GCIRC CAMBRIDGE

Digitizing the plant trialing Industry With High-throughput Phenotyping: A Decade Of Progress And Perspectives

DATE

April 2025

AUTHOR(S)

Alexis Comar, Nicolas Cheviet, Hugh Frost & All Hiphen team

The history of Hiphen

Alexis Comar, PhD

CEO & founder of Hiphen

Following his thesis on high-throughput digital phenotyping, Alexis decided to found Hiphen, the contraction of “**Hi**gh-throughput” and “**Phen**otyping”, to bridge the gap between Research and Industry by developing remote sensing applications and image analytics solutions that empower breeders, product developers and crop scientists with greater phenotyping capabilities globally.

2014 – The **company is founded** in Avignon, France on October 22nd

2016 – Winner of a **prestigious french grant** (I-lab) + Launch of the 1st **product** (Airphen camera)

2018 – HIPHEN® joins the **CAPTE Research Unit** (co-founded with INRAe & Arvalis)

2020 – Launch of our homegrown data platform **Cloverfield™**

2021 – HIPHEN® reaches **20 employees**, still based in Avignon, France

2022 – Launch of **HIPHEN Corp.** our subsidiary in the USA

2023 – **Acquisition** of US-based plant phenotyping company **SlantRange®**

2024 – Launch of two new products **Literal** & **PhenoLite®**

2025 – **Acquisition** of **Aurea Imaging** Phenotyping Division

We transform images & signals into decisions

Deep agronomic knowledge

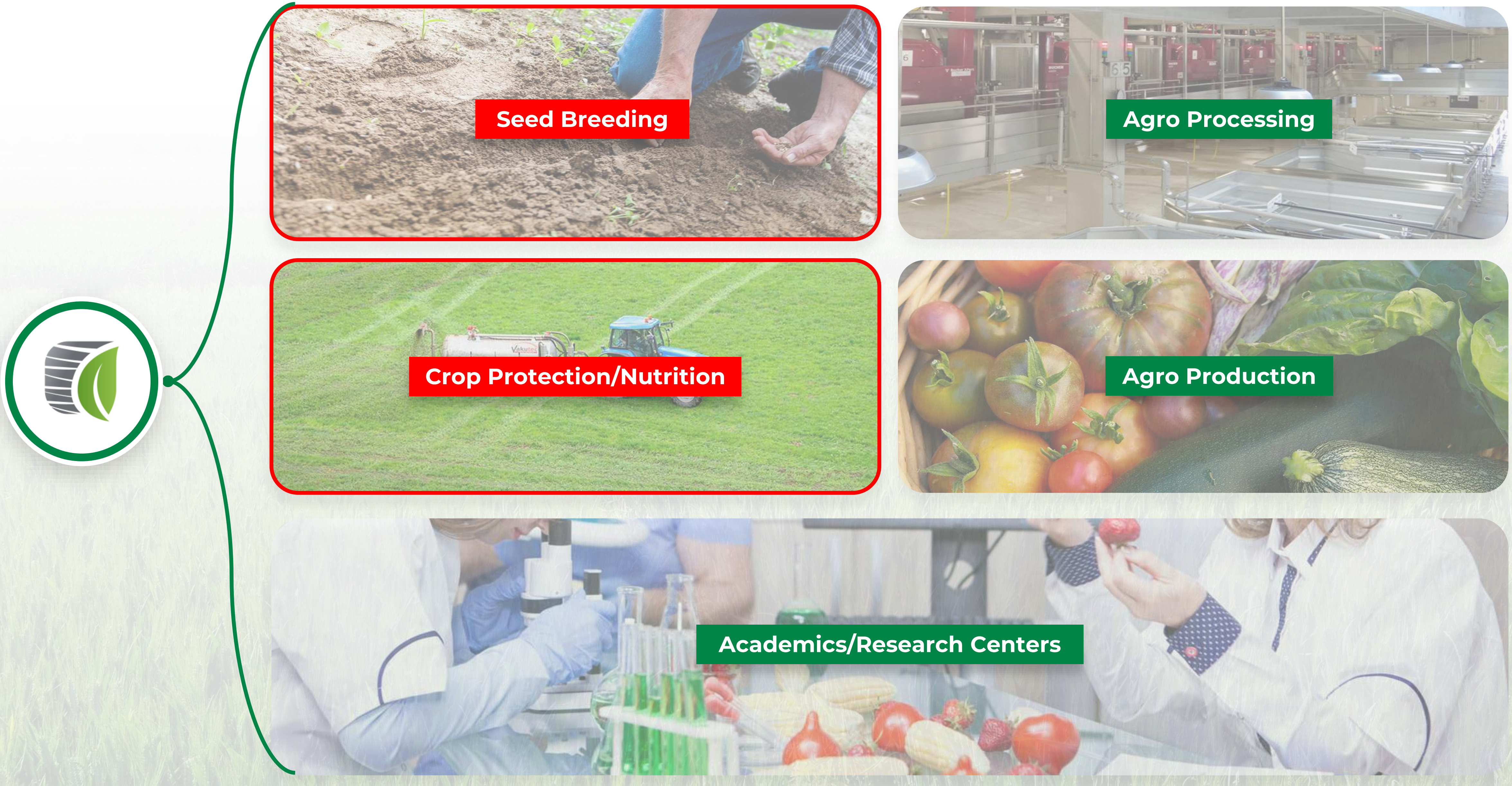
Mechanical engineering

Cloud-based data engine built for scale

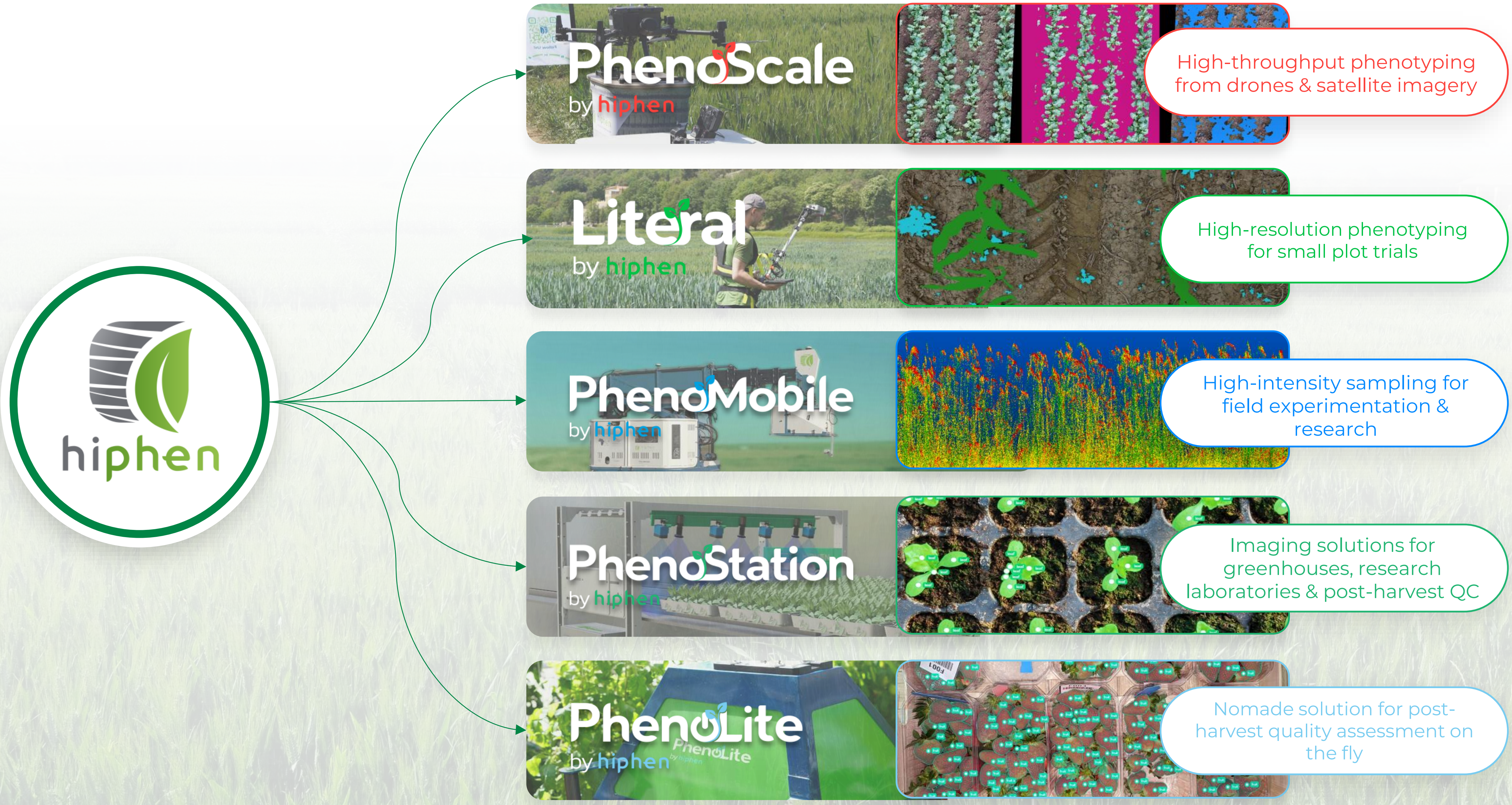
Tailor-made Equipment

Easy-decision making tools

Serving the entire Ag ecosystem



Our products to deliver powerful image analytics



Who is HIPHEN?

