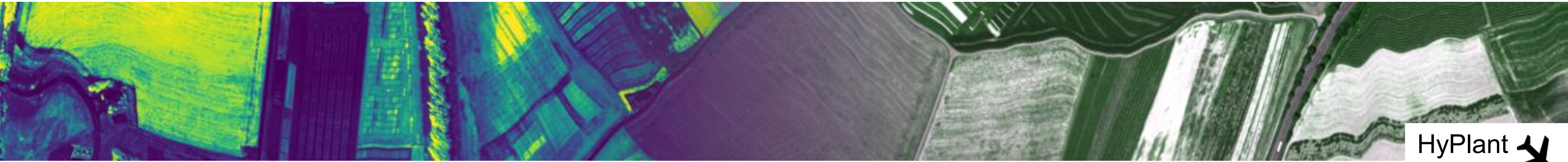
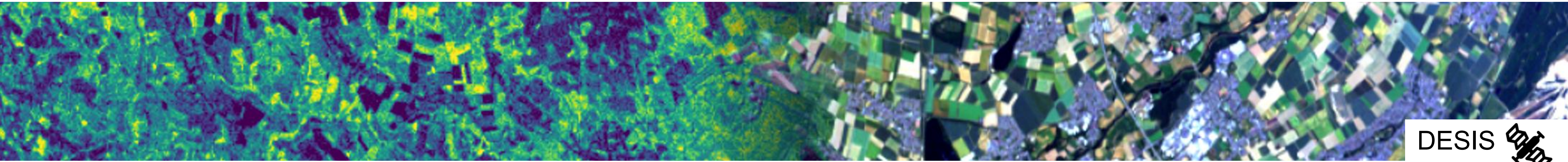




Formulating SIF retrieval as a downstream task to a hyperspectral foundation model

Jim Buffat[1], Miguel Pato[2], Stefan Maier[2], Emiliano Carmona[2], Uwe Rascher[3], Hanno Scharr[1]

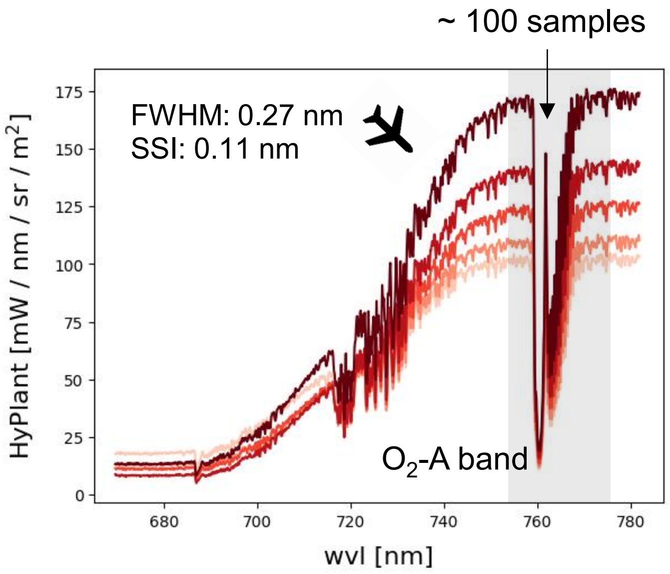
[1] Forschungszentrum Jülich, Institute of Advanced Simulations, IAS-8: Data Analytics and Machine Learning
[2] German Aerospace Center (DLR), Earth Observation Center, Remote Sensing Technology Institute, Oberpfaffenhofen, Germany
[3] Forschungszentrum Jülich, Institute of Bio- and Geosciences, IBG-2: Plant Sciences

Sensor-agnostic retrieval of solar-induced fluorescence (SIF) in the O_2-A band

- Foundation modelling paradigm:
  - Self-supervised **pretext** task on **large data** set
  - **downstream** task finetuning on **specialized data**
- **Explored here: FM-based SIF** retrieval from hyperspectral radiance
- **Long-term goals**
  - Single retrieval **SIF composites** (FLEX, DESIS, S5P, S3, TanSat)
  - **Few-shot** adaption to new sensors
- Necessary: **Feature-based SIF** retrieval methodology

