Knowledge Discovery by Constructing AgriBigData

Masayuki HIRAFUJI

hirafuji@evolution-lab.com

Wei GUO

guowei@isas.a.u-tokyo.ac.jp

International Field Phenomics Research laboratory
The University of Tokyo

Life-span is becoming longer in all countries.
Can we enjoy delicious meals for long-life, e.g. 100 years old?

Four types of global crop yield trends

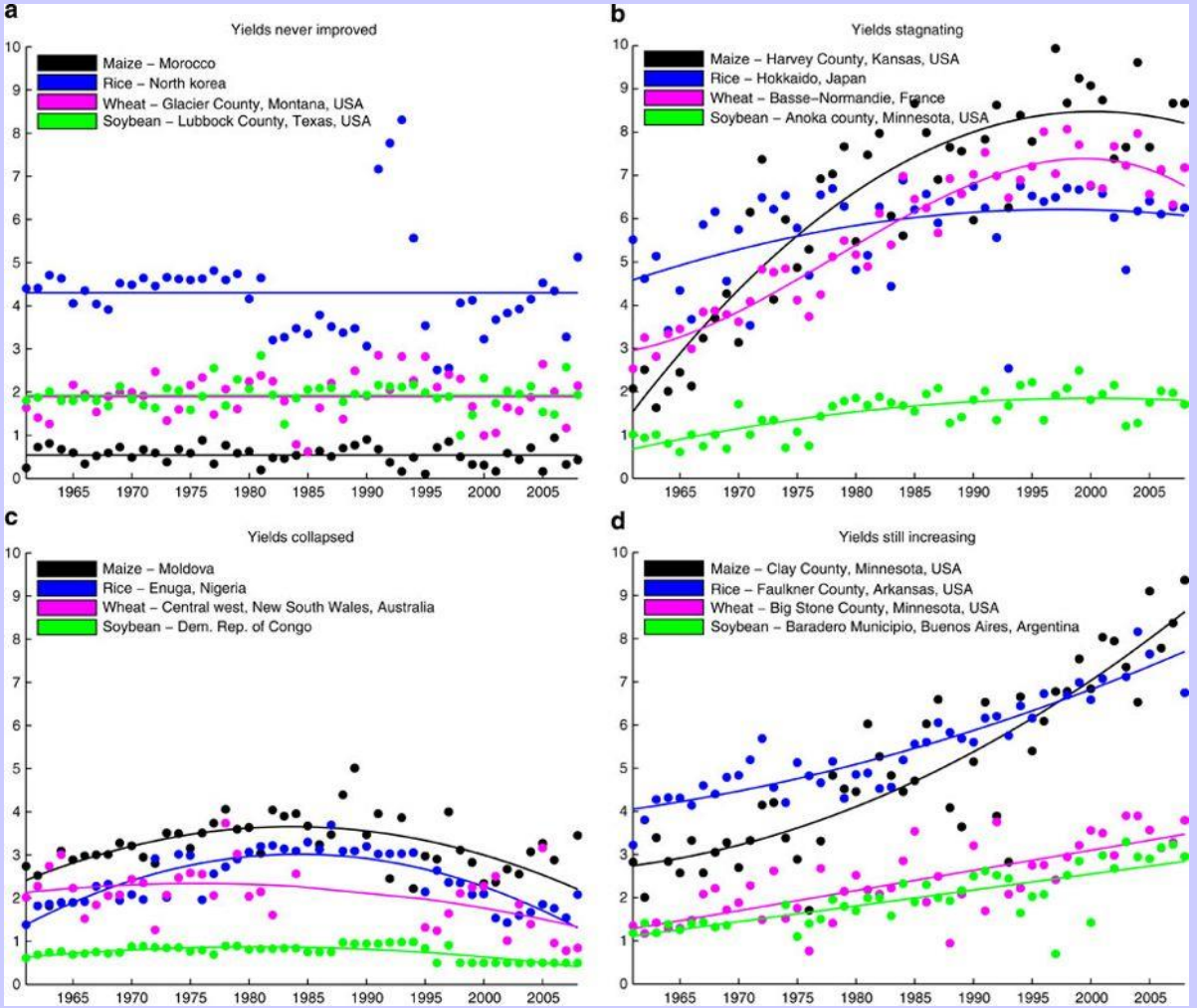
Yield trends are different for crops, cultivars, climates, management methods etc.

