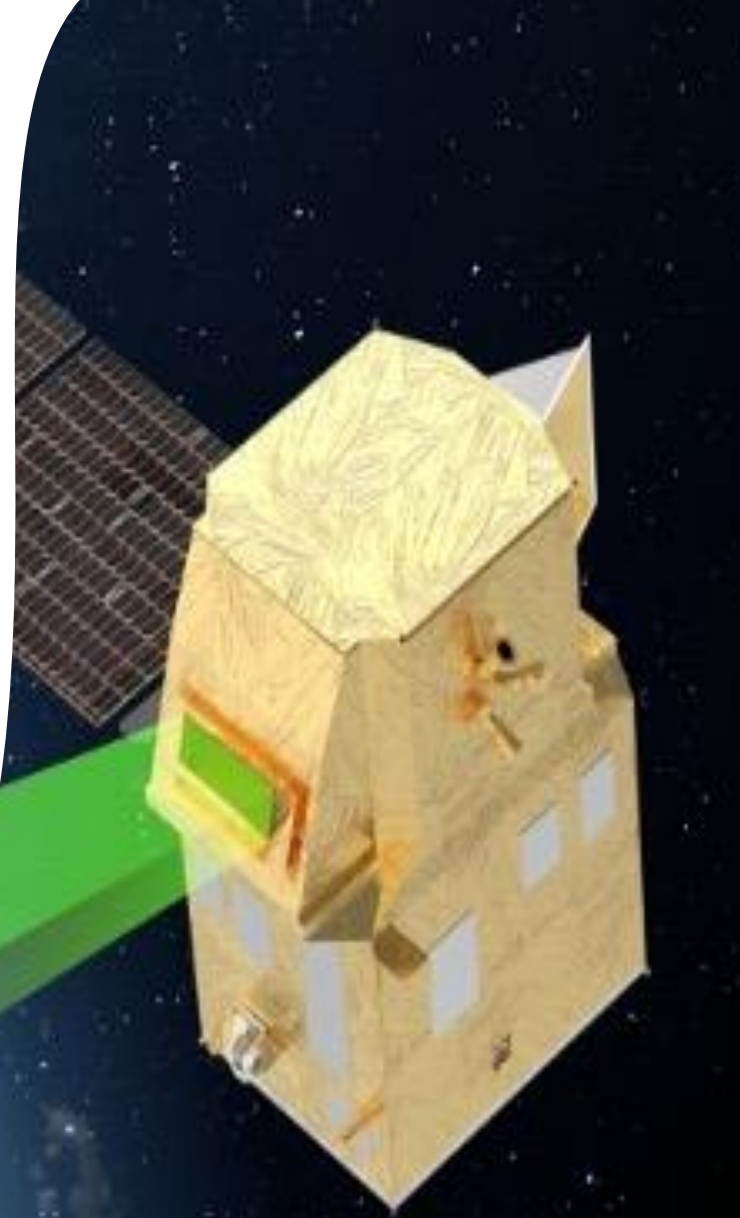
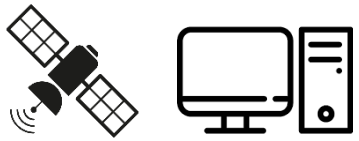
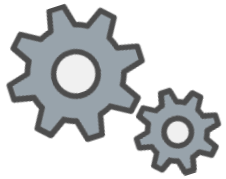
EARSeL València 2024

Latest status of the vegetation traits retrieval processor and models in the context of CHIME mission preparation

José Luis García-Soria, Miguel Morata, Jochem Verrelst, Ana Belén Pascual-Venteo,Katja Berger, Cinzia Panigada, Giulia Tagliabue, Ana María Sánchez Montero



Outline of the presentation

- **CHIME** & **CHEES** 
- Retrieval strategies… why **hybrid**?
- MLRA & **uncertainties**
- How does the module work? 
- CHIME **models evaluation** strategy and results
- **Conclusions**

CHIME: Copernicus Hyperspectral Imaging Mission for the Environment

CHIME will carry a visible to shortwave infrared spectrometer to provide global routine hyperspectral observations.

Supporting:

- Sustainable agricultural and biodiversity management.
- Soil property characterisation.

Technical concept:

- Instrument: Pushbroom Imaging Spectrometer **400 – 2500 nm, Δλ <= 10nm**
- Revisit **11 days**
- GSD (spatial resolution): **30 m**
- **Sun synchronous orbit** (LTDN 10:30 – 11:30)
- **Nadir view** covering land and coastal areas
- High radiometric accuracy, low spectral/spatial misregistration

CORE Data Products:

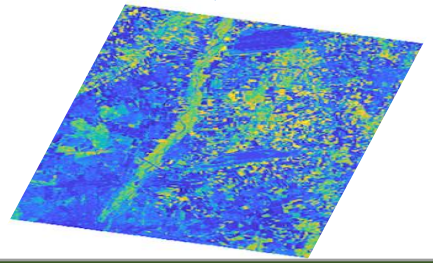
- **L1C:** TOA radiance
- **L2A:** atmospherically corrected and ortho-rectified reflectances
- **L2BM:** Mineral products
- **L2BS:** Soil products
- **L2BV:** Vegetation products

Priority

- **Canopy nitrogen content**
- Canopy chlorophyll content
- Canopy water content
- **Leaf area index**
- **Leaf nitrogen content**
- **Leaf mass area**
- Leaf chlorophyll content
- Leaf water content

What is CHEES? CHime End to End Simulator

- OPSI Geometry Module
- Scene Generator Module
- OPSI Instrument and Processing Module
- L2 GPP Module
- L2B Retrieval Module



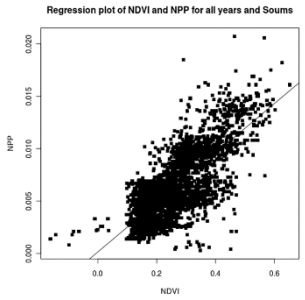
Performance Assessment Module

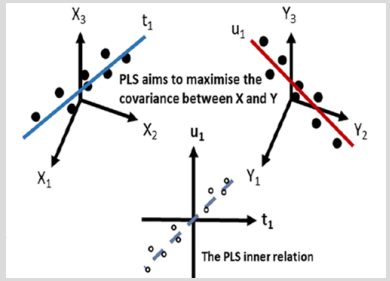
- **OGM**: Simulate CHIME orbit and attitude
- **SGM**: Simulates the scene to be observed at fine spectral and spatial resolution.
- **OIP**: Models the CHIME sensor behavior.
- **L2GPP**: Performs the atmospheric correction
- **L2B**: Implements the algorithms for retrieval of **bio-/geophysical variables**, objective of the mission (e.g., retrieval of mineral content/soil/**vegetation**).
- **PAM**: Evaluates the product performance at different processing levels

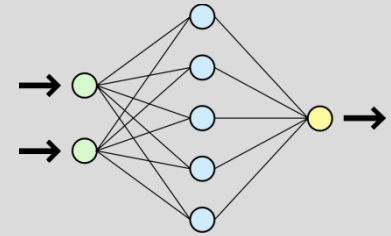
Retrieval methods for vegetation properties mapping

Parametric regression	Non-parametric regression	RTM inversion
Spectral relationships that are sensitive to specific vegetation properties	Data-driven techniques that search for relationships between spectral data and biophysical variables	Models that simulate interactions between vegetation and radiation

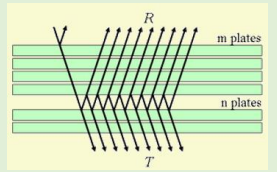
$$NDVI = \frac{(\rho_{NIR} - \rho_{RED})}{(\rho_{NIR} + \rho_{RED})}$$

