FLEXIGROBOTS

D6.1 Pilot 3 objectives, requirements and design

Document Identification			
Status	Final	**Due Date**	31/12/2021
Version	1.0	**Submission Date**	27/12/2021

Related WP	WP6	**Document Reference**	D6.1
Related Deliverable(s)	D2.7, D6.2, D6.3, D6.4	**Dissemination Level (*)**	PU
Lead Participant	BioSense Institute (BIO)	**Lead Author**	Oskar Marko, BIO
Contributors	ART, ZEL	**Reviewers**	Angela Ribeiro, CSIC
			Daniel Calvo, ATOS

Co-funded by the Horizon 2020 Programme of the European Union

Document Information

List of Contributors	
Name	Partner
Oskar Marko	BIO
Marko Panić	BIO
Linas Didžiulevčius	ART
Bojana Ivošević	BIO
Goran Kitić	BIO
Nevena Momirović	ZEL

Document History			
Version	Date	Change editors	Changes
0.1	05/11/2021	Oskar Marko (BIO)	Initial version, document set up with the ToC
0.2	24/11/2021	Oskar Marko (BIO)	High-level architecture added, project objectives.
0.3	9/12/2021	Oskar Marko (BIO)	Diagrams updated, actors edited, formatting improved.
0.4	20/12/2021	Oskar Marko (BIO)	Document updated based on quality control (reviewers' comments)
1.0	22/12/2021	Oskar Marko (BIO)	FINAL VERSION TO BE SUBMITTED

Quality Control		
Role	Who	Approval Date
Deliverable leader	Oskar Marko (BIO)	22/12/2021
Quality manager	Iván Zaldívar (ATOS)	22/12/2021
Project Coordinator	Daniel Calvo (ATOS)	23/12/2021

Table of Contents

List of Tables

List of Figures

List of Acronyms

Abbreviation / acronym	Description
AgTech	Agricultural Technologies
AI	Artificial Intelligence
API	Application Programming Interface
AWS	Amazon Web Services
CC	Creative Commons
CMOS	Complementary Metal-Oxide-Semiconductor
CNN	Convolutional Neural Network
CSV	Comma-separated values
CV	Computer Vision
DOI	Digital Object Identifier
DSS	Decision-Support System
DL	Deep Learning
EC	Electrical Conductivity
ECC	Enhanced Correlation Coefficient
EO	Earth Observation
ESA	European Space Agency
EXIF	Exchangeable image file format
FAIR	Findable, Accessible, Interoperable, Reusable
FMIS	Farm Management Information System
FOV	Field of View
FWHM	Full Width-Half Maximum
GCP	Ground Control Points
GDPR	General Data Protection Regulation
GLI	Green Lead Index
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSD	Ground Sampling Distance
HDMI	High-Definition Multimedia Interface
IEC	International Electrotechnical Commission

Abbreviation / acronym	Description
JPEG	Joint Photographic Experts Group
KPI	Key Performance Indicator
LAI	Leaf Area Index
LIDAR	Light Detection and Ranging o Laser Imaging Detection and Ranging
LORA	Long Range
ML	Machine Learning
MP	Megapixel
NDA	Non-disclosure agreement
NVDI	Normalized Difference Vegetation Index
NIR	Near InfraRed
NTRIP	Network Transport of RTCM via Internet Protocol
PCA	Principal Component Analysis
PPK	Post Processed Kinematic
RGB	Red, Green, Blue
ROI	Region of Interest
ROS	Robot Operating System
RTCM	Radio Technical Commission for Maritime Services
SIPI	Structure Insensitive Pigment Index
SQL	Structured Query Language
SVM	Support Vector Machine
TIFF	Tagged Image File Format
TRL	Technology Readiness Levels
UAV	Unmanned aerial vehicle
UGV	Unmanned ground vehicle
USB	Universal Serial Bus
UV	Ultraviolet
UVL	Ultra-low volume
WP	Work Package
XMP	Extensible Metadata Platform

Executive Summary

This document gives an overview of Pilot 3, its requirements and technological achievements towards developing flexible robotic solutions for use in high-value crops, i.e., blueberries. Firstly, the document outlines the work plan set in the proposal, which has been refined during the first 12 months of the project. In concrete, there are 3 use cases that will be covered in Serbia and Lithuania and these include:

- Data acquisition using UAVs, UGV-mounted equipment, and other sources;
- Data processing using deep learning / image processing algorithms;
- UGV actions - pesticide application and soil sampling.

The document provides details about the UAV/UGV equipment that was used for data acquisition and that will be used for these purposes, along with the description of the datasets acquired using this equipment. Regarding the equipment, the Pilot foresees the development of innovative robotic parts for soil sampling and pesticide application. Their requirements and current status of their development have been elaborated in the corresponding section, while for both UAVs and UGVs, a precise pipeline was defined that the pilot will follow in terms of data acquisition and operation of the robots. The methodology defined in this way will ensure the highest standards of operation and a clear and feasible pipeline for the development of the technology.

All activities planned for this reporting period were conducted according to highest standards. Initial datasets were acquired and basic image processing algorithms developed, which will be further developed in the next phase of the project. Robotic modules for soil sampling and spraying have been designed according to the requirements and their implementation has successfully kicked-off.

1 Introduction

1.1 Purpose of the document

Blueberries can be grown in pots or on the ridges in the field, there are many different varieties and shapes of the plants, different settings require different equipment properties, while the choice of equipment dictates the way that activities are conducted on the field. The purpose of this document is to provide a detailed definition of the farmers' requirements in the field and lay the foundations for the technical development of concrete algorithms and robotic modules. Lastly, the goal is to set the pipeline for field experiments and the future operation of the system.

1.2 Structure of the document

This document is structured in 6 major sections.

Section 1 - Introduction. This section gives an overview of the document and provides a brief overview of the background and its purpose.

Section 2 - Objectives of the pilot. This section defines the concrete objectives of Pilot 3.

Section 3 - Description of use cases. This section covers the use cases that have been defined and that will be covered during project implementation.

Section 4 – Identified datasets. This section presents the datasets that have been produced and the ones that will be acquired during the course of the project.

Section 5 – Description of robots, devices and platforms. This section gives detailed technical specifications of the equipment and algorithms that will be developed within the project.

Section 6 - Conclusions. This section gives the conclusion of the document in terms of the technical specifications, development roadmap and field testing methodology.

2 Objectives of the pilot

FlexiGroBots is a project that aims to showcase the use of aerial and ground robots in agriculture, which are capable of conducting multiple field operations in different crops and regions across Europe. WP6 is concerned with the application of UAVs and UGVs in fruit production, i.e., high-value crops, which usually require a lot of manual labour and quite a few visits to the field. The goal is to conduct the pilot at two locations with contrasting growing conditions - Lithuania and Serbia. These countries are located on the northern and the southern tips of the region in Europe that supports the growing of the more quality type of blueberries - the northern high-bush blueberry. This allows farmers to achieve the highest prices, as their produce is sold at the very beginning and the very end of the season when this fruit is rare.

Concrete Pilot 3 objectives include:

- Early-stage blueberry disease detection
- Yield prediction based on disease mapping
- Automated field soil sampling and analysis
- Targeted and autonomous agrichemical spraying.

3 Description of use cases

In order to showcase the flexibility of robots to perform different field operations and contribute to holistic optimisation of activities in blueberry production, we defined 3 use cases that the pilot will focus on. These have been elaborated thoroughly in the corresponding documents that follow the IEC 62559 standard. Excerpts from these documents are provided below.

3.1 High-level description

Pilot 3 relies on UAVs, UGVs and their interaction, facilitated through the FlexiGroBots platform. The high-level architecture of the pilot is shown in Figure 1.

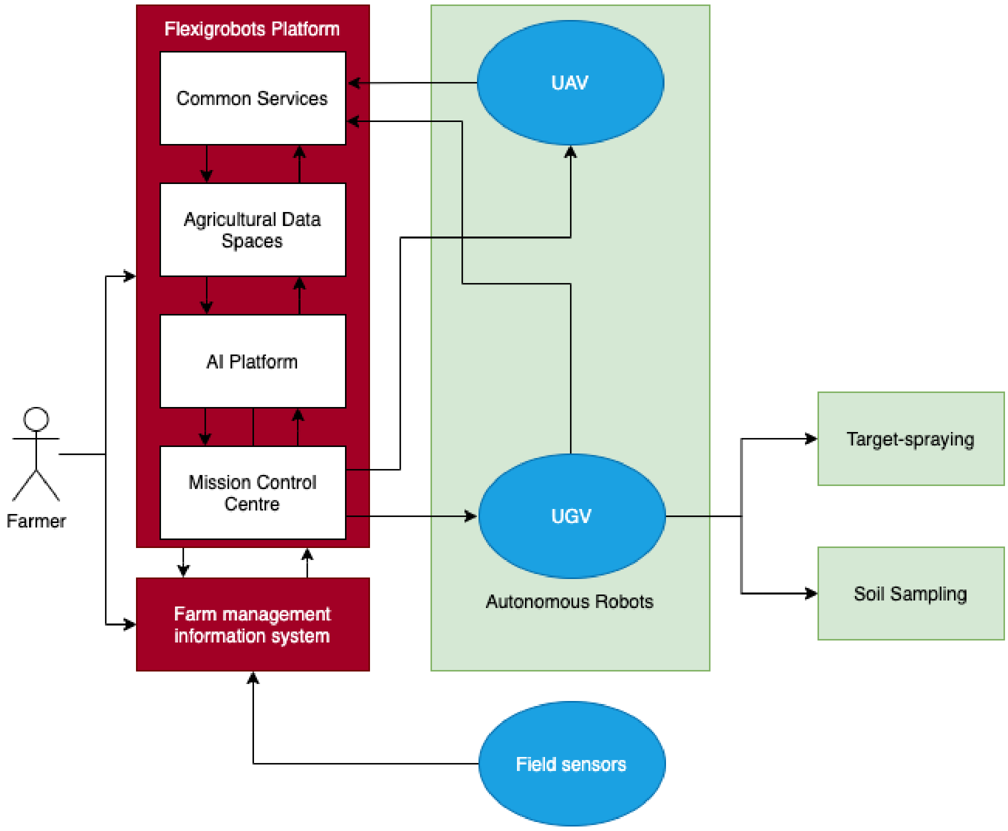


Figure 1: High-level overview of Pilot 3

UAVs are the primary tool for data acquisition and field monitoring. They are used for mapping the weeds and areas affected by diseases. This is a crucial piece of information, as it is an indicator of the expected yield and the maps derived in this way are the basis for UGV operation on the field. There are two main actions that the UGV will execute – soil sampling and target-spraying of pesticides. However, the UGV is also a source of information as the soil analyses, images and lidar data are acquired using the robotic platform. The FlexiGroBots platform is the central point of the pilot, as it encompasses the AI platform, Mission Control Centre and other components crucial for data processing and decision-making. Lastly, other sources of data, mainly field sensors have been installed on the field and they are sending the data to the farm management system that the farmer uses in the daily operations and crop monitoring.

3.2 Use case descriptions

3.2.1 Use case 1: Monitoring and detection of weeds and diseases in blueberries

Blueberries will be scanned using the UAVs regularly throughout the season. We will use vegetation indices to quantify the intensity of photosynthesis, overall health and availability of water and nutrients. By mapping the problematic areas, and assessing the health of crops, we will also predict the yield. UAV images will also be used for the detection of weeds and diseases.

3.2.1.1 Description of the use case

3.2.1.1.1 Name of Use case

Use case identification		
ID	Area / Domain (s) / Zone (s)	Name of use case
UC1	WP6 – Pilot 3	Monitoring and detection of weeds and diseases in blueberries.

Table 1: Use case identification

3.2.1.1.2 Version management

Version management				
Version No	Date	Name of author (s)	Changes	Approval status
0.1	26.5.2021	Oskar Marko		

Table 2: Version management

3.2.1.1.3 Scope and Objectives of Use case

Scope and objectives of use case	
Scope	WP6 – Pilot 3
Objective (s)	1) Monitoring the growth of blueberries. 2) Yield prediction. 3) Detection of problematic areas.

Scope and objectives of use case	
Related business case (s)	1. Informed business planning. 2. Reduction in pesticide/ herbicide amount. 3. More efficient production (lower costs and higher yields).

