SCUOLA DI DOTTORATO
UNIVERSITA' DEGLI STUDI DI MILANO-BICOCCA

Department of Earth and Environmental Sciences

Ph.D. program: **Chemical, Geological and Environmental Sciences** Cycle: **XXXIV**
Curriculum in: **Environmental Sciences**

Solar-induced chlorophyll fluorescence signal retrieval in terrestrial vegetation and inland waters from hyperspectral proximal sensing

Surname: **Cesana**
Name: **Ilaria**
Registration number: **727177**

Tutor: **Prof. Antonio Finizio**
Supervisor: **Dr. Mariano Bresciani**
Supervisor: **Dr. Sergio Cogliati**
Supervisor: **Dr. Claudia Giardino**
Coordinator: **Prof. Marco G. Malusà**

ACADEMIC YEAR 2020-2021

Abstract

The PhD research aimed to develop novel strategies able to better retrieve and interpret the chlorophyll Solar-Induced Fluorescence (SIF) signal emitted by terrestrial vegetation and inland waters at ground level, to advance the understanding of ecosystems structure and functioning.

SIF metrics were defined taking advantage of the full SIF spectrum available from the recently developed *"spectrum-fitting"* algorithm (SpecFit). The metrics were designed to characterize the SIF spectrum in terms of red and far-red peaks maximum values (SIF_{red}, $SIF_{far-red}$), corresponding wavelengths and the spectrally integrated value (SIF_{INT}). SIF typically evaluated in the O_2-A (SIF_{760}) and O_2-B (SIF_{687}) bands and reflectance indices (used as proxies for canopy biophysical parameters) were compared to the SIF spectrum. The reflectance indices selected were the $NDVI_{red-edge}$, $CI_{red-edge}$, NIR_v and PRI. The above-mentioned analysis were carried out at seasonal and diurnal scales, exploiting top-of-canopy (TOC) spectral measurements acquired over three different crops. The SIF evaluated at the peaks always shows a strong correlation with the corresponding O_2 bands values, while the SIF_{INT} represents a more complete parameter characterized by peculiar dynamics. At diurnal scale, the combined use of reflectance indices and TOC SIF metrics allows to gain a better knowledge of the crops dynamics. Seasonally, the SIF and reflectance indices show more similar temporal evolution along the growth-phases because they are mainly driven by changes in the overall canopy biomass, chlorophyll content and incident light. The reabsorption of the SIF within the canopy-leaf system affects the overall SIF spectral shape and magnitude at this temporal scale. As demonstrated on the synthetic dataset, the reabsorption effect prevents an accurate evaluation of the fluorescence quantum yield (SIF_{yield}). Correcting the TOC SIF spectrum for the reabsorption is pivotal. In this regard, two different approaches were developed and tested. The parametric method enables to correct SIF for the reabsorption (SIF^{RC}) establishing parametric relationships with spectral variables routinely measured at TOC. The method accuracy depends on the plant growth phase, showing better results for medium-dense canopies. This behavior compromises the application of the method on the full seasonal analysis. The second approach, based on Fourier-Machine Learning algorithm, retrieves the SIF^{RC} along with biophysical parameters

of interest (i.e. LAI, C_{ab}, SIF_{yield}, aPAR) with a better accuracy for all the conditions investigated. The two approaches outcomes were compared by considering synthetic simulations and real field measurements. These methods were developed and tested starting from different assumptions: the parametric method can be used in a simpler way but it lacks accuracy for sparse conditions, while the Fourier-Machine Learning algorithm is more complex but offer better results.

Regarding clear lake waters, a novel version of the Fluorescence Line Height approach was implemented. The SIF proxy obtained agrees with the temporal evolution of other conventional spectral indices (E_{PAR}, R_{550} and [Chl-a]). Novel phytoplankton primary production models were defined and tested adapting the vegetation Light Use Efficiency model for inland waters. Promising results were achieved when the SIF_{FLH} and a novel photosynthesis efficiency proxy here introduced were considered ($R^2 > 0.95$).

In conclusion, this work allowed to develop new concepts and processing methods towards a better interpretation of SIF at diurnal/seasonal scales in different ecosystems. The approaches were prototyped on synthetic data and employed on real field measurements. The results obtained highlight the relevance to retrieve the SIF spectrum and the importance to employ SIF reabsorption correction methods to obtain relevant parameters (i.e. reabsorption corrected SIF_{yield}) better related with terrestrial vegetation functioning and less affected from canopy structure. Concerning clear lake waters, this study has demonstrated that hyperspectral and high frequency measurements allow to follow the phytoplankton dynamics, particularly in clear sky days. Furthermore, the use of parameters linked to the SIF represents a promising approach in the phytoplankton primary production monitoring and estimation in lakes.

Contents

Page

Acronyms and Symbols **xvii**

1 Introduction **1**
1.1 Objectives . . . 5
1.2 Thesis Outlines . . . 6

2 Characterizing the SIF full spectrum at Seasonal/Diurnal scales for agricultural crops **9**
2.1 Theoretical background . . . 9
2.2 Material and Methods . . . 10
2.2.1 Field measurements over different crops . . . 10
2.2.2 Vegetation Reflectance Indices . . . 13
2.2.3 SIF spectrum metrics . . . 15
2.3 Results . . . 17
2.3.1 Seasonal analysis . . . 17
2.3.1.1 Reflectance Indices . . . 17
2.3.1.2 SIF full spectrum metrics . . . 18
2.3.1.3 SIF metrics and Reflectance indices relationships . . . 23
2.3.2 Diurnal analysis . . . 27
2.3.2.1 Reflectance Indices . . . 27
2.3.2.2 SIF full spectrum metrics . . . 30
2.3.2.3 SIF metrics and Reflectance indices relationships . . . 31
2.4 Discussion . . . 33
2.4.1 Seasonal analysis . . . 33
2.4.2 Diurnal analysis . . . 36
2.5 Conclusions . . . 38

3 Correcting SIF for canopy reabsorption for quantum yield estimations 41
3.1 Theoretical background . . . 41
3.2 Material and Methods . . . 43
3.2.1 Radiative Transfer simulations . . . 43
3.2.2 Estimation of quantum yield . . . 47
3.2.3 Estimation of the amount of light absorbed by the canopy (aPAR) . . . 48
3.2.4 Method to correct the SIF spectrum: parametric method . . . 52
3.2.5 Fourier-ML algorithm . . . 57
3.3 Results . . . 58
3.3.1 Characterizing the reabsorption effect . . . 58
3.3.2 aPAR parametric method accuracy (RT simulations) . . . 64
3.3.3 SIF^{RC} retrieval accuracy (RT simulations) . . . 66
3.3.3.1 Parametric method . . . 66
3.3.3.2 Fourier-ML algorithm . . . 74
3.3.4 Qualitative comparison of the methods on field measurements . . . 76
3.4 Discussion . . . 77
3.4.1 Reabsorption effect on the Top-Of-Canopy SIF . . . 77
3.4.2 Parametric method . . . 79
3.4.3 Fourier-ML algorithm . . . 82
3.5 Conclusions . . . 82

4 Advancing SIF retrieval in clear lake waters and development of primary production models 85
4.1 Theoretical Background . . . 85
4.2 Material and Methods . . . 87
4.2.1 Measurements description . . . 87
4.2.1.1 Study Area . . . 87
4.2.1.2 Water samples collection . . . 87
4.2.1.3 Hyperspectral measurements acquisition . . . 88
4.2.2 Measurements analysis . . . 90
4.2.2.1 Water samples laboratory analysis . . . 90
4.2.2.2 Hyperspectral measurements analysis . . . 91
4.2.2.3 Novel FLH parametrization exploiting hyperspectral measurements . . . 92
4.2.2.4 Spectral indices evaluated on hyperspectral measurements . . . 94
4.2.3 Phytoplankton primary production (PP) in inland waters . . . 95
4.2.3.1 PP definition exploiting in-situ quantities . . . 95
4.2.3.2 PP models exploiting hyperspectral measurements . . . 97
4.3 Results . . . 98
4.3.1 Water body phytoplankton characterization . . . 98
4.3.2 Phytoplankton temporal dynamics . . . 99

4.3.2.1 Comparison between water sampling and hyperspectral quantities . . . 99
4.3.2.2 Hyperspectral measurements analysis . . . 103
4.3.3 Test of phytoplankton primary production models . . . 105
4.4 Discussion . . . 107
4.4.1 Water body phytoplankton characterization . . . 107
4.4.2 Phytoplankton temporal dynamics . . . 107
4.4.3 Test of phytoplankton primary production models . . . 109
4.5 Conclusions . . . 110

5 Conclusions **111**
5.1 Main Results . . . 111
5.2 Concluding remarks and future perspectives . . . 116

References **117**

List of Figures

2.1 Location of the FLEXSense campaign site in Italy. Panels A, B and C show the three crops investigated (forage, alfalfa and corn), the experimental set-up and the acquisition time intervals. 12

2.2 Fluorescence metrics defined on the full spectrum. Black line corresponds to the incoming radiance (L_d), red line to the top of canopy SIF, grey bands to the O_2 absorption bands, grey dashed lines to the wavelengths' positions corresponding to the peaks in the visible red and far-red, respectively. Black dots highlight the SIF metrics position on the full spectrum. 16

2.3 Seasonal trend corresponding to the reflectance-based indices. Panel A contains the PAR, Panel B shows the $NDVI_{red-edge}$ (in blue) and the NIR_v (in red), Panel C displays the $CI_{red-edge}$. The temporal information is showed on the x axes as Day of the Year (DOY). Data showed have been averaged over the 30 minutes interval around the solar noon. Dashed lines divide the different crop time series: forage, alfalfa and corn, respectively. . . 17

2.4 Seasonal trend corresponding to the SIF full spectrum metrics. Panel A contains the PAR (in blue) plotted together the SIF_{INT} (in red), Panel B shows metrics evaluated in the far-red spectral region (in blue SIF_{760} and in red $SIF_{far-red}$), Panel C displays the metrics defined in the visible red spectral region (in blue SIF_{687} and in red SIF_{red}). The temporal information is showed on the x axes as Day of the Year (DOY). Data showed were averaged over the 30 minutes interval around the solar noon. Dashed lines divide the different crop time series: forage, alfalfa and corn, respectively. . . 19

2.5 (top) Comparison between the SIF metrics evaluated in the far-red spectral range. (bottom) comparison between the metrics evaluated in the red spectral range. All the crops are displayed together; the colored dots represent the whole acquired time series. Color bar contains the temporal information. Black dots correspond to the spectral measurements averaged around the solar noon, while the error bars are the standard deviations. Red lines are the linear regressions performed on the averaged values. 20

2.6 Experimental SIF peaks position trends for different $CI_{red-edge}$ values. (left) λ_{red}. (right) $\lambda_{far-red}$.Data showed have been averaged over the 30 minute interval around the solar noon. Blue dots correspond to forage, red triangles to alfalfa and green square to corn. 22
2.7 Comparison between selected SIF metrics and R indices. Blue dots correspond to forage, red triangles to alfalfa and green square to corn. Values displayed correspond to the values acquired around the solar noon (not averaged). 23
2.8 Principal component analysis performed at seasonal scale. (A) Scores on the first two principal components. Blue dots correspond to forage, red triangles to alfalfa and green square to corn. Values displayed correspond to the values acquired around the solar noon. (B) Loadings on the first two principal components. The seven variables on which the PCA has been performed are displayed here. 25
2.9 Diurnal trends of reflectance-based indices in three selected clear sky days (forage: DOY 112, alfalfa: DOY 145, corn: DOY 217). Blue dots correspond to forage, red triangles to alfalfa, green squares to corn. Data shown correspond to half-hour average. Error bars correspond to standard deviations. The temporal information is reported in fraction of the day, where 0.5 corresponds to the solar noon. 28
2.10 Diurnal trends of the SIF metrics in three selected clear sky days (forage: DOY 112, alfalfa: DOY 145, corn: DOY 217). Blue dots correspond to forage, red triangles to alfalfa, green squares to corn. Data shown correspond to half-hour average. Error bars correspond to the standard deviations. The temporal information is reported in fractions of the day, where 0.5 corresponds to the solar noon. 30
2.11 Principal component analysis performed at diurnal scale. (A) Scores on the first two principal components. Blue dots correspond to forage (DOY 112), red triangles to alfalfa (DOY 145) and green square to corn (DOY 217). Values displayed correspond to values acquired during the day. (B) Loadings on the first two principal components. The seven variables on which the PCA was performed are displayed here. 32

3.1 MODTRAN5-SCOPE code. Step 1: generation of the t_n. Step 2: t_n given as input in SCOPE, that run for a selected SZA, LAI, C_{ab}, Fqe combination. This module gives as output the TOC reflectance components and SIF (SIF^{TOC}), the SIF at the photosystems level (SIF^{RC}) and the aPAR. Step 3: TOC reflectance components, the SIF are coupled to the t_n obtaining L_d, L_u, R_{app} and R. Step 4: SIF^{TOC}, SIF^{RC}, L_d, L_u, R_{app}, R are convolute with the FLoX spectral resolution. Step2-4 are iterated for all the SZA, LAI, C_{ab}, Fqe combinations identified. 45

3.2 The 10000 C_{ab} (in $\mu g cm^{-2}$) and LAI ($m^2 m^{-2}$) couples simulated. LAI-C_{ab} values can be qualitatively grouped in four categories: bare soil (S0) in brown, sparse (V1) in orange, medium (V2) in light green and dense (V3) in dark green vegetation. 46

3.3 Absorbed Photosynthetically Active Radiation (aPAR). Panel A shows the aPAR evaluated for increasing LAIxC_{ab} values (SZA=20° and Fqe = 0.013). The color code allows to discern between bare soil (brown), sparse (orange), medium (light green) and dense (dark green) vegetation. Panel B collects the aPAR plotted for all the SZA simulated (Fqe = 0.013). Panel C display the aPAR for variable Fqe (SZA = 20°). 48

3.4 Comparison between the aPAR and the reflectance-based indices, namely Panel A: $NDVI_{red-edge}$, Panel B: $CI_{red-edge}$ and Panel C: NIR_v. Data displayed corresponds to the values obtained evaluating the spectral indices on the simulated reflectance in agreement with the equations collected in Table 2.1. Each point plotted corresponds to a LAIxC_{ab} couple. SZA= 20°. . 49

3.5 Results of the fit performed over the aPAR-NIR_v comparison displayed for increasing SZA values. Panel A: R^2, Panel B: RMSE, Panel C: k_1 and corresponding margin of error, Panel D: k_2 and corresponding margin of error. 51

3.6 Panel A: a_1 and Panel B: a_2 behaviours for increasing LAIxC_{ab} values and different simulated Fqe. Panel C: ratio between a_2 and a_1. Dashed black line corresponds to the constant values of 0.2546. This ratio is independent the LAIxC_{ab} and Fqe values considered. Panel D: comparison between TOC SIF evaluated at 760 nm and a_1. All the values displayed were simulated considering a SZA equal to 20°. Color code allows to discern between the different Fqe values. 53

3.7 Results of the fitting performed on the SIF_{760}^{TOC} and a_1 comparison for SZA= 20°. Each point (black dot) corresponds to a specific LAIxC_{ab} values. In red is represented the fit carried out. 54

3.8 Panel A: coefficients a obtained from the fit using equation 3.2.9 displayed respect to the Fqe. Panel B: coefficients b obtained from the fit using equation 3.2.9 displayed with respect to the Fqe. Each point corresponds to the values obtained from the fit, while the error bar to the correspond margin of error (see Figure 3.7 as reference). The color code indicates the different SZA considered. 55

3.9 Trend between the a_{avg} values and the cos(SZA). For a fixed angle (i.e. SZA, the a_{avg} were obtained averaging over all the Fqe considered. Therefore, the data displayed (white dots) corresponds to mean value estimated, while the error bar to the standard deviations. 56

3.10 Comparison between SIF spectra simulated at TOC scale (orange) and corrected for the reabsorption (SIF^{RC} in blue). They are both measured in $mW m^{-2} nm^{-1} sr^{-1}$. Each column corresponds to a different LAI (in $m^2 m^{-2}$) and C_{ab} (in $\mu g cm^{-2}$) couple, increasing from left to right. Each row corresponds to different SZA and Fqe values. 58

3.11 (top): Fluorescence metrics evaluated on the TOC SIF spectra. (bottom): Fluorescence metrics evaluated on the SIF corrected for the re-absorption (i.e. at the photosystems level). The trends showed are evaluated on the synthetic dataset. Each point corresponds to a single LAIxC_{ab} value. Measurements referred to the bare soil were not considered. Color code (white rather than grey) allows to discern between the metrics evaluated in the same spectral region. Except for the SIF_{INT} cases, grey dots correspond to the SIF estimated in the O_2 bands, white dots to the fluorescence at the peaks. Specifically, SZA= 20°. Fqe = 0.013. 60

3.12 (top): peaks positions in the far-red (left) and visible red (right) evaluated on the TOC SIF for different SZA (Fqe = 0.013). (bottom): peak positions in the far-red (left) and visible red (right) evaluated on the TOC SIF for different Fqe (SZA= 20°). Each point corresponds to a specific LAIxC_{ab} value. The color code helps to discern between the whole SZA and Fqe considered. . . . 62

3.13 Panel A: SIF^{TOC}_{yield} for different SZA values. Fqe = 0.013. Panel B: Comparison between the Fqe (SCOPE input) and the SIF^{TOC}_{yield} evaluated according to eq. 3.2.6. Panel C: SIF_{yield} for different SZA values. Fqe = 0.013. Panel D: Comparison between the Fqe (SCOPE input) and the SIF_{yield} evaluated according to eq. 3.2.5. Color code in Panel A and C corresponds to the several SZA considered. Data displayed in Panel B and D correspond to the average performed over all the LAIxC_{ab} and SZA values for a fixed Fqe. The error bars are the standard deviations. In red the linear regression performed. The dashed line is the 1:1. 63

3.14 Comparison between the aPAR generated by SCOPE ($aPAR_{true}$) and the aPAR evaluated with the method proposed in Section 3.2.3 ($aPAR_{retr}$). Points corresponds to the LAIxC_{ab}. Dashed grey line is the 1:1. In red is represented the linear regression performed. 65

3.15 Comparison between SIF^{RC} spectra simulated with SCOPE (in blue) and using the parametric method (in red). Case A: spectra evaluated with approximation on the coefficient a (a_{avg}). Case B: spectra evaluated without approximation on the coefficient a (a(Fqe)). SZA= 20°. Fqe= 0.013. 67

3.16 RRMSE% obtained comparing SIF simulated with SCOPE and SIF spectra corrected for reabsorption with the parametric method. 68

3.17 (first row): scatterplot between the SIF^{RC} metrics evaluated on spectra simulated and corrected for reabsorption (SZA= 20°; Fqe= 0.013). (rows 2-5): outcomes of the linear regression for all the cases investigated. A detailed Figure description is reported in the text. Dots colors are linked to the whole Fqe values investigated (blue low, green medium, orange high Fqe). 70

3.18 SIF_{yield} evaluated on the synthetic dataset. Each row corresponds to a case described in Table 3.2. White dots correspond to the several $LAIxC_{ab}$ values. The red line is the Fqe given as input in SCOPE. SZA = 20°. 72

3.19 RMSE and RRMSE% evaluate on the comparison between Fqe and the SIF_{yield} for the six cases summarized in Table 3.2. The values are displayed for increasing SZA. Color code allows to discern between the several Fqe investigated. 74

3.20 Comparisons between estimated and simulated values. The first row collects the biophysical parameters, while the second row the SIF metrics. The dashed grey line represents the 1:1. In white is displayed the linear fit. SZA = 40°. 75

3.21 Panel A: seasonal trends obtained with the parametric method. Panel B: seasonal trends obtained with the Fourier-ML algorithm using two different approaches (considering the O_2 spectral bands, in red and excluding the O_2 spectral bands in blue). Points displayed correspond to the average value on the time interval investigated ($\pm$ 15 minutes around the solar noon). The error bars are the standard deviations. Dashed line separates the two crops investigated. 77

4.1 Lake Maggiore and experimental location. (top right) Buoy on which ROX has been mounted on. It was left around 50 meters far from the lake shore to avoid the bottom contribution in the spectra collected. (bottom right) Experimental set-up. 89

4.2 (A): R_{rs} spectra times series. (B): λ_C, λ_L, λ_R positions in the R_{rs} spectrum (red points). The two arrows highlight the artefacts due to O_2 absorption bands. (C): the grey area shows the spectral range in which the λ_L was searched for. The red line represents the second-degree polynomial used as fit function. (D): the grey area shows the spectral range in which the λ_C was searched for. The red line corresponds to the Gaussian used as fit function. . 93

4.3 Composition of the phytoplankton major taxa expressed in percentage respect to biovolume (A) and density (B). Each color corresponds to a different sample (see nomenclature in Table 4.1). 99

4.4 Dots represent variables measured close to the surface (z_0), while triangles refer to the sampling performed at the Secchi Disk depth (z_{SD}). Time is shown as Day-Of-the Year (DOY): 183 is the 02.07.2019 while 184 is the 03.07.2019. First line (Panel A left and right) shows the E_{PAR} (in blue) and SIF_{FLH} (in red) obtained from the ROX hyperspectral measurements. (B) and (D) are the $[Chl\text{-}a]_{HPLC}$ and the biovolume, respectively, evaluated from laboratory analysis. (C) and (E) corresponding to F_A and F_{SIF} have been obtained combining the laboratory outcomes and the Satlantic measurements. All the quantities displayed, except the biovolume (D), are mean values. The error bars correspond to the standard deviations. 100
4.5 Values displayed correspond to measurements evaluated at the surface. The colors (grey scale) help to discern between the several samples considered. All the displayed measurements correspond to mean values and the error bars show the standard deviations. 101
4.6 (A): Φ_{SIF} values obtained from the laboratory analysis. (B): relative Φ_{SIF} ratio. (C) comparison between the SIF_{FLH}/F_A and Φ_{SIF}. (D) comparison between SIF_{FLH}/F_A and Φ'_C. 102
4.7 (A): E_{PAR}; (B) R^{app}_{550}; (C) SIF_{FLH} and (D) $[Chl\text{-}a]_{OC4}$. Data shown represent the mean values (n_{max} per interval ~10), averaged on a time interval of 10 minutes, while the error bars correspond to the standard deviations. 104
4.8 (A): comparison between E_{PAR} and SIF_{FLH}. Each symbol corresponds to a different acquisition day. Data shown represent the mean values (n_{max} per interval ~10), averaged on a time interval of 10 minutes, while the error bars correspond to the standard deviations. (B) comparison between the $[Chl\text{-}a]_{OC4}$ and $[Chl\text{-}a]_{HPLC}$. Data showed corresponds to $[Chl\text{-}a]_{HPLC}$ collected close to the surface. 105
4.9 First row collects the cases in which the Φ_C proxy has been kept constant, while the F_A has been replaced by indices and metrics obtained from the hyperspectral measurements. Second row collects the cases in which also the Φ_C has been replaced by a proxy defined from RS quantities. The red lines correspond to the linear regression performed on the measurements. The scale for F_{C-RS} have been omitted on purpose. Due to the approximations taken only a qualitative comparison was possible. 106

List of Tables

2.1 Reflectance-based indices. The PRI index used have been evaluated automatically by the FLoX routine. The spectral bands used to estimate the other indices are 670-680 nm (red), 705-715 nm (red-edge) and 770-780 nm (far-red). 14
2.2 Fluorescence metrics evaluated on the full spectrum. 16
2.3 Coefficients obtained from the linear regressions performed on the comparison between the SIF full spectrum metrics. The fits have been evaluated only on the values averaged around the solar noon. The Case ID corresponds to the several Panels collected in Figure 2.5. 21

3.1 Coefficients and statistics obtained from the liner regression shown in Figure 3.13B (Case ID TOC) and Figure 3.13D (Case ID RC). p1 corresponds to the linear regression slope, while p2 to the intercept. 64
3.2 SIF_{yield} evaluated on the synthetic dataset exploiting six different strategies. 71
3.3 RRMSE% evaluated on the comparison between the variables (biophysical parameters and SIF metrics) simulated and the values estimated with the Fourier-ML algorithm. The RRMSE% corresponds to the average value over all the SZA in the dataset, while the margin of errors to the standard deviations. 76

4.1 Nomenclature used in Chapter 4 to discern between the different water sample collected. The temporal information is reported in local time (UTC+1). 88
4.2 Values obtained from the laboratory analysis. The $[\text{Chl-a}]_{HPLC}$ values shown in the table have been evaluated averaging the two replica available. The errors associated correspond to the standard deviation. 91
4.3 Phytoplankton primary production models. 97

Acronyms and Symbols

$\mathbf{a_{CDOM}(440)}$	Colored Dissolved Organic Matter absorption coefficient at 440 nm
$\mathbf{a_{NAP}}$	Non-Algal Particles spectral absorption coefficient
$\mathbf{a_{phy}}$	phytoplankton spectral absorption coefficient
$\mathbf{a^*_{phy}}$	mean phytoplankton specific absorption coefficient
aPAR	absorbed Photosynthetically Active Radiation
BRDF	Bidirectional Reflectance Distribution Function
C3,C4	photosynthesis pathways
$\mathbf{C_{ab}}$	chlorophyll content
CDOM	Colored Dissolved Organic Matter
Chl-a	chlorophyll-a pigment
[Chl-a]	chlorphyll-a concentration
$\mathbf{[Chl\text{-}a]_{HPLC}}$	chlorophyll-a concentration estimated using the HPLC
$\mathbf{[Chl\text{-}a]_{OC4}}$	chlorophyll-a concentration evaluated with the Ocean Color algorithm using four spectral bands
$\mathbf{CI_{red-edge}}$	Chlorophyll Index exploiting the red-edge spectral region
DAYfrac	fraction of the day
DOY	Day Of the Year
$\mathbf{E_d}$	downwelling irradiance
$\mathbf{E_{PAR}}$	incoming irradiance over the Photosynthetically Active Radiation spectral region

ESA	European Space Agency
Φ_c	carbon fixation yield
$\Phi_c^{'}$	carbon fixation yield proxy (ratio between biovolume ad F_A)
Φ_{SIF}	Chl-a SIF emission quantum yield
fAPAR	fraction of light absorbed in the Photosynthetically Active Radiation
$\mathbf{F}_A$	light absorbed by the phytoplankton contained in a unit of volume
$\mathbf{F}_C$	carbon assimilation due to the phytoplankton contained in a unit of water volume
$\mathbf{F}_{C-RS}$	carbon assimilation using spectral indices and SIF
$\mathbf{F}_{SIF}$	SIF emitted by the phytoplankton contained in a unit of water volume
FLEX	FLuorescence EXplorer
FLH	Fluorescence Line Height
FOV	Field Of View
Fqe	SCOPE Fluorescence Quantum emission Efficiency
FWHM	Full Width Half Maximum
GPP	Gross Primary Production
HPLC	High Performance Liquid Chromatography
k	coefficient linked to the light attenuation in the water column
$\lambda_{far-red}$	SIF far-red wavelength position
$\lambda_{L,C,R}$	left, central and right wavelengths used in the FLH
λ_{red}	SIF red wavelength position
LAI	Leaf Area Index
LAIxC$_{ab}$	canopy chlorophyll content
$\mathbf{L}_d$	downwelling radiance
LIDF	Leaf Inclination Distribution Function
LIDF$_a$	leaf inclination
LIDF$_b$	variation in leaf inclination

$\mathbf{L_u}$	upwelling radiance
LUE	Light Use Efficiency
$\mathbf{L_w}$	leaving water radiance
$\mathbf{L_{wL,C,R}}$	left, central and right water leaving radiances used in the FLH
ML	Machine Learning
MODTRAN	MODerate resolution atmospheric TRANsmission
NDVI	Normalized Difference Vegetation Index
$\mathbf{NDVI_{red-edge}}$	Normalized Difference Vegetation Index exploiting the red-edge spectral band
$\mathbf{NIR_v}$	Near-Infrared Reflectance Vegetation index
NPQ	Non-Photochemical Quenching
$\mathbf{O_2}$**-A**	atmospheric oxygen absorption band at 760 nm
$\mathbf{O_2}$**-B**	atmospheric oxygen absorption band at 687 nm
OC	Ocean Color
PSI	photosystem I
PSII	photosystem II
PAR	incoming radiance over the Photosynthetically Active Radiation spectral region
PC	Principal Component
PC1	first Principal Component
PC2	second Principal Component
PCA	Principal Component Analysis
PP	phytoplankton Primary Production
PQ	Photochemical Quenching
PRI	Photochemical Reflectance Index
rdd	reflectance diffuse contribution in the hemispherical term due to the surrounding
rdo	reflectance diffuse contribution in the viewing direction
rsd	reflectance direct contribution in the hemispherical term due to the surrounding

rso	reflectance direct contribution in the viewing direction
R	Reflectance
$\mathbf{R}^{app}$	apparent Reflectance
$\mathbf{R}^{app}_{550}$	apparent Reflectance evaluated at 550 nm
$\mathbf{R}_{rs}$	remote sensing Reflectance
RC	Correction for the Reabsorption
RMSE	Root Mean Square Error
RRMSE%	Relative Root Mean Square Error in percentage
RS	Remote Sensing
RT	Radiative Transfer
S0	bare soil
SAIL	Scattering by Arbitrary Inclined Leaves
SCOPE	Soil-Canopy Observation of Photosynthesis and Energy
SFM	Spectral Fitting Method
SIF	chlorophyll Solar-Induced Fluorescence
$\mathbf{SIF}^{RC}$	SIF corrected for the reabsorption (photosystems level)
$\mathbf{SIF}^{TOC}$	SIF evaluated at TOC scale
$\mathbf{SIF}_{687}$	SIF values at 687 nm (O2-B band)
$\mathbf{SIF}_{760}$	SIF value at 760 nm (O2-A band)
$\mathbf{SIF}_{far-red}$	SIF maximum emission in the far-red
$\mathbf{SIF}_{FLH}$	SIF proxy evaluated with the FLH method
$\mathbf{SIF}_{INT}$	spectrally integrated SIF over 670-780 nm
$\mathbf{SIF}_{red}$	SIF maximum emission in the red
$\mathbf{SIF}_{so}$	directional TOC SIF
$\mathbf{SIF}_{yield}$	fluorescence efficiency evaluated as the ratio between SIF_{INT} and aPAR
$\mathbf{SIF}^{TOC}_{yield}$	fluorescence efficiency at TOC scale
SNR	Signal to Noise Ratio
SpecFit	Spectral Fitting method to retrieve the SIF and R full spectrum
SZA	local Solar Zenith Angle
t1	extraterrestrial solar radiance normalized by the cosine of the local SZA

t3	spherical albedo
t4	downward direct transmittance
t5	downward diffuse transmittance
t12	product between downward direct transmittance and the spherical albedo
tn	atmospheric transfer functions
TOC	Top OF Canopy
TSM	Total Suspended Matter
V1	sparse vegetation
V2	medium vegetation
V3	dense vegetation
VIS-NIR	VISible and Near-InfraRed spectral region
VGPM	Vertical Generalized Productivity Model
$\mathbf{z}_0$	depth just below the water surface
$\mathbf{z}_{SD}$	Secchi Disk depth

Chapter 1

Introduction

Chlorophyll-a (Chl-a) is a photosynthetic pigment contained in the chloroplasts of terrestrial plants and in many aquatic phytoplankton taxa. Chl-a absorbs the incoming solar light in the Photosynthetically Active Radiation spectral region between 400-700 nm. Under non-stressed conditions, the absorbed light is primarily used to drive photosynthesis in a process known as Photochemical Quenching (PQ). Photosynthesis is a highly regulated process that dynamically adapts itself in order to optimize the use of light and avoid potential damage to the photosynthetic apparatus (Pinto et al., 2020). The excess of absorbed energy is rapidly dissipated through non-radiative (non-photochemical quenching, NPQ) and radiative (i.e. Solar Induced Fluorescence, SIF) pathways (Porcar-Castell et al., 2014, Hendrickson et al., 2004). Since PQ, NPQ and SIF emission are in direct competition for the same excitation energy, the efficiency variations of each process affect the others (Baker, 2008). Therefore, SIF is closely related to the light harvesting process and responds promptly to rapid changes in photosynthesis (Krause and Weis, 1991; Baker, 2008; Papageorgiou and Govindjee, 2004). Unlike the other energy usage processes, SIF is an optical signal and its spectrum covers a wavelength range in the VIS-NIR. It is emitted between 640-850 nm and is characterized by two peaks around 685 nm and 740 nm, respectively (Meroni et al., 2009). As an optical signal, SIF can be opportunely detected by remote sensing techniques by means of sensors mounted on different platforms (e.g., ground-based measurements, drone, airborne and satellite). In this regard, I refer to the reviews of Mohammed et al., (2019) and Gupana et al., (2021) for the terrestrial and aquatic ecosystems, respectively.

SIF represents a well-established and non-invasive approach in ecosystems monitoring at several spatio-temporal scales. The Remote Sensing (RS) of SIF represents a consolidated technique, widely employed in water bodies monitoring since 1960s (Mohammed et al., 2019, Gower 2016). The Chl-a SIF emission from the oceanic phytoplankton is routinely used to quantify phytoplankton abundance or biomass and for primary production estimations (Falkowski and Kiefer, 1985, Hout and Babin, 2010; Maritorena et al., 2000). Furthermore, it can be also exploited to detect potentially harmful algal blooms (Ruddick et al., 2019). On

the other hand, the RS of SIF from terrestrial vegetation is more recent (Mohammed et al., 2019), but it is rapidly emerging as a novel and promising approach to measuring plant function (Porcar-Castell et al., 2014) due to its link to the photosynthesis.
The analysis of the RS SIF is more advantageous compared to more established methods relying on the remote sensed reflectance (Verrelst et al., 2015). Historically, the analysis of the reflectance spectrum provides estimations of plant biophysical parameters, such as the leaf area index (LAI), chlorophyll (Chl) content and the fraction of the solar light absorbed by the Chl-a pigments (fAPAR) (e.g. Baret et al., 2007; Malenovsky et al., 2009). Similarly, in the aquatic ecosystems, the remote sensing Reflectance (R_{rs}) is used to infer the abundance of phytoplankton (e.g. Smith et al., 1989, Brewin et al., 2017). Regardless of the ecosystem considered, this information can be exploited to deduce the plant photosynthetic rates (Meroni et al., 2009), but also to evaluate phytoplankton biomass and productivity (Hout and Babin, 2010, Deng et al., 2017). However, photosynthesis is an actively regulated process and its efficiency is highly variable. It adjusts according to the environmental conditions by altering or rearranging the Chl-a pigments without any detectable changes in reflectance spectrum (Meroni et al., 2009). In this sense, the SIF represents a more direct and reliable proxy for photosynthesis, because it is emitted by the photosynthetic machinery itself. In fact, the emission of light as SIF is in competition with photochemical conversion and therefore SIF may allow a more accurate productivity estimation and earlier stress detection than is possible from reflectance data alone (Campbell et al., 2008; Cesana et al., 2021; Mohammed et al., 2019; Deng et al. 2017; Porcar-Castell et al., 2014).

Nevertheless, SIF is a complicated process and its retrieval from proximal and remote sensed acquisitions is challenging. The SIF signal cannot be easily measured because it is very weak and superimposed onto the radiance reflected by the target (Maxwell and Johnson, 2000; Zhao et al., 2018). Thus, disentangling these two contributions is not trivial. Despite the above-mentioned difficulties, the last few decades have seen great strides in measurement techniques, retrieval algorithms, modelling of fluorescence-photosynthesis and radiative transfer process (Mohammed et al., 2019).

Concerning the terrestrial vegetation, most of the implemented strategies retrieve the SIF only in selected solar or telluric absorption bands and require hyperspectral measurements. The oxygen absorption bands A (O_2-A at 760 nm) and B (O_2-B at 687 nm) are typically exploited to detect far-red and red SIF (e.g., Celesti et al., 2018; Mazzoni et al., 2008; Meroni et al., 2010; Mazzoni et al., 2010; Mazzoni et al., 2012, Verhoef et al., 2018; Cogliati et al., 2015), whereas several Fraunhofer Lines in the far-red (740–755 nm) (e.g., Guanter et al., 2012; Joiner et al., 2014) and near-infrared (771 nm) wavelengths (e.g., Sun et al., 2018; Frankenberg et al., 2011) are examined to retrieve far-red SIF. In these features, indeed, the irradiance is strongly attenuated and the contribution of the SIF emission in the radiance is higher. The SIF retrieval in the far-red has been proved to be feasible from field spectroscopy (e.g. Damm et al., 2010; Guanter et al., 2013; Rossini et al., 2016, Chang et al., 2020), airborne imaging spectrometers (e.g. Colombo et al., 2018; Damm et al., 2014; Rascher et al., 2015), and satellite remote sensing measurements (e.g. Guanter et al., 2014; Sun et al., 2018). Conversely, the red SIF retrieval is more challenging because the signal observed from the

spectrometer is much lower due to the strong reabsorption occurring within the leaves and canopy. Specifically, the Chl-a's absorption and emission spectra overlap in the visible red spectral region, therefore, the SIF emitted in the red could be re absorbed in its path within the canopy. The reabsorption process is strictly linked to several biophysical parameters such as the amount of chlorophyll contained in the canopy (C_{ab}) along with the leaf shapes (canopy geometry) and area (LAI) that also influence the amount of incident light absorbed from the canopy (aPAR). Specifically, the denser and more developed the canopy is, the greater the red SIF attenuation. Nevertheless, recent studies demonstrate the possibility to evaluate this value at different scales (e.g. Fournier et al., 2012; Liu et al., 2017; Rossini et al., 2015). A combined analysis of red and far-red SIF would offer additional and more direct information about plant physiology. The SIF spectrum, is the result of the emissions of photosystems I and II (PSI and PSII) involved in the early stages of photosynthesis (Mohammed et al., 2019). The major contribution in the red is attributable to PSII, while the broad shoulder from 700 to 750 nm and the peak at 740 nm are due to both PSII and PSI. PSII is directly linked with the first reactions, therefore, the corresponding SIF in the red better follows the plant functional response to stress situations. Furthermore, as Verrelst, et al., (2015) suggests, even more information related to photosynthetic activity could be found in the full SIF emission spectrum than in the individual peaks. The SIF full spectrum allows to properly estimate the fluorescence yield and then to obtain more information about the photosynthetic efficiency. This concept motivated the recent development of novel retrieval algorithms, specifically designed to evaluate the entire SIF spectrum. Until now, the available methods have been mainly implemented and tested on ground level measurements (e.g. Cogliati et al., 2019, Zhao et al., 2014). Although the full SIF spectrum retrieval is feasible, its interpretation is currently under investigation. The processes of reabsorption and scattering, occurring in the leaf-canopy system, affect the Top Of Canopy (TOC) SIF magnitude and spectral shape (Guanter et al., 2014; Yang et al., 2020; Liu et al., 2019; Siegmann et al., 2021). Thus, the SIF yield evaluated at TOC level (i.e. SIF^{TOC}_{yield}) might not be completely representative for the actual photosynthetic efficiency (SIF_{yield}). This makes the physiological interpretation even more challenging.

Concerning aquatic ecosystems, the SIF emitted by the phytoplankton contained in natural waters is characterized by a predominant contribution in the visible red spectral region, with a peak around 685 nm (Gupana et al., 2021). Conversely, the far-red emission is completely diminished by the strong water absorption (Gupana et al., 2021). The main difficulty in the phytoplankton SIF retrieval is represented by the high variability of the water body and its constituent bio-optical properties (Tenjo et al., 2021). The substances contained in the water body, in addition to the phytoplankton, can affect the signal acquired by the sensor in the spectral region where the SIF emission occurs (Gilerson et al., 2007). Water bodies can be practically distinguished in: i) Case-1 waters, in which the optical properties are determined primarily by the phytoplankton and they are related to the Colored Dissolved Organic Matter (CDOM) and detritus degradation products and ii) Case-2 waters where the optical properties are significantly influenced by constituents (e.g. mineral particles, CDOM and microbubbles) whose concentrations do not covary with the

phytoplankton (Mobley et al., 2004; Morel and Prieur, 1977). Oceanic waters usually belong to Case-1, while inland and coastal waters are generally considered Case-2. Lakes, which belong to Case-2, are typically characterized by a high degree of spatio-temporal variability of the inherent optical properties and related compounds (Cesana et al., 2021). In this framework, the phytoplankton SIF could be beneficial for monitoring water quality due to the dynamic nature of such environments and their often intensively used shores (Gupana et al., 2021). The Fluorescence Line Height (FLH) algorithm, developed by Gower (1980), is still frequently used to retrieve the SIF from the water leaving radiance. The FLH method combines three spectral bands opportunely selected in the red and far-red obtaining a proxy that correlates well with the actual SIF amplitude (Gilerson et al., 2007). The accuracy mainly depends on the position of the wavebands and the signal-to-noise of the sensor used (Tenjo et al., 2021). At satellite scale, the FLH is routinely applied on measurements acquired by ocean color sensors such as MODIS and MERIS/OLCI. In Case-1 water, the proxy obtained is linked to the SIF signal and the Chl-a concentration, hence the SIF can be used as a proxy for the phytoplankton biomass. Conversely, in more optically complex waters, the few wavebands available, along with the medium spectral resolution, impair the SIF retrieval (Gilerson et al., 2007). For this reason, the practical use of the SIF generated from inland and coastal waters is still under-exploited (Gupana et al., 2021). Nevertheless, hyperspectral measurements could potentially facilitate the decoupling of the different optical components signal contributions, improving the SIF retrieval. Obtaining a valid SIF proxy for optically complex water will help inland water monitoring and understanding. Furthermore, it could be used to estimate the phytoplankton primary production (PP) also in these environments.

Understanding the link between the SIF full spectrum and photosynthesis in terrestrial vegetation and improving the SIF retrieval for primary production estimation from inland waters are fundamental, specifically in view of the upcoming FLEX (FLuorescence EXplorer) satellite mission of the European Space Agency (ESA), the launch of which is planned for 2025. FLEX is the first satellite mission specifically designed for terrestrial SIF measurements (Drusch et al., 2016). It will provide measurements at a spectral and spatial resolution enabling the retrieval and interpretation of the full chlorophyll fluorescence spectrum emitted by terrestrial vegetation (Drusch et al., 2016). Although FLEX is mainly directed toward over land applications the acquired hyperspectral measurements could be used for scientific applications in coastal areas and inland waters, supporting the interpretation of Sentinel-3 observations (Drusch et al., 2016).

1.1 Objectives

SIF represents a promising tool to monitor both terrestrial and aquatic vegetation. However, as previously mentioned, its retrieval and interpretation are challenging and several complexities have been highlighted depending on the ecosystem considered. At present, in terrestrial vegetation, a novel generation of algorithms allow to obtain the TOC SIF full spectrum (e.g. Zhao et al., 2018; Liu et al., 2015; Cogliati et al., 2019; Verhoef et al, 2018; Celesti et al., 2018). Nevertheless, the possibility to gain more information about plant functioning and productivity from the SIF full spectrum, with respect to the traditional approaches (i.e. SIF evaluated in the O_2 bands and reflectance-based indices) need to be assessed. Current retrieval methods evaluate the TOC SIF, but the obtained spectrum is severely affected by distortions due to the SIF reabsorption process, occurring within the leaf-canopy system. This phenomenon further complicates the link between SIF and photosynthesis. Therefore, correcting the TOC SIF spectrum for the reabsorption could improve the evaluation of the SIF yield. On the other hand, aquatic ecosystems are very dynamic and the signal acquired by proximal and remote sensors is strongly influenced by the water body constituents. Specifically, Case-2 waters are currently under-studied because of the strong presence of optically active pigments, other than phytoplankton, that complicates the SIF retrieval. The advent of VIS-NIR hyperspectral observations could potentially help to disentangle biotic (particle bearing Chl-a pigments) and abiotic (suspended and dissolved matter) contributions in the signal acquired by the sensor. This will contribute to improving the SIF retrieval algorithms, obtaining a valid proxy for phytoplankton production estimation in optically complex waters.

In this context, the aim of my research consists in developing novel vegetation monitoring strategies, based on the SIF signal, exploiting hyperspectral field measurements and radiative transfer models at ground level. The specific objectives of this research can be summarised as follows:

- characterizing the SIF full spectrum behavior at seasonal/diurnal scale in comparison to SIF evaluated in selected bands and reflectance-based indices used as proxies for different biophysical parameters (e.g, canopy chlorophyll content, leaf area index, fraction of light absorbed by the canopy);
- investigating whether the SIF^{TOC}_{yield} approximates the SIF yield evaluated on the spectrum corrected for the canopy reabsorption (SIF_{yield}). Developing methodology to correct the TOC SIF spectrum for reabsorption toward the estimation of the SIF_{yield};
- improving the phytoplankton diurnal monitoring exploiting hyperspectral measurements and a novel FLH parametrization specifically developed for clear lake waters. Testing several phytoplankton primary production models exploiting hypersperctral measurements and *in-situ* values in clear lake waters.

1.2 Thesis Outlines

The objectives of the thesis are investigated in dedicated Chapters. Each Chapter is organized as a scientific paper in which the theoretical background, the material and methods, the results, the discussion and conclusions are presented.

In Chapter 2, novel SIF metrics defined on the full spectrum were compared with the SIF evaluated in the two oxygen absorption bands (i.e. O_2-A and O_2-B) and to opportunely selected reflectance-based indices. The analyses were carried out seasonally and diurnally to better characterize the SIF full spectrum behaviour at different temporal scales. The experimental SIF full spectra used were retrieved from hyperspectral measurements collected on the Italian site (Tuscany) of the ESA funded FLEXSense campaign, carried out in the 2018. In this framework, the spectral signal was acquired on three different crops, namely forage, alfalfa and corn. Therefore, the SIF's metrics potential variability due to the crop considered was also investigated.

Chapter 3 focuses on the impact that reabsorption has on the TOC SIF spectral shape and magnitude. A preliminary study was carried out comparing the TOC SIF to the fluorescence generated at the photosystems level and therefore not affected by the reabsorption. The analysis were performed on a synthetic dataset opportunely generated coupling the MODTRAN5 and SCOPE radiative transfer models (RT). The reabsorption impact on the TOC SIF was evaluated in different illumination conditions, canopy development stages and plant physiological state exploiting the SIF metrics defined in Chapter 2. Furthermore, the fluorescence yield evaluated at TOC scale (SIF^{TOC}_{yield}) was compared to the corresponding values calculated on the photosystems level SIF spectra (SIF_{yield}). The same synthetic dataset was exploited to develop strategies able to correct the TOC SIF for the reabsorption process (SIF^{RC}). The first approach investigated was the parametric method in which the SIF^{RC} spectral shape was parametrized combining the SIF wavelengths emission with constant values, TOC SIF metrics and regression coefficients. The second approach exploits a Fourier based method coupled with a supervised Machine Learning (ML) algorithm. It was developed in collaboration with the Laboratory of Advanced Bio-spectroscopy (Physics Department "G. Occhialini") of the University of Milano-Bicocca. The methods' accuracy and robustness were evaluated on the synthetic dataset and then they were applied on the experimental spectral measurements investigated in Chapter 2.

