

## In-Situ Data Collection and Preparation

### Novel Approaches for Agricultural EO Applications

**Raphaël D'ANDRIMONT, Ph.D.**

✉ [raphael.dandrimont@gmail.com](mailto:raphael.dandrimont@gmail.com)    🐦 @rdandrimont

🌐 <https://github.com/rdandrimont/>

🌐 <http://www.linkedin.com/in/dandrimont/>

Raphael D'ANDRIMONT

22/11/2022



# The European Commission's science and knowledge service

Joint Research Centre



European  
Commission

# JRC

## Joint Research Centre



# The JRC is the European Commission's knowledge and science service

**€ 386** million Budget annually,  
plus **€ 62** million earned income

**Independent** of private,  
commercial or national interests



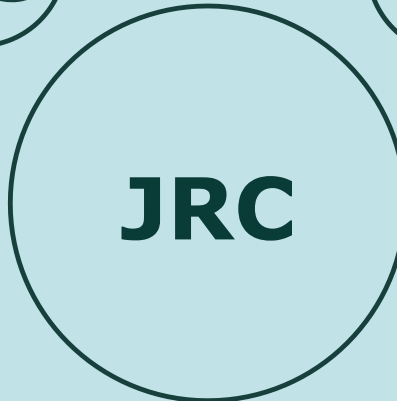
**Policy neutral:**  
has no policy agenda of its own

**42** large scale research facilities,  
more than 110 online databases

**30%** of activities in policy preparation,  
**70%** in implementation

**125** instances of support  
to the EU policy-maker annually

More than **100** economic,  
bio-physical and nuclear models



**1500** core research staff, **3000** total staff  
**83%** of core research staff  
with PhD's



**Over 1,400**  
scientific publications per year



# In-situ

## The holy grail





In situ

# IN - SITU COMPONENT: OVERVIEW

- *In situ* data = "observation data from ground-, sea-, or air-borne sensors, reference and ancillary data licensed or provided for use in Copernicus" (*Copernicus regulation – article 3*)
- Use of *In situ* data:
  - Validate & calibrate Copernicus products
  - Reliable information services





# In-situ data is still the holy grail of remote sensing

In-situ data for EO agriculture could be

- Crop type
- Phenology information
- Agricultural practices
- Biophysical variables (e.g.LAI, LAIeff, FAPAR or FCOVER, soil moisture, yield, crop height, density )

and other data collected on the **ground** or obtained from data analysis.

They are used for **training** / **calibration** of algorithm and for **validation**.

As remote sensing **data** arise « free and open »,  
Processing **capabilities** are available,  
Analytical **method** are mature,

the **last frontier is the in-situ data** to generate high quality information.

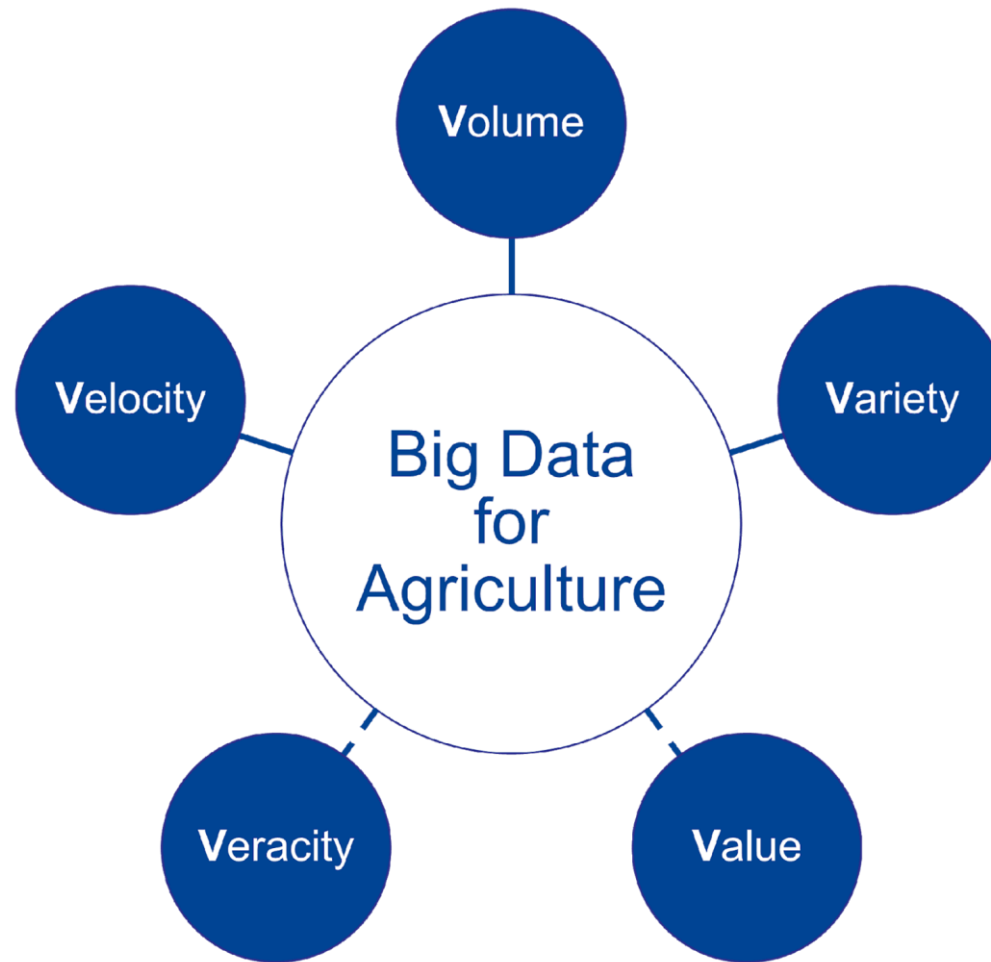
# Big Data and Agriculture

## Context





# Agriculture in the data age



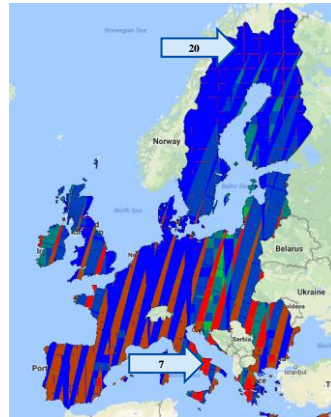
# Velocity, Volume and Variety are now the state of the art

70.000.000 EU agricultural parcels observed more 300 times/year with Copernicus (<20 m)



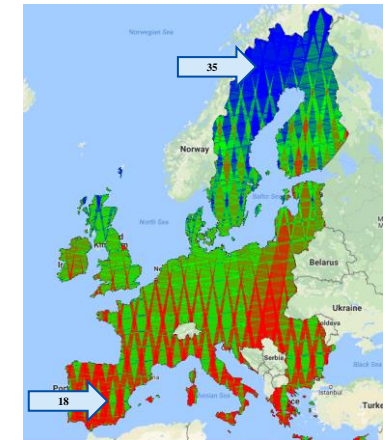
## sentinel-1

+/- 240 observations per year (half of it with 1 sensor)



## sentinel-2

+/- 90 observations per year



## sentinel-3

+/- 365 observations per year

Cloud processing



Google Earth Engine

### Novel technology



Volume

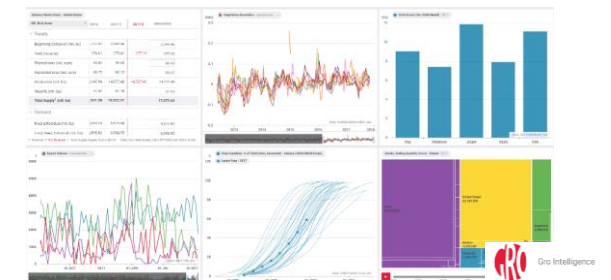
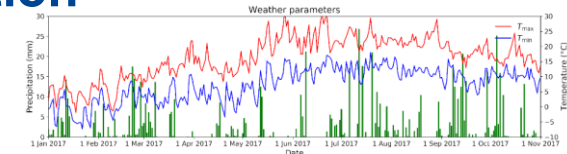
Velocity

Big Data for Agriculture

Variety

### Weather information

daily



### Market and farm information



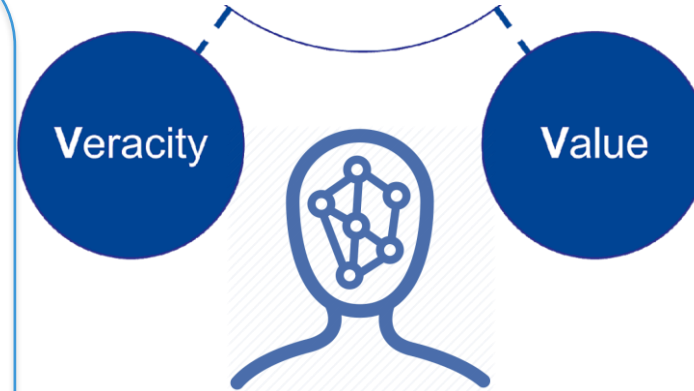
# Disruptive ways to bring Veracity and Value?

## Administration data

- Yearly parcel crop type information (LPIS-GSAA)
- Statistics (LUCAS)



Farmers' declarations



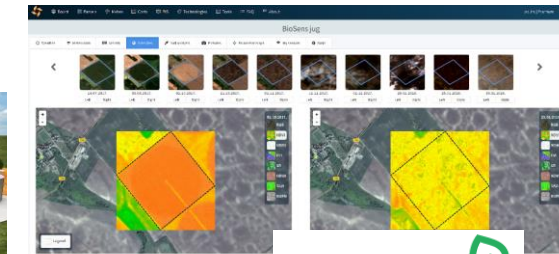
## Farmers data

Farm management tools



**LandSense**  
A Citizen Observatory and Innovation Marketplace for Land Use and Land Cover Monitoring

Digital Agriculture of Serbia  
AgroSense



OneSoil

Farm sensors and machinery



## Survey



## Crowdsourcing

Active



Geo-tagged street level imagery



Opportunistic



Mapillary

OpenStreetMap



Social Networks



European Commission

# Disruptive ways to bring Veracity and Value?

## Administration data

- Yearly parcel crop type information (LPIS-GSAA)
- Statistics (LUCAS)



Farmers' declarations



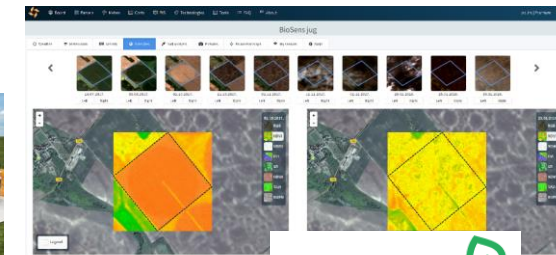
## Farmers data

Farm management tools



**LandSense**  
A Citizen Observatory and Innovation Marketplace for Land Use and Land Cover Monitoring

Digital Agriculture of Serbia  
AgroSense



OneSoil

Farm sensors and machinery



## Survey



## Crowdsourcing

Active



Geo-tagged street level imagery



Opportunistic



Mapillary



OpenStreetMap



Social Networks

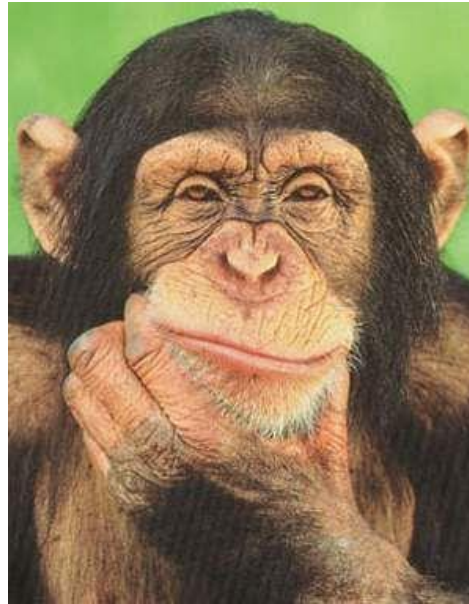


European Commission



# Roadside survey with camera

## Proof of concept study



# Survey with roadside data collection

Roadside data collection is an easy and cheap type of survey to collect crop information

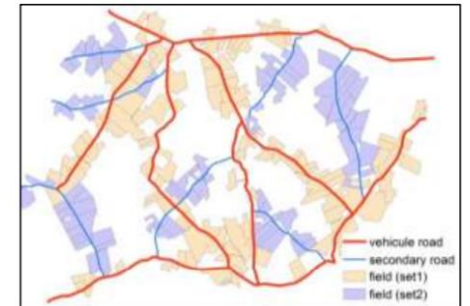
(see JECAM guidelines )

Why looking at new ways to collect roadside data?

- Traditional surveying lacks the **scale** and possibility for **automated** integration using big data analytics
- Prone to sampling **errors**
- Require a considerable organizational **effort** and **money**,
- Difficult to achieve **periodic re-sampling** to assess changes in dynamic agricultural phenomena

**JECAM**  
Joint Experiment for Crop Assessment and Monitoring

**GEO** GROUP ON  
EARTH OBSERVATIONS



JECAM guidelines 'windshield' strategy

# Which sampling strategy?

Roadside and Transect data are significantly less representative of the population compared to random data.

Differences in representativeness do not systematically translate into marked accuracy differences (<2%).

Transect sampling is subject to drops in accuracy as large as 15% and its accuracy levels highly variable.

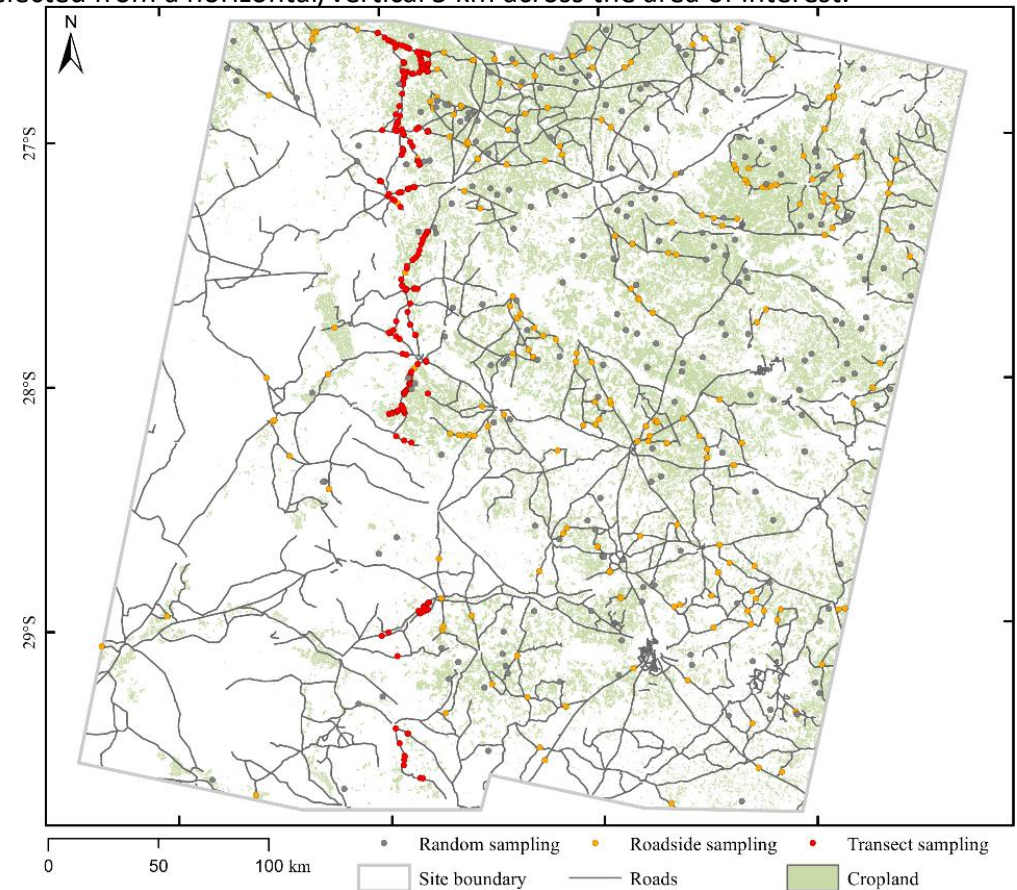
**Random and Roadside** training sets with similar representativeness yield comparable accuracy.

Roadside sampling for cropland mapping is a valuable source of calibration data when the range of environmental and management gradients are surveyed.

**Random sampling** selects cropland training samples irrespective of their location.

**Roadside sampling** emulates what can be collected from a roadside survey. Cropland pixels <250 m away from a road are randomly selected.

**Transect sampling** is a spatially-constrained variant of Roadside sampling. Cropland pixels are selected from a horizontal/vertical 5 km across the area of interest.



Waldner et al. (2019)



# CAP Checks by monitoring

## Sentinels and LPIS

- *From sampled approach to 100%*
- *Detailed area managed in LPIS+GSAA*
- *Sentinels confirm declared crop*
- *Sentinels monitor agricultural activity (e.g. grass mowing, catch crops)*
- *In EU(2018)/746 since May 2018*

- *Outlier identification preferably < 5%*
- *Marker analysis for activity detection*
- *Requires Big Data Analytics*



*Crop type at parcel level with Sentinel 1*



# Grasslands are key biotopes to monitor and were used for a case study

Key biotopes in **climate change mitigation** (support carbon sequestration)

Habitats for plants and animals supporting **biodiversity**

**Declining** since the 1960s in EU → **EU Policies** developed

- Limit declines in the ratio of permanent grassland to total agricultural area below 10 % and then 5%
- Management of at least 5% of the arable land as Ecological Focus Areas

→ **Grassland monitoring needed**



# Specific research questions

- (1) Define efficient **markers for grassland** monitoring combining **Sentinel-1** SAR and **Sentinel-2** multi-spectral observations
- (2) Evaluate the efficiency and effectiveness of **street-level imagery** as a source of ground truth



# A 1-day survey in the Netherlands

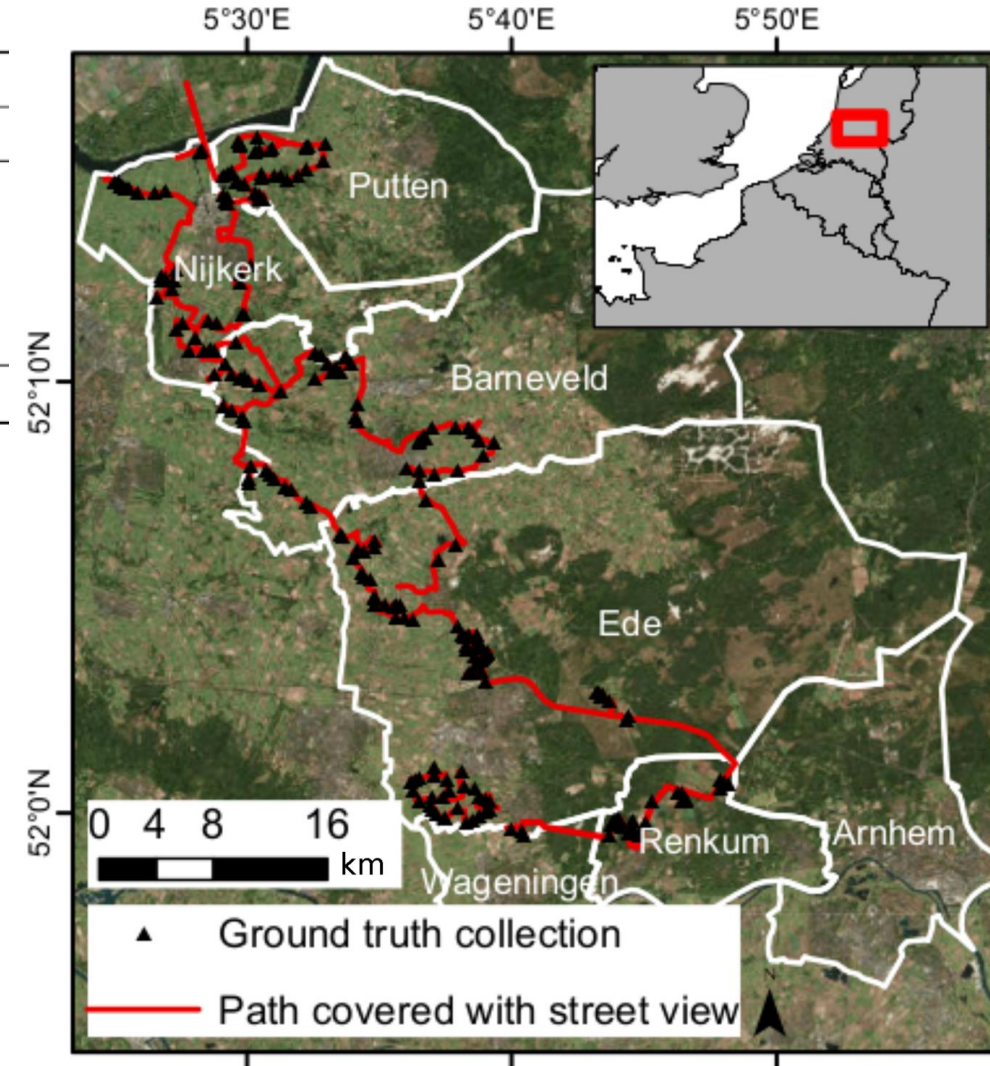
## Study site

- 15395 parcels

## Survey

- 231 parcels observed by surveyor
- 1411 parcels surveyed by cameras

Class	Parcel Area		Parcel Count	
	Area (Ha)	Area (%)	Number	Number (%)
GRA	18,629	68.90	11,773	76.47
MAI	507	18.75	2637	17.13
CER	451	1.67	198	1.29
POT	371	1.37	116	0.75
OTH	2517	9.31	671	4.36
TOTAL	27,039	100.00	15,395	100.00



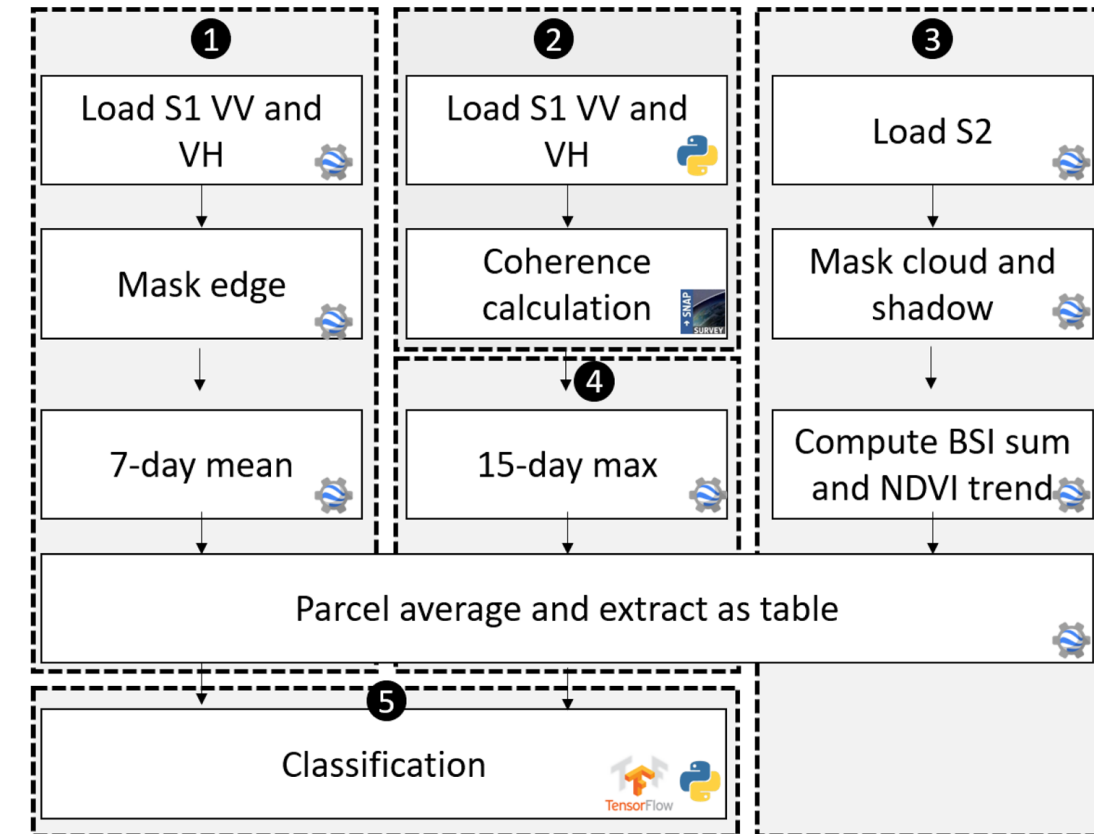
Field survey combining ground truth and roof cameras

Study site

# 'Mark' the non-grassland parcels with the Sentinels

- 15000 parcels from BRP
- S1 VV, VH, COHERENCE TensorFlow classification
- S2 Bare Soil Index (BSI) and NDVI trend

$$BSI = \frac{(\rho_{SWIR1} + \rho_{Red}) - (\rho_{NIR} + \rho_{Blue})}{(\rho_{SWIR1} + \rho_{Red}) + (\rho_{NIR} + \rho_{Blue})}$$



*Main processing steps*

→ Target potential 'non grasslands' parcels



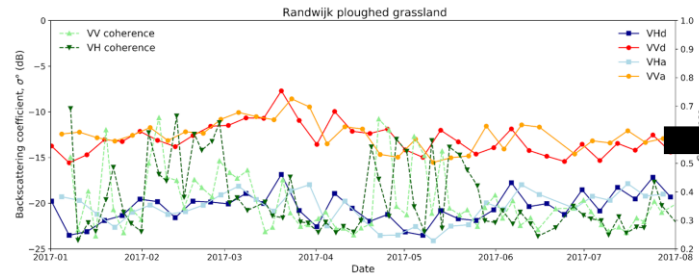
<https://github.com/rdandrimont/AGREE>



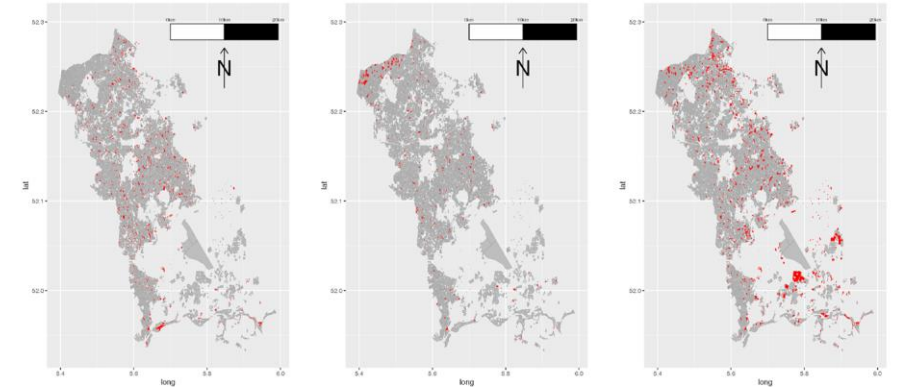
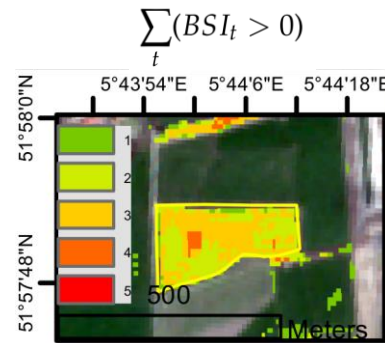
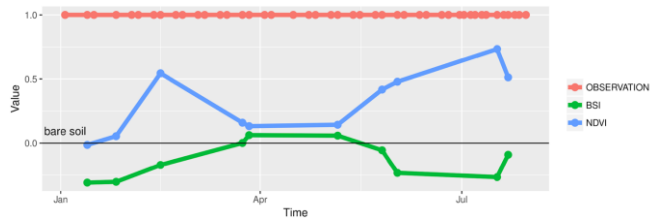
European  
Commission



