Juhan Ross Legacy Symposium

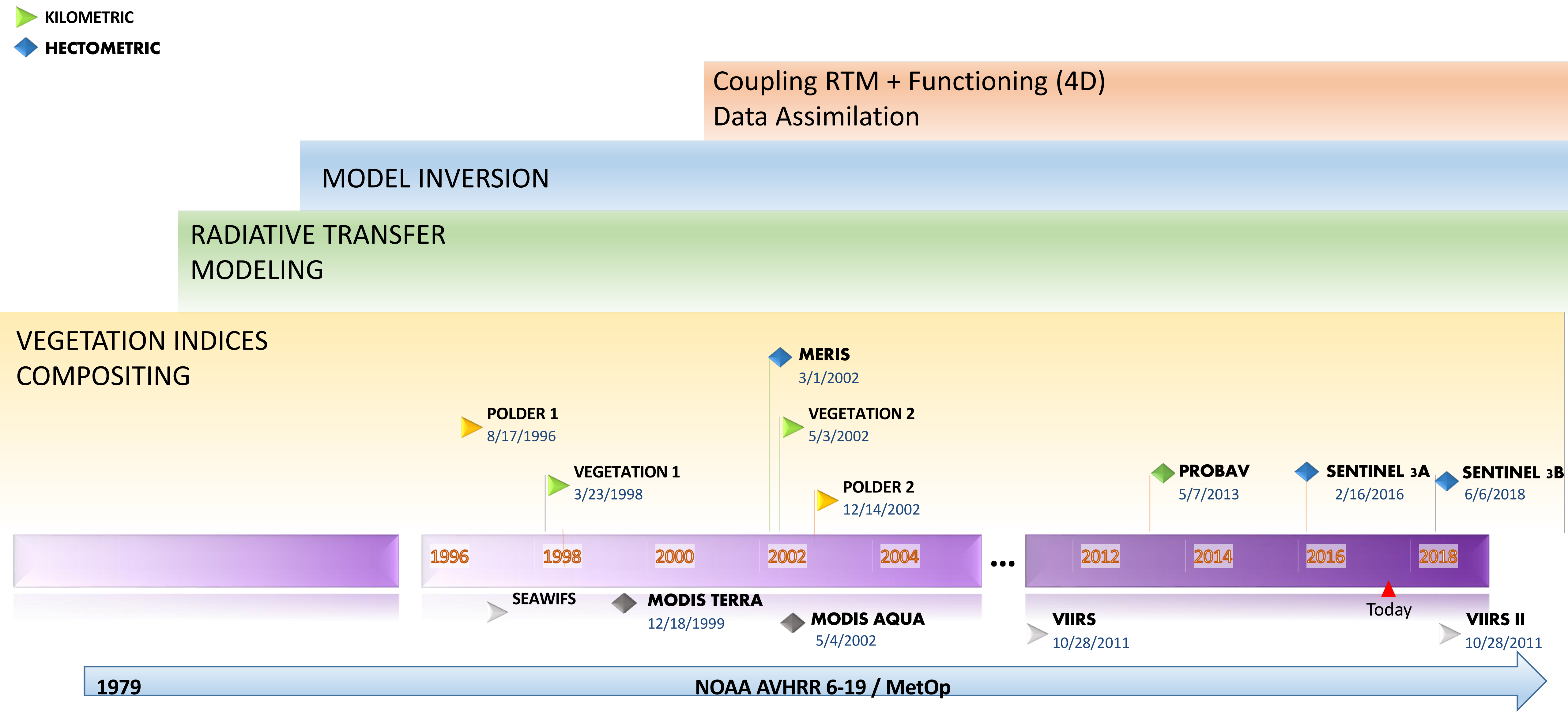
24-25 August 2017

20 years of algorithms for the derivation of global vegetation products from European medium resolution sensors

