



Mapping Crops and their Biophysical Characteristics with Polarimetric Synthetic Aperture Radar and Optical Remote Sensing

Part 3: Sen4Stat Open-Source Toolbox for Large Scale Crop Mapping (Theory and Practical)

Pr. Pierre Defourny (UCLouvain)

26 April, 2022

### **Training Outline**

12 April

Part 1: SAR
Polarimetry for
Agriculture (Theory
and Practice)

19 April

Part 2: Practical SAR
Polarimetry with
Sentinel-1, RCM, &
SAOCOM Imagery for
Agriculture

April 26

Part 3 - Sen4Stat Open-Source Toolbox (Theory and Practical) May 3

Part 4: Crop-Specific Time Series Analysis for Growth Monitoring

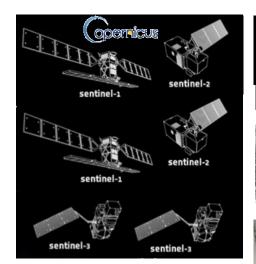
#### **Table of Contents**



- Section 1: Challenges for large scale crop mapping at high spatial resolution
- Section 2: Large scale crop mapping using Sen4Stat toolbox
- Section 3: Practical using Sen4Stat toolbox for crop type mapping
- Section 4: Q&A session



#### New satellite missions and IT (r)evolutions have changed the game for agriculture



























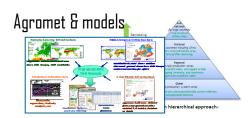




tools using

ML & Al algo.



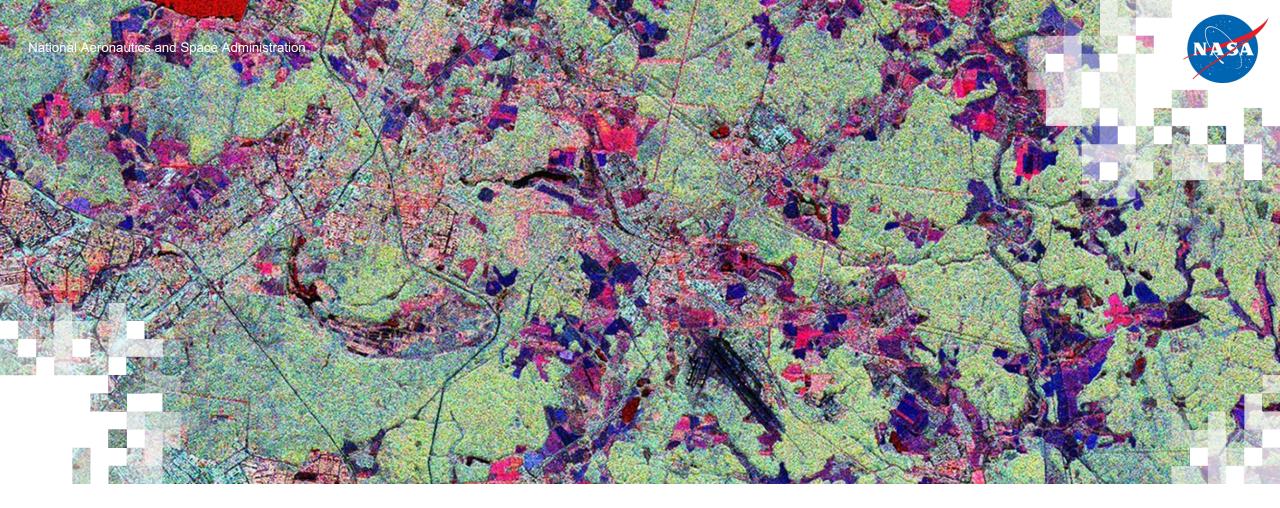


EO ag. products => (intra-) field level

- 1. Cropland mapping
- 2. Crop type mapping
- 3. Biophys. variables (fCover, LAI, fAPAR,...)
- 4. Crop condition monitoring
- 5. Ag. practices monitoring

How to take advantage of Copernicus Sentinel-1 and Sentinel-2 data for crop mapping at national scale?



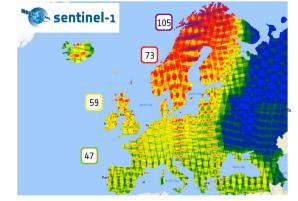




Section 1: Challenges for large scale crop mapping at high spatial resolution

In addition to large data volume and the IT requirements

- Observing satellite systems at high resolution do not cover large area at once
  - Sentinel-2 instrument swath: 290 km
  - Sentinel-1 interferometric wide swath mode: 250 km
  - Overlap between images acquired from adjacent orbit
  - Increasing overlap due to converging orbit to North Pole
  - ⇒ Heterogeneity of observation density across the area



Nbr of observation dates by S-1A & -1B (July-Sept)

- Cloud coverage frequency quite variable (optical systems) and nonsystematic acquisition plan over space and time (optical and SAR systems)
  - ⇒ Heterogeneity of valid observation density across the area
- Significant agro-climatic gradients introducing large variability in crop calendars and vegetation development across the area

## Challenges for large scale crop mapping – why is it so special?

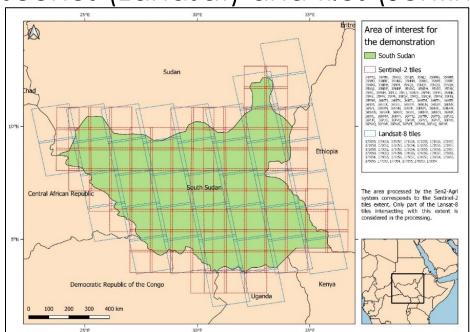
3 challenges already identified

temporal resolution critical for crop discrimination

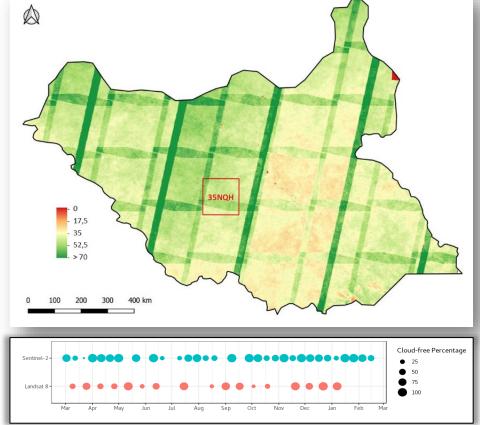
#### **Example of South Sudan** (619.745 km²)

Nbr of granules for satellite data:

scenes (Landsat) and tiles (Sentinels)



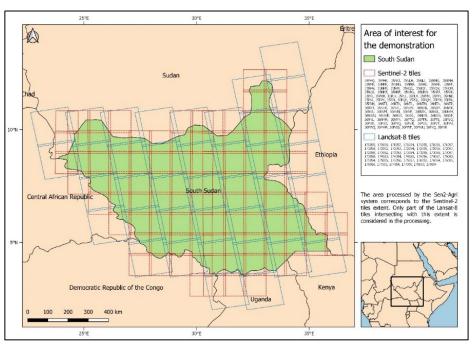
Nbr of valid images (cloud cover < 10%) from March 2017 to Feb. 2018



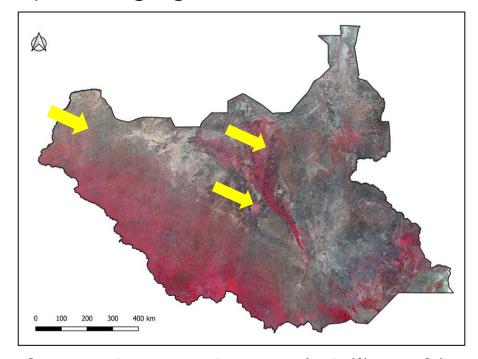
-072

Example of South Sudan – first solution to cope with heterogeneous valid observ. density

Nbr of granules for satellite data: scene (Landsat) and tile (Sentinels)



**Monthly synthesis** (June 2017) by averaging all cloud free reflectances

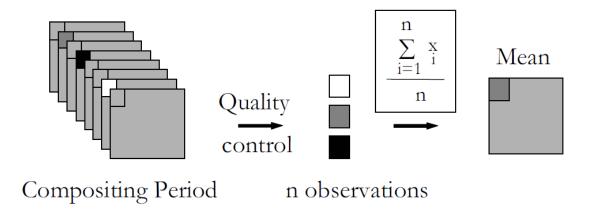


- => artefacts due to the variability of input data
- => reduction of temporal resolution to monthly





#### Temporal synthesis methods averaging all cloud free reflectances acquired during the period



#### Sen4Stat: Weighted Average Synthesis Processor

- weight as function of date (central date distance)
- weight as function of Aerosol Optical Thickness
- weight as function of distance to clouds
- weight as function of sensor (S2>L8)
- => reduction of the compositing artefacts
- => monthly temporal resolution not really suitable for crop discrimination

