Crop Classification with RapidEye and Radar Data

U. Schulthess, M. Kunze and H. Weichelt
DGPF Hannover - 2010/11/18

<www.rapideye.de>

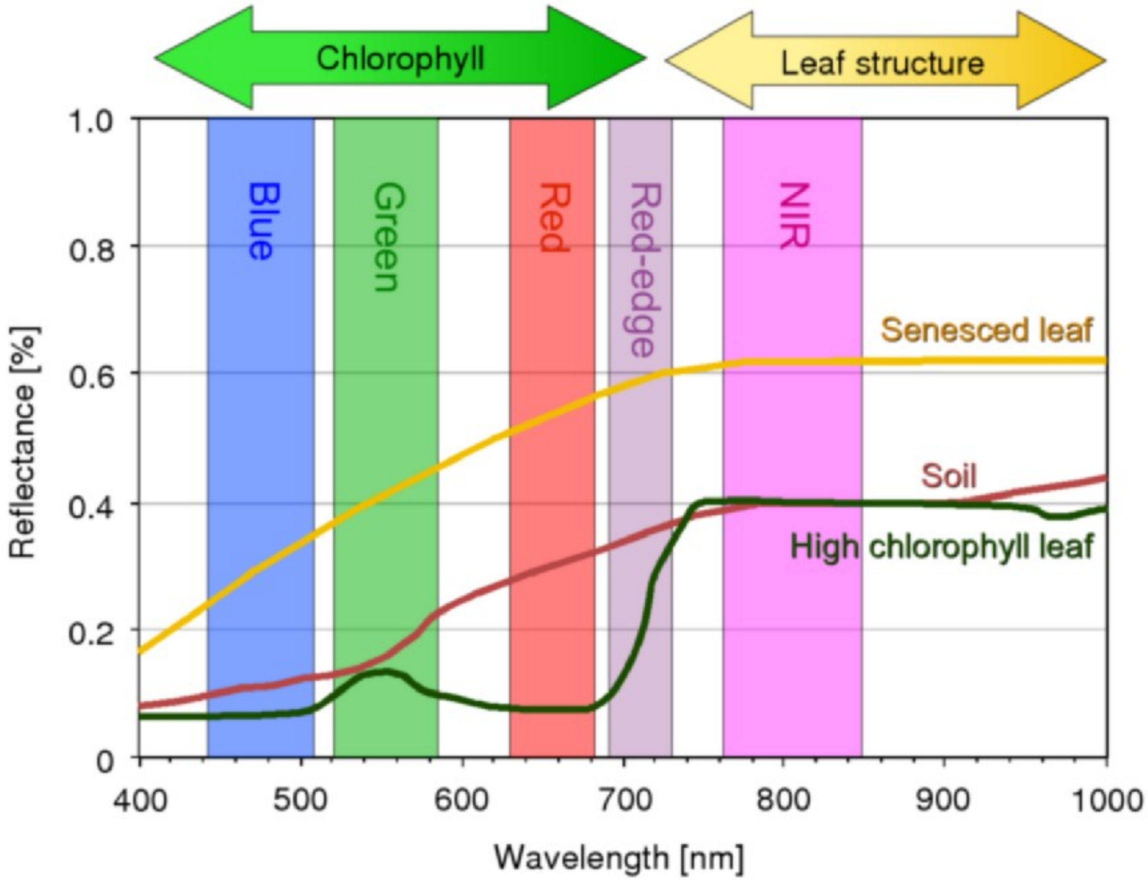
Big Brother or Useful Stuff?

Source: The Economist, Nov 5, 2009
http://www.economist.com/sciencetechnology/displaystory.cfm?story_id=14793411

Overview

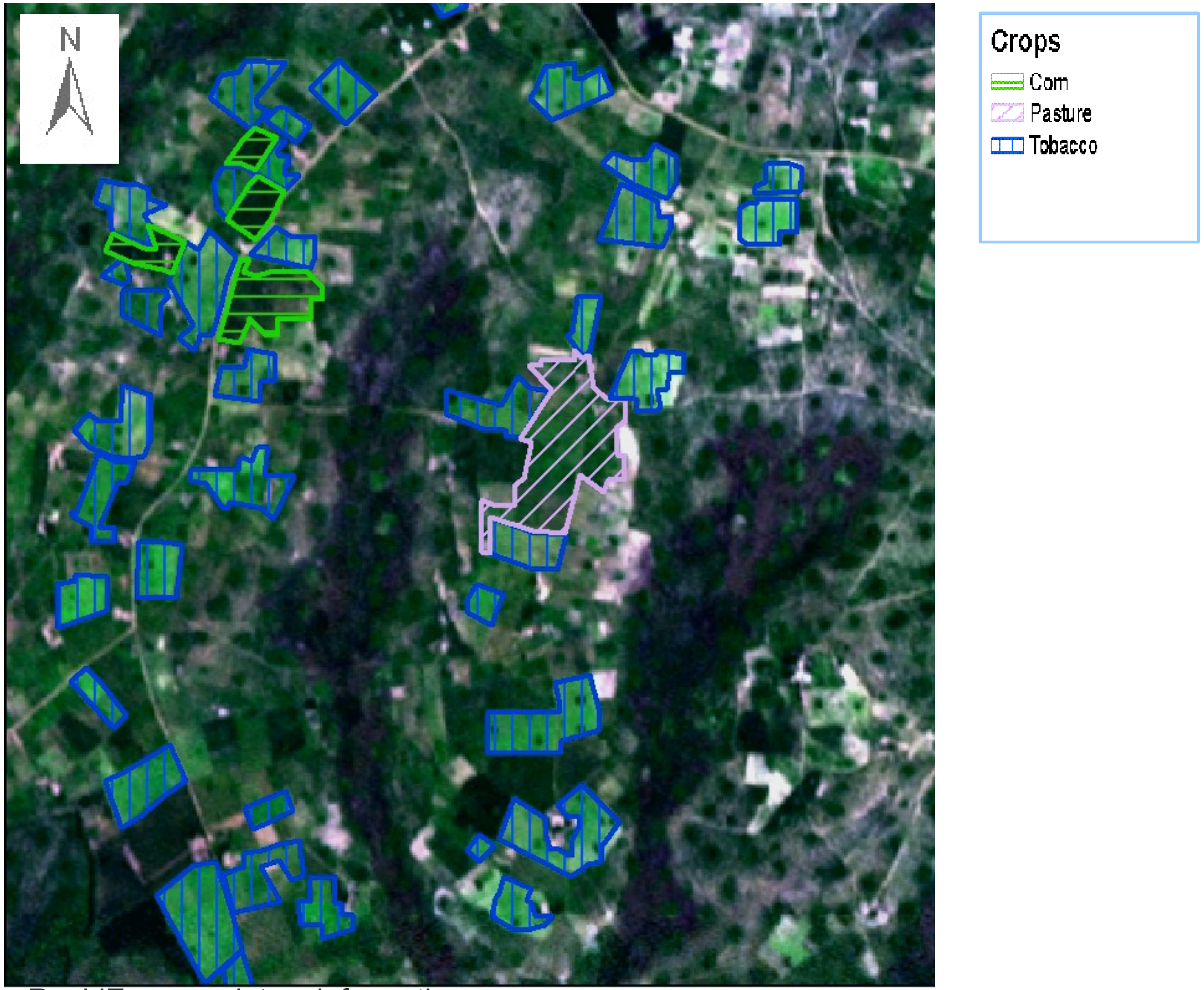
- 5 bands – examples of various band combinations to visualize specific crops
- Crop classification with C5.0 Data Mining tool
  - Procedure
  - Examples