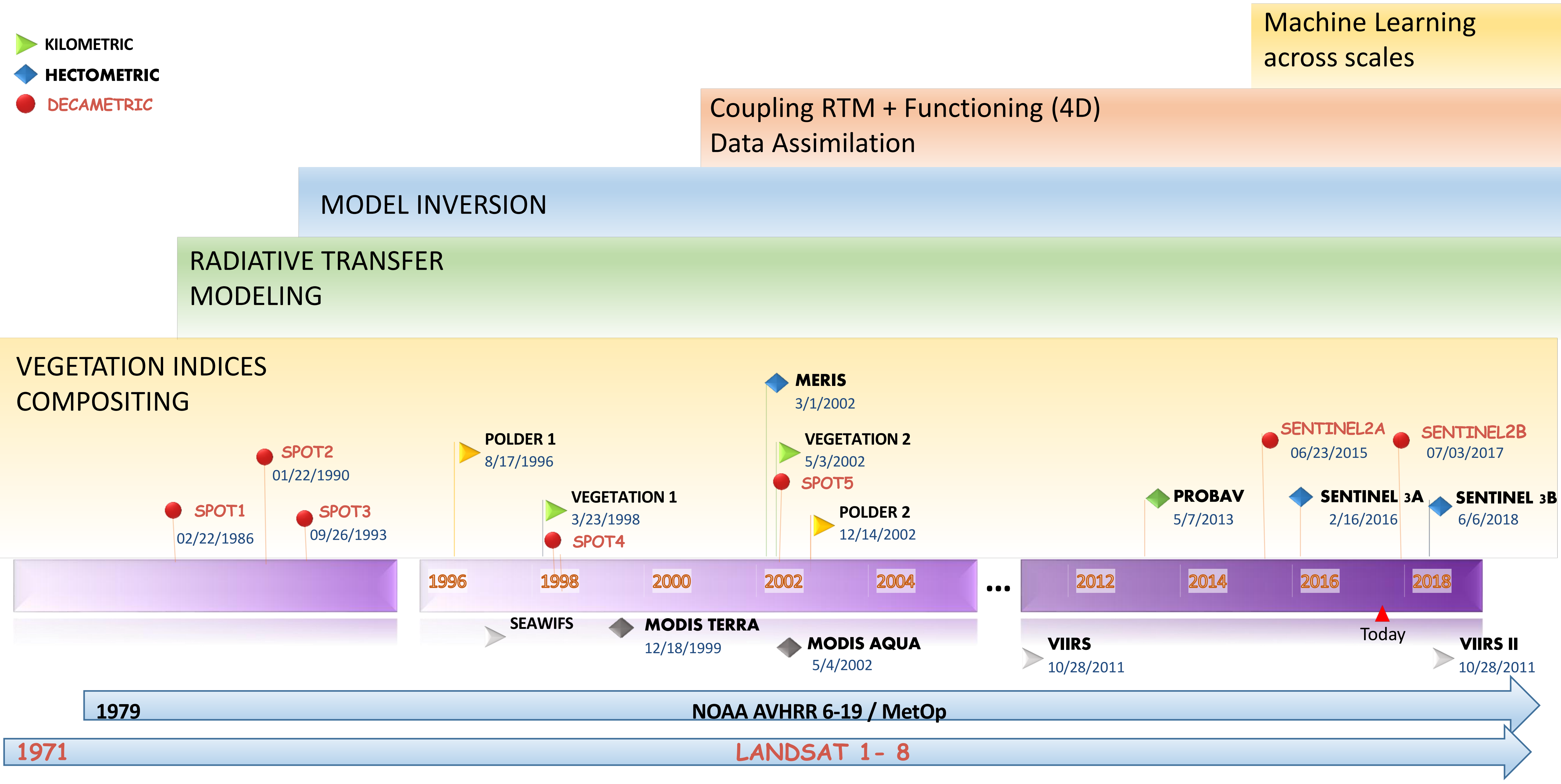
M. Weiss, F. Baret

A.Verger	M. Leroy	R. Myneni	V. Bruniquel	F. Jacob	S. Buis
F. Camacho	P. Bicheron	Y. Knyazikhin	O. Hagolle	P. Pacholcyzk	R. Lercerf
R. Lacaze	B. Berthelot	B. Tan	C. Di Bella	H. Makhmara	S. Kandasamy
B. Smets	P. Rossello	W. Yang	M.E.Beget	G. Duveiller	E. Vermote
K. Pavageau	H. Eerens	A. Pragnère	D. Raymaekers	P. Defourny	W. Li
C. Bacour	T. Block	B. Combal	J. Demarty	M. Claverie	D. Allard
S. Garrigues	B. Scholze	O. Hautecoeur	E. Fillol	V.Demarez	C. Lattore …

MEDIUM RESOLUTION: GLOBAL COVERAGE – HIGH TEMPORAL FREQUENCY



MEDIUM RESOLUTION: GLOBAL COVERAGE – HIGH TEMPORAL FREQUENCY



REQUIREMENTS FOR GLOBAL MEDIUM RESOLUTION PRODUCTS

⇨ GMES/Copernicus context:

- global monitoring for environment and security

⇨ Meet the user needs

- Accuracy : Quality Flags / Confidence Intervals / Validation
- Consistency : through time & between sensors
- No Gap

USER ALWAYS ASSOCIATED TO THE PRODUCT DEVELOPMENT

⇨ Meet technical requirements

- Operational context + Near Real Time
- Easy access to the community (ESA does not deliver Level2 & 3 data)

TECHNICAL CENTERS (VITO/CNES) INVOLVED IN PROJECTS

Biophysical Algorithms: Machine Learning

⇨ Principle:

- calibrate non linear relationships between inputs (reflectance) and outputs (biophysical variable)

⇨ Machine learning:

- currently neural networks
- Generic algorithm

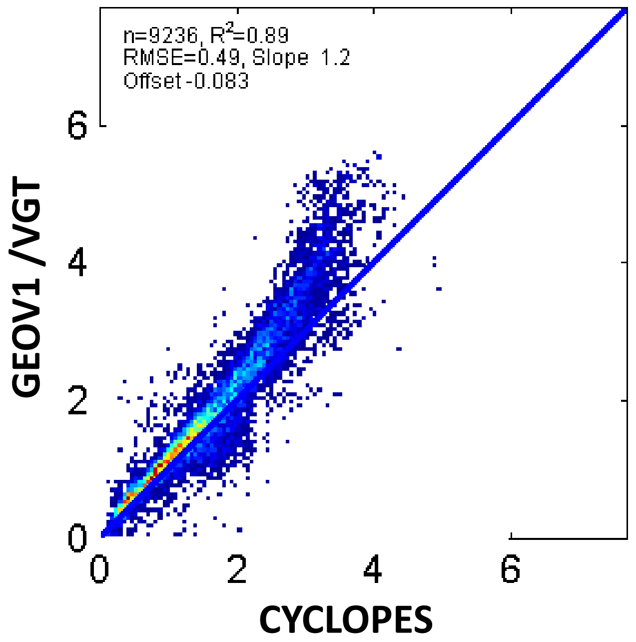
⇨ Setting up of the learning dataset is crucial: representativeness

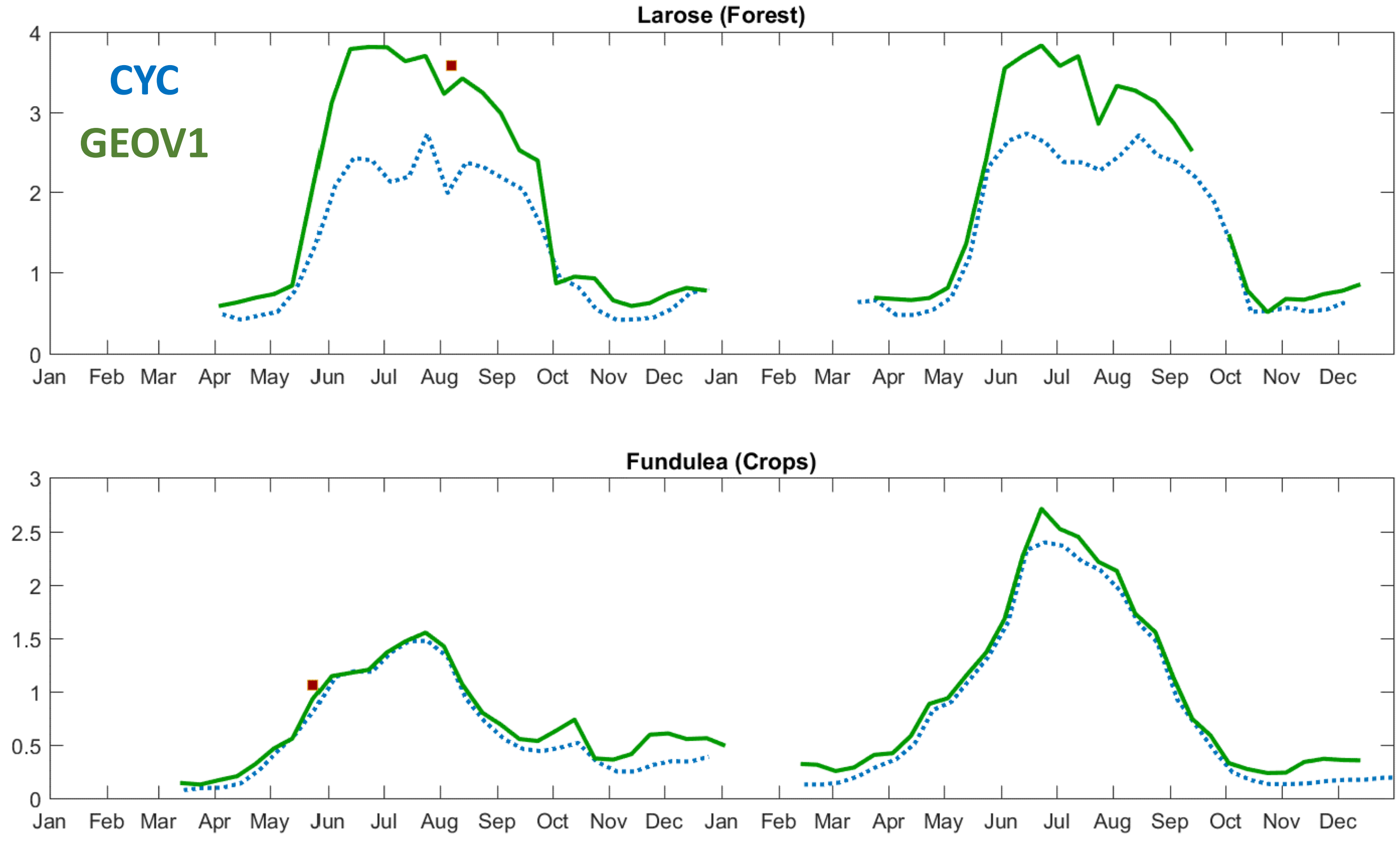
- Vegetation types, development stages & conditions
- Radiometric Noise
- Observational configuration

