Master Thesis

submitted within the UNIGIS MSc programme
Interfaculty Department of Geoinformatics - Z_GIS
University of Salzburg

Distantly Trained Random Forest Classifiers for Supervised Crop Type Classification

An Experimental Study Across Multiple European Regions

by

M.Sc. Pia Ferenci
Student number 106880

A thesis submitted in partial fulfilment of the requirements of the degree of
Master of Science – MSc

Advisor:

Dr. Lorenz Wendt

Merching, 15.11.2023

Contents

List of Tables

List of Figures

Statuary Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources and resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Pia Ferenci

Merching, 15.11.2023

Acknowledgments

I would like to extend my sincere appreciation to the individuals whose support has been instrumental in the completion of this master thesis.

First and foremost, I express my deepest gratitude to my husband and best friend, Dejan, and my children, Lana and Luke. Their understanding, patience, and encouragement provided the foundation upon which I could build my academic pursuits. Their sacrifices and consistent support have been a constant source of strength throughout this journey.

I am profoundly thankful to my advisor, Dr. Lorenz Wendt, for his guidance, advice and invaluable insights.

I would also like to express gratitude to my employer, TAUW GmbH, for financially supporting this educational adventure. I appreciate the opportunities for professional development that they have provided.

Additionally I'd like to acknowledge the vibrant StackExchange community for its contributions to the development of my coding abilities. The insightful shared knowledge has broadened my perspectives and enabled me to tackle all technical obstacles which I encountered.

Lastly, my appreciation extends to the entire UniGIS team and classmates who have shared their knowledge and experiences, contributing to a rich and collaborative learning environment.

Thank you all for being pivotal in this journey. Your support has been immeasurable, and I am truly grateful for the impact you have had on my academic and personal growth.

Pia Ferenci

Merching, 15.11.2023

Abstract

Crop type classification using remote sensing data has gained substantial importance in various fields. While supervised machine learning approaches have shown high accuracy in crop type mapping, the availability of training data can be a significant constraint, especially in areas with limited access or conflicts. This study aims to explore the spatial transferability of supervised crop type classification models trained on data from regions differing from the classification regions, addressing the challenge of insufficient ground truth data.

This research concentrates on optical data and a multi-temporal approach. Specifically, the maximum value of the Normalized Difference Vegetation Index (NDVI) calculated over 2-week intervals from April 1, 2018, to October 31, 2018 was utilized in this work due to its often demonstrated suitability in distinguishing between different crop types. The focus of this study lies at investigating whether the spatial distribution and similarities between training and classification regions, including altitude, climate conditions, soil properties, and spatial distance, influence classification results.

The study employed Sentinel-2 optical imagery with its high resolution and frequent revisits for multi-temporal crop classifications based on 2018 field data. It utilized the European Soil Database to identify regions by soil types, Corine Land Cover data to select suitable areas, and Ground Truth Data from the LUCAS survey for training and accuracy assessment. Additionally, elevation information from the Global Mid-resolution Terrain Elevation Data 2010 (GMTED2010) was used to evaluate regional altitudes.

The results of the conducted experiment reveal significant differences in classification success among various crop types, with 'Common Wheat' and 'Maize' showing more promising outcomes compared to 'Barley'. Despite extensive analysis, no clear correlations were found between methodic parameters, region parameters, and classification accuracies. Notably, some regions consistently outperformed others as either training region, classification region or both, suggesting that region-specific conditions may influence classification success. The study highlights the potential of using remotely trained RF classifiers for classification of certain crop types. It promotes the idea of creating representative training datasets for suitable crop types to enable supervised classifications independent of the availability of on-site training data.

Introduction

Crop type classification using remote sensing data is a widely used technique with numerous applications in areas like agriculture, food security or climate modelling (Blickensdörfer et al. (2022), Heupel, Spengler, and Itzerott (2018)). There are abundant approaches for crop type mapping, using varying sensor data, methods and algorithms (Pluto-Kossakowska 2021). Especially machine learning approaches and multi-temporal analysis as well as their combination became increasingly popular during the last years for showing highly accurate results (Benos et al. 2021).

The most basic distinction of methods is the distinction whether a classification is supervised, semi-supervised or unsupervised (Grira, Crucianu, and Boujemaa 2004). Whereas supervised classifications rely on training data for the algorithm to identify different crop types by their different spectral signatures (Perumal and Bhaskaran 2010), unsupervised classifications use statistical clustering methods to group pixels with similar spectral signatures into unlabeled classes (Grira, Crucianu, and Boujemaa 2004). Ma et al. (2020) compared the accuracy of several unsupervised classification algorithms with overall accuracies (OA) between 0.74 and 0.82 (Ma et al. 2020). Pluto-Kossakowska (2021) compared the results of > 50 studies applying different supervised machine learning (ML) algorithms on different sensor data, paying particular attention to the achieved OAs. The different supervised approaches reach OAs of 0.70 - 0.98 % with 73% of them showing an OA of 0.85 or higher (Pluto-Kossakowska 2021).

These findings suggest, that supervised classifications tend to yield higher OAs than unsupervised approaches.

Even though using a supervised classification approach apparently seems to be reasonable, considering the generally higher accuracies to be reached, in some cases it is not possible to collect the absolutely necessary training and test samples for a supervised classification. For crop type classification, this data is usually gathered in the form of in-situ data directly on the fields (Fowler, Waldner, and Hochman 2020). Some areas are naturally difficult to reach or temporarily not available for field data collection due to conflicts or otherwise unsafe conditions. Wang, Azzari, and Lobell (2019) discussed the challenges of crop type-mapping without field-level labels and suggest two possibilities, dealing with a lack of training data: "(1) applying a supervised model trained elsewhere or (2) using regional statistics and an unsupervised learning algorithm." (Wang, Azzari, and Lobell 2019). Wang, Azzari, and Lobell (2019) pointed out that using available training data derived from a region similar to the area of interest might yield in more consistently accurate classifications than an unsupervised approach.

Ringrose et al. (1994) compared the spectral reflectance of green vegetation throughout different climate zones and under different soil conditions. According to this study, spectral signatures of vegetation are correlated to these factors (Ringrose et al. 1994). This can be used to determine, what makes a region "similar" to another region in terms of comparability of the spectral reflectance of the crop types.

Orynbaikyzy, Gessner, and Conrad (2022) assessed the spatial transferability of Random Forest models for crop type classification combining Sentinel-1 and Sentinel-2 data in seven different regions in Germany. The study focused on a comparison of optical-only, radar-only or a combined approach and the reached accuracies for different crop types when transferred to different regions within Germany. Besides the finding that in general best results are reached using Sentinel-1 and Sentinel-2 data in combination, Orynbaikyzy, Gessner, and Conrad (2022) found indications that soil properties - in this case the classification according to the Müncheberger soil quality rating (SQR) (Mueller et al. 2014) - as well as altitude of the sample locations might have an influence on the transferability of the corresponding model.

This thesis aims, similar to the study of Orynbaikyzy, Gessner, and Conrad (2022) and following the suggestion of Wang, Azzari, and Lobell (2019), to overcome a lack of ground truth data by "applying a supervised model trained elsewhere", to further investigate the spatial transferability of supervised crop type classification models, trained on field data from areas differing from the classification regions. In contrast to Orynbaikyzy, Gessner, and Conrad (2022) who focused on evaluation of spatial transferability comparing different sensors, the experiment conducted within the frame of this master thesis concentrated on the influence of spatial distribution and similarities between regions. Hence it did not cover several sensors but focused on optical data only. Particular attention was paid to the question, if similarities between training and classification region in terms of altitude, climate conditions and soil properties as well as their spatial distance have an influence on the achieved classification results. The SQR, which was identified as potentially influencing model transferability between regions by Orynbaikyzy, Gessner, and Conrad (2022), is not available for whole Europe. Instead the soil type of different training and classification regions was taken into account in this experiment.

To meet the objective of this study, the following steps were carried out:

- Identifying several regions throughout Europe with distinct climate conditions and soil properties, which serve as training and classification regions.
- Identifying most abundant crop types in the previously selected regions that serve as target crop types
- Multi-temporal cross-classification: Training of several classification models using multi-temporal training data from the identified regions. Subsequent crop type classification of fields from regions spatially distant to the training regions and accuracy assessment via ground truth data of the respective classification regions.
- Statistical analysis of the achieved classification accuracies, with regard to the similarity of the individual training and classification regions in terms of spatial distance, climate conditions and soil properties.

Methods

To evaluate the spatial transferability of classification models, existing training data from several regions in Europe was used to train several classifiers and subsequently apply it for classifications of all regions individually.

Bannari et al. (1995) compared and summarized 35 different vegetation indices in regard to their field of application and environments for which they are particularly suitable. They concluded that the choice of one vegetation index over another is a complicated task. As the study regions classified in the course of this work are located in very different environments, the most widely used vegetation index NDVI (Rouse 1974) was used to distinguish between different crop types.

The NDVI shows the normalized difference between near-infrared and red reflectance of the surface (Rouse 1974) (equation 1):

$$NDVI = (NIR - Red)/(NIR + Red) \tag{1}$$

Values near 1 indicate strong "greenness" of vegetation whereas values « 1 usually are associated with non-vegetation land cover. As the NDVI reflects "greenness" of the landsurface, it is widely used to evaluate vegetation health (Kinyanjui 2011) but also a common measure for crop type classification (Orynbaikyzy, Gessner, and Conrad 2019).

Zhang et al. (2020) investigated which classification algorithms provide the best classification results using multi-temporal Sentinel-2 imagery. Comparing classifications, using classification and regression tree (CART) decision tree, Support Vector Machine (SVM), and random forest (RF), showed best overall accuracy for RF classifiers (Zhang et al. 2020). Classifications in this thesis were therefore implemented using a set of RF classifiers.

Used Data

Satellite imagery

To calculate the NDVI, optical data is needed which limits the choice of suitable imagery products to optical sensors. Sentinel-2 imagery combines relatively high spatial resolution (10 m for B04 (Red) and B8 (NIR) required for NDVI calculation) and a revisit time of 5 days (ESA 2023) in a multi-spectral sensor and is freely available dating back to June 2015. Therefore it is very suitable for the multi-temporal classifications performed in the course of this experiment, conducted based on field data from the year 2018. Spectral information was abstracted from atmospherically corrected surface reflectance images (SR) from the Sentinel 2 mission.

Soil Data

The European Soil Database v2.0 (European Commission and European Soil Bureau Network 2004) was used to identify regions throughout Europe based on their soil types. Among other information, the data contains a code ("WRBFU"), classifying the soil types according to the World Reference Base for Soil Resources (WRB) (FAO 1998). This attribute built the base for identification of homogeneous regions respecting the soil type.

Land Use and Land Cover Data

LULC data was extracted from the the Corine Land Cover Dataset (European Union 2018). Selecting soil regions exclusively located in areas classified as 211 ("non-irrigated arable land") or 231 ("Pastures, meadows and other permanent grasslands under agricultural use") according to CLC classification (European Union 2018) narrowed down the potential study areas.

Ground Truth Data

Ground Truth Data to train the classification algorithms and test classification results, originates from the Land Use and Coverage Area frame Survey (LUCAS) (Eurostat (2018a)), which is carried out in 28 European countries each three years. LUCAS data provides information on the actual crop types of thousands of fields throughout Europe and therefore can be used as basic information for training and accuracy assessment. Standardization of methodology of data collection and description ensures comparability of the obtained data throughout Europe (Eurostat 2018a). LUCAS data contains information on land use and land cover and additional information concerning the data collection and reliability of the data itself. This descriptive information was used to identify suitable and reliable data points.

Elevation Data

To evaluate the results regarding altitude of the different regions, the publicly available *Global Mid-resolution Terain Elevation Data 2010 (GMTED2010)* (Danielson and Gesch 2011) was used. It provides elevation information with a spatial resolution of 250 m with world-wide coverage.

Data Preparation

As spectral reflectance of crops not only depends on the crop type itself, but among other factors, it is linked to the soil background in the examined areas (Prudnikova et al. 2019), regions for training and classification were selected in a way, that the study regions are located on one for each region homogeneous soil type. As crop classification by definition deals with the classification of agricultural areas, any non-agricultural areas were excluded from the potential study regions. Adding ground data, containing information on land use and cultivation allowed for narrowing down the potential study areas to several regions with each showing an homogenous soil type, located on agricultural area and containing a suitable number of ground data describing different crop types. As spectral reflectance of vegetation is sensitive to climatic conditions (Ringrose et al. 1994), the spatial distribution of potential areas is taken into account, to allow for comparison of classification results between distant respectively close regions with similar or different climatic conditions.

1. Exclusion of non-agricultural land and mixed land use classes

In a first stage, suitable training regions and target crop types were identified, analyzing the abundance and spatial distribution of soil types and crop types in the available datasets.

To narrow down the potential study regions and reach a manageable amount of data, only areas of specific land use according to CLC 2018 (Figure 1) were taken into account. As crop types in various regions in Europe were to be classified, agricultural areas with a very restricted spatial distribution as vineyards were excluded from the potential study areas. Additionally, agricultural areas with complex cultivation patterns and areas with a high amount of natural vegetation were not taken into account. Finally, as this study focused on the distinction between crop types, also areas classified as "fruit tree and berry plantation" were excluded.This selection of distinct land use classes ensured, that the results of the classifications reflect the ability of the classifiers to distinguish between different crop types resp. grassy areas opposed to a differentiation between crop types and a very broad range of land use classes (Figure 2). Clipping LUCAS and soil data to the remaining agricultural CLC classes 211 and 231 (Figure 1) provided the basis for the further isolation of distinct study regions.