Cameras on board of the satellites use 5 bands to measure light reflected from the surface of the earth

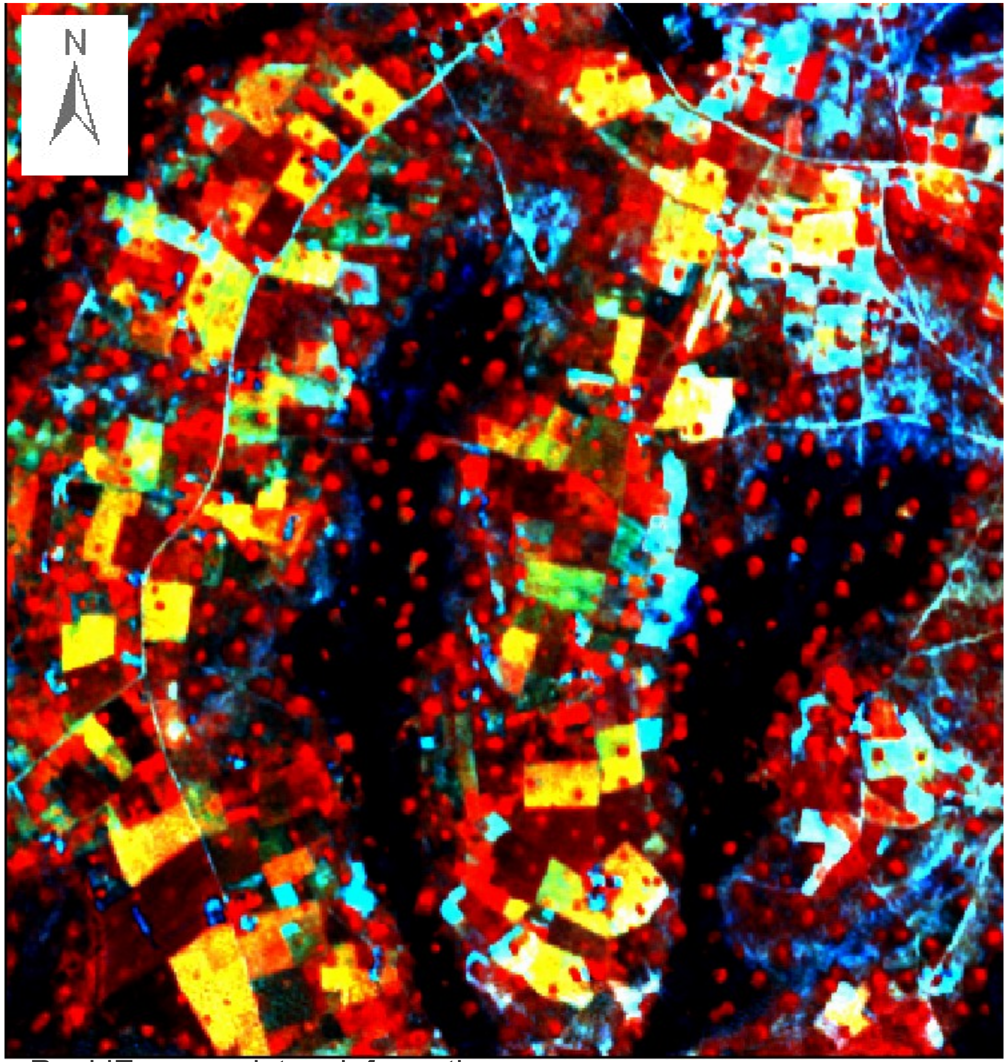


Tobacco identification in Malawi (RGB)

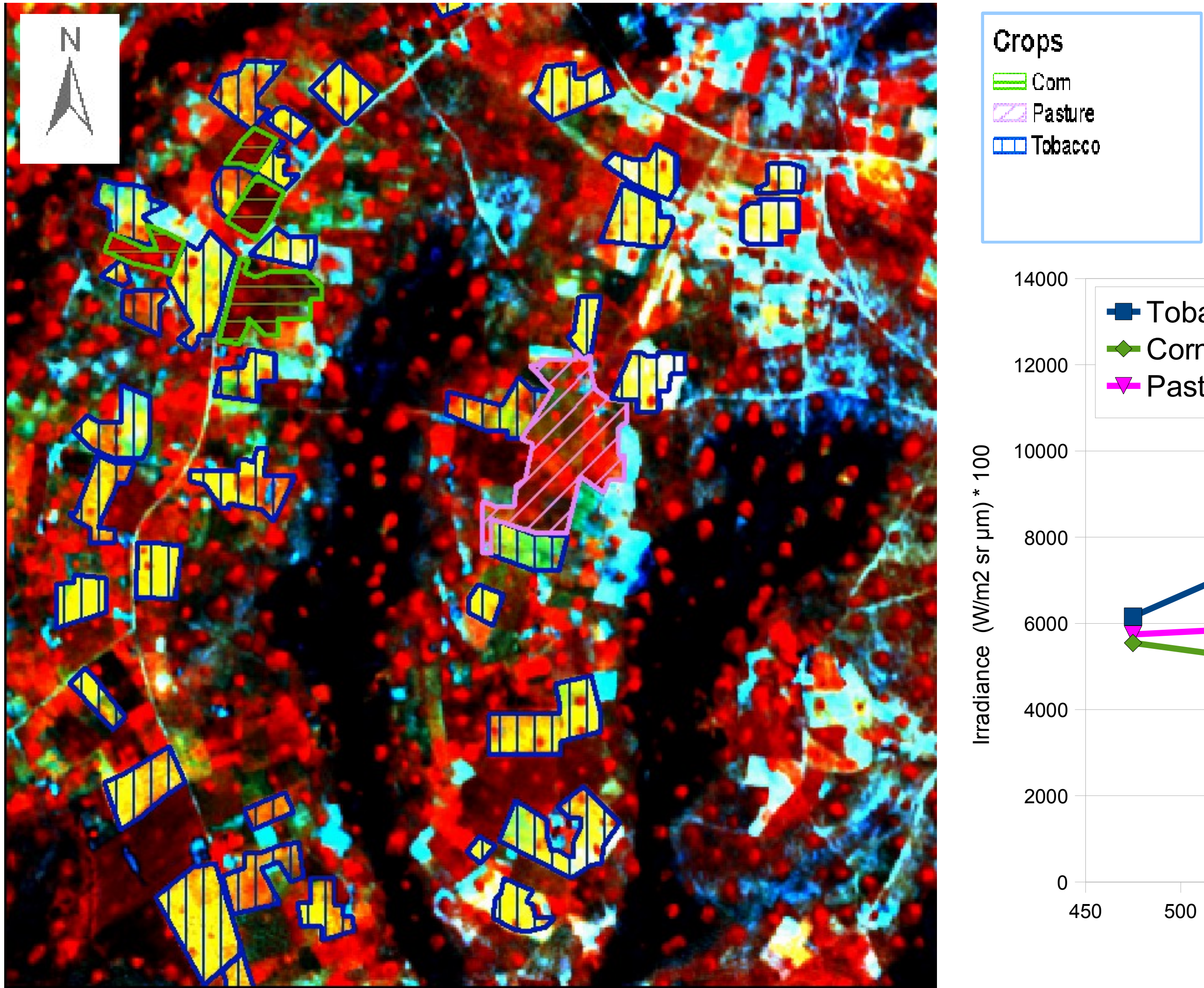
Tobacco identification in Malawi (RGB)

