AI4EO Activities at CommSensLab-UPC

Adriano Camps, Mercè Vall·llosera, Carlos López-Martínez

and students and collaborators

6 Febrero 2023

Outline

→ Earth Observation @ CommSensLab – UPC

→ AI4EO Activities and/or Applications

- Soil Moisture Estimation
- Forest Fire Burned Area Prediction
- Dengue Prediction in Brazil
- Multitemporal SAR Coherence for Land Mapping/Classification
- Forest Heigh Estimation based on SAOCOM L-band SAR Data
- Soil Moisture, Sea Ice Extent, Concentration and Thickness, and Sea Surface Salinity retrievals from FSSCat
- HyperSpectral Imagery Compression

Earth Observation @ CommSensLab – UPC

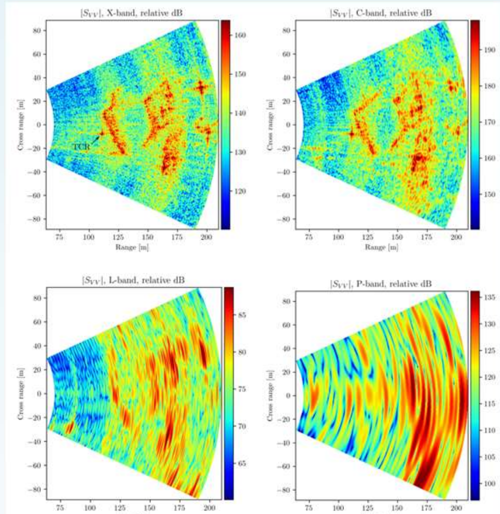
CommSensLab has a large experience in (Microwave) Earth Observation from sensors design and development to data analysis and exploitation in diverse applications

→ Synthetic Aperture Radar (SAR)

- Collaboration with

MWSE

Multifrequency ground-based SAR Sensor

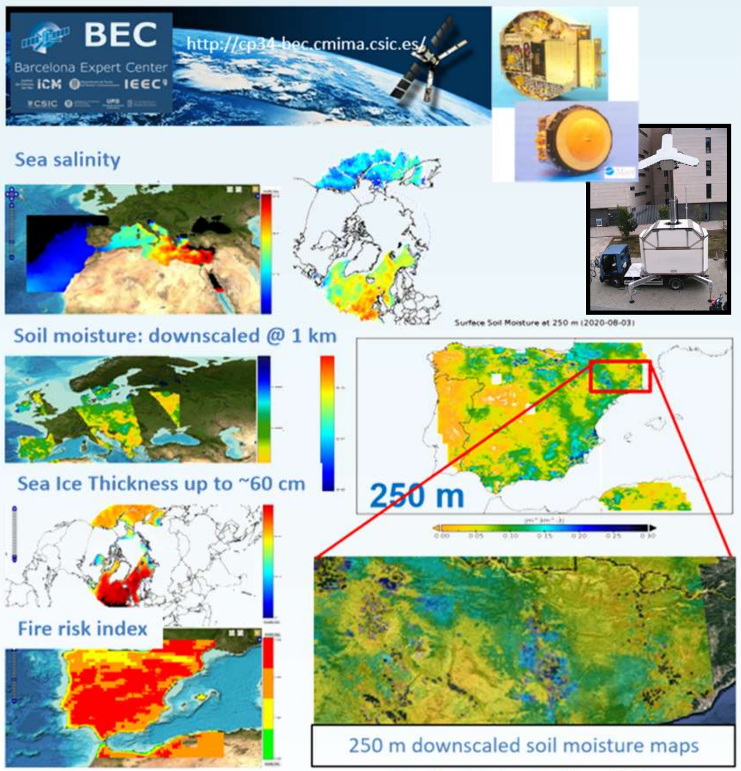


3

Earth Observation @ CommSensLab – UPC

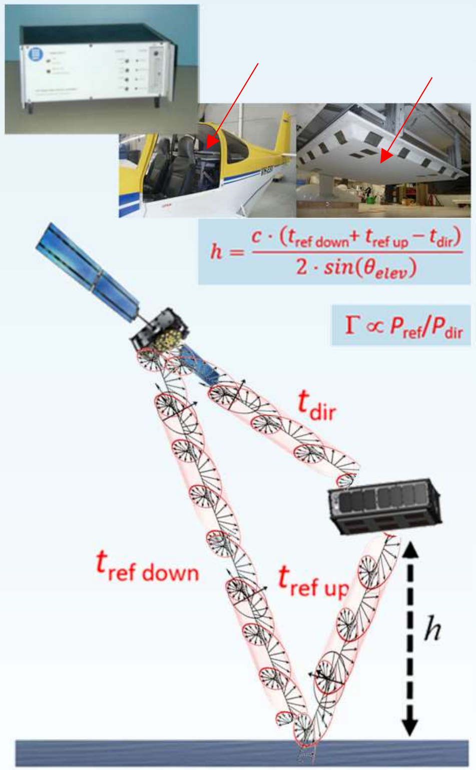
1993 2000 2007 2015

SMOS activities:
from MIRAS instrument to novel applications…

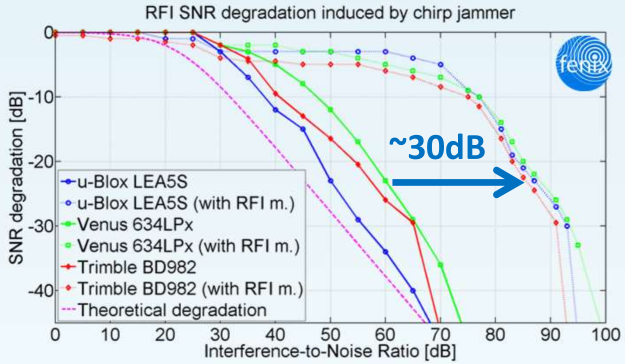


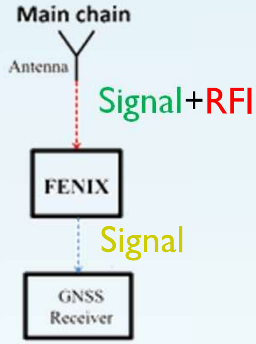
Credits: Miriam Pablos, BEC

GNSS-R:
instruments dev. and applications



RFI detection & mitigation:
MWR & GNSS-R





NanoSats: test bed of new remote sensors

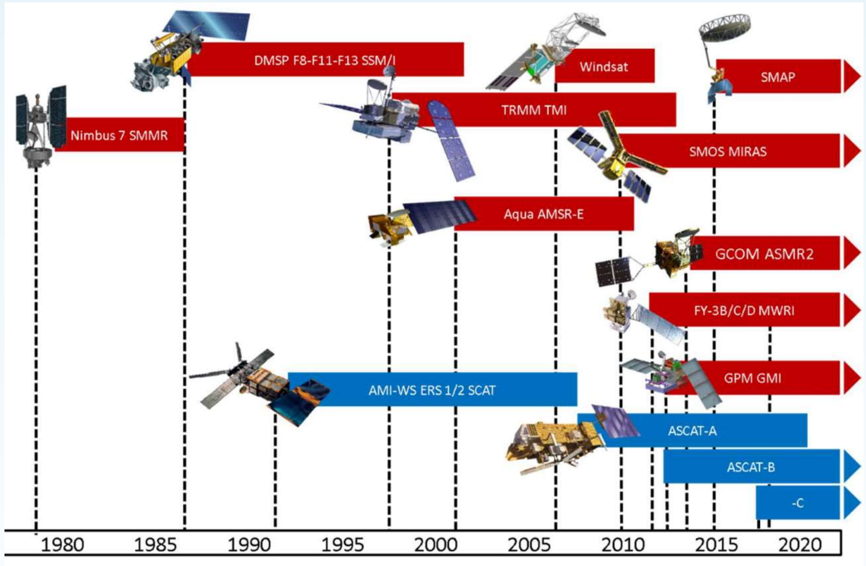
Soil Moisture Estimation: Random Forest (Data)

Predictors

Variable	Periodicity	Resolution
Sentinel 2		
12 bands	**5 days**	60 m
10 indices	**5 days**	60 m
Ancillary data		
ERA5-land Skin Temp.	daily	9 km
DEM	static	60 m
Slope	static	60 m
Aspect	static	60 m
Hillshade	static	60 m
Month	-	-

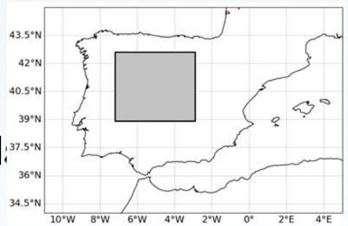
Target Variable

Variable	Periodicity	Resolution
Climate Change Initiative (CCI)		
Soil Moisture (SM)	daily	**25 km**



Microwave Instruments used for the generation of the CCI SM[1]

- Study **period**: **2018-2019**
- Study **area**: central part of the **Iberian Peninsul**



- **Sentinel 2** (A/B) data is the most **temporally constraining**
- **Clouds** can **mask Sentinel 2** data
- **ESA CCI** provides a **combined SM product** (active + passive)
- **ESA CCI SM** has the **lowest resolution**

[1]"**ESA Climate Change Initiative Plus Soil Moisture, Product User Guide (PUG), Supporting Product Version v07.1.**" Earth Observation Data Centre for Water Resources Monitoring (EODC) GmbH, May 05, 2022.

