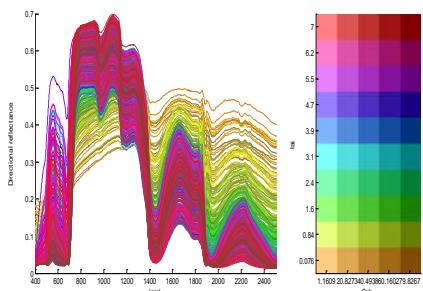


From model simulations towards vegetation properties mapping:

automating, optimizing & simplifying



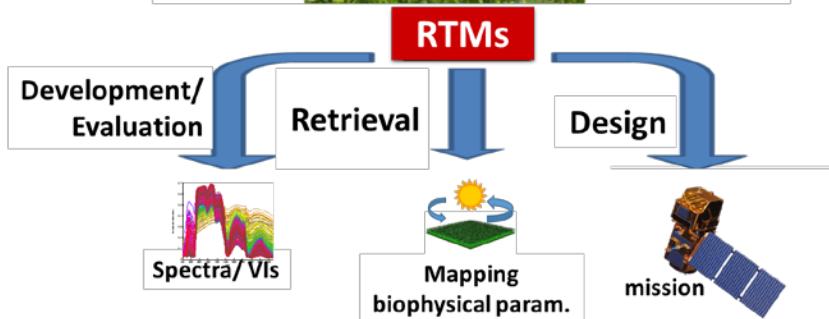
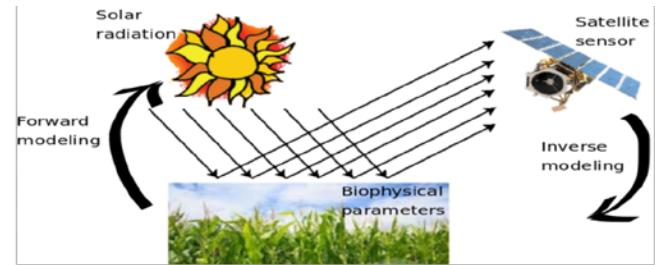
J. Verrelst, J.P. Rivera & J. Moreno

ISSI Workshop – 21-25 Nov2016

Background

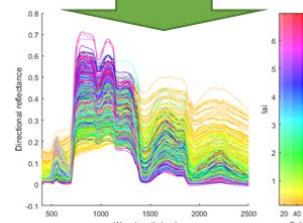
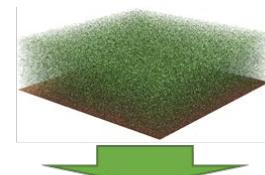
AUTOMATING RTMs

- RTMs
- ARTMO/forward
- Retrieval toolboxes



OPTIMIZING exploiting spectroscopy data

- Retrieval
- Band selection
- Dimensionality & sample reduction



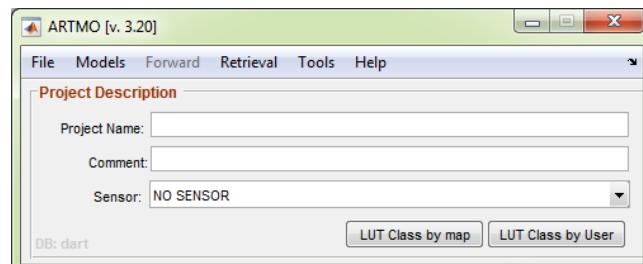
SIMPLIFYING RTMs

- Global sensitivity analysis
- Emulation
- Retrieval



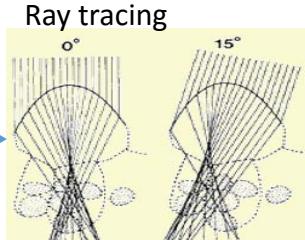
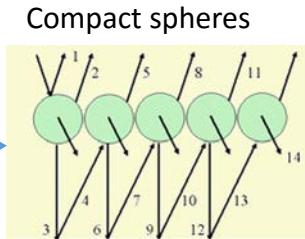
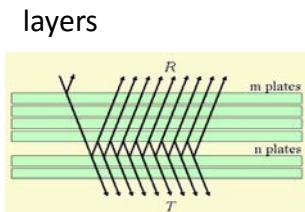
AUTOMATING

- RTMs
- ARTMO/forward
- Retrieval toolboxes



Radiative transfer models (RTMs)

Leaf RTMs



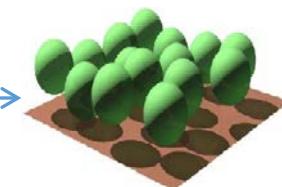
Canopy RTMs



Turbid medium



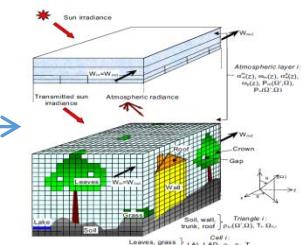
Geometric



Hybrid



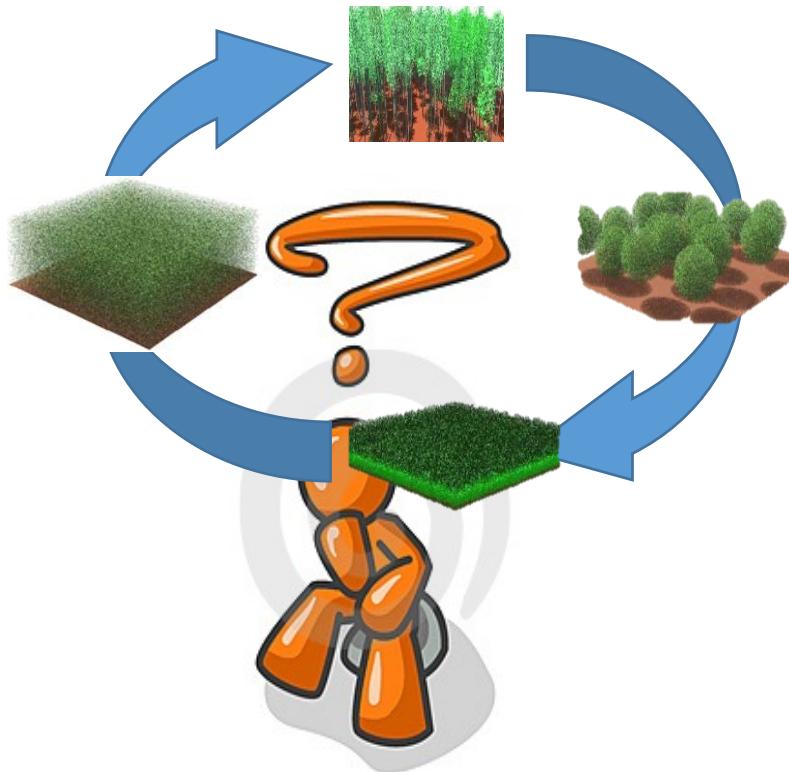
Volumetric



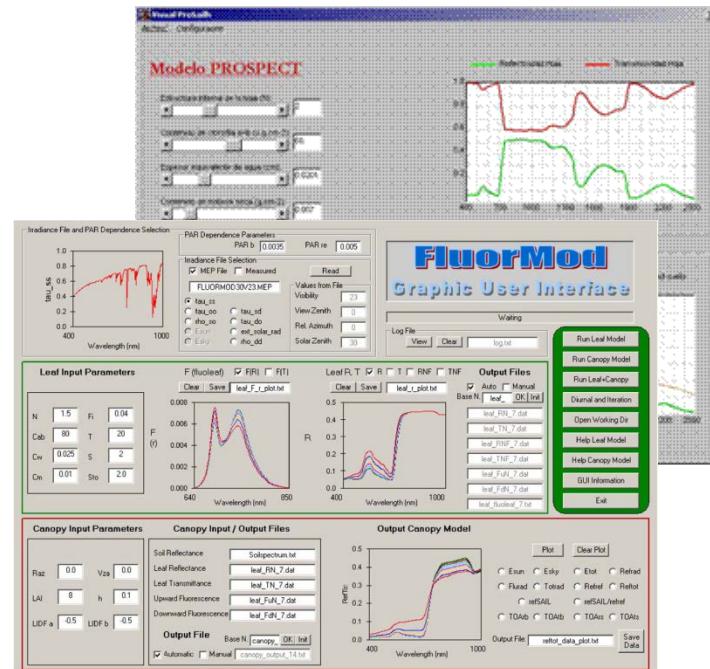
A diversity of RTMs exist with different complexity.

RTMs are important tools in EO research, but not always easy to use. Only very few of them offer user-friendly interfaces.

Which RTM to choose?



Only very few offer a GUI.



- No interface exists that brings multiple RTMs together in one GUI.
- None of existing (publicly available) GUIs provide post-processing tools.

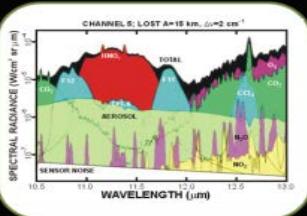
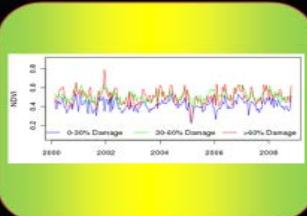
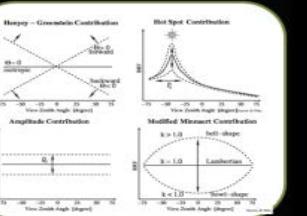
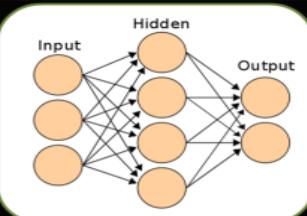
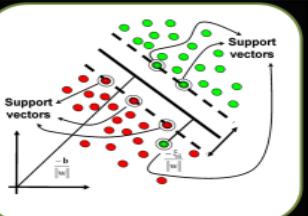
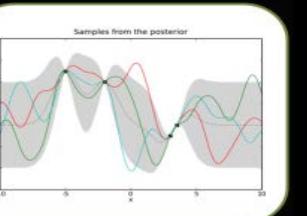
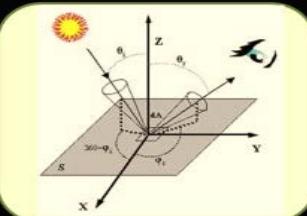
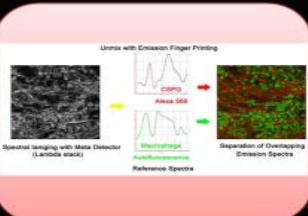
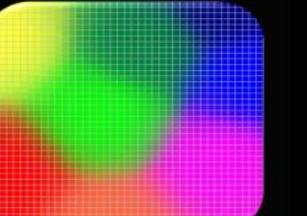
Toolbox for EO applications:



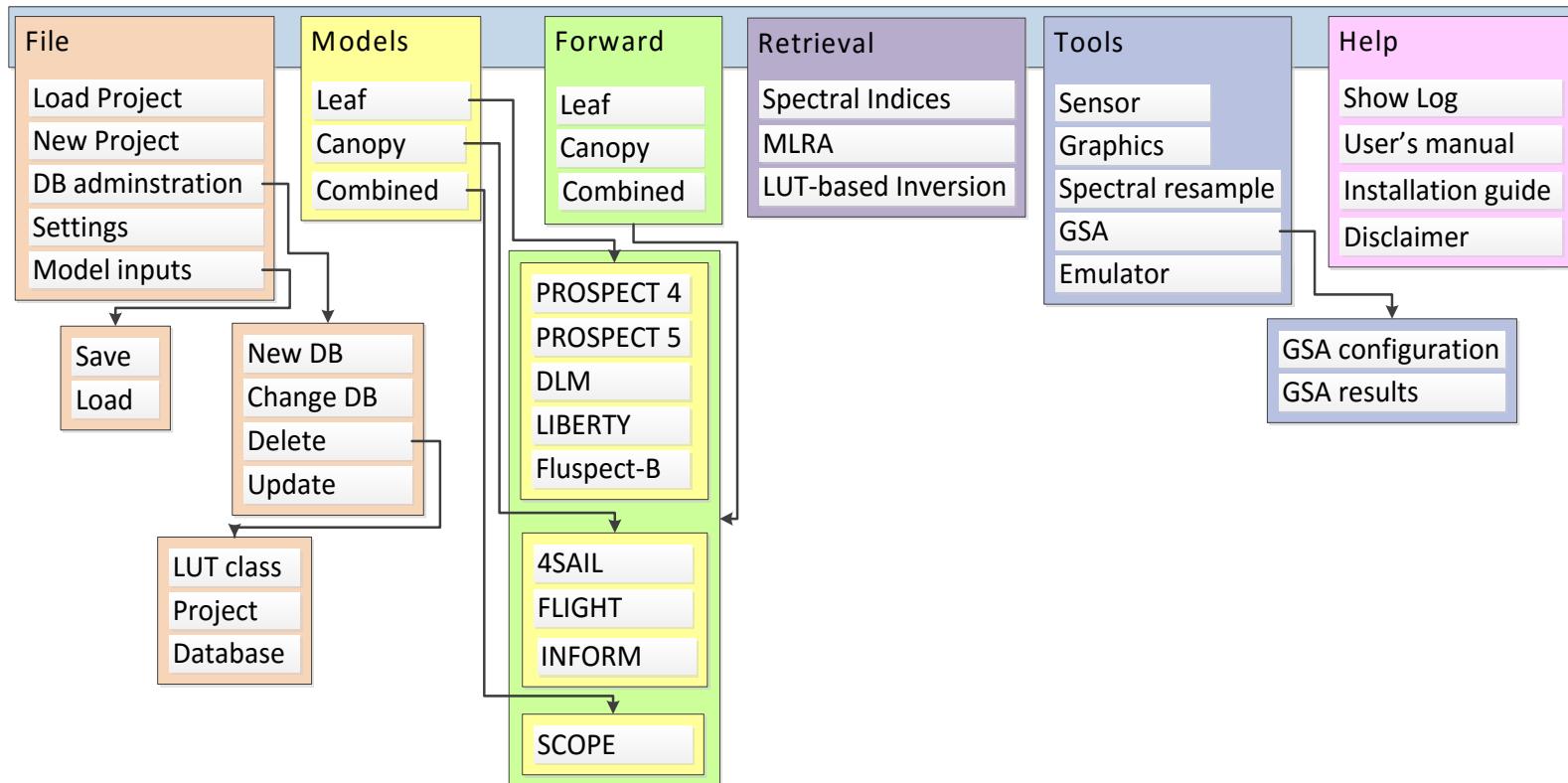
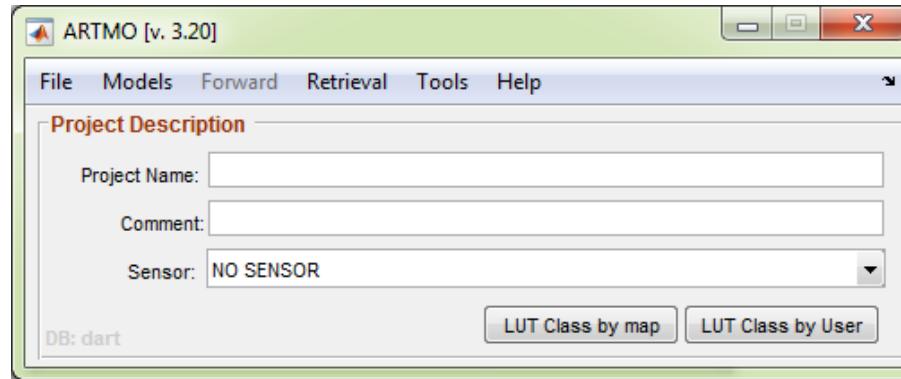
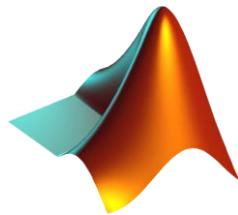
Automated
Radiative
Transfer
Models
Operator



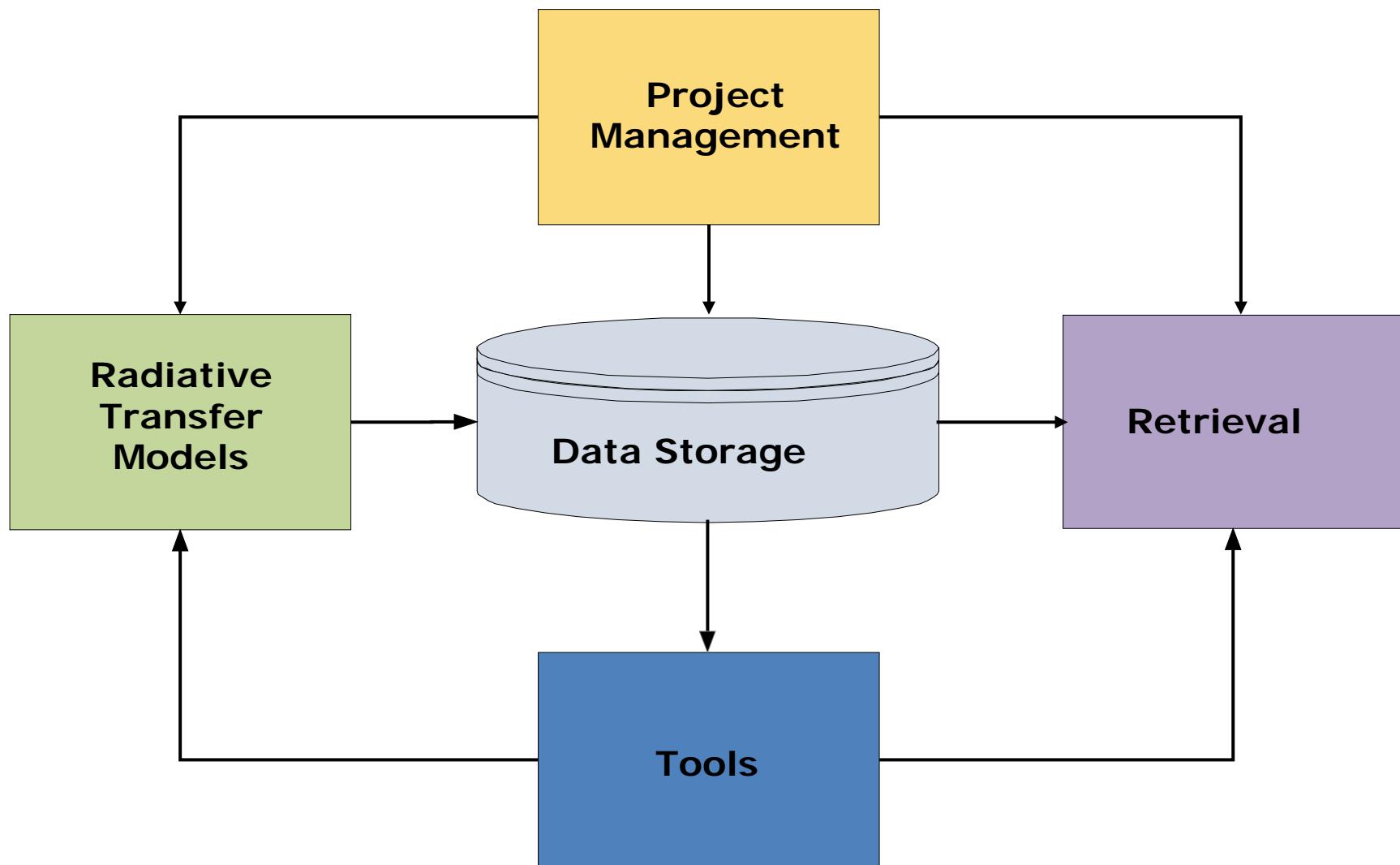
ARTMO

 <p>Atmospheric models</p>	 <p>MODTRAN</p>	 <p>Vegetation indices</p>
 <p>Time series analysis</p>	 <p>Ray tracing model</p>	 <p>RPV model</p>
 <p>Neural nets</p>	 <p>Support vectors</p>	 <p>Gaussian Processes</p>
 <p>BRDF apps</p>	 <p>Spectral unmixing</p>	 <p>Classifiers</p>

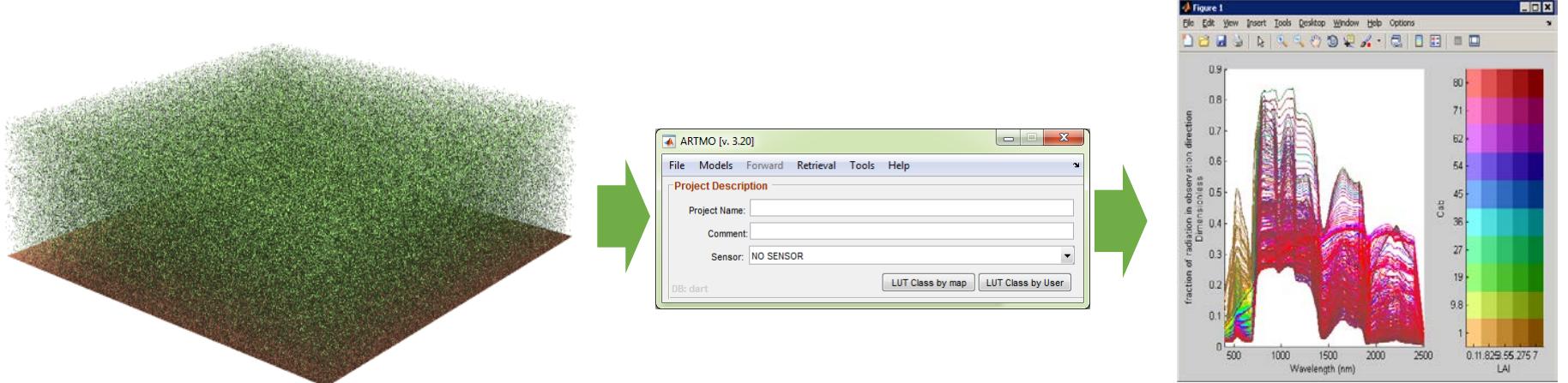
ARTMO v. 3: modular design



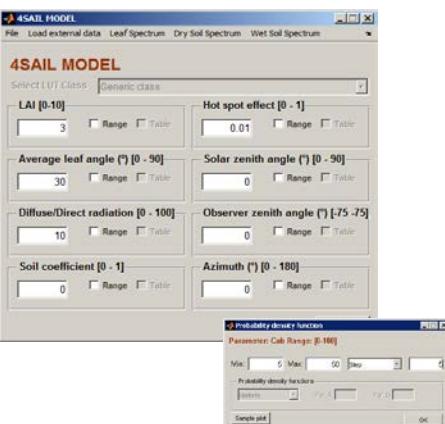
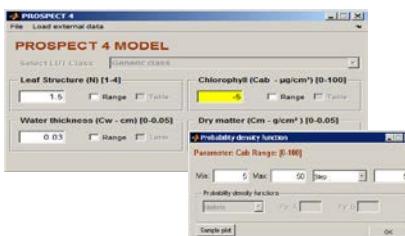
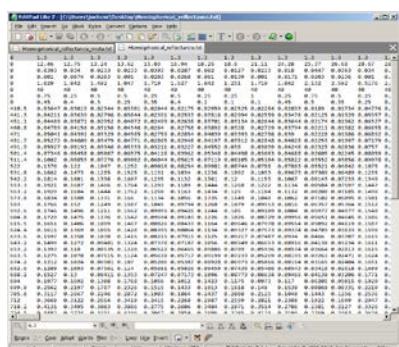
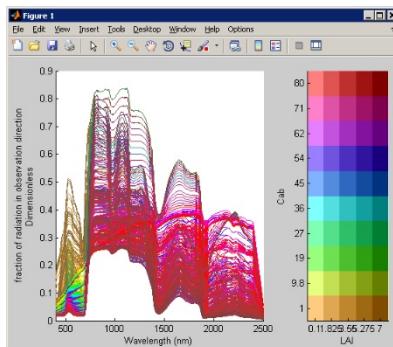
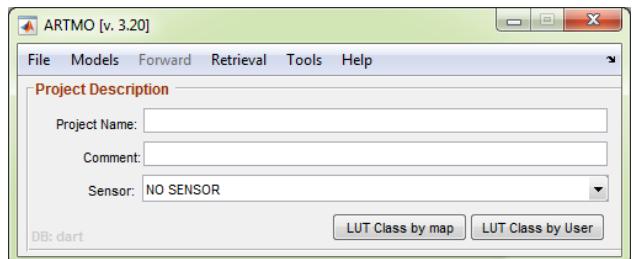
Conceptual architecture ARTMO



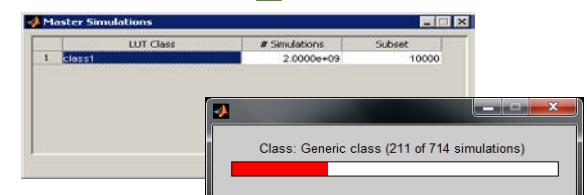
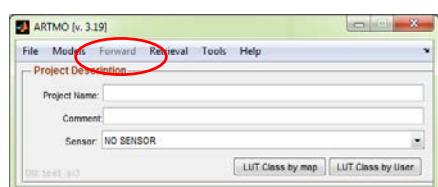
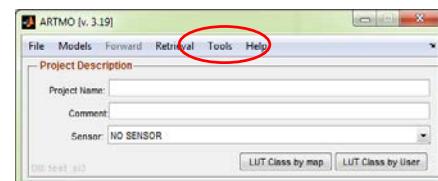
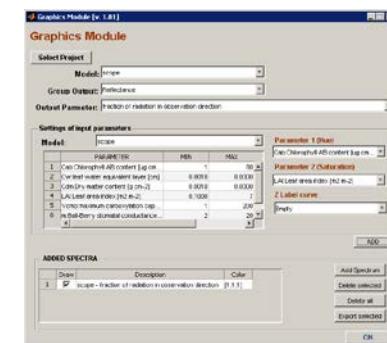
Forward



RTM outputs only a few clicks away...



