RETRIEVAL OF SUN INDUCED FLUORESCENCE USING ADVANCED SPECTRAL FITTING METHODS FROM RADIATIVE TRANSFER SIMULATIONS AND HYPLANT IMAGERY


(1) Remote Sensing of Environmental Dynamics Lab, DISAT, Università degli Studi Milano-Bicocca, P.zza della Scienza 1, Milano, Italy, Email: sergio.cogliati@unimib.it
(2) Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
(3) Department of Earth Physics and Thermodynamics, University of Valencia, Dr Moliner, 50, 46100 Burjassot - Valencia, Spain
(4) University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), Hengelosestraat 99, P.O. Box 217, 7500 AE Enschede, The Netherlands

ABSTRACT
In the framework of the candidate EE-8 ESA/FLEX mission, different Solar-Induced Fluorescence (SIF) retrieval algorithms are under implementation and testing. The Spectral Fitting Method is the baseline for the fluorescence retrievals for FLEX, since it is a powerful technique to decouple canopy reflected radiance and SIF by exploiting the high-resolution observations in the red/far-red spectral regions. Two main approaches are investigated in this work: i) the SIF retrieval around the oxygen absorption bands (O\textsubscript{2}-A and O\textsubscript{2}-B); ii) the SpecFit algorithm to retrieve the full SIF emission spectrum in the 670-780 nm spectral range. The accuracy of the retrieval algorithms is assessed by employing radiative transfer simulations representing the FLEX/FLORIS radiances. The proposed algorithms are under testing on “real datasets” collected by the novel airborne HyPlant imaging spectrometer, which is well suited as airborne demonstrator of the FLEX mission. Currently, the results on radiative transfer simulations show that the relative root mean square error (RRMSE) of the fluorescence retrieval at the O\textsubscript{2}-A and O\textsubscript{2}-B bands is about 3%; the RRMSE for the full spectrum retrieval (integral of fluorescence emission) is around 7%. The preliminary map of fluorescence derived from HyPlant imagery over an agriculture area in Germany is presented. The accuracy of the fluorescence maps is assessed by comparison to ground-based measurements collected simultaneously to the airborne surveys.

1. INTRODUCTION
The Earth Explorer-8 candidate ESA/FLEX mission is aimed at the detection of the Sun-Induced Fluorescence (SIF) emitted by plants at a unique spatial resolution of 300m [1]. The SIF is a small amount of light that plant canopies continuously add to the reflected radiance. Although the spectral behaviours of these contributions are different, the decoupling is difficult because the fluorescence emission represents only a few percent of the total signal. The Spectral Fitting Methods (SFM) [2], [3] are advanced techniques able to retrieve the fluorescence and reflected radiance from high-resolution observations. Several factors affect the accuracy of SFM, like the mathematical functions used to model reflectance (R) and SIF, the fitting window and the cost function used in the numerical inversion. The optimization of SFM to retrieve accurate SIF values depends also on the technical characteristics of the spectrometer in terms of spectral sampling, resolution and noise level. Currently, several projects initiated by ESA are in progress to consolidate the FLEX mission concept, the instrument specifications, the retrieval algorithms and the application of SIF to infer vegetation health and vitality. In this framework, the Performance Analysis and Requirements Consolidation Study (PARCS) project aims at the prototyping of a SIF retrieval algorithm optimized for the space-borne FLuOResence Imaging Spectrometer (FLORIS). The sensor consists of two spectrometers [4]: i) the Narrow Band Spectrometer (NBS), which provides high resolution radiance (FWHM=0.3nm) in limited spectral ranges around the O\textsubscript{2}-A and O\textsubscript{2}-B absorption bands; ii) the Wide Band Spectrometer (WBS), characterized by a broader spectral coverage from 500 to 740nm with a maximum spectral resolution of 0.6 nm. The general SIF retrieval scheme is based on a two step approach that considers a preliminary atmospheric correction followed by the SIF and R decoupling. The object of the present work is to present the status and the results achieved on the SIF and R decoupling algorithms developed employing different implementations of spectral fitting. Two main algorithms are under development: i) the optimization
of SIF for the retrieval of SIF around the oxygen absorption bands (O$_2$-A and O$_2$-B); ii) the SpecFit algorithm to retrieve the full SIF emission spectrum in the 670-780 nm spectral range. Both the algorithms exploit the filling-in of fluorescence within the oxygen absorption bands and the solar Fraunhofer lines present in the spectral region where fluorescence emission occurs. The coupled canopy-atmosphere radiative transfer (RT) simulations resampled to the FLORIS technical specifications are used to evaluate the retrieval algorithms.