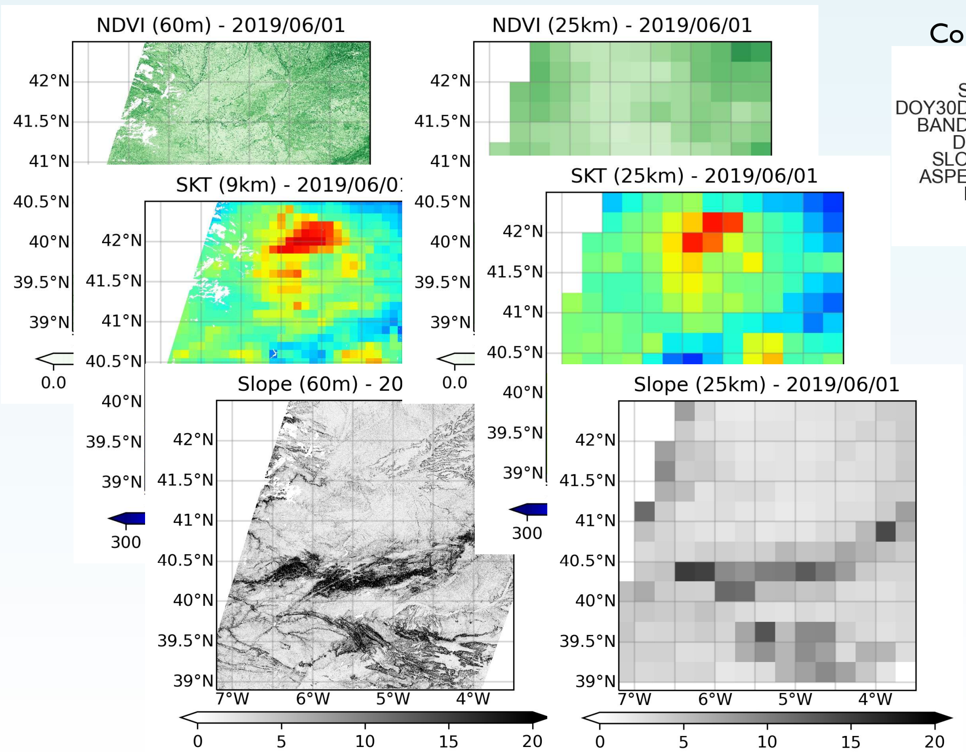
Soil Moisture Estimation: Random Forest (Methodology & Results)

Predictors

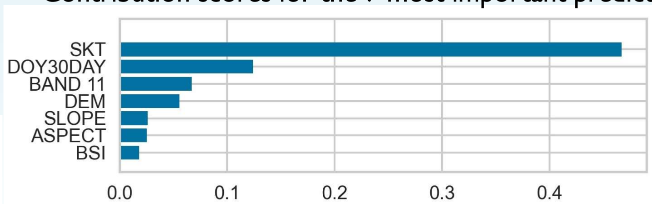
Aggregation to the target resolution (25 km)

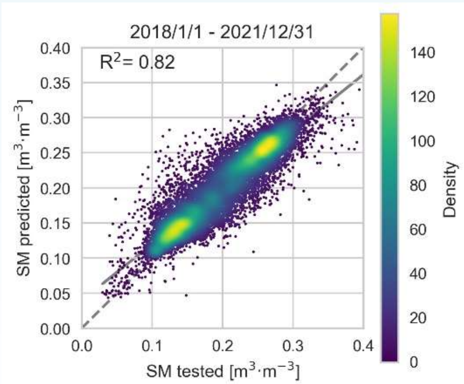
Random Forest
80% training
20% testing

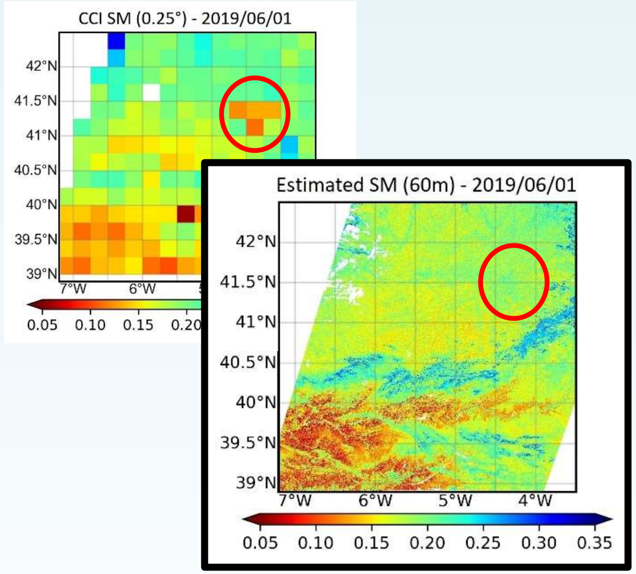
Estimation of the **60-m SM** maps



Contribution scores for the 7 most important predictors



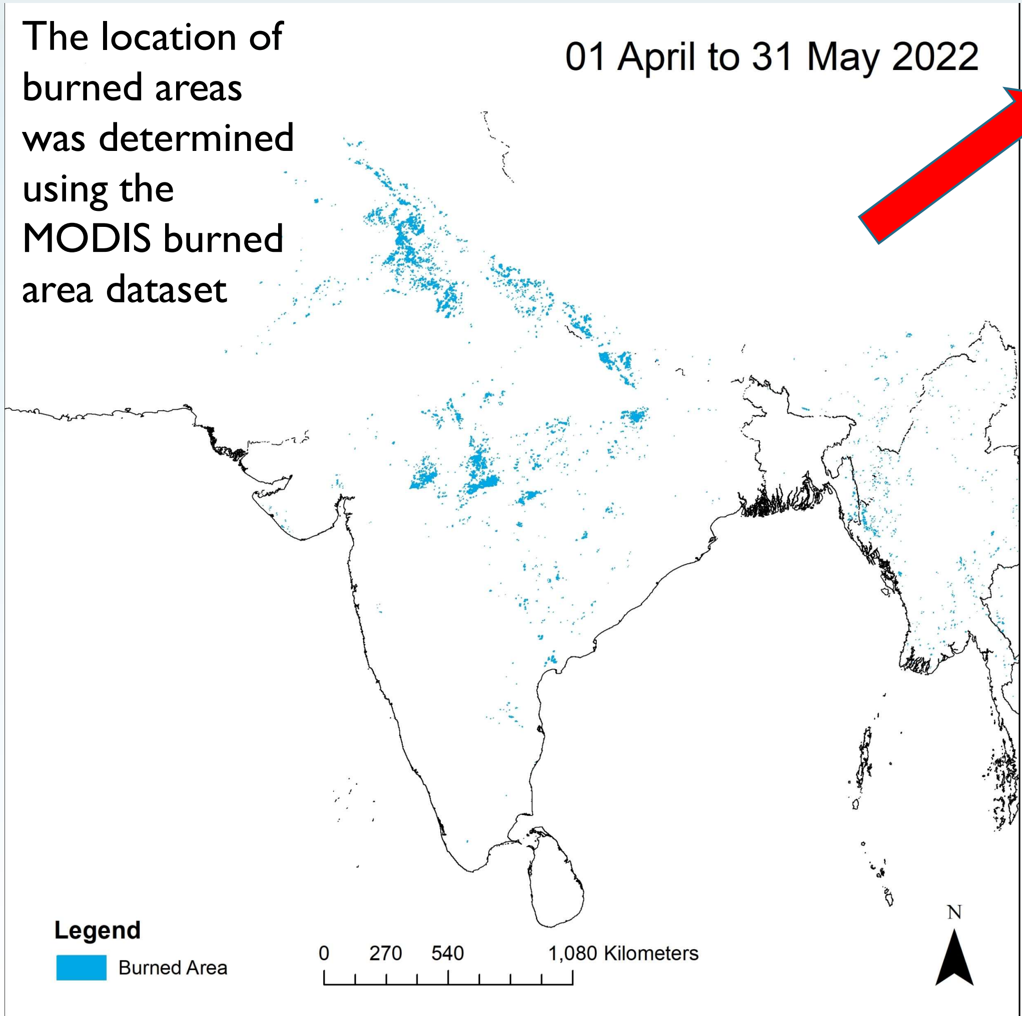




- R^2 = 0,82 and RMSE = 0,026 m^3m^{-3}
- High spatial heterogeneity
- Difficulty to detect extremes

Forest Fire Burned Area Prediction (Data)

The location of burned areas was determined using the MODIS burned area dataset



ANÀLISI

Com ens arribarà l'efecte dominó de l'onada de calor extrema a l'Índia i el Pakistan

Temperatures sense precedents en 120 anys de registres col·lapsen els serveis bàsics i la salut de tots dos països, un tast del que està venint pels efectes de l'escalfament global

JORDI VILARDELL GÓMEZ

Periodista de TV3 especialitzat en crisi climàtica i de biodiversitat

@JordiVilardell

11/05/2022 - 23.13 | Actualitzat 20/06/2022 - 09.18

Incendis forestals? Milers

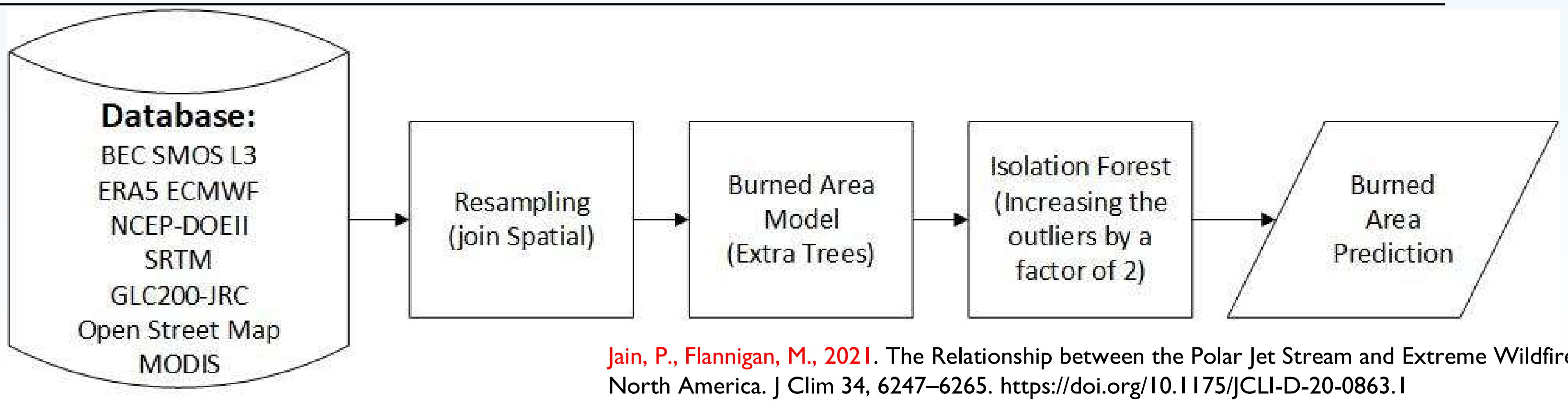
A finals d'abril es van detectar, en tres dies, més de 7.800 incendis forestals a l'Índia.

https://www.ccma.cat/324/com-ens-arribara-lefecte-domino-de-lonada-de-calor-extrema-a-lindia-i-el-pakistan/noticia/3162928/

Forest Fire Burned Area Prediction (Data)

Source	Parameter	Resolution	Explanation
BEC SMOS L3	SM	25 km	Extent of the study by Chaparro et al. (2016), who used remotely observed SM and LST to predict the fires extent
	VOD	25 km	
ERA5 ECMWF	VPD	0.25°	Indicate the aridity conditions in the surface air
(https://cds.climate.copernicus.eu/)	LST	0.25°	-
NCEP–DOE II	u_{300} and v_{300}	2.5°	To determine jet stream characteristics in relation to a very big fire using Spatiotemporal Composite technique. Further insight can be found in Jain & Flannigan (2021)
(https://psl.noaa.gov)	ΔZ_{500}	2.5°	
SRTM (https://portal.opentopography.org/)	Elevation	90 m	Obtained through the United States Geological Survey (USGS)
GLC2000-JRC (https://forobs.jrc.ec.europa.eu/)	Land use	1 km	Global Land Cover Product (GLC) coordinated by Forest Resources and Carbon Emissions (IFORCE)
Open Street Map (https://www.openstreetmap.org/)	Distance to road	-	Calculated in GIS by Euclidean Distance
MODIS Land Product (https://lpdaac.usgs.gov/)	Burned Area	1 km	MODIS burned area datasets (MC64A1) is obtained through sftp (Server: fuoco.geog.umd.edu, Login name: fire, Password: burnt)

