HARVEST 2025 INTERNATIONAL CONFERENCE ON PARALLEL PROCESSING

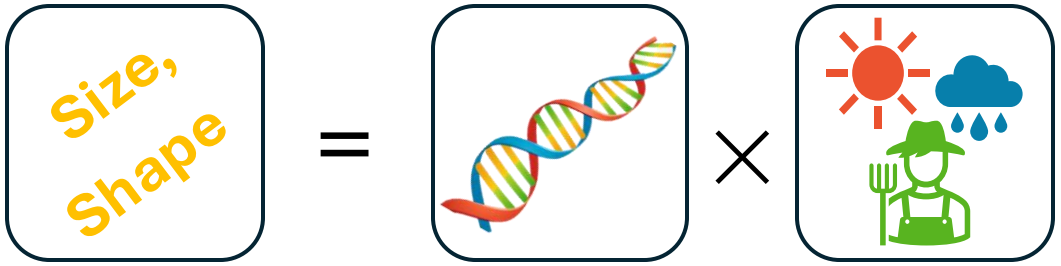
An HPC Framework for **Multi-Modal Plant Phenotyping** Integrating Controlled Environment & Open Field Studies

Sun **Oh**, Libo **Zhang**, Yang **Yang**, Mitch **Tuinstra**
Institute for Plant Sciences

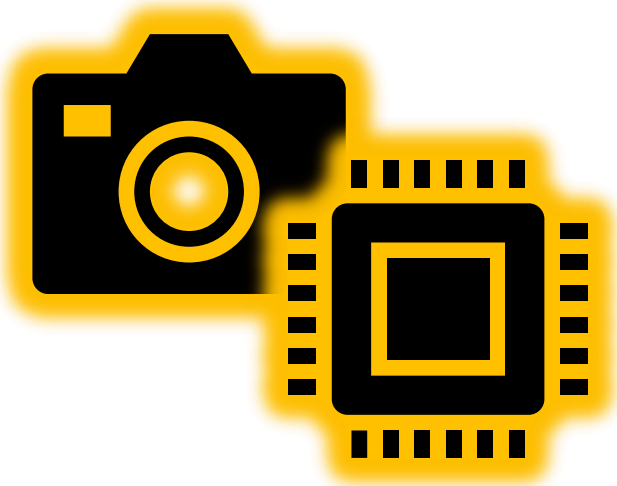
PURDUE UNIVERSITY

Plant Phenotyping

- Phenotype



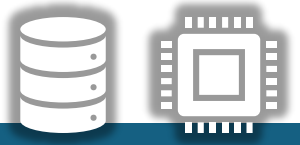
- Imaging sensors, high-throughput



Phenotyping Research

Grow plants

- Treatment



Plant phenotyping

- Data collection
- Measurements

Data analysis

- Modeling
- Interpretation

G	E x M	Msmt 1	Msmt 2	Msmt 3
A	Control	0.1	0.5	1.5
A	Treatment	0.2	0.5	1.8
B	Control	0.2	1.0	3.0
B	Treatment	0.4	1.1	3.6

