TRUSTWORTHY, COGNITIVE AND AI-DRIVEN COLLABORATIVE ASSOCIATIONS OF IOT DEVICES AND EDGE RESOURCES FOR DATA PROCESSING

Grant Agreement no. 101136024

Deliverable D5.1
Use cases technological developments

Programme:	HORIZON-CL4-2023-DATA-01-04
Project number:	101136024
Project acronym:	EMPYREAN
Start/End date:	01/02/2024 – 31/01/2027

Deliverable type:	Report
Related WP:	WP5
Responsible Editor:	EV ILVO
Due date:	31/07/2025
Actual submission date:	31/08/2025 (initial submission 1.0) 29/10/2025 (revised version)

Dissemination level:	Public
Revision:	FINAL

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101136024

Revision History

Date	Editor	Status	Version	Changes
30.04.2025	Jan Bauwens (ILVO) Theodoros Chalazas (ILVO)	Initial Draft	0.1	ToC definition
30.05.2025	All partners	Draft	0.2	First round of input collection
30.06.2025	All partners	Complete deliverable	0.5	First complete deliverable
15.07.2025	Reviewers feedback	Reviewers feedback	0.8	Reviewers provide feedback
30.07.2025	Jan Bauwens (ILVO) Theodoros Chalazas (ILVO)	Final version	1.0	Final Deliverable version
29.10.2025	Jan Bauwens (ILVO), Javier Martín (IDEKO), Eduardo Canovas (UMU), Panagiotis Kokkinos (ICCS)	Revised	1.1	Address first project review comments

Author List

Organization	Author
ILVO	Jan Bauwens, Theodoros Chalazas
IDEKO	Aitor Fernández, Javier Martín
UMU	Antonio Skarmeta, Eduardo Canovas
ICCS	Aristotelis Kretsis, Panagiotis Kokkinos, Manos Varvarigos
RYAX	Yiannis Georgiou, Pedro Velho
ZSCALE	Ivan Paez
NVIDIA	Dimitris Syrivelis
CC	Marton Sipos, Marcell Feher, Daniel E. Lucani
NUBIS	Anastassios Nanos, Christos Panagiotou
NEC	Roberto González, Jaime Fúster

Internal Reviewers

- CC (Marton Sipos)
- UMU (Alonso)
- ICCS (Aristotelis Kretsis, Panagiotis Kokkinos)

Abstract: Deliverable D5.1, *Use Cases Technological Developments*, documents the initial implementation and integration of EMPYREAN's AI-driven, secure, and federated architecture across multiple real-world use cases. These include anomaly detection in robotic machining, dynamic soil organic carbon assessment in agriculture via UAVs and robotics, and 5G-enabled vehicle-assisted edge computing. The report outlines the technological progress made in embedding these applications into the EMPYREAN platform, detailing the development of edge-cloud workflows, stakeholder engagement, integration activities, and preliminary testing. It marks a foundational step toward validating EMPYREAN's vision of trustworthy, cognitive associations of IoT and edge devices for distributed data processing, setting the stage for more advanced iterations in future phases.

Keywords: Edge Computing, IoT, Precision Agriculture, Anomaly Detection, 5G Services, AI Orchestration

Disclaimer: *The information, documentation and figures available in this deliverable are written by the EMPYREAN Consortium partners under EC co-financing (project HORIZON-CL4-2023-DATA-01-04-*101136024*) and do not necessarily reflect the view of the European Commission. The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The reader uses the information at his/her sole risk and liability.*

Table of Contents

List of Figures

Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
ARTOF	Agricultural Robot Taskmap Operation Framework
CTI	Cyber Threat Intelligence
DRS	Data Recording System
DSS	Decision Support System
EC	European Commission
EO	Earth Observation
FCS	Fingerprint Comparison System
FGS	Fingerprint Generation System
GNSS	Global Navigation Satellite System
GPU	Graphics Processing Unit
IF	Isolation Forest
IoT	Internet of Things
KNN	K-Nearest Neighbors
KDE	Kernel Density Estimation
LOF	Local Outlier Factor
ML	Machine Learning
NDVI	Normalized Difference Vegetation Index
NFV	Network Function Virtualization
NIR	Near-Infrared
OBD	On-Board Diagnostics
ODM	OpenDroneMap
OCSVM	One-Class Support Vector Machine
PSDK	Payload Software Development Kit
RTK	Real-Time Kinematic
SDK	Software Development Kit
SOC	Soil Organic Carbon
SWIR	Short-Wave Infrared
UAV	Unmanned Aerial Vehicle
UC	Use Case
vRSU	Virtual Road-Side Unit
VNF	Virtual Network Function

1 Executive Summary

This deliverable, D5.1, entitled “Use Cases Technological Developments”, presents the first iteration of the technological work performed in the EMPYREAN project to implement and validate a set of complex, real-world use cases. These use cases will demonstrate the potential of EMPYREAN’s novel approach to trustworthy, cognitive, and AI-driven collaborative associations of IoT devices and edge resources for distributed data processing. The deliverable documents the progress made towards creating advanced technological solutions that integrate heterogeneous devices and computational resources at the edge and cloud, while maintaining security, trust, and adaptability throughout the processing continuum.

Specifically, this report details the project’s initial technological developments across key domains. The first involves anomaly detection in robotic machining cells, where the focus has been on creating robust, real-time process monitoring applications capable of detecting deviations during composite manufacturing operations. These developments aim to enhance the reliability and efficiency of robotic machining by combining high-frequency sensor data collection with intelligent processing close to the source. The second domain concerns proximal sensing in agricultural fields, where advanced unmanned aerial vehicles (UAVs) and ground robotics are deployed to assess soil organic carbon levels dynamically. This effort seeks to enable precision agriculture through detailed spatial mapping of soil properties, combining UAV-based remote sensing and ground-based spectroscopic analysis with machine learning models operating at the edge.

Additionally, this deliverable includes an exploration of secure orchestration in smart factory environments, particularly through international collaboration efforts in South Korea. These developments focus on designing mechanisms that safeguard data integrity and system availability in distributed manufacturing environments exposed to cyber threats. Furthermore, the deliverable addresses the unique challenges of vehicle-assisted edge computing in 5G-enabled environments, where orchestration must dynamically adapt to the mobility of connected vehicles while maintaining low-latency data processing and resilience against potential security breaches.

This deliverable outlines the architectural choices, integration activities, stakeholder engagement, and testing strategies applied to each use case. It demonstrates the feasibility of implementing EMPYREAN’s Association-based architecture, which federates IoT and edge resources and manages them autonomously using AI-driven orchestration. The deliverable also identifies and discusses the challenges faced in integrating heterogeneous technologies and ensuring seamless operation across distributed computing environments.

By reporting on the progress of these use cases, Deliverable D5.1 marks an important milestone in the EMPYREAN project, laying the foundation for future iterations that will expand functionality, enhance interoperability, and increase the maturity of the technological solutions. The insights gained from this iteration will feed directly into the refinement of the EMPYREAN integrated platform, guiding the development of more advanced releases capable of supporting a wider range of industrial and societal applications.

2 Introduction

The EMPYREAN project addresses the emerging challenges of managing hyper-distributed computing environments, where large numbers of heterogeneous IoT devices, edge resources, and cloud infrastructures must operate together seamlessly to deliver reliable, secure, and efficient services. Central to EMPYREAN's vision is the concept of Associations: federated clusters of IoT devices and edge computing resources that collaborate autonomously and intelligently. These Associations enable flexible, dynamic orchestration of computational and data processing tasks across the cloud-to-edge continuum, supported by cognitive, AI-driven decision-making mechanisms that adapt in real-time to changing operating conditions.

This deliverable, D5.1, documents the first development iteration of the EMPYREAN project's use cases, which represent practical, real-world scenarios designed to validate the project's architectural approach and core functionalities. It provides a comprehensive account of the technical progress achieved in implementing initial versions of the use cases, including robotic machining process monitoring, proximal sensing in precision agriculture, security orchestration in smart factory and 5G vehicle-assisted environments, and advanced edge-based workflows for dynamic service deployment.

Note that, starting from June 2025 a project amendment is in effect, providing a new description for EMPYREAN's third use case, namely: 5G-Enabled Vehicle-Assisted Services. This use case is currently in design phase as described in the present document (Section 5). Also, in light of these changes, the new 3rd use case has been aligned to support and reflect ongoing and future collaboration with Korean colleagues under EMPYREAN, through the "Korea International Collaboration Use Case".

A central focus of this deliverable is the integration of the developed workflows with EMPYREAN components, demonstrating how the use cases are adapted to operate within the EMPYREAN Association-based architecture. The document details how workflows have been decomposed, redesigned, or extended to run on the distributed, federated system envisioned by EMPYREAN. It also describes how these workflows are being integrated with critical platform services such as orchestration, telemetry, AI-enabled decision-making, and secure data management.

By presenting the results of this first iteration, the deliverable highlights the feasibility of applying EMPYREAN's concepts to diverse application domains, the architectural and practical challenges encountered during the integration process, and the initial validation activities performed in collaboration with stakeholders. These insights set the stage for refining and expanding both the use cases and the EMPYREAN platform in future iterations, ultimately moving towards a robust, flexible, and secure IoT-edge-cloud ecosystem capable of meeting the demands of complex, real-world industrial and societal applications.

2.1 Purpose of this document

The purpose of Deliverable D5.1 is to provide a detailed and structured account of the first iteration of technological developments in the EMPYREAN use cases, documenting how initial implementations of key applications have been created and how these workflows are being integrated into EMPYREAN's distributed system architecture. The document outlines the design choices made, the implementation strategies followed, the integration with EMPYREAN components, and the results of initial tests and stakeholder feedback activities.

Through this deliverable, the consortium provides evidence of technical progress towards the objectives defined for Work Package 5, demonstrating the practical application of EMPYREAN's innovative concepts in real use cases. Additionally, the document serves as a communication tool for the European Commission, project stakeholders, and the wider community, providing the current status of relevant development efforts, the challenges identified during the first iteration, and the planned steps for the refinement and extension of the use cases and the integrated platform.

2.2 Document structure

This deliverable is structured to ensure clarity and transparency in presenting the work conducted during the first iteration of the EMPYREAN use cases. Following this introduction, the document provides detailed sections for each use case, describing their objectives, technical background, initial development efforts, integration with EMPYREAN's components and workflows, stakeholder engagement activities, testing strategies, and the planned next steps.

After the comprehensive presentation of individual use cases, the document includes a section dedicated to the broader integration of the EMPYREAN platform. This section outlines the adopted integration approach to ensure interoperability between the developed workflows and the core platform components, describes the current status of the platform's implementation, and provides an overview of the verification testing process and timeline for future releases.

The document concludes with a summary of the technological progress achieved during this iteration, the key challenges encountered, and the next steps planned for advancing the use cases and platform development towards the project's objectives. References to relevant standards, scientific literature, and related deliverables are also provided.

2.3 Audience

The primary audience for this deliverable comprises the technical and research partners of the EMPYREAN consortium, including developers, engineers, software architects, and project managers directly involved in the design, implementation, and integration of the use cases and platform components. Additionally, the deliverable is intended for the European

Commission project officers and reviewers overseeing the progress, quality, and impact of the EMPYREAN project within the Horizon Europe program.

Beyond the consortium and funding bodies, the document may also be of interest to external stakeholders such as industrial practitioners, technology providers, and researchers working in fields related to distributed edge-cloud computing, trustworthy AI, IoT system orchestration, and cyber-physical system security. Policymakers, standardization bodies, and organizations considering the adoption of federated edge computing solutions may also benefit from the insights presented, as they illustrate both the opportunities and challenges of deploying AI-driven Associations of IoT and edge resources in real-world environments.

3 Anomaly Detection in Robotic Machining Cells Technological Developments

3.1 Overview

The integration of robotic systems in machining processes offers clear advantages in the manufacturing sector, including greater flexibility and cost reduction when compared to traditional machine tools. Robots enable rapid adjustments in both production lines and product designs, providing a high degree of adaptability. This adaptability is a crucial advantage in rapidly changing markets, where manufacturers must respond swiftly to evolving consumer demands and technological advances while maintaining precision and efficiency.

Nevertheless, robotic machining presents specific challenges, especially when working with composite materials. These challenges include issues such as loss of precision, tool breakage, and process instability. To address them effectively, robust real-time process monitoring is essential in order to ensure consistent product quality and operational efficiency.

This use case aims to develop a solution that can detect anomalies during machining by monitoring robotic machining cells in near real time during composite manufacturing operations. These operations include turning, milling, and drilling. By leveraging mid- to high-frequency sensor data, the system will generate timely alerts, enabling quick intervention and minimizing production losses.

Implementing such a solution involves several technical challenges. One major difficulty stems from the large volume of data generated, which must be processed despite the limited computational capacity typically available at the edge.

3.2 Initial development

The initial development of this use case within the EMPYREAN project builds upon a layered architecture commonly adopted by machine tool clients. This architecture is composed of two main computational layers: (i) deep edge devices, which are physically integrated with the robotic systems, and (ii) far edge resources, which are hosted on the client's premises (see Figure 1). This separation allows for an efficient distribution of computational workloads, enabling tasks to be assigned to the most suitable layer based on their processing requirements.

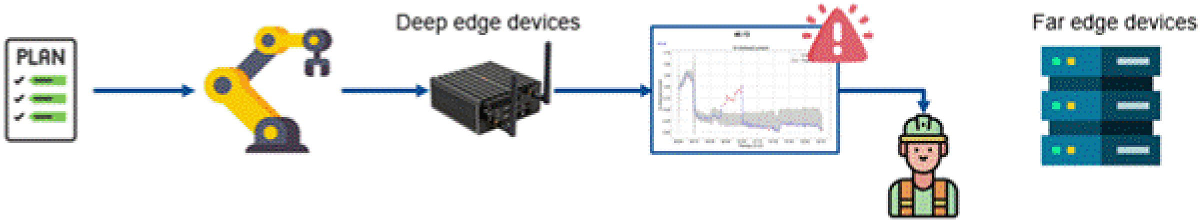

Figure 1: Layered architecture with deep-edge and far-edge components for distributed processing in robotic machining.

IDEKO has developed a solution for anomaly detection in robotic machining processes (see figure 2). In the current setup, manufacturing orders are planned and divided into discrete operations to be executed by a robotic system. Once a machining task is completed, data from selected operations is analysed offline through a workflow application specifically developed for this purpose. This application relies on low-frequency data and applies statistical methods to generate a unique fingerprint for each operation. If an anomaly is detected, the system notifies the operator. The segmentation of manufacturing processes into individual operations enables grouped analysis of similar tasks, improving the reliability of the anomaly detection process. The data processing pipeline is implemented using a microservices architecture, ensuring modularity and scalability.

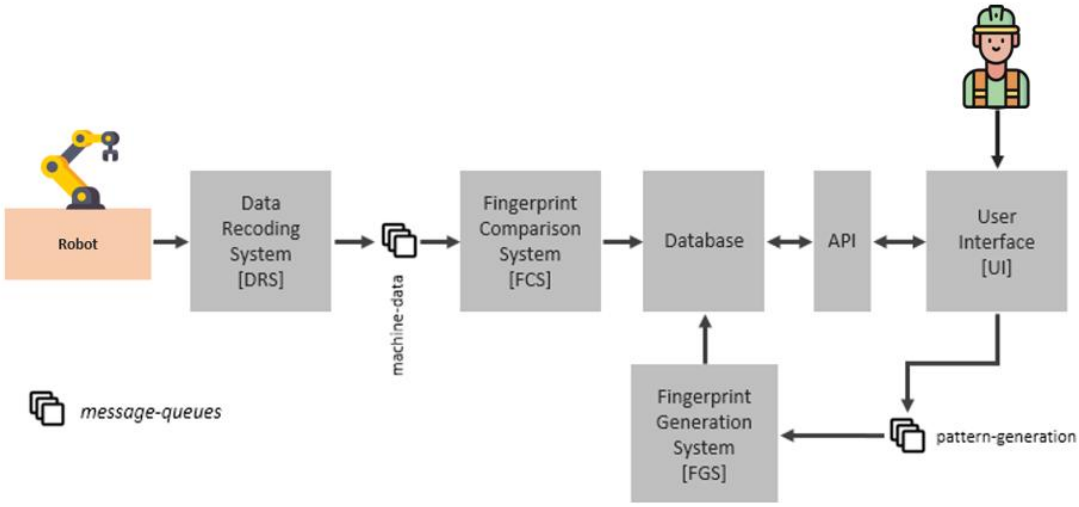

Figure 2: Offline monitoring workflow using statistical fingerprinting for anomaly detection.

As illustrated in Figure 2, the robot sends data to the Data Recording System (DRS), where it is compared against optimal operation patterns stored in the Fingerprint Comparison System (FCS) using statistical metrics. If deviations are detected, an alert is triggered for the operator. Additionally, the Fingerprint Generation System (FGS) allows for the continuous updating of reference patterns, ensuring the system remains adaptive to process variations.

While this current implementation has proven effective, the project aims to evolve it towards a more advanced and responsive monitoring system. The transition involves addressing three key objectives that define the future direction of development.

Objective 1: Use mid-to-high frequency data

The current system operates with low-frequency data, which limits the granularity of process monitoring. To overcome this, efforts are underway to adapt the monitoring application to support mid- to high- frequency data acquisition. This requires upgrading the robot's data acquisition devices (Figure 3) to capture more detailed signals, enabling a finer analysis of the machining process. In parallel, enhancements to the DRS will enable handling this higher-resolution data streams, allowing the FCS to perform more accurate and timely anomaly detection, all while maintaining efficient and scalable processing within the distributed deep – far edge architecture.