Data flow



Radiative Transfer Models

ARTMO's leaf models

PROSPECT-4

PROSPECT 4

Select LUT Class Generic class

Leaf Structure (N) [1-4] Chlorophyll (Cab - $\mu\text{g}/\text{cm}^2$) [0-0.6]

Water thickness (Cw - cm) [0-0.05] Dry matter (Cm - g/cm²) [0-0.05]

New!

DLM

Dorsiventral Leaf Model - DLM [v. 2.10]

Select LUT Class Generic class

MODE Normal reflectance

Biochemical contents

- Chlorophyll ($\mu\text{g}/\text{cm}^2$) [0-150]: 50
- Carotenoids ($\mu\text{g}/\text{cm}^2$): 0
- Brown Pigments [0-5]: 0
- Water thickness (cm⁻¹) [0-0.05]: 0.02
- Dry matter (g/cm²) [0-0.05]: 0.01

Structure parameters

- f. air spaces [0-1]: 0.5
- f. total mass in pal. [0-1]: 0.012
- Abaxial scattering [0 - 0.99]: 0
- f. Pigm. in palisade [0-1]: 0.003
- roughness factor [0-2]: 0

OK

LIBERTY

LIBERTY

Select LUT Class Generic class

Cell Diameter [20-100]: 40

Chlorophyll content ($\mu\text{g}/\text{cm}^2$) [0-600]: 200

Intercellular air space [0.01-0.1]: 0.045

Water content (g/m²) [0-500]: 100

Leaf thickness [1-10]: 1.6

Lignin & Cellulose (g/m²) [10-80]: 40

Baseline [Fresh: 0.0006 - Dry: 0.0004]: 0.0006

Nitrogen content (g/m²) [0.3-2]: 1

Albino absorption [0-4]: 2

OK

PROSPECT-5

PROSPECT 5

Select LUT Class Generic class

Leaf Structure (N) [1-4] Chlorophyll (Cab- $\mu\text{g}/\text{cm}^2$) [0-100]

Carotenoids ($\mu\text{g}/\text{cm}^2$) Brown Pigments

Water thickness (Cw-cm) [0-0.05] Dry matter (Cm - g/cm²) [0-0.05]

OK

Fluspect-B

Fluspect-B [v. 1.00]

Select LUT Generic class

Weather conditions Leaf optical Leaf Biochemical

New!

OK

Meteorological data

File Load external data Irradiance spectrum

Meteorological

Incoming shortwave radiation [W m⁻²]: 600

Oxygen pressure [hPa]: 970

Leaf temperature [°C]: 20

Vapour pressure [hPa]: 15

CO concentration in the air [ppm]: 380

Air pressure [hPa]: 209

OK

Leaf parameters

File Load external data Optical Parameters

Leaf optical

Leaf Structure (N) [1-3]: 1.4

Chlorophyll (Cab - $\mu\text{g}/\text{cm}^2$) [0-100]: 80

Water thickness (Cw - cm) [0-0.5]: 0.009

Dry matter (Cm - g/cm²) [0-0.05]: 0.012

Senescent material [0-0.3]: 0

Carotenoid (Cca - $\mu\text{g}/\text{cm}^2$) [0-100]: 20

OK

Leaf biochemical

File Load external data

Leaf biochemical

Maximum carboxylation capacity [0.1 - 200]: 80

PSI 0.002 PSII 0.01

Ball-Berry stomatal conductance [2 - 20]: 6

Temperature correction coefficients

sB 0.2 sH 0.3

Th 281 Th 308 Tr 328

Rdparam [0.01 - 1]: 0.015

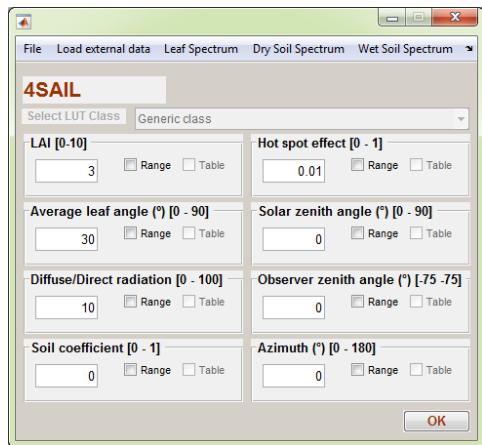
Photosynthetic Fluorescence

C3 C4

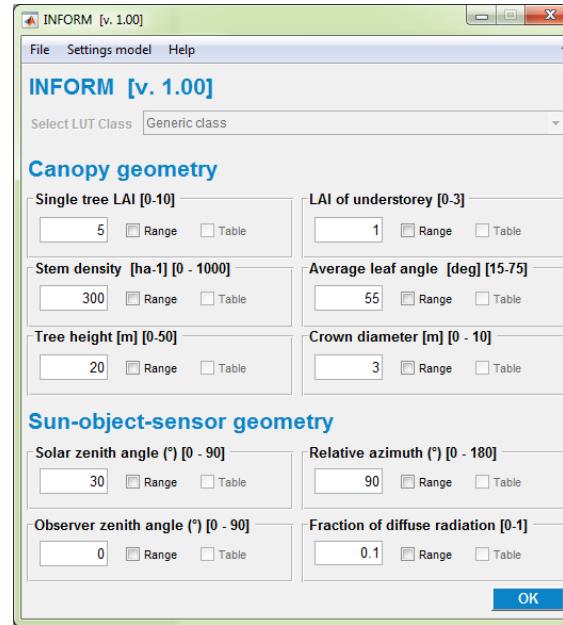
Drought fit OK

ARTMO's canopy models

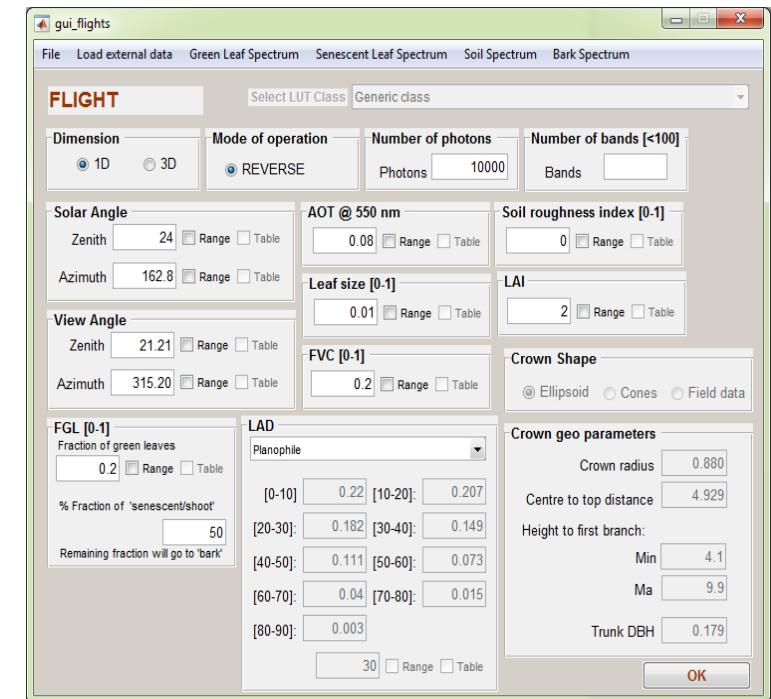
SAIL



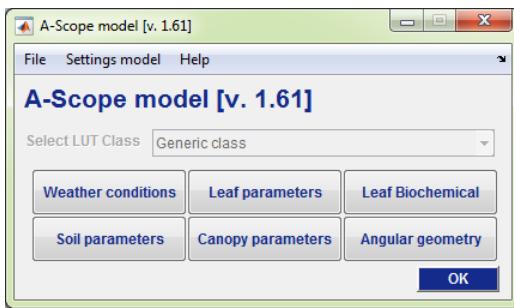
INFORM



FLIGHT



ARTMO's combined models: SCOPE



Weather conditions

File Load external data

Weather conditions

Incoming shortwave radiation [W m⁻²] 600 Range Table

Incoming longwave radiation [W m⁻²] 300 Range Table

Air temperature [°C] 20 Range Table

Air pressure [hPa] 970 Range Table

Atmospheric vapour pressure [hPa] 15 Range Table

Wind speed [m s⁻¹] 2 Range Table

CO₂ concentration in the air [ppm] 380 Range Table

O₂ concentration in the air [ppm] 209 Range Table

Measurement height of meteorological data [m] 10 Range Table

OK

Leaf parameters

File Load external data

Leaf parameters

Green Leaf

Leaf Structure (N) [1.3] 1.4 Range Table

Chlorophyll (Cab - µg/cm²) [0-100] 80 Range Table

Water thickness (Cw - cm) [0-0.5] 0.009 Range Table

Dry matter (Cm - g/cm²) [0-0.05] 0.012 Range Table

Senescent material [0-0.3] 0 Range Table

Carotenoid (Cca - µg/cm²) [0-100] 20 Range Table

Broadband thermal reflectance 0.01 Range Table

Broadband thermal transmittance 0.01 Range Table

OK

Leaf - Biochemical [Von Caemmerer-MD12]

File Load external data

Leaf - Biochemical [Von Caemmerer-MD12]

Photosynthetic pathway

Maximum carboxylation capacity [0.1 - 200] 30 Range Table

Rdparam [0.01 - 0.1] 0.015 Range Table

Ball-Berry stomatal conductance [2 - 20] 8 Range Table

Temperature correction coefficients

slt: 0.2 Range Table

sht: 0.3 Range Table

Thl: 281 Th: 308 Tr: 328 Range Table

Extinction coefficient for Vcmax [0 - 0.8] 0.6396 Range Table

Photosynthetic pathway C3 Range Table

Fluorescence quantum yield efficiency at photosystem level 0.02 Range Table

OK

Soil Parameters

File Load external data

Soil Parameters

ID_soil

Soil_ColumnID1
Soil_ColumnID2
Soil_ColumnID3

Soil resistance for evaporation [2000 s m⁻¹] 500 Range Table

Volumetric soil moisture content [0.01 - 0.7] 0.25 Range Table

Broadband soil reflectance in the thermal range 0.06 Range Table

Volumetric heat capacity of the soil [J m⁻² K⁻¹] 1180 Range Table

Specific mass of the soil [kg m⁻³] 1800 Range Table

Heat conductivity of the soil [W m⁻¹ K⁻¹] 1.55 Range Table

OK

Canopy geometry

File Load external data

Canopy geometry

LAI [0-10+]

3 Range Table

Vegetation height (h) [m] [0.05 - 100]

1 Range Table

Leaf width [m] [0.01 - 2]

0.1 Range Table

Aerodynamic Parameters

Soil boundary layer resistance [s m⁻¹] [5-30]

10 Range Table

Within canopy layer resistance [s m⁻¹] [0-20]

0 Range Table

Leaf boundary resistance [s m⁻¹] [5-20]

10 Range Table

Leaf Inclination distribution function

Monomorphic
Erectophile
Plagiophile
Extremophile
Spherical
Uniform
User

Leaf drag coefficient (Cd) 0.3 Range Table

Roughness length for momentum canopy [zo] 0.123 Range Table

Displacement height (d) 0.67 Range Table

OK

Angular geometry

File Load external data

Angular geometry

Solar zenith angle (°) [0 - 90]

30 Range Table

Observer zenith angle (°) [0 - 90]

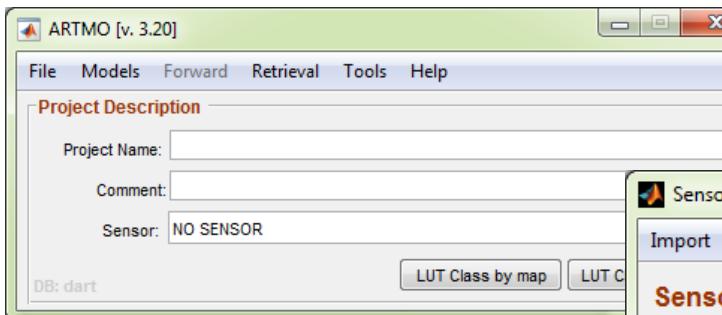
0 Range Table

Relative azimuth (°) [0 - 180]

90 Range Table

OK

Sensor



Simulations can be generated according to band settings of a selected sensor.

The screenshot shows the Sensor Module software interface. The title bar reads "Sensor Module [v. 1.04]". The menu bar includes "Import", "Edit", "Spectral Filter", and "Help". A "Sensor Information" section shows "Sentinel-2" selected. A dropdown menu on the left lists "Micrometers", "Nanometers" (which is selected), "Wavenumber", "GHz", "MHz", "Index", and "Unknown". A table titled "Band details" lists 10 bands with their properties:

	Band name	Min	Max	Center	FWHM
1	Band1	433	453	443	20
2	Band2	457.5000	522.5000	490	65
3	Band3	542.5000	577.5000	560	35
4	Band4	650	680	665	30
5	Band5	697.5000	712.5000	705	15
6	Band6	732.5000	747.5000	740	15
7	Band7	773	793	783	20
8	Band8	784.5000	899.5000	842	115
9	Band8a	855	875	865	20
10	Band9	925	945	935	20

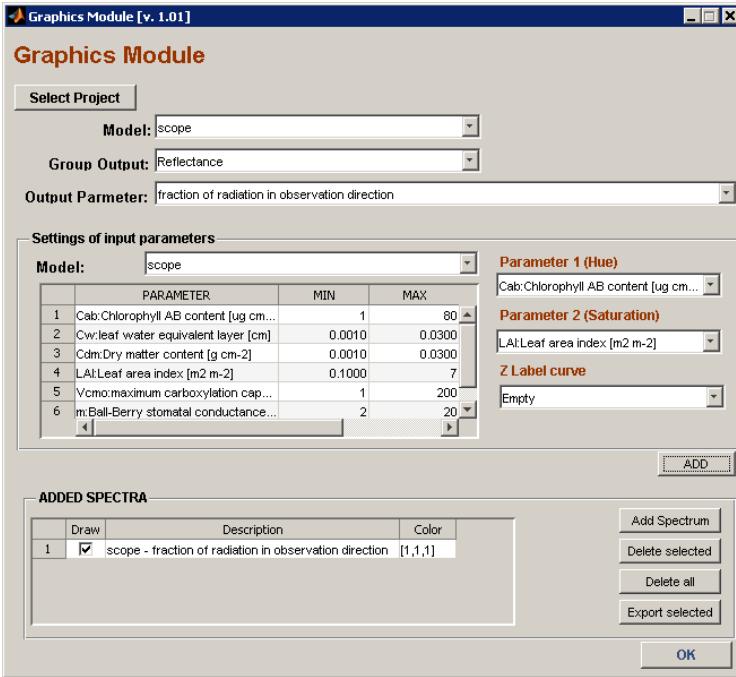
An "OK" button is located at the bottom right of the table area.

- New sensor settings can be imported by clicking on the 'Import' button in the top bar.
- Existing band settings can be modified or new ones can be added by clicking on the 'Edit' button.
- Also a spectral filter of a sensor can be imported or viewed by clicking on the 'Spectral Filter' button.

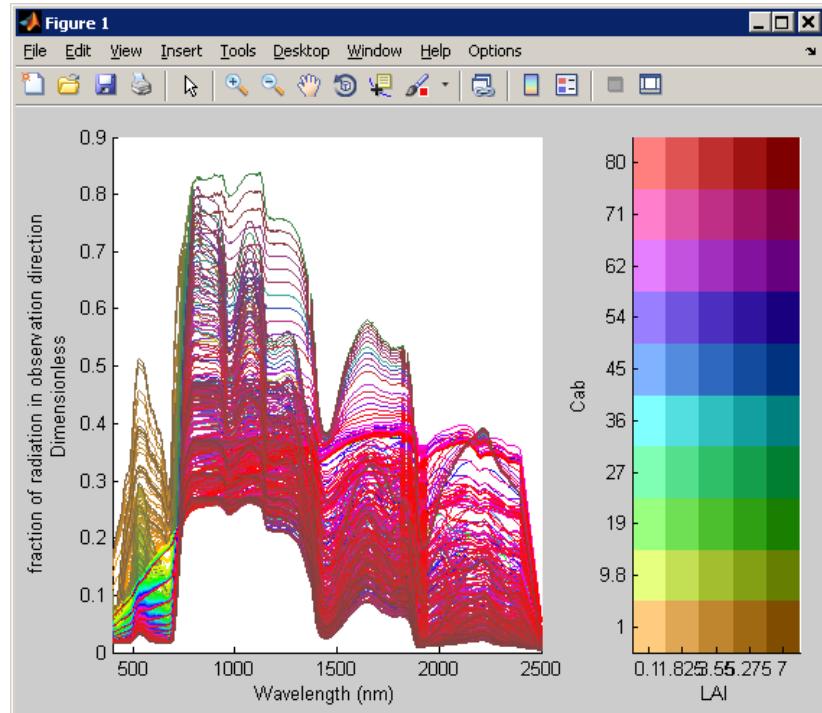
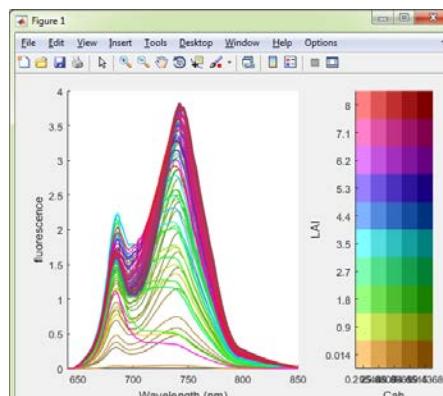
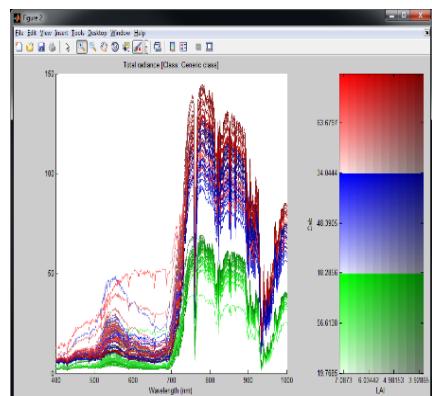
Default sensors:

- Landsat 7 TM
- Landsat 7 ETM+
- SPOT-4 VMI
- SPOT-4 HRVIR
- CHRIS Mode-3
- MODIS
- MERIS
- Sentinel-2
- Sentinel-3 OLCI
- Sentinel-3 SLSTR
- Landsat 8
- Pleiades-1A
- Quickbird

Graphics

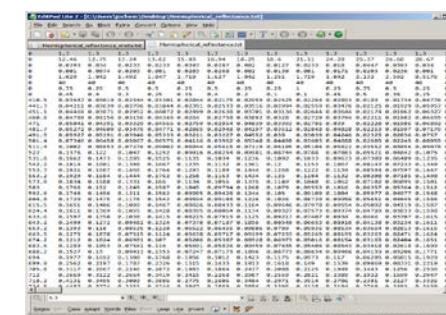


Visualization options:

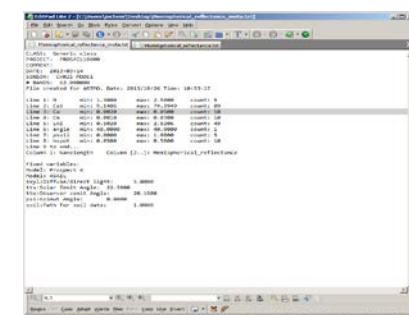


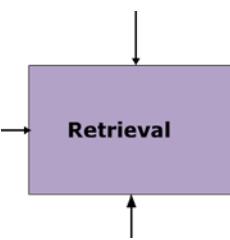
Export:

Spectral data

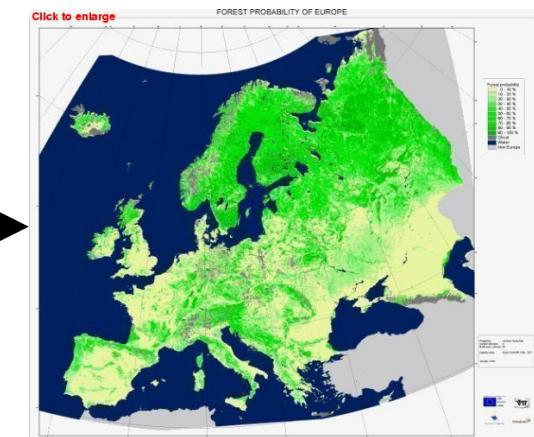
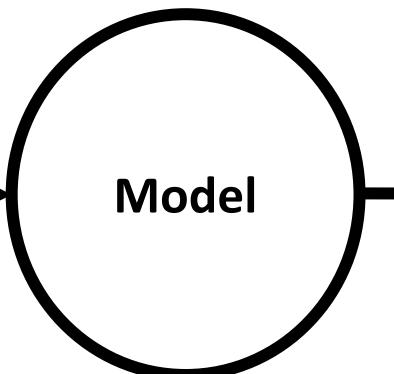


Associated metadata

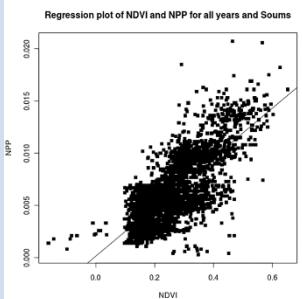
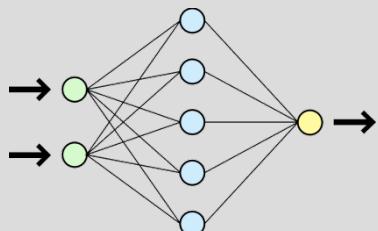
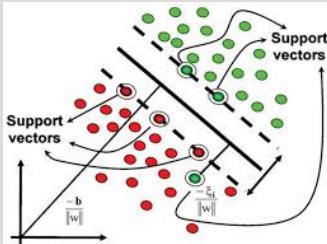
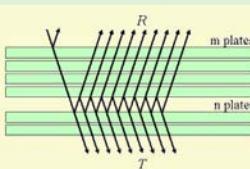
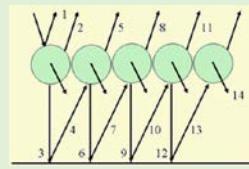
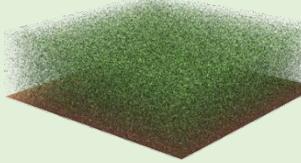




