science & technology
Department: Science and Technology
REPUBLIC OF SOUTH AFRICA

agriculture & rural development
Department: Agriculture and Rural Development
PROVINCE OF KWAZULU-NATAL

sansa
SOUTH AFRICAN NATIONAL SPACE AGENCY

Sentinel 2 derived Biophysical parameters for yield modelling

Clement Adjorlolo,
cadjorlolo@sansa.org.za

Clement Adjorlolo
SANSA Senior Scientist: Remote Sensing Research
http://www.sansa.org.za

Mahlatse Kganyago,
SANSA Scientist: Remote Sensing Research
http://www.sansa.org.za

Nosiseko Mashinyi,
SANSA Scientist: Remote Sensing Research
http://www.sansa.org.za

Jonathan Atkinson
Professional Agricultural Scientist: Soils, GIS & RS (Pr. Sci. Nat)(PrGISc)
KZN Dept. Agric & Rural Development
Jon.Atkinson@kzndard.gov.za

Kurt Barichievy
Professional Agricultural Scientist: Soil & Hydrology (Pr. Sci. Nat)
KZN Dept. Agric & Rural Development
Kurt.barichievy@kzndard.gov.za

INTRODUCTION

- **Objective**: to validate Biophysical products retrieved from satellite observation for crop yield modelling,

  - from radiance to reflectance, to Leaf Area Index (**LAI**) and leaf/canopy chlorophyll content (**Cab**), FAPAR.

- LAI relates to standing biomass, leaf properties - chlorophyll, NPP and canopy water (e.g. site-specific ET,..).

- LAI and Cab:
  - are important variables for agro-ecological applications (crops/grasslands/forests, etc.)
  - can be inferred from both satellite observations or ground-based indirect radiative transfer (RT) approaches

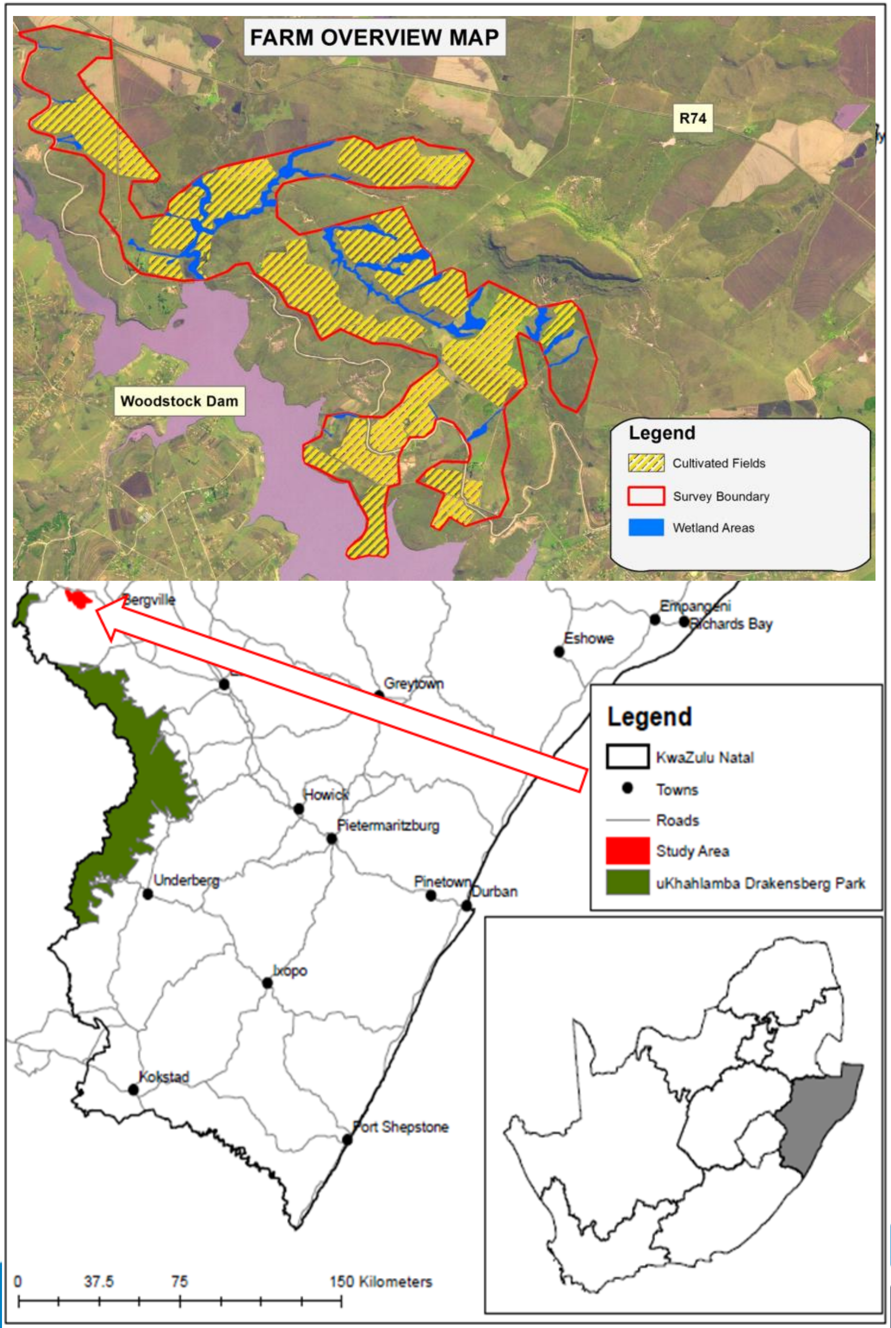
- LAI inferred is a one-sided leaf area (m^2) per unit area (m^2) of ground
  - dimensionless.

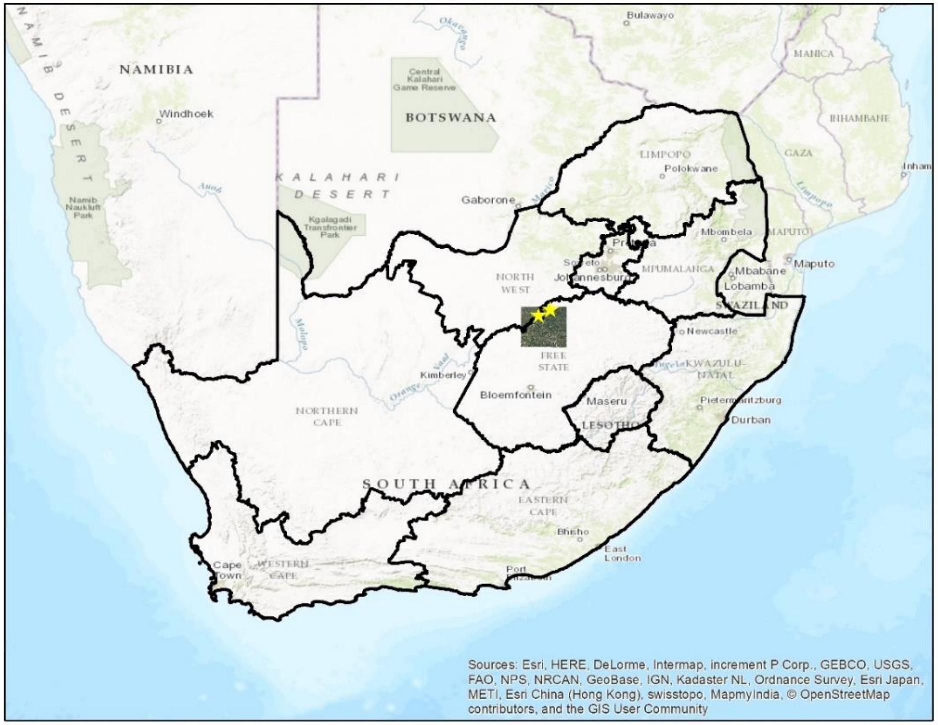
Important Notes

- LAI from EO or ground-based in situ instruments is a function of the radiative transfer (RT) approach used to retrieve it,
  - LAI (EO, 1D RT) ≠ LAI (EO, 3D RT) ≠ LAI (field) ≠ LAI real!!!!
  - scaling is an important factor
  - So must be calibrated against field observation

- ESA SNAP Toolbox
- Biophysical Processor: LAI/Cab
- Sen2cor pre-processing script for retrieving surface reflectances.