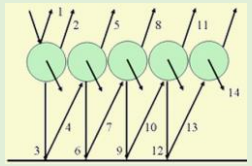




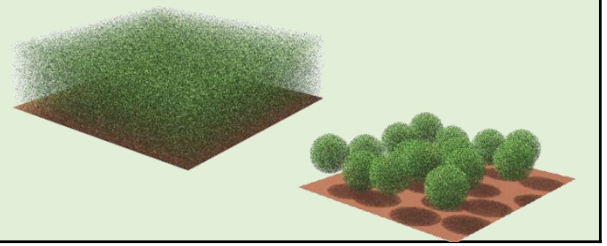


leaf

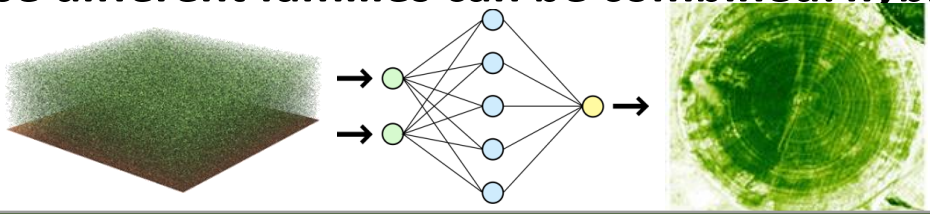




canopy



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



Key capabilities in **operational** context

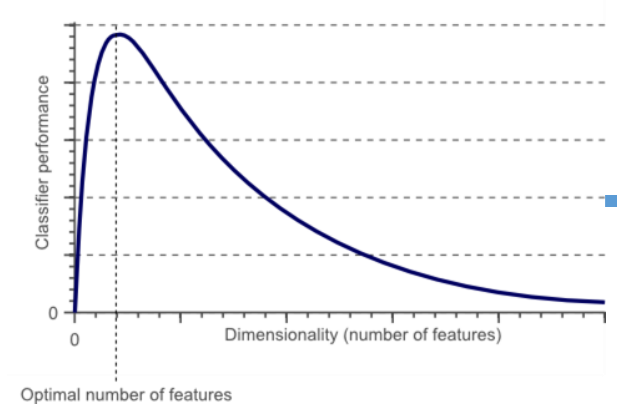
Characteristic	Parametric	Non-parametric	RTM-based	Hybrid
Generalization capacity	--	-	++	++
Mapping Speed	++	+	--	+
Uncertainties	--	++*	+	++*
Accuracy	+	++	+	++

* Some machine learning methods (e.g. probabilistic or ensemble methods)

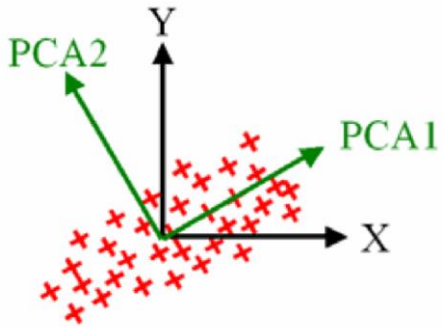
The training data matters in hybrid models…

Spectral domain…

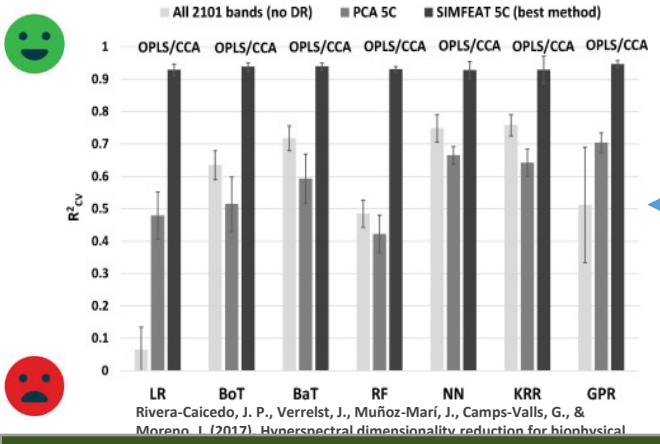
Curse of dimensionality



Principal Components analysis



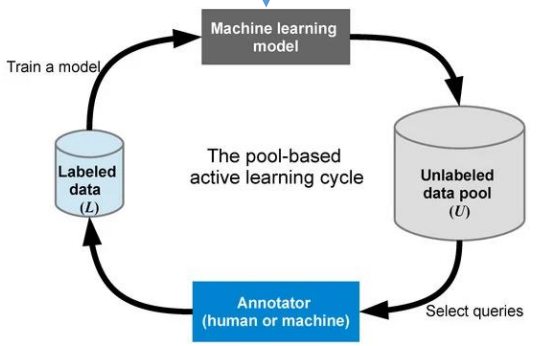
Reduced number of PC



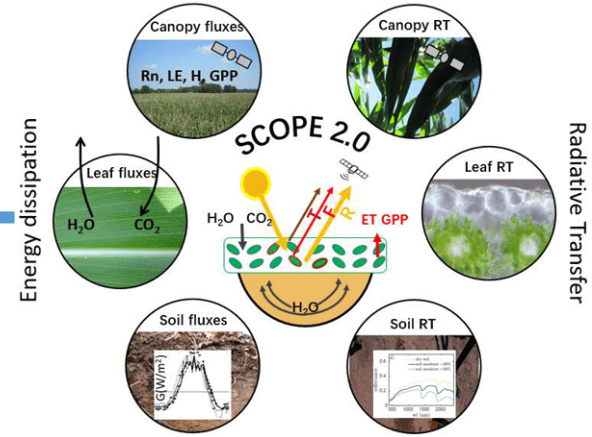
Rivera-Caicedo, J. P., Verrelst, J., Muñoz-Marí, J., Camps-Valls, G., & Moreno, J. (2017). Hyperspectral dimensionality reduction for biophysical

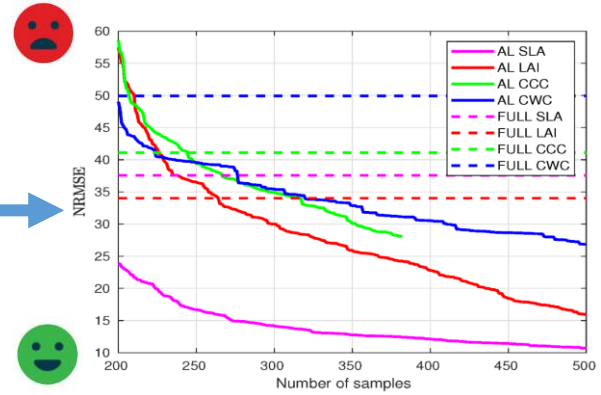
Samples domain…

Active Learning



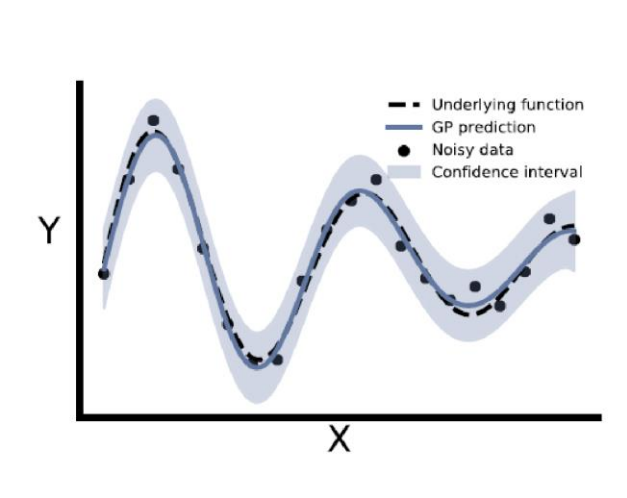
Optimal samples + randomness



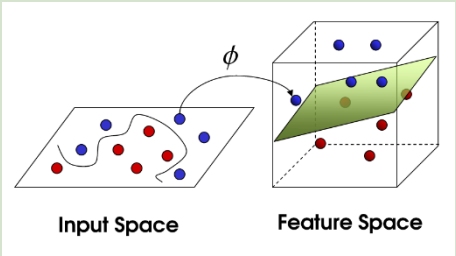


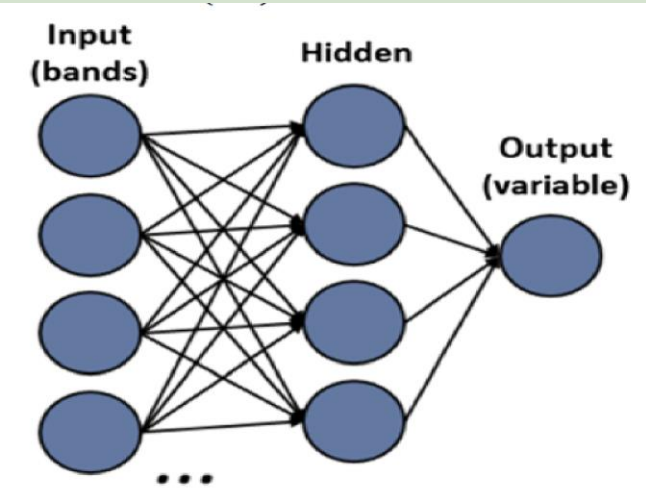
Which MLRA to use?