Table 3: Scope and objectives of use case

3.2.1.1.4 Narrative of Use case

Narrative of use case
Short description
Using UAVs to monitor the growth of crops and detect weeds and diseases.
Complete description
Blueberries will be scanned using the UAVs regularly throughout the season. We will use vegetation indices to quantify the intensity of photosynthesis, overall health and availability of water and nutrients. By mapping the problematic areas, and assessing the health of crops, we will also predict the yield. UAV images will also be used for the detection of weeds and diseases.

Table 4: Narrative of use case

3.2.1.1.5 Key Performance Indicators (KPIs)

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
1	Comprehensive monitoring of crops	Lower cost of soil sampling for 35%	Zone delineation through UAV analysis reduces the number of required soil locations.
2	Optimal and data-driven decision-making	25% less damages from sub-optimal pesticide application	Vegetation indices signify the occurrence of the diseases and are the first step.
3	Targeted herbicide and pesticide application	40% cut in pesticide/ herbicide amount	Vegetation indices signify the occurrence of the diseases and are the first step.

Table 5: Key performance indicators

3.2.1.1.6 Use case Conditions

Use case conditions
Assumptions
Blueberries are mostly grown in the open and not in the glasshouses.
Weeds and plant diseases pose severe problems in blueberry growing.
Soil sampling is necessary for optimal field management.
Prerequisites

Test fields defined in both Serbia and Lithuania.
Blueberries on the test field are healthy and old enough to give a representative yield.
Plant diseases visible on the leaves (in contrast to the root system or stems).

Table 6: Use case conditions

3.2.1.1.7 Further Information to the Use case for Classification / Mapping

Classification information
Relation to other use cases
Monitoring is the first step and a prerequisite for further development of decision-support systems, management zone delineation and precision farming in general.
Level of depth
Detailed Use Case
Prioritisation
High
Generic, regional or national relation
Generic
Nature of the use case
Technical

Table 7: Classification information

3.2.1.2 Diagrams of use case

Diagram(s) of use case

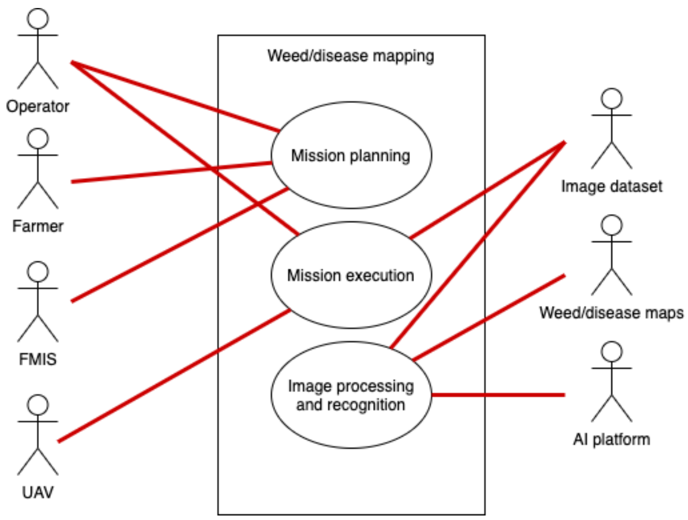


Figure 2: Diagram of weed/disease mapping use case

3.2.1.3 Technical Details

3.2.1.3.1 Actors

Actors			
Actor name	Actor type	Actor description	Further information specific to this use case
Operator	Human	Person controlling the UAV	Company's internal/ external operator
Farmer	Human	Farm owner/ agronomist	Person leading the production
FMIS	System	Internet platform/ app	AgroSense/ AgroSmart
UAV	Robot	Autonomous robot	UAV with multi or hyper-spectral camera
Image dataset	Data	Raw images from UAV	Multi/ hyper-spectral images
Weed/disease maps	Data	Georeferenced images	Weeds/ diseases areas highlighted
AI platform	System	AI algorithms for image recognition	DL algorithms for clustering/ zone delineation

Table 8: Actors

3.2.1.4 Step by Step Analysis of Use Case

3.2.1.4.1 Overview of the Scenarios

Scenario conditions						
No.	Scenario name	Scenario description	Primary actor	Triggering event	Pre-condition	Post-condition
1	Mission planning	Mission planning	Operator	Regular intervals	Unknown scanning area	UAV route/ height planned
2	Mission execution	Mission execution	UAV	Operator flies the UAV	No field images	Field images acquired
3	AI processing and recognition	Image processing and recognition	AI platform	UAV images acquired	Unknown distribution of weeds/ diseases	Problematic areas mapped

Table 9: Scenario conditions

3.2.1.4.2 Steps – Scenarios

Scenario							
Step No.	Event	Name of process / activity	Description of process / activity	Service	Information producer (actor)	Information receiver (actor)	Information exchanged (IDs)
1	Regular scanning scheduled	Mission preparation	Getting field boundaries	GET	FMIS	Operator	UC1-01
2	Field boundaries set	Mission planning	Boundaries and flying height planned	SET	Operator	UAV	UC1-02
3	Mission kicked-off	UAV scanning	UAV flying over crops	RUN	UAV	Dataset	UC1-03
4	Images acquired	Image processing	Ortho-mosaicing, vegetation indices	GET	Dataset	AI platform	UC1-03, UC1-04
5	Images pre-processed	Image segmentation	Detection of problematic areas	RUN	AI platform	Weed/disease maps	UC1-05

Table 10: Scenario

3.2.1.5 Information exchanged

Information exchanged		
Information exchanged, ID	Name of information	Description of information exchanged
UC1-01	Field boundaries	Shapefile of the field
UC1-02	Mission parameters	Mission height, area
UC1-03	Raw UAV images	Individual images from UAV
UC1-04	Pre-processed UAV images	Orthomosaic and vegetation indices
UC1-05	Weed/disease maps	Problematic areas mapped

Table 11: Information exchanged

3.2.2 Use case 2: Smart soil sampling

Homogeneous zones will be defined in the UAV images and the electrical conductivity map. The centroids of these zones are the locations where the UGV will be sent to perform automatic soil sampling and on-the-go lab analysis. Precision soil maps will be generated as a result.

3.2.2.1 Description of the use case

3.2.2.1.1 Name of Use case

Use case identification		
ID	Area / Domain (s) / Zone (s)	Name of use case
UC2	WP6 – Pilot 3	Smart Soil Sampling

Table 12: Use case identification

3.2.2.1.2 Version management

Version management				
Version No	Date	Name of author (s)	Changes	Approval status
0.1	26.5.2021	Oskar Marko		

Table 13: Version management

3.2.2.1.3 Scope and Objectives of Use case

Scope and objectives of use case	
Scope	WP6 – Pilot 3
Objective (s)	1) Perform zone delineation from UAV images- 2) Defining the sampling points. 3) Conducting soil sampling with the UGV.
Related business case (s)	1. Precision soil mapping for optimisation of fertiliser application and irrigation. 2. Production of more quality produce. 3. No need for human labour.

Table 14: Scope and objectives of use case

3.2.2.1.4 Narrative of Use case

Narrative of use case
Short description
Using UAV images to delineate zones and sampling points where the UGV is sent to perform sampling.
Complete description

Homogeneous zones will be defined in the UAV images and the electrical conductivity map. The centroids of these zones are the locations where the UGV will be sent to perform automatic soil sampling and on-the-go lab analysis. Precision soil maps will be generated as a result.

Table 15: Narrative of use case

3.2.2.1.5 Key Performance Indicators (KPIs)

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
1	Comprehensive monitoring of crops	Lower cost of soil sampling for 35%	Targeted soil sampling means more precise maps with fewer samples
2	Optimal and data-driven decision-making	25% less damages from sub-optimal pesticide application	Optimal decision-making regarding fertiliser application increases the overall health status of the plants which are hence more resistant.
3	Targeted herbicide and pesticide application	40% cut in pesticide/ herbicide amount	Healthier plants are less likely to be affected by pests, and optimal fertiliser application contributes to their health.

Table 16: Key performance indicators

3.2.2.1.6 Use case Conditions

Use case conditions
Assumptions
Blueberries are mostly grown in the open and not in the glasshouses
Weeds and plant diseases pose severe problems in blueberry growing
Soil sampling is necessary for optimal field management
Prerequisites
Test fields defined in both Serbia and Lithuania
Blueberries on the test field are healthy and old enough to give a representative yield
Plant diseases visible on the leaves (in contrast to the root system or stems)

Table 17: Use case conditions

3.2.2.1.7 Further Information to the Use case for Classification / Mapping

Classification information
Relation to other use cases
Soil sampling is performed based on the UAV images processed in the Monitoring use case (Use Case 1).

Level of depth
Specific
Prioritisation
High
Generic, regional or national relation
Generic
Nature of the use case
Technical

Table 18: Classification information

3.2.2.2 Diagrams of use case

Diagram(s) of use case

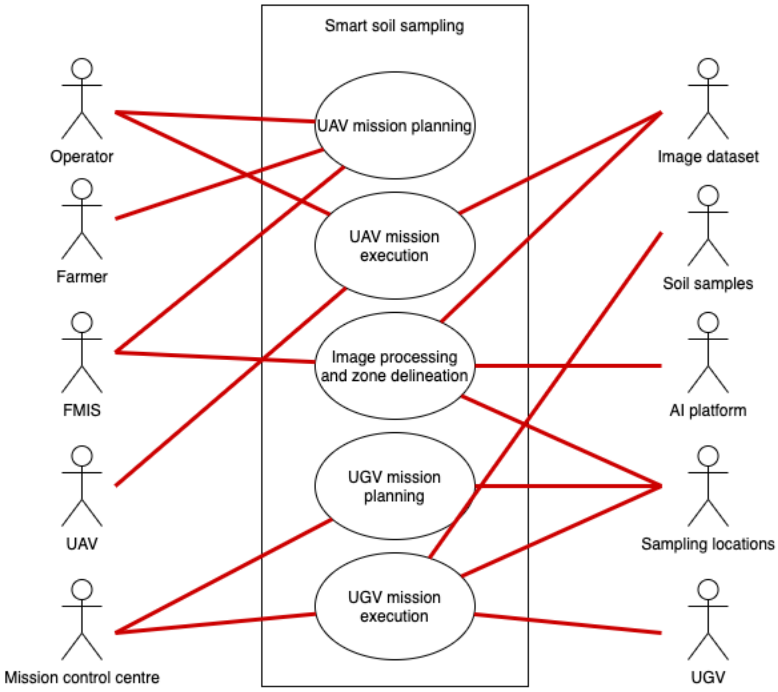


Figure 3 – Diagram of smart soil sampling use case

3.2.2.3 Technical Details

3.2.2.3.1 Actors

Actors			
Actor name	Actor type	Actor description	Further information specific to this use case
Operator	Human	Person controlling the UAV	Company's internal/ external operator
Farmer	Human	Farm owner/ agronomist	Person leading the production
FMIS	System	Internet platform/ app	AgroSense/ AgroSmart
UAV	Robot	Autonomous robot	UAV with multi or hyper-spectral camera
Mission Control Centre	System	Algorithms for routing and movement planning	Guidance through crops
Image dataset	Data	Raw images from UAV	Multi/ hyper-spectral images
Soil samples	Data	Chemical analysis of soil	N, P, K, pH…
AI platform	System	AI algorithms for image recognition	DL algorithms for clustering/ zone delineation
Sampling locations	Data	Centres of management zones	Typically, 12 locations per ha
UGV	Robot	UGV equipped with soil sampling module	Mobile lab on the robot

Table 19: Actors

3.2.2.4 Step by Step Analysis of Use case

3.2.2.4.1 Overview of the Scenarios

Scenario conditions						
No.	Scenario name	Scenario description	Primary actor	Triggering event	Pre-condition	Post-condition
1	Mission planning	Mission planning	Operator	Regular intervals	Unknown scanning area	UAV route/ height planned
2	Mission execution	Mission execution	UAV	Operator flies the UAV	No field images	Field images acquired

Scenario conditions						
No.	Scenario name	Scenario description	Primary actor	Triggering event	Pre-condition	Post-condition
3	Image processing and zone delineation	Clustering of pixels	AI platform	UAV images acquired	Unknown distribution of zones	Problematic areas mapped
4	UGV mission planning	Mission planning	Mission Control Centre	Sampling locations known	Unknown route	Route defined
5	UGV mission execution	Field sampling	UGV	Route defined	Unknown soil properties	Soil analysis conducted