211	Non-irrigated arable land
221	Vineyards
222	Fruit tree and berry plantations
231	Pastures, meadows and other permanent grasslands under agricultural use
242	Complex cultivation patterns
243	Land principally occupied by agriculture, with significant areas of natural vegetation

Figure 1: Nomenclature and code of different agricultural areas according to CLC 2018 (European Union 2018)

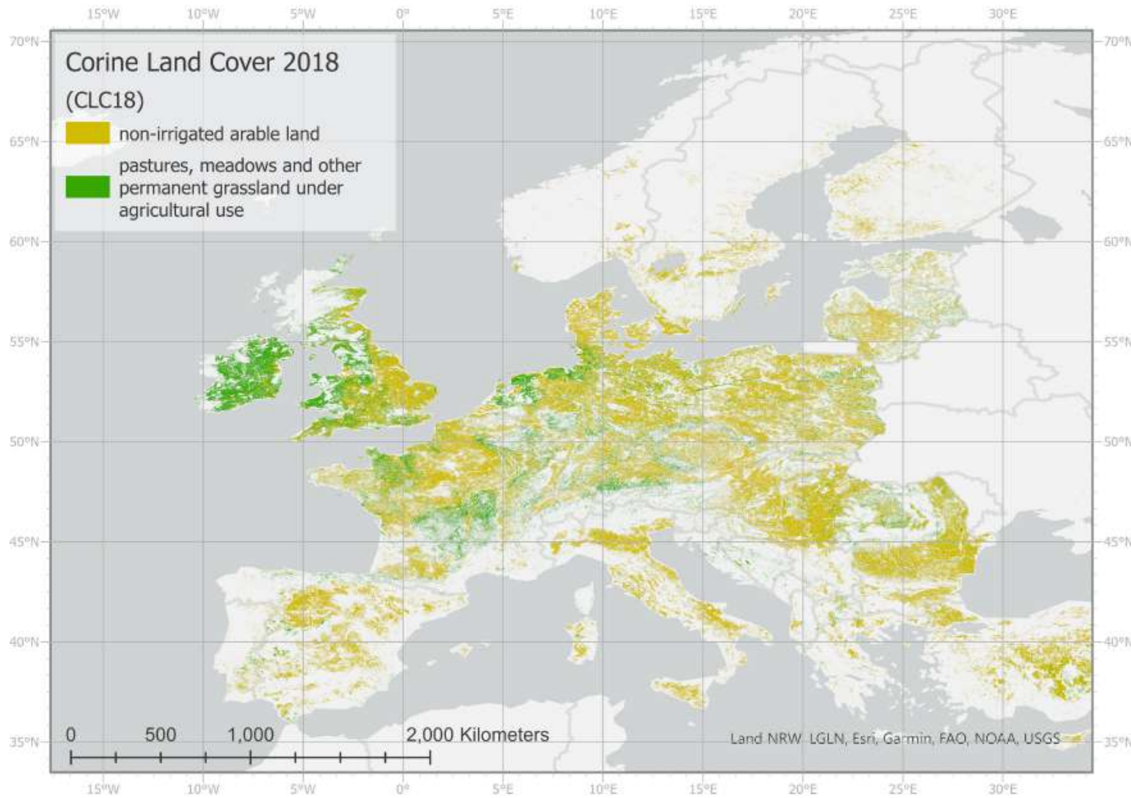

Figure 2: Areas suitable for crop type classification (CLC18-Code 211 resp. 231) (European Union 2018)

2. Selection of target crop types

This experiment was designed to evaluate if different crop types throughout Europe can be distinguished using remote sensing, with classifiers trained in areas differing from the classification areas. Suitable crop types therefore need to be abundant in ideally all of the eventually selected study regions. The three most abundant crop types in the LUCAS data set clipped to the suitable CLC land use types 211 and 231 are 'Common Wheat', 'Barley' and 'Maize' (Eurostat 2018b) (Figure 3).

3. Preparation of LUCAS data

The following information abstracted from LUCAS micro data was used to filter the LUCAS data set and identify suitable and reliable data points:

- GPS distance to point [m]: ≤ 10
- GPS coordinate system: WGS84
- GPS precision [m]: ≤ 5 (i.e. an inaccuracy of max. 5 m)
- Type of Observation: Field survey, point visible, ≤ 100 m

As the classifications used the NDVI calculated from Sentinel-2 bands B04 and B08 with a spatial resolution of 10 m (ESA 2023), the accuracy of the GPS coordinates and the distance between data point and GPS position point needed to be in a similar order of precision. Yet the decision on exclusion limits is a compromise between gaining high reliability of the location of data points and the necessity to keep a sufficient amount of data points for training and testing during classification.

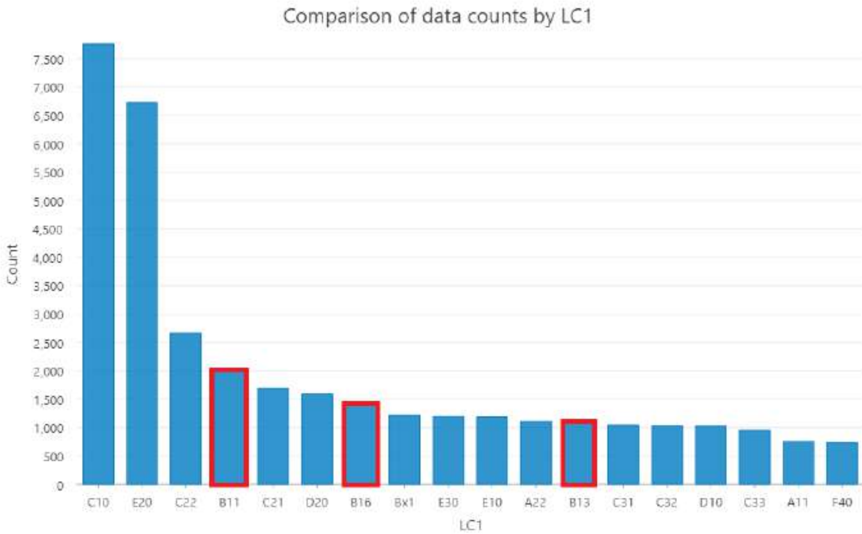

Figure 3: Number of data points by land cover class (LC1) with B11 = 'Common Wheat', B13 = 'Barley' and B16 = 'Maize' and non-crop landcover occupying classes <> 'Bx' (based on data of European Union 2018)

Using "Type of Observation" as filter increased confidence in the correctness of the crop classification during the field survey, which is the very basis for successful classification using remote sensing. To ensure high validity of the training data, data points classified from a distance > 100 m or photo-interpreted were excluded from the data set.

The parameter "Point Longitude E/W" was used to identify data points which are located west of the prime meridian. The GPS coordinates recorded during field survey show values > 0 which implies, the longitude of points with "Point Longitude E/W" = 2 (Eurostat 2018c) (which are located west of the prime meridian) has to be multiplied by -1 to position the corresponding point at the correct location of data acquisition.

The data set, filtered and modified as described, was used to identified areas with a high amount of data points representing crop types 'Common Wheat', 'Barley' and 'Maize'. Extraction of areas meeting this criterion was accomplished by creating buffer zones of 50 km around all point data representing the selected target crops (Figure 4), clipping them to the identified areas with CLC code 211 respectively CLC code 231 (Figure 4) and subsequently matching them with the most abundant soil regions, identified in the next step of data preparation.

4. Analyzing abundance of soil types

To create an experimental setting, where results of classifications with the same respectively different soil types in training and classification regions can be compared, abundance, area size and spatial distribution of soil types according to FAO (1998) were analyzed. The objective was to find several soil types with the following characteristics:

1. large connected areas of one distinct soil type
2. spatial distribution of soil types in several distant regions in Europe
3. containing a sufficient amount of data points representing the identified target crop types 'Common

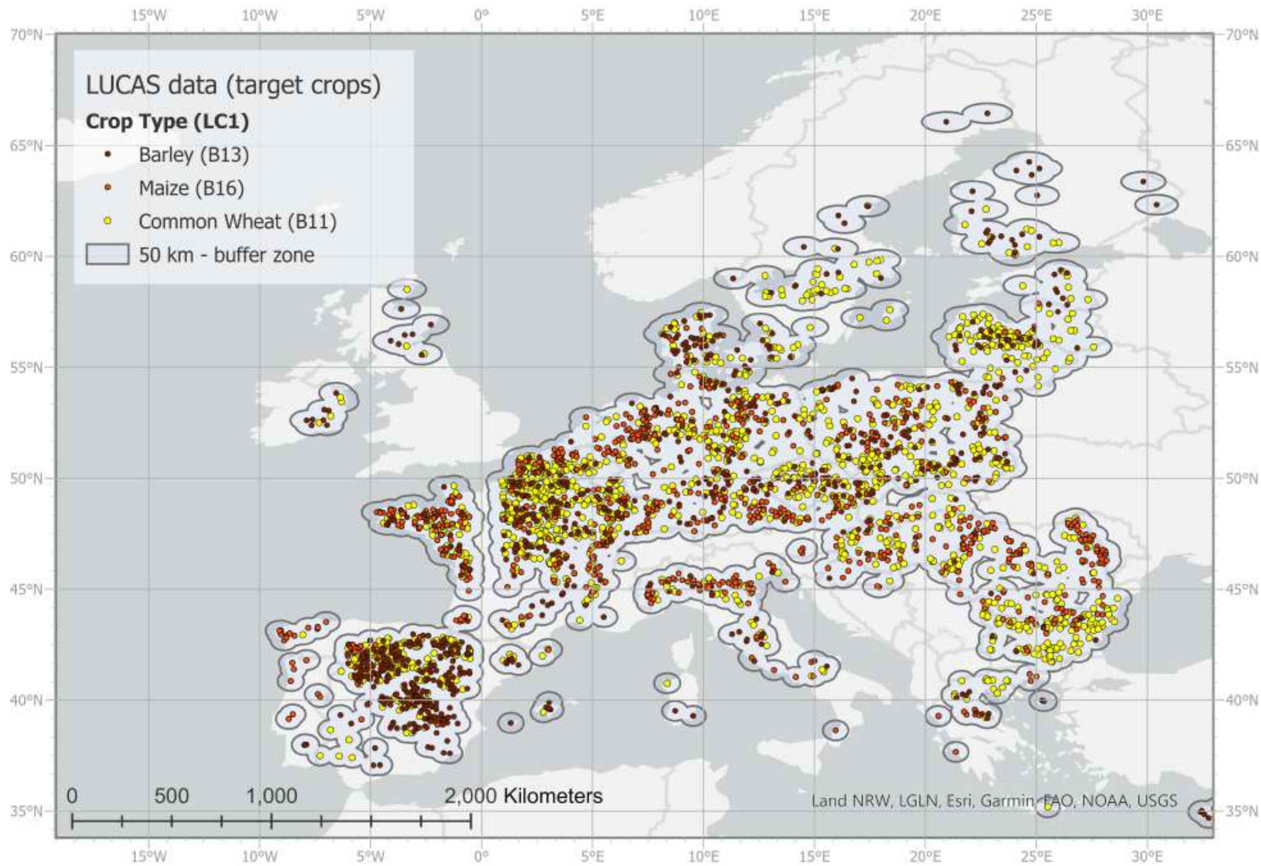

Figure 4: Extraction of zones around LUCAS data points of target crops (European Union 2018)

Wheat', 'Barley' and 'Maize'

Data of the European Soil Database v2.0 (European Commission and European Soil Bureau Network 2004) was clipped to the previously created buffer area around the LUCAS points (clipped to CLC 211 and 231). Borders between features of the same soil type ("WRBFU") were dissolved and the resulting areas of homogeneous soil types filtered by size, taking only areas > 1.000 km² into account. Comparing the area sum of those homogeneous soil regions of substantial size grouped by soil type showed the most prevalent soil types to be *Calcaric Cambisol (CMca)*, *Dystric Cambisol (CMdy)*, *Eutric Cambisol (CMeu)*, *Haplic Luvisol (LVha)* and *Haplic Podzol (PZha)* (Figure 5).

The identified "soil type regions" were compared against criterion 2 and 3 mentioned above. The limitation of data points during preparation of ground truth data lead to difficulties in implementation of criterion 3 ("containing a sufficient amount of data points representing the identified target crop types 'Common Wheat', 'Barley' and 'Maize') for a sufficient number of regions. Considering the partially sparse number of data points covering the regions, a number of two data points was considered sufficient for training a classifier in one region. To enable statistically evaluable results, a number of approximately 100 cross-classifications was targeted which requires a number of at least 11 regions (as each of n regions was used as training region for ideally n-1 classification regions).

Among those regions fulfilling criterion 1 and 2, 14 areas were identified as at least fairly fulfilling criterion 3

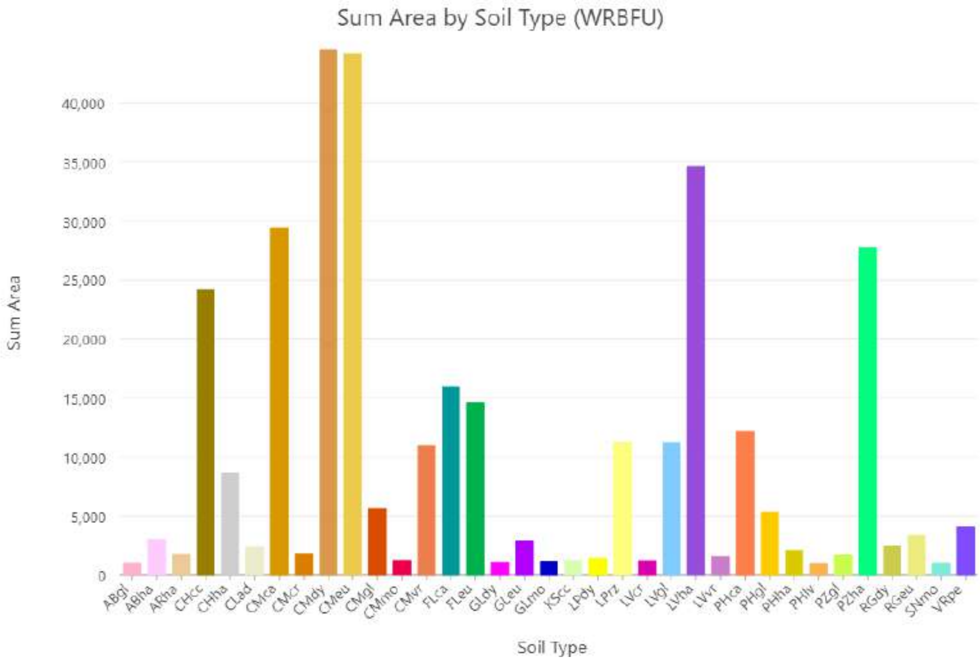

Figure 5: Accumulated area by soil type, taking into account only soil regions of min. 1.000 km² area. (based on data of European Commission and European Soil Bureau Network 2004)

(Figure 6), (Table 1).

As displayed in Table 1, eight out of 14 regions were not covering all three classes 'Common Wheat', 'Barley' and 'Maize'. Regions with a sufficient number of data points for each class can be used to train classifiers for classification of all other regions. Regions with a lack of data may only be utilized as training data for classification of regions which do not contain data points of other classes than the training regions, as this would have a direct impact on the classification results in a negative way. Table 2 shows all possible combinations of training and classification regions, which results in a number of 92 classifications in total.

Each data point represents one agricultural field of specific crop type resp. land cover. To provide a sufficient number of training pixels, for each data point of the target crop types, a polygon was manually drawn in GIS to raise the amount of training pixels per class per region. As a result, for each region, areas of altogether approximately 300.000 - 1.100.000 m² were assigned one of the classes 'Common Wheat', 'Barley', 'Maize' respectively 'Other' (Figure 7).