Gaussian Processes Regression (GPR)	Kernel Ridge Regression (KRR)	Neural Network (NN)
-Bayesian method. -Outputs mean and variance.	-Ridge regression in a higher dimensional space.	-Artifical neural network.

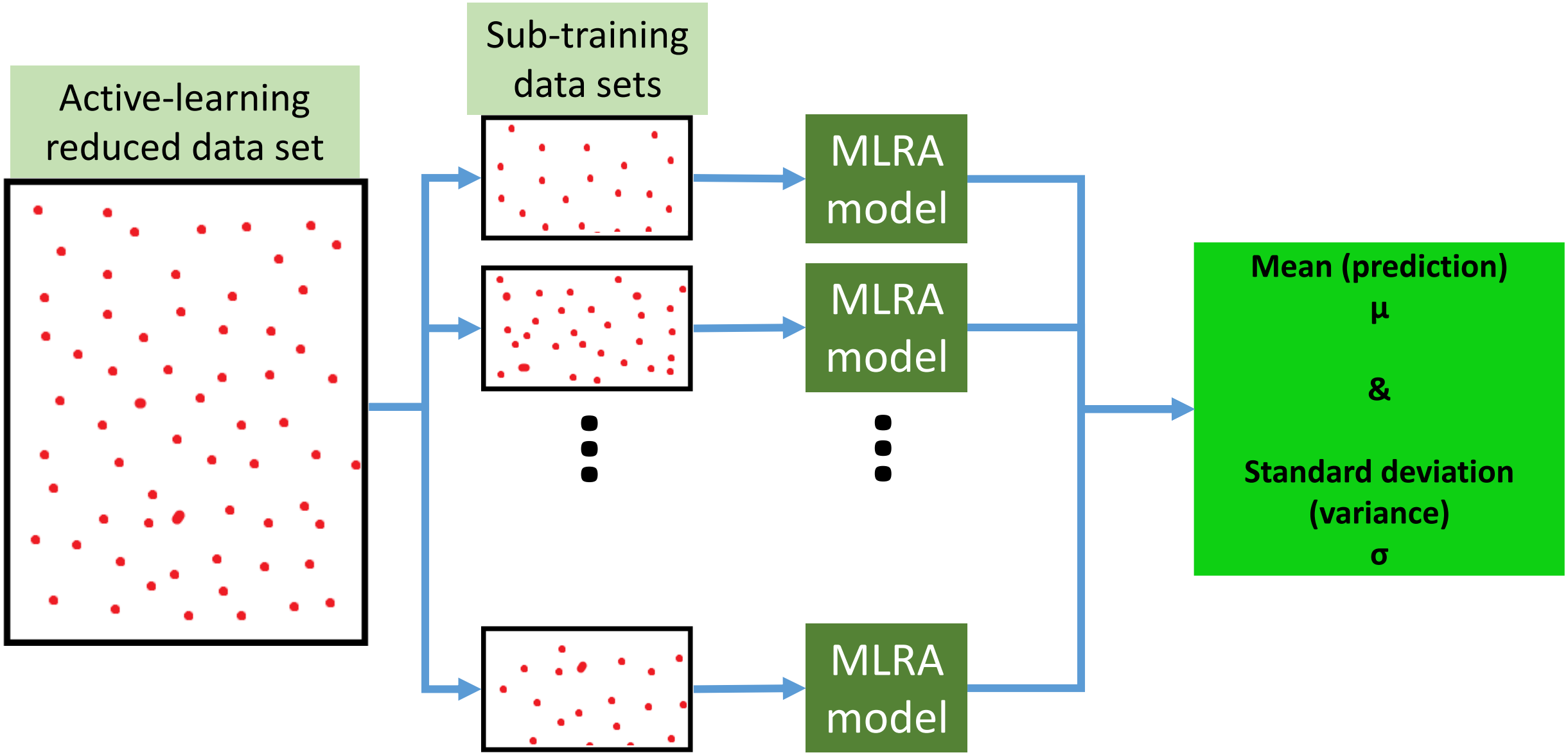


$$C = \frac{1}{2}\sum_i (y_i - w^T x_i)^2 + \frac{1}{2}\tau\|w\|^2$$

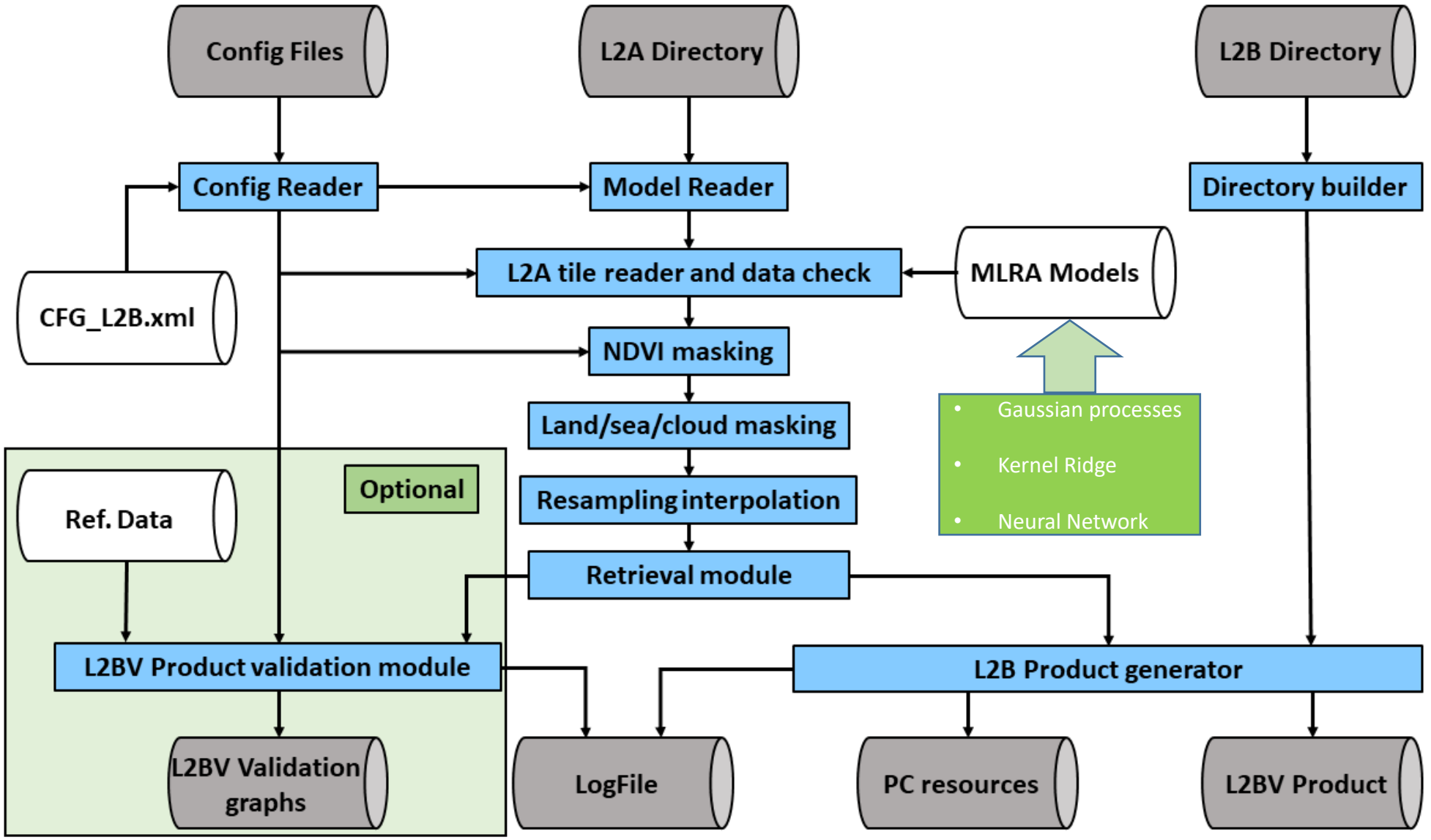




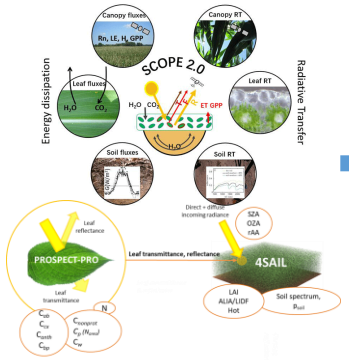
Uncertainty of the models bootstrapping…