15M Plots Screened

>6k Flights Processed

70+ Traits

>18 Countries

PhenoMobile by hiphen

Literal by hiphen

PhenoScale by hiphen

PhenoStation by hiphen

On-time Phenotyping with the Cloverfield Data Platform

All Your Traits Only a Click Away

Upload Datasets Easily

Tailor-made Analytics Dashboards

Browse Our Trait Glossary

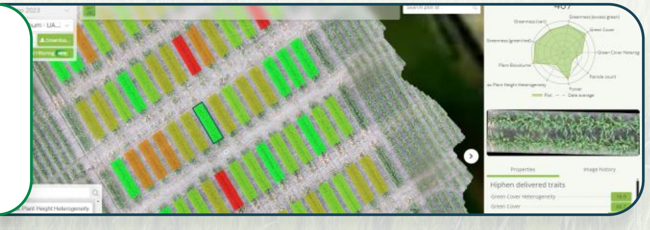
Monitor campaign Progress

The different use cases in the industry

Use case 1:
Field phenotyping in nurseries

Use case 2:
Field phenotyping for yield correction

Use case 3: Digitizing traits used in regulatory trials for crop protection



Use case 4 and perspective: towards PHENOMIC PREDICTION

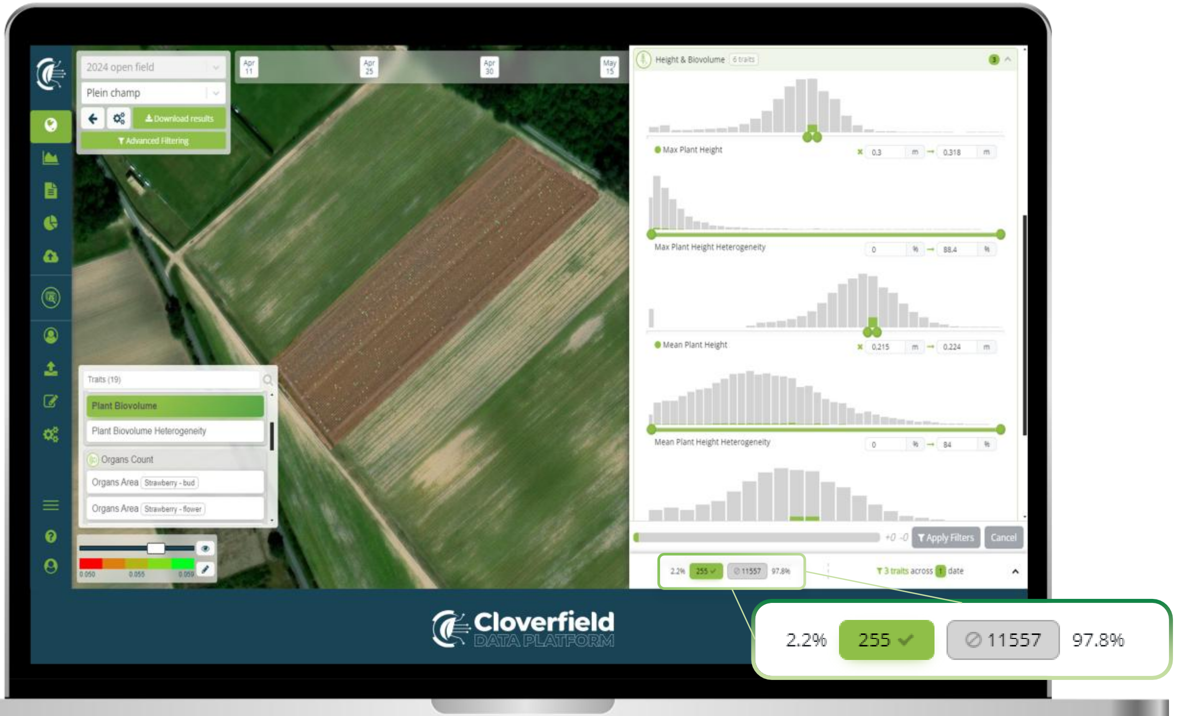
Phenotyping Strawberries at 'Nursery' Stage

1. Complex filters
2. Diversity on traits
3. Diversity on images

Diversity

Selected traits

- Precocity
- Biovolume
- Number of fruits
- Size of the flowers

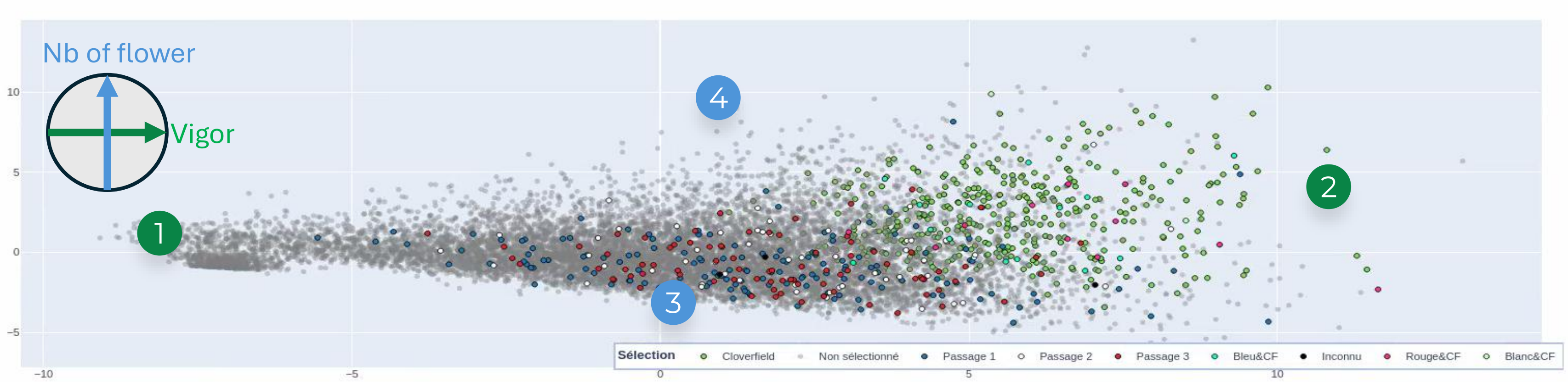


Phenomic Breeding values → Sensor data (Phenoscale) → Processing tech (Image to traits) → Mapping diversity (Traits histogram) → **Selection** (Which plant to advance ?)

Step 2 mapping diversity with traits

Phenomic Breeding values — Sensor data (Phenoscale) — Processing tech (Image to traits) — Mapping diversity (Traits diversity map) → **Selection** (Which plant to advance ?)

Diversity map of 12 000 strawberries plant



1

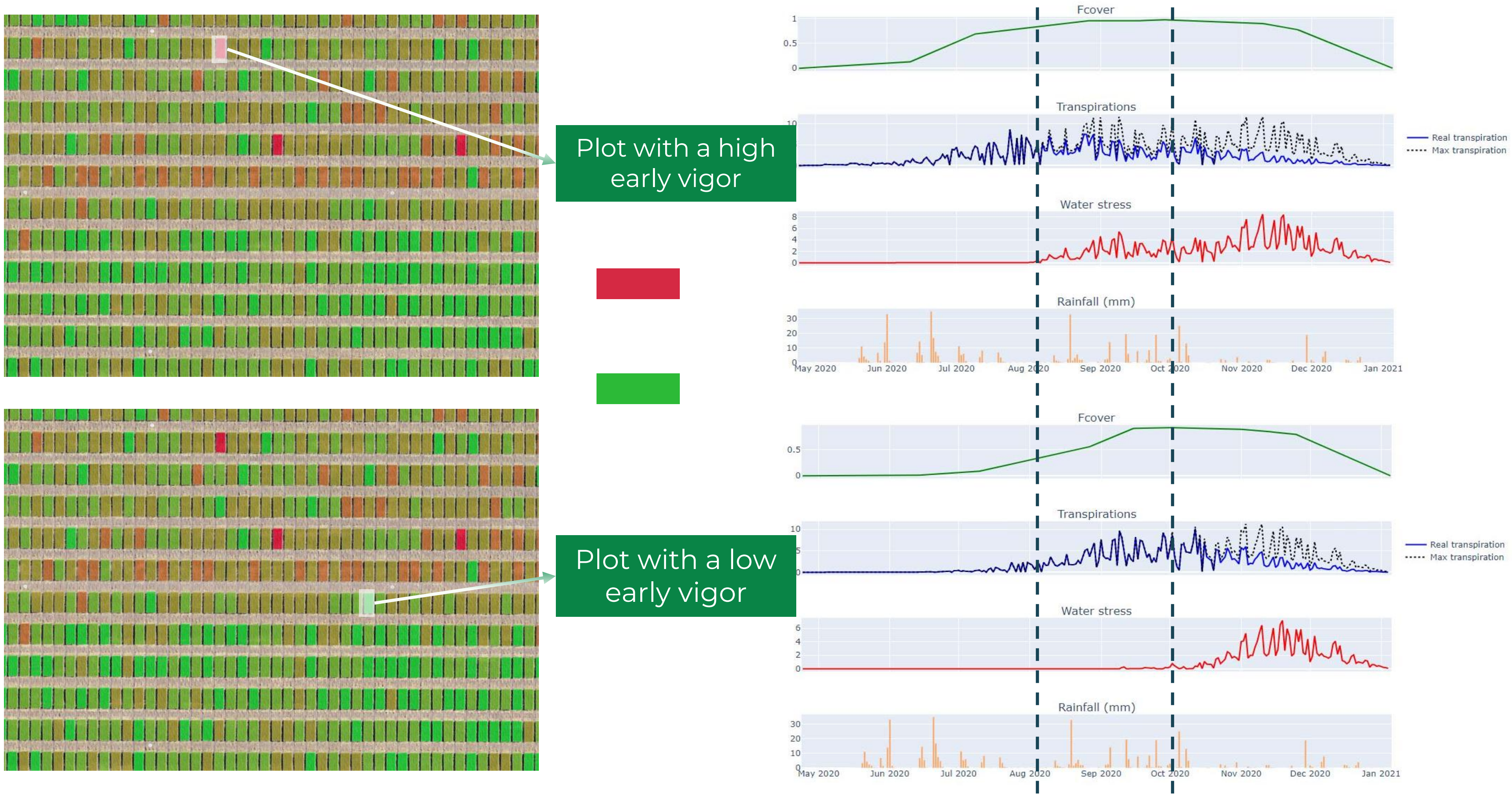

3


2


4

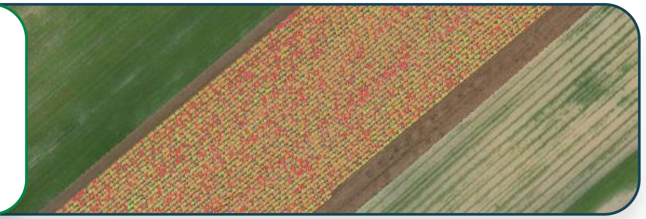

Screening for drought resistance through modeling

1. Water balance model
2. Disentangling the signal with instantaneous stresses
3. Calculating “efficiencies” (WUE – LUE, etc..)