The inland waters' objective is addressed in Chapter 4. All the analysis described were carried out on experimental measurements and water samples collected on Lake Maggiore during the field campaign organized in 2019. Continuous hyperspectral measurements were exploited to implement the FLH algorithm for inland waters applications. The phytoplankton temporal evolution was analyzed exploiting remote sensed indices and metrics linked to the amount of light reaching the target (E_{PAR}), the chlorophyll-a concentration ($[\text{Chl-a}]_{OC4}$) and the fluorescence emission proxy (SIF_{FLH}). Furthermore, the relations between the remote sensed quantities and the laboratory analysis outcomes were employed to develop and test several phytoplankton primary productivity (PP) models. The content of this chapter is published on *Sensors* as: Cesana, I., Bresciani, M., Cogliati, S., Giardino, C.,

Gupana, R., Manca, D., Santabarbara, S., Pinardi, M., Austoni, M., Lami, A. and Colombo, R., 2021. Preliminary Investigation on Phytoplankton Dynamics and Primary Production Models in an Oligotrophic Lake from Remote Sensing Measurements. *Sensors*, 21(15), p.5072.

Chapter 5 is dedicated to the conclusions in which the main findings of the PhD research, along with the future perspective, are summarized.

Chapter 2

Characterizing the SIF full spectrum at Seasonal/Diurnal scales for agricultural crops

2.1 Theoretical background

Indices evaluated from the remote sensed reflectance have been historically employed to characterize and follow biophysical and phenology changes occurring in terrestrial vegetation at different spatio-temporal scales (Grace et al., 2007, Zarco-Tejada et al., 2013, Pettorelli et al., 2005). Nevertheless, the analysis of the remote sensed SIF emitted by the Chl-a pigments is emerged as a novel and promising strategy to monitor terrestrial vegetation dynamics (Mohammed et al., 2019). SIF offers more advantages compared to the reflectance-based approach, because this signal is originating uniquely from the vegetation and it is more related to plant physiology (Verrelst, Rivera, et al., 2015) inasmuch as strongly linked to the photosynthesis process. Most of the reflectance indices have been implemented in order to minimize the variability linked to the solar-target-sensor geometry, along with the BRDF effects, and maximize the relation to the biophysical parameters, such as the canopy chlorophyll content and structure, leaf area and composition and fraction of solar light absorbed by the Chl pigments. Therefore, reflectance indices are mainly characterized by slow changes in the temporal domain (Rautiainen et al., 2010; Verrelst et al., 2012). SIF is also driven by the same biophysical parameters influencing the reflectance, but in addition, it also depends on physiological fluctuations occurring during the day, therefore it shows a diurnal dynamic (Amoros-Lopez et al., 2008; Zarco-Tejada et al., 2013).

As previously introduced in Chapter 1, the SIF retrieval is challenging regardless the scale considered. Moreover, the strategies currently developed requires hyperspectral

measurements and retrieve mainly the SIF in the far-red. Nevertheless, in the last years, a new generation of SIF retrieval approaches have been developed to evaluate the fluorescence emission between 640-780 nm from hyperspectral ground measurements (e.g., Zhao et al., 2014; Zhao et al., 2018; Liu et al., 2015; Cogliati et al., 2015; Verhoef et al., 2018; Celesti et al., 2018). In this Chapter, the SpecFit algorithm proposed by Cogliati et al., (2019) has been applied on the experimental measurements collected during the ESA-funded FLEXSense campaign carried out in the 2018. The algorithm retrieves the SIF and reflectance (R) full spectra between 670 and 780 nm starting from the downwelling (L_d) and upwelling (L_u) radiances. The SpecFit accuracy and robustness have been already tested on both synthetic and field measurements in Cogliati et al., (2019). In particular, the experimental dataset used to evaluate the SpecFit performances over different vegetations and illumination conditions is the same used in this work. On the experimental full spectra obtained, conventional and novel SIF metrics and reflectance-based indices have been defined.

The aim of the work described in this chapter is to investigate whether metrics defined on the SIF full spectrum are more informative with respect to the SIF evaluated in the two oxygen absorption bands. These novel metrics have been also compared to reflectance-based indices used as proxies for specific crop's biophysical parameters (i.e. fraction of light absorbed by the canopy, chlorophyll content, photosynthetic activity) to assess their behaviour at different temporal scales (i.e. seasonally and daily) and for different crop types.

2.2 Material and Methods

2.2.1 Field measurements over different crops

The field hyperspectral measurements analysed were acquired in an agricultural area placed in central Italy, in the province of Grosseto (Tuscany, Italy), during the ESA-funded FLEXSense campaign carried out in the 2018. Hyperspectral incoming (L_d) and upwelling (L_u) radiances were systematically collected over three different targets: forage, alfalfa and corn. The forage was characterized by several growing stages, from bare soil up to dense canopy (21^{st} of February - 24^{th} of May). Alfalfa acquisitions were performed from the 25^{th} of May to the 12^{nd} of July over an already well-developed canopy. The only evident change in the crop chlorophyll amount was linked to a cut around the end of June. Hyperspectral measurements over corn started on 13^{th} July and concluded on 31^{th} August. Similarly, to the forage, also corn was characterized by a continuous growth during the acquisition time frame.

Field measurements were collected with the same instrument, a FLoX spectroradiometer (by JB-Hyperspectral Devices, Germany). FLoX is particularly recommended for continuous and long-lasting field measurements because requires only minimal user inputs and maintenance. Moreover, its SNR, spectral coverage and resolution are designed in agreement with the ESA FLEX mission instrument specification (Cogliati et al., 2019). FLoX

is equipped with two spectrometers:

- QEPro optic (Ocean Optics, USA) characterized by high spectral resolution in the wavelengths range in which the SIF emission occurs (650-800 nm), a FWHM of 0.30 nm and sampling interval (SSI) of 0.15 nm suitable for the top of canopy SIF retrieval;
- FLAME-s optic (Ocean Optics, USA) covering a spectral interval in the VIS-NIR with a FWHM of 1.7 nm and SSI of 0.6 nm.

Both the spectrometers were housed in a Peltier thermally regulated box, keeping the internal temperature lower than 25°C in order to reduce the dark current drift (Campbell et al. 2019). The FLoX is designed for high temporal frequency acquisitions of about 5 minutes time interval.

The experimental set-up is schematically shown in Figure 2.1. The sensors collecting the upwelling radiance were set at 1-3 meters height. Since the sensors were placed few meters above the target, atmospheric corrections have been not applied on the spectral measurements acquired. The experimental set-up were kept unchanged once installed on the selected crop. The optic pointing downward was characterized by a FOV of 25° and was placed perpendicular to the target (nadir viewing angle). Therefore, the average circular region acquired by the sensor was about 0.25 m^2. This optic consists into a bare fiber measuring the target upwelling radiance in which both the crops reflected and emitted contributions are present. Conversely, the incoming solar irradiance was collected using a cosine receptor. The spectral measurements were acquired in accordance with the protocol described in Cogliati et al., (2019). Each acquisition cycle consists into consecutive measurements of the L_d, followed by L_u arising from the target and an additional L_d. The signal is automatically optimized for each channel at the beginning of each measurement cycle (lasting in 5 minutes) and two associated dark spectra are collected as well. Metadata such as spectrometer temperature, detector temperature and humidity, GPS position and time are also stored in the SD memory of the system. A quality criterion was used to preliminary select measurements acquired in stable illumination conditions. Specifically, the quality flag was automatically evaluated by the FLoX as the difference between the two consecutive L_d acquired within the same collection cycle. Only measurements characterized by a difference lower than the 1% have been considered for the analysis.

FLoX data collected were processed applying spectral and radiometric calibration procedures to convert row digital counts to at-sensor calibrated radiances. Specifically, two open-source R packages (https://github.com/tommasojulitta) were used to obtain the incoming radiance to the surface (L_d), the TOC upwelling radiance (L_u) and the apparent reflectance spectra. These packages also estimate the SIF in the O_2-A and O_2-B bands using the SFM method (Meroni et al., 2009) and the PRI reflectance index. The radiances and the SIF evaluated are measured in $mWm^{-2}nm^{-1}sr^{-1}$.

All the analysis presented in this work were carried out on clear sky days selected by analyzing the diurnal trend of the incoming radiance at 750 nm. The seasonal analysis has been performed on spectra acquired in a 30 minutes interval centered in the solar noon.

Conversely, in the diurnal analysis, spectra collected between the 07:00 and 17:00 (UTC time) were used.

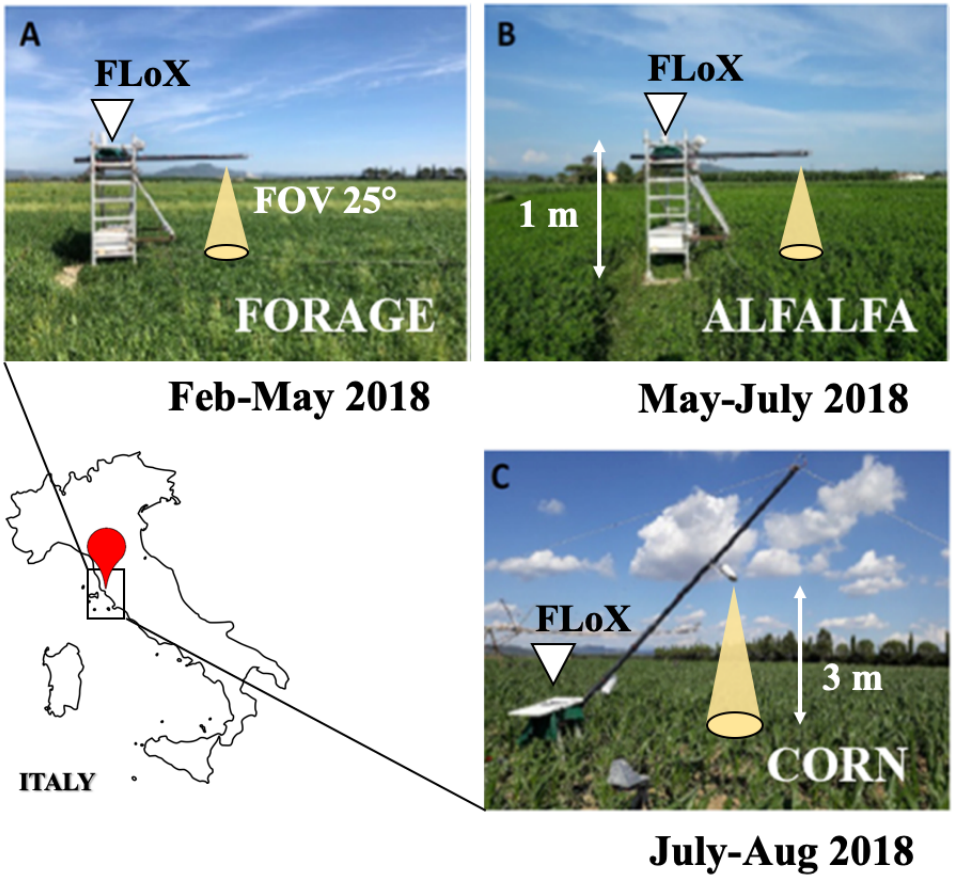

Figure 2.1: Location of the FLEXSense campaign site in Italy. Panels A, B and C show the three crops investigated (forage, alfalfa and corn), the experimental set-up and the acquisition time intervals.

2.2.2 Vegetation Reflectance Indices

The most widespread type of indices defined from remote acquisition are based on mathematical combinations of reflectance bands selected within visible and near infrared (NIR) spectral ranges. The main purpose of these indices is to synthetize the information contained in the reflectance bands, to infer the vegetation biophysical characteristics, minimizing soil, atmospheric and sun-target-sensor geometry sources of errors. In general, this kind of indices is able to describe more than 95% of the vegetation biophysical variability observed during the entire crop-growing season and they are closely related to the green leaf biomass (Gitelson et al., 2003a; Gitelson et al., 2003b; Jiang et al., 2006). In this work, canonical and novel reflectance-based indices have been investigated.

The $NDVI_{red-edge}$ (Normalized Difference Vegetation Index), introduced for the first time by Viña and Gitelson (2005), was used because it is well correlated to the fraction of light absorbed by the chlorophyll pigments (fAPAR), according to Li et al., (2020).

Most of the vegetation indices tend to be species specific and therefore are not robust when applied across different crops, characterized by different canopy architectures and leaf optical properties. Therefore, $CI_{red-edge}$ (Chlorophyll Index) was chosen because insensitive to the crop type and exhibits a low sensitivity to soil background effects. Unlike its canonical equation, defined by Gitelson et al., (2005), the $CI_{red-edge}$ used in this works exploits the NIR reflectance evaluated in spectral range corresponding to 770-780 nm instead of 790-800 nm, and the red-edge interval starting from 705 nm up to 715 nm.

To capture the changes in the photosynthetic activities, the Photochemical Reflectance Index (PRI) was evaluated (Gamon et al., 1992; Campbell et al., 2019). PRI is typically calculated as a normalized difference in the visible spectral range, utilizing reflectance at the 531 nm, as compared to a reference band at 570 nm (Campbell et al., 2019). Diurnal dynamics in PRI are mainly driven by changes in the xanthophyll cycle and illumination levels (Peñuelas et al., 1995; Gamon et al., 1997). Therefore, at diurnal scale, the PRI correlates well with the NPQ (Non-Photochemical Quenching). For this reason, it was compared to the SIF metrics. Conversely, the PRI response over seasonal timescales is more linked to changes in the photosynthetic pigment contents (e.g., chlorophyll and carotenoids), in the canopy structure and background signal. Under these circumstances, the relationship between PRI and the xanthophyll cycle is modified (Alonso et al., 2017). Furthermore, the PRI is strongly correlated with the $CI_{red-edge}$ because they are both mostly driven by the canopy chlorophyll, as reported by Campbell et al., (2019). Since, these two reflectance indices brought the same information at seasonal scale, only the $CI_{red-edge}$ was investigate.

The last index selected is the NIR_v (near-infrared reflectance vegetation index), introduced for the first time by Badgley et al., (2017), and originally used to evaluate the reflectance fraction attributable to vegetation in mixed pixel. This novel index is computed as the product between the NIR reflectance and the NDVI, using the spectral bands defined in Wu et al., (2020). According with Sellers and coworkers, cited in Badgley et al., (2017), NIR reflectance should be a more robust proxy for the fAPAR than NDVI, if it were only

possible to disentangle the vegetation signal from the variation in non-vegetated background reflectance (Badgley et al., 2017). However, since NDVI itself is used to estimate the vegetation cover, the NIR_v is able to capture the fraction of NIR reflectance attributable to the only vegetation because it better normalizes the variation in background reflectance than NDVI and NIR alone. Moreover, NIR_v correlates well to fluorescence because this latter parameter is insensitive to background contamination and is strongly linked to the GPP (Gross Primary Production), a fundamental proxy used to monitor the photosynthesis activity. However, these strong relationships are true only at large time scales (i.e. monthly), but require a modified NIR_v definition in the diurnal analysis framework. Therefore, the NIR_v index was investigated only at seasonal scale. In this case, the NIR_v correlates well with the SIF as long as the observed canopy is not stressed. Furthermore, the NIR_v only gives the information related to canopy structure but it does not provide the physiological information carried by the SIF.

Table 2.1 collects the indices used and the corresponding equations.

R-based spectral indices definition	References
$NDVI_{red-edge} = \frac{R_{NIR} - R_{red-edge}}{R_{NIR} + R_{red-edge}}$	Viña and Gitelson (2005)
$CI_{red-edge} = \left(\frac{R_{NIR}}{R_{red-edge}}\right) - 1$	Gitelson et al., (2005)
$PRI = \frac{R_{531} - R_{570}}{R_{531} + R_{570}}$	Gamon et al., (1992)
$NDVI = \frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}}$	Tucker et al., (1975)
$NIR_v = NDVI * R_{NIR}$	Bagley et al., (2017)

Table 2.1: Reflectance-based indices. The PRI index used have been evaluated automatically by the FLoX routine. The spectral bands used to estimate the other indices are 670-680 nm (red), 705-715 nm (red-edge) and 770-780 nm (far-red).

2.2.3 SIF spectrum metrics

The SIF metrics were defined on the full spectra obtained using the SpecFit algorithm proposed by Cogliati et al., (2019). This method allows to disentangle the SIF and R spectra from the total radiance evaluated at TOC level in the 670-780 nm spectral window. SIF and R are retrieved employing general-purpose mathematical functions that parametrize their shapes in the VIS-NIR spectral region. Specifically, the surface reflectance spectrum is modeled as a cubic-spline with 20 knots, while the SIF spectrum is approximated as a linear combination of two Lorentzian functions centered in the red and far-red, respectively. The algorithm combines the modeled SIF and R functions with the experimental L_d, obtaining the upwelling radiance (L_u). The L_u modeled in this way is iteratively compared to the corresponding experimental spectrum. In each iteration, the coefficients used to define the SIF and R spectral shapes are optimized until the best matching between the modeled and experimental L_u is reached. The Matlab source code of the present algorithm is available for download thought to the git repository https://gitlab.com/ltda/flox-specfit.

The SIF metrics defined on the SpecFit full spectrum and investigated hereinafter can be divided in two groups:

- conventional metrics, in which the SIF is evaluated in the two oxygen absorption bands (i.e. O_2-B at 687 nm and O_2-A at 760 nm);
- novel fluorescence metrics, specifically the SIF evaluated at the two peaks (i.e. in the red and far-red), the corresponding wavelength positions and the spectrally integrated SIF over 670-780 nm. The SIF at the peaks was selected because linked to the two photosystems PSI and PSII maximum emission. For the first time also, the corresponding wavelength positions were investigated at seasonal and diurnal scales. The spectrally integrated SIF, by definition, corresponds to the total amount of light emitted as fluorescence in the FOV acquired by the sensor.

Table 2.2 collects the SIF metrics previously defined, while Figure 2.2 shown the metrics over the full spectrum.

Fluorescence metrics	Description
SIF_{687}	SIF value at 687 nm (O_2-B band)
SIF_{760}	SIF value at 760 nm (O_2-A band)
SIF_{red}	SIF maximum emission in the red
$SIF_{far-red}$	SIF maximum emission in the far-red
λ_{red}	SIF_{red} wavelength position
$\lambda_{far-red}$	$SIF_{far-red}$ wavelength position
SIF_{INT}	spectrally integrated SIF over 670-780 nm

Table 2.2: Fluorescence metrics evaluated on the full spectrum.

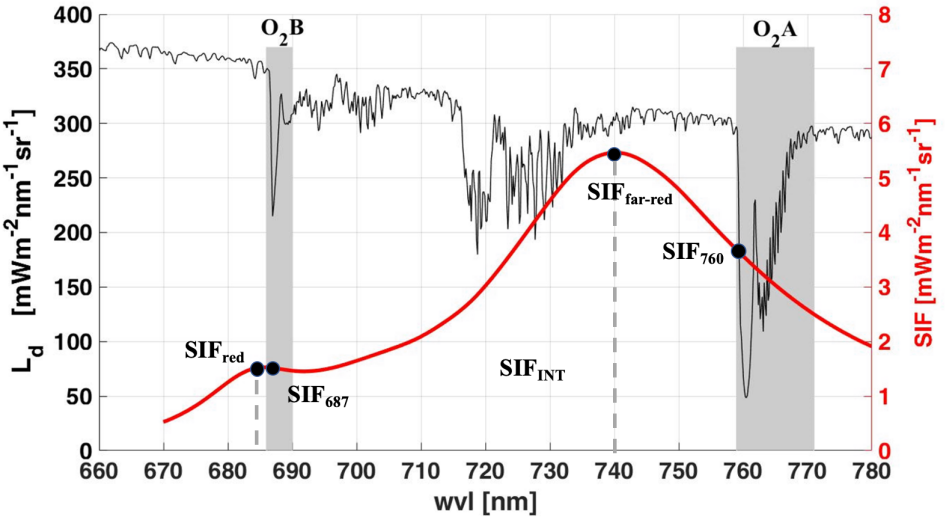

Figure 2.2: Fluorescence metrics defined on the full spectrum. Black line corresponds to the incoming radiance (L_d), red line to the top of canopy SIF, grey bands to the O_2 absorption bands, grey dashed lines to the wavelengths' positions corresponding to the peaks in the visible red and far-red, respectively. Black dots highlight the SIF metrics position on the full spectrum.

2.3 Results

2.3.1 Seasonal analysis

2.3.1.1 Reflectance Indices

Figure 2.3 collects the seasonal trends of the reflectance indices. The PAR (Figure 2.3A), defined as the integral of the incoming radiance L_d over the Photosynthetically Active Radiation spectral interval (400-700) nm, was used to track the illumination seasonal changes.

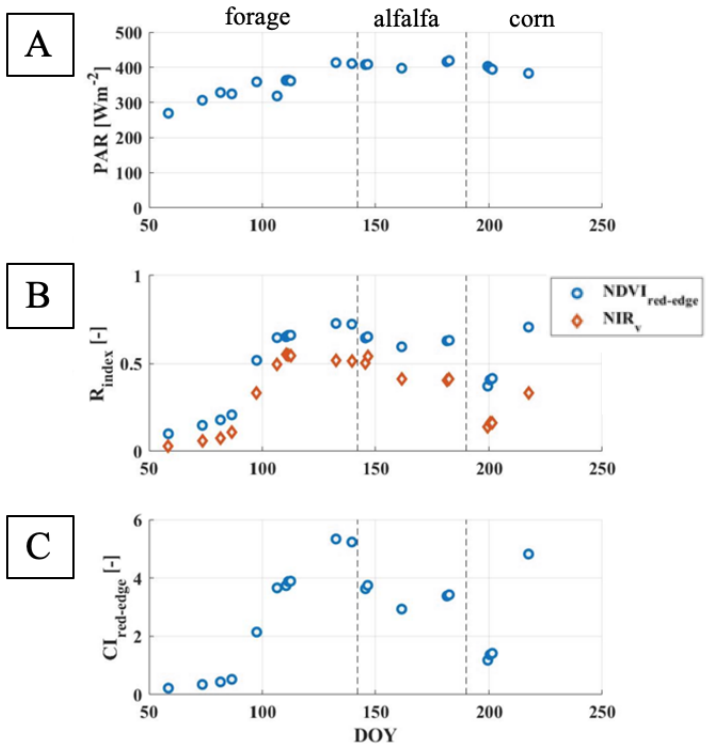

Figure 2.3: Seasonal trend corresponding to the reflectance-based indices. Panel A contains the PAR, Panel B shows the $NDVI_{red-edge}$ (in blue) and the NIR_v (in red), Panel C displays the $CI_{red-edge}$. The temporal information is showed on the x axes as Day of the Year (DOY). Data showed have been averaged over the 30 minutes interval around the solar noon. Dashed lines divide the different crop time series: forage, alfalfa and corn, respectively.

The indices investigated globally follow the PAR evolution in time. The only exception is represented by the first three measurements acquired on corn, in which the PAR slightly decreases in time, while the spectral indices increase. For all the crops investigated, the trends associated to $NDVI_{red-edge}$ (Figure 2.3B), NIR_v (Figure 2.3B) and $CI_{red-edge}$ (Figure 2.3C) agree to each other. The forage time series consider the largest number of clear sky days (10 days) and shows an initial growth, then followed by a saturation of the indices' values. Lower values correspond to acquisitions carried out around March and over a crop in the first growing stages. The increase in the indices is clearly observed starting from DOY 100 (Day Of the Year). This behaviour is probably due to an enhancement of the light reaching the target (PAR), but also to a gradual development of the crop, as suggested by the $CI_{red-edge}$ (linked to the canopy Chl content), NIR_v (linked to the vegetation contribution in the reflectance) and $NDVI_{red-edge}$ (linked to the fAPAR) trends. Considering alfalfa, only five clear sky days were selected. All the indices show a minimum around DOY 160. This trend is probably attributable to a change in the crop chlorophyll content, rather than to a variation in the illumination reaching the target because the PAR is almost constant in the temporal interval investigated. I remind that the alfalfa was cut around the end of June, therefore, the indices trends agree with this last observation. Finally, the corn times series is composed by four clear sky days. As observed for the alfalfa, the PAR show values always around 400 Wm^{-2}, but the indices investigated are characterized by a monotonic increase in time consistently with an crop under development.

2.3.1.2 SIF full spectrum metrics

The seasonal trends of the SIF metrics linked to the fluorescence magnitude are collected in Figure 2.4. As observed for the reflectance-based indices, also the SIF metrics values increase with the PAR, because the greater the amount of light reaching the target, the greater the fraction of light absorbed by the Chl-a pigments and then dissipated as fluorescence (in not stressed conditions). However, both the increase of the incoming light and the amount of chlorophyll content contribute to the growth of the SIF signal in the entire spectral region in which fluorescence occurs. Therefore, lower values are observed in correspondence of crops in the first stages of growing, such as forage (between DOY 50 and DOY 100) and corn (DOY around 200). Conversely, values greater than 2 $mWm^{-2}nm^{-1}sr^{-1}$ (SIF_{760} and $SIF_{far-red}$) and 100 $mWm^{-2}sr^{-1}$ (SIF_{INT}) were obtained over canopy well developed. The same trend is observed also in the the SIF metrics evaluated in visible red spectral region (Figure 2.4C). The metrics' values always lie below 2 $mWm^{-2}nm^{-1}sr^{-1}$, because the SIF signal is here strongly diminished by the canopy reabsorption effect. Therefore, the transition between sparse and dense vegetation, along with the corresponding chlorophyll content increases, is less clear compared to the trends observed in Figure 2.4A and Figure 2.4B.

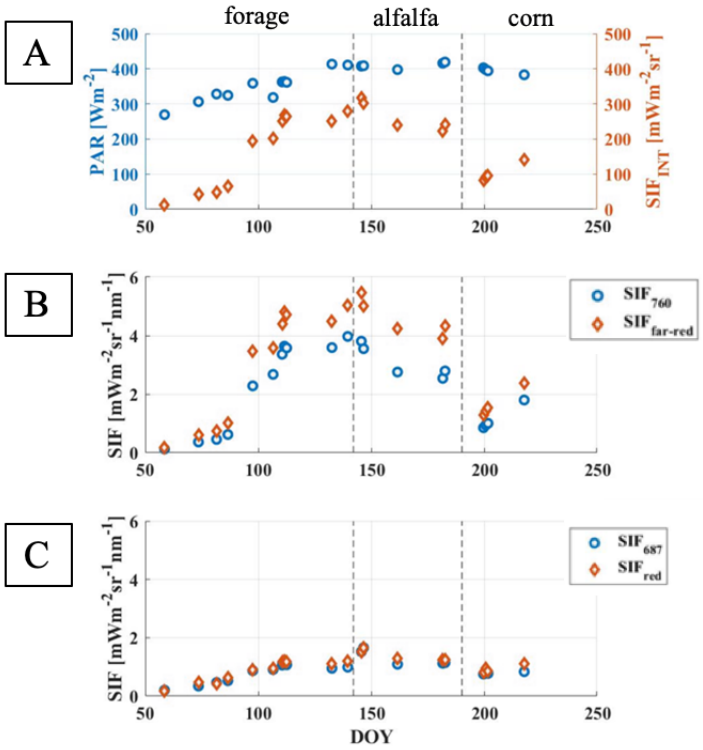

Figure 2.4: Seasonal trend corresponding to the SIF full spectrum metrics. Panel A contains the PAR (in blue) plotted together the SIF_{INT} (in red), Panel B shows metrics evaluated in the far-red spectral region (in blue SIF_{760} and in red $SIF_{far-red}$), Panel C displays the metrics defined in the visible red spectral region (in blue SIF_{687} and in red SIF_{red}). The temporal information is showed on the x axes as Day of the Year (DOY). Data showed were averaged over the 30 minutes interval around the solar noon. Dashed lines divide the different crop time series: forage, alfalfa and corn, respectively.

Focusing on the SIF_{760} and $SIF_{far-red}$ metrics, they show the same seasonal trend, regardless the crop considered. In particular, values referred to sparse canopies almost overlap, as observed at the beginning of forage and corn time series. The variability between the SIF evaluated at a specific wavelength (i.e. at 760 nm) and the fluorescence at the maximum emission peak in the far-red is clearer when crops already developed are considered. In this case, the increase of the chlorophyll content, and then of the canopy biomass, determines a growth of the SIF signal in the far-red spectral region due to the reabsorption and scattering processes.

Similarly to the fluorescence metrics defined in the far-red, the SIF_{687} and SIF_{red} show

the same temporal pattern regardless the crop considered. However, the differences in terms of absolute values are smoothed by the SIF reabsorption occurring in the visible red spectral region.

The qualitative correlations previously highlighted between the several SIF full spectrum metrics were quantitatively investigated in Table 2.3 and Figure 2.5. The linear regressions showed in Figure 2.5 and the corresponding statistics were evaluated on the values averaged around midday and for all the crops together.

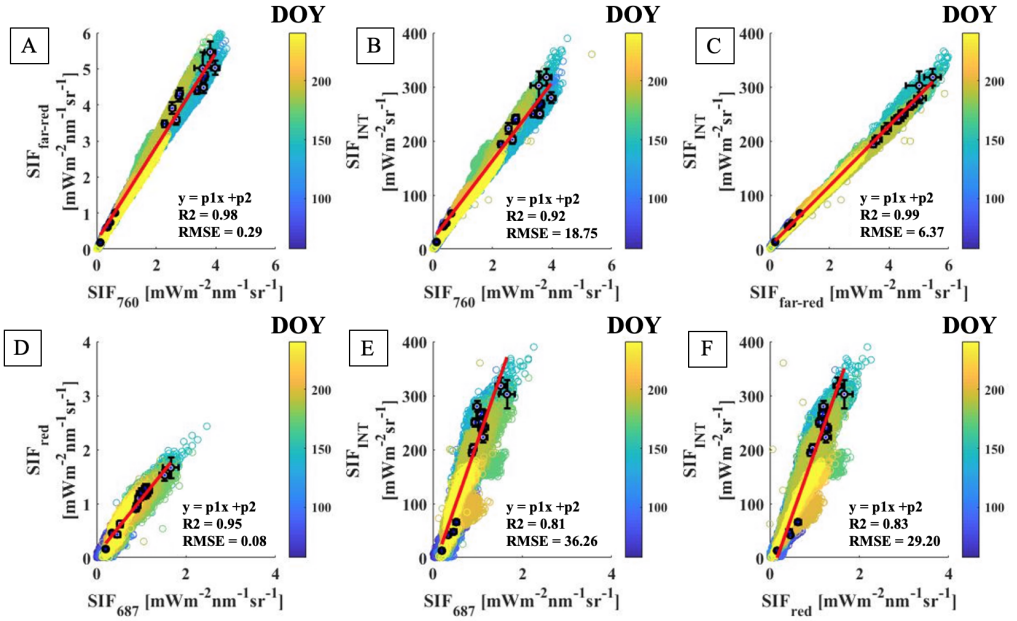

Figure 2.5: (top) Comparison between the SIF metrics evaluated in the far-red spectral range. (bottom) comparison between the metrics evaluated in the red spectral range. All the crops are displayed together; the colored dots represent the whole acquired time series. Color bar contains the temporal information. Black dots correspond to the spectral measurements averaged around the solar noon, while the error bars are the standard deviations. Red lines are the linear regressions performed on the averaged values.

CASE ID	Coefficients	Fitting values	Interval of confidence
A	**p1**	1.31	(1.19 ; 1.42)
	p2	0.23	(-0.11 ; 0.57)
B	**p1**	72.44	(64.78 ; 80.10)
	p2	18.51	(-3.16 ; 40.19)
C	**p1**	55.64	(53.66 ; 57.61)
	p2	5.36	(-2.34 ; 13.06)
D	**p1**	1.03	(0.92 ; 1.14)
	p2	0.06	(-0.06 ; 0.17)
E	**p1**	237.50	(186.40 ; 288.50)
	p2	-22.78	(-74.46 ; 28.90)
F	**p1**	233.10	(193.8 ; 273.4)
	p2	-38.11	(-81.21 ; 5.00)

Table 2.3: Coefficients obtained from the linear regressions performed on the comparison between the SIF full spectrum metrics. The fits have been evaluated only on the values averaged around the solar noon. The Case ID corresponds to the several Panels collected in Figure 2.5.

Strong linear correlations were observed when the SIF evaluated in the two O_2 bands (O_2-A and O_2-B) are compared to the corresponding metrics estimated at the SIF maximum emission in the far-red and red. In particular, the regression line performed in the scatterplot between SIF_{760} and $SIF_{far-red}$ shows a slope (p1) equal to 1.31, R^2 of 0.98 and RMSE equal to 0.29 (Figure 2.5A). From the SIF_{687}-SIF_{red} comparison, a lower R^2 has been estimated (R^2=0.95), but the values obtained are closer to the 1:1 (not showed in the figure) with a p1 equal to 1.03 and RMSE of 0.08 (Figure 2.5D).

The SIF metrics were also compared to the SIF_{INT} values to verify whether the full spectrum carries different information compared to the SIF evaluated at selected wavelengths. Both the SIF_{760} and the $SIF_{far-red}$ linearly increase with the SIF_{INT} (Figure 2.5B and Figure 2.5C). However the $SIF_{far-red}$ is more correlated to the SIF_{INT} respect to the SIF_{760}, with a R^2= 0.99 and a RMSE = 6.37. Concerning the visible red spectral region, the link between SIF_{687}, SIF_{red} and SIF_{INT} is more complex to interpret. When sparse vegetation is considered (bottom left portion of the plot in Figure 2.5E and Figure 2.5F), the SIF_{INT} increase linearly with the other two metrics. For more developed canopies (top right portion of the plot in Figure 2.5E and Figure 2.5F), the SIF_{687} and SIF_{red} values are soften by the reabsorption process. Their growth is then less pronounced compared to the SIF_{INT} one. These behaviours determine a change in the slope between the SIF metrics at a specific wavelength and the SIF_{INT} when different canopy structure are considered (e.g. Figure 2.5F). Considering the whole time series, SIF_{687} and SIF_{red} are then less correlated to the SIF_{INT} respect to SIF_{760} and $SIF_{far-red}$. Specifically, the linear regressions performed on

the scatterplot in Figure 2.5E and Figure 2.5F, show R^2 equal to 0.81 (SIF_{687}) and 0.83 ($SIF_{far-red}$).

For the first time, the wavelengths position corresponding to the SIF maximum emission in the red and far-red were investigated (Figure 2.6). λ_{red} and $\lambda_{far-red}$ were compared to the $CI_{red-edge}$ spectral index to evaluate their behavior for increasing chlorophyll content values.

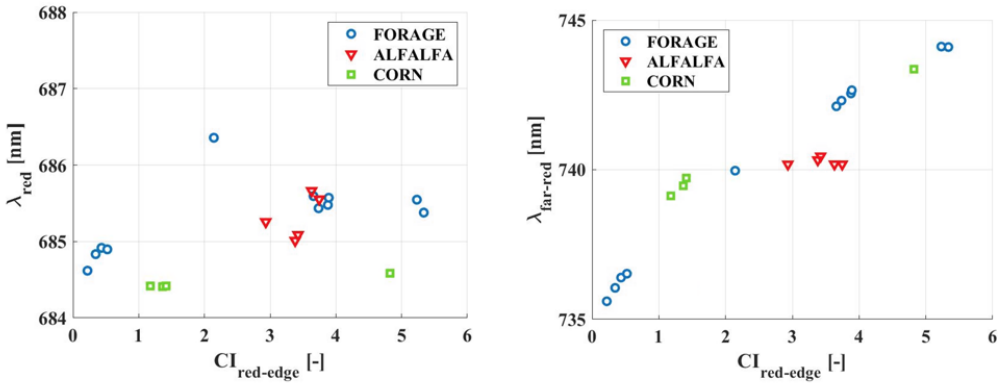

Figure 2.6: Experimental SIF peaks position trends for different $CI_{red-edge}$ values. (left) λ_{red}. (right) $\lambda_{far-red}$.Data showed have been averaged over the 30 minute interval around the solar noon. Blue dots correspond to forage, red triangles to alfalfa and green square to corn.

All the λ_{red} estimated lie between 684 nm and 687 nm (Figure 2.6, left Panel). In the forage case (blue dots) a clear variability with the $CI_{red-edge}$ is observed. Specifically, an initial increase is then followed by constant λ_{red} values. The alfalfa peak positions (red triangles) are grouped into a cluster probably because also the $CI_{red-edge}$ vary only between 3 and 4. Finally, the corn values (green squares) are constant, regardless the $CI_{red-edge}$ considered and show the lowest values compared to the other two crops. Conversely, the $\lambda_{far-red}$ undergoes to a shift toward longer wavelengths for increasing $CI_{red-edge}$, with values lying between 735 nm and 745 nm (Figure 2.6, right Panel). $CI_{red-edge}$ lower than 1 are linked to almost bare soil/sparse vegetation conditions. However, also not considering the points characterized by $CI_{red-edge}$ below 1, the overall $\lambda_{far-red}$ shift toward longer wavelengths is clear, specifically in the forage and corn cases. Conversely, all the far-red peak positions evaluated for alfalfa are constant and lie around 740 nm. Focusing on the forage time series, because composed by the higher numbers of days, the λ_{red} and $\lambda_{far-red}$ values mutually increase up to $CI_{red-edge}$ equal to 2. After

this point, the λ_{red} declines, while $\lambda_{far-red}$ goes on shift toward longer wavelengths. These different behaviours with the $CI_{red-edge}$ (and with the canopy chlorophyll content) are probably linked to the SIF reabsorption. In case of sparse vegetation, the temporal patter is dominated by the increase of the photosynthetically pigments concentration, while in dense canopies the reabsorption prevails and strongly affects the SIF spectral shape. Therefore, the emission in the visible red is uniformly soften and the λ_{red} are almost constant. Conversely, the SIF contributions in the far-red increases with the $CI_{red-edge}$ determining a $\lambda_{far-red}$ shift toward longer wavelengths.

2.3.1.3 SIF metrics and Reflectance indices relationships

Figure 2.7 shows the qualitative comparison between the reflectance-based indices and the SIF metrics in order to evaluate their mutual evolution at seasonal scale and to highlight links between couples of selected parameters. Each panel displays all the crops together in order to examine the potential inter-crop variability. SIF_{red} and $SIF_{far-red}$ are strongly correlated to SIF_{687} and SIF_{760}, as demonstrated in Figure 2.5. Therefore, only the two metrics evaluated in the O_2 bands were selected for the comparisons. The lines showed in Figure 2.7 help to follow the variables evolution in time and do not have a statistical meaning.

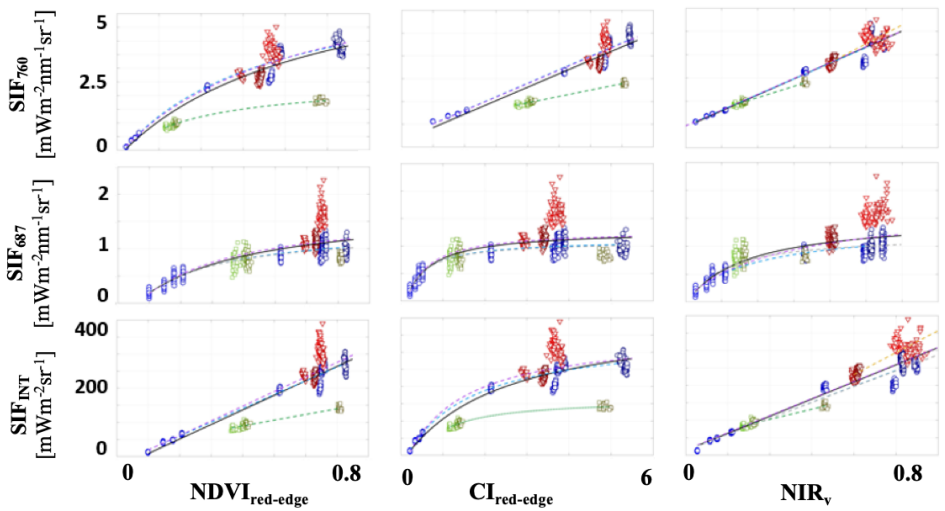

Figure 2.7: Comparison between selected SIF metrics and R indices. Blue dots correspond to forage, red triangles to alfalfa and green square to corn. Values displayed correspond to the values acquired around the solar noon (not averaged).

In all the cases investigated, the spectral indices and metrics evaluated on forage (blue dots) and alfalfa (red triangles) have comparable values. Corn (green squares) shows, in general, lower values except when the SIF_{687} row and the NIR_v column are considered. In these cases, there is no longer a difference between the three crops absolute values. The SIF_{687} is strongly affected by the reabsorption and this effect smooths the potential inter-crop variability, regardless the reflenctance index considered (Figure 2.7, second row). In forage and corn, the SIF_{687} shows a saturated pattern, in which the lower values correspond to the measurements acquired over crops not already developed. Concerning the NIR_v (Figure 2.7, third column), it has been selected because more linked to the signal emitted by the vegetation respect to the NDVI. For this reason, it has been expected to be correlated to the SIF. Linear correlations have been observed between the NIR_v, the SIF_{760} and SIF_{INT}, while the link with SIF_{687} is probably affected by the reabsorption. The $NDVI_{red-edge}$ was used as a proxy for the fAPAR and it increases with all the SIF metrics considered at seasonal scale (Figure 2.7, first column). In particular, a linear relation was observed between $NDVI_{red-edge}$ and SIF_{INT}. Finally, the $CI_{red-edge}$ was selected because linked to the canopy chlorophyll content and less affected by the vegetation types respect to the $NDVI_{red-edge}$ (Figure 2.7, second column). However, the $CI_{red-edge}$ shows an inter-crop variability, except in the SIF_{687} case. In general, this spectral index increase with the SIF metrics, but a linear relation has been observed only in the comparison with the SIF_{760}.

A statistical study was performed by means of a Principal Component Analysis (PCA). PCA is a highly suitable approach to explaining variations and potential links between variables in a multivariate dataset. The tool (toolbox 1.5) used in this framework has been developed and provided by the Chemometrics and QSAR Research Group of the University of Milano-Bicocca. Seven variables were investigated, specifically the six parameters showed in Figure 2.7 (i.e. $NDVI_{red-edge}$, $CI_{red-edge}$, NIR_v, SIF_{760}, SIF_{687}, SIF_{INT}) and the PAR. The cumulative variance, in percentage, highlighted that more than the 90% of the dataset information is contained in the first two components (see Figure 2.8). Figure 2.8 summarizes the PCA outcomes. Specifically, Figure 2.8A shows the scores on the first two components, while Figure 2.8B the loadings

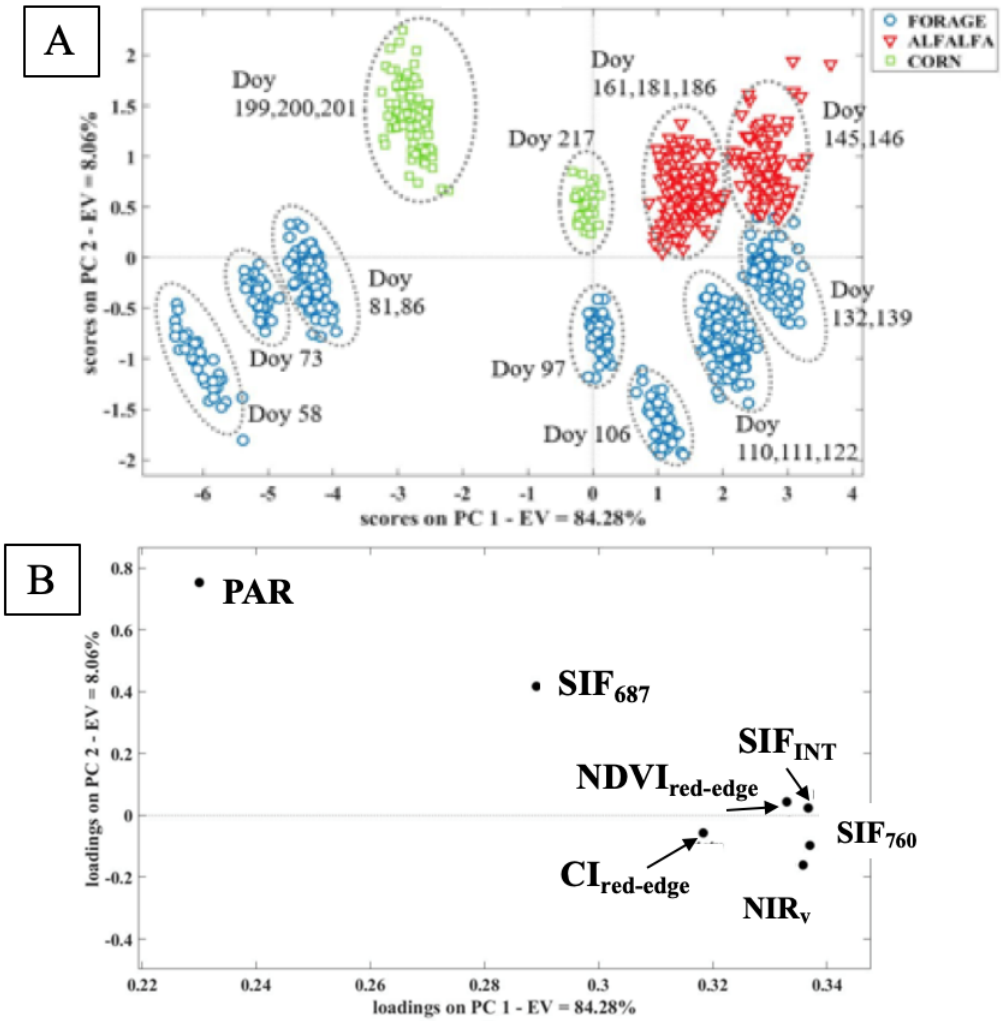

Figure 2.8: Principal component analysis performed at seasonal scale. (A) Scores on the first two principal components. Blue dots correspond to forage, red triangles to alfalfa and green square to corn. Values displayed correspond to the values acquired around the solar noon. (B) Loadings on the first two principal components. The seven variables on which the PCA has been performed are displayed here.

Figure 2.8A collects all the measurements acquired in the 30 minutes interval around the solar noon. The temporal information is highlighted as Day Of the Year (DOY), while the different symbols and colors used allow to discern between the three crops investigated. Each cluster, identified in Figure 2.8A, collects measurements characterized by similar behaviours in time. For example, corn values acquired in consecutive days (i.e. DOY 199, 200 and 201) fall in the same cluster because the variables selected for the PCA were almost constant in the time interval considered (see Figure 2.3 and Figure 2.4 as reference). Concerning the first principal component (PC1) it seems to be related to the canopy chlorophyll content. Following the forage time series (blue dots), the values acquired on sparse vegetation (i.e., DOY lower than 90) are characterized by negative PC1. Conversely, data linked to a denser canopy (i.e. measurements collected in DOY greater than 90) are characterized by positive PC1 values. A similar pattern is observed for the corn case (green square). Although the few clear sky days available (only four), a clear shift of the PC1 toward positive values is observed in the temporal interval considered. Concerning alfalfa (red triangles), all the measurements show PC1 values greater than 0. However, the lower PC1 value (close to zero) is observed for DOY 161, around the harvesting day. Again, this trend supports the hypothesis that the first principal component is linked to the chlorophyll amount, because the alfalfa was cut around DOY 160. Focusing on the second principal component (PC2), there is a clear separation between measurements acquired on forage, characterized by negative values, and the other two crops. Exploiting the temporal analysis reported in Figure 2.3 and Figure 2.4, the PAR is the only parameter that varies for forage, while is almost constant for alfalfa and corn. Therefore, the PC2 is probably linked to changes in the illumination reaching the target.