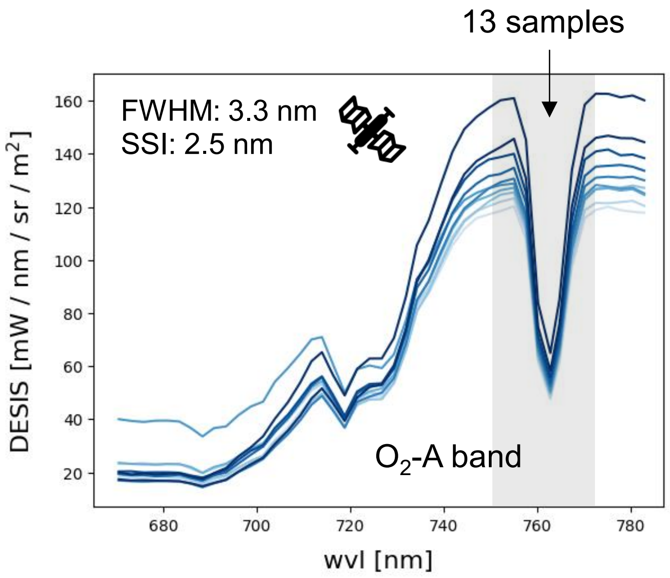
HyPlant FLUO

GSD < 5m, FWHM: 0.27 nm

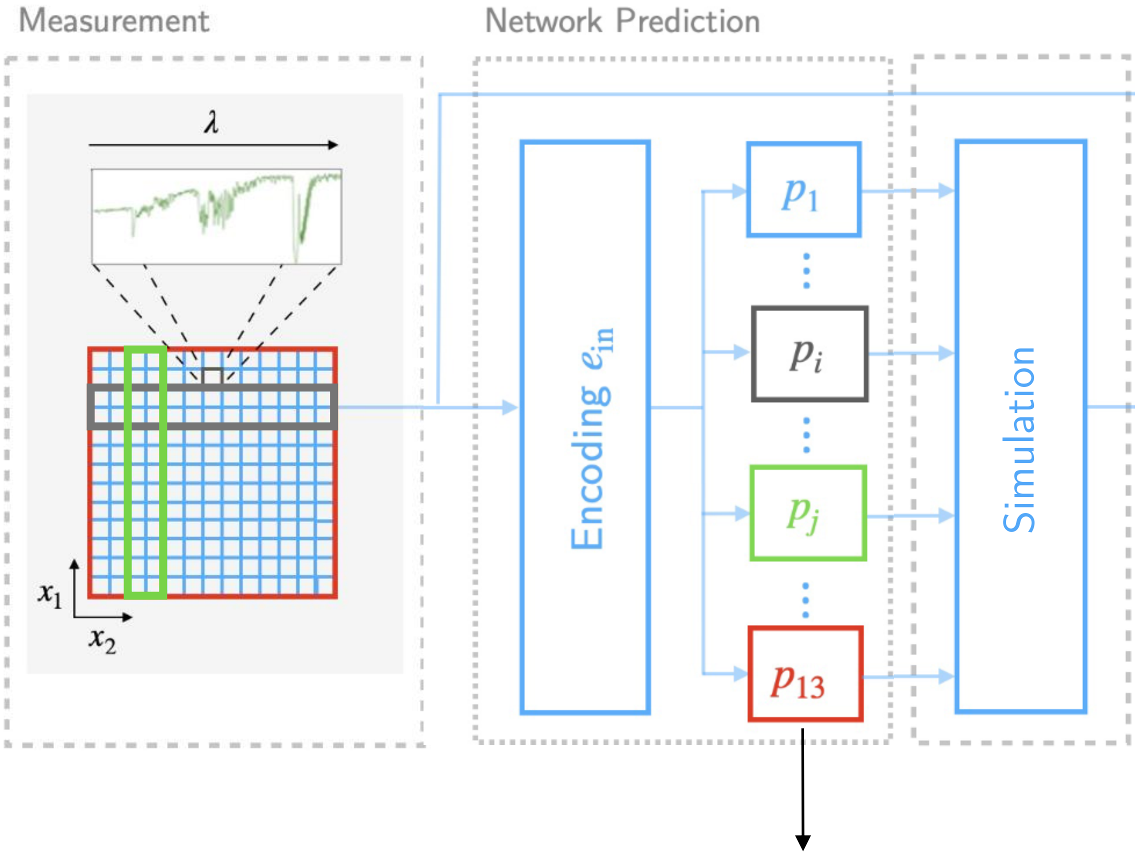


DESIS

GSD = 30 m, FWHM: 3.5 nm

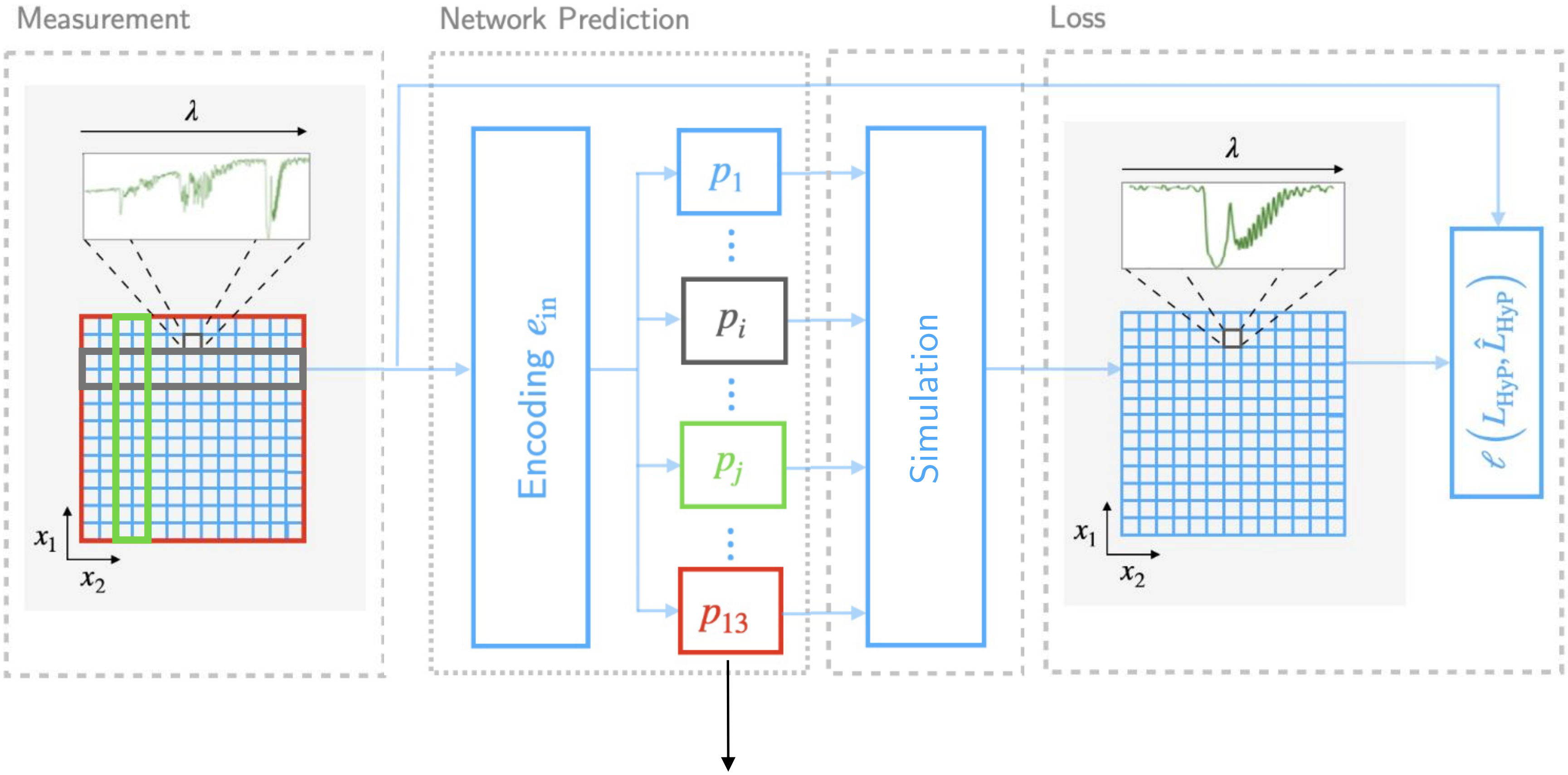


Outline of the Spectral Fitting Method Neural Network (SFMNN): Replacing the optimization in in SFM



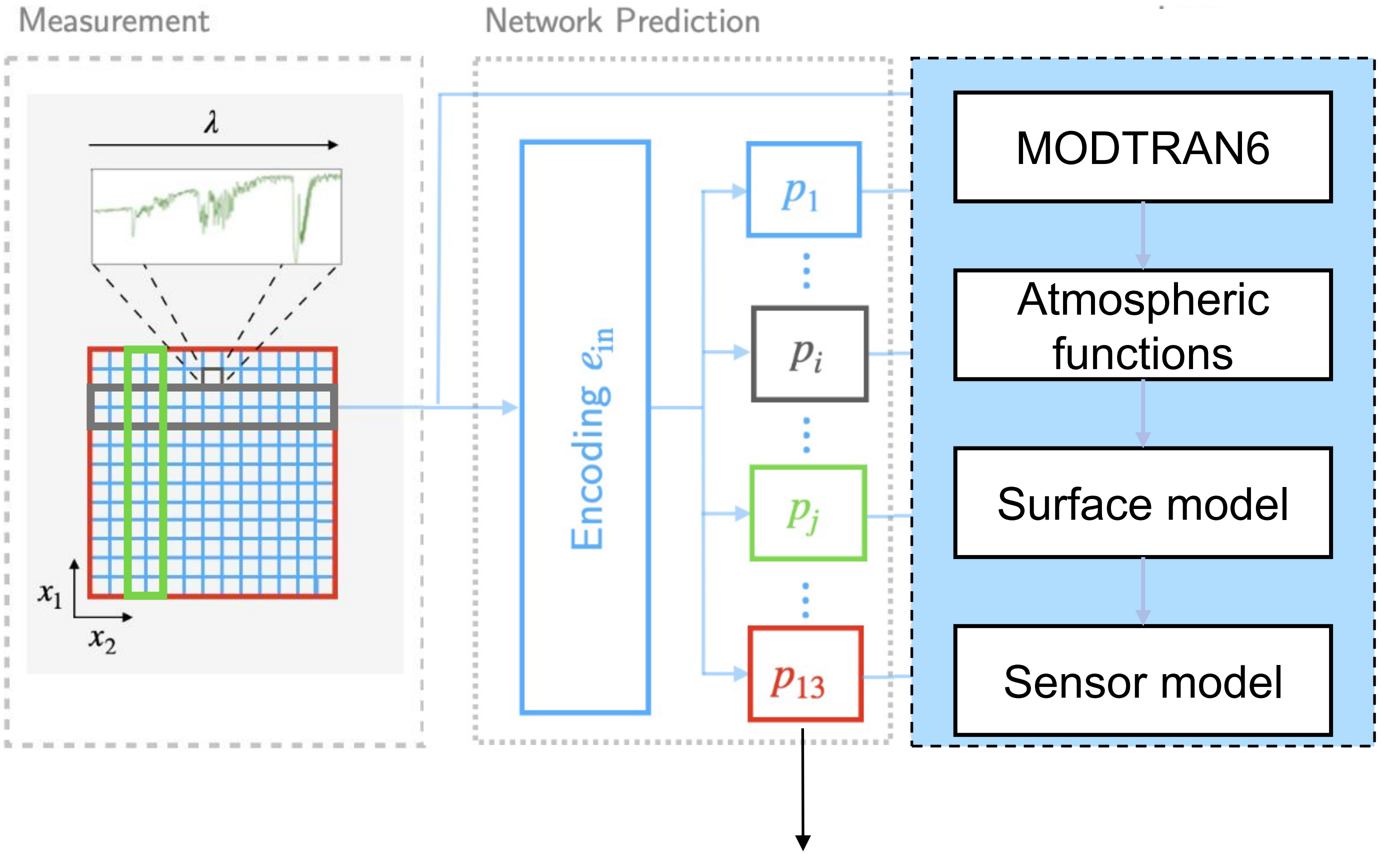
e.g. AOT, reflectance or fluorescence

Outline of the Spectral Fitting Method Neural Network (SFMNN): Replacing the optimization in in SFM