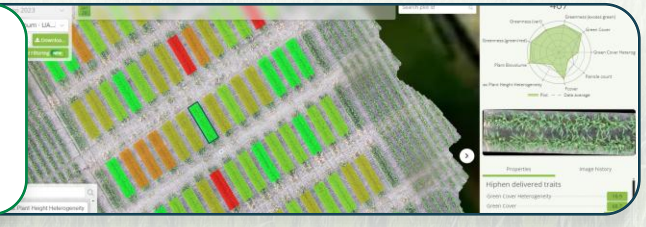
The different use cases in the industry

Use case 1:
Field phenotyping in nurseries



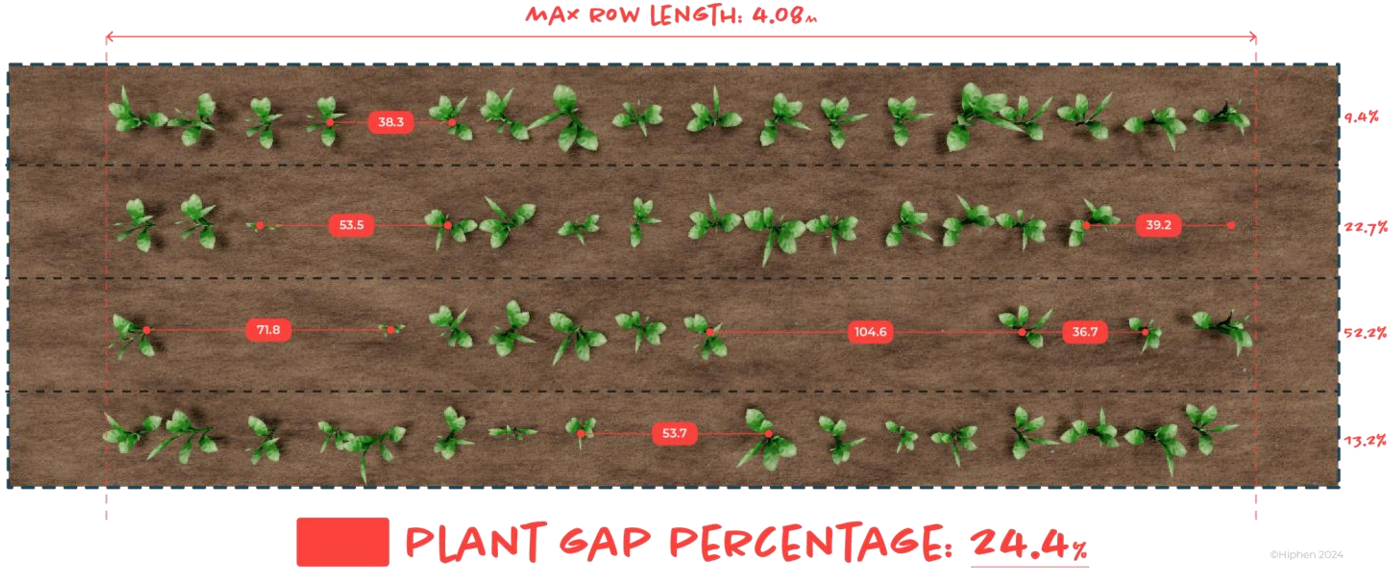
Use case 2:
Field phenotyping for yield correction

Use case 3: Digitizing traits used in regulatory trials for crop protection



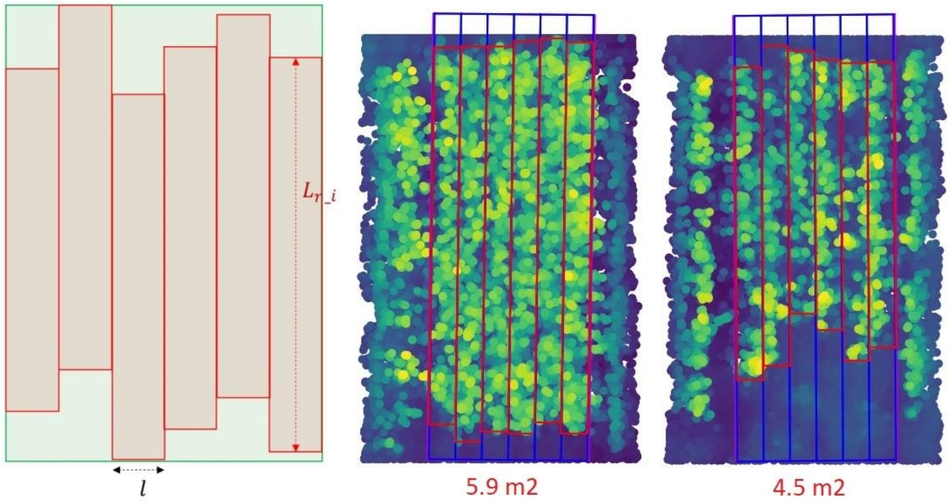
Use case 4 and perspective: towards PHENOMIC PREDICTION

Yield correction example – ***In Theory***



$$Yield_{raw} = \frac{Weight\ harvested\ (T)}{Surface\ (ha)}$$

$$Yield_{corrected} = \frac{Weight\ harvested\ (T)}{Measured\ Surface\ (ha)}$$



l: row spacing

Lr_i: Length of row i

$$Plot\ Area = \sum_{i=1}^{nb\_rows} L_{r\_i} \times l$$

Yield correction example – ***In Practice***

Plot 1 - 100% harvestable

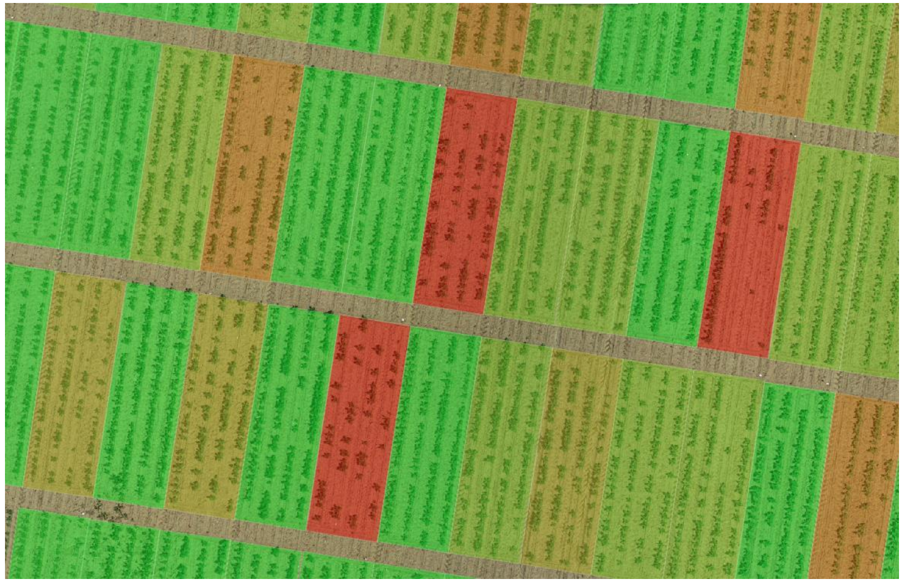
$$Yield_{corrected}\ P1 = Y_{harvested}$$

Plot 2 - 72% harvestable

$$Yield_{corrected}\ P2 = \frac{Y_{harvested}}{0{,}72}$$

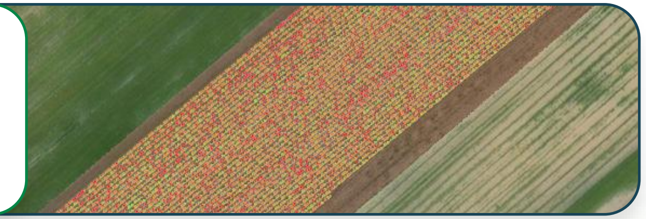
Microplot quality assessment

Emergence issues evaluation at the plot level – from Drone imagery



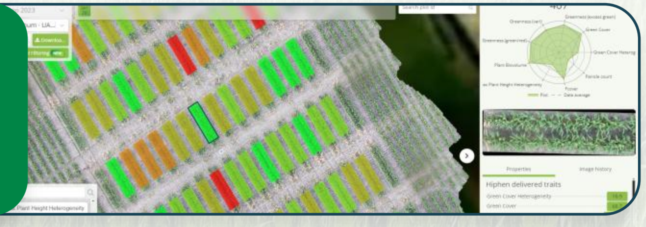
The different use cases in the industry

Use case 1:
Field phenotyping in nurseries



Use case 2:
Field phenotyping for yield correction

Use case 3: Digitizing traits used in regulatory trials for crop protection



Use case 4 and perspective: towards PHENOMIC PREDICTION

How to assess the accuracy & reliability of digital technologies for plant phenotyping?

