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

SCALEAGDATA
MAKING DATA COUNT

Deliverable D4.5
"Sensor-integrated data products v2.1

Revised version, 13 May 2026

Author		
Name	**Organization**	**Signature**
Adam Fojud	WODR	
Annimari Hartikainen	LUKE	
Dainis Jakovels	IES	
Emmanouela Prassou	ICCS	
Frederico Tupinamba Simões	VITO	
George Charvalis	NP	
Giorgia Milli	VITO	
Giorgos Tsilimanis	ICCS	
Gunnar Grosse Hovest	ATB	
Hanna Huitu	LUKE	
Haris Ampas	AUTh	
Idan Kopler	MIGAL	
Isabelle Piccard	VITO	
Kristof Van Tricht	VITO	
Maria Pat Gonzalez	IFAPA	
Mariapina Castelli	EURAC	
Matti Pastell	LUKE	
Michal Blaszczak	PSNC	
Nick Berkvens	EV ILVO	
Nick Gutkin	VITO	
Nuno Grosso	Deimos	
Paolo Cosmo Silvestro	Deimos	
Rado Guzinski	DHI	
Sarah Verbeke	UGent	
Stephanie Delalieux	VITO	
Thomas Papakosmas	NP	
Valentina Manstretta	HORTA	
Walid Ghariani	DHI	

Review on behalf of the Executive Board		
Name	**Organization**	**Review date**
Stephanie Delalieux	VITO	10/12/2025
Giorgos Tsilimanis	ICCS	10/12/2025
Wannes De Man	EV ILVO	10/12/2025
Isabelle Piccard	VITO	13/05/2026

Revision records			
Version	**Date**	**Changes**	**Authors**
1.0	30/06/2025	D4.3	VITO
2.0	15/12/2025	Original version D4.5	VITO
2.1	13/05/2026	Revised version D4.5	VITO

Acronyms and Abbreviations	
6s	Second Simulation of the Satellite Signal in the Solar Spectrum
ADAM	Adaptive Moment Estimation
AGINS	Agro Insurance International
AI	Artificial Intelligence
AIM	Artificial Intelligence and Machine Learning
ANN	Artificial Neural Network
AOI	Area of Interest
API	Application Programming Interface
APSIM	Agricultural Production Systems sIMulator
ATB	Institut für angewandte Systemtechnik Bremen GmbH
AUTh	Aristotle University of Thessaloniki
AVR	AVR BVBA (Belgium)
CA	Consortium Agreement
CDSE	Copernicus Data Space Ecosystem
CMAAS	Crop Model as a Service
CNHi	CNH Industrial Belgium
CNN	Convolutional Neural Network
CRS	Coordinate Reference System
DA	Data Assimilation
DEM	Digital Elevation Model
DES	Deimos Spain
DHI	DHI A/S (Denmark)
DL	Deep Learning
DME	DEIMOS ENGENHARIA SA
DMK	DMK Deutsches Milchkontor GmbH
DMS	Data Mining Sharpener
DSS	Decision Support System
DVC	Data Versioning Control
EA	Actual vapour pressure
EC	Electrical Conductivity
EC-EU	European Commission
EEAB	External Expert Advisory Board
EGM	Easy Global Market SAS
EO	Earth Observation
EOD	Earth Observation Data
ERGAS	Relative Dimensionless Global Error
ET	Evapotranspiration
Etref	Reference evapotranspiration
EURAC	Accademia Europea di Bolzano (Eurac Research)
EV ILVO	Flanders Research Institute for Agriculture, Fisheries and Food
ExBo	Executive Board
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FEU	Farm Europe

FL	Federated Learning
GA	General Assembly
GDPR	General Data Protection Regulation
GPI	Grassland Production Index
GPP	Gross Primary Productivity
GRD	Ground Range Detected
HE	Homomorphic Encryption
HGBR	Histogram-based Gradient Boosting Regression
HPC	High Performance Computing
HR	Relative Humidity
HSI	Hyperspectral imaging
ICCS	Institute of Communication and Computer Systems
IFAPA	Instituto Andaluz de Investigación y Formación Agraria, Pesquera y Alimentaria
IoT	Internet of Things
IPR	Intellectual Property Rights
ISMN	International Soil Moisture Network
KMI	Royal Meteorological Institute of Belgium
LAI	Leaf Area Index
LAU	Local Administrative Unit
LST	Land Surface Temperature
LSTM	Long Short-Term Memory
LUKE	Natural Resources Institute Finland
MAE	Mean Absolute Error
MIGAL	MIGAL Galilee Research Institute
ML	Machine Learning
MNDWI	Modified Normalized Difference Water Index
MPC	Multi-Party Computation
MSE	Mean Squared Error
MSI	Multispectral Instrument
MST	Management Support Team
N	Nitrogen
NDVI	Normalized Difference Vegetation Index
NMDI	Normalized Multi-band Drought Index
NP	Neuropublic SA
NPP	Net Primary Productivity
NRT	Near Real Time
NRT	Near Real Time
OC	Organic Carbon
OHB DS	OHB Digital Services GmbH, Bremen, Germany
PAR	Photosynthetically Active Radiation
PEF	Product Environmental Footprint
PET	Privacy Enhancing Technologies
PO	Project Officer
PPI	Plant Phenology Index

PRESTO	Pretrained Remote Sensing Transformer
PSNC	Instytut Chemii Bioorganicznej Polskiej Akademii Nauk
PSNR	Peak Signal-to-Noise Ratio
R&D	Research and Development
Rad	Solar Radiation
RBF	Radial Basis Function
RECO	Ecosystem respiration
RF	Random Forest
RGB	Red, Green Blue
RH	Relative Humidity
RIE	Research and Innovation Environment
RIL	Research and Innovation Lab
RMSE	Root Mean Square Error
ROI	Region Of Interest
RUE	Radiation Use Efficiency
RZSM	Root Zone Soil Moisture
S1	Sentinel-1
S2	Sentinel-2
S3	Sentinel-3
SAM	Spectral Angle Mapper
SCL	Scene Classification Layer
SGD	Stochastic Gradient Descent
SLSTR	Sea and Land Surface Temperature Radiometer
SM	Soil Moisture
SME	Small and Mid-size Enterprise
SOC	Soil Organic Carbon
SR	Super Resolution
SRE	Signal-to-Reconstruction error
SSIM	Structural Similarity Index Measure
SSL	Secure Sockets Layer
SVM	Support Vector Machine
T	Temperature
TIR	Thermal Infrared
Ugent	University Gent
UAV	Unmanned Aerial Vehicle
UQI	Universal Image Quality Index
VI	Vegetation Index
VITO	Vlaamse Instelling voor Technologische Onderzoek
VNIR	Visual Near Infrared
VPD	Vapor Pressure Deficit
VRI IES	Foundation "Institute for Environmental Solutions"
VTT	Technical Research Centre of Finland Ltd.
WODR	Wielkopolski Osrodek Doradztwa Rolniczego w Poznaniu
WP	Work Package

Table of Contents

List of tables

List of figures

1. Introduction

1.1 Project overview

ScaleAgData addresses the HORIZON-CL6-2022-GOVERNANCE-01-11 call, focusing on upscaling (real-time) sensor data for EU-wide monitoring of agricultural production and agri-environmental conditions. Running from January 2023 to December 2026, the project brings together a consortium of 26 partners from 14 countries.

The project's dual vision is to (1) develop robust governance and organizational models for managing complex, multi-source agricultural data streams, and (2) advance data technologies that enable scaling of farm-level sensor data to regional, national, and European datasets for improved agri-environmental monitoring and agricultural management.

ScaleAgData pursues five core objectives:

- Develop innovative approaches for in-situ data collection and advanced data technologies.
- Enable and promote data sharing across the entire agricultural data value chain.
- Demonstrate how sensor data can be scaled to agri-environmental data products at national, regional, and European levels.
- Showcase the benefits of improved monitoring capacities in precision farming contexts.
- Demonstrate the added value of upscaled regional datasets for the broader agricultural sector.
- Throughout its lifecycle, ScaleAgData explores seven innovation areas: innovative sensor technology, edge processing, data sharing architecture and governance, satellite data augmentation, data assimilation to service development, privacy-preserving technologies, and data integration methodologies.

Six Research and Innovation Labs (RILs) have been established across diverse biogeographical regions in Europe. These RILs, covering water productivity, crop management, yield monitoring, soil health, grasslands, and dairy, serve as testbeds for evaluating and demonstrating different data upscaling and integration models. The project will deliver recommendations on how integrated datasets can support national and regional policymaking, strengthening both the competitiveness and sustainability of European agriculture.

1.2 Scope of the document

This document is an update to Deliverable 4.3, building on the foundations and progress reported in the previous version. It reflects the latest developments and lessons learned in Task 4.2 of the ScaleAgData project, which focuses on developing and refining methodological frameworks for integrating sensor data into agri-environmental data products.

The primary aim remains to enhance monitoring capabilities at both local and regional scales, delivering improved information and services to farmers, and supporting public authorities and private stakeholders with actionable, high-quality data. Sensor data integration is addressed at multiple stages of the monitoring workflow: (1) directly improving model performance with sensor inputs, (2) adapting or correcting model outputs using real-time data; (3) enabling the creation of new data products, even when only limited sample data is available.

Since the previous deliverable, the frameworks and prototypes developed in Task 4.2 have continued to be evaluated and validated by the RILs as part of Task 4.4. All methodological frameworks (including code, notebooks, and documentation) and resulting data products are made accessible via the Research & Innovation Environment (RIE), alongside existing project outputs. The project continues to follow an iterative approach, with the goal of reaching Technology Readiness Level 5 (TRL5). Through ongoing validation, the RILs can determine which data and methods are most suitable for their specific applications and use cases.

This updated deliverable provides an overview of the frameworks developed and improved since the previous version, details the data products made available on the RIE during the first and ongoing iteration rounds, and outlines the process for further updates and partner contributions as the project progresses.

1.3 Document structure

This updated deliverable is structured as follows:

- Chapter 1 provides an overview of the ScaleAgData project, describes the scope and responsibilities of this deliverable, and highlights the updates compared to the previous version.
- Chapter 2 presents an updated overview of the methodological frameworks developed and refined to facilitate and promote the integration of sensor data into agri-environmental data products, including recent advances and lessons learned.
- Chapter 3 details the agri-environmental data products developed and validated in the RILs during the first and ongoing iteration rounds, with a focus on new results, improvements, and partner contributions since the last deliverable.
- Chapter 4 explains how and where the methodological frameworks and data products are made available on the Research & Innovation Environment (RIE)

1.4 Evolution of the document

Deliverable 4.3 described the initial methodological frameworks and sensor-integrated data products developed during the first iteration round of the ScaleAgData project (M7–M18).

This Deliverable 4.5, which is the second version of the "Sensor-integrated data products" deliverable builds on that foundation, providing a comprehensive overview of the progress made during the second iteration round. It includes new and refined frameworks, additional data products, validation results, and lessons learned since the previous version.

Following the review of the first periodic report, following changes were applied to the original document.

History of changes	
Section	**Nature of change and reason (if applicable)**
2 – Methodological frameworks	Updated introduction providing more details on the work done during the second iteration.
3 – Agri-environmental data products	Table 9 has been added. This table provides an overview of the EO and sensor data inputs used to generate the agri-environmental data products, the methods used for sensor data integration as well as how

	these data products are validated within the RILs. The introduction of each data product now also explicitly mentions which sensor and EO data have been combined to generate the data product. Section 3.2.2 has been updated to reflect recent changes.
4 – Research and Innovation Environment	This section is updated with a clarification on the role and set up of the RIE. Two tables have been included: Table 70 provides a timeline for the ingestion of the sensor-integrated data products into the RIE catalogue, while Table 71 presents an overview of the methodological frameworks and models available in the GitHub repository. For those elements that are not yet available, a timeline is provided.

2. Methodological frameworks

ScaleAgData continues to strengthen and expand its methodological frameworks that enable the integration and effective use of sensor data within agri-environmental data products. These frameworks support improved monitoring and decision-making across both local (farm and field) and regional (landscape and policy) scales.

This deliverable presents the advancements achieved during the second iteration round, building on the foundations laid in the first iteration and addressing the needs and requirements identified through the co-design process (WP2) with RILs and relevant stakeholders. The methodological work in WP4 is structured around several key pillars:

- **Federated learning**: Development of privacy-preserving methods that enable partners to collaboratively train machine learning (ML) models using locally held sensor data, without data sharing.
- **Data assimilation**: Integration of in situ sensor and Earth Observation (EO) data with crop and environmental models to enhance Digital Twin capabilities.
- **Data integration:**
    - Solutions for training models with limited labelled data (e.g., few-shot learning).
    - Methods for fusing local sensor data with spatially explicit EO-based products, improving model accuracy, and enabling near-real-time updates.
- **Super resolution tools**: Techniques for increasing the spatial or temporal resolution of EO S2/S3 datasets to enhance product detail and usability.
- **Meteorological time series forecasting:** Techniques that use sensor observations, EO data, and model simulations to forecast key meteorological and agri-environmental variables

Since the first iteration, these frameworks have been further developed, tested, and validated using specific input datasets and RIL use cases (Task 4.4). Several methods have been refined and piloted in operational settings, while new collaborations between RILs have also been established. The current iteration has focused on improving the scalability, interoperability, and operational deployment readiness of the developed frameworks, while progressively linking the methodologies, workflows, APIs, and generated products to the Research and Innovation Environment (RIE). This includes integration of selected tools, services, catalogues, and processing components to support testing, validation, sharing, and reuse across the consortium. Further cross-RIL validation activities, including validation on newly acquired in-situ datasets and assessment of framework transferability across different RIL contexts, will be carried out in WP5.

The following sections describe the current status, recent developments, and lessons learned from this second iteration, highlighting the progress made towards operational, sensor-integrated data products.

2.1 Privacy-preserving technologies

As EO-based analytics increasingly combine satellite observations with in-situ sensor data, privacy considerations become more relevant. Modern EO missions generate large volumes of up-to-date imagery, and when these data are integrated with micro-data (e.g., field-level measurements, management logs, IoT sensor streams), new risks related to data protection and ethical use can arise.

To address these concerns, ScaleAgData investigates Privacy Enhancing Technologies (PETs) that enable data-driven modelling while keeping sensitive information protected.

The project considers three main families of PETs relevant to sensor-integrated agricultural data products:

- Model-based approaches, e.g. Federated Learning, continual learning, or neural-network encoding, where models are trained collaboratively without sharing raw data.
- Data-based approaches, e.g. Differential Privacy, where data are transformed to prevent the identification of individual contributors.
- Encryption-based approaches, e.g. Multi-Party Computation or Homomorphic Encryption, enable processing of encrypted data without revealing the underlying values.

Feedback from RILs (collected through WP2 co-design activities and ongoing consultations) guides which techniques are most relevant for deployment within ScaleAgData.

In the **first iteration**, developments centered on requirements from the Soil Health RIL, preparing a federated learning experiment to train an AI model for soil carbon estimation from sensor-integrated datasets. Federated Learning (FL) was prioritized because it enables collaborative model training across partners without moving or exposing raw local data, thus meeting strong privacy constraints.

Based on workshops with the Soil Health RIL, a set of technical requirements was defined, mainly extending the capabilities of the Flower.ai federated learning framework to support ScaleAgData needs. Key requirements included:

- Automated configuration of FL clients and servers within the virtual lab environment
- Additional metrics for model evaluation
- Support for local validation datasets
- Dataset-balancing options and weighted model updates
- Exploration of pretraining strategies and compatibility with classical ML models
- Tutorials and documentation for integrating new/custom models

During the **second iteration**, developments focused on the automated setup of Flower.ai servers and clients within the ScaleAgData environment. This functionality simplifies deployment by handling networking, connections, and configuration without requiring users to manually define technical details.

A working version of this module is available in the project's GitHub repository: https://github.com//ScaleAgData

2.2 Data assimilation methodologies

Data assimilation combines models with observations. In ScaleAgData, it is used within Digital Twins (see Deliverable D4.2) to merge remote-sensing products with crop simulation models. The goal is to improve yield and quality forecasts, support precision nitrogen management, and assist in irrigation scheduling.

Traditionally, crop models assimilate observed leaf area index (LAI) by adjusting simulated biomass using methods such as the Ensemble Kalman Filter. Biomass in different plant organs is rebalanced to match the observed LAI. This approach assumes that model calibration, soil parameters, and nutrient

status are correct, and that any mismatch with observations stems from external factors such as pests or unobserved rainfall.
However, these assumptions often fail with on-farm data, where cultivar-specific parameters or detailed soil profiles are limited. ScaleAgData therefore incorporates uncertainty in both model parameters and initial conditions. Ensembles of crop models are generated, and the most likely model given the observations is selected using a modified particle filter. In the project, we also tested the simultaneous assimilation of new EO products, such as canopy chlorophyll, canopy water content, and evapotranspiration, to improve modeling of nitrogen and water stress during the growing season.

2.2.1 Data assimilation service development

During the **first iteration**, a Python library and a C# extension were developed to run model ensembles using the Agricultural Production Systems sIMulator (APSIM) crop simulation model, along with an example Jupyter notebook for running the code. Past experimental data and Terrascope EO data products from wheat fields in Finland and Belgium were collected to support algorithm development and testing. Method performance was and is being explored in collaboration with the RILs (see also Chapter 3).

In the **second iteration**, the APSIM crop model was replaced with the simplified wheat model developed in T4.2 and implemented within the Crop Model as a Service (CMAAS). This model was designed specifically for the data assimilation use case, providing a clear causal link between photosynthesis-driven biomass accumulation and leaf area index development. In tests using Yield Lab datasets, the CMAAS model performs comparably to the more complex APSIM model, but with substantially better computational efficiency.

Additionally, an Application Programming Interface (API) endpoint was developed in Kotlin to support the CMAAS data assimilation use case. The crop model and API are described in more detail in Deliverable D4.4.

Users provide daily weather data, fertilization records, crop model parameters, and LAI time series. The service assimilates the LAI observations with the model forecasts and returns yield and nitrogen uptake predictions (Figure 1). The spatial resolution is defined by the API user and depends on the precision of the input data. For API development, Yield Lab data was aggregated to an H3 hexagonal grid https://h3geo.org) at ~0.2 ha resolution. LAI data was sourced from TerraScope, management data from the Yield Lab via CNH, and daily weather data from the Open-Meteo API (https://open-meteo.com/).

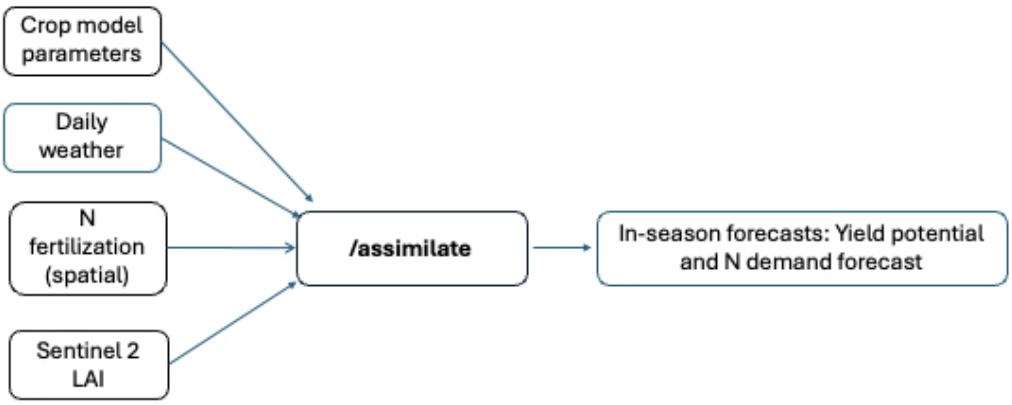


Figure 1: Schematic of the inputs and outputs of the Crop Model as a Service assimilation endpoint.

A Python library for interfacing with the CMAAS API was also developed, along with example notebooks to support integration of the service into the ScaleAgData RILs. Compiled versions of the CMAAS service and Jupyter notebooks for (1) model calibration and (2) data assimilation was developed. The API is documented using the OpenAPI 3.1.0 specification (Figure 2). The documentation is bundled with the application, and models are automatically generated from the CMAAS source code.

The data assimilation framework was designed to be extensible: additional endpoints for assimilating forecasts using other observations, such as biomass or crop nitrogen content, can be added easily or combined with LAI-based assimilation.

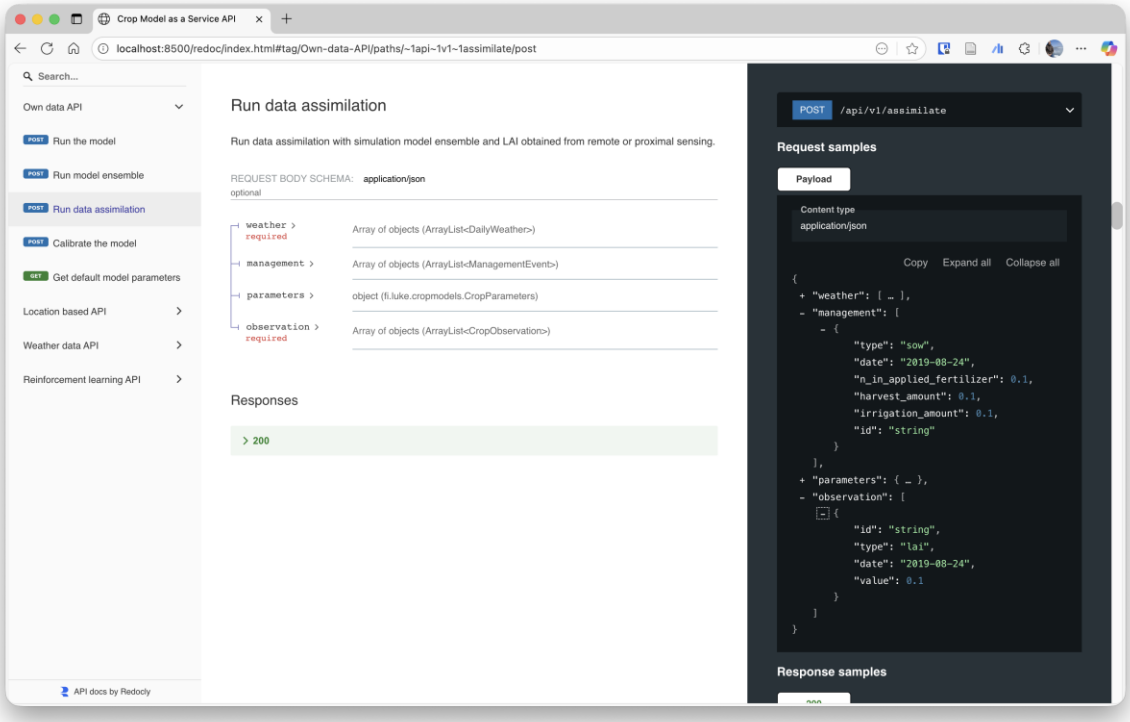


Figure 2: Documentation for the data assimilation endpoint in Crop Model as Service API.

2.2.2 Data assimilation methodology

The data assimilation methodology in the CMAAS model uses a four-dimensional ensemble-variational (4DEnVar) data assimilation. The method seeks the initial condition such that the forecast best fits the observations within the assimilation interval. In the current CMAAS assimilation endpoint, the growth of the crop is simulated using an ensemble of 200 members where the radiation use efficiency (RUE) parameter is varied evenly between 0.5 and 2.5, where 0.5 represents extremely stressed crop and 2.5 is above the normal unlimited growth calibrated from the yield lab field experiment data. The RUE is affected by both N and water stress, which allows estimating compound stressors with a single ensemble. Model forecasts are calculated by averaging the weighted crop model ensemble time series with the provided LAI time series, the weights w_i^* are based on the likelihood of the Euclidean distance (e_i) between the observed LAI and simulated LAI at each observation time point t following the normal distribution:

$$e_i = \Sigma_{t=0}^{T}\left(LAI_t^{obs} - LAI_t^{sim}\right)$$

$$l_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}*(e_i/\sigma)^2}$$

$$w_i^* = \frac{l_i}{\Sigma_{i=0}^{N} l_i}$$

where e_i = error between observed and measured LAI at time t, LAI^{obs} = observed LAI, LAI^{sim} simulated LAI. l_i likelihood for the ensemble member i, w_i^* normalized weight for the ensemble member. Equal weight was given to observations at each time point as in the current implementation all ensemble members start from the same initial conditions. The test case for the methodology for yield forecasting in the yield lab is presented under section 3.3.3.

2.3 Data integration methodologies

2.3.1 Solutions for limited availability of labeled data for model training

A key challenge in agricultural remote sensing is the limited availability of high-quality labeled data for supervised ML and deep learning (DL) models. Although remote sensing provides abundant information, privacy constraints, annotation costs, and heterogeneous data sources often limit the amount of labelled data available for training.

Recent approaches developed to overcome these limitations involve leveraging pretrained models that have learned either in a supervised manner from large, annotated datasets or in an unsupervised or self-supervised manner from unlabeled datasets. In the latter scenario, when the training dataset is large and diverse enough to comprehensively cover the variations in the input distribution, the resulting model exhibits strong generalization capabilities which can be exploited for wide range of applications dealing with similar types of input data. These models are known as Foundation Models.

In remote sensing, for instance, a model trained on extensive unlabeled satellite data which learned to capture and interpret the variations of different type of sensor signals, could be adapted for multiple downstream tasks that make use of similar input signals. Given a dataset with only a few annotated examples, a pretrained model can either be used to create compressed versions of the inputs for supervised ML training or provided with a classification/regression layer and retrained for the specific task. Few-shot learning aims to develop models that can learn from a small number of labelled instances while enhancing generalization and performance on new, unseen examples.

As part of ScaleAgData, few-shot learning was implemented using a robust Foundation Model pre-trained on multi-sensor data time-series. This model was fine-tuned with a limited number of application-specific annotated examples to be specialized ad hoc for the target downstream task. Figure 3 offers an overview of the pipeline applied to a yield prediction use case.

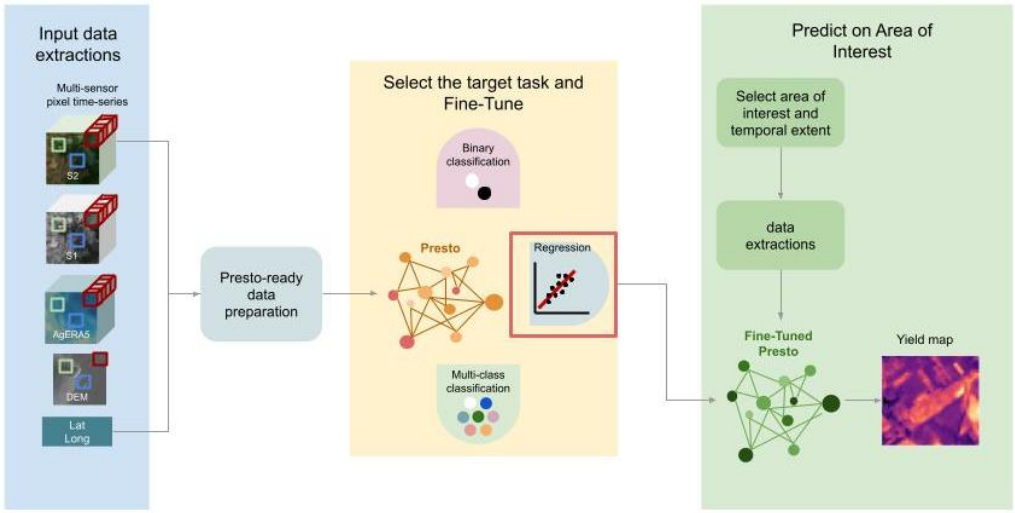


Figure 3: Overview of the Presto-based few-shot learning pipeline fine-tuned for yield estimation. The model was fine-tuned for yield prediction (regression task) on a limited amount of Sentinel-1, Sentinel-2, meteorological and topographic annotated pixel time-series. After training, the model was deployed over a selected Area of Interest to produce a yield map on unseen data.

To achieve our objectives, we built upon the Presto (**Pre**trained **Re**mote **S**ensing **T**rans**fo**rmer) framework as our foundation model (Tsjeng et al., 2023). Presto was originally trained on a large unlabeled dataset (~21M examples) of Sentinel-2 (S2), Sentinel-1 (S1), Meteorological (precipitation and temperature) and Topography (slope and elevation) pixel-timeseries, with each time-step corresponding to the monthly aggregation of the sensor observations (Tsjeng et al., 2021). The model integrates multi-sensor data and, for each pixel-timeseries, compresses the information into 1D embeddings, effectively capturing long-range relationships across the temporal and radiometric & sensor dimensions. By utilizing Presto, we were able to eliminate redundant details from the raw input data, to significantly increase the signal-to-noise ratio, and to generate a concise yet highly informative representation of the inputs. In essence, the label-free self-supervised learning step leads to a significant head start for a downstream supervised learning task compared to any other approach that would have to start learning from scratch.

For the **first iteration**, we selected the Presto model version developed in collaboration with WorldCereal[1]. This updated architecture was expanded to handle the sensor inputs used in ScaleAgData and to cope with missing time steps or sensor observations, an improvement over the original Presto design. Within this project, we further refined the architecture so that Presto can ingest and interpret data at a 10-day temporal resolution (instead of only monthly) and be fine-tuned for both regression and multi-label classification tasks. These enhancements ensure that the few-shot learning methodology developed here can effectively support the diverse use cases across the research labs.

In the **second iteration**, ScaleAgData has successfully developed and demonstrated an operational few-shot learning pipeline based on foundation models, specifically leveraging the Presto framework. Presto, which was trained on a large, unlabeled dataset of S2, S1, meteorological, and topographic pixel time series, is designed to compress such multi-dimensional data into rich and informative embeddings. Due to its strong generalization capabilities, the model was applicable to a wide range of downstream tasks and demonstrated particularly robust addressing scenarios with limited annotated data.

[1] https://github.com/WorldCereal/presto-worldcereal/

The end-to-end workflow begins with data preparation, where users provide annotated datasets in formats such as parquet, GeoJSON, shapefile, or GPKG. Each geometry is linked to a date, unique ID, and annotation. To improve the reliability of target values and reduce noise, good practices include removing polygons near field borders and subdividing large polygons into 20 m × 20 m subfields for regression tasks. Quality control plays a critical role in this step, as invalid geometries, polygons located on non-field areas, and outliers can have a significant impact on model performance. Ensuring data integrity and curation is therefore a central responsibility for users.

Once the dataset is validated, feature extraction is carried out using the Copernicus Data Space Ecosystem (CDSE). Presto-ready features are generated for the user's valid geometries and period of interest, with the process supported by a dedicated Jupyter Notebook that guides users through parameter selection and data extraction. The foundation model is then fine-tuned on the user's annotated dataset to address the specific application. Demonstrations to date have included potato yield estimation (regression) and crop/no-crop classification (binary classification), with the framework also supporting multiclass classification tasks. Following fine-tuning, the model can be applied to unseen areas, enabling inference and the generation of spatial prediction maps for selected regions and time ranges.

In April 2025, VITO showcased the full pipeline in a live demonstration that covered all stages from data preparation and feature extraction to fine-tuning and inference, using real-world datasets. The workflow, accessible through a Jupyter Notebook and supported by resources in the project's GitHub repository, provided concrete evidence of the approach's effectiveness. The demonstration underlined three key findings: first, that the generalization capabilities of foundation models can be successfully exploited for a wide range of applications, even when only limited labeled data are available; second, that few-shot learning achieves better performance than traditional machine learning models in low-label scenarios; and third, that high data quality remains the most important factor for ensuring reliable and robust results.

The first internal tests of the framework within ScaleAgData were carried out using the Yield Lab datasets (see section 3.3), demonstrating its practical value. The Grassland Lab (section 3.5) and Crop Management Lab (section 3.2) are also preparing to adopt the framework for their use cases.

2.3.2 Modeling: development of improved data products using sensor data

EO products offer broad spatio-temporal coverage but often lack local precision, while in-situ measurements provide accurate point information with limited spatial extent. Combining both can produce higher-quality Root Zone Soil Moisture (RZSM) estimates.

In the **first iteration**, the methodological focus was on improving the RZSM product developed by DHI, which is currently based on physics-based modelling combined with EO inputs. Local measurements, such as soil moisture, soil properties, weather data, and irrigation records, can help reduce uncertainties in these models and improve product accuracy.

In the **second iteration**, work has focused on refining and improving the RZSM data product by enhancing the integration of EO and in-situ soil moisture measurements. Whereas D4.3 concentrated on outlining the conceptual framework and describing possible integration approaches, the current iteration has advanced this work into practical implementation, tool development, and testing with RILs and project partners.

To support testing and partner engagement, EO-derived evapotranspiration (ET) and soil moisture products were distributed to six partners across three RILs. These datasets are being used both for benchmarking and as inputs to local-scale models of yield, irrigation, or crop status. In parallel, a dedicated experiment with WP3 was conducted to determine how EO-derived spatial variability can

guide optimal placement of soil moisture sensors. This work supports sensor network design and ensures that in-situ measurements are representative of the key soil and vegetation patterns observed from EO.

A major milestone of this iteration was the presentation and release of the machine learning toolbox for EO–in-situ data fusion during an online workshop in October 2025 and released as a Python library and Jupyter Notebook (https://github.com/ScaleAGData/dhi-mleo). The toolbox is now available for RILs to test, evaluate, and provide feedback on. Testing will continue until April 2026, after which the framework will be refined and redeployed in the RIL environment for final assessment between September and December 2026.

Overall, the second iteration has resulted in significant progress compared to D4.3. The methodological approaches outlined previously have now been translated into practical tools, calibrated models and collaborative testing workflows. Although soil moisture has been the primary focus, the tools and methods developed during this phase are broadly applicable to other EO-based variables and types of in-situ measurements.

Case study: EO and In-situ data integration for improved Soil Moisture: HOBE case study

In-situ SM measurements illustrated in Figure 4 were gathered from about 30 stations across HOBE Network over Denmark for the year of 2018 and were originally provided by HOBE center of Hydrology (http://www.hobe.dk/) and made available through the download service of the International Soil Moisture Network (ISMN, https://ismn.earth/) using an API URL request. At each station, Decagon 5TE sensors were installed at 0-5, 20-25 and 50-55 cm depth. The 5TE sensors simultaneously measure soil moisture, temperature, and electrical conductivity. The data includes information such as the station ID, the latitude, the longitude, the sensor ID and corresponding depth, and the measurements of SM.

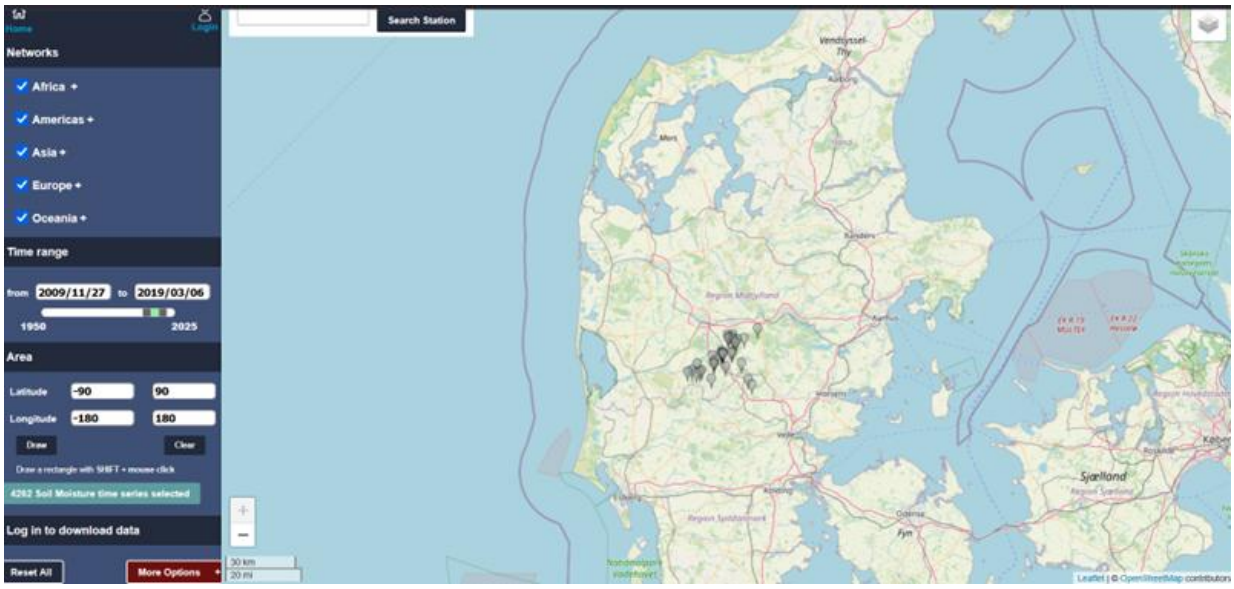


Figure 4: The International Soil Moisture Network (ISMN) data viewer https://ismn.earth/en/dataviewer/

Methodological Workflow

Figure 5 represents the workflow of this case study including the following stages:

- Acquisition of earth observation datasets: Collection of multi-sources EO data from S2, ERA5, along with DHI EO derived products (E.g; ETref Reference Evapotranspiration, fAPAR (Fraction of Absorbed Photosynthetically Active Radiation), Potential Evapotranspiration)
- Data preprocessing: Spatial and temporal harmonization of EO datasets, georeferencing, resampling, and gap-filling procedures. Sensor data filtering and outliers' detection.
- In situ and EO data integration: Integration of sensor-based soil moisture data with EO-derived variables.
- Machine learning data preparation: Feature engineering, outliers' detection and feature exploration.
- Model development and validation: Applied ML tree-based algorithms.

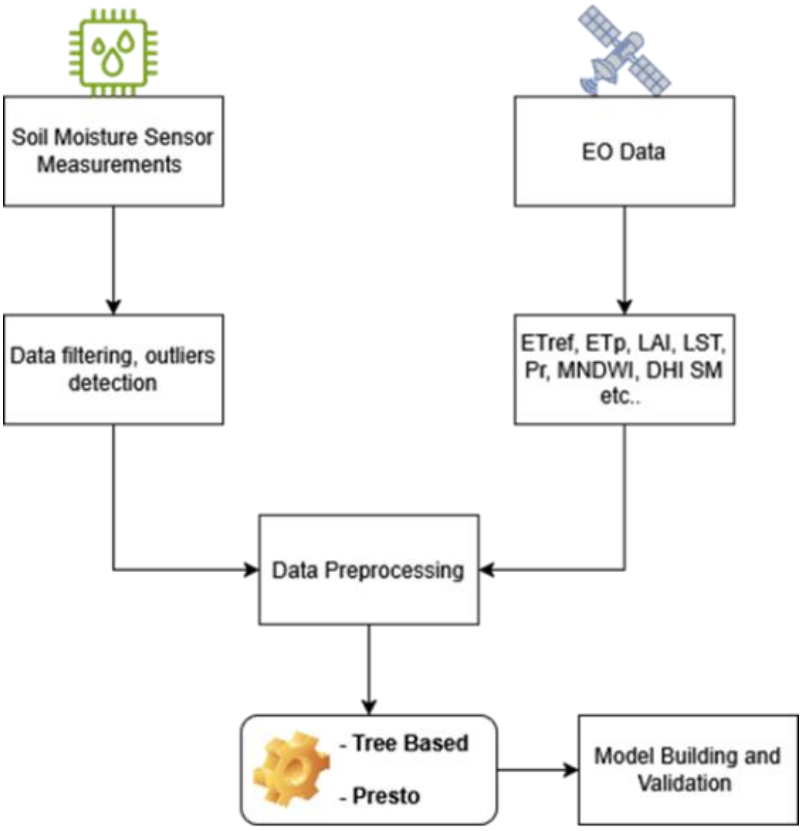


Figure 5: Workflow representing the different stages implemented for the in situ-EO-ML integration.

Machine learning

As a baseline ML model to test our methodology, we used a tree-based model called Histogram-based Gradient Boosting Regression (HGBR) which is part of the scikit-learn library. HGBR is an ensemble method that builds regression trees sequentially, where each tree is trained to correct the errors of the previous ones. Unlike standard gradient boosting (Friedman 2001), HGBR discretizes continuous features into histograms to speed up split search and decrease memory usage. This algorithm allows HGBR to efficiently handle large data while keeping strong predictive performance. Moreover, it is also worth noting that the model handles missing values during training and prediction, which is particularly useful for EO applications where missing values are common.

Results:

- Descriptive statistics

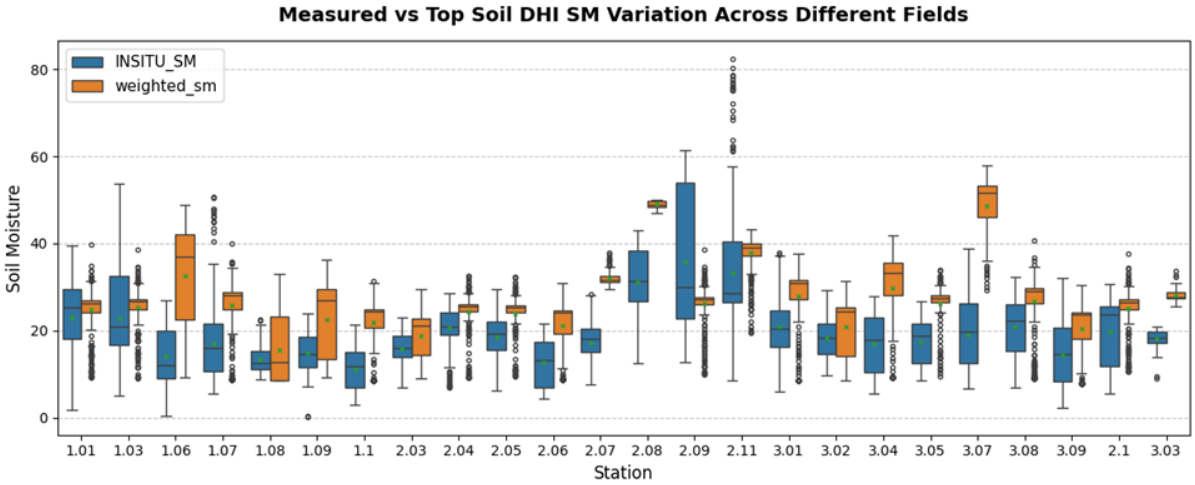


Figure 6: Measured vs topsoil RZSM Variation across different stations.

Figure 6 shows the histograms of soil moisture across different stations as compared to the produced weighted RZSM. First, the ground truth soil moisture recorded at different stations and fields shows a significant variability. Second, the produced weighted RZSM generally overestimates the true SM values at most stations.

- Feature correlation

Figure 7 shows the Pearson correlation between SM and the 28 selected predictor variables on the training dataset. From these variables, 10 variables showed a negative correlation with SM. The variables with the highest negative correlation were PT_20_ETp_day_gf (Potential ET gap-filled) with a Pearson correlation coefficient of r = -0.47, followed by the month with r = -0.41. DOY (day of the year), ETref, had correlation coefficients of -0.403, -0.401, respectively. While 10 variables showed a positive correlation, with the highest positive correlation recorded for the 20_RZSM_band1 (topsoil RZSM at 10cm) with r = 0.34, followed by the weighted SM and Modified Normalized Difference Water Index (MNDWI) with correlation of 0.33 and 0.21 respectively.

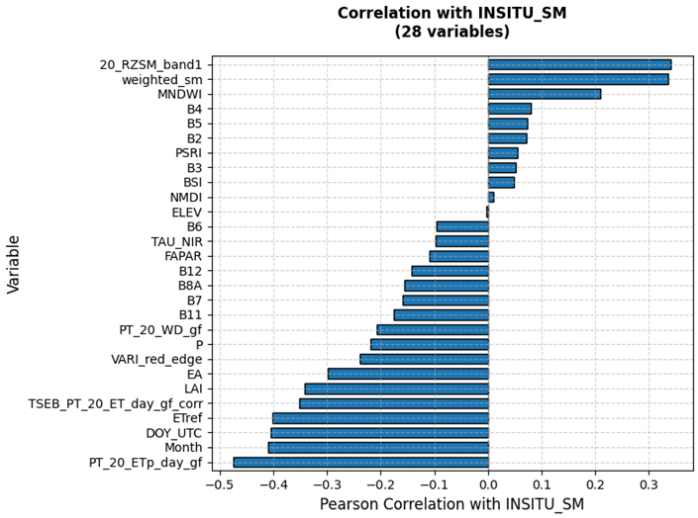


Figure 7: Bar plot with Pearson correlation between ground truth SM and the selected variables

- Model performance

  Figure 8 shows the model performance with a scatter plot of the actual SM and the predicted SM on the test dataset. The model achieved an average performance with Pearson correlation of 0.66, R^2 = 0.40, RMSE = 5.56, and MAE = 4.43. The green line indicates the regression fit, while the 1 to 1 line represents the ideal fit, highlighting the difference between the predicted and actual values.

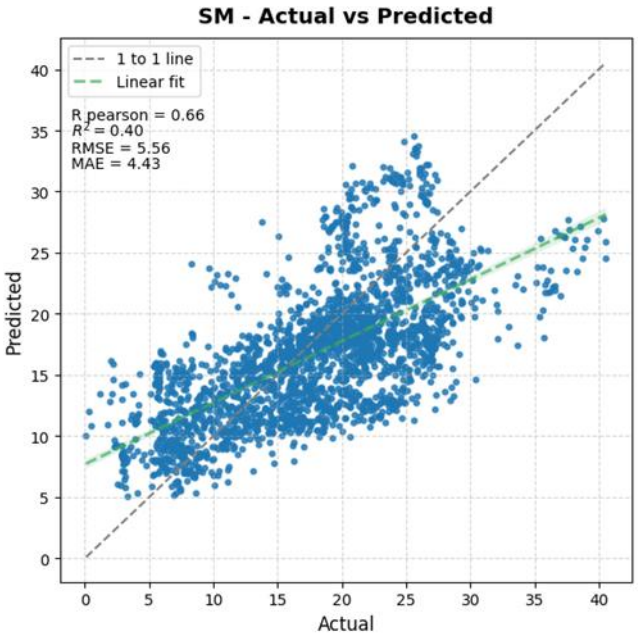


Figure 8: Scatter plot comparison of the actual and predicted daily SM on the test dataset using HGBR model.

- Feature importance

  Figure 9 presents a bar-plot illustrating the features importance in a descending order for HGBR. The importance was generated using the permutation approach to measures the contribution of each feature to fitted model's statistical performance and to overcome the limitations of the impurity-based feature importance. The results show that elevation, weighted SM, and DOY were the highest important ranked features followed by Normalized Multi-band Drought Index (NMDI) and Actual vapour pressure (EA).

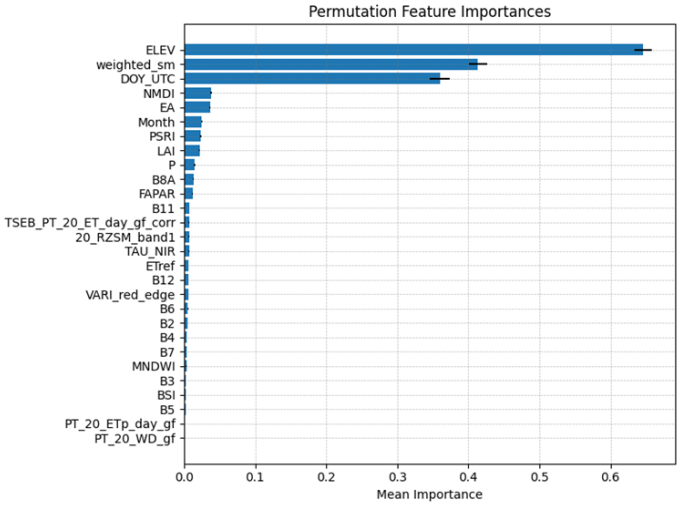


Figure 9: Variable importance of HGBR.