Comparison of temporal synthesis methods

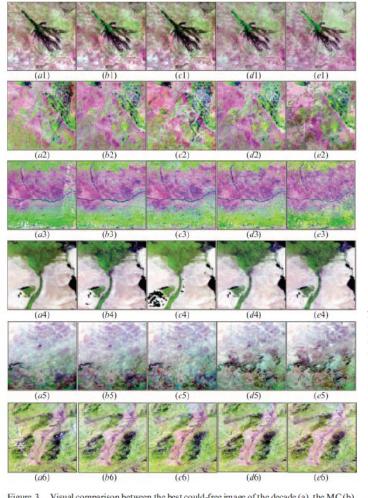
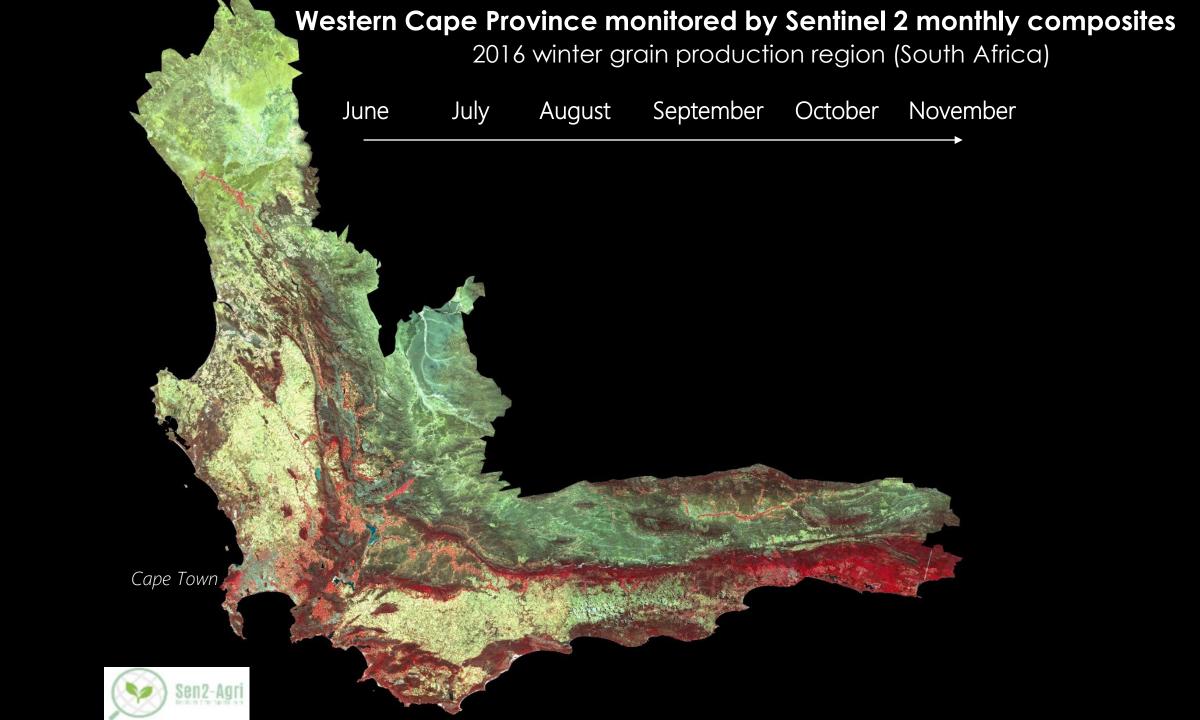
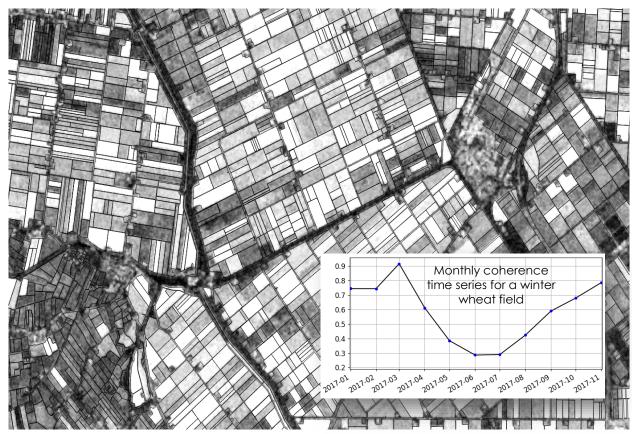


Figure 3. Visual comparison between the best could-free image of the decade (a), the MC (b), the BDC (c), the AVG (d) and the MVC (e) colour composites (SWIR, NIR, red), for 6 samples: (1) Botswana (1st decade of June 2002), (2) Sudan (1st decade of November 2002), (3) Romania (2nd decade of June 2002), (4) Egypt (1st decade of November 2002), (5) Senegal (1st decade of June 2002), and (6) Germany (2nd decade of June 2002). Available in colour online.



#### SAR temporal synthesis by averaging all SAR observations over the period for noise reduction



Mean of the coherence values as derived from Sentinel-1 interferometric pairs for March 2017 over Netherlands



Sentinel-1 color composite (blue: mean of the July coherence; green: mean of the March coherence; red: seasonal standard deviation)



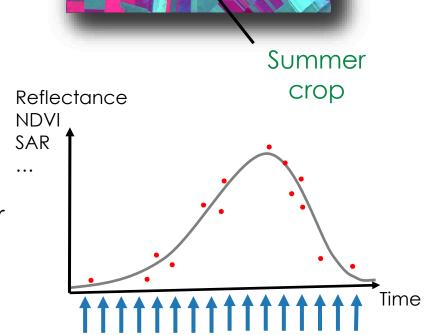
Temporal information very critical for crop discrimination Second solution to cope with heterogeneous obs. density over large area





Production of an optical or SAR time series which is:

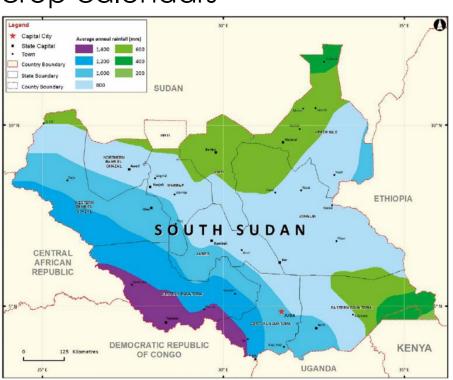
- gap-filled using a linear interpolation method with respect to missing data (clouds, cloud shadows, no data, ...)
- temporally resampled at regular interval
- ⇒ This approach provides the same "observation" dates over the whole area independently from the sensor orbits and cloud cover which is very critical to run most of the classification methods.





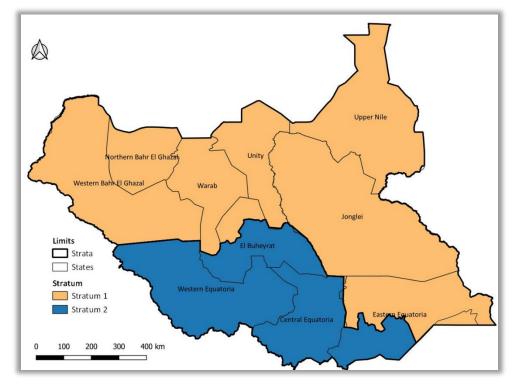
#### **Example of South Sudan**

Major rainfall gradient across the country defining two different crop calendars





#### Stratification in zones to be classified separately

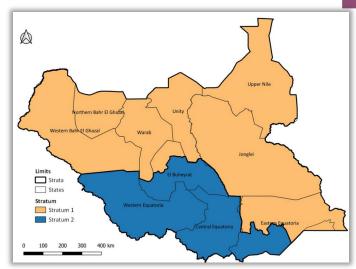


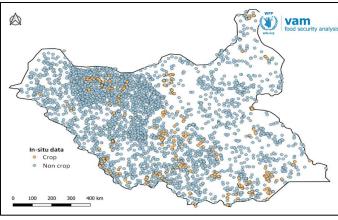
#### Example of South Sudan – stratification

A model to calibrate for each stratum requires collection of enough ground samples for each zone covering all crop types but also all non-cropland classes.

Crop / Non crop	Land cover class	Nr of samples	Proportion in crop / no crop categories	Proportion in total
Non crop	Bush	1504	44%	33%
	Forests	566	17%	12%
	Grasslands and meadows	565	16%	12%
	Shrubland	408	12%	9%
	Wetland	93	3%	2%
	Water bodies	78	2%	2%
	Fallow	76	2%	2%
	Built-up surface	71	2%	2%
	Bare soil	64	2%	1%

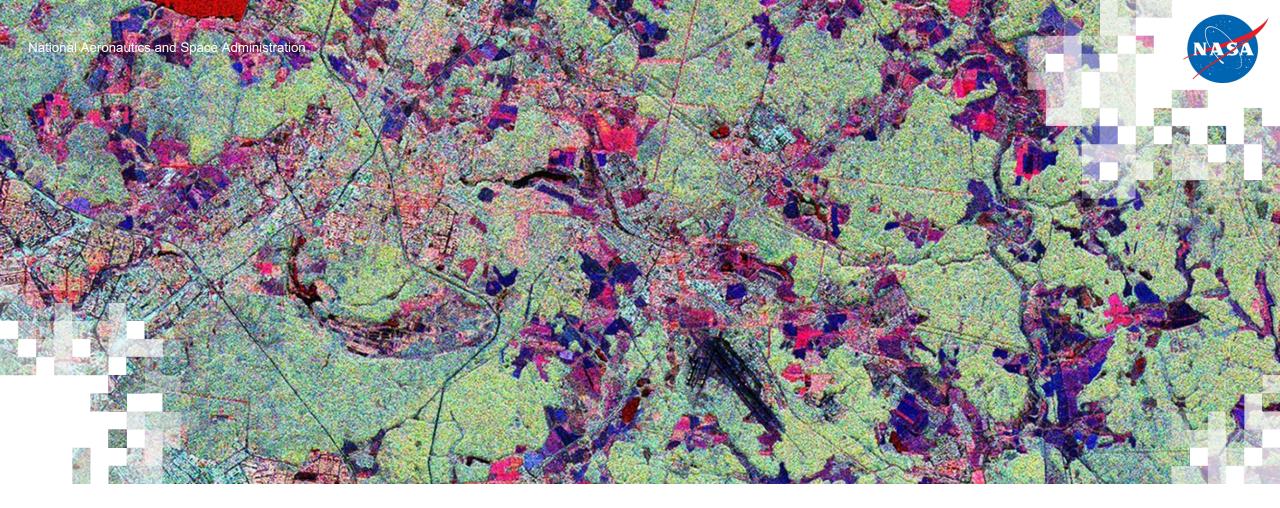
Crop type		Nr of samples		Proportion in crop / no crop categories	n in total	
Early Sorghum		369		31%	8%	
Late Sorghum	244			21%	5%	
Maize	122			10%	3%	
Groundnut	116		10%		3%	
Early Sorghum Mixed	105		9%		2%	
Cassava	79		7%		2%	
Late Sorghum Mixed	5	55		5%	1%	
Maize Mixed	4	8		4%	1%	
Sesame	1	3		1%	0%	
Rice	1	0	1%		0%	
Bulrush Millet 9		9		1%	0%	
Beans	7		1%		0%	





=> 4602 samples well distributed







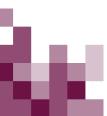
Section 2: Large Scale Crop Mapping Using Sen4Stat Toolbox



### ESA open-source SAR and optical toolbox for operational crop type mapping and monitoring over very large areas



- Process Sentinel-1, Sentinel-2 and Landsat-8&9 time series according to the stateof-the-art including advanced SAR products (coherence, gamma naught,...)
- Deliver automatically or on request 5 types of products (processors) in near realtime along the satellite data acquisition or off-line:
  - 1. 10m optical cloud free temporal synthesis and SAR temporal synthesis
  - 2. Time series of spectral indices (NDVI, coherence,...) and biophysical variables (LAI, fCover, fAPAR)
  - 3. 10m crop type maps along the season based on in situ dataset and stratification
  - 4. Large set of crop growth conditions metrics (including meteorological data)
  - 5. Crop yield estimation at various aggregation levels (national, regional, ...)

