A Spectral Fitting Algorithm to Retrieve the Fluorescence Spectrum from Canopy Radiance

Sergio Cogliati 1,*, Marco Celesti 1, Ilaria Cesana 1, Franco Miglietta 2, Lorenzo Genesio 2, Tommaso Julitta 3, Dirk Schuettemeyer 4, Matthias Drusch 4, Uwe Rascher 5, Pedro Jurado 4 and Roberto Colombo 1

1 Remote Sensing of Environmental Dynamics Lab., DISAT, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milan, Italy
2 Institute of Biometeorology, National Research Council (CNR-IBIMET), Via Caproni 8, 50145 Florence, Italy
3 JB Hyperspectral Devices UG, 40225 Düsseldorf, Germany
4 ESA-ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
5 Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, Leo-Brandt-Str., 52425 Jülich, Germany

* Correspondence: sergio.cogliati@unimib.it; Tel.: +39-026-448-2913

Received: 12 June 2019; Accepted: 2 August 2019; Published: 7 August 2019

Abstract: Retrieval of Sun-Induced Chlorophyll Fluorescence (F) spectrum is one of the challenging perspectives for further advancing F studies towards a better characterization of vegetation structure and functioning. In this study, a simplified Spectral Fitting retrieval algorithm suitable for retrieving the F spectrum with a limited number of parameters is proposed (two parameters for F). The novel algorithm is developed and tested on a set of radiative transfer simulations obtained by coupling SCOPE and MODTRAN5 codes, considering different chlorophyll content, leaf area index and noise levels to produce a large variability in fluorescence and reflectance spectra. The retrieval accuracy is quantified based on several metrics derived from the F spectrum (i.e., red and far-red peaks, O_2 bands and spectrally-integrated values). Further, the algorithm is employed to process experimental field spectroscopy measurements collected over different crops during a long-lasting field campaign. The reliability of the retrieval algorithm on experimental measurements is evaluated by cross-comparison with F values computed by an independent retrieval method (i.e., SFM at O_2 bands). For the first time, the evolution of the F spectrum along the entire growing season for a forage crop is analyzed and three diverse F spectra are identified at different growing stages. The results show that red F is larger for young canopy; while red and far-red F have similar intensity in an intermediate stage; finally, far-red F is significantly larger for the rest of the season.

Keywords: sun-induced chlorophyll fluorescence; spectral fitting method; SFM; fluorescence spectrum; field spectroscopy; FLEX

1. Introduction

Sun-Induced Chlorophyll Fluorescence (F) is becoming a relevant tool for observing vegetation with remote sensing techniques. The interest around F is motivated by its inherent link to plant functioning that makes it a useful tool in understanding and modeling atmospheric-biosphere interactions at different spatiotemporal scales [1–3]. The F emission of plants (650–800 nm) is characterized by a complex spectrum with two distinct peaks in the red (685 nm) and far-red (740 nm) wavelengths. These peaks originate from two different photosystems and they are subjected to extremely diverse absorption and scattering processes along their path through the leaf tissues and the canopy. Typically, red F is strongly reabsorbed by chlorophyll pigments while far-red F is mainly scattered before escaping...
the canopy and being remotely sensed. Numerous studies have shown the capability to observe terrestrial \( F \) from space-borne instruments, but most of the efforts were focused on observing and analyzing the far-red fluorescence only [4–9]. Joiner et al., (2016) [10] and Wolanin et al., (2015) [11] were the only attempts to retrieve red \( F \) from GOSAT and SCIAMACHY, respectively. Conversely, studies based on measurements collected by field spectrometers [12–16] and airborne high-resolution imaging spectroscopy [17–22] more frequently considered both the red and far-red \( F \).

Currently, only a few studies have analyzed the potential value to exploit the entire \( F \) spectrum compared to red and far-red \( F \) peaks. Verrelst et al., (2015) [23] performed a global sensitivity analysis of the Soil-Canopy Observation of Photosynthesis and Energy (SCOPE) [24,25] radiative transfer model (RTM) and suggested that information related to photosynthetic activity is found more often in the red than in the far-red fluorescence, and even more often in the full \( F \) emission than in individual bands. However, similar considerations were not yet fully tested on experimental observations. The advent of the European Space Agency (ESA) FLuorescence EXplorer (FLEX) [26] mission, dedicated to observing global photosynthesis through the analysis of the \( F \) emission spectrum, will provide relevant measurements to further investigate such theoretical findings.

However, the retrieval of \( F \) is not trivial since the radiance flux emitted is added to the canopy reflected radiance, while the latter represents a large part of the total signal emerging from the canopy. In general, the \( F \) retrieval approaches exploit the contrast between wavelengths where radiance is dominated by canopy reflectance and narrow absorption features in the incoming irradiance where \( F \) contribution is larger. In fact, the \( F \) retrieval is enabled by observing high-resolution spectra around either Telluric or Solar absorption features. The oxygen absorption bands A (760 nm) and B (687 nm) are typically exploited to detect far-red and red \( F \), respectively [13,27–32]; whereas several Fraunhofer Lines (FLs) in the far-red (740–755 nm) [5,33] and near-infrared (771 nm) wavelengths [6,34] are examined to retrieve far-red \( F \). The approaches based on FLs are often employed to retrieve \( F \) from currently available atmospheric-chemistry satellites, in fact they rely on simpler atmospheric compensation methods that can also be performed on such instruments that are not specifically designed to observe \( F \). Conversely, the \( O_2 \) bands are considered to be more suitable to detect \( F \) because these spectral features are broader and deeper (darker), but a complex atmospheric correction is required to retrieve \( F \). This is challenging for satellite and airborne remote sensing data, but it does not represent an issue for field spectroscopy measurements.

Several approaches were developed to retrieve fluorescence in selected absorption bands from top-of-canopy (TOC) observations [35,36]. The FLD method [37], and successive versions like 3FLD [38] and iFLD [39], disentangle reflected radiance and \( F \), making use of few discrete spectral bands inside and outside the absorption feature. In contrast, the Spectral Fitting Methods (SFMs) use all the contiguous wavelengths to quantify fluorescence and reflectance over a defined spectral window centered at the absorption bands. Recently, a novel set of algorithms designed to estimate the entire \( F \) spectrum were proposed by different authors. The SpecFit algorithm [31] represents an extension of the SFM concept over the entire spectral range where the fluorescence emission occurs. Similarly, the Fluorescence Spectrum Reconstruction (FSR) [40], the Full-spectrum Spectral Fitting (F-SFM) [41], and the advanced FSR (aFSR) [42] methods use combinations of basis spectra to model the \( F \) spectrum. These basis spectra are derived from principal component analysis or singular vector decomposition techniques on training dataset obtained by canopy RTMs. Alternatively, the numerical inversion of canopy RTMs is an interesting and emerging approach for quantifying \( F \) together with a set of canopy biophysical parameters (e.g., chlorophyll content, leaf area index, etc.). Verhoef et al. (2018) [32] and Celesti et al. (2018) [13] proposed model inversion schemes in which the SCOPE model is used to reproduce canopy reflectance and retrieve the \( F \) spectrum. Although various algorithms were proposed to retrieve fluorescence based on different approaches, it is still a growing area of research toward more effective and robust solutions, as recently highlighted in the review from Mohammed et al., (2019) [43].

In this framework, we introduce and test a novel full \( F \) spectrum retrieval algorithm and we evaluate its retrieval performance using simulated and field data in the framework of ongoing studies.
to support the FLEX mission development. The proposed approach represents an evolution of the algorithm presented in reference [31] in which several improvements are introduced with the main aim to develop a simpler and efficient version of the retrieval algorithm. The remaining part of the manuscript is organized as follows: Section 2 delignates the theoretical rationale and actual implementation of the algorithm; the simulated and experimental data analyzed are described in Section 3; the results observed based on defined evaluation metrics (Section 4) are reported in Section 5; finally, areas of discussion and conclusions are provided in Sections 6 and 7.

