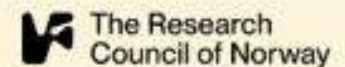


From Phenotyping Experiment to FAIR Data

11-13 March 2026 | Ås
NMBU, Norway



NBPPN



11th of March [Day 1]

Location: Innsikten, VET building

Morning (Hybrid, 9:00 – 12:00 CET)

Lectures:

8:30 – 9:00: *Welcome and coffee*

09:00-09:20: Introducing Nordic phenotyping infrastructures and networks

Morten Lillemo (PheNo, NMBU), Kristiina Himanen (NaPPI, UH), Svend Christensen (NBPPN, PhenoLab, UCPH), Xiang Su (NordPheno)

09:20-09:30: The Importance of Data Management in Plant Phenotyping – *Morten Lillemo (NMBU, NO)*

09:50-10:25: Data management benefit, common practice and potential tools – *Jesper Cairo Westergaard (University of Copenhagen, DK)*

10:25–10:40: *Coffee break*

10:40-11:25: The Road to Precision: Spectral Imaging for Abiotic Stress Responses—Are We There Yet? – *Sónia Negrão (University College Dublin, IE)*

11:30-12:00: Principles of handling UAV data – *Sahameh Shafiee (NMBU, NO)*

12th of March [Day 2]

Location: Innsikten, VET building

Morning (Hybrid, 8:45 – 12:00 CET)

Lectures:

8:30 – 8:45: *Welcome and coffee*

08:45-09:15: Making Data FAIR compliant – *Donald Hobern* (APPN, AU)

09:30-10:05: Image- based Phenotype of Fusarium head blight in barley and wheat –
Brian Steffenson (University of Minnesota, USA)

10:10–10:20: *Coffee break*

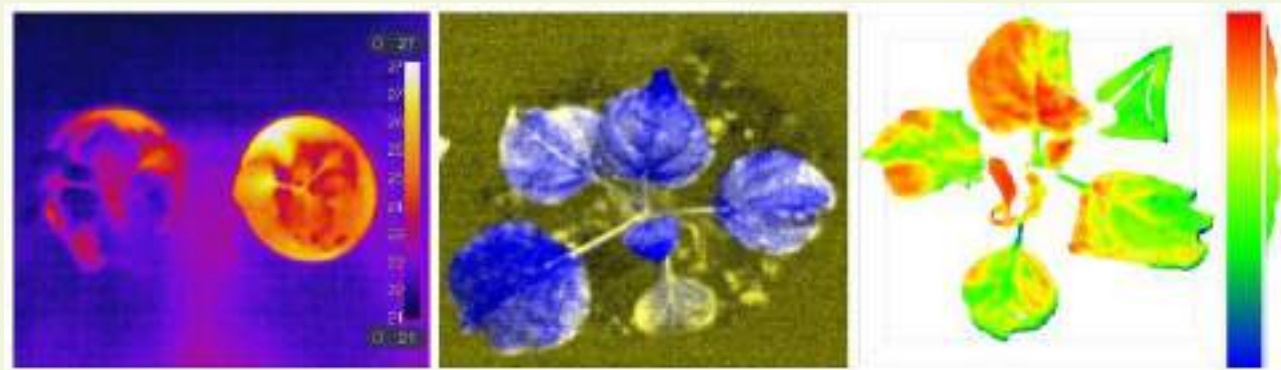
10:20-10:40: TraitFinder: 3D Multispectral Phenotyping for Plant Research – *Rita Armonienè* (LAMMC, LT)

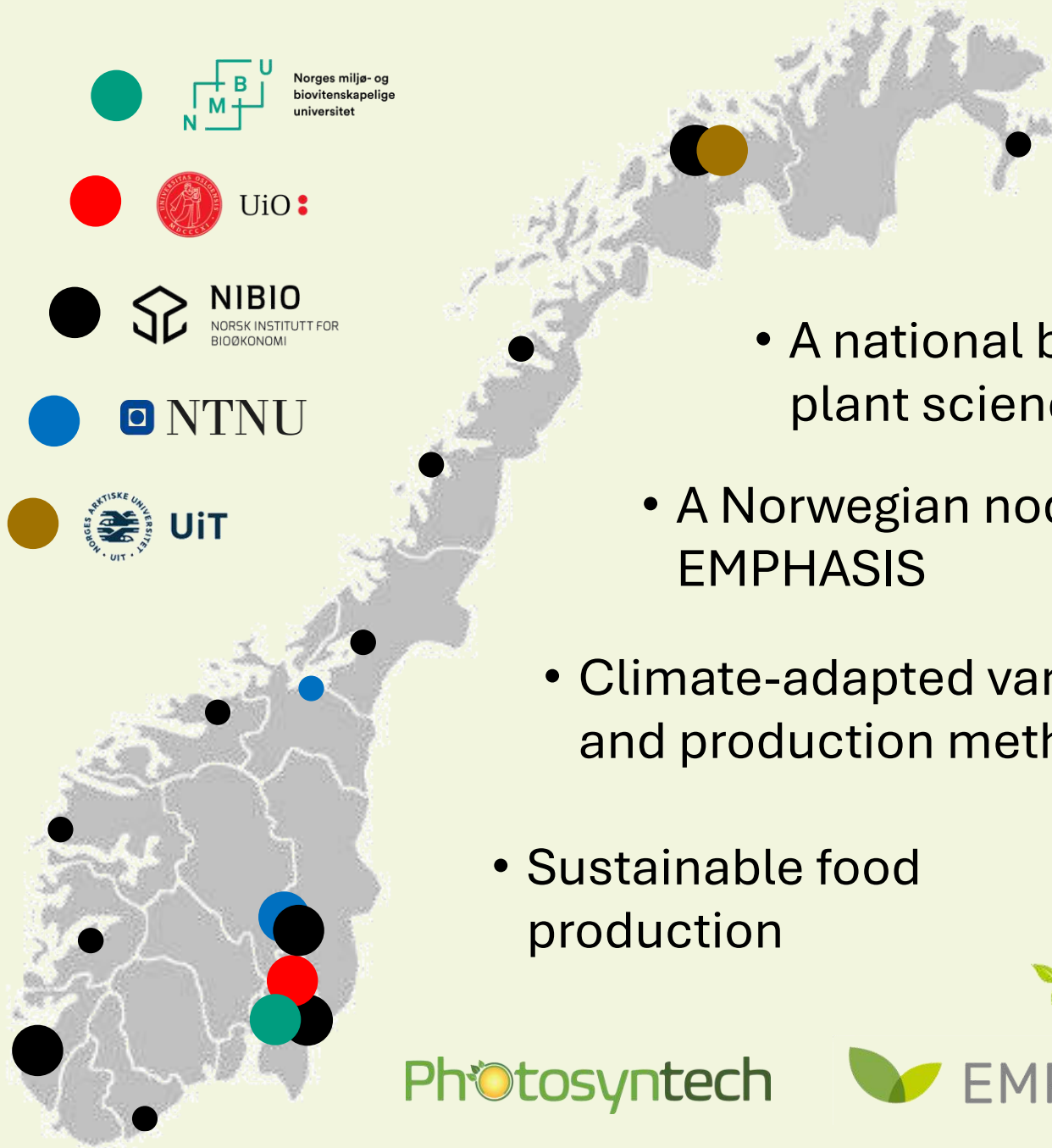
10:40-11:10: Focus on PHIS main principles, architecture – *Farzaneh Kazemipour*
(INRAE, FR)

11:10-11:30: *Closing online session - Kahoot*

PheNo - The Norwegian Plant Phenotyping Infrastructure

Morten Lillemo





Norges miljø- og
 biovitenskapelige
 universitet

UiO

NIBIO
 NORSK INSTITUTT FOR
 BIOØKONOMI

NTNU

UiT

• A national boost for
 plant science

• A Norwegian node in
 EMPHASIS

• Climate-adapted varieties
 and production methods

• Sustainable food
 production

Photosyntech

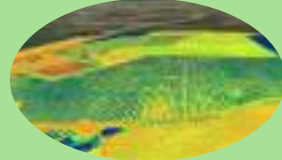
IPPN

EMPHASIS

Research facilities and
 services in:



Controlled
 environment
 phenotyping



Field phenotyping



Seed phenotyping



Data analysis



Funded by
 The Research
 Council of Norway



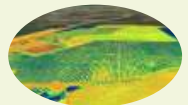
PheNo objectives

To provide the plant science community in Norway with **state-of-the art research facilities** and equipment for **precise fenotyping** under **controlled environment** and **in the field** as well as **seed phenotyping**

PheNo is a **distributed infrastructure** with installations across the country:



- Controlled environment phenotyping: NMBU, UiO, UiT, NIBIO



- Field phenotyping: NMBU, NIBIO



- Seed phenotyping: NMBU, UiO



- Data analysis and data management services: NTNU, UiT



National and international networks

- A Norwegian node in EMPHASIS



- International Plant Phenotyping Network



- Close connections to industry



- National research school in plant sciences



Visit our website at <https://pheno.no/>



- Overview of services
- Contact info
- Project leader:
Morten Lillemo
morten.lillemo@nmbu.no
- Coordinator:
Sara Laranjeira
sara.laranjeira@nmbu.no



Finnish National Plant Phenotyping Infrastructure NaPPI

Research coordinator

Kristiina Himanen

University of Helsinki, Finland

11.3.2026



High throughput Phenotyping Infrastructure at the University of Helsinki, Viikki campus, Finland



Small plant unit

Large plant unit

Capacity

1080 plants

270 plants

Cameras

PAM Chl Fluorescence
Thermal imaging
RGB/visible light

PAM Chl Fluorescence
RGB/visible light

Plant types

Plants up to 40 cm,
Arabidopsis, herbs, in vitro
plates, red berries, GMO

Crops, legumes, cereals, rosette
plants, trees (spruce, birch), up to
120 cm, GMO

Benefits of High Throughput Plant Phenotyping



- High throughput system for (relatively) high sample numbers
 - Hundreds of samples
- Automation of growth, morphology and physiology measurements
 - Multisensor imaging
- Time series resolution
 - from seedling to seed set
- Critical mass of data
 - Effective data management



Barley germplasm screen in drought (Prof. A. Chawade, SLU)

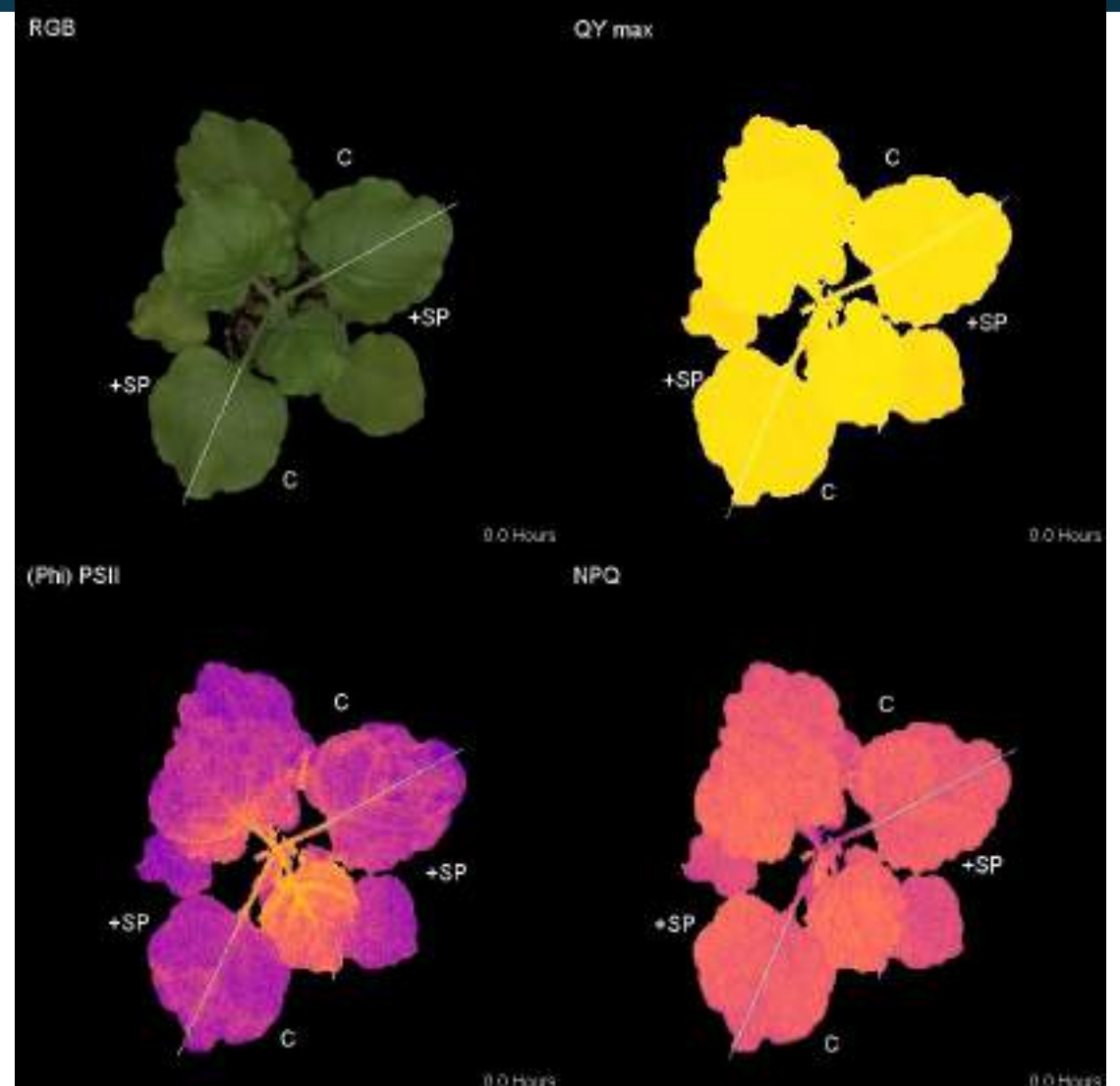
Research examples at NaPPI



Multisensor imaging for high throughput analysis of disease progression

Chlorophyll fluorescence imaging for monitoring effects of *Heterobasidion parviporum* small secreted protein induced cell death and in planta defense gene expression

Wen Z., Raffaello T., Zeng Z., Pavicic M. & Asiegbu F. (2019). *Fungal Genetics and Biology*. 126, p. 37-49

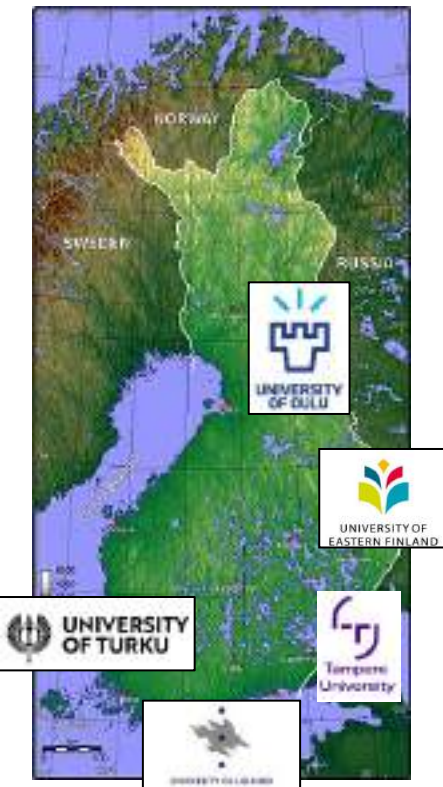


National Biocenter Finland network

**Biocenter
Finland**

Network of universities
with Biocenters

- Yearly evaluations
- Users, **user groups**, national, international
- **Outputs**, publications, **education**, patents etc
- Coordinates Research Council of Finland **funding** and **FIRI roadmap** applications



“ BF provides state-of-the-art technology services that cover key technologies in life sciences and biomedicine. The infrastructure services are organized into 17 Technology Platforms, which provide services annually for ~3000 research groups in academy, healthcare, and industry. All services are provided with an open-access policy.”

Helsinki Institute of Life Science



University of Helsinki research infrastructure network

Yearly **evaluations**

Develops **evaluation criteria:**

utilization rate, user base (excellency), cost efficiency

Funding for personnel and instruments


RI career paths with RI specific job descriptions

Has established **Life Science Data Competence Center**

Future Topics to promote in the long term:

- Multiomics approaches
- Platform collaboration
- Implementation of AI tools





Helsinki

Viikki

NaPPI

NaPPI partners and networks



National:



Maa- ja metsätalousministeriö



**Biocenter
Finland**

Nordic:



NBPPN

NOVA
UNIVERSITY NETWORK

European:



International:



IPPN

NaPPI team



PhD **Sylvain Poque**, NaPPI, sylvain.poque@helsinki.fi
Project planner at the National Plant Phenotyping Infrastructure
Department of Agricultural Sciences, University of Helsinki, Finland
Till June 2026



PhD **Kristiina Himanen**, Docent in developmental plant biology
Research coordinator at the National Plant Phenotyping Infrastructure,
Organismal and Evolutionary Biology Research Program and Dept. of
Agricultural Sciences, University of Helsinki, Finland
kristiina.himanen@helsinki.fi



Nordic Baltic Plant



Phenotyping Network

- The Nordic Baltic Plant Phenotyping Network (NBPPN) is a collaborative research network connecting plant phenotyping infrastructures and experts across the Nordic and Baltic countries.
- NBPPN was developed in response to the outcomes and momentum generated by two earlier Nordic initiatives that concluded in 2023.
- Organization:
 - an annual symposium (currently hosted in Båstad, Sweden)
 - one thematic workshop per year
 - a regular newsletter
 - opportunities to network, initiate collaborations and share expertise

Members NBPPN 2026

The current Nordic Baltic Plant Phenotyping Network partners consist of:

- Graminor AS
- Lithuanian Research Centre for Agriculture and Forestry (LAMMC)
- Findus
- Natural Resources Institute Finland (LUKE)
- SEJET
- University of Helsinki
- Centre of Estonian Rural Research and Knowledge (METK)
- Norwegian University of Life Sciences (NMBU)
- The Arctic University of Norway (UiT)
- Danespo
- Swedish University of Agricultural Sciences (SLU)
- Lantmännen
- University of Copenhagen
- TystofteFonden
- Lund University
- DLF
- Technological Institute Denmark



Organization 2025



Prof. Svend Christensen, University of Copenhagen
NBPPN chairperson



Jon Arne Dieseth, Graminor
NBPPN Vice- Chairperson



Johan Lundmark, Lantmännen
NBPPN Vice-chairperson



Katja Annette Willrodt, University of Copenhagen
NBPPN Project Mgr.

NBPPN Steering Committee 2025:

The members of the steering group are:

1. Svend Christensen (KU) Chair
2. Jon Arne Dieseth (Graminor) Vice-Chair
3. Johan Lundmark (Lantmannen) Vice-Chair
4. Kristiina Himanen (University of Helsinki)
5. Mathias Timmermann Christiansen (Danespo)
6. Gintaras Brazauskas (LAMMC)
7. Morten Lillemo (NMBU)

Upcoming events



NPPN Field Day 2019.

2025

- [Register here](#) for the Annual Symposium in Båstad, Sweden: 20th-21st November 2025 at [Hotel Skansen](#)



Nordic RI Hub for Digital Plant Phenotyping (NordPheno)

NordPheno positions the Nordic region as a leader in sustainable innovation,
creating a unified, data-centric RI for the future of agriculture.



The Need for Plant Phenotyping Digitalization



Unique Climate: Nordic crops grow at the edge of cultivation zones with increasing weather fluctuations.



Green Deal Goals: 20% reduction in fertilizers and 50% in pesticides by 2030.



HTP Integration: Coupling High Throughput Phenotyping with predictive modeling for genomic selection.



Interdisciplinary Gap: Bridging the skill void between plant scientists and computer scientists.



Overcome RI Fragmentation

Establishing the **Nordic Plant Data Hub**—a federated ecosystem linking servers and cloud services for trustworthy data sharing.



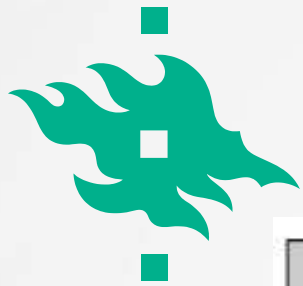
IoT & Distributed AI

Revolutionizing phenotyping via edge computing, autonomous sensors, and **Digital Twins** for simulation and real-time insights.

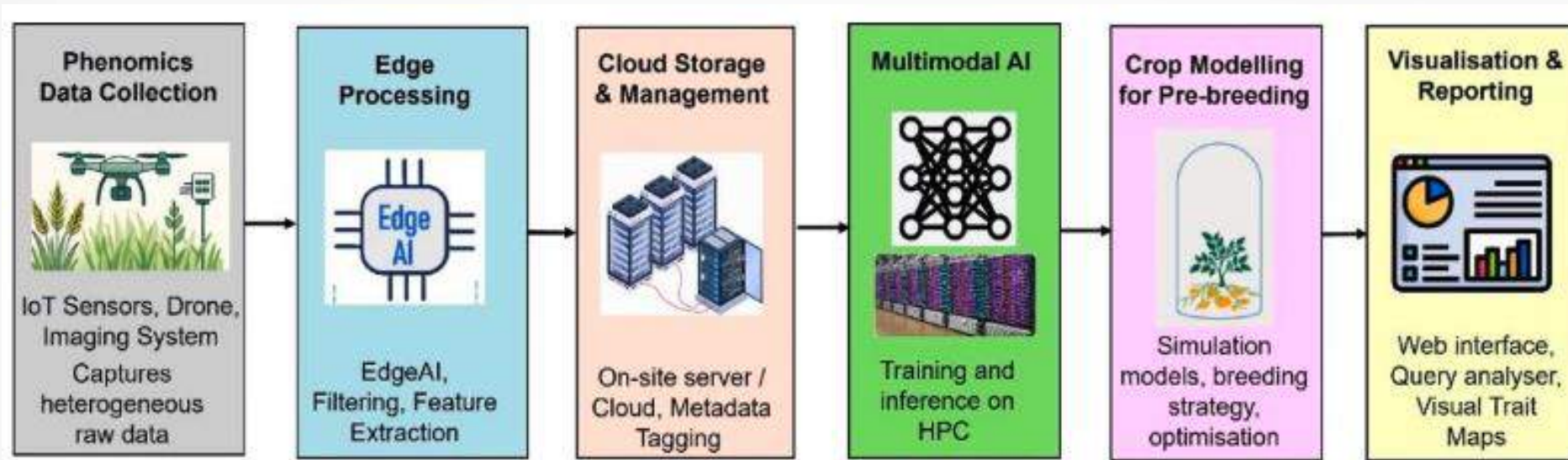


Training & Networking

Establishing intensive educational activities to bridge the knowledge gap between biology and computer science methodologies.



Digital Pipeline for Plant Phenotyping



Infrastructure Gaps

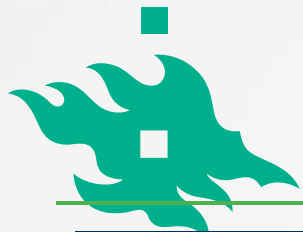
Current RIs rely on siloed data systems with limited interoperability. Manual Excel-based workflows for time-series data delay insights by weeks.

Target: Automated segmentation and integrated Python pipelines.

Knowledge Gaps

A disconnect exists in implementing AI-based phenotyping due to limited interdisciplinary training in Computer Science areas like IoT and Edge AI.

Target: Educating a new generation of bilingual (Bio+CS) researchers.



Consortium Partners & Expertise

Partner Institution	Core Contribution / Role	Country
Univ. of Helsinki	AI, Edge computing, Indoor phenotyping (NaPPI)	Finland
NTNU / NMBU	IoT, AI for Agriculture, Plant breeding, Field trials	Norway
Univ. of Copenhagen	HTP Indoor & Field phenotyping (PhenoLab)	Denmark
NIBIO / UiT / UiO	Arctic crop adaptation, Precision agriculture, Genomics	Norway
Swedish Univ. of Agri. Sci.	Sustainable agriculture policy & Field trials	Sweden

Work Packages (WP)

 **WP1:** Network Coordination & Strategic Alignment.

 **WP2:** Research Mobilities & Joint Publications.

 **WP3:** Workshops and Intensive Training Courses.

 **WP4:** Dissemination, Communication & Outreach.

 **FG A:** FAIR Data Management & Repositories.

 **FG B:** IoT Sensors & Edge Computing Strategies.

 **FG C:** AI Analytics & Plant Digital Twins.

 **FG D:** Emerging Tech (Gene Editing, Bio-pesticides).

The importance of data management

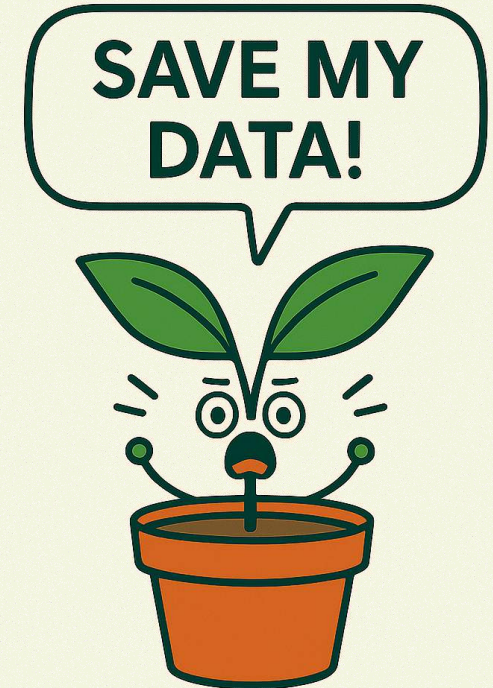
Morten Lillemo

Data Management Workshop at NMBU 11.03.2026



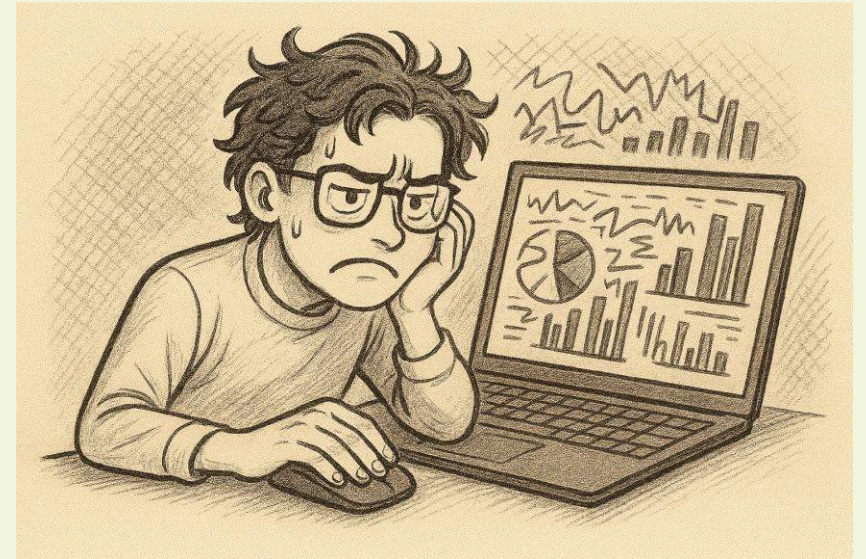
90% of science is lost!

- Of **100** datasets created...
 - **80** remain in the lab (unshared or hidden)
 - **20** are shared but rarely reused
 - **Fewer than two** meet FAIR principles
 - **Only one** typically drives new findings
-
- **The result:** delayed research findings, meta-studies short on evidence, and research that cannot be reproduced



Why should we care?

- Unpublished data from master students, PhD student, postdocs etc. are difficult to re-use if not properly organized and described
- In worst case, experiments have to be done again by your new student



Why should we care?

- Properly organized and described datasets is a pre-requisite for meta-studies



Tansley review

A meta-analysis of plant responses to light intensity for 70 traits ranging from molecules to whole plant performance

Author for correspondence:
Hendrik Poorter
Tel: +49 2461 67 8684
Email: h.poorter@fhnw.nl


Received: 4 October 2018
Accepted: 4 February 2019

Hendrik Poorter^{1,2}, **Ülo Niinemets**^{3,4}, **Nikolaos Nagler**¹,
Alrun Siebenkäs⁵, **Maarit Mäenpää**^{1,6}, **Shizue Matsubara**¹ and
Thijs L. Pons⁷

¹Max Planck Institute for Plant Physiology, Potsdam, Germany; ²Department of Biological Sciences, Massey University, North Ryde, NSW 2119, Australia; ³Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Kreutzwaldi 1, Tartu 91006, Estonia; ⁴Estonian Academy of Sciences, Tallinn 10130, Estonia; ⁵Department for Nature Conservation and Landscape Planning, Aachen University of Applied Sciences, Steinbaldstr. 29, 09104 Dresden, Germany; ⁶Department of Environmental and Biological Sciences, University of Jyväskylä, FI-40100 Jyväskylä, Finland; ⁷Plant Morphology, Institute of Environmental Biology, Utrecht University, 3512 PR Utrecht, the Netherlands

500 experiments
760 plant species

<https://doi.org/10.1111%2Fnp.15754>



EVANS REVIEW
<https://doi.org/10.1071/FP22294>

Functional Plant Biology

A meta-analysis of plant tissue O₂ dynamics

Max Herzog^{A,B}, Elisa Pellegrini^{A,B} and Ole Pedersen^{A,C}

ABSTRACT

Adequate tissue O₂ supply is crucial for plant function. We aimed to identify the environmental conditions and plant characteristics that affect plant tissue O₂ status. We extracted data and performed meta-analysis on >1500 published tissue O₂ measurements from 112 species. Tissue O₂ status ranged from anoxic conditions in roots to >53 kPa in submerged, photosynthesizing shoots. Using information-theoretic model selection, we identified 'submergence', 'light', 'tissue type' as well as 'light × submergence' interaction as significant drivers of tissue O₂ status. Median O₂ status were especially low (<50% of atmospheric equilibrium) in belowground rhizomes, potato (*Solanum tuberosum*) tubers and root nodules. Mean shoot and root O₂ were ~25% higher in light than in dark when shoots had atmospheric contact. However, light showed a significant interaction with submergence on plant O₂, with a submergence-induced 44% increase in light, compared with a 42% decline in dark relative to plants with atmospheric contact. During submergence, ambient water column O₂ and shoot tissue O₂ correlated stronger in darkness than in light conditions. Although use of miniaturised Clark-type O₂ electrodes has enhanced understanding of plant O₂ dynamics, application of non-invasive methods in plants is still lacking behind its widespread use in mammalian tissues.

For full list of author affiliations and disclosures see end of paper.

Correspondence to:
Max Herzog
The Freshwater Biological Laboratory,
Department of Biology, University of
Copenhagen, Universitetsparken 4,
3rd Floor, Copenhagen 2108, Denmark
Email: herzogm@bio.ku.dk

Handling Editor:
Faming Dong

129 studies
112 plant species

<https://doi.org/10.1071/FP22294>



Metadata – data about the data

- What's in this can?



Metadata – data about the data

- What do the numbers in this table mean?

1101	45	60	67	0.800	1101	45	60	67	0.800
1102	3	65	60	0.830	1102	3	65	60	0.830
1103	1313	70	70	0.868	1103	1313	70	70	0.868
1104	1617	70	68	0.864	1104	1617	70	68	0.864
1105	1528	55	70	0.888	1105	1528	55	70	0.888

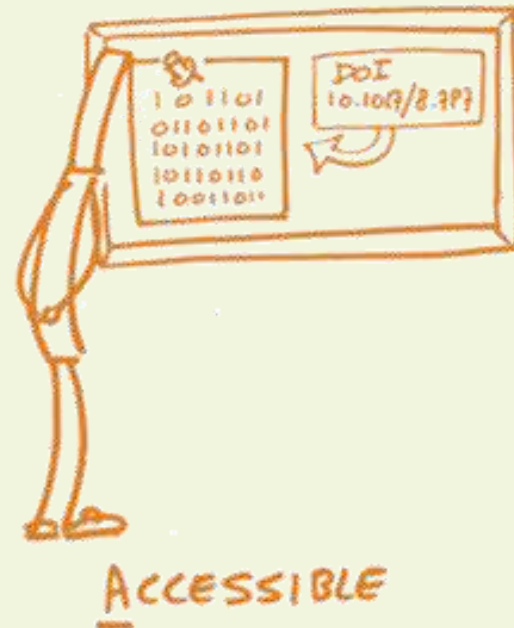
Metadata

Description of the growing conditions
How the traits were measured, which
sensors were used, etc.



The FAIR data principles

FAIR DATA PRINCIPLES



<https://book.fosteropenscience.eu/>



The four basics of FAIR data

Principle	What it means in practice
F indable	discoverable with metadata, identifiable and locatable by means of a standard identification mechanism
A ccessible	always available and obtainable; even if the data is restricted, the metadata is open
I nteroperable	both syntactically parseable and semantically understandable, allowing data exchange and reuse between researchers, institutions, organisations or countries
R eusable	sufficiently described and shared with the least restrictive licences, allowing the widest reuse possible and the least cumbersome integration with other data sources



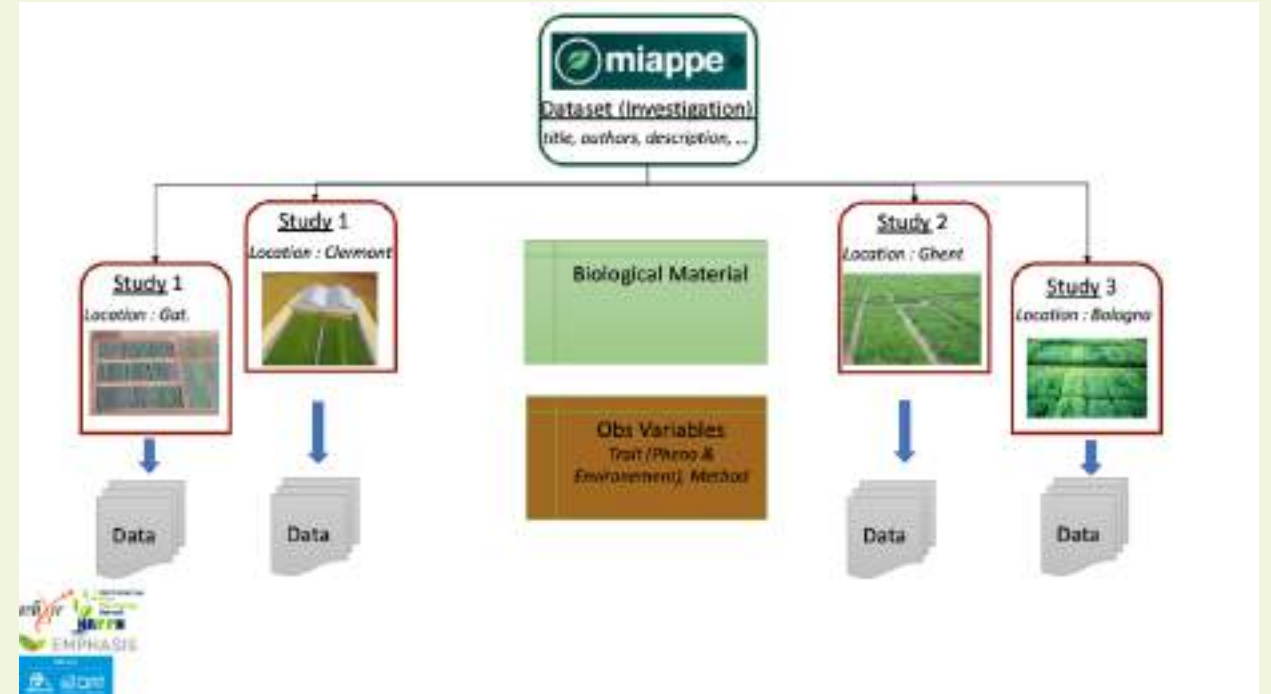
Things to remember

- FAIR is a set of principles; not a standard
- Does following the FAIR principles mean that your data has to be shared openly with everyone? **NO**
- Data can be FAIR but not open. For example, data could meet the FAIR principles, but be private or only shared under certain restrictions
- Open data may not be FAIR. For example, publicly available data may lack sufficient documentation to meet the FAIR principles, such as licensing for clear reuse



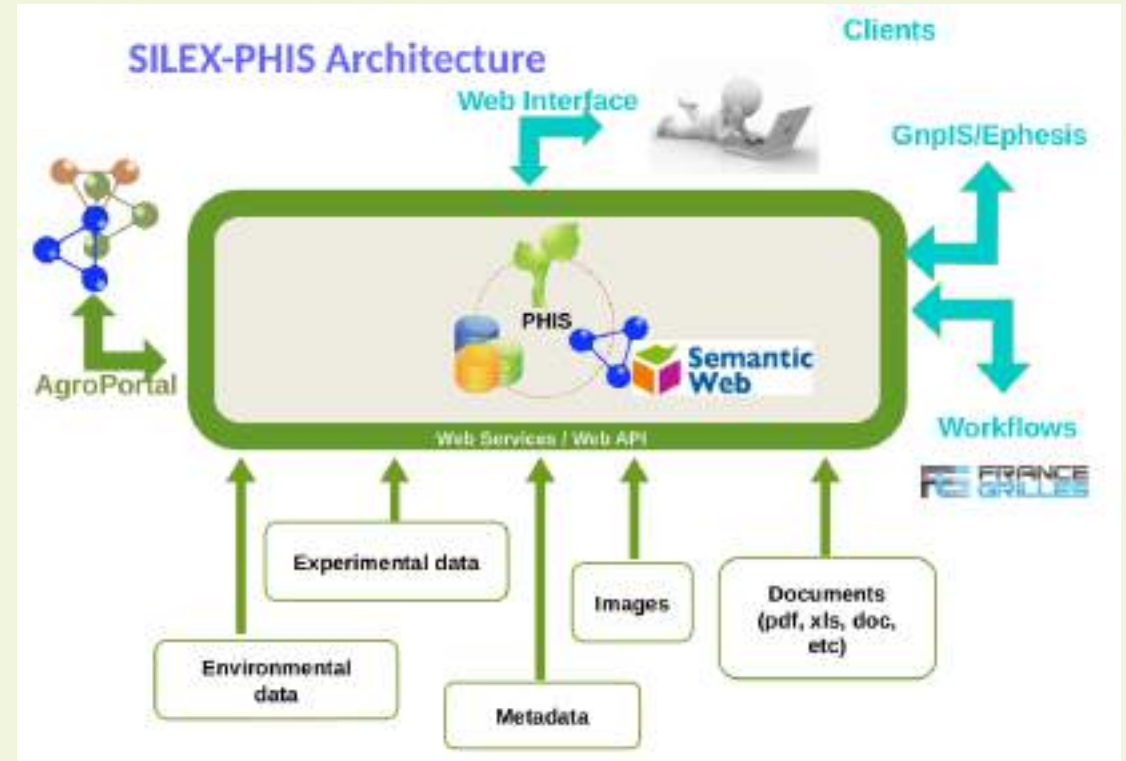
MIAPPE

- The Minimum Information About Plant Phenotyping Experiments (MIAPPE)
- A data standard designed to harmonize the description of experimental and computed data to enable its sharing, publication and reuse.



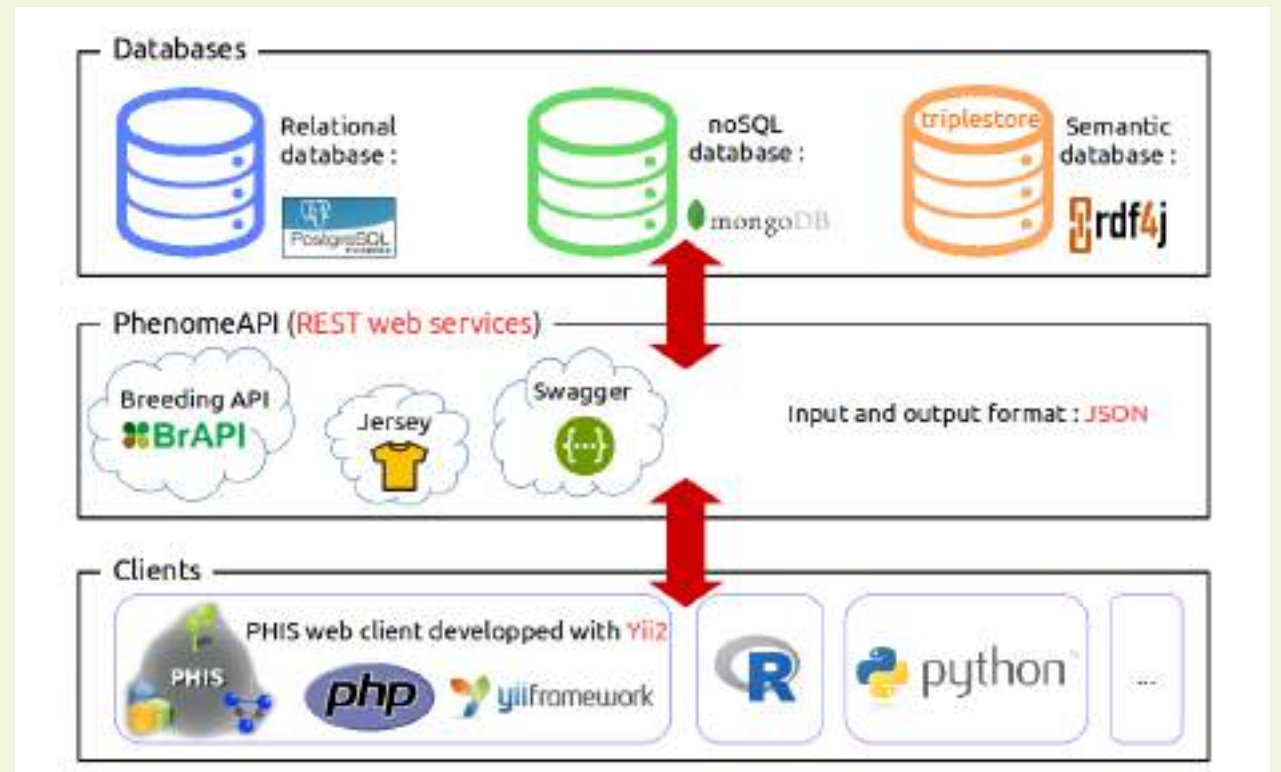
PHIS

- Phenotyping Hybrid Information System (PHIS)
- A database system for plant phenotyping data



BrAPI

- The Breeding API (BrAPI)
- A standardized web service API specification for exchange of scientific data
- Enables interoperability among databases and tools



Enjoy the workshop!

