



Simulation framework for solar-induced fluorescence retrieval and application to DESIS and HyPlant

Miguel Pato ^{a,*}, Kevin Alonso ^b, Jim Buffat ^{c,d}, Stefan Auer ^a, Emiliano Carmona ^a, Stefan Maier ^a, Rupert Müller ^a, Patrick Rademske ^c, Uwe Rascher ^{c,e}, Hanno Scharr ^d

^a Remote Sensing Technology Institute, German Aerospace Center (DLR), Oberpfaffenhofen, Germany

^b European Space Agency (ESA), Largo Galileo Galilei, Frascati 00044, Italy

^c Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Jülich, Germany

^d Forschungszentrum Jülich GmbH, Institute of Advanced Simulations, IAS-8: Data Analytics and Machine Learning, Jülich, Germany

^e University of Bonn, Faculty of Agricultural, Nutritional and Engineering Sciences, Institute of Crop Science and Resource Conservation, Bonn, Germany

ARTICLE INFO

Edited by Jing M. Chen

Keywords:

Solar-induced fluorescence
Hyperspectral sensors
Radiative transfer
Machine learning
Simulation

ABSTRACT

Fluorescence light emitted by chlorophyll in plants is a direct probe of the photosynthetic process and can be used to continuously monitor vegetation status. Retrieving solar-induced fluorescence (SIF) using a machine learning (ML) approach promises to take full advantage of airborne and satellite-based instruments to map expected vegetation function over wide areas on a regular basis. This work takes a first step towards developing a ML-based SIF retrieval method. A general-purpose framework for the simulation of at-sensor radiances is introduced and applied to the case of SIF retrieval in the oxygen absorption band O₂-A with the spaceborne DESIS and airborne HyPlant spectrometers. The sensor characteristics are modelled carefully based on calibration and in-flight data and can be extended to other instruments including the upcoming FLEX mission. A comprehensive dataset of simulated at-sensor radiance spectra is then assembled encompassing the most important atmosphere, geometry, surface and sensor properties. The simulated dataset is employed to train emulators capable of generating at-sensor radiances with sub-percent errors in tens of μ s, opening the way for their routine use in SIF retrieval. The simulated spectra are shown to closely reproduce real data acquired by DESIS and HyPlant and can ultimately be used to develop a robust ML-based SIF retrieval scheme for these and other remote sensing spectrometers. Finally, the SIF retrieval performance of the 3FLD method is quantitatively assessed for different on- and off-band configurations in order to identify the best band combinations. This highlights how our simulation framework enables the optimization of SIF retrieval methods to achieve the best possible performance for a given instrument.

1. Introduction

Methods to quantify solar-induced fluorescence (SIF) have gained great interest in the remote sensing community over the last years. This emerging attention was triggered by the selection of the FLEX mission of the European Space Agency (Drusch et al., 2017) to become the first dedicated Earth Explorer satellite mission to quantify SIF globally in moderate spatial resolution (300 m) and by the achievements in retrieving SIF from existing satellite-borne instruments (see Joiner et al. (2013), Köhler et al. (2020), Guanter et al. (2021) and references therein). The fluorescence emitted by green vegetation originates as a weak signal from the chlorophyll pigments of photosynthetic active plant material in the wavelength range between 670 nm and 780 nm. Although the signal is much less intense than reflected light at these

wavelengths, it is directly dependent on the efficiency of light reactions during photosynthesis (see Mohammed et al. (2019), Porcar-Castell et al. (2021) for recent reviews). SIF has therefore been shown to have high potential not only to monitor the greenness of vegetation but also for the early detection of stress-induced impairment of photosynthesis (Ač et al., 2015; Pinto et al., 2020; Zeng et al., 2022) and to help estimate carbon and water fluxes within the green vegetation layer (Damm et al., 2018; Maes et al., 2020; Martini et al., 2022; Norton et al., 2019).

Despite its great potential to track the actual state of photosynthesis, it remains challenging to measure top-of-canopy SIF accurately without any hidden bias and to relate it to physiologically relevant leaf level SIF. It was thus proposed to combine SIF with novel reflectance-based

* Corresponding author.

E-mail address: Miguel.FigueiredoVazPato@dlr.de (M. Pato).

indices (Zeng et al., 2021, 2022). The challenges of detecting SIF from remote platforms is firstly related to the intrinsic weakness of the fluorescence signal that needs to be separated from the quantitatively larger intensity of reflected photons (the SIF signal is approximately only 2–5% of the total photons in this spectral region). The separation of SIF from reflectance is generally done in solar and/or atmospheric absorption regions, where the intensity of incoming and reflected light is greatly reduced. However, resolving these fine absorption bands requires high spectral resolution sensors, which in turn generally suffer from a low signal-to-noise ratio (SNR) if high spatial resolution is desired. Finally, on its path to the sensor, the top-of-canopy SIF signal itself is partly absorbed by the atmosphere, rendering the retrieval of SIF from airborne or satellite data a complex task.

Dedicated SIF instrumentation has been developed with an optimized balance between spectral resolution and SNR (Cendrero-Mateo et al., 2019), but in parallel several attempts have been undertaken to also use medium resolution instruments to retrieve SIF or at least signals that contain parts of the fluorescence emission (Damm et al., 2014). Nowadays, we have different systems which enable continuous recording of SIF from towers and ground-based systems (Rossini et al., 2016; Julitta et al., 2016; Peng et al., 2022). Additionally, dedicated airborne sensors for SIF measurements have become available (Rascher et al., 2015; Frankenberg et al., 2018; Siegmann et al., 2019) and smaller sensors which can be mounted on small unmanned aircrafts are currently on the fringe of becoming available (Garzonio et al., 2017; Bendig et al., 2020; Vargas et al., 2020; Wang et al., 2022). Retrieval of the SIF signal is generally done using the Fraunhofer Line Discrimination method (FLD) and an expansion of it (Maier et al., 2003), which was subsequently further modified to account for the non-linearity of the reflectance and fluorescence signals in the red spectral region (Cogliati et al., 2015b, 2019). However, it remains a challenge for these retrieval methods to account for the modulation and reabsorption of the signal in the downwelling and upwelling path of photons through the atmosphere (Sabater et al., 2018). Even though great advances have been made in the past years to constrain the atmospheric absorption of the signal (Sabater et al., 2021), several activities are currently ongoing to use machine learning (ML) and deep learning approaches to handle atmospheric effects and retrieve the SIF signal (Scodellaro et al., 2022; Vicent Servera et al., 2022; Buffat et al., 2023, 2025a,b). For these ML approaches to reach the accuracy required for SIF retrieval it is crucial to have abundant high-quality training data with accurate detector characterization.

SIF retrieval can be seen as a complex, multi-dimensional inversion problem. We aim to eventually develop a ML method for SIF retrieval capable of mapping fluorescence over extended regions and characterizing its uncertainty. This is currently not possible with other methods. The ML model will need to learn and abstract away the complicated influence of the atmosphere on the measured signal and the sensor spectral uncertainties in order to extract the SIF signal. The present work takes the first step in this effort, namely the preparation of a meaningful and thorough body of training data for the learning process. We start by implementing a general-purpose simulation framework for at-sensor radiances in any reflective spectral range and for any instrument. Our focus here is on two different classes of spectrometers of relevance for SIF retrieval, represented by DESIS (Krutz et al., 2019) and HyPlant (Siegmann et al., 2019). On the one hand, DESIS is an example of a space-based spectrometer capable of covering wide areas around the globe at moderate spatial and spectral resolutions. On the other hand, HyPlant typifies a group of airborne spectrometers of high spectral resolution that can cover a handful of small regions at high spatial resolution. Although other classes of instruments may be employed in the study of fluorescence, DESIS and HyPlant encompass a representative range of current and future capabilities. A precise sensor model is then developed to simulate high-fidelity DESIS and HyPlant data that closely resemble real measurements.

Two applications of relevance for both traditional and novel SIF retrieval methodologies are explored using the assembled simulated data. First, we show that it is possible to train fast and accurate at-sensor radiance emulators, thereby providing a fast alternative to computationally demanding radiative transfer modelling in SIF retrieval methods. Second, the SIF retrieval performance of the 3FLD method (Maier et al., 2003) is evaluated using our simulations, illustrating the assessment of a SIF retrieval method with the simulation dataset presented in this contribution. The framework and results presented here have been the starting point for several of our recent works towards a robust ML-based SIF retrieval method, including the development of fast and accurate emulators (Pato et al., 2023, 2024) and a novel ML-based SIF retrieval method for HyPlant (Buffat et al., 2025c, 2023, 2025b) and DESIS (Buffat et al., 2025a).

It is important to explain our choice of the 3FLD method for SIF retrieval assessment in the current study as opposed to other state-of-the-art algorithms (Alonso et al., 2008; Mazzoni et al., 2012; Meroni et al., 2010; Cogliati et al., 2015b, 2018, 2019; Buffat et al., 2025c). Although more modern approaches are nowadays routinely used for the derivation of fluorescence from ground-based or airborne instruments (see e.g. Siegmann et al. (2019)), the 3FLD retrieval is straightforward, spectrum-based and not explicitly coupled to atmosphere conditions or observation geometries through radiative transfer modelling. In addition, the method is typically robust and does not create retrieval artifacts. These features make 3FLD the ideal choice for our purposes of illustrating the assessment of SIF retrieval accuracy with our simulated datasets in a simple fashion. The evaluation of more modern methods with the simulated data is also a relevant line of research which is left for future work.

The paper is organized as follows. We start by describing the simulation framework and its application to the case of SIF retrieval in Section 2, including radiative transfer, surface properties and sensor specifics. The dataset of simulated at-sensor radiance spectra for DESIS and HyPlant is presented in Section 2.3 including the adopted input parameter ranges and different sampling strategies. Section 3 then documents in detail our results and findings. Two direct applications of the simulated data are showcased, namely the training of forward emulators of at-sensor radiances (Section 3.1) and the evaluation of the SIF retrieval performance of the 3FLD method (Section 3.2). The simulated data are shown to be high-fidelity proxies of real data routinely acquired by DESIS and HyPlant in Section 3.1. Section 4 discusses the results and further elaborates on how the simulated datasets can form the basis for a ML-based SIF retrieval method. Our concluding remarks are given in Section 5. In Appendix technical details are given regarding the sensitivity analysis that informed our simulations.

2. Data and methods

2.1. Simulation framework

The radiance measured by an airborne or satellite-based sensor depends in general on the atmospheric conditions, observation geometry, surface properties and sensor characteristics:

$$L_s = L_s(x_{\text{atm}}, x_{\text{geo}}, x_{\text{sur}}, x_{\text{sen}}) \quad (1)$$

where L_s is the at-sensor radiance spectrum and x_{atm} , x_{geo} , x_{sur} and x_{sen} represent the atmosphere, geometry, surface and sensor parameters, respectively. The complicated multi-dimensional dependence has to be accounted for in any simulation of at-sensor radiances. Fortunately, the problem can be decoupled into a part that depends only on atmosphere and geometry and another part that depends only on surface and sensor properties. This decoupling permits a more efficient simulation and a more thorough exploration of the high-dimensional input parameter space. We have accordingly developed a two-module simulation tool to deliver at-sensor radiances in a defined spectral range given a set of atmosphere, geometry, surface and sensor parameters. Note that the

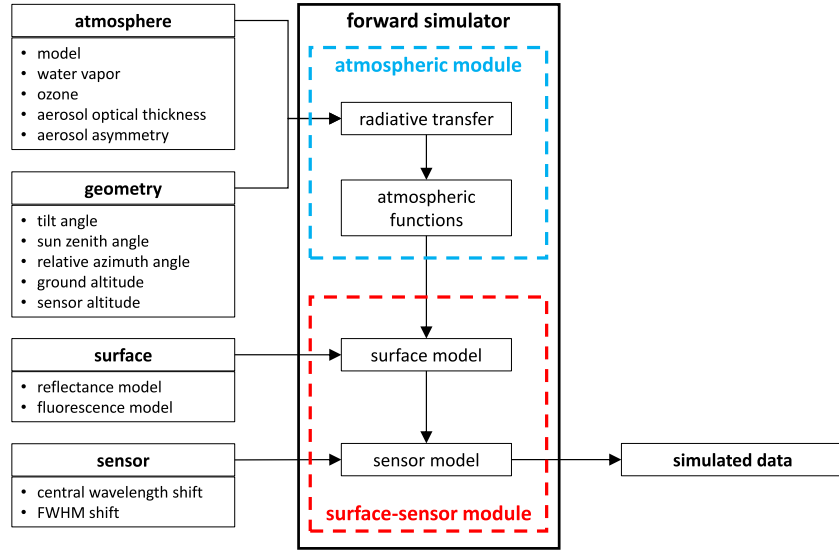


Fig. 1. Schematic diagram of the software tool developed to simulate at-sensor radiances depending on atmosphere, geometry, surface and sensor parameters.

framework is general and may be easily applied to any optical remote sensing application, wavelength range or sensor. A diagram of our simulation setup is shown in Fig. 1. In the following we describe the two modules of the tool: the atmospheric module and the surface-sensor module.

2.1.1. Atmospheric module

The at-sensor radiance may be modelled for our purposes as the combination of three components, namely surface-reflected light, surface emission and atmosphere-scattered light (generically called path radiance). Accordingly, we write the radiance measured by a detector element at a given wavelength as

$$L = L_p + (L_R(\rho) + L_F)T^\dagger, \quad (2)$$

where ρ is the surface reflectance, L_R is the reflected on-ground radiance, L_F is the SIF on-ground radiance, L_p is the path radiance and $T^\dagger$ is the total transmission coefficient from surface to sensor (typically decomposed into direct and diffuse components, $T^\dagger = T_{\text{dir}}^\dagger + T_{\text{dif}}^\dagger$). Following Guanter et al. (2009), the reflectance component L_R may be further expressed in terms of the global solar irradiance on the ground E_g^0 (comprising direct and diffuse components) and the spherical albedo of the atmosphere S :

$$L_R(\rho) = \frac{E_g^0 \rho}{\pi(1 - \rho S)}, \quad (3)$$

with the implicit assumption of a Lambertian surface. Notice that all variables in the above equations are wavelength dependent and that any multiple scattering effects are implicitly included in the definition of L_p , E_g^0 and $T^\dagger$. The dependence on the top-of-atmosphere solar irradiance enters in E_g^0 and L_p . Any solar model can be selected in our tool. In this work, we have used the recent TSIS-1 solar model (Coddington et al., 2023).

The five so-called atmospheric functions L_p , E_g^0 , $T_{\text{dir}}^\dagger$, $T_{\text{dif}}^\dagger$ and S depend solely on atmosphere and observation geometry, not on surface nor sensor specifics. Therefore, the atmospheric functions can be computed once at high spectral resolution and then applied for multiple surface and sensor configurations, as long noticed and done in the remote sensing literature (Guanter et al., 2009; de los Reyes et al., 2020). This computation typically requires a radiative transfer code to account for the absorption and scattering of light throughout its path in the atmosphere. We choose to work with MODTRAN6 (Berk et al., 2014, 2015) for this step given its wide use, versatility and high spectral resolution, but other atmosphere radiative transfer models

could be used in our general simulation framework. MODTRAN, as other models, does not directly provide the atmospheric functions; these have to be derived from the outputs of two distinct MODTRAN runs with different (ideally non-zero) reflectance values and no surface emission. A full account of this procedure can be found in Guanter et al. (2009). Overall, as sketched in Fig. 1, the atmospheric module takes as input atmosphere and observation geometry parameters (virtually any specification that can be passed to MODTRAN), runs MODTRAN6 twice, and derives the five atmospheric functions L_p , E_g^0 , $T_{\text{dir}}^\dagger$, $T_{\text{dif}}^\dagger$ and S in a specified range of wavelengths at high spectral resolution. These data are stored in an internal database, which we shall call ATM DB, and constitute the starting point for the second module of the simulation tool.

2.1.2. Surface-sensor module

The surface-sensor module expects as input the above mentioned ATM DB as well as surface and sensor properties to deliver the final simulated spectra. Although conceptually distinct, the surface and sensor models are incorporated together into a fast module in our simulator framework.

The surface is characterized by providing models for the wavelength dependence of reflectance ρ and SIF emission L_F . Together with the five atmospheric functions computed by the atmospheric module, this completes the definition of the high-resolution at-sensor radiance spectrum L in Eq. (2). Any parametric model for $\rho(\lambda)$ and $L_F(\lambda)$ may be implemented in the surface-sensor module. The very general definition of the surface properties is designed to support a wide range of applications. The particular models used for our purposes in this work are described in Section 2.2.1.

With the high-resolution at-sensor radiance spectrum $L(\lambda)$ fully specified, the sensor spectral characteristics of the instrument can now be factored in. The spectral response function (SRF) for a spectral band b is modelled as a Gaussian parameterized by a central wavelength CW_b and full width at half maximum $FWHM_b$. In order to accommodate any imperfections in the spectral calibration of the instrument, we allow for small additive shifts for CW and changes in FWHM: $CW'_b = CW_b + \delta_{CW}$ and $FWHM'_b = FWHM_b + \delta_{FWHM}$. The shifts of CW and changes in FWHM are global, i.e. the same for all bands. The spectral response function for the particular band b is given by

$$SRF_b(\lambda) = \frac{1}{\sigma'_b \sqrt{2\pi}} \exp\left(-\frac{(\lambda - CW'_b)^2}{2\sigma'^2_b}\right), \quad (4)$$

where $\sigma'_b = \text{FWHM}'_b / (2\sqrt{\ln 2})$ is the modified standard deviation. The spectral response function and corresponding shifts applied in our work for DESIS and HyPlant are specified in Section 2.2.2. The sensor-resolution at-sensor radiance for band b then reads

$$L_{s,b} = \int d\lambda L(\lambda) \text{SRF}_b(\lambda) \quad (5)$$

There is one final subtlety in the simulation of at-sensor radiances related to the medium where the instrument operates. While MODTRAN reports vacuum wavelengths λ_0 , the spectral response function refers typically to wavelengths in the measurement medium $\lambda = \lambda_0/n_m$, where n_m is the refractive index of the relevant medium. Thus we have to modify Eq. (5) by scaling the vacuum (simulation) wavelengths appropriately:

$$L_{s,b} = \int d\lambda L(\lambda n_m) \text{SRF}_b(\lambda) \quad (6)$$

For instruments operating in vacuum, as in the case of DESIS, $n_m = 1$ and the wavelength difference is irrelevant. The effect is also negligible for sensors with moderate spectral resolutions, because the spectral shift due to refraction is relatively small (0.22 nm at $\lambda = 760$ nm for air at 0 °C and 1 atm). However, for high-resolution instruments operating in air such as HyPlant, it is crucial to consider the air-to-vacuum shift. For HyPlant, we specify the refractive index of the measurement medium as $n_m = n_{\text{air}} = 1.000293$ (air at 0 °C and 1 atm) to model the spectral shift. There is a slight flight altitude dependence that cannot be accounted for with this simple model, but it can with the central wavelength shifts implemented in the surface-sensor module (cf. Eq. (4)). This approach enables a very good match between the simulated at-sensor radiance spectra and the real HyPlant measurements (cf. Section 3.1).