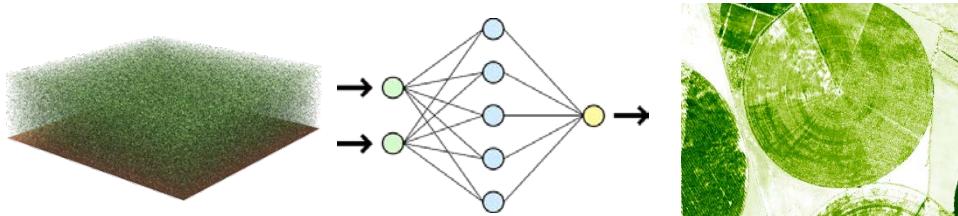
Retrieval



Retrieval families

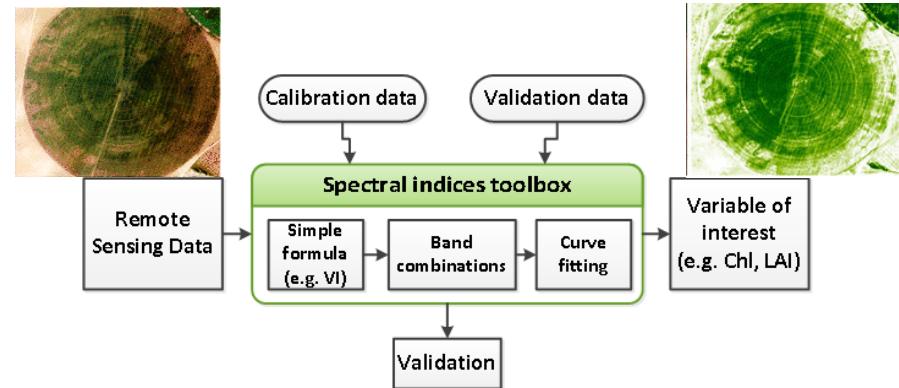
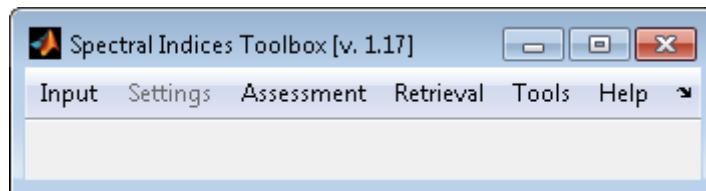
Parametric regression	Non-parametric regression	RTM inversion
<p>Spectral relationships that are sensitive to specific vegetation properties</p> $NDVI = \frac{(\rho_{NIR} - \rho_{RED})}{(\rho_{NIR} + \rho_{RED})}$ <p>Normalized Difference Vegetation Index</p> 	<p>Advanced techniques that search for relationships between spectral data and biophysical variables</p>  	<p>Models that simulate interactions between vegetation and radiation</p> <p>leaf</p>   <p>canopy</p>  

Methods of these different families can be combined: *hybrid methods*

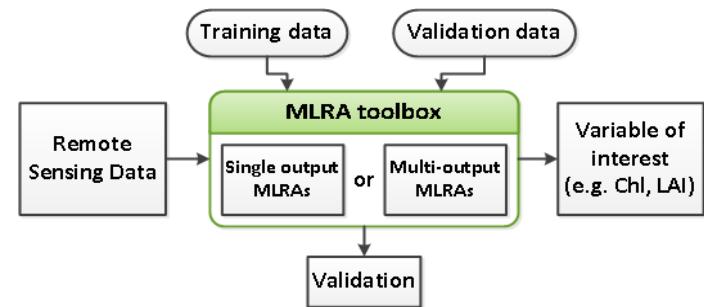


ARTMO's retrieval toolboxes:

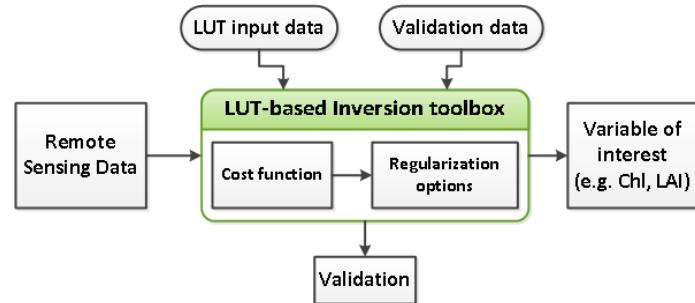
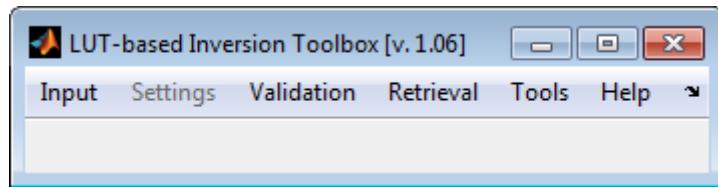
Spectral indices toolbox



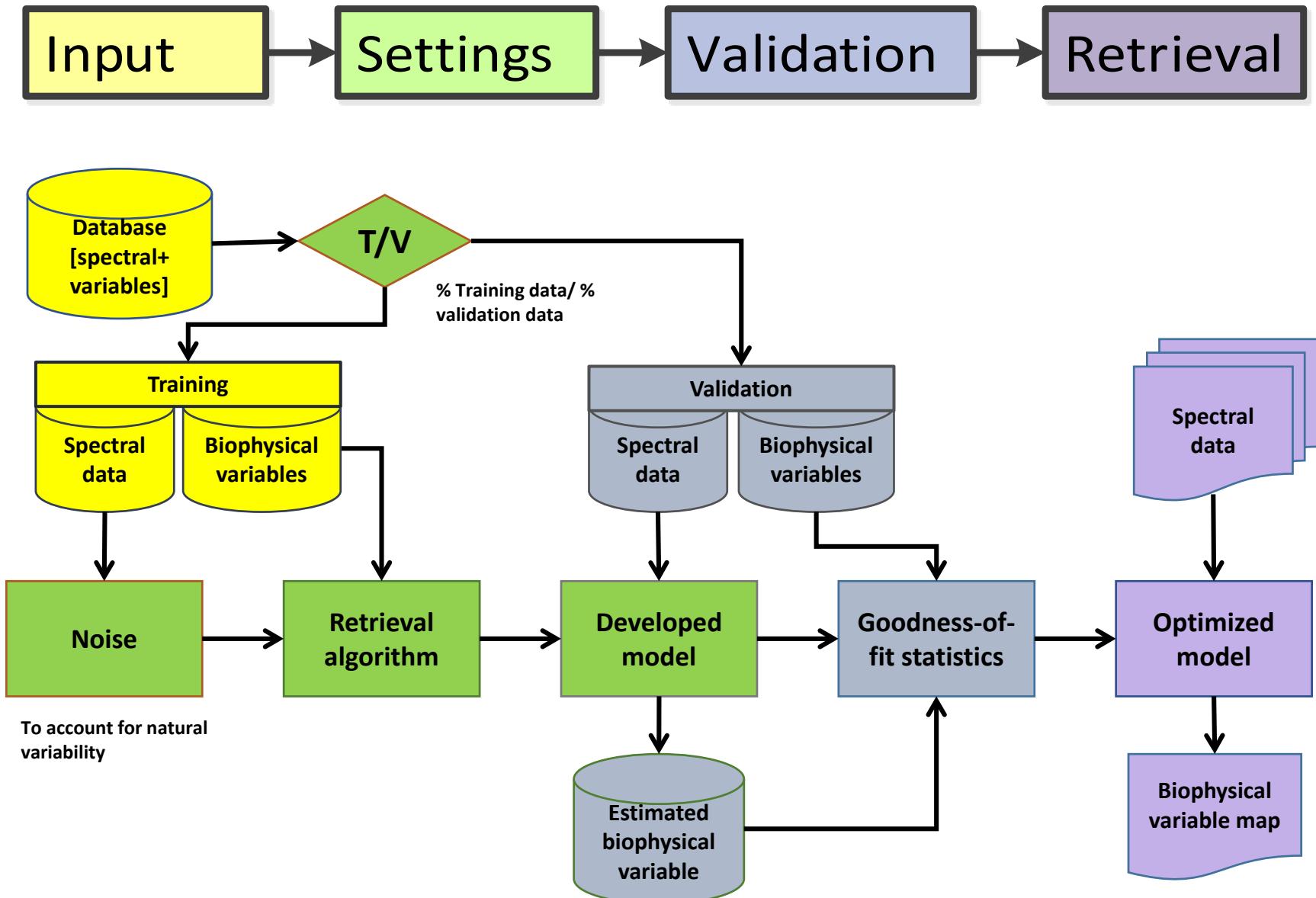
Machine learning regression algorithm toolbox



LUT-based inversion toolbox

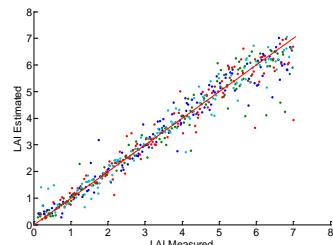


General structure:

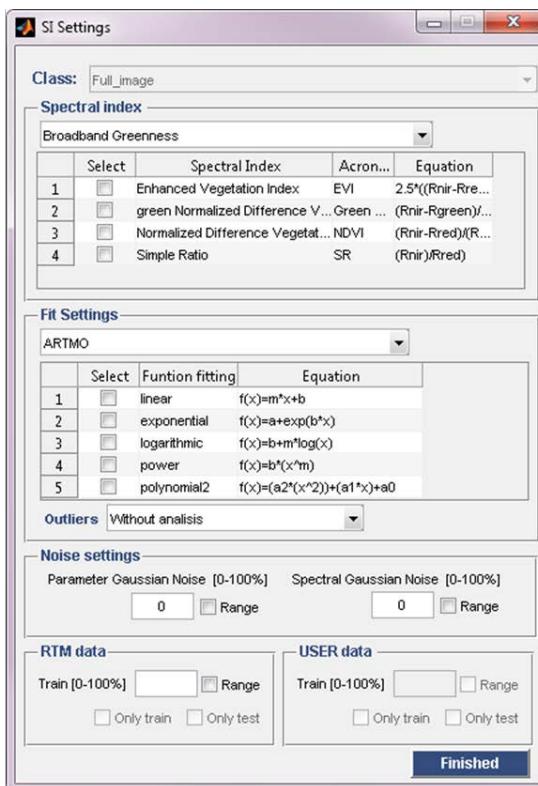
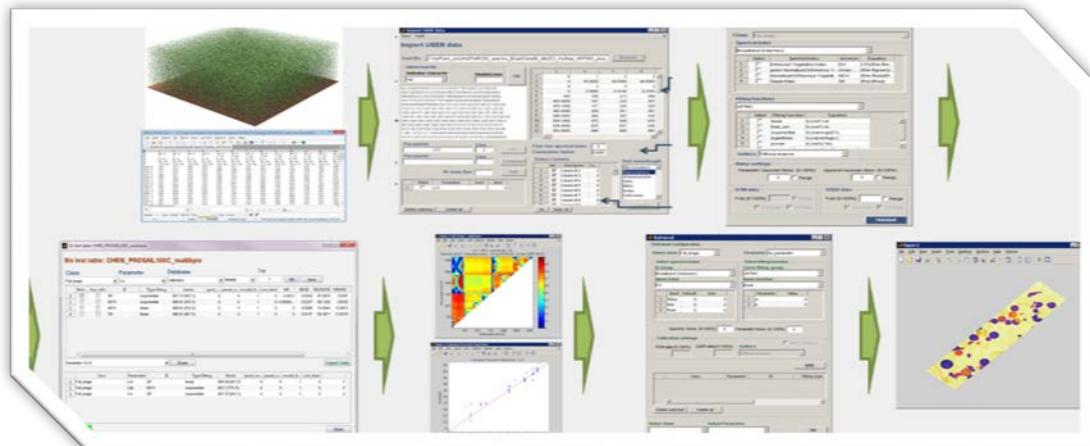
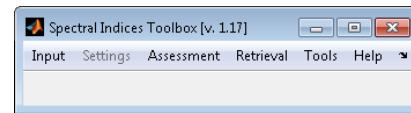


OPTIMIZING

- Retrieval: parametric/non-parametric/inversion
- Band selection
- Dimensionality & sample reduction



Spectral indices toolbox:



If active, configure per land cover class.

Select an Index group

Select one or multiple indices

Select one or multiple curve fittings

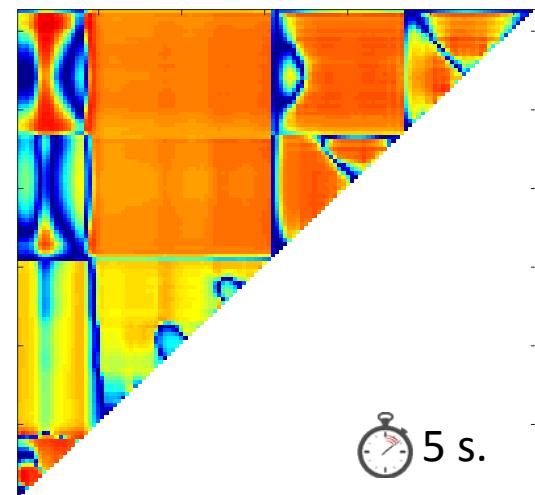
Options to add noise

Option to mix RTM with field observations

Properties:

- Calculates all possible band combinations.
- For index formulations with up to 10-band indices ($\#b^{10}$, for a 10-band sensor that would be 10 billion combinations)
- Includes multiple fitting functions (linear, exponential, logarithmic, power, polynomial)
- Noise & Cross-validation options
- Results stored in MySQL
- Top-performing indices per formulation and fitting function are given.
- Can process both image or individual spectra.

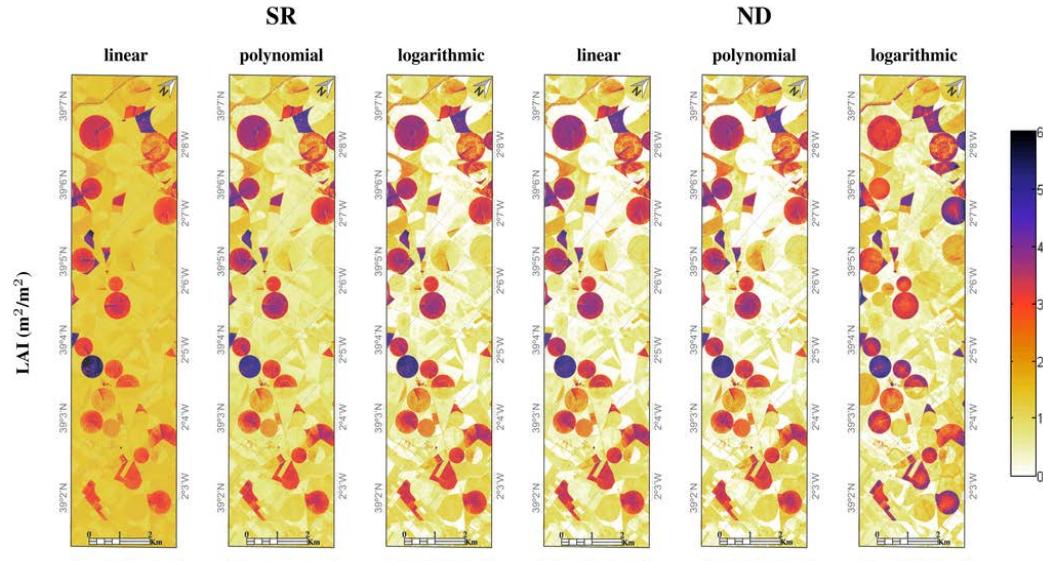
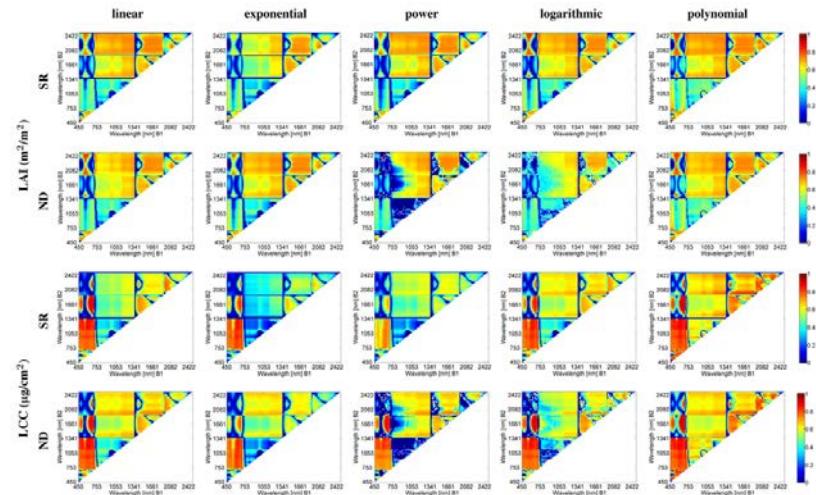
SPARC – HyMap - LAI



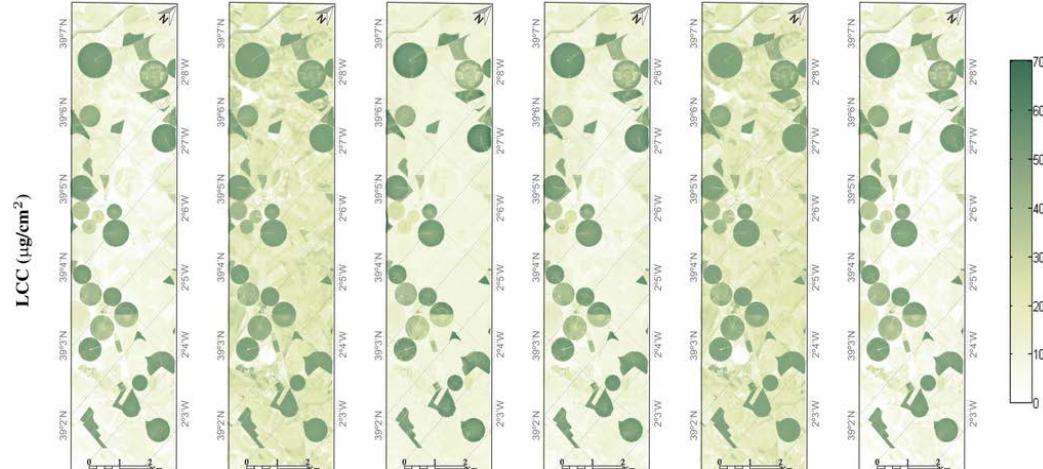
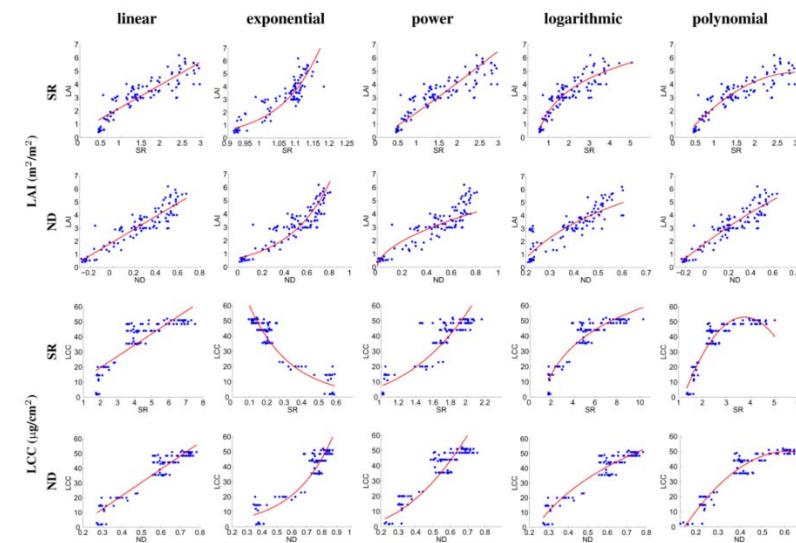
Best-performing index can be applied to an image.

SPARC dataset (Barrax Spain); HyMap data

SR and ND for different fitting functions

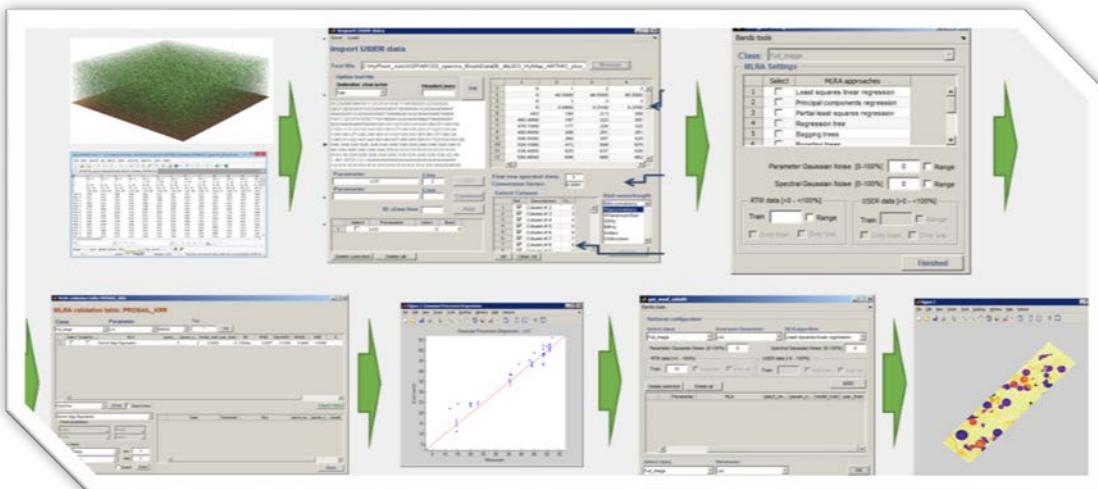
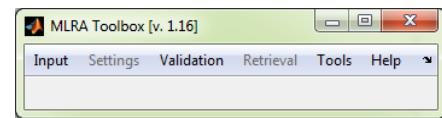


Best performing model.



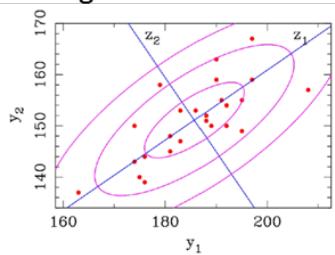
- Which VI method is most correct?
- Why restricting to a few bands only?

Machine learning regression algorithm toolbox

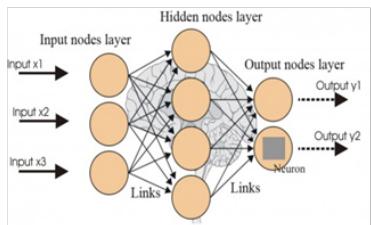


Simpler to execute than SI: no band selection needed.