- Mapping SM
  Once validated and trained, the HGBR model was then applied to raster data to generate SM predictions at 20 m resolution. Figure 10 shows an example of the spatial prediction over the HOBE region for July 1, 2018. The map demonstrates a clear variability in soil moisture across different land cover types, highlighting the spatial dynamics of SM.
  This spatial variability reflects the influence of various factors such as vegetation density, soil type, and the topography on moisture retention and distribution. It is worth noting that the soil moisture is also temporally dynamic.

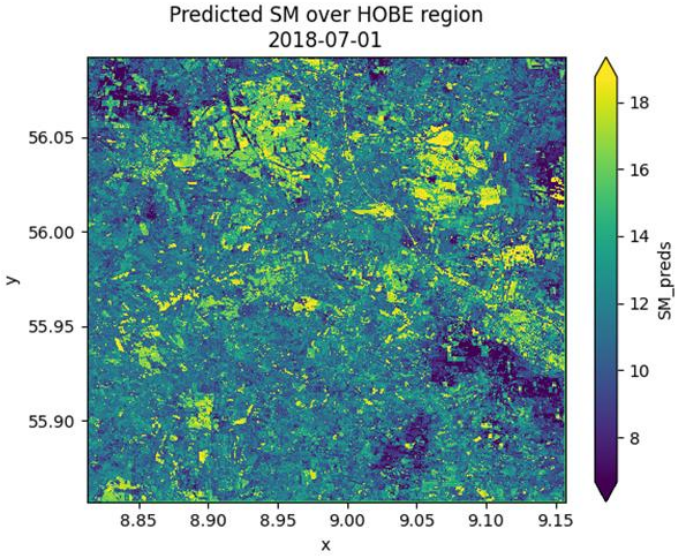


Figure 10: The spatial prediction of SM using HGBR over the HOBE region for 2018-07-01

Summary

This case study demonstrated a successful development of integrating earth observation and in-situ data coupled with machine learning to improve soil moisture modelling. As a baseline the HGBR model was trained and validated both spatially and temporally. The model achieved an average Pearson correlation of 0.66, R^2 = 0.40, RMSE = 5.56, and MAE = 4.43, outperforming the baseline modelled RZSM, which recorded a Pearson correlation of 0.7126, Test R^2: -0.4646, Test RMSE: 8.7114, and Test MAE: 7.0416.

The spatial cross-validation revealed a lack of reliability across different stations, which could be explained by the significant variability in ground soil moisture due to agricultural practices, land cover types, and micro-environmental factors. This presents some challenges in applying the ML approach for soil moisture modelling.

During this development, we released a python library called dhi-mleo (https://github.com/ScaleAGData/dhi-mleo/) that could be used for several applications to integrate in-situ measurements and earth observation data with machine learning workflows. It offers spatial prediction with Xarray, robust spatial and temporal cross-validation strategies, efficient handling of time series data, and a flexible workflow compatible with any machine learning model or task.

2.4 Super resolution methodologies

A key advancement in this D4.5, not yet available at the time of D4.3, is the development of a suite of super-resolution (SR) tools designed to enhance the spatial detail and analytical value of satellite-based EO products. These new tools provide RILs with the ability to derive higher-resolution

information from existing satellite missions, thereby improving the precision of downstream analyses and modelling workflows.

- **Sentinel-2 Super Resolution tool:** This tool improves the spatial resolution of Sentinel-2 (S2) data by combining the coarser 20 m and 60 m spectral bands with the finer 10 m bands. The result is sharper, more detailed imagery that preserves the original spectral quality, enabling RILs to monitor vegetation, land cover, and agricultural features with greater accuracy.
- **Sentinel-3 SLSTR Thermal Sharpening tool**: By integrating thermal data from Sentinel-3 (S3) SLSTR with super-resolved Sentinel-2 imagery, this tool produces high-resolution land surface temperature maps at 10 m resolution. This supports precise environmental monitoring and heat-related analyses, such as evapotranspiration modelling, helping RILs to make data-driven decisions.

Both tools are accessible via an API, allowing users to define specific Areas of Interest (AOI) and time windows. The primary constraint remains the availability of the necessary satellite imagery from the providers, but within these limits, the SR tools offer a powerful means to generate high-quality actionable EO data.

2.4.1 Technical implementation

2.4.1.1 Sentinel-2 Super Resolution Tool

The performance of super-resolution (SR) networks can be strongly influenced by both architectural differences and training hyperparameters. To systematically investigate this, three representative deep residual network architectures: DSen2[2], ResNet[3], and SPRNet[4]—were trained and evaluated using a single large training dataset. In addition, the effects of various training hyperparameters were examined to optimize both the quantitative and qualitative performance of SR models.

Based on this evaluation, DSen2 was selected as the final S2 SR tool since it achieved the second-best results and required a shorter training time. This tool aims to enhance the 20 m and 60 m resolution bands to a uniform 10 m resolution.

DSen2 is a CNN-based supervised DL model, which requires paired datasets containing both low-resolution inputs and their corresponding high-resolution ground truth targets. However, since S2 lacks 10 m reference data for the 20 m and 60 m bands, this limitation is addressed through a synthetic data generation strategy. The input images are downsampled according to the scaling ratio between the coarse and the fine bands prior to training. This procedure follows the principles of Wald's protocol[5] which assumes that the spectral correlation of an image is self-similar within a limited range of scales. Therefore, downsampling from 20 m to 10 m is simulated by resampling images from 40 m to 20 m, while the 60 m to 10 m case is represented by resampling from 360 m to 60 m. Under this assumption, an unlimited amount of training data can be produced by synthetically degrading original S2 images by the desired scaling factor (e.g., s=2,6) and using the original higher-resolution data as the training targets.

[2] Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E., & Schindler, K. (2018). Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. ISPRS Journal of Photogrammetry and Remote Sensing, 146, 305-319.
[3] Palsson, F., Sveinsson, J. R., & Ulfarsson, M. O. (2018). Sentinel-2 image fusion using a deep residual network. Remote Sensing, 10(8), 1290.
[4] Wu, J., He, Z., & Hu, J. (2020). Sentinel-2 sharpening via parallel residual network. Remote Sensing, 12(2), 279.
[5] Wald, L., Ranchin, T., & Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. Photogrammetric engineering and remote sensing, 63(6), 691-699.

Given that training datasets can be generated synthetically without requiring high-resolution reference images, two main training strategies have been developed. The first approach involves training the neural network on a large and globally distributed dataset, enabling it to generalize across diverse regions and to super-resolve any S2 scene without the need of retraining. The second approach trains and tests the model on the same image, which makes the process faster but limits its applicability. The first approach is considered more efficient since it requires a single training phase.

Dataset Creation

To train a deep neural network that is capable of super-resolving arbitrary S2 images without the need for retraining, we created a large and diverse training dataset. Atmospherically corrected S2 Level-2A tiles from 45 randomly selected locations (X2) were obtained from the Copernicus Data Space Ecosystem (CDSE), ensuring broad global coverage across different climate zones and land-cover types. Additionally, five S2 Level-2A tiles were collected for testing purposes. To simplify both training and evaluation, only images with low cloud coverage and without undefined ("black background") pixels were included. Of the total dataset, 80% of the images were used for training, and the remaining 20% served as a validation dataset.

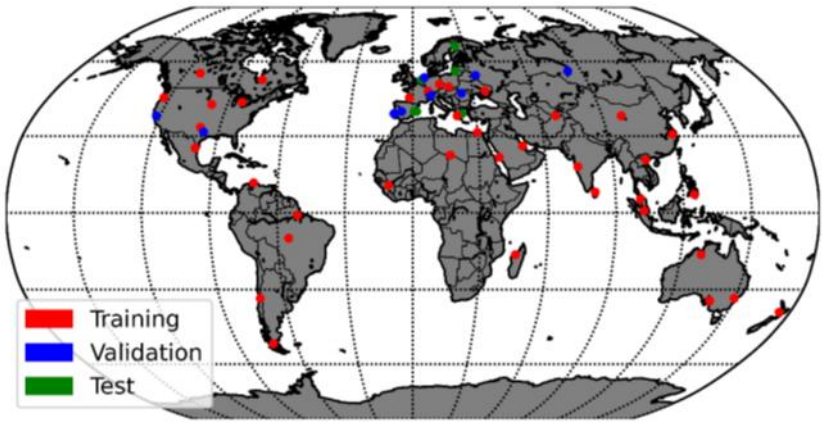


Figure 11: Locations of Sentinel-2 tiles selected for training and testing.

Due to the scale-invariance property, the mappings between 20->10 m and 40->20 m (for the 2x network), as well as between 60->10 m and 360->60 m (for the 6x network), can be considered roughly equivalent. Based on this property, two training datasets were created for training, as described Table 1.

The first dataset contains fine-resolution images at 20 m and coarse-resolution images at 40 m, obtained by down-sampling the original 10 m and 20 m bands by a factor of 2. This dataset was used for training a network of 2x super resolution. The second dataset comprises images at 60 m, 120 m, and 360 m resolution, generated by downsampling the original 10 m, 20 m, and 60 m bands, and was employed to train a network of 6x super resolution. After training on these synthetically downsampled data (40 m/20 m and 360 m/120 m/60 m, respectively), the CNN models were then applied to the original S2 images at their native resolutions (20 m/10 m and 60 m/20 m/10 m, respectively). It should be noted that, due to unavailability of 10 m ground truth data for the coarser bands, quantitative evaluation of the results was also performed at the reduced resolutions, and the resulting super-resolved 10 m bands were assessed only through visual inspection.

Table 1. Dataset Resolution Specifications

Scale Factor	Original Band Resolutio	Training Dataset Resolution
2x	B2, B3, B4, B8 (10 m) B5, B6, B7, B8A, B12 (20 m)	B2, B3, B4, B8 (20m) B5, B6, B7, B8A, B12 (40 m)
6x	B2, B3, B4, B8 (10 m) B5, B6, B7, B8A, B12 (20 m) B1, B9 (60 m)	B2, B3, B4, B8 (60 m) B5, B6, B7, B8A, B12 (120 m) B1, B9 (360 m)

For a clearer understanding of the dataset creation process, we introduce the following notation:

Let the observed fine resolution bands be denoted by $Y \in R^{d_1 \times d_2 \times L_1}$ and the observed coarse resolution bands by $X \in R^{\frac{d_1}{s} \times \frac{d_2}{s} \times L_2}$ where $d_1 \times d_2$ represents the spatial dimensions of the fine resolution bands, L_1 and L_2 are the numbers of fine and coarse bands respectively and s is the scale ratio between fine and coarse bands. Down-sampling a factor s is denoted by the operator *D*, and up-sampling a factor s is denoted by the operator *U*. Finally, we denote the spatially degraded fine and coarse bands by *DY* and *DX* respectively. Since *DX* is smaller than *DY* by a factor of s along each spatial dimension, it must be interpolated to match the size of *DY*, before being combined elementwise with the fine resolution information and compared to DY during training. The interpolated degraded coarse bands are thus denoted by $X^D \in R^{\frac{d_1}{s} \times \frac{d_2}{s} \times L_2} = UDX$, while the degraded fine bands are denoted by $Y^D \in R^{\frac{d_1}{s} \times \frac{d_2}{s} \times L_2} = DY$. After this processing, *XD, YD* and *X* share the same spatial dimensions enabling elementwise operations during the training of the super resolution network.

In practice, to generate training data with the desired scale factors of 2 and 6, the original S2 images were first blurred using a Gaussian filter with a standard deviation of $\sigma = \frac{1}{s}$ pixels. The images were then downsampled by averaging over s × s windows, where s = 2 and s = 6, respectively.

To make the training of the networks computationally feasible, each input image was divided into many small patches of appropriate size. The network therefore takes input patches that combine the degraded coarse and fine resolution bands. Each patch corresponds to the same spatial region across all bands, and the corresponding target patch is taken from the original high-resolution image. In practice, 1,000 random patches were sampled from each training image. To determine the optimal patch size during hyperparameter tuning, three separate training sets of 45,000 patches each were created, with patch sizes of 32x32, 64x64, and 128x128 pixels.

The dataset processing was carried out using Python. Gaussian blurring was applied via the gaussian filter function from the scipy.ndimage package for multidimensional image processing. Downscaling by averaging was performed using the skimage.measure.block_reduce function. Coarse image patches were subsequently upscaled to match the dimensions of the fine resolution patches using the skimage.transform.resize function, which employs spline interpolation with a default order of 1. Finally, the splitting of the images into patches was achieved using the gdal.Translate function.

CNN architecture

Three different CNN architectures were evaluated to identify the model achieving the fastest convergence and best performance: i) DSen2 proposed by Lanaras et al. [2], ii) ResNet proposed by Palsson et al. [3], and iii) SPRNet proposed by Wu et al. [4]. The three neural network architectures are presented in Figure 3. All networks were implemented in the Keras framework with TensorFlow as the backend, using the same training dataset and hyperparameter configuration for fair comparison.

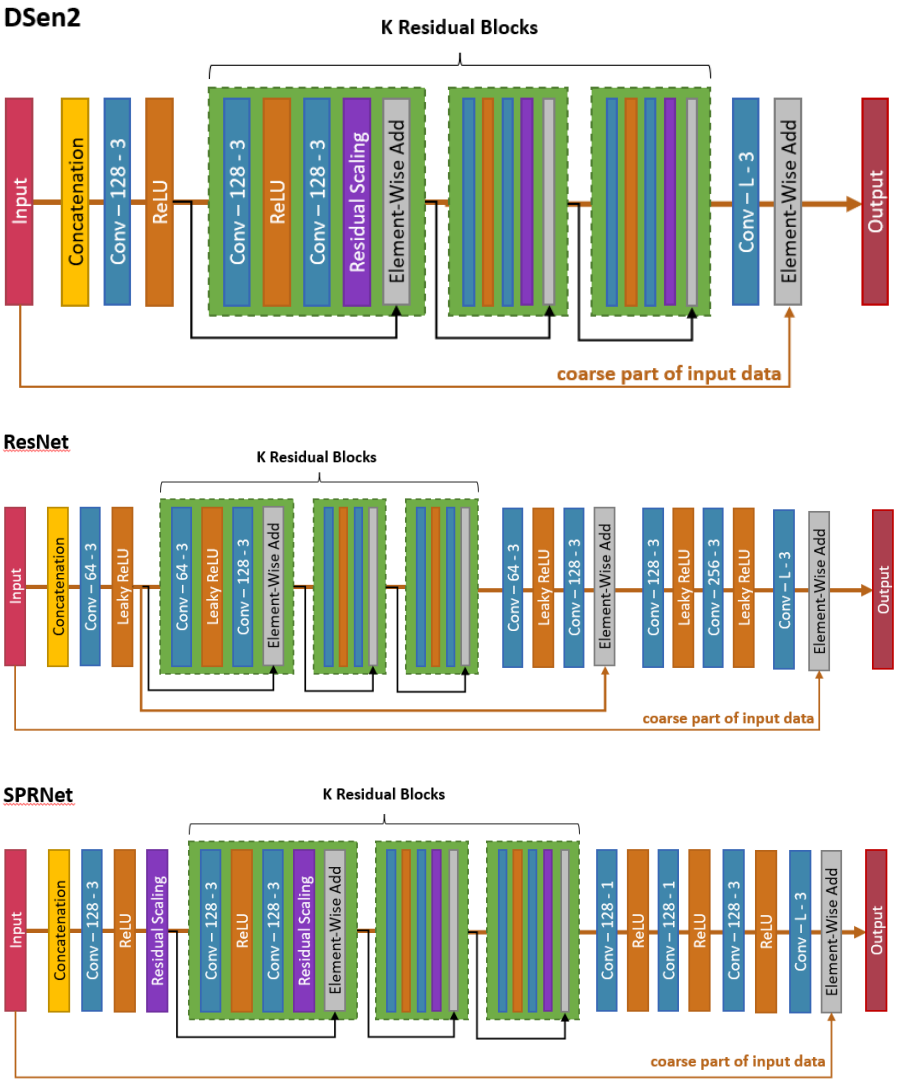


Figure 12: Sentinel-2 SR Neural Network Architectures

Training and hyperparameter tuning

The training dataset consisted of 45 S2 images, which were divided into 45,000 patches. For numerical stability, the raw reflectance values of each image patch were normalized by dividing by 2,000 before training.

The networks were initially trained for 30 epochs using a mini-batch size of 128 patches and a patch size of 32x32 pixels. Mean Absolute Error (MAE) was employed as the loss function, while the ADAM variant of Stochastic Gradient Descent (SGD) with Nesterov momentum was selected as the optimizer. The network weights were initialized to small random values using the He uniform method[6]. The initial learning rate was set to 1e-4 and reduced by a factor of 2 whenever the validation loss remained stable for 5 consecutive epochs.

To achieve optimal performance, we systematically investigated the effects of key hyperparameters on both quantitative and qualitative results. First, four different batch sizes were tested while keeping

[6] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision (pp. 1026-1034).

the other parameters constant, including a patch size of 32 pixels, MAE as the loss function, 6 residual blocks, ADAM optimizer with Nesterov momentum, an initial learning rate of 1e-4, and 50 training epochs. Figure 13 demonstrates the Spectral Angle Mapper (SAM) values achieved with each batch size configuration.

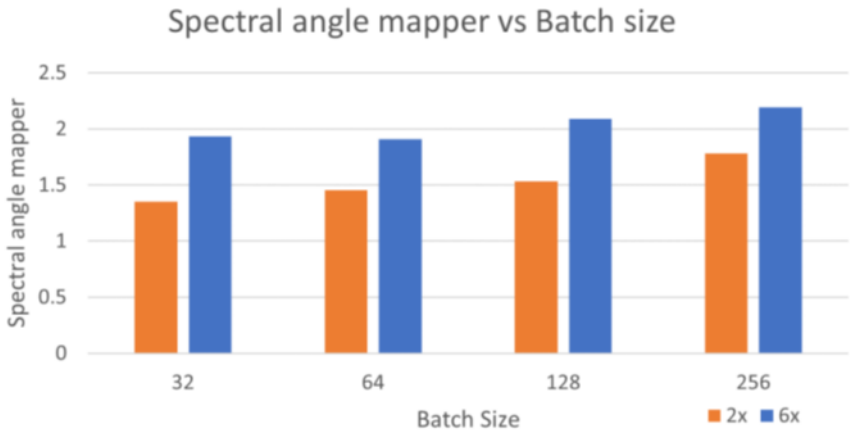


Figure 13: Spectral Angle Mapper achieved with different batch sizes

Subsequently, three different patch sizes were evaluated: 16, 32, and 64 pixels. This investigation aimed to determine the optimal spatial context for training while maintaining consistent settings for the other hyperparameters. Figure 14 presents the SAM values obtained with each patch size configuration.

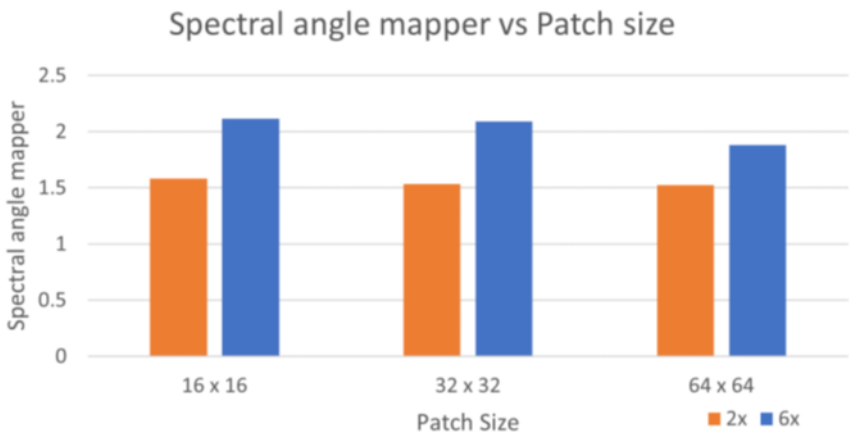


Figure 14: Spectral Angle Mapper achieved with different patch sizes

Finally, four loss functions were compared: Mean Absolute Error (MAE), Mean Squared Error (MSE), Huber Loss, and Relative Dimensionless Global Error (ERGAS). Figure 15 illustrates the SAM values achieved with each loss function under the same training conditions.

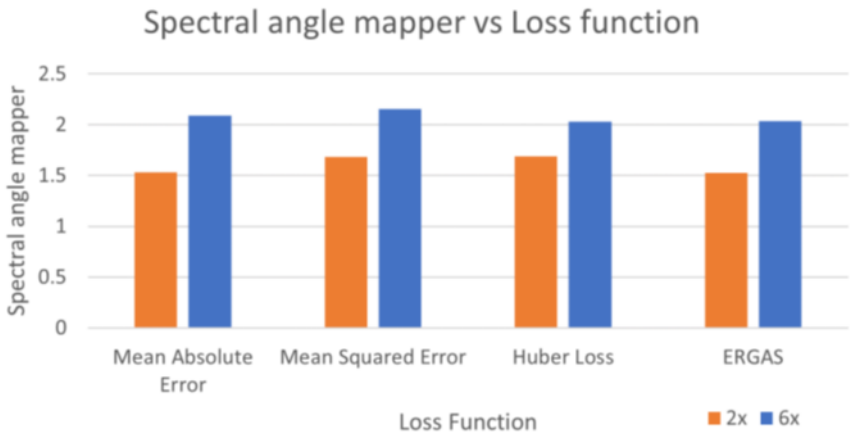


Figure 15: Spectral Angle Mapper achieved with different loss functions

Based on the hyperparameter tuning experiments, distinct optimal configurations were identified for each scale factor (2x, 6x). For 2x super resolution (20 m to 10 m), the optimal configuration employed patches of 32x32 pixels, while for the 6x super resolution (60 m to 10 m), patches of 64x64 pixels were used. Both configurations utilized a batch size of 32 patches, the MAE loss function, and the ADAM optimizer with Nesterov momentum. The learning rate was set to 1e-4 and reduced by a factor of 2 when validation loss was stable for 5 epochs. The models were trained with this optimal setting for 100 epochs. The training curves for both scale factors are presented in Figure 16 and Figure 17, respectively.

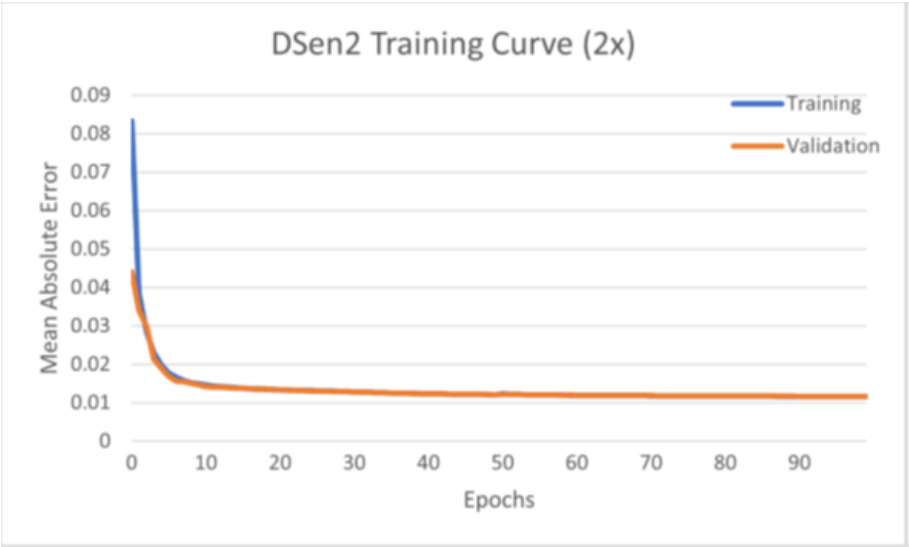


Figure 16: Training curve of DSen2 for the 2x network

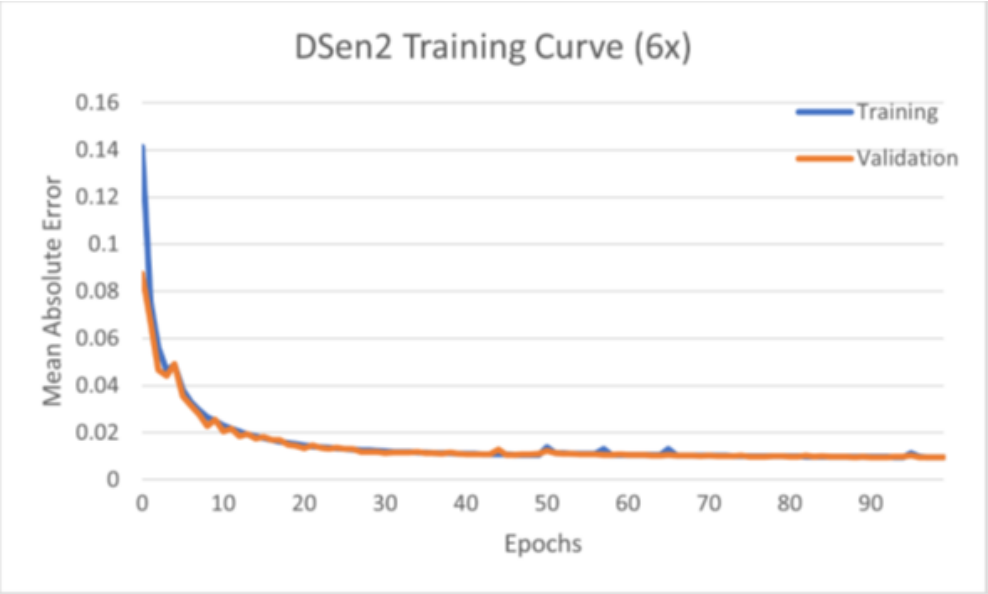


Figure 17: Training curve of DSen2 for the 6x network

2.4.1.2 Thermal Sharpening of Sentinel-3 SLSTR

Thermal sharpening methods typically enhance TIR imagery by using empirical relationships between Land Surface Temperature (LST) and spectral signals (e.g., the relationship between temperature and the Normalized Difference Vegetation Index NDVI). This relationship is derived at the coarser thermal resolution and then applied at a finer shortwave resolution to create the sharpened image. Most approaches rely on the assumption that this relation is unique across a given imaging scene and for various land-cover types; therefore, they are applicable mainly on homogeneous landscapes.

A thermal sharpening tool was developed based on the Data Mining Sharpener (DMS) technique proposed by Gao et al.[7]. Unlike traditional approaches, DMS does not rely on a predefined relation between thermal infrared (TIR) data and shortwave spectral reflectance; instead, it learns it directly from the data. This data-driven method provides strong adaptability across various land-cover types and heterogeneous scenes. It constructs regression trees between (TIR) data and shortwave spectral reflectance. Building on this concept, the model that was developed extends the DMS framework to enhance S3 thermal data from 1000 m to 10 m spatial resolution by integrating information from S2 images.

The implemented workflow requires two primary inputs: a coarse-resolution S3 SLSTR thermal image and a fine-resolution, atmospherically corrected (Level 2A) S2 image. Both images must be cloud-free and temporally close. The S2 data are first super-resolved using the SR tool described in section 2.4.1, resulting in all spectral bands being resampled to a uniform 10 m spatial resolution. The S3 data are then re-projected to the S2 coordinate reference system using the GDAL Python library, and only the overlapping regions between the two datasets are retained for further processing.

In the sharpening process, the high-resolution Sentinel-2 reflectance bands are used as independent variables, while the S3 thermal data serve as the target variable. Within each leaf of the regression tree, a linear regression is fitted between the thermal and reflectance variables, allowing the model to capture locally linear relationships. Although higher-order functions could model temperature

[7] Gao, F., Kustas, W. P., & Anderson, M. C. (2012). A data mining approach for sharpening thermal satellite imagery over land. Remote sensing, 4(11), 3287-3319.

variations more precisely, they risk overfitting and noise sensitivity. To ensure compatibility, Sentinel-2 reflectance are aggregated to match the coarse SLSTR spatial resolution. Sub-pixel variability within each SLSTR pixel, derived from the aggregated Sentinel-2 data, is used to identify representative training samples that link LST with fine resolution reflectance. These samples are used to train an ensemble regression model that captures the statistical relationships between thermal and reflectance signals.

Both local and global regressions are performed using the scikit-learn library, through an ensemble of decision trees trained on random subsets of the data, following the Bagging approach[8]. The local model is implemented using a moving window: within each predefined window, samples selected are used to train local regressions that are then applied to the coarse pixels inside that window to make predictions on the fine resolution. The global model is trained on the entire image. Predictions from both models are compared to the original LST values to compute residuals, and a combined prediction is formed by weighing the local and global results according to their residuals (smaller residual ~ larger weight). Finally, residual and bias corrections are applied to ensure consistency between the sharpened high-resolution imagery.

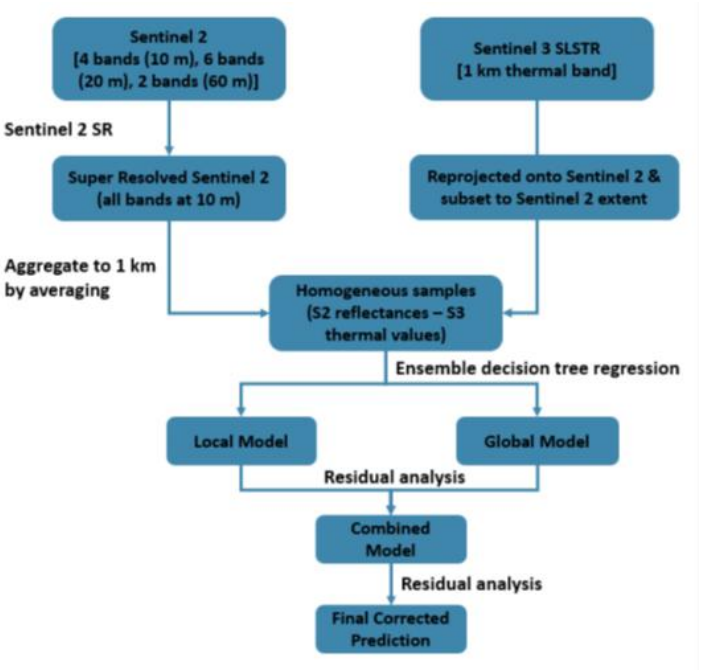


Figure 18: Sentinel-3 SLSTR sharpening workflow

2.4.2 Validation

2.4.2.1 Sentinel-2 SR tool

After determining the optimal network configuration and training setup, the three models were evaluated using multiple quantitative metrics. Evaluation metrics were computed both during validation (at the end of each training epoch) and in the final testing phase, including the Mean Absolute Error (MAE), Peak Signal-to-Noise Ratio (PSNR), Spectral Angle Mapper (SAM), Relative Dimensionless Global Error (ERGAS), and Universal Image Quality Index (UQI).

[8] Breiman, L. (1996). Bagging predictors. Machine learning, 24(2), 123-140.

The evaluation results summarized in Table 2 and Table 3 indicate that the SPRNet architecture achieved the highest overall accuracy but required significantly longer training time and higher computational complexity. The DSen2 model provided comparable accuracy while being considerably more efficient in terms of training time and inference speed.

Based on this trade-off between accuracy and computational efficiency, DSen2 was selected as the optimal architecture for integration into the Super-Resolution tool and subsequent validation with Sentinel-2 images.

Table 2. Evaluation results for the three architectures

	ERGAS	MAE	PSNR	SAM	SRE	UIQ
SPRNet	0.991924	0.013503	33.457920	0.937818	40.766830	0.999572
ResNet	1.376152	0.018745	30.612630	1.301025	39.407833	0.999177
DSen2	1.130744	0.015817	32.318603	1.064090	39.190921	0.999445

Table 3. Training time comparison of the evaluated models

	Epochs	Time per Epoch	Training Duration
SPRNet	30	1h 49 min	54h 30 min
ResNet	30	26 min	13h
DSen2	30	45 min	22h 30min

The final evaluation of the selected DSen2 model was carried out separately for the two resolution scales, 2x and 6x on the independent test dataset.

Across both scaling factors, DSen2 achieved consistently good performance across all evaluation metrics, as shown in Table 4. The 2x super-resolution provides clear improvements in spatial resolution with minimal artefacts, while the more challenging 6x case still maintains robust performance, accurately recovering fine structures and textures.

Table 4. Evaluation results of the optimal DSen2 (2x, 6x) networks

Scale Ratio	ERGAS	MAE	PSNR	SAM	UIQ
2x	1.4961	0.0121	34.7523	1.3308	0.9991
6x	1.4763	0.0159	32.2598	1.9371	0.9986

To further assess the quality of super-resolution results, a qualitative evaluation was performed on selected Sentinel-2 scenes using the SR tool. The tool is capable of super-resolving entire Level-2A images in .SAFE format, processing each 20 m and 60 m band to 10 m resolution. The final super-resolved image includes 12 bands at 10 m spatial resolution, and it is stored in the output file prefixed with 'SR_'. The results of the Sentinel-2 super-resolution tool can be visually assessed in Figure 19, Figure 20 and Figure 21 where the original and super-resolved Sentinel-2 bands are shown for qualitative comparison.

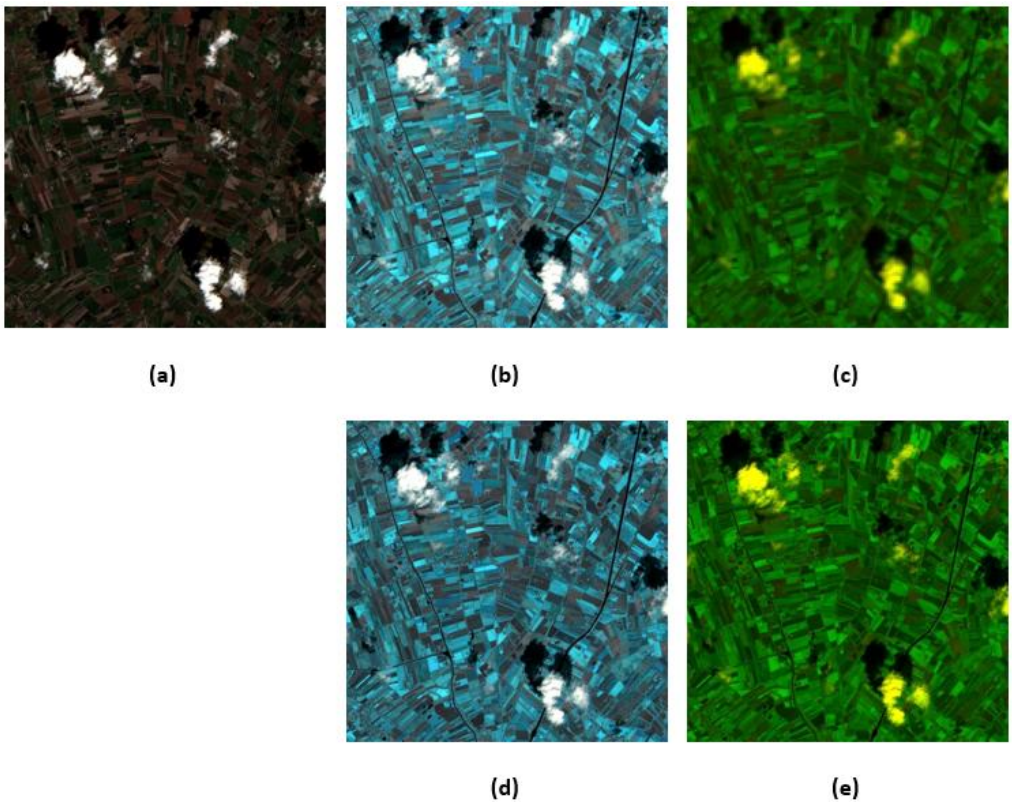


Figure 19: Top row: Original Sentinel-2 bands at (a) 10m (B4, B3, B2), (b) 20m (B5, B6, B7), (c) 60m (B1, B9). Bottom row: Super-resolved bands at 10 m resolution (d) B5, B6, and B7 and (e) B1 and B9.

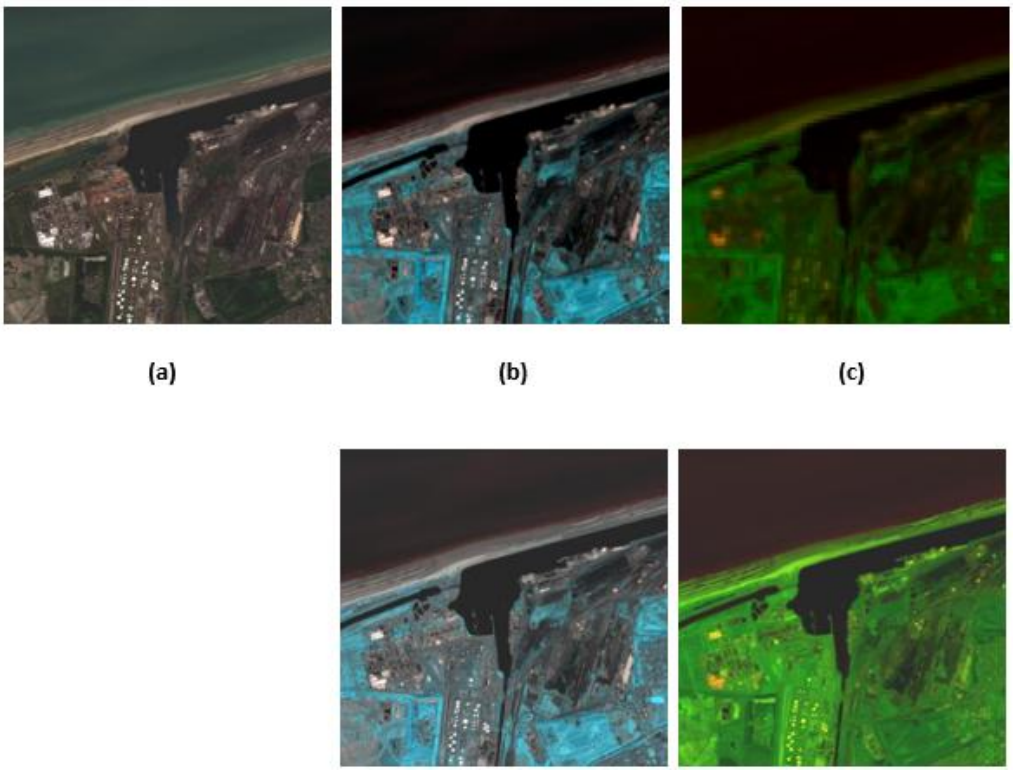


Figure 20: Top row: Original Sentinel-2 bands at (a) 10 m (B4, B3, B2), (b) 20 m (B5, B6, B7), (c) 60 m (B1, B9). Bottom row: Super-resolved bands at 10 m resolution B5, B6, B7 and B1, B9.

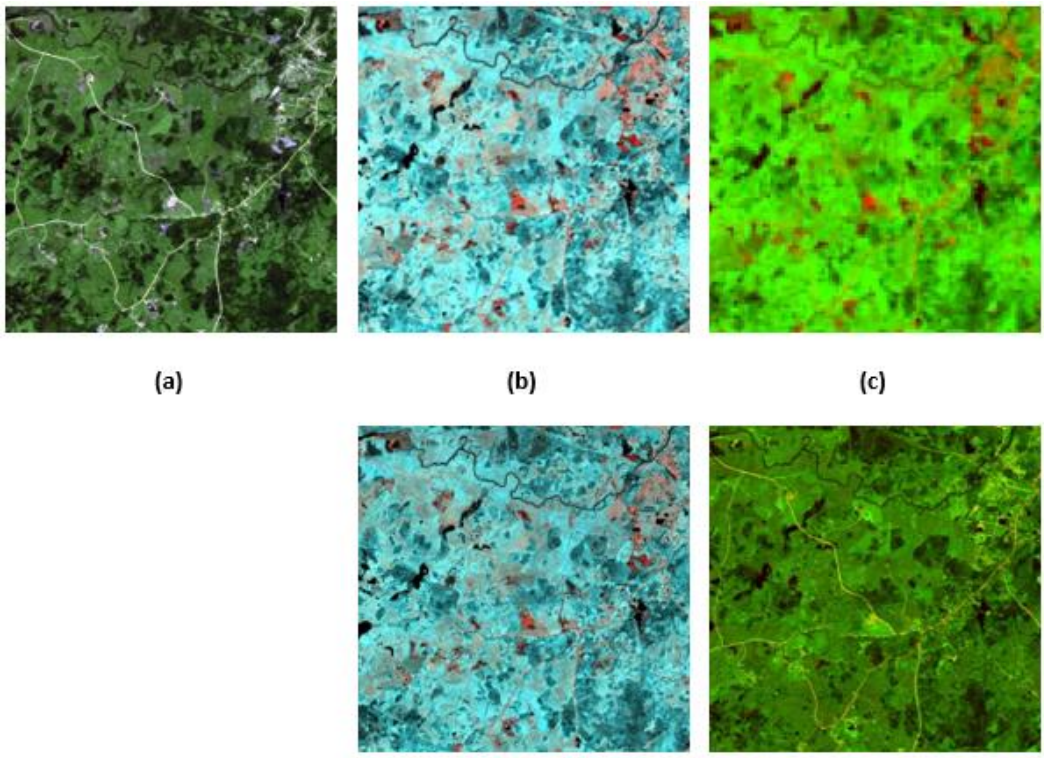


Figure 21. Top row: Original Sentinel-2 bands at (a) 10m (B4, B3, B2), (b) 20m (B5, B6, B7) (C) 60m (B1, B9) Bottom row: Super-resolved bands at 10m resolution B5, B6, B7, and B1, B9

Validation with data provided by the Water Productivity RIL

To validate the performance of the super resolution tool using very high-resolution data, a hyperspectral image with 48 channels and a spatial resolution of 0.5 m/ pixel provided by the Water Productivity RIL was utilized for pixel wise comparison. A dedicated pre-processing class was developed to prepare both the super-resolved Sentinel-2 image and the hyperspectral image. The first step in the pre-processing involved atmospheric correction of the hyperspectral image, converting the original radiance values (in milli-SRU) into reflectance values (ranging for 0 to 1), to match the scale of the super-resolved Sentinel-2 image. Next steps included spatial and spectral alignment, which will be described in detail below.

The atmospheric correction model was implemented using Py6S[9], a Python interface of the 6S radiative transfer model. The model was defined for each specific wavelength (ranging from 407.256 nm to 990.653 nm), date, sensor altitude and ground elevation. The solar geometry and sensor's viewing angle were manually set for the date of acquisition. The atmospheric profile was selected based on the latitude and acquisition date, while the aerosol effects were defined using a continental aerosol profile, with a specified atmospheric visibility. Ground reflectance for the model initialization was assumed to be homogeneous Lambertian surface. From each band, the input radiance was first converted from milli-SRU to SI units ($W/m^2 \cdot sr \cdot \mu m$). The model outputs atmospheric correction coefficients, which were then applied to compute surface reflectance using the Lambertian formula.

To validate the atmospheric correction, the mean radiance for each band was compared with the mean reflectance after correction for the corresponding band. As observed in the plot, the corrected values generally follow the original trend and both curves exhibit a similar trend.

[9] https://py6s.readthedocs.io/en/latest/

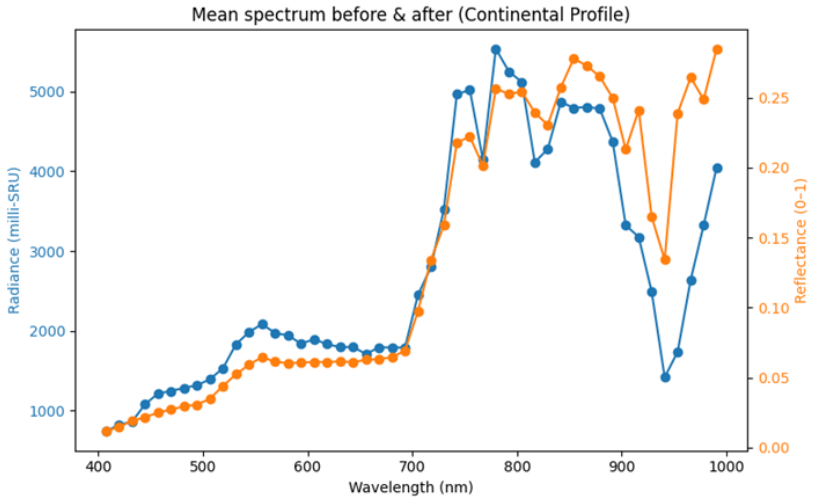


Figure 22. Comparison of mean radiance and mean reflectance across spectral bands after atmospheric correction

To further validate this process, three random pixels were selected from the most representative classes in the image (vegetation, soil, urban), and their spectral signatures were plotted. As seen in the plot, the spectral signatures curves generally follow the expected patterns for each class, with vegetation showing the typical red-edge, and soil and urban exhibiting relatively flat spectra. However, in the extreme bands at the lower and upper ends of the spectrum, the patterns start to deviate, showing increased fluctuations. The validation metrics for these bands were also low, so they were excluded from the final evaluation.

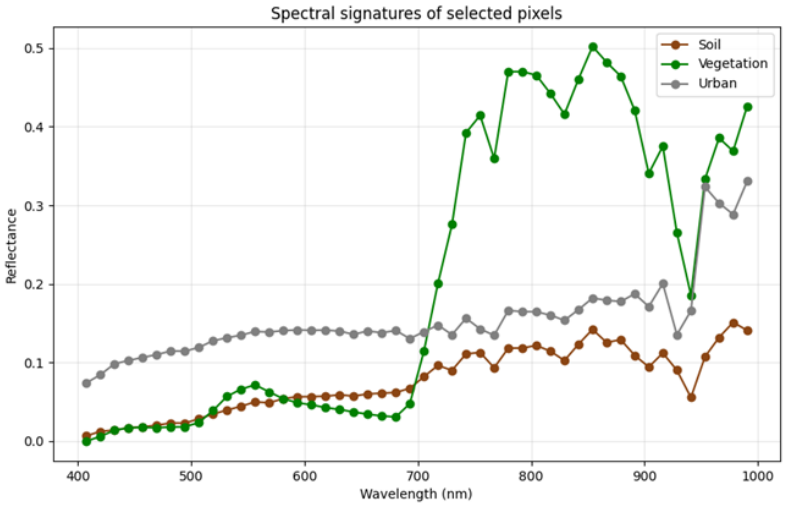


Figure 23. Spectral signatures of selected pixels after atmospheric correction

After performing atmospheric correction on the hyperspectral data, a pre-processing framework was developed to spatially and spectrally align the Sentinel-2 and hyperspectral images for pixel-wise comparison. A band mapping procedure was employed to identify the hyperspectral bands that best correspond to Sentinel-2 bands based on central wavelength proximity. In hyperspectral imaging, the extreme bands at the lower and upper ends of the spectrum are often less reliable due to sensor sensitivity and higher noise levels. These bands tend to be more affected by atmospheric effects, so

for the purpose of this analysis, the extreme bands (443-445 nm and 945-941nm) were excluded, and the focus was placed on the mid-range bands. The bands that were chosen for the final comparison are presented in Table 5.

