This project has received funding from the European Union's Horizon 2022 Research & Innovation Actions - Project. 101086355 – HORIZON-CL6-2022-GOVERNANCE-01

SENSOR AND EDGE PROCESSING SELECTION, DEVELOPMENT, SPATIAL PLANNING AND DATA COLLECTION, V1

Revised version, 27 January 2024

Authors		
Name	**Organization**	**Draft release date**
Heikki Astola	VTT	12/12/2023
Alexander Kokka	VTT	12/12/2023
Matti Mõttus	VTT	12/12/2023
Valantis Tsiakos	ICCS	12/12/2023

Approval on behalf of the Executive Board		
Name	**Organization**	**Date of approval**
Valantis Tsiakos	ICCS	30/01/2024
Nick Berkvens	EV ILVO	30/01/2024

Revision records			
Version	**Date**	**Changes**	**Authors**
Draft 00	12/12/2023	Original document (draft)	All Partners
Draft 0.1	17/01/2024	Review draft	ICCS, ILVO, VTT
Issue 1.0	31/01/2024	Version 1	All Partners
Issue 1.1	27/01/2025	Revised version that takes the comments of the EC and external reviewers into account	All Partners

Acronyms and Abbreviations

Acronyms and Abbreviations	
ADC	Analogue-to-Digital Converter
AGINS	AgroInsurance International
AI	Artificial Intelligence
ATB	Institut für angewandte Systemtechnik Bremen GmbH
AUTh	Aristotle University of Thessaloniki
AVR	AVR BVBA (Belgium)
B	Biomass
BTLE	Bluetooth Low Energy
CA	Consortium Agreement
CC	Chlorophyll Content
CDC	Capacitance to Digital Converter
CNH	CNH INDUSTRIAL BELGIUM
CO2	CO2 flux
DAC	Digital-to-Analogue Converter
DES	Deimos Spain
DME	DEIMOS ENGENHARIA SA
DMK	DMK Deutsches Milchkontor GmbH
DSS	Decision Support System
EC	European Commission
EEAB	External Expert Advisory Board
EGM	Easy Global Market SAS
EO	Earth Observation
EURAC	Accademia Europea di Bolzano (Eurac Research)
EV ILVO	Eigen Vermogen van het Instituut voor Landbouw en Visserij Onderzoek
ExBo	Executive Board
FEU	Farm Europe
FPI	Fabry-Pérot interferometer
FUOTA	Firmware Update Over-The-Air
ICCS	Institute of Communication and Computer Systems
G	soil heat flux
GA	General Assembly
GDPR	General Data Protection Regulation
GPP	Gross Primary Production
H	sensible heat
HE	Homomorphic Encryption

HPC	High Performance Computing
IFAPA	Instituto Andaluz de Investigación y Formación Agraria, Pesquera y Alimentaria
InGaAs	Indium Gallium Arsenide
IoT	Internet of Things
IPR	Intellectual Property Rights
KUVA	Kuva Space Oy
LAI	Leaf Area Index
LE	Latent Heat
LUKE	Natural Resources Institute Finland
MCU	Microcontroller Unit
MIGAL	MIGAL Galilee Research Institute
ML	Machine Learning
MPC	Multi-Party Computation
MST	Management Support Team
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NP	Neuropublic SA
NPP	Net Primary Production
OEM	Original Equipment Manufacturer
OHB DS	OHB Digital Services GmbH, Bremen, Germany
PAR	Photosynthetically Active Radiation
PBMA	Poly (butyl methacrylate)
PCB	Printed Circuit Board
PEMA	Poly (ethyl methacrylate)
PET	Privacy Enhancing Technologies
PHEMA	Poly (2-hydroxyethyl methacrylate)
PIBMA	Poly (isobutyl methacrylate)
PISEC	Platform for Integrated Sensing and Edge Computing
PO	Project Officer
PSNC	Instytut Chemii Bioorganicznej Polskiej Akademii Nauk
Pt	Platinum
R&D	Research and Development
RGB	Red Green Blue
RIL	Research and Innovation Lab
Rn	net Radiation
SDK	Software Development Kit
SM	Soil Moisture

SME	Small and Mid-size Enterprise
TBC	To Be Confirmed
UAV	Unmanned Arial Vehicle
UGent	Universiteit Gent
VITO	Vlaamse Instelling voor Technologische Onderzoek
VRI IES	Foundation "Institute for Environmental Solutions"
VTT	Technical Research Centre of Finland Ltd.
WASM	WebAssembly
WODR	Wielkopolski Osrodek Doradztwa Rolniczego w Poznaniu
WP	Work Package

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Project overview

ScaleAgData is a response to the call HORIZON-CL6-2022-GOVERNANCE-01-11 Upscaling (real-time) sensor data for EU-wide monitoring of production and agri-environmental conditions. The ScaleAgData project will run from January 2023 until December 2026 and consists of a consortium of twenty-six partners from fourteen countries. The vision of ScaleAgData is two-fold. On one hand, it wants to obtain insights in how the complex data streams related to sensors and Earth Observation (EO) should be governed and organized (governance call). On the other hand, it aims to develop the data technology needed to scale data collected at the farm level to regional datasets, agri-environmental monitoring, and the management of agricultural production.

To do so, ScaleAgData has five objectives:

- Developing innovative approaches for collecting in-situ data and applying data technologies.
- Enabling and promoting data sharing along the entire data value chain.
- Demonstrating how the sensor data can be scaled to agri-environmental data products at the national, regional, or European level.
- Demonstrating the benefit of the improved monitoring capacities in a precision farming context.
- Demonstrating the benefit of upscaled regional datasets for the agricultural sector in general.

During its lifecycle, the project will explore seven innovation areas: innovative sensor technology, edge processing, data sharing architecture and data governance, satellite data augmentation, from data assimilation to service development, privacy-preserving technology, and data integration methodologies.

Six Research and Innovation Labs (RIL) have been identified within the project, across various bio-geographical regions of Europe, where different data upscaling and integration models or approaches will be evaluated and demonstrated. The six RI labs each have their own thematic focus, being Water productivity, Crop management, Yield monitoring, Soil Health, Grasslands, and Sustain Dairy. With this extensive thematic coverage, sensor data will be collected on a wide variety of agri-environmental conditions, allowing ScaleAgData to cover the four dimensions of the environment: Soil, Water, Air and Living organisms (crops and livestock in this case).

1.2 Scope of the document

This deliverable summarizes the work done in WP3 Task3.1 (later: T3.1), including the sensor selection and their data acquisition planning in ScaleAgData within the respective RILs (i.e. which data will be collected, where and when), and the development of new sensor technology and edge processing enabling technologies.

A major part of the document will be the Sensor Data Catalogue, that will record all the sensor data sets collected during the project. This catalogue is a living document that will be updated constantly along the sensor data acquisition campaigns by the RIL leaders. The Sensor Data Catalogue will be implemented as an Excel spreadsheet document, and it will be added as an annex to this deliverable.

1.3 Document structure

This document is structured as follows:

- Chapter 1 provides a project overview and then goes on to describe the scope, responsibilities, and structure of this deliverable.
- Chapter 2 presents the data acquisition planning in each RIL.
- Chapter 3 describes the development of new sensor technology.
- Chapter 4 presents the ICCS approach to optimal sensor placement.
- Chapter 5 presents the developments in edge processing technologies.
- Chapter 6 presents the sensor data catalogue structure and gives a link to the catalogue that is a separate document (Excel file).
- Chapter 7 draws the conclusions of the work in WP 3 T3.1.

1.4 Evolution of the document

Version 1.0 of this document, submitted on 31 January 2024, provides a description of the sensor selection and their data acquisition planning in ScaleAgData within the respective RILs, and the development of new sensor technology and edge processing enabling technologies.

The present version of this document, version 1.1, submitted on 27 January 2025, includes minor changes, taking the comments of the EC and external reviewers on this deliverable into account.

More details regarding the activities performed to develop the PISEC (now renamed to YaraSense) platform and sensors, have been provided in Chapter 5. The excel file with the Sensor Data Catalogue has been embedded in Chapter 6. Clarifications about the sensor data constraints and interpretation of the sensor data have been added to sections 2.2.2.5, 2.2.5.2 and 2.6.2.1. Finally, section 3.1.5 has been added to describe the status of the nanoparticle gas-sensing array as of 31 January 2024.

Version 2.0 of the document, deliverable D3.5, is foreseen for September 2025. This version will provide an update on the sensor data acquisitions, the new sensor developments and technology implementations and on the results obtained.

Additional updates will take place if necessary.

2 Sensor Data Acquisition Planning

The different data upscaling and integration models or approaches, that will be developed in ScaleAgdata, will be evaluated, and demonstrated in the six (RILs. The chapters below will present the plan for collecting the essential sensor data for the RIL at hand, listing the data sources, the data acquisition schedule and the resources needed for the actual data acquisition work.

2.1 Approach

The work in WP3 T3.1 inherits from the work done in the co-design workshops, and their outcomes, which are documented in the deliverable D2.1 Vision scenarios, requirements, and innovative governance models. WP2 produced a high-level backlog for each RIL with the list of requirements and their time-related prioritization, as well as the rolling plan for facilitating the upcoming co-operation in developing the services and product in the following work packages WP3, WP4 and WP5.

The outcomes of WP2 T2.1 are the starting point for the selection of the set of sensors needed in each RIL to fulfil the objectives and requirements obtained from the co-design phase. The RIL leaders and technology providers will verify the requirement lists and derive/refine the specifications and technical requirements and will adapt to this information in planning the data acquisition/development work. Figure 1 depicts the overall approach of sensor data acquisition planning and sensor data collection in ScaleAgdata.

This work in WP3 T3.1 will follow the principles and aspects described in D1.2 Open Science and Data Management Plan, where applicable.

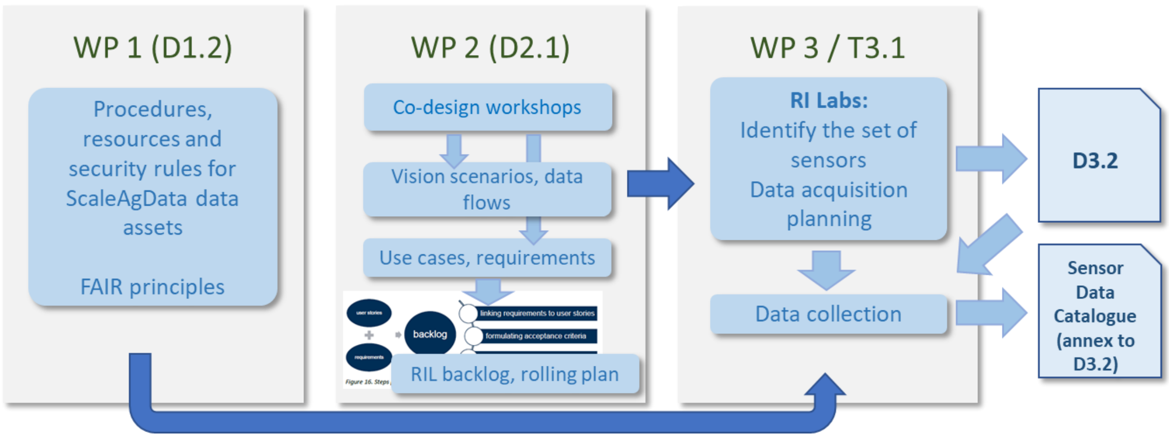

Figure 1. The overall approach of the sensor data acquisition planning and sensor data collection in ScaleAgdata

2.2 RIL Yield monitoring

2.2.1 Sensor data

The following sensors will be used in the Yield monitoring Lab:

- AVR harvester yield sensor (see 2.2.2 for more information)
- AVR harvester machine parameters (traction, fuel usage)
- CNH harvester yield sensor (see 2.2.3 for more information)
- CNH harvester machine parameters

- Weather station (for some test fields)
- Soil scanner (for some test fields)
- Crop scanner (for some test fields)
- RGB camera
- Hyperspectral scanner (from VTT)

2.2.2 AVR yield sensor

2.2.2.1 Sensor Selection

AVR has implemented a yield measurement system on its AVR potato harvester. This consists out of 2 weighing cells mounted on the top belt of the harvester connected to a signal amplifier, after which the data is sent each second to the cloud platform (AVRConnect). On this cloud platform, the yield information is visualized on a field view showing the variability of the yield measured on the field (see Figure 2).