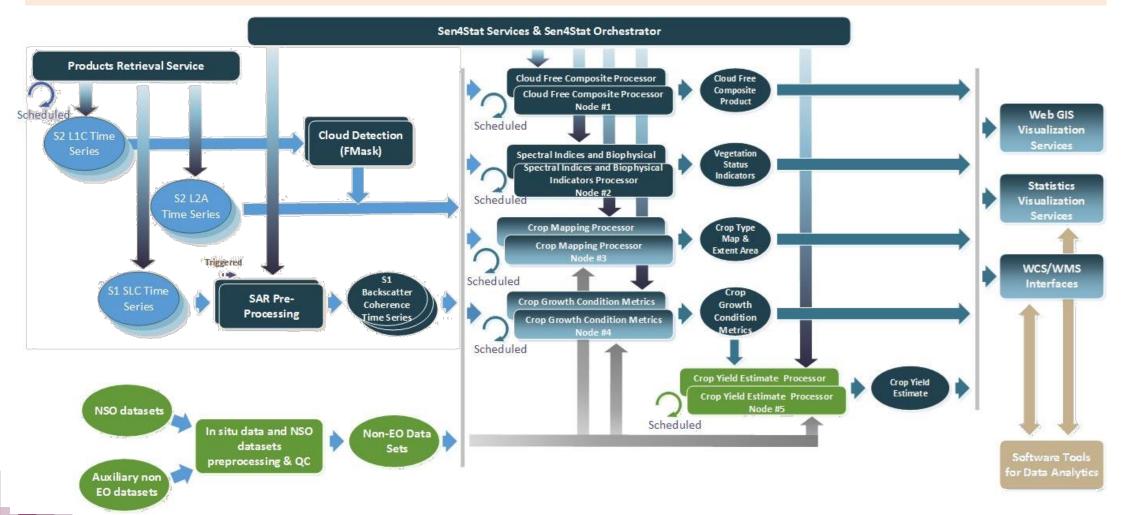




### Sen4Stat - ESA open-source SAR and optical toolbox for operational crop mapping and monitoring over very large areas

Data access /download

**Pre-processing** In situ quality control **Processors** of products **Analytics and** visualization















### ESA open-source SAR and optical toolbox with a graphical user interface to configurate the system to launch the production and monitor the processes



System configuration

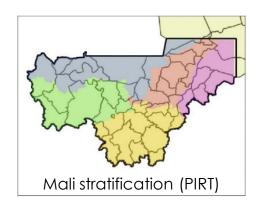
Start of the season

End of the season.





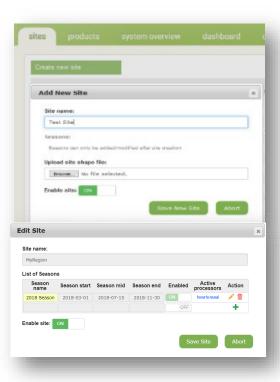
Sen4Stat:	en4Stat: parameters settings			
Area of Interest	Shapefile to be uploaded			
Monitoring period	Start and end dates to be defined			
S1 - S2 - L8&9	To be selected			
Data sources	ESA&USGS – AWS – DIAS – local storage			





Sen4St	Sen4Stat: tiela campaign			
Sampling design	Stratification and sampling			
Field visit	In situ data collection – early survey			
	In situ data collection – mid-season survey			
Data upload	Field data quality control and formatting			











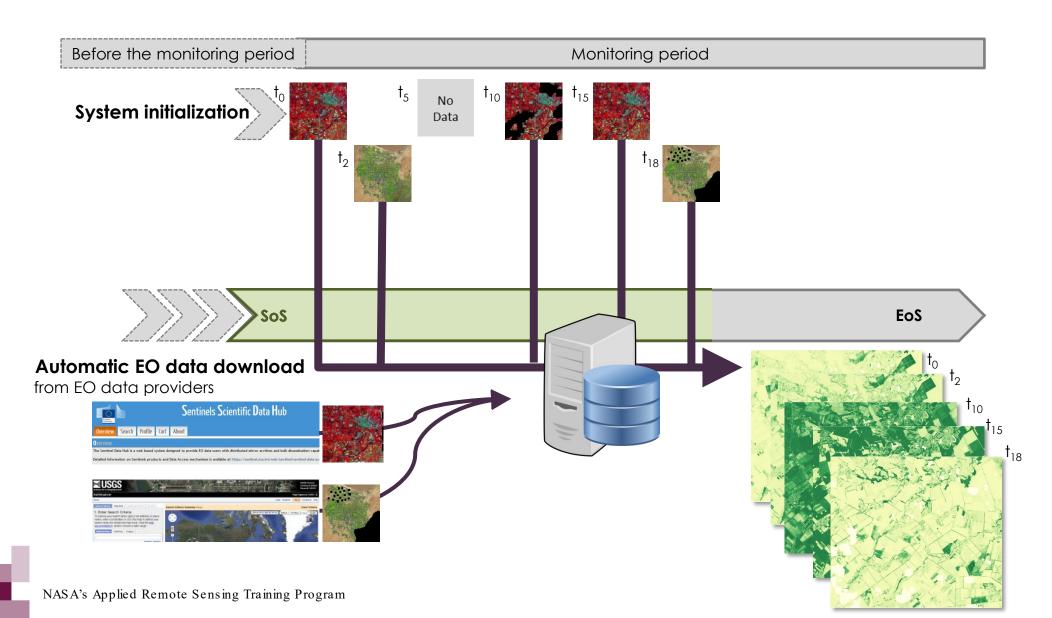






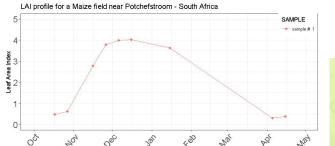
# Sen4Stat – system operation for crop growth monitoring:

Leaf Area Index (LAI) production in near-real time along the season

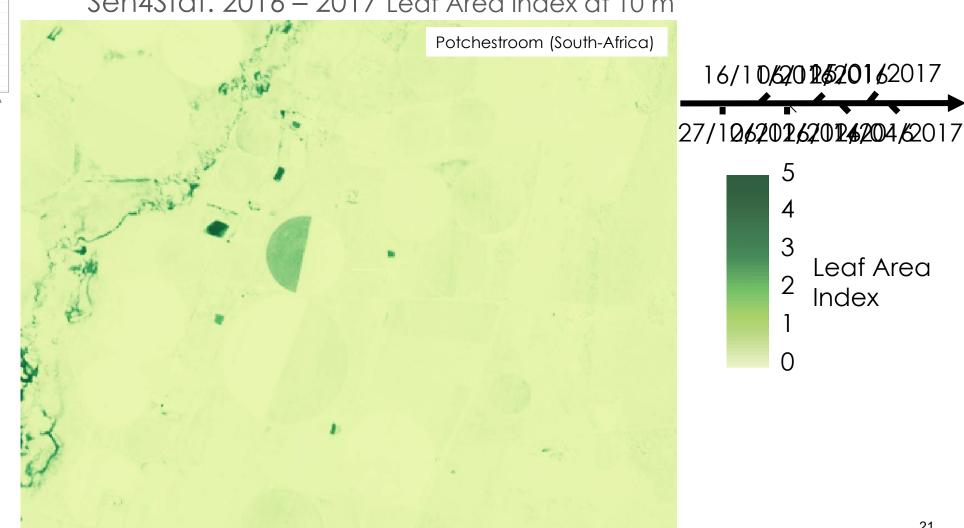




#### Sen4Stat - System operation for crop growth monitoring: Leaf Area Index (LAI) production in near real-time along the season







Leaf Area

Index







#### Automated mode through the web graphical user interface (GUI)

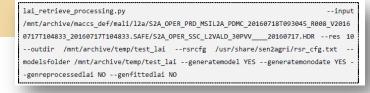
- a) based on the Orchestrator with by-default parameterization, automatic data download and processing until the end of the season, on-time delivery => operational production in near real time (NRT)
- b) Processor execution on user request, with by-default parameterization



#### Manual mode: to run processor independently, with custom parameters

- through the GUI, with the Custom job approach
- through SNAP software (only processor of Level 3 and 4)
- in command line through a Linux terminal









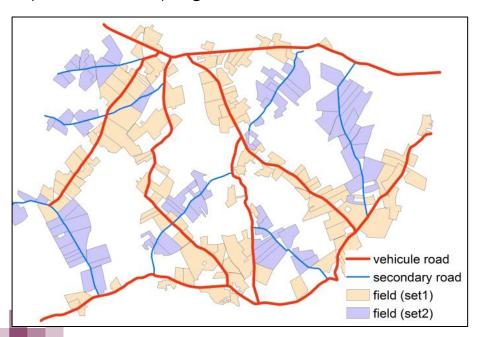


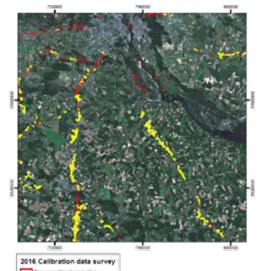




#### Sen4Stat – In situ data collection required for annual crop type mapping

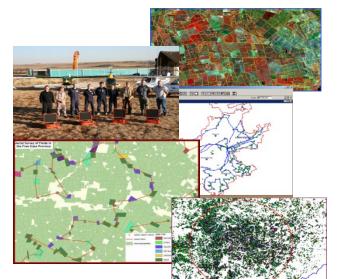
- Available reference dataset collected according to an area or list sampling frame (national statistical surveys, Ministry of Agriculture,...)
- JECAM Guidelines ad hoc "windshield SUrvey" by motorized vehicles (car, motorbike) selecting a set of appropriate roads (set 1) and complemented by regular additional transect (set 2) using secondary roads and tracks to try to reduce the spatial bias brought about by roadside sampling.