Method:



Jain, P., Flannigan, M., 2021. The Relationship between the Polar Jet Stream and Extreme Wildfire Events in North America. J Clim 34, 6247–6265. https://doi.org/10.1175/JCLI-D-20-0863.1

Forest Fire Burned Area Prediction (Methodology)

Random Forest (RF) and Extremely randomized Trees (Extra Trees) are both methods that use multiple decision trees to make predictions

Capability	**Extra Trees**	**RF**
Level of randomness	High, as it uses random thresholds for each feature	Moderate, as just the best split point is used
Decision tree constructed	All samples are considered for each split	By bootstrap aggregating the samples (that is with replacement)
Features that are considered for each split	All features	A random subset of features
Handling noisy and extreme data	Better	Good
Conclusion:	**Extra Trees can be useful when the data is more extreme or contains more noise because the higher level of randomness in the construction of the decision trees**	

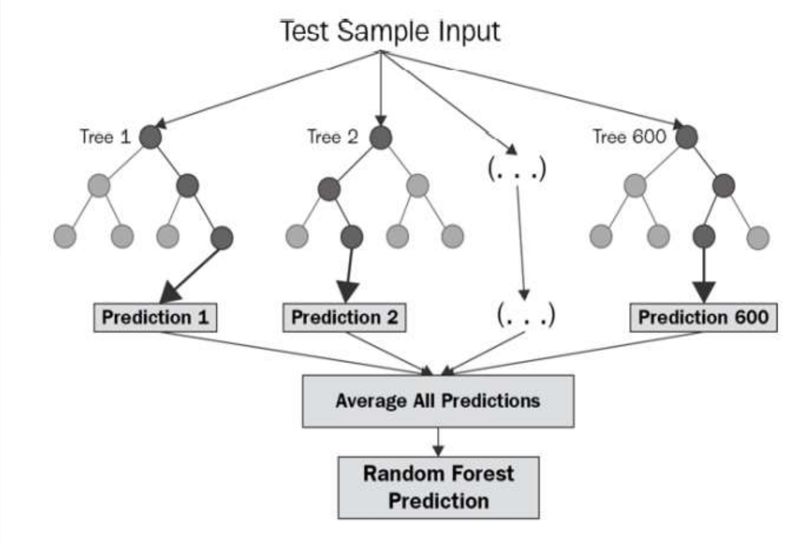


Illustration of an RF algorithm structure obtained from Serra, (2021)

Serra, A.P., 2021. Classification Strategies for Multi-Temporal Sentinel-1 Data. Universitat Politècnica de Catalunya, Barcelona.

Forest Fire Burned Area Prediction (Methodology)

Isolation Forest

It uses the concept of decision tree to identify the anomalies (or extreme values)

The method is based on the assumption that anomalies are less dense and isolated compared to the normal data points, so the fewer splits required to isolate an instance, the more likely it is to be an anomaly

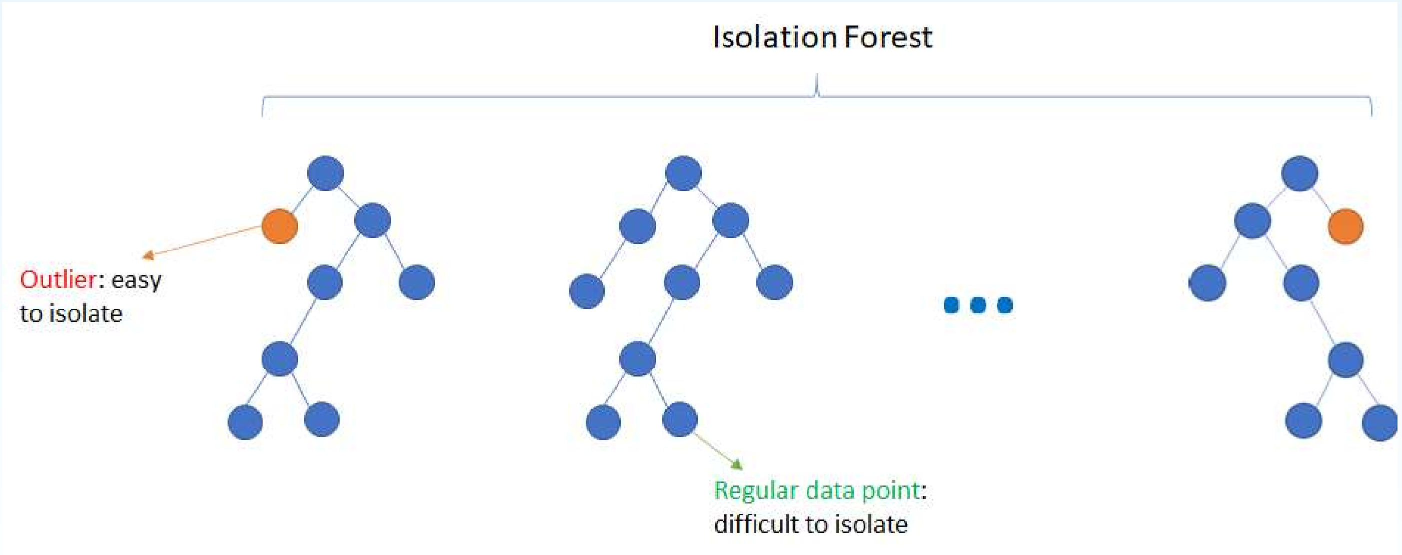
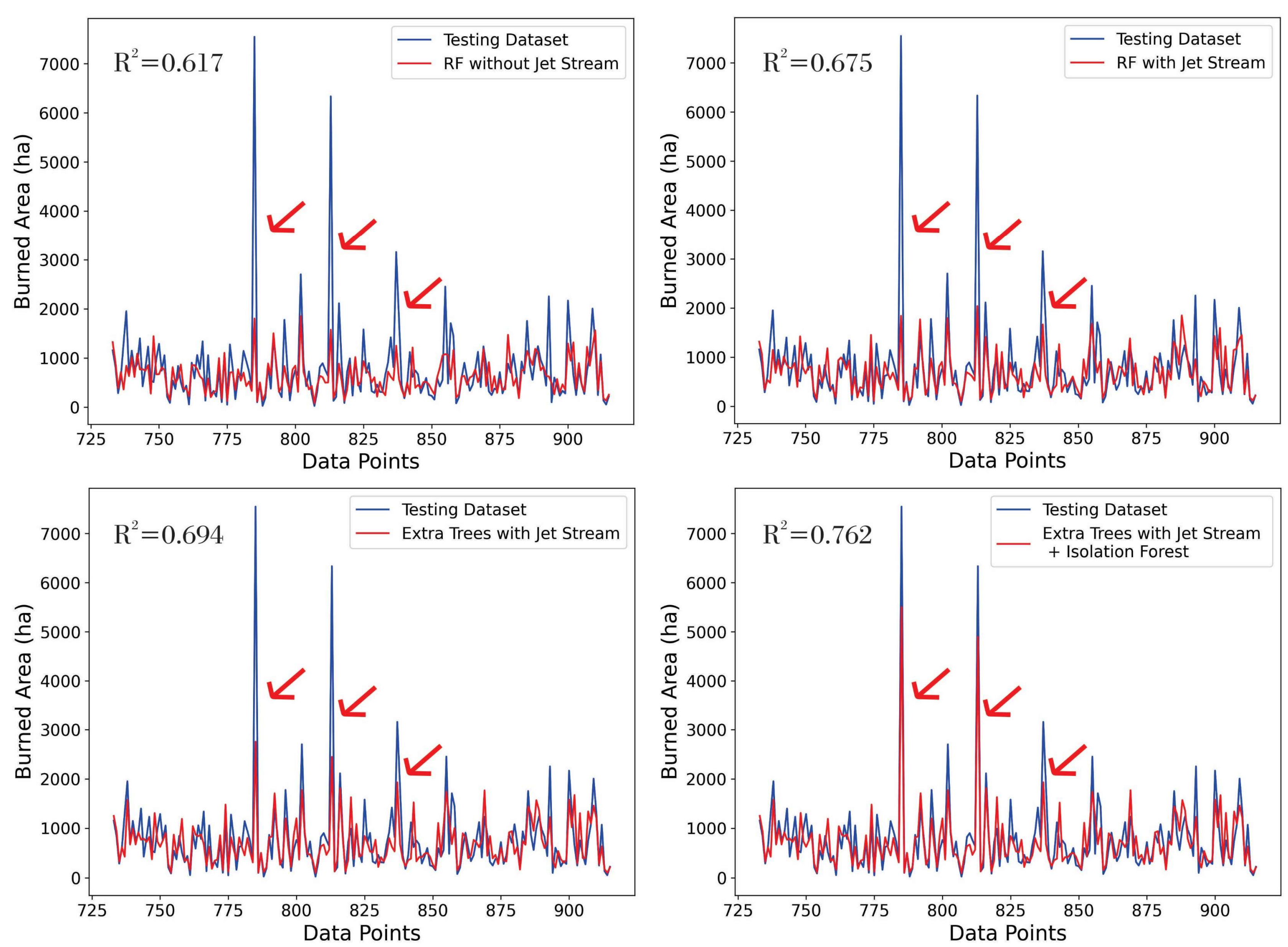


Illustration of an Isolation Forest

Forest Fire Burned Area Prediction (Results)

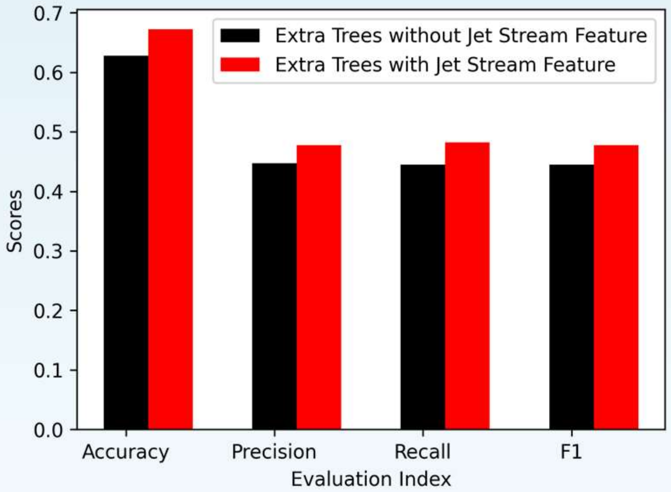
Results: Burned Area Prediction

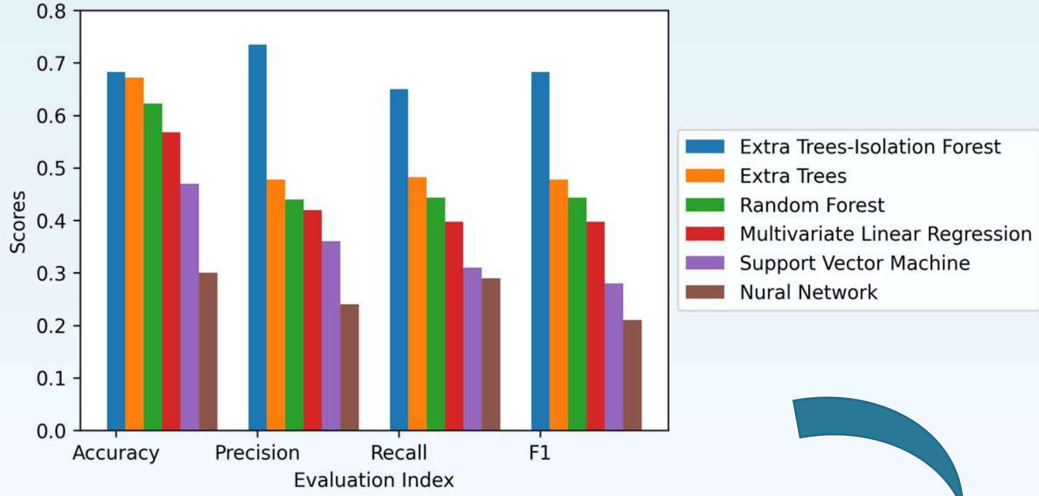


Forest Fire Burned Area Prediction (Results)

Model Comparison with other ML Algorithms

We separated the fire class categories into low, medium, large, and very large for model comparison and then calculated the evaluation index for all models.