Figure 3: Mid-frequency data acquisition setup integrated into robotic machining systems.

Objective 2: Apply machine learning models for time-series comparison

In parallel with architectural improvements, the use case aims to detect anomalies in the robotic machining process through data acquisition and the application of artificial intelligence models. To achieve this, a series of tasks have been defined and executed, including the extraction of relevant features from the acquired data, the tuning of novelty detection models, and the evaluation of their performance (Figure 4).

The general approach is as follows. An anomaly detection model is trained using existing data to learn a "normality pattern." It is important to emphasize that this is done exclusively with machining operations that meet quality requirements, commonly referred to as "good parts." When a new part is machined, the system infers the machining status, using the trained model, and classifies it as anomalous or non-anomalous. Therefore, we are approaching the task as a novelty detection problem.

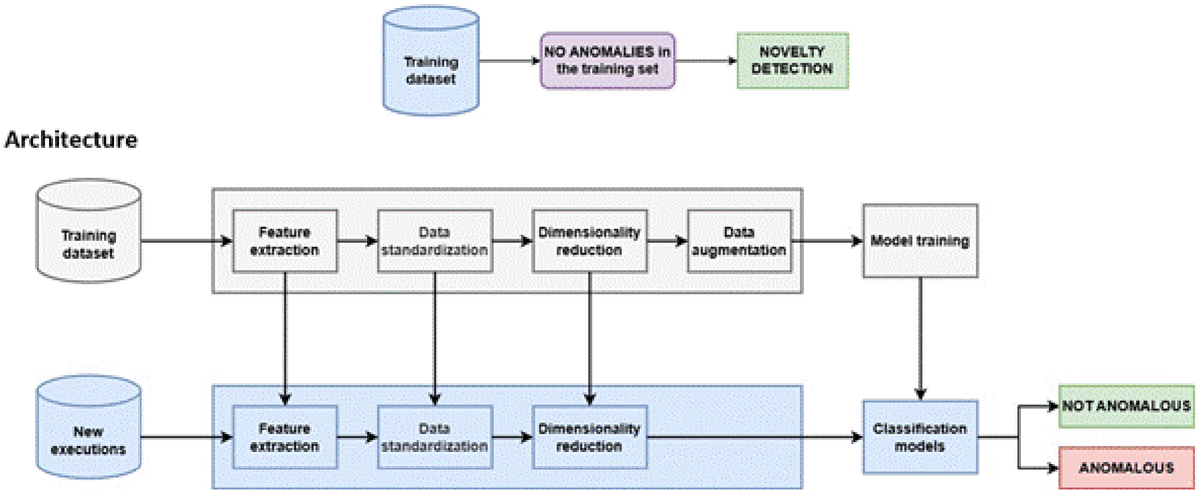

Figure 4: Machine learning architecture for anomaly detection in time-series data.

Various models are currently being evaluated using Scikit-learn[1] library in Python, with the objective of identifying the most effective approach for the specific characteristics of robotic machining data. This includes exploring different anomaly detection strategies, such as density estimation, decision boundary methods, isolation techniques, and clustering. The models under evaluation include:

- Density-based models:
  - Kernel Density Estimation (KDE)
  - Gaussian Mixture Models (GMM)
- Decision boundary-based models:
  - One-Class Support Vector Machine (OCSVM)
- Isolation-based models:
  - Isolation Forest (IF)
- Clustering-based models:
  - Local Outlier Factor (LOF)
  - K-Nearest Neighbors (KNN)

Each model is being trained using normalized and dimensionally reduced data (Figure 5).

[1] https://scikit-learn.org/stable

Models	GMM	KDE	OCSVM	IF	LOF	KNN
Testing	✓	✓	✓	✓	✓	✓

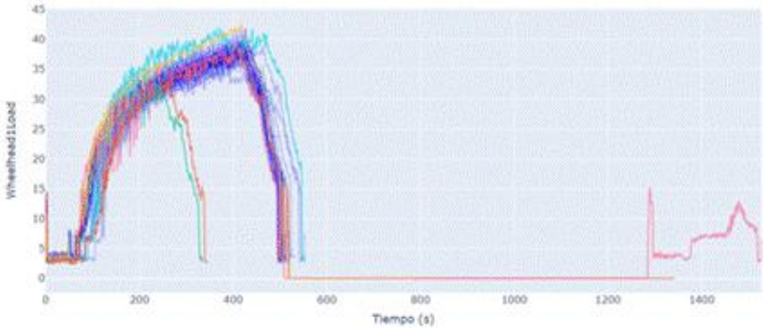

Figure 5: Model testing for time-series anomaly detection.

Objective 3: Detect anomalies online while operations are running

The third objective, currently in its conceptual phase, focuses on enabling near real-time anomaly detection during machining operations. The proposed approach involves segmenting each operation into smaller blocks, allowing the system to evaluate each segment immediately after completion (Figure 6). By applying machine learning techniques to these segments, the system can identify deviations from expected behaviour and notify the operator as early as possible. This capability would significantly reduce response times and minimize the impact of anomalies on production quality and operational efficiency.

Operation X	Machine learning model
Segment 1	✓
Segment 2	✓
Segment 3	✓
Segment 4	✖
Segment 5	
Segment N	

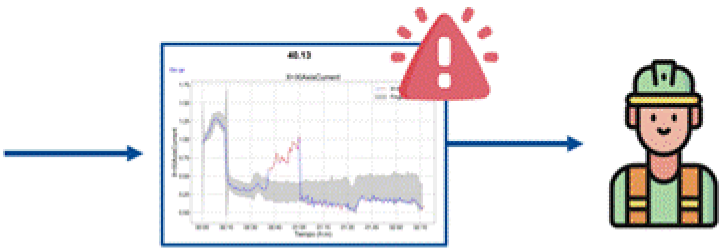

Figure 6. Conceptual segmentation of machining operations for real-time anomaly detection.

In summary, the progress described above represents the technical development of the three objectives focused on advancing anomaly detection in robotic machining processes. In parallel, efforts are being made to integrate this solution within the EMPYREAN architecture, as addressing the identified key challenges would not be feasible without such integration. The following section outlines the initial steps taken toward embedding our anomaly detection system into the EMPYREAN framework.

3.3 Integration with the EMPYREAN components

The first step toward integrating this use case into the EMPYREAN architecture focuses on an initial configuration centred on a single robot. This approach is intended to validate the core architecture and its components. The robot is equipped with a deep edge device , which may vary in CPU, memory, and storage capacity. Although these devices do not include dedicated graphical processing units (GPUs), they are capable of running containerized applications. This enables the deployment of EMPYREAN components close to the data source, reducing latency and improving responsiveness. Tasks requiring greater computational power are handled by far edge resources located on the client's premises, which support more demanding operations such as real-time analytics and advanced data processing (see Figure 6).

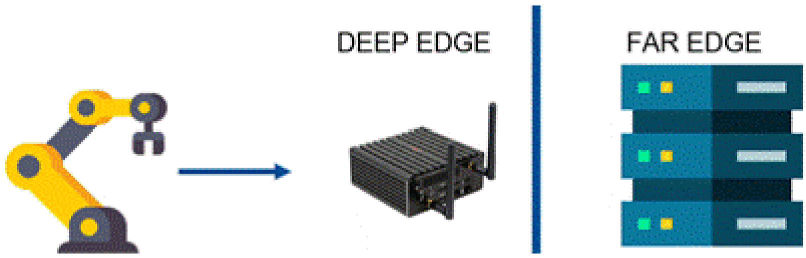

Figure 6: The robot is equipped with a deep edge device and supported by a far edge node for high-computation tasks.

As part of the transition to this architecture, the current application developed by IDEKO for anomaly detection in robotic machining cells (Figure 2) is being restructured to align with EMPYREAN's distributed and modular design. To facilitate this integration, it has been agreed to decompose the solution into three distinct RYAX workflows, each addressing a specific objective. Although these workflows are still in the design phase, their definition provides a clear roadmap for aligning the system with EMPYREAN's capabilities and ensuring future scalability and maintainability. The three planned workflows are as follows (Figure 7).

Workflow 1: Data Recording System (DRS)

This workflow is designed to operate across both the deep edge and far edge layers. The DRS will be deployed at the deep edge, where it will retrieve high-frequency data from on-premises edge devices and IoT sensors. It will compute a set of indicators for first-level processing and then forward the processed data to a message queue backed by a storage system located either at the far edge or within IDEKO's production infrastructure. This setup is intended to ensure that relevant data is captured and pre-processed as close to the source as possible, minimizing data transfer volumes and enabling faster response times.

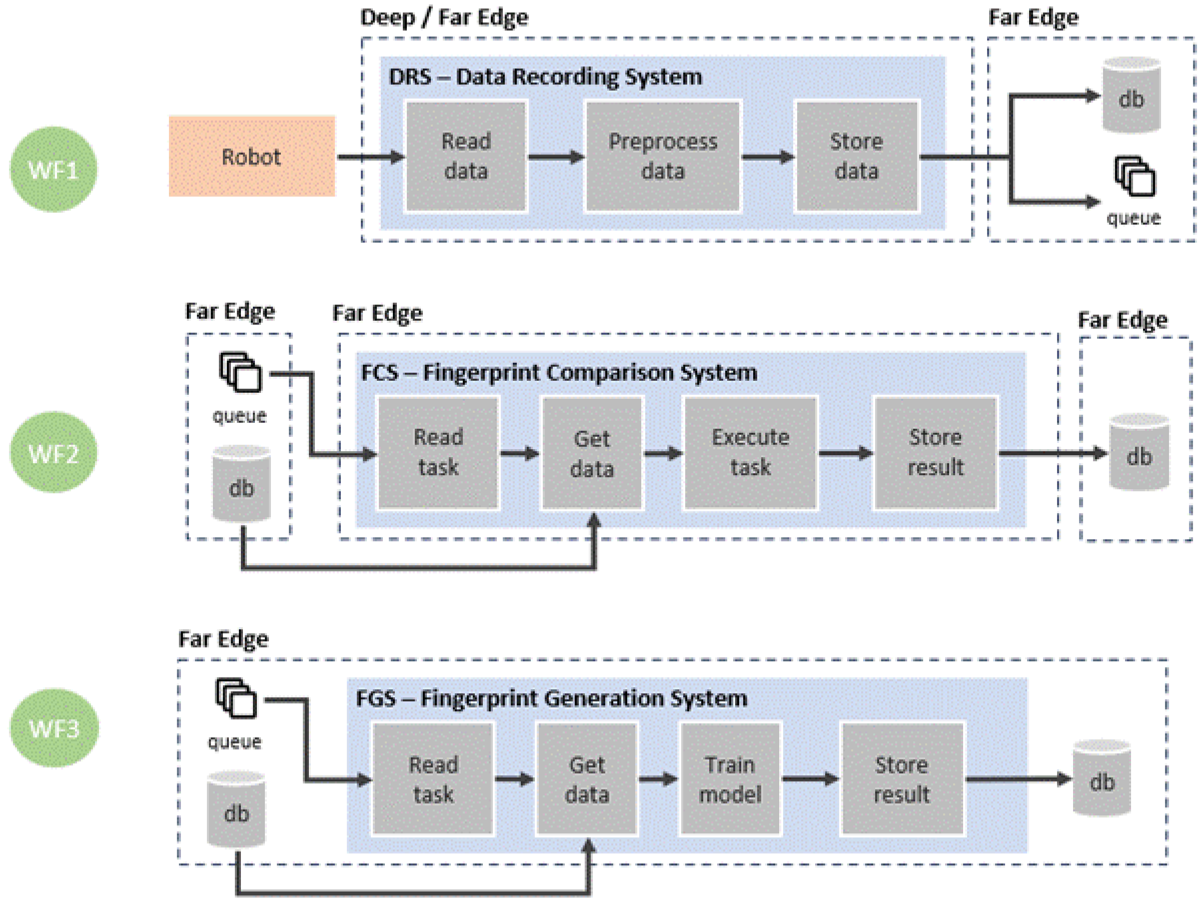

Figure 7: Planned decomposition of the anomaly detection solution into three RYAX workflows for integration with the EMPYREAN architecture.

Workflow 2: Fingerprint Comparison System (FCS)

Designed to run primarily at the far edge, this workflow will be triggered upon the arrival of new data in the message queue. It will execute the FCS, which will initially use statistical simulations to compare the incoming data against existing patterns. In future iterations, this component is expected to incorporate machine learning inference capabilities. The workflow will also trigger online alerts when anomalies are detected and store the results in the database for further analysis. This workflow is essential for enabling near real-time monitoring and decision-making.

Workflow 3: Fingerprint Generation System (FGS)

This workflow is intended to run periodically on the most powerful computational resources available at the far edge. It will retrieve historical data from the database and execute the FGS to generate new patterns through machine learning training. These models will then be stored and used by Workflow 2 to enhance the detection accuracy and robustness of the anomaly detection process. This workflow supports continuous learning and system adaptation to evolving production conditions.

While it is designed to run primarily at the far edge, its deployment may vary depending on the available hardware resources. GPU acceleration is not strictly required in the initial version, but may be beneficial for future iterations involving machine learning inference and training. Intelligent scaling and orchestration will be supported through the Ryax platform, ensuring adaptability to different execution contexts.

Once the workflows were defined, the RYAX Workflow Engine Worker (RYAX engine) was deployed on Kubernetes clusters running on edge devices, both at the deep edge and the far edge (Figure 8). This setup enables RYAX to orchestrate and manage workflows across the entire edge layer, control executions, and maintain communication with the main RYAX server. By leveraging Kubernetes, the system gains flexibility, scalability, and resilience, supporting dynamic resource allocation and fault-tolerant execution of containerized services. This deployment represents a foundational step toward the integration with the EMPYREAN architecture, ensuring support for distributed processing.

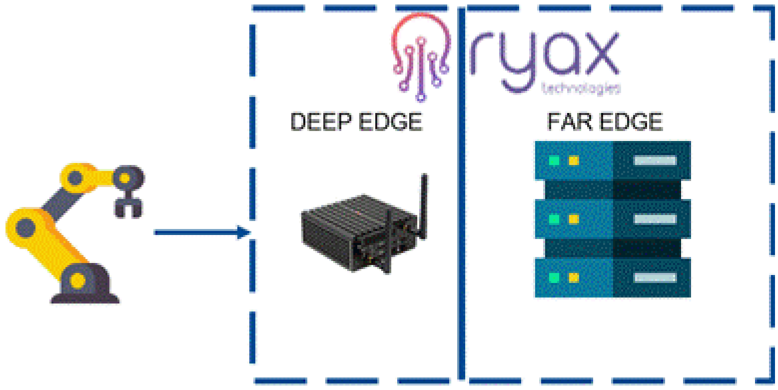

Figure 8: RYAX engine components deployed on Kubernetes clusters at edge devices

Looking ahead, the RYAX engine is expected to serve as the execution environment for the three workflows comprising the anomaly detection solution (Figure 9). Each workflow will be deployed as a modular, containerized service, enabling independent updates, monitoring, and scaling. This modularity is essential for maintaining system robustness and facilitating future enhancements, such as the integration of new data sources or advanced analytics capabilities.

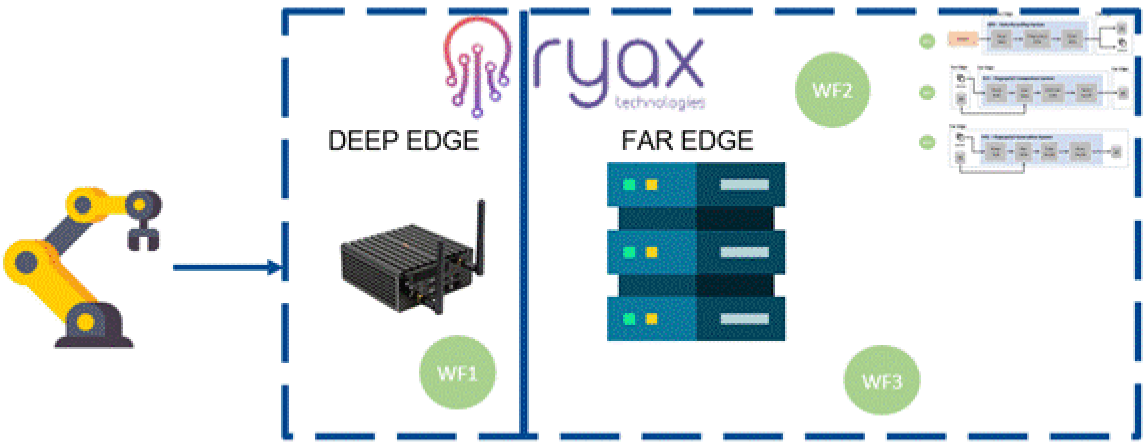

Figure 9: Distribution of the three workflows within the RYAX engine.

To initiate this process, development began with Workflow 1 (DRS). A dedicated Docker image was created for this workflow, encapsulating the logic for data acquisition and pre-processing. This image was then prepared for deployment within the RYAX engine, allowing the system to begin testing the first stage of the anomaly detection pipeline in a real-world edge computing environment. This milestone represents the first concrete step toward full deployment of the solution within the EMPYREAN ecosystem and lays the groundwork for the integration of the remaining workflows.

With the integration of Workflow 1 completed, the next step focused on validating its functionality through a dedicated testing phase, as described in the following section 3.5.

3.4 Stakeholders engagement

As part of the stakeholder engagement strategy, initial contact has been established with key actors within the robotic machining ecosystem, including machine manufacturers and end users of robotic machining cells. Each of these stakeholder groups plays a critical role in shaping the direction and relevance of the use case.

