Practical uses of machine learning in agriculture

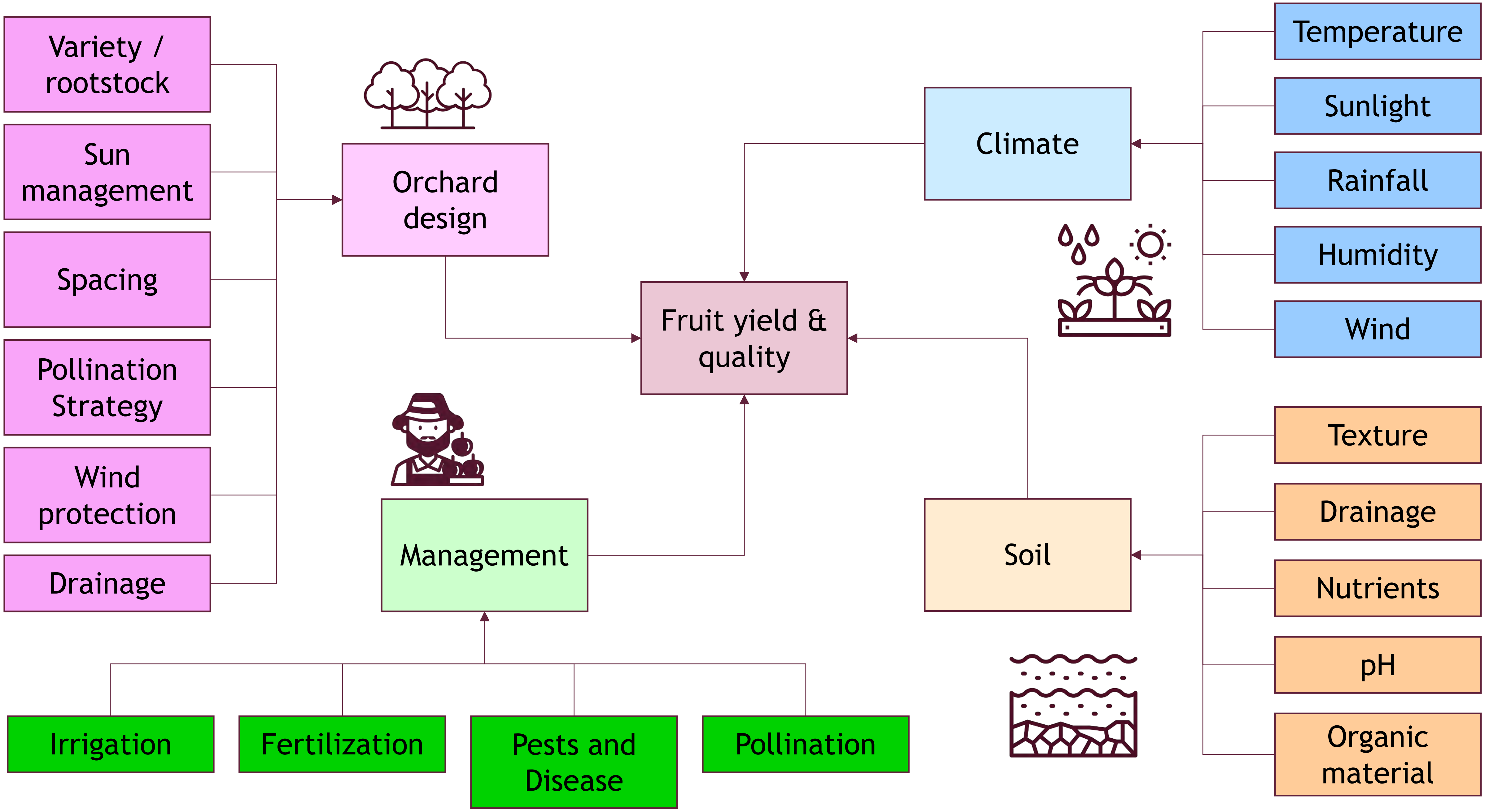
Prof Adriaan van Niekerk

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

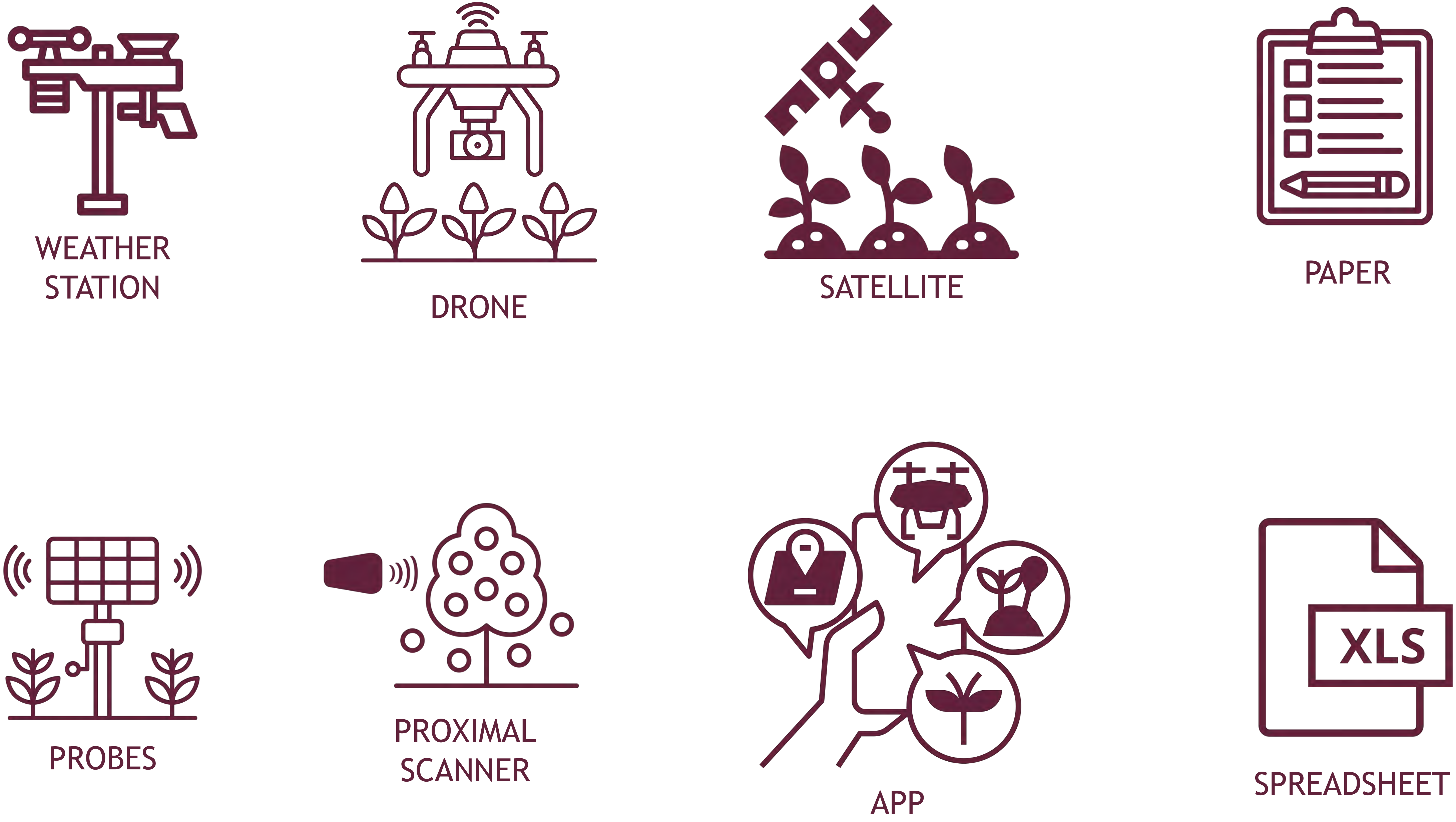
Photo by Stefan Els

Structure

1. Machine learning?
2. Examples



Variety / rootstock
Sun management
Spacing
Pollination Strategy
Wind protection
Drainage
Orchard design
Fruit yield & quality
Climate
Temperature
Sunlight
Rainfall
Humidity
Wind
Management
Soil
Texture
Drainage
Nutrients
pH
Organic material
Irrigation
Fertilization
Pests and Disease
Pollination


WEATHER STATION
DRONE
SATELLITE
PAPER
PROBES
PROXIMAL SCANNER
APP
SPREADSHEET
XLS

GIS, GNSS*
&
Remote Sensing

offer a set of **geographical tools** and techniques to bring all of these data together

* e.g. GPS

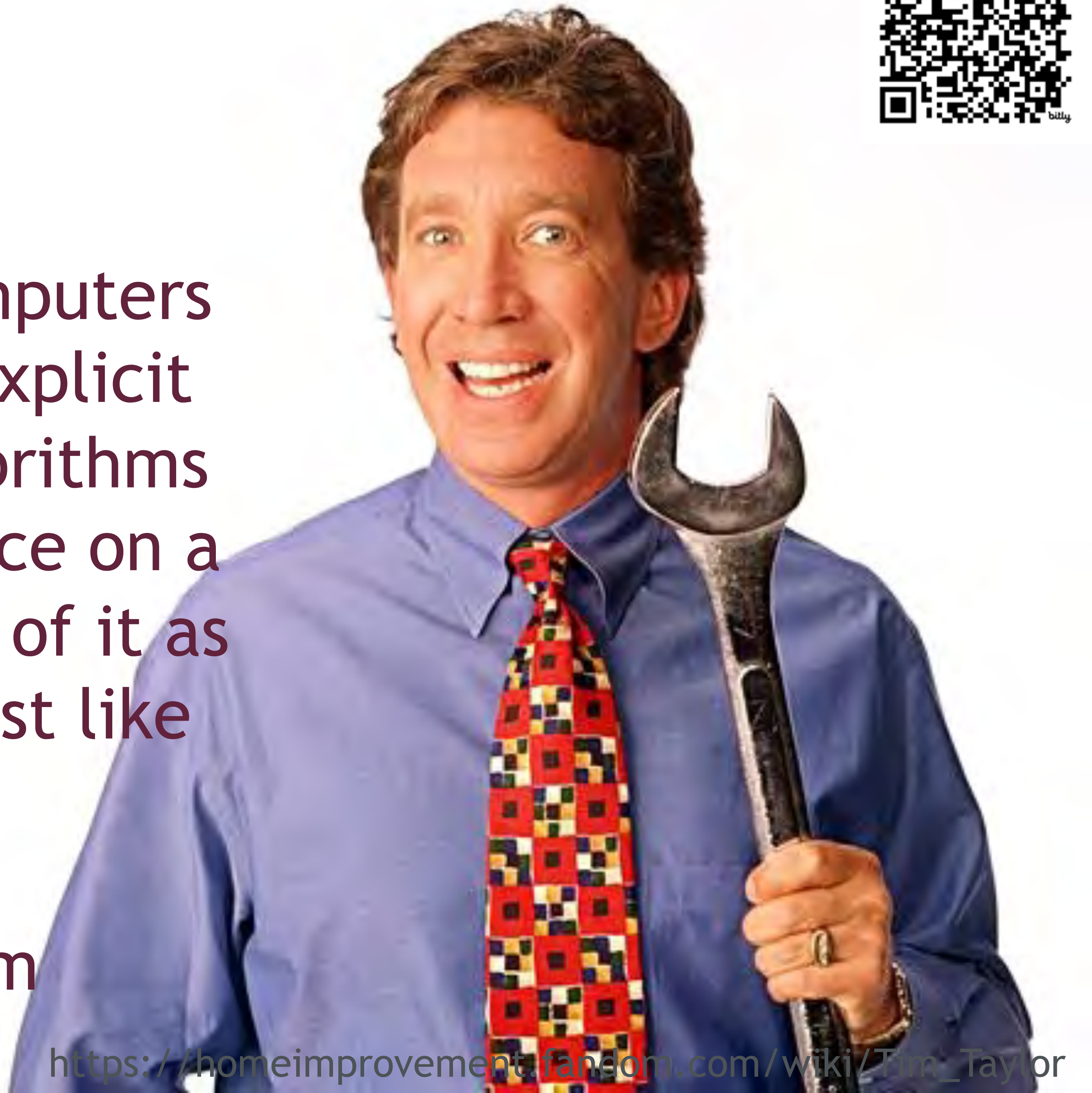
Problem

Too much data to make sense of!

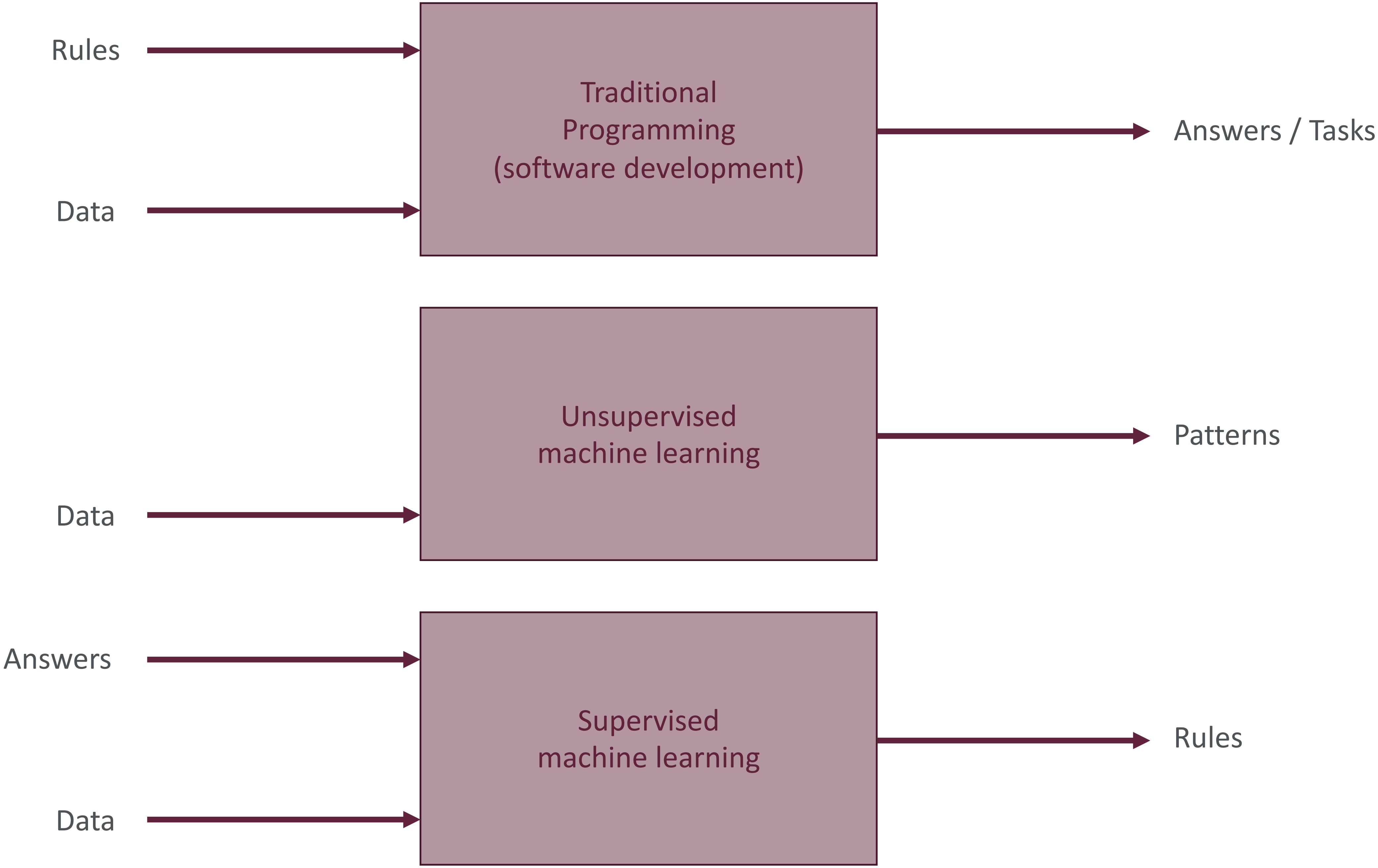
New tool! (kinda)

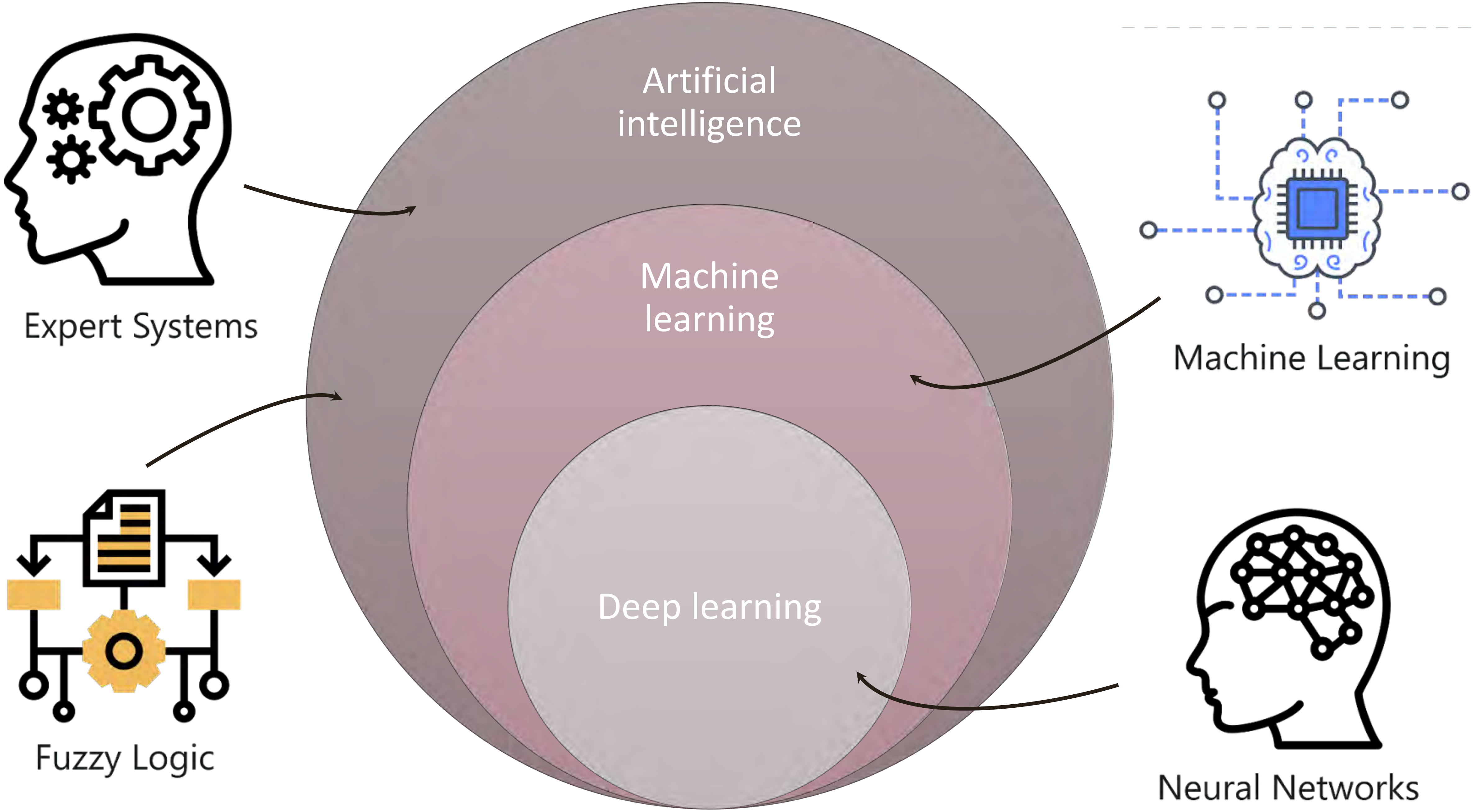
"Machine learning allows computers to learn from data without explicit programming. It involves algorithms that improve their performance on a specific task over time. Think of it as learning from experience, just like humans do.."

