NIVA – NEW IACS VISION IN ACTION

WP3 Harmonisation and Interoperability

D3.5 Recommendations for standardised connections between IACS and other applications

Deliverable Lead: OPEKEPE

Deliverable due date: 30/05/2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 842009.

Document Control Page

Document Control Page	
Title	D3.5 Recommendations for standardised connections between IACS and other applications
Creator	Michael Manolalakis (OPEKEPE)
Description	D3.5 deals with 2 key topics for Paying Agencies in the context of new CAP: access and preprocess of EO data and data exchanges between IACS and FMIS.
Publisher	"NIVA - New IACS Vision in Action" Consortium
Contributors	Nikos Kalatzis (Neuropublic) Yorgos Efsthathiou (Neuropublic) Dominique Laurent (IGN) Benoit More (IGN)
Date of delivery	27/07/2021
Type	Text
Language	EN-GB
Rights	Copyright "NIVA - New IACS Vision in Action"
Audience	☒ Public ☐ Confidential ☐ Classified
Status	☒ In Progress ☐ For Review ☐ For Approval ☐ Approved

Revision History			
Version	**Date**	**Modified by**	**Comments**
1.0.0	27/07/2021	Dominique Laurent	Version for internal review by NIVA partners
1.0	15/09/2021	Dominique Laurent	Comments from NIVA reviewers taken into account

Disclaimer

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1. General context

As part of its on-going move to simplify and modernize the EU's Common Agricultural Policy (CAP), the European Commission has adopted new rules that will for the first time expressly allow a range of modern technologies to be used when carrying out checks for area-based CAP payments. This includes the possibility to completely replace physical checks on farms with a system of automated checks based on analysis of Earth Observation (EO) data and completed by self-certification processes.

The CAP monitoring through EO data has been made possible by the Sentinel program that delivers as open data high resolution images with frequent visits: the time series of Sentinel or other satellite images are a powerful mean to check farmer declarations. However, accessing and pre-processing these big volumes of images in order to get ARD (Analysis Ready Data) is a complex task, raising many issues.

In the new monitoring system, the analysis of EO data provides should provide traffic lights (green, yellow or red). In case of yellow lights, i.e. in case of doubts, alternative evidences have to be provided by farmers. A possible source of these alternative evidences is the FMIS (Farm Management Information System).

In addition to support the new CAP monitoring system, the NIVA project also aims to reduce the administrative burden on farmers. The farmer declaration might be simplified by data exchange between the information system of Paying Agencies (IACS) and the one of farmers (FMIS). Use of IACS data would facilitate farmer declaration regarding parcel geometry (e.g. ensure the agricultural parcel is within the reference parcel) and the direct import of FMIS data within GSAA would avoid fastidious and error-prone manual work from farmer.

1.2. Deliverable scope

This deliverable is part of the work related to technical interoperability. It was initially planned to be result of Task 3.5 whose main purposes were

- to investigate the requirements and existing standards to connect to other systems (Farm Management Information System (FMIS), farming machinery, advisory services, external registries) and to other applications (e.g. food processing/business, environment, climate)
- to propose relevant standards or recommendations for connections with other applications.

During the progress of NIVA project, it has appeared that two external systems were of key interest but raising lots of difficulties for Paying Agencies (Earth Observation data sources and Farm Management Information System) whereas most other external applications are just requiring delivery of IACS data according to more generic rules about data formats or exchange protocols.

Therefore, this deliverable “Recommendations for standardised connections between IACS and other applications” is focusing on the access to Earth Observation data and on exchanges between IACS and FMIS. The topic of connexions with other external systems has been considered under deliverable D3.4 “Recommendations for IACS data flows”.

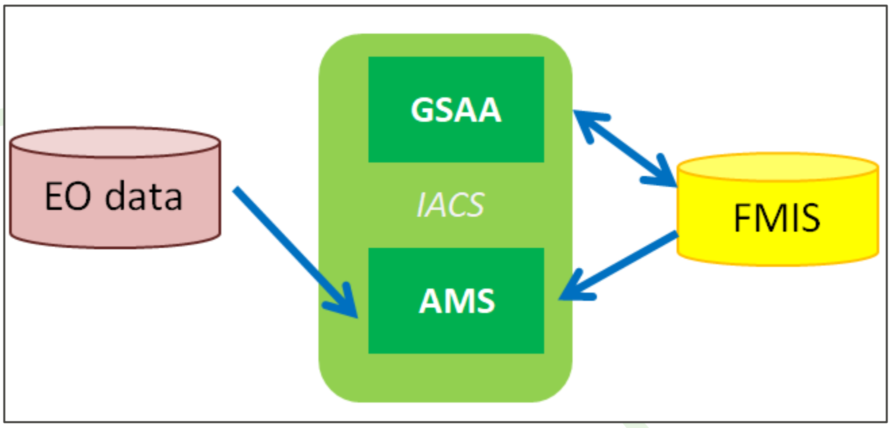


Figure 1 the data flows considered in D3.5

NOTE: Earth Observation data is already used in the current control system of CAP (CwRM: Control with Remote sensing); CwRS may be mentioned in some chapters of this deliverable but it is not the purpose of this deliverable. This deliverable is focusing on the EO data used for the new monitoring system.

1.3. Objectives and content

At the beginning of the NIVA project, access to big volumes of EO data and derived products was considered as a very complex topic that nobody can perfectly understand and the situation regarding potential exchanges between IACS and FMIS was even worse, with very limited knowledge about FMIS available in the project. This deliverable does not pretend to provide perfect understanding about these two new, valuable but still a bit mysterious data sources for the new CAP; its aim is more simply to provide at least some common basic knowledge, enabling readers to be aware of the available data and of the potential exchange solution options with their advantages and drawbacks.

The two main chapters (EO data and FMIs) are organised in a similar way:

- introductory paragraphs that provide basic information such as the available data description

- central paragraphs that capitalise the NIVA experiences
- final paragraphs that summarise the main learning's and that propose a few recommendations.

This document targets mainly Paying Agencies; it aims to support them in the preparation of the new Area monitoring System (and more generally of the new CAP), whether they set up it by themselves or they prefer to sub-contract it.

However, this document may also be of interest for any other stakeholder involved in EO monitoring or in FMIS data exchange, such as PA technical partners, researchers, FMIS editors or farmer organisations.

In practice, the introduction and conclusion chapters are presenting the information in a simple way and are expected to be of easy reading by anyone having interest in D3.5 topics whereas the central paragraphs capitalising the NIVA experiences include more complex content and so may be more adapted to more technical or motivated readers.

1.4. Methodology

This deliverable is based on various sources of information and knowledge:

- personal expertise of the document writers
- bibliographic research, attendance to conferences and webinars
- Direct experience of NIVA partners mainly through the feed-back provided by the developing teams of the project.
- other sources of experience
    - experience from previous projects, such as Sen4CAP for EO data or IoF2020 (Internet of Fruits and Farms), ATLAS, DEMETER, Mef4CAP for data exchanges between IACS and FMIS)
    - national experiences of NIVA partners (even if not conducted for the project)

To complement these various sources, a questionnaire was sent, to Paying Agencies and to FMIS editors regarding respectively access to EO data and data exchange between FMIS and IACS.

More details about these questionnaires may be found in related chapters.

1.5. Glossary

AOI: Area of Interest

AMS: Area Monitoring System

API: Application Programming Interface

ARD: Analysis Ready Data

BOA: Bottom of Atmosphere

CAP: Common Agricultural Policy

CbM: Check by Monitoring

CNES: Centre National d'Etudes Spatiales (National Centre of Spatial Studies)

DIAS: Data and Information Access Services

DTM: Digital terrain model

EO: Earth Observation

EU: European Union

GPS: Global Positioning System

GRD: Ground Range Detected

HHR: High High Resolution

HR: High Resolution

IaaS: Infrastructure as a Service

IACS: Integrated Administration and Control System

FCOVER: Fractional of Vegetation Cover

FMIS: Farm Management Information System

LAI: Leaf Area Index

LTA: Long Term Archive

LUT: Lookup Table

MAJA: Maccs-Atcor Joint Algorithm

MR: Medium Resolution

MS: Member State

NDVI: Normalized Difference Vegetation Index

NIR: Near-infrared

PA: Paying Agency

PPI: Plant Phenology Index

REST: Representational State Transfer

S-1: Sentinel-1

S-2: Sentinel-2

SAFE: Standard Archive Format for Europe

SAR: Synthetic Aperture Radar

SI: Spectral Indices

SLC: Single Look Complex

SNAP: Sentinel Application Platform

SRTM: Shuttle Radar Topography Mission

SWIR: Short-wave infrared

TOA: Top of Atmosphere

UC: Use Case

VHR: Very High Resolution

VM: Virtual machine

VPP: Vegetation Phenological and Productivity

WMS: Web Map Service

WCS: Web Coverage Service

WPS: Web Processing Service

CSW: Catalogue Service for the Web

2. Earth Observation data and services

2.1. Satellite images

In what follows, we briefly describe the various types of satellite data that the Paying Agencies use or may use for EO monitoring under the EU's Common Agricultural Policy (CAP).

There are many different types of satellite images. They vary in sensor type (optical, radar), spatial, temporal & spectral resolution, spectral bands, tile size, orbit type, geographical coverage and licence type (open or commercial).

- The geographical coverage is the part of Earth covered by the set of satellite images. For EO monitoring, images should cover EU territory.
- The spatial resolution is documented by the ground pixel size (e.g. 10 m). To be monitorable by EO processing, a given agricultural parcel should have a minimum number of pixels, i.e. an agricultural territory with small parcels will require finer resolution satellite images. In addition, the spatial resolution also impacts the type of analysis that can be done regarding the spatial texture of the images (e.g. detection of rows of trees or other crop). However, finer resolution implies bigger data volume to be processed
- The sensor type and the spectral resolution (number of relevant bands for optical imagery) are the image semantic; they determine the type of analysis that may be conducted. For instance, the near-infrared band of optical images is very useful for vegetation analysis.
- The temporal resolution is similar to the visit frequency of satellite image capture (time elapsed between observations of the same point on Earth). EO monitoring is generally not conducted on a single image but on temporal series. Having reliable and dense enough temporal series may be quite necessary for event detection (mowing, tillage, harvest, catch-crop growth)
- The licensing conditions (open or commercial) have obvious economic impact. When possible, open data is best solution.

In practice, only MR-HR images (resolution equal or coarser than 10 m) are available as open data whereas finer resolution images are delivered as commercial products.

In this document, we present first the open data HR images (Sentinel 2, Sentinel 1 and Landsat 8) and then the HHR or VHR commercial products, with main focus being on Sentinel satellite data, as they are the most widely used in the checks by monitoring approach.

All these satellite image sensors are covering the whole (political) Europe.

2.1.1. Sentinel – 2 (optical)

Sentinel-2 is a constellation with two twin satellites, Sentinel-2A and Sentinel-2B. They acquire optical imagery at a spatial resolution of 10, 20 and 60 m. The 13 different spectral bands of the Sentinel-2 sensors are shown in table 1. The revisit frequency of the combined constellation is 5 days (2-3 days at mid-latitudes).

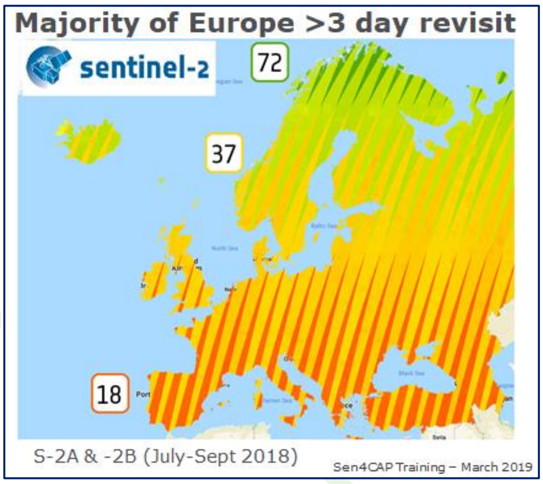


Figure 2 Sentinel-2 revisit frequency

Sentinel-2 satellites give new perspectives of land and vegetation. Their multispectral imager (13 spectral bands) provides high-quality image layers that can be used to derive various spectral indices and ratios, useful for instance for crop monitoring. More specifically, the spectral bands in red-edge and infrared regions of light spectra are used to calculate various spectral indices (SI) identifying the signal of certain vegetation, crop and soil characteristics.

Sentinel-2 spectral bands	Spatial resolution (m)
Band 1 – Coastal aerosol	60
Band 2 – Blue	10
Band 3 – Green	10
Band 4 – Red	10
Band 5 – Vegetation red edge	20
Band 6 – Vegetation red edge	20
Band 7 – Vegetation red edge	20
Band 8 – NIR	10
Band 8A – Narrow NIR	20
Band 9 – Water vapour	60
Band 10 – SWIR – Cirrus	60
Band 11 – SWIR	20
Band 12 – SWIR	20

Table 1 Spectral bands for the Sentinel-2

The spatial resolution of SENTINEL-2 is dependent on the particular spectral band: (https://sentinels.copernicus.eu/).

The bands B1, B9 and B10 do not provide information about the soil but about the atmosphere. They are used for atmospheric correction (see chapter 2.2.1.1) but they are not of direct interest for CAP monitoring.

There is some overlap between adjacent orbits entailing most frequent visits on the overlapping areas. S-2 products are a compilation of elementary granules of fixed size, along with a single orbit. A granule is the minimum indivisible partition of a product (containing all possible spectral bands). For Level-1C and Level-2A (see chapter 2.2.1.2), the granules, also called tiles, are 100x100 km^2 ortho-images in UTM/WGS84 projection. There is also some overlap between these tiles.

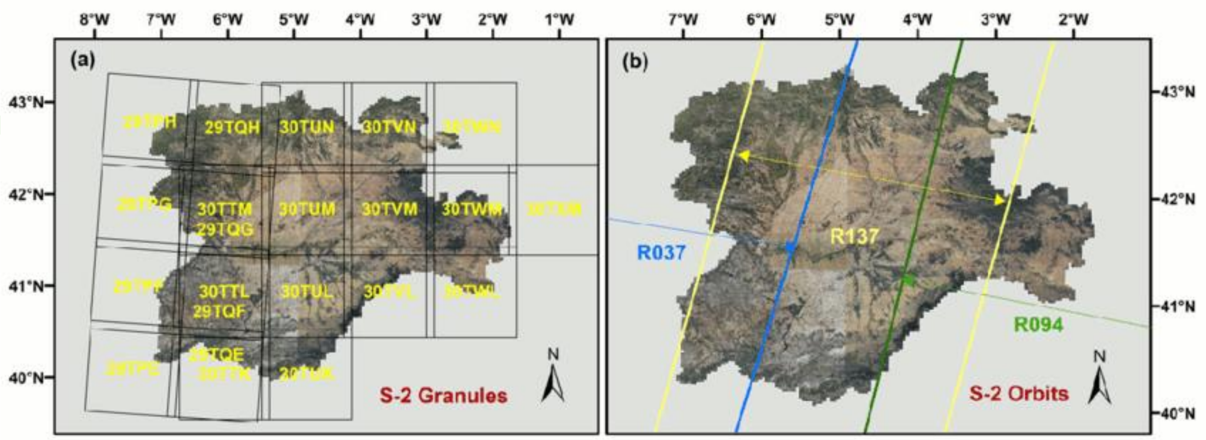


Figure 3 All Sentinel-2 granules and orbits covering Castile and León

NOTE (source of figure 3): Paredes-Gómez V, Gutiérrez A, Del Blanco V, Nafría DA. A Methodological Approach for Irrigation Detection in the Frame of Common Agricultural Policy Checks by Monitoring. Agronomy. 2020; 10(6):867. https://doi.org/10.3390/agronomy10060867)

Sentinel 2 images are the best candidates for EO monitoring because they are optical images (easy for interpretation) with rich semantics (13 bands), relatively good spatial resolution (up to 10 m) and because they are open data (freely available).

In theory, they have a good revisit frequency (around 5 days in Europe) however in practice there may be clouds in the sky making the images non exploitable. The cloud issue depends on geographic location; in Europe, clouds are generally more frequent in northern and central EU countries or in mountainous areas.

More detailed information about S2 may be found on:

https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook

2.1.2. Landsat 8 (optical)

NASAs Landsat-8 satellite sensor is an American Earth observation satellite developed by NASA and the U.S. Geological Survey (USGS). It has two main sensors: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) acquiring images in various wavelengths consisting of nine spectral bands with a spatial resolution of 30m, one panchromatic band at 15m and two thermal bands collected at 100m. Its revisit cycle (temporal resolution) is 16 days and the data is freely available.

Landsat data is considered of high quality and very stable products as the Landsat satellite program is the longest continuous Earth imaging program in history. Since 1972, Landsat satellites have collected huge amounts of consistent spectral imagery.

Landsat-8 spectral bands	Spatial resolution (m)
Band 1 – Coastal aerosol	30
Band 2 – Blue	30
Band 3 – Green	30
Band 4 – Red	30
Band 5 – Near Infrared (NIR)	30
Band 6 – SWIR1	30
Band 7 – SWIR2	30
Band 8 – Panchromatic	15
Band 9 – Cirrus	30
Band 10 – Thermal Infrared 1	100
Band 11 – Thermal Infrared 2	100

Table 2 Spectral bands for the Landsat 8 images

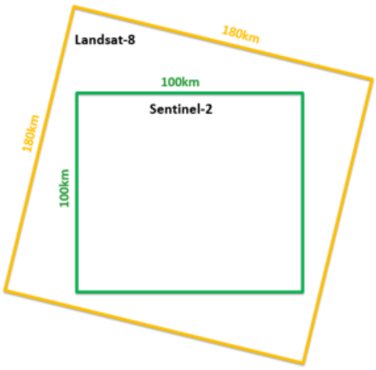


Figure 4 Comparison between Sentinel-2 and Landsat 8 images

Landsat 8 has a coarser spatial resolution (and so, largest tiles as shown in figure 3) than Sentinel-2 and that makes the images less suitable for EO monitoring. However, Landsat data is often used as supplementary data to Sentinel-2 images for constructing denser satellite time-series (combining Landsat and Sentinel the revisit time can drop below 3 days) or to mitigate the cloud issue.

This approach is used by several Paying Agencies, specifically in the northern European countries where the cloud cover is generally higher. For example the Danish Paying Agency uses an approach that takes full advantage of all available optical imagery over the agricultural season to create consistent optical time series. Further, Landsat-8 images are widely used in the crop classification tasks.

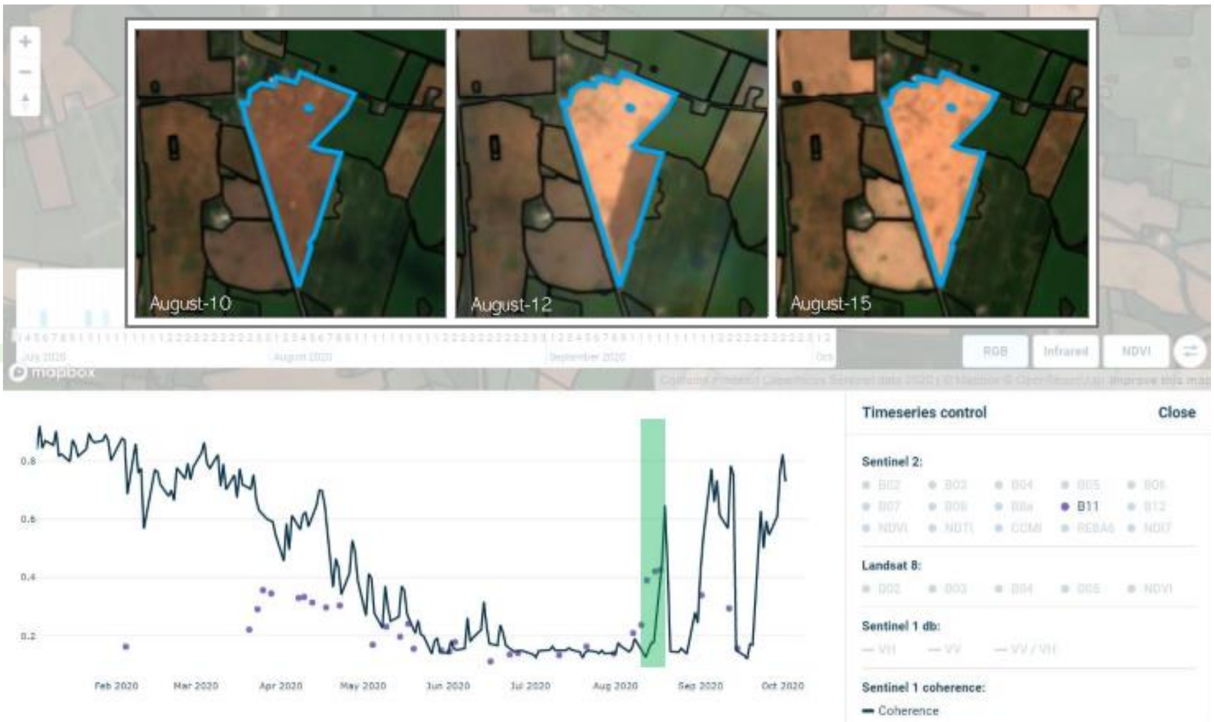


Figure 4 Example Time-series markers combining S2, L8 and S1 imagery in Denmark

NOTE: Figure 4 is coming from DHI GRAS (EO for Agriculture under Pressure workshop - Session 1)

2.1.3. Sentinel-1 - Radar

The Sentinel-1 radar mission comprises a constellation of two polar-orbiting satellites (Sentinel-1A and Sentinel-1B,) which share the same orbital plane, operating day and night and acquiring imagery regardless of the weather. Sentinel-1 is performing Synthetic Aperture Radar (SAR) imaging, i.e. The sensor aperture is managed by digital means. SAR is a type of active data collection which means that the sensor emits its own energy, in the form of a signal (in a given band - interval of frequencies) and then records the amount of that energy reflected back (reflected backscatter) after interacting with the Earth. SAR data require a different way of thinking compared to the optical data because the signal is instead responsive to surface characteristics like structure and moisture. Radar images contain information divided into the phase and the amplitude of the wave.

Sentinel-1 emits in band C to a central frequency is 5,405 GHz and to a wavelength of around 5 cm. The wavelength indicates what the radar is able to detect. Sentinel-1 wavelength is well-adapted to detection of agricultural activities.

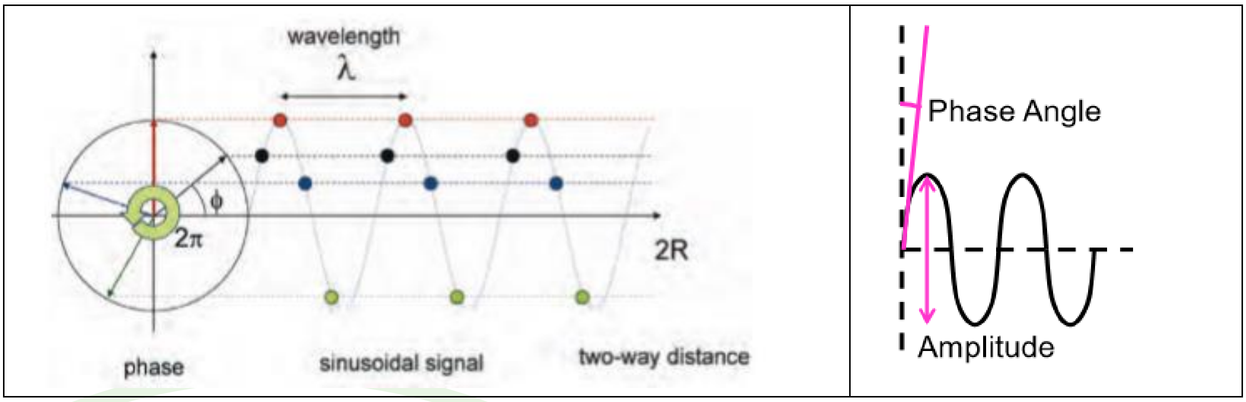


Figure 5 Radar signal is a wave characterised by its amplitude and phase

Sentinel-1 is a polarised radar; in theory, a polarised radar can emit a signal according two different directions: horizontal and vertical. Some materials react more to horizontal polarisation than to vertical one or vice versa. Therefore, radar polarised signal are powerful for detecting the nature of soil material or soil cover.

In practice, Sentinel-1 is emitting only according the vertical direction. The reflecting signal (called backscatter) is then reflected with some deviation angle, depending on the material met by the signal. The reflecting signal may be decomposed into a horizontal and a vertical components.

There is a coding convention:

- First letter gives the polarisation of the emitted signal
- Second letter gives the polarisation of the reflected signal component

For instance, VH is the horizontal component of the reflected signal providing from a vertical emitted signal.

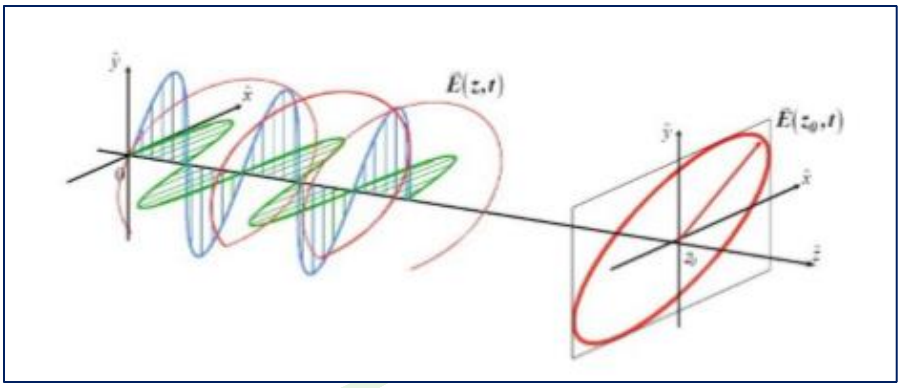


Figure 6 Vertical and horizontal polarisation of Sentinel 1 images

In addition, Sentinel-1 may have an ascending orbit (from South to North) or a descending one (from North to South).

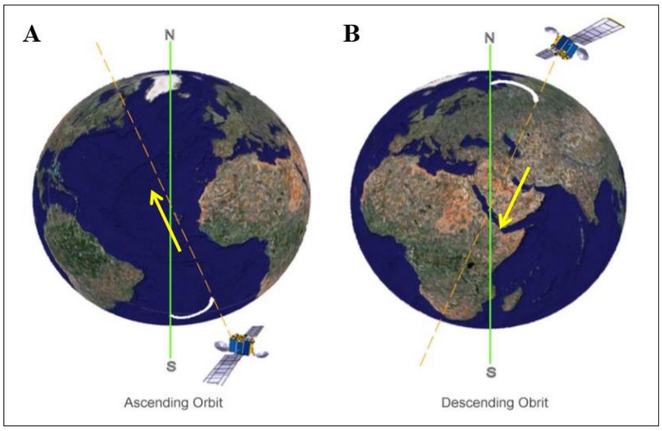


Figure 7 Ascending and descending orbits of Sentinel 1

The difference of the reflected signal according to the orbit direction depends on the terrain roughness: it is very weak in case of smooth terrain but high if the terrain is ruguous (sand beach, ploughed field ….).

More explanations about radar may be found on:

https://gis1.servirglobal.net/TrainingMaterials/SAR/SARHB_FullRes.pdf

Sentinel-1 sensors have a good revisit frequency (around 3 days in Europe).

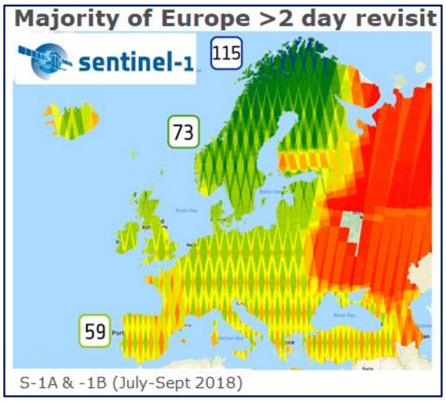


Figure 8 Sentinel-1 visit frequency

Radar can penetrate clouds and because of this, these data provide an advantage over optical imagery. Radar signatures from the VV, VH and VH/VV backscatter ratio (in decibels) from Sentinel-1 can be used for vegetation monitoring as polarisation temporal profiles can be created for inspecting the vegetation signals.

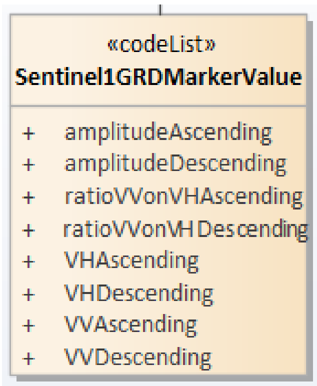


Figure 9 The markers usually derived from a single Sentinel 1 image

2.1.4. HHR-VHR imagery

In the NIVA context, HHR (High High Resolution) applies to ground pixel size less than 5 m and equal to or more than 1 m whereas VHR (Very High Resolution) applies to ground pixel size less than 1 m.

- **Possible sources of images**

These images are characterised by a spatial resolution smaller than 5 meters to even less than 1 meter. There are many commercial providers offering high-resolution satellite data. These are usually tasked imagery, taking an image over a given Area of Interest on specific day.

HHR			VHR		
Optical imagery					
Satellite/ Sensor name	**Spectral bands**	**Spatial resolution (m)**	**Satellite/ Sensor name**	**Spectral bands**	**Spatial resolution(m)**
Spot-6/7	B,G,R,NIR, PAN	1.5	Worldview constellation	Coastal Blue, B,G,Y,R, Red-edge, NIR, PAN, SWIR	0.3-0.5
Planet (planetscope)	B,G,R,NIR	3.125	Geoeye	B,G,R,NIR, PAN	0.5
Jilin-1	B,G,R,NIR, PAN	1	Triplesat	B,G,R,NIR, PAN	0.8
			Kompsat constellation	B,G,R,NIR, PAN	0.5-0.7
			Pleiades	B,G,R,NIR, PAN	
			SupeView constellation	B,G,R,NIR, PAN	0.5
			Planet (Skysat)	B,G,R,NIR, PAN	0.5
			Pleiades Neo	Deep-Blue,B,G,R,Red-edge, NIR, PAN	0.30

Radar images					
Satellite/ Sensor name	**Polarisation**	**Spatial resolution (m)**	**Satellite/ Sensor name**	**Polarisation**	**Spatial resolution (m)**
Radarsat	HH; HV; VH; VV; HH+HV; VV+VH; HH+VV+HV+VH	3-100	Capella SAR	HH	0.5-1.0
			ICEYE SAR	VV	0.25-0.5
			TerraSAR-X	VV, HH, HV, VH	0.24-0.6
			Kompsat-5	HH, HV, VH, VV	<0.85
			COSMO SkyMed	HH, VV, HV, VH	0.9-1.0

Table 3 HHR and VHR satellite images

HHR and VHR imagery are or may be used by Paying Agencies for two main purposes: the current control system and the new Area Monitoring System or EO Monitoring.