The algorithms developed for the space-borne FLORIS sensor are under implementation and testing at image level on “real datasets” collected with the novel airborne Hyperspectral Plant Imaging Spectrometer (HyPlant) sensor. Although the HyPlant sensor does not present exactly the same characteristics as FLORIS, especially in terms of noise, it is well suited as airborne demonstrator of FLEX. Preliminary maps of SIF derived from selected flight lines are presented and the accuracy of the fluorescence maps is assessed by comparing airborne retrieved values against field spectroscopy measurements collected simultaneously to the airborne survey.

2. FLUORESCENCE RETRIEVAL

In the two step retrieval scheme, the decoupling of SIF and R from the radiance observations is performed after the preliminary compensation of atmospheric effects. Once the radiance at the bottom of atmosphere is calculated, it consists of the additive contribution of SIF and the reflected radiance (Eq. (1)). The typical spectra of upwelling radiance, reflectance, apparent reflectance (with fluorescence contributes) and SIF in the spectral region involved in the fluorescence emission process are depicted in Figure 1.

\[
L_{BOA}(\lambda) = SIF(\lambda) + R(\lambda) \frac{E_g(\lambda)}{\pi}
\]

(1)

The decoupling of the two radiance terms can be achieved by using the spectral fitting approach. These methods rely on the modelling of canopy reflectance and fluorescence at the different wavelengths (\(\lambda\)). The parameters of the mathematical functions used to represent the SIF and R spectra are estimated through a least square nonlinear curve-fitting optimization technique that minimizes the sum of squared differences (Eq. 2).

\[
\min \sum \left( L_{BOA}(\lambda) - SIF(\lambda) - R(\lambda) \frac{E_g(\lambda)}{\pi} \right)^2
\]

(2)

The exploitation of a large number of spectral bands has many advantages allowing to estimate a larger number of model parameters describing the spectral variables behaviour and reducing the impact of the instrumental noise. This approach results in an accurate estimation of SIF not only in a few wavelengths within the major absorption features (i.e. O$_2$ bands; solar Fraunhofer lines), but it models the SIF spectrum within the spectral range considered. Beyond the common bases of the spectral fitting techniques, two different SIF retrieval algorithms are presented in this section.

![Figure 1: Typical spectra of: canopy radiance (grey); SIF (red line), apparent reflectance (light green); reflectance (dark green).](image)

2.1. SIF retrieval around the O$_2$ bands (SFM)

The rationale behind the exploitation of spectral bands characterized by stronger absorption resides in the higher contribution of SIF with respect to the total radiance. The use of broader windows including additional regions outside the main absorption bands permits to exploit additional nearby absorptions features (i.e. the solar Fraunhofer lines) and reducing the impact of the instrumental noise. In such regions, the modelling of SIF and R as a function of wavelength is easier. This is particularly true at the O$_2$-A band where reflectance presents a smooth spectral behaviour, while it is more difficult at the O$_2$-B band due to the sharp rising of the reflectance in that region. One of the sources of error in the SIF retrieval are the wrong assumptions that underlie the mathematical functions used to represent SIF and R spectral shapes. Functions with a large number of parameters would provide better results, but it could lead to ill-posed problems or to a higher sensitivity to noise. To select the proper configuration of the SFM algorithm, several combination of functions (Tab. 1) were tested systematically. The R models are based on polynomial functions (P), Legendre (L) polynomials and piecewise cubic splines (S). The SIF is represented with different classes of polynomial functions used for R (model IDs 1-9) and, additionally, by using of Gaussian, Lorentzian and Voigt profiles (model IDs 10-18). The models 1-9 are obtained by combining polynomial functions of different degree (i.e. from $1^{st}$ to $3^{rd}$ degree) or different number of knots for both R and SIF. The models 10-18 are built considering polynomials for R and the Gauss, Lorentz and Voigt profiles to model SIF. All the combinations of
mathematical functions were tested over two fitting windows centred at the O₂ absorption bands: i) exploits the absorption band only (i.e. 759-769 nm at O₂-A; 696-691 nm at O₂-B); ii) includes spectral regions outside the absorption bands (i.e. 740-780 nm at O₂-A; 686-696 nm at O₂-B) where the solar Fraunhofer lines are present.

Table 1: Combinations of mathematical functions tested in the SFM algorithms to retrieve R and SIF.