~gemini.google.com

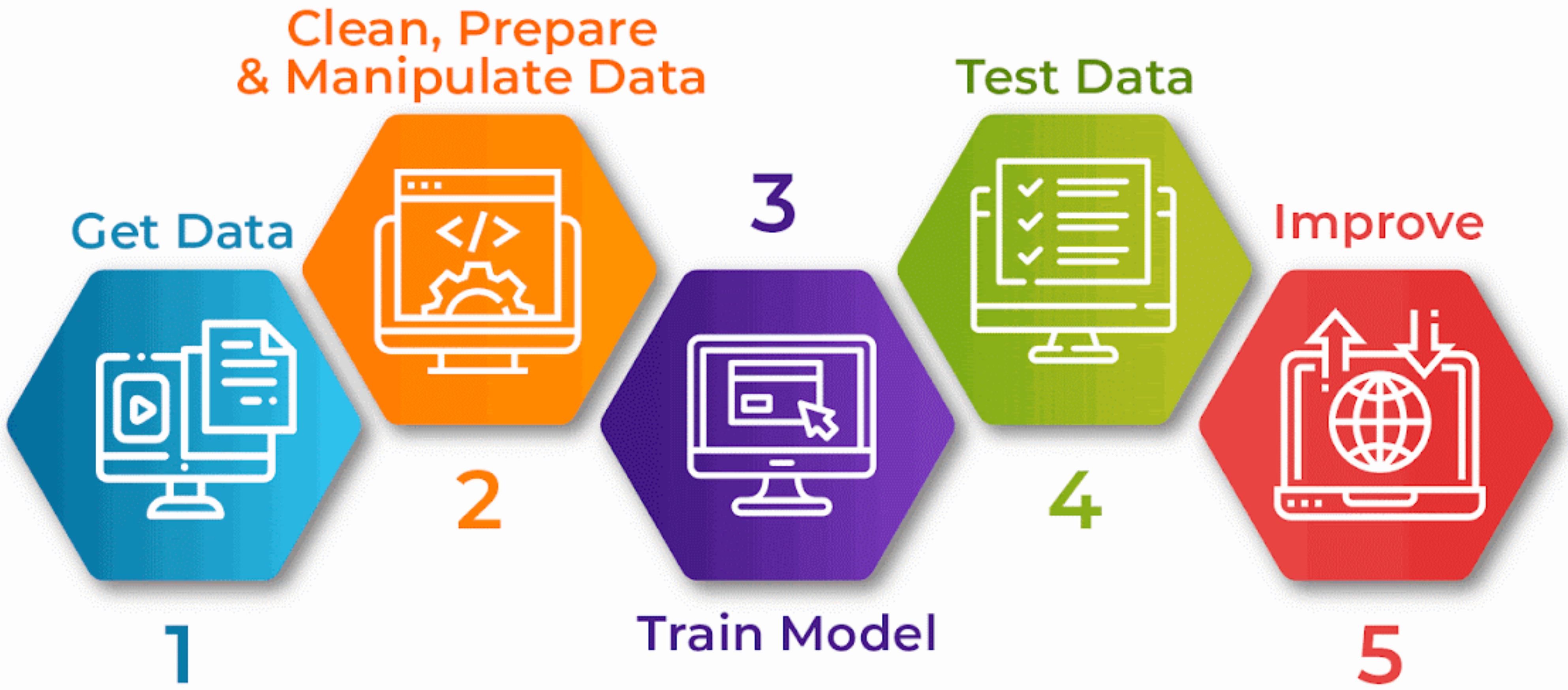


https://homeimprovement.fandom.com/wiki/Tim_Taylor


Rules
Data
Traditional Programming (software development)
Answers / Tasks
Unsupervised machine learning
Data
Patterns
Answers
Data
Supervised machine learning
Rules

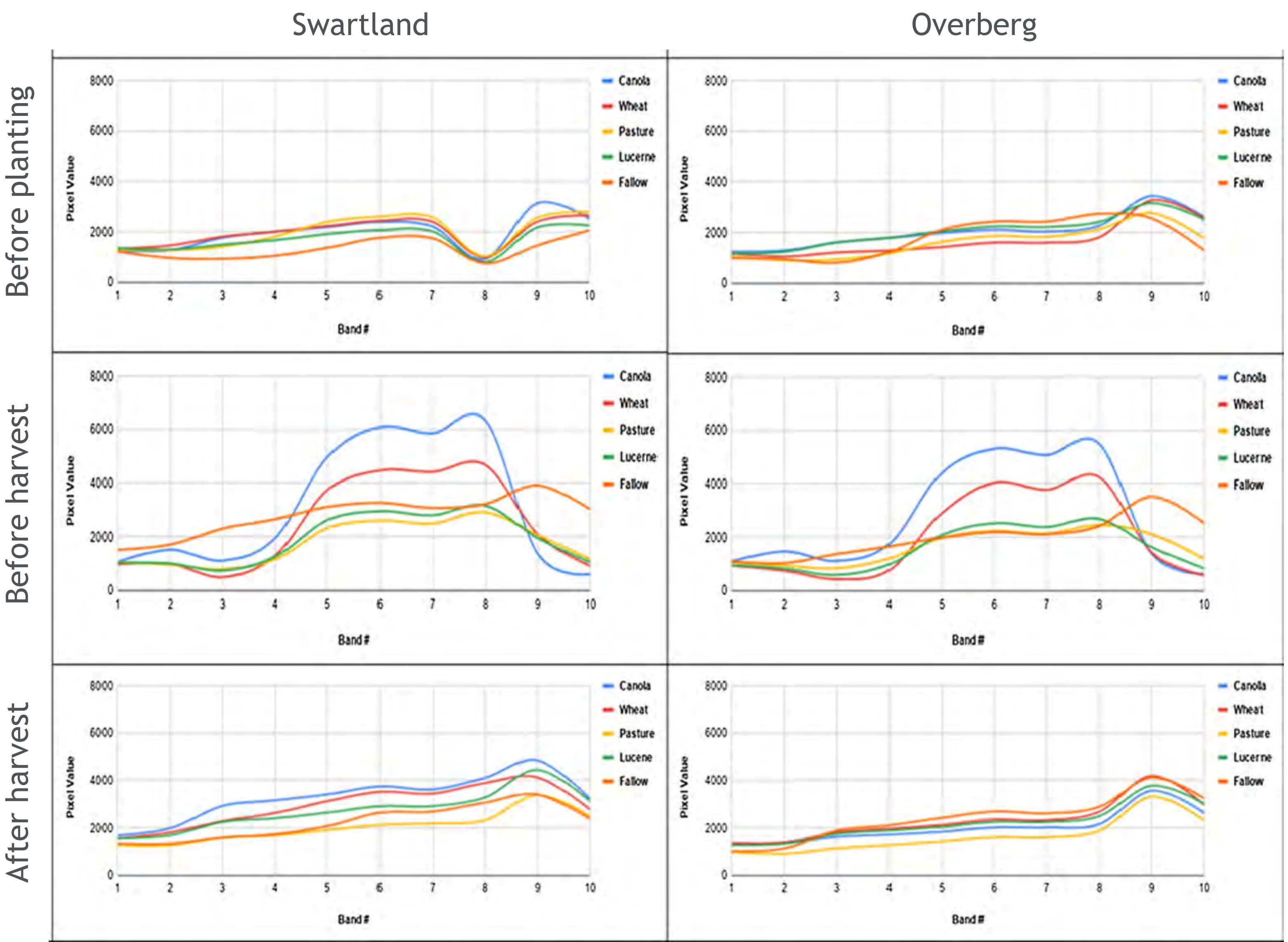

Artificial intelligence
Machine learning
Deep learning
Expert Systems
Fuzzy Logic
Machine Learning
Neural Networks

Supervised machine learning process



Example 1: Crop type/cultivar mapping

Field crop type differentiation by exploiting phenological differences



Site	Image	Dates	SVM
			OA
A	1 – 2	6 Apr – 26 Apr	51.7
	1 – 3	6 Apr – 16 May	56.0
	1 – 4	6 Apr – 6 Jun	61.5
	1 – 5	6 Apr – 4 Aug	79.6
	1 – 6	6 Apr – 3 Sept	80.3
	1 – 7	6 Apr – 3 Oct	79.2
	1 – 8	6 Apr – 13 Oct	78.0
	1 – 9	6 Apr – 23 Oct	78.8
	All	6 Apr – 31 Jan	81.1
		AVG	69.2
		SD	11.8
B	1 – 2	3 Apr – 22 Jun	56.8
	1 – 3	3 Apr – 2 Jul	68.8
	1 – 4	3 Apr – 1 Aug	77.2
	1 – 5	3 Apr – 11 Aug	79.0
	1 – 6	3 Apr – 21 Aug	78.4
	1 – 7	3 Apr – 20 Oct	80.2
	1 – 8	3 Apr – 29 Dec	80.2
	1 – 9	3 Apr – 8 Jan	80.8
	All	3 Apr – 18 Jan	80.5
		AVG	72.7
		SD	8.02

Maponya, M. G., Van Niekerk, A., & Mashimbye, Z. E. (2020). Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Computers and electronics in agriculture*, *169*, 105164.

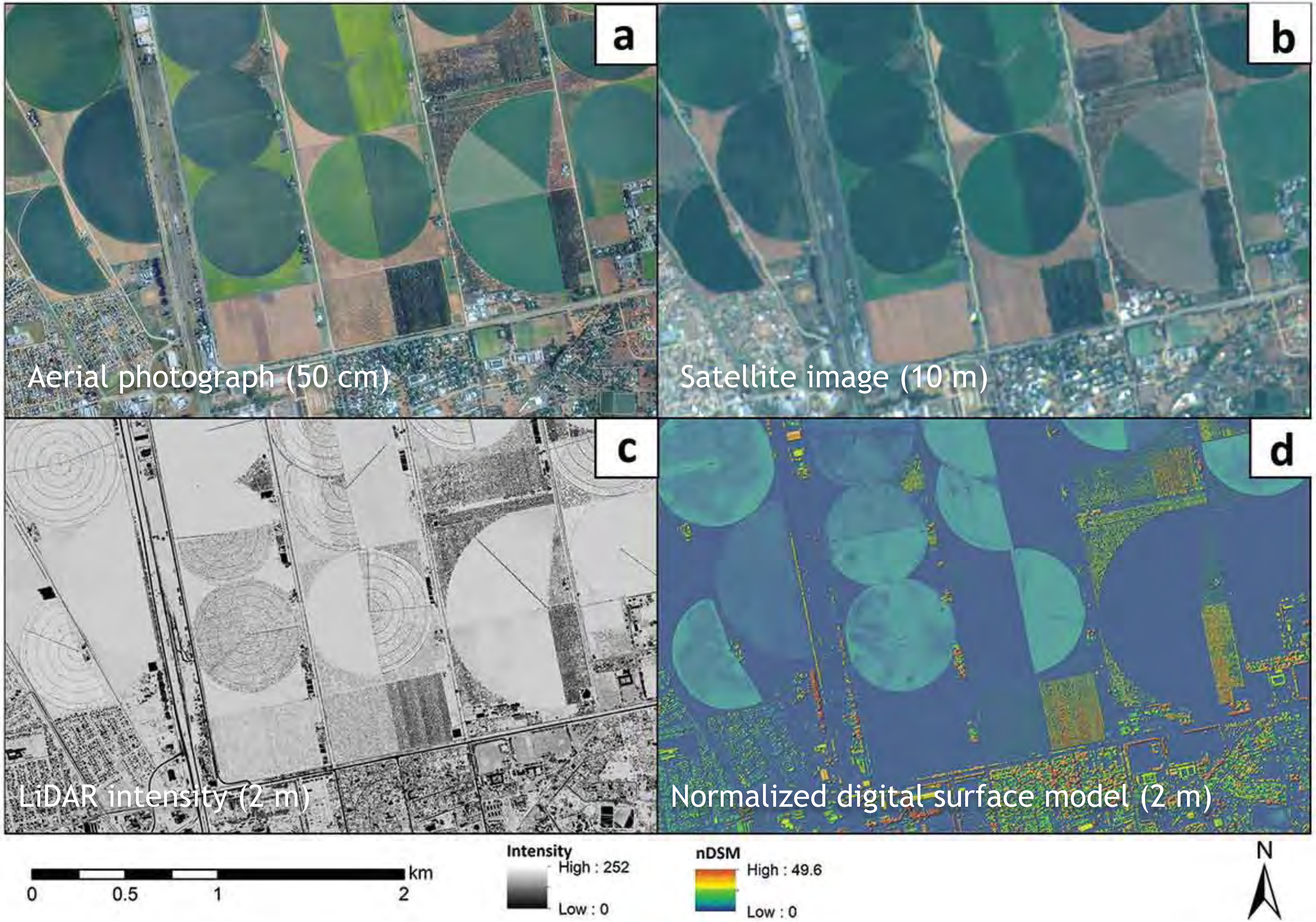
Broad crop type classes

Non-agriculture
Maize
Orchard (Pecan)
Groundnut
Cotton

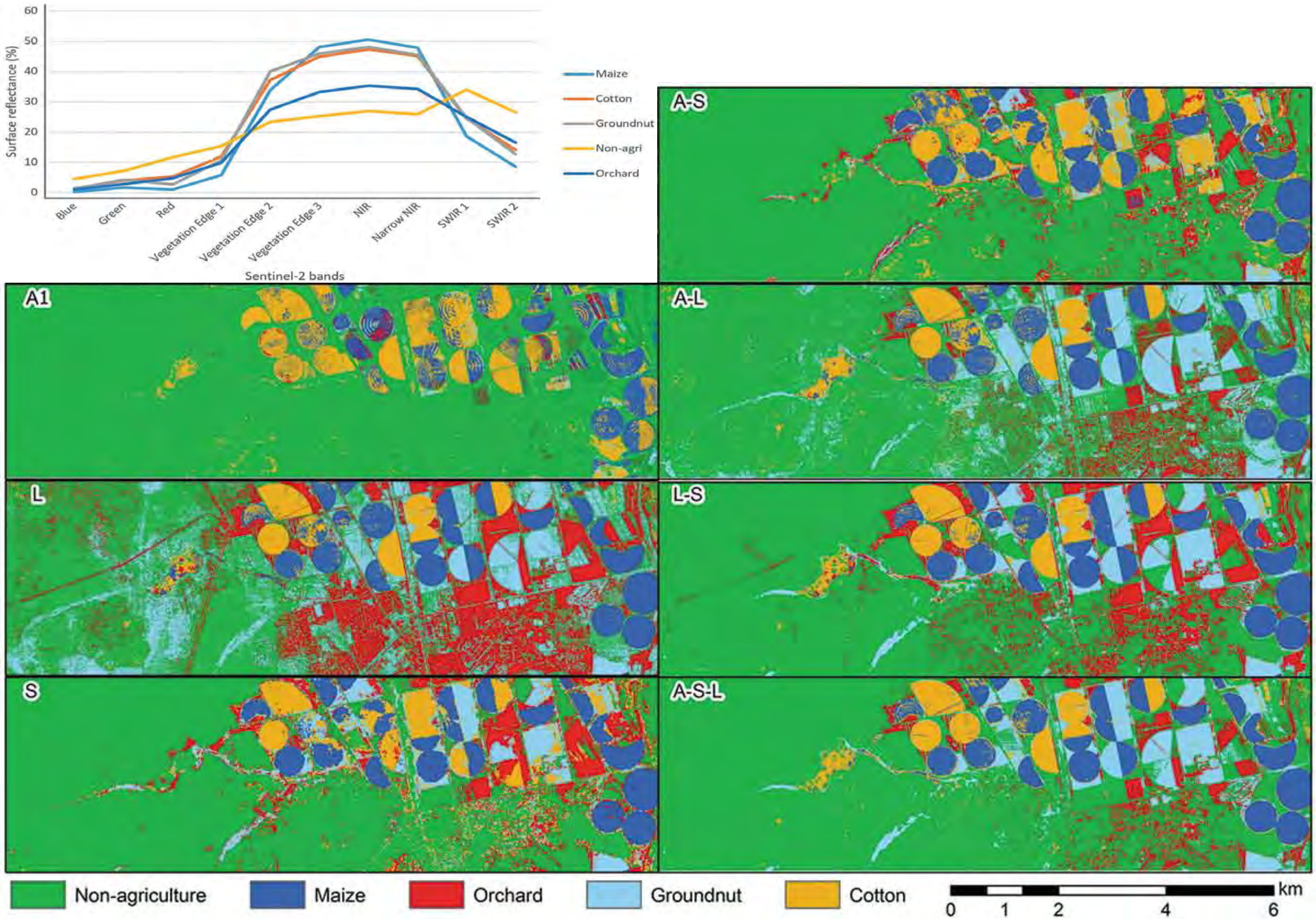
Compared three sources of data:

1. Aerial photograph
2. Satellite image
3. LiDAR data

Compared various machine learning algorithms



Prins, A. J., & Van Niekerk, A. (2021). Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. *Geo-Spatial Information Science*, *24*(2), 215-227.



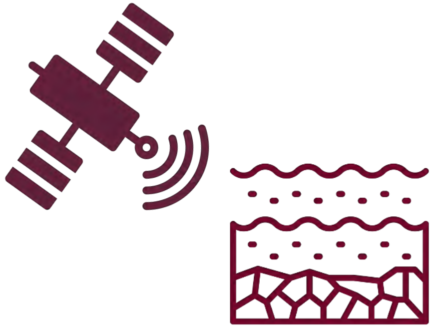
Random forest outperformed (95%) all the other classifier algorithms.