Cropland sample



VIA used for aerial observations





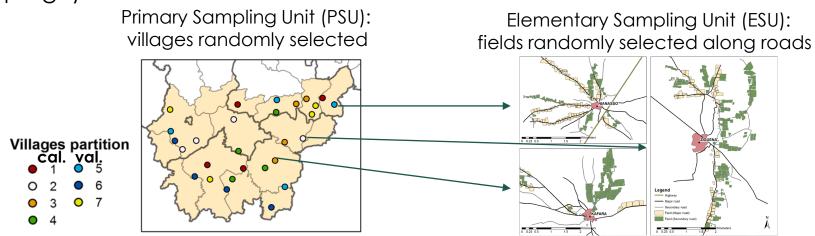




#### Sen48tat – Requirements for the in situ data collection



- In situ data for **calibration** (**training**): sampling to cover the diversity of situations existing in the study area (possibly a national territory) in order to cover the range of possible signatures for the different elements of interest (i.e. most crop types and the main non cropland classes).
- In situ data for validation to estimate the products accuracy (with a confidence interval) using a statistically-sound sampling to be objective and independent; for logistic reasons, sampling not strictly random but 2-stage sampling (with PSU and ESU) to assess the crop types (one field campaign)



⇒ calibration and validation field campaigns for crop type mapping can be combined but the sampling design should be statistically sound to obtain fully independent dataset.

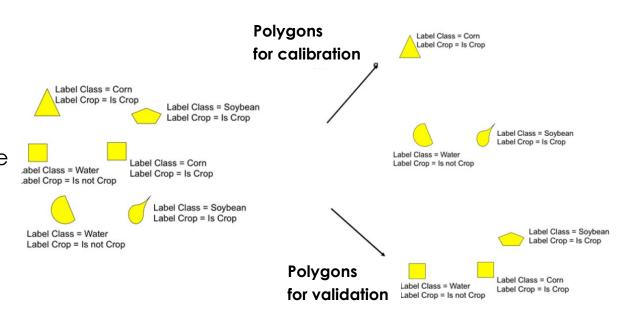






#### Sen48tat – In situ data quality control process and dataset preparation

- High quality reference data collection:
  - Crop type acquired from field campaigns
  - Non-crop data obtained by visual interpretation of very high spatial resolution imagery available online
- Quality control: geometry consistency, typology standardisation, homogeneity check...
- 3) Splitting polygons for calibration (25%) and validation (75%) for each class (alternative strategies might be required in case on poorly balanced in situ dataset)
- 4) Enhancing the samples number for marginal crop type using Synthetic Minority Oversampling Technique (SMOTE)





#### **Sen4Stat** – Accuracy assessment using four metrics



•The **Overall Accuracy** (OA) is calculated as the total number of correctly classified pixels  $(n_{ii} \sim \text{diagonal elements of the confusion matrix})$  divided by the total number of test pixels:

$$OA = \frac{\sum_{i=1}^{r} n_{ii}}{\sum_{i=1}^{r} \sum_{j=1}^{r} n_{ij}}$$

• **Precision** or **User's Accuracy** (UA) for the class *i* is the fraction of correctly classified pixels with regard to all pixels classified as this class *i* in the classified image:  $UA_i = \sum_{j=1}^r \frac{n_{ii}}{n}$ 

• **Recall** or **Producer's Accuracy** (PA) for the class *i* is the fraction of correctly classified pixels with regard to all pixels of that ground truth class *i*:

$$PA_i = \sum_{j=1}^r \frac{n_{ii}}{n_{ji}}$$

		Refer	ence Data	а				
	W	S	F	U	С	Н	Row Tota	
Classification data								
W	226	0	0	12	0	1	239	
S	0	216	0	92	1	0	309	
F	3	0	360	228	3	5	599	
U	2	108	2	397	8	4	521	
C	1	4	48	132	190	78	453	
H	1	0	19	84	36	219	359	
Column total	233	328	429	945	238	307	2480	
Producer's Accuracy				User'	s Accura	acy		
W = 226/233 = 97%				W = 226/239 = 94%				
S = 216/328 = 66%				S = 216/309 = 70%				
F = 360/429 = 84%				F = 360/599 = 60%				
U = 397/945 = 42%				U = 397/521 = 76%				
C = 190/238 = 80%				C =	190/453	= 42%		
H = 219/307 = 71%				H = 1	219/359 :	= 61%		
Overall accuracy $= (220)$	6 + 216	+ 360 +	397 + 190	+ 219)/	2480 = 6	5%		

W, water; S, sand; F, forest; U, urban; C, corn; H, hay.

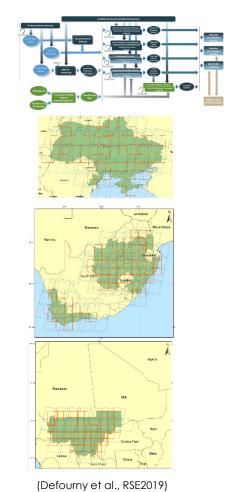
F-Score (also known as F-1 score or F-measure) is the harmonic mean of the Precision and Recall and reaches its best value at 1 and worst score at 0:

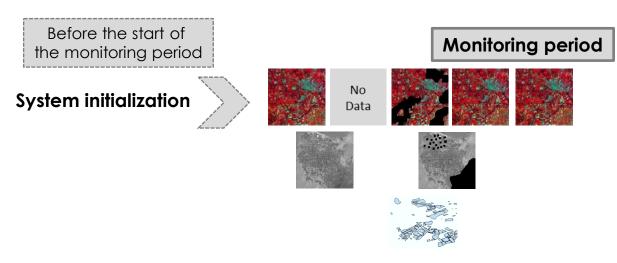
$$FScore = 2x \frac{Precison * Recall}{Precision + Recall}$$



### Sen4Stat – System operation in automated mode to deliver in NRT 10-m crop type maps at national scale

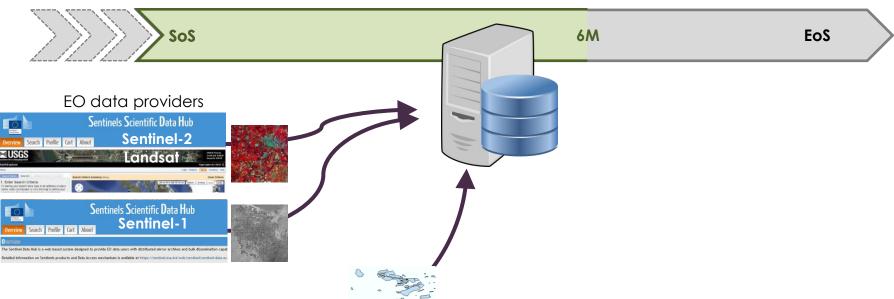






**Operators** 

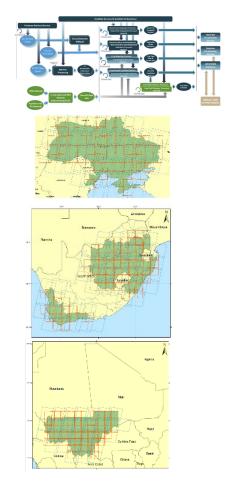
- Cloud computing or automatic EO data download
- Manual in situ data upload

