Physics-Based Machine Learning Emulator of at-Sensor Radiances for Solar-Induced Fluorescence Retrieval in the O_2-A Absorption Band

Miguel Pato, Jim Buffat, Kevin Alonso, Stefan Auer, *Member, IEEE*, Emiliano Carmona, Stefan Maier, Rupert Müller, Patrick Rademske, Uwe Rascher, and Hanno Scharr

***Abstract*—The successful operation of airborne and space-based spectrometers in recent years holds the promise to map solar-induced fluorescence (SIF) accurately across the globe. Machine learning (ML) can play an important role in this effort, but its application to SIF retrieval methods is in part hindered by the need for time-consuming radiative transfer modeling to account for atmospheric effects. In this work, we address this difficulty and develop a fast and accurate physics-based ML emulator of at-sensor radiances around the O_2-A absorption band for the space-based DESIS and the airborne HyPlant spectrometers. Different ML models are trained on an extensive set of simulated spectra encompassing a wide range of atmosphere, geometry, surface, and sensor configurations. A fourth-degree polynomial model is found to perform best, presenting errors at or below 10% of typical SIF at-sensor radiances and a prediction time per sample spectrum of $10-20$ μs. Using data acquired with the HyPlant instrument, the proposed model is also shown to be able to match very closely the measured spectra. We illustrate how to improve further the accuracy of the emulator and how to generalize it to other sensors using the particular case of ESA's FLEX space mission. Our findings suggest that physics-based emulators can be efficiently used for the development of ML-based SIF retrieval methods by generating large training datasets in short time and by enabling a fast simulation step for self-supervised retrieval schemes.**

***Index Terms*—Hyperspectral sensors, machine learning, radiative transfer, solar-induced fluorescence.**

I. Introduction

REMOTELY sensed optical data collected by airborne or space-based Earth observation instruments inevitably carry the imprint of the atmosphere. These so-called atmospheric effects depend on wavelength range of interest, spectral resolution, and observation conditions, but usually they are substantial and cannot be ignored when analyzing the spectrometric data. Therefore, the processing of remote sensing data requires complex radiative transfer modeling in order to account or correct for atmospheric effects before surface-related observables can be derived. This step typically relies in part on specialized radiative transfer packages (such as MODTRAN [1], [2], libRadtran [3], [4], or 6S/6SV [5], [6], [7]), which are crucial in accurately modeling the light path through the atmosphere but are often too slow to be included in online correction or retrieval schemes. In many applications, the time-consuming radiative transfer calculations are performed offline and saved in look-up tables that can be swiftly accessed during correction or retrieval.

One such application is the retrieval of solar-induced fluorescence (SIF) emitted by plants between 650 and 800 nm during the photosynthetic process [8], [9]. SIF represents an important indicator of the efficiency of photosynthesis and can be used to monitor vegetation status. Thus, there is considerable interest in measuring SIF from remote platforms in order to study extended areas of vegetation. The SIF radiance is however much smaller than the surface reflected radiance [8], making its measurement particularly challenging even under favorable conditions. The relative strength of the SIF signal is significantly improved in regions of the spectrum where the reflected signal is reduced as in solar Fraunhofer lines or in atmospheric absorption bands. To resolve these features requires in turn a high spectral resolution instrument with sufficient signal-to-noise ratio. Despite the challenges, SIF has been successfully retrieved from selected airborne instruments both in solar Fraunhofer lines and in oxygen absorption bands [10], [11], [12] (see also [13]). Retrieval schemes such as the spectral fitting method [14], [15] make use of precomputed look-up tables based on offline radiative transfer computations. Because the latter are slow, the look-up tables are necessarily restricted to a limited range of atmosphere and observation geometries. This limitation could be efficiently addressed by machine learning (ML) models since they are well suited to speed up the forward simulation step and thereby provide a fast alternative to full-fledged radiative transfer modeling (see, e.g., [16], [17], [18]). Besides extending the scope of traditional SIF retrieval methods, an ML emulator can also be seamlessly integrated into the recent efforts to develop ML-based SIF retrieval schemes [19], [20], [21], [22].

Received 29 April 2024; revised 2 August 2024; accepted 27 August 2024. Date of publication 10 September 2024; date of current version 18 October 2024. *(Corresponding author: Miguel Pato.)*

Miguel Pato, Stefan Auer, Emiliano Carmona, Stefan Maier, and Rupert Müller are with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Oberpfaffenhofen, Germany (e-mail: Miguel.Figueiredo VazPato@dlr.de).

Jim Buffat, Patrick Rademske, and Uwe Rascher are with the Institute of Bio-and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany.

Kevin Alonso is with the Starion Group c/o European Space Agency, Largo Galileo Galilei, 00044 Frascati, Italy.

Hanno Scharr is with the Institute of Advanced Simulations, IAS-8: Data Analytics and Machine Learning, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany (e-mail: h.Scharr@fz-juelich.de).

Digital Object Identifier 10.1109/JSTARS.2024.3457231

The goal of the present work is to develop a physics-based ML emulator of at-sensor radiances in the vicinity of the O_2-A absorption band ($\lambda = 740 - 780$ nm) for use in SIF retrieval methods. The choice of the O_2-A band is motivated by the significant depth of the absorption feature and its wide use for SIF retrieval. The narrow range also helps reducing the complexity of the problem and enables the generation of large datasets of simulated spectra. In the future, the O_2-B absorption band (around 687 nm) and other spectral ranges should also be considered. We focus on two representative instruments currently in operation, namely the DESIS spectrometer [23], [24] operating on the International Space Station since 2018 and the airborne HyPlant spectrometer [12] which is the demonstrator for the FLEX space mission. The proposed emulator can also be generalized to other sensors in a straightforward manner.

The article is organized as follows. Section II introduces the simulated data and methodology adopted, while Section III reports the performance of all tested physics-based ML emulators and explores strategies to improve upon our proposed ML model. Section IV discusses the relevance of our results and the concluding remarks are given in Section V. Note that this manuscript is an extended and improved version of our previous work presented in [25]. In particular, we use here newly simulated datasets (see Section II-A), compare the emulator to data acquired by the HyPlant instrument (see Section IV), and implement emulator variants with improved accuracy performance (see Section III-C).

II. Data and Methods

A. Simulated Data

The first step in developing our physics-based ML emulator is to assemble a body of simulated at-sensor radiance spectra capturing a meaningful range of atmosphere conditions, observation geometries, surface properties and sensor characteristics. For our purposes, the at-sensor radiance signal has contributions due to light scattered by the atmosphere as well as light reflected and emitted by the surface [26]:

$$L_s = L_p + \frac{E_g^0 \rho T^\uparrow}{\pi(1-\rho S)} + L_\mathrm{F} T^\uparrow \tag{1}$$

where L_p is the path radiance, E_g^0 is the global solar irradiance on the ground, $T^\uparrow$ is the total transmission coefficient from surface to sensor (comprising direct and diffuse components, $T^\uparrow = T^\uparrow_\mathrm{dir} + T^\uparrow_\mathrm{dif}$), and S is the spherical albedo of the atmosphere. The surface reflectance and SIF on-ground radiance are represented by ρ and L_F, respectively. The at-sensor radiance spectrum in sensor resolution is then obtained by convoluting the above expression with the appropriate spectral response functions. It is noteworthy that the atmospheric functions L_p, E_g^0, $T^\uparrow_\mathrm{dir}$, $T^\uparrow_\mathrm{dif}$, and S depend solely on atmosphere and geometry and can thus be computed separately and applied for different surface and sensor configurations. Accordingly, a two-module software tool was developed to simulate at-sensor radiance spectra. The atmosphere and geometry parameters are passed to the first module, which derives the atmospheric functions L_p, E_g^0, $T^\uparrow_\mathrm{dir}$, $T^\uparrow_\mathrm{dif}$, and S at very high spectral resolution. We modeled radiative transfer through the atmosphere with the line-by-line algorithm of MODTRAN6 [2], but other radiative transfer software can also be used in our generic approach. The surface and sensor properties are handled by the second module of our tool, which outputs the at-sensor radiance spectrum in sensor resolution for any given atmosphere, geometry, surface, and sensor configuration.