From Phenotyping Experiment to FAIR Data

11-13 March 2026 | Ås
NMBU, Norway



The illustration shows three potted plants in the foreground. Above them is a sensor on a stand. To the right, a computer monitor displays a bar chart and a plant icon. The background is filled with various icons representing data, networks, and scientific processes.

PheNo
NordPlant UPSCALE
SLU
Swedish University of Agricultural Science
NaPPI
European National Plant Phenotyping Infrastructure
NKJ
UNIVERSITY OF HELSINKI

UNIVERSITY OF COPENHAGEN
NBPPN
PHIS
EMPHASIS
INRAE
The Research Council of Norway
NordForsk



Data management

- Benefits
- Common practices
- Potential tools

Jesper Cairo Westergaard,
UCPH, PLEN, Crop Sciences, Biostatistics Group

Presented at the NMBU DM workshop, 2026-03-11

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE



This presentation

Title given: **Data management benefit, common practice and potential tools**

So, I have split this up into exactly those three

- Benefits
- Common practices
- Potential tools

There will be some slides reused from earlier work by Isabelle Alic, Sylvain Poque and me.

And then at the end...

- Questions? Comments?
- and sharing of your experiences with DM - good or bad

PheNo: Data management workshop - From Phenotyping Experiment to FAIR Data

We invite researchers, students, and professionals interested in plant phenotyping and data management to participate in a hands-on workshop focused on metadata analysis and data reusability. The activity will emphasize best practices for organizing, storing, and sharing data to maximize its long-term value.

About me

Jesper Cairo Westergaard

Master of Science (MSc) in Agriculture

Part of **the Biostatistics Group** at the Section for Crop Sciences, Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen

My office hours, greenhouse and field work, mainly takes place at our Taastrup Campus, Højbakkegård Allé 13, DK-2630 Taastrup

Networks: NBPPN – the Nordic Baltic Plant Phenotyping Network

Project involvement currently:

IBIS: Initiative for Biofertilizer Innovation and Science

RIC: Robotic Intercropping – From monoculture toward diverse cropping systems

OCH: One Crop Health: Next generation crop protection

MATRIX: Microbiome Assisted Triticum Resilience In X-dimensions

RootAccess: a root phenotyping field facility



Some of our research platforms at Taastrup Campus

RadiMax

root research

RootAccess

root research

Weather Station in the field
+ environmental sensors in climate
chambers & greenhouses

Mobile climate stations

Robotti

autonomous robot

Drones

PhenoLab

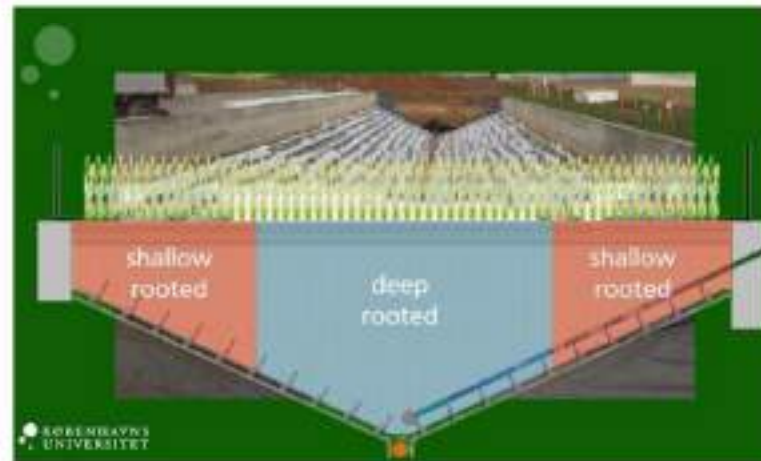
automated greenhouse

Person observing in field or greenhouse

Some of our research platforms at Taastrup Campus

RadiMax: Root research

- Above and below ground
- Cameras inserted to tubes
 - one per two rows of plants
 - angling down
- Isotope insertion in watering
- Imaging from drones
- Rainout shelters (manually)



Some of our research platforms at Taastrup Campus

RootAccess: Root research

- Tunnel, 3m deep
- In total 1296 inspection windows
- Split into 24 panels
 - 12 north, 12 south
- access to the full root zone of the crop in real field conditions.
- Camera insertions
- Permanent sensors for soil sensing
- Sampling access for soil and roots



Some of our research platforms at Taastrup Campus

Weather Station in the field

+ environmental sensors climate chambers & greenhouses

- Soil
 - Temperature
 - Humidity
- Precipitation
- Wind
- Radiation



Some of our research platforms at Taastrup Campus

Mobile climate stations

- We can move them from one field to another
- We can also access data near a farm in a project
- Data on cloud > manual download to local PC
- Data also in a phone app



Some of our research platforms at Taastrup Campus

Robotti

- Autonomous operation
 - Imaging
 - Sowing
 - Tilling
-
- We added extra high-quality camera with a very powerful flash, to get homogenous lighting for modelling images



Some of our research platforms at Taastrup Campus

Drones



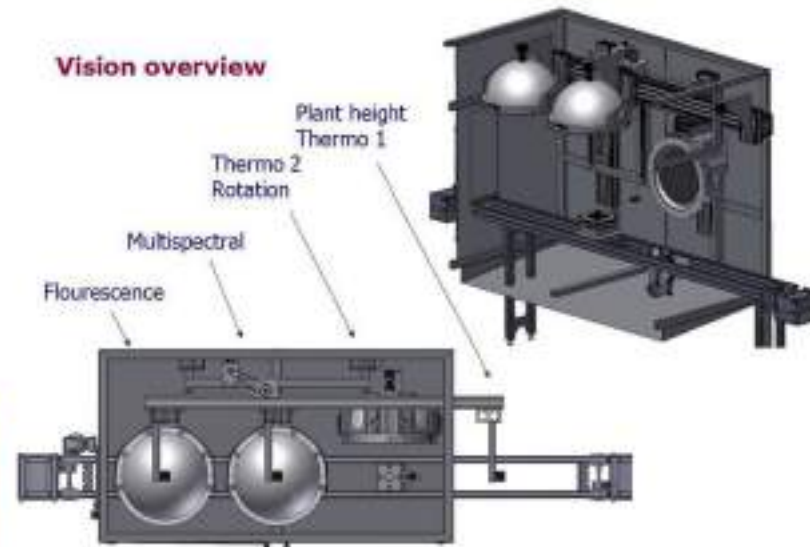
- Autonomous operation
- Imaging
 - RGB
 - Multispectral
 - Thermal
 - LiDAR
- Spraying herbicides & sowing/dropping seeds
 - Still in process with permission (SORA, very time-intensive)



Some of our research platforms at Taastrup Campus

PhenoLab

- 117 fixtures/plants
- Automatic watering , Weighing, height, and imaging (multispectral and thermal)
- 4 watering sources (different nutrition)



Some of our research platforms at Taastrup Campus

Person observing in field or greenhouse

Mergin Maps

- App for phone/tablet
- Imaging, text, lists, checkboxes, dropdowns, maps, track a path, save a location, arrange views in tabs for better user experience
- Interface designed in QGIS > pushed to cloud > pulled from phone
- Phone synchronizes to cloud
- Data the pulled to QGIS from cloud for further handling, visualization, analysis, and for data management



Both support QR codes



E.g. for pots, plots, fields, sensors, machines...

Ensures data validity: You scan the item before you start your annotating, sampling, imaging/sensing

Licor LI600

- Plant physiology
- Single leaf
- Stomatal conductance
- Chlorophyll fluorescence



Some of our research platforms at Taastrup Campus

Overview of systems ... currently

	Collection		Storage		Integration	Comments
	Automated, incl. dataloggers	Manually	Department only		Also, with partners, sometimes public	
			Before release (non-curated)	Availability		
RadiMax root research	Soil and air sensing	Imaging, isotope insertion,	networked drive	Once added to resource		
RootAccess root research		Imaging, isotope insertion, soil samples,		Once added to resource		
Weather Station in the field + environmental sensors in climate chambers & greenhouses	Timed daily pull from dataloggers to PC and onwards to networked drive			Raw: Upon request	Curated: Yearly (Jan/Feb), web with DOI, frozen archive.	
Mobile climate stations	Data live to cloud (proprietary)	We pull from cloud to us, with temporal and geographic range	networked drive			Cordulus
Robotti autonomous robot	Imaging, sowing, tilling					Robotti
Drones	Imaging	Mission planning, data transfers	Networked drives			Dronevolt, Pix4D, Agisoft, QGIS w. PlotCut
PhenoLab automated greenhouse	Imaging, watering, weighing (twice daily).					Proinvent, Videometer
Person observing in field or greenhouse		Phone/tablet : Text images, position/track				Mergin Maps
		Leaf physiology, incl. position				Licor LI600

But, enough about us...

What is Data Management in research

“Research data management refers to the handling of research data (collection, organisation, storage, and documentation) during and after a research activity. Good data management helps ensure that researchers share their data in a FAIR way (findable, accessible, interoperable, and re-useable). Research organisations increasingly require their researchers to develop a data management plan to ensure that all aspects are considered from the start of an activity on.”

- Collection
- Organisation
- Storage
- Documentation



What is Data Management in research

Data Management involves the practice of planning, collecting, keeping, and using data securely, efficiently, and cost-effectively.

Key Components:

- Data planning
- Data Collection
- Data Storage
- Data Security
- Data Governance
- Data Integration
- Data Analysis



<https://rdmkit.elixir-europe.org/>

Benefits of Data Management

Mismanagement of data

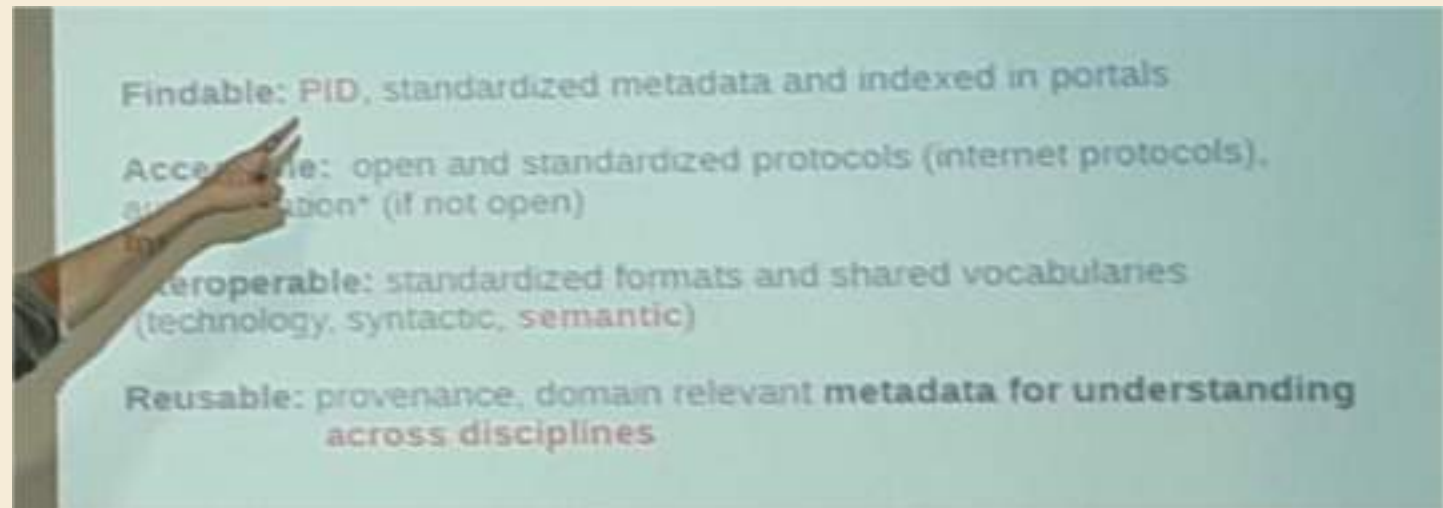
- Late arrival of data to “next in line”
- Incorrect metadata or parts of the data
 - Example: Bungled up CRS (coordinate reference system) or other location data, e.g. north/south reversed
- Insufficient metadata
 - Missing time-of-day. May not have seem relevant but it ended up being.
- **Does anyone else have some examples?**

Benefits of Data Management

- **Living up to standards or at least realistic expectations**
 - Helps you and your colleagues
 - Helps others
 - Is probably a demand put upon you already
- **FAIR data**
 - Findable
 - Accessible
 - Interoperable
 - Reusable



Pascal Neveu @InraMontpellier explaining FAIR implementation at an international plan in @OpenSilex and @PHISphenomics



Why should you care?

... about sharing and especially keeping data FAIR



Questions? Comments?

- _____
- _____
- _____
- _____

Common practices of Data Management

- When applying for a project
 - Demands from funding bodies:
 - Example for a project, in a moment
- With the project starting
- Ongoing
- Wrapping it up
- Afterwards... The lifespan of your data (?)

Common practices of Data Management

When applying for a project:
An example from the project **Robotic Intercropping**

Researchers will teach robots to grow wheat and clover grass on the same field

NEWS

ROBOTICS Autonomous robots could help farmers reduce the use of pesticides and grow food more sustainably. A new project led by University of Copenhagen aims to pave the way for robotics in agriculture by exploring how multiple crops can be cultivated efficiently on the same plot of land. The Novo Nordisk Foundation is supporting the project with a grant of 75 million DKK (10 million EUR).

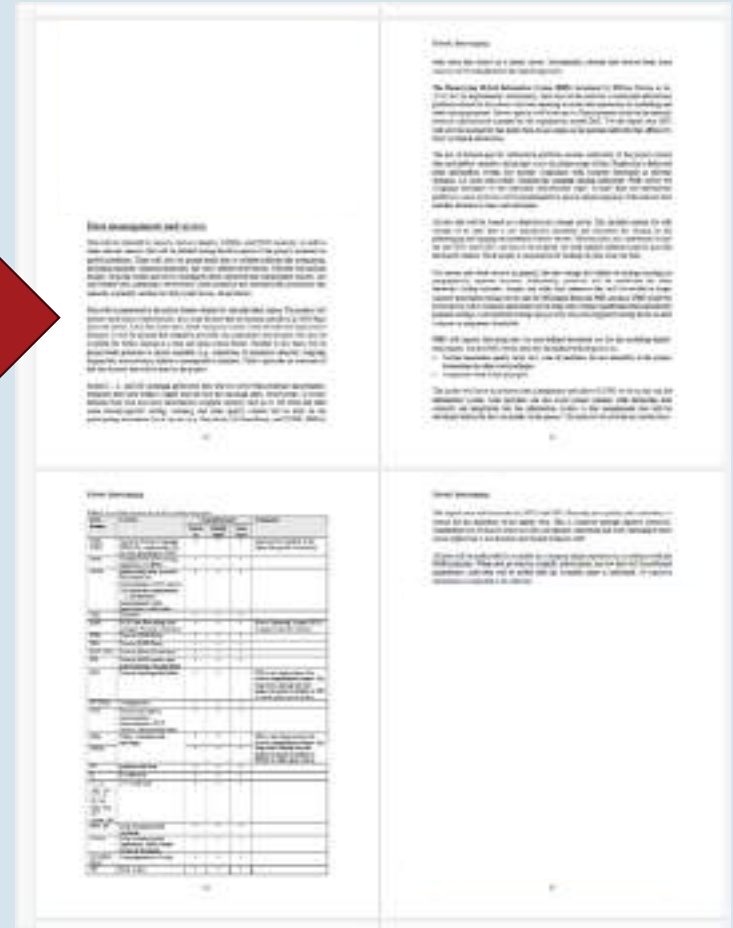


Illustration: Jeppe Skovgaard/DTU

Imagine a field without a tractor. Instead, the heavy machinery is replaced by smaller, self-driving robots.

That's the vision behind the new Robotic Intercropping project, led by Professor Svend Christensen from the Department of Plant and Environmental Sciences, the ambitious project is a collaboration between the University of Copenhagen, the Technical University of Denmark (DTU), the University of Southern Denmark, and Wageningen University in the Netherlands.

DM as a part of the application



Read more here: <https://science.ku.dk/english/press/news/2025/researchers-will-teach-robots-to-grow-wheat-and-clover-grass-on-the-same-field/>

Common practices of Data Management

When applying for a project:
An example from the project **Robotic Intercropping**

Data management and access

Data will be collected by sensors, such as cameras, LIDARs, and GNSS receivers, as well as other relevant sensors (that will be defined during the first period of the project) mounted on mobile platforms. There will also be ground-truth data to validate/calibrate the automation, including manually sampled phenotypic and stress-related observations. The data will include images, 3D point clouds and GNSS coordinates from calibrated and synchronized sensors, soil and weather data, phenotypic observations (both manually and automatically, preciously and remotely acquired), and harvest data (yield, losses, observations).

Data will be generated in the native format adopted by the individual sensor. The project will utilize a wide array of data formats, incl. some formats that are domain-specific (e.g. ROS Bags from the robots, LAZ files from the LIDAR and point clouds; both are free and open-source formats). It will be ensured that whenever possible, any proprietary data formats will also be available for future sharing in a free and open-source format. Parallel to this data will be project-wide protocols to ensure metadata (e.g., annotation of anomalies detected, sampling responsible, access/events) adhere to interoperable standards. Table 4 provides an overview of the data formats that will be used in the project.

Spatial 1-, 2-, and 3D including geolocated data will be saved with positional uncertainties. Temporal data both within a single trial but also for monitoring units, observations, or events between trials will also have uncertainties assigned. Initially, data in its raw form and after some domain-specific sorting, cleaning, and other quality control will be held on the participating institutions' local servers (e.g. Nextcloud, MS SharePoint, and UCPH's ERDA)

with some data stored on a shared server. Subsequently, relevant data derived from these sources will be transferred to the central repository.

The Phenotyping Hybrid Information System (PHIS) developed by INRAe (Neyens, et al., 2019) will be implemented. Additionally, there may be the need for a centralized information platform tailored to the robotic activities ensuring accurate data annotation for modelling and other sharing purposes. Server capacity will be set up at a Danish partner based on the national research infrastructure managed by the organization named DeIC. For the digital twin SDU will provide hosting for the digital twin board, again on the national infrastructure offered by DeIC to Danish universities.

The use of domain-specific information platforms ensures uniformity of the project-generated data and enables seamless and pre-empt cross-discipline usage of data. Employing a dedicated plant information system also ensures compliance with accepted ontologies in relevant domains: e.g. plant physiology, morphology, imaging/sensing, pathology. PHIS allows for assigning ontologies to the individual measurement types. If more than one information platform is used, protocols will be implemented to ensure unique mapping of the relevant data matches between systems and ontologies.

All raw data will be stored on a shared-access storage server. This includes options for safe storage of all data that is not specifically annotated and formatted for sharing on the phenotyping and imaging and potential robotics servers. Mission plans, incl. operational scripts for the UGVs and UAVs, are also to be archived, for both internal reference and for possible third-party interest. Each partner is responsible for backing up their own raw data.

For servers and cloud services in general, the data storage will adhere to backups residing on geographically separate locations. Additionally, protocols will be established for other necessary backup routines. Images and other data intensive files will be resided in large-capacity prioritized storage servers and be URI-linked from the PHIS instance. PHIS itself will be hosted on a fast-response prioritized server with both systems benefiting from periodically planned backups, user-initiated backup and possibly also auto-triggered backup based on data volumes or magnitude thresholds.

PHIS will support delivering data via user-defined download sets for the modelling/digital-twin experts. On the PHIS server, data will be annotated with respect to its:

- Current annotation quality level incl. state of readiness for use internally in the project downstream in other work packages.
- Compliance with FAIR principles.

The project will have an in-house data management specialist at UCPH, to set up and run the information system. Said specialist can also assist project partners with formatting data correctly and integration into the information system. A data management plan will be developed within the first six months of the project. This plan will be revised on a yearly basis.

The digital twin will be hosted by DTU and SDU. Ensuring data quality and consistency is crucial for the operation of the digital twin. This is achieved through rigorous protocols, standardized use of servers across all sites (or rigorous calibration and cross-checking if exact sensor replication is not feasible) and trained technical staff.

All data will be made publicly available in a common online repository in accordance with the FAIR principles. When data are used in scientific publications, the raw data will be published immediately, non-data will be added after the scientific paper is published. No sensitive information is expected to be collected.

- With the project starting
- Ongoing
- Wrapping it up
- Afterwards... The lifespan of your data

Common practices of Data Management

When applying for a project:
An example from the project **Robotic Intercropping**

What data formats?

Here a list of the data formats that we expect to use in the project →

Table 3: list of data formats that will be used in the project

Data format	Used in	Considerations			Comments
		Universal	Standardized	Open source	
KML, KMZ	Keyhole Markup Language (KMZ for compressed). For mission planning for UAV's	+	+	+	international standard of the Open Geospatial Consortium
XML	Configuration and ontology definitions for PHIS	+	+	+	
JSON	phenotyping data; potential file formats for measurements of IoT-sensors (soil humidity, temperature, ...), navigational measurements, data annotations, config files	+	+	+	
SQL	Database	+	+	+	
BAG	ROS Data Recording (aka. rosbags). For data collection	+	+	+	Robot Operating System (ROS) is open-source for robotics.
PNG	Dataset (RGB Data)	+	+	+	

Table 3: list of data formats that will be used in the project

Data format	Used in	Considerations			Comments
		Universal	Standardized	Open source	
KML, KMZ	Keyhole Markup Language (KMZ for compressed). For mission planning for UAV's	+	+	+	international standard of the Open Geospatial Consortium
XML	Configuration and ontology definitions for PHIS	+	+	+	
JSON	phenotyping data; potential file formats for measurements of IoT-sensors (soil humidity, temperature, ...), navigational measurements, data annotations, config files	+	+	+	
SQL	Database	+	+	+	
BAG	ROS Data Recording (aka. rosbags). For data collection	+	+	+	Robot Operating System (ROS) is open-source for robotics.
PNG	Dataset (RGB Data)	+	+	+	
PCD, LAS	Dataset (Point Cloud data)	+	+	+	
TIF	Dataset (RGB, multi- and hyperspectral, height data)	+	+	+	
PSD	Dataset (multispectral data)	+	+	-	PSD is not open source, but used as intermediate format. For long-term sharing relevant media are made available as TIF or other open source format.
PYVALE	Configuration	+	+	+	
CSV	Dataset (navigation measurements, measurements of IoT-sensors, phenotyping data)	+	+	+	
MP4	Video: Container and encoding	+	+	-	MP4 is not open source, but used as intermediate format. For long-term sharing relevant media are made available as WebM or other open source.
WebM		+	+	+	
PY	python code base	+	+	+	
R	R code base	+	+	+	
C, cc, app, cxx, c++, k, H, Hk, app, fxx, k++	C++ code base	+	+	+	
cppt, .so		+	+	+	
PTH, PT	Deep learning model (pytorch)	+	+	+	
ONNX	Deep learning model (universal), Open Neural Network Exchange	+	+	+	
DOCKER, FILE	Containerization of code	+	+	+	
SH	Batch scripts	+	+	+	

Questions? Comments?

- _____
- _____
- _____
- _____

Potential tools of Data Management

- The two main types of Data Management tool
 - The DM plan “Just talk”
 - Necessary,
 - but doesn’t “get the job done
 - The action
 - Putting the data somewhere
 - Ensuring its quality
- And the hidden tool:
 - Persistence..
 - Revisiting your data. OFTEN
 - Having others involved
 - Consider: How many people can deliver data to those who need?

Potential tools of Data Management

- Feeling the need for “something better”
- But better is not always “perfect”,
 - it can also be something that “gets the job done”

Potential tools of Data Management

- Professional shared systems like PHIS – actually has data stored in it (or linked to other locations where the data is, eg large images)
- The data about the data
 - Example: MIAPPE
- Homemade systems just used in a research group or a university etc.

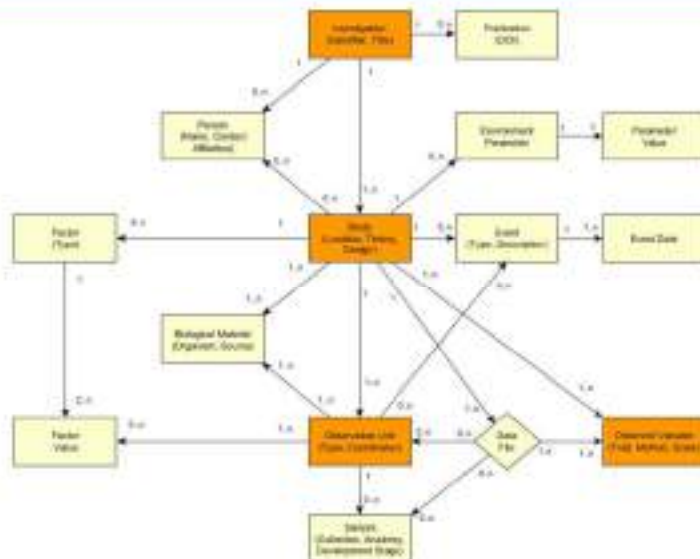
Potential tools of Data Management: MIAPPE

Structuring your experiment and associated metadata



Minimum Information About Plant Phenotyping Experiments (MIAPPE, <https://www.miappe.org/>)

- An open, community driven project to harmonize data from plant phenotyping experiments.
- Comprises both:
 - Conceptual data representation
 - Recommended checklist of metadata required to adequately describe a plant phenotyping experiment.

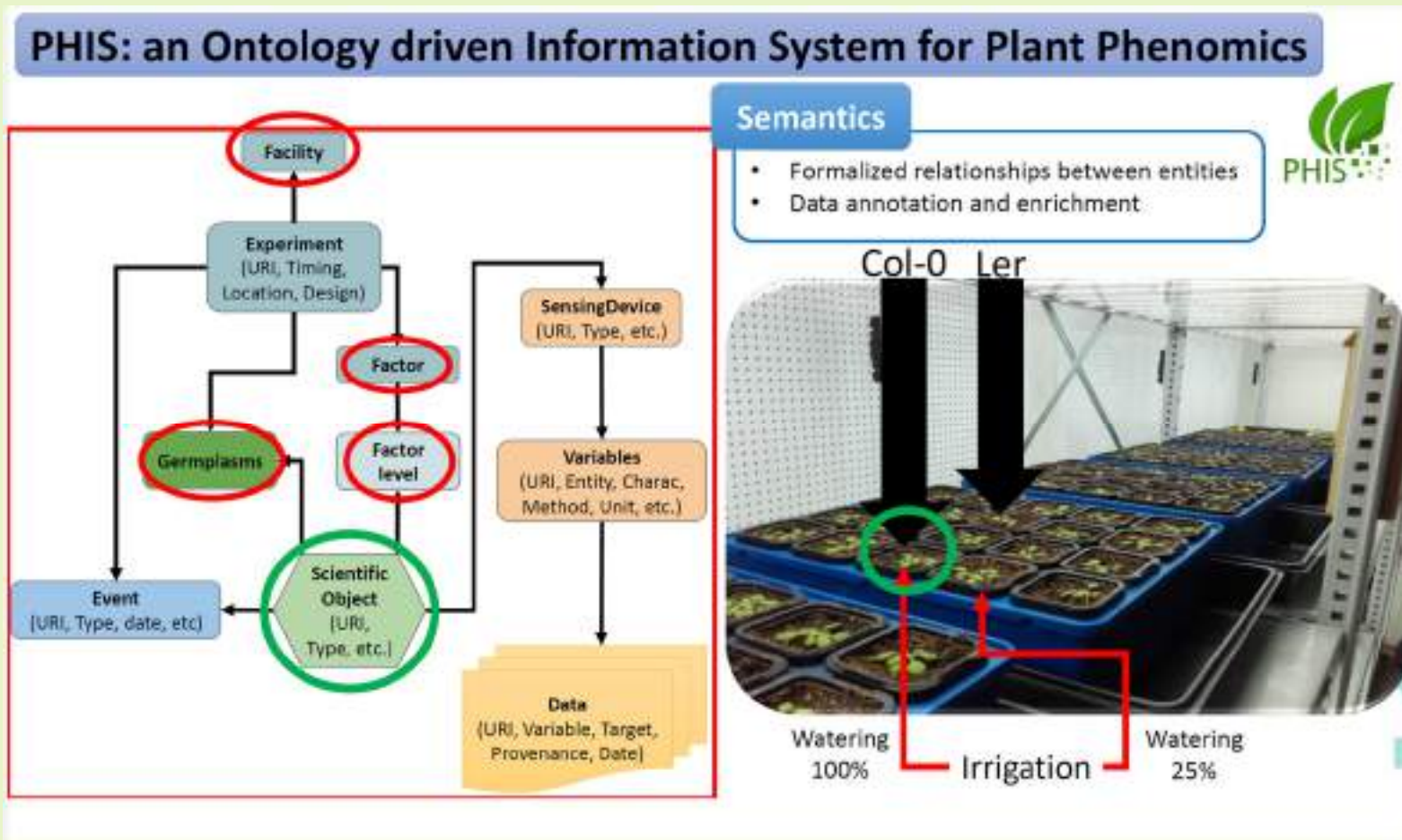


Each Elements have:

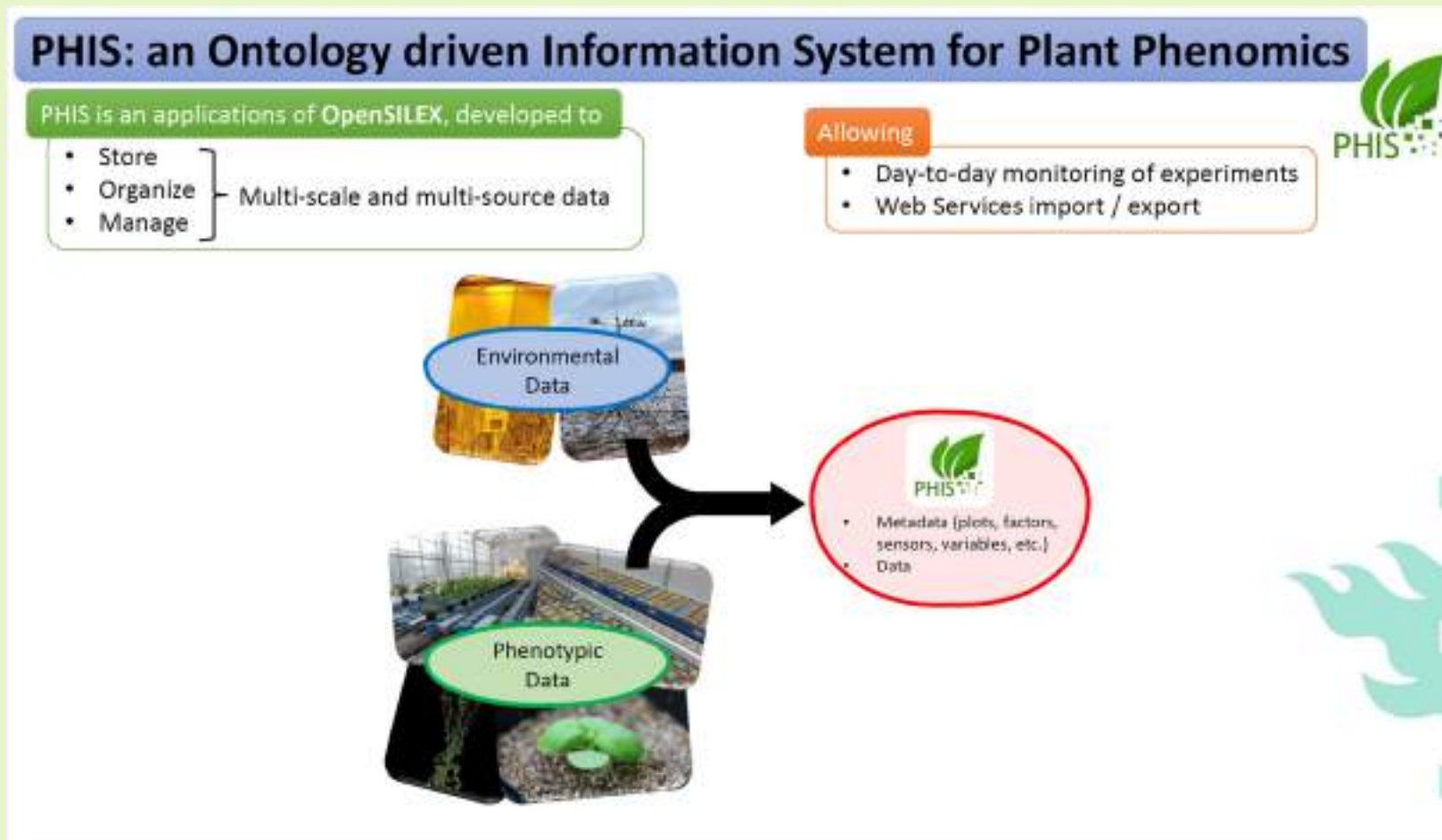
- Precise definition
- Precise description

To help you structure your experiment and ensures its comprehensiveness

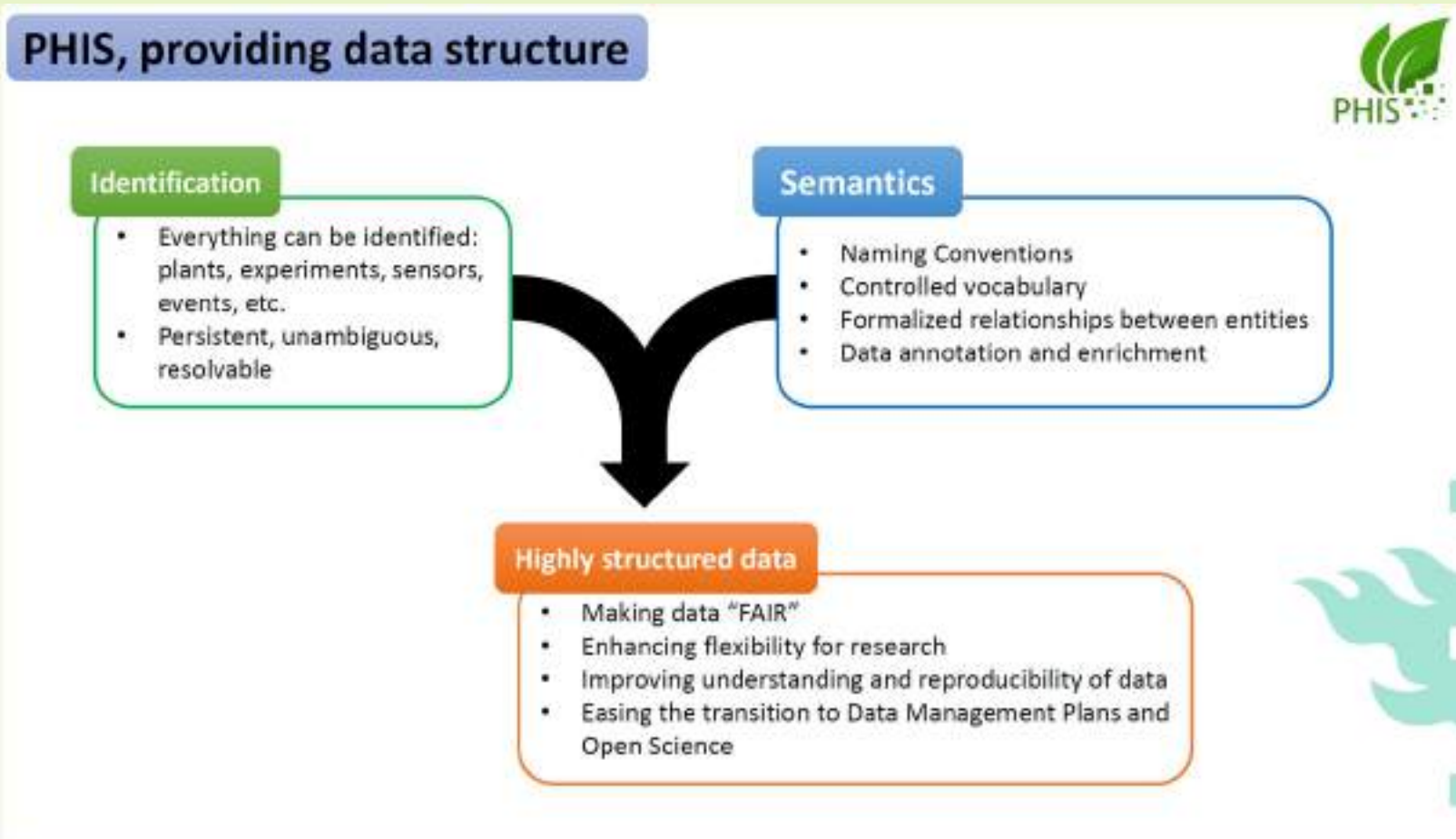
Potential tools of Data Management: PHIS



Potential tools of Data Management: PHIS



Potential tools of Data Management: PHIS



Potential tools of Data Management: PHIS

PHIS, providing data structure



Identification

- Everything can be identified: plants, experiments, sensors, events, etc.
- Persistent, unambiguous, resolvable

URI of plant:
<<http://phenome.fr/arch/2017/c17000118>>

URI of pot:
<<http://phenome.fr/arch/2013/pc13001542>>

URI of cart:
<<http://phenome.fr/arch/2013/ct1300123>>

URI of cabin:
<<http://phenome.fr/arch/2018/ac180015>>

URI of camera:
<<http://phenome.fr/arch/2018/ac180019>>



URI of image: <<http://phenome.fr/arch/2017/ic17002295855>>

Potential tools of Data Management: PHIS

PHIS, providing data structure

Semantics

- Naming Conventions
- Controlled vocabulary
- Formalized relationships between entities
- Data annotation and enrichment

Variables names follow precise structure:

Entity + **Characteristic** + **Method** + **Scale (unit)**

ShootSystem *Height* *ImageProcessing* *MM*

Morphological Param				
Variable name	Entity	Characteristic	Method	Units
ShootSystem_Height_ImageProcessing_Millimeter	ShootSystem	Height	ImageProcessing	Millimeter
ShootSystem_Width_ImageProcessing_Millimeter	ShootSystem	Width	ImageProcessing	Millimeter
ShootSystem_Perimeter_ImageProcessing_Millimeter	ShootSystem	Perimeter	ImageProcessing	Millimeter
ShootSystem_ProjectedArea_ImageProcessing_SquareMillimeter	ShootSystem	ProjectedArea	ImageProcessing	SquareMillimeter
ShootSystem_Compactness_Computing_unitless	ShootSystem	Compactness	Computing	unitless
ShootSystem_Height_LaserScanning_Millimeter	ShootSystem	Height	LaserScanning	Millimeter
ShootSystem_FreshBiomass_Weighing_Gram	ShootSystem	FreshBiomass	Weighing	Gram
ShootSystem_DryBiomass_Weighing_Gram	ShootSystem	DryBiomass	Weighing	Gram
Environmental Param				
Variable name	Entity	Characteristic	Method	Units
Air_Temperature_Datalogging_DegreeCelsius	Air	Temperature	Datalogging	DegreeCelsius
PAR_PPFD_Datalogging_MicroEinsteins	PAR	PPFD	Datalogging	MicroEinsteins
Air_RelativeHumidity_Datalogging_PCT	Air	RelativeHumidity	Datalogging	PCT

Potential tools of Data Management

Example of a domain for DM: UAV (drone) data

- PlotCut for the advanced stuff: Can handle some DM aspects too
- But a simple stand-alone homemade program for the easy-access to data
- And NMBU building something amazing with UAV data managing

Potential tools

of Data Management

PlotCut v.3 - A plugin for QGIS, to aid image-based plant research and breeding

The project's layers menu
From the top down Polygon (vectors) for defining extent of trial plots and extraction areas

The current layers.
An orthomosaic image (raster) from a drone mission of potato emergence. On top the trial (turquoise) and sam plot (beige) and extraction (purple) polygons.

The PlotCut v.3 plugin.
Here shown active on the tab "Image, template, layout" where the polygons for the trial are set within the trials extent, the plots size and a potential buffer to avoid plots' edge-effects being part of the extracts and subsequent analyses.

Matching image and rectangles.
The drone image is paired with trial info. There can be multiple polygon templates for each trial. Existing templates can be applied to a new mission's drone-based image, or a new template can be created

Defining the X and Y
The number of rows and columns of the desired extraction is defined. Either by entering the numbers or importing a CSV-file.

Specifics of a plot
The size of the plots and potential buffer is defined.

PlotCut v.3 is designed by Jesper Cairo Westergaard of UCPH, programmed by Bo Victor Thomsen of Aestas.dk and financed by grants for the GP3 project by NordGen.

Potential tools of Data Management – at the Danish e-infrastructure Consortium DEiC



The screenshot shows the DeIC DMP website. The header includes the DeIC logo, navigation links for Services, Get help, About DeIC, DA, and Contact, and a search icon. A secondary navigation bar lists Security, Forskingsnettet, Data Management, Supercomputing (HPC), and Quantum Infrastructure. The main content area is titled "DeiC DMP" and includes a description of the tool, a list of links for account creation and guides, and sections for Research data management, Templates, and FAIR.

DeiC

Services Get help About DeIC DA Contact

Security Forskingsnettet Data Management Supercomputing (HPC) Quantum Infrastructure

Home / Data Management / DeiC DMP

DeiC DMP

The Online tool DeiC DMP helps researchers write data management plans - plans describing how project data are generated, collected, stored, used and reused, and how others can access the data at a later stage.

Some institutions and funders have a requirement to prepare a data management plan and recommend the use of the DeiC DMP.

[Create DeiC DMP account](#) →

[DeiC DMP guides](#) →

Research data management

Data management plans provide an overview of how research data is collected, stored, processed and shared with others. Thus, the data management plan is an important element in the planning of research projects to ensure that they can be carried out in a safe and responsible manner.

Templates

The DeiC DMP can be customized and contains guides and various templates from a number of research institutions and funders, such as EU Horizon 2020.

FAIR

A data management plan can also help you establish FAIR data management practices in your project. FAIR stands for "Findable, Accessible, Interoperable, Reusable" and describes how to make research data searchable, accessible and interoperable for reuse. Applying the FAIR principles in your research project may be a requirement from selected funders.