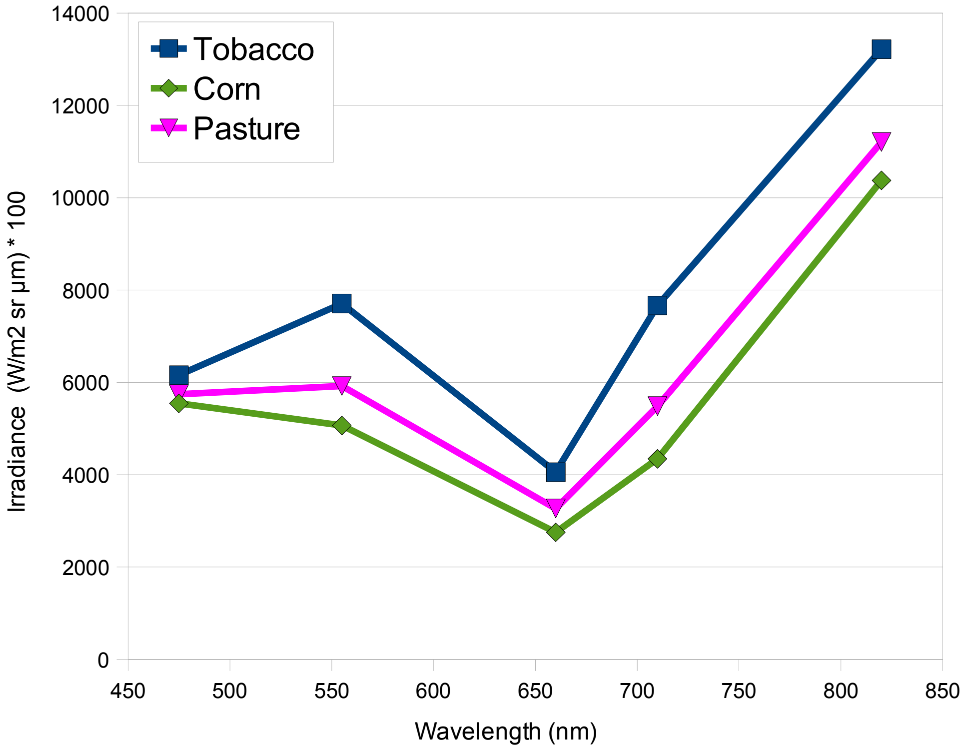


Tobacco identification in Malawi with band combination NIR-green-blue

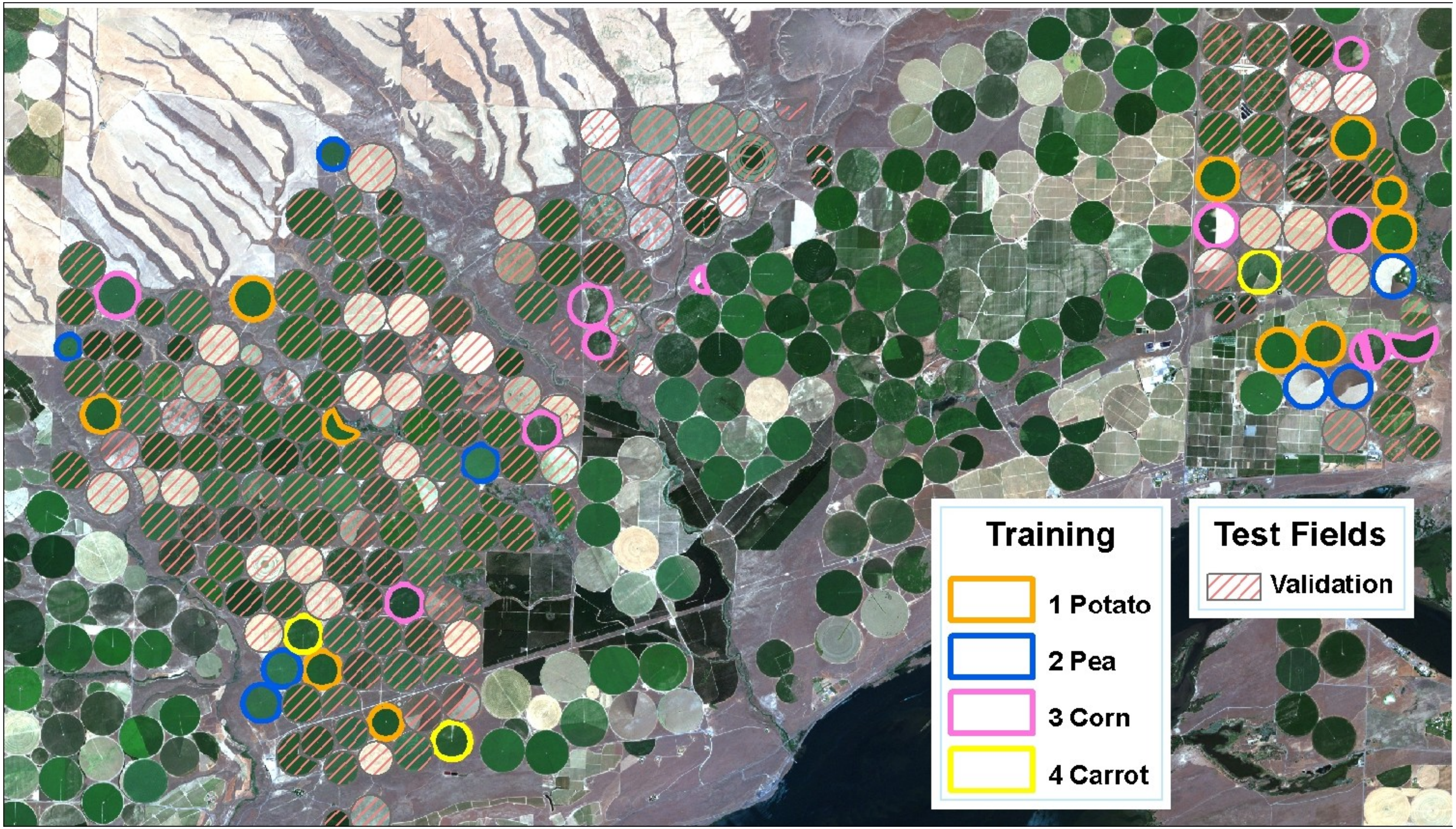


Tobacco identification in Malawi with band combination NIR-green-blue

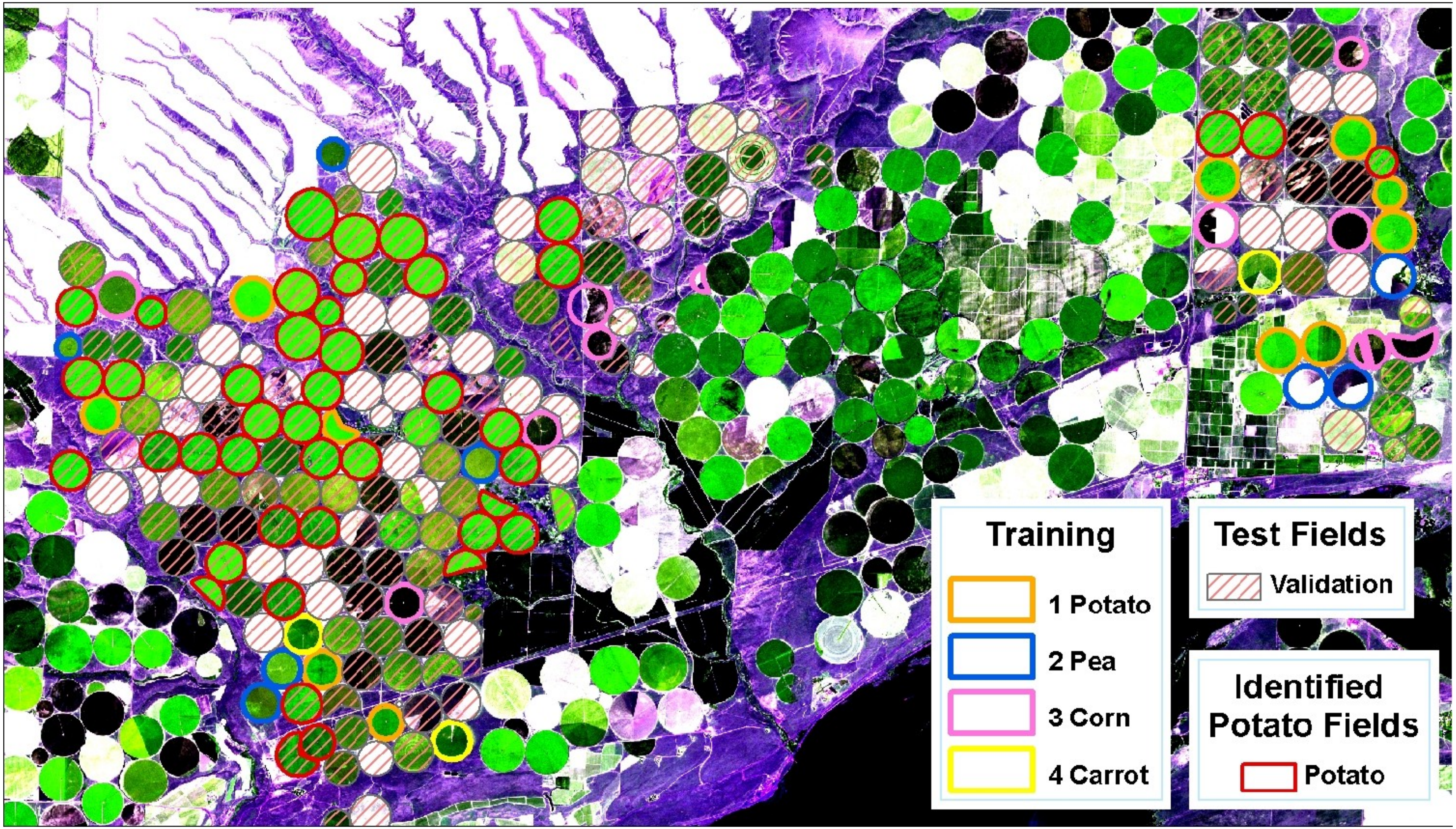




Potato Identification (RGB)



Potato Identification with band combination red-edge, green, blue



Crop Identification Tool

- Use of data mining tool C5.0 (RuleQuest)
  - Supervised learning algorithm
    - Decision Tree
    - Rule Set
  - **Recursively partitions a data set of records until all data belong to a particular class**
  - Advantages:
    - Prior knowledge of the class labels of data records makes feature/attribute selection easy (as opposed to unsupervised classification)
    - Non-normal distribution of data is not a problem (important for field crops, where dates of sowing can vary greatly)
    - Handles missing data

Example of Decision Tree

RapidEye™

```
I1B4i <= 0.12324 (0.125065):
:...I2B1i >= 0.0564 (0.053405): 3 (601.9/7.3)

