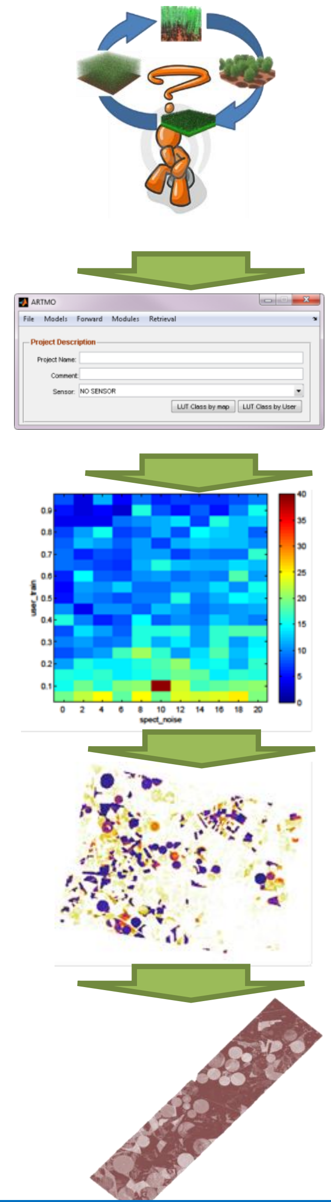
ARTMO's new Machine Learning Regression Algorithms (MLRA) module for mapping biophysical parameters

Jochem Verrelst, Juan Pablo Rivera, Jordi Muñoz-Mari, Jose Moreno, Gustavo Camps-Valls

Outlook

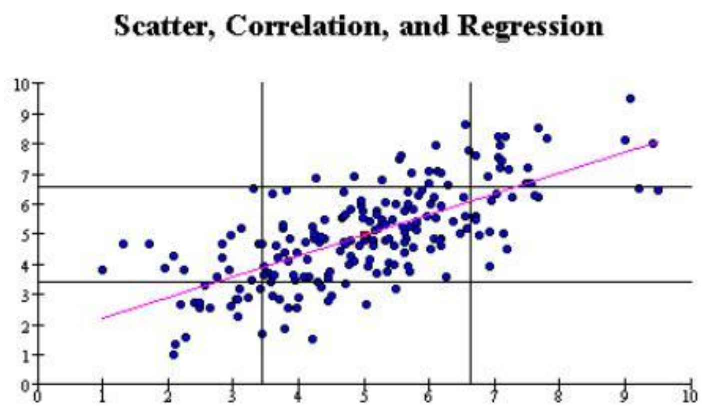


- Background
  - Biophysical parameter retrieval
  - Nonparametric regression for retrieval of biophysical parameters
  - ARTMO
- MLRA toolbox
  - MLRAs
  - MLRA settings
  - Results tests
  - Retrievals
  - Multi-output
  - Coupling with RTMs
- Conclusions

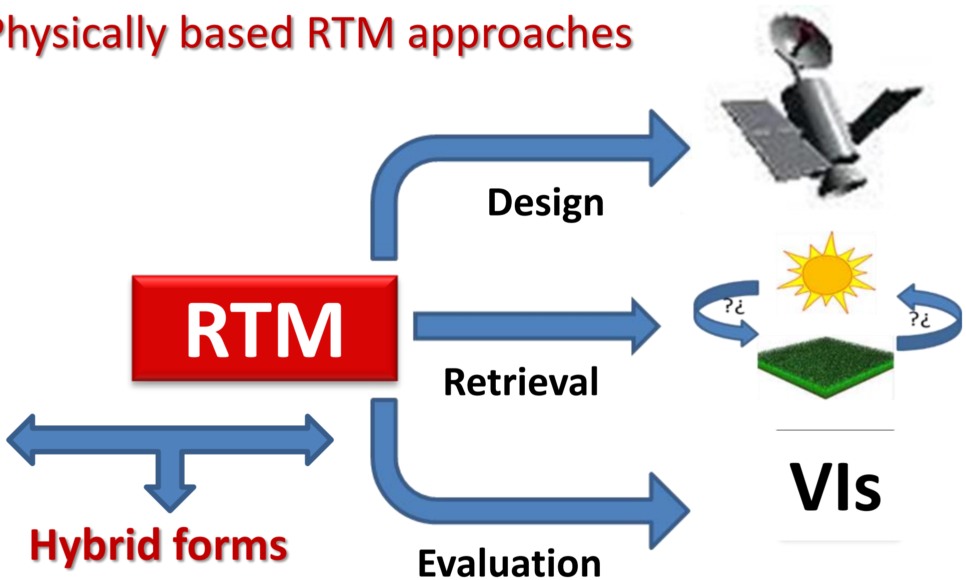
Basics biophysical parameter retrieval

Retrieval of biophysical parameters from optical EO data **always occurs through a model**; e.g. through **statistical models**, through **inversion** of physically-based **radiative transfer models** (RTM), or through **hybrid forms**.

Statistical approaches



Physically based RTM approaches



Parametric – Nonparametric – Physically-based inversion

- **Parametric regression:** Some constraints introduced
- **Nonparametric regression**: No constraints in developing models
- **Physically-based approaches**: Inversion of RTMs using parametric or non-parametric inversion techniques (i.e., ***hybrid forms***).

Nonparametric regression is a form of regression analysis in which the predictor **does not take a predetermined form** but is constructed according to information derived from the data. Nonparametric regression requires larger sample sizes than regression based on parametric models because the data must supply the model structure as well as the model estimates.

Nonparametric regression based on machine learning algorithms:

- Linear transformations: PCR, PLS
- Non-linear transformations: NN, GPR, KRR, SVR.

Nonparametric regression

Pros ☺

- Powerful
- Relatively fast
- Additional features
- Multi-output

Cons ☹

- Portability?
- Black-box (for the majority)
- Difficult to use

Operational retrieval of biophysical parameters

- **Neural networks coupled with RTMs** widely used but face limitations: black box, unstable, difficult and slow in training.

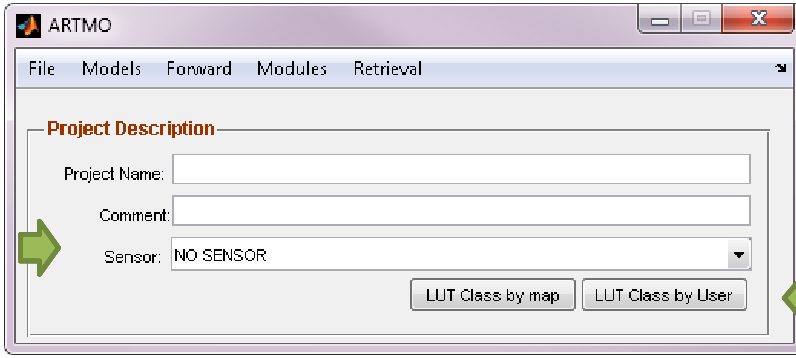
  ✓ MERIS
  ✓ SPOT-VGT
  ✓ Sentinel-2,-3

- The ability of NN to process hyperspectral data?

- Towards new generation of MLRA regressors for operational use.