# What is available at parcel level?



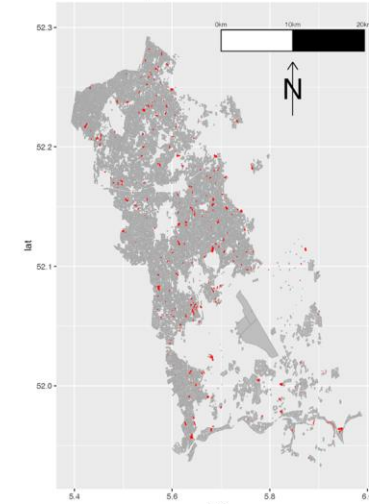
Probability  
of being  
outliers



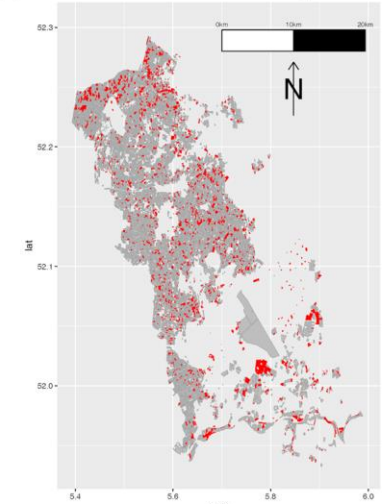
S1 backscatter

S1 coherence

S2 BSI



S1 AND S2

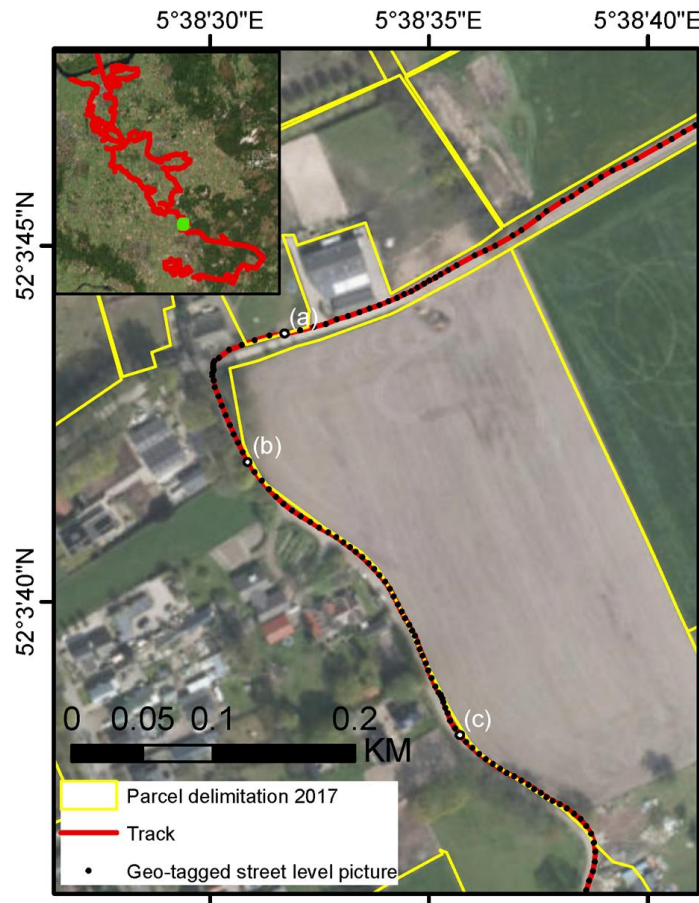


S1 OR S2

*Marked fields mapped*



# Crop types from street-level imagery



*Parcel observation with street-level pictures*



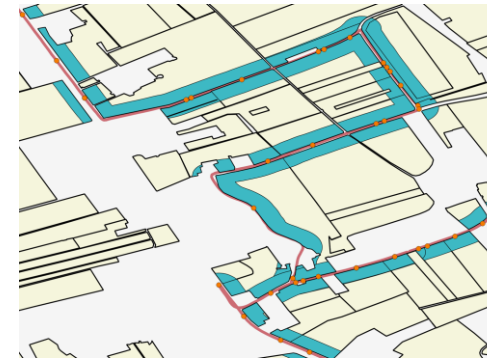
(a)



(b)



(c)



*Along road sampling*

Class	Street-Level Observation (N Fields)	Percentage of the Total (%)
CER	11	0.78
GRA	1157	82.00
MAI	192	13.61
OTH	48	3.40
POT	3	0.21
TOTAL	1411	100



(a)



(b)

Interpreter	1	2	3
Overall accuracy	92.06	92.06	88.79
Kappa	84.94	84.85	79.59

*Interpretation of the street-level pictures*

# 3 key results

- The number of parcels identified by **S1 and S2** as non grassland ranged from **2.57% to 17.12%** of the total of 11,773 parcels declared as grassland
- Parcels flagged by the combined S1 and S2 marker were **robustly detected** as non-grassland parcels with **ground-truth** (F-score = 0.9)
- **Street-level imagery** collection could improve collection **efficiency** by a 7 factor compared to field visits (1411 parcels/day vs. 217 parcels/day) while keeping an overall accuracy of about 90% compared to the ground-truth

Farmers' declaration versus the ground-truth

	Ground-Truth						Metrics		
	CER	GRA	MAI	OTH	POT	TOTAL	UA	PA	F-Score
CER	1	0	1	2	0	4	1.00	0.25	0.40
GRA	0	130	9	5	0	144	0.99	0.90	0.95
MAI	0	0	78	0	0	78	0.89	1.00	0.94
OTH	0	1	0	2	0	3	0.22	0.67	0.33
POT	0	0	0	0	2	2	1.00	1.00	1.00
TOTAL	1	131	88	9	2	231	-	-	-

Markers versus ground-truth

Markers	TP	FP	FN	TN	Sensitivity	Specificity	Precision	Accuracy	F-Score
S1 backscatter	90	2	40	12	0.69	0.86	0.98	0.71	0.81
S1 coherence	91	7	39	7	0.70	0.50	0.93	0.68	0.80
S2 BSI	72	2	58	12	0.55	0.86	0.97	0.58	0.71
S1 AND S2	109	2	21	12	0.84	0.86	0.98	0.84	0.90
S1 OR S2	20	1	110	13	0.15	0.93	0.95	0.23	0.26

Markers versus street-level pictures

Markers	TP	FP	FN	TN	Sensitivity	Specificity	Precision	Accuracy	F-Score
S1 backscatter	92	0	82	10	0.53	1.00	1.00	0.55	0.69
S1 coherence	47	2	44	5	0.52	0.71	0.96	0.53	0.67
S2 BSI	95	0	84	11	0.53	1.00	1.00	0.56	0.69
S1 AND S2	32	0	23	9	0.58	1.00	1.00	0.64	0.74
S1 OR S2	193	0	180	13	0.52	1.00	1.00	0.53	0.68

(d'Andrimont et al., 2018)

# Computer Vision

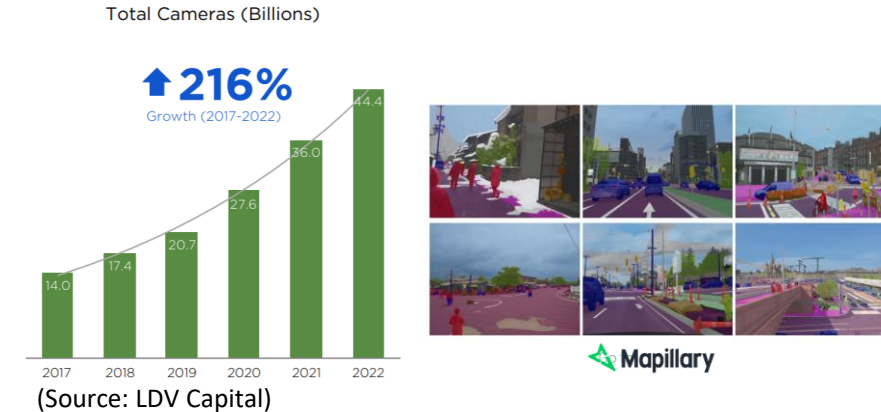
**From pictures to knowledge**





# A paradigm shift in visual data capture

- Capture evolution
  - 44 billion of **cameras** by 2022
  - 700 trillion of **pictures** taken every day by 2050
- **Computer vision** deep learning recent developments
  - Deep learning and google street view to estimate voting patterns in US (Gebru et al., 2017)
  - Treepedia (Seiferling et al., 2017)
- Improved **processing** capabilities
  - cloud, GPU
- **Crowdsourcing** potential
  - Mapillary open platform (1,800 M pictures)



Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States

Timnit Gebru, Ji Li Fei-Fei  
PNAS December 12  
<https://doi.org/10.1073/pnas.1711111115>

Ian Seiferling<sup>a,b,\*</sup>, Nikhil Naik<sup>c</sup>, Carlo Ratti<sup>a</sup>, Raphaël Proulx<sup>b</sup>

<sup>a</sup> Senseable City Laboratory, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Room 10-485, 77 Massachusetts Avenue, Cambridge, MA, 02139, United States

<sup>b</sup> Canada Research Chair in Ecological Integrity, Centre de Recherche sur les Interactions Bassins Versants-Ecosystèmes Aquatiques, Université du Québec à Trois-Rivières, 3351 Boulevard des Forges, Trois-Rivières, Québec, G9A 5H7, Canada

<sup>c</sup> MIT Media Lab, 75 Amherst St, Cambridge, MA 02139, United States



# CAP Checks by Monitoring

## IACS & geo-tagged pictures

- *Complimentary information provided by farmers and controllers*
  - *Based on a monitoring alert, or to document specific measures (e.g. grassland mix)*
- 
- ***Farmers may supply millions of pictures...***
  - ***In-situ data needed for a sample of parcels subject to CAP CbM***
  - ***Computers may help to recognize crops on 90% of the pictures***





# What can geo-tagged pics be used for?

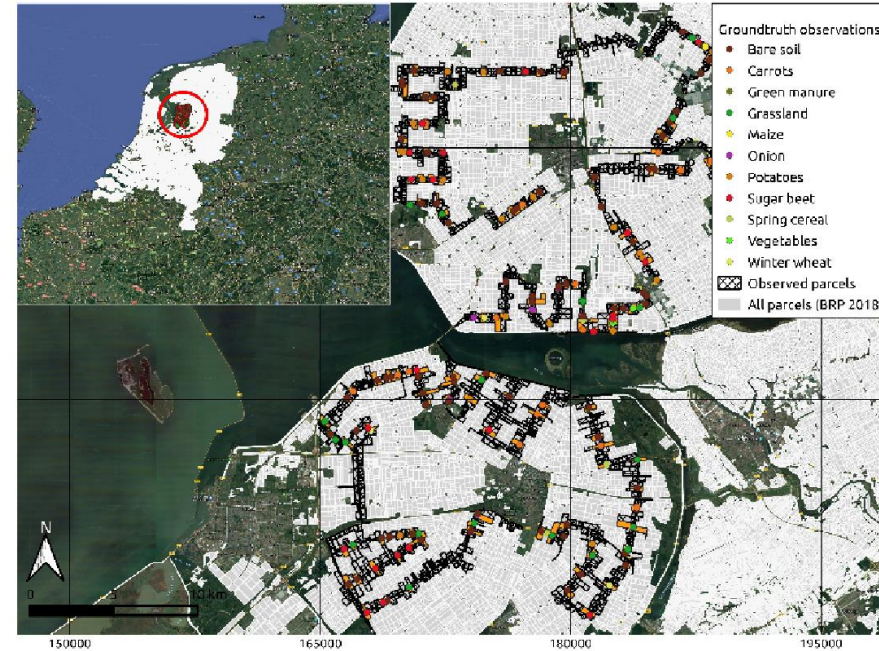
Evidence for practices we cannot determine with remote sensing





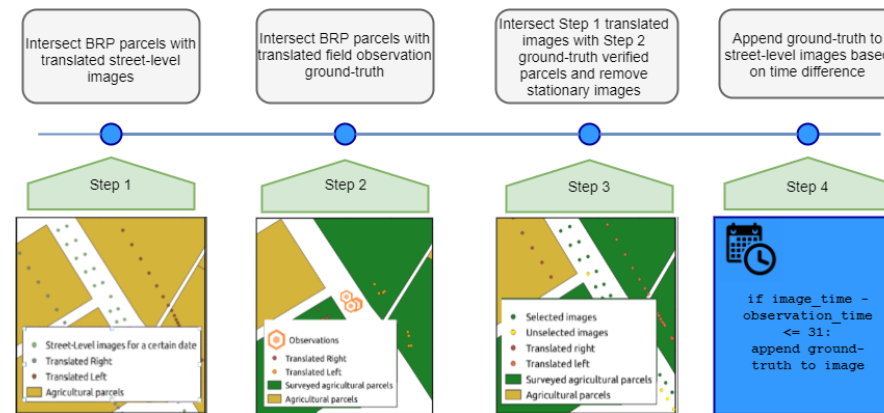
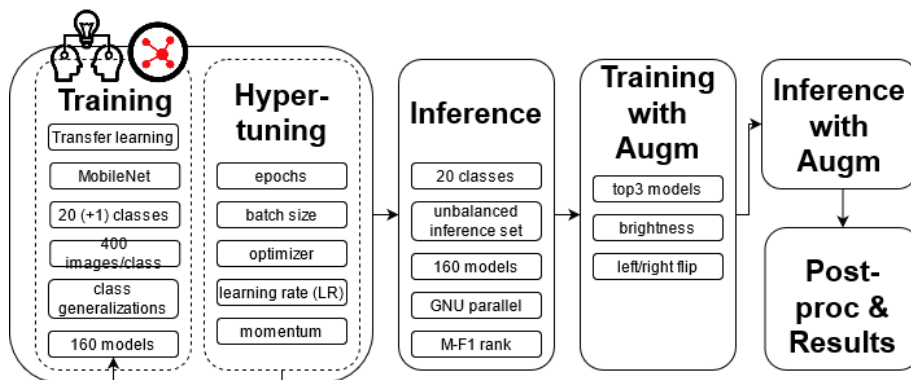
# Monitoring crop phenology with street-level imagery using computer vision

- Street-level imagery (SLI) for scaling up in-situ data.
- Crop type and phenology (BBCH-scale) monitoring.
- 2018 field campaign
- Geo-processing workflow to append BBCH info to parcels and SLIs.
- Hyper-parametarizing



## Field data collection:

- 2 side looking Sony action cameras
- 1 Nikon high resolution camera
- 8 field visits (March-October 2018)
- 200 km route
- 49,000 SLI/visit == ~ 400,000 SLIs
- 220 in-situ pheno observed parcels
- 17 observed crops
- 82 crop-pheno combinations



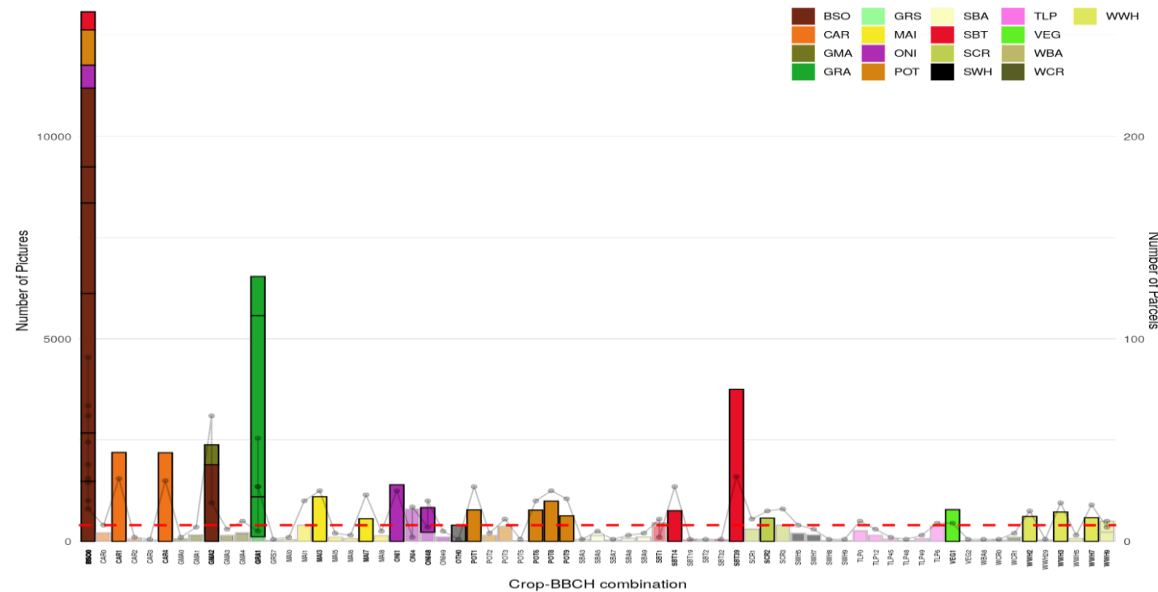
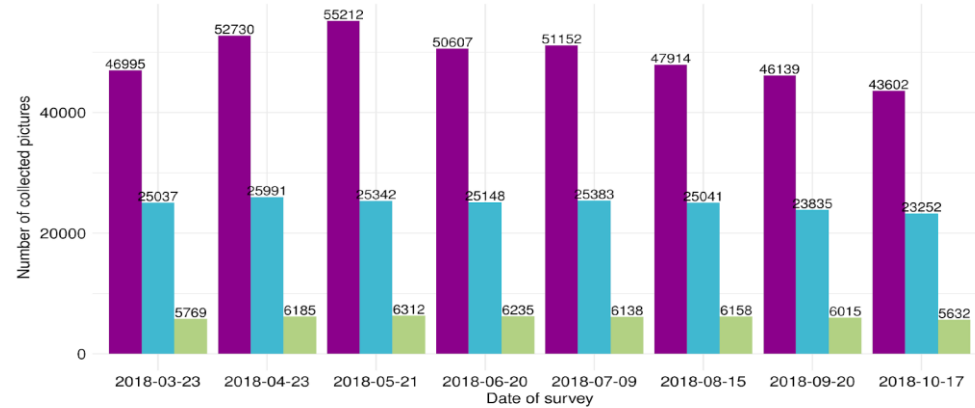


MAIO

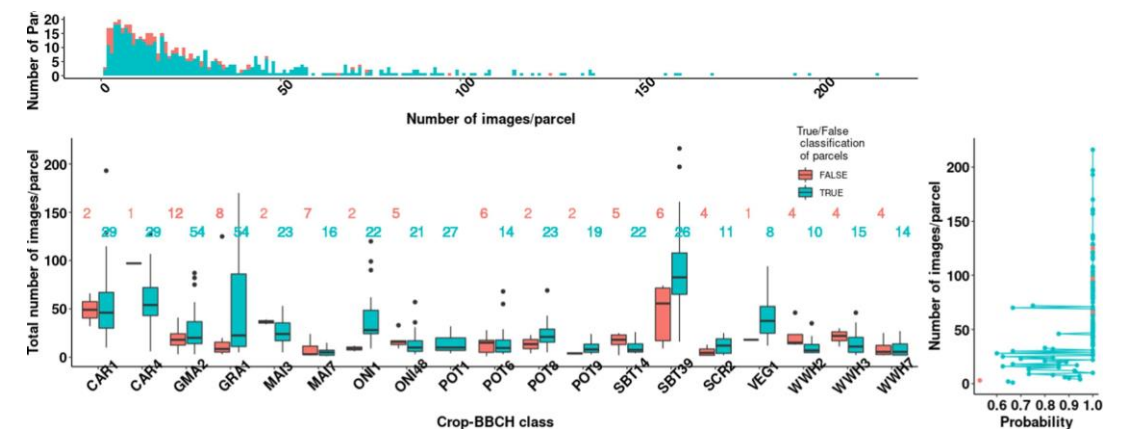
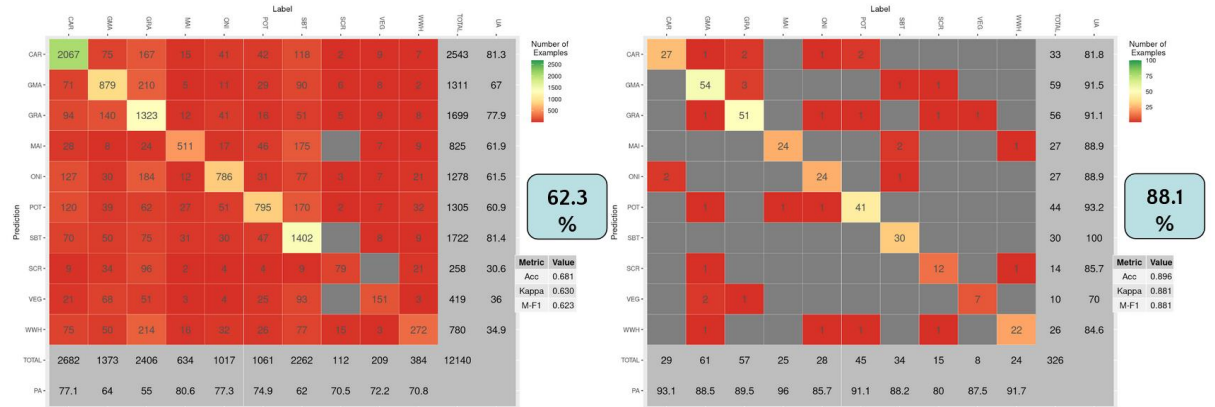
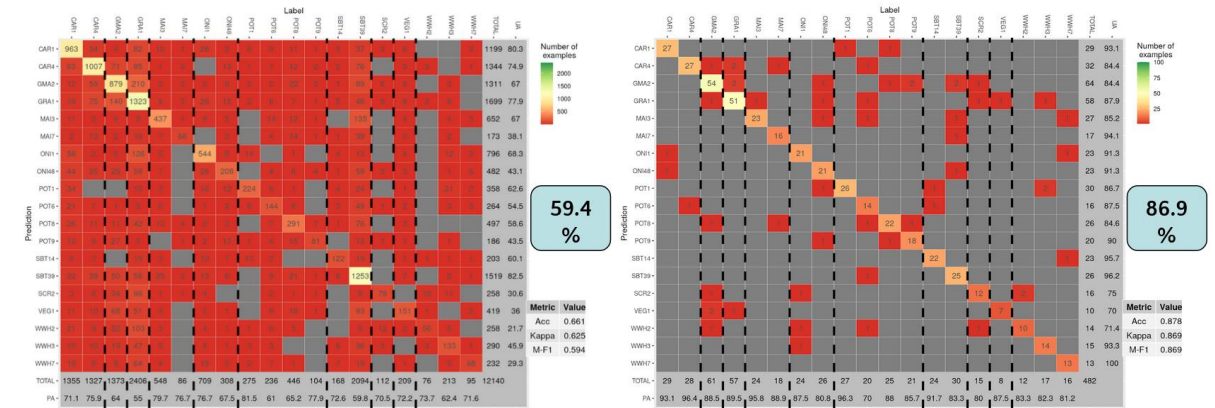




# Results



- Limitations include: detail of data lost due to CNN resampling; CNN input resolution; "OTH" and "BSO" class handling; BBCH distribution along parcel assumption;





# FlevoVision v2 - Towards Hierarchical multinet classification



1. Geolocation, association with parcel, and orientation towards parcel.
2. Automatically segment a cropped parcel on street level imagery.
3. Crop type classification on segmented window?
4. Use full scope of pheno (BBCH) information in the data.

# FlevoVision v2 – Tiling and tile filtering




```
In [3]: from pigeon import annotate
        from IPython.display import display, Image
        from os import walk

In [14]: f = []
        mypath = '/Dropbox/backup/work/Ispra/FlevoLand/heuristic4IS/postTilingFilteringPresi/tiles/'
        for (dirpath, dirnames, filenames) in walk(mypath):
            f.extend(filenames)
            break
        tileList = [mypath + s for s in f]

In [15]: annotations = annotate(
        tileList,
        options=['goodTile', 'badTile'],
        display_fn=lambda filename: display(Image(filename))
    )
```


0 examples annotated, 12 examples left

☒ goodTile ☐ badTile ☐ skip



1 examples annotated, 11 examples left

☐ goodTile ☒ badTile ☐ skip



Methods for assessing fitness of tile – find difference between summary statistic (SumStat) of tile and SumStat of Detectron parcel mask.  
SumStats tried – RGB\_mean, RGB\_mode, RGB\_quantiles, histogram comparison.



## And other way ...





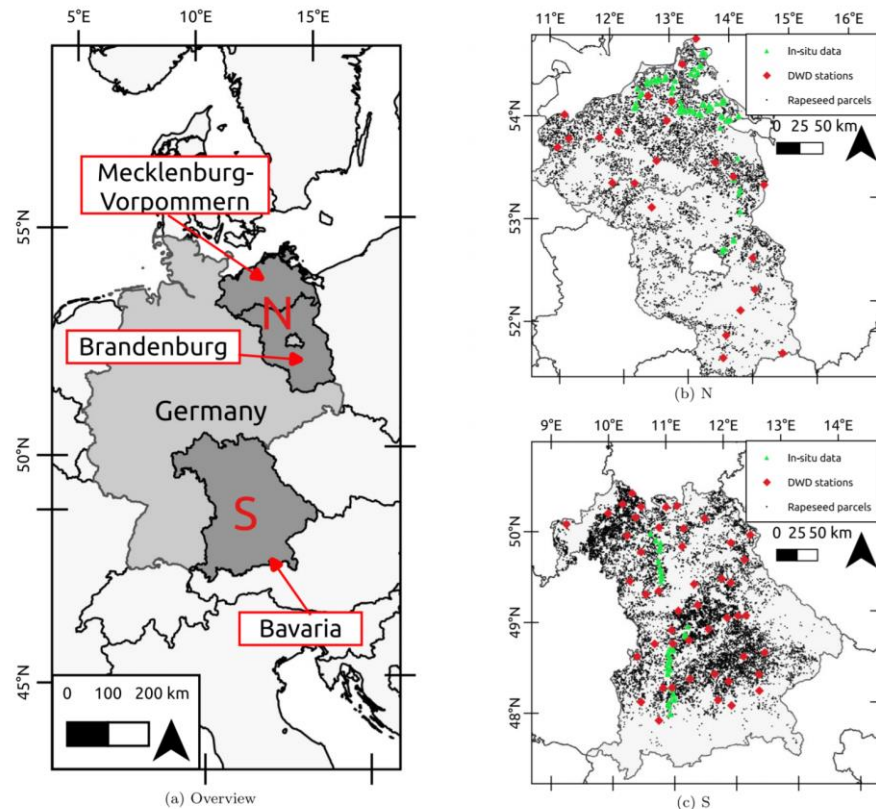
# Roadside surveys

## Other examples



# Can we accurately detect the peak of rapeseed flowering with S1 and S2 time-series at parcel level?

Study site and in-situ data



	<i>In-situ</i>	<i>Regions</i>	<i>In-situ BBCH observations</i>							<i>DWD observations</i>	
	<i># parcels</i>	<i># parcels</i>	<i>61</i>	<i>62</i>	<i>63</i>	<i>64</i>	<i>65</i>	<i>67</i>	<i>69</i>	<i>61</i>	<i>69</i>
<i>North</i>	115	10693	8	9	19	44	112	131	11	23	23
<i>South</i>	114	21662	30	29	26	9	99	109	6	47	47