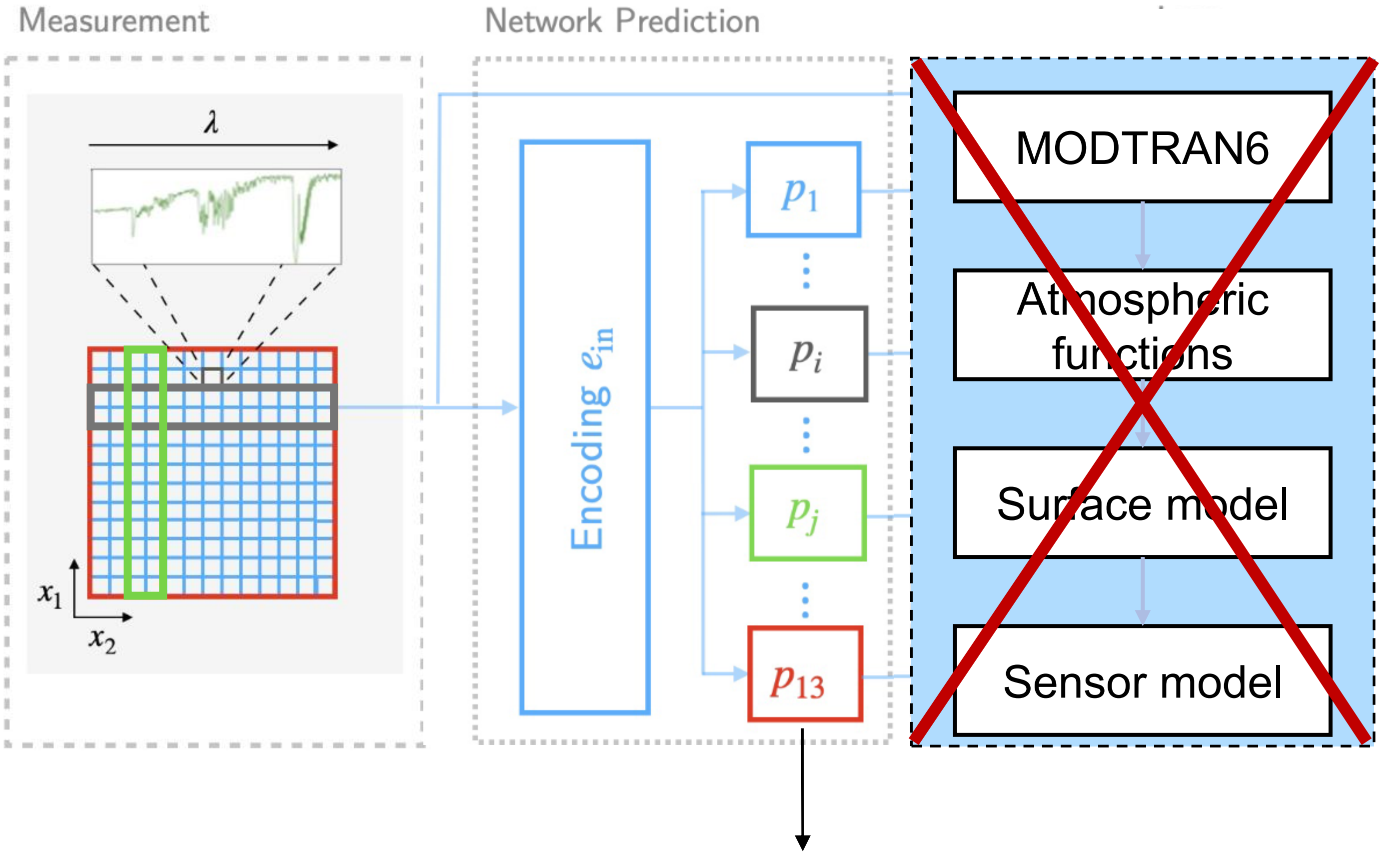
e.g. AOT, reflectance or fluorescence

Outline of the Spectral Fitting Method Neural Network (SFMNN): Replacing the optimization in in SFM



e.g. AOT, reflectance or fluorescence

Outline of the Spectral Fitting Method Neural Network (SFMNN): Replacing the optimization in in SFM



e.g. AOT, reflectance or fluorescence

→ **No gradient**

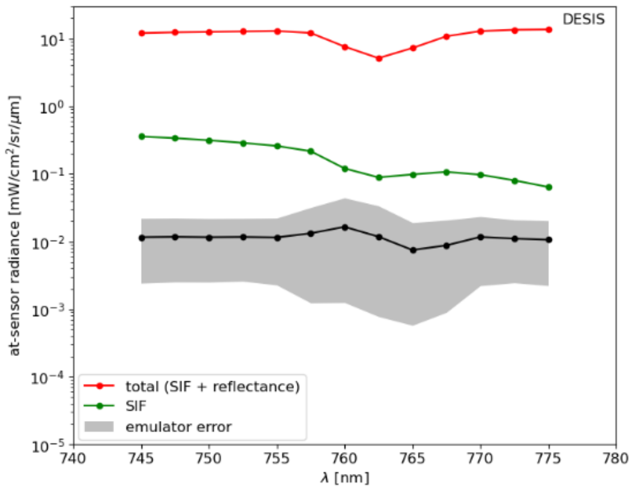
→ **Computationally expensive**

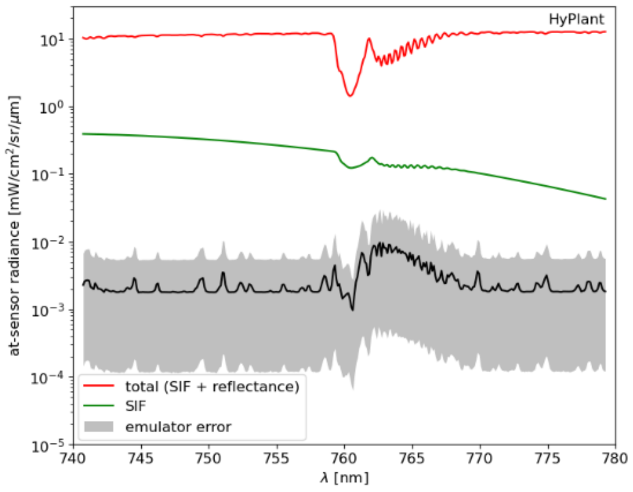
Fast emulation of in DESIS and HyPlant O_2-A absorption bands for NN training

- RTM **simulations** of **typical HyPlant** and **DESIS** observational conditions in O_2-A
- Simulation **approximation** with a **surrogate model**
- Balance three goals
  - Approximation **accuracy**
  - Prediction **speed**
  - **Gradient**
- Polynomial, simple NN (MLPs), kernel and Gaussian ridge regression
- **4th order polynomial** meets requirements

Performance parameter	DESIS							
	OLS	**P2**	**P4**	**P5**	**N1**	**N2**	**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**	**N1**	**N2**	**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