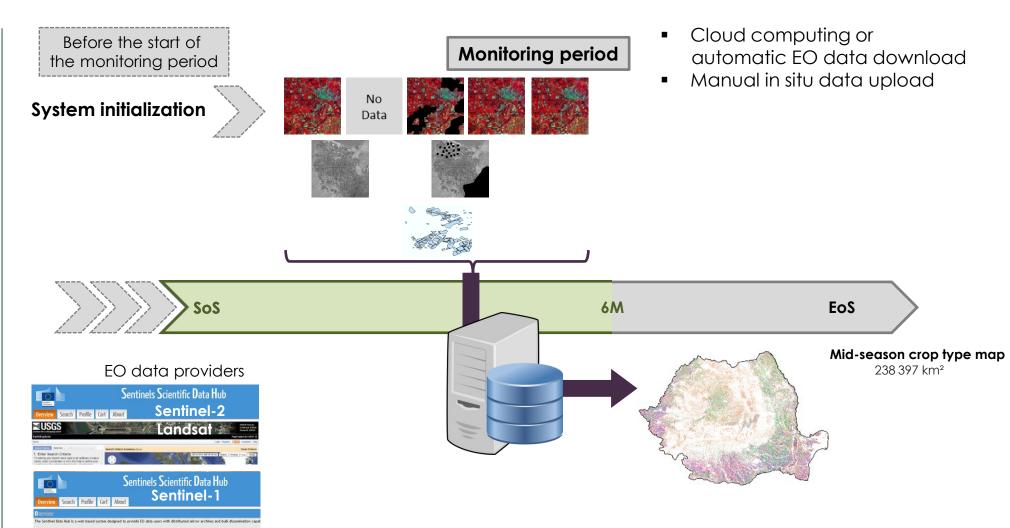




### Sen4Stat – System operation in automated mode to deliver in NRT 10-m crop type maps at national scale



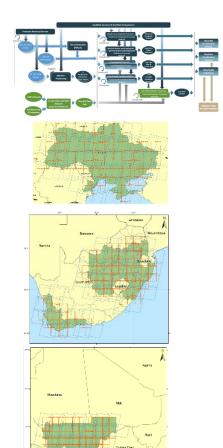


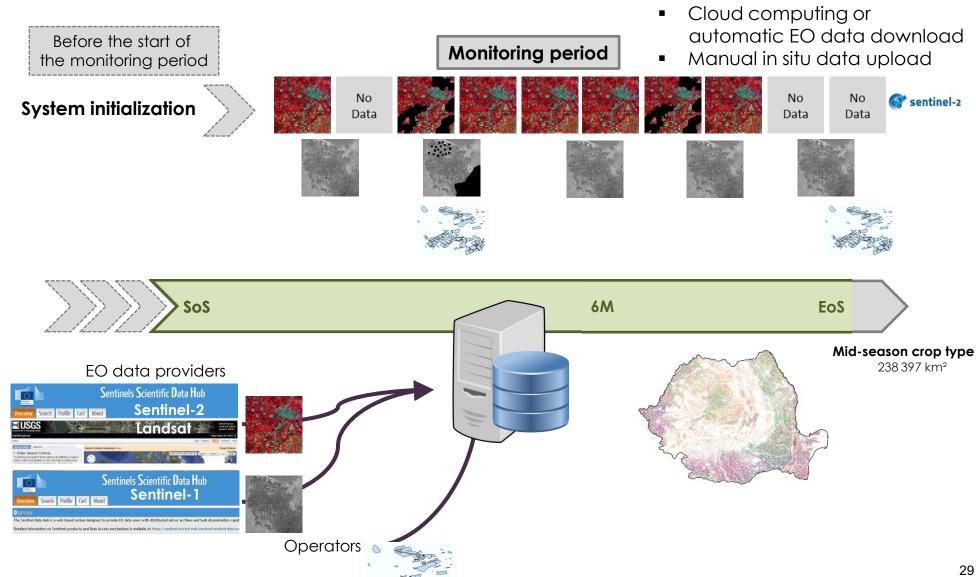




#### Sen4Stat – System operation in automated mode to deliver in NRT 10-m crop type maps at national scale







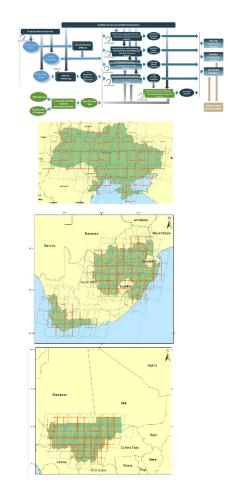
UCLouvain

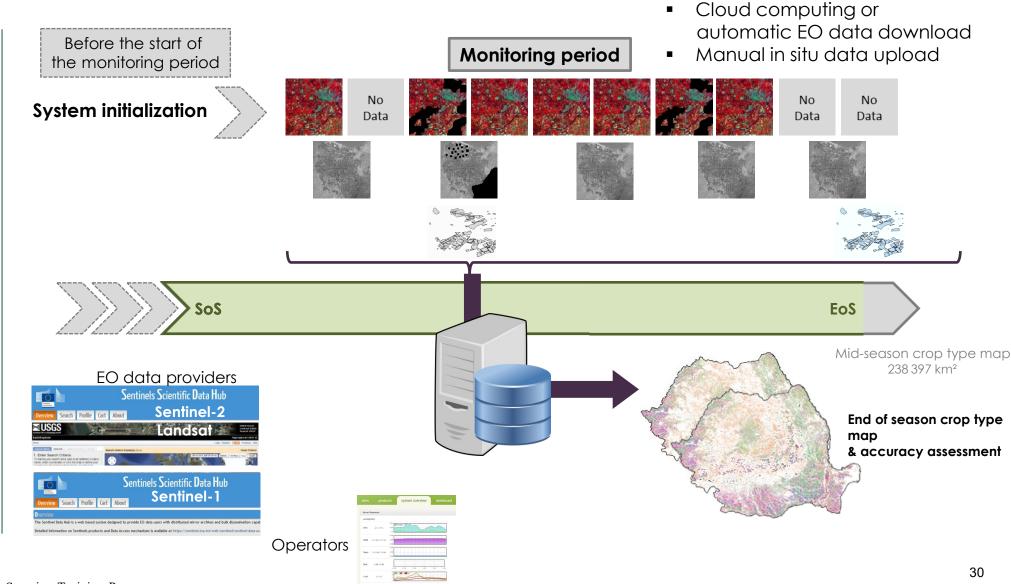


# Sen4Stat – System operation in automated mode to deliver in NRT 10-m crop type maps at national scale



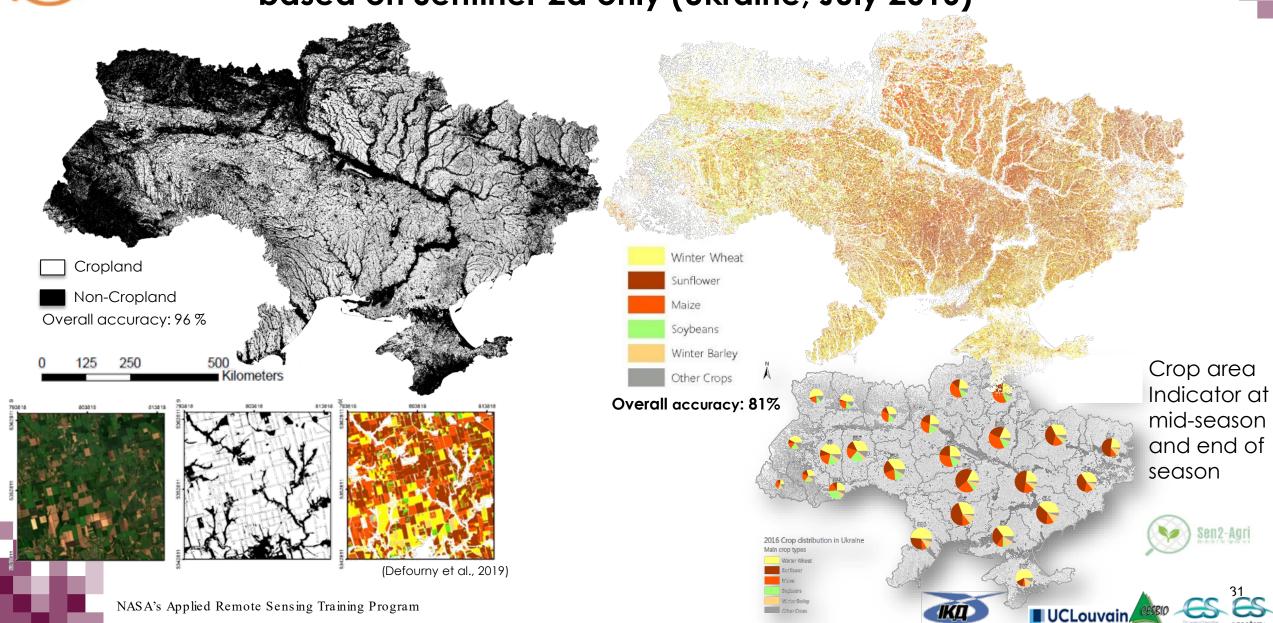
UCLouvain







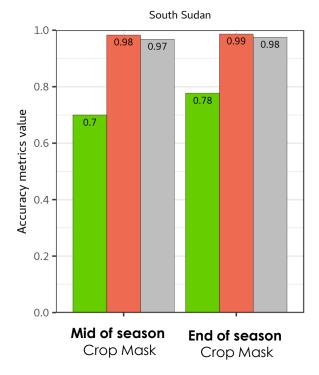
Sen4Stat – Very first 10-m national cropland and crop type maps based on Sentinel-2a only (Ukraine, July 2016)



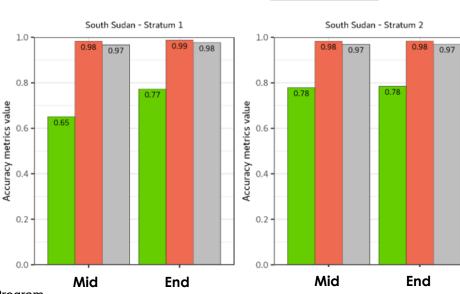
Sen4Stat – Cropland map at mid and end of the season for the

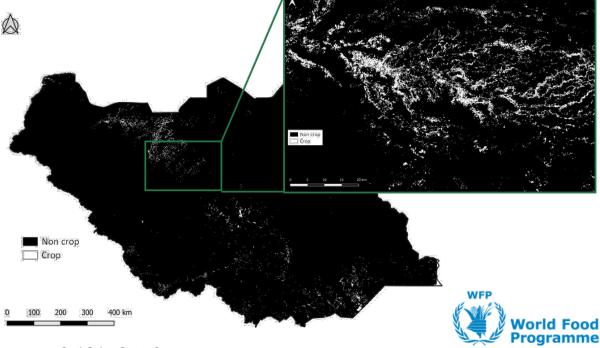
two strata of South Sudan

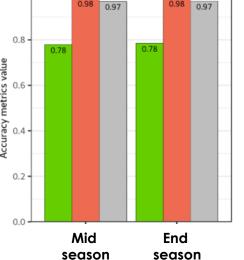
F1-Score of cropland increases with the amount of EO data available along the season









