Impact of Atmospheric Inversion Effects on Solar-Induced Chlorophyll Fluorescence: Exploitation of the Apparent Reflectance as a Quality Indicator

Neus Sabater 1,*, Jorge Vicent 1, Luis Alonso 1, Sergio Cogliati 2, Jochem Verrelst 1 and José Moreno 1

1 Image Processing Laboratory (IPL), Parc Científic, Universitat de València, 46980 Paterna, Spain; jorge.vicent@uv.es (J.V.); luis.alonso@uv.es (L.A.); jochem.verrelst@uv.es (J.V.); jose.moreno@uv.es (J.M.)
2 Remote Sensing of Environmental Dynamics Lab., DISAT, Università di Milano-Bicocca, P.zza della Scienza 1, 20126 Milan, Italy; sergio.cogliati@unimib.it

* Correspondence: m.neus.sabater@uv.es; Tel.: +34-96-354-3110; Fax: +34-96-354-3261

Academic Editor: Lars T. Waser
Received: 3 May 2017; Accepted: 13 June 2017; Published: 16 June 2017

Abstract: In the last decade, significant progress has been made in estimating Solar-Induced chlorophyll Fluorescence (SIF) by passive remote sensing techniques that exploit the oxygen absorption spectral regions. Although the O2–B and the deep O2–A absorption bands present a high sensitivity to detect SIF, these regions are also largely influenced by atmospheric effects. Therefore, an accurate Atmospheric Correction (AC) process is required to measure SIF from oxygen bands. In this regard, the suitability of a two-step approach, i.e., first an AC and second a Spectral Fitting technique to disentangle SIF from reflected light, has been evaluated. One of the advantages of the two-step approach resides in the derived intermediate products provided prior to SIF estimation, such as surface apparent reflectance. Results suggest that errors introduced in the AC, e.g., related to the characterization of aerosol optical properties, are propagated into systematic residual errors in the apparent reflectance. However, of interest is that these errors can be easily detected in the oxygen bands thanks to the high spectral resolution required to measure SIF. To illustrate this, the predictive power of the apparent reflectance spectra to detect and correct inaccuracies in the aerosols characterization is assessed by using a simulated database with SCOPE and MODTRAN radiative transfer models. In 75% of cases, the aerosol optical thickness, the Ångstrom coefficient and the scattering asymmetry factor are corrected with a relative error below of 0.5%, 8% and 3%, respectively. To conclude with, and in view of future SIF monitoring satellite missions such as FLEX, the analysis of the apparent reflectance can entail a valuable quality indicator to detect and correct errors in the AC prior to the SIF estimation.

Keywords: solar-induced chlorophyll fluorescence; atmospheric correction; oxygen bands; apparent reflectance; spectral fitting method; FLEX

1. Introduction

Remote sensing measurement of Solar-Induced chlorophyll Fluorescence (SIF) provides a new optical mean to track plant photosynthesis and gross primary productivity (GPP) of terrestrial ecosystems [1]. SIF consists of photons of red and near infra-red light (650–850 nm) that are emitted by chlorophyll foliar pigments in response to absorption of photosynthetically active radiation. At a satellite level, the SIF signal is around two orders of magnitude lower than the reflected radiance, which makes its detection by remote sensing instruments challenging. However, SIF emission can still be detected by exploiting the fact that SIF is a proportionally larger fraction of the total radiance.
within dark lines and bands of the atmospheric spectrum [2]. These dark features include both solar Fraunhofer and telluric absorption regions such as the \( \text{O}_2 - \text{B} \) and the \( \text{O}_2 - \text{A} \) bands. Taking advantage of this fact, multiple SIF retrieval strategies have been developed in the last decade [3]. For instance, some SIF retrieval strategies exploit the Fraunhofer solar lines, which are generally weak and narrow but not as influenced by atmospheric effects as the oxygen bands \([2,4–8]\). Other SIF retrieval strategies use the \( \text{O}_2 - \text{A} \) absorption band, which is deeper and wider but is more affected by the atmospheric effects \([9–18]\). Some other strategies exploit both, the \( \text{O}_2 - \text{A} \) and some Fraunhofer lines in the retrieval scheme \([19,20]\), and recently, the \( \text{O}_2 - \gamma \) band has also been used in combination with some solar Fraunhofer lines to anchor the \( \text{O}_2 - \text{B} \) band and to provide additional information on red SIF \([21]\).

At global scale, first SIF maps were retrieved by the exploitation of the Fraunhofer lines from atmospheric chemistry satellite missions (e.g., \([2,4–8,20]\)). However, since these missions were not designed for vegetation monitoring, SIF was detected with a coarse spatial resolution e.g., \( \sim 10 \) km for GOSAT-2 and \( \sim 2.25 \) km for OCO-2. Conversely, the future ESA’s Fluorescence Explorer (FLEX) mission \([22]\), specifically designed to monitor the health status of terrestrial vegetation, is dedicated to retrieve the full SIF signal from 650–800 nm at a spatial resolution of 300 m. FLEX will particularly exploit the information at the oxygen absorption bands by measuring these regions at a very high spectral resolution, i.e., \( 0.1 \) nm of Spectral Sampling Interval (SSI) and \( 0.3 \) nm of Spectral Resolution (SR). See Appendix A for more details about the FLuORescence Imaging Spectrometer (FLORIS) on board FLEX.

Although measurements in the \( \text{O}_2 - \text{B} \) and the deep \( \text{O}_2 - \text{A} \) absorption bands possess the maximum sensitivity to detect SIF \([22]\), these bands are also highly affected by atmospheric effects, especially by multiple scattering of aerosols and molecules \([23]\). SIF at satellite level is about \( 1–2 \) orders of magnitude (depending on the spectral region) lower than the reflected radiance, which implies that any inaccuracy in the Atmospheric Correction (AC) can rapidly make SIF estimation prone to errors \([3]\). Thus, given the importance of the AC as a prior step to the application of any SIF retrieval method, e.g., Spectral Fitting (SF) \([24]\), it is important to analyse how errors and assumptions on AC modify the inverted apparent reflectance. Consequently, this paper begins with evaluating the mathematical formulation of the FLEX data processing scheme \([22]\), from the Atmospheric Inversion (AI) and to its coupling with the at-surface level SF method. In the following, the spectral distortions on the apparent reflectance caused by an inaccurate estimation of aerosol optical properties are also evaluated and analysed using simulated data from the Soil Canopy Observation Photochemistry and Energy (SCOPE) \([25]\) and the MODeRate resolution atmospheric TRANsmission (MODTRAN) \([26]\) radiative transfer models. As a result, this work proposes the use of the apparent reflectance as quality indicator to refine the AC. Notice that while AC, atmospheric correction, includes the atmospheric characterization and the mathematical inversion from Top-Of-Atmosphere (TOA) radiance to surface reflectance, AI, atmospheric inversion, refers only to this last step.