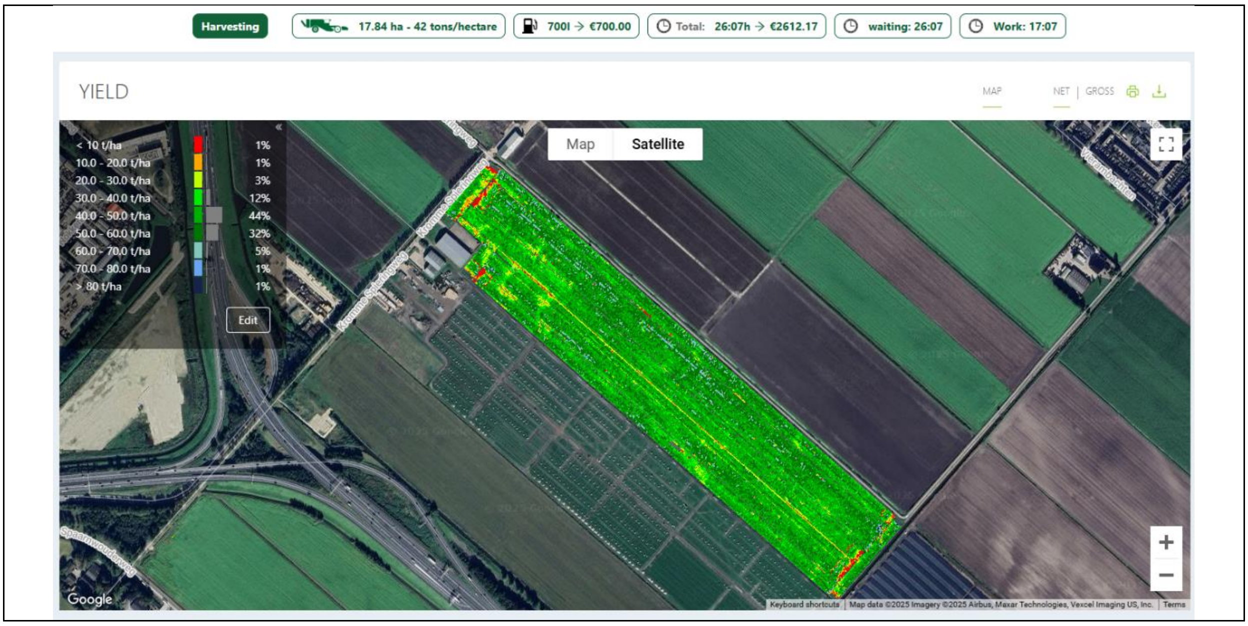

Figure 2. Yield (tons/ha) measured on the field by an AVR potato harvester equipped with a yield measurement system. Lower than average yields are shown in red, average yields in green, higher than average yields in blue.

2.2.2.2 General information

Sensor data is collected from multiple harvesters. AVR has a coverage of 10 machines in Western-Europe.

2.2.2.3 Schedule

Yield information (tons/ha) is measured every second by the harvester. This data is only collected during yield measurement on the field while the potato harvester is harvesting potatoes on the agricultural fields.

2.2.2.4 Resources

AVR has machinery in different countries, so data collection is done in the fields all over Europe.

2.2.2.5 Special data constraints

Special attention is needed to have correct data. Calibration is frequently needed to have "realistic" results. So, we need to be very conscious on the quality of the data. A common source of error, for example, arises from soil clods that resemble potatoes. Additionally, in wet conditions, the absolute yield may be higher due to soil covering the tubers.

2.2.3 CNH yield sensor

2.2.3.1 Sensor Selection

A yield sensor is available in the CNHi portfolio. Each data point is linked to a geospatial location. The data is recorded in 1 second intervals and sent in batches to the cloud (see Figure 3).

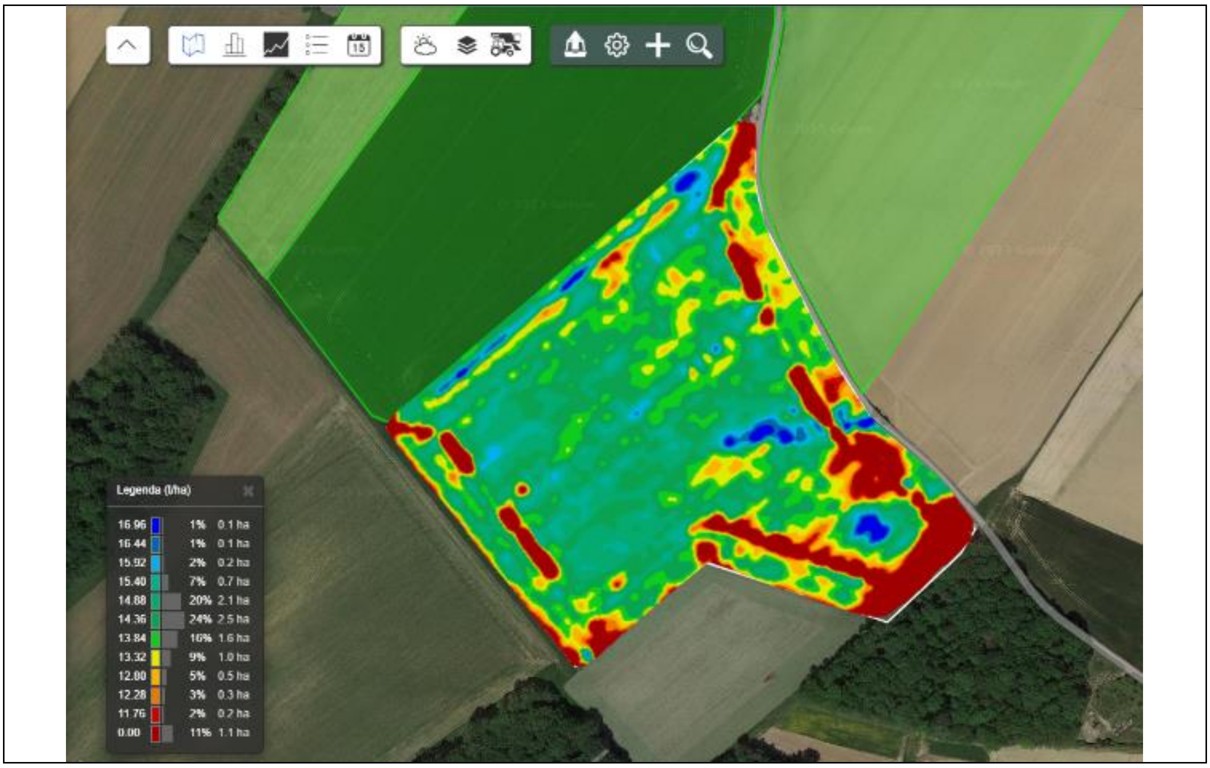

Figure 3. Yield measure in the field with CNH sensor

2.2.3.2 General information

CNH is working with a test farmer located in Braine-le-Château, Belgium. He records data on some of his 2000ha of winter wheat fields.

2.2.3.3 Schedule

Yield data is collected only during harvest time (July-August). This is highly dependent on the weather.

2.2.3.4 Resources

4 test machines collect data in this region.

2.2.3.5 Special data constraints

Data is collected in cn1 formats, this is not a directly readable format, so conversions need to be made. Data quality is also highly dependent on sensor calibration. If this is not done correctly the absolute values will be off.

2.2.4 Other field sensors

CNH's test farmer also collects the following data on selected fields:

- weather data (~10 weather stations, https://smartfarm.nl/en/sensors-en/)
- soil data (electrical conductivity at 4 depths) (https://www.topsoil-mapper.com/en/)
- crop data (augmenta scan ~ biomass index) (https://www.augmenta.ag/)

These data are collected throughout the season.

2.2.5 Sensor use cases

2.2.5.1 Tare weight estimation in potatoes (UGent in collaboration with AVR)

Currently, the tare weight of the potato harvest, i.e. the soil / dirt attached to the tubers and the harvester's conveyer belt, cannot be measured on the field. Farmers will not allow the harvester to pause to collect the potatoes on a set geolocation, to link to the other measurements.

Therefore, in a first phase, **calibration experiments** need to be performed on experimental fields where the harvester can be paused. A test experiment has been planned at the Bottelare research farm (November 2023) but has been delayed due to bad weather. In this calibration experiment, the harvester will be paused to collect the potatoes with the dirt + the dirt attached to the conveyer belt. The weight of the dirt will be linked to RGB images to see if the tare can be estimated. This experiment should be repeated for different soil types and with the hyperspectral camera of VTT.

During the **field experiments**, data will be gathered to estimate the soil moisture first, after which this soil moisture can be used to estimate the tare. Soil sensors and samples will allow a detailed calibration of the soil moisture, but the goal is to estimate the soil moisture with only weather station and EO data (for each soil type). Ideally, the soil attached to the potatoes and harvester band is collected during these field experiments as well, but as an alternative, the tare can be estimated with the ML model from the tare calibration experiment. This is why it is necessary to perform that calibration experiment for different soil types. The soil moisture and machine parameters can then be linked to the estimated tare on the field.

2.2.5.2 Improved winter wheat yield maps from CNHi harvesters (UGent in collaboration with CNHi)

Yield maps from the harvesters are still variable due to missing data, variation between harvesters, ... To fill in the gaps, EO data will be used to simulate spatial variation in yield (simulation with a digital twin or ML-based yield estimation model – in collaboration with LUKE). The spatial variation of previous years (in both EO data and harvester data) can be used to estimate variation in soil quality. The goal is to compare yield variation in a field simulated with the model and measured with the harvester sensors. If this is comparable, then the model can be used to fill in the missing data from the harvesters. However, care must be taken to distinguish the variations caused by other reasons, like seasonal variations in weather, management practices, or seed variety from the (more stable) soil quality variation itself.

2.2.5.3 Schedule

- Oct-Nov 2023: tare calibration experiment (1 soil type, RGB camera)
- Summer 2024: prepare the sensors on the fields of CNH, AVR
- Summer 2024: prepare the digital twin model for wheat.
- Summer 2024: gather EO data of previous years of CNH fields? + use for testing the digital twin model.
- Oct-Nov 2024: tare calibration experiment (more soil types, RGB + hyperspectral camera)

- Oct-Nov 2024: first soil moisture calibration and tare estimation experiment (1 soil type)

2.2.5.4 Resources

The following human resources will be necessary:

- Farm technicians: to perform the tare calibration experiments.
- Data scientists: to read and analyze all the data, to make and assess the models

Support we expect to receive:

- VITO: EO data for a few fields, to test the models
- LUKE: digital twin of wheat (crop model?)
- VTT: support for hyperspectral camera mounting, usage, data analysis
- DHI: support for soil scanner?

2.2.5.5 Special data constraints

If the tare calibration cannot be performed on different soil types, there is a risk of falsely estimating the tare, without even knowing the estimate is incorrect. So, the field measurements are only useful for the soil types where a tare calibration is available.

2.3 RIL Water

2.3.1 Data Acquisition Plan

2.3.1.1 Sensor data

Local (or regional) IoT meteorological station data:

- Precipitation
- Air temperature
- Humidity
- Wind speed and direction
- Air pressure
- Solar radiation

Local IoT soil sensors:

- Soil moisture
- Soil temperature

Local airborne data:

- Spectral data in VIS-NIR (400-1000 nm) spectral range
- Thermal data

Regional satellite data products:

- Evapotranspiration data products (expected from DHI)
- Soil moisture data products (expected from DHI)
- Vegetation and moisture indices (expected from VITO)

2.3.1.2 Schedule

Local IoT meteorological and soil sensor data will be acquired for fields of interest at least once every 30 min for the whole cropping season to ensure a near real-time data flow. Local airborne data will be acquired for fields of interest on request when the situation in the field has significantly changed as to update the spatial distribution of water status and predicted yield. Expected update frequency are once per 1-.4 weeks. Regional satellite data products will be obtained with the highest possible frequency to upscale developed models for water status assessment and yield prediction on a regional scale for fields where local IoT sensor data is not available.

2.3.1.3 Resources

IoT meteorological and soil sensors used in the study are property of project partners – IES (Latvia) and MIGAL (Israel) and will be placed in chosen test fields. Airborne data acquisition equipment is a property of project partners – IES (Latvia) and MIGAL (Israel) and will be used for data acquisition in chosen test fields. It is expected to receive support from project partners in satellite data processing:

- DHI – satellite data-based evapotranspiration and soil moisture data products
- VITO – satellite data-based vegetation and other indices

2.3.1.4 Special data constraints

Cropping seasons might vary in time in Latvia and Israel, therefore, data acquisition will be adapted to agricultural schedule.