Finally, the loadings plot (Figure 2.8B) highlights the correlations between the seven variables investigated. All the points lie in the plot region characterized by positive PC1 values, therefore all the parameters considered are positively correlated to each other. This result is supported by the relations displayed in Figure 2.7. Furthermore, a cluster is clear on the right portion of the plot. The only two outliers are represented by the PAR and SIF_{687}. Points corresponding to $NDVI_{red-edge}$, NIR_v, $CI_{red-edge}$, SIF_{760} and SIF_{INT} are very close to each other, therefore they are strongly positively correlated. As observed in Figure 2.7, opportunely selected couples of these variables, lying in the above-mentioned cluster, are linearly related. Concerning the second principal component in Figure 2.8B, the SIF_{687} and SIF_{760} are negatively correlated. The SIF_{INT}, instead, lies between these two metrics, but it is closer to the fluorescence evaluated at 760 nm.

2.3.2 Diurnal analysis

As introduced in the Theoretical background, both the reflectance-based indices and the SIF signal are strongly influenced by the canopy's biophysical parameters such as the amount of solar light absorbed, the canopy chlorophyll content and geometry. To minimize the variability linked to these parameters, the diurnal analysis presented have been carried out selecting only the clear sky days characterized by similar illumination conditions (i.e. PAR), chlorophyll values (i.e $CI_{red-edge}$) and fAPAR (i.e. $NDVI_{red-edge}$). Exploiting the seasonal trend showed in Figure 2.3, the PAR is almost constant starting from DOY 100. Since the alfalfa was cut, measurements acquired before DOY 160 have been selected. In this time interval, the alfalfa $NDVI_{red-edge}$ and $CI_{red-edge}$ are similar to the corresponding values acquired on forage in DOY 110, 111 and 112. Conversely, the corn time series consists in only four clear sky days. To minimize source of error linked to the soil contributions, the only DOY acquired on an almost developed crop was selected (DOY 217). Of all the possible days highlighted, the diurnal analysis was carried out on DOY 112, DOY 145 and DOY 217, for forage, alfalfa and corn, respectively. The aim is to maximize in this way the inter-crop differences linked to the vegetation types (i.e., C3 for the forage and alfalfa; C4 for the corn) and canopy geometries (i.e. the corn have elongated leaves mainly arranged on the vertical direction, while forage and alfalfa develop horizontally). Furthermore, if the biophysical parameters are comparable, the differences between the reflectance-based indices and SIF metrics are mainly ascribable to the plant physiological status and photosynthetic activity.
The measurements displayed correspond to half-hour average. The error associated to each point represents the standard deviation.

2.3.2.1 Reflectance Indices

The reflectance indices selected for the diurnal comparison are the $NDVI_{red-edge}$ because linked to the fraction on light absorbed by the photosynthetic pigments (fAPAR), the $CI_{red-edge}$ used as proxy for the crop chlorophyll content (C_{ab}) and the PRI because linked to the photosynthetic activity and correlated to the NPQ at diurnal scale. Similarly to the seasonal analysis, the PAR was used to follow the evolution of the incoming light during the day. All the variables trends are displayed in Figure 2.9, in which the temporal information is shown on the x axes as fraction of the day (DAYfrac), while the symbols and colors used help to discern between the three crops investigated.

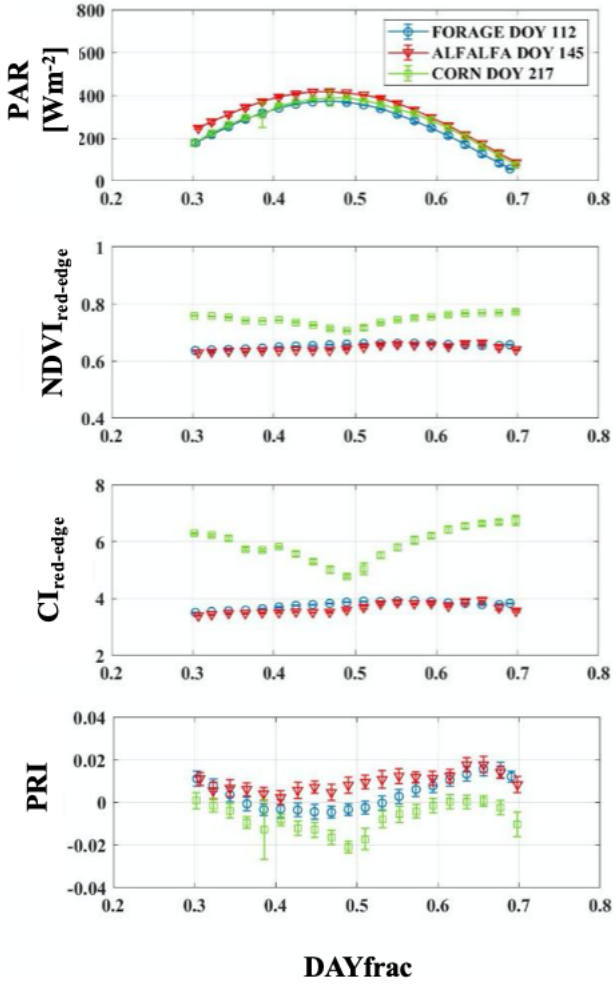

Figure 2.9: Diurnal trends of reflectance-based indices in three selected clear sky days (forage: DOY 112, alfalfa: DOY 145, corn: DOY 217). Blue dots correspond to forage, red triangles to alfalfa, green squares to corn. Data shown correspond to half-hour average. Error bars correspond to standard deviations. The temporal information is reported in fraction of the day, where 0.5 corresponds to the solar noon.

As expected, the PAR diurnal trends are typical of clear sky days, with values proportional to the cosine of the local solar zenith angle (SZA) and a maximum around the solar noon (i.e. DAYfrac of about 0.5). As previously introduced, the DOY selected show similar PAR values, therefore it is possible to state that the variables comparison was carried out under the same illumination conditions, assuming an atmospheric transmittance almost constant in the three days investigated.

Focusing on the reflectance-based indices, an inter-crop variability was observed. While forage and alfalfa always show similar (e.g. in the PRI) or overlapped values (such as in the $NDVI_{red-edge}$ and $CI_{red-edge}$), the corn spectral indices are characterized by different absolute values. This discrepancy is evident when the $CI_{red-edge}$ is considered: forage and alfalfa show almost constant values around 4, whereas in corn the points lie between 4 and 6. However, the described difference is attributable to the DOY chosen for corn. Due to the limited days available, DOY 217 is simply characterized by a chlorophyll content greater compared to the other two crops, that influences also the $NDVI_{red-edge}$ values. In terms of temporal trends, the forage and alfalfa $NDVI_{red-edge}$ and $CI_{red-edge}$ are constant throughout the day. This pattern is reasonable for the $CI_{red-edge}$, because it is linked to the C_{ab}, that does not vary at diurnal scale. The $NDVI_{red-edge}$ is used as a proxy for the fAPAR, that, by definition, is the ratio between the aPAR (i.e. light absorbed between 400-700 nm) and PAR. Under the same assumptions highlighted for the PAR, the aPAR is theoretically characterized by a similar diurnal trend, with a maximum around the solar noon. Therefore, their ratio is theoretically constant at diurnal scale, as observed for the forage and alfalfa. Conversely, the corn shows a depression around midday. Although the reflectance bands are opportunely combined to minimize the variability linked to the solar position respect to the target-sensor and canopy geometries, in the corn case a dependence to the leaves architecture is still evident. The clear depression observed around the solar noon is probably ascribable to the more complex canopy structure and row plantation of corn, respect to the forage and alfalfa cases.

Concerning the PRI, greater negative values occur around the solar noon for all the crops investigated. This temporal interval is characterized by the maximum of the PAR, that potentially provoke a stress condition due to the strong illumination. In this situation, the plants can activate protection mechanisms to dissipate the excess of light absorbed by non photochemical (NPQ) or radiative (SIF) pathways in order to avoid potential damages to the photosynthetic machinery. At diurnal scale, the PRI is related to the NPQ and the trend observed around the solar noon could be ascribable to the plant answer of a stressed condition. The differences in terms of PRI values between crops could be linked to the vegetation type considered. Forage and alfalfa are C3, while corn is a C4, where C3 and C4 represent two of the three possible photosynthetic pathways existing among terrestrial plants. Furthermore, these two photosynthetic pathways respond quite differently to changes in the temperature.

2.3.2.2 SIF full spectrum metrics

Figure 2.10 collects the diurnal trends evaluated for all the SIF metrics defined in Table 2.2. In addition, the PAR is shown in Figure 2.10A.

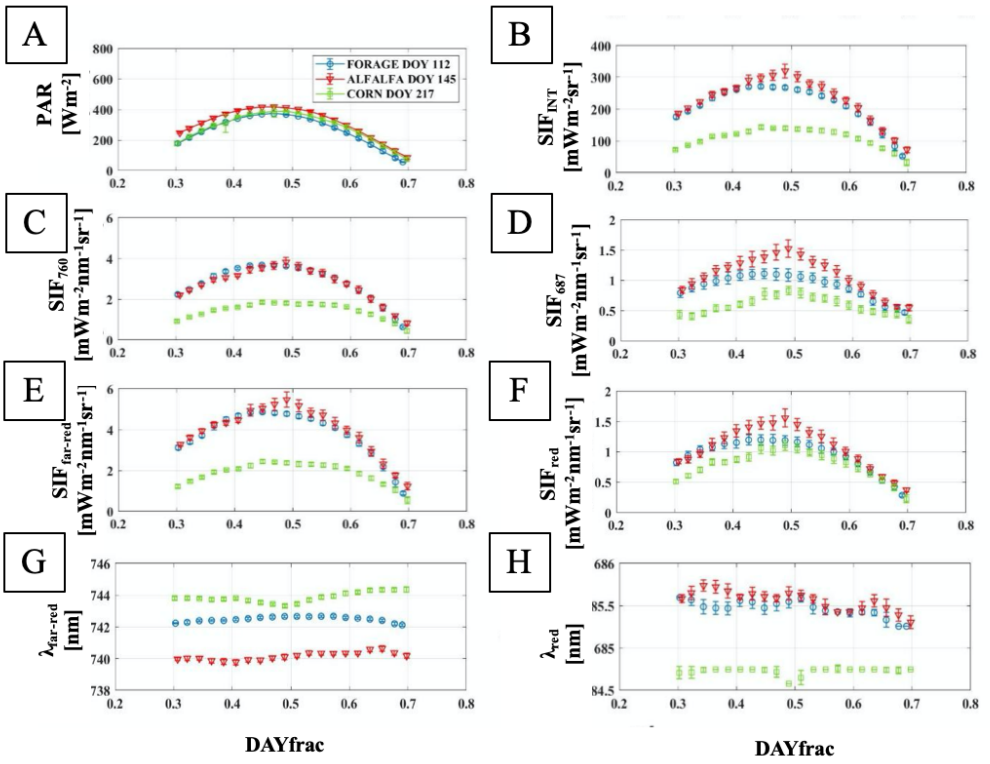

Figure 2.10: Diurnal trends of the SIF metrics in three selected clear sky days (forage: DOY 112, alfalfa: DOY 145, corn: DOY 217). Blue dots correspond to forage, red triangles to alfalfa, green squares to corn. Data shown correspond to half-hour average. Error bars correspond to the standard deviations. The temporal information is reported in fractions of the day, where 0.5 corresponds to the solar noon.

From a qualitative point of view, all the SIF metrics referred to the magnitude (Figure 2.10B-F), show the same diurnal pattern observed for the PAR. As expected, the SIF is strongly driven by the amount of light reaching the target. As observed at seasonal scale, the metrics evaluated in the O_2 bands and at the peaks are strongly correlated to each other. Similarly to the reflectance-based indices, the forage and alfalfa are characterized absolute

values that almost completely overlap when the metrics in the far-red (Figure 2.10C and Figure 2.10E) and the spectrally integrated fluorescence are considered (Figure 2.10B). Concerning the SIF_{687} (Figure 2.10D) and SIF_{red} (Figure 2.10F), they are characterized by different values in the central part of the day, between DAYfrac 0.4 (~ 10:00 UTC time) and DAYfrac 0.6 (~ 14:00 UTC time). Unlike the reflectance indices, the values associated to corn always lie below the other two crops. This behaviour could be linked to the PRI trend and then to the NPQ. In Figure 2.9 the PRI referred to corn shows the greater negative values: in this case the NPQ dissipation is stronger compared to what observed for forage and alfalfa. Since NPQ and SIF are dissipation processes occurring in competition, a predominant non-radiative dissipation could determine a decrease in the SIF efficiency and then to a lower fluorescence signal. Nevertheless, the strong differences between crops observed in the SIF metrics displayed in Figure 2.10B, C and E could be not only be due to the NPQ influence. More likely, also the plants canopy architecture potentially emphasizes the diversity between corn and the other two crops values. Finally, the peaks wavelength position was investigated at diurnal scale (Figure 2.10G and Figure 2.10H). In general, they show an overall constant trend. However, an inter-crop variability was observed. Concerning the λ_{red}, forage and alfalfa are characterized by the same values, lying around 685.5 nm. Conversely, corn shows a lower λ_{red} of about 684.6 nm. Focusing on $\lambda_{far-red}$, each crop is characterized by a different value. Specifically, 740 nm, 742 nm and 744 nm for alfalfa, forage and corn, respectively.

2.3.2.3 SIF metrics and Reflectance indices relationships

The links between SIF metrics and reflectance indices at diurnal scale were assessed statistically by means of a PCA. The variables selected and the results obtained are summarized in Figure 2.11. Seven parameters have been used for the analysis, namely PAR, $NDVI_{red-edge}$, $CI_{red-edge}$, PRI, SIF_{760}, SIF_{687} and SIF_{INT}. The data displayed correspond to all the measurements acquired during the three clear sky days selected (DOY). The different crops investigated are recognizable by means of the colors and symbols used. The cumulative percentage shows that the first two principal components contain more than the 81% of the dataset information.

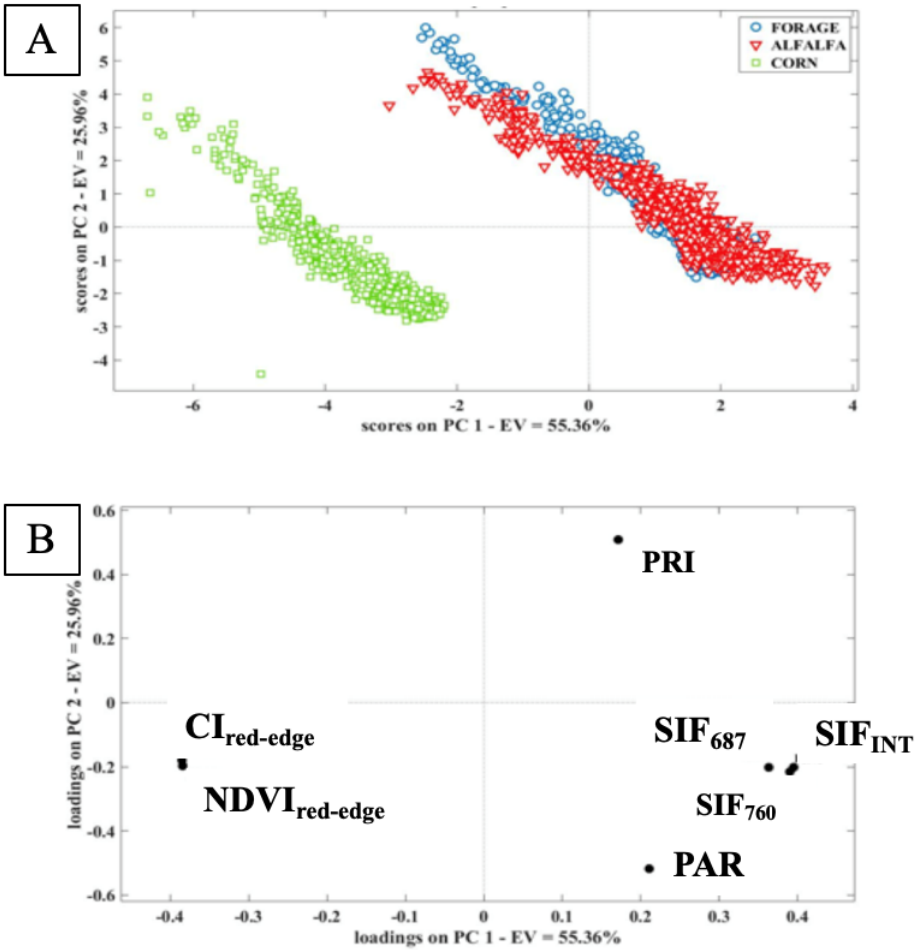

Figure 2.11: Principal component analysis performed at diurnal scale. (A) Scores on the first two principal components. Blue dots correspond to forage (DOY 112), red triangles to alfalfa (DOY 145) and green square to corn (DOY 217). Values displayed correspond to values acquired during the day. (B) Loadings on the first two principal components. The seven variables on which the PCA was performed are displayed here.

Focusing on the scores plot (Figure 2.11A), it is clear that forage and alfalfa follow the same pattern during the day. Specifically, all the measurements almost overlap and are placed in the same portion of the plot. A faint variability between forage and alfalfa is observed for values characterized by negative PC1 and PC2 values. Referring to Figure 2.9 and Figure 2.10, these two crops show different values around the solar noon, when the PRI and the SIF metrics (except SIF_{760}) are considered. Concerning the PC1, corn is characterized by all negative values on the PC1. Focusing on the second principal component, all the crops are characterized by PC2 values lying between -4 and 6. For this reason, the PC2 is probably linked to the PAR (as already occurs at seasonal scale), because similar in the three days selected. Conversely, the PC1 account the different chlorophyll content, canopy geometry and vegetation type.

The plot showing the loadings (Figure 2.11B), highlights the potential links between the seven variables investigated. The PAR and all the SIF metrics lie in the portion of the graph characterized by PC1 and PC2 positive values. Therefore, they are positively correlated on both the two principal components, as already observed in Figure 2.10. These metrics are linked to the PRI. In particular, the PRI and SIF are negatively correlated on the PC2 because they compete for the same PAR energy. Focusing on all the reflectance-based indices, $NDVI_{red-edge}$ and $CI_{red-edge}$ show a similar diurnal profile and therefore are characterized by the same PC1 and PC2 values. They are completely uncorrelated to the PRI. Finally, $NDVI_{red-edge}$ and the $CI_{red-edge}$ are negatively correlated to the PAR (and then to the SIF metrics) on the PC1, because they have been formulated to minimize the effect of the sun-target-sensor geometry.

2.4 Discussion

2.4.1 Seasonal analysis

The reflectance-based indices $NDVI_{red-edge}$, $CI_{red-edge}$ and NIR_v were preliminary investigated to characterize the changes of the crops at seasonal timescale (Figure 2.3). In particular, the analysis of their temporal evolution can be exploited to identify the different crops growing stages because they provide information about the canopy structure and chlorophyll content. According to Campbell et al., (2019), the $NDVI_{red-edge}$ and $CI_{red-edge}$ increase in time is distinctive of a crop under development (i.e. growing stage). In the mature phase, these indices saturate, afterward they tend to decrease during senescence. The forage time series is the only one in which both the growing and mature phases are observed (Figure 2.3B and Figure 2.3C). In particular, the transition between the two stages approximately occurs around DOY 100. This day has been then used to qualitatively distinguish between sparse vegetation (DOY lower 100) and medium-dense vegetation (DOY greater than 100). In this case, the reflectance indices evaluated show values greater than 0.4 ($NDVI_{red-edge}$ and NIR_v) and 2 ($CI_{red-edge}$). All the clear sky days selected for alfalfa show indices values similar to those obtained for a medium-dense

vegetation. Expect for the minimum observed around DOY 160, the small variability between the indices values is coherent with a crop in a mature phase. The faint decline observed is ascribable to a decrease in the canopy chlorophyll content, according to the information carried by the $CI_{red-edge}$ (used as proxy for the C_{ab}) and the NIR_v (linked to the signal attributable to the vegetation). Concerning corn, only four days compose the time series. Nevertheless, the monotonic increase of all the spectral indices suggests a crop in a growing phase. In particular, from a quantitative point of view, values around DOY 200 are typical of sparse vegetation. Conversely, the single measure acquired at DOY 217 was probably collected on the dense vegetation. The novel reflectance index NIR_v, introduced by Badgley et al., (2017), shows seasonal trends in agreement with those observed for the $NDVI_{red-edge}$ and $CI_{red-edge}$, regardless the crop considered. The information inferred by these indices, globally agree with the experimental dataset description carried out in Section 2.2.1. Specifically, forage and corn are characterized by a continuous growth, while alfalfa underwent to a cut around the end of June (i.e. ~ DOY 160). Globally, the seasonal trends of the reflectance indices agree with the ones reported in Campbell et al., (2019).

The SIF metrics evaluated on the full spectrum show seasonal trends in agreement to those observed for the reflectance indices, specifically when the SIF in the far-red spectral region ($SIF_{far-red}$ and SIF_{760}) and SIF_{INT} are considered (Figure 2.4A and B). Therefore, at seasonal scale, the fluorescence is mainly driven by the variations of the crop's biophysical parameters, such as the chlorophyll content and the fraction of light absorbed by the target (fAPAR) (Porcar-Castel et al., 2014). According to the crops characterization carried out, lower SIF values (< 2 $mWm^{-2}nm^{-1}sr^{-1}$ for the $SIF_{far-red}$ and SIF_{760} and < 150 $mWm^{-2}sr^{-1}$ for the SIF_{INT}) are observed over crops in a growing phase and then characterized by a sparse canopy. In this case, the differences between the SIF evaluated at 760 nm and the peak emission in the far-red values are minimal (Figure 2.4B). The seasonal trend observed for the SIF_{760} agree with what found in Campbell et al., (2019) and Yang et al., (2020). The discrepancy between $SIF_{far-red}$ and SIF_{760} grows when medium-dense vegetation is considered. This behaviour, clear at seasonal scale, is probably ascribable to the increase of the SIF reabsorption within the leaf-canopy system. The fluorescence absorption and emission spectra overlap in the visible red spectral region. Therefore, the signal emitted in the red is re absorbed and re emitted at longer wavelengths (i.e. in the far-red). Increasing the amount of pigments that absorb and the canopy density, the probability of the SIF signal to be absorbed increase. Considering the SIF full spectrum, the signal in the visible red is attenuated, while the contribution in the far-red wavelengths grows. The red quenching is clear in Figure 2.4C (Campbell et al., 2019). Exploiting the metrics evaluated at 687 nm (SIF_{687}) and at the peak (SIF_{red}), they show values always lying below 2 $mWm^{-2}nm^{-1}sr^{-1}$ and very close to each other. The reabsorption affects also the SIF spectral shape. It is clear when the SIF maximum emission wavelengths are compared to the $CI_{red-edge}$ (Figure 2.6). Using as example the forage time series, because composed by the greater number of clear sky days, the λ_{red} increases up to $CI_{red-edge}$ equal to 2 and then decreases to a constant value around 685.5 nm. The maximum λ_{red} is observed in correspondence to the transition between sparse and medium-dense canopy.

Conversely, $\lambda_{far-red}$ shows a monotonic growth with the $CI_{red-edge}$ in the two crops characterized by a continuous development, namely forage and corn. From Figure 2.4 is clear that the SIF metrics evaluated in the O_2-A and O_2-B are strictly correlated to the fluorescence values estimated at the maximum emission in the far-red and red, respectively. This qualitative outcome is supported by the quantitative analysis shown in Figure 2.5 and values collected in Table 2.3. In particular, $SIF_{far-red}$ is linearly correlated to the SIF_{760}. The regression line performed gives R^2 = 0.98, RMSE = 0.29, slope (p1) of 1.31 ($\pm$ 0.12) and intercept (p2) close to 0 (p2 = 0.23 $\pm$ 0.34). Similarly, the linear regression performed on the SIF_{red}-SIF_{687} comparison gives R^2 = 0.95, RMSE = 0.08, p1 = 1.03 $\pm$ 0.11 and p2 = 0.23 $\pm$ 0.15. Therefore, at seasonal scale, the SIF evaluated at the peaks emission bring the same information of SIF_{760} and SIF_{687}. The metrics estimated at a specific wavelength were also compared to the SIF_{INT} in order to investigate whether the full spectrum is more informative. SIF_{INT} is linear related to both the SIF_{760} and $SIF_{far-red}$, with R^2 equal to 0.92 and 0.99, respectively. These results are expected because the far-red contribute in the full spectrum prevails compared to the red, especially for dense canopy in which dominates the reabsorption. Similarly, the SIF_{687} and SIF_{red} are linearly correlated to the SIF_{INT}, with R^2 equal to 0.81 and 0.83, respectively. However, the values referred to a sparse canopy (i.e. grouped in bottom left part of the plot in Figure 2.5E and F) show a different slope compared to the metrics evaluated on a more developed canopy (top right of Figure 2.5E and F). During the growing phase, the SIF in the red mutually increase with the SIF_{INT}. Conversely, when the strong reabsorption occurs, the SIF in the red is quenched, while the SIF_{INT} continuously increases due to the far-red contributions to the full spectrum. Therefore, even though the SIF_{INT} is strongly correlated to the metrics evaluated in the far-red at seasonal scale, it also contains the visible red information, especially in the crop initial growing stages.

Finally, the reflectance indices were compared to the SIF metrics. Figure 2.7 collects the qualitatively comparison carried out. In agreement with Badgley et al., (2017), the NIR_v is strongly correlated to the fluorescence emitted at 760 nm. A similar relation was observed considering the SIF_{INT} because its value is dominated by the far-red emission, especially in dense canopy. Unlike the $NDVI_{red-edge}$ and the $CI_{red-edge}$, the NIR_v does not show a dependence to the crop considered. In most of the comparison investigated, the values estimated on corn are lower respect to the forage and alfalfa. This inter-crop variability, probably linked to the canopy geometry, is not observed in the SIF_{687} case. A prevalent saturated trend is clear when SIF_{687} is compared to the reflectance indices. The continuous increase in the fAPAR (represented by the $NDVI_{red-edge}$), canopy chlorophyll content ($CI_{red-edge}$) and signal attributable to the vegetation (NIR_v) is mutual with the SIF_{687} when sparse vegetation is considered. Then the fluorescence in the red is attenuated by reabsorption and an asymptotic growth is observed. Furthermore, the $NDVI_{red-edge}$ is linearly correlated to the SIF_{INT}, while the $CI_{red-edge}$ to the SIF_{760}. The relations previously highlighted are supported by the PCA outcomes (Figure 2.8). Focusing on the loadings plot (Figure 2.8), all the variables selected are positively correlated on the PC1, even though the PAR and SIF_{687} are outliers with respect to the cluster formed by

the $NDVI_{red-edge}$, $CI_{red-edge}$ and SIF metrics. The points corresponding to the NIR_v and SIF_{760} are very close to each other. Therefore, they have a similar seasonal behaviour and are positively related, in agreement with the corresponding scatterplot in Figure 2.7 (linear correlation). Similarly, the spectral indices-SIF metrics couples characterized by linear correlation show similar PC1 and PC2 values. In particular, $NDVI_{red-edge}$-SIF_{INT} has PC1 and PC2 greater than 0, while $CI_{red-edge}$-SIF_{760} shows positive PC1 and negative PC2 values. SIF_{687} and SIF_{760} are negatively correlated to the second principal component, while the SIF_{INT} lie between the two and is closer to the second one. Therefore, it is more correlated to the emission in the far-red, especially when reabsorption occurs.

2.4.2 Diurnal analysis

The diurnal analyses were carried out comparing days in which the three crops were characterized by similar illumination conditions and canopy chlorophyll content. The diurnal trends obtained for the reflectance indices and SIF metrics are summarized in Figure 2.9 and Figure 2.10, respectively.

For all the cases analysed, an inter-crop variability was observed. While forage and alfalfa show similar values, the corn is characterized by greater values when the reflectance indices are considered, lower in the SIF metrics linked to the fluorescence magnitude (Figure 2.10B-F). This outcome is probably ascribable to differences in the canopy structure and row plantation of corn, because the PAR is the same for all the three crops (see Figure 2.10A).

The canopy architecture plays a key role in the reflectance indices diurnal trend. It is clear when the $CI_{red-edge}$ is considered (Figure 2.9). As expected, forage and alfalfa values are constant during the day, while the corn shows a depression around midday. The reflectance indices spectral bands, are generally combined in order to minimize the effects linked to the acquisition geometry and canopy structure (case $NDVI_{red-edge}$ and $CI_{red-edge}$). However, when crops characterized by a predominant development in the vertical direction are considered, a dependence on the leaf architecture is still evident.

Concerning the PRI, it was investigated because able to capture changes in the plant photosynthetic activities and related to the NPQ at diurnal scale. NPQ processes are controlled photoprotective mechanisms which; once activated; strongly control the dynamics of photochemical efficiency (Alonso et al., 2017). With illumination conditions increasing and decreasing during a diurnal cycle; photoprotection mechanisms need to change accordingly (Alonso et al., 2017). The PRI evaluated for the three crops is shown in Figure 2.9. A decrease in the values around the solar noon was observed for all the crops, even though it is more evident in the corn. Concerning the corn, the trend observed corresponds with that found in Campbell et al., (2019). As reported in Xu et al., (2018), TOC PRI is sensitive to the canopy structure, but differences in the values could be also linked to the different vegetation types considered (C3 forage and alfalfa, C4 corn). According to Ehleringer et al., (2002), C3 and C4 plants responds differently to changes in environmental conditions, such as temperature. Regardless of the PRI absolute values, all the crops show a minimum

around the solar noon. Similar behaviours have been observed in Campbell et al., (2019), Chang et al., (2021)and Alonso et al., (2017). In all these cases, the PRI pattern of midday depression could be explained as an actuation of photoprotective mechanisms to prevent damage to the photosynthetic apparatus (Gamon and Bond, 2013).

Focusing on the SIF metrics linked to the fluorescence magnitude (Figure 2.10B-F), their temporal evolution follows the trend observed for the PAR (Figure 2.10A) (Campbell et al., 2019; Yang et al., 2020; Alonso et al., 2017; Yang et al., 2015). At diurnal scale, the SIF is mainly driven by the amount of light reaching the target in which greater values are observed around the solar noon (Chang et al., 2021). Similarly to the reflectance indices cases, an inter-crop variability is clear. Although corn is characterized by greater $CI_{red-edge}$ values, the corresponding SIF metrics always lie below the ones evaluated for forage and alfalfa. It is clearer when SIF_{760}, $SIF_{far-red}$ and SIF_{INT} are considered, in which forage and alfalfa values completely overlap. SIF of corn is also lower because of row structure and the fact that more soil is measured in contrast to forage and alfalfa. Furthermore, this outcome is also attributable to the different canopy geometries characterizing the three crops: planophile for forage and alfalfa, erectophile for the corn. Furthermore, forage and alfalfa show slightly different values in the region around the solar noon when the SIF_{687} and SIF_{red} are considered (Figure 2.10D and Figure 2.10F). A similar pattern was observed in the PRI and it could be linked to the NPQ. Under clear sky conditions, the solar noon is characterized by the maximum value of the PAR. In this framework, the plant could undergo stress due to the strong illumination, activating photoprotective mechanisms other than the SIF emission (i.e. NPQ). The corn outlier values respect to the other two crops could be due to the different crop response under stress situations that depend on the vegetation types (i.e. C3 or C4). Nevertheless, also between crops of the same type (i.e. forage and alfalfa), faint differences are observed (Figure 2.10D and Figure 2.10F). It must be remembered that the fluorescence evaluated in the visible red is mainly attributable to the PSII emission. As reported in Maxwell et al., (2000), PSII is accepted to be the most vulnerable part of the photosynthetic apparatus to light-induced damage and therefore it could be exploited to detect the first manifestation of stress in leaf. The differences between the forage and alfalfa values are also observed in the SIF_{INT} diurnal trend (Figure 2.10B). On the SIF metrics, a clear midday depression, as the PRI case, was not highlighted. However, focusing on the corn, the SIF_{760}, $SIF_{far-red}$ and SIF_{INT} diurnal trend is more flat in the time interval around the solar noon.

For the first time, also the SIF metrics linked to the maximum emission were investigated at diurnal scale. In both the cases analysed, $\lambda_{far-red}$ and λ_{red} positions are almost constant during the day. These outcomes are expected because the two metrics are more linked to changes in the SIF spectral shape due to the reabsorption (clearer at larger timescale). Conversely, at diurnal scale and for a single crop, the canopy chlorophyll content and geometry do not vary and the fluctuations observed in Figure 2.10G-H are more linked to the illumination geometry. Furthermore, the peak positions show an inter-crop variability. The canopy architecture mainly affects the λ_{red} values, while the different crops considered the $\lambda_{far-red}$.

The mutual relations between reflectance indices and SIF metrics were investigated by means of a statistical analysis (PCA) (Figure 2.11). Concerning the scores plot (Figure 2.11A), all the variables referred to the forage and alfalfa show the same diurnal profile, with data that almost overlaps in all the measurements considered. However, faint differences, in terms of absolute values, are observed for PC1 and PC2 lower than 0. In this case, it is possible to discern between the forage and alfalfa. Referring to the diurnal trends reported in Figure 2.9 and Figure 2.10, these two crops differ only in the interval around the solar noon and then under a stress situation due to the strong illumination. The conclusions reached analysing separately the reflectance indices and SIF metrics diurnal trends are immediately clear from the PCA outcome. The corn values show the same qualitative profile as the other two crops, with data decreasing for greater PC1 value. Although the PC2 values are similar, the corn is characterized by only negative PC1. For this reason, the first principal component probably accounts for the differences in terms of canopy structure, chlorophyll content but also physiological response to changes in the illumination occurring at diurnal scale. While score plot summarizes the overall temporal crops crops behaviours, the loadings highlight the relations between the variables considered in the PCA (Figure 2.11B). SIF metrics and PAR are characterized by positive values on the first principal component and negative on the second, therefore they are strongly correlated to each other. This result is expected because the SIF is mainly driven by the light reaching the target (Chang et al., 2021). Assuming the PRI as a proxy for the Non-Photochemical Quenching, NPQ and SIF compete for the same energy usage: PRI and SIF metrics are negative correlated on the second principal components. Points referred to $CI_{red-edge}$ and $NDVI_{red-edge}$ lying close in the loadings plot. Since they are mainly driven by crop biophysical parameters (such as the fluorescence), $CI_{red-edge}$, $NDVI_{red-edge}$ and SIF metrics show similar negative PC2 values. On the other hand, the SIF also depends on the PAR and the crop physiological status, therefore opposite PC1 values are observed. Finally, the PRI is completely uncorrelated to the other two reflectance indices (opposite PC1 and PC2).

2.5 Conclusions

For the first time, the SIF full spectrum metrics were analysed at seasonal and diurnal scale, also in comparison to reflectance indices that are assumed as proxies of the canopy biophysical variables.

From a qualitative point of view, reflectance and SIF metrics show comparable seasonal patterns. In particular, the PAR and chlorophyll content increase during the crop growth determines an increase in both reflectance-based indices and SIF metrics during the growing phase, while a saturated trend is reached when the crop enters into the mature stage. As expected, the biophysical and illumination changes are the main drivers for the reflectance and fluorescence values at seasonal scale. Nevertheless, when the spectral indices and the SIF metric are compared to each other, interesting relations are observed. In particular,

the SIF metric evaluated at 687 nm shows an initial linear growth then followed by a saturation trend. This pattern was observed for all the reflectance indices considered and it is attributable to the reabsorption affecting the SIF signal in the visible red. At seasonal scale, the $NDVI_{red-edge}$ is linearly correlated to the SIF_{INT}, while the $CI_{red-edge}$ to the SIF_{760}. In these cases, an inter-crop variability was observed. Specifically, corn is always characterized by lower values compared to the forage and alfalfa. The novel NIR_v index was investigated at seasonal scale. As expected, it is linearly correlated to both the SIF_{760} and the SIF_{INT}. Furthermore, it is not affected by crop type. All the relations between reflectance indices and SIF metrics are supported by the results obtained applying the PCA on the seasonal time series.

The fluorescence metrics defined on the full spectrum were investigated. Regardless of the timescale considered, the SIF evaluated at peak emissions are strongly correlated to the values estimated in the O_2 bands. Therefore, no clear additional information is obtained by the use of the $SIF_{far-red}$ and SIF_{red}. On the other hand, the SIF_{INT} gives more complete information respect to the fluorescence evaluated at a selected wavelength. By definition, the SIF is estimated over the spectral interval in which the fluorescence occurs, therefore the wavelength dependence is removed. It represents the amount of energy radiatively dissipated by the crop and then it is influenced by the variations in the SIF efficiency respect to the other light usage processes (NPQ and PQ). Therefore, it could better correlate to the GPP compared to the SIF evaluated at single wavelength. It must be pointed out that, at seasonal scale, the NIR_v is linked to the Gross Primary Production, and SIF_{INT} and NIR_v are linearly correlated. Furthermore, the SIF_{INT} could be used together with the aPAR, to evaluated the SIF_{yield}.

At diurnal scale, all the SIF metrics show the same evolution in time, but differences in terms of discrete values are observed between crops. In the far-red the variability is mainly linked to the canopy geometry, with forage and alfalfa values that overlap and the corn lying below. In the visible red, an additional shift between the forage and alfalfa is observed in the temporal interval around the solar noon. The SIF_{INT} shows both these behaviours, therefore it simultaneously accounts for the different behaviours observed separately in the red and far-red. All the SIF metrics are correlated to the PRI and highlighted peculiar trends around the solar noon, probably ascribable to photoprotection mechanisms actuated under strong illumination conditions. Since the SIF is linked to the energy dissipation by radiative pathways, while the PRI to the NQP, their combined analysis allows to better describe plants response under stress situations.

Finally, for the first time, the SIF emission peaks positions were investigated at seasonal and diurnal scales. In the first case, a shift toward longer wavelengths is observed for the $\lambda_{far-red}$. This trend is due to the reabsorption that affects both the SIF magnitude and spectral shape. Conversely, at diurnal scale, the λ_{red} and $\lambda_{far-red}$ are almost constant, but an inter-crop variability is clear.

Chapter 3

Correcting SIF for canopy reabsorption for quantum yield estimations

3.1 Theoretical background

Spectral measurements acquired in field are influenced by the light changes that naturally occurs daily and seasonally (Romero et al., 2020). These variations affect photosynthesis that is a process actively regulated by various mechanisms aimed to protect the plant from potential damage (Romero et al., 2020). Instantaneous fluctuations in photosynthesis efficiency can be inferred by means of the chlorophyll-emitted SIF, as discussed in Chapter 2. SIF contains information about plant physiology. However, the fluorescence signal measured at Top-Of-Canopy scale is significantly influenced by the reabsorption affecting both the SIF magnitude and its spectral shape. Hence, an uncertainty in the relationship between fluorescence and plant physiological state is introduced. SIF reabsorption occurs both within the leaf (Agati et al., 1993; Gitelson et al., 1998; Ramos and Lagorio, 2004; Cordón and Lagorio, 2006) and the canopy (Porcar-Castell et al., 2014; Liu et al., 2019; Romero et al., 2018), causing an enormous alteration in the spectral distribution of the fluorescence measured above the canopy itself (Romero et al., 2020). Reabsorption mainly affects the visible red spectral region connected to the PSII emission. PSII represents the part of the photosynthetic apparatus most vulnerable to light-induced damage, therefore its emission is often used to early highlight the stress in leaf (Maxwell and Johnson, 2000). Conversely, the SIF signal emitted in the far-red is scattered multiple times within the canopy leaves (Yang et al., 2020). The parameter SIF yield accounts for the amount of the solar light absorbed (aPAR) actually dissipated as fluorescence. It can be calculated

dividing the SIF, at a specific wavelength (i.e. oxygen absorption bands as reported in Campbell et al., 2019) or the SIF spectrally integrated between 650-800 nm, by the aPAR. However, due to the reabsorption and the scattering processes previously mentioned, the SIF yield evaluated at canopy level might be very different from the one estimated at the photosystems scale. The potential differences due to the scales considered might lead to an inaccurate physiological interpretation. Being able to correct the TOC SIF for reabsorption is pivotal to infer reliable physiological information.

In this Chapter, the impact of reabsorption on the SIF spectrum was qualititively investigated exploiting a synthetic dataset generated coupling MODTRAN (MODerate resolution atmospheric TRANsmission) and SCOPE (Soil-Canopy Observation of Photosynthesis and Energy) radiative transfer models. TOC and at-the-photosystems level SIF spectra were simulated for different illumination geometries (i.e. SZA), canopy chlorophyll content (obtained multiplying the C_{ab} by the Leaf Area Index, LAI) and fluorescence emission efficiencies (namely Fqe when it is given as input in SCOPE or SIF_{yield} when the SIF spectrally integrated is divided by the aPAR). These parameters represent only a subset of the whole key variables that drive the SIF emission, as highlighted in the global sensitivity analysis described in Verrelst et al., (2015). They were selected because the canopy biophysical parameters such as, C_{ab} and LAI, determine the amount of incident light absorbed by the canopy (aPAR) and then the amount of energy that can be potentially dissipated as fluorescence. They are connected to the reabsorption, because in denser canopies the probability of the SIF signal to be absorbed and scattered is greater compared to sparse vegetation. Furthermore, the aPAR is influenced by the sun-target position (SZA). Finally, the amount of light dissipated as fluorescence depends on the efficiency of this process (Fqe) that is intrinsically linked to the physiological state of the plant. In this framework, a specific attention was paid to the SIF emitted in the visible red spectral region because it is strongly affected by reabsorption and to the SIF yield evaluated at both TOC (SIF^{TOC}) and photosystems scales. Specifically, one of the aims of the work proposed is to assess whether the SIF_{yield} evaluated using the TOC fluorescence (i.e. SIF^{TOC}) approximates the SIF yield estimated exploiting the SIF corrected for the reabsorption and then the full spectrum emitted at the photosystems level.

Two methods aimed to retrieve the SIF spectra corrected for the reabsorption (SIF^{RC}) are here proposed. The first approach consists in a parametric method implemented on a subset of synthetic TOC hyperspectral measurements. In this case, the SIF^{RC} spectral shape was modeled as a sum of two Gaussian. The coefficients used to parametrize this mathematical function were obtained exploiting relations between the simulated spectra and variables potentially measurable in field, such as the SZA and the SIF at 760 nm. Similarly, a strategy to evaluate the aPAR was developed in which the coefficients used are linked to the SZA and the NIR_v. The aPAR represents a fundamental parameter in the SIF yield estimation. On the other hand, the second approach couples the Fourier based method with a supervised machine learning (ML) algorithm. Unlike the parametric method, the Fourier-ML algorithm retrieves the SIF at both canopy and photosystem levels, simultaneously with other biophysical variables, namely C_{ab}, LAI, aPAR and Fqe (equal

to the SIF_{yield} evaluated at the photosystems level). In this case, the simulated apparent reflectance (R_{app}) was used to train and test the algorithm. The above-mentioned approach was developed in collaboration with the Laboratory of Advanced Bio-spectroscopy (Physics Department "G. Occhialini") of the University of Milano-Bicocca. The Fourier-ML method development and description are topic of a paper and another PhD thesis currently not published. Therefore, I report here only a few of the meaningful results obtained and a brief description of the method.

The accuracy of these two methods were investigated on the synthetic dataset, then they were separately applied on the field hyperspectral measurements described in Chapter 2. Concerning the results obtained on the experimental dataset, only a qualitative comparison was possible between the parametric and the Fourier-ML approaches. Although the collaboration carried out, the numerical outcomes of the Fourier-ML method will be provided and available for further analysis only after the corresponding paper submission.

3.2 Material and Methods

3.2.1 Radiative Transfer simulations

The synthetic dataset was generated coupling the MODTRAN and SCOPE radiative transfer models (RT). MODTRAN5, developed by Berk et al., (2006), models the solar light propagation within the atmosphere. The surface-atmosphere coupling is based on the four-stream radiative transfer theory with the addition of the fluorescence flux (Cogliati et al., 2019; Verhoef et al., 2018; Verhoef and Bach, 2012; Verhoef and Bach, 2007). Conversely, the interaction occurring between the solar light and the photosystems complex, along with the leaf-canopy interactions, were simulated by SCOPE (v.1.73) (Van der Tol et al., 2014; Van der Tol et al., 2009). SCOPE consists of several RT equations, originally based on the SAIL scheme (Scattering by Arbitrary Inclined Leaves), in which the target surface is assumed to be homogeneous, infinitely extended and then processed with a 1-D turbid medium model. The SCOPE RT model simulates the fluorescence and reflectance spectra, the TOC radiance with a number of canopy biophysical, photosynthesis and energy balance parameters (Cogliati et al., 2019). In the recent versions released, the leaf RT are calculated with Fluspect (Vilfan et al., 2016) that provides not only the TOC SIF, but also the signal emitted at the photosystems level and then corrected for the reabsorption. In this work the SCOPE version 1.73 has been coupled to MODTRAN5.