- **Current control system**

High and Very High Resolution imagery is provided to Paying Agencies by the European Commission during the CwRS Campaign for specific acquisition windows for each agricultural control zone. That relates to the yearly CAP image aquisiton work programme where each Member State requests the EU to obtain the satellite imagery for the controls via DG JRC.

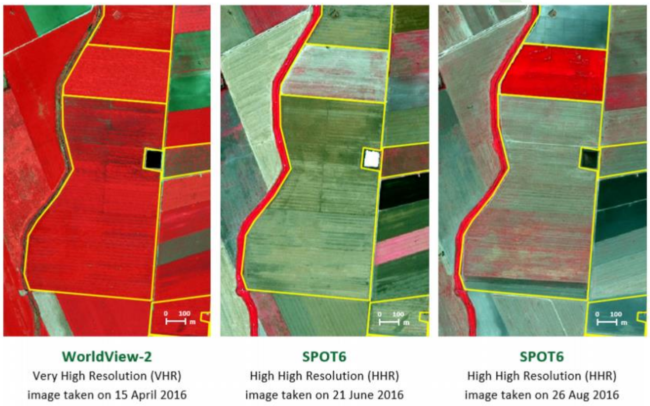


Figure 10 Examples of HHR –VHR images for 2020 CwRS campaign – JRC source

These images are traditionally used for the on the spot checks (OTSC) controls for Basic Payment Scheme / Single Area Basic Payment Scheme and crop diversification. This usually applies to a sample of around 5% of farmers to verify farmers' declarations and adherence to eligibility rules.

- **New Area Monitoring System**

Also, some countries (e.g. Malta) use some HHR imagery for checking the small parcels as the spatial resolution Sentinel is not sufficient for concluding on the eligibility status (insufficient number of pixels falling entirely within the field boundaries).

Use of VHR or HHR imagery might also be used for validation on sample areas as providing an independent source of data.

2.2. Images pre-processing

Satellite image processing plays a crucial role in the core processes, workflows and methodologies for CAP satellite monitoring. However, the images coming directly from the satellites are not directly suitable for analysis. A number of pre-processes are required to get ARD (Analysis Ready Data).

2.2.1. Optical images

2.2.1.1 Basic processes

Basic processes are the processes that are required before performing EO monitoring.

- **Unzipping/ extraction**

Usually EO data are provided in an archived format (.zip, .tar.gz or else). Thus, it is needed to be extracted (decompressed) by decompressing software. That process opens the contents of the file where the individual satellite bands are stored in ***tiff*** or ***jp2* file** format together with the metadata of the image.

- **Geometric correction & orthorectification**

Raw EO data are provided in the sensor geometry but they are not georeferenced i.e., they are not associated with an Earth-related coordinate reference system. The ortho-rectification aims to provide georeferenced EO images, by correcting the discrepancies due to the sensor position and orientation during the image capture (using knowledge of the satellite orbit and Ground Control Points) and the discrepancies due to the orography (using a Digital Terrain Model or a Digital Elevation Model).

Global NASA's SRTM (Shuttle Radar Topography Mission) 30 meter resolution is the DEM often used for orthorectification but some countries have their own datasets, generally with finer resolution.

The accuracy of the ortho-rectification process is generally varying between ½ pixel size and 1 pixel size. It is an indication about the uncertainty of the true location of the boundary pixels of a parcel.

For Sentinel-2 imagery the geolocation accuracy is evaluated by ESA, having subpixel accuracy (<10m). For Landsat 8 images the geolocation accuracy is 12 m. The geometric correction is usually applied to HHR/VHR imagery using a DEM and ground reference data (Ground Control Points (GCP)). The JRC provides to the Member States or to their contractors the image acquisition specifications concerning the geometric error in the output images.

- **Atmospheric correction**

When acquiring EO data, the atmospheric composition can create noise to the acquired imagery, due to the scattering and absorption effects occurring on the returning radiation. These factors can be created through models that try to eliminate such noise. These models are classified as Top of the Atmosphere (TOA - measuring the ratio of the radiation reflected and recorded by the sensor to the incident solar radiation on a given surface) or Bottom of the Atmosphere (BOA) which employ atmospheric models that correct the scattering and absorption effects of the atmosphere.

- **Mask creation**

Optical satellite imagery is not always reflecting the earth's surface (what is of interest for EO monitoring) as there may be natural obstacles (clouds, shadows, haze and snow) preventing to capture an unobstructed view of the ground.

Clouds (and other artifacts) can create significant issues during EO data processing, thus, based on the desired processing, one can use a cloudless image, or employ a cloud mask algorithm to eliminate cloudy areas. Clouds make the spectral signatures of the satellite imagery change, which can cause erroneous or misleading patterns of land cover characteristics. Thus, accurate cloud and shadow detection in optical satellite imagery is of high importance for CbM analysis processes such as crop classification and image interpretation. To quickly process large amounts of satellite imagery data such as the Sentinel-2, various cloud detection algorithms have been used and incorporated into the automated satellite image processing chains.

- Sen2Cor and MAJA are the cloud detection algorithms most frequently used for Sentinel-2 images. Both are correcting Sentinel-2 Top-Of-Atmosphere (TOA) products from the effects of the atmosphere and deliver a Bottom-Of-Atmosphere (BOA) reflectance product. They also provide Aerosol Optical Thickness (AOT) map and Water Vapour (WV) map
  - Sen2Cor was developed by Telespazio on behalf of ESA. It provides Scene Classification (SCL) map with Quality Indicators for cloud and snow probabilities, in JPEG 2000 image format. The BOA data production runs on the systematic basis over Europe with dissemination through the Copernicus Open Access Hub since May 2017
  - MAJA was developed by French and German spatial agencies (CNES, DLR). Its multi-temporal algorithm is based on the assumption that surface reflectance tends to change slower in time than cloud cover. It provides in addition: set of masks for clouds, cloud shadows, topographic shadows, snow and water. The MAJA chain was used by the Sen4CAP project.
- Fmask, Tmask, LaSRC (Land Surface Reflectance Code) are rather used for Landsat images
- ATCOR is provided by PCI Geomatica and applies to various satellite sensors (MR/HR/HHR/VHR)

For further reading regarding atmospheric correction methods please refer to
https://www.sciencedirect.com/science/article/pii/S2666017220300092

NIVA
NEW IACS VISION IN ACTION

2.2.1.2 Advanced processes

- **Mosaicking**

Mosaicking consists in merging several acquired images into a single seamless image. Mosaicking enables to perform EO monitoring on the desired area of interest, e.g. on an administrative unit or on a bio-geographical region instead of on a satellite image tile.

- **Cross-calibration**

This process consists in making compatible the spatial and spectral calibration between sensors (e.g. Sentinel2, Landsat8). That process is important to create consistent time-series among various sensors for continuous monitoring. For further reading, refer to

https://earthdata.nasa.gov/esds/harmonized-landsat-sentinel-2

- **Super-resolution & fusion**

For sensors that do not acquire a panchromatic band to perform pan-sharpening but provide different-resolution images (e.g. Sentinel 2 provides multispectral imagery in 10m, 20m and 60m resolution for the same scene), advanced fusion techniques can be developed to downscale the lower-resolution bands. This process usually includes the development of advanced deep learning algorithms to super-resolve the lower-resolution bands of various sensors while preserving their spectral characteristics. Down sampling all bands to the finest band resolution is very useful for reaching a high degree of detail in the analysis.

Fusion and super-resolution may also be applied to spectral bands coming from different image sensors but this makes the process even more difficult.

- **Vegetation indices calculation**

Spectral indices are (usually) pixel-based operations between at least two bands of the same satellite aiming to enhance the spectral signature of a certain material. There are more than two hundred spectral indices reported in bibliography (e.g. https://www.indexdatabase.de/). The Normalized Difference Vegetation Index (NDVI) designed to enhance the spectral signature of vegetated surface is the most famous one. There are various other Vegetation indices that are used such as Green Vegetation Index (VIg), LAI (Leaf Area Index), FAPAR (Fraction of Absorbed Photo synthetically Active Radiation), FCOVER (Fraction of Vegetation Cover), Normalised difference tillage index (NDTI), BSI (Bare Soil Index) ...

2.2.1.3 Temporal series

EO monitoring is generally not using single date satellite images but temporal profiles. These temporal profiles may be provided according several options:

- Supporting geometry : pixel or Feature of Interest (e.g. agricultural parcel)
- Semantic: image band or derived index (see advanced processing's).
- Frequency: raw data or regular frequency.

The temporal series may be provided as raw data, i.e. by providing the list of valid observations (without clouds) with their capture date, implying a relatively high data volume (as all the dates have to be provided).

The temporal series may also be provided as regular frequency data, generally by interpolation of the missing observations. This gap filled temporal series may be simpler to be used (more compact format, regular frequency) but the interpolation is not always fully reliable (see chapter about quality).

NOTE: OGC is working on a standard for temporal profiles: TimeseriesML 1.0 defines an XML encoding that implements the OGC Timeseries Profile of Observations and Measurements [OGC 15-043r3], with the intent of allowing the exchange of such data sets across information systems.

2.2.1.4 Various levels of image products

Sentinel 2 images are provided according various levels of product (ESA classification), according to the performed preprocesses.

The first levels address single-date images:

- L0: raw data (in sensor geometry – Top Of Atmosphere data)
- L1A : decompressed data (internal product – not available)
- L1B : basic radiometric corrections (still Top Of Atmosphere data)
- L1C: ortho-rectified data (in terrain geometry - still Top Of Atmosphere data)
- L2A : atmospheric correction (terrain geometry – Bottom of Atmosphere data)

Generally, data at level L2A include the Sentinel images themselves (orthorectified and BOA) and the mask image(s).

For EO monitoring, the Analysis Ready Data are the S-2 images at level L2A.

However, the atmospheric correction may be done according different models (e.g. various cloud mask detection algorithms). In case a Paying Agency or its technical partner have expert(s) in atmospheric corrections, it may be better to use as input data L1C images and to choose the most adapted process, for instance to adapt to local conditions.

Products of level L3A correspond to regular frequency temporal series of S-2 images and products of level L4A correspond to regular frequency temporal series of derived indexes (NDVI, LAI, FCOVER).

For instance, Sen4CAP is providing L4A products (from S-2 images) with a 10 days frequency.

2.2.2. Radar Imagery

2.2.2.1 Levels of basic (single date) products

Sentinel 1 images are provided in the form of different product levels depending on the performed pre-processes. According to the ESA classification these products are:

- Level 0 (L0): compressed raw data that include noise
- Level 1 (L1): decompressed data (Doppler centroid estimation, sensor geometry and precise satellite orbit, single look complex focusing). They are processed into the following:

  - Ground Range Detected (GRD) products contain amplitude information of return signal
  - Single Look Complex (SLC) products: preserve phase information of return signal
- Level 2 (L2): Ocean (OCN) products that include:
  - Ocean Wind field (OWI)
  - Ocean Swell spectra (OSW)
  - Surface Radial Velocity (RVL)

Obviously, OCN products are not of interest for CAP monitoring.

2.2.2.1 Basic processes on GRD images

The basic pre-process presented in following paragraph apply on a single date GRD image. The explanations below are coming from the SNAP (ESA Sentinel Application Platform).

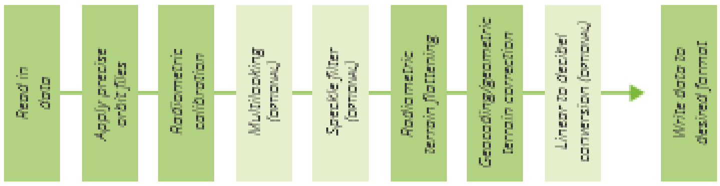


Figure 11 General data flow of S-1 (GRD) images

Apply precise orbit files	Applying precise orbit files enables first geolocation. In general, real time orbit vectors contained in the metadata of SAR products are not accurate. Precise orbit files are issued by the European Space Agency within weeks after the acquisition of a data set. The operation of applying a precise orbit available in SNAP (Sentinel Application Platform) provides accurate satellite position and velocity information. The precise orbits are not annotated in the image data directly but are rather provided as a separate file.
Radiometric calibration	Radiometric calibration converts backscatter intensity - as received by the radar sensor - to the normalized radar cross section as a calibrated measure, taking into account the sensor-specific characteristics (global incidence angle of the image, etc.). This makes radar images of different dates, sensors, and/or imaging geometries comparable.
Multilooking	If the DEM is of lower resolution than the SAR data, the SAR data may have to be multilooked (reduced in resolution) before processing. This step is not required if the DEM is of similar resolution as Sentinel-1 data. This is the case of the SRTM DEM that is widely used to ortho-rectify S-1 data.
Speckle Filtering	Sentinel-1 images suffer from a 'salt-and-pepper' effect called "speckle". Speckle, appears in SAR images as noise and it is due to the interference of waves reflected from many elementary scatterers. Low pass filters are usually applied to images for the reduction of radiometric noise .Various filters may be selected according to the data type, the spatial resolution and the speckle degree of the input radar image.
Range Doppler terrain correction	Range Doppler terrain correction is a correction of geometric distortions caused by topography, such as foreshortening and shadows. These discrepancies are corrected by using a Digital Elevation Model. This operation is also called "Radiometric Terrain Flattening".
Geocoding	The captured geometries are geo-transformed and projected to a map coordinate system, making them

	also comparable to other geospatial products
Linear to decibel conversion	When the backscatter intensity is represented in a linear scale, the image interpretation is usually hard to be implemented due to low contrast. To achieve a better contrast of the values, a transform of the radar image values into a logarithmic scale is applied. This image enhancement process stretches the current values so that better distinction among them is accomplished.

Table 4 Basic steps of GRD radar image (pre-) processing

NOTE: The process of ortho-rectification is not done in the same way for S-1 and S-2 images; ortho-rectification of S-1 is mainly based on precise orbit files (no Ground Control Points). This implies some delays to get the precise orbits from ESA and a georeferencing accuracy between 1 and 2 pixels.

2.2.2.3 SLC images and their processes

The Interferometric Wide (IW) swath mode is the main acquisition mode over land for Sentinel-1. It acquires data with a 250-km swath at 5-x-20-m spatial resolution (single look). Interferometric wide mode captures three sub-swaths using the TOPSAR acquisition principle.

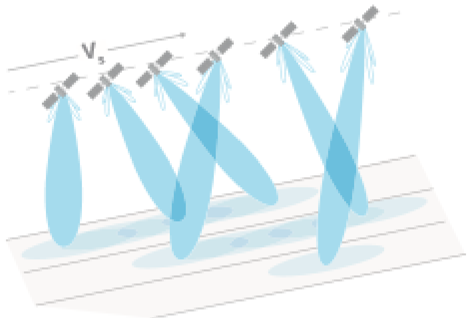


Figure 12 TOPSAR acquisition principle

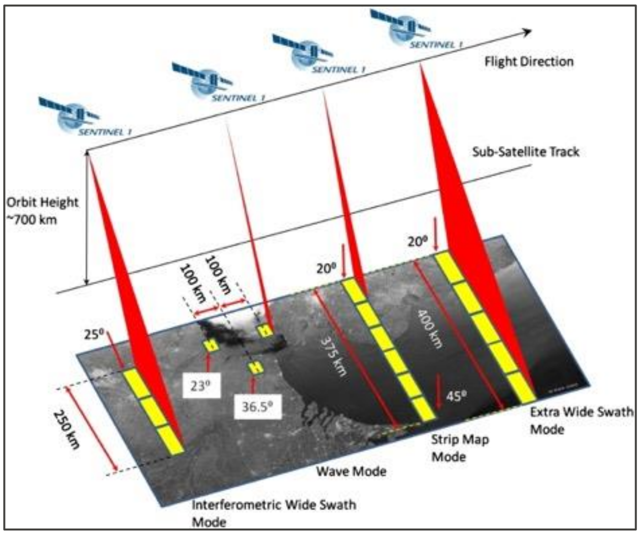


Figure 13 Sentinel-1 Product Modes (Credit: ESA)

Interferometric wide SLC products contain one image per sub-swath and one per polarization channel, for a total of three (single-polarization) or six (dual-polarization) images in an IW product.

SLC are complex products; their main interest is the fact they keep the phase information (in opposite to GRD products that contain only the amplitude information).

For EO monitoring, the main process consists of the coherence computation, i.e. the phase comparison between 2 products at 2 different dates but on same area and taken on the same orbit. Coherence estimation is very important in radar imaging, as it attempts to measure the similarity between two (complex) synthetic aperture radar images. The measure of dissimilarities indicates potential change regions. Maintaining coherence means we maintain consistency in the phase relationship between signals. The coherence is useful to detect the changes of surface properties. Low values of coherence indicate something that is not changing, such as bare soil.

The computation of coherence implies a set of pre-processing:

- Choice of the input images
- First georeferencement of the 2 input images (knowledge of accurate satellite orbits)
- Choice of the interesting part of each input image
- Co-registration (the 2 input images are combined to form only one output image)
- Coherence computation
- Processing of emitting bands (to ensure a single continuous image)
- More accurate georeferencing (terrain correction).

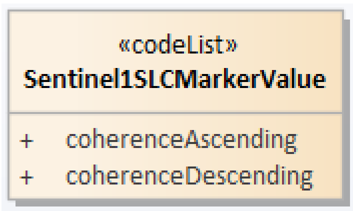


Figure 14 Usual markers from Sentinel-1 SLC products

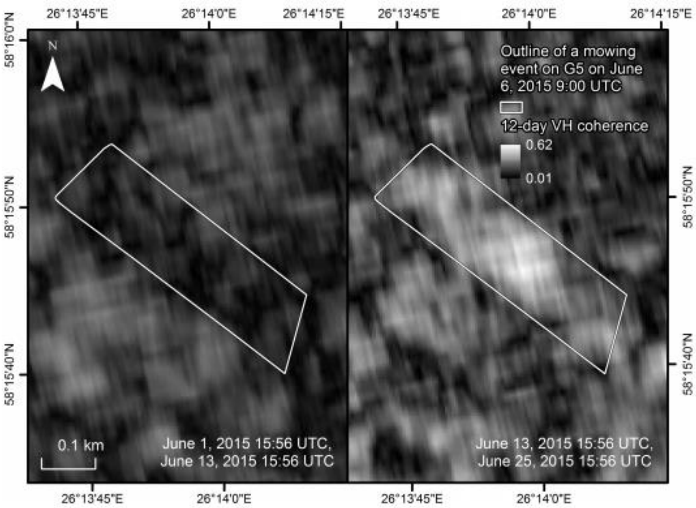


Figure 15 Sentinel 1A 12-day VH coherence before (left) and after (right) a mowing event

NOTE (source of figure 15): Tamm T, Zalite K, Voormansik K, Talgre L. Relating Sentinel-1 Interferometric Coherence to Mowing Events on Grasslands. Remote Sensing. 2016; 8(10):802. https://www.mdpi.com/2072-4292/8/10/802

2.2.2.4 Temporal series

EO monitoring is often performed on temporal series of S-1 data (as with S-2 images) but without the issue of missing observations due to clouds. Radar image time-series may be used to monitor crop dynamics for a variety of crop types. VV and VH backscatters and the ratio VH/VV can be reliable crop growth indicators for monitoring agricultural practices such as plough or harvesting (figure 16).

For instance, Sen4CAP is providing weekly and/or monthly temporal series of amplitude (according the various polarities) and of coherence.

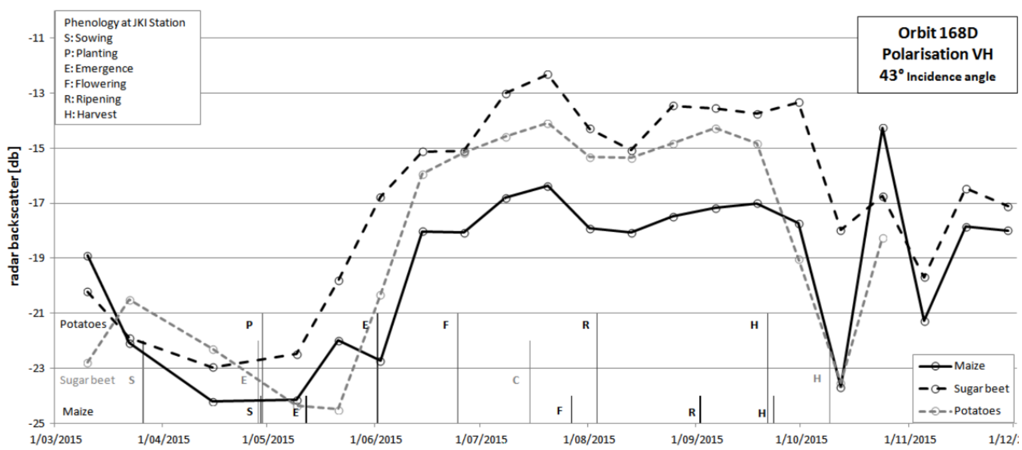


Figure 16 Radar backscatter signatures for maize, sugar beet and potatoes

NOTE (source of figure 16): Lilienthal H, Gerighausen H and Schnug E 2016. First experiences with the European remote sensing satellites Sentinel-1A / -2A for agricultural research. In: 13th International conference on Precision Agriculture, ISPA, Monticello, IL, USA, pp. 1–11

2.3.Solutions to deal with Sentinel images

2.3.1. Overview

Four main functionalities are expected from the EO monitoring infrastructure:

-quick and easy access to Sentinel data

- access to data storage capabilities

- computation power for the EO monitoring processes (e.g. crop classification)

- possibly, data delivery capabilities (view …)

- **Taking images from their source: the natural but not so easy solution**

The Sentinel images come from ESA whose official site is Copernicus Open Access Hub (https://scihub.copernicus.eu/). The previous name was Sentinels Scientific Data Hub; many people are calling it ESA's scihub.

The images may be downloaded, either manually or through a REST API that conforms to the Open Search/ open data search standards.

The ESA Hub offers as available data, S-2 images at levels L1C or L2A (obtained from Sen2cor processor) and S-1 images at level L1 (GRD and SLC).

Currently, the most common method to access Sentinel data is to download the Sentinel archives on the area and period of interest, to uncompress the images and then to read the images.

This method has two main constraints: first, the download of Sentinel archives is time-consuming as open-data API are limiting the output flow and number of downloads in parallel to 2 images. The second point is that the method requires significant storage place (1GB per Sentinel-2 zipped file and 2GB per uncompressed, imported Sentinel-2 image).

There is another constraint related to the images located in long term archives (LTS). These images acquired more than one year ago are stored on magnetic bands (or other non-direct access support) and the procedure to get them is more complex than for recent images. First, one has to make a demand to transform the images from Long Term Storage to temporary storage, accessible for download (this may take 1 day) and then one has to do the download demand itself.

In addition, some images pre-processing in order to get ARD (Analysis Ready Data) would be of great interest. For instance, Sentinel-1 images are currently provided as open-data almost as raw data and require a lot pre-processing before they can be used as input for crop classification or any other EO monitoring process. These pre-processings are consuming both data storage place (raw S-1 images + pre-processed data) and of computation power to perform the pre-processings.

Regarding Sentinel-2, the preparation of temporal series interpolated at regular frequencies (e.g. every week) would simplify the processing chain of EO monitoring.

- **Possible solutions**

The main possible solutions include:

- Use of national portals: these portals are generally providing Sentinel images (at least) on the national territory, usually working as mirrors of the ESA Hub; they may also provide added value by offering better pre-processed Sentinel images or additional satellite images. They may facilitate the access to satellite images but they don't offer any storage or computation resources.
- "home" made infrastructure: the Paying Agency (or its technical partner) builds its own infrastructure, ensuring enough storage and computation power for accessing, pre-processing and analyzing the necessary images;
- Copernicus Data and Information Access Service (DIAS): they provide centralised access to Copernicus data and information (including Sentinel images), as well as to processing tools
- Other cloud providers may also provide similar functionalities.

The following paragraphs are describing the experience collected by the NIVA project about these issues, trying to provide more details about the above solutions but also about the Sen4CAP project and derived initiatives.

2.3.2. Use of national portal

2.3.2.1 Overview of existing portals

Country	National Mirrors and Collaborative Ground Segments
Austria	https://data.sentinel.zamg.ac.at/dhus/#/home
Belgium	https://www.terrascope.be/
Czech Republic	https://dhr1.cesnet.cz/#/home
France	https://peps.cnes.fr/rocket/
France	https://theia.cnes.fr/atdistrib/rocket/#/search?collection=SENTINEL2
Finland	https://finhub.nsdc.fmi.fi/#/home
Greece	https://sentinels.space.noa.gr/
Germany	https://code-de.org/en/about
Italy	http://ec2-3-208-162-171.compute-1.amazonaws.com/fdai/platforms/details/14
UK	https://sedas.satapps.org/

Sweden	Swedish Space Data Lab
the Netherlands	Dutch National Satellite Data Portal https://satellietdataportaal.nl/ * Satellietdataportaal.nl is only accessible from within The Netherlands
Portugal	https://ipsentinel.ipma.pt/dhus/#/home
Norway	https://colhub.met.no/#/home
Luxemburg	https://www.collgs.lu/

Table 5 List of national portals

2.3.2.2 Use of THEIA by UC1b

- **Context**

UC1b is developing agro-environmental indicators whose computation requires as input data temporal series of NDVI coming from Sentinel-2 images. The default solution is to get this input data from Sen4CAP but other means to get NDVI series are also possible, as illustrated by figure below.

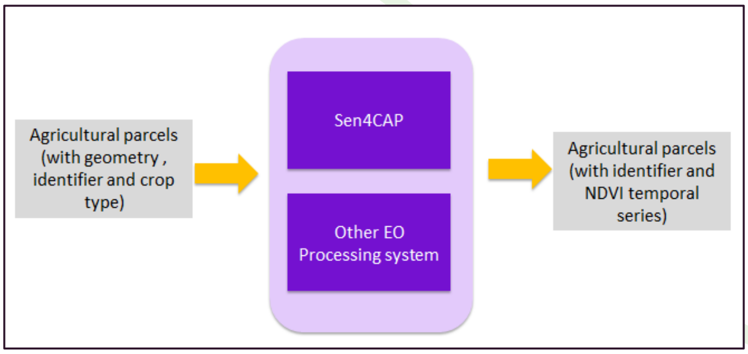


Figure 17 How to get NDVI temporal series for agro-environmental indicators

The tool was developed using sample data provided by the Sen4CAP project on a part of French territory (as France was one of the pilot partners of the project) during some experimentation on crop monitoring, i.e. in another context and for a different objective.

In order to test the tool at larger scale, i.e. on other areas and other campaign year, the UC1b team had to get more NDVI temporal series. The first attempt was to try to install Sen4CAP to derive the necessary input data but the Sen4CAP installation was more difficult than expected (see chapter 2.3.5.2). This is why it was decided to adopt an alternative solution: the use of the national portal THEIA.

- **What is THEIA?**

In the mythology, THEIA is the daughter of Ouranos (the Sky) and of Gaïa (the Earth). This why this name was chosen by a consortium of several French institutions working in the domain of Earth

Observation and environmental sciences. THEIA was created in 2012 in order to provide centralised access point for EO satellite images and derived products, making easier the use of EO data.

The portal offers access to optical satellite images at least on France and sometimes on other countries, depending on the product. Sentinel-2, Landsat, Venus images are provided as value-added products whereas Pleiades, Spot 6/7 , SWH & RapidEye are provided as more basic images (only orthorectification). The portal doesn't provide access to Sentinel 1 images that is more difficult to pre-process. The portal also offers access to thematic products, such as humidity or land cover maps.

- **How was THEIA used by UC1b?**

THEIA provides S-2 images at level L2A. The images are first downloaded from THEIA, from a query about date and area of interest. User gets a zip that has to be decompressed.

In practice, for each band of S-2, there is both the image itself and the related mask, obtained from the MAJA treatment. Both are orthorectified in same Coordinate Reference System.

The values of the mask are coded on 8 bits, each bit corresponding to a case where the soil can't be seen in good conditions (cloud, water, snow, shadows, irrelevant sun position …): the bit is set à 0 if there is good condition and at 1 if there is bad condition.

To compute the NDVI temporal series, the UC1b team has selected the relevant bands of S-2 images (B4 and B8, both at 10m resolution), computed the NDVI (that is a combination of these 2 bands) and withdraw the mask pixels, i.e. the pixels for which there was at least one bit with "1" as value.

The choice of considering that the only valid pixels are those with (0, 0, 0, 0, 0, 0, 0, 0) mask value is also the choice of Sen4CAP.

The UC1b team has developed a convenient tool that takes care automatically of all those steps.

2.3.2.3 Dutch National Satellite Data Portal.

For several years now, the Dutch government is acquiring and publishing commercial VHR-HR satellite data of the Netherlands as open data on the National Satellite Data Portal which is maintained by the Netherlands Space Office, the space agency of the Dutch government. However, the portal does not provide Sentinel images, it just refers to the ESA's scihub.

The satellite imagery is freely available to all users within the Netherlands. This is highly beneficial for many sectors including agriculture. More specifically, farmers and agri-cooperatives use the images to increase the sustainability and efficiency of their production. Among the users of that portal, there are also many governmental organisations such as RVO (Netherlands Enterprise Agency), the Dutch Paying Agency.

2.3.3. "Home" infrastructure

2.3.3.1 Experience from ITACYL (Spain)

- **Context**

One of the key issues met by technology companies operating with satellite data is the access and pre-process to large volumes of satellite images, especially due to the context of new EO monitoring system.