Manual Assessments

Tedious, subjective, No data interoperability

Official breeding trials

Digital Literal® Assessments

Consistent, Repeatable, Precise Assessments

Validation needed for official breeding trials

DOI: 10.1111/epp.13037

EPPO STANDARD ON EFFICACY EVALUATION OF PLANT PROTECTION PRODUCTS

PP 1/333 (1) Adoption of digital technology for data generation for the efficacy evaluation of plant protection products

Specific scope: This Standard describes the validation, verification, and calibration of digital technologies that may be used to assess the efficacy of plant protection products (PPP). Currently, efficacy data are collected through human observation or other documented meth-

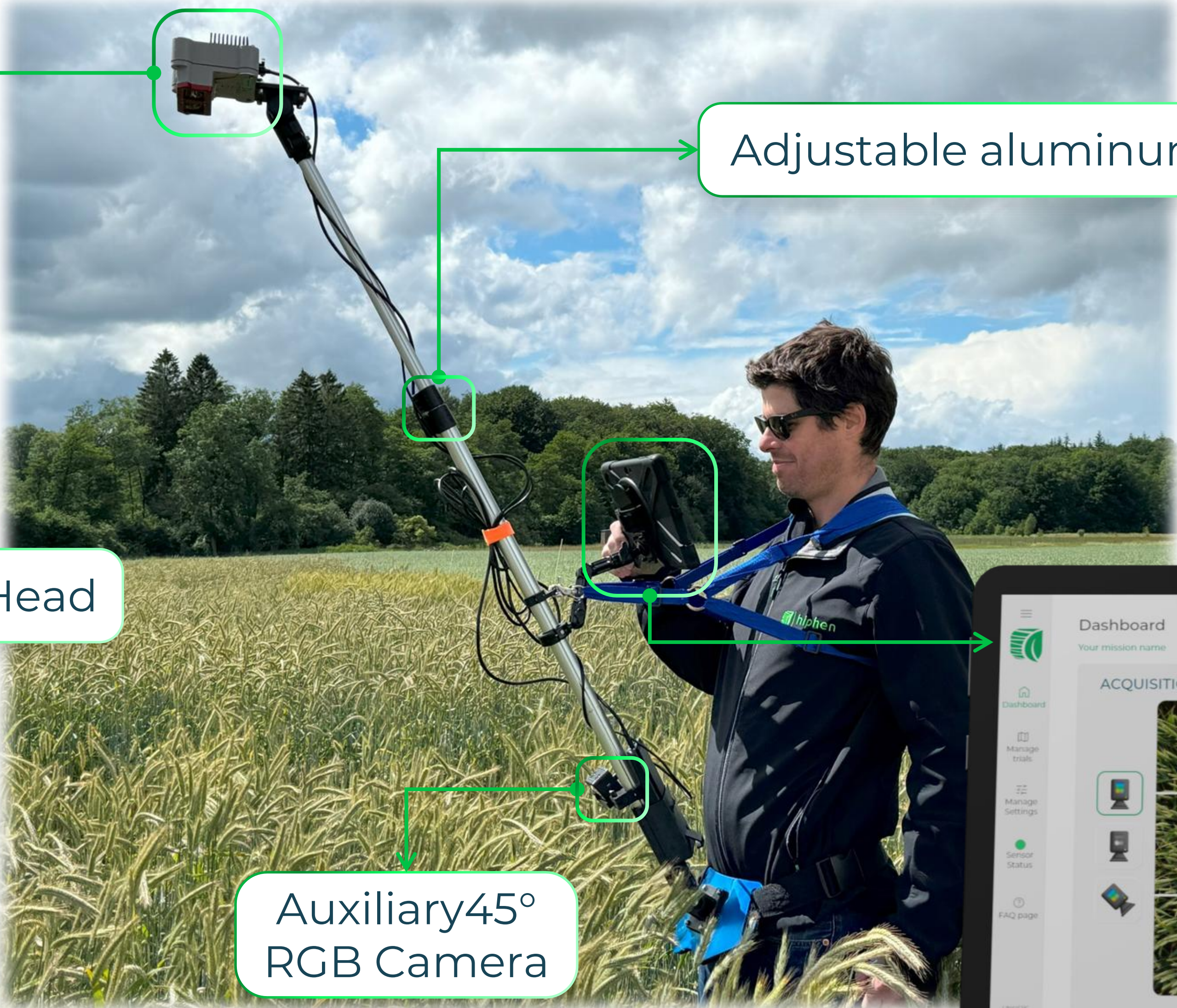
- Software: any algorithm approach (e.g. object recognition and classification, regression, segmentation, classical approaches, machine or deep learning as well as simple mathematical indexes such as the Normalized Difference Vegetation Index).

EPPO guidelines to adapt digital technologies for product testing

Literal® - High-resolution imaging device for small plot trials

TriCam Sensor Head

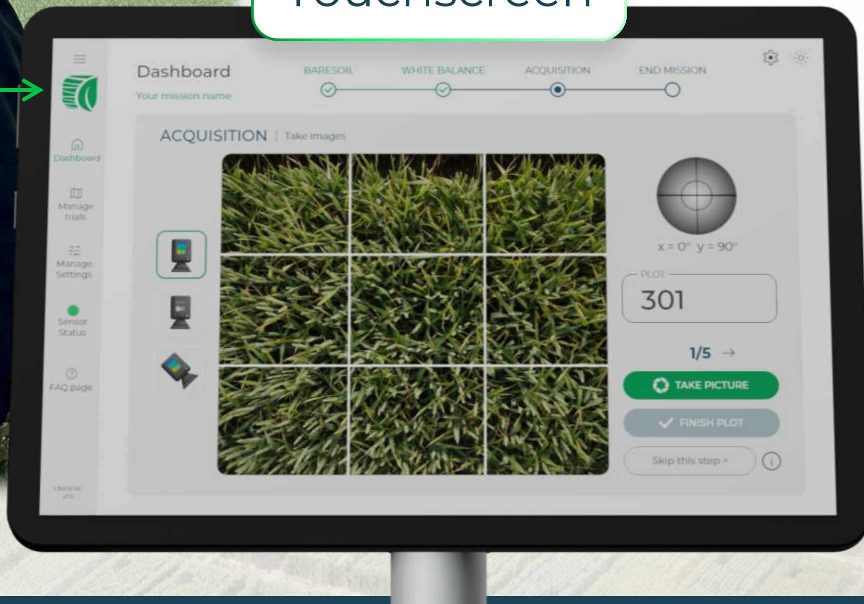
- **2 RGB Cameras** (8.3 MP resolution)
- **1 NIR Camera** (715 – 900 nm)



Adjustable aluminum stick

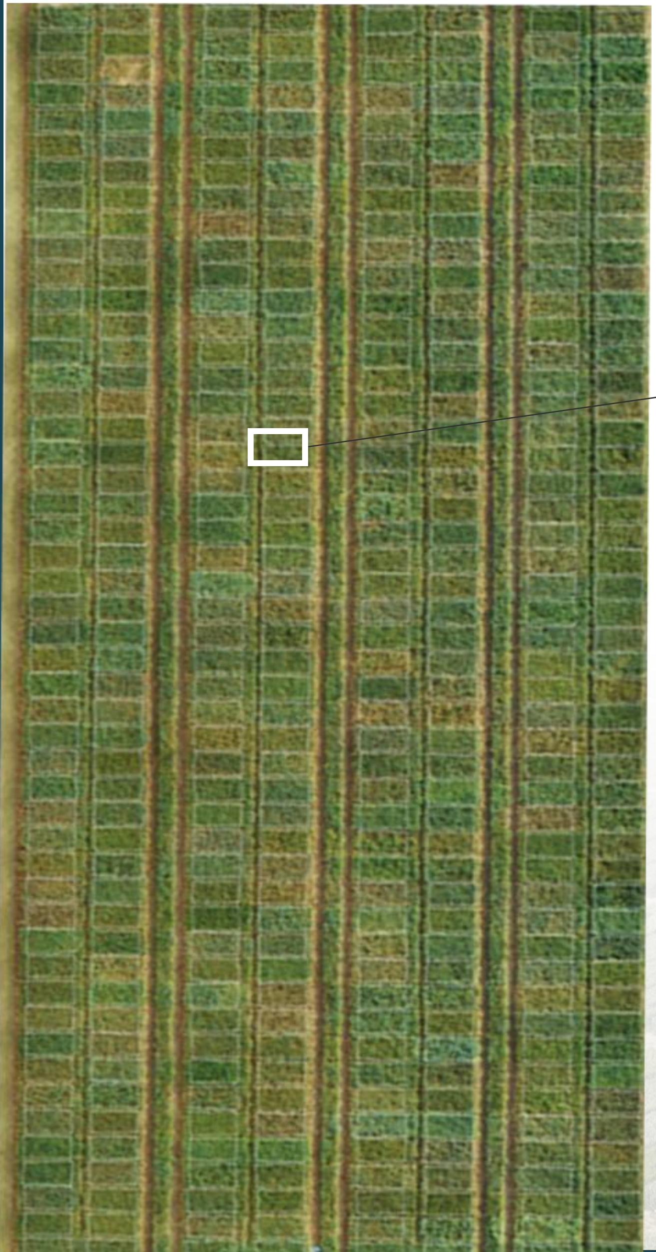
Auxiliary45° RGB Camera

Touchscreen



Designed by ARVALIS

Barely yellow dwarth virus standard method



Plot sampling

Intensity : 7
Area : 4

Intensity : 7
Area : 5

Intensity : 9
Area : 5

Intensity : 2
Area : 5

Typical trial design:

- 3 trial sites
- 200 genotypes
- 2 replicates
- 20 controls
- 2 m x 1 m plots

Visual scoring:

- Infestation rate: 1 (no symptoms) to 9
- Surface affected: 1 (limited area affected) to 5

Image analysis for Barley Yellow Dwarf Symptom Quantification

- Literal® image acquisition for a wide diversity of genotypes and controls
- Automatic analysis pipeline:

RGB Image

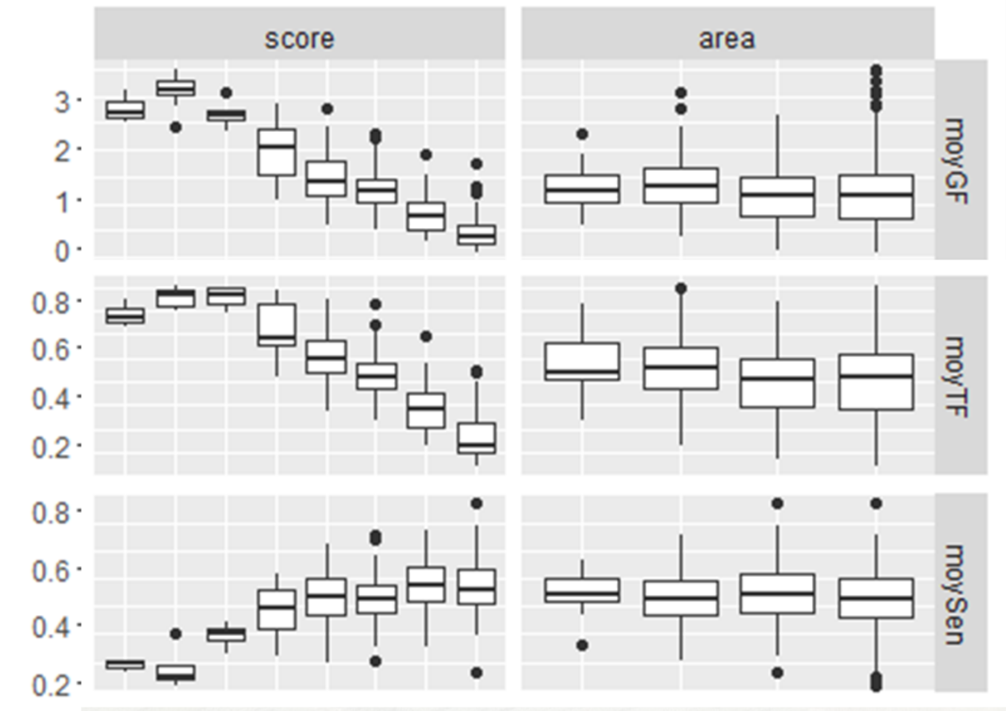
Segmentation into green / chlorotic / necrotic

Rating / patch

Comparing 420 plots (ca. 200 genotypes) with visual scores

Visual scores

LITERAL variables



CanopyGreen_CoverFraction (moyGF): Fraction of soil covered by green vegetation. It is the product of the two previous variables.

CanopyTotal_CoverFraction (moyTF): Total canopy fraction of the vegetation

CanopyTotal_SenescentFraction (moySen): Senescence fraction of the vegetation

- Good agreement between the green or the total canopy cover and the score of intensity
- The senescence fraction brings additional information
- **Heritability LITERAL (H^2 = 0.75) vs Heritability visual rating (H^2 = 0.65)**

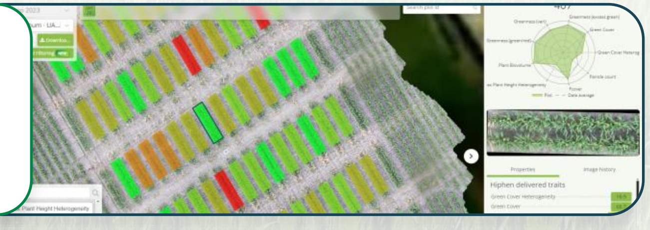
hiphen

The different use cases in the industry

Use case 1:
Field phenotyping in nurseries

Use case 2:
Field phenotyping for yield correction

Use case 3: Digitizing traits used in regulatory trials for crop protection



Use case 4 and perspective: towards PHENOMIC PREDICTION

Hunting for predictions through classical phenotyping or phenomic prediction

$$P = GxExM$$

Sensor

Trait Phenotyping

- Based on direct interpretation of the sensor signal
- Validated to ground truth
- We try to capture what is directly quantifiable by the signal

Phenomic Prediction

- Based on a training dataset independent of the sensor acquired dataset
- We try to capture the heritable part of the trait estimation
- We try to **remove the E x M component**

Expl => on 2024 data I predict the theoretical yield I would have had in 2023

Traditional – crop physiologist / remote sensing /academic approach

…
Generally, Modelling

Breeders Target

…
Generally, Diversity Mapping

New more breeder-oriented approach

Phenomic Phenotyping started with NIRs

G3 Genes | Genomes | Genetics

Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar

Renaud Rincent,* Jean-Paul Charpentier,†,‡ Patricia Faivre-Rampant,§ Etienne Paux,* Jacques Le Gouis,* Catherine Bastien,† and Vincent Segura†,1
*GDEC, INRA, UCA, 63000 Clermont-Ferrand, France, †BioForA, INRA, ONF, 45075 Orléans, France, ‡GenoBois analytical platform, INRA, 45075 Orléans, France, and §EPGV, INRA, CEA-IG/CNG, 91057 Evry, France

The concept of Phenomic selection

Scenario A

Scenario B

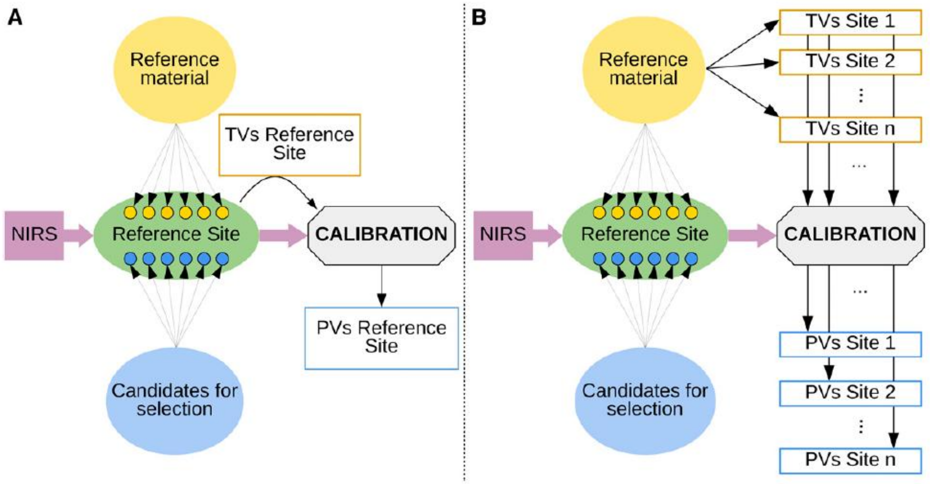
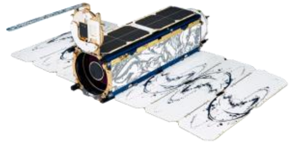


Figure 1 Schematic representation of the concept of phenomic selection, including the two scenarios tested in the present work: (a) S1, where the calibration model is trained with true values (TVs) and NIRS data collected at the same - reference - site and (b) S2, where the calibration model is trained with NIRS data collected at the reference site and TVs from other(s) environment(s). In both scenarios, the outcome of the prediction consists of predicted values (PVs).

True values (TVs) → Predicted values (PVs)

The 6 dimensions of Phenotyping



Satellite
5,000 to 100,000 plots

Largest footprint for production fields

Drones
1,000 to 50,000 plots

Standard traits at scale

Rovers (PhenoMobile)
1,000 to 10,000 plots

Enhanced precision at scale

Handheld (Literal)
30 to 1,000 plots

High-res plant features segmentation

The 6 dimensions of phenotyping

$M\ (measurements)$

x	y	z	λ	θ	t
Scanner	Imagery	3D imagery	Wavelength	Direction	Time

PhenoMobile by hiphen

Literal by hiphen

PhenoStation by hiphen

PhenoScale by hiphen

PhenoLite by hiphen

How to synthetize the raw data, 2 approaches are possible

Signal / Trait Approach

2018

frontiers in Plant Science
ORIGINAL RESEARCH
published: 06 June 2019
doi: 10.3389/fpls.2019.00685

A High-Throughput Model-Assisted Method for Phenotyping Maize Green Leaf Area Index Dynamics Using Unmanned Aerial Vehicle Imagery

Justin Blancon[1], Dan Dutartre[2], Marie-Hélène Tixier[1], Marie Weiss[3], Alexis Comar[2], Sébastien Praud[1] and Frédéric Baret[3]*

[1] Biogemma, Centre de Recherche de Chappes, Chappes, France, [2] HIPHEN SAS, Avignon, France, [3] INRA UMR 114 EMMAH, UMT CAPTE, Domaine Saint-Paul, Avignon, France

(…)

2023

OXFORD G3 Genes | Genomes | Genetics
G3, 2023, 13(1), jkac294
https://doi.org/10.1093/g3journal/jkac294
Advance Access Publication Date: 29 November 2022
Investigation

Temporal phenomic predictions from unoccupied aerial systems can outperform genomic predictions

Alper Adak[1], Seth C. Murray[1,*], Steven L. Anderson[2]
[1]Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843-2474, USA
[2]Syngenta, Naples, FL, 34114, USA
*Corresponding author: Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843-2474, USA. Email: sethmurray@tamu.edu

