→ REMOTE SENSING OF FLUORESCENCE, PHOTOSYNTHESIS AND VEGETATION STATUS

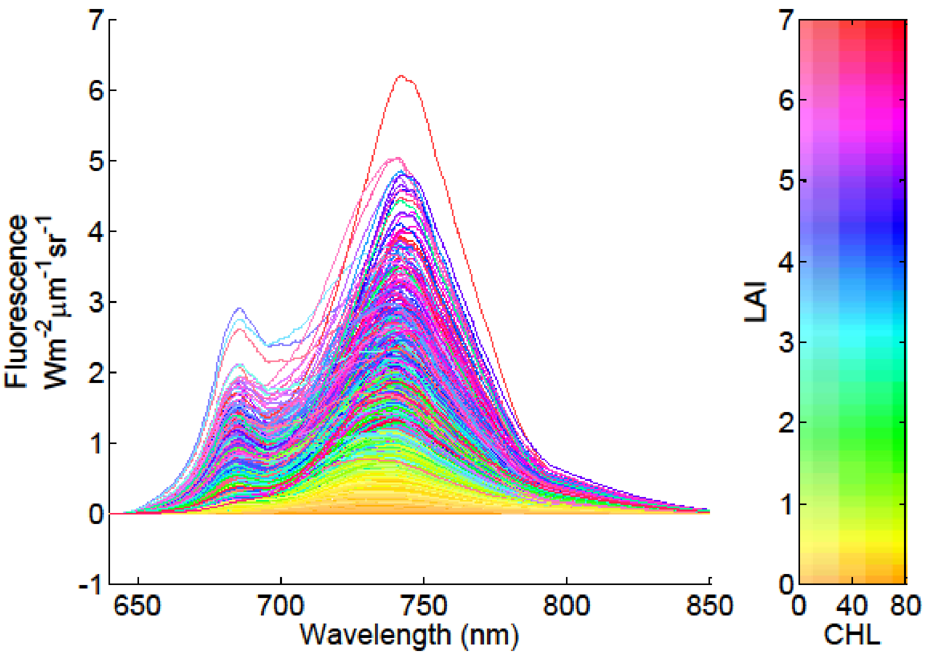
17–19 January 2017 | ESA–ESRIN | Frascati (Rome), Italy

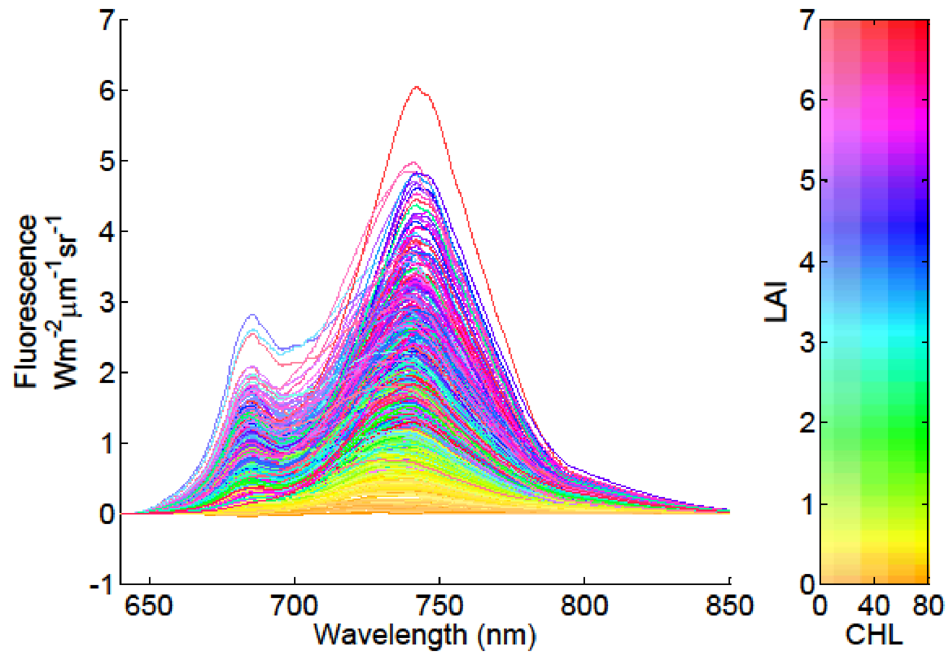
SPEEDING UP THE SIMULATION OF VEGETATION FLUORESCENCE THROUGH EMULATION: *PRACTICAL APPLICATIONS FOR FLEX DATA PROCESSING*

Jochem Verrelst, Juan Pablo Rivera & Jose Moreno
Image Processing Laboratory, Univ. of Valencia (Spain)

Any difference?

Which model would you choose?





~13 min

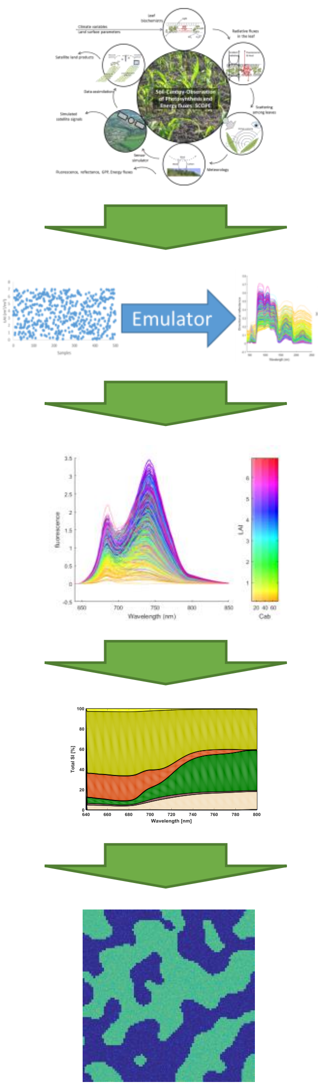
1 s.

Radiative transfer model (SCOPE)

Metamodel SCOPE (emulator)

Outline:

- **Advanced RTMs - SCOPE**
- **Metamodeling: Emulation**
- **Emulation of SCOPE**
- **Applications emulation for FLEX/SIF data processing**

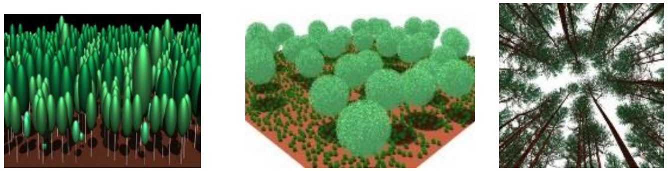


Advanced RTMs: *more realistic but slow*

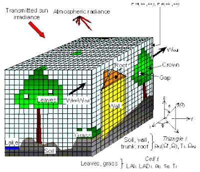
- **Radiative transfer models (RTMs)** are widely used in remote sensing science, e.g. for development of new missions and retrieval (**inversion**).
- When choosing an **RTM**, **a trade-of between invertibility and realism** has to be made: **simpler models** are easier to invert but **less realistic**, while **advanced models** more realistic but require a large amount of variables to be configured.