The paper is structured as follows: Section 2 introduces and assesses the impact of the mathematical assumptions considered in the FLEX AI, as well the steps to couple SF fluorescence retrieval with inverted Top-of-Canopy (TOC) apparent reflectance. Section 3 analyses the errors in the apparent reflectance caused by an inaccurate AC process, essentially by an inaccurate aerosol characterization. Additionally, the predictive power of the apparent reflectance to detect and correct the retrieved aerosols optical properties from the AC is also evaluated. Section 4 discusses the suitability of the mathematical approximations assumed in the AI, highlights the limitations of the presented study and subsequently proposes future in-depth analyses to establish the use of the apparent reflectance as an indicator of inconsistencies in the AC. Finally, Section 5 provides the main conclusions.
2. The AI Process: From TOA Radiance to SIF through the Apparent Reflectance Inversion

2.1. Assessment of Mathematical Approximations

The radiance reaching the sensor at satellite level can be simplified as the contribution of three different fluxes (Figure 1) [27]:

1. The scattered light from the atmosphere or path radiance (thick dashed-dotted line).
2. The reflected light from the observed target which is transmitted to the sensor (solid line).
3. The light coming from multiple reflections between the surface and the atmosphere, not necessarily produced at the observed target, but finally reaching the sensor (thin dashed line).

![Figure 1](image-url)

**Figure 1.** Contribution of distinct fluxes acquired at satellite level assuming a Lambertian surface reflectance behaviour. The thick black area at the Top-of-Canopy (TOC) level represents the observed target.

Assuming a Lambertian surface reflectance $\rho$, the radiance reaching the sensor ($L_{sen}$) can consequently be formulated as the following (Equation (1)):

$$L_{sen} = L_0 + \frac{E}{\pi} \cdot T^{\uparrow} \cdot \rho + \frac{E}{\pi} \cdot S \cdot T^{\uparrow} \cdot \rho^2 + \frac{E}{\pi} \cdot S^2 \cdot T^{\uparrow} \cdot \rho^3 + ...,\quad (1)$$

where $L_0$ is the path radiance (scattered light from the atmosphere), $E$ is the total solar irradiance reaching the surface, $T^{\uparrow}$ is the total upward transmittance and $S$ is the atmospheric spherical albedo, which accounts for the reflections occurred from the atmosphere to the surface. Likewise, $E$ can be understood as the contribution of the solar irradiance diffusely transmitted to the surface ($E_{diff}$) and the solar irradiance directly transmitted to the surface ($E_{dir}$), corrected by the solar zenith angle ($\theta$):

$$E = E_{dir} \cdot \cos(\theta) + E_{diff}.\quad (2)$$

Equation (1) is mathematically known as a geometrical series expression, which converges into Equation (3):

$$L_{sen} = L_0 + \frac{E \cdot \rho \cdot T^{\uparrow}}{\pi(1 - S \cdot \rho)}.\quad (3)$$

Furthermore, in the particular case of vegetation, the SIF emission ($F$ in the formulation for brevity) can be included in Equation (3) as an additional source of energy. Since SIF is orders of magnitude lower than the TOA radiance reaching the sensor, it can be physically considered as a small perturbation of Equation (3):

$$L_{sen} = L_0 + \frac{(\frac{E}{\pi} \cdot \rho + F \cdot \rho) T^{\uparrow}}{(1 - S \cdot \rho)} = L_0 + \frac{E \cdot \rho_{app} T^{\uparrow}}{\pi(1 - S \cdot \rho)},\quad (4)$$
where $\rho_{app}$ is the so-called apparent reflectance defined as Equation (5), which accounts for the reflectance and the SIF emission normalized by the solar irradiance at surface level (Figure 2):

$$\rho_{app} = \rho + \frac{\pi \cdot F}{E}.$$  

Figure 2. Reference reflectance ($\rho$), fluorescence ($F$), and apparent reflectance ($\rho_{app}$) spectra marked as black dashed line, red solid line and black solid line, respectively. High resolution (HR) $\rho_{app}$, at 0.1 cm$^{-1}$, is also shown as a thin grey solid line. The MODerate resolution atmospheric TRANsmission (MODTRAN) inputs parameters used in the simulation and values used for the spectral convolution are detailed in Appendix B. Reflectance and Solar-Induced Chlorophyll FLuorescence (SIF) spectra are mean reference values defined in an internal FLEX technical note [28].

However, the impossibility of retrieving the reflectance ($\rho$) and the apparent reflectance ($\rho_{app}$) terms simultaneously from Equation (4) should be noted. To overcome this limitation, FLEX AI assumes that $S \cdot \rho \approx S \cdot \rho_{app}$, which leads to the following Equation (6):

$$L_{sen} = L_0 + \frac{E \cdot \rho_{app} T_\uparrow}{\pi (1 - S \cdot \rho_{app})}.$$  

Thus, assuming the atmospheric transfer functions ($L_0$, $S$, $E$ and $T_\uparrow$) as known, it becomes now possible to invert the apparent reflectance from the radiance acquired by a sensor, $L_{sen}$. This approximation, however, comes with an associated error, which is illustrated in Figure 3a,b. Here, the relative difference on TOA radiance (%) between using Equations (4) and (6) is evaluated for the reference fluorescence and reflectance spectra used in Figure 2 and various atmospheric conditions (low to high aerosol load). As it is demonstrated, the impact of this approximation on TOA radiance is one order of magnitude lower than the impact caused by an increase of 10% in the SIF emission. The selection of such a threshold for comparison, 10% of the emitted SIF, follows the criteria defined by the FLEX mission [22].
Figure 3. Relative difference on Top-Of-Atmosphere (TOA) radiance (%) between Equations (4) and (6) for a range of aerosol loads in the O$_2$–B (a) and the O$_2$–A (b) spectral regions, i.e. $100 \cdot \left( \frac{L_{\text{sen}}(\text{Eq.4})}{L_{\text{sen}}(\text{Eq.6})} \right) - 1$. Relative difference on TOA radiance (%) between using the series expansion (2nd or 3rd order) and the TOA radiance expression (Equation (6)) in the O$_2$–B (c) and the O$_2$–A (d) spectral regions, i.e. $100 \cdot \left( \frac{L_{\text{sen}}(2\text{nd},3\text{rd})}{L_{\text{sen}}(\text{Eq.6})} \right) / L_{\text{sen}}(\text{Eq.6})$. MODTRAN inputs parameters used in the simulation and values used for the spectral convolution are detailed in Appendix B.