Although SCOPE already considers the vegetation-atmosphere scheme, the atmospheric transfer functions (t_n) were first simulated with MODTRAN5 and then provided as input in SCOPE. Disentangling the atmosphere module from the vegetation one (i.e. SCOPE) is advantageous because determines a decrease in the overall computational time. In this way, MODTRAN5 runs only at the beginning of the routine, generating immediately all the t_n functions needed and not for each scenario simulated with SCOPE. The SZA is a fundamental variable in the t_n evaluation and represents the only parameter made to vary

in MODTRAN5. Therefore, only a single type of atmosphere was simulated. Although it represents an approximation that does not reflect the actual experimental conditions, this option was chosen because the aim of the work proposed is more oriented to evaluating the impact of the sun-target-geometry and the canopy biophysical parameters on the SIF signal, rather than the atmosphere composition variability. All the possible vegetation cases are simulated with SCOPE, combining the selected C_{ab}, LAI and Fqe values for a fixed SZA. The correct t_n given as input in SCOPE are opportunely selected by means of the SZA value. This codes architecture also allows to exploit the hyperspectral resolution, typical of the t_n spectral functions obtained from MODTRAN5 (0.1 cm^{-1}), to evaluate the radiances and the SIF spectra. The MODTRAN5-SCOPE coupling rationale implemented is similar to the one described in Cogliati et al., (2019). In particular, the simulated L_d and L_u have been obtained combining the MODTRAN5's t_n with the fluorescence and reflectance contributions generated by SCOPE as follow:

$$L_d = t_1 * \left(t_4 + \frac{t_5 + t_{12} * r_{sd}}{1 - t_3 * r_{dd}} \right) \tag{3.2.1}$$

$$L_u = t_1 * \left(t_4 * r_{so} + r_{do} * \frac{t_5 + t_{12} * r_{sd}}{1 - t_3 * r_{dd}} \right) + SIF_{so} \tag{3.2.2}$$

where the L_d and L_u are evaluated at TOC level. t_1 is the extraterrestrial solar radiance normalized by cosΘ, where Θ corresponds to the local Solar Zenit Angle (SZA). t_3 is the spherical albedo. t_4 and t_5 are the downward direct and diffuse transmittance, while t_{12} is the product of the downward direct transmittance and the spherical albedo. Concerning the canopy reflectance, r_{so} and r_{do} correspond to the direct and diffuse contributions in the viewing direction, while r_{sd} and r_{dd} account the direct and diffuse hemispherical terms, due to the surrounding. These terms are generated by SCOPE modeling the canopy reflectance by a four Bidirectional Reflectance Distribution Function (BRDF). SIF_{so}, also generated by SCOPE, corresponds to the directional TOC SIF. The R_{app} (apparent reflectance) has been obtained by the ratio between L_u and L_d, while R (reflectance) as the difference between R_{app} and the SIF_{so} normalized by the L_d:

$$R_{app} = \frac{L_u}{L_d} \tag{3.2.3}$$

$$R = R_{app} - \frac{SIF_{so}}{L_d} \tag{3.2.4}$$

Summarizing, for each variables' combination (C_{ab}, LAI, Fqe, SZA), the L_d, L_u, SIF TOC (SIF^{TOC}), SIF at the photosystems (SIF^{RC}), R_{app} and R spectra were simulated. Furthermore, also the aPAR represents one of the code's outcomes and was saved in the output storage. Finally, to match the experimental hyperspectral measurements, the simulated spectra were convoluted with the FLoX (QEPro) spectral resolution. The code architecture is shown in Figure 3.1.

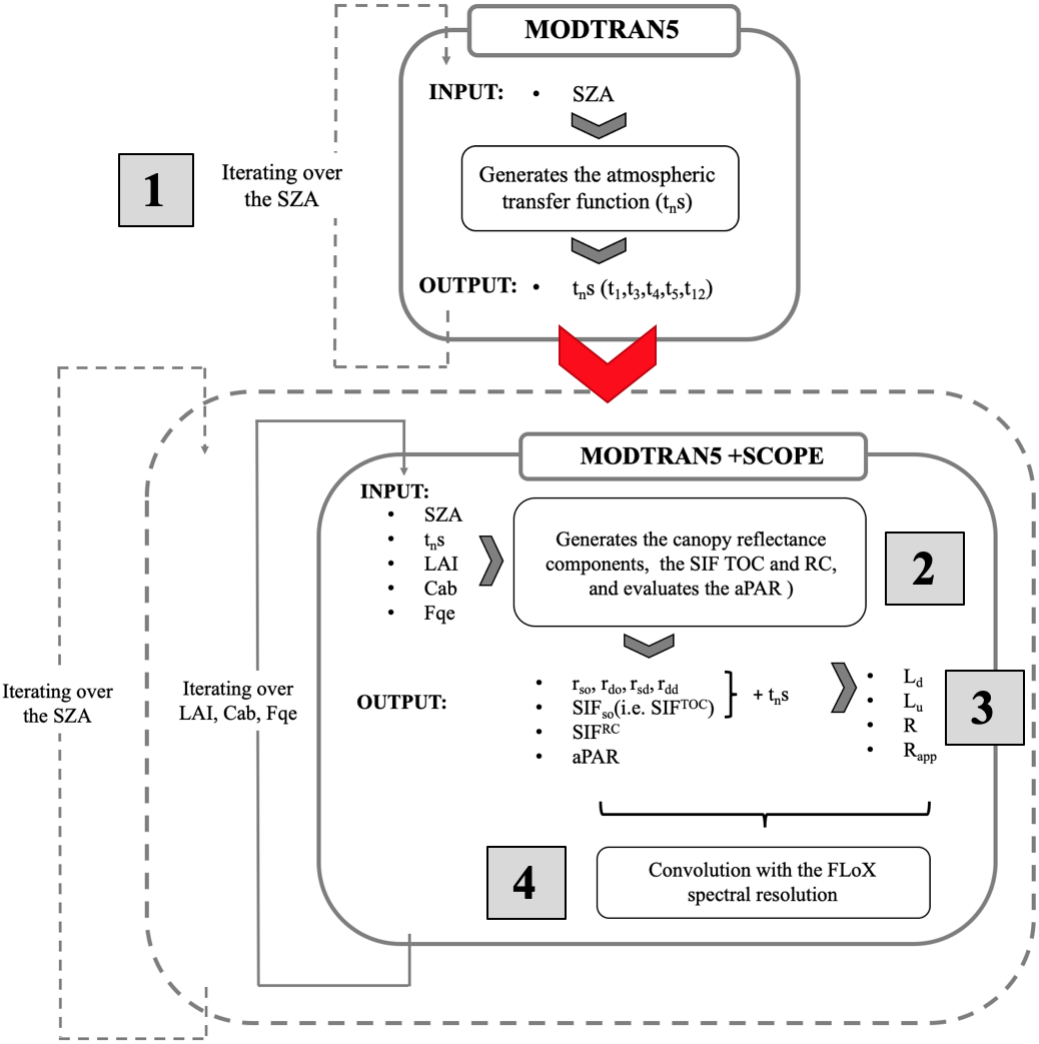

Figure 3.1: MODTRAN5-SCOPE code. Step 1: generation of the t_n. Step 2: t_n given as input in SCOPE, that run for a selected SZA, LAI, C_{ab}, Fqe combination. This module gives as output the TOC reflectance components and SIF (SIF^{TOC}), the SIF at the photosystems level (SIF^{RC}) and the aPAR. Step 3: TOC reflectance components, the SIF are coupled to the t_n obtaining L_d, L_u, R_{app} and R. Step 4: SIF^{TOC}, SIF^{RC}, L_d, L_u, R_{app}, R are convolute with the FLoX spectral resolution. Step2-4 are iterated for all the SZA, LAI, C_{ab}, Fqe combinations identified.

To generate a simulated dataset as general as possible, different illumination conditions, canopy growing stages and plants' physiological status were considered. Concerning the sun-target geometry, the SZA was made to vary between 20°-60° with a step of 2°. Concerning the vegetation, SCOPE allows to simulate a wide range of plant, characterized by different physiology, growing stages and canopy geometries. The LAI and C_{ab} vary between 1-8 m^2m^{-2} and 1-80 μgcm^{-2} respectively and they were coupled obtaining 10000 possible combinations according to De Grave et al., (2020). The LAI-C_{ab} can be qualitatively grouped in four categories: bare soil (S0), sparse (V1), medium (V2) and dense (V3) vegetation. In Figure 3.2 they have been differentiated by color code.

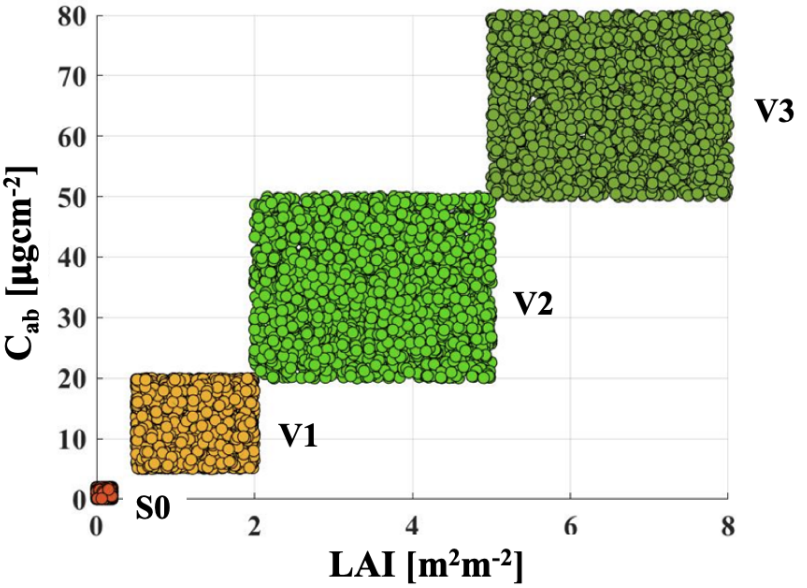

Figure 3.2: The 10000 C_{ab} (in μgcm^{-2}) and LAI (m^2m^{-2}) couples simulated. LAI-C_{ab} values can be qualitatively grouped in four categories: bare soil (S0) in brown, sparse (V1) in orange, medium (V2) in light green and dense (V3) in dark green vegetation.

According to Porcar-Castell et al., (2014), the SIF quantum yield *in-vivo* is usually lower than the 10%, with typical values lying between 0.05% and 3% under steady-state illumination and unpaired electron transport (Mohammed et al., 2019). Therefore, the SCOPE's Fqe has made it vary between 1% and 2.4%, with a step of 0.1%. The canopy geometry and the vegetation type also influence the SIF emitted and they can be opportunely selected in SCOPE. However, due to the high number of cases simulated (more than 2900000), these two variables have been kept fixed. The vegetation type selected was

C3, while the canopy geometry was assumed being spherical, setting the parameters linked to the leaf's inclinations $LIDF_a$ and $LIDF_b$ equal to -0.35 and -0.15, respectively. These variables have been chosen to match two of the experimental crops investigated, namely forage and alfalfa. As previously observed, forage and alfalfa show similar characteristics and their measurements compose most of the experimental dataset described in Chapter 2. Therefore, the parametric and the Fourier-ML methods were applied only on the forage and alfalfa spectral measurements.

3.2.2 Estimation of quantum yield

The differences between the fluorescence at TOC and at the photosystems level were investigated exploiting the metrics described in Chapter 2 (see Table 2.2). In addition, also the SIF yields were considered. In order to adequately characterize the SIF TOC and RC behaviours it is recommended to consider all the parameters affecting the fluorescence emission, namely illumination conditions (i.e. SZA), canopy biophysical parameters (e.g., C_{ab}, LAI, aPAR) and plant physiological status (linked to the SIF efficiency). Furthermore, to achieve as global an understanding as possible, these parameters need to vary in wide (and reasonable) intervals. This analysis is not feasible on experimental measurements, because many of the variables previously described are not currently available (such as the fluorescence yields or the SIF spectra at the photosystems level). Therefore, the only possible strategy consists in exploiting theoretical models (i.e., RT simulations), assuming that they accurately simulate the plant behaviour and features occurring under natural experimental conditions. In this framework, the synthetic dataset described in Section 3.2.1 were used. In particular, the SIF^{RC} simulated with SCOPE, corresponds to fluorescence evaluated at the photosystems level and then not affected by the reabsorption. The Fqe (namely the Fluorescence quantum efficiency) coincides with the true SIF yield evaluated on the spectrum corrected for the leaf-canopy reabsorption (SIF^{RC}). Although they represent the same parameter, with Fqe I identify the SCOPE input variable, while with SIF_{yield} the efficiency evaluated exploiting the MODTRAN5-SCOPE outputs (i.e. SIF^{RC} and aPAR). The SIF^{TOC}_{yield} and SIF_{yield} have been evaluated as follow:

$$SIF_{yield} = \frac{\int_{670}^{780} SIF^{RC}(\lambda)d\lambda}{aPAR} \tag{3.2.5}$$

$$SIF^{TOC}_{yield} = \frac{\int_{670}^{780} SIF^{TOC}(\lambda)d\lambda}{aPAR} \tag{3.2.6}$$

All the fluorescence metrics were investigated for increasing LAIxC_{ab} values to mimic the crop/canopy growth in time. For this reason, the plant changes occurring during the senescence phase were not considered. The product between LAI and C_{ab} represents the canopy chlorophyll content. The metrics analyses were carried out over a subset of SZA and Fqe values. In particular, the SZA was made it vary between 20°and 48°, while the Fqe in the interval 0.013-0.018. For each case, all the C_{ab} and LAI couples were analysed but

neglecting the values linked to bare soil, identified with S0 in Figure 3.2. The comparison between SIF^{TOC}_{yield} and SIF_{yield} was carried out in order to verify whether the TOC yield approximates the true fluorescence efficiency.

3.2.3 Estimation of the amount of light absorbed by the canopy (aPAR)

The aPAR is fundamental to evaluate the fluorescence yield. For this reason, a preliminary analysis of this parameter was carried out. The aPAR was studied for various $LAIxC_{ab}$ (Figure 3.3A), SZA (Figure 3.3B) and Fqe (Figure 3.3C).

In general, the aPAR shows a saturated trend for increasing canopy chlorophyll content ($LAIxC_{ab}$) values, regardless of the SZA and Fqe considered. Focusing on Figure 3.3A, as expected, the amount of light absorbed by the chlorophyll pigments is null when evaluated over bare soil surfaces (brown dots). The aPAR increases linearly in sparse vegetation (orange dots) and starts saturating when medium vegetation is considered (light green dots). In this case, the coupling between LAI and C_{ab} plays a key role in the aPAR value. Different combinations of these two parameters (i.e. different leaves shapes) affect the amount of light absorbed. Conversely, when the vegetation is dense and uniform (dark green dots), the above-mentioned variability is attenuated and the aPAR is almost constant regardless of the LAI-C_{ab} couple considered. Globally, the relationship between the canopy chlorophyll content and light absorbed is non-linear because the increment in light absorption per unit of chlorophyll decreases at high chlorophyll contents (Porcar-Castell et al., 2014; Gitelson et al., 1998; Adams et al., 1990).

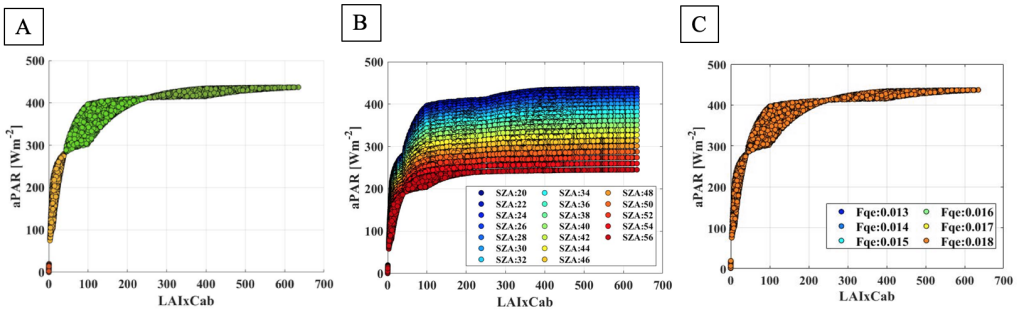

Figure 3.3: Absorbed Photosynthetically Active Radiation (aPAR). Panel A shows the aPAR evaluated for increasing $LAIxC_{ab}$ values (SZA=20° and Fqe = 0.013). The color code allows to discern between bare soil (brown), sparse (orange), medium (light green) and dense (dark green) vegetation. Panel B collects the aPAR plotted for all the SZA simulated (Fqe = 0.013). Panel C display the aPAR for variable Fqe (SZA = 20°).

The aPAR absolute values are also influenced by the SZA (Figure 3.3B). Greater values occur for lower SZA (in blue), therefore when the sun is almost perpendicular to the target. Conversely, lower aPAR are observed for higher SZA (in red). This behaviour is mainly driven by the path that the solar light follows in the atmosphere. The lower the path, the fewer the interactions with the atmosphere components, the less the energy dissipated and the greater the light reaching the Earth surface. Summarizing the aPAR is inversely correlated to the SZA. Finally, the amount of energy absorbed by the vegetation does not depend on the Fqe (Figure 3.3C): all the values simulated overlap, regardless of the fluorescence emission considered.

Since the simulated aPAR is mainly driven by biophysical parameters, such as $LAIxC_{ab}$ and SZA (linked to the PAR), it was compared to the reflectance-based indices investigated in Chapter 2. Specifically, the $NDVI_{red-edge}$, $CI_{red-edge}$ and NIR_v. The PRI was not selected because linked to the plant photosynthetic activity. Furthermore, the spectra investigated were simulated only in the spectral range in which the SIF emission occurs to save computational time. Since the PRI is evaluated combining the reflectance at 531 nm and 570 nm, it was not possible to calculate this index. This analysis was carried out in order to investigate whether the $NDVI_{red-edge}$, $CI_{red-edge}$, NIR_v could be used to estimate the amount of light absorbed by the canopy. The comparisons between the aPAR and reflectance indices were carried out for all the SZA and $LAIxC_{ab}$ simulated. Figure 3.4 collects the trends obtained for a SZA equal to 20°.

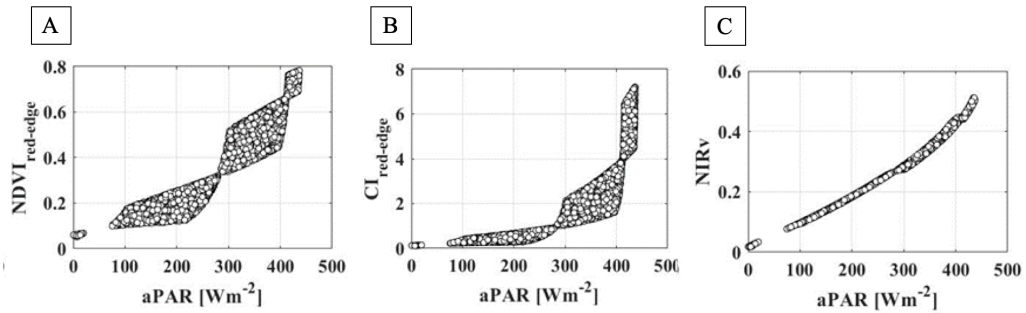

Figure 3.4: Comparison between the aPAR and the reflectance-based indices, namely Panel A: $NDVI_{red-edge}$, Panel B: $CI_{red-edge}$ and Panel C: NIR_v. Data displayed corresponds to the values obtained evaluating the spectral indices on the simulated reflectance in agreement with the equations collected in Table 2.1. Each point plotted corresponds to a $LAIxC_{ab}$ couple. SZA= 20°.

The $NDVI_{red-edge}$ and $CI_{red-edge}$ show a strong variability with the $LAIxC_{ab}$ values (Figure 3.4A and Figure 3.4B). Furthermore, a clear vertical asymptote is observed for aPAR greater than 400 Wm^{-2} in both the cases investigated. Conversely, an interesting relationship appeared between the aPAR and the NIR_v (Figure 3.4C). The NIR_v index smooths the variability linked to the vegetation considered and it is clearly less influenced by the LAI-C_{ab} combination. As reported in Badgley et al., (2017), the NIR_v is able to isolate the vegetation contribute in the reflectance, eliminating the source of error linked to the soil. Furthermore, it shows a growth almost linear with the aPAR. All these characteristics make the NIR_v a promising proxy for the aPAR. However, this peculiar relationship is probability related to the approximation used in the synthetic datased creation. In particular, the canopy is assumed to be spherical. It is plausible that the almost linear trend between aPAR and NIR_v will be modified when a more complex canopy structure is considered (e.g. erectophile). In that case, a different aPAR-NIR_v parametrization might be necessary. Bearing in mind this potential limit, a method to calculate the aPAR has been developed. In general, evaluating the amount of light absorbed by the canopy from field measurements is not trivial. This parameter can be, for example, estimated comparing the incoming, upwelling and transmitted-through-the-canopy radiances (e.g. Goward and Huemmerich, 1992). An alternative approach consists in estimating the aPAR by inversion, as described in Celesti et al., (2018). In this framework, I decided to exploit the relation between the aPAR and the NIR_v to implement a novel retrieval parametric method. The trend observed in Figure 3.4C was fitted with the following equation:

$$aPAR = k_1 sin(k_2 NIR_v) \tag{3.2.7}$$

This fitting equation was selected because it gives the best statistics (R^2= 0.996 - 0.999; RMSE = 3.4 - 4.4) and the lower margins of errors on the two coefficients k_1 and k_2. All these results are summarized in Figure 3.5.

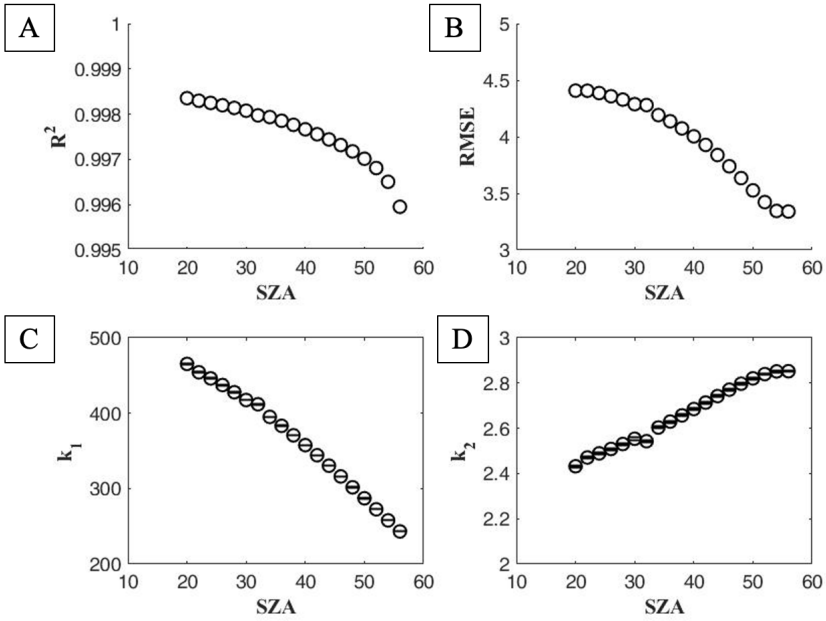

Figure 3.5: Results of the fit performed over the aPAR-NIR_v comparison displayed for increasing SZA values. Panel A: R^2, Panel B: RMSE, Panel C: k_1 and corresponding margin of error, Panel D: k_2 and corresponding margin of error.

Since under clear sky conditions, the PAR is proportional to the cosine of the local SZA, k_1 and k_2 were separately compared to the cos(SZA). k_1 shows a linear relation with cos(SZA). The fit performed (not shown) gives a slope of 584, intercept of -88, R^2 = 1 and RMSE = 2.27. Therefore, knowing the SZA value (in radiant) and the regression coefficients, it is possible to estimate k_1. Conversely, for the cos(SZA)-k_2 comparison, all the fitting equations investigated give high margins of error on the coefficients and not good statistics. For these reasons, the k_2 and SZA values have been saved in a table. Unknown and intermediate k_2 and SZA can be obtained by means of linear interception between the two closest points simulated. The aPAR is then obtained combining the NIR_v, the cos(SZA), k_1 from the linear regression coefficient and k_2 from the table created.

3.2.4 Method to correct the SIF spectrum: parametric method

The rationale exploited to parametrize the aPAR-NIR$_v$ relation was implemented to obtain the SIF spectrum corrected for the reabsorption. Several fitting equations were tested to model the simulated photosystems SIF spectral shape. However, the sum of two Gaussian (eq. 3.2.8) gave the best statistic and the lower margin of errors on the fitting parameters defined.

$$SIF = a_1 e^{-\left(\frac{wvl-b_1}{c1}\right)^2} + a_2 e^{-\left(\frac{wvl-b_2}{c2}\right)^2} \quad (3.2.8)$$

in which wvl corresponds to the wavelengths lying in the spectral range considered (i.e. 670-780 nm), a_1 and a_2 are linked to the Gaussian amplitude, b_1 and b_2 represent the two peaks positions observed in the SIF spectrum in the red and far-red respectively, while c_1 and c_2 the corresponding FWHM. All the simulated SIF spectra lying in the subset chosen (i.e. SZA=20°-48° and Fqe= 0.013-0.018) have been fitted with the equation 3.2.8. The six coefficients obtained by the fit (namely a_1, a_2, b_1, b_2, c_1, c_2) were plotted for increasing LAIxC_{ab} values and studied respect to the SZA and Fqe. In all the cases investigated, b_1, b_2, c_1 and c_2 resulted constant, regardless of the variables combination considered. In particular, b_1 is equal to 682 nm, b_2 to 706.5 nm, whereas c_1 and c_2 are equal to 11.5 nm and 54.5 nm, respectively. I want to point out that the two Gaussian used do not represent the photosystems (PSI and PSII) SIF emissions, but they fit only the overall SIF spectral shape. For this reason, b_2 and c_2 do not correspond to the usual PSI features described in literature (e.g. Maxwell and Johnson, 2000).

Conversely, a_1 and a_2 vary with LAIxC_{ab}, SZA and Fqe. Their behaviours with respect to the canopy chlorophyll content and Fqe are displayed in Figure 3.6A and Figure 3.6B. The trends obtained are similar to those observed for the aPAR in Figure 3.3. It is expected because a_1 and a_2 are linked to the SIF magnitude that intrinsically depend on the amount of light absorbed by the canopy. Nevertheless, the aPAR does not depend on the Fqe (Figure 3.3C), while a_1 and a_2 rightly show a variability respect to the fluorescence emission efficiency (Figure 3.6A and Figure 3.6B). As expected, their values increase with the Fqe. Focusing on the coefficient discrete values, the values referred to the visible red emission (i.e. a_1) are greater compared to the ones linked to the far-red (i.e. a_2).

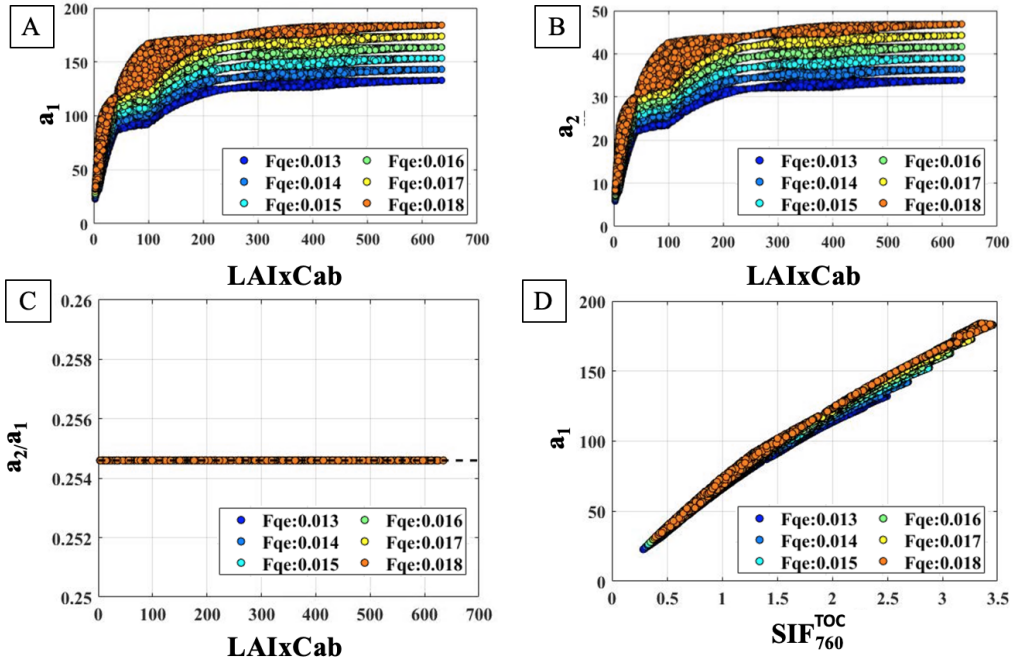

Figure 3.6: Panel A: a_1 and Panel B: a_2 behaviours for increasing $LAIxC_{ab}$ values and different simulated Fqe. Panel C: ratio between a_2 and a_1. Dashed black line corresponds to the constant values of 0.2546. This ratio is independent the $LAIxC_{ab}$ and Fqe values considered. Panel D: comparison between TOC SIF evaluated at 760 nm and a_1. All the values displayed were simulated considering a SZA equal to 20°. Color code allows to discern between the different Fqe values.

This outcome is reasonable because a_1 and a_2 are evaluated on the SIF spectra corrected for the reabsorption. Furthermore, an interesting relation links a_1 to a_2. In particular, the ratio a_2/a_1 is constant respect to $LAIxC_{ab}$, SZA, Fqe and it is equal to 0.2546 (Figure 3.6C). This result is probably a consequence of the SCOPE fluorescence parametrization and does not potentially reflect the real field situation. Nevertheless, since the RT models represent the only option to carry out this analysis, I considered valid the relationship between a_1 and a_2. This approximation is convenient because knowing one of them it is possible to estimate the other. Summarizing, b_1, b_2, c_1, c_2 are constant and their values are obtained by the fits performed. Conversely, a_2 and a_1 vary with both biophysical and physiological parameters, but their ratio is constant.

Similarly to the aPAR case, the parametric method rationale consists in linking a_1 (or a_2) to quantities measurable in field (TOC level) in order to retrieve the SIF spectral shape

corrected for the reabsorption. A preliminary analysis carried out plotting separately the a_1 coefficients and the SIF metrics for increasing $LAIxC_{ab}$ values highlighted how the a_1's trends are qualitatively similar to those observed for the TOC SIF_{760} (see Figure 3.3A and Figure 3.11 as reference). This observation is supported by the comparison between these two variables in Figure 3.6D and Figure 3.7. The variability linked to the $LAIxC_{ab}$ values is attenuated and a_1 almost linearly increases with SIF_{760}^{TOC}. A faint dependence on the Fqe is observed. The SIF_{760}^{TOC}-a_1 relationship was fitted using the following equation:

$$a_1 = a * (SIF_{760}^{TOC})^b \tag{3.2.9}$$

The equation used has been chosen because it gives the better statistics (R^2 around 0.998 $\pm$0.001; RMSE = 1.6 $\pm$0.4) and the lower margins of error on the coefficients obtained, namely a and b (see Figure 3.7).

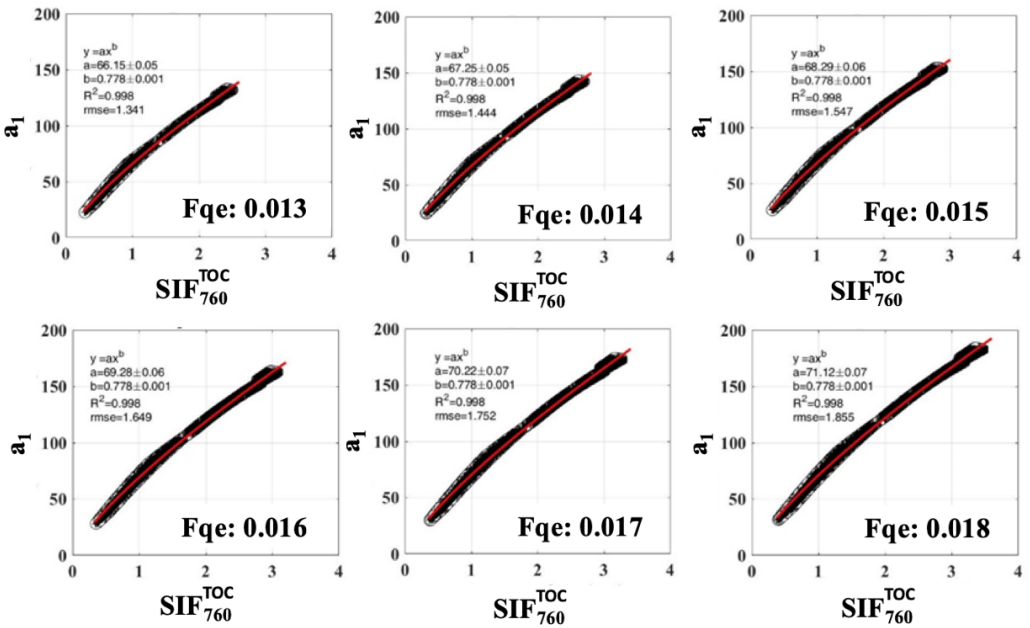

Figure 3.7: Results of the fitting performed on the SIF_{760}^{TOC} and a_1 comparison for SZA= 20°. Each point (black dot) corresponds to a specific $LAIxC_{ab}$ values. In red is represented the fit carried out.

This analysis was iterated for all the LAIxC$_{ab}$, SZA and Fqe values composing the simulated subset. All the coefficients a and b values and the corresponding errors obtained by the fit are collected in Figure 3.8. As expected they both depend on the SZA. The coefficient a shows an unclear variability with the SZA and increases with respect to the Fqe (Figure 3.8A). Conversely, the coefficient b decreases for greater SZA (Figure 3.8B) regardless of the fluorescence emission efficiency.

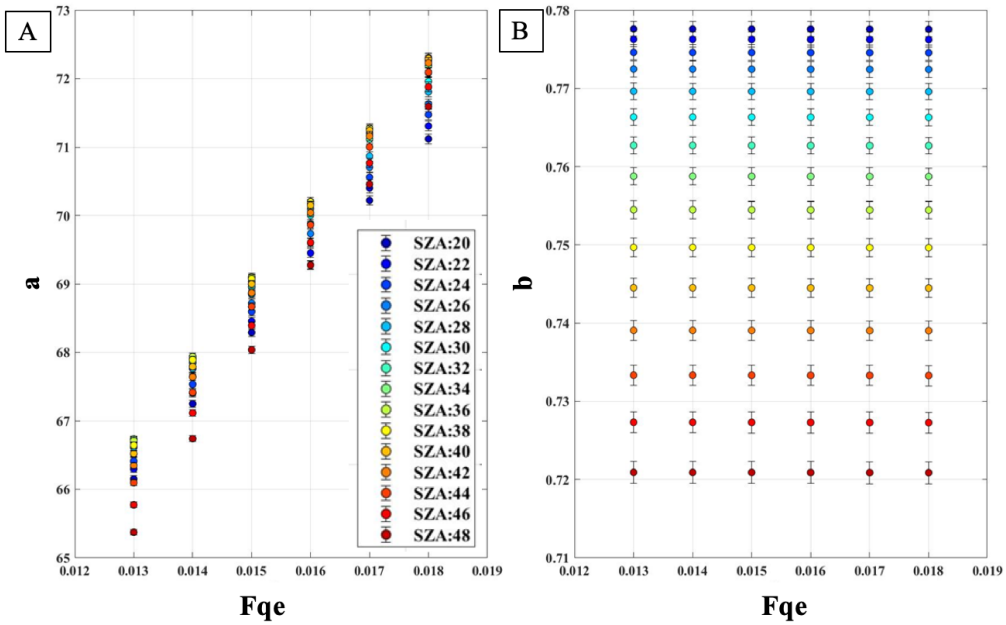

Figure 3.8: Panel A: coefficients a obtained from the fit using equation 3.2.9 displayed respect to the Fqe. Panel B: coefficients b obtained from the fit using equation 3.2.9 displayed with respect to the Fqe. Each point corresponds to the values obtained from the fit, while the error bar to the correspond margin of error (see Figure 3.7 as reference). The color code indicates the different SZA considered.

Since b depends only on the SZA, it was compared to the cos(SZA) used as proxy for the incoming radiance reaching the target (PAR). The trend obtained was fitted with a second-degree function (R^2 = 0.99; RMSE = 0.0005). The regression coefficients evaluated are p1 = -0.23 $\pm$ 0.99, p2 = 0.58 $\pm$ 0.14 and p3 = 0.43 $\pm$ 0.06 (sorted by decreasing degree). Conversely, the coefficient a case is more complicated to address because it shows a not insignificant link with the SZA and intrinsically depends on the Fqe. It must be noted that Fqe corresponds to the SIF$_{yield}$ that is the unknown parameter that I want to evaluate correcting the fluorescence spectrum for the reabsorption. Here a strong approximation

was introduced: for each SZA, the a(Fqe) coefficient was replaced by the corresponding mean value (a_{avg}) obtained averaging all over the Fqe. In this way, I am assuming that the a_1 variability linked to the Fqe (see equation 3.2.9) is only contained in the SIF_{760}^{TOC} term. Similarly to the coefficient b case, the a_{avg} values were then compared with the cos(SZA). The trend obtained is shown in Figure 3.9. It was fitted with different functions, but as happened for the k_2 term in the aPAR, the statistics obtained were not good and the variability on the regression coefficients too wide. Therefore, also in this case, the a_{avg} and cos(SZA) values were collected in a table and the intermediate measures obtained by linear interpolation.

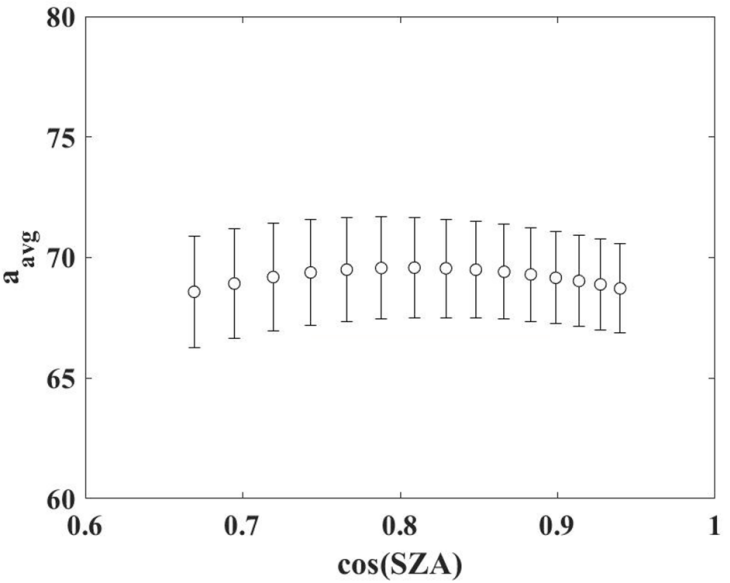

Figure 3.9: Trend between the a_{avg} values and the cos(SZA). For a fixed angle (i.e. SZA, the a_{avg} were obtained averaging over all the Fqe considered. Therefore, the data displayed (white dots) corresponds to mean value estimated, while the error bar to the standard deviations.

Summarizing the SIF spectral shape corrected for the reabsorption is obtained combining the wavelengths with the b_1, b_2, c_1, c_2 constant values, the a_2 estimated multiplying a_1 by 0.2546 and a_1 evaluated by means of the SIF_{760}^{TOC}, cos(SZA) and regression coefficient values, namely b and a_{avg}.

The method accuracy has been quantitatively assessed on the synthetic dataset evaluating the RMSE and RRMSE% as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (val_{retr} - val_{true})^2}{n}} \tag{3.2.10}$$

$$RRMSE\% = \sqrt{\frac{\sum_{i=1}^{n} \left(\frac{val_{retr} - val_{true}}{val_{true}}\right)^2}{n}} * 100 \tag{3.2.11}$$

where n is the number of points on which the comparison has been carried out. When they are evaluated on the spectrum, n corresponds to the wavelengths used. Conversely, when the analysis is carried out on the single variables (i.e. SIF metrics, aPAR or SIF yield), n is the number of LAIxC_{ab} simulated. With the label "true" I refer to the simulations outcomes, whereas "retr" are the results obtained applying the parametric approaches proposed.

To evaluate the error introduced by the approximation on the coefficient a, both the SIF spectra (and metrics) obtained with a(Fqe) and a_{avg} were compared to the corresponding true values (simulation outcomes).

3.2.5 Fourier-ML algorithm

In this section I have briefly described the Fourier-ML method used to retrieve both the SIF at TOC and photosystems scales, along with many relevant biophysical parameters, namely the C_{ab}, LAI, aPAR and Fqe. A Fourier transform is applied on the apparent reflectance (R_{app}), generating a discrete point in the 2D complex (or phasor) plane. Since the R_{app} spectral shape is influenced by the canopy structure and biophysical parameters, each point lying in the phasor plot contains information about the signal spectral properties. Spectra characterized by similar parameters lie closely within the 2D complex space. The algorithm proposed is trained on the points obtained performing the Fourier transform on the simulated spectra described in Section 3.2.1. In this framework, the above-mentioned biophysical parameters, along with the SIF spectra are known. Therefore, each point is uniquely associated to this information. The discrete Fourier transform is then applied on the experimental R_{app} in which all the biophysical parameters and SIF spectra need to be estimated. The simulated points used as reference are compared to the experimental ones. As previously assessed, points characterized by similar parameters lie close in the phasor plane. Therefore, exploiting both a supervised machine learning approach, coupled with a statistical analysis, the SIF spectrum at TOC, the SIF spectrum corrected for reabsorption, the SIF quantum efficiency (Fqe) and biophysical parameters are retrieved starting from the R_{app} spectra.

3.3 Results

3.3.1 Characterizing the reabsorption effect

The comparison between the SIF TOC and SIF corrected for reabsorption was preliminarily carried qualitatively on the MODTRAN5-SCOPE simulated spectra. A few examples are displayed in Figure 3.10.

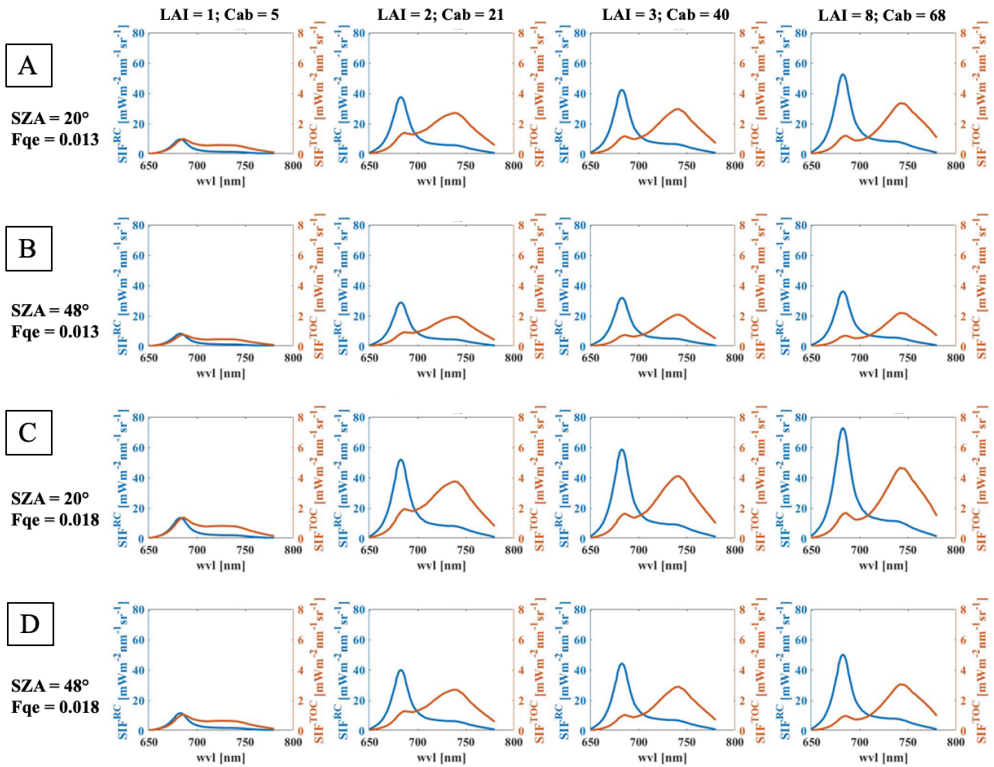

Figure 3.10: Comparison between SIF spectra simulated at TOC scale (orange) and corrected for the reabsorption (SIF^{RC} in blue). They are both measured in $mWm^{-2}nm^{-1}sr^{-1}$. Each column corresponds to a different LAI (in m^2m^{-2}) and C_{ab} (in μgcm^{-2}) couple, increasing from left to right. Each row corresponds to different SZA and Fqe values.

In terms of discrete values, the fluorescence emitted at the photosystems level (orange spectra) is always greater compared to the one evaluated at top-of-canopy scale (blue spectra). This result is expected because part of the SIF emitted is re absorbed by the leaf-canopy system. When re absorbed, the energy could be used again for PQ, NPQ or SIF. Therefore, only a fraction of the SIF absorbed is re emitted radiatively. Furthermore, in the path within the canopy, the fluorescence could be scattered outside the sensor field of view. The reabsorption also has a strong influence on the SIF spectral shape. All the spectra simulated at the photosystems level are characterized by a clear peak in the visible red and a broad shoulder in the far-red. Conversely, at TOC scale, the SIF peaks are more or less evident depending on the LAI and C_{ab} considered. For very low LAI and C_{ab} (e.g. the first column in Figure 3.10), red and far-red emissions are comparable. Increasing the canopy chlorophyll content (Figure 3.10 from left to right), the fluorescence signal is dominated by the far-red, even though at the photosystems the SIF is emitted mainly in visible red. In terms of spectral shape, the LAI and C_{ab} values influence only the SIF^{RC} magnitude, leaving unchanged the position of the peaks. Conversely, at TOC the $\lambda_{far-red}$ shifts toward longer wavelength is here observed, as discussed in Chapter 2 on the experimental measurements.

This shift is again linked to reabsorption. Comparing column 2 and 4 in Figure 3.10, a peculiar trend is observed in the region between the two peaks when the SIF^{TOC} are considered. In particular, the local minimum identified becomes deeper for greater LAI-C_{ab} values. This trend could be explained analysing simultaneously the SIF^{RC} and SIF^{TOC} behaviors. The emission at the photosystems level in the red increase with the LAI-C_{ab} and it is not completely re absorbed. Therefore, at TOC scale, a residual contribution in the red is not re absorbed. If the red and far-red both increase and the far-red peak shifts toward longer wavelengths, the gap between the two emission peaks becomes clearer (at TOC). The SIF changes were also evaluated for different SZA and Fqe values. In Figure 3.10 are shown only the spectra simulated with SZA equal to 20° and 48° and Fqe equal to 0.013 and 0.018, respectively. For a fixed SZA, the increasing Fqe values determine and overall increase of the SIF magnitude, but the spectral shape seems unaffected, regardless of the scale considered (e.g. Figure 3.10A and Figure 3.10C). The same trend was observed for fixed Fqe and variable SZA (e.g. Figure 3.10A and Figure 3.10B). Summarizing, the SIF magnitude depends on the LAIxC_{ab}, SZA and Fqe values regardless of the scale considered. Furthermore, the SIF spectral shape is strongly affected by the LAIxC_{ab} when the signal is upscaled from the photosystems level to Top-Of-Canopy scale. This trend is attributable to reabsorption.

A similar analysis was carried out by means of the SIF metrics evaluated on the full spectrum. Figure 3.11 collects the results obtained, in which the first row considers the SIF^{TOC} metrics, while the second row the ones estimated on the corrected spectra (SIF^{RC}).