The developed simulation tool is generic and can be used for virtually any optical remote sensing application. In this work, the tool is employed for the simulation of DESIS and HyPlant at-sensor radiances between 740 and 780 nm in view of SIF retrieval applications in the O_2-A absorption band. The modeling of surface and sensor properties is tailored for our specific case. The surface reflectance spectrum is parameterized as a second-order polynomial, while the SIF emission spectrum is taken as a Gaussian of variable normalization F_{737} at $\lambda = 737$ nm and with 20 nm standard deviation. Using real data acquired by DESIS and HyPlant, these models were found to be realistic for different vegetation and soil land covers in the wavelength range around the O_2-A band, but are not intended for other applications or spectral regions. In addition, the DESIS and HyPlant sensor properties were carefully considered based on expert knowledge of their performance and calibration over recent years. With typical spectral sampling distances of 2.55 nm for DESIS and 0.11 nm for HyPlant, there are 13 DESIS and 349 HyPlant spectral bands in the range $\lambda = 740 - 780$ nm with full width at half maximum (FWHM) of approximately 3.5 nm for DESIS and 0.24 nm for HyPlant. For detailed instrument specifications, please refer to [12], [23], [24]. We characterize the pixels of each band by an average Gaussian spectral response function and allow for small additive changes of central wavelength (CW) and FWHM. In this way, the simulated spectra for DESIS and HyPlant are very realistic and closely follow the spectral performance of the instruments.

We aim at simulating at-sensor radiances precisely enough to be useful to retrieve the weak SIF signal. For typical SIF on-ground outputs of $F_{737} = 0.1 - 0.4$ mW/cm^2/sr/μm at $\lambda =$ 737 nm, the corresponding SIF at-sensor radiance amounts to $0.01 - 0.4$ mW/cm^2/sr/μm in the range $\lambda = 740 - 780$ nm. This sets the sensitivity goal for our simulations. We therefore conducted a sensitivity analysis with the objective of pinning down all parameters of interest for our case. The resulting set of parameters define atmosphere (water vapor content H_2O, aerosol optical thickness at 550 nm AOT $_{550}$), geometry (tilt angle TA, sun zenith angle SZA, relative azimuth angle RAA, ground altitude h_gnd, sensor altitude h_sen), surface (reflectance at 740 nm ρ_{740}, reflectance slope at 740 nm s, ratio of reflectance slopes at 780 and 740 nm e) and sensor (CW shift δ_CW, FWHM change δ_FWHM) properties. The corresponding ranges are listed in Table I, where the cases of DESIS and HyPlant were treated separately given their operation specificities. For clarification, the tilt angle is the angle between the line-of-sight and the nadir direction and is essentially equal to the view zenith angle in

TABLE I
SPECIFICATION OF THE DESIS AND HYPLANT SIMULATED DATASETS

Specification		DESIS	HyPlant
Atmosphere	H_2O [cm]	0.3–5.0	0.3–3.0
	AOT_{550} []	0.02–0.30	0.02–0.30
Geometry	TA [°]	0–25	0–20
	SZA [°]	0–55	20–55
	RAA [°]	0–180	0–180
	h_{gnd} [m]	0–600	0–300
	h_{sen} [km]	100	0.659–0.691
Surface	ρ_{740} []	0.05–0.60	0.05–0.60
	s [nm^{-1}]	0–0.012	0–0.012
	e []	0–1	0–1
	F_{737}/F_0	0–0.8	0–0.8
Sensor	δ_{CW} [nm]	$[-1.75, +1.25]$	$[-0.080, +0.080]$
	δ_{FWHM} [nm]	$[-0.3, +0.3]$	$[-0.040, +0.040]$
Input dimensions		12	13
Number of bands		13	349
Number of samples		4.3×10^6	6.3×10^6
- uniform grid		1×10^6	3×10^6
- random		3×10^5	3×10^5
- Halton		3×10^6	3×10^6

our cases of interest. Note that in the table the sensor altitude is given with respect to the ground in the case of HyPlant and the SIF on-ground radiance is expressed in terms of $F_0 = 1$ mW/cm^2/sr/μm. Beyond the parameters specified in Table I, we adopted throughout the mid-latitude summer atmosphere, the rural aerosol model, an ozone content of 332 DU, and the TSIS-1 solar model [27].

Table I effectively defines the input parameter space for our simulated dataset, encompassing 12 dimensions for DESIS and 13 for HyPlant. It is crucial to uniformly cover these high-dimensional spaces in order to appropriately represent the influence of atmosphere, geometry, surface and sensor properties on the at-sensor radiance spectra. Uniform grid sampling, frequently used for the generation of look-up tables, is not efficient for more than a few dimensions leaving large chunks of the parameter space unexplored. The random and Halton [28], [29] methods provide much more uniform sampling in high dimensions with a limited number of samples (for other methods, see, e.g., [30]). Therefore, we sampled the input space with a combination of the uniform grid, random, and Halton methods and then ran the simulation tool to generate the at-sensor radiance spectra. The key specifications of the DESIS and HyPlant simulated datasets are detailed in Table I; a full account of the simulations used here is presented in [31]. For later comparison, the simulation of a single spectrum typically takes 1–4 min depending on the input configuration and this time is dominated by the radiative transfer part. The simulated datasets, generated using ten cores on a dedicated virtual machine (Intel(R) Xeon(R) Gold 6132 CPU @ 2.60 GHz, 64 GB RAM) and leveraging the decoupling of the two modules of the simulation tool, contain 4.3 million DESIS and 6.3 million HyPlant spectra. This body of simulated data constitutes the basis for training a physics-based ML emulator.

B. Methodology

The influence of atmosphere, geometry, surface, and sensor properties on the at-sensor radiance spectrum can be learned by a ML model trained on the simulated datasets presented in the previous section. In fact, this constitutes a regression problem in multiple dimensions:

$$L_s = F(x) \tag{2}$$

where the feature vector $x \in \mathbb{R}^p$ contains the parameters identified in Table I, the target vector $L_s \in \mathbb{R}^b$ is the at-sensor radiance spectrum in sensor resolution, and $F : \mathbb{R}^p \to \mathbb{R}^b$ is the mapping function to be learned ($F : \mathbb{R}^{12} \to \mathbb{R}^{13}$ for DESIS, $F : \mathbb{R}^{13} \to \mathbb{R}^{349}$ for HyPlant). Given the sheer number of input dimensions in our problem, simple models for F can capture a significant degree of complexity while being very fast at evaluation. With this observation in mind, we attempted to learn $F(x)$ with basic benchmark models, namely linear functions, polynomials, and shallow neural networks. For clarity and later reference, a polynomial of degree d is defined as

$$P_d(x) = \sum_{\mathbf{k} \in \mathcal{K}_d} w_{\mathbf{k}} \prod_{i=1}^{p} x_i^{k_i}$$
$$\mathcal{K}_d = \left\{ \mathbf{k} \in \mathbb{N}^p : \sum_{i=1}^{p} k_i \leq d \right\} \tag{3}$$

with $w_{\mathbf{k}} \in \mathbb{R}^b$ the polynomial coefficients. For the models above, we use least squares as loss function. The case of a linear function (polynomial of degree $d = 1$) is simply ordinary least squares (OLS). In addition, kernel ridge regression (KRR), Gaussian process regression (GPR), support vector regression, and k nearest neighbors were also applied to learn $F(x)$. No attempt was made to realize more complex ML models (including deep learning). This choice is justified by the very encouraging results obtained with the very simple ML models mentioned above (see Section III). A comparative analysis of the different models studied here with traditional look-up table interpolation methods is a relevant line of research that is left for future work. Note however that, unlike interpolation methods, the emulators studied here are very lightweight in terms of memory usage.