An additional factor to consider is that the radiance acquired by an instrument, $L_{\text{sen}}$, is spatially and spectrally convolved by its instrumental response function. Consequently, the isolation of $\rho_{\text{app}}$ term from Equation (6) would imply the convolution of each atmospheric transfer function independently, which would derive a wrong apparent reflectance estimation:

$$\rho_{\text{app}} \neq \frac{L_{\text{sen}} - \langle L_0 \rangle}{\langle E \cdot T_\uparrow \rangle + \langle L_{\text{sen}} - \langle L_0 \rangle \rangle T_\uparrow}.$$

This inequality exists because the convolution of the product of two given functions is not mathematically equivalent to the product of these functions convolved, i.e., $\langle a \cdot b \rangle \neq \langle a \rangle \cdot \langle b \rangle$. This fact becomes especially relevant when working with non-smooth functions, such as the atmospheric transfer functions within the absorption regions and at high spectral resolution. In order to derive $\rho_{\text{app}}$ and thereby avoid errors associated with the convolution of each individual atmospheric transfer function, we can take the original series expansion in Equation (8). As such, the high spectral resolution atmospheric transfer functions provided by a radiative transfer model (e.g., MODTRAN [26], 6S [29]) can now be multiplied and then convolved with the corresponding Instrument Spectral Response Function (ISRF):

$$L_{\text{sen}} = \langle L_0 \rangle + \langle E \cdot T_\uparrow \rangle \cdot \rho_{\text{app}} + \langle E \cdot S \cdot T_\uparrow \rangle \cdot \rho_{\text{app}}^2 + \langle E \cdot S^2 \cdot T_\uparrow \rangle \cdot \rho_{\text{app}}^3 + \ldots$$

Hence, it is necessary to evaluate until which order the series expansion in Equation (8) must be taken so that the errors in the inverted apparent reflectance are below a required threshold. In this case, the threshold has been defined as the relative difference in $\rho_{\text{app}}$ caused by increasing the SIF emission 10% of its value. Figure 3c,d evaluates the relative difference (%) on TOA radiance between taking the series expansion expression (Equation (8)) until the 2nd and 3rd order and its convergence (Equation (6)) in the O$_2$–B and the O$_2$–A regions, respectively. As can be observed, the error derived from the 2nd order approximation is one order of magnitude lower than the impact caused by the defined threshold. Consequently, FLEX AI inverts $\rho_{\text{app}}$ by resolving a second order equation:
\begin{equation}
\rho_{\text{app(atm)}} = -\langle E \cdot T^\uparrow \rangle + \sqrt{\langle E \cdot T^\uparrow \rangle^2 - 4\langle E \cdot T^\uparrow \cdot S \rangle \pi (\langle L_0 \rangle - L_{\text{sen}})}.
\end{equation}

Note that the suffix \((\text{atm})\) in Equation (9) distinguishes between the apparent reflectance obtained from an AI process, \(\rho_{\text{app(atm)}}\), and from the definition presented in Equation (5), \(\rho_{\text{app}}\).

2.2. Coupling Spectral Fitting and Apparent Reflectance

In order to couple consistently the AI with the SF, it is necessary to understand the difference between \(\rho_{\text{app}}\) (introduced in Equation (4)) and \(\rho_{\text{app(atm)}}\) (after the AI from Equation (8)) [30]. Even in the ideal case of a perfectly characterized atmosphere, the mathematical approximations presented in Section 2.1 introduce discrepancies between these two apparent reflectance expressions (see Figure 4).

\textbf{Figure 4.} Apparent reflectance \(\rho_{\text{app}}\) (black) computed as \(\rho + \pi F/E\), and apparent reflectance \(\rho_{\text{app(atm)}}\) (red) from an atmospheric inversion process where all the atmospheric parameters are perfectly known for the \(O_2-B\) (a) and \(O_2-A\) (b) spectral regions. HR \(\rho_{\text{app}}\), at 0.1 cm\(^{-1}\), is also shown as a thin grey solid line. MODTRAN inputs parameters used in the simulation and the values used for the spectral convolution are detailed in Appendix B.

SF methods such as [17,24] are generally based on the minimization between: (1) the sensor acquired radiance, measured at surface level; and (2) the simulated radiance, built from 
\[ L_{\text{loc}} = \frac{E \rho}{\pi} + F \]
with parametric expressions for \(\rho\) and \(F\). However, in the case of airborne or satellite scale, radiance at surface level comes from an AI procedure. Thus, in order to couple consistently AI with SF, two alternatives can be considered (Figure 5):

- **Option 1:** To perform the SF method at TOA level, thereby minimizing the difference between the radiance acquired by the spaceborne instrument and the simulated radiance using Equation (4).
- **Option 2:** To perform the SF method at TOC level, thereby minimizing the difference between the atmospherically corrected radiance (or apparent reflectance) and the simulated radiance (or apparent reflectance) accounting for the AI procedure.

Note that in Figure 5 the inconsistent coupling process is also described and labelled as Wrong Option 2. In this case, the discrepancy between the simulated, \(\rho_{\text{app}}\), and the inverted, \(\rho_{\text{app(atm)}}\), shown in Figure 4 would therefore lead to an inconsistent minimization process, and hence to an erroneous SIF estimation.
Figure 5. Processing scheme detailing the two possible consistent options described to couple Atmospheric Correction (AC) and Spectral Fitting (SF) methods: (Option 1) at TOA, and (Option 2) at the TOC level. In Option 2, the suffix \( \text{sen} \) and \( \text{sim} \) on \( \rho_{\text{app(atm)}} \) indicates if radiance in the AI, \( L_{\text{sen}} \), is measured by the sensor or simulated according to Equation (4). Additionally, the inconsistent coupling process is also described at the TOC level (Wrong Option 2).