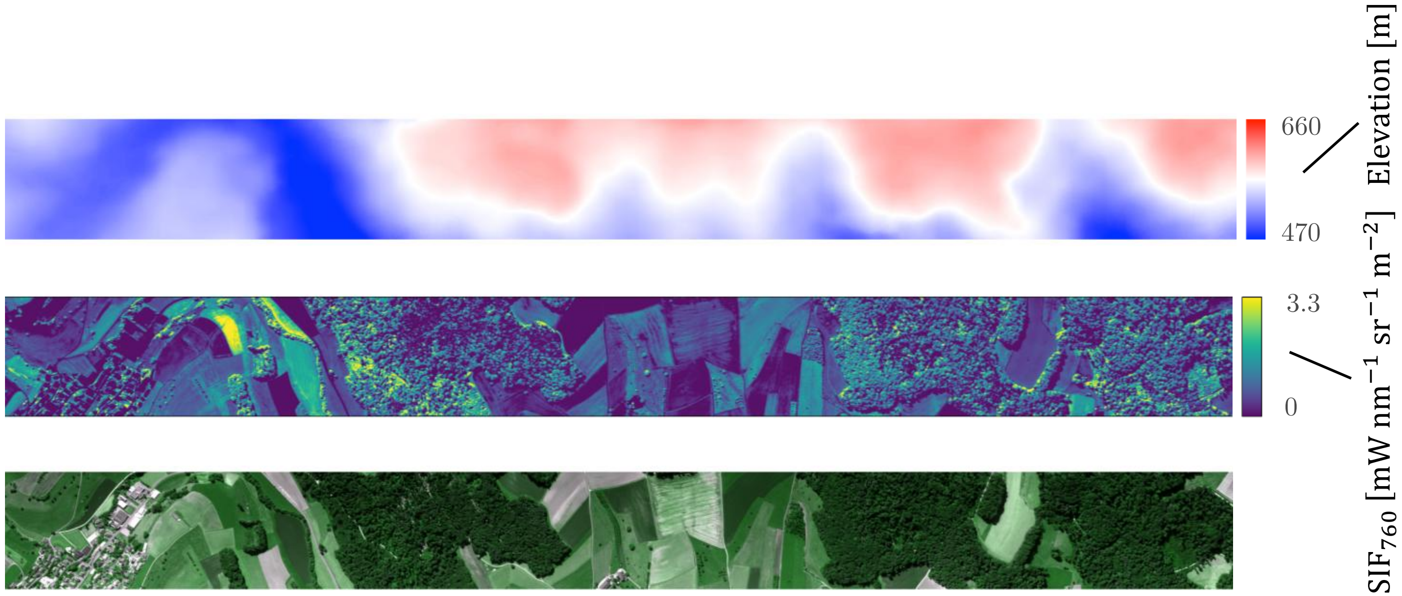




Pato et al. 2025, JSTARS, doi: 10.1109/JSTARS.2024.3457231

SFMNN yields state-of-the-art SIF retrieval in HyPlant

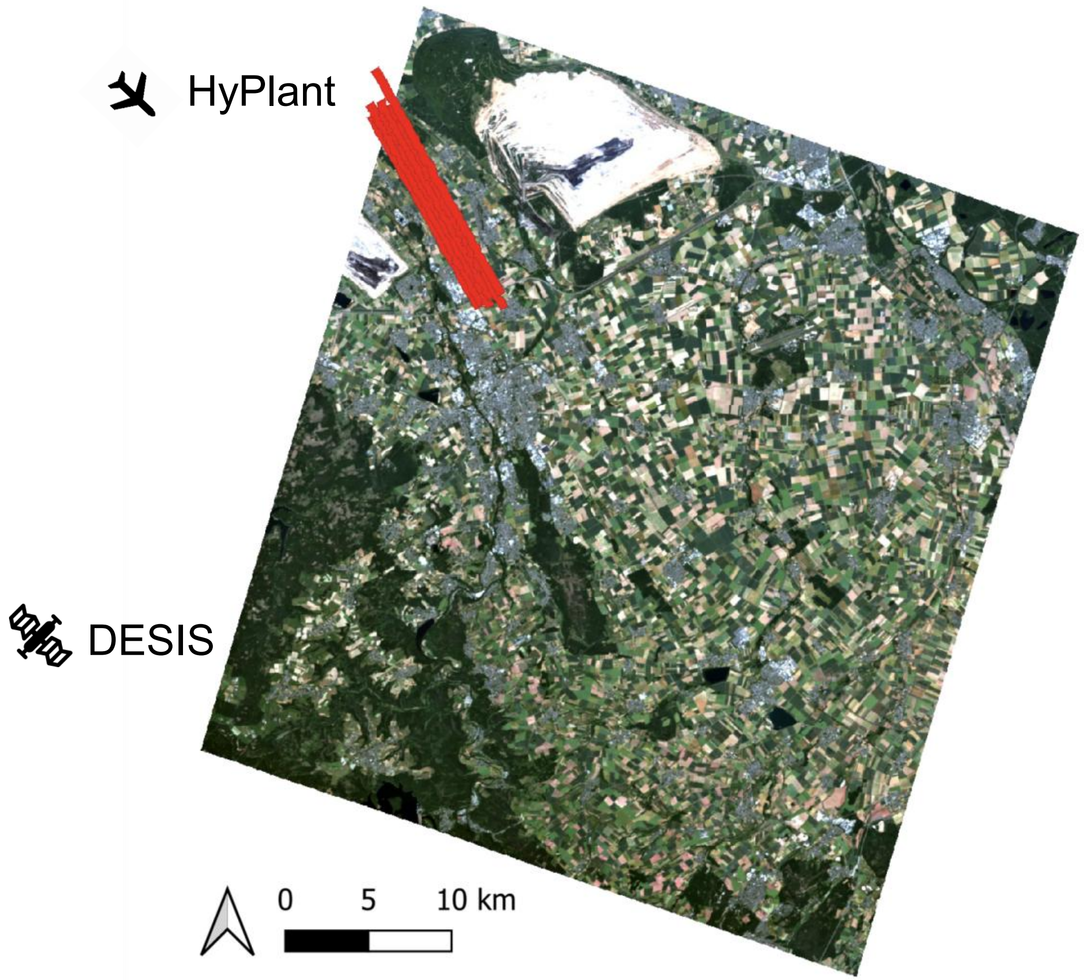
- **Encoder** and **decoder** defined as **MLPs**
- **Training** on individual **campaign** data sets
- Validation with **FLoX SIF** estimates
- **State-of-the-art** SIF retrieval accuracy
- **Straight forward** SIF prediction in **topographically variable** terrain



Buffat et al. 2026, RSE, doi: 10.1016/j.rse.2025.115203

Quasi-simultaneous matchups for SFMNN performance validation in DESIS

- 2023 **match-ups** between DESIS and HyPlant
- Small time difference of **4 - 25 minutes**
- SFMNN with **additional regularization**

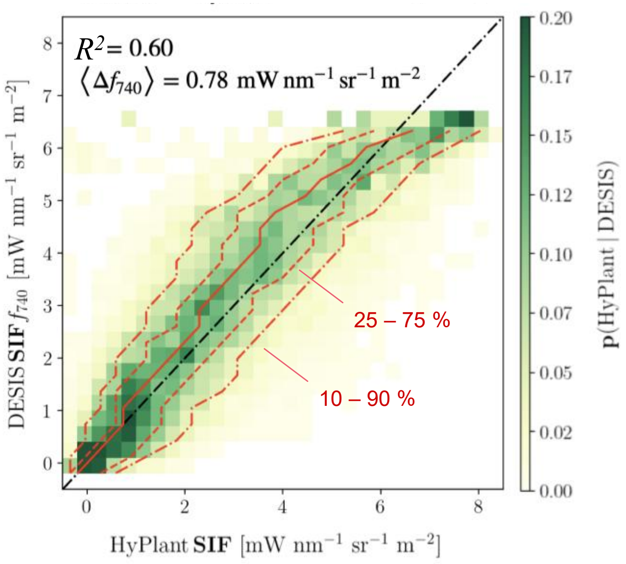


Buffat et al. 2024, Selected HyPlant, DESIS and OCO-3 acquisitions for the cross-comparison of sun-induced fluorescence products
doi: 10.26165/JUELICH-DATA/YNYLFQ