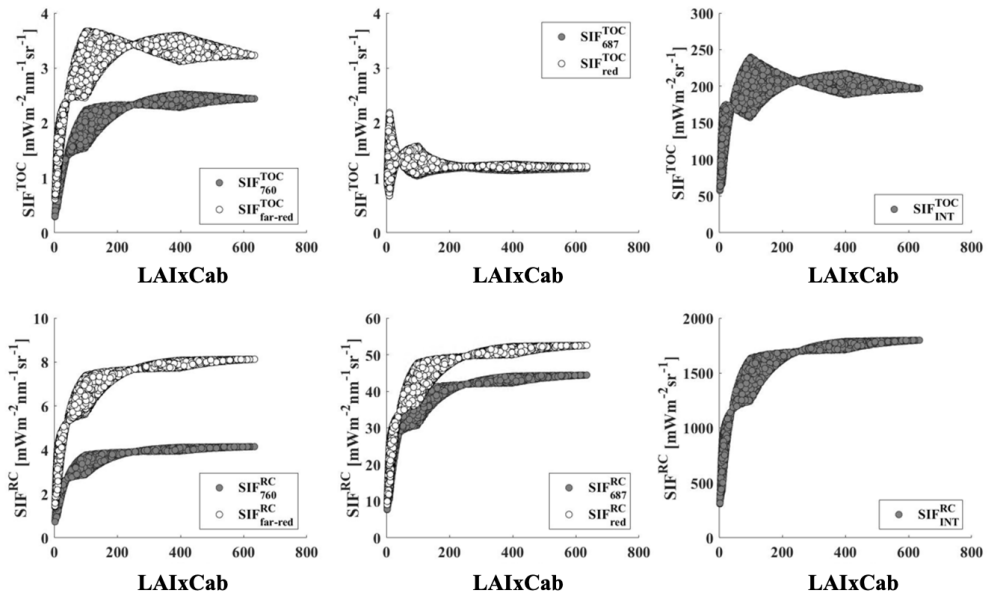

Figure 3.11: (top): Fluorescence metrics evaluated on the TOC SIF spectra. (bottom): Fluorescence metrics evaluated on the SIF corrected for the re-absorption (i.e. at the photosystems level). The trends showed are evaluated on the synthetic dataset. Each point corresponds to a single $LAIxC_{ab}$ value. Measurements referred to the bare soil were not considered. Color code (white rather than grey) allows to discern between the metrics evaluated in the same spectral region. Except for the SIF_{INT} cases, grey dots correspond to the SIF estimated in the O_2 bands, white dots to the fluorescence at the peaks. Specifically, SZA= 20°. Fqe = 0.013.

This qualitative comparison clearly highlights the strong effect that the reabsorption has on the SIF^{TOC} in the red spectral region. While all the SIF^{RC} metrics show a saturated trend for increasing $LAIxC_{ab}$, the SIF^{TOC} is characterized by different behaviors depending on the spectral region considered. In particular, the fluorescence metrics defined in the far-red and the spectrally integrated SIF saturate with respect the $LAIxC_{ab}$. Globally, SIF^{TOC}_{760}, $SIF^{TOC}_{far-red}$, SIF^{TOC}_{INT} show the same behaviour because when reabsorption occurs, the far-red contribution in the SIF spectrum is predominant. This theoretical result corresponds with the observations obtained on the experimental measurements (Chapter 2, Figure 2.4A-B and Figure 2.5 first row). On the other hand, the TOC fluorescence in the visible red is characterized by an initial variability (values referred to sparse vegetation), then followed by a decrease when a denser canopy is considered. This trend is attributable to the reabsorption that quenches the fluorescence signal in the red and prevails for greater

canopy chlorophyll content. In this case, SIF_{687}^{TOC} and SIF_{red}^{TOC} are characterized by the same values, in in accordance with what was observed experimentally in Chapter 2 (Figure 2.4C and Figure 2.5D).
At TOC scale, the $LAIxC_{ab}$ coupling plays a key role in the metrics values. In case of medium vegetation ($LAIxC_{ab}$ between 50 and 200), $SIF_{far-red}^{TOC}$ and SIF_{INT}^{TOC} show a clear variability. This behaviour is less evident at the photosystems level.
In general, all the metrics values (expect SIF_{red}^{TOC} and SIF_{687}^{TOC}) saturate moving from sparse to dense vegetation. The same trend was also observed for the aPAR (Figure 3.3A). This result is expected because, the amount of energy dissipated as fluorescence strongly depends on the amount of light absorbed by the plant. Therefore, one of the SIF drivers is the aPAR. The SIF metrics were also investigated with respect to the SZA and Fqe (not shown). Both these parameters mainly affect the metrics magnitude. Concerning the SZA, greater values are obtained for lower angles, similarly to the aPAR case (Figure 3.3B). Since the Fqe determines the efficiency of the fluorescence process, the SIF values increase mutually with the Fqe.

The changes in terms of SIF spectral shape, qualitatively highlighted in Figure 3.10, were investigated by means of the SIF metrics linked to the maximum emission wavelengths. At the photosystems level the peak position in the red and far-red are always constant, regardless of the $LAIxC_{ab}$, SZA and Fqe considered (not shown). Specifically, λ_{red}^{RC} lies around 682 nm, while $\lambda_{far-red}^{RC}$ around 740 nm. Conversely, at TOC scale, the pattern observed on the field measurements is found also on the synthetic dataset (see Figure 2.6 and Figure 3.12). Figure 3.12 collects the metrics values evaluated at TOC scale. In general, λ_{red} lies around 685 nm and slightly decreases with increasing $LAIxC_{ab}$. Conversely, the far-red peak (i.e. $\lambda_{far-red}$) undergoes to a clearer shift toward longer wavelengths. As expected, the peak positions do not depend on the SZA and Fqe values.

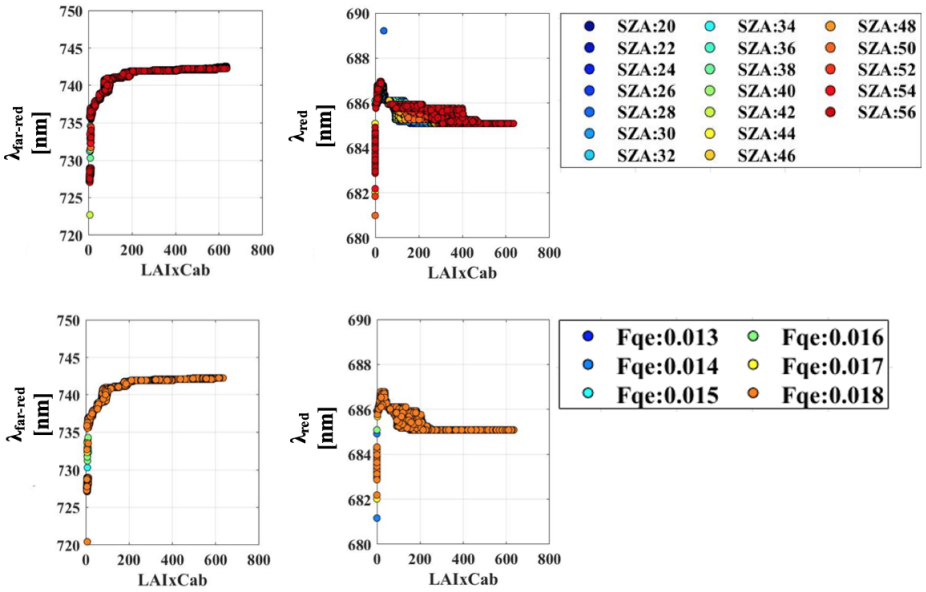

Figure 3.12: (top): peaks positions in the far-red (left) and visible red (right) evaluated on the TOC SIF for different SZA (Fqe = 0.013). (bottom): peak positions in the far-red (left) and visible red (right) evaluated on the TOC SIF for different Fqe (SZA= 20°). Each point corresponds to a specific $LAIxC_{ab}$ value. The color code helps to discern between the whole SZA and Fqe considered.

The last two metrics investigated are the SIF yields evaluated at TOC scale and at photosystems level (see equations 3.2.6 and 3.2.5). The fluorescence yield represents a pivotal parameter in the vegetation monitoring because strictly linked to the efficiency of the photosynthetic process. In this framework, I want to investigate whether the SIF^{TOC}_{yield} approximates the true value obtained from the spectrum corrected for reabsorption (SIF_{yield}). The SIF^{TOC}_{yield} and SIF_{yield} trends were analysed for different SZA and Fqe. From a qualitative point of view, a clear difference appears between the yield evaluated at TOC scale and on the RC spectrum (Figure 3.13; case Fqe = 0.013). The first one shows a strong dependence on the $LAIxC_{ab}$ values (Figure 3.13A). The decreasing trend observed is probably attributable to the reabsorption effect. Conversely, at the photosystems level the SIF_{yield} is constant regardless of the $LAIxC_{ab}$ and SZA values. The SIF yields evaluated for a fixed fluorescence efficiency were averaged over all the $LAIxC_{ab}$ and SZA simulated and compared to corresponding Fqe (Figure 3.13B and D). In general, a linear relation was observed in both the cases investigated. At TOC, the average values show an evident margin of errors linked to the strong variability of the measurements observed in Figure

3.13B. The liner regression performed gives p1 = 0.131, p2 = $8.7e^{-8}$, R^2= 1 and RMSE= $4e^{-10}$. Conversely, at the photosystems level (Figure 3.13D), all the average SIF_{yield} lie on the 1:1 (p1 = 0.995, p2 = $7e^{-11}$, R^2= 1 and RMSE= $3e^{-11}$). As expected, the SIF_{yield} evaluated on the fluorescence spectrum corrected for the reabsorption match with the theoretical quantum efficiency settled in SCOPE.

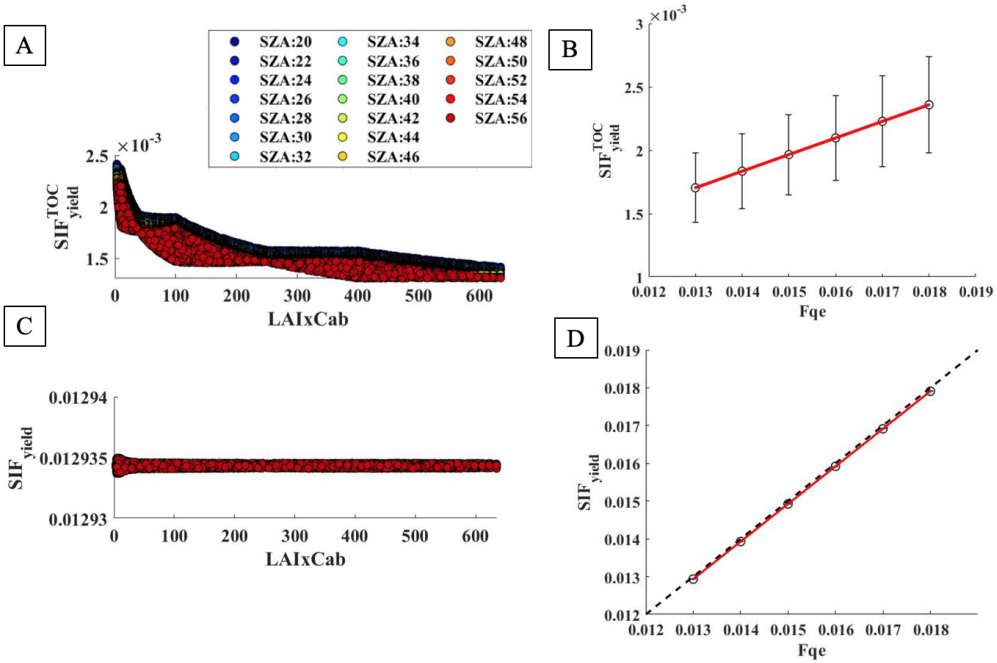

Figure 3.13: Panel A: SIF^{TOC}_{yield} for different SZA values. Fqe = 0.013. Panel B: Comparison between the Fqe (SCOPE input) and the SIF^{TOC}_{yield} evaluated according to eq. 3.2.6. Panel C: SIF_{yield} for different SZA values. Fqe = 0.013. Panel D: Comparison between the Fqe (SCOPE input) and the SIF_{yield} evaluated according to eq. 3.2.5. Color code in Panel A and C corresponds to the several SZA considered. Data displayed in Panel B and D correspond to the average performed over all the $LAIxC_{ab}$ and SZA values for a fixed Fqe. The error bars are the standard deviations. In red the linear regression performed. The dashed line is the 1:1.

CASE ID	Parameters	Parameters values	Interval of confidence
TOC	**p1**	0.131	(0.131 ; 0.131)
	p2	$8.7e^{-8}$	($8.3e^{-8}$; $9.0e^{-8}$)
	$\mathbf{R^2}$	1	-
	RMSE	$4e^{-10}$	-
RC	**p1**	0.995	(0.995 ; 0.995)
	p2	$7e^{-11}$	($-3e^{-10}$; $4e^{-10}$)
	$\mathbf{R^2}$	1	-
	RMSE	$3e^{-11}$	-

Table 3.1: Coefficients and statistics obtained from the liner regression shown in Figure 3.13B (Case ID TOC) and Figure 3.13D (Case ID RC). p1 corresponds to the linear regression slope, while p2 to the intercept.

3.3.2 aPAR parametric method accuracy (RT simulations)

The performances linked to the aPAR retrieval algorithm by means of the NIR_v reflectance index were investigated. The accuracy of the method proposed was assessed on the synthetic dataset. A comparison between the "true" aPAR (SCOPE outcomes) and the corresponding values obtained with the method described in Section 3.2.3 was carried out. The accuracy was verified from the results of the linear regression performed on the scatterplots displayed in Figure 3.14.

In all the cases investigated, the values estimated lie on the 1:1 (dashed grey line) and show R^2 equal to 1. Moreover, the RMSEs evaluated are always lower than 4.5 (see Figure 3.14 for the regression results and statistics).

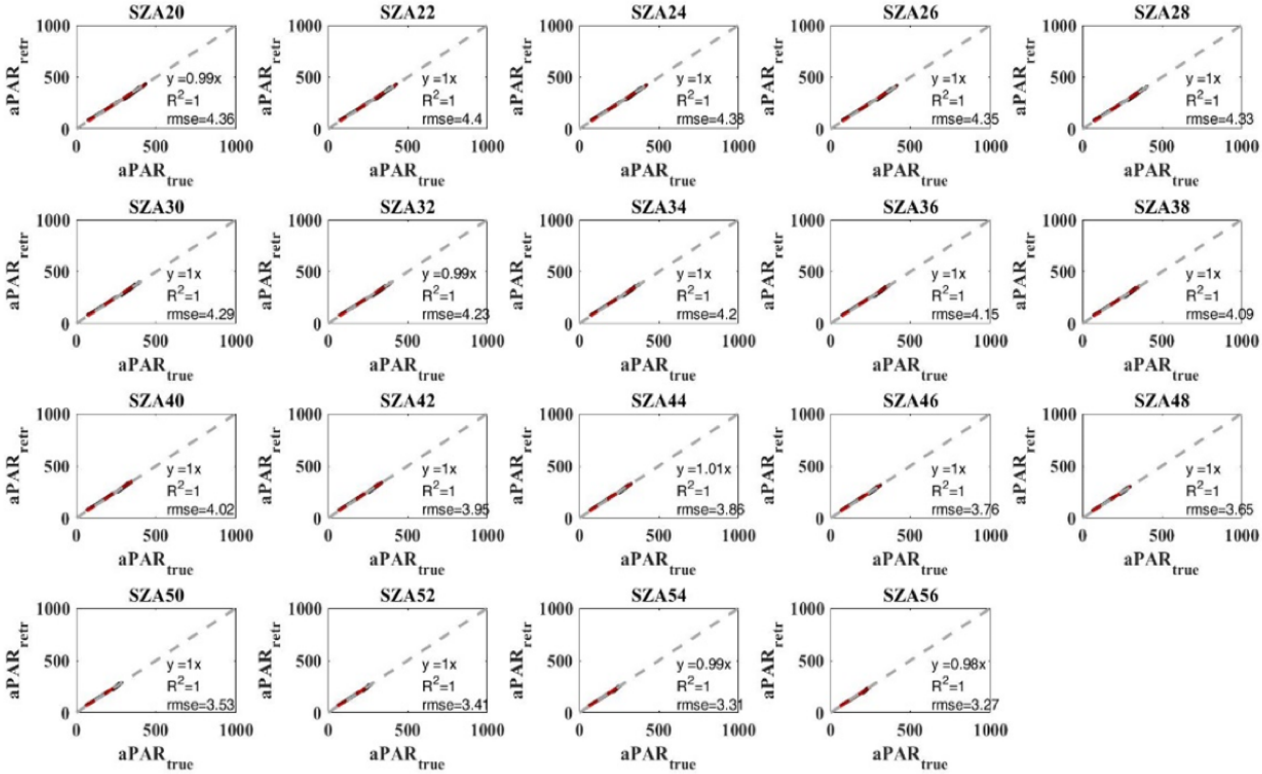

Figure 3.14: Comparison between the aPAR generated by SCOPE ($aPAR_{true}$) and the aPAR evaluated with the method proposed in Section 3.2.3 ($aPAR_{retr}$). Points corresponds to the $LAIxC_{ab}$. Dashed grey line is the 1:1. In red is represented the linear regression performed.

3.3.3 SIF^{RC} retrieval accuracy (RT simulations)

3.3.3.1 Parametric method

The accuracy of the parametric method was first evaluated on the simulated measurements. The test was carried out on the same subset used in Section 3.3.1, namely SZA=20°-48° and Fqe= 0.013-0-018. Furthermore, the values corresponding to bare soil were not considered.

The SIF simulated at the photosystems level by SCOPE was compared to the spectra obtained by the parametric method (with and without the approximation on the a coefficient), with the same input variables (i.e. SZA, LAI, C_{ab} and Fqe). The accuracy was quantitatively estimated evaluating the RMSE and RRMSE% in agreement with eq. 3.2.10 and eq. 3.2.11, respectively. This comparison allows to quantify how much the retrieved spectral shapes differ from the simulated one. Figure 3.15 reports few examples of the spectra obtained using a_{avg} (Figure 3.15A) and a(Fqe) (Figure 3.15A), respectively. The spectra simulated with SCOPE are shown in blue, in red the ones estimated with the parametric method. In all the six LAI and C_{ab} cases shown, the approach proposed seems to model correctly the theoretical SIF^{RC} spectral shape, regardless of the approximation on the coefficient a. As expected, better performances are obtained using a(Fqe), as highlighted by the RMSE shown in Figure 3.15B. The only exception is observed for low LAI and C_{ab} values (see case LAI=2 and C_{ab}=10).

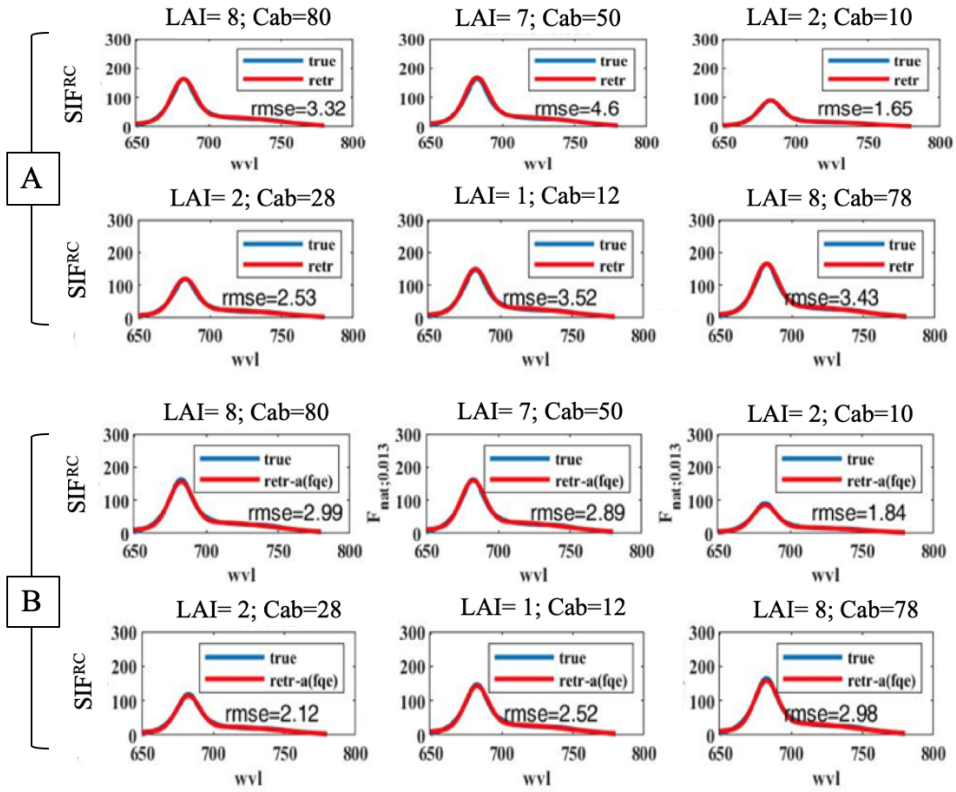

Figure 3.15: Comparison between SIF^{RC} spectra simulated with SCOPE (in blue) and using the parametric method (in red). Case A: spectra evaluated with approximation on the coefficient a (a_{avg}). Case B: spectra evaluated without approximation on the coefficient a (a(Fqe)). SZA= 20°. Fqe= 0.013.

This analysis was carried out for all the LAI, C_{ab}, SZA and Fqe values composing the subset. Figure 3.16 summarize the results obtained in terms of RRMSE% for all the cases investigated with and without the approximation on the a coefficient. The method always retrieves the SIF^{RC} spectra with an RRMSE% lower than the 40% (Case a_{avg}) and 35% (Case a(Fqe)). A strong dependence to the LAIxC_{ab} value is clear. Greater RRMSE% are observed for sparse vegetation (LAIxC_{ab} lower than 20). In this case the low chlorophyll concentration, along with the background contributions, negatively influence the SIF^{RC} retrieval. Conversely, the RRMSE% are less variable for the other vegetation simulated (i.e. medium and dense). When a(Fqe) is considered, the RRMSE% are independent of the Fqe and are less affected by the SZA. On the other hand, when the a_{avg} is used, a

dependence on both the SZA and Fqe is clear. In particular, a_{avg} better approximates the SIF^{RC} simulated with higher Fqe. Not considering the outcomes associated to the sparse vegetation, the parametric method models the SIF^{RC} spectral shape with an RRMSE% always lower than the 25% in case of a(Fqe) and 27% for a_{avg}.

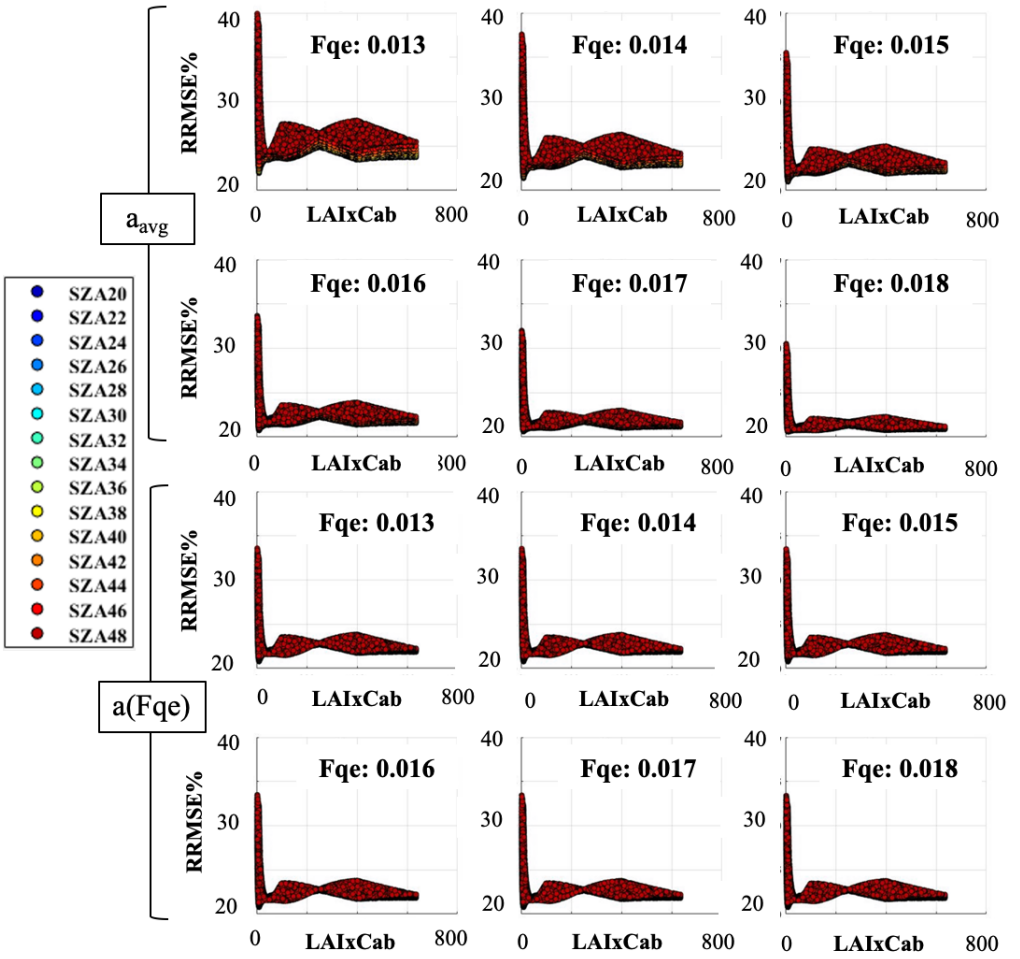

Figure 3.16: RRMSE% obtained comparing SIF simulated with SCOPE and SIF spectra corrected for reabsorption with the parametric method.

From an experimental point of view, the variables of interest are the SIF^{RC} evaluated at the peaks, because linked to the photosystems emission, and the integral of SIF^{RC} for the fluorescence yield evaluation. The corresponding metrics were evaluated on the spectrum corrected for reabsorption and compared to the values estimated on the SCOPE simulations. Similarly to the previous analysis, the SIF spectra were corrected with and without the approximation on the coefficient a. The three metrics investigated are SIF^{RC}_{687}, SIF^{RC}_{760} and SIF^{RC}_{INT}. A linear regression was performed on the scatterplot between true and retrieved metrics for all the $LAIxC_{ab}$ couples, but fixed SZA and Fqe values (Figure 3.17). First row of Figure 3.17 shown the above-mentioned scatterplots for SZA= 20° and Fqe= 0.013. On the x the true SIF metrics estimated on the SCOPE spectra are reported, while on the y the corresponding values obtained with the parametric method. Each point corresponds to a single $LAIxC_{ab}$ value, the linear regression performed is shown in red. The dashed line represents the 1:1. These comparisons were carried out for all the SZA and Fqe composing the subset. The regression outcomes are summarized in Figure 3.17 (rows from 2 up to 5) and shown for increasing SZA. The color code (the same used in Figure 3.12) allows to discern between the several Fqe considered. Lower Fqe values are in blue, medium in green, higher in orange.

It is clear that the method accuracy increases when only selected metrics are considered instead of the whole spectra. Using the first row of Figure 3.17 as reference, all the metrics investigated lie on the 1:1 (dashed grey line) regardless the $LAIxC_{ab}$ considered and the approximation applied. All the slopes of the regression line are collected in Figure 3.17B. As expected, a variability with the SZA and Fqe is observed when a_{avg} is used (column 1-3). Nevertheless, the slopes lie between 0.9 and 1.1 for all the cases considered. Conversely, when the SIF spectrum is corrected with the coefficient a(Fqe), the slopes are constant and equal to 1, except for the SIF^{RC}_{687} case characterized by value lower than 1.

The RMSE depends on the Fqe values, regardless of the approximation applied on the coefficient a (Figure 3.17C). In particular, it increases with the fluorescence efficiency. Conversely, the RRMSE% evaluated with a(Fqe) are constant with the SZA and always lie below 5% for all the metrics considered (Figure 3.17D). Slightly greater values are observed for the metric evaluated at 687 nm. Conversely, when the a_{avg} case is considered, there is more variability in the results obtained. A strong dependence to the SZA and the Fqe is clear, but in general the RRMSE% evaluated always lies below 10%. Finally, the R^2 always lies between 0.994 and 1 (Figure 3.17E). Decreasing values are obtained for greater SZA, regardless of the Fqe and the coefficient approximation used.

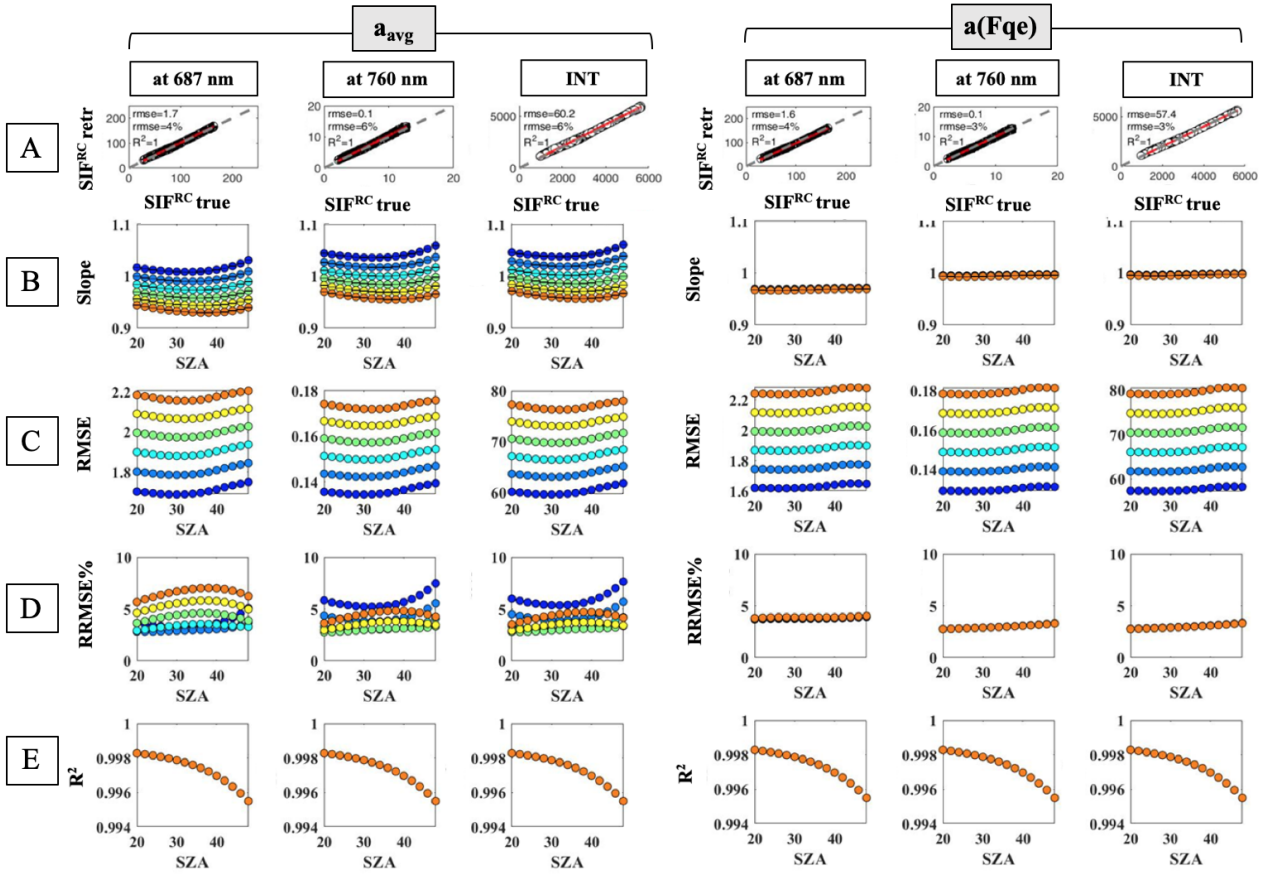

Figure 3.17: (first row): scatterplot between the SIF^{RC} metrics evaluated on spectra simulated and corrected for reabsorption (SZA= 20°; Fqe= 0.013). (rows 2-5): outcomes of the linear regression for all the cases investigated. A detailed Figure description is reported in the text. Dots colors are linked to the whole Fqe values investigated (blue low, green medium, orange high Fqe).

The last parameter studied is the SIF yield. The reliability of the SIF_{yield} evaluated using the aPAR and SIF^{RC} from the corresponding parametric methods is investigated. The SIF_{yield} (at the photosystems level) was evaluated following different approaches. Six cases were highlighted and compared, namely:

- **Case 0**: it represents the reference because the SIF_{yield} is evaluated using the aPAR and SIF^{RC}_{INT} originated with SCOPE;
- **Case 1**: the SIF_{yield} is evaluated using the aPAR generated by SCOPE and the SIF^{RC}_{INT} obtained using the coefficient a_{avg};
- **Case 2**: the SIF_{yield} is estimated using the aPAR simulated by SCOPE and the SIF^{RC}_{INT} obtained using the a(Fqe) values;
- **Case 3**: the SIF_{yield} is evaluated using the aPAR retrieved by means of the NIR_v index and the SCOPE SIF^{RC}_{INT};
- **Case 4**: the SIF_{yield} is obtained using the aPAR retrieved by means of the NIR_v index and the the spectrally integrated SIF_{RC} retrieved using the approximation a_{avg};
- **Case 5**: the SIF_{yield} is evaluated using the aPAR retrieved by means of the NIR_v index, combined with the spectrally integrated SIF_{RC} obtained using the a(Fqe) values.

The cases summarized in Table 3.2 were compared to the Fqe values settled as input in SCOPE. The comparison was carried out for all the $LAIxC_{ab}$ and SZA composing the dataset.

CASE ID	Outcome	aPAR	SIF^{RC}_{INT}
0	SIF_{yield}	SCOPE	SCOPE
1	SIF_{yield}	SCOPE	param. method with a_{avg}
2	SIF_{yield}	SCOPE	param. method with a(Fqe)
3	SIF_{yield}	param. method	SCOPE
4	SIF_{yield}	param. method	param. method with a_{avg}
5	SIF_{yield}	param. method	method with a(Fqe)

Table 3.2: SIF_{yield} evaluated on the synthetic dataset exploiting six different strategies.

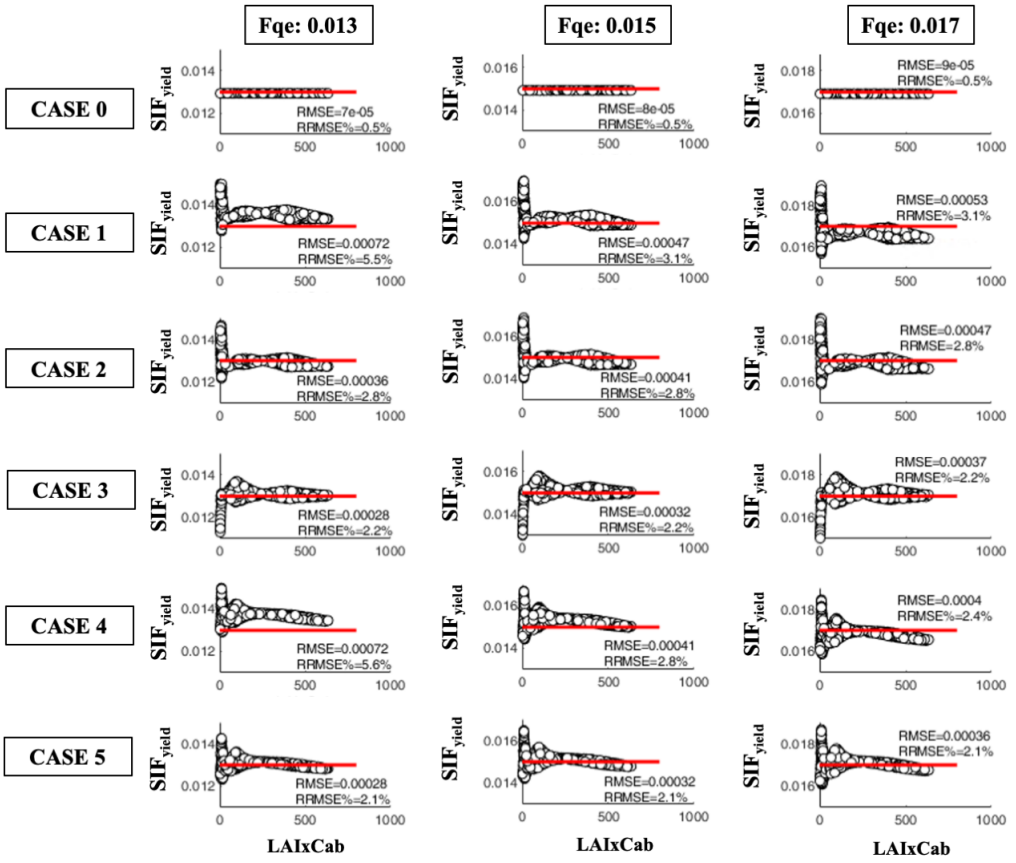

Figure 3.18: SIF_{yield} evaluated on the synthetic dataset. Each row corresponds to a case described in Table 3.2. White dots correspond to the several $LAIxC_{ab}$ values. The red line is the Fqe given as input in SCOPE. SZA = 20°.

A representative example of the analyses carried out is reported in Figure 3.18 for a SZA of 20° and Fqe equal to 0.013, 0.015 and 0.017. The red line represents the Fqe imposed as input is SCOPE, while the points to each SIF_{yield} evaluated in the six summarized in Table 3.2 for all the LAI-C_{ab} composing the subset.
Except for Case 0, that represents the reference, in all the cases investigated a strong dependence on the LAI and C_{ab} values is clear. In particular, a great variability is observed when sparse vegetation is considered ($LAIxC_{ab} < 20$).
Case 1 and Case 2 (true aPAR and SIF^{RC} parametrized) allow to highlight the impact of the approximation of the a coefficient on the SIF_{yield} discrete values. When a_{avg} is considered

(Case 1), the outcomes shift respect to the reference (red line). Specifically, the SIF_{yield} is overestimated for low Fqe, almost correctly estimated for medium Fqe and underestimated for greater Fqe. This trend is attributable to the average performed on the a coefficient because in Case 2 all the values obtained lie around the reference, regardless of the Fqe considered.
Case 3 allows evaluation of the impact of the aPAR parametrization on the SIF_{yield} value. There is a clear variability linked to the $LAIxC_{ab}$, specifically when sparse vegetation is considered. Nevertheless, the SIF_{yield} computed always lies around the true Fqe value.
Case 4 mimics a potential experimental situation, in which the Fqe is unknown and then the a_{avg} is used to correct the SIF spectra for reabsorption. The aPAR is evaluated exploiting the NIR_v reflectance index. Similarly to Case 1, the SIF_{yield} moves around the reference accordingly to the Fqe considered.
Conversely, Case 5 exploits the a(Fqe) values and the SIF_{yield} is more stable with respect to the reference.

The qualitative comparison previously described, was iterated over all the SZA and Fqe composing the subset. The RMSE and RRMSE% obtained were collected in Figure 3.19. As expected, lower RMSE and RRMSE% are computed for Case 0, which corresponds to the reference. The a_{avg} approximation clearly introduces a variability linked to both the SZA and Fqe, regardless of the aPAR used to evaluate the SIF_{yield} (Case 1 and Case 4). The RMSE and RRMSE% are lower than 0.0011 and 8%, respectively. Conversely, when the true SIF^{RC} or the a(Fqe) are used, the dependence on the Fqe disappears (see the RRMSE% trend). In these cases, the RMSE and RRMSE% always lie below 0.0006 and 4% (Case 2, Case 3 and Case 4). Case 4 shows a faint dependence on the SZA, with values that mutually increase with the SZA.

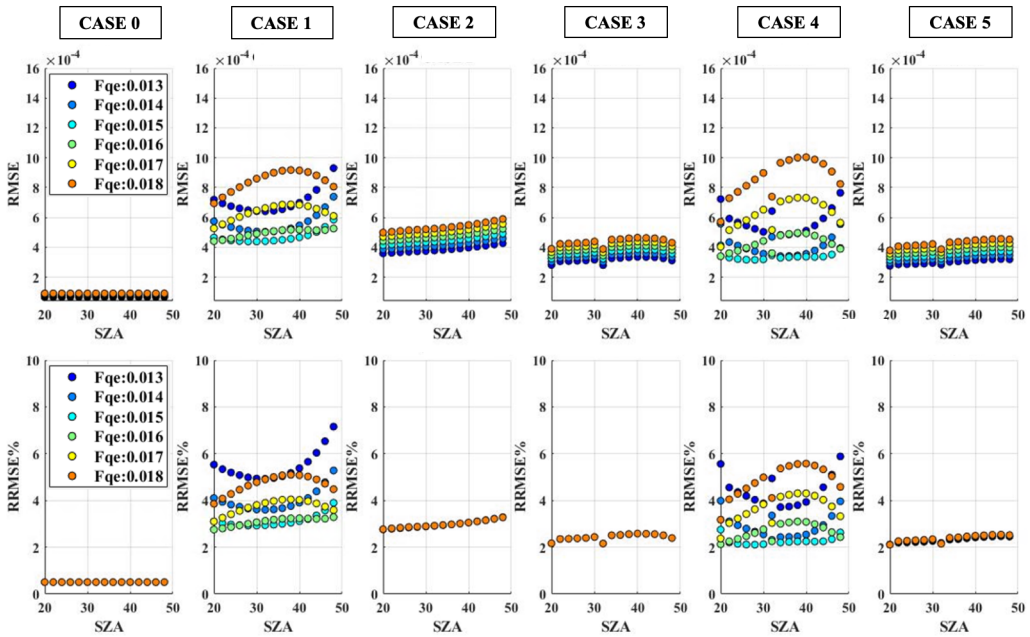

Figure 3.19: RMSE and RRMSE% evaluate on the comparison between Fqe and the SIF_{yield} for the six cases summarized in Table 3.2. The values are displayed for increasing SZA. Color code allows to discern between the several Fqe investigated.

3.3.3.2 Fourier-ML algorithm

The Fourier-ML algorithm accuracy was investigated dividing the whole synthetic dataset described in Section 3.2.1 in 80% training and 20% test set, with a 5-fold cross-validation. The retrieving accuracy was evaluated on both the biophysical parameters and the selected SIF metrics estimated on the full spectra (TOC and at the photosystems level). In particular, the fluorescence was estimated in the two oxygen absorption bands at 760 nm (O_2-A) and 687 (O_2-B) nm. The true values were compared with the corresponding one estimated with the Fourier-ML algorithm and a linear least square fit was performed. The results obtained are summarized in Figure 3.20.

All the parameters investigated show an almost perfect linear trend and the values lie on the 1:1, as demonstrated by the slopes of the linear regression performed (between 0.986 and 1) and the intercepts close to zero. Furthermore, the R^2 are always greater than 0.99. A peculiar trend is observed for the LAI, in which a spread of the points around the reference is clear for values between 6-8. Conversely, the less continuous pattern characterizing the Fqe is due to the low variability of this parameters with respect to the other variables (i.e.

23 cases for the Fqe vs 10000 for the LAI and C_{ab}). The RRMSE% estimated are collected in Table 3.3. As expected, the LAI is characterized by the higher error. Nevertheless, the RRMSE% is always lower than the 2%.

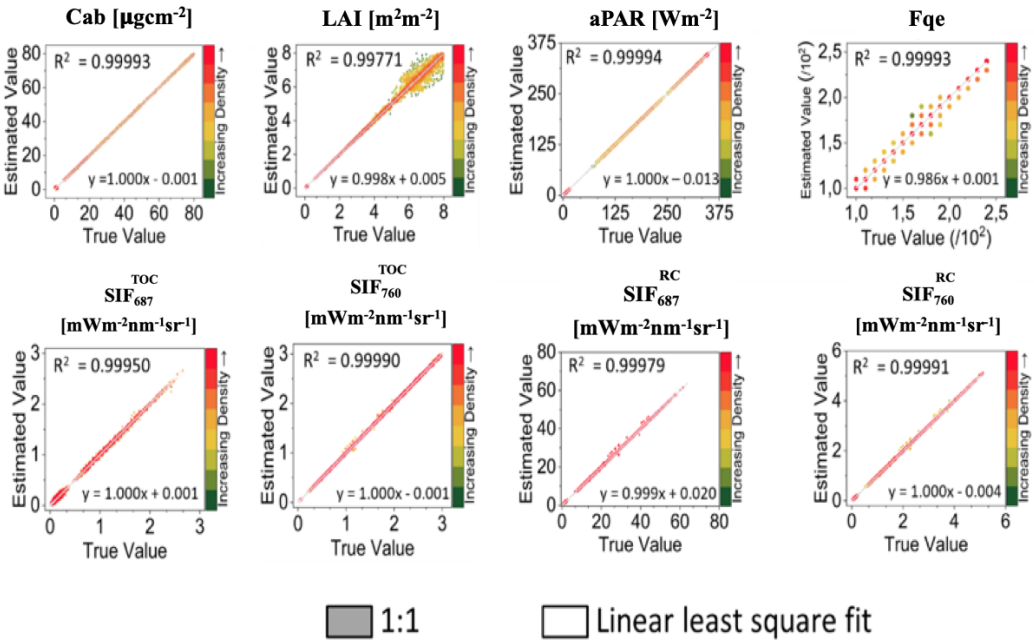

Figure 3.20: Comparisons between estimated and simulated values. The first row collects the biophysical parameters, while the second row the SIF metrics. The dashed grey line represents the 1:1. In white is displayed the linear fit. SZA = 40°.

Parameter	RRMSE%	Parameter	RRMSE%
C_{ab}	1.0 ± 0.7	SIF_{687}^{TOC}	0.7 ± 0.7
LAI	2.0 ± 0.7	SIF_{760}^{TOC}	1.0 ± 0.8
aPAR	0.7 ± 1.0	SIF_{687}^{RC}	0.9 ± 0.8
Fqe	0.8 ± 0.5	SIF_{760}^{RC}	0.9 ± 0.8

Table 3.3: RRMSE% evaluated on the comparison between the variables (biophysical parameters and SIF metrics) simulated and the values estimated with the Fourier-ML algorithm. The RRMSE% corresponds to the average value over all the SZA in the dataset, while the margin of errors to the standard deviations.

3.3.4 Qualitative comparison of the methods on field measurements

The two methods developed to correct the SIF for reabsorption were applied on the hyperspectral measurements acquired in field (experimental dataset described in Section 2.2.1). Concerning the parametric method, the local SZA, the NIR_v and the SIF_{760}^{TOC} were combined with the constant values and the regression coefficients highlighted in Section 3.2.4 to estimate the SIF^{RC} and the aPAR. In particular, the SIF^{RC} was evaluated using the a_{avg} approximation. SIF^{RC} and aPAR were then combined to estimate the SIF_{yield}. On the other hand, the experimental R_{app} represents the only input in the Fourier-ML algorithm. Its values were compared to the synthetic dataset to compute the crop biophysical parameters (C_{ab}, LAI, Fqe, aPAR) and the SIF spectra (TOC and RC). Both these approaches were applied only on the spectral measurement acquired in a $\pm$ 15 minutes interval around the solar noon. The outcomes were displayed in time to evaluate the parameters and SIF metrics behaviours at seasonal scale. Since the parametric method does not retrieve correctly the SIF^{RC} spectra for sparse vegetation, only measurements acquired on medium-dense canopies were used (DOY greater than 100). Conversely, the Fourier-ML algorithm were applied on the whole time series, opportunely expanded with additional days characterized by clear sky conditions around the solar noon. Here I would reiterate that the corn crop was not considered, as per what is displayed in Section 3.2.1. The seasonal trends obtained are displayed in Figure 3.21, in which the common metrics were compared, namely SIF_{760}^{RC}, SIF_{687}^{RC}, aPAR and SIF_{yield}. Figure 3.21A shows the parametric method results, while Figure 3.21B the Fourier-ML algorithm ones.
From a qualitative point of view, the outcomes obtained separately with the two approaches lie in the same numerical intervals. However, a more quantitative comparison was not possible. In the common days available, the trends referred to the SIF metrics evaluated on the RC full spectrum correspond to each other. In particular, the forage time series is characterized by an initial increase of the SIF values, followed by a saturation. As observed in Chapter 2, the alfalfa shown a minimum around DOY 160. The aPAR seasonal trends are completely comparable. As observed in the preliminary analysis carried out on the

synthetic dataset, the aPAR saturates when denser canopies are considered. This behaviour is clear in the last days of the forage time series and for all the alfalfa days. Here I will point out that in these temporal intervals, both these crops were in mature stage and were characterized by a well-developed canopy. Finally, the two SIF_{yield} were compared. In this case, the outcomes obtained with the parametric method differ from the SIF_{yield} computed with the Fourier-ML algorithm. The shift between the two is clear when DOY close to 100 are considered. The parametric method anomalous results are probably attributable to the strong variability affecting the SIF_{yield} for low $LAIxC_{ab}$ values, as observed in Figure 3.18 (Case 4).