Photo collection with roof cameras



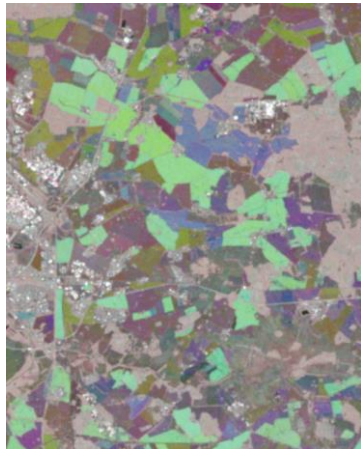
Close-up photos to identify BBCH stage



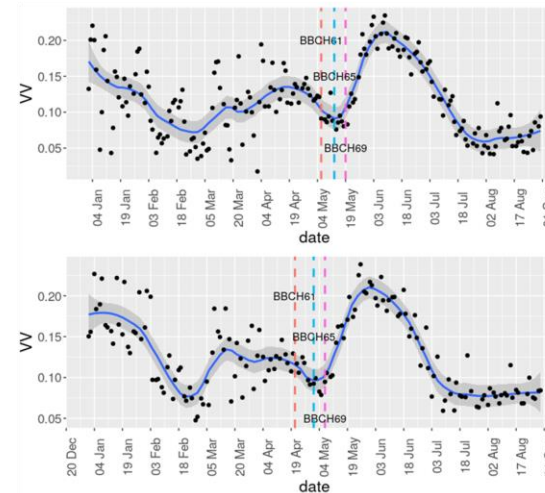


# Satellite time series detected parcel flowering date with an accuracy of 1 to 4 days combining Sentinel-1 & -2

Sentinel-1 : structure and morphology



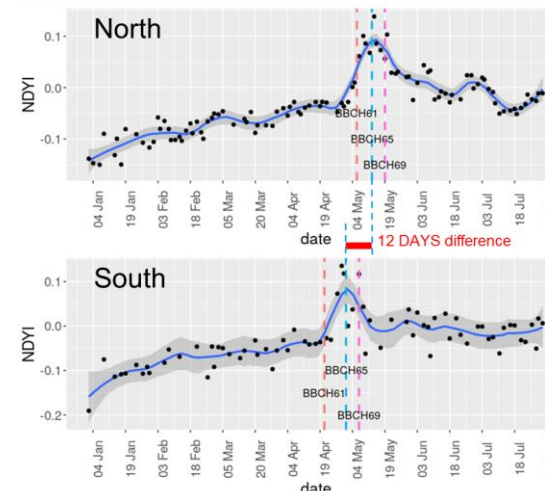
VV drop during flowering



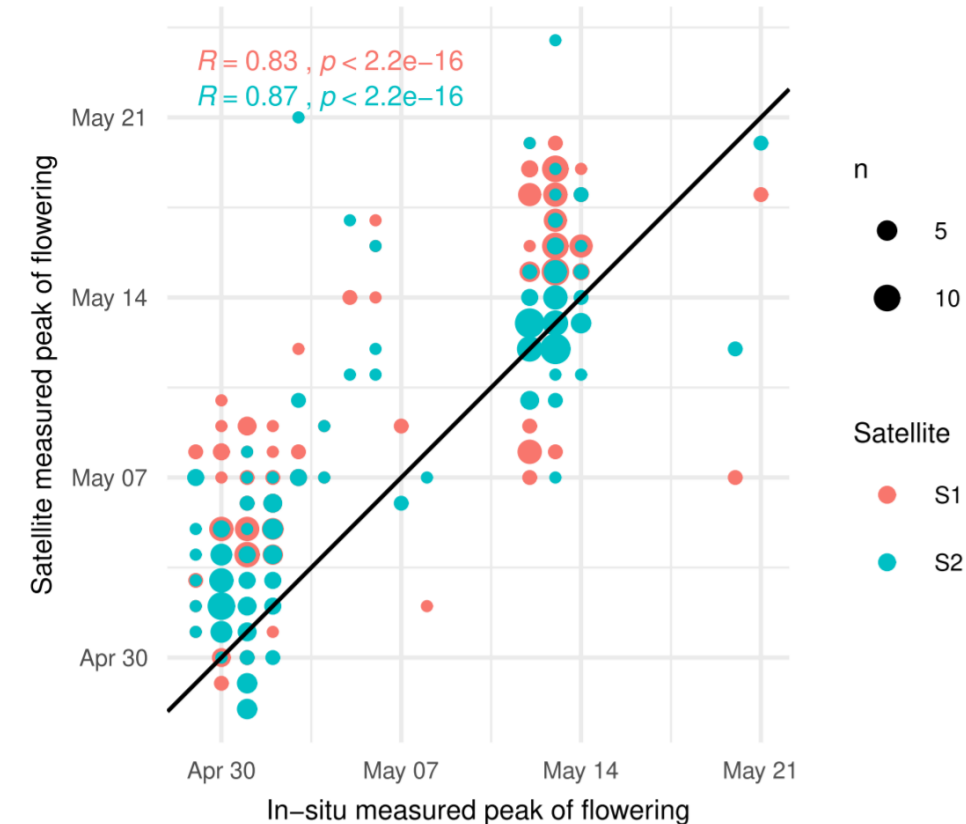
Sentinel-2 : yellow index



NDVI peak during flowering



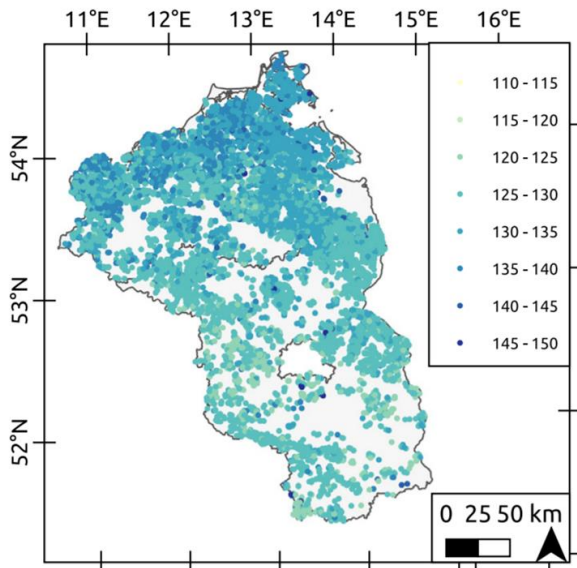
Parcel in-situ and Sentinels



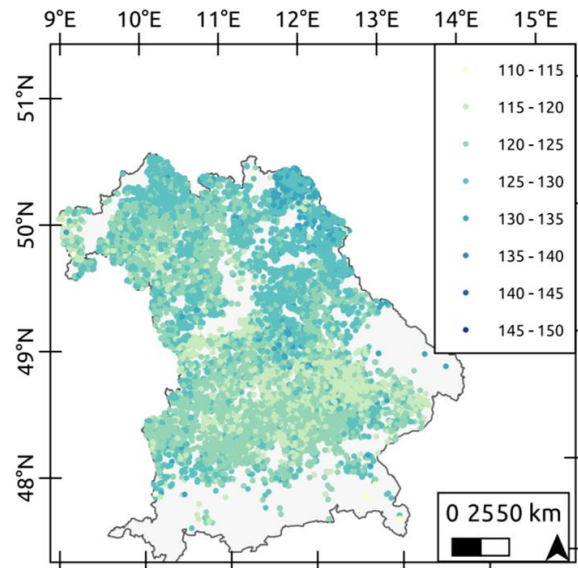


# Mapping at regional scale shows expected climatic gradient in flowering from South to North

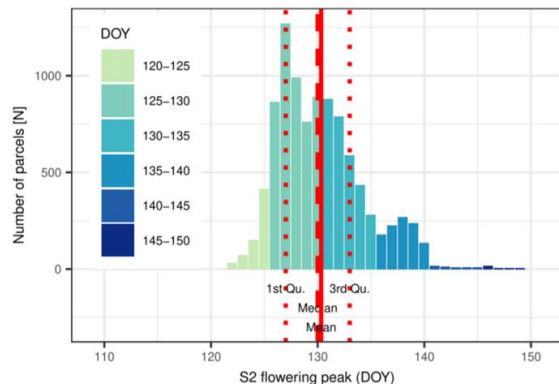
Parcel flowering peak (DOY) at regional level



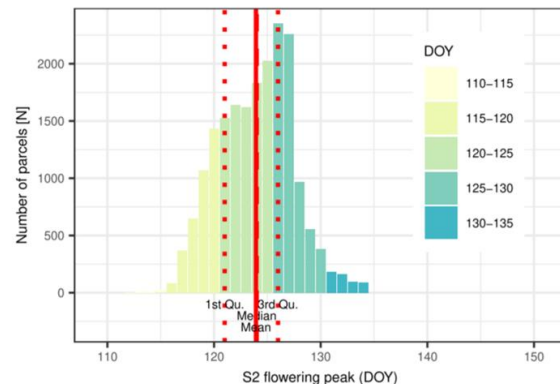
(a) N



(b) S



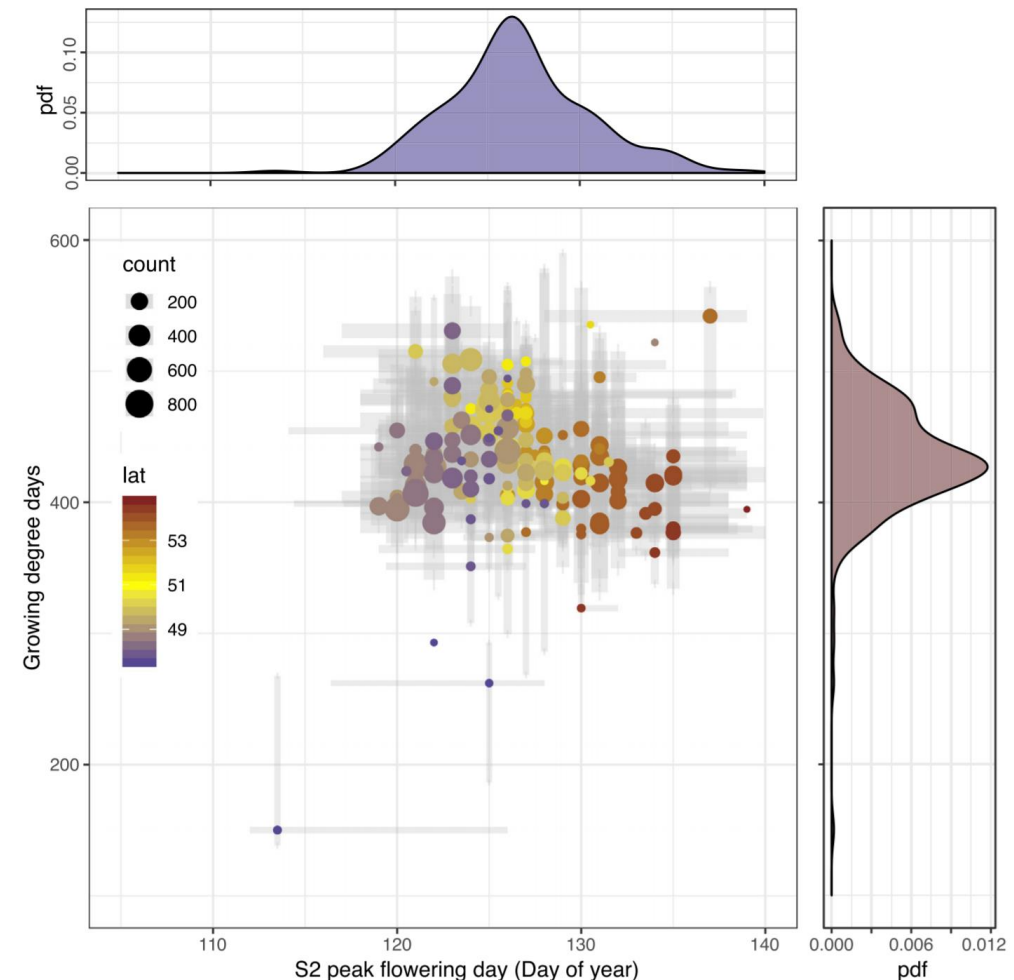
(c) N



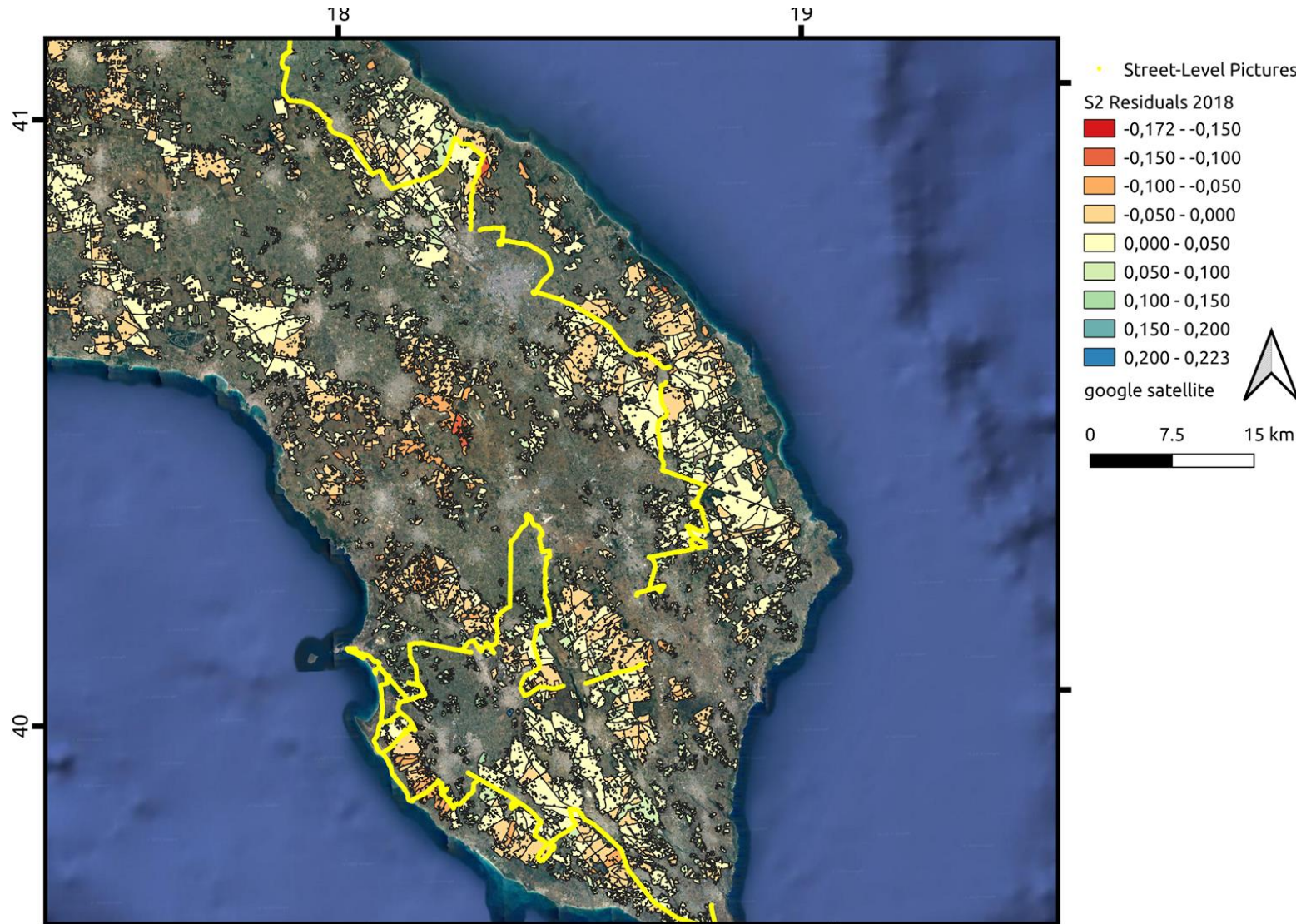
(d) S

25-km aggregation  
and thermal time

Crop model-relevant thermal time to flowering based on satellite observations.



# Can we quantify the damage of *Xylela fastidiosa* in Puglia?



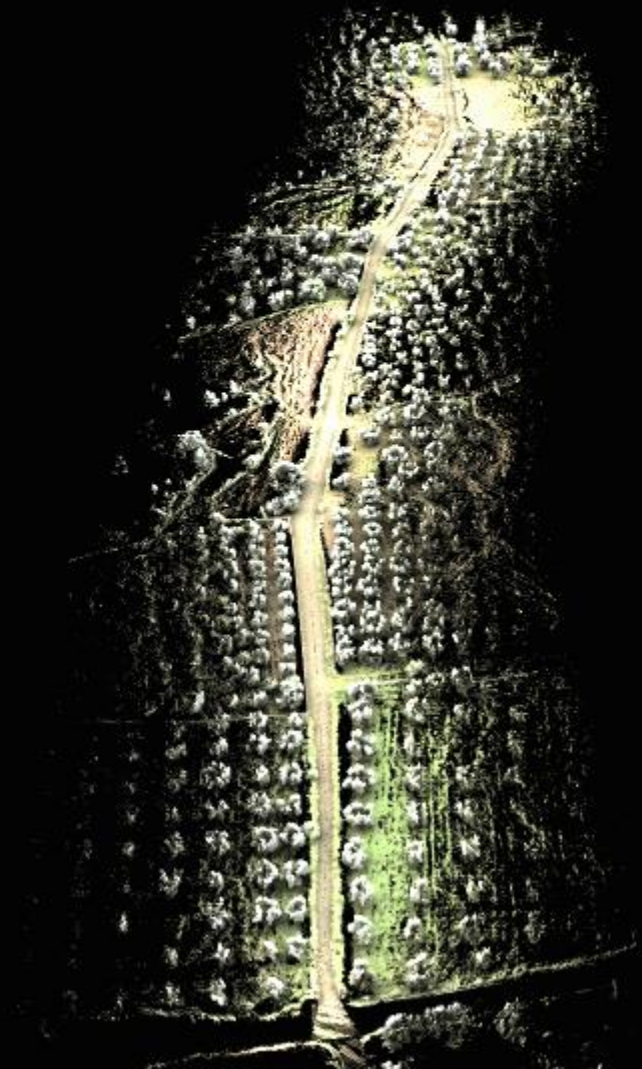
300 km  
2 days



# Data overview







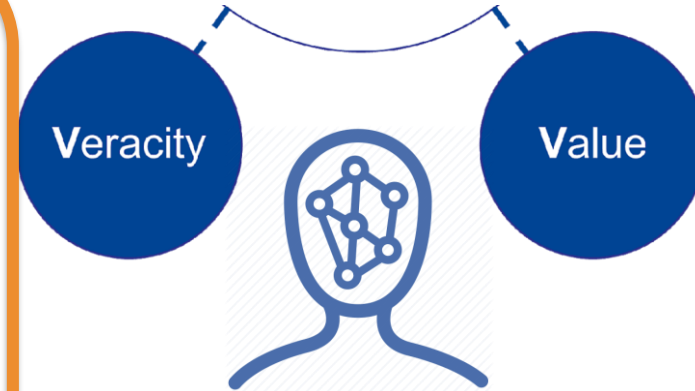
# Disruptive ways to bring Veracity and Value?

## Administration data

- Yearly parcel crop type information (LPIS-GSAA)
- Statistics (LUCAS)



Farmers' declarations



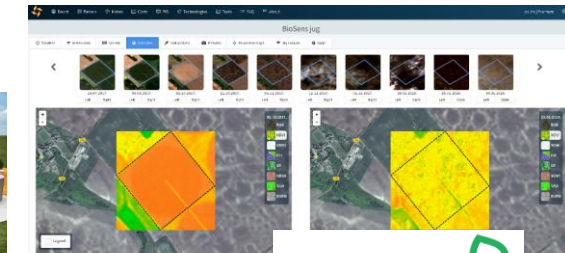
## Farmers data

Farm management tools



**LandSense**  
A Citizen Observatory and Innovation Marketplace  
for Land Use and Land Cover Monitoring

Digital Agriculture of Serbia  
AgroSense



OneSoil

Farm sensors and machinery



## Survey



## Crowdsourcing

Active

Geo-tagged street level imagery



Pl@ntNet

iNaturalist



Mapillary

OpenStreetMap



Social Networks



European Commission



# Farmers' declarations as “in situ” data?



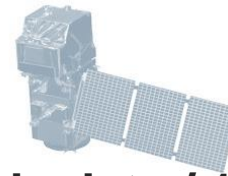
# AI4Boundaries - open data set for parcel extraction and standards' evaluation

## Parcels

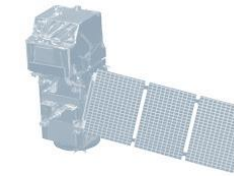
2019 LPIS data for those countries with open data licenses following a stratified sample.

## Images

Images of 256x256 pixels will be selected based on a stratified sampling (eg, using the total segment length) to ensure that a diversity of landscapes are selected.



**Single date / 10 m**  
a set of cloud free  
images centred  
around June

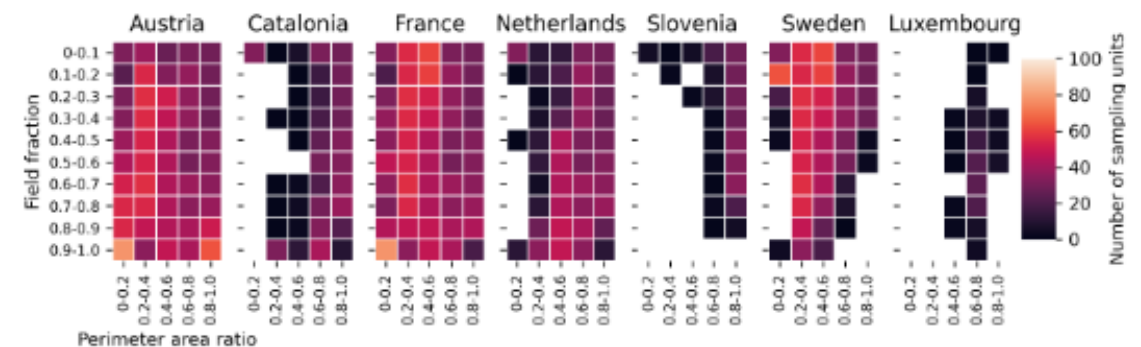
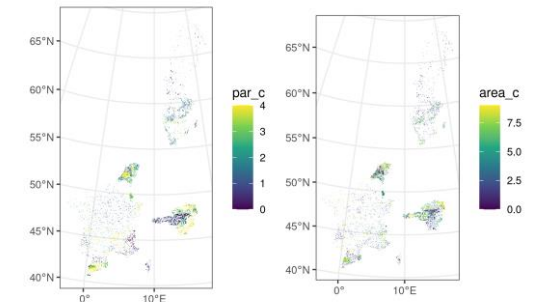


**Multi date / 10 m**  
a set of monthly  
composites

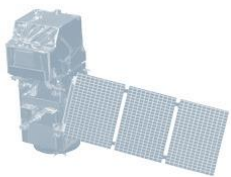


**Single date /  $\leq 1$  m**  
a set of cloud free  
orthophotos

Country/Region	Number of sampling units
Austria	2091
Catalonia	652
France	2078
Luxembourg	132
Netherlands	1157
Slovenia	301
Sweden	1420
Total	7831







**Multi date / 10 m**  
a set of monthly  
composites



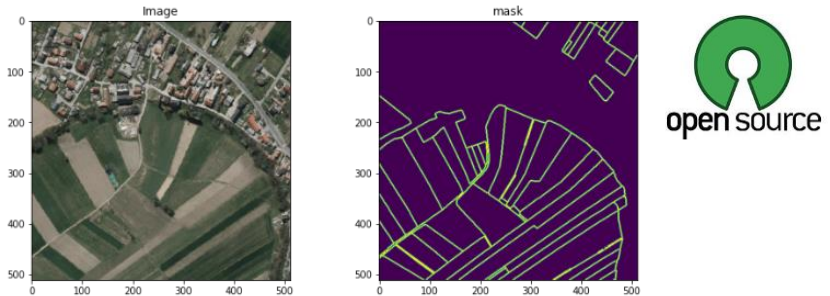
**Single date /  $\leq 1$  m**  
a set of cloud free  
orthophotos



# Main Purpose: A Reusable Framework

1

An open **AI-ready data set** to map field boundaries with Sentinel-2 and aerial photography (*AI4Boundaries*).



2

**Open-source** TensorFlow scripts for creating an optimised data set, coding and running Deep Learning (DL) **architectures**.

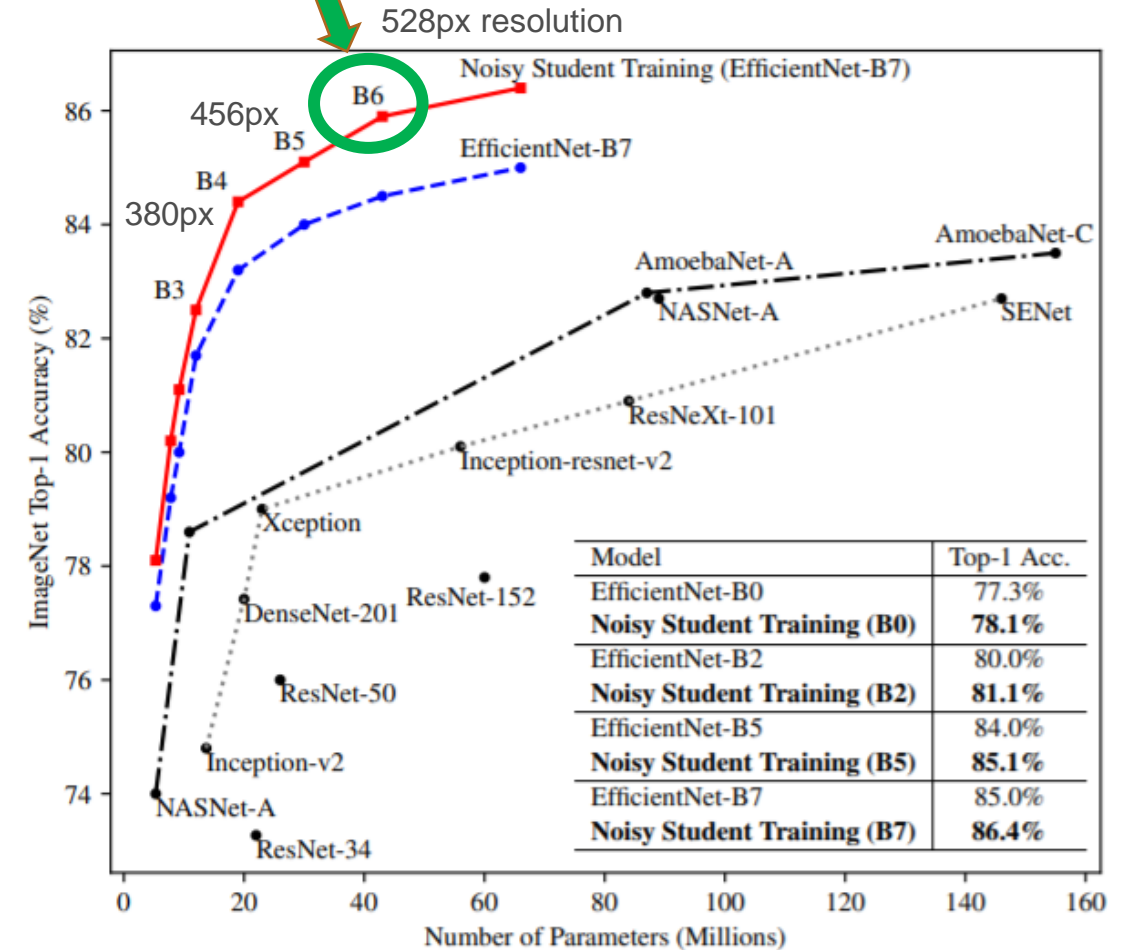
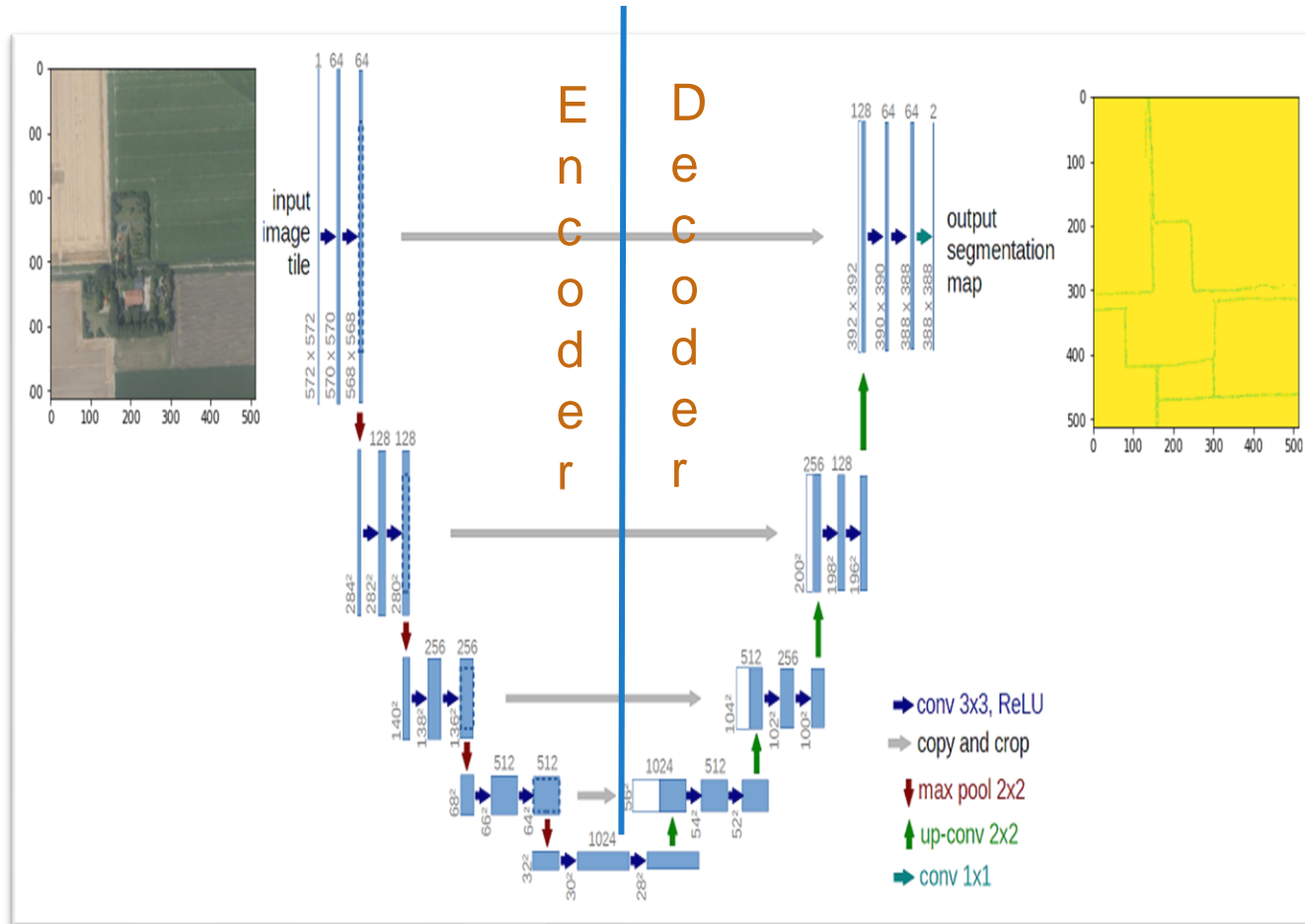
3

Flexible, high-performance **serving system** for deep learning models.





# Base Model: U-Net + EfficientNetB6 + Noisy Student Learning Approach



- T. Agrawal, et al., *EfficientUNet: Modified encoder-decoder architecture for the lung segmentation in chest x-ray images*, Expert Systems, April 2022
- B Baheti et al., *Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment*, Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 2020.