Table 5. Mapping of Sentinel-2 bands to corresponding hyperspectral wavelengths

Sentinel-2 wavelength	Hyperspectral wavelength
443 nm	444.5 nm
490 nm	494.2 nm
560 nm	556.3 nm
665 nm	668 nm
705 nm	705.3 nm
740 nm	742.5 nm
783 nm	779.7 nm
842 nm	841.8 nm
865 nm	866.6 nm
945 nm	941.0 nm

To ensure spatial consistency with Sentinel-2 data, the hyperspectral image (originally acquired at spatial resolution of 0.5 m/pixel) was downsampled to 10 m/pixel using an averaging resampling method implemented with the *rasterio* library. Subsequently, the super-resolved Sentinel-2 image was spatially aligned and cropped to match the extent and resolution of the resampled hyperspectral data. All spatial metadata, including the coordinate reference system (CRS) and geotransform parameters, were preserved during the alignment process.

After the pre-processing and atmospheric correction, the validation was performed using MAE, RMSE, Pearson and Spearman correlation coefficients, UQI and SSIM metrics. The results are presented in Table 6.

Table 6. Pixel-wise validation metrics for Sentinel-2 vs. hyperspectral reference

S2 wavelength	HS wavelength	MAE	RMSE	Pearson_r	Spearman_r	UQI	SSIM
490 nm	494.2 nm	0.145	0.145	0.804	0.841	0.280	0.319
560 nm	556.3 nm	0.134	0.135	0.534	0.471	0.322	0.493
665 nm	668 nm	0.106	0.109	0.795	0.822	0.407	0.494
705 nm	705.3 nm	0.153	0.156	0.614	0.537	0.422	0.520
740 nm	742.5 nm	0.147	0.157	0.901	0.859	0.784	0.614
783 nm	779.7 nm	0.151	0.163	0.909	0.865	0.814	0.618
865 nm	866.6 nm	0.155	0.168	0.904	0.864	0.815	0.632

Although the pixel-wise comparison cannot fully capture spectral consistency due to sensor limitations and empirical pre-processing, the results still indicate that the applied corrections achieved a reasonable level of agreement between the two images. The MAE and RMSE values are generally low across the mid-range bands, indicating that super-resolved Sentinel-2 images closely match the hyperspectral reference in terms of pixel-wise numerical agreement. Both Pearson and Spearman

correlation coefficients are moderate to high across mid-range wavelengths, suggesting that the relationship between the two images is primarily monotonic, but not perfectly linear. Finally, UQI and SSIM values are low to moderate for the shorter wavelengths, while the longer wavelength bands are relatively high in values, indicating improved structural similarity.

2.4.2.2 Sentinel-3 SLSTR sharpening tool

The developed thermal sharpening tool was applied to S3 and S2 image pairs collected over 5 AOIs. Table 7 summarizes the percentage of homogeneous samples extracted for each image pair, along with the RMSE obtained from the residual analysis between the sharpened (fused) products and the original coarse resolution S3 thermal image.

Table 7. Evaluation results of the thermal sharpening tool over AOIs

Case Area	Homogeneous Samples	RMSE
Netherlands	40.6%	1.85857
Lithuania	86.8%	1.97244
Spain	63.6%	1.11956
Greece	86.4%	1.88751
Finland	96.6%	0.65909

Figure 24, Figure 25, and Figure 26 present qualitative evaluation of the sharpening tool for three AOIs, based on Sentinel-2 and Sentinel-3 acquisitions from September 2024.

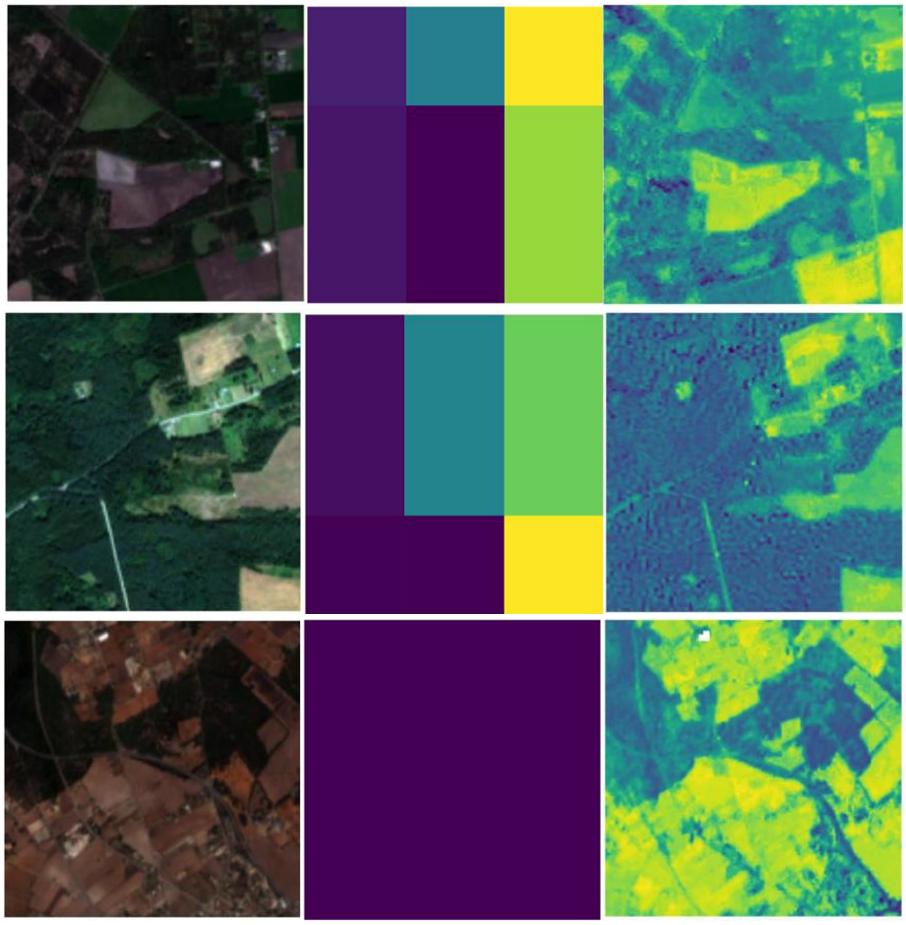


Figure 24: Illustrates representative subsets of the Sentinel-2 and Sentinel-3 pairs, together with the corresponding predicted images generated by the thermal sharpening tool.

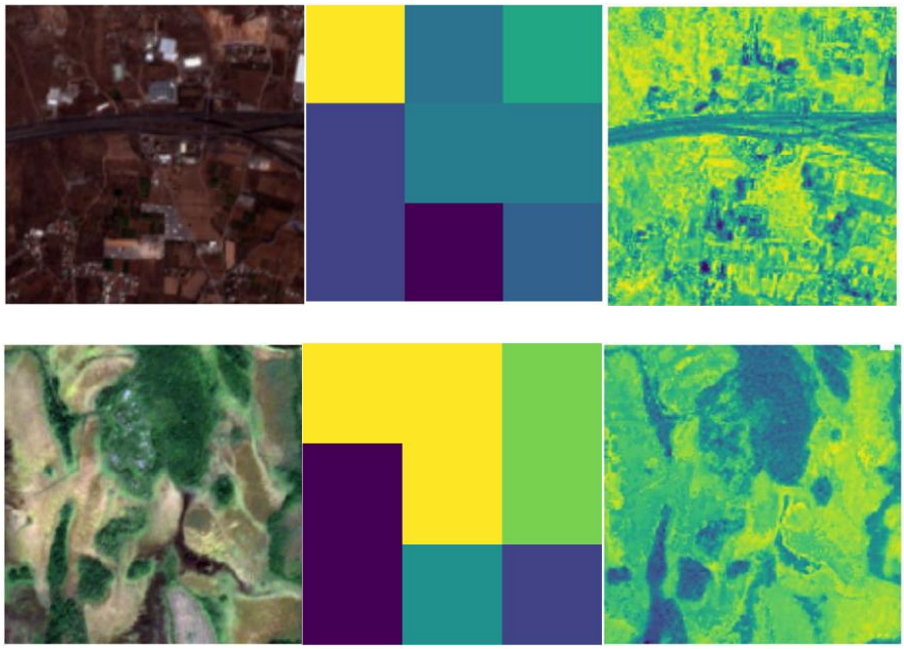


Figure 25: Top to bottom: Sentinel-2 super resolved RGB image, Sentinel-3 SLSTR thermal image at 1km resolution, Thermal sharpened image at 10m resolution

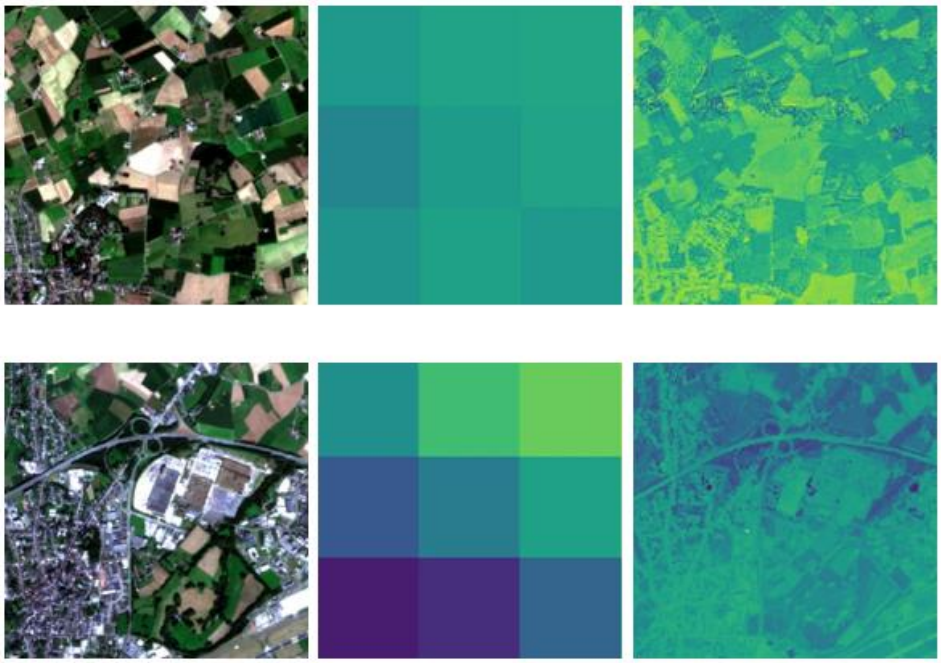


Figure 26. Top to bottom: Sentinel 2 super-resolved RGB image, Sentinel-3 SLSTR thermal image at 1km resolution, Thermal sharpened image at 10 m resolution.

Use cases

In the Yield Monitoring RIL the Sentinel-2 Super Resolution and Sentinel-3 Thermal Sharpening products will be evaluated for:

- Potato tare estimation: to check within-field variability, correlation with soil moisture and tare weight. Two potato fields, monitored in 2024, have been selected for the test.

- Potato monitoring & yield estimation: as additional input to the potato yield model that is being developed in the Yield Monitoring RIL (see section 3.3.1), and/or for product validation, to check within-field variability, possible causes of variations in yield.

2.5 Timeseries forecasting framework

A major addition in this second iteration, not yet included in D4.3, is the development of a dedicated time series forecasting framework. Time series forecasting plays an important role in enabling data-driven and proactive decision-making in agricultural and environmental systems. By analysing historical patterns and trends from meteorological data or model outputs, satellite observations and sensor-based measurements, forecasting models can predict future states of key parameters. These predictions support informed actions in domains such as resource management, strategic planning, and operational decision-making, ultimately enhancing efficiency, resilience, and sustainability.
Based on the correlation analysis of parameters along the dairy chain conducted in joint work of WP4 and WP5 in the Sustain Dairy RIL during the first iteration cycle of the project, the involved stakeholders and project partners identified the need for a framework for time series forecasting. Following an empirical validation of its feasibility in the Sustain Dairy RIL context, the development of this framework started with the second iteration project cycle. As time series forecasting models are already well established, offering a wide range of proven statistical and machine-learning approaches for capturing temporal dynamics, the focus of the development was not on new model architectures but on designing a framework for Near Real Time (NRT) predictions, using ensembles of established and customized models with time lagged model set selection.

2.5.1 Technical implementation and innovations

The implementation of the time series forecasting framework wraps around the python-based library darts[10] that itself wraps and expands time series forecast models of well-known libraries such as pytorch, scikit-learn, scipy, statsmodels and many more. Darts already offers the option to run a set of models as an ensemble. However, all these libraries lack automated model selection, an approach which could improve the forecast if the performance of the models in the ensemble set varies over time. Therefore, the set of models in the ensemble may be optimized and updated, so that for a final forecast only a subset of all individual models is used for the ensemble. Especially for NRT forecast in Sustain Dairy RIL, it was found to be beneficial to recurrently run ensemble forecasts only with the best subset of all defined models, based on the criteria which individual models performed best a lagged time interval before the actual forecast time. Therefore, the core component of the time series forecasting framework is a methodology for automated selection of optimal model sets for ensemble forecasts in an NRT context based on short term historical model performance. Furthermore, the development encompassed the customization of pre-implemented models to account for seasonal effects (e.g. in linear regression models).

The definition of the ensemble of the models is the key aspect in the framework, which are specified by the following parameters:

- A set of n time series forecasting models
- A time step interval t to evaluate model performances retrospectively
- An error metric for performance evaluation
- A model fusion metric

[10] Time Series Made Easy in Python — darts documentation

Figure 27: Methodology of the time series forecasting framework building ensemble models from best performing members by time lagged validation.

Figure 27 depicts the methodology of model selection in the framework. Each ensemble model in the framework takes a set of n time series forecasting models covering the model type and all hyper parameters of the individual models. For each forecast time, all individual models in the set are separately trained to include the most recent data, as the inclusion of the last time steps highly impacts the forecast's accuracy. Each forecast time, the model outputs are evaluated against validation data using diverse error metrics (e.g. RMSE, MAE, …). In an NRT application, this validation is conducted as soon as validation data becomes available. The ensemble models of the framework evaluate the performance of all the individual models for t timesteps before the actual forecast time and select the most performant model parameters as a subset of the n models. Based on the fusion metric, the outputs of the chosen individual models at the actual forecast time are merged into one single time series as the forecast of an ensemble.

Generally, the framework can be run on any set of variables, provided that enough training data for the target and covariates required by each individual model is available. However, the focus of this framework is on joint usage with the meteorological datasets in Zarr. format, optimized for time series extraction, which were developed in close connection and synergy under WP3 activities. Therefore, meteorological variables can easily be integrated as future and past covariates. The processing pipeline can be recurrently triggered and therefore process NRT forecasts when deployed on OHB resources as implemented in the Sustain Dairy RIL.

2.5.2 Use cases, validation and results

The time series forecasting framework is successfully applied in the Sustain Dairy RIL and further developed based on the feedback of partners and in synergy with WP3, WP4, and WP5 activities. Within the Sustain Dairy RIL, the framework enables NRT forecasts of milk quality on the regional level using ECMWF IFS and ERA5 meteorological datasets developed in WP3 as covariates.

Figure 28 depicts the processing pipeline of the Sustain Dairy RIL milk quality forecast use case. The purple box shows the processing steps of the overall time series forecasting framework, starting with the target and covariate datasets optimized for time series extraction. The extracted time series are merged in the processing step of running the individual forecast models, indicated as the red box. Individual models are run, and their results are saved in one common data cube. Next, the results are validated against historical values of the target data cube with diverse metrics and results saved in a separate data cube. As the last step of the processing in the framework, the ensemble models generate their outputs based on the recent individual model outputs and the lagged validation results.

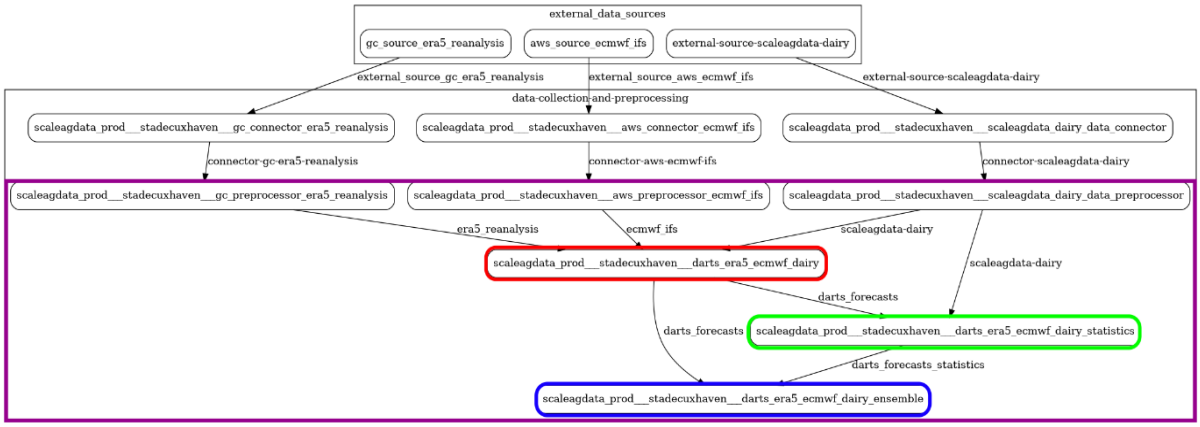


Figure 28: Processing pipeline in the Sustain Dairy RIL use case. The purple box indicates processing steps of the time series forecasting framework.

With milk quality data being available from 2018 onwards, and historical ECMWF IFS 360h forecast data being available starting 13th Nov 2024 twice daily, the performance of the framework could be tested and evaluated for a sufficiently long-time interval. As the forecast horizon depends on the forecast horizon of ECMWF IFS forecasts, which is 15 days, individual model and ensemble model outputs cover this time interval. A typical ensemble forecast with validation data is shown in Figure 29.

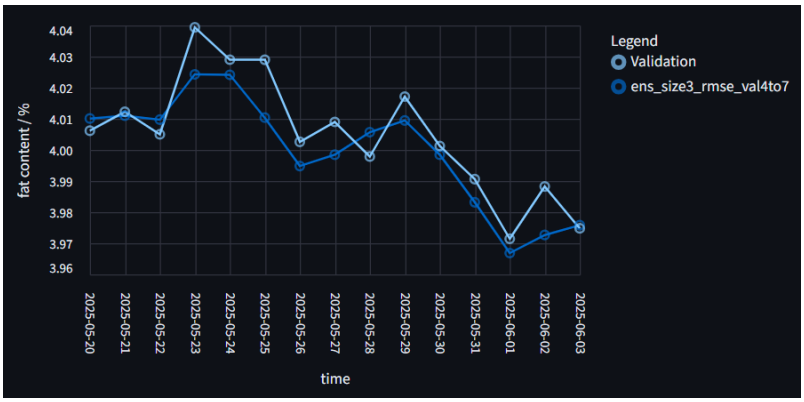


Figure 29: Example of a forecast output including validation data of an ensemble model based on time lagged model set selection in RIL6.

Further, the performance of the ensemble forecasts was investigated by comparing RMSE of average individual models and average ensemble models, as shown in Figure 30**Error! Reference source not found.**. RMSE values for the predicted milk quality parameter "fat percentage" are shown over time and as a histogram covering the period from 4th Dec 2024 - 9th Aug 2025. The evolution over time exhibits that the ensemble models of the framework outrun the simple ensemble mean of all implemented individual models for most periods. However, there are intervals in which the ensembles lack performance, but the histogram shows that the ensemble forecasts of the framework show lower RMSE values for most of the dates.

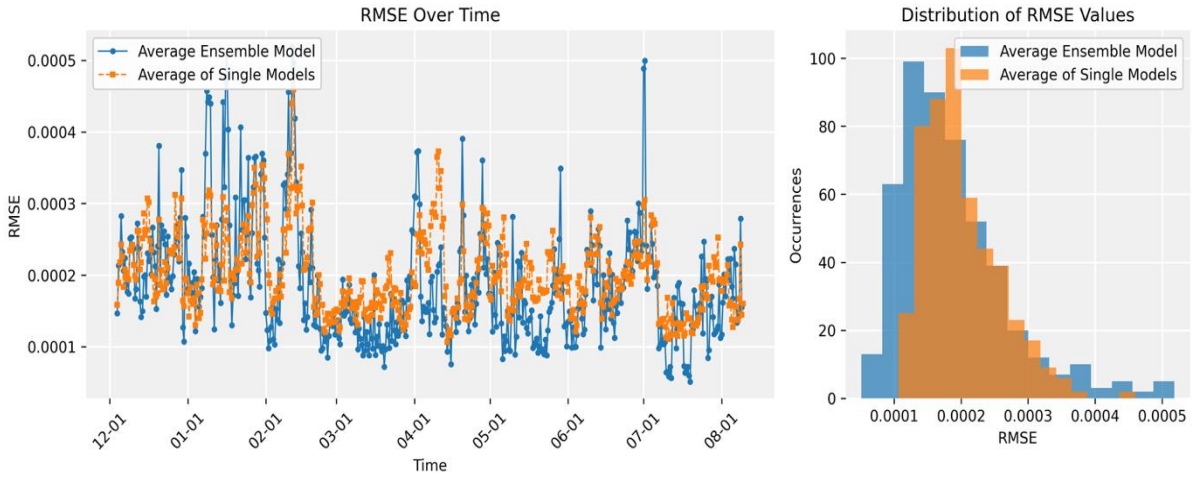


Figure 30: Performance (RMSE) of average ensemble models with time lagged model set selection in the Sustain Dairy RIL, covering the time interval Dec 2024 to Aug 2025. The histogram (right) depicts the distribution of RMSE values.

The RMSE over time reveals that both individual models and ensemble models vary in their precision over time, with average ensemble models being more performant for most time intervals. The histogram depicts the distribution of RMSE values illustrating the enhanced performance of ensemble forecasts with the time lagged model selection over the mean ensemble of individual models.

As early validation results were promising, the framework is used in the Sustain Dairy RIL for milk quality forecasts, and the outputs serve as a data product of milk quality forecasts on the regional level, as elaborated in the next chapter. An NRT application of this framework is conducted in WP5 activities.

2.5.3 Ongoing and future work

Current and upcoming developments seek to enhance insights into the areas of the framework that offer potential for further improvement. Besides others, possible optimizations encompass the balance of the number of individual models and number of ensemble models or the influence of error and fusion metrics on the ensemble forecast performance. Furthermore, also the NRT capability of the framework could be tested in the Sustain Dairy RIL, as the dairy milk quality datasets are recurrently and automatically updated by DMK. Lastly, and if requested due to other potential application scenarios, Jupyter Notebook scripts may be included in the RIE to run a correlation analysis between variables from other labs and meteorological variables (from the provided Zarr files). These scripts would serve as a decision criterion if an application of the framework may be useful to further labs.

3. Agri-environmental data products

This chapter provides an overview of the agri-environmental data products that have been developed and refined within the ScaleAgData RILs during the second iteration. By delivering sensor-integrated data products, the project continues to strengthen monitoring capacities at multiple scales. At the local level, these products provide farmers and agricultural advisors with actionable insights and tools that can support decision-making and farm management. On the regional scale, they serve as valuable information sources for public authorities and private companies, contributing to evidence-based policies and services in the agricultural sector.

During the first iteration round (M7-M18), the RILs mainly focused on developing basic models to generate agri-environmental data products. Sensor data, EO data, and other reference data were collected. Collaborations were set up with technology providers in the context of WP3, related to new sensor developments and sensor data standardization, and in WP4, to co-develop and test the various methodological frameworks (see Chapter 2). In some Labs, this already resulted in initial data products. For most RILs, however, model development was still ongoing.

Building on these foundations, the second iteration has focused on refining the models and integrating improved sensors and EO data sources. Methodological frameworks from WP4 have been tested, and implementations started in the RILs, allowing for more advanced and robust data products to be produced. Several Labs have now moved from prototype stage to generating consolidated products, which are being deployed through the RIE (see Chapter 4). The emphasis has been on ensuring that these products are operationally relevant, scalable, and transferable across different agricultural and environmental contexts.

The table below provides an updated overview of the agri-environmental data products developed in the RILs.

Table 8. Overview of agri-environmental data products developed by the RILs

RIL	Data product
Water productivity	field water status and predicted yield for target crops
	satellite based field water status and predicted yield for target crops
Crop management / Agri-environmental monitoring for policy makers	soil moisture and evapotranspiration aggregation at LAU / Commune Level
	calculated indicators (aggregates) based only on the ground truth evidence
	calculated indicators based on data assimilation mechanisms along with the respective annotations
	aggregated pesticide uses for policy makers
Crop management / Sustainability performance	Calculated sustainability indicators
	DSS model outputs
Crop management / Early pest detection	statistical data on the accuracy of observations of the occurrence of agrophages
	improved predicted agrophage occurrence data based on geolocation
	predicted overall level of agrophage occurrence risk for the selected region
Yield monitoring	Potato yield estimates (subfield level)
	Improved tare yield estimates for potatoes
	Winter wheat yield estimates (LUKE)
	Winter wheat yield estimates (VITO)
Soil health	EO based regional soil organic carbon map

	Soil health indicator estimates
Grasslands	Gap-filled grasslands LAI maps at parcel level
	Estimated grassland yield at parcel level
	Improved grassland GPP maps based on flux tower sensors
Dairy	Regional productivity of dairy farms
	Deviation of milk quality & quantity
	Assessment of grass yield at regional level
	Milk quality forecast on the regional level

Table 9 provides an overview of the EO and sensor data inputs used to generate these data products, the methods used for sensor data integration as well as how these data products are validated within the RILs.

In the sections that follow, each data product and its underlying models are described in more detail, with attention given to the required input data, methodological approach, and progress achieved so far.

Table 9: EO and sensor data inputs, integration and validation methods for each data product generated within the RILs

RILab	Data products	EO product input	Sensor input	Integration method	Data product validation
RIL1 – Water Productivity	• field water status and predicted yield for target crops • satellite based field water status and predicted yield for target crops	Sentinel-2 (NDVI, LAI, fAPAR, RGB); Sentinel-2/3 evapotranspiration; Sentinel-2/3 soil moisture; sensor-integrated soil moisture	IoT meteorological + in-soil sensor networks; YaraSense IoT; Phytech stations; airborne hyperspectral + thermal	Physical models + EO upscaling (coupling field-calibrated water-status indicators with EO-based soil-moisture + ET products)	Compared with ground measurements and soil water deficit models
RIL2a – Agri-environmental monitoring	• soil moisture and evapotranspiration aggregation at LAU / Commune Level • calculated indicators (aggregates) based only on the ground truth evidence • calculated indicators based on data assimilation mechanisms along with the respective annotations • aggregated pesticide uses for policy makers	Sentinel-2 crop classification; Sentinel-2/3 soil moisture	NP gaiasense logs, pesticide gas sensor	Aggregation pipeline, incl. AIM transformation; classification (EO + ground truth)	Accuracy assessment using field labels; Consistency checks of aggregated data
RIL2b – Sustainability performance	• Calculated sustainability indicators • DSS model outputs	Sentinel-2 (NDVI, yield maps); Sentinel-2/3 evapotranspiration; Sentinel-2/3 soil moisture; sensor-integrated soil moisture	10 weather stations	EO data assimilation into fertilization model	Comparison with 'standard' DSS outputs
RIL2c – Early Pest Detection	• statistical data on the accuracy of observations of the occurrence of agrophages • improved predicted agrophage occurrence data based on geolocation • predicted overall level of agrophage occurrence risk for the selected region	Sentinel-2 NDVI; meteorological inputs	eDWIN platform (advisor observations, meteo sensors)	Interpolation of meteo data; data fusion: ML models integrating EO + observations to estimate agrophage-occurrence probability; aggregation to higher admin level	Validated by consistency between field observations and disease models; ML model validated against advisor observations on subset of fields

RILab	Data products	EO product input	Sensor input	Integration method	Data product validation
RIL3 – Yield Monitoring	• Potato yield estimates (subfield level) • Improved tare yield estimates for potatoes • Winter wheat yield estimates (LUKE) • Winter wheat yield estimates (VITO)	Sentinel-2; Sentinel-1/2 fused (CropSAR); Sentinel-2/3 evapotranspiration; Sentinel-2/3 soil moisture; super-resolution products	RGB camera on harvester; soil moisture sampling; EC scanning; field campaigns (N, LAI, length, development stage of wheat; potato and soil weight to calculate tare)	Data assimilation (CMAAS) for winter wheat yield estimation, ML models and few-shot learning for potato and yield estimation; tare-correction model using RGB harvester imagery	Estimated yields compared with observed yield data (RMSE, correlation); estimated tare compared with tare from classified RGB images and field samples
RIL4 – Soil Health	• EO based regional soil organic carbon map • Soil health indicator estimates	Sentinel-2 bare soil indices; Hyperfield data; hyperspectral EO	UAV hyperspectral; soil sampling campaigns; spatial sampling algorithm	Federated learning for SOC estimation in Flanders + C-Macedonia without centralizing data; data fusion of UAV + satellite hyperspectral data	Validation against in-situ soil measurements
RIL5 – Grasslands	• Gap-filled grasslands LAI maps at parcel level • Estimated grassland yield at parcel level • Improved grassland GPP maps based on flux tower sensors	Sentinel-1/2 fused products; Sentinel-2 indices; Sentinel-1 backscatter; meteorological reanalysis	Flux towers; LAI/fAPAR/biomass sampling; meteo + soil moisture monitoring	DL to predict LAI from S1 or S2 + auxiliary data; LUE model for NPP estimation; GPP and grassland yield estimates refined using EO + meteorology + multi-season field data	Validation with flux tower, field LAI and biomass data
RIL6 – Sustain Dairy	• Regional productivity of dairy farms • Deviation of milk quality & quantity • Assessment of grass yield at regional level • Milk quality forecast on the regional level	Sentinel-1/2 biomass; Sentinel-2 NDVI; ERA5 Reanalysis; ECMWF IFS	Harvester-based sensors; OEM telemetry aggregation	Time-series forecasting framework meteorological + farm data	Validation against production records

3.1 RIL Water management

3.1.1 Field water status and predicted yield for target crops

This product, developed by IES and MIGAL, includes field level target crop water status and potential yield estimates from in situ meteorological and soil sensor data in combination with high resolution (<1 m/pix) airborne VNIR spectral and thermal data for more detailed distribution insight into the field level. Two target crops requiring intensive irrigation are chosen: Peppermint in Latvia and Quinoa in Israel. The proposed framework integrates in-situ soil moisture probes, local meteorological data, evapotranspiration-based water deficit indicators, irrigation treatments, and crop/yield reference measurements with EO image processing outputs, including satellite-derived soil moisture and evapotranspiration products and high-resolution airborne VNIR spectral and thermal imagery. In this integration, in-situ data provide ground reference and validation for crop water status and yield responses, while EO image processing enables the spatial upscaling of point-based observations into field-scale water status and potential yield prediction maps.

Sensor input data:

In situ sensor data was collected from local meteorological stations and soil sensor probes for two seasons of 2024 and 2025. Additional reference data were collected manually. In the case of peppermint, the summer of 2025 was very wet (precipitation exceeded climatic norm twice) and relatively cool (average temperature was below climatic norm) which affected the development of peppermint. Therefore, it wasn't possible to apply variable irrigation and to harvest the yield. This resulted in no reference data available on peppermint yield.

Table 10: Overview of sensor input data (water status & yield)

Sensor data	Source	Data provider	AOI/test sites	Nr. Fields/ season	Meas. frequency	Season(s)	Nr. training data
Precipitation (mm) Air temperature (°C) Solar radiation (W/m2) Air humidity (%) Air pressure (mm Hg) Wind speed (m/s) Wind direction ET, mm day)	Local meteorological station	IES and MIGAL	Peppermint test fields in Latvia, Quinoa test fields in Israel	1 station for 4 Peppermint fields in Latvia, 1 station for 3 Quinoa fields in Israel	Once every 30 min	2023*, 2024, 2025	1440 per month

Soil temperature (°C) Soil moisture (cB or %)	Soil sensors (~10 cm depth)	IES and MIGAL		4 sensors in Peppermint fields in Latvia, 12 sensors in Quinoa fields in Israel	Once every 30 min	2023*, 2024, 2025	5760 per month
Irrigation water supply (mm)	Irrigation monitors	IES and MIGAL		3-4 different irrigation regimes	Once per irrigation action	2023*, 2024, 2025	TBD
Yield (t/ha)	Weighting	MIGAL		4 fields	Once per season	2023*, 2024, 2025	4
Plant status (healthy / stressed)	Visual assessment	IES and MIGAL		4 fields	Once per week	2023*,2024, 2025	16 per month

*only in Israel

Test fields were in two distant regions: Peppermint fields in Latvia and Quinoa fields in Israel. In the case of peppermint, there were four test fields (~1 ha each) with different irrigation regimes located in one place. The local meteorological station was in proximity to the fields (see image below) and had a mobile data transfer module. Soil sensors were located within the fields and were wirelessly connected to the meteorological station (Figure 31). Meteorological and soil sensor data were acquired continuously (once every 30 minutes). Irrigation water supply was measured manually during the watering process using manual rain gauges. Plant status (healthy / stressed) was assessed visually once a week during on-site visits. Yield assessment was performed at the end of each season. Leaf samples were collected at the end of the season to measure qualitative yield indicators, e.g. concentration of essential oils, because peppermint didn't develop enough to be harvested.

Figure 31: Local meteorological station near peppermint fields in Latvia

Figure 32: Node with soil moisture and temperature sensors in peppermint field in Latvia

In the case of Quinoa, there were three test fields (1.44 ha each) with different irrigation regimes located within the perimeters of MIGAL's "Mataim" experimental farm. The local meteorological station was located 5 km from the test site, and its data was automatically stored in a designated, yet open to public site (http://www.mop-zafon.net/). Like Latvia, meteorological and soil sensor data were acquired continuously (once every 30 minutes). Yet the soil moisture sensors were rented for the duration of the growing season from an AgTech company, Phytec.com, which provided installment and maintenance of the sensors and storage of their data on a designated cloud. The access to the data was admitted only to specific users (MIGAL crew). Irrigation water supply was measured automatically via an irrigation computer. The amount of irrigation was concluded each morning according to the Penman–Monteith ET measure supplied by the weather station. The irrigation method was drip irrigation. As the site was on MIGAL's experimental farm, plant status (healthy / stressed) was visually assessed daily. Yield assessment was performed once at the end of the season – plant density, plant height, stem diameter, and biomass were measured.

Figure 33: Young Quinoa in Israel - test field

EO input data:

Airborne remote sensing data was gathered using drones or a small, manned aircraft.

Table 11: Overview of EO data (water status & yield)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Thermal image	IES and MIGAL	Peppermint test fields in Latvia, Quinoa test fields in Israel	Jun-Sep 2024 and 2025 for Peppermint, June-Oct 2024 and 2025 for Quinoa	Monthly	1 m
VNIR hyperspectral	IES and MIGAL			Monthly	1 m

Despite the different platforms and spatial resolutions, the airborne campaigns were expected to deliver comparable datasets suitable for assessing vegetation conditions, evapotranspiration indices, and within-field plant status. Monthly airborne acquisitions were made across two growing seasons (2024 and 2025).

For peppermint, data were collected using a small, manned aircraft equipped with a broadband thermal sensor (one spectral band) and a VNIR hyperspectral camera (24 bands, 400–1000 nm), achieving a spatial resolution of 1 m/pixel or better.

For quinoa, data were acquired using drones equipped with sensors covering the visible, near infrared, and thermal domains.

Methodology and validation results:

Data preparation is the first step to ensure standardized inputs for the models. At present, in-situ sensor data must be downloaded manually, but options for automated data access via an API are being explored; an essential requirement for implementing the Digital Twin concept. Meteorological and soil sensors operate at the same frequency, which reduces the need for preprocessing.

Soil moisture sensor data were used as a reference for evaluating alternative approaches. Correlation between the reference measurements and the alternative datasets was selected as the evaluation metric.

Sensor readings can be translated into field water status using a decision-tree approach based on standard soil moisture ranges:

- 0–10 cb: wet soil
- 10–30 cb: optimal soil moisture
- 30–60 cb: moderately dry; irrigation may be applied but is not yet required
- 60–100 cb: dry; irrigation is required to avoid drought stress
- >100 cb: very dry; crop likely under severe drought stress

A translation of soil moisture readings from peppermint fields in Latvia into water-status classes is shown in the figure below, with a corresponding map visualization on the right.

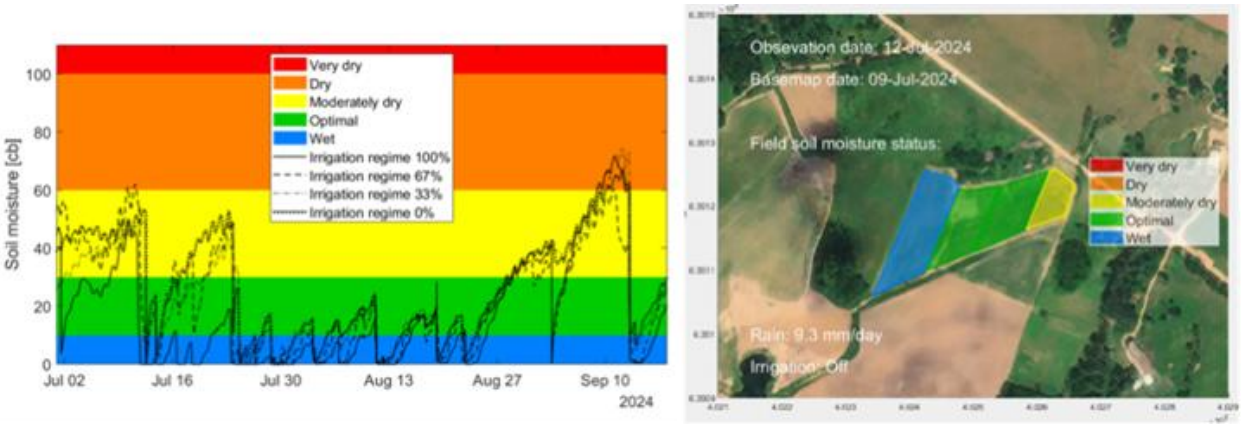


Figure 34: A translation of soil moisture readings from peppermint fields in Latvia into water-status classes (left), with a corresponding map visualization (right).

The soil water deficit method was selected as an alternative approach for assessing field water status. A simplified soil water balance equation was used: Dc = Dp + ET – P – Irr, where Dc is the current day's soil water deficit, Dp is the previous day's deficit, ET is evapotranspiration, P is precipitation, and Irr is the applied irrigation. This approach does not require soil sensors in every field and can rely instead on local meteorological data and recorded irrigation amounts.

A strong correlation (R^2 = 0.78) was found between soil moisture measurements (cb) and soil water deficit estimates (mm) derived using the evapotranspiration-based method (figure below).

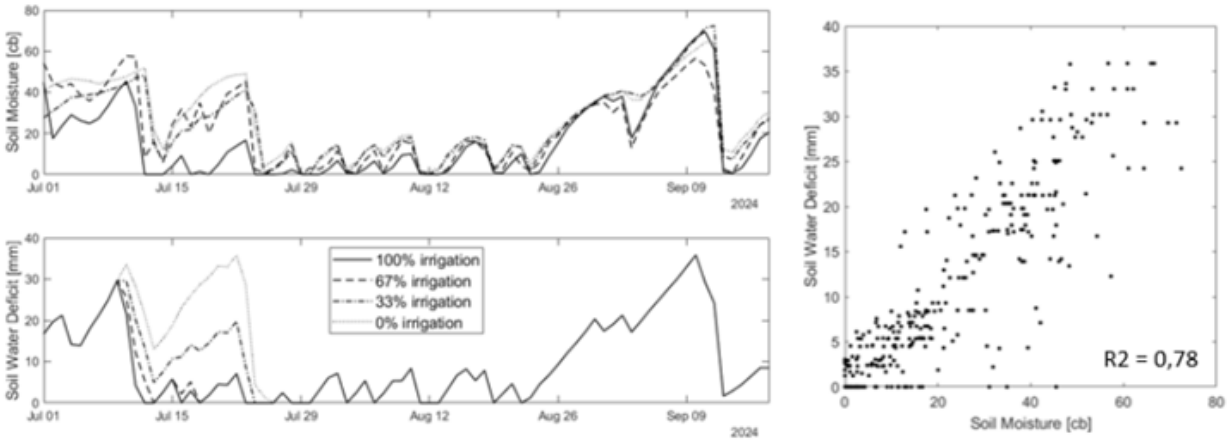


Figure 35: Correlation between soil moisture measurements and soil water deficit estimates using the evapotranspiration-based method

For airborne data, field boundaries were used to define the spatial extent of the raster datasets. VNIR spectral data was resampled to match the spatial resolution of the thermal imagery. Soil sensor locations serve as sampling points to extract airborne sensor values, supporting the upscaling of in-situ measurements across the field. Airborne data therefore enables refinement of model outputs from single point estimates to a spatial resolution of 1 m/pixel within the field.

An empirical yield-prediction model was planned to be built, using yield reference data from different irrigation regimes together with sensor inputs (meteorological, soil, and airborne). However, at least two years of data are required to develop and test such a model due to the limited sample size. Yield

data from the quinoa pilot site in Israel under different irrigation regimes are shown in the figure below.

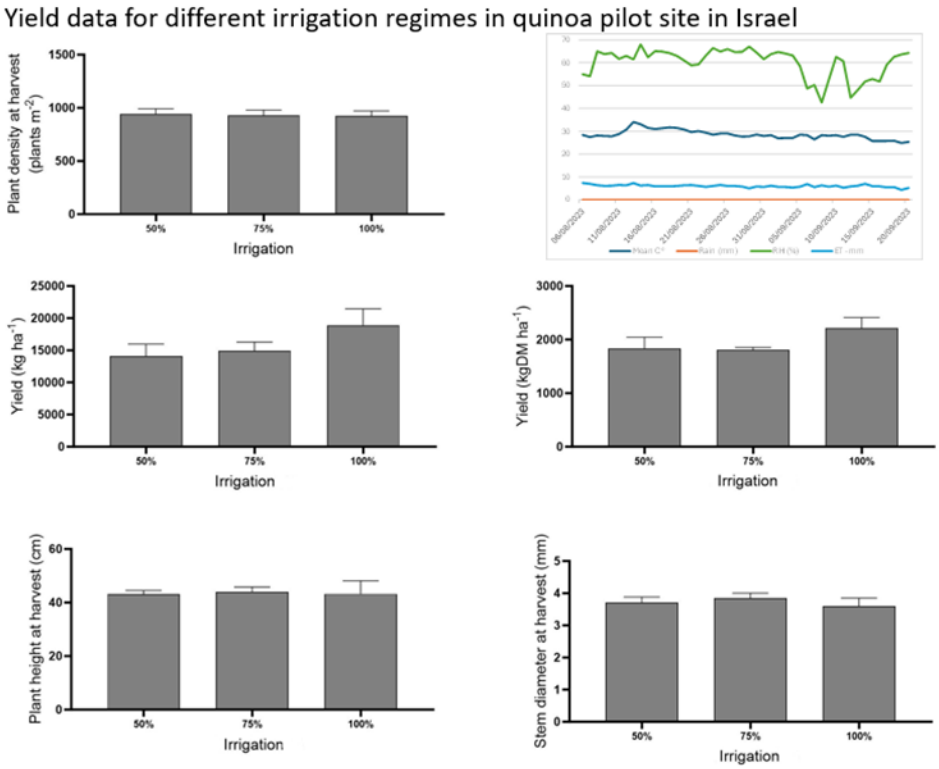


Figure 36: Yield data from the quinoa pilot site in Israel under different irrigation regimes

Sensor-integrated data products:

Two sensor-integrated data products were pursued: field water status and predicted yield. Field water status was successfully demonstrated using two alternative approaches:

- a soil-moisture–sensor-based model, when in-situ soil moisture data available
- a soil water deficit model, when only local meteorological data available

Both methods were demonstrated with daily update frequency. Further testing will explore increasing spatial resolution to 1 m/pixel using the acquired airborne data.

The collected yield reference data is used to develop the yield prediction model, incorporating meteorological data, IoT sensor measurements, and airborne observations as inputs.

Table 12: Overview of sensor-integrated data products (water status & yield)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Field water status (normal / stressed)	Peppermint in Latvia, Quinoa in Israel	Jun – Aug for Peppermint in Latvia, June – Oct for Quinoa in Israel	Daily	1 m/pix	TBD
Predicted yield (t/ha)			Daily	1 m/pix	TBD

Use case(s):

The pilot sites were implemented in close collaboration with local farmers, who evaluated the results and provided valuable feedback. In Latvia, the peppermint test fields were established together with SIA Field and Forest, a regional leader in the production of medicinal and aromatic plants. In Israel, the chief cereal guide of the Ministry of Agriculture played a key role in designing the quinoa experiment and considers it an important step toward advancing quinoa cultivation in the country. Several quinoa growers, mainly in northern Israel, were informed about the trial, and knowledge exchange is ongoing with some of them. The Galilee Agriculture Company is a major stakeholder, actively supporting the experiment and contributing its own resources.

3.1.2 Satellite based field water status and predicted yield for target crop

This product, that is being developed by IES and MIGAL, monitors crop water status and potential yield using in situ meteorological and soil sensor data combined with high-resolution airborne VNIR spectral and thermal imagery. Focused on Peppermint in Latvia and Quinoa in Israel, the models use similar principles and data but are tailored for each crop. The proposed framework integrates in-situ soil moisture probes, local meteorological data, evapotranspiration-based water deficit indicators, irrigation treatments, and crop/yield reference measurements with EO image processing outputs, including satellite-derived soil moisture and evapotranspiration products and high-resolution airborne VNIR spectral and thermal imagery. In this integration, in-situ data provide ground reference and validation for crop water status and yield responses, while EO image processing enables the spatial upscaling of point-based observations into field-scale water status and potential yield prediction maps.

In situ data were collected using IoT sensors (e.g., soil moisture) as well as during onsite visits (e.g., soil composition, crop yield) and were subsequently used as reference data for model training and validation. Field water status can be directly assessed from soil moisture sensor data; however, deploying IoT sensors across all fields would be prohibitively expensive. Therefore, indirect field water status assessment approaches were also tested, utilizing evapotranspiration (ET) derived from local meteorological stations, as well as ET and soil moisture obtained from EO data. In the case of yield prediction, only a limited number of reference samples were available — insufficient for empirical model training but nonetheless providing insights into the impact of different irrigation systems on the potential yield of quinoa.

EO input data:

EO input data is provided by partners DHI and VITO:

Table 13: Overview of EO input data (water status & yield)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Evapotranspiration	DHI	Peppermint test fields in Latvia, Quinoa test fields in Israel	Jun-Sep 2024-2025 for Peppermint, June-Oct 2024-2025 for Quinoa	Daily	20 m/pix
Soil moisture	DHI			Daily	20 m/pix
NDVI	VITO			Weekly	20 m/pix
LAI	VITO			Weekly	20 m/pix
PPI	VITO			Weekly	20 m/pix
fAPAR	VITO			Weekly	20 m/pix

Methodology and validation results:

Potential upscaling of field water status was tested using soil moisture and evapotranspiration EO data products provided by DHI group.

In the case of soil moisture data products, reference data was converted from tensiometer readings to soil moisture percentage. Comparison of IoT soil moisture probe and EO data product readings is shown in figure below, where R2 was < 0,5. Nevertheless, soil moisture drops showed similar patterns, thus indicating potential application for irrigation triggering.