Instituto Tecnologico Agrario de Castilla y Leon (ITACYL) has as its main activity the technological development of the food and agriculture sector. Within ITACYL, the Infrastructures department is focused in providing services to the government companies and farmers through the use of geo-technologies.

In this context, ITACYL processes systematically Sentinel images in order to produce every year a land cover map with crop identification. The product is a layer with multiples uses related to the control of CAP subsidies.

- **Experience with home infrastructure**

Among the various solutions to access to satellite images, ITACYL uses home infrastructure. ITACYL has developed its own infrastructure based on the download of Sentinel Images through Copernicus Open Access Hub. Sentinel-2 satellite images are mainly obtained, specifically L2A level images, i.e. atmospherically corrected. Then different geoprocessing (masks computation, index computation …) are carrying out depending on the result it wants to obtain.

Due to the territorial extension of Castile and Leon, it is necessary to download the images that concern the R137, R093 and R037 orbits Sentinel 2; a total of 33 granules corresponding to these orbits are analysed. The download of the granules depends on the cloud coverage and the composition of the orbit.

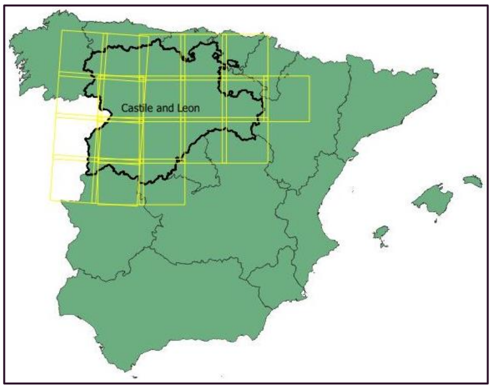


Figure 18 Granules used for EO monitoring in Castilla-Leon

As a sample, in the last campaign year, 74 orbits have been processed and 805 granules has been downloaded, i.e., the 33% of the granules available. Processing of these images generates 5400 GB of data that are stored in the network.

In the context of NIVA project, ITACYL is involved in WP2 in the outscaling of Use Cases. Until this moment, ITACYL has worked to estimate empirically the net annual CO2 flux at both the parcel level. and the pixel level (UC1b: Agro-environmental monitoring). To carry out this process, it has been taken advantage of our processed data, located in the system. From these data ITACYL has obtained the net annual CO_2 flux.

- **Benefits of home infrastructure**

ITACYL considers that having its own infrastructure to access to satellite images and its own processing to obtain results is an added value. In addition, this developed workflow is completely free and independent, thus it is beneficial for ITACYL. It is important to say that when ITACYL began to work with a huge amount of satellite data (It has been making Castile and León crops and natural land maps since 2011) there were no other alternatives such us DIAS or other cloud infrastructures, therefore its workflow is adapted to work with home infrastructure.

2.3.4. Use of DIAS

In 2018, the European Commission launched the DIAS initiative. DIAS (Data and Image Access Services) are commercial solutions aiming to facilitate the access to Sentinel images for Paying Agencies in order to support them for setting up the new EO monitoring system. However, PAs are free to consider alternative implementations of CAP monitoring, on the basis of technical and commercial criteria.

DIAS are expected to include access to Sentinel data and Copernicus services outputs, (logical) co-location of cloud computing resources for scalable fast parallel processing and access to relevant processing modules for raster and vector data.

The European Commission defined a set of minimum requirements for DIAS, such as the expected available data products and services.

2.3.4.1 DIAS assessment by NIVA

NIVA conducted by end 2019 a comparative analysis of the available DIAS in order to choose the DIAS to be used by the project. This analysis was updated in May 2020 in order to provide a state-of-play as recent as possible.

Currently there are 5 DIAS available. Their development started in 2016 for CREODIAS, SOBLOO, MUNDI, ONDA and in 2018 for WEKEO. They all provide data and cloud resources in various flavours:

- Data offer: Copernicus, Landsat, other open products, commercial EO data
- Cloud resources: IaaS (Infrastructure as a Service) as a baseline, advanced/customized services may also be available

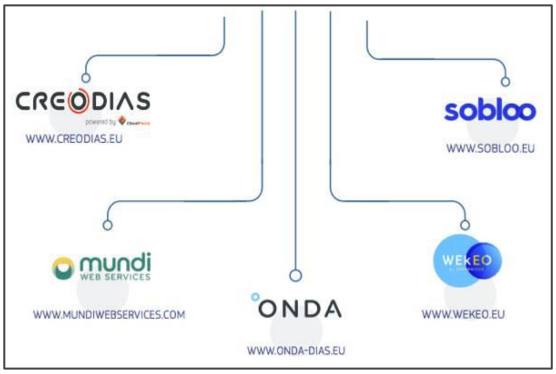


Figure 19 The 5 available DIAS

The comparative analysis has been conducted identifying key characteristics about the image product offer and the computation power. The results of this analysis are presented on the 2 next pages.

Regarding the table about DIAS data offer:

NOTE 1: Missing data retrieval is necessary in case (that occurs sometimes) the synchronization between the scihub and DIAS is not working well, in this case DIAS normally re-trying to download original products from the source (scihub). This requires more time than the case of a download from local copy.

NOTE 2: The ordering mechanism efficiency depends on the scihub availability. It can be minutes if data is on-line but also days if data is off-line (for Long Term Storage data as scihub has also a data policy of 12 month rolling-archive).

NOTE 3: Some of the DIAS (Mundi, Sobloo) offer precalculated masks

Regarding the table about DIAS cloud service comparison:

NOTE 1: OpenStack is a Cloud API that can be used to use Cloud IaaS resources, such as Create/destroy a VM or Create/destroy Storage volumes or Object Storage buckets.

NOTE 2: Sen4CAP has been natively developed on CREODIAS and there is a MI (Machine Image) already configured on CREODIAS that can be used. Without CREODIAS, the installation of Sen4CAP requires a system administrator with strong Linux experience

NOTE 3: Cold storage is a storage “tier” that costs less because data has to be retrieved with some latency (normally hours).

NOTE 4: A cost comparison was done on by end 2019 based on specific use cases that were submitted to the different DIAS for quotation. Time was unfortunately missing to refresh this cost analysis. In any case, prices are really similar among the cloud providers. High prices normally are from advanced cloud providers (Mundi/Telekom Cloud) while other providers have lower prices with less Cloud capabilities.

	CREODIAS	Mundi	ONDA	Sobloo	WEKEO
Sentinel 2	L1C: full archive L2A: Orderable (also non-ESA) rolling cache 1PetaByte	L1C: last 12 months L2A: last 48 months (only Europe data)	L1C: full ESA archive L2A: full ESA archive	L1C, L2A: orderable, available last 9 months	L1C: full metadata, orderable
Sentinel 1	SLC: full archive in EU, 6 month worldwide, GRD: full archive		Full archive for SLC and GRD. Part of the archive are on cold storage (delayed retrieval available)	SLC, GRD: orderable, available last 9 months	GRD, SLC: full metadata orderable
Landsat 5/7/8	Landsat 5/7/8 full archive over Europe	Landsat 7/8 orderable	Available since 04/2018 (for Europe)	Landsat 8 On-demand	-
Missing/other data retrieval	Ordering/Caching mechanism available	Missing L2A can be retrieved from ESA or processed if not available	Missing data can be Retrieved and hosted in native format. Available VHR commercial data (orderable)	Spot sample data available/orderable	Many datasets from Climate/Meteorology

Table 6 DIAS data offer

	CREODIAS	Mundi	ONDA	Sobloo	WEKEO
Availability of Cloud API	Openstack API	Openstack API	Openstack API	Different API Available	Morpheus API
Pricing Mechanism	On-demand, IaaS and various automation to create personal clusters (e.g. Kubernetes)	On-demand, advanced IaaS services available (e.g. Container service)	On-demand, IaaS and some EO processing tools	On-demand, IaaS and various Managed Software (Databases, Middleware)	Based on subscription or on-demand. On-demand is also available on request.
Cost report	Available	Available	Available using an API	Available the day after	Not clear, probably yes
SEN4CAP Support	Supported in Sen4CAP, a configured Machine Image is available	Supported in Sen4CAP	Supported in Sen4CAP	Supported in Sen4CAP	-
Apps and other	WMS/WCS/Data Cube, Configured Virtual Machines	Configured Virtual Machines	Configured Virtual Machines	Configured Virtual Machines	Jupyter Notebook,
Cold storage	Not Available	Available	Available	Available	Not Clear

Table 7 DIAS cloud service comparison

The choice of a DIAS for NIVA has been done according the following main characteristics: data more suitable for Agriculture applications, minimal Cloud services available, easiness to configure Sen4CAP.

As a result, CREODIAS was the DIAS chosen by the NIVA project.

2.3.4.2 IGN study about interest of using DIAS.

This study was conducted by end of 2019 by IGN (French Mapping Agency) in order to help the French Paying Agency to decide on the infrastructure to be adopted for performing EO monitoring experimentations.

The study consisted in identifying the expected functionalities and to compare the costs and benefits of DIAS (in general) with other potential solutions.

- **Quick and easy access to Sentinel data**

The main purpose is to get a direct access to images rather than having to download and uncompress them. The constraints linked to the download of archives (uncompressed images) can be mitigated by several solutions:

- Get access not to a catalogue of archives but to the uncompressed images themselves. In France, this may be done by using the national portal called THEIA
- Get access to a data cube or a solution like Google earth engine. Their related API enables to handle Sentinel-2 data in a direct and very simple way: https://openeo.org/
- Get very quick access to archives without parallel download constraints. This would enable to implement solutions that download only necessary input data, launch processing's and throw the downloaded data at the end of computation. This would significantly decrease the need for data storage. However, this method requires specifically designed code. This code is not yet available for operational EO monitoring (it exists only as research results).
- Get quicker access to the download of image part (e.g. only the bands or area of interest) through a dedicated protocol or API. This might be done using storage at object level and an API type Amazon 3 that provides access to each Sentinel band in tiff/geotiff format. This might also be done using a (paying) DIAS service.
- Get access data to images through a download API of WCS type. In these cases, the issue would be to adapt existing research code to ensure it may use the WCS image flows without too many queries.

In conclusion, there is a set of solutions to decrease the requirements of data storage, some being service offers from DIAS (in this case the effort comes from DIAS and makes client life easier) and some requiring adaptation of research codes and the use of a classical cloud infrastructure (in this case, the main effort comes from the client).

- **Computation power**

The volume of Sentinel data in input of EO monitoring is so big that the data processes are often shared between several machines in order to get quick results. S-2 data being supplied as tiles (100 km x 100 km), a simple solution is to spread the tiles on various machines and to process them independently.

A better solution is to use libraries and software dedicated to computation distribution. This software enables to optimise the use of computation resources and to decrease the infrastructure costs. They also enable more flexibility in the computation (organisation of tiles).

A priori, there are two main families for scientific computation distribution:

- scientific computation clusters of HPC type: the infrastructure is supplied with a set of software/API facilitating its use. The user splits the planned computations in more or less elementary tasks and assess the necessary resources (time, processor number, RAM requirements) and then submits the pile of computation tasks (possibly ordered) through a standardised protocol to the computation cluster that has in charge to launch the computation on the relevant machines and to warn in case of issue. In France, this solution is accessible through the Iota2 chain (satellite image processing chain) that uses the CNES cluster
- solutions based on a cloud infrastructure. The user has to set and configure the various elements enabling to make the computation distribution work. This second solution is more flexible but requires advanced IT skills (netwok infrastructure – cloud).

DIAS offer only the second type and let users define the whole infrastructure, to configure and manage it. They may provide help to do so, as paying services (what may be also proposed by other cloud infrastructure providers).

DIAS solutions were considered as less user-friendly than computation clusters solutions in the context of EO monitoring experimentations. They may be more adapted when the process is made operational as offering more flexibility to optimize the work process chain.

- **Cost/benefit assessment**

During these experimentations, the DIAS offer did not look mature enough and did not seem to present significant advantage compared to a classical cloud infrastructure.

The documentation about the various DIAS that is available on Internet was rather poor : no examples of use, no white paper for setting up the infrastructure or for basic S-2 preprocessings. Apparently, DIAS may propose more advanced solutions than simple image download but these offers were not publicly described and presented with their related costs and clear conditions of use.

2.3.4. Use of cloud infrastructure (other than DIAS)

Several commercial providers propose cloud solutions offering more or less the same functionalities as DIAS. Some examples are provides below:

Google: the public Google Cloud Storage bucket offers SAFE format and covers the EU region

https://cloud.google.com/storage/docs/public-datasets/sentinel-2

The SAFE format (Standard Archive Format for Europe) has been designed to act as a common format for archiving and conveying data within ESA Earth Observation archiving facilities. The SAFE format wraps a folder containing image data in a binary data format and product metadata in XML. This flexibility allows the format to be scalable enough to represent all levels of SENTINEL products. For more info access: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/data-formats

Amazon AWS: https://registry.opendata.aws/sentinel-2/

Planet Sentinel-2 access is included in commercial API

https://developers.planet.com/docs/data/sentinel2l1c/

In addition, some Member States offer also national cloud infrastructure. For instance, CODE-DE is a German cloud very much alike DIAS conforming in addition to the security requirement of Germany (https://code-de.org/).

2.3.5. Use of Sen4CAP

2.3.5.1 Generalities

According to the context, the term "Sen4CAP" may be used either for the project or for the system developed by the project.

Sen4CAP is a project funded by ESA with the purpose to develop tools dedicated to EO monitoring and based on Sentinel images. The project ran from 2017 to March 2021, it was conducted by a consortium of technical partners, with the contribution of pilot Paying Agencies expressing their requirements for EO monitoring tools and testing the tool and its results.

Sen4CAP has developed a system that may be installed locally or using the CREODIAS virtual machine; the system may then automatically download images on the area and period of interest, compute derived products described in this document (such as S-1 Analysis Ready Data, temporal series, vegetation indexes) and more advanced ones directly related to the monitoring of agricultural practices, such as results of crop classification or event detection (mowing, tillage …).

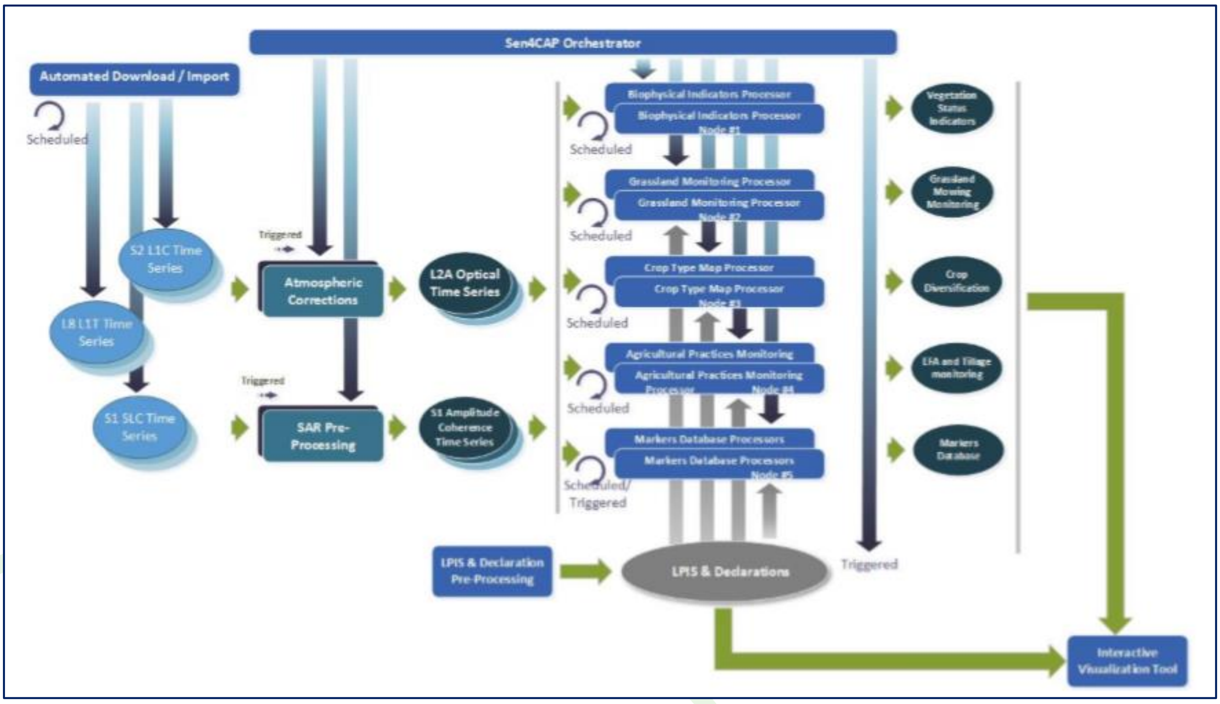


Figure 20 Logical data flow of the EO Sen4CAP processing system

The general strategy of the NIVA project has been to build its EO monitoring related tools by re-using as much as possible what was already done by the Sen4CAP project. The two projects have worked in parallel during around 2 years; the NIVA teams tested several versions of the Sen4CAP system, i.e. some of the difficulties described in the next paragraphs may have been solved by a new version of the software.

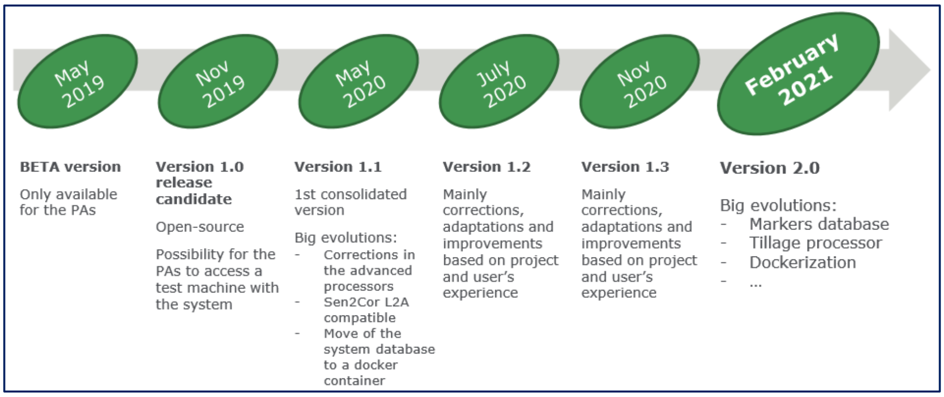


Figure 21 The various versions of Sen4CAP

2.3.5.2 Experience from UC1a (EO monitoring and Traffic Lights)

- **Context**

UC1a, as member of the consortium of the European project NIVA, has exploited Sen4CAP as the main EO classification engine. The output of Sen4CAP is used as input to the NIVA Decision Support System (DSS). NIVA-DSS concludes on a decision regarding traffic light color codes at parcel level, based on specific predefined criteria/rules that are structured and must be met according to the case/scenario under the control target.

- **Various steps in experiencing Sen4CAP**

In the beginning of 2020, the UC1a team requested access to the Sen4CAP platform via a virtual machine for a limited period. This service is provided by the Sen4CAP consortium for a first contact of the Member States' Paying Agencies with the platform.

Then we performed a local installation of Sen4CAP version 1.1 (June2020). In September 2020 Sen4CAPlocal installation was upgraded (in memory and disk, 5.5 TB).

In Sen4CAP2.0 version (released Jan2021), UC1a team investigates the tillage detector. This is not a separate processor, but an additional marker that is integrated in the L4c processor. UC1a team is also currently investigating 'L4b' processor, focusing not on pasture areas but on fallow land. UC1a interest lies on testing 'L4b' in fallow land parcels where a prerequisite of one ploughing/year must be checked and identify the presence of a crop (no production/harvest marker).

- **Feed-back from first use of Sen4CAP (with CREODIAS)**

We made attempts to get familiarized with the platform and the processes. The difficulty was to understand how the input data had to be structured (their appropriate format) so that the LPIS GSAA processor could run properly and consequently the L4A processor to drive a crop type mapping product.

At the level of input data (farmer's declarations) the format of the required fields is very specific. The LUT table has also a specific structure i.e. the field of crops must be an integer while the fields of holdings and parcels must be strings.

Interpreting and resolving error messages was a difficult procedure and most of them could not be corrected without an email communication. The errors that appear in the monitoring table are not detailed enough to help the user understand the actual error in the procedure; a status message 'error' appears instead of 'finished'. If the LPIS GSAA processor hasn't run properly and in the meantime L4A processor has started running, a status message 'finished' is displayed which doesn't correspond to reality (no product is produced). We communicated with Sen4CAP consortium members via mail to resolve those issues.

The Support User Manual helped us, but mainly the online webinars and the dedicated Forum page were very supportive. For example, through Sen4CAP team's response to the chat of a webinar's session, we learned how the LPIS GSAA processor runs.

The downloaded S1 & S2 images refer to the tile-s in which our area of interest is located. Some images are saved that are not needed and cannot be deleted 'easily' (absence of a delete button). So, in each test site the user must decide on which images to use and there is not a ready-to-use tool to select and import them directly into 'L3a' or 'L4a' for further processing. The user must take a note of the tile name (which contains 60 characters) and then identify and select the specific tiles in another tab) so that the processors 'L3a' and 'L4a' run only on the desired images. Such procedure requires a lot of effort and is prone to errors.

- **Experience from the installation process in local system**

The installation trial took place in June 2020 on version 1.1 of Sen4CAP; it was conducted by a software engineer of the OPEKEPE (Greek Paying Agency) UC1a team. A virtual machine specific for the needs of Sen4CAP was created. Since the initial goal was to explore Sen4CAP possibilities and not use it in a production capacity, requirements on the VM were laxed.

The initial VM's characteristics were as follows: 92GB RAM, 8 Cores, 2 TB of storage space; The VM was running in a 16 CPUs x Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz machine.

The operating system was one proposed by Sen4CAP (CentOs 7.7). The latest of the CentOS versions supported was selected in order to be as close as possible to CentOS development.

Installation itself proved problematic due to the fact that some of the CentOS repositories required for installation had changed place and the script could not find the available software. This mishap was due to the fact that the CentOS version required by Sen4CAP was not current; according to CentOS policies older versions of the operating system fetch the required software from specific locations that store older (not current) versions. The problem was quickly fixed by updating the repository list and installation proceeded with no problems.

The second major problem and one that was difficult to solve was the expected MAJA version. Sen4CAP expects MAJA version 3.2.2; this again is not the current version of MAJA that is available for download. MAJA provides 3.2.2 as a download that, as it turned out is not the 3.2.2 version that Sen4CAP expects. When the problem was uncovered, the exact MAJA version required by Sen4CAP was provided to us by the Sen4CAP team and finally it was possible to run Sen4CAP classifications.

However, it was quickly obvious that even for testing purposes a larger storage space would be necessary; the next step was therefore to upgrade the available storage space to 8TB. This is the maximum that can be provided with our current hardware setup.

After the first tests and the release of Sen4CAP version 2.0 it was decided to upgrade our current installation to that version.

Upgrade moved smoothly with no reported errors. After a couple of tests though it was clear that some problems not reported in the upgrade script were present related mainly to the way the *docker* subsystem was installed (see reported issue https://github.com/Sen4CAP/Sen4CAP/issues/8). In this step as in the previous steps too the Sen4CAP team and forum has helped us a lot and provided quick responses on all of our problems.

In order to run classifications spanning back a year it is necessary to request the images to be moved back from the archived section of *SciHub*. This move is not spontaneous and there may be a delay from the time of the request to the actual provision of the image;

These delays seem to affect Sen4CAP operation to the point of not producing results since it is impossible to actually download the EO images required for the analysis.

As a conclusion, Sen4CAP had no problems being installed on a local server. However for production purposes the amount of storage required for the satellite images have made using Sen4CAP on a local server problematic.

2.3.5.3 Experience from UC1b (agro-environmental monitoring)

- **Context**

UC1b is developing agro-environmental indicators whose computation requires as input data temporal series of NDVI coming from Sentinel-2 images. The default and recommended solution is to get this input data from Sen4CAP, as Sen4CAP looked the obvious solution: it is an open-source system, devoted to EO monitoring and it was expected that most Paying Agencies would be willing to install and use it as the required efforts would benefit them both for the future Area Monitoring System to be set up and for the testing of the UC1b agro-environmental computation tools.

For more details, see chapter 2.3.4.1.

- **Experience from the French team**

The installation trial took place in October 2020 on version 1.2 of Sen4CAP; it was conducted by the main developer of the UC1b tools, using a server of CNES (National Centre of Spatial Studies) as the French Paying Agency could not provide a relevant server at least on short delay. The minimum configuration necessary to install Sen4CAP includes 64 Go of memory and 8 CPU; the cost was estimated to 15 000 €.

The installation was not successfully achieved because of strong difficulties and because an easier alternative solution was found (use of national portal THEIA). The difficulties were due to security reasons. The installation of Sen4CAP supposed to write in private folders of the exploitation system, what is forbidden as it entails big risks on security. It is possible to bypass the issue by changing the access path of the installation folders. A script was initiated to change these paths but due to the many dependencies in the Sen4CAP installation, it became too complex and was abandoned.

There is a scheduler because Sen4CAP may run continuously (to download images) implying a strong control on the machine, with password required. As the installer of Sen4CAP was not the machine administrator, this raised an additional issue. Working on a virtual machine as with CREODIAS is less dangerous (the DIAS may propose safeguard).

As a conclusion, Sen4CAP seems to be designed to be installed on a DIAS rather than on a local server. The installation on a local server would be feasible with relevant and significant resources (IT engineer, own server with administrator rights).

- **Experience from testing countries**

Spain, Netherlands and Denmark are the testing countries of the UC1b tools. Spain has produced the NDVI temporal series on a Spanish sample area from S-2 images using its own infrastructure and algorithms. Netherlands has produced the NDVI temporal series on whole country using Sen4CAP.

Until now, Denmark has not yet tested the UC1b tool due to difficulties in getting the input data; the envisaged solution is the use of Sen4CAP running on the virtual machine of CREODIAS but the uncertainty about budget funding is slowing the process.

2.3.6. Open EO API solution

- **Open EO API**

Open API is specification created to have a common way to describe any kind of API. Open EO API is a specification for modern API, based on Open API and dedicated to searching, viewing and processing EO data (images, time series …). Open EO API is for developers, it is a standardised way to describe the API end points and the parameters.

Open EO API offers an interface to process data on a remote server (what is quite useful as EO data may be too big to be easily downloaded and processed on local site). In practice, images may be requested, using either a list of images (i.e. providing the image identifiers) or by using a query (select images on a given area and period of interest).

Some parts of Open EO API will be integrated to OGC new generation of services. There is some overlap mainly with CSW (Catalogue Service for the Web), WMS (Web Map Service), WCS (Web Coverage Service) and WPS (Web Processing Service).

Data from several EO providers is available through Open EO API, for instance the Copernicus services mirroring the ESA Hub (and serving Sentinel-1 and Sentinel-2 data). In general, discovery is free of charge whereas the processing services that consume cloud resources are charged.

More details on Open EO API may be found in Annex 2 of this document and in NIVA deliverable D4.2 Design and activation of basic SW stack for running test.

- **The NIVA Open EO APIs**

The NIVA project is developing a set of micro services based mainly on the marker database coming from Sen4CAP (e.g. temporal series of mean values and standard deviation of S-1 or S-2 bands or derived indexes, at parcel level)

To benefit from these services, user has to install the Sen4CAP system (e.g. in own premises or with CREODIAS). It should be noted that the project ended in April 2021 but ESA will ensure at least basic

Sen4CAP system maintenance. There are still last decisions to be taken about this maintenance. Until now, installing Sen4CAP is the safer solution to ensure that the NIVA micro services will work.

In summary, using the NIVA Open EO APIs micro services won't relax Paying Agencies or their technical partners to install Sen4CAP (what is the main difficulty) but it will help to make workflows more automatic, more integrated (no longer need of manual data file transfer). These APIs are of interest if the Paying Agency is not using Sen4CAP alone but combined with other tools. For instance, they might be quite useful for combining in an operational way Sen4CAP with the Decision Support System of the NIVA UC1a.

2.3.7. Data as a service

- **EO-widget project**

EO-Widget is a project supported by ESA; it was launched in January 2021. The project aims to provide commercial services based on the processors developed by Sen4CAP.

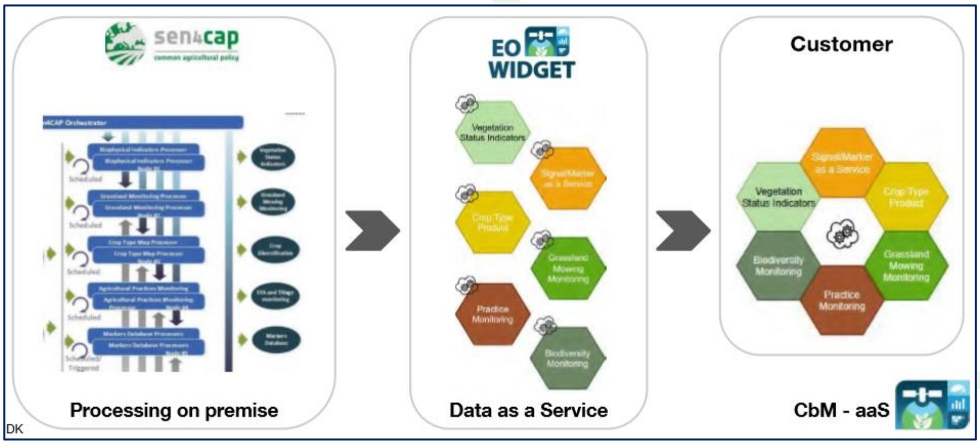


Figure 22 The EO-Widget principle

The final objective is to propose Checks by Monitoring as a service. The EO monitoring products are generated on-the-fly: the Paying Agency (or its technical partner) provides the input data, such as agricultural parcels with farmer's declaration and the service provides the requested markers or the results from more advanced processes such as crop classification or event detection (e.g. mowing).

It will be opened to several data or service providers, enabling other services to be developed. For instance, the project is proposing a quality assessment tool. The "market place" will enable users to find what the available services are.

The EO-Widget services are very promising because they would alleviate the burden of costly satellite data management and signal processing: as the results may be provided at parcel level, there will be no need of handling big data volumes and no need to install Sen4CAP (on-premise or using a DIAS).

The above information comes from a presentation given during a Sen4CAP webinar (18/05/2021). Unfortunately, at this date, the business model was not yet prepared, i.e. there was no idea how much these services might cost.