COPERNICUS PRODUCTS

	Name	Sensors	Resol.	Learning	Input	Temporal compositing			Improvement
						Smoothing	Gap Fil.	NRT	
	CYCLOPES	VGT1	1km	Sim. Generic	TOC Red/NIR/ SWIR	Reflectance Weighted 30 days	✗	✗	
	GEOV1	VGT1/VGT2 /PROBAV	1km	**Meas.** Generic	TOC Red/NIR/ SWIR	Reflectance Weighted 30 days	✗	✗	**Accuracy (High LAI)**
	GEOV2	VGT1/VGT2	1km	Meas. Generic	**TOA** Red/NIR/ SWIR	**Product** 30 days	✗	✗ **1-2 day lag**	**Temporal consistency**
	GEOV3 1km	VGT1/VGT2 /PROBAV	1km	Meas. **EBF/Non EBF**	TOC RED/NIR Red/NIR/ SWIR	Product **Variable temp window**	**Climato**	✓	**NRT, temporal consistency, completeness**
	GEOV3 300m	PROBAV	**300m**	Meas. EBF/Non EBF	TOC **Red/NIR**	Product Variable temp window	**Data**	✓	**NRT, spatial resolution**

RESULTS : PRODUCT COMPARISON CYCLOPES vs GEOV1

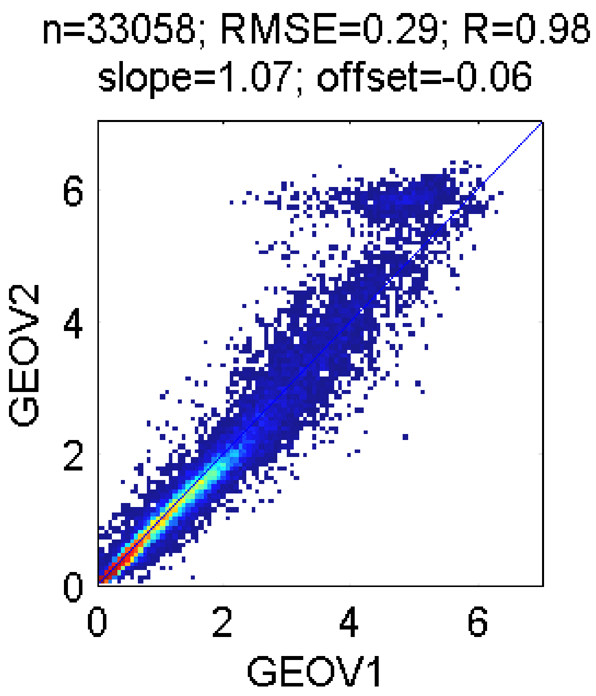


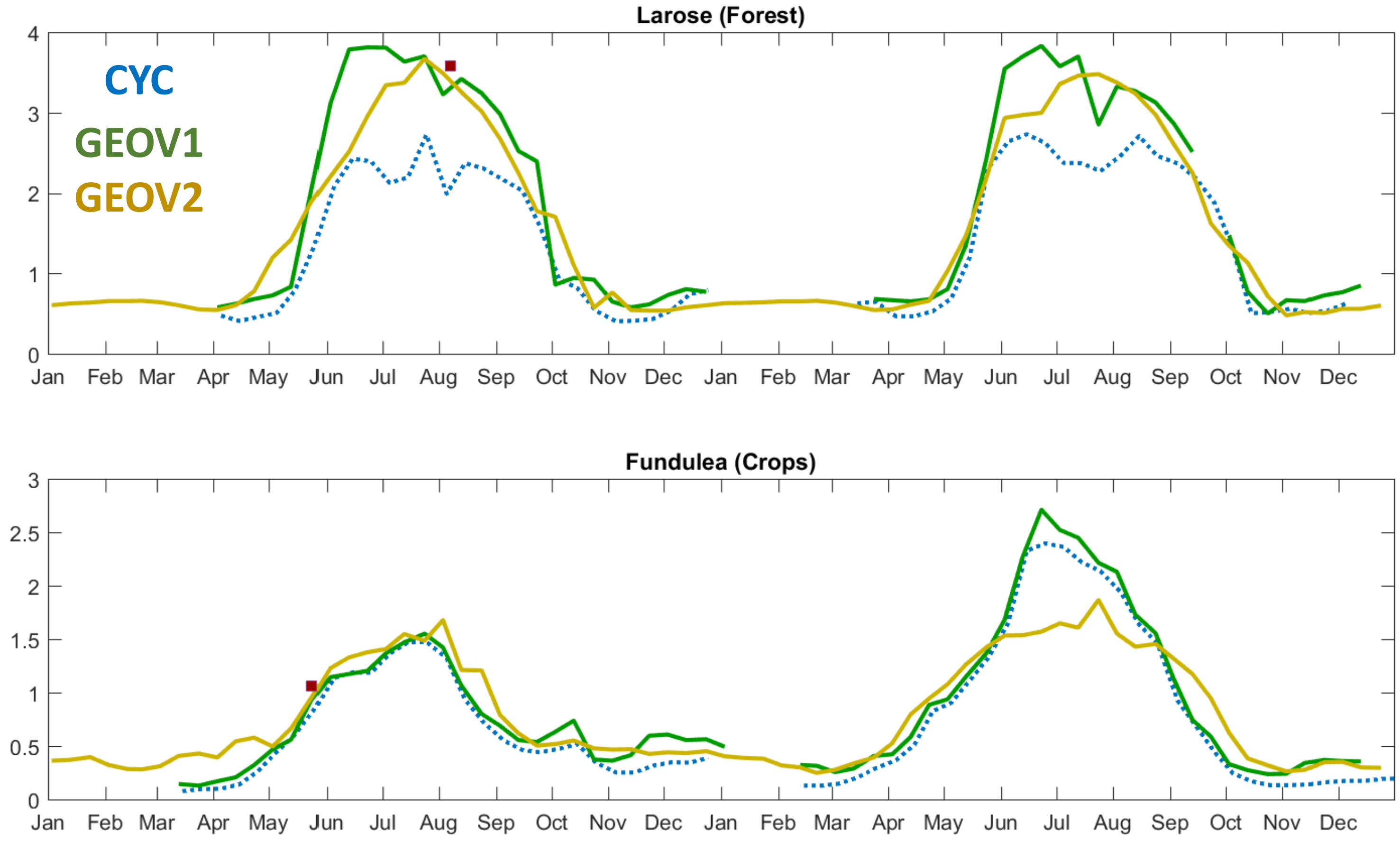


CYCLOPES/GEOV1:

1D simulations vs actual refl + fused MOD+CYC

RESULTS : PRODUCT COMPARISON GEOV2 vs GEOV1

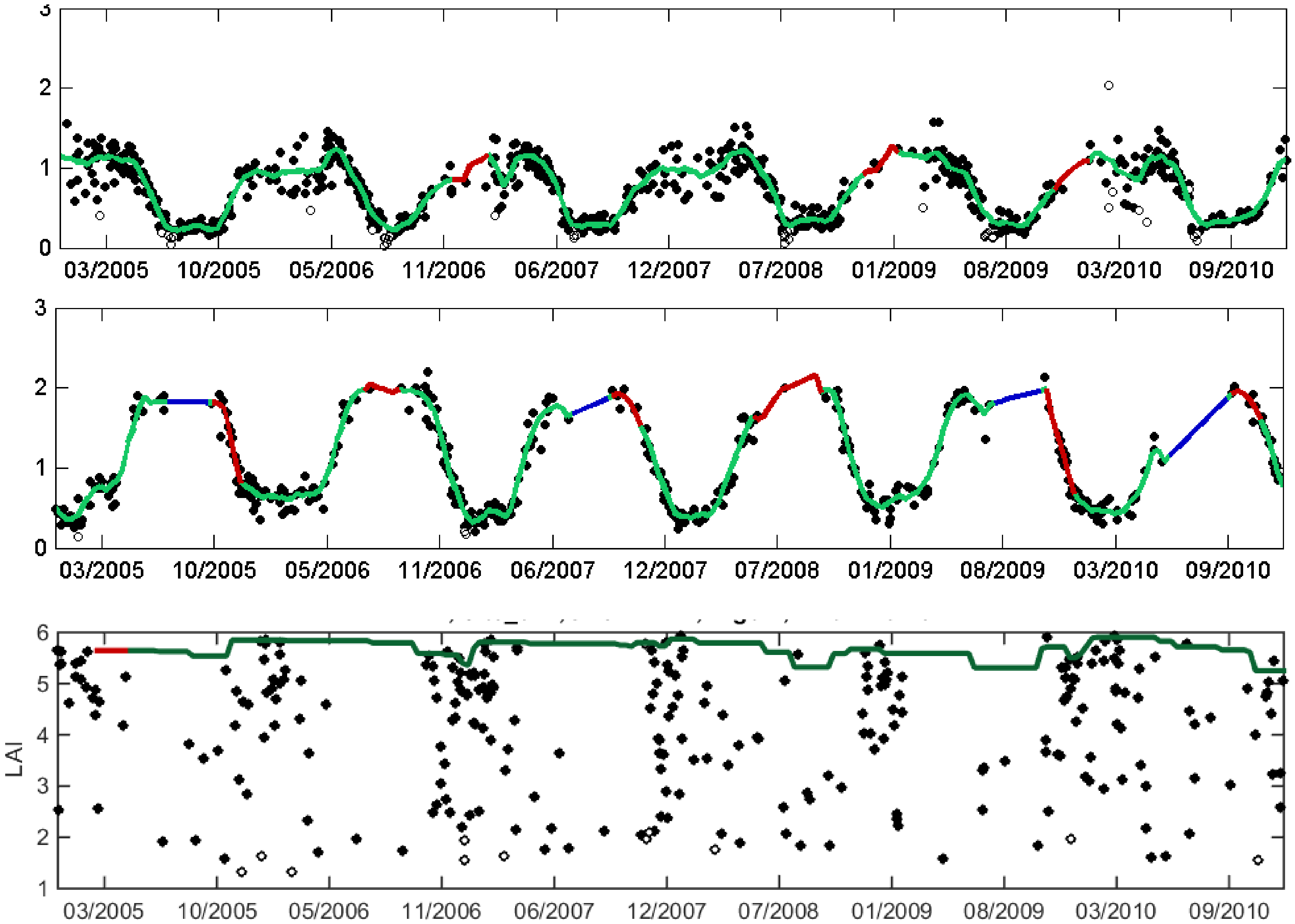




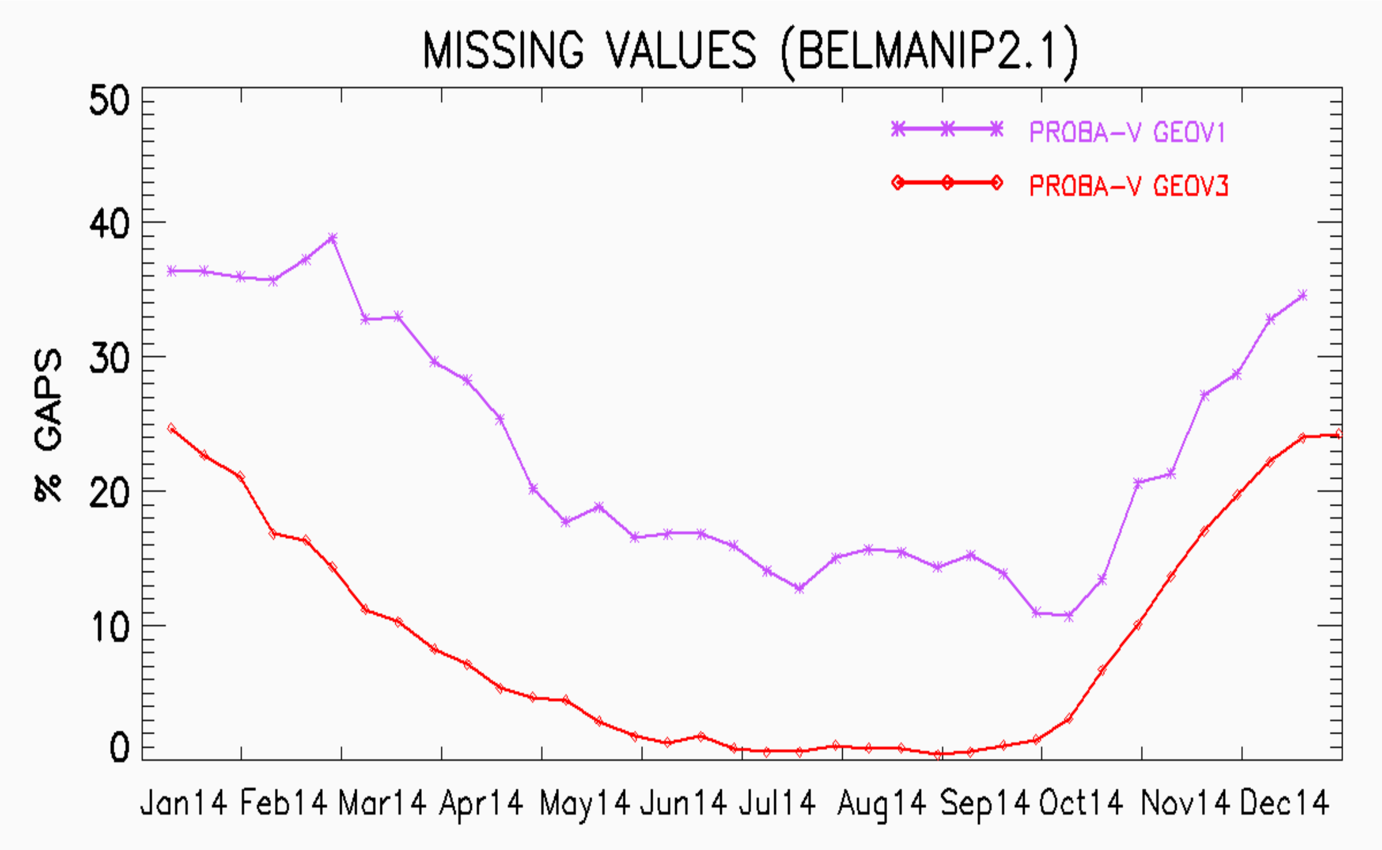
⇨ **GEOV2 /GEOV1**

- Temporal compositing at the product level
- TOA vs TOC

RESULTS : GEOV3



RESULTS – GAP FILLING



- ⇨ Polynomial fitting in GEOV3 reduce the %gaps