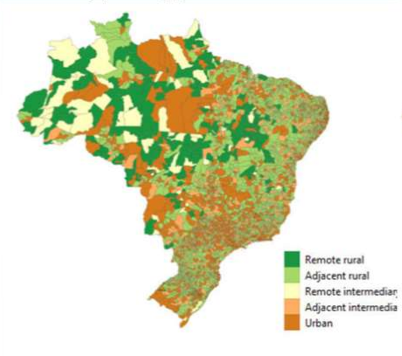


Best Model	Evaluation Index	Low fire (≤500 ha)	Medium fire (>500 - 1000 ha)	Large fire (>1000 - 3000ha)	Very large fire (>3000 ha)
Extra Trees + Isolation Forest	Accuracy	0.86	0.61	0.46	0.67
	Precision	0.79	0.63	0.52	1.0
	Recall	0.86	0.61	0.46	0.67
	F1-Score	0.82	0.62	0.49	0.80

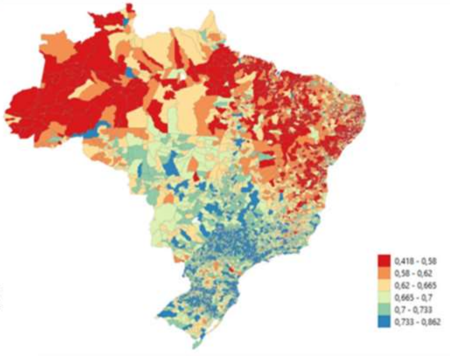
Dengue Prediction in Brazil

- **Environmental data from satellite**: NDVI, NDWI and LST, accumulated precipitation (in mm), Soil moisture
- **Dengue episodes distribution data:** Notifiable Diseases Information System (SINAN), developed by Ministry of Health of Brazil and available at DATASUS
- The actual risk

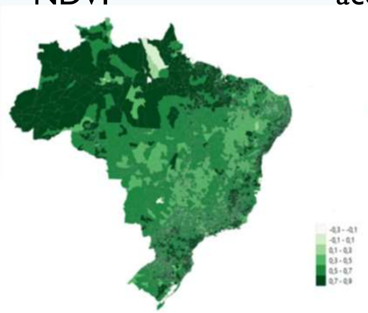
Urban typology classification



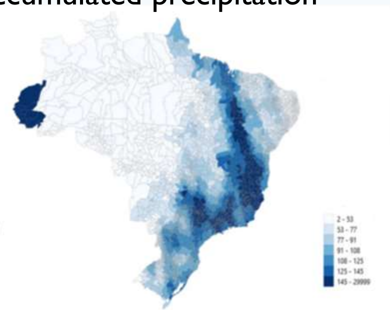
Municipal Human Development Index



NDVI



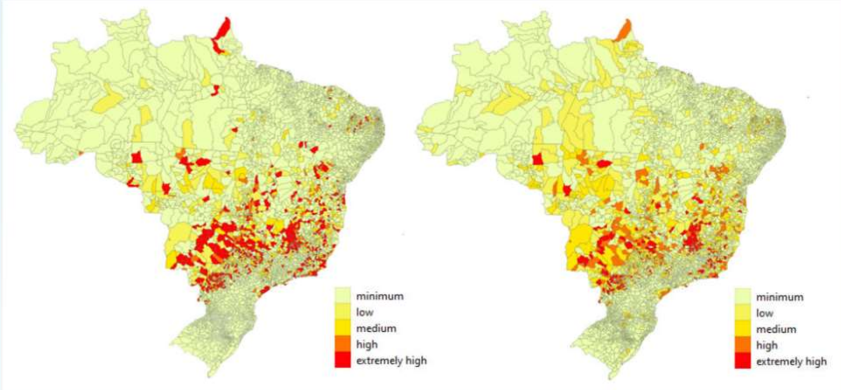
accumulated precipitation



10 most decisive parameters

Random Forest
Current Risk
Precipitation
Day Temperature
Night Temperature
NDWI
Soil Moisture
Federal state
NDVI
IDHM
Urban typology

Risk index

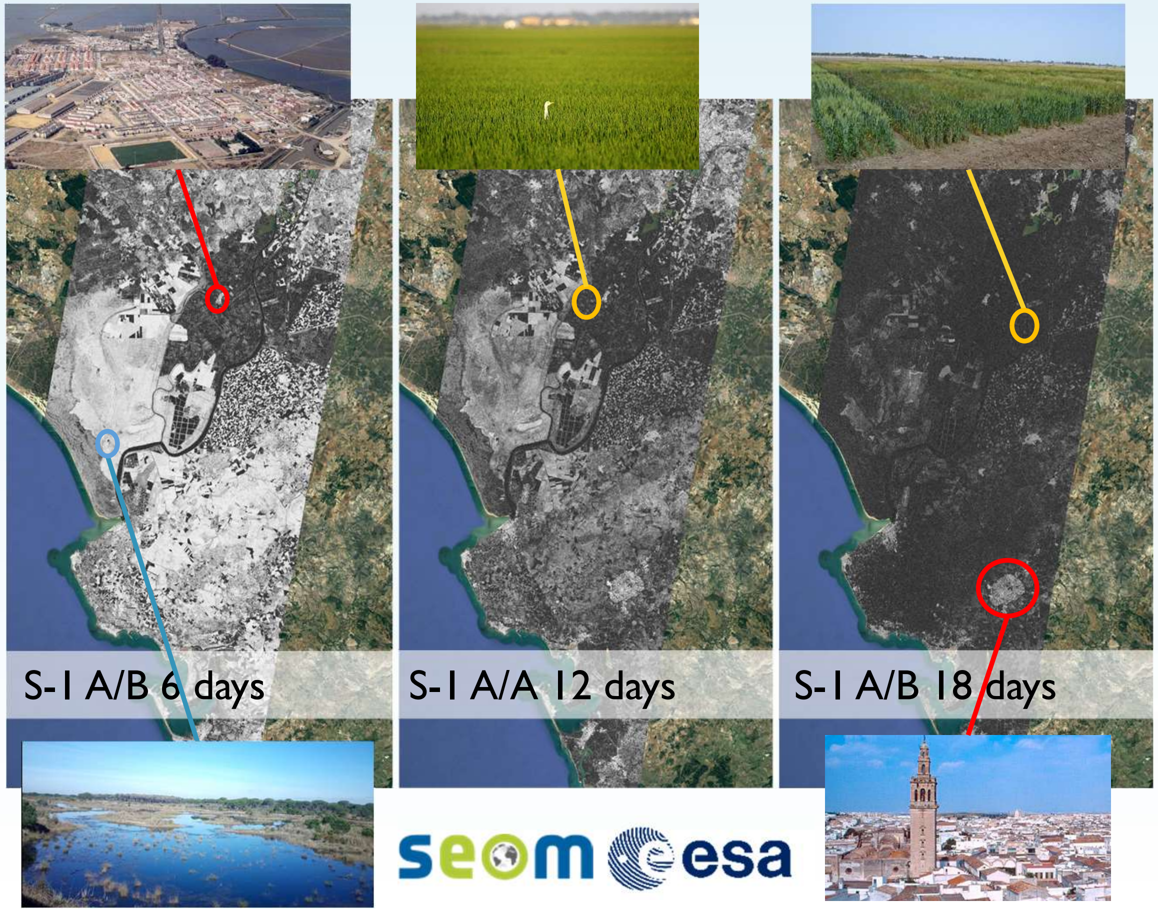


Risk indexes maps obtained for March 2013 using the registered data (left) and the random forest model (right)

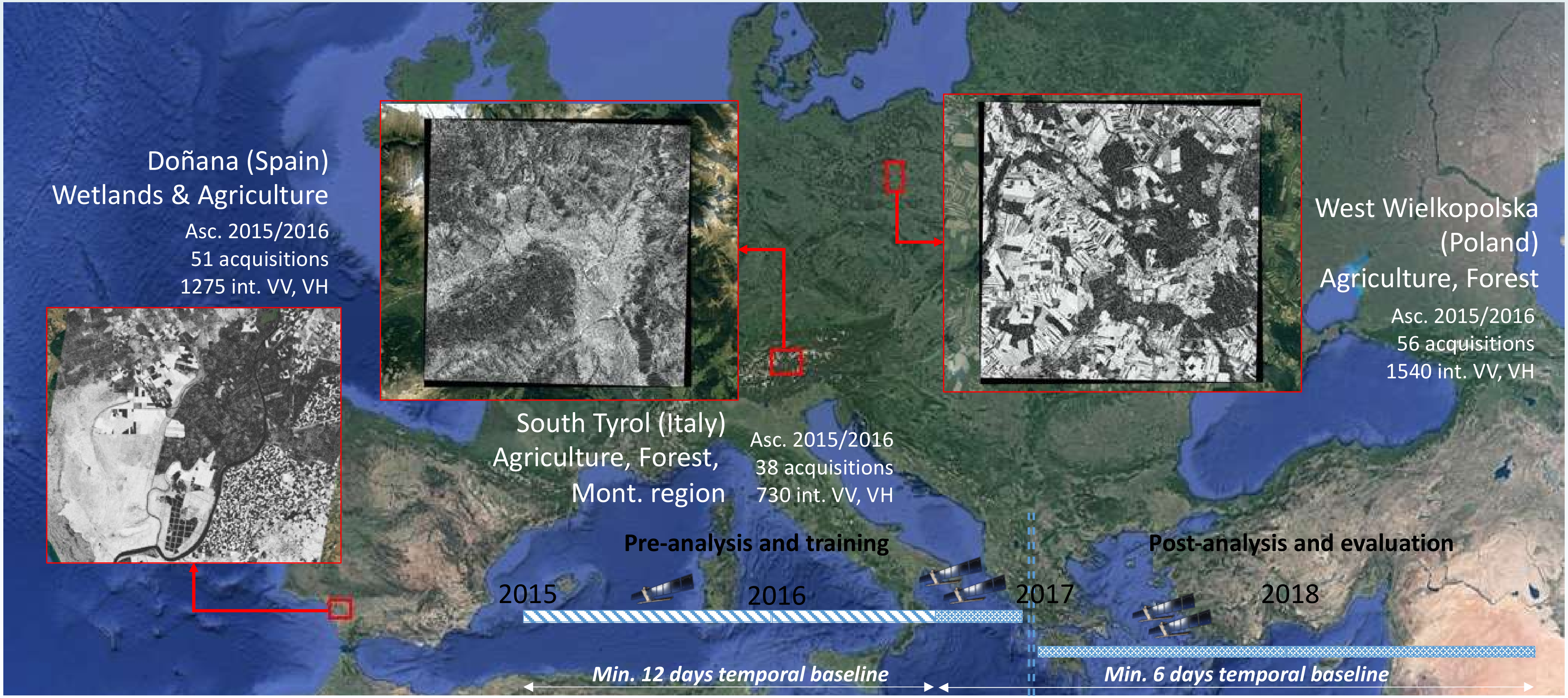
INDEX A	VALUES *Cases per 100000 inhab*
Minimun	*< 100*
Low	>100, <200
Medium	>200, <300
High	>300, <400
Very High	>400