- **Use of Copernicus phenology services**

Copernicus Land Monitoring Service offers over the 39 EEA member states the High Resolution Vegetation Phenology and Productivity (HR-VPP) Copernicus product that is derived from Sentinel-2 imagery.

Product Themes	Product	Main characteristics
Basic Vegetation Index	NDVI	10 m spatial resolution, Sentinel-2 UTM grid daily, available within 12 hours
	LAI	10 m spatial resolution, Sentinel-2 UTM grid daily, available within 12 hours
	FAPAR	10 m spatial resolution, Sentinel-2 UTM grid daily, available within 12 hours
	PPI	10 m spatial resolution, Sentinel-2 UTM grid daily, available within 12 hours
Seasonal Trajectories	PPI	10 m spatial resolution, Sentinel-2 UTM and ETRS89-LAEA grid 100 m spatial resolution, ETRS89-LAEA grid 10-daily, available after year-end
Vegetation Phenological and Productivity	VPP	10 m spatial resolution, Sentinel-2 UTM and ETRS89-LAEA grid 100 m spatial resolution, ETRS89-LAEA grid yearly up to two seasons, available after year-end

Table 8 Copernicus phenology products (from Technical Specification Issue 1.1 on 2020.11.25)

The plant phenology index (PPI) is a physically based vegetation index and has a linear relationship with green leaf area index. It is the only index that is available as temporal series (or seasonal trajectory).

The “Vegetation Phenological and Productivity” layer includes a set of phenological, and productivity parameters such as the start-of-season, end-of-season, max-season-value, etc. It provides these parameters on a yearly basis.

For EO monitoring, the added value of the Copernicus phenology services is about its offer on vegetation indexes and about the more advanced results provided by the VPP layer.

The image files are provided in COG (Cloud Optimized Geotiff) format. The Vegetation Indices (VI) are provided in the original Sentinel-2 tiling grid, with UTM projection, without any spatial resampling. The results may be available through various services.

Service Category	Service	Main characteristics
Discovery	CSW	OGC Catalogue Search Web service
Viewing	WM(T)S	OGC Web Mapping Tile Service
Machine2Machine	WCS	OGC Web Coverage Service
	REST	Stateless File Download Service
Direct Data access	IDP	Identification Service
	HTTPS	Secure File Download Service
	GeoTrellis	TimeSeries analysis service (see https://proba-v-mep.esa.int/api/timeseries/apidocs/)

Table 9 The Copernicus phenology services

2.3.8. Monitoring experiences

- **Survey results (questionnaire)**

One of the key issues met by Paying Agencies is the accessing and pre-processing of big volumes of satellite images in the context of the new EO monitoring system approach. For this reason, we have performed a short survey to get more feedback about PAs experiences with Earth observation technologies for monitoring. The survey was sent out to the NIVA partners and to the official monitoring countries out of NIVA (Belgium-Flanders, Malta), asking various questions on the use of EO data and technologies in monitoring controls for area-based CAP payments. We received responses from 7 paying agencies (out of 11 – 63.6% response rate).

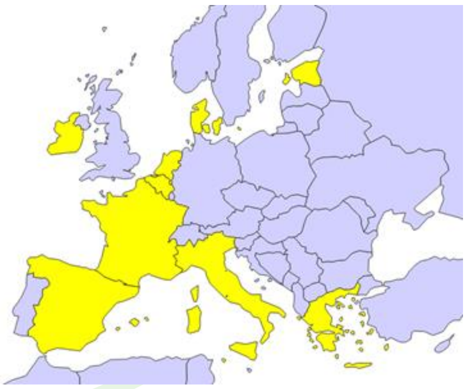


Figure 23 The Member States who participated in the survey

- **EO monitoring**

Our survey indicates that at the moment, there are no specific payment schemes officially checked by the PAs in the monitoring approach. However, there are various schemes partially applied by some PAs in the EO approach such as single area payment scheme (SAPS), Basic Payment Scheme (BPS), small farmers scheme (SFS) and coupled support (VCS).

As regards the size of the area of EO monitoring application, the respondents implement checks by monitoring in various extents ranging from nationwide coverage to couple of thousands parcels.

- **Accessing the satellite images**

Our survey shows that Sentinel images (both optical and radar) are heavily used by all paying agencies. On the contrary, this doesn't apply to NASA’s Landsat-8 images where only 1 PA out of 7 answered that it uses them.

Figure 24 The satellite images most widely used for EO monitoring

Concerning accessing satellite images by the PAs, interestingly enough, the most used way for downloading/accessing the images is via a commercial web service (e.g. Amazon, Google). DIAS platform is the second most widely used method with 29%. The Copernicus open access hub is only used by 1 PA (out of 7). 71% of the respondents obtain S2 Level-2A products (pre-processed and atmospherically corrected by Sen2Cor processor). Further, 57% of PAs are interested in accessing old Copernicus Sentinel data (older than 1 year) via LTA.

- **(Pre)-processing**

According to the results of our survey, 5 out of 7 PAs perform in-house satellite image processing using open-source software. GDAL and SNAP are the most widely used toolsets. Only one PA uses commercial software (i.e., PCI GEOMATICA) for image processing. Also, 2 paying agencies use DIAS platform for image processing. As regards the processing steps, our survey shows that almost all PAs perform cloud-masking (filtering the clouds) and NDVI calculation. The presence of cloud cover affects the temporal and spatial availability of the observations. That is reflected in our survey results, as 85% of the Paying Agencies use masking methods to remove cloud cover. The most dominant methods used are MAJA and Sen2Core. However, 5 out of 7 PAs have not assessed the cloud cover issues regarding the impact on the completeness of time-series observations.

Concerning S1 images, all PAs use the L1 GRD/SLC products of S1. More than 70% of PAs perform in-house processing using SNAP or other open source software. Many PAs report facing various technical challenges when dealing with radar imagery such as difficulties in signal interpretation and missing or duplicated slices on ESA-hub. Finally, our survey shows that more than half PAs use a web-portal for visualization of the results of the controls.

- **Dealing with small parcels**

Small parcels are usually assigned yellow color-codes, mainly due to insufficient number of Sentinel 2 pixels falling within the parcel boundaries. Regarding the definition of small parcels, our survey showed that PAs have different definitions. The size of small fields ranges from 0.1-1ha. Also, it was reported that the polygon's (parcel) shape plays a significant role in characterizing inconclusive small parcels. For example, various elongated, irregular or concave polygons may be flagged as yellow due to their particular shape, even though they may extend a certain parcel size.

For tackling the small parcel issues, most PAs consider acquiring higher resolution satellite imagery (e.g., Planet, SPOT-6/7, Worldview). One PA considers using advanced satellite image fusion methods to downscale Sentinel2 images.

- **IT system**

Adopting checks by EO monitoring suggests dealing with bigger and more complex data. For this reason, PAs should consider upgrading and modifying their IT systems. According to our survey, PAs applied various changes to their systems as they can be seen in figure below.

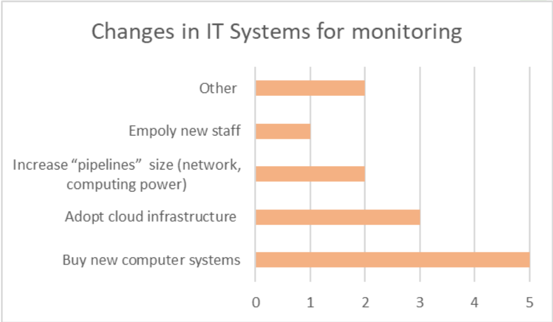


Figure 25 The changes in IT system required to perform EO monitoring

Overall, through the EO monitoring exercise the MS gained valuable knowledge and experience as they familiarized with sentinel data and cloud environments as well as with innovative concepts and developments. Although the positive experience PAs had, various issues concerning specific EO knowledge and skills, sentinel image quality, time needed for complex processing and IT infrastructure were reported by the PAs

2.3.9. EO data quality issues and quality assessment

2.3.12.1 Introduction

The quality of EO monitoring depends mainly on the relevant choice of input images and of the data analysis, of the EO monitoring itself (i.e. of the methods used for crop classification or agricultural event detection). Chapter 2.1 of this document aims to provide relevant information to help readers to satisfy the first condition (relevant input images). The EO monitoring quality assessment is out of scope of D3.5.

However, the preliminary steps (image pre-processing) may also influence the quality of the EO monitoring results. The next chapters provide some considerations on the main quality issues identified during the NIVA project.

2.3.12.2 Temporal series of S2 images

The main quality issue is due to the cloud impact on S-2 images, providing gaps (missing observations) in the temporal series. The impact of these missing observations is especially disturbing if the aim is to detect an agricultural event (mowing, tillage ….).

- **Mosaicking of S-2 images**

The principle is to combine S-2 images on a 10 days or one month period or even more in order to get images free of clouds.

This service (Sentinel 2 Global Mosaic) is proposed by the Sentinel hub (https://apps.sentinel-hub.com/mosaic-hub/#/) and also by CREODIAS, EGI (Advanced Computing Services for Research) Mundi, COIH (Commercial Operator Identity Hub), CODE-DE.

For CAP monitoring, the mosaicking of S-2 images looks of interest if the mosaic is performed on a 10 days period; in case a longer period is required, there is significant risk that the time series becomes not dense enough for event detection.

- **Use only S-2 data and document the uncertainty**

First example is provided by the NIVA Use Case on agro-environmental indicators (UC1b). The carbon indicator (CO_2 flow) is computed, for an agricultural campaign period, using as input NDVI temporal series from S-2 images and applying then a threshold value between bare soil and active vegetation. The missing values have been filled by linear interpolation which is the simplest but not most reliable method (as shown in next figure).

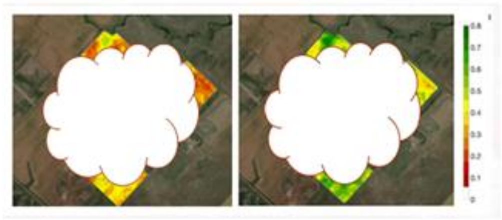
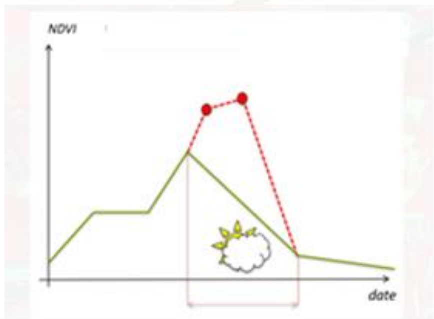
	
2 missing dates due to clouds	Linear interpolation may be not valid

Figure 26 The risks of temporal series interpolation

To mitigate this issue, the UC1b is providing the carbon indicator together with a set of metadata attributes (number of valid observation, width of the longest hole …) that aims to document the reliability of the CO_2 flow. However, this solution is not considered as very user-friendly and the UC1b team has envisaged but not yet succeeded to provide a single aggregated quality indicator: there are remaining issues. For instance, should this quality indicator measure the intrinsic temporal series quality (what would make it usable for any application) or should it be tailored to measure only the reliability of the CO_2 flow (i.e. it would then depend on the chosen threshold between bare soil and active vegetation)? The second option looks the most appropriate but its benefit would be limited to the very specific case of CO_2 flow indicator.

The second example is coming from the Copernicus phenology services. The PPI Seasonal Trajectories is derived from a function fitting of the time-series of the raw PPI values thereby acting as a time-series filtering technique. The fitting function is not described but from next figure, it looks more elaborated than a simple linear interpolation.

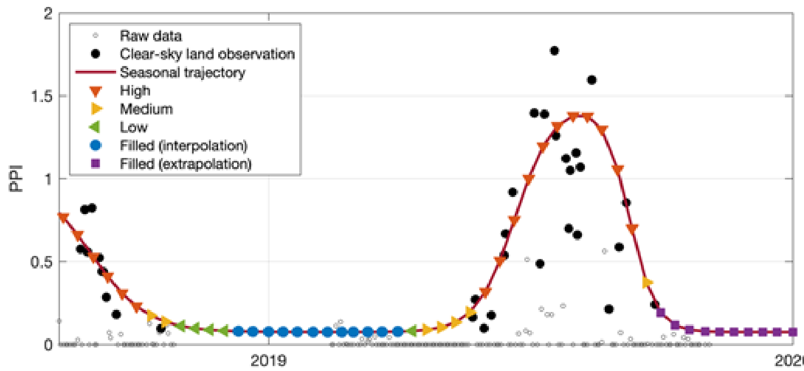


Figure 2 Seasonal trajectory and its quality flags. The seasonal trajectory values are extracted from the regular trajectory on dates 1st, 11th, and 21st in each month.

Figure 27 The temporal series of Copernicus phenology services

The quality of each observation is coded according to the following table

Value	Quality	Definition
5	High	More than 8 Clear-sky land observations found in a 91-day-window
4	Medium	3 to 8 Clear-sky land observations found in a 91-day-window
3	Low	1 to 2 Clear-sky land observations found in a 91-day-window
2	Filled (interpolation)	Clear-sky land observation(s) found on both left and right sides but outside a 91-day-window
1	Filled (extrapolation)	Clear-sky land observation found on one side (left or right) but outside a 91-day-window
0	No data	The time series was not processed.

Table 10 Reliability of temporal series coding (Copernicus Phenology services)

- **Use markers both from S-2 and S-1**

This is the choice of Sen4CAP. For instance, the grassland mowing detection product (L4B) is based on the processing of S1 and S2 derived products, which are respectively coherences, calibrated amplitude backscatter (from S-1 images) and three basic indicators (from S-2 images) which are the NDVI, the FAPAR and the LAI. It is reminded that the L4B are temporal series at regular frequencies (holes being filled by linear interpolation)

The final result is decided by combining the reliability result of the set of indicators.

- **Fill the S-2 holes by S-1 data**

The most advanced way to solve this issue is to use S-1 data in order to complement the temporal series of S-2. IGN (French Mapping Agency) is conducting some research on it, for mowing detection.

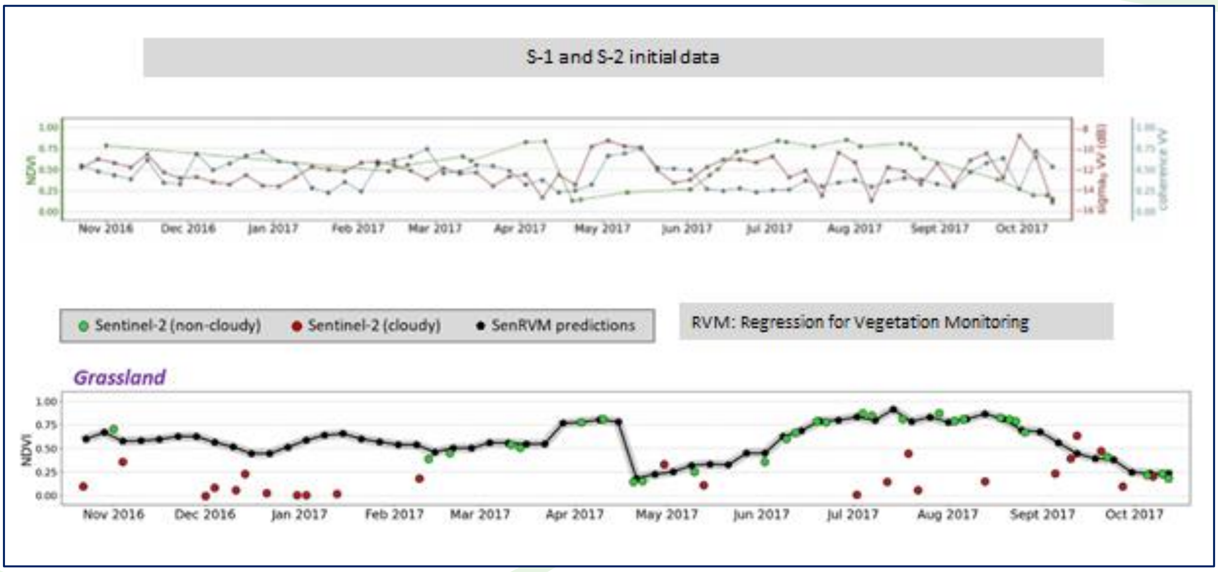


Figure 28 The gap filling of S-2 temporal series using S-1 data

The principle is to use temporal series of NDVI coming from S-2 images considered as the most convenient marker.

- When available, the S-2 observations are used as training data; the aim is to “push” S-1 data to simulate the behaviour of NDVI
- Once the model is trained, the S-1 data are used to fill the gaps of the S-2 temporal series of NDVI.

The publication may be found on ResearchGate.

2.3.12.3 Aggregation at parcel level

The quality of the EO monitoring results is also impacted by the decisions taken about how the pixels have been selected and how their values have been aggregated.

- **Boundary pixels**

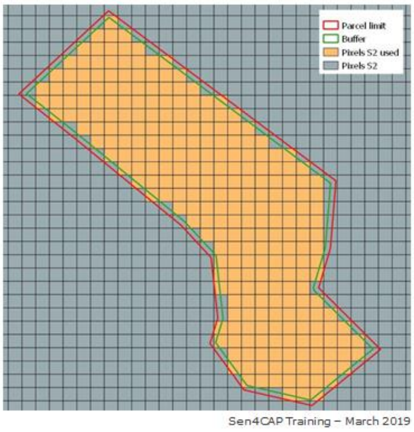
	The rules used to exclude the boundary pixels should be clearly documented. The exclusion of boundary pixels should aim to select only: - The pixels that are semantically significant, i.e. those that are fully or at least partly within the given parcel - The pixels that are geometrically significant, i.e. correctly located within the given parcel. Taking into account the uncertainty due to image georeferencing, it may be worth to keep only the reliable pixels fully within the parcel or to apply an inner buffer.

Figure 29 The exclusion of parcel boundary pixels

		
The initial geometry of agricultural parcel	The landscape features (trees, pond) or ineligible features (lane) may be in separate layers.	The parcel geometry that should ideally be considered for EO monitoring

Figure 30 The exclusion of landscape or ineligible features

- **Minimum number of pixels**

Obviously, a parcel can be EO monitored only if it contains at least one pixel of the selected image type (ex: S-1, S-2) on which the data analysis may be performed. It may be chosen (or not) to decide on more demanding constraints: there is no consensus but various practices.

JRC conducted a study about "Applicability limits of Sentinel-2 data compared to higher resolution imagery for CAP checks by monitoring".

https://publications.jrc.ec.europa.eu/repository/bitstream/JRC115564/small_parcels_case_study_tech_report_pubsy_last.pdf

The evidence from this study suggests that the size of the parcel expressed as an area in metric units is not an appropriate measure to define the 'small parcel' in the context of CAP checks by monitoring. Another two geospatial parcel parameters instead seem to manifest the correlation with the ability of the sensor to provide interpretable NDVI signal providing expected information. These parameters are: 1) number of full pixels remaining after application of a buffer, expressing both the size of the parcel and the position of the parcel regards to the raster; 2) the percentage of pixels lost after application of the same buffer, considering the shape and position of the parcels regards to the raster.

There is a strong probability that the Sentinel-2 NDVI time series could deviate from the Planet Scope NDVI time series above the acceptable threshold for parcels that:

- contain less than <u>8 full pixels</u> and at the same time
- the percentage of S2 pixels lost after application of 5m negative buffer is higher than 60%.

Another example is provided by Sen4CAP that is imposing a minimum of <u>3 pixels of S-2 images</u> within a given parcel.

However, a reasonable balance between the reliability of the results and the efficiency of the EO monitoring process has to be found. In case of agricultural landscape with lots of small parcels, imposing this minimum of 3 pixels of S-2 images may conduct to an exaggerate number of yellow lights. In this case, Paying Agencies may decide to decrease this <u>minimum number to 1 pixel.</u> This choice was mentioned by Slovenia during the Panta Rhei webinar (07/10/2021).

This issue was also met by OPEKEPE in the NIVA project when reusing Sen4CAP results for the Decision Support System of UC1a: "SEN4CAP has a limit on the size and shape of the parcels being assessed. The centroids of 3 pixels at least of S2 images must be located within the respective polygon and 1 pixel of S1 images. Decreasing the number of non-assessed parcels was an important issue for us. The maximum number of parcels that can be assessed is achieved by using the 'L4a' processor initially with both S1&S2 and then, for the remaining non assessed parcels, a second 'running' is performed only by using S2. To reach this conclusion, lots of attempts were made and Sen4CAP Forum provided us a great help. Given the fact that Greece has a high diverse landscape, the testing sites were selected to cover a variety of crop types and agronomic practices. As expected, the results differ. In general, there are several non-assessed parcels. i.e. in a control zone, with predominant land cover arable (Thessaly region):

- Using both S1&S2 resulted in 23% non-assessed parcels.
- Using only S2 decreased the percentage of non-assessed parcels to 9% (but with lower confidence levels)"

This issue of small parcels that can't be monitored using Sentinel data may be mitigated by applying the JRC guidelines that recommend to aggregate adjacent parcels with same declared crop or practice under FOI (Feature of Interest).

- **Valid pixels**

Pixels may be correctly located within a given parcel but observation may be missing in case of optical sensor (cloud issue). Following user requests, Sen4CAP will document the number of valid pixels used to compute an average value of a band or vegetation index on a given parcel.

More generally, the exact formula used to compute the average or median value at parcel level should be clearly documented.

- **Parcel heterogeneity**

The main risk is to launch an EO monitoring process on a parcel that the farmer has forgotten to split. This risk may be mitigated by applying some preliminary checks. For instance, Sinergise is proposing a homogeneity marker [from Panta Rhei webinar – April 2021]. In a similar way, NIVA UC2 (Prefilled application) has developed a methodology for preliminary parcel delineation that can detect this type of error as soon as the farmer makes his/her declaration.

The tool with its code and documentation is available on the NIVA GitLab.

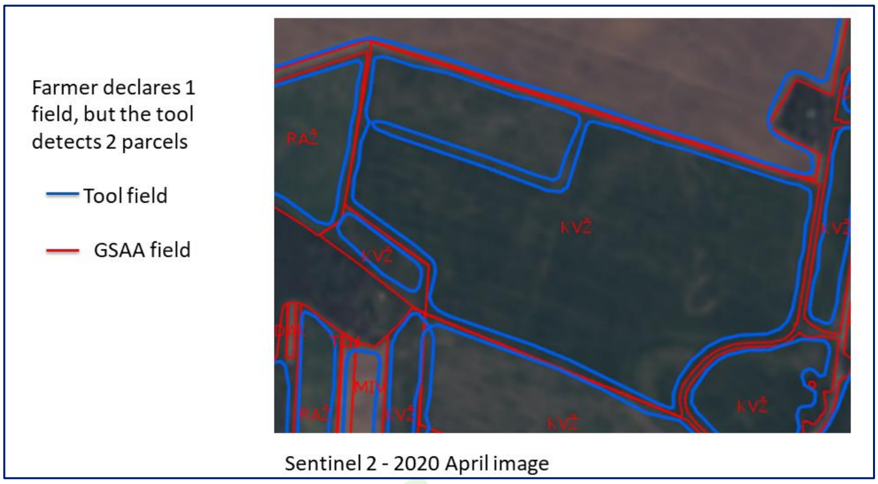
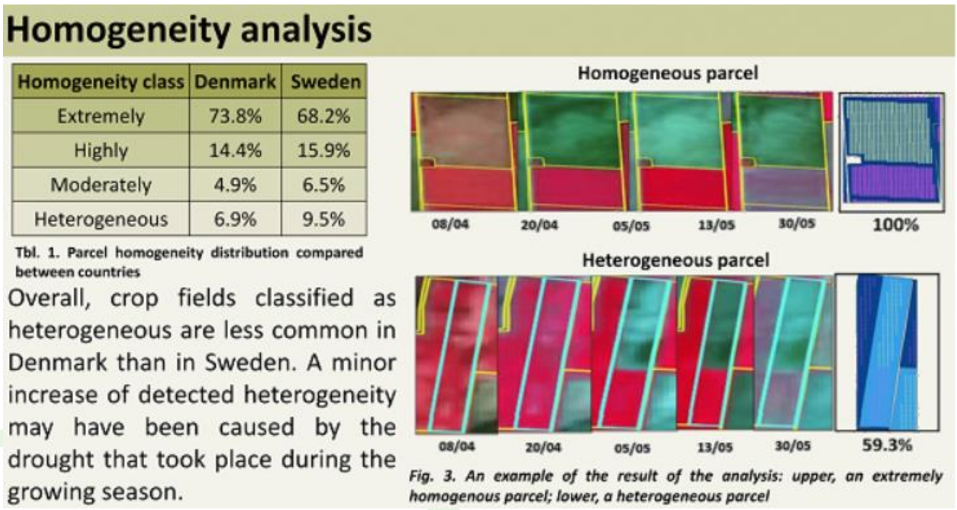


Figure 31 Preliminary parcel delineation

Also in 2018, the Danish Agricultural Agency and the Swedish Board of Agriculture performed a homogeneity analysis to flag the heterogeneous parcels. The work was performed by NEO BV using Sentinel data and automated machine learning techniques.



Homogeneity class	Denmark	Sweden
Extremely	73.8%	68.2%
Highly	14.4%	15.9%
Moderately	4.9%	6.5%
Heterogeneous	6.9%	9.5%

Figure 32 Automated detection of parcel homogeneity in Denmark and Sweden. Source: NEO BV, 24th MARS Conference

More generally, it may be worth to know if the selected pixels of a given parcel have more or less similar behaviour regarding EO monitoring markers in order to analyse if the average (or median) value at parcel level is really meaningful.

The Sen4CAP statistics products at parcel level include the standard deviation that gives an idea about the parcel heterogeneity.

The Open EO API developed by NIVA are including micro-services related to input data homogeneity ; the intrafield service enables to detect the various heterogeneities occuring on a parcel and these results may be used to decide if more boundary pixels should be excluded (because influenced by neighbour parcels).

2.4. Main learning's

- **Using S-2 images**

S-2 images have rich information, they are easy to be interpreted and the pre-processed products are available as open data. They are the first candidates for EO monitoring. However, they suffer from the cloud issues (clouds are always problem of optical remote sensing data) that depend on location and are more disturbing in northern or mountainous areas.

Paying Agencies are generally accessing L2A products (Bottom of Atmosphere); however, dealing with L1C products (Top of Atmosphere) may be envisaged if the PA has image experts. Even for the “standardised” product L2A, there are various options for atmospheric correction according to the way to access images, MAJA being used by Sen4CAP and Sen2Cor being used by ESA.

Paying Agencies often require data more than one year old, i.e. Long Term Archive images that are more difficult to be accessed.

- **Using S-1 images**

S-1 images are powerful but complex products that are more difficult to be interpreted. The ARD (Analysis Ready Data) products are not available as open data from the ESA hub. However, the ESA provides the SNAP tool that can be used to perform the necessary pre-processes. Despite of these difficulties, S-1 images are also widely used by Paying Agencies due to their power to provide useful information, even in case of bad weather conditions.

Opposite to S-2 images, the georeferencing is up to the user. As EO monitoring processes often imply to combine S-1 and S-2 images, the best solution is to georeference S-2 images choosing the same Coordinate Reference System used for S-2 images.

Sentinel-1 data are usually used for 2 main tasks regarding EO monitoring; these are crop classification and agricultural activity detection such as grassland mowing and crop harvest (e.g. maize harvest). Many case studies showed that radar imagery can improve the crop classification accuracy when used in combination with optical data. Further, radar images are proved to be very efficient in detecting mowing events in various agricultural grasslands. There are several methods based on SAR backscattering and interferometric coherence time series that can detect mowing with very high accuracy. These methods are used by several Paying Agencies mainly in northern European countries such as the Netherlands, Germany, Denmark and Estonia.

- **Dealing with parcels with few pixels**

From the questionnaire, the issue occurs not only for small parcels but also for parcels of specific shape (such as narrow parcels). Using HHR images such as Planet and SPOT or even VHR images such as WorldView have been mentioned as potential solutions in the answers to the questionnaire.

It should be noted that in NIVA related events, Planet products have been presented several times, claiming to be adapted for agriculture in general including the EO monitoring. More especially, a “fusion” product has been developed, with the aim to be compatible with Sentinel 2 (with similar calibration of the common spectral bands).

Conditions of commercial offers haven’t been investigated. This investigation would have been difficult as these conditions are not always publicly available.

The access to the HHR or VHR imagery used to deal with these specific parcels should be granted not only to the Paying Agencies (or their sub-contractors) but for transparency reasons also to the concerned farmer in order to enable him/her to understand the rationale behind the traffic lights.

- **Use of Sen4CAP**

Sen4CAP may be installed locally or on a virtual machine with CREODIAS.

Local installation requires resources that users may be reluctant to provide just for testing purposes during a project. For testing purposes, it looks simpler to use rather CREODIAS.

However, the experience of UC1a in NIVA has proved the feasibility of the installation on a local server.

- **No free solution**

The access to Sentinel images is free but in practice, the storage and handling of these big volumes of data require significant IT infrastructure that has to be paid in a way or another.

Getting the information at pixel level means that users will get image or "raster" data and will have to deal with huge volume of data. Until now, this is the only available solution on the market. This solution at pixel level is available according various options as the investment on preprocessing and IT infrastructure may be done in the Paying Agency itself (through buying ICT material and training staff) or it may be outsourced (through buying predefined services and computation power).

Getting the information at feature level implies that the user is invited to provide a layer of features of interest and gets in return an average or median value and possibly standard deviation from the values of the pixel within the feature. This implies a quite smaller data volume to be handled but also less choice in all the preliminary steps. Until now, only the EO-Widget project aims to provide such services but the solution is not yet available on the market.

- **Main points to decide on relevant solution**

It is mainly between being easy (lots of preliminary steps already done before PA accesses the image data) and being flexible (doing things yourself require more expertise but enable to decide on each step of the process). Solution may also depend on what is the most relevant investment; for instance, is it easier to invest in external solutions or on internal capacity building?

Security has also to be taken into account.

2.5.Recommendations

2.5.1. Recommendations for Paying Agencies

- **Get practical experiences**

EO monitoring and the preliminary steps of getting Analysis Ready Data is a complex topic. It is impossible to make definitive decision for whole country without some preliminary experimentation.

Installing Sen4CAP with CREODIAS (and VM) and using it with a few S2 tiles may be a small but easy beginning.