The sensor-resolution simulated spectra $L_{s,b}$ are saved in a dedicated database called SENSOR DB. This final database serves as the basis for the results presented in Section 3.

2.2. Application to SIF

2.2.1. Surface model

The key surface parameters for our purposes are reflectance ρ and fluorescence emission L_F . We use simplified parametric models for the wavelength dependence of both ρ and L_F having in mind the final application to SIF retrieval in the O₂-A band. The choice of simplified models in our framework is intentional since it keeps the number of parameters to a minimum and enables the assembly of very large simulated datasets, which would not be feasible with very detailed radiative transfer models (e.g., SCOPE (van der Tol et al., 2009; Yang et al., 2021)). Furthermore, we are not modelling any relationships between incident solar radiation, vegetation constituents, structure, physiological state, reflectance and fluorescence. This is an important feature of our approach since the absence of any such relationships in our simulated data is necessary for a SIF retrieval method to extract the actual SIF emission signal, and not a correlation between surface reflectance and underlying physical parameters.

Reflectance could be described using a full, coupled leaf and canopy radiative transfer model. However, these models require a large number of parameters which would substantially increase the dimensionality of the input parameter space. A simple reflectance model for vegetated surfaces can be created through linear combination of vegetation and soil spectra. Fortunately, in the spectral range investigated here (740–780 nm) leaf pigments and water only have a minor influence on reflectance (Maier, 2000), leading to spectrally flat spectra. Furthermore, due to the small spectral range, soil reflectance spectra can be described by an affine equation. Consequently, we describe the surface reflectance ρ as

$$\rho(\lambda) = \rho_{740} + s(\lambda - \lambda_1) + \frac{s(e-1)}{2(\lambda_2 - \lambda_1)}(\lambda - \lambda_1)^2 \quad (7)$$

with $\lambda_1 = 740$ nm and $\lambda_2 = 780$ nm. The three parameters ρ_{740} , s and e have clear meaning: ρ_{740} and s are the reflectance and its spectral slope at 740 nm, while e is the ratio of reflectance spectral slopes at 780 nm and 740 nm.

The model in Eq. (7) encompasses a limited but useful range of land covers including vegetation, soil, concrete and bitumen, but not necessarily other artificial materials. This can be seen explicitly by analysing observational reflectances in the spectral range of interest. In order to test our reflectance parameterization, we made use of top-of-canopy (TOC) reflectances derived from the DUAL sensor integrated in HyPlant (cf. Section 2.2.2). In particular, we gathered a dataset of TOC reflectance estimates comprising all DUAL acquisitions (Siegmann et al., 2019) from the years 2018–2022 (469 acquisitions, 41 campaigns) and sampled random subsets consisting of a fixed number of 30,000 spectra. Fig. 2 shows the maximum absolute error obtained by fitting a linear reflectance model ($e = 1$) and a quadratic model ($e > 0$) to the observed reflectances. We excluded the wavelength interval 755–770 nm from the fit since in this range the DUAL reflectances are less accurate due to the influence of the O₂-A absorption feature. Clearly, the linear model fails to provide an accurate description of the reflectance for typical land covers in HyPlant acquisitions. The underlying reason is primarily the curvature of the red edge in the range 740–750 nm, as depicted with an exemplary reflectance curve in Fig. 2 (right). The quadratic model does capture the curvature of typical reflectance spectra around the O₂-A band since the parameter e effectively controls the degree of curvature of the reflectance spectrum. With both s and e positive, this formulation guarantees monotonically increasing reflectance spectra in the considered wavelength range, as expected for vegetation and soil, which we aim to simulate. This is important to avoid introducing unrealistic reflectance spectra in the simulated dataset which could bias a ML SIF retrieval method. Naturally, our simulations should not be used to interpret data from land covers different from those considered in our framework. As can be observed in Fig. 2 (left), the quadratic model has a maximum error of less than 1.2% in the analysed dataset, which makes this model appropriate for our purposes.

Next we determine appropriate parameterization ranges for ρ_{740} , s and e by matching cubes to the fitted distributions of the parameters. While the ranges are derived from the given data distribution, we are confident that the formulation of the reflectance model as well as the large margins around the parameter space covered by this particular dataset are general enough to be valid for other datasets. Fig. 3 shows the marginal distributions $p(s, e)$ and $p(\rho_{740})$. These distributions motivate the parameter space delimited by the red lines in the figure, encompassing $\rho_{740} = [0.05, 0.6]$, $s = [0, 0.012] \text{ nm}^{-1}$ and $e = [0, 1]$.

Solar-induced fluorescence in vegetation shows two peaks at around 690 nm and 737 nm. Fig. 4 illustrates solar-induced fluorescence spectra modelled using the leaf radiative transfer model SLOPE (Maier, 2000) and measured under natural conditions (Amoros-Lopez et al., 2008; Krämer, 2024). Besides a single spectrum from Amoros-Lopez et al. (2008), we compare our model with 161 leaf-level on-ground SIF measurements at high spectral resolution (0.75 nm) collected with the FLUOWat/LOX measurement system (Van Wittenberghe et al., 2013) from different crops in dedicated campaigns in Campus Klein-Altendorf in 2022 and 2023 (Krämer, 2024). Below 730 nm an approximation based on Gaussian functions is not satisfactory. In contrast, above 730 nm, a Gaussian model is a very good approximation for the solar-induced fluorescence spectra. We therefore model SIF emission in the spectral range 740–780 nm as

$$L_F(\lambda) = F_{737} \exp\left(-\frac{(\lambda - \lambda_F)^2}{2\sigma_F^2}\right) \quad (8)$$

with $\lambda_F = 737$ nm and $\sigma_F = 20$ nm. The only free parameter is F_{737} , the SIF on-ground radiance at 737 nm, and we consider the range $F_{737} = [0, 0.8] \text{ mW/cm}^2/\text{sr}/\mu\text{m}$ for our simulations. Note that our simplified SIF model is well motivated for our purposes here given the narrow range 740–780 nm considered around the O₂-A band (where the SIF spectrum

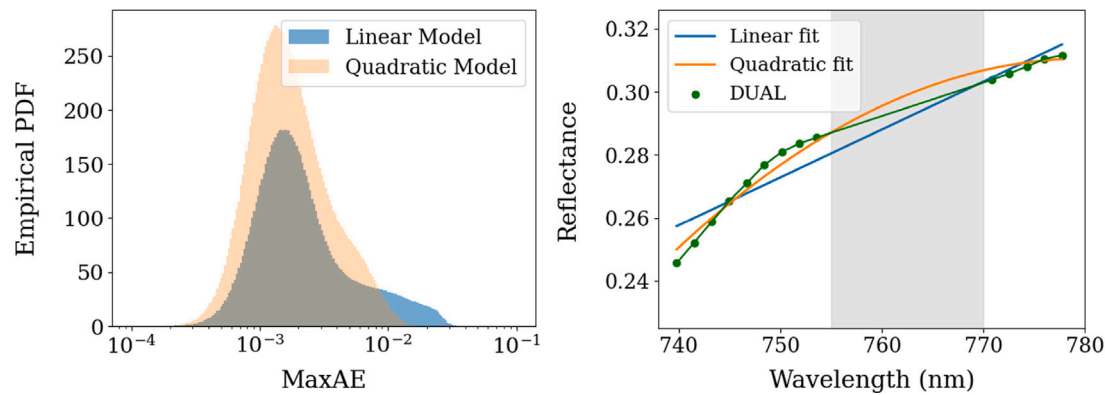


Fig. 2. Reflectance model fits to HyPlant DUAL data in the spectral region around the O_2 -A band. The left panel shows the distribution of the maximum absolute error (MaxAE) in spectral dimension of linear and quadratic fits over 30,000 reflectance spectra collected from DUAL acquisitions between 2018 and 2022. In the right panel, an exemplary DUAL reflectance spectrum (green) is compared to the corresponding linear fit (blue) and quadratic fit (orange). The grey band indicates the spectral range of the O_2 -A absorption feature which was not used in the reflectance model fits.

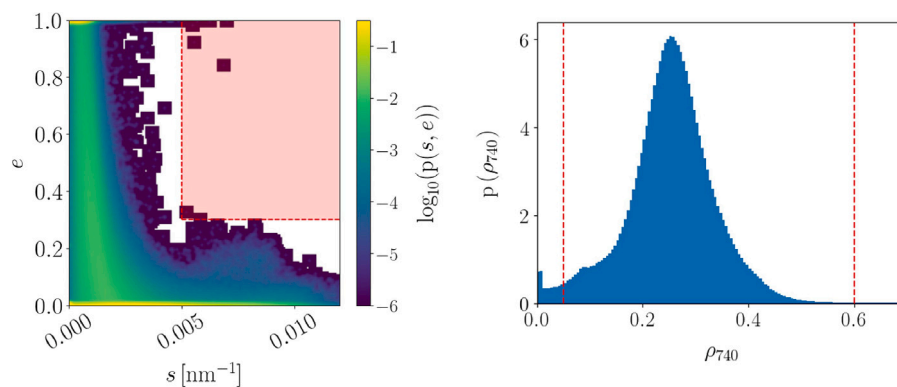


Fig. 3. Distribution of fitted parameters for the reflectance model applied to HyPlant DUAL data. The marginal distribution over the fitted reflectance slope s and slope ratio e is shown on the left, while the marginal distribution over the fitted reflectance ρ_{740} is reported on the right. Dashed red lines mark the borders of the selected parameter space to generate the simulated dataset.

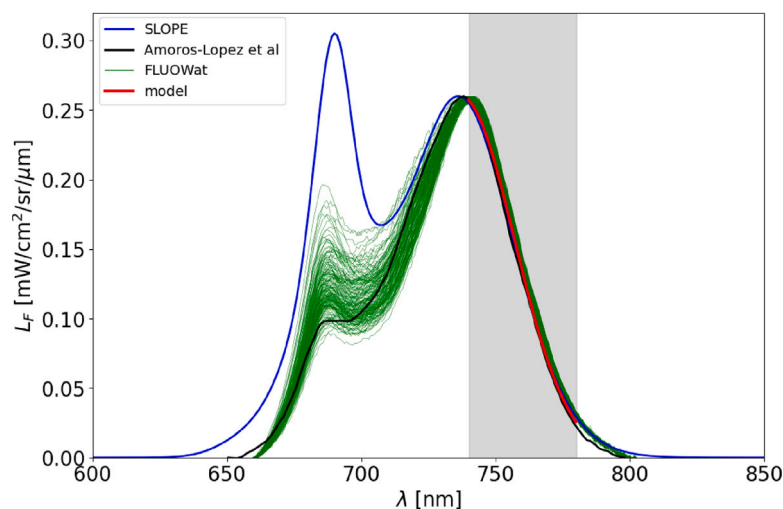


Fig. 4. Solar-induced fluorescence spectra modelled using the leaf radiative transfer model SLOPE (Maier, 2000), measured under natural conditions in Amoroso-Lopez et al. (2008) (one spectrum) and in Krämer (2024) (161 spectra) and a Gaussian model in the considered wavelength range 740–780 nm. All spectra have been scaled to the spectrum from Amoroso-Lopez et al. (2008) for visualization. The measured spectra from Amoroso-Lopez et al. (2008), Krämer (2024) have been baseline corrected to ensure zero fluorescence outside of the fluorescence emission range. The grey shaded area depicts the wavelength range considered in this study.

consists of the decaying tail of the second fluorescence peak, cf. Fig. 4) and the interest in medium spectral resolution instruments such as DESIS. For SIF retrieval at the O_2 -B band or for a full SIF spectrum reconstruction as aimed by the upcoming FLEX, a more complex model is

needed, possibly including parameters for the position, width, intensity and asymmetry of the fluorescence peaks. Our simulation framework may be used for such studies and we briefly comment on this possibility in Section 4.

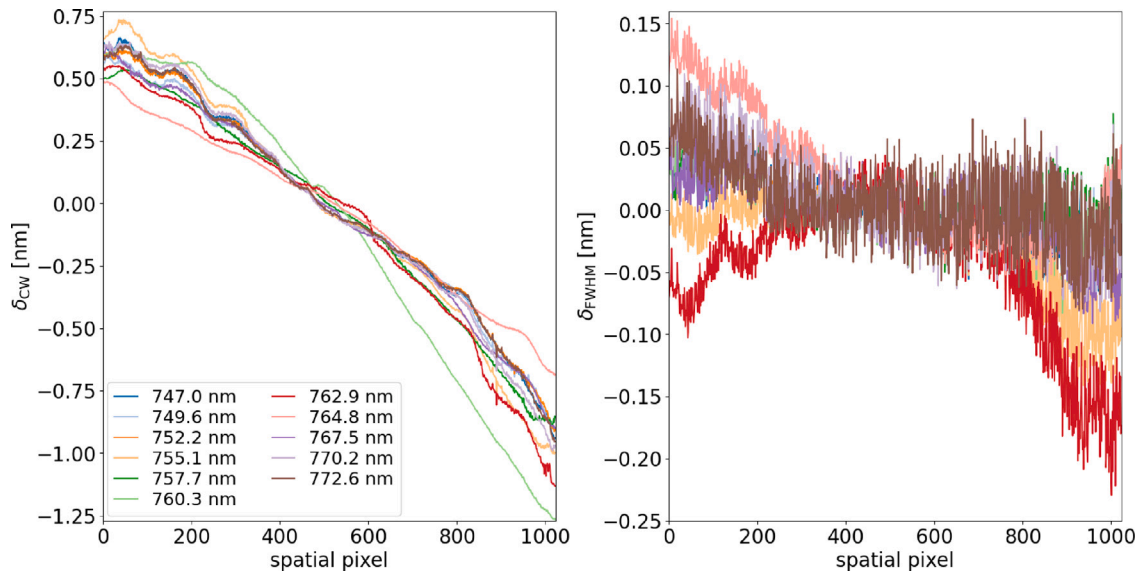


Fig. 5. The variation of the DESIS spectral response function in terms of central wavelength (CW, left) and full width at half maximum (FWHM, right) in the across-track direction for spectral bands near the O₂-A absorption band. The differences δ_{CW} and δ_{FWHM} shown in the plots are defined with respect to the corresponding average over the 400 central pixels. The DESIS spectral response function shown here is based on the spectral calibration table with version 0207 and valid from May 1, 2020 to Dec 1, 2020.

2.2.2. Instruments and data

A concise overview of the key specifications of DESIS and HyPlant around the O₂-A band (specifically, in the spectral range 740–780 nm) is presented next. Our focus is on the precise characterization of the sensor spectral properties, while a detailed model for sensor radiometric noise is deferred to future work. Note however that the handling of sensor noise can be implicitly included in recent SIF retrieval schemes (cf. Buffat et al. (2025a)).

The DLR Earth Sensing Imaging Spectrometer (DESI) is a space-based instrument installed and operated on the International Space Station (ISS) (Krutz et al., 2019). DESIS was launched to the ISS in June 2018, began initial operation in November 2018 and was declared fully operational in September 2019. DESIS image products consist of 235 spectral bands ranging from 400 nm to 1000 nm and 1024x1024 spatial pixels with a ground sampling distance of 30 m. The overall absolute radiometric calibration is within ~5% at the top-of-atmosphere radiance level outside strong atmospheric absorption regions when validated against RadCalNet (Bouvet et al., 2019), Sentinel-2 and Landsat-8. The spectral calibration is better than ~0.5 nm for the majority of DESIS observations. For a small fraction of observations, though, DESIS shows larger spectral deviations that are within 0.85 nm (~1/3 of a spectral pixel). The spectral sampling distance is nominally stated as 2.55 nm, while the full width at half maximum of the bands is nominally 3.55 nm. A complete description of the DESIS data products and DESIS performance can be found in Alonso et al. (2019).

Several DESIS acquisitions were analysed in our effort to generate realistic simulations. In particular, we used a DESIS scene of 30 × 30 km² southwest of Cologne (centre-west Germany with the geographical WGS84 centre coordinates of latitude 50.7744° and longitude 6.8593°) to gauge differences between simulation and real data. The scene was acquired on June 23, 2020 at 10:10 UTC with the following geometry parameters: sun zenith angle of 32.05°, sun azimuth angle of 140.57°, scene incidence angle of 21.79°, scene azimuth angle of 205.60° and a geometrical accuracy of ~20 m linear deviation in east and north directions. In addition, we studied multiple DESIS scenes acquired between 2018 and 2021 over two sites in Germany (Jülich, Munich) and one in Australia (Litchfield) in order to set realistic DESIS observation geometries in our simulations. More details are given in Section 2.3.

The DESIS spectral response functions were measured on-ground during the instrument characterization campaigns. The central wavelengths can be refined post-launch using internal coloured LEDs, but

the full widths at half maximum are not updated during the whole mission lifetime. Our simulation uses the averaged band CW and FWHM around the O₂-A feature for a particular spectral calibration table (corresponding to the DESIS acquisitions described above) together with the existing dependency in sensor across-track direction. The differences in CW and FWHM with respect to the band averages in the central pixels are shown in Fig. 5 for the wavelengths of interest and the mentioned spectral calibration table. This is sufficient to account for smile (i.e., variation of band central wavelength in the across-track direction) during a limited time interval. However, based on numerous DESIS calibration and Earth datatakes acquired over several years, it is known that under certain conditions a global spectral shift of up to about 0.5 nm can arise (Carmona et al., 2021). Thus, we fix the range of δ_{CW} to [−1.75, +1.25] nm for our simulated dataset. The range of δ_{FWHM} is likewise fixed to [−0.30, +0.30] nm to account for possible inaccuracies in the knowledge of the spectral bandwidth. The precise characterization of the DESIS spectral response function presented here ensures that our simulation is realistic and capable of reproducing DESIS spectra in a wide variety of conditions.