Open Source

The DeiC DMP is based on the open source codebase (DMPRoadmap) jointly developed by the Digital Curation Centre (DCC) and the University of California Libraries Center (UCS).

Potential tools of Data Management – at the Danish provider DEiC

Create a new plan

Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.

* What research project are you planning?

mock project for testing, practice, or educational purposes

* Select the primary research organisation

Organisation

- or - No research organisation associated with this plan or my research organisation is not listed

* Select the primary funding organisation

Funder

- or - No funder associated with this plan or my funder is not listed

Which DMP template would you like to use?

We found multiple DMP templates corresponding to your funder.

Potential tools of Data Management – at the Danish provider DEiC

The screenshot displays the 'OCH - One Crop Health' data management tool interface. At the top, there is a navigation bar with tabs for 'Project Details', 'Contributors', 'Plan overview', 'Write Plan', 'Share', 'Request feedback', and 'Download'. Below the navigation bar, there are links for 'expand all | collapse all' and a progress indicator '0/41'. The main content area consists of a list of nine sections, each with a plus sign on the right side, indicating that they can be expanded. The sections are:

- 0. Project information (0 / 3)
- 1. Data description (0 / 2)
- 2. Rights to research data (0 / 5)
- 3. Ethical and legal approvals (0 / 2)
- 4. Collection, processing and documentation (0 / 5)
- 5. Storage and information security (0 / 6)
- 6. Data sharing (0 / 11)
- 7. Long term preservation (0 / 3)
- 8. Resources and responsibilities (0 / 4)

Potential tools of Data Management – at the Danish provider DEiC

OCH - One Crop Health

Project Details Contributors Plan overview Write Plan Share Request feedback Download

Set plan visibility

Public or organisational visibility is intended for finished plans. You must answer at least 50% of the questions to enable these options. Note: test plans are set to private visibility by default.

- Private: restricted to you and your collaborators.
- Organisational: all registered DEiC DMP users from your organisation can view.*
- Public: anyone can view your plan in the Public DMPs list.

* Note that the affiliation is defined by the users themselves and their profile and not controlled by the organisation.

Manage collaborators

Invite specific people to read, edit, or administer your plan. Invitees will receive an email notification that they have access to this plan.

Email address	Permissions
wnk239@ku.dk	Owner

Invite collaborators

* Email

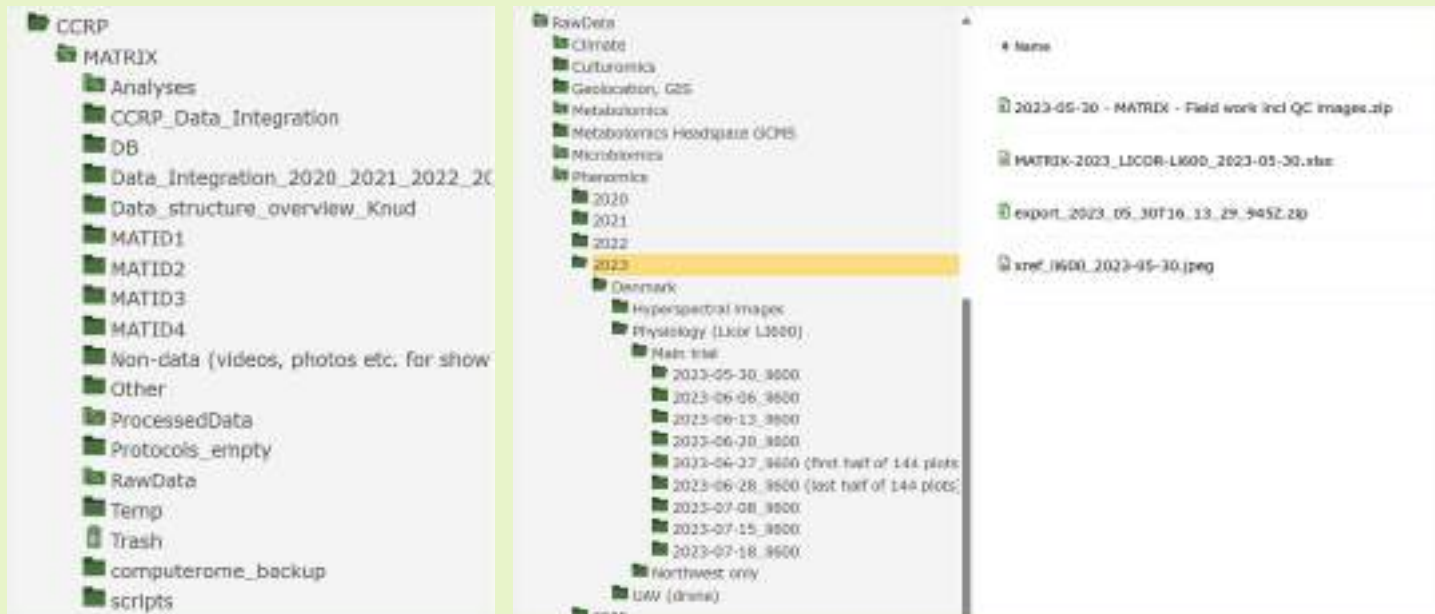
* Permissions

- Co-owner
- Editor
- Read only

Submit

Potential tools of Data Management

The MATRIX project at the ERDA portal of UCPH as an example

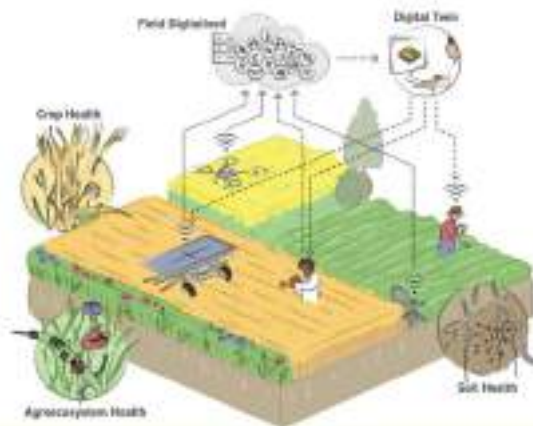


Potential tools of Data Management

The OCH project Margin Maps/QGIS as an example

One Crop Health: Next generation crop protection

The One Crop Health project integrates the latest advances in agri-tech, biotechnology, ecology, and data sciences to reduce reliance on pesticides in agroecosystems.



One Crop Health for Next-Generation Crop Protection is a research consortium led by the University of Copenhagen, with partners at Aarhus University, Rothamsted Research, and the University of Sheffield. Funded by the Novo Nordisk Foundation and running from 2024 to 2029, the project aims to reduce pesticide use in cropping systems through innovative, sustainable approaches.

Inspired by the One Health concept from human health, the project promotes a holistic One Crop Health framework that connects plant, soil, and agroecosystem health. By combining agroecological practices with cutting-edge technologies in data science, robotics, and ecology, the team will develop tools for early detection and prediction of pest, weed, and disease outbreaks.

The goal is to minimize chemical interventions, enhance biodiversity, improve soil health, and build resilient cropping systems for a more sustainable agricultural future.

Potential tools of Data Management

The OCH project Mergin Maps/QGIS as an example

Let's jump into QGIS and take a look

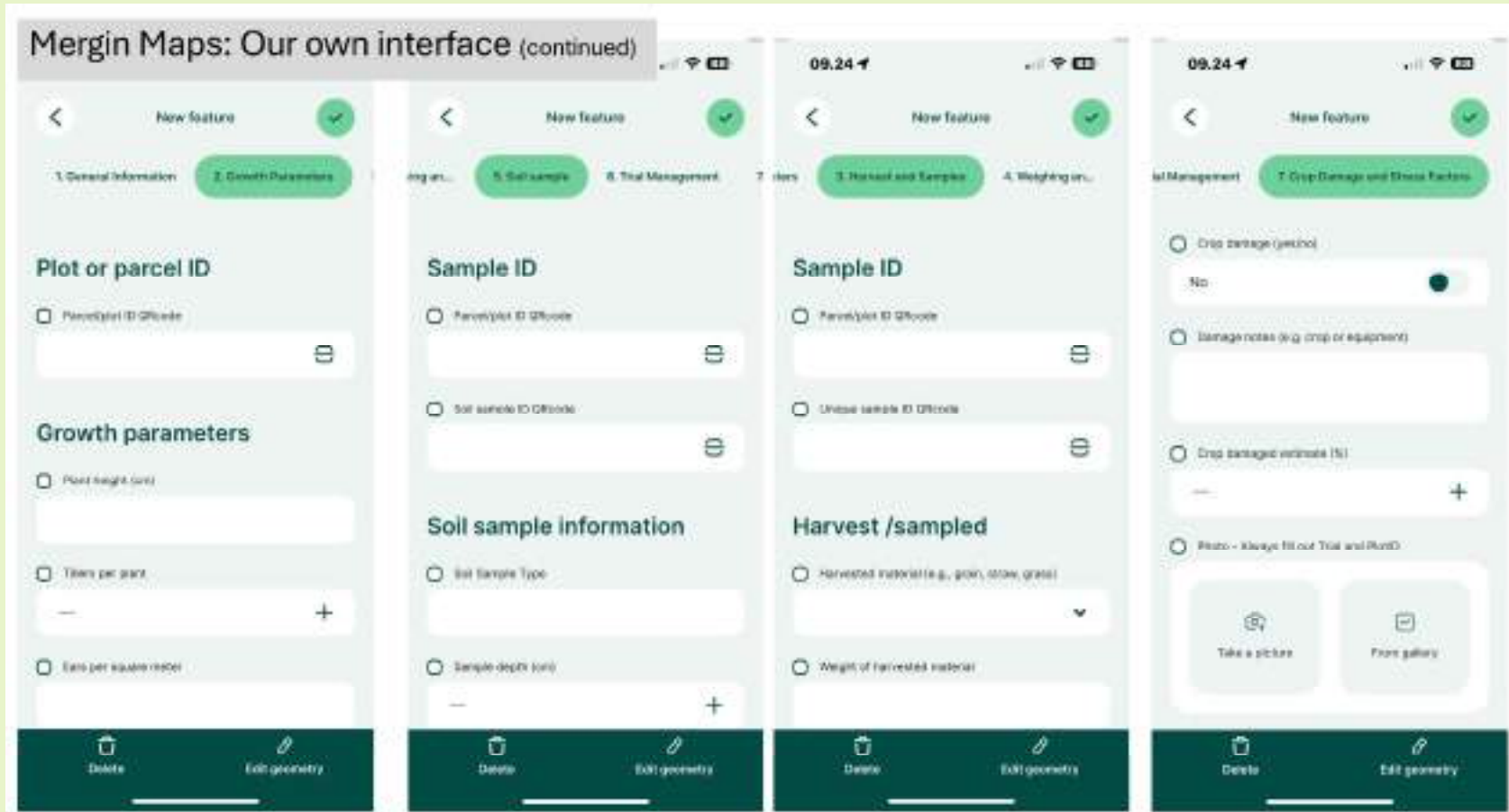
Notice for for post-meeting
viewers of these slides:
Here was a live view showing
of a Mergin Maps project
as they look in QGIS

Potential tools of Data Management

Mergin Maps: Our own interface



Potential tools of Data Management



Questions? Comments?

and sharing of experiences with DM - good or bad

- Anyone?

- _____

- _____

- _____

- _____

Thank you for listening

Feel free to reach out to me: jcw@plen.ku.dk

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE



**BECAUSE WE TRY TO ANSWER EVEN
THE CRAZIEST QUESTIONS**



science.ku.dk

The Road to Precision: Spectral Imaging for Abiotic Stress Responses- Are We There Yet?

*Assoc. Prof. Sónia Negrão
University College Dublin
School of Biology & Environmental Science*



sonia.negrao@ucd.ie



An Roinn Talmhaíochta,
Bia agus Mara
Department of Agriculture,
Food and the Marine

**2022 Agrobiodiversity
154**



Taighde Éireann
Research Ireland

18/FRL/6197

What is the plant phenome?

Environment



Interaction



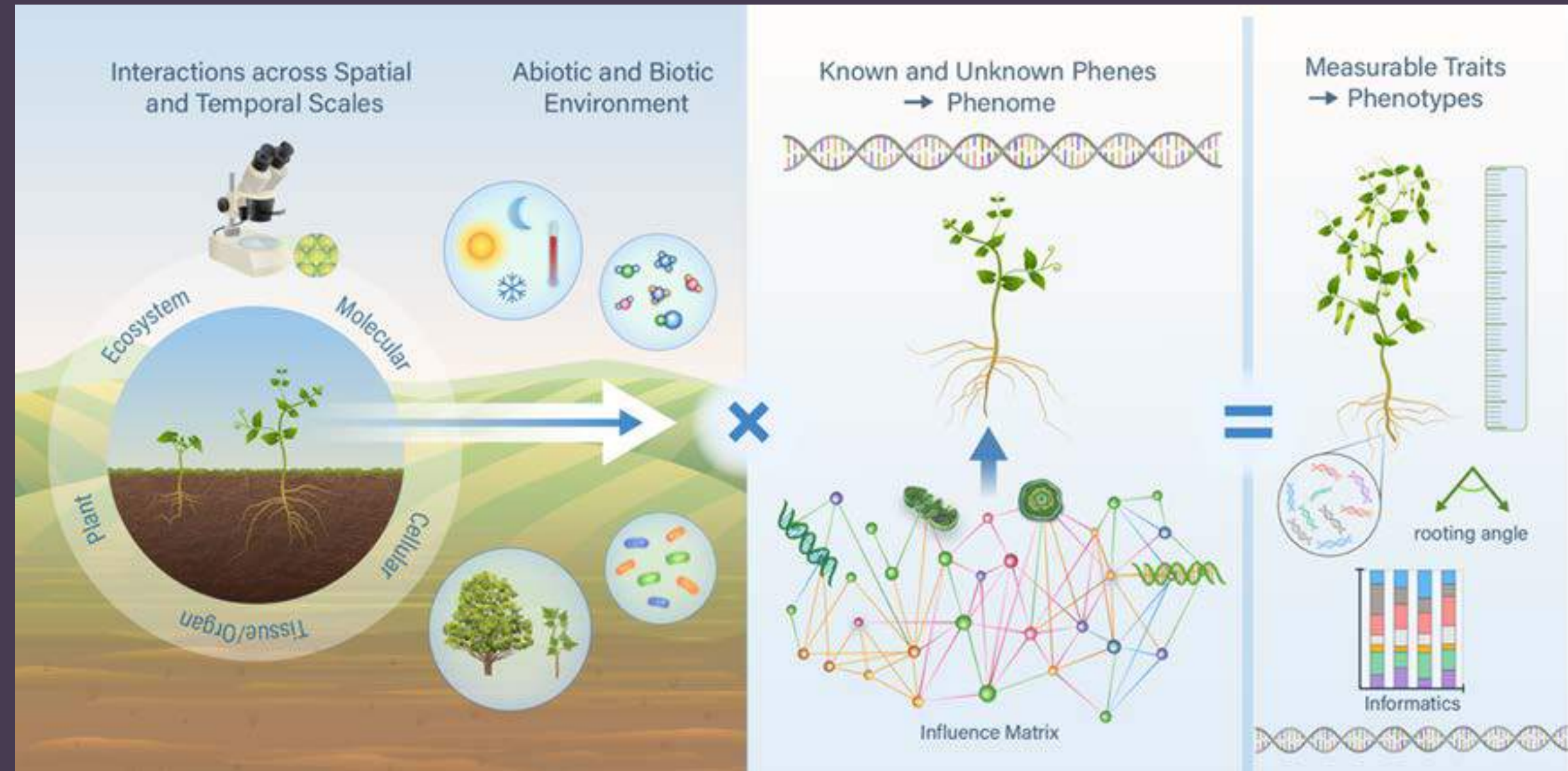
Genotype

Phenotype

“The phenome quantifies how an organism interprets an environment with a set of phenes to form an optimal multi-scale phenotype.”

What is Plant Phenomics?

Multidisciplinary scientific discipline that systematically studies the "phenome"—the set of a plant's possible phenotypes and its spatio-temporal interactions with the environment



Who is a Plant Phenomicist?

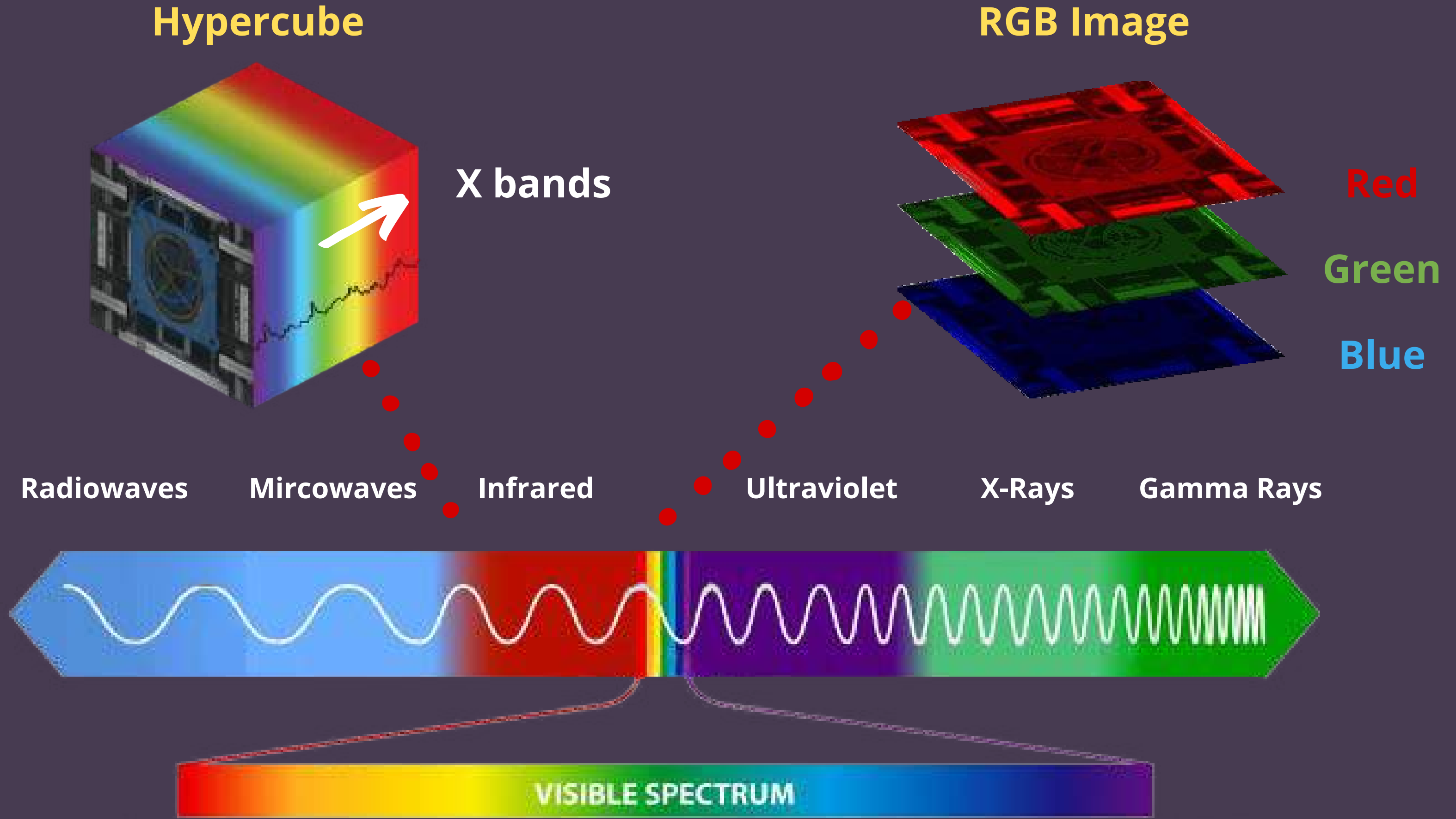


“A plant phenomicist defines formal systems to discover the dynamics of plant-environment interactions”

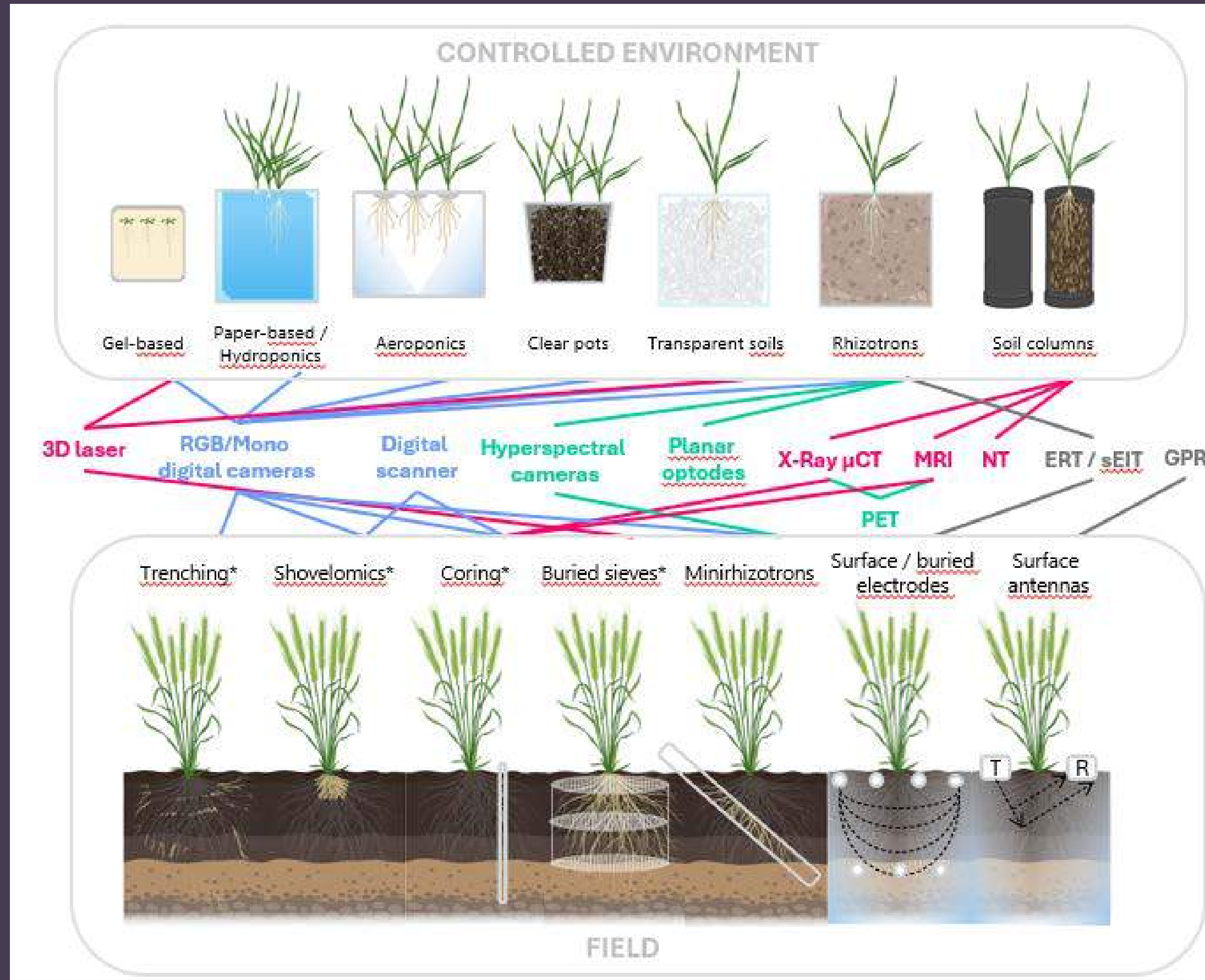
How can we assess abiotic stress using imaging?



RGB vs Hyperspectral



What about the hidden half?






GUINNESS
JAMES'S GATE BREWERY

FEELIN
WHISKEY

**Ireland- the land of Guinness
and Irish Whiskey**



Barley in Ireland

First cereal crop

€1.4 B value in drinks

Barley

Fourth most important cereal crop in the world

Human food, animal feed and production of alcohol

SDG 9 Industry, Innovation and Infrastructure



IRISH WHISKEY DISTILLERIES MAP

● Planned
● New
● Existing

IRISH WHISKEY DISTILLERIES

- 01 Alltech Distillery - Dublin
- 02 Ballykeefe Distillery - Kilkenny
- 03 Blackwater Distillery - Waterford
- 04 Boann Distillery - Louth
- 05 Bushmills Distillery - Antrim
- 06 Cape Clear Distillery - Cape Clear Island
- 07 Cooley Distillery - Louth
- 08 Dingle Distillery - Kerry
- 09 Dublin Whiskey Company - Dublin
- 10 Echlinville Distillery - Down
- 11 Glendalough Distillery - Wicklow
- 12 Great Northern Distillery - Louth
- 13 Irish Distillers - Cork
- 14 Kilbeggan Distillery - West Meath
- 15 Lough Gill Distillery - Sligo
- 16 Nephin Irish Whiskey - Mayo
- 17 Powerscourt Whiskey - Wicklow
- 18 Rademon Estate Distillery - Down
- 19 Renegade Spirits - Waterford
- 20 Slane Castle Distillery - Meath
- 21 Sliabh Liag Distillery - Donegal
- 22 Teeling Whiskey Company - Dublin
- 23 The Connacht Whiskey Company - Mayo
- 24 The Shed Distillery - Leitrim
- 25 Tullamore Distillery - Offaly
- 26 Walsh Whiskey Distillery - Carlow
- 27 West Cork Distillers - Cork



Source: agriculture.gov.ie

The Impact of Terroir on the Flavour of Single Malt Whisk(e)y New Make Spirit

by [Maria Kyraleou](#)¹, [Dustin Herb](#)², [Grace O'Reilly](#)³, [Neil Conway](#)³, [Tom Bryan](#)⁴ and [Kieran N. Kilcawley](#)^{1,*}

Foods 2021, 10(2), 443; <https://doi.org/10.3390/foods10020443>

SDG 13 Climate Action



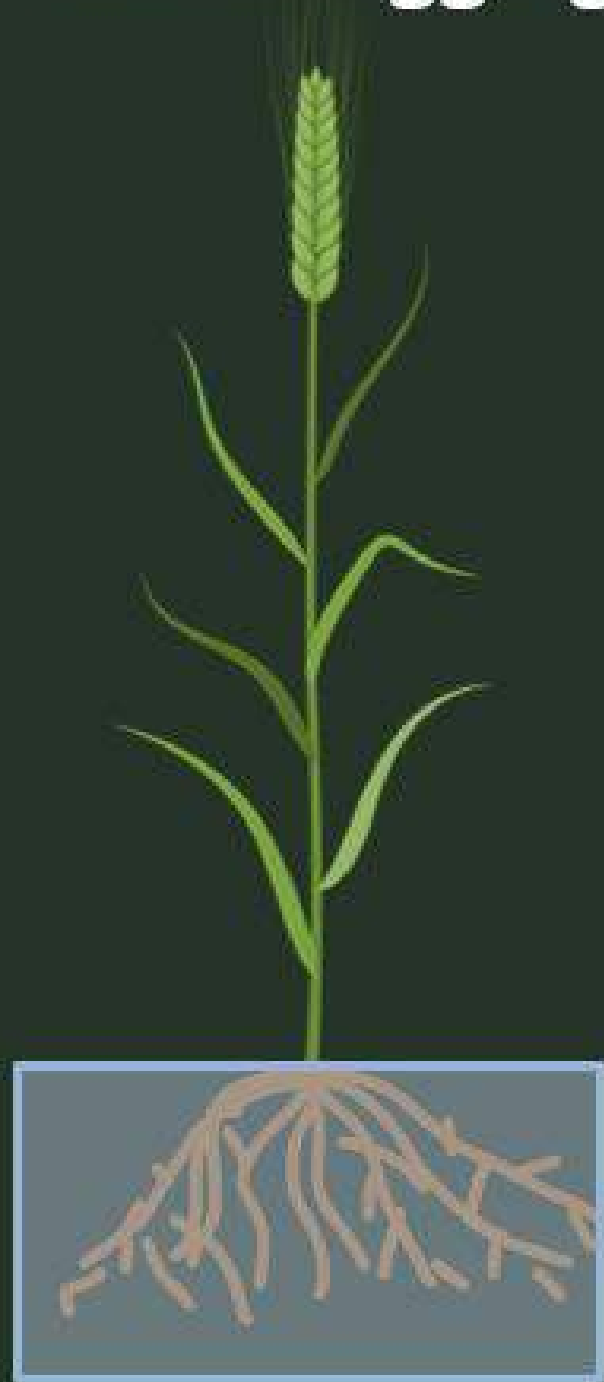
THE SEASONS OF IRELAND

Waterford Distillery 2019



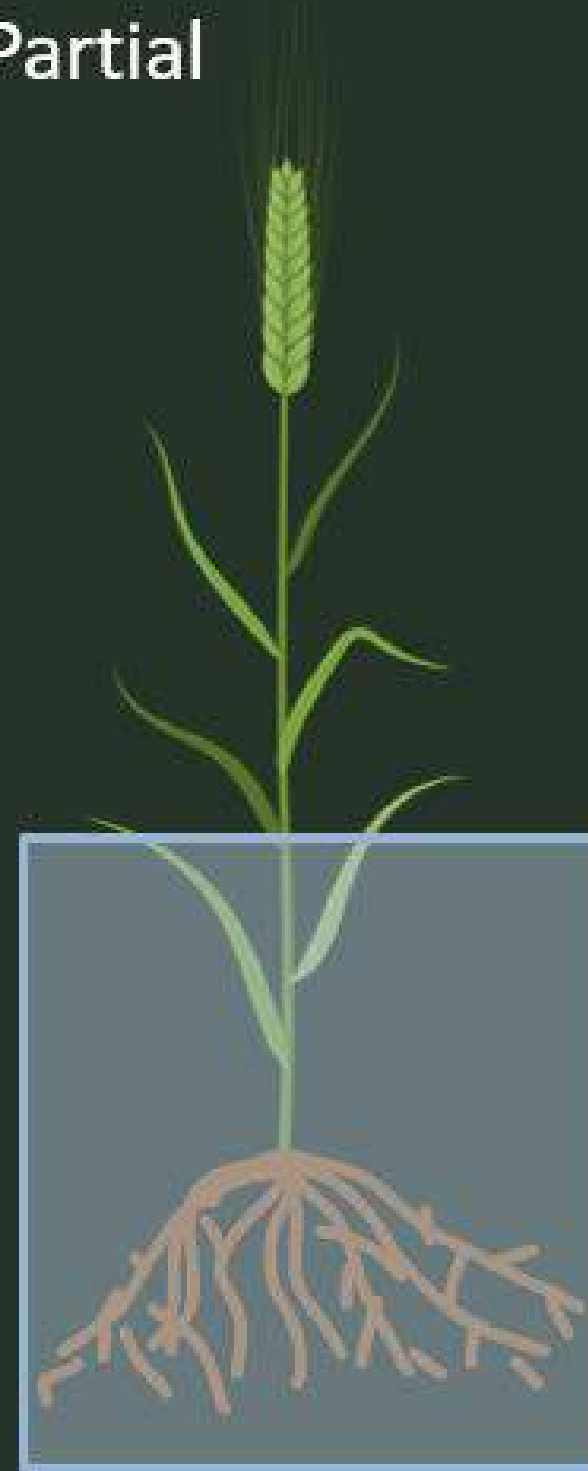
What is waterlogging?

Waterlogging



Submergence

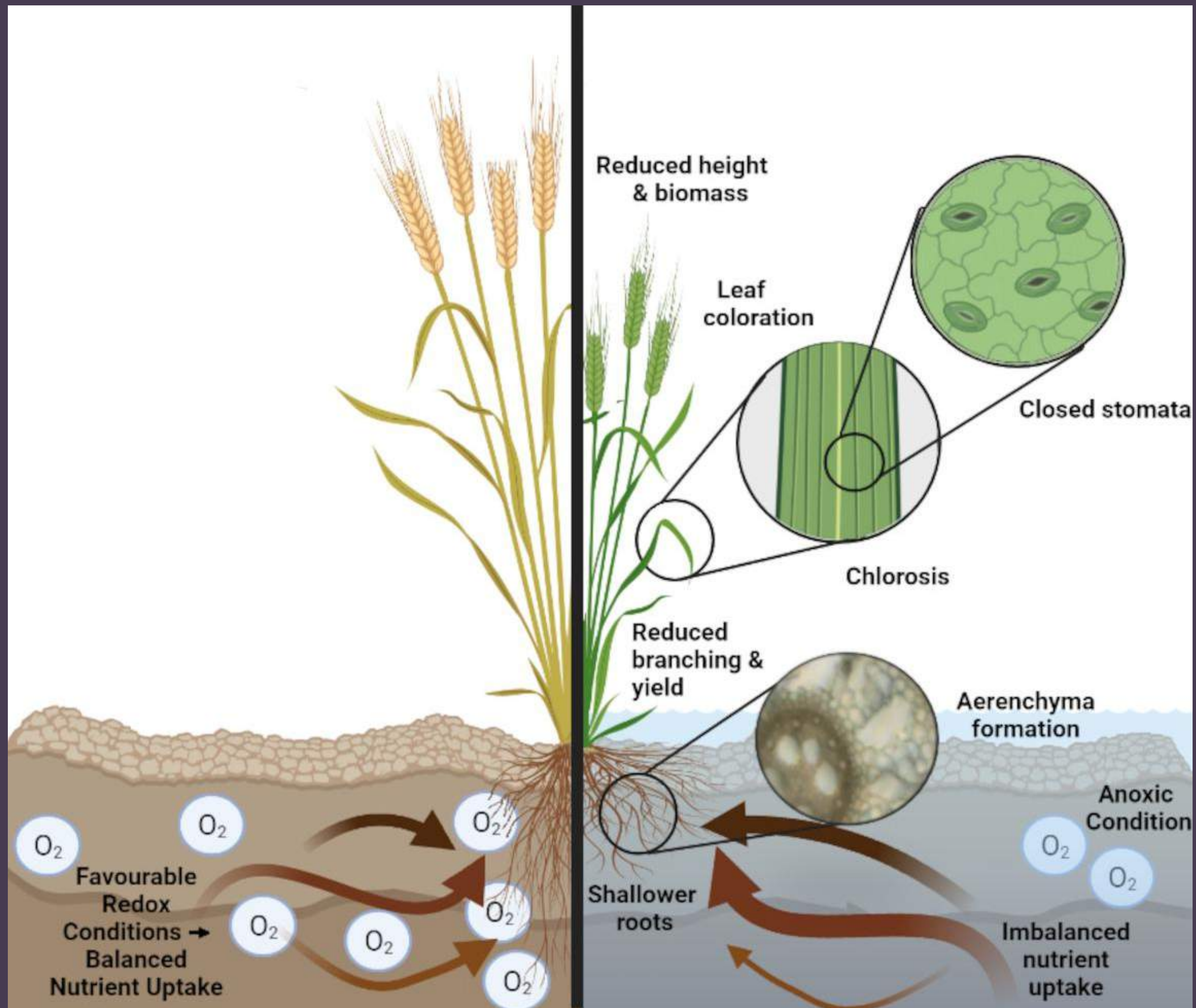
Partial



Complete



Air: 21% O₂
Hypoxia: < 21% O₂
Anoxia: 0% O₂



Genetic diversity is key for stress tolerance

Hypothesis:

Modern cultivars were bred under optimal conditions, favouring yield-traits

Heritage germplasm have higher stress tolerance, but remains largely uncharacterized



ExHIBiT

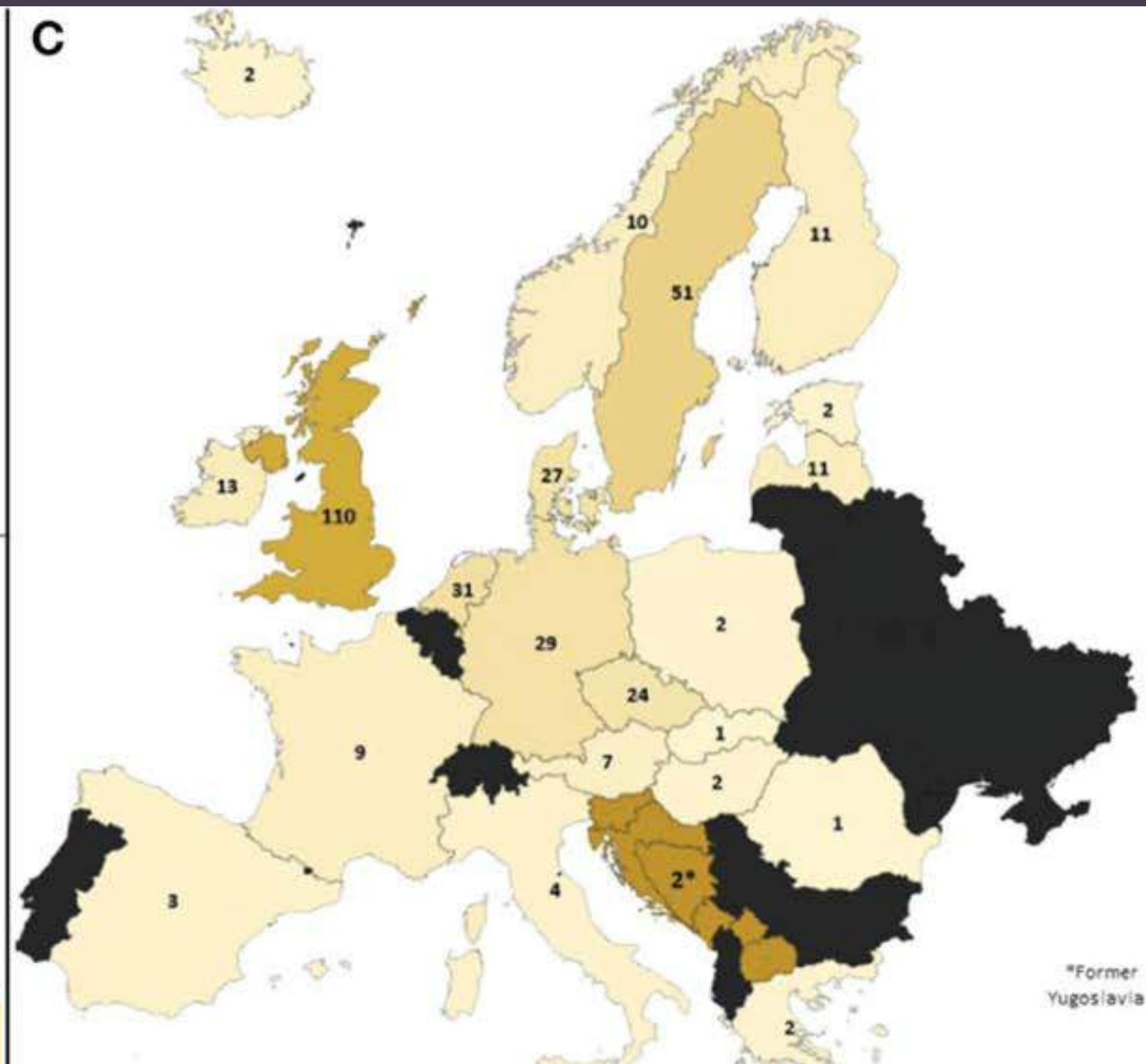
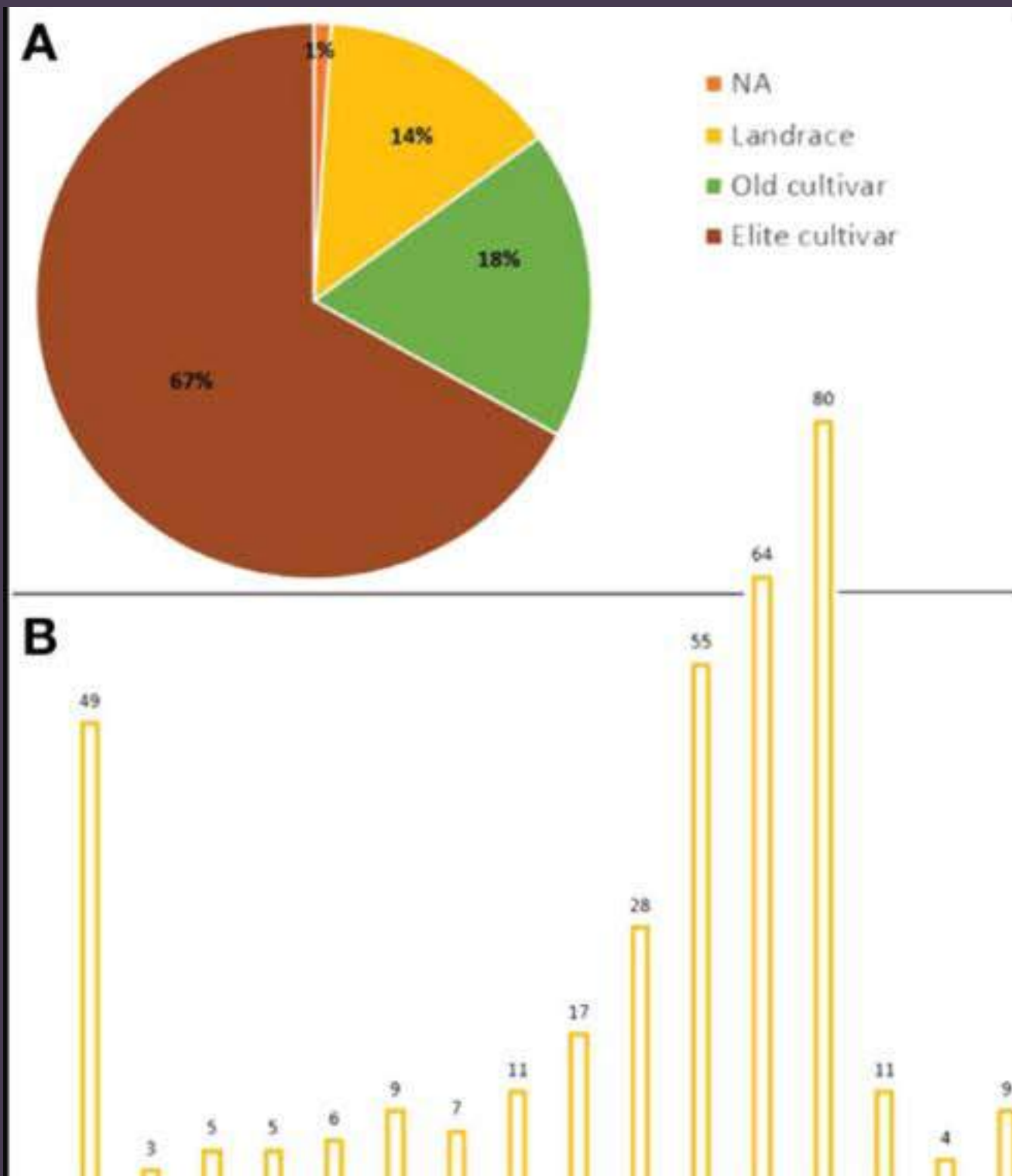
European Heritage Barley collection