- **Assess your national (or regional) requirements and possibilities**

The first assessment concerns the agricultural landscape: are there many parcels? many small parcels? Is it a cloudy area? The answers to these basic questions will provide a first idea of the type of images to be chosen and of the necessary number of images; this estimated volume of data will help to decide on the IT infrastructure size.

The second assessment is more about what is the best way to invest efforts : it is by using (as much as possible) ready-to-use but paying data and services such as DIAS or other cloud infrastructures or is it by managing images at the Paying Agencies premises, which implies also costs in buying computation power and in staff capacity building.

In practice, the EO monitoring and its preliminary steps may be delegated to a technical partner but the Paying Agency has to drive or at least to be aware of the technical choices.

- **Conduct pilots to evaluate the feasibility and performance of the EO Monitoring approach**

Prepare small scale preliminary studies to familiarise with new technologies and improve on the process before they would be introduced by the new CAP.

- **Take care of quality**

Most of the pre-processes of satellite images imply some choices, with their advantages and limits. These choices should ideally be done based on a comparative analysis, taking into account both the technical quality and the simplicity of use. They should also be well documented (lineage information of the input data)

- **Take care of security**

EO monitoring is dealing with farmer data (whose privacy should be protected) and with payments (which requires data security). The necessity to ensure data privacy and security may be a reason to give preference to home infrastructure or to national cloud.

- **Keep aware of what is going on**

EO monitoring is a fast evolving domain; there are many technical initiatives going-on that may provide new services. Therefore, technologic watchdog is required, e.g. through attendance to conferences or similar events.

- **Invest in capacity building**

Whatever the decision about the way to access and preprocess the satellite images and even if most of the EO monitoring work is done externally, PA should get minimum knowledge about all the issues raised by this disruptive method.

2.5.2. Recommendations for European Commission

- **Provide pre-processed S-1 images**

Until now, Analysis Ready Data from S-2 images is available as open data whereas users have to take care (in a way or another) of the pre-processes of S-1 images.

Some solutions exist: for instance, ESA is offering the SNAP tool to perform these pre-processes or these pre-processes may be ensured by DIAS or other cloud providers. However, this lack of open and pre-processed S-1 data leads to a duplication of efforts (that is likely not the most efficient way to spend public money) and add some difficulty in the adoption of EO monitoring by Paying Agencies.

- **Promote work about quality of EO monitoring**

JRC has published draft guidelines about quality assurance of the EO monitoring: https://marswiki.jrc.ec.europa.eu/wikicap/index.php/CbMQA_TG_v1_1

These guidelines are mainly based on the visual check by human beings of the results of the automatic analysis of Sentinel images. This is for sure a necessary step but the quality of EO monitoring results depend on other factors: the main one is the relevance of the decision rules but the quality of input data (as described in this deliverable) might also impact the reliability of final results. Encouraging some research and/or knowledge exchange to assess this potential impact should be envisaged.

3. Exchanges between IACS and FMIS

This section provides an analysis on how and in what extend the various technologies and digital farming systems that are supporting every day cultivation practices on a farm level can also act as a new source of information for the various objectives that are related with IACS operations.

3.1. Introduction

- **What is an FMIS?**

The concept of "Farm Management Information Systems" (FMIS) is an umbrella term that refers to a set of computer based information systems operating at a farm level which are able to receive data streams, store and process them and provide output useful to the various stakeholders (individual farmers, farmers associations, advisors, etc.). The FMISs aim to keep track of farm activities and to manage the large amount of information generated through the use of ICT solutions. There are currently hundreds of FMISs available on the market that can support decision making by finding the best practices for farm management (Fountas et al. 2015). According to Sørensen et al. (2010) the FMIS is defined as: "*a planned system for the collecting, processing, storing and disseminating of data in the form of information needed to carry out the operations functions of the farm*." Examples of commercial FMISs with an already significant user based are: Agworld[1] , Dacom[2], Farmnet365[3], Akkerweb[4] and Gaiasense[5] . Each FMIS may focus on one or multiple domains of the agricultural sector, for example, livestock or arable farming. In general, a general purpose computer based Information System aims to support decision making by providing timely information about the planning, control and operational functions of an organization (Watson et al. 1991). In a similar manner, the FMIS does the same for the agricultural domain usually aiming to reduce the production costs, maintain high quality and to comply with the agricultural standards. In a nutshell, a FMIS can be considered as a typical ERP[6] system that can manage/share the following type of information (Fountas et al. 2015):

Function title	**Function description**
Field operations management	Includes the recording of farm activities. It can also be considered as a digital calendar where applied practices and the overall status of the cultivation are recorded. This function also helps the farmer to optimize crop production by planning future activities

[1] https://www.agworld.com/us/
[2] https://www.dacom.nl/en/
[3] https://www.365farmnet.com/en/
[4] https://akkerweb.eu/en-gb/
[5] https://www.gaiasense.gr/en/gaiasense
[6] ERP (Enterprise Resource Planning) - refers to business management software, typically a suite of integrated applications that an organization can use to collect, store, manage, and interpret data from many business activities

	and observing the actual execution of planned tasks. Furthermore, preventive measures may be initiated based on the monitored data.
Best practice (including yield estimation)	Includes production tasks and methods related to applying best practices according to agricultural standards (e.g. organic standards, integrated crop management requirements). A yield estimate is feasible through the comparison of actual demands and alternative possibilities, given hypothetical scenarios of best practices.
Finance	Includes the estimation of the cost of every farm activity, input–outputs calculations, labour requirements, and so on, per unit area. Projected and actual costs are also compared and input into the final evaluation of the farm's economic viability.
Inventory of materials	Includes the monitoring and management of all production materials, equipment, chemicals, fertilizers, and seeding and planting materials. The quantities are adjusted according to the farmer's plans and customer orders. A traceability record is also an important feature of this function
Traceability	Includes crop recall, using an ID labeling system to control the produce of each production section. Traceability records related to the use of materials, employees, and equipment can be easily archived for rapid recall.
Reporting	Generally includes the creation of farming reports, such as planning and management, work progress, work sheets and instructions, orders purchases cost reporting, and plant information.
Site specific	Includes the mapping of the features of the field. The analysis of the collected data can be used as a guide for applying inputs with variable rates. The goal of this function is to reduce or optimize input and increase output
Sales	Includes the management of orders, the packing management and accounting systems, and the transfer of expenses between enterprises, charges for services, and the costing system for labour, supplies, and equipment charge-outs
Machinery Management	Includes the details of equipment usage, the average cost per work-hour or per unit area. It also includes fleet management and logistics.
Human resource	Includes employee management, including, for example, the

management	availability of employees in time and space. The goal is the rapid, structured handling of issues concerning employees, such as work times, payment, qualifications, training, performance, and expertise
Quality assurance	Includes process monitoring and the production evaluation according to current legislative standards
Decision support on applied cultivation practices.	This functionality goes a step further that simple recording. It refers to the provision of recommendations on cultivation practices such as fertilization, pest management, irrigation based on various parameters such as environmental recordings, cultivation type and scientific algorithms. The extraction of recommendations can be supported by advanced data processing algorithms based on machine learning and/or artificial intelligence approaches.

Table 11. Overview of functionalities offered by FMISs

- **Evolution of FMIS towards smart farming**

During the recent years Information and Communication Technologies (ICT) are continuously being introduced within everyday farming practices on an increased pace. Several cutting-edge technologies in agricultural production, from GPS and remote sensing to big data, artificial intelligence and machine learning, robotics, and the Internet of Things (IoT), are gradually integrated in support of every day farming activities aiming to optimize inputs, lower costs, and reduced environmental impact. All these technologies that are described under the umbrella term of “smart farming” are aiming to support end-users (e.g. farmers, advisors) to benefit from data that can be gathered at the farm-level. The core objective of smart farming technologies is to collect, classify, and analyse vast amounts of data to detect patterns and support decision making. According to (Balafoutis et al., 2017) a conceptual categorisation of the overall process is: i) data acquisition, ii) data analysis and evaluation and iii) precision applications. The FMISs are the core entities with which all these technologies are integrated in order information to be collected at a single point and hence to be further processed and rendered to the end-user in order to be utilised for advanced decision making.

- **Potential use of FMIS for CAP monitoring**

Within the context of the CAP monitoring and evaluation framework the introduced technologies provide a potential new source of information. The fact that with the help of digital technologies it is feasible to monitor new data items and create new information streams allows the introduction of new CAP indicators[7] but also the more efficient monitoring of existing ones.

[7] For more information on this topic refer to H2020 CSA “Monitoring and Evaluation Frameworks for the Common Agricultural Policy (CAP) Mef4CAP” project: https://mef4cap.eu/

Although the primary functional objective of FMISs is to support farming activities they are also a valuable source of information for the needs of CAP monitoring. In this deliverable we evaluate what kind of information is feasible to be retrieved and how.

3.2. Field data requirements

In this section the main information items related with agricultural practices and applied inputs are identified that are considered as useful in the context of CAP processes.

Given the functional characteristics of the FMISs a list of indicative information items that are required (or are of interest) by IACS follows:

- **Information on agricultural land use**: The parcel's area (e.g. in hectares) and location (e.g. polygon coordinates) is among the core information entities that are useful for IACS. In this case the use of satellite technologies and LPIS are particular useful and have been analysed in the previous sections.
- **Agricultural Inputs** : The type, amount and time of applied inputs (e.g. pesticides, fertilisers, irrigation) at a parcel is among the most significant information items for IACS; it is expected to be required data for the CAP post-2020.
- **Crop type and yield**: The actual type of crop cultivated for a specific time period and the harvested yield is of interest by IACS.
- **Applied agricultural practices – Planting, Harvesting, Mowing, Ploughing, etc.:** This category includes the type of applied practice and the respective time period that are applied. The use of machinery data is also included within this category.
- **Organic cultivation practices:** This information item refers to wether a cultivation is treated with a manner approved for organic agricultural products.
- **Livestock - Herd management**: Total number of animals, type of animals, annual births/deaths, medicines utilised, animal feed utilized, etc.
- **Livestock Pasture management**: Conditions of pastures.
- **Financial Inputs/Outputs**: This refers to financial related information items (cost, amounts, etc.) related with purchased agricultural products (fuels, chemicals, seeds, equipment etc.) and their respective consumption/use. Also the potential income from selling the production is of interest.

The presented list is only indicative on the primary information items that are handled by agricultural information systems and that may be of interest for evaluating current and future CAP targets/indicators. It is obvious that this list is not exhaustive given the plethora of existing and upcoming ICT technologies for agricultural systems.

On a similar manner IACS are also able to provide useful information to farmers through FMIS:

- IACS through the Land Parcel Identification System can provide to FMIS data on parcels geometries and unique identifiers through user friendly means of rendering.
- Information on registered animals in the context of livestock management.

- Given that IACS act as a centralized repository of agricultural data it is feasible to provide information on a regional bases on the following:
  - Aggregates on pesticides/fertilisers use for the area that the parcel is located.
  - Pest infestation early warnings for the area that the parcel is located along with recommendations for pest management related actions.
  - Carbon footprint performance for the area that the parcel is located.
  - Soil quality and soil erosion for the area that the parcel is located.
  - Other agricultural statistics (e.g. harvested yields, harvest dates, crop types) for the area that the parcel is located.

This list is not exhaustive and it might be expanded based on the respective integration of additional information sources. The following sections elaborate on the technical details of the various current and potential information sources.

3.3. Overview of FMIS data and data sources

In this section an overview of data sources are presented that are currently utilized by FMIS but can potential also be utilized for gathering ground truth evidences in the context of CAP monitoring and evaluation.

3.3.1. Farmers digital calendar

The “farmer’s calendar” is the place where all agricultural practices are recorded. Farmers usually maintain their own calendar even without the help of digital tools. Nowadays the farmer’s calendar is usually maintained as a simple spreads sheet or as part of an ICT system with a wider scope like an FMIS. A digital farm calendar contains all the activities that are taking place at the farm. In tables below, indicative example snapshots of farm calendar entries are presented focusing on different agricultural activities (e.g. cultivation growth stages, fertilization inputs, pesticides inputs, irrigation).

Start Date	End Date	Growth-Stage
20/4/2018	21/4/2018	Planting of tuber seed or "potato seed"
24/5/2018	27/5/2018	Stems growing towards soil surface, formation of scale leaves in the axils of which stolons will develop later
1/6/2018	15/6/2018	Emergence: stems break through soil surface
16/6/2018	20/6/2018	4th-6th basal side shoot visible (> 5 cm)
20/6/2018	21/6/2018	Crop cover complete: about 90% of plants meet between rows / First individual buds (1–2 mm) of first inflorescence visible (main stem) / Tuber initiation: swelling of first stolon tips to twice the diameter of subtending stolon

21/6/2018	1/7/2018	First flower petals of first inflorescence visible / First open flowers in population

Table 12. Example of recorded growth stages for a potato cultivation

Start Date	End Date	Type	Application method	Commercial Name	Dose	Unit
17/3/2018	17/3/2018	basal fertilisation	Broadcasting	Macro Speed Gren 25%MgO + 50%SO3	266	kg/hectare
4/4/2018	4/4/2018	basal fertilisation	Broadcasting	Wapno Nordkalk (Organic manure)	2600	kg/hectare
18/4/2018	21/4/2018	basal fertilisation	Broadcasting	RSM (26%)	558	kg/hectare

Table 13. Example of recorded fertilisation inputs

Start Date	End Date	Commercial Name	Active Substance	Dose	Unit
18/4/2018	21/4/2018	MONCUT 46 SC	flutolanil 460g/lt	513	mL
21/5/2018	21/5/2018	Boa 360 CS	clomazone 360 g/l	0.16	L
28/5/2018	28/5/2018	Bandur 600 SC	aclonifen 600 g/l	2	L
15/6/2018	15/6/2018	Citation 70 WG	metribuzin 700 g/kg	0.45	kg
10/7/2018	10/7/2018	Stomp 400 SC	pendimethalin 400 g/l	4	L

Table 14. Example of pesticides applications calendar

Starting Date - Time	Ending Date- Time	Water quantity estimation		
29/07/2020	29/07/2020	6	mm	hectare
30/07/2020	30/07/2020	15.6	mm	hectare

4/08/2020	4/08/2020	9	mm	hectare
11/08/2020	11/08/2020	8.1	mm	hectare
21/09/2029	22/09/2020	12	mm	hectare

Table 15. Example of irrigation entries at a farmer's calendar

3.3.2. Photograph applications

The use of photographs and/or video in order to remotely monitor cultivation's status is an important tool that is gaining more ground during the recent years. As it is analysed in Colwell et al. (2021) even from 1960 aerial photos were utilized in order to identify in a field-by-filed manner how vigorous a cultivation is, identity of crop-damaging agents, and probable crop yield. In our days, many FMISs have already introduced the use of images in their operational mode. For example "PESSL Instrument[8]" offers an environmental monitoring station which is integrated with cameras able to capture images of the cultivation (Figure 30).

Figure 33. CropView by PESSL

Based on the images taken it is feasible to deduce the phenological growth status of the cultivation, potential yield and potential pest infestations.

[8] https://metos.at/cropview/

Figure 34. Remote fruit monitoring – AI system for fruit measurement (Pessl)

Figure 35. Remote field monitoring – check growth of your crop (Pessl)

Figure 36. Remote crop monitoring – check phenophase of crop (Pessl)

As it is indicated in Cerro et al (2021), cameras installed on aircrafts (including UAVs) performing missions over agricultural fields allow the extraction of various information items. The respective remote sensing sensors can be mainly clustered in several sets: RGB (visible spectrum) sensors, multi-spectral, and hyper spectral cameras. The main difference between multispectral and hyper spectral imaging is the number of wavebands being imaged and how narrow the bands are. Multispectral imagery generally refers to 3 to 10 discrete “broader” bands, whereas hyper spectral imagery relies on much narrower bands. Based on the state of the art review presented in Cerro et al (2021), the following information, among others, can be extracted with the use of photographs:

- Nutrients Evaluation and Health Assessment
- Water Stress Analysis
- Yield and Biomass Estimate
- Soil Monitoring
- Weeds Detection
- Environmental Monitoring

Beside cameras integrated to fixed stations or at UAVs, there is available a plethora of photo applications for smart mobile phones. As it is reported in the state of the art review by Mendes et. al (2020) many of these applications are focusing in identifying diseases based on image analytics of leaf photos. The techniques utilized for disease recognition can be performed in a standalone mode, combined with remote data (cloud) or portable molecular analysis equipment.

In order the disease estimation through image analytics to be feasible there is a need for trained algorithms usually based on Machine Learning or Artificial Intelligence techniques. Such an effort is currently conducted by “Eden Library[9]” an initiative started from Agricultural University of Athens; end evolved to a multidisciplinary team of AI experts, agronomists, and software engineers. The “Eden Library” offers collections of high value plant datasets embedding agricultural domain knowledge produced in an academic environment. As it is stated, the goal is to use AI capacity in agriculture for food safety and quality control, while adding value to the dataset users. The offered datasets of images are exhaustively annotated and integrate quality control frameworks in order to produce i) problem-specific datasets ii) rich in metadata, iii) clustered based on multiple criteria and iv) standardized in ways that facilitate researchers, tech industry and agrifood business experts to use them in various contexts based on their needs.

As a final note are photo apps through mobile devices that are integrating security mechanisms in order to ensure the actual location (coordinates including altitude in some cases) and the direction of the device. These technologies are classified under the term “geotagged photos” and each captured photograph file is escorted by a set of metadata ensuring the actual location of the mobile phone, the direction of the device and the time when the photo was taken encoded as parameters according to Exchangeable Image File Data (EXIF) standard. In some cases, these photos are integrated within the overall data records referring to a specific parcel that are maintained by a FMIS. With these photos it is feasible for the farmer advisor to have a more complete overview of the current status of a cultivation, inferring for example the phenological growth stage of the crop, and avoiding an actual

[9] https://edenlibrary.ai/

visit to the field. Table below provides an example of the actual metadata (EXIF) that are feasible to be recorded by a geotagged photo application for mobile phones.

Photo	Location (coordinates)	Gyroscopic	Parcel Id	Date & Time	Device type
	25.269230615119, 40.962504122927	x:0.7733285427093506, y:0.02837530151009559, z:-0.3877149224281311	12345	15/6/2020 10:37:16 AM	Samsung
	25.269230615119, 40.962504122927	x:0.7733285427093506, y:0.02837530151009559, z:-0.3877149224281311	12345	17/7/2020 10:00:16 AM	Samsung
	25.269230615119, 40.962504122927	x:0.7733285427093506, y:0.02837530151009559, z:-0.3877149224281311	12345	17/8/2020 10:15:16 AM	Samsung

Table 16. Examples of crop's geotagged photos

3.3.3. Financial information

Various FMISs are already supporting the management of invoice documents in an integrated manner with the rest of the available agricultural related information. The invoices of purchased products utilised as inputs at a farm level consist a valuable source of information within the context of CAP monitoring. For example the invoice issued during the purchase of pesticides can act as escorting evidence on the recorded action of pesticide application at the "Field Book" of a FMIS. The same may apply for purchased fertilisers and recorded actions of fertilisation.

In some cases it is necessary to utilise Optical Character Recognition[10] (OCR) technologies in order to scan, digitise and store paper invoices within the FMIS.

Recent years there are on-going efforts on MS level to harmonise the issuing process of invoices through the introduction of digital processes. To this end, the European Standard for eInvoicing through the Directive 2014/55/EU is introduced in a response to the many eInvoice formats used across the EU. The rational is to move towards electronic invoicing mechanisms that comply with the European norm. Such an approach will make easier the integration of financial related data within the FMISs and their respective processing from various systems.

[10] https://en.wikipedia.org/wiki/Optical_character_recognition

3.3.4. In-field sensors

There are various state of the art reviews (Brewster et al., 2017) that are extensively analyzing a plethora of field sensing technologies that are able to monitor the conditions around the areas that are deployed. With regards to CAP performance monitoring, individual field sensors may not be directly considered as a particularly useful source of information given that the individual recordings need further interpretation. However, usually these sensors are part of greater precision farming system or FMIS and provide the necessary input in order to proceed with decision making that will guide the applied agricultural practices. In some cases sensor recordings can be considered as additional evidences of recorded farm practices. For example, sensed alterations in soil moisture can be considered as a ground truth evidence that escort a recorded irrigation event on the farmers field book also allowing to infer the actual amount of water applied. This is the case for example in next figure where an increase of soil moisture in different depths is detected after an irrigation event (indicated as blue water drop in graph).

The following list contains a set of sensor technologies that are directly or indirectly related with CAP monitoring objectives:

- Soil moisture sensors
- Soil salinity/conductivity sensors
- Camera sensors, ultra violet, multispectral sensor
- Liquid level sensors
- Temperature and gas monitoring of silo/granary
- Gas (e.g. CO2, Ammonia, Oxygen) monitoring sensors
- Active Substance (Pesticide) Sensors
- Digital (camera-equipped) insect pest traps
- Machinery data

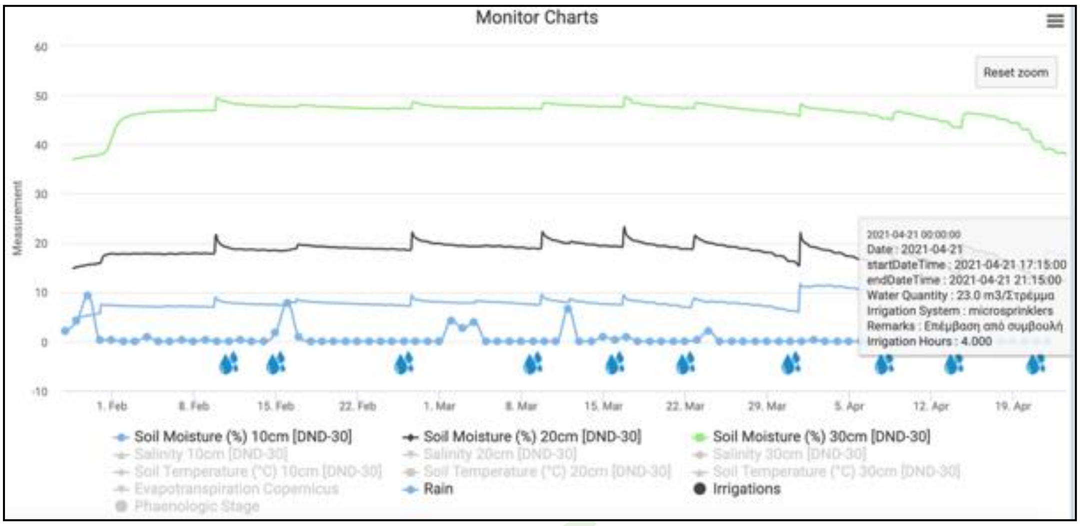


Figure 37. Soil moisture measurements can be utilized as evidences of irrigation events.

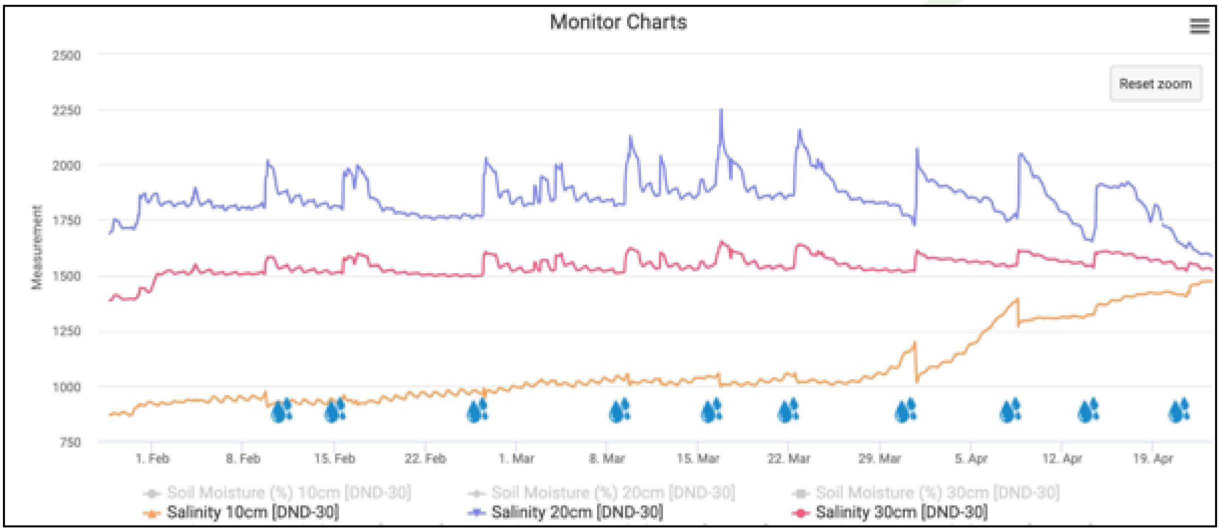


Figure 38. Variations of soil salinity measurements may be associated with fertilisers diffusion and irrigation.

3.3.5. Evolution of FMIS data

Until now, the use of FMIS in Europe is not accurately known as there are no good census held in Europe, but only anecdotical and small surveys results.

The situation looks quite heterogeneous, as one can state that in northern countries, most of arable land is cultivated with some kind of FMIS whereas in other countries in south or central Europe, it is much lower – although it varies strongly between regions. The situation varies also regarding the type of agriculture, precision farming being generally more spread for arable land than for grassland.

More detailed information about adoption of precision agriculture may be found in following papers:

https://www.researchgate.net/publication/330049521_Exploring_the_adoption_of_precision_agricultural_technologies_A_cross_regional_study_of_EU_farmers/citations?latestCitations=PB%3A342233870

https://acsess.onlinelibrary.wiley.com/doi/full/10.2134/agronj2018.12.0779

However, the situation might evolve in Europe as for some types of agricultural inputs, record keeping is already mandatory in EU. This is the case for the applied pesticides where based on Directive 2009/128/EC every Member State (MS) must provide support to farmers in order to apply best practices on pesticides management and utilisation. All MS shall therefore collect data on usage of pesticides for monitoring and statistics extraction purposes. MS must comply with this directive through national plans. For example in Italy it is mandatory for farmers to maintain a “Field Book” where the respective details of pesticides applications are recorded as a way to collect the data required. In addition, the Italian Ministry of Agriculture, Food and Forestry has published technical/operational guidelines[11] with logbook forms templates, data to be collected and digital data exchange formats for digital logbooks.

In another example, Greece implemented the respective directive on 21.09.2020 where it is mandatory the agricultural stores to keep digital evidences when selling or purchasing pesticides. This means that the ERP system of the agricultural store is connected on-line with the respective Information System of the Ministry of Agriculture in order to assist the process of pesticides tracking. Farmers are obliged to keep in a specific form of the "Field Book" the interventions with pesticides per plot, crop, disease / enemy, dosage, etc. In addition, among the next steps is the wider use of digital subscription scheme on a national level for pesticides trading where pesticides will be sold only after the issuing of the respective recommendation by authorised agronomists. Each subscription will be maintained as digital evidence for a long time period and it will include among others the following information items: active substance, cause, crop, plot, etc.

Estonia is also mandating a Field Book to record a wide range of data that may impact water quality, such as data about crops and agricultural practices, use of fertilizers and plant protection products, manure stacks, livestock grazing, drainage system, results of soil analysis, yields; until now, this recording may be done using digital means but also just using paper registers.

Similar policies and tools are currently under development for the collection of data on utilized irrigation water and fertilisers (e.g. FaST platform[12]).

In summary, farmers are submitted to more and more legal incentives pushing them to record data of interest and the general trend is to encourage this recording through digital tools.

3.4. FMIS and data exchange

In the previous sections the basic operating principles and the respective information items maintained by FMISs and other agricultural digital technologies were presented. One of the main issues hindering the sharing of agricultural data items is the fact that these systems have been developed in parallel without considering among their design principles the exchange of information with remote/third party systems. There is a lack of interoperable-by-design functional features on current FMISs which is also the case for the IoT systems landscape in general. (Liu et al., 2015). In this section, we elaborate on FMIS interoperability by initially introducing the basic terms and concepts, then the various approaches for harmonising agricultural data on a semantics level and finally mechanisms and best practices on applying interoperability mechanisms.

3.4.1. Definition of Interoperability

Currently there are various FMISs and other agricultural data management platforms using diverse data models and APIs resulting in a highly fragmented ecosystem. This situation makes particular

[11] https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/3296
[12] https://fastplatform.eu/

difficult the sharing of data from FMIS with various national IACS which are also utilise diverse data models and formats. Figure 39 illustrates the current status.

Data interoperability is the ability of a data set to be reused by any system without special effort. There are different layers to data interoperability: data can be technically interoperable thanks to a machine-readable format (e.g. CSV) and an easy-to-parse structure (e.g. JSON) but, in order for an external system to perform more operations on a dataset, the 'meaning' of data in the structure has to be explicit, and this is achieved through semantic interoperability by using metadata and values that have been previously assigned an unambiguous meaning and an identifier that can be used across systems. (FAO, 2021)

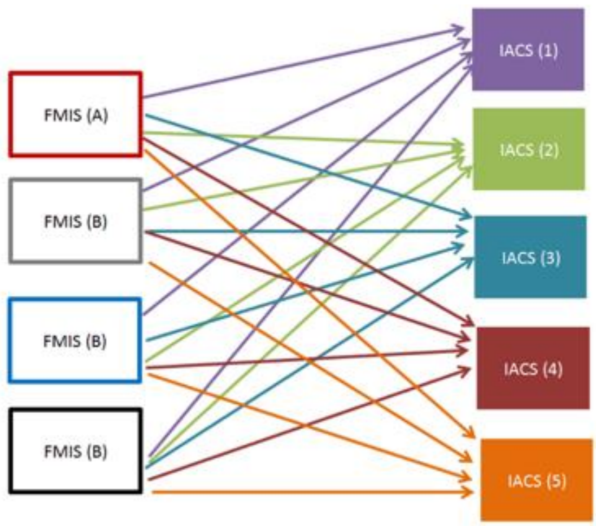


Figure 39. Non-interoperable data exchange.