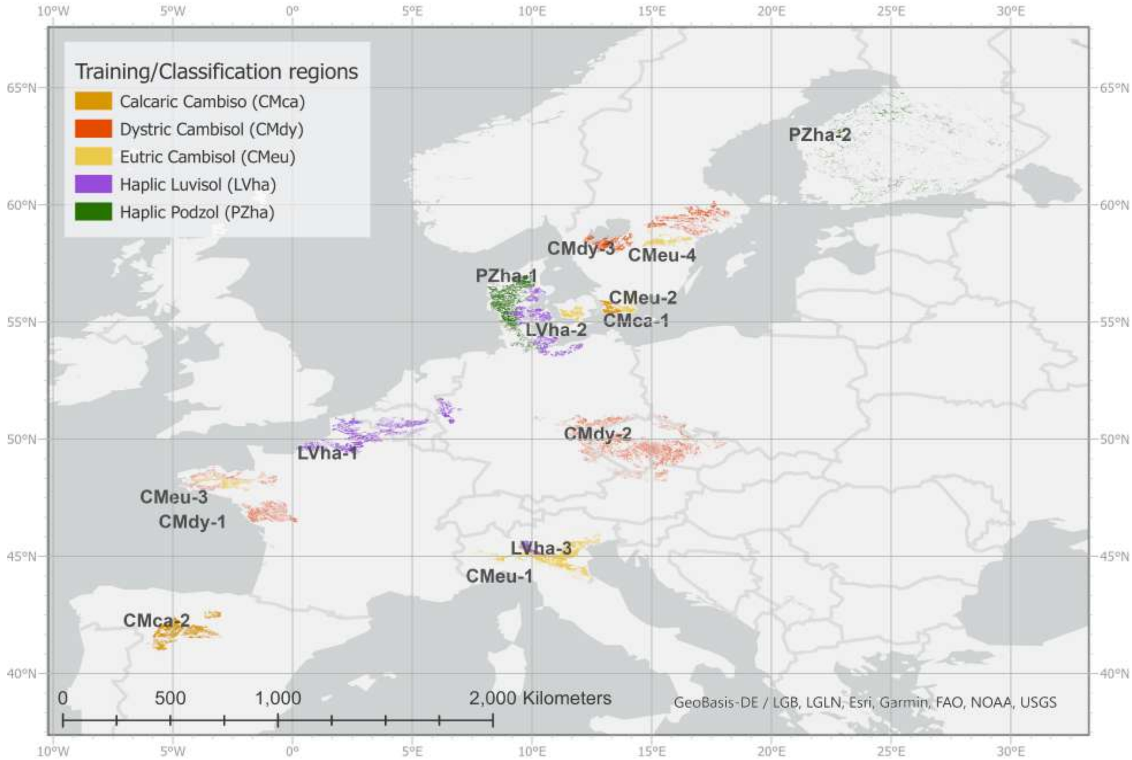

Figure 6: Regions of homogenous soil type selected for crop type classification

Table 1: Identified regions of homogenous soil type with number of ground truth data points for classes 'Other', 'Common Wheat', 'Barely' and 'Maize'

region	sum	other	common wheat	barley	maize
CMca-1	17	7	6	4	0
CMca-2	146	86	29	31	0
CMdy-1	68	29	11	6	22
CMdy-2	137	87	19	16	15
CMdy-3	79	58	16	5	0
CMeu-1	29	6	12	0	11
CMeu-2	26	20	4	2	0
CMeu-3	23	8	4	2	9
CMeu-4	30	18	12	0	0
LVha-1	95	27	50	7	11
LVha-2	38	15	9	10	4
LVha-3	15	7	1	0	7
PZha-1	64	39	1	12	12
PZha-2	78	61	5	12	0

Table 2: Combination of training and classification regions, based on the availability of data points per class. Classifiers may only be used for classification of regions which do not contain different classes than those covered by the training region.

training region	common wheat	barley	maize	classification regions	number of classifications
CMdy-1	11	6	22	all 13	13
CMdy-2	19	16	15	all 13	13
CMeu-3	4	2	9	all 13	13
LVha-1	50	7	11	all 13	13
LVha-2	9	10	4	all 13	13
PZha-2	5	12	0	CMca-1, CMca-2, Cmdy-3, Cmeu-2, Cmeu-4	5
CMdy-3	16	5	0	CMca-1, CMca-2, Pzha-2,Cmeu-2, Cmeu-4	5
CMca-2	29	31	0	CMca-1, Cmdy-3, Pzha-2,Cmeu-2, Cmeu-4	5
CMeu-2	4	2	0	CMca-1, Cmdy-3, Pzha-2,Cmeu-2, Cmeu-4	5
CMca-1	6	4	0	CMca-2, Cmdy-3, Pzha-2,Cmeu-2, Cmeu-4	5
CMeu-1	12	0	11	Cmeu-4, Lvha-3,	2
CMeu-4	12	0	0	no training	0
LVha-3	1	0	7	no training	0
PZha-1	1	12	12	no training	0
Total	179	107	91	-	92

Depending on the availability of data points, the reached ground truth areas vary in total size. A potential influence of the number of ground truth pixel on classification results is going to be discussed at a later stage. As class-imbalance in training data is considered to tendentially decrease classification accuracy (Mellor et al. 2015), particular attention was drawn to create a preferably balanced data set concerning class coverage during field digitization. Due to the under-respresentation of classes ‘Barley’ and ‘Maize’, corresponding agricultural fields have been digitized almost in their entire field sizes, whereas individual ground truth fields for ‘Common Wheat’ and ‘Other’ generally are smaller in size but larger in number.

The digitized agricultural fields, each assigned as one of the classes 0 - 3 (with ‘Other’ = 0, ‘Common Wheat’ = 1, ‘Barley’ = 2, ‘Maize’ = 3) were subsequently used as training respectively test data for the classifications performed in the course of this experiment.

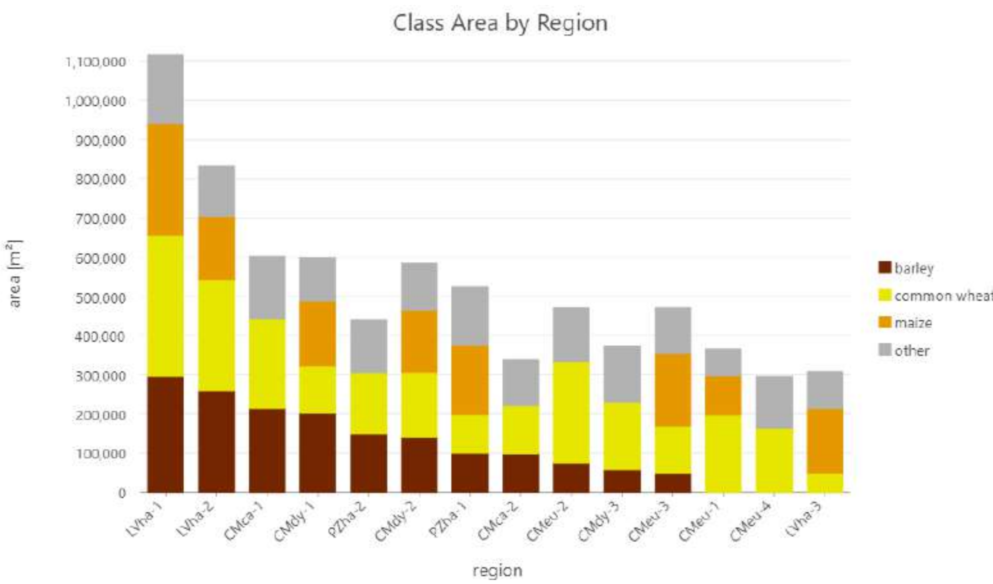

Figure 7: Accumulated area per class and soil type region

Crop Type Cross-Classifications

Classifications were performed using Google Earth Engine (GEE).

Input Data

Training data was provided as polygon shapefile, created from LUCAS data points as described in the previous chapter. Another shapefile containing training resp. classification regions was used to limit calculations to those areas.

Data Preprocessing

Data sets containing ground truth data and region extents were uploaded to the GEE platform. Sentinel 2 imagery from a date range of 2018-04-01 to 2018-10-31 was accessed via GEE. A simple geometry, covering all study regions and directly drawn in GEE was used as spatial filter. The provided regions from the input data set as spatial filter contain too many edges to be processed in GEE so that using a simple geometry was more effective in this case. The Normalized Difference Vegetation Index (NDVI) was calculated from bands B4 (Red) and B8 (VNIR) for all obtained images in the Sentinel image collection. A new image collection was created, containing only bands “NDVI” and “SCL”. The “SCL” band provides a pixel-wise classification which was used as cloud filter in the further course.

To enable multi-temporal classifications, the image collection containing data from 2018-04-01 to 2018-10-31 and bands “NDVI” and “SCL” was divided into time slots of 14 days. From all available images of each 14-day time interval, the highest reached “NDVI” of each pixel was retained and a new image was assembled from the thereby identified “greenest pixels”. The resulting 14 greenest-pixel images were composited to one resulting image, containing bands NDVI 01-14 and SCL 01-14. Those bands contain information on the maximum NDVI value from each 14-day time interval and the corresponding pixel class according to “SCL”. The classifications in this experiment are based on bands NDVI 01 - NDVI 14, which depict the course of the

NDVI in each pixel throughout vegetation period.

The described steps of data preparation resulted in a single image, containing 28 bands (NDVI 01-14, SCL 01-14). The image was clipped to all 14 regions to create image subsets for subsequent classifications.

Training and Classification

Ground truth data containing class information (class 0-3) and imagery from the corresponding region containing multi-temporal NDVI values were combined in one *FeatureCollection* per region to train one RF classifier for each region.

Using the maximum NDVI value served the goal to obtain pixels from a relatively cloud-free date of recording. Yet a cloud filter had to be applied to minimize negative effects of cloudiness on the classification results. For this purpose, features from the resulting *FeatureCollections* with SCL = 0 (no data), 8 (medium-probability clouds) and 9 (high-probability clouds) were excluded from classification.

As training and classification did not take place at the same location, there was no need to distinguish between training and test data. All available ground truth data for one region was used for training of the corresponding classifier and all available ground truth data for all classified regions could be used for accuracy assessment. Training and test data for each individual region were extracted from the above created and cloud-filtered *FeatureCollection*.

Reached OAs as well as number of training pixels per class and contents of the confusing matrices of all classifications were stored in a table for subsequent analysis of the results.

Due to processing limitations of the free version of GEE, not all classifications could be calculated in one run. The corresponding script therefore had to be split up and individually modified to run about five classifications at once. The used scripts are accessible on GEE and additionally attached to this thesis in Appendix A1-2. All classification results and derived data is shown in Appendix B1-3.

Multi-class classifications

Results

Overall Accuracy of the classification results ranges from very low values (min. OA = 0.13) to very good results (max. OA = 0.92) (Figure 8) (Appendix C: Table 5). To extend insight in the classification qualities, User's and Producer's Accuracies (UAs and PAs) per class were calculated from the confusion matrices.

The classifications show OAs nearly normally distributed and slightly skewed to the left (Figure 9). Regarding the individual classes 0 to 3 however the reached average accuracies (avg. from User's Accuracy and Producer's Accuracy) show different distributions. With a mean average accuracy of 0.24, and a strongly right skewed distribution (Figure 10c)), identification of 'Barley' showed poorest results. The mean average accuracy of 'Common Wheat' classification lies at 0.51 with a bimodal distribution peeking at ≤ 0.05 respectively between 0.5 and 0.65 (Figure 10b)). The class 'Other' shows an average accuracy of 0.69 and a left skewed distribution (Figure 10a)). Half of classifications show accuracies of > 0.72 for class 'Other'. Best class recognition was achieved for class 'Maize' with an average mean accuracy of 0.83 and a median of 0.84 (Figure 10d)).

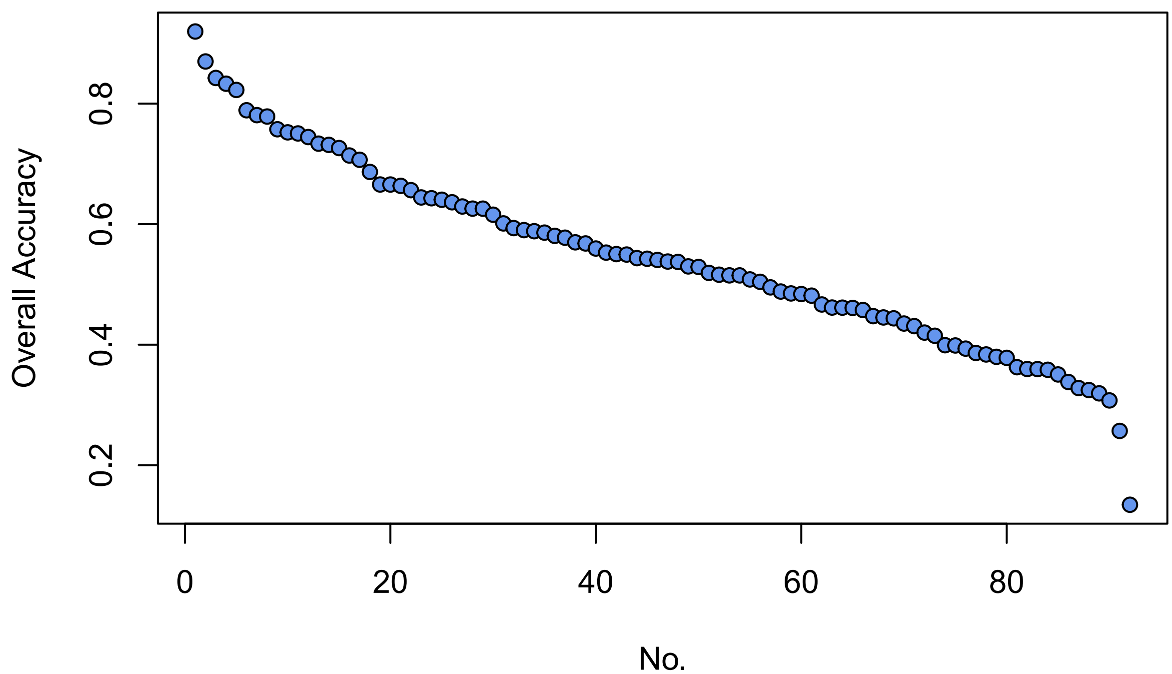

Figure 8: Overview of reached Overall Accoracies (OAs) for all multi-class classifications

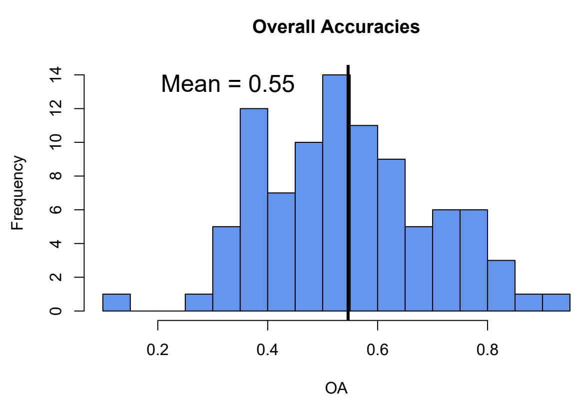

Figure 9: Overview of reached Overall Accoracies (OAs) for all multi-class classifications

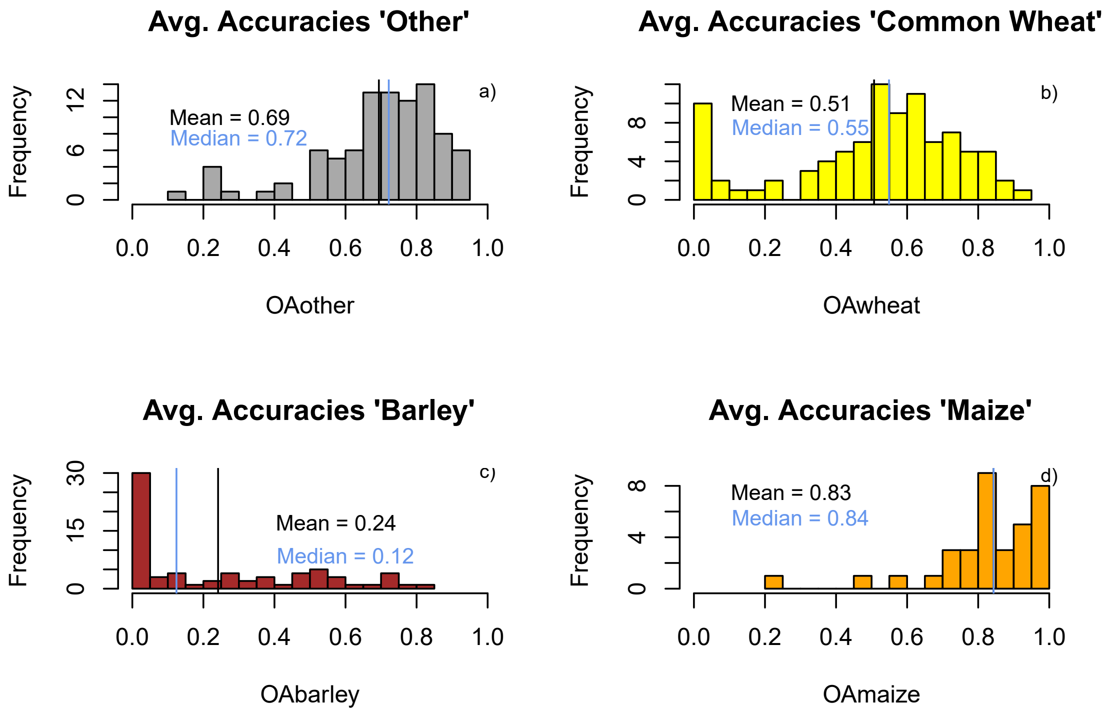

Figure 10: Reached Overall Accoracies (OAs) separated by class

Influence of class balance

To evaluate the influence of class balance as described by Mellor et al. (2015), an index to measure imbalance of class representation in the used training data was developed.