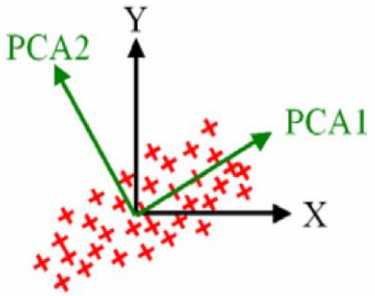


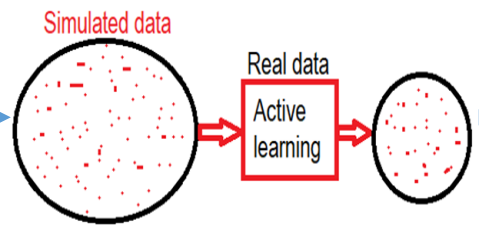
How does the L2BV module work?

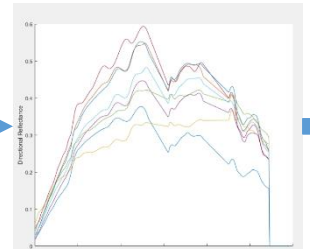


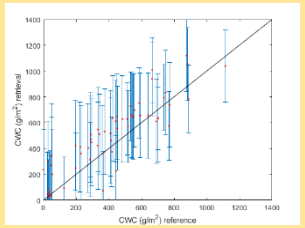
Evaluation of vegetation models…





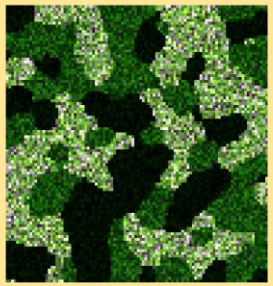




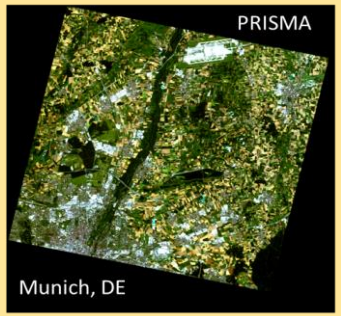


Experimental

Output: reflectance spectrum in observation direction
R: 650nm G:550nm B: 449nm



Synthetic



Satellite images

Gaussian Processess

Kernel Ridge Regression

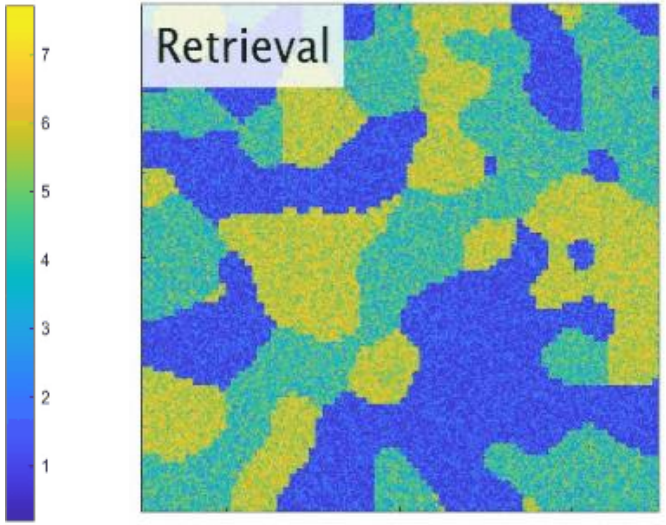
Neureal Networks

Retrievals from the L2BV: Canopy variables

LAI (m^2m^{-2})

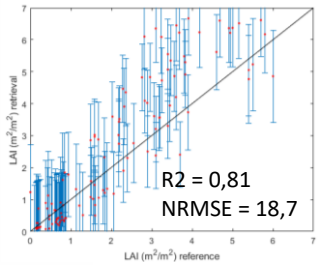
GPR

Synthetic



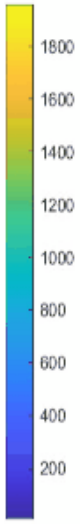
R^2 = 0,78
NRMSE = 15,4

Experimental

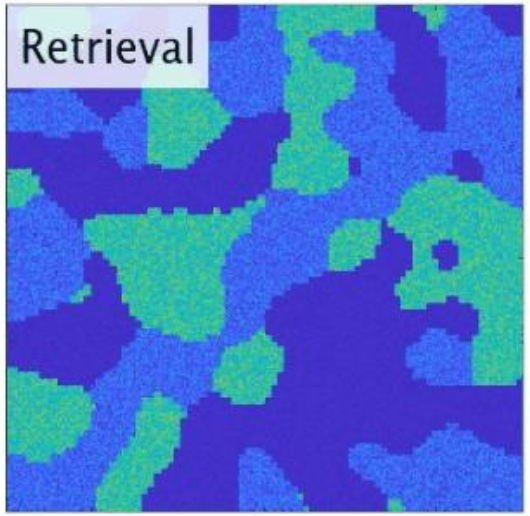


CWC ($g\ m^{-2}$)

GPR

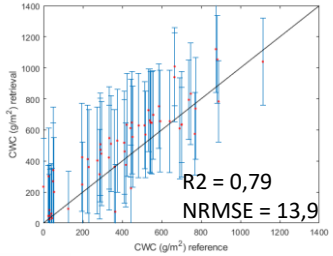


Synthetic



R^2 = 0,89
NRMSE = 10,2

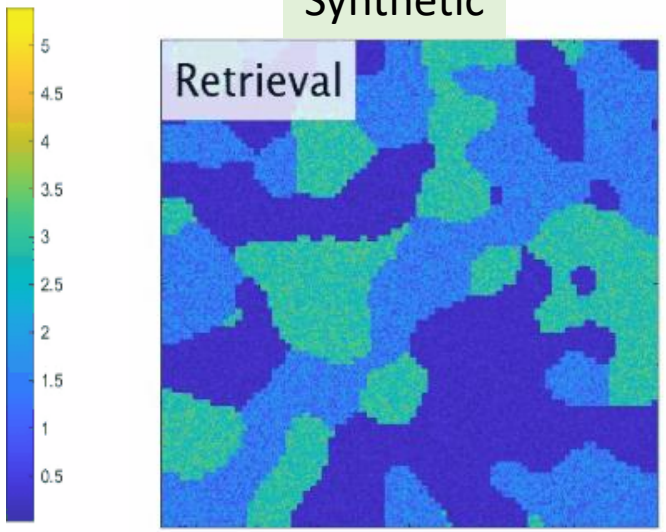
Experimental



CCC (gm^{-2})