2. F Spectrum Retrieval Approach

The novel F spectrum retrieval algorithm relies on the Spectral Fitting technique. The rationale consists of employing general-purpose mathematical functions to represent fluorescence and reflectance spectra in the 670–780 nm spectral window. The F and R spectra are disentangled from total radiance emerging from the top of the canopy ($L^\uparrow$) on the basis of their spectrally variable additive contributions according to Equation (1):

$$L^\uparrow(\Omega_0, \Omega, \lambda) = R(\Omega_0, \Omega, \lambda) L^\downarrow(\Omega_0, \Omega, \lambda) + F(\Omega_0, \Omega, \lambda)$$

where $R$ represents the canopy reflectance, $L^\downarrow$ the at-surface downwelling radiance and $F$ the sun-induced chlorophyll fluorescence spectrum. All these spectral variables depend on the distribution of the incident illumination field ($\Omega_0$) and the sensor view direction ($\Omega$). Additionally, the Equation (1) considers isotropic $R$ and $F$ and neglects a possible adjacency contribution from surrounding, which is a reasonable assumption for TOC observations. The main idea underlying the proposed algorithm consists in a simplified modeling of the $F$ spectrum (two parameters), leveraging on the spectral information included in the canopy $R$ as a proxy of the $F$ reabsorption and scattering. Figure 1 summarizes the main steps involved in the retrieval algorithm that basically consists in: (i) modeling the reflectance; (ii) modeling the fluorescence; (iii) estimating actual $R$ and $F$ spectra by means of an iterative optimization technique.

![Flow chart of the fluorescence spectrum retrieval algorithm based on the Spectral Fitting technique for top-of-canopy observations (field spectroscopy).](image)

The $F$ retrieval is enabled at the top-of-canopy by measured variables such as the downwelling and upwelling radiance spectra. The surface reflectance spectrum is modeled in the retrieval
algorithm by a piece-wise cubic spline with 20 knots \( \mathcal{R}(\mathbf{y}_j; \lambda) \); where \( \mathbf{y}_j \) is the set of unknown spline coefficients associated to each knot. The locations (wavelength) of knots is settled statistically by the routine (according Schoenberg-Whitney conditions [44]) and left constant in the successive numerical optimization where the \( \mathbf{y}_j \) parameters are estimated. The usage of a spline for modeling canopy reflectance is motivated by the fact that canopy reflectance is in general spectrally smooth. The spline function is first computed on surface apparent reflectance (Figure 2), to obtain a first guess for the numerical optimization to estimate the true canopy reflectance \( R \) (i.e., without \( F \) emission). For reference, apparent reflectance (\( R^* \)) refers to the ratio between total canopy radiance (i.e., reflected and \( F \) radiances) and downwelling radiance. The characteristic peaks in \( R^* \), which were caused by \( F \) filling-in at the \( \mathrm{O}_2 \) bands, are excluded in this first stage, since the spline is in the end targeted at modeling reflectance (free of \( F \)). Thus, a smooth function (\( R^* \) fit) is initially obtained, considering \( R^* \) but avoiding \( \mathrm{O}_2 \) bands and, from this, the reflectance is finally estimated iteratively in the optimization algorithm (consistently with \( F \)).

The \( F \) spectrum is instead modeled as a linear combination of two Lorentzian functions (Equation (2)) to mimic the red (\( F_{\text{RED}} \)) and far-red (\( F_{\text{FAR-RED}} \)) fluorescence emission (Equations (3)–(6) respectively). Each Lorentzian peak is characterized by one single state parameter only (\( x_1, x_2 \)) which modulates the overall peak intensity. The other two parameters required to describe the spectral distribution of each peak (\( u_{\text{RED}} \) and \( u_{\text{FAR-RED}} \)) are left constant to their characteristic value (i.e., wavelength of maximum emission and spectral width). The \( F \) spectrum obtained in a first step by combining \( F_{\text{RED}} \) and \( F_{\text{FAR-RED}} \) is further adjusted by a spectrally-variable correction factor (1-\( R \)) (Figure 3) which is intended as a proxy of canopy reabsorption (red wavelengths) and scattering (far-red wavelengths). The [1-(1-\( R \))] term could be simply reduced to \( R \), but the extended notation is maintained in Equation (2) because the element-wise product (Hadamard product) between the sum of \( F_{\text{RED}} \) and \( F_{\text{FAR-RED}} \) and the (1-\( R \)) term explicitly provides the \( F \) radiance reabsorbed by the canopy. The correction factor is mainly introduced to simplify the overall modeling of the red and far-red peaks, including the valley in between. This spectral region is particularly critical and difficult to model because: (i) there is a substantial contribution of both red and far-red peaks; and (ii) it corresponds to reflectance red-edge wavelengths characterized by a sharp transition from strong absorption (red) to scattering (far-red). Similarly, it is reasonable to assume that \( F \) reabsorption has a progressive decrease from red toward far-red wavelengths, introducing a spectrally-variable effect that has a complex behavior in this spectral region.

\[
F(\lambda) = (F_{\text{RED}}(x_1; \lambda) + F_{\text{FAR-RED}}(x_2; \lambda)) \circ (1 - (1 - R(\mathbf{y}_j; \lambda)))
\]

\[
u_{\text{RED}}(\lambda) = \frac{(\lambda - 684)}{10}
\]

Figure 2. Apparent reflectance (\( R^* \), dark green line), spline fit of apparent reflectance (\( R^* \) fit, blue line) and true reflectance (\( R \), light green). Inner boxes show details at \( \mathrm{O}_2 \) bands.
soil reflectance play a secondary role. Yang et al., (2018) [48] developed a theoretical study on the
indicates that most of the information is included in the canopy reflectance, while transmittance and
The approach is verified on few empirical measurements and a successive sensitivity analysis
framework to correct reabsorption based on canopy reflectance, transmittance and soil reflectance.
reabsorption effect at leaf scale. Recently, Romero et al., (2018) [47] proposed a more theoretical
from this work suggests that leaf reflectance can be successfully used for correcting chlorophyll
of Gitelson et al., (1998) [49] was one of the first attempts on this relevant topic; the main outcome
theoretical [48] studies aimed at analyzing canopy reabsorption and scattering. The pioneering work
impact on the overall canopy reflectance, it might introduce undesired effects. However, the 1-
reabsorption and scattering and eventually $x_1$ and $x_2$ coefficients are properly
adjusted in the iterative optimization to disentangle $F$ from the reflected radiance.
However, the correction function is not intended to quantify and correct reabsorption and scattering, as it would be desirable to quantitatively link TOC fluorescence with leaf or photosystem level $F$ emission [45,46]. Conversely, it is aimed as a simple and practical approach which allows to model $F$ spectrum with a limited number of parameters. An additional consideration must be discussed about the usage of 1-R term that, in practice, represents a proxy for canopy absorption (reabsorption) assuming a black albedo background underneath the canopy. This is motivated by the assumption that the spectral information from $R$ is mainly related to the canopy and therefore the correction of the $F$ emitted spectrum is more efficient. Conversely, when the soil effect has a larger impact on the overall canopy reflectance, it might introduce undesired effects. However, the 1-R term is used as a first guess for reabsorption and scattering and eventually $x_1$ and $x_2$ coefficients are properly adjusted in the iterative optimization to disentangle $F$ from the reflected radiance.
The correction function introduced in the algorithm is also motivated by experimental [47] and theoretical [48] studies aimed at analyzing canopy reabsorption and scattering. The pioneering work of Gitelson et al., (1998) [49] was one of the first attempts on this relevant topic; the main outcome from this work suggests that leaf reflectance can be successfully used for correcting chlorophyll reabsorption effect at leaf scale. Recently, Romero et al., (2018) [47] proposed a more theoretical framework to correct reabsorption based on canopy reflectance, transmittance and soil reflectance. The approach is verified on few empirical measurements and a successive sensitivity analysis indicates that most of the information is included in the canopy reflectance, while transmittance and soil reflectance play a secondary role. Yang et al., (2018) [48] developed a theoretical study on the SCOPE model to analytically derive the relationship between scattering of $F$ caused by canopy and $R$. The study indicates that canopy scattering of far-red SIF can be expressed as a simple function of canopy interceptance, TOC reflectance and leaf albedo. All these studies have thus motivated our approach, with the intent to simplify the $F$ spectrum retrieval approach.
The state parameters in the forward model are estimated by means of a non-linear least square technique. The retrieval algorithm iteratively computes (simulates) canopy upwelling radiance ($L^\text{mod}$) based on Equation (1) and optimizes $F$ and $R$ spectral functions coefficients ($x_i$, $y_j$), for a total of 22 parameters, until the best match with observed spectra ($L^\text{obs}$) is obtained (Figure 1). The iterative optimization searches the minimum value of the cost function evaluated as spectral sum of square
difference between forward modeled and experimentally measured radiances (Equation (7)). There are no minimum and maximum boundaries on model parameters for the current implementation, or first guesses for red and far-red $F$. Typically, the stopping criteria of the iterative optimization are satisfied within 3 iterations. In the future, these additional constraints and initial information can be introduced to speed-up calculations or to restrict the variables domain to physical solutions only.