Following the definitions provided by Serrano et. al, (2015) the different aspects of interoperability on computer systems can be defined as follows:

- Technical Interoperability: usually associated with communication protocols and the infrastructure needed for those protocols to operate.
- Syntactic Interoperability: usually associated with data formats and encodings, e.g., XML, JSON and RDF.
- Semantic Interoperability: associated with a common understanding of the underlying meaning of the exchanged content (information).
- Organisational Interoperability: associated with the ability of organisations to effectively communicate and transfer information even across different information systems, infrastructures or geographic regions and cultures.

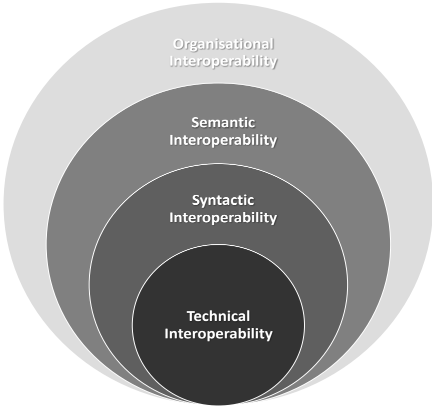


Figure 40. Different levels of interoperability (Serrano et. al, 2019)

3.4.2. Farm observation data standards

In FAO's report (FAO, 2021) on standards for farm observations the following categories of data standards are presented according to the information entities that are focusing on:

• **Crop basic data** (from germplasm descriptors to official names to product classifications) have been standardised by normative bodies (e.g. FAO, International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRA). The European Food Safety Authority (EFSA), the United States Department of Agriculture (USDA)) and their core properties have been modeled in ontologies by research institutions (e.g. Consultative Group on International Agricultural Research (CGIAR), and the Institut national de la recherche agronomique (INRA)).

• Some data standards for **crop growth** data and **crop growth models** have been developed by research institutions that wanted to share or reuse models (e.g. the Agricultural Model Intercomparison and Improvement Project (AgMIP) and the Global Agricultural Trial Repository and Database (AgTrials)).

• Data standards for **soil observations, soil profiles and soil properties** (chemical properties, physical properties) exist (from USDA and FAO classifications to Infrastructure for Spatial Data in Europe (INSPIRE) data specifications).

• **Weather data standards**, have been created by meteorology agencies.

However, other data used in FMIS (data about machinery, sensors, agricultural input like fertilisers and pesticides, in some cases sales management data) partly follow industry standards and partly are just encoded in closed proprietary formats. This is a crucial gap given the objective of modeling FMIS data with uniform manner in order to share with IACS.

An approach to resolve the issue of interoperability among FMIS systems and among FMIS and IACS might be realized through the use of ontologies and semantic technologies (Drury et al, 2019).

Data model	Domain	More details
UN-eCrop	Data exchange for FMIS	https://unece.org/fileadmin/DAM/cefact/brs/BRS_eCROP_v1.pdf
ISOBUS	Farm machinery	https://www.aef-online.org/about-us/isobus.html
ADAPT -	Data exchange for FMIS	https://adaptframework.org/
ICAR - International Committee for Animal Recording	Livestock	https://www.icar.org/
GS1- EPCIS	Supply chain	https://www.gs1.org/standards/epcis
ETSI- SAREF4AGRI	Agricultural IoT devices	https://saref.etsi.org/saref4agri/v1.1.2/
ETSI-NGSI-LD	Context information for sensors and other IoT devices.	https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.01.01_60/gs_CIM009v010101p.pdf
OGC - Agriculture Domain Working Group	Geospatial data and geographic analysis focusing on agriculture	https://www.ogc.org/projects/groups/agriculturedwg

Table 17. List of standardised data models focusing on different agricultural domains.

Table 17 provides a list of dominant data modeling approaches referring to various agricultural domains. These data models provide a potential solution to semantic interoperability however there are still many unresolved challenges. For example, there are parallel standards for the same application domain that are not harmonized introducing heterogeneities on semantic level even though they have been standardised. In other cases, standardised data model might not be able to model every aspect of agricultural concepts, especially given the introduction of new information items that need to be modeled due to adaptation of new sensing approaches/devices.

3.4.3. Enabling Interoperability on Digital Agricultural systems

As it was presented in the previous section standardisation of agricultural data models can significantly contribute in resolving the issue of semantic interoperability. But there are still issues on how to apply these harmonised data models especially for already deployed operational FMIS serving farmers in 24/7 mode. It is not feasible to expect existing agricultural systems to change drastically their internal mode of operation – potentially causing sever disruptions- in order to achieve interoperability with remote systems (Kalatzis et. al, 2019).

A current best practice in order to enable interoperability is considered the use of lightweight software modules that act as translators from custom modelled datasets to a standardised data model. (Brewster et. al, 2018). These translators are also called (inbound-outbound) interoperability enablers and are integrated in a seamless manner with existing systems (e.g. FMIS). A set of design principles for interoperability enablers are defined in Kalatzis et. al (2019) and Freire et. al (2019) and the overall objective is not to affect or to affect as minimal as possible the existing systems. As it is stated, it is the interoperability enabling solution that needs to be versatile and easy to adapt to the underlying system. To this end, a Plug-and-Play and light-weight (as possible) software design is necessary for seamless integration with existing systems.

Recent years there are significant on-going efforts on enabling interoperability for agricultural systems. On 2017 the "Internet of Fruits and Farms (IoF2020)[13]" large scale pilot was initiated. The IoF Innovation Actions was funded with 30M euros and on the next 4 years more than 25 use cases where realised in almost all EU countries making IoF2020 one of the largest piloting programs in smart agriculture. Among the core objectives of this project was the establishment of system interoperability among agricultural systems aiming to realise a "system of systems" approach. The work presented in Public Deliverable "Opportunities and barriers in the present regulatory situation for system development[14]" provides a layered architectural approach (figure 38) modelling the mode of operation of various digital agricultural systems along with the Interoperability Points (IOP) where data sharing through the use of minimum interoperability mechanisms (data mappers/translators) can be applied. It should be noted that within IoF2020 the CEF Orion Context Broker[15] was utilised in combination with SmartAgriFood[16] data models in order to enable interoperability in semantic and syntactic level.

The effort on establishing interoperability mechanisms for agricultural systems in EU level is on-going through the H2020 DEMETER[17] and ATLAS[18] projects. As it is stated "*The H2020 DEMETER project is a large-scale deployment of farmer-driven, interoperable smart farming-IoT (Internet of Things) based platforms, delivered through a series of 20 pilots across 18 countries (15 EU countries).*" On a similar manner "*The goal of ATLAS is the development of an open interoperability network for agricultural applications and to build up a sustainable ecosystem for innovative data-driven agriculture*". Although these two projects are aiming on establishing interoperability mechanisms among existing systems (including FMISs) the overall effort is of interest also from the perspective of the IACS.

[13] https://www.iof2020.eu/

[14] https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5bb2a8e03&appId=PPGMS

[15] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/Orion+Context+Broker

[16] https://github.com/smart-data-models/SmartAgrifood

[17] https://h2020-demeter.eu/

[18] https://www.atlas-h2020.eu/

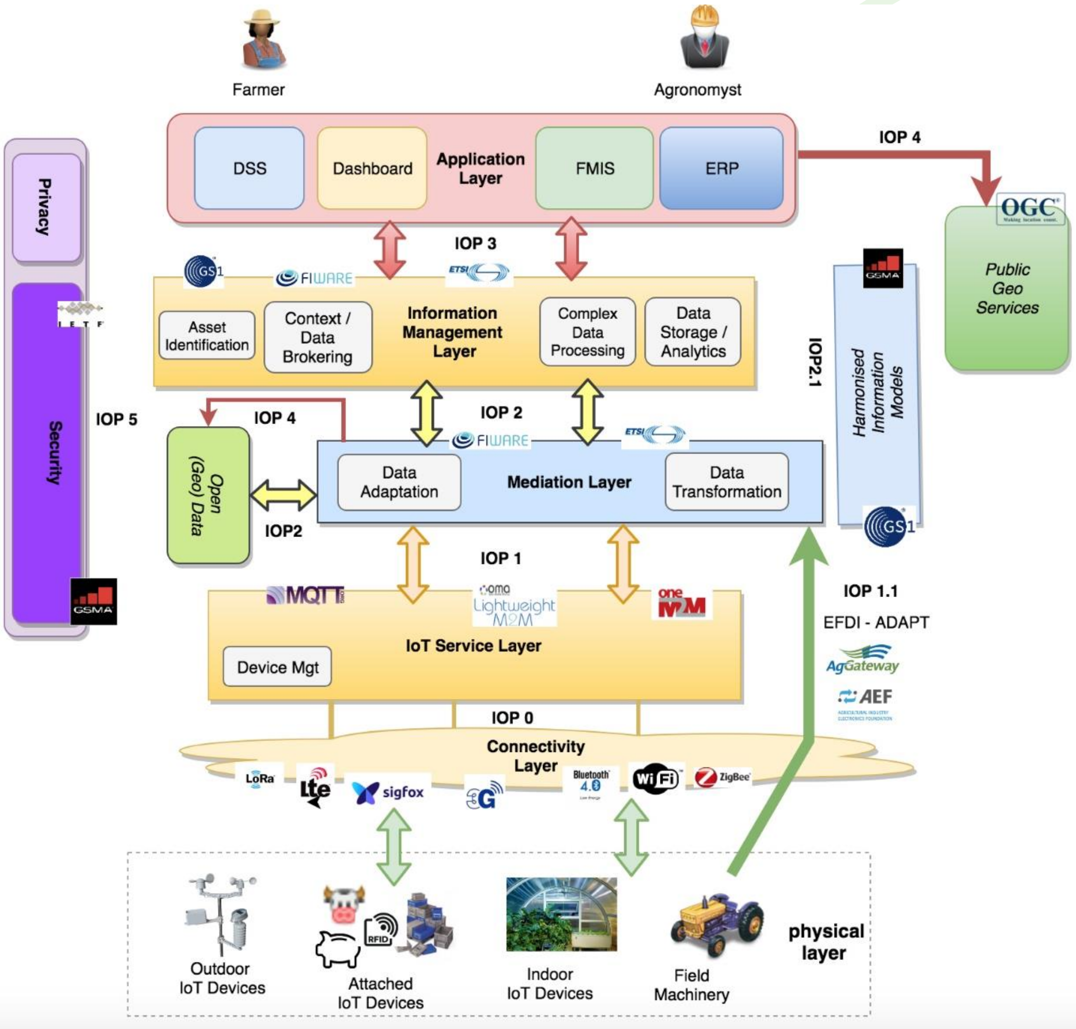


Figure 41. The IoF2020 approach on enabling interoperability

3.5. Experiences from NIVA Use Cases

3.5.1. Use Case 1c: Farmer Performance

3.5.1.1 Use Case objective

In order to design effective policy measures, data is needed to evaluate farmers' impact on environment, climate and sustainability (farmer performance). Valuable source of such data is IACS. However, although IACS already contains a lot of data, there are still data gaps which must be filled by getting additional data about farming activities from other sources.

Valuable source of information about farming activities is FMIS, type of (commercial or non-commercial) software used by farmers to manage farm data. By exchanging bi-directionally data between IACS and FMIS-type of applications, agricultural activities data collected in a farm will become an additional input for monitoring farmer performance.

Data exchange between IACS and FMIS-type of applications provides an opportunity to reduce administrative burden for farmers (data already existing in the FMIS can be shared with IACS and vi*ce versa*), also PA can use such exchange of data to support farmers' compliance with regulatory requirements.

Entering data manually and/or manual import/export of data files is time consuming, risky and therefore automatic system to system exchange of data is desirable solution (less administrative burden).

Objective of the NIVA use case UC1c is to enhance assessment of farmer performance in the context of CAP post-2020, combining IACS and FMIS data.

3.5.1.2 Context of experience

As a first step, prototype technical solution of bi-directional exchange of data between IACS and FMIS had to be established. UC1c design and development is led by ARIB (Estonian Agricultural Registers and Information Board), i.e. the Estonian PA.

An Estonian private company, eAgronom https://eagronom.com/en/ offering FMIS software was involved to the process on a voluntary basis.

The main characteristics of eAgronom are the following:

- Target farmers are grain producers;
- Approximate number of customers – over 1500 paying clients, managing around 1 000 000 hectares of land under cereal production;
- FMIS is used in Estonia, Latvia, Lithuania, Poland, Czech Republic, Romania and Australia;
- Main aims and functionalities of the FMIS:
    - Agronomical planning – crop rotation, inventory, task maps, and financial analysis;
    - Personnel management – tracking working hours, giving work orders;
    - Artificial Intelligence based alerts and suggestions – AI crop planner, business appraisals, task timing;
    - Integrated soil and air sensors;
    - Normalized Difference Vegetation Index (NDVI) calculation;
    - Consultation – agribusiness consultancy services.
- Main technical characteristics of the FMIS:
    - Software as a service (SaaS) platform;
    - desktop and mobile application (both iOS and Android);
    - offline mode (for Android only).

The dataset concerned is mostly GSAA data (source – IACS), data about agricultural activities performed on a parcel, fertiliser and plant protection product use data (source – FMIS).

3.5.1.3 Technical solution

Current version of UC1c prototype is a micro service for exchanging data between IACS and FMIS type of software. Data exchange protocol is based on the standard proposed by the UN/CEFACT for electronic exchange of crop cultivation data along the supply chain - eCROP (with some modifications).

The main features of the product are the following:

- Converting IACS data into eCROP based format;
- Making IACS data available to FMIS by returning the data as a response to a request sent by FMIS;
- Validating FMIS data compliance with eCROP based format;
- Accepting compliant data (at this moment prototype solution will not actually save and store data);
- Rejecting non-compliant data.

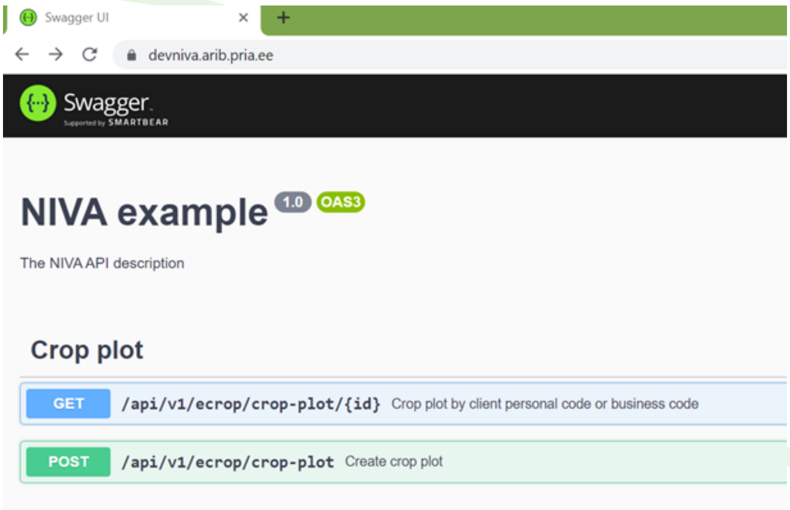


Figure 42 UC1c IACS-FMIS data exchange API deployment in ARIB development environment

Technical description:

- REST API with 2 endpoints:
  - GET /api/v1/ecrop/crop-plot (FMIS can request data about agricultural parcels from IACS);
  - POST /api/v1/ecrop/crop-plot (FMIS can send data about agricultural parcels to IACS).
- Node.js and Typescript application;
- run in Docker;
- User Interface – Swagger.

Source code is available in NIVA GitLab https://gitlab.com/nivaeu/uc1c-public-api

The development of UC1c prototype is on going, with additional components being added to REST API – database, calculation library and graphic user interface.

3.5.1.4 Results and lessons learnt

Conclusions and remarks from the Single Member State testing phase in Estonia:

- Standardized data message structure is of vital importance (there are many FMIS providers, also IACS systems differ). In general, the eCROP standard was found to be suitable for standardized data exchange messages between Paying Agency's IACS system and FMIS. However, some modifications to the eCROP standard were necessary due to specific characteristics of the IACS dataset (some data is collected in payment application process which does not have exact match in eCROP).
- Some complications in establishing system-to-system exchange of data between IACS and FMIS arise because of different purposes of these information systems. IACS is a tool for Paying Agency and it is oriented for claim, control and payment process. FMIS is a tool for farmer and is oriented for farm management; farmers' main purpose of using FMIS is to better manage their farming activities. Some key concepts of IACS are not so important in the context of FMIS (e.g. reference parcel, land cover type), but the IACS data model is strongly based on these.
- Suitable code lists for pan-European use have to be agreed. Currently each Paying Agency needs to use their own code lists for crop types. For agricultural activities, fertilizers, plant protection products, the Estonian Paying Agency does not have code lists at all. For real-life exchange of data Paying Agency should have a code list service available, enabling all FMIS software providers who want to exchange data, to adopt these code lists.
- FMIS data in IACS infrastructure – current IACS data model does not include data potentially to be received from FMIS (agricultural activities, fertilizer use etc).
- The development tool was done only one-way (from IACS to FMIS); for the opposite side, the tool has to be developed by the FMIS editor. In the Estonian experience, unfortunately, eAgronom did not have the necessary resources to develop the POST direction data exchange module
- Legislation – current legislation does not explicitly address sharing FMIS data with IACS.
- User acceptance. Farmers may be reluctant to share data unless clear assurance that this data exchange is used for their benefit only and not for penalties (e.g. PA supporting them to be compliant with CAP requirements, without any penalties arising from FMIS data shared with PA). FMIS software providers will be interested in establishing technical readiness for data exchange only when their customers (farmers) request such functionality.
- Deployment of the IACS-FMIS data exchange application in real life conditions requires that acceptable and compatible authentication/authorisation system is in place at both sides (IACS and FMIS).

Prototype has been handed over for Multi Member State testing in Italy; no testing results are available yet.

3.5.2. Use Case 3: Farm Registry

3.5.2.1 Use Case objective

The new IACS, as a management system for CAP post 2020 applications, needs the information of all the farmers updated in real-time in order to have guarantees of having correct data at the time of the corresponding aid calculation. Thus, direct relations with other systems that can provide updated information are needed. This information can be provided by the farmer or other agencies or entities that ensure the reliability of the data. So that, the IACS should be able to receive information from systems such as FMIS for these controls.

Moreover, it is also interesting that FMIS can receive the information of the surfaces for each farmer. UC3 proposes to receive this information from the Farm Registry, which farmers can update throughout the year. UC3 has developed a Farm Registry data model and its associated database structure.

The data to be exchanged will be those related to surfaces, such as parcel identification, product, rain-fed or irrigated system, geometry and surface.

3.5.2.2 Context of experience

UC4b (machine data) as a tester of UC3, has carried out different tests to ensure its proper functioning as well as a successful integration between UCs. UC4b team has provided information directly from the concerned FMIS, so UC3 has studied how to reconcile it in the UC3 data model to have all the details about how does the farmer produce integrated in the Farm Registry.

Nevertheless, specific adaptations have been noticed when other UC wants to export some specific data to the Farm Registry due to different work units. The picture below shows the different adaptations mentioned:

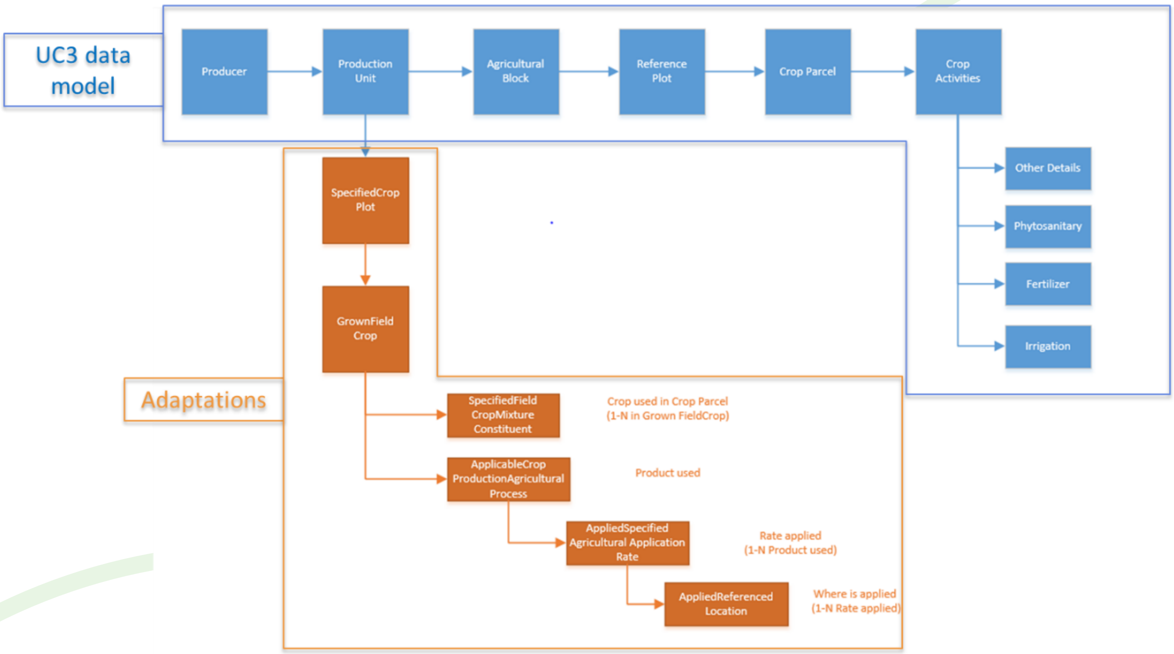


Figure 43 Adaptations from UC4b (Machine data) to feed the data model of UC3 (Farm Registry)

Thus, specific developments are needed for the use of information coming from the FMIS in order to be stored in the Farm Registry.

3.5.2.3 Results and lessons learnt

Regarding the lessons learnt due to this testing experience besides the importance of common working units (mentioned above), we have the codes issue, that must be taken into account, because codes not converted to a NIVA standard will not be valuable for the rest of the participants, as they will not know the meaning.

3.5.3. Use Case 4b: Machine data

3.5.3.1 Use Case objective

UC4b aims to reduce the administrative burden, because the data will flow fluently from a machine to administration and to simplify governance due to the increased data quality (precision in time, location and activity).

This use case is intended to let data flow from farm machine to the Payment Agency (PA). To make the case more concrete a specific example was chosen. In this UC, the first focus was on the sowing of the catch crop. In the Multimember State testing phase other field activities will be tested.

3.5.3.2 Context of experience

Till now the direction tested by UC4b was from FMIS to IACS, automatic data exchange from IACS to FMIS being already operational in the Netherlands.

The concerned data involves **information from the machine** such as geometry (type, polygon), location and rate of application (kg/m2) and **information from FMIS:** farm, farmer, farm plot, catch crop, physical process (when, how), mixture (percentage), crop (botanical), specific crop (species)

In UC4b, four focus groups (stakeholder groups) are distinguished. The two main ones are the farmer and the Paying Agency. In the end these parties have to exchange data. The farmer to apply for CAP payments, the Paying Agency (PA) to process the data so that it can conclude a payment can be done. In the data environment two other stakeholders are of great importance; the machine builder which builds machines which capture the data where we are interested in in this use case, and the farm management information system (FMIS). Most farmers have a system where they organise their data. The smart phone (a mobile version of a computer) could also be an instrument for a farmer to manage his/her parcel information.

- Release 1 is implemented with a Bogballe machine and the Seges FMIS (Denmark). Seges is also a farmers union. The additional software for creating an eCrop message from machine and FMIS data, validating the message and sending it to the PA is implemented in open source on a Seges platform.
- Release 1 is implemented with open source infrastructure at the WUR (Wageningen University, Netherlands)
- Release 1 cannot be implemented at other FMIS's than Seges
- Release 1 supports only the Danish Bogballe machine
- JSON message is stored in PA database as one object.

In release 2 the 'entrance' will be made more generic and the end point storage at the PA will be a more detailed relational database.

3.5.3.2 Technical solution

The farmer plans/specifies the activities for the machine on the parcel by making a 'Task map' on the Farming Management Information System (FMIS). First the farmer transfers the 'Task map' via an app of the farming machine manufactory (in our single Member State test a Bogballe) to the data hub (OEM) of the machine supplier.

The machine records a field treatment in an 'As applied map' by logging the executed activities of the machine. First the farmer could start executing activities along the parcel contour so delivering an accurate treatment boundary in the 'As applied map'. After the parcel boundaries have been stored in the app, the inner part of the field can be treated (e.g. sowing fertilizer, catching crop, mowing around bird nests).

The data hub of the machine supplier transfers the treatment data to the data hub of the FMIS of the corresponding country (e.g. Denmark: FMIS SEGES; SEGES Cropmanager cloud). The Farming Management Information System (FMIS) receives and validates the data file. The FMIS connector sends the validated data as an e-Crop message to the webserver of the national Paying Agency (e.g. RVO for the Netherlands).

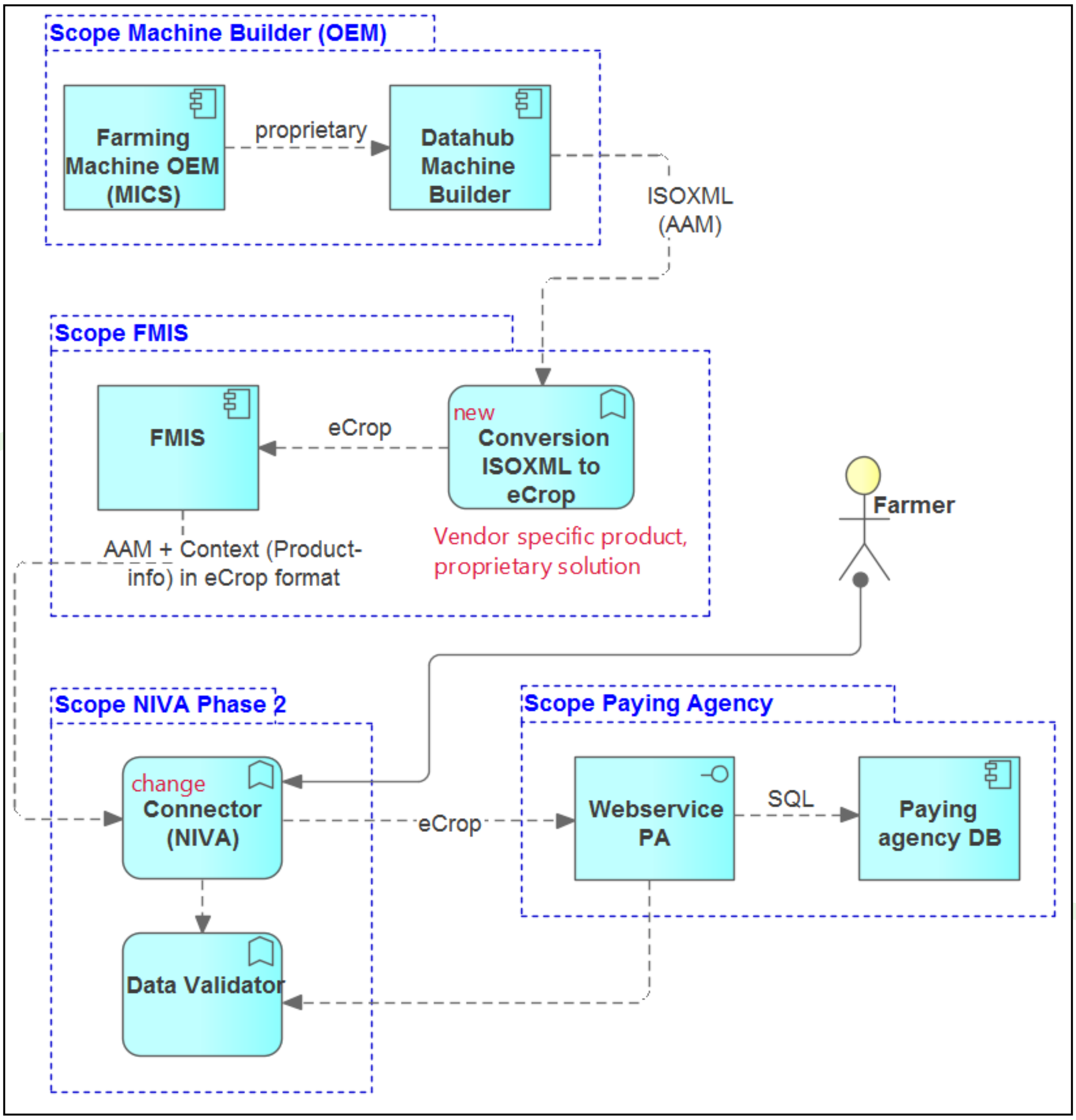


Figure 44 Technical solution for the exchange of machine data

The communication between Farm machine and FMIS is using ISOXML standard. The communication between FMIS – Connector – PA is via eCrop message. The farm machine provides source of rate, location, geometry whereas the FMIS provides source of farmer, plot, crop, activity, product (this data is not present in the machine).

Efforts have been done to make the tool as reusable as possible: functions of the application components are open source software, use of standards (Use of eCrop message and ISOXML message, Exchange protocol JSON. The accounts used for the PA are UC4b-pilot-accounts. So each PA can implement its own security standard.

However, some parts are specific: the Read functions of the FMIS are proprietary software, because the database of the FMIS is FMIS specific, the FMIS implementation was done at Seges and at akkerweb, there is need to implement different Farm machines.

3.5.3.3 Results and lessons learnt

The central issue, "can we upload data from machines to a Paying Agency", has been proved positive.

The main learning's and difficulties have been the following:

- It is not possible to send a complete eCrop message directly from the machine to the PA. Information from the FMIS has to be added to complete the message (farmer, product, crop, type of activity).
- The new data set has to be integrated in the working process of the PA. We are ahead of the PA.
- The FMIS organizations are reluctant to invest in new connections when there is not yet a direct commercial interest.
- The software has to be developed further for more user friendliness.
- Software for reading machine data has to be developed further for more universal use. Seges works on this (ISOxml)
- The focus was on the implementation of the data flow. The quality of the data needs more attention. At this moment the message can be adapted.
- We do need EU and national regulations in order to implement these possibilities in the working processes of the PA

3.5.4 CAP Markers & Data Signals Sharing Component

3.5.4.1 Use Case objective

CAP Markers & Data Signals Sharing Component has been developed under WP4 of NIVA project and aims to act as a single point of access, capable to provide heterogeneous information items derived by multiple sources through a standardised OGC compatible API. Information sources examples include EO crop classification engines outcomes (e.g. predicted crop types escorted by a confidence level probability), geotagged photos (digital photographs with spatial information) and **farm management information systems (digital crop calendars).** This system facilitates agricultural data collection and integration from various sources aiming to support public administrators, farmers and agricultural consultants towards the monitoring of various indicators including (new) CAP monitoring and evaluation.