2.4 RIL Soil

The RIL Soil will focus on using innovative hyperspectral imaging monitoring tools to improve on the contemporary products that rely on remote sensing data of multi-spectral data. The goal is to both i) provide enhanced estimation accuracy, as the more detailed hyperspectral signal can assist AI models to better estimate the soil properties by exposing more absorption bands, and ii) provide soil maps at unprecedented detail using UAV data (i.e., <1m) which is particularly important for more effective land management practices and to provide maps in orchards where the current spatial resolution (10 m) is insufficient to differentiate between tree and exposed soil.

The overall approach plans on using two forthcoming hyperspectral products, namely the VTT HSI system and the KUVA HS satellite mission of Hyperfield-1. Recognizing that there is an inherent risk on relying on future sensors whose delivery might be delayed, and to ensure that there will be adequate data for the RIL Soil, we have developed a **contingency plan** as follows:

- For UAV data, both ILVO and AUTH have their own UAV platforms and previously procured hyperspectral cameras, albeit at a lower spectral range than the VTT developed HSI. The plan is to use these devices if, for any reason, the VTT HSI fails to become operational.
- With respect to spaceborne data, in case Hyperfield-1 for any reason does not become operational or due to the limitations of the revisit time and weather conditions (e.g., consistent cloud coverage at each satellite overpass during the critical period when bare soil is visible from above), the plan is to use other HS satellite data. To this end, both ILVO and AUTH will make sure to place orders for PRISMA and/or EnMAP. It should be noted that the same limitations apply to PRISMA and EnMAP data as well.

2.4.1 Sensor Selection

Two sensors will be used to further improve existing Soil Organic Carbon (SOC) models, i.e., that predict the topsoil SOC content from visible and infrared reflectance spectra.

- VTT HSI camera
- KUVA HS satellite Hyperfield-1

Currently, multispectral EO (Sentinel 2) is used for the SOC models; with hyperspectral imagery we plan to gain more fine-grained spectral measurements and a higher spectral resolution for the agricultural soils we are monitoring.

The spaceborne data from Hyperfield-1 provide much larger coverage which is particularly suitable for annual crops which exhibit wide areas of bare soil in the critical period. The UAV data, on the other hand, are more useful to detect bare soil in orchards and tree crops (e.g., olive trees, vineyards, that are very prevalent in the Mediterranean and apple trees in Northern Europe). Therefore, we aim to use both data sources to provide maps of SOC.

In addition, we combine spectral data with datasets providing more contextual information about the soils of the agricultural fields e.g. the 'bodemassociatiekaart' of Flanders (https://www.dov.vlaanderen.be/page/bodemkaarten, Digital vectorial dataset with an overview of the occurrence and classification of soil associations in Flanders) various soil property maps provided by European Soil Data Centre (ESDAC, https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data).

2.4.2 Data Acquisition Plan

2.4.2.1 General information

VTT HSI sensor: One sensor, we plan to attach the sensor to a UAV (Matrice 600) and potentially a robotic platform (e.g., Thorvald by SAGA), the sensor should be ready for experiments in 2024. The sensor will be transferred to both ILVO and AUTH to conduct the experiments.

KUVA HS satellite Hyperfield-1: current plans are the data should be available for the whole of Flanders and Greece by the end of 2024/beginning 2025.

2.4.2.2 Schedule

Bare soil is needed on the agricultural fields to be able to use satellite/sensor images to predict the topsoil SOC (VTT HSI sensor).
In Flanders we plan to perform experiments on potatoes fields before harvest (March-May) or after harvest (Aug-Sep), in Greece we plan to perform experiments in tree crops (e.g., grapes, olives, apricot, peaches) in which case bare soil may be identified throughout the year (in best cases). Experiments will be on fields of 1-2ha.
- KUVA HS satellite: results of an already executed soil campaign collecting and analyzing soil samples throughout the whole of Flanders will be used, subsequently satellite imagery for the whole of Flanders and available throughout the whole year will be used (we will use as many satellite images as available and suitable for spectral data related to bare soils).

2.4.2.3 Resources

The following human resources will be necessary:
Farm technicians, UAV and robotic engineers, data engineers, GIS scientists and data scientists.
We expect to receive support from project partners VTT and KUVA in satellite data processing.

2.4.2.4 Special data constraints

The weather conditions must be good for UAV-flights and capturing data with the VTT sensor, in addition clouds prevent KUVA HSI satellite to capture spectral measurements of the agricultural fields.

2.5 RIL Grasslands

Monitoring key biophysical parameters and exchange processes in grasslands allows for a comprehensive assessment of grassland ecosystem health. Changes in LAI, PAR, biomass, chlorophyll content, soil moisture, and CO2, energy and water fluxes can indicate stress, environmental changes, or the impact of management practices. The data obtained from these sensors can inform land managers and researchers about the need for adjustments in land management practices, such as altering grazing patterns, adjusting irrigation schedules, or addressing nutrient deficiencies. In the study areas of the RIL Grasslands we focus on the following parameters (see Table 1):

Table 1. The key biophysical parameters selected for RIL Grasslands

Parameter	Importance	Use
Leaf Area Index (LAI)	LAI is significant for understanding the density and vigor of vegetation.	LAI helps estimate the amount of solar radiation intercepted by the vegetation, which is critical for understanding photosynthesis, evapotranspiration, and overall plant health.
Photosynthetically Active Radiation (PAR)	PAR represents the portion of sunlight that plants use for photosynthesis. It is crucial for understanding the energy available for plant growth.	It helps assess the efficiency of light capture by the vegetation, which, in turn, relates to plant growth, productivity, and ecosystem health.
Chlorophyll Content (CC)	Chlorophyll is a pigment crucial for photosynthesis, and its content is indicative of the plant's ability to convert light into energy.	provides information about the physiological status of plants. It helps assess plant stress, nutrient deficiencies, and overall plant health.
Soil Moisture (SM)	SM is critical for plant growth as it directly affects water availability to roots.	It enables efficient water use and preventing over-irrigation or drought stress.
Biomass (B)	Essential for understanding plant growing patterns and productivity.	Provides information about the growth of the vegetation and the aboveground Net Primary Production (NPP)
CO2 flux (CO2)	Important for understanding plant growth patterns and productivity	Provides information about the growth of the vegetation and the Gross Primary Production (GPP)
Energy fluxes: latent heat (LE), sensible heat (H, net radiation (Rn) and soil heat flux (G)	It is required to assess water use and system functioning	It enables the measurement of the energy balance of the surface and plant water consumption
Meteorological data (Rad, Ta, u, DPV, P)	Required to evaluate the influence of climate on plant growth and water status	Data input to growth and water models

2.5.1 Sensor Selection

The planned sensors are shown in Table 2.

Table 2. The sensors to be used in RIL Grasslands

Sensor	Producer/Model	Location	Sampling freq.	Repository
LAI	Licor, LAI 2200C	South Tyrol, Trento, Italy	Bi-weekly	DOI
PAR	Apogee, MQ-301X	South Tyrol, Trento, Italy	Bi-weekly	DOI
CC	Konica 502, SPAD 502	South Tyrol, Trento, Italy	Bi-weekly	DOI
SM	Campell, HydroSense II	South Tyrol, Trento, Italy	Bi-weekly	DOI
LAI	Licor, LAI 2200C	Cordoba, Spain	Monthly	DOI
fPAR	Accupar LP-80 Ceptometer	Cordoba, Spain	Monthly	DOI
B	Grasmasster Pro, Jenquip EC20 Platemeter	Cordoba, Spain	Monthly	DOI
CO2, H, LE	*Gas analyzers:* (1) LICOR 7500, (1) LI-7500DS, (2) Cambell IRGASON *3D sonic anemometer:* (1) Campbell CSAT, (1) Gill Windmaster	Cordoba, Spain	Continuous (half hour)	DOI
Rn, G	Rn: (1) Kipp&Zonen CNR4, (3) Campbell NR01 G: Hukseflux HFP01 (6 measurement point), (6) TCAV	Cordoba, Spain	Continuous (half hour)	DOI
SM	(7 probes,3depths) Sentek Environscan (3) Stevens Hydraprobe	Cordoba, Spain	Continuous (half hour)	DOI
Rad, Ta, u, DVP, P	(1) Campbell LP02, (3) Vaisala HMP155, (1) Vaisala45A, (1) Campbell, TR525, (2) Campbell ARG100	Cordoba, Spain	Continuous (half hour)	DOI

2.5.2 Data Acquisition Plan

2.5.2.1 General information

The data is acquired in two study areas located in Italy and Spain. Eight different alpine meadows have been selected in the Provinces of Trentino and South Tyrol, North Italy. Ten Mediterranean pasture fields, corresponding to five oak savanna farms, were chosen in Córdoba, southern Spain (see Figure 4).

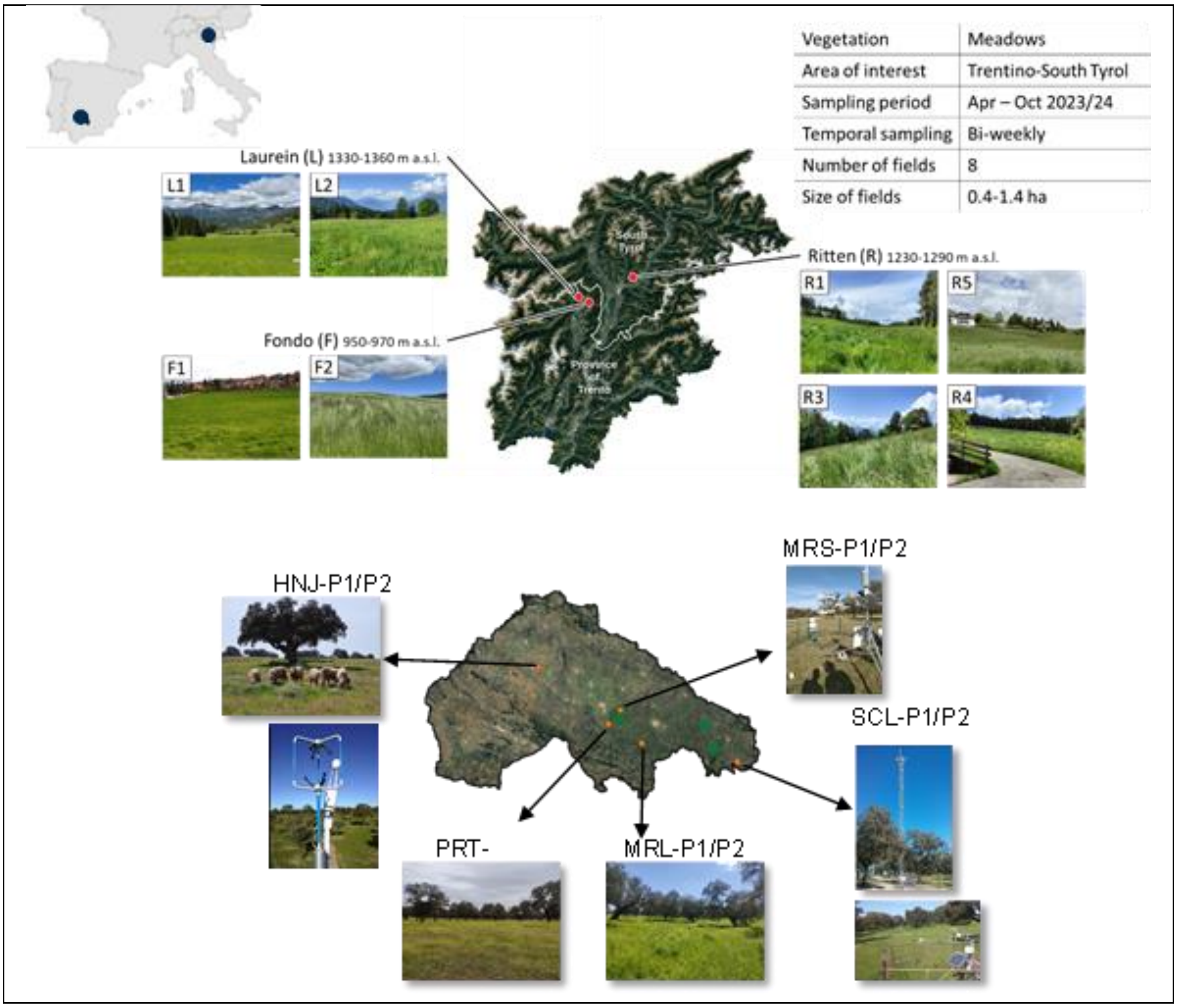

Figure 4. RIL Grasslands test sites

2.5.2.2 Schedule

The sampling frequency in Italy is every two weeks. Taking measurements over different pixel distributed area that corresponds to the spatial resolution of Sentinel2, to later assimilate such information. Apart from data collected from sensors, the height and composition are registered using a quadrat technique, reporting conditions (e.g. lodging, fertilization) and collection biomass. A similar procedure is followed in Spain with a monthly frequency. Information about grazing activities is also collected.