Table 20: Scenario conditions

3.2.2.4.2 Steps – Scenarios

Scenario							
Step No.	Event	Name of process / activity	Description of process / activity	Service	Information producer (actor)	Information receiver (actor)	Information exchanged (IDs)
1	Regular scanning scheduled	Mission preparation	Getting field boundaries	GET	FMIS	Operator	UC2-01
2	Field boundaries set	Mission planning	Boundaries and flying height planned	SET	Operator	UAV	UC2-02
3	Mission kicked-off	UAV scanning	UAV flying over crops	RUN	UAV	Dataset	UC2-03
4	Images acquired	Image processing	Ortho-mosaicing, vegetation indices	GET	Dataset	AI platform	UC2-04
5	Images pre-processed	Zone delineation	Zones delineated and centroids (sampling locations) found	RUN	AI platform	Mission control centre	UC2-05
6	Sampling locations defined	Route planning	Route set according	SET	Mission control centre	UGV	UC2-06

Scenario							
Step No.	Event	Name of process / activity	Description of process / activity	Service	Information producer (actor)	Information receiver (actor)	Information exchanged (IDs)
			to locations				
7	Route planned	UGV mission execution	Sampling performed	RUN	UGV	Soil samples	UC2-07

Table 21: Scenario

3.2.2.5 Information Exchanged

Information exchanged		
Information exchanged, ID	Name of information	Description of information exchanged
UC2-01	Field boundaries	Shapefile of the field
UC2-02	Mission parameters	Mission height, area
UC2-03	Raw UAV images	Individual images from UAV
UC2-04	Pre-processed UAV images	Orthomosaic and vegetation indices
UC2-05	Management zone centroids	Homogeneous zones within the field
UC2-06	Route information	Trajectory of the UGV
UC2-07	Analysis of soil samples	N, P, K, pH…

Table 22: Information exchanged

3.2.3 Use case 3: Precision spraying

In order to perform precision spraying, we are using weed/ disease maps produced by UAV image processing. The UGV is then navigated to the problematic areas and sprays these areas only, thus reducing the amount of pesticide/ herbicide spent

3.2.3.1 Description of the use case

3.2.3.1.1 Name of use case

Use case identification		
ID	Area / Domain (s) / Zone (s)	Name of use case
3	WP6 – Pilot 3	Precision spraying

Table 23: Use case identification

3.2.3.1.2 Version management

Version management				
Version No	Date	Name of author (s)	Changes	Approval status
0.1	26.5.2021	Oskar Marko		

Table 24: Version management

3.2.3.1.3 Scope and Objectives of Use case

Scope and objectives of use case	
Scope	WP6 – Pilot 3
Objective (s)	1) Navigate UGV to problematic locations. 2) Spray the plants.
Related business case (s)	1. Reduction in the use of pesticides/ herbicides. 2. No need for human labour.

Table 25: Scope and objectives of use case

3.2.3.1.4 Narrative of Use case

Narrative of use case
Short description
Parts of the field with weeds and diseases are mapped and the UGV is sent to spray these areas.
Complete description
In order to perform precision spraying, we are using weed/ disease maps produced by UAV image processing. The UGV is then navigated to the problematic areas and sprays these areas only, thus reducing the amount of pesticide/ herbicide spent.

Table 26: Narrative of use case

3.2.3.1.5 Key Performance Indicators (KPIs)

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
1	Comprehensive monitoring of crops	Lower cost of soil sampling for 35%	-
2	Optimal and data-driven decision-making	25% less damages from sub-optimal pesticide application	Weeds/ diseased areas are sprayed as soon as the problems are detected, thus reducing the damages.

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
3	Targeted herbicide and pesticide application	40% cut in pesticide/ herbicide amount	The amount of required chemicals is cut drastically through timely and targeted spraying practice.

Table 27: Key performance indicators

3.2.3.1.6 Use case Conditions

Use case conditions
Assumptions
Blueberries are mostly grown in the open and not in the glasshouses.
Weeds and plant diseases pose severe problems in blueberry growing.
Soil sampling is necessary for optimal field management.
Prerequisites
Test fields defined in both Serbia and Lithuania.
Blueberries on the test field are healthy and old enough to give a representative yield.
Plant diseases visible on the leaves (in contrast to the root system or stems).

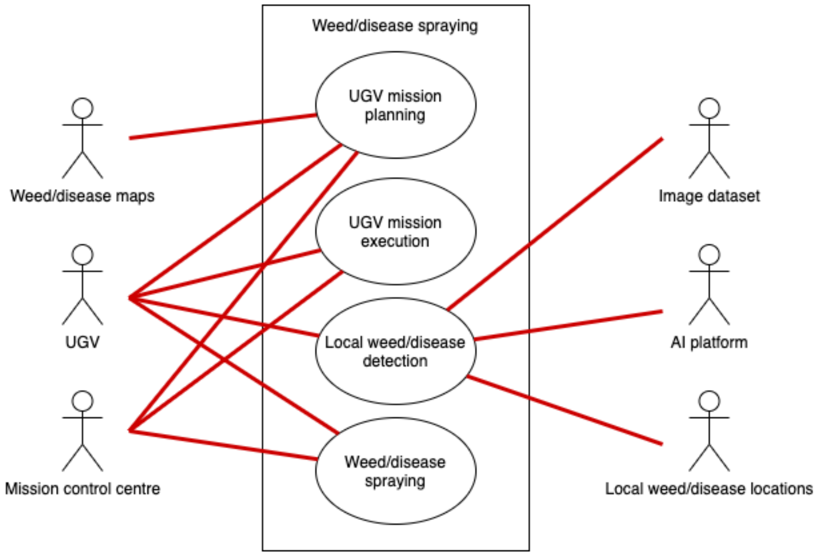
Table 28: Use case conditions

3.2.3.1.7 Further Information to the Use case for Classification / Mapping

Classification information
Relation to other use cases
Soil sampling is performed based on the UAV images processed in the Monitoring use case (Use Case 1).
Level of depth
Specific
Prioritisation
High
Generic, regional or national relation
Generic
Nature of the use case
Technical

Table 29: Classification information

3.2.3.2 Diagrams of use case

Diagram(s) of use case
 Figure 4 – Diagram of weed/disease spraying use case

3.2.3.3 Technical Details

3.2.3.3.1 Actors

Actors			
Actor name	Actor type	Actor description	Further information specific to this use case
Weed/disease maps	Data	Initial maps derived from UAVs	Problematic zones mapped
UGV	Robot	UGV equipped with soil sampling module	Mobile lab on the robot
Mission Control Centre	System	Algorithms for routing and movement planning	Guidance through crops
Image dataset	Data	Raw images from UAV	Multi/ hyper-spectral images
AI platform	System	AI algorithms for image recognition	DL algorithms for clustering/ zone delineation
Local weed/disease locations	Data	Weed/disease location extracted by a UGV camera	On-board localisation

Table 30: Actors

3.2.3.4 Step by Step Analysis of Use case

3.2.3.4.1 Overview of the Scenarios

Scenario conditions						
No.	Scenario name	Scenario description	Primary actor	Triggering event	Pre-condition	Post-condition
1	Mission planning	Mission planning	Mission control centre	Weeds/disease found in UAV images	Weeds/disease present on the field	Mission planned how to tackle them
2	Mission execution	Mission execution	UGV	UGV route planned	UGV not in field	UGV moves through the field
3	Local weed/disease detection	Weeds/disease detected from a UGV camera	UGV	UGV in the field	Unknown presence and location of weeds and diseases regarding the UGV	UGV presence and location known
4	Weed/disease spraying	Mission planning	UGV	Sampling locations known	Weeds and diseases growing undisrupted	Weeds and diseases sprayed

Table 31: Scenario conditions

3.2.3.4.2 Steps – Scenarios

Scenario							
Step No.	Event	Name of process / activity	Description of process / activity	Service	Information producer (actor)	Information receiver (actor)	Information exchanged (IDs)
1	Weeds and diseases detected from the UAV	Mission preparation	Getting the location of problematic areas	GET	Weeds and diseases maps	Mission control centre	UC3-01
2	Problematic areas known	Mission planning	Route setup	SET	Mission control centre	UGV	UC3-02

Scenario							
Step No.	Event	Name of process / activity	Description of process / activity	Service	Information producer (actor)	Information receiver (actor)	Information exchanged (IDs)
3	Mission kicked-off	Mission start	UGV passing through the field and detecting local weeds and diseases with its own camera	RUN	UGV	AI platform	UC3-03
4	UGV onboard camera images acquired	Detection of local weeds and diseases	Image processing	GET	AI platform	Local weed and diseases locations	UC3-04
5	Local weed and disease locations	UGV spraying	Sprayer directed towards weeds and diseases	RUN	Mission control centre	UGV	UC3-05

Table 32: Scenario

3.2.3.5 Information exchanged

Information exchanged		
Information exchanged, ID	Name of information	Description of information exchanged
UC3-01	Locations of weeds and diseases	Georeferenced locations
UC3-02	Route information	Trajectory through the field
UC3-03	UGV images	Images taken by the UGV's onboard camera
UC3-04	Segmented images	Images with weeds/diseases recognised
UC3-05	Precise location of weeds/diseases in 3D space	Location for navigating the sprayer

Table 33: Information exchanged

4 Description of datasets

For use cases 1 and 2, a dataset of multi/hyperspectral images using UAV was created following the procedure [1] recommended by DJI, the UAV producer. In Lithuania, DJI Matrice Pro with BaySpec OCI™-F hyperspectral camera was used for the pilot. Since the newly acquired Quantum system Trinity F90+ with MicaSense dual-camera arrived after the growing season, BIO's DJI P4 Multispectral drone was used for the first year of the Serbian pilot. Before BIO's researchers start actively using the new drone, they will receive extensive training from the producer, in which they will learn to set up missions and choose optimal flight parameters. In both pilot locations, a strict procedure was followed for image acquisition and data pre-processing.

4.1 Data Summary

The purpose of the data is to monitor the growth of blueberries and relate the observed parameters with natural phenomena. This has mostly to do with the occurrence of weed, plant diseases or changes in plants due to malnutrition or the lack of water. The dataset will be further processed using state-of-the-art machine/deep learning techniques and image processing tools.

The main types of data are:

1. UAV images
2. Ground conductivity maps
3. Satellite images
4. Sensory data (soil moisture / weather)
5. Soil samples
6. Field books
7. Plant-o-Meter readings

During M1-12 we focused on acquisition of UAV images and sensory data. UAV images are the key input for the pilot. After the image acquisition procedure, the following steps are conducted before the creation of orthomosaic:

- Multispectral images have phase differences caused by different multispectral camera locations and by different exposure times. They are aligned according to the selected reference image (NIR-band) using information stored in EXIF and XMP data of each multispectral image and using high-frequency information within the image (edges) within the Enhanced Correlation Coefficient (ECC) maximization method.
- The next steps after alignment are vignetting correction and distortion calibration, with the vignetting coefficients and calibration parameters stored in the XMP data of each image.

- After correction steps, the multispectral image is normalized using sensor gain and exposure time stored in XMP data and normalized black level value which is stored in EXIF data. Multispectral images are in 16-bit format; hence the normalized black level value is obtained after division with 65535.
- The final pre-processing step involves the correction of the obtained multispectral image from the previous step using the irradiance information from the sunlight sensor stored in the XMP data of each image.

Although most of the previously mentioned pre-processing steps can be done in Pix4D software during the creation of orthomosaic, the advantage of using a procedure implemented in Python 3 using OpenCV library is the possibility to properly determine which multispectral images might be problematic for the orthomosaic creation due to variation in lighting conditions during the acquisition phase. Therefore, an initial version of the method for detection and correction of uneven brightness of the acquired multispectral images based on overall image statistics is created. Figure 5 shows orthomosaics for all spectral bands with a GSD of 2 cm/pix. The orthomosaics presented are generated from the flight missions at the pilot sites, i.e., blueberries fields in Babe, Serbia.