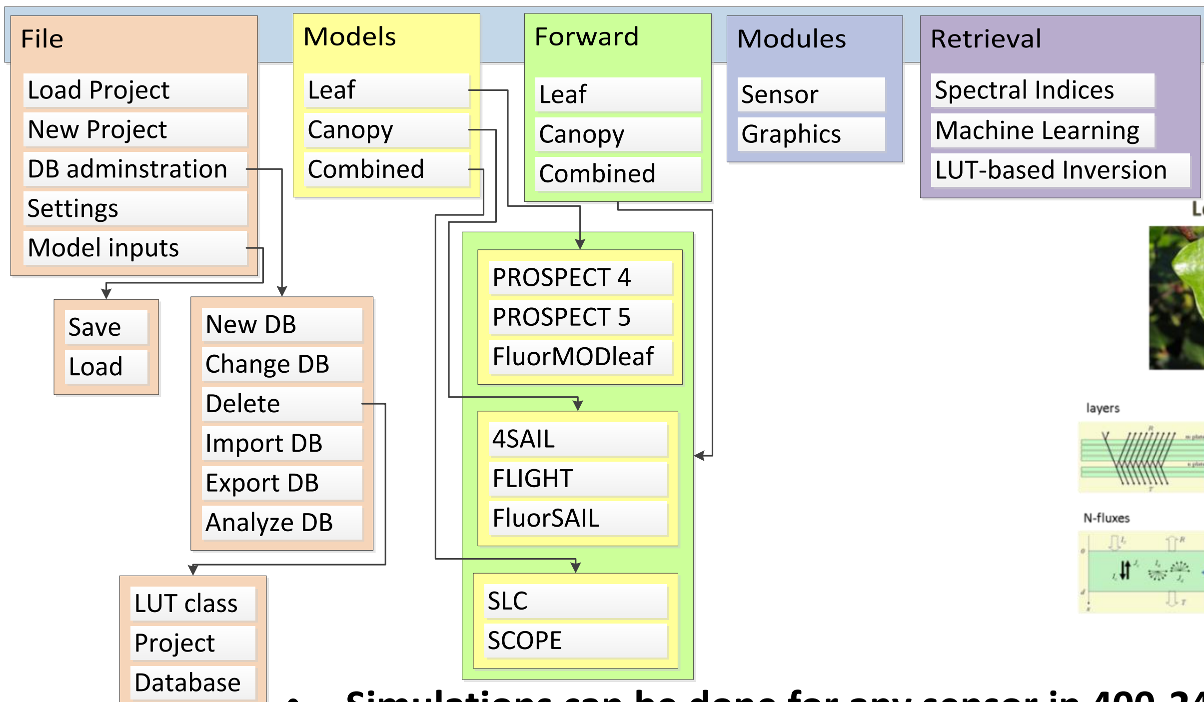
V3: Modular design

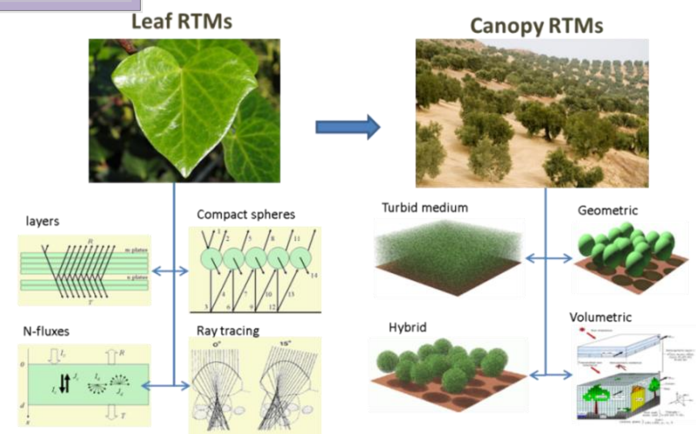
Simulations according to a predefined sensor setting



All models and modules can be accessed from the Menu bar

LUTs can be configured per land cover class or defined by user.





- **Simulations can be done for any sensor in 400-2400 nm range.**
- **Input, output and metadata stored in MySQL running underneath.**

MLRA toolbox

- **Load Image & Class**
  - Load Image
  - Load Class
- **Input**
  - RTM data
    - Select project
    - Edit settings
  - User data
- **MLRA setting**
  - Single-output
  - Multi-output
- **Test MLRAs**
  - Load test
  - New test
- **Retrieval**
- **Tools**
  - Save
  - Load
  - Test database
    - Rename
    - Delete
  - View maps
  - Settings

When loading a land cover map then retrieval strategies can be optimized per class.

Input data can come either from RTMs, from field observations or from both.

In '**MLRA setting**' multiple MLRA retrieval strategies can configured, either single-output or multi-output

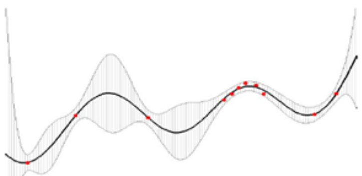
In **Test MLRAs,** run tests with the predefined strategies or load an existing test

Manually set up a **MLRA retrieval strategy** or select an earlier evaluated strategy.

'Tools' offer various options to manage the MLRA Module.

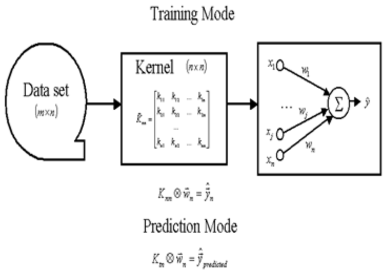
Implemented MLRAs

Gaussian Processes Regression (GPR)



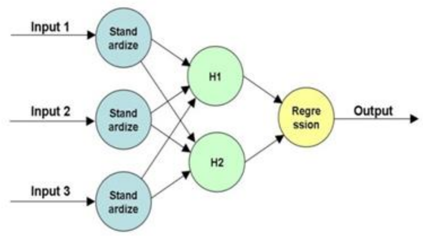
- ☺ Robust regressor
- ☺ Transparent: provides insight in relevant bands and samples
- ☺ Provides addtional confidences
- ☹ Difficulty with many training samples, e.g. > 2000

Kernel ridge regression (KRR)



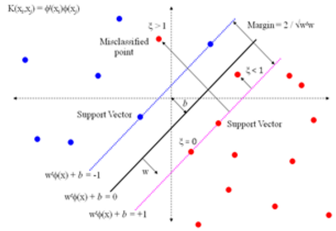
- ☺ Robust regressor
- ☺ Rather fast
- ☹ Prone to outliers

Artificial Neural Networks (NN)



- ☺ Robust regressor
- ☺ Ability to detect complex nonlinear relationships
- ☺ Once trained, fast in applying to images
- ☹ Lack of transparency
- ☹ prone to overfitting
- ☹ Computational demanding in the training phase

Support vector regression (SVR)



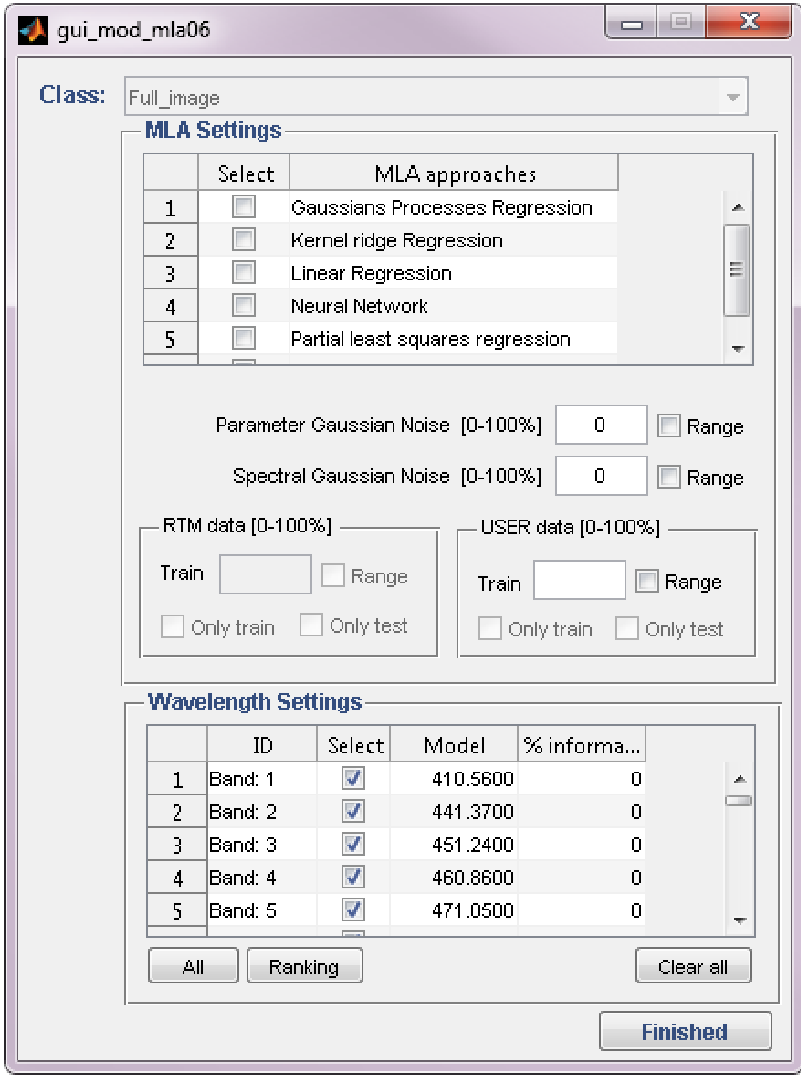
- ☺ Robust regressor
- ☺ Robust to outliers
- ☺ Provides some information through support vectors
- ☹ Computational demanding

Linear nonparametric regressors:

- Linear regression (LR)
- Partial least squares regression (PLS)

Alternative regressors are planned to be added: PCR, LASSO

MLRA settings



If active, configure per land cover class.