Here, we focussed on analysing and assessing the coupling process performed at TOC. While both options are valid (Option 1 and Option 2), coupling at TOC implies the retrieval of the \( \rho_{\text{app(atm)}} \) as an intermediate product, which has the additional advantage that it can be used as a quality indicator of the atmospheric correction process (see Section 3.2). Therefore, we evaluated the impact of coupling consistently (Option 2) and inconsistently (Wrong Option 2, from Figure 5) the AI process and the SF method into the retrieved SIF in the O\(_2\)-A absorption region. This evaluation was conducted in different wavelength intervals labelled as R1, R2, R3 and R4 corresponding to (759.3–765) nm, (759.3–768) nm, (759.3–770) nm and (759.3–772) nm, according to [24]. In addition, a range of atmospheric aerosol load was simulated to test its influence under both coupling processes. Surface reflectance and SIF spectra, \( \rho \) and \( F \), were modelled as quadratic functions. The first guess of the \( \rho \) polynomial coefficients was estimated by fitting the \( \rho_{\text{app}} \) with a quadratic function in the spectral region around the O\(_2\)-A but avoiding the absorption band. The first guess of the \( F \) polynomial coefficients was estimated by fitting the reference \( F \) spectrum to a quadratic function.

Figure 6 shows the relative error between the retrieved and the reference SIF spectrum evaluated at the bottom of the O\(_2\)-A absorption band for:

- An inconsistent coupling process, i.e., modelling \( \rho_{\text{app}} \) at TOC as \( \rho_{\text{app}} = \rho + \frac{\pi F}{E} \). The inconsistent coupling causes a large relative error between the retrieved and the reference SIF values for all of the minimization wavelength intervals considered (R1–R4). In addition, a slight dependency on the Aerosol Optical Thickness (AOT) can also be observed.
- A consistent coupling between AI and SF as introduced in Figure 5 (Option 2). In this case, all of the SIF retrieved values (i.e., for the different AOT) are overlapping, meaning that when the
atmospheric state is perfectly known, the retrieved SIF is completely independent from the AOT value. Consequently, the accuracy of the SF minimization process only depends on the wavelength interval considered.

Figure 6. Relative error ($\epsilon_r$) between retrieved and reference SIF values at the bottom of the O$_2$–A absorption band, 760.7 nm, covering different values of AOT at 550 nm and wavelength intervals (R1–R4) for an inconsistent coupling process (a) and a consistent coupling process (b). The following error statistics are provided: median (horizontal red line), 25th and 75th percentiles (blue boxes) and extreme min/max values (black dashed lines).

Note that in both cases atmospheric transfer functions were assumed as known. Thus, the larger error derived in the inconsistent coupling case is mainly caused by the different $\rho_{app}$ and $\rho_{app(atm)}$ definitions. Consequently, because of the different definitions used, estimated SIF strongly depends on AOT only in the inconsistent coupling process. Conversely, in both coupling strategies, residual relative errors appear according to the wavelength interval defined.

3. Apparent Reflectance Error Analysis and Its Predictive Power

3.1. Spectral Error Analysis on Apparent Reflectance

This section analyses the spectral distortions in $\rho_{app(atm)}$ caused by an inaccurate AC, fundamentally due to errors in the estimation of aerosol optical properties. We emphasized the aerosols’ optical properties since aerosol characterizations can be considered as one of the main sources of uncertainty when retrieving SIF from space [13]. However, it must be noted that distortions in $\rho_{app(atm)}$ can also be caused by other factors such as: an inaccurate instrumental spectral characterization [31], instrumental radiometric calibration errors or an inaccurate surface pressure estimation. While the characterization of other key atmospheric parameters such as the Water Vapour (WV) or the ozone (O$_3$) are also crucial in any AC algorithm [32–34], their radiometric effect can be neglected in the O$_2$ absorption regions, whereas the aerosol radiometric effect cannot. Thus, we studied the spectral distortions derived from an inaccurate estimation of:

- the AOT, which is directly related to the total aerosol load in a vertical atmospheric column, becoming a greater aerosol load the higher the AOT value.
- the Angstrom exponent ($\alpha$) [35], which accounts for the AOT spectral variation and is associated with the aerosol size, becoming a higher ($\alpha$) exponent with a smaller aerosol size.
- the asymmetry parameter ($g$) of the Henyey–Greenstein (HG) scattering phase function [36], which indicates the anisotropy of the scattering pattern, this parameter being limited to $[-1, 1]$. 
The $g$ parameter takes the value $-1$ and $1$ for a full backward and forward scattering respectively, while taking the value $0$ for an isotropic scattering pattern.

These three possible causes of distortions are evaluated over three different surfaces types (see Figure 7): (a) a bare soil, (b) a mixture of vegetation and bare soil and (c) a dense vegetation. We selected this variety of spectra to ascertain if errors derived from the AC produced a systematic error in the surface apparent reflectance, $\rho_{\text{app(atm)}}$, regardless of the nature and shape of the underlying surface.

This section thus performs an assessment of how an erroneous estimation of aerosol optical properties leads to spectral distortion on the $\rho_{\text{app(atm)}}$. In particular, we want to quantify (a) how the spectral distortion pattern depends on the evaluated aerosol optical variable, and more specifically, (b) if these distortions can be totally disentangled from the underlying surface reflectance and SIF emission. To account for that, TOA radiance spectra were simulated following Equation (4) using the three surface reflectance and SIF emission spectra from Figure 7, and assuming a reference atmospheric state defined by $\text{AOT} = 0.16$, $\alpha = 1.39$ and $g = 0.75$. Afterwards, by using Equation (9), the AI was performed by overestimating and underestimating the AOT, the $\alpha$ and the $g$ values used as a reference in the simulation. In Figures 8 and 9, the over-/under-estimation is represented as a percentage of the initial values on the $y$-axis, while the differences in $\rho_{\text{app(atm)}}$ are shown in a color scale. Main findings for the O$_2$–B region are summarized as follows:

• The over-/under-estimation of the AOT causes an under-/over-estimation of $\rho_{\text{app(atm)}}$ in SIF emitting surfaces, i.e., partially-mixed or dense vegetation, while the opposite effect is observed in non-SIF emitting surfaces, i.e., bare soil.