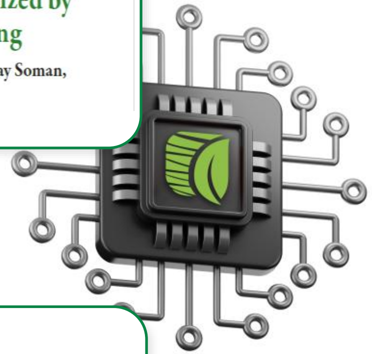
Latent space Approach

2018

The Plant Phenome Journal
Original Research

In-Field Whole-Plant Maize Architecture Characterized by Subcanopy Rovers and Latent Space Phenotyping

Joseph L. Gage,* Elliot Richards, Nicholas Lepak, Nicholas Kaczmar, Chinmay Soman, Girish Chowdhary, Michael A. Gore, and Edward S. Buckler



(…)

2023

Plant Phenomics
A SCIENCE PARTNER JOURNAL
RESEARCH ARTICLE

GenoDrawing: An Autoencoder Framework for Image Prediction from SNP Markers

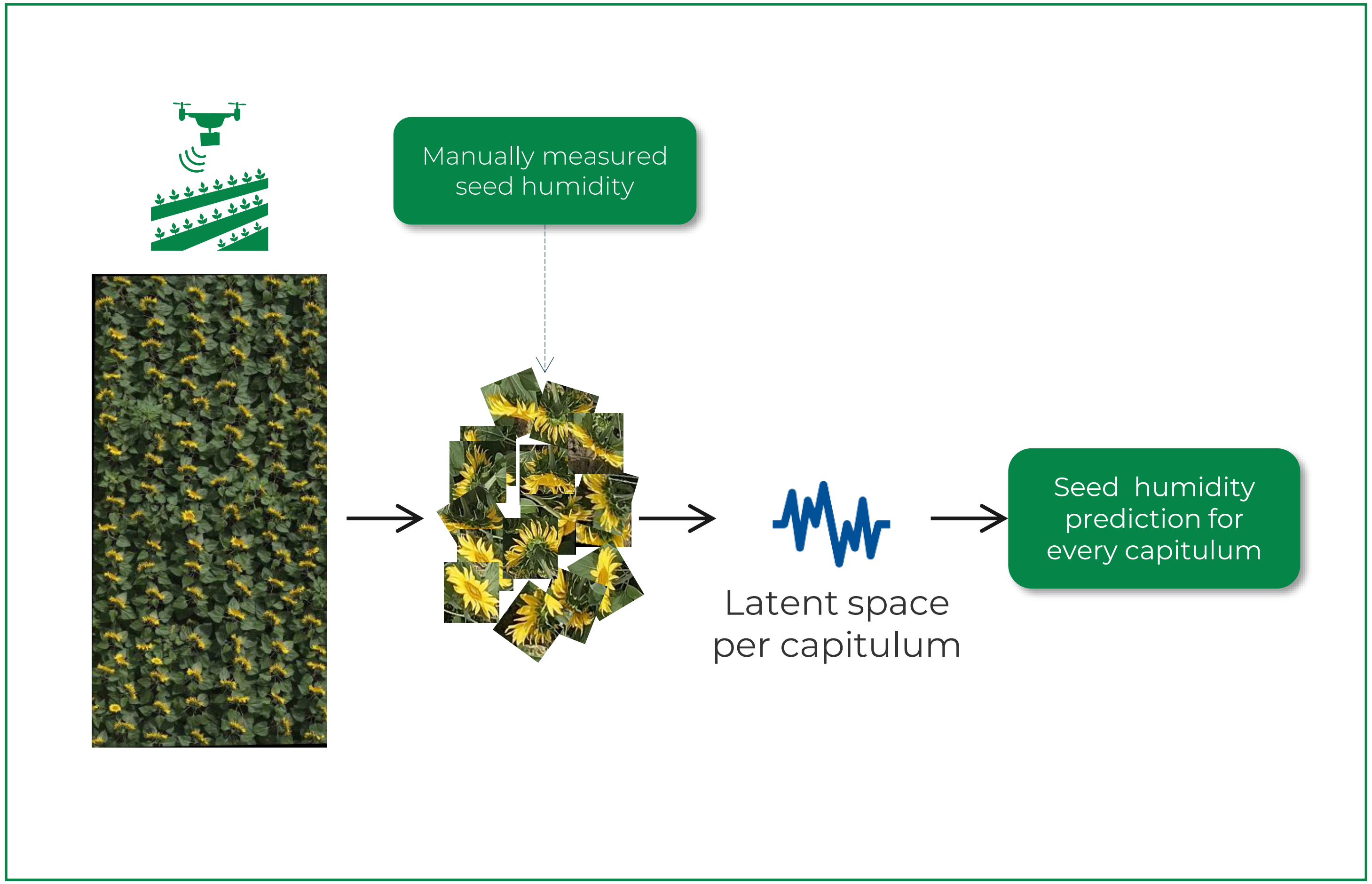
Federico Jurado-Ruiz[1], David Rousseau[2], Juan A. Botía[3], and Maria José Aranzana[1,4*]

[1]Center for Research in Agricultural Genomics (CRAG), 08193 Barcelona, Cerdanyola, Spain. [2]Université d'Angers, LARIS, INRAe UMR IRHS, 49000 Angers, France [3]Department of Information and Communication Engineering, University of Murcia, 30071 Murcia, Spain [4]IRTA (Institut de Recerca i Tecnologia Agroalimentàries), Barcelona, Spain

Sunflower

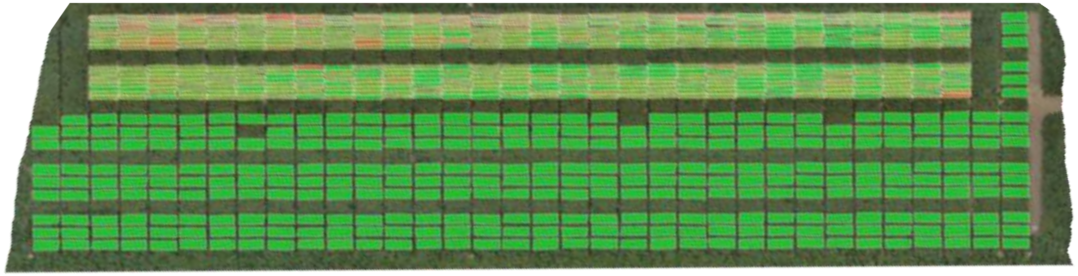
Use Case

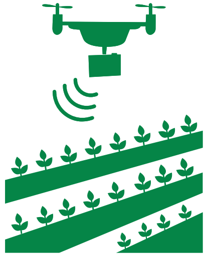
Seed humidity estimation at any date on any site



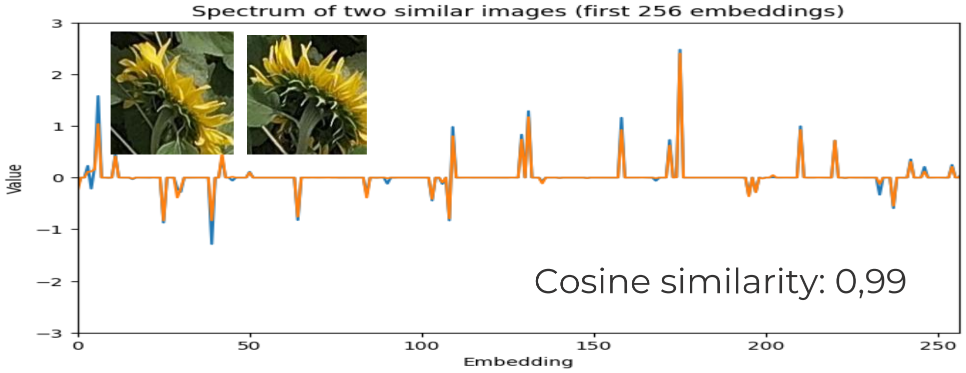
A glimpse of the pipeline

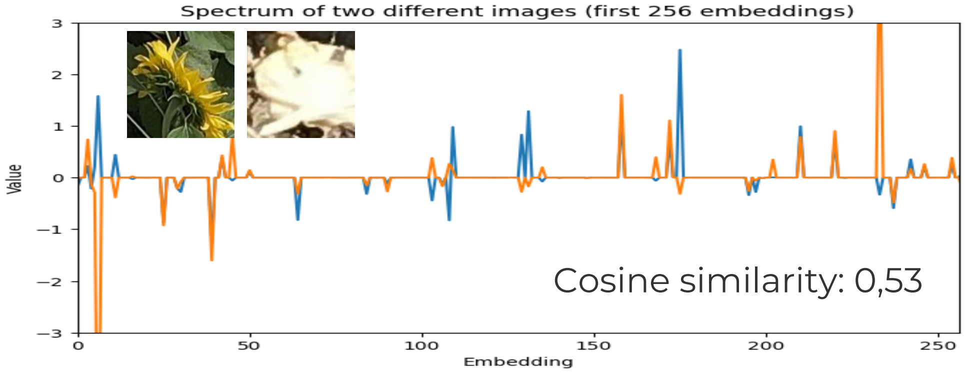
Orthomosaic and plot boundaries generation