Multitemporal SAR Coherence for Land Mapping/Classification

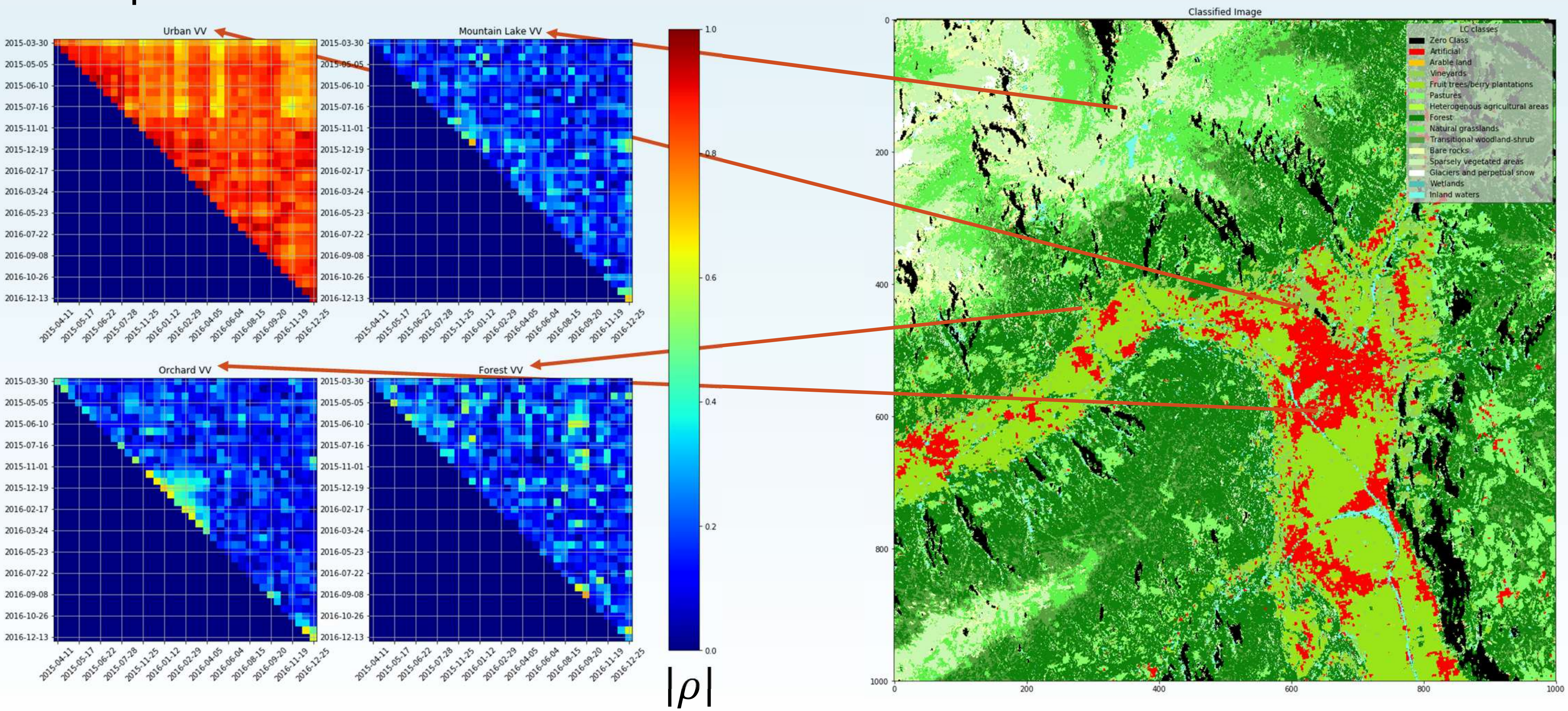
SinCohMap Project: Develop, Analyse and Validate Novel Methodologies for Land Cover & Vegetation Mapping/Classification Using Sentinel-1 **Interferometric Coherence Evolution**



Multitemporal SAR Coherence for Land Mapping/Classification



Multi-temporal Coherence Evolution



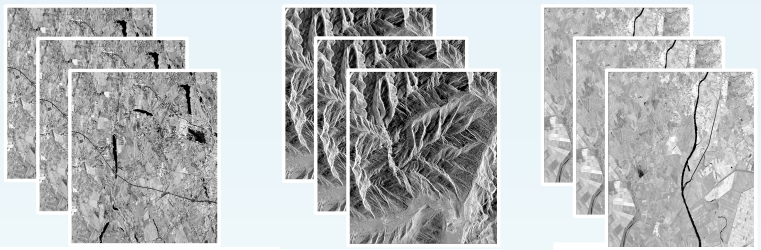
30/03/2015-13/12/2016

South Tyrol (Italy)

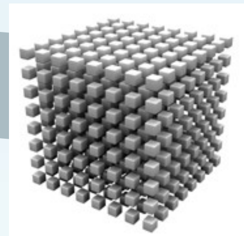
Multitemporal SAR Coherence for Land Mapping/Classification

Round-Robin Experience

Pre-processed InSAR stack sites

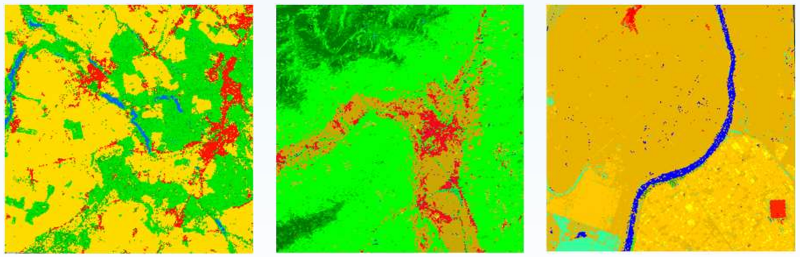


West Wielkopolska (Poland) South Tyrol (Italy) Doñana (Spain)



WCS/WCPS

Land cover maps



IGiK
Universitat d'Alacant Universidad de Alicante
Universitat Politècnica de Catalunya BarcelonaTech
eurac research
SInCohMap
LIST
Université de Rennes 1
DARES Looking Forward
eurac research | Sentinel Alpine Observatory
rasdaman raster data management
OGC Making location count.
jupyter
Aalto University
Bundesamt für Kartographie und Geodäsie

Multitemporal SAR Coherence for Land Mapping/Classification

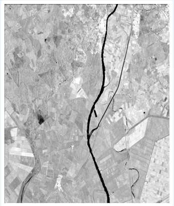
Classification Methodologies Comparison: Classical & ML/DL

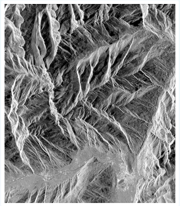
- Role of the methodology and role of the input features

Methodology	Space Object	Decision Type	Temporal Baselines	Polarization	Intensity
Random Forest	Pixel	ML Classifier	shortest	VV & VH	Yes
Eigen-value Decomposition + RF	Pixel	ML Classifier	all	VV & VH	Yes
Temporal Dynamic Indices + RF	Pixel	ML Classifier	all	VV & VH	Yes
Object-based (KTH-SEG) SVM	Object	ML Classifier	two shortest	VV & VH	Yes
Super-pixel (SLIC) + kNN	Object	ML Classifier	all	VV & VH	No
Expert Knowledge Decision Tree	Object	Decision Tree	shortest	VV & VH	No
Data Adaptive Rule-Based	Object	Threshold	selection	VV & VH	No

Multitemporal SAR Coherence for Land Mapping/Classification

Impact of Polarimetric & Interferometric SAR Information





		VV	VH	VV + VH
Doñana (Spain)	**Coherence**	**78.2**	**74.5**	**79.8**
	Intensity	**74.6**	**74.0**	**77.8**
	Coherence + Intensity	**81.9**	**79.8**	**83.3**
South Tyrol (Italy)	**Coherence**	**70.9**	**68.8**	**72.5**
	Intensity	**54.9**	**56.1**	**58.9**
	Coherence + Intensity	**72.0**	**70.9**	**73.8**
West Wielkopolska (Poland)	**Coherence**	**71.0**	**67.4**	**71.7**
	Intensity	**64.0**	**66.9**	**69.5**
	Coherence + Intensity	**73.2**	**71.7**	**75.0**