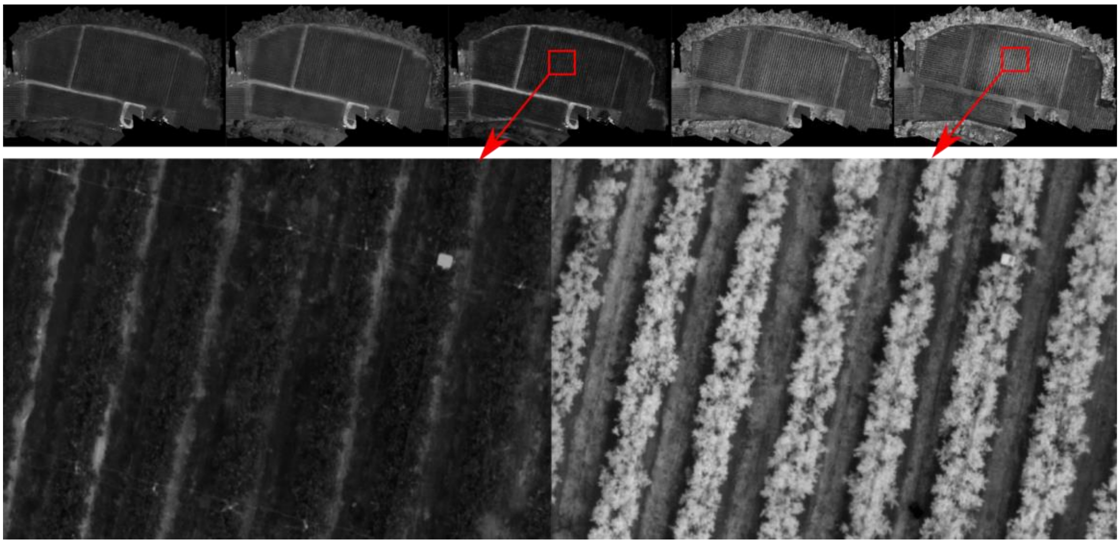


Figure 5: First row: Created orthomosaics for each spectral band in the following order from left to right: blue, green, red, red-edge and NIR. Second row: Zoomed ROI (denoted with a red rectangle) from red and NIR orthomosaic.

Hyperspectral images undergo multiple steps of pre-processing and analysis before the classification maps of chemical components are produced. Pre-processing and data analysis is based on comprehensive state-of-the-art denoising algorithms and classical classification methods like Neural Networks, SVM, etc. During the pre-processing stage the initial hyperspectral images are being cropped and denoised which implies the usage of substantial computing and time resources. The whole data pre-processing and analysis pipeline ran on Amazon Cloud Computing Services (AWS) while the Amazon S3 Cloud Storage was used as the main operational and archival data storage. This will be running on the FlexiGroBots platform, where pre-processing will be one of the common methodologies and services. Figure 6 shows an example of a raw and a pre-processed image.

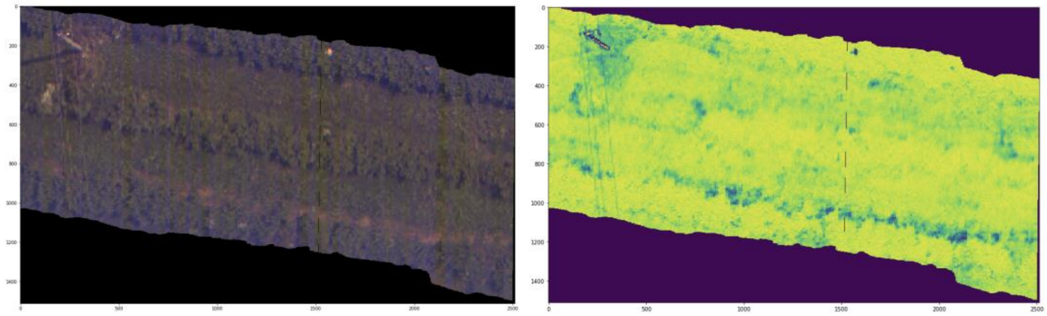


Figure 6: Raw and pre-processed hyperspectral images.

UAV image datasets include hyperspectral imaging data cubes where spectral resolution of images is 3 nm, spectral range – 400-1000 nm, pixel count - 1024*1280.

Specifications of the hyperspectral UAV dataset		
Flight date	Raw hyperspectral image count	Raw hyperspectral image data size
2021-07-07	48398	63,43 GB
2021-07-16	42321	55,47 GB
2021-09-15	59765	78,34 GB

Table 34: Specifications of the hyperspectral UAV dataset

Regarding the sensors, local weather data is crucial for optimal decision-making regarding irrigation, as the level of applied water and timing of application is directly correlated to the actual conditions on the field. In order to monitor the real-time data on the field, we installed a meteorologic station that monitors rainfall, solar radiance, UV light, air temperature and pressure. The station is accompanied by a soil moisture sensor for monitoring soil conditions in which the plants grow (Figure 7).

Figure 7: Weather station (left) and the LORA communication module (right)

Field sensors are connected to the cloud through a communication module. The communications module transmits/receives messages to/from the cloud through a LORA network. LORA stands for LOng RAnge communications and it is ideal for agricultural operations where fields can be located far away from the cable/optical internet and power supply. LORA signal has been provided in the region by BIO’s network of LORA base stations installed in collaboration with a Serbian telecom operator. The data is displayed in AgroSense, BIO’s farm management information system, as in Figure 8.

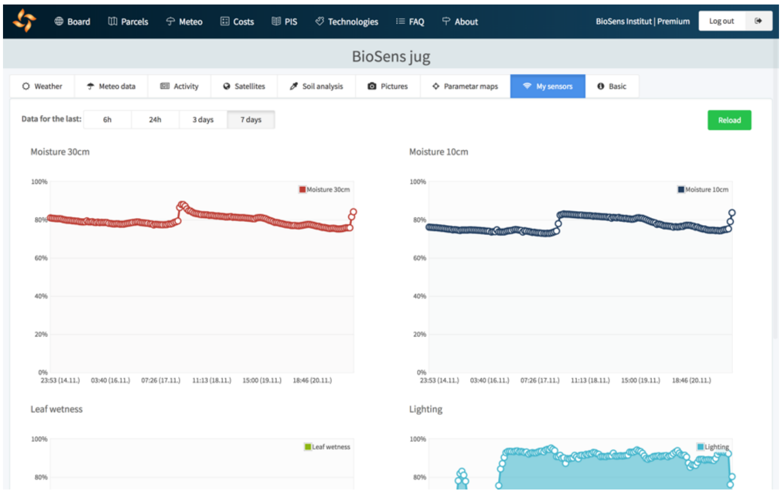


Figure 8: Sensory readings are displayed in the AgroSense platform

4.1.1 Data Acquisition

Describe how the data will be acquired (devices used, communication protocols, real-time or acquisition in batches, APIs…)

1. The data will be logged from the tablet used for UAV mission control and the resulting images will be downloaded as a batch from the tablet. 2. During the scanning with the ground conductivity probe, data is logged onto a laptop and saved as a single N38 file. 3. New satellite images will be downloaded as a batch as soon as they are published, through Sentinel Data Hub and its API. 4. The data will be logged from the field meteorological station and soil moisture sensors. Data will be acquired in real-time through the LORA network and stored in an internal database. 5. Soil samples will be retrieved from the lab or from the robot's soil analysis. In any case, they will be retrieved as a batch. 6. Farmers will send the researchers their field books as a batch. 7. Field scanning using the Plant-o-Meter, data acquisition in batches

4.1.1.1 Format and structure of the data

Describe what formats of data the Task will generate/collect (e.g., pdf, xlsx, SQL, etc.). What is the structure of the data within these files? If there are multiple files, how are they organised?

1. Tiff format. There will be a folder for each scanning date with all individual photos and a joint mosaic image. 2. N38 format. The Geonics EM38 MK2 probe logs the GPS location and ground conductivity measurements simultaneously and stores them in its own N38 format. The format, which resembles CSV with a heading is converted to shapefile / pandas data frame using BIO's internal Python procedure. 3. Tiff format. Sentinel images come in 12 different bands, with a separate monochromatic image for each band. All images from one date will be located together in the same folder. 4. CSV format. There are 7 parameters logged by the meteorologic station and 1 logged by the field sensor. These will be located in different columns and followed by a timestamp. 5. xlsx format with predefined columns for different substances in the soil. 6. xlsx format with separate columns for fertiliser and pesticide application, planting/harvest dates and the yield. One row denotes one season at one field. Different season-field combinations will be located in separate rows. 7. xlsx format with predefined columns for different spectral channels / vegetation indices

4.1.1.2 Existing data

Will the task re-use any existing data and how? (See also Chapter 7 of the CESSDA guide)

We have already made a preliminary drone scanning in Lithuania, and drone and soil conductivity scanning of the pilot fields in Serbia, to get familiar with the peculiarities of blueberries and the terrain on which they will be grown.

4.1.1.3 Origin of the existing data (if any)?

What is the origin of the data?

The existing data has been acquired by BIO's researchers during the project preparation phase, before proposal submission, while the subsequent data will be acquired during the project implementation.

4.1.1.4 Size of the data

What is the expected size of the data? (in Gb)

1. Drone images ~1TB 2. Ground conductivity maps ~50MB 3. Satellite images ~10GB 4. Sensory data (soil moisture / weather) ~10MB 5. Soil samples ~10MB 6. Field books ~10MB

4.1.1.5 Usage of the data

To whom it might be useful ("data utility")?

All would be useful for farmers to better understand the state of crops and the spatial variability within the field, as well as to map the problematic areas within the field.

4.1.1.6 Where and how are the data stored?

All data will be stored at BIO's and ART's servers.

4.1.1.7 What are the risks for the data?

1. Drone images: change in sunlight/cloudiness during the drone flight (mission)

2. Ground conductivity maps: EC probe battery flat, readings missing 3. Satellite images: cloud occlusion (missing data) 4. Sensory data (soil moisture / weather): connection loss (missing data) 5. Soil samples: bad lab results, lab analysis repeated 6. Field books: missing activities 7. Plant-o-Meter measurements: bad connection with mobile phone / errors with logging

4.2 Data management

4.2.1 FAIR Data (Findability, Accessibility, Interoperability, Reusability)

4.2.1.1 Making data FINDABLE

(Chapter 2 of CESSDA Guide)

4.2.1.2 Metadata provision

Are the data produced and/or used in the Task discoverable with metadata?

Yes

4.2.1.3 Identification of data

Does the task plan to make use of persistent and unique identifiers such as Digital Object Identifiers?

Yes

4.2.1.4 Naming conventions

What naming conventions do you follow? (*Folders and files conventions*)

Top-level - location name Middle level - Type of data Bottom level - Date of acquisition

4.2.1.5 Search keywords

Will search keywords be provided that optimize possibilities for re-use? If yes, please list the preliminary keywords

e.g., Tagging items (i.e., datasets, documents, codes, etc.) with relevant keywords that are automatically indexed by the search

Yes Blueberries, precision agriculture, AgTech, image processing, deep learning, machine learning, UAVs, UGVs

4.2.1.6 Versioning

Do you provide clear version numbers?

e.g., Might be taken in charge by a tool, might only upload one version of each dataset

Yes. We are using _vX ending, where X is the number of the version

4.2.1.7 Standards for metadata creation

What metadata will be created? In case metadata standards do not exist in your specific discipline, please outline what type of metadata will be created and how.

e.g., Description, ownership, date etc.

e.g., Standard e.g., Dublin Core metadata standard

e.g., readme.txt file

1. Source 2. Region 3. Owner / producer 4. Licence 5. Sensitive / personal data 6. Frequency of acquisition 7. Date / time span 8. Spatial resolution 9. Total size

4.2.1.8 Making data ACCESSIBLE

See Chapter 4 and Chapter 6 of CESSDA Guide

4.2.1.8.1 Open available datasets

Which data produced and/or used in the Task will be made openly available as the default?

Follow the principle "as open as possible, as closed as necessary".

We will make the drone images available to the scientific community, as they are the most important data that will be generated within the project.

4.2.1.8.2 Closed datasets

If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g., ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.).

Farmer's field books will only be shared with farmer's permission, and not by default.

4.2.1.8.3 Repository

How will the data be made accessible (e.g., by deposition in a repository)?

Can you provide us with the link and D.O.I. (digital object identifier)?

e.g., Deposited in open access repository (e.g., OSF, Zenodo)

DOI for datasets differs from DOI for publications. You might find some useful information under: https://academia.stackexchange.com/questions/52032/how-do-i-get-a-doi-for-a-dataset

If you do not have/or will obtain a DOI, just mention that here.

The data collected so far was deposited at public repositories (DOI: 10.5281/zenodo.5712792). As soon as new data is acquired, it will also be uploaded and assigned a link and DOI.