A fast and accurate emulator of at-sensor radiances needs to be complex enough to handle the range of simulated configurations and simple enough to be computationally efficient. It turns out that basic ML models are very effective in our application case of simulating DESIS and HyPlant data around the O_2-A absorption band. We trained and assessed in detail several benchmark ML models: OLS, second-, fourth- and fifth-order polynomials (P2, P4, P5) and neural networks with 128 × 64 (N2) and 128 × 64 × 32 (N3) hidden nodes, batch size 16, and rectified linear unit activations. The polynomial models were trained with ridge regression, while for the neural network stochastic gradient descent was used. KRR and GPR were also considered since they have shown good performance for the emulation task [17], [18], [20]. ML models based on support vector regression and k nearest neighbors (k=1,5,10) led to poor results in terms of

TABLE II
TIME AND ACCURACY PERFORMANCE OF DIFFERENT ML MODELS TRAINED ON THE DESIS AND HYPLANT SIMULATED DATA

Performance parameter	**DESIS**							
	OLS	**P2**	**P4**	**P5**	**N2**	**N3**	**KRR**	**GPR**
Test set MAE [mW/cm^2/sr/μm]	0.77	0.17	0.011	0.0032	0.040	0.037	0.043	0.039
Total training time	1.6 s	19 s	1.8 min	1.2 min	1.2 h	1.2 h	1.1 min	2.7 s
Prediction time per sample	0.06 μs	1.1 μs	11 μs	28 μs	56 μs	45 μs	0.1 ms	0.2 ms

Performance parameter	**HyPlant**							
	OLS	**P2**	**P4**	**P5**	**N2**	**N3**	**KRR**	**GPR**
Test set MAE [mW/cm^2/sr/μm]	0.59	0.092	0.0027	0.0013	0.027	0.020	0.018	0.025
Total training time	1.7 min	45 s	1.3 min	1.5 min	2.0 h	2.2 h	6.3 s	13 s
Prediction time per sample	1.3 μs	2.2 μs	17 μs	47 μs	44 μs	43 μs	72 μs	0.7 ms

accuracy and speed early on in our analysis and were accordingly disregarded for in-depth evaluation. The performance of each of the benchmark models was explored by considering a limited set of the relevant hyperparameters (including L2 norm penalty for OLS, P2, P4, P5, N2, and N3; hidden nodes, batch size, training epochs for N2 and N3; kernel function for KRR) in the validation dataset. Since the figures did not change dramatically by fine-tuning the hyperparameters, in the following we opted to focus on the simplest representative configuration for each model. In particular, the models shown all have zero L2 norm penalty.

The simulated datasets specified in Table I were split into training, validation, and test sets considering the properties of the different sampling techniques used. Uniform grid samples populate the input space very sparsely but include the borders, while Halton and random samples lie uniformly across the space but not at its borders. Ultimately, we are interested in realizing a physics-based ML emulator with good performance in the bulk of the input parameter space, not necessarily on its borders. The borders are however important for the training phase. Therefore, we assigned all uniform grid and Halton samples to the training set and split the random samples into validation and test sets. The training, validation, and test sets contain 4×10^6 (93.0%), 2.1×10^5 (4.9%), and 9×10^4 (2.1%) samples, respectively, for DESIS and 6×10^6 (95.2%), 2.1×10^5 (3.3%), and 9×10^4 (1.4%) samples, respectively, for HyPlant. Note that the baseline emulators presented in the next section do not require the use of the full training set to deliver excellent accuracy performance (see Section III-B for details). This should be taken into account when evaluating the apparent imbalance of the training, validation, and test sets.

The performance of each ML model was evaluated in terms of accuracy using the at-sensor radiance mean absolute error (MAE) and in terms of computational time using the prediction time per sample. These performance metrics are to be compared to the at-sensor radiance corresponding to a typical SIF output of $F_{737} = 0.4$ mW/cm^2/sr/μm and the simulation time of $1-4$ min per spectrum. The training, validation, and test phases for each model were conducted on the same standard virtual machine as mentioned in Section II-A using one core only. No effort was made to speed up the performance of the models by parallelizing the implementation.

III. RESULTS

A. Evaluation of Different Models

The test set accuracy, total training time, and prediction time per sample for all benchmark models are reported in Table II. Comparing the accuracy of the different models in terms of MAE with the typical SIF at-sensor radiances $0.04-0.4$mW/cm^2/sr/μm at $\lambda = 740 - 780$ nm (corresponding to a SIF on-ground radiance of $F_{737} = 0.4$ mW/cm^2/sr/μm at $\lambda = 737$ nm), it is clear that linear or quadratic functions (OLS, P2) are not complex enough to approximate the mapping function F and are therefore not appropriate for our purposes. The accuracies attained by N2, N3, KRR, and GPR are significantly better, but still of the order of the small fluorescence signal we aim to characterize. The results shown for KRR correspond to a third-order polynomial kernel; similar but slightly worse results were obtained when using a radial basis function kernel. Note that limited effort was put into exploring all possibilities for neural networks, KRR, and GPR. It is well possible that better results can be obtained with these models by fine-tuning their configuration and/or hyperparameters. Fourth-order polynomials seem to be adequate to represent the interplay between simulated data and atmosphere, geometry, sensor, and surface parameters. In fact, with an MAE of 0.011 mW/cm^2/sr/μm for DESIS and 0.0027 mW/cm^2/sr/μm for HyPlant, the P4 model provides an average accuracy well below typical SIF at-sensor radiances. Higher order polynomials perform even better at the cost of prediction time as illustrated by the P5 results in Table II, but we decided to use P4 as our baseline model since it is faster and its performance is already excellent. Our results suggest in particular that in our specific case deep learning models are not required to realize a very accurate ML emulator.

The training phase is relatively fast for all models with total training times of seconds for OLS, P2, and GPR, minutes for P4, P5, and KRR and hours for N2 and N3. Note that due to memory limitations some models were trained on a subset of the training data. The neural networks were trained on CPU,