Never improved

Stagnating

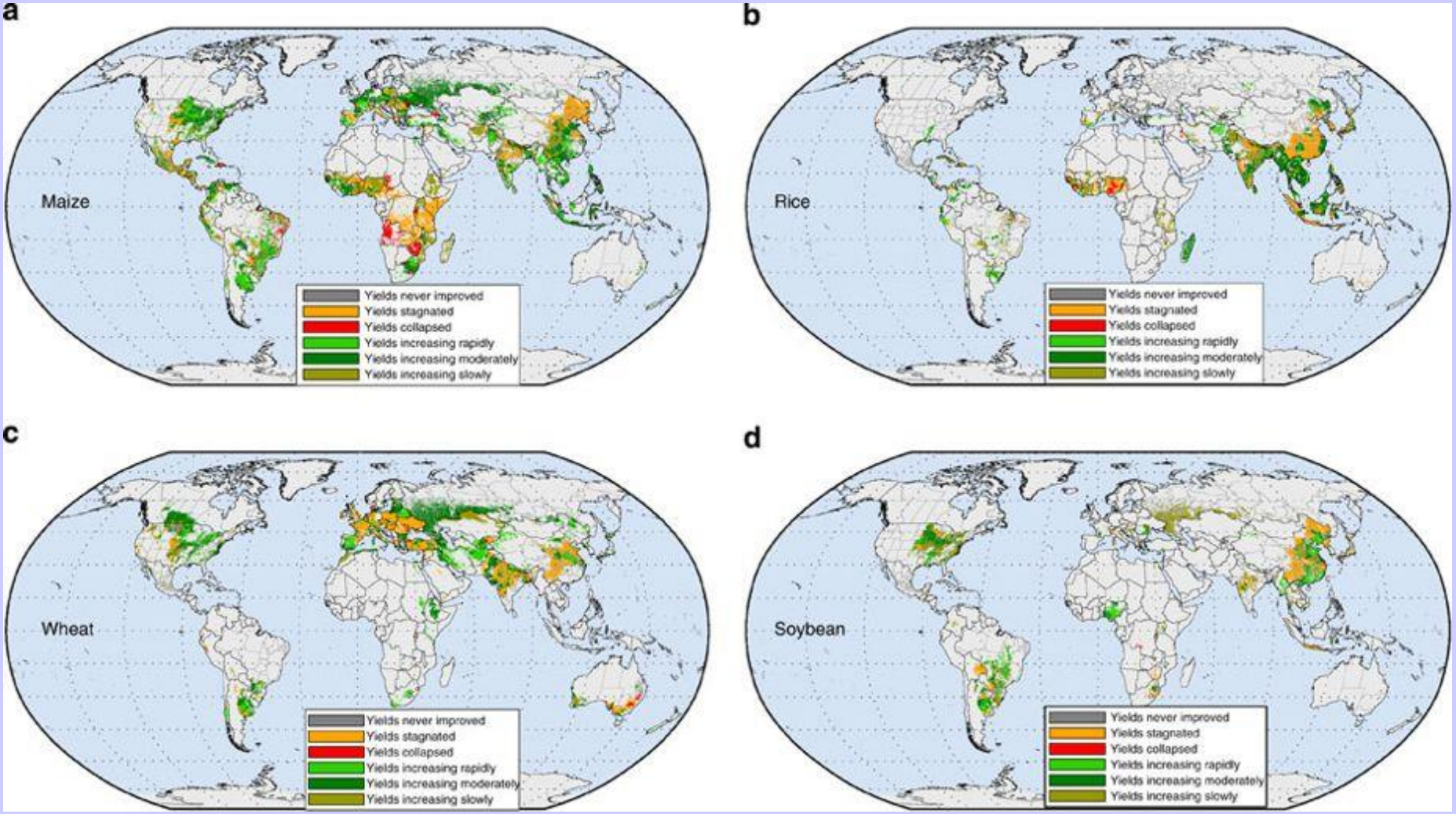
Collapsed

Still increasing



https://www.nature.com/articles/ncomms2296/figures/1

Recent patterns of crop yield growth and stagnation



https://www.nature.com/articles/ncomms2296/figures/2

If smart agriculture can improve food production system dramatically, we will be able to live long peacefully without hard-works.

AgriBigData can accelerate advance of smart agriculture by producing knowledge and its apps.

Knowledge & Apps

Big data

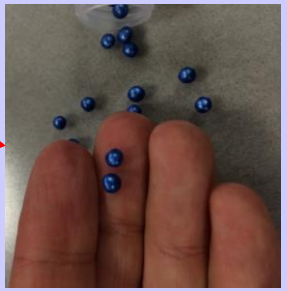
DB DB DB DB SNS Open Data

Farm work data

Machinery data

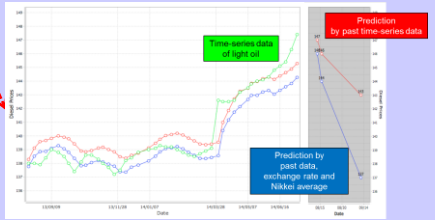
Field image data

Field sensing data



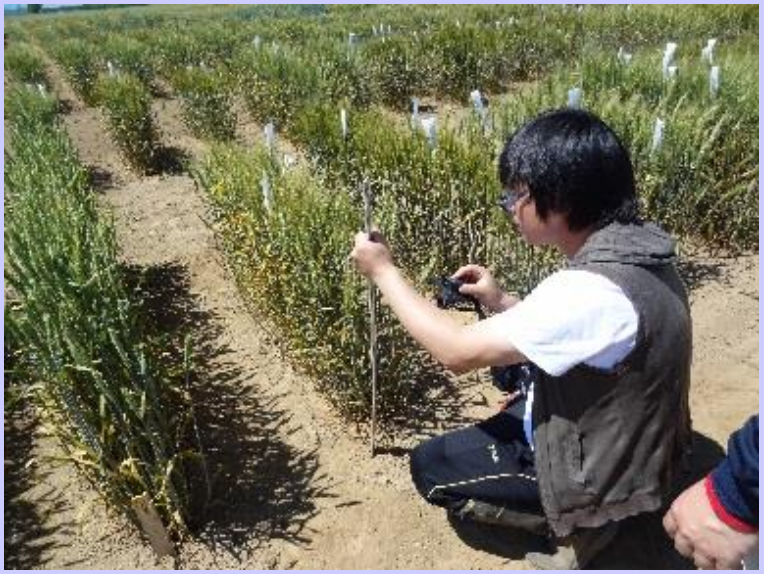
For breeding

For farming



For management

One of the bottle-necks of smart agriculture is data collection in real fields.