Figure 5. Two weeksling strategy and Data collected

2.5.2.3 Resources

Instrumentation described in Table 2. Two trained persons in the use of instruments and field data collection.

2.5.2.4 Special data constraints

If the day(s) before the data campaign was raining, the sampling is shifted for a few days to avoid water saturation in biomass and SM. In Spain, the duration and calendar of the growing cycle varies each year depending on climatic conditions, mainly rainfall distribution. Production can partly or entirely be lost during drought events (e.g. 2023). In those cases, the data collection period must be extended (for example, adding 2025 to complete two seasons).

2.6 RIL Dairy

2.6.1 Sensor Selection

The main objective of the dairy lab is to investigate potentials for improving the forecasting of data concerning milk quality and quantity, which is relevant for the production facilities of the involved dairy cooperative. If this will be feasible, there might be future potentials to provide additional feedback to farmers.

The lab aims at combining the use of the following sensors/ data sources:

- Sentinel Data 10m archive
- Milk samples and subsequent laboratory devices/analysis
- Sensors installed in forage harvesters

The Sentinel data is currently considered as a basis to analyze the NDVI. It is also planned to use further EO-data-based services that provide information about the available biomass. The combination of in-situ sensor data with EO data shall be further analyzed, also deriving further requirements for the refinement of the sensor selection.

Since the dairy cooperative aims at covering a large land area, due to their 20 production facilities and 4,700 milk producers it is targeted at a combination of in-situ data with EO data, where the in-situ data will serve for a calibration of EO data-based analysis of yields.

The list of sensors could grow as the project progresses, including operational satellite or institutional data where the data products contribute to the correlation.

2.6.2 Data Acquisition Plan

2.6.2.1 General information

All milk that is delivered to the dairy cooperative is analyzed by taking samples from each milk delivery of each farmer. This is generally done for product safety as well as for determining the specific milk parameters that are also defining the revenue of farmers. The milk samples are taken at the farm when the milk is put into the milk transporter. Specific laboratories are used to perform the related analysis. Therefore, several milk samples per week are taken for each of the 4,700 farmers in the overall area covered by the dairy cooperative in Germany, Italy and the Netherlands.

In the first phase of the project, the dairy lab is investigating the available EO data that could be used for understanding the development of the vegetation in the area of interest (using e.g. NDVI or possibly also the NDMI) over the year. This work will also analyze the amount of data available (i.e. revisit frequency and resolution).

The in-situ data measured during harvest on the field shall provide details about the e.g. harvested grass or maize. The yield per hectare and the percentage of dry matter is measured by the farm equipment. This was used to analyze the amount of available feed over time in specific geographical areas. In the project, a specific aim is to collect aggregated data that can be accessed via data intermediaries that are anonymizing the data, thus avoiding the need for detailed interaction with individual farmers. Therefore, additional parameters like ADF and NDF can not be taken into account in this project phase.

In a later phase, it will be further analyzed how to consider additional parameters from reference farms that could be used for further validation of the results that are identified, based on the analysis of the aggregated data, aiming to minimize the burden for farmers accordingly.

2.6.2.2 Schedule

In the first two years of the project, the aim is to use historic data available for the last 5 years to analyze potential correlations of EO data and in-situ data. If the correlation analysis will lead to positive results, further details and optimization potentials will be investigated in cooperation with the milk quantity and milk quality forecasting department as well as the production facilities in the second two years' cycle of the project.

2.6.2.3 Resources

The basic EO data analysis is based on available Sentinel data that can be accessed for the analyzed regions for free. However, a more detailed and fine-grained analysis might require the usage of paid services to facilitate real-time access to required EO data.

Based on initial discussions, there is also a potential to receive support from the project partner VITO with respect to satellite data processing, considering VITO's satellite-data-based vegetation and other indices.

The in-situ data shall be collected from forage harvesters that are widely used by most farmers in the area concerned. The required data is collected by the telemetry of the forage harvesters and collected by the OEM in a proprietary cloud-based system. There will be efforts required to access related data, as well as potential costs to acquire the data accordingly.

2.6.2.4 Special Data Constraints

The satellite data might not be available in the expected frequency, due to cloud cover. It needs to be analyzed if this might represent a specific constraint for certain regions. At the same time, due to specific limitations of data protection, there is no exhaustive list of farmers' field parcels that could serve for a one-to-one analysis of the fields based on EO data. Hence, the analysis needs to consider the provision of average results for each region.

The milk quality and quantity data are available in high detail at the dairy cooperative, covering every sample of every farmer for the related delivery days. However, a related analysis needs to consider requirements for data protection by aggregating data. It shall be analyzed how to balance the required level of detail and the required anonymization of data.

The data of forage harvesters needs to be aggregated to enable the anonymization of data if they are provided by the OEM. This will require further work to understand the needed level of aggregation of data as well as the level of detail that is required to enable a meaningful analysis. The lab will also analyze conditions for experimental use for research purposes as well as future potentials for commercial exploitation. The analysis of the vegetation will also have to be aligned with the related cropping seasons accordingly.

2.7 RIL Crop management – Sub lab Agri-environmental monitoring for Policy Makers

2.7.1 Sensor Selection

The focus of this RIL is crop monitoring and management through the integration of IoT meteorological station data, EO data, farm log data and field observations. The sensors that will be utilized are IoT weather stations that provide measurements on an hourly basis regarding the following parameters listed in Table 3.

Table 3. The parameters measured in RIL Crop management (Sub lab NP)

Atmospheric Parameters	Measurement Unit
Air temperature	◦ C
Precipitation	mm
Relative Humidity	%
Wind speed	km/h
Wind direction	°
Air pressure	hPa
Solar radiation	W/m2
Soil Parameters	**Measurement Unit**
Soil temperature (in depths of 10, 20, 30, 40, 50, 60 and 70 cm)	◦ C
Soil moisture (in depths of 10, 20, 30, 40, 50, 60 and 70 cm)	%
Soil salinity (in depths of 10, 20, 30, 40, 50, 60 and 70 cm)	dS/m
Vegetation Parameters	**Measurement Unit**
Foliage temperature	◦ C
Leaf wetness	n/a
Foliage humidity	%

2.7.2 Data Acquisition Plan

2.7.2.1 General information

In this RIL, cotton, tomato and potato fields located in the pilot areas of Thessaly and Crete will be monitored. During the growing period of 2023, 8 IoT sensors were installed in selected parcels of the pilot areas and were providing historical data concerning the prevailing field conditions. After the end of the growing period, in October 2023, the IoT stations were uninstalled.

Due to crop rotation practices, these parcels may not be cultivated with the same crops, so parcels with similar characteristics will be selected for the growing period of 2024, when the first iteration of the RIL will take place. The exact identity of the pilot parcels will be defined during the beginning of spring 2024, and immediately after that, the IoT stations will be installed on the filed again.

2.7.2.2 Schedule

The IoT stations are installed on the field during the growing period, so sensor data collection is taking place during the growing season, which includes the period from March to October.

2.7.2.3 Resources

It is important that the connectivity of the sensors is stable, and that data collected are safely transferred and stored for further processing.

2.7.2.4 Special data constraints

The inspection of the IoT sensors on a regular basis is essential to ensure that they are fully functionable, and no malfunctions are occurring. Especially during extreme weather events it is critical to ensure that the sensors have not been negatively affected.

2.8 RIL Crop management – Sub lab Sustainability performance

2.8.1 Sensor Selection

The need of this RI Sub lab is to have weather data that can feed models and rules which are present in the various functionalities of the Horta's Decision Support System grano.net®. This Decision Support System (DSS) is an online tool that supports farmers in the sustainable management of their wheat crop, providing advice on the main crop operations to be performed in field (i.e. sowing, fertilization, crop protection interventions, yield forecasts, …). Several kinds of inputs are used by the models and algorithms in the DSS, such as weather data, soil data, crop operations data, remote sensing data. Weather data are collected by means of in-situ weather stations, which have sensors to register air temperature, rain amount, relative humidity, and leaf wetness.

2.8.2 Data Acquisition Plan

2.8.2.1 General information

Horta will monitor 10 wheat fields in the North of Italy. Each field will be associated to a representative weather station, which will collect data of temperature, relative humidity, ari and leaf wetness. A total of 10 weather stations will then be used in the RI Sub Lab (see Figure 6).

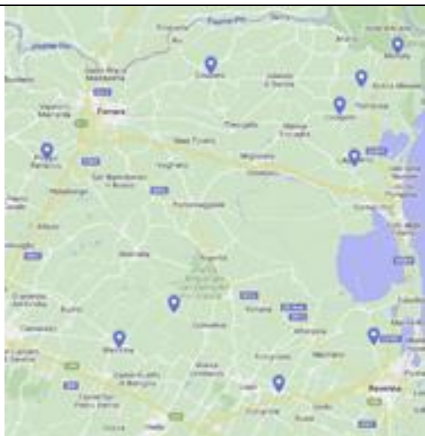

Figure 6. Location of weather stations in the North of Italy monitored in the RIL Crop management (Sub Lab Sustainability performance)

2.8.2.2 Schedule

Weather stations considered in this RI Sub lab are already installed. Wheat cropping season in north Italy start in October and ends in July, data relevant for the crop will then be collected in these months. The first season considered is the 2023-24 cropping season.

2.8.2.3 Resources

The data flow for data collection from weather station is automatized and is up and running. Resources are needed to maintain the data flow, and to fix eventual problems (sensor failing, anomalous data are automatically corrected).

2.8.2.4 Special data constraints

Additional data expected to be used in this RI Sub lab are Sentinel-2 data, which are being retrieved from a commercial provider. In addition to this, this RI Sub lab has enquired with VITO and DHI for their data provision, which is expected to be further discussed.

2.9 RIL Crop management – Sub lab Early Pest Detection

2.9.1 Sensor Selection

The system for organizing agrophage signaling in public agricultural advisory centers working within the eDWIN Advisory Platform would be expanded to include algorithms for optimizing data collection. Various data sets will be used, based on analysis of which eDWIN will suggest to observers and those organizing the observation system the optimal observation dates for specific agrophages. The result of the introduced changes will be the carrying out of observations at times of increased risk of occurrence of given agrophages in a given area.

Data from the following sources including sensors will be analyzed:

- Meteorological data from the public network of agro-meteorological stations equipped with the following sensors:
  - Air temperature
  - Humidity
  - Wind speed and direction
  - Amount of precipitation
  - Value of leaf wetting
- Observation and phenological stations – these stations are equipped with PTZ cameras and cloud software allowing remote management and control of actions. The stations allow observation of fields for the presence of disease and pests on them, and are able to accurately track the state of plant development.
- Agrophage mathematical models and its results based on field data (entered by farmers and advisors in eDWIN system) and meteorological station network.
- NDVI data (value) for the smallest possible area, reduced to the point and it trend line in time
- Structured field observations collected by advisors and farmers within the eDWIN system as e-service.

2.9.2 Data Acquisition Plan

2.9.2.1 General information

We will use about 100 meteorological stations from a network of nearly 600 stations installed all over Poland (see Figure 7). The stations used will be in the Wielkopolska region. The number of fields and plants to be used in the analysis will depend on the order of the regional agrophage observation coordinator. Usually during the growing season, it is about 100 – 120 fields (observation points).

Figure 7. The meteorological stations used in RIL Crop management (Sub lab Early Pest Detection)

Eight phenological stations will also be used for automatic observations. These stations are in the Wielkopolska region as well. Some of the stations may be located in fields where, in addition, manual validation of observations by a qualified agricultural advisor will be caried out.