The airborne imaging spectrometer HyPlant is a hyperspectral instrument that was developed within a cooperation between the Finish company SPECIM (SPECIM, 2024) and Forschungszentrum Jülich. HyPlant is composed of two modules, called DUAL and FLUO, which are co-aligned on a rigid baseplate and operated in a synchronized manner. The DUAL module measures radiance from 400 nm to 2500 nm with a spectral resolution of 3–4 nm in the visible and near-infrared spectral region (400–1050 nm) and with a spectral resolution of 10 nm in the short wave spectral region (1050–2500 nm). The FLUO module was designed to enable SIF retrieval and measures radiance in high spectral resolution at the O₂-A and O₂-B oxygen absorption bands. The sensor is fully temperature controlled to ensure stable operation under flight conditions and operates at an effective spectral resolution between 0.23 nm and 0.26 nm (FWHM under flight conditions, see Siegmann et al. (2019) for a detailed description of the radiometric performance of this sensor). The spatial resolution of the HyPlant system depends on the flight altitude and ranges between 0.5 m and 5 m per pixel for the nominal flight altitudes, which were used in the past years. Top-of-canopy reflectance and SIF emission are calculated by applying a dedicated atmospheric correction and physically based retrieval algorithms, which were developed in preparation of the FLEX

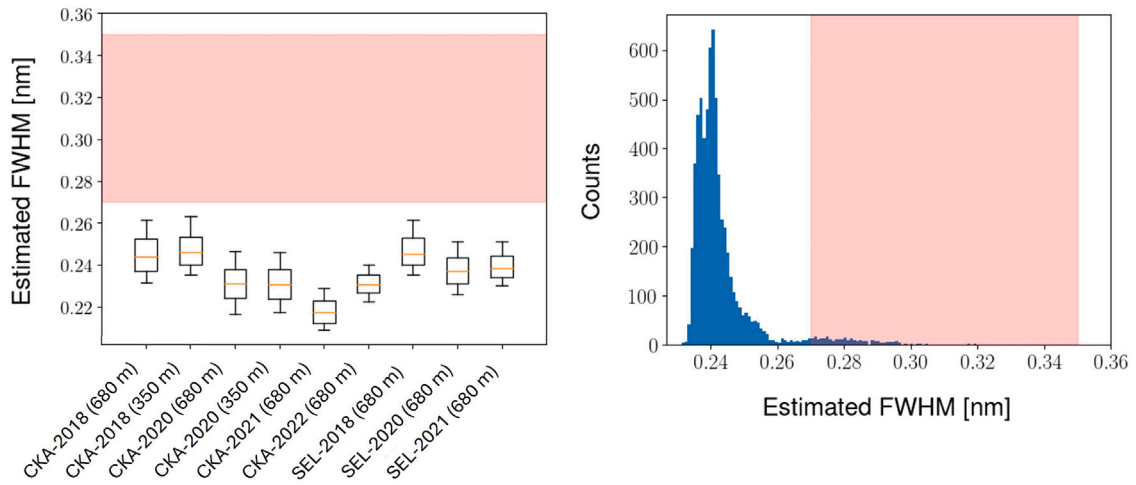


Fig. 6. Re-evaluation of HyPlant spectral resolution around O_2 -A absorption band. The left panel shows the 10%, 25%, 75% and 90% percentiles of the distribution of FWHM values estimated with HyPlant measurements in the range 763–767 nm from campaigns over Campus Klein-Altendorf (CKA) and Selhausen (SEL) between 2018 and 2022 and acquired at 350 m and 680 m flight altitudes. The right panel shows the distribution of fitted FWHM values for the range 740–780 nm derived from two acquisitions with gas emission lines performed during the HyPlant spectral calibration campaign in 2022. In both panels, the range of FWHM based on HyPlant calibration data from 2018 and 2019 is marked by the red filling.

satellite mission (Siegmann et al., 2019; Cogliati et al., 2019; Sabater et al., 2021).

The ready access to recent HyPlant acquisitions (Buffat et al., 2024b) has greatly helped in improving our simulation setup. Throughout our work we used data from multiple HyPlant campaigns in the period 2018–2022 over two sites in Germany (Selhausen and Campus Klein Altendorf) at different flight altitudes (350 m and 680 m above ground level). For each campaign there are several flight lines available, each covering 384 pixels in across-track direction and typical lengths of 3–12 km. In total, several million spectra are collected per flight line with spatial resolutions of 0.5 m and 1.0 m per pixel for flight altitudes of 350 m and 680 m, respectively. We have mostly used the at-sensor radiance spectra measured by the HyPlant FLUO sensor, but in some specific cases (cf. Section 2.2.1) the reflectance derived from the DUAL sensor was also analysed.

The spectral response functions of the HyPlant FLUO sensor are measured in terms of both CW and FWHM during yearly routine calibrations under laboratory conditions by the manufacturer SPECIM. Calibration data from 2018 and 2019 suggest that CW differences with respect to the band average lie in the range $[-0.012, 0.012]$ nm and that the FWHM varies in the range $[0.27, 0.35]$ nm for the wavelengths around the O_2 -A absorption band. Since the spectral response function can be very sensitive to thermal, atmospheric pressure and geometrical instabilities under typical in-flight conditions, we make use of real HyPlant data to inform our choice for the CW and FWHM ranges. We compare a representative set of HyPlant spectra from campaigns over Selhausen and Campus Klein-Altendorf in 2018–2022 at flight altitudes of 350 m and 680 m to dedicated simulations with variable CW and FWHM. The resulting CW shifts vary essentially within ± 0.05 nm, but in some cases shifts can reach slightly higher values, so we fix the range of δ_{CW} to $[-0.080, 0.080]$ nm. The estimated FWHM values are shown in Fig. 6 (left) and most lie between 0.20 nm and 0.27 nm. This result is further cross-checked with the help of two acquisitions with gas emission lines performed during the HyPlant spectral calibration in 2022. The fitted FWHM values, reported in Fig. 6 (right), lie between 0.23 nm and 0.28 nm. The results from imaging and calibration acquisitions are consistent with each other and overall they call for a HyPlant FWHM range $[0.20, 0.28]$ nm, or equivalently δ_{FWHM} in the range $[-0.04, 0.04]$ nm for a baseline FWHM of 0.24 nm.

2.3. Simulated datasets

The simulation framework presented in Section 2.1 is completely general and can be used for virtually any spectral range and application

with airborne or satellite optical instruments. In this work, we restrict ourselves to SIF retrieval in the O_2 -A band with the DESIS and HyPlant instruments (cf. Section 2.2.2).

The first step in generating synthetic data is to single out the ranges of atmosphere, geometry, surface and sensor parameters relevant for the particular application at hand. This is done with the help of the sensitivity analysis documented in Appendix along with expert knowledge of the instruments under consideration. Several results from the sensitivity analysis are worth mentioning here. First, the simulation of the percent-level effect of SIF in the at-sensor radiance spectra requires the use of the line-by-line radiative transfer algorithm with the highest resolution available in MODTRAN6 (0.1 cm^{-1}) and a precise multiple scattering treatment. Second, while aerosol optical thickness and aerosol model are the most relevant atmosphere parameters as expected, also water vapor content plays a role at the extremes of the spectral range considered. Third, the modelling of solar zenith angle, surface and sensor heights as well as reflectance spectrum is indispensable for a precise simulation including SIF. Finally, our sensitivity analysis also shows that the at-sensor radiance signal is particularly sensitive to shifts of central wavelength across the sensor, so an accurate retrieval method needs to implicitly or explicitly account for such shifts when estimating SIF.

Table 1 outlines the key parameters and their ranges for the generation of DESIS and HyPlant simulated spectra. For comparison, a column also shows the values used in the sensitivity analysis. The atmosphere and geometry parameters were defined based on existing DESIS and HyPlant acquisitions over selected locations. For DESIS, 20 acquisitions over Munich, Germany and 16 over Litchfield, Australia in the period 2018–2021 were used to encompass typical observation geometries, while atmospheric conditions for both mid-latitude and tropical sites were gauged considering three sites: Munich and Jülich, Germany and Litchfield, Australia. There are separate sub-databases for DESIS corresponding to mid-latitude summer and tropical atmosphere models. For HyPlant, the focus is on mid-latitude European sites and accordingly the typical atmosphere and geometry parameters were set based on 13 available acquisitions from three campaigns over Selhausen in 2018 at nominal flight altitude of 680 m. In addition, the range of aerosol optical thickness at 550 nm for both DESIS and HyPlant databases was fixed to $[0.02, 0.30]$ based on both global distributions (Kinne, 2019) and on measurements of the aerosol optical thickness (at 675 nm, 870 nm and 1020 nm) made by the CIMEL instrument at the AERONET station in FZJ-JOYCE (AERONET FZJ-JOYCE, 2024). The adopted values for

Table 1

Ranges of input atmosphere, geometry, surface and sensor parameters used for the generation of DESIS and HyPlant simulated data as well as for the sensitivity analysis (cf. Appendix). Atmosphere parameters include atmosphere model, water vapor (H_2O), ozone content (O_3), aerosol optical thickness at 550 nm (AOT_{550}) and aerosol model, while the geometry parameters are tilt angle (TA), sun zenith angle (SZA), relative azimuth angle (RAA) between sun and line of sight, ground altitude h_{gnd} and sensor altitude h_{sen} . The surface is modelled by the reflectance at 740 nm, reflectance slope at 740 nm, ratio of reflectance slopes at 780 nm and 740 nm and SIF on-ground radiance at 737 nm, while the sensor parameters include the shift of central wavelength δ_{CW} and of full width at half maximum δ_{FWHM} . The sensor altitude for HyPlant databases refers to altitude in km above ground level (agl). The radiance unit used to specify the fluorescence output F_{737} is $F_0 = 1 \text{ mW/cm}^2/\text{sr}/\mu\text{m}$.

Parameter		Sensitivity analysis	DESI DB	HyPlant DB
Atmosphere	model	mid-latitude summer, tropical	mid-latitude summer, tropical	mid-latitude summer
	H_2O [cm]	0.3–5.0	0.3–5.0	0.3–3.0
	O_3 [DU]	200–500	332	332
	AOT_{550} []	0.05–0.50	0.02–0.30	0.02–0.30
	aerosol model	rural, maritime, desert, urban, none	rural	rural
Geometry	TA [°]	0–15	0–25	0–20
	SZA [°]	0–45	0–55	20–55
	RAA [°]	0–180	0–180	0–180
	h_{gnd} [m]	0–4000	0–600	0–300
	h_{sen} [km]	0.001–100	100	0.659–0.691 agl
Surface	ρ_{740} []	0.00–0.60	0.05–0.60	0.05–0.60
	s [nm ⁻¹]	0–0.0008	0–0.012	0–0.012
	e []	1	0–1	0–1
	F_{737}/F_0	0–0.8	0–0.8	0–0.8
Sensor	δ_{CW} [nm]	[–1.00, +1.00]	[–1.75, +1.25]	[–0.080, +0.080]
	δ_{FWHM} [nm]	[–0.25, +0.15]	[–0.3, +0.3]	[–0.040, +0.040]

surface and sensor properties were based on the discussion in Sections 2.2.1 and 2.2.2, respectively. It is important to stress that the ranges in Table 1 are representative of real DESIS and HyPlant data to be used in the future in conjunction with the simulated datasets presented here. This point is crucial for the success of any SIF retrieval method trained on simulated data.

The next step is to sample the input parameter space defined above. Three sampling methods are considered: uniform grid, random sampling and Halton sampling with Owen scrambling (Halton, 1960; Owen, 2017). The multi-dimensionality of the input space (> 10 dimensions) requires millions of samples in order to achieve a reasonable sampling uniformity. This is currently beyond our available computational resources because the time required to simulate one at-sensor radiance spectrum is several minutes, dominated by MODTRAN running time. However, as seen in Section 2.1, the simulation can be decoupled into a (slow) atmospheric module, that depends only on atmosphere and geometry parameters, and a (fast) surface-sensor module, that depends only on surface and sensor properties, see Fig. 1. Instead of running both steps in sequence for each sample of the full input space, one may first execute the atmospheric module in a sample set of the atmosphere/geometry input space and then, for each run, execute the surface-sensor module in a sample set of the surface/sensor input space. This strategy represents a reasonable compromise between sampling uniformity and running time, making it feasible to simulate a reasonably uniform body of millions of high-dimensional samples in useful time with the available resources.

The sampling specifications chosen for the DESIS and HyPlant databases are reported in Table 2. In addition to random and Halton samples, a set of uniform grid sampling points is also included in order to populate the databases with the borders and corners of the input parameter space as well as special values (e.g., zero sensor shifts), which are not necessarily picked up by random or Halton sampling. The use of the three sampling methods for the generation of the databases will be useful in the future to study the effect of sampling strategies in the training of a ML-based SIF retrieval method. In total,

Table 2

Specification of the generated DESIS and HyPlant databases. The input space and number of samples are reported for each database employing uniform grid (UG), random (R) and Halton (H) sampling techniques.

DB	Specification	DESI DB			HyPlant DB		
ATM	Input space	6d			7d		
	Sampling	UG	R	H	UG	R	H
	Nr. samples	2×3^6	2×10^3	2×10^4	3^7	10^3	10^4
SENSOR	Input space	6d			6d		
	Sampling	UG	R	H	UG	R	H
	Nr. samples	2×3^6	300	300	2×3^6	300	300
Total samples		2×10^6	6×10^5	6×10^6	3×10^6	3×10^5	3×10^6

approximately 15 million spectra were simulated distributed across the different databases and sampling strategies. Several more sets with tens of million of spectra overall were simulated in intermediate datasets leading up to the final databases presented here.

The databases of simulated spectra were generated with a dedicated virtual machine using 10 cores. The total running time for the final datasets amounted to approximately one week. The single-core running time for the atmospheric module varies between 1 min/sample for nadir geometries and about 4 min/sample for off-nadir geometries, while the same time needed for the surface-sensor module amounts to approximately 0.02 s/sample for both DESIS and HyPlant.

We conclude this section by showing in Fig. 7 selected at-sensor radiance spectra from the generated databases for both DESIS and HyPlant. The different resolutions of DESIS and HyPlant can be fully appreciated as well as the complex interplay between atmosphere, geometry, surface and sensor properties. Although limited, the set of spectra in Fig. 7 serves to illustrate the power of an extensive body of simulated data that can be used to analyse the signal measured by hyperspectral sensors and help in the SIF retrieval task.

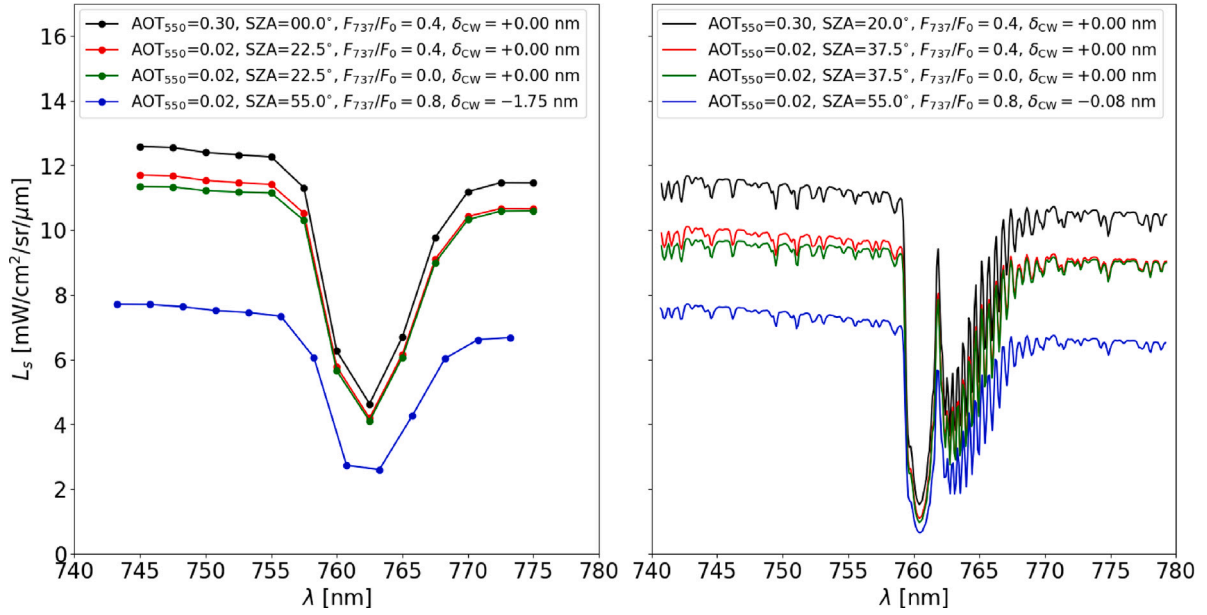


Fig. 7. Selected simulated at-sensor radiance spectra surrounding the O₂-A feature (740–780 nm) for different atmosphere, geometry, surface and sensor properties. The spectra are extracted from DESIS DB (uniform grid sampling, mid-latitude atmosphere model) in the left panel and from HyPlant DB (uniform grid sampling) in the right panel. All cases correspond to a constant 30% surface reflectance. The number of spectral bands in the wavelength range considered is 13 for DESIS and 349 for HyPlant.

3. Results

In this section we use the simulated dataset described above for two specific applications. The analysis presented here demonstrates explicitly the usefulness of our simulations for SIF retrieval and motivates further applications to be pursued in the near future.

3.1. At-sensor radiance emulators

The assembled set of simulated data effectively establishes a direct connection between a combination of atmosphere, geometry, surface and sensor properties and the corresponding at-sensor radiance spectrum in the wavelength range of interest for DESIS and HyPlant (cf. Eq. (1)). While the dataset is originally intended to aid in the ill-posed inverse task of retrieving one surface property (in our case, SIF) from the at-sensor radiance spectrum, the same dataset can also be used for the direct task of predicting the at-sensor radiance spectrum given a set of atmosphere, geometry, surface and sensor properties:

$$L_s = F(x_{\text{atm}}, x_{\text{geo}}, x_{\text{sur}}, x_{\text{sen}}) \quad (9)$$

where $F: \mathbb{R}^p \rightarrow \mathbb{R}^b$ is a multi-dimensional function, p is the number of atmosphere, geometry, surface and sensor parameters ($p = 12$ for DESIS, $p = 13$ for HyPlant, cf. Table 1) and b is the number of spectral bands in the considered range around the O₂-A band ($b = 13$ for DESIS, $b = 349$ for HyPlant). The learning of the function F with the help of the simulated dataset is a well-defined regression problem and provides an at-sensor radiance emulator which mimics the simulation step. Several works in the literature have proposed and studied emulators in different contexts and applications (see e.g. Vicent Servera et al. (2022)). For our regression problem, different parametric functions for F may be chosen; a comprehensive study of different emulators and their performance is presented by us in Pato et al. (2024). Fourth degree polynomials turn out to be fast and accurate emulators for our purposes in this work. We therefore train a fourth degree polynomial emulator separately for DESIS and HyPlant using the uniform grid and Halton samples of simulated spectra (cf. Table 2) as training set, while leaving the random datasets for validation and testing.