Select a MLRA

Options to add noise

Option to mix RTM with field data

Option to select bands (manually or automatically through mutual information: band with most information first)

Data:

SPARC campaign, Barrax, Spain

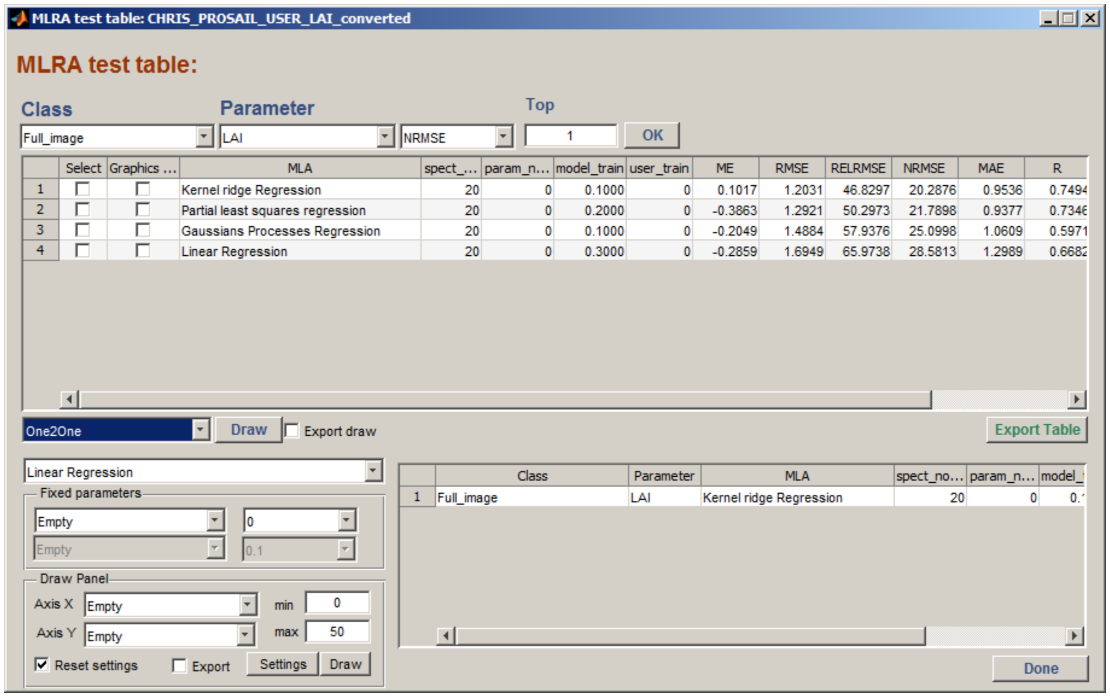
Field data:

- **LCC** measured with **CCM-200**
- **LAI** measured with LiCor LAI-2000

Spectral data:

- **CHRIS mode 1** (62 bands; 34m) nadir spectra
- **HyMap** (5 m resolution;
- 125 bands ; 450-2500 nm

Results test



	Select	Graphics ...	MLA	spect_...	param_n...	model_train	user_train	ME	RMSE	RELRMSE	NRMSE	MAE	R
1			Kernel ridge Regression	20	0	0.1000	0	0.1017	1.2031	46.8297	20.2876	0.9536	0.7494
2			Partial least squares regression	20	0	0.2000	0	-0.3863	1.2921	50.2973	21.7898	0.9377	0.7346
3			Gaussians Processes Regression	20	0	0.1000	0	-0.2049	1.4884	57.9376	25.0998	1.0609	0.5971
4			Linear Regression	20	0	0.3000	0	-0.2859	1.6949	65.9738	28.5813	1.2989	0.6682

	Class	Parameter	MLA	spect_no...	param_n...	model_
1	Full_image	LAI	Kernel ridge Regression	20	0	0.

Results can be organized according to land cover class, parameter, cal/val, and statistical output

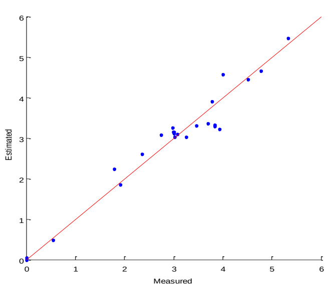
Overview of results. Here, best results per SI andcurve fitting

Options to plot all kinds of output and export results

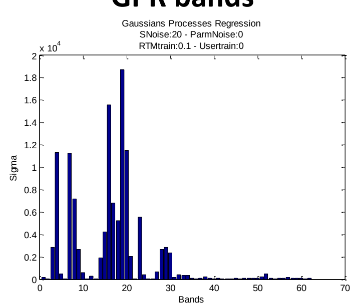
Selected strategies appear here and can be transported to the Retrieval module.

output

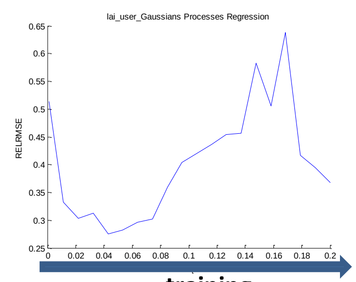
1:1



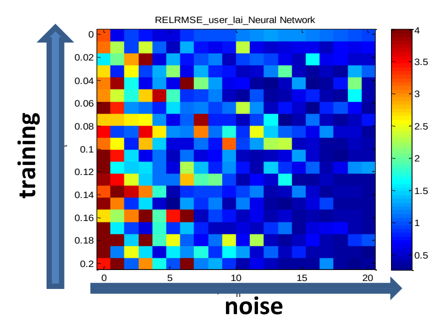
GPR bands



1D results



2D results

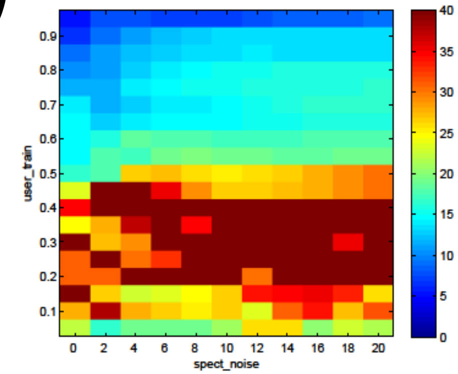


Case study SPARC-CHRIS

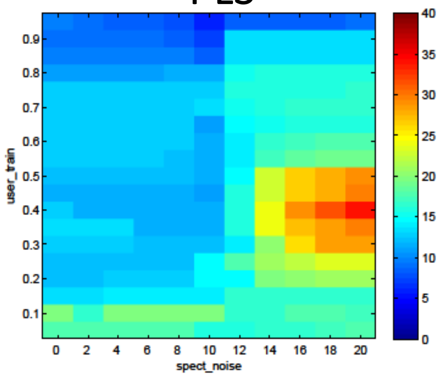
LCC – User data (SPARC - CHRIS)

MLRA	Spectral noise [%]	training [%]	RMSE	NRMSE [%]	R2
Kernel ridge Regression	0	95	0.97	1.89	0.998
Gaussians Processes Regression	0	90	1.03	2.02	0.997
Neural Network	6	90	1.50	2.95	0.995
Linear Regression	0	95	2.71	5.31	0.988
Partial least squares regression	10	95	2.90	5.69	0.991