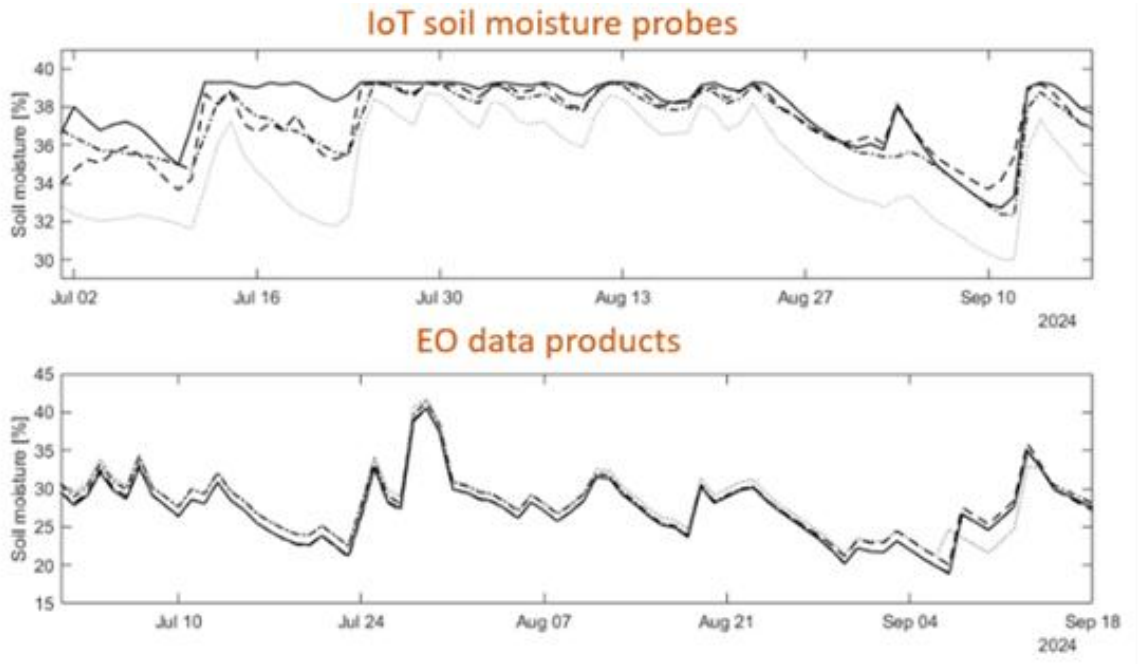


Figure 37: Comparison of IoT soil moisture probe and EO data product readings

For the evapotranspiration data products, a reasonably strong correlation (**R^2 = 0.62**) was found between local meteorological measurements and the DHI model outputs (see figure below). Based on this agreement, the EO-based evapotranspiration products were used to calculate soil water deficit and to generate corresponding soil water status maps.

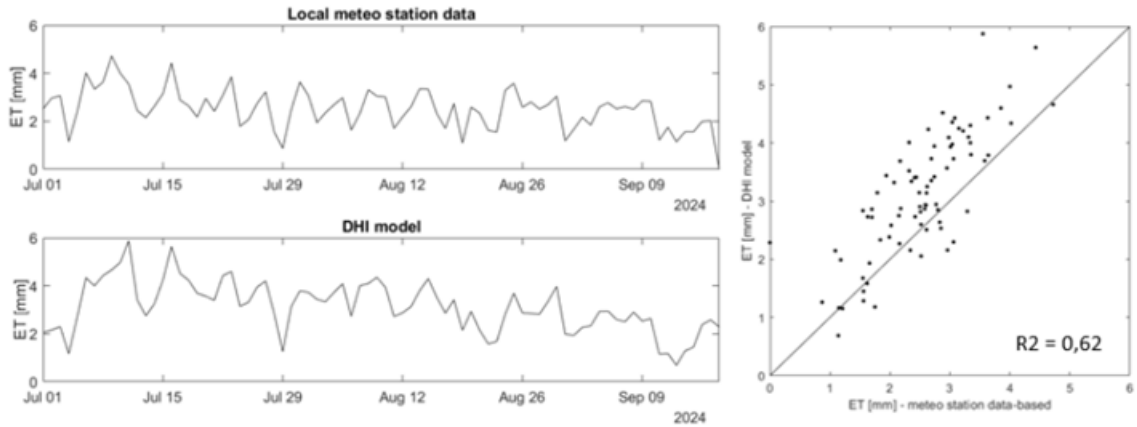


Figure 38: Correlation between local meteorological and DHI model

Comparison of the visualizations from the IoT soil moisture probe-based approach and the evapotranspiration EO data-based approach are shown in figure below.

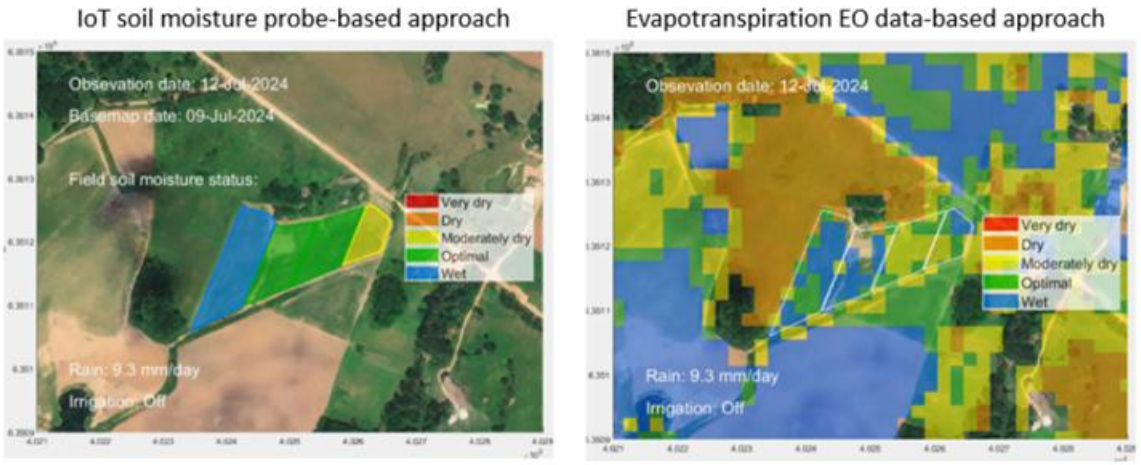


Figure 39: Comparison of visualizations from the IoT soil moisture probe-based approach and the evapotranspiration EO data-based approach

Yield prediction will be tested using Sentinel spectral data, including vegetation indices such as NDVI, LAI, PPI, and FAPAR. Various machine-learning approaches will be evaluated for developing empirical models. Feature-importance analysis will be used to identify the most relevant inputs and exclude less informative EO products, thereby simplifying and optimizing the model.

Sensor-integrated data products:

Two sensor-integrated data products were expected as an output – field water status as a vegetation health status assessment and predicted yield as a forecast. Potential upscaling of field water status using SM and ET data products was demonstrated with data from 2024 and will be repeated with data from 2025. Yield prediction model testing will be started in 2026 when all necessary inputs will be gathered.

Table 14: Overview of sensor-integrated data products (water status & yield)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Field water status (normal / stressed)	Peppermint in Latvia, Quinoa in Israel	Jun – Aug for Peppermint in Latvia, June – Oct for Quinoa in Israel	Weekly	20 m/pix	TBD
Predicted yield (t/ha)			Weekly	20 m/pix	TBD

Use case(s):

The pilot sites were implemented in collaboration with local farmers who evaluated the result and provided feedback. In Latvia, Peppermint test fields were implemented in collaboration with the farming company SIA Field and Forest, regional leader in production of medical and aromatical plants. In Israel, the Galilee Agriculture Company was collaborating in the implementation of the experiment.

3.2 RIL Crop management

3.2.1 Calculated indicators (aggregates) based only on the ground truth evidence

This product is produced by Neuropublic (NP) and provides an aggregated estimation of the applied agrochemical and water quantities per hectare based on farm-log data that are collected from the pilot parcels of potato, tomato and cotton in Thessaly, Greece. During the two iterations of the project, the aggregation was implemented on two different scales, as presented in the following figure.

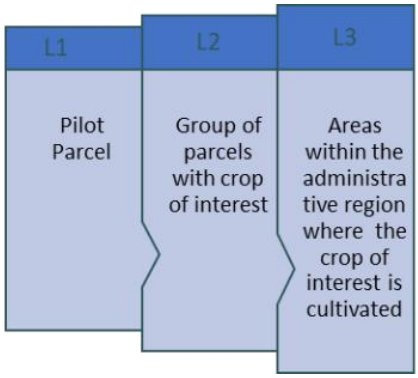


Figure 40: Aggregation levels for the two iterations of the project

The pilot parcel's measurements are aggregated to the level of group of parcels. This product integrates in-situ ground-truth evidence (farm-management records and sensor-based parcel data harmonised through AIM, including parcel-level practice events such as sprays, fertilisations, irrigations and harvests) with EO-derived crop-classification masks generated from Sentinel-2 imagery (Bands B2–B6, processed through an SVM-RBF classifier trained on 2023 reference data). The Sentinel-2 crop masks are used to identify the spatial extent of each crop type at LAU/NUTS level, against which the parcel-level ground-truth indicators are aggregated to produce regional-scale practice indicators per crop and per period.

Methodological framework

Our methodological approach (Figure 41) uses as inputs in-situ farm data from the pilot parcels to generate related regional agricultural statistics. The workflow also utilizes additional groups of parcels for creating an intermediate aggregate level that will be used for evaluation of the quality of pilot farm data to the regional level. The collection of farm-level input data within a region of interest (ROI), utilizes a Local Administrative Unit (LAU) area (https://ec.europa.eu/eurostat/web/gisco/geodata/statistical-units/local-administrative-units) or Commune area (https://ec.europa.eu/eurostat/web/gisco/geodata/administrative-units/communes) including the farm calendar data and parcel polygons that fall within the region of interest.

Both pilot parcel data and group parcel data are used to calculate specific indicators per hectare, such as the amount of pesticides used or the yield produced. Additional, regional-level input data for the ROI, including crop type classification binary mask and EO data, or other data from external data sources (e.g. Eurostat) are gathered. The calculated farm-level indicators are then aggregated to the regional level based on the regional crop type area, resulting in comprehensive regional indicators such as pesticides used per hectare, irrigation used per hectare, yield per hectare, and financial inputs per hectare. Regional land use data are also used as additional source of information to provide context to the end users regarding the environmental importance and rural overview of the region of interest.

Work during the second iteration has focused on improving the robustness and usability of the EO-based processing pipeline. Advanced EO-assisted scaling and confidence flagging were introduced for Level-3 aggregates to better manage parcel-level variability and inter-annual data gaps. Crop-type classification was further developed using a Support Vector Machine (SVM) classifier with a Radial Basis Function (RBF) model trained on 2023 Sentinel-2 bands (B2–B6) and applied to the 2024 season. The resulting layers have been packaged for dashboard integration, enabling new Region+Crop filtering options and per-year/per-region visualisation. In parallel, backend services were stabilised, including AIM (Artificial Intelligence and Machine Learning) semantic translation components for farm-log harmonisation and a Geolocation API that automatically assigns LAU/NUTS identifiers to all generated outputs.
To support location-based filtering, NP developed a Geolocation API module that performs geospatial queries to associate any geographic point with its corresponding Greek administrative unit and NUTS statistical region. The system receives input in the form of WGS84 coordinates (longitude and latitude), and uses a pre-processed GeoDataFrame of Greek commune polygons (from Eurostat's GISCO dataset) to determine the exact spatial unit the point resides in.

By using the GeoPandas library, the module performs a spatial query to identify the commune polygon that contains the given point. Once a match is found, the API returns not only the name and ID of the respective commune but also the corresponding NUTS 1, NUTS 2, and NUTS 3 region codes and names, including their Latinized versions. This "geographic tagging" enables the dynamic aggregation and filtering of data by spatial regions, improving the flexibility and granularity of statistical analysis within the platform.

The geospatial functionality is fully integrated into a standalone REST API built with FastAPI, allowing batch queries, polygon-based matching (via centroid lookup), and optional visualization using Folium maps.

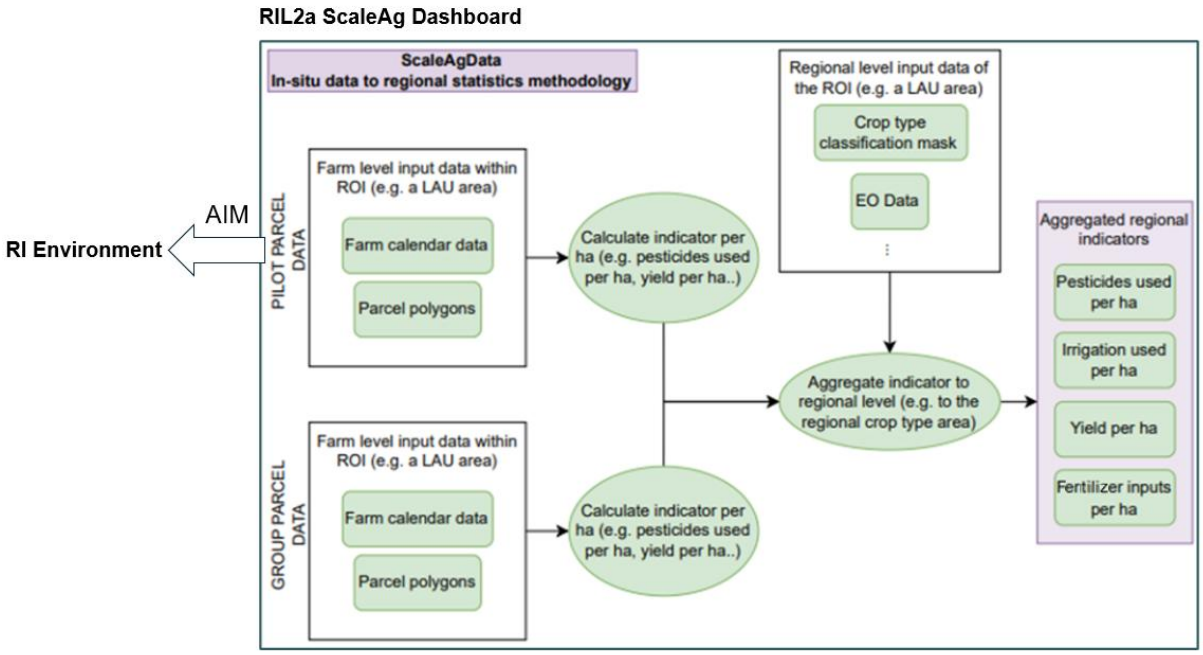


Figure 41: RIL crop management Methodological framework

Pilot Parcels

In the **first iteration**, the pilot parcel measurements (L1) were aggregated to the level of group of parcels (L2).

During the **second iteration** we progressed from parcel-level (L1) and group-level (L2) analytics to regional-level (L3) aggregation where the crop is cultivated. At pilot-parcel level, the system is now able to collect a structured response for each parcel (including parcel ID, startDate and endDate) and to summarize key indicators:

- Sprays
- Fertilizations
- Irrigations
- Harvests

Consequent derivatives of those parcel level data include pesticides applied/ha, irrigation used m3/ha, yield kg/ha, and financial inputs euro/ha. To mitigate sparse and scattered ground data when aggregating to LAU/region, we leaned more on EO-derived indices with confidence flags attached to L3 aggregates.

The original 8 pilot parcels in Thessaly (tomato, cotton) and Crete (potato) were kept but in the next seasons additional parcels were included. Parcel rotation in 2024 required updated deployments but preserved comparable agronomic profiles. The initial three crop types (potato, tomato and cotton) during the first iteration are now expanded with peach and wheat parcels.
The total pilot's area per crop in hectares (for years 2022-2024) are:

- Tomato: 24.46 ha
- Cotton: 21.62 ha
- Wheat: 13.92 ha
- Peach: 7.07 ha
- Potato: 1.09 ha

A summary table with all distinctive parcels' IDs used for the second iteration activities can be viewed below (Table 15):

Table 15. RIL Crop Management - Parcels used in 2nd iteration

NUTS1	NUTS2	NUTS3	Commune region	parcel ID	crop	year
Voreia Elláda	Dytiki Makedonia	Grevena, Kozani	Dimos Velventoy	270393	Peach	2022
Voreia Elláda	Dytiki Makedonia	Grevena, Kozani	Dimos Velventoy	270393	Peach	2023
Voreia Elláda	Dytiki Makedonia	Grevena, Kozani	Dimos Velventoy	269968	Peach	2023
Voreia Elláda	Dytiki Makedonia	Grevena, Kozani	Dimos Velventoy	270393	Peach	2024
Voreia Elláda	Dytiki Makedonia	Grevena, Kozani	Dimos Velventoy	269968	Peach	2024
Voreia Elláda	Kentriki Makedonia	Imathia	Dimos Alexandreias	270195	Peach	2023
Voreia Elláda	Kentriki Makedonia	Imathia	Dimos Alexandreias	345638	Peach	2023
Voreia Elláda	Kentriki Makedonia	Imathia	Dimos Alexandreias	270194	Peach	2023
Voreia Elláda	Kentriki Makedonia	Imathia	Dimos Alexandreias	270195	Peach	2024
Voreia Elláda	Kentriki Makedonia	Imathia	Dimos Alexandreias	345638	Peach	2024
Kentriki Elláda	Thessalia	Larisa	Dimos Kileler	254816	Cotton	2022
Kentriki Elláda	Thessalia	Larisa	Dimos Kileler	254627	Cotton	2022
Kentriki Elláda	Thessalia	Larisa	Dimos Kileler	254626	Cotton	2022
Kentriki Elláda	Thessalia	Larisa	Dimos Kileler	254627	Cotton	2023
Kentriki Elláda	Thessalia	Larisa	Dimos Kileler	254626	Cotton	2023
Kentriki Elláda	Thessalia	Larisa	Dimos Farsalon	254944	Tomato	2022
Kentriki Elláda	Thessalia	Larisa	Dimos Farsalon	271384	Tomato	2023
Nisia Aigaiou, Kriti	Kriti	Lasithi	Dimos Oropedioy Lasithioy	119700	Potato	2022
Nisia Aigaiou, Kriti	Kriti	Lasithi	Dimos Oropedioy Lasithioy	167914	Potato	2022
Nisia Aigaiou, Kriti	Kriti	Lasithi	Dimos Oropedioy Lasithioy	273368	Potato	2023
Nisia Aigaiou, Kriti	Kriti	Lasithi	Dimos Oropedioy Lasithioy	274471	Potato	2023
Kentriki Elláda	Thessalia	Karditsa, Trikala	Dimos Sofadon	254952	Tomato	2022
Kentriki Elláda	Thessalia	Karditsa, Trikala	Dimos Sofadon	270773	Tomato	2023
Kentriki Elláda	Sterea Elláda	Fthiotida	Dimos Domokoy	271208	Tomato	2023
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	327660	Wheat	2023
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	299955	Wheat	2023
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	299860	Wheat	2023
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	299846	Wheat	2023
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	327660	Wheat	2024
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	299955	Wheat	2024
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	299860	Wheat	2024
Voreia Elláda	Kentriki Makedonia	Kilkis	Dimos Kilkis	299846	Wheat	2024
3	**5**	**7**	**8**	**21**	**5**	**3**

In total, 21 pilot parcels are included, distributed across 3 NUTS 1, 5 NUTS 2 and 7 NUTS 3 regions, and 8 commune regions. The parcels cover the 5 different crops and span 3 reference years 2022-2024. Below are general and detailed visualizations of the parcel data availability per region, year, and crop type that were employed in the second iteration.

Table 16. General (above) and detailed (below) data availability for parcels used in the 2nd iteration

crop	Cotton			Peach			Tomato			Potato			Wheat		
year	2022	2023	2024	2022	2023	2024	2022	2023	2024	2022	2023	2024	2022	2023	2024
Commune name															
Alexandrias					✓	✓									
Domokou								✓	✓						
Kileler	✓	✓	✓				✓	✓	✓						
Kilkis														✓	✓
Oropediou Lasithiou										✓	✓	✓			
Servion - Velventou				✓	✓	✓									
Sofadon							✓	✓	✓						
Farsalon	✓	✓	✓				✓	✓	✓						

crop	Cotton						Peach						Tomato						Potato						Wheat					
year	2022		2023		2024		2022		2023		2024		2022		2023		2024		2022		2023		2024		2022		2023		2024	
level	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2
Commune name																														
Alexandrias	0	0	0	0	0	0	0	0	3	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Domokou	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	0	17	0	0	0	0	0	0	0	0	0	0	0	0
Kileler	3	0	2	58	0	1	0	0	0	0	0	0	0	1	0	8	0	71	0	0	0	0	0	0	0	0	0	0	0	0
Kilkis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	4	0
Oropediou Lasithiou	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	1	0	4	0	0	0	0	0	0
Servion - Velventou	0	0	0	0	0	0	1	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sofadon	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
Farsalon	0	2	0	3	0	5	0	0	0	0	0	0	1	0	1	14	0	41	0	0	0	0	0	0	0	0	0	0	0	0

The numbers in the table correspond to the available parcels in the region depending on if they belong to pilot parcels (L1) or group of parcels within the region (L2).

In the maps below, the commune boundaries are shown as light–yellow polygons and the pilot parcels as coloured points (Figure 42). The first map provides a zoomed-in view of the commune in the Lasithi plateau (Crete) with the potato cultivation locations (brown).

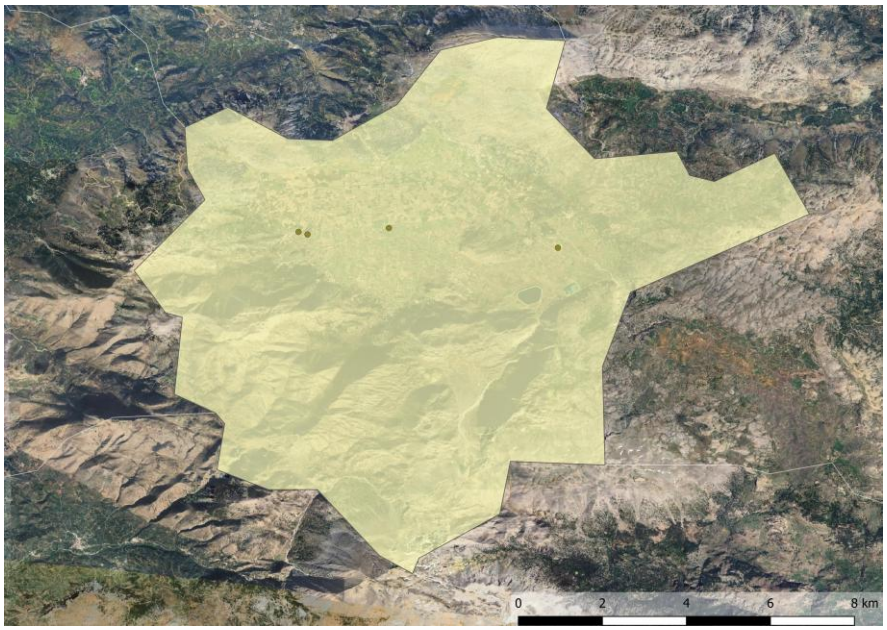


Figure 42: Visualization of the commune boundaries (yellow) in the Lasithi plateau and the pilot parcels (green points)

The second map shows the three selected communes in northern Greece, illustrating peach (orange) and wheat (purple) locations within the wider agricultural landscape and the surrounding coastal area.

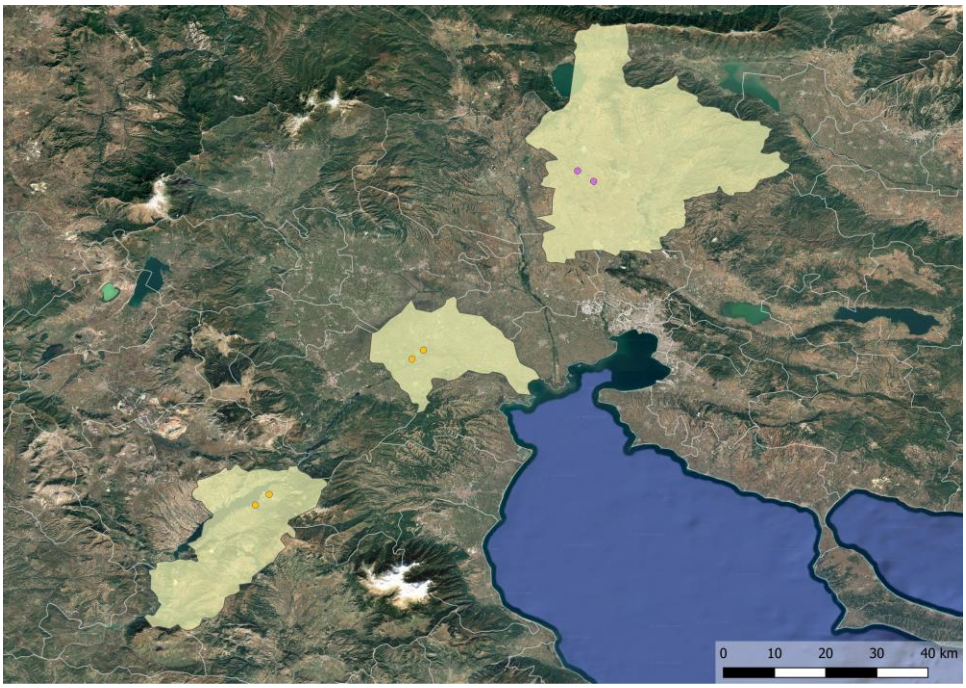


Figure 43: Visualization of the three selected communes in northern Greece, illustrating peach (orange) and wheat (purple) locations

The third map showcases the main agricultural area of central Greece and the broader region of Thessaly with the cultivations locations for tomato (red) and cotton (green).

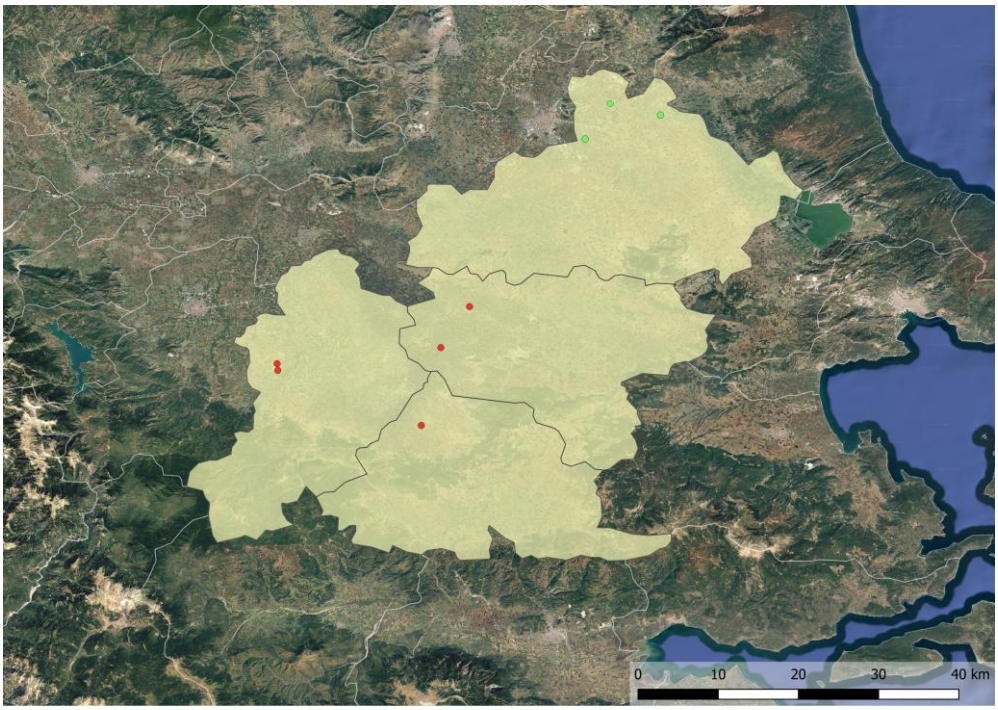


Figure 44: Visualization of the main agricultural area of central Greece and the broader region of Thessaly with the cultivations locations for tomato (red) and cotton (green)

Regarding the 2025 growing season pilot integration, we currently have 7 parcels integrated into the pipeline, covering three administrative regions and three crops (potato, wheat, and tomato). As

additional farm records are ingested and validated, the number of parcels and associated events is expected to increase, expanding the spatial and crop coverage of the pilot dataset for 2025.

Crop Classification Product

To support the L3 aggregation level, a regional crop classification product is required. To perform the crop classification, we used a SVM classifier with a RBF kernel, trained on parcel-level reference data.

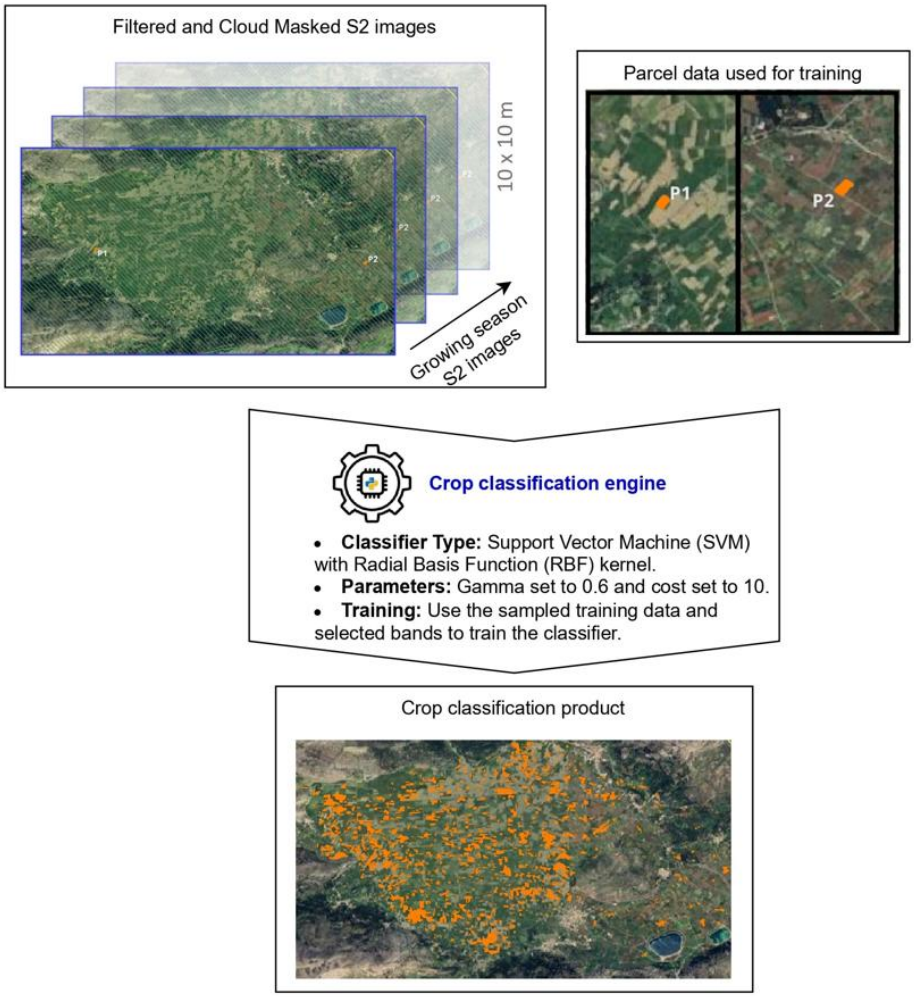


Figure 45: RIL Crop Management - Crop classification pipeline scheme

As input, we used:

i. parcel data (polygon geometries, declared crop type, start and end of the growing season)
ii. pre-processed Sentinel-2 imagery (cloud-masked images over the growing season and selected bands B2–B6)
iii. Auxiliary layers such as Local Administrative Unit (LAU) boundaries and land-use masks used to exclude non-agricultural areas (water, bare land, forest, built-up).

The training samples are automatically generated by intersecting the parcel geometries with the time series of Sentinel-2 pixels, extracting spectral values and labelling each sample with the corresponding crop type.

Once trained, the microservice applies the SVM model to all eligible 10 × 10 m pixels within the selected LAU(s), generating a wall-to-wall crop classification product. The output is returned as a geospatial raster (or vectorized polygons), together with metadata on the input data, parameter

settings (e.g. gamma = 0.6, cost = 10) and processing date, enabling traceability and reproducibility within the wider system architecture.

An example result for potato crop classification product is illustrated below (blue areas).

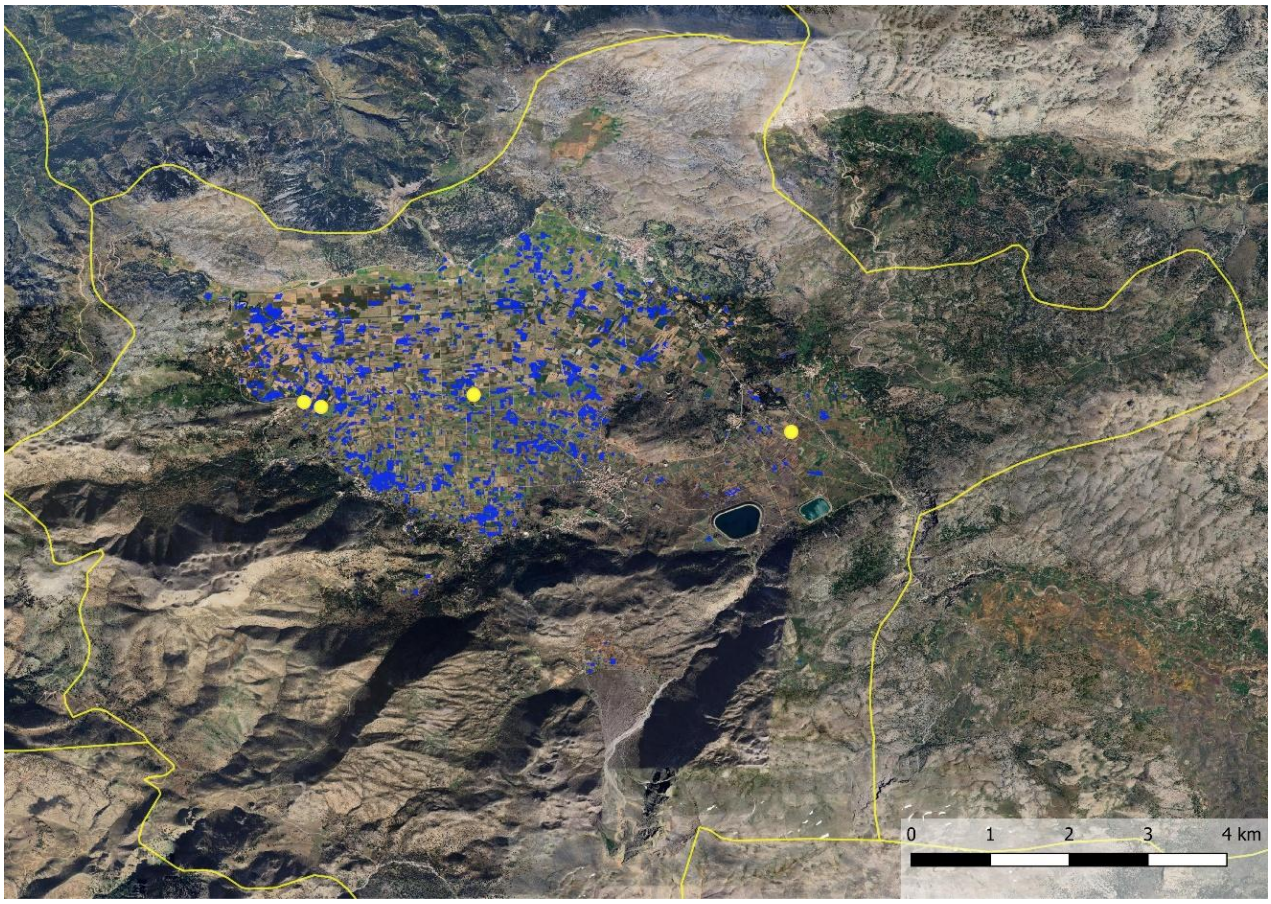


Figure 46: Example of potato crops classification (blue polygons) in Thessaly, Greece using Sentinel-2 imagery

Sensor input data:

To develop this product, hourly data measurements were collected from multiple sensors across the various pilot parcels and seasons.

- Meteorological parameters including air temperature, wind speed, solar radiation, rainfall, humidity
- Crop/vegetation properties (e.g. type, height, root-depth)
- Irrigation measurements derived from available farm calendars

In detail, the sensor data are described in Table 17.

Table 17: Sensor data description (aggregated estimation of the applied agrochemical and water quantities per hectare)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
• Leaf Relative Humidity (%) • Relative Humidity (%) • Soil Moisture 10 cm – 90 cm (%) • UV • Atmospheric Pressure (mbar) • Soil Salinity 10cm – 90 cm • Average Wind Speed (km/h) • Wind Direction • Rainfall (mm) • Leaf Temperature (°C) • External Temperature (°C) • Soil Temperature 10 cm – 90 cm (°C) • Leaf Wetness (h) • Solar Radiation (wh/m²)	IoT stations	NP	Thessaly and Crete, Greece	21 (more to be inserted for 2025)	hourly	2022-2024
Pesticide Spraying Events	IoT stations	NP	2 parcels with tomato in Thessaly, Greece	1	hourly	2024, 2025

During the second iteration, significant progress was made on the pesticide sensor (nanoparticle gas-sensing array) workflow. Following the controlled spray tests carried out in 2024, the sensor station was repaired and redeployed in 2025. Data collection is ongoing, and the current focus is on validating the new measurements before integrating them into the full processing pipeline. Once validation is complete, the sensor data will be made accessible through a REST-based interface to ensure interoperability with the RIE and other project components. Figure 47 illustrates the API developed to request pesticide sensor data in a standardized and interoperable manner.

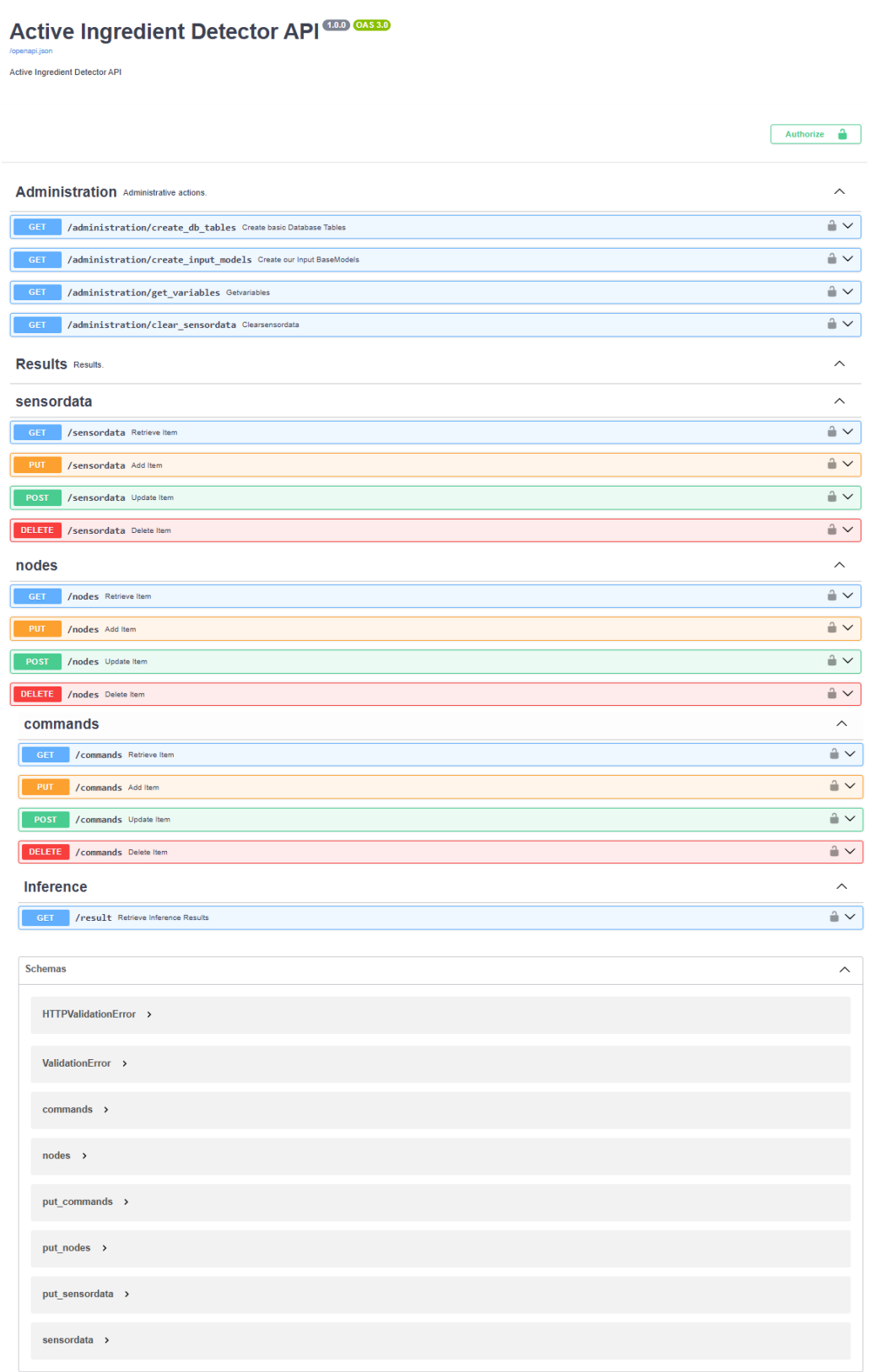


Figure 47: Illustration of the API developed to request pesticide sensor data in a standardized and interoperable manner.

The main endpoint exposed is a GET /sensordata request. Through this call, a user—either a machine client or a human operator— can provide standard query parameters such as:

- node (sensor node ID),
- datetimestamp_from_date / datetimestamp_to_date (time window for the requested measurements),
- optional limit and offset (for pagination), and
- optionally a specific id (to retrieve an individual record).

In response, the API returns a JSON object containing time-stamped pesticide measurements from the selected node(s), together with the supporting variables.

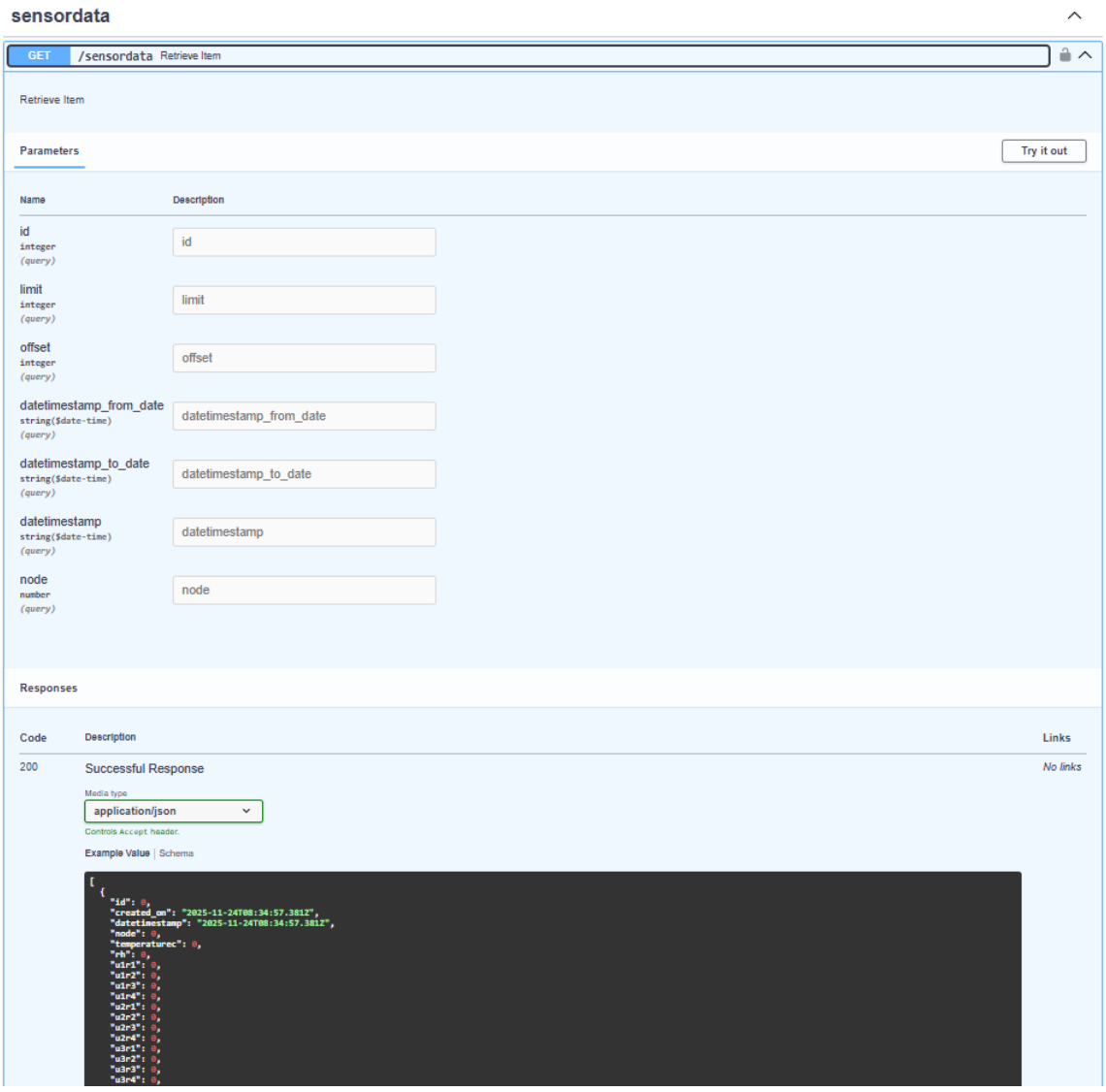


Figure 48: Illustration of API – Get Sensor data

Data from the 2025 deployment are currently undergoing quality control and validation (sensor drift checks, outlier screening, and comparison with reference measurements) prior to full integration into the project's data processing pipeline. Once this validation is completed end-to-end, the new data will be available through a RESTful service, ensuring scalable and interoperable access to pesticide exposure data.

Table 18: Overview of the EO input data (aggregated estimation of the applied agrochemical and water quantities per hectare)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel 2 - Level-1C, Level 2A	Google Earth Engine (GEE)	Four communes in Thessaly, (Tiles 34SEJ, 34SFJ)	2022-2024	~5 days	10 m
Binary crop classification based on Sentinel 2 Level-2A products	Google Earth Engine (GEE)	21 pilot parcel Communes	Early 2023 – Early 2024	~5 days	10 m

During the second iteration, work progressed on operationalising additional spectral indices, more specifically the Normalized Difference Water Index and Pigment Sensitive Ratio Index, extending the workflow beyond the Normalized Difference Vegetation Index. These indices are now generated using Evalscript-based masking to handle zero-denominator cases, exclude water pixels via Sentinel-2 scene classification layer classes, and apply dataMask filtering to retain only valid observations. The outputs include both the computed index values and an accompanying binary dataMask, which is used for quality assurance and scaling procedures.

Sensor-integrated data products:

The different data products that will be generated are shown in the table below. These products are available for 2022-2024 datasets and are currently being updated for the 2025 growing season.