• Errors in the the Angstrom parameter ($\alpha$) hardly lead to errors in $\rho_{\text{app(atm)}}$. However, the spectral distortion appears to be driven by the underlying surface reflectance, the spectral distortions being more abrupt for the bare soil and the mixed vegetation and bare soil than in a full vegetation spectrum.

• The asymmetry parameter ($g$) of the HG scattering phase function is clearly the driving parameter that causes the strongest distortion in the $\rho_{\text{app(atm)}}$. Additionally, the spectral distortions mostly follow a similar spectral pattern regardless of the surface reflectance spectra.

![Figure 7](https://example.com/figure7.png)
Figure 8. Apparent reflectance spectral distortion caused by an over-/under-estimation of the AOT, \( \alpha \) and \( g \) aerosol optical properties for three different surfaces at the \( \text{O}_2\text{--B} \) absorption band at FLORIS spectral resolution. The dashed horizontal black line over the \( g \) parameter figures delimits the area where the \( \rho_{\text{app}} \) spectral distortions have been computed by extrapolating \( g \) values higher than 1.

The main findings for the \( \text{O}_2\text{--A} \) region are summarized as follows (see Figure 9):

- Spectral distortions in the \( \text{O}_2\text{--A} \) caused by each aerosol optical property follow a similar distortion pattern regardless of the surface reflectance and the fluorescence emission. Due to the deepest absorption in the \( \text{O}_2\text{--A} \) band, this region seems spectrally more sensitive to aerosol over-/under-estimation.
- Although distortions on \( \rho_{\text{app(atm)}} \) produced by changes in the AOT and \( g \), which are the driving parameters, are approximately on the same order of magnitude, the spectral distortion pattern is slightly different.
- As in the \( \text{O}_2\text{--B} \) region, the over-/under-estimation of the \( \alpha \) parameter produces a weaker spectral distortion than those produced by the AOT and the \( g \) parameters.

Note that the spectral distortion patterns observed in both \( \text{O}_2 \) regions are the result of: (1) the specific shape of the reflectance and the fluorescence emission spectra, and (2) the distortions in the atmospheric transfer functions.
Figure 9. Apparent reflectance spectral distortion caused by an over-/under-estimation of the Aerosol Optical Thickness (AOT), \(\alpha\) and \(g\) aerosol optical properties for three different surfaces at the \(O_2\)-A absorption band for the FLORIS spectral resolution. The dashed horizontal black line over the \(g\) parameter figures delimits the area, where the \(\rho_{\text{app}}\) spectral distortions have been computed by extrapolating \(g\) values higher than 1.

3.2. Spectral Distortions in the Apparent Reflectance as Quality Indicator of the Atmospheric Correction

In Section 3.1, we demonstrated that errors in the retrieved aerosol optical properties lead to specific spectral distortions in the \(\rho_{\text{app(atm)}}\). In this section, we explore the idea of using the retrieved \(\rho_{\text{app(atm)}}\) spectra to infer errors in the aerosol optical properties, i.e., \(\Delta\text{AOT, } \Delta\alpha\) and \(\Delta g\). In this way, for a given known instrumental response, \(\rho_{\text{app(atm)}}\) could be used to refine the characterization of atmospheric parameters. As a proof of concept, a database (DB) containing a set of simulated TOA radiance spectra was created using MODTRAN and SCOPE. At the surface level, five reflectance and SIF spectra were used (three of them shown in Figure 7 and two additional spectra between the mixed vegetation and bare soil cases (b) and (c)). The corresponding LAI and \(\text{Chl}\) of the five spectra selected are detailed in Appendix C. For the atmospheric transfer functions simulations, we covered all possible combinations of reference values of the atmospheric parameters \(\text{AOT} = 0.16, \alpha = [1.39, 1.54, 1.74], g = [0.75, 0.8], \text{WV} = [1.2, 2.4] \text{g/cm}^2\) and surface elevation \(h = [100, 1500]\) m, assuming the viewing and illumination geometric conditions detailed in Appendix B. The motivation of these reference values is as follows:

- to select a medium AOT value to evaluate the aerosols effect but avoiding extreme cases [37];
- to range variations around the \(\alpha\) and \(g\) values of 1.54 and 0.8, respectively, which corresponds to typical values found for continental aerosol types [38];
- to simulate the effect of the surface pressure by including two different altitudes;
- to determine if the different WV content impacts the predictive power of the \(\rho_{\text{app(atm)}}\) on the \(O_2-B\) absorption region.
For each TOA radiance spectra, the surface apparent reflectance was atmospherically inverted by over-/under-estimating the corresponding reference values of AOT, $\alpha$ and $g$ by a factor of 0%,±5%,±15%,±25% and ±50% (all possible combinations). Accordingly, the DB comprised a collection of more than $\sim 8 \times 10^4$ wrongly corrected $\rho'_\text{app}(\text{atm})$ spectra and its corresponding wrongly estimated values of AOT', $\alpha'$ and $g'$. As an extra parameter, the Normalized Difference Vegetation Index (NDVI') was also computed from $\rho'_\text{app}$ and stored in the DB. The NDVI' values were included to add an extra source of information related to the surface reflectance spectral shape out of the oxygen absorption regions. Therefore, following the multivariate linear regression formulation, $Y = MX$, the used input values ($X$) were: $\rho'_\text{app}$, AOT', $\alpha'$, $g'$, and the NDVI' index. The output values ($Y$) were the $\Delta\text{AOT}$, $\Delta\alpha$ and $\Delta g$, i.e., the over-/under-estimation of each aerosol parameter in percentage [%]. In order to accelerate the regression process, the spectral DB dimensionality was reduced by means of applying a Principal Component Analysis (PCA) technique, taking the first 50 components. Then, 30% of the DB was randomly selected to train, i.e., to obtain the multivariate linear regression coefficient matrix $M = (X^\top \cdot X)^{-1} \cdot X^\top \cdot Y$. Afterwards, the remaining 70% of the DB was used to test the retrieved outputs ($\hat{\Delta}\text{AOT}$, $\hat{\Delta}\alpha$ and $\hat{\Delta}g$) versus their reference counterparts ($\Delta\text{AOT}$, $\Delta\alpha$ and $\Delta g$). For clarification, all steps concerning the DB simulation, regression training and testing are summarized in Figure 10.