The time and accuracy performance of the trained emulators for DESIS and HyPlant is shown in Fig. 8. Two main conclusions are

evident from our results. First, as seen in the left panel of Fig. 8, the emulators have prediction times per sample spectrum ranging from 1–2 ms (for a single sample) down to 10–20 μ s (for a million samples), with the DESIS emulator slightly faster due its lower dimensionality. The prediction is faster (per sample) when done in bulk since it involves only matrix multiplications, which are highly optimized in standard software packages. The single-sample time of a few ms is relevant for the sequential emulation of spectra (as needed in the simulation step of some SIF retrieval algorithms (Buffat et al., 2023, 2025c)), while the million-sample time of 10–20 μ s is to be considered for emulation in bulk (as in the generation of large datasets for training SIF retrieval algorithms). Either way, the emulator is several orders of magnitude faster than the 1–4 min needed for the full physics-based simulation of one spectrum.

Second, the emulators are very accurate in reproducing the simulated spectra both for DESIS and HyPlant, as shown by the distributions of the band-averaged relative errors in the right panel of Fig. 8. The errors cluster around approximately 0.1% for DESIS and 0.02% for HyPlant and there are few cases where they exceed 1%. Such impressive sub-percent accuracy is adequate to characterize the SIF signal, which usually amounts to a few percent of the total at-sensor radiance. In fact, we employed the emulators to reproduce the whole DESIS DB (mid-latitude atmosphere model) and HyPlant DB presented in Section 2.3 with $\sim 0.1\%$ average error or better in just under 3 min. This highlights the power of emulation and its excellent combined time and accuracy performance.

Finally, we show that the emulators accurately reproduce real measurements, thereby providing the first step for a robust SIF retrieval and the estimation of its uncertainty. The comparison between simulated and measured spectra over a representative land cover is reported in Fig. 9 for a DESIS scene and in Fig. 10 for a HyPlant scene. We employ the emulators for DESIS and HyPlant to fit every spectrum in the scenes. An unconstrained least squares optimization is performed for each individual spectrum to retrieve the atmosphere, geometry and surface parameters and column-wise per band to retrieve the sensor shifts. The deliberately simplified scheme is not intended for SIF retrieval, but it is adequate to interpret DESIS and HyPlant spectra. The results of this procedure for the example pixels highlighted in Figs. 9(a) and 10(a) are shown in Fig. 9(b) and Fig. 10(b), respectively. The distribution of

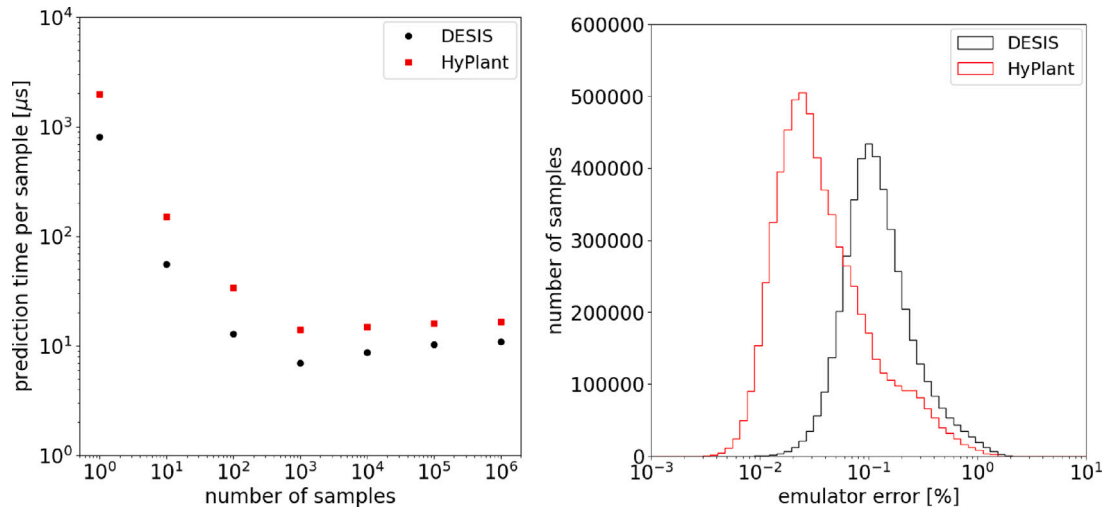


Fig. 8. Time and accuracy performance of fourth degree polynomial emulators of at-sensor radiances for DESIS and HyPlant around the O₂-A band. The left panel displays typical prediction times per sample for different number of samples, while the right panel shows the distributions of the band-averaged relative errors in the entire DESIS DB (mid-latitude atmosphere model) and HyPlant DB.

the residuals between data and simulation for the full patches and all bands, reported in Figs. 9(c) and 10(c), indicates that typical residuals lie in the interval $9 \times 10^{-4} \pm 0.07$ mW/cm²/sr/μm for DESIS and $5 \times 10^{-3} \pm 0.09$ mW/cm²/sr/μm for HyPlant.

We shortly comment on the magnitude of the residuals between simulation and data. While it is true that some residuals are larger than about 0.1 mW/cm²/sr/μm, we do not see this as problematic for two main reasons. First, the distribution of residuals is shown for the full patch, including surface compositions not represented in our simulation. In order to be conservative we have not filtered these regions out for Figs. 9(c) and 10(c). Second, our simulation is noise-free while the measured spectra naturally contain noise that ends up in the quantification of the residual. Notice in particular that, for a sensor such as DESIS, a radiance noise of 0.1 mW/cm²/sr/μm is very small. Overall, the key message of the residuals shown in Figs. 9 and 10 is that their distribution is not biased and that our simulation setup reproduces DESIS and HyPlant at-sensor radiance spectra to within better than 0.1 mW/cm²/sr/μm across the considered wavelength range. These findings demonstrate the realistic nature of the simulations and show that our body of simulated data constitutes a solid basis for training ML-based SIF retrieval methods.

3.2. SIF retrieval performance

The simulated dataset presented in our work can be used in a straightforward fashion to test the performance of any SIF retrieval method for various atmosphere, geometry, surface and sensor conditions. Such studies may also inform the actual design of retrieval algorithms before application to real data. We demonstrate this here, making use of a small set of simulated DESIS spectra to evaluate the SIF retrieval performance of different variants of the 3FLD method. This method, initially proposed in Maier et al. (2003), estimates a relative SIF signal by using a band or combination of bands in the O₂-A absorption region (on-band) and a virtual band consisting of two or more bands on either side of the absorption region (off-band). The procedure may be applied to each sensor element separately and relies on the existence of non-fluorescent pixels in the image to factor out the influence of atmosphere and geometry. In order to evaluate the SIF retrieval performance of 3FLD we collected small subsets of 2000 DESIS and HyPlant simulated spectra with fixed atmosphere, geometry and sensor properties and varied surface reflectance and fluorescence. In such a dataset the SIF retrieval accuracy can be assessed directly

by comparing the retrieved fluorescence for each spectra with the simulated value.

The choice of the on- and off-bands in the 3FLD method depends sensitively on the properties of the sensor and is better done guided by simulated data. We start by considering only simulated spectra with linear reflectances ($e = 1$) and comment on non-linear reflectances afterwards. Fig. 11 shows the correlation between the 3FLD retrieved fluorescence and the actual simulated fluorescence for all simulated DESIS spectra using three different on- and off-band configurations. Since the algorithm retrieves only relative SIF, an absolute error estimation is not possible, but the correlations in Fig. 11 can be used to assess the 3FLD accuracy. DESIS has 13 spectral bands in the region around the O₂-A absorption feature (cf. Fig. 7), so different on- and off-band configurations are possible. One reasonable choice is to select band 8 (762.5 nm) sitting deep in the absorption feature as on-band and combine bands 5 and 11 (755 nm and 770 nm) outside the feature to form the off-band. This baseline 3FLD band configuration leads to a good retrieval performance with a Pearson correlation coefficient of 0.68 and a very small bias as evident from the top panel of Fig. 11. The other panels show the effect of using alternative bands on either side of the absorption feature (bands 3–5 and 11–13, cf. middle panel) and additionally inside the absorption feature (bands 7–9, cf. bottom panel). The addition of bands seems to slightly reduce the retrieval accuracy with the correlation coefficient dropping to 0.65 (middle panel) and 0.63 (bottom panel) while the bias is kept negligible in both cases. Notice however that our simulated data does not include sensor noise. When using real data, adding bands in the 3FLD method should reduce the impact of sensor noise and therefore lead to an increase in retrieval quality. Other combinations of on- and off-band configurations lead to poor retrievals (not shown), highlighting that this choice of wavelengths is essential for the success of 3FLD in estimating SIF.

The case of HyPlant is qualitatively different from DESIS given the smaller path between target and sensor and especially its higher spectral resolution. In particular, it is possible with HyPlant to place a narrow on-band for the 3FLD method deep inside the oxygen absorption feature, thereby reducing dramatically the down-welling irradiance and reflectance signal in that band and consequently increasing the relative importance of the fluorescence signal. This facilitates SIF retrieval with a high spectral resolution instrument as HyPlant when compared to a medium spectral resolution spectrometer as DESIS. Our results confirm this well-known expectation. The 3FLD retrieval performance for the HyPlant simulated spectra (with linear reflectance spectra, $e = 1$) is reported in Fig. 12 when using HyPlant band 179 (760.42 nm) as on-band and bands 120 and 238 (753.90 nm and 766.95 nm) as off-band.

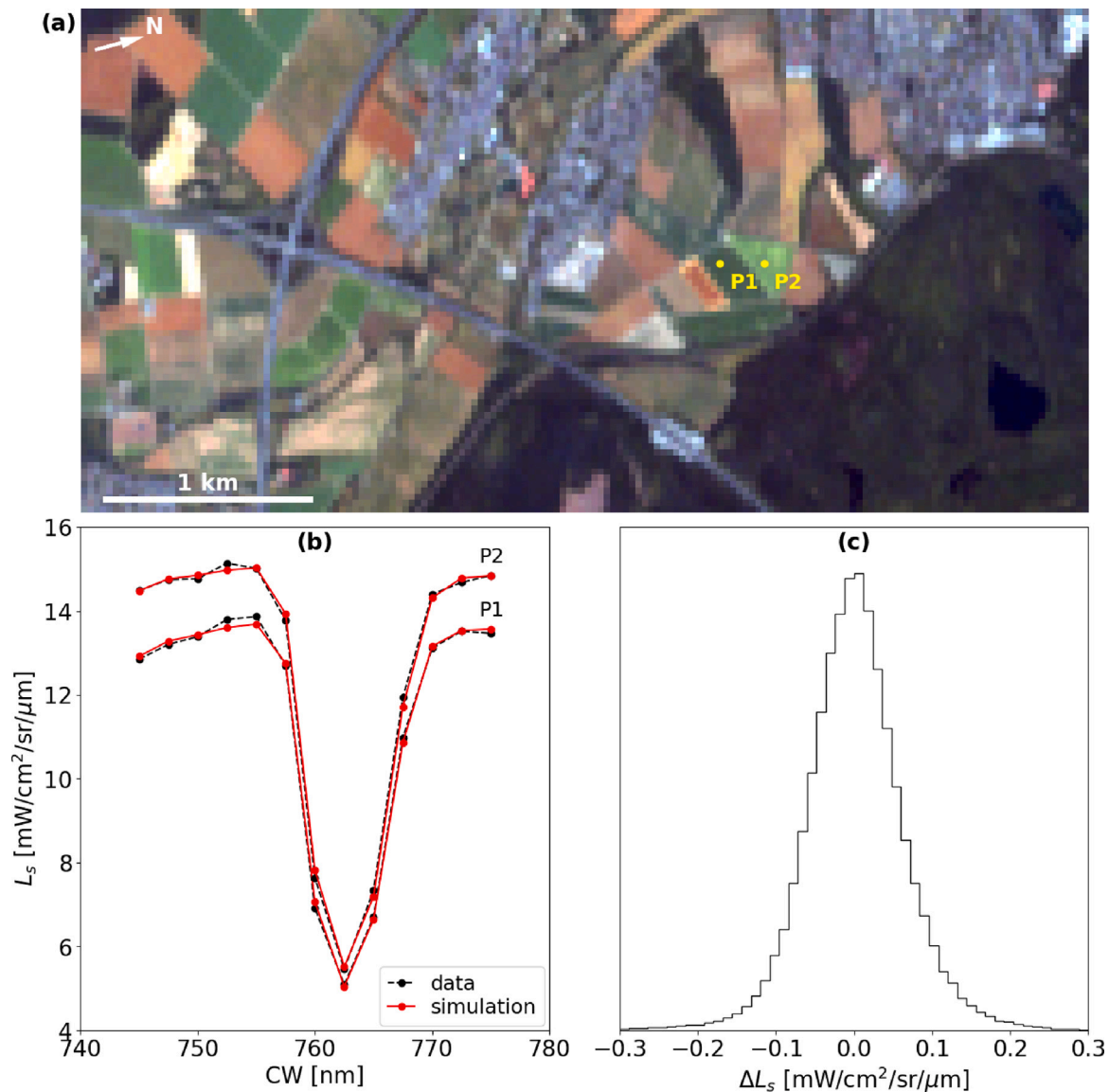


Fig. 9. Comparison between DESIS data and simulation. The top panel (a) displays the true colour composite of a heterogeneous region of approximately $2.4 \times 4.8 \text{ km}^2$ (80×160 pixels) based on a DESIS L1B product (at-sensor radiances) without smile correction from an acquisition southwest of Cologne, Germany on June 23, 2020 at 10:10 UTC. Two pixels with representative vegetation covers are highlighted in yellow and the corresponding comparison between simulated and measured spectra is shown in panel (b). Panel (c) reports the distribution of the residuals between measured and simulated at-sensor radiances across the patch for all bands.

The retrieval accuracy achieved with HyPlant is very good attaining a correlation coefficient of 0.99, clearly superior to the figures obtained for DESIS. In addition, averaging up to three HyPlant bands to form the 3FLD on- and off-bands does not appreciably change the results in Fig. 12, always leading to a correlation coefficient of 0.99. This situation is different from DESIS (cf. Fig. 11) and can be understood by the high spectral resolution and small spectral sampling distance of HyPlant.

Lastly, we comment briefly on one important limitation of the 3FLD method, namely the assumption of a linear reflectance model ($e = 1$). In fact, the SIF retrieval accuracy obtained by the 3FLD approach can be significantly reduced for simulated spectra having non-linear reflectances. For instance, using the DESIS band configuration shown in the bottom panel of Fig. 11 (bands 7–9 as on-band, bands 3–5 and 11–13 as off-band), we find a positive bias growing from 3.3×10^{-4} for $e = 1$ (cf. bottom panel of Fig. 11) to 4.2×10^{-3} for $e > 0.8$ and 9.9×10^{-3} for $e > 0.5$. Accordingly, the correlation coefficient drops from 0.63 for $e = 1$ to 0.45 for $e > 0.8$ and 0.30 for $e > 0.5$. This loss of accuracy using DESIS spectra can be mitigated by restricting the spectral range of the off-band: using only bands 5 and 11 as off-band brings the

bias to 4.2×10^{-3} and the correlation coefficient to 0.53 for $e > 0.5$. For HyPlant, we find instead a very good retrieval accuracy with a correlation coefficient of 0.99 for both linear and non-linear reflectance spectra. This is a direct consequence of the high spectral resolution of HyPlant and the possibility of using a narrow 3FLD on-band deep in the absorption feature, which significantly weakens the impact of the reflectance spectrum on the retrieval. The linear reflectance limitation of 3FLD can be addressed with other SIF retrieval algorithms, but this example serves to illustrate how tailored simulated data can be applied to study in detail the performance of a SIF retrieval method.

4. Discussion

Simulated at-sensor radiance spectra may be used in several ways to help combine physics-based SIF retrieval methods with ML algorithms, as pursued in our recent studies for HyPlant (Buffat et al., 2023, 2025c,b) and DESIS (Buffat et al., 2025a). The present work explored two specific applications of the simulated data.