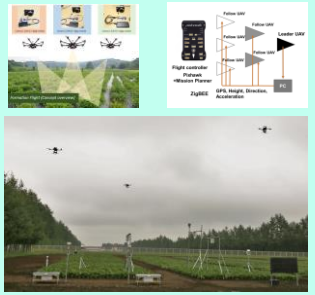
Observation by
manually measured data

Sensor networks (Field Server)
and drones

Especially time-series data collection is difficult in large scale fields.

Collect Data in Fields

Field Sensor Network

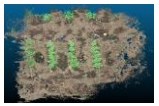


Swarm of Drones

New Sensing Methods

Sensor Data

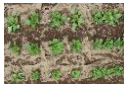
Get numerical data from Images



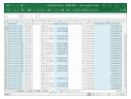
3D-reconstruction

Detecting flowering

Physiological event time



Identifying plants



Numerical data (coverage, height)

Numerical Data

Big Data

Discover Knowledge

Apps

Developing data collection methods with sensor networks and easy/affordable deployment technologies in real fields

The shorter node is the better for spraying, but electro-magnetic field is shielded by plant canopy.

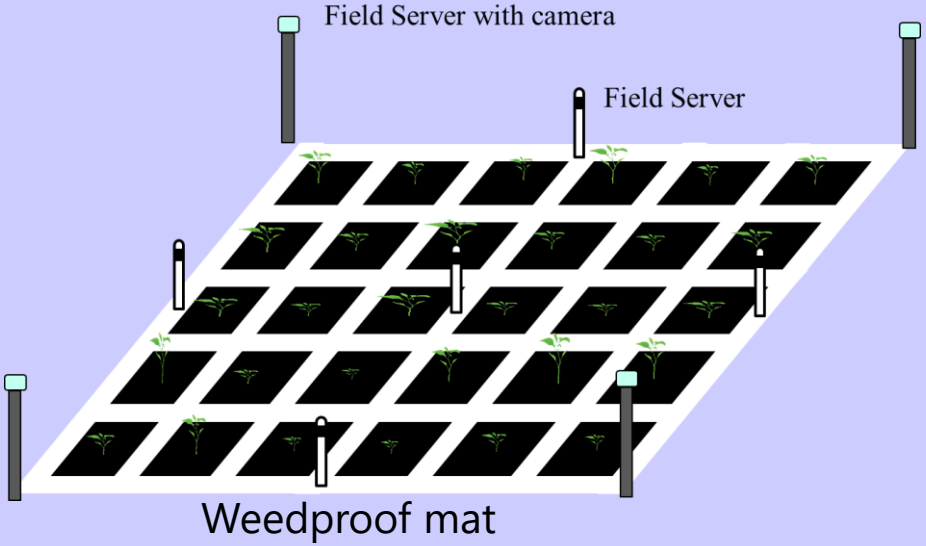
LPWA (Low power, Low bit-rate, Long distance) could send data through the canopy.

Base node for:

- 4G/LTE/Wi-Fi gateway
- Rich sensors
- Maker for drones
- Calibration for 3D-reconstruction
- RTK-GPS base station
- Edge computing