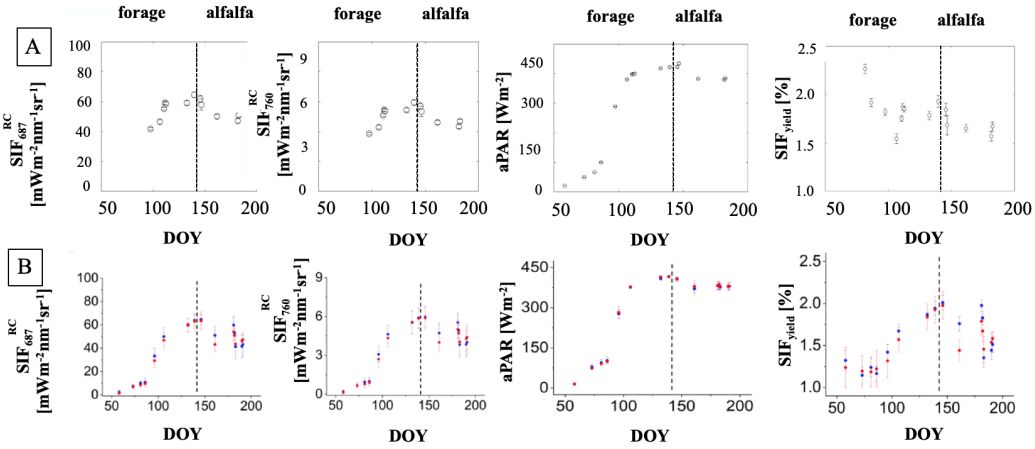

Figure 3.21: Panel A: seasonal trends obtained with the parametric method. Panel B: seasonal trends obtained with the Fourier-ML algorithm using two different approaches (considering the O_2 spectral bands, in red and excluding the O_2 spectral bands in blue). Points displayed correspond to the average value on the time interval investigated ($\pm$ 15 minutes around the solar noon). The error bars are the standard deviations. Dashed line separates the two crops investigated.

3.4 Discussion

3.4.1 Reabsorption effect on the Top-Of-Canopy SIF

Results shown in Section 3.3.1 had widely demonstrated how reabsorption process affects both the SIF^{TOC} magnitude and spectral shape. Concerning the magnitude, the fluorescence at the photosystems level is always greater compared to emission estimated at Top-Of-Canopy scale. The reabsorption and the leaf-canopy scattering determine a decrease on the SIF signal of about 10 times. Using Figure 3.10 as reference, SIF^{RC} always lies

between 0 and 80 $mWm^{-2}nm^{-1}sr^{-1}$, while SIF^{TOC} between 0 and 8 $mWm^{-2}nm^{-1}sr^{-1}$. The SIF reabsorbed in the visible red is not completely re emitted following a radiative pathway, determining a first decrease on the fluorescence signal. Furthermore, the SIF emission always occurs at longer wavelengths with respect to the absorbed ones. Therefore, the fluorescence reabsorbed in the visible red is emitted in the far-red. These wavelengths are strongly scattered by the leaf-canopy systems; therefore, a portion of the SIF signal is probably scattered outside the sensor FOV or backscattered. In terms of SIF spectral shape, the SIF^{RC} is always characterized by a predominant maximum emission in the red and a broad shoulder in the far-red, regardless of the LAI, C_{ab}, SZA, Fqe considered (see Figure 3.10 in orange). The SIF^{RC} spectral shape corresponds to the description reported in Maxwell and Johnson, 2000. Conversely, the SIF^{TOC} strongly depends on the LAI- C_{ab} values that determine a different proportion between the emission in the red and far-red (see Figure 3.10 in blue). In terms of discrete values, the emissions in the red and far-red are comparable for very low LAI-C_{ab} (e.g. first column in Figure 3.10). On the other hand, the peak magnitude mutually increases with the LAI-C_{ab}, even though this behaviour is clearer when the SIF in the far-red is considered. This trend is attributable the reabsorption, because for fixed LAI-C_{ab} the SIF^{TOC} spectral shape is unaltered, regardless of the SZA and Fqe values.

A similar qualitative analysis was carried out comparing the SIF metrics defined on the full spectra simulated at TOC and photosystems level (Figure 3.11). Regardless of the scale considered, all the metrics follow a trend similar to the one observed for the aPAR. A first almost linear increase in the SIF magnitude is observed for the values referred to sparse vegetation, then followed by saturation for medium and dense canopies. The only exception is represented by the SIF metrics defined in the visible red and at 687 nm at TOC scale (i.e. SIF^{TOC}_{red} and SIF^{TOC}_{687}). In this case, the fluorescence decreases for greater LAIxC_{ab}. The trend observed is attributable to the reabsorption that reduces the fluorescence emission in the visible red. In particular, the differences in terms of discrete values between SIF^{TOC}_{red} and SIF^{TOC}_{687} are attenuated, as previously highlighted on the experimental case reported in Chapter 2 (see Figure 2.4C). Similarly, the theoretical behaviour observed for the SIF estimated in the far-red (i.e. SIF_{760} and $SIF_{far-red}$) corresponds to the field outcomes (Figure 3.11 and Figure 2.4B). In particular, the differences in magnitude between SIF_{760} and $SIF_{far-red}$ become clearer when the vegetation changes from sparse to dense.

At the photosystems level, the peak positions in the red and far-red are unaltered regardless of the LAI, C_{ab}, SZA and Fqe. Conversely, at TOC scale peculiar behaviours are observed (Figure 3.12). As expected, the reabsorption also influences the TOC SIF spectral shape. In general, the $\lambda_{far-red}$ is characterized by a clear shift toward longer wavelengths. On the other hand, the λ_{red} initially increases when sparse vegetation is considered and then decreases toward constant value around 685 nm for higher LAIxC_{ab}. These trends estimated on the synthetic dataset correspond to the results obtained on the experimental measurements (Chapter 2, Figure 2.6). When reabsorption is less probable (low LAI and C_{ab}), $\lambda_{far-red}$ and λ_{red} increase mutually. Conversely, the probability that the SIF emitted in the red to be re absorbed increases in denser canopy. Therefore, the emission in the

far-red prevails and the $\lambda_{far-red}$ shifts toward longer wavelengths. On the other hand, the portion of the spectrum emitted in the visible red is attenuated and an opposite trend is observed. In particular, λ_{red} moves toward lower wavelengths. Constant values are reached in well developed canopies because the aPAR also saturates. In this framework, the amount of light absorbed and then dissipated radiatively is almost the same regardless of the LAIxC_{ab} considered. However, faint differences in the spectral shape are observed in the visible red, for different SZA (Figure 3.12 top right). Conversely, the Fqe affect only the SIF magnitude, but not the peak position (Figure 3.12, second row).

Finally, the SIF yields evaluated on the TOC and at the photosystems level were investigated (Figure 3.13). They show a different behaviour with increasing LAIxC_{ab} and fixed SZA and Fqe. In particular, SIF^{TOC}_{yield} decreases from sparse to denser canopies. This trend is again attributable to reabsorption because at the photosystems level the SIF_{yield} is constant (Figure 3.13A and C). The SIF^{TOC}_{yield} decreases because when reabsorption occurs a portion of the SIF is dissipated non radiatively or the signal emitted is scattered outside the sensor FOV. Averaging the SIF yields over all the LAIxC_{ab} and SZA values, a linear trend with the theoretical Fqe is observed regardless of the scale considered (Figure 3.13B and D). Despite the similar behaviour, at TOC scale the SIF^{TOC}_{yield} values lie far from the 1:1 (not shown). In particular, the slope of the linear regression performed is 0.131. Conversely, the SIF_{yield} perfectly corresponds to the theoretical Fqe and all the values lie on the 1:1. In this case, the linear regression shows a slope equal to 0.995, R^2= 1 and RMSE= $3e^{-11}$. Summarizing, the SIF^{TOC}_{yield} does not approximate sufficiently the SIF yield true values. Therefore, correcting the fluorescence spectrum for reabsorption is mandatory to improve the understanding of the relationship between SIF and plant physiological status.

3.4.2 Parametric method

A novel method to correct the TOC SIF spectrum for the reabsorption was developed and tested. This approach was implemented starting from the SIF simulated at the photosystems level obtained by means of the coupled RT models MODTRAN5-SCOPE. A preliminary analysis highlighted how a two Gaussian function adequately approximates the SIF^{RC} spectral shape. The parametric method proposed links the eq. 3.2.8 unknown parameters (a_1, a_2, b_1, b_2, c_1, c_2) to constant values, regression coefficients and TOC experimental quantities (i.e. SZA ans SIF^{TOC}_{760}). In this framework, two cases were investigated in which i) the a_1 dependence to the Fqe is completely contained in the SIF^{TOC}_{760} term (i.e. case a_{avg} corresponding to the actual experimental situation); ii) the a_1 dependence to the Fqe is contained in both SIF^{TOC}_{760} and the coefficient a (i.e. case a(Fqe) used as reference). The SIF^{RC} spectra estimated with these two approaches were compared to the simulated values computed with SCOPE.

A first qualitative analysis was carried out comparing the SIF spectral shapes (Figure 3.15). In general, a good correspondence between the simulated and estimated spectra were observed, regardless the LAI-C_{ab} considered, for fixed Fqe and SZA values. From a quantitative point of view, the RRMSE% was evaluated between the true and retrieved

cases for all the Fqe, SZA, LAI and C_{ab} composing the subset investigated. The results are summarized in Figure 3.16. Regardless of the approximation used on the coefficient a, greater RRMSE% are observed for LAIxC_{ab} values typical of sparse vegetation. In particular, the RRMSE% is up to the 40% in the a_{avg} case, 35% with the a(Fqe). Therefore, the parametric method is unable to model correctly the SIF^{RC} spectral shape when the canopies considered are not already well developed. This represents a limit of the approach proposed because unreliable outcomes are obtained under specific conditions, affecting the method application on analysis at seasonal scale. Furthermore, the RRMSE% values in the a_{avg} case clearly vary with the Fqe. These results are expected because the a coefficient shows a strong dependence on the fluorescence efficiency, with values lying between 65 and 73, as reported in Figure 3.8A. Therefore, the a_{avg} value is not completely representative of the coefficient variability with the Fqe. Nevertheless, the RRMSE% decreases with the Fqe (Figure 3.16). This trend is attributable to the a_{avg} use because in the a(Fqe) case the values obtained are always the same, regardless of the Fqe considered. Globally, not considering the sparse vegetation outcomes, the SIF^{RC} spectral shape is approximated with a relative mean square error lower than the 27% (with a_{avg}) and 25% (with a(Fqe)). Obtaining a SIF^{RC} spectrum as accurate as possible is recommended, but it is not mandatory. From an experimental point of view, it is more important that the SIF at the photosystems level is accurate around the maximum, specifically in the visible red because linked to the PSII. Here I will point out that the PSII emission is affected by the potential photodamages occurring in strong illumination conditions. On the other hand, the spectrally integrated SIF (SIF^{RC}_{INT}) is fundamental because used to evaluate the SIF_{yield}. Since SIF^{RC}_{INT} is obtained integrating over the fluorescence emission spectral interval, the discrete value estimated is more important than the SIF spectral shape. Bearing in mind these considerations, the parametric method accuracy was also investigated by means of three SIF metrics, namely SIF^{RC}_{760}, SIF^{RC}_{687} and SIF^{RC}_{INT}. The comparisons between the true values (simulated with SCOPE) and the estimated ones are displayed in Figure 3.17A. On the scatterplots, a linear regression was performed. The fitting results and statistics are collected in Figure 3.17B-E. Again, both the a_{avg} and a(Fqe) cases were investigated. When only the SIF metrics were considered, the values always fall on the 1:1 (dashed grey line). For all the cases investigated, the regression line is characterized by slopes between 0.9 and 1.1, RRMSE% between 0% and 10%, R^2 between 0.994 and 1. As expected, in case a_{avg} a dependence on the Fqe is observed, highlighted by the color code used in the aforementioned Figure. However, these outcomes are referred to all the LAIxC_{ab} values. Therefore, the weak performance of the method previously observed on the full spectrum for low LAI and C_{ab} could be potentially hidden by the average trend of the metrics. This observation is supported by the outcomes obtained on the SIF_{yield}. The SIF yield at the photosystems level was evaluated following six different approaches, as shown in Table 3.2, and they were compared to the true value, namely Fqe (see Figure 3.18). Except for Case 0, which represents our reference, in all the other cases considered a dependence on the LAIxC_{ab} is clear. In particular, a strong variability around the true value is observed when sparse vegetation is considered. This trend is obtained when at least one of the parametric methods

proposed (i.e. for the aPAR and SIF^{RC}) is used. On the other hand, the SIF_{yield} evaluated on medium-dense vegetation show less variability in terms of discrete values. Furthermore, a peculiar trend is observed when a(Fqe) and a_{avg} are considered. In particular, when the SIF^{RC} generated with SCOPE or the SIF^{RC} estimated with a(Fqe) are used, the SIF_{yield} always lie around the red line used as reference, regardless of the Fqe considered (Case 2,3,5 in Figure 3.18). Conversely, with a_{avg} the yield values are overestimated, accurately estimated and underestimated for increasing Fqe. It is expected because, by definition, a_{avg} better approximates medium Fqe (in this case Fqe= 0.015). The differences between Fqe and these six SIF yields were quantified evaluating the RMSE and RRMSE%. Figure 3.19 summarize all the results obtained displayed for increasing SZA and Fqe (color code). Focusing on the RRMSE%, as previously assessed, Case 2, 3 and 5 do not depend on the Fqe. Conversely, when a_{avg} is used, not insignificant behaviours with the Fqe and SZA are observed (Case 1 and 4). I will point out that Case 4 corresponds to the potential experimental situation in which both SIR^{RC} and aPAR are unknown. In general, averaging over all the LAI and C_{ab} the SIF yields are estimated with RMSE and RRMSE% values lower than 0.0011 and 8%, respectively.
In general, the parametric method works quite well on the synthetic dataset, even though there is a clear dependence on the LAIxC_{ab} and Fqe values. It must be pointed out that the canopy architecture simulated is the default one (spherical) and the soil contribution is very low even for medium vegetation. Conversely, for more complex canopies (i.e. erectophile like corn) the soil contribution in the signal acquired and the leaf inclination will play a key role in the relations described in the parametric method. It is plausible that the coefficients obtained in this framework will be different because, as already observed in Chapter 2, the corn SIF values are always lower compared to forage and alfalfa cases (see Figure 2.7 and Figure 2.10 as reference). Furthermore, it is possible that the relations between aPAR-NIR_v and a_1-SIF_{760}^{TOC} will be modified. I would have expected less linear trends and more scattered values, specifically for lower LAIxC_{ab}. Conversely, I think that trees characterized by well developed canopy (e.g. Oaks) would show patterns similar to those observed in the spherical case. Although I would expect comparable behaviours, the coefficient values will be probably different. Referring to Yang et al., 2015, the TOC SIF values to deciduous forest are generally lower compared to those of crops.

Finally, the parametric approaches were applied to the spectral measurements described in Chapter 2. Due to the low performances obtained on the synthetic dataset for sparse vegetation, only acquisitions performed after DOY 100 were considered. The values estimated lie in the same intervals obtained with the Fourier-ML method. However, a more quantitative comparison was not possible. In general, SIF_{760}^{RC} and SIF_{687}^{RC} follow the same seasonal trends observed at TOC scale. The aPAR evolution in time is reasonable because lower values are estimated for low PAR and underdeveloped crops (beginning of the forage time series). Conversely, the SIF_{yield} shows anomalous behaviour specifically around DOY 100. Since both aPAR and SIF^{RC} are affected by higher errors for low LAI and C_{ab}, the yield obtained by their combination is not reliable.

3.4.3 Fourier-ML algorithm

The Fourier-ML algorithm allows to retrieve both biophysical parameters of interest along with the SIF spectrum at TOC and corrected for the reabsorption. The method accuracy was quantified on the synthetic dataset. Each parameter estimated is retrieved with an average RRMSE% lower than the 2% (see Table 3.3). The linear correlation found corresponds to the outcomes reported in Liu et al., (2019) in which the SCOPE simulations were used to train a random forest to predict SIF at photosystem, leaf and canopy level. Focusing on the photosystem scale only, the Fourier-ML algorithm shows best performances compared to Liu et al., (2019). Regardless of the wavelength considered (far-red or red), Fourier-ML outcomes always lie in the 1:1 (slope around 1, intercept close to 0 and R^2 equal to 0.9998). Conversely, with the random forest model, in the far-red the scatterplot between the reference and the retrieved SIF is characterized by a slope equal to 0.96, intercept 0f 0.561 and R^2=0.95. In the visible red, the slope is equal to 0.983, the intercept to 1.237 and the R^2=0.861. Concerning the Fourier-ML algorithm application on the experimental spectral measurements, the seasonal trends obtained are reasonable (Figure 3.21). As previously observed with the parametric method, lower aPAR values are observed in the first days of the forage time series, then it saturates to a constant value. This trend corresponds to the aPAR theoretically evolution highlighted on the synthetic dataset. In particular, it increases mutually with the canopy chlorophyll content and then saturates when denser canopies are considered. As demonstrated in Chapter 2, forage after DOY 112 and the alfalfa were characterized by a well developed canopy. Concerning the SIF^{RC} metrics, they both follow the aPAR seasonal trend, as already highlighted in Figure 3.11. Furthermore, since they are evaluated on the SIF spectrum at the photosystems level, they are not affected by reabsorption. Therefore they share the same seasonal evolution, regardless of the spectral interval considered. Concerning the SIF_{yield}, its qualitative trends corresponds to the SIF^{RC} ones. Lower Fqe are observed in correspondence to lower SIF^{RC} values and vice versa. This behaviour is reasonable because, by definition, the SIF_{yield} represents the efficiency of the fluorescence process. Moreover, while in the parametric method these two variables (i.e. SIF^{RC} and SIF_{yield}) are obtained one from the other, in the Fourier-ML algorithm they are estimated simultaneously and separately. For this reason, the qualitative agreement between their seasonal trend (second row in Figure 3.21) supports the robustness of the Fourier-ML algorithm.

3.5 Conclusions

The analysis carried out on the synthetic dataset was highlighted how the reabsorption process strongly affects the TOC SIF magnitude and spectral shape. As expected the main differences between the fluorescence signal at the photosystems level and Top-Of-Canopy scale are clearer in the visible red spectral region. As a consequence, the physiological information linked to the PSII inferred at TOC scale is dramatically influenced by the

reabsorption effect. Furthermore, the SIF^{TOC}_{yield} does not approximate adequately the true yield evaluated at the photosystems. For all these reasons, correcting the SIF spectrum for the reabsorption is pivotal in order to exploit the fluorescence emission as a reliable physiological proxy. In this regard, two different approaches were developed and tested. The parametric method performances depend on the canopy development stage. In particular, for low LAI and C_{ab}, corresponding to sparse vegetation, the SIF^{RC} obtained is not reliable. This behaviour limits the method application for seasonal scale analysis, as demonstrated in the experimental case investigated. Therefore, the method needs to be further implemented in order to minimize the error associated to the LAIxC_{ab} variability. Conversely, the Fourier-ML algorithm represents a valid approach to retrieve both biophysical parameters and the SIF spectra (TOC and RC) with a good accuracy. However, its performances strongly depend on the dataset used to train the algorithm. In the case shown here, the canopy geometry, crop type and atmosphere composition were kept fixed. Future studies could be devoted investigating whether these three parameters influence the method accuracy.

Chapter 4

Advancing SIF retrieval in clear lake waters and development of primary production models

4.1 Theoretical Background

Chl-a, contained in many phytoplankton taxa, represents a valid proxy to evaluate the trophic status of water bodies and to detect potentially harmful blooms (Ruddick et al., 2019). As commonly known, Chl-a is also a photosynthetic pigment. During the photosynthesis, the solar light absorbed by the Chl-a, is converted in chemical energy (light reaction process). This energy is then used in the carbon assimilation reaction, in which the atmospheric carbon dioxide (CO_2) is used to produce sugar (Porcar-Castell et al., 2014). This process is known as primary production (PP). Chl-a concentration ([Chl-a]) is therefore correlated to the phytoplankton biomass and it can be used for primary production (PP) estimations (Falkowski and Kiefer, 1985). Appraisals of the phytoplankton PP are pivotal because this parameter is generally associated to the ecosystems food webs and the global carbon cycle (Lakshmi et al., 2014).

For over four decades, the [Chl-a] has been estimated from remote sensing measurements (Smith et al., 1989) exploiting Ocean Color (OC) algorithms. Water bodies are dynamic ecosystems therefore, the signal arising from its surface depends on the water itself and its compounds inherent optical properties (Lakshmi et al., 2014; Bukata et al., 2018; Gitelson et al., 2009). OC algorithm predicts the near-surface [Chl-a] combining empirical coefficients with the remote sensing Reflectance (R_{rs}), the latter one evaluated in specific spectral intervals. Wavebands positions are selected considering the spectral features of the acquired signal. Specifically, for OC algorithm, the R_{rs} is evaluated in the

blue range, where Chl-a pigments efficiently absorb the solar light, and in the green range, because less influenced by the Chl-a contributions. Although [Chl-a], obtained from OC algorithm, is fundamental to monitor the water body's quality, it must be remembered that it is calculated from the R_{rs} that does not carry information about the photosynthesis and it is less linked to the phytoplankton biomass. In this regard, an alternative approach is represented by the use of phytoplankton SIF emission. SIF is a side product of the photosynthesis. It is closely interrelated to the [Chl-a] because the emitted signal is proportional to the pigments that actually absorb the solar light. Moreover, it arises from the core of the photosynthetic machinery, then it is more linked to the phytoplankton biomass and primary production. SIF is an optical signal, so it can be potentially detected by sensors mounted on different platforms (Mohammed et al., 2019). As highlighted in Chapter 1, in aquatic ecosystem the SIF retrieval strongly depends on the investigated water body, along with the sensor used characteristics (e.g. spectral resolution, signal-to-noise ratio etc.). Focusing on Case-2 waters, lakes are typically characterized by a high degree of spatio-temporal variability of the inherent optical properties and related compounds (Cesana et al., 2021). At hourly and daily scales, the acquired signal is mainly determined by the phytoplankton dynamics, whose trend depends on light and nutrients availability and variations in the suspended matter concentrations, linked to the wind-induced re-suspension of sediments, that is typical of shallow lake (Cesana et al., 2021). To discern between abiotic and biotic contributions, hyperspectral and high frequency measurements are fundamental. From one side, hyperspectral resolution potentially improves the SIF retrieval in optically complex waters, to the other, high frequency measurements better follow the spatio-temporal dynamism, typical of inland waters. Obtaining a reliable SIF proxy in Case-2 waters is pivotal in terms of phytoplankton primary production (PP) estimations. In literature, there are examples of PP estimations carried out in lakes and exploiting remote sensed quantities. For instance, Deng et al., (2017), applied the Vertical Generalized Productivity Model (VGPM) on MODIS data to investigate the long-term variations in PP in Lake Taihu (China) (Cesana et al., 2021). Although the good results achieved, the method described does not include the SIF in its formulation.

To better understand the potentiality of the SIF use in the inland water characterization a one-week experiment was conceived in the Lake Maggiore (Italy) in July 2019. Hyperspectral measurements were automatically collected during the day by an optical sensor mounted on a buoy left far enough from the lake shore (around 50 meters). The obtained time series were then exploited to implement the FLH algorithm for inland waters applications. Furthermore, indices and metrics evaluated on the hyperspectral measurements were combined with the values obtained from water samples analysis in order to define potential phytoplankton PP models. The models reliability was preliminary assessed by comparison with the biovolume.

4.2 Material and Methods

4.2.1 Measurements description

Measurements and water samples analysed hereinafter were collected in Ghiffa (VB), a site located on the Lake Maggiore western shore. The field campaign was carried out in July 2019. Hyperspectral and high frequency measurements were collected from the 02^{nd} up to the 07^{th} of July. Furthermore, water samples were collected *in-situ* the 02^{nd} and 03^{rd} of July. A detailed description of the investigated water body, the sampling performed and the acquired hyperspectral measurements are reported in Section 4.2.1.1, Section 4.2.1.2 and Section 4.2.1.3, respectively.

4.2.1.1 Study Area

Lake Maggiore is a large oligo-mesotrophic lake situated between Italy (ca. 80%) and Switzerland (ca 20%). Placed at the south of the Alps, its maximum depth reaches the 370 meters. The lake has 33 tributaries and only one emissary, the River Ticino (Cesana et al., 2021). Formed by glacial erosion in a pre-existing fluvial valley, it is considered holo-oligomictic and rarely undergoes to a complete mixing (Barbanti and Ambrosetti, 1989). Like most lakes in Italy and Central Europe, Lake Maggiore underwent to anthropogenic eutrophication during the second half of the 20^{th} century (Salmaso and Mosello, 2010) with peak of phosphorus (P) loads at mid-seventies (Ruggiu et al., 1998). In the following decades, sewage treatment plants were improved and total phosphorus in detergents were reduced, until values close to pre-industrial concentrations were reached (Marchetto et al., 2004).
Lake Maggiore can be considered a clear waters lake. It has been monitored since 1970s and a comprehensive long-term dataset of the phytoplankton records and environmental variables is currently available (Morabito et al., 2018). As reported in Morabito et al., (2018), the phytoplankton response to weather conditions depends on the groups considered. Concerning Cyanobacteria, rainfall positively affects growth, while higher wind speed has a negative impact on their development. Diatoms reach a maximum growth in spring, when the wind influences the mixed regime and the nutrient replenishment. Finally, also water temperature and light intensity are pivotal parameters influencing the *Mougeotia sp.* (Chlorophytes) and Cyanobacteria growth.

4.2.1.2 Water samples collection

The water sampling was carried out with the purpose of characterize the water body bio-optical properties and to apprise the phytoplankton composition. In order to evaluate the phytoplankton response to different illumination conditions and its behavior within the water column, several samples were collected in different depths and moments of the day. The sampling was repeated in correspondence of the solar noon and close to the sunset.

For each sampling performed, water samples were collected just below the surface z_0 (0.2-0.5 m) and at the Secchi Disk depth z_{SD} (around 5.5-6 m). The following nomenclature was used to discern between the different samples collected: S1 and S2 correspond to sampling on 02^{nd} of July, while S3 and S4 on 03^{rd} of July. Odd numbers refer to the measurements undertaken close to the solar noon, even numbers refer to the sunset ones. The additional labels, z_0 and z_{SD} (i.e. $S1z_0$ and $S1z_{SD}$), indicate the depths at which the samples were collected. The nomenclature is also summarized in Table 4.1. All the samples were filtered *in-situ* with a GF/F glass fiber filters and subsequently analyzed in laboratory. Moreover, to improve the matching with the hyperspectral measurements, water samples were collected close enough to the buoy on which the sensors were mounted on.

Sample ID	Sampling day	Sampling time	Sampling depth
$\mathbf{S1}_{z0}$	02/07/2019	~11:44	below the surface
$\mathbf{S1}_{zSD}$	02/07/2019	~11:44	Secchi Disk depth
$\mathbf{S2}_{z0}$	02/07/2019	~16:25	below the surface
$\mathbf{S2}_{zSD}$	02/07/2019	~16:25	Secchi Disk depth
$\mathbf{S3}_{z0}$	03/07/2019	~11:20	below the surface
$\mathbf{S3}_{zSD}$	03/07/2019	~11:20	Secchi Disk depth
$\mathbf{S}_{z0}$	03/07/2019	~15:10	below the surface
$\mathbf{S}_{zSD}$	03/07/2019	~15:10	Secchi Disk depth

Table 4.1: Nomenclature used in Chapter 4 to discern between the different water sample collected. The temporal information is reported in local time (UTC+1).

4.2.1.3 Hyperspectral measurements acquisition

Hyperspectral measurements were collected by a ROX spectrometer manufactured by the JB Hyperspectral Devices (Germany). This system is equipped with sensors specifically designed for hyperspectral and continuous measurements in the VIS-NIR. The ROX employs an Ocean Optics spectroradiometer collecting the incoming and the upwelling irradiance/radiance almost simultaneously in the wavelength interval between 400-950 nm, with a spectral resolution of 1.5 nm and acquisition time of about 1 minute. The ROX system consists in two separated probes. The first one pointed upward to collect the downwelling irradiance (E_d) reaching the target. The E_d optic was mounted on a goniometer to keep the probe as perpendicular as possible to the water surface regardless of the buoy oscillations. The second sensor was pointed downward to measure the upwelling radiance (L_u) rising from the water body. To avoid the solar glint, the sensor was placed

below the water surface (ca. 15 cm) and pointed toward the lake bottom. The instrument was mounted on a flouting buoy left far enough from the coastal zone (around 50 meters) to avoid the bottom contribution in the spectral signal acquired.
The experiment location and the experimental set-up is shown in Figure 4.1.

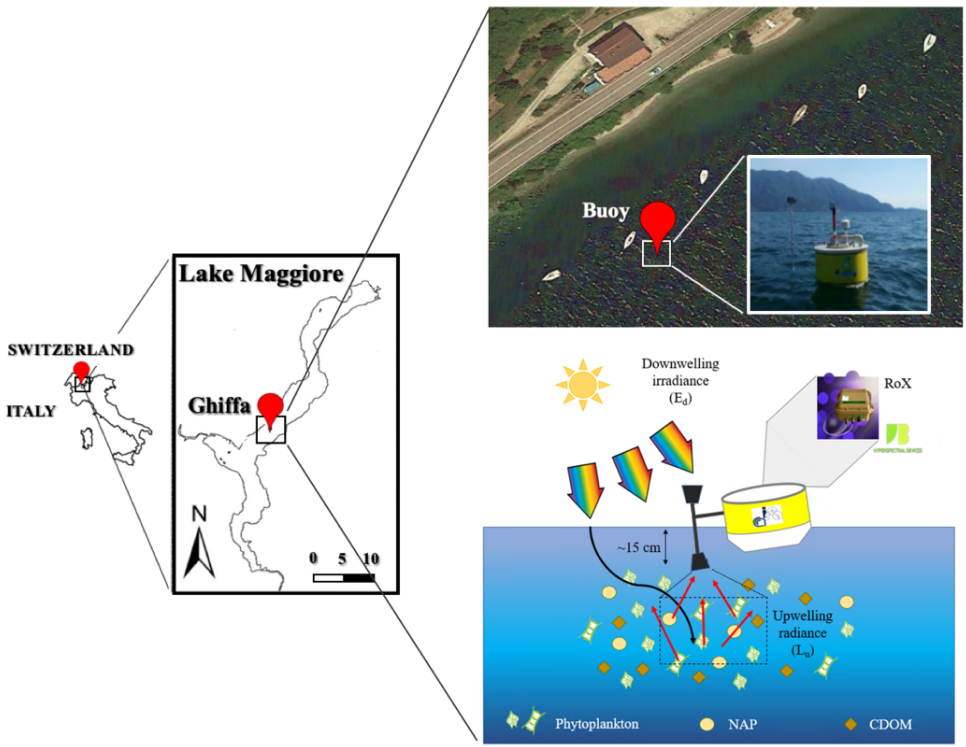

Figure 4.1: Lake Maggiore and experimental location. (top right) Buoy on which ROX has been mounted on. It was left around 50 meters far from the lake shore to avoid the bottom contribution in the spectra collected. (bottom right) Experimental set-up.

To optically characterize the water column, a vertical profile of the downwelling irradiance and upwelling radiances was acquired by means of the Satlantic radiometer. Satlantic is equipped with radiometers that collect the E_d and L_u between 350.5-796.5 nm (137 channels, 3.3 nm) at several depths with an acquisition step of 0.1 m. In the upcast mode, used in this work, the sensor starts collecting the incoming irradiance from the surface up to greater depths. The vertical profile acquisition was carried out in conjunction with the water samples collection, therefore the 02^{nd} and 03^{rd} of July.

4.2.2 Measurements analysis

4.2.2.1 Water samples laboratory analysis

As described in Bresciani et al., (2016), the photosynthetic pigments for high-performance liquid chromatography (HPLC) analysis were extracted in 90% acetone, overnight in the dark, under nitrogen. The obtained extract wasn used to quantify the Chlorophyll-a concentration ($[\text{Chl-a}]_{HPLC}$, in $\mu g/L$), its derivatives and the total carotenoids by spectrophotometry (Lorenzen, 1967). Individual carotenoids were detected by revers-phase HPLC with an Ultimate 3000 (Thermo Scientific). Specific pigments were identified by ion pairing, reverse-phase HPLC described in (Guilizzoni et al., 2011). Total Suspended Matter (TSM) was obtained gravimetrically (Strömbeck and Pierson, 2001). Backscattering was measured using a Hobi Labs Hydroscat-6. The spectral absorption coefficients of phytoplankton (a_{phy}) and non-algal particles (a_{NAP}) (Kishino et al., 1985) were obtained spectrophotometrically using the filter pad technique (Trüper and Yentsch, 1967). CDOM was measured spectrophotometrically immediately after filtration through Whatman Nucleopore membrane filters (diameter 47 mm, pore size 0.2 μm). The CDOM absorption coefficient at 440 nm ($a_{CDOM}(440)$) was derived according to Kirk (2011).

Phytoplankton samples were collected and analyzed for the purposes of species identification and cell count under an inverted microscope (400x magnification; Utermöhl 1958).

The Chl-a SIF emission quantum yields (Φ_{SIF}) were estimated by means of an analysis of the excited state decay relaxation, employing a laboratory-assembled time-correlated single photon counting apparatus, as described in Remelli and Santabarbara (2018). In brief, excitation is provided by a pulsed laser diode (PicoQuant 800B), centred at 632 nm, at a repletion rate of 20 MHz, and an intensity of 1 pJ/pulse. Emitted photons are collected with right-angle geometry, through a monochromator (Jasco JT-10) and multichannel-plate photomultiplier (Hamamatsu, R5916U-51). Acquisition electronic, consisting in Time-to-Amplitude Converter (TAC), Constant-Fraction timing Discriminator (CFD) and multichannel timing analyser (MCA) are embedded and controlled by PC-mounted acquisition board (Becker & Hinkl, SPC-330). Samples were resuspended from the sampling-filters, in Bold's basal growth media, and diluted to an OD equivalent to 0.05 cm^{-1} at 680 nm, before the measurements, placed in 3 mm path-length cuvette. To attain the Φ_{SIF}, 10 μM of the inhibitor DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea) was added to the samples. The measuring conditions avoid artefacts due to reabsorption of emitted photons. Therefore, the values obtained correspond to the real fluorescence yield/efficiency. All decay traces were collected at 682 nm (FWHM 3 nm) to obtain at least 2X104 counts at the peak channel. Signal are fitted with an iteration-reconvolution routine, accounting for the instrument response function (120 ps), which is measured using a scattering solution (Ludox), and using a linear combination of exponential decay as the model kinetic function by means of a laboratory written software, as described in Santabarbara et. al., (2017). The SIF yield is retrieved from the estimation of the mean

Sample ID	$[\text{Chl-a}]_{HPLC}$ [µg/L]	biovolume [mm^3m^{-3}]	Φ_{SIF} [-]
$S1_{z0}$	1.68±0.55	1740	0.062±0.003
$S1_{zSD}$	1.15±0.42	2137	0.065±0.002
$S2_{z0}$	1.92±0.28	659	0.068±0.003
$S2_{zSD}$	2.56±1.00	2266	0.077±0.004
$S3_{z0}$	1.60±0.69	1020	0.068±0.002
$S3_{zSD}$	2.07±0.77	1158	0.060±0.004
$S4_{z0}$	1.74±0.76	526	0.060±0.003
$S4_{zSD}$	2.46±0.59	968	0.081±0.004

Table 4.2: Values obtained from the laboratory analysis. The $[\text{Chl-a}]_{HPLC}$ values shown in the table have been evaluated averaging the two replica available. The errors associated correspond to the standard deviation.

decay lifetime starting from the fit parameter ($\tau_m = \sum_{i=1} A_i\tau_i^2 / \sum_{i=1} A_i\tau_i$, where τ_i and A_i are the lifetimes and associated amplitude, respectively) and using the decay of pure Chl-a dissolved in dry methanol as a reference (monoexponential τ = 4.1 ± 0.2 ns).

The laboratory values used in this chapter are summarized in Table 4.2.

4.2.2.2 Hyperspectral measurements analysis

As described in Section 4.2.1.3, L_u was acquired below the water surface in order to avoid the solar glint contribution in the acquired spectra measurements. In view of future RS applications, in which the signal is collected above the water surface, L_u was corrected to obtain the water leaving radiance (L_w). L_u measured by the ROX, was multiplied by the constant factor of 0.543 (Zibordi et al., 2012) to consider the water-air interface. According to Austin (1974), this coefficient is assumed be wavelength independent. The L_w obtained with this procedure was used hereinafter. The R_{rs} was evaluated dividing the E_d by L_w. Since L_w is measured in $Wm^{-2}nm^{-1}sr^{-1}$ and E_d in $Wm^{-2}nm^{-1}$, the R_{rs} is reported in sr^{-1}.

A quality criterion was used to remove from the E_d, L_w and R_{Rs} times series the measurements collected under not stable illumination. Similarly to the FLoX, the ROX automatically evaluates the differences between two consecutive L_d acquisition lying in the same collection cycle. As previously described in Section 2.2.1, measurements characterized by differences greater than the 1% were deleted from the time series. The filtered time series obtained was then averaged over a temporal interval of 10 minutes. The errors associated to the single averaged value corresponds to standard deviations.

4.2.2.3 Novel FLH parametrization exploiting hyperspectral measurements

The FLH method developed by Gower (1980), is still frequently used today (Gupana et al., 2021) as a valid approach to evaluate a proxy for the fluorescence emission in aquatic ecosystems. A virtual baseline is obtained connecting two wavebands selected in the spectral regions not influenced by the SIF (Letelier and Abbott, 1996; Xing et al., 2007). In particular, they are respectively placed on the left and on the right of the central band linked to the SIF emission. The virtual linear baseline obtained corrects the SIF for the scattering contribution (Gower et al., 2004). The SIF proxy is then evaluated at a selected wavelength, subtracting the above-mentioned baseline from the central band linked to the SIF. The central band position and width need to be chosen carefully in order to exclude the absorption feature linked to the atmospheric oxygen band at 687 nm (O_2-B) (Letelier and Abbott, 1996; Xing et al., 2007). The SIF proxy obtained with the FLH algorithm (SIF_{FLH}) is usually evaluated on the L_w spectrum and can be summarized in two leading equations:

$$SIF_{FLH} = Lw_C - Lw_{baseline} \tag{4.2.1}$$

$$Lw_{baseline} = Lw_L - (Lw_R - Lw_L)\frac{\lambda_C - \lambda_R}{\lambda_L - \lambda_R} \tag{4.2.2}$$

where λ_C corresponds to the SIF central waveband, while λ_L and λ_R are the left and right ones. L_{wC}, L_{wL}, L_{wR} are the corresponding water leaving radiance.

As highlighted in Chapter 1, the water body's composition heavily influences the SIF retrieval, specifically in optically complex waters. In trophic lakes, the scattering produced by high Chl-a and TSM concentrations dominates the signal acquired in the visible red, hiding the SIF contribution at 685 nm (Gilerson et al., 2007). Conversely, as occurs in Lake Maggiore, under more oligotrophic conditions (i.e., [Chl-a] lower than 3-5 mgm^{-3}), the Chl-a SIF signal in the red-NIR is less affected by CDOM and NAP (Non Algal Particles) residual absorption (Dall'Olmo et al., 2003; Gilerson et al., 2010). In this case, the FLH assumption of a straight baseline is valid (Gower et al., 2004). For this reason, the FLH algorithm was selected for this case of study.

When applied on RS measurements, the FLH accuracy is strongly influenced by the considered water-body (i.e. Case-1 rather than Case-2 waters), the wavebands positions (fixed and linked to sensors characteristics, e.g., MODIS, MERIS, OLCI), the spectral (usually medium) and spatial (pixel size ~km) resolutions. Since the performed analysis were carried out on proximal sensed measurements and on a single water body, only the wavebands' position limit was investigated. Hyperspectral sensors provide hundreds of spectral bands (Gupana et al., 2021). Thus, the wavebands in eq.4.2.2 can be dynamically chosen in agreement with the changes observed on the experimental spectra at diurnal scale (see Figure 4.2A). Spectral changes can be due to different illumination conditions occurring during the day but also to variations in the water body composition (e.g. phytoplankton migration within the water column).

The novel FLH parametrization here proposed, evaluates a SIF proxy selecting dynamically the wavebands' positions. Spectral changes are clearer in the R_{rs} spectrum than in

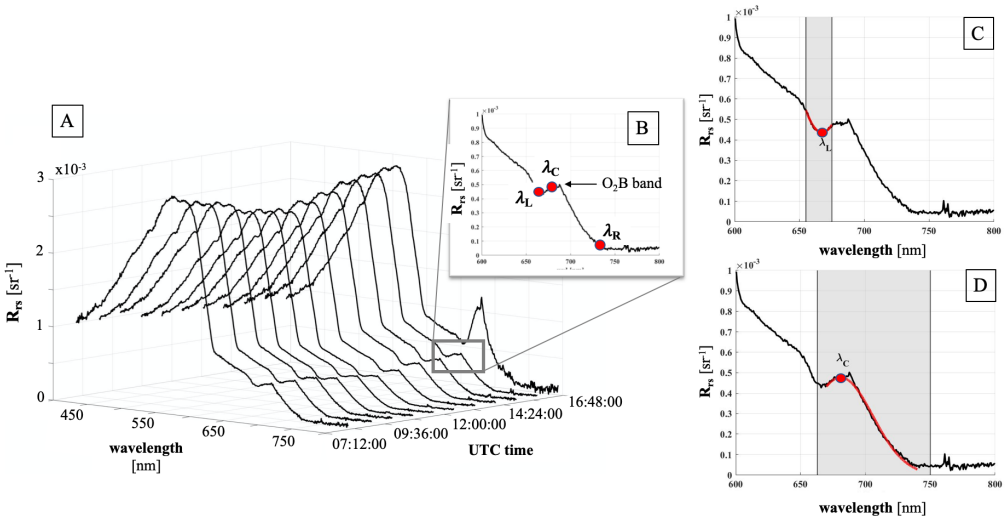

Figure 4.2: (A): R_{rs} spectra times series. (B): λ_C, λ_L, λ_R positions in the R_{rs} spectrum (red points). The two arrows highlight the artefacts due to O_2 absorption bands. (C): the grey area shows the spectral range in which the λ_L was searched for. The red line represents the second-degree polynomial used as fit function. (D): the grey area shows the spectral range in which the λ_C was searched for. The red line corresponds to the Gaussian used as fit function.

the L_w. Focusing on the R_{rs} spectra (shown in Figure 4.2A), it is clear that the R_{rs} changes during the day. Thus, the wavebands' positions were preliminary selected on the R_{rs}.

λ_C and λ_L, used in eq. 4.2.2, are obtained following a two steps procedure. At first, the corresponding "wide" wavebands were opportunely identified on the R_{rs} spectrum (grey regions in Figure 4.2C and 4.2D). λ_C should be as close as possible to λ_L, to remove the interference due to the CDOM and NAP but, at the same time, should be in a spectral region minimally affected by the Chl-a absorption. The interval between 655-675 nm was selected for λ_L, while the interval from 663-750 nm for λ_C. Then, λ_C and λ_L positions were retrieved fitting the R_{rs} spectrum in the previously identified intervals, with appropriate functions. λ_L corresponds to the minimum due to the Chl-a absorption and it was obtained fitting the R_{rs} spectrum with a second-degree polynomial. Conversely, λ_C is related to the SIF maximum emission. Its position was obtained fitting the spectrum with a Gaussian function. The wide spectral interval used in this framework allows the Gaussian to fit also the R_{rs} far-red tail and then overlook the artefact around 687 nm liked to O_2-B band.
This procedure was iterated on each experimental spectrum acquired, but only the outcomes showing a R^2s greater than 0.70 were considered reliable. The remaining values, and corresponding spectra, were deleted from the time series. λ_R was kept fix to 730 nm.

According to Kritten et al., (2020) at this wavelength both the CDOM absorption and the NAP scattering are almost negligible.
SIF_{FLH} was obtained replacing the λ_C, λ_L, λ_R values in eq. 4.2.2. The corresponding L_ws values were obtained averaging the radiance on the wavebands centered in λ_C, λ_L, λ_R, respectively. The bands' widths were calculated as the standard deviation of all theλ_L and λ_C evaluated, while for the λ_R, a nominal band width of 4 nm was chosen. The method implemented is schematically presented in Figure 4.2.

4.2.2.4 Spectral indices evaluated on hyperspectral measurements

To characterize the phytoplankton diurnal dynamics, ancillary indices were evaluated on the hyperspectral measurements.

To account the amount of light reaching the target, the incoming irradiance (E_d) was spectrally integrated over the PAR (i.e. 400-700 nm). The obtained values are indicated as E_{PAR}.

The apparent reflectance evaluated at 550 nm (R^{app}_{550}) was used to approximate the amount of light reflected by the target. R^{app} was obtained as the ratio between the incoming (L_d) and the water leaving radiance (L_w).