First, physics-based emulators trained with the databases were shown to provide a very fast and precise alternative to a full radiative

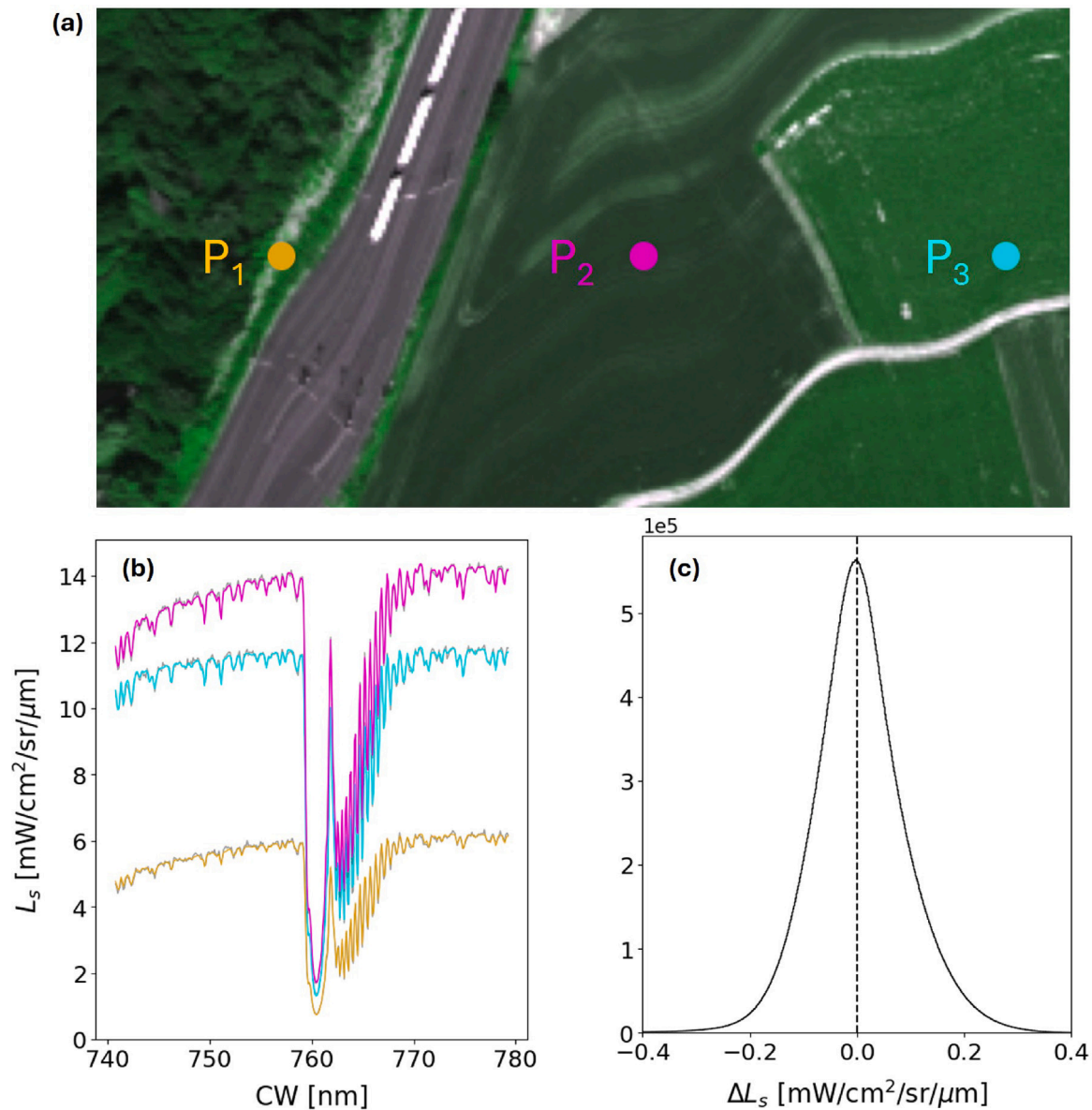


Fig. 10. Comparison between HyPlant data and simulation. Panel (a) displays the true colour composite of a subset of a HyPlant acquisition from Jülich, Germany on June 26, 2018 at 15:30 CEST. Three pixels with representative vegetation covers are highlighted in colour and the corresponding comparison between simulated (coloured) and measured (black) spectra is shown in panel (b). Panel (c) reports the distribution of the residuals between measured and simulated at-sensor radiances across the patch for all bands.

transfer model (cf. Section 3.1). The emulators can complement – and in some cases replace – the time-consuming simulations including radiative transfer. Our findings have at least two important consequences for SIF studies. Most important, the emulators can be straightforwardly incorporated into retrieval methods (see e.g. Buffat et al. (2025a)) to provide a fast online simulation step and thereby improve the reconstruction of the measured signal and the retrieval of fluorescence. But an additional advantage is that extensive tailored datasets of simulated spectra can be emulated at very modest computational cost. These datasets may serve as training sets for developing novel ML-based SIF retrieval methods and as reference data to evaluate the SIF retrieval accuracy of any method. Both points illustrate how our emulators may aid in specific tasks which are simply not feasible with full physics-based simulations. It is worth cautioning that the emulators proposed here are valid only for DESIS and HyPlant in the spectral region around the O₂-A band and for the considered range of input parameters. However, our procedure can be easily applied to derive emulators for other applications.

Second, the databases constitute a testing ground which we exploited to assess the SIF retrieval performance of the 3FLD method (cf. Section 3.2). As explained in detail in Section 1, our preference for 3FLD over state-of-the-art SIF retrieval methods in this assessment is grounded on its simplicity and robustness. The performance study presented is deliberately brief and it serves to illustrate the usefulness of simulated datasets, such as the one introduced in our work for SIF research. The study may certainly be extended in several respects, including for instance the inter-comparison of different retrieval methods and the retrieval sensitivity to atmosphere, geometry and sensor properties. In particular, the SIF retrieval uncertainty may be quantified for different observation conditions. We defer the exploration of these possibilities to future work. The simulated dataset can also be employed as a labelled training dataset for supervised learning methods. Incidentally, we have recently developed a novel SIF retrieval scheme combining physics-based modelling and ML algorithms (Buffat et al., 2025c) and have extended it to incorporate our simulations in a separate work (Buffat et al., 2025a,b).

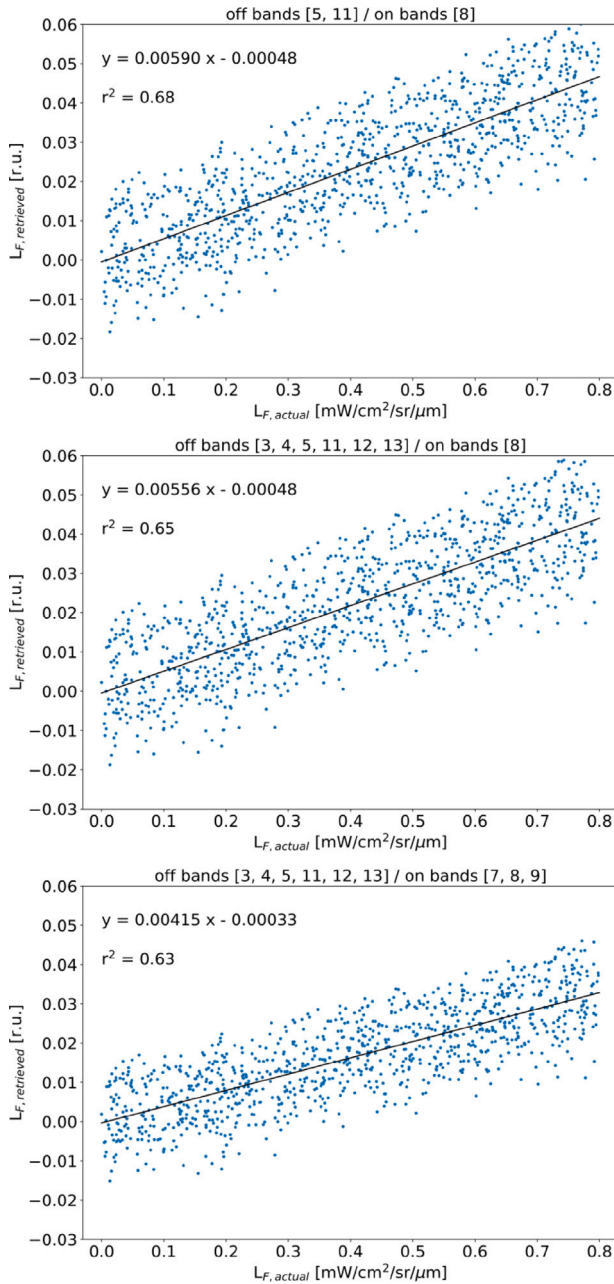


Fig. 11. SIF retrieval performance of the 3FLD method using different on- and off-band configurations evaluated on DESIS simulated data. All panels show the correlation between the relative SIF signal retrieved by 3FLD and the actual simulated SIF on-ground radiance. The upper panel corresponds to the baseline 3FLD band configuration with DESIS band 8 (762.5 nm) used for the on-band and bands 5 and 11 (755 nm and 770 nm) for the off-band. In the middle panel the on-band is formed by DESIS band 8 (762.5 nm) and the off-band by DESIS bands 3–5 (750–755 nm) and 11–13 (770–775 nm), while in the bottom panel the on-band is formed by DESIS bands 7–9 (760–765 nm) and the off-band by DESIS bands 3–5 (750–755 nm) and 11–13 (770–775 nm). The best linear fit and corresponding Pearson correlation coefficient are indicated in each case.

Finally, we comment on the applicability of our work to other sensors and spectral ranges. Here we focused on the SIF retrieval potential of DESIS and HyPlant in the O₂-A band by implementing a precise sensor characterization to simulate realistic spectra, which differentiates our simulated dataset from others in the literature. While our results are necessarily specific to DESIS and HyPlant at 740–780 nm, the proposed framework can be applied to other instruments and wavelength ranges. For instance, one may examine the case of FLEX (Drusch et al., 2017)

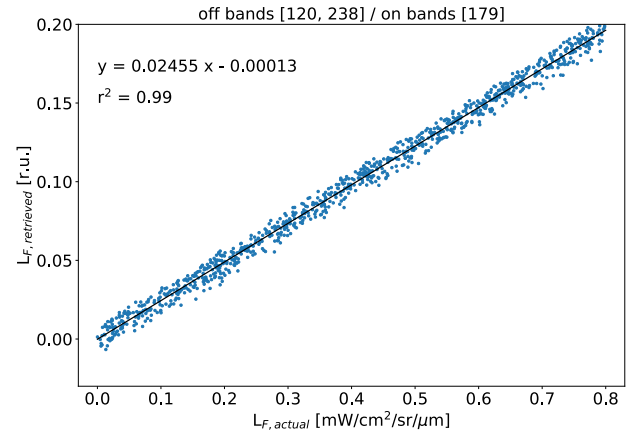


Fig. 12. SIF retrieval performance of the 3FLD method evaluated on HyPlant simulated data. The plot shows the correlation between the relative SIF signal retrieved by 3FLD and the actual simulated SIF on-ground radiance along with the best linear fit and corresponding Pearson correlation coefficient. The 3FLD band configuration used here consisted of HyPlant band 179 (760.42 nm) as on-band and bands 120 and 238 (753.90 nm and 766.95 nm) as off-band.

at the O₂-A band, provided a precise characterization of the instrument is available. The case of the O₂-B band (at around 687 nm) is more challenging for SIF retrieval due its moderate depth and would require the definition of different reflectance and fluorescence models in our framework. Either at the O₂-A band, O₂-B band and/or full spectral range, it is crucial to represent the spectral characteristics of the FLEX instruments as realistically as possible (eventually with the help of on-ground calibration data) in order to generate high fidelity simulated spectra. An extensive dataset of such spectra is essential for the development of ML methods for SIF retrieval with FLEX and it could also be used to train a fast forward simulator and evaluate the SIF retrieval performance of existing methods. This research direction is of particular relevance for the exploitation of FLEX data, but it lies outside the scope of the present contribution and is left for future work.

5. Conclusion

The retrieval of solar-induced fluorescence from airborne and satellite-based hyperspectral measurements is a complex inversion problem that requires an intimate knowledge of the spectral performance of the instrument and a precise correction for atmospheric effects. The combination of physics-based retrieval methods with ML algorithms has the potential to tackle such problems efficiently and thus provide a promising avenue for retrieving SIF from a multitude of hyperspectral sensors of diverse spectral and spatial resolutions. However, any ML model can be only as good as the body of training simulated data available to it. Furthermore, ML models are proficient at learning subtle features in the data and thus may fail to properly interpret real measurements if the simulated training data are not reliable and representative of the instrument. It is therefore crucial that the simulated data used for training are representative of the real data. In this work, we attempted to provide a sound starting point for the learning process of ML-based SIF methods by assembling an extensive set of highly realistic simulated at-sensor radiance spectra around the O₂-A band for two representative classes of currently operating imaging spectrometers, namely DESIS and HyPlant. Our approach leverages on the expert knowledge about the calibration and characterization of DESIS and HyPlant in order to provide high-fidelity simulated spectra that closely reproduce actual measurements. The same approach can be applied to existing or future instruments such as FLEX in the O₂-A band, or other spectral ranges. In addition, such simulated datasets can easily be used to train fast and accurate emulators of at-sensor radiances as

well as to evaluate the performance of SIF retrieval methods. These are two important applications whose results will aid in the ongoing effort towards developing robust ML-based retrieval methods for SIF.

CRedit authorship contribution statement

Miguel Pato: Writing – review & editing, Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Kevin Alonso:** Writing – review & editing, Software, Methodology, Data curation, Conceptualization. **Jim Buffat:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Stefan Auer:** Writing – review & editing, Supervision, Conceptualization. **Emiliano Carmona:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Stefan Maier:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Rupert Müller:** Supervision, Methodology, Conceptualization. **Patrick Rademske:** Conceptualization. **Uwe Rascher:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Hanno Scharr:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Miguel Pato, Kevin Alonso, Jim Buffat, Stefan Auer, Emiliano Carmona and Rupert Mueller reports financial support was provided by Helmholtz Artificial Intelligence Cooperation Unit. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Rudolf Richter for discussions on atmospheric correction, Neus Sabater for insights on global distributions of aerosol parameters and the MODTRAN development team for support in the use of MODTRAN6. This work was supported by the project “FluoMap” (Impulsfonds-Förderkennzeichen ZT-I-PF-5-12) funded by the Helmholtz Initiative and Networking Fund, Germany, Helmholtz AI, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Germany and Forschungszentrum Jülich GmbH (FZJ), Germany. The authors gratefully acknowledge computing time on the supercomputer JURECA (Jülich Supercomputing Centre, 2021) at Forschungszentrum Jülich under grant no. fluomap-ct. We gratefully acknowledge the financial support by the European Space Agency (ESA), France for airborne data acquisition and data analysis in the frame of the FLEXSense campaign (ESA Contract No. 4000125402/ 18/NL/NA) and the Photoproxy project (ESA contract No. 4000125731/19/NL/LF). Additionally, HyPlant data acquisition has partially been funded by the German Federal Ministry of Education and Research, Germany within the German-Plant-Phenotyping Network (DPPN) (project identification number: 031A053), the ‘Strukturwandel-Projekt Bioökonomie REVER’, which is funded by the German Federal Ministry of Education and Research, Germany (project identification number 031B0918A), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Germany under Germany’s Excellence Strategy – EXC 2070–390732324 and the project Land surface Interactions with the Atmosphere over the Iberian Semi-arid Environment (LIAISE) funded by the Centre national de la recherche scientifique (CNRS), France.

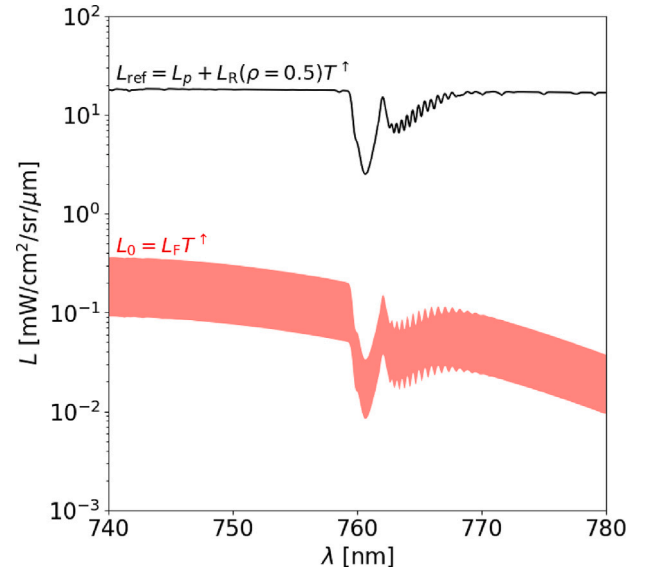


Fig. A.1. Sensitivity goal of the simulation of at-sensor radiances for SIF retrieval in the O₂-A band. The simulation should be at least as accurate as the effect we aim to study, in our case SIF. The sensitivity goal L_0 shown in red is set by the SIF at-sensor radiances corresponding to the typical SIF outputs $F_{737} = (0.1 - 0.4) \text{ mW/cm}^2/\text{sr}/\mu\text{m}$, while the reference at-sensor radiance L_{ref} for a reflectance $\rho = 0.5$ is reported in black. The curves were resampled to FWHM of 0.3 nm and correspond to the reference configuration (nadir observation from space, sea-level surface, sun zenith angle of 30°, mid-latitude summer atmosphere, rural aerosol profile, 23 km visibility, accurate radiative transfer baseline A1).

Appendix. Sensitivity analysis

The SIF on-ground radiance spectrum has two characteristic peaks at approximately 690 nm and 737 nm, so the narrow spectral range we consider (740–780 nm) essentially comprises the falling tail of the feature at 737 nm. Typical on-ground radiances for this second fluorescence peak amount to $F_{737} = (0.1 - 0.4) \text{ mW/cm}^2/\text{sr}/\mu\text{m}$ (Meroni et al., 2009; Cogliati et al., 2015a; Rascher et al., 2015). The corresponding range of SIF at-sensor radiances around the O₂-A band is illustrated in Fig. A.1 by the red band, which effectively sets the sensitivity goal L_0 for our study. In other words, our simulations are designed to be accurate enough to represent SIF signals down to $F_{737} = 0.1 \text{ mW/cm}^2/\text{sr}/\mu\text{m}$, but not below that value. Also shown in the figure is a reference at-sensor radiance spectrum L_{ref} for a surface reflectance $\rho = 0.5$ and no fluorescence emission.

It is clear from Fig. A.1 that SIF emission by vegetation leads to a percent-level effect in the measured at-sensor radiance, which is largely dominated by the reflectance signal. Any simulation of use for SIF retrieval must be at least as accurate as the order of magnitude of the fluorescence signal itself, which in our case implies a percent-level accuracy. For our purposes, the relevant accuracy measure is the difference to the reference at-sensor radiance: $\Delta L = L - L_{\text{ref}}$. We therefore carried out an extensive sensitivity analysis to identify all parameters or configurations with an impact on the at-sensor radiance spectrum larger than the SIF sensitivity goal, namely $|\Delta L| > L_0$. The exact sensitivity of the retrieval to these parameters depends on the SIF retrieval algorithm.

The reference configuration used for our sensitivity analysis corresponds to a nadir-looking sensor in space, sea-level surface, sun zenith angle of 30°, mid-latitude summer atmosphere, rural aerosol profile, 23 km visibility (corresponding to an aerosol optical thickness

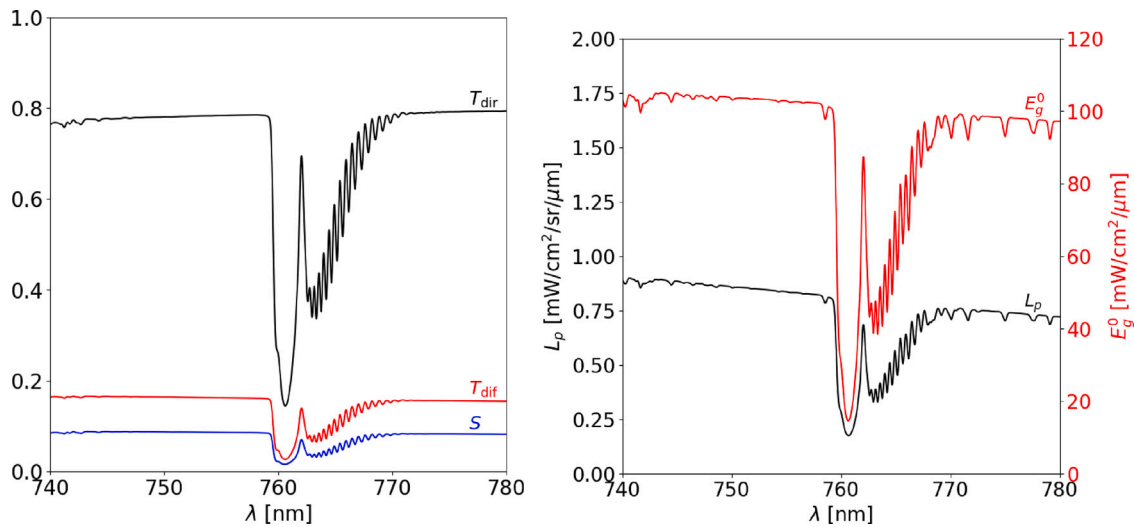


Fig. A.2. Atmospheric functions derived for the reference configuration. The different functions shown quantify the transmission and scattering of light through the atmosphere and their dependence on wavelength. The left panel shows the transmission coefficients $T_{\text{dir}}^{\uparrow}$, $T_{\text{dif}}^{\uparrow}$ and the spherical albedo of the atmosphere S , while path radiance L_p and global solar irradiance on the ground E_g^0 are displayed in the right panel. The high-resolution atmospheric functions were here resampled to FWHM of 0.3 nm for visualization purposes.