Phenotyping Facilities: Open-field & Indoor

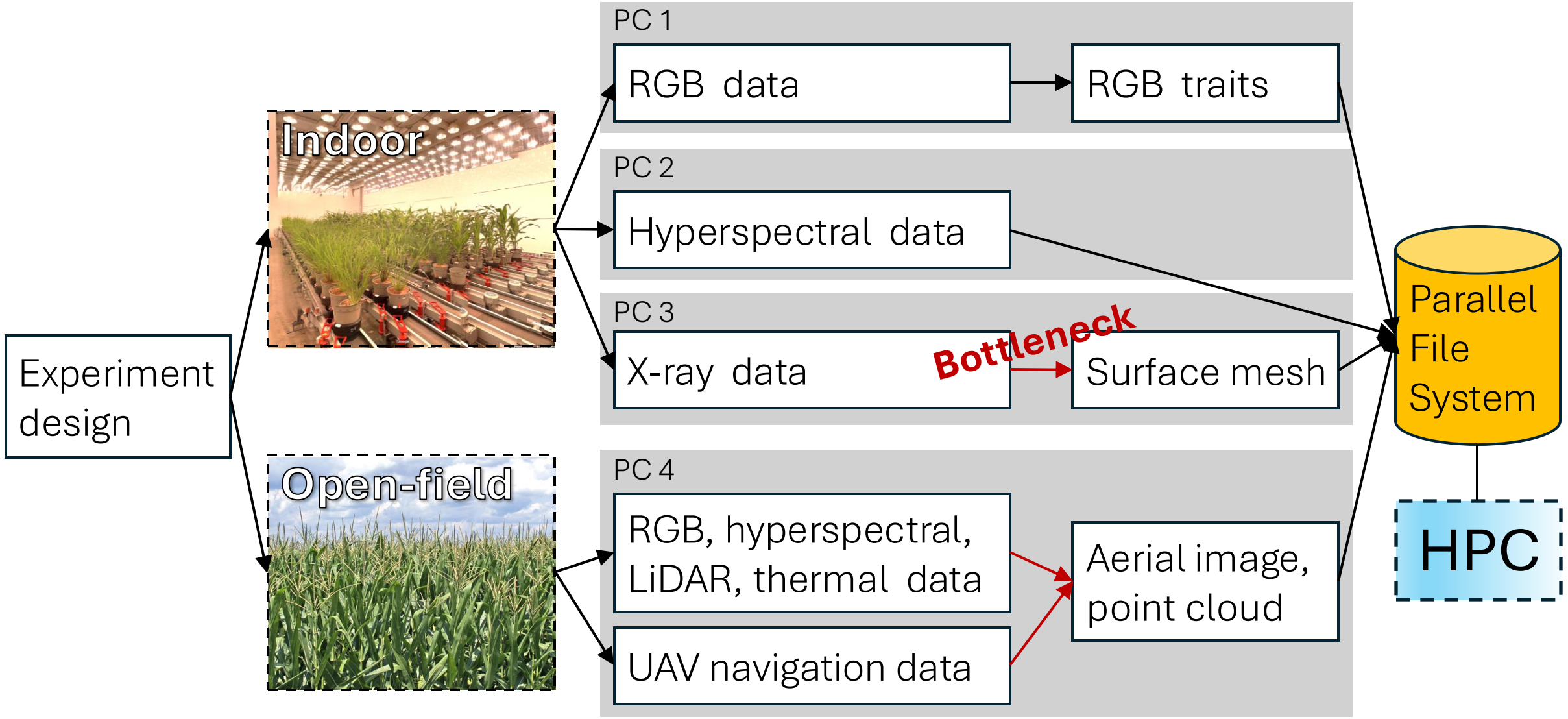
Experimental unit	Plot ✅	Pot
Environment	Real-world ✅	Simulated
Environment control	Unpredictable	Possible ✅
Time constraint	During the season	No ✅
Sensors	RGB, hyperspectral, LiDAR, thermal	RGB, hyperspectral, X-ray CT

Objectives

1. **Multi-modal** sensor data workflows

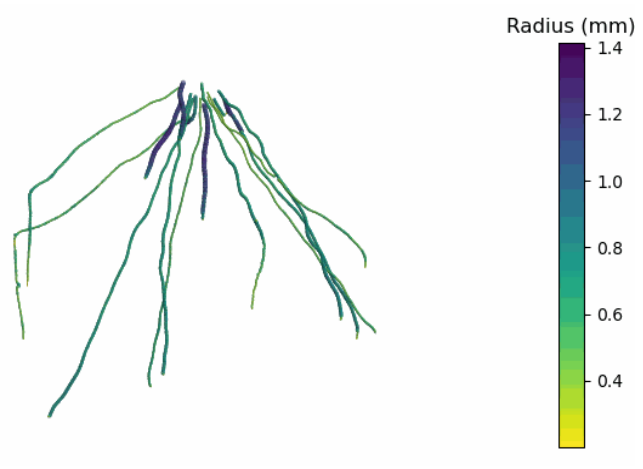
2. Node-level, thread-level **parallelization**

Data Workflows (1/2, PC-based)