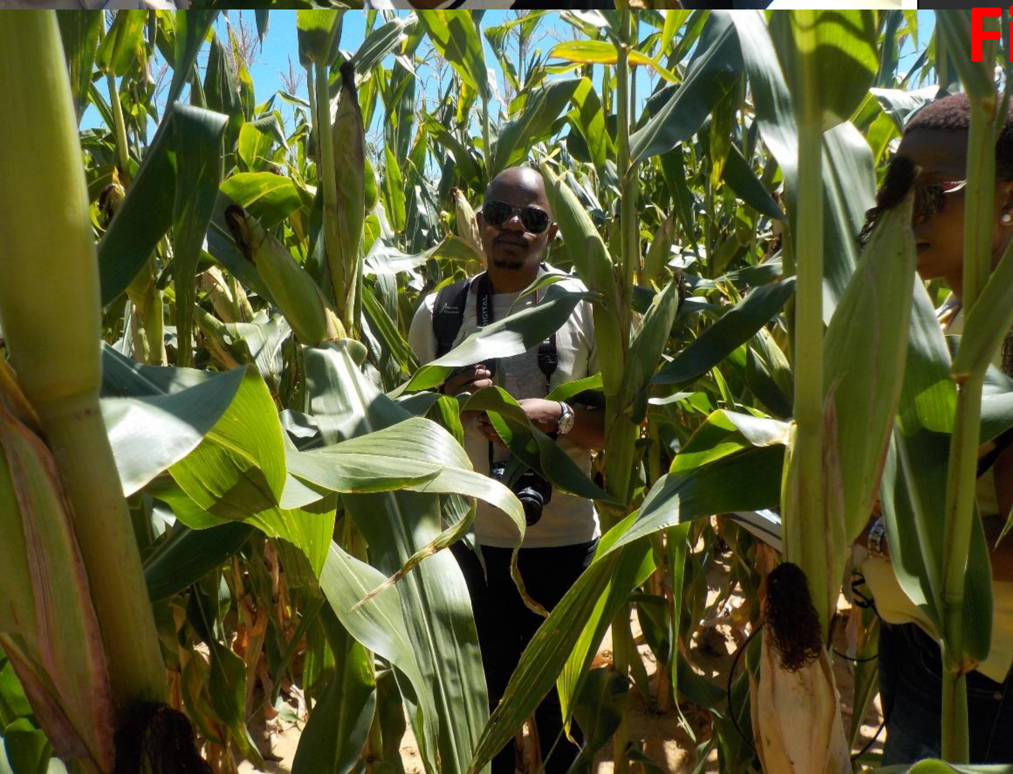
Calibration/Validation Test Sites

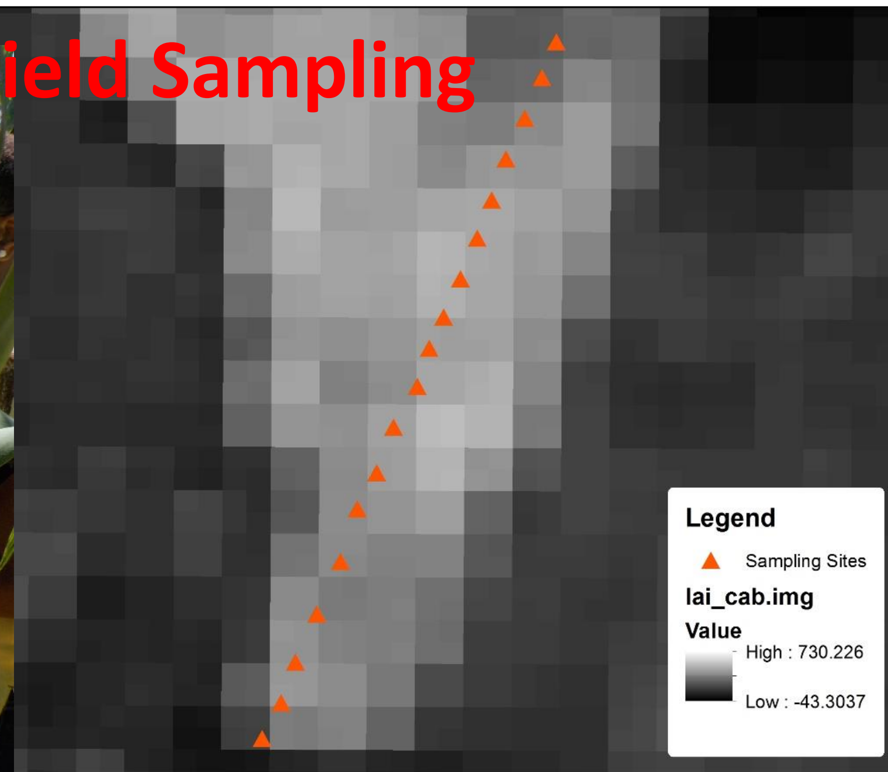




Methods: Fieldwork

Field Sampling





Back in the office: Analyses

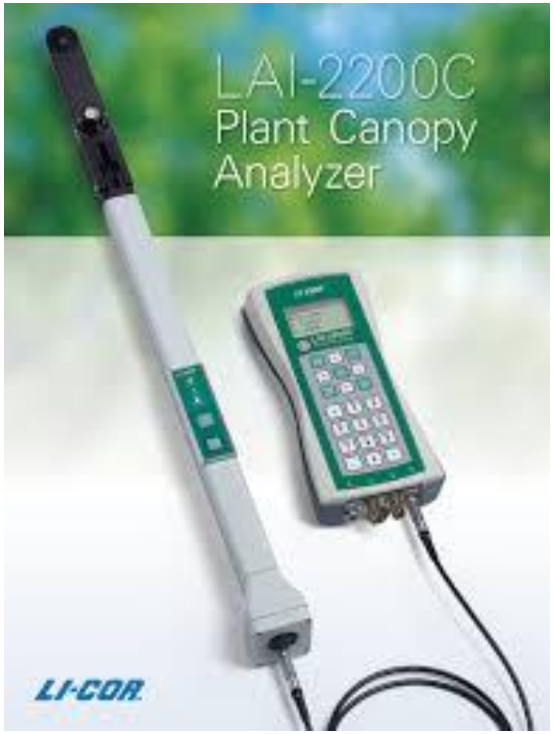
Download field data

- Canopy Spectra
- LAI
- CChl
- Precision GPS

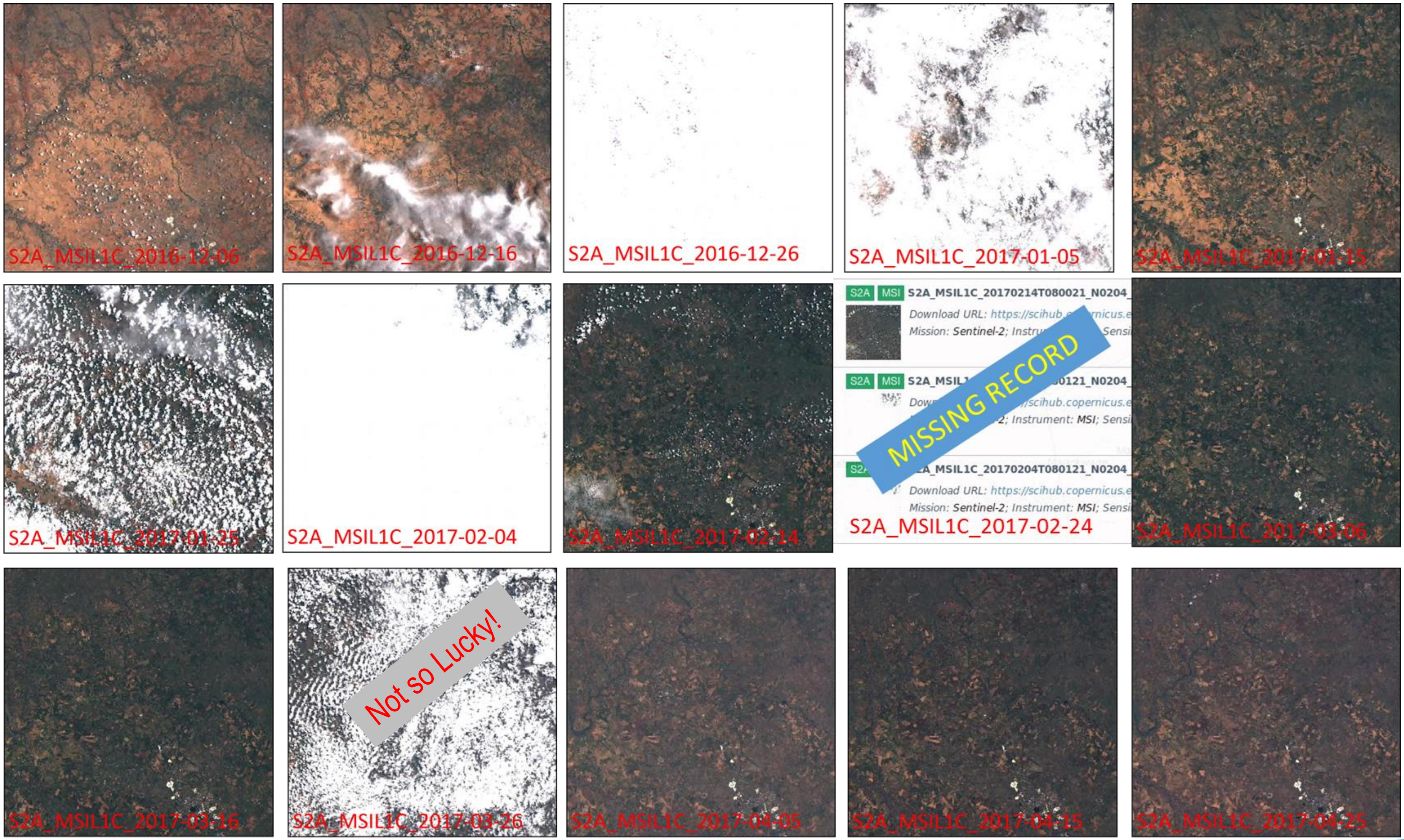
Pre-processing:

- resample field spectra to Sentinel-2 bands using the Sentinel-2A MSI Spectral Responses from ESA