For this purpose, in a first step, the share of pixels per class was calculated and standardized by number of classes (equation 2):

$$cb0...3 = ((px0...3/(px0 + px1 + px2 + px3))/(1/n)) \tag{2}$$

with

cb0. . . 3 = class specific imbalance of classes 0 to 3

px0. . . 3 = pixel count of classes 0 to 3

n = number of classes

This calculation returns values close to 1 for classes with a pixel count share near to 1/n of the total of used training data. Values > 1 indicate over-representation of pixels for the corresponding class and values < 1 accordingly indicate under-respresentation of the class in the data set.

The resulting class-specific values were used to determine the overall imbalance of data for each classification by summing up their absolute deviation from 1 and subsequently dividing the resulting value by number of classes n (equation 3):

$$imbalanceIndex = (|cb0 - 1| + |cb1 - 1| + |cb2 - 1| + |cb3 - 1|)/n \quad (3)$$

with

cb0. . . 3 = class specific imbalance

n = number of classes

Low values of the *Imbalance Index* indicate equally sized classes whereas high values imply higher imbalance regarding pixel counts of the individual classes.

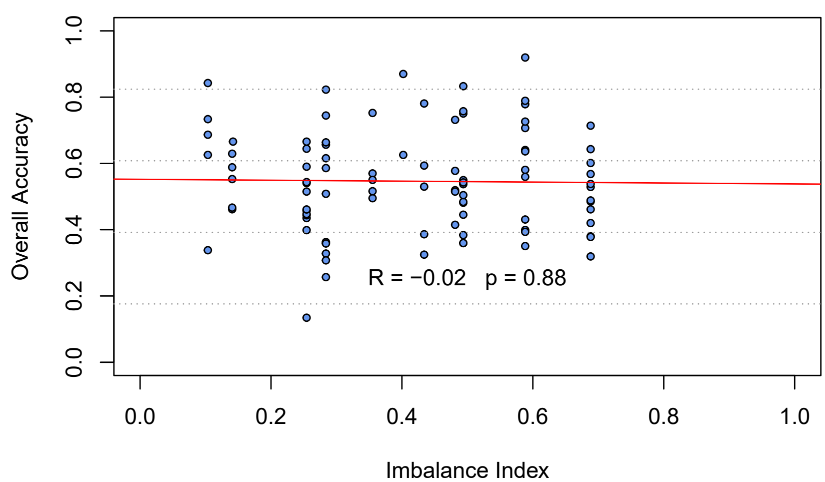

Figure 11: Correlation between Imbalance Index and reached accuracies

The evaluation of a possible correlation between class balance and reached OAs results in a Pearson's R of -0.02 with a p-value of 0.88. The high p-value, which can be translated to a low confidence interval of 12%, verifies that the relation between the variables is to be considered random. Consequently the results show *no significant correlation* between class balance and reached OAs (Figure 11).

Binary classifications

The results of the multiclass-classifications indicate a strong variance of identifiability of the different crop types using the depicted approach.

Overall Accuracy is a measure of the ability of a classifier to correctly classify test pixels of all present classes. Thus it represents a mixed quality measure which does not allow for drawing further inferences from it in this case of strongly differing accuracies for the individual classes. Therefore, in the further course, classification results of barley were excluded from analysis. Instead of OAs of the multi-class classifications, reached accuracies for 'Common Wheat' and 'Maize' were examined separately in binary classifications.

There are generally two possible approaches for separating the achieved accuracies per class:

- pre-processing approach:

Before executing the classifications, the input data set containing training and test data is modified in a way that it is suitable for a binary classification. For this purpose, data which represents 'Barley' and 'Maize' resp. 'Barley' and 'Common Wheat' is assigned to class 0 'Other' whereas the crop type to be identified ('Common Wheat' resp. 'Maize') is assigned to class 1. This approach is time intensive as all classifications have to be run again for both classes.

A less time consuming approach is the re-calculation of accuracies from the multi-class classification results as followed:

- post- processing approach:

Overall Accuracies are calculated as the total of all correctly classified pixels out of all classified test pixels. To calculate OAs from the multi-class results without taking into account class 'Barley' and 'Maize' resp. 'Common Wheat', the confusion matrices are consulted. The confusion matrix of a classification gives information on which test pixels were classified as which class. Calculating OAs for a 'Common Wheat' (= class 1) classification therefore means summing up all class 1 test pixels correctly classified and all test pixels from class 0, 2 and 3 not classified as class 1. The number of in this sense correctly classified pixels divided by the total number of classified test pixels gives the post-processing calculated OA for 'Common Wheat' classifications. OA for 'Maize' (class 3) can be calculated analogously, summing up classes 0, 1 and 2. This approach is less time-consuming as re-classification is not necessary. Nonetheless, this approach is based on the assumption, that those crop types which are assigned to class 0 ('Other') in post-processing were also correctly classified as class 0 in a regular binary classification (i.e. the pre-processing approach).

The decision which approach to follow in the further course of this study was based on a brief comparison of both approaches using class 'Maize' for classification. Only eight out of the 14 regions have a suitable number of 'Maize' training sites, which means a total of 56 re-classifications, deploying the pre-processing approach. The hereby reached Overall Accuracies were compared to the post-processing calculated OAs for classification of 'Maize'. The mean absolute deviation between pre- and post-processing OAs lies at 0.026 with pre-processing OAs by an average of 0.011 lower than the corresponding results of the OAs calculated in the post-processing approach (Table 3). Pre-processing allows for a higher number of classifications, as in the post-processing approach, some region combinations which are possible for 'Maize' classification were not executed due to a lack of data points for at least one of the other crop classes investigated in the multi-class classification. Nevertheless, using the *post-processing approach for 'Common Wheat'* is reasonable as both approaches result in similar OAs and all 92 classifications show results for classification of 'Common Wheat'. As the differences in reached accuracies between pre- and post-processing approach are neglectable and re-classification regarding class 'Maize' allows for a higher number of classifications and thus a more robust basis for statistical analysis of results, *pre-processing results of 'Maize' classifications* are evaluated in the further course.

Table 3: Comparision of reached Overall Accuracies for classification of class 'Maize' using the pre- resp. post-processing approach

No.	Training Region	Classified Region	OA pre-processing	OA post-processing	|OA pre - OA post|	OA pre - OA post
1	LVha3	CMeu3	1.000	0.997	0.003	0.003
2	LVha2	CMeu3	1.000	1.000	0.000	0.000
3	CMdy1	CMeu3	1.000	1.000	0.000	0.000
4	PZha1	CMeu3	1.000	-	-	-
5	LVha2	CMdy1	0.997	0.996	0.001	0.001
6	CMeu3	CMdy1	0.994	0.997	0.002	-0.002
7	LVha1	CMdy1	0.992	0.999	0.007	-0.007
8	LVha1	CMeu3	0.991	-	-	-
9	CMeu3	LVha1	0.989	0.991	0.002	-0.002
10	CMeu1	CMdy1	0.988	-	-	-
11	PZha1	CMeu1	0.986	-	-	-
12	CMeu1	LVha3	0.984	0.999	0.015	-0.015
13	PZha1	CMdy1	0.981	-	-	-
14	LVha2	LVha1	0.979	0.995	0.016	-0.016
15	CMdy1	LVha1	0.973	0.997	0.024	-0.024
16	LVha3	CMdy1	0.967	-	-	-
17	CMdy1	LVha2	0.963	0.963	0.000	0.000
18	CMeu3	LVha2	0.963	0.953	0.010	0.010
19	CMdy1	PZha1	0.963	0.963	0.000	0.000
20	PZha1	LVha2	0.963	-	-	-
21	PZha1	LVha1	0.962	-	-	-
22	CMdy2	LVha2	0.962	0.960	0.002	0.002
23	LVha1	LVha2	0.962	0.962	0.000	0.000
24	CMeu1	LVha1	0.962	-	-	-
25	CMdy2	CMdy1	0.961	0.933	0.029	0.029
26	LVha1	CMeu1	0.952	0.952	0.000	0.000
27	CMeu1	LVha2	0.950	-	-	-
28	LVha2	PZha1	0.949	0.958	0.008	-0.008
29	CMeu3	PZha1	0.949	0.918	0.031	0.031
30	LVha3	LVha1	0.947	-	-	-
31	PZha1	CMdy2	0.946	-	-	-
32	CMdy2	PZha1	0.932	0.916	0.015	0.015
33	LVha3	CMeu1	0.929	-	-	-
34	LVha1	PZha1	0.924	0.924	0.000	0.000
35	CMeu3	CMeu1	0.920	0.930	0.010	-0.010
36	CMeu1	CMdy2	0.918	-	-	-
37	CMeu1	CMeu3	0.918	-	-	-
38	CMeu1	PZha1	0.917	-	-	-

No.	Training Region	Classified Region	OA pre-processing	OA post-processing	|OA pre - OA post|	OA pre - OA post
39	CMdy2	LVha1	0.911	0.935	0.024	-0.024
40	CMdy1	CMdy2	0.907	0.911	0.004	-0.004
41	CMdy2	CMeu1	0.904	0.896	0.008	0.008
42	CMdy1	CMeu1	0.899	0.955	0.056	-0.056
43	CMeu3	CMdy2	0.894	0.824	0.070	0.070
44	LVha2	CMdy2	0.893	0.925	0.033	-0.033
45	LVha2	CMeu1	0.893	0.898	0.005	-0.005
46	LVha3	LVha2	0.890	-	-	-
47	CMeu3	LVha3	0.874	0.773	0.101	0.101
48	LVha1	CMdy2	0.872	0.922	0.050	-0.050
49	PZha1	LVha3	0.857	-	-	-
50	LVha1	LVha3	0.853	0.916	0.063	-0.063
51	CMdy1	LVha3	0.834	0.966	0.133	-0.133
52	LVha3	CMdy2	0.821	-	-	-
53	LVha3	PZha1	0.808	-	-	-
54	CMdy2	CMeu3	0.762	0.756	0.005	0.005
55	LVha2	LVha3	0.714	0.925	0.211	-0.211
56	CMdy2	LVha3	0.684	0.696	0.012	-0.012
57	Average	-	-	-	0.026	-0.011

Results for binary 'Maize' and 'Common Wheat' classifications

To answer the research question, for both, 'Maize' and 'Common Wheat' classifications, Overall Accuracies were calculated according to the post-processing ('Common Wheat') resp. pre-processing ('Maize') approach and used for further analysis. OAs provide a general measure for the quality of the classification results regarding both classes, class 0 ('Other') and class 1 ('Maize' resp. 'Common Wheat'). As the main interest is the ability of the classifiers to identify 'Maize' resp. 'Common Wheat', User's and Producer's Accuracy for those target classes were additionally evaluated.

Overall accuracies for 'Maize' classifications show a mean of 0.93 and median of 0.95 (Figure 12) with 53 out of 56 classifications (95 %) reaching an OA > 0.8 (Appendix D: Figure 23). User's Accuracies for 'Maize' show a mean of 0.93 and a median of 0.98 (Figure 13c)) with 50 out of 56 classifications (89 %) reaching an UA > 0.8 (Appendix D: Figure 31)). Producer's Accuracies for 'Maize' show a mean of 0.72 and median of 0.75 (Figure 13a)) with 25 out of 56 classifications (45 %) reaching an PA > 0.8 (Appendix D: Figure 30). 'Maize' classifications which show good results (> 0.8) for both, PA and UA, account for a total number of 25 (45 %) of all 'Maize' classifications (Appendix D: Figure 25 and Figure 24.

Overall accuracies for 'Common Wheat' classifications show a mean of 0.66 and median of 0.69 (Figure 12) with 18 out of 92 classifications (20 %) reaching an OA > 0.8 (Appendix D: Figure 26. User's Accuracies for 'Common Wheat' show a mean of 0.69 and median of 0.80 (Figure 13d)) with 48 out of 92 classifications (52 %) reaching an UA > 0.8 (Appendix D: Figure 34). Producer's Accuracies for 'Common Wheat' show a mean of 0.45 and median of 0.46 (Figure 13b)) with 18 out of 92 classifications (20 %) reaching an PA > 0.8

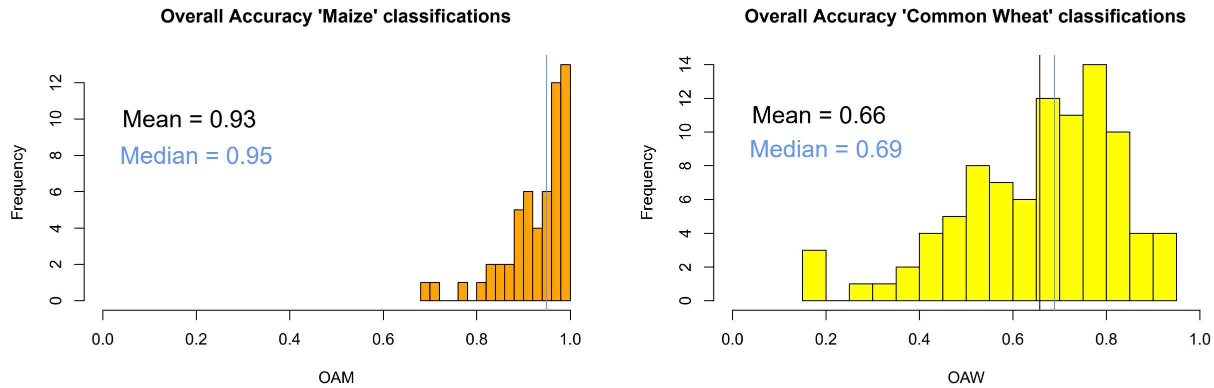

Figure 12: Reached Overall Accuracies (OAs) for binary Maize resp. Common Wheat classifications

(Appendix D: Figure 33). ‘Common Wheat’ classifications which show good results (> 0.8) for both, PA and UA, account for a total number of 8 (9 %) of all ‘Common Wheat’ classifications (Appendix D: Figure 28 and Figure 27).