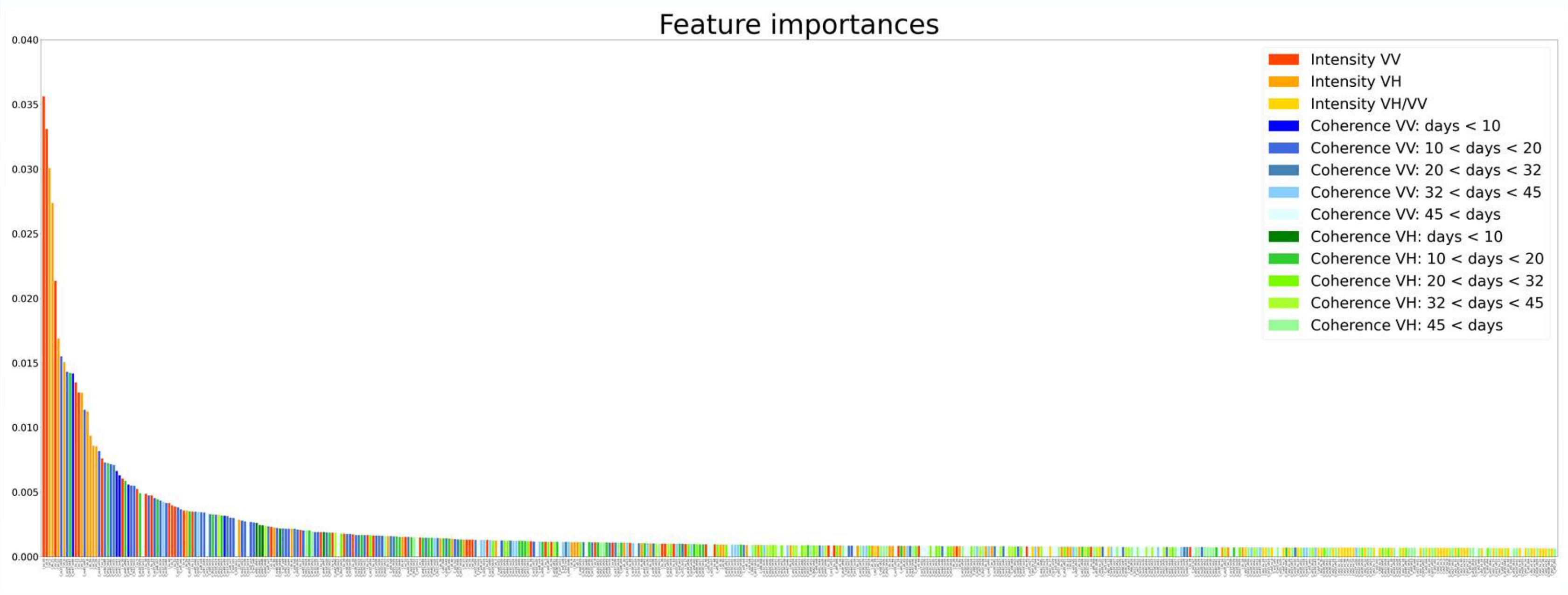
Overall Accuracy

Methodology: Eigen-decomposition + Random Forest

Multitemporal SAR Coherence for Land Mapping/Classification

Analysis of Temporal Features in a Random Forest Approach

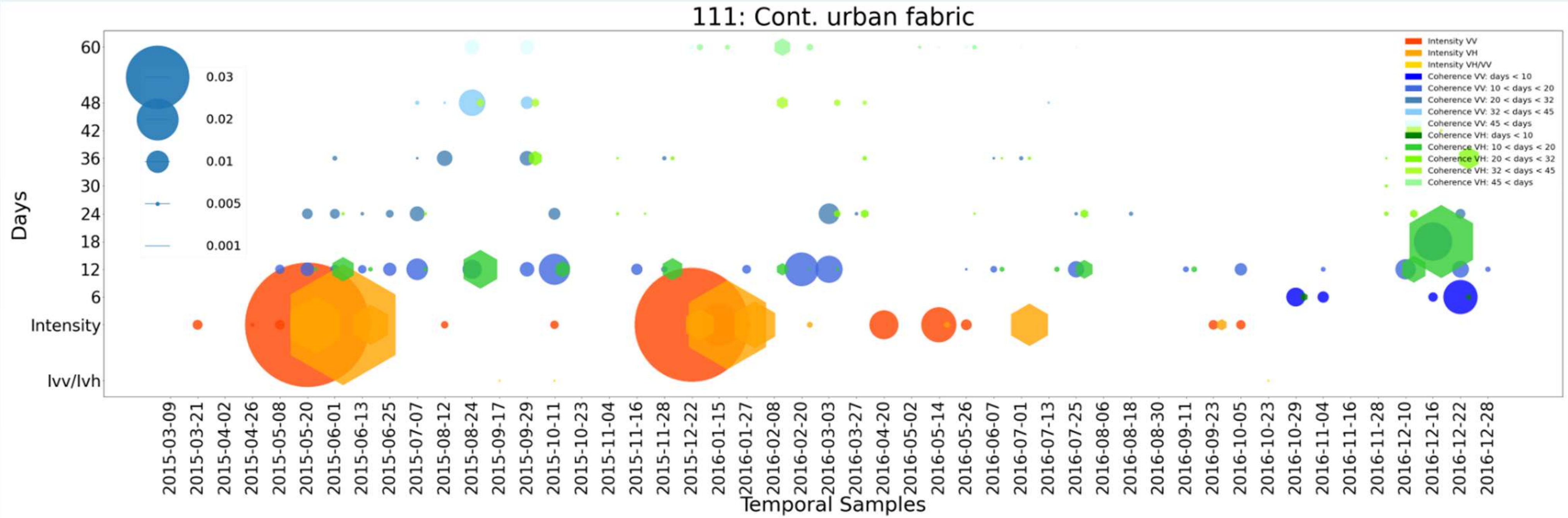
Global feature importance: Importance of each feature in the final classification accuracy



Lack of information about how much each feature contributes to each class classification accuracy and lack of temporal information

Multitemporal SAR Coherence for Land Mapping/Classification

Feature importance per class: Importance of each feature in the final classification accuracy



Multitemporal SAR Coherence for Land Mapping/Classification

Temporal features are selected specifically:

- Yearly
- Minimum temporal samples 10 & 20 in identified times detected as important
- Minimum temporal samples 10 & 20 in random times

Type	Number of estimators	Max. num. features	Number of Images	Number of features	Overall Accuracy	Macro Average	Weighted Average
Default	100	23	49	519	88.99%	75.86%	87.27%
Optimum	400	50	49	519	89.89%	78.02%	88.44%
2015*	400	50	20	200	86.10%	67.31%	83.80%
2016*	400	50	29	319	88.97%	75.01%	87.36%
10**	400	50	10	70	87.06%	68.89%	85.29%
20**	400	50	20	138	88.75%	74.77%	87.11%
10 random***	400	50	10	66	82.69%	57.03%	80.05%
20 random***	400	50	20	136	86.98%	69.89%	84.89%

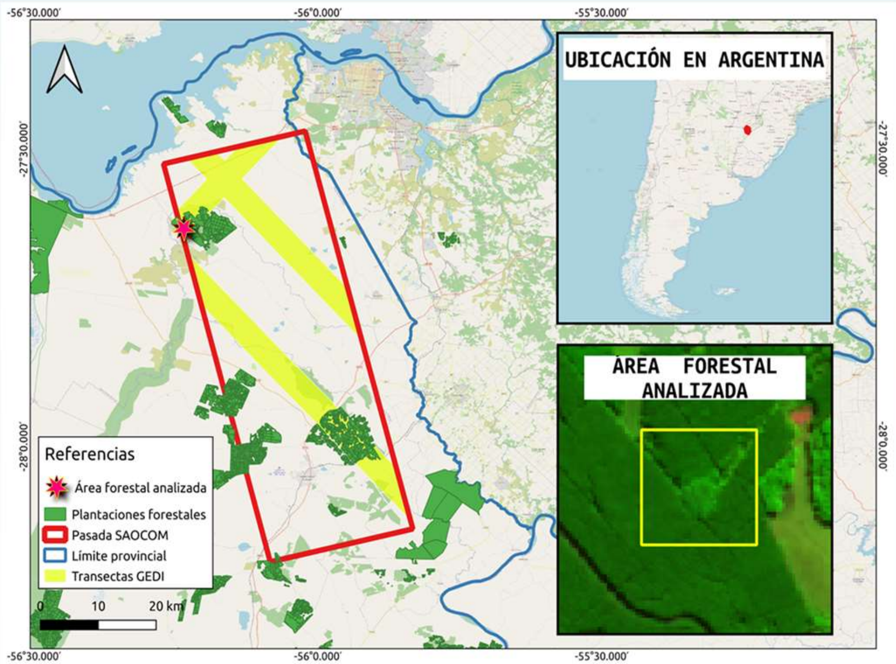
* Year representation **Set of relevant temporal samples ***Set of random temporal samples

Temporal samples reduced by 80%
but classification accuracy only drops 3%
Reduction of computation power & storage

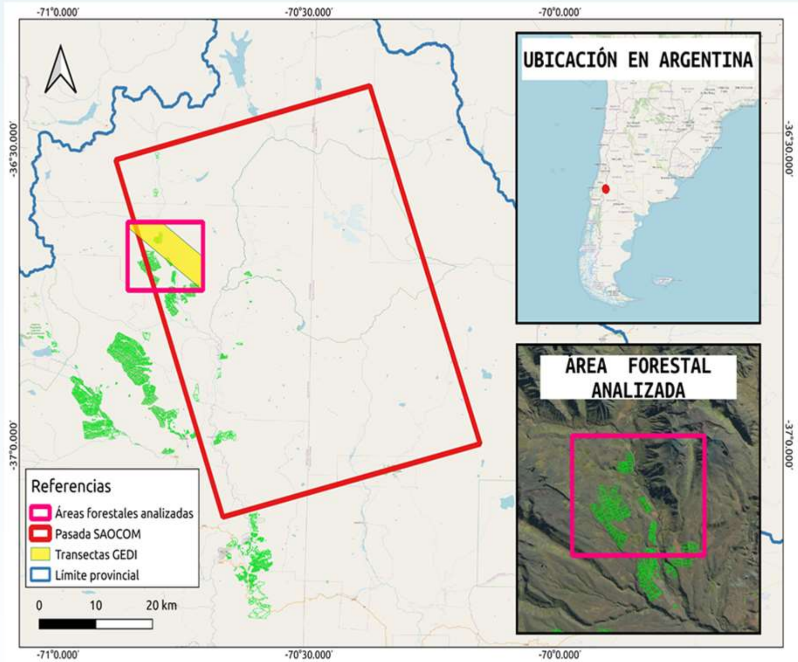
Forest Heigh Estimation based on SAOCOM L-band SAR Data