Complicated coverage and standardisation at European level due to

→ Not all countries publish their data for research purposes

→ Demand for a uniform European taxonomy

Based on EU regulation and the EAGLE matrix

Organises all classes into a 6-level hierarchy

Any granularity obtained from the countries can be reflected and stored





# Administration data

## A hidden gold mine

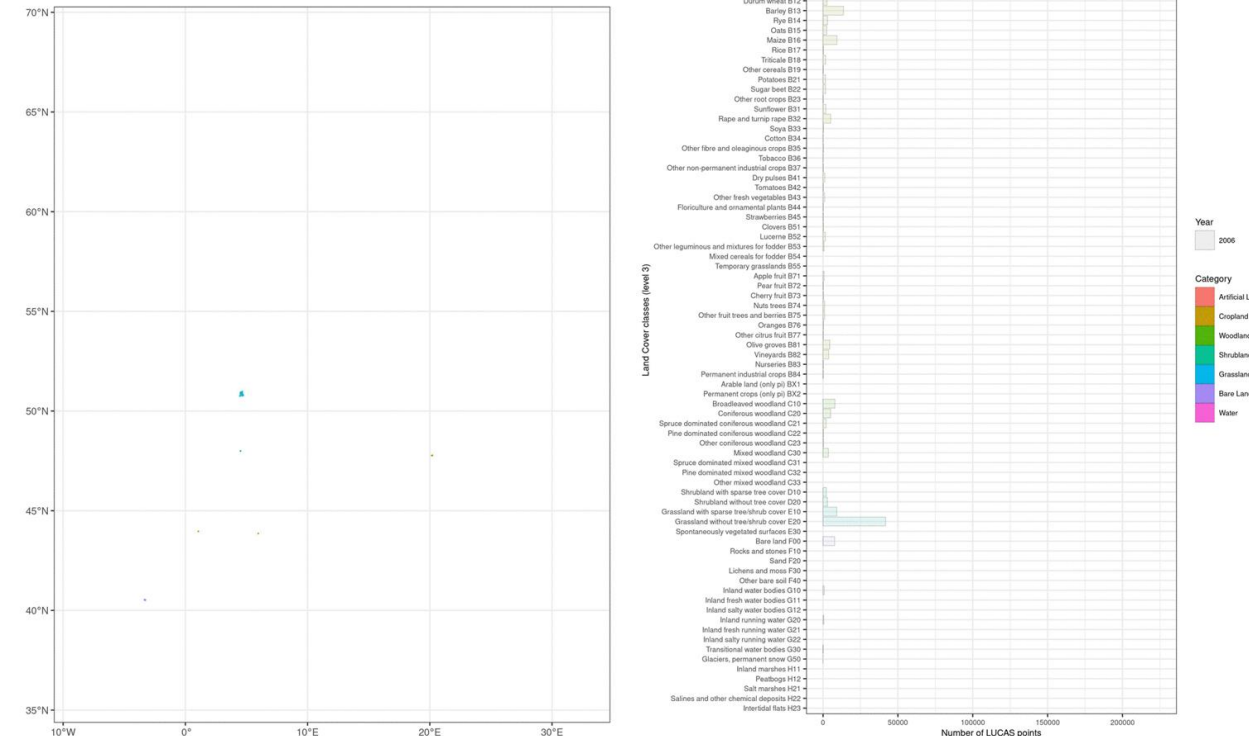


# Harmonising LUCAS - the biggest collection of in-situ data in Europe

**LUCAS** is the **L**and **U**se/**C**over **A**rea frame **S**urvey

- sample every 2km
- 100 variables
- 15 years (2006, 2009, 2012, 2015, 2018)
- 10 K surveyors
- 1.3 M points and 5.4 M photos
- Designed for statistics not for EO

2006-2



1.3 M points



5.4 M photos



# 5.4 M LUCAS PUBLIC PHOTOS (P, N, E, S, W)

Year	East	North	Point	South	West	TOTAL	Location [%]	Orientation [%]
2006	137461	137426	134538	137368	137179	683972	0	0
2009	199208	199264	171165	199129	199117	967883	5.4	0
2012	269329	269286	243074	269277	269205	1320171	34.2	15.3
2015	265421	265392	242772	265368	265285	1304238	68.5	22
2018	237259	237529	215190	237262	236955	1164195	72.9	6.7
Total	1108678	1108897	1006739	1108404	1107741	5440459		

TAB - Number of LUCAS photos per year, per type (N, E, S, W, P) with proportions that have EXIF geo-location (Location [%]) and orientation information (Orientation [%]).

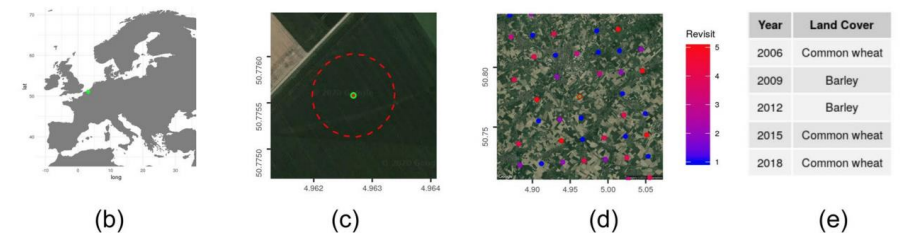
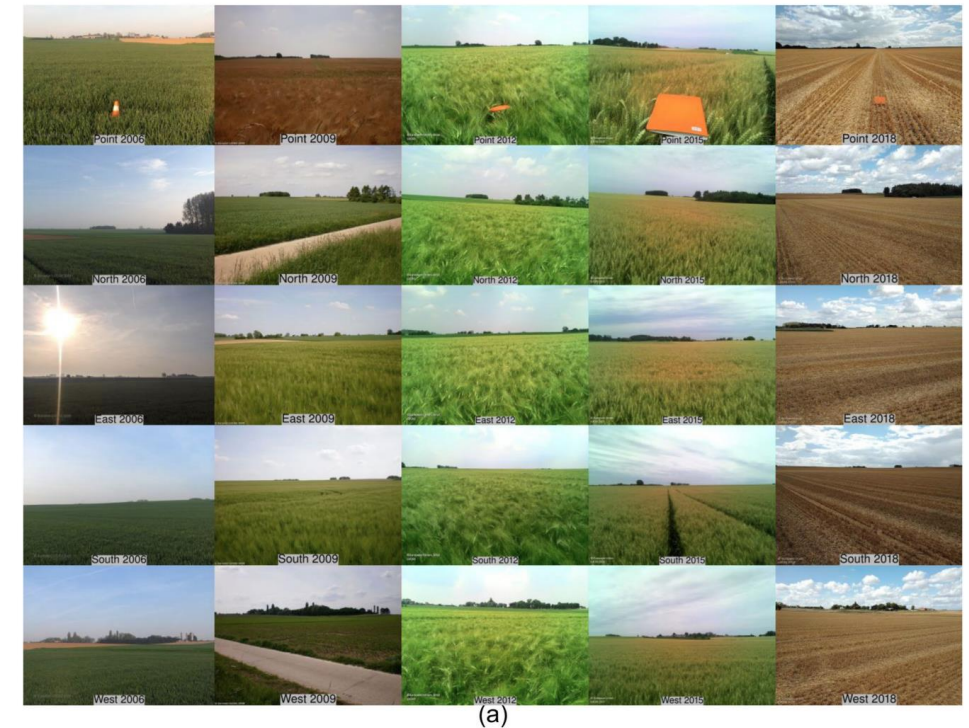
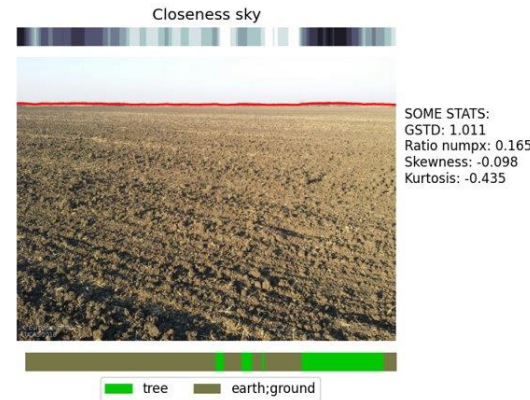
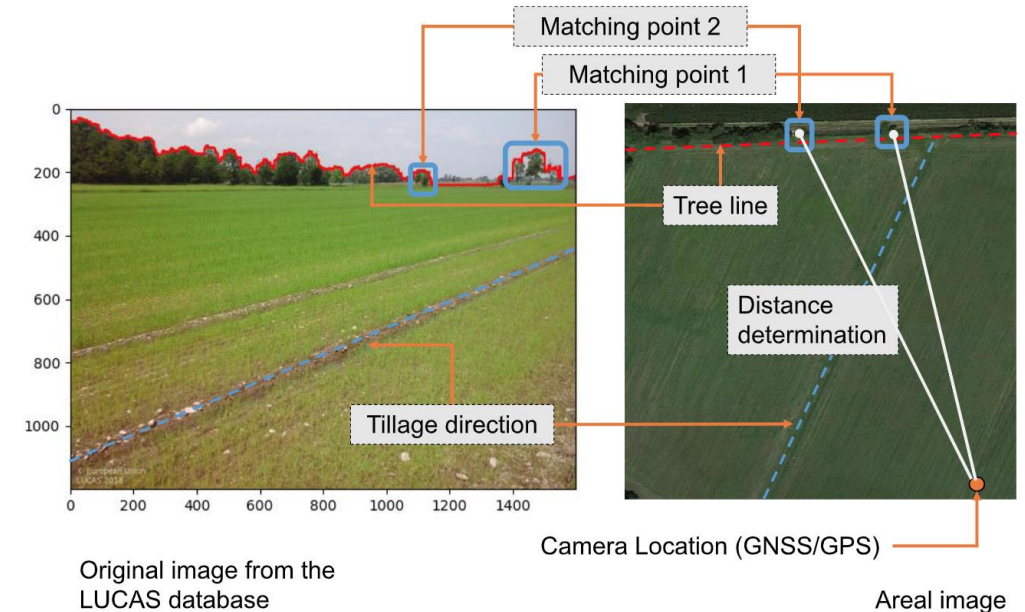


FIG - Overview of the data available for a LUCAS point that was visited five times: (a) Point, North, East, South and West photos for 2006, 2009, 2012, 2015 and 2018, (b) Location of the point in the EU, (c) Zoom showing the point (3-m diameter in green, 50-m diameter in dashed red), (d) Visit frequency on a 20 by 20 km square centered on the point, and (e) In-situ land cover observation of the point for the different years

# Landscape openness and distance estimation thanks to computer vision

- Natural objects that are farther away will have less pixels representing the object
- Approximate the real distance on an object in an image
- Derive the openness of an image relative to the objects in the horizon
- Use of signal processing methods to derive distances on an image

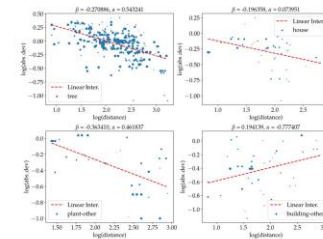


The second metric considered for the analysis is the sample  $\hat{y}_{sky}[x]$  defined as [9]:

$$m_{sv,i} = \frac{1}{D_i - 1} \sum_{x=x_{c,i}+1}^{x_{c,i}+D_i-1} (\hat{y}_{sky}[x] - \bar{y}_{sky})^2$$

where  $\bar{y}_{sky}$  is the sample mean:

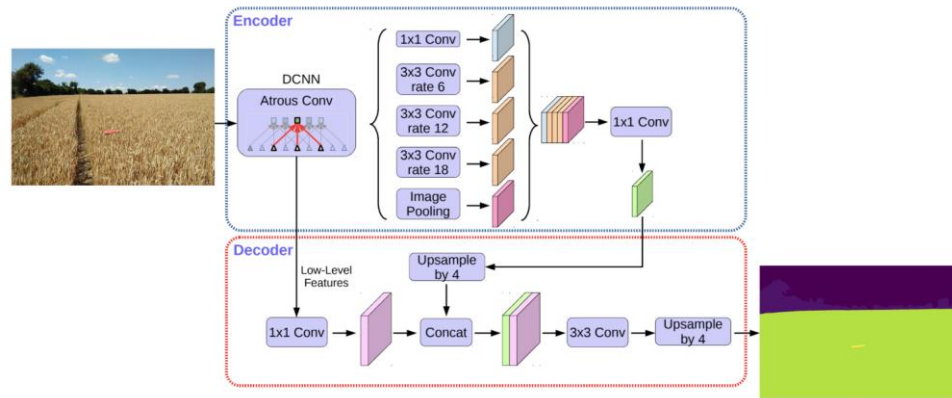
$$\bar{y}_{sky} = \frac{1}{D_i - 1} \sum_{x=x_{c,i}+1}^{x_{c,i}+D_i-1} \hat{y}_{sky}[x].$$



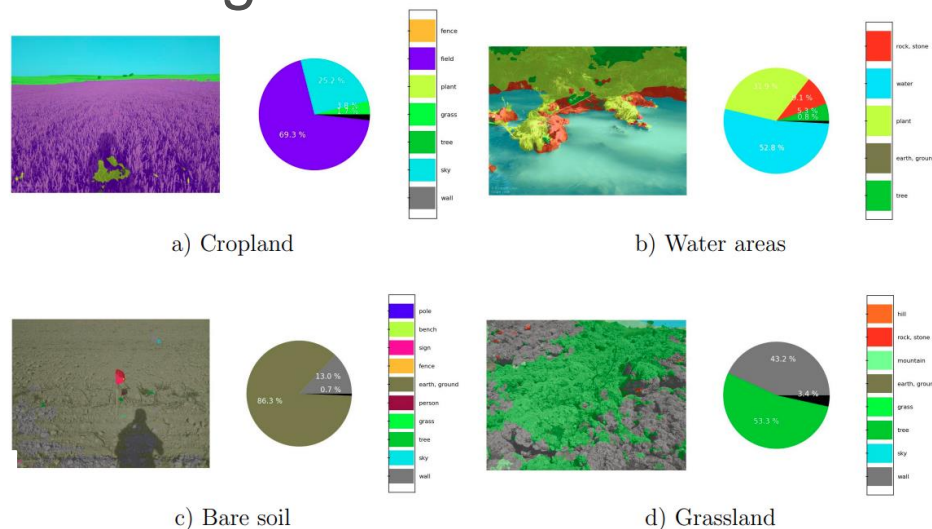


# Semantic segmentation of landscape photos could tell us about land cover

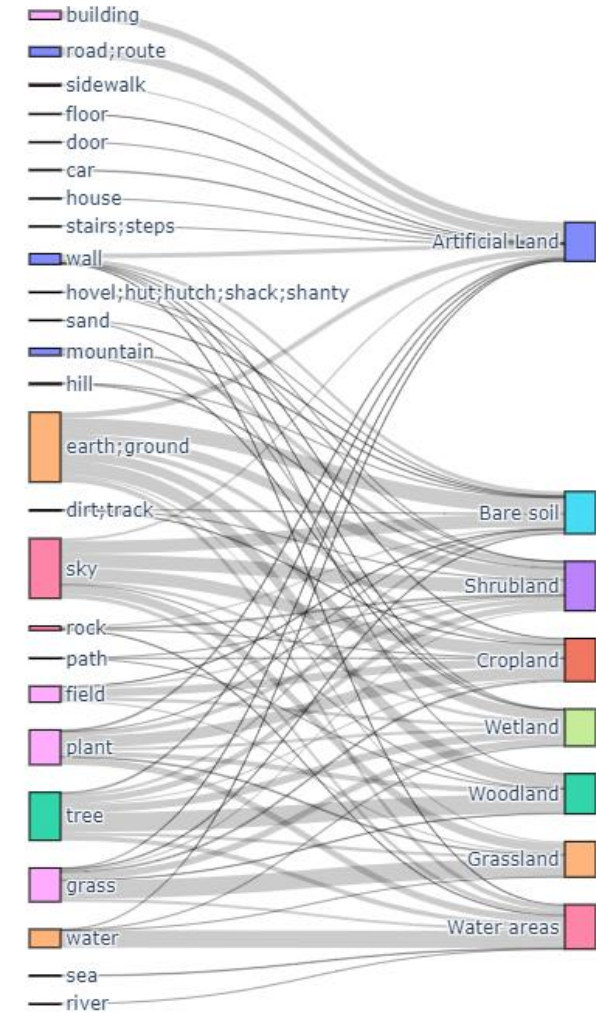
DeepLab v.3 with ADE20K



## Semantic segmentation of LUCAS images



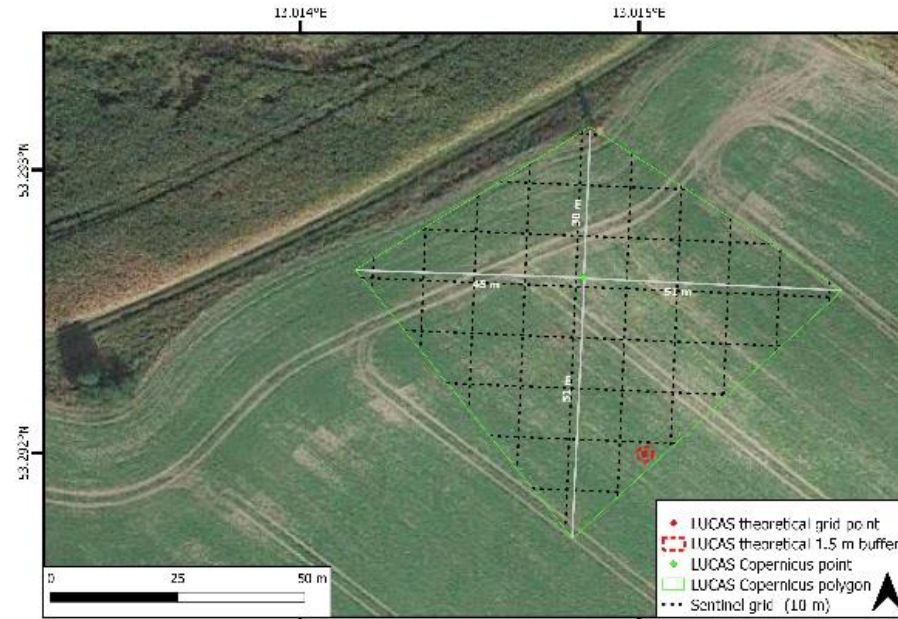
AI OBJECTS FROM LUCAS PHOTOS



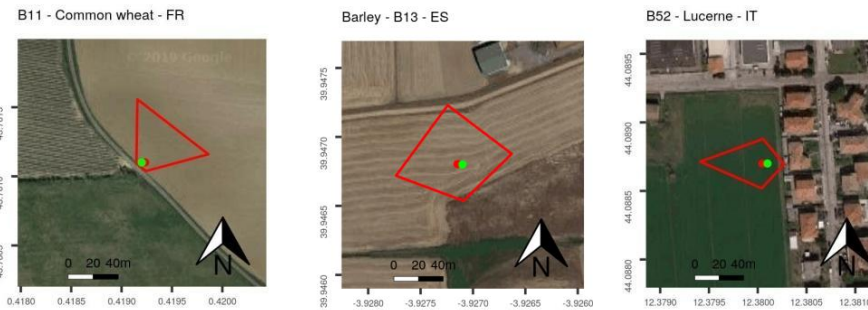
LUCAS LAND COVER

# LUCAS Copernicus an in-situ data for EO training and validation

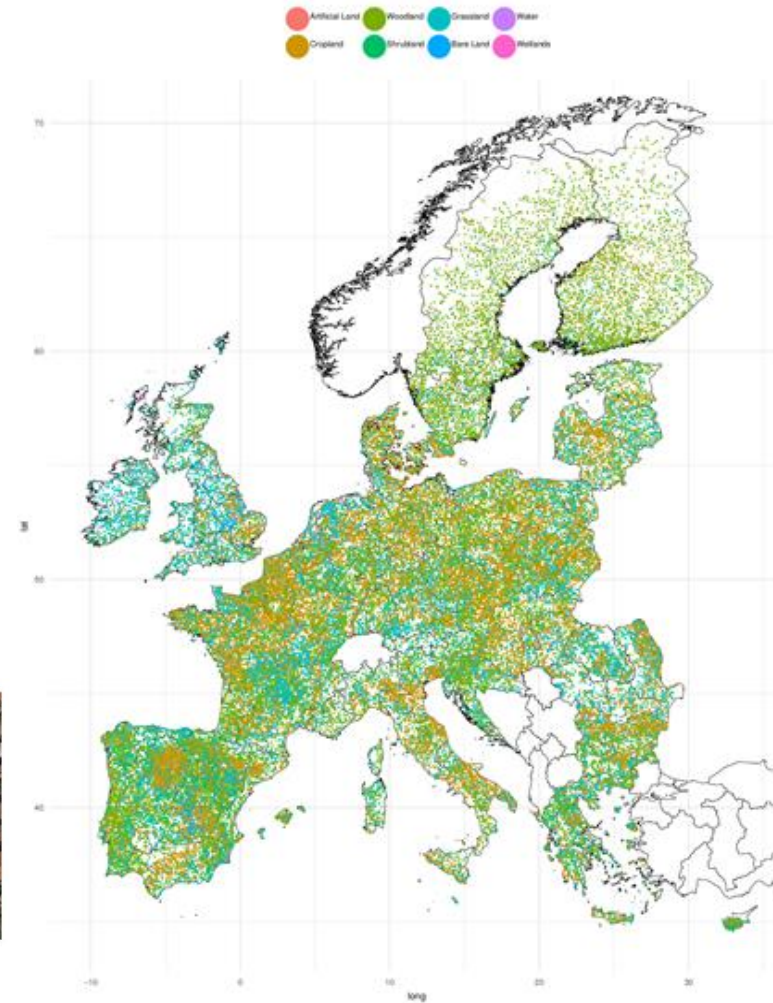
- Copernicus Programme
- Need for in-situ data : ground truth data is still the holy grail of remote sensing
- EO LUCAS module introduced in 2018 to collect surface relevant for EO instead of points
- 1.5-m points not suited for EO
- 60 K “pure” polygons collected
- Average area of 0.3 Ha
- Open Data set in ESSD\*



Example of a LUCAS Copernicus polygon and Sentinel 10-m grid VS LUCAS point

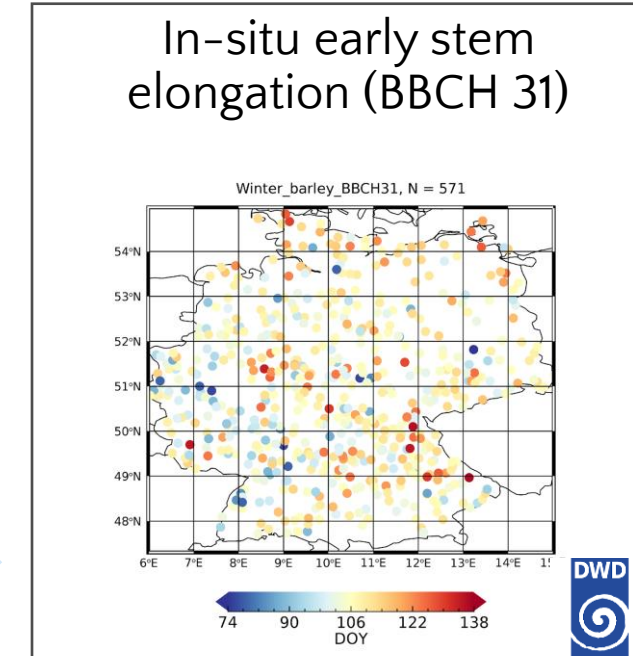
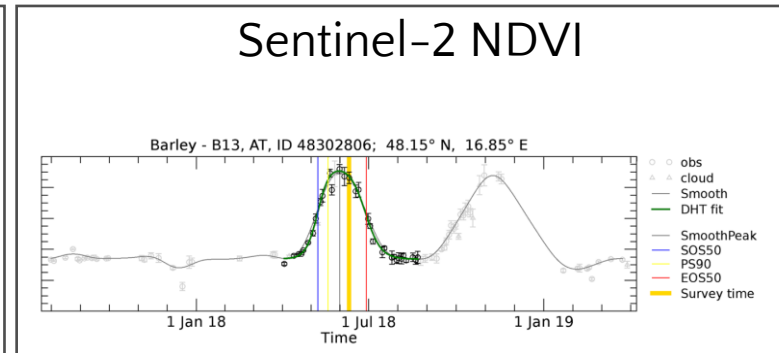
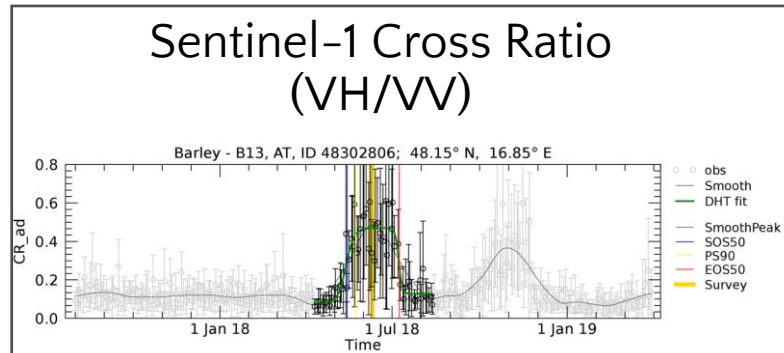


Examples of a LUCAS Copernicus polygon





# Satellite time series are used to obtain Land Surface Phenology and compared with in-situ data



Satellite time series fitting provides Land Surface Phenology metrics (start, peak and end of the season)

Satellite Start of season (SOS)




In-situ start of season

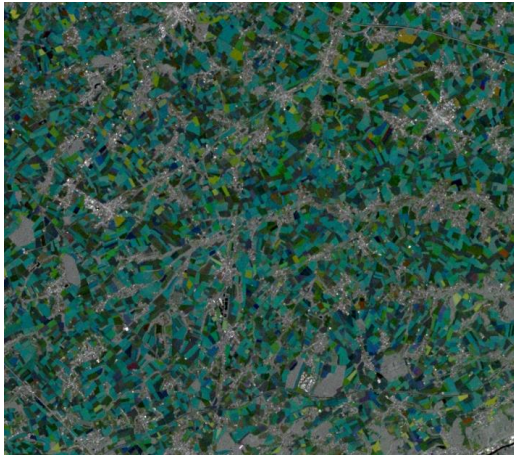
Crop	NDVI						CR					
	SOS50		PS90		EOS50		SOS50		PS90		EOS50	
	BBCH	Δ	BBCH	Δ	BBCH	Δ	BBCH	Δ	BBCH	Δ	BBCH	Δ
Wheat	31	10.0	31	5.5	87	-7.2	31	-4.7	31	6.2	87	-5.7
Barley (winter)	31	-0.1	51	-6.1	87	8.3	31	5.2	51	-6.7	99	-5.4
Barley (spring)	10	-0.3	31	-10.0	87	-8.2	10	5.0	31	-10.6	87	8.2
Rapeseed	51	1.4	61	-1.2	89	-7.5	61	4.7	61	21.4	99	-3.3
Maize	31	4.1	53	-15.1	83	6.5	31	0.1	53	-9.0	99	-1.2
Sugar beet	35	-6.5	35	6.9	99	-56.3	35	-15.6	35	-4.3	99	-34.4

Δ is the difference in days between the average timing of the LSP metric and BBCH stage

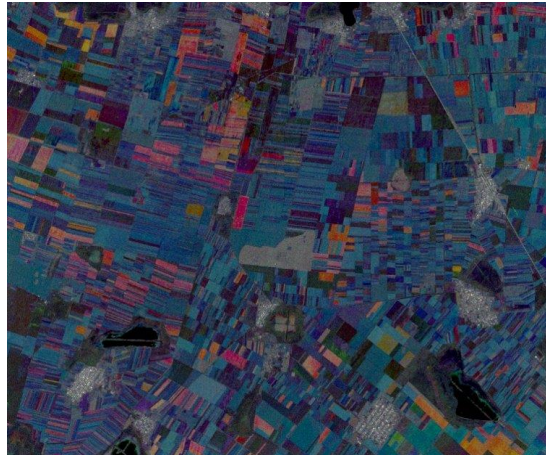
# Sentinel-1 data are consistent at EU and national level

Sentinel-1 GRD pre-processed

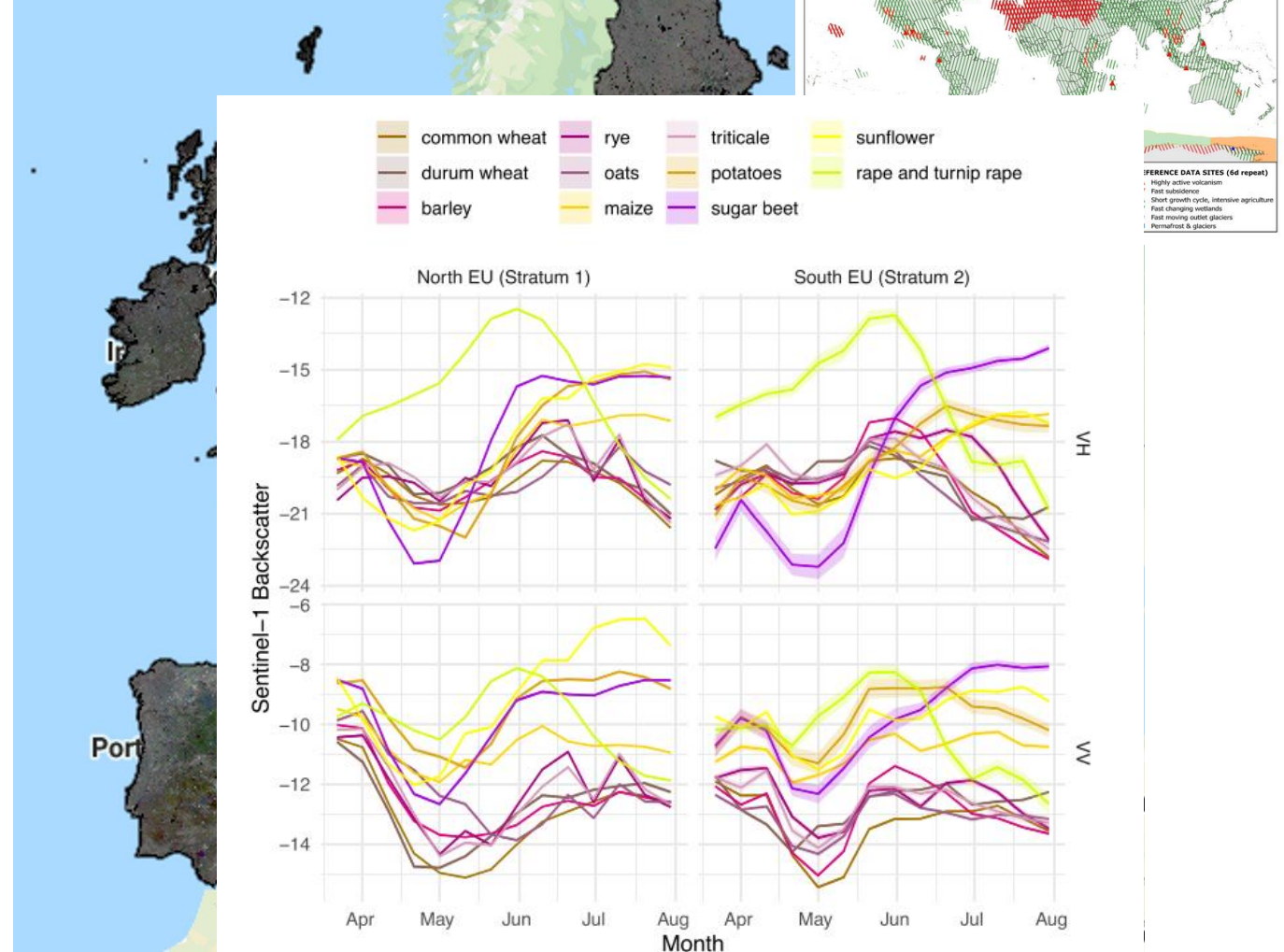
- Google Earth Engine 
- 10-meter pixel spacing (~ 4 ENL)
- Averaged over 10 day periods
- VV and VH backscatter  $\sigma^0$
- 2018