JHI team

Photo by Villő Bernád

ExHIBiT



Northern Europe
2-row Spring barley
363 accessions



Bernád et al. (2024) *FiPS*: [e1268847](https://doi.org/10.1111/fips.12688)

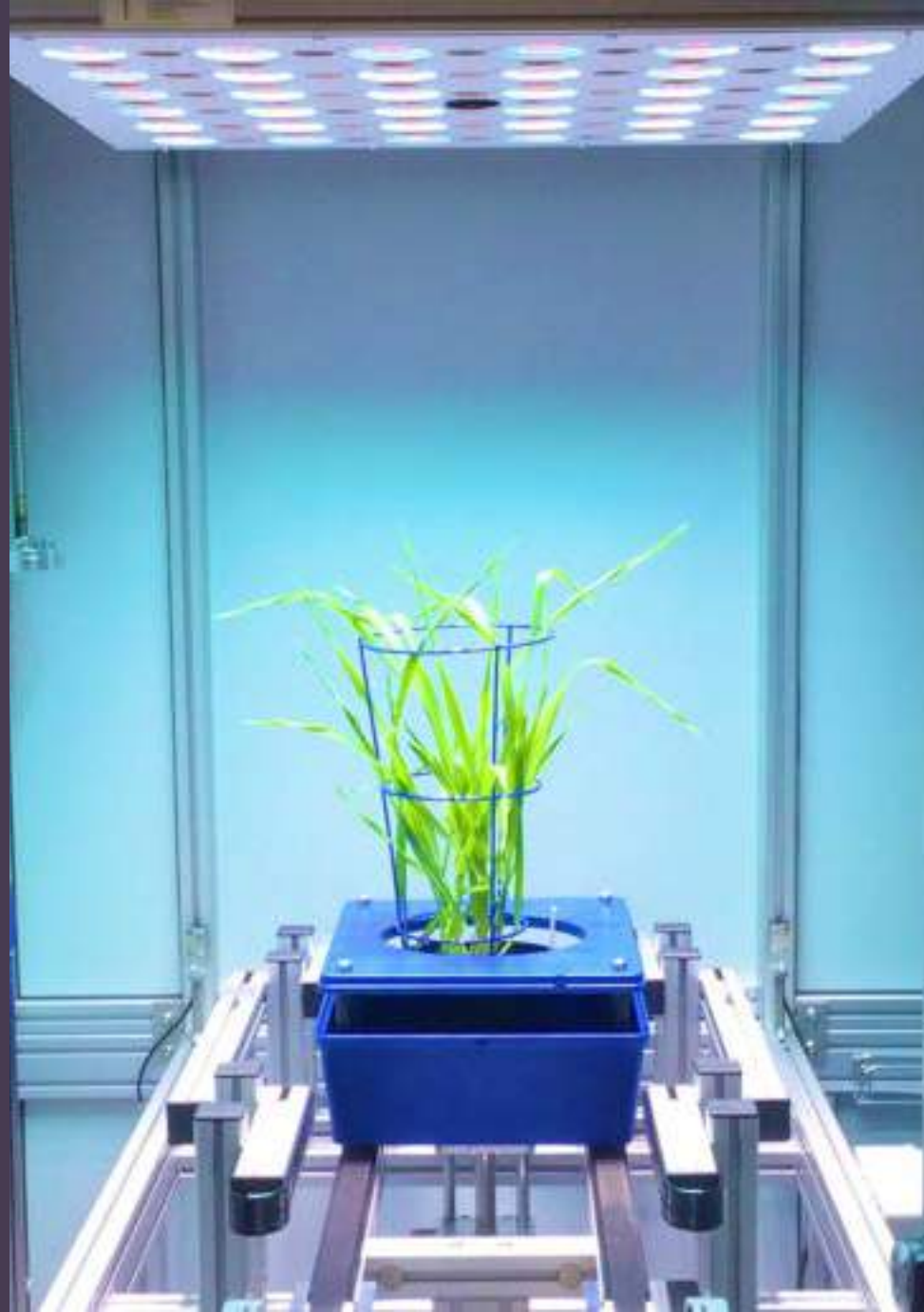
Establishment of
the ExHIBiT core
collection:

*From 363 to 230
lines*



Two phenotyping conditions

Under controlled conditions

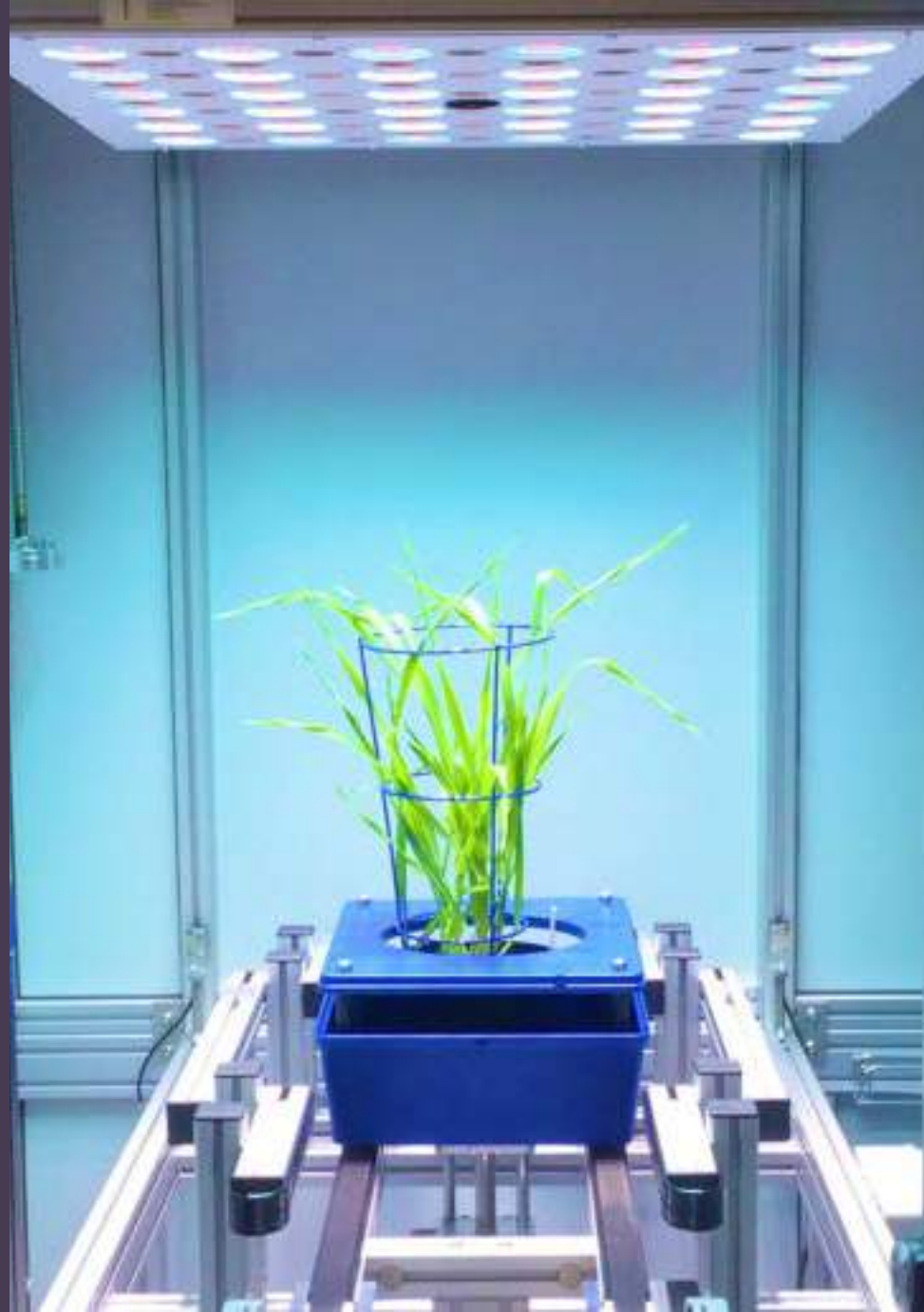


Under field conditions



Two phenotyping conditions

Under controlled conditions



Under field conditions



Shoot Data

UCD

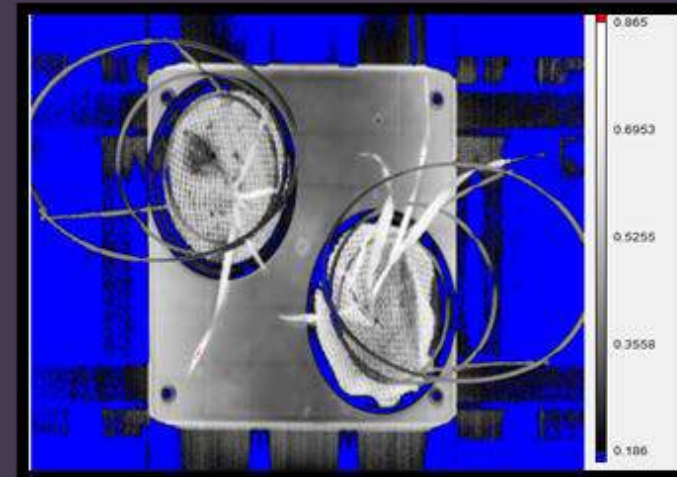


UPJV



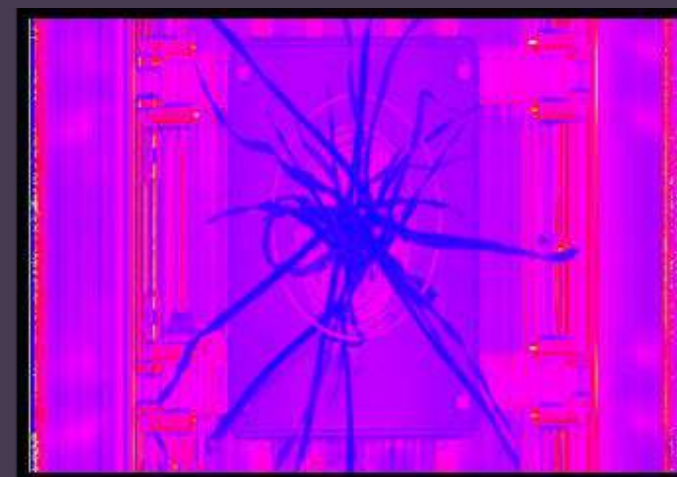
RGB Imaging Sensor

Provides biomass and shoot architecture.
(465 - 630nm)



Visible Near Infrared

Pigment content and health status.
(380 - 900nm)



Short Wave Infrared

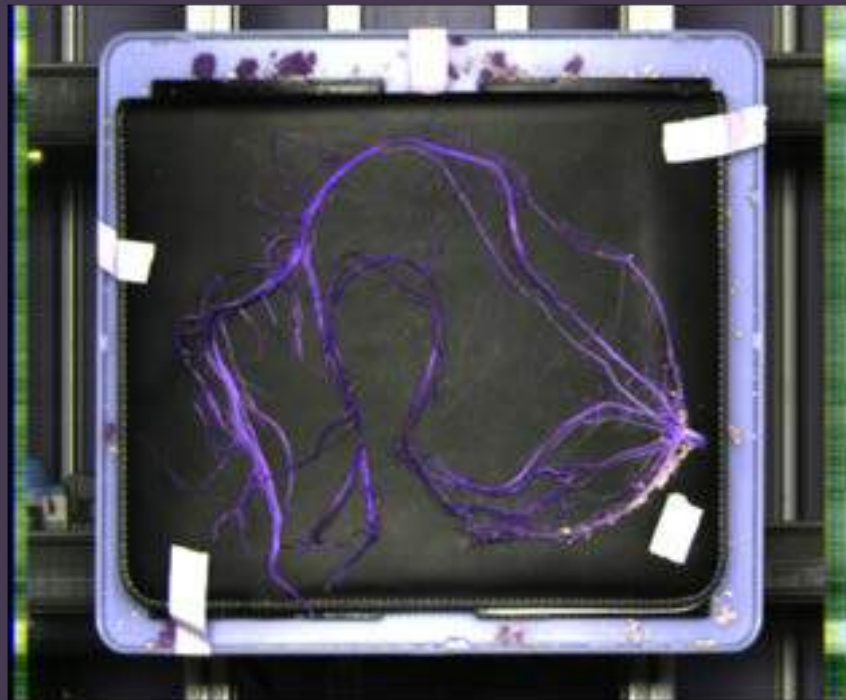
Highlight moisture content, water distribution
(900 - 1700nm)

Root Data

UPJV



Daily measurements



Time-point specific

Monochromatic

Growth, angle, diameter etc...

Visible Near Infrared

Pigment content and health status.
(380 – 900nm)

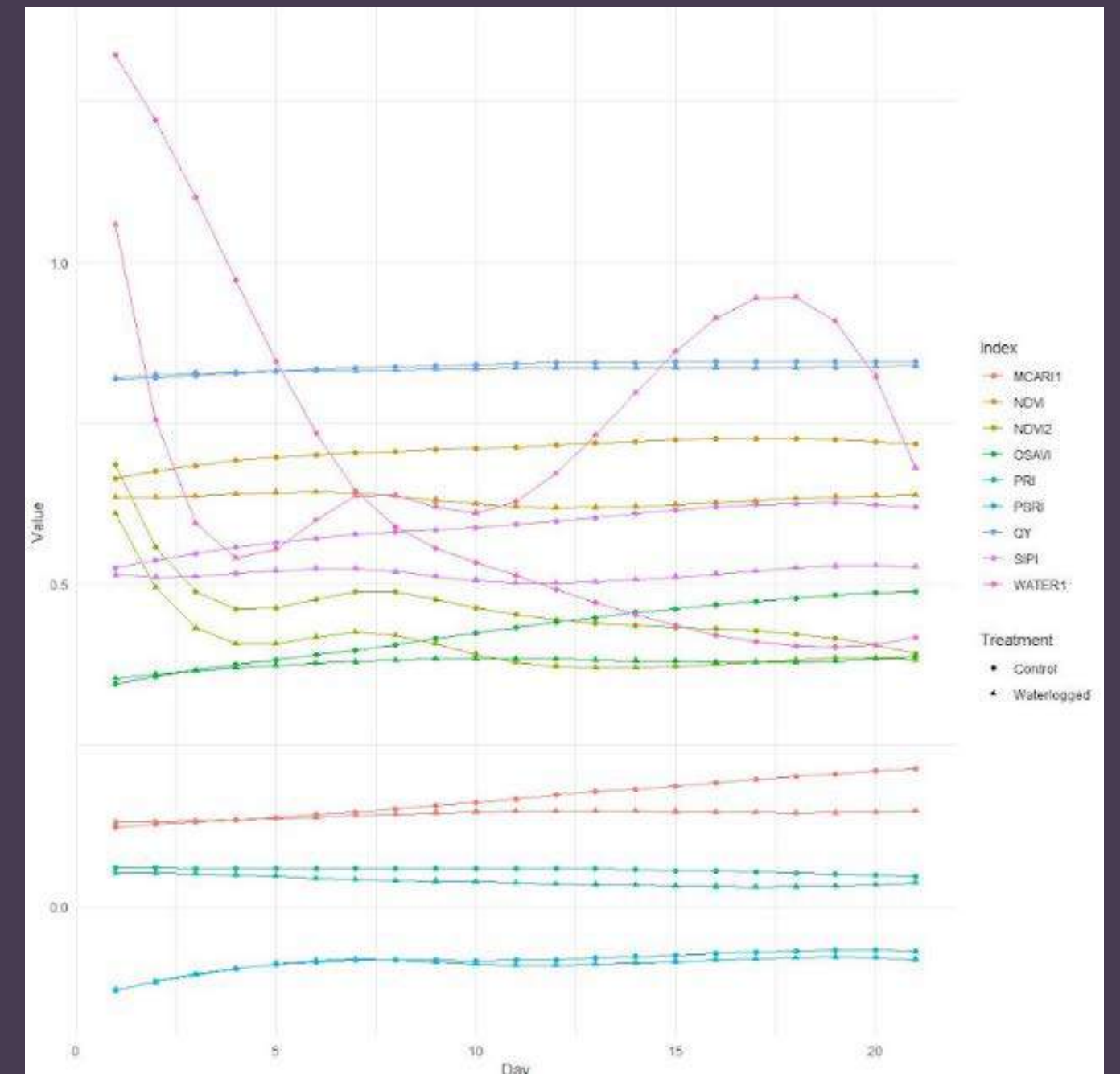
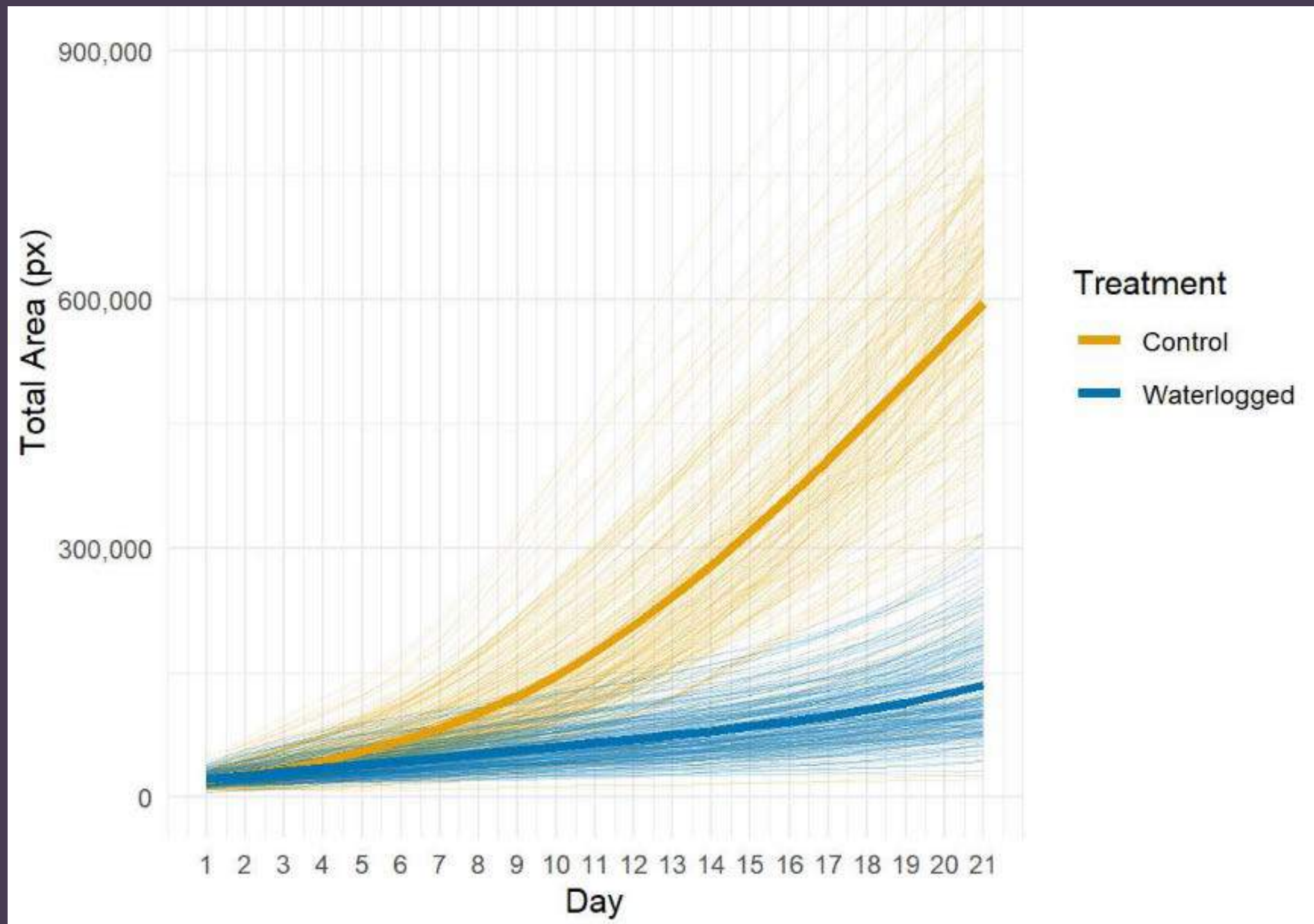
Short Wave Infrared

Highlight moisture content, water
distribution
(900 - 1700nm)

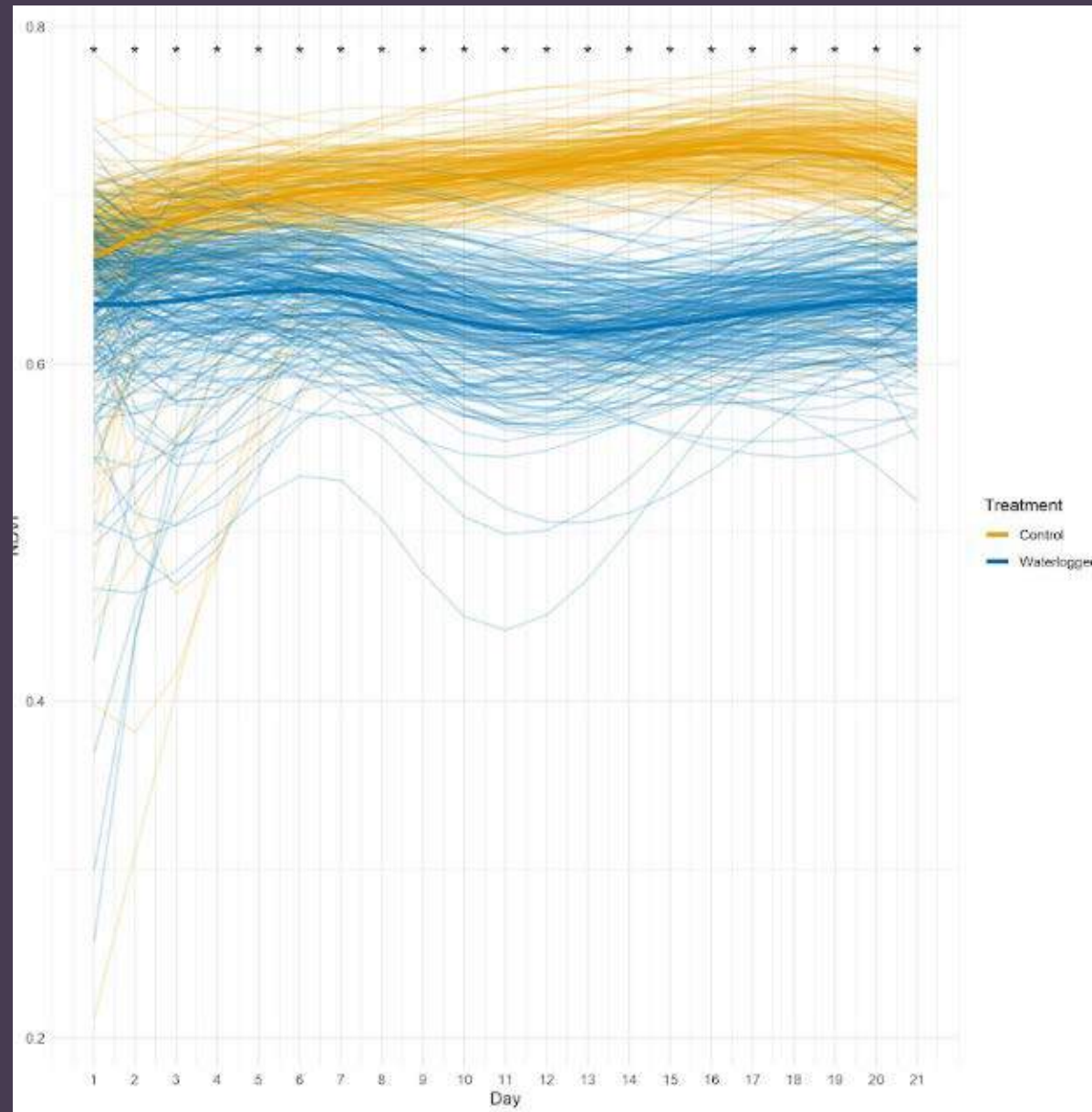
Challenges ahead



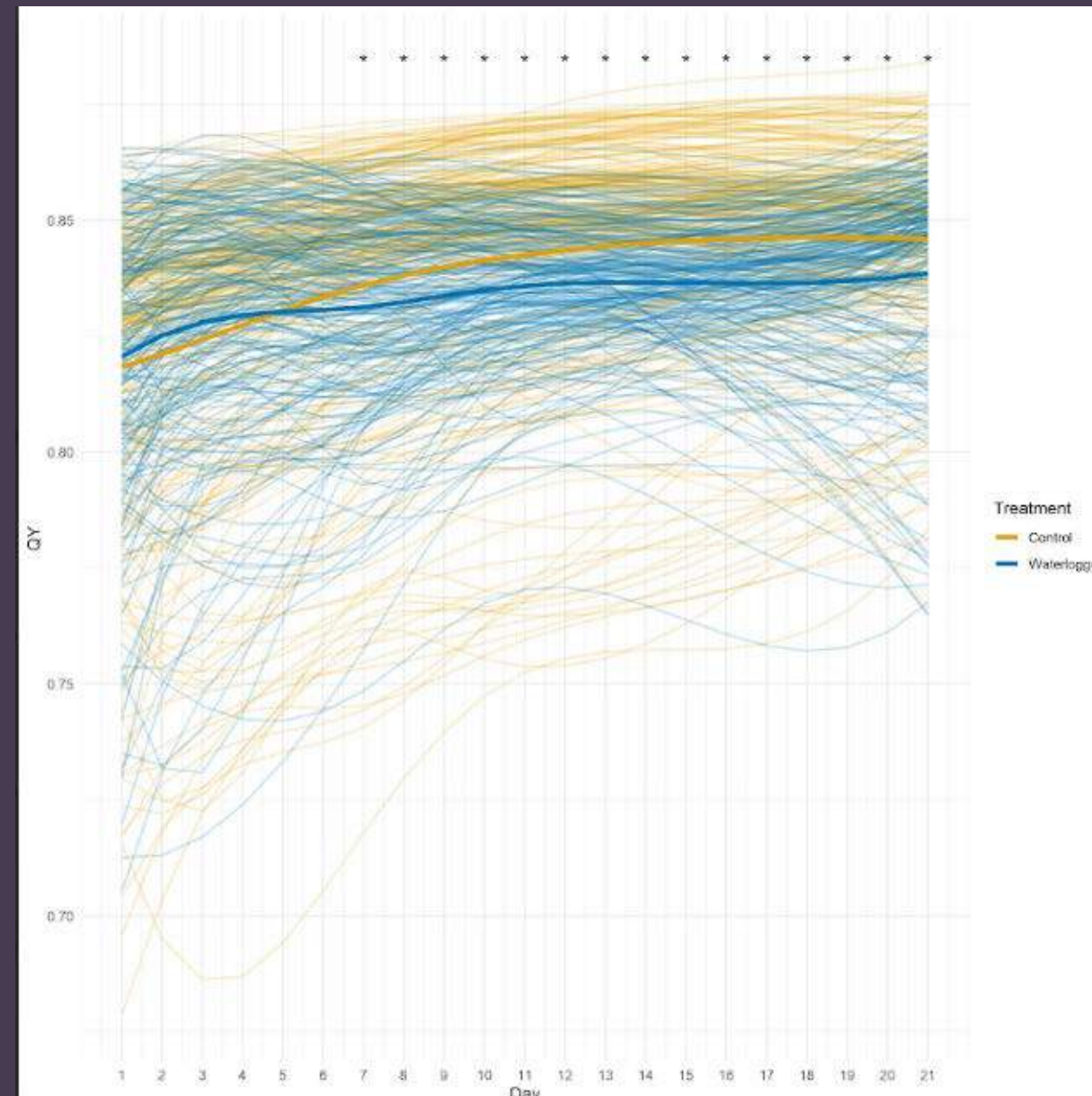
Plant responses to stress are dynamic



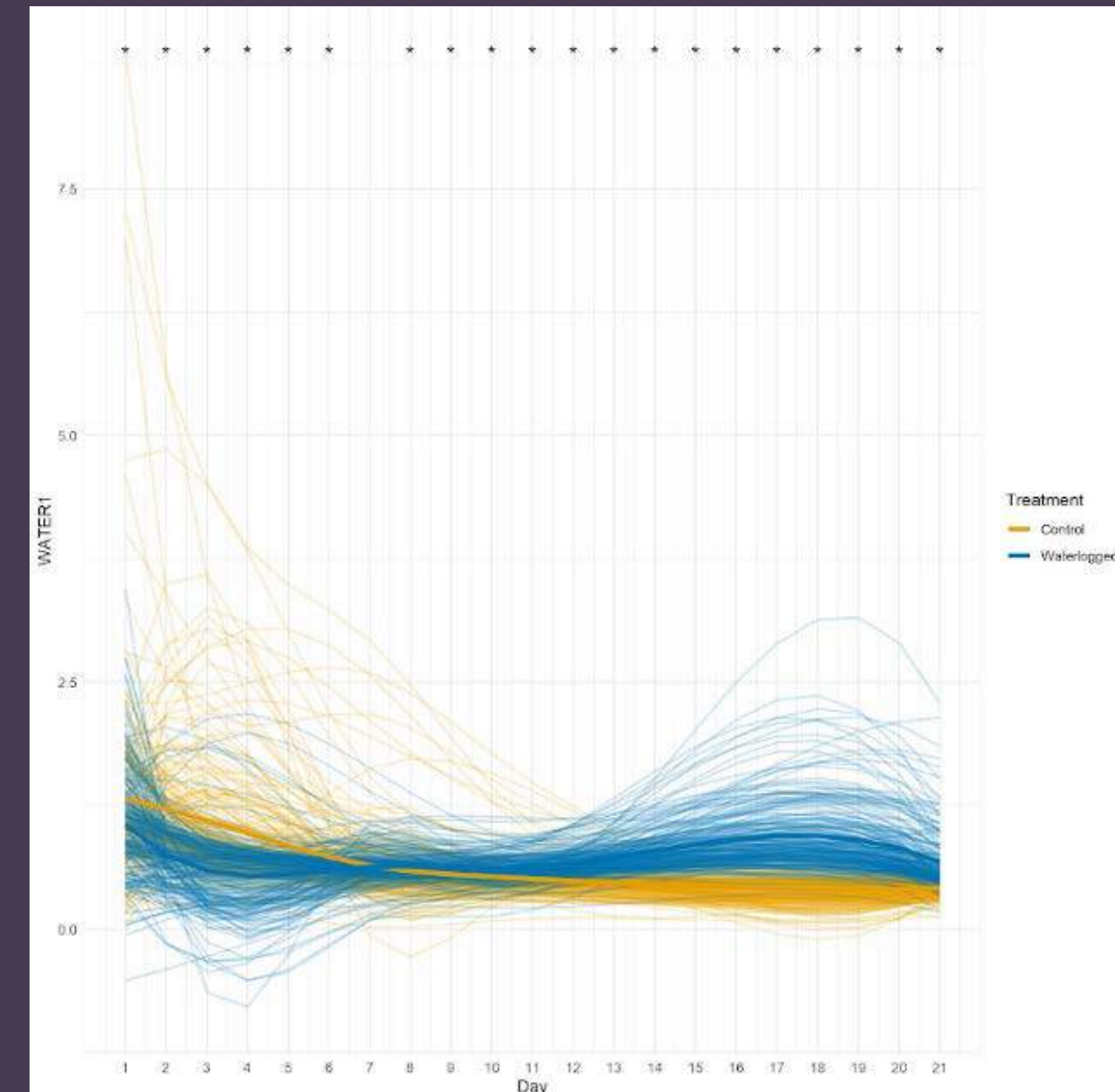
How can we interpret complex data?



**Normalized Difference
Vegetation Index (NDVI)**



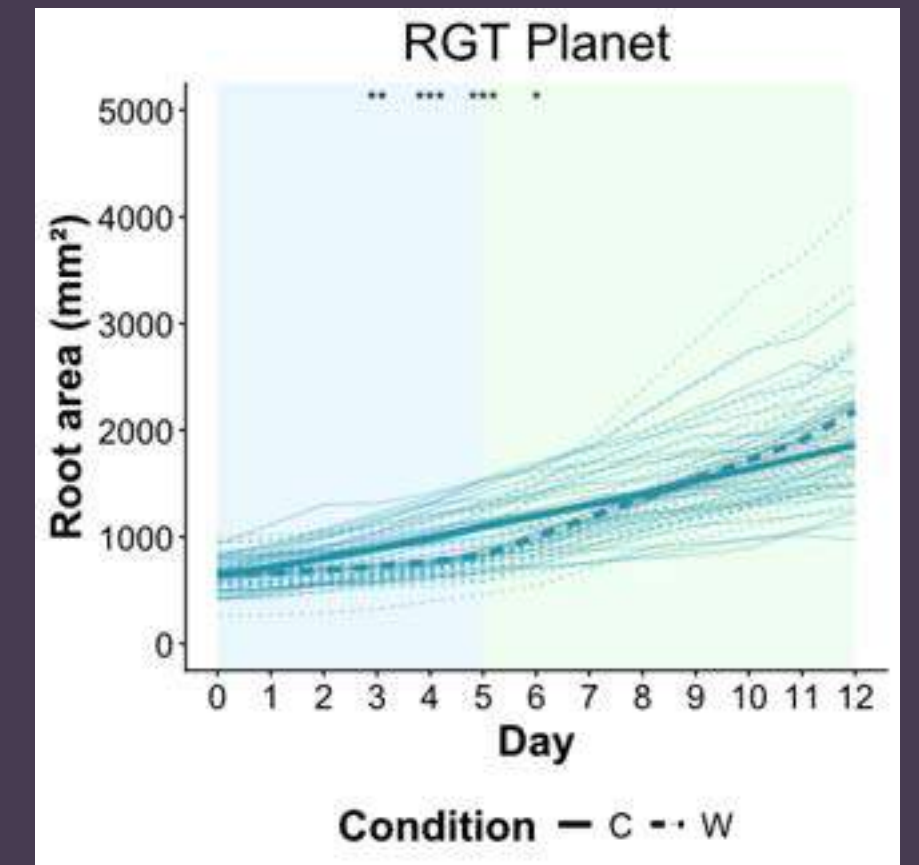
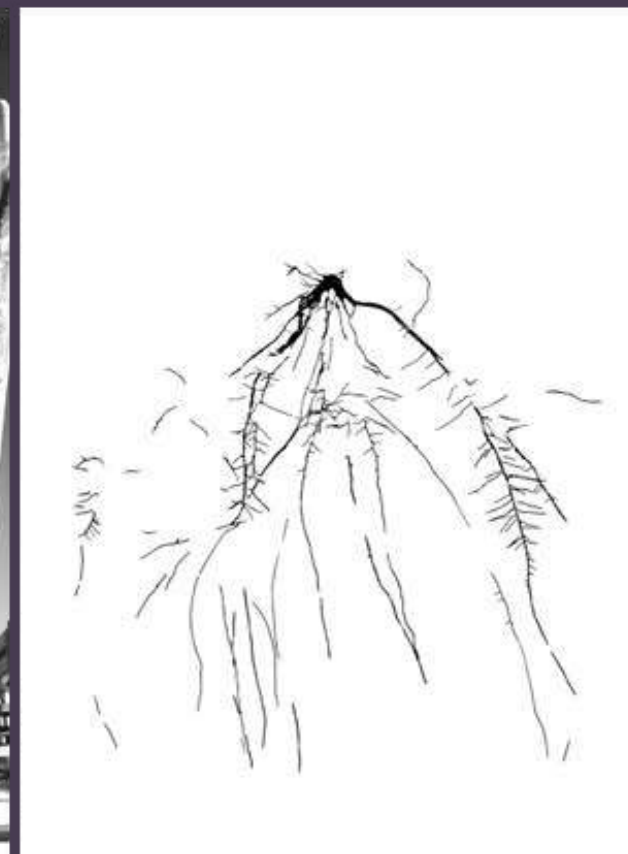
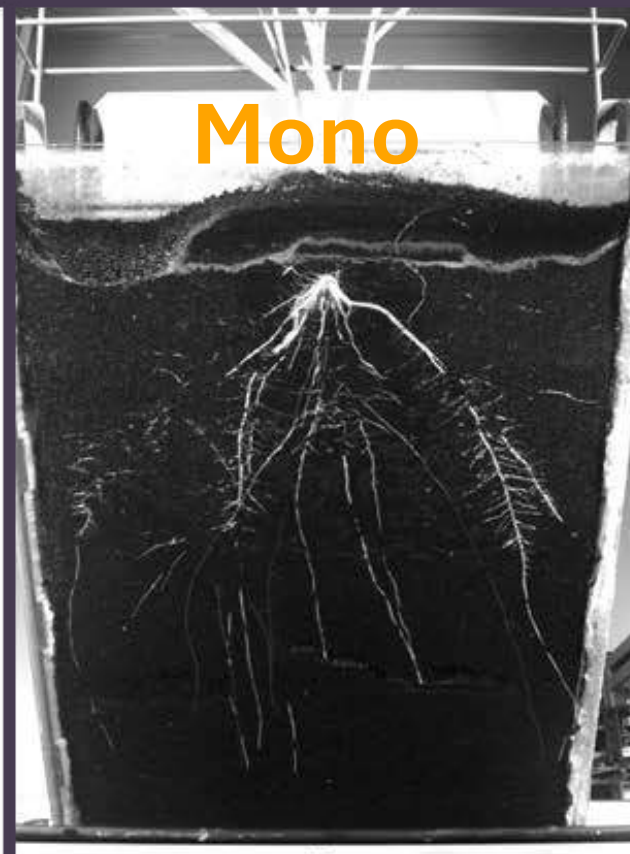
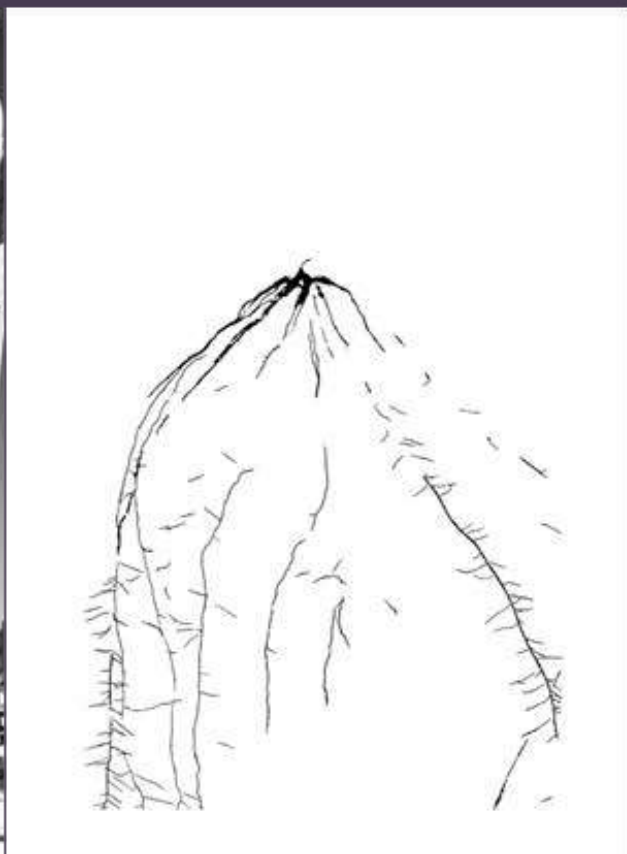
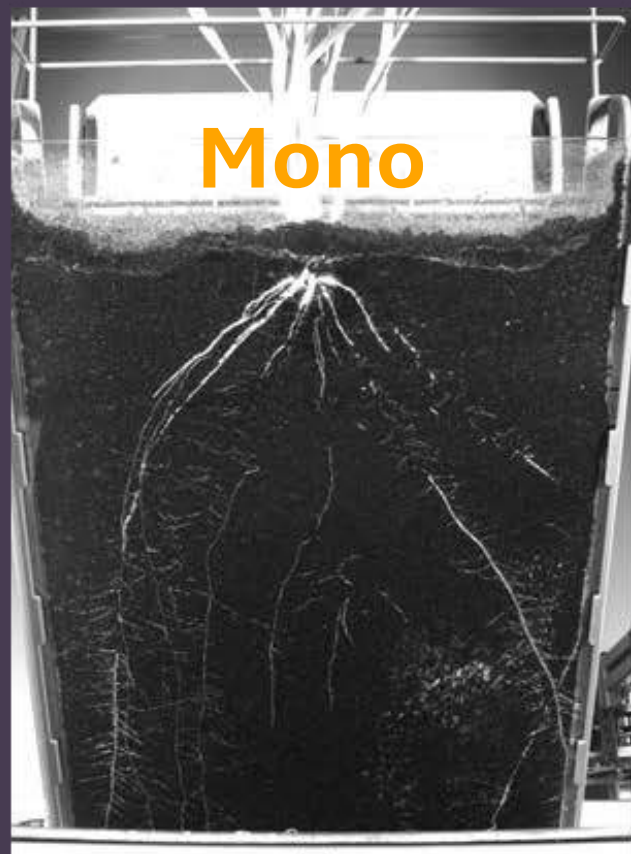
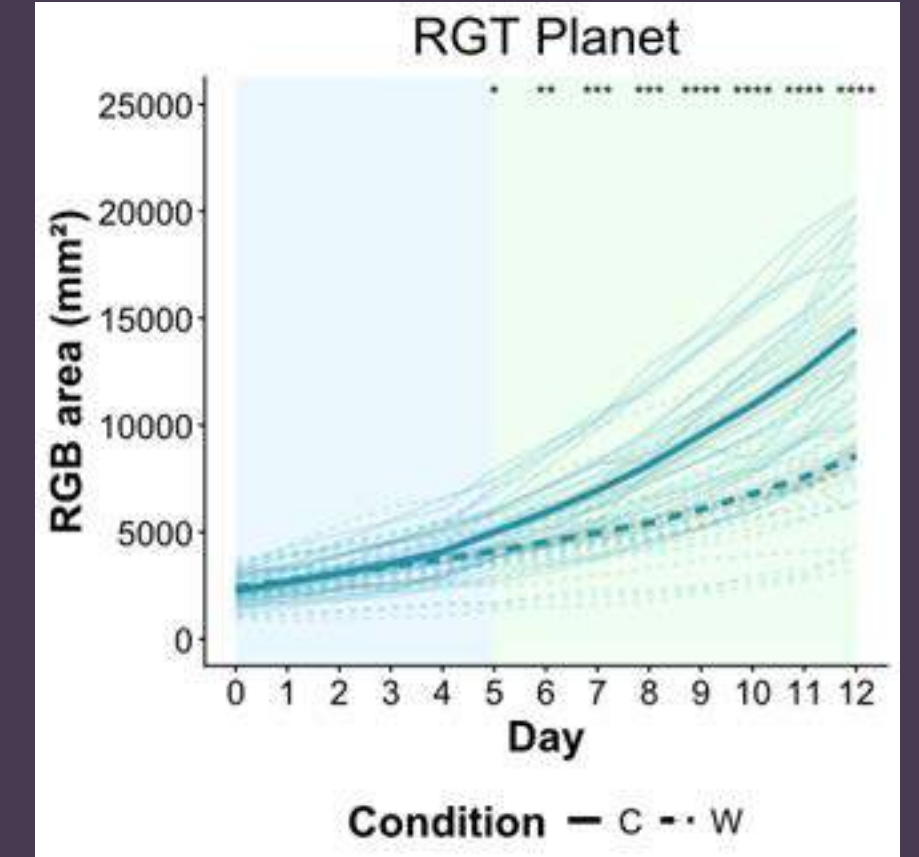
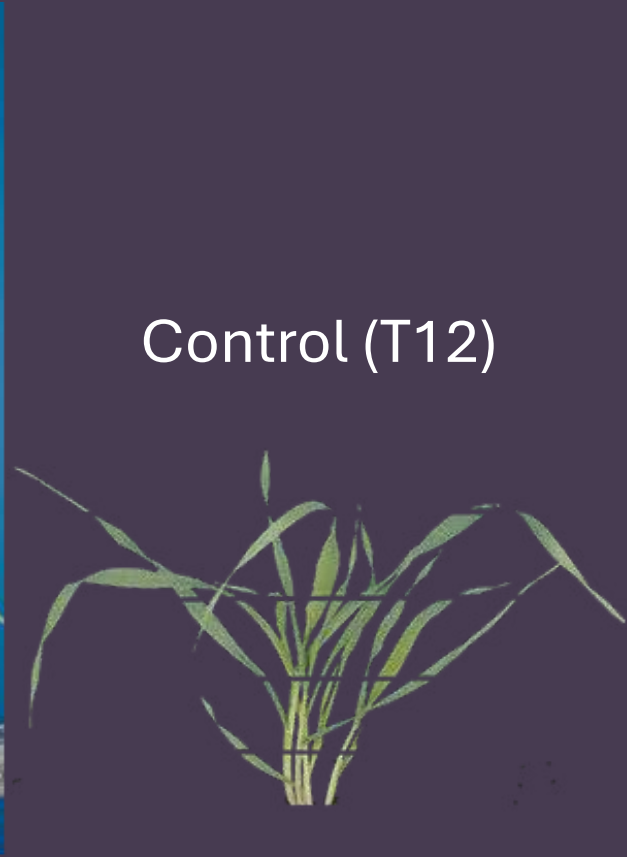
Quantum Yield (QY)



Water1

Bernád et al. (in preparation)

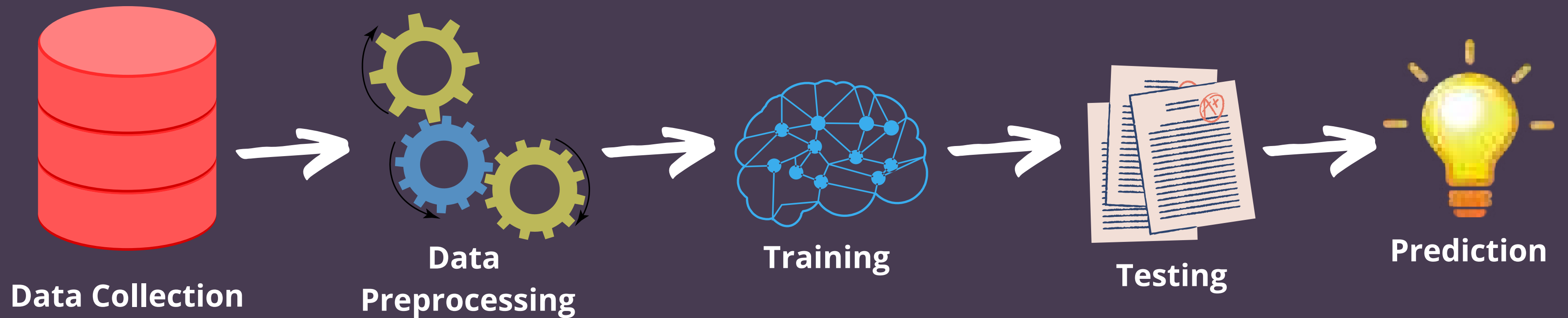
Shoot and root behave differently



The image features a complex, glowing blue circuit board pattern that fills the entire frame. The lines of the circuit are intricate and vary in thickness, creating a sense of depth and connectivity. In the center of the image, the words "BIG DATA" are prominently displayed in a bold, white, sans-serif font. The text is slightly shadowed, giving it a three-dimensional appearance as if it's floating above or attached to the circuitry. The overall aesthetic is high-tech and digital, representing the vastness and complexity of big data.