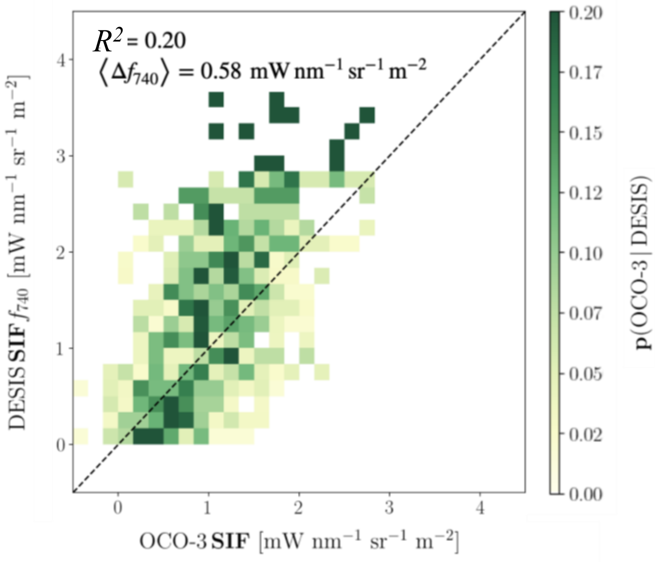
DESIS: the first SIF estimates from space at 30 m

- Additional constraints:
  - **Atmospheric** quantities
  - **Residual prediction** wrt. supervised predictor
- Comparison in match-ups reveals **significant agreement**
- Performance **skillful** under **exclusion** of eventual **reflectance correlation**

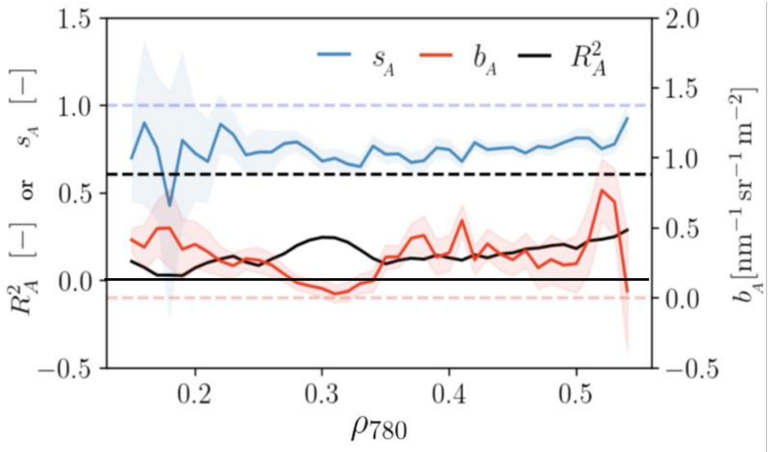
DESIS – HyPlant comparison



DESIS – OCO3 comparison



R^2 controlled for reflectance at 780 nm



$$\langle R^2 \rangle_A = |P|^{-1} \sum R^2(A_{\rho_{780}})$$

$$A_{\rho_{780}} = p\left(f_{740} \,\middle|\, \hat{f}_{740}, |\rho - \rho_{780}| < d\rho\right)$$

Buffat et al. 2025, ECCV CVPPA 2024, doi: 10.1007/978-3-031-91835-3_6

Setup of FM-based multi-sensor SFMNN training

- **Limitations** of SFMNN:
  - Sensor-specific emulation and training
  - Finetuning necessary
- **Targeted solution** (SpecFound): FM + SFMNN downstream training
- FM pretraining over **varied set** of hyperspectral data sources

Dataset	#Rad	Ref.	SIF	Size	Bands	GSD
DESIS	20×10^3	✓	✓	256×256	235	30m
EnMAP	20×10^3	✓	–	256×256	202	30m
FLUO	27×10^3	–	✓	128×128	1024	<5m
DUAL	21×10^3	✓	–	128×128	626	<5m
EMIT	65×10^3	✓	–	256×256	240	30m
AvirisNG	8×10^3	✓	–	256×256	220-430	<25m
APEX	10×10^3	✓	–	128×128	312	<10m
Sentinel-3 (*)	27×10^3	–	–	256×256	11	500 m
Sentinel-5P (*)	22×10^3	–	✓	256×256	200	7 km
E2E FLEX (*)	120	✓	✓	128×128	409	300m

prepared

planned

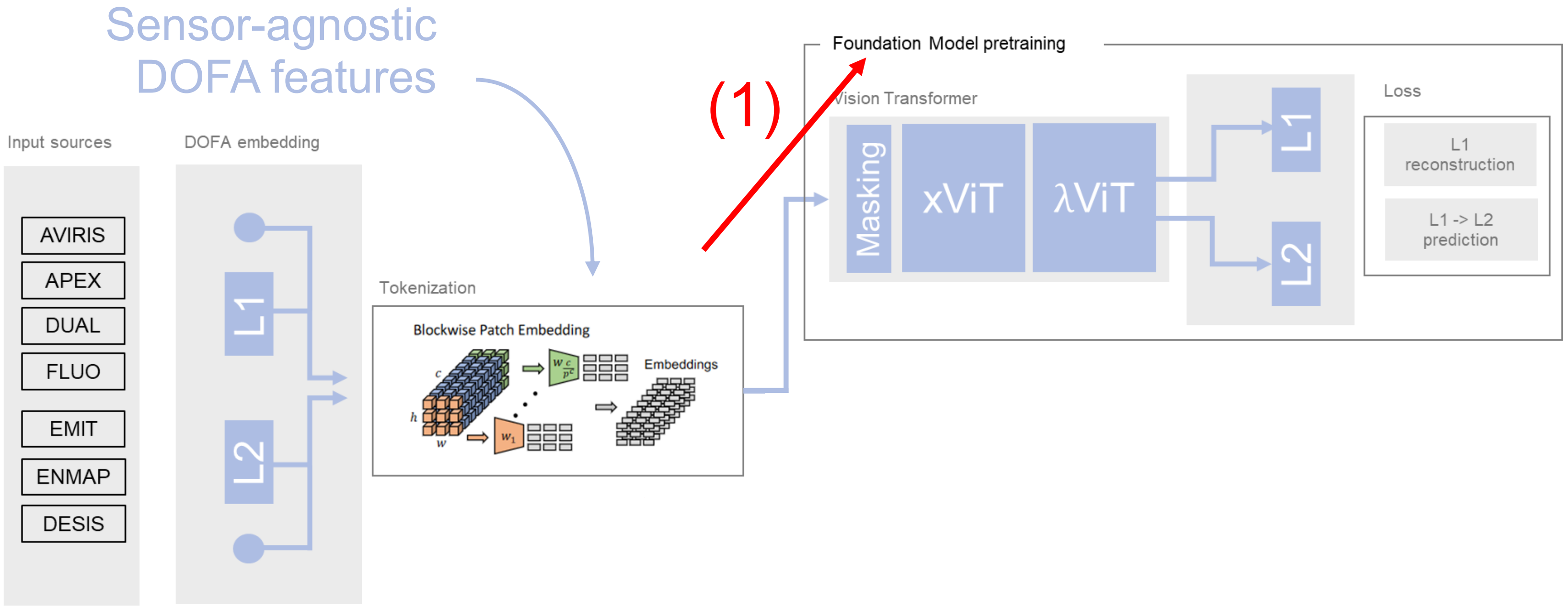
Setup of FM-based multi-sensor SFMNN training (SpecFound)

Pretraining

- **Dynamic-One-For-All** : sensor-agnostic spectral representation (DOFA, *Xiong et al. 2024*)
- **Separation** of **spatial** and **spectral** features (MaskedSST, *Scheibenreif et al. 2023*)
- **Masked image modelling** and **Rad. to Ref.**

SFMNN downstream finetuning

- **Downstream training:** SFMNN on DOFA features
- **Joint** training on DESIS and HyPlant