Study test sites with Multitemporal L-band SAR data, GEDI data and Ground-Truth

Area 1: Corrientes

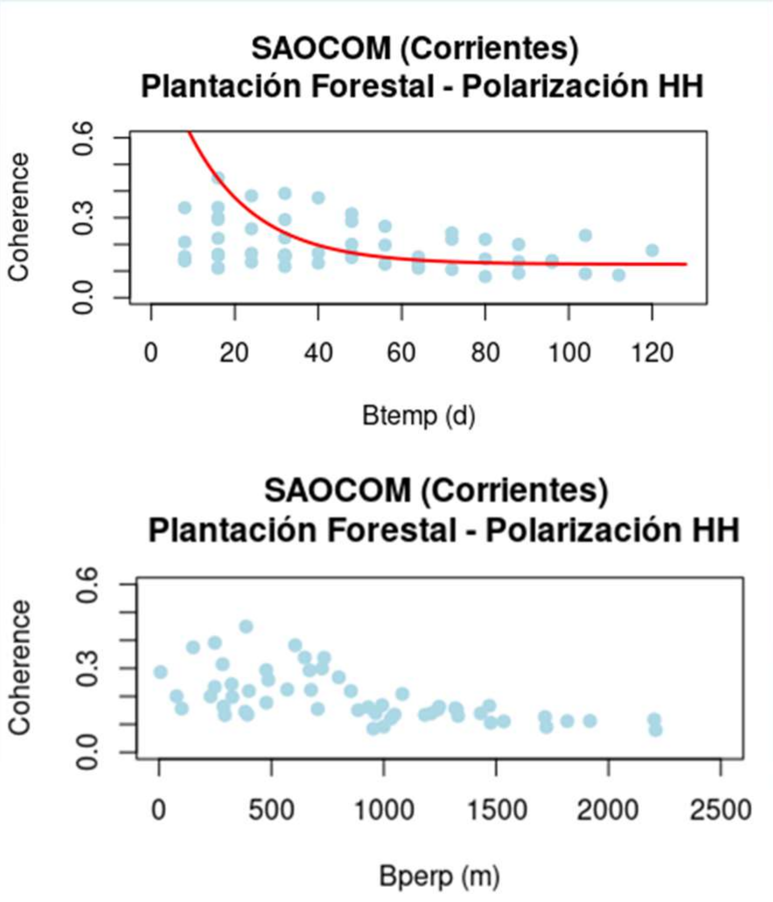


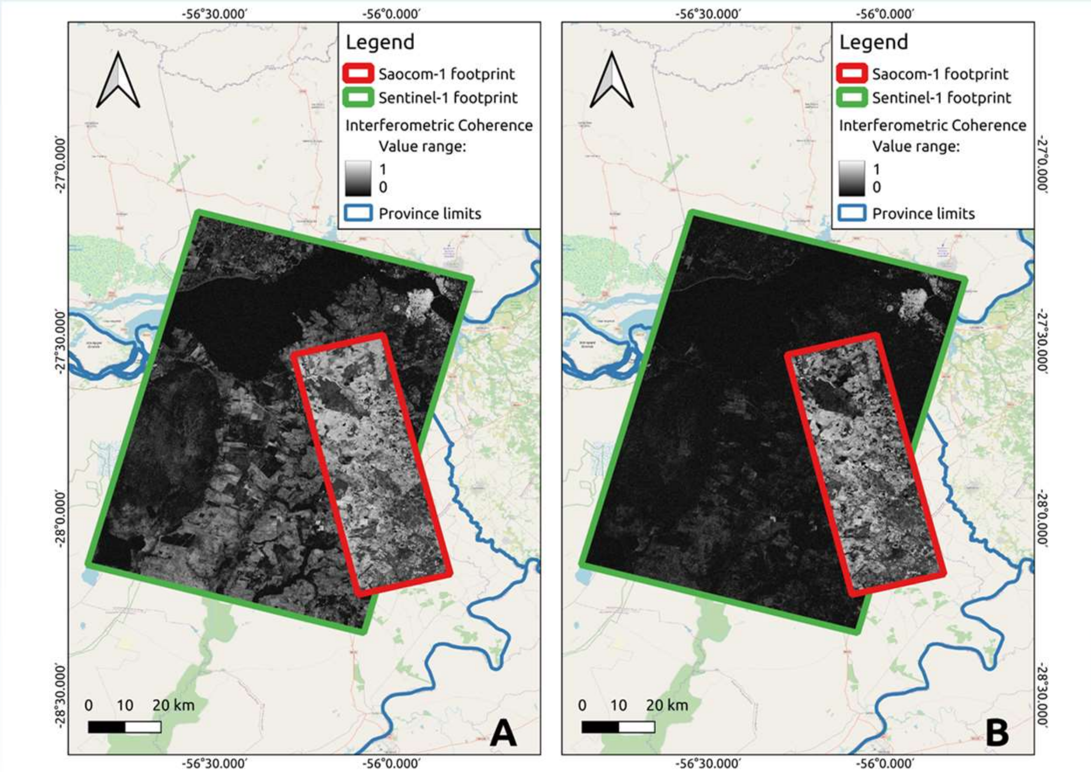
Area 2: Neuquén



Forest Height Estimation based on SAOCOM L-band SAR Data

Interferometric baseline & system effects (C vs L-band)