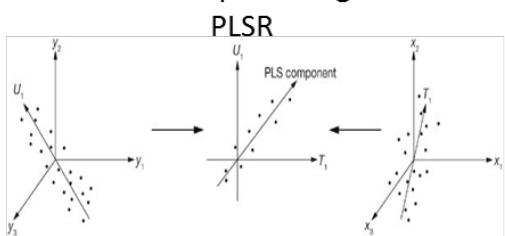
Principal component regression – PCR



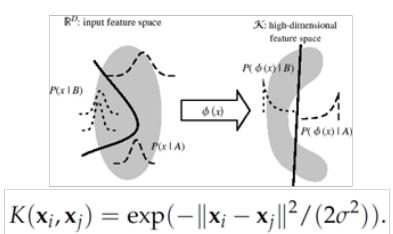
Neural networks
NN



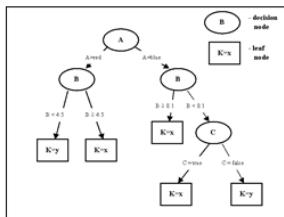
Partial least squares regression



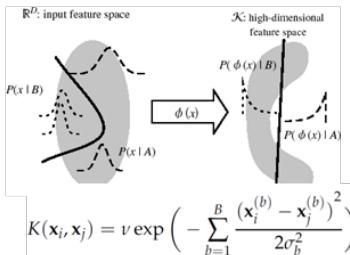
Kernel ridge regression
KRR



Decision Trees – DT



Gaussian processes regression
GPR



Properties:

- About 15 MLRAs implemented
- Single-output & multi-output
- Noise & Cross-validation options
- Dimensionality reduction options
- Results stored in MySQL
- GPR properties: band relevance & uncertainties
- Can process both images or individual spectra.
- Active learning, GPR-BAT, dim. reduction

Non-parametric models:

- SimpleR [Camps-Valls et al., 2013]
- <http://www.uv.es/gcamps/code/simpleR.html>

Also:

- Bagging trees (BAGTREE)
- Boosting trees (BOOST)
- Neural networks (NN)
- Extreme Learning Machines (ELM)
- Support Vector Regression (SVR)
- Relevance Vector Machine (RVM)
- Variational Heteroscedastic Gaussian Process Regression (VHGPR)

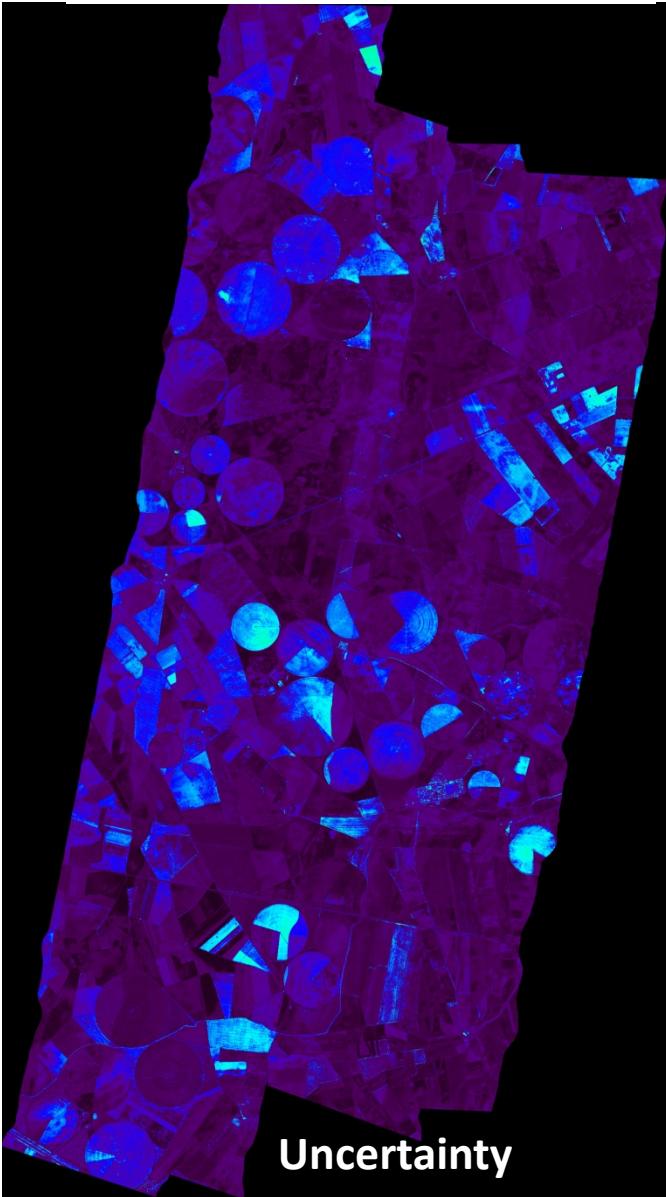
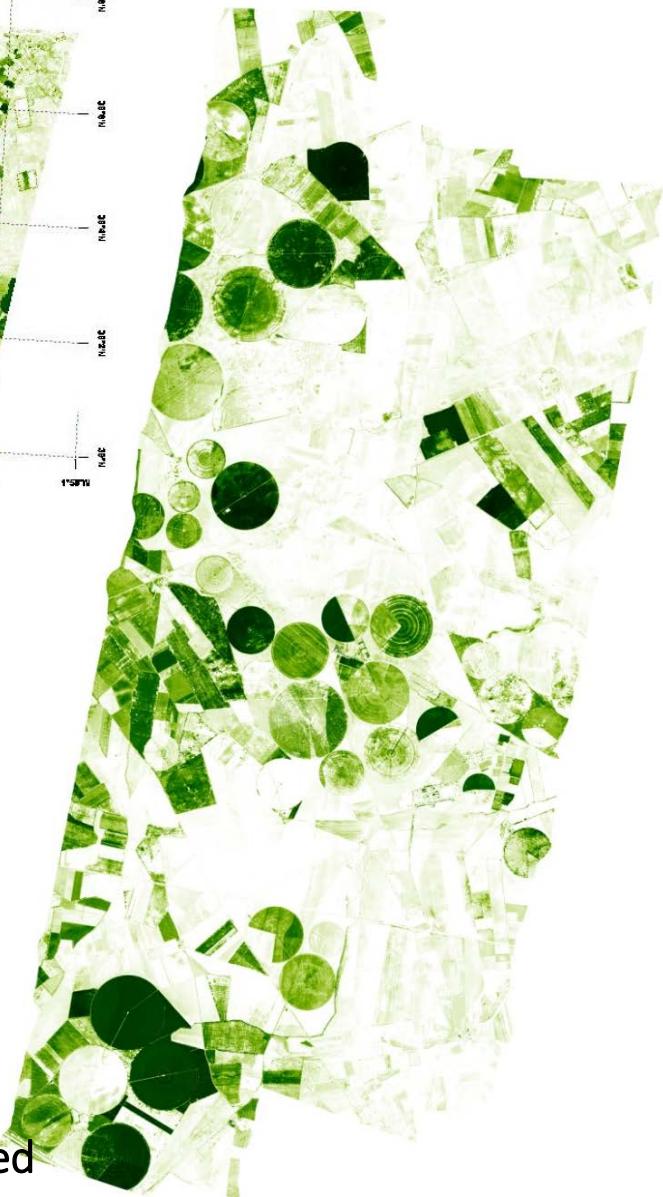
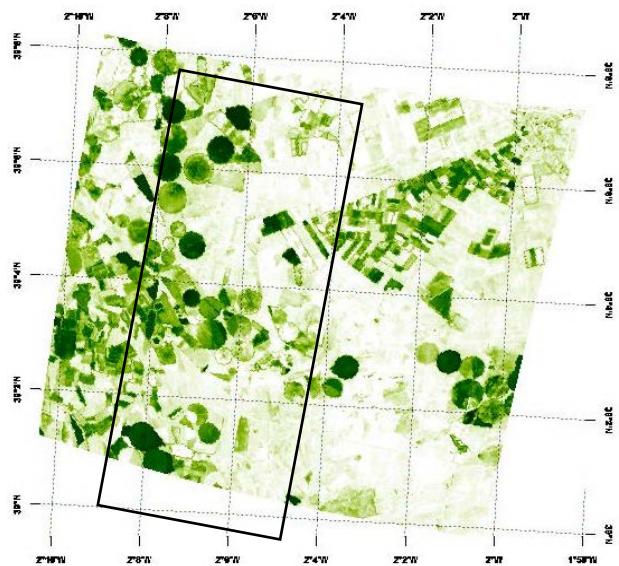
GPR in Bayesian framework also provides:

- Band relevance
- Uncertainty estimates

GPR maps

LCC [$\mu\text{g}/\text{cm}^2$]

CHRIS

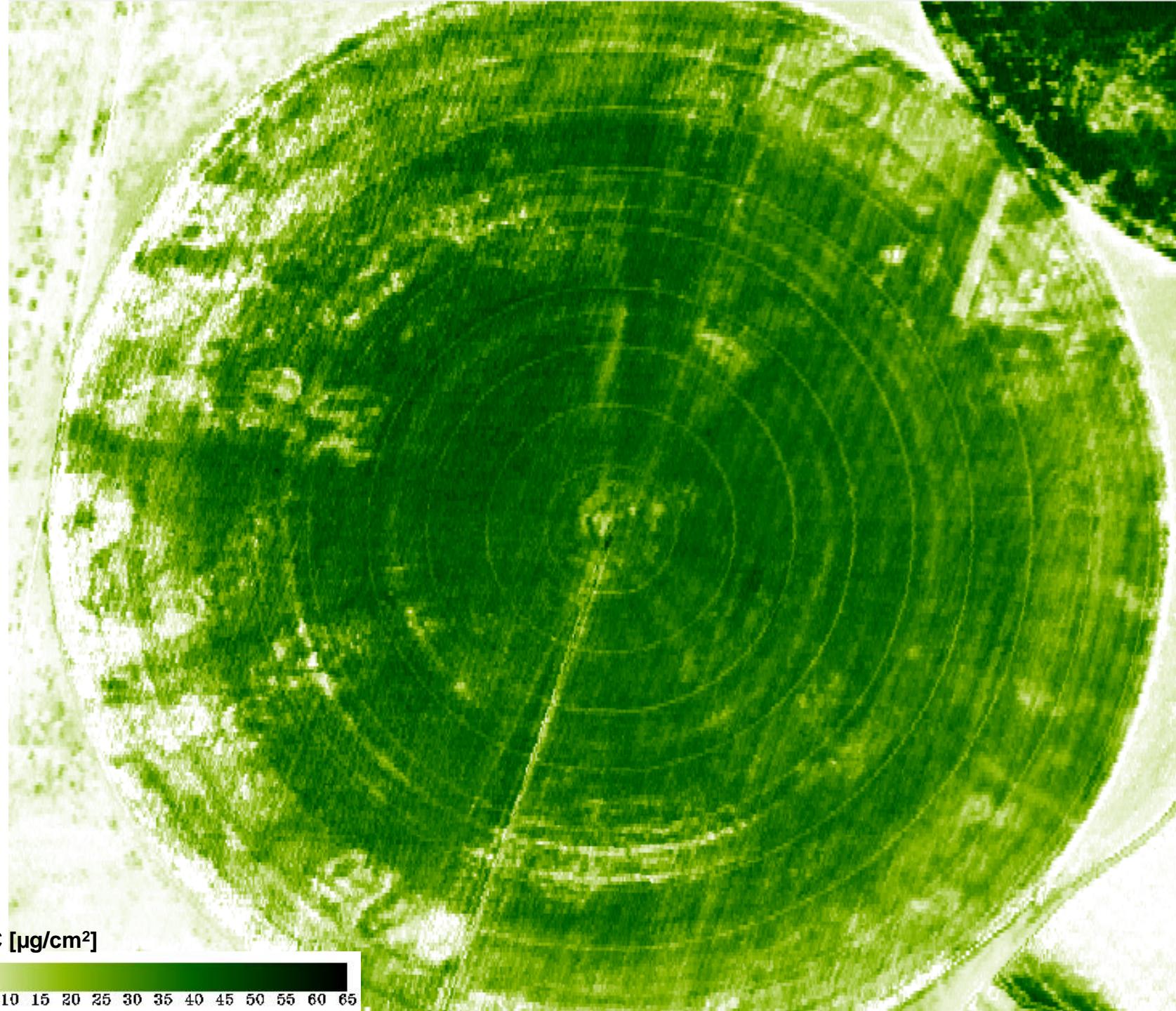


CASI1500

- pixel size: 1.4m
- 288 bands

Same GP model was applied

Uncertainty



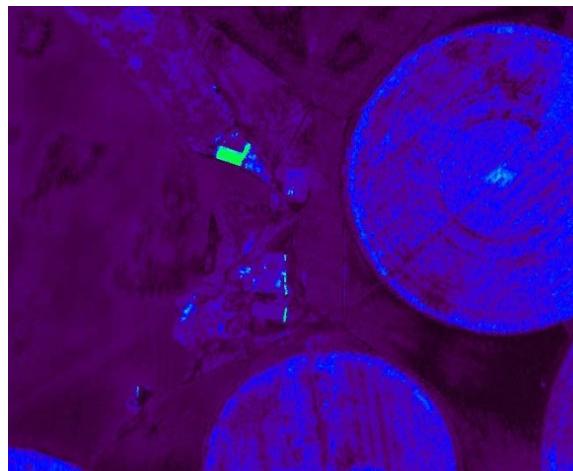
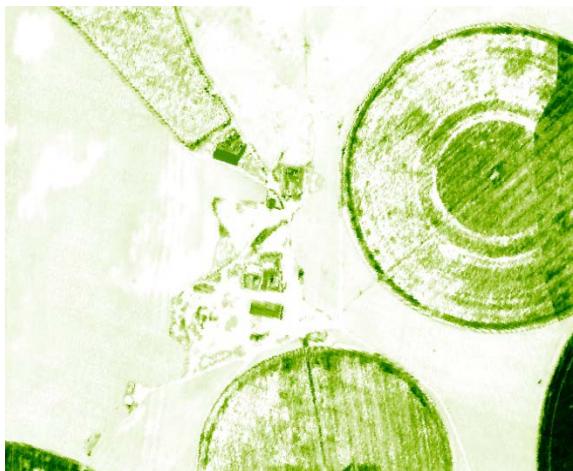
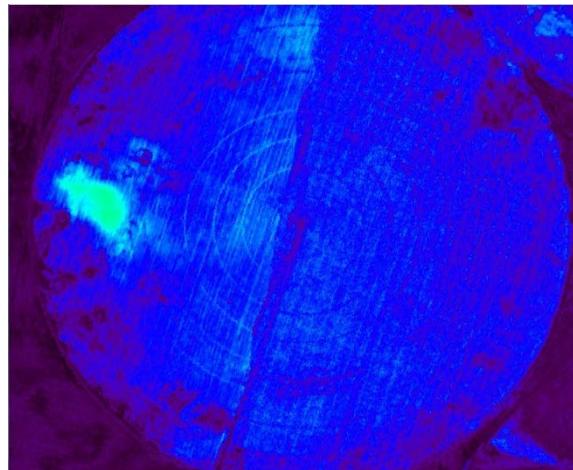
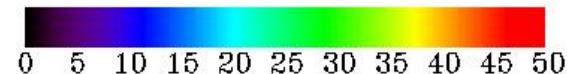
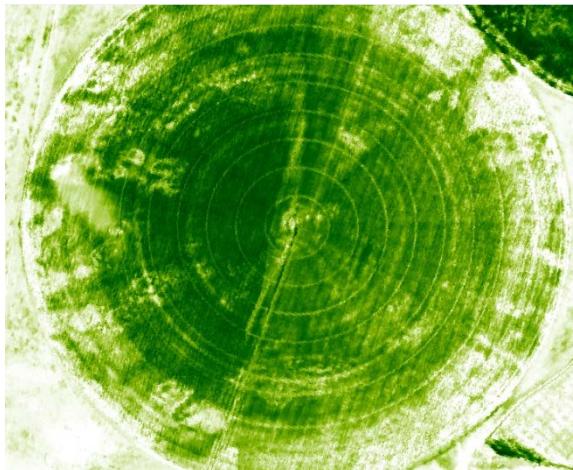
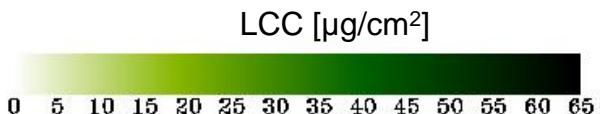
LCC [$\mu\text{g}/\text{cm}^2$]

0 5 10 15 20 25 30 35 40 45 50 55 60 65

St Dev

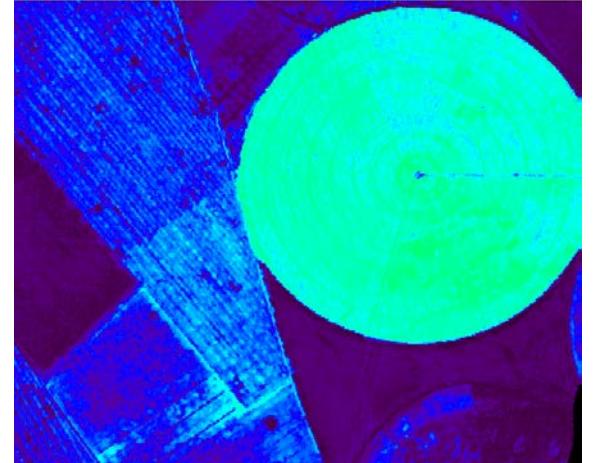
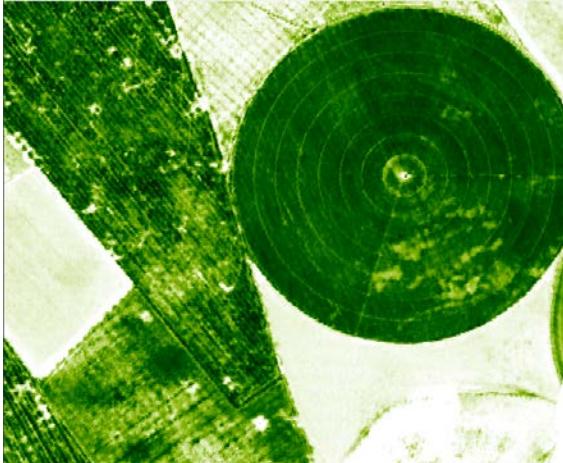
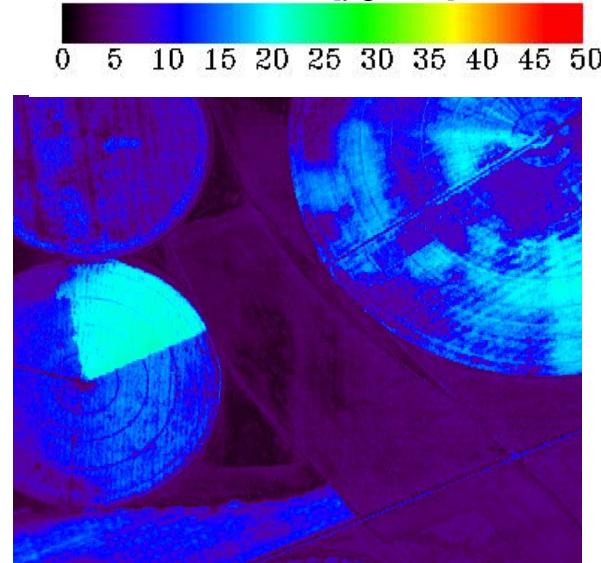
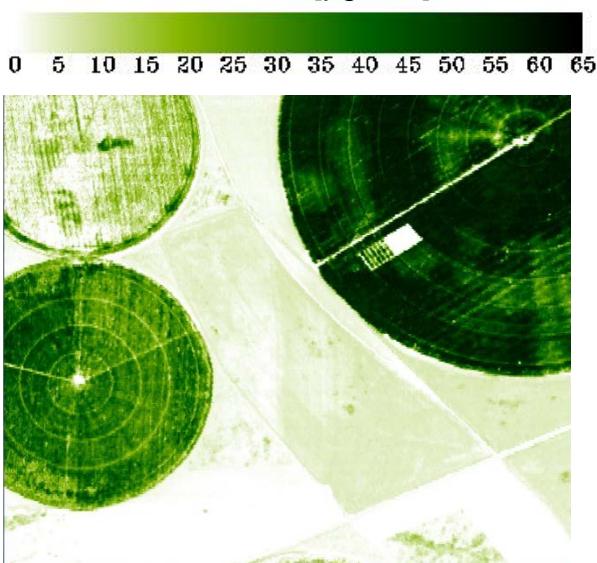
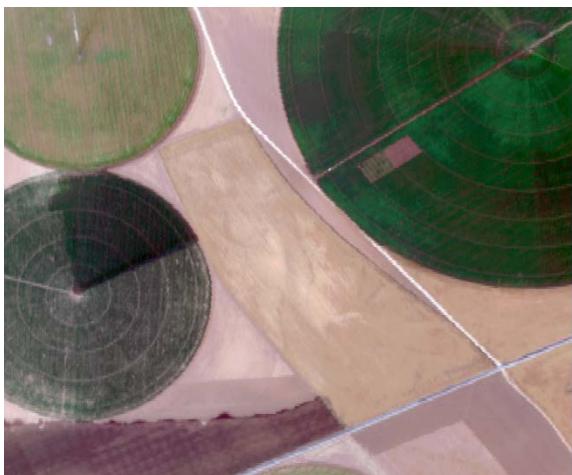
LCC [$\mu\text{g}/\text{cm}^2$]

RGB CASI



- **Uncertainty maps** provide **additional info** which may be hidden on the images.
- **Implausible estimations are detected.**

RGB CASI



- In turn, despite low uncertainties also good estimations. No impact on recently irrigated areas (other methods have difficulties with wet soils).
- For operational applications, of interest to flag/mask regions with high uncertainties.



The challenge of machine learning applied to imaging spectroscopy

- *Machine learning methods are adaptive and can be very powerful. However that goes a computational cost. This can be problematic when large datasets are involved, either in the sampling or in the spectral domain (e.g. for hybrid methods).*
- *Moreover, when many bands are involved: multicollinearity leads to statistical problems (suboptimal performance).*

Solutions to deal with large datasets:

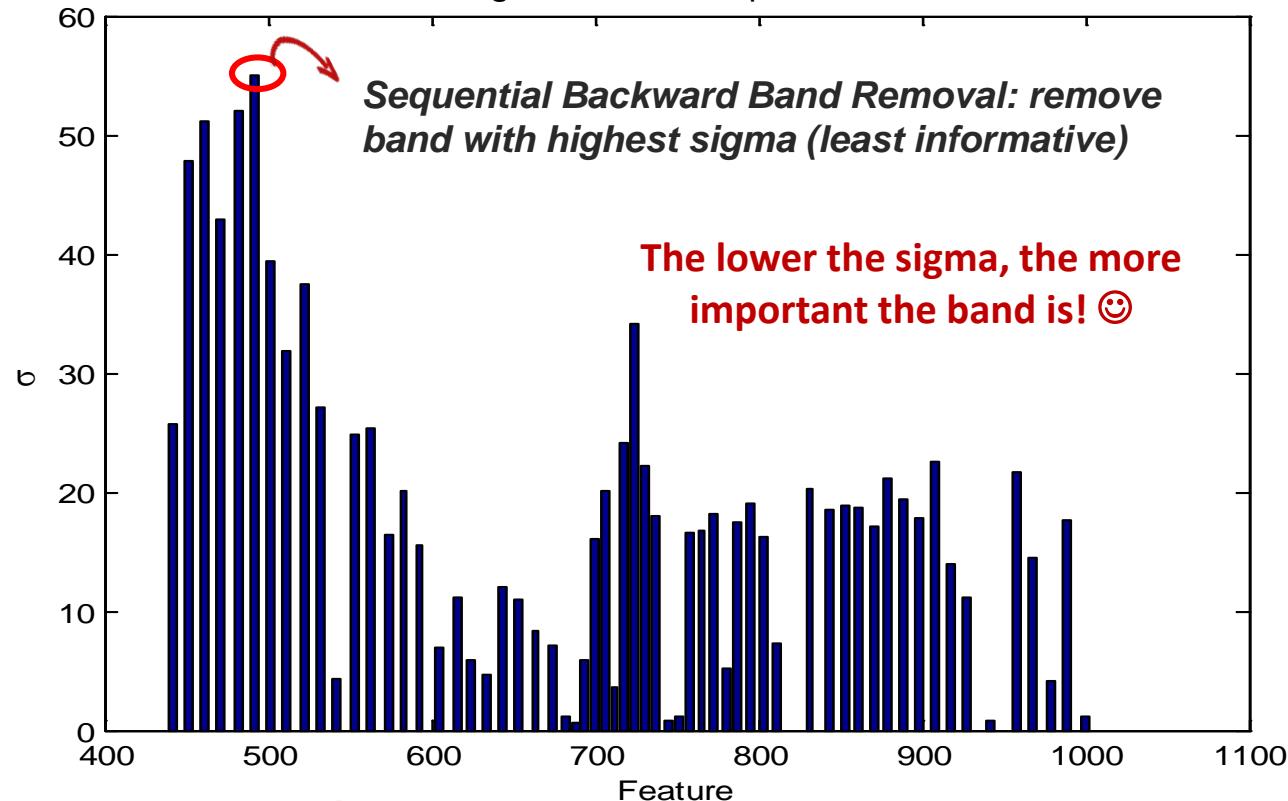
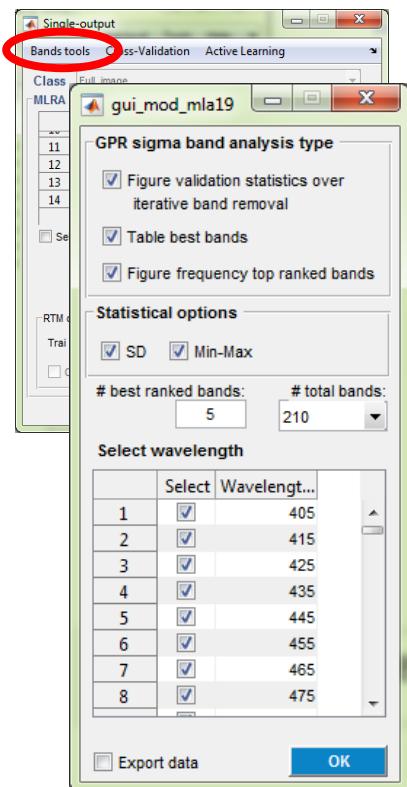
1. Reducing spectral data:

- I. band selection (GPR-BAT),
- II. dimensionality reduction

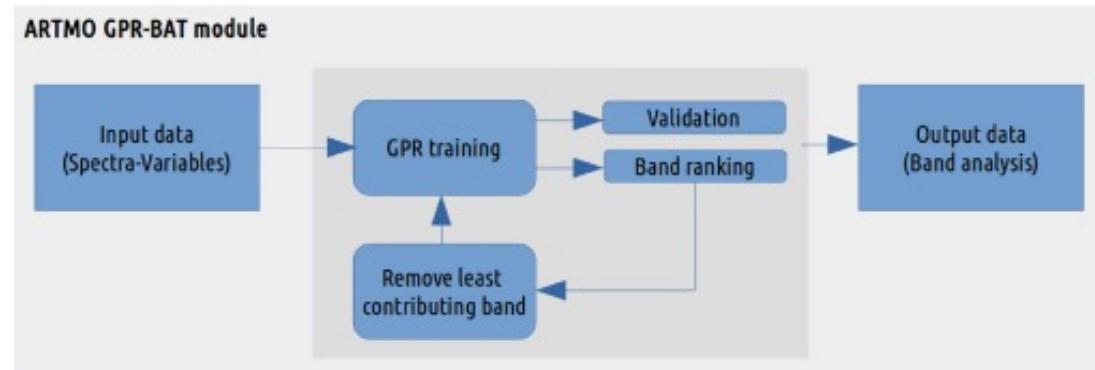
2. Samples reducing : Active learning

I) Band selection: GPR-BAT

Gaussian processes regression – Band analysis Tool (GPR-BAT).



automated

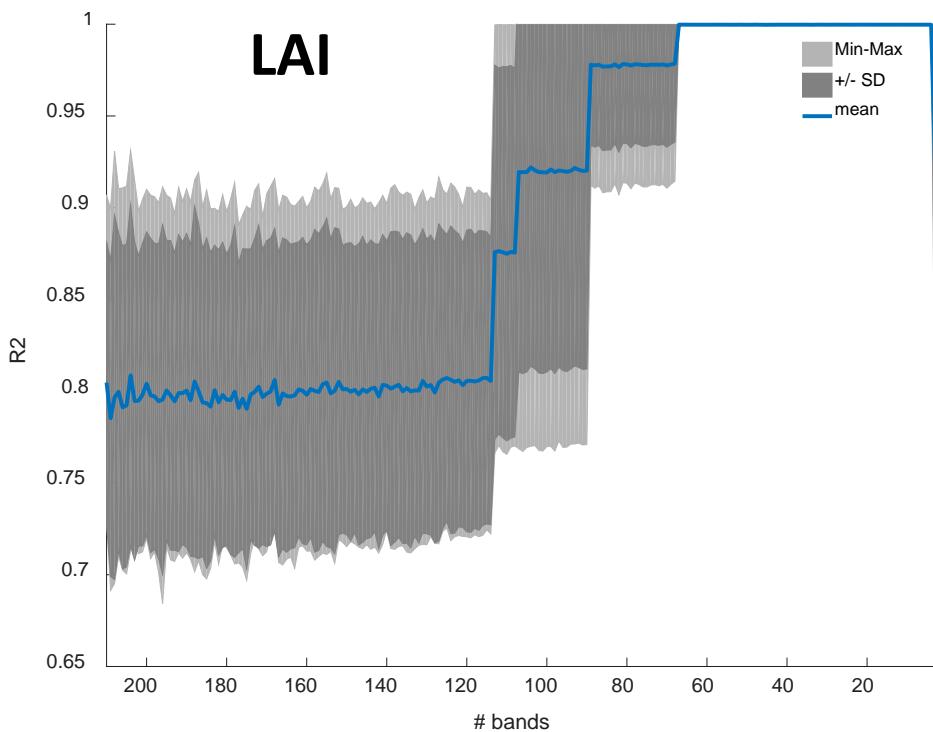
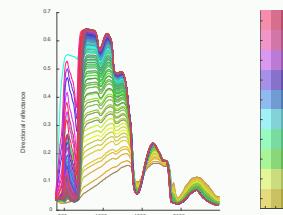


Best-performing method can be applied to an image.