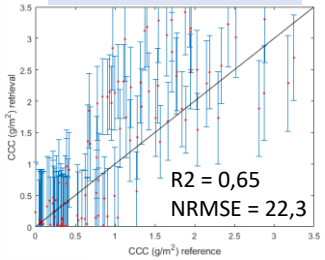
GPR

Synthetic



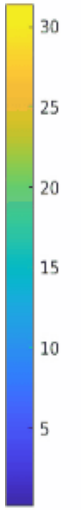
R^2 = 0,82
NRMSE = 12,5

Experimental

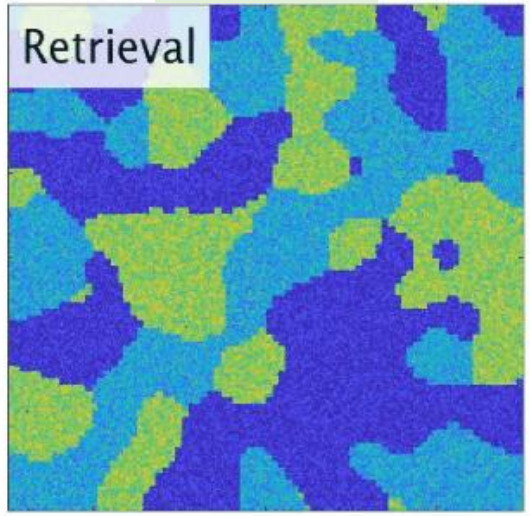


CNC ($g\ m^{-2}$)

GPR

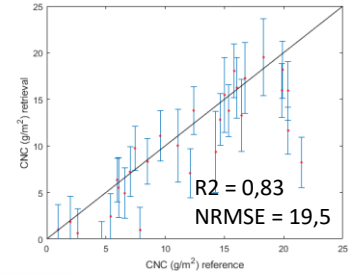


Synthetic



R^2 = 0,79
NRMSE = 12,9

Experimental

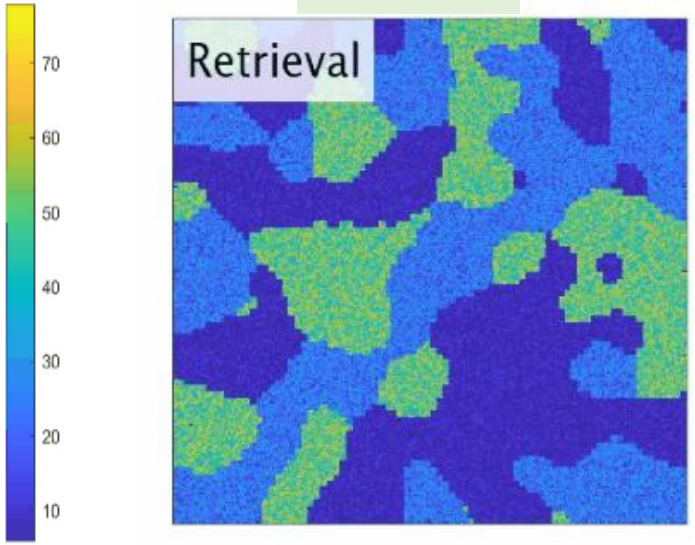


Retrievals from the L2BV: Leaf variables

LCC ($\mu g\ cm^{-2}$)

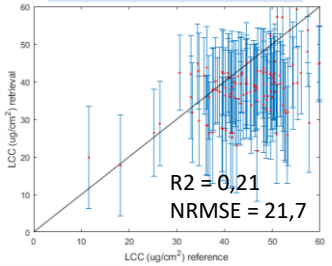
GPR

Synthetic



R^2 = 0,86
NRMSE = 13,2

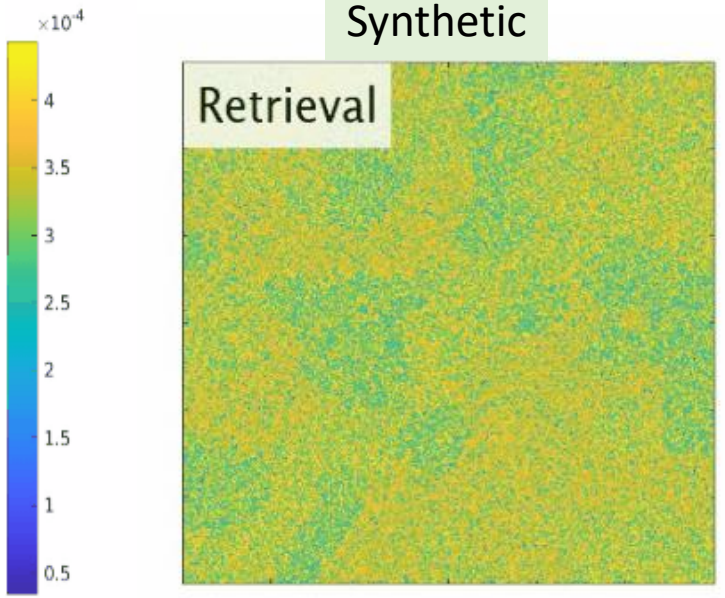
Experimental



LNC ($g\ cm^{-2}$)

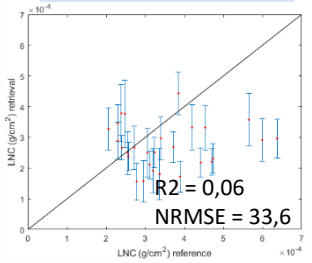
KRR

Synthetic



R^2 = 0,00
NRMSE = 35,0

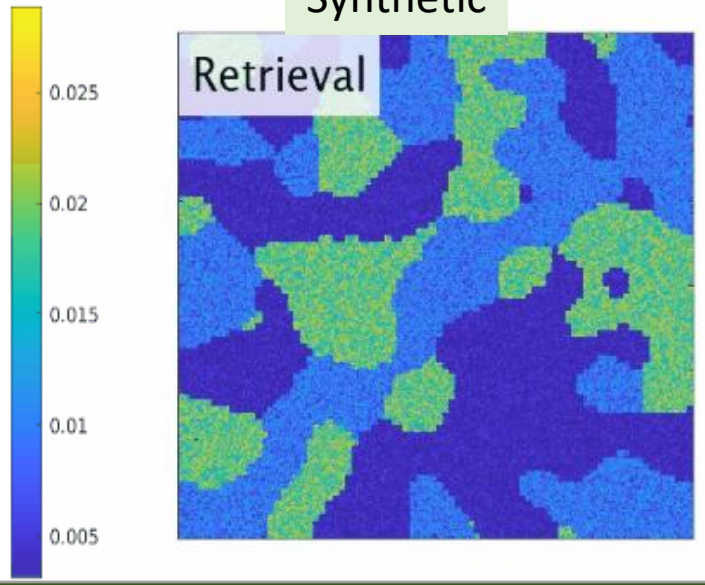
Experimental



LWC ($g\ cm^{-2}$)

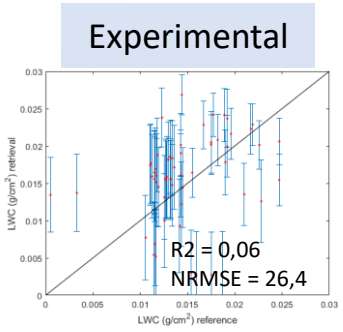
GPR

Synthetic



R^2 = 0,94
NRMSE = 10,0

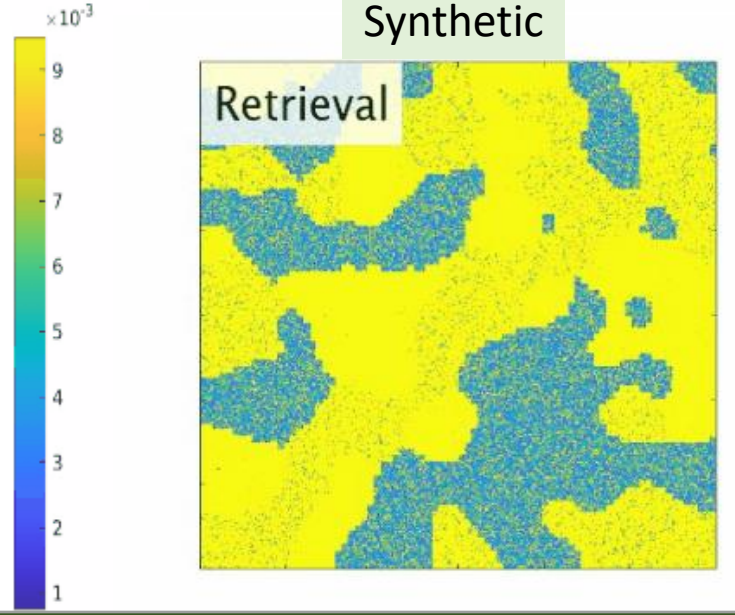
Experimental



LMA ($g\ cm^{-2}$)