Latent space





Plot image

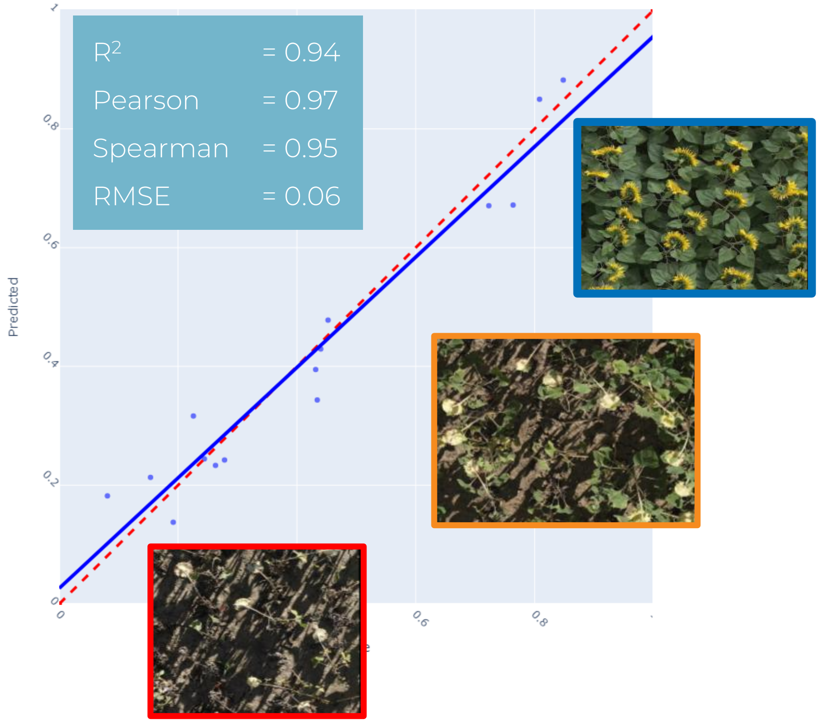
Organ detection

Organ pre-processing

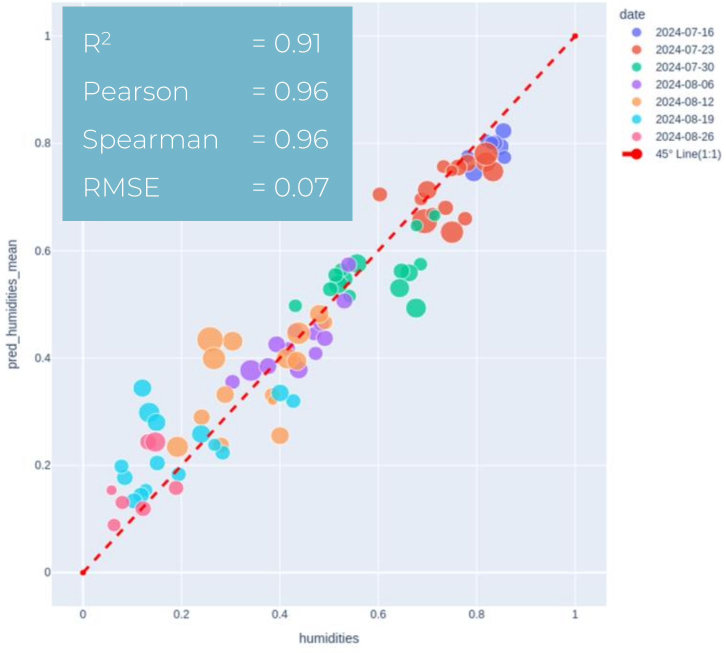
Seed humidity prediction

Model trained in France 2023

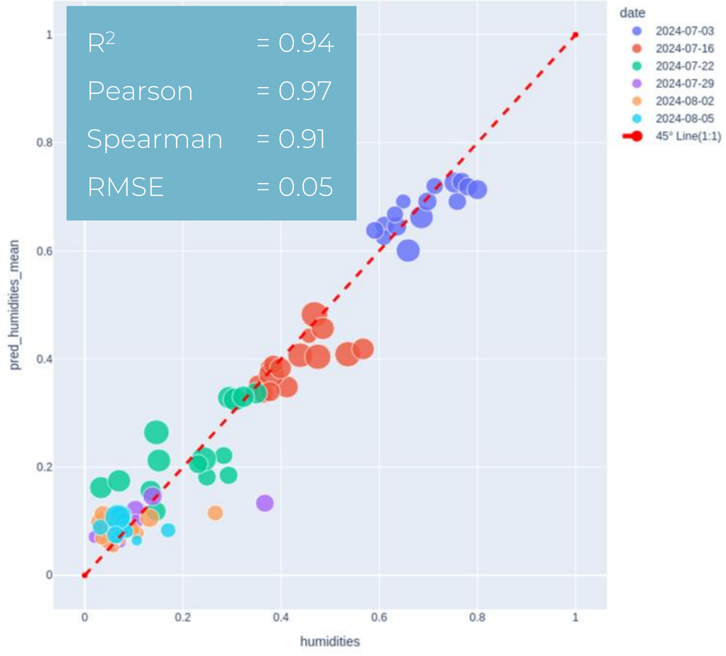
Prediction in France 2023



Prediction in France 2024



Prediction in Spain 2024



Every plant at every date has a seed humidity

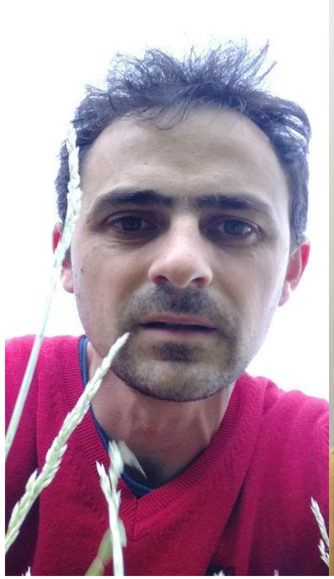
Corn
Use Case

Predictions of hybrides behavior based on parents measurements

With whom?

INRAE

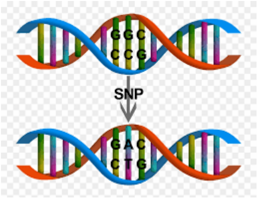
Génétique Quantitative et Évolution Le Moulon



Renaud Rincent **Laurence Moreau**

Which methods to compare?

Genomic Prediction



Phenomic Prediction

Trait

Latent space

The experimental design and objectives

Contract details

Location:
INRAE Moulon

Nb of trials	Nb of flights	Nb of plots
1	**6**	**210**

Acquisition

DJI M3M | RGB | Airphen | M3T* | Lidar*

Number of traits

33 / dates (~180)

The experiment

4 Flint founders

821 Flint lines

180 hybrids Incomplete factorial design

4 Dent founders

801 Dent lines

A. Lorenzi & Al (2022)

The pipeline

Target "Phenomic B value" → Sensor data → Processing tech → Prediction

- Target: Yield, DM content, FLOF
- Sensor data: Phenoscale, NIRS
- Processing tech: Image to traits, Latent Space

The traits on 6 dates (180 points)

8 Vegetation indices
- NDVI
- ....

5 PROSAIL estimation
- LAI
- Cab
- ....

3 SfM Algorithm
- height
- biovolume
- ....

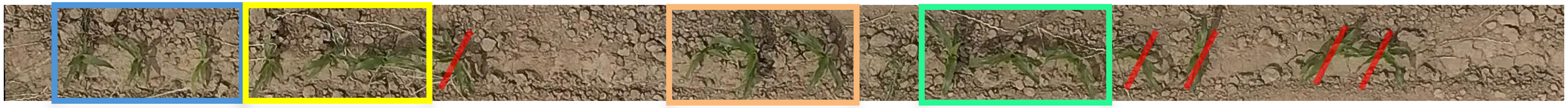
4 AI segmentation
- Green cover
- Greeness
- ....

2 AI counts
- Plant count
- Presence of panicules
- ....

Latent space Environmental effects removal:

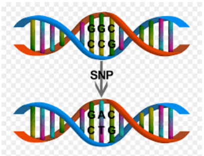
Background / soil removal

Selection of corn experimenting light competition



Plot_X3_Y13_A Plot_X3_Y13_B Plot_X3_Y13_C Plot_X3_Y13_D

The results so far



		Genomic Prediction	Phenomic Prediction		
Random cross validation	Yield	0.84	0.73	0.79	0.76
	DM Content	0.83	0.43	0.59	0.61
	FLOF	0.85	0.57	0.76	0.74
New Dent Flint	Yield	0.23	0.28	0.46	-0.05
	DM Content	0.08	-0.1	0.31	-0.18
	FLOF	0.06	-0.06	0.06	0.02

The next steps: taking even more the environnement into account

$$P = GxExM$$

- Remove more environemental effect
- Adapt the size of the latent space to the dataset

- Add dynamic traits
- Add agroclimatic traits
- Test the predictive power of each trait category
- Make easy UI to test different predictors

Disentangling Genotype and Environment Specific Latent Features for Improved Trait Prediction using a Compositional Autoencoder

Anirudha Powadi [1], Talukder Zaki Jubery, [2], Michael C. Tross, [3,5], James C. Schnable, [3,5,*] and Baskar Ganapathysubramanian, [1,2,4]