:   I2B1i <= 0.04981 (0.053405):
:   :...RD20B1i <= 0.07048 (0.08438499): 2 (8)
:       RD20B1i >= 0.09829 (0.08438499):
:       :...RD30B9i <= -0.6019 (-0.53972): 1 (5)
:           RD30B9i >= -0.47754 (-0.53972): 6 (3)
I1B4i >= 0.12536 (0.125065):
:...RD25B8i >= 0.087432 (0.0818705):
    :...RD28B9i >= -0.60599 (-0.61268):
    :   :...I2B4bysB1235 <= 0.23126 (0.243475): 7 (61.2/0.1)
    :   :   I2B4bysB1235 >= 0.2683 (0.243475): 9 (4.4/0.4)
    :   RD28B9i <= -0.64981 (-0.61268):
    :   :...RA29B3i <= 0.02161 (0.02405): 6 (9.2/1.2)
    :       RA29B3i >= 0.0252 (0.02405):
    :       :...RA26B9i >= -0.54376 (-0.58566): 7 (4.4/0.4)
    :           RA26B9i <= -0.59557 (-0.58566):
    :           :...I2B4bysB1235 <= 0.17295 (0.187775): 6 (4)
    :               I2B4bysB1235 >= 0.20167 (0.187775): 9 (419.2/4.1)
    RD25B8i <= 0.07929093 (0.0818705):
    :...RD22B2i >= 0.0455 (0.04481):
        :...I2B2bysB1345 >= 0.14971 (0.13222):
        :   :...RA34B3i >= 0.12778 (0.104845): 19 (6)
```