GPR-BAT example with simulated data (PROSAIL)

Experimental setup:

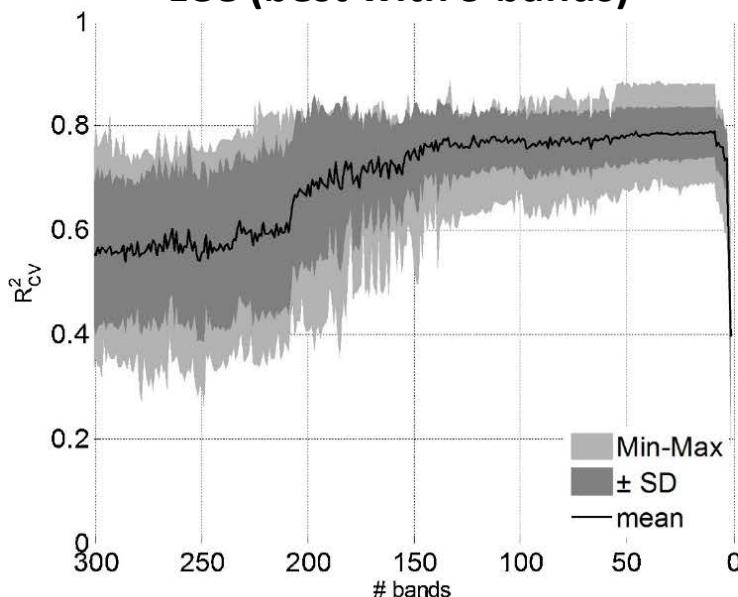
- PROSAIL: LHS 100# @ 10 nm, 220b; Cab, LAI
- 4k cross-var sampling



# band	R2	wavelengths
5	0.9997	815, 1145, 1205, 122, 1245
4	0.9997	815, 1145, 1205, 1245
3	0.9213	815, 1145, 1205
2	0.8104	815, 1145
1	0.8104	815

**Best performances achieved between 70 and 4 bands
(using all bands or <3 bands not recommended)**

LCC (best with 9 bands)



All bands

482, 500, 564, 566, 710, 712, 714, 878, 966, 980
482, 500, 564, 710, 712, 714, 878, 966, 980

482, 500, 564, 710, 712, 714, 878, 966

482, 500, 564, 710, 714, 878, 966

482, 500, 710, 714, 878, 966

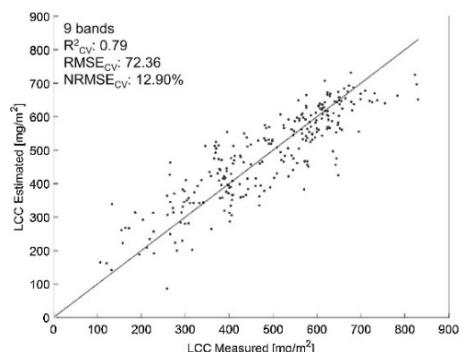
500, 710, 714, 878, 966

500, 710, 714, 878

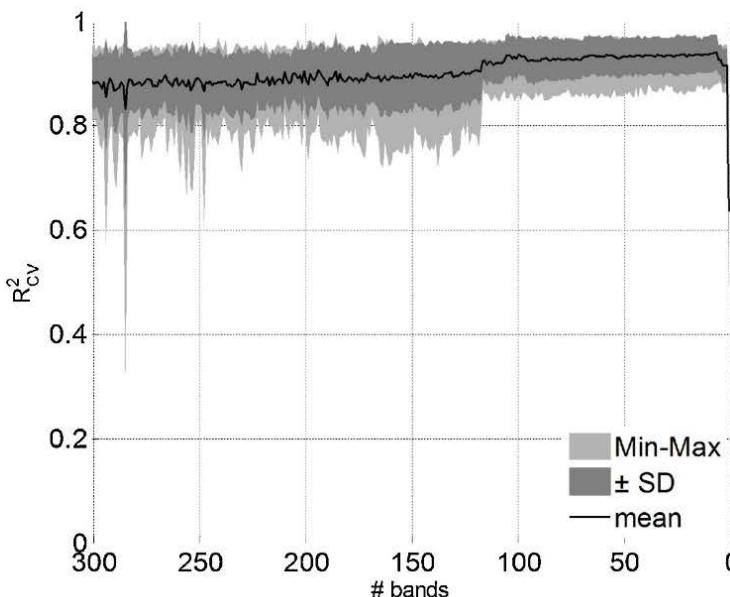
500, 710, 878

500, 710

710



gLAI (best with 7 bands)



All bands

406, 746, 770, 790, 792, 794, 798, 808, 858, 878

406, 746, 790, 792, 794, 798, 808, 858, 878

406, 746, 790, 792, 794, 798, 858, 878

406, 746, 792, 794, 798, 858, 878

746, 792, 794, 798, 851

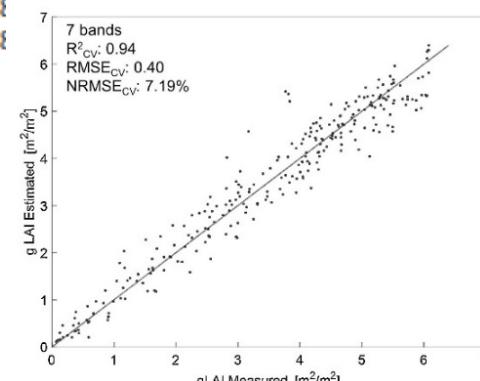
746, 792, 794, 798, 871

746, 792, 794, 798

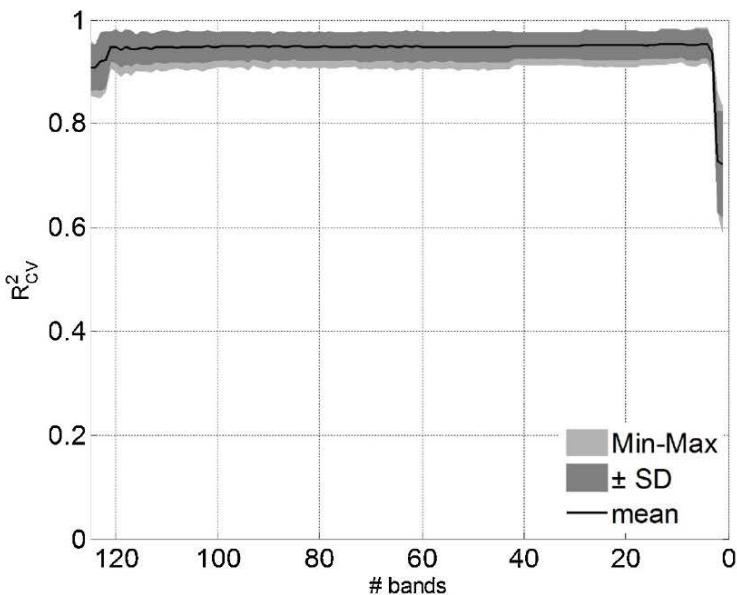
746, 792, 794

746, 792

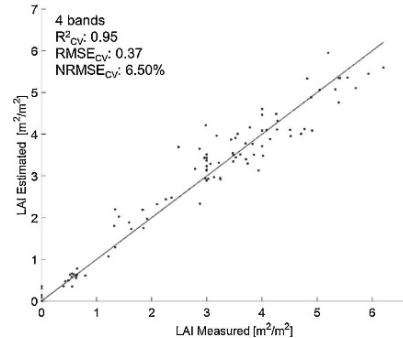
792



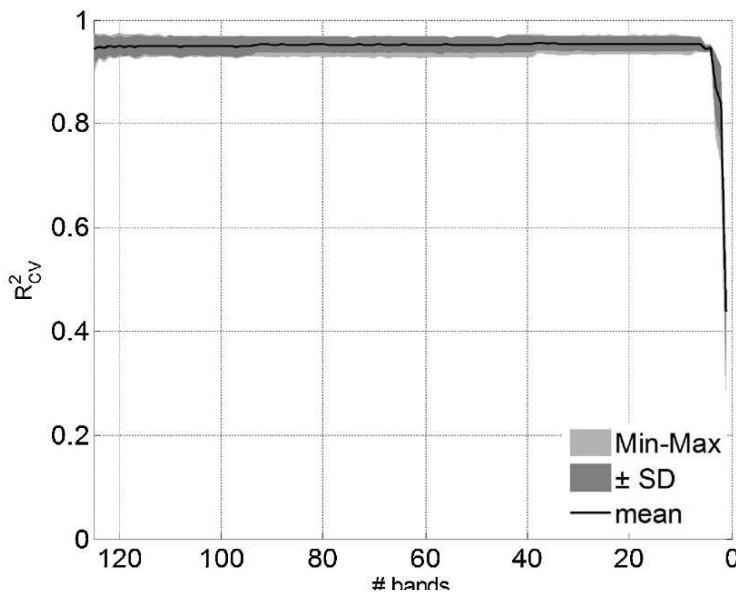
LAI (best with 4 bands)

*All bands*

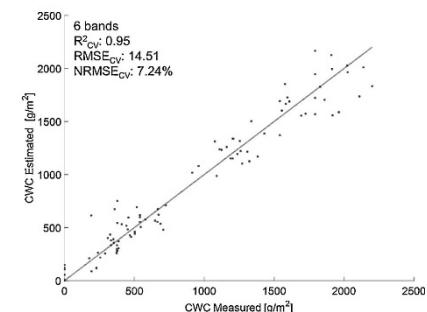
- :
- 462, 478, 708, 723, 1215, 1243, 1272, 1327, 1635, 2483
- 462, 478, 708, 723, 1215, 1243, 1272, 1327, 2483
- 462, 478, 708, 723, 1215, 1243, 1272, 1327
- 462, 478, 708, 723, 1215, 1272, 1327
- 462, 478, 708, 723, 1215, 1327
- 462, 478, 708, 723, 1327**
- 462, 708, 723, 1327**
- 462, 708, 1327
- 462, 1327
- 462



CWC (best with 6 bands)

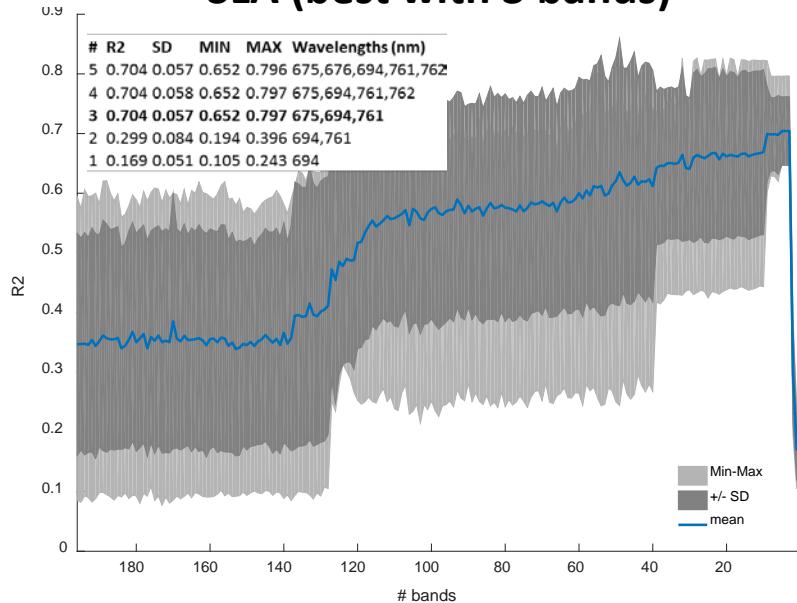
*All bands*

- :
- 462, 723, 1128, 1157, 1272, 1286, 1299, 1327, 1419, 2483
- 723, 1128, 1157, 1272, 1286, 1299, 1327, 1419, 2483
- 723, 1128, 1157, 1272, 1286, 1327, 1419, 2483
- 723, 1128, 1157, 1272, 1286, 1327, 1419
- 723, 1157, 1272, 1286, 1327, 1419**
- 723, 1157, 1272, 1286, 1327
- 723, 1157, 1272, 1286
- 1157, 1272, 1286
- 1157, 1286
- 1286

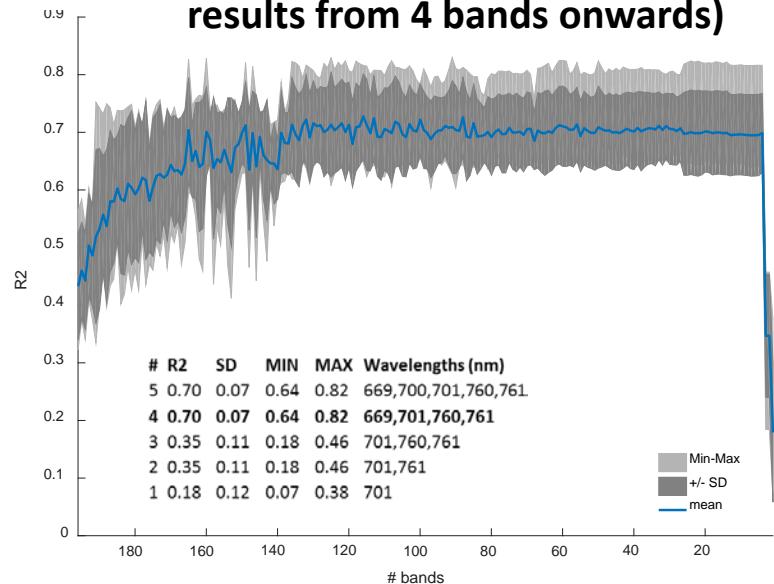


BIOHYPE & SCOPE: SIF (200#b)

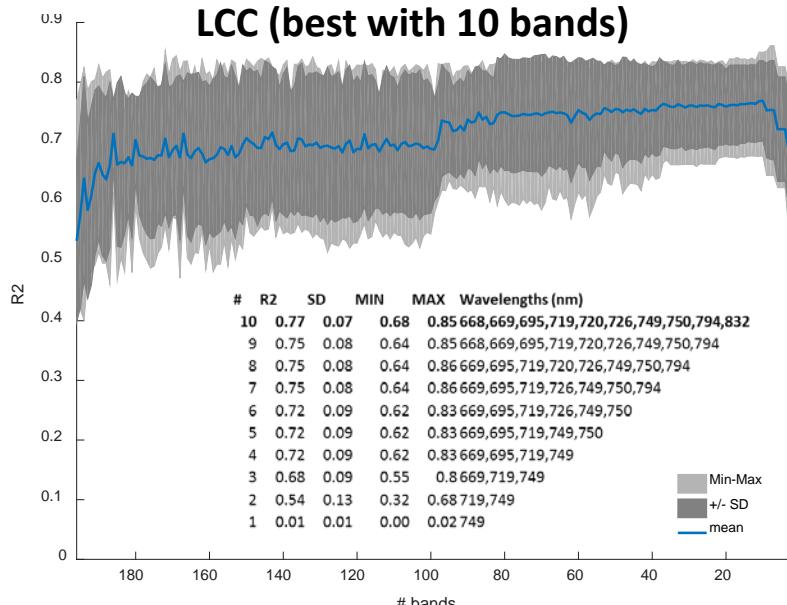
SLA (best with 3 bands)



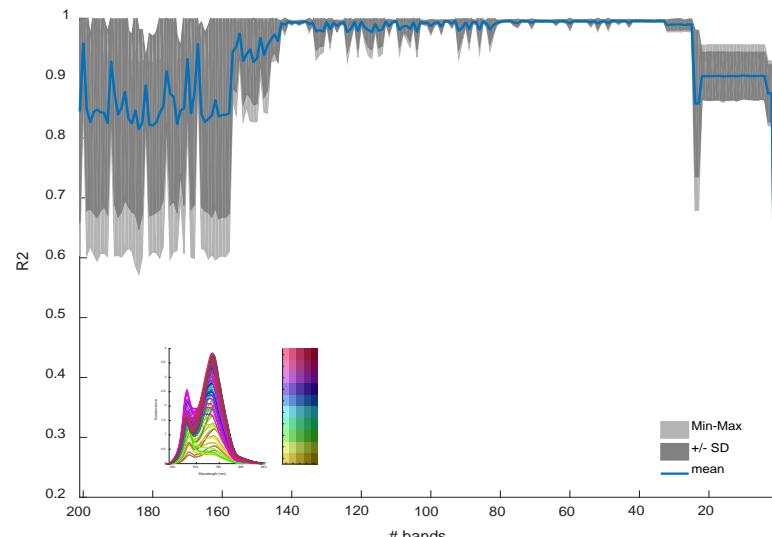
LWC (best with 116 bands, stable results from 4 bands onwards)



LCC (best with 10 bands)



SCOPE: LAI (best at 44 bands)



II) Dimensionality reduction: SIMFEAT

13 dimensionality reduction methods implemented.

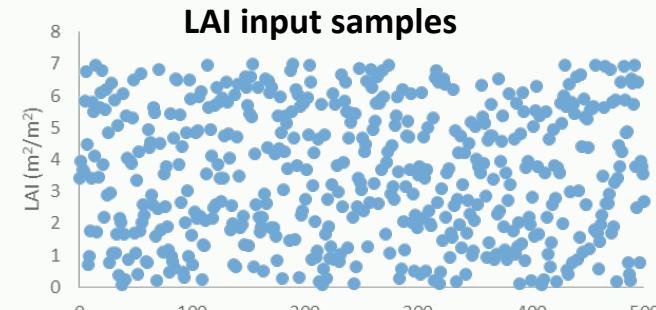
Method	Description	Kernel type
Principal component analysis (PCA)	Empty	Empty
Partial least squares (PLS)	Empty	Empty
Primal Partial least squares (PPLS)	Empty	Empty
Ortho-normalized PLS (OPLS)	Empty	Empty
Canonical correlation analysis (CCA)	Empty	Empty
Minimum Noise Fraction (MNF)	Empty	Empty
Principal Component of KECA method (KECA)	rbf	Mean
Kernel Principal Component Analysis (KMNF)	rbf	Mean
Kernel Principal Component Analysis (KPCA)	rbf	Mean
Kernel dual partial least squares (KDPLS)	rbf	Mean
Kernel partial least squares (KPLS)	rbf	Mean
Kernel Orthonormalized Partial Least Squares (KOP...)	rbf	Mean
Kernel Canonical Correlation Analysis (KCCA)	rbf	Mean

Feature: 5 OK

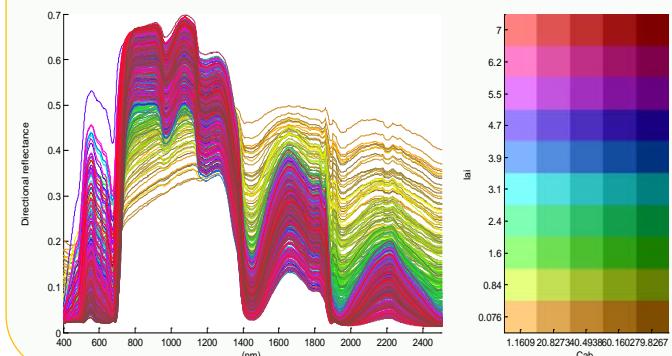
Experimental setup:

PROSAIL: 500 random samples

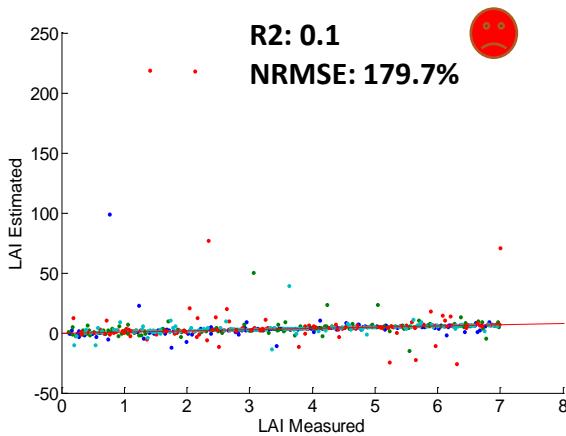
Variable	Min	Max
N	1.3	2.5
Cab	1	80
Cw	0.002	0.05
Cm	0.002	0.05
LAD	0	90
LAI	0.01	7



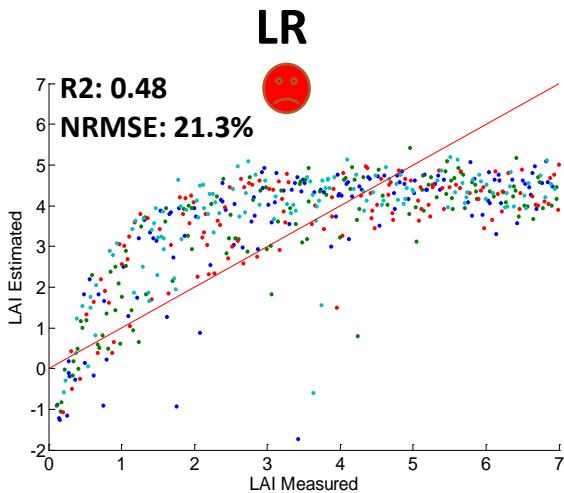
directional reflectance (2101 bands)