Best results (92%) were obtained when all three sets of data were used in combination.

New project: Using hyperspectral (5 m) imagery to differentiate crop types and cultivars

Example 3: Soil property mapping (salinity)

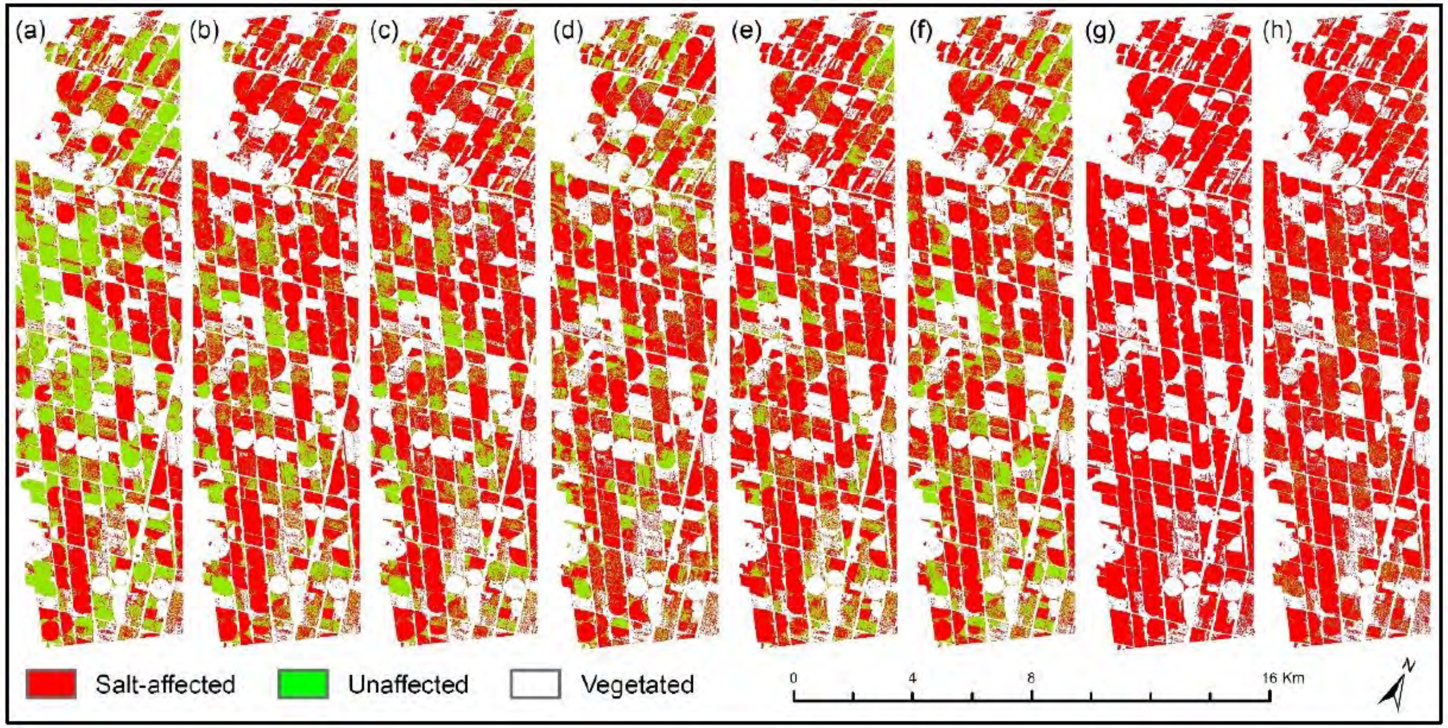
Direct approach

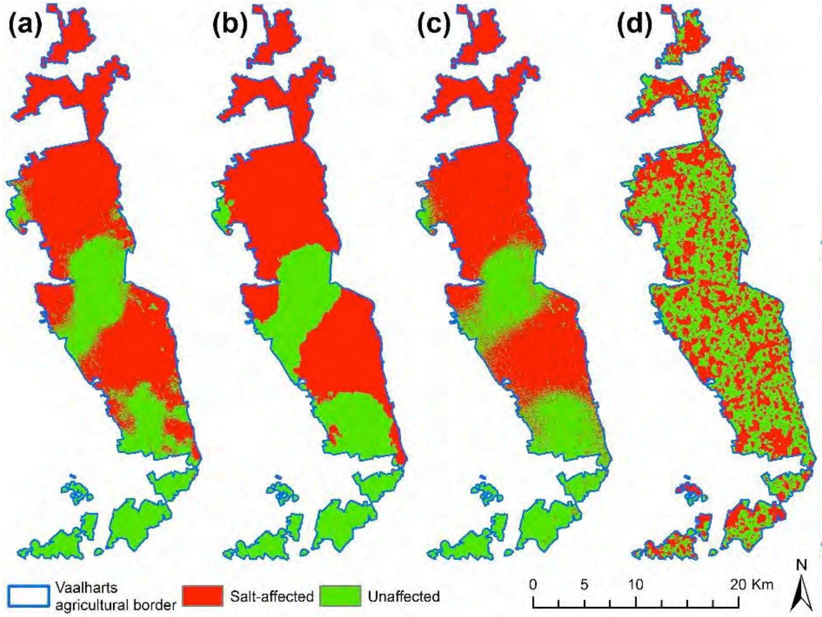


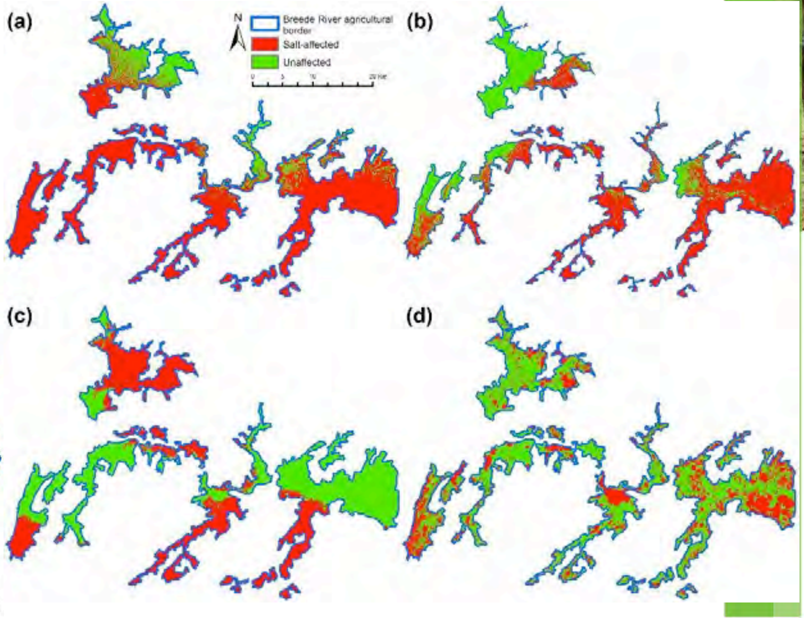
1. Collect & analyze soil samples
2. Acquire satellite images when field is bare (after ploughing)
3. Combine soil analyses results and satellite imagery at sampled points
4. Train machine learning model
5. Apply model to entire satellite image to produce maps

Indirect approach

1. Collect & analyze soil samples
2. Acquire satellite images when crops are in vegetative phase
3. Combine soil analyses results and satellite imagery at sampled points
4. Train machine learning model
5. Apply model to entire satellite image to produce maps


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
Salt-affected
Unaffected
Vegetated
0
4
8
16 Km
N

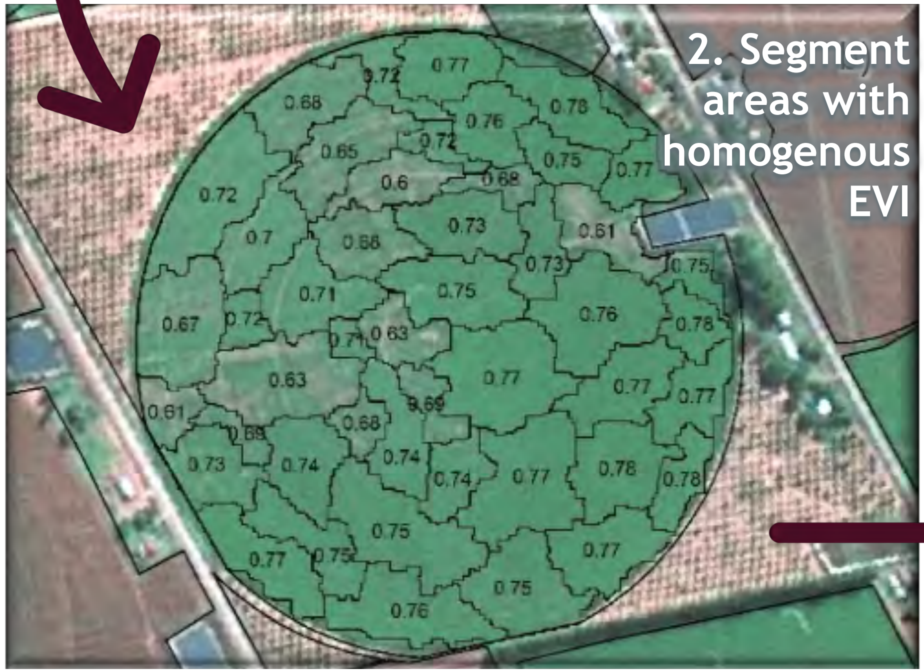

(a)
(b)
(c)
(d)
Vaalharts agricultural border
Salt-affected
Unaffected
0
5
10
20 Km
N

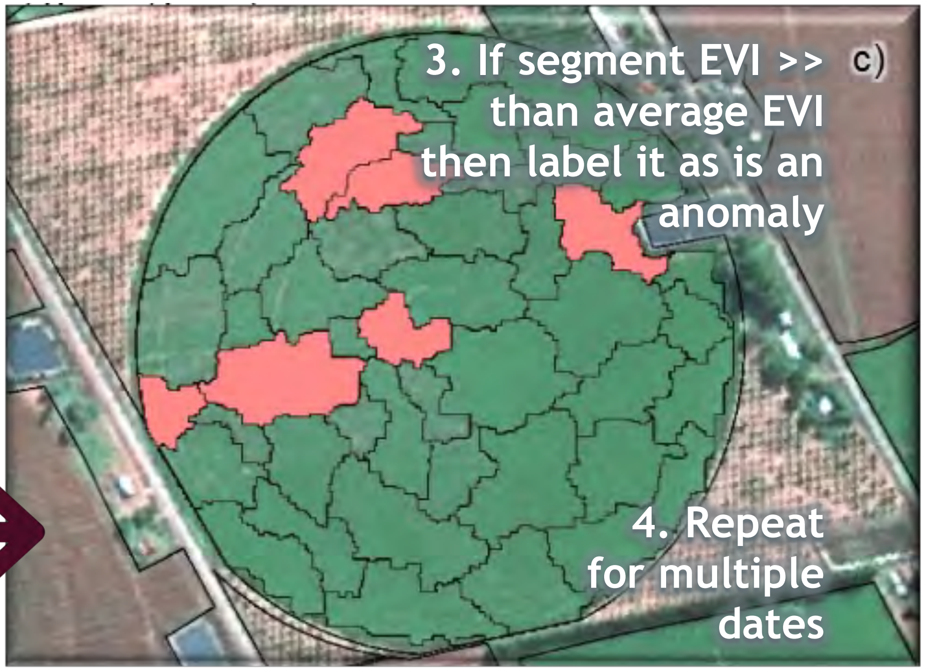

(a)
(b)
(c)
(d)
Breede River agricultural border
Salt-affected
Unaffected


LEGEND
Anomalies
0
250
500
CGA
bitly
Methodology for monitoring waterlogging and salt accumulation on selected irrigation schemes in South Africa
JP Nell, A van Niekerk, SJ Muller, D Vermeulen, T Pauw, G Stephenson, J Kemp
WATER RESEARCH COMMISSION
TT 648/15

Knowledge-based within field anomaly detection

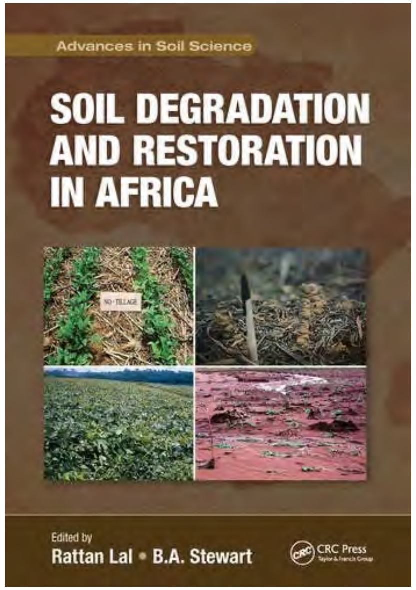
(for salt accumulation monitoring)



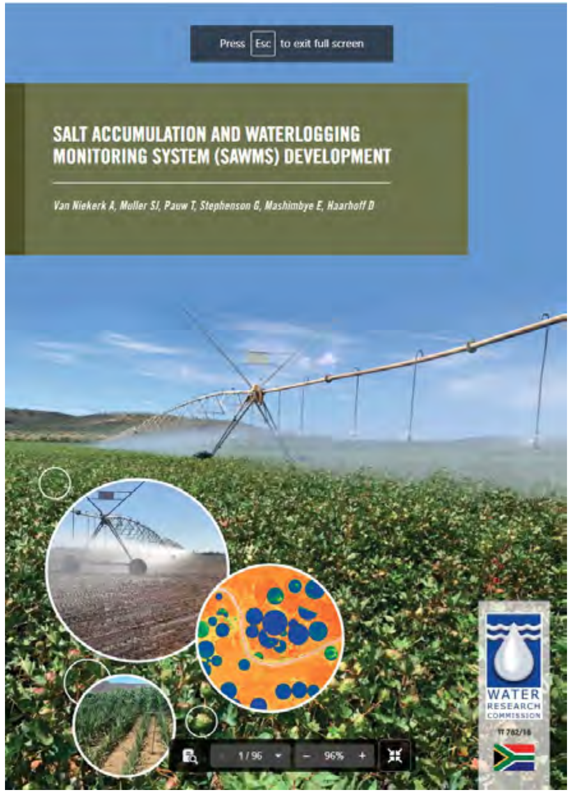


WATER RESEARCH COMMISSION

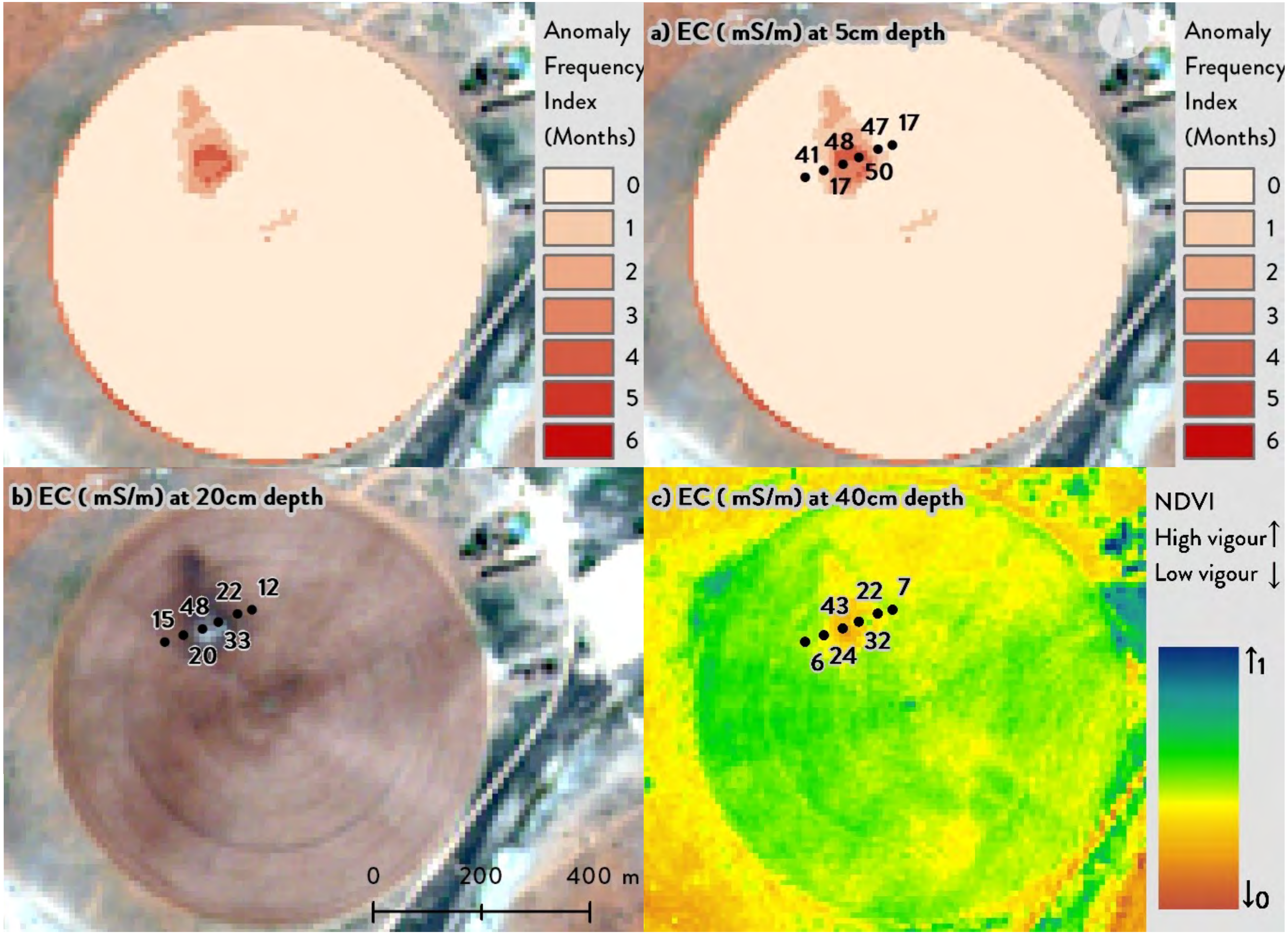
Muller, S. and van Niekerk, A., 2019. Within-field monitoring of secondary salinity in irrigated areas of South Africa. In Soil Degradation and Restoration in Africa (pp. 88-110). CRC Press.