<u>**Examples of advanced models:**</u>

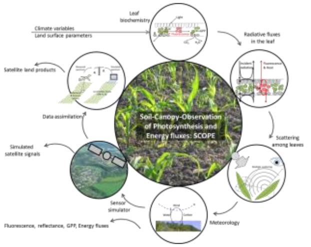
- **Ray tracing models** (e.g. FLIGHT, RAYTRAN, DRAT)



- **Voxel models**: DART



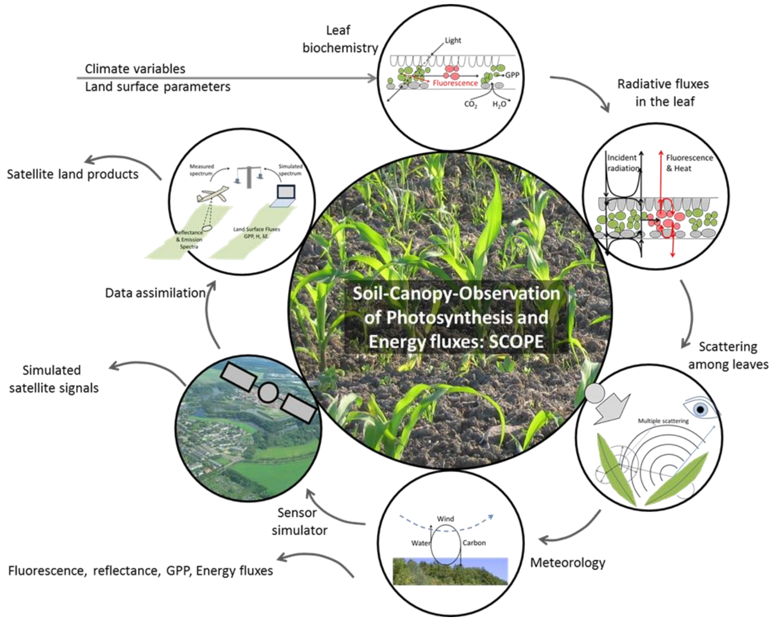
- **Soil-Vegetation-Atmosphere-Transfer (SVAT) models:** e.g. SCOPE



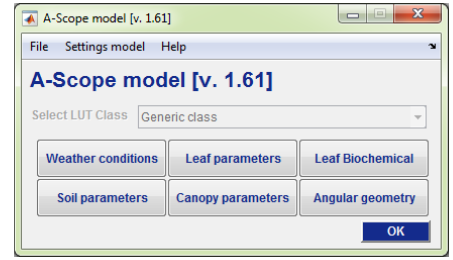
- **Main drawback of complex models involves their long processing speed**: ***the more computationally expensive, the longer it takes to generate output.***
- Long processing time makes that **advanced RTMs are of little use** for operational tasks, e.g., pixel-by-pixel retrieval schemes.

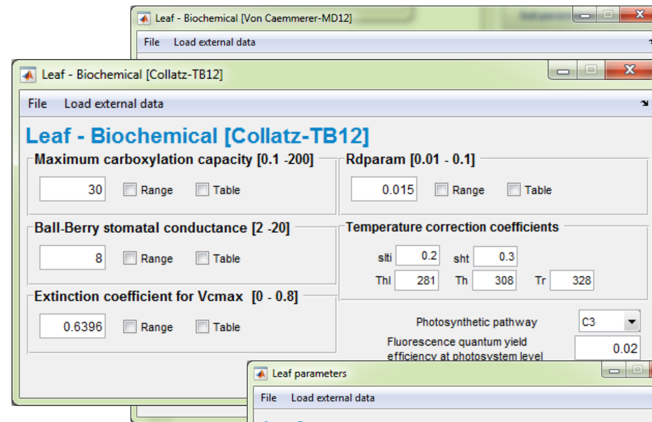
SCOPE (C. Van der Tol)

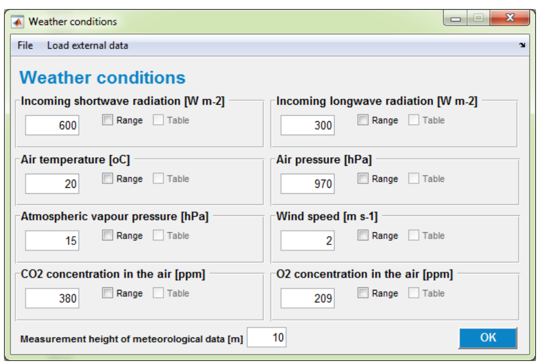
→ REMOTE SENSING OF FLUORESCENCE, PHOTOSYNTHESIS AND VEGETATION STATUS

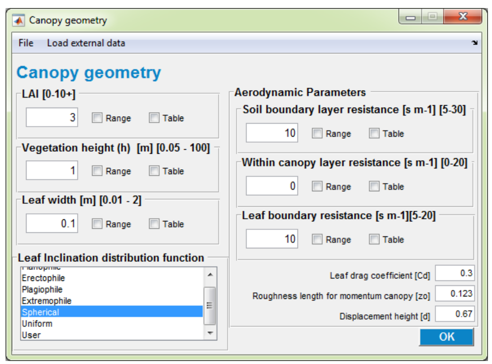


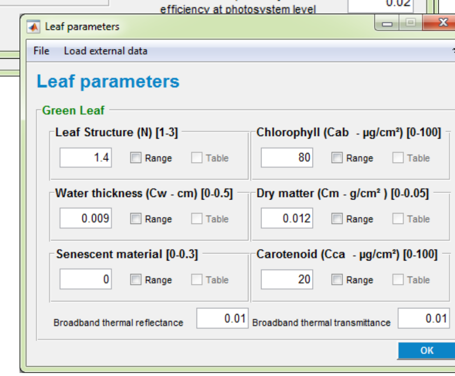
A-SCOPE:









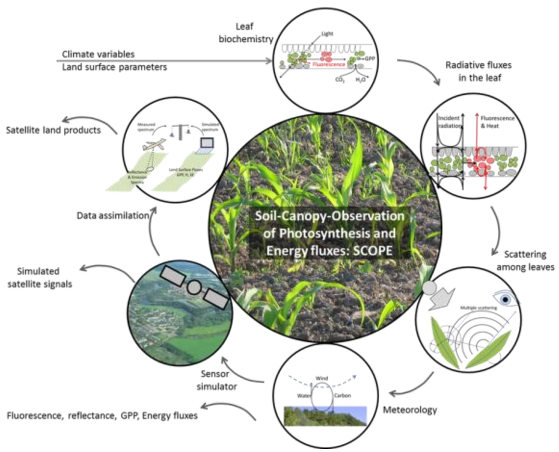


- **SCOPE generates multiple outputs, including canopy-leaving SIF. However, for operational use it is rather slow (>2 min for 100#).**
- Recently, it has been proposed **to approximate RTMs through machine learning** (Rivera et al., 2015; Gomez-Dans and Lewis, 2016).