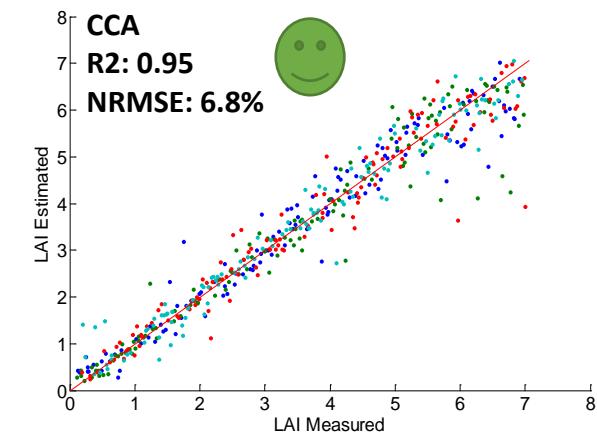
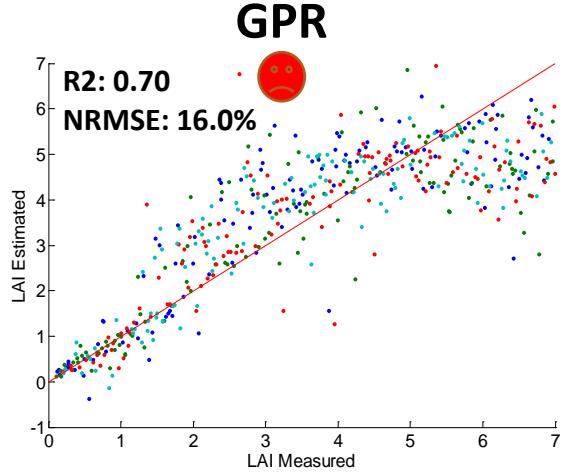
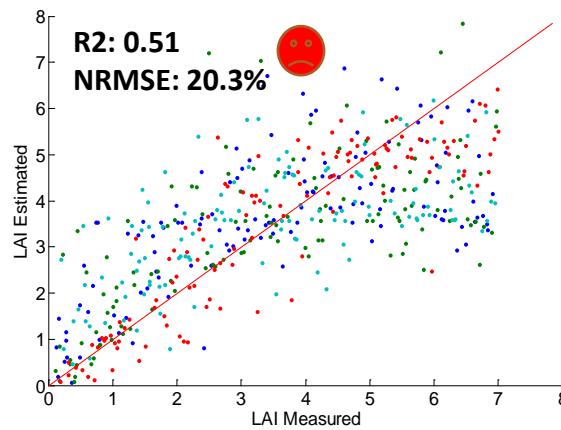
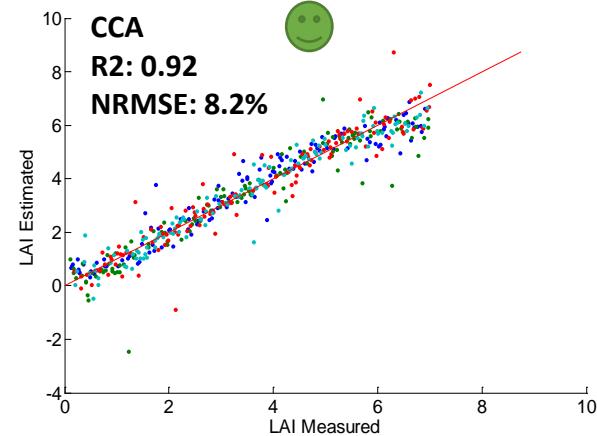
No DR (2101#)



PCA (5#)

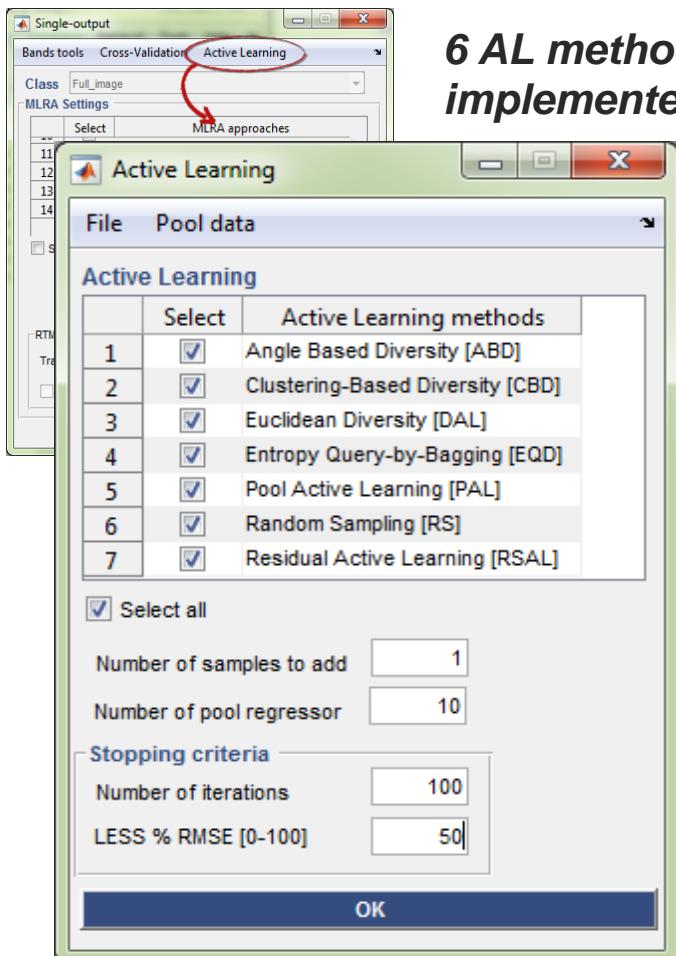


Best DR method (5#)



By combining advanced DR methods with (advanced) regression methods, hyperspectral data can be exploited to the fullest. Or, full spectral dataset into regression require (advanced) DR methods.

2) Sample reduction: Active learning (AL)

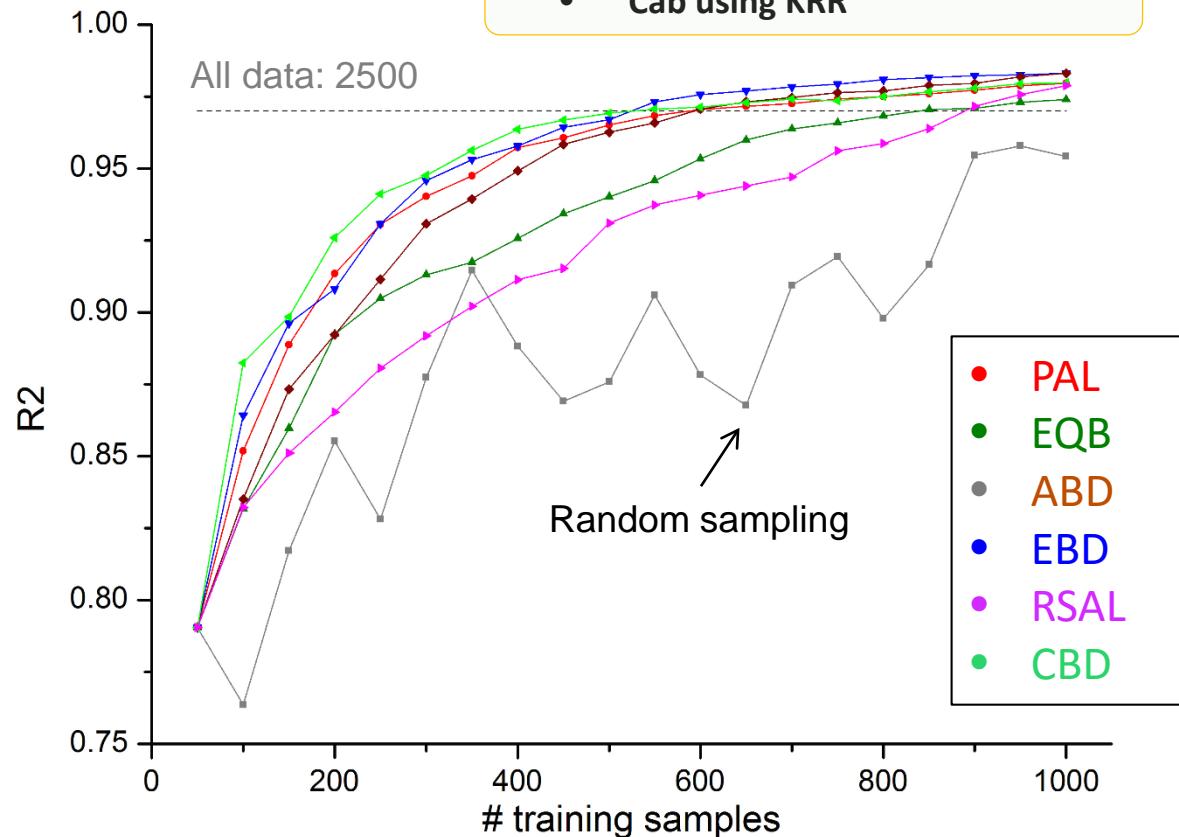


6 AL methods implemented.

Experimental setup:

PROSAIL: 5000 samples

- 2500 training; 2500 validation
- Cab using KRR

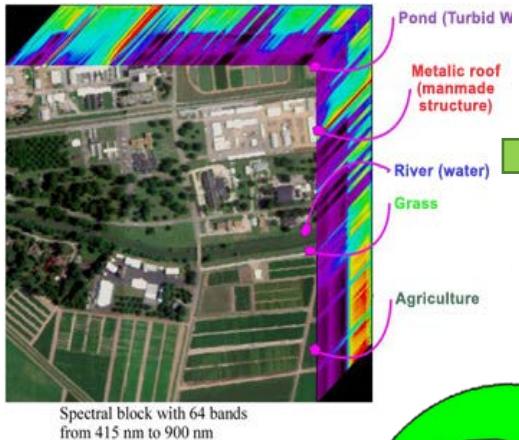


- Active learning (AL) searches for new samples from a data pool based on **uncertainty** (PAL, EQB, RSAL) and **diversity** (ABD, CBD, EBD).
- AL method search more efficiently for relevant samples than random sampling or when using all data.

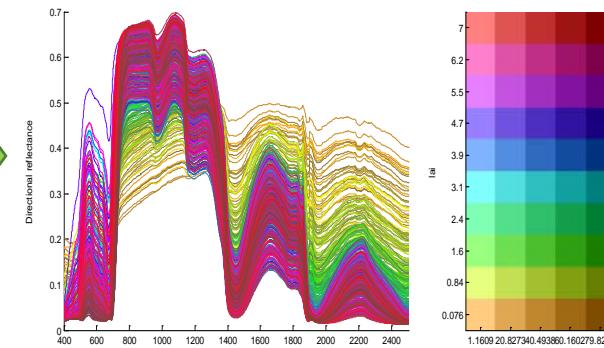
Best-performing method can be applied to an image.

Background LUT-based inversion

RS imagery



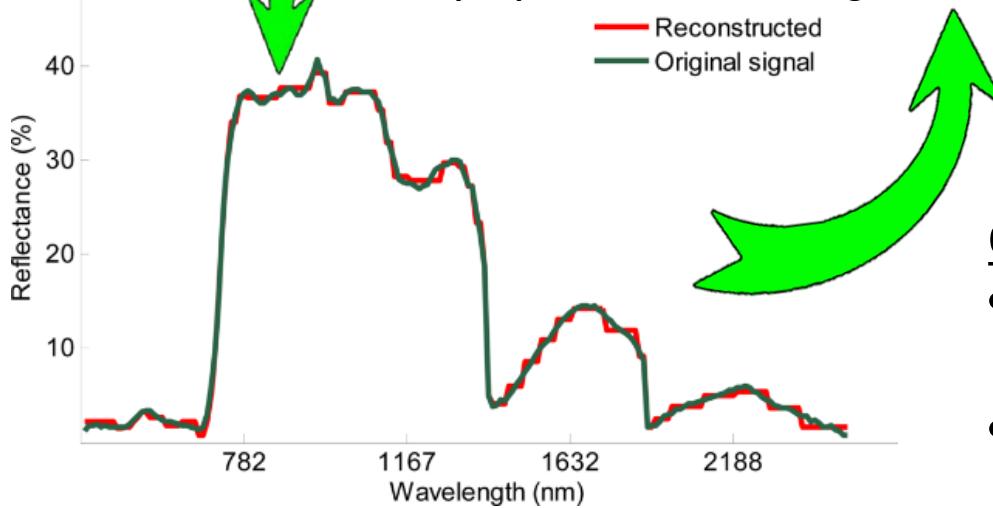
PROSAIL LUT



Best match is obtained through a ‘Cost function’, or ‘Minimum distance function’.

CostFunction	Definition	Minimum	Maximum
C^{LS}	$\sum(Y - X)^2$	0	∞
C^{NC}	$\frac{\sum(X \cdot Y)}{\sqrt{\sum X^2} \sqrt{\sum Y^2}}$	-1	1
C^W	$\sum_k \frac{n_k}{N} \frac{\sqrt{\text{Var}(Y_k)}}{\mu(Y_k)}$	0	∞
C^{CR}	$\frac{1}{\text{Var}(Y)} \sum_k \frac{n_k}{N} \text{Var}(Y_k)$	0	1

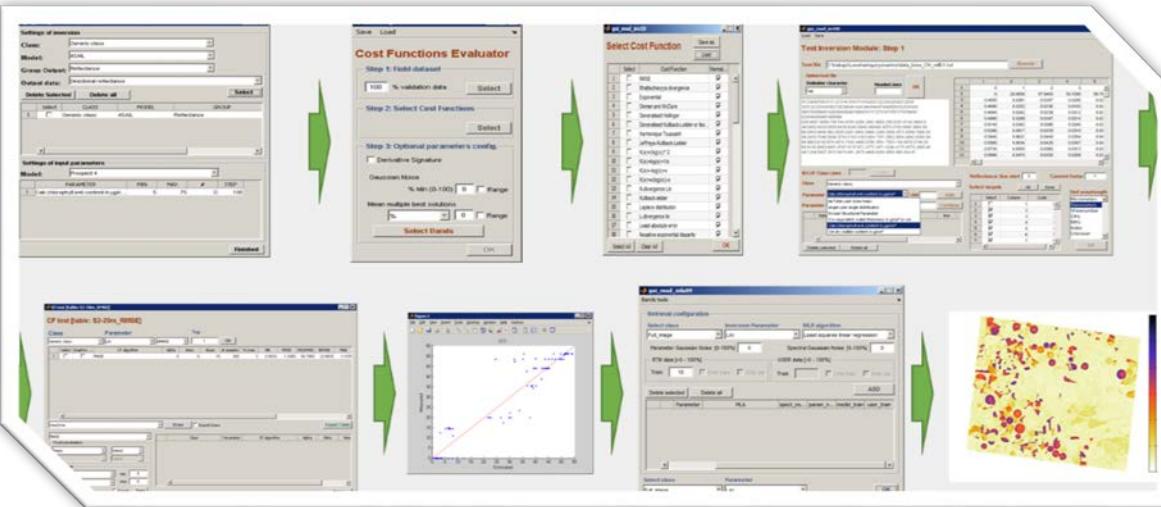
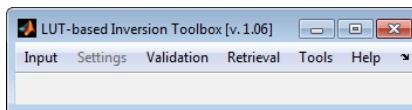
Search per pixel for best match against LUT



Other important factors:

- Adding **noise** (to account for natural variability)
- Selecting mean/median of **multiple solutions**

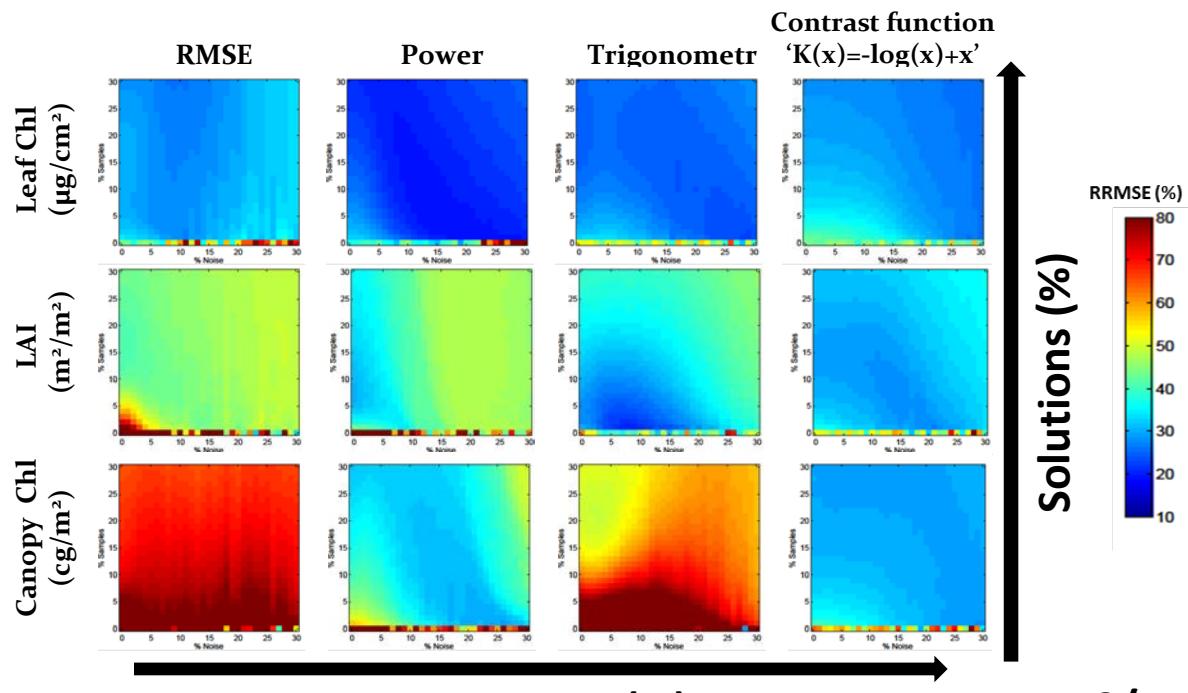
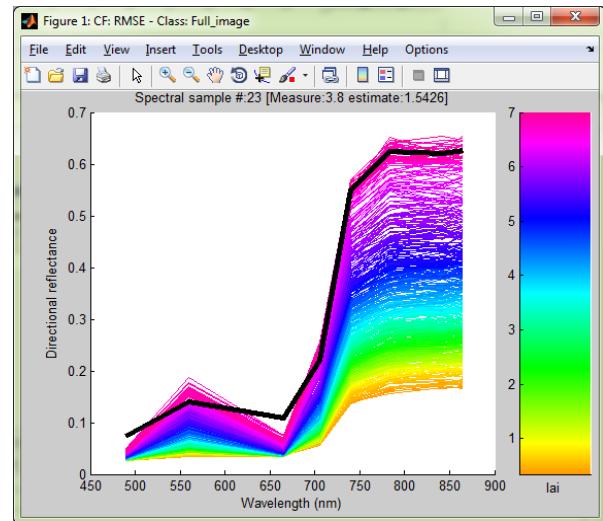
LUT-based inversion toolbox:



Properties:

- LUT ARTMO RTMs or external LUT
- Over 60 different cost functions
- Noise & multiple solutions
- Results stored in MySQL
- Top-performing inversion strategies are given.
- Can apply inversion to both image or individual spectra.

Matching a pixel against a part of the LUT.



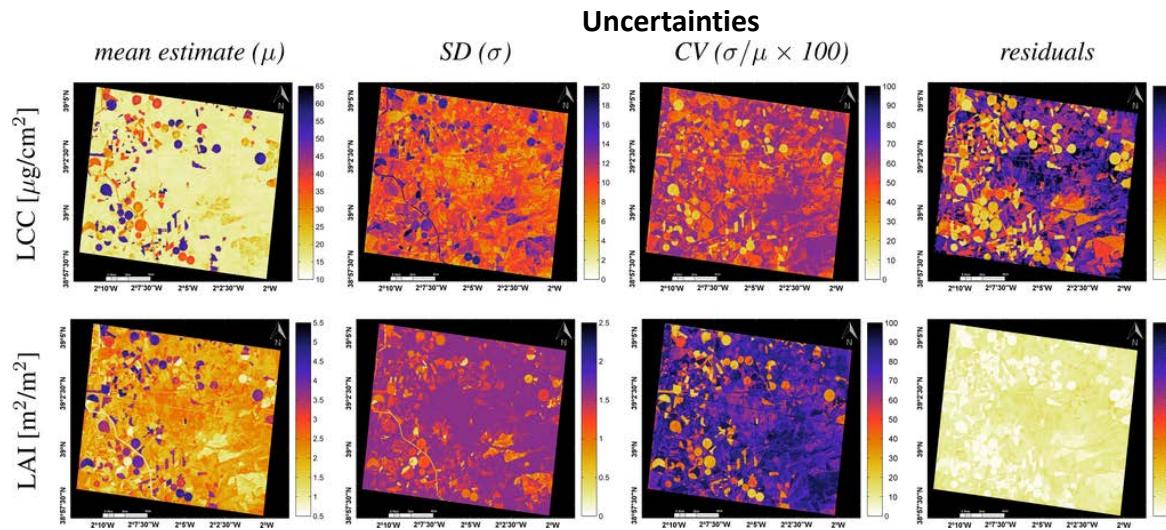
Best-performing method can be applied to an image.

Noise (%)

40/55

SPARC dataset, CHRIS resampled to S2, S3

S2, 20m

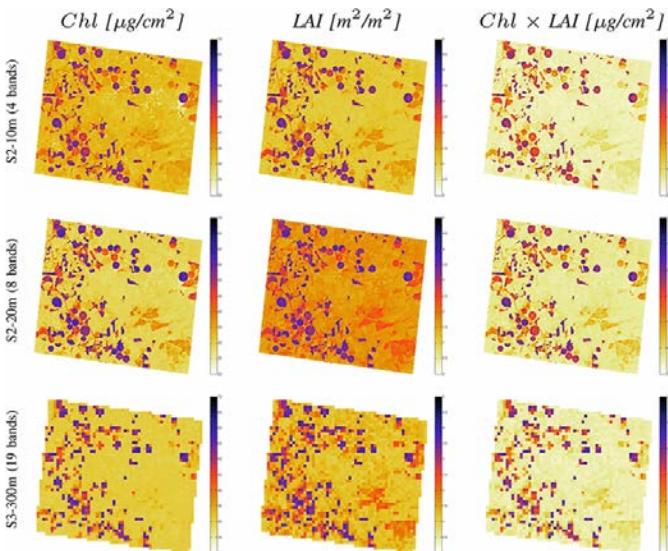


Optimized cost function for each variable.

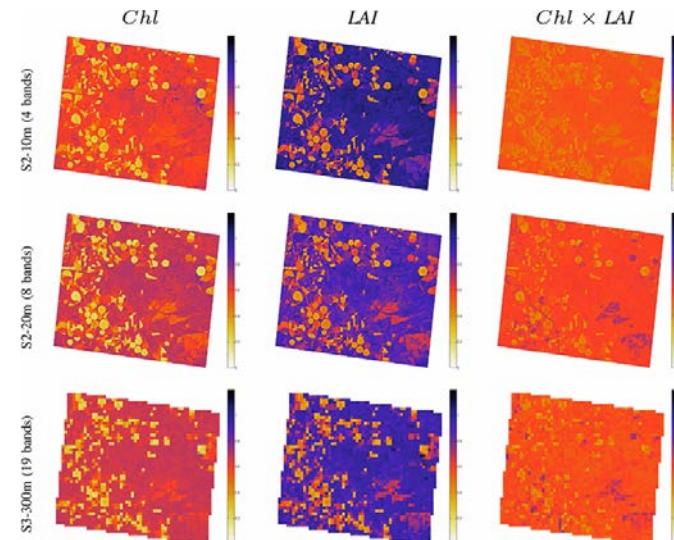
One optimized cost function for simultaneous retrieval of multiple variables.

Sentinel	S2-10m	S2-20m	S3-300m
Spatial resolution [m]	10	20	300
# bands	4	8 (4 + 4 at < 20 m)	19
Band position	B2, B3, B4 and B8	B2 to B8a	O2 to O20
Wavelengths [nm]	490-665 and 842	490-865	413-940

Estimates



Uncertainties (CV)



SIMPLIFYING

- Global sensitivity analysis
- Emulation
- Retrieval



Global sensitivity analysis

Global sensitivity analysis: explores the full input parameter space, i.e. all input parameters are changed together.