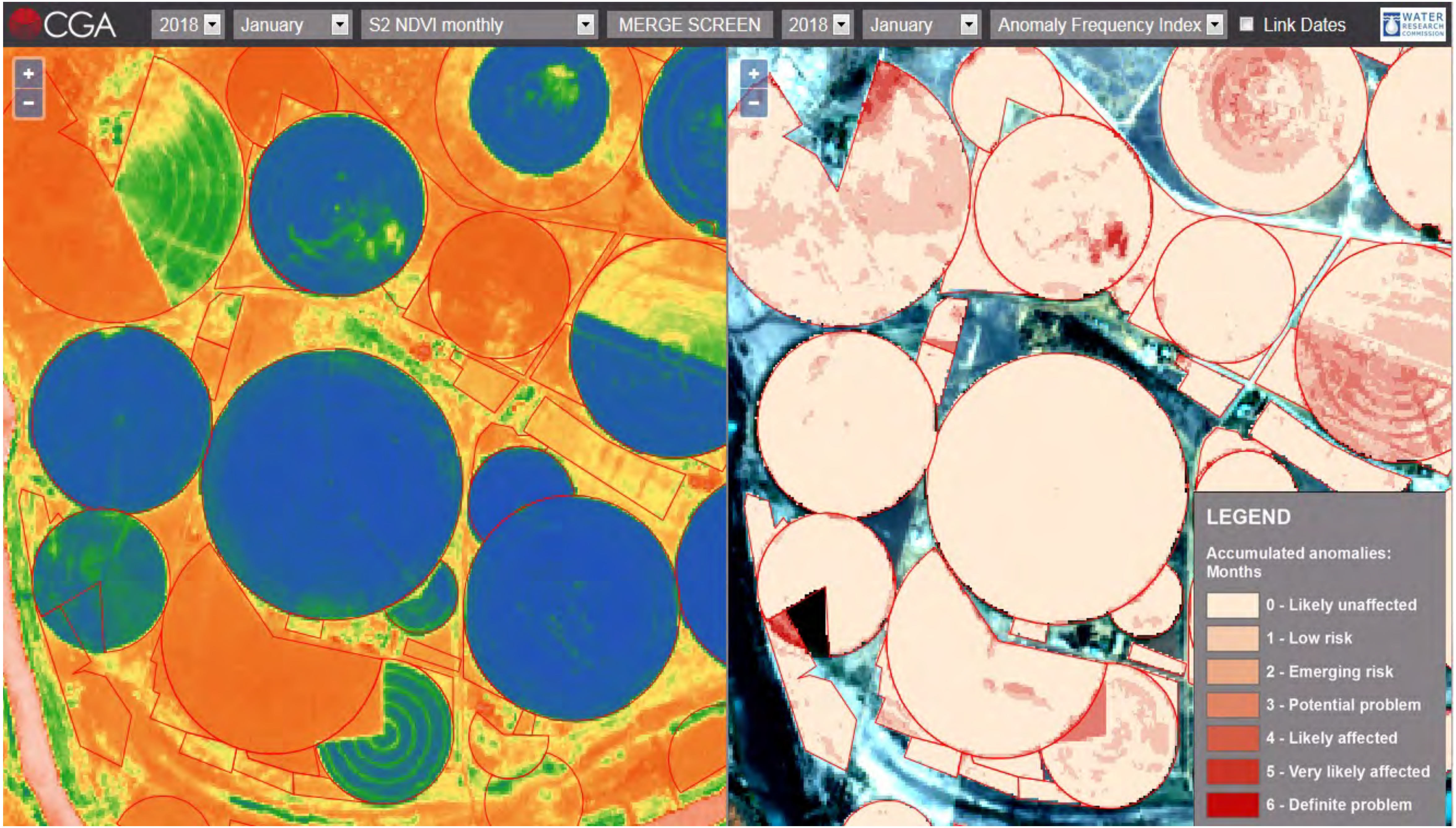


bitly



Van Niekerk, A., Muller, S.J., Pauw, T., Stephenson, G., Mashimbye, E. and Haarhoff, D., 2019. Salt accumulation and waterlogging monitoring system (SAWMS) development. WRC Report TT782/18. Pretoria, Water Research Commission. Pp. 78.





Salt Accumulation and Waterlogging Monitoring System (SAWMS)

Example 4: Crop/cultivar suitability analysis (TerraClim)

TerraClim Suitability Tool | Deciduous Fruit

Field: 1

Suitability Score	Cultivar
100	Apple
86	Cherry
64	Pear
6	Nectarine
0	Plum

Field: 2

Suitability Score	Cultivar
100	Cherry
94	Apple
64	Pear
19	Plum
0	Nectarine

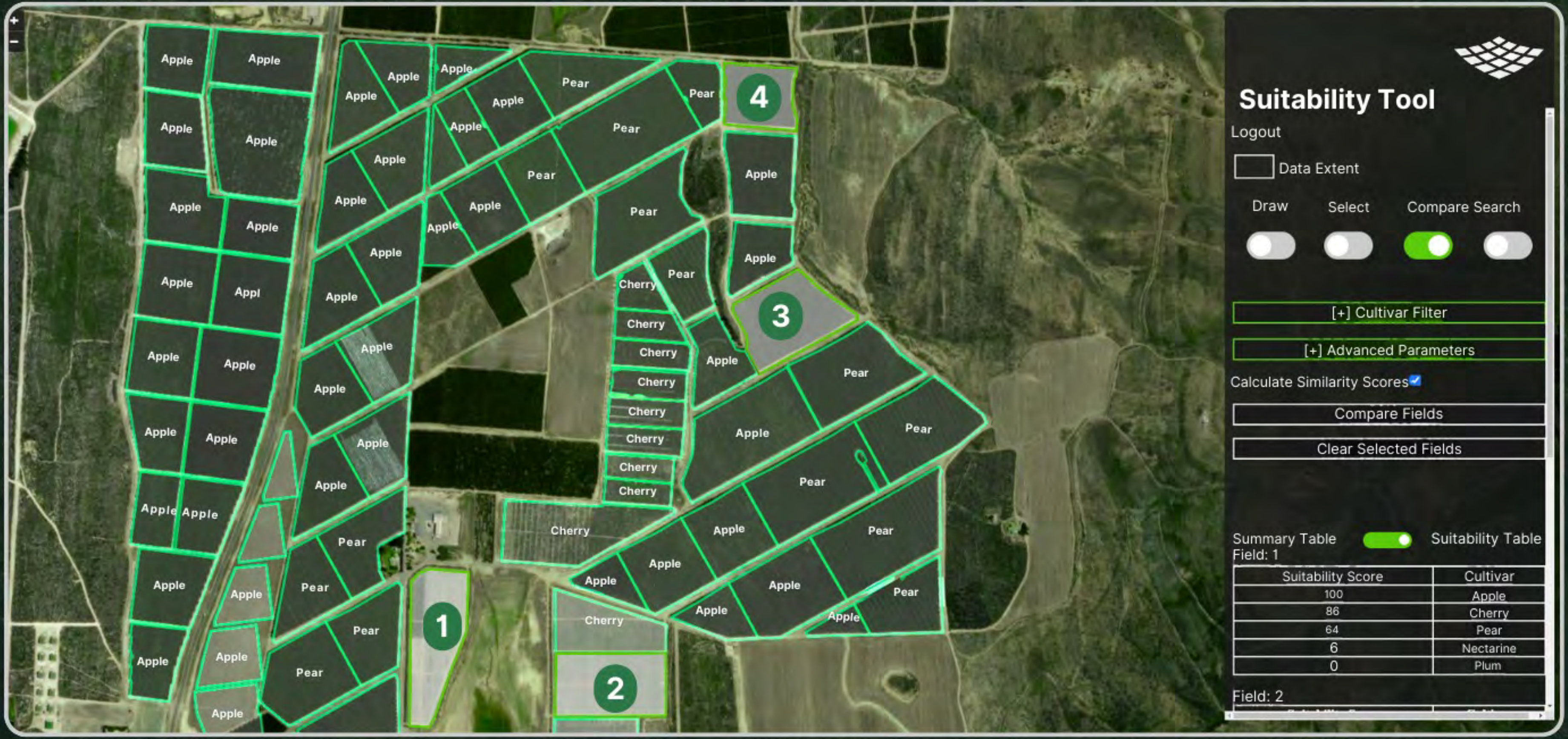
Field: 3

Suitability Score	Cultivar
100	Apple
87	Cherry
70	Pear
6	Plum
0	Nectarine

Field: 4

Suitability Score	Cultivar
100	Apple
97	Cherry
78	Pear
14	Plum
0	Nectarine

AI-assisted Suitability Tool with Machine Learning



TerraClim Suitability Tool | Deciduous Fruit

Field 1:

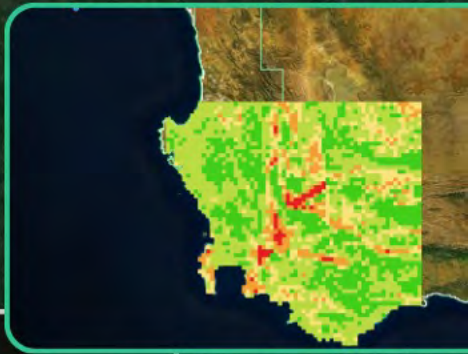
Suitability Score	Cultivar
100	Apple
86	Cherry
64	Pear
6	Nectarine
0	Plum

Field 2:

Suitability Score	Cultivar
100	Cherry
88	Apple
67	Pear
24	Plum
0	Nectarine

Feature Engineering involves extracting **43 important climate, soil and terrain variables** for each vineyard to create a detailed environmental profile for each cultivar.

The algorithm learns **cultivar profiles** to **differentiate cultivars** based on **unique patterns and ideal terroir "fingerprint".**



TerraClim Suitability Tool | Analysis of the Suitability of Wine Grape Cultivars (Terroir)

Decoding terroir data:
Analyse key climate, soil and terrain variables to better understand cultivar suitability.

Suitability Tool

Summary Table | Suitability Table

Field: 1

Average Value	Variable
4.43	Slope (%)
2823.80	Growing Degree Days
21.35	Growing Season Temperature
295.50	Elevation (m)
41231.06	Distance To Coast (m)
1426205	Solar Radiation
5.05	Mean Wind Speed (m/s)
North	Aspect
402.90	Mean Soil Depth
11.50	Mean Clay

Field: 2

Average Value	Variable
15.67	Slope (%)
2605.70	Growing Degree Days
20.95	Growing Season Temperature
297.68	Elevation (m)
40853.39	Distance To Coast (m)
1387948	Solar Radiation
4.71	Mean Wind Speed (m/s)
West	Aspect
402.90	Mean Soil Depth
11.50	Mean Clay

TERRACLIM
FORWARD FARMING

SOUTH AFRICA WINE
discover diversity in a glass

TerraClim Suitability Tool | Analysis of the Suitability of Wine Grape Cultivars (Terroir)

Decoding terroir data: Analyse key climate, soil and terrain variables to better understand cultivar suitability.

Suitability Tool

Summary Table / Suitability Table

Field: 1

Suitability Score	Cultivar
100	Sauvignon Blanc
70	Cabernet Sauvignon
12	Shiraz
11	Chenin Blanc
9	Pinotage
1	Chardonnay
0	Merlot

Field: 2

Suitability Score	Cultivar
100	Sauvignon Blanc
99	Cabernet Sauvignon
31	Shiraz
24	Pinotage
23	Chenin Blanc
20	Merlot
0	Chardonnay

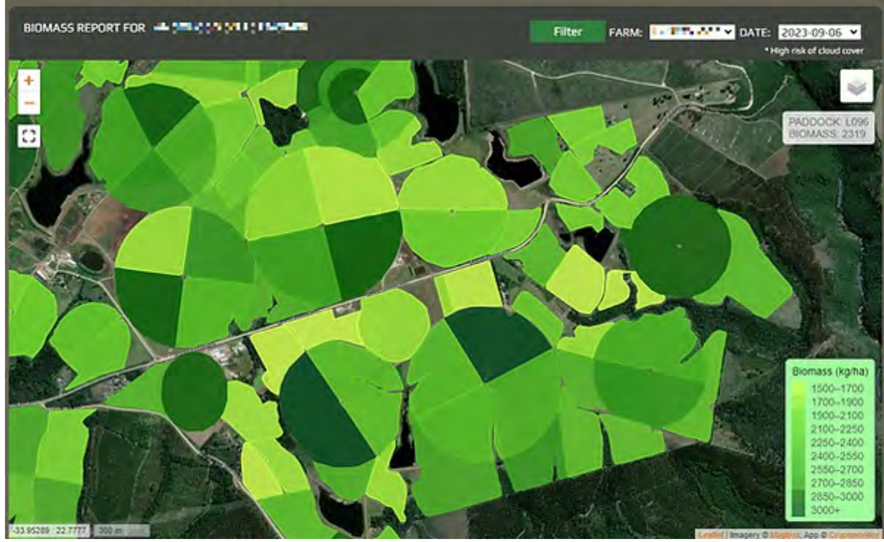
View Profile

TERRACLIM
FORWARD FARMING

SOUTH AFRICA WINE
discover diversity in a glass

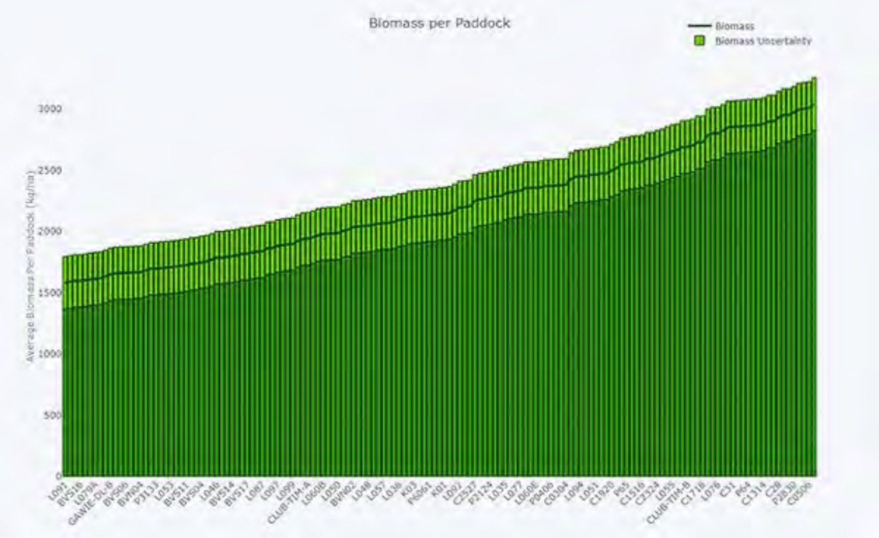
Example 3: Biomass monitoring for pasture management (CropsMonitor)

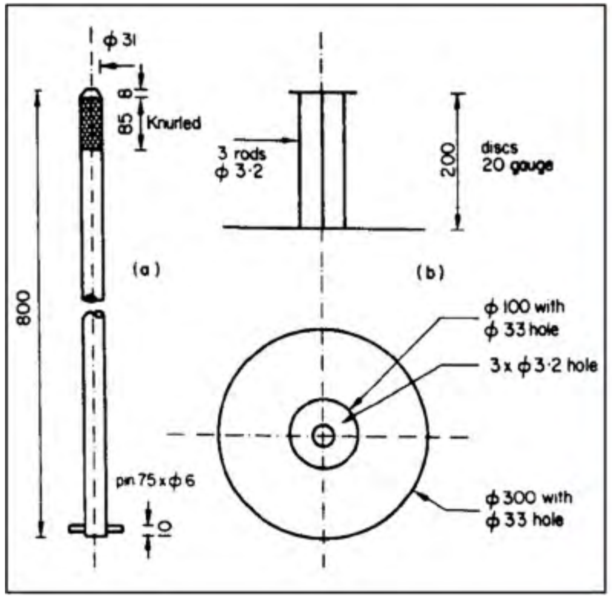
Pasture biomass monitoring



Random forest regression model to estimate pasture biomass per paddock (using satellite imagery) to within 200 kg/ha accuracy.

Used >200000 plate meter (biomass) readings to train and test the model!





TRACE & SAVE SUSTAINABILITY™

FOURTH

cropsmonitor

Plate meter photo credit: https://link.springer.com/article/10.1007/s41064-018-0055-2

Example 4: Yield modelling

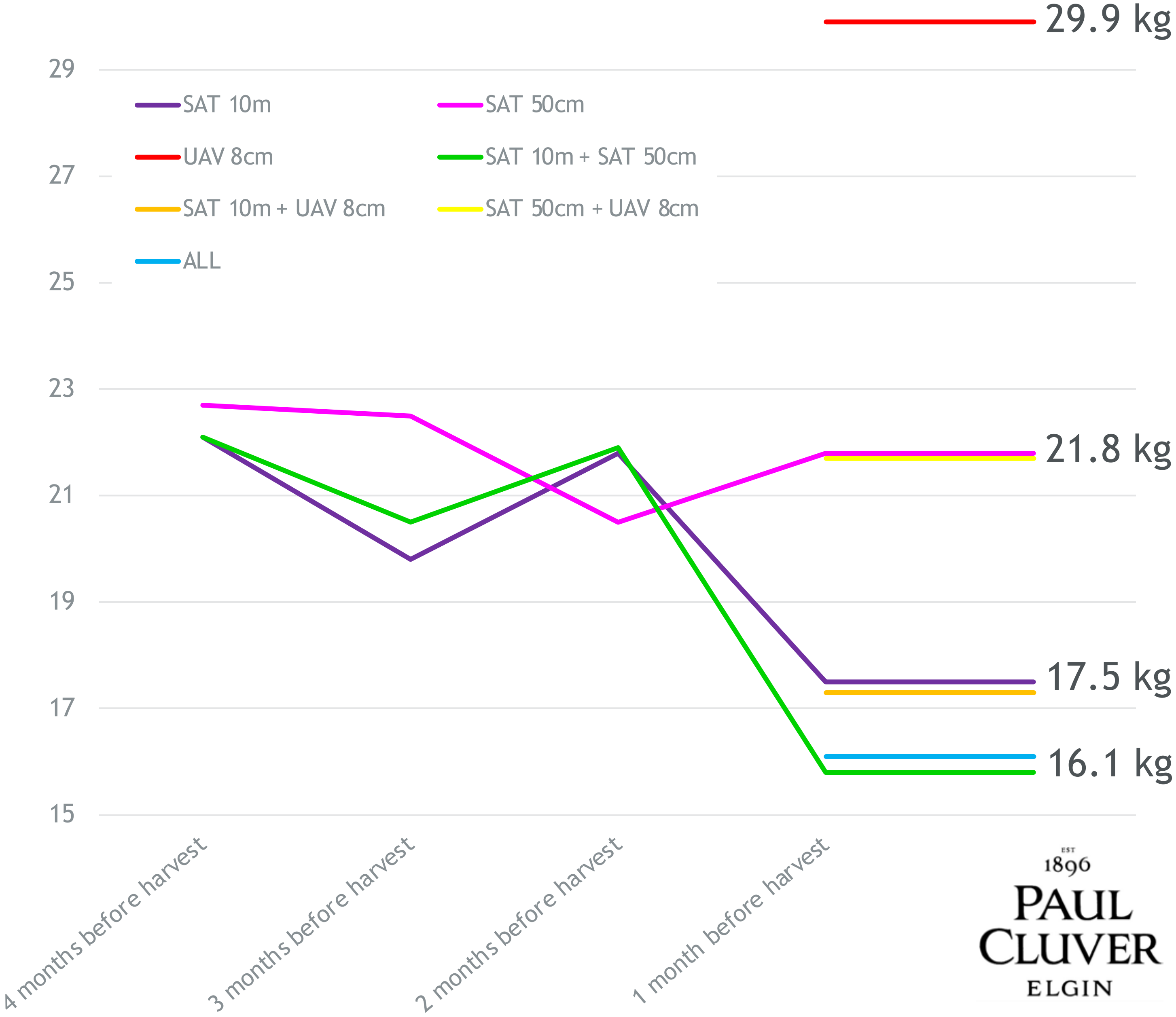
Apple yield modelling

How early in the season can we predict apple yield (kg/tree) using remotely sensed data only?