3.5.4.2 Context of experience

Up to day the technical approach of using a GIS system for rendering data related with agricultural practices where more focused on relatively static information items e.g. crop type and parcel's polygon, parcel's area, parcel's administrators (farmer) details and for data derived by Earth

Observation sources (e.g. satellite derived data products like NDVI). The “CAP Markers & Data Signals Sharing Component” aims to extend/enhance these existing approaches and support the integration of data provided by in-situ sources (e.g. farm calendars, sensors deployed at fields).

As it is illustrated in next figure, the use of OGC API allows the federated use of additional-existing and open geo-data repositories that utilize standardised OGC compliant APIs for data provision. Such repositories are maintained by national or regional administrations, research organisations, statistical agencies, etc. and provide data on areas of particular interest (e.g. Natura2000, wildlife sanctuaries, water bodies, carbon footprint, etc.)

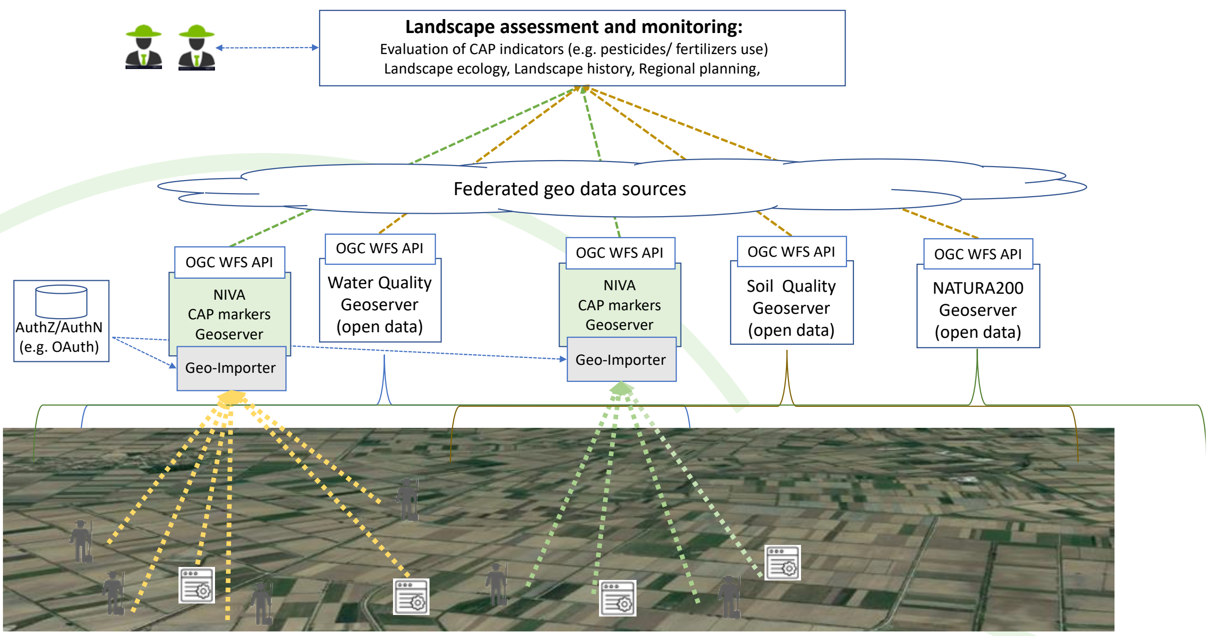


Figure 45. A conceptual view on the utilization of “NIVA - CAP Markers & Data Signals Sharing Component” for landscape monitoring.

3.5.4.3 Technical solution

This approach provides the technical means for capturing up-to-date information of applied inputs, pest infestations, cultivation’s phenological growth stages, etc. on a farm/region/national level towards the realisation of landscape monitoring concept.

In order to avoid developing an additional data management platform, the GeoServer (http://geoserver.org/) open source server is reused. As it is stated[19]: “*GeoServer implements industry standard OGC protocols such as Web Feature Service (WFS), Web Map Service (WMS), and Web Coverage Service (WCS). Additional formats and publication options are available as extensions including Web Processing Service (WPS), and Web Map Tile Service (WMTS).*”

[19] http://geoserver.org/

As it is illustrated in following figure, data import is facilitated through the simple use of web service implemented as a RESTfull API ("Geo-Importer") which is easily expandable for additional data types. Data export of collected data is realized through standardised OGC WFS/WMS API calls ensuring syntactic interoperability while semantic interoperability is satisfied with the use of dominant agricultural data modelling approaches e.g. eCrop. Access control on OGC WFS/WMS API is realised though the mechanisms that are already available by geoserver implementation.

The "Geo-Importer" API is expected to be consumed by existing systems (e.g. FMISs) that collect field data and have the appropriate authorisation/authentication credentials in order to provide selections of these data sets (e.g. calendar of agricultural practices) referring to a specific parcel/farmer. The "Geo-Importer" API supports basic CRUD (create, retrieve, update, delete) operations for the various data providers and role based access control. It should be noted that management and administrative rights on the provided data sets are foreseen only for the initial data provider.

More technical details on this component are available at NIVA gitlab code repository: https://gitlab.com/nivaeu/WP4_cap-markers-data-signals_ogc_api

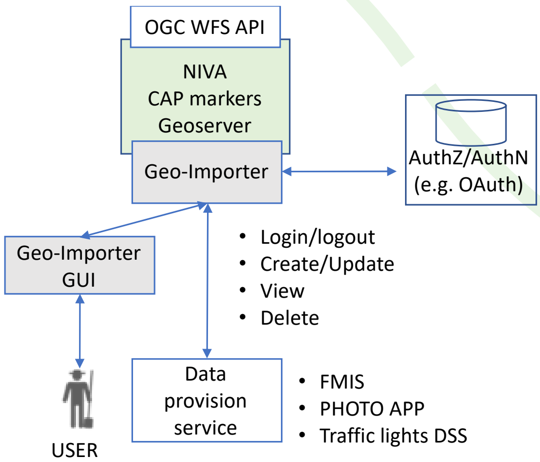


Figure 46. A functional view of the NIVA - CAP Markers & Data Signals Sharing Component

3.5.4.4 Results and lessons learnt

The implementation of this component is completed and has already been presented to NIVA partners. The first impression is that it attempts to address an existing gap on enabling the interconnection of heterogeneous information sources, especially FMISs, through the use of standardised data exchange and modelling mechanisms. The component is currently in testing phase by NIVA partners.

3.6. Feedback from FMIS providers on data sharing with IACS

In order to expand our information collection process - even beyond NIVA consortium - a questionnaire was specified and circulated to various FMIS providers. The questionnaire template is available at “Annex 3. Template for FMIS providers”. The questionnaire was replied on June 2021 by FMIS providers - mainly SMEs – that are listed in Table 18.

FMIS provider (SME)	FMIS solution	Link
ABACO	ABACO Farmer	https://www.abacofarmer.com/
Farmnet365	Farmnet365	https://www.365farmnet.com/en/
Neuropublic	Gaiasense	https://www.gaiasense.gr/en/gaiasense-smart-farming
Hispatec	ERPagro	https://www.erpagro.com/
Horta	Grano.net; granoduro.net; orzobirra.net; mais.net; girasole.net; pomodoro.net; legumi.net; vite.net; uva.net; olivo.net	https://www.horta-srl.it/
Seges	SEGES Crop Manager	https://cropmanager.dk/#/?currentLanguage=%22en%22
Smag	Smag FARMER / Smag EXPERT	www.smag.tech

Table 18. FMIS providers replied to FMIS-IACS data exchange questionnaire.

The full set of the replied questionnaires is available for review and further processing to NIVA partners through the NIVA project data repository. For any other interested party the replied questionnaires are available upon request. The following paragraphs present the main outcomes that are extracted based on an analysis of the received replies. It should be noted that the names of the FMIS providers (SMEs) are not explicitly included in the analysis to follow in order to avoid issues related with the commercial promotion of the offered services.

3.6.1. Analysis of replies

- **FMIS use**

The first set of questions focused on the use of the FMISs, including the number of farmers that are supported, the area of cultivated land that the services are currently supporting, and the number of countries that the FMIS operates. As it is deduced by the replies, our review has managed to capture a representative sample of FMISs that are currently in use in EU level ranging from FMISs that are operating in a national level with relatively small focused user base (e.g. FMIS that supports 800 farmers in one country) to FMISs that are operating in number of countries even outside EU (FMIS that supports 800,000 associates/users in more than 10 countries). Similarly, with regards to the total land covered by the reviewed FMIS the replies range from 70.000 ha to 49.400.000 ha. It should be noted that in many cases, beside the numerical reply there were comments that the FMIS providers were not in a position to define the exact actual number of active users and that the provided numbers are only an approximation.

The next question was on the type of the end users that the FMIS provides services. The respective replies indicate that most of the FMISs are mainly aiming to support individual farmers through their agricultural advisors. Some of the FMISs, but less than the majority, are also offering their functionalities through the farmers' cooperatives.

- **Functionalities offered by FMISs**

With regards to the supported functionalities the FMISs representatives had to select from a predefined list: Parcels and farmers profiles Inventory, Farmer's calendar or Field book, Exporting of basic reports, Exporting of reports for specific purposes, Advanced decision support on applied cultivation practices, Machinery Management, Finance management, Human resource management, Quality assurance and advanced monitoring. More details on each functional category were available within the provided questionnaire and are also available at the Annex of this document. Figure 47 illustrates the recorded replies.

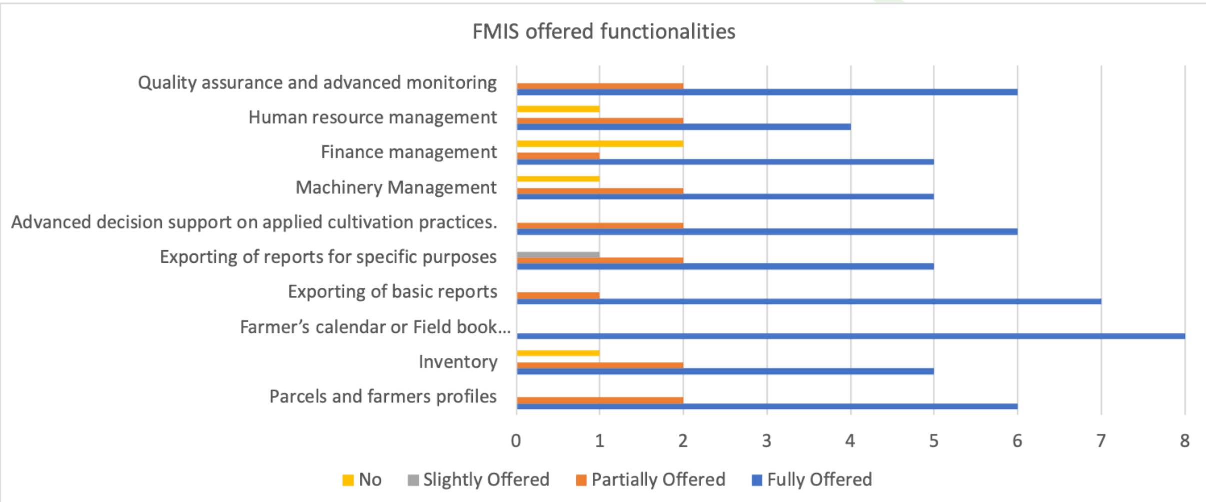


Figure 47. Reported replies on FMIS offered functionalities.

Within the context of this deliverable's research objectives (FMIS-IACS data exchange) we are particularly interested if the FMIS maintain the following information items:

- Details of parcels (location, area, polygon) and farmers profiles (name, age, national identification number, etc.).
- Field operations Management (including operations of machinery) and Farmer's calendar (also known as Field book).
- Exporting of reports for generic and specific purposes.

The existence of such functionalities is associated with the ability of the FMIS to directly provide information entities that are of interest of IACS as these have been identified in section 3.2

With regards to **"Parcels and farmers profiles"** all FMISs replied that are fully or partially supporting it**.** With regards to **"Field operations Management and Farmer's calendar or Field book"**, all of the 8 FMISs are keeping a digital record of the operations applied at the field. However, there are various differentiations reported on the type of recorded information and the way that this information is imported (e.g. automated functions, manually recording, use of 3rd party services for live tracking).

With regards to **Exporting of basic reports** all FMIS provide this functionality while **Exporting of specific reports** is offered to some extend by 7. As it is commented by FMIS provider specific reports are related with managerial and agronomic purposes, for execution and management of the activities to be carried out on the basis of the obligations and compliance established by both internal and external (e.g., organic code of rules, product specifications) and in another case for automatic calculation of KPIs (e.g. for sustainability targets is supported).

- **Information items maintained by FMISs**

Within the next set of questions the FMIS representatives were asked to report explicitly the information items that their systems is handling, the source for these items and also to comment whether the process of recording these items derives from any kind of regulatory obligation. Figure 42 illustrates the replies for the more significant (for IACS) information items.

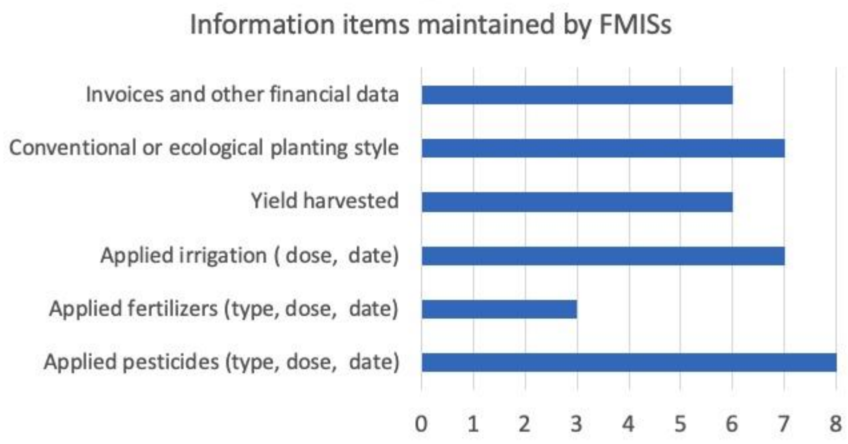


Figure 48. Information items maintained by FMISs

As it is obvious all the FMISs are recording pesticides application related activities which are related with the respective national regulation. On the contrary the application of fertilisers is not recorded by the majority of the FMIS.

- **Background BaseMap utilised by FMIS**

The next question was on the BaseMap type that potentially is utilized by the FMIS. The received replies are reported in next, where each row corresponds to one FMIS.

FMIS	BaseMap type	GIS Information
1	Google Maps	SHAPE Files or ISOXML-Datasets
2	Manual capture	WGS84 / EGSA87
3	GeoServers and WFS Services provided by third parties	SHP, GPX, KML or other type of DATAfiles
4	Orthofoto and cadaster-data and quality-checked against official parcel-maps.	Data is stored in UTM Zone 32 (official danish zone).
5	GoogleMaps	Proprietary, export in GeoJSON, ShapeFiles.
6	*MAXAR SecureWatch*	Graphical digitisation of plots and/or automatic recognition on orthorectified BaseMap and/or using national registries provided by third party services with differentiated reference systems including spherical marker and geographic (EPSG 3857-4326-32632-32633).

Table 19. GIS information in FMISs.

Not all FMIS replied to this question and the various systems are utilizing heterogeneous approaches.

- **APIs provided by FMISs for data exchange**

This question is focusing in one of the most important technical issues, the availability of APIs along with the respective technical specifications of the FMIS. The existence or not of FMIS APIs for data exchange with 3rd party systems is considered as a stepping stone towards the specification of technically sound manner for interconnecting with IACS. Given the importance of the issue we present in next table the full set of the received replies.

FMIS	API Type	Security mechanism	Exchange format	Exchanged data model (Semantics)	Comment
1	REST	OAUTH2.0	JSON	"Harmonized Data Bases (example: Homologa Catalog by Lexiagri)" "Standard codes for Fertilizers and Seeds depending on the market/customer" "ISO11783-Datasets for Ag Machinery"	
2a	REST API	Token based	JSON	Custom JSON structure, using mainly digits in order to reduce the size of the transferred data.	The API is only utilized for internal purposes or with predefined 3rd party service.
2b	REST API – Orion Context Broker	Token based	JSON	NGSI v2, Agriculture related smartdatamodels provided by FIWARE.	This service only provides data for selected pilot fields that have been part of H2020 projects eg. IoF2020.
3	REST API	SSL	JSON	"Currently, FMIS is following the DATA normalization based on Ministry of Agriculture in Spain."	
4				The EPPO glossary is utilized for plant and pest types https://gd.eppo.int/taxon/ as well as other standards	Data exchange with third parties have been performed in the frame of research project. Custom formats have been used.
5	REST API	OAuth	JSON	ISOBUS and eCrop.	
6a	REST API for Farmers Log /calendar data export	OAuth	JSON	In general custom key, values. Agrovoc taxonomy is used for crop type and pests.	
6b	REST API for weather forecast	token based	JSON		No

	data import				
7	Rest API		JSON	•BBCH standard scale - Phenology phases •OGC-WFS standard – Geometric objects •OGC-WMS standard – Raster maps	Comment: the system can host several standards.

Table 20. FMIS APIs for data exchange

As it is obvious from the received replies, almost all of the FMISs use REST API technology for exchanging data using the JSON format. With regards to applied security mechanisms the use of OAuth[20] is a popular approach along with custom token based mechanisms. As it was expected, there are various heterogeneous approaches on the semantic models utilised for formulating the various exchanged data items. There is an obvious effort from FMIS providers to confront with standards and dominant modeling approaches (e.g. use of Agrovoc taxonomy, EPPO glossary, BBCH standard scale for Phenology phases, ISOXML, OGC-WFS/WMS, FIWARE data models for agriculture).

- **Export of reports**

In this question the FMIS providers where asked if their systems are capable to export the maintained data to file. All add the FMIS replied positively supporting popular file formats like: PDF, CSV, XML. Some of the FMISs are also capable to export dataset in ISOXML, SHP and GeoJSON formats.

The set of questions that followed were less technical and focused on the opinion of FMIS representatives on various challenges that need to be addressed in order to establish a sound data exchange approach between FMIS and IACS.

- **Data validity of FMIS data.**

The first question was on data validity and the issue of intentionally or unintentionally recording inaccurate data within the FMIS.

"There is currently no concrete mechanism to validate 100% of the information that farmers have documented/provided within the platform…. However, the main challenges can be addressed by two central points: 1. achieve farmer's trust to use digital technology and 2. achieve certain level of technical understanding to use digital tools in a profitable way."

"Overall it is easy for the farmer, advisor, FMIS operator to "cheat". However if the data set is rich enough "cheating" is more difficult as inconsistencies will occur. For example a parcel with poor soil that produces high yields with low quantities of fertilisers is an irregular situation. Another approach

[20] https://oauth.net/

is the farm calendar records to be escorted with additional hard evidence, for example the entry referring to an application of fertilisers to be escorted with the records of the respective invoices of purchase for the chemical and even changes on soil salinity that might be recorded by a soil sensor."

"Technical tools for monitoring follow-up actions per each planned activity which can monitor the actions done vs. the planned actions in order to measure the deviations."

"… there is no way to prevent fraud from farmer. A parallel certification system has been developed to verify data input by farmers (e.g., malting barley)…"

"… In my view, the 'no fraud from farmer' can be handled only with on-site checks/audits."

"… The challenge is always about integration and standardization."

- **Existing integration of the FMIS with the national/local IACS**

The next question was on existing integration channels of the questioned FMISs with IACS e.g. in the scope of CAP related payments. Two of the FMISs reported that are fully integrated with IACS through the use of web services. Three reported that they have no connection established. Three reported that have a partial integration with IACS (e.g. for retrieving the geometries of the parcels, for accomplishing the official Field Notebook required by Ministry of Agriculture, and for indirectly assisting the declaration process through the exporting of the required forms).

- **FMIS dataset that are useful for IACS**

The next question was on which FMIS dataset types are considered as useful for the IACS. In this question not all FMIS representatives replied which can be considered as an indication that they are not really familiar with the IACS processes. However, datasets related with Crop, Farmers calendar, and Earth Observation are among the most popular. Figure 49 illustrates the received replies.

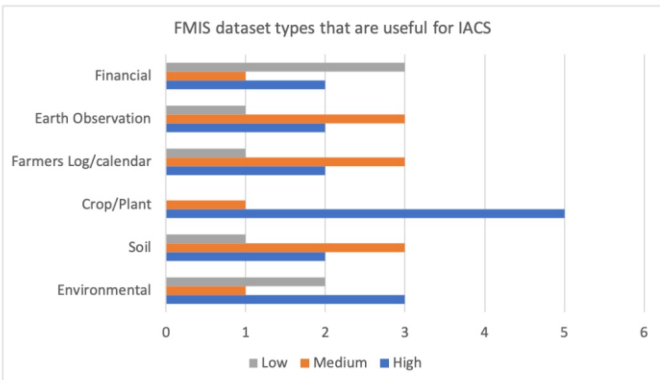


Figure 49. FMIS data types that are useful for IACS.

- **FMIS Data types that are difficulty to be shared**

The last question focused on the sensitivity and reluctance for sharing for the various data types. FMISs were asked to on which of the data categories that are handled by their systems are considered more difficult to be shared because of various issues (e.g. legal, business & competition, farmers not trusting the recipient of the data, etc.). Figure 50 illustrates the received replies.

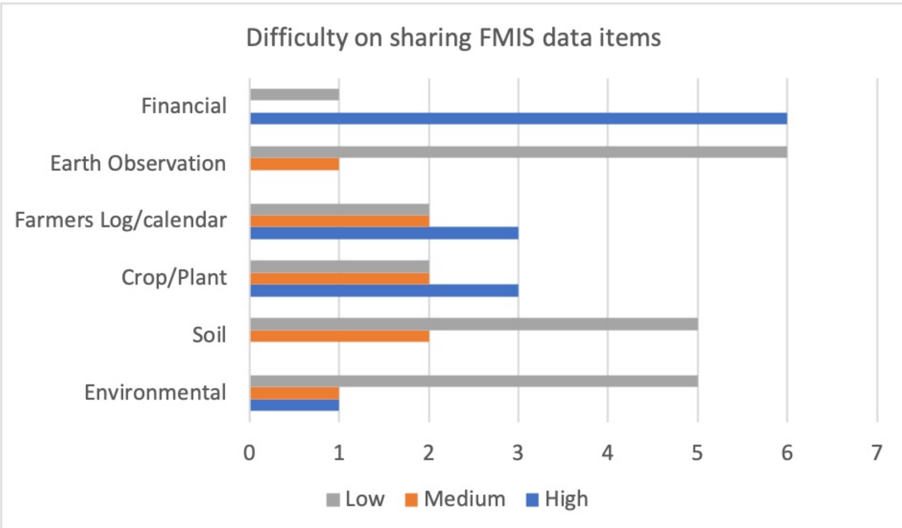


Figure 50. Difficulty on sharing various data items maintained by FMIS.

As it was expected data that are related with financial outcomes are the most difficult to be shared. Farmer's calendar and plant related data are also considered as sensitive.

3.6.2. Consolidated outcomes of the review

The key conclusions from the survey analysis are included in the following section:

- There are currently various FMIS systems that demonstrate different heterogeneous characteristics. The FMIS may focus on specific domains of farming activities.
- One important outcome is that in most of the cases the individual farmer needs the assistance of the farmer advisor in order to exploit the services offered by the digital solutions provided by the FMIS. This is an important parameter that needs to be considered in a future scenario where the FMISs will have the role of farm's data mediation to local IACS. Given the current conditions the role of the farmer's advisor will remain significant in order to assist the overall process of selecting the appropriate data sets and transfer them to the IACS.
- Based on the received replies the majority of the FMISs are not considering -yet- the public administrative agencies (e.g. Paying Agencies) ad their information system (e.g. IACS) as potential 3rd party entities that are useful to interact with, either to provide data or to retrieve data. This means that currently there are few or even no data exchange mechanisms by FMISs explicitly established for connecting with IACS systems. This will also be analysed on dedicated replies on data exchange.
- FMISs will continue to extend their field of operation including additional, currently, unforeseen areas.

- With regards to FMIS interaction with other systems the most popular approach is the use of REST API with JSON format payloads secured by existing dominant/interoperable approaches (e.g. OAuth). The respective semantics are still heterogeneous however there is an effort towards harmonization based on existing approaches with a proven record (e.g. Agrovoc[21] taxonomy, EPPO[22] glossary, BBCH[23] scale, ISOXML).
- There is no simple solution on addressing the issue of inaccurate data. Interesting approaches have been recorded including the “building of trust with farmers”, “on the spot checks”, “provision of rich data sets for cross-checking”.
- Data items related with applied cultivation practices and financial activities remain the most sensitive to share. However environmental, earth observation and soil related data are easier to share. These data sets are useful for monitoring.

3.7. Conclusions

3.7.1. Potential benefits of data exchanges (FMIS to IACS)

- **Data requirement analysis to be updated**

This deliverable is trying to identify the FMIS data that may be used for IACS, based on the work done by the NIVA Use Cases.

FMIS data can be used to check current eligibility rules according to the new EO monitoring system. The investigation of potential useful FMIS data has been partly done by UC4b (machine data) and it might be refined (if considered as relevant) before the end of the NIVA project.

FMIS data can also be used to check the future eligibility rules of CAP post-2020; UC1b, UC1c, UC3 have explored which FMIS data would be required. However, this new CAP is still draft regulation and in addition, it will let more flexibility to MS (national strategic plans). There will be need to identify accurately which FMIS data will be necessary or at least useful in the context of this new CAP. For instance, the UC3 (farm registry) data model is an attempt, a reasonable guess about what will be required for CAP post-2020 but it will have to be revised and updated once definitive regulations are in place.

In summary, NIVA has performed a good analysis of the main data requirements but extra work will be required to get more detailed and updated understanding of the expected FMIS data for the monitoring of CAP post-2020.

[21] http://www.fao.org/agrovoc/
[22] https://gd.eppo.int/
[23] https://www.julius-kuehn.de/en/jki-publication-series/bbch-scale/

- **Data availability by FMIS**

"Digital field books" - usually as part of a FMIS- is potentially one of the most valuable information sources for IACS especially in the context of CAP monitoring. Unfortunately, until now, FMIS are not yet used by a wide range of farmers, the situation varying a lot depending on countries and regions.

However, the use of digital means (e.g. FMIS) for recording applied practices is expected to be continuously adopted by more farmers during the next years due to introduced legislation but also because it is more efficient for the everyday activities of the farmers.

Typically, recordings of pesticides applications are more mature to be captured and shared with various administrative entities (e.g. IACS and other CAP monitoring services. Similar policies and tools are currently under development for the collection of data on utilized irrigation water and fertilisers/soil nutrition status (e.g. FaST).

3.7.2. Data exchange & interoperability issues and solutions

- **Semantic interoperability**

The semantics utilized for recording the various information items maintained by FMISs are still heterogeneous. However there are various (in parallel) efforts towards harmonization based on existing well known approaches especially with code lists (e.g. Agrovoc taxonomy, EPPO glossary, BBCH standard, code lists for fertilizers and PPP).

Other EU initiatives are working towards the harmonization/standardisation of agricultural data semantics (e.g. H2020 Demeter, H2020 Atlas, FIWARE smart data models) following more complete data modeling approaches based on Semantic Web technologies (RDF/RDFS/OWL and JSON-LD) (section 3.4.2). However, according to our review semantic web technologies are not inherently utilised by FMIS systems (section 3.6.2) or at least there is no dominant data modeling approach. The use of ontologies for harmonising agricultural data can be utilised in combination with software based modules (also known as "Outbound-Inbound interoperability enablers") which act as data translators and mediators. Such an approach is also followed by the NIVA "CAP Markers & Data Signals Sharing Component" common component (section 3.5.4).

A suitable code lists uniquely identifying crop types for pan-European use have to be agreed. The identification of such a list will be of particular importance especially if an approach of reaching a minimum-interoperability-level and then start building on top of it is followed (for more details on this interoperability applying approach see section 3.4.3). The Common Semantic Model of NIVA is proposing a code list combining the species and the product classification based respectively on EPPO and GPC standards but this approach should be presented and discussed with FMIS editors.

NIVA project also utilized the eCrop Standard for modelling of agricultural data with promising results. In general, the eCROP standard was found to be suitable for standardized data exchange messages between Paying Agency's IACS system and FMIS. However, some modifications/extensions to the eCROP standard are necessary: typically, the standard is rather complex and has to be profiled for specific applications; in addition, it does not include any code list.

- **Quality and reliability of FMIS data**

Although in some cases record keeping in “Field Book” is even mandatory the recording of information is still fragmented and prone to intentional or unintentional errors. Data related with cultivation practices is not reliable given that it is manually imported to FMISs (by farmer or advisor). An approach to mitigate inconsistency of manually imported records is to escort them with data derived from additional sources (e.g. farm machinery, geotagged photos, environmental sensors and hard copies of invoices). As it is indicated in section 3.6.2 there is no simple solution on addressing the issue of inaccurate data. Interesting approaches have been recorded including the “building of trust with farmers”, the combined use of FMIS data with “on the spot checks”, and the “provision of rich data sets and additional evidences for cross-checking including data from farm machine, sensors and scanned copies of invoices”.

- **Technical interoperability**

With regards to data exchange the use of web services based on REST API is a well-accepted approach among FMIS. Preferred security (authorization/authentication) mechanisms are varying but well established solutions (e.g. OAuth) are well accepted. The use of (outbound-inbound) interoperability enablers deployed on top of existing FMIS is a well-established practice which is currently further developed within various EU projects. The “FMIS outbound” enabler can operate as a “data export module” that will on demand transfer data to selected IACS.

- **Organisational interoperability**

One of the NIVA objectives is to reduce the administrative burden on farmers. Common choices between FMIS and IACS would contribute to make farmer life (and data exchange) easier.