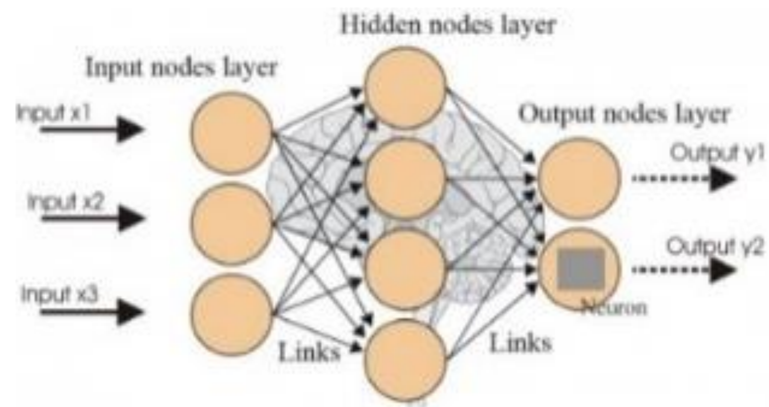
<u>Emulators</u> are surrogate statistical models that are able to approximate the processing of an RTM - <u>at a fraction of the computational cost</u>:

making a statistical model of a physical model

RTM

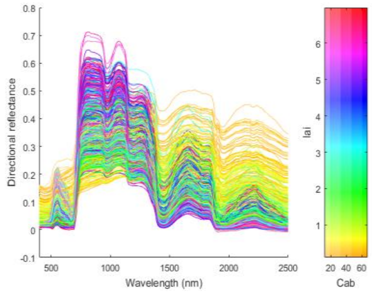


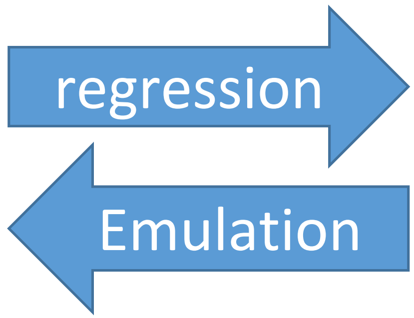
Machine learning



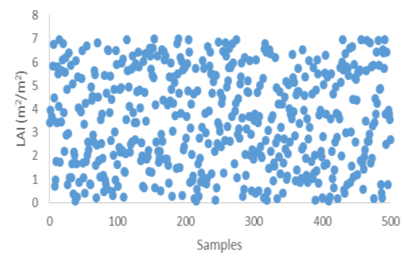
Emulation applied to RTMs:

Spectra





Variables
(e.g. LAI, chlorophyll,…)



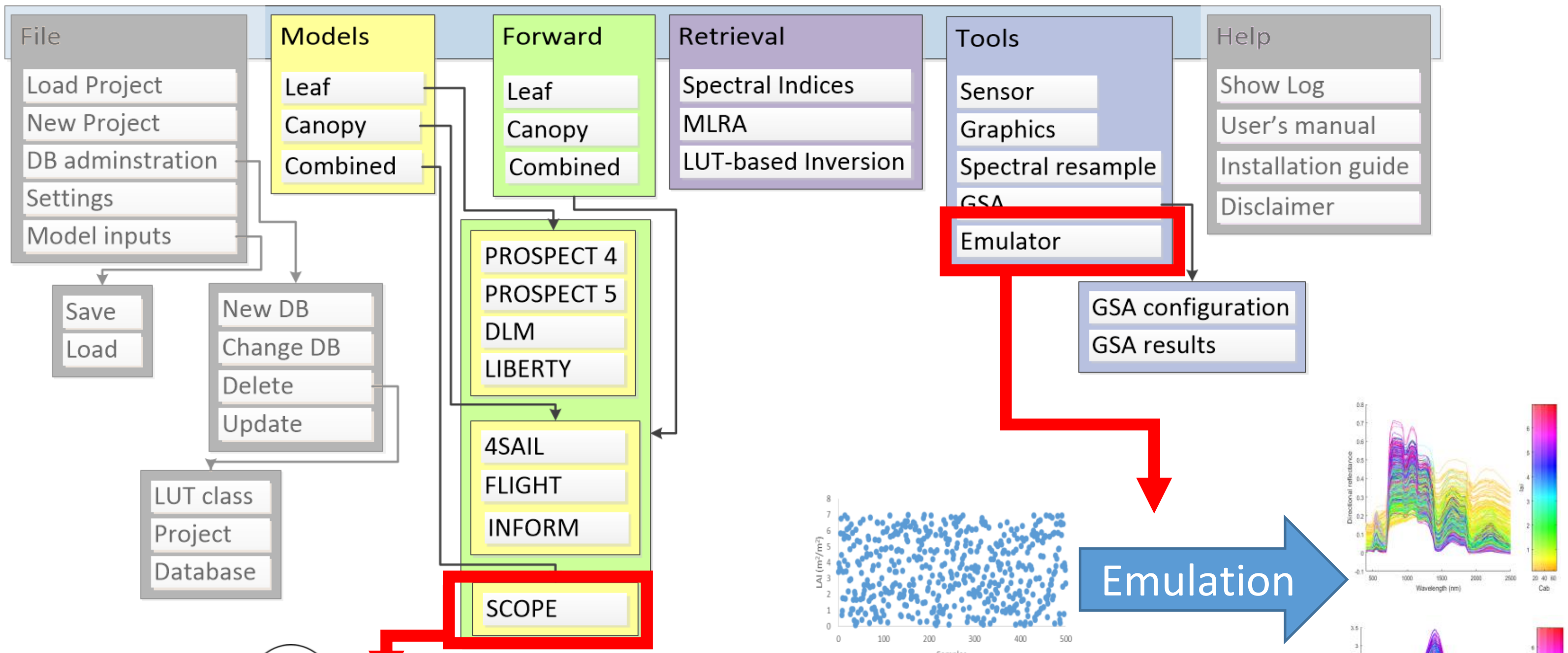
- In principle any **nonlinear, adaptive machine learning regression algorithm (MLRAs) can serve as emulator**.
- However**, to emulate RTM spectral output,** the MLRA should have the capability to reconstruct **multiple outputs**, i.e. the complete spectrum: **resolved with dimensionality reduction** techniques (e.g. PCA).

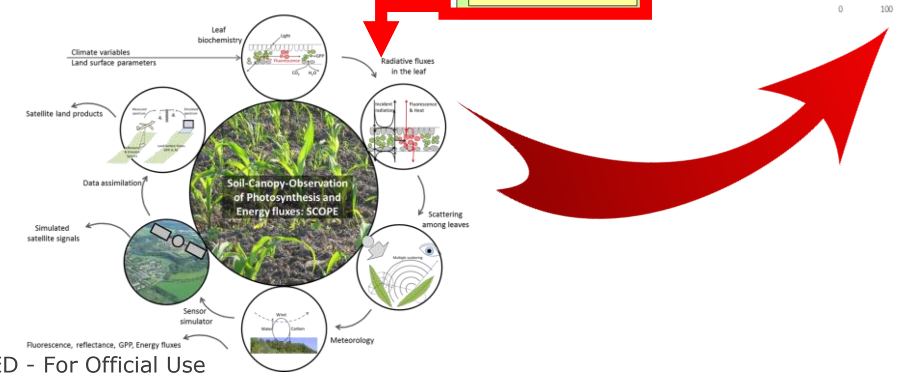
Processing steps *(Rivera et al, 2015; Verrelst et al., 2016)*

Input (variables + spectra) → Splitting into training/validation → PCA on spectra → MLRA training looping over components → Prediction of components → Reconstruction of full spectrum → Validation → Emulator

With this processing chain any RTM can be converted into an accurate emulator.

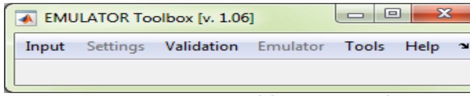
As part of *FLEX scientific studies*, the scientific software package ARTMO has been expanded with SCOPE and an Emulator toolbox.