Machine manufacturers provide essential insights regarding the integration of robotic systems into machining environments, contributing to the technical feasibility and industrial compatibility of the solutions under development. Meanwhile, end users offer valuable and practical perspectives on usability, operational requirements, and real-world challenges, which are crucial for aligning project outcomes with actual industry needs.

Engagement with these stakeholders is ongoing and will continue throughout the project's lifecycle. Their input is expected to inform key decisions, guide technical development, and support the co-creation of solutions that are both innovative and applicable in real-world robotic manufacturing environments.

3.5 Testing and evaluation

Following the integration of Workflow 1 (DRS) into the edge devices within the RYAX engine cluster, the testing phase was initiated. To simulate continuous machining activity, a robot simulator was developed using a real dataset collected from an industrial robotic cell. This simulator was configured to send low-frequency data streams to Workflow 1, enabling the DRS to read, process, and store the data in the target database (Figure 10).

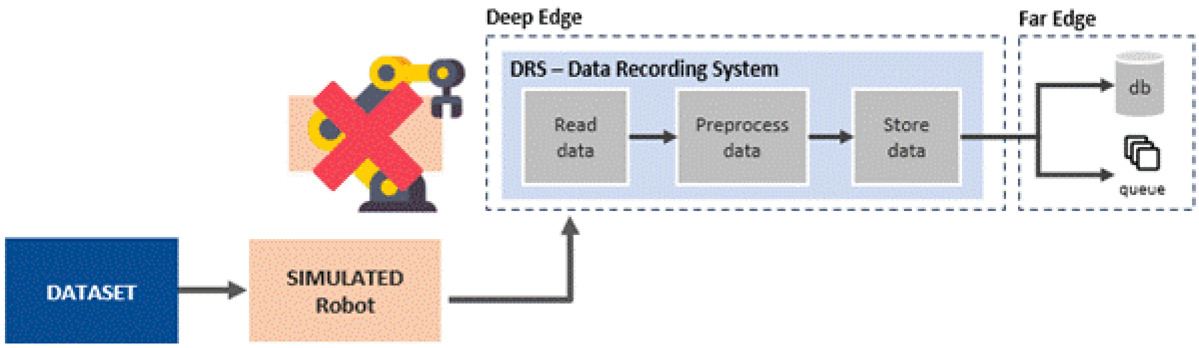

Figure 10: Using a real robot dataset, a simulator was developed to send continuous data to Workflow 1 (DRS).

This setup enabled a full end-to-end test of Workflow 1 within the RYAX engine, which orchestrates distributed workloads and workflows in the EMPYREAN architecture. The test environment consisted of three main components: the robot simulator, the DRS, and a MongoDB database. The evaluation involved sending data from IDEKO's infrastructure to the DRS deployed on the edge cluster, where it was processed and stored (Figure 11).

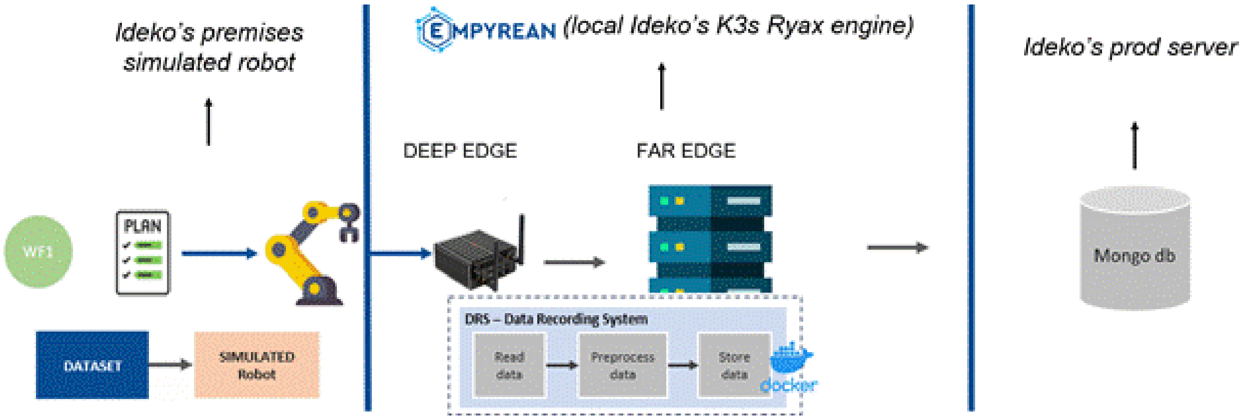

Figure 11: Initial setup for testing and evaluating the integration of Workflow 1 into the EMPYREAN architecture.

The initial results were positive. The DRS application was successfully deployed on demand across IDEKO's edge infrastructure, including both deep edge and far edge devices through the RYAX engine. System logs (Figure 12) confirmed that the DRS continuously received data from the simulator, processed it correctly, and stored it in the database, thus validating the intended functionality of Workflow 1.

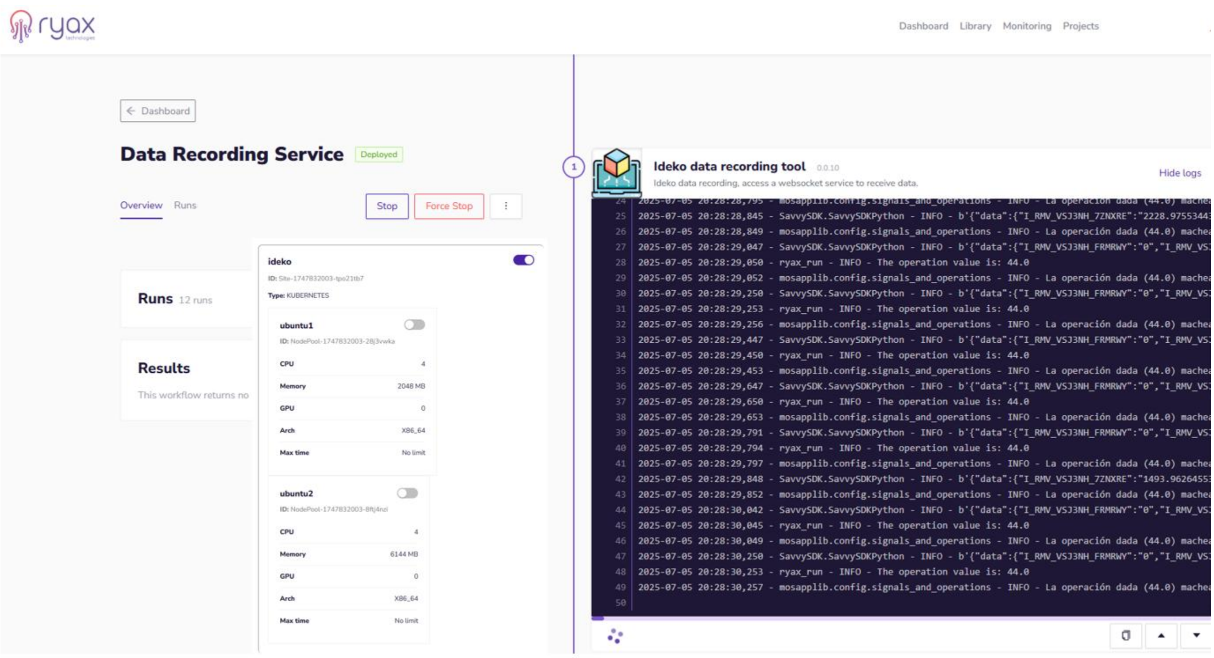

Figure 12: RYAX interface used to deploy the DRS on IDEKO's edge devices.

This testing phase not only confirmed the technical feasibility of deploying the DRS within the EMPYREAN framework but also demonstrated the effectiveness of the RYAX engine in managing edge-based workloads. It marks a significant milestone in the progressive integration of the anomaly detection solution and lays the foundation for testing the remaining workflows in subsequent phases.

3.6 KPI progress

In deliverable D2.1, we have defined three KPIs and respective metrics to track the progress of the use case development.

No	Indicator	Success Criteria
1	Transition from offline operation analysis to real-time operation fingerprint analysis	-
2	Ability to process real-time data	3 robots / 200 operations
3	Ability to alert an abnormal operation	máx. 2 sec after it occurs

1. **Transition from offline operation analysis to real-time operation fingerprint analysis**

Anomaly detection is currently performed offline, after operations are completed. The goal is to enable real-time fingerprint analysis during machining. To support this transition, operations are being segmented into meaningful phases, allowing each segment to be assessed immediately after completion. This enables near real-time anomaly detection and early alerting.

2. **Ability to process real-time data**

This KPI is being addressed by adapting data monitoring devices and defining new data workflows. The solution has initially been structured into distinct workflows to ensure low latency and high-volume data processing. The target is to support 3 robots and 200 operations, maintaining data integrity and responsiveness throughout the system.

3. **Ability to alert an abnormal operation**

To meet this KPI, machine learning models for time-series comparison are being analysed and tested. These models aim to improve detection accuracy and adaptability to varying machining conditions. The system must be capable of issuing alerts within a maximum of 2 seconds after an anomaly is detected, enabling timely intervention before part quality is affected.

3.7 Next Iteration

Building on the successful deployment and validation of Workflow 1 (DRS), the next iteration of development will focus on adapting and extending all three workflows, DRS, FCS, and FGS, to meet the internal objectives of the use case and ensure full integration within the EMPYREAN architecture through the RYAX engine.

The primary goal of this phase is to modify, containerize, and optimize the workflows for efficient deployment as modular services across IDEKO's multi-layer edge infrastructure. These modifications are essential to support the following enhancements:

- ***Integration of mid-high frequency data***: Workflow 1 will be extended to handle higher-resolution data streams. This requires updating the data acquisition logic and optimizing preprocessing steps to maintain performance within the constraints of edge devices.
- ***Machine learning-based inference and training***: Workflows 2 and 3 will be adapted to incorporate machine learning models for anomaly detection. Workflow 2 (FCS) will include inference capabilities for real-time decision-making, while Workflow 3 (FGS) will support periodic training using historical data to improve model accuracy (Figure 13).
- ***Workflow decomposition and modularization***: All workflows will be further divided into smaller, more manageable components to improve maintainability and scalability. This will also facilitate more granular deployment and monitoring within the RYAX engine.
- ***Containerization using Docker***: Each workflow will be packaged as a Docker container to ensure portability, reproducibility, and seamless orchestration by the RYAX engine. This step is critical for enabling deployment across heterogeneous edge environments.

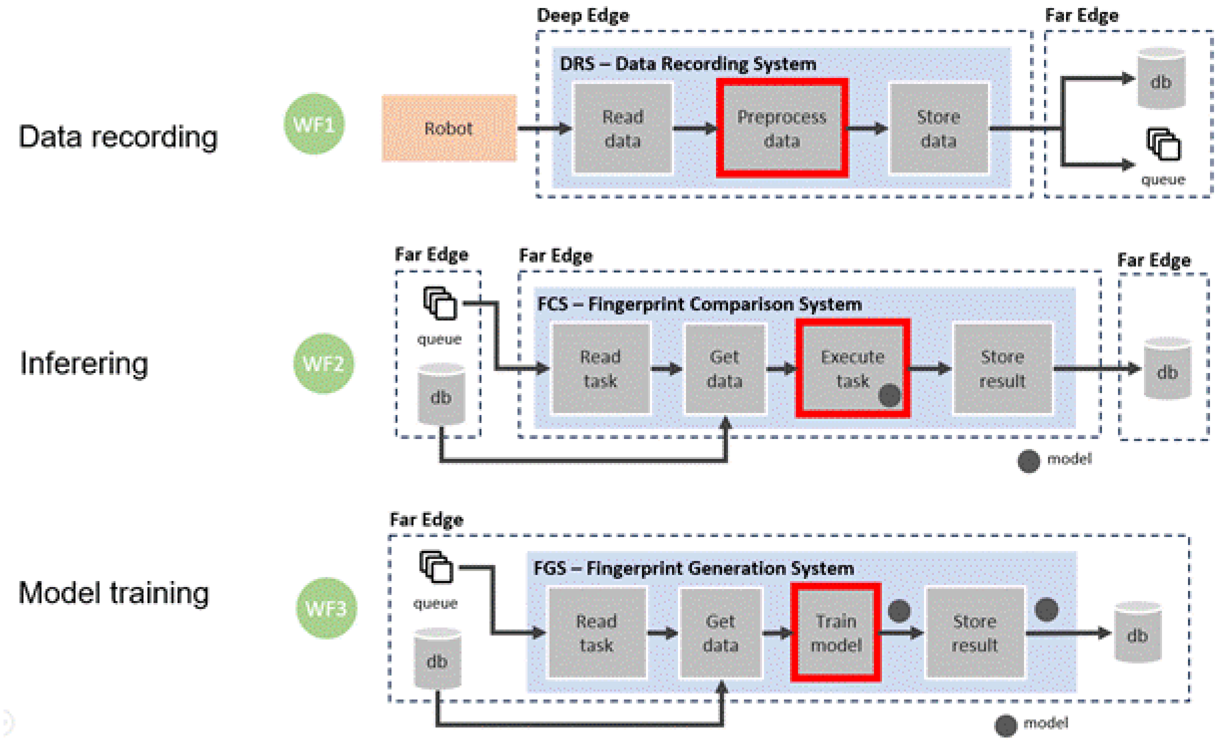

Figure 13: Workflow modifications to support internal objectives and enable deployment via the RYAX engine.

These enhancements will allow the anomaly detection system to evolve into a fully distributed, intelligent monitoring solution capable of operating in near real time across IDEKO's robotic machining cells. The modular and containerized design will also ensure that the system remains adaptable to future requirements, including the integration of new data sources or edge hardware.

The upcoming development cycle will begin with the restructuring and containerization of Workflow 2 (FCS), followed by Workflow 3 (FGS). Once completed, all three workflows will be orchestrated through the RYAX engine and deployed across the deep and far edge layers, completing their integration into the EMPYREAN ecosystem.

4 Proximal Sensing in Agriculture Fields Technological Developments

This use case focuses on the dynamic assessment of Soil Organic Carbon (SOC) to evaluate and manage soil conditions in agricultural fields. By combining proximal sensing technology with edge computing, the system enables real-time SOC assessment without relying on centralized data processing. This innovative approach supports integrative farm management and sustainable agricultural practices by providing timely and actionable insights for soil health.

4.1 Overview

The use case is structured in two phases. In Phase 1, a UAV surveys an agricultural field to assess soil organic carbon (SOC) levels. The output is a map that divides the field into distinct zones based on SOC variability. In each zone, several sample points are selected for further investigation. Phase 2 involves detailed data collection at these points using an autonomous robot equipped with proximal sensing technology. The data gathered from both phases is then integrated and presented to the end user via a web-based dashboard.

4.1.1 UAV

The UAV workflows are designed to estimate SOC levels through remote sensing techniques. A UAV equipped with a multispectral camera captures high-resolution imagery across multiple distinct spectral bands during field surveys. These images form the basis for SOC classification and are processed through a series of advanced steps, including image stitching, reflectance calibration, and the generation of management zones.

To enable efficient and timely data processing, portions of the workflows can be deployed directly on the drone’s onboard computing system. This deep edge computing capability allows for real-time data compression, storage, and preliminary analysis directly on the UAV, minimizing latency and reducing the volume of data transmitted for further processing. Such an approach optimizes the overall data pipeline and enables faster feedback loops.

The processed imagery is used to develop an SOC classification model, which categorizes field areas into low, medium, or high SOC levels. These classifications are then translated into management zones that support precision agriculture decision-making. Additionally, these zones can guide more targeted and detailed soil analysis using proximal sensing technologies mounted on the ILVO robotic platform.

<u>Hardware set-up</u>

The platform used to support the UAV SOC assessment workflows is the DJI Matrice 350 RTK (M350 RTK), a professional-grade drone engineered for high-precision applications such as mapping, surveying, and infrastructure inspection. Known for its robust payload capacity, extended flight time, and centimeter-level positioning accuracy via RTK GNSS, the M350 RTK also offers features such as obstacle avoidance and compatibility with a variety of advanced payloads. These capabilities make it well-suited for demanding remote sensing tasks in agricultural settings.