$$\min \sum_{\lambda=1}^{n} \left( L_{\text{mod}}(x_i, y_j; \Omega_0, \Omega, \lambda) - L_{\uparrow}(\Omega_0, \Omega, \lambda) \right)^2$$

Once the optimization algorithm converges to solution, the state parameters of the forward model are determined. The outputs of the retrieval algorithm consist in the surface reflectance and the full fluorescence spectrum in the spectral window 670–780 nm. A number of $F$ spectrum derived metrics (Figure 4) are routinely computed from the full $F$ spectrum, facilitating the exploitation of the spectral information i.e., $F$ values at red and far-red peaks maximum; values at the $O_2$ bands and the spectrally-integrated value. The Matlab source code of the present algorithm is available for download through the git repository https://gitlab.com/ltda/flox-specfit.

3. Theoretical and Experimental Data
3.1. RT Simulations

A set of radiative transfer simulations is generated to test the retrieval algorithm and to quantify its accuracy. The high-resolution simulated spectra were calculated by coupling the SCOPE (version 1.61) with the MODerate resolution atmospheric TRANsmission (MODTRAN5) codes. The surface-atmosphere coupling is based on the four-stream radiative transfer theory [50,51] with the addition of the fluorescence flux. It represents an accurate and efficient way to describe the radiative transfer interactions of the Earth’s surface-atmosphere system. The simulations were designed to reproduce TOC observations in which canopy upwelling radiance (reflected radiance and fluorescence) and incoming downwelling radiance were stored as output of the simulations. The formalism introduced in references [31,32] is used hereafter, but several main elements are recalled to facilitate the understanding of the manuscript.

The SCOPE RTM simulates surface reflectance and fluorescence spectra, linking top-of-canopy radiance with a number of canopy biophysical, photosynthesis and energy balance parameters. The canopy RT equations are inherited from the widely used SAIL model (Scattering by Arbitrary Inclined Leaves), which is a 1-D turbid medium model that considers homogenous and infinitely extended target surface. The spectra computed by SCOPE, with a spectral sampling of 1 nm, are successively coupled with high-resolution MODTRAN5 spectra to compute TOC upwelling and downwelling radiance spectral fluxes in high-resolution. The canopy reflectance is modeled in SCOPE by four Bidirectional Reflectance Distribution Function (BRDF) terms: $r_{so}$ and $r_{do}$ are direct
and diffuse reflectance in the viewing direction; whereas $\overline{r}_{ad}$ and $\overline{r}_{dd}$ represent the direct and diffuse hemispherical contribution from surrounding. Similarly, $F_{so}$ and $F_{hem}$ represent fluorescence radiance from the target in the viewing direction and fluorescence emitted in the whole hemisphere from the surrounding respectively.

The atmospheric functions i.e., direct/diffuse and upward/downward transmittances, spherical albedo and atmospheric bi-directional reflectance; are derived from MODTRAN5 according to the interrogation technique described in reference [51]. The atmospheric functions are re-arranged in high-resolution according to the so-called T-14 system to avoid the fine spectral band effect described in Verhoef et al., (2018) [32]. In fact, the average of the product between two atmospheric functions over a finite spectral band is not equal to the product of the functions when the bands are largely correlated. This condition occurs at the atmospheric absorption bands when downward and upward transmittances are multiplied in the RT equations. This effect leads to an apparent violation of Beer’s law in presence of strong absorption lines that it might be confused with fluorescence filling-in effect.

Specifically, the atmospheric transfer functions employed for TOC simulations are: $t_1$ extraterrestrial solar radiance normalized by $\cos \theta$ ($\theta$ is the local solar zenith angle); $t_3$ spherical albedo; $t_4$ and $t_5$ downward direct and diffuse transmittance; $t_{12}$ product of downward direct transmittance and spherical albedo.

The TOC radiance ($L^{\uparrow}$) is calculated according to Equation (8). The first term represents the combined contribution of direct ($t_1 t_4 r_{so}$) and diffuse canopy radiance, the second term is the fluorescence in the observer viewing direction and third term is the adjacency (reflected and emitted fluxes).

$$L^{\uparrow} = t_1 \left[ t_4 r_{so} + \frac{t_5 + t_{12} \overline{r}_{ad}}{1 - \overline{r}_{dd}} r_{do} \right] + F_{so} + \left[ \frac{F_{hem} t_3 r_{do}}{1 - \overline{r}_{dd} t_3} \right]$$ (8)

Consistently, Equation (8) can be re-arranged by replacing surface reflectance terms to unity and thus simulating downwelling radiance at-surface level (Equation (9)).

$$L^{\downarrow} = t_1 \left[ t_4 + \frac{t_5 + t_{12} \overline{r}_{ad}}{1 - \overline{r}_{dd} t_3} \right] + \left[ \frac{F_{hem} t_3}{1 - \overline{r}_{dd} t_3} \right]$$ (9)

The total TOC reflectance ($R$), which is a weighted contribution of the four BRDF reflectance terms, is computed in SCOPE by rationing upwelling radiance in Equation (8) (excluding fluorescence terms) and downwelling radiance spectra (Equation (9)). In our simulations, the BRDF reflectance terms are replaced by $R$ in Equation (8) with the aim to simulate an isotropic canopy. Later on, the $R$ term is used for comparison with the reflectance spectrum estimated by the retrieval algorithm.