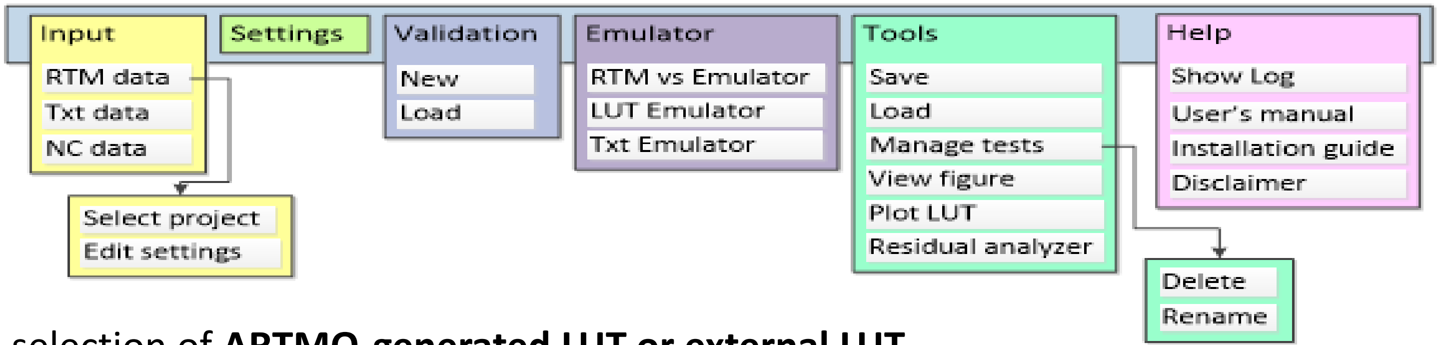
http://ipl.uv.es/artmo/

EMULATOR toolbox v. 1.06



http://ipl.uv.es/artmo/

Input → Settings → Validation → Emulator



1. <u>Input</u>: selection of **ARTMO-generated LUT or external LUT**
2. <u>Settings</u>: **selection of a MLRAs & DR method**
3. <u>Validation</u>: This step **validates the configured MLRAs** through various **goodness-of-fit-statistics**
4. <u>Emulator:</u>
   - <u>RTM vs Emulator</u>: The **chosen emulator can be tested against the actual RTM for a given input**. This module visualizes both outputs and calculates the accuracy and gain in processing speed.
   - <u>LUT Emulator</u>: This module allows the **emulation of a LUT through a chosen emulator and ranges of input variables**. The input values are then randomly selected.
   - <u>Txt Emulator</u>: This module **generates a LUT** based on input values coming **from a TXT file**.

Emulators can be exported to other ARTMO toolboxes: global sensitivity analysis, scene generation, retrieval (inversion)

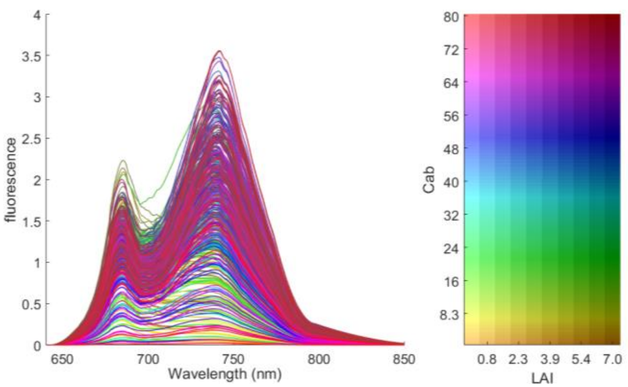
Experimental setup: emulating SCOPE

Experimental setup:

- SCOPE TB12-D: LUT 500# @ 1 nm; 12 variables
- 6 machine learning methods tested: RF, KRR, NN, GPR, SVR, VHGPR
- # 10 PCA components tested; 70/30% training/validation

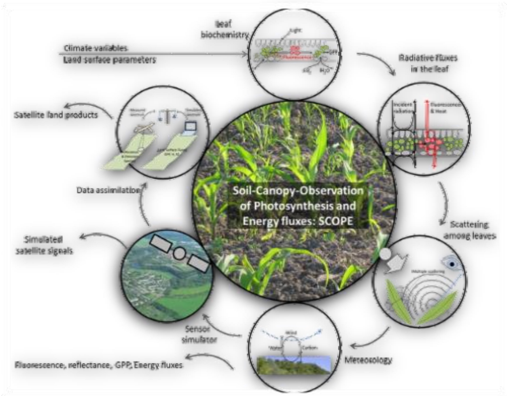
- N
- LWC
- DMC
- Cs
- LCC
- m
- kV
- Rdparam
- Vcmo
- Leaf width
- LAI
- hc

SCOPE: ⏱ 12 min (500#)



Emulator

Emulated SCOPE

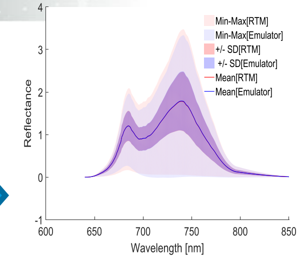


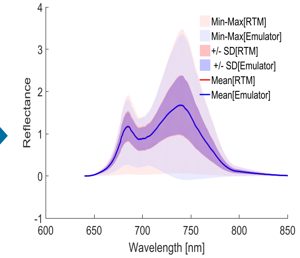
- Random Forests (RF)
- Kernel Ridge Regression (KRR)
- Neural Networks (NN)
- Support Vector Regression (SVR)
- Gaussian Processes Regression (GPR)
- Variational Heteroscedastic GPR (VH-GPR)

esa

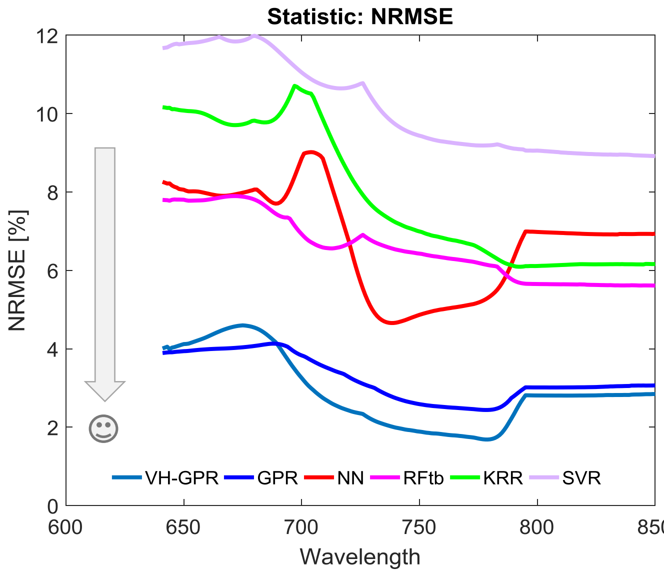
Goodness-of-fit results (30%)

Method	RMSE	NRMSE (%)	CPU (s)
VH-GPR	**0.59**	**0.25**	**91**
GPR	0.70	0.30	33
NN	1.38	0.57	16
RF	1.62	0.68	23
KRR	**1.80**	**0.76**	**1**
SVR	2.42	1.02	235