**Figure 10.** Database (DB) generation process: TOA radiance simulation (1st) and atmospheric correction process (2nd). DB multivariate linear regression training (3rd) and validation (4th) processes. Grey-shaded boxes correspond to the multivariate linear regression inputs.
Regarding the predictive power of the $\rho_{app}$ to correct the estimation of the aerosols optical properties retrieval, Figure 11 summarizes the results obtained for the O$_2$–B and the O$_2$–A absorption bands.

![Figure 11](image)

**Figure 11.** (upper row) Scatter plots for the retrieved ($\Delta AOT$, $\Delta \alpha$, and $\Delta g$) versus the reference ($\Delta AOT$, $\Delta \alpha$, and $\Delta g$) parameters. (bottom row) Histograms showing the difference between the retrieved and the reference parameters, i.e., $\Delta AOT - \Delta AOT$, $\Delta \alpha - \Delta \alpha$, and $\Delta g - \Delta g$.

For both oxygen regions, the applied method reduces in 75% of the validation cases the errors below 0.5%, 8% and 3% for aerosol parameters AOT, $\alpha$ and $g$, respectively. As expected, and in agreement with the results in Section 3.1, the predictions of the $\Delta \alpha$ coefficient are much less sensitive than predictions of the $\Delta AOT$ and the $\Delta g$ parameters. In addition, results suggest that the linear regression predicts the over-/under-estimation of the aerosol optical properties precise enough despite including different input values of WV and surface pressure. However, to quantify this effect, we computed the same statistics as showed in Figure 11, but only considering one WV and surface pressure value in the DB (not shown). As expected, obtained results were in 75% of the validation cases slightly better than in the full DB case, deriving relative errors below 0.2%, 6% and 2% for aerosol parameters AOT, $\alpha$ and $g$, respectively.

4. Discussion

This section discusses the most important messages derived from this work in the context of current remote sensing fluorescence retrieval strategies. Limitations identified and proposed improvements for futures studies are also addressed.

4.1. Suitability of the Mathematical Approximations Assumed in the FLEX AI

Over the last few years, various SIF retrieval strategies have been developed based on the exploitation of different spectral regions, the solar Fraunhofer lines [2,4–8] or the telluric oxygen absorption bands [9–18]. Alternatively, remote sensing SIF retrieval strategies can also be categorized according to those which retrieve SIF directly from TOA level, e.g., [2,4,5], or those which retrieve SIF from TOC after the AC of the signal, e.g., [9,10,17,18]. The latter are the so-called two-step approximations and involve: (1) performing the AC, and (2) coupling an SF method to estimate SIF at surface level. Two-step approaches have been used in the past to estimate SIF from airborne data [39,40], and also a two-step strategy has been proposed in the context of the forthcoming FLEX/Sentinel-3 tandem mission [17,22]. Although there are many advantages about the use of this approximation,
e.g., the derivation of the $\rho_{\text{app(\text{atm})}}$ as an intermediate product to be analysed (discussed in Section 4.2), it requires an accurate AI procedure.

To start with, when it comes to inverting non-smooth high spectral resolution functions, such as the atmospheric transfer functions with their characteristic absorption regions, an important aspect implies dealing with the convolution to the sensor instrumental spectral response function. On the one hand, all the assumptions considered for the AI procedure of the FLEX mission were evaluated. Although in general terms it follows the classical full physics atmospheric inversion procedures [32–34], FLEX AI introduces two main approximations: (1) considering $S \cdot \rho \approx S \cdot \rho_{\text{app}}$, and (2) using the second order series expansion. Assuming both, the independent convolution of each of the atmospheric transfer functions is avoided, which would otherwise rapidly derive into errors in the estimated SIF. Results suggest that the impact of these two approximations is at least one order of magnitude lower than the impact on TOA radiance caused by varying 10% of the SIF signal. This implies that FLEX AI formulation becomes suitable for fluorescence retrieval.

On the other hand, at the surface level, apparent reflectance ($\rho_{\text{app}}$) must be disentangled from emitted SIF and reflectance. As it was earlier noted by Verhoef et al. [30], $\rho_{\text{app}} = \rho + \frac{\pi \cdot F \cdot E}{4}$ and $\rho_{\text{app(\text{atm})}}$ from the AI are substantially different, particularly in the absorption regions. This bears consequences to the SF method. In essence, SF strategies are based on minimizing the difference between simulated and measured spectra to decouple $\rho$ and SIF. In the case of spaceborne or even airborne scale, measured data at the surface level is always derived from an AI procedure. Thus, in order to consistently couple the atmospherically inverted and the simulated signal, the formulation used should be consistent in both processes. Conversely, if a different formulation was used in the simulation and in the AI, then the SF method would minimize the differences between two inconsistently defined spectra. As demonstrated in Section 2.2, this would lead to errors depending on the atmospheric transfer functions, i.e., depending on the atmospheric state.

To close this section, we should remark that reported error values are estimated under the assumption of a Lambertian surface without considering instrument noise. Evaluating the formulation under ideal conditions allows any other source of uncertainty to be discarded. Therefore, the errors identified here must be understood as an error baseline, intrinsic to the AI and to the coupling processes.

4.2. Apparent Reflectance Spectral Distortion Analysis and Exploitation

An initial analysis was designed to evaluate the apparent reflectance spectral distortions produced in the oxygen regions by over-/under-estimating three different aerosol optical properties. The selected varying aerosol optical properties were the AOT, the Angstrom exponent ($\alpha$) and the asymmetry parameter of the HG scattering phase function ($g$), related to the aerosol load, aerosol size, and aerosol scattering anisotropy. Among them, the AOT and the $g$ are the driving aerosol optical parameters, as they have a higher impact on the spectral distortions of the surface apparent reflectance ($\rho_{\text{app(\text{atm})}}$). These results are in agreement with the aerosol sensitivity analyses performed in [41,42].