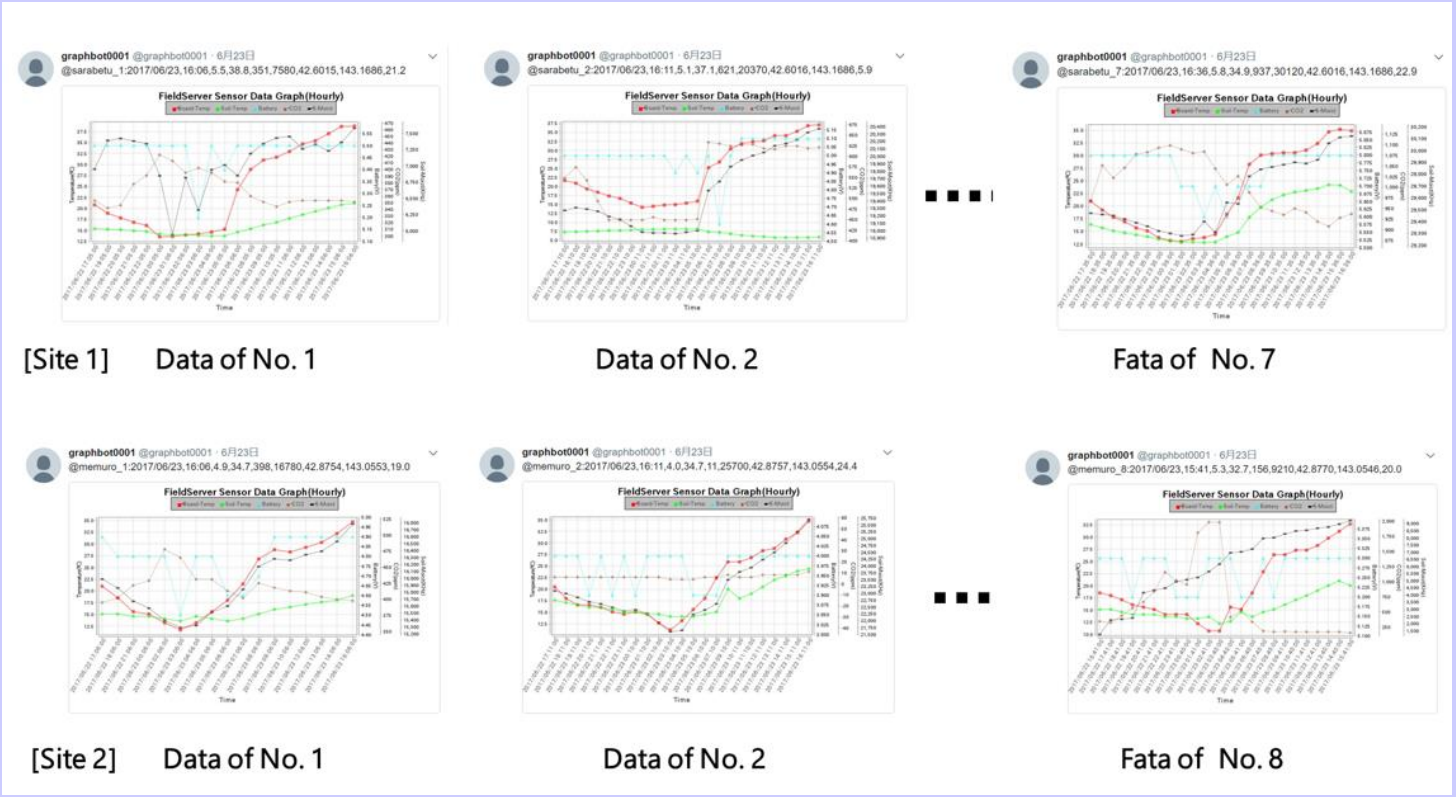
DATA-FARM: Reinvention of the experimental farm

Designed to collect ground truth data easily and to harvest sufficient energy



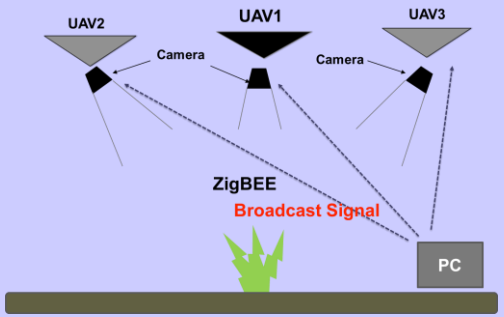
Collected Environment Time-series Data:

Soil moisture, Soil temperature, CO_2 concentration, etc.