<table>
<thead>
<tr>
<th>ID</th>
<th>Polynomial</th>
<th>Legendre</th>
<th>Spline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P linear</td>
<td>L linear</td>
<td>K knots</td>
</tr>
<tr>
<td>2</td>
<td>P quadratic</td>
<td>L quadratic</td>
<td>K knots</td>
</tr>
<tr>
<td>4</td>
<td>L linear</td>
<td>L quadratic</td>
<td>K knots</td>
</tr>
<tr>
<td>5</td>
<td>L quadratic</td>
<td>L quadratic</td>
<td>K knots</td>
</tr>
<tr>
<td>6</td>
<td>L cubic</td>
<td>L cubic</td>
<td>K knots</td>
</tr>
<tr>
<td>7</td>
<td>L linear</td>
<td>L cubic</td>
<td>K knots</td>
</tr>
<tr>
<td>8</td>
<td>L quadratic</td>
<td>L cubic</td>
<td>K knots</td>
</tr>
<tr>
<td>9</td>
<td>L cubic</td>
<td>L cubic</td>
<td>K knots</td>
</tr>
<tr>
<td>10</td>
<td>L linear</td>
<td>G gaussian</td>
<td>K knots</td>
</tr>
<tr>
<td>11</td>
<td>L quadratic</td>
<td>G gaussian</td>
<td>K knots</td>
</tr>
<tr>
<td>12</td>
<td>L cubic</td>
<td>G gaussian</td>
<td>K knots</td>
</tr>
<tr>
<td>13</td>
<td>L linear</td>
<td>L Lorentzian</td>
<td>K knots</td>
</tr>
<tr>
<td>14</td>
<td>L quadratic</td>
<td>L Lorentzian</td>
<td>K knots</td>
</tr>
<tr>
<td>15</td>
<td>L cubic</td>
<td>L Lorentzian</td>
<td>K knots</td>
</tr>
<tr>
<td>16</td>
<td>L linear</td>
<td>V Voigt</td>
<td>K knots</td>
</tr>
<tr>
<td>17</td>
<td>L quadratic</td>
<td>V Voigt</td>
<td>K knots</td>
</tr>
<tr>
<td>18</td>
<td>L cubic</td>
<td>V Voigt</td>
<td>K knots</td>
</tr>
</tbody>
</table>

\[ V_{FarRed}(\lambda) = f \frac{\mu}{(\lambda - \lambda_0)^2} + (1 - f) \mu \exp \left( \frac{-(\lambda - \lambda_0)^2}{2\sigma_{asym}(\lambda)^2} \right) \] (5)

\[ \sigma_{asym}(\lambda) = \frac{2\sigma}{(1 + \exp(a(\lambda - \lambda_0)))} \] (6)

3. DATASETS

3.1. RT simulations

RT simulations are calculated coupling the state-of-art canopy fluorescence model SCOPE 1.4 [6] and the atmospheric code MODTRAN 5.2.1 [7]. The RT simulations at 1 cm⁻¹ spectral resolutions (in the 400-2400 nm range) are convolved to the FLORIS NBS and WBS instruments spectral response functions and binned according to the current instrument specifications. A realistic noise is added to the TOA radiance simulations. The RT database consists of 31 cases simulated with different soil, leaf, canopy and atmospheric parameters. The case n° 19 is the “standard case” which stands for a typical scenario in terms of vegetation and atmospheric parameters. The cases 1-18 are deviations of soil and canopy parameters (i.e. bright and dark soil, leaf chlorophyll, water and dry matter, canopy LAI and LIDF, etc.), the cases 19-31 cover deviations in atmospheric parameters (i.e. surface altitude, visibility, water vapour, aerosol type, vertical profile and solar zenith angle, etc.). The output layers stored in the database consist of TOA radiance with sensor noise; TOA radiance without noise; TOC radiance; TOC radiance with noise, the WLR (White Lambertian Reference); the canopy R (without fluorescence contribution) and the SIF. The latter two layers are referred hereafter as “true” because they represent the target variables of the retrieval algorithm. The TOC radiance with noise and the WLR are the two input layers used to emulate the two steps retrieval approach. However, it must be noted that an almost “perfect atmospheric correction” which uses the correct a-priori knowledge of atmospheric parameters was applied on the dataset. A fundamental advancement included for the first time in this dataset is the inclusion of the canopy-atmosphere directional effects.