Variance-based methods: the output variance is decomposed to the sum of contributions of each individual input parameter and the interactions (coupling terms) between different parameters.

Based on the work of Sobol', variance-based sensitivity measures are represented as follows:

$$1 = \sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \dots + S_{12,\dots,k}$$

in this equation, $S_i, S_{ij}, \dots, S_{12,\dots,k}$ are **Sobol's global sensitivity indices**.

The **first order sensitivity index** S_i measures and quantifies the sensitivity of model **output Y** to the input parameter X_i (without interaction terms), whereas, $S_{ij}, \dots, S_{12,\dots,k}$ are the sensitivity measures for the higher order terms (interaction terms).

The **total effect sensitivity index** S_{Ti} measures the whole effect of the variable X_i , i.e. the first order effect as well as its coupling terms with the other input variables:

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123}$$

GSA toolbox



ARTMO [v. 3.19]

File Models Forward Retrieval Tools Help

Project Description

GSA configuration

Save Load

Project PRO4SAIL_1000

Type SI Saltelli

subsamples 1000

RT model 4SAIL-Prospect 4

Sensor NO SENSOR

RTM input settings

Group Canopy

Parameter Total Leaf Area Index

External Parameters Empty

Range min 0 max 10 Distribution Sobol

Select Parameter m

1	<input type="checkbox"/> Leaf Structural Parameter	
2	<input type="checkbox"/> chlorophyll a+b content i...	
3	<input type="checkbox"/> equivalent water thickne...	1.000
4	<input type="checkbox"/> dry matter content in g/c...	1.000
5	<input type="checkbox"/> Total Leaf Area Index	
6	<input type="checkbox"/> Leaf angle distribution	
7	<input type="checkbox"/> Diffuse/direct light	

All Add parameter

RTM outputs

Group Canopy

Parameter Directional reflectance

Add output

Select Parameters Group

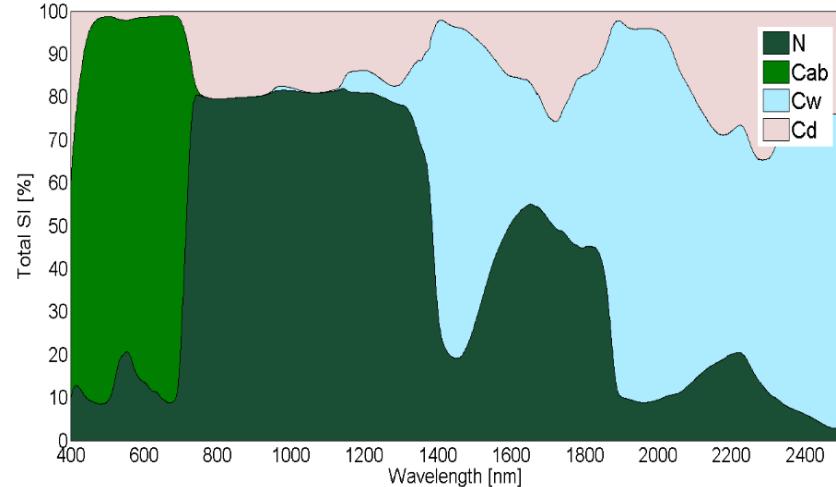
1	<input type="checkbox"/> Directional reflectance	Canopy
---	--	--------

Run

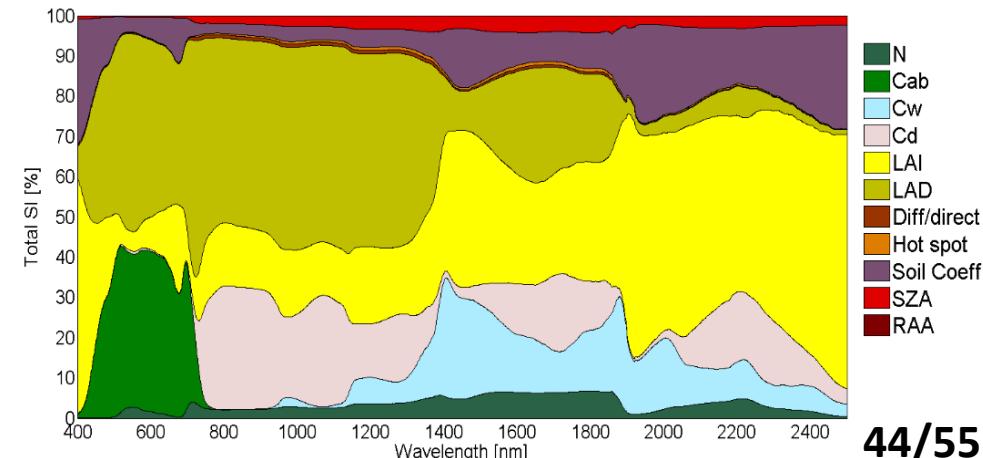
Properties:

- ARTMO RTMs
- *Saltelli 2010* GSA method
- Various sample distributions
- Results stored in MySQL
- First order or total order Sobol Sensitivity indices
- Can process multiple RTM outputs.

PROSPECT-4 Reflectance (#1000)

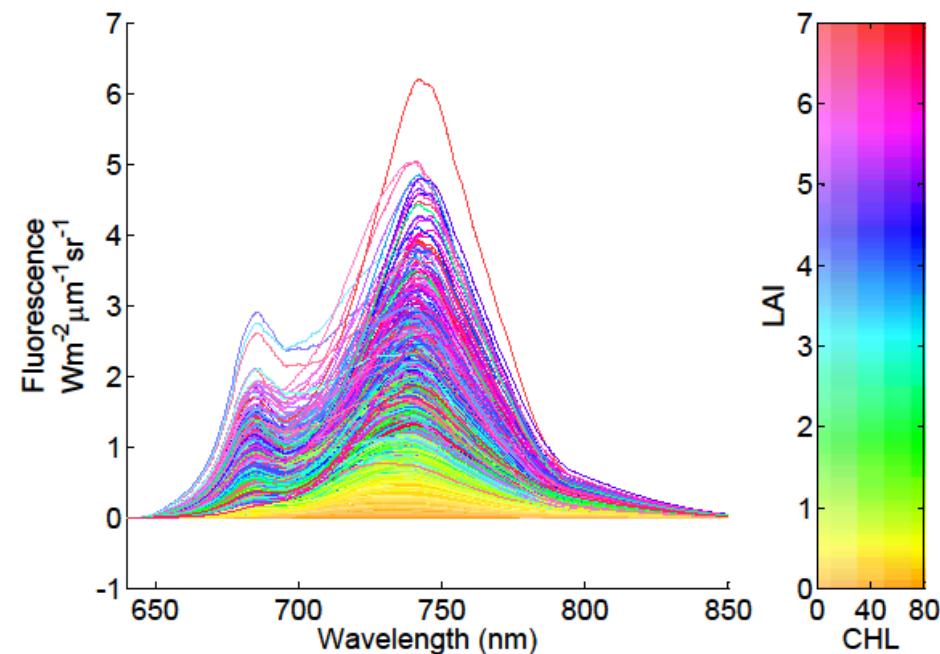


PROSAIL Directional Reflectance (#1000)



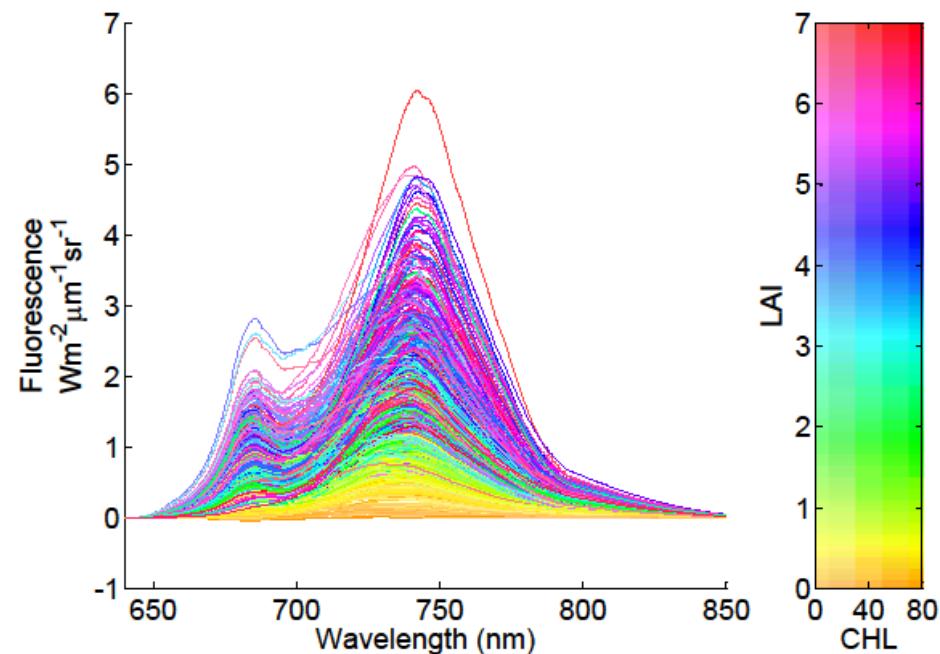
Any difference?

Which one would you choose?



12 min 54 s.

SCOPE



1 s.

Metamodel (emulator)

Emulation

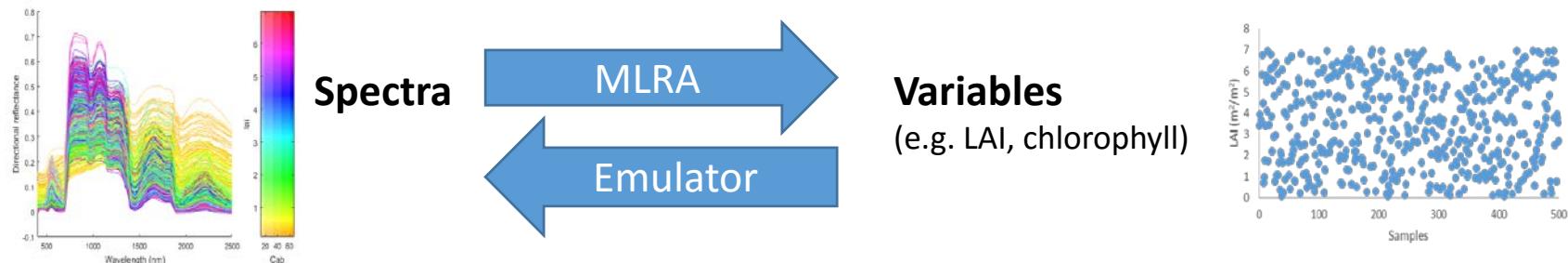
Emulators are regression models that are able to approximate the processing of an RTM, at a fraction of the computational cost:



making a statistical model of a physical model

Emulators applied to RTMs:

- In principle any nonlinear, adaptive **machine learning regression algorithms (MLRAs)** can serve as emulators.

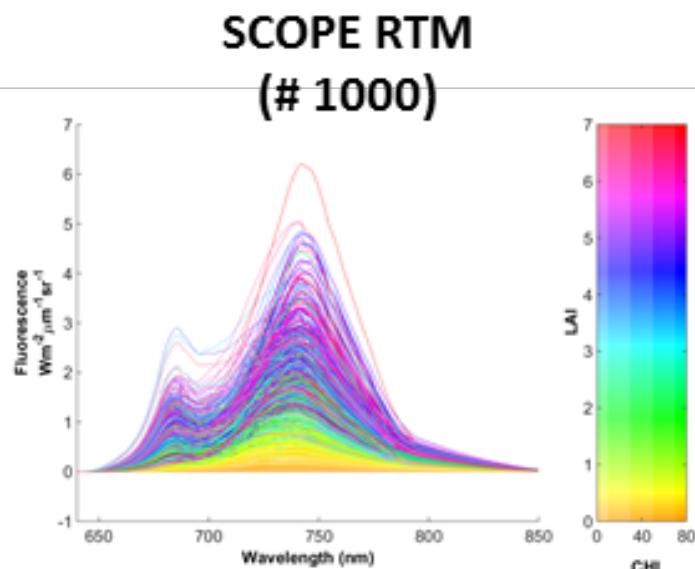


- To emulate RTMs, the emulator should have the capability to reconstruct multiple outputs, i.e. the complete spectrum: resolved with **dimensionality reduction** techniques (e.g. PCA).

Processing steps:



Emulating SCOPE fluorescence outputs



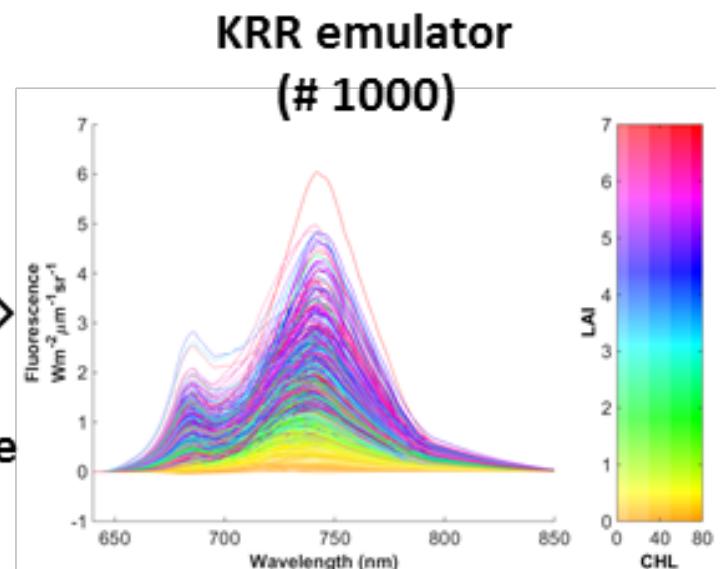
Emulation

Training validating →

Multi-output machine learning regression algorithm (e.g. KRR)



774 s.



1 s.

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



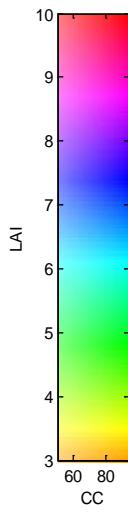
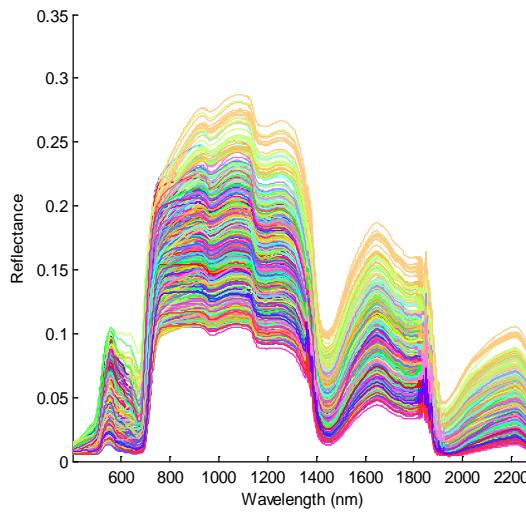
In Emulation, physical models go hand in hand with machine learning

Emulating a complex 3D RTM: DART

Experimental setup:

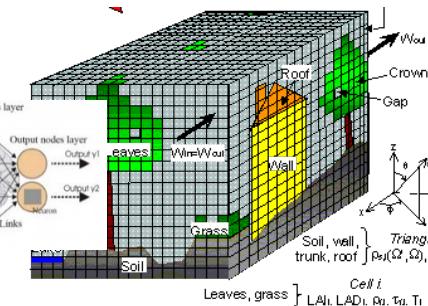
- DART: LUT1000# @ 1 nm; 7 variables
- 3 MLRAs tested: KRR, NN, GPR
- Various # PCA components tested (5, 10, 20, 30)

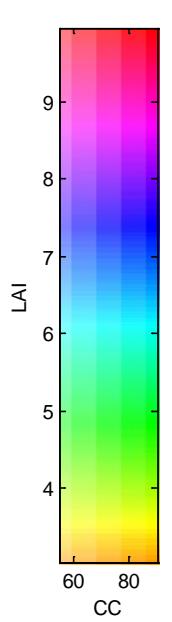
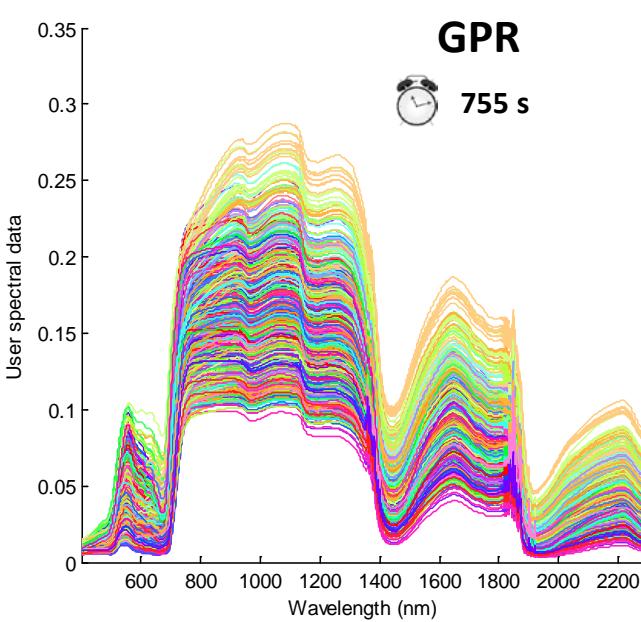
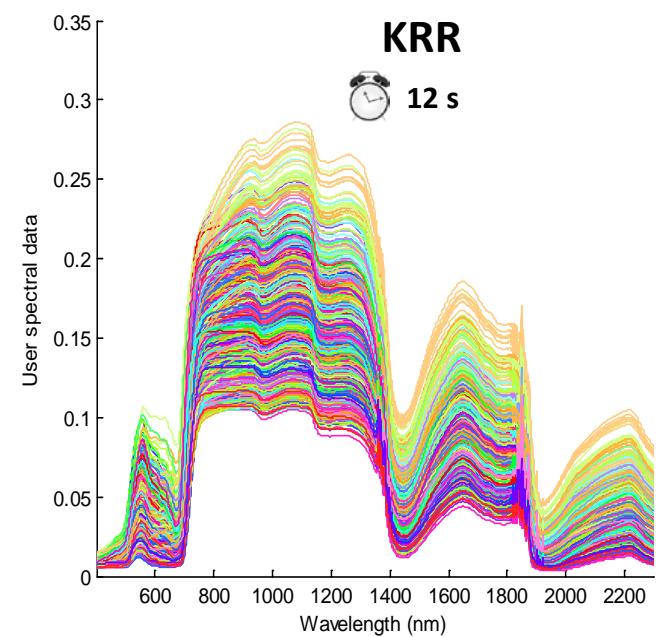
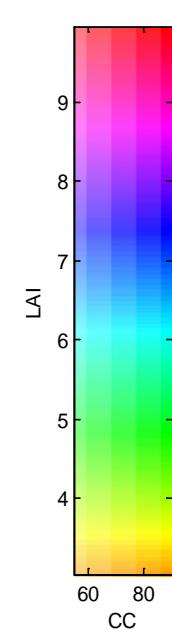
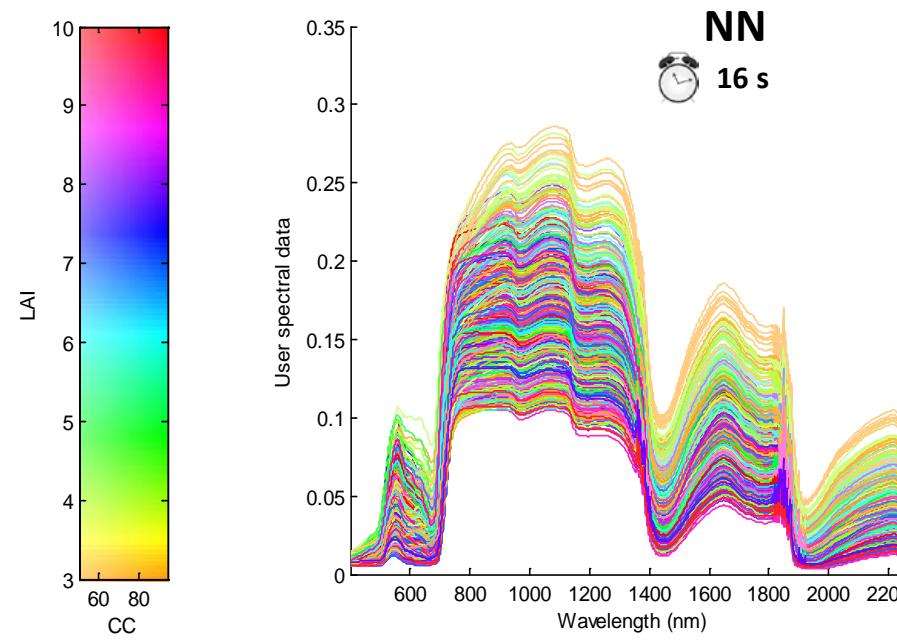
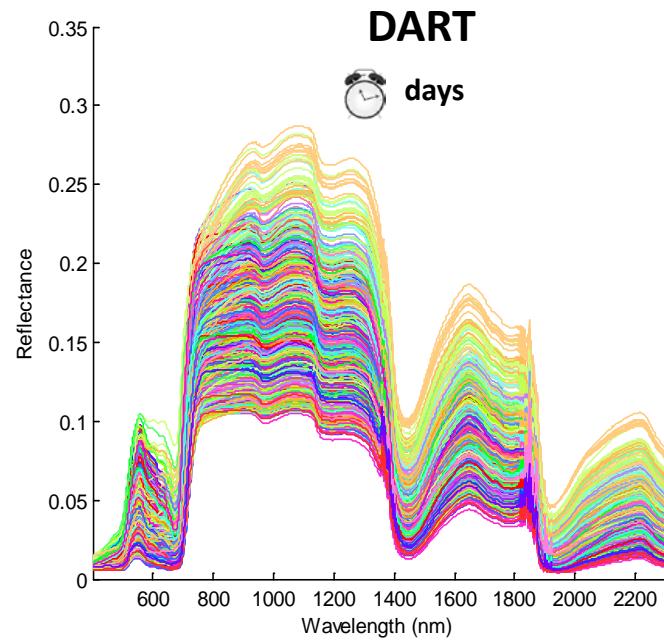
- N
- LWC
- DMC
- Carc
- LCC
- CC
- LAI
- TOPO



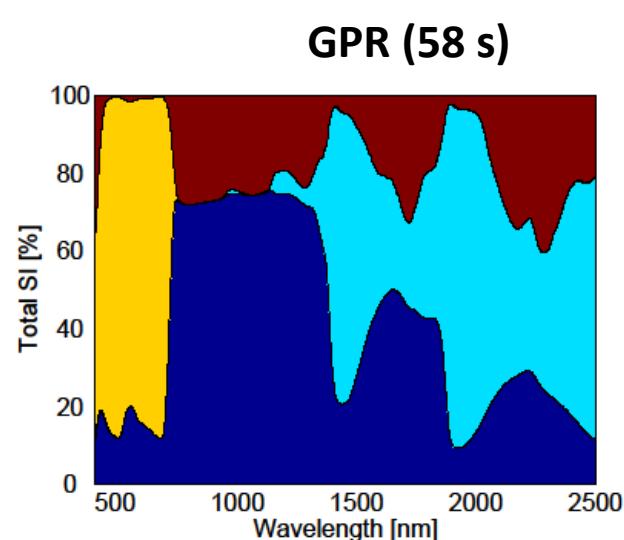
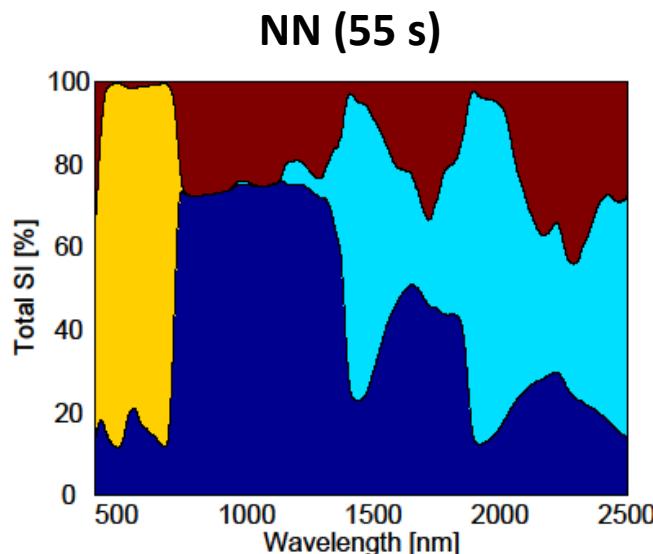
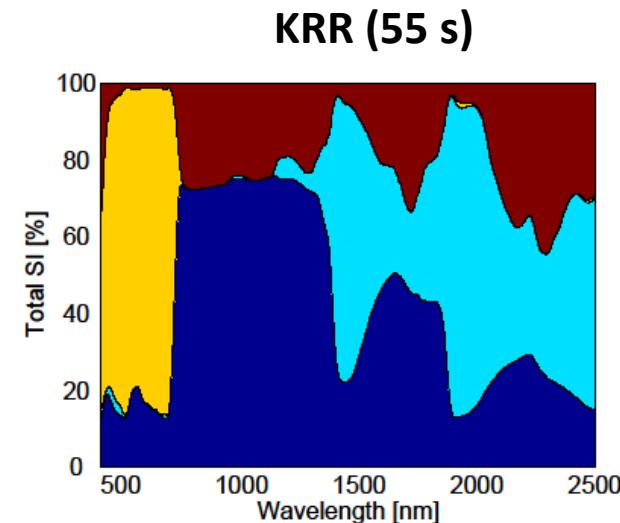
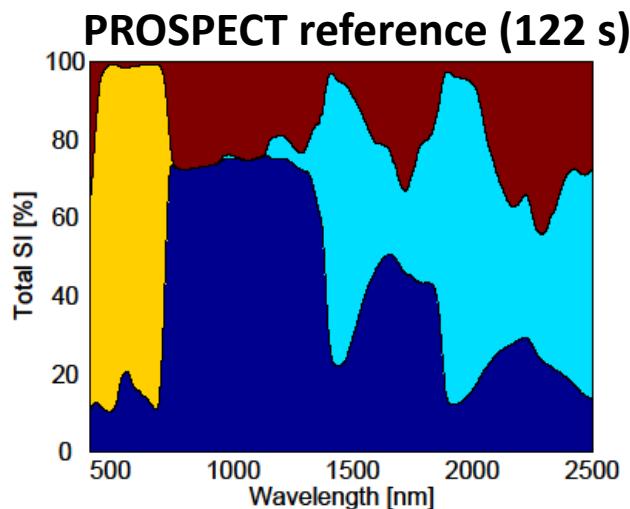
Emulator