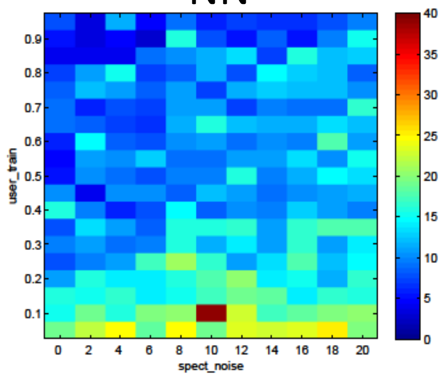
LR



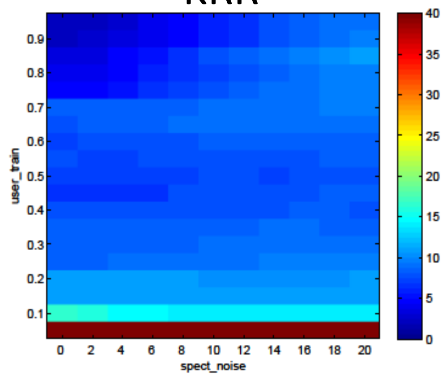
PLS



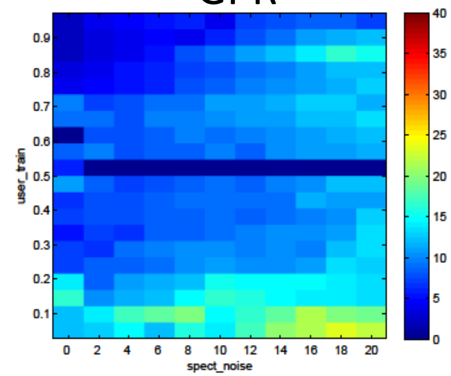
NN



KRR



GPR



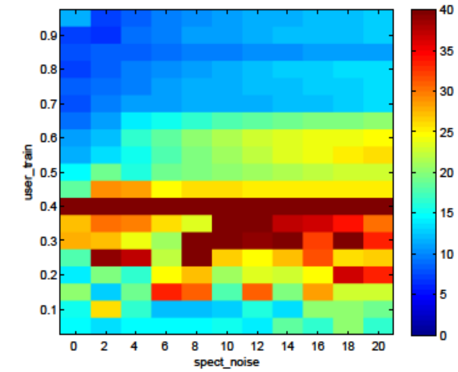
- **LR poor**, only suitable with high training.
- **PLS somewhat better**, but not excellent performances. Needs noise.
- **NN behaves erratic**: can lead to good performances but unstable.
- **KKR: Excellent performances**, very robust.
- **GPR: Excellent performances**, robust.

SI 3-bands LR:
R2: 0.91
NRMSE: 8,25

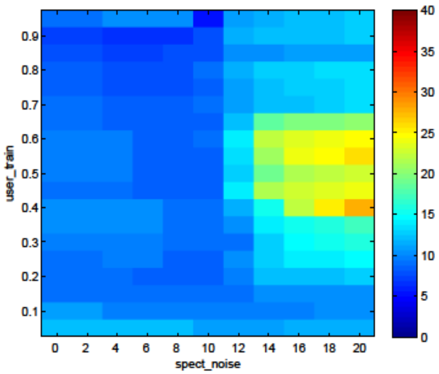
LAI – User data (SPARC - CHRIS)

MLRA	Spectral noise [%]	training [%]	RMSE	NRMSE [%]	R2
Kernel ridge Regression	0	90	0.15	2.75	0.99
Neural Network	0	90	0.19	3.42	0.99
Gaussians Processes Regression	0	90	0.22	3.98	0.99
Partial least squares regression	10	95	0.21	5.55	0.99
Linear Regression	2	90	0.36	6.65	0.96

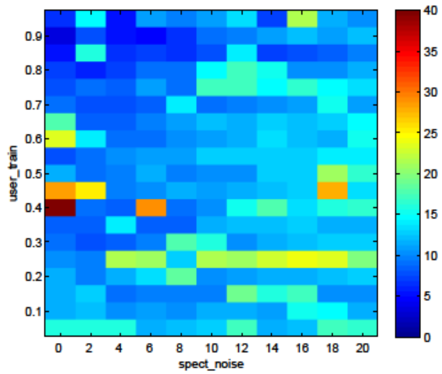
LR



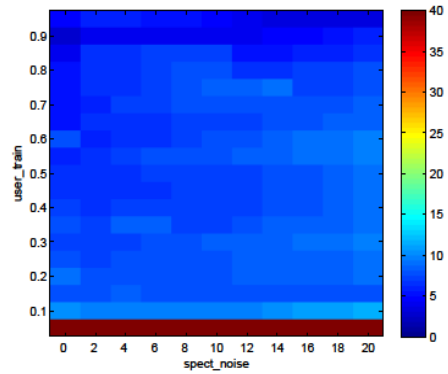
PLS



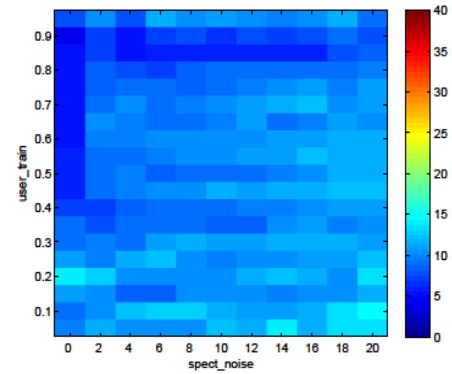
NN



KRR



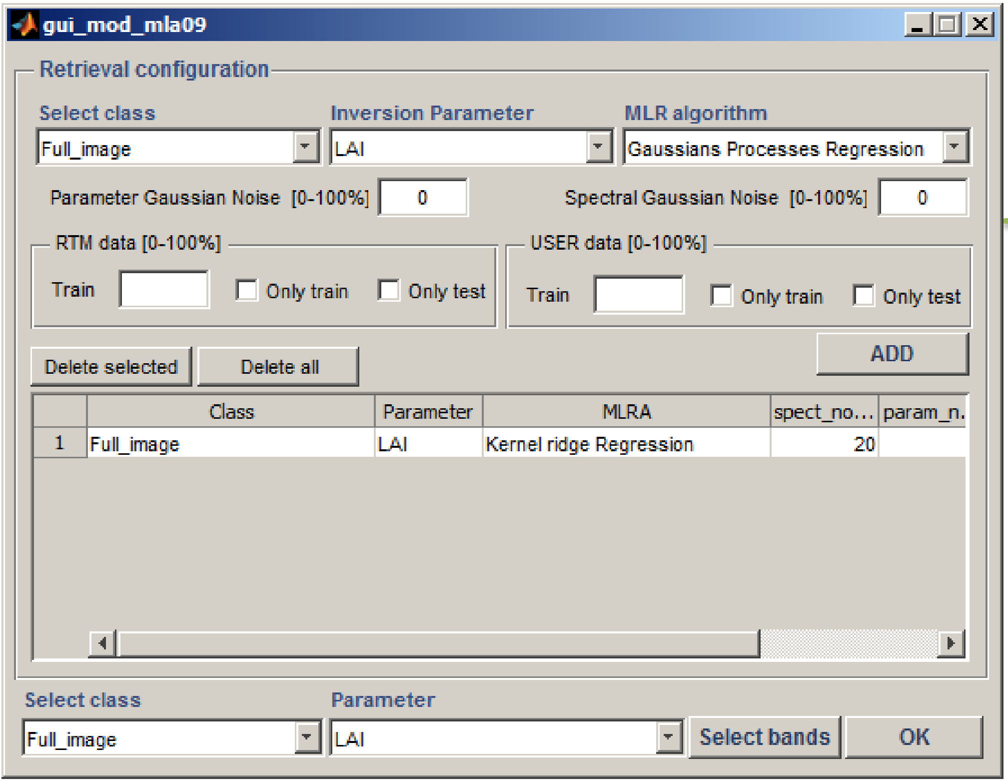
GPR



- **LR poor**, only acceptable with high training.
- **PLS somewhat better**, but not excellent performances. Needs noise.
- **NN behaves erratic**: can lead to good performances but unstable.
- **KKR: Excellent performances**, very robust.
- **GPR: Excellent performances**, robust.

