Leveraging a large-scale radiative transfer simulation for an emulator based retrieval scheme of sun-induced fluorescence in HyPlant imagery

EARSeL Valencia 2024
Abstract
Corresponding Author: j.buffat@fz-juelich.de

Jim Buffat[1], Miguel Pato[2], Stefan Auer[2], Kevin Alonso[3], Emiliano Carmona[2], Stefan Maier[2], Rupert Müller[2], Patrick Rademske[1], Uwe Rascher[1], Hanno Scharr[4]

[1] Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Jülich, Germany
[2] German Aerospace Center (DLR), Earth Observation Center, Remote Sensing Technology Institute, Oberpfaffenhofen, Germany
[3] RHEA Group c/o European Space Agency (ESA), Largo Galileo Galilei, Frascati, Italy
[4] Forschungszentrum Jülich GmbH, Institute of Advanced Simulations, IAS-8: Data Analytics and Machine Learning, Jülich, Germany

Keywords (5): Earth Observation, Sun-induced Fluorescence, Hyperspectral, Neural Network, Machine Learning

Challenge (800 - 1000 characters incl. spaces)

The prediction of sun-induced fluorescence (SIF) from hyperspectral radiance has been identified as a corner stone to assess plants' photosynthetic efficiency remotely. It is widely accepted that remotely sensed SIF offers great potential for a variety of applications. To provide such estimates, top-of-canopy SIF products derived from passively sensed radiance measurements of various airborne and spaceborne sensors have been developed over the last decades. To date, however, physically based SIF retrieval schemes require a prohibitive use of computationally costly radiative transfer simulations especially when used in complex observational conditions such as in hilly terrain. In this contribution we report on our on-going work to develop a lightweight self-supervised neural network to retrieve SIF in the O_2-A absorption band of HyPlant acquisitions. We aim at a tight integration of a physical radiative transfer model with the network to ensure physically sound predictions by leveraging large scale simulation and emulation of HyPlant at-sensor radiance observations. We report on first results that we achieve on a dedicated data set.

Methodology (1200 – 1500 characters incl. spaces)

We formulate the SIF retrieval scheme as a feature based optimization performed by a neural network. A specifically designed neural network is trained to spectrally fit a physical model of the at-sensor radiance to observed data. The loss formulation for training involves the fitting residuals and additional regularizers guaranteeing physiological and physical constraints. We moreover take advantage of variability constraints in the architectural set up. Contrarily to other spectral fitting methods the physical model is, however, not implemented explicitly. Instead we make use of a fast and accurate emulator built from a large database of simulated HyPlant spectra mapping the simulation parameters to the hyperspectral at-sensor radiance (cf. Fig. 1). The presented approach is completely self-supervised, differently to supervised approaches where the network would be trained to directly predict a SIF label. A supervised approach for SIF prediction from DESIS data based on a similarly constructed data set was developed in parallel and has been submitted by Pato et al.

We model the radiative transfer around the O_2-A absorption band (740 - 780 nm) with a general-purpose simulation tool based on MODTRAN6. The simulation tool parameterizes surface properties, geometrical conditions and sensor characteristics affecting the at-sensor radiance with 13 parameters. Detailed analysis of the parameters' sensitivities with respect to the at-sensor fluorescence signal and a precise determination of observed parameter ranges allow us to realistically model HyPlant at-sensor radiance. We generated hyperspectral simulation databases (349 bands, spectral sampling interval of 0.11 nm) by extensively sampling the input parameter ranges.

Results (1200 – 1500 characters incl. spaces)

Our work explores a self-supervised neural network based SIF prediction scheme leveraging an emulator trained on a large-scale hyperspectral database. We show that the implementation of the emulator as a 4^{th} degree polynomial yields a mean radiance error of less than 0.1 $mW/m^2/sr/nm$ across the O_2-A band which is less than 10% of a typical SIF at-sensor signal. We furthermore detail that the emulator yields highly realistic spectra by showing that the mean fitting residual of the neural network over the whole fitted spectrum is less than 1.5 $mW/m^2/sr/nm$ in a typical HyPlant flight line.

The proposed network was trained on a preliminary compilation of HyPlant acquisitions from 2018 - 2023 (30 acquisitions) and tested on a data set of HyPlant campaigns from 2018 and 2020 where suitable top-of-canopy SIF measurements exist that can be used as ground truth in a validation study. We find our approach to yield comparable results to the state-of-the-art Spectral Fitting Method for HyPlant (Cogliati et al. 2019) with significant ($p < 0.01$) correlation scores larger than 0.85 with respect to the top-of-canopy SIF measurements.

Furthermore we acquired additional HyPlant data over hilly terrain in 2022 and 2023 to be able to test the retrieval under conditions that are not suitable for operational state-of-the-art SIF retrieval schemes. In Fig. 2 we show an exemplary HyPlant acquisition and corresponding SIF predictions. Although we cannot conduct a validation study with ground measurements in these cases, we are able to assert that the resulting SIF predictions are not correlated to the topography as is the case with other methods. Further work will focus on establishing a better base line to validate the SIF prediction under these challenging conditions.

Outlook for the future (800 - 1000 characters incl. spaces)

The presented network was trained and tested on a particular compilation of HyPlant acquisitions. Our results indicate that the training leads to good performance with respect to ground measurements and suggested consistent performance in a variety of topographic and observational conditions. Further work will address specifically the consistency of our approach across larger HyPlant data sets and the potential benefit of larger training data sets for the generalization capability of the network.