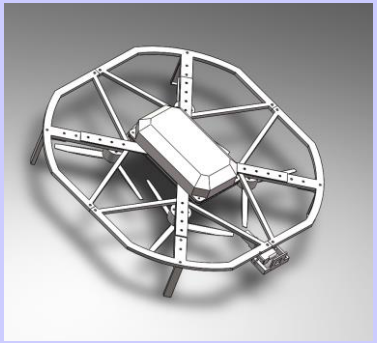


Tweeted data on Twitter

Developing practical observation methods by drones



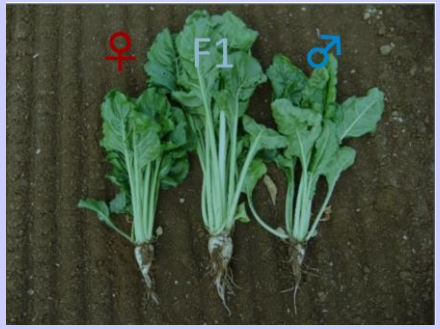
Formation flight control methods

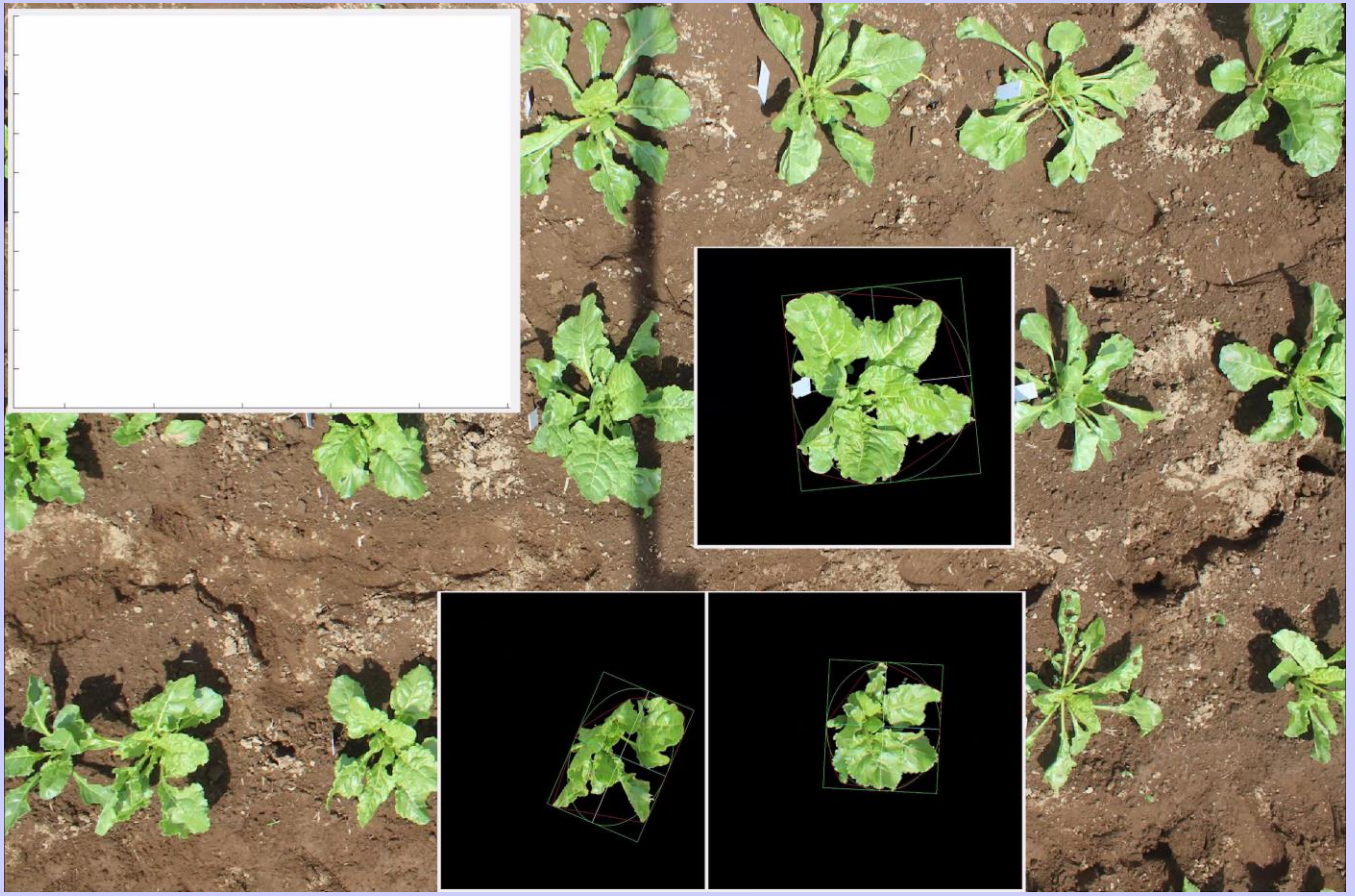


Safety drone by 3D printing

Plant's 3D data can be measured precisely by formation flight against drones' downwash and natural winds.

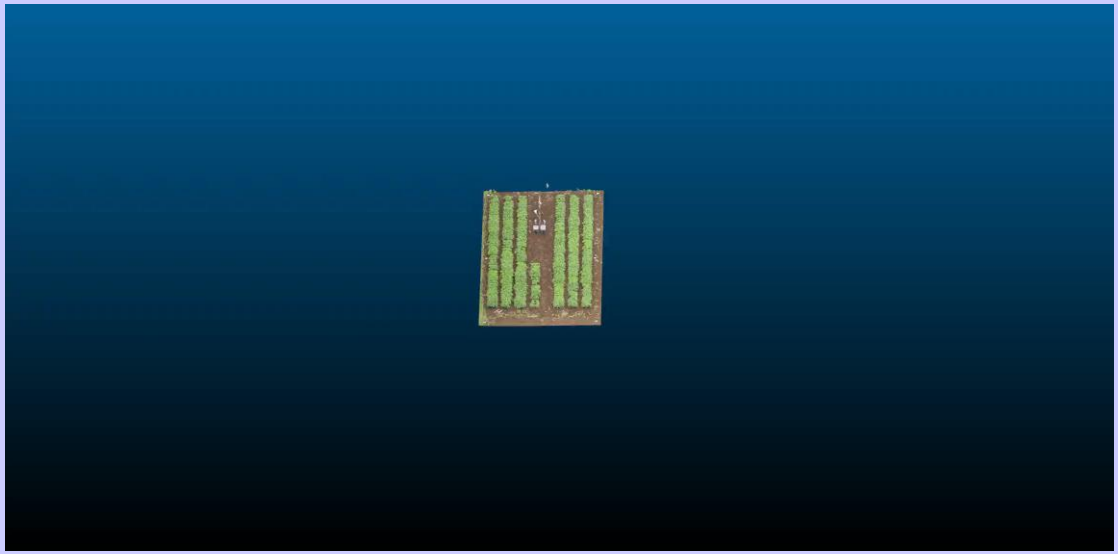
Developing robust segmentation methods against complex background and diverse lighting