- ⇨ Winter period: gaps are too large because of snow & bad weather

…. AND VALIDATION

⇨2000-2005:

- 73 campaigns
- Similar to BigFoot/MODLAND
- Main limitations: spatial sampling vs man power
- Use of HR data to spatially interpolate local measurements

⇨2006:

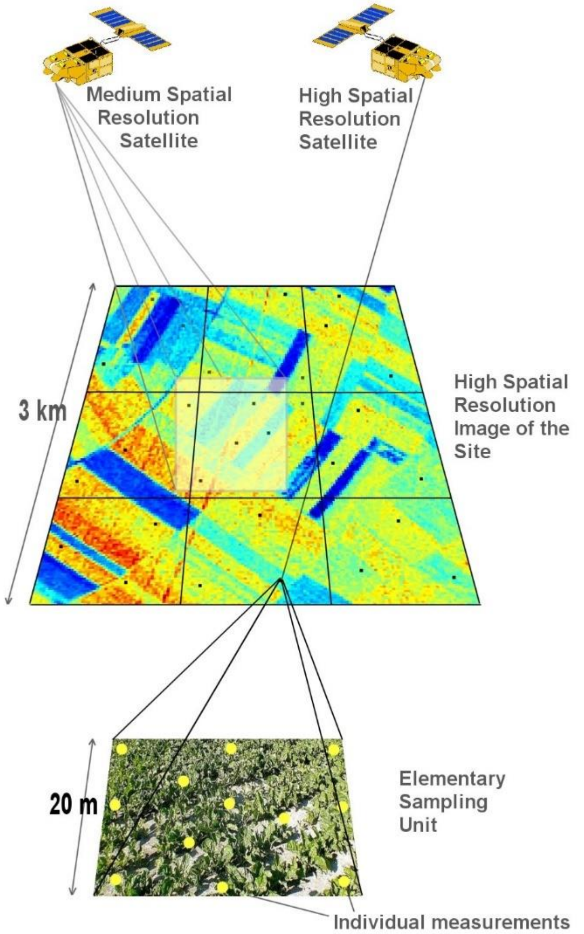
- Ground measurements: not enough
- Product inter-comparison database + machine learning