A range of different canopy reflectance and fluorescence spectra were simulated by varying two of the most relevant parameters affecting the overall spectrum magnitude and shape (Figure 5). The leaf area index (LAI) and leaf chlorophyll content (Cab) are varied covering a large range of values for a total of 49 different simulated canopies. Notably, LAI values 1-2-3-4-5-6-7 and the Cab values of 20-30-40-50-60-70-80 ($\mu g$ cm$^{-2}$) are considered, while all the possible combinations between the two variables levels are simulated. The other model parameters are left to their default values as suggested in the input_data_default.xlsx file available in the SCOPE model software package. This permits us to obtain a wide range of different $F$ spectra, characterized by low to high $F$ values and different spectral behaviors. Typical atmospheric conditions for clear-sky days during summer at mid-latitudes are considered to run MODTRAN5, specifically the following values are used: aerosol optical thickness of 0.1, a mid-latitude summer atmospheric model, rural aerosol model, a water content of 2.0 (g/cm$^2$), and a spectral resolution of 0.1 cm$^{-1}$. The line-of-sight parameters are settled with a solar zenith angle of 30° and a nadir viewing angle, since most of the ground-based measurements, considered later on in this study, are collected with this nadiral configuration. A single atmospheric condition is simulated since atmospheric influence is mainly removed by the direct measurement of downwelling radiance when considering field spectroscopy measurements.
Afterward, the simulated spectra are convolved with the instruments spectral response functions (ISRFs) which are assumed gaussian with a full width at half maximum (FWHM) of 0.3 nm. This set-up was selected because it closely represents the spectral configuration of the FLORIS [52] instrument aboard of FLEX, as well the FLOX field spectrometer (Section 3.2). Finally, a gaussian distributed noise with variance linearly related to the radiance intensity at the different wavelengths is added according to reference [32], this noise model considers both additive and multiplicative photon noise terms. To test the robustness of the retrieval algorithm and thus the opportunity of processing data collected by instrument characterized with diverse signal-to-noise (SNR), the simulated dataset includes signatures with different noise levels from “noise-free” to SNR values about of 1000, 200, 50. The noise-free spectra are relevant to understand the performance of the retrieval algorithm in absence of instrument disturbance, basically the intrinsic accuracy that can be gained with the proposed $R$ and $F$ functions is quantified from these data. More realistic conditions are instead considered when SNR = 1000 is applied, in fact it corresponds to the performances expected from FLORIS and FLOX spectrometers. Similarly, SNR level of 200 is considered because it roughly represents the noise level in imagery collected with HyPlant [19] that is a widely used imaging spectrometer for airborne remote sensing of vegetation fluorescence. Lastly, SNR of 50 is also analyzed as a general example to understand whether the algorithm could be successfully used with spectral measurements recorded by spectrometers with poor SNR or measurements collected with not optimal conditions i.e., low radiance levels (low solar zenith angles) as it can occurs in the early/late daylight hours during morning/afternoon when long-lasting field spectroscopy measurements are considered. The resulting data set consists of fluorescence, reflectance, total upwelling and downwelling radiance fluxes at the top-of-canopy.

3.2. Field Spectroscopy

The FLOX (JB Hyperspectral Devices, Düsseldorf, Germany) is a field spectrometer designed for continuous and long-lasting high-resolution spectral measurements for $F$ retrieval at the top-of-canopy. The spectrometer technical specifications in terms of spectral coverage, resolution and SNR were designed on the base of the FLEX mission instrument specifications [26,52]. The FLOX is equipped with two grating spectrometers: (i) QEPro (Ocean Optics, Largo FL, USA) with high spectral resolution (FWHM=0.3 nm; SSI=0.15 nm) in the fluorescence emission range 650 nm–800 nm; (ii) FLAMES (Ocean Optics, Largo FL, USA) covering the full range of VIS-NIR (FWHM=1.7 nm; SSI=0.6 nm). The spectrometers entrance is split towards two optical fibers that lead to a cosine receptor measuring the downwelling radiance and a bare fiber measuring the canopy upwelling radiance. The spectrometers are housed in a Peltier thermally regulated box, keeping the internal temperature constant at 25 °C in order to avoid dark current drift and spectral shifts related with temperature changes. The thermoelectric cooler (TEC) of QEPro is set to 20 °C to control the back thinned CCD detector SNR (nominal SNR > 1000:1). The spectrometers integration time is optimized for each
channel (down- and up-looking channels) at the beginning of each automatic measurement cycle and two associated dark spectra are systematically recorded [53].

The spectral measurements for this study were collected in an agricultural area in central Italy (Braccagni, province of Grosseto) in 2018 during the ESA-funded FLEXSense campaign. This area is located ~20 km from the coastline in central Tuscany. It consists of a patchy agricultural landscape with a variety of different crops. High-resolution spectra were systematically collected over different targets during the campaign as reported in Table 1. The data set covers different growing phases of some of the crops investigated, from young and sparse canopies (beginning of the growing season) to well-developed and full-cover canopies (end of the season). The typical measurement set-up follows the measurement protocol described in previous publications [12,53–56]. In short, it consists of canopy measurements collected at few meters above the canopy (1–3 m) to limit potential issues with the atmospheric correction [57,58] (Figure 6), but on the other hand allowing us to sample a reasonably significative area of the canopy (about 1–2 m circular spot depending on the canopy height). The data were systematically processed, applying spectral and radiometric calibration procedures, to convert raw digital counts to at-sensor calibrated radiances. Finally, time series were routinely filtered to remove spectral measurements collected with not stable illumination conditions according to a number of data quality criteria described in reference [53].

Table 1. Field spectroscopy measurements collected on different crops. The crop species, FLOX serial number, start/end dates of measurements for different targets and latitude/longitude of the different point measured are indicated.

<table>
<thead>
<tr>
<th>Target</th>
<th>Instrument Id</th>
<th>Measurements Date</th>
<th>Location (Lat/Lon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forage</td>
<td>JB-009-ESA</td>
<td>21 February–25 May</td>
<td>42.828 N; 11.069 E</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>JB-009-ESA</td>
<td>26 May–12 July</td>
<td>42.828 N; 11.076 E</td>
</tr>
<tr>
<td>Corn</td>
<td>JB-009-ESA</td>
<td>13 July–31 August</td>
<td>42.825 N; 11.068 E</td>
</tr>
<tr>
<td>Chickpea</td>
<td>JB-013-ESA</td>
<td>10 June</td>
<td>42.818 N; 11.078 E</td>
</tr>
</tbody>
</table>

Figure 6. Field measurement set-up over the different targets investigated: (A) forage; (B) alfalfa; (C) corn; (D) chickpea.

4. Evaluation Metrics

The accuracy of the proposed retrieval algorithm is evaluated through numerical RT simulations. The retrieved $F$ and $R$ spectra, as presented in Section 2, are compared with the corresponding reference spectra in the simulate data set (SCOPE). The agreement between retrieved and reference values considering the entire simulated data set is evaluated in terms of linear regression analysis.
and statistical indices. The root mean square error (RMSE) (Equation (10)) quantifies the amount by which an estimation differs from the true value of the quantity, and the relative RMSE (RRMSE) that represents the percentage of error with respect to the actual values (Equation (11)). The statistics are specifically computed for all the full F spectrum derived metrics corresponding to red and far-red peaks maximum; values at the O_2 bands and the spectrally-integrated value.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (F - F_{SCOPE})^2}{n}} 
\]  
\[
RRMSE = \sqrt{\frac{\sum_{i=1}^{n} \left( \frac{F - F_{SCOPE}}{F_{SCOPE}} \right)^2}{n}} \times 100 
\]  

The retrieval algorithm is further exploited to process experimental measurements, but a direct quantitative evaluation of the F retrieval accuracy is not feasible in this case. This is a common issue for F measurements at TOC because independent measurements do not exist. To overcome this limitation, F spectrum estimated by the proposed algorithm is compared with values obtained from an independent retrieval method. Specifically, we used the F values calculated at the O_2 bands by the much more consolidated and widely used SFM approach. The SFM is intrinsically less prone to errors since the approach exploits only narrow spectral windows around the O_2 bands: (i) the mathematical functions used to fit F and R are simpler than those required by the full spectrum algorithm; (ii) the F contribution to the total signal is larger and therefore less prone to uncertainties.

5. Results

5.1. Accuracy Quantification

Figure 7 shows a typical example of reflectance (and apparent reflectance), fluorescence and upwelling radiance spectra retrieved by the algorithm and compared to the corresponding reference signatures from the simulated data base. The TOC upwelling radiance optimized by the inversion procedure accurately matches the reference signal (residuals in general lower than ±0.05 mW m^{-2}sr^{-1}nm^{-1}), and a similar agreement is also observed for surface reflectance and fluorescence.