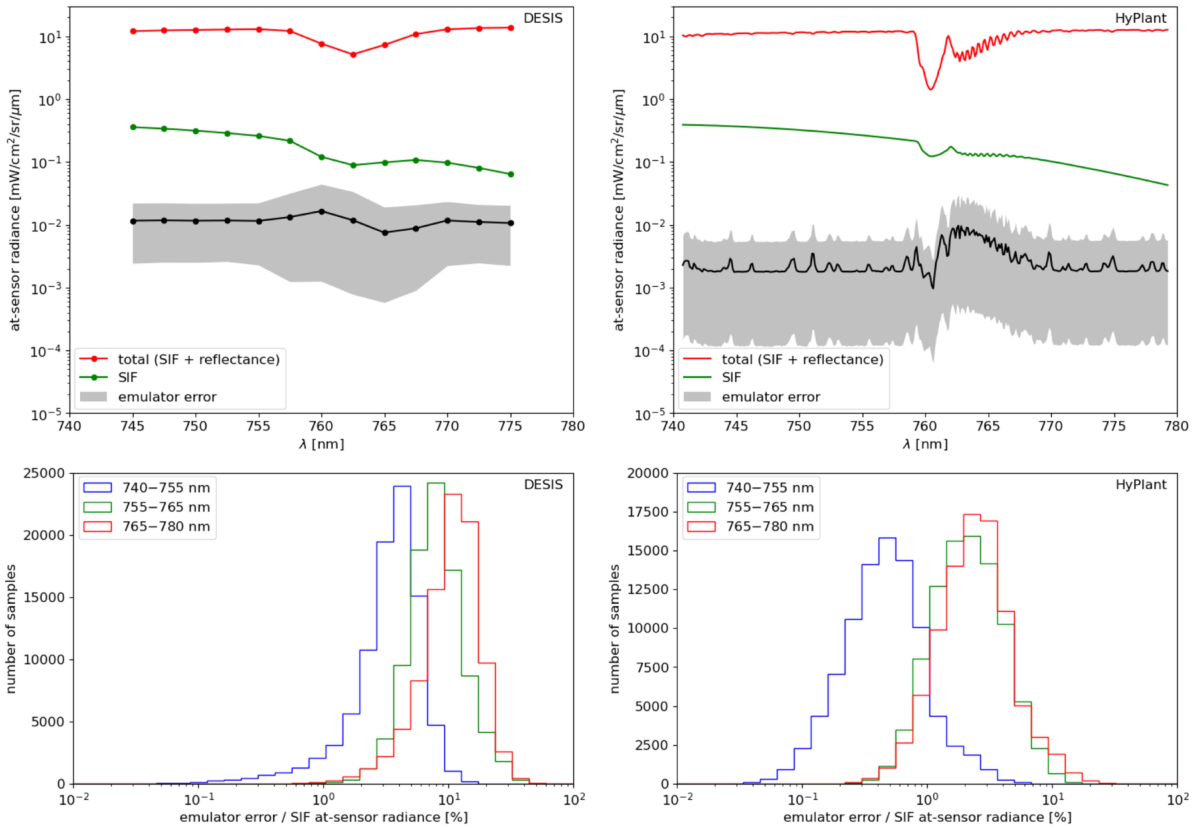


Fig. 1. Accuracy performance of the baseline P4 emulators trained on the DESIS (left) and HyPlant (right) simulated data. The top panels show the mean (black) and 5th/95th percentiles (gray) of the absolute error per band in the test set in direct comparison to the test set average total at-sensor radiance (red) and the SIF at-sensor radiance corresponding to $F_{737} = 0.4$ mW/cm^2/sr/μm (green). The distribution of the absolute error in the test set is plotted in the bottom panels relative to the SIF at-sensor radiance (top, green) for the spectral ranges around the O_2-A absorption band: $\lambda = 740 - 755$ nm (blue), $\lambda = 755 - 765$ nm (green), and $\lambda = 765 - 780$ nm (red).

so their time performance can likely be improved by the use of GPUs. Except for OLS, the times are similar for the cases of DESIS and HyPlant and do not scale with the number of output features (13 bands for DESIS, 349 for HyPlant). It is fair to point out that the total training time is not critical since training just occurs once; one can easily afford a longer training if a better accuracy can be achieved (see Section III-C for specific examples). More important is the prediction time per sample spectrum, also reported in Table II and computed as the mean time per spectrum when predicting on the training, validation, and test sets. OLS and P2 predictions take around 2 μs or less, while N2 and N3 prediction times are of order 50 μs (using CPU only). KRR and GPR prediction times are slightly higher at about 0.1 ms. But these models are not particularly useful for our purposes due to their reduced accuracy as shown above. The prediction time per spectrum for the proposed P4 model is of order 10–20 μs. This is about 10^7 times faster than the simulation, which takes roughly 1–4 min. The extreme speed up can be explained by the simplicity of the polynomial model used and the fact that prediction involves only matrix multiplications, which can be made very fast on a normal CPU. Note as well that for these models the predictions can be performed in bulk (e.g., for the whole test set). The sequential prediction of one spectrum at a time might differ from the values in Table II.

B. Baseline Emulator

It is clear from the discussion above that our best ML emulator is the fourth-order polynomial P4. The errors per band for this baseline model are shown in Fig. 1 separately for the DESIS and HyPlant cases. The upper plots compare the average total at-sensor radiances in the test set (red), the at-sensor radiance for a typical SIF signal with $F_{737} = 0.4$ mW/cm^2/sr/μm (green), and the P4 model error in the test set (black and gray) across the wavelength range $\lambda = 740 - 780$ nm. The mean test set error is represented by the black line, while the 5th and 95th percentiles are marked by the gray band and encompass the interval $5 \times 10^{-4} - 5 \times 10^{-2}$ mW/cm^2/sr/μm for DESIS and $10^{-4} - 10^{-2}$ mW/cm^2/sr/μm for HyPlant. Such error levels correspond to $\lesssim 0.1\%$ of the total at-sensor radiance, attesting the excellent accuracy of the P4 model. As evident from the bottom panels of Fig. 1, the model error is consistently at or below 10% of the typical SIF signal with $F_{737} = 0.4$ mW/cm^2/sr/μm for different wavelength ranges. Note in particular that there are virtually no test samples for which the error exceeds the typical SIF signal. The model performs similarly at and above the O_2-A band ($\lambda = 755 - 765$ nm and $\lambda = 765 - 780$ nm) and slightly better below the absorption band ($\lambda = 740 - 755$ nm). This is because the absolute errors are mostly constant across the

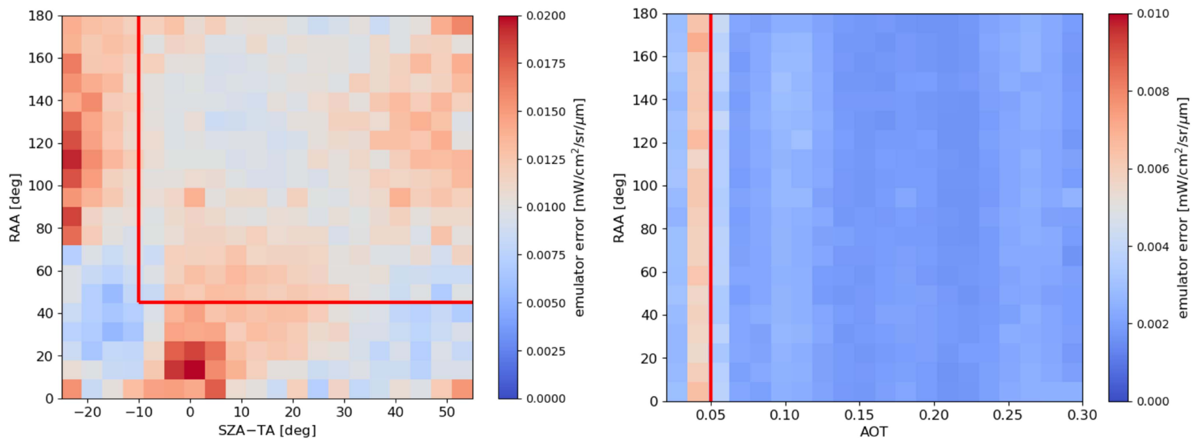


Fig. 2. Dependence of the accuracy of the baseline P4 emulator on observation conditions for DESIS (left) and HyPlant (right). The plots color-code the band-averaged absolute error in the training set as a function of SZA − TA and RAA for DESIS and of AOT_{550} and RAA for HyPlant. The regions to the left of the red lines encompass the largest errors and were masked out for the training and test of the P4A emulator variant.

spectral range considered (cf., upper panels of Fig. 1), while the SIF signal gradually decreases toward larger wavelengths. The main message conveyed by our results is that the P4 emulator is very accurate for the purposes of simulating the feeble SIF signal in the O_2-A band for DESIS and HyPlant.

For completeness, the performance of the baseline emulator is evaluated for different training set sizes. It turns out the P4 model is very accurate even when trained on small datasets. In fact, the accuracy of the emulator is the same or better than in Table II down to 5×10^3 training samples for DESIS and 5×10^4 training samples for HyPlant. For instance, using 5×10^4 randomly selected samples from the full training set, the test set MAE of the P4 model is 0.0061 mW/cm^2/sr/μm for DESIS (better than previously) and 0.0027 mW/cm^2/sr/μm for HyPlant (the same as previously). In this context, our test set with 9×10^4 samples, although a small fraction of the whole simulated dataset, is appropriate for testing the benchmark models. In addition, the results suggest that a significantly smaller training dataset would be sufficient for our particular application with the corresponding speed up of the simulation step and training of the emulator.