Yield (kg) from 30 trees in three Fugi orchards

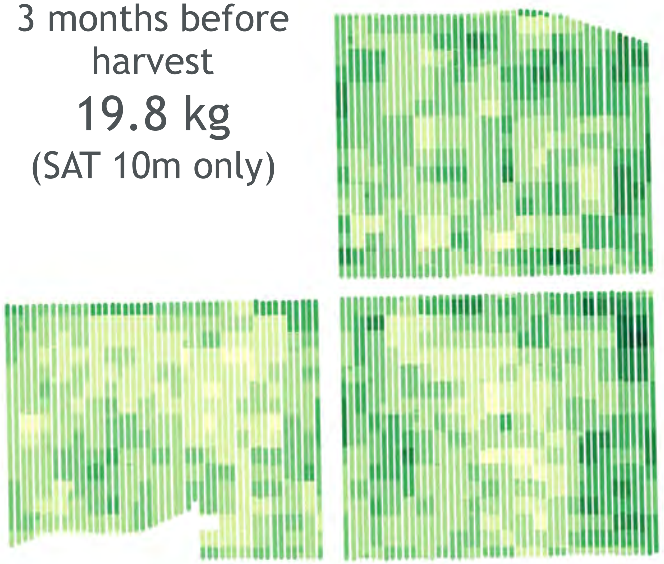
Imagery:
- Drone (UAV) image (x1)
- 50cm satellite images (x4)
- 10 m satellite images (x15)

Algorithm:
Random forest regression

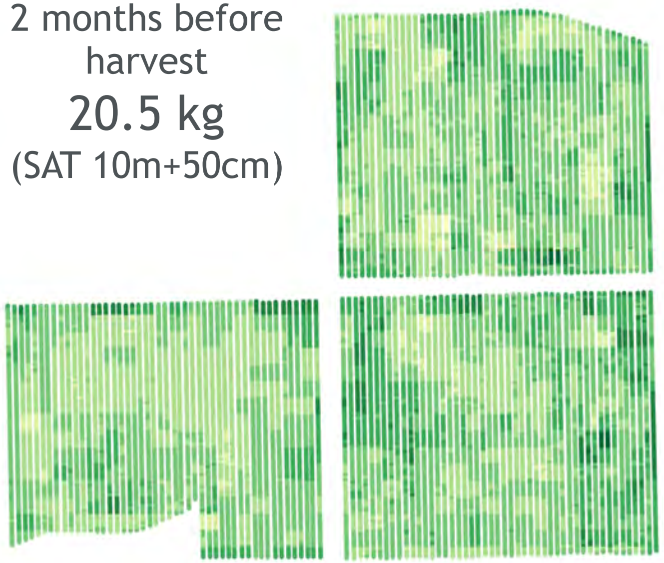


Apple yield (kg/tree) modelling results

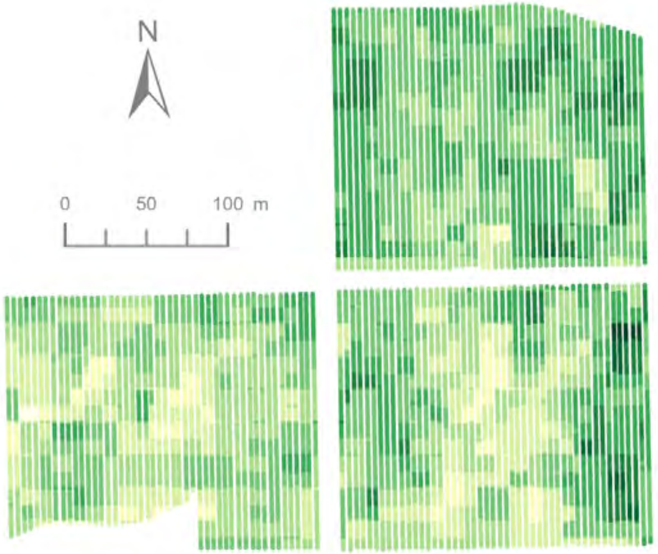
3 months before harvest
19.8 kg
(SAT 10m only)



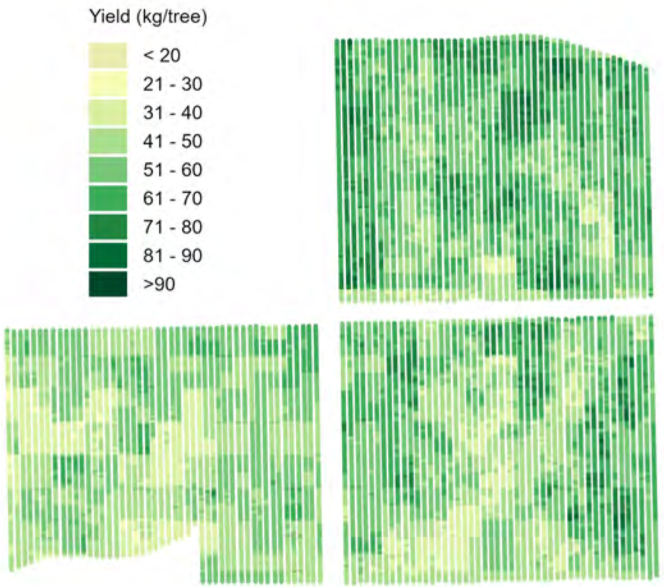
2 months before harvest
20.5 kg
(SAT 10m+50cm)



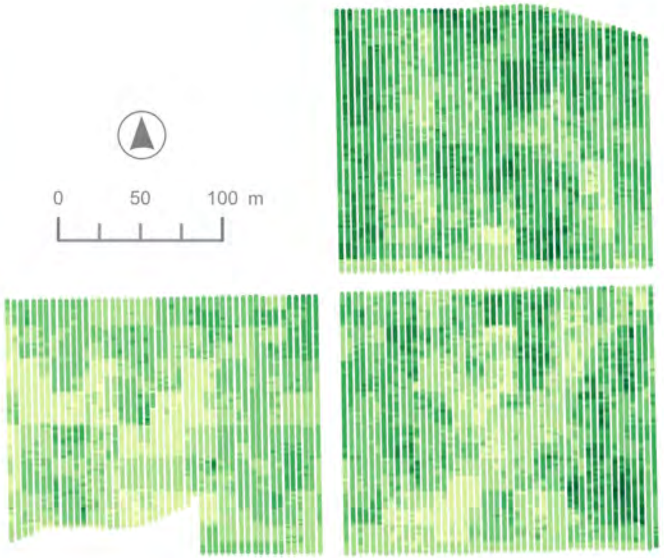
1 month before harvest



17.5 kg (SAT 10m only)



15.8 kg (SAT 10m+50cm)



16.1 kg (SAT 10m+50cm+UAV)

HORTGRO
Growing Fruit IQ

Apple fruit size modelling

Predictor variables:
- Drone images (x4)
- 50cm satellite images (x4)
- 10 m satellite images (x15)

Target classes:
- Soil properties
- Biomass
- Bud counts
- Growth
- Yield (mass, size, quality)

Algorithms:
- Decision trees
- Random forests (RF)
- RF-Regression

HORTGRO Growing Fruit IQ

AGRIMOTION SCIENTIFIC | INDEPENDENT | PRECISE

Pecan nut yield/tree modelling

How accurately can pecan nut yields be predicted using remotely sensed and in situ data?

Farm: Green Valley (Northern Cape)

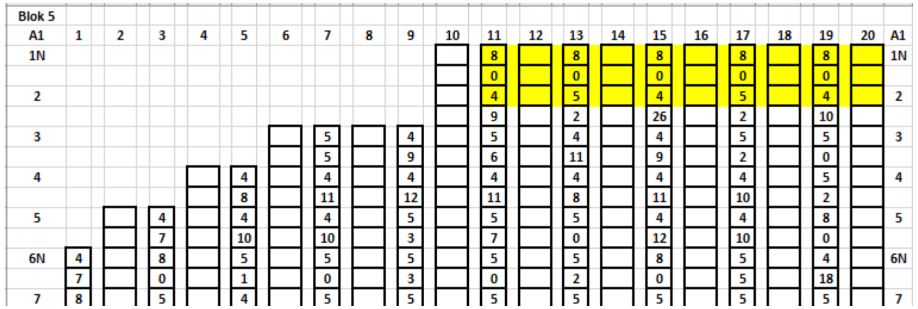
Predictor variables:
- Drone images (x3)
- 10 m satellite images (x15)
- Tree grading (visual assessment)

Target:
- Yield (kg/tree)

Algorithm:
- Random forest (RF) regression

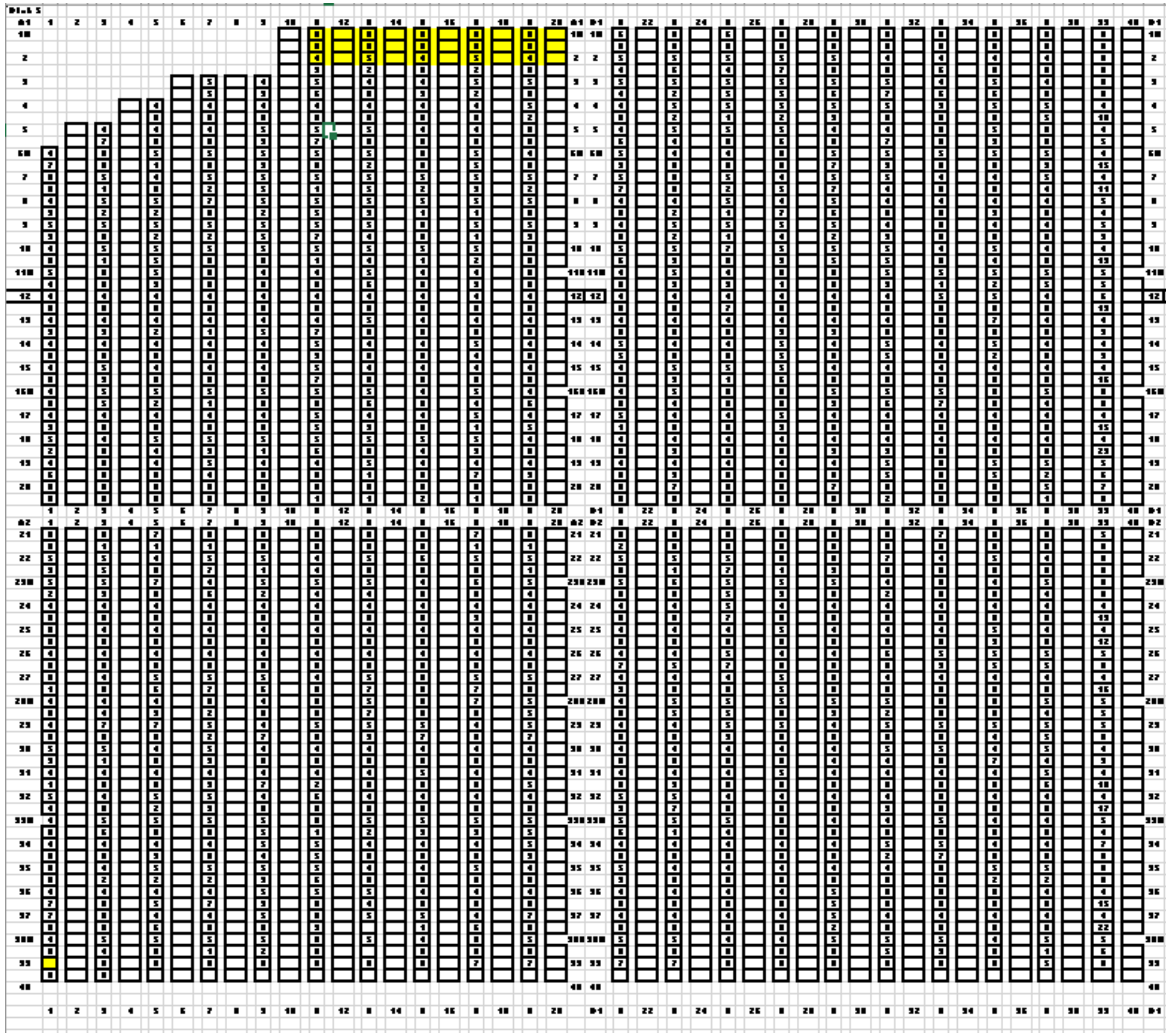
Sample size: 32 424 of 63 056 (51%)