![Figure 7](image-url)

Figure 7. Retrieval of full F spectrum in the 650–780 nm spectral window from RT simulations: (left) apparent reflectance (dark green), reference reflectance from SCOPE (dashed blue) and retrieved reflectance (light green); (middle) reference fluorescence from SCOPE (blue) and retrieved fluorescence (red); (right) reference and modeled upwelling radiance ((dark and light blue) respectively).

The F spectrum retrieved for all simulated cases with SNR = 1000 and the reference spectra from SCOPE are depicted in Figure 8. The subplots are organized as follows: LAI values increase from top to bottom, whereas Cab values from left to right. The canopy reflectance is also reported for reference, since the 1-R term represents the spectrally-variable correction function in the retrieval. Accurate results are observed and the F spectrum retrieved by the algorithm closely matches the reference
spectrum at all different wavelengths. The \( F \) values corresponding to red and far-red peaks maximum are accurately estimated in all cases, as well the value at the \( O_2 \)-A and \( O_2 \)-B bands. In particular, simulated cases with intermediate and high values of \( \text{Cabs} \) (40–80 \( \mu \text{g cm}^{-2} \)) and LAI (3–7) show the better results at all wavelengths (cases 17–49). Conversely, few cases are instead characterized by a slight spectral mismatch in the valley in between the two \( F \) emission peaks, corresponding to the spectral region among 690–730 nm. This issue can be observed for cases \( n^\circ \) 1–14 characterized by LAI values of 1–2. This behavior can be caused by a slight overcorrection introduced in the \( F \) spectrum by the retrieval algorithm, probably caused from a larger contribution of soil onto the canopy reflectance signature. The cases characterized by lower chlorophyll content (\( \text{Cabs} \) of 20, 30 \( \mu \text{g cm}^{-2} \)) also show an imperfect match of the red peak and a subsequent slight overestimation of the minimum of the valley.

A quantitative analysis on the accuracy of the retrieval algorithm is reported in Figure 9 for all of the 49 simulated cases. The scatterplots between \( F \) spectrum derived metrics and reference values show good linear relationships. The \( R^2 \) ranges between 0.99 for \( F_{\text{FAR-RED}}, F_{\text{760}} \) and \( F_{\text{INT}} \) to 0.94 in case of \( F_{\text{RED}} \). The slopes estimated by ordinary least square linear regression model for the different metrics are almost close to 1:1 line (range 0.87 to 1.10) and intercepts are close to zero for all the different metrics. The \( F_{\text{RED}} \) is estimated with \( \text{RMSE} = 0.027 \text{ mW m}^{-2} \text{sr}^{-1} \text{nm}^{-1} \) (\( \text{RRMSE} = 2.3\% \)), the \( F_{\text{FAR-RED}} \) with \( \text{RMSE} = 0.061 \text{ mW m}^{-2} \text{sr}^{-1} \text{nm}^{-1} \) (\( \text{RRMSE} = 2.3\% \)) and the spectrally-integrated value \( F_{\text{INT}} \) with \( \text{RMSE} = 3.308 \text{ mW m}^{-2} \text{sr}^{-1} \text{nm}^{-1} \) (\( \text{RRMSE} = 1.8\% \)). Similarly, the \( F \) values at the \( O_2 \) bands are estimated with \( \text{RMSE} = 0.023 \text{ mW m}^{-2} \text{sr}^{-1} \text{nm}^{-1} \) (\( \text{RRMSE} = 1.9\% \)) at 687 nm; and \( \text{RMSE} = 0.011 \text{ mW m}^{-2} \text{sr}^{-1} \text{nm}^{-1} \) (\( \text{RRMSE} = 0.5\% \)) at 760 nm.

The robustness of the retrieval algorithm is further evaluated on data characterized by diverse instrumental noise levels. The results observed confirm that \( F_{\text{FAR-RED}}, F_{\text{760}}, \) and \( F_{\text{INT}} \) are retrieved more accurately compared to \( F_{\text{RED}} \) and \( F_{687} \) and this behavior is retained for the diverse SNRs considered (Table 2). In particular, the retrieval accuracy for all the \( F \) spectrum derived metrics resulted almost unaffected when considering “noise-free” and SNR = 1000 scenarios. The \( F_{\text{FAR-RED}}, F_{\text{760}} \) and \( F_{\text{INT}} \) are estimated with similar accuracy also from spectra with SNR = 200 (\( \text{RRMSE} = 2.3\%, 0.5\%, 1.9\% \) respectively), while retrieval error slightly increases when considering data with SNR = 50 (\( \text{RRMSE} = 2.7\%, 1.3\%, \text{and 2.9}\% \)). In contrast, the retrieval accuracy for \( F_{\text{RED}} \) and \( F_{687} \) appear more affected by noise, in fact a first negligible decrease of accuracy is found for SNR = 200 (\( \text{RRMSE} = 2.6\%, 2.3\% \)) and a larger significative drop is observed when spectra with SNR = 50 are analyzed (\( \text{RRMSE} = 8.5\%, 8.7\% \))

**Table 2.** Statistical metrics in terms of slope, intercept, coefficient of determination (\( R^2 \)), Root Mean Square Error (\( \text{RMSE} \)), Relative Root Mean Square Error (\( \text{RRMSE} \)) of the linear regression model between reference and retrieved values considering different noise levels.

<table>
<thead>
<tr>
<th>Noise</th>
<th>Statistics</th>
<th>( F_{\text{RED}} )</th>
<th>( F_{\text{FAR-RED}} )</th>
<th>( F_{\text{INT}} )</th>
<th>( F_{687} )</th>
<th>( F_{760} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No noise</td>
<td>slope</td>
<td>0.87</td>
<td>0.94</td>
<td>1.10</td>
<td>0.89</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>intercept</td>
<td>0.15</td>
<td>0.21</td>
<td>-19.20</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.940</td>
<td>0.995</td>
<td>0.995</td>
<td>0.985</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>( \text{RMSE} )</td>
<td>0.027</td>
<td>0.061</td>
<td>3.307</td>
<td>0.023</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>( \text{RRMSE} % )</td>
<td>2.3</td>
<td>2.3</td>
<td>1.9</td>
<td>1.9</td>
<td>0.5</td>
</tr>
<tr>
<td>SNR = 1000</td>
<td>slope</td>
<td>0.87</td>
<td>0.94</td>
<td>1.10</td>
<td>0.89</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>intercept</td>
<td>0.15</td>
<td>0.21</td>
<td>-19.11</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.940</td>
<td>0.995</td>
<td>0.995</td>
<td>0.985</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>( \text{RMSE} )</td>
<td>0.027</td>
<td>0.061</td>
<td>3.307</td>
<td>0.024</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>( \text{RRMSE} % )</td>
<td>2.3</td>
<td>2.3</td>
<td>1.9</td>
<td>1.9</td>
<td>0.5</td>
</tr>
<tr>
<td>SNR = 200</td>
<td>slope</td>
<td>0.88</td>
<td>0.95</td>
<td>1.11</td>
<td>0.90</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>intercept</td>
<td>0.13</td>
<td>0.21</td>
<td>-19.80</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.930</td>
<td>0.995</td>
<td>0.996</td>
<td>0.981</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>( \text{RMSE} )</td>
<td>0.031</td>
<td>0.061</td>
<td>3.267</td>
<td>0.027</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>( \text{RRMSE} % )</td>
<td>2.6</td>
<td>2.3</td>
<td>1.9</td>
<td>2.3</td>
<td>0.5</td>
</tr>
<tr>
<td>SNR = 50</td>
<td>slope</td>
<td>0.69</td>
<td>0.95</td>
<td>1.08</td>
<td>0.71</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>intercept</td>
<td>0.28</td>
<td>0.17</td>
<td>-18.39</td>
<td>0.24</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.770</td>
<td>0.995</td>
<td>0.995</td>
<td>0.857</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>( \text{RMSE} )</td>
<td>0.105</td>
<td>0.048</td>
<td>4.762</td>
<td>0.106</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>( \text{RRMSE} % )</td>
<td>8.5</td>
<td>2.7</td>
<td>2.9</td>
<td>8.7</td>
<td>1.3</td>
</tr>
</tbody>
</table>
Figure 8. Comparison between $F$ spectrum retrieved (red) and reference (blue) for all the simulated cases.
and F respectively), while retrieval error slightly increases when considering data with SNR = 50 (RRMSE estimated with similar accuracy also from spectra with SNR = 200 (RRMSE of 2.3%, 0.5%, 1.9% − accurately compared to F RED and F687 and this behavior is retained for the diverse SNRs considered (Table 2). In particular, the retrieval accuracy for all the