2.9.2.2 Schedule

The methodology for conducting field observations has been developed and has been operating effectively since 2020 in eDWIN, and much earlier outside of it (one of the elements of the eDWIN implementation was the optimization of the process of planning and execution of field observations and integration with the national agrofagi.com.pl system). Data will continue to be collected and used for the acquisition optimization and scale process for future years.

The methodology consists of:

1. Designating individuals for monitoring (the coordinator does this based on knowledge of the advisor's place of work)
2. Designation of the crops that a particular advisor will monitor (the coordinator does this based on knowledge of the crops in the area),
3. Indication of the exact points where the observations will be carried out (the advisor does this in a dedicated system, enters, e.g. coordinates and data on the crop)
4. Implementation of observations at the designated point through a dedicated tool
5. Notification of threats to farmers in areas where agrophages have been confirmed.

In parallel to this process, we will conduct an analysis of the possibility of optimizing the planning and execution of field observations based on the integration of observation data with data from meteorological stations, observation stations and NDVI.

2.9.2.3 Resources

Most of the resources are provided by partner (PSNC), i.e. computing power, data space.

Data is being collected on server infrastructure using continuous aggregation services. Most repositories have an extended retention policy, using disk resource scaling, maintaining the entire history of measurements while regularly creating data backups. Complex data acquisition services are used from various sources simultaneously - FTP servers, API interfaces, reading directly from the database. The data has registers describing point data sources and a series of sensor sets. Access to the data is done by selecting the appropriate measuring point in the interface intended for reading. Data from various sources are unified and described with as much metadata as possible.

There were first algorithms of data validation introduced, with further potential to extend scope data quality supervision. Technologies used: REST API, Kafka streams, Logstash, time series DB, virtualization, containerization, microservices.

2.9.2.4 Special data constraints

To the data we already have, the methodology for acquiring NDVI data at specific points where agrophage occurrence data will be collected (crop fields) must be developed and positioned. Also, for historical field observations, for which we have e.g. meteorological data and mathematically modelled threat data, it will be necessary to acquire NDVI data.

3 Sensor Development

3.1 Nanoparticle Gas-Sensing Array

The development of nanoparticle gas sensors in RIL Crop management (Sub lab Agri-environmental monitoring for Policy Makers) have the potential to provide valuable information through the detection of pesticides in crops and the estimation of pest application on the fields.

3.1.1 Chemical sensors

Chemical sensors convert a chemical input signal into a measurable electrical output signal and are widely used in biomedicine and environmental sciences. Also, they are often used for the detection of potentially toxic substances. They consist of two functional parts: an identification medium that selectively reacts with the measured chemical substance and a transducer medium that converts the reaction of the first medium into an electrical signal. The performance of the chemical sensors is judged by their response time, signal-to-noise ratio, selectivity, and detection limits, among others. All these factors depend on the two functional parts already mentioned, i.e. the identification and the transducer medium. Therefore, the construction of efficient chemical sensors is directly linked to the development of materials that have good recognition and mutagenic properties. Nanoparticles can perform both recognition and mutation roles and are widely used for the development of chemical sensors.

3.1.2 Electrochemical sensors

Electrochemical sensors are a class of chemical sensors whose output signal is expressed by a change in their capacitance or resistance. This is achieved as their sensing medium is a dielectric material (membrane) between the mutating medium, which is conductive, effectively creating electrodes that are separated by a dielectric surface.

In the case of electrochemical resistance sensors, a predetermined voltage is applied to the electrodes. The sensing membrane reacts to the chemical signal by changing its conductivity resulting in a change in the measured resistance in the circuit. In the case where resistive sensors incorporate nanoparticles, their conductivity is achieved by coating the electrodes with two-dimensional or three-dimensional films of metallic nanoparticles. The sensor film can be polymer, ions, surfactants, or biomolecules and surrounds the nanoparticles and electrodes. The sensing membrane responds to the chemical signal by changing its morphology, which results in a change in the distance between the nanoparticles and thus their resistance/conductivity. Thus, the sensor responds to the chemical stimulus by changing its resistance in a constant potential difference circuit.

3.1.3 Description of the sensing array to be deployed

The sensing array that will be used by in RIL Crop management (Sub lab Agri-environmental monitoring for Policy Makers) will have as a goal to discern between four widely used pesticides and humidity, which is a gas that is commonly present in real conditions of the fields. The sensing array consists of gold electrodes on which two-dimensional conductive films of platinum nanoparticles have been deposited. On top of the nanoparticles layer there are four polymer layers which are sensitive to volatile organic compounds and will be used as gas-sensitive layer. The transducing layer absorbs gases that are present in short vicinity whereas each one of the four polymer layers are susceptible of different gases. The absorption of the gases leads to the swelling of the four polymer layers. As a result, the nanoparticle layer is deformed and the mean value of the inter nanoparticle distance is increased. This leads to an increase in the resistance of the sensor due to its exposure to the pesticides.

The sensing array is integrated as an autonomous data source within the IoT station. Data collection on pesticides detection are initially locally processed and then transferred to the cloud data-repository for integration and further processing. Figure 8 provides an illustration on the integration of the sensing array with the gaiasense IoT gateway. More information for the gaiasense IoT station is available in section 2.7 RIL Crop management.

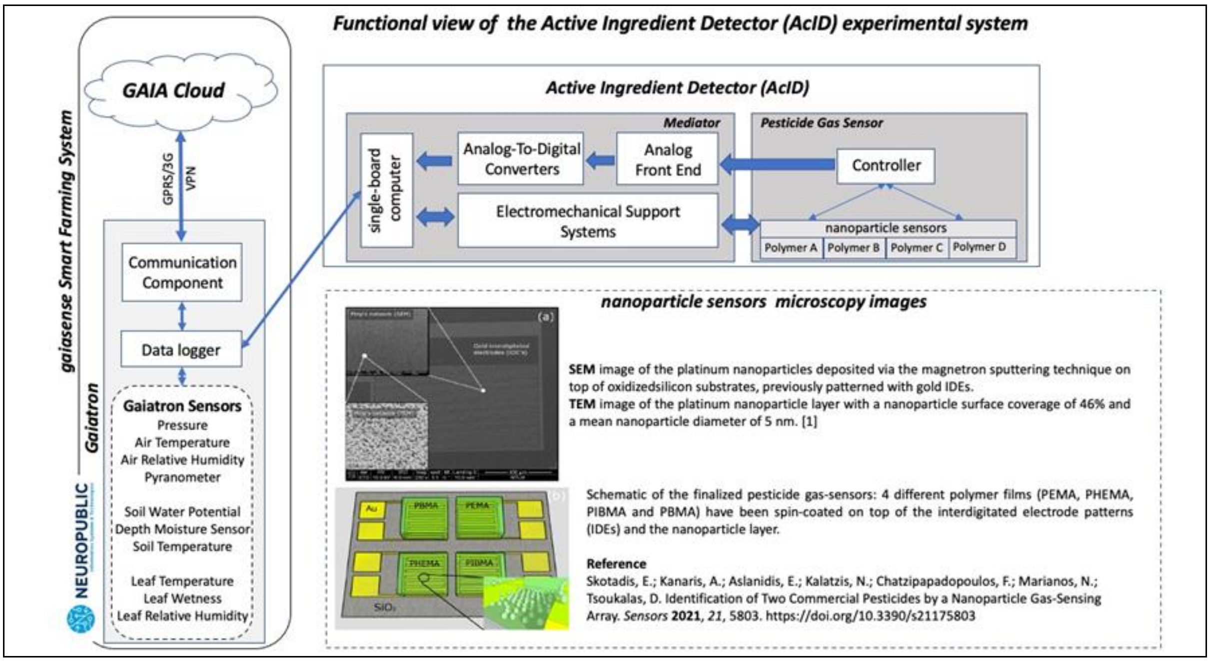

Figure 8. Integration of the nanoparticles sensing array with the IoT gateway.

3.1.4 Fabrication of the Sensors

As already mentioned, the array of sensors is based on the combination of the polymer films and platinum nanoparticles. Their fabrication is achieved through the following steps:

- On top of 10 cm diameter oxidized silicon substrate, gold interdigitated electrodes are produced through the e-gun technique and lift-off techniques.
- The wafers are broken down into wafers which contain 4 individual sensors which form a sensor array.
- The platinum nanoparticles, which have a diameter of 5nm, are deposited on individual wafers.
- The nanoparticles are produced and deposited on the electrodes, through the DC sputtering technique.
- Four different polymers are spin coated on top of the nanoparticle substrates and a polymer layer was obtained with a thickness of 500 nm.

Each sensor array/chip consists of a total of four individual sensors with the same polymer film. For each of the four different polymeric films, a total of two sensor arrays (eight individual sensors) were fabricated and characterized (Figure 8) [1].

[1] Skotadis, E., Kanaris, A., Aslanidis, E., Michalis, P., Kalatzis, N., Chatzipapadopoulos, F., ... & Tsoukalas, D. (2020). A sensing approach for automated and real-time pesticide detection in the scope of smart-farming. *Computers and Electronics in Agriculture*, *178*, 105759.

The arrays of sensors that will be developed with films from the following four different polymers are:

- Poly (ethyl methacrylate) (PEMA)
- Poly (2-hydroxyethyl methacrylate) (PHEMA)
- Poly (isobutyl methacrylate) (PIBMA)
- Poly (butyl methacrylate) (PBMA)

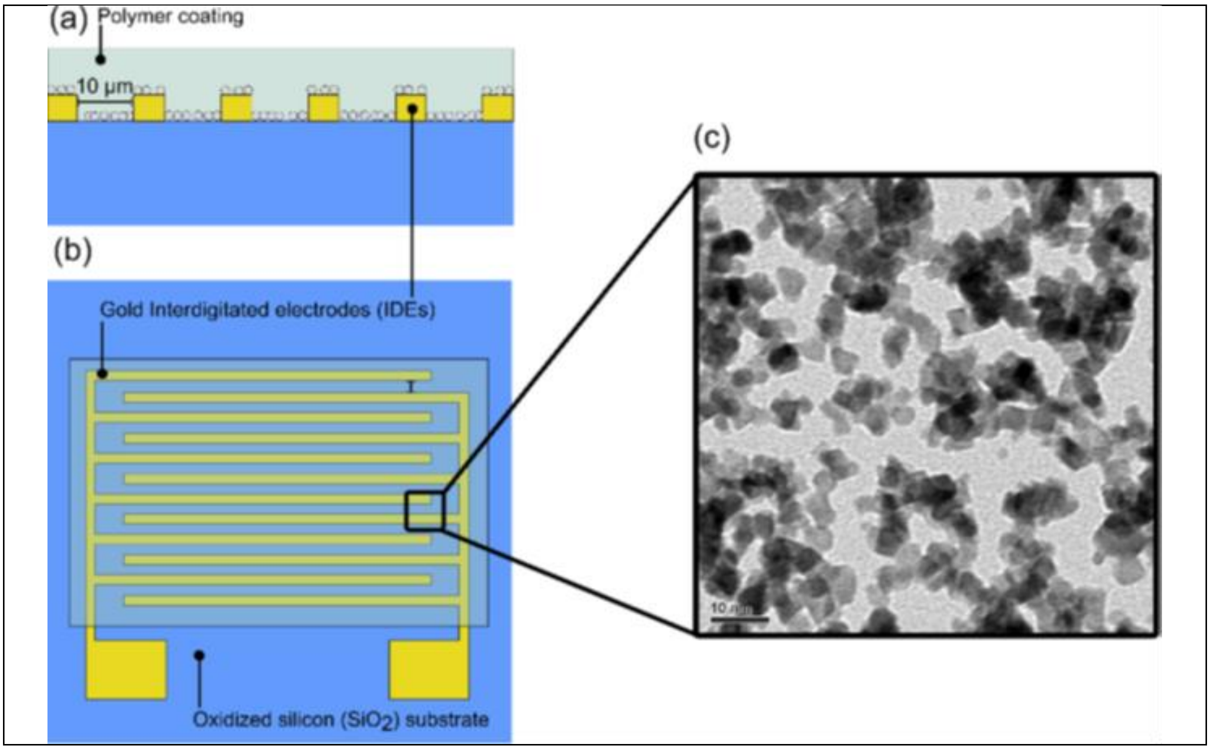

Figure 9. Sensor schematic and microscopy images (a) cross section of the sensing device (b) top-down view of the sensor (c) transmission electron microscopy image, the Pt nanoparticle distribution and size can be seen (source: Skotadis et al., 2020)

3.1.5 Current status

As of 31 January 2024, the sensor array is fully developed and ready for installation. Its deployment has been scheduled to coincide with the period of pesticide application in tomato cultivation, which typically occurs during the summer months. This strategic timing ensures that the sensor array is operational during a critical phase for data collection on pesticide detection and application monitoring.