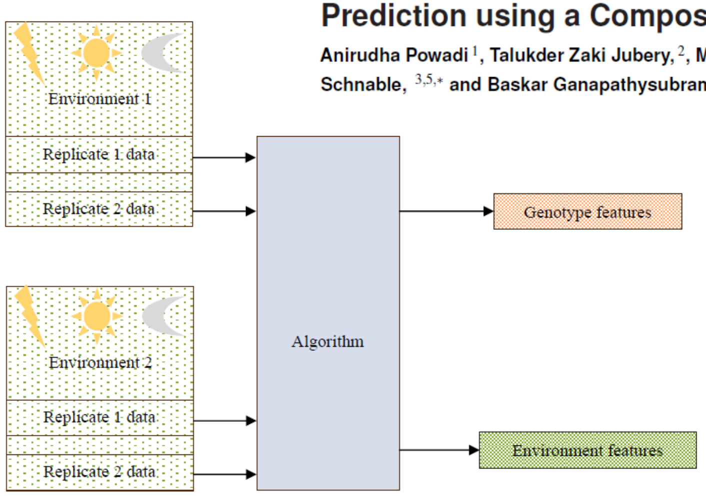
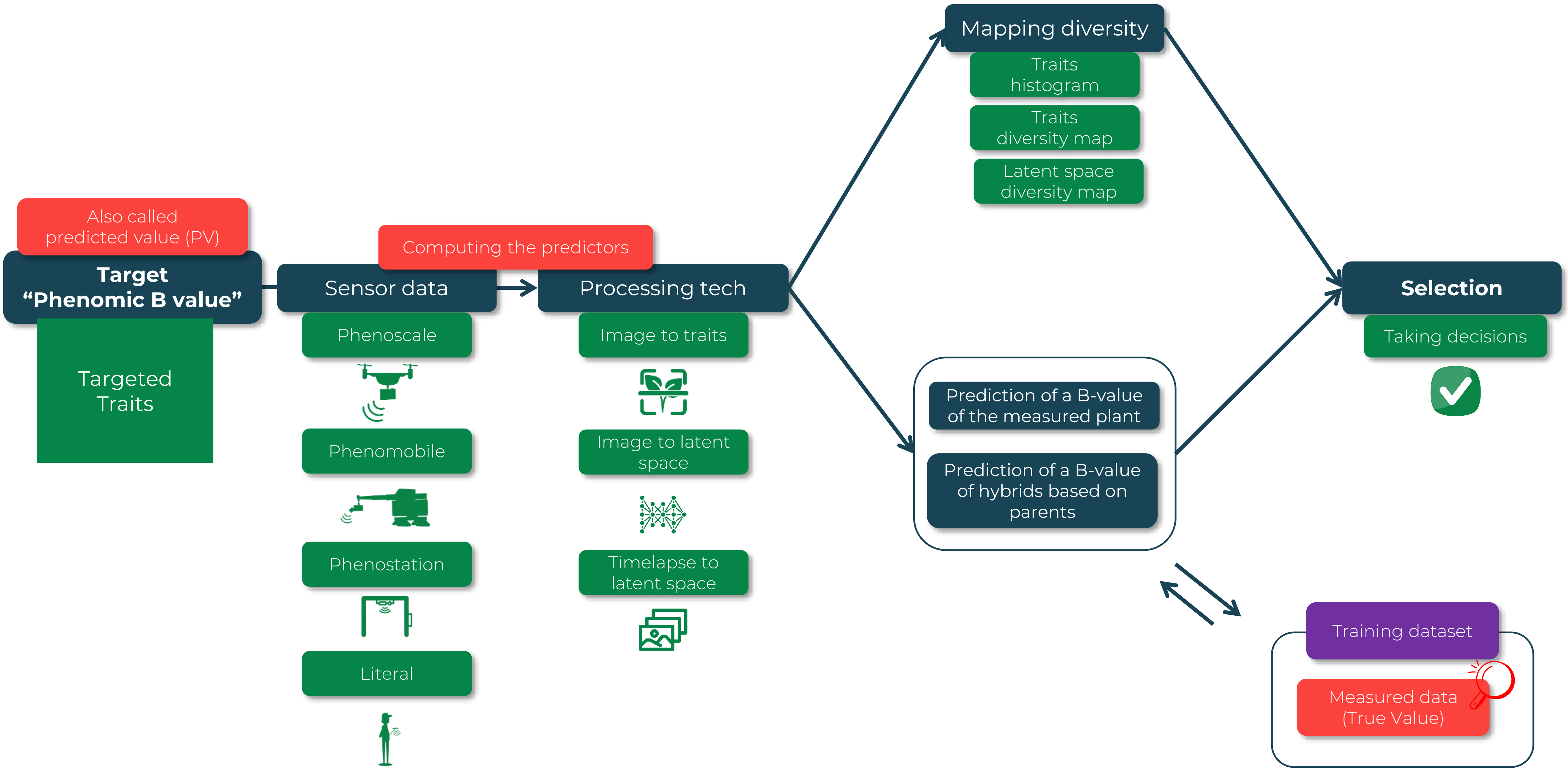
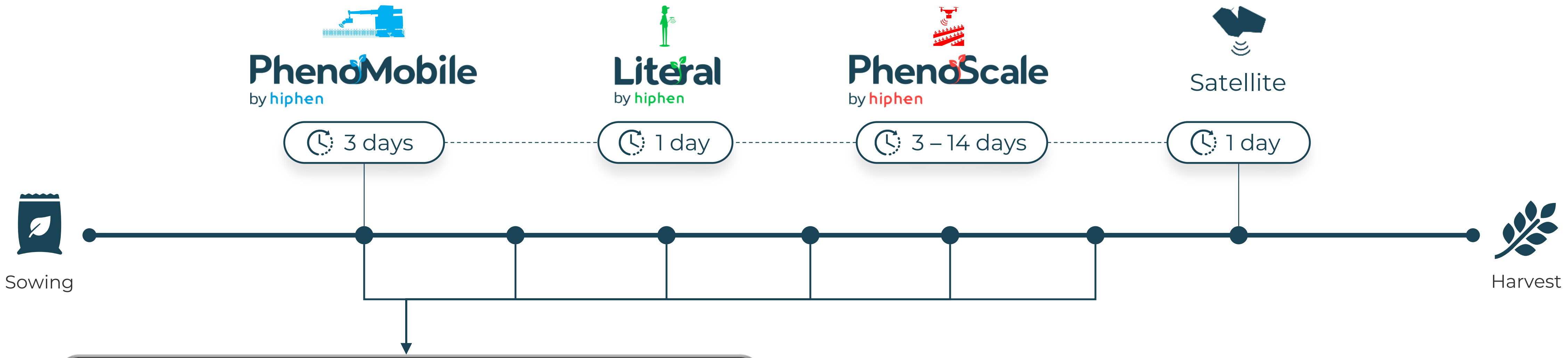


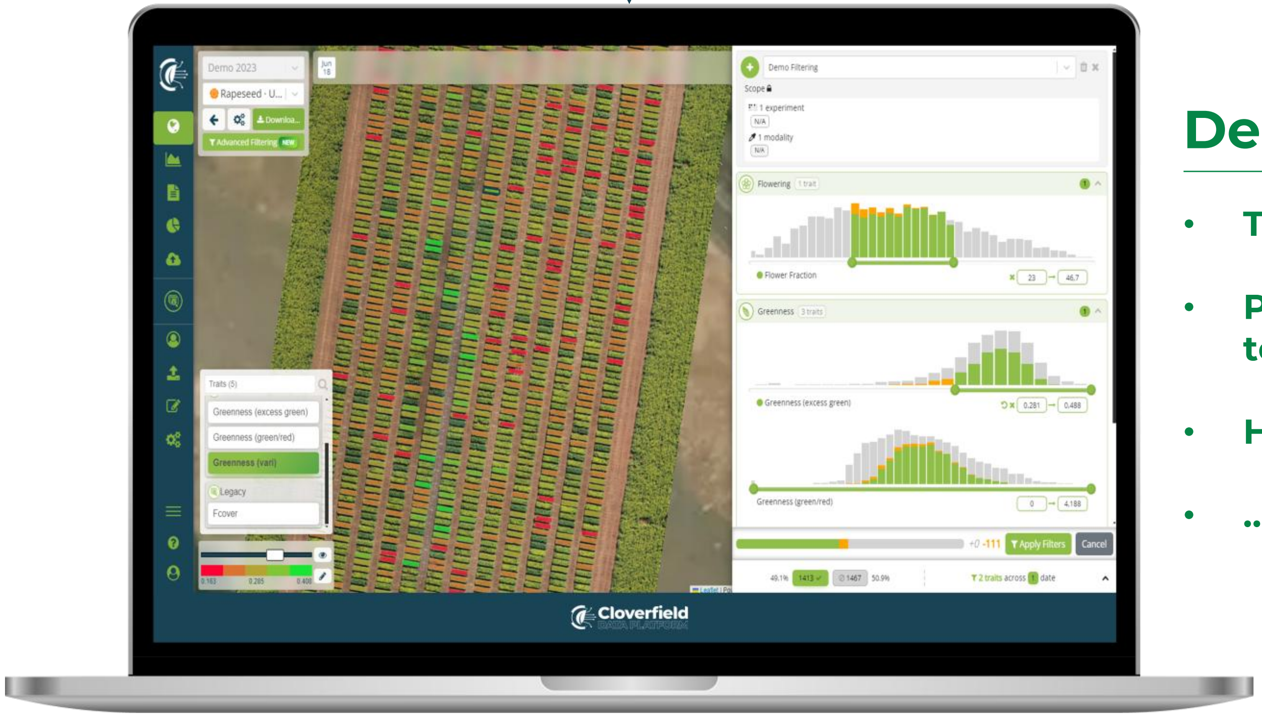
Figure 1. The problem definition: Extract and disentangle the effects of genotype and environment for a given type of sensor data, assuming multiple observations of each genotype in each environment. Our specific test dataset was a set of hyperspectral reflectance data collected from 578 distinct genotypes of maize in two distinct environments, with two replicates of each genotype in each environment (four total replicates per genotype and 1,156 total replicates per environment).

HIPHEN view on Phenomic prediction pipeline



On Time and Low-Cost Phenotyping





Decisions

- **Trial quality**
- **Plot assessment to select or not**
- **Harvest decisions**
- **...**

Outsourcing trait processing is cheaper and quicker than internalizing

Professional pipelines are arriving to maturity

Analytics (Phenomic Prediction) help making decisions

Hiphen

120 rue Jean Dausset

Agroparc – Bâtiment technicité

84140 Avignon

France

+33 (0)7 84 14 31 63

hiphen-plant.com | contact@hiphen-plant.com