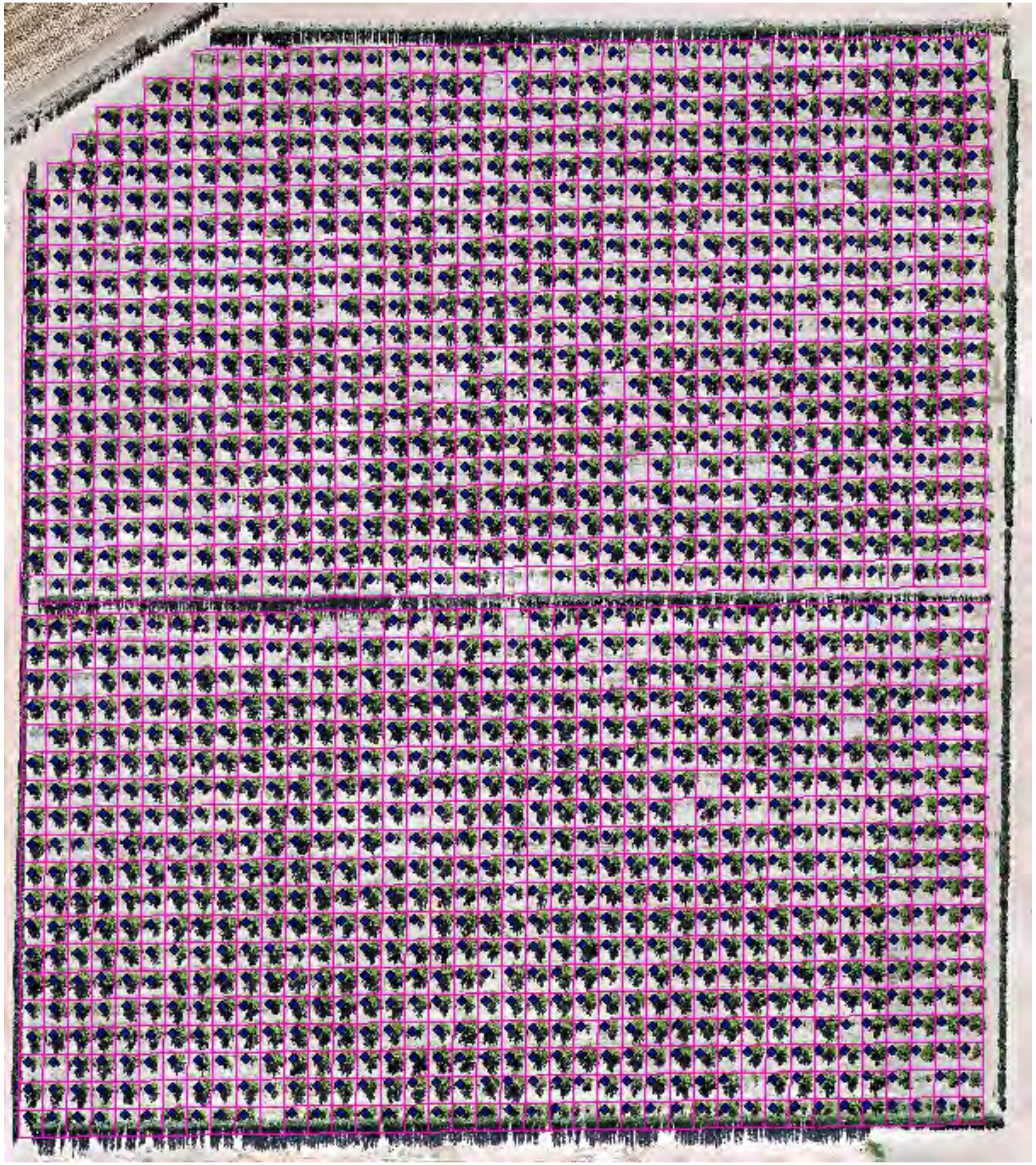

0.2
0.4 Km
0 0.10.2 0.4 Km

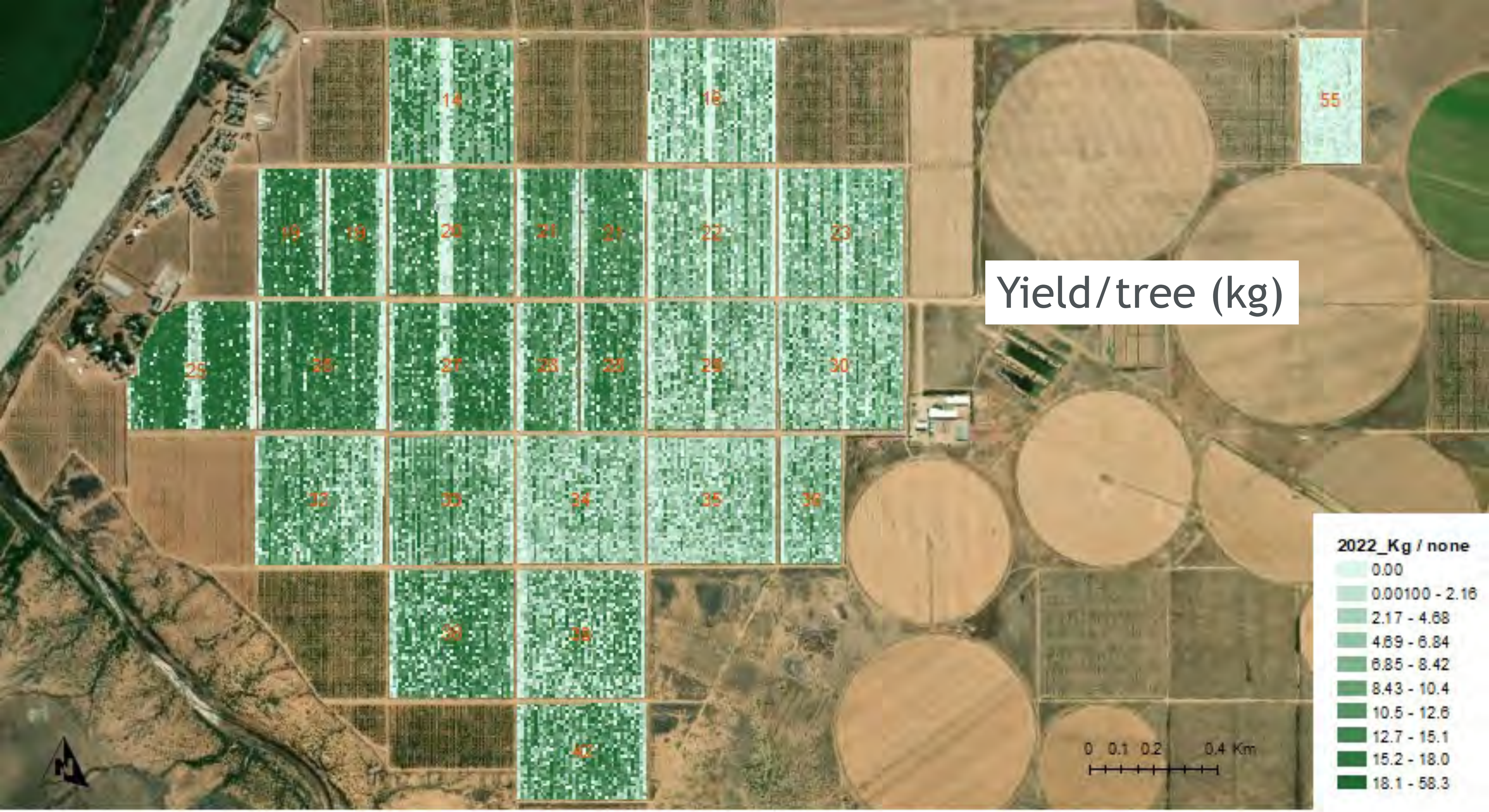


Blok 5

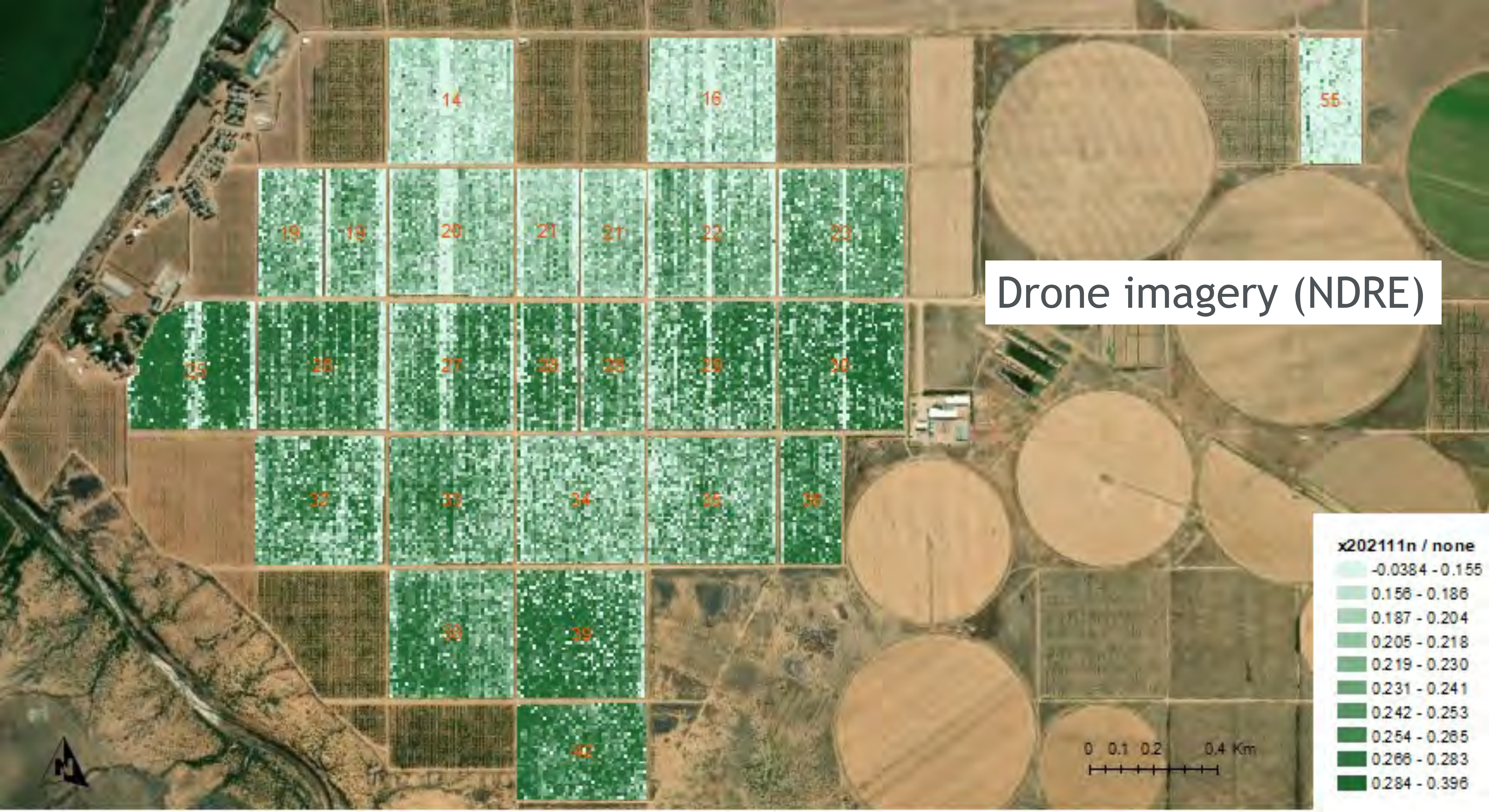
A1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	A1
1N											8		8		8		8		8		1N
											0		0		0		0		0		
2											4		5		4		5		4		2
											9		2		26		2		10		
3							5		4		5		4		4		5		5		3
							5		9		6		11		9		2		0		
4					4		4		4		4		4		4		4		5		4
					8		11		12		11		8		11		10		2		
5			4		4		4		5		5		5		4		4		8		5
			7		10		10		3		7		0		12		10		0		
6N	4		8		5		5		5		5		5		8		5		4		6N
	7		0		1		0		3		0		2		0		5		18		
7	8		5		4		5		5		5		5		5		5		5		7



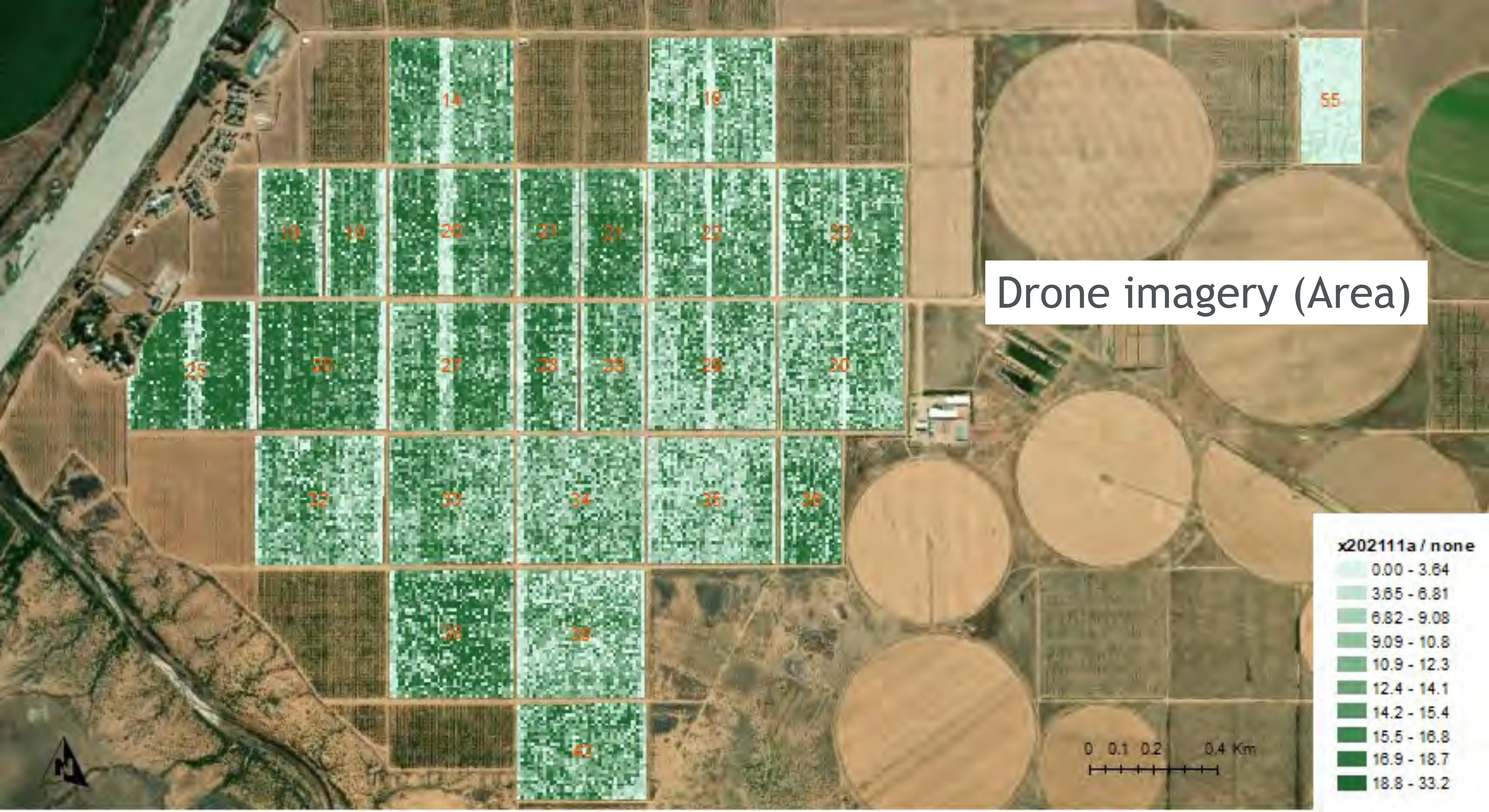



Yield/tree (kg)
14
16
55
19
19
20
21
21
22
23
25
26
27
28
28
29
30
32
33
34
35
36
0 0.1 0.2 0.4 Km
2022_Kg / none
0.00
0.00100 - 2.16
2.17 - 4.68
4.69 - 6.84
6.85 - 8.42
8.43 - 10.4
10.5 - 12.6
12.7 - 15.1
15.2 - 18.0
18.1 - 58.3

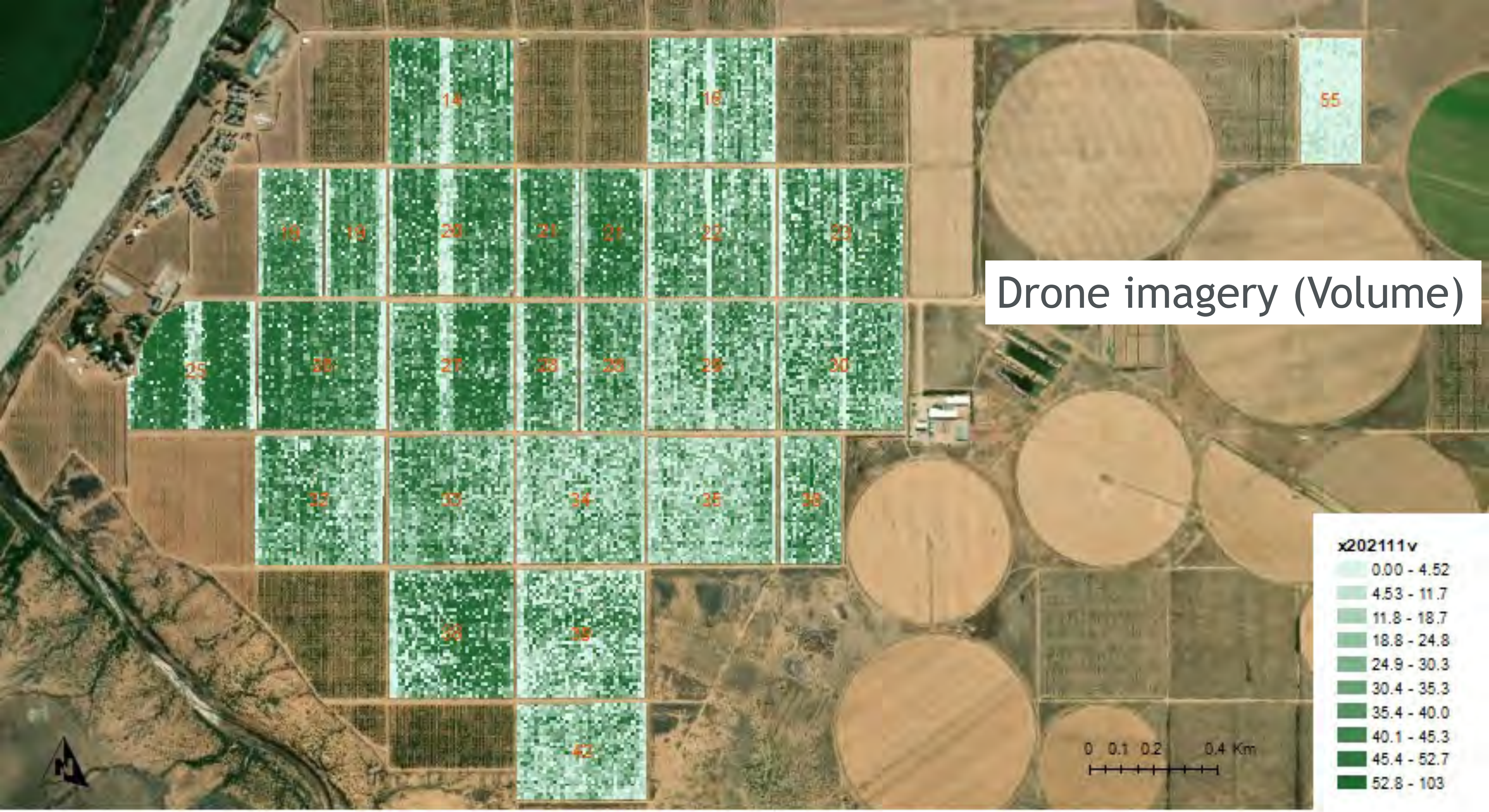
Drone imagery (NDRE)



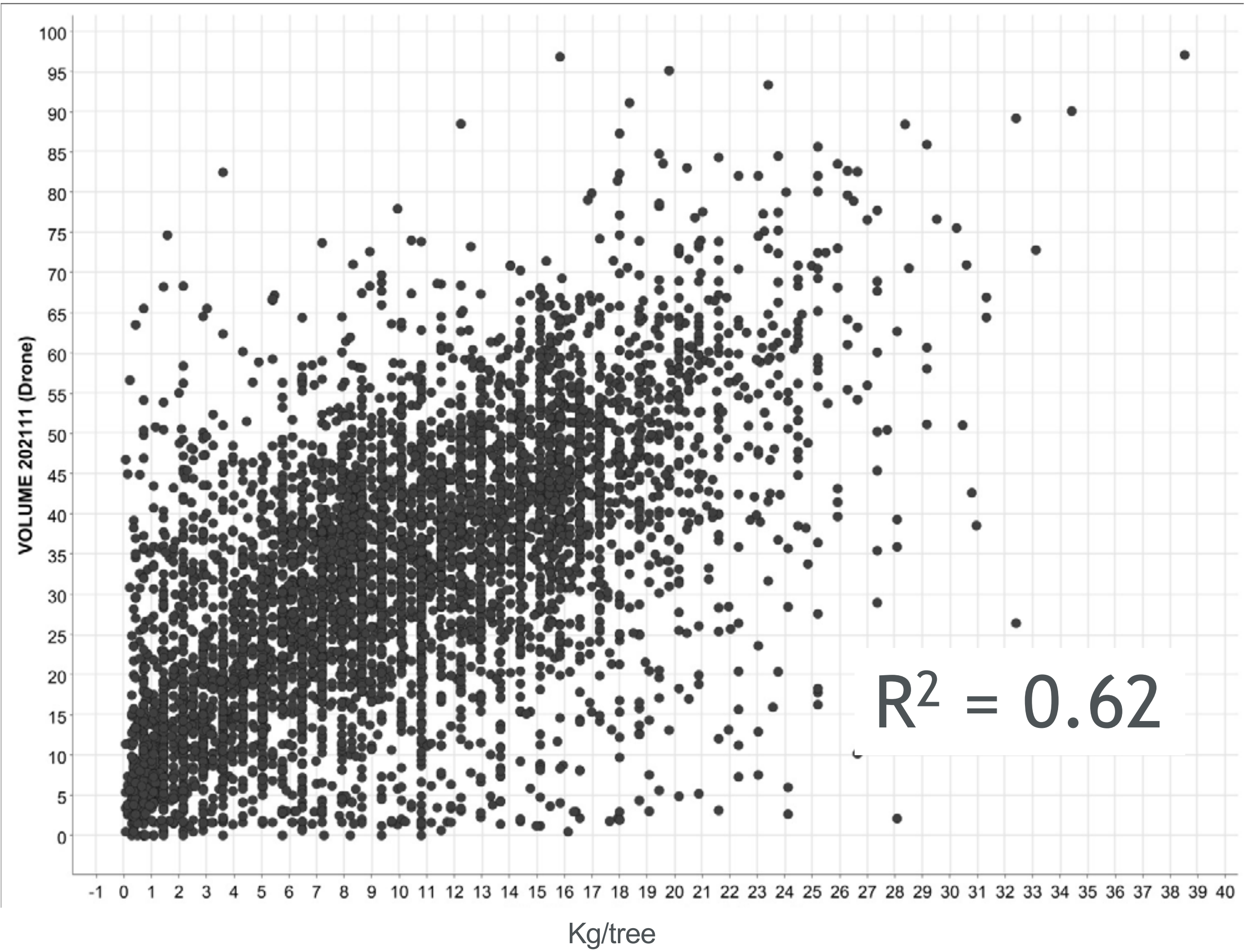
Drone imagery (Area)