Figures 24 and 25 (Appendix D) show classifications of region ‘CMdy1’ and ‘CMeu1’ achieving very high accuracies in nearly all ‘Maize’ classifications. Contrary to this, ‘Maize’ classifications in region ‘LVha2’ do in no case simultanously show PAs and UAs > 0.8 which is solely attributable to low to medium PAs in all classifications. CMdy2 does not yield high accuracies in most of the classificationsn, neither as training nor as classification region. Training data from region ‘PZha1’ is most successfully used for ‘Maize’ classification in the other regions with OAs between 0.86 and 1.00.

‘Common Wheat’ classifications with simultanously high OA, Pa and UA are executed in region ‘CMeu4’. Training data from region ‘CMeu3’ results in highest accuracies for ‘Common Wheat’ classification in other regions (Appendix D: Figures 28 and 27). Due to the low to medium PAs in most of the ‘Common Wheat’ classifications, none of the regions continuously shows simultaously high PAs and UAs.

Even though class ‘Other’ shows naturally more diversity in its spectral profiles, spectral profiles of classes ‘Common Wheat’ and ‘Other’ (Appendix E: Figure 37 - 39) show similarities in the general shape with samples peaking in the first and last third of the investigated time range and a depression in between. Only regions ‘CMca2’ and ‘LVha3’ only show high NDVI values in the first half of this time period but still showing similarities between ‘Common Wheat’ and ‘Other’ spectral profiles.

Spectral profiles of classes ‘Maize’ and ‘Other’ (Appendix E: Figure 35 - 36) show less similarities in the general shape as ‘Maize’ is peaking in roughly the time range of lower values of class ‘Other’ samples.

Influence of model characteristics

Besides potential influences from differing climate and soil conditions, quality of classifications also depends on characteristics of the classifications itself.
Even though RF classifiers are considered relatively insensitive to training set size (Rodriguez-Galiano et al. 2012), the reached accuracies were tested for correlation with methodological characteristics of the individual classifications to identify potential influences by them.

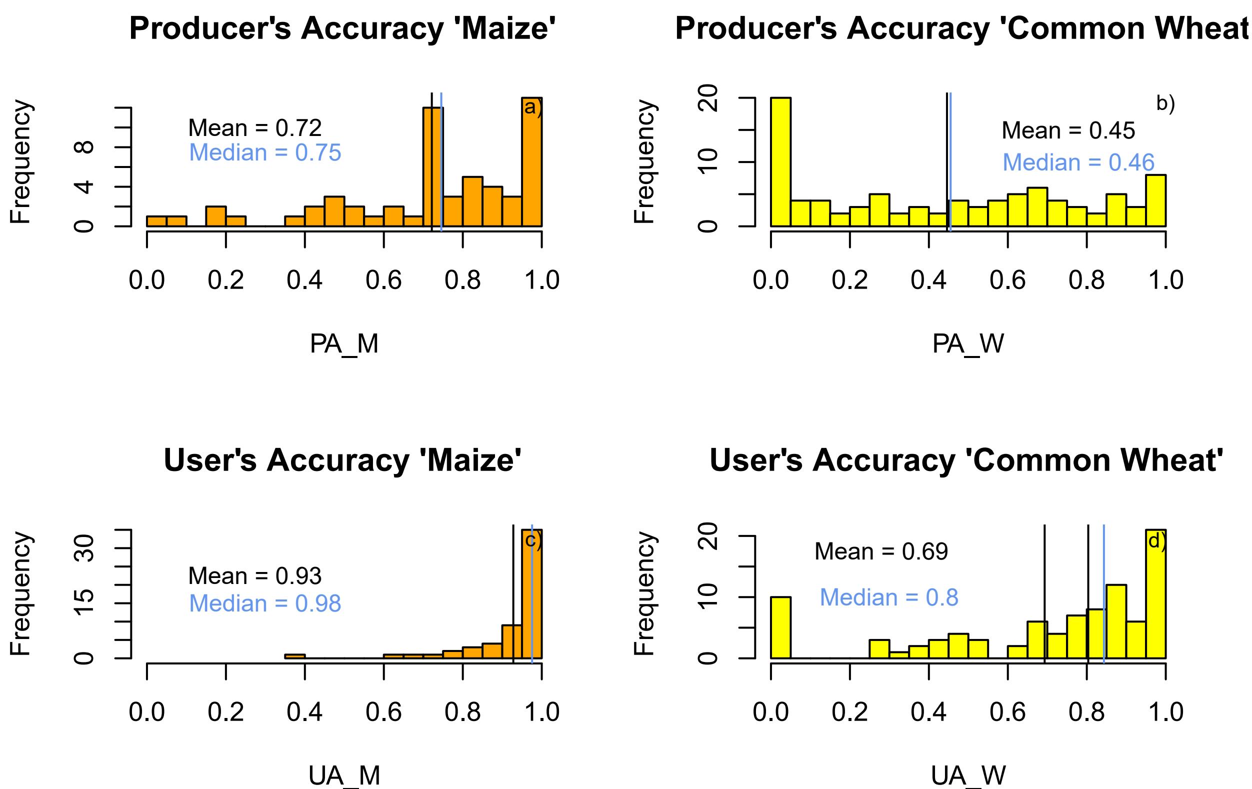

Figure 13: Producer's (PA) and User's (UA) Accuracies for classes Maize resp. Common Wheat in binary classifications

Class imbalance

Due to the initially planned multi-class approach, the prepared training and test data set showed a high imbalance of class representation in regard to pixel count per class for the binary classifications. Even though no significant relation between class imbalance and classification accuracies was detected in the multi-class classifications - with its comparably evenly distributed training sets - a potential correlation between imbalance and results was evaluated in the stronger imbalanced binary classifications (Figure 14).

Correlation tests show Pearson's R-values between -0.39 and 0.29 and the associated p-values from 0.00 to 0.71.

For 'Common Wheat' classifications, the analysis revealed a weak to medium negative correlation, which was highly significant for Producer's Accuracy (PA) and Overall Accuracy (OA) with p-values of 0.0 (confidence level of 100%). However, this negative correlation was not significant when assessing User's Accuracy (UA) and the Imbalance Index.

On the other hand, 'Maize' classifications exhibited a weak positive correlation for OA and PA, with corresponding R-values of 0.05 and 0.29. This positive correlation was significant for PA, with a confidence level of 97%. There was a weak but not significant correlation observed between UA and the Imbalance Index.

In summary, the results indicate a significant, weak to medium negative correlation between the Imbalance

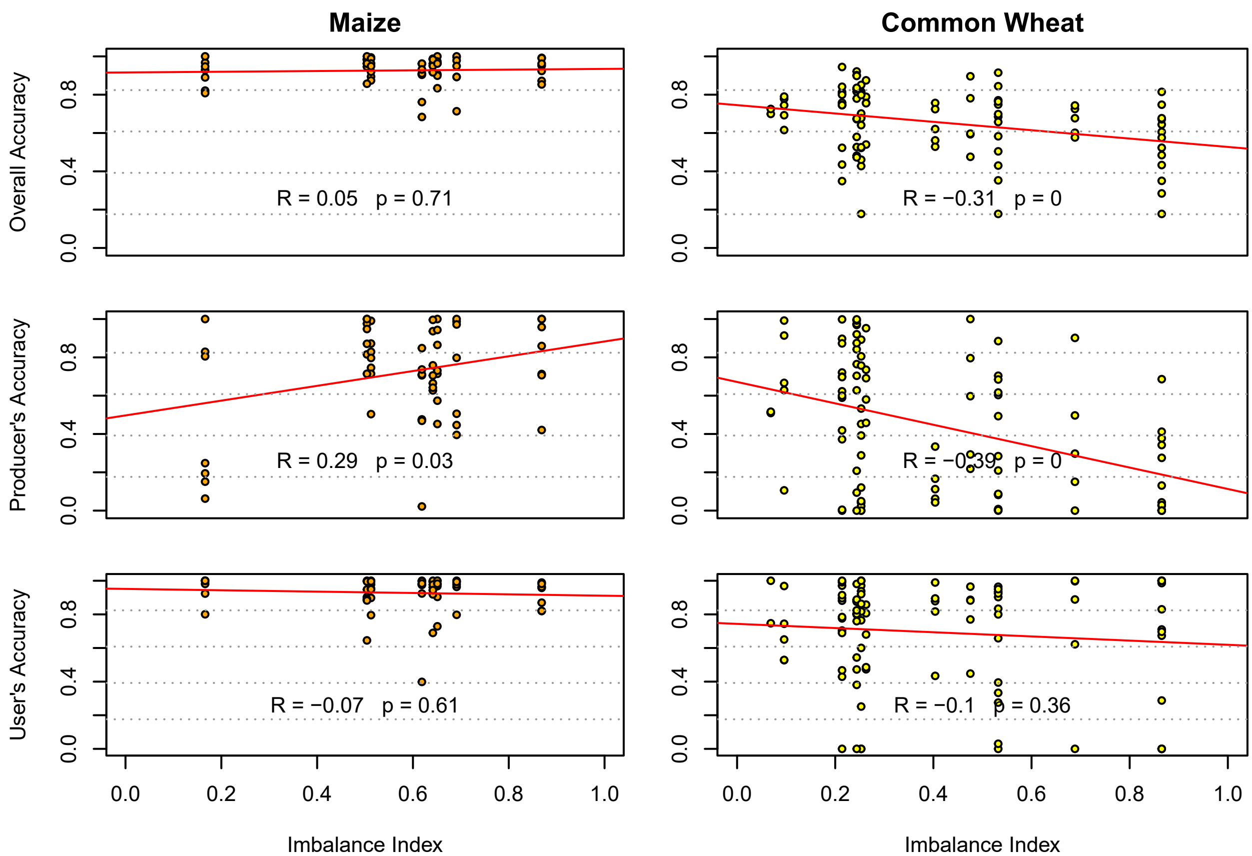

Figure 14: Correlation between class imbalance and reached accuracies

Index and PA/OA for 'Common Wheat' classifications. Additionally, there is a significant, medium positive correlation between the Imbalance Index and PA for 'Maize' classifications. However, no significant correlations were observed between class imbalance and OA/UA for 'Maize' classifications and UA for 'Common Wheat' classifications (Figure 14).

Number of training pixels

Especially the impact of amount and quality of training data is discussed in literature, with among others Rodriguez-Galiano et al. (2012) researching their influence on random forest land-cover classifications. Pal and Mather (2003) emphasize the effect of training set size, especially for heterogenous classes to ensure representation of all the variability present in one class (Rodriguez-Galiano et al. (2012), Pal and Mather (2003)).

Evaluating the potential correlation between the number of training pixels and achieved accuracies, Pearson's R-values were found to range from -0.09 to 0.11, with corresponding p-values ranging from 0.4 to 0.99. These R-values indicate a generally weak, mostly positive relationship between reached accuracies and the number of training pixels across various datasets. The high p-values associated with these relationships, which translate to low confidence intervals ranging from 1% to 60%, suggest that the observed connections between the variables are likely random and lack statistical significance.

In summary, the results revealed that there is no significant correlation between the count of training pixels and the achieved accuracies. (Figure 15).

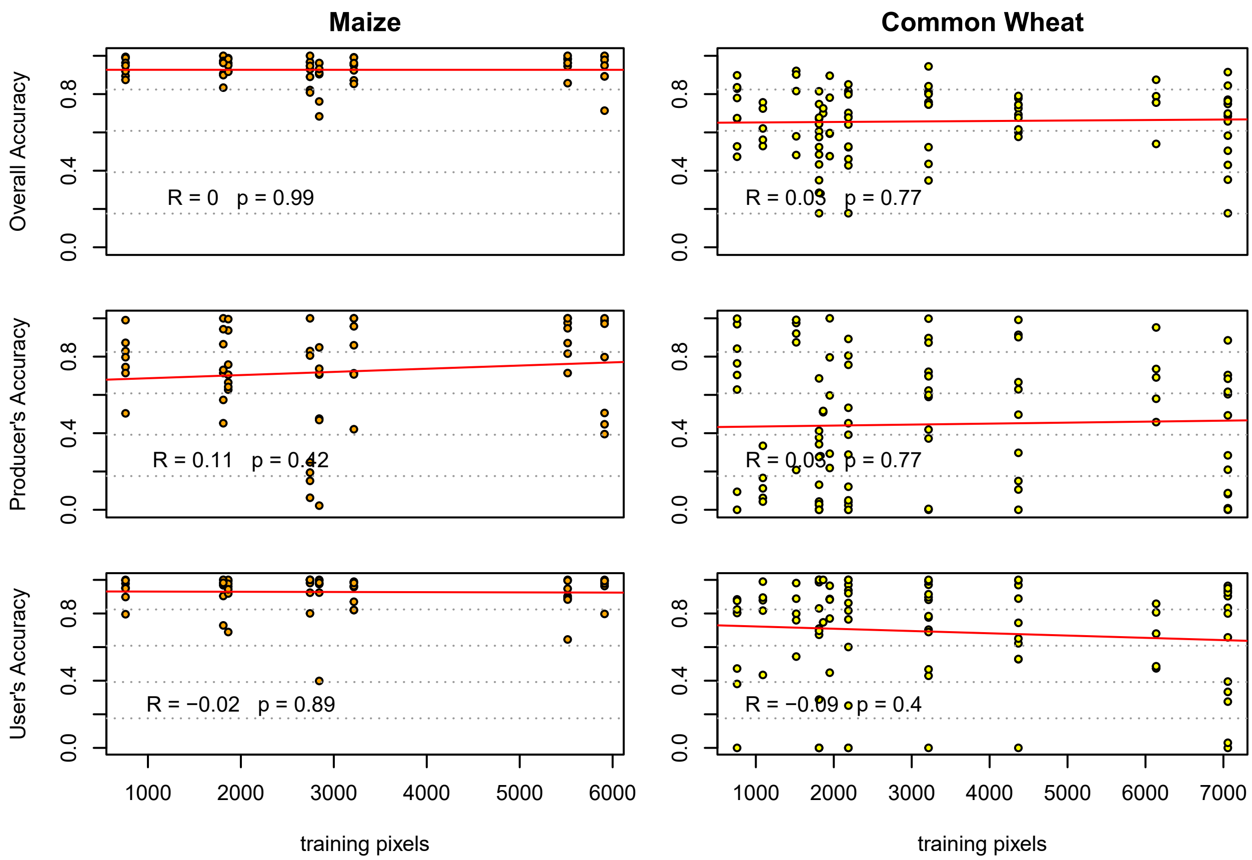

Figure 15: Correlation between number of training pixels and reached accuracies

Number of test pixels/Number of training pixels

As heterogeneity potentially increases with a larger number of classified (test) pixels, not only the total number of pixels was taken into account but also the ratio of test to training pixels for each classification. The value of the resulting quotient provides a measure for potential heterogeneity of test pixels in comparison to the potential heterogeneity of training pixels. Large values indicate higher potential heterogeneity of classified test areas compared to the training data the classification is based on.