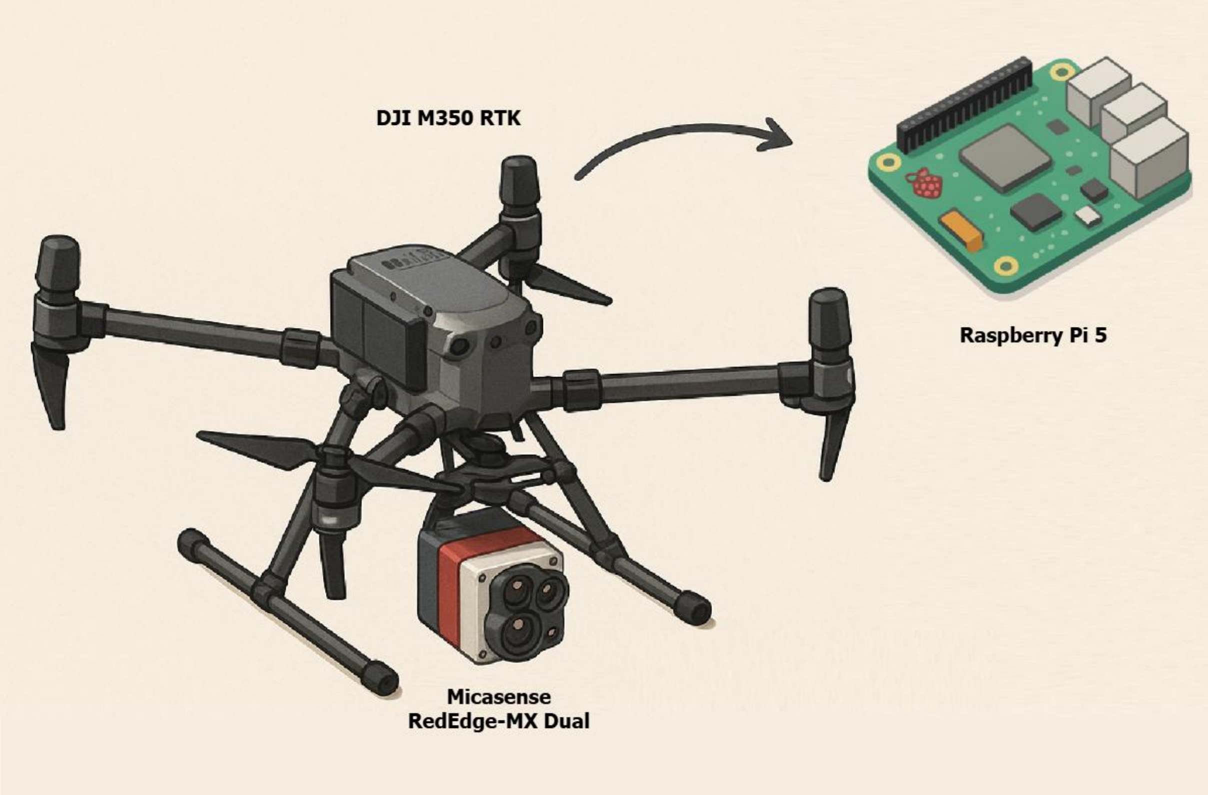

Figure 14: Hardware set-up for the UAV workflow in UC2

To support advanced soil analysis, the following components will be integrated with the UAV:

1. **Raspberry Pi 5**: A Raspberry Pi 5 will act as the onboard computing unit and central controller. It handles communication between the UAV and the multispectral camera, manages data logging, and performs preliminary pre-processing of the imagery. Data collected during flight missions is wirelessly transmitted to a cloud infrastructure for further analysis and long-term storage.
2. **Multispectral Camera**: A MicaSense RedEdge-MX Dual camera system will be mounted on the underside of the UAV. This dual-sensor system captures synchronized images across 10 spectral bands, covering the 400 nm to 900 nm range. It is specifically designed for advanced agricultural and remote sensing applications, including soil and vegetation analysis. The camera exposes an HTTP-based interface for accessing imagery and metadata, which the Raspberry Pi can retrieve.
3. **DJI E-Port Development Kit**: The DJI E-Port development kit serves as the hardware interface between the drone and external payloads. It converts the proprietary E-Port output into standard hardware interfaces, supplying power to the Raspberry Pi and enabling data exchange. Through the PSDK (Payload Software Development Kit) API, the Raspberry Pi can also access drone telemetry, such as GPS position, battery status, and flight statistics, further enriching the dataset with contextual information.

Installation

To protect the Raspberry Pi, control electronics, and the DJI E-Port development kit during flight, a custom enclosure was designed to securely house all these components. The case was 3D printed, so it is very lightweight. Internal compartments or mounts ensure that each component is held firmly in place, preventing movement or disconnection during flight.

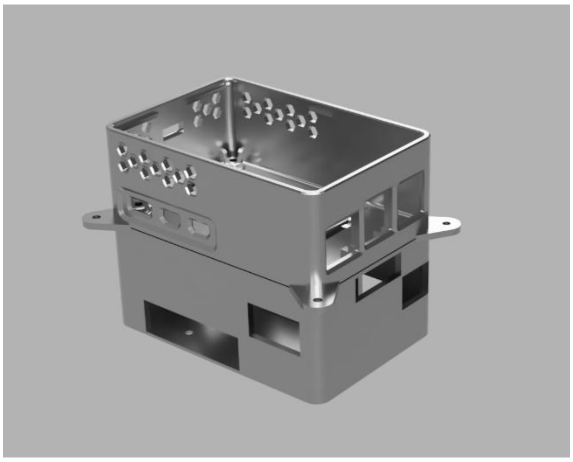

Figure 15: Casing for the Raspberry Pi and control electronics

The enclosure is designed for ease of integration with the DJI M350 RTK drone and can be mounted directly onto the frame using the pre-existing mounting screw holes located on the top of the drone. This ensures a stable and vibration-resistant connection.

Figure 16: Installation of the Raspberry Pi on the UAV

Furthermore, the design takes into account airflow and heat dissipation requirements, especially for the Raspberry Pi and power modules, which can generate significant heat during operation. Ventilation slots or heat sinks may be included to prevent thermal throttling and ensure stable system performance throughout the duration of a mission.

4.1.2 Robot

The ILVO CIMAT robot conducts complementary ground operations, focusing on detailed SOC analysis at designated points of interest. Equipped with a portable spectrometer and a moisture sensor, the robot gathers highly precise soil data. This data contributes to the generation of a comprehensive and accurate SOC map.

The SOC Prediction Model (Robot) relies on data from a portable spectrometer and a soil moisture sensor mounted on the robot. These devices capture precise measurements at specific points of interest. The model's objective is to provide highly accurate SOC values for particular locations within the field, offering granular insights and serving as a benchmark for validating and refining the classification model. For training, soil samples will be collected from ILVO fields and undergo detailed spectrometer analysis and laboratory testing to determine their SOC levels and other relevant properties. The spectrometer readings are correlated with the lab-analysed SOC values to train the model, with environmental factors such as soil moisture levels included to ensure robustness under varying field conditions. After training, the prediction model will operate on the robot's computing infrastructure at the deep edge, enabling real-time SOC predictions during field operations.

Hardware set-up

The robot platform used for this project is the ILVO CIMAT robot, an autonomously driven, multifunctional unit designed to support a wide range of agricultural tools and sensors. It can be equipped with various standard implements, as well as cameras and other instruments for data acquisition.

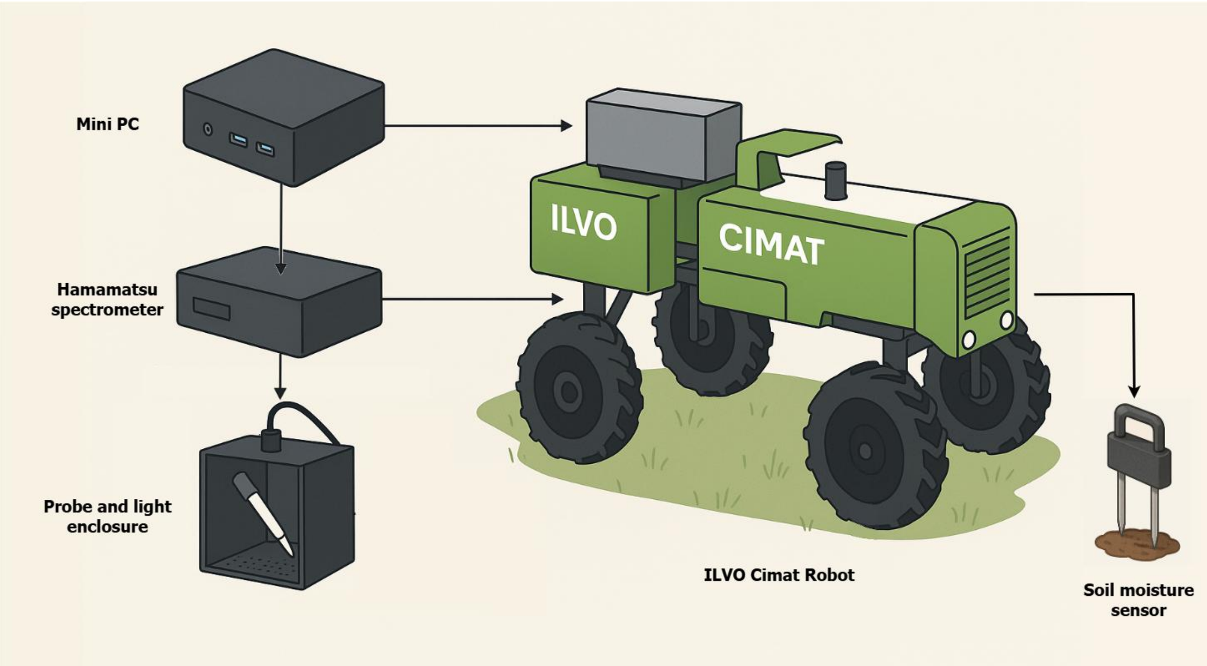

Figure 17: Hardware set-up for the robot workflow in UC2

For the purpose of soil analysis, several components will be mounted on the robot:

1. **Mini PC:** A compact PC will serve as the central controller for the system. It will manage communication between the robot, the spectrometer, and the soil moisture sensor, while also handling data logging and pre-processing. Collected data will be transmitted to a cloud-based infrastructure via a wireless uplink for further analysis and storage.
2. **Spectrometer:** A Hamamatsu C11118GA spectrometer will be used to capture the reflectance of the soil. With a spectral range spanning from 850 nm to 2550 nm, this device is well-suited for detecting features relevant to soil organic carbon assessment, making it a key component of the sensing system.
3. **Probe enclosure:** An enclosed housing will be mounted at the rear of the robot to contain the spectrometer probe and an integrated light source. The enclosure is essential to block ambient light from entering the optical path, ensuring that reflectance measurements are not affected by external factors.
4. **Soil moisture sensor:** The ML3 ThetaProbe soil moisture sensor will be used to measure the water content of the soil. Since soil moisture significantly influences spectral reflectance, these measurements will be used to calibrate and correct the spectrometer data, improving the reliability of soil organic carbon estimations.

4.2 Initial developments

This section outlines the initial technical developments carried out to support the use case. The primary focus was on developing the individual building blocks that will serve as the foundation for future iterations. These components will later be integrated into a complete workflow aligned with the Empyrean framework.

4.2.1 Drone integration

To enable seamless integration between our edge node and the UAV, we utilize the DJI E-Port Development Kit. This development kit transforms the proprietary E-Port of the DJI drone into several standardized hardware interfaces, including the XT30 Power Output Ports, USB 2.0 Port, and UART/PPS Signal Port. This flexibility significantly simplifies the development and debugging process, as it allows various onboard modules to interface directly with the drone's ecosystem. The XT30 power output from the E-Port is particularly important in our setup, as it supplies power to the Raspberry Pi 5 and additional peripherals such as the AI accelerator module. This eliminates the need for separate power sources, reducing overall system complexity.

DJI also provides a Payload Software Development Kit (PSDK), which is used to interface custom payloads with the UAV’s onboard systems. The PSDK offers a range of APIs that enable the payload device (in our case the Raspberry Pi) to interact with and retrieve data from the drone. The information management function of PSDK includes information acquisition and message subscription functions. The edge device developed based on PSDK can actively obtain information such as drone model, hardware platform type, power management, camera functionality, flight control, etc. Various components on the drone will generate a large amount of real-time data according to the actual flight status of the drone and will be pushed to other modules by the drone. The user can specify the data information to be subscribed to by using the payload device with the message subscription function.

Energy monitoring

Energy monitoring for the UAV platform is achieved through two distinct but complementary data sources, allowing for both high-level and component-level insights. The first source is the DJI Payload SDK (PSDK)[2], which provides system-wide telemetry data. This includes the remaining battery capacity, overall voltage and current, as well as detailed per-cell voltage, current and capacity information. These measurements reflect the total energy consumption of the entire UAV system, which includes not only the drone itself and its propulsion system, but also the onboard edge device and the AI accelerator.

Figure 18: First installation of the USB power consumption meter on the UAV

In addition to the PSDK data, a second energy monitoring mechanism was added by integrating a USB power consumption meter (Fnirsi FNB58). This meter is placed in-line with the power supply to the edge computing infrastructure, which consists of the Raspberry Pi and the AI accelerator. It allows for real-time measurement of voltage and current specific to these components. By isolating the energy usage of the edge device and its associated hardware, we gain a clearer understanding of how much power is consumed solely by the computational tasks running onboard. This distinction is particularly useful for optimizing edge workloads and quantifying whether the edge computing has any impact on flight duration.

An initial version of the energy consumption acquisition script has been developed and deployed to the following Git repository: https://gitlab.ilvo.be/empyrean/energy-monitor/ . This script is responsible for collecting power-related data from the USB power consumption

[2] https://github.com/dji-sdk/Payload-SDK

meter. The data is intended to be stored in a structured database environment to support post-flight analysis and long-term monitoring. For initial testing and development purposes, the script currently writes data to a timeseries database, specifically InfluxDB[3]. This choice allows for efficient handling of high-frequency, timestamped data streams and offers compatibility with common dashboarding tools for visualization and inspection. In future iterations, energy consumption information will be offloaded using the Empyrean telemetry component.

Radiometric calibration

The raw values captured with the MicaSense multispectral camera must first be calibrated to ensure their usability in further analysis. Radiometric calibration converts the raw pixel values into absolute spectral radiance values, correcting for several factors such as the sensor's black-level offset, gain, exposure time, and lens vignetting. This step is essential for producing consistent and comparable data across different flights and lighting conditions.

In a typical workflow, the calibration process begins with capturing an image of a reflectance calibration panel. These panels have known reflectance values, which are used to estimate the incident irradiance based on the measured radiance in the image. The calculated irradiance is then used to correct the images in subsequent captures, in combination with measurements from the Downwelling Light Sensor (DLS) integrated into the MicaSense system. This approach compensates for differences in ambient lighting between the time the panel image and the flight images are taken.

Because each spectral band is captured by a separate sensor in the camera, slight geometric misalignments between the bands must also be corrected. Image alignment is performed by estimating homography transformations that align all bands to a common reference. This ensures that pixel-level correspondence across bands is maintained, which is critical for any pixel-wise analysis.

This workflow ensures that the captured data are both radiometrically accurate and geometrically consistent, forming a robust foundation for reliable multispectral analysis. The Ryax workflow code is available at https://gitlab.ilvo.be/empyrean/ryax-workflows.

Stitching

To prepare the multispectral data captured by the MicaSense sensor for further analysis, the individual geotagged calibrated images must be stitched into a consistent orthomosaic. This process is essential for creating a spatially coherent dataset where each pixel corresponds to a real-world location and is accurately aligned across spectral bands.

The stitching is performed using an automated workflow built on ClusterODM, an API layer that orchestrates processing tasks using OpenDroneMap (ODM). The process begins by collecting all relevant image files from the data capture phase. A task is then submitted to the ClusterODM node with predefined settings optimized for the specific characteristics of the MicaSense sensor imagery. Once processing is completed, the resulting assets, typically

[3] https://github.com/influxdata/influxdb

including the orthophoto and associated metadata, are uploaded to the Skyflok storage server.

The Ryax workflow code is available at https://gitlab.ilvo.be/empyrean/ryax-workflows.

4.2.2 Robot integration

To interact with the ILVO CIMAT robot, we utilize the Agricultural Robot Taskmap Operation Framework (ARTOF) [4], which was developed at ILVO. ARTOF provides the common functionality of task map execution for a wide range of robot configurations and farming applications based on Global Navigation Satellite System (GNSS) positioning. The ARTOF framework is composed of two layers. The mechatronic layer provides motion control and machine safety and the operational layer enables the autonomous task map execution. The add-ons provide additional functionality, behind the scope of the core functionality. The ARTOF Redis API maintains the communication between the operational layer, the add-ons and other network devices or implements. The framework architecture is shown in the figure below.

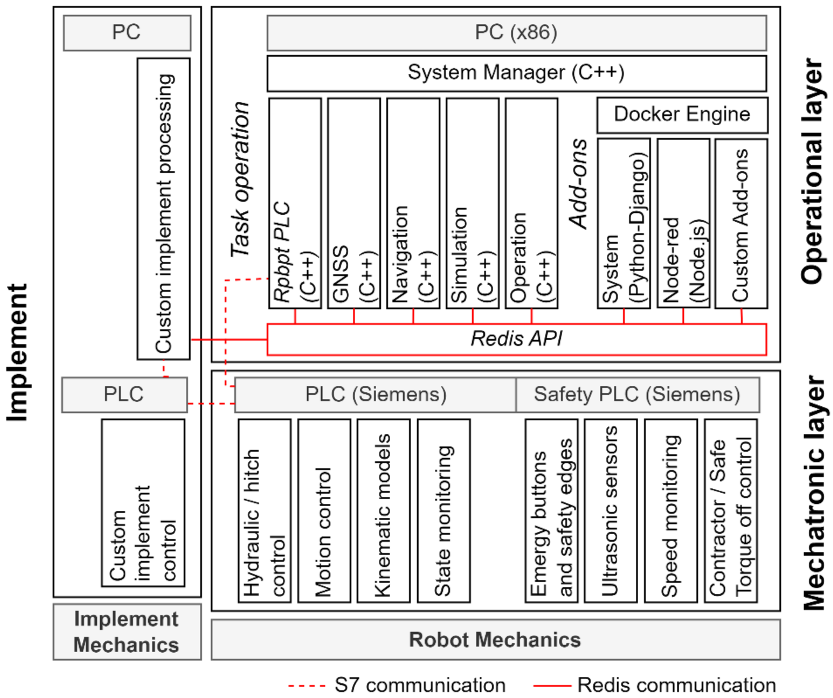

Figure 19: ARTOF framework architecture

The EMPYREAN mini-PC, which will be installed on the CIMAT robot, functions as an auxiliary processing and control unit. It extends the robot's capabilities by handling on-board data collection and processing related to soil condition monitoring. The mini-PC interacts directly with the ARTOF framework using a Redis-based interface, enabling seamless integration and real-time data exchange with the robot's core systems.