3.2 Hyperspectral Fabry-Pérot interferometer camera

3.2.1 Overview

For developing a hyperspectral sensor for this project, the main objectives were set to be compatibility with airborne measurements with drones and fast data acquisition. The acquisition speed requirement, in turn, demands powerful light gathering optics. To enable this, a new type of Fabry-

Pérot interferometer (FPI) filter, with a large clear aperture, has been developed. The hyperspectral imager designed and built in this study is shown in Figure 10 below.

Figure 10: Hyperspectral imager developed in this project.

In addition to an FPI as the adjustable bandpass filter, the main components of the imager are:

- Industrial camera module with an InGaAs sensor
- Control electronics and low-pressure housing for the FPI
- Off-the-shelf commercial camera lens with static long-pass and short-pass filters

The main specifications of the imager are presented in Table 4.

Table 4. Key specifications of the hyperspectral instrument

Parameter	Value
Pixel resolution	1280 x 1024
Pixel size	5 µm
Field of view	7.4° x 5.9°
Wavelength range	650–950 nm
Spectral bandwidth	20–25 nm
FPI aperture diameter	43 mm
Focal length	50 mm
f/N	1.4
Mass	1.9 kg

3.2.2 Fabry-Pérot interferometer

To increase the ability of the instrument to collect signal, a large FPI filter prototype has been developed. The filter is based on two 43-mm clear aperture metal mirrors separated by a small gap.

By adjusting this distance between the mirrors, the wavelength of the transmitted radiation can be tuned. The size of the gap is altered by varying the control voltage applied to the three piezoelectric elements attached to the mirror substrates. The structure of the FPI developed for the imager is schematically illustrated in Figure 11.

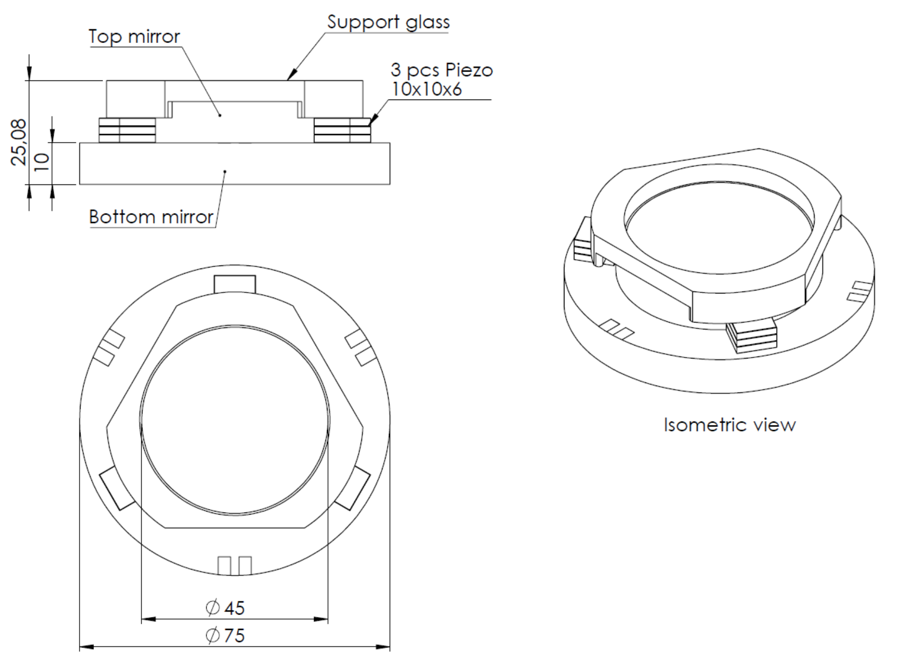

Figure 11. The simplified schematic of the FPI filter.

The wavelength λ transmitted by the FPI depends on the gap according to the equation $\lambda = 2d / n$, where d is the distance between the mirrors and n is a positive integer indicating the order of the constructive interference.

The atmospheric pressure resists and damps the movement of the mirror substrates required for adjusting the distance between the mirrors. This damping effect is especially significant when reducing the gap, as the air pressure between the mirrors rises.

3.2.3 Electronics

The distance of the mirrors of the FPI is continuously controlled by measuring the capacitance of the electrodes etched on the surface of the mirror substrates. The capacitance measurement feedback loop is also used to maintain the mirror surfaces parallel to each other to keep the gap constant throughout the whole mirror area.

The FPI controller electronics are implemented as a ring-shaped printed circuit board (PCB) around the FPI. The controller is installed in the same low-pressure chamber with the FPI to make the disturbance-sensitive connections between the controller and the FPI as short as possible. This solution also reduces the number of signals required to be fed through the low-pressure chamber.

In addition to controlling the FPI, the electronics also designed to trigger the image acquisition, once the FPI has settled on the desired gap, and to actuate the mirrors to the next gap once the exposure of each frame is ready.

The key components of the controller implementation are capacitance to digital converter (CDC), microcontroller unit (MCU), piezo drivers with integrated high-voltage boost-converter. The implementation also contains a digital-to-analogue converter (DAC) to generate control signals for piezo drivers, an analogue-to-digital converter (ADC) to convert piezo voltages to digital format for the MCU and environmental sensors to measure the temperature and pressure inside the FPI housing. The block diagram of the implementation is shown in Figure 12.

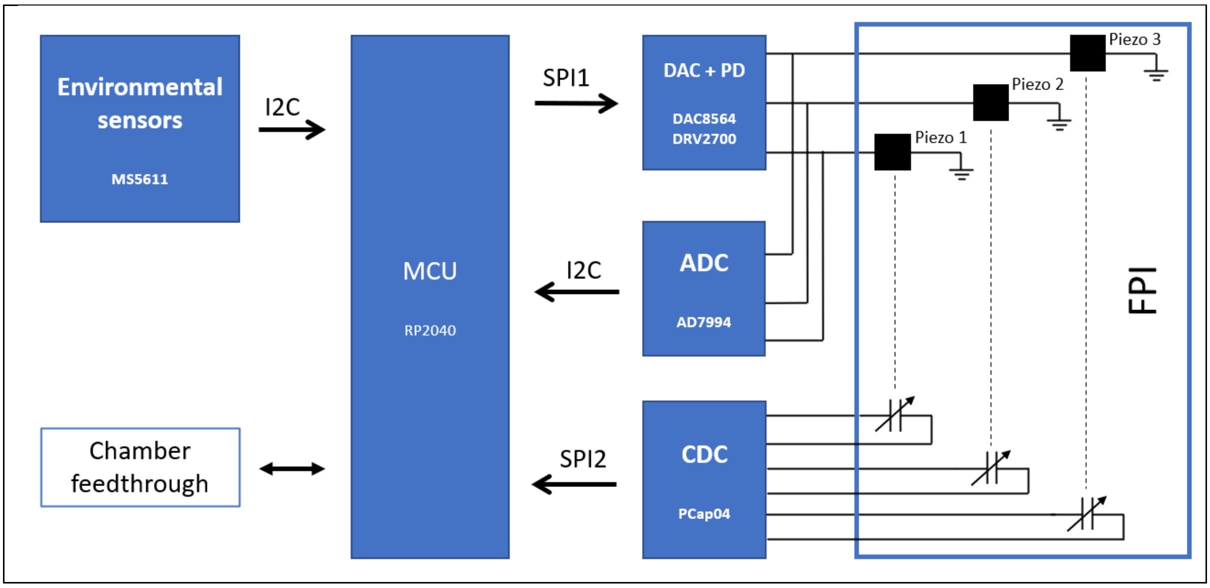

Figure 12. Block diagram of the implemented controller. Bus arrows present directions of data transfer in normal operation conditions.

The PCB layout was implemented in a ring-shaped PCB with a 100 mm outer diameter and 66 mm inner diameter. The shape of the PCB posed a challenge for isolating the switching mode voltage controller noise from the sensitive capacitance measurement. To minimize the disturbances, capacitance measurement circuit and switching mode voltage converters are laid on the opposite sides of the PCB, and ground fills were added on the upper and lower copper layer with capacitance measurement tracks. The layout of the FPI control PCB is shown in Figure 13.

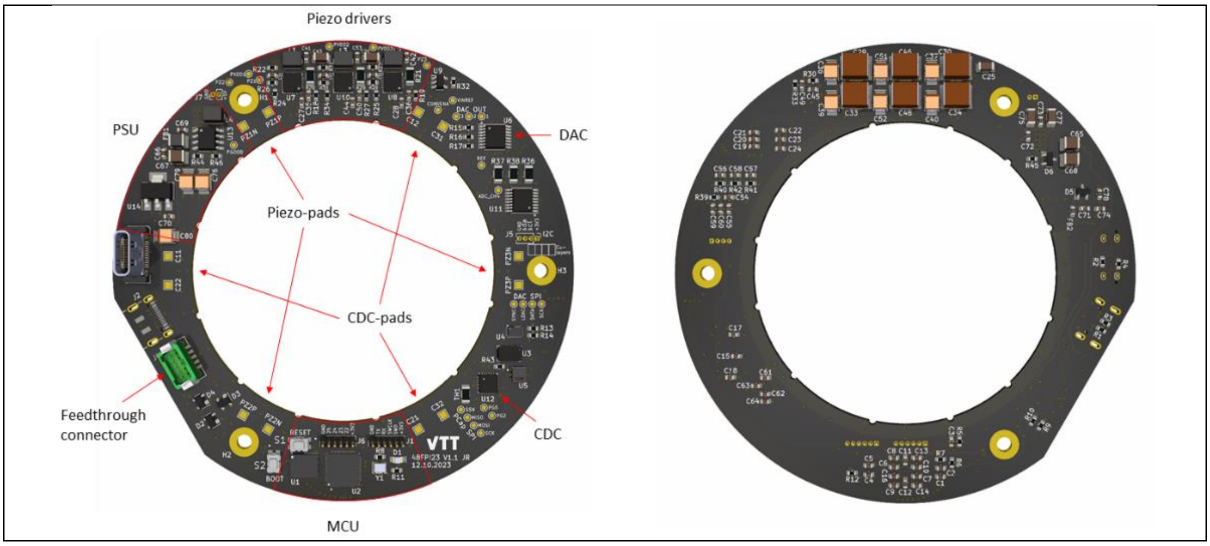

Figure 13. Layout overview of the implemented FPI-controller.

The firmware of the controller is written in C-language. The firmware utilizes both cores of the MCU, using one for execution of the gap control loop and tasks synchronized with it. The other core is used for handling serial communication with the master device and executing other tasks related to received serial commands.

The gap controller is a digital PI controller. The error signal for the controller is filtered with a moving median filter of adjustable length for both proportional and integral terms. The execution frequency of the control loop is approximately 8 kHz. The controller employs variable controlling parameters for the settled and dynamic states.

3.2.4 Optics and mechanics

To reduce the impact of air pressure on the actuation speed of the FPI, the filter is enclosed in a low-pressure housing. A short-pass and a long-pass filters installed on the optical axis on both sides of the housing act as the windows of the chamber. A ten-pin connector is used to feedthrough the supply power and electrical signals of the FPI, and a one-way valve is used to create, upkeep, and depressurize the vacuum inside the chamber. The mechanical design schematic is shown in Figure 14.