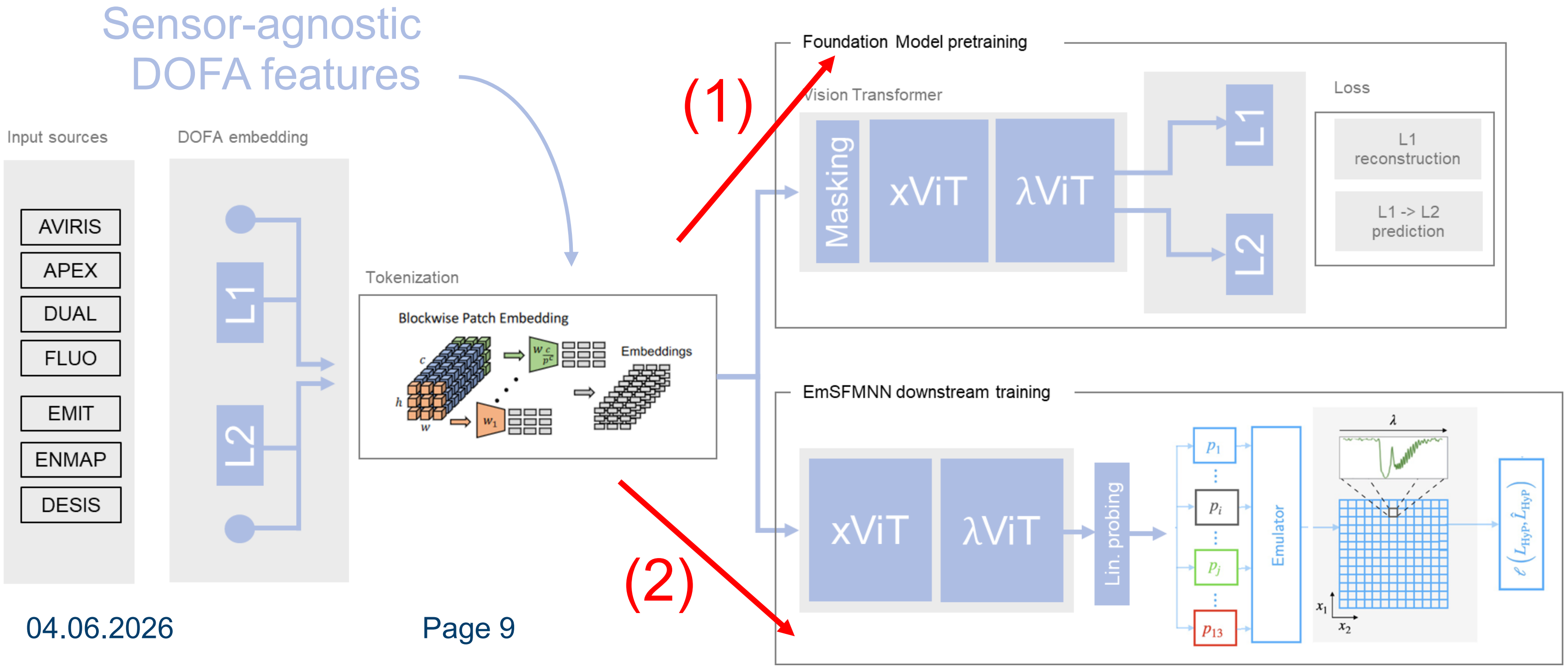
Setup of FM-based multi-sensor SFMNN training (SpecFound)

Pretraining

- **Dynamic-One-For-All** : sensor-agnostic spectral representation (DOFA, *Xiong et al. 2024*)
- **Separation** of **spatial** and **spectral** features (MaskedSST, *Scheibenreif et al. 2023*)
- **Masked image modelling** and **Rad. to Ref.**

SFMNN downstream finetuning

- **Downstream training:** SFMNN on DOFA features
- **Joint** training on DESIS and HyPlant



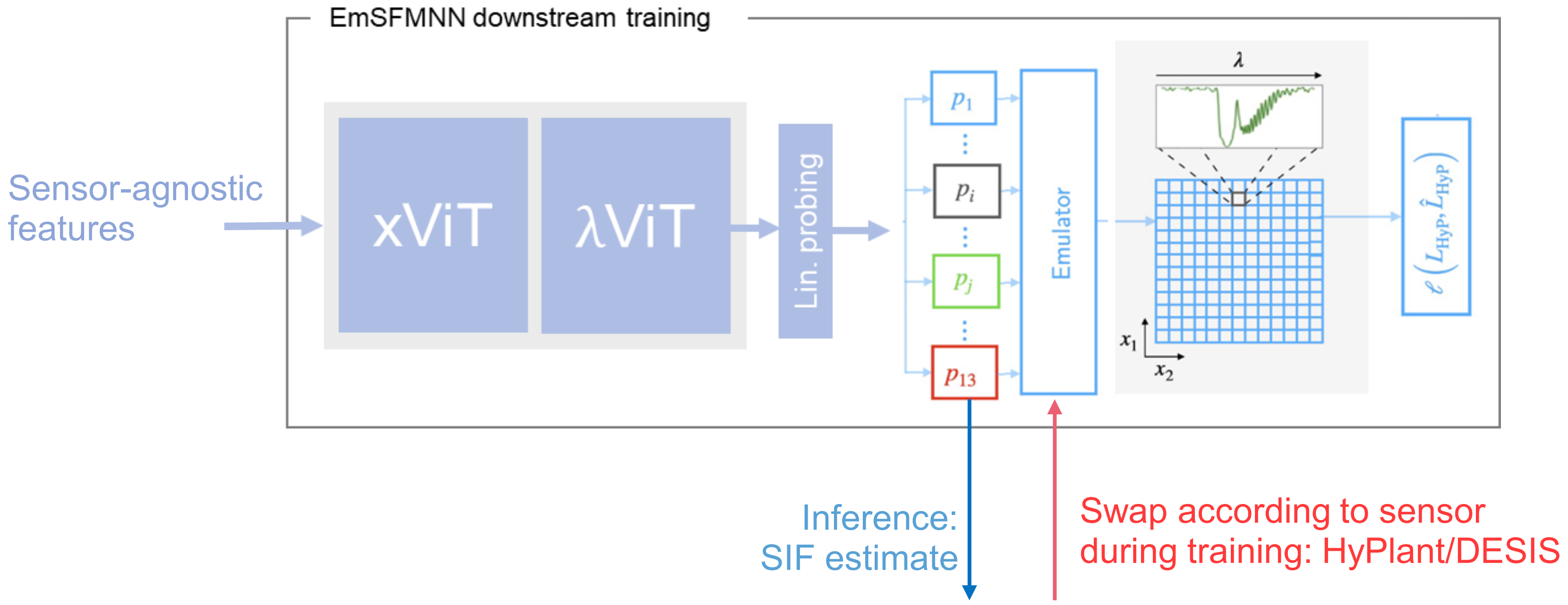
Setup of FM-based multi-sensor SFMNN training (SpecFound)

Pretraining

- **Dynamic-One-For-All** : sensor-agnostic spectral representation (DOFA, *Xiong et al. 2024*)
- **Separation** of **spatial** and **spectral** features (MaskedSST, *Scheibenreif et al. 2023*)
- **Masked image modelling** and **Rad. to Ref.**

SFMNN downstream finetuning

- **Downstream training:** SFMNN on DOFA features
- **Joint** training on DESIS and HyPlant



HyPlant: SpecFound SIF retrieval performance is subpar to SOTA but sensitive to in-situ variation

- Explained variance (R^2) similar to SFM and iFLD baselines
- Absolute Error (MAE) is elevated with respect to SFMNN