[4] https://artof-ilvo.github.io/

Figure 20: ILVO CIMAT robot operating in one of ILVO's fields on bare soil

The robot's operational path is generated from a shapefile that can be uploaded through a dedicated web platform. This shapefile contains a sequence of geo-referenced points, which are defined based on insights derived from a prior drone-based survey of the field. Each point corresponds to a targeted sampling location where the robot is programmed to stop and perform measurements using the Hamamatsu spectrometer and the ThetaProbe soil moisture sensor.

Figure 21: Example drone path created on an ILVO field

The ARTOF framework continuously pushes a wide array of monitoring data to a Redis database. For the purposes of the EMPYREAN project, two of the most critical parameters are GNSS location and the battery capacity. This information is available through Redis using standard key-value access. In addition to basic telemetry, ARTOF provides state information for various robot components and its attached tools. These state updates are represented internally using 4×4 affine transformation matrices, which allow for efficient computation of positions and orientations in 3D space.

To validate and demonstrate this integration, a sample application was developed that periodically polls the Redis database to retrieve real-time monitoring values. The application collects various metrics such as:

- Power output to each of the robot's wheels
- Velocity and linear acceleration
- Angular velocity and rotational orientation
- GNSS position and estimated state of platform components

The source code has been made available on GitLab at:
https://gitlab.ilvo.be/empyrean/cimat

```
Connection to redis server host: robotframework:6379 done!
Monitoring values:
{'plc.monitor.drive_fl.power': '0',
 'plc.monitor.drive_fr.power': '0',
 'plc.monitor.drive_rl.power': '0',
 'plc.monitor.drive_rr.power': '0',
 'plc.monitor.imu.acceleration.x': '0',
 'plc.monitor.imu.acceleration.y': '0',
 'plc.monitor.imu.acceleration.z': '0',
 'plc.monitor.imu.rotation.x': '0',
 'plc.monitor.imu.rotation.y': '0',
 'plc.monitor.imu.rotation.z': '0',
 'plc.monitor.navigation.velocity.angular': '0',
 'plc.monitor.navigation.velocity.lateral': '0',
 'plc.monitor.navigation.velocity.longitudinal': '0'}
---
Robot reference state:
{'R': [0.0, 0.0, 300.5926177266613],
 'R_cov': [[1.0, 0.0, 0.0], [0.0, 1.0, 0.0], [0.0, 0.0, 1.0]],
 'T': [555146.5782715841, 5648555.232477879, 0.0],
 'T_cov': [[1.0, 0.0, 0.0], [0.0, 1.0, 0.0], [0.0, 0.0, 1.0]],
 'point': {'latlng': [50.98594070735032, 3.7856718100950073],
           'xy': [555146.5782715841, 5648555.232477879]}}
```

Figure 22: Sample output of the CIMAT monitoring application

4.2.3 Spectrometer

After careful evaluation of available options, we selected the Hamamatsu C11118GA as the spectrometer for conducting soil organic carbon (SOC) assessments. This device offers a compact and robust design well-suited for integration into mobile platforms such as our robotic system.

One of the key advantages of the Hamamatsu C11118GA is its broad spectral range, spanning from 850 nm to 2550 nm. This range includes the near-infrared (NIR) and shortwave infrared (SWIR) regions that are known to contain key absorption features associated with organic matter and mineral content in soil, making it highly suitable for SOC analysis.

In addition to its hardware capabilities, the spectrometer is supported by sample software that allows users to configure measurement conditions, acquire and save spectral data, and visualize it through built-in graphing tools. This provides a useful starting point for testing and development.

To enable integration with our data collection workflows, the device also comes with a Windows-only Software Development Kit (SDK) that supports several programming languages, including Visual Basic, Visual C++, and Visual C#. This SDK enables the development of custom software to control the spectrometer and retrieve data directly from a mini PC deployed on the robotic platform.

<u>SDK implementation</u>

The Visual C++ version of the SDK was used to evaluate both the functionality of the Hamamatsu C11118GA SDK and the feasibility of integrating it into our software platforms. During this assessment, the SDK performed as expected, allowing successful communication with the spectrometer, data acquisition, and basic configuration of measurement parameters. This confirms that integration into our mini PC-based edge computing workflow is technically feasible and aligns well with our requirements.

```
C:\Users\jbauwens\OneDrive - ILVO\Documenten\Projects\Empyrean\Spectrometer\Minispectrometer_USB2.0 edition_ver1.1.0.0\DeveloperTools\ForVisualCppCLI\SampleApl\bin\x64\Debug\SampleApl.exe
index(250-)  8387  8568  7178  9373 11346  9106
Stop Capture.........
Start Capture.............

index(  0-) 12198  8569  8949 17421 12909 22285 11368 11018  6097 10049
index( 10-)  8589  9233  9409  6593  7376 10918  6994  9011  7810  8200
index( 20-)  8158  9394  9121 11991 15770 13114 10605 11285 27181  7779
index( 30-) 31541 13083 12611  8782  9216  9071  8503  8837  8319  7707
index( 40-)  9951 26805  8690  8881  7884 30776  8560 12597  8606  5057
index( 50-)  8948 13286  6752  9418  8154  7223 10859 10137 13307  8356
index( 60-) 12528 11214  8482  8142  7446 10028 15045  7579  8766  7370
index( 70-) 10850  9260  8915  6187 10934  6059 11735  5510  8727  7196
index( 80-) 10881  8850  6341  7161  9276  8379  5372  9296  8118  5545
index( 90-)  8664  9756  7957  7559  9753  6225  7925  9920  7089 11075
index(100-)  8455  9130  8086 10184 10062  6419  7337 10249  8460 10460
index(110-)  6097  9131  8013  8038  8515  9078  9988  8528  9244  9601
index(120-) 10407 10057  7945  9661  6858 10578  6630 14347  3892  9403
index(130-)  7832  9948  7156  8062  6995 10819  7228  8509  7924  9701
index(140-) 11065  5699 10453  7666  7683  8916 10440  6914 10478  9523
index(150-)  9698  6825  7285 10339  9027 10536 10416  8309 20010  9103
index(160-)  7955  8830  8517 11838 11270  7388  9839  7159  9077 12683
index(170-)  8936 18090 10215  7565 12537 10007  7668  9307  9367  9956
index(180-)  6661 22884 10618  7575 10512  6852  5770 10659  7477  8103
index(190-) 10890  7152  7213  8260  8730  7450  9043 10613  8832 10946
index(200-) 15941  9654  6888  7603  9299  6577  7324 11195  7405  7886
index(210-)  8647  8824  7707  9323  8137  8683  8978  7925  7122  6640
index(220-)  6804  7031  7027  9442  9232  8759  7374  9959  8474  9203
index(230-)  7642 11305  6716  8554  6424  7988 11301  8235  6806  9221
index(240-)  8294  6433  9544  8480  4378 12232  5351  9188  7750  9159
index(250-)  8378  8574  7165  9370 11316  9119
Stop Capture.........
Stop Capture.........
Releasing Buffer......
End of measurement
>
>
```

Figure 23: Output of a sample application written with the Visual C++ SDK

The final implementation will wrap the SDK functions in a higher-level API for communication with other modules. An initial implementation has been made available in the following repository: https://gitlab.ilvo.be/empyrean/hamamatsu.

Testing

To verify that the spectrometer gives the expected outputs, tests were performed with a demo kit provided by Hamamatsu. Only after this verification, we made the decision to order this spectrometer. The tests were performed on a soil sample collected from outside the ILVO building. At the time of collection, the soil was notably dry, due to an extended period of low rainfall in recent months. For the measurements, a 15-watt bulb was used as the light source.

Figure 24: Testing set-up for spectrometer

First, we tried to find out the optimal integration time of the spectrometer, which refers to the duration a spectrometer's detector collects light before recording a spectrum. This is analogous to the exposure time in a camera. A longer integration time allows more light to be collected, which is beneficial for low-intensity light sources or improving the signal-to-noise ratio. A long integration time in a spectrometer can cause saturation, leading to clipped peaks and making the measurement impractical. It's important to find a balance between increasing the signal and avoiding saturation.

The different exposure time settings produced similar results up to approximately 2100 nm. Beyond this wavelength, the 1000 µs setting exhibited noticeably higher noise levels, making it less suitable for reliable measurements (Figure 25). In contrast, the 5000 µs and 7000 µs exposures yielded visibly smoother and cleaner spectra, with reduced noise. To objectively quantify the noise across configurations, the average of the second derivative of each spectrum was calculated. We excluded the 2300 nm to 2550 nm region, as this spectrum is inherently noisy. This metric serves as an indication for signal smoothness: lower values indicate less noise. Based on this analysis, the 5000 µs exposure time showed to be the optimal setting among the five tested.

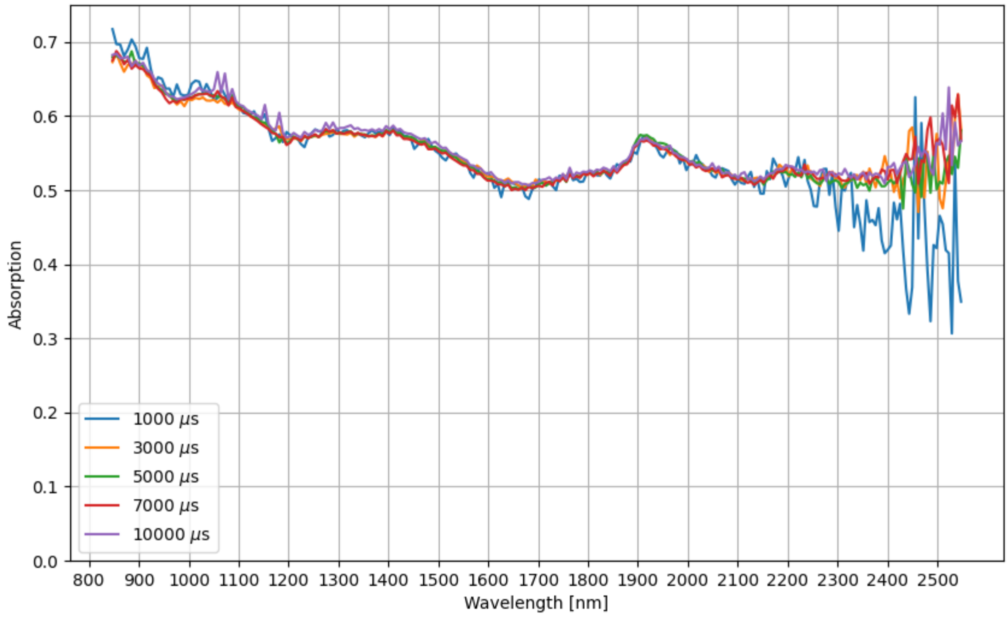

Figure 25: Soil absorption measurements with different exposure times

To further improve the quality of the spectral data, the measurement was repeated 50 times using an exposure time of 5000 µs. By averaging the results over 50 iterations, the noise present in individual scans was significantly reduced (Figure 26). This approach enhanced the clarity of the spectral signal, particularly in the higher wavelength regions where noise was previously more pronounced.

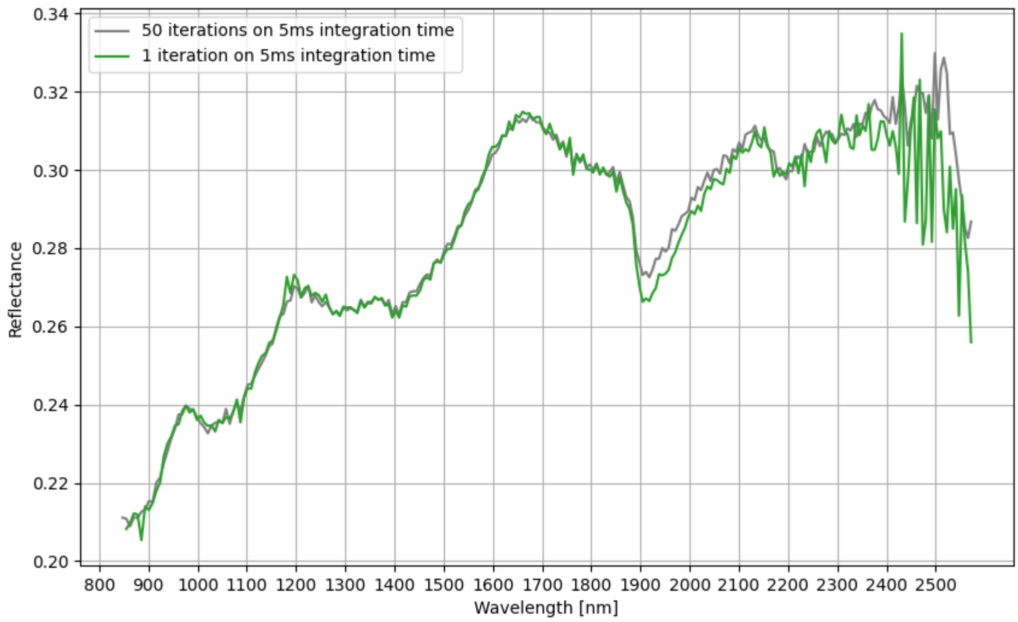

Figure 26: Reflection measurement comparison of 1 iteration and average of 50 iterations

To further evaluate the performance of the Hamamatsu spectrometer, we compared its output to that of the high-quality Spectral Evolution PSR+ spectrometer. It is important to note that the PSR+ measurements were performed on a different soil sample, so a direct one-to-one comparison is not appropriate. Instead, the focus was on identifying whether both spectrometers capture similar spectral patterns, particularly the presence and location of peaks and valleys. As shown in Figure 27, the spectral trends from both devices align closely, with comparable features across the wavelength range.

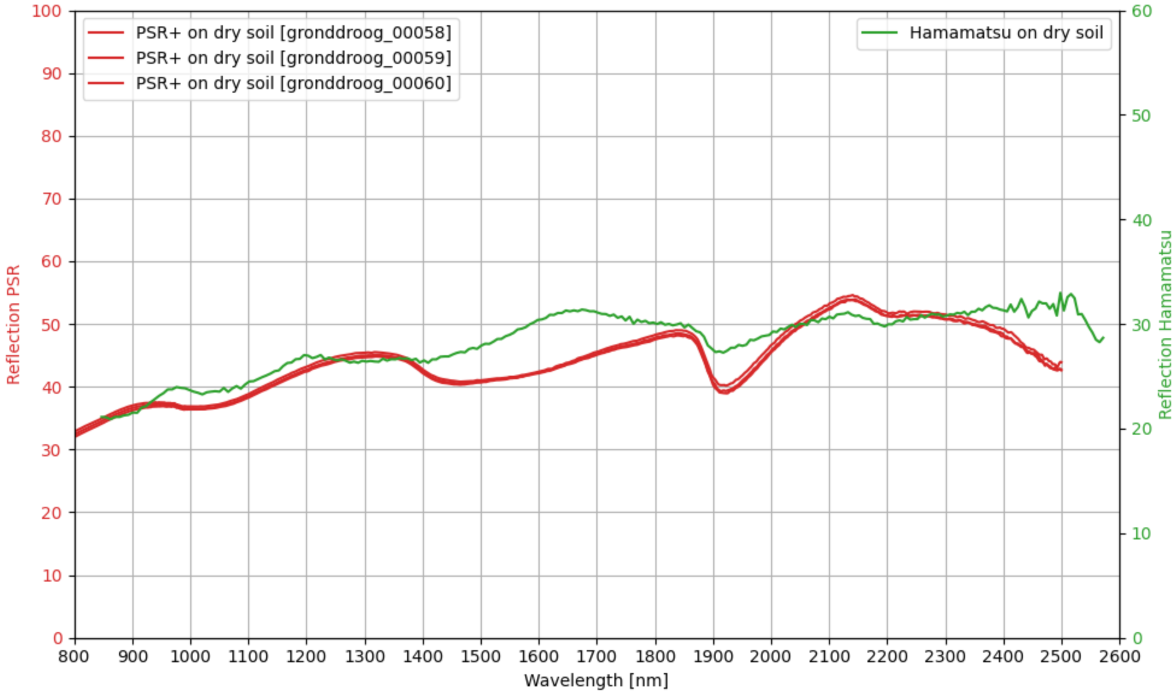

Figure 27: Comparison of reflectance between the PSR+ (red, spectra on three different mostly dry soil samples) and the Hamamatsu (Green, other and dryer soil sample)

4.2.4 Soil moisture sensor

The soil moisture sensor, a Delta-T ML3 ThetaProbe (Figure 28), will be integrated into the SOC assessment workflows to measure volumetric soil water content at each sampling point. This sensor provides precise, in-situ measurements that are critical for interpreting and correcting reflectance data from both the UAV multispectral camera and the spectrometer. Moisture levels significantly influence soil spectral properties; therefore, incorporating moisture data will enable the calibration of SOC predictions under varying field conditions. In the next iteration, workflows will be developed to acquire soil moisture measurements synchronously with spectrometer readings, store these measurements, and feed them into the SOC prediction models for improved accuracy. Initially, measurements will be performed manually, with plans to automate acquisition through the CIMAT robot in future phases.