To predict the near-surface [Chl-a], an Ocean Color (OC) algorithm was used. Conventionally, the OC algorithm consists in a fourth-order polynomial relationship between a ratio of R_{rs} and [Chl-a] (Cesana et al., 2021). Specifically, the ratios of the R_{rs}, evaluated in opportunely selected spectral regions, are combined with four empirical regression coefficients. According to the number of bands used, the OC can be called OC_2, OC_3 and OC_4, respectively. Even though OC approach was originally formulated for mainly Case-1 applications, it has been used on the hyperspectral measurements acquired. This approximation is valid because Lake Maggiore is considered as a clear water lake. Laboratory analysis highlighted [Chl-a] and TSM concentrations less then 3 μgL^{-1} and 1.50 mgL^{-1} respectively and a CDOM value of 0.04 m^{-1}. To better resolve the phytoplankton temporal variability, the OC_4 was used. The selected bands are 443 nm, 490 nm, 510 nm and 555 nm. According to O'Reilly et al., (2000) procedure, the maximum band ratio (MBR) was determined as the greater of the $R_{rs}(443)/R_{rs}(555)$, $R_{rs}(490)/R_{rs}(555)$ and $R_{rs}(510)/R_{rs}(555)$ (see eq. 4.2.4). Empirical coefficients (a_0-a_4) used in eq. 4.2.3 are 0.1731, -3.9630, -0.5620, 4.5008, -3.0020. In agreement with the nomenclature described in O'Reilly et al., (2000), the following equations were used to evaluate the [Chl-a]:

$$[Chl-a]_{OC4} = 10^{a_0+a_1R_4+a_2R_4^2+a_3R_4^3+a_4R_4^4} \tag{4.2.3}$$

where:

$$R_4 = \log_{10}[max(R_{rs}{}^{443}_{555}, R_{rs}{}^{490}_{555}, R_{rs}{}^{510}_{555})] \tag{4.2.4}$$

4.2.3 Phytoplankton primary production (PP) in inland waters

Phytoplankton are the primary producers that serve as the base of the aquatic food web (Li et al., 2021). Through the biochemical process of photosynthesis, they contribute roughly half of the primary production (PP) on Earth (Buitenhuis et al., 2013; Field et al., 1998). Phytoplankton photosynthesis in inland lakes has nontrivial feedback effects on the global carbon cycle via regulation of regional carbon exchanges, which are profoundly impacted by anthropogenic factors (Li et al., 2021). Gross primary production (GPP) is defined as the total amount of carbon fixed during the photosynthetic process. However, the proposed analysis was carried out at local scale. All the quantities evaluated in this Chapter are referred to the investigated water volume. They are indicated with the letter F, in agreement with the nomenclature proposed by Kiefer et al. (1989) and Morrison (2003).

4.2.3.1 PP definition exploiting in-situ quantities

According to Kiefer et al. (1989) and Morrison (2003), the SIF emitted (F_{SIF}) and the carbon assimilated (F_C) by the phytoplankton contained in a unit of water volume can be evaluated as follows:

$$F_{SIF} = \Phi_{SIF} * F_A \tag{4.2.5}$$

$$F_C = \Phi_C * F_A \tag{4.2.6}$$

where (Φ_{SIF}) is the quantum yield of fluorescence, F_A is the light absorbed by the phytoplankton and Φ_C is the yield of carbon fixation. The F_C calculated with eq. 4.2.6 is in molcm^{-3}s^{-1} and can be considered as a proxy of the water body PP. Similarly to the terrestrial vegetation, the Chl-a contained in the phytoplankton absorbs the solar light in the PAR. This information can be used to rewrite the F_A term. Specifically, the amount of light absorbed by the water body in the PAR can be approximated as the product between the phytoplankton spectral absorption coefficient a_{phy} and the irradiance reaching the target (E_d). Both of them depend on the wavelength and the depth. Therefore, for a generic depth z:

$$F_A(z) = \int_{400}^{700} a_{phy}(\lambda, z) E_0(\lambda, 0^-) e^{-k(\lambda)z} d\lambda \tag{4.2.7}$$

where a_{phy} is measured in m^{-1}, $E_0(\lambda, 0^-)$ represents the irradiance evaluated just below the water surface (in Wm^{-2}nm$-$1), while k is the coefficient accounting the light attenuation with the depth (in m^{-1}). Replacing eq. 4.2.7 in eq. 4.2.5, an alternative expression for F_{SIF} is obtained:

$$F_{SIF}(z) = \Phi_F(z) \int_{400}^{700} a_{phy}(\lambda, z) E_0(\lambda, 0^-) e^{-k(\lambda)z} d\lambda \tag{4.2.8}$$

The right side of eq. 4.2.8 can be evaluated exploiting the measurements acquired *in-situ*. Specifically, Φ_{SIF} and a_{phy} were obtained from the laboratory analysis described in Section 4.2.2.1. Conversely, E_0 and k can be retrieved by the irradiance vertical profile carried out with the Satlantic radiometer. The experimental k was obtained following the protocol described by Mishra et al., (2005). Briefly, assuming the illumination on the water body stable during each Satlantic acquisition (lasting in less than 10 minutes) and k independent with the depth, then E_d decrease at a specific wavelength within a general water column can be written as follows:

$$E_d(\lambda) = \int_{z_0}^{z_m} E_0(\lambda, z) e^{-k(\lambda)z} dz \tag{4.2.9}$$

where z_0 represents the depth (in meters) closest to the surface (but below the water-air interface) while z_m is chosen deeper than z_0. Resolving the integral respect to z:

$$-k(\lambda)(z_m - z_0) = \ln\left(\frac{E_d(\lambda, z_m)}{E_d(\lambda, z_0)}\right) \tag{4.2.10}$$

$E_d(z_0)$ and $E_d(z_m)$ correspond to the downwelling irradiances acquired by the Satlantic at z_0 and z_m, respectively at a specific wavelength. k is obtained as the slope of the linear regression performed on the comparison between (z_m – z_0) and the logarithm of the irradiance ratio showed on the right side of eq. 4.2.10. The k spectrum over the PAR was obtained experimentally iterating the procedure for the different wavelengths in the spectral interval from 400 to 700 nm. The F_{SIF} at z_0 and z_{SD} was obtained replacing the corresponding Φ_{SIF}, a_{phy}, E_d in eq.4.2.8 and integrating the result over the PAR. Since Φ_F is unitless, a_{phy} is expressed in m^{-1} and E_d is an irradiance in $Wm^{-2}nm^{-1}$, it turns out the F_{SIF} is measured in Wm^{-3}.

The same rationale was applied on F_C obtaining:

$$F_C(z) = \Phi_C(z) \int_{400}^{700} a_{phy}(\lambda, z) E_0(\lambda, 0^-) e^{-k(\lambda)z} d\lambda \tag{4.2.11}$$

Unlike the F_{SIF} case, in eq. 4.2.11 two parameters are unknown, namely F_C and Φ_C. Nevertheless, F_C was replaced by the phytoplankton biovolume (in mm^3m^{-3}), assuming that they are positively correlated to each other. This approximation is supported by studies (e.g. Deng et al., 2017; Bergamino et al., 2010) demonstrating that the phytoplankton biomass (and then the biovolume) is the most important factor affecting the temporal variations of phytoplankton primary production (Cesana et al., 2021). At last, an experimental proxy for the carbon fixation ($\Phi_C^{'}$) was evaluated from eq. 4.2.11:

$$\Phi'_C(z) = \frac{biovolume(z)}{\int_{400}^{700} a_{phy}(\lambda, z) E_0(\lambda, 0^-) e^{-k(\lambda)z} d\lambda} \tag{4.2.12}$$

where, again, the two depths considered are z_0 and z_{SD}. Due to the approximation previously introduced, $\Phi_C^{'}$ is measured in mm^3W^{-1}.

4.2.3.2 PP models exploiting hyperspectral measurements

Eq. 4.2.6 exploited to evaluate the carbon assimilation in the water body is similar to the one used to calculated the GPP in the terrestrial vegetation (e.g. Porcar-Castell et al., 2014; Rossini et al., 2010) according to the LUE (Light Use Efficiency) model:

$$GPP = LUE * aPAR \tag{4.2.13}$$

where aPAR is the absorbed light in the Photosynthetically Active Radiation spectral region. In this model, GPP is proportional to the amount of light absorbed by the Chl-a pigments actually used to drive photosynthesis. In Rossini et al., (2010), several Light Use Efficiency schemes have been tested replacing the LUE and aPAR terms with spectral indices and SIF metrics. The same approach was here re-proposed and adapted for aquatic applications. Different models were defined incorporating water spectral indices and fluorescence metric, as proxies for the carbon fixation efficiency (Φ_C) and the absorbed light (F_A) in eq. 4.2.6. I decided to name the novel F_C formulation by means of RS quantities F_{C-RS}. All the cases investigated and tested in this work are summarized in Table 4.3.

Case ID	Φ_C	F_A
1	constant	$[\text{Chl-a}]_{OC4}$
2	constant	E_{PAR}
3	constant	SIF_{FLH}
4	constant	$[\text{Chl-a}]_{OC4}$*E_{PAR}
5	constant	$[\text{Chl-a}]_{OC4}$*SIF_{FLH}
6	SIF_{FLH}/F_A	$[\text{Chl-a}]_{OC4}$
7	SIF_{FLH}/F_A	E_{PAR}
8	SIF_{FLH}/F_A	SIF_{FLH}
9	SIF_{FLH}/F_A	$[\text{Chl-a}]_{OC4}$*E_{PAR}
10	SIF_{FLH}/F_A	$[\text{Chl-a}]_{OC4}$*SIF_{FLH}

Table 4.3: Phytoplankton primary production models.

In the first 5 cases (1-5), (Φ_C) is kept constant. Φ_C was not experimentally measured, therefore it was replaced by the corresponding proxy (Φ'_C), evaluated according to eq. 4.2.12. Specifically, the used constant value corresponds to the average of all the Φ'_C estimated from the water samples analysis (Φ'_C = 331.69 $mm^3 W^{-1}$).
Conversely, in cases from 6 to 10, Φ_C was evaluated as the ratio between SIF_{FLH} and F_A. This formulation contains the Φ_C dependence to the light availability and the phytoplankton status. The light availability affects the amount of light absorbed by the phytoplankton,

while the phytoplankton status is accounted by the SIF proxy evaluated with the novel FLH parametrization.
F_A, is obtained by the product between the amount of light reaching the target (E_{PAR}) and the phytoplankton absorption spectra a_{phy}, according with eq. 4.2.7. E_{PAR} is routinely measured or retrieved in the remote sensing framework, while a_{phy} is usually estimated by laboratory analysis. Nevertheless, a_{phy} depends on [Chl-a] because it is the product between the mean phytoplankton specific absorption coefficient (a^*_{phy}) and [Chl-a]. Furthermore, a linear relationship exist between [Chl-a] and the phytoplankton primary production, as observed by Deng et al. (2017) in the Lake Taihu. Thus, F_A was at first replaced by [Chl-a], and specifically by the spectral index $[Chl\text{-}a]_{OC4}$. Similarly, F_A affects the SIF emission, so SIF_{FLH} was considered as a proxy for F_A.
Finally, combinations of $[Chl\text{-}a]_{OC4}$, E_{PAR} and SIF_{FLH} were investigated. The ten F_{C-RS} parametrizations were compared to the *in-situ* phytoplankton biovolume to assess the models robustness at local scale and for inland waters. Due to the few sample values available, only a qualitative interpretation of the results was possible.

4.3 Results

4.3.1 Water body phytoplankton characterization

The phytoplankton taxa's composition in the water column was evaluated by microscope counting of the samples collected at z_0 and z_{SD} respectively. Taxa were identified to species level and counts converted to biovolume.
For all the samples analysed, a clear dominance of Bacillariophyta (diatoms) was observed with a percentage above 60% of the total algal biovolume, regardless of the depth considered, the collection time and the meteorological conditions (Figure 4.3A).

Considering the algal density instead of the biovolume (Figure 4.3B), the diatoms are still the dominant taxa in all the samples, with the exception of the deep sample collected on the morning of 02^{nd} July ($S1_{zSD}$) where the most abundant class was the Cyanobacteria (*Aphanothece minutissima*), a group of non-vacuolate, small-celled and colonial with mucilage. The genera *Aphanocapsa/Aphanothece* are usually found in large colonies. Their cells are characterized by very small biovolume, compared with other taxa. This results in a low biovolume contribution despite relatively high cell concentrations and a different content of chlorophyll-a. On the 03^{rd} July (S3 and S4), a codominance between Bacillariophyta and Chrysophyta (*Chrysochromulina sp.*) was observed.

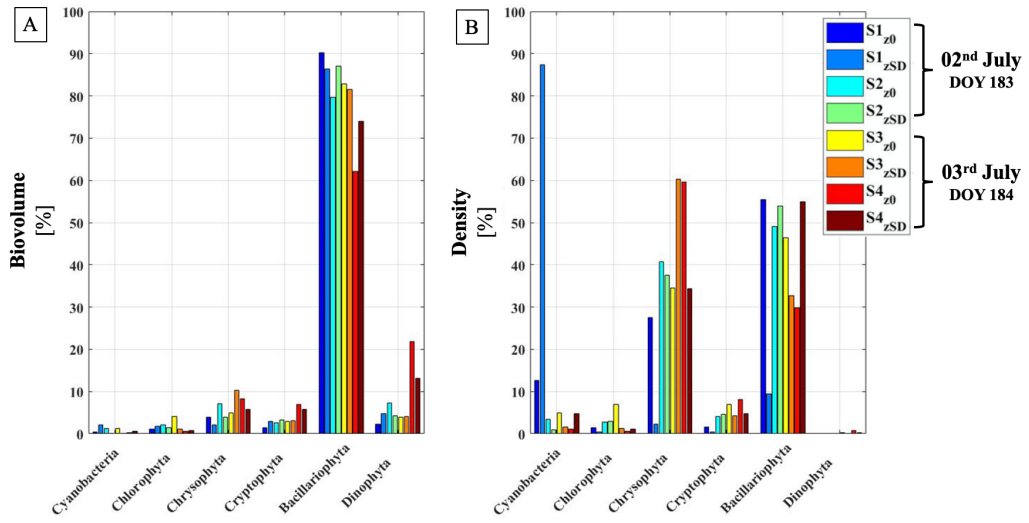

Figure 4.3: Composition of the phytoplankton major taxa expressed in percentage respect to biovolume (A) and density (B). Each color corresponds to a different sample (see nomenclature in Table 4.1).

4.3.2 Phytoplankton temporal dynamics

The phytoplankton dynamics in different illumination conditions and within the water column are hereafter presented. The results of the laboratory analysis were compared to the indices and metrics defined on hyperspectral measurements. All this data refers to the sampling performed on the 02^{nd} and 03^{rd} of July. Furthermore, also the spectral indices and metrics temporal evolution during the whole acquisition week were analysed.

4.3.2.1 Comparison between water sampling and hyperspectral quantities

Values obtained from laboratory analysis were combined with the information estimated on the collected hyperspectral measurements. Specifically, $[\text{Chl-a}]_{HPLC}$ and the algal biovolume (see Table 4.2), were compared to F_A and F_{SIF}, evaluated exploiting eq. 4.2.7 and 4.2.8, respectively. The comparison was carried out both on the depth and temporal domains. The acronym DOY stays for Day Of the Year. Specifically, the 02^{nd} of July corresponds to DOY 183, while the 03^{rd} of July to DOY 184.

Hyperspectral measurements, collected by the RoX simultaneously with the water sampling were extrapolated from the whole week-long time series. E_{PAR} and SIF_{FLH} were

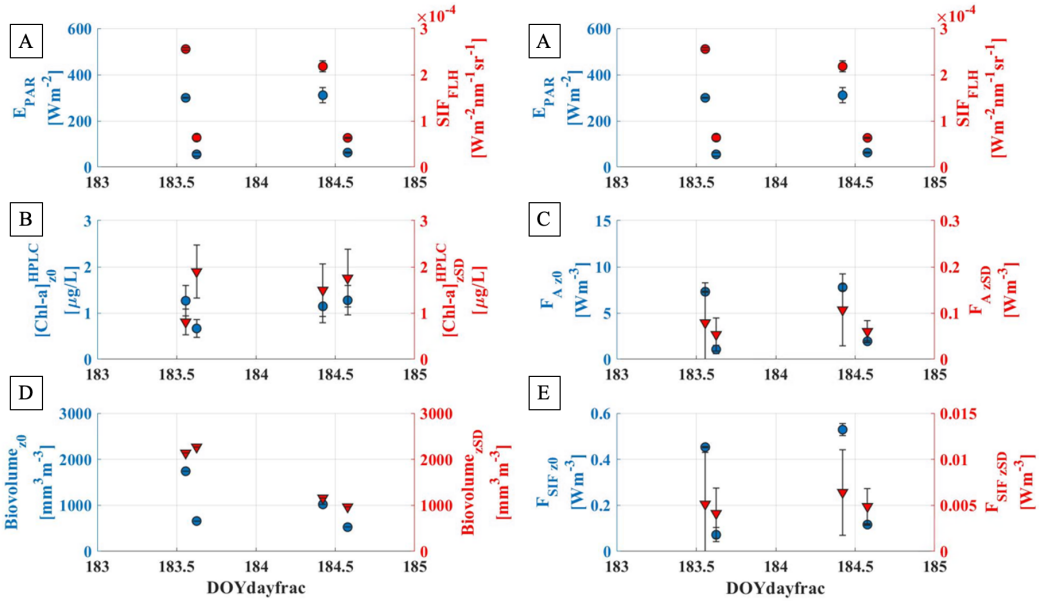

Figure 4.4: Dots represent variables measured close to the surface (z_0), while triangles refer to the sampling performed at the Secchi Disk depth (z_{SD}). Time is shown as Day-Of-the Year (DOY): 183 is the 02.07.2019 while 184 is the 03.07.2019. First line (Panel A left and right) shows the E_{PAR} (in blue) and SIF_{FLH} (in red) obtained from the ROX hyperspectral measurements. (B) and (D) are the $[Chl\text{-}a]_{HPLC}$ and the biovolume, respectively, evaluated from laboratory analysis. (C) and (E) corresponding to F_A and F_{SIF} have been obtained combining the laboratory outcomes and the Satlantic measurements. All the quantities displayed, except the biovolume (D), are mean values. The error bars correspond to the standard deviations.

selected as proxies of the amount of light reaching the target and the SIF emitted by the layer close to the surface. The trends observed are displayed in Figure 4.4. Two waters sampling were performed per day: the first around the solar noon, the second at the sunset.

The first line of Figure 4.4 shows the trends of E_{PAR} and SIF_{FLH}. Unlike the other parameters in Figure 4.4, they are evaluated only at the surface. As expected, E_{PAR} decreases during the day. The same trend was observed for the SIF_{FLH}.

The surface $[Chl\text{-}a]_{HPLC}$ shows opposite diurnal trends in the two days investigated (Figure 4.4B, blue dots). The 02^{nd} of July (DOY 183) was characterized by clear sky conditions up to the solar noon then clouds appeared, the 03^{rd} of July (DOY 184) was completely cloudy. Thus, under almost clear sky conditions, the amount of $[Chl\text{-}a]_{HPLC}$ at z_0 decreases during the day, following the E_{PAR} and SIF_{FLH} trends. Conversely, in predominant cloudy sky conditions, the surface $[Chl\text{-}a]_{HPLC}$ values are more stable throughout the day. The

differences between DOY 183 and DOY 184 in the $[\text{Chl-a}]_{HPLC}$ trend could be linked to the anomalous peak of Cyanobacteria (observed in the $S1_{zSD}$ sample) as highlighted in Figure 4.3. Concerning the Secchi Disk depth (Figure 4.4B, red triangles), the $[\text{Chl-a}]_{HPLC}$ seems less affected by the weather variability. In both cases evaluated, there is a common increase of the $[\text{Chl-a}]_{HPLC}$ from the solar noon till the afternoon.

Focusing on the biovolume, the observed trends do not always agree with the ones highlighted for the $[\text{Chl-a}]_{HPLC}$ (Figure 4.4D). This is probably linked to the different phytoplankton species contained in the samples. Under clear sky conditions (DOY 183), the biovolume at the Secchi Disk depth was greater than the one at the surface. Conversely, under cloudy conditions (DOY 184) the biovolume shows almost constant values, regardless the depth considered.

Figure 4.4C and E show the F_A and F_{SIF} trends. According to its parametrization (see eq. 4.2.7), F_A is strictly linked to the irradiance reaching the target. Thus, the F_A diurnal evolution agrees with the one observed for the E_{PAR}, regardless the depth considered. Differences between surface and depth are restricted only to discrete values; specifically, F_A decreases with the depth. Similarly, F_{SIF} shows the same E_{PAR}, F_A and SIF_{FLH} diurnal trend. In both days investigated, at the surface, F_{SIF} decreases from the solar noon to the sunset. Conversely, at the z_{SD} (Figure 4.4E, red triangles), F_{SIF} seems less affected by the illumination conditions. In general, values linked to z_{SD} are characterized by a greater variability, compared to the surface values, according to the standard deviations evaluated.

Interesting relationships between water samples analysis outcomes and parameters obtained from the hyperspectral measurements, were qualitative observed (Figure 4.4). A more detailed investigation was carried out comparing E_{PAR} and SIF_{FLH} to at-the-surface F_A and F_{SIF}. The link between these quantities is summarized in Figure 4.5.

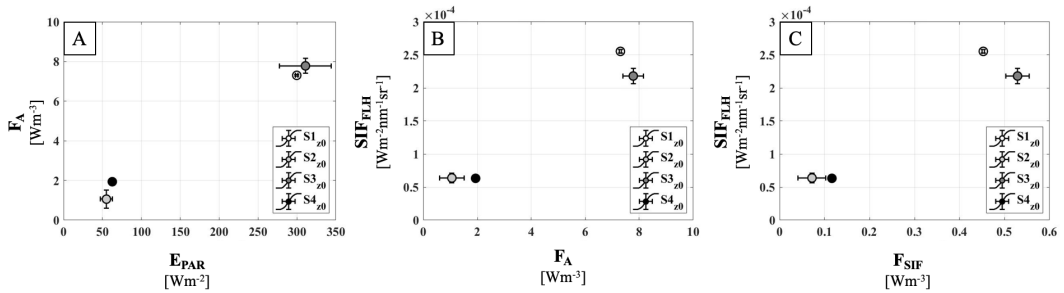

Figure 4.5: Values displayed correspond to measurements evaluated at the surface. The colors (grey scale) help to discern between the several samples considered. All the displayed measurements correspond to mean values and the error bars show the standard deviations.

In detail, Figure 4.5A and B show a linear relationship between E_{PAR}, F_A and SIF_{FLH}. Thus, both E_{PAR} and SIF_{FLH} can be considered F_A proxies. This last observation supports many of the assumptions considered in the phytoplankton PP models (Table 4.3, Cases 2,3,6,7).
The SIF_{FLH}, obtained with the novel FLH parametrization, was compared to F_{SIF} (Figure 4.5C). Due to the few points available, it was not possible to quantify the FLH accuracy. Nevertheless, SIF_{FLH} and F_{SIF} seem correlated one to each other. Therefore, SIF_{FLH}, evaluated from hyperspectral measurements, can be considered a proxy for F_{SIF}.

Eventually, an analysis of the SIF yield (Φ_{SIF}) was performed. Laboratory outcomes are summarized in Figure 4.6A.

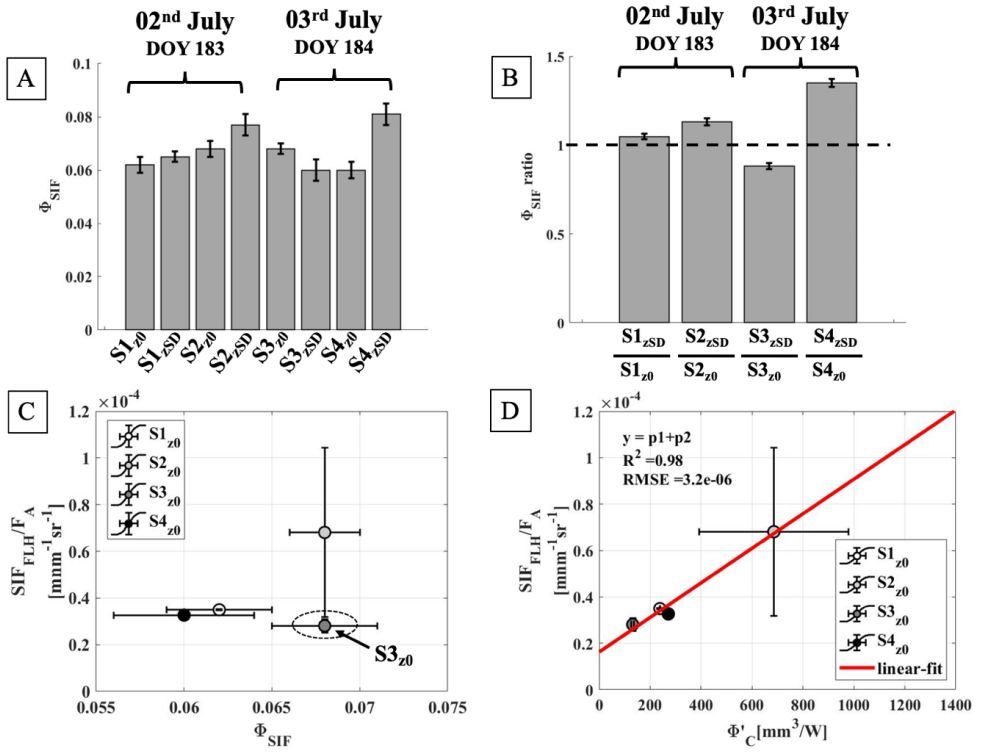

Figure 4.6: (A): Φ_{SIF} values obtained from the laboratory analysis. (B): relative Φ_{SIF} ratio. (C) comparison between the SIF_{FLH}/F_A and Φ_{SIF}. (D) comparison between SIF_{FLH}/F_A and Φ'_C.

As shown, the maximal values of Φ_{SIF} (i.e. when the Chl-a re-absorption is inhibited) are almost homogeneous in the different samples collected. To characterize the Φ_{SIF} behaviour with the depth, a relative Φ_{SIF} ratio was evaluated dividing the z_{SD} values by the z_0 ones. In general, the ratio gives values close to one. Case S3 is an exception because it is characterized by a decrease of the yield from the surface to the Secchi Disk depth (Figure 4.6B) and then by a ratio minor than one. Laboratory Φ_{SIF} were also compared to the equivalents obtained exploiting the hyperspectral measurements (Figure 4.6C). In details, the experimental SIF yield was evaluated by the ratio between SIF_{FLH} and F_A. The comparison was carried out only between surface values. While data referred to S1, S2 and S4 seems linear correlated, S3 is an outlier. Conversely, a very good correlation was observed between SIF_{FLH}/F_A and Φ'_C (Figure 4.6D). The regression performed shows a R^2 equal to 0.98. Again, this last observation agrees with the approximations introduced in the phytoplankton PP parametrization.

4.3.2.2 Hyperspectral measurements analysis

In order to resolve the phytoplankton dynamics in time, high frequency hyperspectral measurements analysis were used. The results shown in Figure 4.7 refer to the whole week acquisition.

To better characterize the several illuminations conditions during the day, the E_{PAR} was evaluated. This parameter, indeed, account for the amount of light reaching the target. To evaluate the portion of light reflected by the water body, the R_{550}^{app} index was used. F_{SIF} and $[\text{Chl-a}]_{OC4}$ are linked to the SIF signal arising from the water body and the phytoplankton concentration, respectively.

In Figure 4.7A, DOY 185 and DOY 186 are characterized by the typical trend of the clear sky days, in which the E_{PAR} approximately goes as the cosine of the solar zenith angle. The small standard deviations evaluated indicate stable illumination conditions during the time interval. During DOY 184, a larger variability was observed due to the clouds appeared in the afternoon. DOY 183 and DOY 187 show a mixed sky condition, with a drastic drop of the light intensity after the solar noon.

R_{550}^{app} is shown in Figure 4.7B. In the two identified clear sky days (DOY 185 and DOY 186), a minimum is reached in correspondence with the E_{PAR} maximum. Conversely, the R_{550}^{app} diurnal trend is more variable under cloud sky conditions.

The SIF_{FLH}, evaluated choosing dynamically the wavebands positions, represents a proxy for the SIF emitted by the phytoplankton. As shown in Figure 4.7C, the SIF_{FLH} is characterized by a peculiar trend in the two clear sky days (DOY 185 and DOY 186), with a drop around the solar noon. In this moment of the day, the amount of energy available to be absorbed by the target is greater, as arises from the E_{PAR} diurnal trend. Thus, strong light conditions seem to affect the SIF emission. From a qualitative point of view, the SIF_{FLH} temporal evolution is similar to the one observed for the E_{PAR}. The correlation between the two variables is clear and it was demonstrated in Figure 4.8A. In this case, a linear relationship between the E_{PAR} and the SIF_{FLH} is highlighted for all the

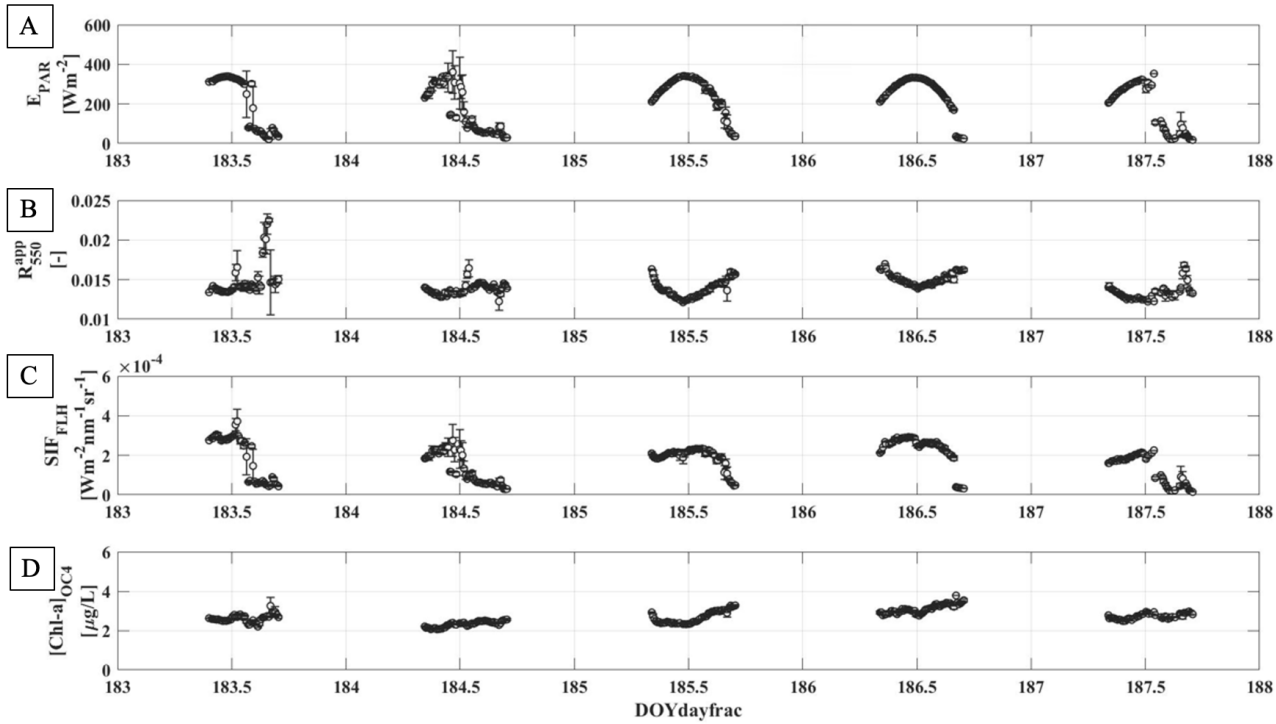

Figure 4.7: (A): E_{PAR}; (B) R_{550}^{app}; (C) SIF_{FLH} and (D) $[Chl\text{-}a]_{OC4}$. Data shown represent the mean values (n_{max} per interval ~10), averaged on a time interval of 10 minutes, while the error bars correspond to the standard deviations.

days investigated. This pattern was also observed in the terrestrial vegetation, when the E_{PAR} was compared to the fluorescence evaluated at 760 nm.

The near-surface [Chl-a]$_{OC4}$ is shown in Figure 4.7D. Almost all the investigated days are characterized by a diurnal, monotonic, growth of this index. Again, the two clear sky days show a peculiar trend around the solar noon where a minimum is reached. Focusing on the [Chl-a]$_{OC4}$, the algorithm should give as outcome a chlorophyll concentration in µg/L. These values were compared to those obtained from the water samples analysis Figure 4.8B. The [Chl-a]$_{OC4}$ overestimate the *in-situ* values. This mismatch is probably attributable to the regression coefficients (a_0-a_4) used in the eq. 4.2.3 that probably are not suitable for the investigated water body. Although a strong correlation between [Chl-a]$_{OC4}$ and [Chl-a]$_{HPLC}$ was not observed, [Chl-a]$_{OC4}$ was used as a proxy for the water body [Chl-a].

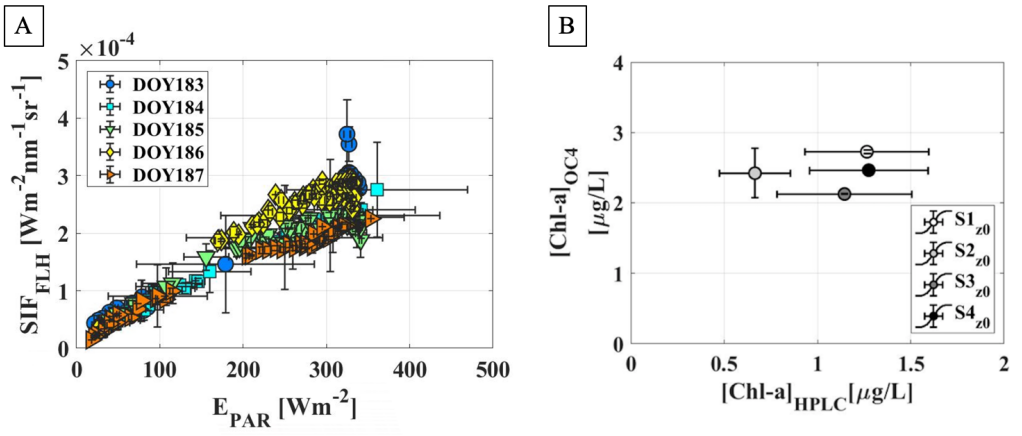

Figure 4.8: (A): comparison between E_{PAR} and SIF_{FLH}. Each symbol corresponds to a different acquisition day. Data shown represent the mean values (n_{max} per interval ~10), averaged on a time interval of 10 minutes, while the error bars correspond to the standard deviations. (B) comparison between the [Chl-a]$_{OC4}$ and [Chl-a]$_{HPLC}$. Data showed corresponds to [Chl-a]$_{HPLC}$ collected close to the surface.

4.3.3 Test of phytoplankton primary production models

Several results obtained in Section 4.3.2 support the approximations applied in the phytoplankton PP models (Table 4.3). Specifically, both E_{PAR} and SIF_{FLH} can be used as a proxy for the F_A terms. Moreover, the Φ_C proxy (Φ'_C) is strictly related to the ratio between SIF_{FLH} and F_A. Considering all these outcomes, the phytoplankton PP models

were tested. Since only four points, corresponding to the surface samples, were available, the carried out analysis was mainly qualitative. A numerical comparison was not possible because of the many approximations taken. Specifically: i) the constant used as *in-situ* value (Cases 1-5) is a proxy for Φ_C; ii) the replacement of the F_C parameter in the models validations with the biovolume and iii) the overestimation of the $[\text{Chl-a}]_{OC4}$ with reference to the true field values. The results of the comparison between F_{C-RS} and the biovolume are summarized in Figure 4.9.

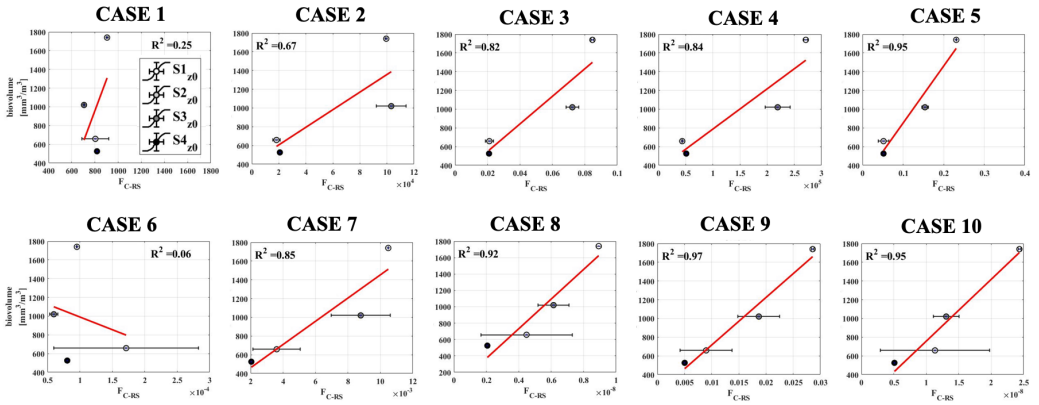

Figure 4.9: First row collects the cases in which the Φ_C proxy has been kept constant, while the F_A has been replaced by indices and metrics obtained from the hyperspectral measurements. Second row collects the cases in which also the Φ_C has been replaced by a proxy defined from RS quantities. The red lines correspond to the linear regression performed on the measurements. The scale for F_{C-RS} have been omitted on purpose. Due to the approximations taken only a qualitative comparison was possible.

If only the spectral index for $[\text{Chl-a}]_{OC4}$ is used to replace F_A, the model fails, regardless of the Φ_C parametrizations (Case 1 and Case 6). Conversely, if F_A is replaced by E_{PAR}, the F_{C-RS} better follows the field biovolume (R^2 equal to 0.67 and 0.85 for Case 2 and Case 7, respectively). The statistics improve when the SIF_{FLH} is used, because this parameter is strictly linked to the amount of light actually absorbed by the Chl-a pigments. In this case, R^2 greater than 0.80 were observed in Case 3 and 8.
When F_A is parametrized with two terms, one linked to [Chl-a] and the other to the available light, the parametrization of Φ_C plays a key role. If Φ_C is constant (Case 5), the F_{C-RS}, obtained multiplying $[\text{Chl-a}]_{OC4}$ and SIF_{FLH}, shows a very good correlation with the biovolume. The evaluated R^2 is equal to 0.95. A further improvement is obtained when Φ_C is replaced by the SIF yield proxy (i.e. SIF_{FLH}/F_A). Case 9 parametrization gives the highest performance with an R^2 equal to 0.97.

4.4 Discussion

4.4.1 Water body phytoplankton characterization

Lake Maggiore phytoplankton composition was characterized by microscopic counting. Specifically, the biovolume and the density associated to the different taxa was summarized in Figure 4.3.
In all the analysed samples, it is clear a dominance of the Bacillariophyta (diatoms). This result agrees with the observations previously reported in Morabito et al., (2007). That is important because diatoms are responsible for 20% to 25% of global carbon fixation (Smetacek 1999) and serve as the basis for pelagic food webs (Znachor et al., 2012). Diatom cells, indeed, are narrower and contain more chlorophyll compared to the other taxa. The Bacillariophyta's growth (in particular of the *Fragilaria*) is constrained by the light availability (Znachor and Nedoma, 2008). In the water samples collected around the solar noon, namely S1 and S3, was observed a decreasing trend of the Bacillariophyta biovolume and density from surface (z_0) to depth (z_{SD}) (Figure 4.3). Conversely, an opposite trend was observed in samples collected in the afternoon (S2 and S4) with greater values at z_{SD}. These results support the strong link between the Bacillariophyta stratification in the water column driven by the light availability.
Finally, a peak of Cyanobacteria was observed at z_{SD} the 02^{nd} of July, in the sample collected close to the solar noon ($S1_{zSD}$). This anomalous value could be linked to the copious rainfalls registered on the 1^{st} of July in the afternoon, that could affect the Cyanobacteria growth.

4.4.2 Phytoplankton temporal dynamics

The phytoplankton dynamics during the day and at different depths were investigated. The values obtained from the water samples analysis and from hyperspectral measurements were compared in Figure 4.4.
The $[\text{Chl-a}]_{HPLC}$ depends on the weather variability. In the two studied days, characterized by different sky conditions, the $[\text{Chl-a}]_{HPLC}$ shows opposite trends from the solar noon to the sunset. Conversely, the biovolume is less affected by the illumination changes. In general, the $[\text{Chl-a}]_{HPLC}$ and the biovolume values evaluated at the Secchi Disk depth are greater compared to the surface's ones. The only exception is represented by the first point that corresponds to the sample S1. In this case, the $[\text{Chl-a}]_{HPLC}$ at the surface is greater compared to the values obtained at the Secchi Disk depth. This anomalous outcome can be explained to the peak of Cyanobacteria observed in the sample $S1_{zSD}$ (Figure 4.3B). These taxa are characterized by high density and low chlorophyll content. Concerning F_A and F_{SIF}, their trends agree with the ones observed for E_{PAR} and SIF_{FLH}, regardless of the depth considered.

A more detailed analysis about the links between E_{PAR}, SIF_{FLH}, F_A and F_{SIF} was carried out on the surface's values. The comparisons are summarized in Figure 4.5. F_A seems to be

positively related to both E_{PAR} and SIF_{FLH}. The linear correlation between F_A and SIF_{FLH} observed in Figure 4.4 is supported by the Φ_{SIF} values obtained in the laboratory analysis. Largely invariant values of Φ_{SIF} were observed in the water samples analysed (Figure 4.6A). Strong variations in the Φ_{SIF}, due to quenching process of non-photochemical nature, are expected to lead a breakdown in the linear correlation between F_A and SIF. Because of the relatively small differences in terms of Φ_{SIF} in all the samples analyzed, it is possible to assess that in the investigated days (DOY 183 and DOY 184) the photochemical quenching was almost negligible.
A comparison between SIF_{FLH} and F_{SIF} was investigated. The few available points available did not allow us to evaluate quantitatively the novel FLH algorithm accuracy. However, a linear trend between SIF_{FLH} and F_{SIF} was observed.

The phytoplankton behaviour was also analysed from a physiological point of view, exploiting the Φ_{SIF} parameter. Values referred to the Secchi Disk depth were divided by the surface Φ_{SIF} (Figure 4.6B). In all the cases, except for the S3 case, the evaluated ratios always show values equal or greater than 1, therefore, the surface values are weakly attenuated in comparison to the depth ones. This trend is probably due to the amount of light reaching the water body that decreases with the depth. Φ_{SIF} was also compared to an experimental SIF yield obtained from the ratio between SIF_{FLH} and F_A (Figure 4.6C). While points referred to the samples S1, S2 and S4 seem linear correlated, S3 is an outlier. I remind that S3 is the case in which the surface Φ_{SIF} was not quenched. Finally, Φ'_C and SIF_{FLH}/F_A are linearly correlated for all the samples investigated, therefore, in first approximation, the SIF yield can be used as a proxy for the the light use efficiency parameter (Figure 4.6D).

Hyperspectral and high frequency measurements were analysed to understand the phytoplankton temporal dynamics. In Figure 4.2A are shown the R_{rs} spectra acquired during the day. The overall spectrum magnitude is greater in the morning and evening, while the minimum values are reached around the solar noon. At diurnal scale, the R_{rs} spectra (and then the L_w) are strongly influenced by sun-target position. Furthermore, a peculiar trend appears in the spectra close to the sunset. In the region where the SIF emission occurs, the R_{rs} magnitude anomaly increases. The observed behaviour could be linked to changes in the SIF yield. With reference to Table 4.2, indeed, greater Φ_{SIF} values are observed in the afternoon.

A more detailed analysis was carried out by means of indices and metrics defined on the hyperspectral measurements, specifically E_{PAR}, R^{app}_{550}, SIF_{FLH} and $[\text{Chl-a}]_{OC4}$ (see Figure 4.7). A peculiar trend was observed around the solar noon in the two clear sky days (DOY 185 and DOY 186). When E_{PAR} reaches the maximum, both the SIF_{FLH} and the $[\text{Chl-a}]_{OC4}$ drop. The most likely explanation of the SIF_{FLH}'s decrease around the solar noon is the occurrence of non-photochemical quenching of the fluorescence under very bright natural light. On the other hand, according to Reynolds et al., (2006), the phytoplankton act strategies to escape the harmful photoinhibition caused by excessive insolation near the surface. In order to avoid damages caused by oxidative stress, the phytoplankton migrates at higher depths. This last hypothesis correspond to the diurnal $[\text{Chl-a}]_{OC4}$ trend characterized by lower values around the solar noon. This index is

evaluated at wavelengths not influenced by the SIF emission, therefore the $[\text{Chl-a}]_{OC4}$ trend can be ascribable to a shift of the phytoplankton organisms deeper in the water column. Nevertheless, a weak surface quenching was previously highlighted. It is hence likely that both movement of the phytoplankton through the water column and changes in Φ_{SIF} explain the experimental trends. The high frequency spectral measurements were also exploited to evaluate the link between the amount of light reaching the target (E_{PAR}) and the SIF_{FLH} metric (Figure 4.8A). The linear relation observed for the whole time series supports again the hypothesis of an almost invariant Φ_{SIF}. Furthermore, the SIF_{FLH} could be used as a proxy for the incoming irradiance (E_d).

4.4.3 Test of phytoplankton primary production models

Results from Section 4.3.2 clearly demonstrate the strong link between *in-situ* and hyperspectral-based quantities and how the hyperspectral resolution is suitable to better follow phytoplankton's temporal dynamics. Briefly, E_{PAR} and SIF_{FLH} can be used as proxies for the F_A. Φ_C proxy (Φ'_C) is strictly related to the ratio between SIF_{FLH} and F_A. E_{PAR} and SIF_{FLH} are linearly correlated. $[\text{Chl-a}]_{OC4}$ can be used as a proxy for the phytoplankton Chl-a concentration in the water body. Therefore, hyperspectral indices and metrics can be potentially used to predict the phytoplankton PP in inland waters.
The models summarized in Table 4.3 were investigated comparing F_{C-RS} to the biovolume. Since only few field measurements were available for the test, the preliminary analysis carried out was mainly qualitative and restricted to local scale. The highlighted correlations have no statistical meaning, but they were used only to asses which F_{C-RS} formulation better follows the biovolume's evolution in time (Cesana et al., 2021).

When F_A is replaced by $[\text{Chl-a}]_{OC4}$, the obtained F_{C-RS} does not correlate to the biovolume, regardless of the adopted Φ_C parametrization (Case 1 and Case 6). In first approximation, this outcome seems to be in contrast with the study carried out by Deng et al., (2017). However, it must be remembered that in Deng et al., (2017) the biomass has been compared to the [Chl-a], while in Lake Maggiore I analysed the biovolume and the chlorophyll concentration related spectral index. Nevertheless, the model's failure in Cases 1 and 6 could mean that the $[\text{Chl-a}]_{OC4}$ alone is not enough to describe the F_A term and then the biovolume diurnal evolution. The [Chl-a], obtained with the OC_4 algorithm, is evaluated on the reflectance's spectral region not interested by the SIF. So, it is possible to infer that the [Chl-a] does not contain any the physiological information linked to the photosynthesis.
Conversely, promising results were reached when the F_A was replaced by the E_{PAR}, (Case 2 and 7) with R^2 of 0.67 and 0.85, respectively. Statistics further improve the SIF_{FLH} is considered, as in Case 3 and 8, where R^2 is greater than 0.80. This is probably consistent with an intrinsically dependence of the SIF_{FLH} to the phytoplankton concentration. So it is possible to state that the SIF_{FLH} is strictly linked to the biovolume and then more interconnected to the photosynthesis. Consequently, fluorescence's proxy could be exploited to obtain a reliable phytoplankton productivity model. These outcomes agree with the

study by Barnes et al., (2014), in which it is stated that the rate of phytoplankton primary production is primarily a function of the incident irradiance (E_{PAR}), light absorption efficiency (linked to SIF_{FLH}) and the quantum efficiency of carbon fixation (Φ'_C from field values or the spectral proxy SIF_{FLH}/F_A).
Finally, a brief consideration about the chlorophyll-a parameter. [Chl-a] is the main driver of variability in primary production in the global ocean and thus simple empirical relationships, that directly relate [Chl-a] to primary production, were used in marine sciences. However, the existing models, developed for ocean waters, are not suitable for lakes (Cesana et al., 2021). Indeed, the parametrizations used in Cases 1 and 6 fail when only the $[\text{Chl-a}]_{OC4}$ is considered. Conversely, when the $[\text{Chl-a}]_{OC4}$ is coupled with the E_{PAR} and SIF_{FLH}, respectively, the statistics improve. In this context, the F_{C-RS} parametrization with the best result is the one developed for Case 9 (R^2 = 0.97) where the Φ_C was chosen to be non-constant in order to account the dependence to the light availability and the phytoplankton status. Φ_C was replaced by the ratio between SIF_{FLH} and F_A. Finally, F_A was obtained by the product between E_{PAR} and $[\text{Chl-a}]_{OC4}$ (used as proxy for a_{phy}), consistently with the F_A definition showed in eq. 4.2.7.