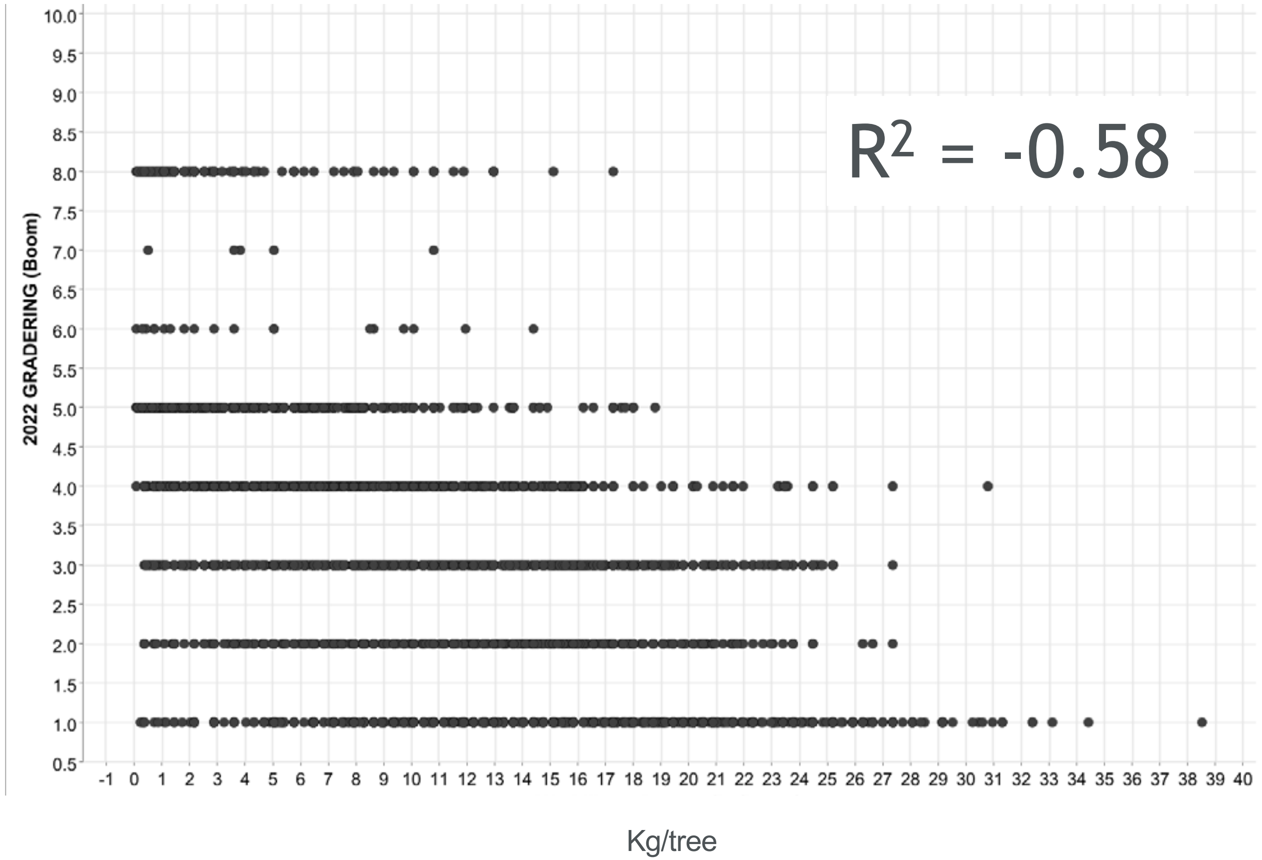
Drone imagery (Volume)