From this analysis, it was found that the spectral distortions produced by each evaluated aerosol property differ not only in magnitude but also in the spectral shape. This suggests that spectral distortions in the $\rho_{\text{app}}$ could be used to refine the estimation of atmospheric parameters in the AC process.

Therefore, we investigated the possibility of using the apparent reflectance to detect inaccuracies in the AC. A simple exercise was performed as a proof of concept, creating a DB of erroneously corrected $\rho_{\text{app(\text{atm})}}$: Using a simple multivariate linear regression algorithm, 30% of the DB was used for training and the remaining 70% of the DB was used for validation. Results suggest that the driving aerosol optical properties, i.e., AOT and $g$, can be predicted with an error below 0.5% (for the AOT) and 3% (for $g$) for 75% of the cases. Despite the low sensitivity to the Angstrom exponent, the proposed method can still reduce its errors below 8% for 75% of the cases. However, some limitations are worth noting. Although these results support the hypothesis that the retrieval of aerosol optical properties from an inaccurate AC could be refined due to the distinct spectral pattern produced, the study should
be extended to a more complex DB in a future work. Regarding expected improvements related to atmospheric simulations, they should not be limited to the spectral distortions produced by the AOT, \(a\) and \(g\) aerosol properties, but should be extended to other aerosol parameters such as the aerosol vertical distribution or the single scattering albedo. In the particular case of the aerosol vertical distribution, its radiometric effect on TOA radiance will strongly depend on the mathematical function assumed to model the vertical distribution, e.g., Gaussian \([13]\) or exponential \([42]\). Thus, a more extensive aerosol sensitivity analysis would be beneficial to identify the driving aerosol parameters to be included in a future DB simulation, i.e., to identify those parameters that highly distort the TOA radiance at the oxygen regions under realistic scenarios. With respect to how other atmospheric parameters affect the DB training statistics, e.g., WV or surface pressure, this will strongly depend on the range of the values covered. While the used DB included two different WV and surface pressure (surface height) values, in a new DB design, a wider range and a larger number of grid-points values should be incorporated. Also regarding the surface level, not only simulated but real surface reflectance and SIF spectra should be considered. Additionally, cases where each reflectance spectrum is associated to a range of SIF spectra should also be incorporated. At the same time, more efforts are required to study the impact of an inaccurate ISRF characterization on the predictive power of \(\rho_{app}\), as well as other instrumental effects, such as radiometric noises, residual polarisation effects, or stray-light contamination. In the particular case of the FLEX mission, a polarization scramble will be located in front of the telescope to cover the full pupil. Thus, while the remaining polarisation effects are expected to be within 1% of the threshold limit established by the FLEX mission \([22]\), other effects such as stray-light contamination can deserve more attention when the observed image presents bright objects. In this respect, to make the apparent reflectance exploitation a solid strategy to improve the accuracy of the atmospheric correction process, future work should also consider the spectral distortions generated by each of these instrumental effects. In addition, this study assumed a Lambertian surface reflectance to create the simulated DB. Future work should therefore consider using a non-Lambertian surface reflectance. A more extensive analysis should also ascertain the most suitable regression algorithm to avoid local minima or divergent solutions. In addition, with respect to the use of a dimensionality reduction technique, it is necessary to determine which is the best method and the optimal number of components to be considered that keep the key spectral features intact.

As a final remark, the regression algorithm was trained and validated here for the \(O_2–A\) and \(O_2–B\) bands separately. In the future, the refinement of the atmospheric correction by using both regions simultaneously to disentangle atmospheric distortions should also be investigated. For instance, while spectral changes in the scattering phase function are typically smooth, others such as the stray-light contamination will probably become stronger in the \(O_2–A\) band than in the \(O_2–B\) due to its dependency on the radiance spectral contrast.

To the best of our knowledge, the use of the apparent reflectance is proposed here for the first time as a quality indicator. Although this opens up possibilities to implement the quality check into the processing of the FLEX/Sentinel-3 tandem mission, it requires a more in-depth inspection. Therefore, follow-up analyses are foreseen by using the FLEX End-to-End simulator \([43]\), which has the versatility to assess the impact of all of the proposed improvements and subsequently quantify the impact on the estimated SIF.

5. Conclusions

Remote sensing retrieval of Solar Induced chlorophyll Fluorescence (SIF) from Top-Of-Atmosphere (TOA) is challenging due to the tininess of the SIF signal. For this reason, the multiple strategies developed in the last few years to retrieve SIF from satellite data make use of the solar and terrestrial absorption regions where radiance at TOA is considerably reduced. In the particular case of the \(O_2–B\) and \(O_2–A\) bands, these are conveniently located in the SIF spectrum, i.e., close to each of the typical SIF emission peaks, being wider and deeper than the solar lines and thus more sensitive for SIF detection. However, oxygen bands are highly affected by atmospheric
effects, especially by the aerosols scattering and surface pressure. For this reason, the FLuorescence EXplorer (FLEX)/Sentinel-3 tandem space mission proposes an SIF retrieval strategy based on a two-step approach: (1) performing the Atmospheric Correction (AC), and (2) applying a Spectral Fitting (SF) technique to retrieve SIF. Given the tininess of the SIF signal, it is well-known that any inaccuracy in the atmospheric state characterization can lead to errors in the estimated SIF. However, just as important as the atmospheric parameters characterization is the use of an appropriate formulation to infer surface apparent reflectance from TOA radiance and to apply a consistent coupling process between the AC and SF strategies. Thus, assuming the atmospheric state as known, we evaluated the mathematical formulation of the FLEX data processing scheme, from the TOA radiance inversion to its coupling with the at-surface level SF method. It was shown that the impact on TOA radiance of the approximations assumed by FLEX is around one order of magnitude lower than an expected variation caused by increasing 10% of the value of a typically SIF emission spectrum. It was also demonstrated that an inconsistent AC-SF coupling process derives errors that depend on the atmospheric transfer functions because of the formulation involved in the process. One of the advantages of using a two-step approach resides in the intermediate products generated, such as the apparent reflectance ($\rho_{\text{app}}$). Using simulated data with SCOPE and MODTRAN radiative transfer models, we subsequently analysed the spectral distortions propagated to $\rho_{\text{app}}$ due to inaccuracies in the estimation of aerosol optical properties. As it was expected, different distortion patterns appear in the $\rho_{\text{app}}$ due to the over-/under-estimation of each of the key aerosol optical properties: the AOT, the Angstrom exponent ($\alpha$) and the Henyey–Greenstein (HG) scattering anisotropy parameter ($g$). This leads us to proposing to exploit the use of the $\rho_{\text{app}}$ as a quality indicator of the AC. As a proof of concept, a simple exercise was performed over $\sim8 \cdot 10^4$ incorrectly atmospherically corrected $\rho_{\text{app}}$ database by using a multivariate linear regression technique to predict the over-/under-estimation of the aerosol optical properties. First, results showed that the AOT, $\alpha$ and the $g$ parameters can be corrected with a relative error lower than 0.5%, 8% and 3%, respectively, for 75% of the evaluated data. In the context of the FLEX/Sentinel-3 tandem mission, the exploitation of the apparent reflectance spectral distortions can open up new opportunities to refine the atmospheric parameters obtained from the AC process.