SI 3-bands LR:
R2: 0.91
NRMSE: 6,73

Retrieval



Manual options

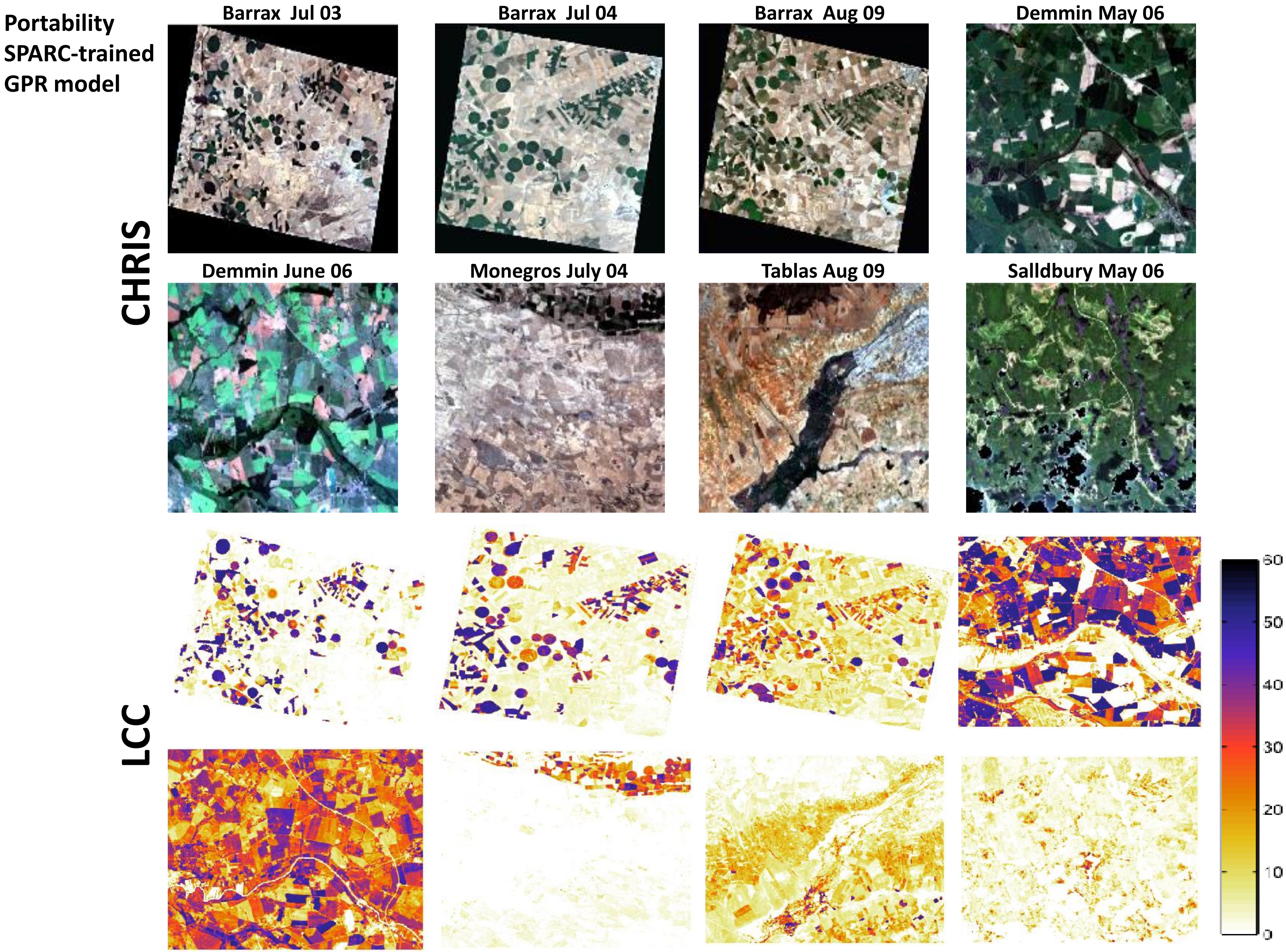
- Options to select land cover class, parameter and algorithm.
- Options to add noise, select user/RTM and train/test data distribution.

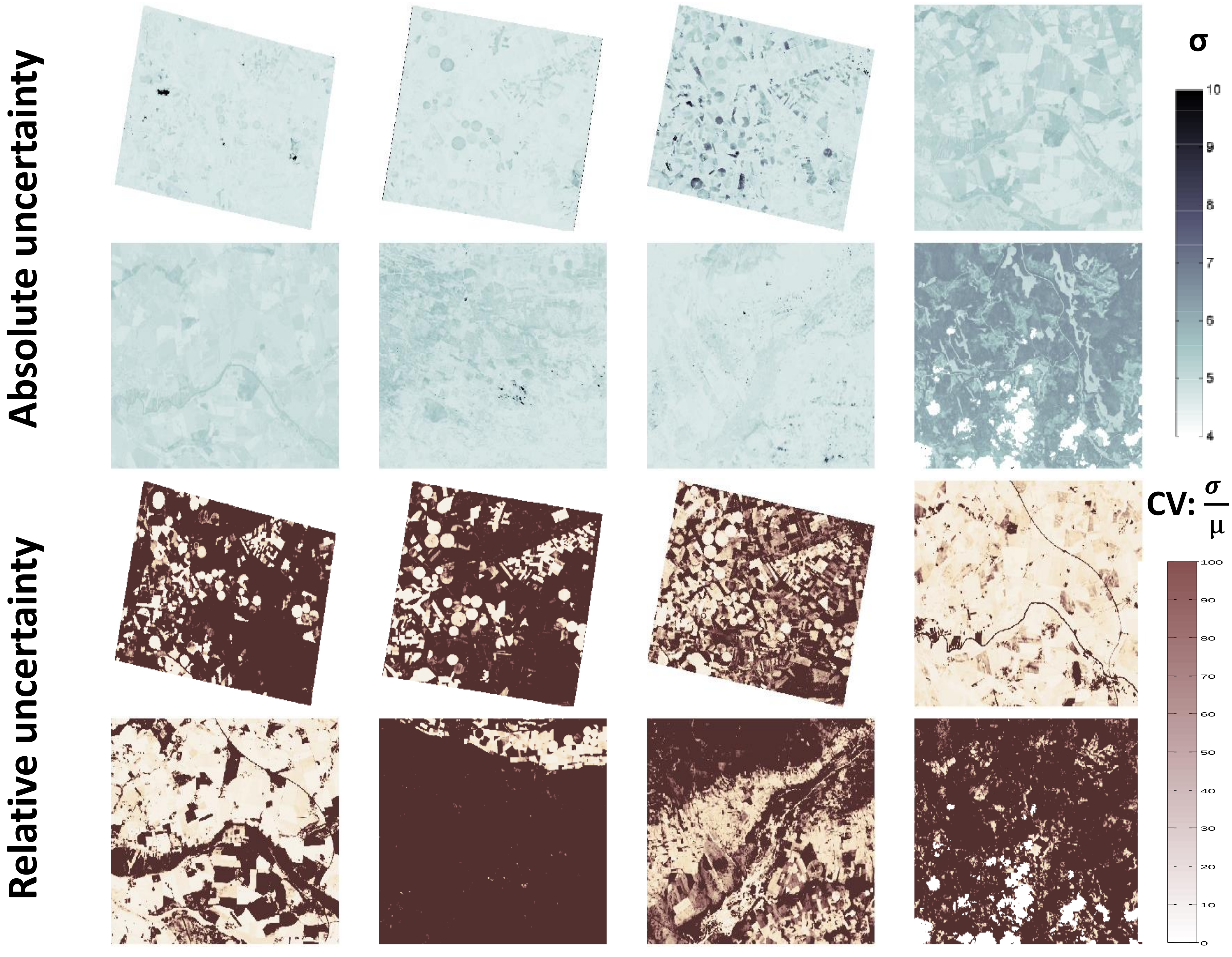
Selected strategies, from above or imported from earlier test.

Plotting options

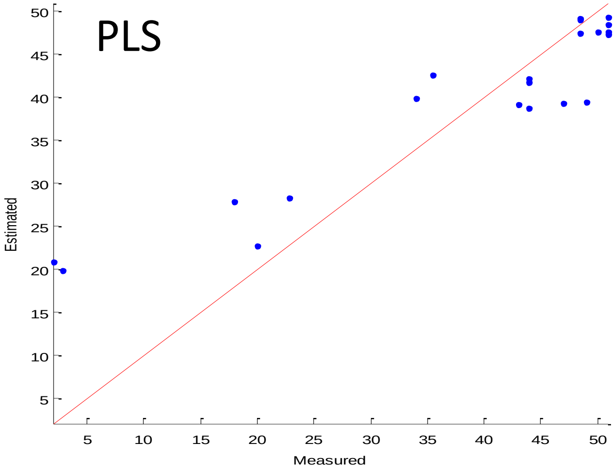
Case studies:

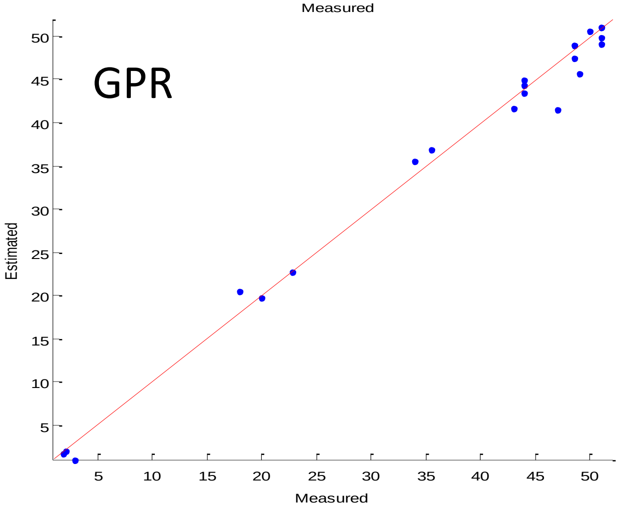
- Applying GPR to images because of additional features - LCC
- The same SPARC-trained model has been applied to various CHRIS images.
- Also HyMap data was processed