Belgium



Romania



R:VH20180511, G: VH20180610, B:VH20180710



# Crop type classification

•**Training dataset:** Extraction of S1 10-day time series over 58,423 polygons (1,337,682 pixels )

•**Supervised classification:** Random Forest

- Two phase - classification:
- Level 1) land cover
- Level 2) crop types



- **Stratification: 2 strata** (North and South)
- **Parametrization and features selection:**
- Indices and temporal
- 1st January to 31st July 2018
- 4 RF models

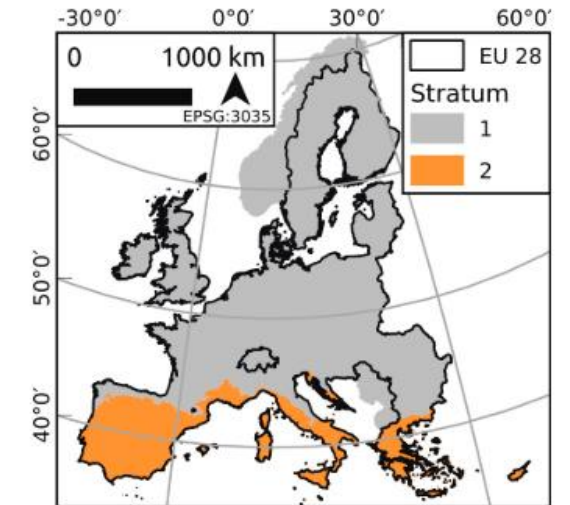
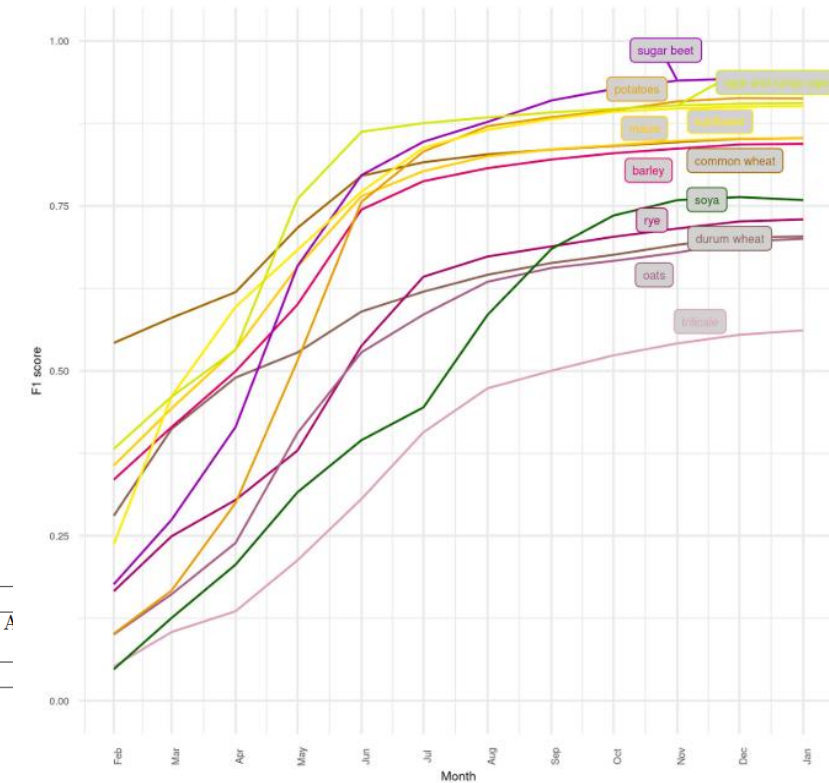
Overall accuracy for S1 indices  
(Jan to July)

Indice	Overall accuracy
VV	71.96
VH	72.70
VH/VV	50.90
VV and VH	79.89
VV and VH and VH/VV	77.91

Legend of the EU crop map

Level1	Level2	
100	Artificial land	A11, A12, A22
200	Arable land	See below
Cereals	211 Common wheat	B11
	212 Durum wheat	B12
	213 Barley	B13
	214 Rye	B14
	215 Oats	B15
	216 Maize	B16
	217 Rice	B17
	218 Triticale	B18
	219 Other cereals	B19
	220 Other roots crops	B23
Root crops	221 Potatoes	B21
	222 Sugar beet	B22
	223 Other roots crops	B23
Non permanent industrial crops	230 Other non permanent industrial crops	B34, B35, B36, B37
	231 Sunflower	B31
	232 Rape and turnip rape	B32
	233 Soya	B33
Dry pulses, vegetables and flowers	240 Dry pulses, vegetables and flowers	B41, B42, B43, B44, B45
	250 Other fodder crops (excl. temp. grasslands)	B51, B52, B53, B54
Bare arable land	290 Bare arable land	F40*
300	Woodland and Shrubland type of vegetation	B71-B77, B81-B84, C10, C21, C22, C23, C31, C32, C33, D10, D20
500	Grassland (permanent and temporary)	B55, E10, E20, E30
600	Bare land and lichens/moss	F10, F20, F30, F40**

\*U111/112/113 (agriculture), \*\*other than U111/U112/U113



# The EU crop map



From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations

Raphaël d'Andrimont<sup>1,\*</sup>, Astrid Verhegghen<sup>1</sup>, Guido Lemoine, Pieter Kempeneers, Michele Meroni, Marijn van der Velde

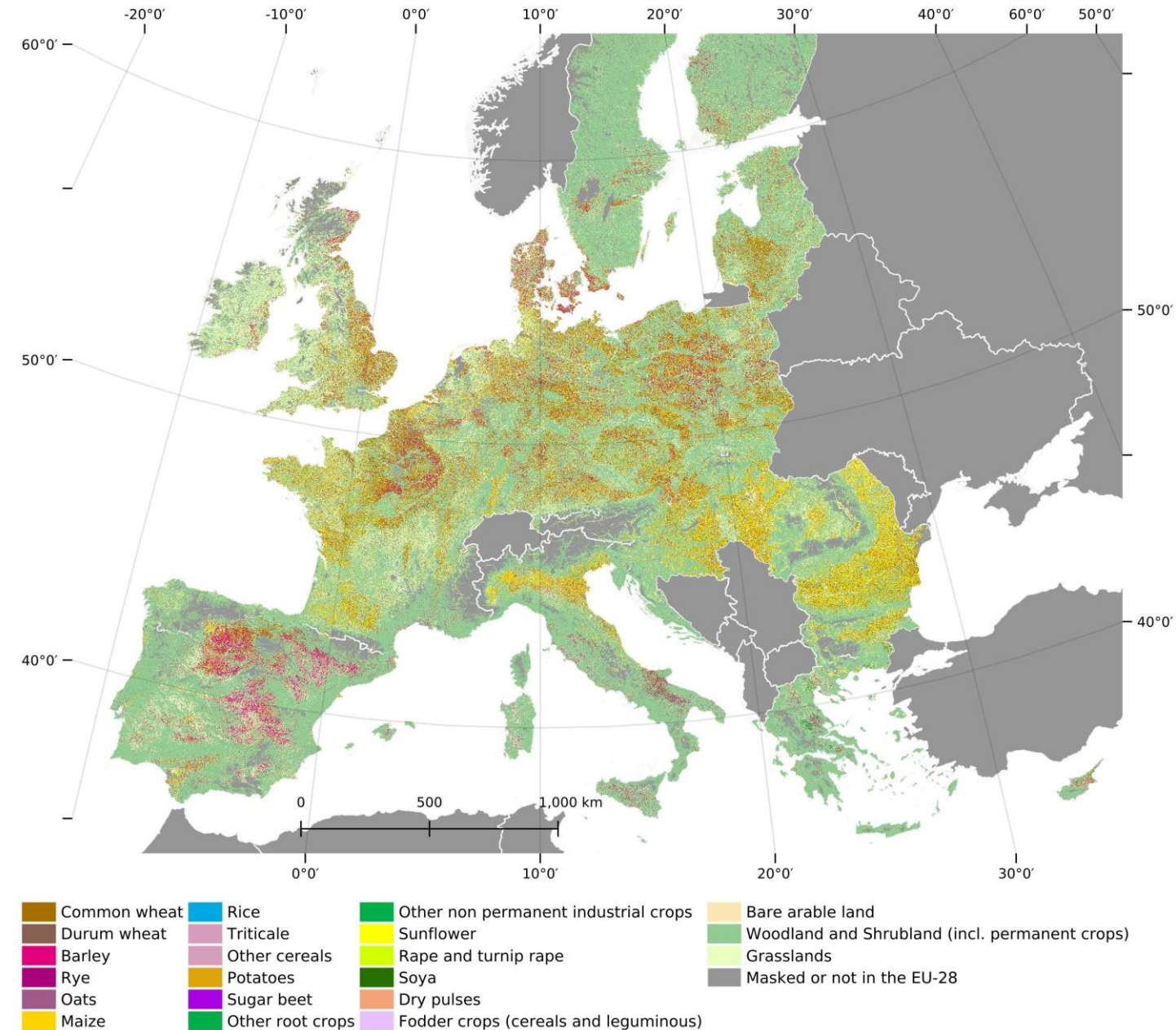
<sup>1</sup>European Commission, Joint Research Centre (JRC), Ispra, Italy

## Data access

EU CROP MAP 2018 FTP			
	Use conditions	<a href="#">European Commission reuse notice &gt;</a>	<a href="#">URL</a>
	Access conditions	<a href="#">No limitations &gt;</a>	
WMS			
	Use conditions	<a href="#">European Commission reuse notice &gt;</a>	<a href="#">URL</a>
	Access conditions	<a href="#">No limitations &gt;</a>	

## [Map download and WMS](#) [Code on Github](#)

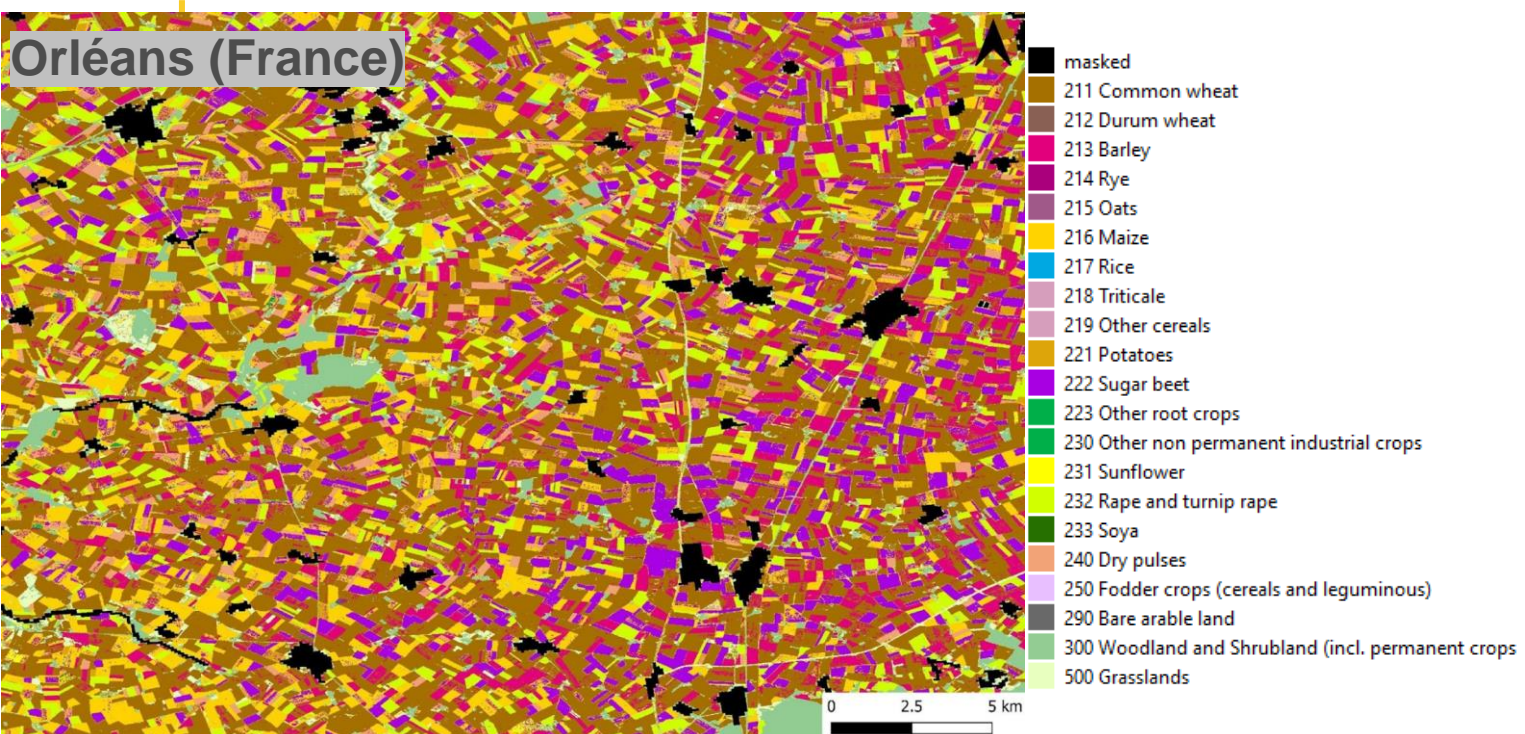
*EU 2018 crop map at 10m pixel spacing (1st January to 31st July 2018). Masked with non-vegetation Corine Land Cover classes.*



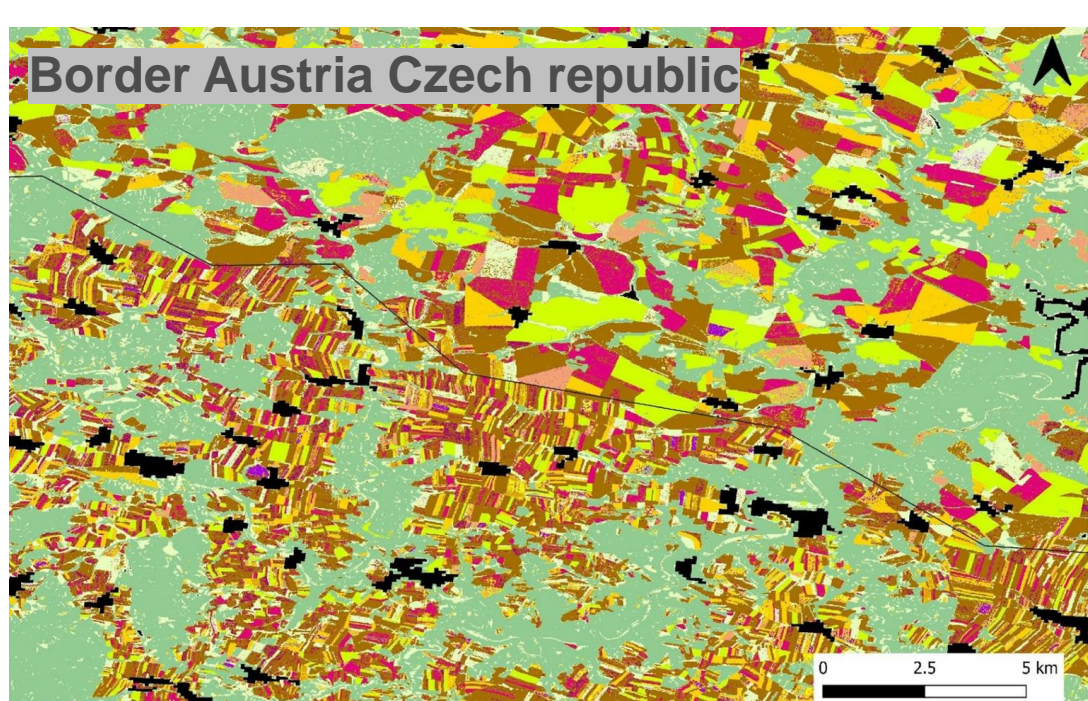
d'Andrimont, R.; Verhegghen, A.; Lemoine, G.; Kempeneers, P.; Meroni, M.; van der Velde, M. [From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations](#). *Remote Sens. Environ.* 2021, 266, 112708.



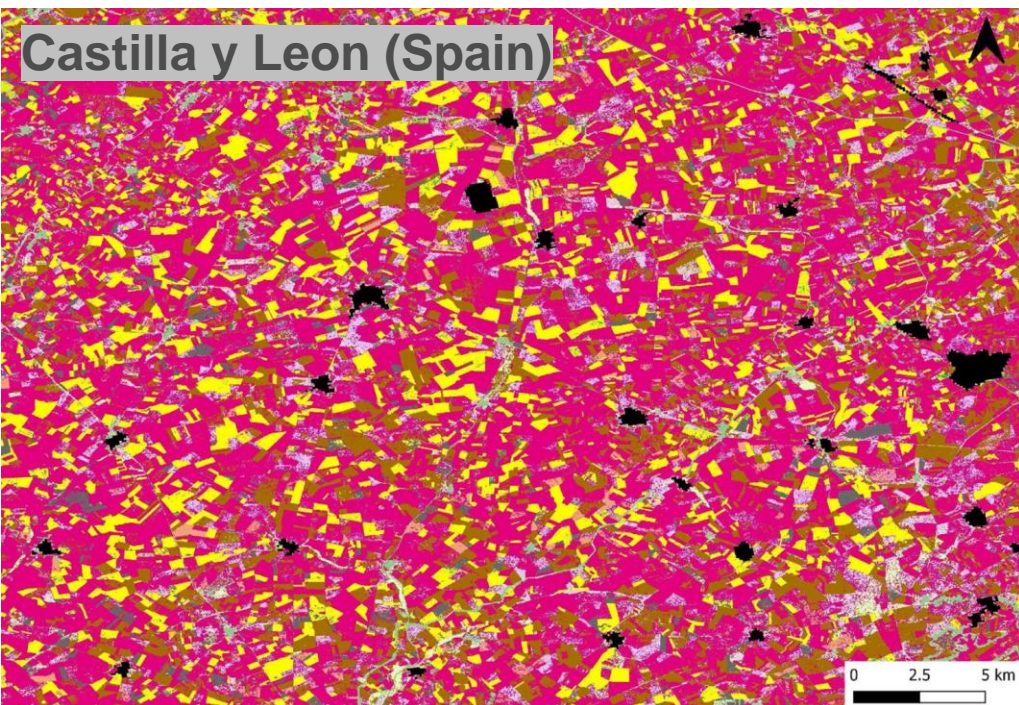
Orléans (France)



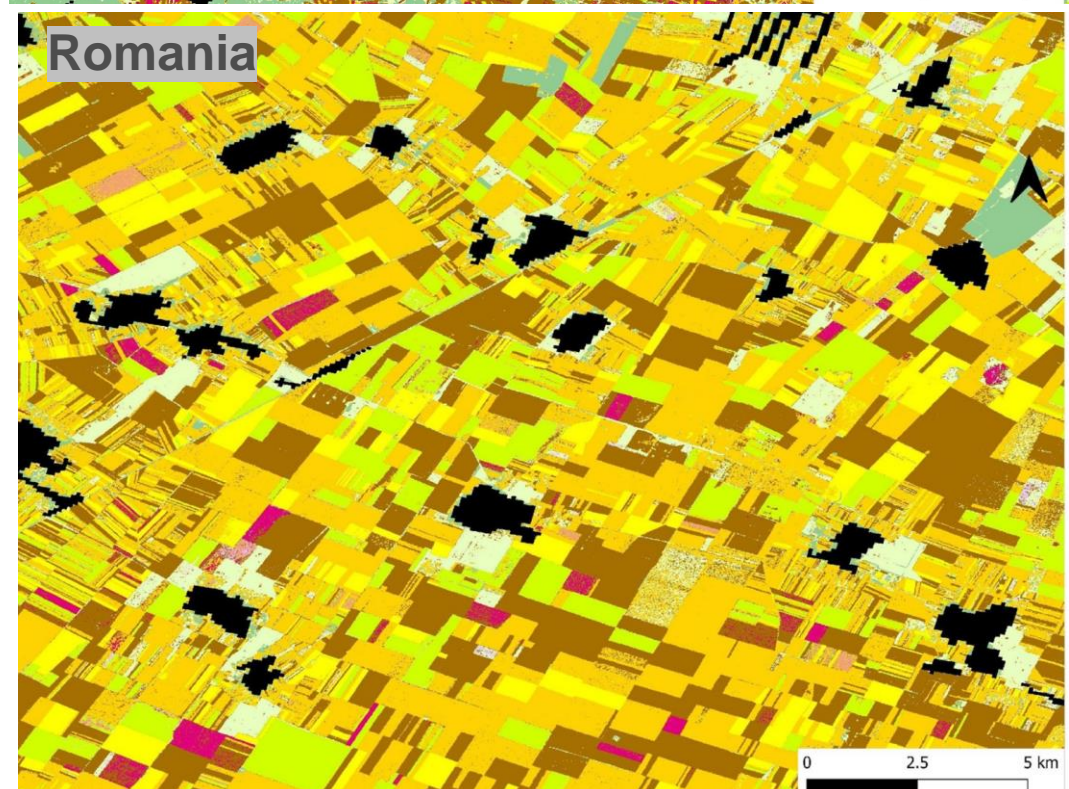
Border Austria Czech republic



Castilla y Leon (Spain)



Romania

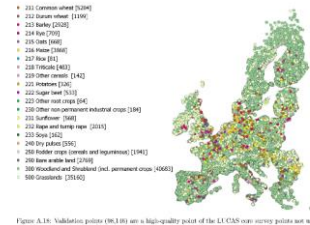




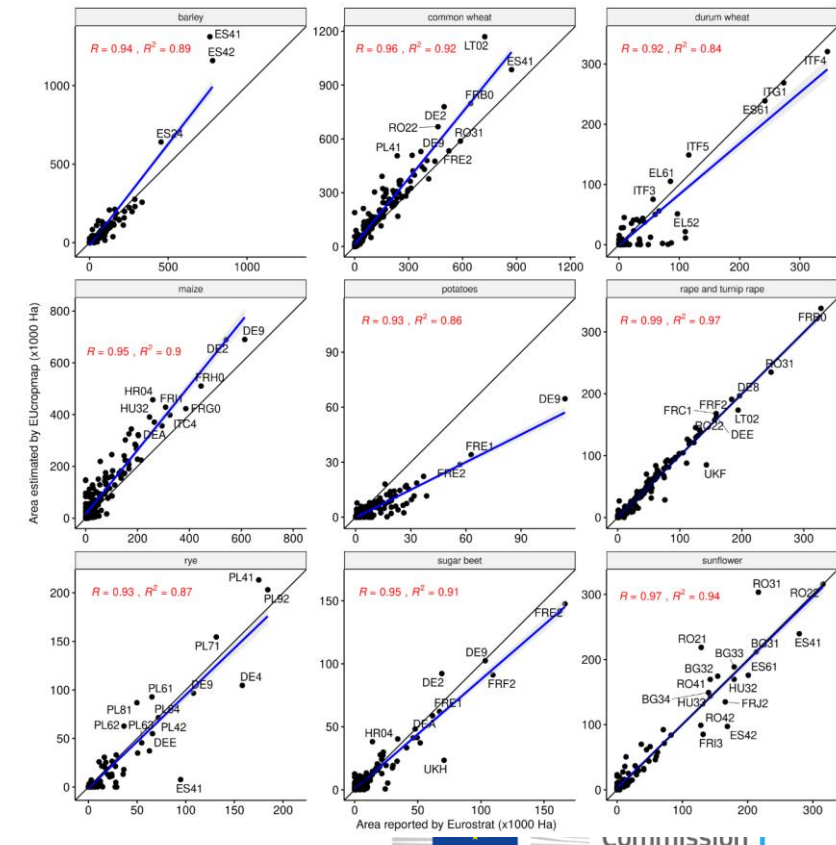
# Robust assessment

- Validation with 87,853 points : accuracy of 80.3% (main crop type groups) and 76% (19 crop types)

Map Class	Reference class (LUCAS point)									UA (%)	SE (%)	OA(%)
	210	220	230	240	250	290	300	500				
210 Cereals	0.13	0.003	0.002	0.002	0.005	0.009	0.002	0.026	72.3	0.6		80.3
220 Root Crops	0	0.006	0	0	0	0	0	0	90.1	2		
230 Non permanent industrial crops	0.001	0	0.024	0.002	0	0.001	0	0.001	79.5	6.2		
240 Dry pulses, Vegetables and Flowers	0	0	0	0.002	0	0	0	0	43.6	4.4		
250 Fodder Crops	0.001	0	0	0	0.001	0	0	0.002	28.8	3.3		
290 Bare Arable Land	0.001	0	0	0	0	0.006	0.001	0.003	50.3	2.7		
300 Tree and Shrub Cover	0.01	0	0	0.001	0.004	0.004	0.483	0.077	83.4	0.4		
500 Grassland	0.01	0	0	0	0.009	0.001	0.013	0.151	82.1	0.5		
Producer accuracy (%)	84.4	58.6	89.3	23.6	6.9	25.7	96.9	58.1				
Standard Error (%)	0.8	3	1.3	3.9	0.8	1.7	0.1	0.6				



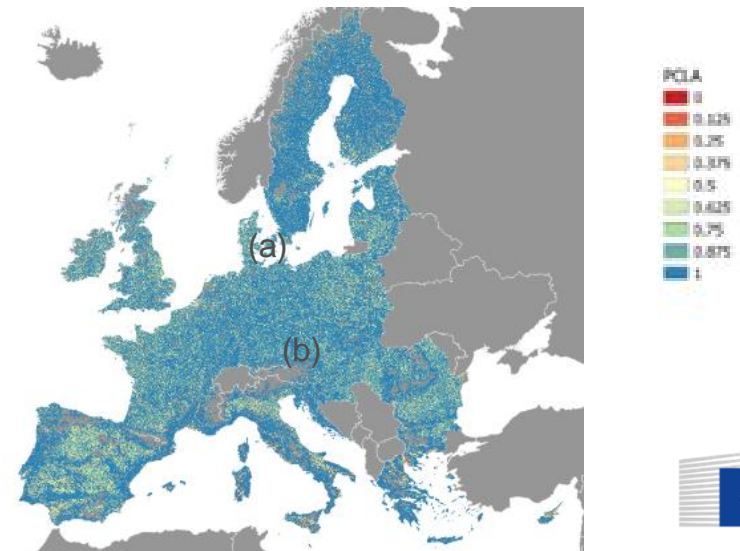
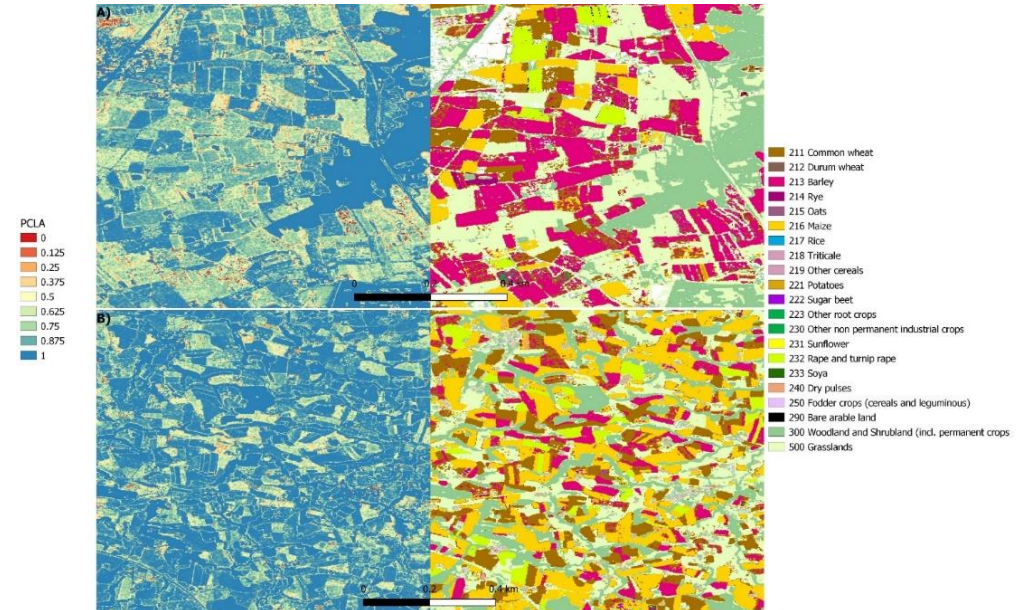
- Comparison with farmers' declarations 3.1 M parcels  
BE fl, DK, FR cv, NL, DE nrw, SI
- Comparison with Eurostat area reporting at nuts-2 level : R from 0.93 (potatoes) to 0.99 (rapeseed)





# Per Pixel Land Cover Accuracy (PLCA)

- Methodology by Ebrahimi et al. (2021) to derive a "per-pixel land cover accuracy" (PLCA)
- Implemented prediction for the EU crop map
- 90% of the 87,853 LUCAS 2018 filtered core point
- For each class: RF to establish a non-linear relationship between the LUCAS dataset (converted to binary) and the VV and VH time series
- Processing at the EU level at 10 meter
- One PLCA value is available for each 10-m pixel of the EU crop map ([download](#))



Verhegghen, A.; D'Andrimont, R.; Waldner, F.; Van der Velde, M. Accuracy assessment of the first EU-wide crop type map with LUCAS data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2021; 2021.