4.5 Conclusions

This study demonstrates how hyperspectral resolution and high frequency temporal measurements are suitable to follow the phytoplankton dynamics, particularly in clear sky days. The calculated spectral indices are strictly linked to the lake water's characteristics inferred by laboratory analysis. Furthermore, several phytoplankton primary production models, driven from remotely sensed data, were tested on Lake Maggiore. The preliminary outcomes demonstrate that, at local scale, hyperspectral-based quantities represent a sensitive tool for monitoring temporal variations in phytoplankton PP. In this work, I originally introduced the fluorescence yield as a proxy of the light use efficiency parameter. The few points used to test the ten PP parametrizations do not allow us to assess which phytoplankton production model is suitable for Lake Maggiore. Nevertheless, this preliminary analysis highlighted how the statistics improve when the SIF_{FLH} replaces the F_A term. In summary, it seems that Case 5 and Case 9 are the best models for inferring productivity in our study area, particularly when fluorescence yield is used as a proxy of carbon fixation efficiency. This cannot be probably extended to other lakes and further investigations and validations are needed. Even though this represents only a preliminary study carried out at local scale, the future goal is to exploit RS fluorescence to retrieve the phytoplankton PP in order to overcome the limits related to the sparse measurements of typical *in-situ* sampling and to improve the inland waters spatio-temporal monitoring and understanding.

Chapter 5

Conclusions

The PhD research aims to implement novel strategies able to retrieve and interpret the sun-induced fluorescence signal emitted by terrestrial vegetation and inland waters in order to advance the understanding of these environments. The approaches developed were investigated and tested on experimental hyperspectral measurements and Radiative Transfer simulations. All the analysis presented were carried out considering the measurements at ground level. Two main ecosystems were investigated: agricultural crops (forage, alfalfa, corn) and clear lake waters. For each study case, different objectives were highlighted, addressed in the corresponding Chapter (i.e. Chapter 2, 3 and 4). The main results achieved are summarized below, along with the concluding remarks and the future research perspectives.

5.1 Main Results

Characterizing the SIF full spectrum at Seasonal/Diurnal scales for agricultural crops

Reflectance-based indices have been historically employed to monitor the terrestrial vegetation because able to describe more than 95% of the vegetation variability of biophysical parameters occurring during the entire crop-growing season (Gitelson et al., 2003a; Gitelson et al., 2003b; Jiang et al., 2006). Nevertheless, the SIF signal has emerged as a more exhaustive parameter in the vegetation monitoring inasmuch it is emitted by the vegetation only and strongly linked to the photosynthesis. Therefore, it is more related to the plant physiology compared to the reflectance indices. In Chapter 2, novel SIF metrics defined on the full spectrum were compared to the fluorescence evaluated in the O_2-A and O_2-B bands, and reflectance-based indices used as proxies for several biophysical parameters. This study was carried out at seasonal/diurnal scales on experimental hyperspectral measurements acquired over three different crops, namely forage, alfalfa and corn. The SIF

at the maximum emission in the red and far-red, the corresponding peak positions and the spectrally integrated SIF were investigated. On the other hand, the reflectance-based indices selected are the $NDVI_{red-edge}$, $CI_{red-edge}$, NIR_v and PRI because linked to the fraction of light absorbed between 400-700 nm (fAPAR), chlorophyll content (C_{ab} and LAI), SIF far-red and photosynthetic activity (NPQ proxy).

At first, the novel and conventional fluorescence metrics were compared. Regardless of the temporal scale considered, the SIF evaluated at the peaks is strongly correlated with the corresponding O_2 bands values. A quantitative analysis was carried out at seasonal scale. The linear regression between SIF_{760} and $SIF_{far-red}$ gives slope= 1.31, intercept= 0.23, R^2= 0.98 and RMSE= 0.29. Similarly, SIF_{687} and SIF_{red} lie closer to the 1:1 with slope= 0.95, intercept= 1.03, R^2= 0.95, RMSE= 0.08. Therefore, clear improvements in the use of the SIF at the peaks than the canonical metrics were not observed. The SIF evaluated at a selected wavelength was compared to the fluorescence emitted between 670-780 nm. SIF_{INT} is strongly correlated to SIF_{760} (R^2= 0.92) and $SIF_{far-red}$ (R^2= 0.99). This result is expected because for well-developed crops, the emission in the visible red is re absorbed by the canopy-leaf system and re emitted in the far-red spectral region. The SIF_{687} and SIF_{red} increase linearly with the SIF_{INT}, with R^2 equal to 0.81 and 0.83, respectively. The lower statistics compared to the far-red cases are attributable to a different slope between the SIF_{INT} and SIF_{red} (or SIF_{687}), when sparse and medium-dense canopies are considered. The initial increase is linear, then the reabsorption occurs and the SIF_{red} is diminished while the SIF_{INT} continues to grow. Although the SIF_{INT} seasonal trend mainly agree with the fluorescence in the far-red ($SIF_{far-red}$ and SIF_{760}), it contains also information about the visible red spectral region. Therefore, it represents a more complete parameter respect to the fluorescence evaluated at a single wavelength.

The reabsorption effect on the TOC SIF spectrum is clear when the peak positions are investigated at seasonal scale. In particular, the $\lambda_{far-red}$ undergoes to a shift toward longer wavelengths, while λ_{red}, after an initial growth, decreases to a constant value.

In general, the reflectance-based indices ($NDVI_{red-edge}$, $CI_{red-edge}$, NIR_v) and the SIF metrics show similar seasonal evolutions. This result is expected because they are both driven by the same biophysical parameters, such as LAI, C_{ab} and fAPAR. All the relations qualitatively highlighted were supported by the PCA outcomes. Focusing on the loadings plot, all the variables considered lie close to each other (except for the PAR and SIF_{687}) and are positively correlated on the first principal component.

Conversely, at diurnal scale reflectance indices and SIF metrics show different patterns. $NDVI_{red-edge}$ and $CI_{red-edge}$ are almost constant during the day (except the $CI_{red-edge}$ evaluated for the corn) because they are defined in order to minimized the variability linked to the sun-target-sensor geometry and background sources of errors. On the other hand, the PRI shows a minimum around the solar noon. This behavior is probably ascribable to the mechanisms activated in order to protect the plant photosystems from potential photodamages. At diurnal scale, indeed, the PRI well correlates with the NPQ. Concerning the fluorescence, all the metrics linked to the SIF magnitude follow the PAR diurnal trend. However, around the solar noon, the signal is flatter compared to the PAR probably because

in this temporal interval, the SIF is in competition with the NPQ. Finally, an inter-crop variability is observed, specifically between corn and forage-alfalfa. All these qualitative observations are supported by the PCA outcomes (scores and loadings plots).

In general, at diurnal scale the combined use of SIF metrics and PRI allows to gain a more completed description of the vegetation behavior.

Correcting SIF for canopy reabsorption for of quantum yield estimations

The analysis carried out on the experimental TOC SIF at seasonal scale in Chapter 2 highlighted how the red spectral region is affected by the reabsorption occurring within the canopy-leaf system. To better characterize the SIF changes in terms of magnitude and spectral shape, the at the photosystems level (corrected for the reabsorption, RC) and TOC fluorescence were compared. In particular, different illumination conditions (SZA), canopy chlorophyll content (C_{ab} and LAI) and plants physiological conditions (Fqe) were considered. The analyses described are currently feasible only exploiting synthetic measurement. In this framework, the dataset used was generated coupling the MODTRAN5-SCOPE radiative transfer models and changing the aforementioned parameters (i.e. SZA, LAI, C_{ab}, Fqe). The comparisons between RC and TOC spectra were carried out by means of metrics defined on the full spectrum (the same used in Chapter 2). In addition, also the SIF yields were evaluated and investigated. The trends were analyzed for increasing LAIxC_{ab} values in order to mimic the canopy/crop growth.

In general, all the metrics evaluated at the photosystems level show a pattern similar to the one observed for the aPAR. Conversely, at TOC scale, differences were observed between the far-red and red spectral regions. While SIF_{760}^{TOC}, $SIF_{far-red}^{TOC}$, SIF_{INT}^{TOC} globally corresponds to the aPAR and the photosystems level cases, the fluorescence emission in the visible red is re absorbed for increasing LAIxC_{ab} (i.e. SIF_{red}^{TOC} and SIF_{687}^{TOC}). Since the only parameters made it vary is the canopy chlorophyll content (Fqe and SZA constant), the trend observed is only attributable to reabsorption. This process affects both the TOC SIF magnitude and spectral shape. In particular, the metrics linked to the peak positions show clear trends with the LAIxC_{ab}, a small variability with the SZA and no evident changes with the Fqe. Summarizing, the TOC SIF in the red undergoes to dramatically variations due to the reabsorption, preventing a correct interpretation of the SIF from a physiological point of view (e.g. link to the PSII emission). Another pivotal parameter for the vegetation monitoring is the SIF yield (or fluorescence efficiency). PQ, NPQ and SIF processes are interconnected to each other, therefore knowing the SIF yield and inferring NPQ from the PRI, allows to gain information about the photosynthesis and then the plant physiological status. However, the SIF yield estimated on the TOC spectrum is strongly affected by the reabsorption. From the comparison between the TOC and RC simulated spectra emerged that the SIF_{yield}^{TOC} is unable to approximate correctly the true fluorescence efficiency. Although SIF_{yield}^{TOC} increases linearly with the Fqe, the values computed lie far from the 1:1 (slope= 1.3, intercept= $8.7e^{-8}$, R^2= 1, RMSE= $4e^{-10}$). Conversely, the SIF yield evaluated on the SIF corrected for the reabsorption perfectly

match with the theoretical values (slopes= 0.995, intercept= 7^{-11}, R^2= 1, RMSE=$3e^{-11}$). Therefore, developing approaches able to correct the TOC SIF spectrum for reabsorption is fundamental to infer correctly the SIF behavior during the canopy/crop growing season and to obtain SIF_{yield} values as accurate as possible.

In this regard, two methods to correct the SIF for reabsorption were developed and test. The first approach was implemented modeling the simulated SIF^{RC} spectral shape and magnitude as sum of two Gaussian. The six coefficients used to parametrize the SIF^{RC} are linked to constants, regression values and quantity measurable experimentally. The method's accuracy carried out on the synthetic dataset shows a strong dependence to the LAI-C_{ab} parameters. In particular, the comparison between the spectrum simulated with SCOPE (used as reference) and the one estimated with the parametric method proposed gives RRMSE% up to 40% when sparse vegetation are considered. The RRMSE% decreases for LAIxC_{ab} values corresponding to medium and dense canopies (RRMSE% < 27%). Furthermore, due to the approximation introduced on one of the coefficients used to parametrize SIF^{RC}, the RRMSE% depend on the Fqe values. The variability linked to the LAIxC_{ab} is observed also when the SIF_{yield} is evaluated using the aPAR and SIF^{RC} estimated with the methods proposed. Averaging over all the LAI and C_{ab}, the SIF^{RC} differs from the true value up to 6%. Nevertheless, the strong dependence to the canopy chlorophyll content represents a strong limit of this approach because preclude its applications at seasonal scale (i.e. over underdevelopment crops). It is clear when the method is applied on experimental measurements. While the SIF metrics and aPAR trend (where estimated) globally show reasonable trends and values, the seasonal SIF_{yield} behavior is anomalous. A more sophisticated approach was then implemented. The Fourier-ML algorithm was developed in collaborations with the Laboratory of Advanced Bio-spectroscopy (Physics Department "G. Occhialini") of the University of Milano-Bicocca. This method allows to retrieve independently and simultaneously biophysical (C_{ab}, LAI, aPAR) and physiological (SIF TOC and RC, Fqe or SIF_{yield}) parameters from the top of canopy apparent reflectance spectra. Its accuracy was first quantified on the synthetic dataset. When compared to the reference values, all the variables and SIF metrics lie on the 1:1. In particular, the linear regressions performed always show slopes between 0.986 and 1, intercepts close to zero and R^2 greater than 0.99. Moreover, all the outcomes are characterized by RRMSE% lower than 2%. Finally, the trends obtained at seasonal scale are reasonable and in agreement with the behavior expected from the synthetic dataset characterization.

Advancing SIF retrieval in clear lake waters and development of primary production models

The phytoplankton SIF emission is closely interrelated to the chlorophyll-a concentration [Chl-a] and it is linked to the phytoplankton biomass and primary production. These links are generally valid in oceanic waters. Conversely, in inland waters (e.g. lakes) the SIF retrieval and interpretation is complicated by the presence in the water body of optically active constituents, other than the phytoplankton. Hyperspectral measurements could potentially allow to disentangle the abiotic and biotic contributions in the optical signal

acquired. On the other hand, inland waters are very dynamics environments, characterized by a high spatio-temporal variability. Therefore, high frequency measurements could help to better resolve the changes that naturally occur in these water bodies. In Chapter 4, the hyperspectral and high frequency measurements acquired on a clear lake water (i.e. Lake Maggiore) were exploited to optimize the FLH method for inland waters applications. The results obtained were investigated along with other spectral indices to characterize the water body dynamics in the temporal domain. Furthermore, several phytoplankton PP models were defined and tested combining spectral indices with the novel SIF proxy proposed.

Concerning the first point, the FLH wavebands used to estimate the SIF proxy were selected automatically exploiting the variation in the R_{rs} profile occurring at diurnal scale. These changes are mainly depending on the light reaching the water surface. However, an anomalous increase of the R_{rs} magnitude in the region in which SIF emission occurs was observed in spectra acquired close to the sunset. Therefore, also a dependence to the water body phytoplankton composition (within the water column) and/or changes in the SIF efficiency emission affect the R_{rs}. Once selected dynamically the wavebands positions, the SIF_{FLH} is evaluated on the water leaving radiance. The SIF_{FLH} reliability was qualitative verified comparing its temporal trend with other spectral quantities, such as the E_{PAR}, R_{550}^{app} and $[\text{Chl-a}]_{OC4}$. A peculiar trend was observed in clear sky days, in which a drop on the fluorescence is observed around the solar noon. Similarly to the terrestrial vegetation case, the SIF emission in the red could be attenuated by mechanisms activated to protect the photosystems from photodamages under strong illumination conditions. However, in aquatic environments, the phytoplankton could also drop to greater depth at which the solar light is attenuated by the water column. Combining the information inferred by the spectral quantities investigated and the SIF efficiency values estimated in laboratory, it is plausible that both these mechanisms are actuated around the solar noon. Spectral measurements were also compared to values inferred form the water samples laboratory analysis. These analyses highlighted that the E_{PAR}, SIF_{FLH}, F_{SIF} and F_A are linearly correlated to each other. Moreover, a proxy for the photosynthesis efficiency obtained by means of the SIF_{FLH} (i.e. SIF_{FLH}/F_A) was here defined for the first time. All these relations were exploited to develop and test several phytoplankton PP models. In particular, the LUE (Light Use Efficiency) model, defined for the terrestrial vegetation, was implemented for inland waters applications. The models outcomes were compared to the *in-situ* biovolume in order to qualitatively assess their reliability. When only the $[\text{Chl-a}]_{OC4}$ is used as F_A approximation, the model fails. Conversely, the matching with the biovolume improves when the E_{PAR}, SIF_{FLH} and the photosynthesis efficiency proxy are used (R^2 greater than 0.95). The analysis carried out is mainly explorative because only four measurements were available. Nevertheless, the preliminary outcomes obtained highlighted how the use of the fluorescence improves the phytoplankton PP estimations.

5.2 Concluding remarks and future perspectives

This PhD research widely demonstrated how the SIF signal can be efficiently exploited to monitor and characterize the terrestrial vegetation and inland waters temporal dynamics. The combined use of reflectance-based indices and TOC SIF metrics allows to gain a better knowledge of the terrestrial vegetation dynamics, specifically at diurnal scale (SIF vs PRI). Furthermore, the SIF_{INT} account for both the red and far-red behaviors therefore it represents a more complete metric respect to the SIF evaluated at a selected wavelength. In this framework, a future work could be devoted to perform a similar analysis but in induced stressed conditions (e.g. water or illumination stress).

At TOC scale, the link between SIF and photosynthesis is strongly affected by reabsorption of the fluorescence signal occurring within the canopy-leaf system. The reabsorption occurring in the visible red spectral region also prevents an accurate evaluation of the SIF_{yield}. Correcting the TOC SIF spectrum for reabsorption is pivotal in order to improve the use of the fluorescence as a valid proxy for the photosynthesis. In this regard, two strategies were developed and investigated. The parametric method shows a strong dependence to the LAI and C_{ab}, therefore it needs to be improved to minimize the error evaluated on sparse vegetation. On the other hand, the Fourier-ML algorithm gives more accurate outcomes. Furthermore, it retrieves not only the SIF^{RC}, but also other biophysical parameters of interested fundamental in the vegetation monitoring. However, the Fourier-ML accuracy and reliability depends on the synthetic dataset used to train the algorithm. In this work, the atmosphere composition, along with the canopy geometry and crop type were kept fixed. Future studies could be devoted to investigate whether and how much these parameters influence the retrieval performances. Furthermore, pivotal could be implementation of the retrieval algorithm for application at airplane and satellite acquired measurements.

Concerning future experimental applications, the SIF TOC and RC metrics, together with the Fqe and aPAR estimated with the Fourier-ML algorithm could be compared to the GPP in order to investigate which of these parameters is the most efficient proxy for the Gross Primary Production.

In the inland waters, the study carried out highlighted how the novel SIF_{FLH} evaluated for hyperspectral and high frequency measurements is able to follow the water body dynamics at diurnal scale in clear lake waters. Since inland waters are, by definition, highly variable environments, the SIF_{FLH} reliability should be investigated also in other water bodies characterized by different bio-optical properties and constituents. Furthermore, the use of parameters linked to the SIF represents a promising approach for monitoring the phytoplankton primary production in lakes. In this framework, only four points were used to evaluated the models accuracy. To quantitatively investigate the models performances, a greater and more variable dataset is needed.

References

Adams III, William W et al. (1990). "Photosynthesis and chlorophyll fluorescence characteristics in relationship to changes in pigment and element composition of leaves of Platanus occidentalis L. during autumnal leaf senescence". In: *Plant Physiology* 92.4, pp. 1184–1190.

Agati, Giovanni et al. (1993). "A simple approach to the evaluation of the reabsorption of chlorophyll fluorescence spectra in intact leaves". In: *Journal of Photochemistry and Photobiology B: Biology* 17.2, pp. 163–171.

Alonso, Luis et al. (2017). "Diurnal cycle relationships between passive fluorescence, PRI and NPQ of vegetation in a controlled stress experiment". In: *Remote Sensing* 9.8, p. 770.

Amoros-Lopez, J et al. (2008). "Evaluation of remote sensing of vegetation fluorescence by the analysis of diurnal cycles". In: *International Journal of Remote Sensing* 29.17-18, pp. 5423–5436.

Austin, Roswell W (1974). "The remote sensing of spectral radiance from below the ocean surface". In: *Optical aspects of oceanography*, pp. 317–344.

Badgley, Grayson, Christopher B Field, and Joseph A Berry (2017). "Canopy near-infrared reflectance and terrestrial photosynthesis". In: *Science advances* 3.3, e1602244.

Baker, Neil R (2008). "Chlorophyll fluorescence: a probe of photosynthesis in vivo". In: *Annu. Rev. Plant Biol.* 59, pp. 89–113.

Barbanti, L and W Ambrosetti (1989). "The physical limnology of Lago Maggiore: a review". In: *Mem. Ist. Ital. Idrobiol* 46, pp. 41–68.

Baret, Frédéric, Vincent Houlès, and Martine Guerif (2007). "Quantification of plant stress using remote sensing observations and crop models: the case of nitrogen management". In: *Journal of experimental botany* 58.4, pp. 869–880.

Barnes, Morvan K et al. (2014). "Absorption-based algorithm of primary production for total and size-fractionated phytoplankton in coastal waters". In: *Marine Ecology Progress Series* 504, pp. 73–89.

Bergamino, Nadia et al. (2010). "Spatio-temporal dynamics of phytoplankton and primary production in Lake Tanganyika using a MODIS based bio-optical time series". In: *Remote sensing of environment* 114.4, pp. 772–780.

Berk, Alexander et al. (2006). "MODTRAN5: 2006 update". In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII*. Vol. 6233. International Society for Optics and Photonics, 62331F.

Bresciani, Mariano et al. (2017). "Earth observation for monitoring and mapping of cyanobacteria blooms. Case studies on five Italian lakes". In: *J. Limnol* 76.Suppl 1, pp. 127–139.

Brewin, Robert JW et al. (2017). "Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups". In: *Frontiers in Marine Science* 4, p. 104.

Buitenhuis, Erik T, Taketo Hashioka, and Corinne Le Quéré (2013). "Combined constraints on global ocean primary production using observations and models". In: *Global Biogeochemical Cycles* 27.3, pp. 847–858.

Bukata, Robert P et al. (2018). *Optical properties and remote sensing of inland and coastal waters*. CRC press.

Campbell, Petya KE et al. (2019). "Diurnal and seasonal variations in chlorophyll fluorescence associated with photosynthesis at leaf and canopy scales". In: *Remote Sensing* 11.5, p. 488.

Campbell, PK Entcheva et al. (2008). "Contribution of chlorophyll fluorescence to the apparent vegetation reflectance". In: *Science of the total environment* 404.2-3, pp. 433–439.

Celesti, Marco et al. (2018). "Exploring the physiological information of Sun-induced chlorophyll fluorescence through radiative transfer model inversion". In: *Remote sensing of environment* 215, pp. 97–108.

Cesana, Ilaria et al. (2021). "Preliminary Investigation on Phytoplankton Dynamics and Primary Production Models in an Oligotrophic Lake from Remote Sensing Measurements". In: *Sensors* 21.15, p. 5072.

Chang, Christine Y, Luis Guanter, et al. (2020). "Systematic assessment of retrieval methods for canopy far-red solar-induced chlorophyll fluorescence using high-frequency automated field spectroscopy". In: *Journal of Geophysical Research: Biogeosciences* 125.7, e2019JG005533.

Chang, Christine Y, Jiaming Wen, et al. (2021). "Unpacking the drivers of diurnal dynamics of sun-induced chlorophyll fluorescence (SIF): Canopy structure, plant physiology, instrument configuration and retrieval methods". In: *Remote Sensing of Environment* 265, p. 112672.

Cogliati, S, M Rossini, et al. (2015). "Continuous and long-term measurements of reflectance and sun-induced chlorophyll fluorescence by using novel automated field spectroscopy systems". In: *Remote Sensing of Environment* 164, pp. 270–281.

Cogliati, S, W Verhoef, et al. (2015). "Retrieval of sun-induced fluorescence using advanced spectral fitting methods". In: *Remote sensing of environment* 169, pp. 344–357.

Cogliati, Sergio et al. (2019). "A spectral fitting algorithm to retrieve the fluorescence spectrum from canopy radiance". In: *Remote sensing* 11.16, p. 1840.

Colombo, Roberto et al. (2018). "Variability of sun-induced chlorophyll fluorescence according to stand age-related processes in a managed loblolly pine forest". In: *Global change biology* 24.7, pp. 2980–2996.

Cordón, Gabriela B and Maria G Lagorio (2006). "Re-absorption of chlorophyll fluorescence in leaves revisited. A comparison of correction models". In: *Photochemical & Photobiological Sciences* 5.8, pp. 735–740.

Dall'Olmo, Giorgio, Anatoly A Gitelson, and Donald C Rundquist (2003). "Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters". In: *Geophysical Research Letters* 30.18.

Damm, Alexander, JAN Elbers, et al. (2010). "Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP)". In: *Global Change Biology* 16.1, pp. 171–186.

Damm, Alexander, Luis Guanter, et al. (2014). "FLD-based retrieval of sun-induced chlorophyll fluorescence from medium spectral resolution airborne spectroscopy data". In: *Remote Sensing of Environment* 147, pp. 256–266.

De Grave, Charlotte et al. (2020). "Quantifying vegetation biophysical variables from the Sentinel-3/FLEX tandem mission: Evaluation of the synergy of OLCI and FLORIS data sources". In: *Remote Sensing of Environment* 251, p. 112101.

Deng, Yubing et al. (2017). "Temporal and spatial dynamics of phytoplankton primary production in Lake Taihu derived from MODIS data". In: *Remote Sensing* 9.3, p. 195.

Drusch, Matthias et al. (2016). "The fluorescence explorer mission concept—ESA's earth explorer 8". In: *IEEE Transactions on Geoscience and Remote Sensing* 55.3, pp. 1273–1284.

Ehleringer, James R and Thure E Cerling (2002). "C3 and C4 photosynthesis". In: *Encyclopedia of global environmental change* 2.4.

Falkowski, Paul and Dale A Kiefer (1985). "Chlorophyll a fluorescence in phytoplankton: relationship to photosynthesis and biomass". In: *Journal of Plankton Research* 7.5, pp. 715–731.

Field, Christopher B et al. (1998). "Primary production of the biosphere: integrating terrestrial and oceanic components". In: *science* 281.5374, pp. 237–240.

Fournier, A et al. (2012). "Effect of canopy structure on sun-induced chlorophyll fluorescence". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 68, pp. 112–120.

Frankenberg, C, A Butz, and GC Toon (2011). "Disentangling chlorophyll fluorescence from atmospheric scattering effects in O2 A-band spectra of reflected sun-light". In: *Geophysical Research Letters* 38.3.

Frankenberg, Christian et al. (2011). "New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity". In: *Geophysical Research Letters* 38.17.

Gamon, JA, Josep Penuelas, and CB Field (1992). "A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency". In: *Remote Sensing of environment* 41.1, pp. 35–44.

Gamon, JAm, L Serrano, and JS Surfus (1997). "The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels". In: *Oecologia* 112.4, pp. 492–501.

Gamon, John A and Barbara Bond (2013). "Effects of irradiance and photosynthetic down-regulation on the photochemical reflectance index in Douglas-fir and ponderosa pine". In: *Remote sensing of environment* 135, pp. 141–149.

Gilerson, A et al. (2007). "Fluorescence component in the reflectance spectra from coastal waters. Dependence on water composition". In: *Optics Express* 15.24, pp. 15702–15721.

Gilerson, Alexander A et al. (2010). "Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands". In: *Optics express* 18.23, pp. 24109–24125.

Gitelson, Anatoly A, Claus Buschmann, and Hartmut K Lichtenthaler (1998). "Leaf chlorophyll fluorescence corrected for re-absorption by means of absorption and reflectance measurements". In: *Journal of plant physiology* 152.2-3, pp. 283–296.

Gitelson, Anatoly A, Yuri Gritz, and Mark N Merzlyak (2003). "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves". In: *Journal of plant physiology* 160.3, pp. 271–282.

Gitelson, Anatoly A, Daniela Gurlin, et al. (2009). "A bio-optical algorithm for the remote estimation of the chlorophyll-a concentration in case 2 waters". In: *Environmental Research Letters* 4.4, p. 045003.

Gitelson, Anatoly A, Andrés Vina, et al. (2005). "Remote estimation of canopy chlorophyll content in crops". In: *Geophysical Research Letters* 32.8.

Gitelson, Anatoly A, Andrés Viña, et al. (2003). "Remote estimation of leaf area index and green leaf biomass in maize canopies". In: *Geophysical research letters* 30.5.

Goward, Samuel N and Karl F Huemmrich (1992). "Vegetation canopy PAR absorptance and the normalized difference vegetation index: an assessment using the SAIL model". In: *Remote sensing of environment* 39.2, pp. 119–140.

Gower, J FR, L Brown, and GA Borstad (2004). "Observation of chlorophyll fluorescence in west coast waters of Canada using the MODIS satellite sensor". In: *Canadian Journal of Remote Sensing* 30.1, pp. 17–25.

Gower, JFR (1980). "Observations of in situ fluorescence of chlorophyll-a in Saanich Inlet". In: *Boundary-Layer Meteorology* 18.3, pp. 235–245.

Gower, Jim (2016). "On the use of satellite-measured chlorophyll fluorescence for monitoring coastal waters". In: *International Journal of Remote Sensing* 37.9, pp. 2077–2086.

Grace, J et al. (2007). "Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence?" In: *Global Change Biology* 13.7, pp. 1484–1497.

Guanter, Luis, Christian Frankenberg, et al. (2012). "Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements". In: *Remote Sensing of Environment* 121, pp. 236–251.

Guanter, Luis, Micol Rossini, et al. (2013). "Using field spectroscopy to assess the potential of statistical approaches for the retrieval of sun-induced chlorophyll fluorescence from ground and space". In: *Remote Sensing of Environment* 133, pp. 52–61.

Guanter, Luis, Yongguang Zhang, et al. (2014). "Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence". In: *Proceedings of the National Academy of Sciences* 111.14, E1327–E1333.

Guilizzoni, Piero et al. (2011). "Use of sedimentary pigments to infer past phosphorus concentration in lakes". In: *Journal of Paleolimnology* 45.4, pp. 433–445.

Gupana, Remika S et al. (2021). "Remote sensing of sun-induced chlorophyll-a fluorescence in inland and coastal waters: Current state and future prospects". In: *Remote Sensing of Environment* 262, p. 112482.

Hendrickson, Luke, Robert T Furbank, and Wah Soon Chow (2004). "A simple alternative approach to assessing the fate of absorbed light energy using chlorophyll fluorescence". In: *Photosynthesis research* 82.1, pp. 73–81.

Huot, Yannick and Marcel Babin (2010). "Overview of fluorescence protocols: theory, basic concepts, and practice". In: *Chlorophyll a fluorescence in aquatic sciences: Methods and applications*. Springer, pp. 31–74.

Jiang, Zhangyan et al. (2006). "Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction". In: *Remote sensing of environment* 101.3, pp. 366–378.

Joiner, J et al. (2014). "The seasonal cycle of satellite chlorophyll fluorescence observations and its relationship to vegetation phenology and ecosystem atmosphere carbon exchange". In: *Remote Sensing of Environment* 152, pp. 375–391.

Kiefer, DA, WS Chamberlin, and CR Booth (1989). "Natural fluorescence of chlorophyll a: Relationship to photosynthesis and chlorophyll concentration in the western South Pacific gyre". In: *Limnology and Oceanography* 34.5, pp. 868–881.

Kirk, John TO (1994). *Light and photosynthesis in aquatic ecosystems*. Cambridge university press.

Kishino, Motoaki et al. (1985). "Estimation of the spectral absorption coefficients of phytoplankton in the sea". In: *Bulletin of marine science* 37.2, pp. 634–642.

Krause, GH and E1 Weis (1991). "Chlorophyll fluorescence and photosynthesis: the basics". In: *Annual review of plant biology* 42.1, pp. 313–349.

Kritten, Lena, Rene Preusker, and Jürgen Fischer (2020). "A new Retrieval of Sun-induced Chlorophyll Fluorescence in Water from Ocean Colour Measurements applied on OLCI L-1b and L-2". In: *Remote Sensing* 12.23, p. 3949.

Lakshmi, E et al. (2014). "Time Series Analysis Of Primary Productivity Along The East Coast Of India Using Oceansat-2 Ocean Colour Monitor (O cm)". In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40.8, p. 1049.

Letelier, Ricardo M and Mark R Abbott (1996). "An analysis of chlorophyll fluorescence algorithms for the Moderate Resolution Imaging Spectrometer (MODIS)". In: *Remote Sensing of Environment* 58.2, pp. 215–223.

Li, Zhaohui et al. (2020). "Solar-induced chlorophyll fluorescence and its link to canopy photosynthesis in maize from continuous ground measurements". In: *Remote Sensing of Environment* 236, p. 111420.

Liu, Liangyun et al. (2017). "Assessing the wavelength-dependent ability of solar-induced chlorophyll fluorescence to estimate the GPP of winter wheat at the canopy level". In: *International Journal of Remote Sensing* 38.15, pp. 4396–4417.

Liu, Xinjie, Luis Guanter, et al. (2019). "Downscaling of solar-induced chlorophyll fluorescence from canopy level to photosystem level using a random forest model". In: *Remote Sensing of Environment* 231, p. 110772.

Liu, Xinjie, Liangyun Liu, et al. (2015). "New spectral fitting method for full-spectrum solar-induced chlorophyll fluorescence retrieval based on principal components analysis". In: *Remote Sensing* 7.8, pp. 10626–10645.

Lorenzen, Carl J (1967). "Determination of chlorophyll and pheo-pigments: spectrophotometric equations 1". In: *Limnology and oceanography* 12.2, pp. 343–346.

Malenovsky, Zbynvek et al. (2009). "Scientific and technical challenges in remote sensing of plant canopy reflectance and fluorescence". In: *Journal of experimental botany* 60.11, pp. 2987–3004.

Marchetto, Aldo et al. (2004). "Lake Maggiore (N. Italy) trophic history: fossil diatom, plant pigments, and chironomids, and comparison with long-term limnological data". In: *Quaternary International* 113.1, pp. 97–110.

Maritorena, Stéphane, Andre Morel, and Bernard Gentili (2000). "Determination of the fluorescence quantum yield by oceanic phytoplankton in their natural habitat". In: *Applied Optics* 39.36, pp. 6725–6737.

Maxwell, Kate and Giles N Johnson (2000). "Chlorophyll fluorescence—a practical guide". In: *Journal of experimental botany* 51.345, pp. 659–668.

Mazzoni, M et al. (2012). "Retrieval of maize canopy fluorescence and reflectance by spectral fitting in the O2–A absorption band". In: *Remote sensing of environment* 124, pp. 72–82.

Mazzoni, Marina, Pierluigi Falorni, and Samuele Del Bianco (2008). "Sun-induced leaf fluorescence retrieval in the O 2-B atmospheric absorption band". In: *Optics Express* 16.10, pp. 7014–7022.

Mazzoni, Marina, Pierluigi Falorni, and Wouter Verhoef (2010). "High-resolution methods for fluorescence retrieval from space". In: *Optics express* 18.15, pp. 15649–15663.

Meroni, M, L Busetto, et al. (2010). "Performance of spectral fitting methods for vegetation fluorescence quantification". In: *Remote Sensing of Environment* 114.2, pp. 363–374.

Meroni, M, M Rossini, et al. (2009). "Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications". In: *Remote Sensing of Environment* 113.10, pp. 2037–2051.

Mishra, Deepak R et al. (2005). "Characterizing the vertical diffuse attenuation coefficient for downwelling irradiance in coastal waters: Implications for water penetration by high resolution satellite data". In: *ISPRS Journal of photogrammetry and remote sensing* 60.1, pp. 48–64.

Mobley, Curtis D et al. (2004). "Optical modeling of ocean waters: Is the case 1-case 2 classification still useful?" In: *Oceanography* 17.SPL. ISS. 2, p. 60.

Mohammed, Gina H et al. (2019). "Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress". In: *Remote sensing of environment* 231, p. 111177.

Morabito, Giuseppe, Alessandro Oggioni, et al. (2007). "Seasonal morphological plasticity of phytoplankton in Lago Maggiore (N. Italy)". In: *Hydrobiologia* 578.1, pp. 47–57.

Morabito, Giuseppe, Michela Rogora, et al. (2018). "Could the extreme meteorological events in Lake Maggiore watershed determine a climate-driven eutrophication process?" In: *Hydrobiologia* 824.1, pp. 163–175.

Morel, Anclré and Louis Prieur (1977). "Analysis of variations in ocean color 1". In: *Limnology and oceanography* 22.4, pp. 709–722.

Morrison, J Ruairidh (2003). "In situ determination of the quantum yield of phytoplankton chlorophyll a fluorescence: A simple algorithm, observations, and a model". In: *Limnology and Oceanography* 48.2, pp. 618–631.

O'Reilly, John E et al. (2000). "Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4". In: *SeaWiFS postlaunch calibration and validation analyses, Part* 3, pp. 9–23.

Papageorgiou, George C and Rahni Govindjee (2004). *Chlorophyll a fluorescence: a signature of photosynthesis*. Vol. 19. Springer.

Penuelas, Josep, Iolanda Filella, and John A Gamon (1995). "Assessment of photosynthetic radiation-use efficiency with spectral reflectance". In: *New Phytologist* 131.3, pp. 291–296.

PERICOLOSE, INDAGINI SU DDT E SOSTANZE and DEL LAGO MAGGIORE NELL'ECOSISTEMA (n.d.). "Commissione internazionale per la protezione delle acque italo-svizzere". In: ().

Pettorelli, Nathalie et al. (2005). "Using the satellite-derived NDVI to assess ecological responses to environmental change". In: *Trends in ecology & evolution* 20.9, pp. 503–510.

Pinto, Francisco et al. (2020). "Dynamics of sun-induced chlorophyll fluorescence and reflectance to detect stress-induced variations in canopy photosynthesis". In: *Plant, cell & environment* 43.7, pp. 1637–1654.

Porcar-Castell, Albert et al. (2014). "Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges". In: *Journal of experimental botany* 65.15, pp. 4065–4095.

Ramos, Maria Eva and Maria Gabriela Lagorio (2004). "True fluorescence spectra of leaves". In: *Photochemical & Photobiological Sciences* 3.11-12, pp. 1063–1066.

Rascher, Uwe et al. (2015). "Sun-induced fluorescence–a new probe of photosynthesis: First maps from the imaging spectrometer HyPlant". In: *Global change biology* 21.12, pp. 4673–4684.

Rautiainen, Miina et al. (2010). "Ecological applications of physically based remote sensing methods". In: *Scandinavian Journal of Forest Research* 25.4, pp. 325–339.

Remelli, William and Stefano Santabarbara (2018). "Excitation and emission wavelength dependence of fluorescence spectra in whole cells of the cyanobacterium Synechocystis sp. PPC6803: Influence on the estimation of photosystem II maximal quantum efficiency". In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1859.11, pp. 1207–1222.

Reynolds, Colin S (2006). *The ecology of phytoplankton*. Cambridge University Press.

Romero, Juan M, Gabriela B Cordon, and M Gabriela Lagorio (2018). "Modeling re-absorption of fluorescence from the leaf to the canopy level". In: *Remote Sensing of Environment* 204, pp. 138–146.

— (2020). "Re-absorption and scattering of chlorophyll fluorescence in canopies: A revised approach". In: *Remote Sensing of Environment* 246, p. 111860.

Rossini, Micol, Michele Meroni, Marco Celesti, et al. (2016). "Analysis of red and far-red sun-induced chlorophyll fluorescence and their ratio in different canopies based on observed and modeled data". In: *Remote Sensing* 8.5, p. 412.

Rossini, Micol, Michele Meroni, Mirco Migliavacca, et al. (2010). "High resolution field spectroscopy measurements for estimating gross ecosystem production in a rice field". In: *Agricultural and Forest Meteorology* 150.9, pp. 1283–1296.

Rossini, Micol, L Nedbal, et al. (2015). "Red and far red Sun-induced chlorophyll fluorescence as a measure of plant photosynthesis". In: *Geophysical research letters* 42.6, pp. 1632–1639.

Roy-Barman, M (2009). "Modelling the effect of boundary scavenging on Thorium and Protactinium profiles in the ocean". In: *Biogeosciences* 6.12, pp. 3091–3107.

Ruddick, Kevin G et al. (2019). "A review of protocols for fiducial reference measurements of water-leaving radiance for validation of satellite remote-sensing data over water". In: *Remote Sensing* 11.19, p. 2198.

Ruggiu, D et al. (1998). "Trends and relations among basic phytoplankton characteristics in the course of the long-term oligotrophication of Lake Maggiore (Italy)". In: *Phytoplankton and Trophic Gradients*. Springer, pp. 243–257.

Salmaso, Nico and Rosario Mosello (2010). "Limnological research in the deep southern subalpine lakes: synthesis, directions and perspectives". In: *Advances in Oceanography and Limnology* 1.1, pp. 29–66.

Santabarbara, Stefano et al. (2017). "Kinetics and heterogeneity of energy transfer from light harvesting complex II to photosystem I in the supercomplex isolated from Arabidopsis". In: *Physical Chemistry Chemical Physics* 19.13, pp. 9210–9222.

Siegmann, Bastian et al. (2021). "Downscaling of far-red solar-induced chlorophyll fluorescence of different crops from canopy to leaf level using a diurnal data set acquired by the airborne imaging spectrometer HyPlant". In: *Remote sensing of environment* 264, p. 112609.

Smetacek, Victor (1999). "Diatoms and the ocean carbon cycle". In: *Protist* 150.1, pp. 25–32.

Smith, RC et al. (1989). "Bio-optical modeling of photosynthetic production in coastal waters". In: *Limnology and Oceanography* 34.8, pp. 1524–1544.

Strömbeck, Niklas and Donald C Pierson (2001). "The effects of variability in the inherent optical properties on estimations of chlorophyll a by remote sensing in Swedish freshwaters". In: *Science of the total environment* 268.1-3, pp. 123–137.

Sun, Ying et al. (2018). "Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP". In: *Remote Sensing of Environment* 209, pp. 808–823.

Tenjo, Carolina et al. (2021). "A new algorithm for the retrieval of sun induced chlorophyll fluorescence of water bodies exploiting the detailed spectral shape of water-leaving radiance". In: *Remote Sensing* 13.2, p. 329.

Trüper, HG and CHARLES S Yentsch (1967). "Use of glass fiber filters for the rapid preparation of in vivo absorption spectra of photosynthetic bacteria". In: *Journal of bacteriology* 94.4, pp. 1255–1256.

Utermöhl, Hans (1958). "Zur vervollkommnung der quantitativen phytoplankton-methodik: Mit 1 Tabelle und 15 abbildungen im Text und auf 1 Tafel". In: *Internationale Vereinigung für theoretische und angewandte Limnologie: Mitteilungen* 9.1, pp. 1–38.

Van der Tol, C, JA Berry, et al. (2014). "Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence". In: *Journal of Geophysical Research: Biogeosciences* 119.12, pp. 2312–2327.

Van der Tol, C, W Verhoef, et al. (2009). "An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance". In: *Biogeosciences* 6.12, pp. 3109–3129.

Verhoef, Wout and Heike Bach (2007). "Coupled soil–leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data". In: *Remote Sensing of Environment* 109.2, pp. 166–182.

— (2012). "Simulation of Sentinel-3 images by four-stream surface–atmosphere radiative transfer modeling in the optical and thermal domains". In: *Remote sensing of environment* 120, pp. 197–207.

Verhoef, Wouter, Christiaan Van Der Tol, and Elizabeth M Middleton (2018). "Hyperspectral radiative transfer modeling to explore the combined retrieval of biophysical parameters and canopy fluorescence from FLEX–Sentinel-3 tandem mission multi-sensor data". In: *Remote sensing of environment* 204, pp. 942–963.

Verrelst, Jochem, Jordi Muñoz, et al. (2012). "Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3". In: *Remote Sensing of Environment* 118, pp. 127–139.

Verrelst, Jochem, Juan Pablo Rivera, et al. (2015). "Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence?" In: *Remote Sensing of Environment* 166, pp. 8–21.

Vilfan, Nastassia et al. (2016). "Fluspect-B: A model for leaf fluorescence, reflectance and transmittance spectra". In: *Remote Sensing of Environment* 186, pp. 596–615.

Viña, Andrés and Anatoly A Gitelson (2005). "New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops". In: *Geophysical Research Letters* 32.17.

Wu, Genghong et al. (2020). "Radiance-based NIRv as a proxy for GPP of corn and soybean". In: *Environmental Research Letters* 15.3, p. 034009.

Xing, Xiao-Gang et al. (2007). "An overview of remote sensing of chlorophyll fluorescence". In: *Ocean Science Journal* 42.1, pp. 49–59.

Xu, Shan et al. (2018). "Diurnal response of sun-induced fluorescence and PRI to water stress in maize using a near-surface remote sensing platform". In: *Remote Sensing* 10.10, p. 1510.

Yang, Peiqi et al. (2020). "Fluorescence Correction Vegetation Index (FCVI): A physically based reflectance index to separate physiological and non-physiological information in far-red sun-induced chlorophyll fluorescence". In: *Remote sensing of environment* 240, p. 111676.

Yang, Xi et al. (2015). "Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest". In: *Geophysical Research Letters* 42.8, pp. 2977–2987.

Zarco-Tejada, Pablo J, A Catalina, et al. (2013). "Relationships between net photosynthesis and steady-state chlorophyll fluorescence retrieved from airborne hyperspectral imagery". In: *Remote Sensing of Environment* 136, pp. 247–258.

Zarco-Tejada, Pablo J, Victoria González-Dugo, et al. (2013). "A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index". In: *Remote sensing of environment* 138, pp. 38–50.

Zhao, Feng, Yiqing Guo, et al. (2014). "A method to reconstruct the solar-induced canopy fluorescence spectrum from hyperspectral measurements". In: *Remote sensing* 6.10, pp. 10171–10192.

Zhao, Feng, Rong Li, et al. (2018). "Reconstruction of the full spectrum of solar-induced chlorophyll fluorescence: Intercomparison study for a novel method". In: *Remote sensing of environment* 219, pp. 233–246.

Zhaoxin, Li et al. (n.d.). "Remote estimation of phytoplankton gross primary production in clear to turbid waters by integrating an analytical model and a machine-learning algorithms". In: *Remote Sensing of Environment (submitted)* ().

Zibordi, G et al. (2012). "In situ determination of the remote sensing reflectance: an intercomparison". In: *Ocean Science* 8.4, pp. 567–586.

Znachor, Petr and Jiri Nedoma (2008). "APPLICATION OF THE PDMPO TECHNIQUE IN STUDYING SILICA DEPOSITION IN NATURAL POPULATIONS OF FRAGILARIA CROTONENSIS (BACILLARIOPHYCEAE) AT DIFFERENT DEPTHS IN A EUTROPHIC RESERVOIR 1". In: *Journal of phycology* 44.2, pp. 518–525.

Znachor, Petr, Jiri Nedoma, et al. (2012). "Bacterial colonization of the freshwater planktonic diatom Fragilaria crotonensis". In: *Aquatic microbial ecology* 66.1, pp. 87–94.