Table 19: Overview of the generated data products (aggregated estimation of the applied agrochemical and water quantities per hectare)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Estimation per hectare of • yield • pesticide use • water quantities • fertilizer type/quantities	21 pilot parcel Communes	Early 2023-Early 2024	Once, at the end of a growing season	Commune level

Use case(s):

The products are assessed mainly by farmers, agricultural cooperatives and policy makers and their main goal is to promote the adoption of more efficient and sustainable agricultural practices. By providing detailed, evidence-based insights into agricultural inputs and outputs at both the farm and regional levels, the methodology aims to inform decision-making, optimize resource use, monitor and enhance overall agricultural productivity and sustainability at multiple levels.

In the second iteration, the introduction of Region + Crop selection in the dashboard interface has reduced interaction effort for end-users and improved the accessibility of crop-specific insights. The updated views now draw on per-year, and per-region layers derived from the packaged crop-type product, supported by ongoing 2025 data ingestion and quality control. These enhancements strengthen the dashboard's value for policy makers and other stakeholders. A user walkthrough will be developed once the service configuration is finalised.

3.2.2 Soil moisture & evapotranspiration aggregation at LAU / Commune Level

These products, to be developed by NP in collaboration with DHI, include evapotranspiration and soil moisture that are derived from in-situ sensors combined with Sentinel-2/3 data in selected pilot parcels. The products are designed to use as inputs localized soil moisture information (probe that samples moisture at multiple depth profile) and weather parameters from IoT sensors and calculate aggregation statistics at the LAU/Commune level, crucial for precision agriculture and land management.

During the second iteration, NP, DHI and ICCS considered an integrated use of the DHI product as a planning input for in-situ sensor network design. NP provided the pilot parcel (peach orchards, Central Macedonia), DHI provided the pixel-level soil moisture map, and ICCS applied its spatial-planning algorithm to derive a candidate sensor placement layout using the DHI-derived spatial SM variability map to identify locations capturing the most meaningful range of conditions across the parcel. This constitutes a first end-to-end exploration within the project of EO-derived soil moisture information being used to guide sensor placement decisions. Hardware was sourced in December 2025 and installations are underway in Spring 2026.

Sensor input data:

The sensor data are similar to section 3.2.1

EO input data:

The EO data are similar to section 3.2.1

Methodology and validation results:

Once the sensor network is operational, the project will explore the degree to which the spatial behaviour of the DHI product is consistent with ground observations over the pilot parcel. Given that the DHI product is based on multi-year historical modelling and the in-situ network will provide observations from a single growing season over a single crop type in a specific agro-climatic context, this exploration is not intended as a generalizable quantitative validation. Rather, it represents an opportunity to examine product behaviour in a Mediterranean agricultural setting, and to identify conditions under which EO-based SM estimates appear coherent with or divergent from local ground observations. The potential to retrieve current season SM products from DHI for further product validation will be also explored.

In parallel, the behaviour of the DHI evapotranspiration product over the same pilot area will be explored. The DHI ET product provides actual evapotranspiration at dekadal frequency derived from the synergistic use of Sentinel-2 and Sentinel-3 observations. NP's IoT meteorological station network in the pilot area provides locally observed atmospheric variables that may serve as a partial consistency reference, offering an additional perspective on whether the DHI product's temporal dynamics are coherent with local conditions. The scope and form of this exploration will depend on the data available at the time and will be reported in the final period.

Use case(s):

Subject to the insights gained from the spatial and temporal consistency exploration, the project will investigate whether and to what extent the DHI SM and ET products could contribute as inputs to a regional aggregation at LAU / Commune level, complementing the existing RILab2a regional indicator pipeline. Several factors will shape this investigation, including the degree of spatial and temporal consistency observed over the pilot, the diversity of crop types and soil conditions present in the

monitored communes, and the resolution and frequency of the DHI products relative to the aggregation needs. Any findings on regional aggregation will be scoped to what the available evidence supports and will be reported in the final period.

3.2.3 Calculated indicators (LUKE) based on data assimilation mechanisms along with the respective annotations

This activity, initially planned for development by NP in collaboration with LUKE, aimed to generate products related to within-field nutrient and water variation, crop health status and yield estimation. NP was originally expected to contribute to this activity by providing in-situ data from four winter wheat fields in Northern Greece, including time series from meteorological stations installed on the parcels, soil analyses and farm calendars. These datasets would have supported calibration and validation of the Digital Twin model under Mediterranean conditions, before its distribution. The work was expected to follow similar data inputs and methodological principles as outlined in section 3.2.1.

However, during the second iteration, this line of work did not progress beyond the conceptual stage. Due to internal NP legal and data-sharing procedures, it was not feasible within the reporting period to release the Internet of Things (IoT) sensor data needed for the agreed integration. As a result, the planned data flow between NP and LUKE could not be established, preventing further advancement of methodological development, data collection or validation activities. Therefore, no new products or results are reported for this component in the second iteration.

NP will explore a potential evaluation of LUKE's Digital Twin model using their in-situ data, with more info about this potential use case provided in D4.4, chapter 2.5.3.

3.2.4 Calculated sustainability indicators

The functionality for crop sustainability evaluation developed and improved by Horta in previous research projects includes several sustainability indicators, which are calculated for the wheat parcels monitored in the ScaleAgData project.

Sensor input data:

Weather data, measured by in-situ weather sensors, as well as soil data and farm logs data entered by the user in the Decision Support System grano.net® (https://www.horta-srl.it/en/grano-net/) are used as input for the calculation of sustainability indicators.

Table 20: Overview of sensor input data (sustainability indicators)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
Weather data (Temperature, Relative Humidity, Rain)	Weather stations (different models)	Horta	North Italy	10	Hourly	Cropping seasons 2023-24 and 2024-25 are considered in the project

EO input data:

No EO data is used to calculate sustainability indicators.

Methodology and validation results:

Weather data measured from in-situ sensors need to be checked for time series completeness and correctness, to have data ready to be used as input for the models in the grano.net® Decision Support System (DSS), including sustainability indicators calculation. Horta has an internal quality control procedure, that is automatically applied to data retrieved from the weather stations installed in field, with the main aim to fill in gaps in the data and exclude anomalous values. To identify anomalous values, the procedure operates several comparisons among the registered data and weather forecast data, climatological data, historical data of the same weather station, and data of nearby weather stations. The comparison is operated for each hourly value retrieved from the weather station. If the data passes all the checks, it is considered validated and can be part of the time series used as models input. If the data does not pass the checks, the data is considered anomalous, and it is discarded. Discarded, or anomalous data are substituted with forecast data, which are retrieved from an external provider for each of the weather stations.

Sensor- integrated data products:

The DSS grano.net® includes a functionality for calculating sustainability indicators, useful to monitor environmental impact of the activities carried out in the fields. The initial set of 20 indicators was developed by HORTA capitalizing results from the EU-funded projects PURE and INNOVINE, which were then adapted to wheat crop. The set of indicators available on the DSS was then expanded, and presently includes more indicators, such as the Product Environmental Footprint (PEF) ones.

The initial set of indicators is divided into six domains: Air, Biodiversity, Energy, Human Health, Soil, Water, measuring the impact of crop management on each domain. Indicators in the Air compartments are related to the measurement of the amount of greenhouse gases emitted in connection to human activities, and the estimation of the amount of carbon sequestered by plant tissues during the growing season. In the Biodiversity compartment, the indicators evaluate the farm's biodiversity, based on the different types of land use and assess the chemical ecosystem hazard score. In the Energy compartment, indicators consider the amount of fuel used for the mechanized operations carried out in the field, the use of fuel from renewable sources, and the farm's waste management. In the Human Health compartment, indicators evaluate the chemical hazard to humans, the exposure of individuals to chemical products, and the surface treated with pesticides. In the Soil compartment, indicators estimate the risk of soil compaction; the loss of soil due to water-caused erosion; the duration of soil coverage; the percentage of organic matter contained in the soil; evaluate land surface necessary to provide the resources. In the Water compartment, indicators assess the water footprint; evaluate the type of irrigation system and water used for crop irrigation; the emissions of compounds causing acid rain and the effect of nutrient excess on water ecosystems. The PEF indicators assess in a multi-criteria perspective the environmental performance of agricultural activities, throughout the life cycle approach (https://green-business.ec.europa.eu/environmental-footprint-methods-0_en). The indicators assess the global warming potential; the ozone depletion potential; the impact on human health and ecosystem; the effect on freshwater and marine; impact on soil quality; and it evaluates resource depletion. For each compartment, the recommendation for the farmer to decrease the environmental impact of the cultivation phase, choosing among the available cultivation techniques, the ones that lead to a lower impact on environment.

Sustainability indicators are presented to the user in both a graphical way, using a radar graph (Figure 49) and in the form of a table (Table 21). The graph allows the user to visually understand the sustainability performances of the crop management carried out in the field, and to identify the compartments in which there is potential to decrease the impact by the selection of different management actions.

Figure 49: Visualization of sustainability indicators in the DSS grano.net ®

Sustainability indicators were calculated for the parcels monitored in the second Crop Management RIL, for both 2023-24 and 2024-25 cropping season considered in the project.

Table 21: Sensor integrated data products (sustainability indicators)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sustainability indicators	Parcels cropped with wheat in North Italy	Wheat cropping seasons 2023-24 and 2024-25	Calculation is done once per cropping season, after harvest.	Calculation is performed for each parcel, on the base of weather data recorded from a nearby weather station, soil data and farm log manually entered

Use case(s):

In the second Crop Management RIL, sustainability indexes were calculated for the parcels monitored in the Sub Lab activities. Field managers (farmers, technicians, agronomists, can visualize the sustainability performance for each parcel, and then decide to take action to adapt the crop management to improve sustainability.

3.2.5 DSS model outputs

The Decision Support System (DSS) grano.net® is an online tool developed by HORTA in previous years aiming to help farmers in the sustainable management of their wheat crops, allowing them to take informed decisions for all the main steps in the crop management. The DSS is used for monitoring wheat parcels in the ScaleAgData project.

The DSS has available several models and algorithms providing, among the others, indications on fertilization, water balance, disease risk, and crop yield and quality prediction. The models in the DSS run for the parcels selected for monitoring in the RI Crop Management Sub Lab 2b, for both cropping seasons considered in the frame of the project (2023-24 and 2024-25).

Yield and disease models are improved and upscaled in ScaleAgData by integrating EO data. EO data (NDVI) from own provider were used as input for the fertilisation model, allowing it to identify homogeneous zones in the parcel, and to deliver variable rate application maps. The DSS had an existing model based on in situ data of soil and crop management, able to calculate the optimal timing (based on crop phenology) and amount of nitrogen to be distributed to the crop, considering the Nitrogen balance in soil. A new fertilization functionality was developed for the DSS, coupling the outputs of the existing model, and EO (NDVI) data able to define homogeneous zones in the field. The new functionality is able to identify if field variability is high enough to justify a variable rate application of nitrogen, the number of homogenous areas, the suitable amount of nitrogen for each area identified, and provide a prescription map.

Sensor input data:

Weather data, measured by in-situ weather sensors, as well as soil data and farm logs data entered by the user of the Decision Support System grano.net®, are used as input for the calculation of the models and algorithms in the DSS.

Table 22: Overview of sensor input data (DSS)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
Weather data (Temperature, Relative Humidity, Rain, Leaf wetness)	Weather stations (different models)	Horta	North Italy	10	Hourly	Cropping seasons 2023-24 and 2024-25 are considered in the project

Table 23: Overview of EO input data (DSS)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
NDVI	External data provider	10 wheat parcels per season in North Italy	Cropping seasons 2023-24 and 2024-25	3-7 days	20m
ET, SM	DHI	10 parcels cropped with wheat in North Italy	Cropping season 2023-24	Daily	20m
Yield potential index	VITO	10 parcels cropped with wheat in North Italy	Cropping season 2023-24	Once in the season	20m

Methodology and validation results:

Models and algorithms in the DSS use weather data, soil data and farm logs input. Weather data are measured by in-situ sensors and are checked according to the procedure described in section 3.2.5. Soil data and farm log data are manually entered into the system by the user.

The **first iteration** took into consideration the inclusion of EO information (i.e. NDVI indexes) as model input, taking into consideration the existing wheat yield model, which was already implemented in the DSS and relies on weather information measured from in situ sensors. The work carried out was aimed at understanding the potential improvement of the yield model, adding EO vegetation indexes as inputs. EO data can provide the model with NRT vegetation information, so potentially their use can lead to improved model outputs compared to the sole use of weather data. Comparison of yield model output with and without the use of EO derived vegetation indexes highlighted an improved model performance only in cases in which crop canopy was highly reduced during the cropping season, while in average conditions model outputs were not significantly improved.

In the **second iteration**, the focus was posed on the use of EO data for the improvement of the fertilization functionality in the DSS, allowing for the development of an algorithm able to generate variable rate prescription maps. The fertilization advice already present in the DSS provides the user with the total amount of fertilizer to be applied to wheat crops, and the splitting of the nitrogen dose in multiple applications to be realized during the most suitable crop phenological phases. The fertilization advice was improved by using EO observation as an input data, allowing for more precise use of technical inputs. The new functionality developed allows the user to identify if the variability inside the field is relevant enough to justify a variable rate application of the fertilizers. Then, according to the machine characteristics, the system allows us to set a defined number of homogenous zones (3 to 7) in the field and provides the amount of fertilizer to be applied in each of the zones, in the form of a downloadable prescription map. In doing this, the algorithm considers the crop phenological phase and weather data.

Sensor-integrated data products:

DSS models were calculated for each of the parcels monitored in the second Crop Management RIL, for the two cropping seasons considered (2023-24 and 2024-25). The models and algorithms aim to support the farmer in making informed decisions about wheat crop management.

Table 24: Overview of sensor-integrated data products (DSS)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution
DSS models output for yield, diseases, fertilization, water balance.	10 parcels cropped with wheat in North Italy in two cropping seasons	Wheat cropping season 2023-24 and 2024-25	Calculation was done during the cropping season, depending on the type of model or algorithm.	Each monitored parcel relies on data of a nearby weather station, soil data and farm log input from the farmer.

Use case(s):

In the second Crop Management RIL, DSS runs for the parcels monitored in the Sub Lab activities for both the cropping seasons considered (2023-24 and 2024-25). Field managers (farmers, technicians, agronomists) can receive support from the DSS for the management of their wheat crops. The DSS supports their decisions regarding the main operations to be performed in the field (i.e. fertilization, plant protection interventions against diseases) and allows them to have information about the water content in soil and forecasts on the crop yield.

3.2.6 Statistical data on the accuracy of observations of the occurrence of agrophages

This product focuses on assessing the accuracy of field observations of agrophage occurrence by combining in-situ inspection data with meteorological information and disease model outputs, in

order to support pest and disease monitoring and advisory activities. In-situ data collected by agricultural advisors through the eDWIN Agrophages mobile application, including pest and disease occurrence records, BBCH growth stages, pest development stages, severity assessments and plant colonization percentages, are integrated with meteorological sensor data from the eDWIN platform for statistical analysis and validation purposes.

Sensor input data:

The main data source is the regularly collected field-inspection data from the eDWIN Agrophages mobile application, where specialists document pest occurrences at predefined sampling points. These points are set in the eDWIN web platform and are inspected at least once per week. Data from the mobile app is combined with measurements from physical sensors to support statistical analysis.

The reference data comes from test fields in the Wielkopolska district managed by WODR. In selected seasons, these fields were planted with corn, wheat, and sugar beet, and the observation data from these plots serves as validated reference information.

Table 25: Overview of sensor input data (agrophages)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
Meteorological data	eDWIN platform	PSNC/WODR	~100	3	hourly	2021/22/23/24/25

EO input data:

No EO data is being used in this case.

Methodology and validation results:

The methodology follows the same approach used in previous implementations of the DSS models. It relies on combining locally collected field data with disease model outputs. When both datasets align, the monitoring point is considered valuable and should be prioritized for future observations to support pest-model accuracy.

All statistical tools used in the analysis are grounded in physical validation of results, ensuring they accurately describe the real-world phenomenon. These tools help to identify areas where local field measurements and disease model outputs are consistent.

Season-to-season data differences are expected due to varying agricultural practices in the local plant production market.

Field observation reports should include:

- pest or disease name
- plant development stage (BBCH scale)
- pest development stage
- percentage of plant colonization
- pest occurrence severity
- optional notes
- date of the report
- indication of whether chemical treatment is needed

Sensor data products:

The data product combines meteorological statistics with physical observations, synthesizes it spatially and creates simple and useful DSS module with relevant representation on the user interface, which can indicate the accuracy of local indications of pest or disease occurrence compared to previously validated pest or disease models.

Table 26: Overview of sensor-integrated data products

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Accuracy of observations	20 parcels	1 season	Weekly/daily		

Use case(s):

The data products are available as micro services ready to be deployed locally. Each test is organized in expert applications, presenting "experimental data" in experts module concerning pests monitoring. Application administrators or granted experts with proper rights are accessing and accepting results of implemented tools.

Results:

As part of the Crop Management agrophages RIL within the ScaleAgData project, we have designed a set of tools aimed at increasing the effectiveness and flexibility of pest and disease monitoring in agriculture. The following sections describe the key solutions implemented:

The Monitoring Dashboard is a central feature of the system, enabling users to select any reference point, either an existing monitoring location or a custom spot on the map. Once selected, the dashboard displays key information for advisors, including disease model forecasts (using both data from physical weather stations and interpolated data), nearby threat reports, records of plant protection treatments, and NDVI data for surrounding fields. This provides a unified and efficient view of regional crop health.

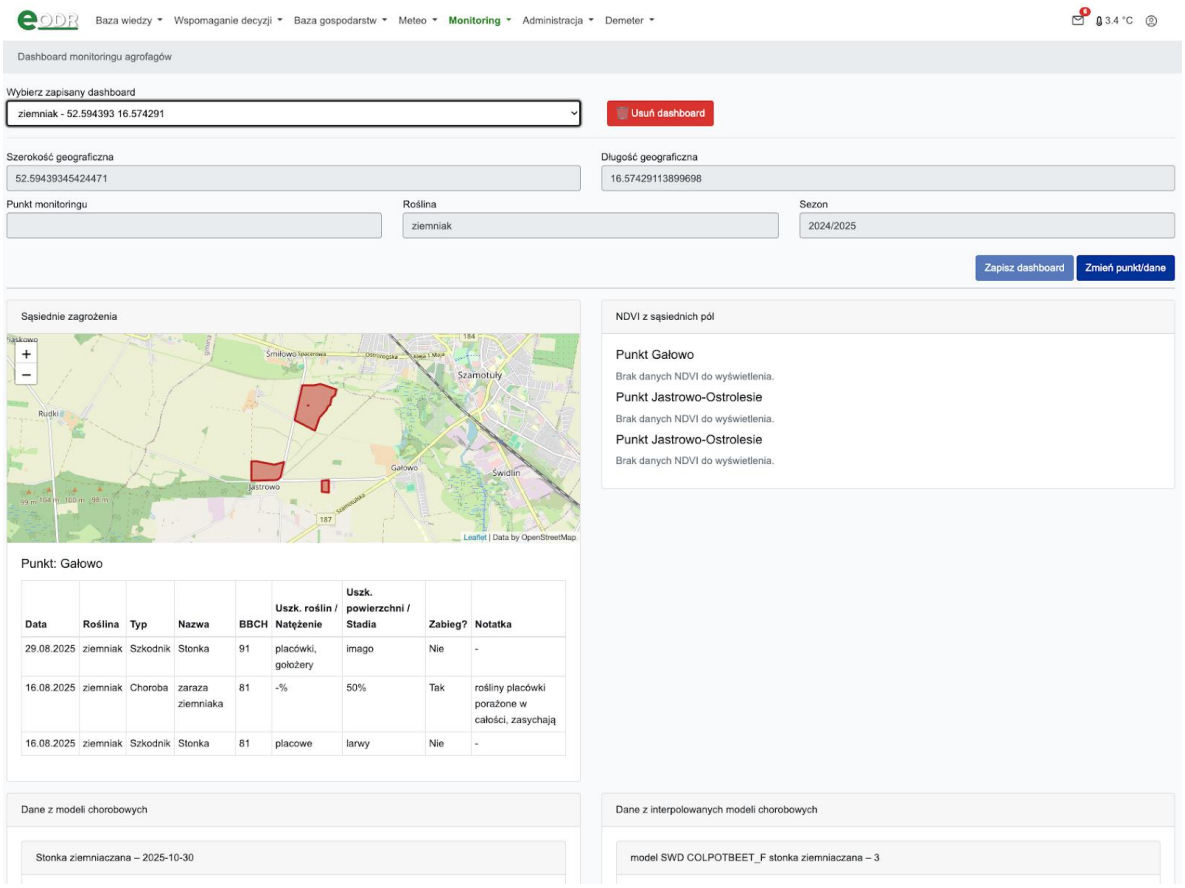


Figure 50: View of the monitoring dashboard with the selected monitoring point for potato

The dashboard generates weekly email notifications, each containing a PDF summary of the same data shown in the interface. These summaries include disease model results, threat reports, treatment information, and NDVI values, along with a link to the online dashboard for deeper analysis.

The system supports ad hoc reporting, allowing advisors to submit observations from any location, not only predefined monitoring points. This flexibility is particularly valuable in regions with low threat levels, where traditional reporting is insufficient.

Weekly regional reports can be created by coordinators to summarize the phytosanitary conditions of a region. The system automates statistical calculations and generates an editable report draft, significantly reducing manual workload.

To enhance accessibility, an automatic notification system alerts users when a new regional report is published. Report data is also available via API, allowing integration with other platforms such as public information screens or agricultural advisory websites.

Monitoring Dashboard Survey Summary:

Feedback from the users has been gathered in the form of a survey with the results presented below.
Visibility:

- Visible to 7 out of 8 advisors; one exception likely due to role/permission or cached settings.

Overall Usefulness:

- Intuitive and highly useful.
- Combines map data, disease models, NDVI, neighbouring threats, and phenology in one interface, improving field readiness and reducing unnecessary reports.

Ad-hoc Reporting:

- Valuable for quick reactions, especially when away from assigned monitoring points.

Coordinators' Feedback:

- Dashboard helps analyse advisor activity and monitoring quality.
- Automatic and ad-hoc reports are logical and necessary.

Key Findings:

- Dashboard works well and improves efficiency.
- Disease models are essential for assessing pressure, planning visits, and interpreting other data.
- Neighbouring threats feature is promising but needs more detail and filtering.
- NDVI and phenology are highly appreciated "bonus" datasets.
- Weekly reports reinforce consistent monitoring.
- Ad-hoc reports increase mobility and reduce low-value visits.

Development Recommendations:

- Map/Data: Add filters for neighbouring threats, enhance popups with BBCH stage, treatment date, crop variety, location accuracy.
- Disease Models: Provide explanations and guidance on interpreting outputs and risk levels.
- NDVI: Allow date switching, show acquisition date, add explanation/legend.
- Phenology: Verify GPS detection and permissions on mobile.
- Ad-hoc Reports: Include all crops with pest/disease categories and allow optional photos.
- Coordinator Module: Add report filters and Excel export capability.

User-Reported Issues:

- Navigation inconsistencies limited neighbouring threat details, incomplete pest/disease tabs, GPS/location problems, missing NDVI context, mailer issues for weekly reports, and lack of interpretation guidance for disease models.

Overall Conclusion:

- The dashboard is positively evaluated, providing meaningful support for advisors' work.
- Most valued: disease models, ad-hoc reports, and combined view of phenology, NDVI, and neighbouring threats.
- Key improvements: extend data detail, improve navigation, refine UI/GPS, ensure feature availability for all crops.

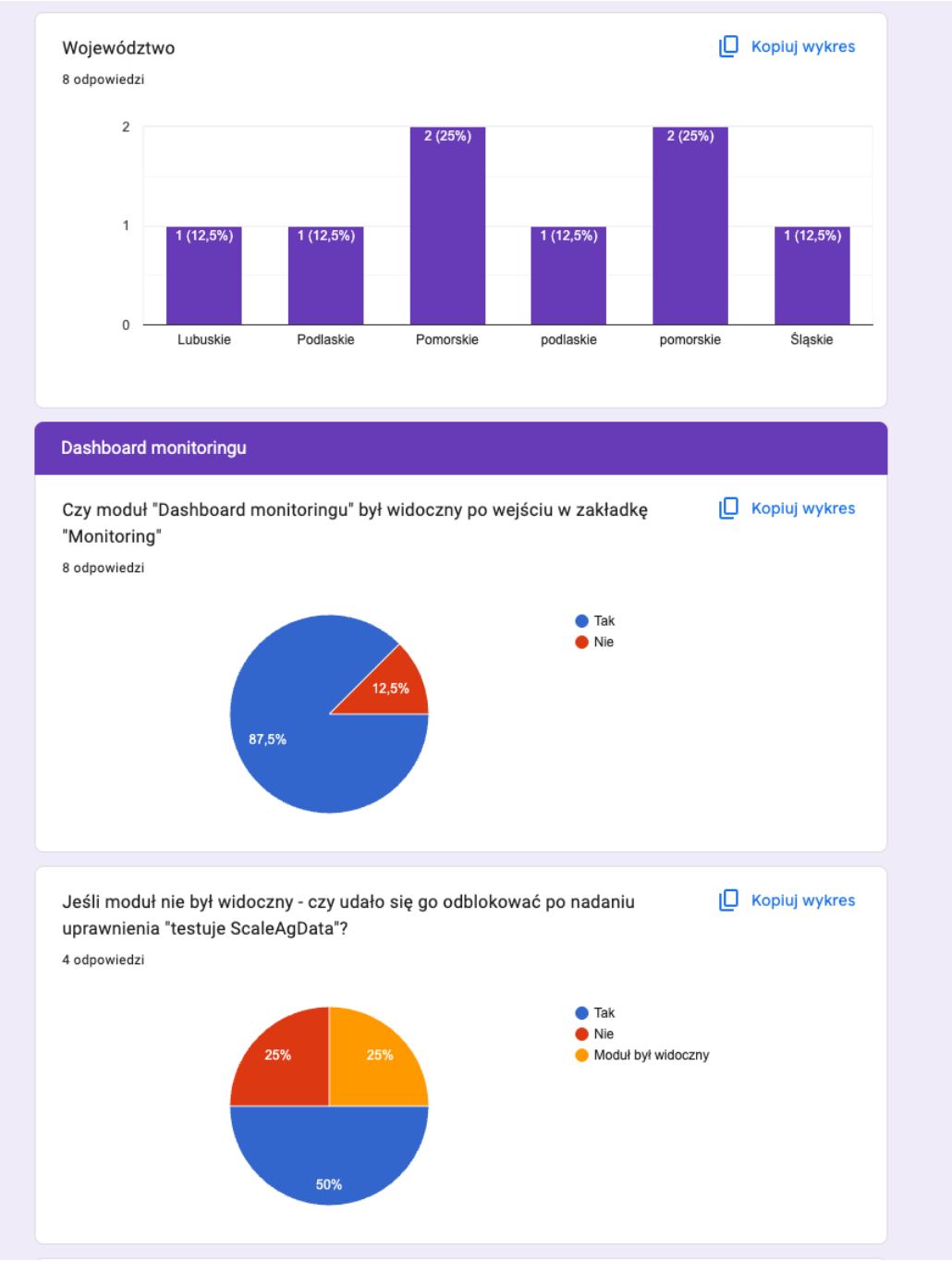


Figure 51: Results of the survey evaluating the new dashboard features

The survey was conducted using the Google Forms platform and distributed to the test users of the monitoring dashboard. It also served as a step-by-step guide, leading users through the new elements of the system while collecting their feedback.

3.2.7 Improved predicted agrophage occurrence data based on geolocation

The previous solution had a limitation: results were based on the nearest meteorological station rather than the specific field. This improved product improves pest and disease risk estimation by introducing geolocation-based predictions that combine interpolated meteorological data, Earth Observation inputs and system data, allowing risk assessments to be generated for any selected location.In-situ

observations collected through the eDWIN Agrophages mobile application, including pest and disease reports, phenological observations and records of plant protection treatments, are integrated with Earth Observation image processing outputs, including Sentinel-2–derived NDVI and high-resolution precipitation products (RainGRS), to improve representation of vegetation status and moisture conditions relevant to pest development.

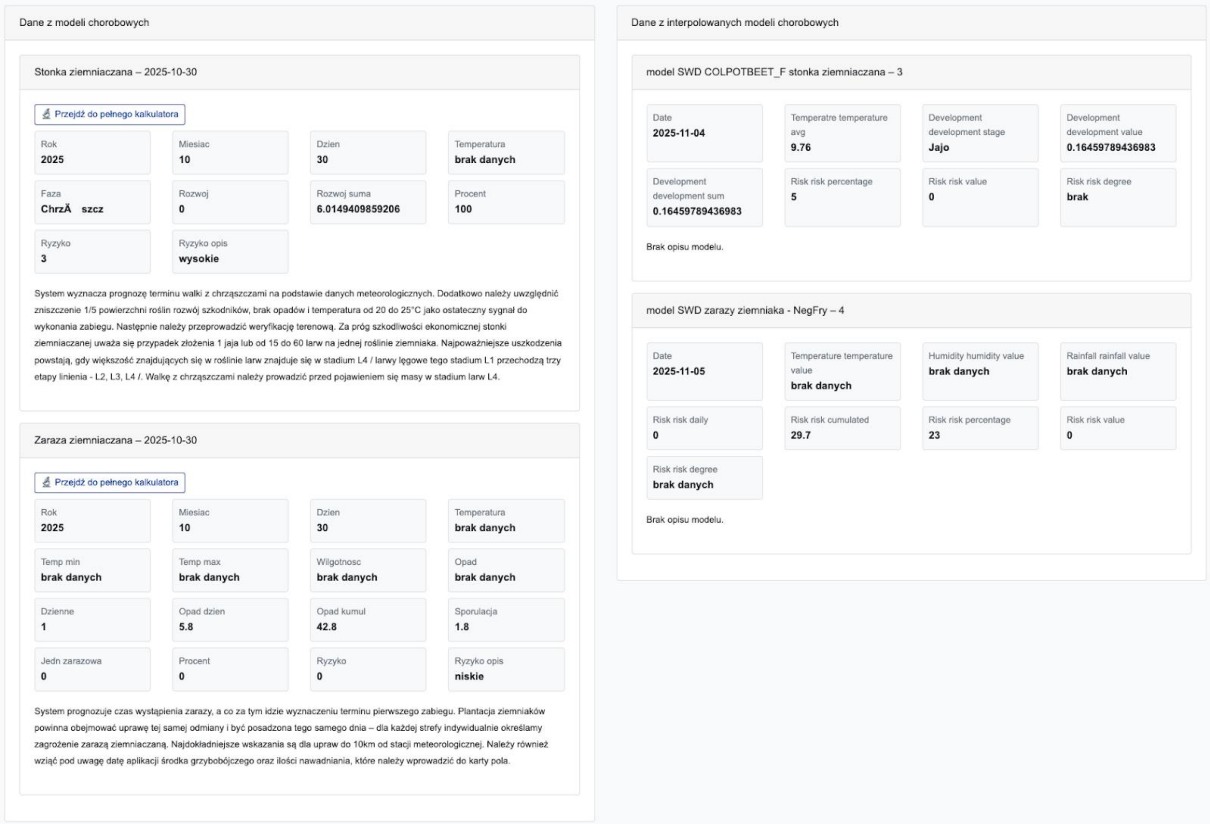


Figure 52: View of the monitoring dashboard summarizing the results of mathematical model calculations based on physical data and interpolated data

The monitoring dashboard now includes an additional section located next to the section with models calculated on physical meteorological data. This new view allows us to compare model results and assess the threat level on the advisor's monitored field with even greater precision.

We also plan to adopt the Few-Shot Learning framework as an additional component of the pest-monitoring dashboard. This approach offers a promising method for identifying visual changes occurring in agricultural fields. Training the model will require combining two key data sources: observations collected through our pest-monitoring activities and satellite imagery from the Copernicus programme. We have already begun analysing the available Copernicus datasets to assess their suitability for this purpose.

The Few-Shot Learning component is expected to support early detection of anomalies by generating alerts whenever relevant visual changes are detected in surrounding fields. In the initial phase, we will apply this method to satellite data for fields identified by our existing algorithm, which locates neighbouring fields monitored by other advisors. These fields are accompanied by extensive metadata, which will further support model development and improve the interpretability of the results.

Data from test fields vary across seasons due to differing local plant production strategies.

Sensor input data:

The data product will use two types of sensor-based information with a common timespan and located in a limited radius (area of interest):

Table 27: Overview of sensor input data

Sensor data	Source	Data provider	AOI/test sites	Nr. season	Meas. frequency	Season(s)
Meteorological data	eDWIN platform	PSNC/WODR	~100	3	hourly	2021/22/23/24/25
Phenological observations	eDWIN platform	PSNC/WODR	8	1	daily	2024/2025

EO input data:

The data product uses two types of EO data. The NDVI is calculated from Sentinel-2A and Sentinel-2B images using data from Band 4 and Band 8. The EO product named RainGRS is a long-term multi-source precipitation estimation with high resolution (reference link: https://amt.copernicus.org/articles/16/4067/2023/amt-16-4067-2023.pdf) creating a possibility to upscale calculations from selected points to complete areas of interest.

Table 28: Overview of EO input data

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
NDVI	PSNC	29 parcels cropped with wheat, beet and corn in West Poland	From seasons 2021/22/23/24/25	5 days	10m
RainGRS	Open Data (IMGW-PIB)	selected area	From seasons 2021/22/23/24/25	hourly	1km

System data:

Users have the possibility to report agrophages observed on the fields and to create a registry of performed protection treatments. Collecting the closest reports and treatments for an observation point gives information about an agrophage that could potentially be a threat to the point. This data is added by users and is not verified, so it may not be correct. The range is determined by the type of the agrophage.

Table 29: Overview of system data

System data	Source	Data provider	AOI/test sites	Seasons
Spraying treatments	eDWIN platform	PSNC	depends on the type of agrophage	2021/22/23/24/25
Agrophage reports	eDWIN platform	PSNC	depends on the type of agrophage	2021/22/23/24/25

Methodology and validation results:

The process starts with interpolation of input data, followed by data fusion. Machine learning is then applied to estimate the probability of agrophage occurrence.

- Kriging with a trend is used to interpolate temperature and humidity.
- User-submitted agrophage reports and protective treatments are point data and are incorporated through data fusion rather than interpolation, using the range and conducive conditions of agrophage development.
- The validation set consists of physical observations from advisors on 29 selected fields.

Sensor-integrated data products:

A risk evaluation model is created to anticipate risks, which makes it possible to get ahead of the validation stage.

Table 30: Overview of sensor integrated data products

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Probability of an agrophage occurrence	29 parcels cropped with wheat, beet and corn in West Poland	In growth season	daily	4km	-

Use case:

The data products are available as micro services ready to be deployed locally. Tests are organized in experts' applications, presenting "experimental data" in experts module concerning pests monitoring. Application administrator or granted with proper rights expert are accessing and accepting results of implemented tools. Advisors receive daily risk assessments with a specific focus on pest occurrences within their designated work areas, facilitating targeted responses. Farmers benefit from this system by receiving precise information about threats in their locality, enabling timely and effective action.

Results:

As part of our Crop Management (agrophages) RIL, we have developed tools to support pest monitoring and operational crop-protection decisions. Data products are integrated into expert applications, where experts and administrators can access and review outputs related to pest occurrences. Advisors receive daily risk assessments for their designated areas, allowing them to respond quickly and effectively to emerging threats.

The system includes a Monitoring Dashboard, where advisors can view pest risk information for specific locations. Ad-hoc reporting allows advisors to submit field observations from any location, ensuring timely updates to the monitoring system. Automated notifications ensure that advisors remain informed about current pest risks within their work areas.

In the latest system update, an interpolation process for meteorological data was introduced. Instead of relying solely on the single, nearest meteorological station, the system now interpolates data from multiple nearby stations. This improves accuracy and ensures that local conditions are represented more realistically for any selected point on the map.

3.2.8 Predicted overall level of agrophage occurrence risk for the selected region

This product focuses on providing a regional-level overview of agrophage occurrence risk by abstracting and aggregating field observation data into a single assessment for selected administrative units, supporting situational awareness for farmers, advisors and supervisory bodies. The product aggregates validated in-situ field observations and geolocation-based predictions enhanced through Earth Observation inputs, including Sentinel-2–derived NDVI and precipitation products, into regional risk indicators.

Input Data:

This data product integrates two key components: (i) statistical information on the accuracy of field observations of agrophage occurrence, and (ii) geolocation-based predictions of agrophage presence. These datasets are merged using interpolation and spatial aggregation techniques to produce consistent and reliable regional estimates.

Sensor-integrated data products:

The resulting output is an estimated level of agrophage risk for each voivodeship (highest administrative unit in Poland) or county. This indicator focuses on a selected disease within a specific crop and represents a synthetic, high-level assessment suitable for regional monitoring. The product is designed for automated dissemination in the form of alerts, issued at least weekly, which can be published through user interfaces or exchanged between systems.

Use case(s):

The primary use case is integration into farm management applications, where the risk indicator is displayed on a map showing conditions for each administrative unit. This enables farmers, advisors, and managing authorities to quickly assess current regional threat levels. In addition, managing and funding bodies can use the aggregated statistical outputs for supervisory actions, reporting, and broader decision-making processes. This ensures that all relevant stakeholders have consistent, minimal essential information to guide plant-health-related decisions.

Results:

Within the RIL we have developed tools that generate regional pest and disease risk assessments to support both operational crop management and oversight functions. The system computes aggregated risk values for selected crops and diseases at voivodeship or county level, combining validated field observation statistics with predicted agrophage occurrence patterns, using interpolation and spatial aggregation techniques to ensure reliable and realistic estimates.

Risk data is presented in a map-based dashboard within farm management applications, allowing managing and funding entities to quickly visualize current conditions across administrative units. Weekly alerts are automatically generated and distributed, ensuring stakeholders receive timely information for supervisory and management decisions.

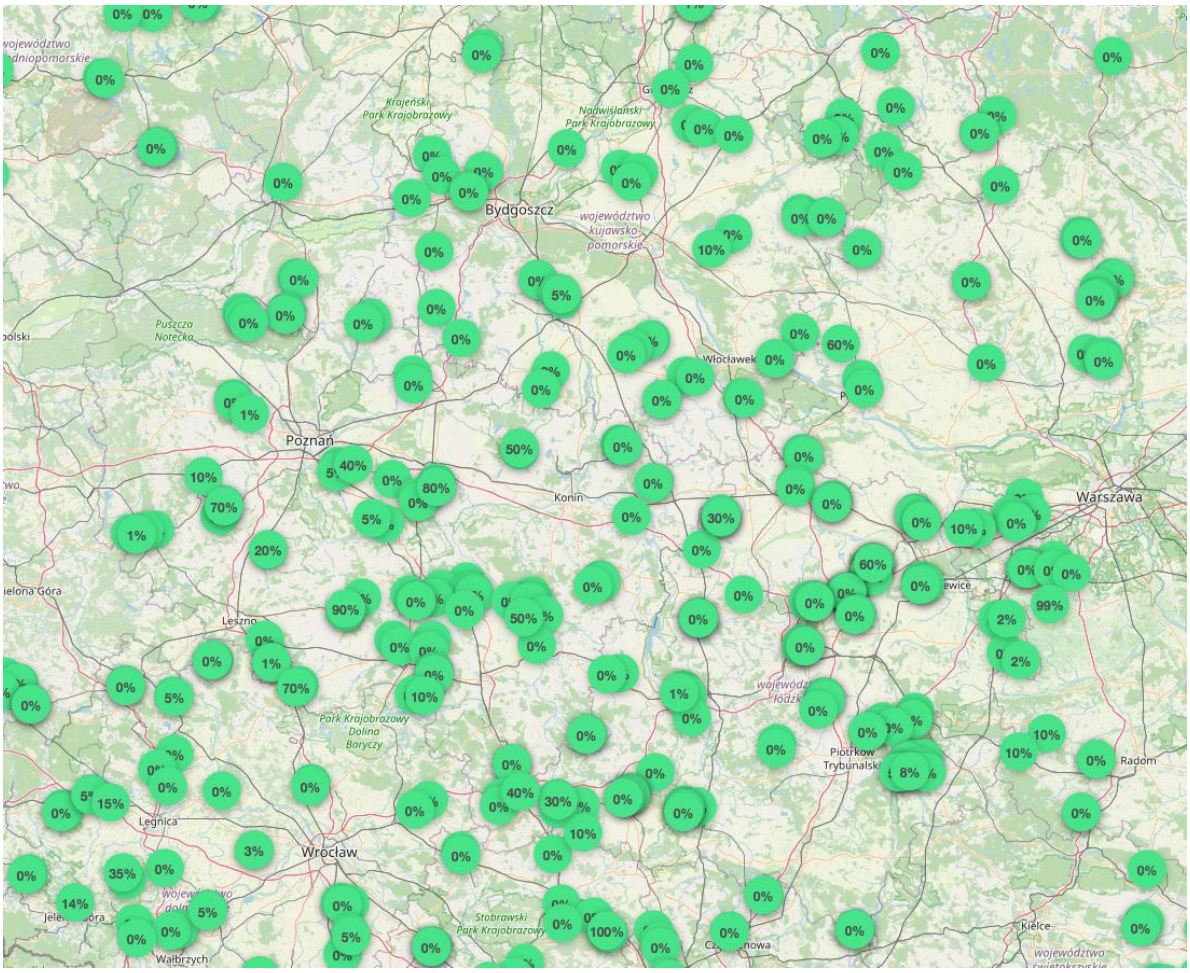


Figure 53: View of the pest monitoring map; this view will also include regional information in the future

The set of points, as shown in the screenshot, will still be visible, but additional information, e.g., in the form of coloured blocks superimposed on selected counties, will be available to the user. They will be able to observe the regional threat level on it.

Experts and administrators can review and validate the aggregated outputs before publication, maintaining the credibility and accuracy of the data. This approach ensures that all relevant stakeholders have the essential information needed to make informed decisions about plant health management at the regional level.

3.3 RIL Yield monitoring

3.3.1 Potato yield estimates

This product, developed by VITO in collaboration with AVR, includes potato yield estimates at subfield-level (10m resolution) at harvest time,showing yield variability within the field. It integrates yield sensor data collected using AVR's potato harvesters with EO (Sentinel-1/2) and meteo data. It therefore makes use of the PRESTO framework for few-shot learning. The final goal is to apply the yield estimation model to other potato fields in the region (without yield sensors) and aggregate the results into regional yield statistics.

Sensor input data:
Gross yield sensor data is collected using AVR's Puma 4.0 harvesters.

Table 31: Overview of in situ input data (yield)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Nr. training data
Potato yield (kg/ha)	Puma 4.0 harvester	AVR	Belgium, Netherlands	+/- 2000	Once, at harvest	3851 Fields, 282844 subfields

Currently, field boundary data is not available across all jurisdictions. As such, fields are automatically calculated based on the harvester data, which is received in CSV format and contains info on yield, collection time and location. The harvesters use an internal calculation to return a yield in kg/ha, based on the area of each collection point. Field boundaries are selected using government-reported Land Parcel Identification System (LPIS) boundaries for each year, filtered to the relevant crop type. Fields from these datasets are selected where there is an overlap with the points in the database. It will then add this grouping as a field and move on to the next points. These fields are divided into 20 m x 20 m subfields, whilst removing any calculated subfields which are not fully contained within the dataset. This resolution is considered sufficient for aggregating satellite data to train and validate the models.

For model development, we focused on Belgium and the Netherlands. We used harvest data collected in 2022 and 2024 – although data from 2023 was made available, there were quality issues preventing its use in the modelling workflow. The model is built in such a way that, in the future, the data can easily be extended to include sites from other European countries as well as different years. The inclusion is dependent on provision by suppliers of both harvester and EO input data for these various regions and timeframes. The dataset is constrained to only values between 0 and 200000, as anything outside of this is considered an outlier. The figure below shows the distribution of values across the dataset.

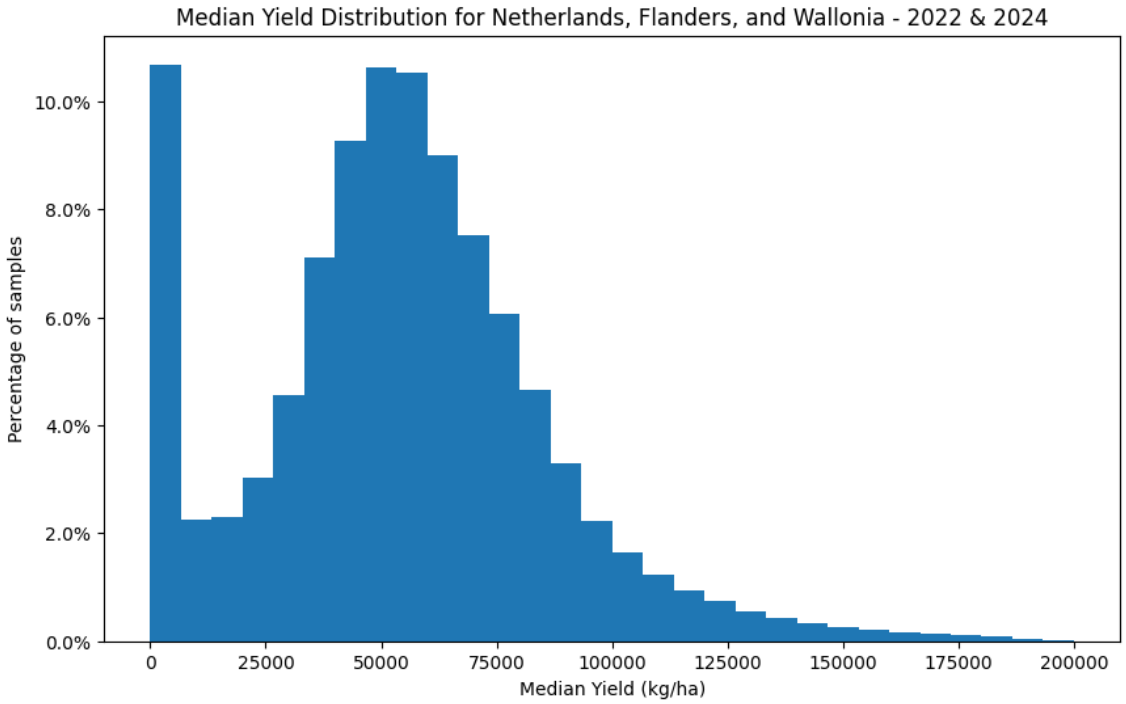


Figure 54: Histogram showing frequency of yield values within Belgium and Netherlands subfields, percentagewise, at 20m resolution. Data is cut to 200000 to remove the impact of large outliers. In this instance, very low values are not yet removed and will affect modelling capability.

Table 32: Overview of the study regions

Nation	Fields	Sub Fields	Harvest Points (Average) Per Subfield	Year
Belgium	1358	117307	6.914.837	2022
Netherlands	582	53487	116.174	2022
Belgium	610	22101	125.0983	2024
Netherlands	769	49727	137.2311	2024

More details on the sensor data can be found in the deliverable D3.2.

EO input data:

The following EO data are used as input for potato yield modeling:

Table 33: Overview of EO input data (yield)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel-1 and Sentinel-2 imagery	VITO	Belgium, Netherlands	All Years	10-Daily composite	10 m (subfield aggregated)
AgERA5 meteo (temperature, rainfall)	VITO	Belgium, Netherlands	All Years	10-Daily composite	10 m (subfield aggregated)

Copernicus DEM (slope, altitude)	VITO	Belgium, Netherlands	All Years	Single moment	20m

More details on EO data, source and processing can be found in deliverable D3.3.