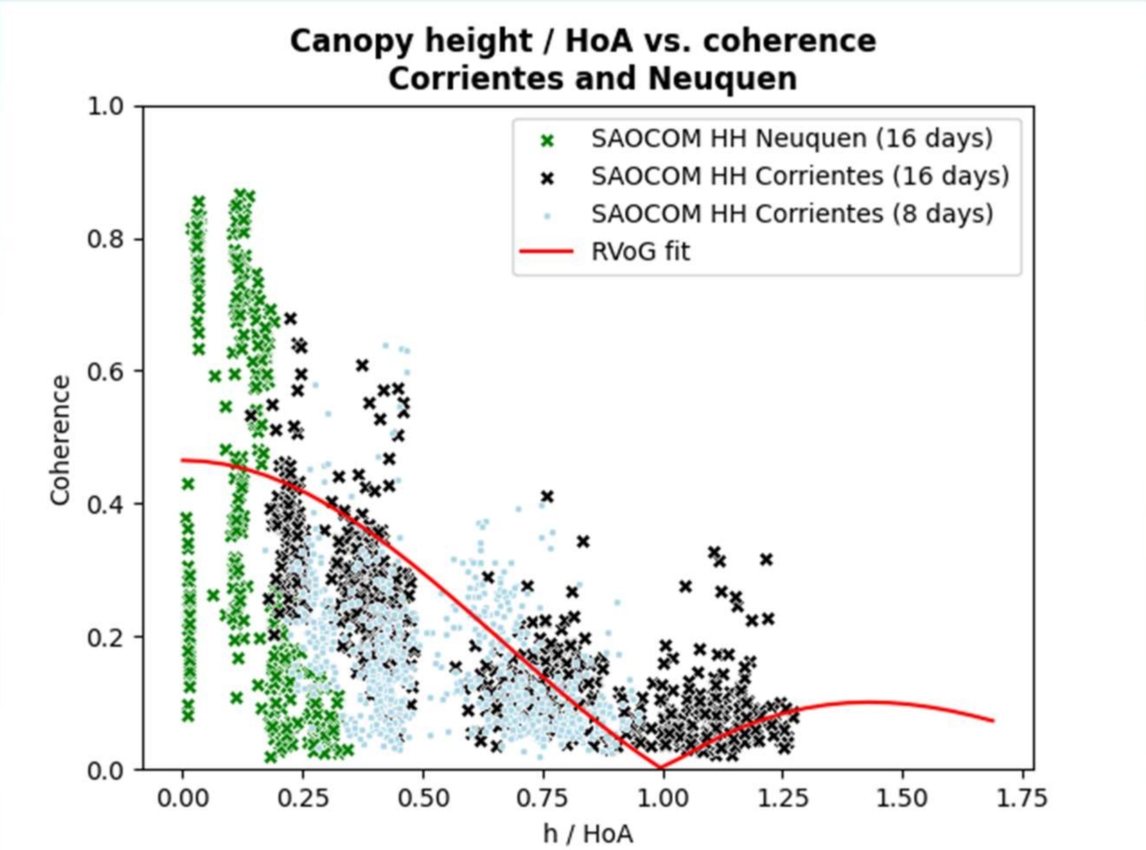


Saocom: 16 días
Sentinel-1: 12 días

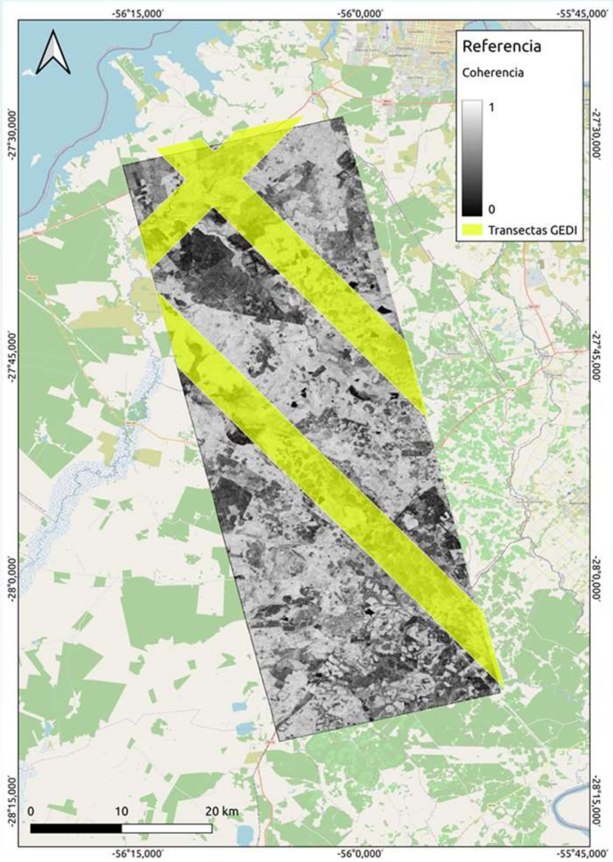
Saocom: 48 días
Sentinel-1: 48 días

Forest Height Estimation based on SAOCOM L-band SAR Data

Coherence vs. GEDI Forest Height



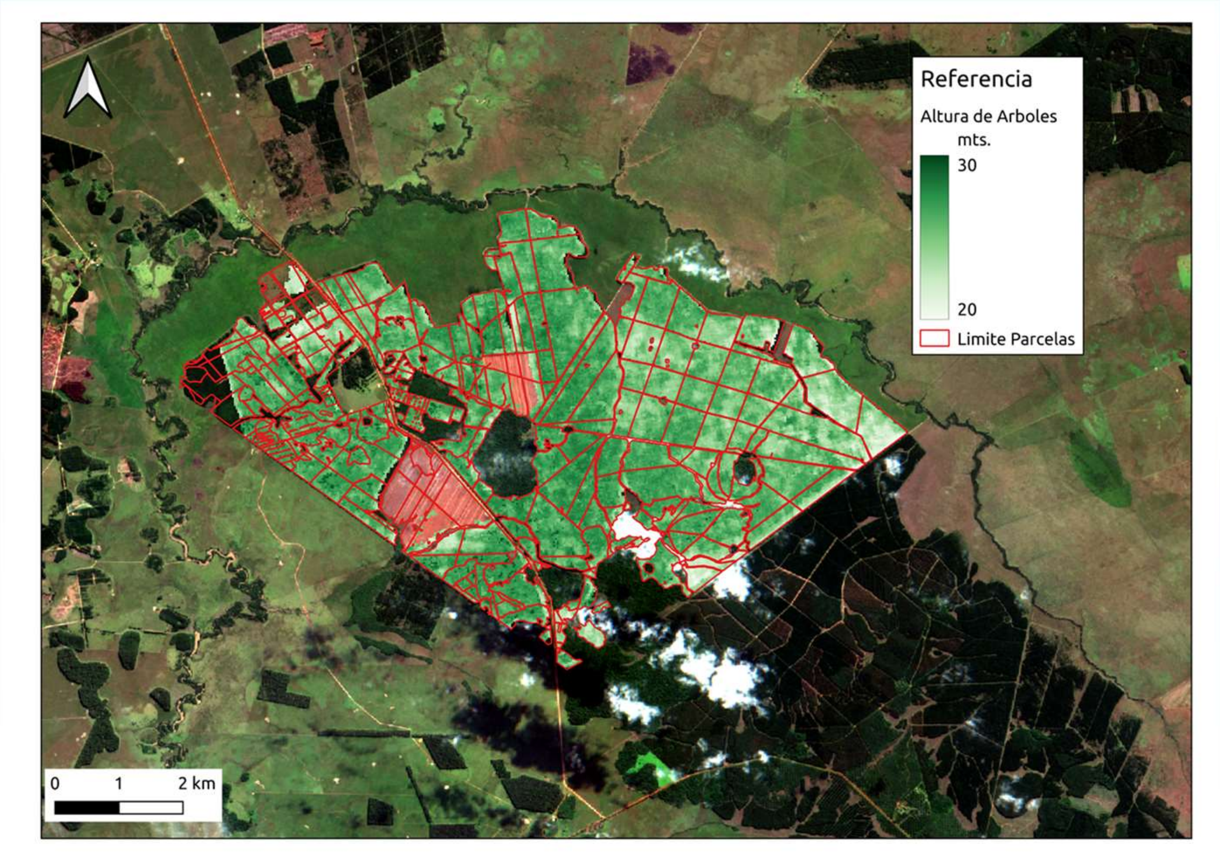
8-day SAOCOM Coherence



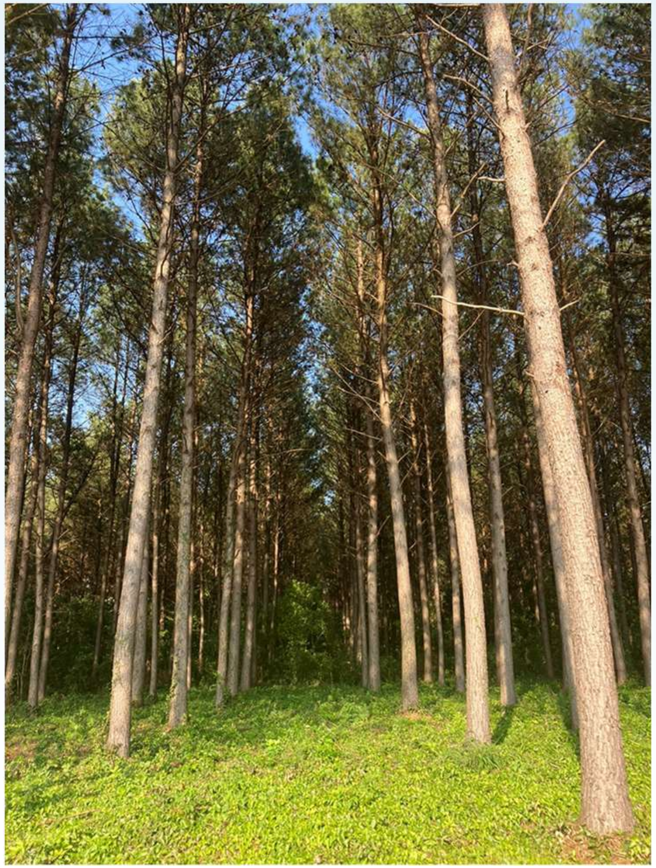
Forest Height Estimation based on SAOCOM L-band SAR Data

First inversion results

Campo Aurora Celeste (BDP) - Corrientes



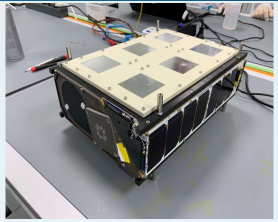
Pinus Taeda - 18 años

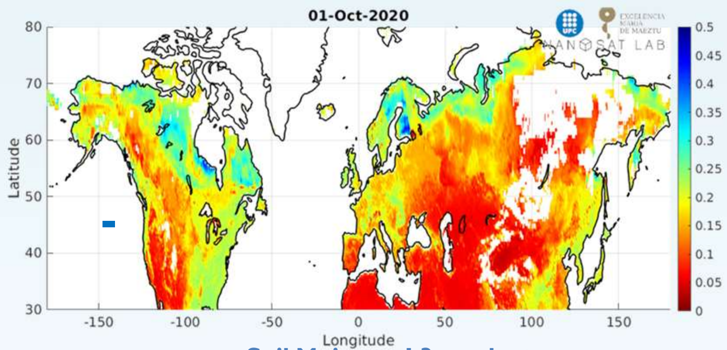


FSSCat Mission (i): SM/SIE/SIC/SIT/SSS retrievals

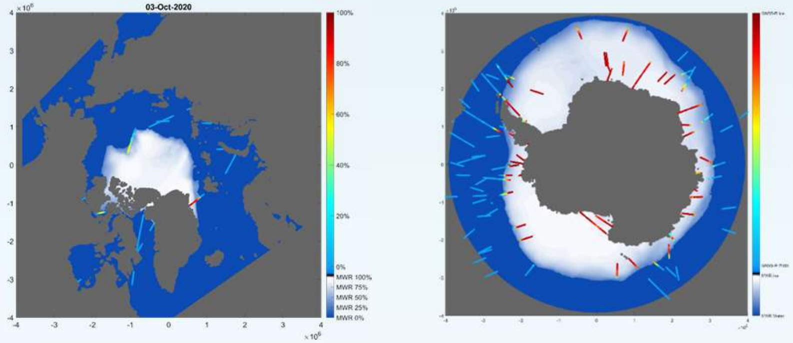
[https://youtu.be/lQAaoYUPluA]

- **FSSCat mission**:
  - [3]Cat-5/A: MWR + GNSS-R (GPS+Gal) + RFI detection/mitigation
  - [3]Cat-5/B: HyperScout-2 (VNIR+TIR Hyperspectral Imager) + AI Proc
  - Both: O-ISL + RF-ISL

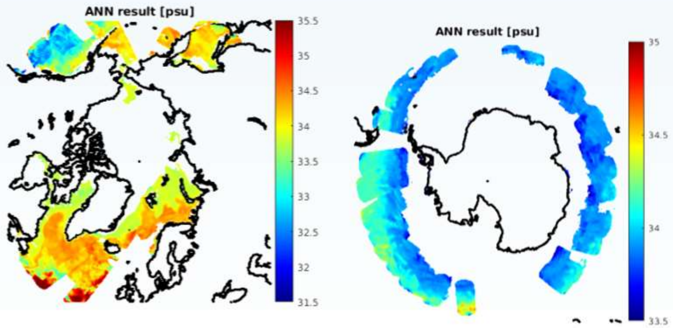




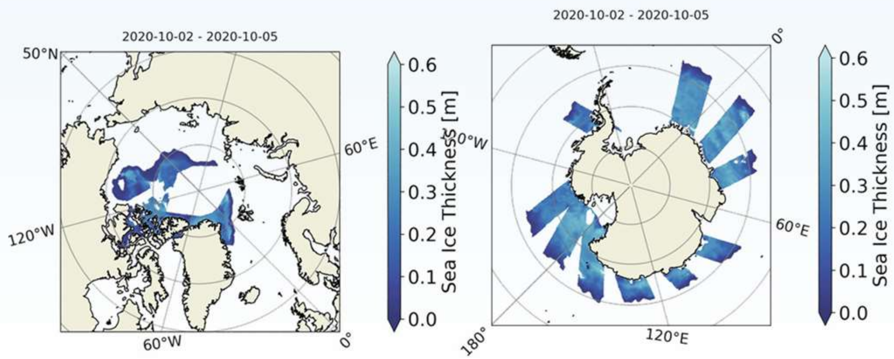
Soil Moisture L3 product



Sea Ice Extent/Coverage L4 product using GNSS-R and MWR maps



Sea Surface Salinity L4 product using GNSS-R and MWR maps



Sea Ice Thickness L4 product using MWR maps

FSSCat Mission (ii): SM/SIE/SIC/SIT/SSS retrievals

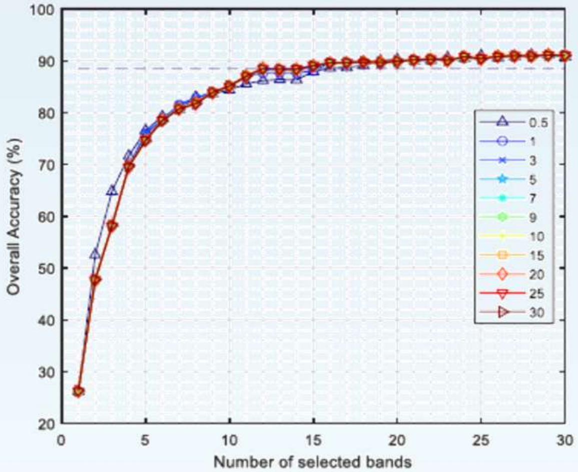
- FMPL-2 onboard ^{3}Cat-5/A simultaneously collected GNSS-R and L-band radiometry data to retrieve:
  - Soil Moisture (4 x ANNs): Optical data only, Optical + L-band MWR data (as in SMOS), GNSS-R data (e.g. NASA CyGNSS), and GNSS-R + L-band MWR
  - Sea Ice Concentration and Extent (2 x ANNs)
  - Sea Ice Thickness (1 x ANN): most difficult to train! Probably because high non-linearities
  - Wind Speed (1 x ANN) and Sea Surface Salinity (1 x ANN)
- All perform well, providing scientific quality data, with a modest budget mission
- ANNs can be applied to retrieve these and other geophysical variables from GNSS-R data where it is difficult to capture the GMF (e.g. vegetation height [1]). However, in some cases they do not outperform classical analysis (i.e. "Natural Intelligence", e.g. snow thickness and sea ice thickness in MOSAIC [2], or altimetry [3]).

[1] J.F. Munoz-Martin *et al*. "Vegetation Canopy Height Retrieval Using L1 and L5 Airborne GNSS-R," in *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022, Art no. 2502405, doi: 10.1109/LGRS.2021.3131263.
[2] J.F. Munoz-Martin, et al. "Snow and Ice Thickness Retrievals Using GNSS-R: Preliminary Results of the MOSAiC Experiment." Remote Sens. 2020, 12, 4038. https://doi.org/10.3390/rs12244038
[3] O. Cervelló-Nogués, et al., "Improved GNSS-R Altimetry Methods: Theory and Experimental Demonstration Using Airborne Dual Frequency Data from the Microwave Interferometric Reflectometer (MIR)." Remote Sens. 2021, 13, 4186. https://doi.org/10.3390/rs13204186,

HyperSpectral Imagery Compression by Sequential Band Selection

GOAL: to reduce the down-link requirements in small sats

Indian Pines Ground Truth



$$S_i = H_i \cdot \prod_1^k \left(1 - \rho_{ij}\right)^w,$$

Looks for maximum entropy and inter-band correlation

Classification using 18 bands out of 220

Error map using 18 bands

Current work: do it automatically using NNs without image georegistration