The performance of the P4 model is excellent across the whole input parameter space, but it is not uniform. In fact, there are parts of the parameter space where the emulator performs slightly worse than on average. This is illustrated in Fig. 2 for DESIS and HyPlant. In the case of DESIS, the most problematic region is related to sun backscatter geometries, where sun, sensor, and target are roughly aligned with the sun behind the sensor (i.e., SZA $\simeq$ TA and RAA $\simeq 0°$). These configurations lead to an increase in sunlight backscattered by aerosols and are typically avoided if possible. For HyPlant, the problematic geometry does not occur since SZA $\geq$ TA (cf., Table I), but the largest residuals arise instead for a narrow range of small AOT $_{550}$ values. In both DESIS and HyPlant cases, the highest errors are very small (of order 0.02 mW/cm^2/sr/μm) and simply indicate where the selected model has a numerical difficulty in representing the simulated data.

C. Accuracy Improvement Strategies

Despite the encouraging findings presented above, it is important to understand the limitations of our proposed P4 emulator and explore ways to improve its performance. In the following, we consider variants of the P4 model with a restricted input space (P4A), bandwise learning (P4B), and atmospheric functions learning (P4C). The performance of the three variants is presented in Table III and Fig. 3.

We start by considering a restriction of the input space as a means to improve the performance of the P4 emulator. In fact, the nonuniform performance shown in Fig. 2 suggests that a better accuracy can be attained by partitioning the input parameter space and developing different models for each partition. In order to assess this possibility, we excluded the regions to the left of the red lines in Fig. 2, namely RAA $< 45°$, SZA $-$ TA $< -10°$ for DESIS and $AOT_{550} < 0.05$ for HyPlant, and retrained and tested a dedicated fourth-order polynomial model (P4A). As can be seen in Table III and Fig. 3, the P4A model improves upon the baseline P4 model on average by 21% for DESIS and 37% for HyPlant while being similarly fast. Although modest, the improvement is significant and illustrates how this low-effort strategy can boost the performance of a model in specific portions of the parameter space. Dedicated emulators for the masked problematic regions could also be developed if needed. Note that the trained emulators are only strictly valid in the corresponding input spaces and should not be used to interpret real data with parameters outside that space. This is especially important due to the polynomial nature of models which makes extrapolations beyond the input parameter space particularly prone to errors.

Next we explore the possibility to train our emulator bandwise. The formulation of the regression problem in Section II-B includes a multidimensional target (the at-sensor radiance spectrum) and implies the learning of all spectral bands in the same model. Alternatively, it is also possible to define a separate regression with a unidimensional target for each band. This

TABLE III
TIME AND ACCURACY PERFORMANCE OF VARIANTS OF THE BASELINE ML EMULATOR

Performance parameter	**DESIS**				**HyPlant**			
	P4	**P4A**	**P4B**	**P4C**	**P4**	**P4A**	**P4B**	**P4C**
Test set MAE [mW/cm^2/sr/μm]	0.011	0.0087	0.011	0.0037	0.0027	0.0017	0.0027	0.0014
Total training time	1.8 min	1.9 min	24 min	2.1 s	1.3 min	1.4 min	8.6 h	3 s
Prediction time per sample	11 μs	11 μs	0.1 ms	6 ms	17 μs	17 μs	4 ms	173 ms

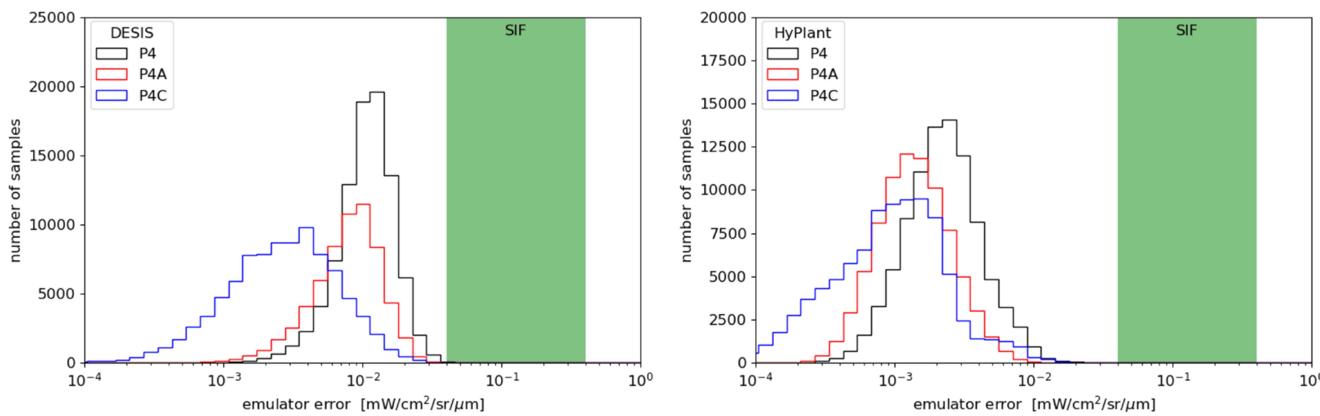


Fig. 3. Error distribution for the baseline P4 emulator and its variants for DESIS (left) and HyPlant (right). The histograms show the band-averaged absolute error in the test set for the baseline P4 model (black) and variants P4A (red, restricted input space) and P4C (blue, atmospheric functions learning). The results for the variant P4B (bandwise learning) are very similar to the ones of the baseline P4 model and are not shown for clarity. The green band indicates the SIF at-sensor radiance values corresponding to a typical SIF on-ground radiance of $F_{737} = 0.4$ mW/cm^2/sr/μm at $\lambda = 737$ nm.

approach decouples the different bands and enables the best possible fitting for each band at the price of additional computation time. The performance of a bandwise fourth-order polynomial model (P4B) is documented in Table III. As expected, training takes several times longer than for the baseline P4 model (especially for HyPlant given the large numbers of bands) and, more importantly, the prediction time per sample spectrum is 0.1 ms for DESIS and 4 ms for HyPlant, about 10 and 200 times slower than P4, respectively. Nevertheless, there is virtually no gain in accuracy of P4B with respect to P4 with MAE changes appearing only after the third significant digit. The results suggest that the overall regression with a multidimensional target used in P4 is very effective for both DESIS and HyPlant and it does not pay off to implement bandwise learning in our specific case.

Finally, we attempt to directly learn the atmospheric functions $L_p, E_g^0, T^{\uparrow}$, and S at very high spectral resolution instead of the sensor-resolution simulated data L_s [cf., (1)]. In other words, our regression problem is no longer as in (2) but it now reads

$$\left(L_p, E_g^0, T^{\uparrow}, S\right) = F(x)\,. \tag{4}$$

The crucial difference is that the new input space of $F(x)$ has a lower dimensionality. In fact, as explained in Section II-A, the atmospheric functions depend only on atmosphere and geometry parameters and can be used with any surface and sensor models for the derivation of the at-sensor radiance spectrum. Therefore, learning the atmospheric functions involves a regression function $F(x)$, where x contains only atmosphere and geometry parameters as opposed to the case of learning the simulated at-sensor radiance spectra in (2) where x contains all atmosphere, geometry, surface, and sensor parameters. This reduces significantly the dimensionality of the input parameter space: from 12 to 6 dimensions for DESIS and from 13 to 7 dimensions for HyPlant (cf., Table I). We trained a fourth-degree polynomial to learn the atmospheric functions (P4C) and combined it with a numerical module to factor in the surface and sensor properties as in (1). The performance of P4C model is detailed in Table III and Fig. 3. The training phase is now much quicker, because the dimensionality and size of the training set are much reduced with respect to the case of the P4 model. The prediction time is approximately 6 ms for DESIS and 173 ms for HyPlant, about three to four orders of magnitude slower than for P4. However, the average error in the test set is decreased by a factor of 3 to 0.0037 mW/cm^2/sr/μm for DESIS and by a factor of 2 to 0.0014 mW/cm^2/sr/μm for HyPlant. This suggests that the learning of atmospheric functions constitutes a promising strategy to realize a more accurate emulator at the price of a somewhat longer prediction. Whether a more accurate but slower ML-based emulator is useful for retrieval schemes depends on the specific application.