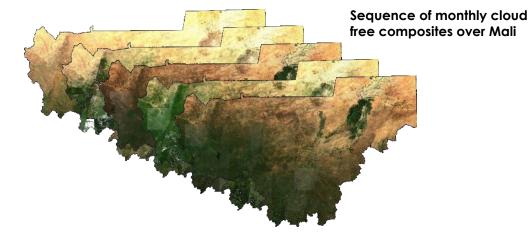




#### - Crop mapping in Mali (2017)

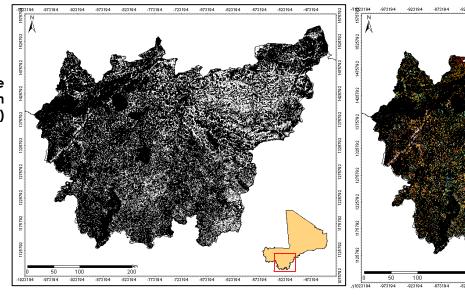
- Successful crop mapping at 10m resolution in collaboration with ICRISAT
- ~4.5 TB of Sentinel-2 data (2152 S2 and 542 L8 images)

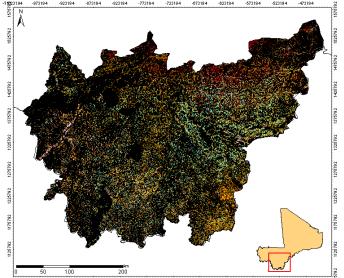




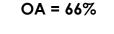
Cropland mask for the 2017 growing season (end-of-season)

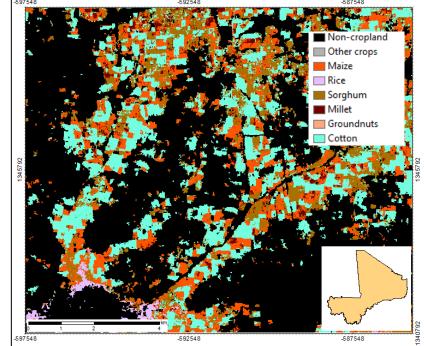
OA = 96%

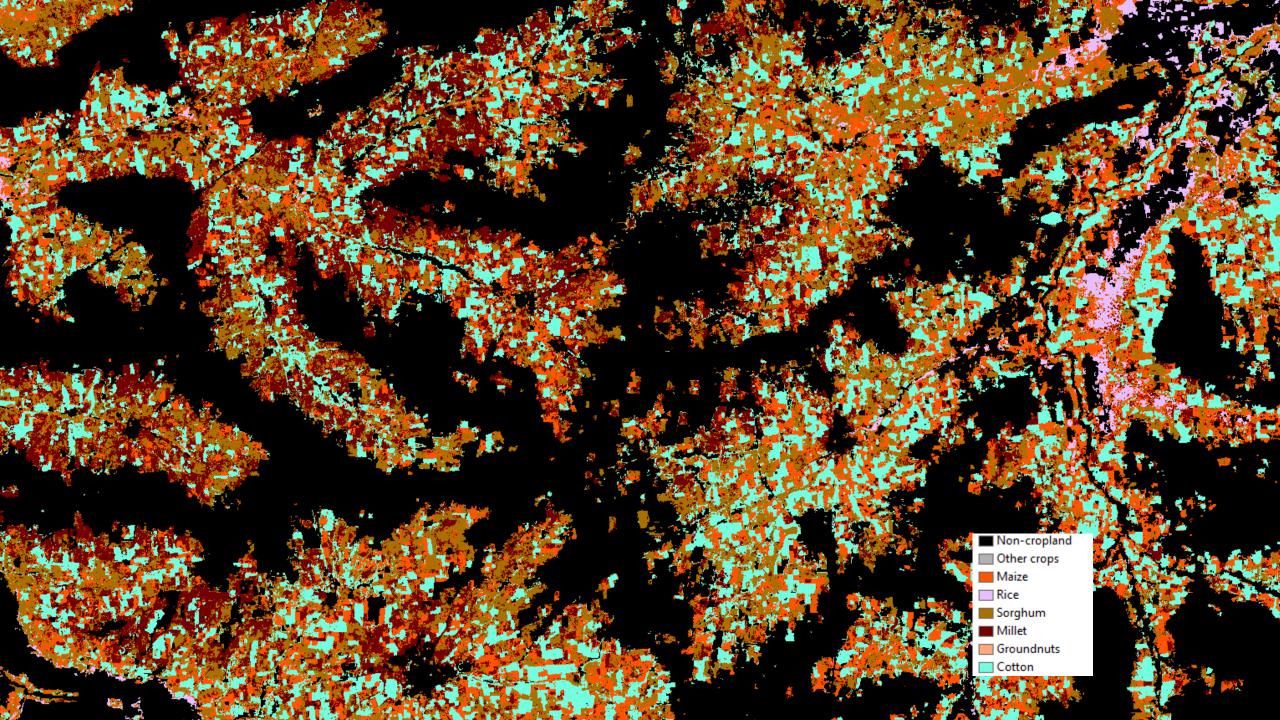




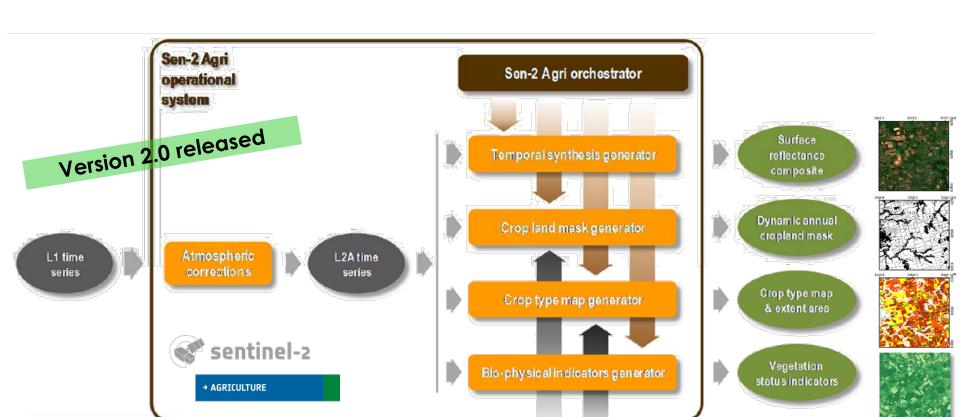
Crop type map for the 2017 growing season (end-of-season)







#### Sen2Agri system: Open source system precursor of Sen4Stat running in NRT or off-line, running locally or in the cloud



Monthly cloud free surface reflectance composite at 10-20 m

Cropland mask at 10m updated every month Crop type map at 10 m for the main regional crops

> **Vegetation status** map at 10 m delivered every



Sen2Agri as a service on the **EODC** cloud (e.g.for WFP)



In situ

information





Documented and downloadable at www.esa-sen2agri.org/resources/software/





# Sen4Stat - Open-source operational solution for national processing in near real time and for local applications



Under GNU-GPL License



Open-source software based on Orfeo ToolBox



Cluster-ready architecture for distributed processing



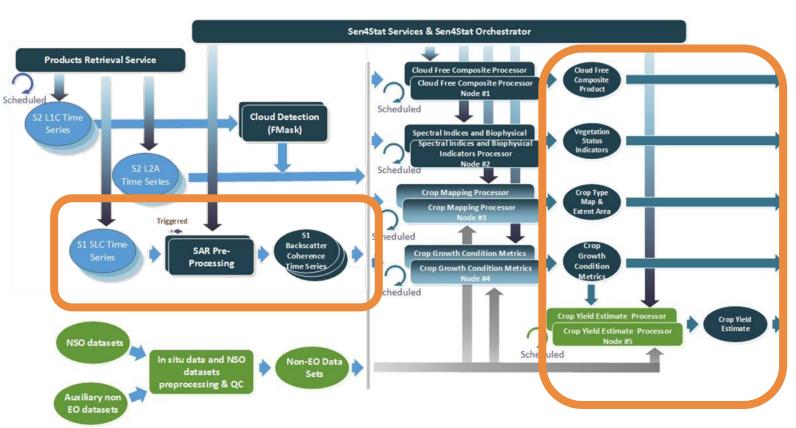
Integration with Sentinel-2 ToolBox



Operational system required: CentOS 7 (GNU/LINUX)



5 main processors to be possibly combined



State-of-the-art SAR processing for high level products









- State-of-the-art SAR processing for high level products

for amplitude and coherence

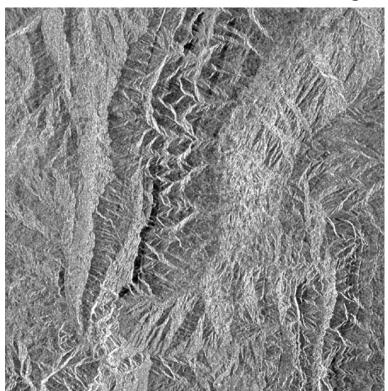


## **beta** nought $\beta_0$

(  $A_{\beta}$  slant range geometry)

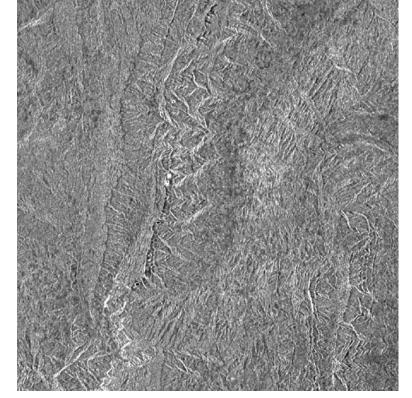


## **sigma** nought $\sigma_n$



## gamma nought

( $\frac{A_0}{A_0}$  corrected for local inc. angle) ( $\frac{A_0}{A_0}$  flattened corrected using DEM)







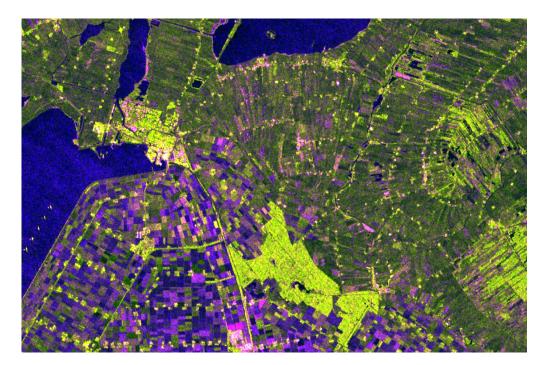
## **Sen4Stat** - Crop type mapping module (S4S\_Crop mapping processor)