The correlation tests revealed Pearson's R-values ranging from -0.06 to 0.24 and corresponding p-values between 0.02 and 0.66. These R-values suggest a generally weak, mostly positive relationship between achieved accuracies and the ratio of test to training pixels across the various datasets.

However, it's important to note that the correlation between User's accuracy (UA) for 'Common Wheat' classifications and the test-to-training pixel ratio is notably weak but statistically significant, with a correlation coefficient of 0.24 and a 98% confidence interval.

The high p-values associated with the other assessed relationships, indicating low confidence intervals ranging from 34% to 83%, suggest that these connections between variables are likely random and not statistically significant.

In summary, the findings indicate a weak yet significant correlation between the test-to-training pixel ratio and User's accuracy in 'Common Wheat' classifications. However, no significant correlations were observed between the test-to-training pixel ratio and User's accuracies for 'Maize' classifications, and for Overall accuracies (OAs) and Producer's accuracies (PAs) of both crop types (Figure 16).

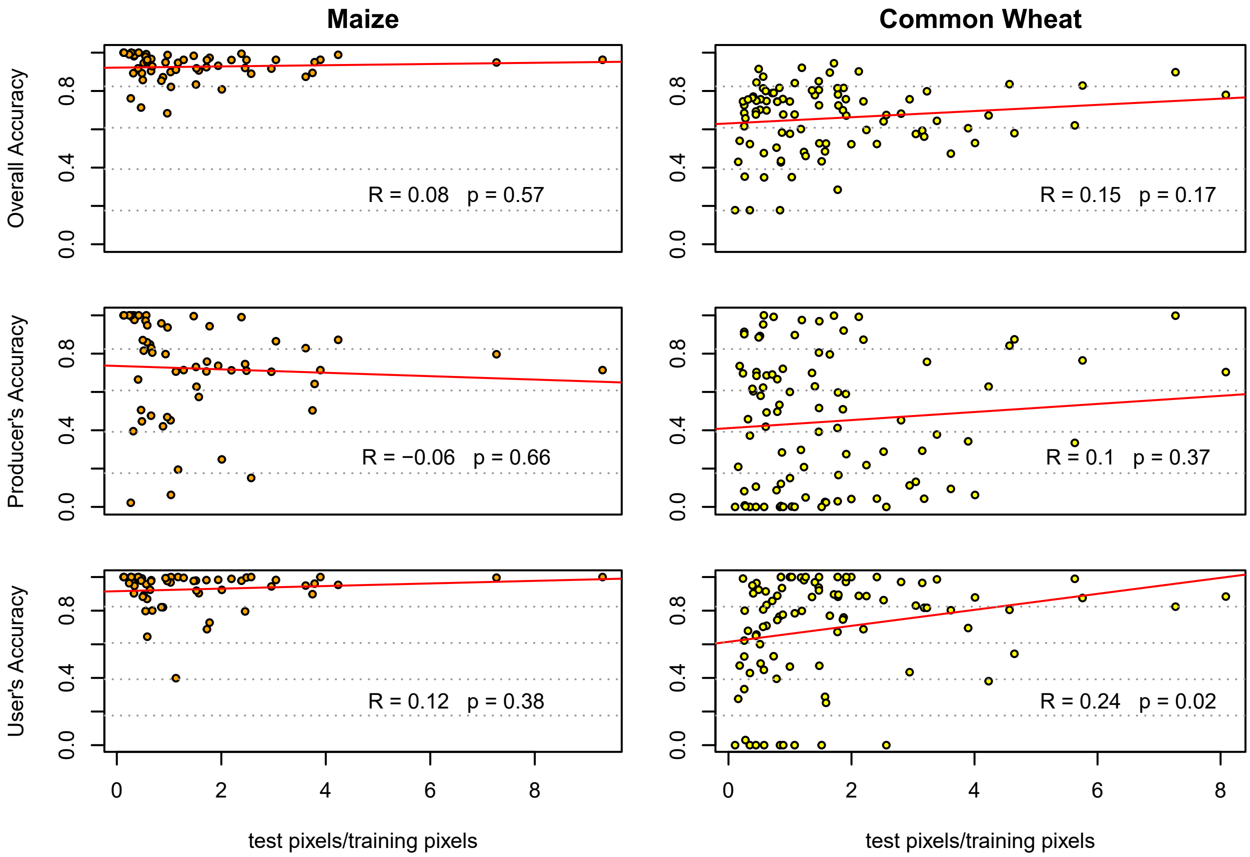

Figure 16: Correlation between ratio of test to training pixels and reached accuracies

Number of training sites

Accounting for the small number of training sites in some regions (Table 1), accuracies were tested for correlation with the number of used training sites. A low number of sites creates higher homogeneity of the training data and therefore might lead to poorer results as representation of all the variability present in the corresponding class might not be given (Pal and Mather 2003).

Correlation tests show Pearson's R-values between -0.1 and 0.07 and p-values of 0.47 to 0.99. The r-values indicate that there is only a weak relation between reached accuracies and the number of training sites in the several datasets which furthermore points in opposite directions. The high p-values, which can be translated to low confidence intervals of 1% to 53%, verify that the relation between the variables is to be considered random. Consequently the results show no significant correlation between number of training sites and reached accuracies (Figure 17).

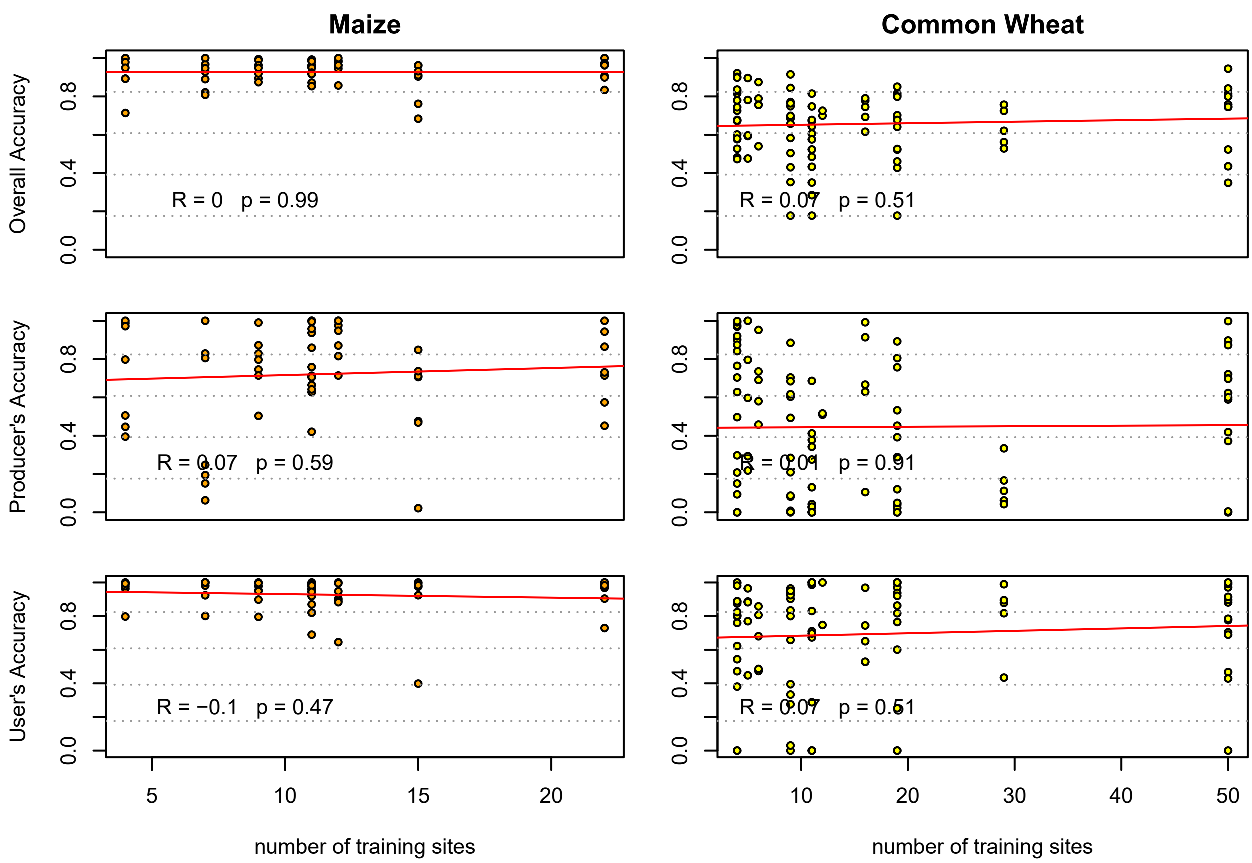

Figure 17: Correlation between number of training sites and reached accuracies

Correlation between classification results and region properties

Soil Type

As already proposed by Ringrose et al. (1994), soil background has an influence on the spectral characteristics of vegetated areas. Recent research, such as the study by Piedallu et al. (2019), delves deeper into the impact of soil parameters on the Normalized Difference Vegetation Index (NDVI) derived from satellite data. Notably, factors like water availability and soil nutrition have been identified as crucial contributors to the greenness and, consequently, the resulting NDVI values (Piedallu et al. 2019). Moreover, Orynbaikyzy, Gessner, and Conrad (2022) identified the Soil Quality Rating (SQR) as a potentially valuable metric for characterizing soil properties, impacting the spatial transferability of Random Forest (RF) models.

Given the limited availability of publicly accessible data with comprehensive information on water availability and soil nutrition, and the uneven availability of the SQR across European regions, the assumption was made that soil falling into the same soil type class according to the World Reference Base for Soil Resources (WRB) (FAO 1998) share similar soil conditions. To validate the classification results, this study assessed the correlation between classification quality metrics, including Overall Accuracy (OA), Producer's Accuracy (PA), and User's Accuracy (UA), and the equality resp. inequality of soil types between the training and classification regions (Appendix D: Figures 29 - 34).

Figures 29 - 34 show that for 'Maize' and 'Common Wheat' classifications, classifications with equal soil type in training and classification region show OAs, PAs and UAs covering the full range from very low values of 0.03 to very high values of 1. There is no correlation between equality of soil type and better classification results.

Total Distance

According to *Tobler's first law of geography* ("everything is related to everything else, but near things are more related than distant things." (Tobler 1970)), the correlation between spatial proximity of training and classification region and the reached classification accuracies was examined. Distances between the regions were approximated by using the center points (identified using GIS) of each region for calculation. The distances were ascertained using equation 4 to calculate the distance of two points located on a sphere, following the laws of spherical trigonometry (Kells, Kern, and Bland 1940):

$$distance = r * acos(sin(lat1) * sin(lat2) + cos(lat1) * cos(lat2) * cos(lon2 - lon1)) \tag{4}$$

with

r = radius (approx. 6.370 km),

lat1/lat2 = latitudes of both locations,

lon1/lon2 = longitudes of both locations

Correlation analyses were conducted, revealing Pearson's R-values ranging from -0.22 to 0.07, with corresponding p-values falling within the range of 0.04 to 0.59. The computed r-values suggest that there exists a generally weak, predominantly negative correlation between the achieved classification accuracies and the total spatial distance between training and classification regions. Specifically, the p-value associated with the correlation analysis between User's Accuracy (UA) for 'Common Wheat' classifications and the total distance

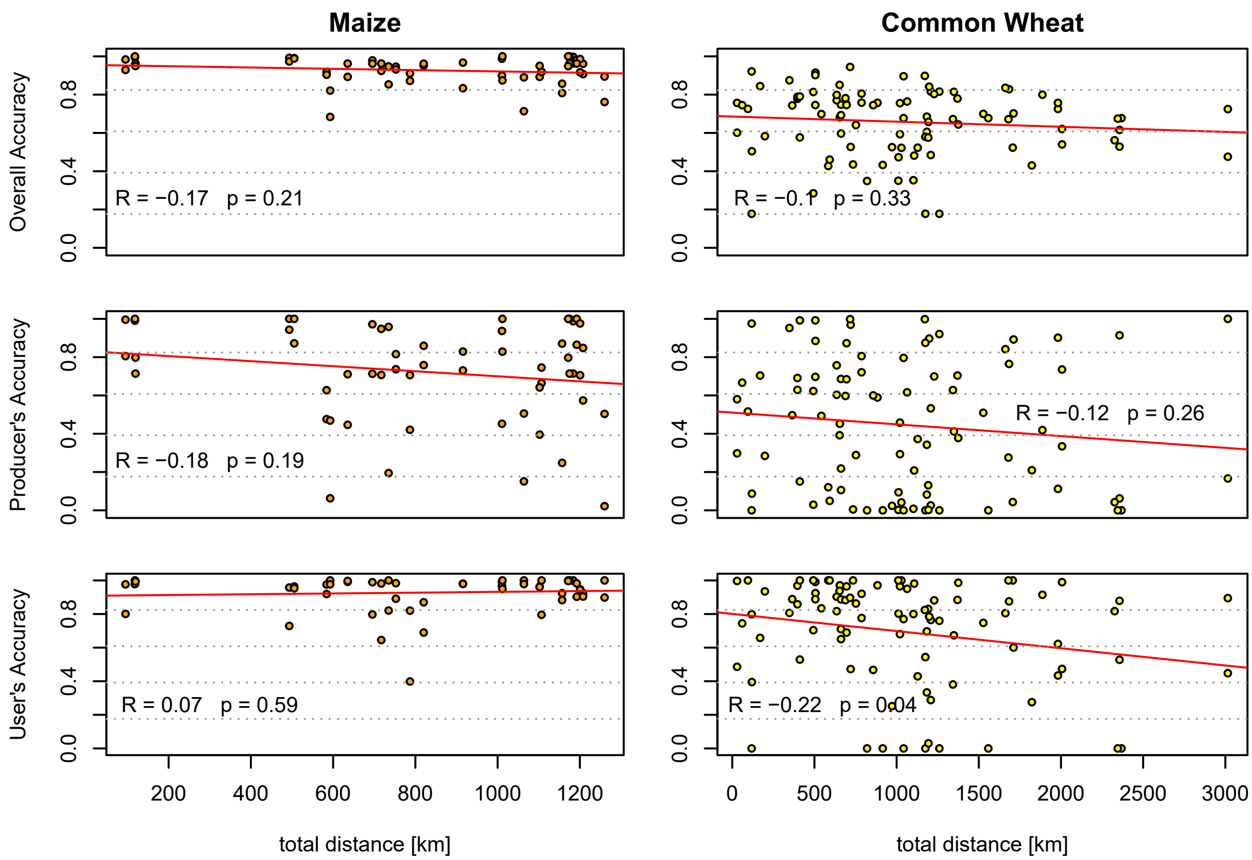

Figure 18: Correlation between total distance of training and classification region and reached accuracies

separating the training and classification regions indicates a statistically significant, albeit weak, negative correlation of -0.22 within a confidence interval of 96%.