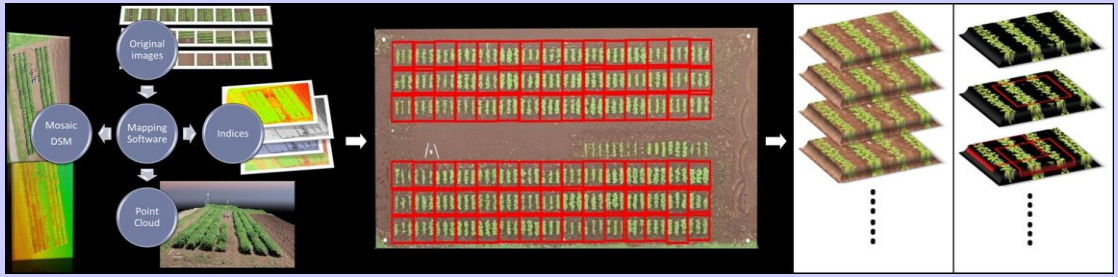


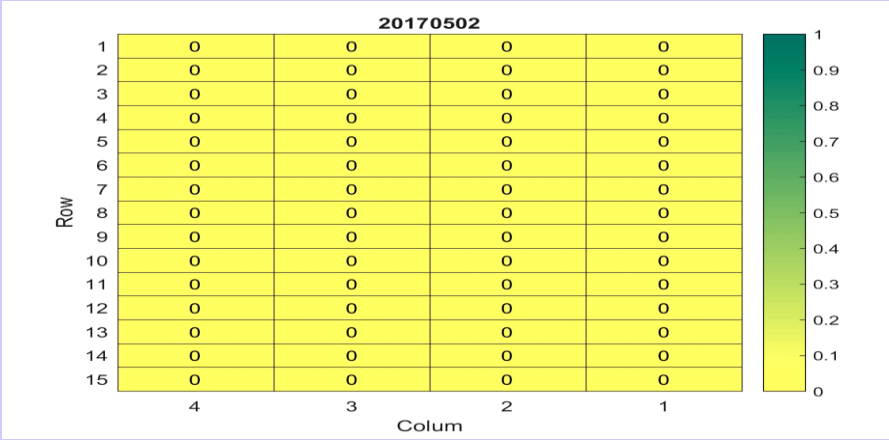


Numerical data (growth rates of parents and F1)
extracted by HyperRecognizer (extended EasyPCC)

Plant phenotyping in fields for soybeans and sugar beets



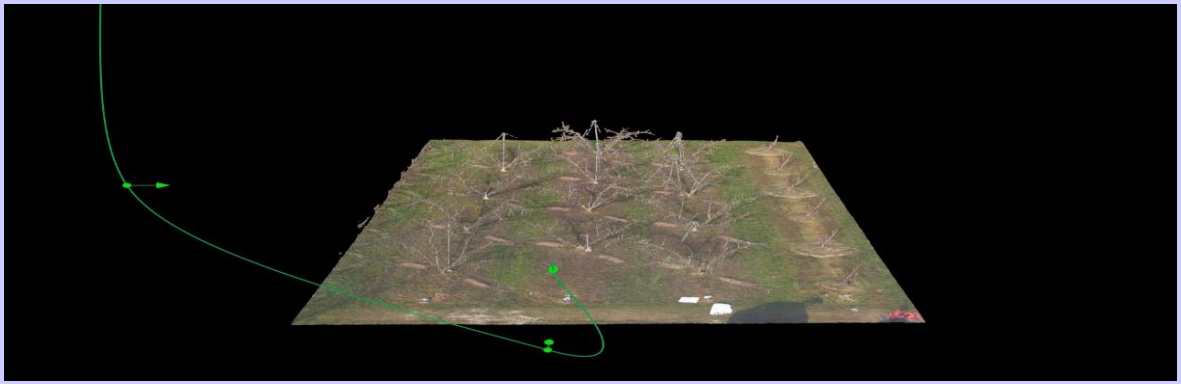


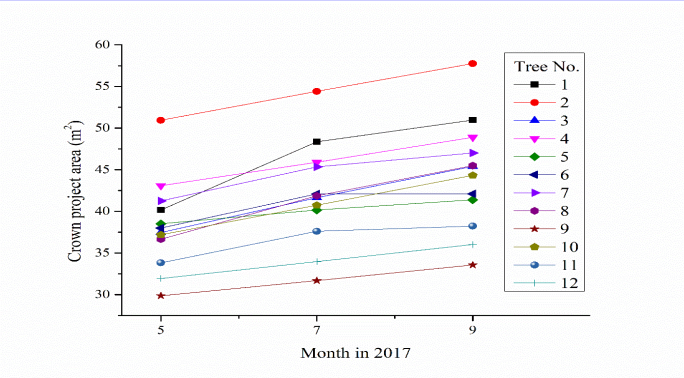


Soybeans (3D data)

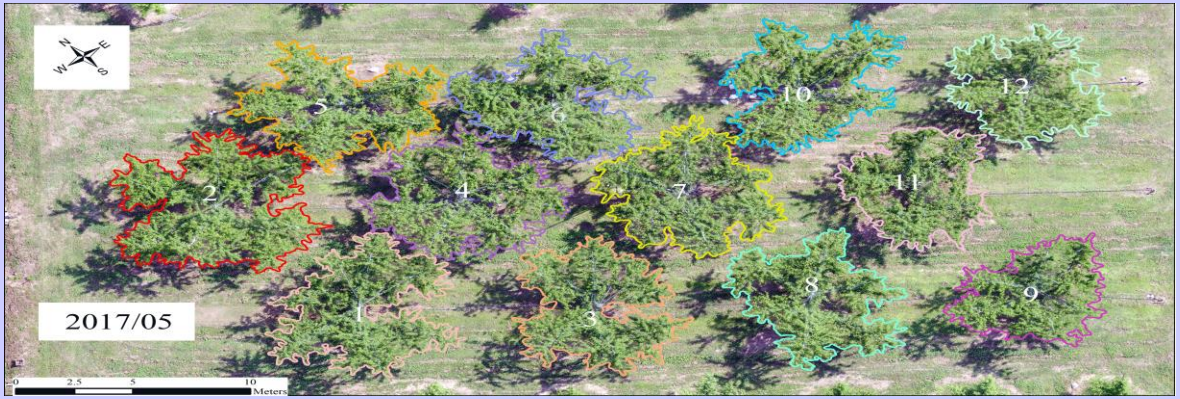
Sugar beets (canopy coverage)