- Create shapefiles of field measured LAI and CChl data



Satellite Data: S2A MSI



Data pre-processing: Sen2Cor

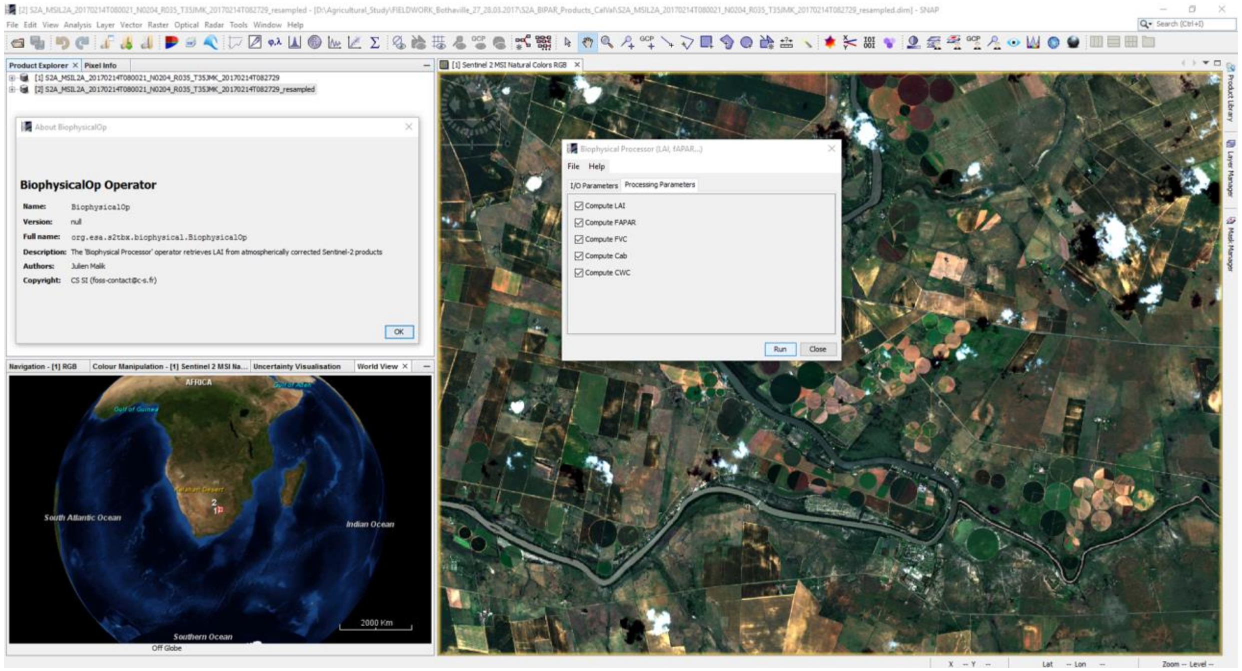
Sen2Cor Configuration and User Manual

Ref. S2-PDGS-MPC-L2A-SUM-V2.3


```xml
<?xml version="1.0" encoding="UTF-8"?>
<Level-2A_Ground_Image_Processing_Parameter xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="L2A_GIPP.xsd">
  <Common_Section>
    <Log_Level>INFO</Log_Level>
    <!-- can be: NOTSET, DEBUG, INFO, WARNING, ERROR, CRITICAL -->
    <Nr_Processes>AUTO</Nr_Processes>
    <!-- can be an unsigned integer value specifying the number or processes you intend to operate in parallel or: AUTO. If AUTO is chosen, the processor determines the number of processes automatically, using cpu_count() -->
    <Target_Directory>DEFAULT</Target_Directory>
    <!-- should be either a directory or 'DEFAULT'. If default, target will be created at root of L1C product -->
    <DEM_Directory>NONE</DEM_Directory>
    <!-- should be either a directory in the sen2cor home folder or 'NONE'. If NONE, no DEM will be used -->
```

Biophysical Processor: LAI and Cab

- “one input layer, made of 11 normalized input data : B3, B4, B5, B6, B7, B8a, B11, B12, cos(viewing_zenith), cos(sun_zenith), cos(relative_azimuth_angle)
- one hidden layer with 5 neurons with tangent sigmoid transfer functions
- one output layer with a linear transfer function”

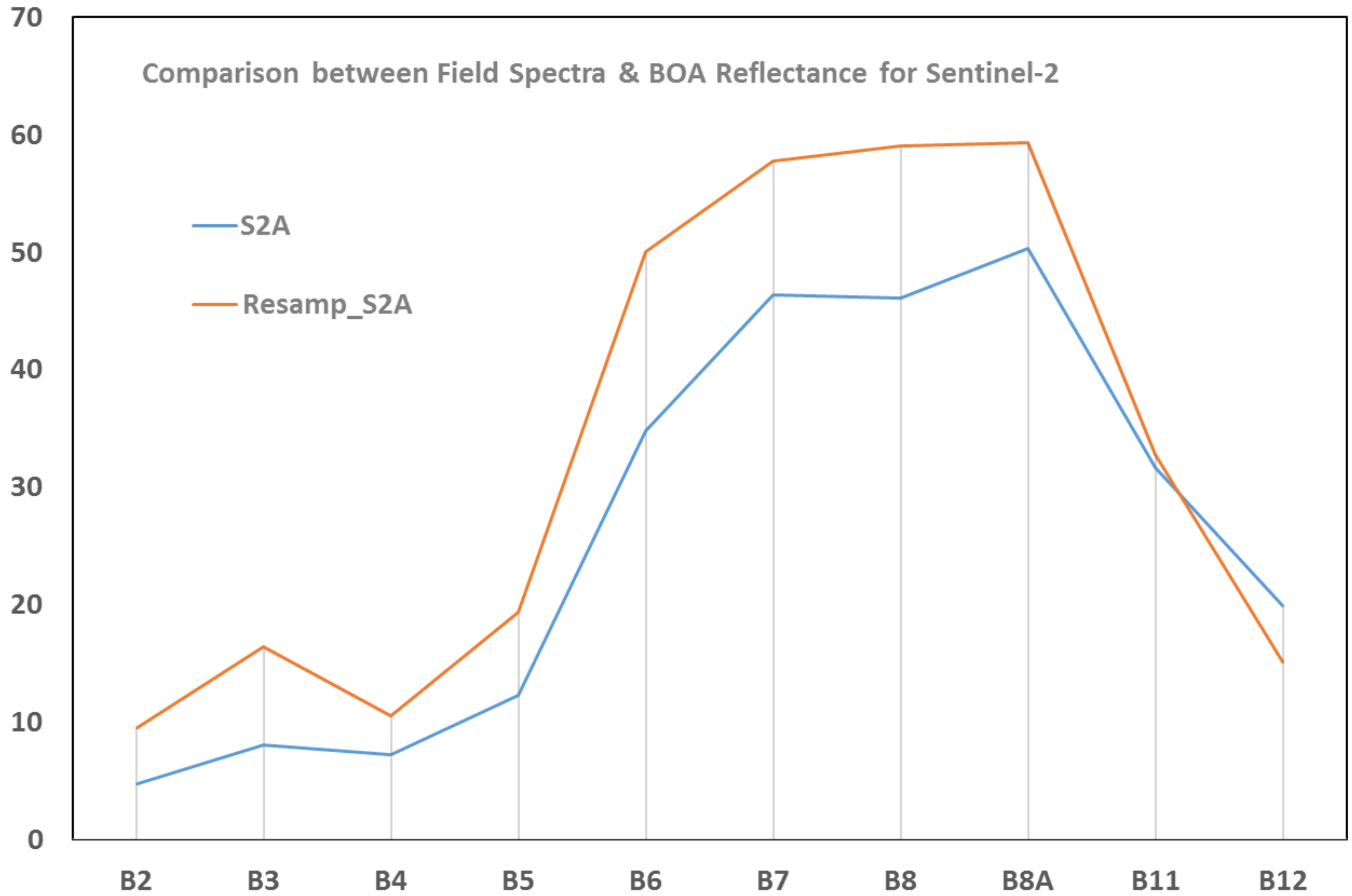


- “The actual algorithm running in SNAP runs the prediction step of the neural network, from the set of precomputed coefficients computed during the training phase.”

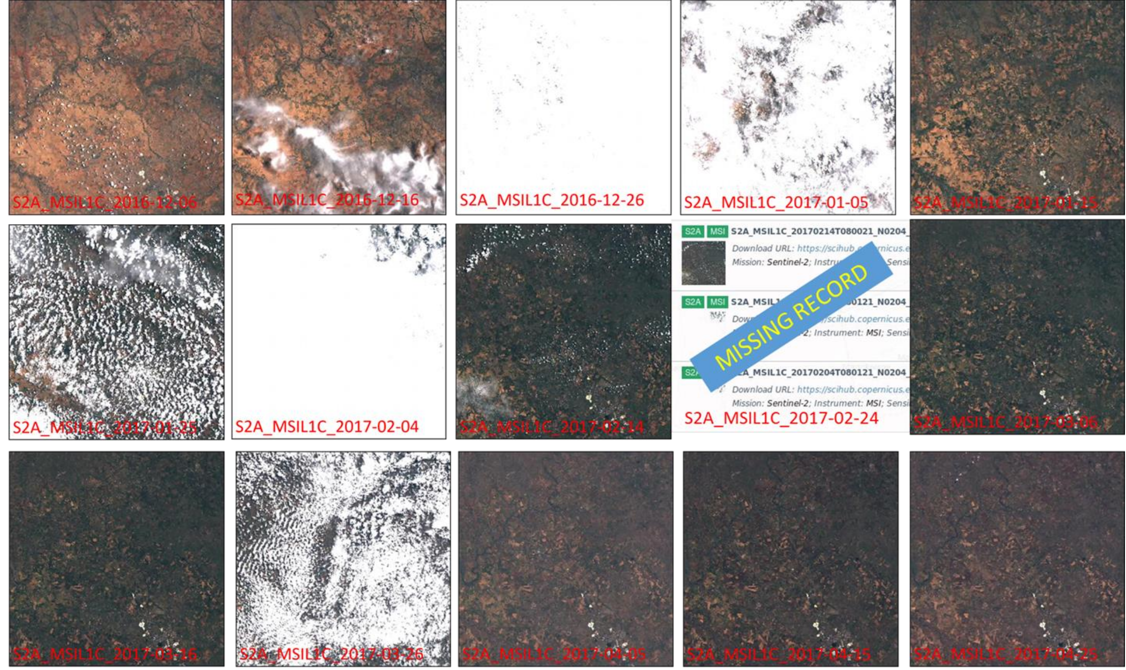
Source: from the ESA SNAP Toolbox – help