# LUCAS Cover dataset – a hidden gold mine?

- Detailed close-up view of sampled tree, crop and plant species.
- 875,661 photos collected between 2006-2018.
- Pure land cover photos – absence of other landscape elements.
- Anonymized.
- Ideal for AI-based workflows.

Year	Total number of photos surveyed			Flagged photos			Total number of photos published
	source	no harmo	no exif	1st step	2nd step	corrupted	
2006	107,140	54	3	1,007	63	1	107,022
2009	150,125	39	9,559	2,239	80	13	149,993
2012	204,944	0	9,652	3,930	88	50	204,806
2015	217,638	0	1,654	4,654	28	1	217,609
2018	195,814	511	2,085	5,050	87	1	195,216
<b>Total</b>	<b>875,661</b>	<b>604</b>	<b>22,953</b>	<b>16,880</b>	<b>346</b>	<b>66</b>	<b>874,646</b>

Limitations include – heterogeneity in terms of field of view, lack of EXIF information.  
Perspectives include – historical analysis from point revisit, (crop) species classification, (crop) organ segmentation.





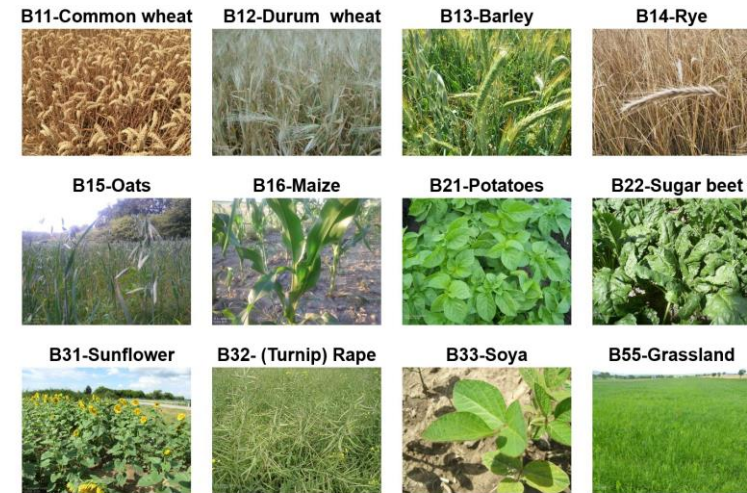
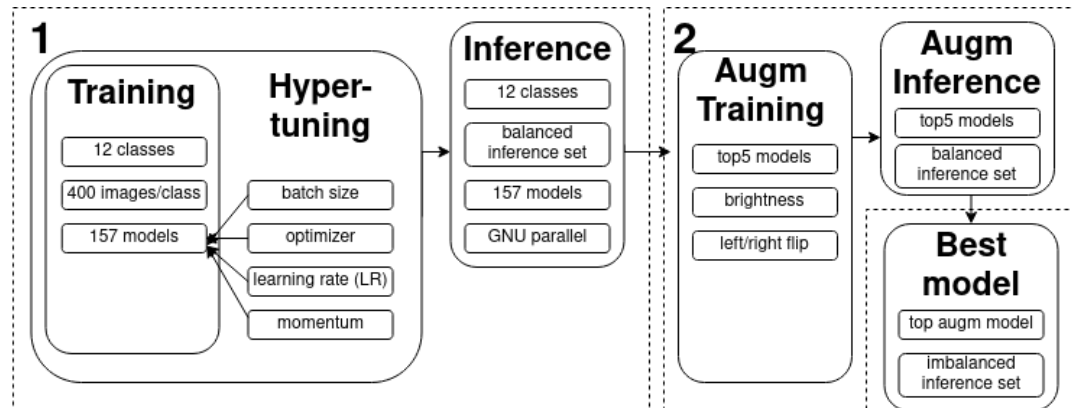
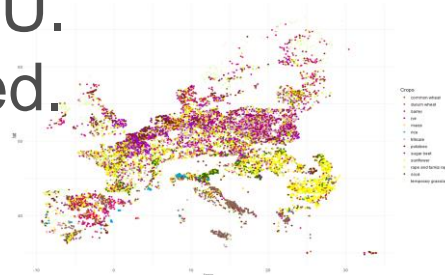
# LUCAS Cover dataset





# Crop type detection

- Crop classification via DL on LUCAS Cover photos.
- Crop calendar harmonization and data extraction for 12 major mature crops in EU.
- 157 models trained.
- M-F1: 0.757
- M-F1\_En: 0.817

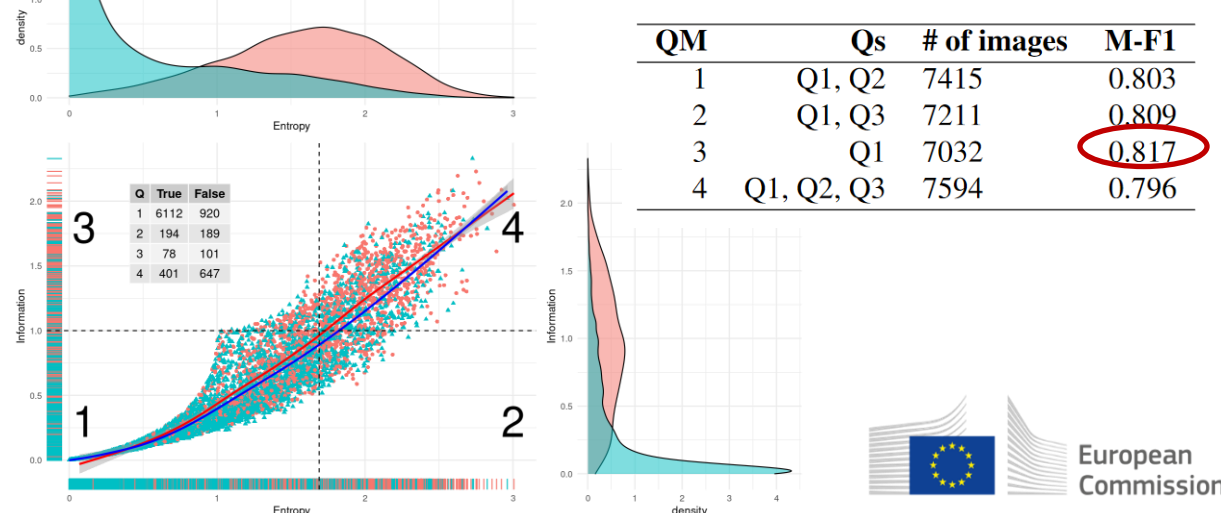


Ranking	1	2	3	Best
Model	78	88	4	78
Level	Augm	Augm	Augm	Best Model
LR	0.0035	0.0073	0.0096	0.0035
BS	1024	512	512	1024
Momentum	0	0	0	0
Optimizer	GD	GD	GD	GD
# of Images	1020	1020	1020	8642
Validation Accuracy	0.7768	0.7789	0.7747	0.7768
Training Accuracy	0.8945	0.8965	0.9238	0.8945
Test Accuracy	0.7941	0.7775	0.7755	0.7854
M-F1	0.7941	0.7775	0.7755	0.7572

- Un-favourable conditions (B11)



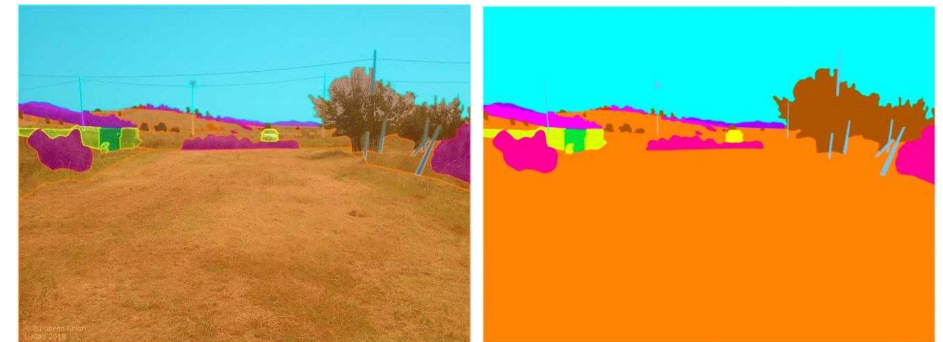
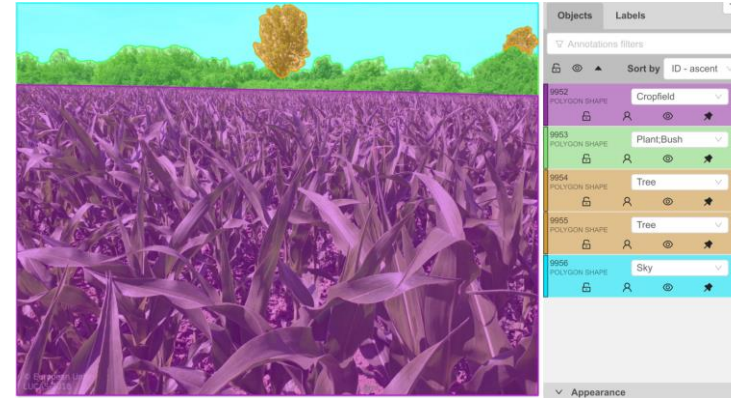
- Entropy-based post-processing





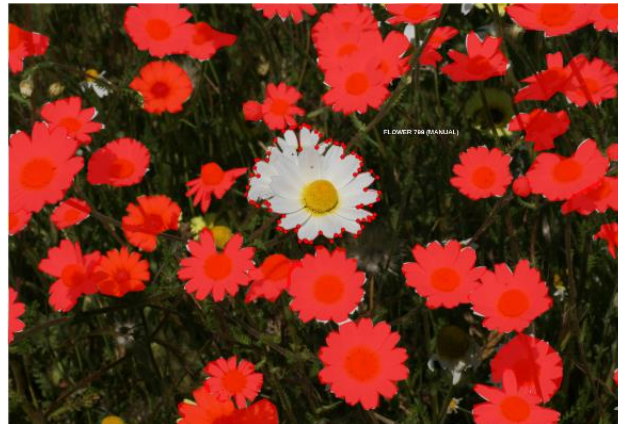
# Semantically segmented dataset focus on agricultural landscapes towards landscape features

- Sampling of LUCAS 3,000 landscape photos
- 29 classes (13 general purpose, 16 to describe the landscape, cropfield, flowerfield, ... )
- 10,519 objects manually delineated
- Quality checks ongoing



# Recognizing flowers on grassland images

- Creating computer vision dataset: 500 images with 9524 manually delineated flowers
- Training and tuning Faster R-CNN to detect flowers
- Using model predictions to extract flower abundance and colors from images
- Identifying individual species with Pl@ntNet API
- Indicator species in relation to pesticide applications?



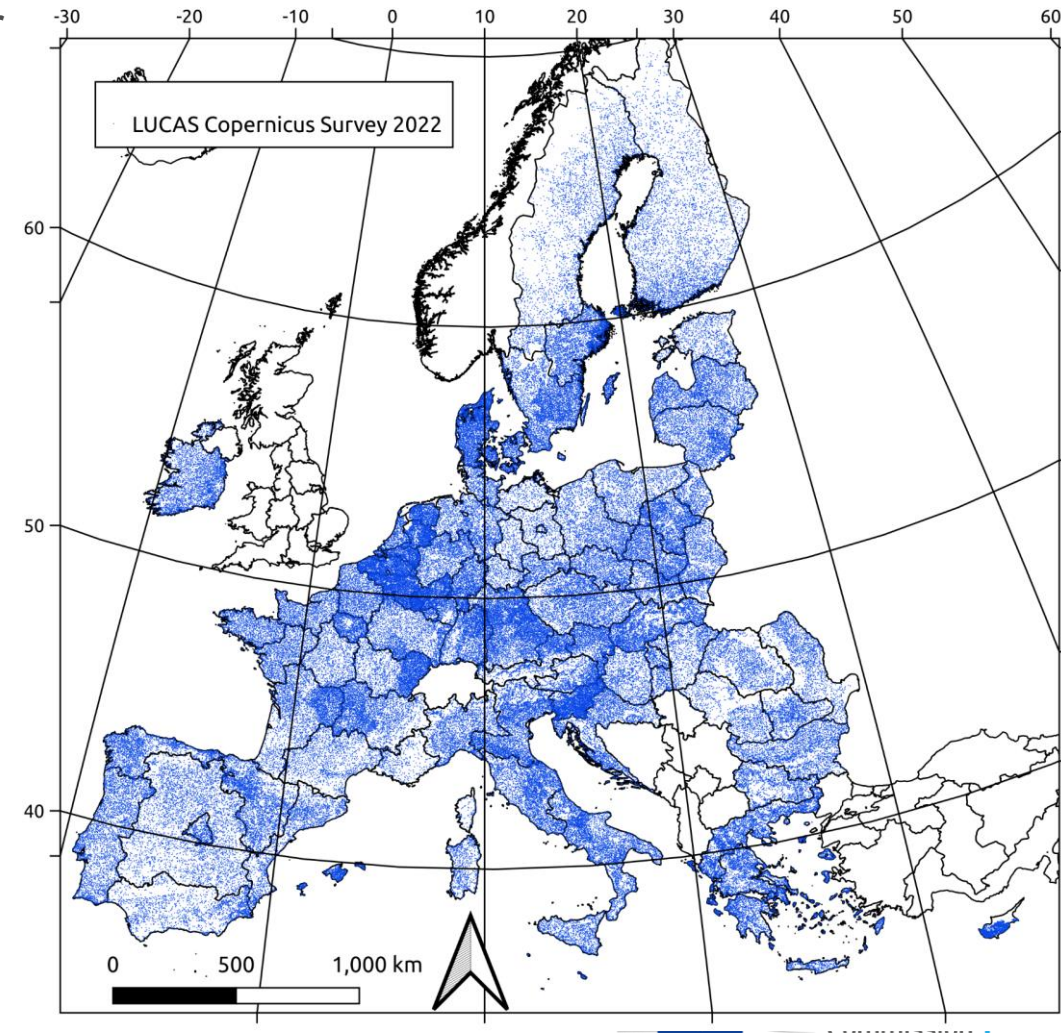


# LUCAS 2022 - 400,000 points

A) Field survey and its components 200.000 (March – November 2022)

- **Copernicus** observation (150.000 pts)
- **Soil** module 41.000 pts
  - Bulk density measures (4.000 pts),
  - Biodiversity sample (2.000 pts)
  - Depth 30 cm
  - Gully erosion on all points
- **Grassland** module (20.000 pts), “20m grassland transect”
- **Landscape Features** (93.000 pts)
- **Extended grassland** parameters (40.000 pts)

B) Photo-Interpretation in office - 200.000 (by March 2023)



# LUCAS LF module

...a **new** element in **LUCAS 2022**...

in **93,000** LUCAS points

(and 41 subpoints: **100x100 m quadrat**)

spatial **representativeness** at MS level

(and possibly also NUTS2),

**consistent quantification** of LFs for the EU and MS level,

with information on different **LF types**,

compatible with data from other sources at EU level

...with a two-stage approach

office-based **photo-interpretation** (PI, phase 1)

**field survey** (SU, phase 2)

**JRC expertise**

**Functional definition** of LF (based on SO6)

Harmonized & **simplified typology**

**Practical guidelines** (for field survey & photointerpretation)

→ goal: a reliable & fit for purpose I.21

*CZUCZ, B. et al.*

Small fragments of non-productive **semi-natural vegetation in agricultural landscape** which provide ecosystem services and support for biodiversity

1 LUCAS base grid

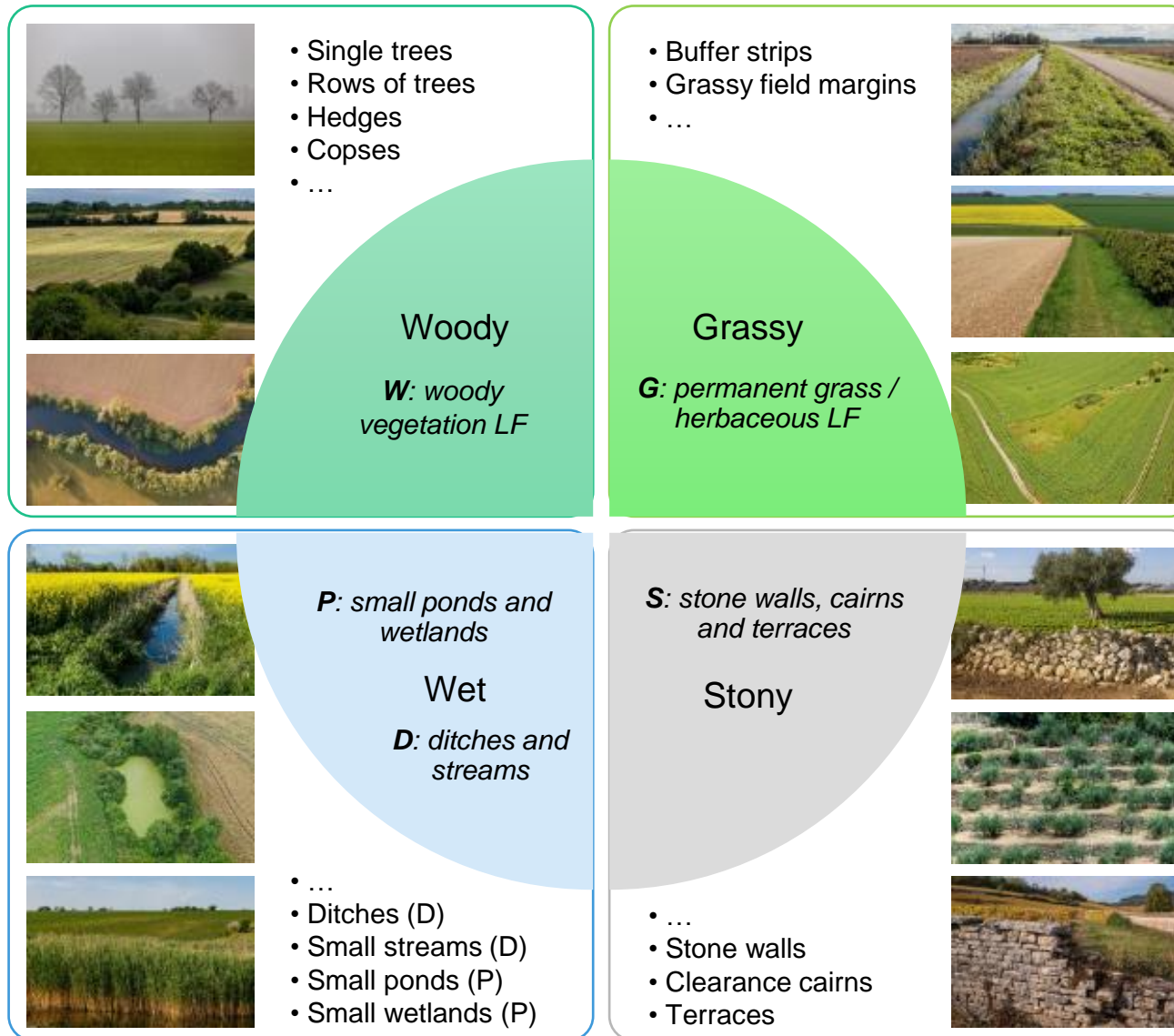
2 Sampling points

3 41 subpoints  
(for the LF module)





# LF types in LUCAS



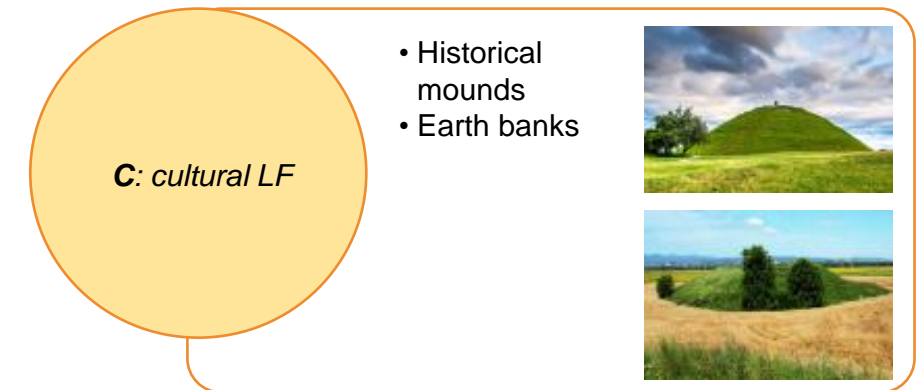
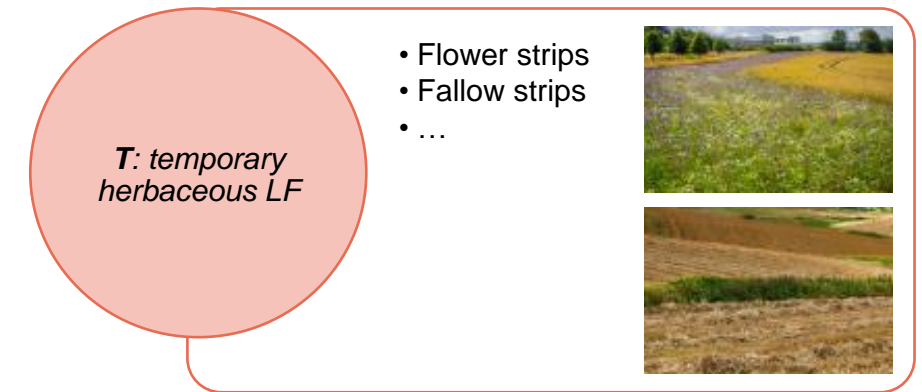
LF in the EU:

<https://publications.jrc.ec.europa.eu/repository/handle/JRC128297>



LF in the MS:

<https://publications.jrc.ec.europa.eu/repository/handle/JRC128876>



# EMBAL: European monitoring of biodiversity in agricultural landscapes

## EMBAL

**Only areas under agricultural use** are surveyed in detail:

- Arable land
- Permanent grassland and permanent pasture
- Permanent crops
- Landscape elements, minimum width of 1m:

### Robust:

- ⇒ Harmonized across EU
- ⇒ Operational / feasible
- ⇒ Repeatable
- ⇒ Meaningful



*EMBAL elements:*

- (1) Plots
- (2) Parcels and landscape elements
- (3) Vegetation transects

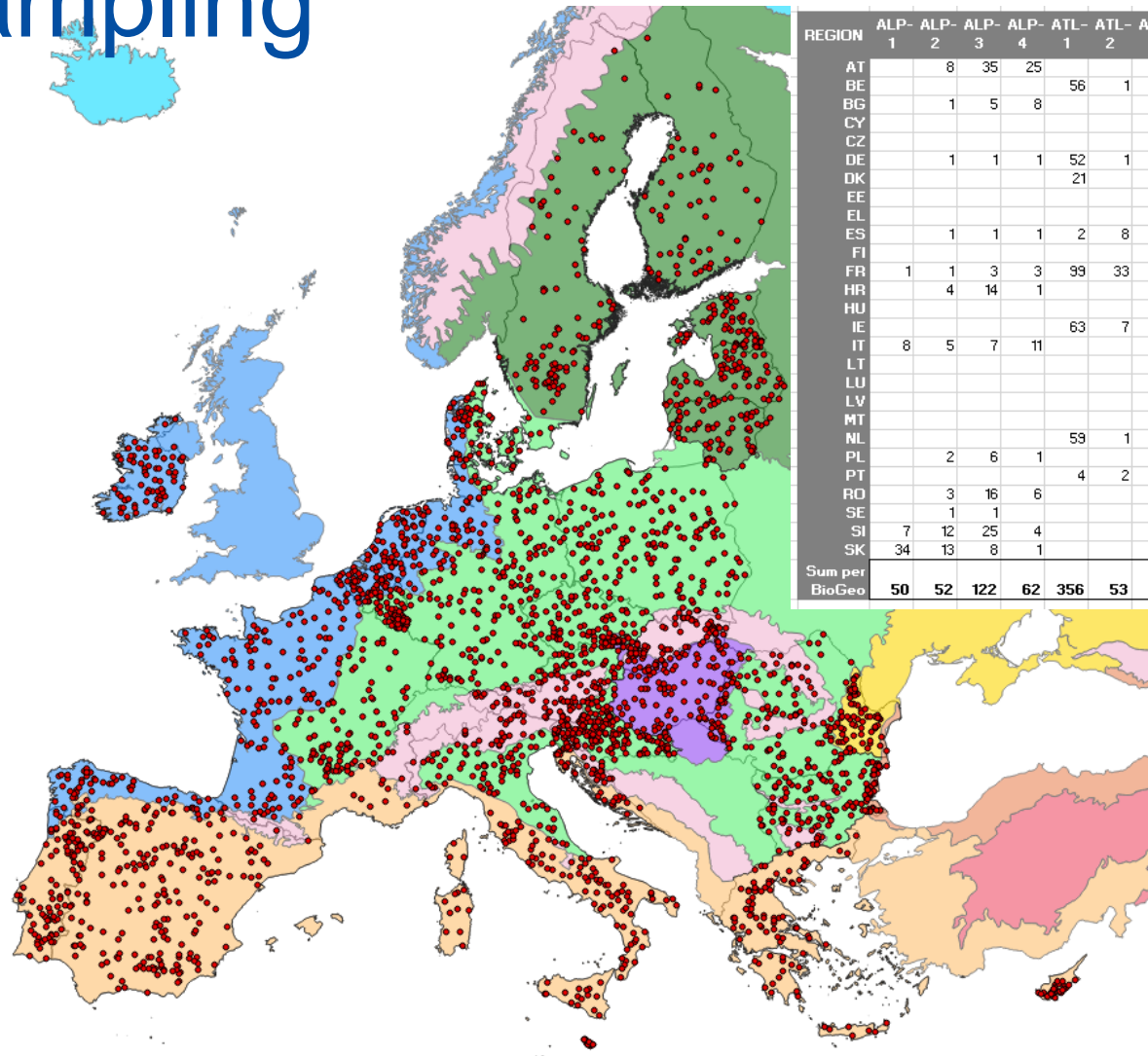
### References are made to:

Oppermann, R., Aguirre, E., Bleil, R., Calabuig, J. D., Šálek, M., Schmotzer, A., & Schraml, A. (2021). A Rapid Method for Monitoring Landscape Structure and Ecological Value in European Farmlands: the LISA approach. *Landscape Online*, 90, 1-24. <https://doi.org/10.3097/LO.202190>.

Sutcliffe, L., Schraml, A., Eiselt, B. & Oppermann, R. (2019). The LUCAS Grassland Module Pilot – qualitative monitoring of grassland in Europe. Eurasian Dry Grassland Group, Scientific Report, p.27. Online: [DOI: 10.21570/EDGG.PG.40.27-31](https://doi.org/10.21570/EDGG.PG.40.27-31).