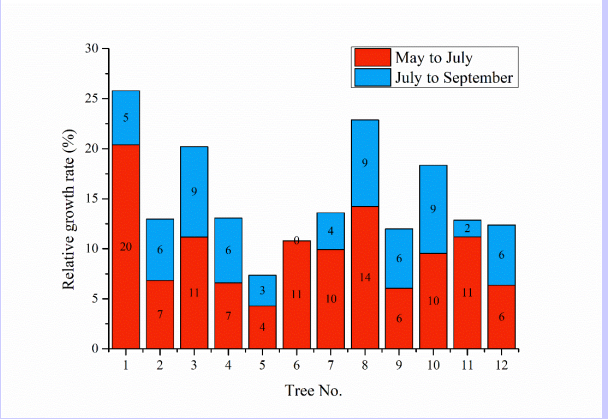
Numerical data (tree's phenotypes) is extracted by HyperRecognizer using time-series images





Crown project area





Growth rate

Breeder Friendly Field Phenotyping System

Image Conversion

Datafication process

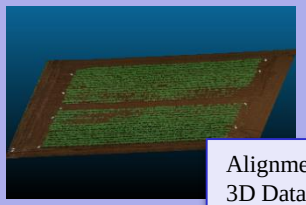
Pipeline A

100m
5ha

3D reconstruction with Opensource software

GCP Information

Alignments and trimming with GCP information.



Alignments 3D Data

Pipeline B

30m
0.5ha

PIX4D

3D images

GeoTiff images

Field design data

Plot segmentation and Phenotyping (calculate coverage and height of plants)

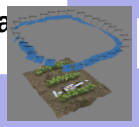
Plant coverage data on each plot

Plant coverage data on each plot

Pipeline C

1.5m
10m²

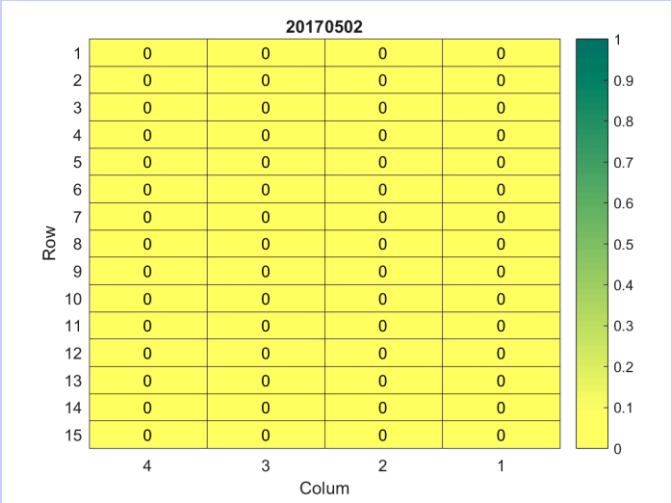
Agisoft Photosca



To be published

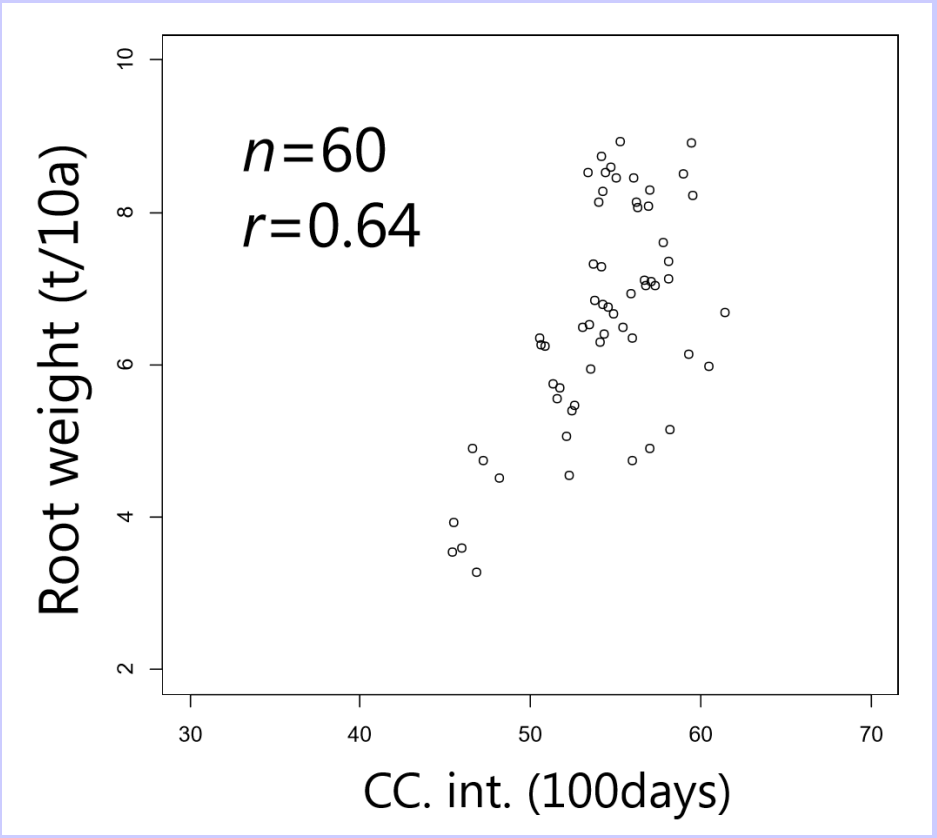
Found new index, 3D score, to predict yield

100 days Canopy Coverage integration (CC. int.)