This is of course the case of semantics (use of common code lists) but other topics might also be considered. For instance, use of common Base Map and of same Coordinate Reference System would facilitate the management of geographic data by farmer. The other example is coming from geotagged photos applications. When a FMIS is using such application, it would be great for farmers to use same application for the new CAP monitoring process.

FMIS editors and Paying Agencies might or even should be invited to cooperate in order to find common solutions.

3.7.3. NIVA achievements regarding data exchanges

- **Exchange from IACS to FMIS**

This data exchange would facilitate the farmer declaration; this would avoid some geometric error (e.g. ensuring that agricultural parcel is within reference parcel). This would reduce the administrative burden for the farmer and lighten the need for some controls for PA.

Developing exchange solutions is feasible at national level, from a specific IACS to a specific FMIS, as proved by the fact that such system is already in place in the Netherlands and that the NIVA Use Case on Farmer performance (UC1c) has developed an exchange tool to exchange data from ARIB (Estonian Paying Agency) to EAgronom (an Estonian FMIS).

- **Data exchange from FMIS to IACS**

The benefits of data exchange from FMIS to IACs have been widely investigated in this document and the clear conclusion is that it would be very useful for the new monitoring system and even more for the new CAP.

The NIVA experiences have proved that developing data exchange solutions is technically possible and such solutions have been proposed by the NIVA Use Case about machine data (UC4b) and by the WP4 CAP Markers & Data Signals Sharing Component.

However, the main issues are not technical but legal and organizational. In practice, the development of such tool has to be done with the strong involvement of the FMIS editor. In the NIVA project, such tools have been developed by SEGES (UC4b tool) and Neuropublic (WP4 tool) who are project partners and have dedicated resources to do so.

It would be difficult to extend this experience during the NIVA project time: FMIS editors develop new functionalities based on the requirements of their users, i.e. of farmers but farmers have not yet the obligation to provide FMIS data to get CAP payments.

More generally, there is still lack of the appropriate regulatory environment and the use of FMIS data by Paying Agencies has to be recognized by legislation.

- **The NIVA steps towards interoperable solutions**

NIVA ambition is to develop generic tools easily reusable in any European country, with reasonable adaptation efforts.

This goal looks more difficult to be achieved as the matching rules between the systems (IACS and FMIS) are quite system specific. Despite of this difficulty, the NIVA project has elaborated a first proposal bringing a lot of improvements as shown by the figure 51.

This proposal is based on the use of the eCrop standard for data exchange between FMIS and IACS and by a proposed common model for the Farm Registry (the Farm Registry is expected to become a component of the new IACS, its role being to record in continuous way the farmer information provided to the Paying Agency).

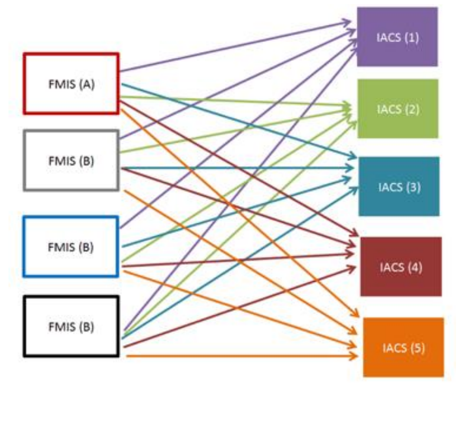
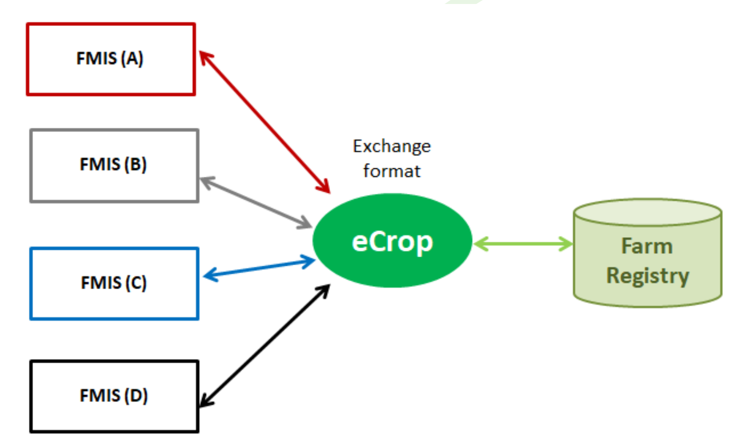
<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Potential data exchanges between FMIS and IACs before the NIVA project</td>
<td>NIVA proposal for more standardized exchanges between (new) IACS and FMIS</td>
</tr>
</table>

Figure 51. Benefits of NIVA proposal for data exchange between FMIS and IACS

Currently, the NIVA proposal is mainly a concept whose feasibility has been proved by the development and testing of prototype tools. It may offer a solution for future but lots difficulties have to be solved and lots of steps have to be done before getting an operational exchange system. The list below reminds the main conditions to be achieved to ensure the proposed scenario may be efficiently put in place but is far from being exhaustive:

- Use of FMIS data for CAP payments should be legally recognized
- Paying Agencies should be convinced to adopt a common data model for the Farm registry component; the one developed by NIVA UC3 can provide a good starting point
- The use of eCrop standard has to be accompanied by harmonization efforts on the code lists; the purpose is that FMIS and IACS use the same or at least compatible code lists
- FMIS editors should be involved in the process as they are the ones who should develop the export modules from their FMIS to the eCrop standard.

4. General conclusions

This general conclusion tries to answer two main questions: what are the achievements regarding harmonization and interoperability? what are the next steps (if any)?

- **Earth Observation data**

D3.5 has focused on technical interoperability, as it was expected. D3.5 is not providing any standardized solution but a state-of-play of possible solutions together with a few recommendations.

The chapter about EO data should provide very useful knowledge to Paying and should hopefully help them to set up their new Area Monitoring System. The work is considered to be final in the context of the NIVA project

However, combined with the on-going work on the data model “Base types for EO monitoring” conducted in deliverable D3.2 Common semantic model , it might be used as a starting point to a standardized way to describe EO monitoring source data and processes. This goal is outside the NIVA objectives but, if considered as relevant by the European Commission, this might be a topic to be investigated by a follow-up project or other kind of initiative.

- **Data exchanges between FMIS and IACS**

D3.5 has investigating all the aspects of data exchange between FMIS and IACS: the experience has shown that it was not relevant to isolate the technical interoperability as initially planned but that it was better to have a wider scope. Clearly, legal context, organizational and semantic interoperability are preliminary conditions or at least strongly related to technical interoperability issues.

The document is not proposing recommendations to the NIVA Use Cases or to the Paying Agencies but is making a state-of-play of the standardization efforts done within NIVA so far.

D3.5 is a one-shot deliverable with only one version to be supplied at M24 (May 2021) but it is definitively not the end of the story regarding the topic of data exchanges with FMIS.

D3.5 provides a good state-of-play about FMIS in general and the NIVA experiences but it has also raised new questions. These new (or remaining) questions might be considered by the NIVA FMIS Working Group and documented according the most relevant means (e.g. last version of common semantic model, as-is analysis of WP2 …).

5. ANNEX 1 : Questionnaire to monitoring countries

5.1. Foreword

WP3 is preparing deliverable D3.5 about data exchanges between IACS and third party applications. One of the key issues met by Paying Agencies is the access and pre-process to big volumes of satellite images due to the context of new EO monitoring system.

In our deliverable, we are willing to present the feed-back from the experience of the NIVA partners and from the official monitoring countries.

This questionnaire addresses technical issues and should be filled by your EO monitoring experts. We expect you are going to fill most of the questions; however, if you have difficulties to answer some specific questions, you can just ignore it (or ask clarifications to us).

The answers to this questionnaire are expected before 21/05/2021.

Many thanks in advance for your help.

5.2. Context

1. Country **: ………………………**

2. What are the schemes applied in the EO monitoring approach?

………

3. On which area do you apply the EO monitoring approach (e.g. whole territory, a region? Provide approximate area and/or number of concerned parcels

………

5.3. Which satellite images are you using?

Sentinel-2 images

1. Have you used Sentinel- 2 images?
   - Yes ☐
   - No ☐

In case of negative answer, go directly to next chapter (Sentinel-1 images).

2. For your EO monitoring process, are you using all the bands of S-2 images?

- Yes ☐
- No ☐

In case of negative answer, explain which the bands you are using and why (or the opposite in case you are not using a small number of them)

…….

3. Which levels of Sentinel -2 images do you access?

- L1B ☐
- L1C ☐
- L2A ☐
- L3A ☐
- L4A ☐

4. Are you performing some preprocessing by yourself?

- Yes ☐
- No ☐

If yes, which ones?

- Orthorectification ☐
- Radiometric correction (Bottom of Atmosphere) ☐
- Mask computation ☐
- Temporal series preparation ☐
- Index computation (NDVI …) ☐
- Other ☐

Provide some more details (especially in case of “other” answer)

…….

5. If yes, what type of software do you use for S2 image pre-processing?

- Open-source ☐
- Commercial ☐
- Home-made ☐

What is the name of this software? …………………………….

6. Which kind of S2 temporal series do you use?

- All available dates ☐
- Resampled series at regular frequency ☐

In case of second answer, which frequency have you chosen? …………………………….

7. Have you assessed the impact of clouds?

- Yes ☐
- No ☐

If yes, can you provide this assessment (e.g. percentage of missing observations)?

…………………………………………………

Do you use cloud or other artefacts masks?

- Yes ☐
- No ☐

If yes, provide some more details (e.g. type of masks, algorithm used to compute these masks, how do you use these masks)

…….

8. Have you applied any advanced fusion techniques to S2 data to downscale its resolution (super-resolve the lower-resolution (20 m and 60 m)?
   - Yes ☐
   - No ☐

If yes, provide more details

……..

9. Do you need to access satellite imagery available on the Long Term Archives (images older than 1 year)? If yes, provide more details (for which purpose? on how many years/months old?).

………

……..

Sentinel 1 images

10. Have you used Sentinel-1 imagery?
    - Yes ☐
    - No ☐
11. If not, why not?

………

In case of negative answer, go directly to next chapter (other HR images).

12. Which S1-bands or markers are you using? Why this choice?

………

13. If yes, for which reasons? What are the main uses of S-1 images?

………

14. Which levels of Sentinel -1 images do you access?
- L0 (Compressed and unfocused SAR raw data) ☐

- Level 1 GRD/SLC ☐
- Level 2 OCN ☐

15. Are you performing some preprocessing by yourself?
    - Yes ☐
    - No ☐

Please, explain which ones

………

16. If yes, what type of software do you use for S1 image pre-processing?
- Open-source ☐
- Commercial ☐
- Home-made ☐

What is the name of this software? …………………………….

17. Have you faced technical challenges in processing and interpreting radar imagery? Please, explain:

………

18. Do you need to access Sentinel1 imagery available on the Long Term Archives? If yes, provide more details (for which purpose? on how many years or months old?).

……

……

Other HR images

19. Have you used Landsat images?
    - Yes ☐
    - No ☐

If yes, please provide some more details (e.g. for which purpose? which kind of Landsat products? ….)

……

20. Have you used any other HR images?
    - Yes ☐
    - No ☐

If yes, please provide some more details (e.g. for which purpose? which kind of images? ….)

……

NIVA
NEW IACS VISION IN ACTION

Dealing with small parcels

21. What is your definition of a “small parcel”?

………..

22. Have you assessed the number of small parcels that have to be monitored?
    - Yes ☐
    - No ☐

If yes, please provide this (approximate) number: ……………………..

23. Which image solution are you using (or envisaging) for monitoring small parcels?
    - Use of PlanetScope images (3m) ☐
    - SPOT6/7 images (1.5m) ☐
    - Other satellite images ☐
    - Super resolution of Sentinel 2 (downsampling) ☐
    - Other solution ☐

Please, provide more details about the reasons for your choice and about the difficulties that you have met (or that you are expecting)

……

5.4. How do you access and preprocess them?

24. In what way do you access/download the open satellite data (e.g. Sentinel, Landsat)?
- We use the API of Open Access Hub (API Hub) ☐
- We use a DIAS ☐
- We use our national platform/hub (sentinel mirror or data portal) ☐
- We use a commercial web service (e.g. AWS, GoogleCloud, Planet Platform) ☐
- We use another solution ☐

In case you are using another solution, please explain which one:

…….

25. How do you deal with the image pre-processing?
- in house solution ☐
- external contractor ☐
- accessing a ready-to-use product platform ☐

What are the reasons for this choice?

………..

26. Have you used any of the DIAS platforms for image preprocessing? If yes, which DIAS have you chosen and what is your experience regarding the technical readiness and user friendliness of the platform?

……

27. How do you deal with the pre-processing of VHR-HHR imagery?

……

28. Do you use a web-portal for the visualisation of the image products (e.g. colour composites, vegetation index layers)?
    - Yes ☐
    - No ☐

29. Have you used Sen4CAP processing system?

- Yes ☐
- No ☐

If yes, are you using it alone or with a DIAS (and which one)?

……

Have you also used other in-house custom algorithms for crop classification/ activity detection?

- Yes ☐
- No ☐

5.5. IT system

The adoption of EO monitoring implies to deal with big data volumes and to have strong computation power. This may strongly impact the Paying Agency IT system.

30. What changes have you applied to your IT systems in order to perform the monitoring?
    - Buy new computer systems ☐
    - Adopt cloud infrastructure ☐
    - Increase "pipelines" size (network, computing power) ☐
    - Employ new staff ☐
    - Other ☐

Please, explain with some more details

……

Is there any other main learning about your experience on access and pre-process of satellite images for EO monitoring that you would like to share?

……

6. ANNEX 2: Open EO API principles

- **The Open EO project**

The project ran from October 2017 to September 2020. It was a H2020 project with following partners.

Partners

1. Technische Universität Wien (Coordinator), AT
2. Westfälische Wilhelms-Universität Münster , DE
3. Wageningen University , NL
4. Vlaamse Instelling Voor Technologisch Onderzoek N.V. , BE
5. Earth Observation Data Centre for Water Resources Monitoring GmbH , AT
6. Mundialis GmbH and Co. KG , DE
7. Sinergise Laboratorij Za Geografske Informacijske Sisteme Doo , SI
8. Accademia Europea di Bolzano (EURAC Research) , IT
9. Solenix Schweiz GmbH , CH
10. Joint Research Centre of the European Commission , IT
11. Google Earth Engine, US/CH - consortium member receiving **no** H2020 funding

The context was the following: Earth Observation data are becoming too large to be downloaded locally for analysis. Also, the way they are organised (as tiles, or *granules*: files containing the imagery for a small part of the Earth and a single observation date) makes it unnecessary complicated to analyse them. The solution to this is to store these data in the cloud, process them ther, and browse the results or download resulting figures or numbers.

OpenEO has developed an open application programming interface (API) that connects clients like R, Python and JavaScript to big Earth observation cloud back-ends in a simple and unified way.

With such an API,each client can work with every back-end and it becomes possible to compare back-ends in terms of capacity, cost and results (validation, reproducibility)

NOTE: OpenAPI are useful for the Paying Agencies that are not willing to download satellite images on their own infrastructure.

- **Open API**

An API is an application programming interface. It *defines* a *language* that two computers (a client and a server) use to communicate.

An Open EO API enables to have a single interface between various clients or servers (also called back-end).

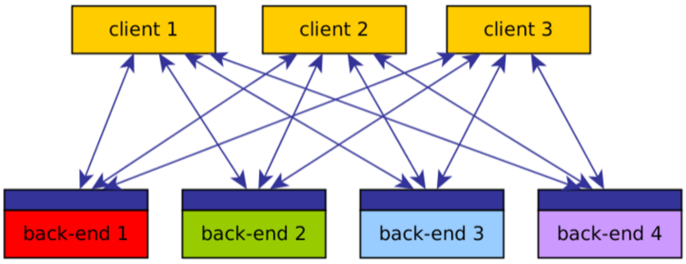


Figure 52. Benefits of open API for data exchange between server & client

However, existing back-ends need to be taught to work with the new API, and clients that interact with back-ends need to be developed.

Open API is a standardised way (in practice, it is a file) to describe end points and parameters, its is dedicated to developers.

- **Open EO API**

The task of the OpenEO project was to design, develop and evaluate an API for cloud-based Earth Observation data processing.

The openEO API defines a HTTP API that lets cloud back-ends with large Earth observation datasets (servers) communicate with front end analysis applications (clients) in an interoperable way.

As an overview, the openEO API specifies how to

- discover which Earth observation data and processes are available at cloud back-ends
- build processing graphs (list of jobs)
- consume such services run the predefined processing graphs)

More detailed information on : https://api.openeo.org/

- **Available data & data discovery**

EO data is discoverable using the STAC (Spatio-Temporal Assets Catalogue) metadata standard. This specification is generic enough to include also non-EO data (e.g. geotagged photos). STAC is a community standard and it is already a de-facto standard. OGC objective is to have STAC integrated in the overall concept of the OGC API (i.e. new WMS, WCS, WFS, CSW …) like it is said here: https://www.ogc.org/blog/4394 .

Figure 53. The STAC community

Most of the services providing Copernicus data (mirrors of ESA Scihub) are providing STAC API / Metadata (Amazon Web Service, Google, Sentinel Hub …). Basically, this is sufficient to be compliant with OpenEO.

- Landsat (catalog)
- CBERS (catalog)
- Sentinel 2 (catalog)
- Google Earth Engine (Collections Only) (catalog), and through OpenEO API at STAC 0.9
- Spacenet (catalog)
- ISERV (catalog)
- EarthSearch is a STAC API for the Earth on AWS datasets that implement STAC (catalog).
- Radiant MLHub hosts datasets for training machine learning algorithms. Authentication is required for the API and an access token can be acquired by visiting the dashboard. See documentation and tutorials.
- Earth OnDemand provides a STAC API for a variety of public datasets.
- Sentinel Hub has a STAC implementation to describe their available datasets. Authentication is required for the STAC compliant Catalog API, provided with paid and free researcher accounts.

Examples of STAC catalogues

The back-end may also be Sen4CAP. This open-source system should be installed somewhere, either locally or using CREODIAS. In the context of NIVA project, WP4 might run the system on behalf of the project in order to enable easier testing of the Open EO components based on Sen4CAP processes.

- **Data process**

The first step is generally the selection of required data. There are two main ways to access the EO data, either by selecting directly the images you want (using their identifiers) or by making a query indicating your area, period and bands of interest => you get a data collection. For EO monitoring, the second case-data collections) is the most useful to get temporal series .

Users can define the tasks to be processed , using either process graph (list of ordered tasks) or a Jupyter notebook. Jupyter notebook is a tool used for experimentation/analysis that requires development of code (python, R, etc.). Installation is super-easy. It is a different approach with respect to a normal Editor of Code as it is more for experimentation, documentation, graphical visualization of results etc. It is a lot used in science community and data analysis community as it allows to see results, changing parameters and document (with markup languages) in a visual way. Earth Observation community is extensively using it as it allows to show also results in map.

OpenEO is rather flexible as it gives freedom to implement any "process" to be applied to data and it also defines a set of processes (https://processes.openeo.org/) even if this set is not mandatory (but of course it improve interoperability among back-ends if all implements the same set of "core" processes). The processes have to be described of course and there is a mechanism similar to the "GetCapabilities of OGC services to do so.

- **Running the services**

The processes defined by users are stored and may be executed. They can for instance be launched in batch mode. Whereas the discovery and the task list elaboration are free, the workbench execution is consuming cloud resources => users have generally to pay for it and to get authorisation. However, there may be some free services.

What has been generated may be used in a view service (WMS), it may be downloaded or used as input in another process. The main interest of Open EO API is to provide an interface that may be integrated into workflows, avoiding manipulation of files.

- **Potential clients**

Open EO has implemented client libraries (Python- JAVA – R). R is a high-level language mostly used in statistics. It is also used a lot with geospatial data. Most of the functions in R are also present (in different versions or libraries) in Python. In general, the libraries can be easily used in a Jupyter Notebook. Being based on Open API, most languages are able to generate automatically most of the interface code (e.g. using a tool like https://github.com/OpenAPITools/openapi-generator).

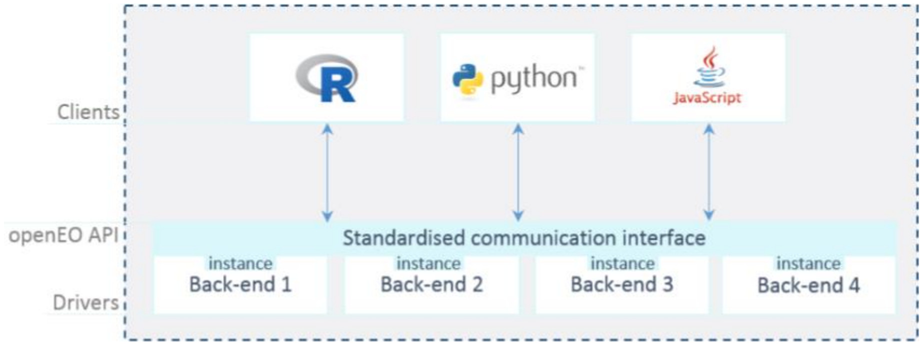


Figure 54. The Open EO architecture with its client libraries

In addition, there are 2 tools, a Web client and a GIS client (QGIS). These tools are for exploration of available data and processes and for the design of processing chains.

7. Annex 3: questionnaire to FMIS providers

7.1. Foreword

Farm Management Information Systems (FMIS) in agriculture have evolved from simple farm recordkeeping into sophisticated and complex systems to support production management. The purpose of current FMIS is to meet the increased demands to reduce production costs, comply with agricultural standards, and maintain high product quality and safety. [FoCa15]

The new Common Agricultural Policy will require continuous monitoring of agricultural practices, mainly based on satellite images and complemented by FMIS data.

The NIVA project is exploring how Paying Agencies might exchange FMIS data with their information system (IACS : Integrated Administraton and Control System) in order to get the necessary data for controlling the payments but also to reduce the burden on farmers by making the declaration easier.

The following questionnaire aims mainly to understand better the data and functionalities of FMIS and its potential of data exchange with IACS.

7.2. Questions:

1. Provide the commercial name of the FMIS and a web reference with more descriptions (if available):
    a. ……

2. Provide an approximation of the FMIS user base: (e.g. number of farmers, number of parcels, area covered)

3. Who is the main user of the FMIS? (farmer, advisor, farmers association, agritech company)

4. In which countries is the FMIS used?

Table 1 presents a set of basic functions that have been identified within the various commercial FMIS software. Most of the categories have specified in the state of the art review presented in [FoCa15]. Please, indicate if and in what extend (fully/partially/slightly) does the FMIS covers the following functionalities.

The FMIS tool may be modular with many existing functionalities but not all of them being widely bought and used (e.g. for cost reasons); please, provide information on this topic. If necessary provide additional descriptions if the FMIS functionalities are not covered by the following table:

Function title	Function description	Yes/No - Comment
Field operations management	Includes the recording of farm activities. This function also helps the farmer to optimize crop production by planning future activities and observing the actual execution of planned tasks. Furthermore, preventive measures may be initiated based on the monitored data	.... Comment: (fully/partially/slightly) ....
Best practice (including yield estimation)	Includes production tasks and methods related to applying best practices according to agricultural standards (e.g. organic standards, integrated crop management requirements). A yield estimate is feasible through the comparison of actual demands and alternative possibilities, given hypothetical scenarios of best practices.	.... Comment: ....
Finance	Includes the estimation of the cost of every farm activity, input–outputs calculations, labour requirements, and so on, per unit area. Projected and actual costs are also compared and input into the final evaluation of the farm's economic viability.	.... Comment: ....
Inventory	Includes the monitoring and management of all production materials, equipment, chemicals, fertilizers, and seeding and planting materials. The quantities are adjusted according to the farmer's plans and customer orders. A traceability record is also an important feature of this function	.... Comment: ....
Traceability	Includes crop recall, using an ID labeling system to control the produce of each production section. Traceability records related to the use of materials, employees, and equipment can be easily archived for rapid recall.	.... Comment: ....
Reporting	Generally includes the creation of farming reports, such as planning and management, work progress, work sheets and instructions, orders purchases cost reporting, and plant information.	.... Comment: ....
Site specific	Includes the mapping of the features of the field. The analysis of the collected data can be used as a	.... Comment:

	guide for applying inputs with variable rates. The goal of this function is to reduce or optimize input and increase output	…………………………………… …………………………………… …………………………………… ……………………………………
Sales	Includes the management of orders, the packing management and accounting systems, and the transfer of expenses between enterprises, charges for services, and the costing system for labour, supplies, and equipment charge-outs	…….. Comment: …………………………………… …………………………………… …………………………………… ……………………………………
Machinery Management	Includes the details of equipment usage, the average cost per work-hour or per unit area. It also includes fleet management and logistics.	…….. Comment: …………………………………… …………………………………… …………………………………… ……………………………………
Human resource management	Includes employee management, including, for example, the availability of employees in time and space. The goal is the rapid, structured handling of issues concerning employees, such as work times, payment, qualifications, training, performance, and expertise	…….. Comment: …………………………………… …………………………………… …………………………………… ……………………………………
Quality assurance	Includes process monitoring and the production evaluation according to current legislative standards	…….. Comment: …………………………………… …………………………………… …………………………………… ……………………………………
Decision support on applied cultivation practices.	Provide recommendations on cultivation practices such as fertilization, pest management, irrigation based on various parameters such as environmental recordings, cultivation type and scientific algorithms.	…….. Comment: …………………………………… …………………………………… …………………………………… ……………………………………
Other	Other	…….. Comment: …………………………………… …………………………………… …………………………………… ……………………………………

5. Provide a description of the various data types that are integrated within the FMIS based on the following categorization (Environmental, Soil, Crop, Farmers Log, Earth Observation,

Financial). *The various data types might be integrated in the FMIS data base structure but more or less used in practice. Please, provide detail on this topic in the comment. Is your FMIS integrating data that farmers have legal obligation to capture (e.g. Field Book)? If yes, please provide us this information (which following data is supposed to be part of the Field Book, if any at column* **Obligatory recordings***)*

6.

Type (Environmental)	**Source**	**Comment**	**Obligatory recordings (Field book)**
(example) Precipitation, Temperature, Humidity, Wind speed, Pressure	(example) Environmental stations integrated with the FMIS. (example) Atmospheric measurements provided by external 3rd party services.		
(example) Weather forecast	(example) 3rd party services and/or forecast algorithms operated by the FMIS		
(example) History logs of atmospheric parameters, daily temperature (min max med)			

Type (Soil)	**Source**	**Comment**	**Obligatory recordings (Field book)**
(example) Soil temperature Soil moisture Soil salinity	(example) Soil sensors		
Nutrients	Sampling and chemical analysis		
Ph	Soil sensors		
Organic matter	Sampling and chemical analysis		

Type (Crop)	Source	Comment	Obligatory recordings (Field book)
(example) Variety	(example) Farmers' observation		
Growth stages	Farmers' observation		
Leaf temperature	Sensor		
Infections	Farmers' observation		
Fruit size	Camera and Algorithm + Farmers' observation		
Yield prediction	Algorithm		

Type (Parcel and Farmers Log/calendar)	Source	Comment	Obligatory recordings (Field book)
(example) Field Location	(example) Farmer, Advisor, FMIS User interface		
Field Area	Farmer		
Irrigation system	Farmer		
Irrigation (dose, date)	Farmer, Irrigation flowmeter		
Planting style	Farmer		
Pruning	Farmer		
Planting day	Farmer		
Harvest day	Farmer		
Yield	Machinery		
Applied pesticides (type, dose, date)			

Type	Source	Comment	Obligatory recordings (Field book)
(example) NDVI	(example) Sentinel II + Algorithm		
LAI	Sentinel II + Algorithm		
Soil moisture	Sentinel I + Algorithm		
Raw images			

Type	Source	Comment	Obligatory recordings (Field book)
(example) Invoices of purchased chemicals (e.g. fertilizers, pesticides,	(example)		
Invoices of sold crops.			
Invoices of other expenses (e.g. fuel)			

Type	Source	Comment
(example) Geotagged photos		

7 What about geographical information? How is the geometry of geographic objects captured? Is your FMIS offering a background BaseMap and which one(s): cadaster, orthoimage …? Which Coordinate Reference System is used in your FMIS?

8 Does the FMIS provide an API in order to interact with 3rd party systems? If yes, what are these 3rd party systems? What security/authentication mechanisms are applied? Please, specify which of the data categories are exported or imported through these APIs. Does the format of the exchanged data comply with any known standards or dominant data models?

API Type	Security	Payload format	Payload Data model (Semantics)	Comment

9. Are you using any standard for the FMIS data exchange? Which one(s)? Please, provide rationale for this choice.

10. Does the FMIS supports the export of data to files?
Which file format is utilized (e.g. xls, pdf)?
Which data are included in the exported file?

11. A key potential issue when using FMIS data for IACS is to get guaranty about the data validity (no fraud from farmer). How can it be ensured by FMIS export functionalities? Is there any experience related to FMIS data export to other stakeholders? Which stakeholders? Which technical solutions were implemented and what are the challenges to be addressed?

12. Does the FMIS currently support any kind of integration with the national/local IACS? Is the FMIS used by farmers when making their declaration to get CAP payments? Have you received request from farmers on this topic? Provide as much detail as possible.

13. According to you opinion, which of the data categories described in question 5 can be potentially useful for an IACS? If necessary provide more details in extra rows or in comments.

Data category	Useful for IACS operations (highly, medium, low)	Comment
Environmental		
Soil		
Crop/Plant		
Farmers Log/calendar		
Earth Observation		

Financial		

14. According to you opinion, which of the data categories managed by the FMIS and described in question 5 are difficult to be shared because of various issues (e.g. legal, business & competition, farmers not trusting the recipient of the data, etc.)

    NOTE: the question is about current status of data sharing in general regarding FMIS data.

Data category	**Difficulty to be shared (highly, medium, low)**	**Comment on reasons**
Environmental		
Soil		
Crop/Plant		
Farmers Log/calendar		
Earth Observation		
Financial		

References

[FoCa15] S. Fountas, G. Carli, C.G. Sørensen, Z. Tsiropoulos, C. Cavalaris, A. Vatsanidou, B. Liakos, M. Canavari, J. Wiebensohn, B. Tisserye, Farm management information systems: Current situation and future perspectives, Computers and Electronics in Agriculture, Volume 115, 2015, Pages 40-50,