4.2.1.8.4 Software tools for access

What methods or software tools are needed to access the data?

No tools, data will be downloaded directly.

Is documentation about the software needed to access the data included?

No

Is it possible to include the relevant software (e.g., in open-source code)?

Yes

4.2.1.8.5 Data depository

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

e.g. Deposited in open access repository (e.g. OSF, Zenodo)

The dataset acquired in Serbia is ~24GB in size and has been published through the Zenodo repository at: https://zenodo.org/record/5712792#.YZtTUOTTUWM The dataset acquired in Lithuania is ~600GB in size and it is located at: http://gofile.me/5G1pV/pbaq96qc5

Have you explored appropriate arrangements with the identified repository?

Yes

If there are restrictions on use, how will access be provided?

No

Is there a need for a data access committee[1]?

No

4.2.1.9 Making data INTEROPERABLE[2]

See Chapter 3 and Chapter 6 of the CESSDA Guide

4.2.1.9.1 Interoperability of the data produced in the Task

Is the Task allowing data exchange and re-use between researchers, institutions, organizations, countries, etc. (i.e., adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

Yes

What data and metadata vocabularies, standards or methodologies will the Task follow to make the data interoperable?

e.g., metadata format is compliant with standard formats (MARCXML, Dublin Core, and DataCite Metadata Schema)

Metadata format is compliant with standard formats (MARCXML, Dublin Core, and DataCite Metadata Schema)

[1] A committee that reviews and authorizes requests for data access and use.

[2] Due to the versatility of the pilots, a number of standards should be used to ensure interoperability and proper communication, protection, and reusability of the generated data. These include standards such as ROS, ISOXML (ISO 11783), ISO 50001, ISOBUS (ISO 11783), ISO 22000, DIN EN 1672-2:2009-07, Machinery Directive 2006/42/EC, FDA 21CFR 174-178, EHEDG Doc. 8 / 13 (+ more depending on the application), ISO 22166, ISO 18497, ISO 17757, ISO 25119, ISO 62443.

4.2.1.9.2 Vocabulary

Will your task use standard vocabularies for the data types present in the datasets to allow inter-disciplinary interoperability?

We will use the common ontologies used in other published scientific datasets

If not, will you provide mapping to more commonly used ontologies?

-

4.2.1.10 Making data RE-USABLE

See Chapter 3 of the CESSDA Guide

4.2.1.10.1 License

How will the data be licensed to permit the widest re-use possible?

e.g., Creative Commons license CC-BY or CC-0 (according to the H2020 guidelines)

Be aware there are different licenses for research data (in comparison with publications), depending on the nature of these data (origin).

Creative Commons licence

4.2.1.11 Availability

When will the data be made available for re-use?

e.g., after upload

After upload

If applicable, specify why and for what period a data embargo is needed.

-

How long is it intended that the data remains re-usable?

e.g., at least 15 years, for the lifetime of the repository

At least 15 years, for the lifetime of the repository.

4.2.1.11.1 Third parties

Are the data produced and/or used in the Task usable by third parties, in particular after the end of the pilot/project? If the re-use of some data is restricted, explain why.

Yes

4.2.1.11.2 Data quality

How is the data quality assured? Are data quality assurance processes described?

We will acquire data using the producers' guidelines and our experience from previous research.

Note: Please note that making data accessible should be the standard/default, but when, for some data, this is not possible, then it should be clearly explained; examples of restricted data could be personal information or consortium confidential data – the latter might be accessible to all the partners but kept within the project for exploitation during some time; this one, when published, normally is made available.

4.2.2 Allocation of resources

See Chapter 1 of the CESSDA Guide.

4.2.2.1 Costs

What are the costs of making data FAIR in your project? And how will these be covered?

e.g., Long-term storage, journal open access costs etc.

e.g., project's budget

No additional costs

4.2.2.2 Responsibility for data management

Who will be responsible for data management in the Task?

Dr Oskar Marko

4.2.2.3 Costs and potential value of long-term preservation

What are the costs of long-term preservation? And who decides how and what data will be kept and for how long?

No costs for long-term preservation. The decision concerning data sharing was made by the Pilot coordinator and fellow researchers.

4.2.3 Data security

Datasets must be preserved/stored during and beyond the lifetime of the Task. This means that each Task must provide a clear description of procedures for short-term and long-term preservation of the datasets.

See Chapter 4 of CESSDA Guide.

4.2.3.1 Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

e.g., data stored in the partners' networks with backups, firewall; in the project's SharePoint accessible with credentials; Basecamp; OwnCloud etc.

Data stored on the server with backup and firewall.

4.2.3.2 Data storage

Is the data safely stored in certified repositories for long term preservation and curation?

Yes

4.2.4 Ethical and legal aspects

In those cases where personal data (GDPR) is involved, detailed information is planned to be provided:

1. On what personal data is collected, stored and processed;
2. On the recruitment process, inclusion/exclusion criteria for participation;
3. On privacy/confidentiality and the procedures that are implemented for data collection, storage, access, sharing policies, protection, retention and destruction during and after the project;
4. On how informed consent is pursued;
5. if application/is needed to be filed with a local/institutional ethics review body (if personal data is being collected) and if yes, which bodies / where / when.

See Chapter 5 of CESSDA Guide.

4.2.4.1 Ethical or legal issues

Are there any ethical or legal issues that can have an impact on data sharing?

e.g., data from 3rd-party that didn't give explicit consent, data that need to comply with the GDPR etc.

In a project like FlexiGroBots, the following ethical issues might arise:

- *privacy and surveillance*
- *data ownership or the right to access data*
- *responsibility for decisions and their consequences (including accidents)*

Farmers' data will not be shared without their consent. Data from ZEL will be given to BIO according to the NDA that will be signed between parties.

4.2.4.2 Data collection in non-EU countries

Does your task involve data collection in non-EU countries[3]?

Yes

In case it does, please specify: Which data are collected in non-EU countries?

A part of data regarding WP6 was collected in Serbia.

Is the research conducted legally in at least one EU Member State[4]?

Yes

4.2.4.3 Data transfer to non-EU countries

Does your task involve a data transfer to non-EU countries[5]?

Yes

In case it does, please specify which data are transferred to non-EU countries.

Drone and sensory data from Lithuania may be transferred to Serbia for research purposes (and vice versa).

4.2.4.4 Personal data

Does your data collection involve the collection of personal data or data that can be traced back to whom it is about? In case not, the remainder of section 5 can be skipped.

Yes

4.2.4.5 Personal data: Information provisions and access

Do you comply with the GDPR concerning information provisions and access to personal data (right to be informed, right to access and informed consent[6] for data sharing and long-term preservation included in questionnaires is given by data providers)?

Yes

[3] EU's ethics requirements apply to all EU-funded research, irrespective of where it takes place. Similarly, the GDPR applies to all data-processing operations conducted by data controllers based in the EU, irrespective of where the processing takes place. This means that, even if you are collecting personal data outside the EU, you must still ensure and be able to demonstrate compliance with EU law.

[4] For activities carried out outside the EU, it is not enough for the activity to be accepted and comply with the legal obligations of a non-EU country; the activities must ALSO be allowed in at least one Member State.

[5] In case personal data is transferred from the EU to an organization located in a non-EU country, then such transfer must comply with Chapter V of the GDPR and must be submitted as a deliverable (https://gdpr-info.eu/chapter-5/)

[6] A template for informed consent can be found on the OwnCloud…

4.2.4.6 Personal data: Rectification and erasure

Do you comply with the GDPR concerning rectification and erasure of personal data (rights to rectification, erasure, restriction of processing, to be notified and data portability)?

Yes

4.2.4.7 Personal data: Right to object and automated individual decision-making

Do you comply with the GDPR concerning the right to object and automated individual decision-making?

Yes

4.2.4.8 Personal data: Data controllers and processors

Do you comply with the GDPR responsibilities for data controllers and processors (the controller and the processor have implemented appropriate technical and organizational measures to ensure a level of security appropriate to the risk and keep records of its processing activities)?

See Chapter 4 of CESSDA Guide.

Yes

4.2.5 Other issues

4.2.5.1 Other procedures for data management

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

See also https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/5.-Protect/Processing-personal-data/Diversity-in-data-protection

No

4.2.5.2 Dissemination of research results

Indicate how the 'research results' of the Task are communicated/disseminated to relevant European and global channels such as OpenAIRE, CIARD, GODAN and Big Data Europe.

The research results will be published in scientific journals.

4.3 Dataset catalogue

1. Drone images

Name	Drone images
Source	UAVs
Region	Serbia, Lithuania
Owner / producer	BIO, ART
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	5 times per season
Date / time span	March - September (2021-2023)
Spatial resolution	2-10 cm
Total size	~1TB

2. Ground conductivity maps

Name	Ground conductivity maps
Source	EC probe
Region	Serbia
Owner / producer	BIO, ART
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	1 time per 5 years

Date / time span	Autumn/winter
Spatial resolution	10m
Total size	100MB

3. Satellite images

Name	Satellite images
Source	Sentinel-2
Region	Serbia, Lithuania
Owner / producer	ESA
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	1 time per 5 years
Date / time span	Spring/Summer 2021-2023
Spatial resolution	10m
Total size	100GB

4. Sensory data (soil moisture / weather)

Name	Sensory data
Source	Field sensors (IoT)
Region	Serbia, Lithuania
Owner / producer	BIO, ART
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	1h
Date / time span	2021-2023
Spatial resolution	Point
Total size	100MB

5. Soil samples

Name	Soil samples
Source	Lab
Region	Serbia, Lithuania
Owner / producer	BIO, ART
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	Yearly (or more frequent)
Date / time span	2021-2023

Spatial resolution	12/ha
Total size	1MB

5. Field books

Name	Field books
Source	Farmer
Region	Serbia, Lithuania
Owner / producer	Zeleni Hit
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	Yearly
Date / time span	2021-2023
Spatial resolution	Field level
Total size	1MB

6. Plant-o-Meter readings

Name	Plant-o-Meter readings
Source	Plant-o-Meter
Region	Serbia, Lithuania
Owner / producer	BIO
Licence	Creative Commons Attribution
Sensitive / personal data	No
Frequency of acquisition	Yearly
Date / time span	2021-2023
Spatial resolution	Field level
Total size	1MB

5 Description of Robots, Devices & Platforms

5.1 Pilot Implementation Architecture

Pilot 3 will showcase the potential of novel robotic solutions for blueberry farming. A new multifunctional and modular robotic platform will be demonstrated during the pilots, with the potential to be utilized for different tasks and at various blueberry growth stages. The solution will consist of:

- **Autonomous ground robotic components** developed by BIO, that will be launched and demonstrated in a real-life blueberry farm operational environment.
- **Soil sampling modular tool**, mounted on the robotic platform, will enable farmers to conduct automated, efficient and on-site soil testing operations that are regularly needed at blueberry farms. The module will offer georeferenced data on N, P, K content in the soil within a 15-20-minute timeframe and upload this data to relevant decision-support tools.
- **Precision spaying modular tool** will be mounted on the robotic platform during fertilization and pesticide or herbicide application periods. Based on field monitoring and decision-support data, the robot will be able to move to problematic areas of a blueberry field and apply agrochemicals precisely, thus reducing farmers' expenses and contributing to global sustainability goals.

Besides the UGV components, the following software parts will be developed:

- **Image segmentation algorithms** based on deep learning. These algorithms will be used for processing UAV images and mapping the problematic areas, those affected by weeds and diseases.
- **Sampling points localisation** tool, which chooses appropriate positions for UGV-driven soil sampling. These points will be the centroids of field management zones that have homogeneous soil.

The high-level planning for the pilot components is detailed in D2.7.

5.2 UGV

5.2.1 UGV Modules

In Pilot 3 regarding UGV, there are three use cases for the application of the UGV system:

- UAV & UGV for disease detection.
- UGV sprayer.
- UGV soil sampling and analysis.

For each of these use cases, a common platform for each system will be UGV, except in the first use case where UAV will also support the task. For each use case, a dedicated module will be developed, which will be carried by the UGV.

Progress in all use cases will be presented here.