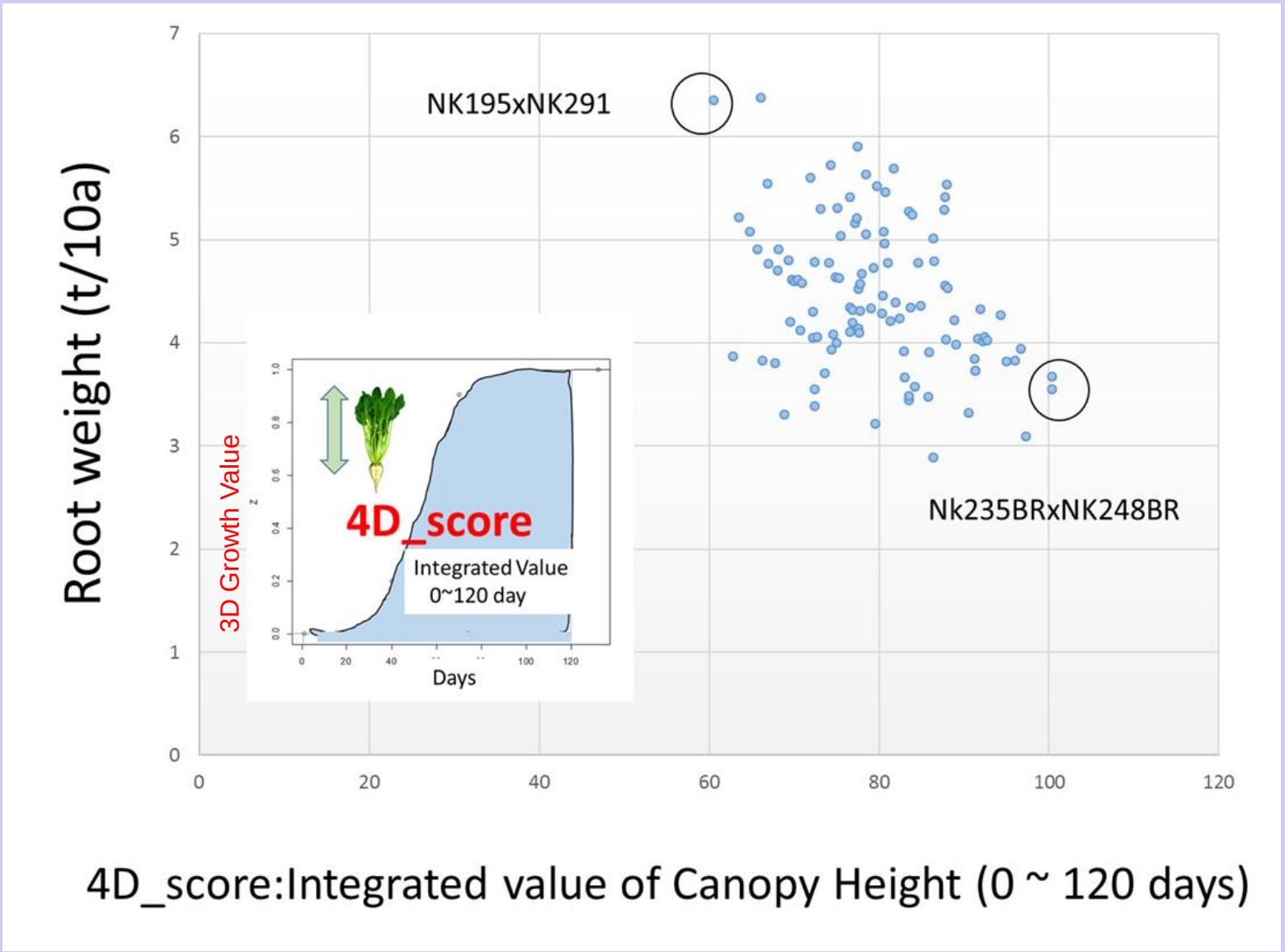
Time-series Canopy Coverage

3D Score
(CC. int.)



Discovered Knowledge:
Heterosis appeared on early stage was measured quantitatively by 3D score as phenotype.

Another Yield-Prediction-Method (4D score) was discovered.



Conclusions and Discussions

1. We proposed a concept of agricultural big data (AgriBigData) which is created by drones and sensor networks to discover knowledge.

2. Software tools and new methods have been developed to construct AgriBigData.

3. We discovered phenotyping index (3D/4D scores) to measure Heterosis quantitively. 3D/4D scores are also useful to predict yield.

4. Collaboration with INRA, ISU, etc. has been very useful.