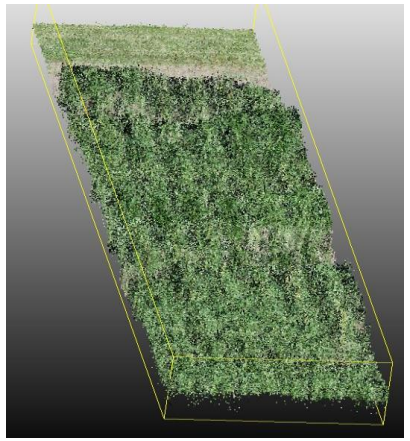


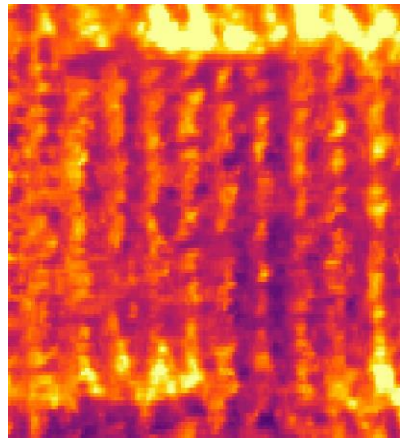
Input Data

Indoor

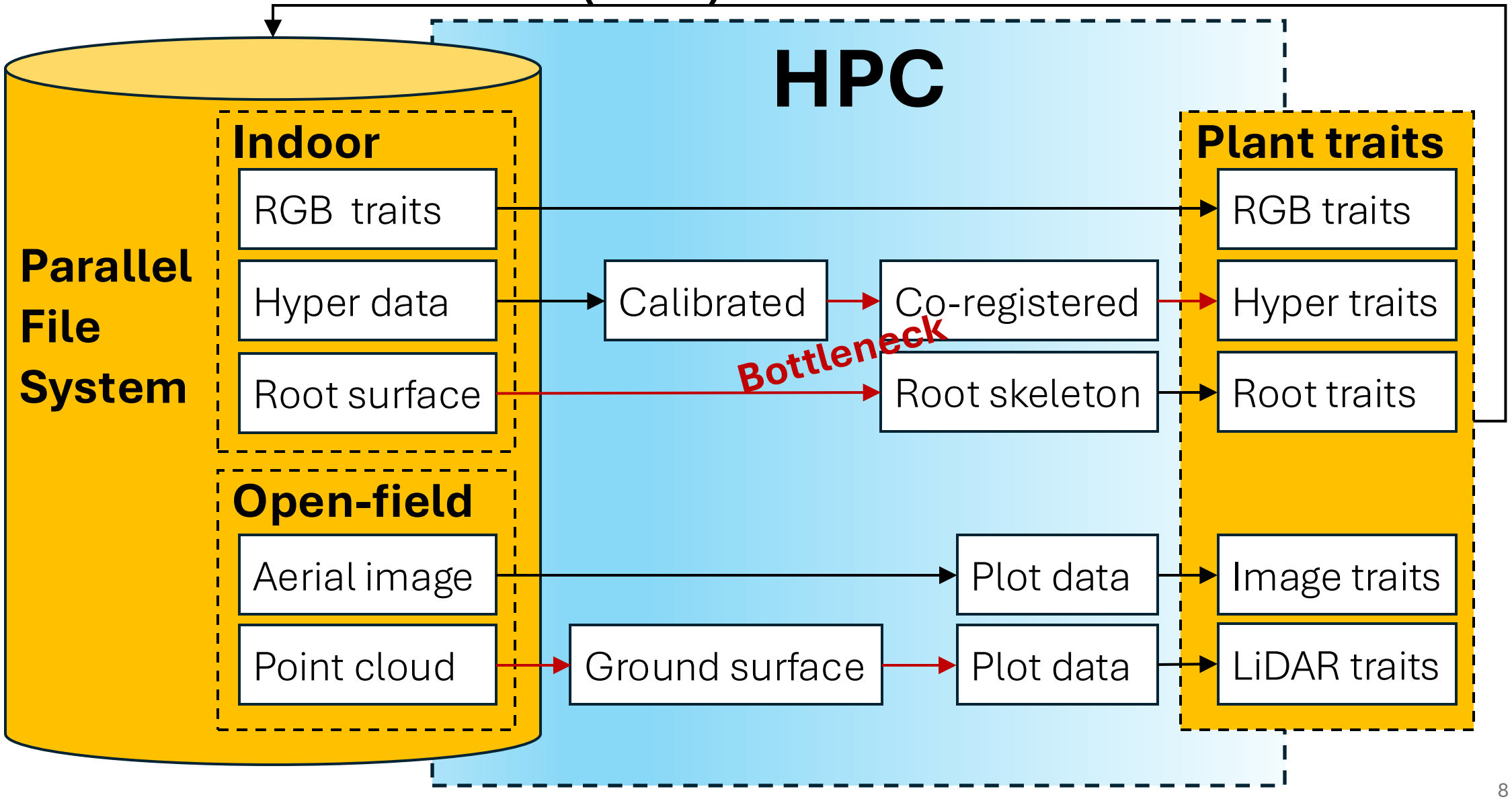


UAV





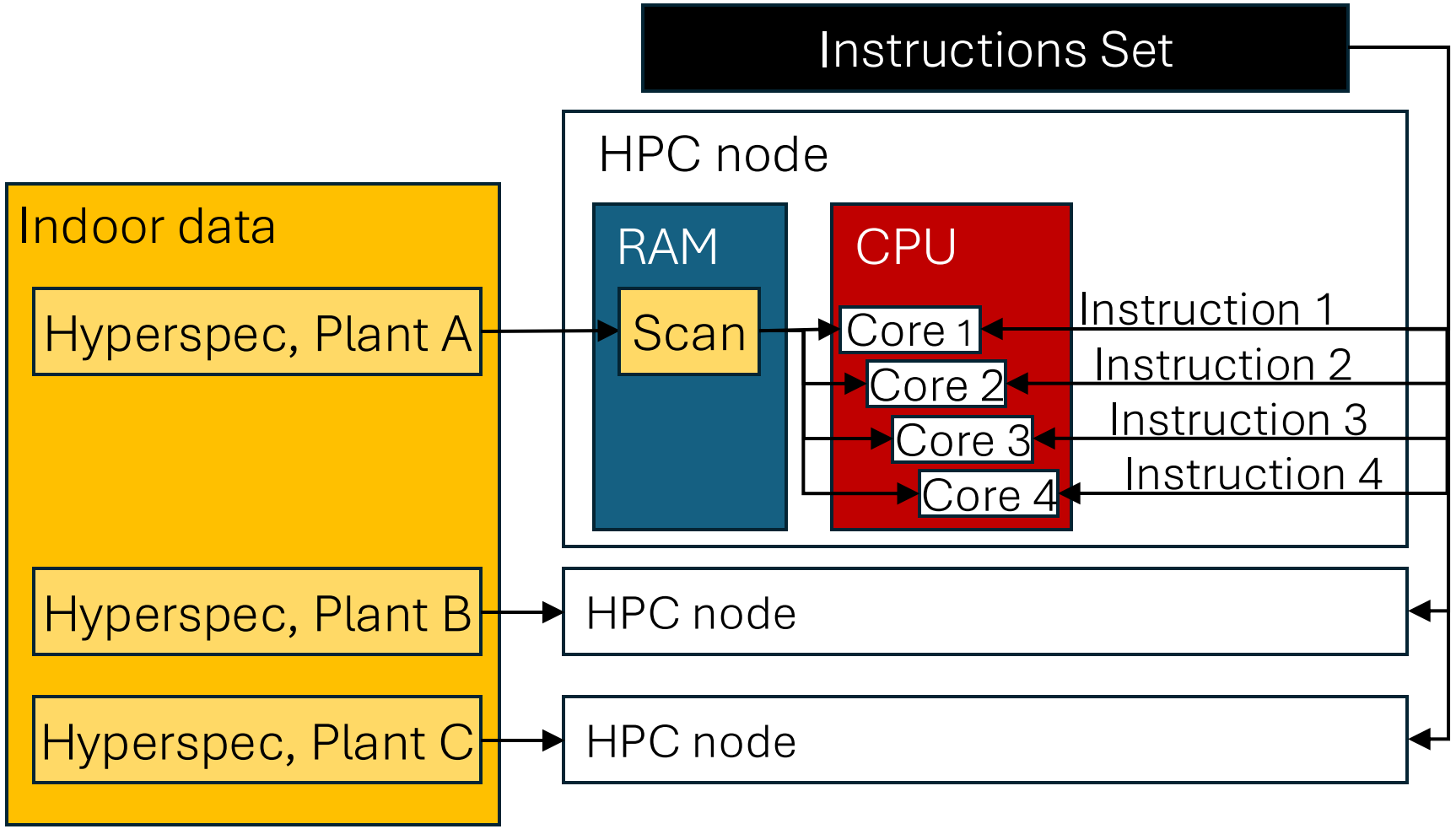
Data Workflows (2/2)