IV. DISCUSSION

In order to assess how well our emulators match observational data, we perform an unconstrained least-squares optimization

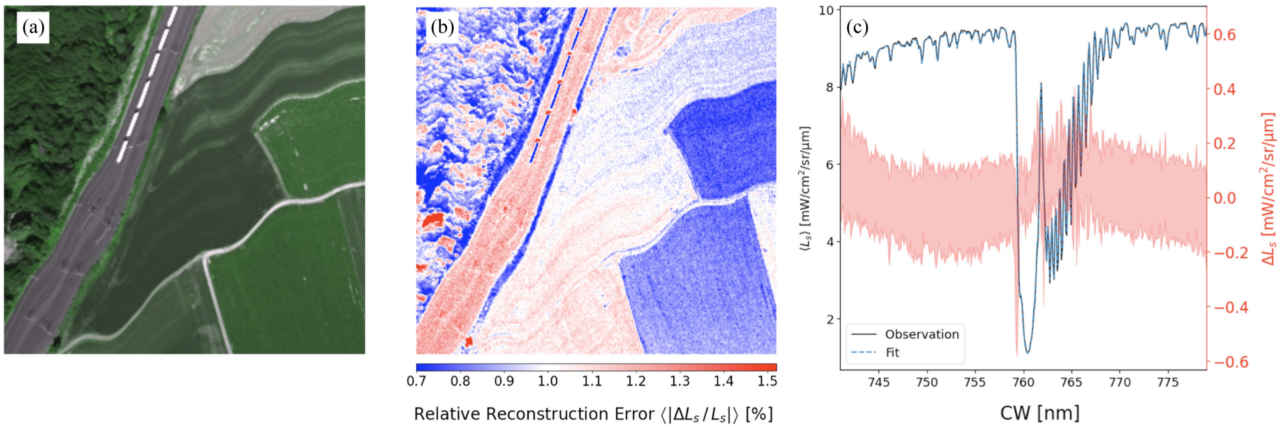


Fig. 4. Reconstruction of HyPlant spectra with the baseline P4 emulator model. The false color composite of a subset of a HyPlant acquisition (2018/06/26, 15:30 CEST, Jülich DE) is shown in (a), while the spectrally averaged relative reconstruction error is displayed in (b). Panel (c) plots the mean observed (black) and fitted (blue) at-sensor radiance within the subset as well as the 1th/99th percentiles of the reconstruction residuals (red).

of the input parameters of P4 on individual pixels of a HyPlant acquisition (cf., Fig. 4). The acquisition was recorded under optimal weather conditions such that its atmospheric state can be assumed to be covered by the simulation input ranges. Furthermore, we could infer geometry parameters from a digital elevation model and the sensor attitude recorded during acquisition which were covered by the simulation input ranges as well. Accordingly, the optimization was restricted to the atmosphere, surface, and sensor parameters (cf., Table I). We find the spectrally averaged reconstruction error to be $\langle|\Delta L_s|\rangle <$ 0.25 mW/cm^2/sr/μm, which corresponds to a relative reconstruction error below 1.4%. It is clear from Fig. 4 that this performance holds across different surface classes commonly found in HyPlant acquisitions (e.g., crops, forested areas, man-made structures, and bare soil). Retrieval of physical parameters from spectral data is ill-posed due to the confounding nature of the parameters affecting the at-sensor radiance. As a consequence, the parameters derived as the optimal parametrization to P4 in each pixel cannot be trusted to approximate well the physical quantities underlying the at-sensor radiance signal generation. However, good reconstruction performance is a necessary precondition for emulators such as P4 to be used in various SIF retrieval approaches. Fig. 4 demonstrates the high accuracy of the baseline P4 emulator when compared to real data from HyPlant and therefore its usefulness for integration into future SIF retrieval methods.

We conclude our discussion by commenting on an additional use of the learned atmospheric functions in the P4C emulator. Since the atmospheric functions are independent of surface and sensor properties (cf., Section II-A), the P4C model effectively opens the possibility to generalize the simulation to any reflectance and fluorescence models and any putative sensor. This point is illustrated explicitly by simulating data with spectral sampling distances and spectral resolutions resembling the near future FLEX mission in the vicinity of the O_2-A absorption band. Loosely based on the specifications for the FLORIS instrument [32] and the learned P4C model, we show in Fig. 5 several simulated at-sensor radiance spectra between 740 and 780 nm for different observation conditions. It is important to emphasize that ours is not a high-fidelity simulation of the performance of the FLORIS instrument, which covers not only the O_2-A band but also the O_2-B band in high resolution and a considerably larger spectral range in lower resolution. In particular, the emulators developed in our work cover a limited spectral range ($\lambda = 740 - 780$ nm) and cannot be employed directly as emulators for the FLEX mission. However, the results in Fig. 5 serve to illustrate how our emulator can be used and extended for sensors other than HyPlant and DESIS. The ability to realistically simulate at-sensor radiance spectra will become crucial to interpret the data collected by FLEX and other missions and ultimately retrieve the underlying SIF spectrum.

V. Conclusion

Physics-based ML emulators have the potential to speed up radiative transfer modelling and thereby widen the reach of retrieval methods in varied remote sensing applications. Our work follows ongoing efforts in the community toward this direction and focuses on the case of SIF retrieval in the O_2-A absorption band. We deliver in particular a simple ML emulator of at-sensor radiances which is both fast and accurate enough to characterize the SIF signal in the DESIS and HyPlant spectrometers. Several strategies are analyzed to improve even more the accuracy of the emulator in exchange for lower speeds at prediction time. The results are not only relevant for DESIS and HyPlant, but are easily transferable to other sensors, including FLEX. A relevant line of research for future work is to extend our emulators to other spectral ranges, especially the O_2-B absorption band. This opens up the prospect of simulating large amounts of realistic spectrometer data quickly and integrating a fast simulation step into SIF retrieval methods. Both possibilities will be instrumental in bringing the power of machine learning to SIF studies over the coming years. In the wider picture, our work constitutes an illustrative example on the use of ML to speed up physics-based calculations in remote sensing applications.

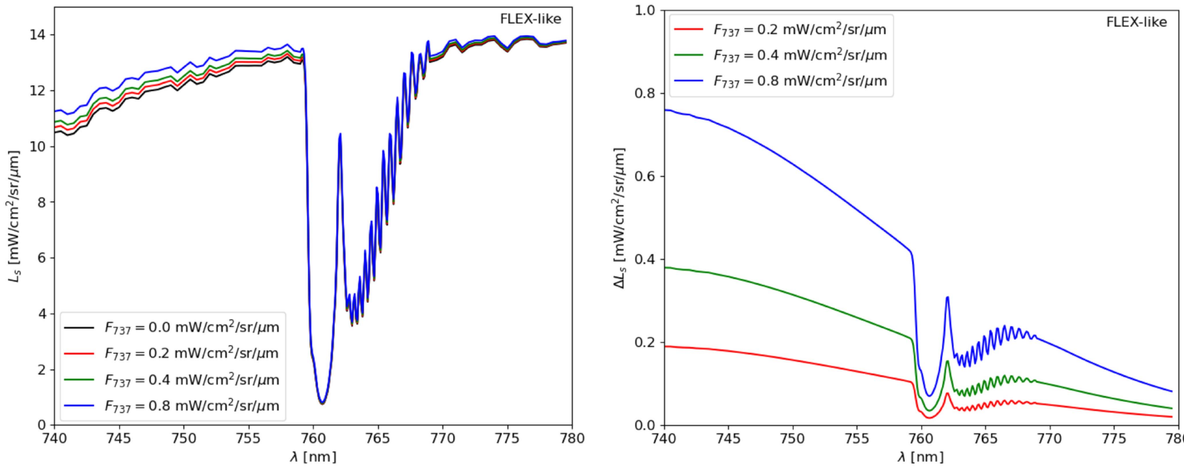


Fig. 5. Simulated at-sensor radiance spectra around the O_2-A absorption band for a FLEX-like instrument. The simulated spectra were obtained with the P4C emulator model for a space-based nadir observation with sun zenith angle of 30°, ground altitude of 0 m, aerosol optical thickness at 550 nm of 0.1, water vapor content of 2.0 cm, a slowly rolling reflectance spectrum ($\rho_{740} = 0.3$, $s = 0.006$ nm^{-1}, $e = 0$) and different SIF on-ground radiances ($F_{737} = 0.0, 0.2, 0.4, 0.8$ mW/cm^2/sr/μm). The spectral response function of the putative FLEX-like instrument was modelled roughly following the specifications of FLORIS [32] in the range $\lambda = 740 - 780$ nm. The left panel shows the at-sensor radiance spectra for different SIF radiances, while the right panel displays the at-sensor radiance differences with respect to the case without fluorescence.