Portability
SPARC-trained
GPR model
CHRIS
LCC
Barrax Jul 03
Barrax Jul 04
Barrax Aug 09
Demmin May 06
Demmin June 06
Monegros July 04
Tablas Aug 09
Salldbury May 06
60
50
40
30
20
10
0

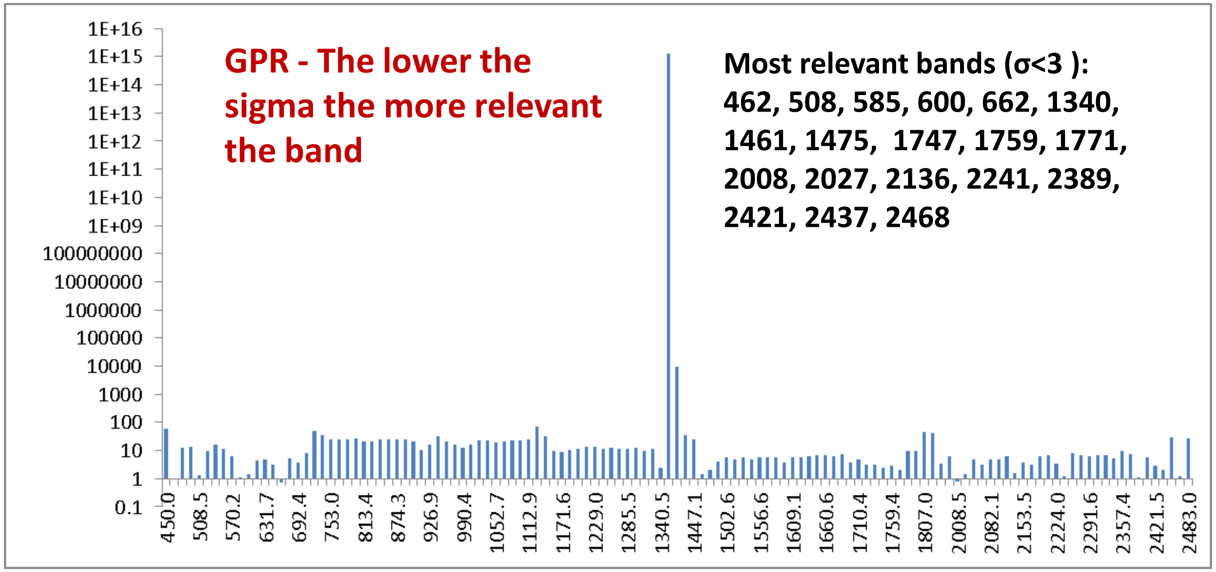

Absolute uncertainty
Relative uncertainty
σ
10
9
8
7
6
5
4
CV: σ/μ
100
90
80
70
60
50
40
30
20
10
0

SPARC – HyMap - LCC





Regressor	Training [%]	RMSE	NRMSE	R2
Gaussians Processes Regression	80	1.73	3.51	0.99
Kernel ridge Regression	80	2.31	4.71	0.99
Neural Network	80	2.71	5.51	0.98
Partial least squares regression	80	8.16	16.61	0.90
Linear Regression	80	23.21	47.23	0.05



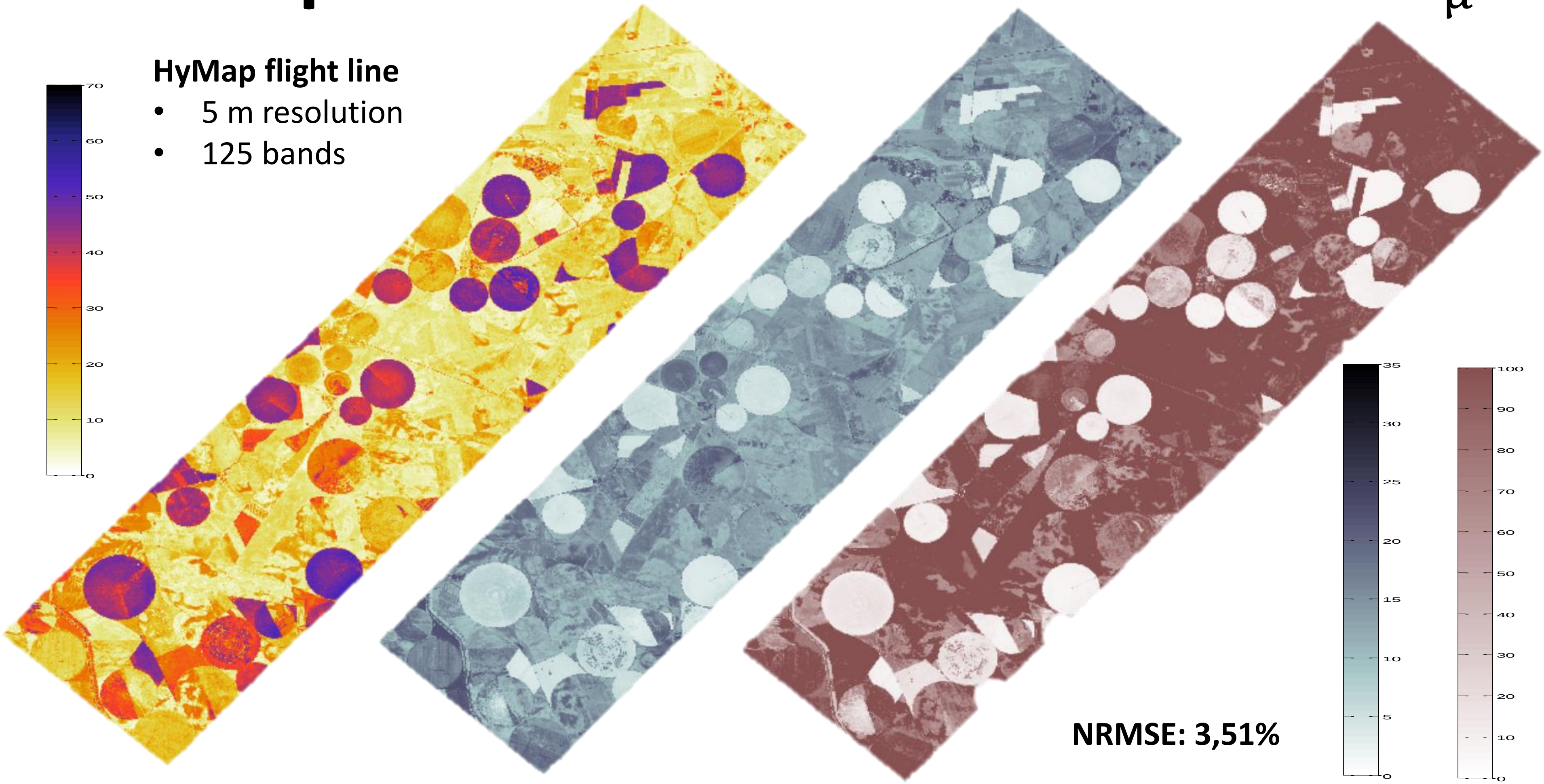
Note that PLS performs considerably poorer. GPR insight in relevant bands