5.2.2 UAVs and UGVs for disease detection

In collaboration with UAV, whose task is to do an aerial detection of possible zones impacted by diseases or weeds, the UGV will be sent to those locations to do an infield weeds/diseases detection. Despite being a very important use case, this is the least challenging task since the module does not involve moving mechanical parts and physical interaction with the environment. It rather incorporates optical inspection of the field, plant, and soil. The module will include an in-house developed device (Plant-O-Meter) for measurement of well-known vegetational indices (like NDVI, SR, DGCI, SIPI, RGR, Ari, etc.), and Intel RealSense Depth Camera D435.

The Plant-O-Meter is initially developed by BIO, up to a TRL6 level. With our collaboration industry partner Bitgear and supported by the Innovation Fund of the Republic of Serbia, the Plant-O-Meter is now a commercial product. More detailed information can be found here: https://dizajn.subotica.ws/plantometer/index.html#.

The latest release of Plant-O-Meter is shown in Figure 9.

Figure 9: The latest release of Plant-O-Meter

5.2.3 UGV Spraying

The UGV sprayer module is still in initial development. Currently, we are brainstorming for a mechanical setup that will be the most promising for the given task. We are also considering a solution for an ultra-low volume (UVL) herbicide applicator.

5.2.4 UGV Soil Sampling

Most of our effort was spent on the UGV soil sampling module. We have good progress in designing the mechanical part of the sampler, which is summarized in the following images.

The main module is divided into submodules. Each submodule is colour coded. These submodules are as follows:

1. Wheeled platform
2. Guide rails with powertrain for lateral movement
3. Main supporting frame
4. Tilting platform
5. Probing/sampling assembly

Figure 10 illustrates UGV for soil sampling divided into submodules. The front and top views are shown as well.

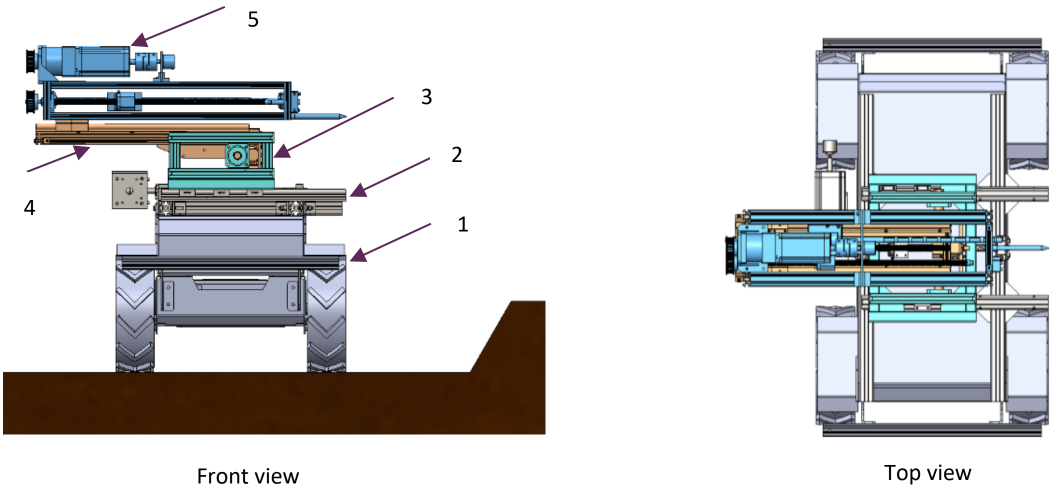
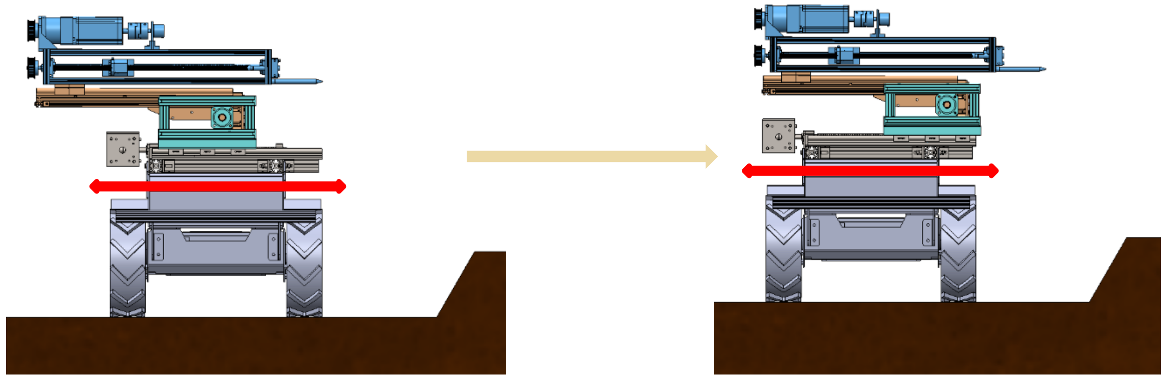


Figure 10: UGV with subdivided sampling module.

The probing/sampling submodule (5) is designed in such a way that its relative position can be adjusted in the horizontal plane (lateral movement), which is illustrated in Figure 11. For that action, the Guide rails with a powertrain for lateral movement (submodule 2) is used. Such extending of the probing/sampling submodule in relation to guide rails is foreseen to compensate for distance variations between the UGV platform and plant row for any reason that may occur (plant treetop size, physical obstacles, uneven terrain etc.)

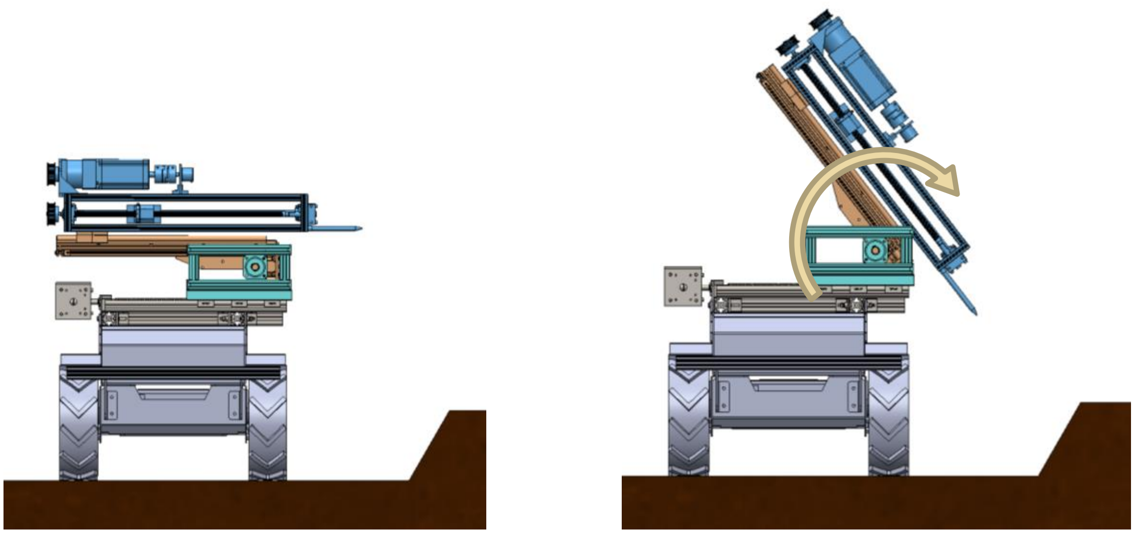


Retracted frame position

Extended frame position

Figure 11: Horizontal slide guides for sampling submodule

The probing/sampling submodule can also be moved in front, or vertical plane, via rotary movement (Figure 12). Tilting whole probing/sampling assembly gives provides benefit of fine adjustment of the angle at which the probe and the sampling tool will be entering the soil.



Horizontal position

Tilted position

Figure 12: Rotation of the sampling submodule

The probing/sampling submodule consists of two actuated linear slides. The bottom slide carries the top slide. The top slide actuates the sampling probe. This way a compact design is accomplished for the purpose of sampling. Furthermore, with the actuation of the bottom slide, along with the top slide, two probes are also moved. These probes are for soil electrical conductivity (EC) measurement, and soil pH measurement. Figure 13 illustrates a situation where only a bottom slide is moved, while the top slide is in contracted state. The detail depicted in the figure shows penetration of EC and pH probe (due to the angle of the picture, only one probe is visible). Such configuration enables independent EC and pH measurement from soil sampling and analysis.

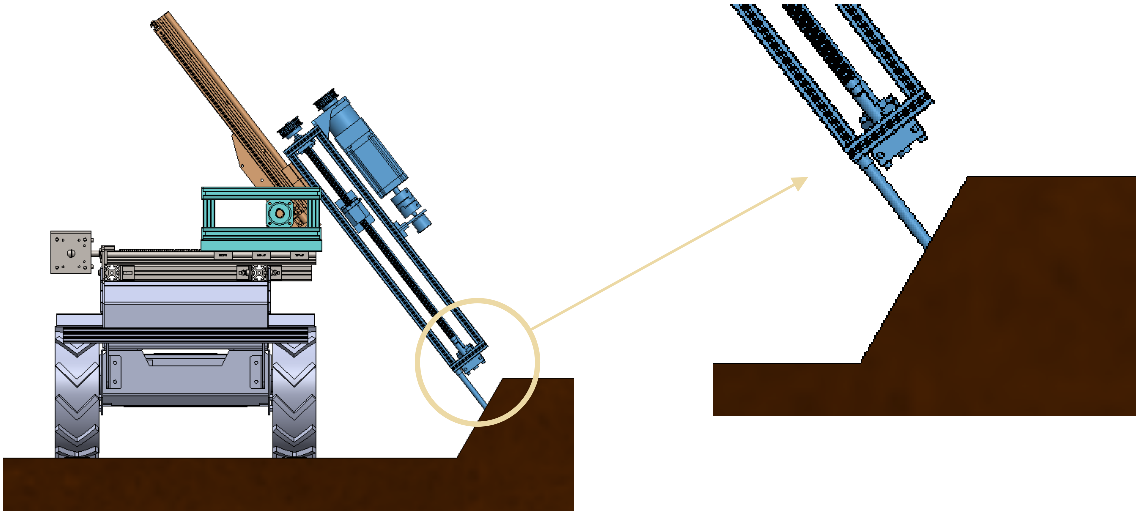


Figure 13: Illustration of probing/sampling submodule for probing

With the extension of the top slide, a soil sampling probe is engaged to take soil samples. This is illustrated in Figure 14. It depicts the situation where both, EC and pH probes, as well as soil sampling probe are inserted into the soil.

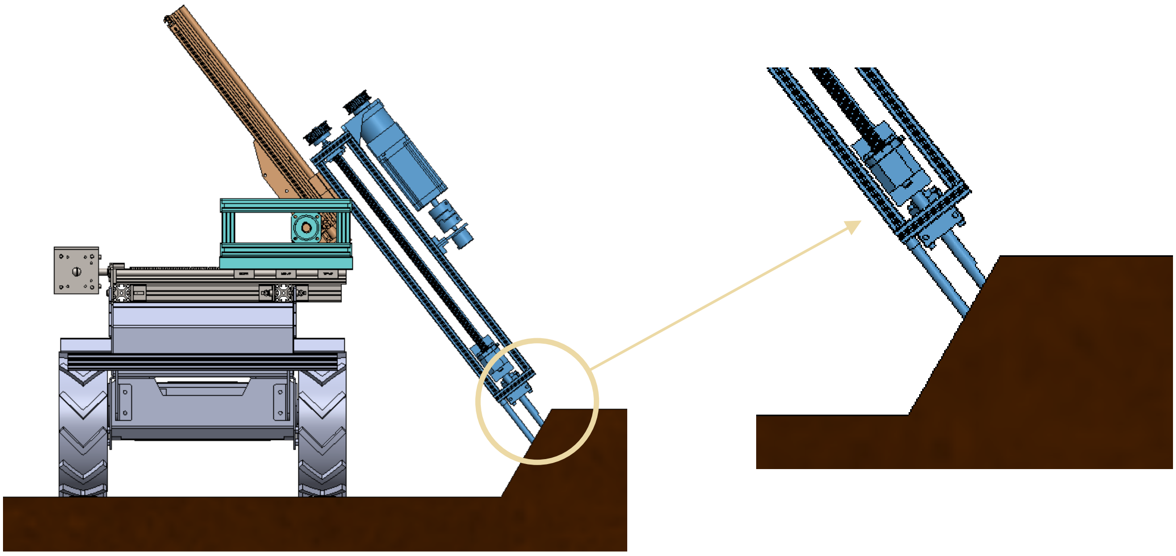


Figure 14: Illustration of probing/sampling submodule for sampling

All above illustrations are simplified for the purpose of clear presentation, they incorporate much more elements.