We believe that the presented approach could be used as a blueprint for computationally efficient SIF retrieval for arbitrary hyperspectral sensors with suitable spatial and spectral resolution. If it can be shown that the network may be trained on a default data set once and then applied for prediction on new data without or only with short additional training times, the presented approach could be interesting for SIF prediction in air and spaceborne sensor missions with high data throughput where efficiency is a key requirement. Therefore, we will extend the validation of this approach to simulated data of the upcoming FLEX mission of ESA.

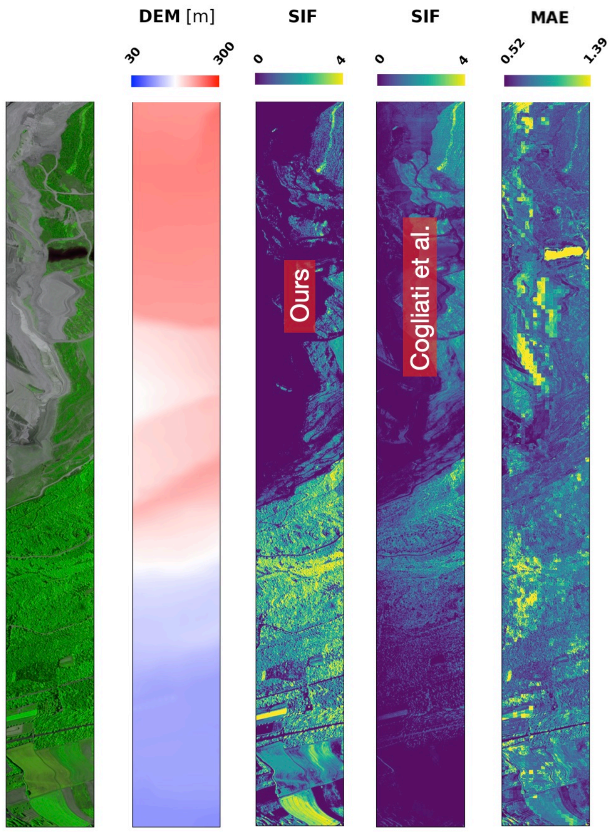

Figure 1: Comparison of SIF prediction of a HyPlant acquisition in hilly terrain (14/06/2022, 09:47 UTC, Jülich DE). Shown are a false color composite of the input data, a DEM warped to sensor geometry, the top-of-canopy SIF prediction following our approach, the top-of-canopy SIF prediction following the Spectral Fitting Method (Cogliati et al. 2019) and the mean absolute error (MAE) of our approach over the fitted HyPlant spectra (left to right). All SIF magnitudes are shown at 761 nm. SIF and MAE are given in units of mW/m²/sr/nm.

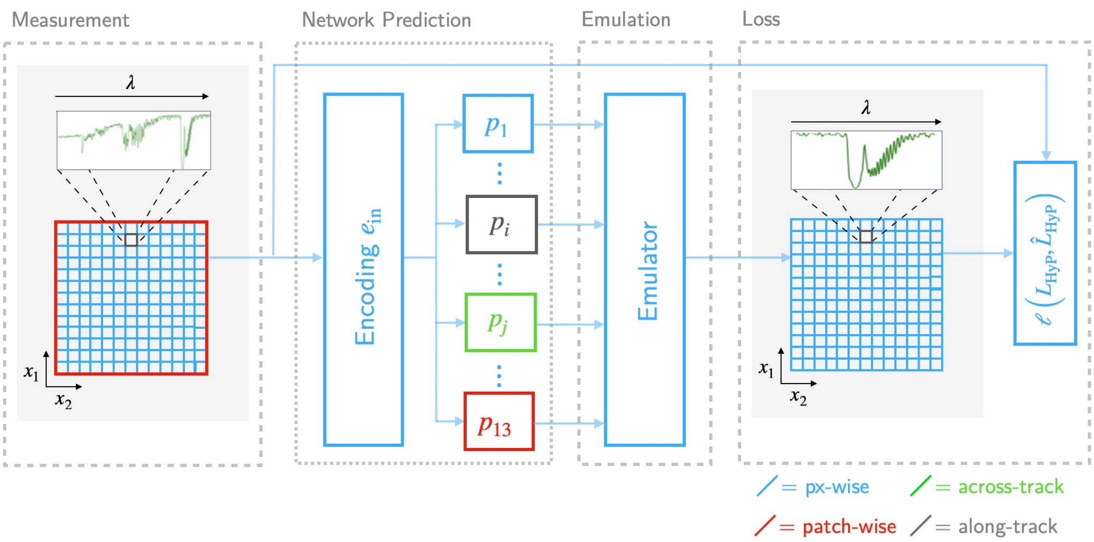

Figure 2: Sketch of the set-up used for training and prediction in this contribution. We provide patches of HyPlant acquisition in sensor geometry, encode it and pass the encoding to submodules p_k predicting physical values for all 13 parameters used in the simulation. Depending on the parameter we estimate a value for each pixel, for a single along-track or across-track position, or a single value for the whole patch. The estimated parameters are then passed to the emulator which provides us simulated HyPlant at-sensor radiance in the range 740 - 780 nm. The fluorescence per pixel is derived from the estimated parameterization.