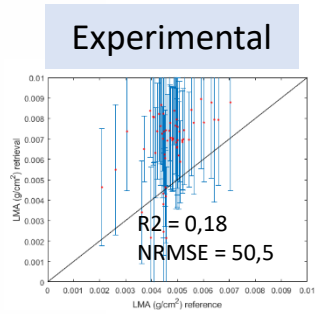
ANN

Synthetic



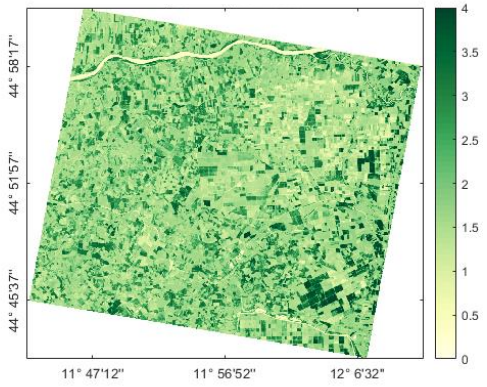
R^2 = 0,51
NRMSE = 56,0

Experimental

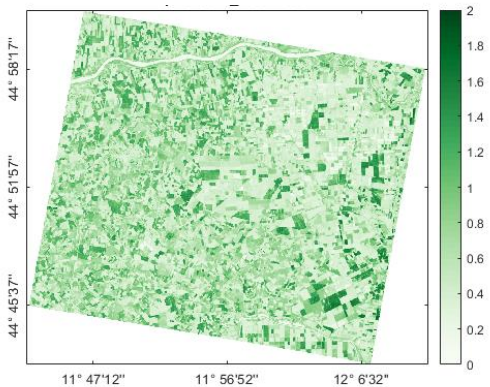


Retrievals from the L2BV: PRISMA scenes

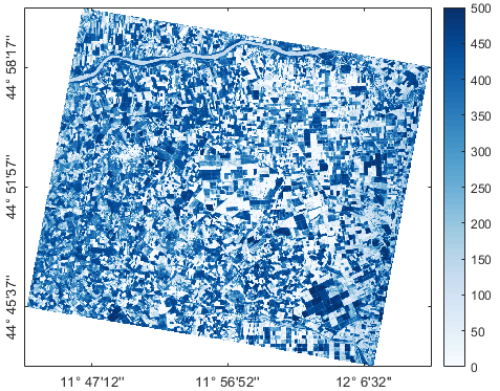
LAI
(m^2/m^2)



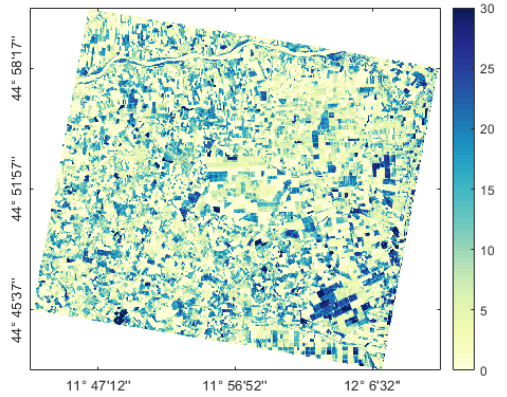
CCC
(g/m^2)



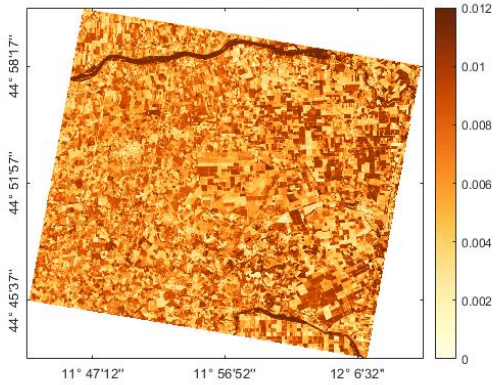
CWC
(g/m^2)



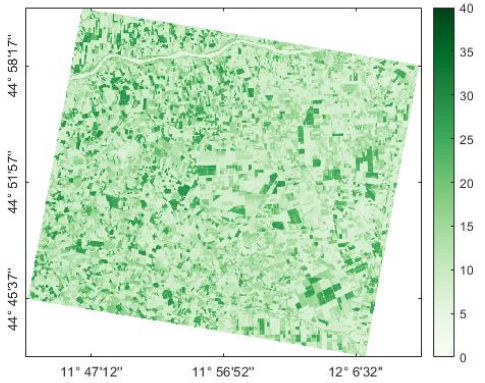
CNC
(g/m^2)



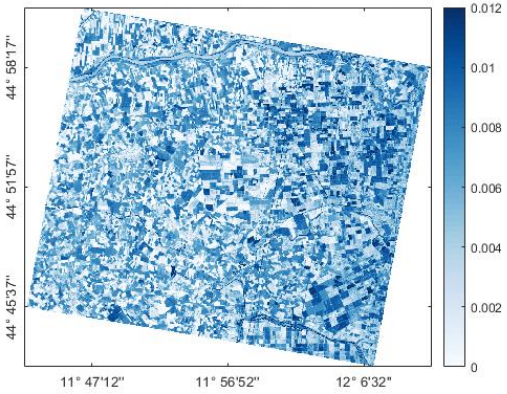
LMA
(g/cm^2)



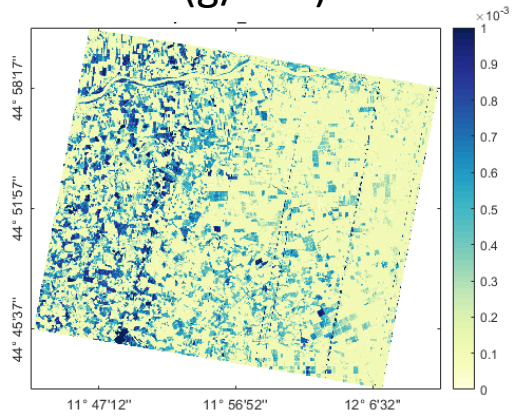
LCC
(ug/cm^2)



LWC
(g/cm^2)



LNC
(g/cm^2)



ToA: Hybrid models applied to ENMap imagery

ARTMO

ALG v3.0 Atmospheric LUT Generator

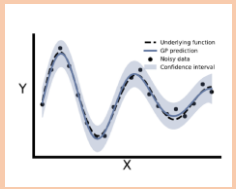
libRadtran

SCOPE
+
Active Learning
+
Bare-soil

Atmospheric synthetic data

1500 samples

GPR
20 PC



Ground Validation

Top of Atmosphere **(TOA)**

Bottom of Atmosphere **(BOA)**

Scatterplot BOA vs TOA

remote sensing
an Open Access Journal by MDPI

Gaussian Process Regression Hybrid Models for the Top-of-Atmosphere Retrieval of Vegetation Traits Applied to PRISMA and EnMAP Imagery

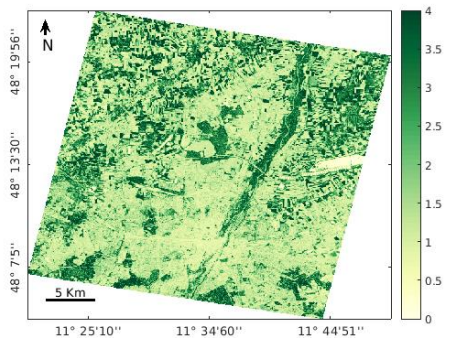
Ana B. Pascual-Venteo; Jose L. Garcia; Katja Berger; José Estévez; Jorge Vicent; Adrián Pérez-Suay; Shari Van Wittenberghe; Jochem Verrelst