\[
\text{RMSE} = 0.023 \text{ mW m}^{-2}.
\]

noise, in fact a first negligible decrease of accuracy is found for SNR = 200 (RRMSE = 2.6%, 2.3%) and

Figure 9. Scatterplots between reference and retrieved values at selected wavelengths: (A) maximum of the red peak; (B) maximum of far-red peak; (C) spectrally integrated F; (D) F at 687 nm (O2-B band); (E) F at 760 nm (O2-A band). The red line is the linear least square fit whereas the dashed blue line represents the 1:1. Data refers to SNR = 1000.

5.2. F Spectrum from Experimental Field Spectroscopy

The typical behaviors of F, R and upwelling radiance observed at noon on the extended data set of field spectroscopy measurements are reported in Figure 10. In general, the spectral variables computed by the retrieval algorithm closely resemble the spectra obtained from the RT simulations, in which the apparent reflectance shows the characteristic two peaks caused by the filling-in of FR and FFAR-RED at the O2 bands.

Figure 10. Retrieval of full F spectrum from FLOX: (left) apparent reflectance (dark green) and retrieved reflectance (light green); (middle) full F spectrum; (right) measured (dark blue) and modeled (light blue) canopy upwelling radiance and downwelling radiance (gray).

Figure 11 reports scatterplots between F and R at 687 nm and 760 nm computed by the proposed retrieval algorithm and the F values at O2 bands estimated by the traditional SFM. The dots represent hourly average measurements (and standard deviation), between morning 10 am until afternoon 5 pm local time. The data refer to clear-sky days (50 days in total) extracted from the entire time series, from February to August, for different crops investigated (see Table 1) which result in a total of 303

<table>
<thead>
<tr>
<th>Band</th>
<th>RRMS</th>
<th>RRMSE</th>
<th>(\chi^2) error</th>
</tr>
</thead>
<tbody>
<tr>
<td>O2-B</td>
<td>0.025</td>
<td>0.5%</td>
<td>0.88</td>
</tr>
<tr>
<td>O2-A</td>
<td>0.027</td>
<td>0.5%</td>
<td>0.89</td>
</tr>
<tr>
<td>O3-B</td>
<td>0.061</td>
<td>0.3%</td>
<td>1.11</td>
</tr>
<tr>
<td>O3-A</td>
<td>0.027</td>
<td>0.5%</td>
<td>0.89</td>
</tr>
</tbody>
</table>

The robustness of the retrieval algorithm is further evaluated on data characterized by diverse instrumental noise levels. The results observed confirm that FFAR-RED, F760, and FINT are retrieved more accurately compared to F RED and F687 and this behavior is retained for the diverse SNRs considered (Table 2). In particular, the retrieval accuracy for all the

\[
\text{RMSE} = 0.023 \text{ mW m}^{-2}.
\]

The typical behaviors of

\[
\text{F}_{\text{RED}} \text{, F}_{\text{FAR-RED}} \text{, F}_{\text{INT}} \text{, F}_{\text{687}} \text{, F}_{\text{760}}
\]

represent the 1:1. Data refers to SNR = 1000.

Square Error (RMSE), Relative Root Mean Square Error (RRMSE) of the linear regression model

\[
\text{R}^2 = 0.99 \text{ for FR, } \text{F760 and FINT to 0.94 in case of FR}
\]

The R2 ranges between 0.99 for FFAR-RED, F760 and FINT to 0.94 in case of FR. Similarly, the

\[
\text{RRMSE} = 1.8\% \text{ at 687 nm; and RMSE} = 0.011 \text{ mW m}^{-2}
\]

of the fitted parameters is computed using Monte Carlo assuming errors are Gaussian and centered.

\[
\text{minimizing } \chi^2 \text{ function calculated using both uncertainties on x and y axes (Deming fit), the uncertainty of the fitted parameters is computed using Monte Carlo assuming errors are Gaussian and centered.}
Specifically, F_{760} shows a coefficient of determination close to one (R^2 = 0.997), the slope of the linear regression model of 0.96 and intercept −0.018, with RMSE = 0.102 mW m^{-2}sr^{-1}nm^{-1} (RRMSE = 16.5%). Analogously, F_{687} values also show a close agreement between the two retrieval methods (R^2 = 0.934), slope and intercept (1.104 and −0.021 respectively) are close to 1:1 line and the overall RMSE = 0.099 (RRMSE = 12.5%). Finally, an even better relationship is found for reflectance R_{760} with RMSE = 0.002 (RRMSE = 0.4%) and for R_{687} RMSE = 0.001 (RRMSE = 1.8%).

![Figure 11](image.png)

**Figure 11.** Scatterplot between fluorescence (top) and reflectance (bottom) at 760 nm (left) and 687 nm (right) estimated by SFM and the novel F spectrum algorithm. The dots represent hourly average values (standard deviation) for the entire time series acquired on different crops during the entire season. The red line is the OLS linear regression model; dashed blue line the 1:1.

For the first time, the temporal evolution of the full F spectrum and R along the entire growing season for one of the crops investigated is depicted in Figure 12. The time series collected over forage canopy was selected because it ranges from almost bare soil (DOY = 52) to full canopy development (DOY = 145). The data reported refer to 28 selected days characterized by clear-sky conditions and the spectra are average values within one hour around solar noon. The analysis of data collected at noon limits canopy directional effects that typically affect diurnal cycle measurements [53], providing an easier inter-day comparison along the growing season. The F and R spectra observed are on the left side, while temporal evolution of F values at specific wavelengths and selected ancillary indices are on the right side. Specifically, F_{RED}, F_{FAR-RED}, F_{INT} values are selected to show the temporal evolution of fluorescence spectrum behavior in comparison with incoming irradiance (L↑) and the MERIS Terrestrial Chlorophyll Index (MTCI) [59].