3 → **Class**; 601.9 → **# of cases**; 7.3 → **# of cases mapped incorrectly**

Example of Rule Set

Rule 0/1: (127, lift 5.3)

127 → # of cases; lift 5.3 → f(accuracy & frequency)

    RD22B2i <= 0.0448
    RA29B3i > 0.03577
    RD46B3i > 0.03379
    RA47B2i <= 0.04854
    I2B4bysB1235 > 0.20167
    -> class 9 [0.992]

[0.992] → level of confidence

Rule 0/2: (419/3, lift 5.3)

    RD25B8i > 0.08185098
    RA26B9i <= -0.58566
    RD28B9i <= -0.61301
    RA29B3i > 0.02396
    I2B4bysB1235 > 0.18735
    -> class 9 [0.990]

Rule 0/3: (377/7, lift 5.2)

    RD25B8i > 0.08185098
    I1B4i > 0.12477
    I2B4bysB1235 > 0.2434
    -> class 9 [0.979]

Rule 0/4: (198, lift 9.8)

    RD30B8i <= 0.1235097
    RA47B2i > 0.04854
    I1B6i <= 22.23273
    -> class 5 [0.995]

C5.0 Features

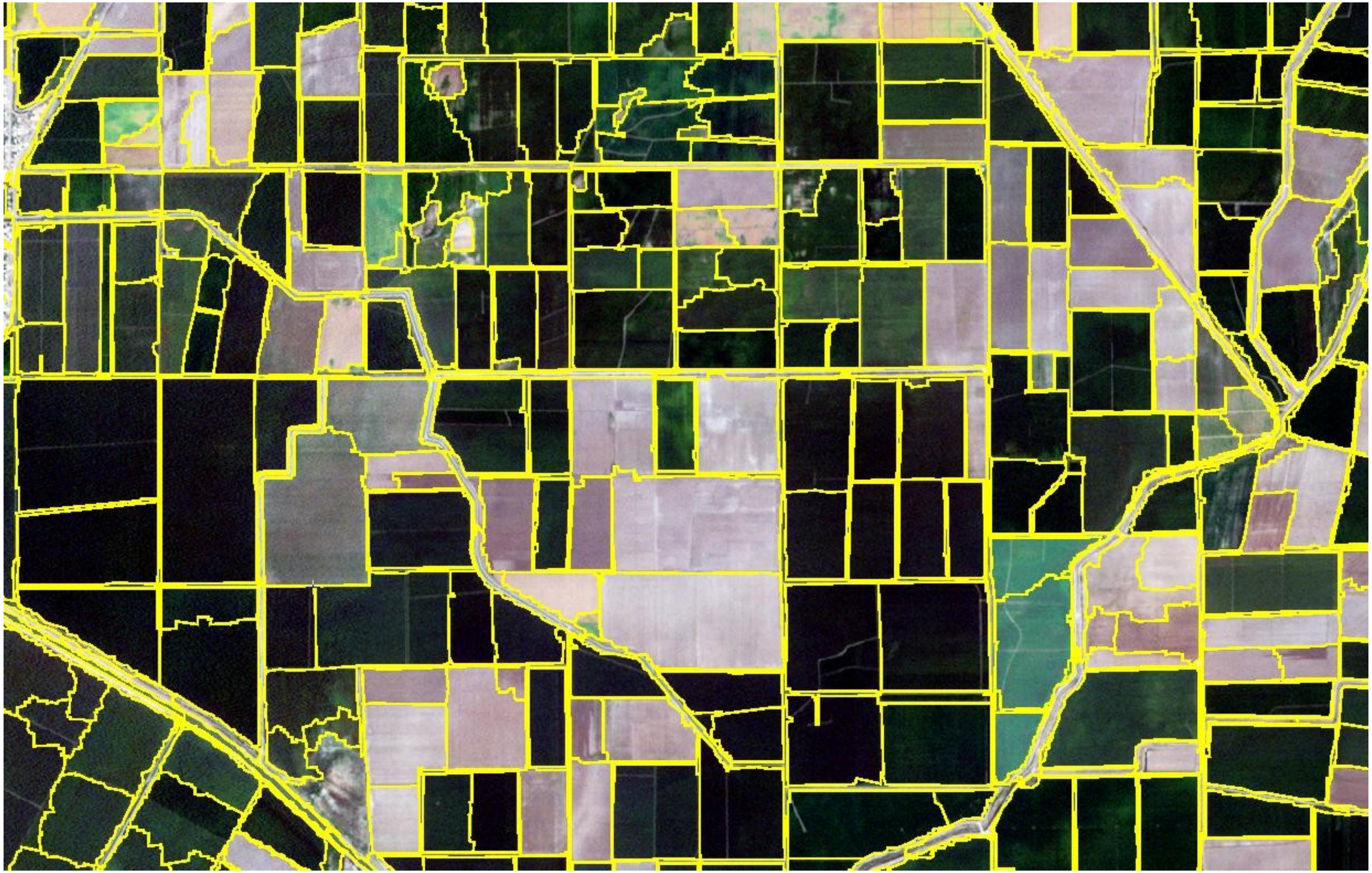
- Boosting: use of several classifiers rather than just one. When a new case is classified, each classifier votes for its predicted class and the votes are counted to determine the final class
- Winnowing attributes: picking and choosing among the predictors
- Pruning: Removing of parts that are predicted to have a relatively high error rate

Classification Process

1. Data preparation (Geo-referencing, cloud/shadow masking)
2. Segmentation (eCognition)
3. Training
4. Extract data
5. Classification
6. Import classification results
7. Quality control