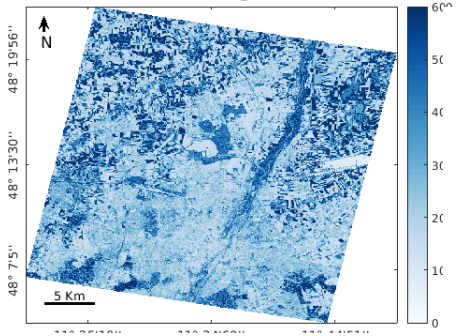
Remote Sens. **2024**, Volume 16, Issue 7, 1211

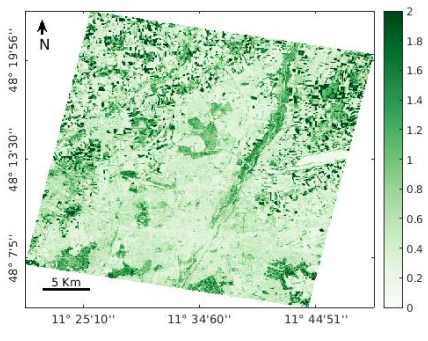
LAI (m^2/m^2)

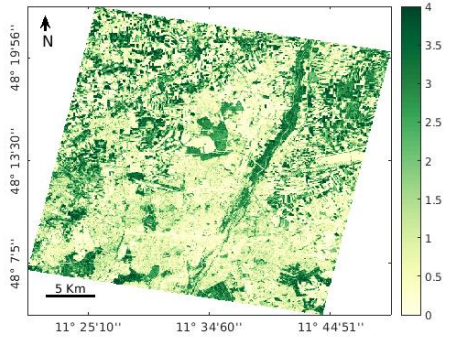
CWC (g/m^2)

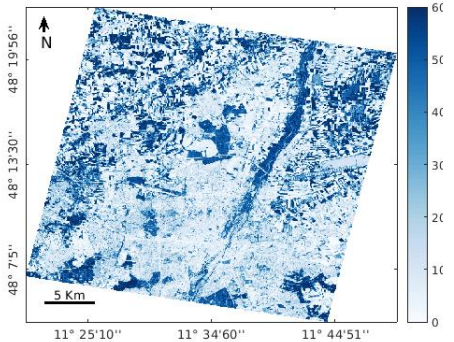
CCC (g/m^2)

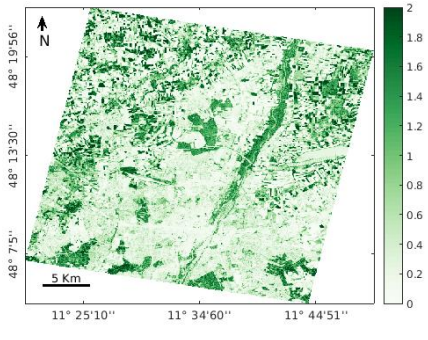


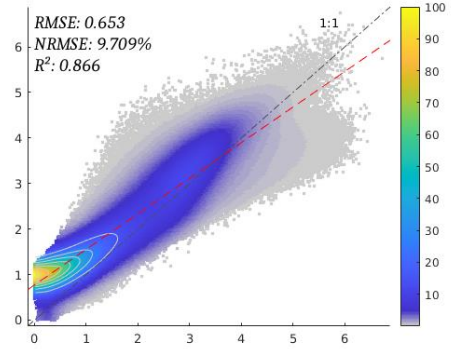


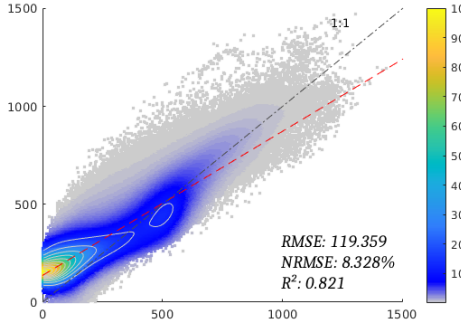


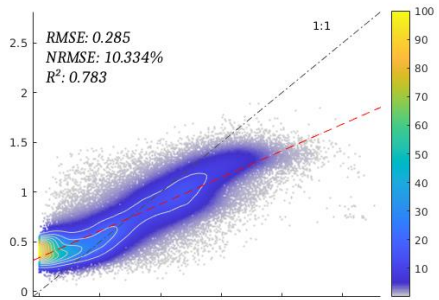








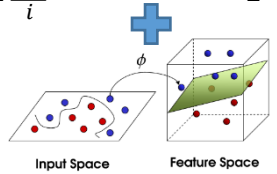


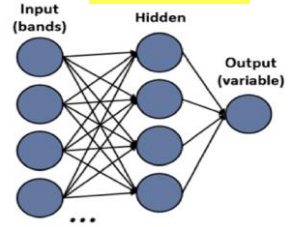


Conclusions…

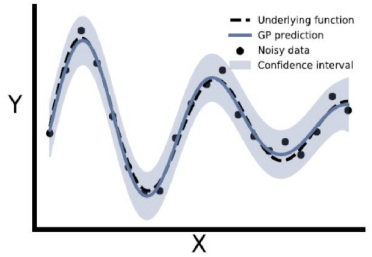
- **New MLRA** have been added to the L2BV module.

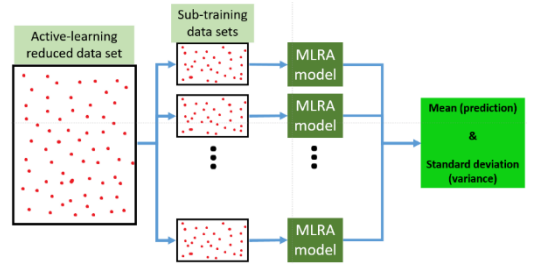
$$C = \frac{1}{2}\sum_i (y_i - w^T x_i)^2 + \frac{1}{2}\tau\|w\|^2$$



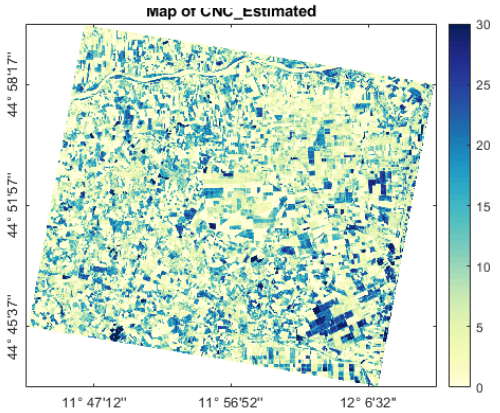


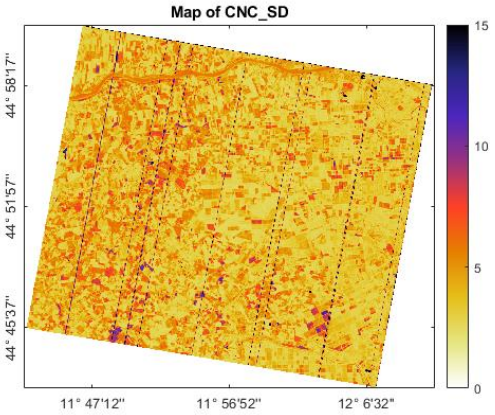
- **Uncertainty** is a requirement **being explored**.





- **Good** performance is achieved at **canopy level**.

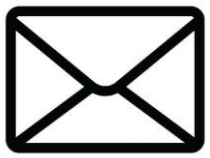




- **Leaf models** require further **research**.

Many thanks for your attention

… any questions?