# Sampling



REGION	ALP-1	ALP-2	ALP-3	ALP-4	ATL-1	ATL-2	ATL-3	BOR-1	BOR-2	CON-1	CON-2	CON-3	CON-4	MED-1	MED-2	MED-3	MED-4	PAN-1	PAN-2	STEBS-1	STEBS-2	Sum per country
AT		8	35	25						6	20	12	1									107
BE					56	1				4	15	2										78
BG		1	5	8						34	29	13	8							9	30	137
CY														25	19	11	5					60
CZ										2	42	29	1					7	20			101
DE		1	1	1	52	1				63	58	33	2									212
DK					21					39												60
EE								56	4													60
EL														30	25	28	6					89
ES		1	1	1	2	8	44							9	31	101	52					250
FI								54	6													60
FR	1	1	3	3	99	33	7			11	30	17	32	4	4	4	1					250
HR		4	14	1						67	10	1		9	12	2						120
HU																		53	16			69
IE					63	7																70
IT	8	5	7	11						42	15	7	5	31	44	34	7					216
LT								60	8	1												69
LU										2	54	4										60
LV								48	12													60
MT														27	3							30
NL					59	1																60
PL		2	6	1						128	30	1										168
PT					4	2	2							29	40	32	3					112
RO		3	16	6						30	29	9	1					35	4	48	26	207
SE		1	1					43	51	3	1											100
SI	7	12	25	4						9	33	11	2									103
SK	34	13	8	1														23	13			92
Sum per BioGeo	50	52	122	62	356	53	53	261	81	441	366	139	52	164	178	212	74	118	53	57	56	3000

EMBAL rollout 2022/2023

3.000 plots across EU-27

# Conclusion and perspectives

- ❖ LUCAS Copernicus component has high potential for remote sensing community to generate timely LC information

and more:

- ❖ Location and validation of Land Cover boundaries
- ❖ Training and validation database for automatic image recognition
- ❖ Extending a professional survey using citizen science approaches
- ❖ Contributing to global free&open in-situ databases for EO analysis
- ❖ Linking sample and areal based Land Cover extent estimates
- ❖ Trigger better integration of statistical and geospatial domains
- ❖ Computer vision could be used to re-engineered legacy data
- ❖ New survey data coming



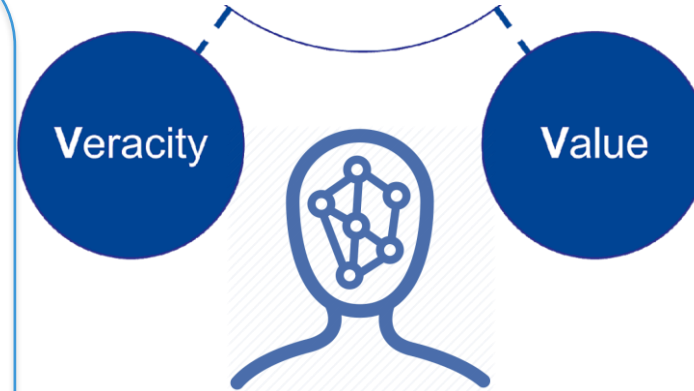
# Disruptive ways to bring Veracity and Value?

## Administration data

- Yearly parcel crop type information (LPIS-GSAA)
- Statistics (LUCAS)



Farmers' declarations



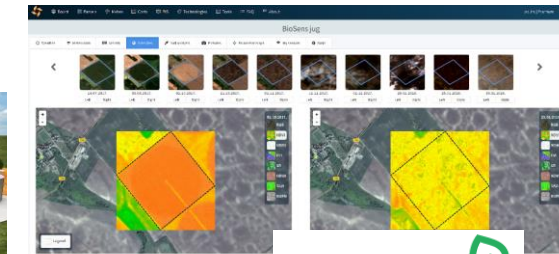
## Farmers data

Farm management tools



**LandSense**  
A Citizen Observatory and Innovation Marketplace  
for Land Use and Land Cover Monitoring

Digital Agriculture of Serbia  
AgroSense



OneSoil

Farm sensors and machinery



## Survey



## Crowdsourcing

Active

Geo-tagged street level imagery



Pl@ntNet

iNaturalist



Opportunistic



Mapillary

OpenStreetMap



Social  
Networks



European  
Commission

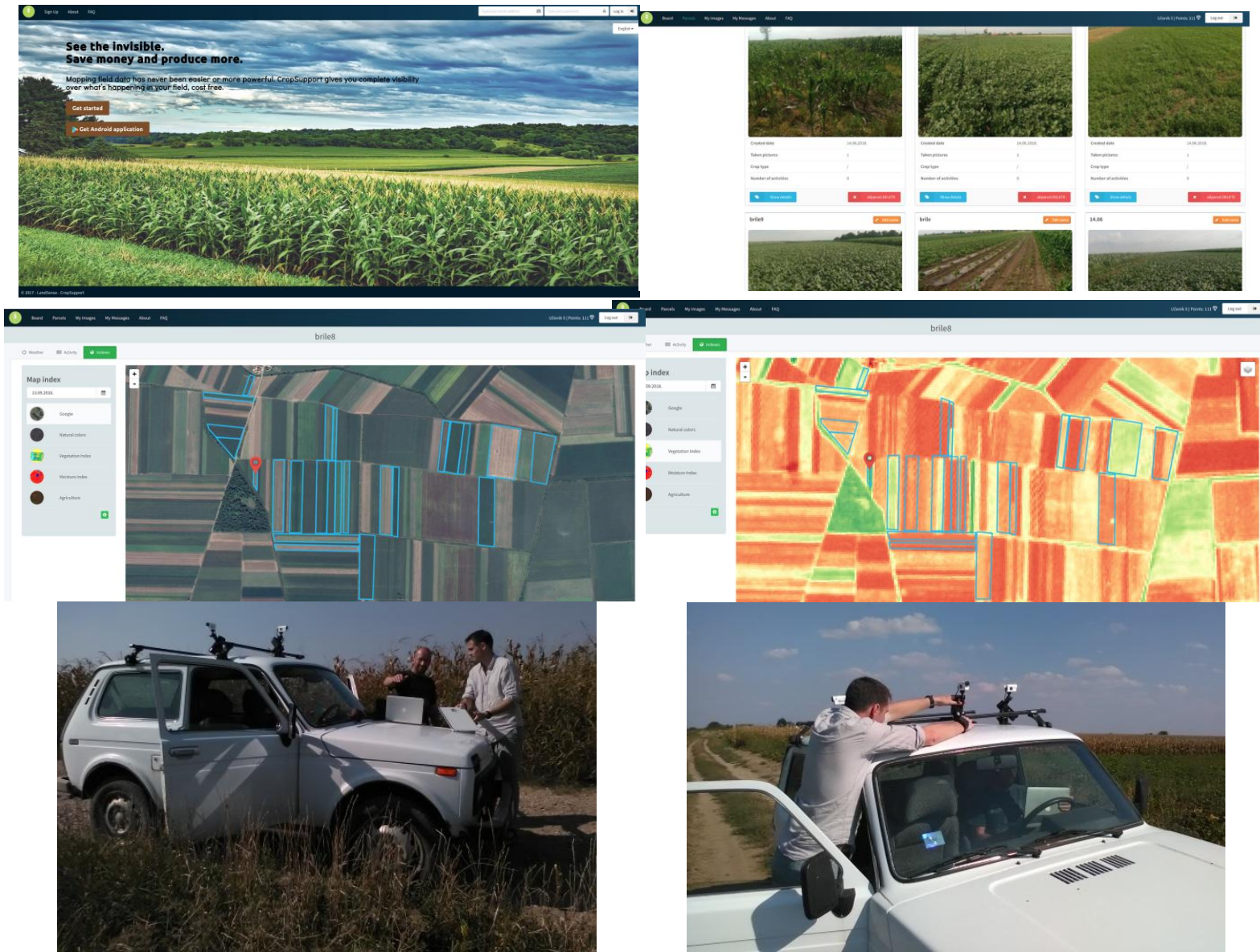
# Active crowd-sourcing

## In-situ data for agriculture





# Can we support Farm Sourcing app in Serbia?



- Crop Support App developed by Inonsens
- 242 parcels monitored by farmers
- Gamification with students
- Corn, Soybean, Wheat
- Very Small parcels (Sentinels)

## Points

Points from activities: 0  
Points from creating parcels: 76  
Points from invite registrations: 0  
Points from uploading images: 35

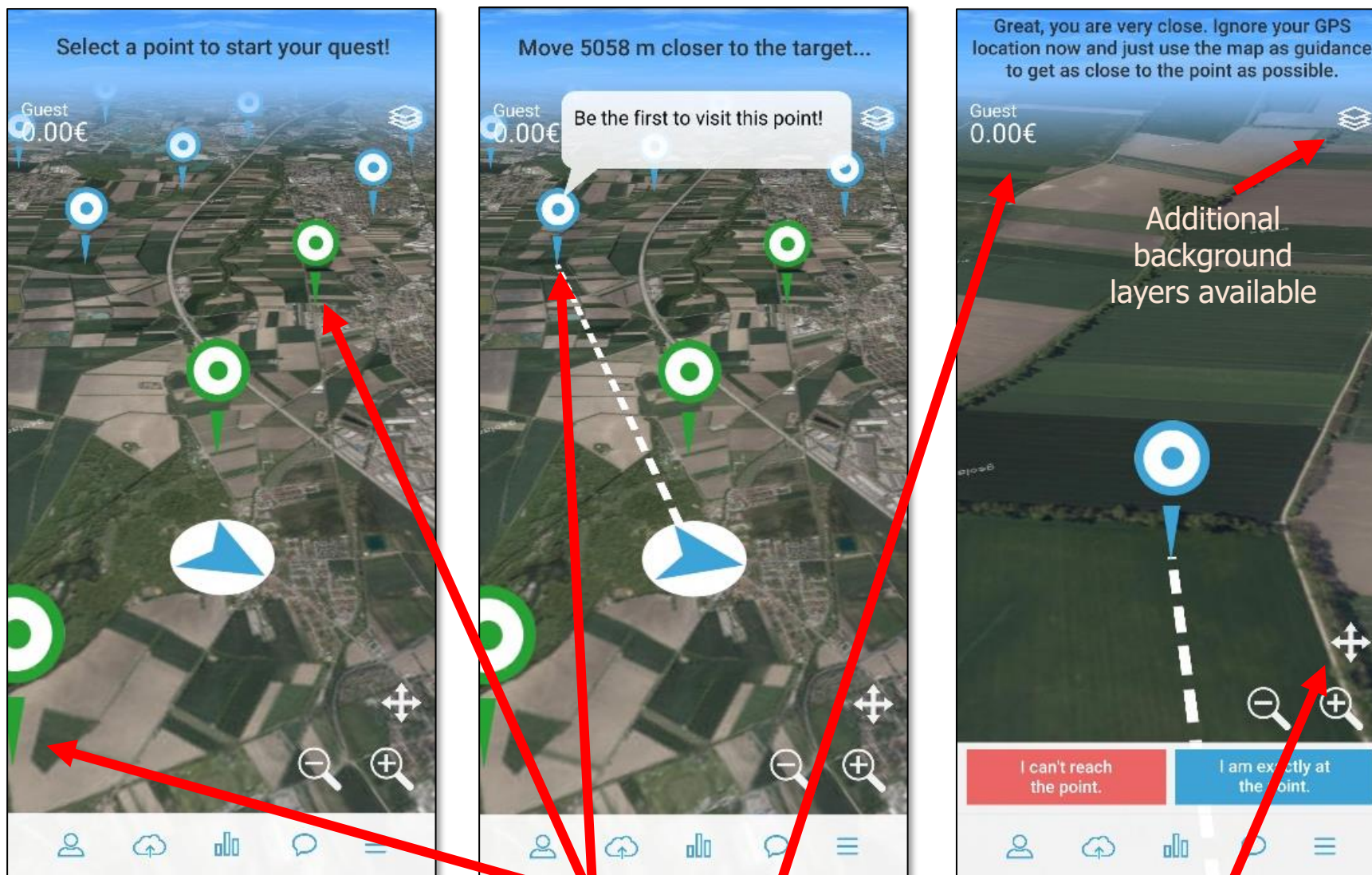
## Legend

Creating activity is: 0.5  
Invitation to register is: 1  
Creating parcel is: 2  
Taking photo of a parcel is: 1

Exit



# Can we collect Land Cover ground truth with crowdsourcing ? FotoQuest



Available locations to visit, money (€) earned, and 2D/3D map view



## Rewards:

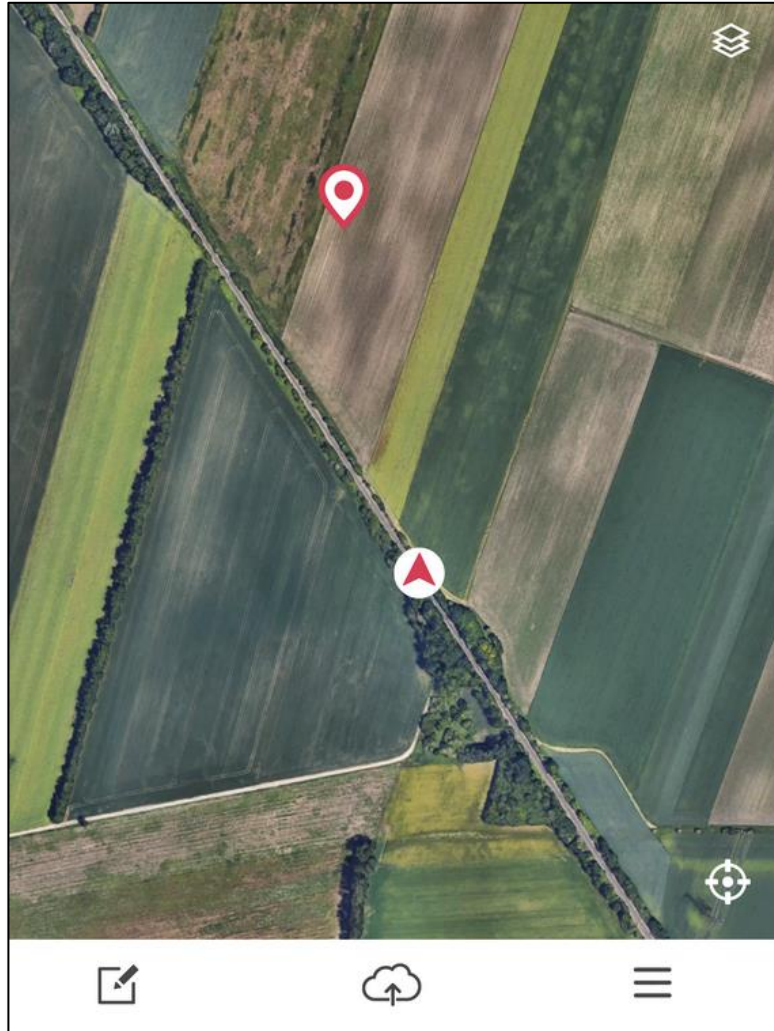
**Each location** visited awarded the participant between **1 to 3 Euros**, depending on the level of difficulty, e.g., €3 for visiting points on sites far from roads

**Weekly challenges** with 1 random point awarded **€30 Euros** to the first visitor





# e-shape CropObserve App



The [e-shape](#) **CropObserve** mobile application was developed to allow anyone to observe agricultural fields anywhere. The app is focused on collecting crop type, phenological stage, visible damage and management practices.



GET IT ON  
Google Play

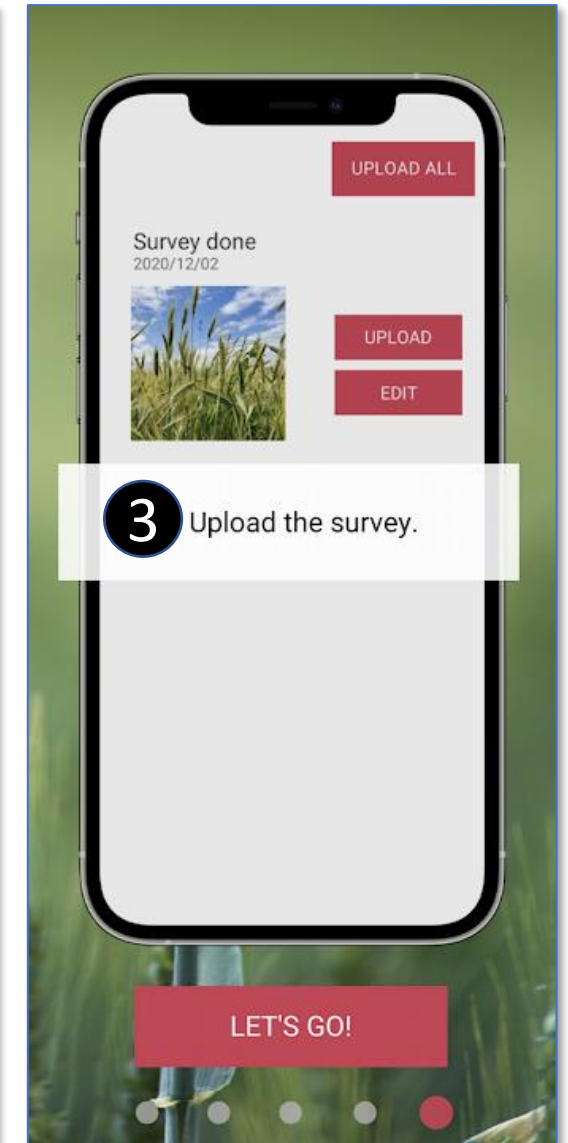
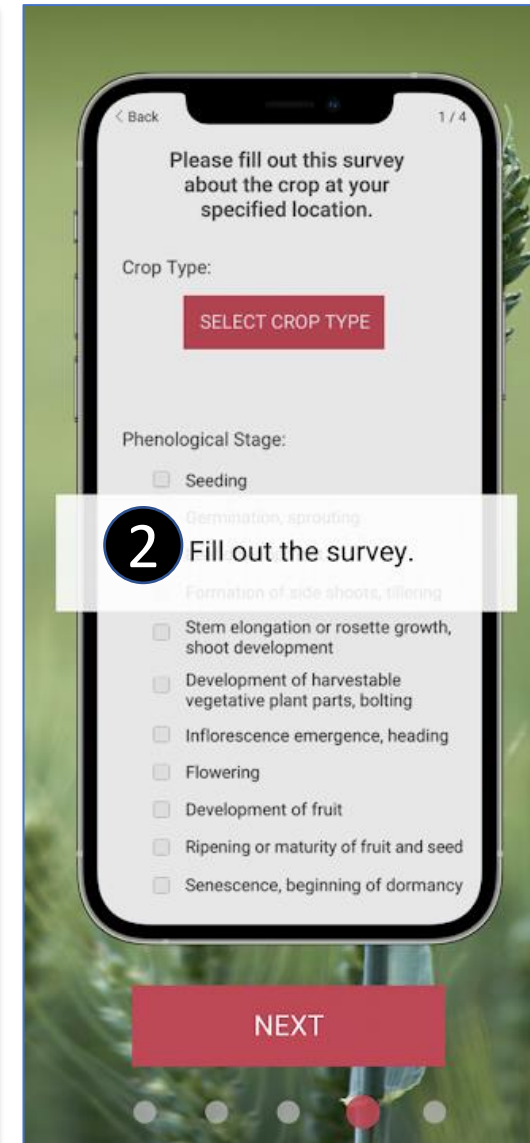
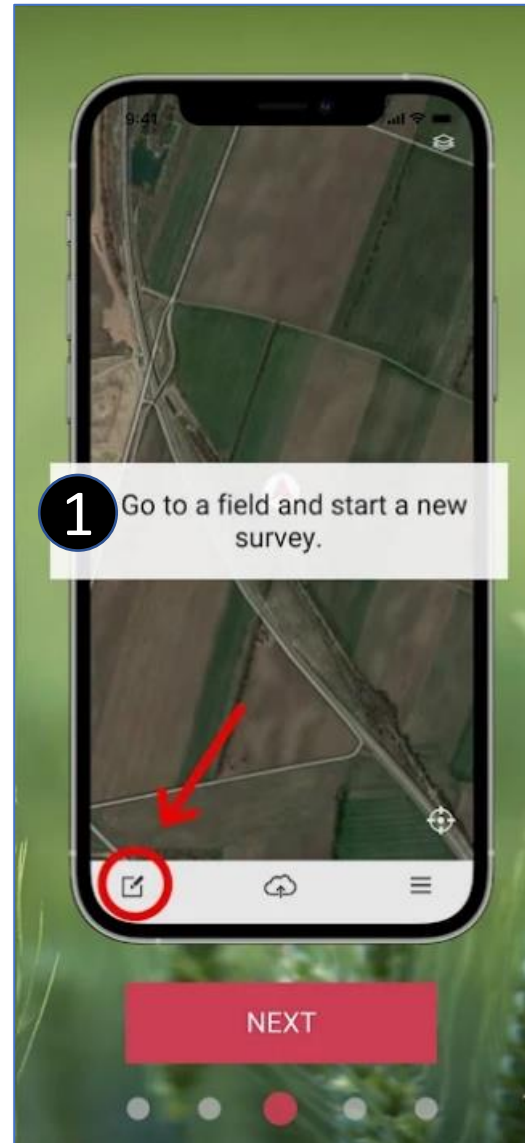
Download on the  
App Store



European  
Commission

# Collection of in-situ data

- Involve non-experts
  - Basic information:
  - Crop type
  - Phenological stage
  - Damage
  - Management activities
- All data is made open





# Pl@ntNet European Crops project



- Develop a branch in Pl@ntNet on European crops
- Using 100.000s LUCAS cover pictures and citizen science
- Further develop deep learning algorithms for European crops
- Deploy app and collect in-situ data on crops across Europe – in-season observations on crop location? Use by citizens, farmers, inspectors, ...?



B11 - Common wheat  
B12 - Durum wheat  
B13 - Barley  
B14 - Rye  
B16 - Maize  
B18 - Triticale  
B21 - Potatoes  
B22 - Sugar beet  
B31 - Sunflower

# Opportunistic crowd-sourcing

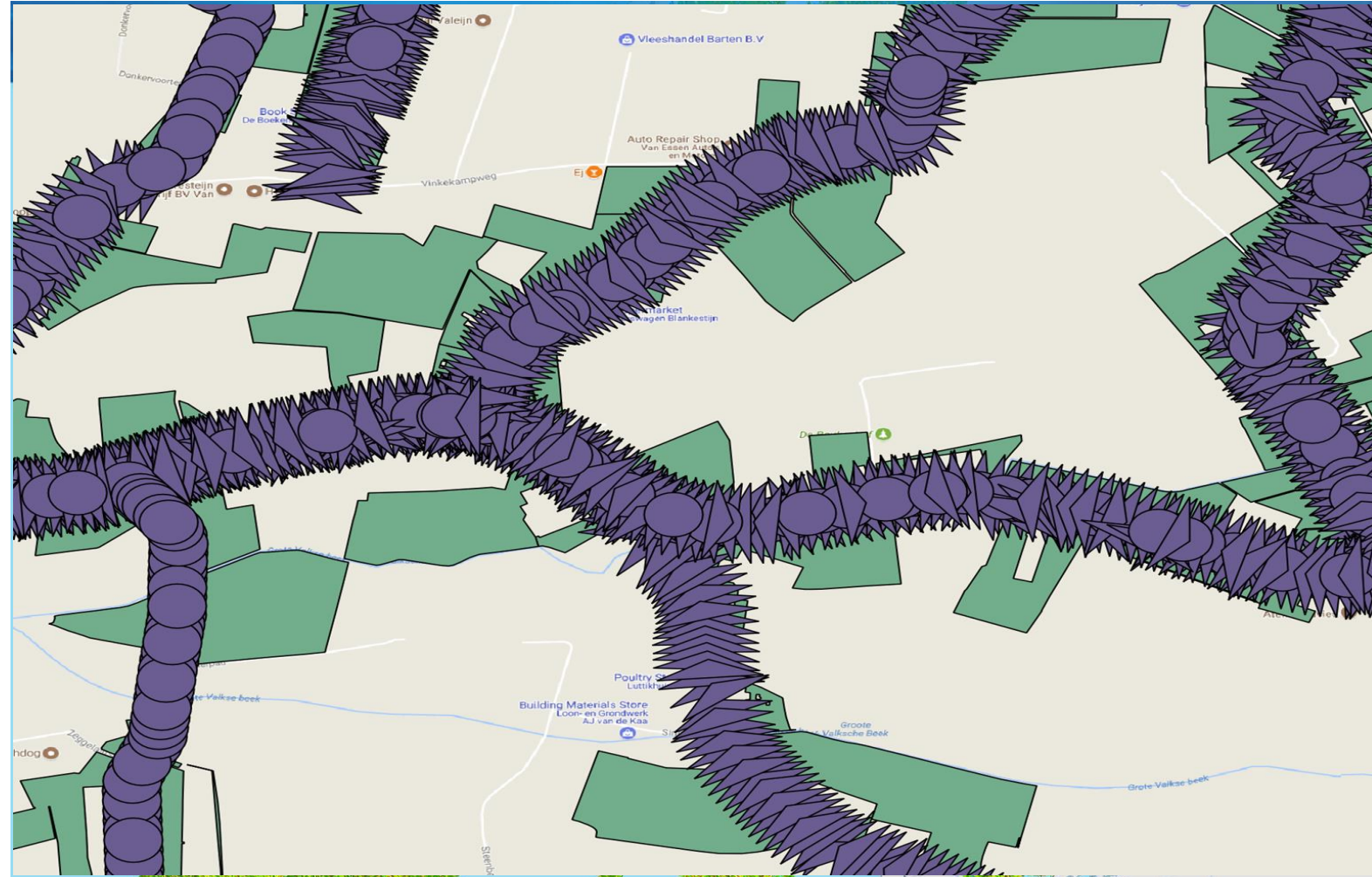
## In-situ data for agriculture





# 1. Introduction

1. Copernicus Sentinels and the need for groundtruth.
2. Crowdsourced, street-level imagery
3. Availability and usability of this data set
4. LUCAS 2018
5. Parcel-level crop identification.



# Research Questions

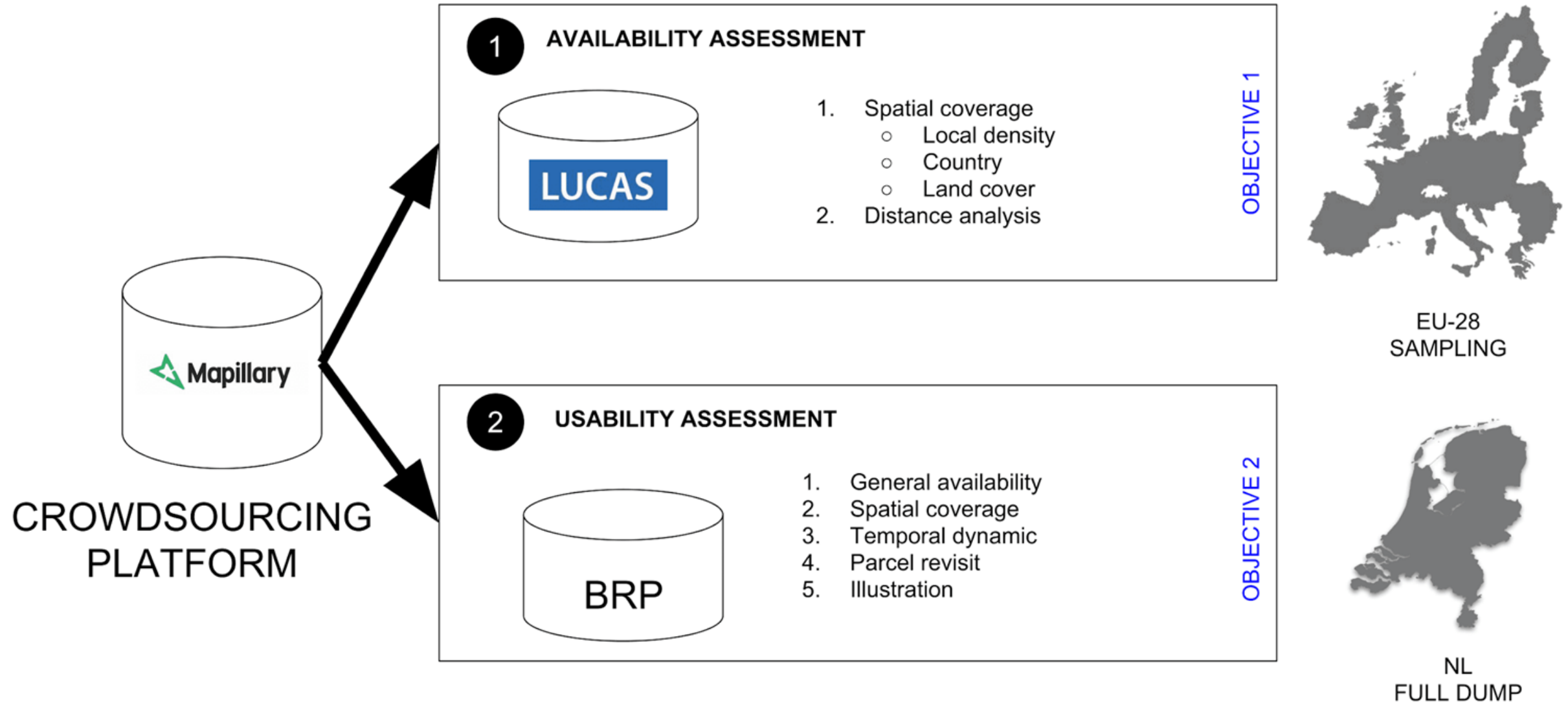
**Is crowd-sourced street-level imagery a potential suitable source of in-site ground-truth data in the field of agricultural Earth Observation?**

**What is the availability of these data sets in terms of their spatial and temporal coverage?**

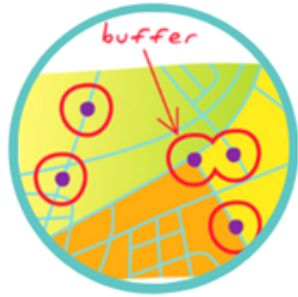
**What is the usability of these data sets in terms of their fitness for purpose for agricultural monitoring?**



# Methodology



# Availability Assessment Methodology - EU



## 1. Image harvest

For each LUCAS 2018 point, the location of the closest crowd-sourced image was collected with a **maximum distance of 2 km**.

$$\hat{\lambda}_i = \frac{\frac{n_{\text{sampling}}}{\alpha}}{\frac{n_{\text{crowdsourced}}}{\alpha}}$$

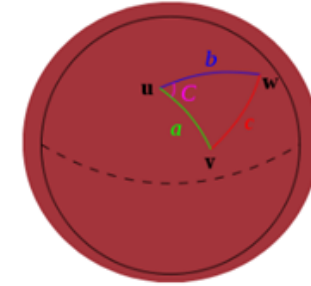
## 2. Local density

The local density, based on the **quadrant density method** is thus defined as the **ratio of the local density of sampling on the local density of crowd-sourced data** where  $\alpha$  corresponds to the area of the grid cell.

$$= \frac{n_{\text{crowdsourcing}}}{n_{\text{sampling}}}$$

## 3. Global density

The global density is the **ratio between number of points available in the crowd-sourcing platform and the point sampling**.