The high p-values of the other evaluated relations, which can be translated to low confidence intervals of 41% to 81%, verify that the relation between the variables is to be considered random.

In summary, the findings point to a weak yet statistically significant negative correlation between the total distance separating regions and User's Accuracy in the context of 'Common Wheat' classifications. However, no statistically significant correlations were identified between the total distance and User's Accuracy for 'Maize' classifications, or between the total distance and Overall Accuracy (OA) or Producer's Accuracy (PA) for both crop types (Figure 18).

Lateral Distance

To explore a potential influence of climate conditions (Ringrose et al. 1994) in training and classification regions on the reached accuracies, correlation between lateral distance and classification results was examined.

Climate is influenced by a multitude of parameters (Stevens 2010). Typically classification of regional climate is accomplished by categorizing regions into climate zones as initially acquired by Köppen (1900), with it's latest version published in 1961 by Geiger (1961). This kind of classification results in nominal classes which each comprises a set of characteristics, qualifying a region for one or another climate zone.

To take into account the possible influence of climatic conditions (Ringrose et al. 1994) prevailing in the individual regions, a statistical analysis of metric values instead of nominal classes is favorable. Therefore, in the presented study, latitude served as proxy for climate conditions of the regions as latitude is one of the important influencing parameters when it comes to regional climate (Stevens 2010).

Analogous to the calculation of the absolute distance between regions, the center point of each region was located using GIS. The lateral distance was then calculated by multiplying the distance between latitudes by 111 km which is the approximate distance between two neighboring latitudinal lines and thereby gave the lateral distance between the two center points (equation 5).

$$latDistance = |lat1 - lat2| * 111km \tag{5}$$

Correlation tests produced a range of Pearson's R-values from -0.14 to 0.17, with corresponding p-values between 0.17 and 0.63 (Figure 19). These results suggest that there is only a weak connection between classification accuracies and the lateral distance between training and classification regions.

The relatively high p-values, which imply confidence levels between 37% and 83%, indicate that this relationship between variables appears to be random.

In summary, the study did not find any statistically significant correlation between the lateral distance separating regions and the achieved classification accuracies (Figure 19).

Altitude difference

Orynbaikyzy, Gessner, and Conrad (2022) found indications, that regions located at higher average altitudes show greater losses in classification accuracies compared to regions at lower altitudes.

To analyze a possible relation between classification accuracies and altitude, for each training resp. classification region, the average altitude was ascertained, combining the region polygons with the publicly available elevation dataset *Global Mid-resolution Terain Elevation Data 2010 (GMTED2010)* (Danielson and Gesch 2011) and calculating local statistics in GIS (Table 4).

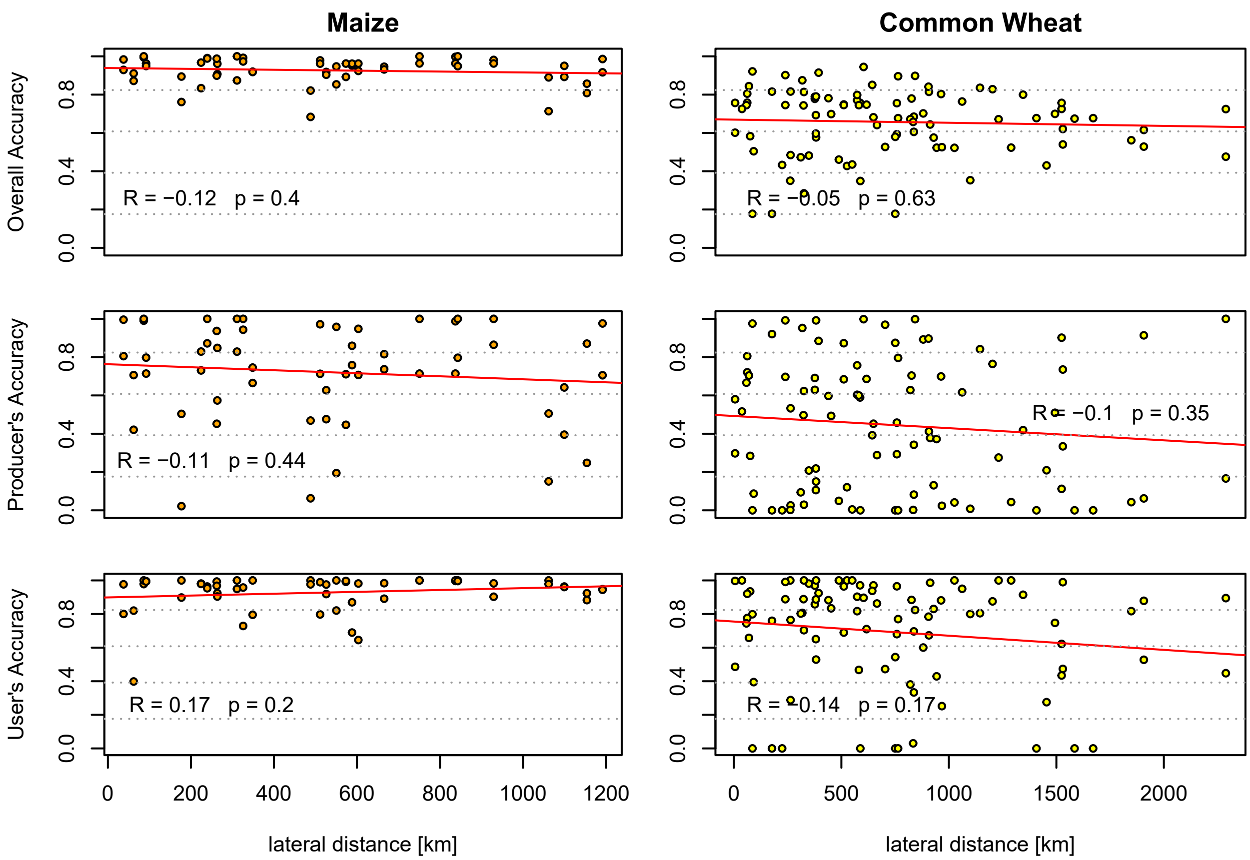

Figure 19: Correlation between lateral distance of training and classification region and reached accuracies

Table 4: Altitude of study regions with minimum, maximum and average elevation height

Region	MIN	MAX	MEAN
PZha-1	-2	154	33.99197
LVha-2	-4	164	41.55583
CMeu-2	-2	166	36.90239
CMca-1	-24	181	35.58366
CMeu-4	0	229	74.62006
LVha-3	17	280	83.45947
CMdy-1	-5	315	105.62546
PZha-2	0	318	88.87766
CMeu-3	2	331	124.35956
CMdy-3	0	332	59.80754
LVha-1	-47	356	113.27975
CMeu-1	-5	951	43.69129
CMdy-2	126	951	481.69786
CMca-2	513	1082	814.38236

Subsequently, the reached classification accuracies were tested for correlation with altitude of training (figure 21) resp. classification region (figure 22) and the height difference between these regions (figure 20).

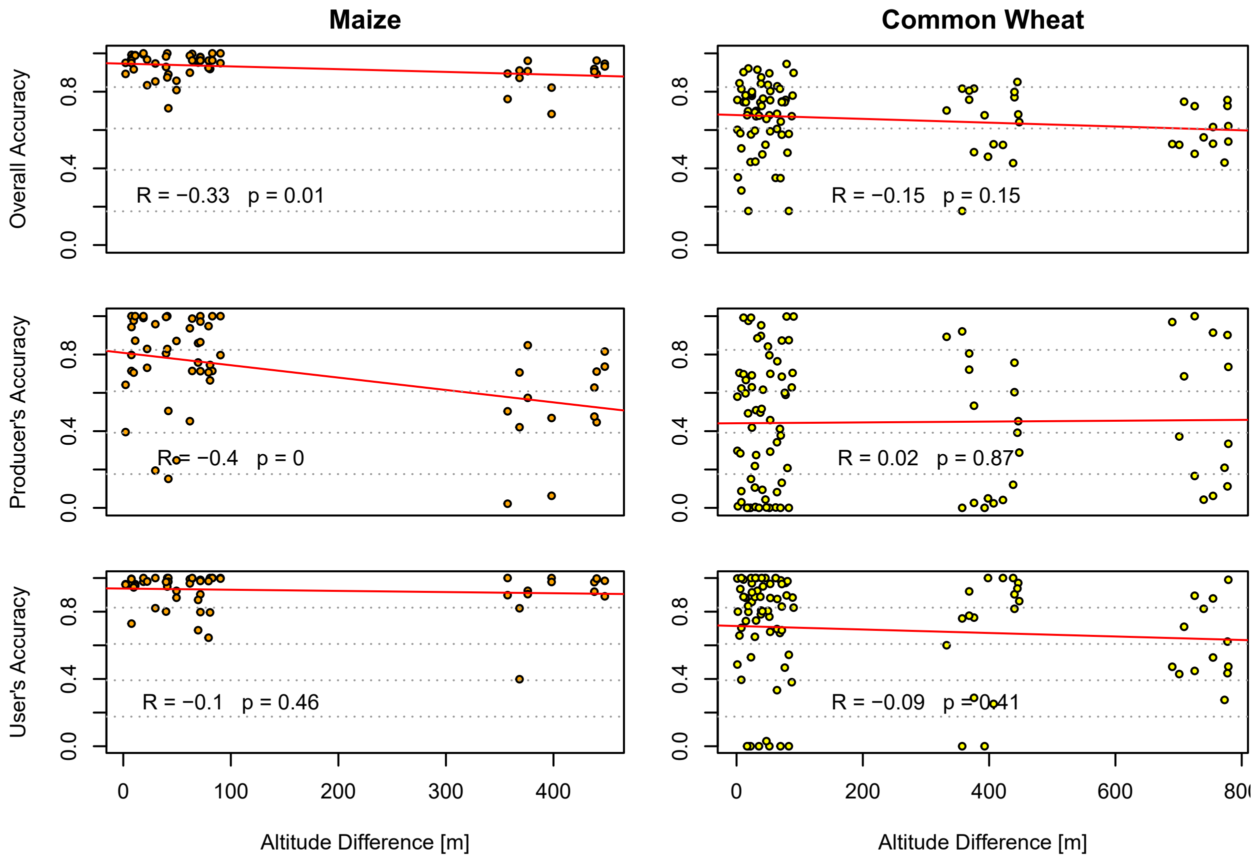

Figure 20: Correlation between altitude difference of training and classification region and reached accuracies

Correlation tests show Pearson's R-values ranging from -0.4 to 0.02, with corresponding p-values between 0.00 and 0.87.

For 'Maize' classifications, a weak to medium negative correlation was detected, particularly notable for Producer's Accuracy (PA) and Overall Accuracy (OA) at confidence intervals of 100% and 99%, respectively. However, this correlation is not significant when comparing User's Accuracy (UA) and altitude difference.

In the case of 'Common Wheat' classifications, there is a weak and mostly negative correlation between reached accuracies and altitude difference. These correlations are not statistically significant, with p-values ranging from 0.15 to 0.87.

In summary, the results show a significant weak to medium negative correlation between altitude difference and PA or OA for 'Maize' classifications, while 'Common Wheat' classifications did not show any significant correlation with altitude difference. (Figure 20).

Altitude Training Region

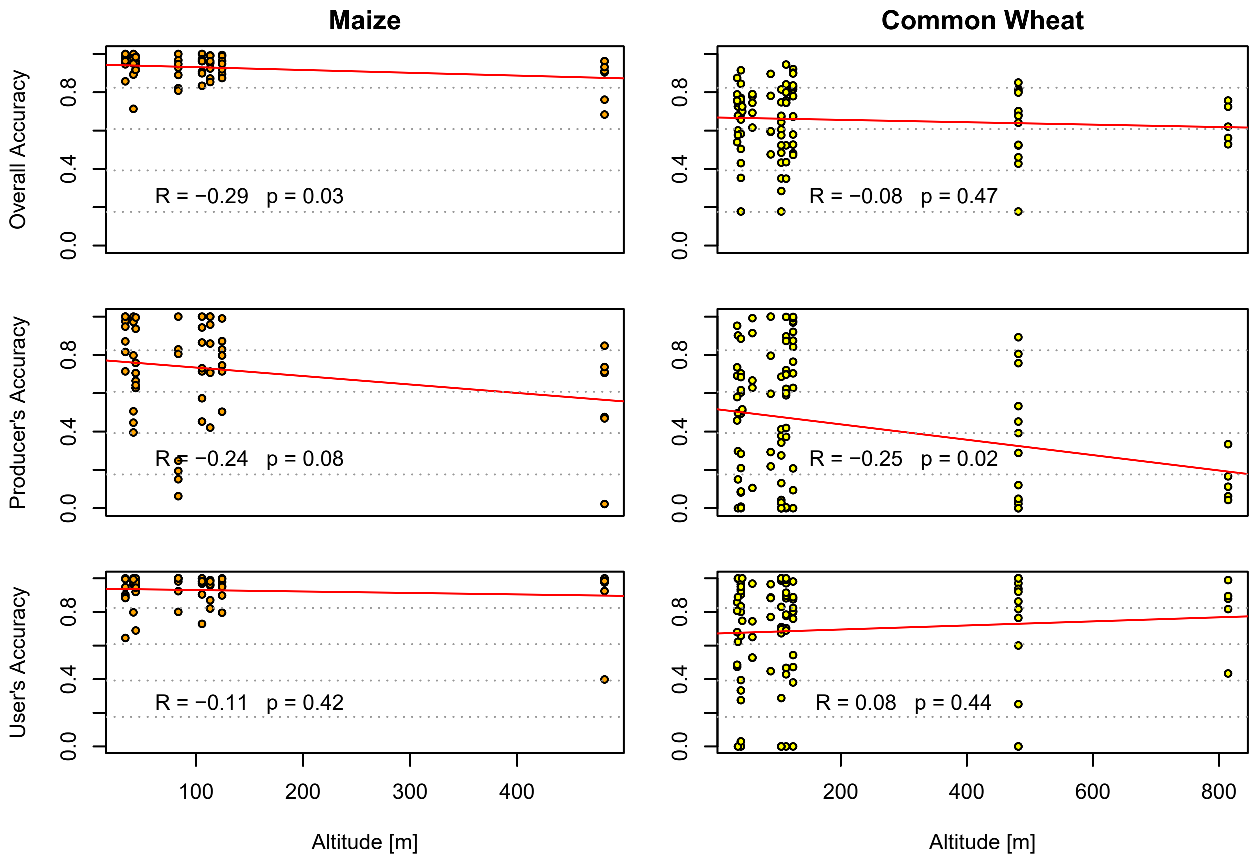

Figure 21: Correlation between altitude of training regions and reached accuracies

The correlation tests yielded Pearson's R-values ranging from -0.29 to 0.08, accompanied by p-values between 0.02 and 0.47.