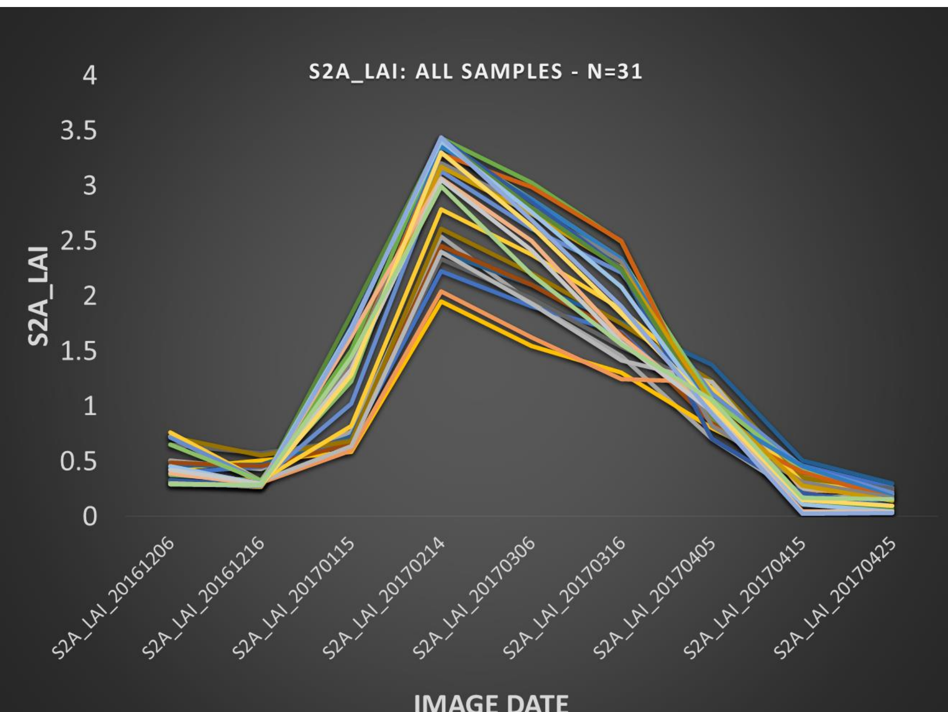
Results

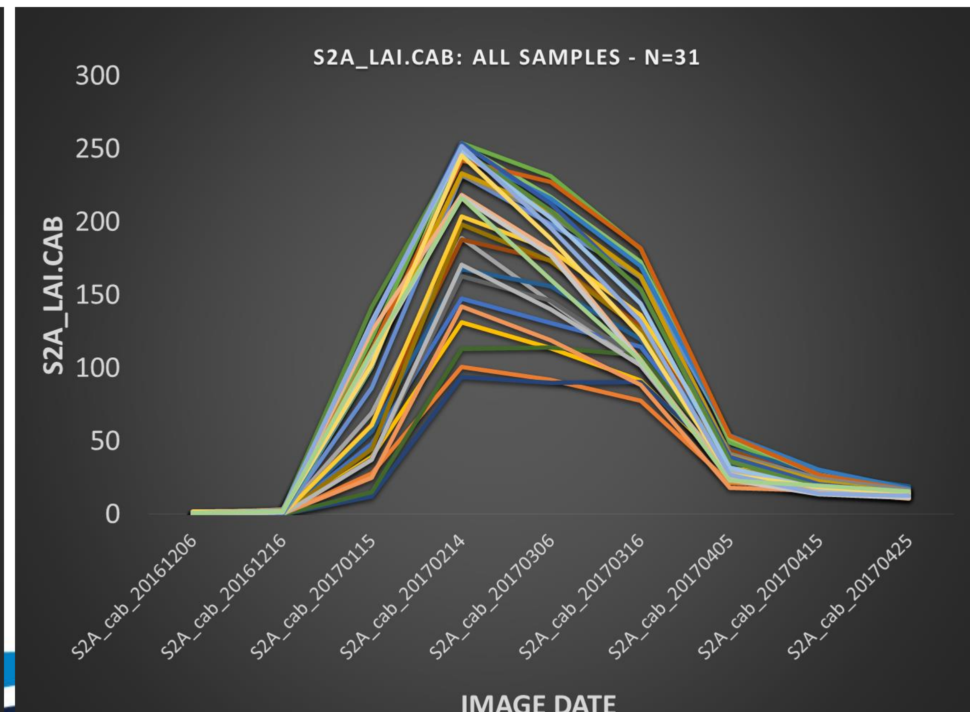
Comparison between Field Spectra & S2A-Bands



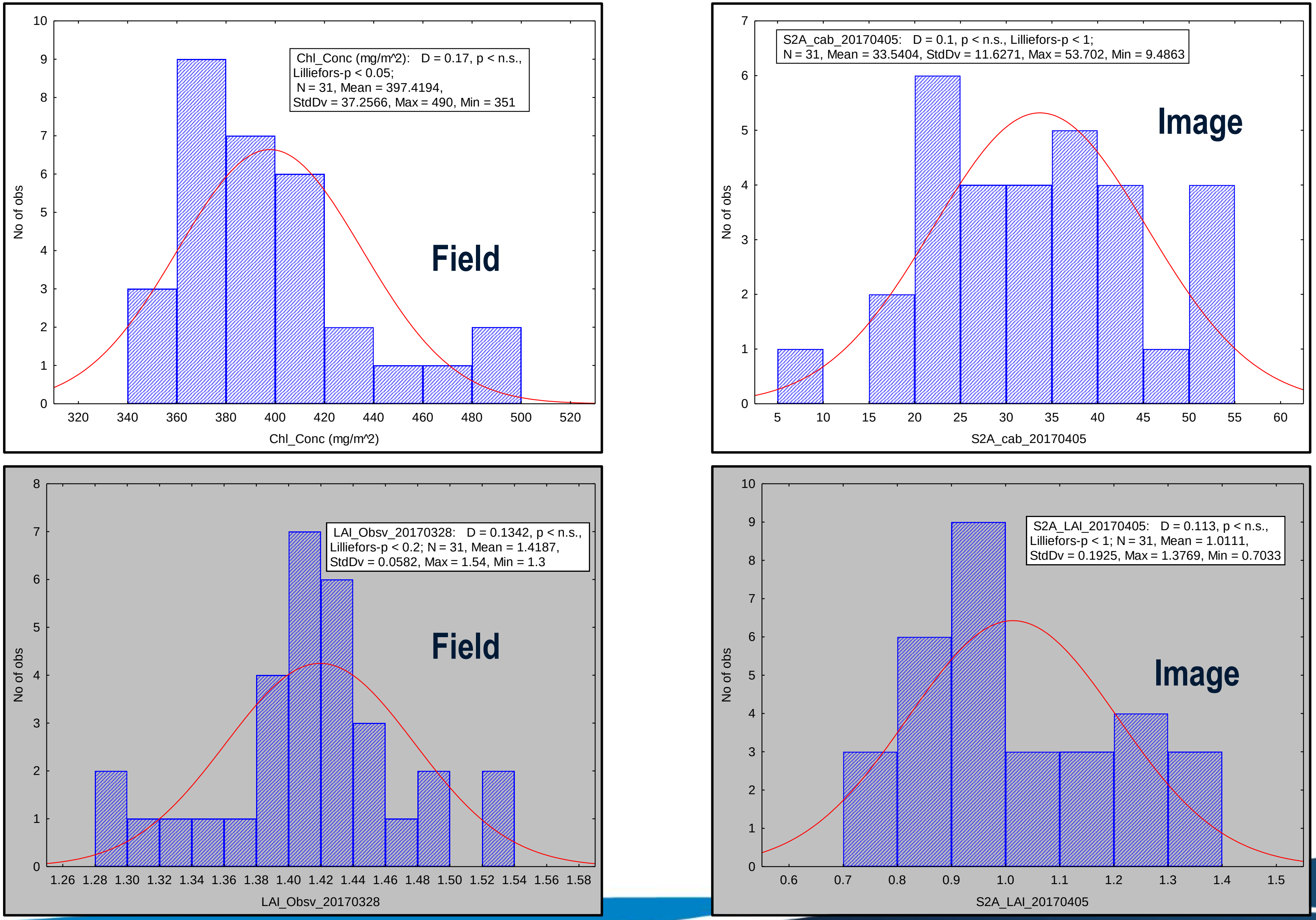
Temporal Profiles: LAI and Cab retrieved from S2A images, - zero cloud cover over target area