Figure 28: Initial testing of the Delta-T ML3 ThetaProbe soil moisture sensor

4.2.5 Decision support system

The decision support system (DSS) will form the final layer of the SOC assessment workflow, transforming raw data and model outputs into actionable information for farmers. The DSS will integrate SOC predictions derived from UAV imagery and spectrometer measurements, alongside other relevant agronomic indicators, into a unified platform. It will generate user-friendly maps, visual summaries, and decision recommendations that support precision soil management. The development of the DSS workflows will advance after the completion of stakeholder interviews, which focus on understanding the data visualizations, levels of detail, and decision guidance that farmers and advisors need. These insights will directly inform the design of DSS workflows, ensuring the system provides practical, easy-to-interpret outputs that align with real-world farming requirements.

SOC map generation

The SOC map generation workflow will produce high-resolution, georeferenced maps showing the spatial distribution of soil organic carbon across the surveyed fields. This workflow represents a key output of both the drone and robot workflows, combining UAV-derived SOC predictions with spectrometer measurements taken at targeted sampling points to refine and validate the maps' accuracy. The SOC maps will integrate information from both aerial and proximal sensing, ensuring they reflect a comprehensive understanding of soil conditions.

These maps will categorize field areas by SOC status (e.g., low, medium, high) and delineate management zones that support precision soil management strategies. Importantly, the SOC maps will serve as the primary visualization layer within the decision support system (DSS), providing farmers with clear, actionable insights into the spatial variability of SOC across their fields. Maps will be designed for compatibility with farm management software and GIS platforms, ensuring that users can easily integrate them into their existing workflows and use them to guide targeted interventions such as variable-rate fertilization or soil amendment applications.

NDVI action

The NDVI action workflow will calculate the Normalized Difference Vegetation Index (NDVI) from Earth Observation sources, such as Sentinel-2 satellite data [5], or UAV multispectral imagery when fields are covered by crops. This workflow will provide valuable information on vegetation Vigor and crop health, allowing users to identify areas of potential stress or variability in plant growth that may relate to underlying soil conditions.

By integrating NDVI maps with SOC assessments in the DSS, the workflow will enable users to correlate crop performance with soil organic carbon distribution. This combined perspective will support a more holistic approach to precision agriculture, where soil and crop indicators can be analysed together to adjust management strategies in a way that optimizes both soil health and crop productivity. The NDVI layers produced by this workflow will be visualized through the DSS alongside SOC and classification maps, ensuring farmers have a comprehensive, integrated view of their fields.

Classification action

The classification action workflow will synthesize information from the SOC maps and NDVI layers, incorporating user-defined parameters or management preferences, to classify each area of the field into practical management zones. By combining these layers, the workflow will account for both soil organic carbon levels and vegetation health, resulting in more meaningful classifications that reflect not just static soil conditions but also dynamic crop performance.

Through the DSS interface, users will have the opportunity to adjust classification settings, such as defining SOC thresholds, selecting which NDVI ranges indicate stress or vigor, or specifying the number of management zones to generate. The workflow will then apply these combined criteria to produce classified maps that segment fields into zones with tailored recommendations for interventions like variable-rate fertilization, irrigation, or soil amendments. These classified maps will be a core component of the DSS outputs, delivering clear, actionable guidance to support data-driven precision agriculture decisions.

The ongoing stakeholder interviews are expected to directly inform and define key user inputs required for the classification workflow and decision support system. These inputs include the thresholds for classifying SOC levels (e.g., what farmers consider low, medium, or high SOC in their specific contexts), acceptable NDVI ranges for distinguishing healthy from stressed crops, and preferences regarding the number and granularity of management zones generated from the combined SOC and NDVI data. Additionally, the interviews will help clarify whether farmers prefer simple, easily interpretable maps or more detailed, layered outputs, as well as identify their expectations for customizing recommendations according to their specific crop types, soil conditions, and management practices. Insights gathered will ensure the workflows and DSS provide flexible, user-aligned tools that match real operational needs in precision agriculture.

[5] https://dataspace.copernicus.eu/explore-data/data-collections/sentinel-data/sentinel-2

4.3 Integration with the EMPYREAN components

For most of the building blocks developed within the use case, Ryax-compatible versions have been implemented. These modular components can be deployed into a complete end-to-end workflow, enabling smooth integration within the broader EMPYREAN architecture.

Figure 29 illustrates the sequence of steps involved in the UAV-based phase of the use case. This includes downloading drone images, followed by image correction and processing. Next, SOC levels, NDVI, and other relevant metrics are computed. Finally, the processed results are uploaded to the Chocolate Cloud storage platform, where they can be accessible for representation through the decision support tool.

This implementation demonstrates that the developed components are not only functional, but also fully operational within the Ryax workflow engine. It confirms that the technical groundwork established during this initial phase supports the scalable, automated processing pipelines envisioned for future iterations, particularly once the complete version of the EMPYEAN framework becomes available.

Figure 29: Workflows and Resources integrated within Ryax engine

4.4 Stakeholders engagement

As part of the first iteration of Task 5.3, we initiated a stakeholder engagement process to validate the direction and applicability of our dynamic Soil Organic Carbon (SOC) assessment solution. The aim is to ensure the system under development aligns with the practical needs of end users, particularly in the agricultural sector. So far, three stakeholder interviews have been conducted: one internally with ILVO's Agrifood and Technology living lab[6], one with the Boerenbond farmer's association[7] in Belgium, and one with Green Supply Chain – Digital Innovation Hub[8], a technology provider specialised in data interoperability and agricultural digital solutions in Greece. Our objective is to complete a total of ten interviews by the conclusion of Work Package 5. The interview questions are provided in Annex 1.

4.4.1 Farmers' association (Boerenbond)

The interview with Boerenbond, a leading agricultural organization in Belgium, provided valuable feedback from a farming perspective. SOC monitoring was recognized as a critical component for improving soil health and optimizing input use. Boerenbond sees clear benefits in transitioning from single-point measurements to spatial SOC mapping, particularly for guiding variable-rate fertilization and identifying problematic field zones such as wet spots. SOC data is viewed as a foundational layer that should be combined with other key variables, such as soil moisture, pH, crop type, weather data, and plant reflectance, all of which influence daily decision-making in the field. Rather than receiving raw data or numeric reports, farmers expressed a preference for receiving actionable insights through maps, ideally accompanied by clear interpretation. While ultra-fine spatial resolution was not considered necessary, the ability to interpolate and delineate management zones within a field was considered essential. In terms of update frequency, Boerenbond referenced standard practices suggesting SOC updates every six to seven years are sufficient for most use cases.

Boerenbond expects SOC maps to support a wide range of agronomic decisions, including fertilization strategies, lime or plaster application, and even irrigation, particularly when field conditions like soil preparation and microclimates are factored in. Integration with digital tools already in use, such as FarmOS[9] is considered important.

They are also expressed interest in combining SOC assessments with Earth Observation indices such as NDVI and NDWI, as well as incorporating shadow and moisture data. Concerns around, trust in data, privacy, and user control were also highlighted as critical factors. Boerenbond emphasized the importance of two-factor authentication, protective measures for drone data, and clear permissions or data-sharing agreements. They also expressed interest in having the ability to adjust task recommendations, such as choosing the number of zones, based on real-world conditions and potentially comparing recommendations with tractor or EO data. This

[6] https://ilvo.vlaanderen.be/en/living-labs/living-lab-agrifood-technology
[7] https://www.boerenbond.be/homepagina
[8] https://www.greensupplychain.eu/el
[9] https://farmos.org/

input from Boerenbond confirms that dynamic SOC assessment is not only technically feasible but highly relevant to farm management. Their feedback will directly inform the development of user-facing features and guide further iterations of the system as additional stakeholder interviews are conducted.

4.4.2 ILVO living lab

An internal interview with the leader of ILVO's robotics living lab provided a scientific and operational perspective on the value and potential applications of SOC monitoring in agriculture. ILVO, as a research institute with an operational farm, has a dual role: advancing the scientific understanding of soil organic carbon while also guiding and supporting farmers in adopting sustainable practices. Monitoring SOC is of high importance, both as a means of tracking carbon stocks at depth and as a tool for managing topsoil variability. Particularly, SOC distribution maps can inform variable-rate composting strategies aimed at homogenizing organic matter levels across the field. This contributes to consistent mineralization rates, more efficient nitrogen use, and reduced nutrient losses to the environment.

Effective use of SOC data requires integration with other parameters such as soil texture, pH, and nitrogen availability in the topsoil. Additionally, crop status data, potentially captured through high-resolution drone imaging, can support precise fertilization planning. These datasets must be considered in conjunction to make meaningful agronomic decisions.

For operational deployment, SOC recommendations should be delivered in compatible formats such as shapefiles or ISO-XML for task map execution, accompanied by explanatory reports. When evaluating carbon stocks, consistency in units, either mass-based or depth-based, was highlighted as critical. The preferred spatial resolution depends on the use case: field-level information is sufficient for overall carbon stock assessment, while finer resolution is needed for variable-rate applications. The resolution should ideally match the working width of application equipment, such as a 6-meter compost spreader.

ILVO sees immediate value in using SOC data primarily for nitrogen management, since carbon accumulation is a slow process. Consequently, annual updates, particularly before the start of the growing season, are considered adequate. While the use of Earth Observation indices is recognized, limitations due to small field sizes in Flanders reduce their practical applicability. On-field measurements via soil sampling or proximal sensing are more reliable under these conditions. In terms of digital infrastructure, ILVO uses platforms such as QGIS[10] and supports integration through open formats and APIs. Trust and data control remain key concerns; all data should remain under the control of the farmer, with explicit permission required for sharing. To ensure the system remains usable and responsive, simple feedback mechanisms that allow farmers to comment on the appropriateness of suggested task maps are suggested, enabling an iterative improvement process. This interview confirmed that robust SOC monitoring is viewed not only as a research interest but also as a practical necessity within ILVO's dual role as a scientific and agricultural actor. Their insights will guide technical

[10] https://qgis.org/

development, especially in ensuring interoperability, data accuracy, and user autonomy in future iterations of the system.

4.4.3 Technology provider (GSC)

The interview with Green Supply Chain (GSC), an organization focused on sustainable agricultural practices, highlighted the value of SOC distribution maps for both operational decision-making and broader policy applications. SOC mapping is seen as instrumental for optimizing fertilizer type and dosage, guiding soil treatment depth and machinery choices, and evaluating soil quality changes pre- and post-cultivation. Beyond agronomy, GSC noted its potential for enabling policy makers to validate soil quality data more automatically and support reporting for organic farming requirements.

GSC emphasized the need to contextualize SOC data with additional information such as soil properties, crop-specific carbon impacts, mowing frequency, and historical field activity. They envision a system that delivers insights through high-resolution maps for prescription applications and printable reports for advisors, with the flexibility to tailor outputs based on user needs. Reports designed to align with CAP-compliant practices were also highlighted as valuable deliverables.

Spatial resolution preferences vary by user role: while machine-generated task maps should support the highest granularity possible, zone-based recommendations are sufficient for human interpretation. In terms of update frequency, GSC suggested SOC assessments at the beginning of the cultivation period, especially if paired with low-cost nitrogen analysis. They expressed strong interest in integrating SOC data with satellite-derived indices, particularly soil moisture data, to explore correlations between moisture and organic matter.

GSC currently uses a range of digital tools, including weather forecasting platforms, UAV-based suitability assessments, land parcel identification systems, and farm calendars. Seamless integration with these tools and with existing soil maps is considered essential. Challenges cited include the high cost of large-scale sampling and the need for harmonized, interoperable data. Importantly, GSC stressed the importance of clear user agency: farmers should have the ability to accept or reject system-generated recommendations and understand the rationale behind them. While privacy and data control were not seen as top concerns at least for farmers in Greece, user adoption and clarity of service were viewed as critical to system success. The feedback from GSC will be a key input in designing flexible, user-centric features that support both field-level decision making and regulatory reporting.

4.5 Testing and evaluation

The Proximal Sensing in Agriculture Fields use case has undergone extensive testing during this first development cycle to verify the functionality, integration, and readiness of its various subsystems. These testing activities focused on validating the correct operation of hardware components, data acquisition pipelines, and preliminary software workflows in preparation for full system integration with the EMPYREAN architecture in future stages.

At the UAV level, testing was carried out to validate integration between the drone platform and the edge computing hardware. The DJI E-Port development kit was used to establish power and data connectivity with the Raspberry Pi-based edge node. Using DJI's Payload SDK (PSDK), software integration was tested by successfully retrieving telemetry data from the drone, including GPS coordinates, battery status, and flight parameters. These early tests confirmed the ability of the payload device to subscribe to real-time data streams, forming the basis for downstream processing and orchestration. To support monitoring of computational energy consumption, a USB power meter was added to the edge stack, and a data collection script was developed and deployed. The script collects voltage and current measurements from the AI-enabled edge device, storing the results in an InfluxDB time-series database to enable post-flight analysis and future performance optimization.

Energy monitoring tests established a dual-source framework for tracking power usage: drone-wide telemetry provided by the PSDK, and component-level insights captured by the USB meter. This allowed the team to isolate the energy demands of the edge computation workload and quantify its impact on total UAV power consumption—an important consideration when deploying AI at the edge in power-constrained environments.

The robotic component of the workflow was also validated through integration with ILVO's CIMAT robot using the ARTOF framework. A mini PC was installed on the robot to function as an auxiliary processing unit, interacting with the robot via ARTOF's Redis interface. A monitoring application was developed to periodically poll key telemetry values such as GNSS position, battery level, and component states. Initial tests confirmed reliable data retrieval and demonstrated the potential for tight coupling between the robot's movement logic and onboard sensing tasks. These tests also showed that the robot could follow paths generated from prior drone-based surveys, establishing a basic feedback loop between UAV and UGV workflows.

Integration testing for the onboard sensors, the Hamamatsu spectrometer, and the ThetaProbe soil moisture sensor, focused on verifying their operational readiness and suitability for field deployment. The spectrometer was tested using soil samples under controlled lighting conditions, with experiments designed to identify optimal exposure times and quantify signal noise. Results showed that longer integration times (e.g., 5000 µs) significantly improved data quality, especially in the near-infrared range. Repeated measurements and averaging techniques further reduced noise and validated the consistency of the spectrometer's output. These spectral readings were then compared to those of a reference spectrometer (Spectral Evolution PSR+), and although samples differed, the overall reflectance patterns were consistent, indicating a good degree of reliability. Spectrometer control and data acquisition were handled via a Windows SDK in C++, which will be encapsulated in a higher-level API during future development.

The ThetaProbe soil moisture sensor was subjected to initial standalone tests to confirm its responsiveness and calibration range. These early trials established confidence in the sensor's precision and confirmed its suitability for integration with SOC prediction models. Plans are in place to synchronize soil moisture readings with spectrometer measurements during robotic sampling, allowing corrections for moisture-related variation in spectral reflectance.

In parallel, initial data workflows have been set up to prepare for higher-level tasks such as radiometric calibration, image stitching, SOC map generation, and decision support. While these components are not yet fully integrated, test environments and data storage solutions (e.g., InfluxDB, local S3-compatible storage) are already in place to support future iterations. Early work on the SOC map and NDVI workflows is ongoing, with planned connections to the decision support system once stakeholder feedback is fully incorporated.

Overall, these testing activities demonstrate that the core technological components of the precision agriculture use case are functioning as expected, and that the supporting software frameworks are ready to handle real-time data collection and analysis. These results establish a strong foundation for the next cycle of development, in which the full EMPYREAN orchestration and data management stack will be integrated with the use case workflows. Future testing will focus on validating these higher-level integrations under operational conditions and confirming the overall robustness and utility of the system in the field.

4.6 KPI progress

In deliverable D2.1, we have defined three KPIs and respective metrics to track the progress of the use case development.

No	Indicator	Success Criteria
1	Development of processes that support the transition from subjective to objective, accurate and harmonised soil health data sets	2 SOC models developed for different sensing technologies
2	Transition to a real- or near real-time assessment of soil, and water parameters, allowing cooperated integrated farm management;	1 SOC model should be able to run on edge hardware in near-real time
3	Reduce the time and effort needed to develop soil data-driven models, compared to training models with data from manual soil sampling campaigns.	by 25%

1. **Development of processes that support the transition from subjective to objective, accurate and harmonised soil health data sets:**

Two complementary SOC prediction models are under development within this use case. The first model is based on UAV multispectral imagery, which estimates SOC levels across the field from aerial data. The second model uses proximal sensing data collected by the ILVO CIMAT robot, relying on spectrometer and soil moisture sensor measurements to provide precise SOC values at specific sampling points.

Training datasets have already been collected for both models. For the UAV-based model, multispectral imagery has been processed and radiometrically calibrated to serve as input for model training. For the robot-based model, soil samples from ILVO fields have been analyzed in the laboratory to determine their SOC levels and corresponding spectral signatures, forming the foundation for the spectrometer-based training dataset. A first version of the AI model for the UAV-based SOC estimation has been developed and is currently undergoing internal validation and refinement.

2. **Transition to a real- or near real-time assessment of soil, and water parameters, allowing cooperated integrated farm management;**

The focus of this KPI is on achieving near-real-time SOC estimation through edge computing. The architecture and data flows required for running the SOC model on the ILVO robot have already been designed and validated. This includes the full data pipeline from the spectrometer and soil sensors to the on-board edge device, local data processing, and optional synchronization with the central EMPYREAN cloud infrastructure. Once the SOC model based on spectrometer data is available during the second development phase, it can be seamlessly deployed within this existing architecture.

3. **Reduce the time and effort needed to develop soil data-driven models, compared to training models with data from manual soil sampling campaigns.**

The assessment of this KPI depends on the availability of the AI model trained on spectrometer data. This model will allow quantification of the improvement in development efficiency compared to traditional SOC model training based on laboratory-analyzed soil samples. Once the spectrometer-based model is operational, the reduction in time and effort for developing accurate SOC prediction models will be measured against the baseline of conventional manual soil sampling and lab analysis workflows.