Besides the presented submodules, one important module that remains to be designed is a minilab. The minilab will collect the soil sample, and with added deionised water, will mix and prepare the sample for the soil nutrient measurement. The sample will be measured with multi-ion selective probe.

As mentioned previously, the major sensors/probes included in this design are:

1. Probe for soil EC measurement.
2. Probe for soil pH measurement.
3. Probe for soil nutrient measurement.

Based on project requirements these probes are carefully selected and acquired.

For EC measurement a rk500-23 probe from Hunan Rika Electronic Tech Co., Ltd was selected (Figure 15) [2]. It is a robust probe that can penetrate easily into the soft soil. The range for the probe is selected to be 2000µS/cm, which will be sufficient for the expected soil type. The output of this probe is current signal from 4 to 20mA, providing substantial noise immunity. Currently, there is ongoing activity to set up an interface and calibrate the probe.

Figure 15: rk500-23 probe

For pH measurement, a Wellinq's LanceFET probe without handle is selected (see Figure 16) [3, 4]. It is a robust probe that can penetrate easily into the soft soil. It covers full range pH measurement (0 to 14). It provides an analogue output signal.

Figure 16: Wellinq's LanceFET pH probe

With the consultation from experts of ZEL, an appropriate probe for soil nutrient measurement is selected. It is purchased as a kit (CleanGrow Multi-ion Nutrient Analyzer kit, Figure 17) [4]. This probe is a multi-ion-selective probe capable of measurement of up to 6 different nutrients (ions). The selected option for individual probe configuration is NPK2 ([Ca2+, K+, Mg2+, NH4+, NO3-, P [HPO42-]). It has a wireless BLE interface. Further details can be found in [3].

Presently, there is ongoing activity to make interface and calibration of the probe. A BLE communication with the probe is implemented.

Figure 17: CleanGrow Multi-ion Nutrient Analyzer kit

5.3 UAV

The UAV quadcopter type DJI P4 Multispectral platform (SZ DJI Technology Co., Shenzhen, China) was used for the image acquisition of the selected field (Figure 18). The UAV equipped with a LiPo 4S battery with a capacity of 5850 mAh, and 15.2V, provides a maximum flight time of 27 minutes and up to 7 km transmission range.

Figure 18: Data collection equipment: Remote Controller, DJI P4 Multispectral platform and DJI D-RTK 2 mobile station. Source: https://droneshopperth.com.au/product/dji-phantom-4-rtk-combo/

DJI Phantom 4 with multispectral camera array and one RGB camera (all with 1/2.9" (4.98mm x 3.74mm), CMOS sensors) was used for the creation of dataset for use cases 1 and 2 (Figure 19). Multispectral camera array consists of 5 cameras which cover the Blue, Green, Red, Red Edge, and Near-Infrared bands – all at 2 MP with global shutter, on a 3-axis stabilized gimbal. The spectral band is defined as follows Blue (B): 450 nm ± 16 nm; Green (G): 560 nm ± 16 nm; Red (R): 650 nm ± 16 nm; Red edge (RE): 730 nm ± 16 nm; Near-infrared (NIR): 840 nm ± 26 nm, where the first number denotes the central wavelength of the filter while the second is ½ of spectral bandwidth Full Width-Half Maximum (FWHM).

Figure 19: Camera system on DJI P4 Multispectral drone. Source: https://talosdrones.com/product/dji-phantom-4-multispectral/

The UAV is equipped with the integrated spectral sunlight sensor on top which captures solar irradiance used to maximize accuracy and consistency of data collection through different times of the day. On the top of the sensors are mounted lenses with the following characteristics: FOV (Field of View): 62.7°, focal length: 5.74 mm (35 mm format equivalent: 40 mm), autofocus set at ∞ and aperture: f/2.2. Image size is 1600×1300 pixels and the RGB image is stored in JPEG format while multispectral monochromatic images are stored in TIFF uncompressed format. The image position accuracy is enhanced by the Real-Time Kinematic (RTK) through the connection of the P4 Multispectral to the D-RTK 2 High Precision GNSS Mobile Station and NTRIP (Network Transport of RTCM via Internet Protocol) (Fig 15). DJI D-RTK 2 mobile station provides real-time corrections and centimetre-level accuracy ensuring that each photo uses the most accurate metadata. The relative positions of the centres of the six cameras' CMOS and the phase centre of the onboard D-RTK antenna have been calibrated and are recorded in the EXIF data of each image. This data is used in the pre-processing step before the creation of the orthomosaic.

In November 2021, the purchased VTOL Quantum Trinity F90+ arrived with a double camera system that provides pixel aligned images from 10 spectral bands in blue, green, red and near-infrared parts of the spectrum with bandwidths shown in Figure 20.

Figure 20: Quantum system Trinity F90+ with MicaSense dual camera kit and Sony UMC RGB camera. Source: https://store.terrestrialimaging.com/MicaSense-Dual-Camera-Complete-Payload-Kit-For-Quantum-Systems-Trinity-F90_p_1135.html

The selection of central wavelengths for the mentioned spectral bands allows for direct comparison between satellite (Sentinel 2A and Landsat 8) and UAV image data. The spectral bands are defined as follows: Coastal blue* 444 nm ± 28 nm, blue 475 nm ± 32 nm, green* 531 nm ± 14 nm, green 560 nm ± 27 nm, red* 650 nm ± 16 nm, red 668 nm ± 14 nm, red edge* 705 nm ± 10 nm, red edge 717 nm ± 12 nm, red edge* 740 nm ± 18 nm, NIR 842 nm ± 57 nm, with sensor resolution**:** 1280 x 960 (1.2 MP per EO band), where bands with * are from RedEdge-MX-blue camera while others are from RedEdge-MX camera. Apart from this, thanks to an easy payload swapping, the Sony UMC-R10C with 20.1 PM sensor and 16 mm fixed focal length (Figure 21) will be available for collecting high-resolution images in visible (RGB) spectrum.

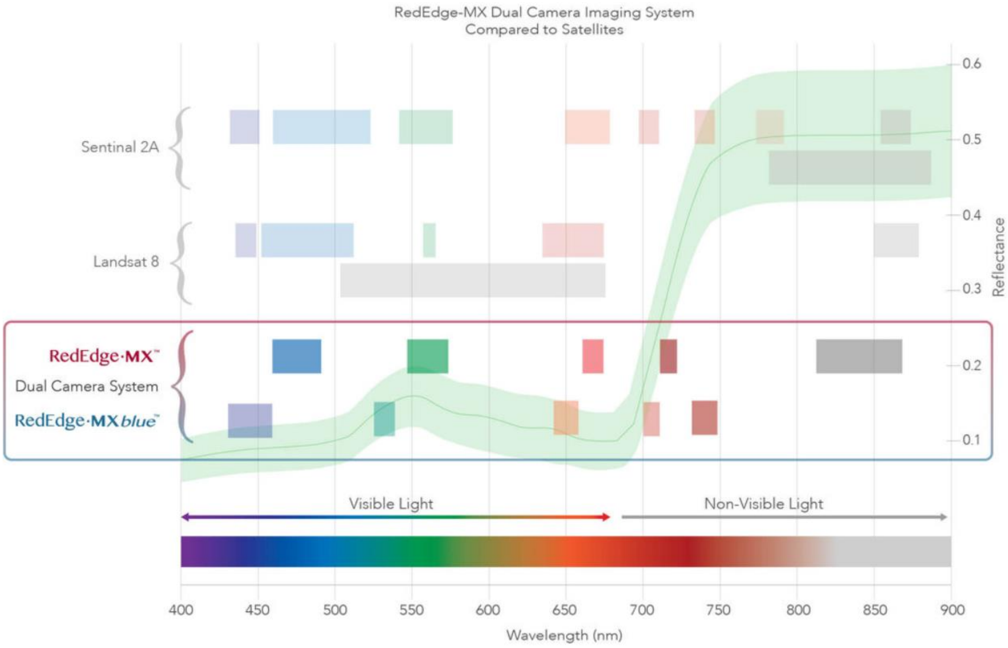


Figure 21: Bandwidths of the dual-camera system (RedEdge-MX and RedEdge-MX blue) spectral bands. Source: https://store.terrestrialimaging.com/MicaSense-Dual-Camera-Complete-Payload-Kit-For-Quantum-Systems-Trinity-F90_p_1135.html

Equipment used in Lithuania is specified in detail in Table 35 and shown in Figure 22.

Figure 22: A hyperspectral camera with computer hardware and software is used.

DJI Matrice 600 Pro UAV	BaySpec OCI™-F camera
· Distance between opposite rotors: 1133 mm; · Size: 1668 mm × 1518 mm × 727 mm; · Weight (with batteries): 10 kg; · Maximum load weight: 5.5 kg; · Maximum allowable wind speed: 8 m/s; · Maximum height: 2500-4500 m (depending on the propellers used); · Maximum speed: 65 km/h; · Unladen flight time: 38 min; · Flight time with 5.5 kg cargo: 18 min; · Flight control system: UgCS ; · Engines: DJI 6010; · Propellers: DJI 2170R; · Operating temperature: -10° C to 40° C; · Frequency of control equipment: 5.725-5.825 GHz; 2.400-2.483 GHz; · Maximum control signal distance in European Union countries: 3.5 km; · Batteries: 4500 mAh. 22.2 V. LiPo 6S, 99.9 Wh.	· Camera spectral range 400-1000 nm; · Spectral resolution – 3 nm; · Frame per second – 50 fps · Spatial Pixels – 800 px X scan-length · Exposure Time – 20 µs – 1 s · Wavelength Calibration – Factory calibrated (calibration fixed permanently) · Objective Lens Interface – C-mount · Operating Temperature – 0°C to 50°C · Weight - ~570 g (including standard lens) · Size – 14 cm x 7 cm x 7 cm (including standard lens
Computer Hardware	Portable Energy (Power Bank) Energizer XP8000A
· Dimension: 4.5 in x 3.5 in. x 1.1 in. · Weight: 1.1 lb. · Power input: 19v (battery pack optional) · Consumption: Peak 50w, typical < 10W · CPU; Intel Core i5 · Memory: 8 G (upgradable) · Storage: 250 (upgradable) · Operating System: Windows 7 Pro · Software: SpecGrabber and CubeCreator · Additional Ports: USB 3.0, HDMI · Monitor with HDMI	· Input: DC19V · Output: DC16V-20V · USB Output · Output: DC9V-12V · Capacity 8000mAh / Max 30Wh 

Table 35 Lithuanian equipment characteristics

5.4 Farm management information systems

Farm management information systems (FMIS) used in this project will be AgroSense and AgroSmart. AgroSense is the Serbian FMIS developed in 2017, which is being used by more than 20k farmers, whose total area equals ¼ of the total Serbian farmland. The weather station and soil moisture sensor were integrated with this platform so that the farmer (Zeleni Hit) could access the information in real time. AgroSense also has a module for UAV image processing (Figure 23) that will be enriched with novel algorithms developed within FlexiGroBots.