## 4. Distance

The shortest distance between the two points is calculated according to the **haversine method** assuming a spherical earth, and thus ignoring ellipsoidal effects.

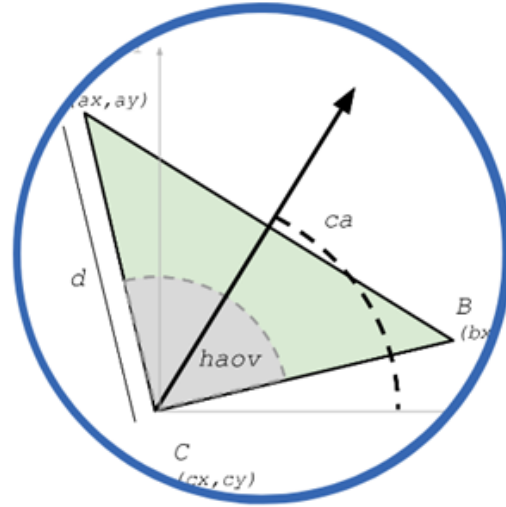


# Availability Assessment Methodology - NL



## 1. Subset Mapillary data set

For each Mapillary point we take only those that fall within a **50m buffer** of a BRP parcel and those that are within the **relevant time window of interest**.



## 2. Triangular or circular IFOV

Third, we generate Instantaneous Fields Of View (IFOV) for the subset of images - either triangular for normal cameras with a fixed horizontal field of view, or circular for panoramic fields of view



## 3. Intersection of polygon data sets

The IFOV polygon data set is intersected with the parcel boundaries while keeping the native geometry and attributes in order to know which parcels are visible on the image.



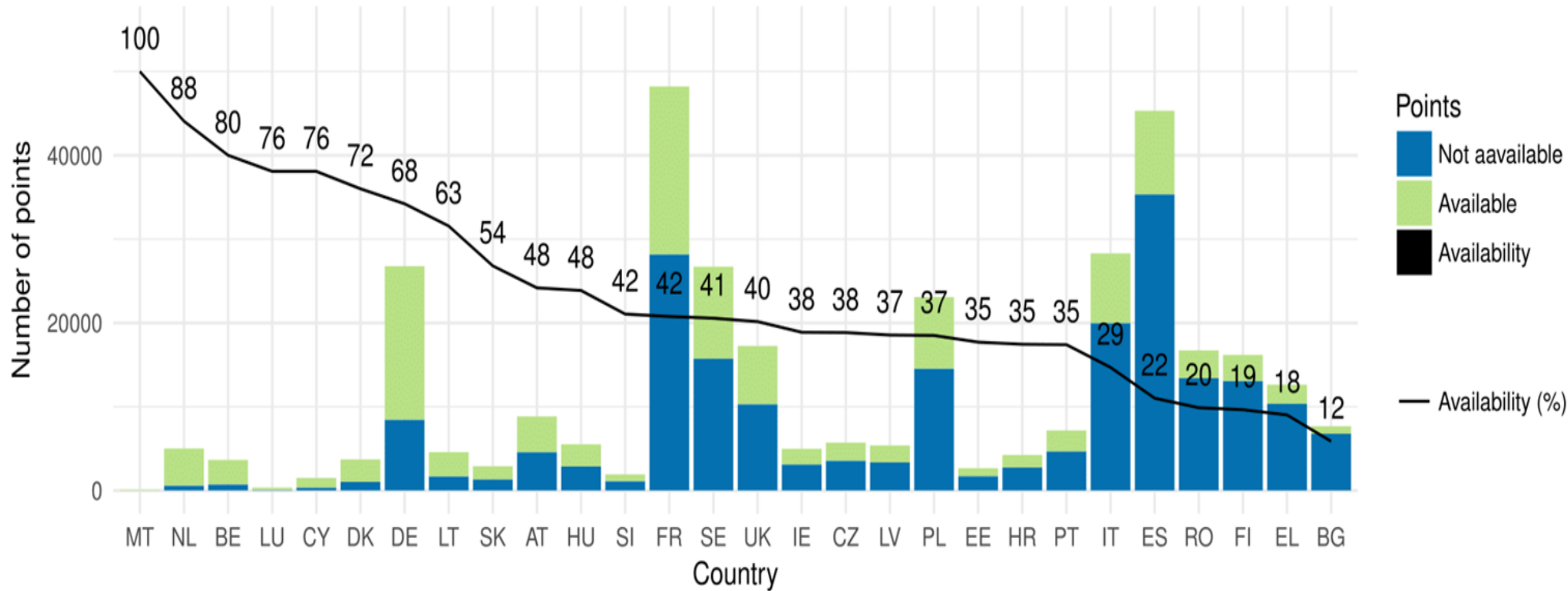
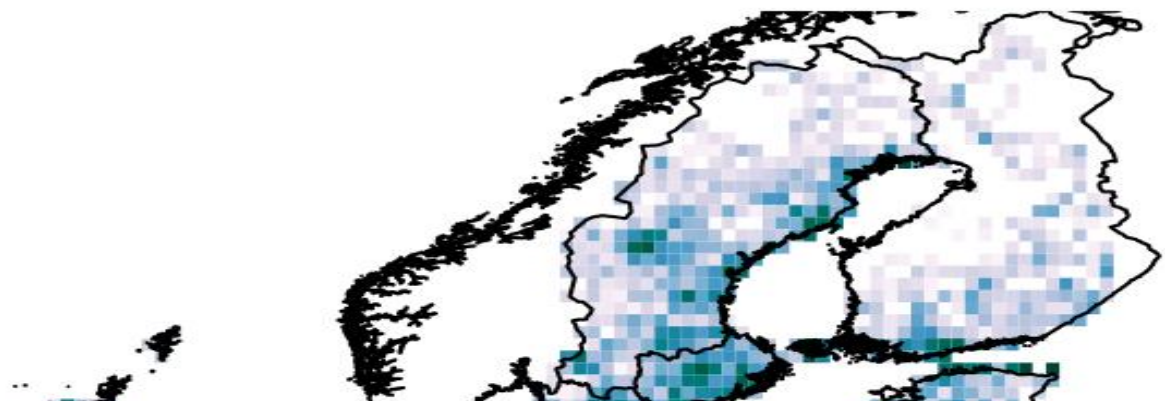
## 4. Final table

Metadata is stored in a table of polygon features which can easily be queried or exported to a desired format.

**4.3. USABILITY**

# Results

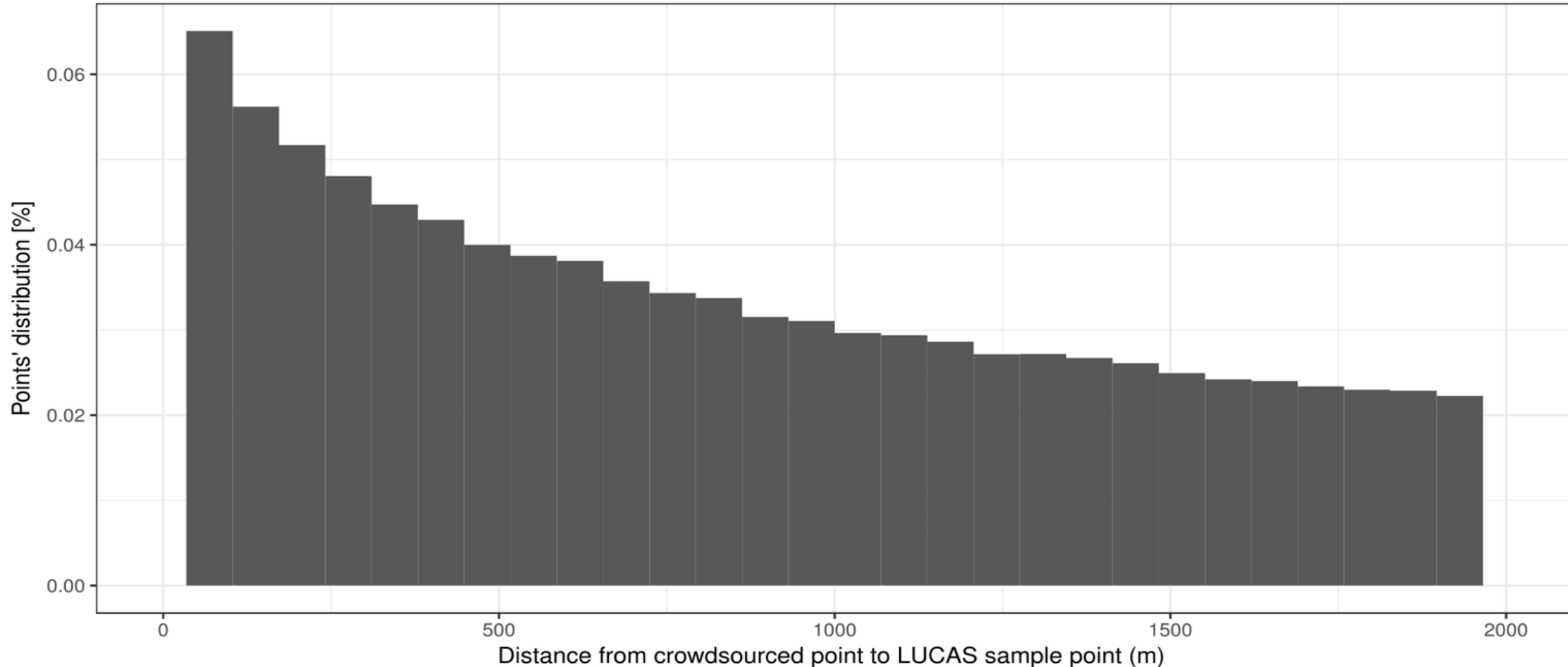
## 5.1. Availability Assessment - EU





# Results: Availability Assessment - EU

Min. 0.03; 1st Qu. 297.11; Median 726.43; Mean 816.11; 3rd Qu. 1289.86; Max. 2007.00; NA's 209685



# Results: Availability Assessment - NL

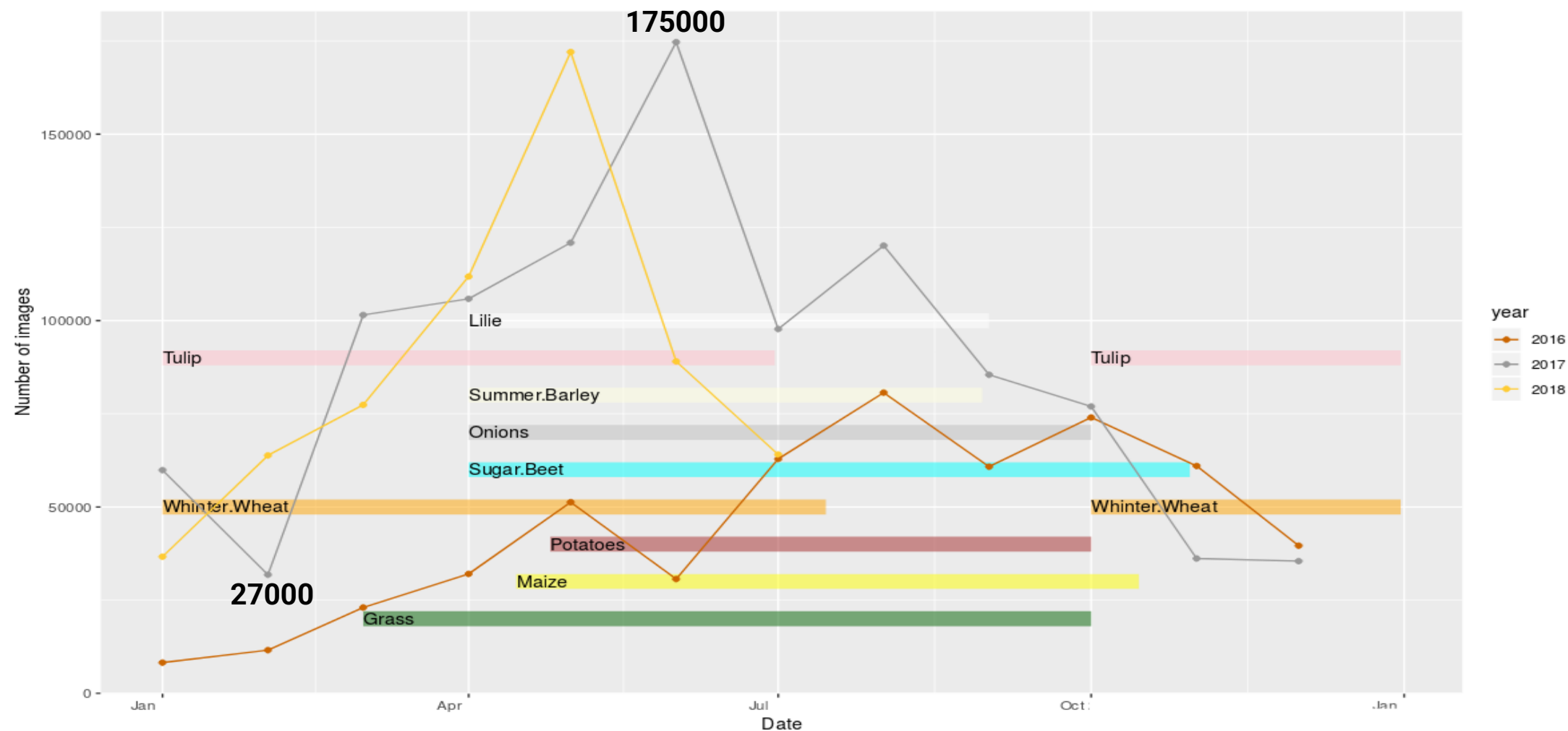
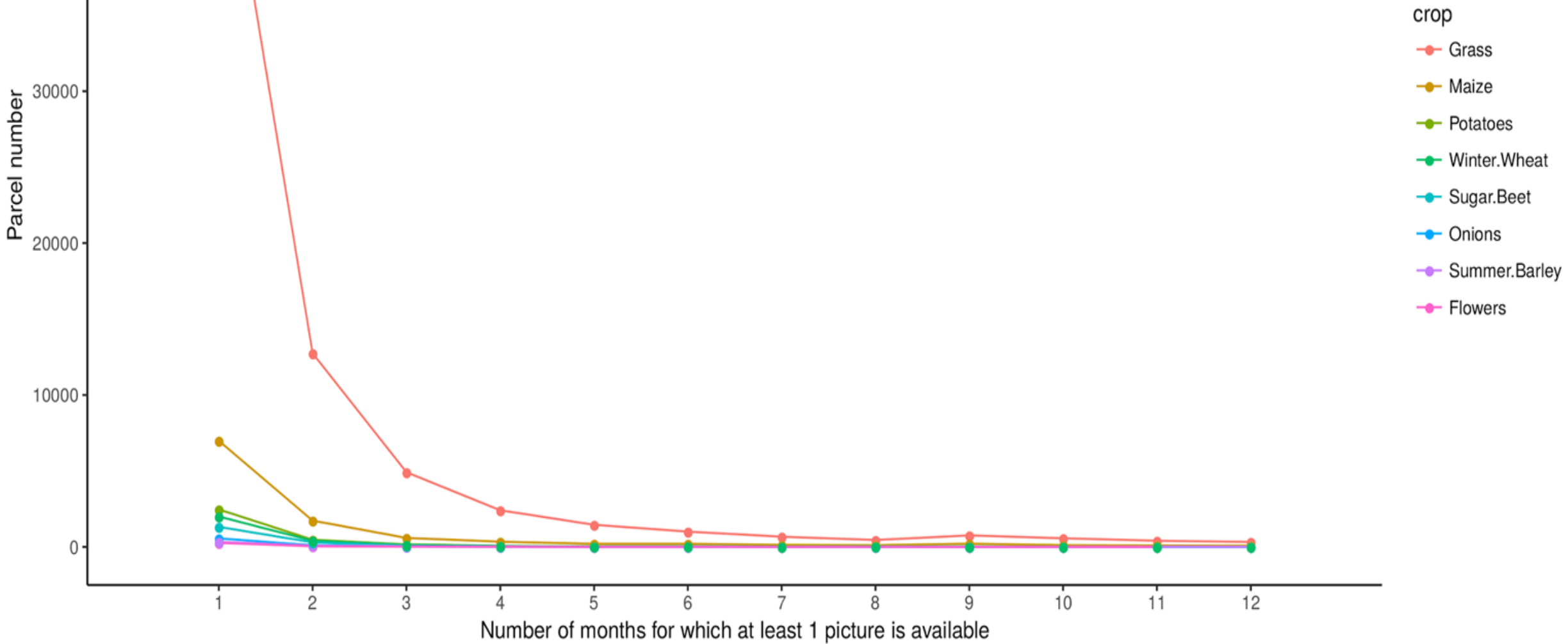


Figure 9. Monthly availability of images throughout the year for 2016-2018



Month	1	2	3	4	5	6	7	8	9	10	11	12
Grass	50249 (9.93%)	12747 (2.52%)	4917 (0.97%)	2411 (0.48%)	1456 (0.29%)	1006 (0.2%)	673 (0.13%)	455 (0.09%)	758 (0.15%)	571 (0.11%)	402 (0.08%)	328 (0.06%)
Maize	6989 (8.21%)	1731 (2.03%)	588 (0.69%)	340 (0.4%)	206 (0.24%)	204 (0.24%)	128 (0.15%)	109 (0.13%)	211 (0.25%)	112 (0.13%)	96 (0.11%)	78 (0.09%)
Potatoes	2467 (7.23%)	471 (1.38%)	159 (0.47%)	72 (0.21%)	40 (0.12%)	39 (0.11%)	4 (0.01%)	16 (0.05%)	15 (0.04%)	14 (0.04%)	9 (0.03%)	7 (0.02%)
Winter Wheat	2001 (9.76%)	406 (1.98%)	124 (0.6%)	59 (0.29%)	24 (0.12%)	43 (0.21%)	15 (0.07%)	33 (0.16%)	20 (0.1%)	12 (0.06%)	5 (0.02%)	8 (0.04%)
Sugar Beet	1318 (7.66%)	300 (1.74%)	76 (0.44%)	34 (0.2%)	28 (0.16%)	28 (0.16%)	13 (0.08%)	22 (0.13%)	15 (0.09%)	8 (0.05%)	8 (0.05%)	4 (0.02%)
Onions	558 (8.28%)	105 (1.56%)	29 (0.43%)	17 (0.25%)	4 (0.06%)	8 (0.12%)	1 (0.01%)	1 (0.01%)	1 (0.01%)	2 (0.03%)	1 (0.01%)	2 (0.03%)
Flowers	327 (5.16%)	75 (1.18%)	32 (0.5%)	9 (0.14%)	8 (0.13%)	6 (0.09%)	4 (0.06%)	11 (0.17%)	3 (0.05%)	1 (0.02%)	-	-



## 5.3. Results: Use cases - NL



(a) 2017.05.06



(b) 2017.07.23



(c) 2017.10.10



(d) 2017.11.05



## 6. Conclusion

- 1) What is the spatial availability of Mapillary street-level imagery across the European Union based on the stratified and systematic LUCAS 2018 sample?
- 2) What is the detailed spatio-temporal availability of these images in relation to crops, crop phenology, and agricultural parcels in the Netherlands?
- 3) Which are the parcels which most likely to be observed by the Mapillary crowd-sourced images using metadata and geo-spatial analysis?
- 4) And finally, what is the potential usefulness of crowd-sourced imagery for different agricultural monitoring use cases?

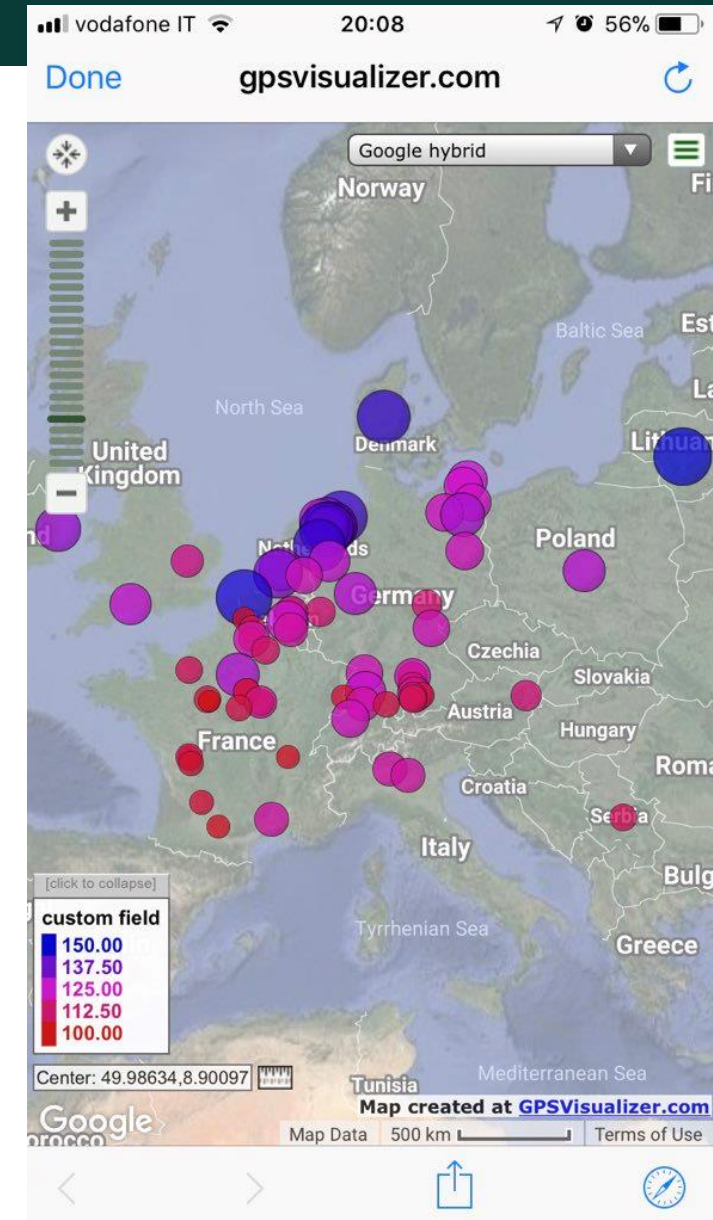
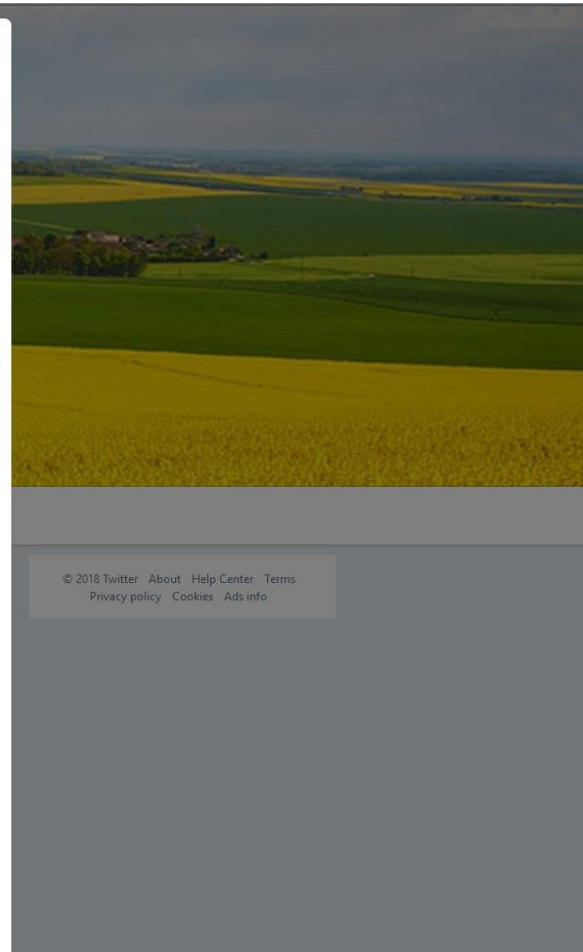
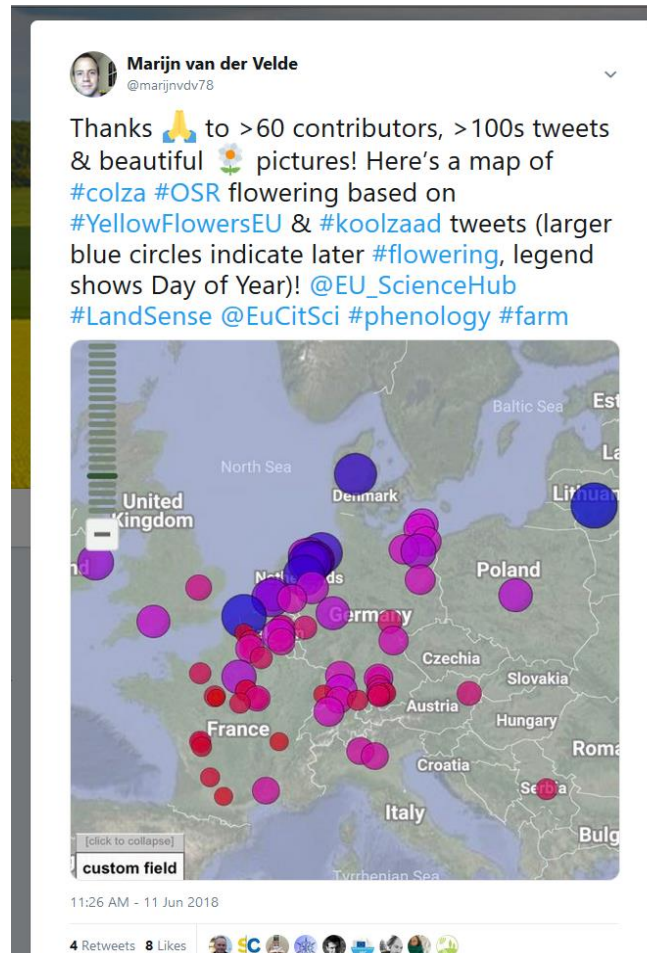
# Social media campaign *#YellowFlowersEU*





# Social media campaign *#YellowFlowersEU*

## Active and opportunistic sourcing...



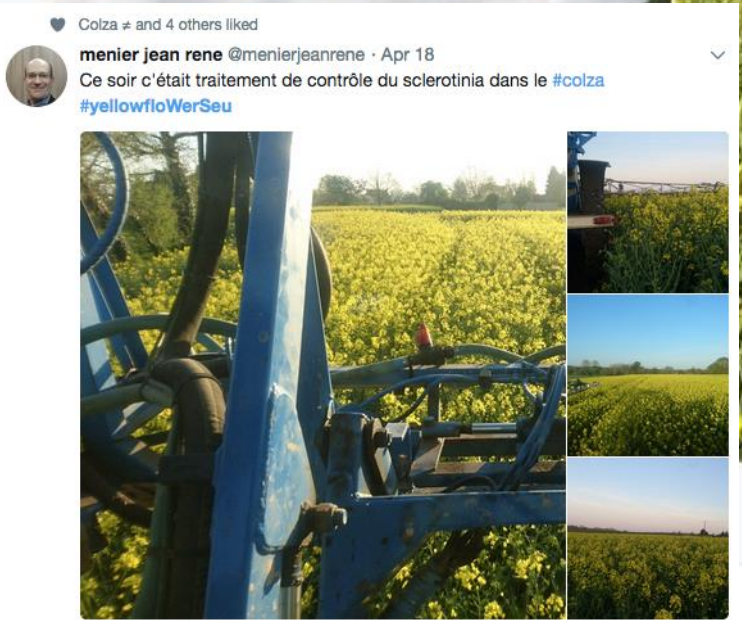


# Social media campaign #YellowFlowersEU

## Citizens



## Farmers



## Photographers





# Social media campaign #YellowFlowersEU

## Rapeseed is controversial: agriculture, pesticides, biofuels, biodiversity & bees



***Inherent bias in some CS activities??  
Connect all stakeholders for policy making...***

# Conclusions & Take home messages

- New paradigm for remote sensing acquisition : free, global, high spatial and temporal resolution
- New (almost free) computing capacities
- Data deluge integration needed
- Money or capacity is not the limit anymore, sky is the limit !

LEARN, LEARN, LEARN....

→ New opportunities for young graduates!!!





# Thanks

## Questions?

You can find me at **[raphael.dandrimont@ec.europa.eu](mailto:raphael.dandrimont@ec.europa.eu)**