4.7 Next Iteration

The first development iteration of the Soil Organic Carbon (SOC) assessment use case focused on establishing the core building blocks of the solution. The second iteration will significantly advance these foundations by enhancing the existing workflows and their integration with EMPYREAN platform components, aiming to create a fully automated, end-to-end process. Building on the initial achievements, this phase will expand functionality, improve accuracy, and prepare the system for demonstration in realistic field conditions.

A key objective for the next iteration is the planned data collection campaign scheduled for September 2025, with a potential follow-up campaign in April 2026. These campaigns will involve the simultaneous acquisition of high-resolution multispectral imagery by the UAV and in-field spectrometer measurements performed manually using the Hamamatsu near-infrared spectrometer. Soil samples will also be collected at selected sampling points for laboratory analysis, establishing ground-truth SOC levels required to train and validate the SOC prediction models. An important constraint for these campaigns is that both the UAV-based imagery and the spectrometer data collection must be performed on bare soil, without vegetation cover, to ensure accurate SOC assessment. The absence of vegetation minimizes interference with reflectance measurements, allowing true soil signatures to be captured. Therefore, the main time windows for these data collection efforts align with periods immediately after harvest or during soil preparation, when fields are free of crops or residue.

For the drone workflows, most core components, including radiometric calibration, image stitching, and prescription map generation, have already been developed in the first iteration. The main focus for the next iteration will be on refining and optimizing these existing workflows to improve robustness, reliability, and efficiency. Particular attention will be given to ensuring consistent results across varying flight conditions, streamlining data transfer to cloud storage, and improving automation within the Ryax orchestration environment to enable smooth integration with downstream processes.

For the robot workflows, a key priority will be the integration of the Hamamatsu near-infrared spectrometer into the ILVO CIMAT robot's autonomous operations. This will involve both the hardware-level integration of the spectrometer with the robot's onboard computing system and the development of new software workflows that allow the robot to acquire reflectance measurements autonomously during its missions. These workflows will include precise control of spectrometer operations, synchronization with the robot's positioning system, and real-time data logging. Additionally, workflows will be created to handle soil moisture sensor measurements, ensuring the collected data can be correlated with SOC predictions to improve model accuracy.

Parallel to the workflow developments, the next iteration will continue the development and refinement of SOC prediction models. Both the multispectral imagery-based model and the spectrometer-based model will be updated with new data collected during the planned bare-soil field campaigns. This work will include retraining and validation to improve model accuracy, reliability, and generalizability to different field conditions. The refined models will form the basis for generating actionable SOC maps for farmers.

For the decision support system (DSS) workflows, development will advance after the completion of the remaining stakeholder interviews, which are focused on understanding what users expect from SOC assessments and how they prefer to visualize and interact with decision support outputs. Based on insights gathered, the DSS workflows will be designed and implemented to transform SOC model outputs into user-friendly, actionable insights, integrating them with other relevant indicators such as NDVI, NDWI, or others. The DSS workflows will also define how data from UAV and robot workflows are aggregated, processed, and presented to farmers through maps or dashboards, ensuring the system aligns with user needs and expectations.

During the current iteration, initial integration of the developed workflows with the EMPYREAN Ryax orchestration engine has been achieved, enabling dynamic execution of UAV and preliminary data processing tasks across edge and cloud resources. In the next iteration, integration efforts will expand to cover all newly developed and refined workflows, including those for the robot's autonomous SOC measurements and the decision support system (DSS). These workflows will be orchestrated within Ryax to ensure coordinated execution across EMPYREAN's distributed architecture.

Furthermore, the next iteration will focus on advancing the integration of telemetry into the EMPYREAN platform, enabling real-time monitoring of key performance metrics from both UAV and robot systems. This will provide insights into device status, energy consumption, and workflow performance, supporting proactive maintenance and optimization.

In addition, the integration will extend to other core EMPYREAN components, including cloud and edge storage solutions, ensuring seamless and secure data transfer and storage across the continuum. Integration with EMPYREAN's security and trust management features will also be prioritized to guarantee data integrity, privacy, and compliance with user-defined permissions, enhancing the trustworthiness of the SOC assessment workflows and the overall platform.

These integration activities are essential for delivering a fully functional and cohesive implementation of the EMPYREAN system, capable of dynamically managing the entire SOC assessment process, from data acquisition and processing to secure and user-friendly decision support for end users.

5 5G-Enabled Vehicle-Assisted Services

5.1 Overview

This use case targets vehicular edge-assisted services operating in dynamic, latency-sensitive 5G environments. The objective is to ensure continuous, secure, and adaptive orchestration of vehicular workloads such as telemetry analysis, camera stream processing by offloading them to nearby edge infrastructure as vehicles move across different network domains.

The implementation and validation will take place at the GAIA Lab testbed of the University of Murcia (UMU), which offers a private 5G network, SDN/NFV-based orchestration, and multiple distributed edge nodes (ATICA, Luis Vives, GAIA, Bellas Artes). This environment provides the ideal setup to test service migration, mobility management, and edge-to-cloud data workflows using the EMPYREAN architecture.

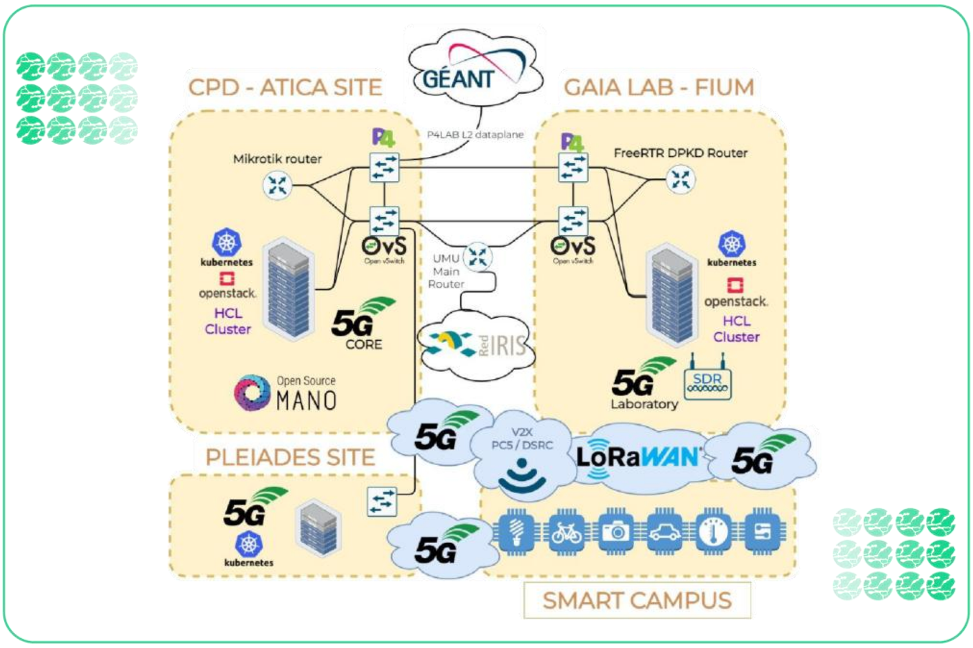

Figure 31: The figure shows the GAIA Lab testbed architecture, with interconnected 5G core and edge sites (ATICA, GAIA, Pleiades) using OpenStack, Kubernetes, and SDN routers to support smart campus mobility and orchestration experiments.locations used for vehicular mobility and orchestration experiments.

5.2 Initial development

The first development phase focused on defining the system architecture and execution workflows that will guide the integration of EMPYREAN components in the vehicular context.

The design considers multiple edge nodes connected to the 5G core network, each acting as part of an Association with distinct computational and storage roles.

The overall system concept includes:

- Virtual Road-Side Units (vRSUs) dynamically deployed close to the vehicle for low-latency processing.
- Orchestration and decision layers at the far edge (ATICA) controlling migration and workload optimization.
- Telemetry-driven anomaly detection mechanisms to ensure resilience.
- Distributed edge-to-cloud storage governed by policy-driven gateways

This initial phase mainly focused on specifying these functional components, their interconnections, and the two workflows that define the operational logic of the use case.

5.3 Integration with the EMPYREAN components

The integration of UC3 within the EMPYREAN architecture is structured around the deployment of a core set of components responsible for orchestration, monitoring, security, and data management. Each component is associated with one or more workflows that define its role in the vehicular service ecosystem.

- **Orchestrator and Decision Engine**
  Responsible for workload placement and migration decisions based on latency, telemetry metrics, and predicted vehicle movement.

  <u>Embedded Workflow – vRSU Deployment and Migration:</u>

  1. The vehicle's On-Board Unit (OBU) connects via 5G to Edge 1 (Luis Vives).
  2. The Orchestrator deploys a vRSU and Telemetry Engine instance at Edge 1.
  3. As the vehicle approaches Edge 2 (GAIA or Bellas Artes), the Decision Engine predicts handover and triggers migration.
  4. The Workflow Manager (RYAX) executes the redeployment of services, ensuring minimal downtime.
  5. DICE and DID/VC mechanisms validate the identity and integrity of components before traffic redirection.
  6. Telemetry confirms the success of the migration, while the CTI Module verifies there are no anomalies during the process.

- **Telemetry Engine**
  Collects and streams real-time operational data (latency, throughput, resource usage, signal quality) to feed orchestration decisions and enable early anomaly detection.

- **Cyber Threat Intelligence (CTI) Module**
  Correlates telemetry with contextual threat data to detect and react to potential attacks, such as spoofing or service disruption attempts during vehicle handovers.

- **Privacy and Security Manager (PSM)**
  Ensures service and node authenticity using Secure Boot, DICE-based attestation, and DID/VC verification during all orchestration events.

- **Workflow Manager (RYAX)**
  Manages distributed service orchestration, data transfer pipelines, and coordination between orchestration, telemetry, and storage components.

- **Skyfolk Edge Storage Gateway**
  Provide policy-driven data management for heterogeneous data types.

  <u>Embedded Workflow – Edge-to-Cloud Data Management:</u>

  1. Each Association includes an Edge Storage Gateway and dedicated S3 buckets.
  2. OBD telemetry data follows a cloud-based policy, automatically pushed to cloud storage via the Association gateway.
  3. Image and perception data follow an edge-based policy, stored locally to minimize latency.
  4. The RYAX Workflow Manager coordinates data movement between storage tiers according to policy and context.

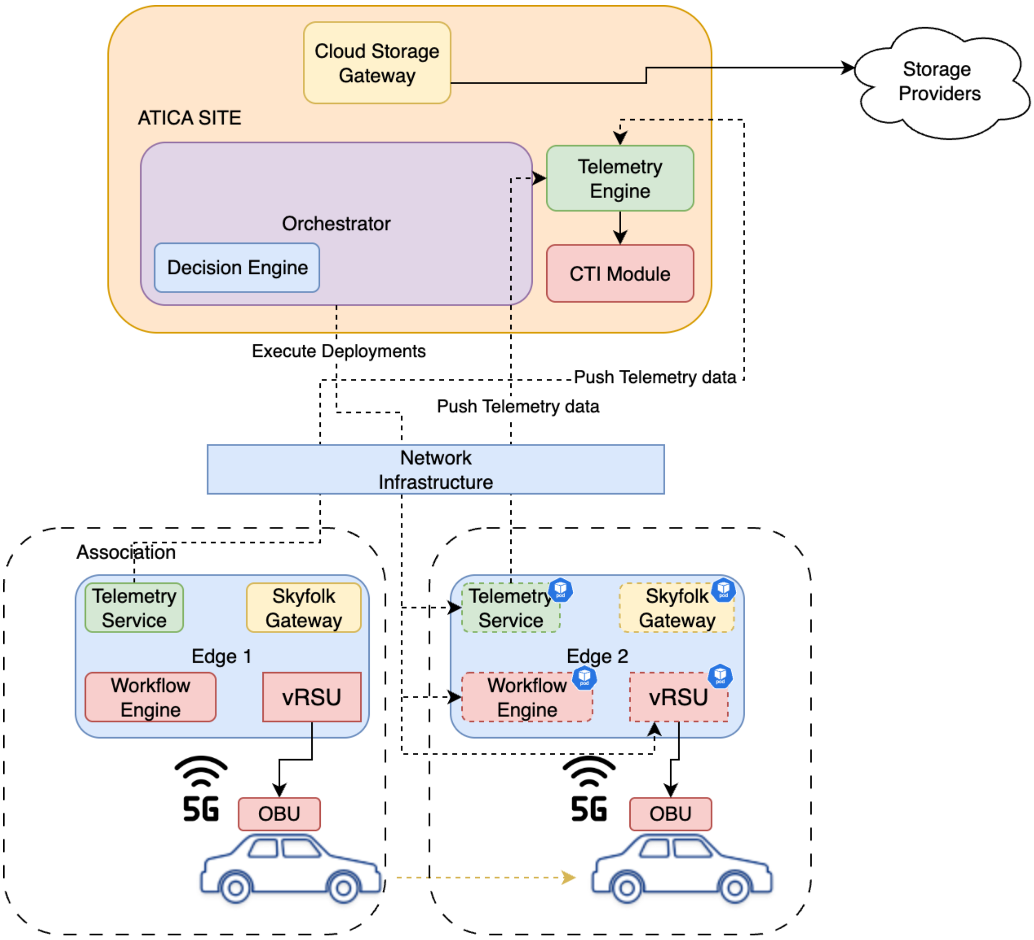

Figure 32: Integration of the vehicle-assisted use case within EMPYREAN combines central orchestration, telemetry, and CTI with edge components like vRSUs, workflow engines, and Skyflok gateways to enable secure, adaptive service migration and storage.

5.4 Stakeholder engagement

The UC3 use case brings together multiple stakeholders from the automotive, telecommunications, and cybersecurity domains. Engagement is ongoing with industrial partners and academic collaborators to ensure technical feasibility and alignment with real-world vehicular orchestration scenarios.

A key part of this engagement is the international collaboration with Kookmin University (South Korea), which contributes expertise in IoT and Cyber-Physical System (CPS) security.

5.5 Testing and Evaluation

At this stage, UC3 remains in the design and pre-integration phase. The infrastructure at GAIA Lab is fully deployed, including 5G connectivity and distributed edge nodes. The next testing phase will focus on:

- Deploying the Workflow Manager and Orchestrator on the 5G edge nodes.
- Establishing OBU–edge communication and simulating basic handovers.
- Validating telemetry collection across edge domains. Testing initial storage policies via Skyfolk Gateways.
- Conducting functional checks of the PSM identity verification process (DICE + DID/VC).

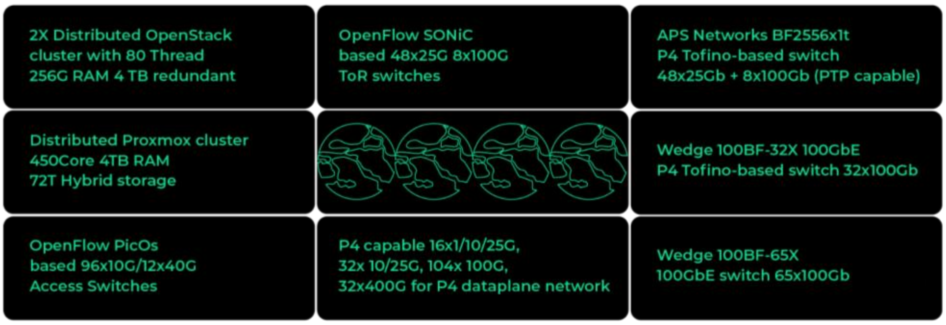

Figure 33: Networking resources at GAIA Lab,

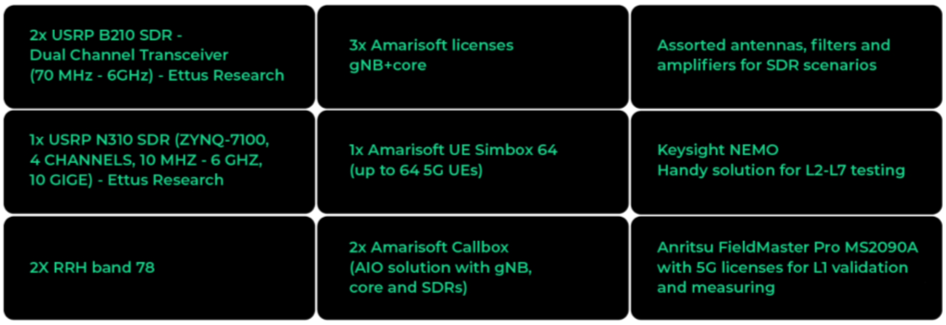

Figure 34: GAIA Lab's radio and measurement infrastructure, featuring SDR transceivers, 5G core and UE simulation tools, and professional-grade equipment for validating and analyzing network behavior and signal quality.

5.6 KPI Progress

As defined in Deliverable D2.1, three KPIs will guide the progress of this use case:

No	Indicator	Success Criteria
1	Service offloading and orchestration decision latency	< 20 seconds
2	Security alert generation and mitigation initiation	< 2 seconds
3	Reduction in false positive reduction using CTI-enhanced detection	≥ 30% improvement vs. telemetry-only baseline

At this stage of the project, the current KPI values remain at their initial baseline, as the use case is still in its early implementation phase.

The infrastructure and architectural design required to measure these indicators have been defined, and future iterations will implement the mechanisms needed to collect, monitor, and validate these metrics once the orchestration and CTI components are deployed.