What is Machine Learning?

"Machine learning is the study of computer algorithms that allow computer programs to automatically improve through experience" – Tom Mitchell



The Machine Learning Process

Supervised Learning

- Split dataset into train/test:
 - Train algorithm using training set.
 - Test performance using test set
- E.g., Support Vector Machine, Random Forest



Dataset



Labels

Unsupervised Learning

- Identify patterns in a unlabeled dataset.
 - Most common method – clustering analysis
- E.g., K-Means, Hierarchical clustering



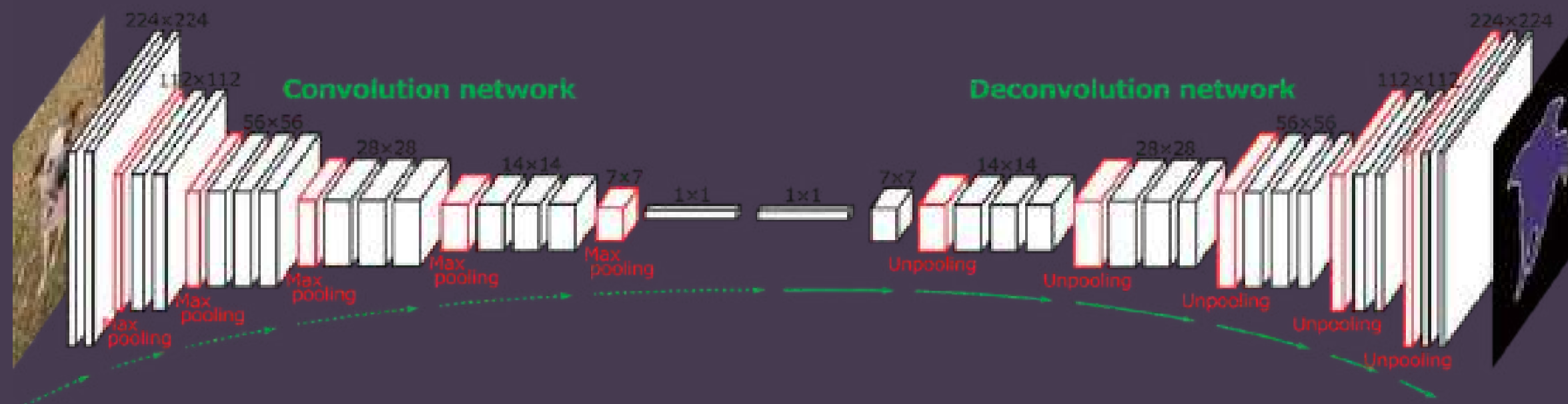
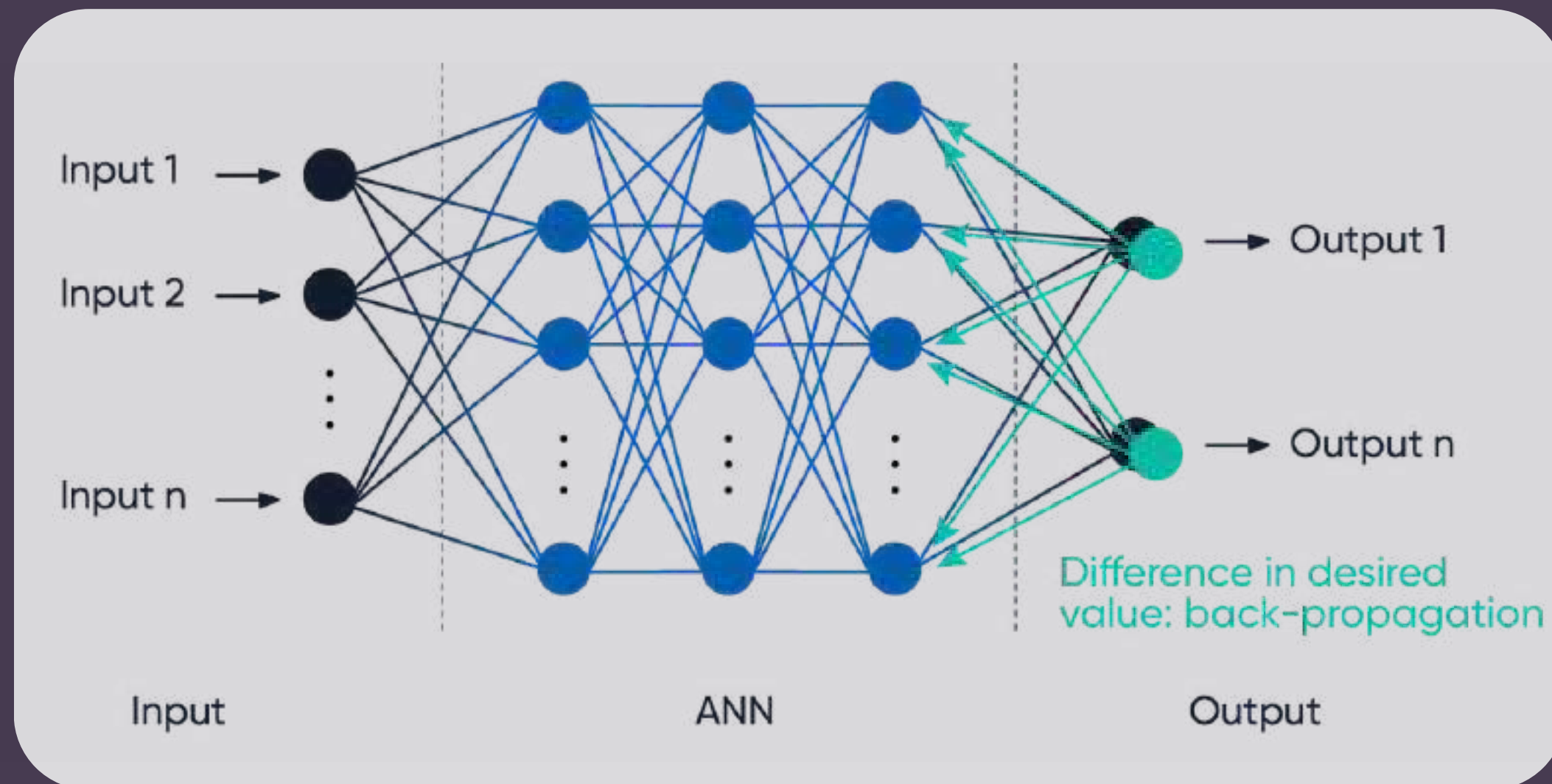
Dataset



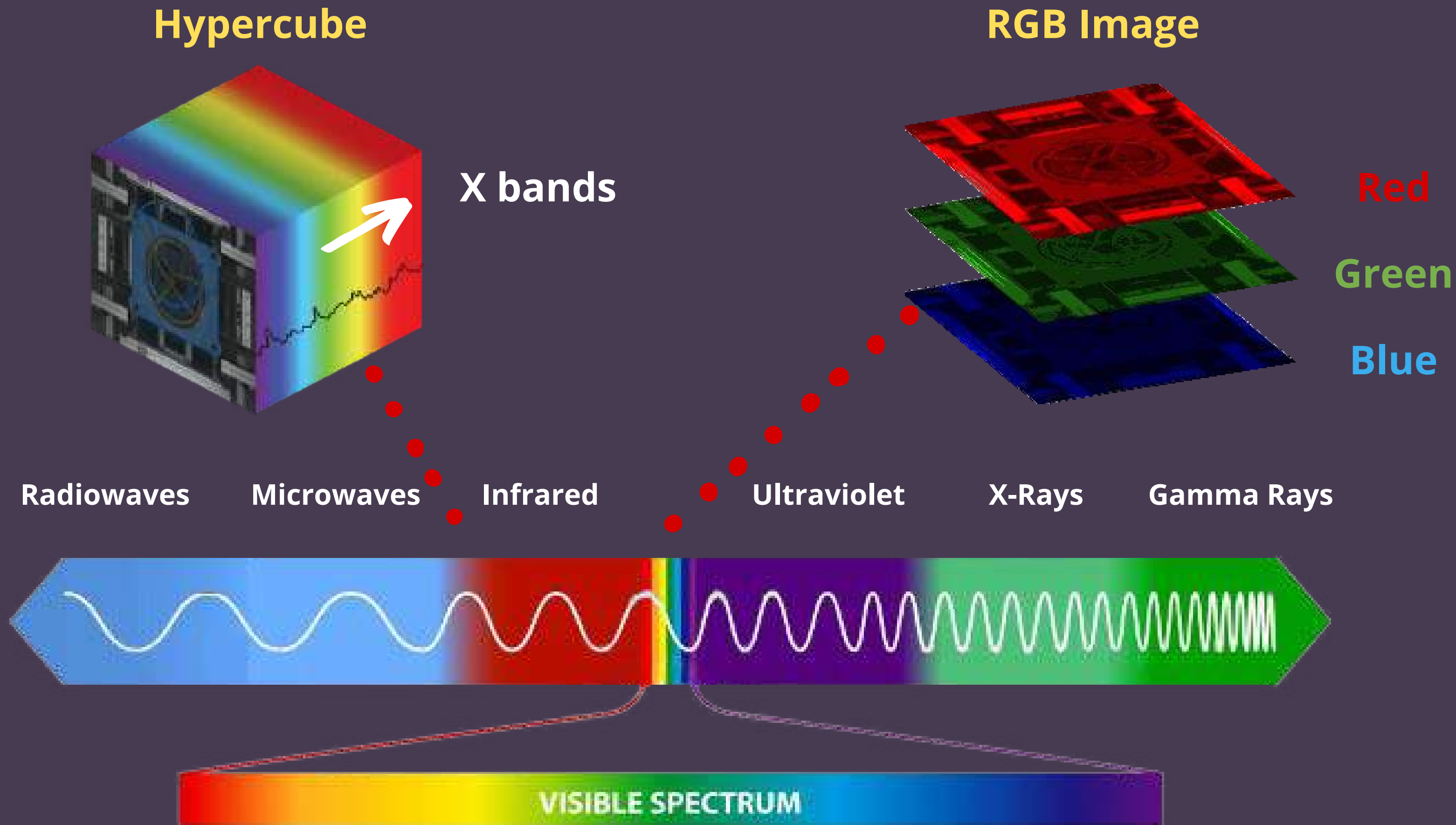
Labels

Deep Learning

- Deep learning – ML subfield
- Cascade of layers
 - Each node extract features from input data.
- Each layer - output of previous layer as input data
- Common applications:
 - Object detection
 - Image classification
 - Semantic segmentation

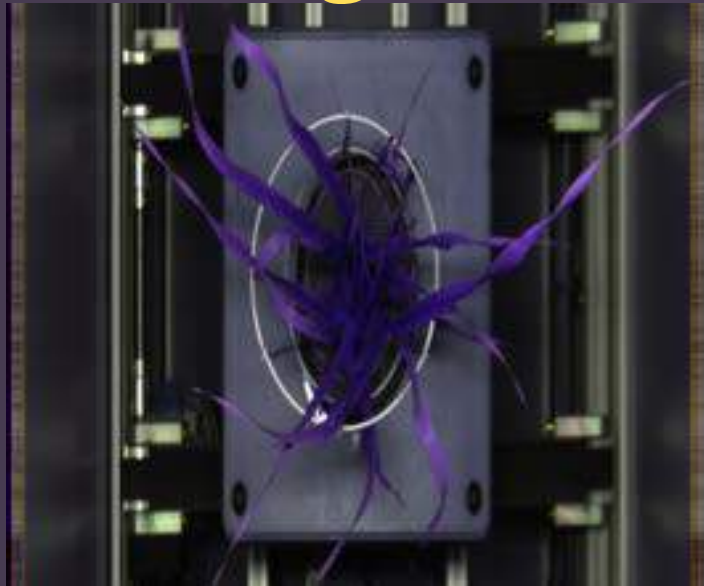


RGB vs Hyperspectral



Plant segmentation in HSI has challenges

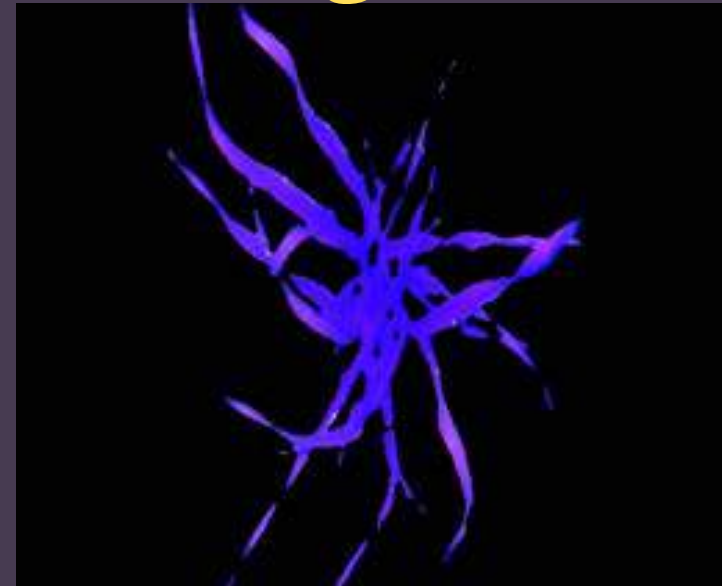
Original



Masked

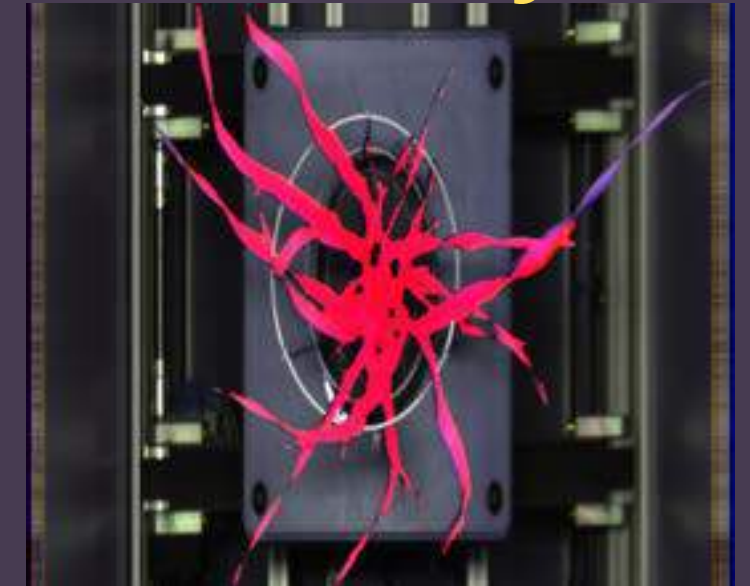


Colour Segmentation



=

Overlay



Current available methodologies

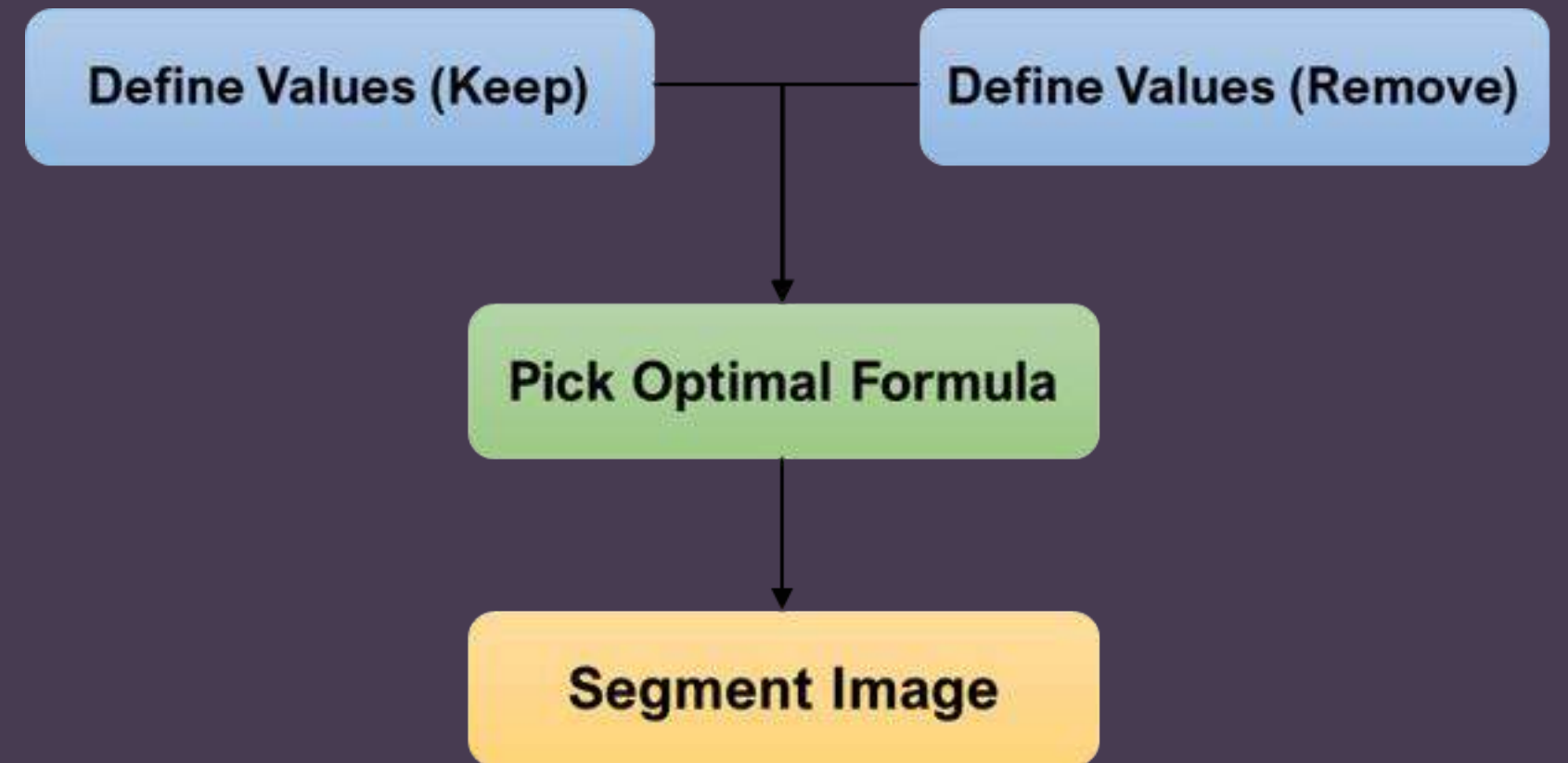
- Imaging systems use straightforward colour segmentation.
- Segmentation output needs improvement for plant biologists
- Binary mask - SWIR - problematic in alignment of all layers

Define Values (Keep)

Define Values (Remove)

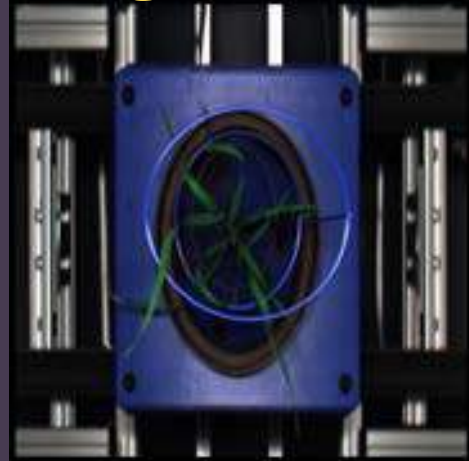
Pick Optimal Formula

Segment Image

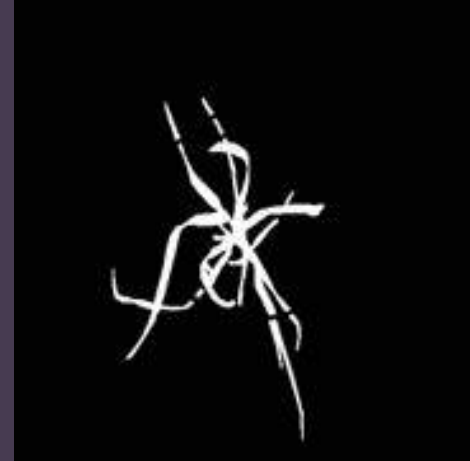


Using Deep Learning for Image Segmentation

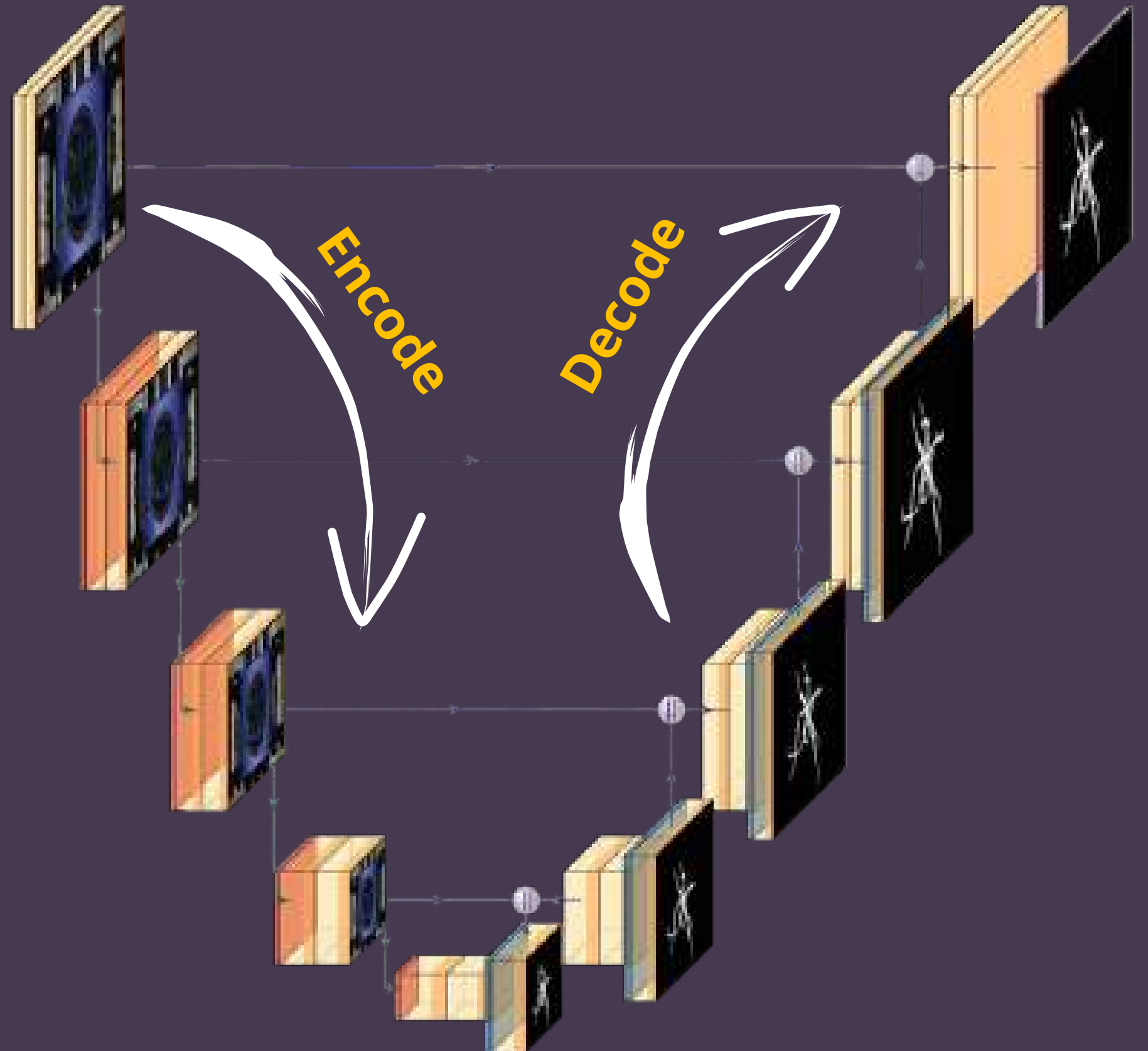
Augmented



Mask



+

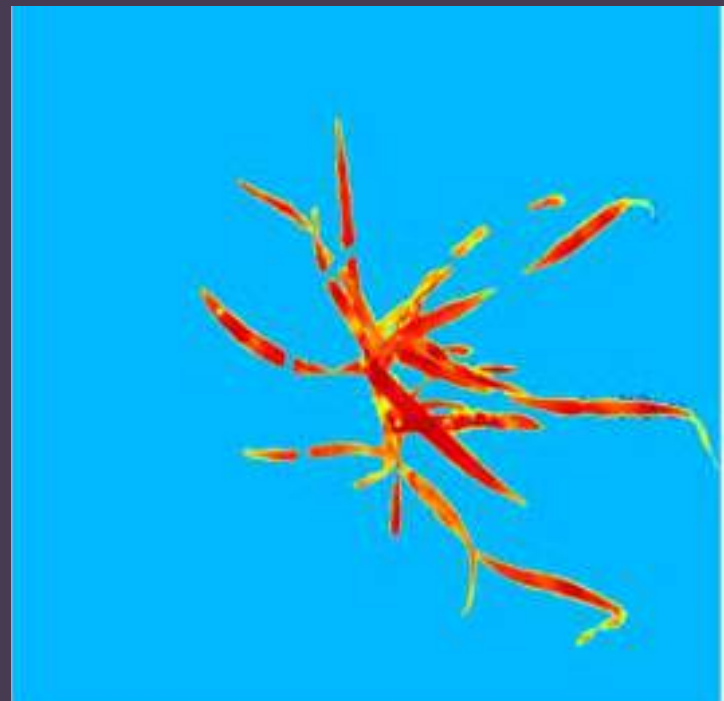
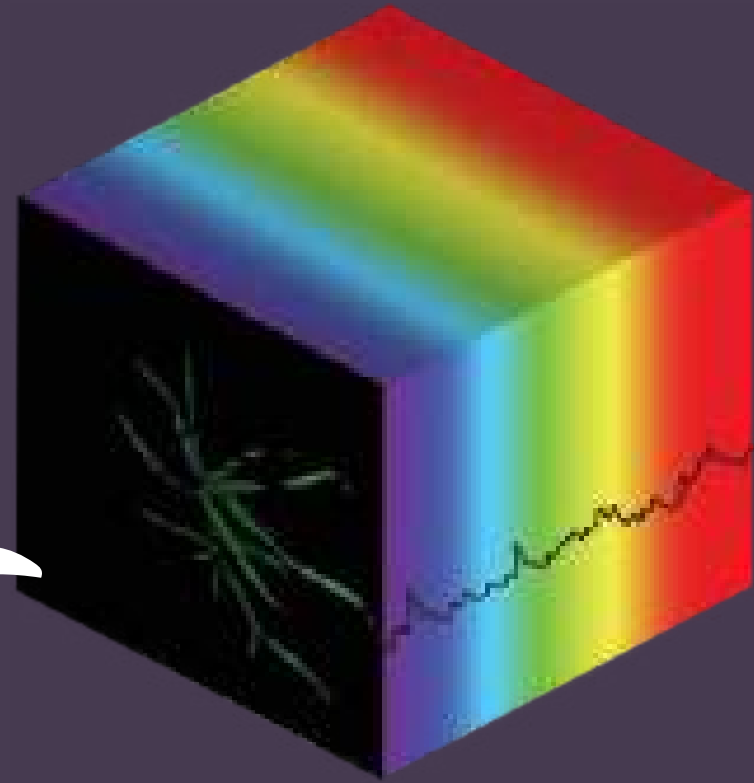


Deep Learning Segmentation

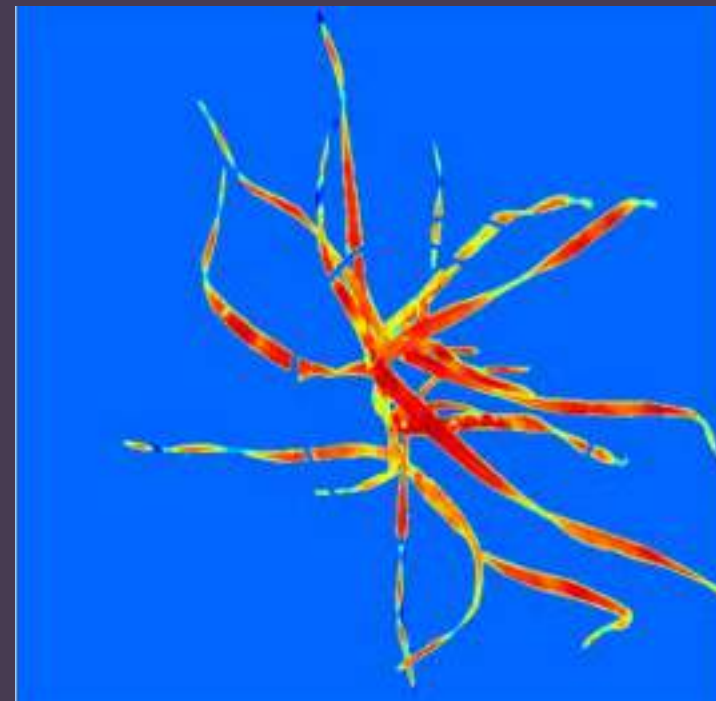
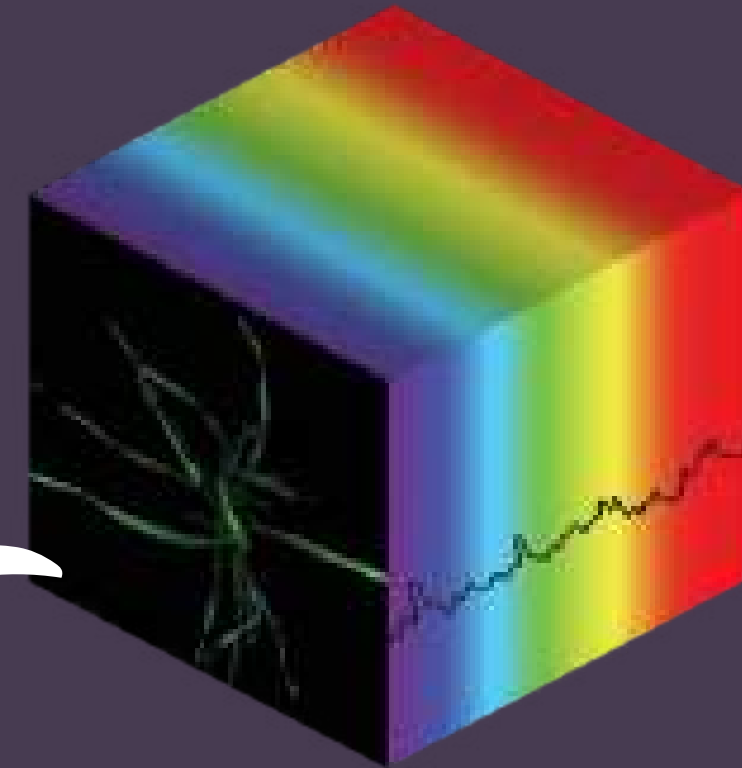
- Lightweight version of U-Net.
- Data aggressively augmented.
- A network for each datatype.

Improving image segmentation using DL

Existing methods



U-Net



Can we improve root segmentation?