LCC map

μ σ CV: $\frac{\sigma}{\mu}$

HyMap flight line

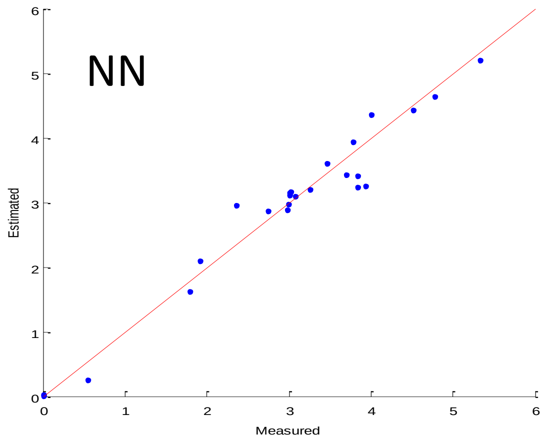
- 5 m resolution
- 125 bands



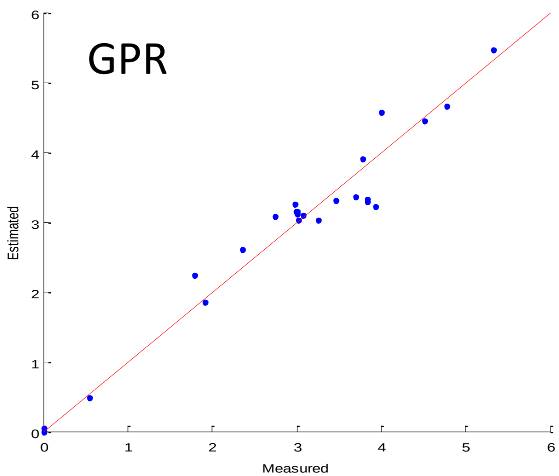
NRMSE: 3,51%

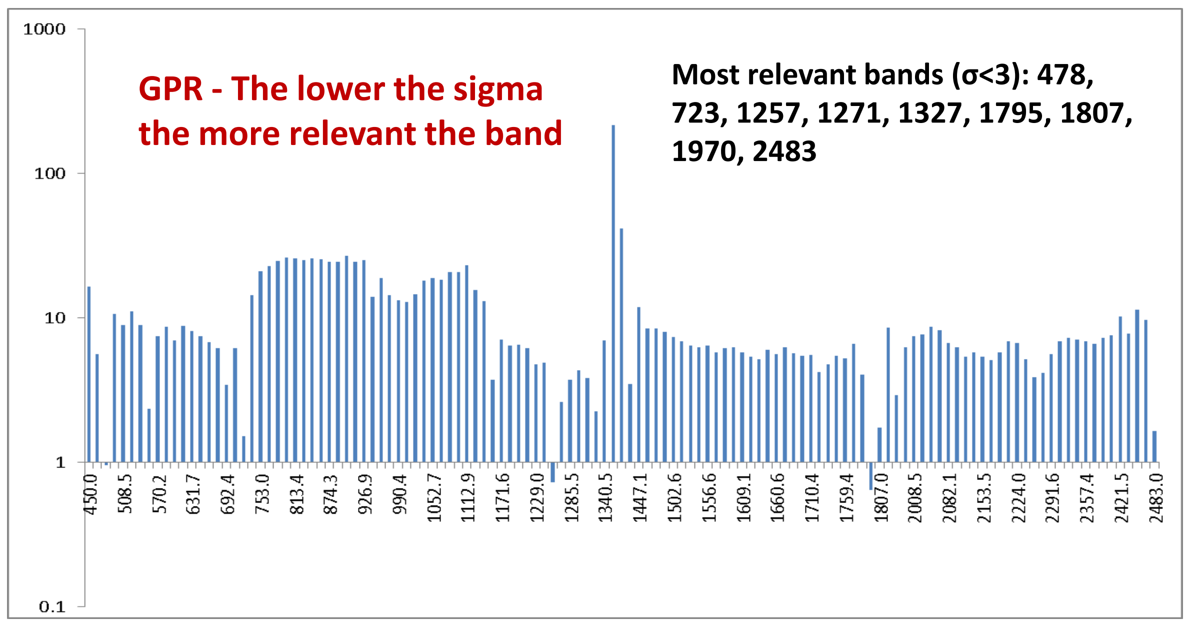
GPR mean estimates, uncertainties and relative uncertainties

SPARC – HyMap - LAI



Regressor	Training [%]	RMSE	NRMSE	R2
Neural Network	80	0.28	5.24	0.96
Gaussians Processes Regression	80	0.30	5.66	0.95
Kernel ridge Regression	80	0.37	6.95	0.93
Partial least squares regression	80	0.67	12.59	0.80
Linear Regression	80	2.72	51.10	0.19





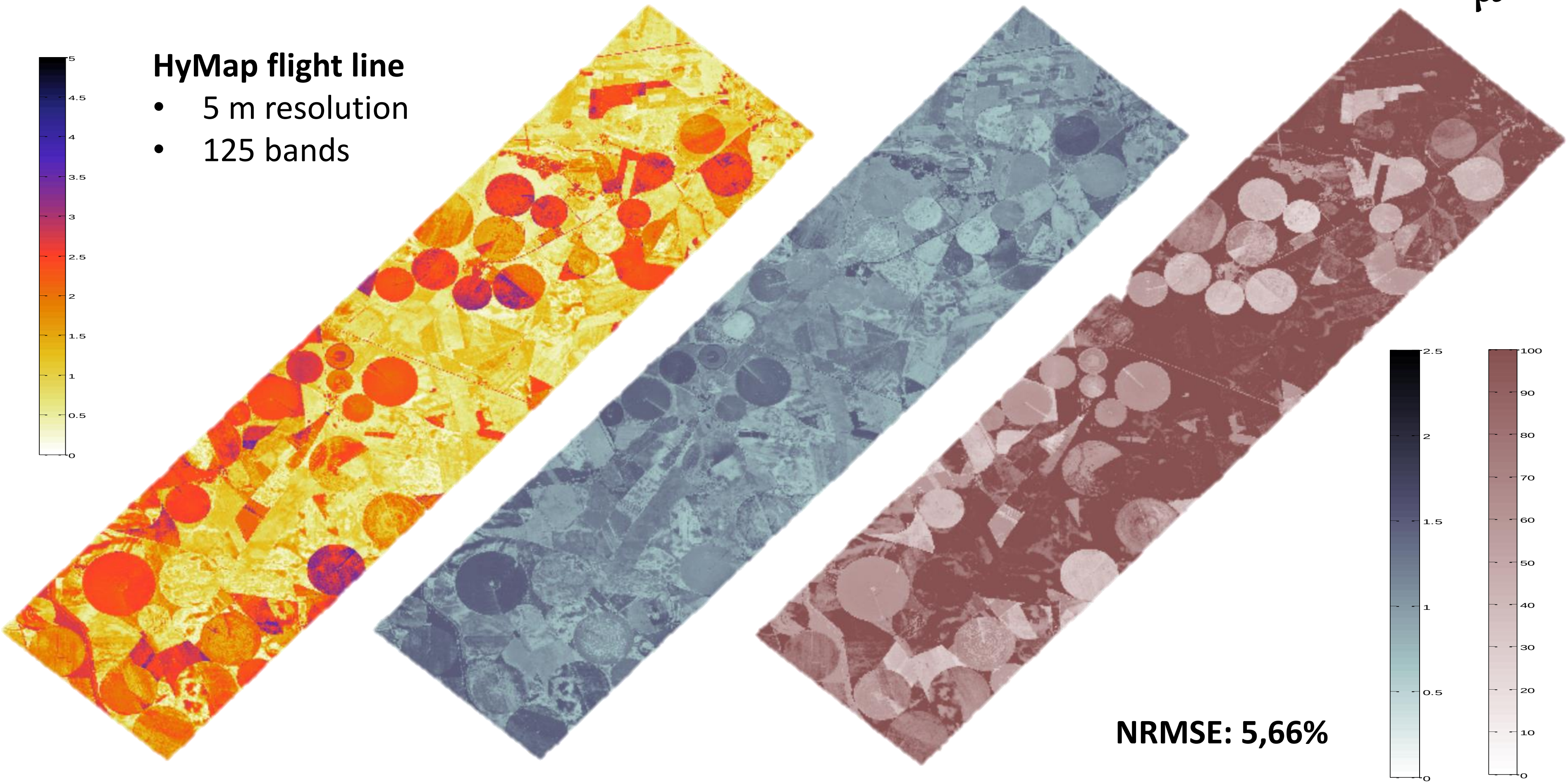
LAI map

μ

σ

CV: $\frac{\sigma}{\mu}$

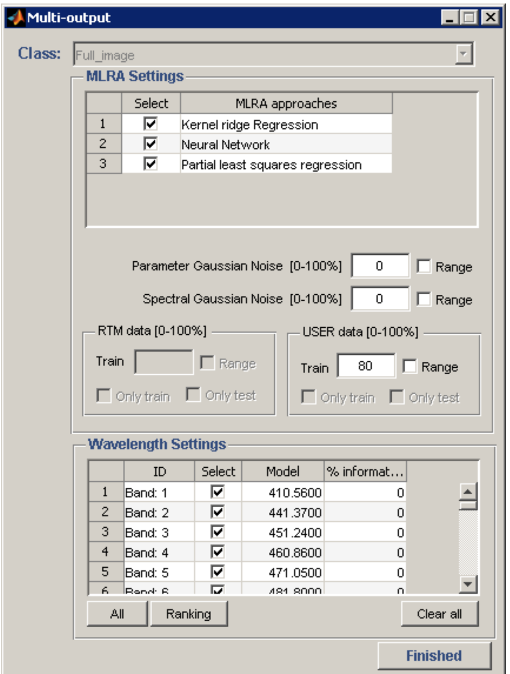
HyMap flight line
- 5 m resolution
- 125 bands