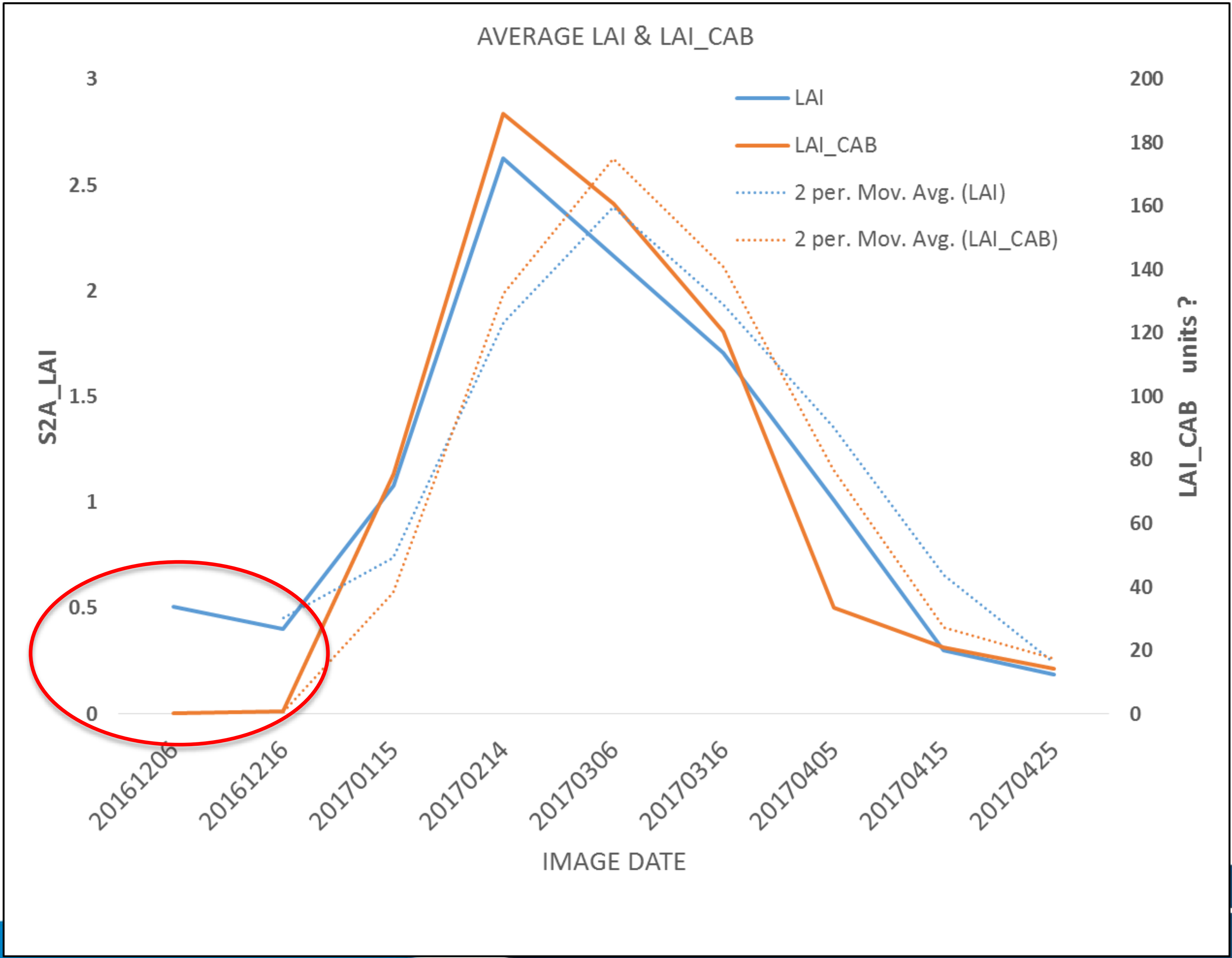


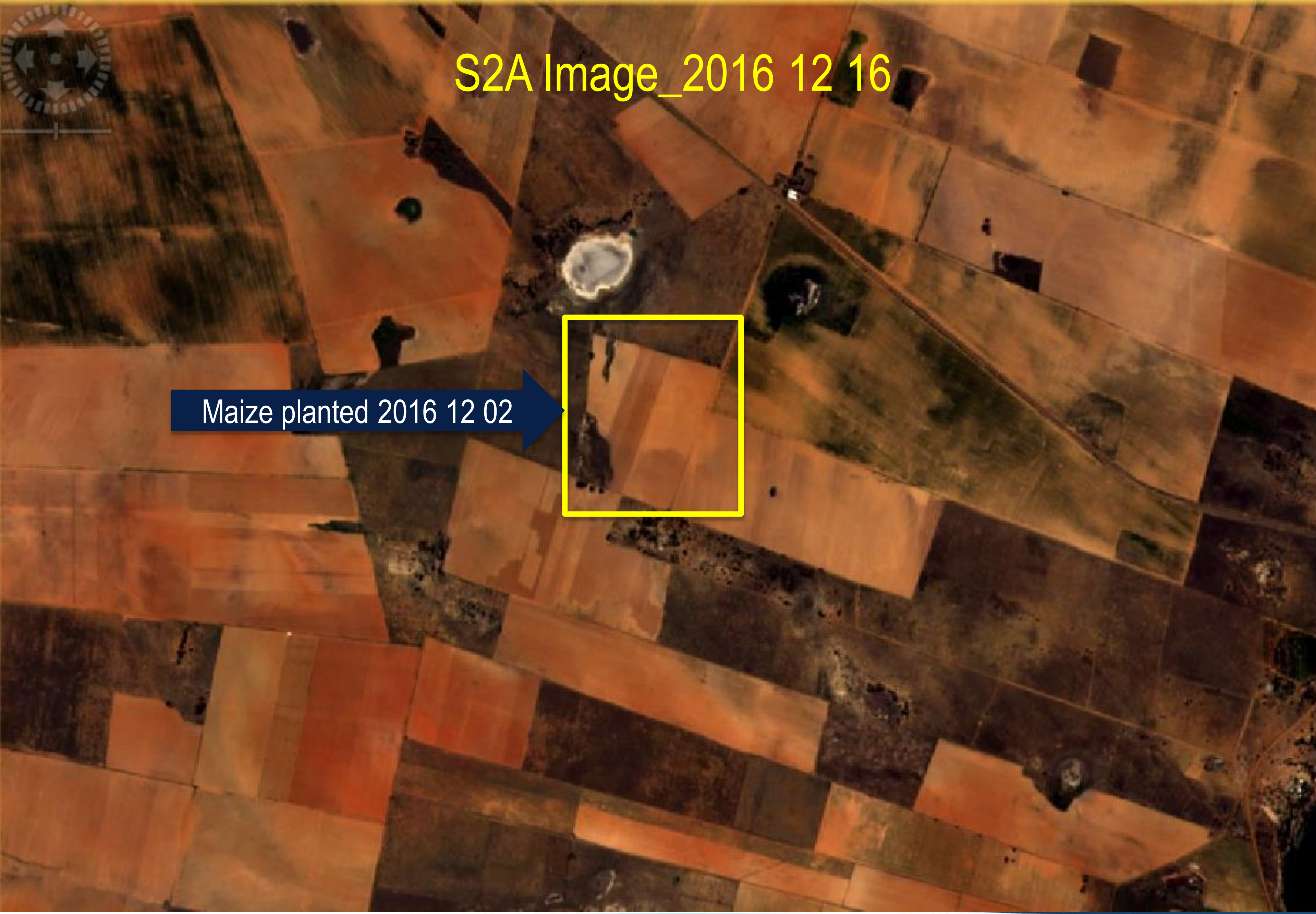


Results: LAI/Cab

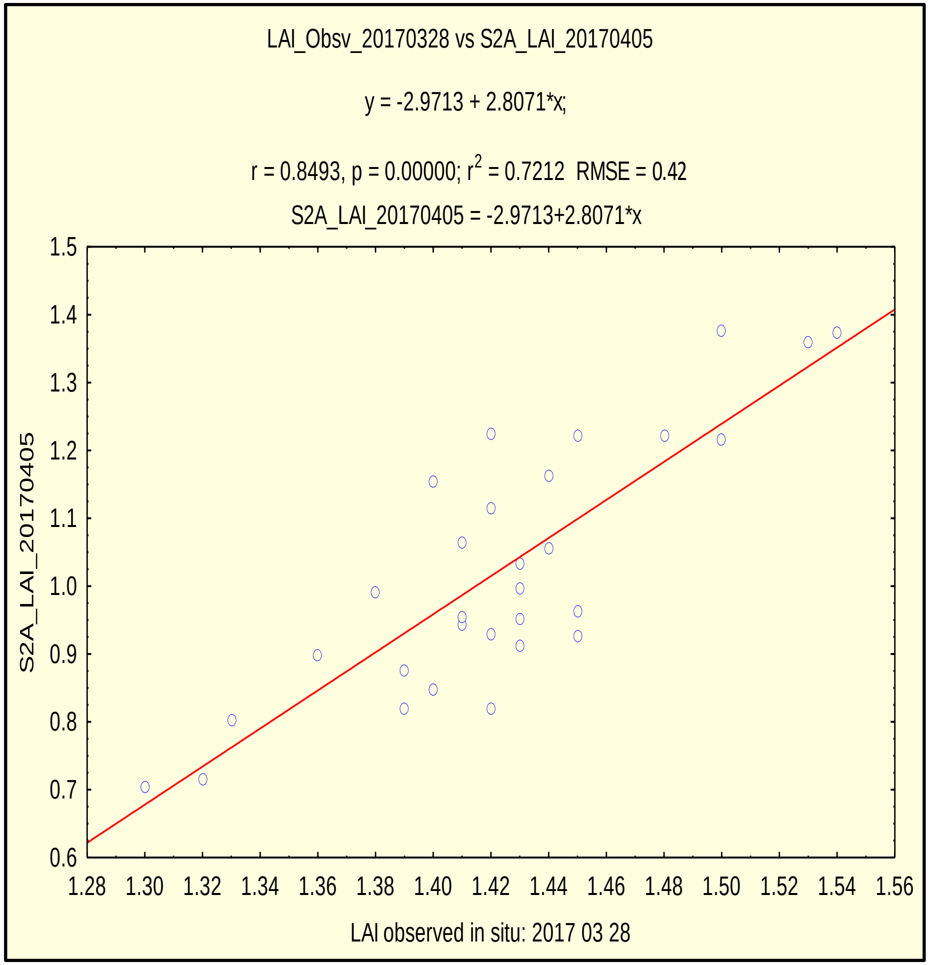


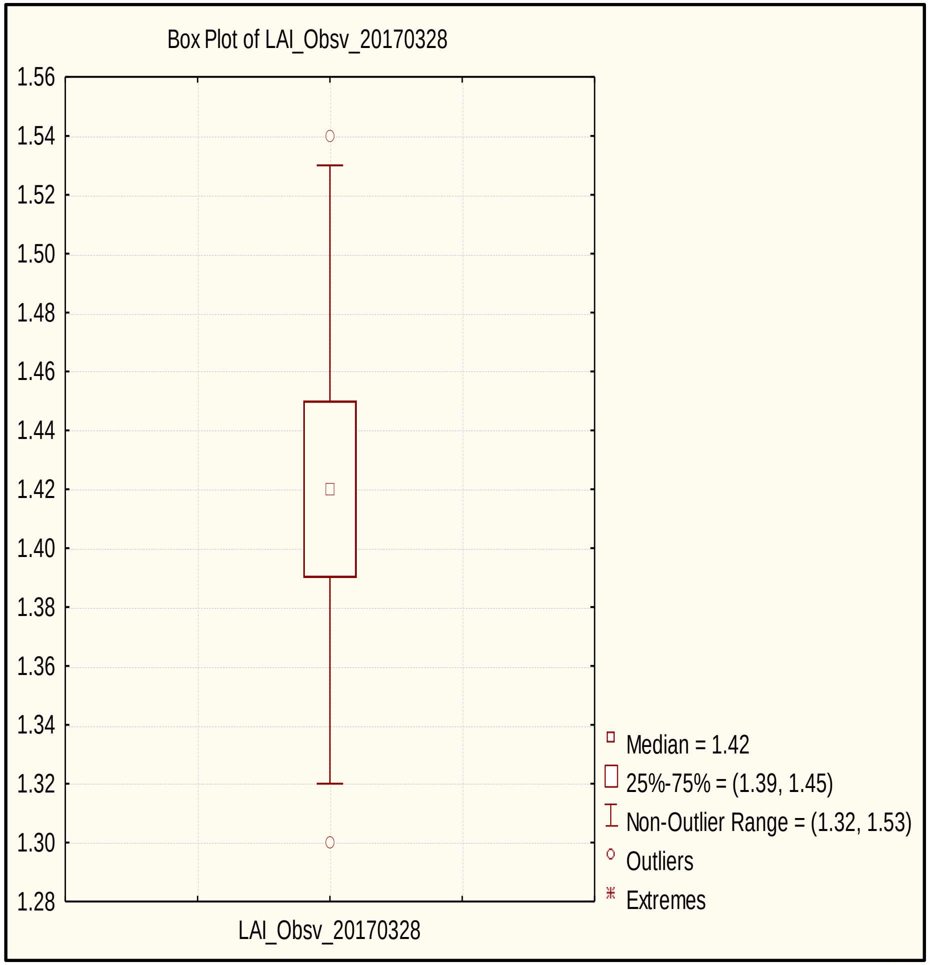
Temporal Profiles: LAI and Cab




S2A Image_2016 12 16
Maize planted 2016 12 02

Comparison between Observed (Field) and S2A LAI/Cab





science
& technology
Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

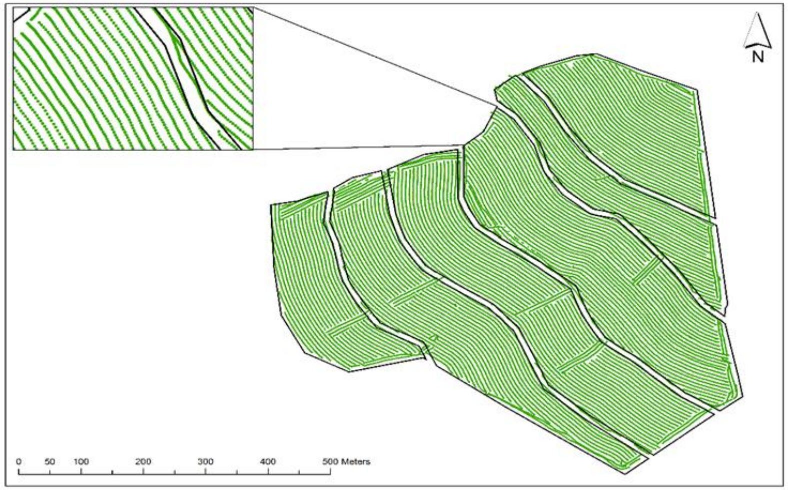
agriculture
& rural development
Department:
Agriculture and Rural Development
PROVINCE OF KWAZULU-NATAL

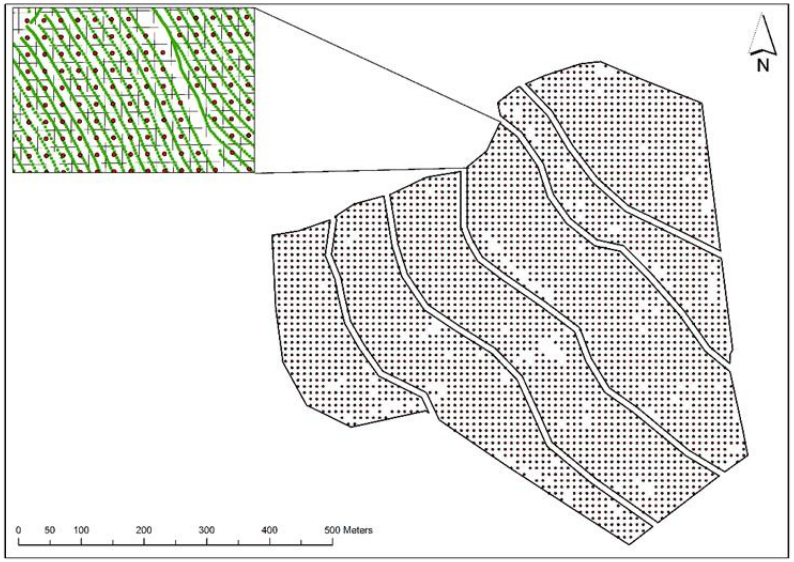
sansa
SOUTH AFRICAN NATIONAL
SPACE AGENCY