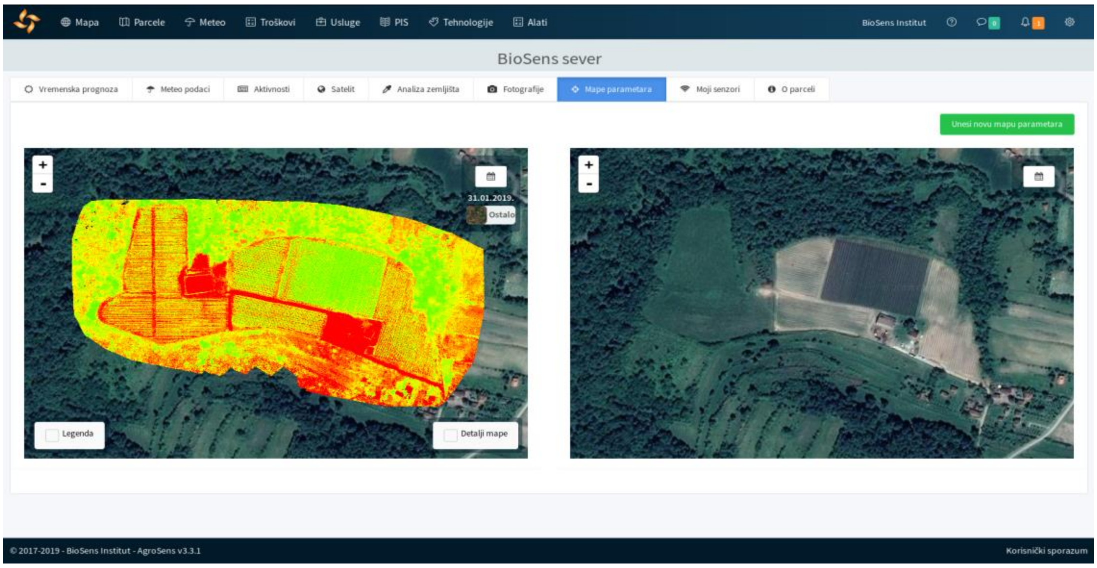


Figure 23: Blueberry field in AgroSense

Agrosmart is a farm management information software designed to help farmers in their daily operations. It is the most used farming software in Lithuania with 1k users, which employ more than 2,700 individual cultivation technologies on 15,000 different fields with a total managed area of more than 120,000 hectares. Some of the functionalities are the following:

- Process satellite data (NDVI and LAI).
- Simulate your future farm vision and help make investment decisions.
- Improve farm’s economic efficiency by increasing productivity and reducing costs.
- Control farm’s operational information in one place easily quick.
- Provide reports to the state bodies, declare your crop fields, according to an agreement exchange data with the private agricultural sector participants.

5.5 Detailed Flight Planning

A selection of flight altitude, during the flight mission planning in the software for DJI P4 Multispectral drone or Quantum Trinity 90+, will influence the distance between two consecutive pixels centres measured on the ground, named Ground Sampling Distance (GSD). The higher altitude of the UAV flight means bigger GSD i.e., the lower the spatial resolution of the acquired image and the less visible details within it. For overcoming possible issues regarding image acquisition with low resolution, first, a minimum size parameter of the objects, which need to be imaged is defined. A necessary flight altitude HF(m)with desired GSD_d(cm/pixel) can be obtained through the inverse equation given in a tool from Pix4D support link [5]:

HF(m) = (GSD_d(cm/pixel) * Focal_length(mm) * Image_width(pixels))/
(Sensor_width(mm)*100)

Furthermore, a double grid flight mission with an altitude between 40-50 m will be conducted. This altitude will ensure a GSD of 2.6-2.1 cm/pix. One flight should have a trajectory of image acquisition parallel while the other should have an orthogonal position to the blueberry rows (see Figure 24). Both horizontal and vertical flight path is needed for a better representation of information within orthomosaics (Figure 25). Images taken from multiple sides, with sufficient overlap of a minimum of 75% (front and side lap) are required for optimal processing during the creation of orthomosaic.

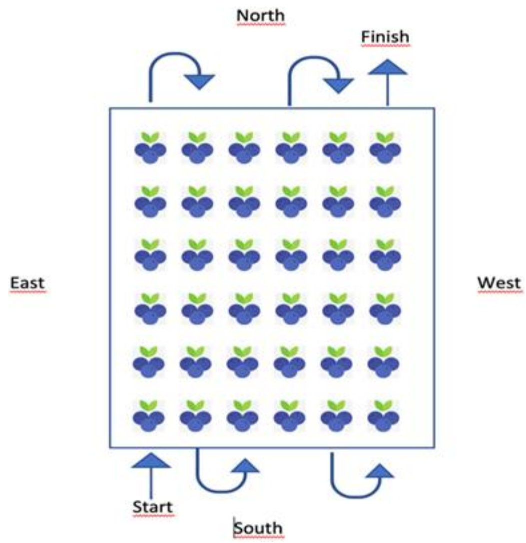


Figure 24: Field scanning procedure (vertical flight path)

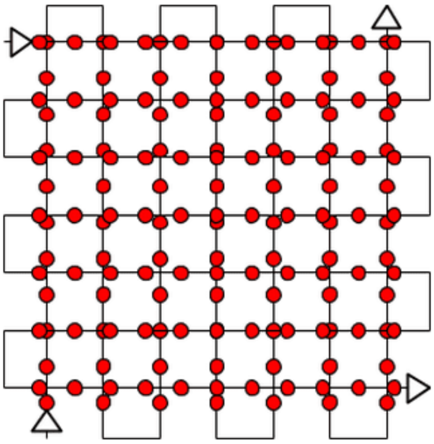


Figure 25: Field scanning procedure (horizontal flight path overlaid)

An example of such a flight plan is given in Figure 26.

Figure 26: Horizontal and vertical flight plans

For the RTK (Real-Time Kinematic) correction of coordinates is happening while the UAV is in the flight mission. However, in the case of PPK (Post Processed Kinematic), the actual processing of that data occurs after a UAV flight is complete. Besides the RTK, PPK is used to make additional corrections before creating a final orthomosaic, if needed. Ground Control Points (GCPs) are marked points on the ground that have a known geographic location. GCPs are required in aerial imaging because they enhance the positioning and accuracy of the mapping outputs and reduce the noise. They should be placed on the ground and be clearly visible in the aerial imagery for example by using high contrast colours. It is also important that the GCPs are evenly distributed over the whole mapping area. These coordinates can be captured via base station or obtained by other sources such as LIDAR, older maps of the area, web map services. Known GCPs should be included in the initial processing of the images right before creating the final orthomosaic.

From the created dataset of orthomosaic images per spectral band, the next step is the segmentation of the region of interest (ROI) within orthomosaic. With the information about the parcel border (polygon) in the form of a shapefile (vector data), provided by the farmer, a mask that denotes pixels within the parcel is created. Given GPS coordinates of points within the polygon from the shapefile are used for the creation of the parcel contour in the raster (image) format. Then all the pixels surrounded by the contour are masked. Further, the masked region within the parcel is divided into small images for more efficient processing. From these images, different maps of vegetation indices (NDVI, GLI, ExG, SAVI dodati reference) and normalized versions of available spectral bands are calculated. Initial semantic segmentation of vegetation and soil (foreground/background binary segmentation) is done using Otsu's method with the supervised tuning of the estimated threshold for reducing the false positives (pixels which are denoted as vegetation while they belong to the soil). These masks will be further regularized in a semi-supervised approach and such that will be used during the learning procedure of the deep neural networks method for semantic segmentation. One of such regularized masks obtained using Otsu's method is presented in Figure 27. Then using an algorithm for finding connected components within the segmented image, coordinates of the pixels that belong to the connected component with the biggest area are obtained. Using these coordinates of the pixels, an orientation of the row is determined using principal component analysis (PCA) and according to the angle of the PC-axis with respect to the original XY axis, the segmented image is rotated and projected into a 1D signal by summing the image values through columns (Figure 27). Estimation of the peaks in created 1D signals reveals the coordinates (in x-axis) of rows which are further used for the creation of bounding rectangles around each row in the segmented image. With this identification of each row is obtained the instance segmentation.

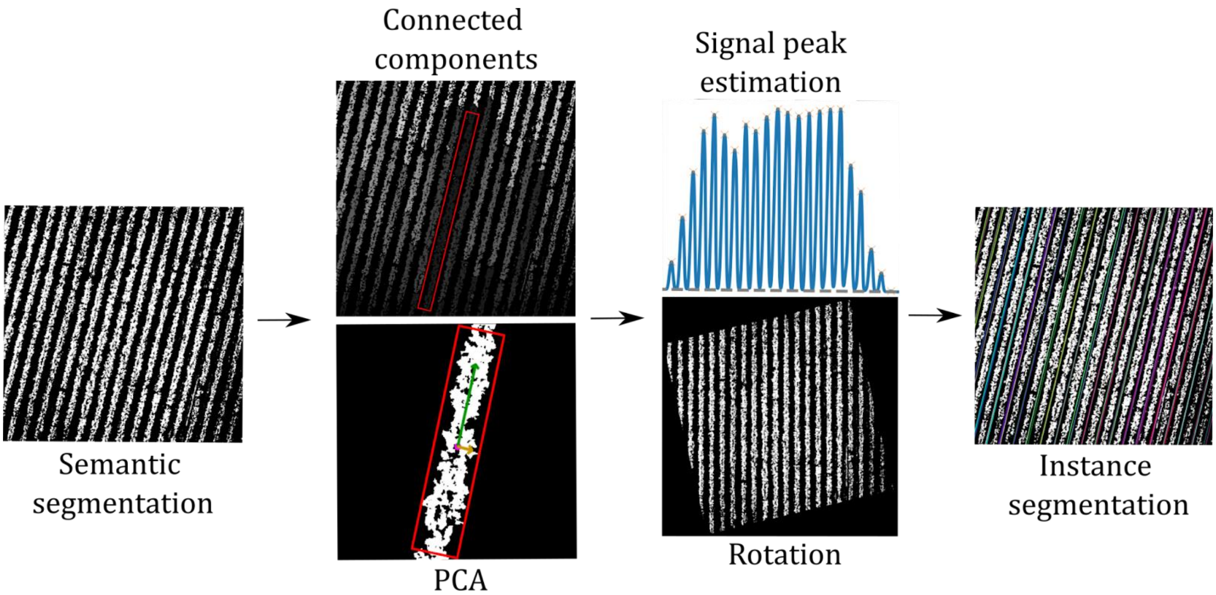


Figure 27: Workflow from semantic segmentation towards instance segmentation (identification of each row within the image).

This initial version of the workflow will be improved in the following period using learning procedures and deep neural network models.

5.6 Further Activities

UGV components development steps:

- Activity to make interface and calibration of the pH, EC and soil nutrient probes.
- Integration of the pH, EC and soil nutrient probes with the robotic platform.
- Design and implementation of guide rails with powertrain for lateral movement, main supporting frame, tilting platform, probing/sampling assembly.
- Design and implementation of spraying module.
- Design and implementation of a test environment that simulates operational environment of a blueberry field.
- Design and integration of disease detection module.
- Implementing ROS procedure for robotic system movement and navigation through the blueberry rows.
- Implementing ROS procedure for robotic system micro positioning with the respect of blueberry bush and blueberry bank.
- Implementing ROS procedure for automation of all robotic system functionalities.
- Development of a smartphone app and server for defining and monitoring robotic tasks and acquiring the measurement results.
- Integration of all modules into a complete system.
- Testing of a complete system in a test environment.
- In-field testing of the complete system.
- Redesign.

Image processing development steps:

- Collection of UAV images.
- Manual labelling of the training set.
- Exploring different CNN architectures.
- Model fine-tuning.
- Testing transferability.
- Comparing the output maps to the ground truth.
- Analysis of UGV-mounted camera and CNN model calibration.

The system will be tested in Serbia and Lithuania. The study will benefit from a variety of data to test the technology in different countries due to different weather conditions and prevailing soil and all other climatic conditions, as well as different cultivars and different growth stages.

6 Conclusions

The first 12 months of the project were used for defining the user requirements with agricultural partners and exploration of technical solutions that could be used for implementation. The pilot takes a holistic approach to the problem of monitoring blueberries and performing field operations. Based on the initial research, we defined the use cases and the roadmap for technological development. The appropriate equipment was procured, and the system prototypes will be tested in the next reporting period (M13-24). All activities have been conducted according to the scope and time frame defined in the project proposal and so far, the work proposed for WP6 has been conducted successfully according to the highest academic and technical standards.

References

[1] "P4 procedure," DJI, [Online]. Available: https://dl.djicdn.com/downloads/p4-multispectral/20200717/P4_Multispectral_Image_Processing_Guide_EN.pdf. [Accessed 21 November 2021].

[2] "rk500-23 probe," Rika Sensor, [Online]. Available: https://www.rikasensor.com/rk500-23-soil-ec-sensor-soil-salinity-sensor.html. [Accessed 21 November 2021].

[3] "LanceFET probe," Wellinq, [Online]. Available: https://www.wellinq.com/shop/product/lancefet-without-handle/. [Accessed 21 November 2021].

[4] "Nutrient Analyzer Kit," CleanGrow, [Online]. Available: https://www.ionselectiveelectrode.com/products/cleangrow-multi-ion-nutrient-analyzer-kit. [Accessed 21 November 2021].

[5] "Pix4D," Pix4D, [Online]. Available: https://support.pix4d.com/hc/en-us/articles/202560249-TOOLS-GSD-. [Accessed 29 Nov 2021].