Email: jose.l.garcia@uv.es

Experimental sites

Grosseto, Italy
N 42º 49.78', E 11º 4.21'

Spectral measurements:
HyPlant

Summer season

LAI: 0 – 6 (m2 m-2)

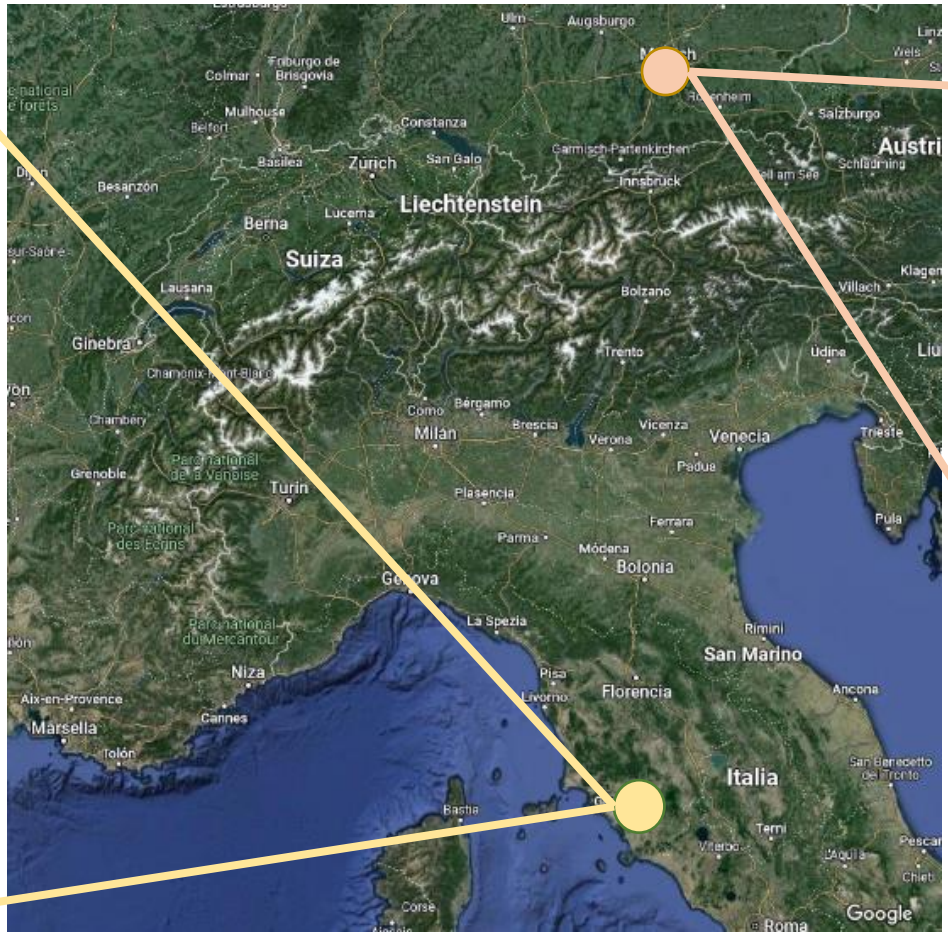
LCC: 11,6 – 60,8 (µg cm-2)

CCC: 0 – 3,2 (g m-2)

LWC: 5e-4 – 2,5e-2 (g cm-2)

CWC: 0 – 1113,6 (g m-2)

LMA: 2,1e-3 – 7,0e-3 (g cm-2)



MNI, Munich North Isar, Germany
N 48º 16', E 11º 42'

Spectral measurements:
FieldSpec4 JR

Growing season

LNC: 1,6e-4 – 6,4e-4 (g cm-2)

CNC: 1,0 – 21,5 (g m-2)

Training data set characteristics

Scope parameter configuration

Model Variables		Units	Range	Distribution
Leaf Variables				
N	Leaf structure parameter	-	1.0-2.7	Normal (μ=1.5, σ=0.5)
Cab	Leaf Chlorophyll Content	μg/cm^2	0-80	Normal (μ=45, σ=35)
Cm	Leaf Dry Matter Content	g/cm^2	0.002-0.02	Normal (μ=7.5e-3, σ=5.0e-3)
Cw	Leaf Water Content	g/cm^2	0.005-0.035	Normal (μ=15e-3, σ=8e-3)
Cp	Leaf Protein Content	g/cm^2	2.3e-5 – 4.5 e-4	Uniform
Cxc	Leaf Carotenoid Content	μg/cm^2	0 - 20	Uniform
Canopy Variables				
LAI	Leaf Area index	m^2/m^2	0.1 - 8	Uniform
LIDFa	Leaf inclination angle	rad	1 - 1	Uniform
LIDFb	Leaf inclination variation angle	rad	-1 - 1	Uniform
Geometry Variables				
SZA	Solar Zenith Angle	º	0 – 80	Uniform
OZA	Observer Zenith Angle	º	0 – 25	Uniform
RAA	Relative Azimuth Angle	º	0 - 180	Uniform
Soil Variables				
SMC	Soil Moisture Content	%	5 – 55	Normal (μ=25, σ=12.5)
BSM	Soil Brightness	-	0 – 0.9	Normal (μ=0.5, σ=0.25)
BSM_{Lat}	Soil Brightness Lat	º	20 – 40	Normal (μ=25, σ=12.5)
BSM_{Lon}	Soil Brightness Lon	º	45 - 65	Normal (μ=50, σ=10)

Reduced by Active Learning training data set

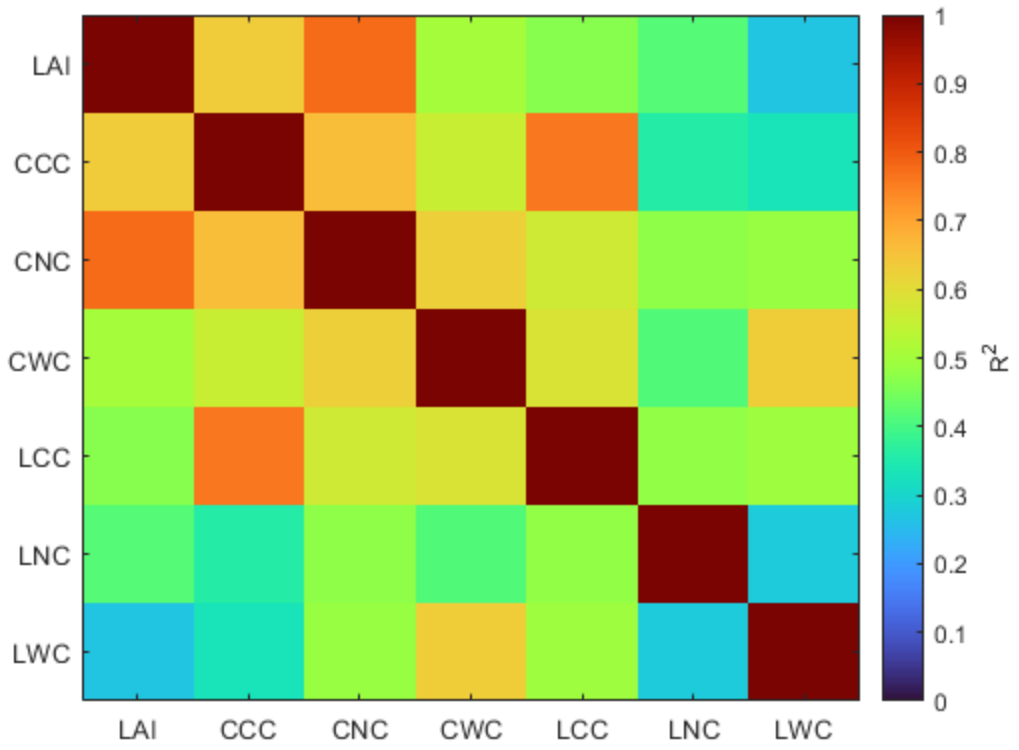
Variable	Min	Max	Units
LMA	0	0,02	g/cm2
LNC	0	5,6e-4	g/cm2
LCC	0	79,89	μg/cm2
LWC	0	0,035	g/cm2
LAI	0	7,9	M2/m2
CNC	0	35,6	g/cm2
CCC	0	6,0	g/cm2
CWC	0	2125,12	g/cm2
FVC	0	0,99	-
FAPAR	0	0,98	-

Next steps on CHIME module and models development

- **Improvement of leaf retrievals**

- Continuum removal.
- Specific regions trait affected spectral regions.
- Wavelets

- **Exploration of heterogeneus vegetated pixels (i.e., forests)**

- 3d RTM (i.e., DART)
- Aggregation/mixturing of soil spectra

Leaf variables correlations