Methodology and validation results

Data Preparation

As data sources are collected at different temporal and spatial resolutions, they must be processed to a common time-series format suitable for modelling. The actual timeseries can be user-defined, which will both restrict the start and end of the timeseries based on the desired inference period, and define the temporal resolution. Timeseries variables are first adjusted to fit the correct temporal resolution by compositing higher resolution data into a standard interval (e.g., 10-day periods). This ensures that all variables align to the same temporal scale. Subsequently, the time-series data is interpolated using nearest-neighbour matching to synchronize the exact temporal points across all variables. For example, even if one dataset originally covers a period from day 1 to day 10 and another from day 2 to day 11, the interpolation process adjusts these periods so that all datasets match precisely to cover consistent intervals, such as from day 1 to day 10. For static variables (e.g. elevation, latitude, longitude), these are simply extrapolated across the length of the timeseries.

These datasets are individually saved and can then be used for testing with several model architectures or used to create feature embeddings for ML models. Datasets are to be stored and versioned in a Data Versioning Control (DVC) system.

An issue with the current dataset is that harvesters are susceptible to errors induced by environmental conditions such as soil moisture or soil type, as well as different adherence to calibration procedures by the farmers who operate the harvesters. Model development during the second iteration focused on improving data confidence. The data provider provided a list of harvesters operated by farmers considered to be strictly adherent to required calibration procedures and thus minimize risk of error in the dataset.

Feature Extraction

A challenge of using multiple sensors across a time-series, and having relatively limited sample sizes, is that the total input features can become too large for model architectures to handle well. Whilst model architecture will create its own embeddings before outputting the final regression or classification layer, these may be suboptimal. “Feature embeddings” aim to first extract useful features, which can then be used as lower-dimensionality model inputs. In the case of yield estimation, PRESTO (section 2.3.1) is used to extract usable features from the Sentinel and METEO (AgERA5) timeseries, transforming them into 128 feature embeddings.

Yield point cleaning

After unsatisfactory results with initial model experiments, a few approaches were attempted to clean up input harvester yield data. A first approach included an azimuth filtration, operating on the assumption that harvesters will pass in parallel lines during actual harvest of the potato fields, thus the points along parallel lines should be considered as valid, while yield data collected while the harvester was turning, or stationary, or crossing fields in a non-aligned pattern should be discarded as they might be erroneous. At a field level, the lines between individual points were calculated, along with their vector direction. The most common 25% vector directions (representing the predominant harvester lines in each field) were kept, while points belonging to other lines were removed from processing (Figure 55).

A second yield point cleaning approach relies on two statistical cleaning methods to identify outlier points. The first outlier detection layer uses quartiles to identify extreme values. It calculates the first quartile (Q_1) and third quartile (Q_3) then computes the interquartile range (IQR). Lower and upper bounds are defined as (Q_1 – 1.5 x IQR) and (Q_3 + 1.5 x IQR), respectively. Any value outside these bounds is flagged as an outlier. The second outlier detection layer calculates the median absolute deviation (MAD) and scales it by 1.4826. The standardized score for each value is (value - median) / MAD. Values with absolute scores exceeding 3.5 are flagged, excluding any zero or missing scores. If a point triggers either of the outlier detection layers, it is flagged as an outlier and removed from the dataset.

Both yield data cleaning methods were tested separately and together on 2022 and 2024 data, as well as combined data for both years. The methods were also tested against raw harvester data.

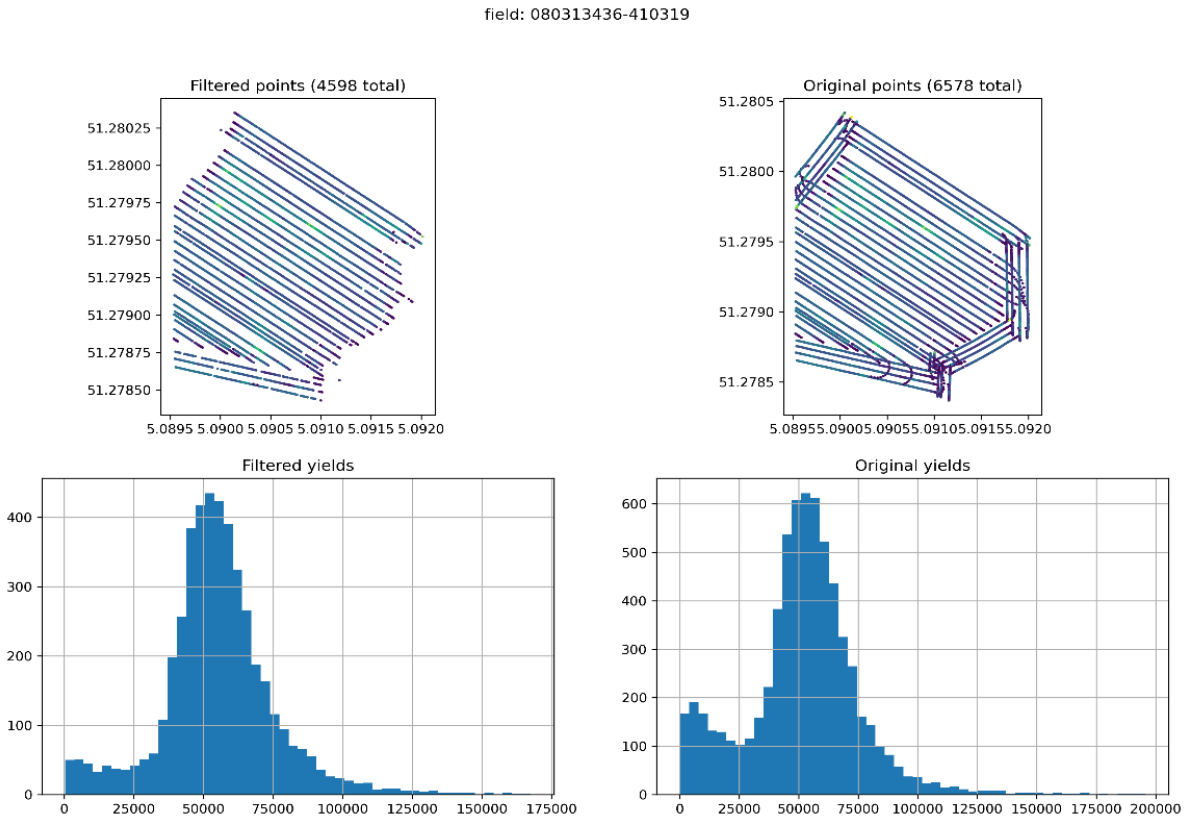


Figure 55: Azimuth filtering method demonstrated on a single field from Netherlands 2024 dataset

Model Framework

As several data sources and feature extraction methods are combined in the context of yield estimation, the model framework is versatile to either timeseries inputs, or static features (embeddings), with the user being provided options to develop and compare different models and their parametrization. The PRESTO model tested and finetuned in this task is also compared against CatBoost, a more traditional machine learning model based on gradient boosting than nonetheless showed promising results with the input data. A third approach is also tested, where the PRESTO model is finetuned on extracted features and input data, and PRESTO encodings are then passed to a CatBoost regressor head to generate predictions. These three approaches are compared for each set of data (covering different years, geographic areas, and yield cleaning methods), allowing us to better understand the accuracy of the PRESTO model predictions relative to other, more traditional methods.

Results: Subfield-level Monitoring (December 2025)

Table 34: Models trained and validated using AVR data extracted in Belgium and the Netherlands for 2022 and 2024, testing different cleaning methods and model types.

Dataset	Model	Cleaning method	RMSE	R^2
Netherlands & Belgium (2024)	19112025-finetuning-LPIS-2024-dk-med-yield	None	24150.4	0.458
	19112025-finetuning-LPIS-2024-dk-med-yield + CatBoost head	None	23938.7	0.468
	Only CatBoost	None	25430.1	0.400
Netherlands & Belgium (2024)	21112025-finetuning-LPIS-2024-dk-med-yield-azimuth	Azimuth	25642.3	0.392
	21112025-finetuning-LPIS-2024-dk-med-yield-azimuth + CatBoost head	Azimuth	25235.1	0.411
	Ony CatBoost	Azimuth	25430.1	0.399
Netherlands & Belgium (2024)	25112025-finetuning-LPIS-2024-dk-med-yield-stats	Stats	25049.3	0.400
	25112025-finetuning-LPIS-2024-dk-med-yield-stats + CatBoost head	Stats	25721.7	0.367
	Only CatBoost	Stats	22502.8	0.515
Netherlands & Belgium (2022)	08122025-finetuning-LPIS-2022-dk-med-yield-orig	None	20338.5	**0.552**
	08122025-finetuning-LPIS-2022-dk-med-yield-orig + CatBoost head	None	19857.9	**0.573**
	Only CatBoost	None	17985.0	**0.655**
Netherlands & Belgium (2022)	10122025-finetuning-LPIS-2022-dk-med-yield-azimuth	Azimuth	23135.3	0.360
	10122025-finetuning-LPIS-2022-dk-med-yield-azimuth + CatBoost head	Azimuth	22996.3	0.368
	Only CatBoost	Azimuth	20822.8	0.481
Netherlands & Belgium (2022)	11122025-finetuning-LPIS-2022-dk-med-yield-stats	Stats	19454.3	0.521
	11122025-finetuning-LPIS-2022-dk-med-yield-stats + CatBoost head	Stats	19364.8	0.526
	Only CatBoost	Stats	17922.7	0.594
Netherlands & Belgium (2022, 2024 combined)	08122025-finetuning-LPIS-2022-2024-dk-med-yield-orig	None	22219.7	0.455
	08122025-finetuning-LPIS-2022-2024-dk-med-yield-orig + CatBoost head	None	22578.8	0.437
	Only CatBoost	None	20805.5	0.522
Netherlands & Belgium (2022, 2024 combined)	11122025-finetuning-LPIS-2022-2024-dk-med-yield-azimuth	Azimuth	22925.5	0.465
	11122025-finetuning-LPIS-2022-2024-dk-med-yield-azimuth + CatBoost head	Azimuth	23131.8	0.455
	Only CatBoost	Azimuth	22139.0	0.500

Netherlands & Belgium (2022, 2024 combined)	12122025-finetuning-LPIS-2022-2024-dk-med-yield-stats	Stats	21863.5	0.4746
	12122025-finetuning-LPIS-2022-2024-dk-med-yield-stats + CatBoost head	Stats	21752.8	0.480
	Only CatBoost	Stats	20518.9	0.537

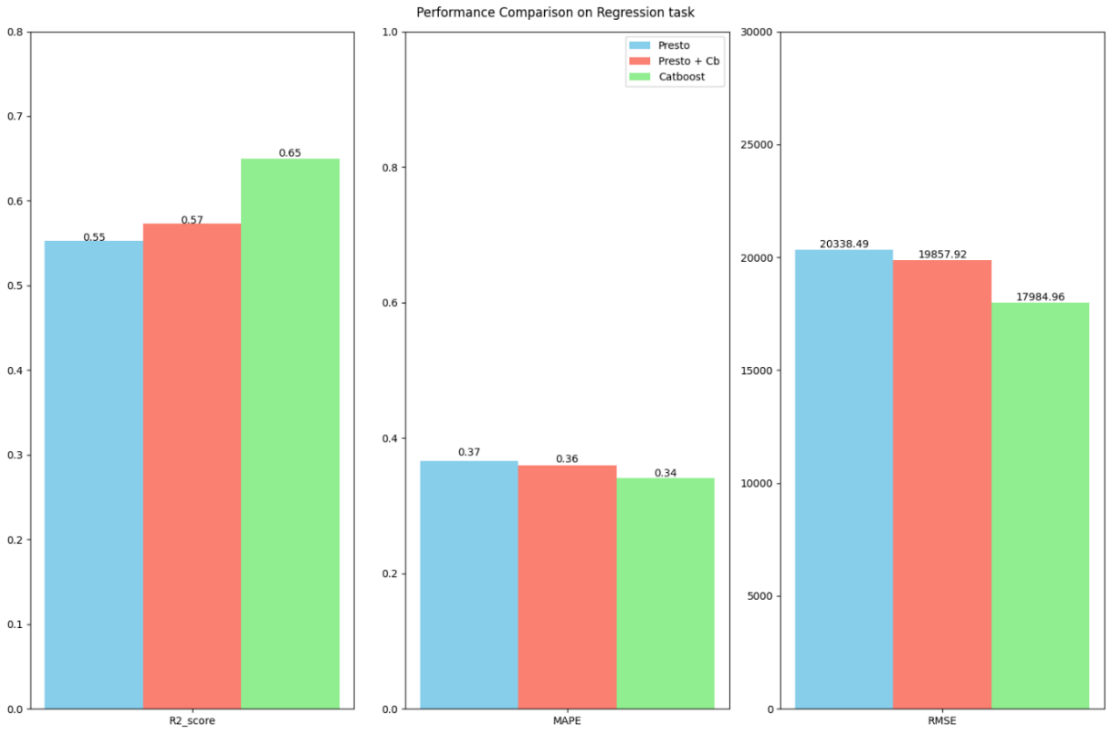


Figure 56: Performance comparison of regression task

Several models were trained using harvester yield data from both Belgium and the Netherlands, from the years 2022 and 2024 and generated PRESTO features. Models were trained using an 80/10/10 split for training, validation, and test datasets, and CatBoost heads/separate CatBoost models were applied on the same datasets in order to compare results against the pure PRESTO model approach. The purely CatBoost models mostly outperformed the pure PRESTO and PRESTO+CatBoost models. The highest performing models were trained on only 2022 data, however the combined 2022+2024 model performed better than the 2024 model alone. Neither of the yield data cleaning approaches improved model results in any of the model sets. Generally, R^2 scores were quite low, while RMSE scores still remained high under all modelling approaches. The poorer results from 2024 may be explained by the fact that it was an unusually wet year, causing issues with potato harvests across both the Netherlands and Belgium which may have had an impact on the final yield data.

The best performing purely PRESTO model, trained on the original (not cleaned) 2022 yield data, shows an acceptable distribution of predicted yields versus actual yields (Fig. 56). Of the independent test dataset, 29% of 20x20m tiles were predicted with residuals smaller than 5000 kg/ha, while 51% were predicted with residuals smaller than 10000 kg/ha. Nonetheless, there still remains significant uncertainty in the model predictions. The retraining of the PRESTO model with multiple additional years of data will be necessary to increase robustness, and future experiments could also focus on identifying whether individual regions show differences in prediction accuracy.

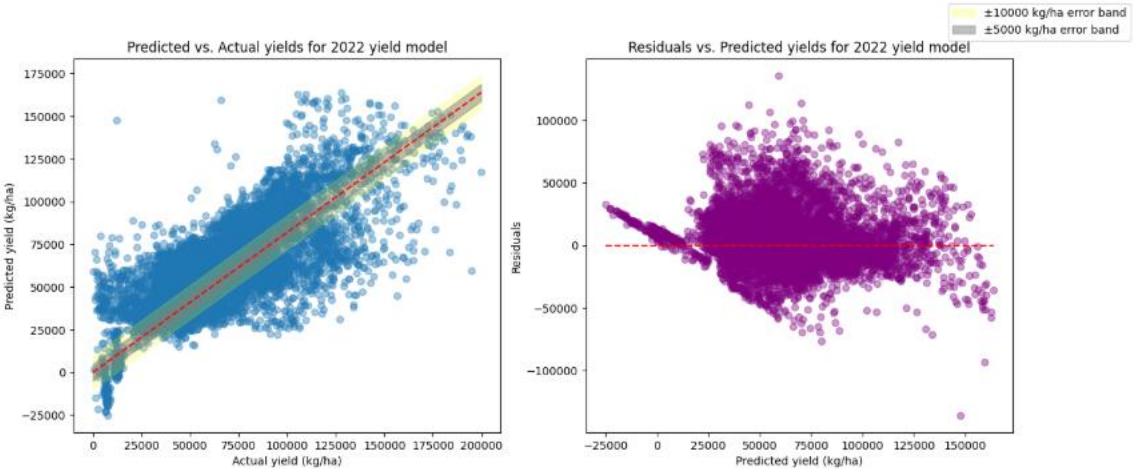


Figure 57: Distribution of actual vs. predicted yields (left) and residuals vs. predicted yields (right) for independent test dataset for best PRESTO model (2022 model with original yields). Error bands represent predictions within 5000 kg/ha (grey) and 10000 kg/ha (yellow) of actual yields.

The model inference workflow was also updated and tested, despite the lower R^2 values of trained models. Summarizing yield values for 20x20m pixels, the predicted yield histograms show similar patterns to actual yield histograms, although the peaks are generally lower and the models have difficulty predicting higher yield values. This is the case in multiple inference fields from the test dataset, and significant differences remain in the yield results. With a remaining high RMSE, there are still challenges to investigate further.

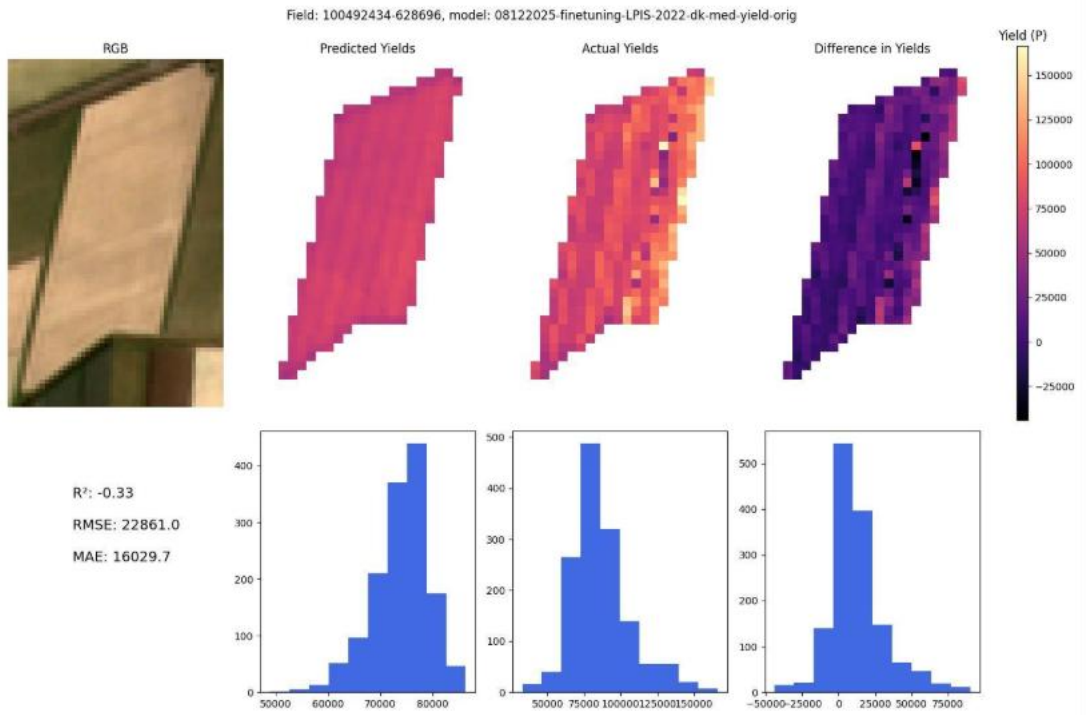


Figure 58: Example inference results for best model (2022 original yields) showing 20x20m yield predictions, actual yields, and yield differences.

Limitations and Future Work

Initial models show ability towards identifying and differentiating different levels of yield based on input sensor data, though the applicability of the model is limited due to high levels of uncertainty. Future model iterations will need to account for this, so that yield estimates can be provided, as well as confidence in these estimates on a pixel, field, and regional basis. Using PRESTO feature extraction improves the results compared to the raw satellite inputs. Such methods should also be applied to other sensor data to allow for further model improvements. These findings were the results previously reported as part of D4.3.

Further work will also focus on testing a multiclass approach rather than using PRESTO as a regression model. From the results obtained during this iteration, it is clear that the functionality of PRESTO as a regression-based model for estimating potato yields is not yet promising. Therefore, while future work will continue to analyse input data and retrain the model with new data sources, we will also try to bin input data in order to attempt a classification-based approach to yield modelling. This is on the basis of previous success with PRESTO as a classification model.

Initial assessment of the model and data distribution also showed that there is a strong spatial autocorrelation of data. Typically, pixels within a field show similar values, as do fields within a region (shown in the graph below, where histograms typically have different peaks, as do median values, which are similar in a local area). Future iterations and up-sampled predictions should account for this, using the location of a sample to improve estimation. The below image displays the histograms of median yield values per field for each region across the dataset, demonstrating that nearby sites typically have similar values. Such information is useful to both improve pixel-level estimation, as well as stabilize model expectations when predicting at field and regional level.

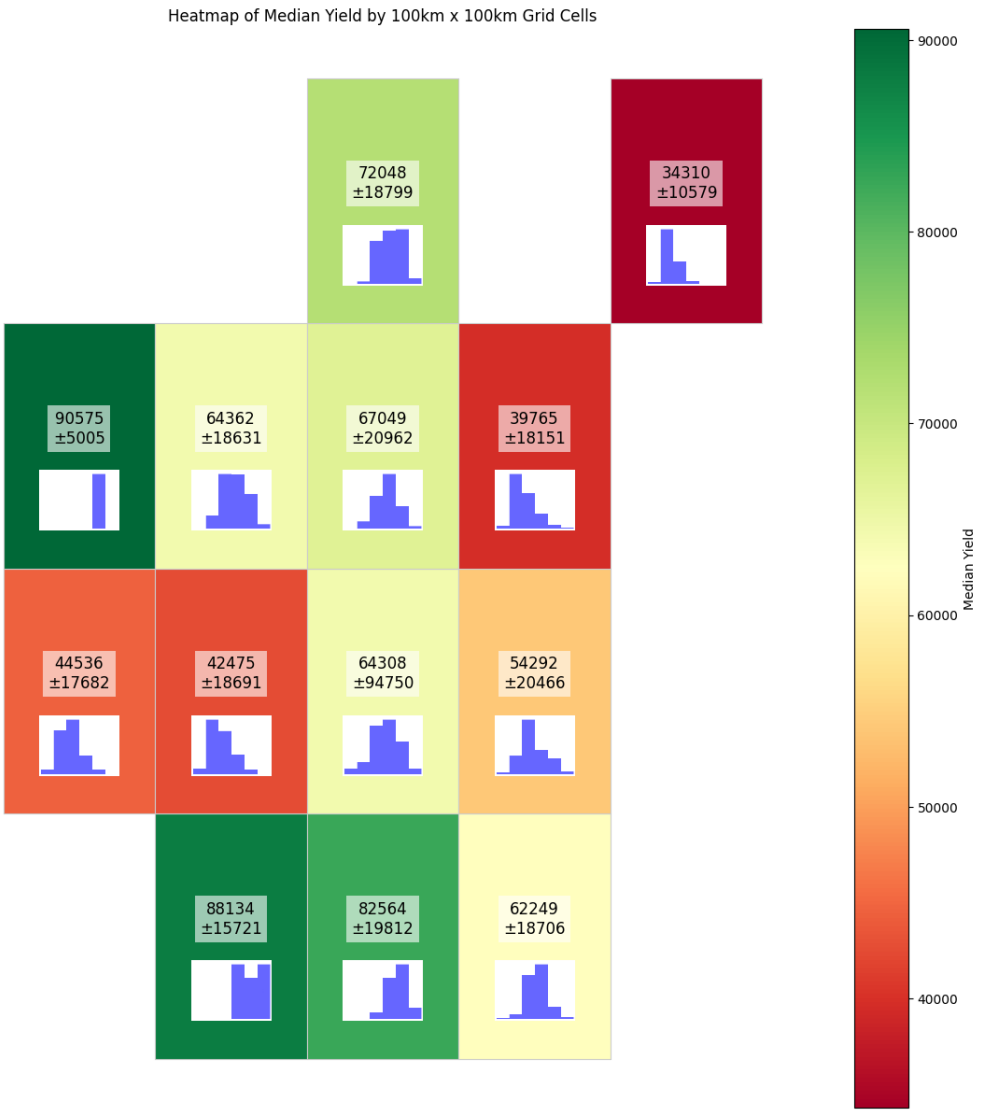


Figure 59: Histograms of median yield values per field for each region across the dataset

Sensor-integrated data products:

Table 35: Overview of sensor integrated data products (yield)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Potato yield (ton/ha) at subfield level	Belgium, Netherlands	2022, 2024	once, at harvest	20m	N/A*

**Confidence metric required to provide users with accurate estimations*

Use case(s):

Subfield (and future pixel)-level yield estimates show variations of crop yield within a field. After diagnosing the causes of the observed yield variations, farmers (whether or not in collaboration with advisors, researchers…) could use the maps to optimize their field practices, e.g., by applying variable rate fertilization.

The maps could also be useful for insurance companies, for damage assessment, or for assessing historical performance of their customers' fields. For this reason, ScaleAgData partner AgroInsurance will also test and evaluate the yield maps.
Further, the maps could also be of interest to input suppliers, such as seed companies, to select homogeneous fields for trials or seed multiplication.

3.3.2 Improved tare yield estimates for potatoes

This product, developed by UGent in collaboration with AVR, provides a pixel-level tare weight estimation based on RGB images. Currently, the tare weight of potato harvests—i.e., soil and dirt attached to tubers and the harvester conveyor belt—is assessed visually, without data-driven input. We hypothesized that tare weight depends on soil type and moisture. The RGB based tare estimate will be used to correct the potato harvester yield data and provide more accurate yield estimates to the farmers. It integrates RGB images captured from the harvester with soil weight and moisture from collected samples and with EO based soil moisture. The final goal is to apply the developed model to correct harvester yield data for tare on other fields in the region.

In autumn 2024, an experiment was conducted on two fields with contrasting soil types (sand-clay and sand). Samples were collected from the harvester to link measured tare weight with estimates derived from RGB images captured by a camera mounted above the conveyer belt on the harvester. Image segmentation (identifying potato tubers and soil clumps) was performed using AVR software. Our measurements aimed to determine whether a reliable relationship exists between identified soil clumps and measured tare weight, and, if so, to enable calibration of a tare weight estimation model. Data from one field (sandy soil) was less reliable because the camera—and thus the potatoes and clumps in view—was inaccessible during harvest. We also examined whether distinguishing soil clumps from soil attached to tubers improves tare weight prediction. Additionally, soil samples were collected to assess the relationship between soil moisture and measured tare weight. Measured soil moisture was compared with alternative estimation methods to evaluate less labor-intensive approaches.

In autumn 2025, four additional fields were sampled on dry loam soils. Data processing will occur next year, providing calibration data for loam soils. This calibration will be compared with that for sand-loam soils to determine whether soil-specific calibration is required or a universal calibration suffices. Furthermore, data from one field will be reserved for validation purposes.

Table 36: Overview of the fields under investigation

Field	Name	Location	Harvest day/ sampling	Soil type	#potato samples	#soil samples	Remarks
1	Zilverleen	Alveringem	10/9/2024	Dry sandy loam	13	51	No raw images
2	Jacob Vandenborne	Retie	16/10/2024	Dry sand	10	33	No match RGB - samples
3	Bartholomé Moyens	Landen	10/10/2025	Dry loam	13	/	
4	Van De Wever	Landen	14/10/2025	Dry loam	13	/	

5	Schuppen Spoorweg	Landen	15/10/2025	Dry loam	13	/	
6	Hombroux Robijns	Landen	20/10/2025	Dry loam	10	/	

Sensor input data:

RGB camera installed in AVR's Puma 4.0 harvesters.

Table 37: Overview of in situ input data (terra estimate)

Sensor data	Data provider	AOI/test sites	Nr. Fields/ season	Meas. frequency	Season(s)	Nr. training data
RGB camera	AVR	Belgium (Alveringem, Retie)	2	Harvest day	2024	10.000 photos
		Belgium (Landen)	4	Harvest day	2025	4.000 photos
Soil scan	EV ILVO	Belgium (Alveringem)	1	Harvest day	2024	Half field

EO input data:

Table 38: Overview of EO input data (tare estimate)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
NDWI	Copernicus	Sentinel Tile 31UFS	2024	On 11/9/2024 and 8/10/2024	10m x 10m
Soil moisture	DHI	Sentinel Tile 31UFS	2024	Harvest day + days after	15m x 25m (subfield aggregated)

Table 39: Other data used (tare estimate)

Data		Data provider	AOI/ test sites	Time period	Temporal frequency	Spatial resolution
In-situ measurements	Soil moisture	UGent	Belgium, (Alveringem, Retie)	2024	Harvest day	30-50 per field
In-situ measurements	Weight of potato, clumps, plants, rocks	UGent	Belgium, (Alveringem, Retie)	2024	Harvest day	10-13 per field
			Belgium (Landen)	2025	Harvest day	10-13 per field

Methodology and results:

The 2024 data was used to test the methodology and to see what measurements are necessary for a proper calibration of the tare weight estimation model. Part of the data was processed within the context of a bachelor's thesis at Ghent University. The results are presented below. The 2025 data still need to be processed but will be used to further improve the model and provide validation data for one soil type.

Figure 60 and Figure 61 show the soil moisture estimates on fields 1 and 2, respectively. The correlation between methods is low. NDWI showed a significant positive correlation with the measured gravimetric soil moisture ($r=0.32$, $p<0.05$) in field 2, but this correlation was absent in field 1. EO and sensor data do not provide an adequate proxy for soil moisture. However, the correlation between measured gravimetric soil moisture and measured tare in the samples was 0.15 for field 1 and 0.16 for field 2 (both non-significant). These results indicate that it is not necessary to include soil moisture estimates when estimating tare weight.

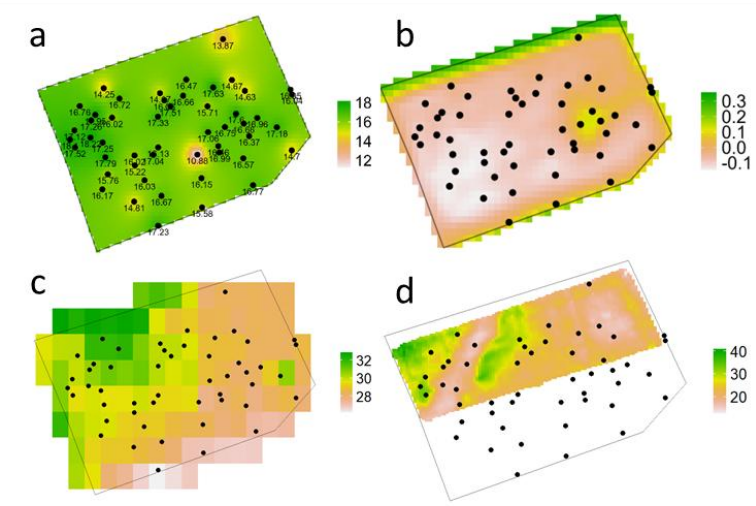


Figure 60: Soil moisture in field 1 with different methods: (a) gravimetric water content from in situ soil samples, (b) NDWI from Sentinel-2 data, (c) simulated volumetric water content provided by DHI, and (d) electrical conductance (proxy for soil moisture) from a soil scan with a CMD Mini explorer 6L (GF Instruments, Brno, Czech Republic).

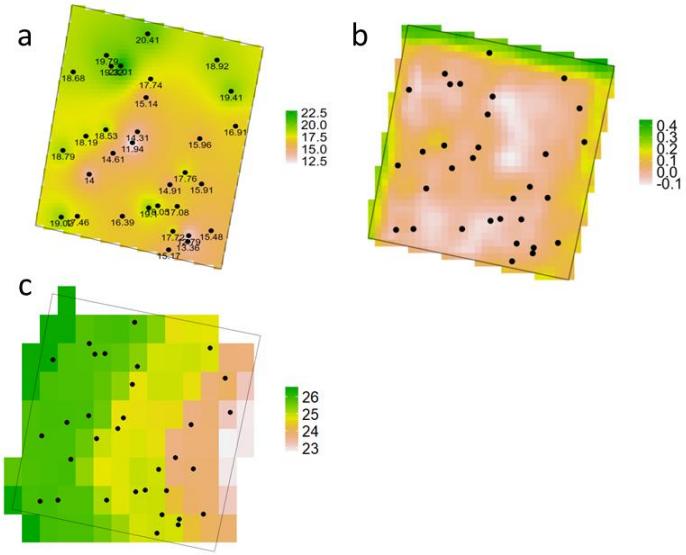


Figure 61: Soil moisture in field 2 with different methods: (a) gravimetric water content from in situ soil samples, (b) NDWI from Sentinel-2 data, and (c) simulated volumetric water content provided by DHI.

Four RGB images per field were used for manual annotation (Figure 62). Three different classes were distinguished: “potato”, “soil on potato” and “soil clumps”. For field 1, all 13 RGB images were used to manually annotate the “soil clump” pixels. This manual annotation had a dual purpose.

Their first purpose was to test the correlation between the real amount of “soil clump” pixels and the measured tare weight of the clumps to see whether RGB image processing contains the information to accurately estimate tare weight. This was also tested for the measured weight of the “soil on potatoes”. For this purpose, only the images of field 1 were used as the images of field 2 did not match the samples (due to inaccessibility of the conveyer belt underneath the camera during harvest). The results are shown in Figure 63 for the “soil clumps” and Figure 64 for the “soil on potatoes”. The regression shows that tare weight can be estimated from image analysis, with a reliability of 95% for the clumps and 93% for the soil on the potatoes. The relation between the pixels and the actual weight differs significantly between the two categories. The soil on the potatoes amounts to 2.6 - 10.8% (95% confidence interval, n=13) of the total tare weight. Most of the inaccuracies of RGB-estimated tare weight will thus come from inaccuracies in identifying “soil clump pixels”.

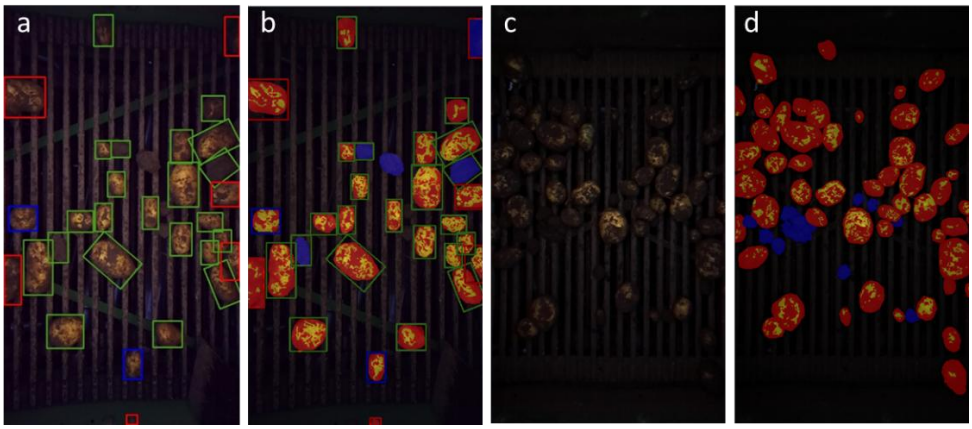


Figure 62: Manual annotation of different classes in the RGB image of field 1 (a,b) and field 2 (c,d). (a, c): raw data, (b,d) manual annotation of potato pixels (yellow), soil on potato pixels (red) and soil clumb pixels (blue). For field 1, only segmented images were available where potatoes were already recognised and marked with a frame.

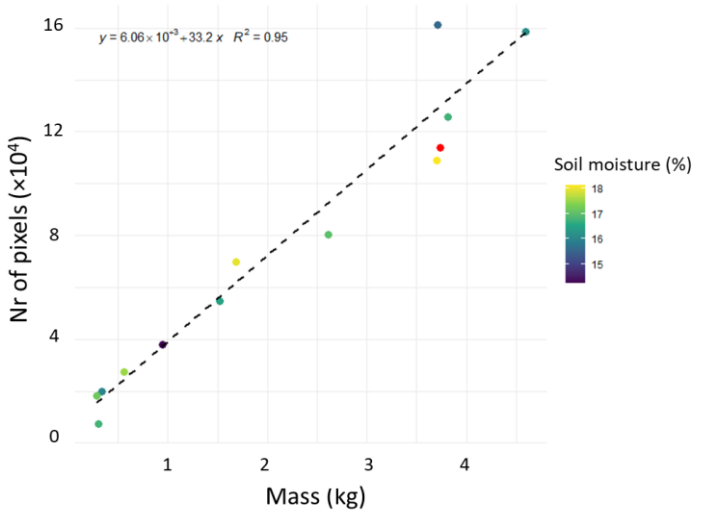


Figure 63: Correlation between mass of soil clumps in field 1 versus the manually annotated pixels of the respective RGB images.

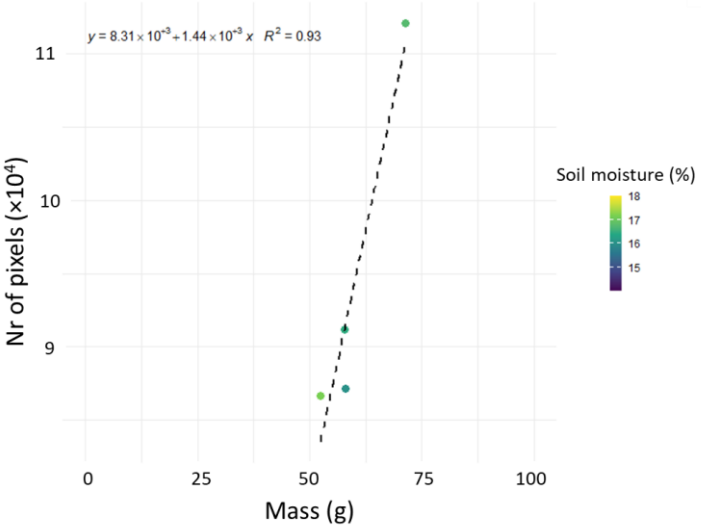


Figure 64: Correlation between mass of the soil on potatoes in field 1 versus the manually annotated pixels of the respective RGB images.

The second purpose of the manually annotated RGB images (four per 2024 field) was to use them as 'ground truth' data for calibrating image processing models. These models were developed to distinguish "potato", "soil on potatoes" and "soil clump" pixels. A first segmentation of the images was performed by an AVR software to eliminate the background. For field 1, pixels within identified rectangles containing potatoes or clumps were analysed. For field 2, pixels within identified potatoes and clumps (already correct shape) were analysed. The processed images were then used to test two models:

Model 1: Classification based on average calibrated RGB values. Classes are: "potato", "soil on potato", "soil clump" and "background". Average RGB values for each class were determined based on the eight manually annotated images. For new images, the pixels within the rectangle-annotated images of field 1, or the potato-and-clump-annotated images were processed. Each pixel was classified according to the shortest Euclidian distance of the separate R, G, and B values to the average RGB

values of the classes. While this method is more precise, its accuracy is strongly affected by image quality and lighting. Field 2 had much darker images (see Figure 62 a, c) compared to Field 1, which affected the average RGB value for each class. A field-specific calibration is required in this case where a few photos are manually annotated to get an average RGB value for each class.

Model 2: Classification based on light/dark differentiation. This method is based on the theory that potato pixels are lighter in color than soil pixels. Pixels lighter (sum of R+G value) than the average value (with a margin) were identified as "potato". This model was mainly useful to identify the soil on the potatoes.

Due to the lack of AVR clump-annotated images for field 1, both methods could only be tested on the "soil on potatoes" for field 1 (Figure 65). This figure shows that there is still a large uncertainty in the model prediction of the tare weight (R^2=0.4) compared to the manual annotation (R^2=0.93). This shows that the model needs further optimization in the next iteration.

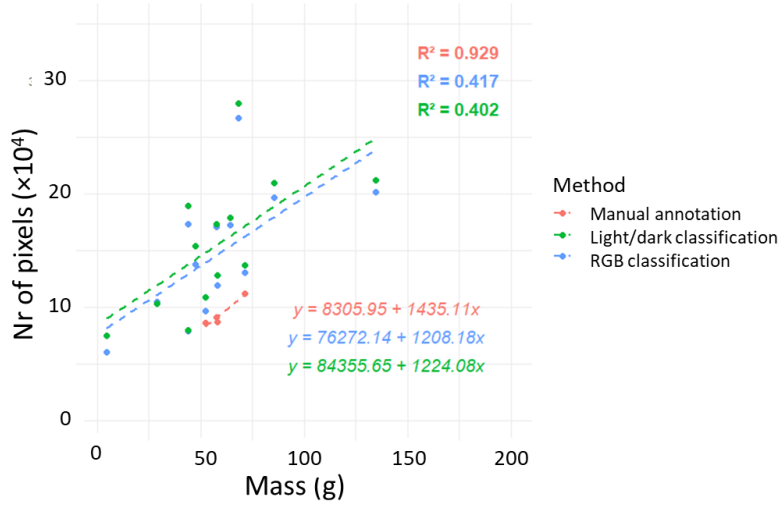


Figure 65: Correlation between mass of the soil on potatoes in field 1 versus classified pixels according to different methods.

Table 40: Overview of sensor integrated products

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Tare estimate based on RGB camera	Belgium	Summer 2024	Once, at harvest	Sub-m² level, dependent on frequency of RGB imaging on field	40%

The experiments conducted in 2024 demonstrate that RGB image analysis can reliably estimate tare weight in potato harvests, with high accuracy when soil clumps and soil on tubers are correctly identified. Manual annotation confirmed strong correlations between pixel classification and measured tare weight (up to 95% reliability), indicating that image-based methods contain sufficient information for calibration of a tare weight estimation model. However, variability in lighting and image quality across fields, and inaccuracies in pixel classification highlight the need for field-specific calibration or improved segmentation/classification algorithms. Soil moisture, initially hypothesized as a key factor, showed no significant correlation with tare weight, suggesting it is unnecessary for tare estimation models. The additional data collected in 2025 will enable calibration for loam soils and

provide validation datasets, supporting the development of a more robust, scalable model for automated tare weight estimation across different soil types.

Use case(s):

If a more accurate model can be developed, this can be used to correct the harvester yield data and provide more accurate yield estimates to the farmers.
The spatial resolution of the RGB based tare estimate is extremely high. Images could be processed at sub-second level during harvest, if so chosen. This would reduce the spatial resolution of the tare estimate to below 1 m^2, although it is hard to pinpoint the exact location of where an RGB image matches in the field, as the potatoes travel over conveyer belts, and not only belt speed can be adapted, but also the harvester speed influences the location match of the RGB image to the field.

3.3.3 Winter wheat yield estimates (LUKE)

This data product, developed by LUKE, in collaboration with UGent and CNHi, focuses on simulating spatial variation in yield estimates of winter wheat to fill gaps in harvester yield maps. Simulation utilizes the in situ data on the soil and crop status, farm management information, weather information, and satellite data.

Yield sensors that are installed on harvesters are often prone to inaccuracies resulting in incomplete yield maps. The calibration of the sensors is not always very accurate, making it difficult to compare data collected from different sensors. When combining data from multiple harvesters/sensors operating on a single field, this results in erroneous yield maps.

LUKE is developing a digital twin in collaboration with UGent, CNH and VITO to model the behavior of winter wheat fields. The aim of using a Digital Twin model is to simulate wheat yield and protein content variation within a field (in a hexagonal grid) based on the spatial variation in EO data. Missing data in yield maps provided by the harvester sensors can then be estimated based on the yield variation estimates from the digital twin model. Modeled yield forecasts, as well as crop and nutrient status during the growing season, can support variable rate fertilization applications.
More detailed information can be found in deliverables D4.2 and D4.4.

During the first iteration round of the ScaleAgData project, the development was targeted at setting up the digital twin modelling framework (see Chapter 2). Aim is to automate digital twin set-up as far as possible for the selected use cases. During the time when data from the lab was not available, development and testing utilized LUKE's test field data collected in Finland. During the second iteration the focus has been to develop the model based on data from Belgium from 2024.

A data assimilation endpoint for the Digital Twin service called CMAAS was developed for forecasting wheat yield and yield variability in a field by assimilating remotely sensed LAI data product from Terrascope to a simple wheat crop growth model was developed. The relationship between LAI, biomass in the canopy during growing season and yield was validated with in-situ canopy samples collected by UGent from 4 wheat fields in Belgium. Validation results are reported in D4.4.

Sensor input data:

Data for the two seasons were used to set up the digital twin. A limited dataset of eight fields in Flanders (region around Nijvel/Nivelle, see map below) monitored in 2023 will be used to test the model first. An extensive dataset of four fields (same region, see map below) monitored in spring/summer 2024 is being gathered to calibrate the model for different winter wheat varieties. More repetitions of sensor data are planned for this dataset.

General data sources including sensor input data are also listed in deliverable D4.2 section 2.3.

Table 41: Overview of in situ input data (yield)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
Yield		CNHi	Belgium	8 fields	once	2023
Yield		CNHi	Belgium	4 fields	once	2024
Yield		CNHi	Belgium	3 fields	once	2025
VI	Augmenta	CNHi	Belgium	8 fields	once	2023
VI	Augmenta	CNHi	Belgium	4 fields	3 times during season	2024
VI	(drone)	UGent	Belgium	4 fields	Twice during season?	2024
Soil EC		CNHi	Belgium	8 fields	once	2023
Soil EC		CNHi	Belgium	4 fields	3 times during season	2024

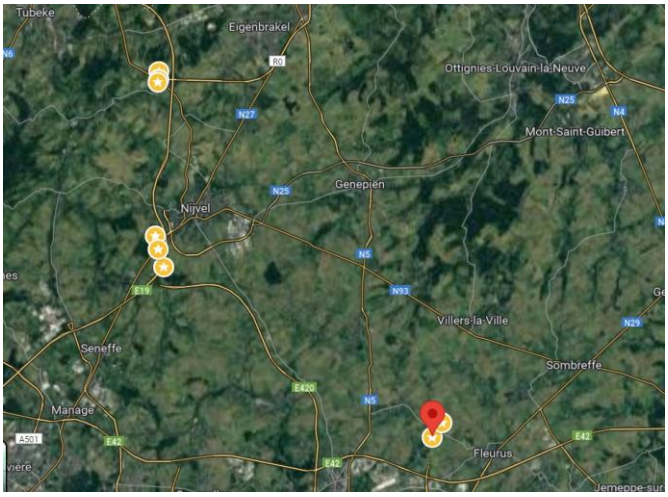


Figure 66: Fields monitored in 2023

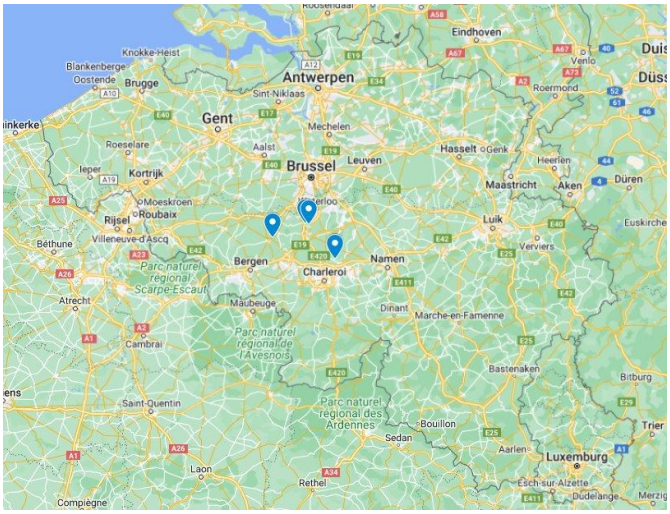


Figure 67: Fields monitored in 2024

Below are listed the EO data that was gathered for the 2023 season. The same dataset will be gathered for the 2024 season after the harvest.