- Random Forest as classifier by default (deep learning alternative: Neural Net [Transformer])
- Specific quality control of in situ dataset to balance the calibration dataset
- Large set of features used by the Random Forest model including:
  - S2 & L8 surface reflectance time series
  - NDVI, NDWI, Red Edge Indices, brightness

and \$1 amplitude and coherence time series and \$1 polarization ratio time series

Temporal metrics from above time series and \$1 temporal metrics from \$AR time series







## Sen4Stat - Random Forest (RF)



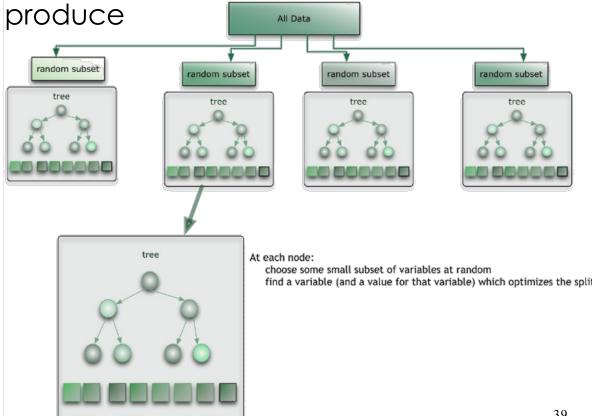
Random Forest ~ ensemble-learning algorithm that combines

multiple classifications of the same data to produce

higher classification accuracies

### Sequence:

- Bagging: individual trees are grown from different subsets of training data (random subsampling of the original data for growing each tree)
- Single majority vote to identify the most frequent class decided from all trees



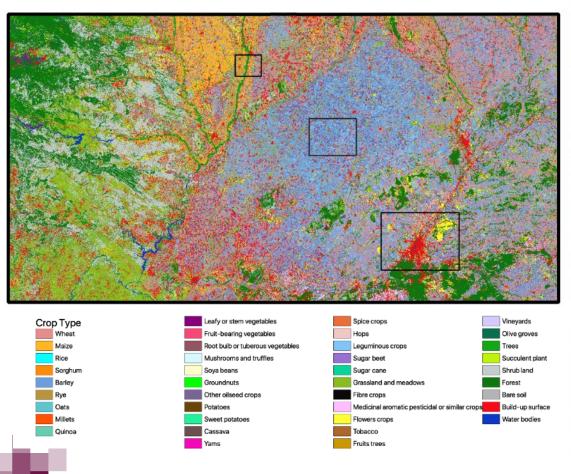
UCLouvain

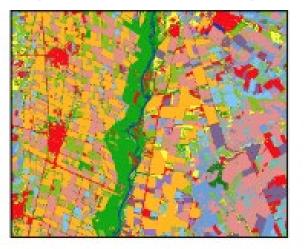


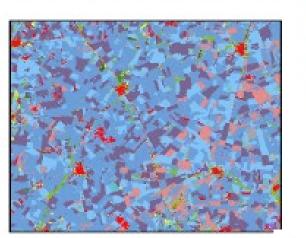
## Sen4Stat - Crop map for Spain area (35 crop types)

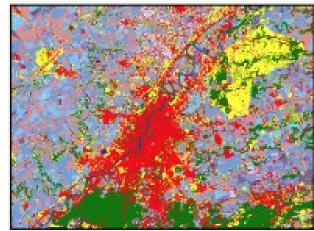
Very similar accuracies for Random Forest and Neural Net (Transformer)

Overall Accuracy (OA): 80% (crop type) & 88% crop group

















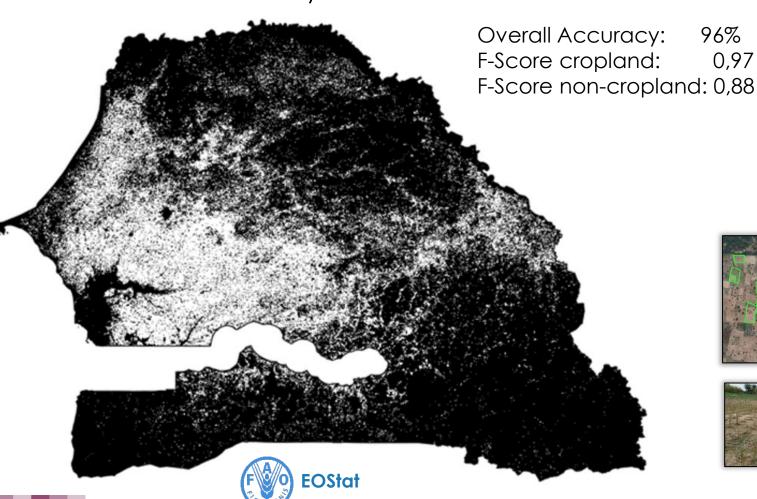


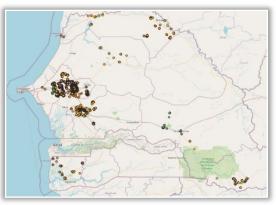
## **Sen4Stat** – National cropland map at 10-m resolution (Senegal, 2020)

96%

0,97

### Random Forest classification based on S2 & L8 time series only







2000 GPS traces for holdings selected through a stratified sampling from the 526.000 holdings in Senegal (~ 0.4%)















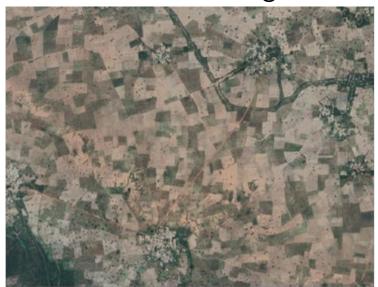






## Sen4Stat – Quality of in situ data as major bottleneck today

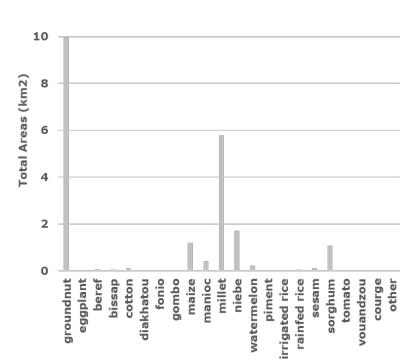






	Arachide	Maïs	Millet	Niébé	Sorgho	Autres	
Arachide	23296	0	6	1101	158	1184	90 %
Maïs	1	0	0	0	0	6	0 %
Millet	178	0	918	93	0	5	77 %
Niébé	3445	0	0	11318	158	1068	71 %
Sorgho	1996	0	0	1200	443	270	11 %
Autres	5037	0	84	2610	191	4256	35 %
	69 %	N/A	91 %	69 %	47 %	63 %	68 %



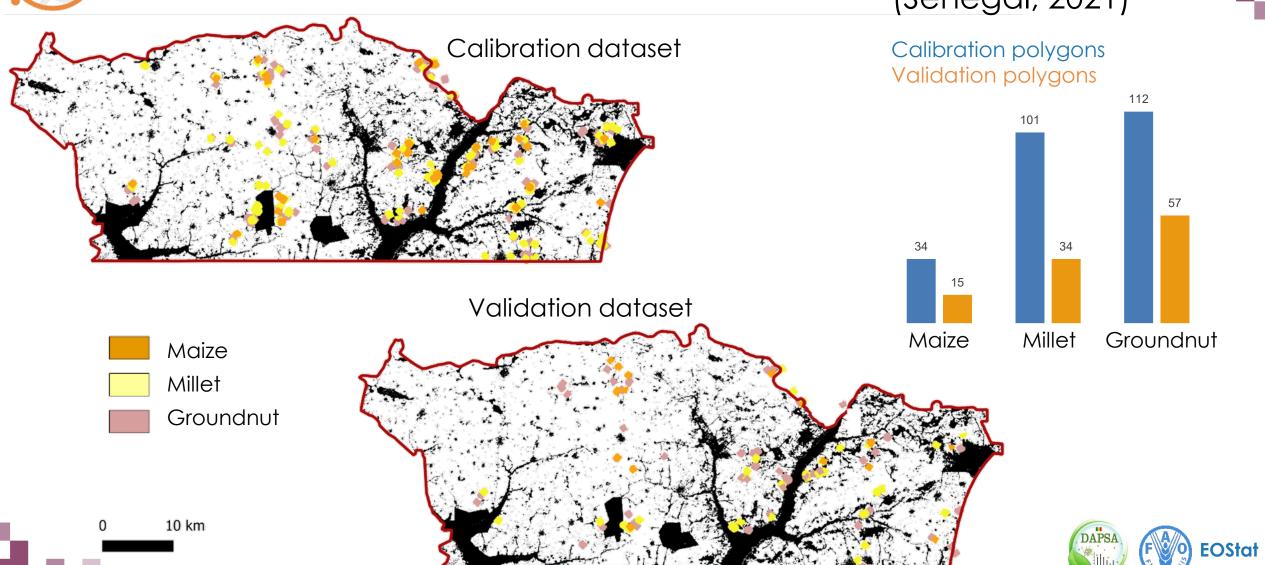






# **Sen4Stat** – Specific in situ data collection for department of Nioro







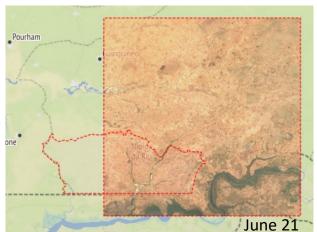




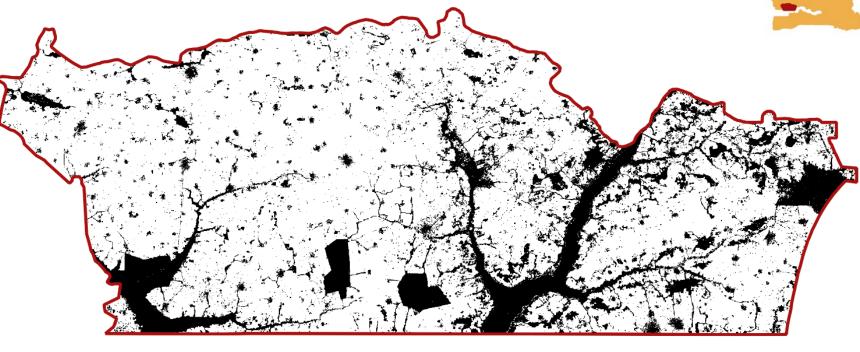


- 10-m cropland map using balanced in situ data collection for the Department of Nioro (Senegal)