Yield modelling

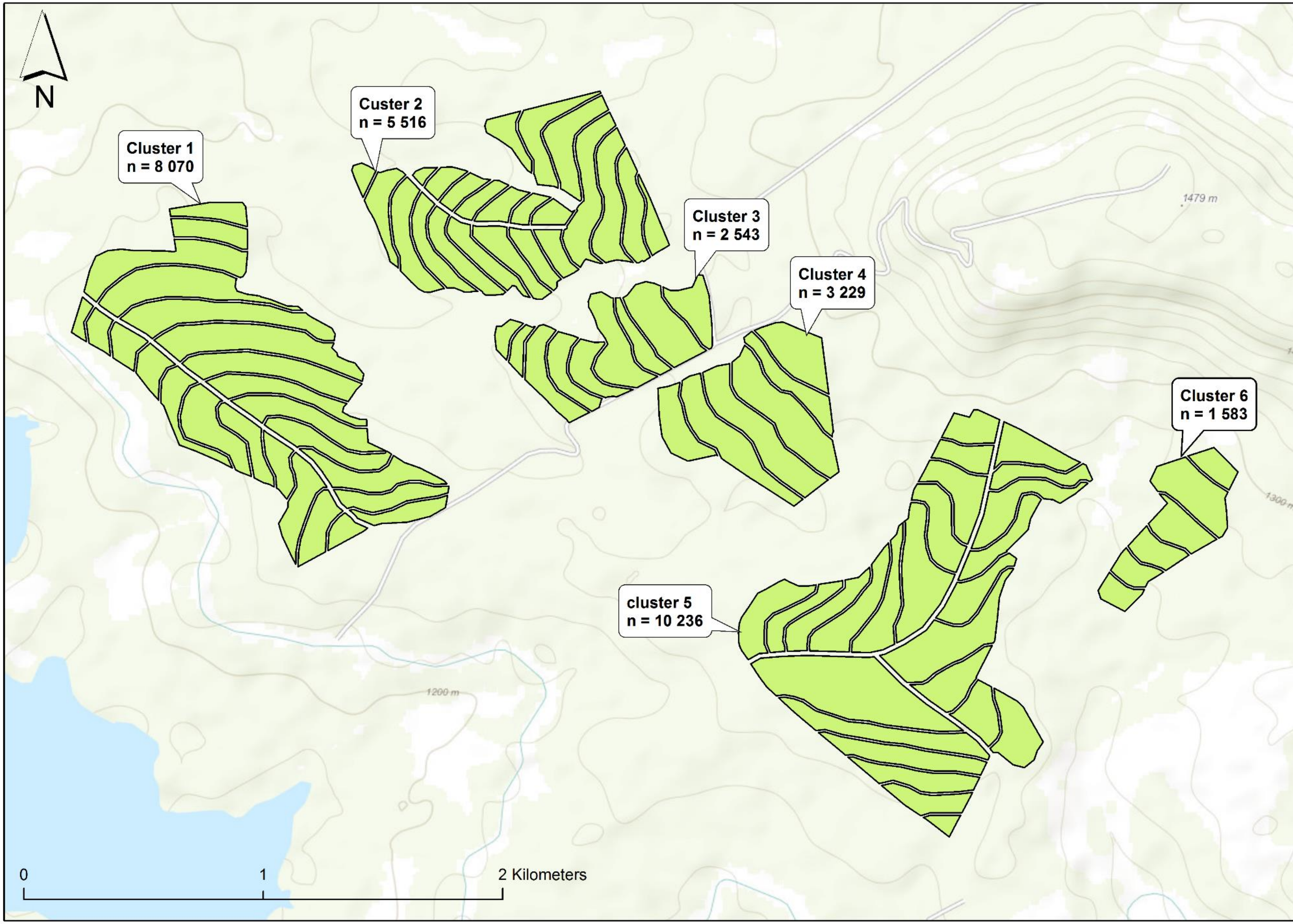
Materials and Methods: Yield Estimation

- Digitize field crop boundaries into 6 fields
- Extract all raw APEX crop data per field boundary
- Mask S2-A & L8 data with field boundaries
- Create 10x10 (S2-A) & 30x30 (L8) fishnet grid with centre point
- Convert all APEX crop data from volume based (Bu/Ac) to Mass Based (Kg/Ha)
- Extract S2-a & L8 LAI & NDVI values using a multi-point grid approach per field
- Summarised crop data per grid: Sum, Mean, Min, Max
- Remove all values of Max Yield/ha > 15 tons – yield on farm approx. 14.5 – 15 t/ha (Clumping error)