HyPlant / FloX measurement match ups

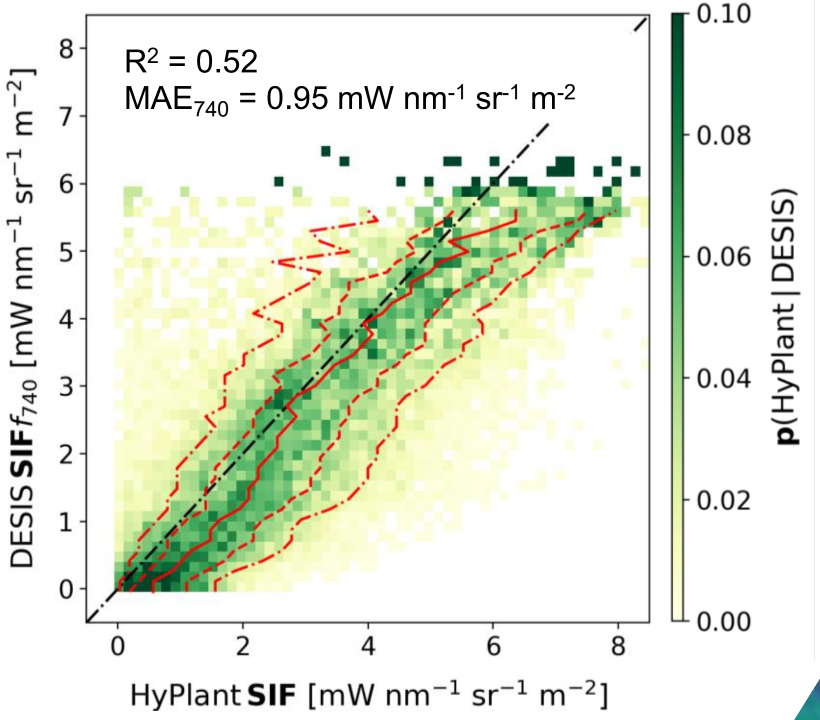
	CKA 2020 (350 m)		CKA 2020 (600 m)		GLO 2021 (1100 m)	
	R^2	MAE_{760}	R^2	MAE_{760}	R^2	MAE_{760}
SpecFound	0.31	0.49	**0.57**	0.38	0.72	**0.20**
SFMNN	**0.62**	**0.17**	0.49	0.25	**0.94**	0.30
SFM	0.15	0.28	-0.48	0.33	0.79	0.21
iFLD	0.40	0.23	0.06	**0.24**	0.56	0.84

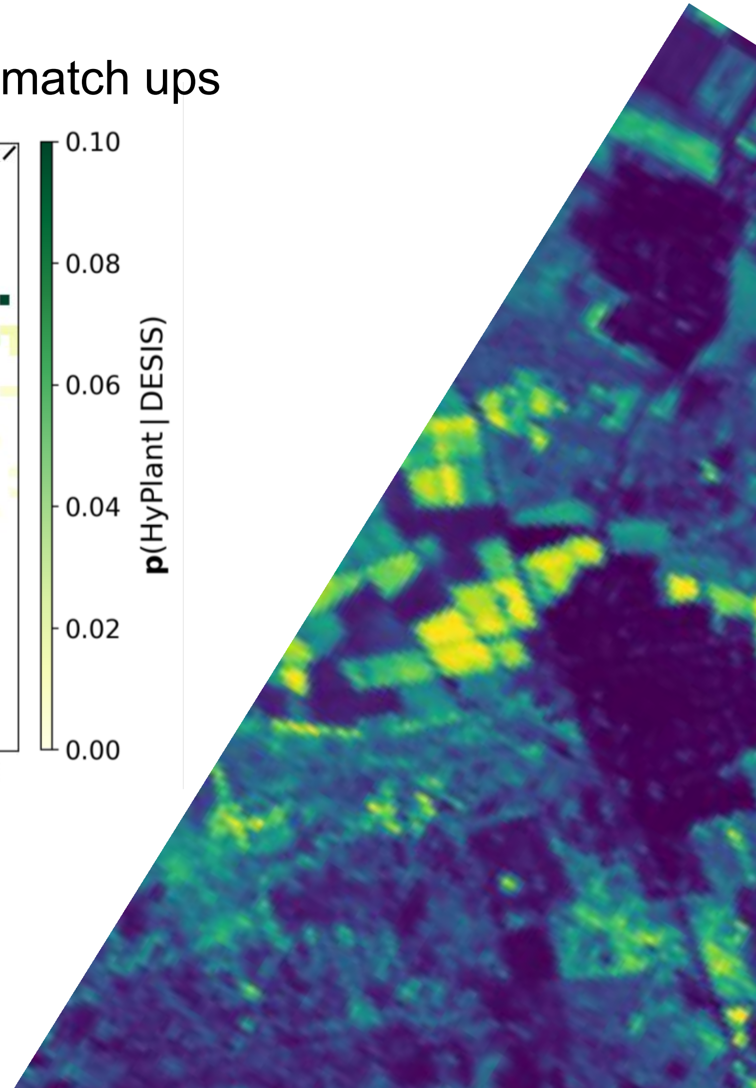
SpecFound DESIS SIF estimate is sensitive to airborne HyPlant SFM product

- SpecFound SIF is **sensitive to HyPlant SFM**
- Absolute **performance lower** than SFMNN
- Achieved **w/o additional** constraints

	MAE_{740} [mW nm^{-1} sr^{-1} m^{-2}]	R^2	$\langle R_A^2 \rangle$
SpecFound	0.95	0.52	0.11
SFMNN	**0.78**	**0.60**	**0.15**

DESIS / HyPlant measurement match ups





Conclusions and Outlook

- **SpecFound: SFMNN** as a downstream task to a sensor-agnostic foundation model.
- **Validation** with in-situ measurements and cross-validation: estimates are **sensitive to high-quality SIF estimates** both in HyPlant and DESIS.

- **SpecFound** SIF retrieval performance is **subpar compared to SFMNN**.
  - Atmospheric constraints
  - Calibration variation

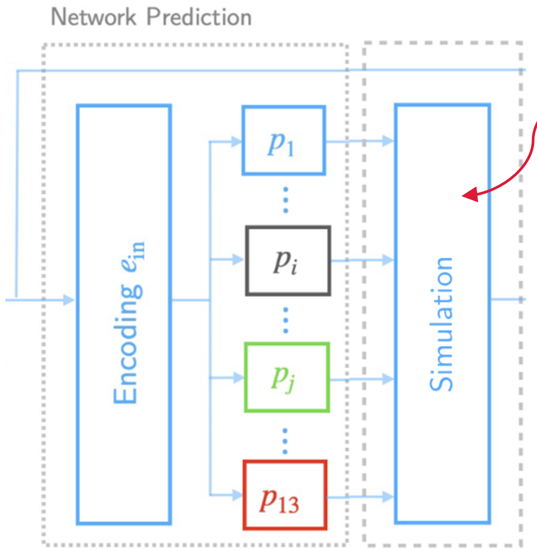
- Few-shot properties in **FLEX** and **Sentinel-5P**
- Additional sensors in data set: performance with **retrained SpecFound** model

The authors gratefully acknowledge computing time on the supercomputer JURECA [1] at Forschungszentrum Jülich under grant no. fluomap-ct.