Acknowledgment

This work was supported by the project "FluoMap" (Impulsfonds-Förderkennzeichen ZT-I-PF-5-12) funded by the Helmholtz Initiative and Networking Fund, Helmholtz AI, Deutsches Zentrum für Luft-und Raumfahrt and Forschungszentrum Jülich GmbH. The authors would like to acknowledge the computing time granted by the JARA Vergabegremium and provided on the JARA Partition part of the supercomputer JURECA [33] at Forschungszentrum Jülich.

References

[1] A. Berk, P. Conforti, R. Kennett, T. Perkins, F. Hawes, and J. van den Bosch, "MODTRAN6: A major upgrade of the MODTRAN radiative transfer code," in *Proc. SPIE*, Jun. 2014, Art. no. 90880H.

[2] A. Berk, P. Conforti, and F. Hawes, "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*, May 2015, Art. no. 947217.

[3] C. Emde et al., "The libRadtran software package for radiative transfer calculations (version 2.0.1)," *Geosci. Model Develop.*, vol. 9, no. 5, pp. 1647–1672, 2016. [Online]. Available: https://gmd.copernicus.org/articles/9/1647/2016/

[4] B. Mayer and A. Kylling, "Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use," *Atmos. Chem. Phys.*, vol. 5, no. 7, pp. 1855–1877, 2005. [Online]. Available: https://acp.copernicus.org/articles/5/1855/2005/

[5] E. Vermote, D. Tanre, J. Deuze, M. Herman, and J.-J. Morcette, "Second simulation of the satellite signal in the solar spectrum, 6S: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 675–686, May 1997.

[6] S. Y. Kotchenova, E. F. Vermote, R. Matarrese, and F. J. Klemm, "Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance," *Appl. Opt.*, vol. 45, no. 26, pp. 6762–6774, Sep. 2006. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-45-26-6762

[7] S. Y. Kotchenova and E. F. Vermote, "Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part II. Homogeneous Lambertian and anisotropic surfaces," *Appl. Opt.*, vol. 46, no. 20, pp. 4455–4464, Jul. 2007. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-46-20-4455

[8] G. H. Mohammed et al., "Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress," *Remote Sens. Environ.*, vol. 231, 2019, Art. no. 111177.

[9] A. Porcar-Castell et al., "Chlorophyll a fluorescence illuminates a path connecting plant molecular biology to earth-system science," *Nature Plants*, vol. 7, pp. 998–1009, 2021.

[10] U. Rascher et al., "Sun-induced fluorescence–A new probe of photosynthesis: First maps from the imaging spectrometer HyPlant," *Glob. Change Biol.*, vol. 21, no. 12, pp. 4673–4684, 2015.

[11] C. Frankenberg et al., "The chlorophyll fluorescence imaging spectrometer (CFIS), mapping far red fluorescence from aircraft," *Remote Sens. Environ.*, vol. 217, pp. 523–536, 2018.

[12] B. Siegmann et al., "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.*, vol. 11, no. 23, 2019, Art. no. 2760.

[13] S. W. Maier, K. P. Günther, and M. Stellmes, *Sun-Induced Fluorescence: A New Tool for Precision Farming*. Madison, WI, USA: American Society of Agronomy, 2003, pp. 209–222.

[14] S. Cogliati et al., "Retrieval of sun-induced fluorescence using advanced spectral fitting methods," *Remote Sens. Environ.*, vol. 169, pp. 344–357, 2015.

[15] S. Cogliati et al., "A spectral fitting algorithm to retrieve the fluorescence spectrum from canopy radiance," *Remote Sens.*, vol. 11, no. 16, 2019, Art. no. 1840.

[16] J. V. Servera et al., "Gradient-based automatic lookup table generator for radiative transfer models," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1040–1048, Feb. 2019.

[17] J. Vicent Servera, L. Martino, J. Verrelst, and G. Camps-Valls, "Multifidelity Gaussian process emulation for atmospheric radiative transfer models," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5519210.

[18] J. Vicent Servera, L. Martino, J. Verrelst, J. P. Rivera-Caicedo, and G. Camps-Valls, "Multioutput feature selection for emulation and sensitivity analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5506411.

[19] R. Scodellaro et al., "A novel hybrid machine learning phasor-based approach to retrieve a full set of solar-induced fluorescence metrics and biophysical parameters," *Remote Sens. Environ.*, vol. 280, 2022, Art. no. 113196.

[20] J. Vicent Servera et al., "Systematic assessment of MODTRAN emulators for atmospheric correction," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4101917.

[21] J. Buffat et al., "Deep learning based prediction of sun-induced fluorescence from hyplant imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 2993–2996.

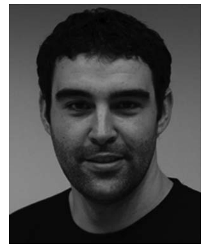
[22] J. Buffat et al., "A multi-layer perceptron approach for SIF retrieval in the O_2-A absorption band from hyperspectral imagery of the airborne hyplant sensor system," *Remote Sens. Environ.*, under review, 2024.
[23] D. Krutz et al., "The instrument design of the DLR earth sensing imaging spectrometer (DESIS)," *Sensors*, vol. 19, no. 7, 2019, Art. no. 1622. [Online]. Available: https://www.mdpi.com/1424-8220/19/7/1622
[24] K. Alonso et al., "Data products, quality and validation of the DLR earth sensing imaging spectrometer (DESIS)," *Sensors*, vol. 19, no. 20, 2019, Art. no. 4471.
[25] M. Pato et al., "Fast machine learning simulator of at-sensor radiances for solar-induced fluorescence retrieval with DESIS and hyplant," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 7563–7566.
[26] L. Guanter, R. Richter, and H. Kaufmann, "On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing," *Int. J. Remote Sens.*, vol. 30, no. 6, pp. 1407–1424, 2009.
[27] O. M. Coddington et al., "Version 2 of the TSIS-1 hybrid solar reference spectrum and extension to the full spectrum," *Earth Space Sci.*, vol. 10, no. 3, 2023, Art. no. e2022EA002637.
[28] J. Halton, "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals," *Numerische Mathematik*, vol. 2, pp. 84–90, 1960.
[29] A. B. Owen, "A randomized Halton algorithm in R," 2017. [Online]. Available: https://arxiv.org/abs/1706.02808
[30] D. G. Loyola, R. M. Pedergnana, and S. Gimeno García, "Smart sampling and incremental function learning for very large high dimensional data," *Neural Netw.*, vol. 78, pp. 75–87, 2016.
[31] M. Pato et al., "Simulated at-sensor radiance data for solar-induced fluorescence retrieval with DESIS and HyPlant in the O2-A absorption band," *Remote Sens. Environ.*, to be submitted, 2024.
[32] M. Drusch et al., "The fluorescence explorer mission concept-ESA's earth explorer 8," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1273–1284, Mar. 2017.
[33] P. Thörnig, "JURECA: Data centric and booster modules implementing the modular supercomputing architecture at Jülich supercomputing centre," *J. Large-Scale Res. Facilities*, vol. 7, 2021, Art. no. A182.