3.2. HyPlant imagery and pre-processing

The HyPlant sensor is a high performance pushbroom imaging spectrometer for vegetation monitoring developed by the Forschungszentrum Jülich (Germany) and SPECIM Spectral Imaging Ltd (Finland). HyPlant consists of two modules: the high-resolution Fluorescence spectrometer (FLUO); the dual-channel VNIR & SWIR imager (DUAL). The SIF signal is retrieved from the FLUO module that produces data at a spectral resolution of about 0.25 nm in the region of the two oxygen absorption bands (670-780 nm). In the framework of several field campaigns funded by ESA, the HyPlant sensor operated over agricultural and forest areas in Germany, Czech Republic, France and the US.
between 2012 and 2013. The imagery collected at 600 (1800) m above ground level (a.g.l.) gaining a spatial resolution of 1 (3) m respectively. The HyPlant imagery recorded on August 23rd 2012 over the Selhausen (Germany) agricultural area are used to implement a SIF retrieval algorithm at image level. Currently, the processing chain under development consists of:

- the removal of flickering pixels present in HyPlant raw data;
- the correction of the spectral Stray Light using a fast correction matrix algorithm [8] based on laboratory stray light characterization;
- the detection of band centres and broadening in the sensor across-track direction using an improved version of SpecCal (Meroni et al. 2010);
- the compensation for atmospheric effects through MODTRAN 5.2.1 calculations at full resolution (0.1 cm\(^{-1}\)) using standard inputs (i.e. Mid Latitude Summer, k-correlated, DISORT 8 streams etc.), survey geometry parameters (i.e. SZA, SAA, ground elevation, sensor altitude etc.) and sun photometer measurements (i.e. aerosol optical density, AOD);
- calculation of SIF maps at O\(_2\)-A band using SFM.

4. ACCURACY EVALUATION

In the case of FLORIS RT simulations, the accuracy of the retrieval algorithms is measured in terms of relative root mean square error (RRMSE %) between the SIF retrieved against the “true” values simulated with SCOPE at the different wavelengths.

The accuracy of SIF maps derived from HyPlant imagery is instead assessed by comparison to ground-based SIF measurements derived from high-resolution field spectroscopy measurements [10]–[12] collected simultaneously to the airborne surveys.

5. RESULTS

5.1. FLORIS RT simulations

The results obtained with SFM and SpecFit algorithms on FLORIS RT simulations are reported in this section. A systematic analysis of the retrieval accuracy of the different SFM versions is done by selecting the most accurate functions. The piecewise cubic splines and the Voigt resulted in higher accuracy to model R and SIF respectively (SFM version 18-S in Table 1). An average RRMSE smaller of 0.1 mWm\(^{-2}\)sr\(^{-1}\)nm\(^{-1}\) is found for almost all the cases, with the exception of simulations 15-18 and 31, where significantly better performances are found at the O\(_2\)-A band (RRMSE < 0.025%). The cases 15, 16 are characterized by LAI values of 0.5 and 6.0 respectively; cases 17, 18 differ in their horizontal and vertical leaf angle distributions, while case 31 presents a solar zenith angle of 60\(^\circ\). The significant differences in performance are related to the canopy-atmosphere directional effects (BRDF anisotropy). In fact, the reflectance signatures from SCOPE show particular oscillations within the O\(_2\) bands due to the coupling of direct and diffuse canopy reflectance with the atmospheric transmittance in most of the cases. This effect is found to be higher at the O\(_2\)-A band because the oxygen absorption is stronger. The directional effects produce most of the errors inducing the overestimation of R and the consecutive underestimation of SIF at the O\(_2\)-A. An example of the good agreement between true and retrieved SIF and R spectra at the O\(_2\)-B (left) and O\(_2\)-A (right) is depicted in Fig. 2.

Figure 2: SFM results at the O\(_2\)-B and O\(_2\)-A bands (left and right respectively) for case n° 17. The upper plots show the true (blu), retrieved (red) and the apparent (green) reflectance; lower plots depict the true and retrieved fluorescence.

An example of the full SIF spectrum retrieval with the SpecFit algorithm is depicted in Fig. 3. Although the difficulties to mathematically model R and SIF over such a broad spectral range, the retrieved variables well represent the true reflectance and fluorescence shapes. As occurred for the estimation at the O\(_2\)-A band using SFM, better performances are found for simulations 15, 16, 17 and 31. This indicates that most of the accuracy of SpecFit is driven by the strong directional effects within the O\(_2\)-A band that affect also the retrieval of the full SIF spectrum.

Figure 3: Comparison between the true (blue lines) and retrieved (red lines) reflectance (left panel) and SIF full emission curve with SpecFit algorithm.

