018.07.008.

Yang, P., Prikaziuk, E., Verhoef, W., Van Der Tol, C., 2021. SCOPE 2.0: A model to simulate vegetated land surface fluxes and satellite signals. Geosci. Model Dev. 14, 4697–4712. https://doi.org/10.5194/gmd-14-4697-2021.

Yang, X., Shi, H., Stovall, A., Guan, K., Miao, G., Zhang, Yongguang, Zhang, Yao, Xiao, X., Ryu, Y., Lee, J.E., 2018. FluoSpec 2—an automated field spectroscopy system to monitor canopy solar-induced fluorescence. Sensors 18. https://doi.org/10.3390/s18072063. Switzerland.

Zarco-Tejada, P.J., González-Dugo, M.V, Fereres, E., 2016. Seasonal stability of chlorophyll fluorescence quantified from airborne hyperspectral imagery as an indicator of net photosynthesis in the context of precision agriculture. Remote Sens. Environ. 179, 89–103. https://doi.org/10.1016/j.rse.2016.03.024.