Processing Bottlenecks

Source	Sensor	Input data (GB)	RAM req. (GB)	Bottlenecks
Indoor	Hyperspectral	1-12	4-64	Band-wise transform, band algebra
Indoor	X-ray	~ 20	32-64	Graph optimization
Open-field	Aerial image	1-8	4-32	Band algebra
Open-field	Point cloud	~ 0.1	4-32	Nearest-neighbor search, Data cropping (w/ boundary)

Parallelization (Single Data Multi Instruction)



Instr 1: Get NDVI

$$\frac{Band_{750} - Band_{705}}{Band_{750} + Band_{705}}$$

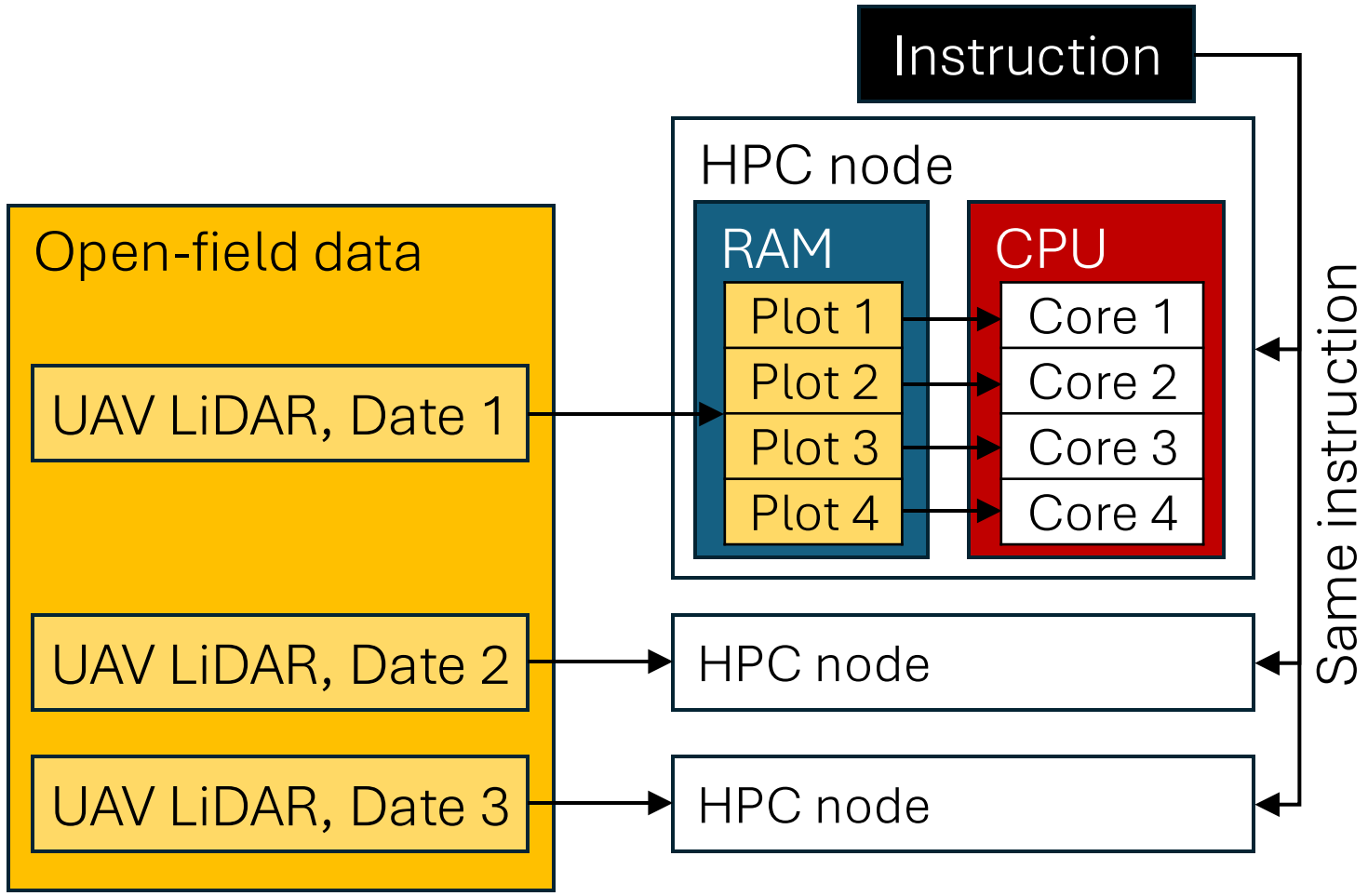
Instr 2: Get WBI

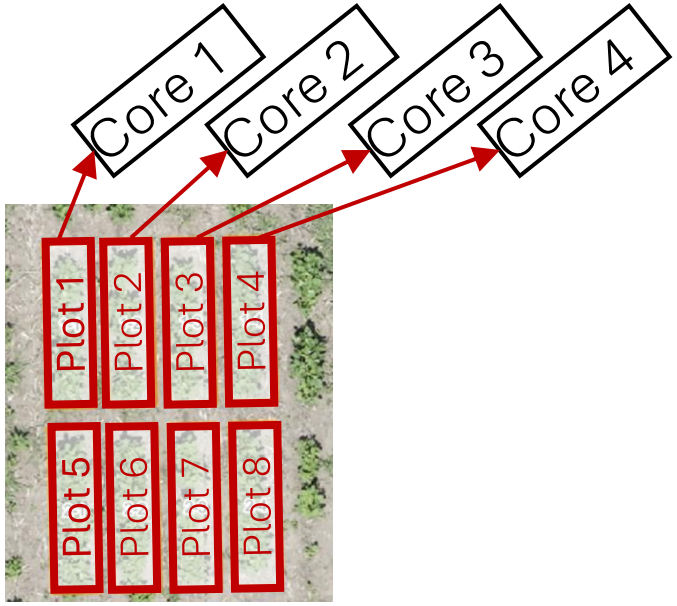
$$\frac{Band_{970}}{Band_{902}}$$

...

500 instructions

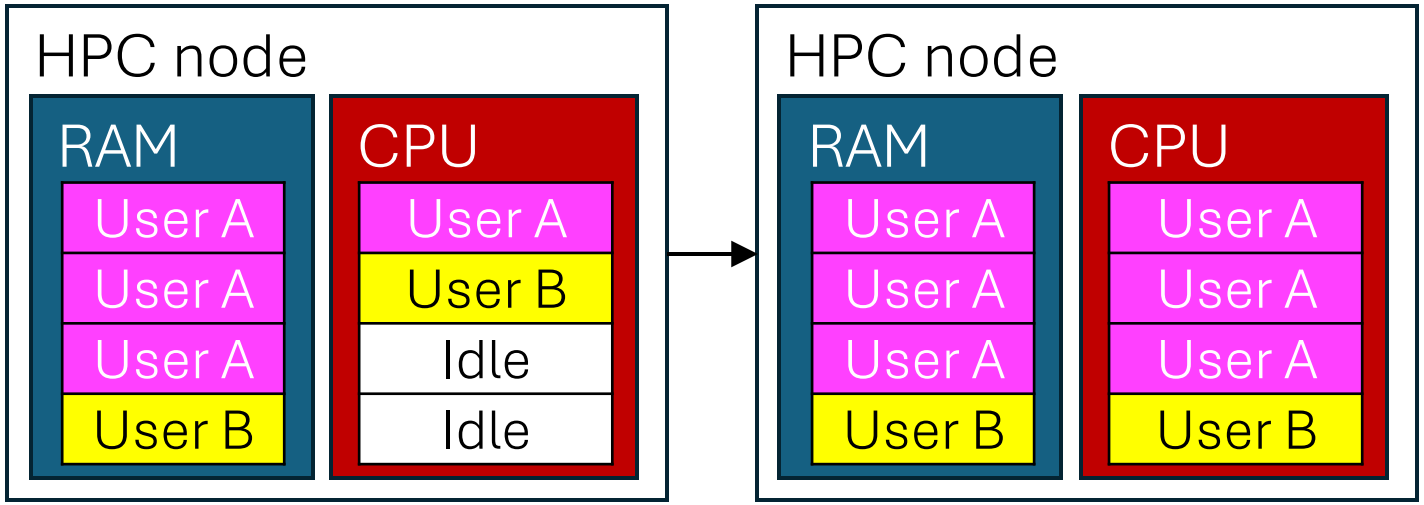
Parallelization (Multi Data Single Instruction)