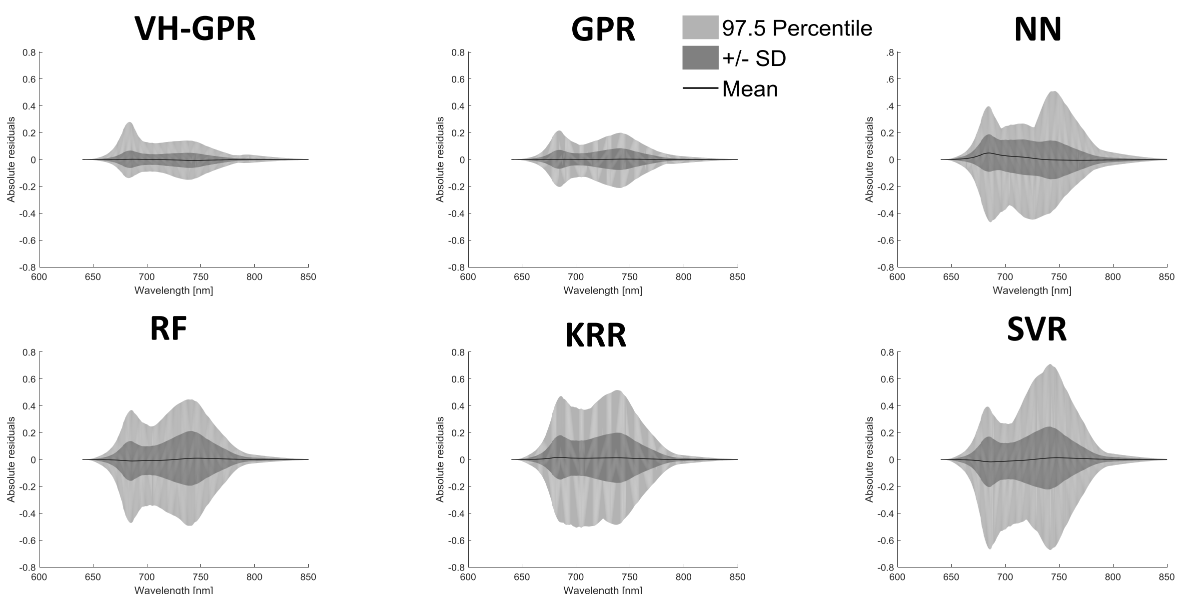




Errrors:

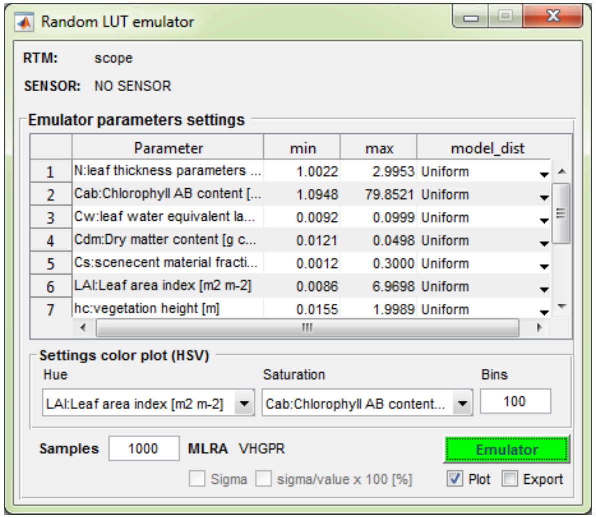


Stats residuals:

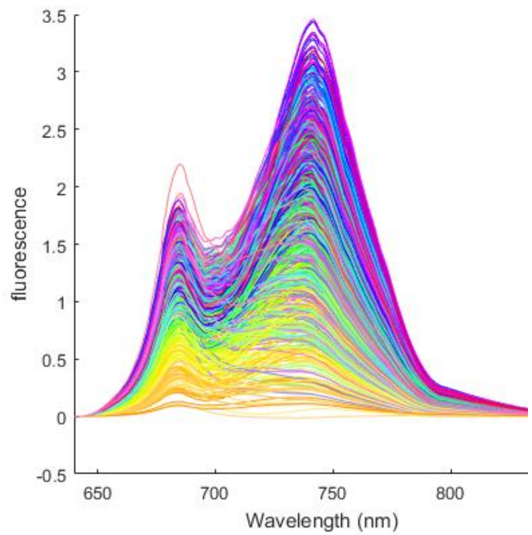


LUT emulator (# 1000)

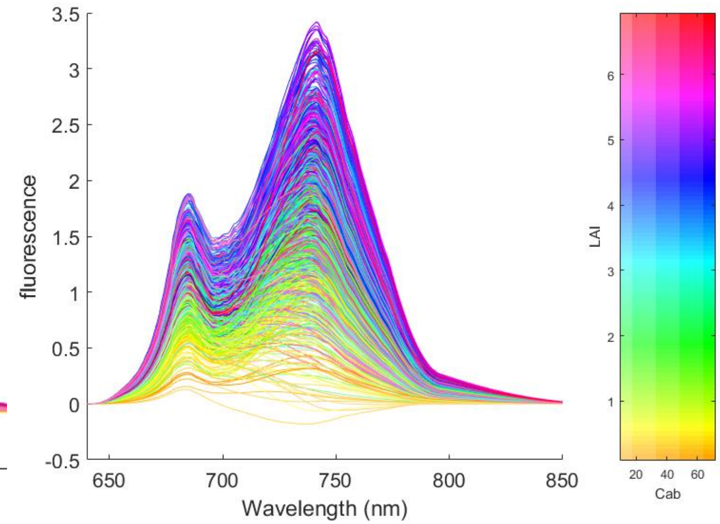
Randomly generated LUT by varying all variables.



VH-GPR: ⏱ 152 s



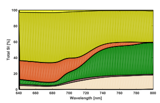
KRR: ⏱ 2 s



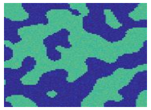
Because of the smooth profiles, SIF spectra are easy to emulate.

In Emulation, physical models go hand in hand with machine learning

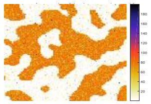
Applications Emulation for SIF Data Processing