- KRR
- NN
- GPR





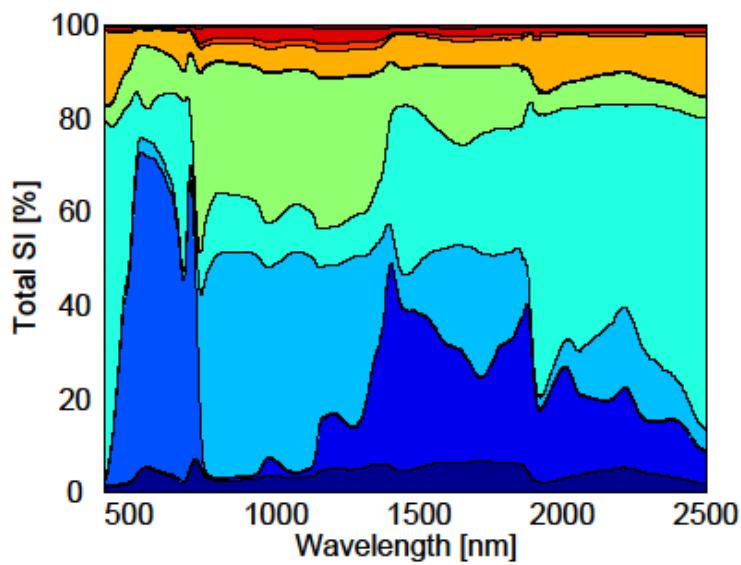
Emulators applied into GSA: PROSPECT-4



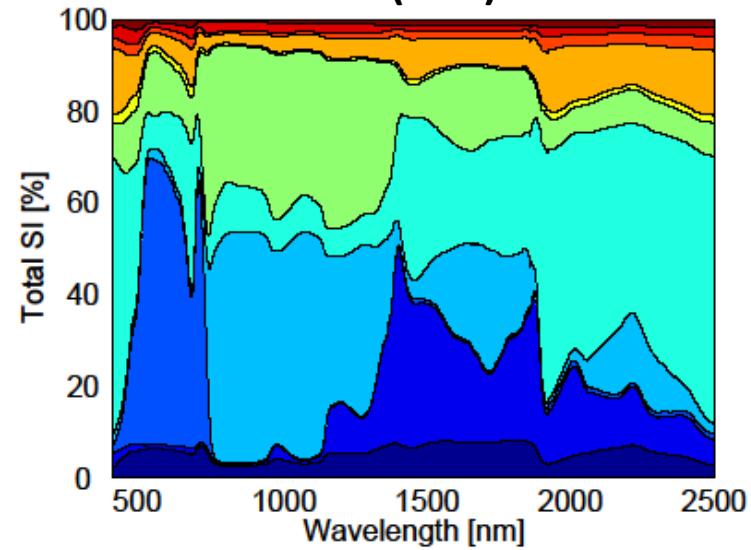
N Cw Cab Cm

PROSAIL emulators

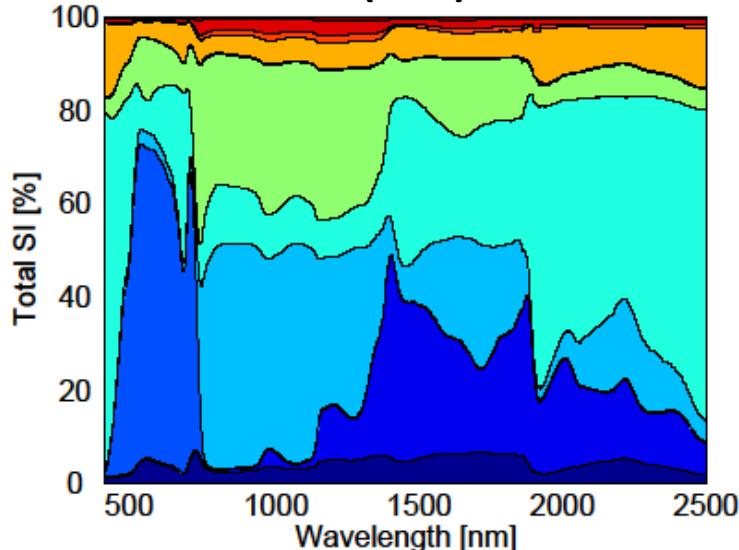
PROSAIL reference (279 s)



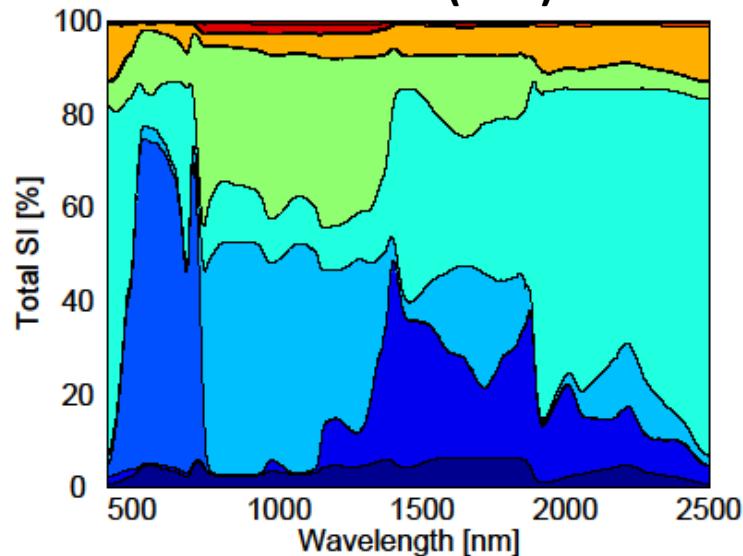
KRR (73 s)



NN (80 s)



GPR (79 s)

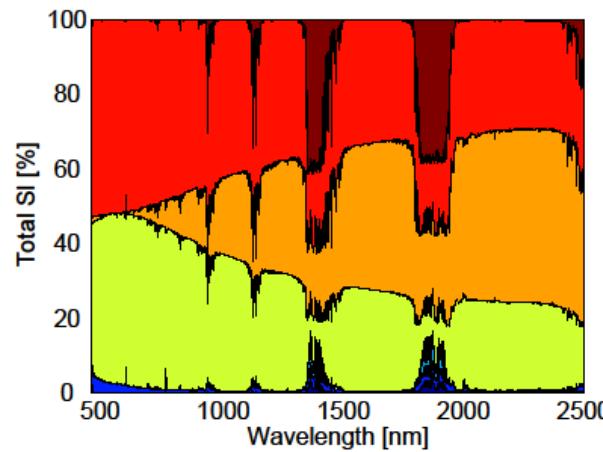


MODTRAN

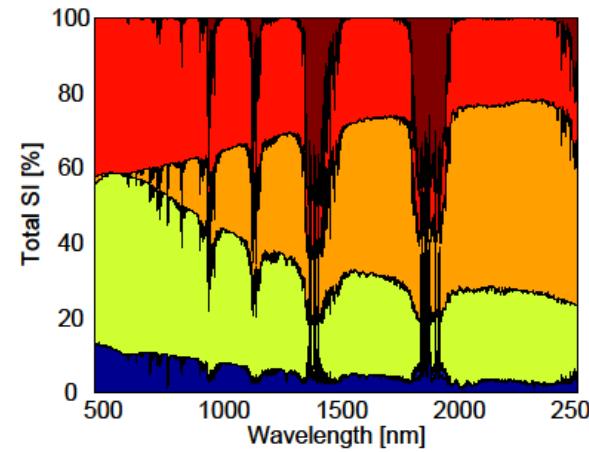
atmospheric transfer functions:

$$L_{TOA} = L_0 + \frac{(E_{dir}\mu_s + E_{dif})(T_{dif} + T_{dir})\rho}{\pi(1 - S\rho)}$$

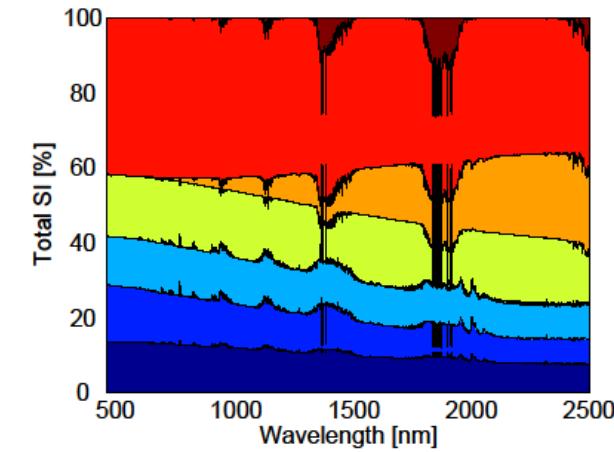
E_{dif} (GPR: 121 s)



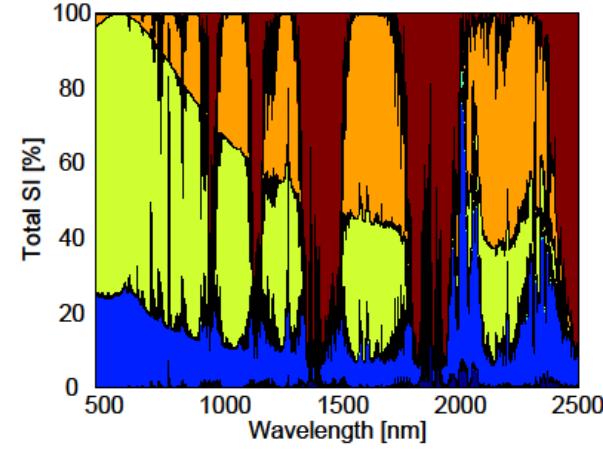
T_{dif} (NN: 157 s)



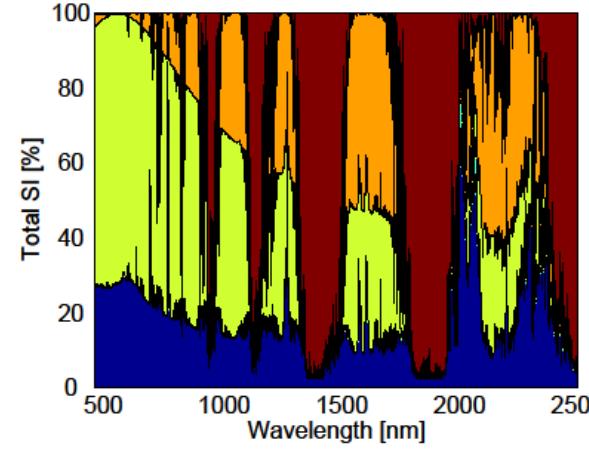
L_0 (GPR: 121 s)



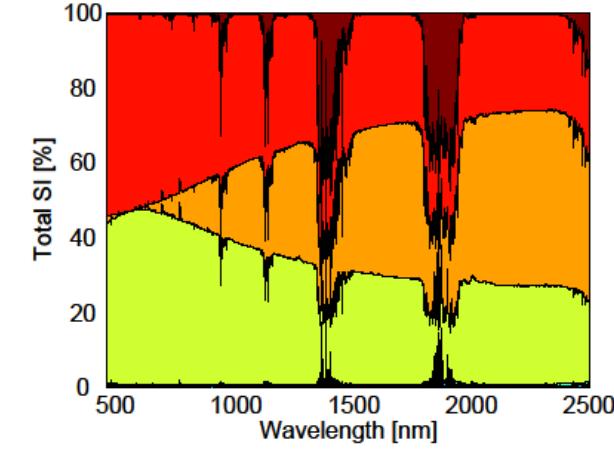
E_{dir} (KRR: 101 s)



T_{dir} (NN: 151 s)



S (GPR: 166 s)



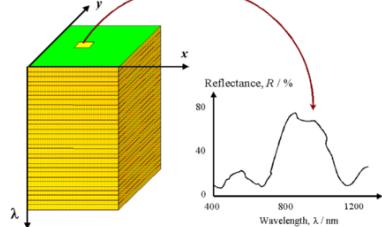
1000#/variable

VZA SZA RAA ELEV AOT AMS G CWV

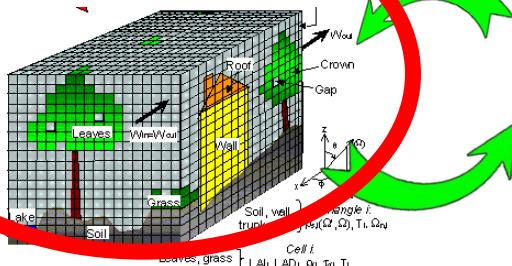
With emulators, processing speed is boosted in the orders of hundred thousand. Using the original MODTRAN simulations would take more than a month.

Emulators into numerical inversion

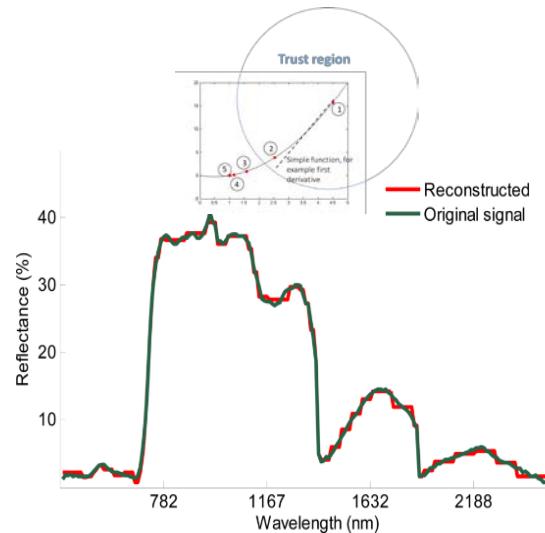
Image



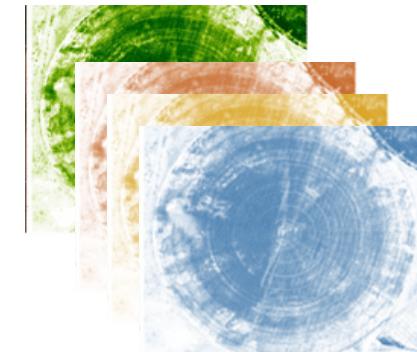
RTM (e.g., DART)



Minimization algorithm: lsqnonlin

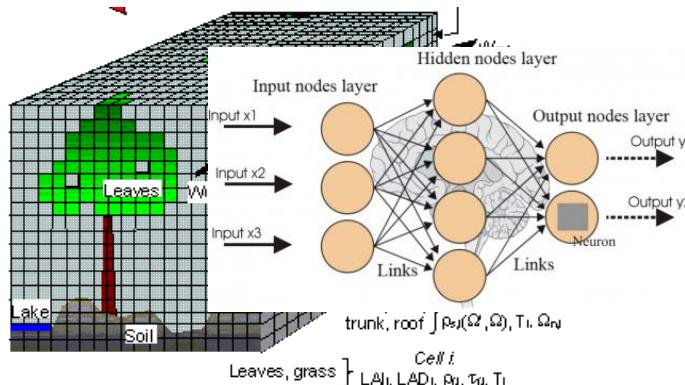


Output maps of RTM variables



Per-pixel RTM iterations: very slow method, inapplicable to computationally expensive RTMs.

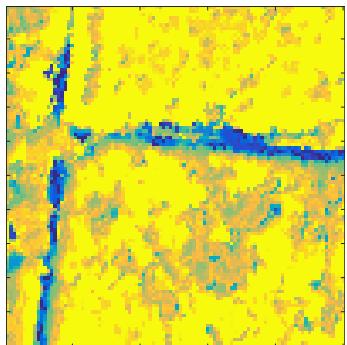
Emulation of an RTM



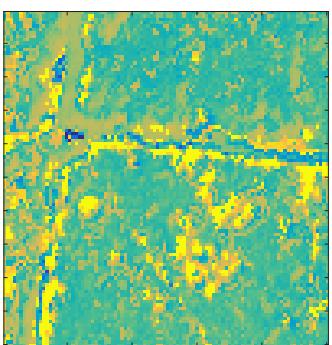


DART KRR emulator applied to
HyPlant DUAL (450-2500 nm)

CC



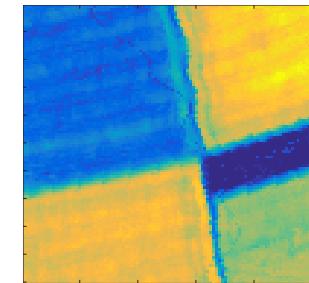
LCC



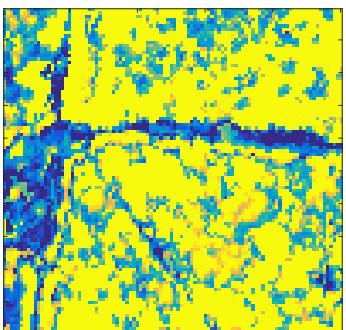
SCOPE KRR emulator applied
HyPlant DUAL (bare soil spectra added)

LAI

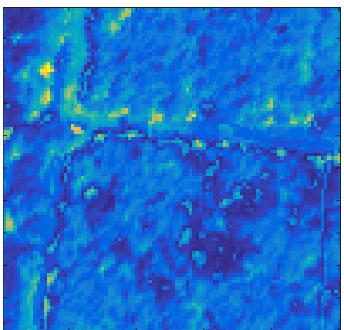
APAR



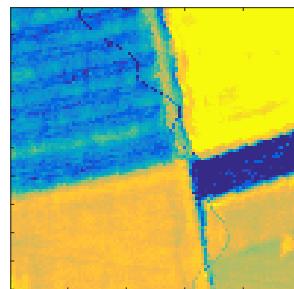
LAI



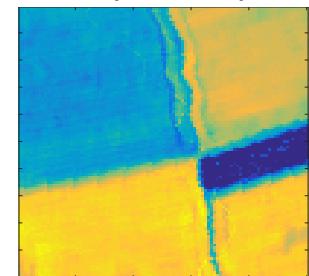
RMSE



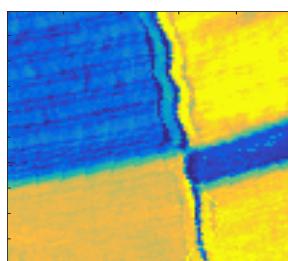
fAPAR



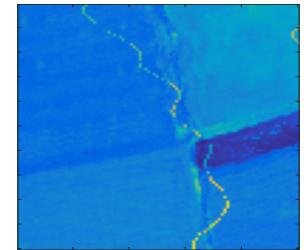
CCC (LCC x LAI)



CWC (Cw x LAI)



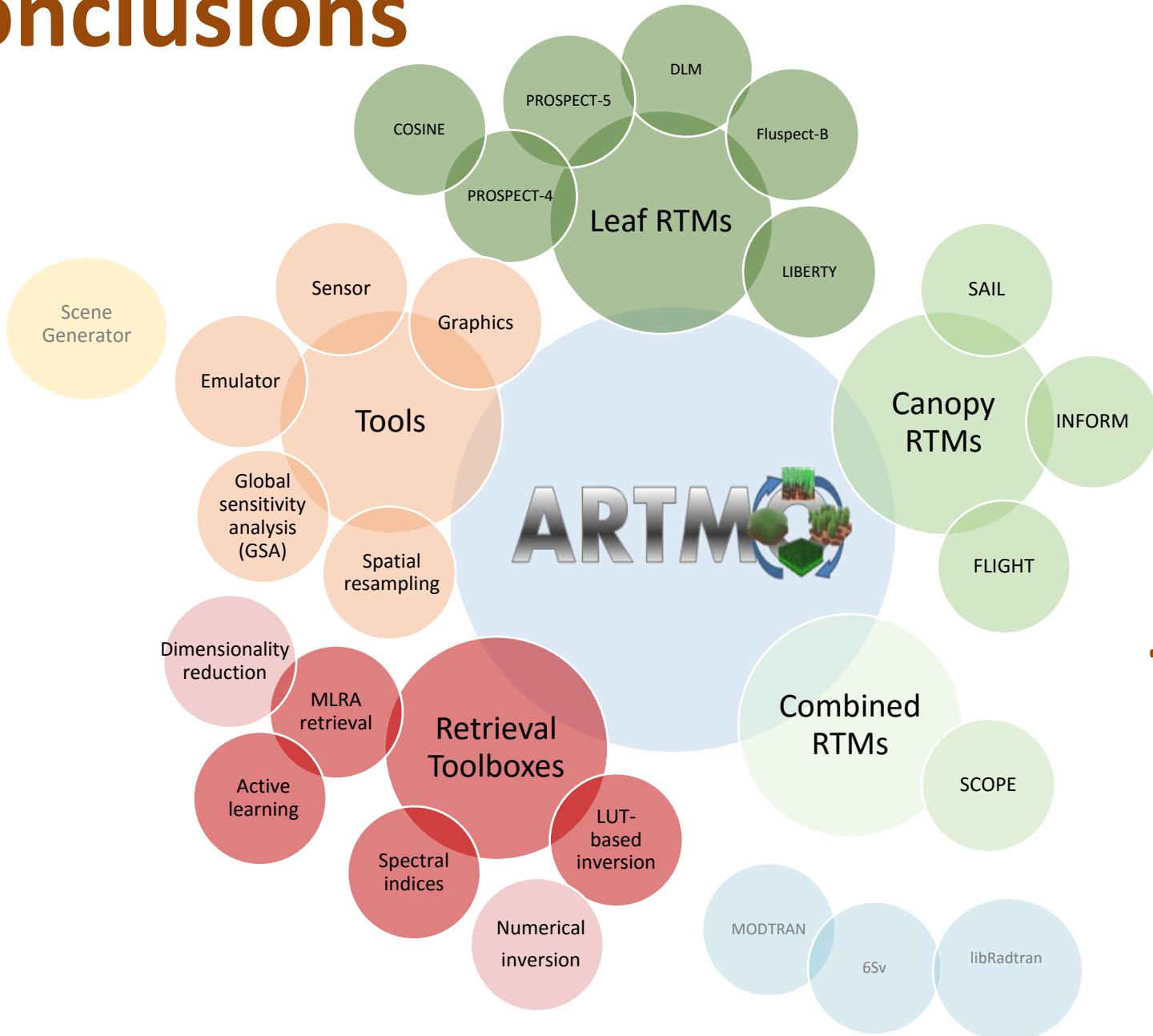
RMSE



Retrieval quality depends on : (1) emulator, (2) number and type of included variables.



Conclusions



Thanks