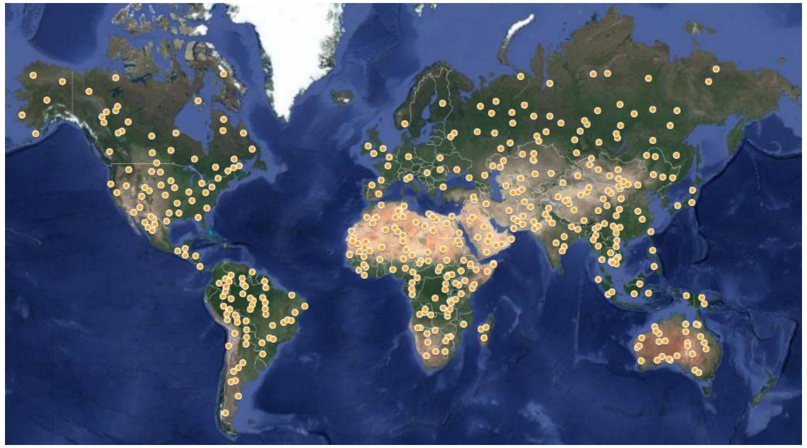
⇨2011:

- Web platform for product intercomparison
- BELMANIP2

⇨2013-2015:

- 20 ground campaigns





WHAT DID WE LEARN: biophysical algorithm

- ⇨ Cloud mask accuracy
- ⇨ TOC/TOA reflectance as inputs
  - Very good performances achieved with TOA as inputs but requires a larger training dataset
- ⇨ Class - Specific processing
  - EBF: can be identified easily and should be processed separately
    - Cloud occurrence
    - Temporal course
  - Other vegetation classes?
    - Dependence on map classification (update frequency? Mis-classification?)
- ⇨ Machine learning
  - Use of actual satellite data is better but limited by the availability of ground data (currently MODIS+CYCLOPES fused products)
- ⇨ Effective/True LAI
- ⇨ Ground truth
  - Limited by man power
  - PSF of medium resolution instruments can not be neglected in heterogeneous conditions

WHAT DID WE LEARN : temporal aspect is fundamental

⇨ Temporal consistency

- Compositing at the product level : better compromise between the temporal smoothness and the data fit
- Adaptive temporal window (function of amount of available data)

⇨Gap filling

- a priori information provides better results than mathematical fitting
- Use of climatology
  - Too long period masks possible recent evolution
  - Too short period more sensitive to atypical previous years
- Polynomial fit
  - Sensitive to noise

⇨Near Real Time :

- Not adequate at the very beginning of growing or senescent phase
- Projection is highly dependent on the amount of available data in the previous temporal window

WHAT TO DO NEXT ?

BENEFIT FROM THE AVAILABILITY OF DECAMETRIC DATA

⇨ Algorithm : Machine learning across scales

- RTM: 3D-4D modelling at very HR resolution (Fred's talk)
  - Deep learning combining accurate measurements & modelling
- Transfer to decametric products

  Machine learning (inverse model)

  Class specific algorithm : better classification accuracy, more frequent updating to capture HR phenology, especially for short vegetation cycle

- Transfer to hecto/kilometric products
  - Learning based on decametric products
  - Going backwards in time (reprocessing archive)

⇨ Validation

- Focus on the validation of decametric products

  Measurements at this scale easy to complete and more and more available

- Temporal monitoring of the vegetation cycle is mandatory

  New sensors are becoming available (IOT)

- Indirect validation of hecto/kilometric resolution through the validation of decametric products (estimate sensor PSF is mandatory)