Acknowledgments: The authors extend their thanks to Valero-Laparra for the contribution to this work. Thanks also go to the Spanish Ministry of Economy and Competitiveness predoctoral grant BES-C-2014-0087 for funding, the AVANFLEX (Advanced Products for the FLEX mission), and the ESP2016-79503-C2-1-P National Program for the Promotion of Scientific and Technical Research of Excellence Ministry of Economy and Competitiveness, Spain.

Author Contributions: Neus Sabater performed the calculus and wrote the paper, Jorge Vicent helped with the calculus performance and the discussion, Luis Alonso supervised and helped with the whole process of understanding, Jochem Verrelst contributed to the paper structure and improved the paper’s English, Sergio Cogliati contributed with the Spectral Fitting coding and José Moreno supervised and guided the full study.

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

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Atmospheric Correction</td>
</tr>
<tr>
<td>AI</td>
<td>Atmospheric Inversion</td>
</tr>
<tr>
<td>AOT</td>
<td>Aerosol Optical Thickness</td>
</tr>
<tr>
<td>DB</td>
<td>DataBase</td>
</tr>
<tr>
<td>FLEX</td>
<td>FLuorescence EXplorer</td>
</tr>
<tr>
<td>FLORIS</td>
<td>FLuORescence Imaging Spectrometer</td>
</tr>
<tr>
<td>HG</td>
<td>Henyey–Greenstein</td>
</tr>
<tr>
<td>HR</td>
<td>High Resolution</td>
</tr>
</tbody>
</table>
Appendix A

Table A1. FLORIS instrument Spectral Resolution (SR) and Spectral Sampling Interval (SSI) for each spectral interval covered [22].

<table>
<thead>
<tr>
<th>Band</th>
<th>PRI Band</th>
<th>Chlorophyll</th>
<th>O₂−B</th>
<th>Red-Edge</th>
<th>O₂−A</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR (nm)</td>
<td>3</td>
<td>3</td>
<td>0.6</td>
<td>0.3</td>
<td>2</td>
</tr>
<tr>
<td>SI (nm)</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>0.1</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Appendix B

Table A2. MODTRAN input parameters used to generate Figures 2–4. In the particular case of Figure 3(a, b) the Aerosol Optical Thickness (AOT) values range from 0.05 (−) to 0.42 (−), as indicated in the figure legend.

<table>
<thead>
<tr>
<th>MODTRAN Input Parameter</th>
<th>Value (Units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospheric parameter</td>
<td></td>
</tr>
<tr>
<td>Model of atmosphere</td>
<td>Mid Latitude Summer</td>
</tr>
<tr>
<td>AOT at 550 nm</td>
<td>0.05 (−)</td>
</tr>
<tr>
<td>Angstrom exponent</td>
<td>0.79 (−)</td>
</tr>
<tr>
<td>Henyey–Greenstein asymmetry (g)</td>
<td>0.8 (−)</td>
</tr>
<tr>
<td>Water vapour</td>
<td>2.4 (g/cm²)</td>
</tr>
<tr>
<td>Geometry parameter</td>
<td></td>
</tr>
<tr>
<td>Surface elevation</td>
<td>100 (m)</td>
</tr>
<tr>
<td>Solar Zenith Angle</td>
<td>45 (°)</td>
</tr>
<tr>
<td>Viewing Zenith Angle</td>
<td>0 (°)</td>
</tr>
<tr>
<td>Relative Azimuth Angle between sun and sensor</td>
<td>0 (°)</td>
</tr>
<tr>
<td>High Spectral Resolution</td>
<td></td>
</tr>
<tr>
<td>Spectral Resolution at O₂−A</td>
<td>0.006 (nm)</td>
</tr>
<tr>
<td>Instrumental Spectral Response</td>
<td></td>
</tr>
<tr>
<td>Spectral bandwidth (c′)</td>
<td>0.3 (nm)</td>
</tr>
</tbody>
</table>

* The double sigmoid function used as the instrumental spectral response function corresponds to Equation (A1):

\[
f_\lambda(c, \sigma, s) = \frac{sgn[-s \cdot (\lambda - c + \sigma/2)] - sgn[-s \cdot (\lambda - c - \sigma/2)]}{sgn[s]} = \frac{1}{1+e^{-c}}, \tag{A1}\]

where \( \sigma \) is the Full Width High Maximum (FWHM), \( s \) the slope and \( c \) is the barycentre in the wavelength domain.
Appendix C

Table A3. Biophysical parameters, Leaf Area Index (LAI) and Chl content used as input in the SCOPE model (v1.53) to generate the simulated DB. Spectra labelled as (a), (b), and (c) corresponds to the spectra showed in Figure 7.

<table>
<thead>
<tr>
<th></th>
<th>LAI (−)</th>
<th>Chl (g/cm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare soil (a)</td>
<td>0</td>
<td>0.47</td>
</tr>
<tr>
<td>Mixed bare soil and dense vegetation</td>
<td>0.5</td>
<td>40</td>
</tr>
<tr>
<td>Mixed bare soil and dense vegetation (b)</td>
<td>1.5</td>
<td>10</td>
</tr>
<tr>
<td>Mixed bare soil and dense vegetation</td>
<td>2.5</td>
<td>40</td>
</tr>
<tr>
<td>Dense vegetation (c)</td>
<td>4</td>
<td>70</td>
</tr>
</tbody>
</table>

References