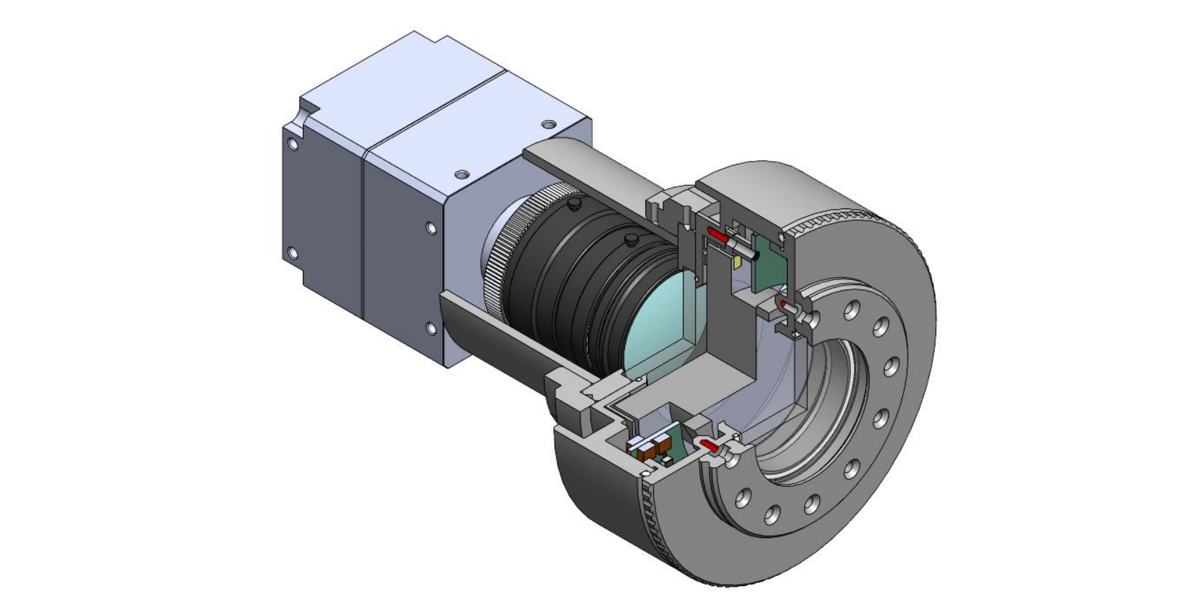

Figure 14. Mechanical design of the hyperspectral imager.

For lightweight structures aluminum is preferred choice due its good mechanical properties and low density. EN AW6082 aluminum alloy was selected as the material as it is commonly available, has good machinability, and is easy to anodize. Each seam interface in the construction increases the rate at which the housing is filled with air. Typically, vacuum sealed flanges are to be installed with multiple screws around the pressure side of the seal. In this case, the main lid was implemented with a single non-standard thread to reduce the diameter of the housing. Altogether, the following interfaces have been identified as the potential points for air leaks:

- the main lid of the housing,
- two 50-mm filters (short-pass and long-pass), functioning as the windows of the chamber,
- custom-sealed electrical feedthrough,
- manual one-way valve for depressurizing the chamber.

To improve the air tightness of the instrument, all the O-ring interfaces are treated with vacuum grease.

The FPI control electronics are enclosed inside vacuum chamber together with the FPI, and the camera lens is installed inside a lens tube to obtain weatherproof construction. The FPI housing is mounted at the end of the lens tube using three stop screws.

The electrical feedthrough connector is sealed using epoxy to reduce air leak rate. The seal was achieved by pumping vacuum at one end of connector while flooding the other end with epoxy. To reduce any possible outgassing, the epoxy was first prepared in a dedicated vacuum chamber before treating the connector. Installation of the connector into housing using epoxy makes the installation permanent. The components of the imager are presented in an exploded view in Figure 15.

Figure 15. Exploded view showing the parts of the instrument.

3.2.5 Current status

As of today (Q4 of the first year of the project), the assembly of the instrument is still in progress. One revision round of the mechanics is completed and the first revision of the electronics has been ordered. The FPI prototype currently installed into the imager is subject to change to better fit the needs of the RILs. The goal is to have the first field prototype by the spring / early summer of the second year of the project (i.e. 2024).

4 Sensor Spatial Planning

In most RILs, in-situ sensors or UAV sensor coverage is required for full data acquisition as standalone solution or complimentary to satellite data. EO data may provide lower resolution than the one needed by the RIL (as an example the resolution by Sentinel 5 is approximately 5x3km). In this case further data collection may be needed from in-situ sensors to augment the data resolution. Similarly, in-situ data may provide different kind of information.

The way the sensors are spatially arranged may affect the quality and adequacy of the collected data. Terrain and cost restrictions as well as data nature itself may prohibit a dense sensor placement in the whole area of interest. A common procedure may be followed by each RIL to select the best possible positioning of the sensors ensuring adequate coverage on area and data while avoiding restrictions bias.

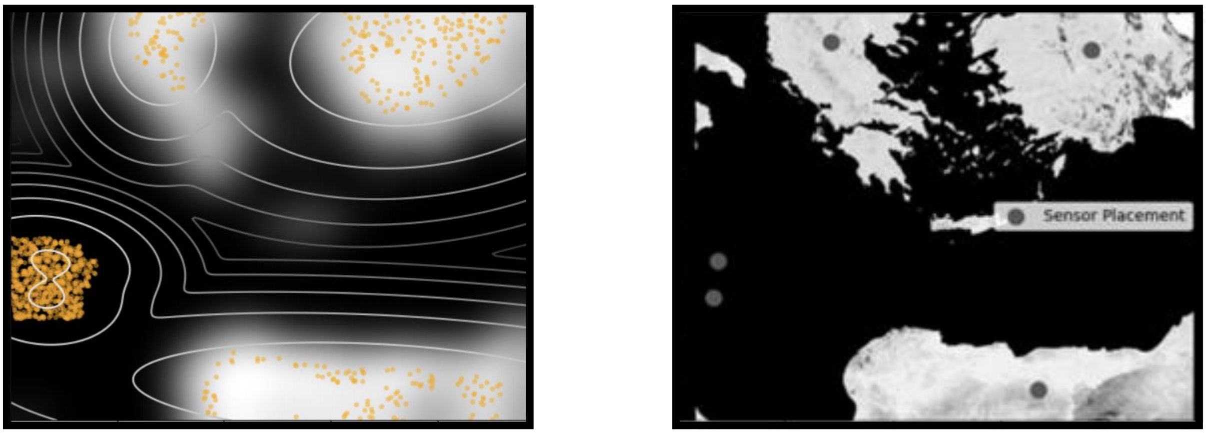

Figure 16. Example of clustering on weighted distribution of low-res map to optimal sensor placement.

A set of techniques is used by ICCS, to define the optimal sensor placement. The procedure combines convex optimization, soft clustering, and cost-minimization on gathered data. Using statistical properties of the dataset (such as maximizing variance and keeping mean value constant), a set of points that can adequately describe the distributed values can be identified. It can be applied with slight modifications to any kind of collected data. The problem is twofold: for known maps, after sub-sampling data with weighted distributions, soft clustering can directly provide optimal sensor positioning. In the cases of unknown maps, an in-house cost-minimization algorithm drive the sensor/UAV placement.

5 Edge processing enabling technologies, real time processing and privacy

Edge computing is a paradigm that enables data processing and analysis near the source of the data, rather than relying on centralized cloud servers. It can reduce the latency, bandwidth, and energy consumption of data transmission, as well as enhance the privacy and security of the data. It can also allow to tailor more resilient applications in case of an unreliable network, since computations and decisions taking are implemented on-site. Edge computing can also enable more complex and intelligent applications that require real-time or near-real-time feedback.

In the case of data for agriculture, which is the topic of this project, edge computing may be of special interest for data quality assessment and improvement, data compression, early alerting systems, or on-the-field interpretation of complex sensors or even for federated learning.

5.1 YaraSense Platform

In the frame of T3.2, we are building a hardware and software platform for integrated sensing and edge computing named YaraSense, tailored to the needs of the agriculture sector (see Table 5).

Table 5. Specifications of the YaraSense platform.

ID	Description
1	The YaraSense hardware can be installed permanently in the middle of a field.
2	The YaraSense hardware is autonomous for its energy supply.
3	The YaraSense hardware can be connected to the cloud using some long-range communication technologies (LoRaWAN, 4G, satellite communications, …).
4	The YaraSense hardware accepts various types of sensors, that can be plugged and unplugged easily on the field.
5	The YaraSense firmware easily accepts new sensor drivers to be installed when it is deployed on the field. (This could be a firmware update)
6	It is possible to connect wired sensors to the YaraSense hardware, using standard communication protocols (I2C, SDI12, SPI, RS485, …).
7	It is possible to connect BTLE sensors, especially using a mesh network.
8	The YaraSense firmware runs an acquisition and processing pipeline (made of processing steps).
9	The YaraSense firmware pipeline can be modified by sending new processing steps or configurations via the long-range communications.
10	The YaraSense enclosure must be weather and dust proof, stable enough to stay on the field in harsh weather conditions and visible enough.

The design of this YaraSense platform will be based on the Edge Spot IoT box, made by EGM. It is a flexible and expandable IoT device capable of connecting to various sensors and actuators, transmitting data across diverse networks. Its versatility stems from the presence of three mikroBUS™ slots, facilitating the integration of extension cards. With over 1000 available options, and the ease of designing custom cards tailored to specific data collection needs, this device offers extensive adaptability. Furthermore, its power management system is highly versatile, enabling energy harvesting from solar panels and efficient battery load management. In scenarios with limited energy resources, the device can enter an ultra-low power mode, consuming mere tens of microamps. See D3.1 for more details.

Figure 17. The Edge Spot hardware.

The EdgeSpot hardware (Figure 17) is stable and has already been used for real-life deployments. Its software is essentially a Software Development Kit (SDK) and an application template that can be both mixed to tailored applications for specific uses. During this project, the SDK will be improved to add new functionalities linked to the needs of the project (edge computing). These developments will allow us to have a complete, deployable platform to fully leverage the potential of the *far* edge computing in this project. These edge computing capabilities will be demonstrated using a multi-sensor integration (see below). This demonstration station will be deployed on at least 2 RI Labs.

A preliminary global architecture for the YaraSense system has been developed, as shown in Figure 18. The YaraSense platform is designed to support three types of serial communication protocols for wired sensors: I2C, RS-485 and SDI-12. Additionally, it will be compatible with BLE (Bluetooth Low Energy) sensors. YaraSense will serve as a centralized hub for collecting data from various sensors, processing it locally through edge processing pipelines, and transmitting the processed data to the cloud for further analysis or integration.

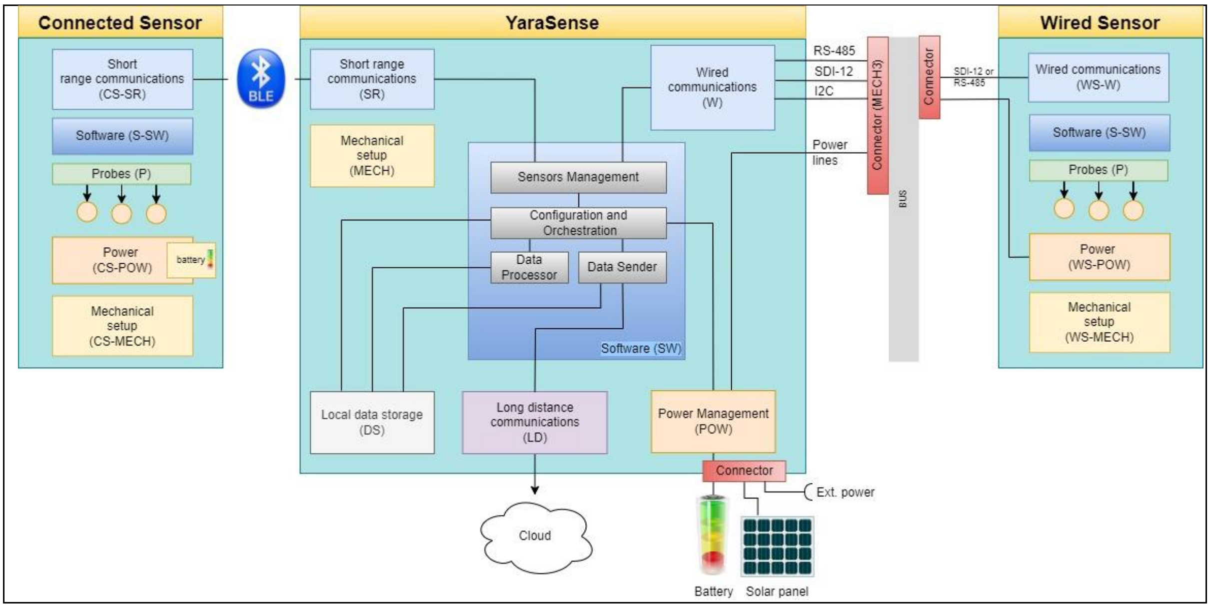

Figure 18. YaraSense architecture.

5.2 YaraSense's sensors

Some sensors have already been selected to meet the specific needs of the RI Labs.