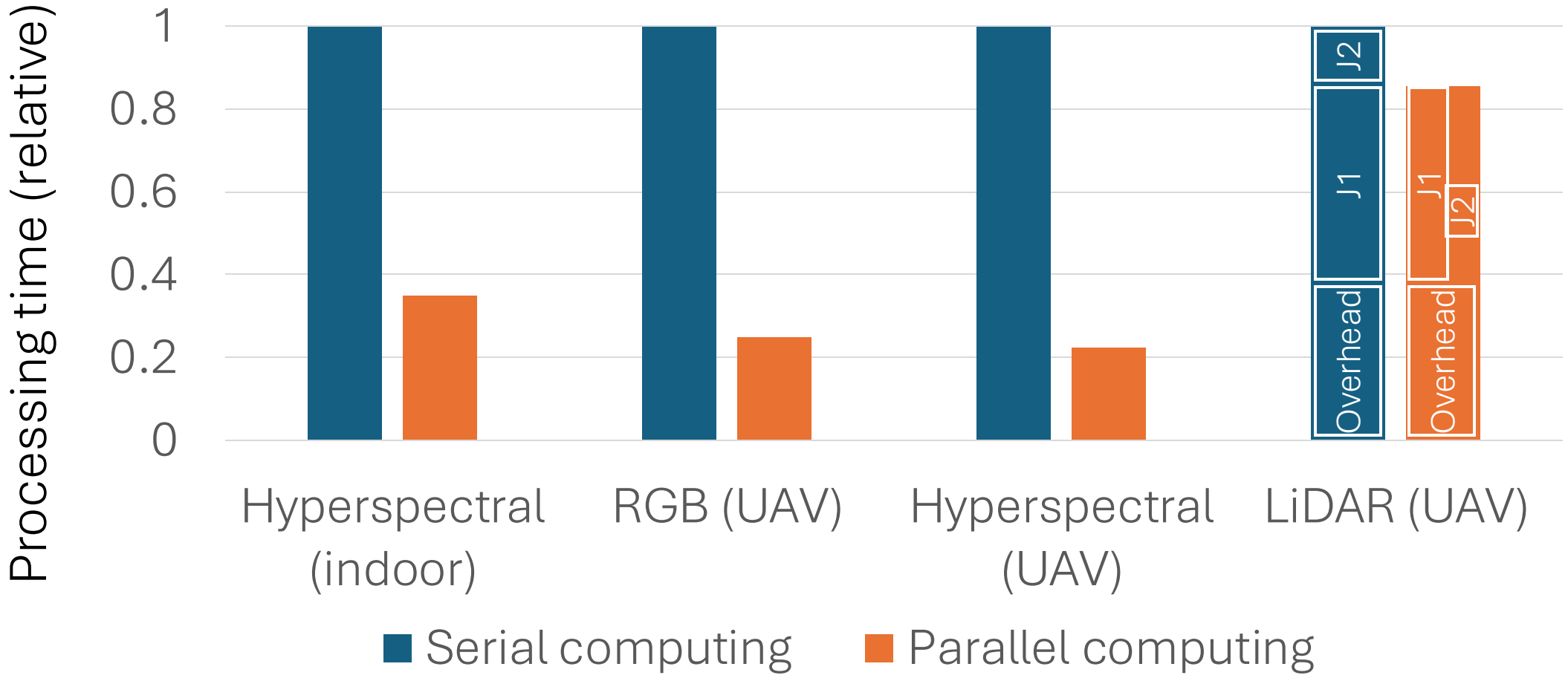
Constraints

- HPC usage rules
  - No. of cores
  - Automatic scaling



- Inherent mem inefficiency
  - Python array
  - Dependencies

Results


Processing time (relative)
1
0.8
0.6
0.4
0.2
0
Hyperspectral (indoor)
RGB (UAV)
Hyperspectral (UAV)
LiDAR (UAV)
J2
J1
Overhead
J1
J2
Overhead
Serial computing
Parallel computing