Each \$2 image covering only a part of the department

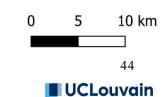






Overall Accuracy: 97,4 % F-Score cropland: 0,98 F-Score non-cropland: 0,95

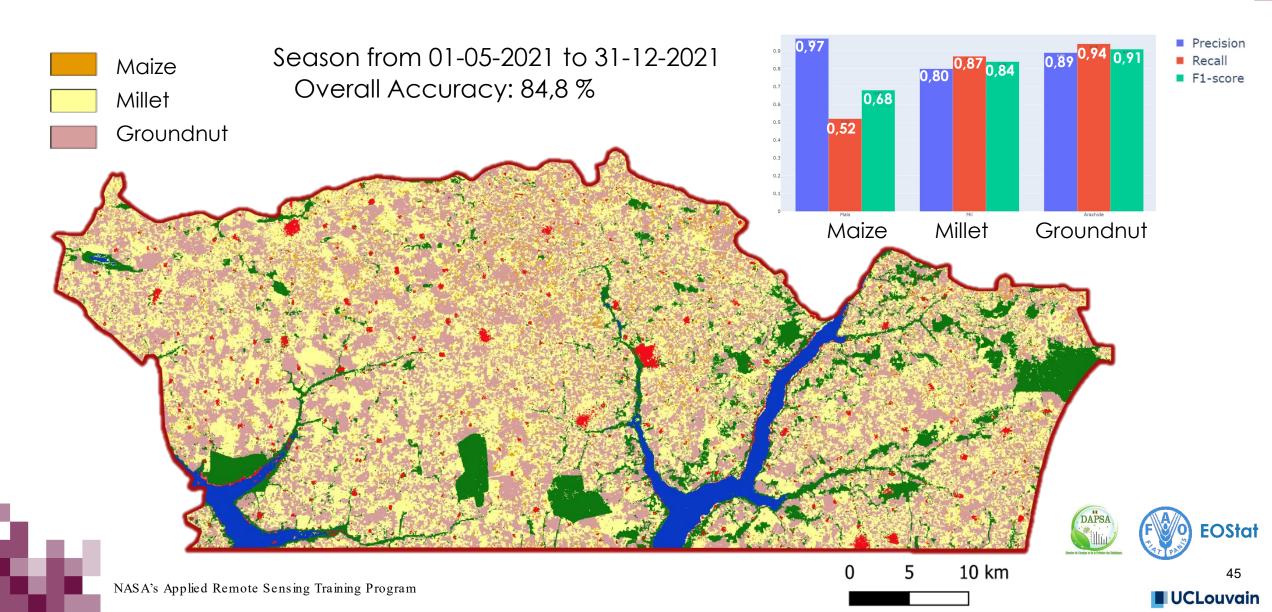






## 10-m crop type map from Sentinel-1 and Sentinel-2 time series acquired over the season and using balanced in situ dataset







## Sen4Stat – Operational exploitation of satellite missions to improve timely information about crop areas, yields and production





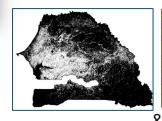


Field surveys

Prendre une photo



Production of crop type maps











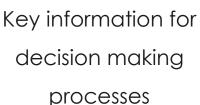


Area sampling frame



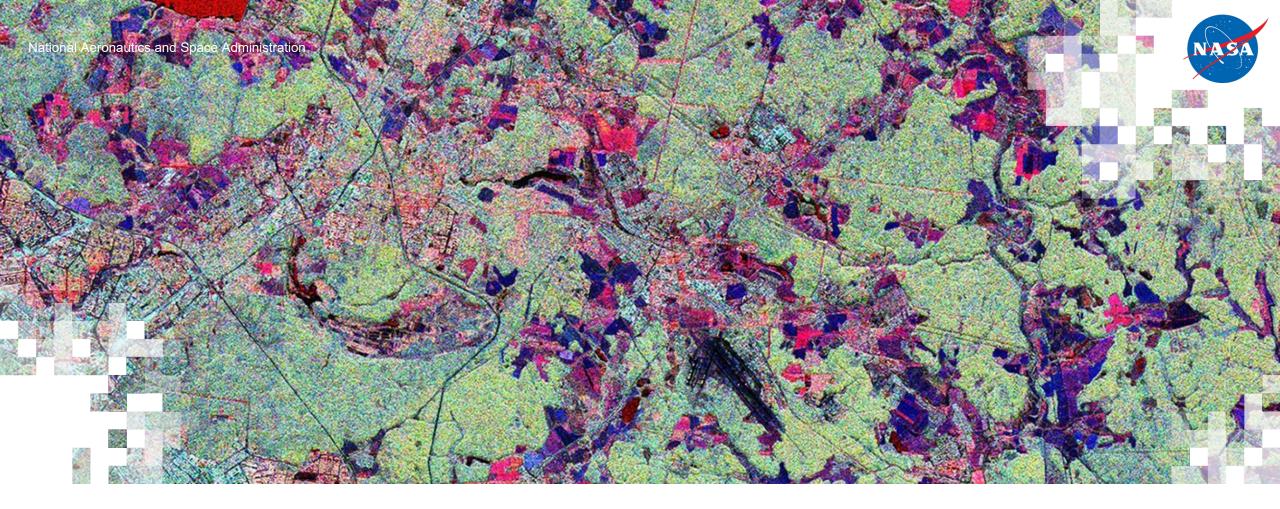














Section 3: Practical using Sen4Stat toolbox for crop type mapping

## Contacts

- Trainers:
  - Pr. Pierre Defourny pierre.defourny@uclouvain.be
- Training Webpage:
  - https://appliedsciences.nasa.gov/join-mission/training/english/arsetagricultural-crop-classification-synthetic-aperture-radar-and
- ESA's EO4Society Website:
  - https://eo4society.esa.int/training-education/
- Twitter: <u>@EOOpenScience</u>



### **Useful links**



Sen4Stat

https://esa-sen4stat.org/

Sen2Agri

https://github.com/Sen2Agri/Sen2Agri-System/

SNAP download

http://step.esa.int/main/download/snap-download/

Copernicus Open Access Hub (download Sentinel data)

https://scihub.copernicus.eu/

 NASA ARSET Agricultural Crop Classification with Synthetic Aperture Radar and Optical Remote Sensing Part 5: Biophysical Variable Retrieval using Optical Imagery to Support Agricultural Monitoring Practices

https://appliedsciences.nasa.gov/join-mission/training/english/arset-agricultural-crop-classification-synthetic-aperture-radar-and

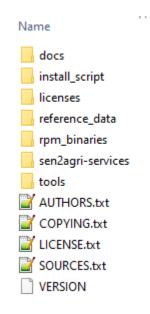
 ESA 'Advanced Training Course on Land Remote Sensing with the Focus on Agriculture' held in Louvainla-Neuve, Belgium, on 16-20 September 2019

https://eo4society.esa.int/resources/advanced-training-course-on-land-remote-sensing-with-the-focus-on-agriculture/



## Sen4Stat Installation Package (available in June 2022)

- Single installation archive containing all files needed
  - Documentation (docs)
  - Installation scripts
  - Sen4Stat rpm files
  - Sen4Stat Services
  - Other tools and configuration files
- CentOS 7 is required for system installation
- Installation performed through a single script
  - sudo ./install.sh
- Two directories with write permissions for all users are needed:
  - /mnt/archive
  - /mnt/upload
- During installation, all needed third party packages are installed via yum.
   Needed docker images are installed
- PostgreSQL database is used for storing system configuration, product information etc.
- The regular user sen2agri-service is created during installation
- System services are created and launched upon installation:
  - sen4stat-executor sen4stat-orchestrator sen4stat-http-listener sen4stat-demmaccs sen4stat-demmaccs.timer sen4stat-monitor-agent sen4stat-scheduler sen4statservices

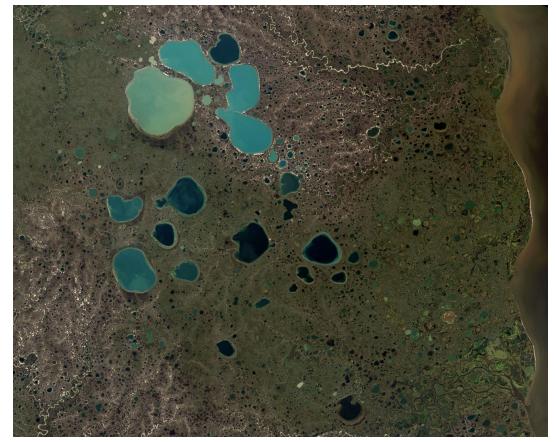


### **Homework and Certificate**

- Homework Assignment:
  - Answers must be submitted via Google Form
  - Due Date: May 17, 2022
- A certificate of completion will be awarded to those who:
  - Attend all live webinars
  - Complete the homework assignment by the deadline (access from website)
  - You will receive a certificate approximately two months after the completion of the course from: <a href="mailto:marines.martins@ssaihq.com">marines.martins@ssaihq.com</a>

## **Questions?**

- Please enter your questions in the Q&A box. We will answer them in the order they were received.
- We will post the Q&A to the training website following the conclusion of the webinar.



https://earthobservatory.nasa.gov/images/6034/pothole-lakes-in-siberia



**Thank You!** 