1. Global Sensitivity Analysis



2. Scene generation

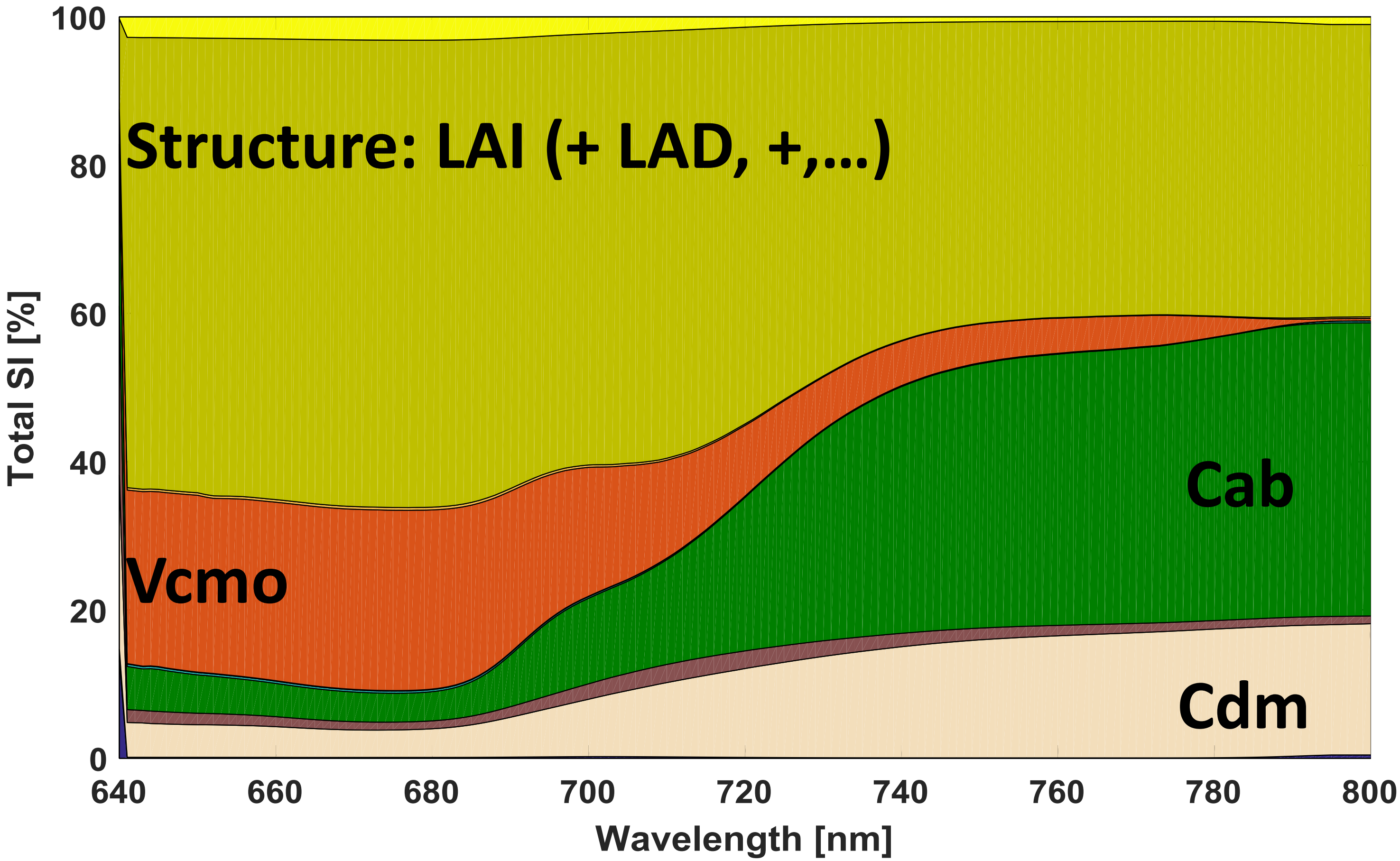


3. Retrieval (inversion)

after Verrelst et al., RSE 2015

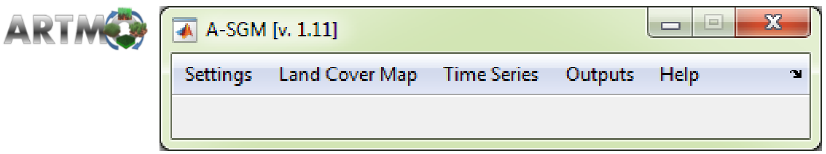
<1 min : (12 min # 500 simulations, 90 s VH-GPR emulator training, 30 s. GSA processing (1000# per variable)

- N
- LWC
- DMC
- Cs
- LCC
- m
- kV
- Rdparam
- Vcmo
- Leaf width
- LAI
- hc



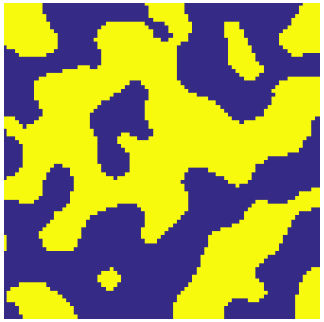
GSA results are consistent with original SCOPE GSA results (Verrelst *et al.*, 2015).

A-SGM: Automated Scene Generation Module

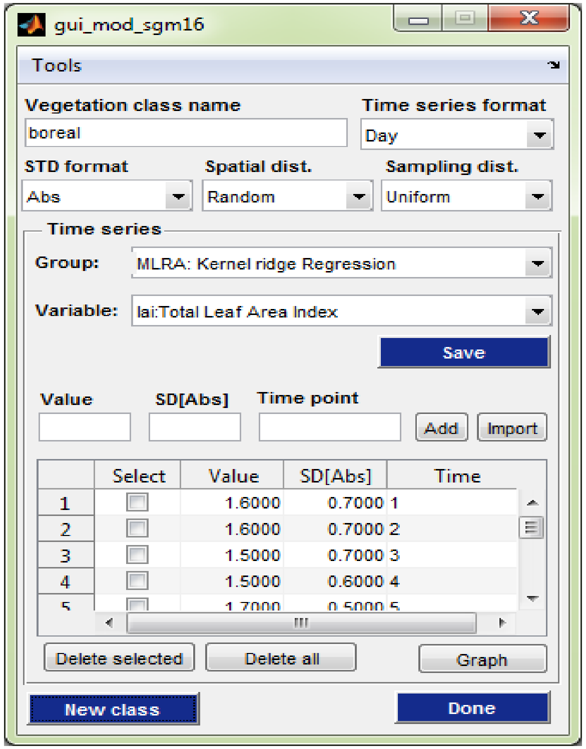


A randomly generated 2-class image

(scene size 10 km^2; pixel: 30 m^2; 334x334=111,556 pixels)



Settings vegetation classes (e.g. PFTs):

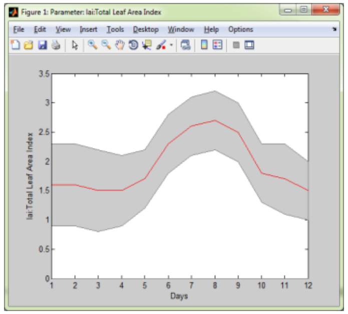


An **RTM** or **Emulator** can be configured with time series of RTM input ranges.

Ranging variables:

- ✓ LAI
- ✓ Cab
- ✓ Vcmo
- ✓ Vegetation height

Input temporal profile: e.g. LAI

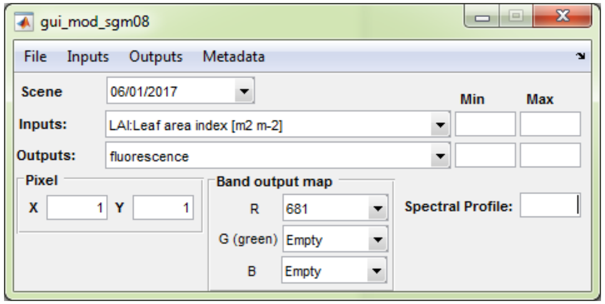


Scene generation: time series input + output

GPR: < 3 hours

KRR: < 4 min.

(SCOPE: > 2days)



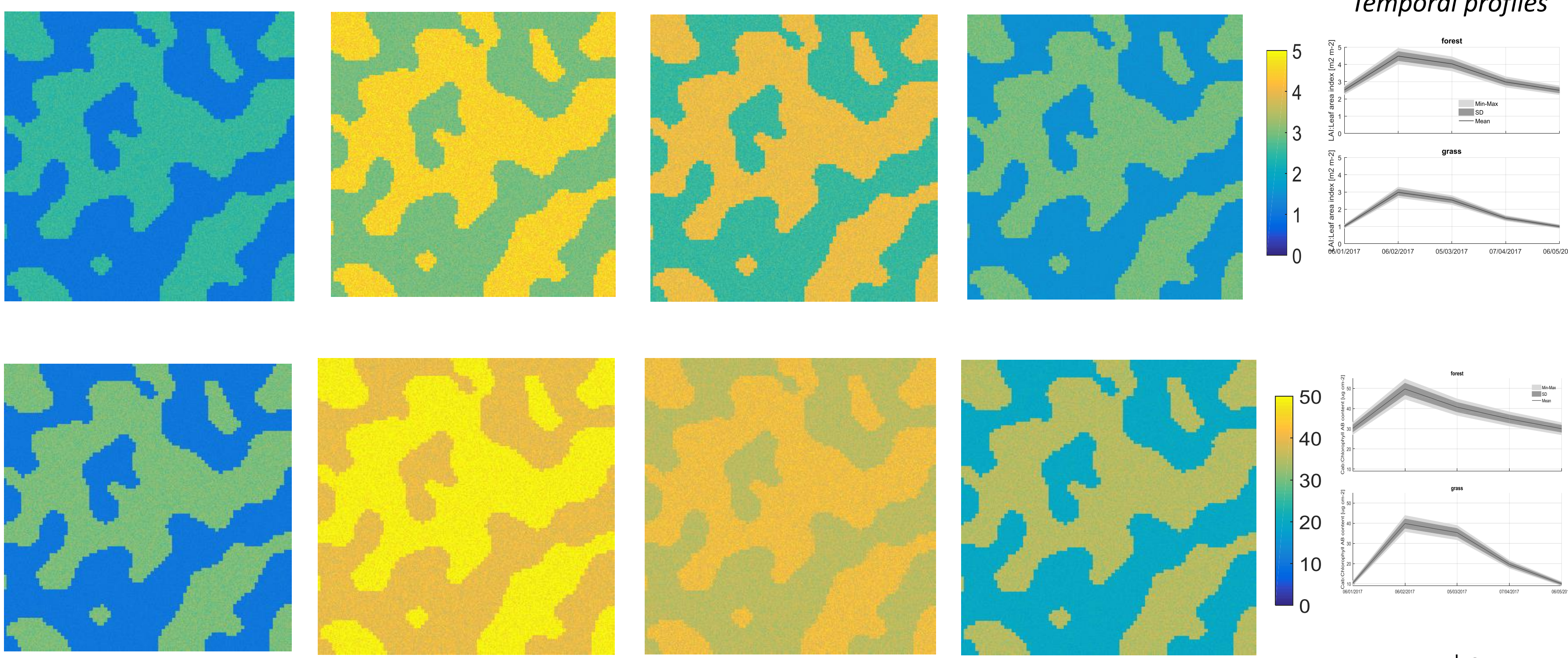
Input maps:

Time 1 | Time 2 | Time 3 | Time 4

Temporal profiles

LAI

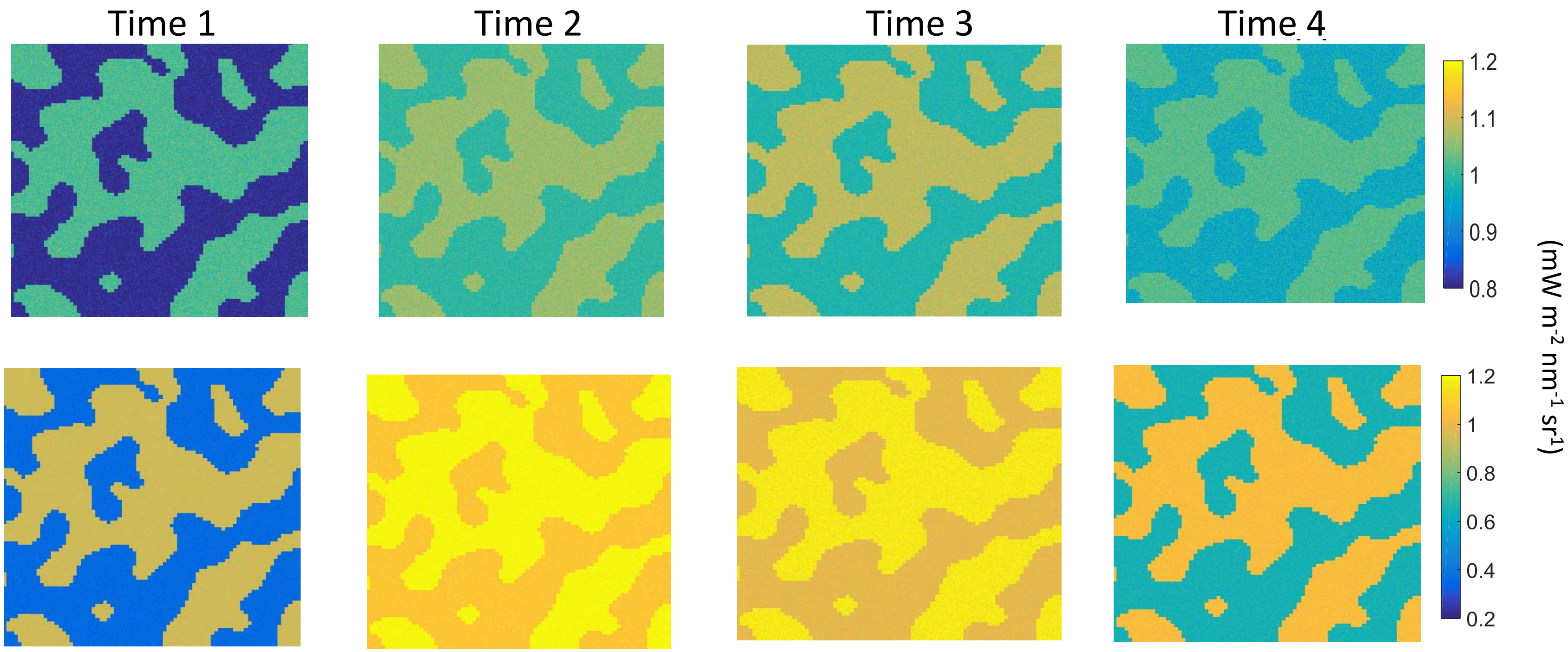
Cab



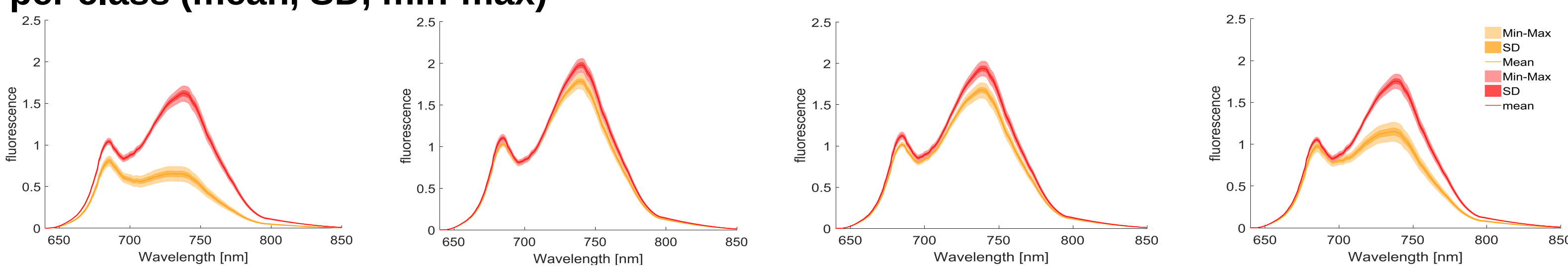
SIF output

O_2-B: 687

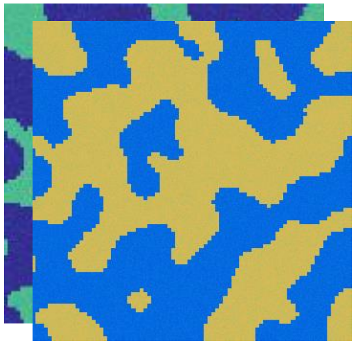
O_2-A: 760



Statistics per class (mean, SD, min-max)



O_2-B, O_2-A SIF bands

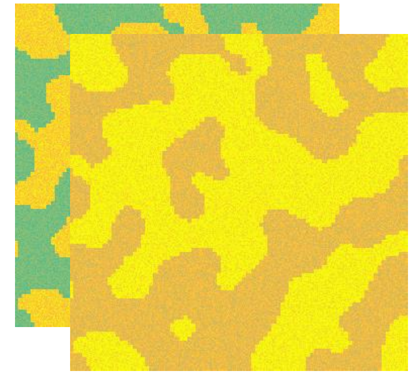


Retrieval (inversion)