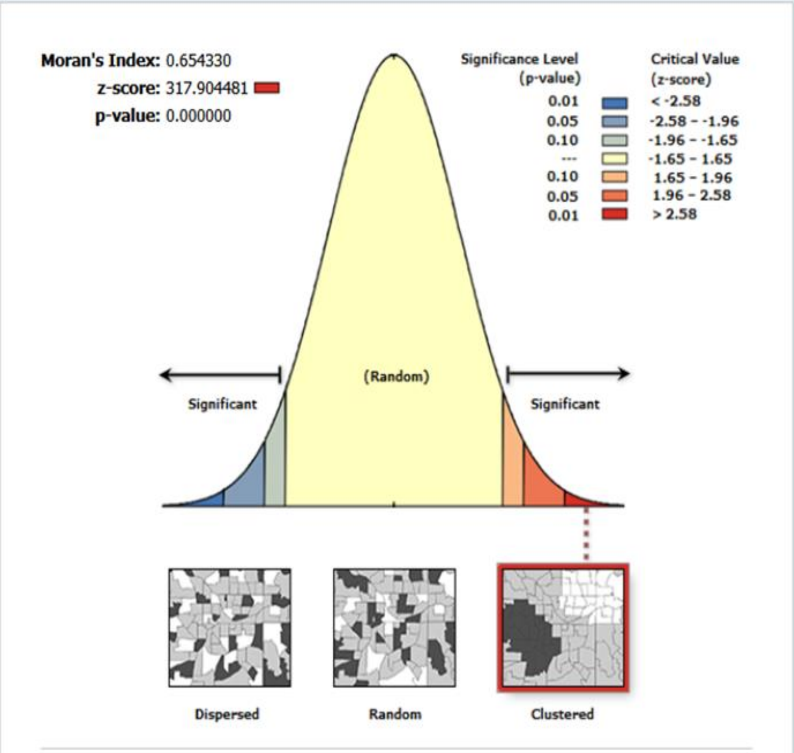


Maize fields



Materials and Methods

- Study assessed utility of three well known classifiers for predicting maize yield using LAI & NDVI from S2A & L8:
  - Random Forest *(Breiman, 2001)(*R-Rattle),
  - Ordinary Least Squares regression (Tanagra) &
  - Classification and Regression Trees *(Breiman et al., 1984)(*Tanagra)

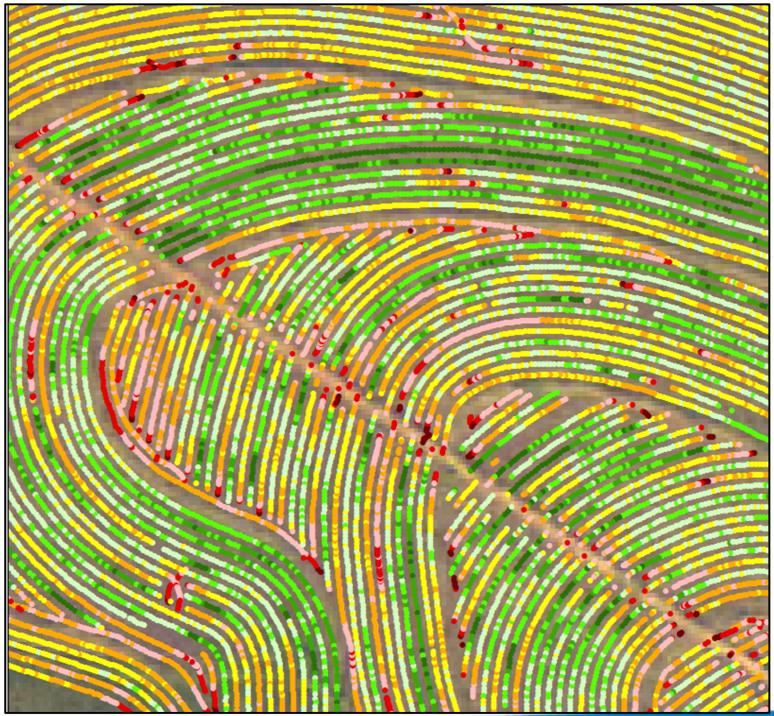


- Coefficient of Variation (R^2) used to describe model performance and residual variability

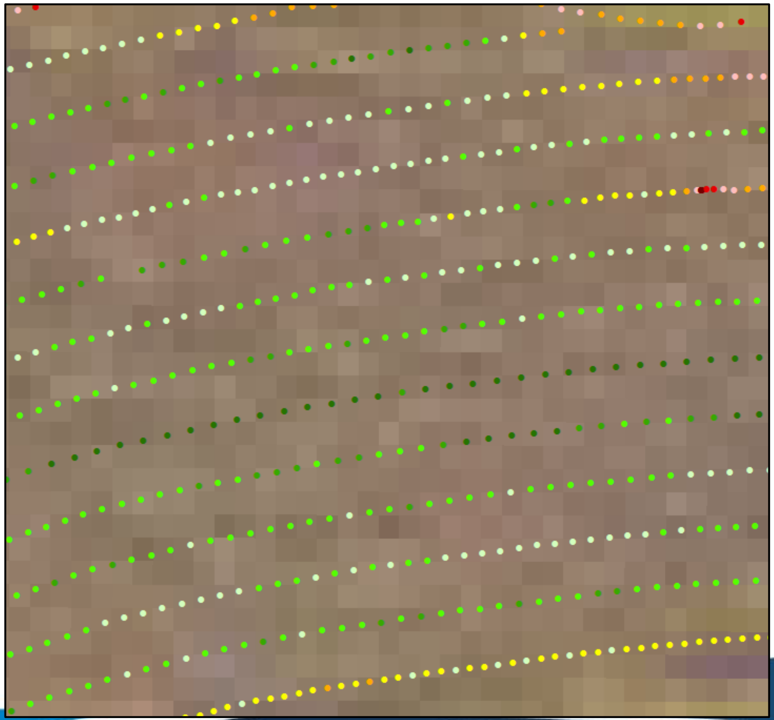
Materials and Methods

High Resolution Yield Analysis

- Yield data for 2015/16 growing season (Precision Harvester)
- 369 ha dry land maize)
- 300 801 yield points for maize (2m x 6m)