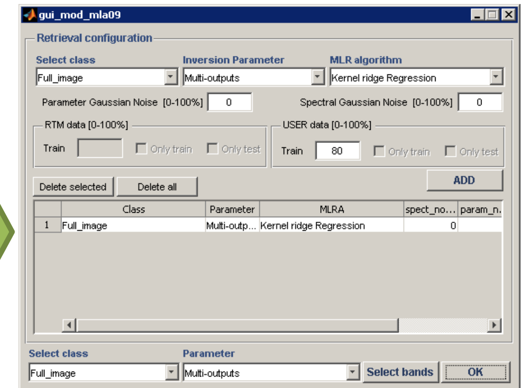
NRMSE: 5,66%

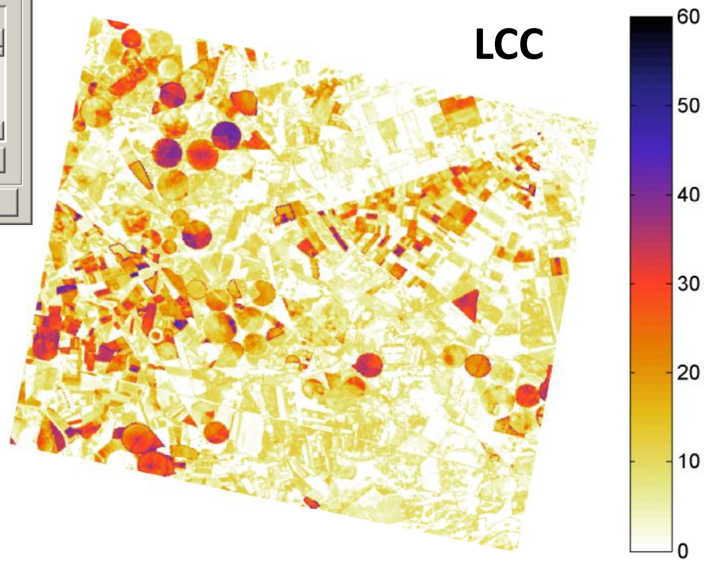
Uncertainties poorer; both over vegetated and non-vegetated surfaces.

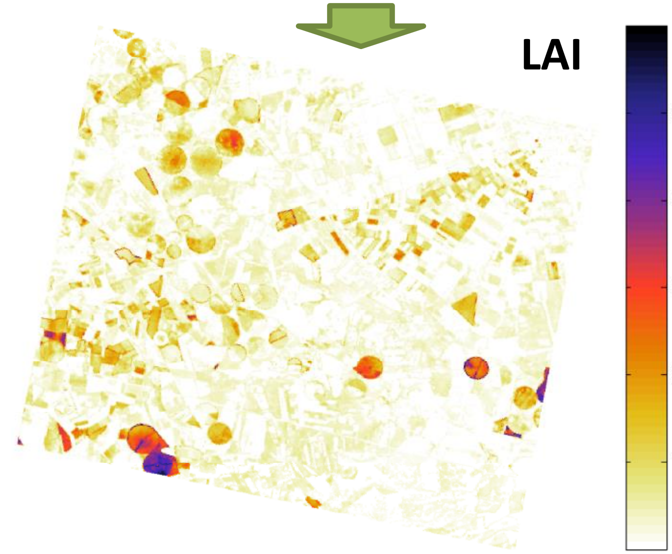
Multi-output CHRIS-SPARC



MLRA	LCC RMSE	LCC NRMSE	LCC R2	LAI RMSE	LAI NRMSE	LAI R2
KRR	4.06	7.96	0.96	0.36	6.43	0.96
NN	3.93	7.70	0.96	0.51	9.06	0.92
PLS	8.26	16.21	0.83	0.60	10.59	0.90







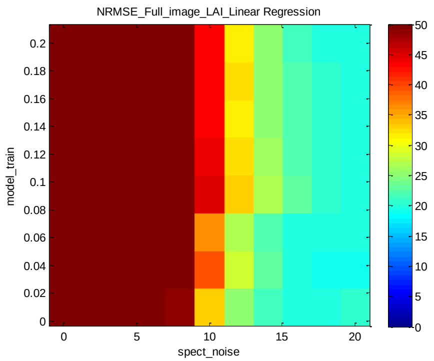
Faster but not necessarily better.

Coupling RTM (PROSAIL) with MLRAs

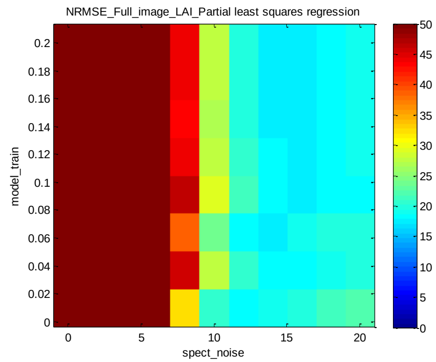
Training: 10000 random simulations; Validation: SPARC dataset (without bare soil)

Sentinel-2 (10m)

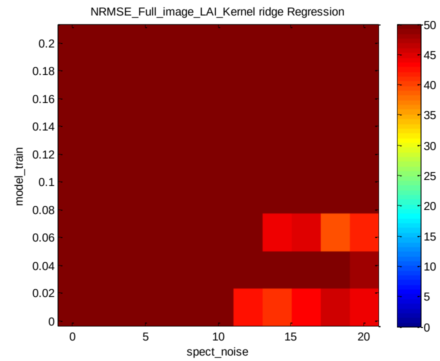
LR



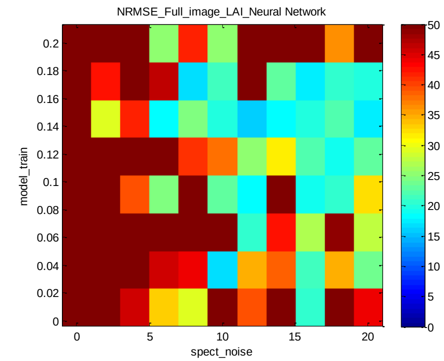
PLS



KRR



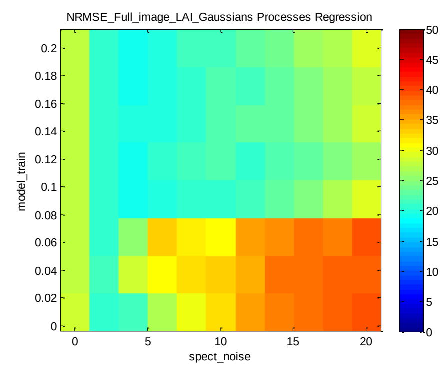
NN



Coupling so far unsuccessful:

- Difficulties to deal with large training samples
- Poor matching: noise needed to enable matching with real data
- Results poor: better configurations needed
- NN rather unstable
- GPR best performing

GPR



Regressor	Training [%]	Noise [%]	RMSE	NRMSE	R2
Neural Network	12	16	0.89	16.24	0.78
Partial least squares regression	16	10	0.97	17.62	0.76
Gaussians Processes Regression	4	13	1.03	18.76	0.62
Linear Regression	20	4	1.06	19.37	0.70
Kernel ridge Regression	18	7	2.15	39.15	0.16

Conclusions

- **Nonparametric regressors powerful retrieval algorithms**. They easily **outperform parametric regressors** (e.g. VI-based).
- PLS not most powerful. **MLRA such as NN, KRR and GPR were best evaluated.**
- **GPR** a Bayesian regressor; **insight in relevant bands and provides uncertainties**.
- **MLRA toolbox** developed in **ARTMO** that guides the user through all necessary processing steps.
- Coupling RTMs with MLRAs possible, but further efforts needed to make it successful.

Thanks

Availability

ARTMO is work in progress - beta version

- Accessible at Valencia University under our supervision.
- Matlab programmers are encouraged to write their own apps. In turn, a copy can be given.
  - Atmospheric models
  - BRDF apps
  - Temporal domain
  - classifiers
- Public available after publication (will take some time – so far unsuccessful)