The overall evolution of F across the growing season primarily responds to the illumination level and to the amount of green biomass in the canopy, especially to the chlorophyll content during the different growing stages [14,53,60]. The average value of downwelling radiance computed in the 400–700 nm spectral windows shows an increasing trend, related to changes in solar zenith angle across the season (from February to May). The standard deviation values are somehow larger in few days (DOY = 91, 113, and 132), probably caused by slight illumination changes due to not fully stable atmospheric status. This condition directly affects F, in fact a larger standard deviation is observed for
in the same days. Conversely, MTCI is less sensitive because illumination changes are intrinsically canceled out when reflectance and reflectance-based indices are calculated. The MTCI seasonal trend is distinguished by values almost close to zero in the early days, typical of sparse canopy. Afterward, it first begins to rise weakly after DOY = 80; followed by a rapid increase at DOY = 90 until it reaches maximum values around DOY = 106. Consistently, the reflectance signature evolves from bare soil to the typical signature of a green, fully developed and healthy canopy. The progressive increase of green biomass (and chlorophyll content) is clearly observed by the decrease of reflectance at the red wavelengths due to chlorophyll absorption. In contrast, far-red reflectance typically increases because a larger amount of incident light is backscattered due to the higher canopy cover and accumulated green-biomass. In the early stage, the F spectrum has an overall faint signal at all the wavelengths, but the spectrum is characterized by a red peak larger than far-red (DOYs = 52 – 68). This behavior is somehow expected, since $F_{\text{RED}}$ at the photosystem level typically has a larger amplitude compared to $F_{\text{FAR-RED}}$ [47]. Furthermore, in these first days, the crop has a lower chlorophyll content, which on one hand implicates a general lower fluorescence emission, but on the other hand also implicates lower reabsorption. After a few days, the F spectrum develops towards an intermediate stage characterized by $F_{\text{RED}}$ and $F_{\text{FAR-RED}}$ with comparable values (DOYs = 68 – 80). In a more developed stage, later in the growing season, the F spectrum assumes a more common spectral behavior characterized by $F_{\text{FAR-RED}}$ significantly higher than $F_{\text{RED}}$ (DOYs = 80 – 145). In this condition, the larger chlorophyll content entails an overall larger amount of F emitted at photosystem level, but in contrast it also causes larger reabsorption at red wavelengths. It results that $F_{\text{RED}}$ develops weakly across the growing season, from 0.29 in the early days to 0.89 mW m$^{-2}$sr$^{-1}$nm$^{-1}$ when the canopy reaches its full development. Analogously $F_{\text{FAR-RED}}$ is also characterized by very low values about of 0.14 mW m$^{-2}$sr$^{-1}$nm$^{-1}$ in the early days, but conversely it quickly rises until reaching values larger than 4.0 mW m$^{-2}$sr$^{-1}$nm$^{-1}$ when the canopy is completely developed. The $F_{\text{INT}}$ ranges between an average value of 14 mW m$^{-2}$sr$^{-1}$ up to more than 250 mW m$^{-2}$sr$^{-1}$nm$^{-1}$ when the canopy is mature.

![Graphs showing temporal evolution of fluorescence and reflectance along growing season of forage crop between February to May: reflectance spectrum, fluorescence spectrum, MTCI and downwelling radiance.](image)

**Figure 12.** Temporal evolution of fluorescence and reflectance along growing season of forage crop between February to May: **(top-left)** reflectance spectrum; **(bottom-left)** fluorescence spectrum; **(top-right)** MTCI and downwelling radiance; **(bottom-right)** $F_{\text{RED}}$, $F_{\text{FAR-RED}}$ and $F_{\text{INT}}$ values.
6. Discussion

6.1. Retrieval Algorithm Assumptions and Limitations

The retrieval of $F$ spectrum based on Spectral Fitting technique requires complex technique to model $F$ and $R$ spectra in the broad spectral window where the $F$ emission occurs (670–780 nm). The simplified assumptions that underlie the mathematical functions used to model $F$ and $R$ is one of the possible error sources in the $F$ retrieval [28]. In general, functions with many parameters would assures a better modeling of the spectral variables of interest. On the other hand, they could lead to undesired problems like data over-fitting (i.e., fitting model does not describe the variable of interest but random error), unstable numerical solution or longer computational time. For these reasons, an optimal balance between the number of model parameters, retrieval accuracy, and stability of the numerical solution needs to be evaluated. For example, the algorithm proposed in reference [31] employs a parametric function with a large number of parameters to fit the $F$ spectrum (9 parameters): three coefficient are employed to model each of the red and far-red peaks (i.e., maximum wavelength, peak intensity and spectral width), two coefficients to compute the Voigt profile (i.e., mix between Gaussian and Lorentzian), and a parameter to account for the asymmetric behavior of the far-red peak. Conversely, this study aimed at introducing a simplified version that is featured by a limited number of model state parameters (2 parameters). This is possible because a spectrally-variable correction function is introduced within the Spectral Fitting inversion scheme. The leaf and canopy radiative transfer involve several extremely complex processes that obviously cannot be easily represented, but the proposed approach is found to be practical to simplify the $F$ spectrum modeling.

The obtained results show that the proposed algorithm is able to fit the $F$ spectra simulated by the SCOPE model with different chlorophyll content and LAI values. The better accuracy values are found on simulated cases with medium and high chlorophyll content and LAI values. Conversely, $F$ spectra corresponding to canopies characterized by low Cab and LAI are not always retrieved accurately at all wavelengths, especially in the red and the valley between red and far-red peaks. This behavior is likely related with the larger contribution of soil onto backscattered reflectance when sparse canopies are analyzed. In fact, the $1-R$ correction term assumes a major contribution from the green canopy elements only (reabsorption) and likely a tiny overcorrection is introduced on the $F$ spectrum when soil reflectance becomes significant. This effect has a larger impact in the spectral regions where the canopy reflectance is lower such as red and red-edge. However, these limitations do not significantly affect estimations of $F$ spectrum and $F$ spectrum derived metrics. In general, $F$ values at the O$_2$ bands are retrieved better than the values corresponding to the $F$ emission peaks maximum. The higher accuracy observed at the O$_2$-A band compared to the O$_2$-B (and $F_{\text{FAR-RED}}$ compared to $F_{\text{RED}}$) is due to the simpler spectral behavior of $F$ and $R$ in this spectral region, together with the fact that the atmospheric absorption feature at 760 nm is broader and therefore more prone to detect fluorescence.

The results observed are consistent also analyzing different noise levels introduced in the RT simulations. The $F_{\text{FAR-RED}}$, $F_{760}$ and $F_{\text{INT}}$ are faintly affected by noise and retrieval accuracy only slightly reduces from “noise-free” to lower SNR levels, whereas $F_{\text{RED}}$ and $F_{687}$ resulted more sensitive. The retrieval accuracy reported in this work (TOC level) cannot be evaluated against the threshold accuracy defined for the FLEX mission. However, we might consider that the 10% error requested for FLEX is defined to enable the further exploitation of $F$ retrievals for plant stress detection and gross primary productivity modeling studies. Therefore, on the basis of this threshold, we might argue that the $F$ spectra retrieved by the presented algorithm (uncertainty largely lower than 10%) can be used for successive physiological analysis.

The accuracy values presented also cannot be directly compared with those reported for the previous version of the algorithm [31] for two reasons: (i) simulations employed in this work are specifically designed to reproduce TOC measurements as collected by field spectrometer; this is in contrast with the previous work where simulations were designed to simulate FLEX satellite in which a preliminary atmospheric correction was performed to obtain TOC spectra; (ii) simulations employed
in this work do not consider canopy directional effects because we preferred an easier scenario for testing the algorithm performance for the moment. However, the proposed algorithm is designed with general purpose mathematical functions to fit the canopy $R$ and $F$, thus this approach should be able to provide a proper fit (adjusting $R$ and $F$) considering anisotropic canopies and different illuminations and viewing geometries. Conversely, the complex surface-atmosphere RT interaction that occurs with anisotropic canopies can significantly affect the retrieval accuracy as previously introduced by Cogliati et al., 2015 [31]. In fact, anisotropic canopies are characterized by a different amount of direct and diffuse reflected radiance. Basically, the direct flux corresponds to the shortest pathlength in the sun via the ground to the sensor, whereas the diffuse fluxes are associated with longer photon paths through the atmosphere and therefore to deeper absorptions. The result of these effects is that the depth of the oxygen absorption bands also depends on the relative amount of direct and diffuse light. Therefore, the directional effects can cause apparent positive or negative filling-in effect that can be confused as fluorescence. For this reason, a more conservative approach based on isotropic assumption was selected for this first testing of the algorithm. Although these effects can in principle be compensated by the retrieval algorithm, these conditions will inevitably complicate the interpretation of $F$ and $R$ in regard to the vegetation physiology. However, further research would focus on testing highly anisotropic canopies as well by introducing uncertainties related with atmospheric correction, but this was not the aim of this work that was mainly focused at the TOC scale. The analysis presented mostly relies on clear-sky conditions since the fluorescence retrieval from field spectroscopy measurements rely on the assumption that the illumination condition does not vary between downwelling and upwelling radiance observations. Illumination changes prevent any further analysis, even on the retrieval of apparent canopy reflectance. Advanced RT technique can be targeted at modeling such unstable conditions and enabling $F$ retrieval.