Table A.1

Radiative transfer configurations used in the sensitivity analysis to balance accuracy and computational time. The last column reports the single-core run time in minutes and seconds needed for the atmospheric module to generate high-resolution atmospheric functions for the reference configuration. Each case is colour-coded in green, orange or red depending on whether its difference with respect to the accurate baseline A1 is below, similar to or above the SIF sensitivity goal, respectively. Case C4 was selected as default for the generation of databases and is marked with an asterisk.

Case	Model	Resolution	Multiple scattering	Run time
00	correlated-k (fast)	1.0 cm^{-1} / –	Isaacs scaled (8S)	00:02
A1	line-by-line	0.1 cm^{-1} / 100	DISORT (8S)	09:13
B2	correlated-k (slow)	0.1 cm^{-1} / –	DISORT (8S)	01:40
B3	correlated-k (fast)	0.1 cm^{-1} / –	DISORT (8S)	00:59
B4	band model	0.1 cm^{-1} / –	DISORT (8S)	00:09
C2	line-by-line	0.1 cm^{-1} / 50	DISORT (8S)	04:34
C3	line-by-line	0.1 cm^{-1} / 20	DISORT (8S)	01:55
C4*	line-by-line	0.1 cm^{-1} / 10	DISORT (8S)	00:57
C5	line-by-line	0.1 cm^{-1} / 5	DISORT (8S)	00:30
C6	line-by-line	0.1 cm^{-1} / 3	DISORT (8S)	00:20
C7	correlated-k (slow)	1.0 cm^{-1} / –	DISORT (8S)	00:20
C8	correlated-k (slow)	5.0 cm^{-1} / –	DISORT (8S)	00:05
D2	line-by-line	0.1 cm^{-1} / 100	Isaacs scaled (8S)	failed
D3	correlated-k (slow)	0.1 cm^{-1} / –	Isaacs scaled (8S)	00:07
D4	correlated-k (fast)	0.1 cm^{-1} / –	Isaacs scaled (8S)	00:06
D5	line-by-line	0.1 cm^{-1} / 100	None	00:05

at 550 nm of 0.30–0.35), the Fontenla et al. 2011 (Fontenla et al., 2011) (mid2) solar model and the accurate radiative transfer baseline A1 (cf. Appendix A.1). The associated atmospheric functions are reported in Fig. A.2. This reference configuration has been used throughout. The sensitivity analysis presented here is performed at the spectral resolution of 0.3 nm to support DESIS (FWHM of 3.5 nm), HyPlant (0.3 nm) and any other instrument of larger FWHM. In the following we focus on the sensitivity to radiative transfer configurations, atmosphere, geometry, surface and sensor parameters.

A.1. Radiative transfer configuration

The requirement of a realistic simulation at high resolution in the spectral range of the O_2 -A band implies a careful choice of the radiative transfer options in MODTRAN6 (Berk et al., 2014). Our strategy was to first set a very accurate baseline configuration and then vary in turn

different radiative transfer options in order to find the best compromise between simulation accuracy and computational time. As accurate baseline, the line-by-line algorithm implemented in MODTRAN6 (Berk et al., 2015) was used with 0.1 cm^{-1} wavenumber resolution,¹ 100 sampling points per spectral bin and the 8-stream DISORT multiple scattering treatment. This baseline is designated as case A1 in Table A.1 and it takes our atmospheric module (Section 2.1.1) about 9 min to run MODTRAN6 twice and derive the five high-resolution atmospheric functions in the wavelength range between 740 nm and 780 nm for the reference configuration (cf. Fig. A.2). We sequentially tested alternative radiative transfer models (correlated-k (slow), correlated-k (fast), band model, see cases B2–B4 in Table A.1), resolutions (1 cm^{-1} ,

¹ For reference, a wavenumber resolution of $\Delta\nu = 0.1 \text{ cm}^{-1}$ corresponds to a wavelength resolution of $\Delta\lambda \approx 0.006 \text{ nm}$ at $\lambda = 760 \text{ nm}$.

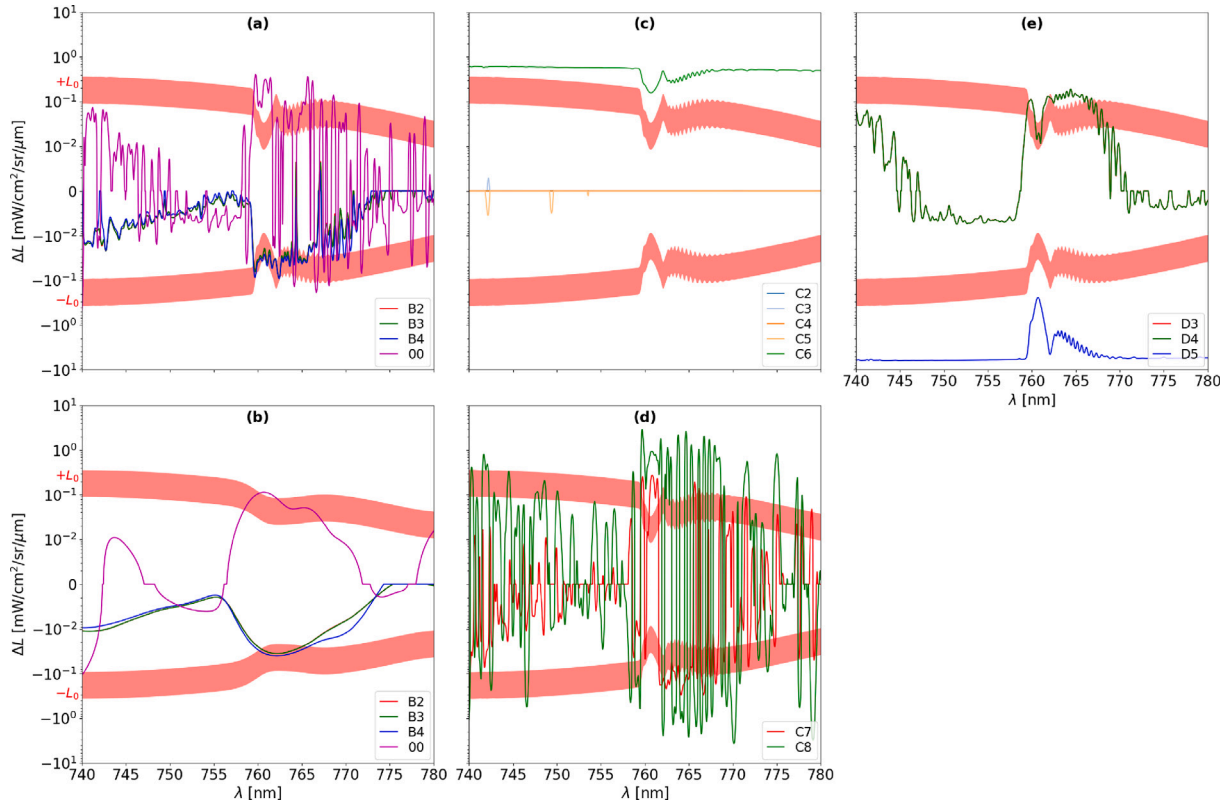


Fig. A.3. Sensitivity to radiative transfer options, namely radiative transfer model (a,b), number of sampling points in line-by-line model (c), resolution in correlated-k (slow) model (d) and multiple scattering (e). The impact of each parameter on the at-sensor radiance (with respect to the reference case A1 in Table A.1) is shown in comparison to the SIF sensitivity goal encompassed by the red bands. Parameter values that overshoot any of the red bands have an influence larger than the SIF signal. The sensitivity to the radiative transfer model is plotted for FWHM of 0.3 nm (a) and 3.5 nm (b), while in the other panels (c,d,e) the FWHM is 0.3 nm. Note that cases B2–3, C2–5 and D3–4 are essentially indistinguishable in the plots.

5 cm^{-1} , different sampling points per spectral bin for the line-by-line algorithm, see models C2–C8) and multiple scattering setups (8-stream Isaacs scaled, no multiple scattering, see models D2–D5). Case 00 using correlated-k (fast), 1 cm^{-1} resolution and 8-stream Isaacs scaled multiple scattering is also included for comparison since it is a common configuration used in MODTRAN5.

Fig. A.3 displays the at-sensor radiance differences ΔL as a function of wavelength for all radiative transfer cases along with our SIF sensitivity goal L_0 . Several important points are implied by this figure. First, as seen in Fig. A.3(a), the alternative radiative transfer models B2–4, namely correlated-k (slow or fast) and band model, are only borderline accurate compared to the sensitivity goal, leaving the line-by-line algorithm as the only viable option for our application. Fig. A.3(b) illustrates that the situation is somewhat relaxed for a DESIS-like FWHM of 3.5 nm, but even then the differences overshoot the expected SIF signal, especially in the oxygen absorption feature between 760 nm and 765 nm. It is also clear that case 00 is not adequate for our purposes.

Second, we are forced to select the finest available resolution of 0.1 cm^{-1} . Resolutions of 1 cm^{-1} or 5 cm^{-1} (only possible with the correlated-k algorithm or band model) are simply too inaccurate as demonstrated by the results of models C7–8 in Fig. A.3(d). However, it is possible to significantly reduce the number of sampling points in the line-by-line algorithm with little changes in the atmospheric functions or at-sensor radiance spectrum, cf. models C2–5 in Fig. A.3(c).

Lastly, the treatment of multiple scattering with the precise 8-stream DISORT option appears unavoidable in view of Fig. A.3(e). In fact, while the deactivation of multiple scattering (model D5) substantially

reduces the at-sensor radiance across the full spectral range, the Isaacs scaled recipe (with the correlated-k algorithm, models D3–4) leads to differences of the order of the expected SIF signal in the bulk of the oxygen absorption band. Note that the Isaacs scaled option is apparently incompatible with the line-by-line algorithm and thus case D2 has failed.

Table A.1 gives an overview of the results regarding attained accuracy (colour coded in the table) and required atmospheric module single-core run time (last column). After careful consideration, we selected case C4 (line-by-line algorithm, 0.1 cm^{-1} , 10 sampling points, 8-stream DISORT, 57 s) as the best compromise between simulation accuracy and running time. While the baseline case A1 is used for the sensitivity analysis, we adopted case C4 as our default radiative transfer configuration for the generation of the simulated datasets in Section 2.3. Note that the running time estimates presented here are only approximate and refer to nadir-looking geometries. For instance, the default case C4, with a running time of about 1 min as reported in Table A.1, may take up to 4 min in off-nadir configurations.

A.2. Atmosphere

The signal ultimately measured by the sensor is necessarily dependent on the atmosphere, namely its constituents, physical properties and conditions at the time of acquisition. Although the simulation framework described in Section 2.1 supports virtually all MODTRAN6 options to model the atmosphere, it is important to identify the key atmosphere parameters for our application and spectral range. We have

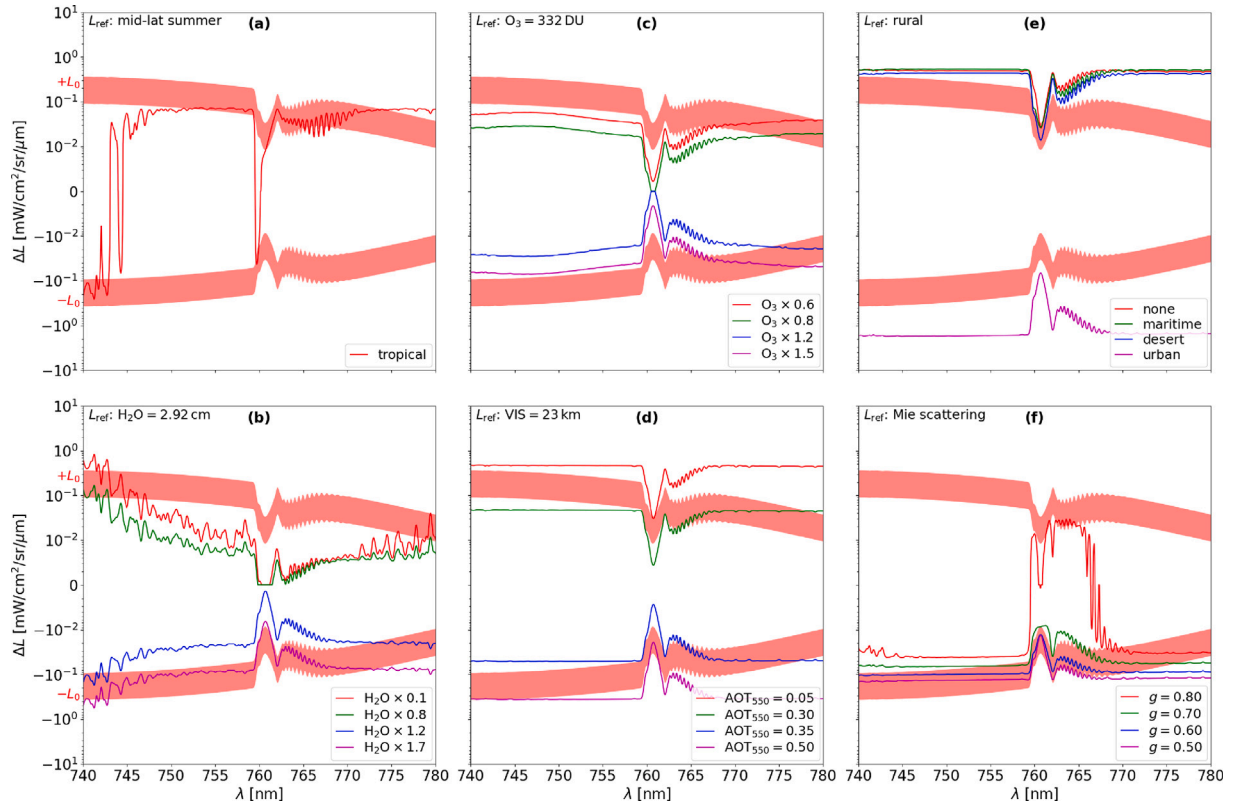


Fig. A.4. Sensitivity to atmosphere parameters, namely atmosphere model (a), water vapor (b), ozone (c), aerosol optical thickness at 550 nm (d), aerosol model (e) and aerosol scattering asymmetry parameter (f). The impact of each parameter on the at-sensor radiance (with respect to the reference case indicated in the top left of each plot) is shown in comparison to the SIF sensitivity goal encompassed by the red bands. Parameter values that overshoot any of the red bands have an influence larger than the SIF signal. The FWHM used for all panels is 0.3 nm.

focused on the following quantities: atmosphere model, water vapor, ozone, aerosol optical thickness at 550 nm, aerosol model and aerosol scattering asymmetry parameter. Fig. A.4 shows the results of our sensitivity analysis for these atmosphere parameters in terms of the comparison between ΔL and L_0 .

The influence of atmosphere and aerosol models in Fig. A.4(a,e) is expectedly larger than the percent-level SIF signal. This implies in particular that these models should be carefully chosen according to the regions and conditions of the measurement place. The water vapor content is another relevant quantity to consider when analysing radiance spectra between 740 nm and 780 nm, cf. Fig. A.4(b). There is actually no strong absorption due to water vapor at these wavelengths, but the tails of two small absorption features just below 740 nm and above 780 nm spill over to the considered range and have an effect of the order of the SIF sensitivity goal. Ozone instead has only a marginal role in the at-sensor radiance spectra around the O_2 -A band, as shown in Fig. A.4(c). We therefore set a default ozone value of 332 Dobson units (DU) and drop ozone from the input parameters to generate the simulated data. Note however that the simulation is really only representative of real data acquired in atmospheres of similar ozone content.

The aerosol optical thickness effectively determines the viewing conditions at the measurement place and naturally affects the at-sensor radiance spectra across the full wavelength range as seen in Fig. A.4(d). For the simulations, we opt to use AOT_{550} instead of visibility since it is a parameter with a clear physical interpretation and easier to sample linearly in a given range of values. It is important to point out that in SIF retrieval applications we will mainly be dealing with data acquired

with good or very good visibilities, so the range of AOT_{550} should not include hazy conditions.

Lastly, the simulated spectrum is sensitive to the aerosol scattering angular distribution, which in MODTRAN6 is computed with Mie scattering for the aerosol profile (default option) or may be parameterized by the quantity g (−1 for pure back scattering, +1 for pure forward scattering). Fig. A.4(f) shows that different scattering angular distributions lead to spectral distortions of the order of our sensitivity goal. Note in addition that there is a complicated dependence on the observation geometry. For instance, if sun, sensor and target are aligned, then any amount of aerosol back scattering can have a huge impact on the measured radiance.