5.2.1 Soil probes

Soil probes are essential tools in modern agriculture, providing critical data about soil conditions that help farmers make informed decisions. These sensors measure various parameters such as moisture, temperature, pH, and nutrient levels, which are vital for optimizing crop growth and yield.
One specific type of soil probe is the Redox probe, which measures the soil's oxidation-reduction potential. This information can provide valuable insights into soil health and microbial activity, both of which influence nutrient availability and plant growth.

To explore the relevance of soil oxidation-reduction potential in agriculture, SWAP Soil Redox probes were selected for testing. These probes require a 12V power supply and communicate using the SDI-12 protocol. SDI-12 is widely used communication standard for environmental sensors. It operates on a single data communication line, allowing multiple sensors to connect to a single data logger. The protocol is designed for low power applications and is ideal for remote or battery-powered systems, making it well-suited for agricultural deployments.

A prototype of the YaraSense system was designed to integrate and evaluate these probes. The EdgeSpot device will supply power to the probes and will operate using a battery supported by a solar panel. The SDI-12 driver will be incorporated into the EdgeSpot SDK to facilitate sensor data acquisition.

To ensure data transmission, a LoRaE5 module will be connected to the EdgeSpot device. The module leverages the LoRaWAN protocol, which is particularly well-suited for this application due to its low-power consumption, long-range communication capabilities, and ability to handle small periodic data transmissions. This makes it an ideal choice for agricultural environments where sensors are often deployed in remote locations with limited access to power and network infrastructure. This enables data to be transferred efficiently to the EGM data platform, Stellio.

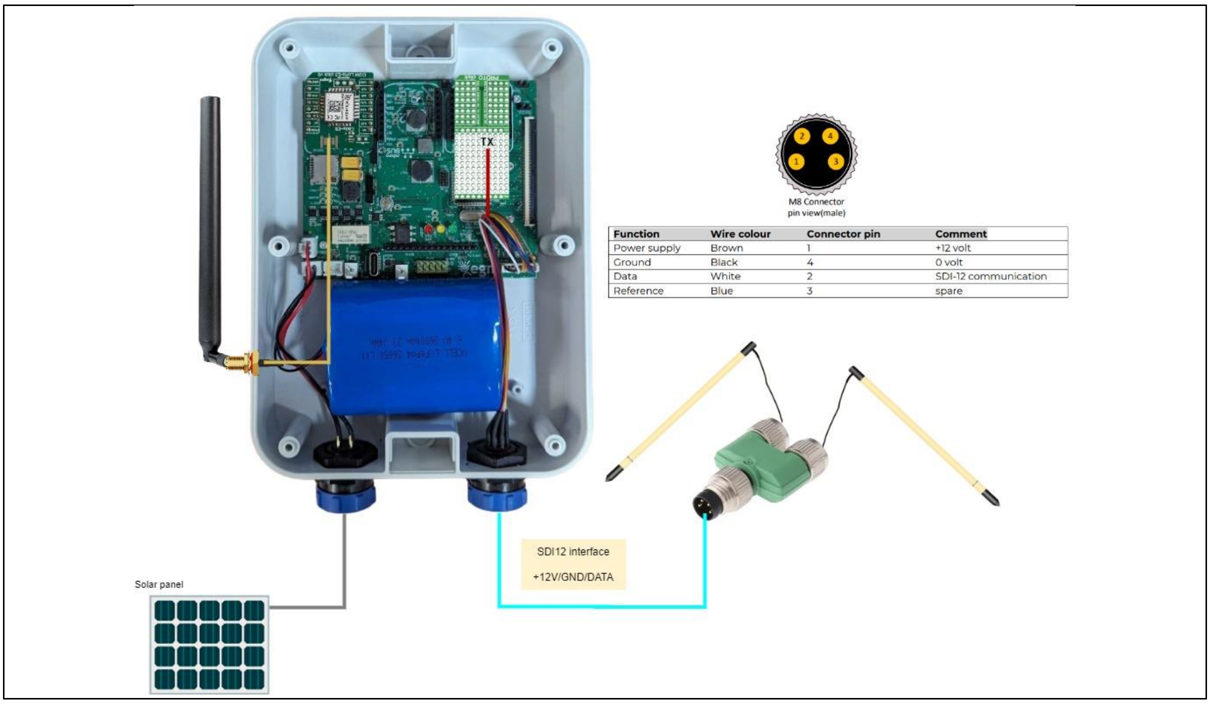

Figure 19. YaraSense prototype with soil probes

The system will be deployed in a vineyard. The probes will be placed in two separate rows of vines, each receiving a different fertilizer treatment. The goal of this test will be to determine whether differences in treatment could be detected through variations in the soil's oxidation-reduction potential.

5.2.2 Spectrometer & IR temperature surface sensors

The Water RI Lab requested a spectrometer sensor to interpolate satellite data and provide more detailed ground-level insights. To meet this need, we selected the SparkFun Triad Spectroscopy Sensor – AS7265x. This sensor covers a wide range of wavelength, making it suitable for detecting and analysing spectral data that can complement satellite imagery. It communicates via the I2C protocol, ensuring straightforward integration with the existing YaraSense system.

In addition, an infrared (IR) temperature surface sensor, the MLX90614 module, will be added to the system. This sensor, which can also be controlled using the I2C protocol, measures the surface temperature of plants and soil without physical contact. Both sensors will be placed next each other, allowing their data to be correlated for deeper insights. The sensors will be powered by the 5V supply from the EdgeSpot device.

5.2.3 eNose

EGM is working on the development of an electronic nose, named eNose, designed to assess various gas emissions, which could provide valuable estimates of organic content or nutrient in the soil. The eNose consists of multiple low-cost gas sensors, each targeting different compounds:

Table 6. eNose gas sensors.

Gas	Sensors
Dioxygen	SEN0219
Ammonia	TGS826-A00
Nitrogen dioxide	SEN0574

Dihydrogen	MQ-8
VOC (Volatile Organic Compounds)	TGS2602
Methane	TGS2611
Butane, propane	TGS2610
Air contaminants	TGS2620

The data from these sensors will be collected by measuring their output voltage, which can then be converted into gas concentration readings. To complement these measurements, a BME280 sensor will be integrated to monitor environmental conditions, including temperature, humidity, and pressure. This sensor communicates using the I2C protocol.

An EdgeSpot device will be utilized to aggregate data from all sensors. A preliminary version of the device has already been developed, enabling data transfer via a serial connection for initial testing and sensor analysis (Figure 19).

The sensor will be integrated as one of the SDI-12 sensors in the YaraSense system. The YaraSense will provide 12V power to the eNose and communicate with it through the SDI-12 protocol. An SDI-12 client driver will be incorporated into the EdgeSpot SDK to transmit the data to the YaraSense SDI-12 server.

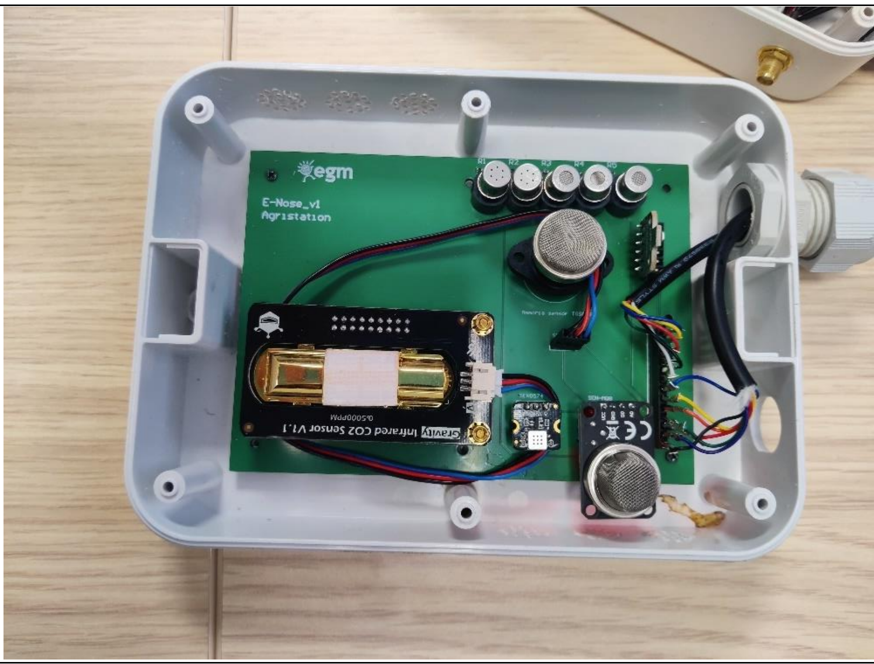

Figure 20. Board containing gas sensors which is connected to the EdgeSpot

5.3 Firmware Update Over-The-Air

Firmware Update Over-The-Air (FUOTA) is a method for remotely updating the firmware of devices, such as sensors or embedded systems, without the need for physical access. In the context of edge processing, FUOTA is particularly useful because it allows updates to be applied to devices that perform data processing locally, at the “edge” of the network. By updating the firmware remotely, it ensures that edge devices can maintain their capabilities without requiring manual intervention. This is crucial in environments where devices are deployed in remote or hard-to-reach locations, such as agricultural fields or industrial sites.

A FUOTA application will be developed to enable remote updates for the YaraSense system. The primary goal of this application is to facilitate the easy update of the edge processing pipelines, allowing for continuous improvement as sensor data is collected and new insights emerge. This will ensure that the system can be optimized without the need for physical intervention, keeping the YaraSense adaptable and efficient in dynamic environments.

6 Sensor Data Catalogue

The information of all sensor data acquired in the RILs will be collected into a common catalogue during the project. This catalogue is an Excel file that will be updated along the data acquisition campaigns. The catalogue will be available at the project cloud repository accessible by the RIL (link below).

As ScaleAgData follows the principle '*as open as possible*', ScaleAgData supports the goals of the Open Science Policy under Horizon Europe and thus, appropriate open science actions will be implemented as an integral part of its proposed methodology as described in deliverable D1.2 Open Science and Data management Plan.

6.1 Data catalogue Structure

The Excel file for cataloguing the data set contains the sheets Definitions & Instructions, Acronyms & Abbreviations, and six sheets named by the RILs. The Definitions & Instructions sheet contain some additional information for filling up the table.

The sensor data set information will be given in the sheets dedicated to each RIL. A simple table format has been selected for inputting the data. The table contains two parts for A) The **Data Sets**, and B) the **Sensors**, defined here as:

Data Set

- A collection of (sensor) data acquired at different times and/or locations. The information recorded should be adequate for 1) **identifying all the data sets acquired and used in ScaleAgdata project**, and for 2) **accessing** (within terms defined by applicable licenses) **those data sets that will be available during/after the project** (e.g. FAIR data).

  Note: A data set may consist of data collected from a set of sensors for a certain time period (e.g. one growing season), as well as for several periods. A data set may also include data from several sensors of different type (e.g weather data).

Sensors

- Sensor(s) acquired to collect the corresponding data set/data sets. The documentation included may consist of one or several sensors per data set.

The data catalogue tables shall be filled up to the level of detail relevant for the above purposes 1) and 2) given in the data set definition above.

For example, in the case of weather station data, containing air temperature, humidity, wind speed and direction, air pressure etc., it is not relevant to list all the separate sensors in the table, if the corresponding data set contains records of all separate sensor data (including location and acquisition time). In this case filling up the table columns for the data set will be enough. However, for tracking the end-to-end performance of a certain data process, the detailed sensor information should be included in the sensor metadata.

Sensor Data Catalogue dated 17.10.2024:

Sensor_Data_Catalogue_v_1.xlsx

7 Conclusion

This deliverable summarizes the planning related to the sensor selection, and data acquisition activities, as well as the development of new sensor technology and edge processing methodologies. This document is thus describing the practical implementation of the early ScaleAgData process step of local data collection design in the development of a generic agricultural data integration concept. The work performed so far is part of the first iteration in the ScaleAgdata co-design development approach and follows directly from the workshops arranged in WP2 and their outcomes, namely the vision scenarios, and user stories, and the derived data flows, use cases, RIL backlogs, and the rolling plan described in the deliverable D2.1.

As this is the first issue of the deliverable, some of the described designs included are in their preliminary stage. Consequently, the presented plans show merely the overall outline for the data acquisition campaigns in the following growing season, while many details in e.g. individual sensor selection and data acquisition may still be missing.

The experiences and the information obtained along the first iteration will play a crucial role in the continuation of the project and in planning the future activities in the RILs and in the hardware design and system implementation.