5.2. HyPlant imagery

HyPlant imagery collected at 600 and 1800 m a.g.l. on the Selhausen area (Germany) on August 23\(^{th}\) were processed. The SIF map derived from the low altitude
flight line recorded at 13:50 local time is shown in Figure 4. It covers a spatial subset of the original image representing different land cover types (i.e. sugar beet field, bare soils and urban area) and the ground-based measurements.

![Subset of the SIF map at 760 nm of Selhausen area is showed on the left. The upper right panel shows the spatial transect of SIF (red) and R (blue) at 760 nm and NDVI (green line). The lower right panel reports the comparison between ground and airborne measurements.](image)

The spatial patterns within the image are well reproduced by the fluorescence map. As expected the agricultural fields with higher green biomass show higher fluorescence values; on the contrary, bare soil and urban areas have lower Fs values.

The spatial profiles of SIF, R and the vegetation index NDVI in the image columns direction at the line corresponding to the ground-based measurements are depicted in Figure 4 (upper-right plot). From left to right, the values representing the bare soil and two distinct sugar beet fields are reported. The boundary between two green fields is evident at spatial pixel 200; the ground-based measurements at pixels 163–165 are indicated with the red bar. Although the NDVI values for the two vegetated fields have similar values, the SIF is different, indicating that the fluorescence signal is complementary to the information gained from the optical remote sensing.

A further comparison between ground-based measurements and SIF maps from HyPlant, for both the low and high altitude flights, is presented in Figure 4 (lower-right plot). The average over selected regions of interest (ROI) of about 3x3 pixels corresponding to ground measurements are shown. The absolute values of SIF are in the range of ground-based measurements for both the low- and high-altitude imagery, but the retrieved values are slightly overestimated by about 0.5 mWm$^{-2}$sr$^{-1}$nm$^{-1}$. The bias is almost of the same magnitude as the residual error found for the non-vegetated surfaces (e.g. bare soils) observed in the SIF spatial profile. This could be due to errors in either the modelling of atmospheric effects or the residual uncertainty related to the spatial stray light correction.

6. CONCLUSIONS

The development and testing of different SIF retrieval algorithms for the candidate ESA/FLEX mission are presented in this work. The retrieval algorithms exploit a new dataset of RT simulations resampled according the current technical specification of the FLORIS space-borne sensor. For the first time, the new dataset includes the full coupling between the canopy and atmospheric directional effects. Several mathematical functions were tested for modelling SIF and R using the SFM algorithm exploiting both the O$_2$ bands and the solar Fraunhofer lines. The piecewise cubic spline and the Voigt functions provided the higher accuracy to model R and SIF respectively. With such SFM implementation, the RRMSE of the SIF retrieval at the oxygen absorption bands is about of 3% at both the O$_2$ bands. Most of the retrieval error is due to the canopy-atmosphere directional effects. The possibility to estimate the full SIF spectrum from the FLORIS radiance using the novel SpecFit algorithm is shown for the first time. A slightly lower accuracy (RRMSE = 7% for the integral of SIF spectrum) is found for the SpecFit algorithm, as it is mainly driven by the more difficult modelling of the full SIF emission spectrum over the broader spectral region. However, the promising result achieved opens new perspectives to investigate further the SIF spectrum in relation to different plant species, photosynthetic rates and stress occurrences. Besides chlorophyll content, the different shapes of the fluorescence emission is related to the different contributions of PSI and PSII to the total fluorescence emission. For this, further studies will investigate the possibility to quantify the PSI and PSII spectra within the spectral fitting approach. Currently, the integration of SFM and/or SpecFit algorithms within the complete FLEX/FLORIS two step retrieval is ongoing.

Part of the algorithms (SFM at the O$_2$-A band) was tested on real imagery collected with the novel HyPlant sensor. The SIF maps retrieved over agricultural area show a consistent spatial pattern and the absolute values are close, albeit slightly higher (maximum difference of 1.0 mWm$^{-2}$sr$^{-1}$nm$^{-1}$) compared to ground validation measurements. Currently, the causes of the remaining biases are still unknown and further investigations concerning the improvements of the atmospheric modelling and stray light correction are needed. Further activities will be focused on the retrieval of SIF at the O$_2$-B band and the retrieval of the full emission spectrum with the SpecFit algorithm from HyPlant data.
ACKNOWLEDGMENTS

This research was supported by the ESA Flex/S3 Tandem Mission Performance Analysis and Requirements Consolidation Study (PARCS)( ESA ESTEC RFQ 3-13397/11/NL/CB); ESA HYFLEX 2012 campaign (ESTEC RFQ-3-13566/12/NL/LF) and ESA Sen2EXP 2013 campaign.

REFERENCES