Image Segmentation

Image Segmentation



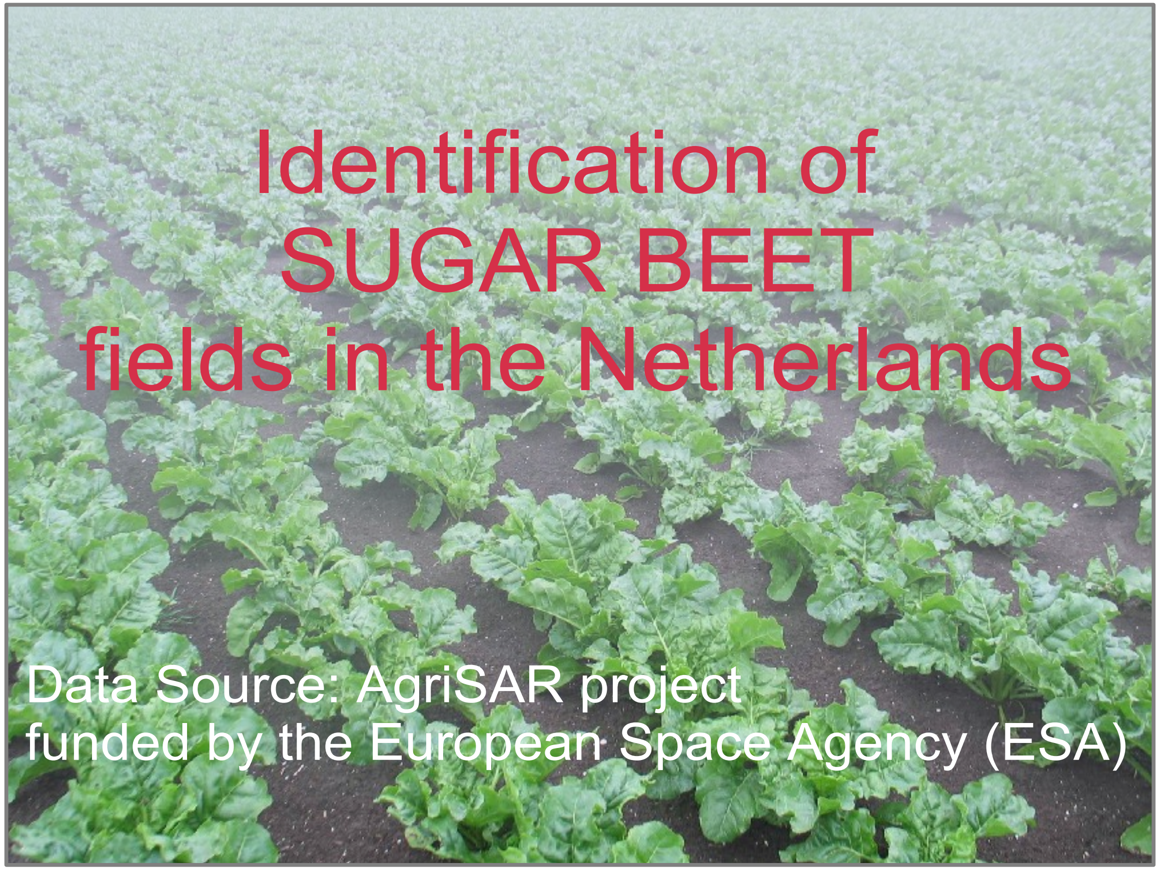
Data Extraction

- Extract statistics for each segment
  - DN average
  - Texture
  -
- Calculation of indices on the fly (within C5.0)
  - NDVI
  - Ground cover
  - ....

Training

- Large training data sets are preferred (> 50 cases)
- Check for consistency
- No clouds!
- ...garbage in, garbage out....

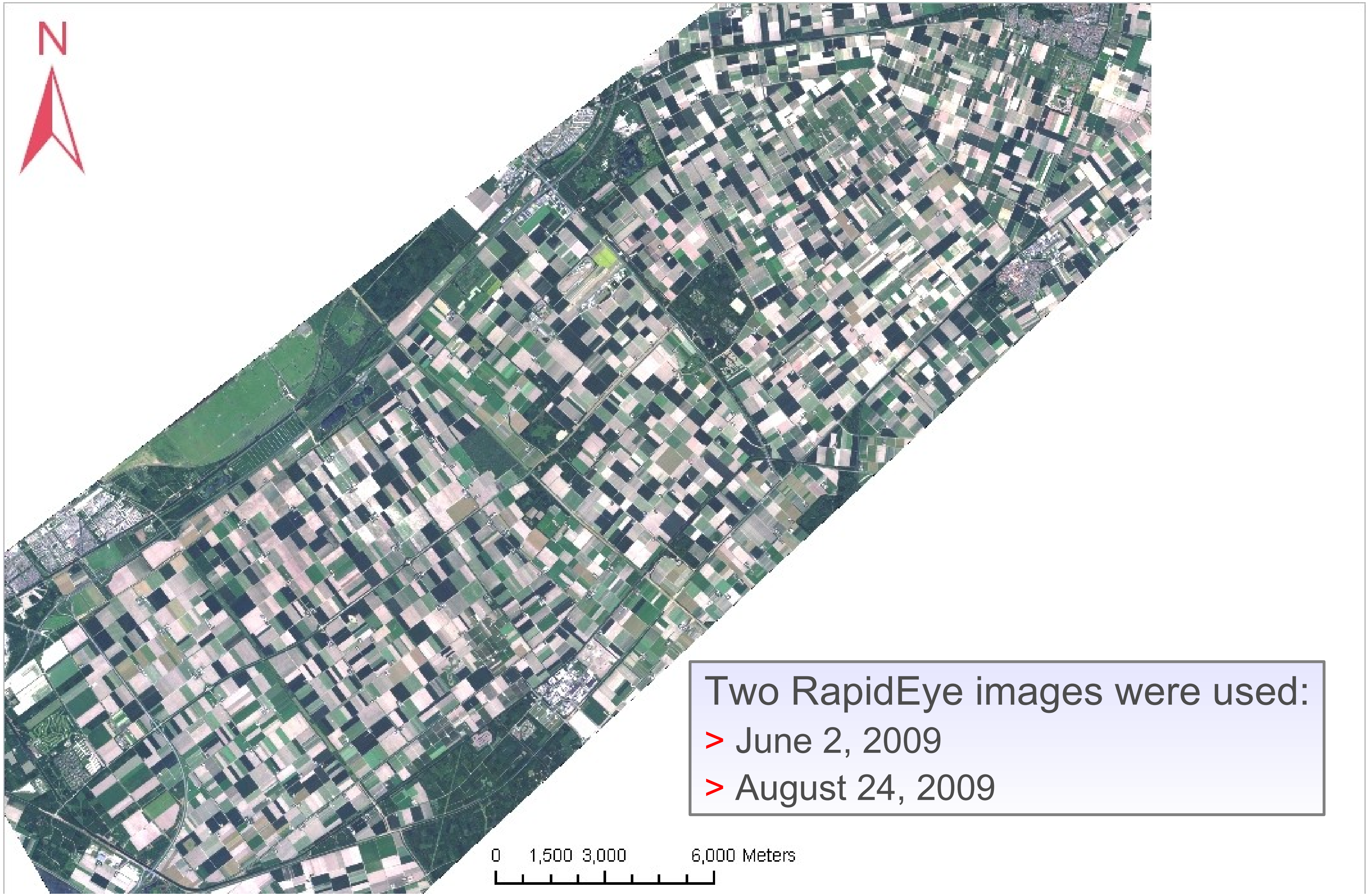
RapidEye™



Identification of SUGAR BEET fields in the Netherlands

Data Source: AgriSAR project
funded by the European Space Agency (ESA)