Conclusions

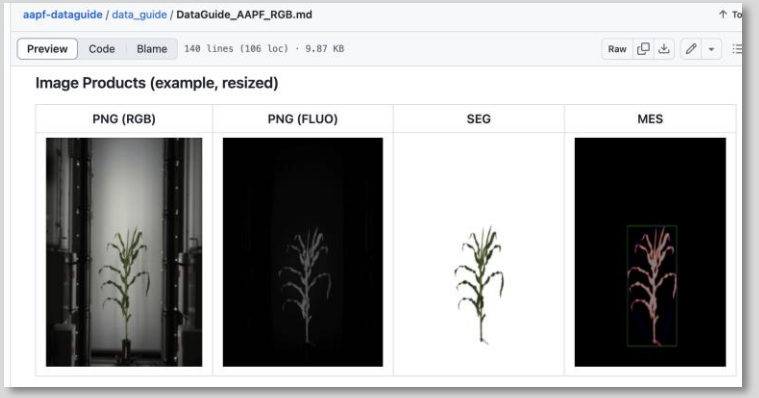
- Comprehensive workflow for HTP, integrating indoor and UAV sensors
- HPC node- and thread-level parallelization
- Some data pipelines still limited by licensed SW and OS
- Our long-term goal is to extend the processing limits and work directly with raw data

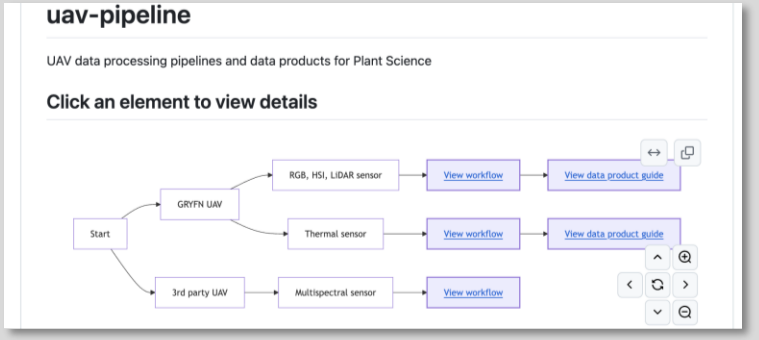
HARVEST 2025 INTERNATIONAL CONFERENCE ON PARALLEL PROCESSING

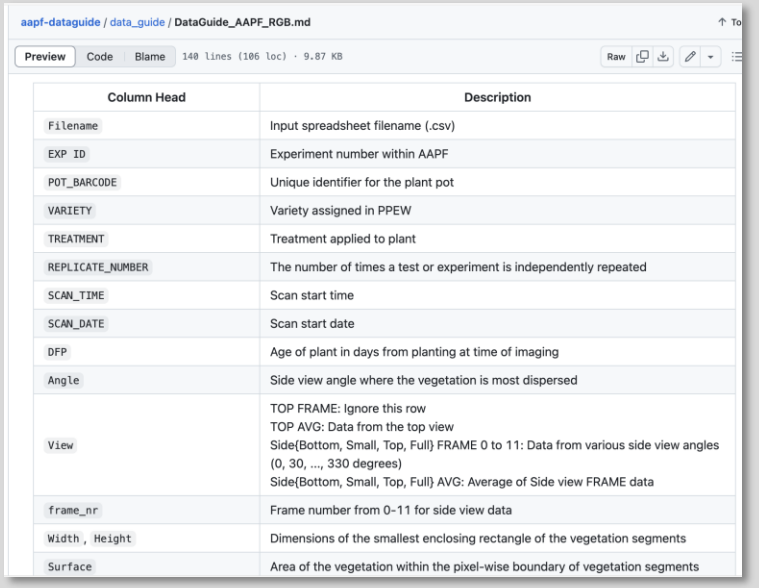
Thank you

github.itap.purdue.edu/plantscience







Column Head	Description
Filename	Input spreadsheet filename (.csv)
EXP ID	Experiment number within AAPF
POT_BARCODE	Unique identifier for the plant pot
VARIETY	Variety assigned in PPEW
TREATMENT	Treatment applied to plant
REPLICATE_NUMBER	The number of times a test or experiment is independently repeated
SCAN_TIME	Scan start time
SCAN_DATE	Scan start date
DFP	Age of plant in days from planting at time of imaging
Angle	Side view angle where the vegetation is most dispersed
View	TOP FRAME: Ignore this row TOP AVG: Data from the top view Side{Bottom, Small, Top, Full} FRAME 0 to 11: Data from various side view angles (0, 30, ..., 330 degrees) Side{Bottom, Small, Top, Full} AVG: Average of Side view FRAME data
frame_nr	Frame number from 0-11 for side view data
Width, Height	Dimensions of the smallest enclosing rectangle of the vegetation segments
Surface	Area of the vegetation within the pixel-wise boundary of vegetation segments