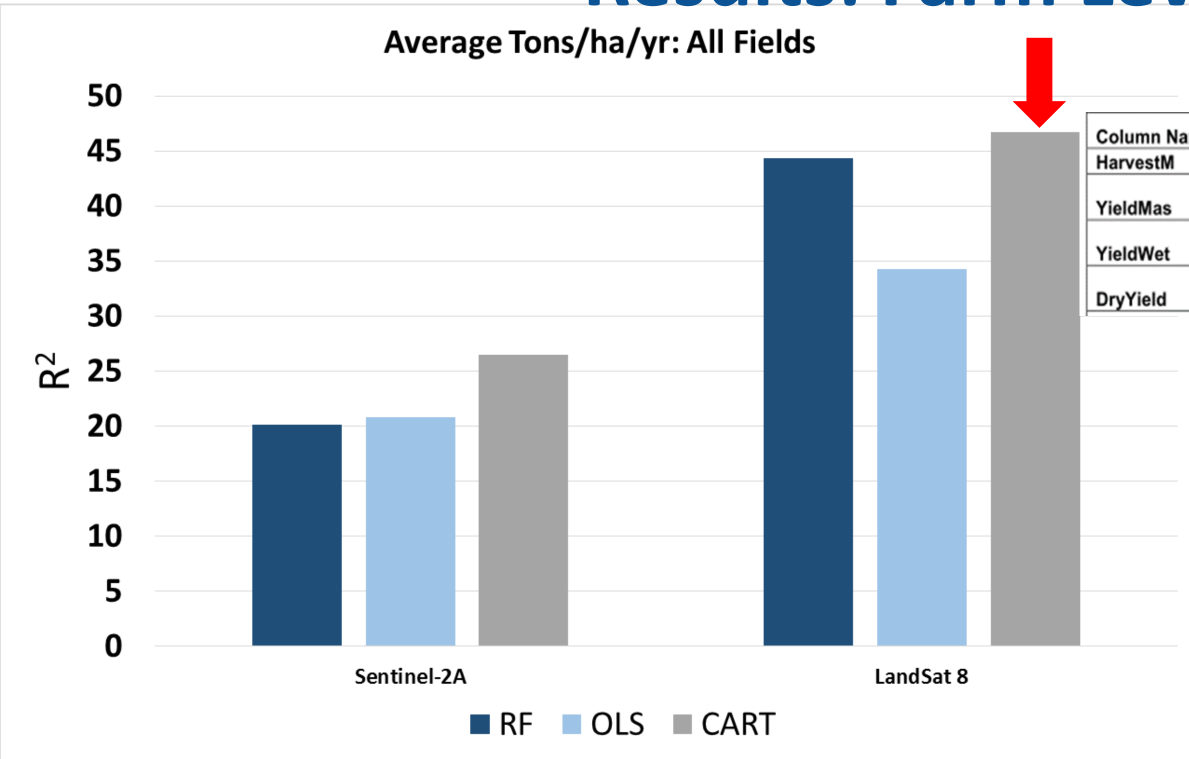


1 : 2 500



1 : 450

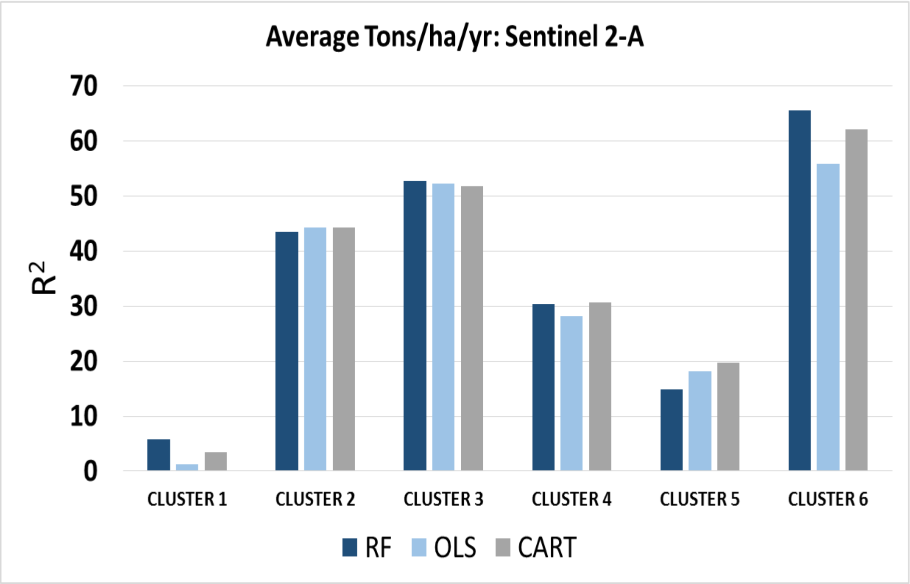
Results: Farm Level

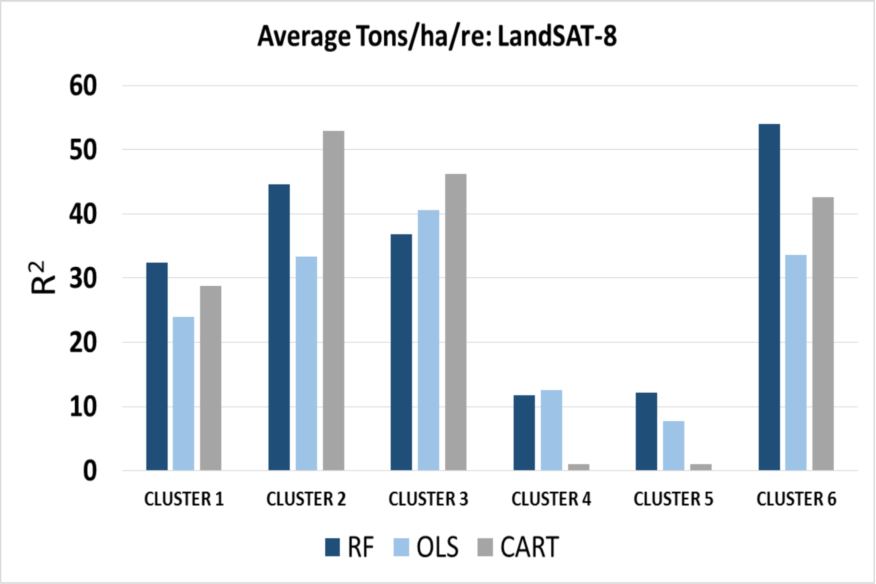


Column Name	Column Description	Units/Values
HarvestM	Harvest Moisture	Percent
YieldMas	Dry Yield Mass	lbs or tons (English), kilograms or metric tons (Metric)
YieldWet	Wet Yield Mass	lbs or tons (English), kilograms or metric tons (Metric)
DryYield	Crop Yield	bu/ac or bales/ac (English), kg/ha or bales/ha (Metric)

- Model performance is marginal when averaged across the entire farm using all sampled LAI, n > 3000
- Best yield prediction using L8 and CART achieved 47 %
- Variability of maize canopy LAI and spacing within fields not adequately represented as a global statistic
- Reducing sample size from farm to field level may improve alignment between canopy homogeneity and LAI values

Results: Field Level





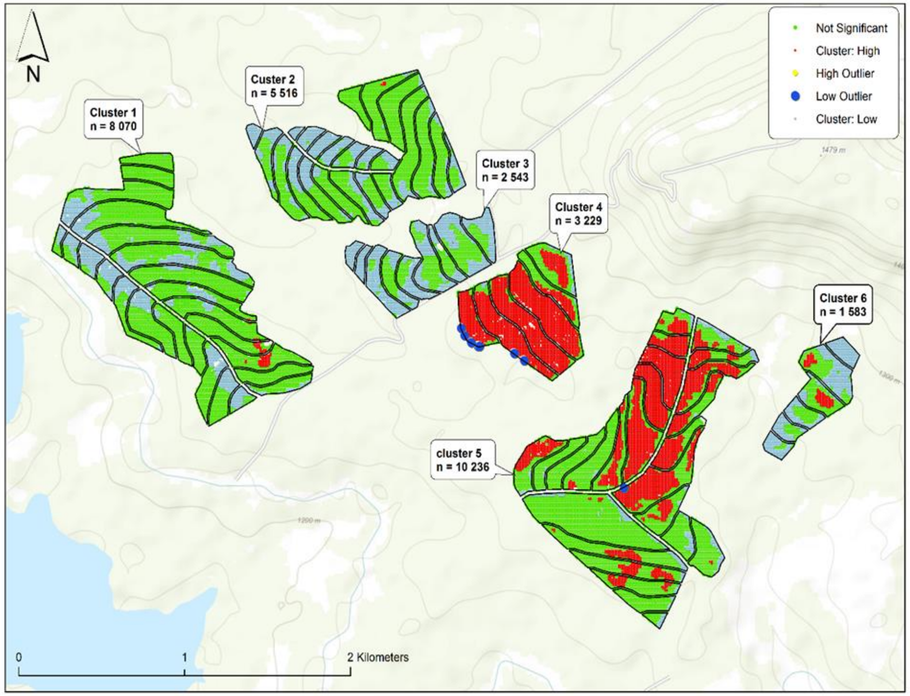
- Maize yield prediction showed significant improvement across both sensors at field level
- Cluster 6 (field 6) showed the best overall model performance in both S2-A & L8 with RF outperforming OLS & CART at 66 & 54% respectively
- Clusters 1, 4 & 5 performed the least accurate for both Sensors with RF the best model at 31% for S2-A and 32% for L8

Results: Field Level

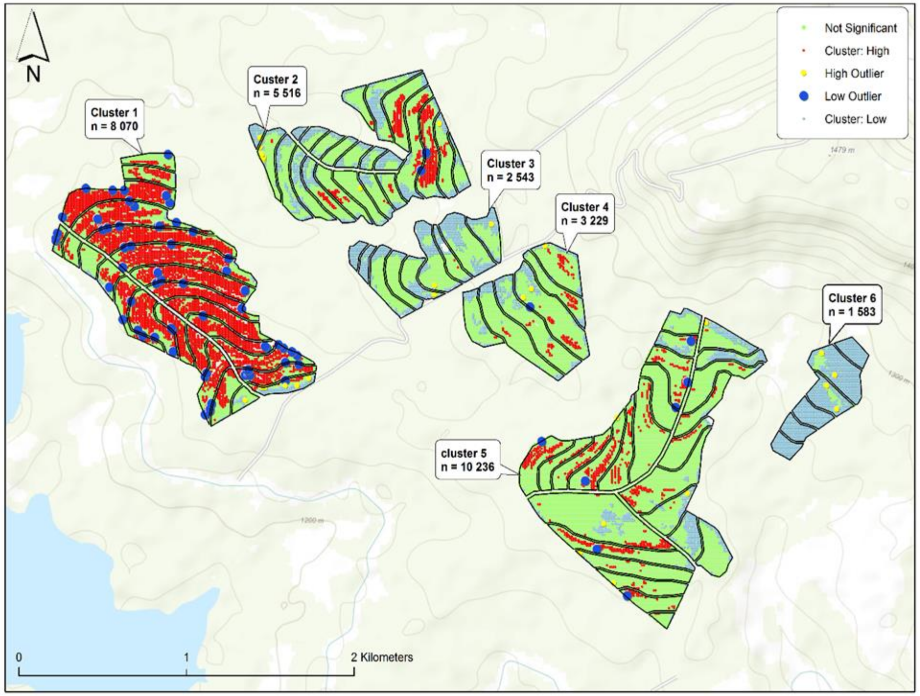
- Despite results there is significant variability between S2-A and L8 performance
- We observation pixel resolution-scale issues when averaging data from grid contributed to variability
- While L8 has coarser resolution than S2-A, averaging over a larger area within the field may result in better representation of field variability
- By running a cluster and outlier analysis we can spatially observe the trends described by the model performance
- **The highest accuracy would be observed in fields with the highest agreement in LAI & yield values, the more heterogeneous the LAI or crop yield the lower the predicted accuracy expected**

Results: Field Level

LAI Outlier Analysis



Crop Yield Outlier Analysis



- From the analysis we, see that in Cluster 1 the Crop yield variation is the main cause for the error
- while in cluster 4 and Cluster 5 it is the LAI variance that is driving down the model performance

Conclusions and further work

- LAI (EO) and LAI (field) are physically-meaningful measure of maize crop canopy properties.
- RMSE between LAI (image) and LAI (field) fairly nominal
  - RMSE 0.42 compared to LAI mean 1.42 (field)
- Need to consider effects of scaling
- Relate measurements at small scales
  - perhaps, extend from 1 pixel to 3 more pixels?
- Measurements and validation at many scales
  - techniques to bridge time/space scales

Conclusions and further work….

- Can LAI and NDVI be calibrated to proximal field crop yield data using Sentinel 2-A and LANDSAT 8 platforms? **YES, but results highly variable**

1. Investigate optimal sampling strategy for the extraction of field data and generalisation of remotely sensed data:
   - 10x10 m gridded approach applied at sub-sampled field level
2. Determine which prediction model is best suited for the description of satellite-to field yield estimates:
   - RandomForest had highest model accuracy with S2-A & L8 LAI and NDVI
3. Determine which sensor platform delivers the most accurate representation of maize yield using LAI and NDVI
   - Maize yield most accurately modelled with S2-A, 66 %

Conclusions and further work….

- Future studies will focus on assessment of other supervised classification/regression methods, such as, SVM and Neural Networks to relate high resolution crop yield data and EO (LAI & NDVI) observations.

- Explore full suite of S2-A thematic land-processing products (LAI, Cab, FAPAR…) for yield prediction

- Apply model(s) to wider range of crops i.e. sugar or soy

<u>Investigate models to fuse pixel space scales for Sentinel and Landsat platforms:</u>

…but how good would be such models…?

Thank You