Overview of test region near Flevoland, NL



Two RapidEye images were used:

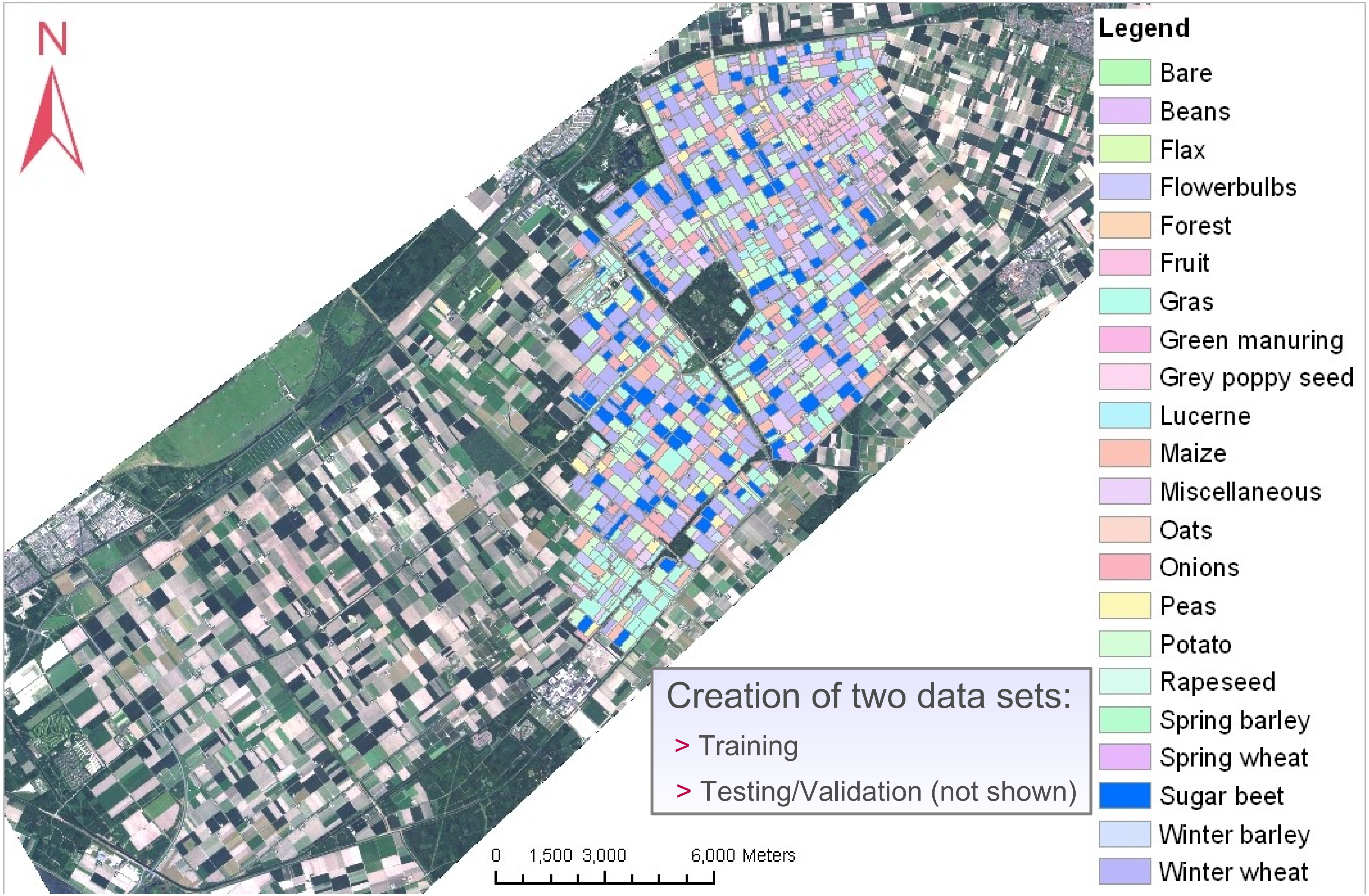
- June 2, 2009
- August 24, 2009

Overview of test region near Flevoland, NL

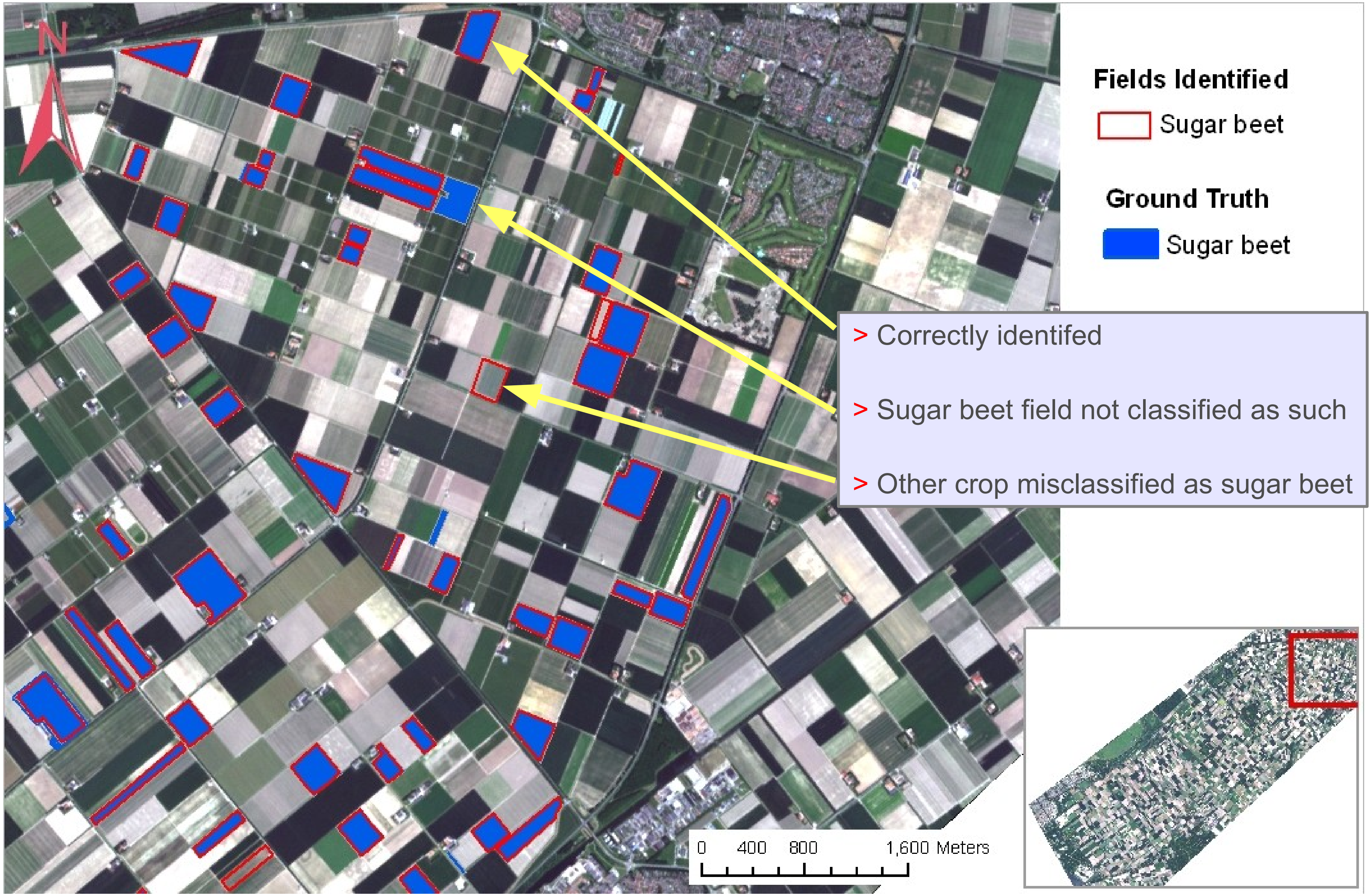
RapidEye™

> More than 4,200 fields
> 22 Crop types

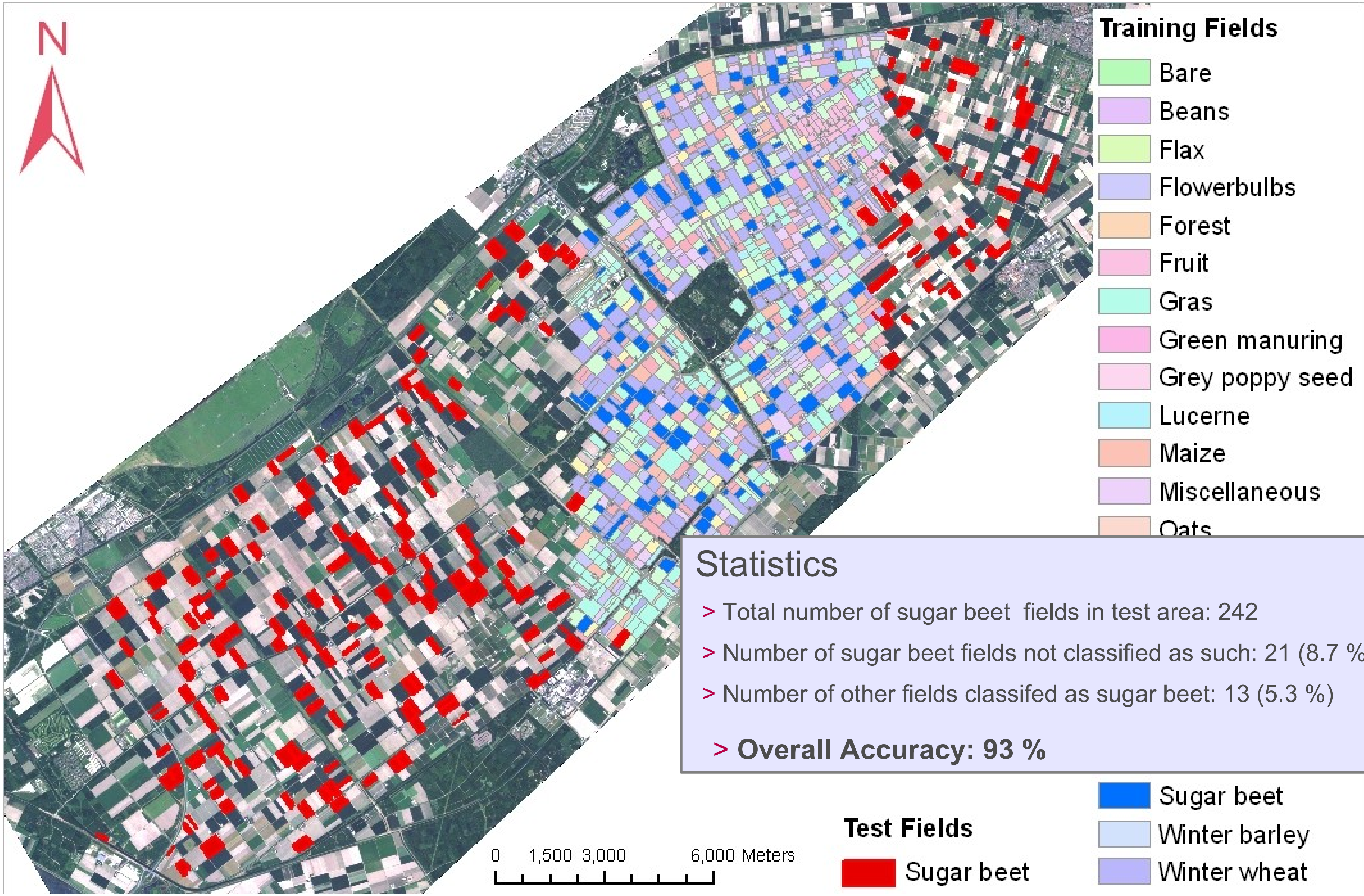
Fields used to train the algorithm



Detailed view of classification results and types of error



Result from applying the trained algorithm to the Test Fields



Statistics

- Total number of sugar beet fields in test area: 242
- Number of sugar beet fields not classified as such: 21 (8.7 %)
- Number of other fields classifed as sugar beet: 13 (5.3 %)

- **Overall Accuracy: 93 %**

Discussion

- RuleSets or DecisionTrees are only applicable for images that were acquired on the same date as the training data
- Accuracy of the input data for training
- Data extraction and import of classification into *.shp file (or raster) is required
- Use of segments allows for fast manual quality control and editing

Vielen Dank!!!

Contact info: schulthess@rapideye.de

<www.rapideye.de>