Miguel Pato received the M.Sc. degree in physics engineering from the Technical University of Lisbon, Lisbon, Portugal, in 2007, and the joint Ph.D. degree in physics from the University of Padua, Padua, Italy, and Paris Diderot University, Paris, France, in 2011.

From 2011 to 2016, he was a Postdoctoral Researcher in the field of astroparticle physics with the University of Zurich, Zurich, Switzerland, Technical University of Munich, Munich, Germany, and Stockholm University, Stockholm, Sweden. Since 2017, he has been with the Department of Photogrammetry and Image Analysis, Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany. He is part of the EnMAP processing and calibration team. His research focuses on hyperspectral image analysis.

Jim Buffat received the M.Sc. degree in geography with specialization in remote sensing from the University of Zurich, Zurich, Switzerland, 2021. He is currently employed at Forschungszentrum Jülich, Jülich, Germany, working toward the Ph.D. in Engineering (Dr.-Ing.) awarded by the Faculty of Agriculture at the University of Bonn, Bonn, Germany.

His thesis centers on deep learning based retrieval of sun-induced plant fluorescence from hyperspectral imagery. His research interests include hyperspectral remote sensing, image processing, and deep learning for environmental applications.



Kevin Alonso received the Technical Telecommunication Engineering bachelor's degree in communication systems and the Telecommunication Engineering master's degree in telematic systems from the Bilbao School of Engineering, Bilbao, Spain, in 2005 and 2007, respectively, and the Ph.D. (Dr.-Ing.) degree, with thesis Heterogeneous Data Mining of Earth Observation Archives: Integration and Fusion of Images, Maps, and In-situ Data, from the Technical University of Munich (TUM), Munich, Germany, in 2017.

From July 2008 to March 2013, he was as a Staff Researcher with Vicomtech-IK4, where he worked in the field of digital television and multimedia services, participating in the European Research Project RUSHES. During 2009 and 2010, he combined his research work on semantic web and geographic information systems (GIS) technologies with the study of the master's degree in space science and technology. From 2013 to 2022, he was with the Department of Photogrammetry and Image Analysis, German Aerospace Center (DLR), Wessling, Germany. In September 2022, he joined Starion Group as a Mission Engineer to support the development of the Copernicus Expansion Missions CHIME and LSTM at ESA-ESRIN, Frascati, Italy.

Stefan Auer (Member, IEEE) received the Dipl.-Ing. degree in geodesy and the Ph.D. degree in remote sensing both from Technical University of Munich, Munich, Germany, in 2005 and 2011, respectively.

Since 2014, he has been a Senior Researcher and Project Manager with the Remote Sensing Technology Institute, German Aerospace Center, Wessling, Germany. Besides coordinating projects in the fields of data processing, image analysis, and machine learning, he pursues research in fusion techniques for multimodal remote sensing data, image interpretation, and scene modeling. He developed the open-source simulator RaySAR, which helps to understand the nature of prominent signatures in synthetic aperture radar images. In the context of his doctoral thesis, he spent three months as a guest Researcher with the Department of Electronic and Telecommunication Engineering, University of Naples "Federico II," Naples, Italy.

Emiliano Carmona received the M.Sc. degree in physics and the Ph.D. degree in astroparticle physics from the University of Valencia, Valencia, Spain, in 1997 and 2004, respectively.

From 2005 to 2013, he was a Postdoctoral Researcher with the Max-Planck-Institute for Physics, Munich, Germany, and a Researcher with CIEMAT-Centre for Energy, Environment and Technology, Madrid, Spain. Since 2014, he has been with the Department of Photogrammetry and Image Analysis, German Aerospace Center, Wessling, Germany. His research interests focus on hyperspectral remote sensing.

Stefan Maier received the B.Eng. and M.Eng. degrees in applied physics from the Scientific Technical Academy Isny, Allgau, Germany, in 1991 and 1994, respectively. He received the Doctorate in engineering from Technical University Munich, Munich, Germany, in 2000.

From 2000 to 2002, he was a Postdoctoral Researcher with the German Remote Sensing Data Center, German Aerospace Center DLR in Oberpfaffenhofen. From 2002 to 2007, he was a Senior Research Officer with Satellite Remote Sensing Services, Western Australian Land Information Authority, Perth, Australia. In 2007, he was a Manager of Satellite Remote Sensing Services, Western Australian Land Information Authority. From 2007 to 2015, he was Principal Research Fellow Remote Sensing with the Research Institute for the Environment and Livelihoods, Charles Darwin University, Darwin, Australia. Since 2015, he has been the Director of maitec—Scientific and Engineering Consulting, Darwin, and since 2019 he has been a Visiting Scientist with the Department of Photogrammetry and Image Analysis, Remote Sensing Technology Institute, German Aerospace Center.

Rupert Müller received the Dipl.-Phys. degree in physics from the Ludwig Maximilians University of Munich, Munich, Germany, in 1984.

He is a senior scientist with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany, and Project Manager of the Ground Segment of MUSES (Multiuser System for Earth Sensing) with the hyperspectral instrument DESIS (DLR Earth Sensing Imaging Spectrometer) installed on the International Space Station. His main research interests include photogrammetric evaluation of optical data from air- and spaceborne sensors, digital image processing, machine learning, and hyperspectral imaging.

Patrick Rademske biography not available at the time of publication.

Uwe Rascher was born in Germany, in 1968. He received the Diploma degree in biology (master's) from University Erlangen-Nürnberg, Erlangen, Germany, in 1996, the Ph.D. degree (Dr.rer.nat.) in biology from the Technical University Darmstadt, Darmstadt, Germany, in 2001, and the Habilitation degree in plant biology from the University of Düsseldorf, Düsseldorf, Germany, in 2009.

From 2001 to 2004, he was working as a Postdoctoral Fellow with the Biosphere 2 Centre, Columbia University, Tucson, AZ, USA. In 2004, he joined the Forschungszentrum Jülich, Jülich, Germany, as a Young Research Scientist, where he was promoted to Group Leader in 2011. In 2014, he was appointed as a Full Professor with the Agricultural Faculty, University of Bonn, Bonn, Germany. He is currently the Head of the research area "Shoot Dynamics" at the Forschungszentrum Jülich, a Principal investigator of the Cluster of Excellence "PhenoRob," and a member of the MAG, European Space Agency (ESA), Darmstadt, Germany, in the preparation of the FLEX satellite mission, which will become the eighth Earth Explorer. His research combines work in plant physiology, remote sensing, ecology, global climate change, and plant phenotyping. His expertise is in characterizing photosynthesis using chlorophyll fluorescence, hyperspectral reflectance, gas exchange, remote sensing, and image analysis techniques. He further focuses on interdisciplinary work using novel remote sensing approaches to better measure photosynthesis and stress, develop novel plant phenotyping concepts, and model plant-mediated exchange processes from the leaf to the region.

Hanno Scharr received the Diploma and Ph.D. degrees in physics, from Heidelberg University, Heidelberg, Germany, in 1996 and 2000, respectively, and the Habilitation in computer science from Frankfurt University, Frankfurt, Germany, in 2014.

His Ph.D. work was on "Optimal Operators in Digital Image Processing." After a first postdoc in Heidelberg, he joined Intel Research, Santa Clara, CA, USA, for a year where he worked on real-time denoising algorithms for nanomachining tools. In 2003, he moved to Forschungszentrum, Jülich, Germany, leading a junior group on "Automatic Imaging and Quantitative Image Processing in Environment Plant Sciences" with a Young Investigator Award. He led the Image Analysis Group of the Institute of Bio- and Geosystems: IBG-2 Plant Sciences, Forschungszentrum Jülich. Since 2021, he has been the Head of the Institute of Advanced Simulation (IAS-8): Data Analytics and Machine Learning, Forschungzentrum Jülich. He has authored and coauthored more than 150 publications and patents. He serves as a Reviewer for all major computer vision journals, and for more than 35 international conference programme committes. His research interests include quantitative image processing, computer vision, and machine learning, with applications to natural sciences.