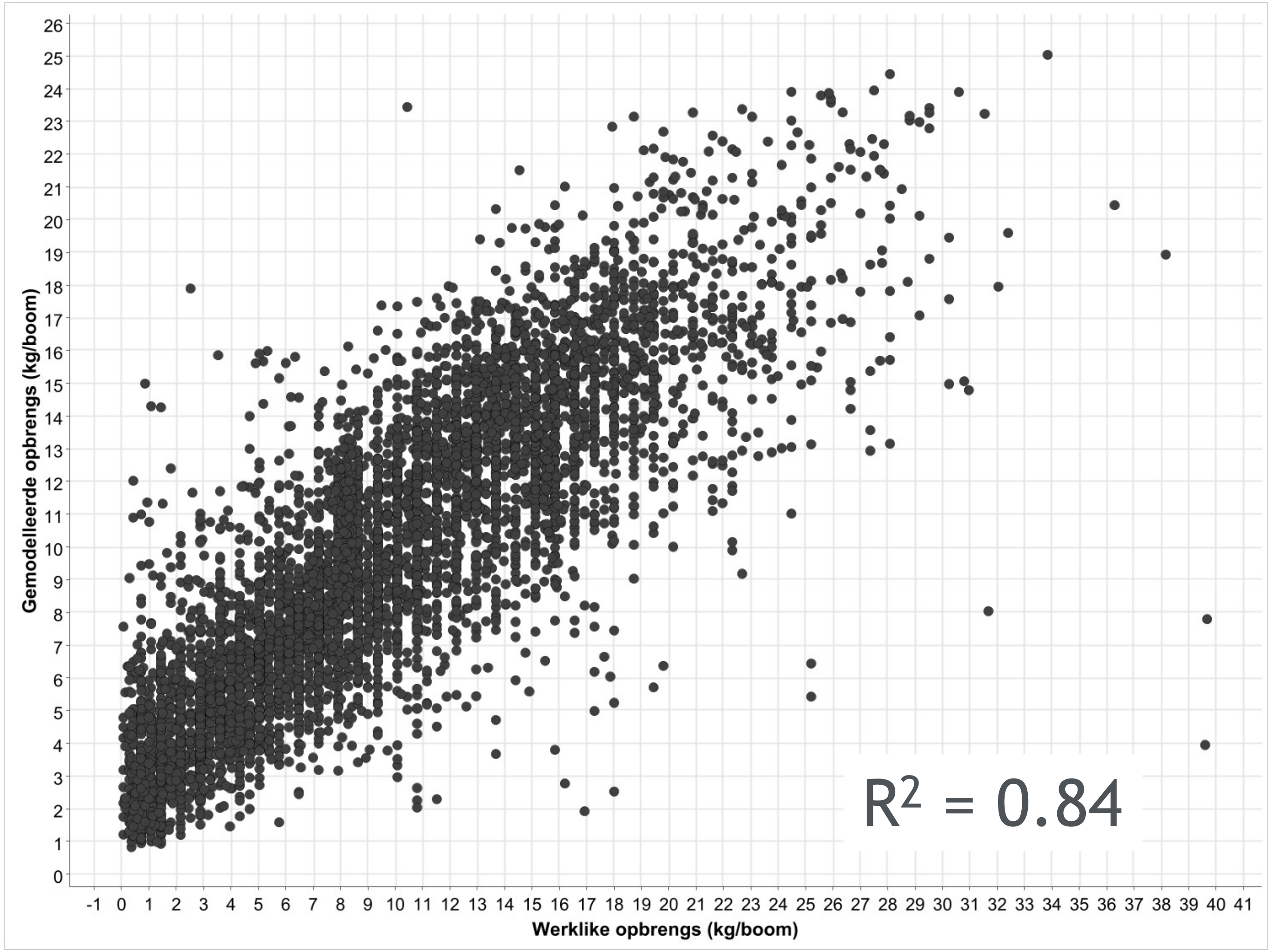
Tree volume (drone) vs kg/tree



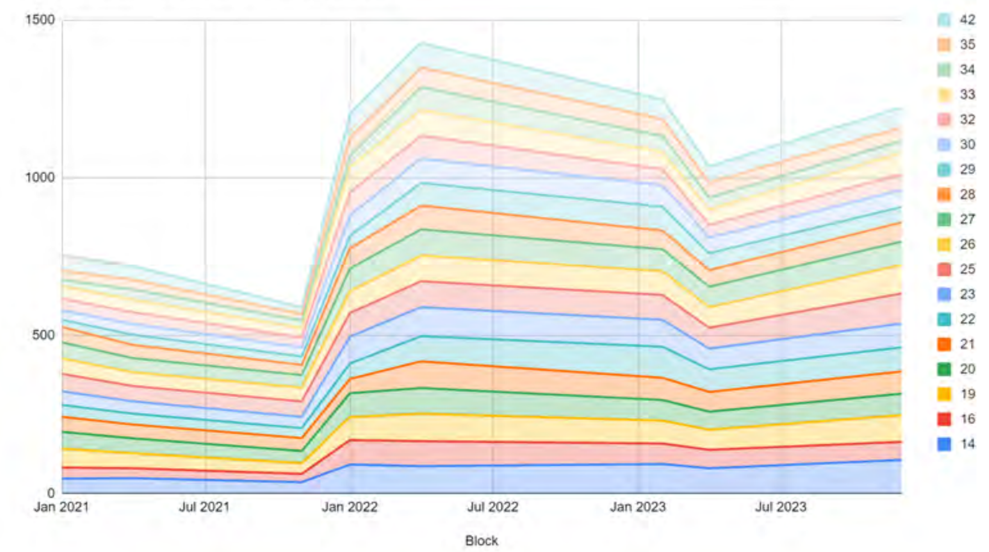
Tree grading (visual) vs kg/tree



Modelled kg/tree vs actual kg/tree

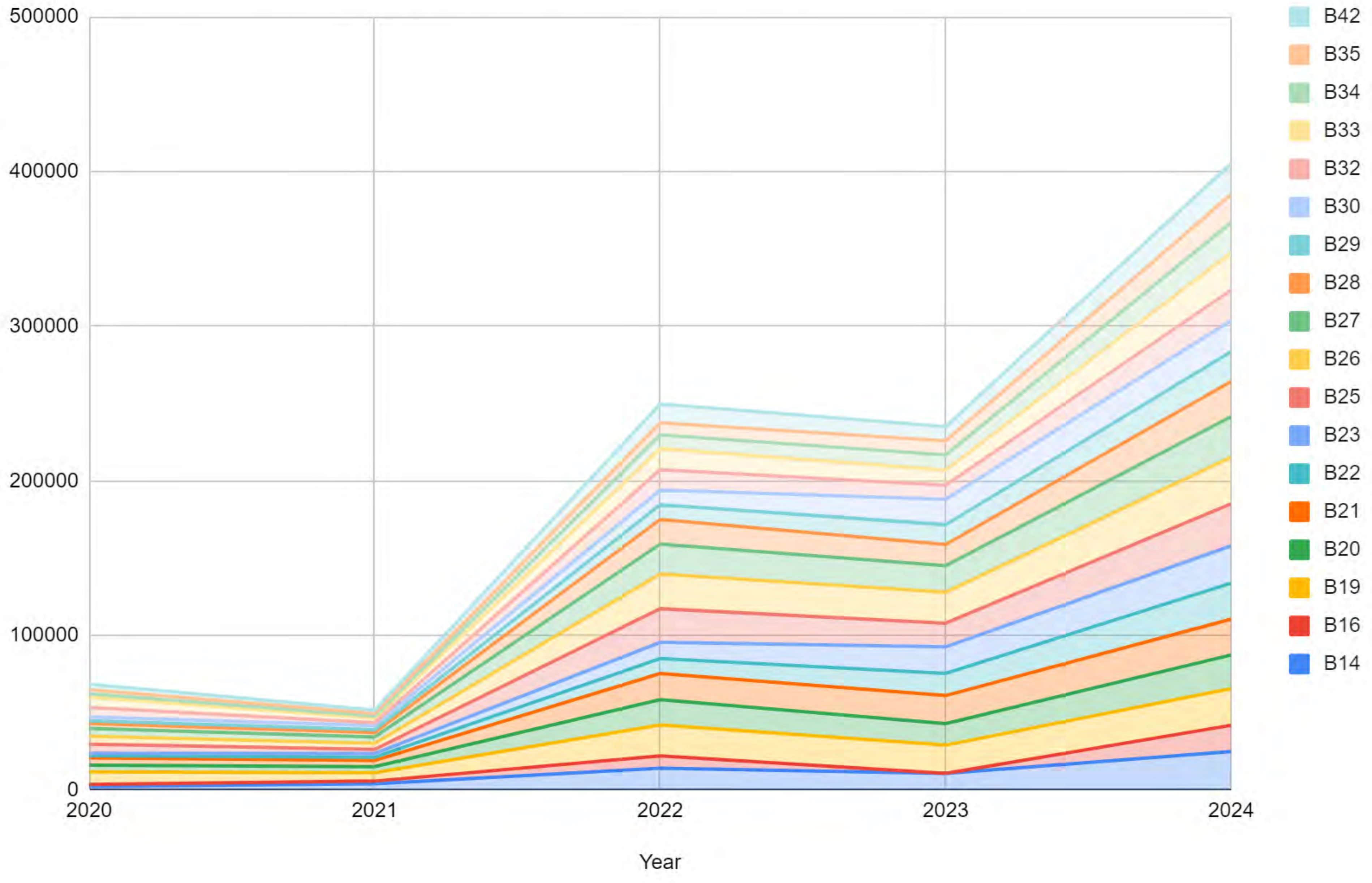


Mean error = 7.8 kg
RMSE = 10.3 kg

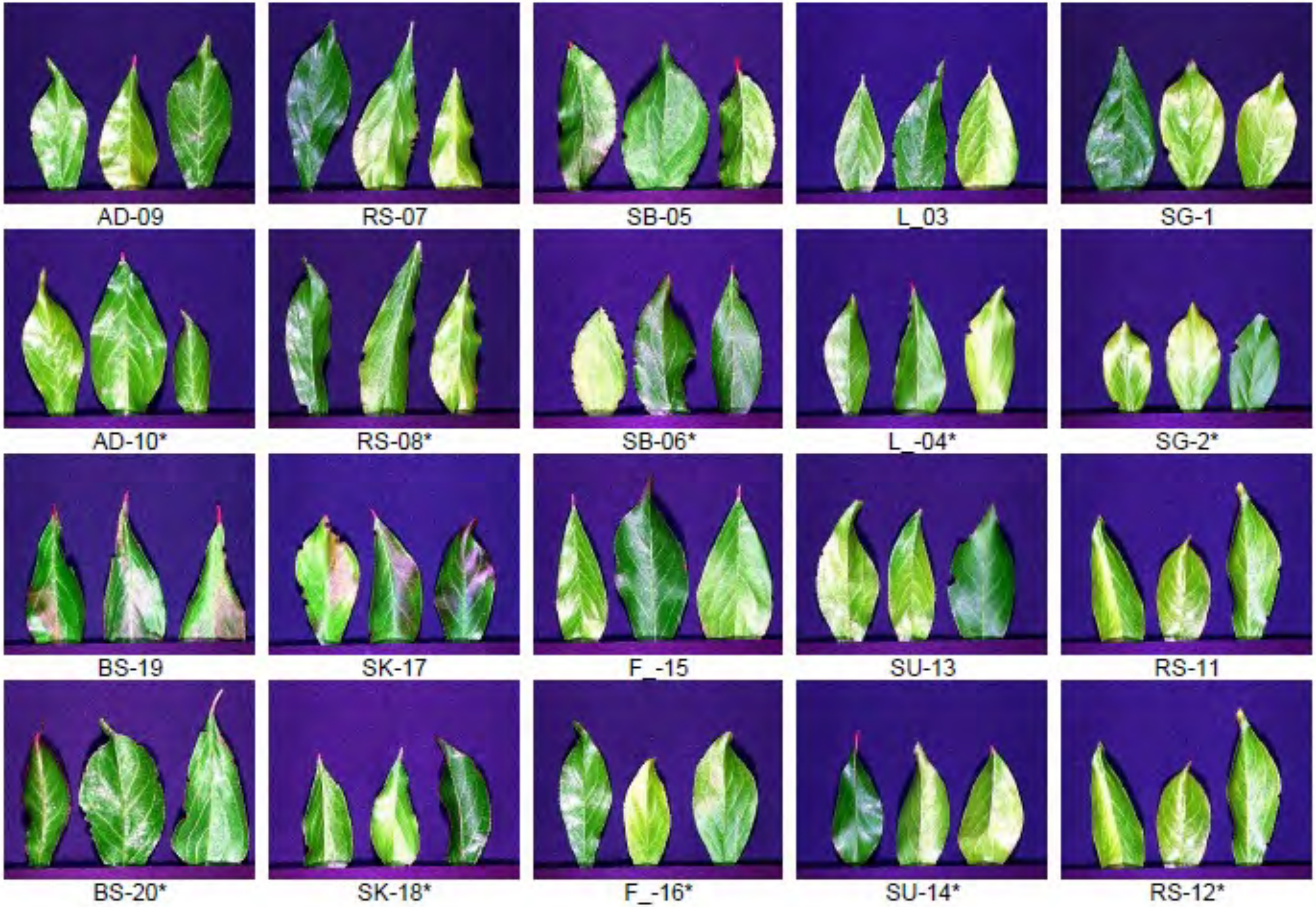


Tree volume

Applying yield model on 2023 drone imagery to predict 2024's yield per block



Example 6: Disease/viroid identification

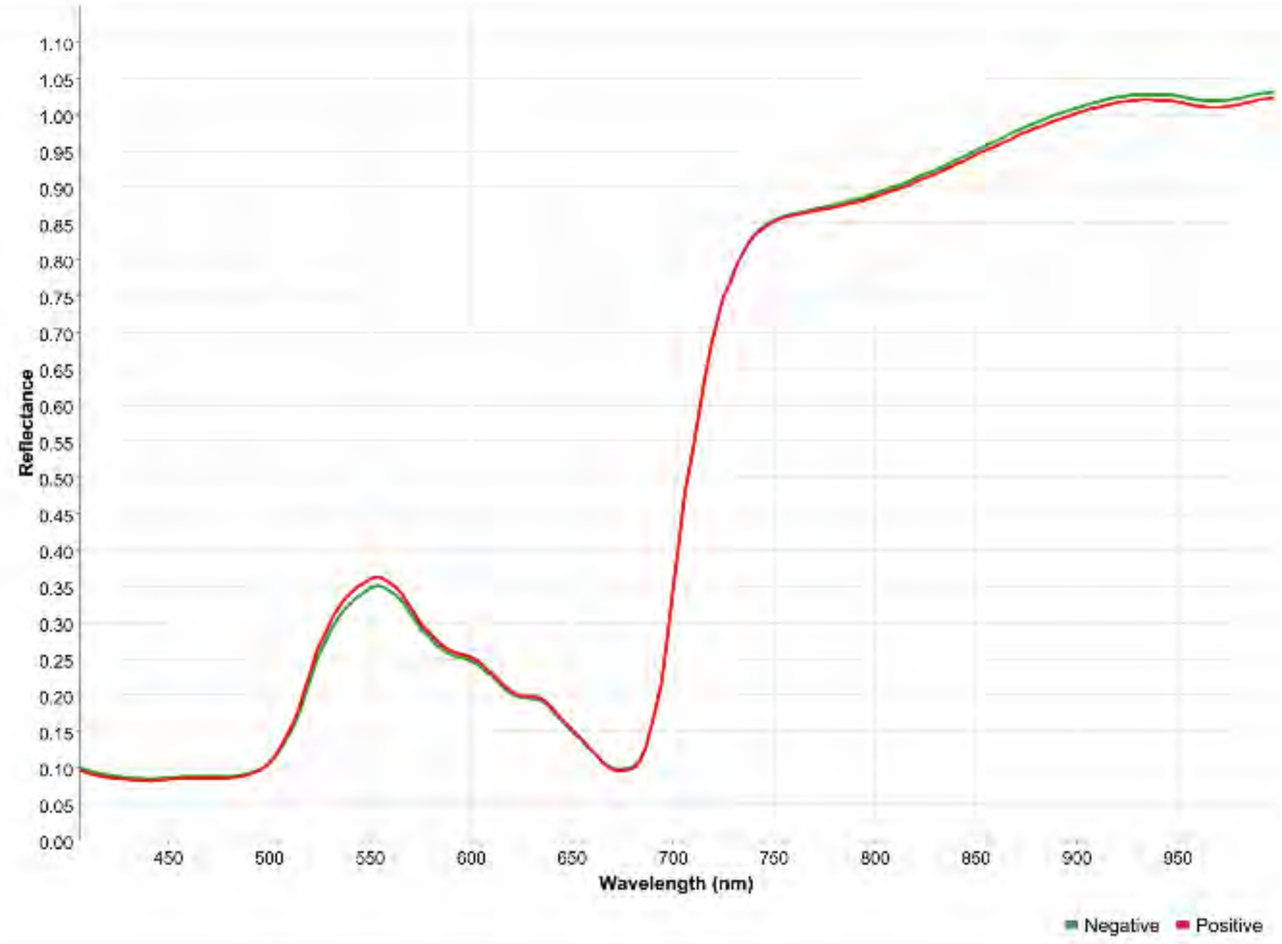


* Inoculated with plum marbling viroid

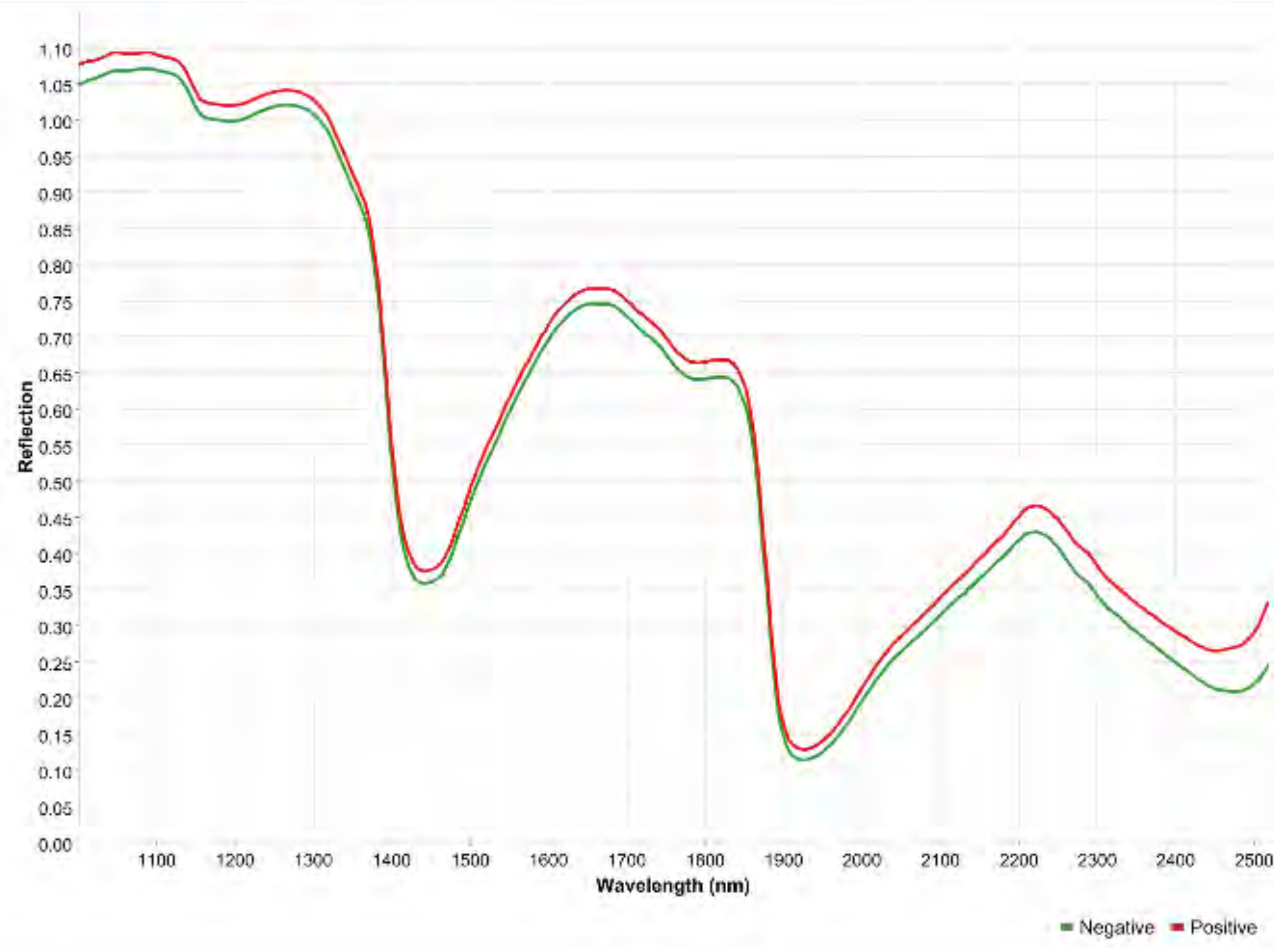
Image segmentation to increase the number of samples

Four different train and test sample selection strategies

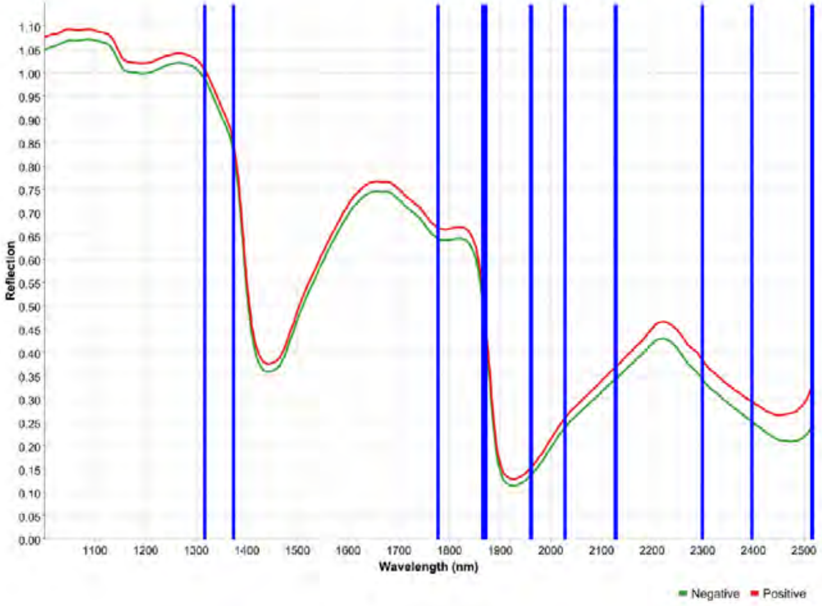
SET	TRAIN / TEST SAMPLE SELECTION STRATEGY	TRAIN# (VNIR)	TRAIN# (SWIR)	TEST# (VNIR)	TEST# (SWIR)
1	Randomly select 20% of segments from all images for independent model testing	2024	1968	506	492
2	Randomly select one leaf per image/bed and use all of the corresponding segments for independent model testing	1704	1658	826	812
3	Use all of the segments from the western column of beds for independent model testing	2160	2112	370	348
4	Use all of the segments from the eastern column of beds for independent model testing	2009	1945	521	515



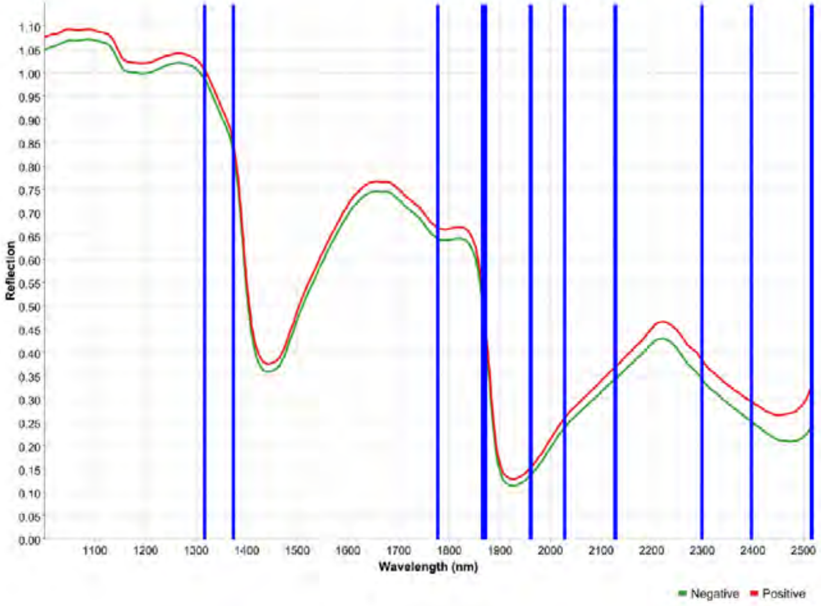
Visible and near-infrared
(best model: 72%)



Short-wave infrared
(best model: 97%)



12 most important bands
(best model: 97%)



4 most important bands
(best model: 95%)

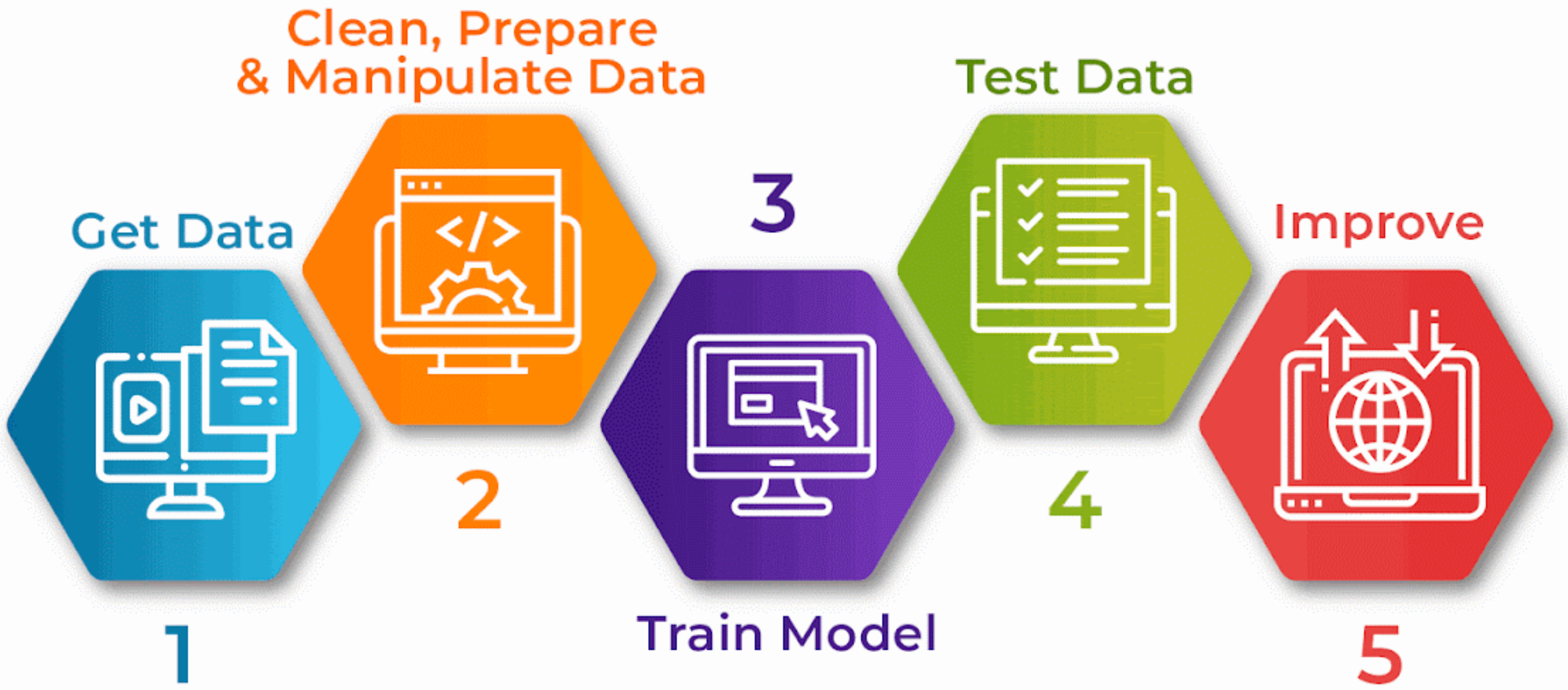
2 most important bands
(best model: 94%)

Conclusion

- Machine learning can be very effective for making sense of huge amounts of (agricultural) data
- It is particularly adept at finding complex relationships between sets of data (e.g. yield and remotely sensed imagery)
- But it requires enough training data from which to learn
- 60% of the effort is in collecting (enough) training data

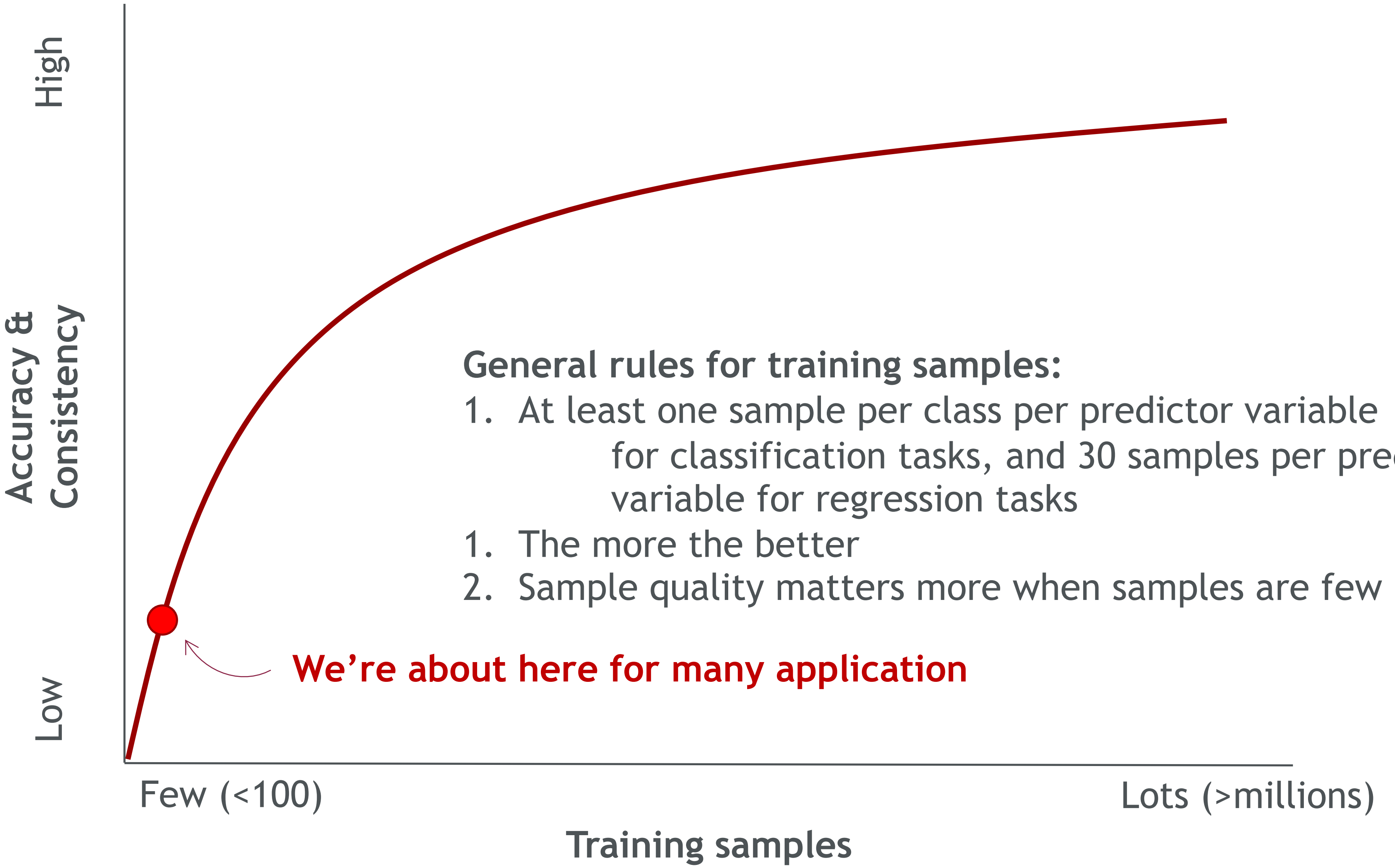
Source: https://www.wurweb.nl/english/killer-robots-far-from-being-a-realistic-threat

Supervised machine learning process



60% 30% 2% 3% 5%

A limitation of machine (and deep) learning is that it requires a large number of training samples (because geographical variation is high)



General rules for training samples:

1. At least one sample per class per predictor variable for classification tasks, and 30 samples per predictor variable for regression tasks
1. The more the better
2. Sample quality matters more when samples are few

https://starsandpopcorn.com/movie-mad-scientists/2/

- We are just scratching the surface of how machine learning can benefit agriculture
- There is still much research to be done
- No doubt that AI holds many benefits for agriculture…

avn@sun.ac.za