Groundtruth



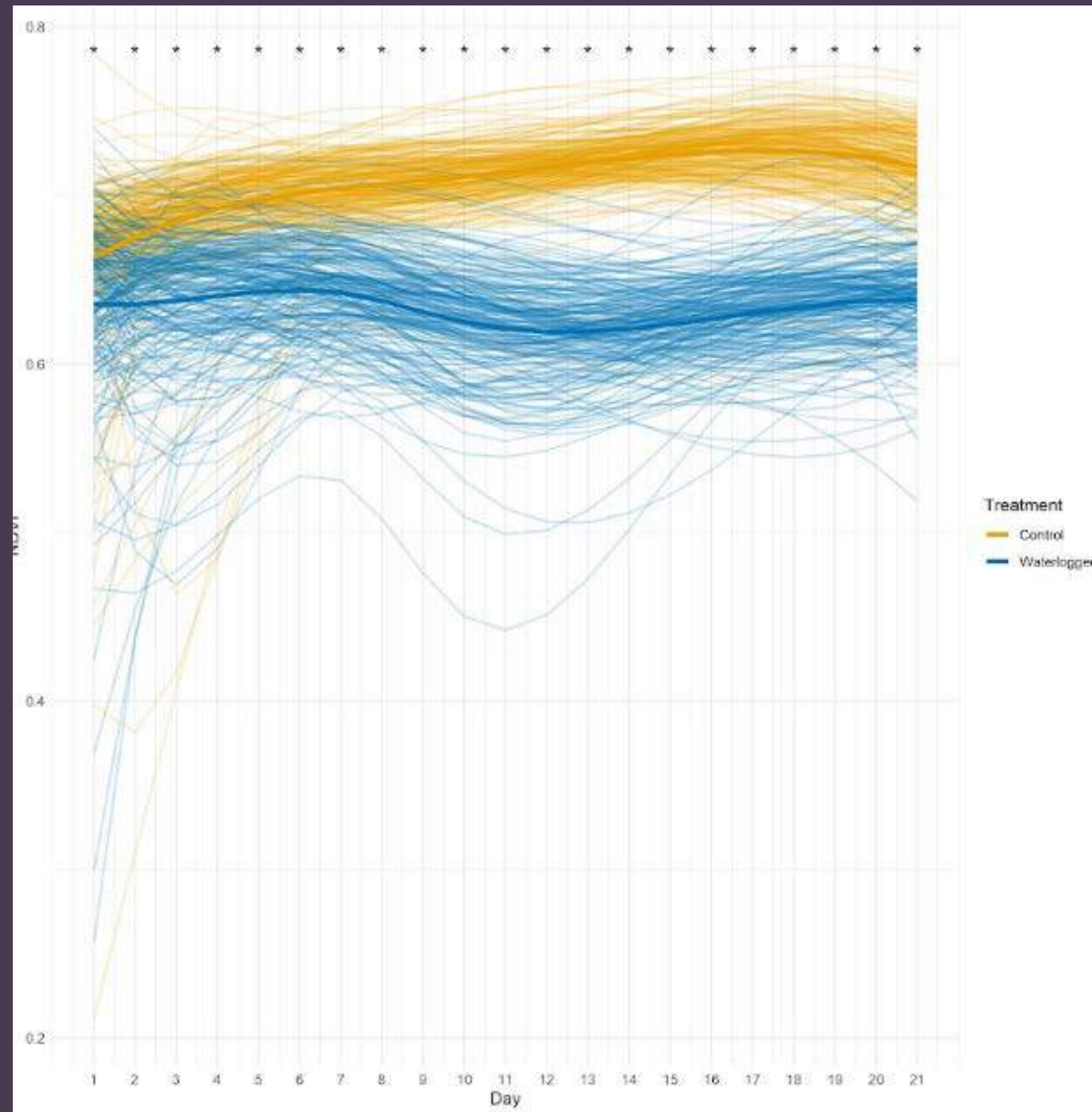
Available



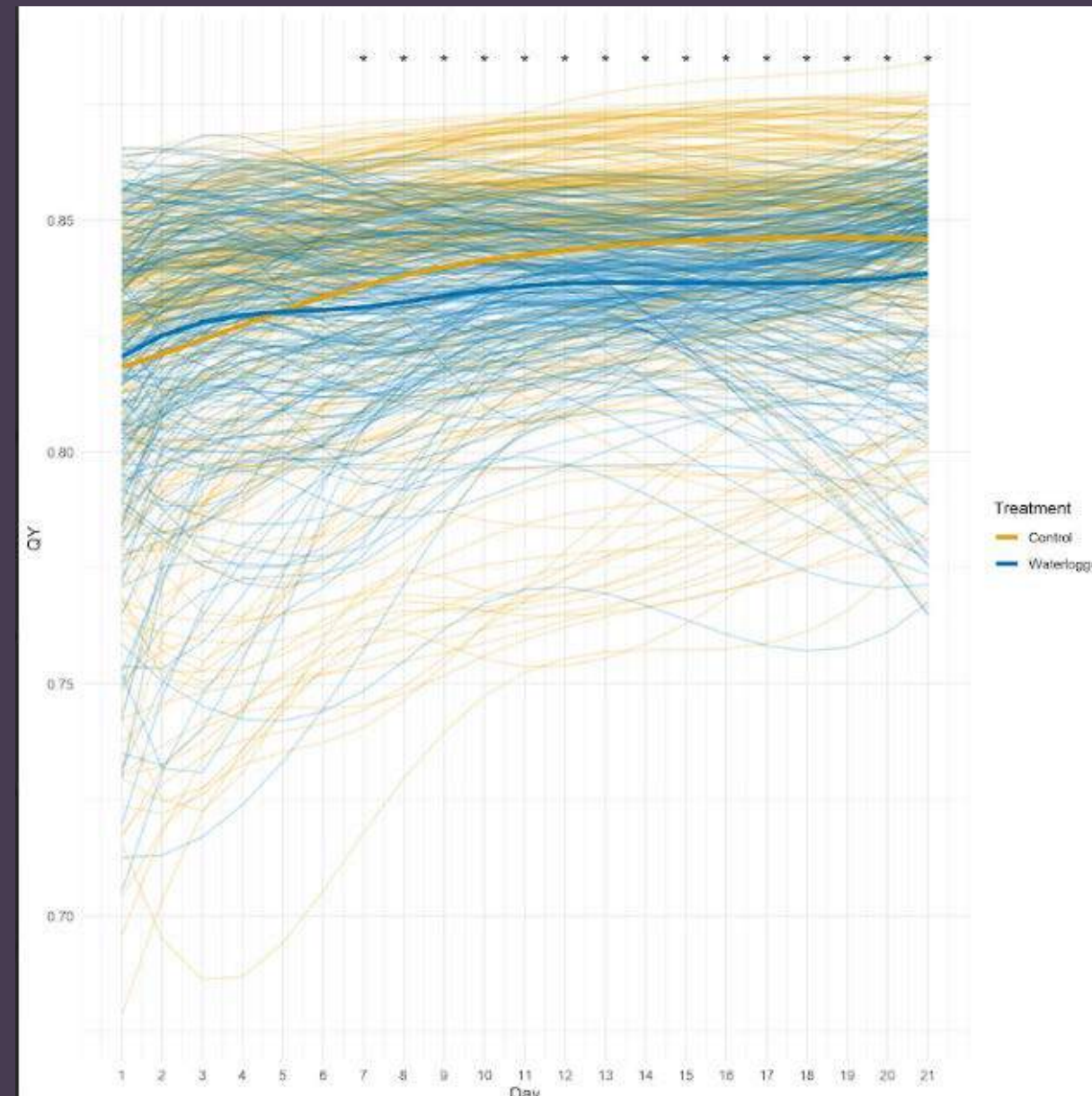
U-Net



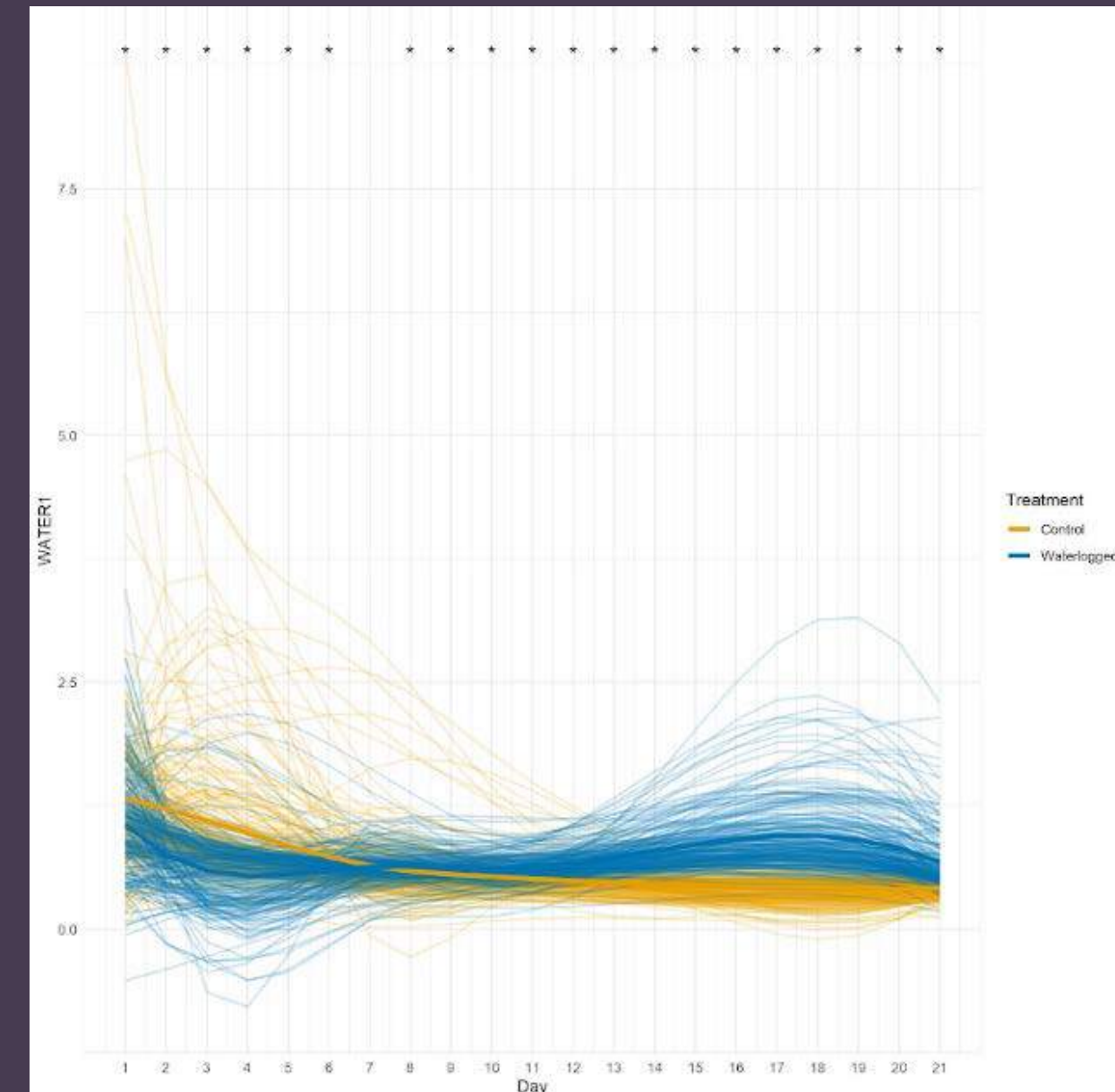
How can we interpret complex data?



**Normalized Difference
Vegetation Index (NDVI)**



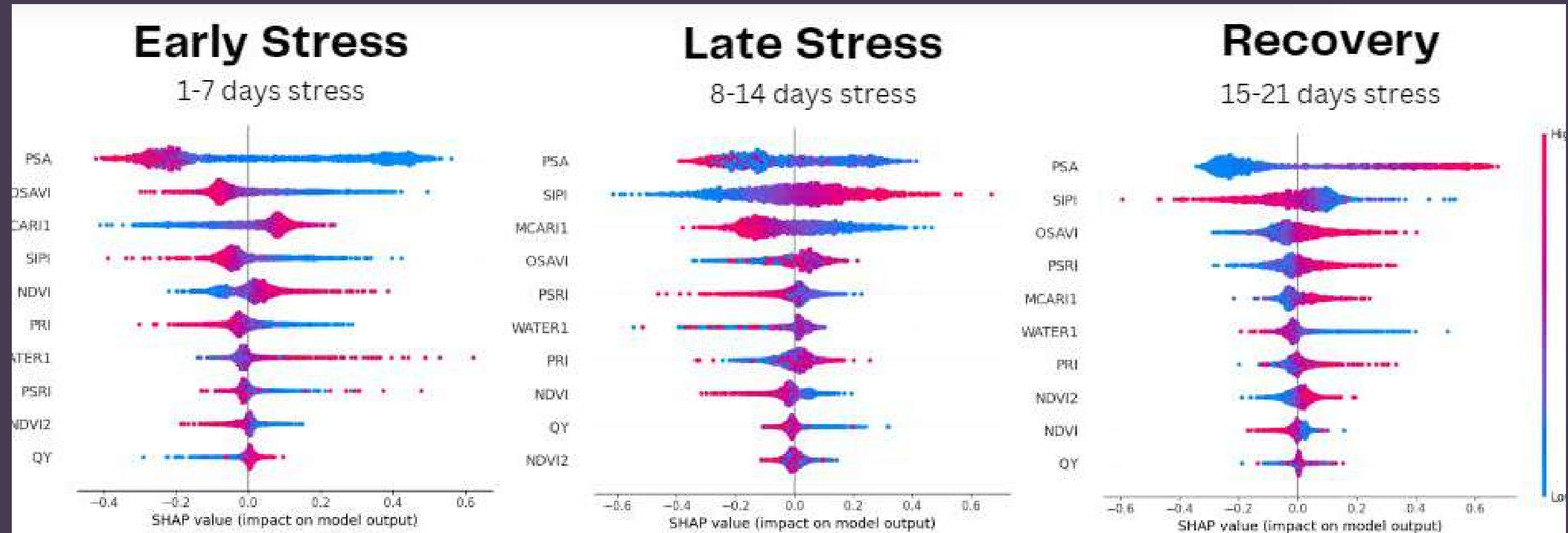
Quantum Yield (QY)



Water1

Bernád et al. (in preparation)

Can we determine main traits to predict stress?



NDVI2	PSRI	SIPI	MCARI1	OSAVI
QY	PSA	NDVI	PRI	WATER1

VEGETATION INDICES (FEATURES)
CLASSIFIED BY AI

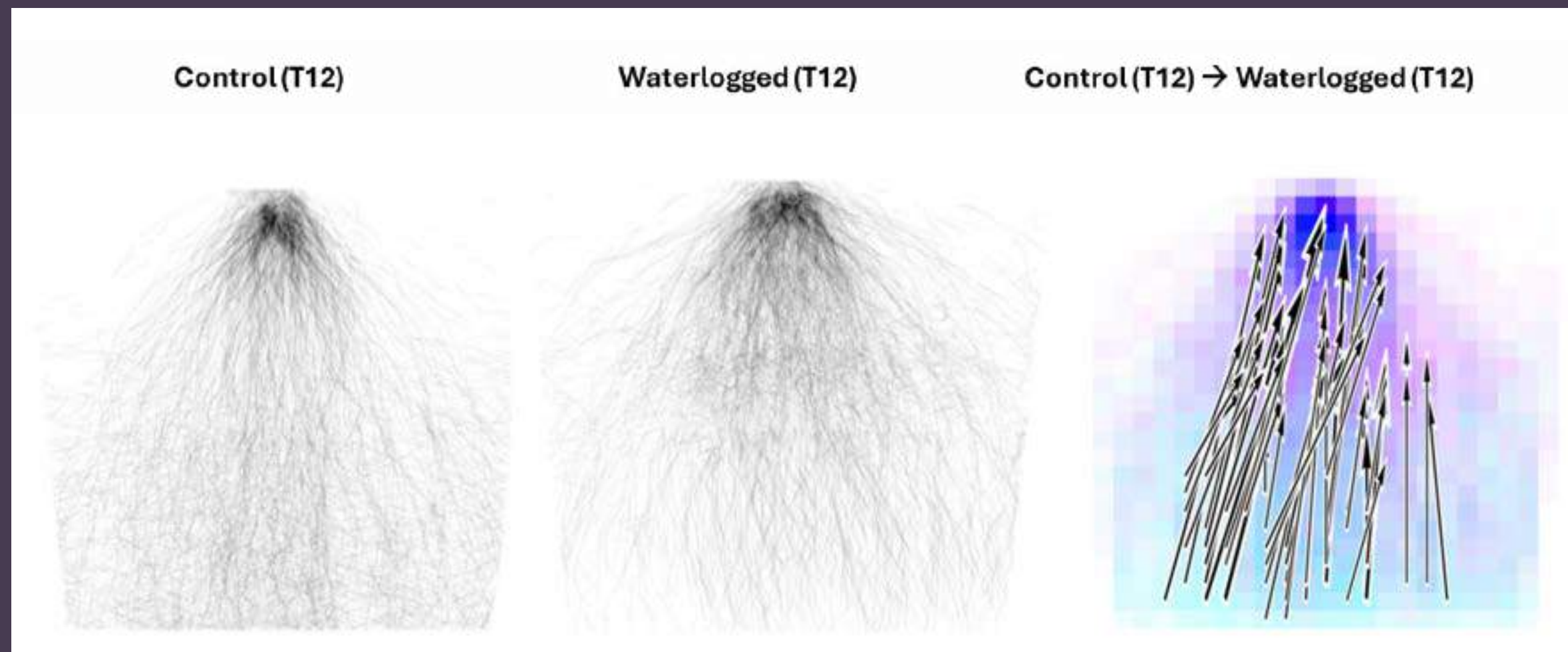
Bernád et al. (in preparation)

Quantum Yield (QY); Project Shoot Area (PSA); Normalized Difference Vegetation Index (NDVI); Structure Insensitive Pigment Index (SIPI); Plant Senescence Reflectance Index (PSRI); Photochemical Reflectance Index (PRI); Improved Modified Chlorophyll Absorption Ratio Index (MCARI); Optimized Soil-Adjusted Vegetation Index (OSAVI)

Can we use ML to understand root dynamic responses to stress?

STARSEED (Peeples et al., 2023) to evaluate treatment effect on root system spatial arrangement :

- Evidence of spatial shift towards superficial layers
- Reduction in lateral spread at depth



What are the genetic regions underlying longitudinal traits?

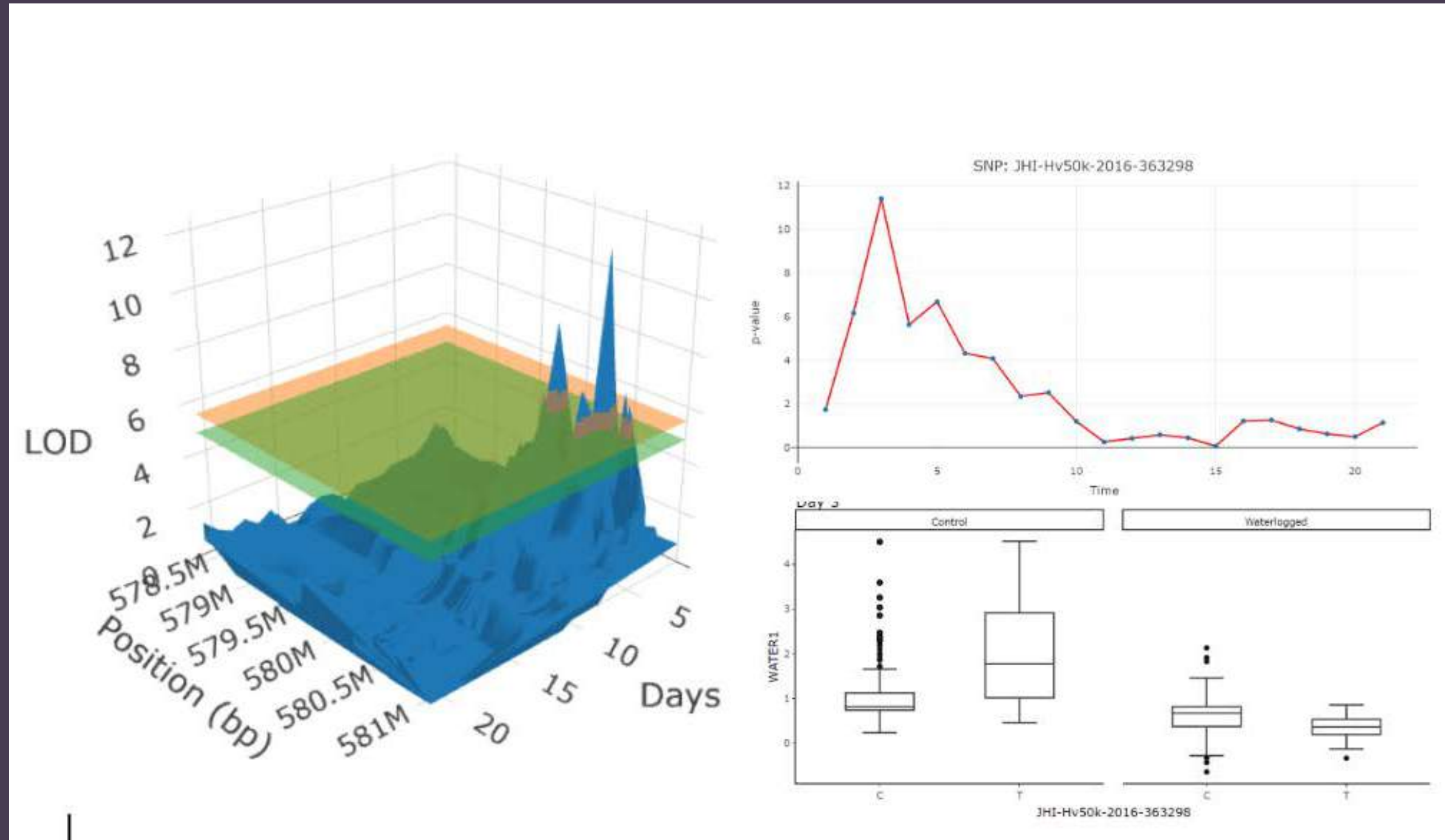
GWAS Interaction Model:

$$Y = G_SNP + Q + Treatment + G_SNP \times Treatment + (1 \mid Taxa) + K$$

What are the genetic regions underlying longitudinal traits?

3D-QTLVis

Water1 throughout time



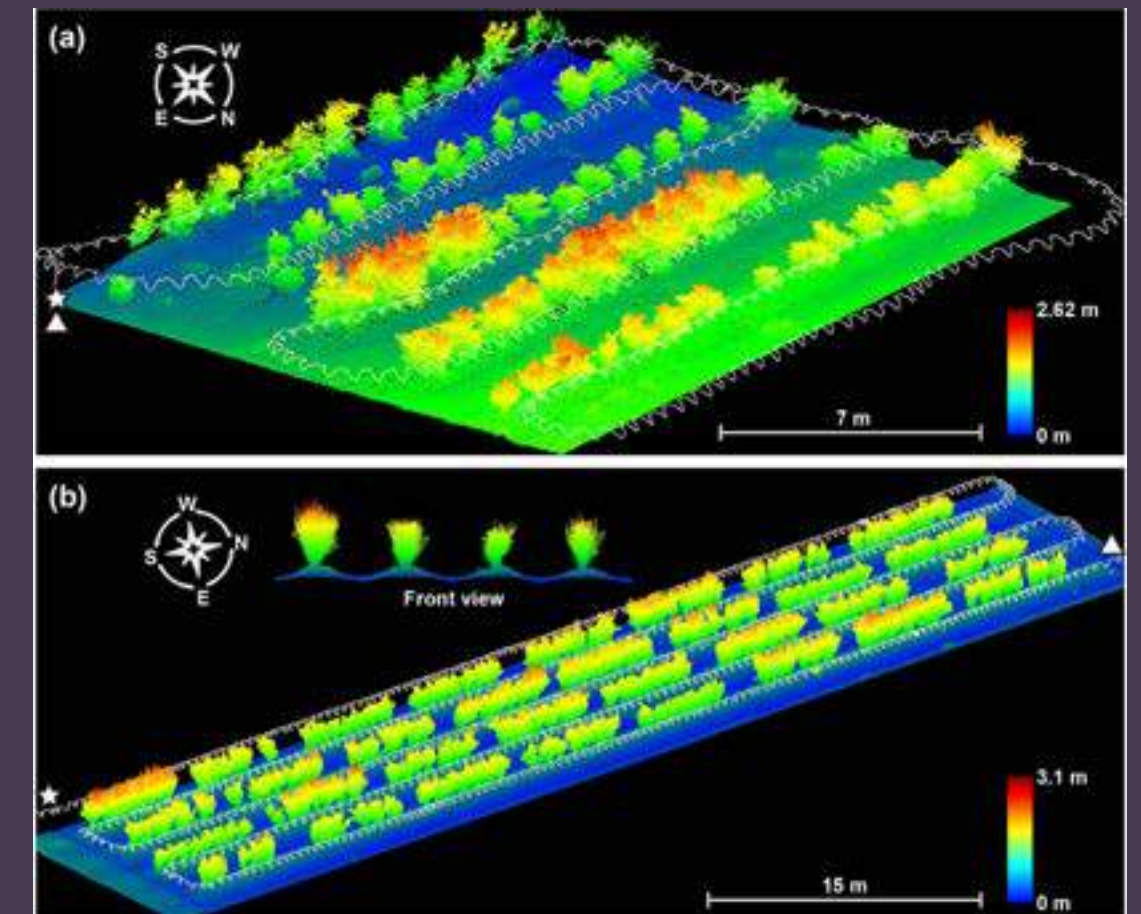
Bernád et al. (in preparation)



Can we further improve imaging in the field?

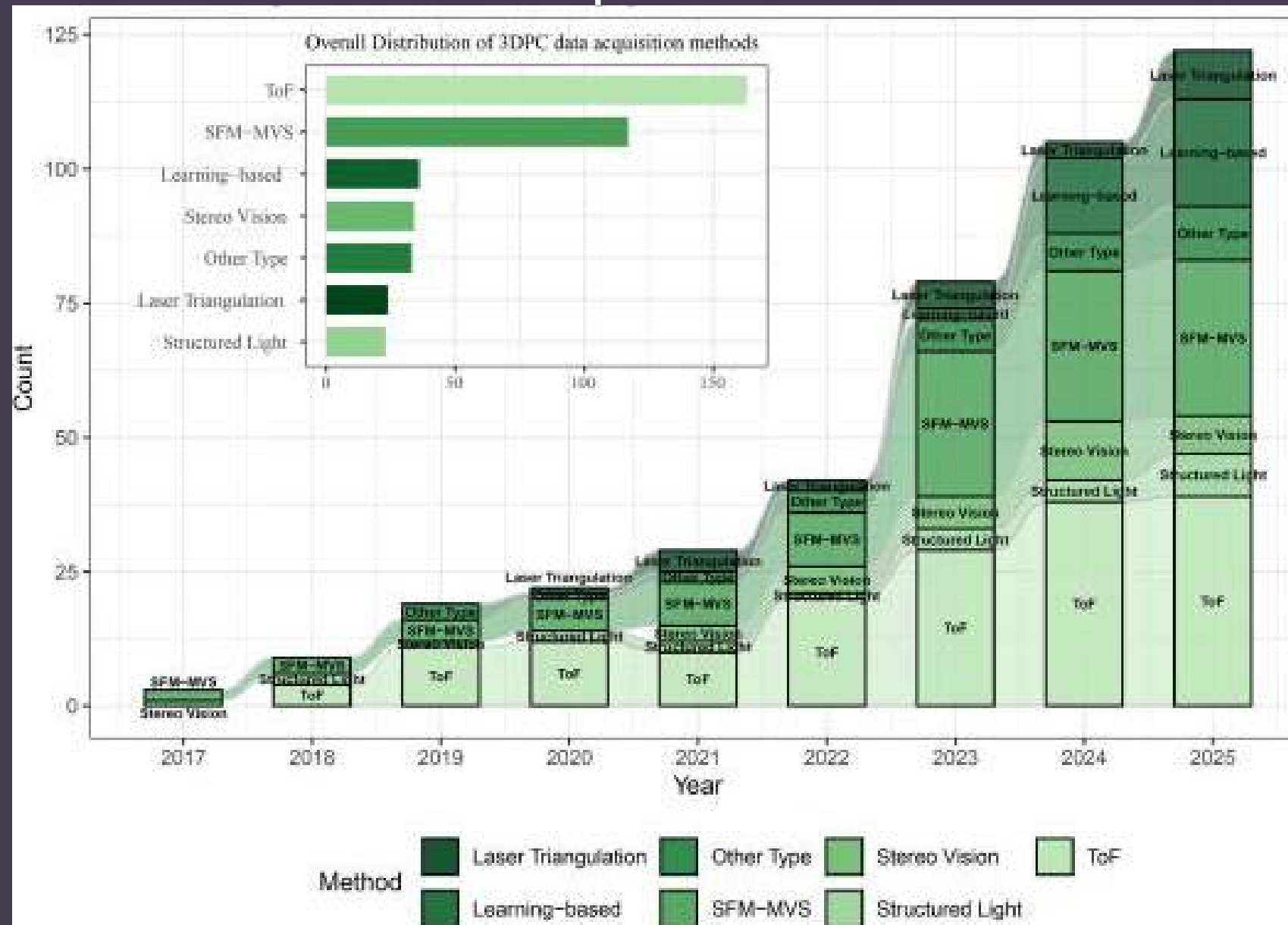
Unlocking 3D Insights: The Power of Point Clouds in Agriculture

- A Three-Dimensional Point cloud (3DPC) is an unstructured representation of data
- Contains more phenotypic information than 2D images (e.g., plant height, canopy structure, etc...)
- Easy to construct 3DPC from RGB or MS images (a combined Structure-from-Motion and Multi-View Stereo photogrammetry technique) □ **Challenge: Big Data processing**

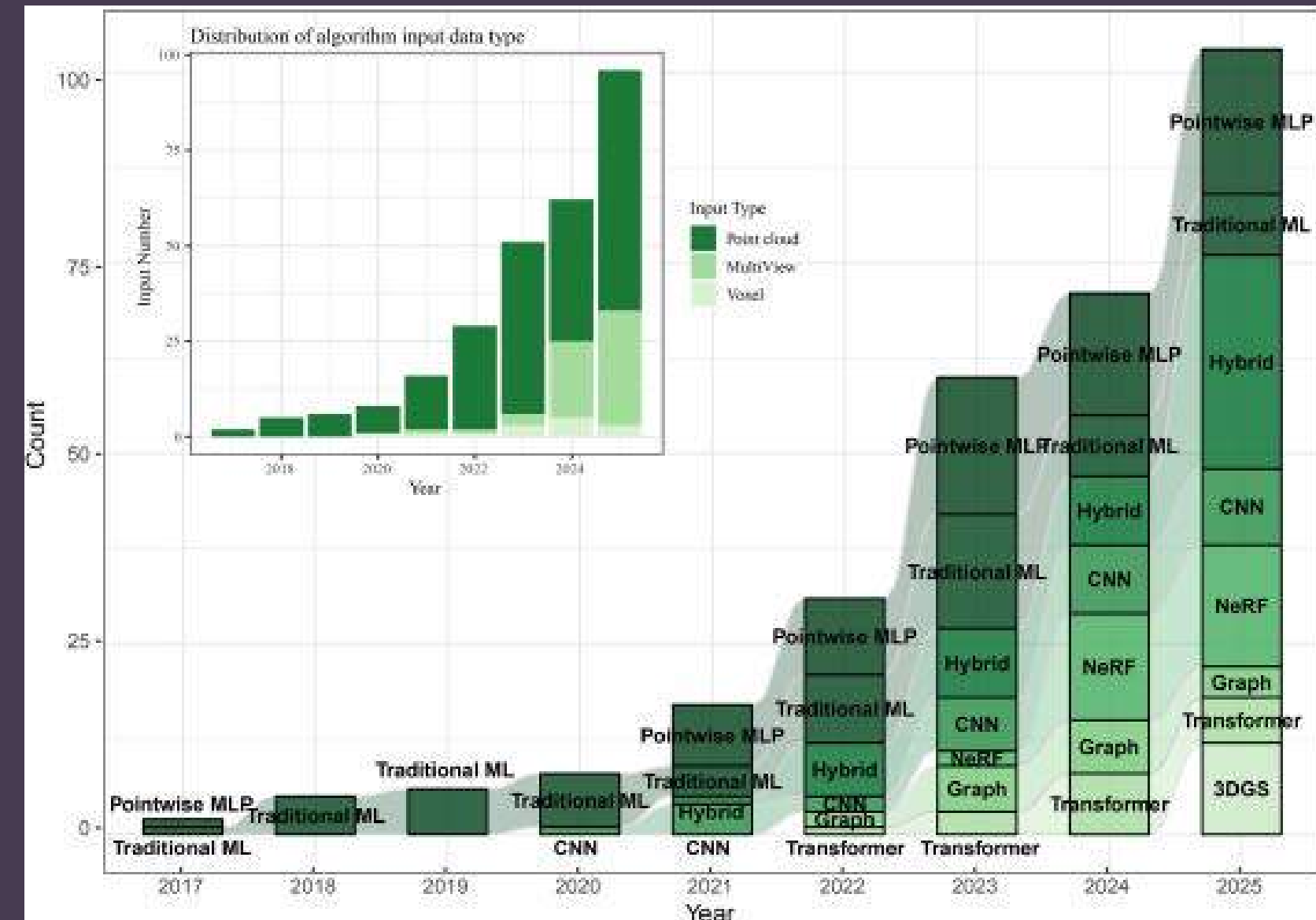


What do we know about 3DPC?

3DPC Data acquisition methods



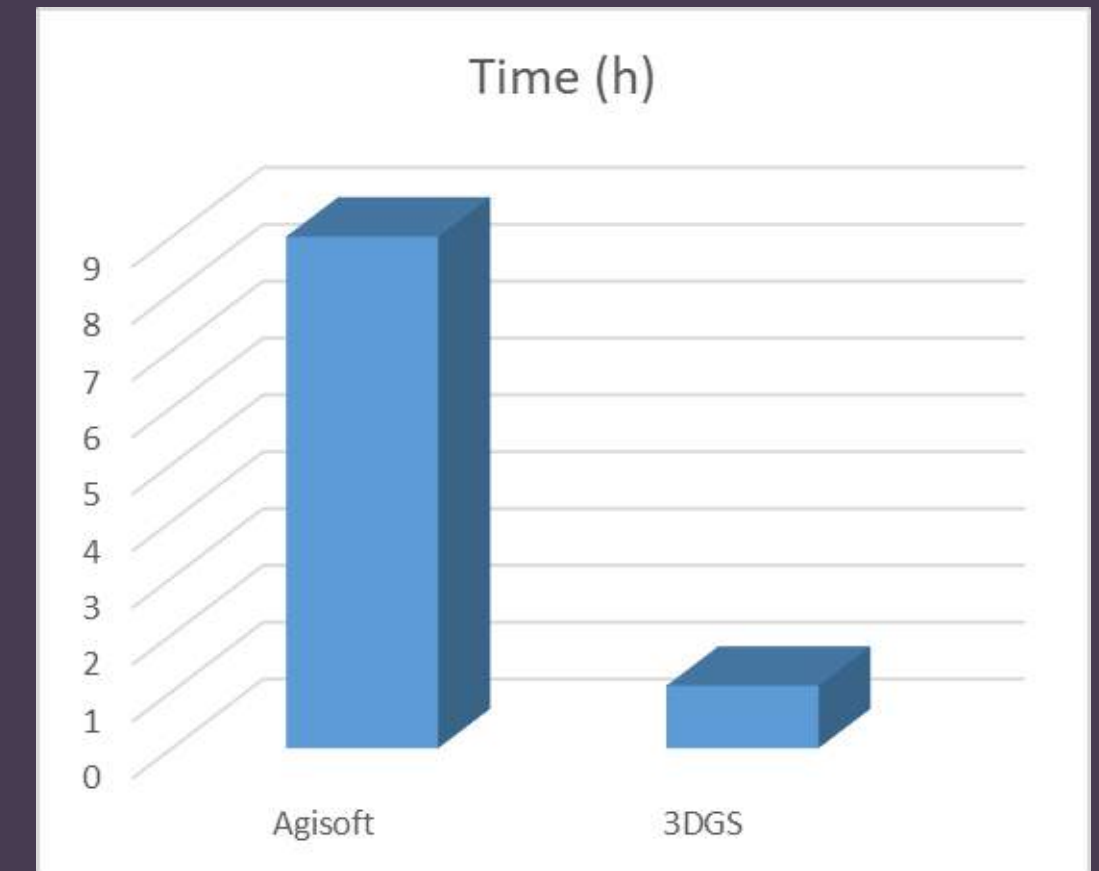
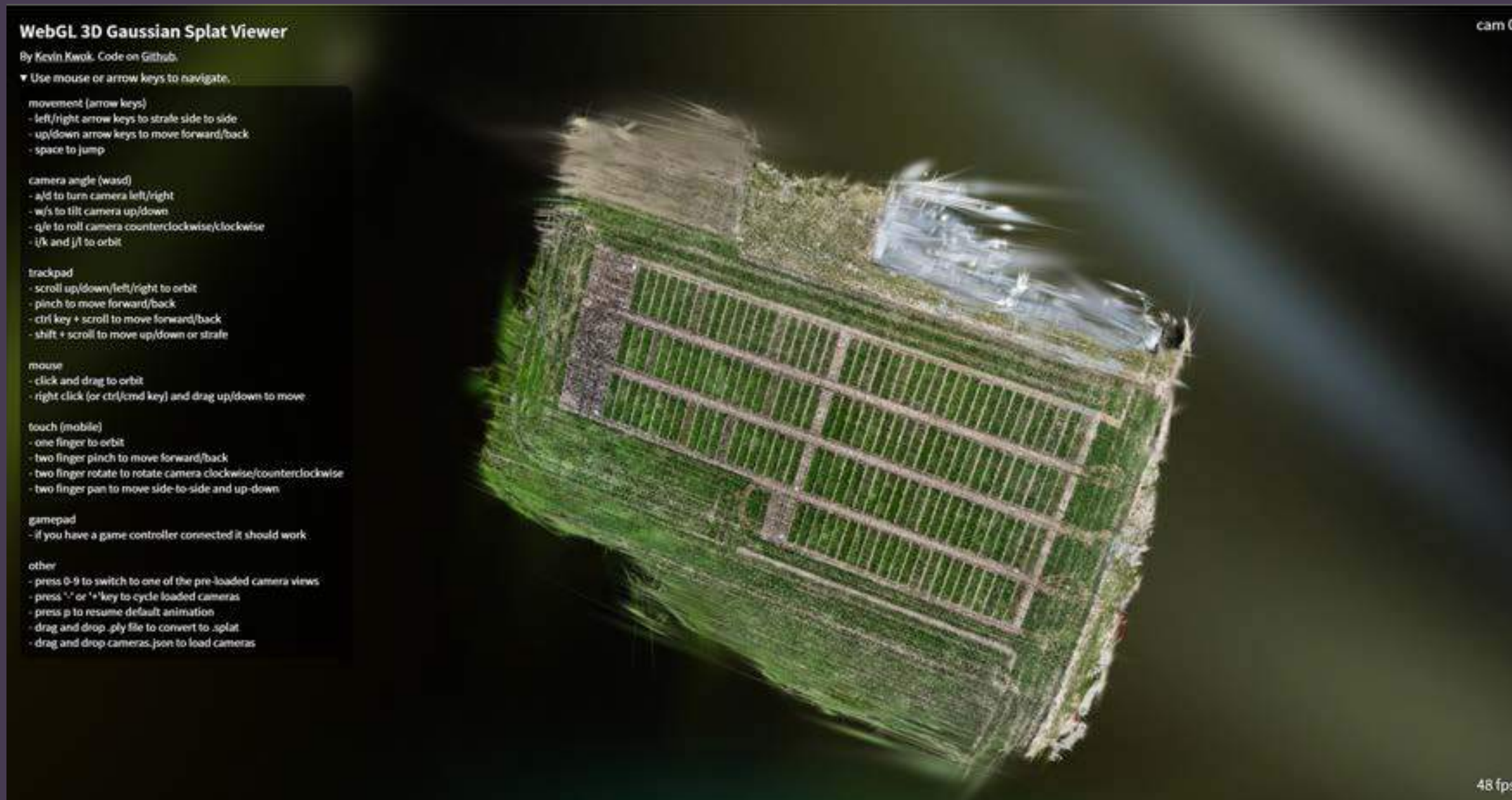
AI Algorithms used in 3DPC analysis



- DL is now dominating, 3DGS and NeRF have broken out recently.
- 3DPC Data acquisition mainly used ToF sensors and SfM-MVS pipelines.



Can we develop a 3DPC from RGB and multispectral images?



Yes, we can! Using a modified 3D Gaussian Splat- 3DGS:

- Established a rapid 3DPC reconstruction method
- Developed an automatic extraction workflow for orthophoto and 3DPC indices

A person in silhouette stands in a field of tall green grass under a cloudy sky. The person is positioned in the upper center of the frame, looking towards the right. The grass is dense and reaches up to the person's waist. In the background, there are some trees and a fence line.

Ongoing research & future perspectives

- 3D-QTLVis manuscript in preparation
- Exploring HSI segmentation in UAVs
- Root data analysis ongoing
- 3DPC data analysis ongoing

Acknowledgments



Dr. Jason Walsh
Dr. Villó Bernád
Victoria Poulain
Jianye Chang
Dr. Patrick Langan
Joey Henchy
Tim Dempsey

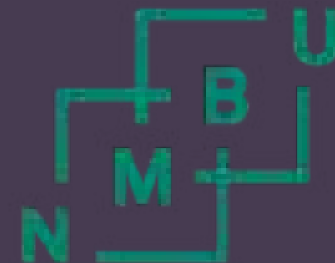


Prof. Laurent Gutierrez
Dr. Emilie Jacob
Gaëlle Mongelard
Hervé Demailly
Stéphanie Guénin



Prof Robbie Waugh
Dr. Joanne Russel
Dr. Kelly Houston
Dr. Luke Ramsay

Prof. Eleni Mangina
Dr. MortazaKhodaeiaminjan
Prof. Kevin McDonnel
Dr. Gary Gillespie
Dr. Mary Harty
Eugene Brennan
Eddie Jordan
Gordon Kavanagh
Eugene Sherry



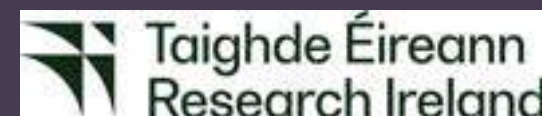
Norwegian
University of
Life Sciences

Morten Lillemo
Sahameh Shafiee
Muhammad Fahad Ijaz



Fergal McCarthy

**Funding: 18/FRL/6197 and
2022Agrobiodiversity154**



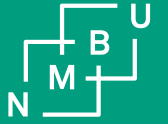


**THANK
YOU**

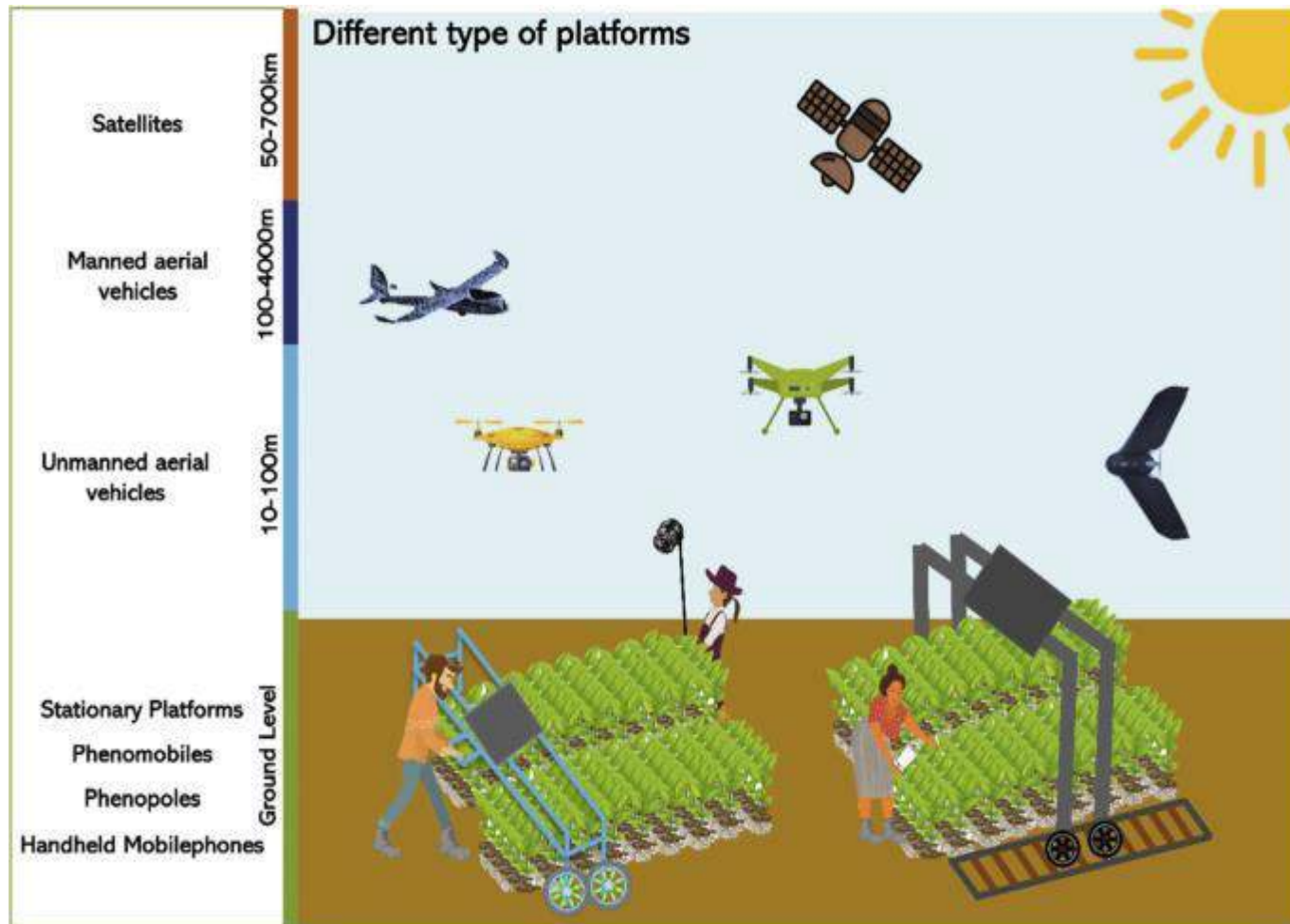
sonia.negrao@ucd.ie

Principles of handling UAV data

Sahameh Shafiee



Field Phenotyping Platforms



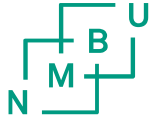
Why Drones??