General data sources including EO input data are also listed in the deliverable D4.2 section 2.3.

Table 42: Overview of EO input data (yield)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
CropSAR index	VITO	Belgium (8 fields)	Jan 1 – Aug 14, 2023	Every 5 days	10 m x 10 m
Evapotranspiration	DHI	Sentinel Tile 31UFS	Sep 2022 - Sep 2023	Every day	15 m x 25 m
Soil moisture	DHI	Sentinel Tile 31UFS	Sep 2022 - Sep 2023	Every day	15 m x 25 m

Additional data is being gathered to improve the model. Weather station for the 2023 season is already available. The same data will be gathered for the 2024 season after harvest. Other data for 2023 and 2024 (that is already available) are also listed below.

Table 43: Overview of additional input data (yield)

Data		Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
Weather station	Min T, max T, average T, RH, Wind speed, Precipiation, Radiation	KMI	Belgium	8 fields	daily	2023, 2024, 2025
Management data	Cultivar, sowing date, harvest date	CNHi	Belgium	8 fields		2023
Management data	Cultivar, sowing date, fertilisation timing and amount	CNHi	Belgium	4 fields		2024
In-situ measurements	Soil moisture, soil nitrogen, plant length, plant fresh weight, plant dry weight, plant nitrogen, leaf chlorophyll, anthocyanin, flavonoid content and nitrogen balance index, LAI	UGent	Belgium	4 fields, 6 subfields	4 times during season	2024
In-situ measurements	plant length, plant fresh weight, plant dry weight, grain weight, LAI	UGent	Belgium	1 field, 8 subfields	Once, just before harvest	2025

Methodology and validation results:

To simulate daily status of a field, yield estimates and within-field variation, the sensor and EO input data are entered into a process-based model and assimilated with remote sensing data. Data assimilation described in section 2.2 updates observed EO input data into models. Flexible use of various modelling approaches and easy change of the simulated field is enabled with digital twin data model. See deliverable D4.2 section 2 for further description of the methodology.

The data assimilation methodology developed for the CMAAS model (section 2.2.2) was developed and validated using the yield lab data. Example inputs and outputs for data assimilation from one field parcel are shown in Figure 68, each black square represents LAI observations from distinct spatial area, and each coloured series (dots) represents an assimilated time series forecast for the canopy of that same region. The spatial representation of the same date forecasted up to harvest is shown in Figure 69. The method has been developed by LUKE using wheat cultivation data from Belgium from the yield lab collected by UGent and CNH during 2024. The results from the development were promising with up to 0.77 correlation between the forecasted crop yield using data assimilation and the harvested crop yield. However, the data from 2024 was also used to calibrate the crop model N response and phenological parameters and the results is therefore not completely independent. The yield lab data from 2025 will be used to further validate the method.

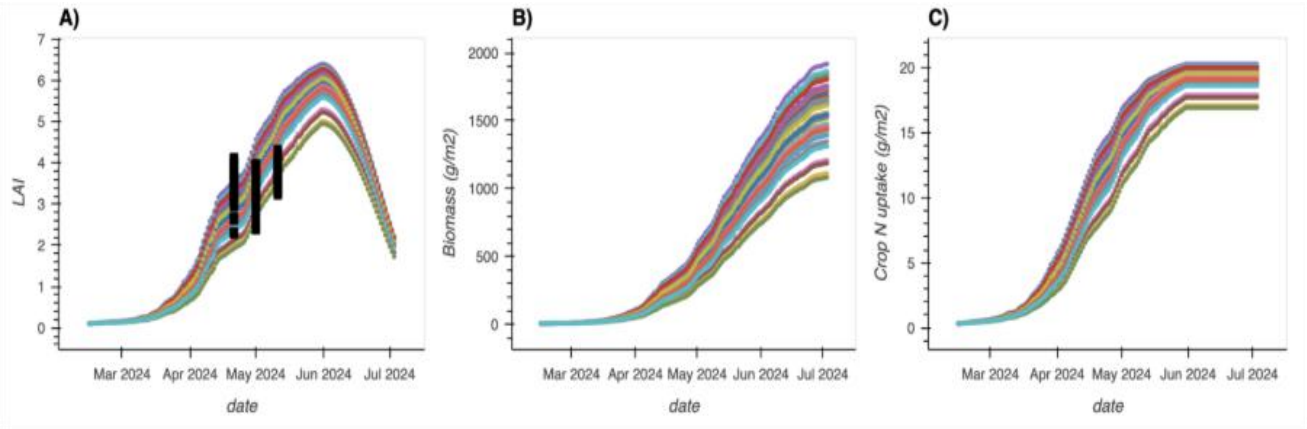


Figure 68 Example of data assimilation results returned by CMAAS. A) Shows the forecasted LAI development (coloured dots) over time and observed LAI from Sentinel 2 (black squares) B) Biomass development forecast C) Crop N uptake forecast.

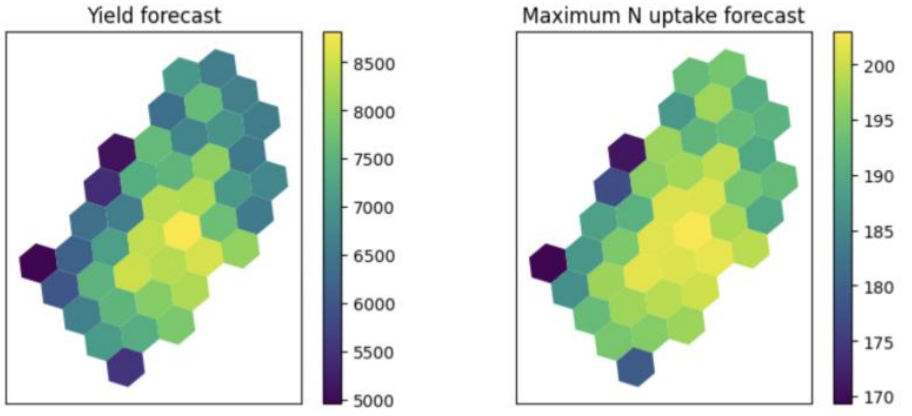


Figure 69. Map representation of data assimilation forecasts. The colour scale on the left represents crop yield (kg/ha) and on the right crop N uptake (kg/ha).

Table 44: Overview of sensor-integrated data products (yield)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Winter wheat yield estimates (field level)	Belgium (8 fields) and Finland (22)	2017 - 2023	Forecast 1 / day	Field level	N/A yet
Winter wheat yield variability maps (pixel-based)	Belgium (8 fields) and Finland (22)	2017 - 2023	Forecast 1 / day	0.01ha to 0.2 ha	N/A yet

Use case(s):

The spatial variations in wheat yield estimated with the digital twin model will be used to correct the harvester yield data and provide more accurate yield estimates to the farmers.

3.3.4 Winter wheat yield estimates (VITO)

Once the yield model framework that is described in section 3.3.1 has been validated for potatoes, it will also be used / finetuned for winter wheat yield estimation. This product, developed by VITO in collaboration with CNH, will include winter wheat yield estimates at subfield-level (10m resolution) at harvest time,showing yield variability within the field. It integrates yield sensor data collected using CNH's wheat harvesters with EO (Sentinel-1/2) and meteo data. It therefore makes use of the PRESTO framework for few-shot learning. The final goal is to apply the yield estimation model to other winter wheat fields in the region (without yield sensors) and aggregate the results into regional yield statistics. Wheat harvester data are provided by CNH.

3.4 RIL Soil Health

3.4.1 EO based regional soil organic carbon map

This product, developed by AUTh and EV ILVO in collaboration with DEIMOS, includes separate pixel-based maps of Flanders and Central Macedonia with a resolution of 10 m. The pixels contain estimations of the SOC in the topsoil at field level and regional level. The SOC map is generated using a SOC prediction DL model that was trained using a federated learning (FL) approach while preserving data privacy for regional datasets. The maps are generated using in-situ soil sample data as training data and processed EO image data as feature inputs for the predictive SOC ML model.

Table 45: Overview of EO input data (soil organic carbon map)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel 2	Google Earth Engine	Flanders	2017-2024	5-daily	10 m to 60 m
Sentinel 2	Copernicus Data Space Ecosystem	Region of Central Macedonia	2017-2024	Every 5 days	10 m to 60 m

Other input data:

Next to EO data, regional and European soil sample data containing lab analyzed SOC values will be used to train DL models (Table 46). For Flanders, a soil campaign performed during 2021 and 2022 will be used, which has provided 210 sample points for topsoil SOC throughout Flanders (Figure 70).For Central Macedonia 2208 sample points are available for model development (Figure 71).

In addition, a European model will be trained as a starting point to further finetune using a federated learning approach. The LUCAS 2018 topsoil dataset (https://esdac.jrc.ec.europa.eu/content/lucas-2018-topsoil-data) containing 7418 SOC values for the whole of Europe was used as training set for the European model.

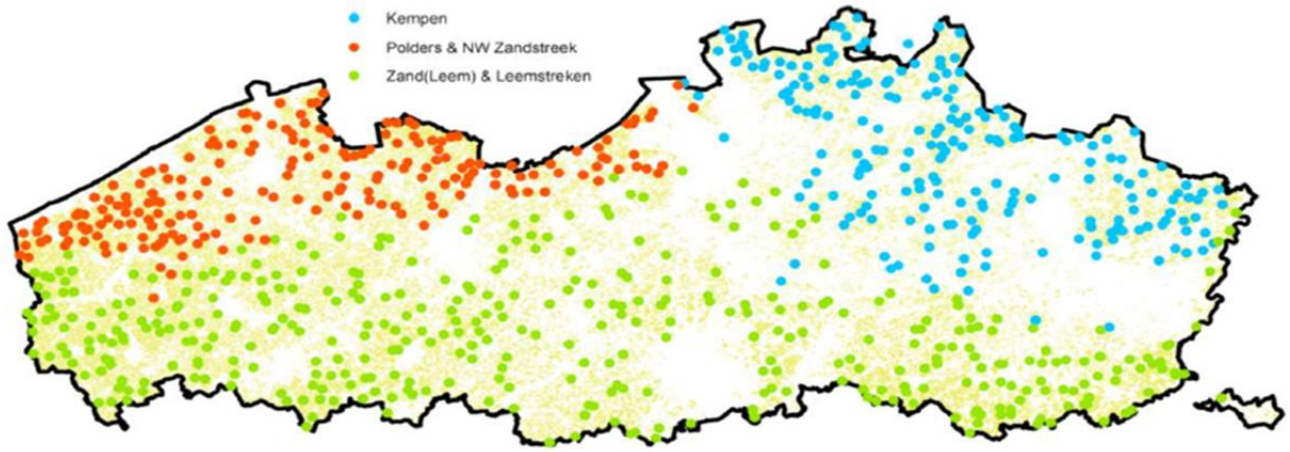


Figure 70: 210 sample points for topsoil SOC throughout Flanders, taken in 2021 and 2022.

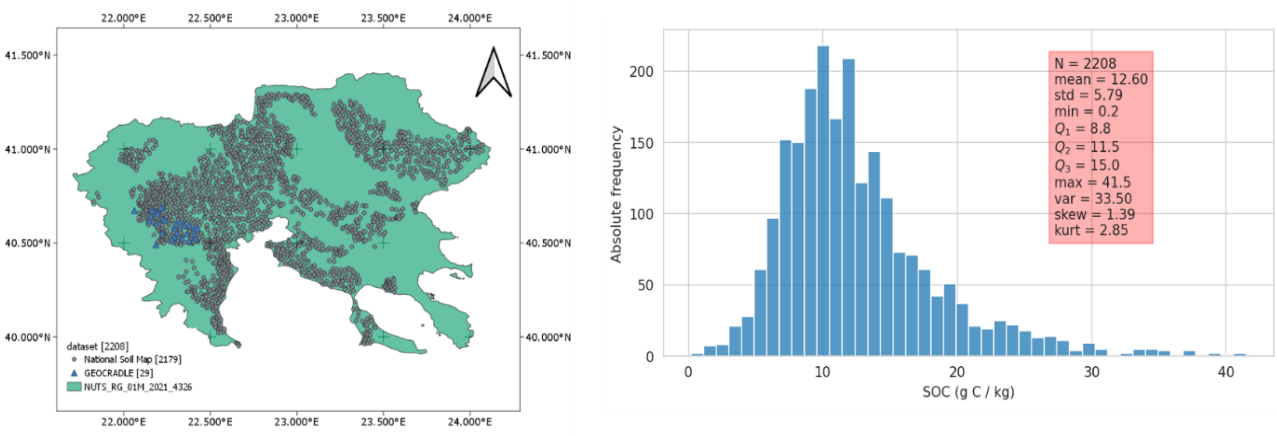


Figure 71: 2208 sample points for topsoil SOC throughout the region of Central Macedonia, and their corresponding topsoil SOC content in the form of a histogram; samples collected between 2014 and 2024.

Table 46. Descriptive statistics for topsoil SOC (g kg−1) concentrations in the mineral croplands of the datasets considered.

Dataset	Acronym	N	Mean	Min	Q1	Q2	Q3	Max	Std	Skew
LUCAS 2018	EUR	7418	17.5	0.0	10.6	14.6	20.5	196.4	13.1	4.8
Flanders	FLA	211	13.7	2.9	10.3	12.8	15.2	36.8	4.9	2.1
Central Macedonia	RCM	2208	12.6	0.2	8.8	11.5	15.0	41.5	5.8	1.4

In addition, ancillary soil property data, bulk density and texture (i.e., percentage of coarse fragments and concentrations of sand, silt, and clay), derived from multiple maps developed by JRC in the ESDAC portal (https://esdac.jrc.ec.europa.eu/resource-type/soil-data-maps) are included as extra contextual data to train the SOC DL models.

Methodology and validation results:

The research develops a privacy-preserving SOC prediction model for croplands in Flanders (Belgium) and Central Macedonia (Greece) using FL. The methodology integrates Sentinel-2 satellite imagery with distributed soil datasets while maintaining data privacy through decentralized machine learning (Figure 72).

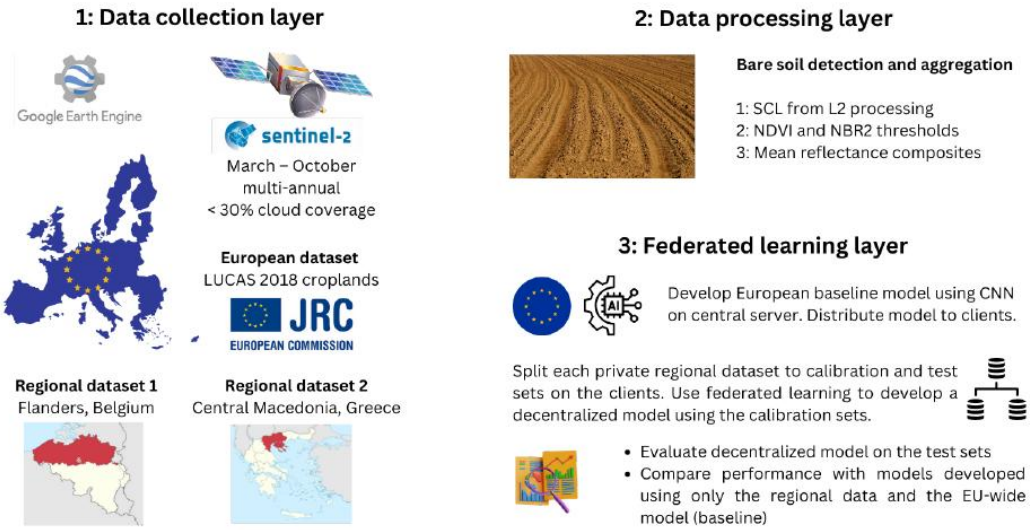


Figure 72. Overview of the federated learning approach used.

Data Processing and Model Architecture:

Sentinel-2 L2A imagery (2018-2020) is processed using Google Earth Engine with cloud coverage below 30%. Bare soil pixels are identified through Scene Classification Layer (SCL) filtering and spectral indices thresholding (NDVI: 0-0.25, NBR2: 0-0.075), creating mean reflectance composites. The approach addresses challenges with perennial crops in Central Macedonia, where approximately 50% of data points were excluded due to permanent vegetation cover. Additional ancillary data including bulk density and soil texture (sand, silt, clay percentages) from JRC-ESDAC are incorporated.

A custom Convolutional Neural Network (CNN) with 1D convolutional layers processes the spectral data through two channels (64 and 32 filters), concatenating deep features with ancillary information through fully connected layers (512, 256, 64 neurons) to predict SOC values (Figure 73).

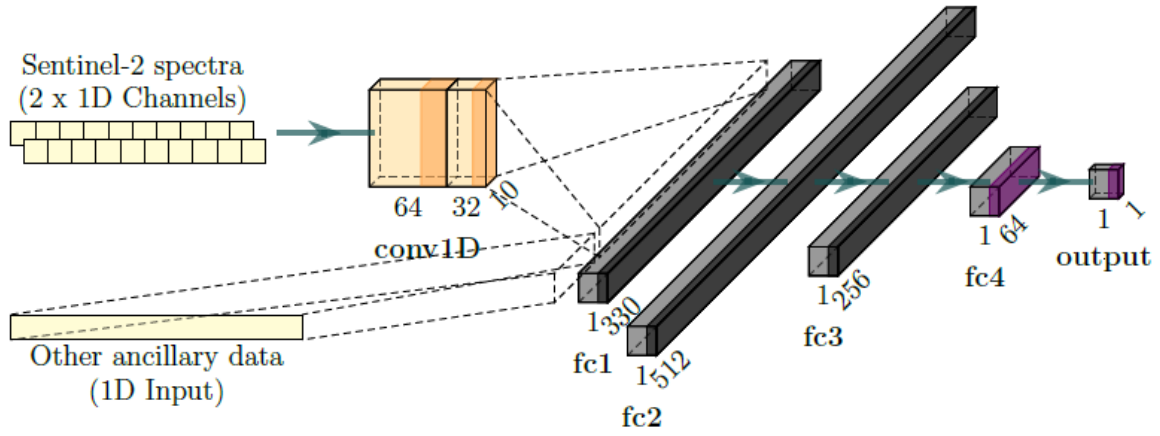


Figure 73. Topology of the multi-input multi-channel convolutional neural network employed in the present study.

Federated Learning Implementation: The FL framework developed in cooperation with DEIMOS based on the flower.ai framework was used here (see section 2.1). The FL framework operates through iterative decentralized training and centralized aggregation. Starting with a continental model trained on LUCAS 2018 data (7,438 points), the system distributes model weights to regional clients (Flanders: 211 samples, Central Macedonia: 2,208 samples). Each client trains locally for one epoch on their private dataset (split 2:5:3 for training, validation, testing), computing validation losses and returning only updated weights to the central server—raw data never leaves local institutions. The server aggregates updates using Federated Averaging, weighted by dataset sizes, producing an improved global model over 100 epochs using RMSprop optimization (learning rate: 0.001, batch size: 64).

Validation Results: The FL model demonstrates region-specific performance improvements compared to the European and regional baseline models (Table 47). The FL model performance was evaluated against multiple baseline models to assess its effectiveness in leveraging distributed datasets while preserving privacy. The baseline models included: (1) a continental model trained solely on LUCAS 2018 European data (EUR), (2) region-specific models trained independently on local datasets from Flanders (FLA) and Central Macedonia (RCM), and (3) the continental model applied directly to regional data without local fine-tuning. For Central Macedonia, the FL model achieved a 3.2% RMSE reduction (3.29 vs 3.40 g kg^{-1} baseline), while Flanders experienced a slight increase of 11.6% RMSE increase (4.71 vs 4.22 g kg^{-1} baseline). This disparity reflects the weighted averaging strategy prioritizing the larger Central Macedonia dataset. The first results using FL demonstrates are promising in using the FL approach to generating regional or even European SOC models using the many siloed soil datasets while guaranteeing data privacy for possible sensitive data content. The FL model successfully generated regional topsoil SOC maps at 10 m resolution for bare soil pixels across both regions (Figure 74).

Table 47. Model performance on the independent test set for predicting topsoil SOC (g kg−1), comparing federated learning with baseline models.

Models	Train on	Predict on	RMSE	MAE	RPIQ
Baseline	EUR	EUR	13.86	7.74	0.95
	EUR	FLA	17.10	8.74	0.30
	EUR	RCM	8.79	5.75	0.65
	FLA	FLA	4.22	3.10	1.21
	RCM	RCM	3.40	2.36	1.68
FL	all	FLA	4.71	3.59	1.08
	all	RCM	3.29	2.32	1.73

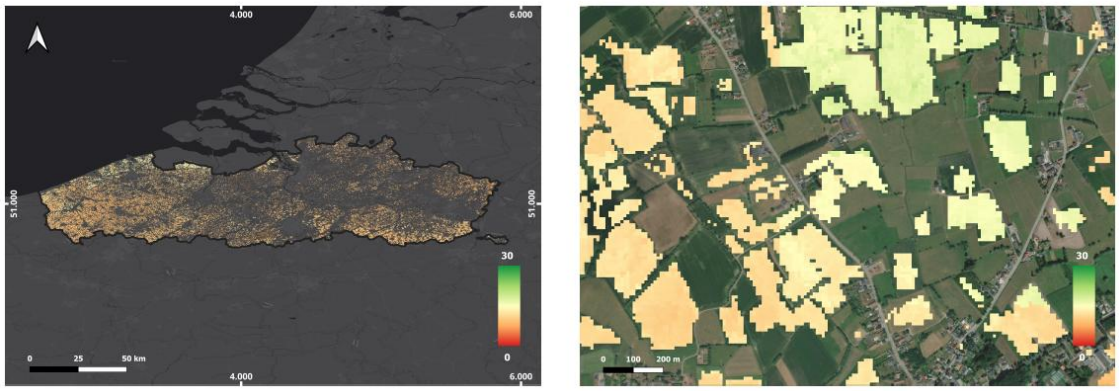


Figure 74. Spatial distribution of predicted topsoil SOC (g kg−1) in croplands in Flanders (left) and detailed view of fields showing intra-field variation in predicted SOC (right). Only bare soil pixels were predicted.

The Flemish SOC maps for 2023 and 2024 are currently provided as RESTful API service at the following endpoint https://scaleagdata-soc-api-1002910116761.europe-west1.run.app , documentation can be found at https://scaleagdata-soc-api-1002910116761.europe-west1.run.app/docs https://scaleagdata-soc-api-1002910116761.europe-west1.run.app. The maps, produced at 10m spatial resolution for bare soil pixels, are hosted on Google Cloud Platform infrastructure. Documentation can be found at https://scaleagdata-soc-api-1002910116761.europe-west1.run.app/docs providing details on available endpoints, request parameters, response formats, and example queries. Metadata for the models and SOC maps are equally provided in an interoperable way (using among others the standards Demeter AIM, DCAT, MLS, …) via one of the endpoints.

Table 48: Overview of generated data products (soil organic carbon map)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
SOC values at pixel level	Flanders, Central Macedonia	2023 and 2024	Annually	10 m	MAE between 2.3 and 3.6 (for values in g/kg)

The map can be assessed by external stakeholders, including farmers, policymakers, research institutes and government institutions and can be used for monitoring and advisory purposes.

3.4.2 Soil health indicator estimates (field level)

This product, developed by AUTh and EV ILVO in collaboration with VTT and Kuva Space, includes separate SOC predictive pixel-based maps of selected fields in Flanders and Central Macedonia with a resolution of < 1 m, developed from UAV data and EO Hyperspectral Imaging (HIS) data. The pixels contain estimations of the SOC in the topsoil. In addition, the soil RIL is combining in-situ VTT HSI data with Kuva Space & Sentinel 2 EO data in a data fusion approach to examine the added value in estimating SOC values at field level. The SOC maps and SOC DL models developed in D3.5 only use VTT HSI data, while in this section we aim to integrate EO data in a data fusion approach to improve predictive capabilities of the SOC DL models.

Table 49: Overview of sensor input data (soil health indicator)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)	Nr. training data
Hyperspectral images of reflectance	VTT camera mounted on UAV	VTT	Selected fields in Flanders, and region of Central Macedonia	Minimum of 2 fields / area / season	1 image / field / year	Growing seasons of 2024, 2025 and 2026	Minimum of 10 samples / field or about 10 samples per ha

The number of training data here refers to new soil samples collected from each field where the UAV flights took place. These were treated similarly as the field samples noted in the previous section (i.e., chemical analysis in the lab). The soil sampling was determined using the sensor spatial planning algorithm discussed in more detail in D3.5.

EO products were used to augment the spectral resolution of the provided VTT hyperspectral camera. Spaceborne sensors have a coarser spatial resolution but bolster a wider spectral range. This fusion helps us develop more robust soil maps.

Table 50: Overview of EO input data (soil health indicator)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel-2	Copernicus Data Space Ecosystem	Same fields as above	2024	1+ cloud-free image(s) / field / year as close as possible to the UAV data collection	10 m to 60 m
Hyperfield	Kuva Space	Same fields as above	2025-2026	1+ cloud-free image(s) / field in 2025 as close as possible to the UAV data collection	As provided by Kuva Space

Methodology and validation results:

We have generated the first version of a SOC DL model and predictive SOC field maps during the first iteration 2024-2025. These are described in D3.5 Sensor and edge processing selection, development, spatial planning and data collection. Examples of the Flemish field maps can be seen in Figure 75.

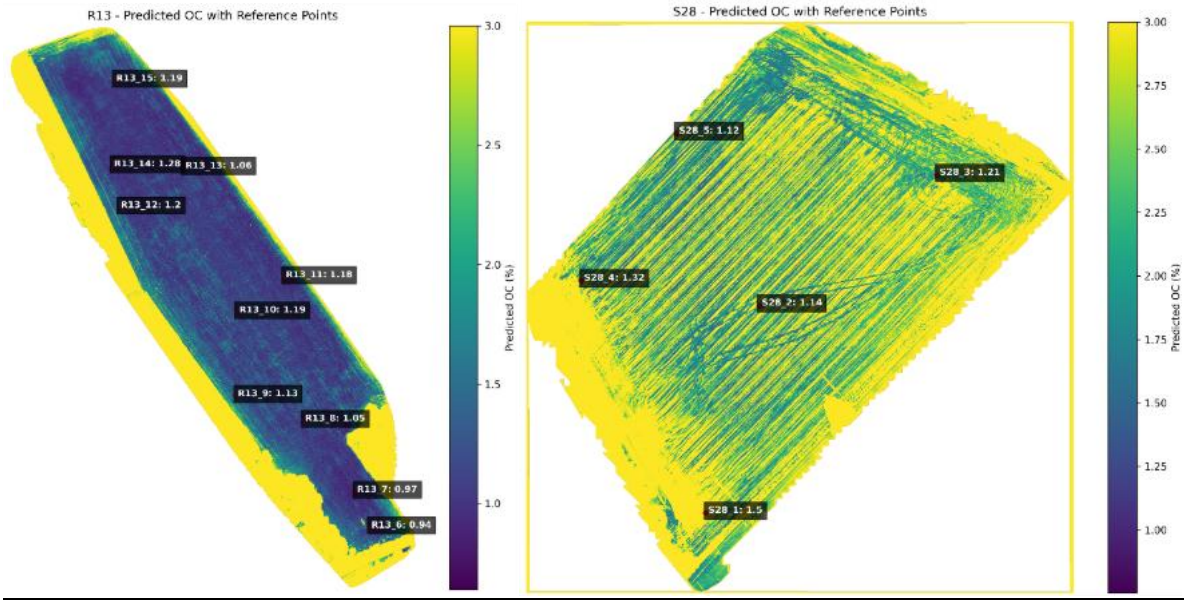


Figure 75. Predicted SOC values by trained DL SOC model for 2 EV EV ILVO fields tested during first soil campaign in 2024

In the first iteration we focused on setting up our experimental framework with the VTT HSI sensor and UAV. During the next iteration in 2025-2026 we plan to collect data for more fields (e.g. in Flanders 20 ha vs the 2 ha of the first iteration). The increase in data will improve the accuracy of our SOC DL models and SOC maps (the DL SOC model for EV ILVO had an MSE of 0.8954 and R^2 of -42.2772).

In WP4 we build improvements in these field level SOC DL models by setting up a data fusion experimental approach by combining the VTT HSI data with HSI EO data provided by Kuva Space. Kuva Space has successfully launched two Hyperfield-1 satellites and provided EV ILVO and AUTh with HSI EO data for the experimental areas via their online platform https://beta.kuvasense.com/login?next=/data-archive (further info can be found in D3.5). We plan to perform these research activities during the beginning of 2026. However, AUTh has already performed some promising experiments combining sentinel-2 and Kuva Space HSI EO data.

For this preliminary fusion test, AUTh used a Sentinel-2 bare-soil composite similar to the one employed in the federated learning component (section 3.4.1) along with the corresponding Kuva Space hyperspectral pixels for the same location. These were aligned to create a fused input combining S2's broad coverage with Kuva Space's finer spectral detail.

Then an ablation study was conducted to compare the effects of using Sentinel-2 alone versus the fused input in SOC estimation. A regression workflow with preprocessing, bootstrapping and k-fold cross-validation was used to compare the two inputs. The fused dataset produced better results, with higher average R^2 (0.40 vs 0.36) and slightly lower prediction error (MSE 0.21 vs 0.22). Although preliminary, these findings indicate that adding Kuva Spaces hyperspectral information systematically improves SOC estimation compared to relying on Sentinel-2 only.

Table 51: Overview of sensor-integrated data products (soil health indicator)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Fused hyperspectral image	Same fields as above	Dec 2024 – March 2025	1 image / field / year	<1 m	N/A

Final Field SOC maps	Same fields as above	Dec 2024 – March 2025	1 image / field / year	<1 m	Target: $R^2 > 0.5$

Final products can be tested by the agronomists and/or farmers who have a knowledge of the fields.

3.5 RIL Grassland

3.5.1 Gap-filled grasslands LAI maps at parcel level

This product, developed by EURAC, includes pixel-based (10m resolution), spatially gap-filled LAI time-series data to estimate grassland yield over the province of Bolzano, in the Italian Alps. The product combines optical Sentinel-2 and SAR Sentinel-1 data, developing and validating different machine learning algorithms (Gaussian Process, Random Forest, CatBoost, Transformer-based Convolutional Regressor) for the improvement of the biopars time-series over the Italian Alps. In situ LAI data collected within the project are used to validate the gap-filled LAI product derived from S1 and S2 data; biomass in situ data are used to assess the impact of using the gap-filled LAI product in the estimation of grassland productivity. It is planned to adapt the framework also to Spain in collaboration with IFAPA.

The gap-filled LAI time-series are validated using ground measurements collected by EURAC over eight grassland parcels throughout the Provinces of Trento and Bolzano. Their location is shown on the map below together with all permanent meadows in the regions that built the AOI. This shows the potential limitation of geographical distribution of the sensor data which might be clustered and is not covering the area equally. The sampling takes place bi-weekly from late April to the end of October of the years 2023 and 2024, spanning the growing period of the grassland before the first until after the latest mowing event of the season. At each field measurement event three replicates are collected per meadow, located in the center of a S2 10 m pixel. They consist of various single measures, spanning from the LAI over soil moisture to the vegetation height. The most important measures for the validation of the data fusion, namely the LAI and yield, can be found in the table below.

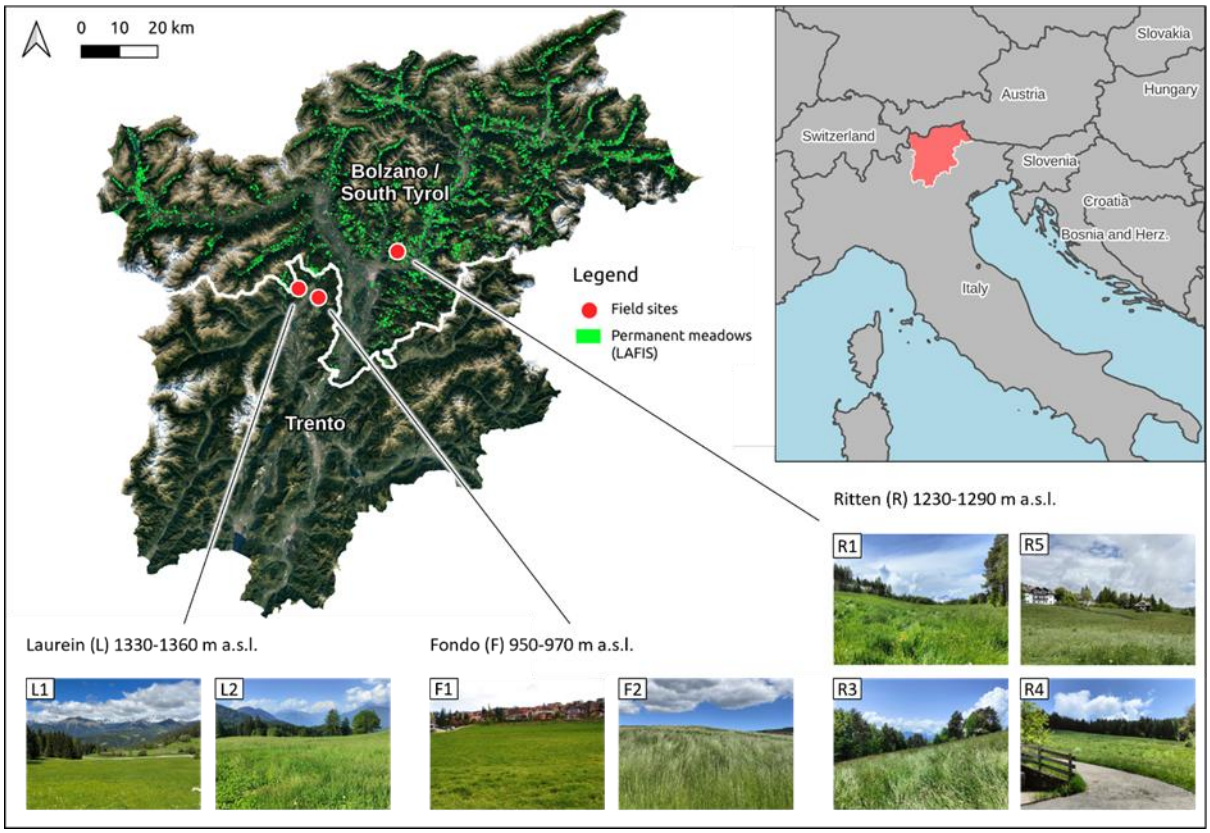


Figure 76: Location of the study area and the field sites. Imagery: Google, © TerraMetrics

Table 52: Overview of sensor input data (Gap-filled grasslands LAI maps at parcel level)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/ season	Meas. frequency	Season (s)	Nr. training data
Grassland LAI	LI-COR LAI-2200C Plant Canopy Analyzer	EURAC	Provinces of Trento and Bolzano, Italy	8	Bi-weekly	2023, 2024	+/- 250 per year
Grassland yield [t/ha]	Destructive samples						

Table 53: Overview of EO input data (Gap-filled grasslands LAI maps at parcel level)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel-2 products (LAI using the SNAP Biophysical Processor)	Copernicus data space ecosystem	Provinces of Trento and Bolzano, Italy	2019-2024	All available scenes (2-3 days)	10m
Sentinel-1 RTC	Microsoft Planetary Computer	Provinces of Trento and Bolzano, Italy	2019-2024	All available scenes (<3 days, since 2023 <6 days)	10m

Daily total precipitation and temperature	EURAC	Provinces of Trento and Bolzano, Italy	2019-2024	daily	250m
Digital elevation model (DEM, classification of the meadows into altitudinal belts)	European Environment Agency	Provinces of Trento and Bolzano, Italy	2019-2024	-	25m
Surface Soil Moisture based on a combination of SAR and optical imagery	EURAC	Provinces of Trento and Bolzano, Italy	2023	3-5 days	20m

Methodology and validation results:

S1 RTC data together with auxiliary data, namely the day of the year, daily total precipitation, daily mean temperature, and the altitudinal class derived from the DEM, are used in a data fusion approach for the spatial gap-filling of S2 LAI over Alpine grassland. For this purpose, different machine learning algorithms (Gaussian Process, Random Forest, CatBoost) are trained over the AOI for 2020-2022. A model selection with cross-validation during the model training ensures that the best-suited predictors amongst the S1 RTC and the auxiliary data are selected and that the hyperparameters are optimized. The validation is done using the in-situ LAI measurements over the eight grassland parcels and using S2 LAI of meadows that were not included in the model training. As the validation targets meadow both temporally and spatially outside of the training data, this can be seen as an independent test of the methodology.

The cross-validation during model training (2020-2022) states RF as best performing model (R^2=0.70, RMSE=0.97). However, all models perform poorly in the validation (on 2023) with CatBoost being the best model (R^2=0.16 and RMSE=1.34 against S2 LAI), and the models underestimate the in-situ LAI strongly. Additionally, the predicted LAI appears temporally and spatially smoothed, and mowing events cannot be identified accurately.

The selected features of all different models include auxiliary data which allows us to disclose that S1 is difficult to interpret for the data fusion approach for S2 LAI over Alpine grassland. Due to only one S1 sensor remaining in orbit, the availability of data is a limitation for this methodology. The strong reliance on auxiliary predictors causes smoothed predicted LAI. Thus, development of the methodology in the second project iteration will include a greater use of S1 data without or less focus on auxiliary predictors. Also, polarimetric decomposition could ease the estimation of LAI from S1. Due to the complex terrain and size of the study area, a more conservative selection of meadows and less generalized algorithms could improve the results significantly.

Building on the outcomes of the previous investigations, we developed a DL approach capable of capturing long-range temporal dependencies in S2–derived LAI time series. The method is based on a domain-specific transformer-convolutional architecture that ingests a set of complementary input features, including Sentinel-1 SAR data, EO-derived soil moisture, a Digital Elevation Model, and the Day of Year.

We selected 3431 parcels classified as permanent meadows and these polygons were stratified across eight altitudinal grassland types and used for training, validation, and testing. LAI was predicted on

the test set over the period from March 2023 to November 2023. From this dataset, 55% of the sample was used for training, 5% for validation, and 40% for testing. The effectiveness of the proposed method for predicting LAI has been validated by testing it against Sentinel-2 derived LAI. In the results, we achieved values of R2, MAE, RMSE respectively of 0.59, 0.63, 0.92 between the predicted LAI and Sentinel-2 LAI on the test sets. We compared the proposed domain-specific model with related DL-based approaches to better assess its feature extraction capabilities. Additionally, we also compared other traditional regressors with the proposed approach. Overall, we found that the proposed approach was more effective in comparison to other methods, especially in accurately predicting extreme values of LAI across the growing season. We analyzed the trend of both predicted and target Sentinel-2 LAI against in situ data collected during the sampling campaign. Specifically, the occurrence of mowing events between June 2023 and November 2023 was considered as one of the forage production field sites in the Province of Trento. From this comparison it emerged that the model is not yet able to predict abrupt changes in LAI time series resulting from mowing events. Further tests will be performed on the parcel scale to overcome the limitations of the current implementation and enhance the accuracy of the product.

During this iteration, we implemented the few-shot learning workflow using the dummy dataset provided by VITO, fine-tuned the model for yield mapping and conducted an initial evaluation. We also prepared point-based vector data for applying the workflow to our study case in collaboration with VITO. Importantly, the team is very motivated to further explore and apply the few-shot learning framework in the upcoming months, with the intention of integrating it into the broader modelling pipeline as soon as the relevant datasets and configurations are in place.

Table 54: Overview of sensor-integrated data products (Gap-filled grasslands LAI maps at parcel level)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Version 1. Spatially gap-filled LAI from S1-S2 data fusion	Permanent grasslands in the Province of Bolzano, Italy	2020-2023	1-4 days	10 m	R^2=0.70, RMSE=0.97 on validation set for the best performing model (RF); R^2=0.16 and RMSE=1.34 on test set for the best model (CatBoost)
Version 2. Temporal Gap-Filled LAI using S1 data	Permanent grasslands in the Province of Bolzano, Italy	2023	1-4 days	10 m	R^2= 0.59 MAE=0.63 RMSE=0.92 on the test set

Use case(s):

The gap-filled LAI maps will be used to generate an improved version of the GPI (Grassland Production Index) that EURAC developed for an index-based insurance for Alpine grasslands in the Province of Bolzano.

3.5.2 Estimated grassland yield at parcel level

This product, developed by IFAPA, includes a daily estimation of grasslands net primary production (NPP) at 10 m. Sentinel-2 images and meteorological information are integrated into a light use efficiency (LUE) model. The in-situ dataset obtained during the 2-growing-season field campaigns, including biomass, fPAR, and spectral radiometry, was used to validate the final NPP estimate and the parameterization of the light-use efficiency coefficient for Mediterranean grasslands, based on meteorological data, and fPAR derived from Sentinel-2 images.

Sensor input data:

The daily estimation of net primary production (NPP) was validated using ground measurements collected by IFAPA during two field growing seasons: 2023/24: 12 validation sites with nine measurement points (S2-pixel size) distributed across five pilot farms from December to June, and 2024/25: 12 validation sites in three pilot farms from November to May. All of them are in the region of Pedroches in the northern part of Córdoba, Spain. These sites were selected to account for the variability of climate, topography, and soil properties within the study region. Their locations are shown on the map below.

Several variables were measured using different, and sometimes redundant, instruments, including grassland biomass (Grasmaster Pro, Jenquip EC20 Platemeter, cutting 30x30 cm quadrants), fraction of photosynthetically active radiation absorbed by the plant (fPAR with the Accupar Ceptometer LP-80) and field radiometry (ASD FieldSpec, only on cloudless days). The most important measures for data fusion validation are listed in the table below.

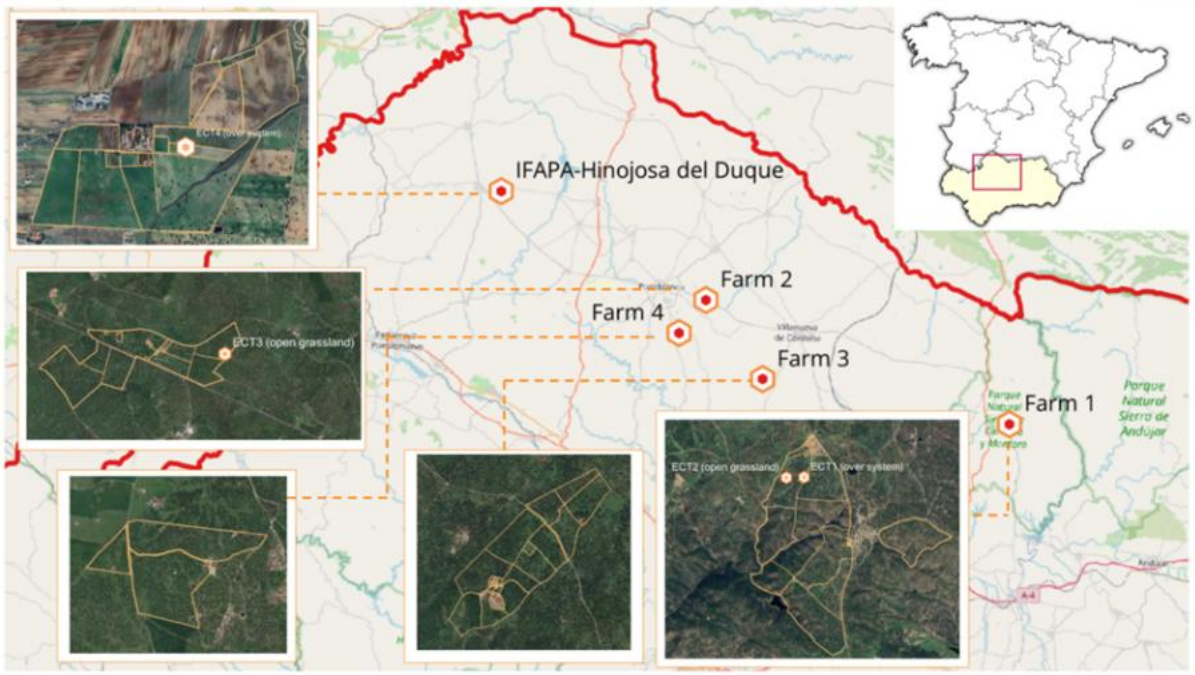


Figure 77: Location of the five pilot farms selected for 2023/2024 field campaign, and the three (farms 1,2 and IFAPA) measured during the 2024/25, and the flux towers (ECT2 and 3) used for section 3.5.3 in the Pedroches region (Spain).

Table 55: Overview of sensor input data (Estimated grassland yield at parcel level)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season (s)	Nr. Training data
Grassland Biomass (Kg MS/ha)	Grassmaster Pro	IFAPA	Pedroches (Spain)	12 parcels (9 S2 pixels) /season	Monthly during the growing season	2023/ 2024 2024/ 2025	+/- 600 per year
	Jenquip EC20 Platemeter						
	Destructive samples						
fPAR	Accupar Ceptometer (LP-80)						
Metorol. Variables (Rad, Ta, u, DPV,..)	Weather stations (different labels)						
Field Radiometr (450-2500nm)	ASD FieldSpec/ Spectralon panel				Only cloudless days		+/- 200 per year

Table 29: Overview of EO input data (Estimated grassland yield at parcel level)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel-2 reflectance	Google Earth Engine	Pedroches (Spain)	2021-2025	All available scenes (3-5 days)	10 m-20 m

Methodology and validation results:

To estimate grassland net primary production (NPP), an adaptation of Monteith crop production model (Monteith, 1977; Gómez-Giraldez et al., 2019) to this ecosystem was used. It focuses on (i) the presence of a variable-density tree layer influencing spectral data and (ii) the estimation of the light-use efficiency parameter for these seminatural grasslands using biomass field measurements. The model calculates net primary production (NPP, kg/ha) using a remote estimation of the fraction of photosynthetically active radiation absorbed by the plant (fAPAR), photosynthetically active radiation (*PAR*) measured by weather stations, and a calibrated value of light use efficiency (ε) following the procedure described in this scheme:

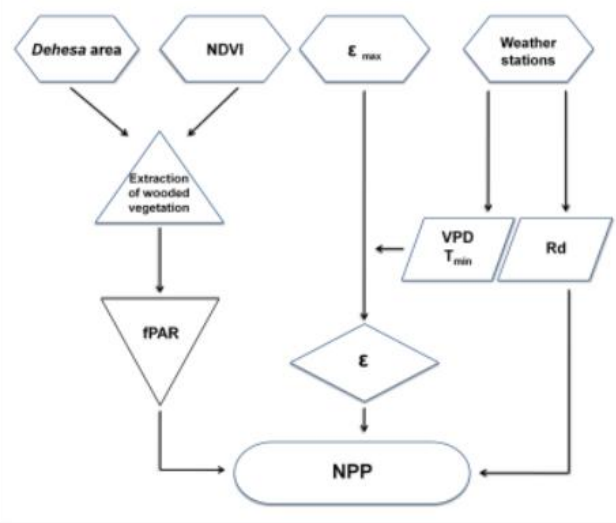


Figure 78: Schematic representation of the adapted Monteith model that is used to calculate NPP

Daily solar radiation (Rad, MJ), minimum and mean air temperature (Tmin, Tmed, ºC), relative humidity (HR, %), and Vapour Pressure Deficit (VPD) for the study period were collected from 10 weather stations around the region. Photosynthetically Active Radiation (PAR) was estimated based on daily solar radiation. Fraction of PAR (fAPAR) values were obtained from the NDVI calculated from S2 using an empirically derived equation tailored to natural Mediterranean grassland. To focus on grassland production, the contribution of oak trees was subtracted using summer NDVI values.