A.3. Geometry

Next we concentrate our attention on the specific observation geometry. Fig. A.5 shows the impact of varying sensor tilt angle, sun zenith angle, relative azimuth angle between sun and line of sight, ground altitude and sensor altitude in reasonable ranges for airborne and space-based spectrometers. The three angles are defined at the sensor point for the purpose of the sensitivity analysis. The influence of the sun zenith angle is the largest and abundantly exceeds the sensitivity goal, cf. Fig. A.5(b). As the sun gets closer to zenith (i.e., as the sun zenith angle decreases), the solar irradiance at the surface increases significantly and so does the at-sensor radiance spectrum in the whole wavelength range. The tilt angle has a smaller influence on the radiance spectrum, but still close or above the expected SIF signal, as seen in Fig. A.5(a). A more inclined line of sight (i.e., a larger tilt angle) leads to

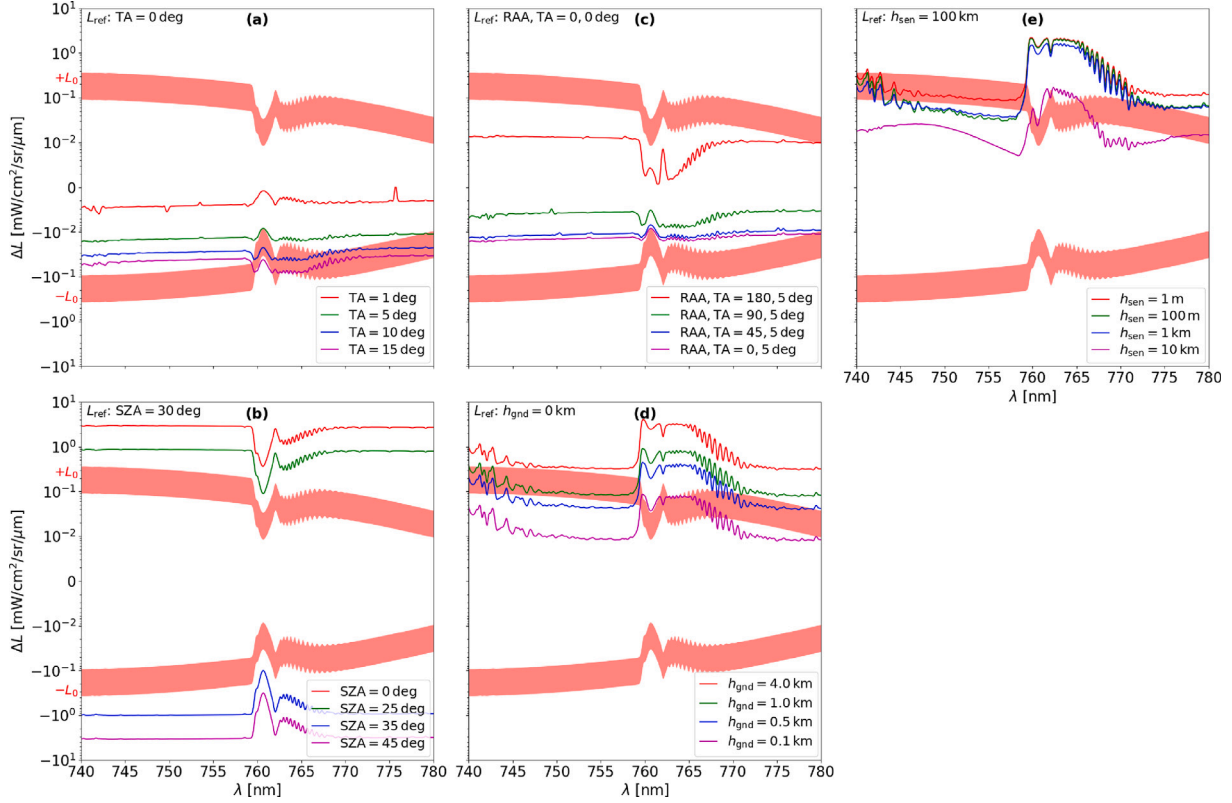


Fig. A.5. Sensitivity to geometry parameters, namely tilt angle (a), sun zenith angle (b), relative azimuth angle (c), ground altitude (d) and sensor altitude (e). The impact of each parameter on the at-sensor radiance (with respect to the reference case indicated in the top left of each plot) is shown in comparison to the SIF sensitivity goal encompassed by the red bands. Parameter values that overshoot any of the red bands have an influence larger than the SIF signal. The FWHM used for all panels is 0.3 nm.

more absorption and more scattering with the net effect of decreasing the at-sensor radiance by up to a few percent for the case where the sun zenith angle at the sensor is fixed to 30°. For nadir observations (zero tilt angle), the relative azimuth angle is irrelevant, but it becomes increasingly important as the tilt angle increases. The effect shown in Fig. A.5(c) is marginal when compared to the sensitivity goal in the case of 5° off-nadir observations. Nevertheless, we shall include the relative azimuth angle in the generation of our databases so that the cases of moderate tilt angles are appropriately covered.

Fig. A.5(d,e) further demonstrates that the simulated spectrum is very sensitive to both ground altitude (for airborne and space-based instruments) and sensor altitude (for airborne instruments only). This is to be expected, because we are focusing on a strong oxygen absorption band and the oxygen content of the atmosphere varies quickly with altitude, especially at low altitudes. In particular, there is less oxygen along the line of sight for higher ground altitudes and smaller sensor altitudes, leading to less absorption and thus a larger at-sensor radiance spectrum. Note that for instruments in space the exact sensor altitude does not play a role, because for our purposes the atmosphere does not extend beyond altitudes of 100 km. Therefore, any space instrument can be simulated assuming a sensor altitude of 100 km and the appropriate adaptation of the geometry viewing angles.

It is important to stress that all geometry parameters studied above are either stored directly in typical data products or can be easily reconstructed from existing metadata. These quantities do not need to be inferred by SIF retrieval methods from the measured spectra. Therefore, in order to accommodate a diverse range of observation geometries for both airborne and space-based instruments, all five geometry parameters are considered in the generation of the simulated

datasets in Section 2.3. This has no significant extra cost apart from the increase of input dimensionality.

A.4. Surface

In the scope of the sensitivity analysis, we have considered the linear reflectance function (ρ_{740} , s , $e = 1$) and the Gaussian SIF emission (F_{737}) described in Section 2.2.1 to model the surface properties of the target. We illustrate the effect of the three parameters ρ_{740} , s and F_{737} on the at-sensor radiance spectrum in Fig. A.6(a,b,c) for the case of DESIS. Clearly, SIF retrieval methods must have a very good handle on both reflectance ρ_{740} and reflectance slope s in order to make a precise measurement of fluorescence.

A.5. Sensor

A precise knowledge of the spectral response function of the instrument is imperative for SIF retrieval. Any slight spectral mischaracterization or sensor instability resulting in shifts of the central wavelength and/or full width at half maximum (i.e., spectral resolution) have an impact greater than the small SIF signal. This is exemplified for DESIS in Fig. A.6(d,e). The sensitivity on central wavelengths in Fig. A.6(d) is particularly noteworthy. The spectral characterization of both DESIS and HyPlant between 740 nm and 780 nm presented in Section 2.2.2 is therefore crucial to define a precise spectral response function for our simulations. On top of a reference spectral response function, we have allowed for shifts on CW and FWHM (δ_{CW} , δ_{FWHM} , cf. Section 2.1.2) so that the simulation can encompass the real performance of the instruments during acquisition. This will allow future retrieval methods to implicitly correct for these shifts while extracting the SIF signal.

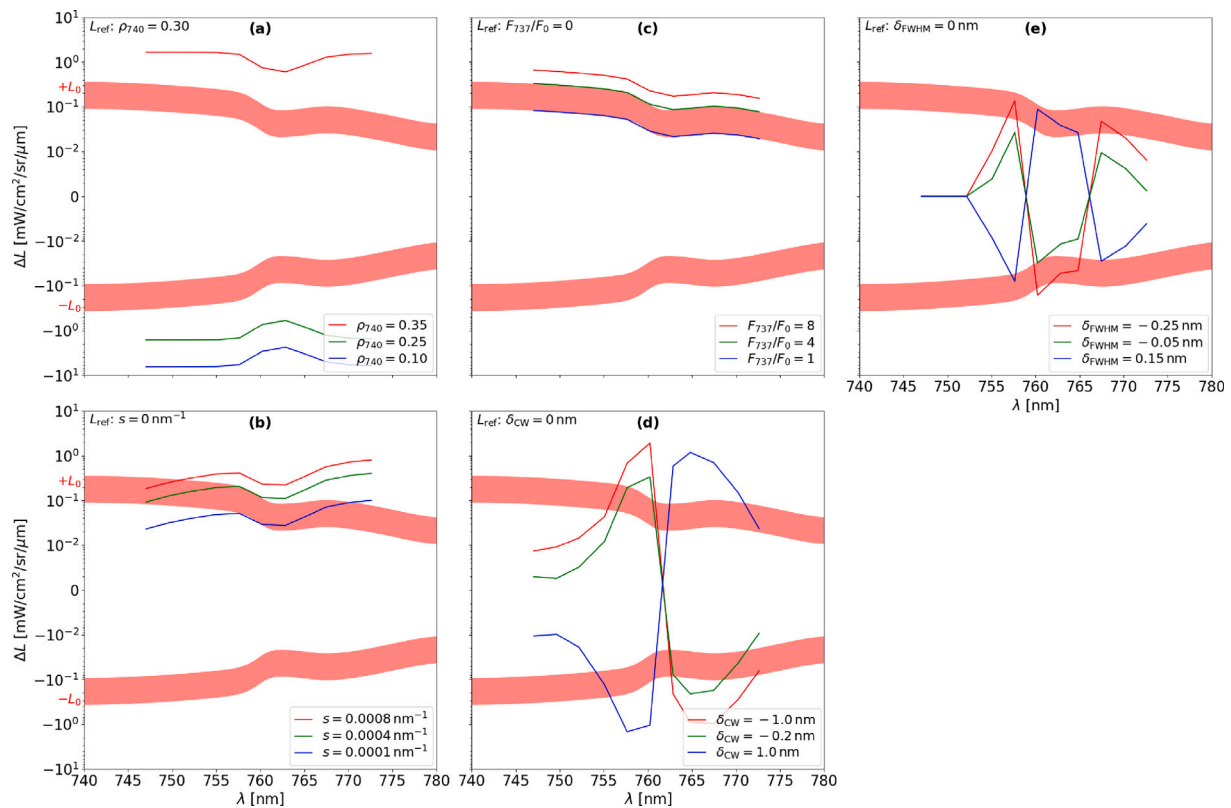


Fig. A.6. Sensitivity to surface and sensor parameters, namely reflectance at 740 nm (a), reflectance spectral slope at 740 nm (b), ground SIF radiance at 737 nm (c), CW shift (d) and FWHM change (e). The impact of each parameter on the at-sensor radiance (with respect to the reference case indicated in the top left of each plot) is shown in comparison to the SIF sensitivity goal encompassed by the red bands. Parameter values that overshoot any of the red bands have an influence larger than the SIF signal. The DESIS spectral response function (with approximately FWHM of 3.5 nm and spectral sampling distance of 2.5 nm) was used for all panels.

Data availability

The authors do not have permission to share data.

References

- Ač, A., Malenovsky, Z., Olejníčková, J., Gallé, A., Rascher, U., Mohammed, G., 2015. Meta-analysis assessing potential of steady-state chlorophyll fluorescence for remote sensing detection of plant water, temperature and nitrogen stress. *Remote Sens. Environ.* 168, 420–436. <http://dx.doi.org/10.1016/j.rse.2015.07.022>.
- AERONET FZJ-JOYCE, 2024. AERONET site information database. URL: https://aeronet.gsfc.nasa.gov/new_web/photo_db_v3/FZJ-JOYCE.html.
- Alonso, K., Bachmann, M., Burch, K., Carmona, E., Cerra, D., de los Reyes, R., Dietrich, D., Heiden, U., Hölderlin, A., Ickes, J., Knodt, U., Krutz, D., Lester, H., Müller, R., Pagnutti, M., Reinartz, P., Richter, R., Ryan, R., Sebastian, I., Tegler, M., 2019. Data products, quality and validation of the DLR earth sensing imaging spectrometer (DESIS). *Sensors* 19 (20), <http://dx.doi.org/10.3390/s19204471>.
- Alonso, L., Gomez-Chova, L., Vila-Frances, J., Amoros-Lopez, J., Guanter, L., Calpe, J., Moreno, J., 2008. Improved Fraunhofer line discrimination method for vegetation fluorescence quantification. *IEEE Geosci. Remote. Sens. Lett.* 5 (4), 620–624. <http://dx.doi.org/10.1109/LGRS.2008.2001180>.
- Amoros-Lopez, J., Gomez-Chova, L., Vila-Frances, J., Alonso, L., Calpe, J., Moreno, J., del Valle-Tascon, S., 2008. Evaluation of remote sensing of vegetation fluorescence by the analysis of diurnal cycles. *Int. J. Remote Sens.* 29 (17–18), 5423–5436. <http://dx.doi.org/10.1080/01431160802036391>.
- Bendig, J., Malenovsky, Z., Gautam, D., Lucier, A., 2020. Solar-induced chlorophyll fluorescence measured from an unmanned aircraft system: Sensor etaloning and platform motion correction. *IEEE Trans. Geosci. Remote Sens.* 58 (5), 3437–3444. <http://dx.doi.org/10.1109/TGRS.2019.2956194>.
- Berk, A., Conforti, P., Hawes, F., 2015. An accelerated line-by-line option for MODTRAN combining on-the-fly generation of line center absorption within 0.1 cm⁻¹ bins and pre-computed line tails. In: *Proc. SPIE 9471, Algorithms and Technologies* for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI. 947217. <http://dx.doi.org/10.1117/12.2177444>.
- Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F., van den Bosch, J., 2014. MODTRAN6: a major upgrade of the MODTRAN radiative transfer code. In: *Proc. SPIE 9088, Algorithms and Technologies for Multi-spectral, Hyperspectral, and Ultraspectral Imagery XX*. p. 90880H. <http://dx.doi.org/10.1117/12.2050433>.
- Bouvet, M., Thome, K., Berthelot, B., Bialek, A., Czaplá-Myers, J., Fox, N.P., Goryl, P., Henry, P., Ma, L., Marcq, S., Meygret, A., Wenny, B.N., Woolliams, E.R., 2019. RadCalNet: A radiometric calibration network for earth observing imagers operating in the visible to shortwave infrared spectral range. *Remote. Sens.* 11 (20), <http://dx.doi.org/10.3390/rs11202401>, URL: <https://www.mdpi.com/2072-4292/11/20/2401>.
- Buffat, J., Pato, M., Alonso, K., Auer, S., Carmona, E., Maier, S., Müller, R., Rademske, P., Rascher, U., Scharr, H., 2023. Deep learning based prediction of sun-induced fluorescence from hyplant imagery. In: *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*. pp. 2993–2996. <http://dx.doi.org/10.1109/IGARSS52108.2023.10282828>.
- Buffat, J., Pato, M., Alonso, K., Auer, S., Carmona, E., Maier, S., Müller, R., Rademske, P., Rascher, U., Scharr, H., 2025a. Retrieval of sun-induced plant fluorescence in the O2-A absorption band from DESIS imagery. In: *Computer Vision – ECCV 2024 Workshops*. pp. 81–100. http://dx.doi.org/10.1007/978-3-031-91835-3_6, Published in the Proceedings of the 9th Computer Vision in Plant Phenotyping and Agriculture (CVPPA) Workshop, European Conference on Computer Vision (ECCV) 2024, 29 Sep 2024, Milan, Italy..
- Buffat, J., Pato, M., Alonso, K., Auer, S., Carmona, E., Maier, S., Müller, R., Rademske, P., Rascher, U., Scharr, H., 2025b. Emulation-based self-supervised SIF retrieval in the O2-A absorption band with HyPlant. <http://dx.doi.org/10.22541/essoar.174000855.50541566/v2>, preprint in ESS Open Archive, February 19, 2025.
- Buffat, J., Pato, M., Alonso, K., Auer, S., Carmona, E., Maier, S., Müller, R., Rademske, P., Siegmund, B., Rascher, U., Scharr, H., 2025c. A multi-layer perceptron approach for SIF retrieval in the O2-A absorption band from hyperspectral imagery of the HyPlant airborne sensor system. *Remote Sens. Environ.* 318, 114596. <http://dx.doi.org/10.1016/j.rse.2024.114596>, URL: <https://www.sciencedirect.com/science/article/pii/S0034425724006229>.