[Field Based High Throughput Phenotyping Rapidly Identifies Genomic - gn.racesociety.com](http://gn.racesociety.com)

First Drone in 1905

The Drones that You have to Feed — Pigeons



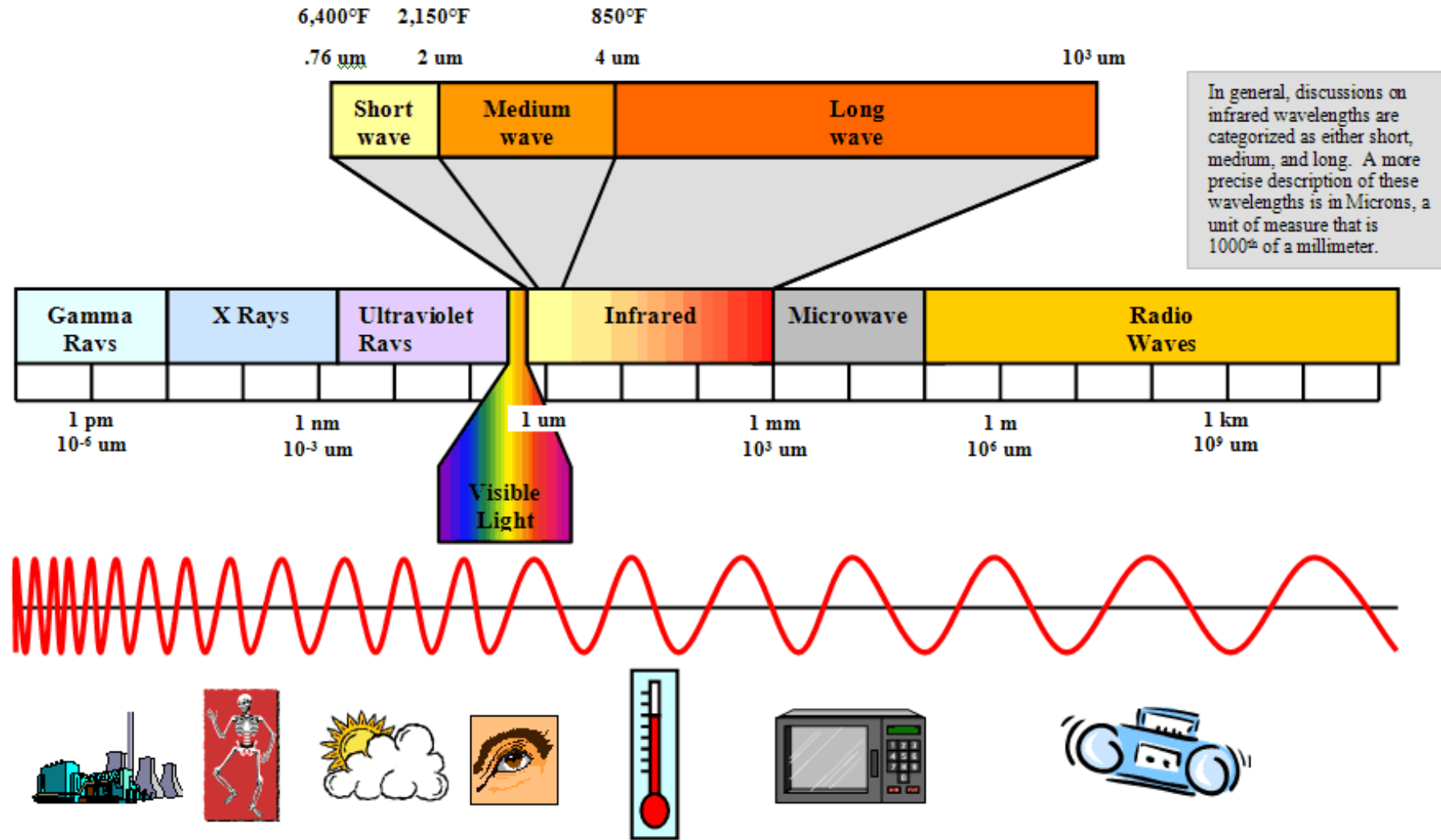
[Before there were Drones, there were ... Pigeons with Cameras? | by Nathan Allen | Frame of Reference | Medium](#)

UAV Applications in Agricultural

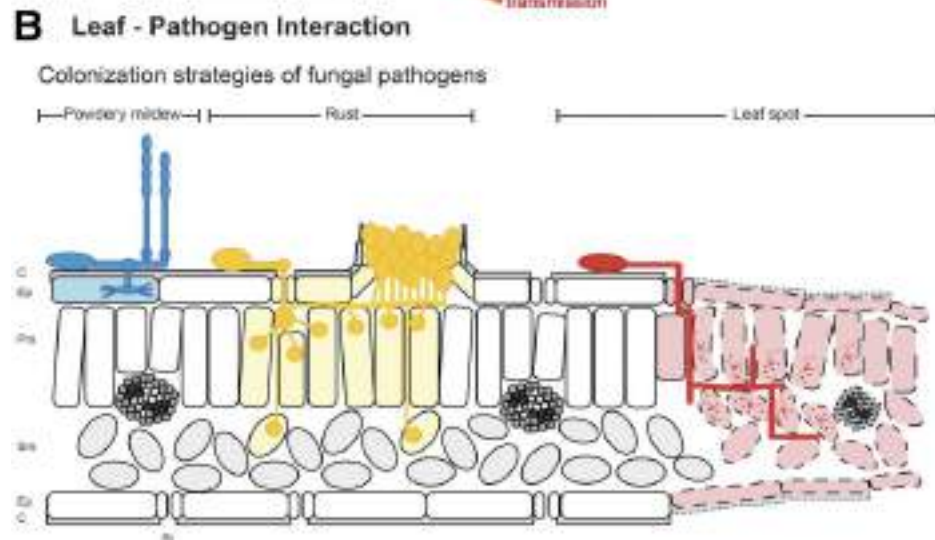
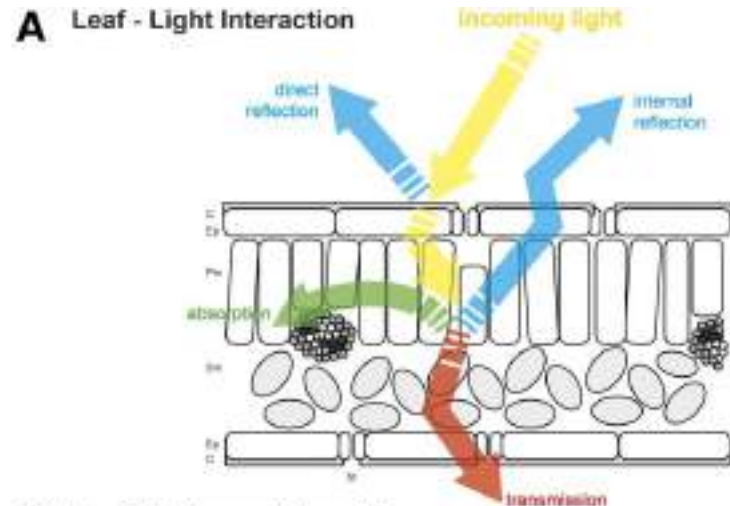
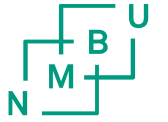
- Plant Phenotyping
- Planting seeds
- Field Monitoring
- Disease Control
- Pest and herbs control
- Crop damage assessment (lodging)
- Livestock management



Light Spectrum and Sensors

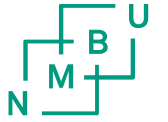


Light Interaction



NDVI and Plant Health

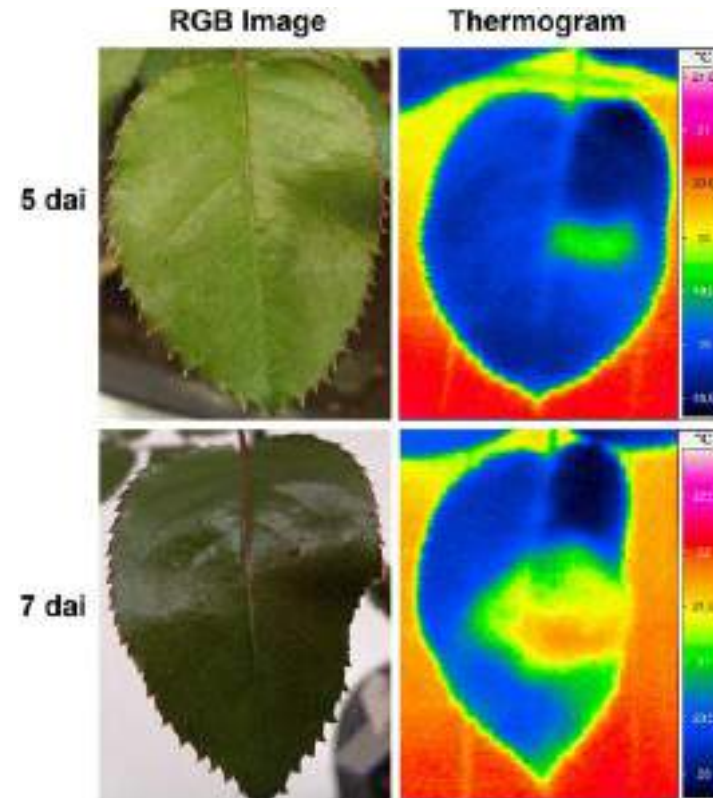
Imaging Sensors



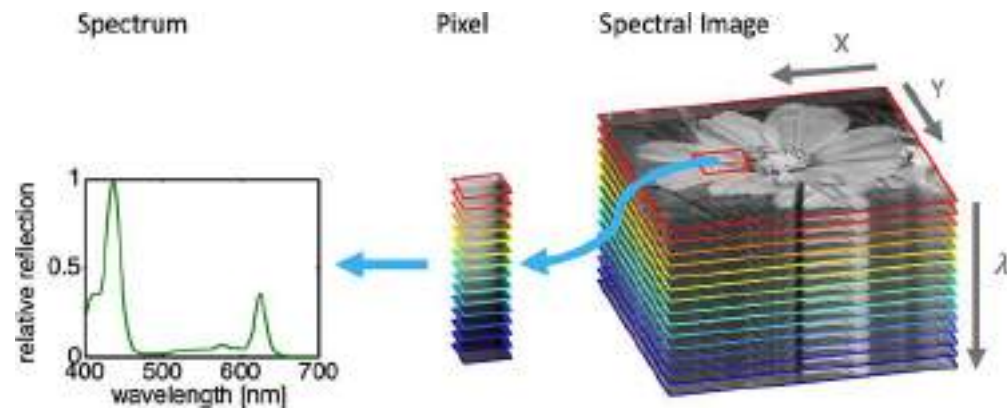
Multispectral



Thermal



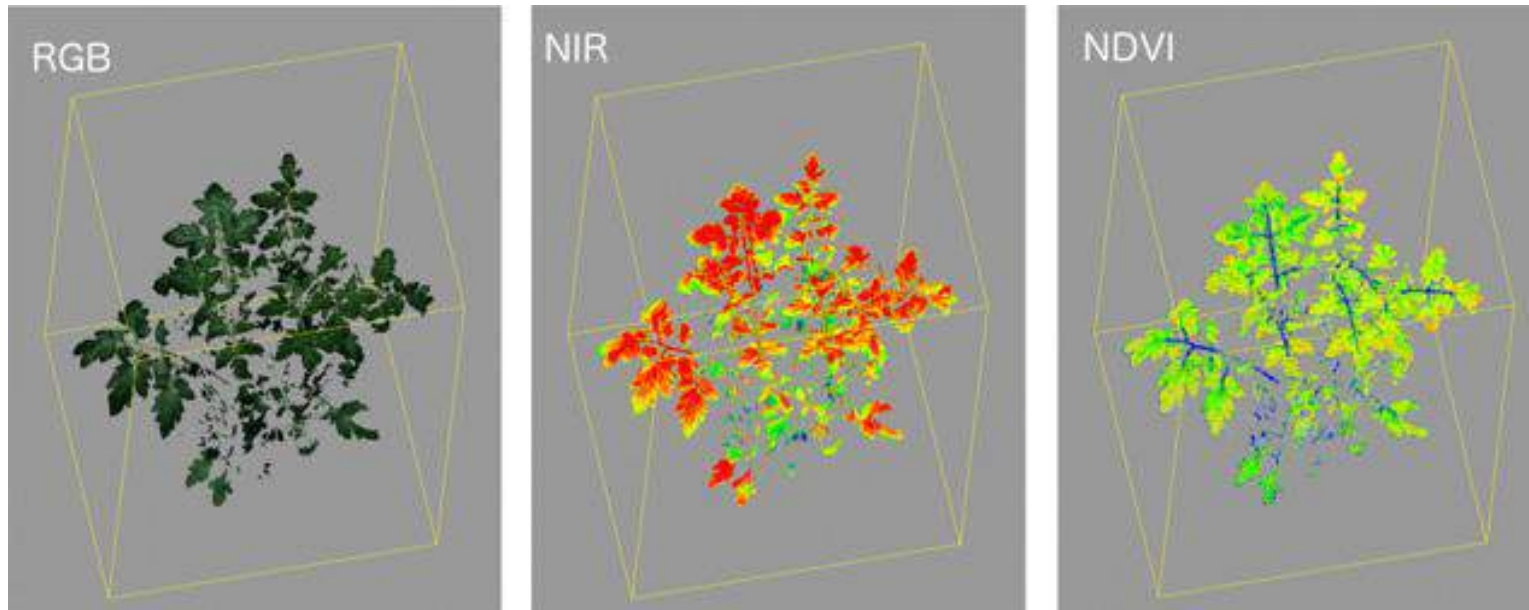
Hyperspectral



3D Imaging

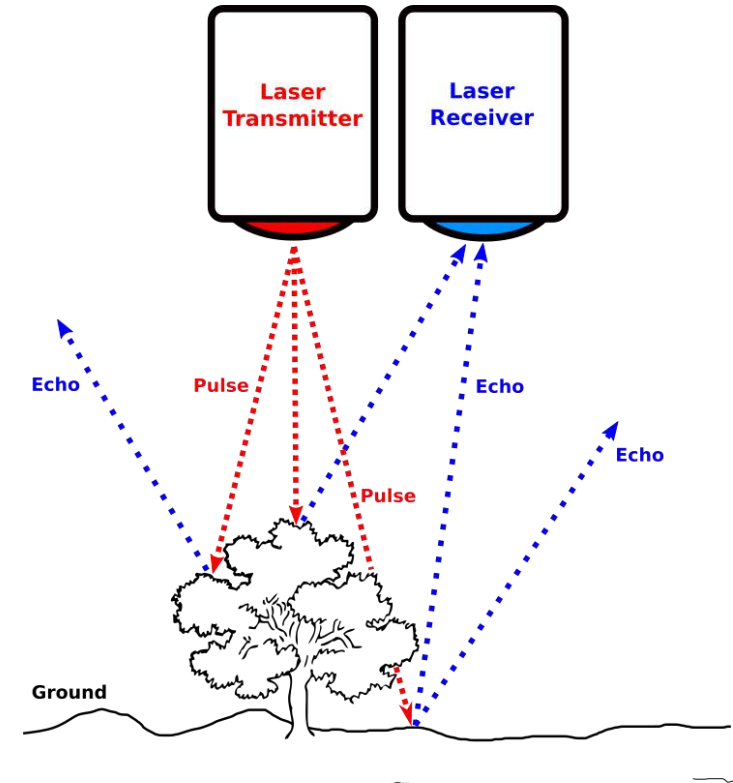
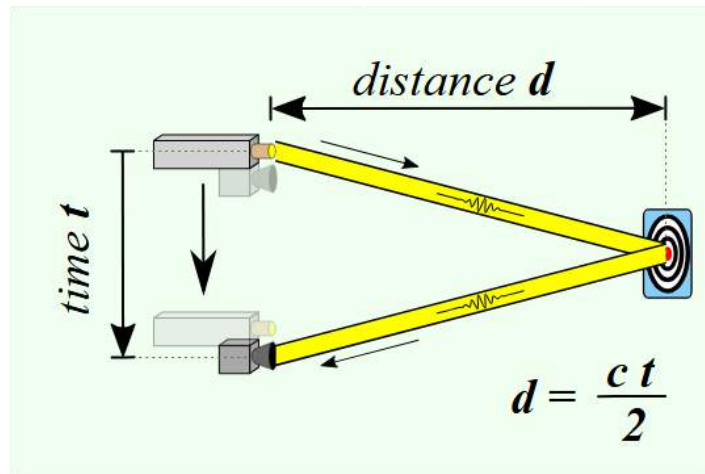


- Many different sensors and techniques are used for the measurement of the distance of objects to the camera, acquisition, and formation of 3D images.
- Optical applications: Passive, such as stereo imaging and structure from motion, or active, such as structured light and time of flight.



Time of flight (ToF) and Light Detection and Ranging (LiDAR)

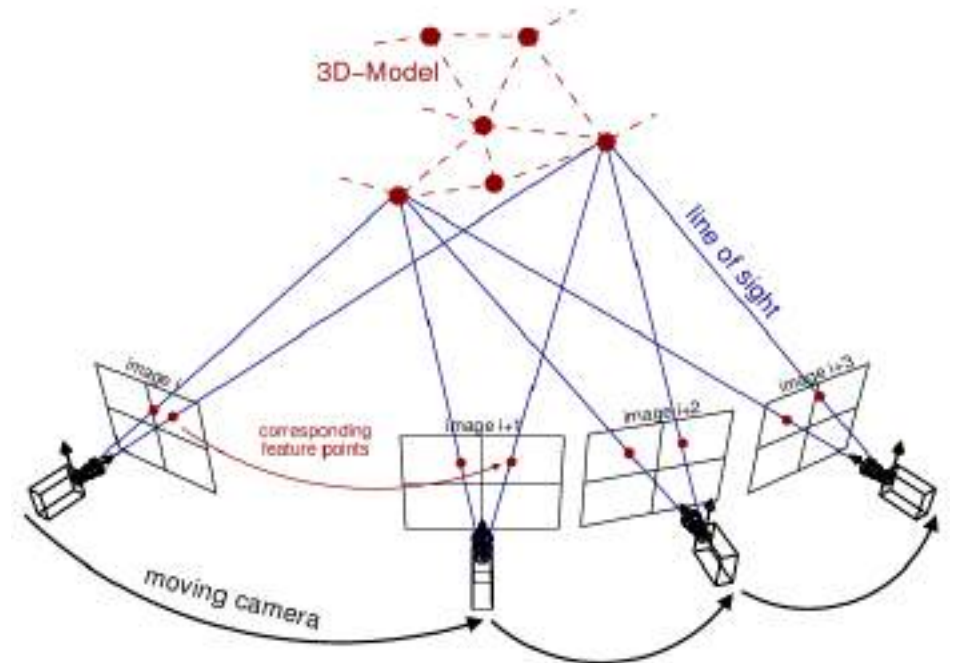
- ToF and LiDAR cameras are based on signal modulation and ranging.
- Measuring the distance between the sensor and a target object by detecting the time difference from the signal emitted by a transmitter, reflected on a target object, and captured back by a receiver.



[The drone LiDAR operation principle - LiDAR drone OnyxScan, UAV 3D laser scanner \(onyxscan-lidar.com\)](https://onyxscan-lidar.com)

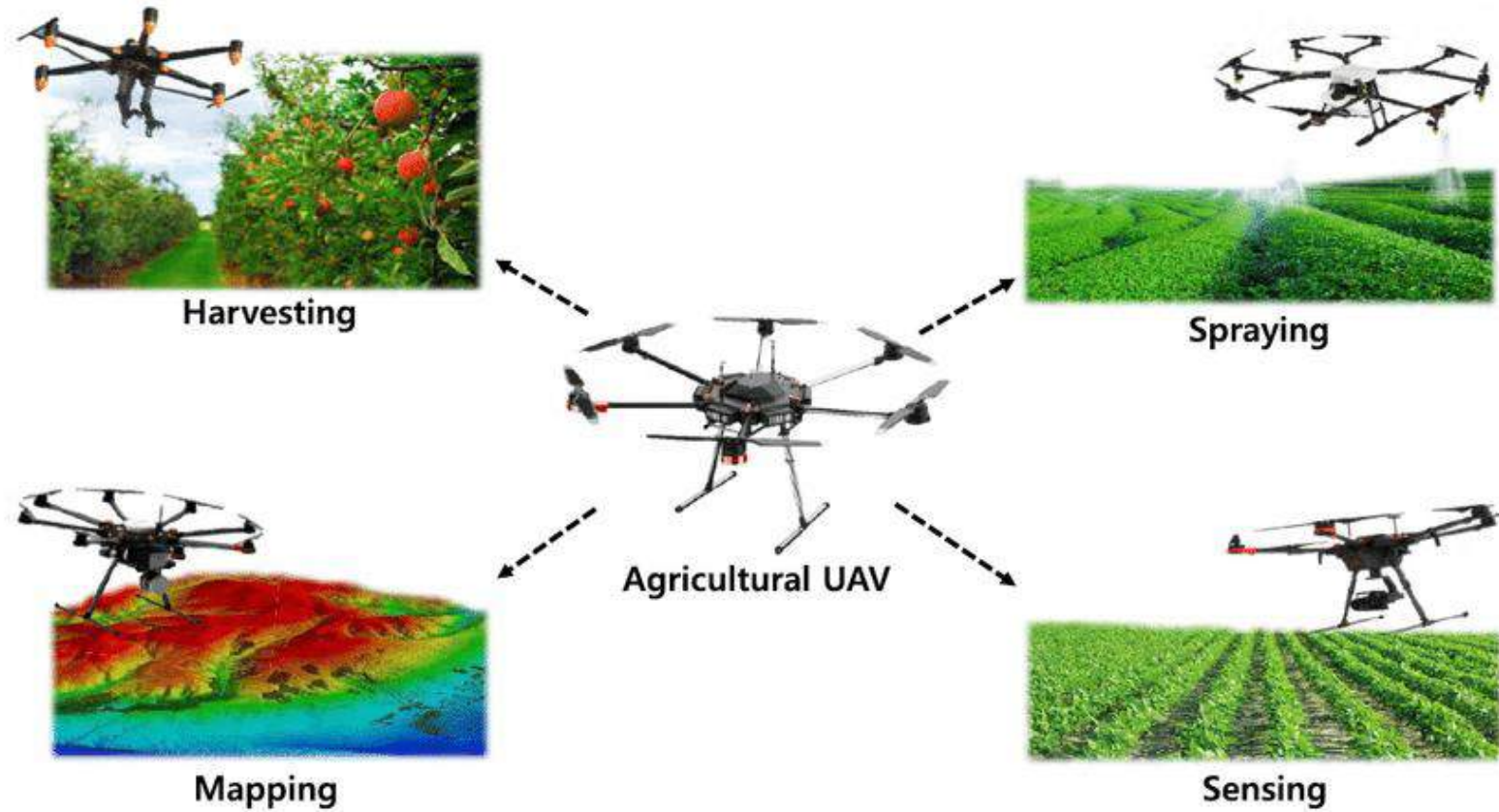
Structure from Motion (SfM)

- Creating a disparity map between the images from a single moving camera.
- The distance between the points where each image was captured by the camera can be used as the distance between “cameras.”
- The main challenge of such a strategy is that the object of interest must remain practically motionless.



Structure from Motion (SfM) photogrammetric principle.
Source: Theia-sfm.org (2016)

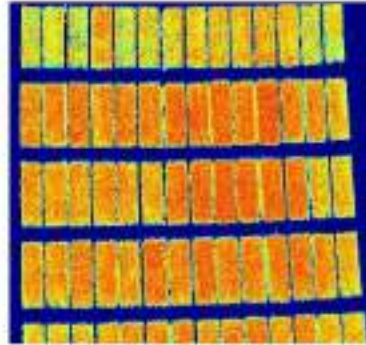
What can we do with a UAV?



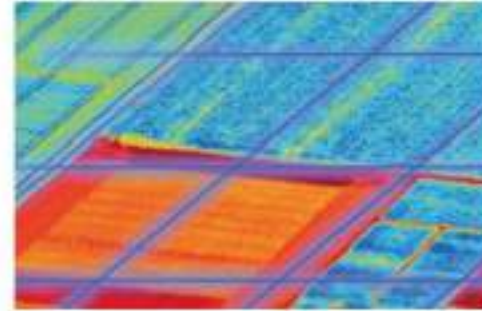
How?



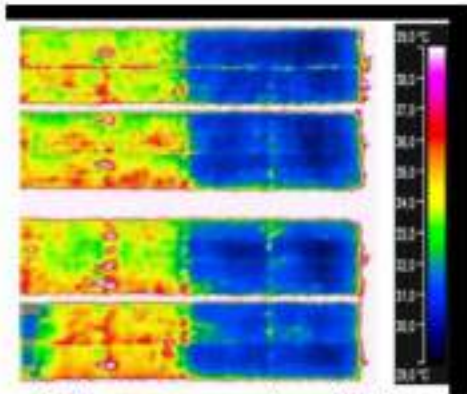
RGB images



Multispectral images



Hyperspectral images

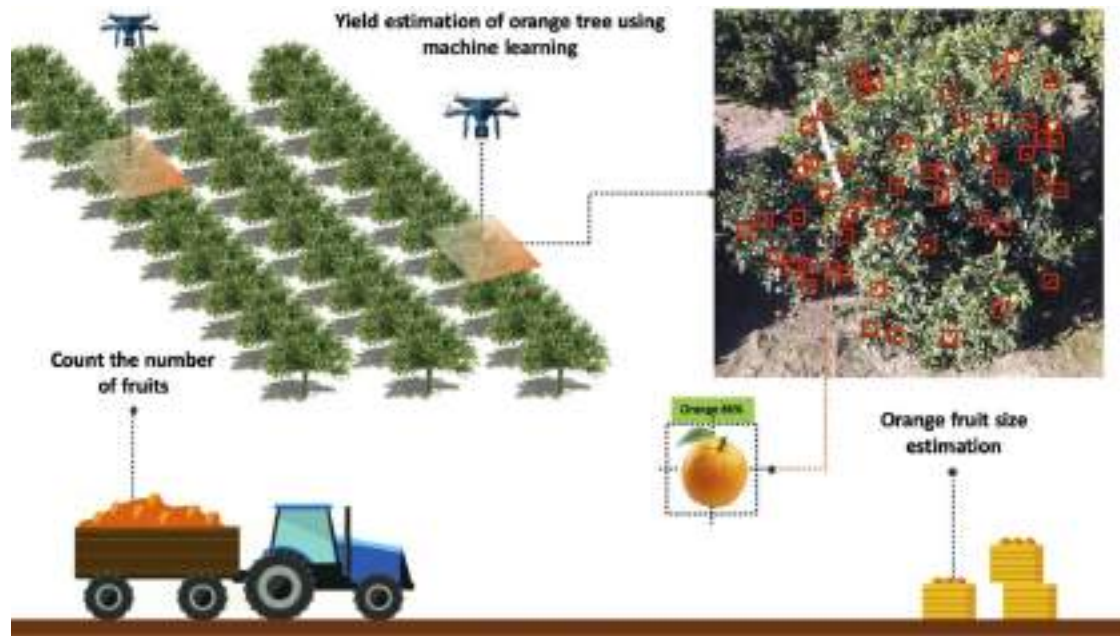


Thermography (IR)

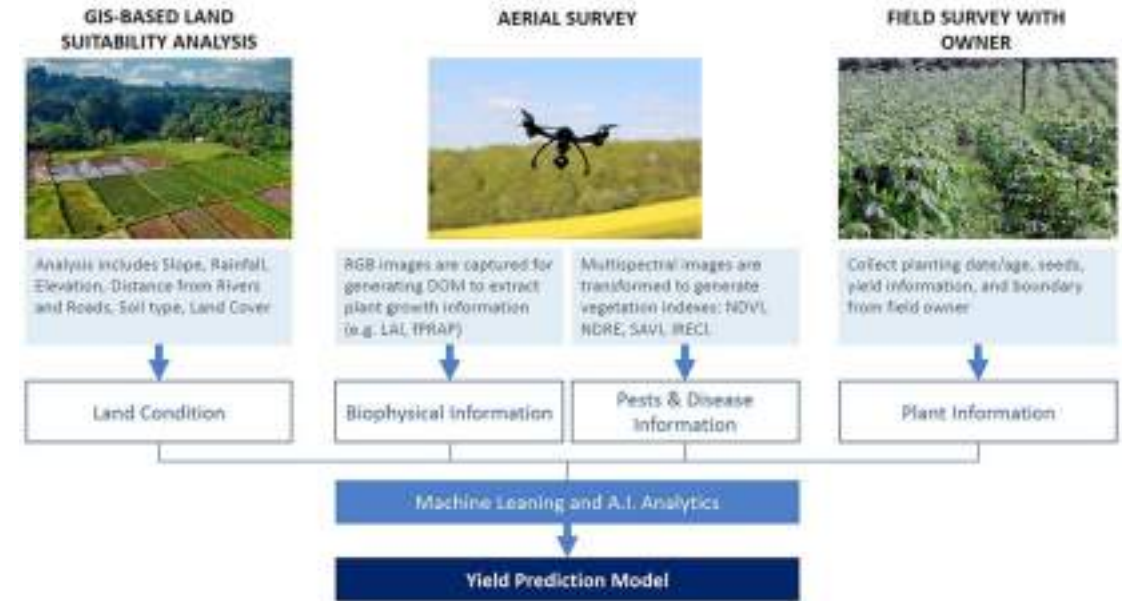


3D models

Yield Estimation

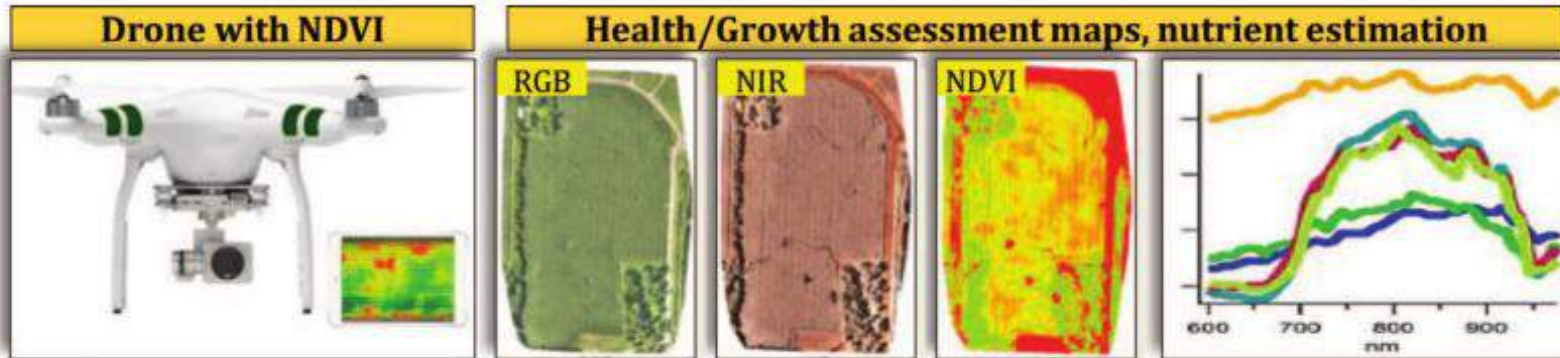


<https://www.sciencedirect.com/science/article/pii/S1161030120300381#fig0055>

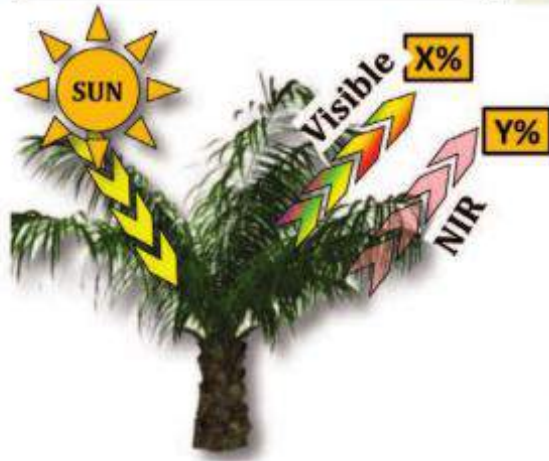


<https://anavision.com/blog/cassava-yield-prediction/>

Disease Detection

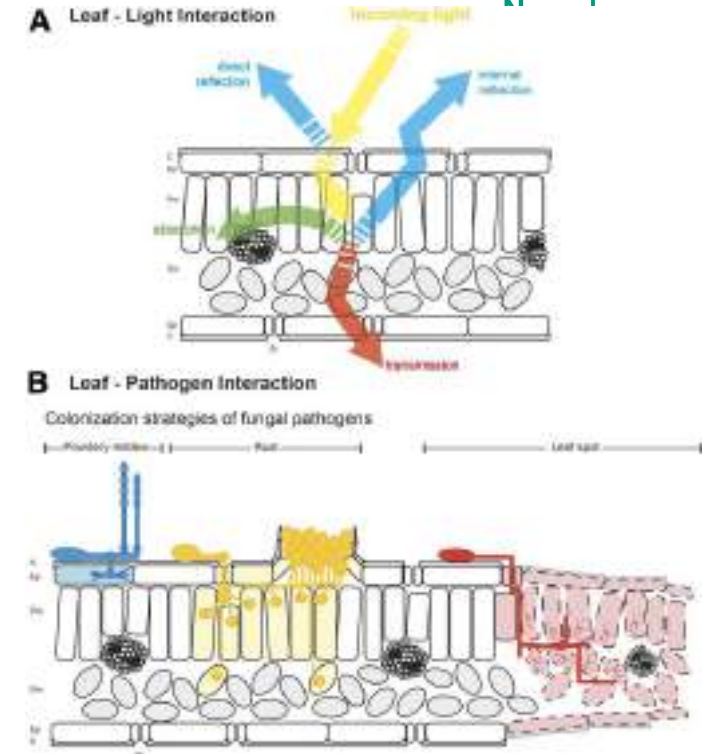


NDVI < 0 : Dead plants, NDVI < 0.33 Unhealthy
 0.33 < NDVI < 0.66: Healthy , NDVI > 0.66 Very healthy



$$-1 \leq NDVI = \frac{X\% - Y\%}{X\% + Y\%} \leq +1$$

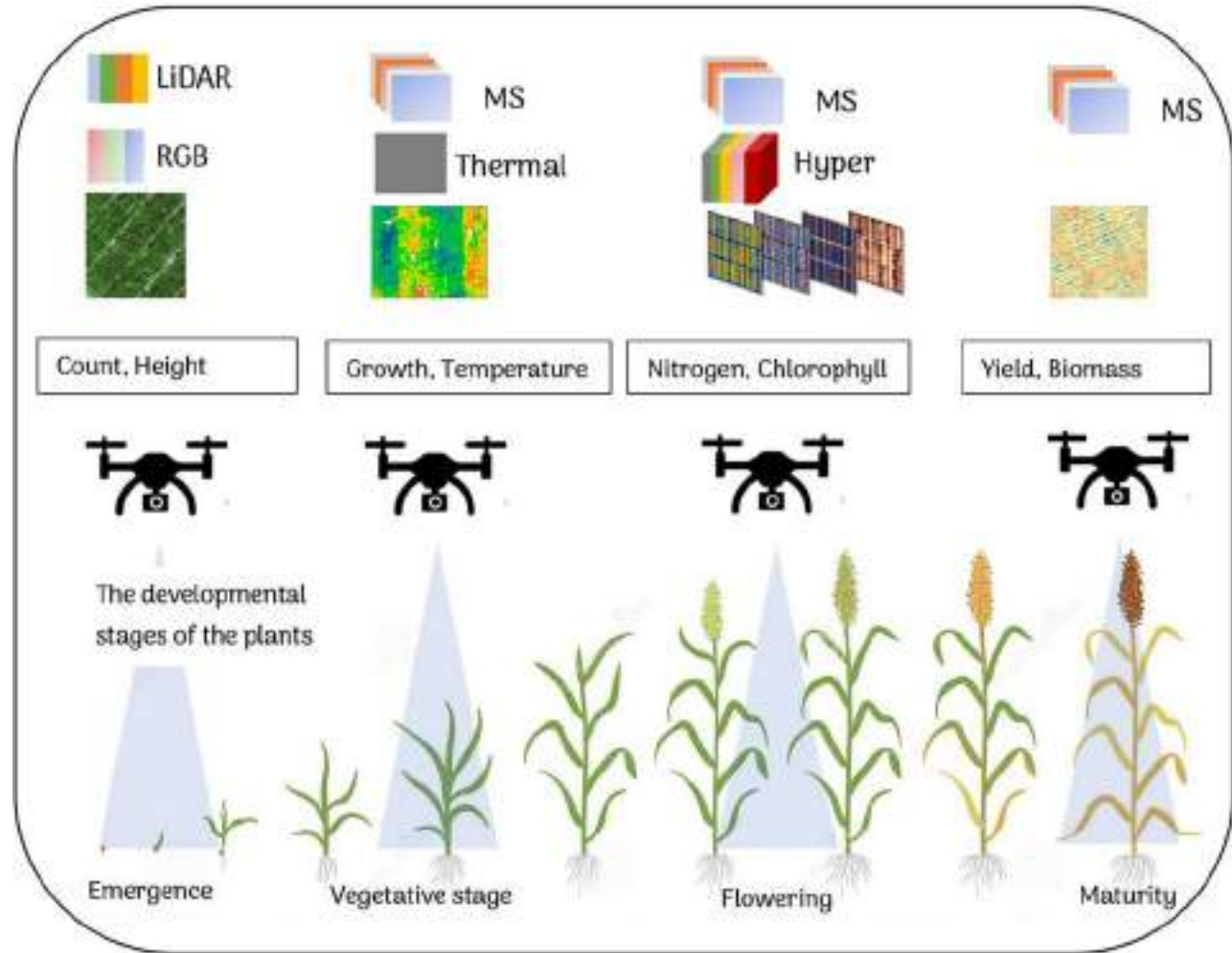
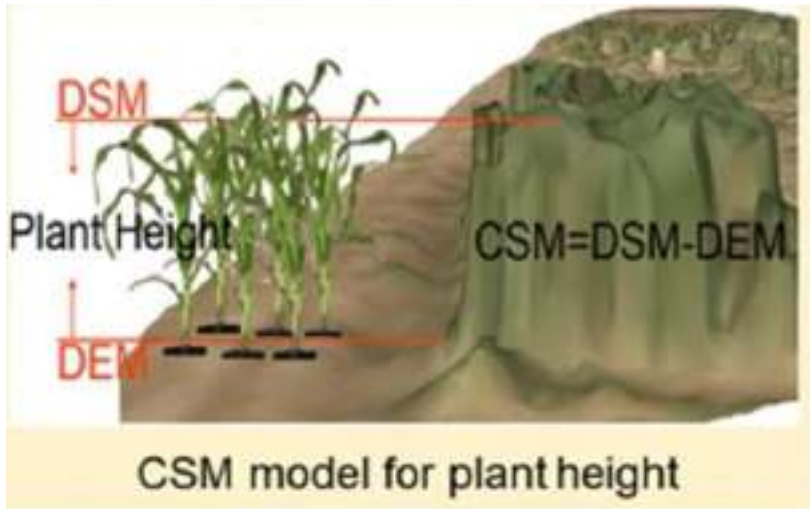
UAV sample NDVI mapping for health assessment and disease detection (Shamshiri et al., 2018).



DOI: 10.1094/PDIS-03-15-0340-FE

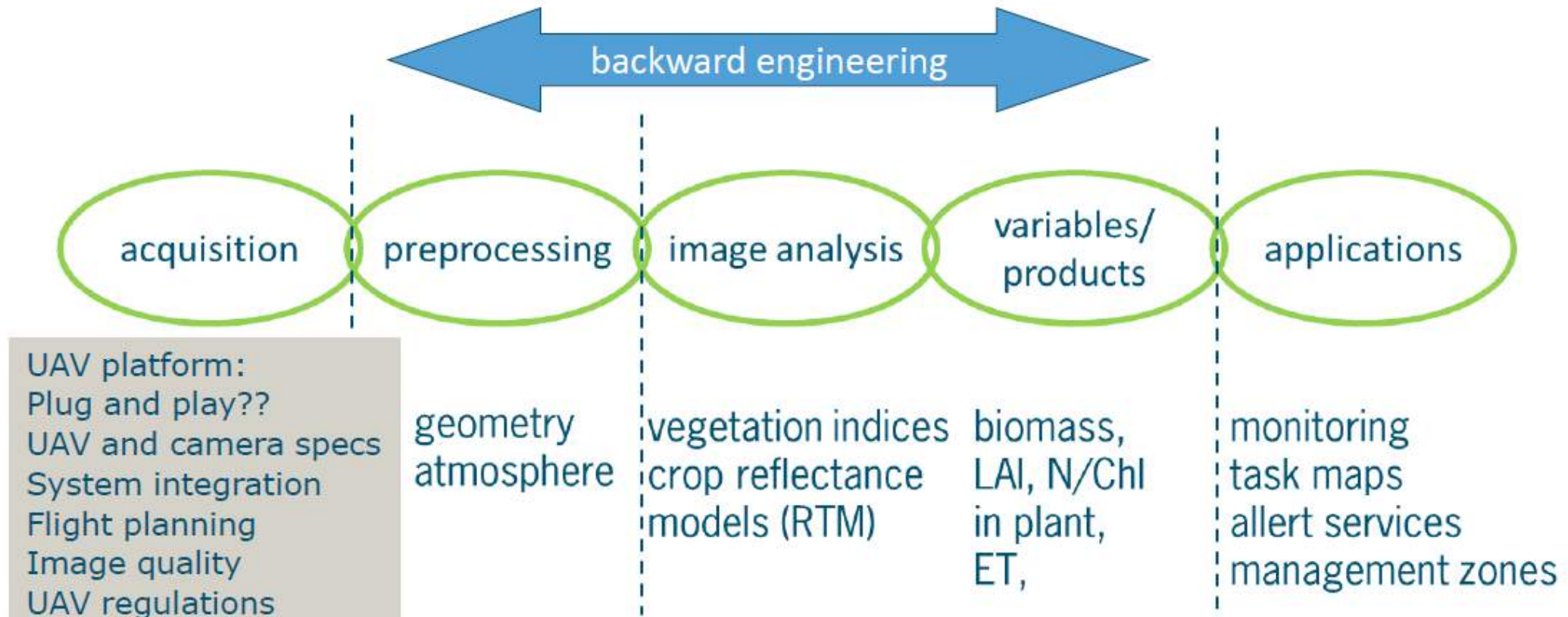


Plant Height:



Schematic overview of the different ways to extract spectral information from plants throughout a growing season(<https://doi.org/10.1002/ppj2.20100>Digital Object Identifier (DOI)).