The light use efficiency parameter (ε) was calculated considering Tmin and VPD, which are factors known to reduce a plant's ability to use light efficiently. The maximum value of ε for the ecosystem (εmax) was adjusted by a scalar minimum temperature and a scalar vapor pressure deficit derived from the daily Tmin and VPD values.

A previous version of the model was validated with good results by Gómez-Giraldez et al. (2019) using ground measurements. However, the model was revised, and a new validation was carried out here to better account for the local scale and evaluate the implementation on the studied farms. Unfortunately, the intensive drought of 2022/2023 led to limited and less variable data collection. The current field campaign, 2023/2024, has been completed, and the data from this campaign is being processed. The outcomes will be shared around August 2024. Another field campaign is planned for 2024/2025 to complete the dataset and compensate for the loss of data caused by drought in 2023. The graph below depicts the validation with limited data for previous campaigns (2021-2023).

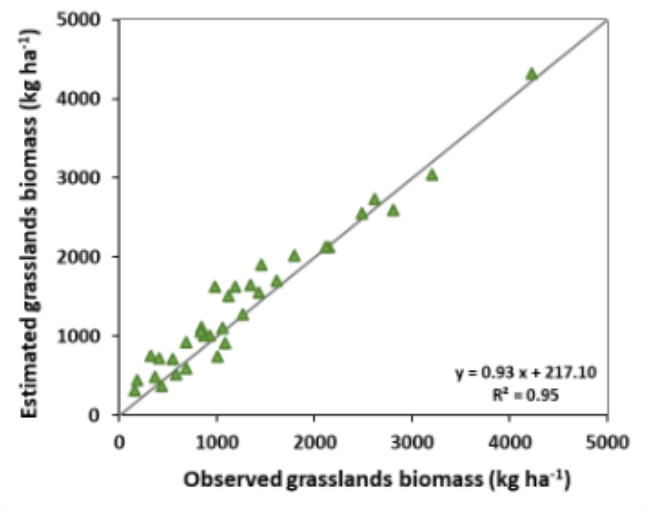


Figure 79: Validation of grassland biomass – results from previous campaigns

Table 56: Overview of sensor integrated data products (Estimated grassland yield at parcel level)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Grassland accumulated NPP	Annual dehesa grasslands in the Pedroches region, Spain	2021-2025	Daily	10 m	RMSE= 224 kg ha-1 (15.5% error) using data from 2021-2023. Yet to be improved.

Use case(s):

The grassland biomass product will be shared with the pilot farms managers and the cooperative technical staff (https://www.covap.es/meet-us/about-covap/the-cooperative) to evaluate the best procedure for transferring the information to the farmers and how it can be used to improve the recommendations on stocking density and grazing rotations.

3.5.3 Improved grassland GPP maps based on flux tower sensors

The grassland GPP maps combine eddy covariance flux sensors, meteorological in-situ data, and EO data (Sentinel-2 and Sentinel-1), and is developed through a collaboration between Deimos and IFAPA. The model being developed is based on a system of Feedforward neural networks trained through GPP data collected in situ by Eddy covariance and temperature sensors. Daily carbon fluxes, together with Sentinel-2 optical imagery and Sentinel-1 SAR data, were integrated into a machine learning model to estimate GPP at high temporal resolution. An independent set of carbon flux observations was used to validate the model outputs and assess their performance, differentiating the results of Sentinel-2 and Sentinel-1 datasets. The algorithm is still under development, and the results obtained so far are preliminary. An algorithm has been tested using historical data from IFAPA . Over the next few months, we will test and improve the algorithm with new EO (Cropsar-2D) and additional in situ data

Sensor input data:

Data from two eddy covariance flux towers installed over grasslands in the *Pedroches* area (see location in section 3.5.2) are being used. Table 57 provides details about the variables and sensors.

Figure 80: Sensors installed in grasslands in the Pedroches area

Table 57: Overview of sensor input data (grassland GPP maps)

Variable	Sensor Data Source	Data Provider	AOI/test sites	Nr of fields/ season	Meas. frequency	Seasons	Nr. Training data
GPP	CSAT 3D anemometer LI 7500A, IRGASON	IFAPA	Pedroches (Spain). Flux towers on open grassland ECT2 and ECT3 (Figure 23)	2/42 months in total	Daily (original 10 Hz averaged 30 min)	Initial historical dataset: 42 (2017-2022)	Initial historical dataset 573 ETC2 + 673 ETC3
ET, H	CSAT 3D anemometer LI 7500A, IRGASON				Daily (original 10 Hz averaged 30 min)		
Rn	4 ways radiometer NR-1				Daily (original 10 Hz averaged 30 min)		
G	HFP01 Huse flux Soil Heat Plates				30 min		
Meteo. Variables: Rad, Ta, u, P, DPV	Weather station (different labels)				30 min		

EO input data:

The data used are, as shown in the table, Copernicus data from Sentinel-1 and Sentinel-2. For Sentinel-1, we used the Level 1 - GRD product. For Sentinel-2, this is MSI, 2-A level data, and Top of Canopy data provided by ESA.

We acquired data from the historical database provided by IFAPA, between 11/20/20/2017 and 10/20/2020, and data from 2022 to 2025 is being processed by IFAPA and will be available in the first months of 2026.

Table 58: Overview of EO input data (grassland GPP maps)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Sentinel-2	Copernicus	Pedroches region	2017/20240	Officially 3 days, really, considering only the cloud free images, on average 10g (with high variance)	10 m, 20 m, 60 m (depending on the bands)
Sentinel-1	Copernicus	Pedroches region	2017/2024	3/6 days	10 m
CropSAR2D	VITO	Pedroches region	2017/2024	5 days	10m

Methodology and validation results:

The developed methodology aims to relate EO data (Copernicus Sentinel-1 and Sentinel-2) with Gross Primary Production (GPP) and Ecosystem Respiration (RECO) using a system of neural networks. This methodology uses data from in-situ sensors (both for training the neural network and for validating the method) and EO data (for training the neural network and subsequently as an independent variable of the model).

The in-situ sensors were Eddy Covariance towers and environmental sensors. The first sensors measure fluctuations in wind speed and gas concentrations on an hourly basis to calculate the mass fluxes of these gases between the ecosystem and the atmosphere. The second sensors measure environmental variables such as: soil temperature, solar radiation, and soil humidity.

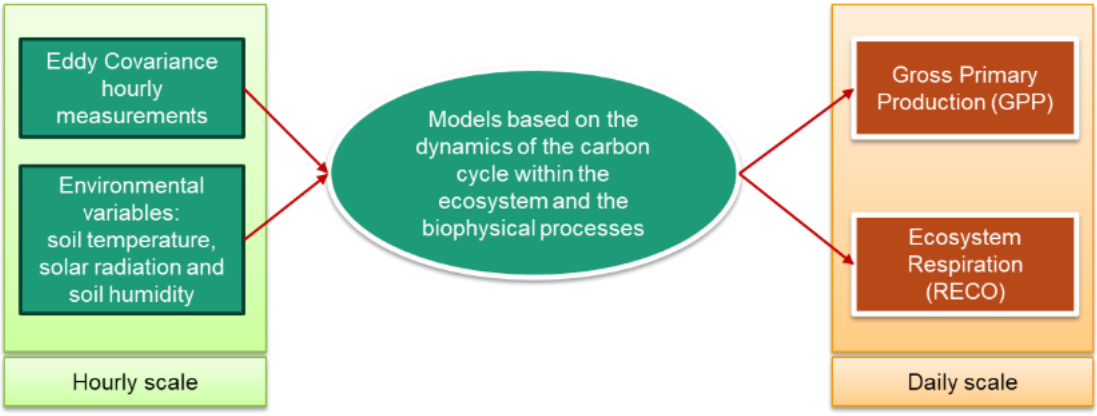


Figure 81: From field measurements to GPP and RECO variables. Data processing scheme.

For the first phase, only S2 data were used to test the algorithm's performance, not only from the results' point of view but also for computational cost and processing times. We are working on introducing S1 data into the algorithm.
The S2 data is MSI sensor data, level 2A (Top of Canopy) for the T30SUH tile, acquired between 10/1/2017 and 8/1/2022. We have excluded from the catalogue all images with cloud cover greater than 20%.

The EO data processing was divided into 6 steps:
1) Selection and acquisition of data from the Copernicus server
2) Raster clipping bat in AOI format
3) Pixel resampling (to 20m) for each band with pixel resolution higher than 20m
4) Pixel extraction inside EC footprint
5) Pixel Average
6) Data Normalization

We thus obtained a database of 294 elements where each element is characterized by: acquisition date, AOI, in situ GPP value, in situ RECO value, reflectance value for each of the 12 available S2 bands. The algorithm used is a Feedforward neural network, composed of an input layer, 2 hidden layers (we also carried out tests for networks with 1 layer), and an output layer. The function tested for hidden layers was tangent-sigmoid, but we are preparing tests with other functions such as ReLU (and its evolutions), Swish and Maxout. A different number of nodes was tested for both layers, setting the maximum number equal to the number of elements in the input layer and trying different combinations.

The Output Layer was constantly defined by a node and the Linear Transfer Function.

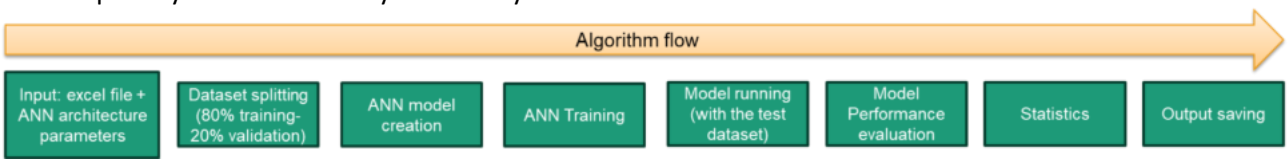


Figure 82: EO and in Situ data for GPP estimations. Algorithm flow chart.

The methodology proposed so far is under development, is being tested and adequately modified to optimize its performance based on the data available and the results obtained from the various tests. The validation of the method was done by randomly dedicating 20% of the dataset to the comparison between variables simulated by the model and variables measured in situ (using EO data as independent variables and GPP/RECO as dependent reference variables). As soon as we have a larger database available, we will structure it as follows: 70% dedicated to training, 15% dedicated to method validation (intended as a tool to avoid overfitting) and 15% dedicated to testing (intended as a tool to quantify and qualify model performance).

Table 59: Overview of sensor integrated data products (grassland GPP maps)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution	Accuracy
Grassland GPP maps	Pedroches/Spain	2017/2024	On average 10 days	10 m	Preliminary results: 75%

Use case(s):
Currently the idea is to have the product tested by IFAPA to use it for research purposes. Moving on to the production phase of the service is not discarded.

3.6 RIL Dairy

The data products explained in this section are intermediate work results, which have been validated in the Dairy RIL during the second iteration round of ScaleAgData. Depending on the results of that validation, decisions were made about further development, validation and potential upscaling of each individual data product, as explained in each of the following subsections. In case these data products are considered useful for users outside the Dairy RIL, further aggregation and anonymization of the data may be required as prerequisite before making them available, because these data products are partly based on confidential data.

3.6.1 Regional productivity of dairy farms

This product, developed by OHB in collaboration with DMK and ATB, includes regionally aggregated time series of milk quality and quantity, allowing analyses of the productivity of dairy farms.

Input data:

Milk quality and quantity parameters are collected for individual farms in the counties of Cuxhaven and Stade in northern Germany all year round from 2018 up to August 2025. The spatial distribution of the farms is visualized in Figure 83 below. Along with the amount of milk as well as the percentage of fat and protein, the dataset contains the classification of respective farms regarding their feeding (indoor housing of cows or cows grazing outside), their geolocations, further quality parameters and metadata.

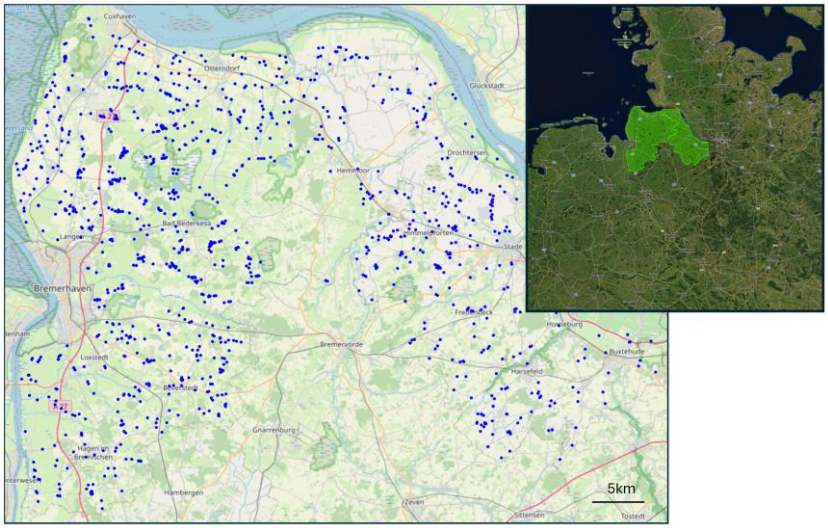


Figure 83: Spatial distribution of dairy farm locations in the sample region in Northern Germany.

Table 60: Overview of input sensor data (Dairy)

Input data	Source	Data provider	AOI/test sites	Nr. Farms/ season	Meas. frequency	Seasons
Milk quantity	Lab analysis	DMK	Counties Stade/ Cuxhaven, Germany	Several hundred farms	2-daily	8
Milk fat percentage	Lab analysis	DMK	Counties Stade/ Cuxhaven, Germany	Several hundred farms	2-daily	8

Milk protein percentage	Lab analysis	DMK	Counties Stade/ Cuxhaven, Germany	Several hundred farms	2-daily	8

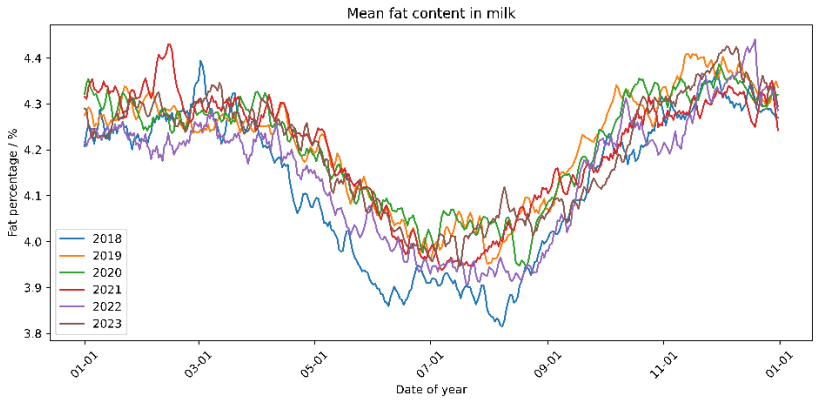
Methodology:

At first, appropriate geographical areas for the initial validation are selected, focusing on regions with a high density/coverage of dairy farms as well as representing homogeneous conditions regarding landscape and weather. For the selected sample regions, raw data is exported from the ERP system of the dairy cooperative. Each data record represents information about one delivery of one farm to a processing plant.

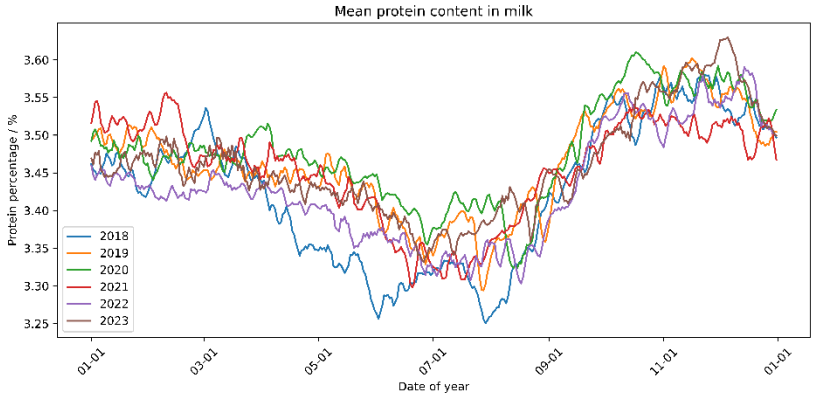
Those raw data are analyzed to identify the relevant parameters:

- the county where the farm is located
- type of farm (cows grazing outside or cows fed in the stable)
- the date of the milk delivery
- the quantity of milk per delivery
- the percentage of fat and protein (results of the lab analysis which is done for each milk delivery)

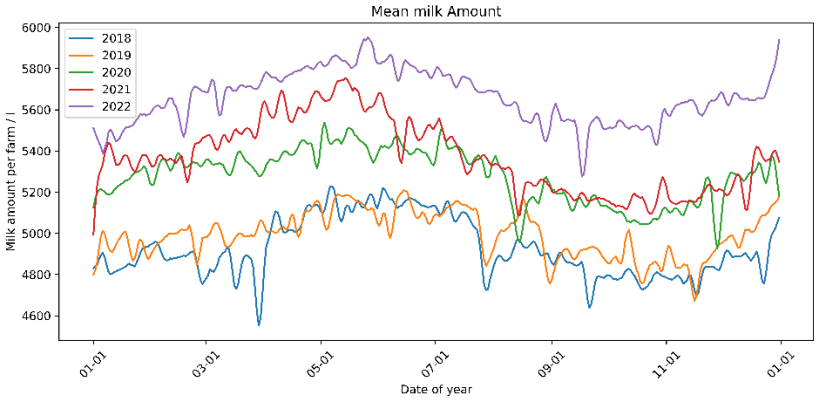
The original data is cleaned/filtered to only contain the relevant data records and fields. The preprocessed basic data set can then be aggregated over different dimensions, summing up the milk quantity and calculating mean values for the quality (percentage of fat and protein). Based on these aggregations, productivity can be compared between different groups of farms, e.g. between farms in different regions or between farms feeding cows in stable vs. farms where cows are grazing outside. Time series of the aggregated data can be extracted from the dataset for further analysis and can also be plotted to allow visual comparison between groups of farms or to show the variation of productivity over the course of each year. As an example, Figure 84 shows time series of daily means of fat and protein content and mean amount of milk delivered by each farm for the years 2018-2023. Recurring annual trends like minimum in the quality parameters and a maximum of milk amount at mid-year can be directly identified. Figure 85 shows a visual comparison of the same parameters for the two sample counties of Cuxhaven and Stade.



a)



b)



c)

Figure 84: Annual time series for the milk quality parameters a) fat and b) protein percentage, as well as c) mean milk quantity per farm. Values are daily means from all farms.

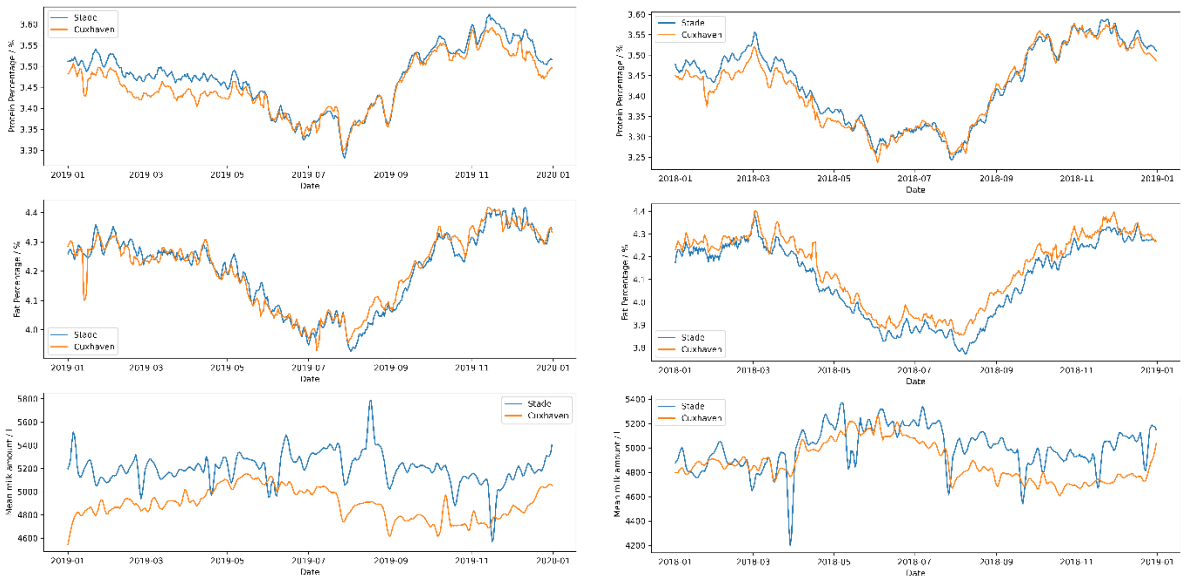


Figure 85: Comparison of milk quality and quantity parameters between counties of Cuxhaven and Stade.

Table 61: Overview of sensor integrated data products (Dairy)

Data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Milk quantity (total kg) per region or per type of farm	Sample counties in Northern Germany (Stade/ Cuxhaven)	2018-2025	2-daily	Region/County
Milk quality – mean % of fat per region or per type of farm				
Milk quality – mean % of protein per region or per type of farm				

Use case(s):

The dairy lab will use this data product as basis to analyze the productivity of farms in different regions and of different types of farms, the development of productivity during each season and the development of productivity over several years.

3.6.2 Deviation of milk quality & quantity

This product, developed by OHB in collaboration with DMK and ATB, includes the deviation of individual (groups of) farms' productivity from the regional productivity, as well as the deviations between the different regions.

Input data:

This builds on top of the data about regional productivity of dairy farms, i.e. the same input data is used.

Table 62: Overview of sensor input data (milk quality & quantity)

Input data	Source	Data provider	AOI/test sites	Nr. Farms/ season	Meas. frequency	Seasons
Milk quantity	Lab analysis	DMK	Counties of Stade/Cuxhaven, Germany	Several hundred farms	2-daily	8
Milk fat percentage	Lab analysis	DMK	Counties of Stade/Cuxhaven, Germany	Several hundred farms	2-daily	8
Milk protein percentage	Lab analysis	DMK	Counties of Stade/Cuxhaven, Germany	Several hundred farms	2-daily	8

Methodology:

Building on top of the time-series data about regional productivity of dairy farms, different deviations are being calculated:

- Deviations of each region's aggregated productivity from the total aggregated productivity over all farms/regions
- Deviations of individual farms' productivity time series compared to the aggregated data
    - Aggregated for all farms in the same region
    - Aggregated for farms of the same type (grazing cows vs. cows fed in the stable)

These deviation time series allow for the assessment of the productivity of individual farms or groups of farms. The resulting data can be plotted to allow visual detection of deviation patterns and can also be used as basis for correlation analyses between individual farms and groups of farms.
Examples regarding the quality parameter "percentage of fat" are shown in the following Figure 86, showing the distribution of individual farms' degree of deviation from the mean value over all farms for two sample years, showing a higher variance in 2019 compared to 2018. Figure 87 shows examples of individual farms' seasonal development of milk quality, compared to the mean value over all farms.

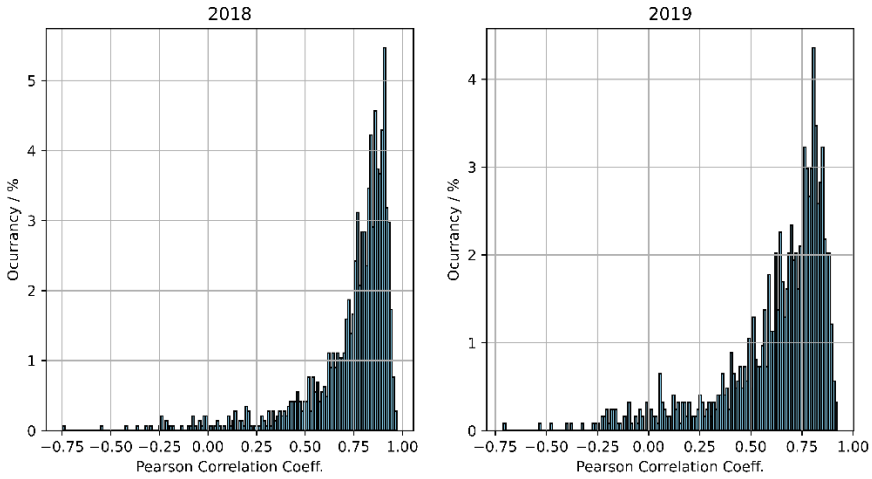


Figure 86: Distribution of Pearson coefficients for the correlation between time series of fat content in milk of individual farms to the time series of mean fat content over all farms.

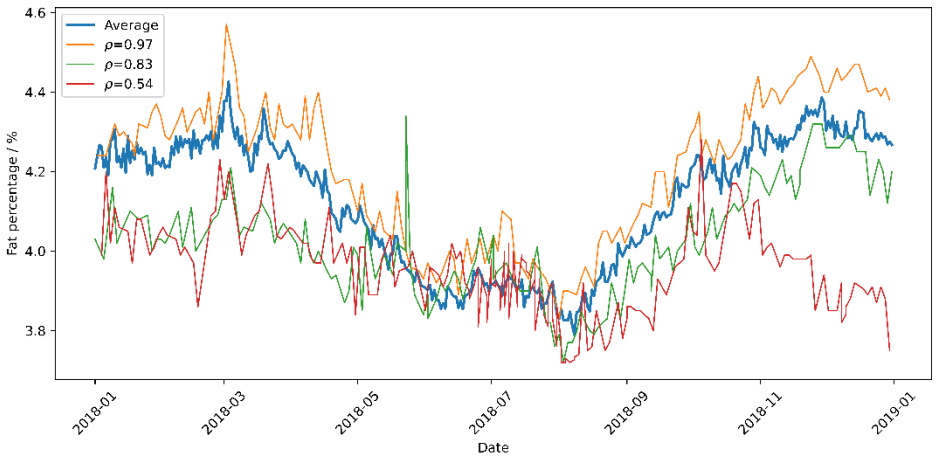


Figure 87: Example for timeseries of fat percentage in milk of individual farms with diverse values of Pearson correlation coefficients to the mean values for year 2018.

Table 63: Overview of sensor integrated data products (milk quality & quantity)

Sensor-integrated data product	Reference values to calculate deviations	AOI/test sites	Time period	Temporal frequen-cy	Spatial resolution
Deviation of milk quality (% of fat, % of protein) and quantity (kg) of individual farms or groups of farms	Aggregate over all farms Aggregate over farms in one region Aggregate over farms of one type	Sample counties in Northern Germany (Stade/ Cuxhaven)	2018-2025	2-daily	Per farm/ region

Use case(s):

The dairy lab will use this data to analyze how far the productivity of individual farms or groups of farms deviates from the mean/totals over different regions or types of farms. This data product may also support milk quality/quantity benchmarking and could offer a baseline to inform the farmers about their specific performance, compared to other producers.

3.6.3 Assessment of grass yield at regional level

This product, developed by OHB in collaboration with ATB, DMK and the external organisation CLAAS, enables an assessment of the grassland productivity on a regional and field-based scale.

Table 64: Overview of sensor input data (regional grass yield)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/ season	Meas. frequency	Seasons
Grass yield per area (ha)	Forage harvester	Machine telemetry	Counties of Stade/ Cuxhaven, Germany	N/A	N/A	6
Grass' dry matter percentage	Forage harvester	Machine telemetry	Counties of Stade/ Cuxhaven, Germany	N/A	N/A	6

Table 65: Overview of EO input data (regional grass yield)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Vegetation Indices based on Sentinel-2	OHB	Full Counties of Stade/Cuxhaven	2018-2023	5-daily	10m
Biomass production	VITO	Representative fields for several individual farms	2018-2023	daily	10m
ESA WorldCover	ESA	Full Counties of Stade/ Cuxhaven	2021	static	10m
Spatial data					
Agricultural land use classification	State ministry	Full Counties of Stade/ Cuxhaven	2023	static	1m

Methodology:

Sensor data collection starts with the export of raw data collected by forage harvesters, encompassing following parameters:

- Machine type
- Fruit classification (manually inserted)
- Vehicle speed
- The timestamp of yield data points
- Yield per area
- Dry matter percentage

Parameters like vehicle speed, fruit classification, and the distribution of yield data are used for data points filtering to ensure that only data points associated with grassland harvest are considered. The aggregation of data points within a county and a given temporal extent enables historical estimates of harvest quality and quantity per area to assess the productivity of grassland on a regional scale, based on analysis of the distribution of yield and dry matter content data.

EO data enables grassland productivity assessment at local and regional scales. For regional scale analysis, the collection of data encompasses automated processing of Level 2 Sentinel-2 data. The processing encompasses re-projecting, clipping and cloud masking of the data products, as well as masking with land cover classification layers and the calculation of vegetation indices, such as NDVI, NDMI or EVI and respective statistics for the full spatio-temporal extent of the AOI and the periods 2018-2023. Furthermore, the collection of growing rates for dry matter productivity and biomass accumulation for field-sized polygons in cooperation with the project partner Vito enables quantitative estimation of grassland productivity of individual fields on the local scale.

Time series and statistics of the EO data products and in-situ harvest data can be provided for the analysis of yearly or seasonal productivity between different regions.

Table 66: Overview of sensor-integrated data products (regional grass yield)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution
In-situ Grassland Productivity (dry matter yield per area)	Counties of Stade/Cuxhaven	2018-2023	Yearly	Per County
EO Grassland Productivity	Counties of Stade/Cuxhaven	2018-2023	Mission-dependent	Index-dependent

Use case(s):

The dairy lab tested this data product to analyze grassland productivity across various regions and scales, as well as to examine the development of grassland productivity over multiple years. This also allows correlation analyses with dairy farms' productivity data, aiming to get new knowledge about the influence of grass yield on the productivity of dairy farms in different regions.

3.6.4 Milk quality forecast on the regional level

This data product, developed by OHB with support by ATB and DMK, is based on outputs from the time series forecasting framework. It enables historical and near real time forecasts of milk fat and milk protein on the regional level with a forecast horizon of up to 15 days.

Table 67: Overview of sensor input data (regional milk quality)

Sensor data	Source	Data provider	AOI/test sites	Nr. fields/season	Meas. frequency	Season(s)
Milk fat / protein percentage	Lab Analysis	DMK	Counties of Stade/Cuxhaven Germany	Several hundred farms	daily	8 (full year)

A customized version of the data product for regional productivity of dairy farms is used as sensor input data. It covers a seamless time series with daily frequency that can be recurrently updated with most recent raw DMK data to account for a NRT application. The data acts as target data in the time series forecasting framework.

Table 68: Overview of EO/ meteorological input data (regional milk quality)

EO products	EO data provider	AOI/test sites	Time period	Temporal frequency	Spatial resolution
	OHB	Counties of Stade/Cuxhaven	2018-2025	daily	0.25° lat/lon
ECMWF IFS Forecast	OHB	Counties of Stade/Cuxhaven	Dec 2024-NRT	Twice daily	0.25° lat/lon

Both input datasets ERA5-Reanalysis and ECMWF IFS are provided by OHB under WP3 activities in Zarr-format, enabling efficient time series extraction. They are recurrently merged into one seamless time series that acts as covariate data in the time series forecasting framework.

Table 69: Overview of sensor-integrated products (regional milk quality)

Sensor-integrated data product	AOI/test sites	Time period	Temporal frequency	Spatial resolution
Milk fat / protein forecast	Counties of Stade/Cuxhaven	Dec 2024-NRT	Twice daily	Per County

Methodology:

The general methodology of the time series forecasting framework is elaborated in chapter 2. Specific aspects of this data product cover the fusion of raw dairy data, preprocessing the meteorological datasets, and the specification of suitable individual forecast models in the time series forecasting framework. As this data product is a Dairy RIL specific output of the time series forecasting framework application, which was presented in detail in section 2.5, Figure 88 repeats its processing pipeline. The purple box covers all necessary preprocessing steps of all input datasets in the RIL6 use case and this data product.

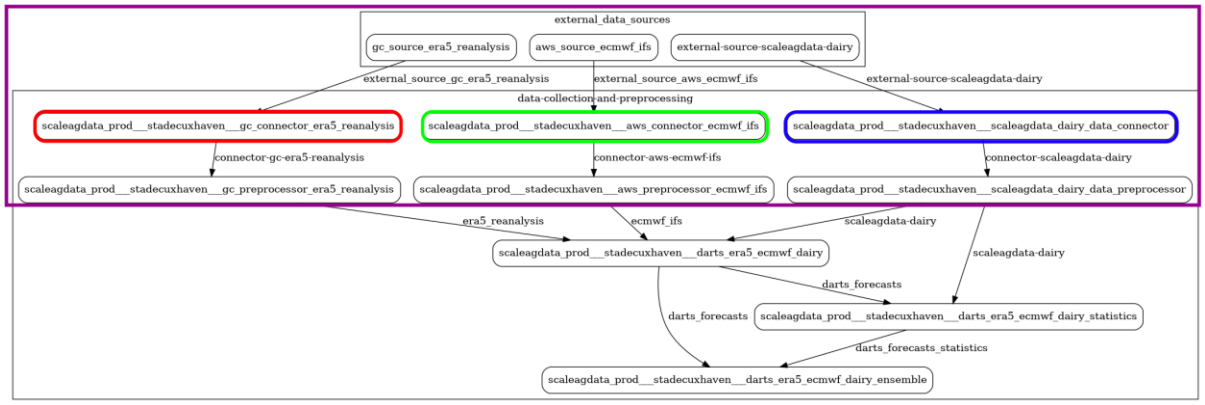


Figure 88: Processing pipeline in the RIL6 Dairy use case. The purple box indicates processing steps of input data sets for the data product of milk quality forecasts on the regional level.

Because the forecasting framework requires continuous, seamless time series, the frequencies of the ERA5-Reanalysis and ECMWF IFS Forecast data (red and green) must be resampled to align with the daily frequency of the sensor input data. Further, for NRT applications, the lack of availability for ERA5-Reanalysis data due to a 5-day latency publishing gap is overcome by stacking multiple ECMWF IFS forecasts.

The preprocessing of raw dairy data produces a seamless time series as target data for the forecasting framework. The processing encompasses the transformation of ERP-data exports into a database that covers the data for every single farm in the AOI. For extracting the milk quality of a region, the milk quality data is weighted by the milk amount per farm, including interpolations for missing data points. Finally, the methodology of the data product of milk quality forecasts on regional level requires a set of suitable individual models with compatible hyper parameters. As a key finding from the conducted correlation analysis in RIL6 during the first iteration cycle of the project, milk quality data exhibits strong seasonal dependencies and variation in combination with high correlation with meteorological parameters. Therefore, models with seasonal components were selected and customized.

An exemplary forecast, the evolution of forecast accuracy over time for the exemplary time interval December 2024 to August 2025 and the histogram of accuracy values can be found in the description of the time series forecast framework in section 2.5.

Use case(s):

The dairy lab will utilize this data for assessments on how NRT prediction of milk quality parameters can influence and support decision making in operations such as milk production planning, milk processing in dairy plants, or logistics.

4. Research and Innovation Environment

The Research and Innovation Environment (RIE) has been established as a cloud-based internal environment to support the consortium's methodology development, integration, and validation activities during the project. In line with the project objectives, the RIE provides a shared technical space in which relevant datasets, intermediate products, analytical components, and development resources can be organised and made available to partners involved in research and prototyping activities. Its primary function is therefore to support the development and testing phase of the project, rather than to operate as a public-facing or production-level service environment.

The RIE has been conceived as an environment where consortium partners can access relevant project data, develop and test analytical workflows, integrate EO and in-situ information, and validate prototype solutions in an iterative manner. In this respect, it supports the technical work associated with the refinement of methods, the testing of data integration approaches, and the assessment of prototype suitability for the RI Labs. Dedicated development resources, including notebook-based environments and supporting software components, are made available to facilitate experimentation and technical validation by partners during implementation.

It is important to clarify that the RIE is not intended to be an operational platform for external users, nor a public environment where datasets and applications would be openly accessible for execution by third parties. Rather, it functions as a consortium-facing research and development workspace, used internally during the implementation period while components, methods, and services are still under active development. For this reason, access to the development environment is restricted to project partners, as this is consistent with its role as a controlled setting for technical development, testing, and integration.

Additionally, the use of the RIE as a development and testing environment is not considered mandatory for all partners and all developments. Instead, the RIE remains available as an optional internal environment, to be used where a shared project setting can provide added value for integration, testing, or validation purposes. In this way, the RIE functions as a flexible support environment within the project, while allowing partners to continue working within their existing technical ecosystems where appropriate.

The role of the RIE should be understood primarily as an internal research, integration, and testing infrastructure that supports the consortium during the iterative development and validation stages of the project. It enables partners to explore technical options, test selected solutions, and assess their suitability for RI Lab needs, but it is not designed to become a permanent operational environment hosting all project services for external use. This distinction is important for understanding both the restricted accessibility of the environment and the level of technical maturity of some of its components during the project period.

At project completion, the elements intended to remain publicly available are more limited and focus on dissemination and reuse of results rather than operational service provision. In particular, the project will maintain a catalogue of input and output datasets and a public GitHub repository containing the applications and software components developed within the project, where these are intended for sharing and reuse. The RIE itself should therefore be considered as the project's internal technical environment for development and validation, while the public legacy is reflected in the documented datasets, software assets, and associated project outputs.

An overview of the RIE and user manual was already provided in D4.1. In addition, the first version of this deliverable complemented the user manual with a set of best practices for data handling within the RIE, which are also summarized below.

- **Data for exclusive use (per person or lab team, not shared):** Each user has personal storage in the RIE, which should be used for data that does not need to be shared with other partners. Team members can use this space like any other drive (create folders and subfolders, add/delete files, etc.). If more space is required, the user can request it via a support ticket at the service desk.
- **Data to be shared between ScaleAgData partners and/or to be shared with external users during or after the project end:** These datasets should be made available through the RIE data catalogue. It is not feasible for the RIE team to anticipate all datasets needed by each RIL. To address this, each RIL team should raise a service desk ticket requesting a given dataset to be added to the catalogue. In response, the service desk team will send a data collection metadata template to be completed with the necessary information to harvest the dataset and define data access permission rules. If no specific access rules are provided, the dataset will by default be accessible to all project partners. Typically, it may take up to two weeks for the dataset to become available in the RIE. Once catalogued, the user will be notified via the corresponding service desk ticket. Details on how to access catalogued datasets are provided in the RIE user manual (D4.1).
- **Data available in external sources (e.g. Google Drive, OneDrive) that must be accessed from the RIE:** There are already examples from the Soil Health RIL where files stored on Google Drive are accessed, and new files are written back to that external drive using specific Python libraries. Therefore, no major limitations are foreseen for this type of access in the RIE. However, if a user encounters difficulty setting this up, they should raise a service desk ticket so that the RIE operations team can investigate.

Efforts have been made to identify all EO datasets required in the RIE environment to support RIL experiments. As a result, a ScaleAgData catalogue has been set up at https://scaleagdata.nextgeoss.eu/, and work is ongoing to make all required datasets discoverable in this catalogue. The complete list of input EO datasets is provided in the updated version of the deliverable D3.3 – EO data collection, spatio-temporal preparation services together with a calendar to have all datasets ingested into the catalogue until end of June 2026. The table below provides an overview of the same information for the output products. It should be noted that in this case only a subset of data will be provided for demonstration purposes.

Table 70: Calendar for the ingestion in the ScaleAgData catalogue of the main project output datasets

Output products	Data provider	RI Lab (partner)	Timeline for ingestion in catalogue
Field water status and predicted yield for target crops	IES/MIGAL	Water productivity	30/06/2026
Satellite based field water status and predicted yield for target crops	IES/MIGAL	Water productivity	30/06/2026
Soil moisture and evapotranspiration aggregation at LAU / Commune Level	NP	Crop management / Agri-environmental monitoring for policy makers	30/06/2026
Calculated indicators (aggregates) based only on the ground truth evidence	NP	Crop management / Agri-environmental monitoring for policy makers	30/06/2026

Calculated indicators based on data assimilation mechanisms along with the respective annotations	NP	Crop management / Agri-environmental monitoring for policy makers	30/06/2026
Aggregated pesticide uses for policy makers	NP	Crop management / Agri-environmental monitoring for policy makers	30/06/2026
Calculated sustainability indicators	Horta	Crop management / Sustainability performance	15/07/2026
DSS model outputs	Horta	Crop management / Sustainability performance	15/07/2026
Statistical data on the accuracy of observations of the occurrence of agrophages	WODR/PSNC	Crop management / Early pest detection	15/07/2026
Improved predicted agrophage occurrence data based on geolocation	WODR/PSNC	Crop management / Early pest detection	15/07/2026
Predicted overall level of agrophage occurrence risk for the selected region	WODR/PSNC	Crop management / Early pest detection	15/07/2026
Potato yield estimates (subfield level)	VITO	Yield monitoring	31/07/2026
Improved tare yield estimates for potatoes	UGent	Yield monitoring	31/07/2026
Winter wheat yield estimates (LUKE)	LUKE	Yield monitoring	31/07/2026
Winter wheat yield estimates (VITO)	VITO	Yield monitoring	31/07/2026
EO based regional soil organic carbon map	ILVO/AUTH	Soil health	31/07/2026
Soil health indicator estimates	ILVO/AUTH	Soil health	31/07/2026
Gap-filled grasslands LAI maps at parcel level	EURAC	Grasslands	14/08/2026
Estimated grassland yield at parcel level	IFAPA	Grasslands	14/08/2026
Improved grassland GPP maps based on flux tower sensors	DEIMOS	Grasslands	14/08/2026
Regional productivity of dairy farms	OHB	Dairy	28/08/2026
Deviation of milk quality & quantity	OHB	Dairy	28/08/2026
Assessment of grass yield at regional level	OHB	Dairy	28/08/2026
Milk quality forecast on the regional level	OHB	Dairy	28/08/2026

Regarding the methodological frameworks/applications, between M18 and M36 of the ScaleAgData project, the RIE operations team set up and began maintaining a GitHub organization at https://github.com/scaleagdata, where partners were requested to commit the latest version of all developed methodological frameworks described in Chapters 2. The version available in this GitHub organization should be considered as the reference version for partners/external users wishing to use a given application.

In the following months an extra effort will be made to bring all the main methodological frameworks to the Github repository, together with the respective documentation that will allow external users to understand the context of development of that application, its purpose and guidelines to be able to replicate results achieved in ScaleAgData.

Table 71: Calendar for complete integration of methodological frameworks/applications in the Github repository

Methodological Framework/ Application	**Developer**	**Link to GIthub repository**	**Foreseen date for integration in Github repository with documentation**
Federated Learning	ILVO, AUTH, Deimos	https://github.com/ScaleAGData/scaleagdata_flower_ai https://github.com/ScaleAGData/federated_ai_soillab	Private repository but to be made public and completed in 30/06/2026
Data Assimilation	LUKE	https://github.com/ScaleAGData/luke-digital-twin/tree/main/CMAAS	Private repository but to be made public and completed by 30/09/2026
Data Integration: Few shot learning	VITO	https://github.com/ScaleAGData/scaleag-vito	Available but to be completed in 30/06/2026
Data Integration: Sensor data integration framework	DHI	https://github.com/ScaleAGData/dhi-mleo	Available but to be completed in 30/07/2026
Super Resolution	ICCS	https://github.com/ScaleAGData/scaleagdata-iccs	Available but to be completed in 30/08/2026
Edge Processing	EGM	https://github.com/ScaleAGData/edgeVM	Private repository but to be made public and completed in 30/08/2026
Time series processing	EGM	https://github.com/ScaleAGData/streamlite	Private repository but to be made public and completed in 30/08/2026
Dairy RIL - ECMWF IFS data processor	OHB	https://github.com/ScaleAGData/DAIRY-RILAB	Available but to be completed in 30/08/2026
Tools for Metadata harmonization	EGM	https://github.com/ScaleAGData/aim-translation	Available but to be completed in 30/06/2026

Implementation of data sharing API	NP	-	30/09/2026
Crop mgt RIL: dashboard and crop classification	NP	https://github.com/ScaleAGData/CROP-RILAB	Available but to be completed in 30/07/2026
Grassland RIL	Deimos, EURAC, IFAPA	https://github.com/ScaleAGData/Grassland-RILAB	Available but to be completed in 30/06/2026
Yield RIL - Potato Tare	UGent	https://github.com/ScaleAGData/YIELD-RILAB	Available but to be completed in 30/07/2026
Soil Health RIL	AUTH, ILVO	https://github.com/ScaleAGData/SOIL-RILAB	Available but to be completed in 30/07/2026

Some of the RI Labs also included in the Github repository more specific models they have developed for generating the data products described in Chapter 3. Others might join in this effort, after internal discussion on the IPR management related to the sharing of these models.

5. References

Gómez-Giráldez, P.J., Aguilar, C., Caño, A.B., García Moreno, A., & González-Dugo, M.P. (2019). Remote sensing estimation of net primary production as monitoring indicator of holm oak savanna management. Ecological Indicators, 106, 105526. https://doi.org/10.1016/j.ecolind.2019.105526

Guo, X., Fang, X., Zhu, Q., Jiang, S., Tian, J., Tian, Q., & Jin, J. (2023). Estimation of root-zone soil moisture in semi-arid areas based on remotely sensed data. *Remote Sensing*, *15*(8), 2003. https://doi.org/10.3390/rs15082003

Kisekka, I., Peddinti, S. R., Kustas, W. P., Mc Elrone, A. J., Bambach-Ortiz, N., McKee, L., & Bastiaanssen, W. (2022). Spatial–temporal modeling of root zone soil moisture dynamics in a vineyard using machine learning and Remote Sensing. *Irrigation Science*, *40*(4–5), 761–777. https://doi.org/10.1007/s00271-022-00775-1

Li, M., Sun, H., & Zhao, R. (2023). A review of root zone soil moisture estimation methods based on remote sensing. *Remote Sensing*, *15*(22), 5361. https://doi.org/10.3390/rs15225361

Friedman, Jerome H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics* 1189-1232. https://doi.org/10.1214/aos/1013203451

Monteith, J.L. (1977). Climate and the efficiency of crop production in Britain. Philos. Trans. R. Soc. London. B, Biol. Sci. 281, 277, LP-294.

Souissi, R., Zribi, M., Corbari, C., Mancini, M., Muddu, S., Tomer, S. K., Upadhyaya, D. B., & Al Bitar, A. (2022). Integrating process-related information into an artificial neural network for root-zone soil moisture prediction. *Hydrology and Earth System Sciences*, *26*(12), 3263–3297. https://doi.org/10.5194/hess-26-3263-2022

Tseng, G., Cartuyvels R., Zvonkov, I., Purohit, M., Rolnick, D., & Kerner, H. (2024). Lightweight, pre-trained transformers for remote sensing timeseries. *arXiv preprint arXiv:2304.14065* (2023). https://arxiv.org/abs/2304.14065

Tseng, G., Zvonkov, I., Nakalembe, C., & Kerner, H. (2021). CropHarvest: a global satellite dataset for crop type classification. *Advances in Neural Information Processing Systems*. CropHarvest: A global dataset for crop-type classification (neurips.cc)

Yinglan, A., Guoqiang, W., Peng, H., Xiaoying, L., Baolin, X., Qingqing, F. (2022). Root-zone soil moisture estimation based on remote sensing data and Deep Learning. *Environmental Research*, *212*, 113278. https://doi.org/10.1016/j.envres.2022.113278