5.7 Next Iteration

The next development cycle will focus on deploying and validating the defined workflows within EMPYREAN's orchestration environment. Planned actions include:

- Deploying RYAX workers on GAIA Lab edge clusters to enable dynamic vRSU instantiation and migration.
- Implementing state-transfer and rollback mechanisms to ensure seamless orchestration during mobility.
- Integrating PSM-based attestation and DID/VC checks into the orchestration workflow for secure service management.
- Testing policy-driven data management via Skyfolk and Chocolate Cloud under real vehicular movement scenarios.
- Activating CTI-assisted anomaly detection and mitigation during handovers.
- Extending collaboration with Kookmin University (South Korea) to perform cross-validation of EMPYREAN's orchestration and trust mechanisms in their smart-factory testbed, ensuring consistency and interoperability between European and Korean infrastructures.
- Collecting the first KPI measurements once orchestration and monitoring components are fully deployed.

The upcoming work will focus on consolidating the current architectural design into a working prototype and validating its operation in realistic 5G mobility scenarios, progressively integrating orchestration, telemetry, storage, and security functions across the distributed GAIA Lab infrastructure.

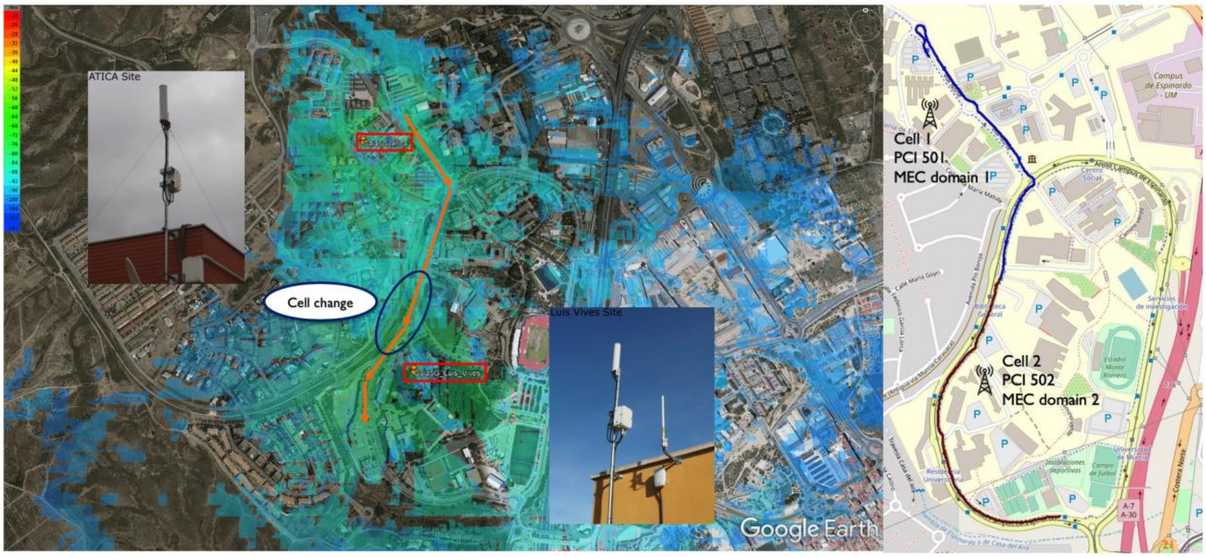

Figure 35: Coverage map showing the signal strength of 5G antennas at Luis Vives and ATICA sites, along with the vehicle route and handover area between two MEC domains, used for testing vRSU migration and service continuity in mobility scenarios.

6 EMPYREAN Integrated Platform and Use Cases Development

6.1 Integration Approach

The integration of diverse components and technologies into a cohesive and functional platform is critical for the EMPYREAN project. Given the distributed, heterogeneous, and multi-layered EMPYREAN architecture, the integration approach has been deliberately designed to ensure modularity, scalability, and flexibility, while supporting cross-cutting functionality across all use cases and architectural layers. Further details on the EMPYREAN CI/CD pipeline, as well as the development, integration, and validation infrastructure, are available in Deliverable D5.2 (M18).

EMPYREAN adopts an incremental integration strategy, aligning platform component integration with use case development across the project's phased release roadmap (initial, full, and final releases). This approach enables early validation and continuous refinement through successive iterations. To ensure coherence across the architectural stack, initial integration efforts have focused on three key areas: (i) device and edge integration, (ii) cross-component integration, and (iii) use case workflow integration.

This deliverable specifically emphasizes the third area, detailing how use case-specific workflows were adapted and extended to interface with core EMPYREAN services and infrastructure. These early integration steps demonstrate how platform functionalities can fulfil the operational needs and requirements of the individual use cases.

During the first development and integration cycle (M4-M18), integration efforts focused on:

- Establishing initial interface definitions and data models between key components.
- Developing lightweight prototypes of key services (e.g., workflow management, orchestration, deployment, telemetry) for early integration with selected use cases.
- Deploying early versions of the platform in controlled testing environments, using containerized services and EMPYREAN-developed orchestration and deployment frameworks.
- Demonstrating basic inter-component communication and coordination through selected workflows, such as robotic process monitoring and UAV-ground data fusion.

Looking ahead, the integration process will expand significantly in the second and third development cycles. Key activities will include: (i) integration and validation of all remaining platform functionalities, (ii) use cases evolvement to more complex workflows involving multi-site orchestration and cross-domain interactions, and (iii) Implementation and evaluation of advanced mechanisms for edge-cloud coordination, secure deployment, monitoring, and trust management.

Through this process, close coordination between technical work packages (WP3, WP4) and the use case development teams (WP5) remains essential to ensuring architectural alignment, operational coherence, and effective integration across all project components.

6.2 EMPYREAN SDK and API Integration

The success and widespread adoption of any platform depend significantly on the functionalities and services it offers to both end users and application developers. EMPYREAN aims to deliver a comprehensive Service Development Kit (SDK) equipped with all necessary tools to support the development and deployment of next-generation applications that fully leverage the platform's core innovations. The SDK and APIs are pivotal in enabling modular integration, extensibility, and seamless interaction between platform components, external services, and use case workflows. Designed with a strong emphasis on interoperability and developer-friendliness, they abstract the complexity of underlying platform services while exposing standardized interfaces to EMPYREAN's key capabilities.

The API layer is built primarily on RESTful and gRPC-based endpoints, allowing robust, language-agnostic integration across diverse and distributed environments. All API interactions are governed by EMPYREAN's security framework, which provides secure authentication, authorization, and data access control to ensure compliance with platform trust policies. Moreover, comprehensive and well-maintained API documentation is a cornerstone of EMPYREAN's development and integration process, especially given the platform's complexity. To this end, EMPYREAN adopts the OpenAPI[11] specification for synchronous APIs and the AsyncAPI[12] specification for event-driven, asynchronous communication. These machine-readable specifications not only streamline internal integration efforts but also simplify the extension and adoption of EMPYREAN components by third-party systems and external developers.

The EMPYREAN SDK will be delivered as a fully-featured Python package, offering developers reusable libraries, utilities, and templates to simplify the integration and extension of applications within the EMPYREAN ecosystem. These tools will facilitate access to critical platform services such as the EMPYREAN Registry, Aggregator, Privacy and Security Manager, Workflow Manager, Telemetry Service, Service Orchestrator, and Decision Engine.

6.3 Platform Release Timeline

The development, integration, validation, and delivery of the EMPYREAN platform follow a well-defined, phased methodology (Figure 36), with a detailed description available in the WP2 deliverables. This structured and iterative approach ensures consistent progress towards the project's overall objectives and the eventual deployment of the final platform. Each phase builds on the outcomes of the preceding one, facilitating smooth transitions from

[11] https://spec.openapis.org/oas/latest.html

[12] https://www.asyncapi.com/docs/reference/specification/v3.0.0

requirements analysis through development and integration to the deployment of a fully functional system. The methodology also incorporates continuous feedback and iterative refinement throughout the development lifecycle.

In alignment with this strategy, the EMPYREAN platform will be released in three major stages:

- **Initial release (Month 18)**: This release incorporates components and functionalities developed during the first implementation cycle (Months 4–15). At this stage, each system component delivers a partial set of its final features, along with the essential interfaces for inter-component communication. The initial release serves as a foundational prototype to showcase the platform's core capabilities and enable initial integration with EMPYREAN use cases, as outlined in the previous sections. A detailed description of the initial release and its covered functionalities is provided in Deliverable D5.2 (M18).
- **Full release (Month 30)**: Scheduled following the completion of technical developments in Work Packages 3 and 4 (by Month 26), this version will incorporate all remaining functionalities. It is intended to deliver a fully integrated and feature-complete platform, ready for pilot deployments and large-scale experimentation using the project's use cases.
- **Final release (Month 36)**: planned for the end of the project, this release will focus on optimizing and refining the platform based on the insights and results gained from final evaluations and demonstrations of the project's use cases (M30-M36).

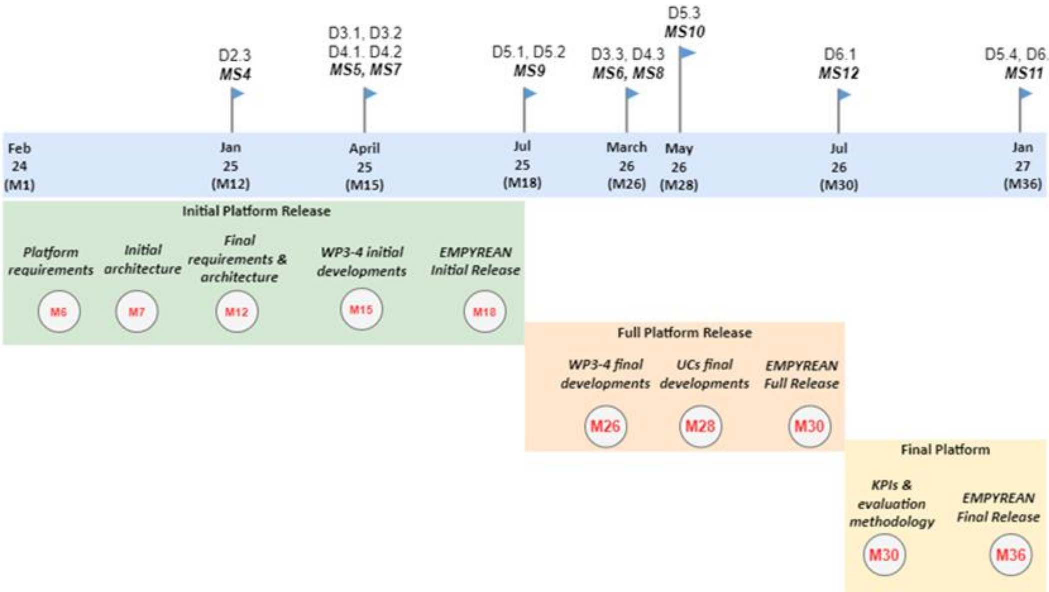

Figure 36: EMPYREAN development roadmap

7 Conclusions

7.1 Summary of Technological Progress

During the first development cycle, the EMPYREAN project has achieved substantial progress in implementing and validating its foundational architecture across three distinct and demanding use cases. These efforts provide early evidence of the platform's capacity to deliver on its vision: enabling trustworthy, secure, and dynamic orchestration of AI-driven services across the edge-cloud continuum. The technological progress spans several core areas, including the deployment of containerized, modular components on heterogeneous IoT and edge devices; the initial execution of orchestration workflows using the Ryax platform; the integration of distributed AI inference mechanisms aligned with specific real-world needs; and the development and implementation of use case-specific workflows that capture end-to-end data processing and service logic.

These workflows, tailored to the specific operational requirements of robotic machining, agricultural monitoring, and vehicular services, have enabled seamless coordination of sensing, processing, inference, and response across the edge-cloud continuum. Their successful deployment has not only validated EMPYREAN's technical components but also helped establish realistic integration pathways for orchestrated AI services in live or semi-realistic environments.

Each use case, robotic machining anomaly detection, proximal sensing in precision agriculture, and 5G-enabled vehicle-assisted services, has demonstrated how EMPYREAN's architectural principles can be effectively applied in practice. By focusing on real-time data collection, local processing, and intelligent decision-making close to the data source, the project has highlighted the feasibility of distributed intelligence and responsive orchestration. These efforts have not only validated the technological direction of the project but also contributed meaningful domain-specific insights that inform ongoing design and development.

While the full integration of EMPYREAN's capabilities is still underway, the initial implementations have already demonstrated how Associations, dynamic workflows, and secure data exchange mechanisms can support the deployment of AI-enhanced services in constrained, variable environments. This work establishes a strong foundation for the subsequent phase of technical refinement and scale-up, confirming both the flexibility and applicability of the EMPYREAN approach.

7.2 Identified Challenges and Mitigation

Throughout the first iteration of the EMPYREAN use cases, several technical and operational challenges have been identified that are inherent to deploying distributed, AI-driven workflows across heterogeneous IoT and edge computing environments. A major challenge lies in ensuring seamless integration between diverse hardware platforms and software components, particularly when dealing with constraints such as limited computational power at the deep edge, varying data formats, and intermittent connectivity. These issues have been partially mitigated through modular, containerized workflow designs and the use of orchestration tools such as the Ryax engine, which facilitate flexible deployment and management of services across the edge-cloud continuum. Additionally, the reliance on standard APIs and microservice architectures has improved interoperability, easing the integration of sensors, devices, and platform components.

Another category of challenges revolves around data quality, synchronization, and privacy. Real-time data acquisition from robotic systems, UAVs, and field sensors can introduce inconsistencies due to environmental noise, variable sampling rates, or hardware calibration issues. To address this, efforts have been made to introduce preprocessing and filtering mechanisms early in the workflows, as well as ongoing testing and validation procedures for both hardware and software components. On the security and privacy side, stakeholder concerns about data control and transparency have been acknowledged, and initial steps are underway to implement secure data pipelines, access control mechanisms, and user-defined permissions. These mitigation strategies will continue to evolve in future iterations as system complexity grows and user needs become more refined.

Looking ahead, the next iteration of development will prioritize refining and scaling the solutions to further mitigate the identified challenges. For integration-related issues, emphasis will be placed on enhancing the robustness of edge-to-cloud communication, improving compatibility between platform components, and optimizing orchestration strategies to support more dynamic and context-aware deployments. From a data perspective, workflows will be extended with enhanced validation, calibration, and synchronization mechanisms to ensure higher fidelity and consistency across sensor modalities. In terms of security and user trust, upcoming releases will incorporate stricter access controls, encryption mechanisms, and user-facing tools that allow for greater transparency and control over data usage. Stakeholder feedback collected during the first phase will directly inform these improvements, ensuring that the technical evolution of the EMPYREAN platform continues to align with practical, real-world needs.

7.3 Next Steps

In the upcoming phase of the project, efforts will be directed toward continuing technological advancements and ultimately finalizing the project use cases developed during this first cycle. While significant progress has already been achieved in establishing their core workflows and validating the early-stage implementation of key features, the next steps will focus on extending and refining these developments to support fully operational and demonstrable scenarios. This includes enhancing responsiveness, reliability, and completeness of the use case logic, deepening the integration of context-specific services, and scaling up to more complex and realistic conditions.

Each use case will continue to evolve technically, incorporating more advanced behaviours, richer data flows, and improved orchestration patterns, based on the foundations laid in the current deliverable. Particular attention will be paid to completing the implementation of domain-specific workflows, from data acquisition and processing to decision support and actuation, ensuring that all operational stages are coherently aligned and technically robust.

At the same time, the next development cycle will support the systematic integration of EMPYREAN's architectural elements within the existing use case frameworks. This will not be treated as a separate activity, but rather as a natural progression of the technological development of the use cases themselves. The orchestration logic, trust mechanisms, observability tools, and security strategies will be embedded and validated directly within the real workflows and execution contexts of each scenario. In this way, the full breadth of the EMPYREAN platform will be exercised through the lens of practical application.

This continued alignment between technological development and component integration will ensure that the use cases remain central to validating the project's overall vision. It will also enable a more accurate evaluation of EMPYREAN's performance under domain-specific requirements and constraints. By the end of the next phase, each use case is expected to demonstrate not only technical completeness, but also the successful embedding of the EMPYREAN architecture into a realistic and relevant operational setting.

8 Annex A

Use case 2 Stakeholders Interviews questions

How important is monitoring Soil Organic Carbon (SOC) for your farm or organization?

What specific gains (e.g., pinpointed fertiliser rates, targeted cover-crop zones, erosion control) do you anticipate from having a map of SOC distribution along the field instead of a single composite measurement?

What additional information, alongside Soil Organic Carbon (SOC) values, would help you make decisions about soil management or crop planning?

In what form would you prefer to receive SOC-related recommendations—maps, numeric reports, alerts, or something else?

At what spatial resolution (field, sub-field, per hectare, per zone) would SOC and task maps be most actionable for you?

What kinds of management decisions (fertilization, tillage, crop rotation) would you expect to base on SOC maps or soil health indicators?

What is the optimal timing or frequency for receiving SOC updates or task maps during a growing season?

Would you be interested in combining SOC data with other satellite indices (e.g., NDVI, NDWI) or sensor data in the same platform? If so, which ones?

What digital tools or platforms do you already use for farm management? How would you like this system to interact with them?

What challenges do you currently face in collecting or interpreting soil data, and how could a decision support system help address them?

How would you like to provide feedback or adjust recommendations if the suggested task maps do not align with on-the-ground realities?

How important are aspects like data trustworthiness, data sharing control, and privacy when using such a system?