Nonlinear least squares optimization

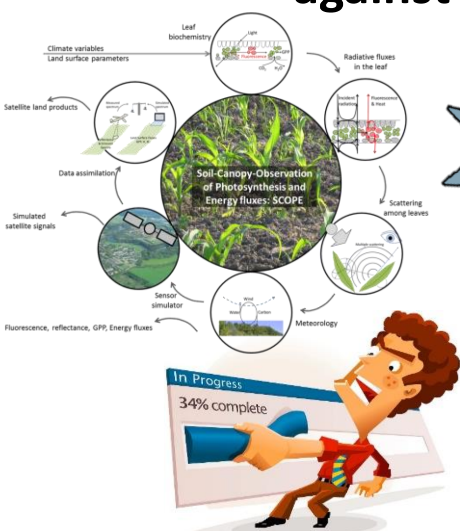
Model

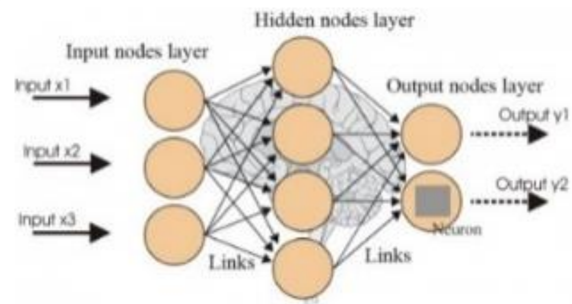
SIF products
(e.g. photosynthesis)



Spectral fitting against <u>RTM</u>

Spectral fitting against <u>emulator</u>



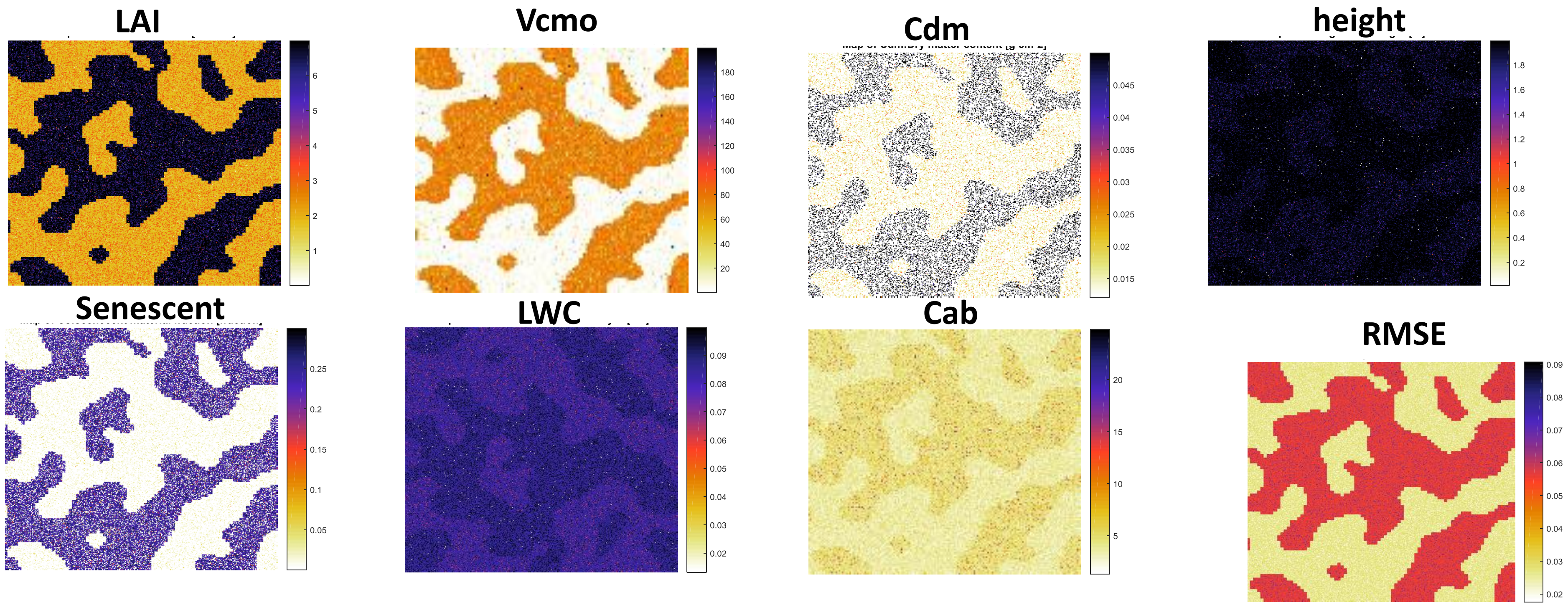


3) NUMERICAL INVERSION: Spectral Fitting (2/2)

→ REMOTE SENSING OF FLUORESCENCE, PHOTOSYNTHESIS AND VEGETATION STATUS

Spectral fitting against O_2B, O_2A and 2 peaks (F_{685}, F_{687}, F_{740}, F_{760})

KRR: < 6 hours.



- **Inversion enables retrieval of all input variables: some variables easier than others: see GSA. *With 2 SIF bands only (O_2B, O_2A), the inversion did not work.* Full SIF signal preferred.**
- **Next : *to include higher-level products into inversion: photosynthesis, GPP, APAR,…***

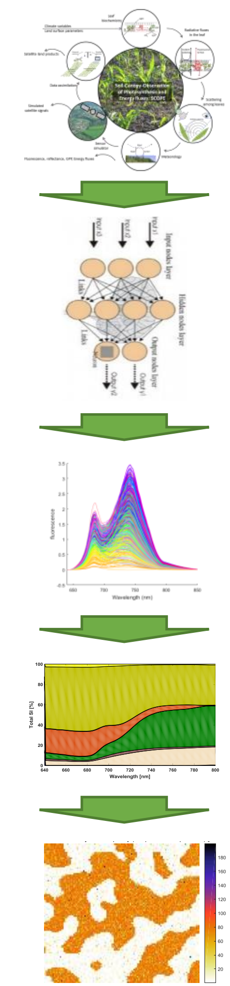
Findings:

- Thanks to machine learning regression algorithms, **RTMs can be emulated**.
- **Trade-off:** **tremendous gain in processing speed** at expense of some loss in accuracy.

- **Emulation allows applying advanced RTMs into tedious, operational processing chains**:
  1. **Global sensitivity analysis** (SCOPE[1], MODTRAN[2])
  2. **Scene generation** (A-SGM[3])
  3. **Retrieval** (numerical inversion)

Future work:

- **Emulation** of SCOPE & MODTRAN will enable **fast processing** of **FLEX E2E simulator** and **retrieval of higher level SIF products**.



[1] Verrelst, J., Rivera, J.P., van der Tol, C., Magnani, F., Mohammed, G., Moreno, J. (2015). Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence? Remote Sensing of Environment, 166, 8-21.

[2] Verrelst, J., Sabater, N., Rivera, J.P., , Muñoz-Marí, J., Vicent, J., Camps-Valls, G., Moreno, J. (2016). Emulation of Leaf, Canopy and Atmosphere Radiative Transfer Models for Fast Global Sensitivity Analysis. Remote Sensing. 8(8), 673.

[3] FLEX Bridge Study, ESA-ESTEC

→ REMOTE SENSING OF FLUORESCENCE, PHOTOSYNTHESIS AND VEGETATION STATUS

17–19 January 2017 | ESA–ESRIN | Frascati (Rome), Italy

Thanks

Jochem Verrelst, Juan Pablo Rivera & Jose Moreno
Image Processing Laboratory, Univ. of Valencia (Spain)