- Buffat, J., Rascher, U., Rademske, P., Siegmann, B., Junker-Frohn, L.V., Emin, D., Grygosch, L., 2024b. HyData: HyPlant FLUO at-sensor radiance data packages and FLOX measurements for SIF retrieval method development from selected campaigns of the years 2018 - 2023. <http://dx.doi.org/10.26165/JUELICH-DATA/QKJKPW>.
- Carmona, E., Alonso, K., Bachmann, M., Burch, K., Cerra, D., De los Reyes, R., Heiden, U., Knodt, U., Krutz, D., Marshall, D., Müller, R., Pagnutti, M., Reinartz, P., Ryan, R., 2021. Vicarious calibration of the DESIS imaging spectrometer. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. pp. 1611–1614. <http://dx.doi.org/10.1109/IGARSS47720.2021.9553866>.
- Cendrero-Mateo, M.P., Wieneke, S., Damm, A., Alonso, L., Pinto, F., Moreno, J., Guanter, L., Celesti, M., Rossini, M., Sabater, N., Cogliati, S., Julitta, T., Rascher, U., Goulas, Y., Aasen, H., Pacheco-Labrador, J., Mac Arthur, A., 2019. Sun-induced chlorophyll fluorescence III: Benchmarking retrieval methods and sensor characteristics for proximal sensing. *Remote Sens.* 11 (8), <http://dx.doi.org/10.3390/rs11080962>.
- Coddington, O.M., Richard, E.C., Harber, D., Pilewskie, P., Woods, T.N., Snow, M., Chance, K., Liu, X., Sun, K., 2023. Version 2 of the TSIS-1 hybrid solar reference spectrum and extension to the full spectrum. *Earth Space Sci.* 10 (3), <http://dx.doi.org/10.1029/2022EA002637>.
- Cogliati, S., Celesti, M., Cesana, I., Miglietta, F., Genesio, L., Julitta, T., Schuettemeyer, D., Drusch, M., Rascher, U., Jurado, P., Colombo, R., 2019. A spectral fitting algorithm to retrieve the fluorescence spectrum from canopy radiance. *Remote Sens.* 11 (16), <http://dx.doi.org/10.3390/rs11161840>.
- Cogliati, S., Colombo, R., Celesti, M., Tagliabue, G., Rascher, U., Schickling, A., Rademske, P., Alonso, L., Sabater, N., Schuettemeyer, D., Drusch, M., 2018. Red and far-red fluorescence emission retrieval from airborne high-resolution spectra collected by the Hyplant-Fluo sensor. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. pp. 3935–3938. <http://dx.doi.org/10.1109/IGARSS.2018.8517758>.
- Cogliati, S., Rossini, M., Julitta, T., Meroni, M., Schickling, A., Burkart, A., Pinto, F., Rascher, U., Colombo, R., 2015a. Continuous and long-term measurements of reflectance and sun-induced chlorophyll fluorescence by using novel automated field spectroscopy systems. *Remote Sens. Environ.* 164, 270–281. <http://dx.doi.org/10.1016/j.rse.2015.03.027>.
- Cogliati, S., Verhoef, W., Kraft, S., Sabater, N., Alonso, L., Vicent, J., Moreno, J., Drusch, M., Colombo, R., 2015b. Retrieval of sun-induced fluorescence using advanced spectral fitting methods. *Remote Sens. Environ.* 169, 344–357. <http://dx.doi.org/10.1016/j.rse.2015.08.022>.
- Damm, A., Guanter, L., Laurent, V., Schaepman, M., Schickling, A., Rascher, U., 2014. FLD-based retrieval of sun-induced chlorophyll fluorescence from medium spectral resolution airborne spectroscopy data. *Remote Sens. Environ.* 147, 256–266. <http://dx.doi.org/10.1016/j.rse.2014.03.009>.
- Damm, A., Paul-Limoges, E., Haghighi, E., Simmer, C., Morsdorf, F., Schneider, F., van der Tol, C., Migliavacca, M., Rascher, U., 2018. Remote sensing of plant-water relations: An overview and future perspectives. *J. Plant Physiol.* 227, 3–19. <http://dx.doi.org/10.1016/j.jplph.2018.04.012>, From aquaporin to ecosystem: Plants in the water cycle.
- de los Reyes, R., Langheinrich, M., Schwind, P., Richter, R., Pflug, B., Bachmann, M., Müller, R., Carmona, E., Zekoll, V., Reinartz, P., 2020. PACO: Python-based atmospheric correction. *Sensors* 20 (5), <http://dx.doi.org/10.3390/s20051428>.
- Drusch, M., Moreno, J., Del Bello, U., Franco, R., Goulas, Y., Huth, A., Kraft, S., Middleton, E.M., Miglietta, F., Mohammed, G., Nedbal, L., Rascher, U., Schuettemeyer, D., Verhoef, W., 2017. The fluorescence explorer mission concept—ESA's earth explorer 8. *IEEE Trans. Geosci. Remote Sens.* 55 (3), 1273–1284. <http://dx.doi.org/10.1109/TGRS.2016.2621820>.
- Fontenla, J.M., Harder, J., Livingston, B., Snow, M., Woods, T., 2011. High-resolution solar spectral irradiance from extreme ultraviolet to far infrared. *J. Geophys. Res.: Atmos.* 116 (D20), <http://dx.doi.org/10.1029/2011JD016032>.
- Frankenberg, C., Köhler, P., Magney, T.S., Geier, S., Lawson, P., Schwachert, M., McDuffie, J., Drewry, D.T., Pavlick, R., Kuhnert, A., 2018. The chlorophyll fluorescence imaging spectrometer (CFIS), mapping far red fluorescence from aircraft. *Remote Sens. Environ.* 217, 523–536. <http://dx.doi.org/10.1016/j.rse.2018.08.032>.
- Garzonio, R., Di Mauro, B., Colombo, R., Cogliati, S., 2017. Surface reflectance and sun-induced fluorescence spectroscopy measurements using a small hyperspectral UAS. *Remote Sens.* 9 (5), <http://dx.doi.org/10.3390/rs9050472>.
- Guanter, L., Bacour, C., Schneider, A., Aben, I., van Kempen, T.A., Maignan, F., Retscher, C., Köhler, P., Frankenberg, C., Joiner, J., Zhang, Y., 2021. The TROPISIF global sun-induced fluorescence dataset from the sentinel-5P TROPOMI mission. *Earth Syst. Sci. Data* 13 (11), 5423–5440. <http://dx.doi.org/10.5194/essd-13-5423-2021>.
- Guanter, L., Richter, R., Kaufmann, H., 2009. On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing. *Int. J. Remote Sens.* 30 (6), 1407–1424. <http://dx.doi.org/10.1080/01431160802438555>.
- Halton, J., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.* 2, 84–90. <http://dx.doi.org/10.1007/BF01386213>.
- Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A.P., Middleton, E.M., Huemmrich, K.F., Yoshida, Y., Frankenberg, C., 2013. Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2. *Atmospheric Meas. Tech.* 6 (10), 2803–2823. <http://dx.doi.org/10.5194/amt-6-2803-2013>.
- Jülich Supercomputing Centre, 2021. JURECA: Data centric and booster modules implementing the modular supercomputing architecture at Jülich supercomputing centre. *J. Large-Scale Res. Facil.* 7 (A182), <http://dx.doi.org/10.17815/jlsrf-7-182>.
- Julitta, T., Corp, L.A., Rossini, M., Burkart, A., Cogliati, S., Davies, N., Hom, M., Mac Arthur, A., Middleton, E.M., Rascher, U., Schickling, A., Colombo, R., 2016. Comparison of sun-induced chlorophyll fluorescence estimates obtained from four portable field spectroradiometers. *Remote Sens.* 8 (2), <http://dx.doi.org/10.3390/rs8020122>.
- Kinne, S., 2019. The MACv2 aerosol climatology. *Tellus B: Chem. Phys. Meteorol.* 71 (1), 1–21. <http://dx.doi.org/10.1080/16000889.2019.1623639>.
- Köhler, P., Behrenfeld, M.J., Landgraf, J., Joiner, J., Magney, T.S., Frankenberg, C., 2020. Global retrievals of solar-induced chlorophyll fluorescence at red wavelengths with TROPOMI. *Geophys. Res. Lett.* 47 (15), <http://dx.doi.org/10.1029/2020GL087541>, e2020GL087541. 10.1029/2020GL087541.
- Krämer, J., 2024. Multi-scale field phenotyping of wheat-bean intercrops: Integrating spectral and agronomic datasets from a three-year trial. <http://dx.doi.org/10.26165/JUELICH-DATA/XWRNHR>.
- Krutz, D., Müller, R., Knodt, U., Günther, B., Walter, I., Sebastian, I., Säuberlich, T., Reulke, R., Carmona, E., Eckardt, A., Venus, H., Fischer, C., Zender, B., Arloth, S., Lieder, M., Neidhardt, M., Grote, U., Schrandt, F., Gelmi, S., Wojtkowiak, A., 2019. The instrument design of the DLR earth sensing imaging spectrometer (DESIS). *Sensors* 19 (7), <http://dx.doi.org/10.3390/s19071622>, URL: <https://www.mdpi.com/1424-8220/19/7/1622>.
- Maes, W.H., Dorigo, B.R., Martens, B., Gentine, P., Guanter, L., Steppe, K., Verhoef, N.E., Dorigo, W., Li, X., Xiao, J., Miralles, D.G., 2020. Sun-induced fluorescence closely linked to ecosystem transpiration as evidenced by satellite data and radiative transfer models. *Remote Sens. Environ.* 249, 112030. <http://dx.doi.org/10.1016/j.rse.2020.112030>.
- Maier, S.W., 2000. Modeling the Radiative Transfer in Leaves in the 300nm to 2.5um Wavelength Region Taking into Consideration Chlorophyll Fluorescence - The Leaf Model SLOPE (Ph.D. thesis). Technische Universität München.
- Maier, S.W., Günther, K.P., Stellmes, M., 2003. Sun-induced fluorescence: a new tool for precision farming. In: Schepers, J., VanToai, T. (Eds.), *Digital Imaging and Spectral Techniques: Applications to Precision Agriculture and Crop Physiology*. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, Wisconsin, USA, pp. 209–222. <http://dx.doi.org/10.2134/asspecpub66.c16>.
- Martini, D., Sakowska, K., Wohlfahrt, G., Pacheco-Labrador, J., van der Tol, C., Porcar-Castell, A., Magney, T.S., Carrara, A., Colombo, R., El-Madany, T.S., Gonzalez-Cascon, R., Martin, M.P., Julitta, T., Moreno, G., Rascher, U., Reichstein, M., Rossini, M., Migliavacca, M., 2022. Heatwave breaks down the linearity between sun-induced fluorescence and gross primary production. *New Phytol.* 233 (6), 2415–2428. <http://dx.doi.org/10.1111/nph.17920>.
- Mazzoni, M., Meroni, M., Fortunato, C., Colombo, R., Verhoef, W., 2012. Retrieval of maize canopy fluorescence and reflectance by spectral fitting in the O2-A absorption band. *Remote Sens. Environ.* 124, 72–82. <http://dx.doi.org/10.1016/j.rse.2012.04.025>, URL: <https://www.sciencedirect.com/science/article/pii/S0034425712002015>.
- Meroni, M., Busetto, L., Colombo, R., Guanter, L., Moreno, J., Verhoef, W., 2010. Performance of spectral fitting methods for vegetation fluorescence quantification. *Remote Sens. Environ.* 114 (2), 363–374. <http://dx.doi.org/10.1016/j.rse.2009.09.010>, URL: <https://www.sciencedirect.com/science/article/pii/S003442570900028X>.
- Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R., Moreno, J., 2009. Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sens. Environ.* 113 (10), 2037–2051. <http://dx.doi.org/10.1016/j.rse.2009.05.003>.
- Mohammed, G.H., Colombo, R., Middleton, E.M., Rascher, U., van der Tol, C., Nedbal, L., Goulas, Y., Pérez-Prigio, O., Damm, A., Meroni, M., Joiner, J., Cogliati, S., Verhoef, W., Malenovsky, Z., Gastellu-Etchegorry, J.-P., Miller, J.R., Guanter, L., Moreno, J., Moya, I., Berry, J.A., Frankenberg, C., Zarco-Tejada, P.J., 2019. Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress. *Remote Sens. Environ.* 231, 111177. <http://dx.doi.org/10.1016/j.rse.2019.04.030>.
- Norton, A.J., Rayner, P.J., Koffi, E.N., Scholze, M., Silver, J.D., Wang, Y.-P., 2019. Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model. *Biogeosciences* 16 (15), 3069–3093. <http://dx.doi.org/10.5194/bg-16-3069-2019>.
- Owen, A.B., 2017. A randomized Halton algorithm in R. <http://dx.doi.org/10.48550/ARXIV.1706.02808>, URL: <https://arxiv.org/abs/1706.02808>.
- Pato, M., Alonso, K., Auer, S., Buffat, J., Carmona, E., Maier, S., Müller, R., Rademske, P., Rascher, U., Scharr, H., 2023. Fast machine learning simulator of at-sensor radiances for solar-induced fluorescence retrieval with DESIS and hyplant. In: IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium. pp. 7563–7566. <http://dx.doi.org/10.1109/IGARSS52108.2023.10281579>.
- Pato, M., Buffat, J., Alonso, K., Auer, S., Carmona, E., Maier, S., Müller, R., Rademske, P., Rascher, U., Scharr, H., 2024. Physics-based machine learning emulator of at-sensor radiances for solar-induced fluorescence retrieval in the O₂-A absorption band. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 17, 18566–18576. <http://dx.doi.org/10.1109/JSTARS.2024.3457231>.

- Peng, H., Cendrero-Mateo, M.P., Bendig, J., Siegmann, B., Acebron, K., Kneer, C., Kataja, K., Muller, O., Rascher, U., 2022. Hyscreen: A ground-based imaging system for high-resolution red and far-red solar-induced chlorophyll fluorescence. *Sensors* 22 (23), <http://dx.doi.org/10.3390/s22239443>.
- Pinto, F., Celesti, M., Acebron, K., Alberti, G., Cogliati, S., Colombo, R., Juszczak, R., Matsubara, S., Miglietta, F., Palombo, A., Panigada, C., Pignatti, S., Rossini, M., Sakowska, K., Schickling, A., Schüttemeyer, D., Stróżecki, M., Tudoroiu, M., Rascher, U., 2020. Dynamics of sun-induced chlorophyll fluorescence and reflectance to detect stress-induced variations in canopy photosynthesis. *Plant Cell Environ.* 43 (7), 1637–1654. <http://dx.doi.org/10.1111/pce.13754>.
- Porcar-Castell, A., Malenovsky, Z., Magney, T., Wittenberghe, S.V., Fernández-Marín, B., Maignan, F., Zhang, Y., Maseyk, K., Atherton, J., Albert, L.P., Robson, T.M., Zhao, F., Garcia-Plazaola, J.-I., Ensminger, I., Rajewicz, P.A., Grebe, S., Tikkanen, M., Kellner, J.R., Ihalainen, J.A., Rascher, U., Logan, B., 2021. Chlorophyll a fluorescence illuminates a path connecting plant molecular biology to earth-system science. *Nat. Plants* 7, 998–1009. <http://dx.doi.org/10.1038/s41477-021-00980-4>.
- Rascher, U., Alonso, L., Burkart, A., Cilia, C., Cogliati, S., Colombo, R., Damm, A., Drusch, M., Guanter, L., Hanus, J., Hyvärinen, T., Julitta, T., Jussila, J., Kataja, K., Kokkalis, P., Kraft, S., Kraska, T., Matveeva, M., Moreno, J., Muller, O., Panigada, C., Píkl, M., Pinto, F., Prey, L., Pude, R., Rossini, M., Schickling, A., Schurr, U., Schüttemeyer, D., Verrelst, J., Zemek, F., 2015. Sun-induced fluorescence – a new probe of photosynthesis: First maps from the imaging spectrometer HyPlant. *Global Change Biol.* 21 (12), 4673–4684. <http://dx.doi.org/10.1111/gcb.13017>.
- Rossini, M., Meroni, M., Celesti, M., Cogliati, S., Julitta, T., Panigada, C., Rascher, U., Van der Tol, C., Colombo, R., 2016. Analysis of red and far-red sun-induced chlorophyll fluorescence and their ratio in different canopies based on observed and modeled data. *Remote Sens.* 8 (5), <http://dx.doi.org/10.3390/rs8050412>.
- Sabater, N., Kolmonen, P., Van Wittenberghe, S., Arola, A., Moreno, J., 2021. Challenges in the atmospheric characterization for the retrieval of spectrally resolved fluorescence and PRI region dynamics from space. *Remote Sens. Environ.* 254, 112226. <http://dx.doi.org/10.1016/j.rse.2020.112226>.
- Sabater, N., Vicent, J., Alonso, L., Verrelst, J., Middleton, E.M., Porcar-Castell, A., Moreno, J., 2018. Compensation of oxygen transmittance effects for proximal sensing retrieval of canopy-leaving sun-Induced chlorophyll fluorescence. *Remote Sens.* 10 (10), <http://dx.doi.org/10.3390/rs10101551>.
- Scodellaro, R., Cesana, I., D'Alfonso, L., Bouzin, M., Collini, M., Chirico, G., Colombo, R., Miglietta, F., Celesti, M., Schuettemeyer, D., Cogliati, S., Sironi, L., 2022. A novel hybrid machine learning phasor-based approach to retrieve a full set of solar-induced fluorescence metrics and biophysical parameters. *Remote Sens. Environ.* 280, 113196. <http://dx.doi.org/10.1016/j.rse.2022.113196>.
- Siegmann, B., Alonso, L., Celesti, M., Cogliati, S., Colombo, R., Damm, A., Douglas, S., Guanter, L., Hanuš, J., Kataja, K., Kraska, T., Matveeva, M., Moreno, J., Muller, O., Píkl, M., Pinto, F., Quirós Vargas, J., Rademske, P., Rodríguez-Moreno, F., Sabater, N., Schickling, A., Schüttemeyer, D., Zemek, F., Rascher, U., 2019. The high-performance airborne imaging spectrometer HyPlant—From raw images to top-of-canopy reflectance and fluorescence products: Introduction of an automatized processing chain. *Remote Sens.* 11 (23), <http://dx.doi.org/10.3390/rs11232760>.
- SPECIM, 2024. Specim, spectral imaging ltd. URL: <https://www.specim.com/>.
- van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., Su, Z., 2009. An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. *Biogeosciences* 6 (12), 3109–3129. <http://dx.doi.org/10.5194/bg-6-3109-2009>, URL: <https://bg.copernicus.org/articles/6/3109/2009/>.
- Van Wittenberghe, S., Alonso, L., Verrelst, J., Hermans, I., Delegido, J., Veroustraete, F., Valcke, R., Moreno, J., Samson, R., 2013. Upward and downward solar-induced chlorophyll fluorescence yield indices of four tree species as indicators of traffic pollution in Valencia. *Environ. Pollut.* 173, 29–37. <http://dx.doi.org/10.1016/j.envpol.2012.10.003>, URL: <https://www.sciencedirect.com/science/article/pii/S0269749112004423>.
- Vargas, J.Q., Bendig, J., Mac Arthur, A., Burkart, A., Julitta, T., Maseyk, K., Thomas, R., Siegmann, B., Rossini, M., Celesti, M., Schüttemeyer, D., Kraska, T., Muller, O., Rascher, U., 2020. Unmanned aerial systems (UAS)-based methods for solar induced chlorophyll fluorescence (SIF) retrieval with non-imaging spectrometers: State of the art. *Remote Sens.* 12 (10), <http://dx.doi.org/10.3390/rs12101624>.
- Vicent Servera, J., Rivera-Caicedo, J.P., Verrelst, J., Muñoz-Marí, J., Sabater, N., Berthelot, B., Camps-Valls, G., Moreno, J., 2022. Systematic assessment of MODTRAN emulators for atmospheric correction. *IEEE Trans. Geosci. Remote Sens.* 60, 1–17. <http://dx.doi.org/10.1109/TGRS.2021.3071376>.
- Wang, N., Siegmann, B., Rascher, U., Clevers, J.G., Muller, O., Bartholomeus, H., Bendig, J., Masiliūnas, D., Pude, R., Kooistra, L., 2022. Comparison of a UAV- and an airborne-based system to acquire far-red sun-induced chlorophyll fluorescence measurements over structurally different crops. *Agric. Forest. Meteorol.* 323, 109081. <http://dx.doi.org/10.1016/j.agrformet.2022.109081>.
- Yang, P., Prikaziuk, E., Verhoef, W., van der Tol, C., 2021. SCOPE 2.0: a model to simulate vegetated land surface fluxes and satellite signals. *Geosci. Model. Dev.* 14 (7), 4697–4712. <http://dx.doi.org/10.5194/gmd-14-4697-2021>, URL: <https://gmd.copernicus.org/articles/14/4697/2021/>.
- Zeng, Y., Chen, M., Hao, D., Damm, A., Badgley, G., Rascher, U., Johnson, J.E., Dechant, B., Siegmann, B., Ryu, Y., Qiu, H., Krieger, V., Panigada, C., Celesti, M., Miglietta, F., Yang, X., Berry, J.A., 2022. Combining near-infrared radiance of vegetation and fluorescence spectroscopy to detect effects of abiotic changes and stresses. *Remote Sens. Environ.* 270, 112856. <http://dx.doi.org/10.1016/j.rse.2021.112856>.
- Zeng, Y., Hao, D., Badgley, G., Damm, A., Rascher, U., Ryu, Y., Johnson, J., Krieger, V., Wu, S., Qiu, H., Liu, Y., Berry, J.A., Chen, M., 2021. Estimating near-infrared reflectance of vegetation from hyperspectral data. *Remote Sens. Environ.* 267, 112723. <http://dx.doi.org/10.1016/j.rse.2021.112723>.