Additionally, it must be noted that we generate a novel set of radiative transfer simulation based on an updated version of the SCOPE (i.e., version 1.61) because several modeling improvements were introduced in the last few years, especially in the Fluspect leaf model [61]. These developments resulted to significantly different $F$ spectra compared with previous versions (i.e., version 1.56) and consequently some of the assumptions employed in the previous retrieval algorithms are not optimal. However, the results achieved in this work show values that are comparable, even slightly better, than results reported in similar works developed with different approaches [13,28,36]. Moreover, a growing number of diverse and more sophisticated RTMs have become available in recent years such as the full 3D RT models FluorWPS [62], FluorFLIGHT [63], and DART SIF [64]. In the future, testing retrieval algorithms on synthetic simulations produced with these advanced models would provide a more realistic quantification of retrieval algorithm performances in the case of heterogenous scenes.

Finally, it should be mentioned that the proposed method uses general-purpose mathematical functions which do not rely on RTMs or specific training data sets. In contrast, similar $F$ spectrum retrieval methods rely on specific basis spectra derived with different techniques from RTMs simulations (i.e., principal components, singular vector decomposition etc.). Similarly, the recently proposed approaches based on canopy RTM inversion have the strong advantages of: (i) offering consistent retrieval of canopy biophysical parameters [32]; (ii) quantifying fluorescence quantum yield [13,46]. Although these complementary variables are extremely useful for further understanding and interpreting the $F$ signal, the routinely exploitation of these methods to process large data sets is not straightforward and computationally demanding until now. However, these promising approaches can be further developed, towards more efficient inversion schemes and optimized compiled codes.

### 6.2. Observations on Experimental Field Data

The reliability and consistency of the $F$ spectrum retrieval algorithm was also evaluated on experimental measurements by comparing $F$ values computed from the proposed algorithm with independent retrievals. A close agreement is observed for $F$ and $R$ at both O$_2$ absorption bands for the
large data set considered that includes several crops observed along the season at different growing stages (young/mature canopies).

For the first time, the temporal evolution of the $F$ spectrum along the entire growing season was investigated. It was interesting to observe three distinct $F$ spectra: (i) $F_{\text{RED}}$ larger than $F_{\text{FAR-RED}}$, (ii) $F_{\text{RED}}$ comparable to $F_{\text{FAR-RED}}$, (iii) $F_{\text{FAR-RED}}$ larger than $F_{\text{RED}}$. These results are expected from a theoretical point of view considering the typical temporal development of agricultural canopies and the absorption, scattering, emission and reabsorption processes which determine $R$ and $F$ spectra. The $F_{\text{RED}}$ is strongly reabsorbed by the chlorophyll within the leaf and the canopy, therefore the amount of $F_{\text{RED}}$ escaping the canopy is typically low, especially when the canopy is fully developed. The results show $F_{\text{RED}}$ with a lower seasonal trend because as soon as the chlorophyll content increases, the red $F$ tends to saturate due to reabsorption. However, it must be considered that $F_{\text{RED}}$ is mainly released from photosystem II and thus it is directly linked to electron transport rate which is one of the key processes driving photosynthesis. Therefore, further studies need to be developed to better understand how to fully exploit $F_{\text{RED}}$ signal and two potential approaches can be designated: (i) correcting $F$ spectrum for reabsorption; (ii) exploiting canopy RTMs in which all these radiative processes are properly modeled. Conversely, $F_{\text{FAR-RED}}$ shows a more pronounced seasonal dynamic that mainly follows the progressive increase of canopy green biomass which in turn determinates a larger amount of absorbed photosynthetic active radiation from the canopy and thus higher $F$ emission. However, $F_{\text{FAR-RED}}$ is composed by fluorescence emitted from both photosystems (photosystems II and I) and consequently its expected relationship with actual photosynthetic rate is weaker compared to $F_{\text{RED}}$ (corrected for reabsorption). All these considerations must be taken into account to understand the behavior of the spectrally-integrated fluorescence. In fact, the temporal dynamic of $F_{\text{INT}}$ results from the relative contributions of $F_{\text{RED}}$ and $F_{\text{FAR-RED}}$ peaks and how they are differently modulated across the season: red peak mostly contributes early, while far-red plays a major role for the remaining part of the growing season.

6.3. Implications for Remote Sensing

The retrieval methodology described in this work is explicitly anchored to TOC observations, avoiding to introduce complex atmospheric correction schemes needed to retrieve $F$ from remote sensing data. Nevertheless, the proposed algorithm can be adapted to process airborne and satellite data. For example, this algorithm can be used in a two-step approach as introduced by Cogliati et al., (2015) [31], in which a preliminary atmospheric correction is performed for converting top-of-atmosphere to TOC radiance spectra before applying the fluorescence retrieval. Alternatively, the approach could be employed in a top-of-atmosphere (TOA) retrieval scheme as suggested by Verhoef et al., (2018) [32]. In this approach, $F$ and $R$ functions are directly estimated from TOA spectral radiance. Lastly, it must be mentioned that the reduced number of retrieval model parameters makes this novel algorithm suitable for even more sophisticated and complete schemes in which atmospheric state parameters required for the atmospheric correction and $F$ are estimated consistently in a single-step.

7. Conclusions

This study aimed at introducing a simplified version of the full $F$ spectrum retrieval algorithm. We found that the proposed algorithm is able to reproduce accurate retrievals of $F$ spectrum when compared with simulated data. A close agreement was also observed on field spectroscopy measurements between $F$ spectra retrieved by the proposed algorithm with the values at the $O_2$ bands estimated by an independent method. The $F$ spectrum of canopies characterized by intermediate and large $\text{Cab}$ and LAI values were estimated with better accuracy, while the soil effect introduces a slight disturbance when retrieving the $F$ spectrum for sparse canopies. Further studies are suggested to improve this specific topic, while also permitting more accurate retrievals for heterogeneous scenarios. The proposed algorithm resulted robust also considering different levels of noise introduced in the RT simulations, suggesting the possibility to employ the algorithm on measurements collected by
Acknowledgments: We thank several persons that enabled our research in particular K. Biriukova for spectral conceptualization, S.C., F.M., D.S., M.D., U.R. and R.C.; Data curation, M.C., I.C., F.M., & editing, M.C., I.C. and P.J.

These interesting observations provide the basis for more detailed studies about the dynamic of the F spectrum at global scale is one of the challenges for the ESA’s upcoming FLEX mission and retrieval algorithms like the one described in this work can be used for the purpose of FLEX.


Funding: This work was supported by the European Space Agency projects “FLEX L1B to L2 Retrieval Study” (AO/1-8897/17/NL/MP), FlexSense campaign (RFP/3-15477/18/NL/NA) and AtmoFLEX (4000122454/17/NL/FF).

Acknowledgments: We thank several persons that enabled our research in particular K. Biriukova for spectral measurements collected over chickpea target; C. Panigada, G. Tagliabue and M. Anelli for their support during the campaign. We also thank you Fondazione per il Clima e la Sostenibilità (FCS) for field work and Giulio Borgia owner of Azienda Agricola Le Rogate for hosting the field experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).