UAV processing chain



Introduction to Photogrammetry



- **Photogrammetry:** Science of obtaining measurements and spatial information from photographs.
- Uses **multiple overlapping images** taken from different angles.
- Applies **geometric principles** to reconstruct objects or terrain in **3D**.
- Images can be captured using **drones, aircraft, satellites, or handheld cameras**.
- Processed with **specialized software** to generate:
 - 3D models
 - Orthophotos
 - Digital Elevation Models (DEMs)
 - Maps
- Widely used in **surveying, mapping, construction, archaeology, engineering, and environmental monitoring**.
- Provides **accurate, efficient, and non-contact data collection**.
- Increasingly accessible due to **advances in drones, digital cameras, and computing power**.

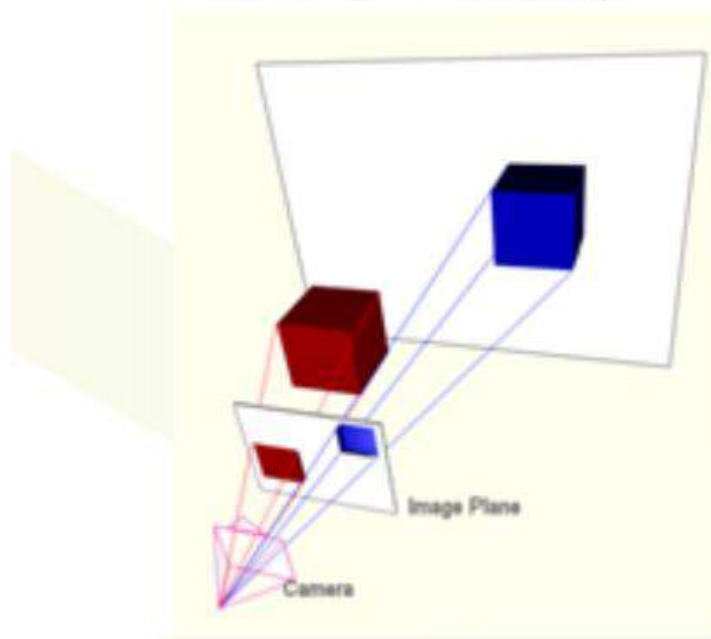


From 2D to 3D-Stereo vision

From 3D to 2D

One camera

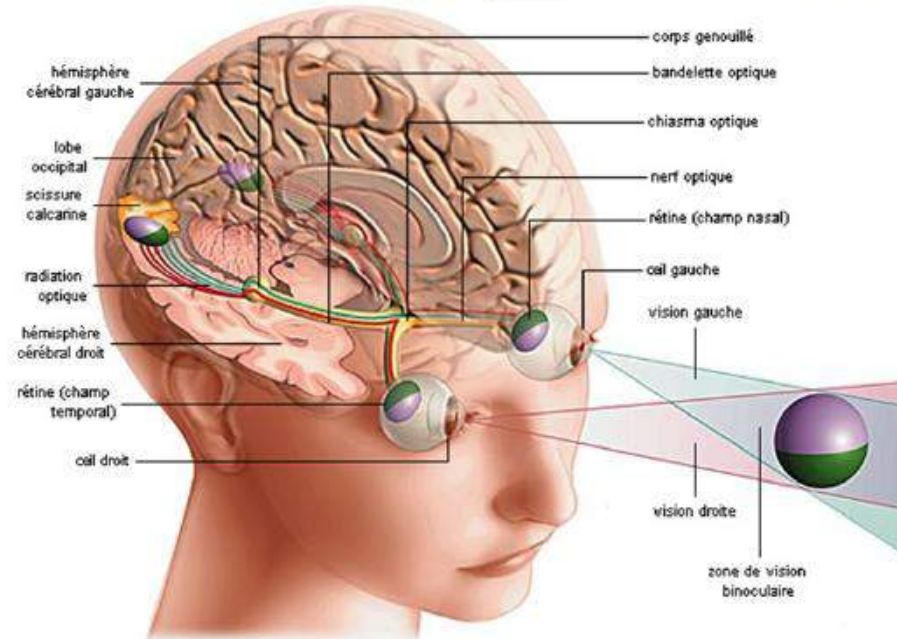
Far appears smaller than near on image



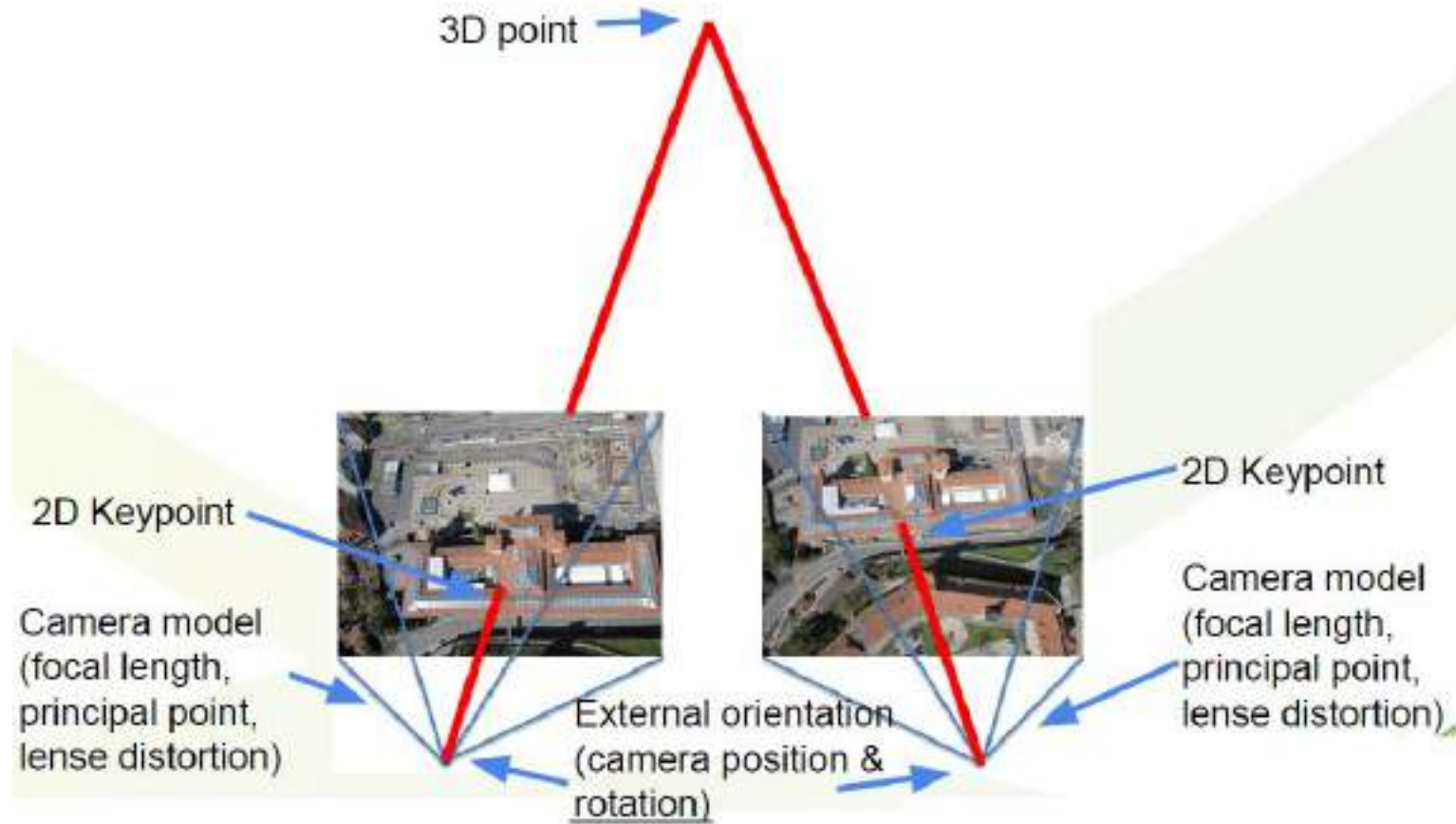
From 2D to 3D

Two images

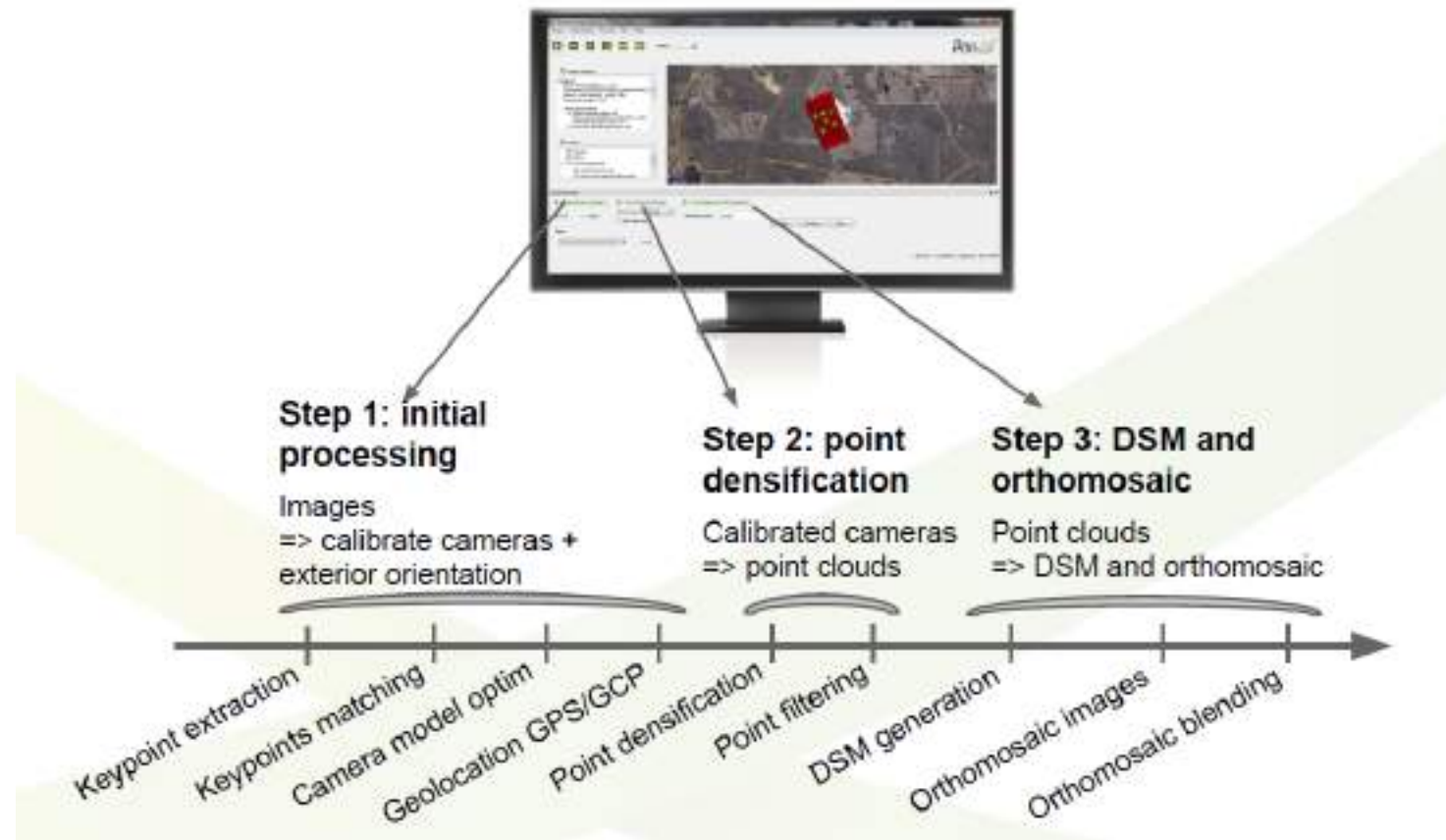
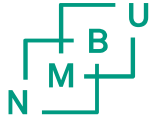
Triangulate to get sense of depth



From 2D to 3D Stereoscopy

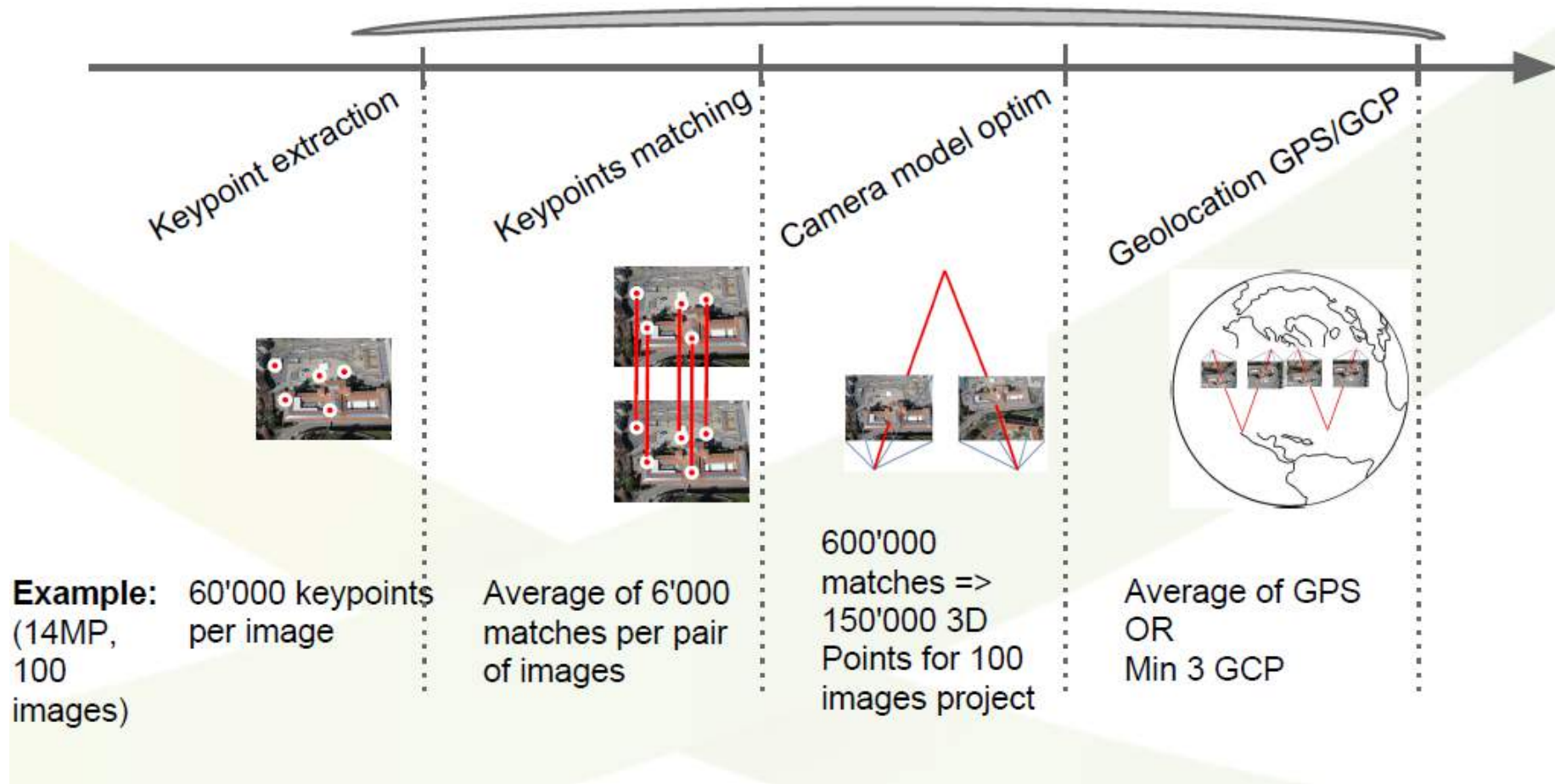


3 main processing steps in Pix4D



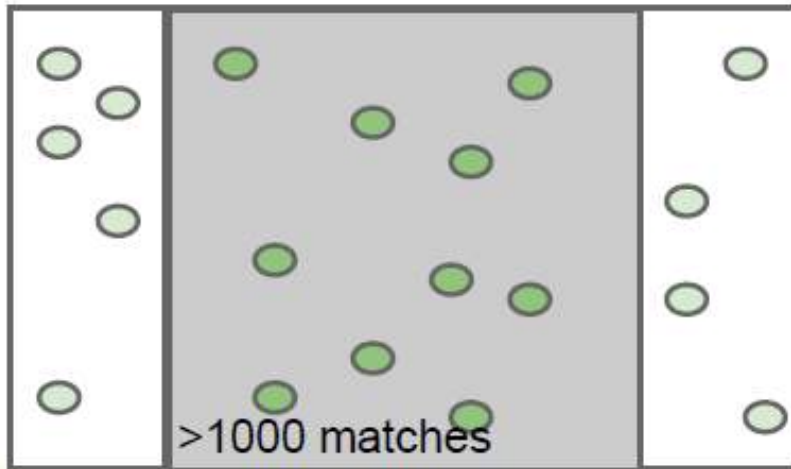
Step 1: initial processing

Images => calibration + exterior orientation

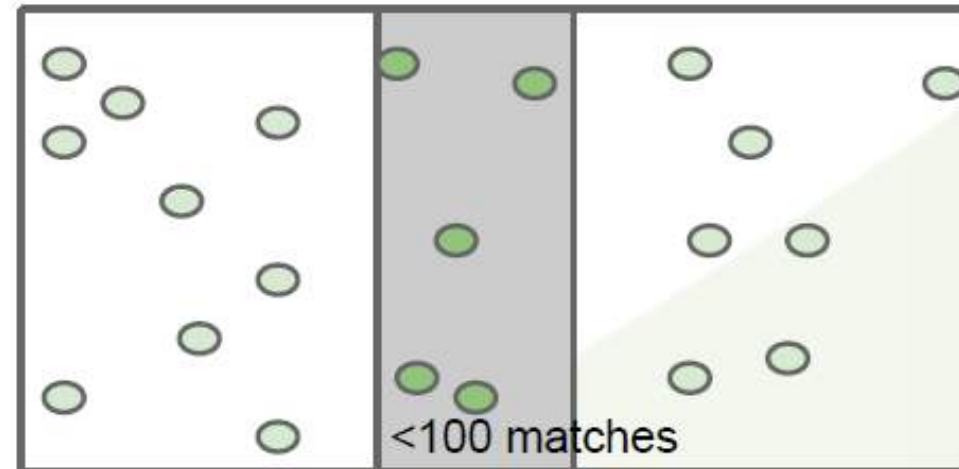


Getting Enough Matches

Enough matches:



75% overlap
enough matches



20% overlap
not enough matches



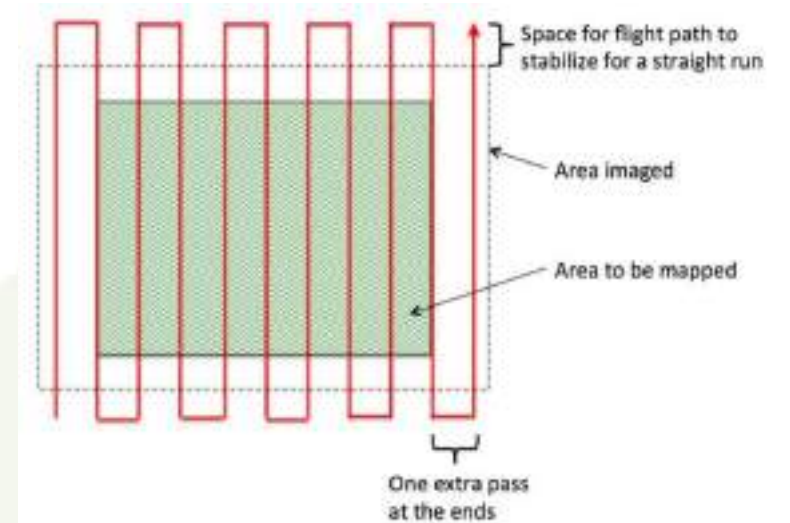
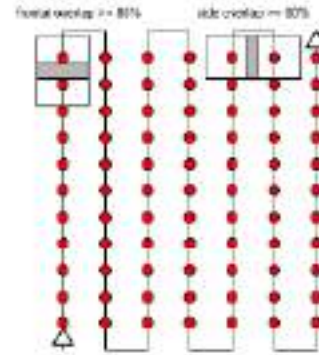
Extract Keypoints

- **Good Images:**
- Natural textures: **rocks, bushes, dirt**
- Urban features: **buildings, streets**
- Resolution: **>10 MP**
- Keypoints: **>10,000 extractable points**
- **Difficult Images:**
- Low-texture surfaces: **sand, snow, fog**
- **Blurry or out-of-focus** images
- **Overexposed or underexposed**
- Low resolution: **<3 MP**
- Keypoints: **<100 extractable points**



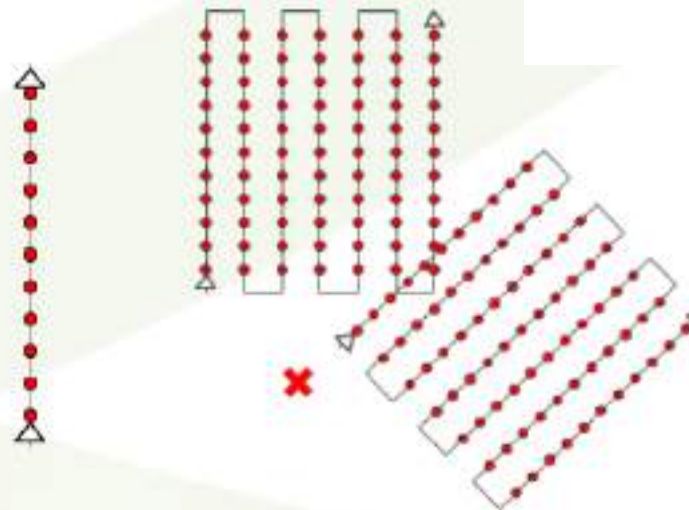
Ideal flight plan

- regular grid flight plan
- easy terrain: 75% frontal, 50% side
- difficult terrain: 85% frontal, 60% side

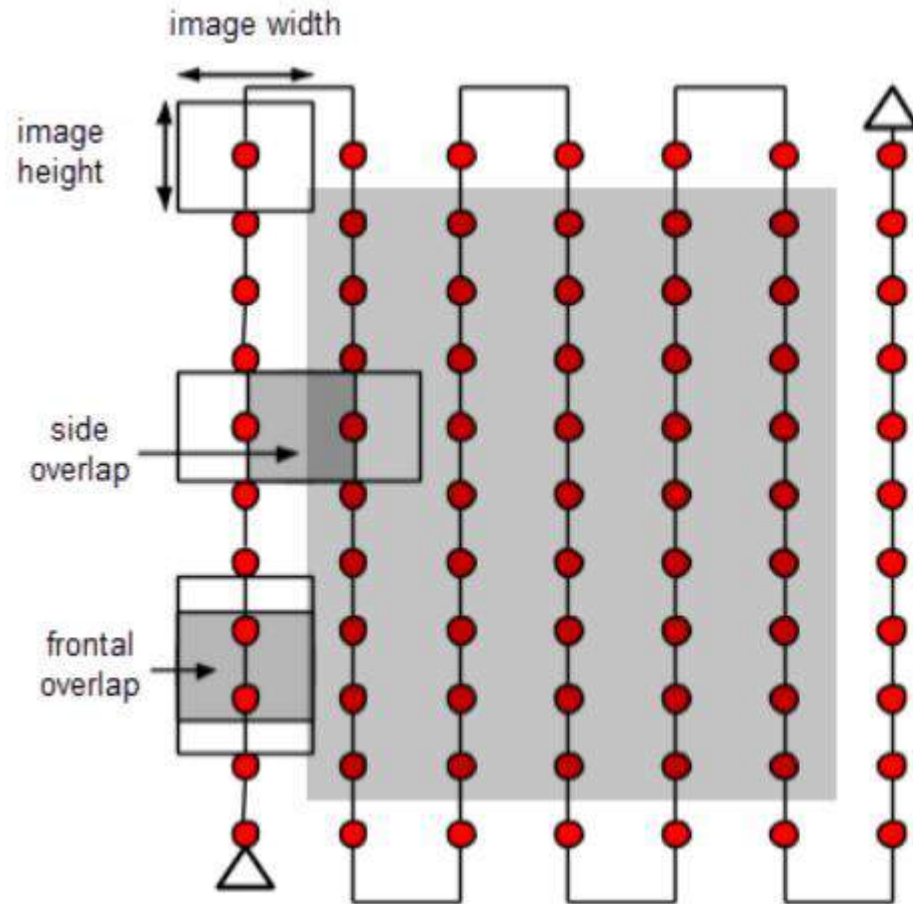


Difficult flight plan

- low overlap
- multiple images at same location
- corridor mapping
- high difference in altitude (>2xGSD)



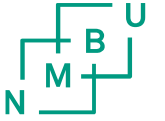
Flat terrain with agriculture fields



- **overlap** between images to at least 85% frontal overlap and at least 70% side overlap.
- **Fly higher.** In most cases, flying higher improves the results.
- Have **accurate image geolocation** and use the *Agriculture* template for processing in *pix4D*.

Designing the Image Acquisition Plan

Computing the Flight Height for a given GSD



- **Ground Sampling Distance (GSD)**
- **Definition:** Distance between the centers of two consecutive pixels on the ground.
- **Importance:**
 - Determines the **accuracy and quality** of the final results.
 - Affects the **level of detail** visible in the orthomosaic.
- **Smaller GSD** → **higher resolution** and more detailed maps.

GSD = desired GSD [cm/pixel]

$$H / F_R = D_W / S_W$$



$$H [m] = (imW * GSD * F_R) / (S_W * 100)$$

$$D_W = (imW * GSD) / 100.$$



S_W = real sensor width [mm]

F_R = real focal length [mm]

H = flight height [m]

D_W = distance covered on the ground by one image in the width direction (footprint width) [m]

Designing the Image Acquisition Plan

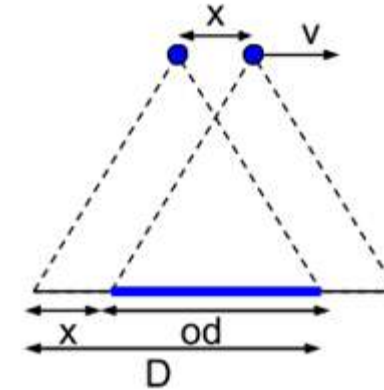
Computing the Image Rate for a given Frontal Overlap



- The image shooting rate to achieve a given frontal overlap depends on the speed of the UAV/plane, the GSD, and the pixel resolution of the camera.

D = distance covered on the ground by one image in the flight direction [m]
 overlap = percentage of desired frontal overlap between two images
 od = overlap between two images in the flight direction [m]
 x = distance between two camera positions in the flight direction [m]
 v = flight speed [m/s]
 t = elapsed time between two images (image rate) [s]

D_h = distance covered on the ground by one image in the height direction (footprint height) [m]
 imH = image height [pixel]
 GSD = desired GSD [cm/pixel]



Camera oriented with the sensor width (long dimension) perpendicular to the flight direction (usual case)

$$od = \text{overlap} * D \quad (1)$$

$$x = D - od \quad (2)$$

$$t = x / v \quad (3)$$

$$D = D_h = (imH * GSD) / 100 \quad (4)$$

Combining Equations (1) and (4) into Equation (2):



$$x = ((imH * GSD) / 100) * (1 - \text{overlap}) \quad (5)$$

$$t = x / v = ((imH * GSD) / 100) * (1 - \text{overlap}) / v \quad (6)$$

Combining the equations (3) and (5):

Flight Height and Image Overlap

- Higher altitude → lower resolution of captured images.
- For mature crops / general variability mapping:
 - High resolution not required → fly higher (20–120 m)
- For young plants / emergence studies:
 - Fly lower (20–60 m)
 - Allows individual plants to be distinguished from the soil in the final map.

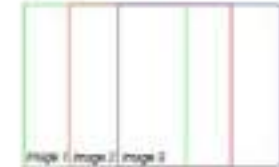
0% overlap



20% overlap



70% overlap

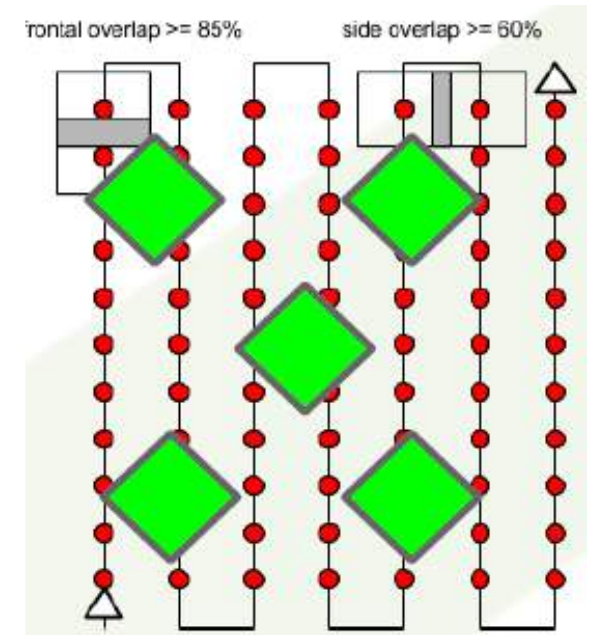


Model Georeferencing

Project size, orientation, and position are initially unknown

GPS Tags

- Provide **approximate position, orientation, and scale** of images
- Require **good synchronization** between camera and GPS
- **Higher number of GPS-tagged images** → **better position approximation**
- Typical GPS accuracy may include **~10 m random positional shift**



Collecting Ground Control Points (GCPs)

Points with known coordinates located within the area of interest.

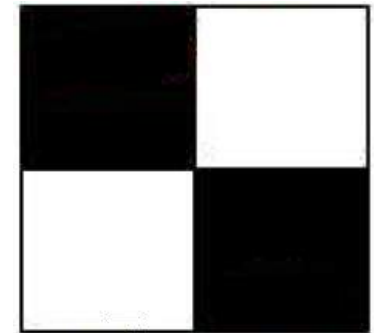
- **Projects with image geolocation:**

- GCPs **improve absolute accuracy** by placing the model in the correct real-world position.

- **Reduce GPS positional errors** from meters to **centimeter-level accuracy.**

- **Projects without image geolocation:**

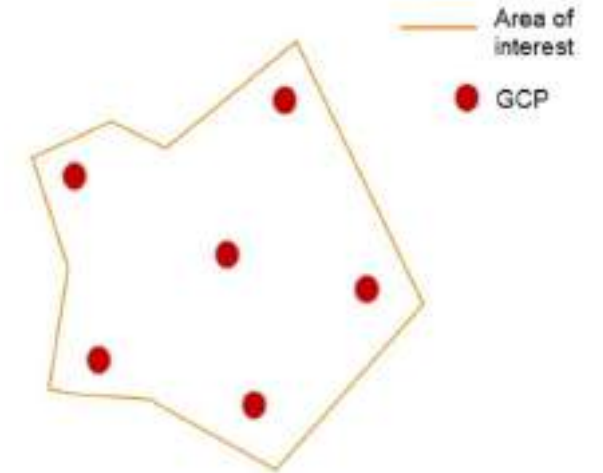
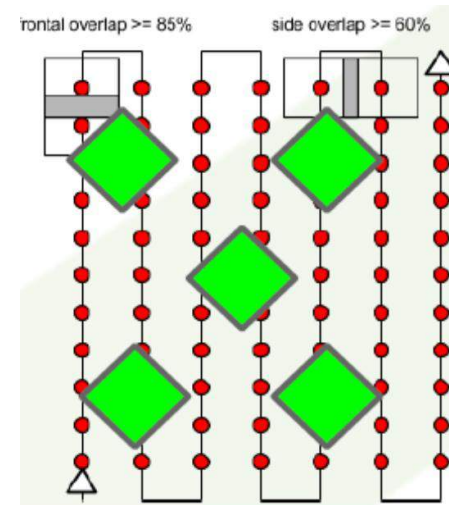
- **GCPs are required** to correctly position and scale the model.





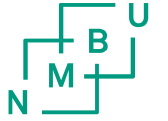
How Many GCPs?

- A minimum number of 5 GCPs is recommended per project.
- The GCPs should be placed evenly on the landscape to minimize the error in Scale and Orientation.
- Do not place the GCPs exactly at the edges of the area
- Try to place one GCP in the center of the area



Distribution of the GCPs.

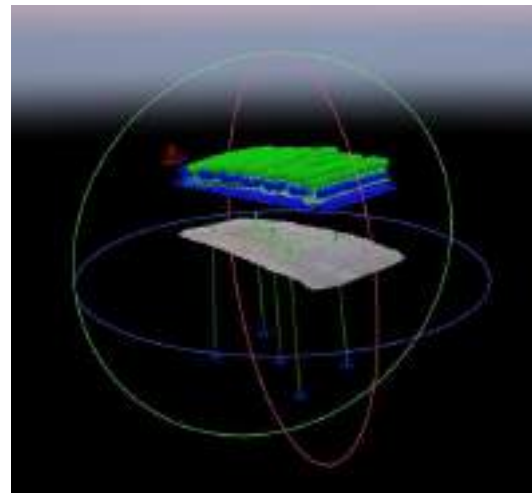
Map Without GCPs and how to correct



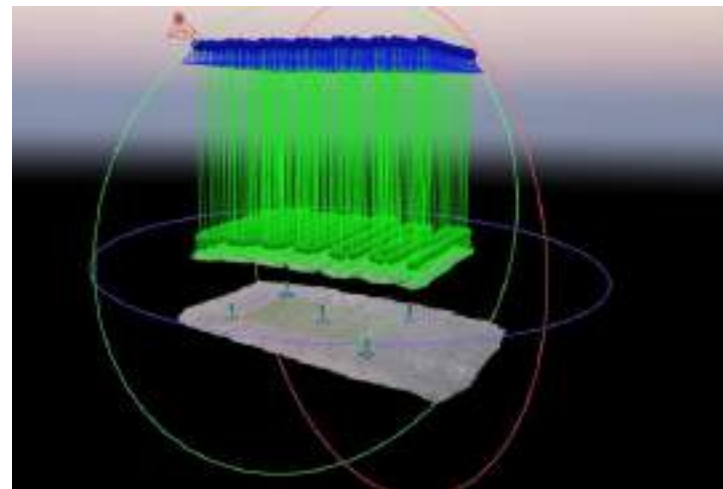
- Map Without GCPs: May have incorrect position, scale, or orientation, GPS errors can cause shifts of several meters, Lower absolute accuracy; suitable only for rough visualization
- How to Correct:



Add (GCPs)

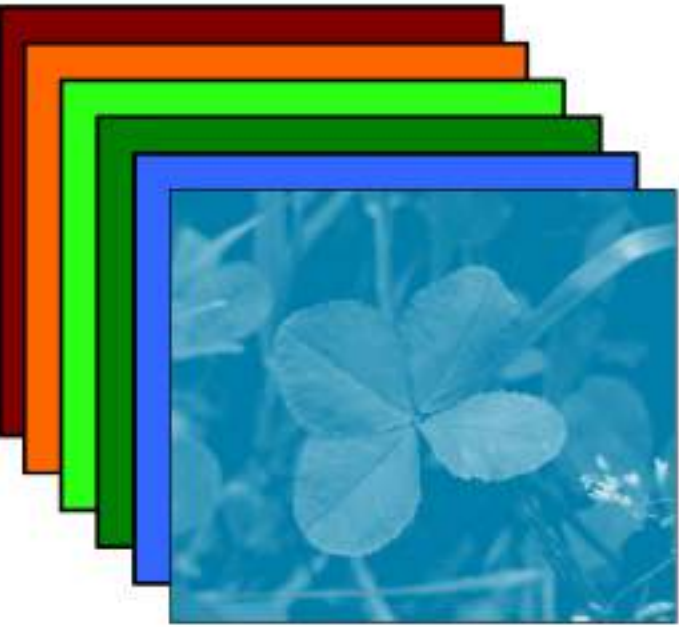
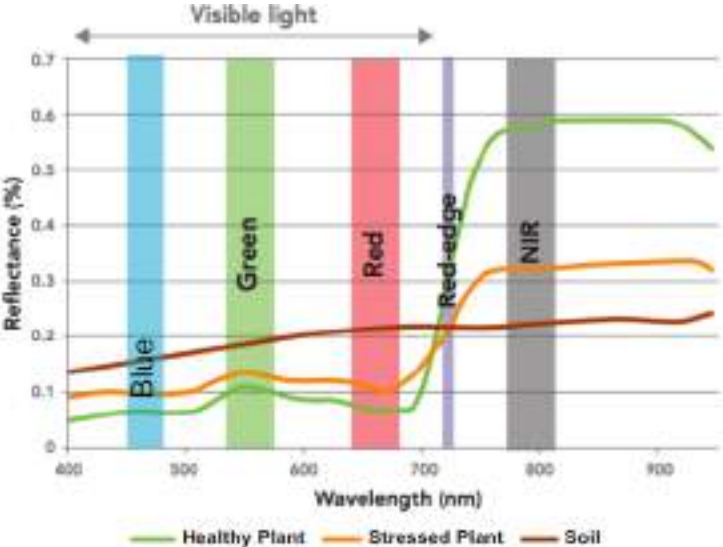


Without GCPs



With GCPs

Multispectral Imaging



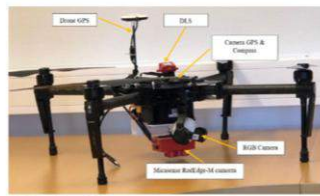
$$NDVI = \frac{NIR - Red}{NIR + Red}$$

$$MTCI = \frac{NIR - Rededge}{Rededge - Red}$$

$$EVI = 2.5 * \frac{(NIR - Rededge)}{(NIR + 6 * Red - 7.5 * Blue + 1)}$$



Micasense RedEdge

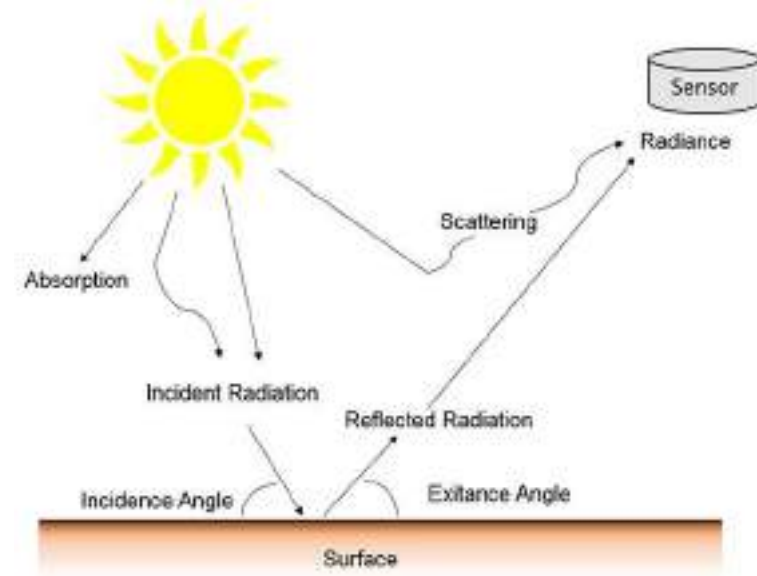
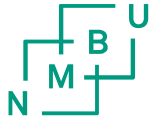


Matrix 100

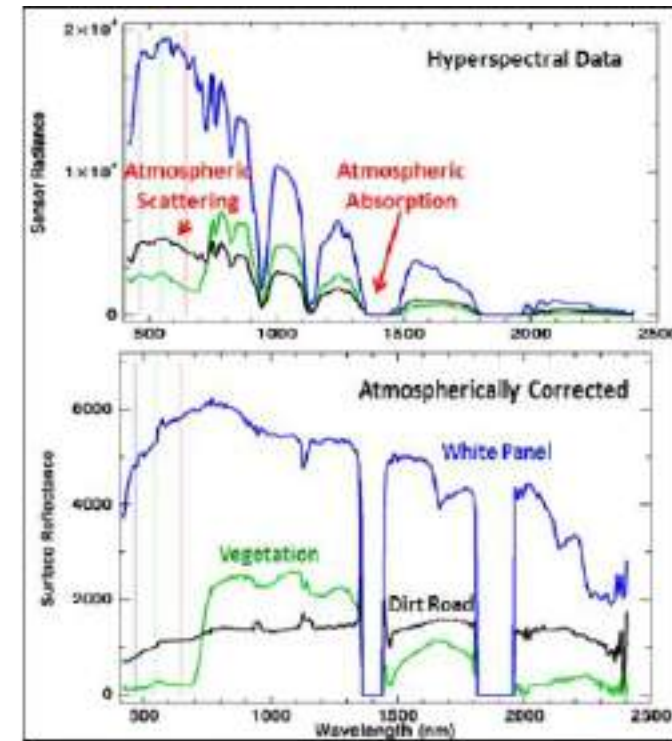
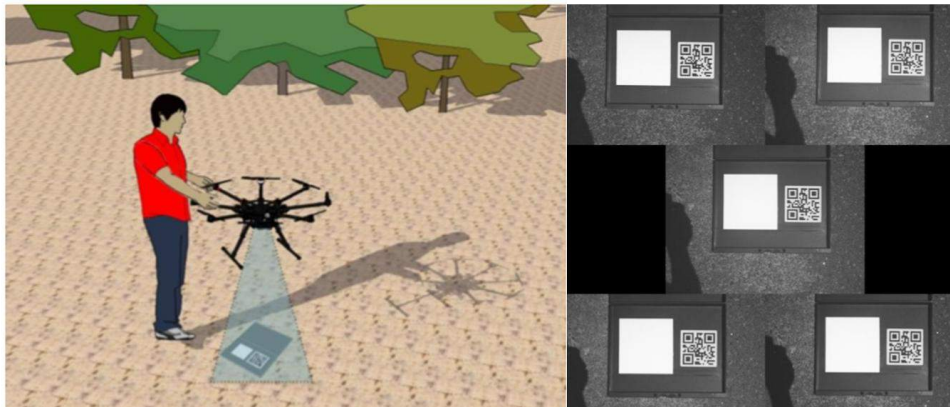


Multispectral cameras

Calibration



Factors that affect radiance (Diagram from Humboldt State University [1])



Spectral signatures showing radiance versus reflectance for a pure white panel, vegetation, and a dirt road. The radiance signatures show some of the main areas where atmospheric effects are a problem. [2]

[The importance of calibrating your remote sensing imagery - Materials Talks \(materials-talks.com\)](https://materials-talks.com)

Questions?

Sahameh.shafiee@nmbu.no