[1] Jülich Supercomputing Centre. (2021). JURECA: Data Centric and Booster Modules implementing the Modular Supercomputing Architecture at Jülich Supercomputing Centre Journal of large-scale research facilities, 7, A182. http://dx.doi.org/10.17815/jlsrf-7-182

Replacing the simulation with an emulator: Spectral Fitting Neural Network (SFMNN)

- SFMNN integration: **use simple emulator** as simulation layer
- Extended emulator with **bandwise sensor miscalibration**



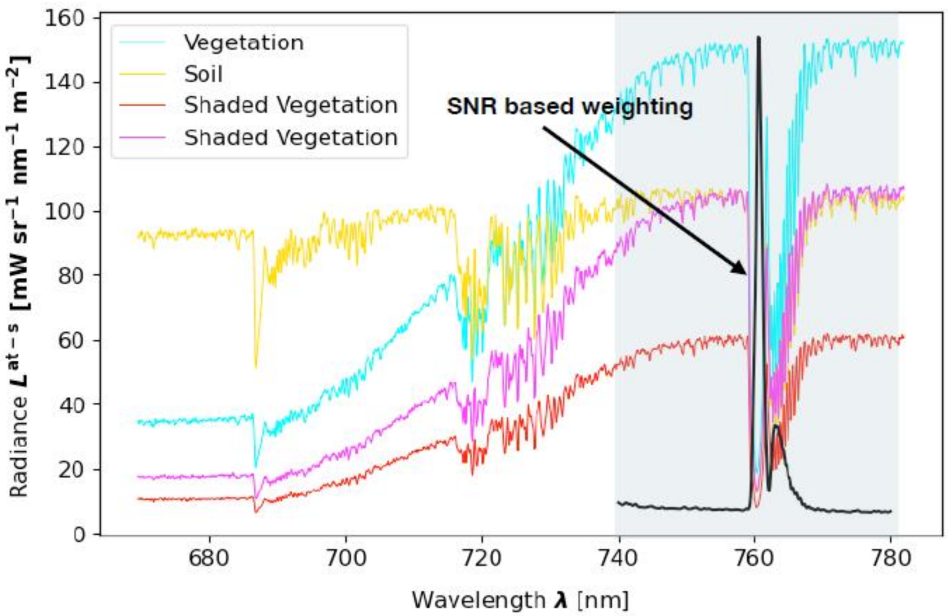
Emulator
(~ 4th order polynomial)

Buffat et al. 2026, RSE, doi: 10.1016/j.rse.2025.115203

JÜLICH Forschungszentrum

SFMNN loss formulation: regularized residual minimization

- Loss **regularization**
- No SIF in **bare soil**
- **Loss weighting** according to modelled signal strength



$$\ell_{\text{res}}\left(L,\hat{L}\right)=\left\langle\left(L-\hat{L}\right)^2\right\rangle_{\lambda,\,x}+\frac{\gamma_f}{|\mathcal{W}_{\text{out}}|}\left\langle\sum_{\lambda\in\mathcal{W}_{\text{out}}}w_\lambda\left(L(\lambda)-\hat{L}(\lambda)\right)^2\right\rangle_x+\gamma_N\,\hat{f}\,\delta\left(\text{NDVI}(L)\leq\tau\right)$$

Overall reconstruction | O_2-A reconstruction | Bare soil constraint

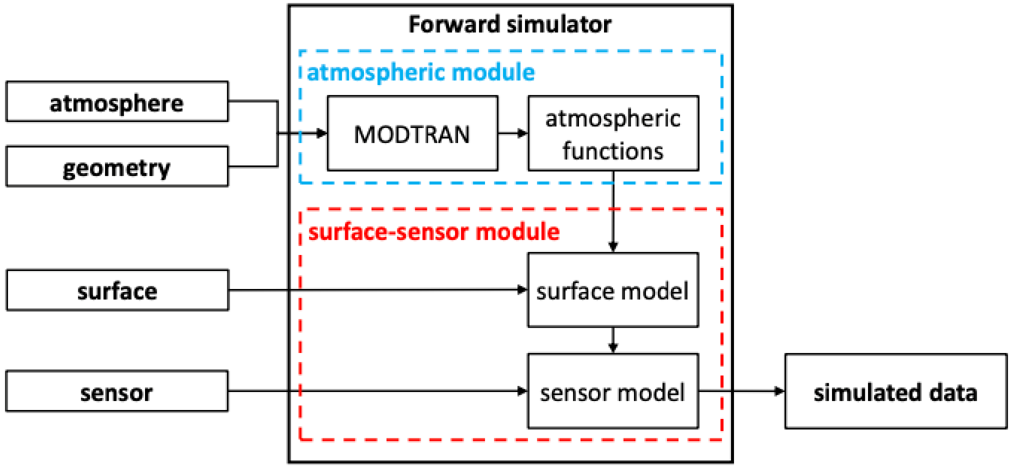
Buffat et al. 2024, RSE, doi: 10.1016/j.rse.2024.114596

JÜLICH Forschungszentrum

High-fidelity simulations of the O_2-A absorption band to allow RTM emulation

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

- MODTRAN6 **simulations** of **typical HyPlant** and **DESIS** observational conditions
- Fine-grained **simulated data** bases



Specification	Databases	
	DESIS	HyPlant
Input dimensions	11	13
Output dimensions	13	349
Number of samples	1.2×10^7	1.5×10^7
Data size [GB]	5.6	64.7

Pato et al. 2025, JSTARS, doi: 10.1016/j.rse.2025.114944

Emulation of bandwise sensor shifts

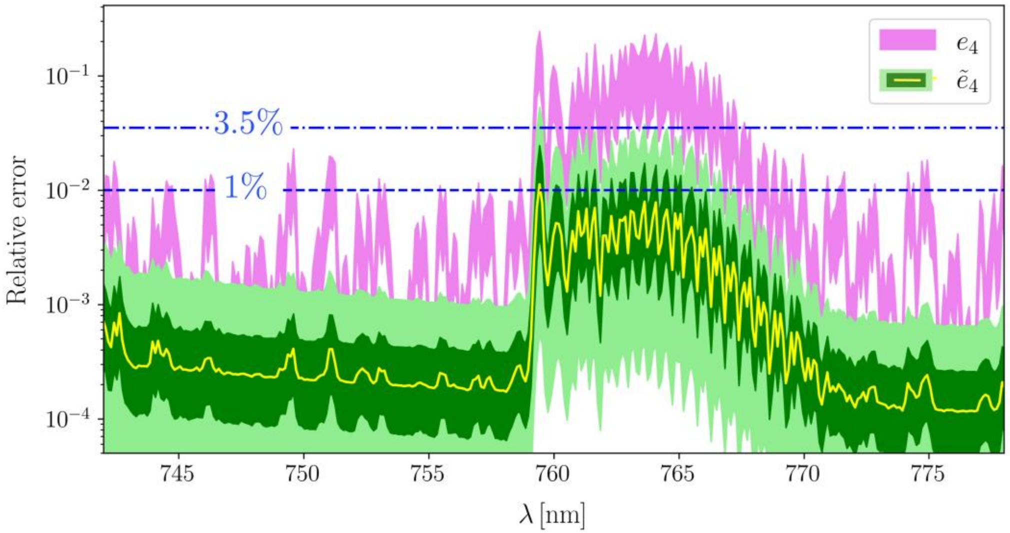
	e_4	e_4^Λ	$\tilde{e}_4$
Prediction time per sample	0.28 μs	55.40 μs	1.93 μs

No emulation

Full emulation

Approximate emulation

- Extend emulation with **bandwise sensor shifts** (wavelength & FWHM)
- **Full** representation is **inefficient**
- Adopting **bandwise approximations**



miscalib. not considered

miscalib. considered

Buffat et al. 2026, RSE, doi: 10.1016/j.rse.2025.115203

Emulation of bandwise sensor shifts

$$w\left(\lambda_i \,|\, \tilde{\mathbf{p}},\, \Delta\lambda_i, \Delta\sigma_i\right) = \frac{e_d\left(\lambda_i \,|\, \tilde{\mathbf{p}},\, \Delta\lambda_i, \Delta\sigma_i\right)}{e_d\left(\lambda_i \,|\, \tilde{\mathbf{p}},\, \Delta\lambda_i = \Delta\sigma_i = 0\right)}$$

em. with shifts

em. w/o shifts

$$m(\lambda_i \,|\, \mathbf{\Delta}\lambda, \mathbf{\Delta}\sigma) = \mathbb{E}\left[w\left(\lambda_i \,|\, \tilde{\mathbf{p}},\, \Delta\lambda_i, \Delta\sigma_i\right)\right]$$

- Extend emulation with **bandwise sensor shifts** (wavelength & FWHM)
- **Full** representation is **inefficient**
- Adopting **bandwise approximations**
  - **Small variation** in weighting factor

Correction weight