For 'Maize' classifications, a medium negative correlation can be observed, particularly noteworthy for Producer's Accuracy (PA) and Overall Accuracy (OA), with significance levels of 92% and 97%, respectively. However, this correlation is not significant when comparing User's Accuracy (UA) and the altitude of the training region.

'Common Wheat' classifications show a medium negative correlation between reached PAs and the altitude of training regions, which is significant at a 92% confidence level. Overall Accuracy (OA) and User's Accuracy (UA) do not exhibit any significant correlation, with p-values of 0.47 and 0.44, respectively.

In summary, these findings reveal a significant negative correlation between the altitude of training regions and PA resp. OA for 'Maize' classifications, as well as a significant negative correlation between the altitude of training regions and reached PAs for 'Common Wheat' classifications. However, there was no significant correlation detected between the altitude of training regions and UA for both 'Maize' and 'Common Wheat' classifications, as well as OA for 'Common Wheat' classifications (Figure 21).

Altitude Classification Region

Correlation tests show Pearson's R-values between -0.26 and 0.29 and corresponding p-values between 0.01 and 0.54.

'Maize' classifications show a weak to medium negative correlation between the altitude of classification regions and reached accuracies. This correlation is significant for Producer's Accuracy (PA) within a 95% confidence interval, but it does not reach significance for User's Accuracies (UAs) resp. Overall Accuracies (OAs).

In the case of 'Common Wheat' classifications a medium positive correlation between reached PAs and altitude of classified regions was detected, as well as weak to medium negative correlations between UAs and altitude of the classification regions. These correlations are significant at confidence levels of 99% and 97%, respectively. However, there was no significant correlation with OAs (p-value of 0.24).

To summarize, the results indicate a significant negative correlation between the altitude of classification regions and PAs for 'Maize' classifications, as well as a significant positive correlation between altitude and reached PAs for 'Common Wheat' classifications. No significant correlations were observed between the altitude of classification regions and UAs resp. OAs for 'Maize,' and OAs for 'Common Wheat' classifications (Figure 22).

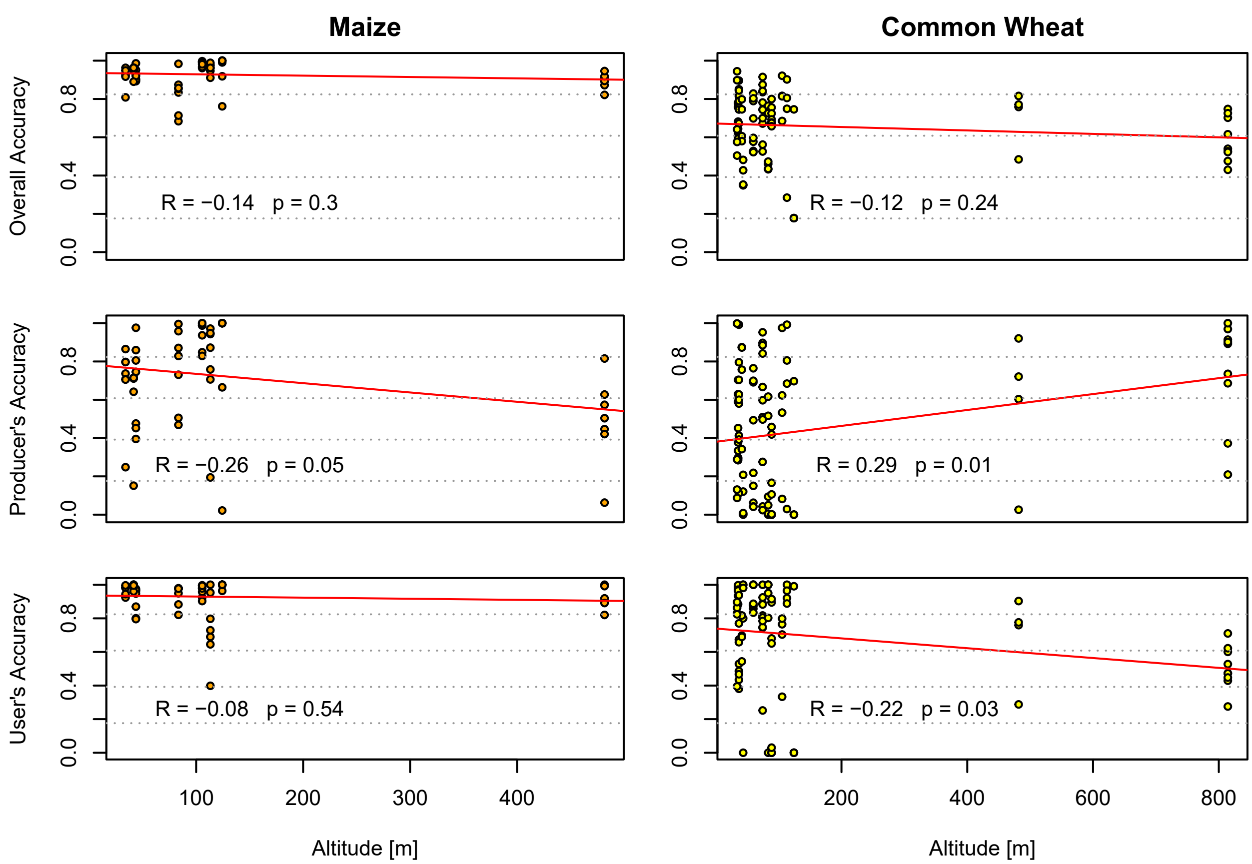

Figure 22: Correlation between altitude of classification region and reached accuracies

Discussion

Classifications which yield high OAs (> 0.8) but only medium to low UAs and PAs (< 0.8) for the target class ('Maize' resp. 'Common Wheat') are not considered successful. This might be the case when class 'Other' shows high accuracies but the target class itself is not reliably identified. This approach ensures that only classifications which were able to correctly identify a high proportion of the target class are counted as success.

The findings of this experiments show that there is no clear answer to the question if a RF classifier can be spatially transferred to classify crop types of a different region.

The multi-class classification results indicate, that it strongly depends on the selected target crop type, if the classification of a remote region can be successful. Class 'Barley' was very poorly identified in most of the classifications, whereas classification of 'Common Wheat' showed mixed results and 'Maize' identification worked very well in many cases. To evaluate the identifiability of the target crop types 'Maize' and 'Common Wheat', a binary approach was followed after this discovery.

Binary classifications for crop type 'Maize' showed very promising results regarding OAs, whereas classifications of 'Common Wheat' resulted in a wider range of OAs from values <0.2 to high OAs of > 0.8.

Regarding User's and Producer's Accuracies for both crop types, UAs showed much better results than PAs. High UAs indicate that the as target crop classified pixels are very likely to belong to the corresponding class in the ground truth data set. Lower PAs on the other hand indicate that areas belonging to the specific target crop type in reality (i.e. in the ground truth data), were not reliably identified as such.

'Maize' classifications showed very good results (> 0.8) for UAs and medium results for the corresponding PAs. 45% of all binary 'Maize' classifications showedd high PA and UA simultaneously and can therefore be considered a success.

Even though half of all 'Common Wheat' classifications yielded in UAs > 0.8, a high number of Wheat fields (as presented in the ground truth data) was not correctly identified as such which lead to low PAs in a majority of the classifications.

Consulting the spectral profiles of 'Maize' and 'Common Wheat' and the corresponding class 'Other', shows higher similarities for 'Common Wheat' and 'Other'. Ergo distinction between those classes is less reliable. This might be one of the main reasons for the high differences of reached accuracies for the target crop types.

Methodical influence on results

As shown in Figure 11, unequally distributed training pixel counts between the classes did not influence the results of the multi-class classifications.

The subsequently executed binary classifications of maize, resp. the post-processing calculated accuracies for binary wheat classifications came along with a high imbalance due to the merging of all non-target classes pixels to one class 'Other'. Evaluation of corresponding correlations in the binary classifications showed different results for 'Maize' and 'Common Wheat'. There is a highly significant negative correlation between class imbalance and reached OAs resp. PAs for 'Common Wheat' classifications. On the other hand, 'Maize' classifications showed a significant positive correlation between imbalance and PA of class 'Maize'. Correlation test results of 'Maize' classifications indicate, that a higher imbalance and more reliable identification (less false negatives) of the corresponding class are correlated. This finding not only contradicts 'Common Wheat' results but also what one would expect. Consulting figure 14 reveals that only a small number of data points with low PA and low imbalance accounted for the positive correlation of PA and Imbalance Index in 'Maize' classifications. Possibly but not conclusively these data points depict outliers that obscure the results. To further investigate the connection between imbalance and reached accuracies, more classifications with widely varying class imbalance should be evaluated.

As shown in figure 15, there is no correlation between the absolute number of training pixels, and reached accuracies which is in accordance with the findings of Rodriguez-Galiano et al. (2012). However, results showed that there is a weak but significant positive correlation between the ratio of test to training pixels and reached UAs of class 'Common Wheat'. This indicates that a higher potential intra-class variance in the test data due to a proportionally high number of test pixels and/or lower potential variance in the training data due to proportionally less training pixels yield in better classification results. This finding is contrary to the expected potential relation.

UAs of 'Maize' classifications and OAs and PAs for 'Maize' and 'Common Wheat' did not show this unexpected relation or at least with p-values of 0.17 to 0.66 no significance of the correlation.

Figure 15 shows that the majority of classifications have a test/training ratio between >0 and 4. Only a few data points show a higher ratio up to 8 (for wheat classification). Thus the significant positive correlation between UAs of wheat and the discussed ratio could be reducible to an insufficient amount of data at the higher end of the scale. A more regular distributed dataset in regard to test/training ratio could give more insights in a possible correlation.

Another measure for the potential variance in the used training data is the number of training sites. Even if there is a high number of training pixels, the inherent variance is potentially lower as they originate from a low number of sites compared to a higher number of training sites. This is based on the assumption that pixels from the same site show a lower spectral variance than pixels of the same class from differing training sites. Figure 17 shows that even classifications based on a very low number of training sites (min. = 2) did not produce significantly poorer results than classifications with a higher or even very high number of training sites.

Region specific influence on results

According to this study, equality of soil types in training and classification region does not increase the capability of classifiers to identify the target crop types. However, soil types according to the World Reference Base for Soil Resources (WRB) (FAO 1998) do not explicitly reflect metric soil characteristics like water availability or soil nutrition which are important factors according to Piedallu et al. (2019). Correlating classification results with individual metric parameters could get a better insight in possible correlations regarding the influence of differences or similarities of soil properties between training and classification region on the one hand and reached classification accuracies on the other hand. As this study aims to provide insights on the potential of classifying crop types in regions without available training data from the corresponding site, using publicly available data, this approach however was not considered due to a lack of comprehensive data. It must be mentioned that proper statistical analysis of correlation of nominal soil type classes and reached accuracies requires a sufficient number of regions per soil type. Due to the exclusion of unreliable data and subsequent filtering during data processing, this criterion can not be met in this study.

Also the associated very high number of classifications resulting from combinations of a higher number of training and classification regions exceeds the means of this thesis.

The results indicate a weak negative correlation between total distance of training and classification region and reached User's Accuracies for classifications of 'Common Wheat'. With increasing total distance, the proportion of falsely as wheat classified pixels increased significantly. This finding however is based on a data set with only a few data points showing distances > 2,000 km.

'Maize' classifications did not show this correlation whereby the maximum distance between "Maize regions" lies at about 1,260 km and the maximum distance between "Wheat regions" lies at about 3,015 km. Consulting Figure 18 shows that in the distance range up to about 2,000 km, also Wheat UAs did not show a significant decrease with increasing distance. Based on this finding, for better understanding the potential correlation, further studies need to be carried out, including more regions with distances > 2,000 km. These could on the one hand bring clarification if the observed correlation for 'Common Wheat' UAs is based on coincidentally low UAs for high distances, rooting it too less data points and on the other hand, if the lack of correlation for 'Maize' classifications is based on a lack of data points for higher distances.

Evaluating the relation between lateral distance and reached accuracies showed no significant correlations. Same as for total distances, the number of data points with higher lateral distances is comparably low. Showing no significant correlation however indicates that the total distance has a higher influence on the classification quality than lateral distance. Similar to the parameter "soil type" which is used as proxy for a variety of not directly available soil parameters, the parameter "lateral distance" serves as proxy for climate conditions. As climate is not only influenced by latitude of a region, this proxy might not sufficiently display differences in prevailing climate conditions in the individual regions. Better understanding could be produced by including additional metric parameters like differences of air temperature and precipitation prevailing in training and classification region. Using climate zones as nominal parameters (analogous to soil types) could also increase the understanding of possible coherencies. However using nominal classes limits the deployment of statistical analysis. To statistically analyse correlation between (in-)equality of climate zones in training and classification region and reached classification accuracies, a sufficient number of regions located in equal resp. unequal climate zones is required and thereby bears a disproportionate effort.

'Maize' classifications showed some training resp. classification regions outperforming others by consistently achieving high PAs and UAs simultaneously. This indicates that region specific conditions - despite the lack of previously discovered correlations - are likely to have an influence on the 'Maize' classification results. On the other hand, no outstandingly well performing training resp. classification regions could be identified in 'Common Wheat' classifications.

Conclusion

This study provides initial insights into the potential utilization of remotely acquired ground truth data for crop type classification in a specific region. Notably, 'Maize' classifications exhibited excellent overall accuracies (OA) and demonstrated promising results in terms of user accuracy (UA) and producer accuracy (PA) in many cases. In contrast, 'Common Wheat' classifications yielded mixed results, while 'Barley' classifications consistently failed to achieve high accuracies.

No conclusive correlations between methodic parameters and reached classification accuracies could be identified. Statistically significant results for correlation of high imbalance with low accuracies turned out partly contradicting each other. Also none of the evaluated region parameters “soil type”, “altitude”, “total distance” and “lateral distance” shows a clear influence on classification results, even though some regions perform distinctively better than others as training resp. classification region for class ‘Maize’.

The finding that some regions perform noticeably better than others for ‘Maize’ classification supports the impression that region specific conditions which have not been discovered in this study might have an substantial influence on the performance as training resp. classification region. The differences between training and classification region however seem to play no considerable role as well performing regions serve as potent training regions even for very different classification regions in terms of soil type and spatial resp. elevation differences. Even though there are no conclusive results regarding importance of class balance, some results indicate that using a well balanced training set can lead to better classification results. Differences in class balance between the regions could partly contribute to the differently well performing regions.

Due to the initially planned multi-class approach and the accordingly tailored ground truths data set, some experimental conditions like the coverage of higher distances, altitude differences and soil types did not ideally fulfill the requirements. Also the objective to use comprehensively available parameters to test them for correlation with the achieved results might have lead to oversimplification and therefore possibly obscured existing correlations between classification results and climate resp. soil conditions.

Further studies should focus on more distinct, preferably metric soil and climate parameters to enable a clear statistical analysis of their influence. Additionally, covering higher spatial distances between the regions with a sufficient number of classifications is necessary to evaluate, if the partly detected and sometimes contradicting correlations are caused by an insufficient dataset, lacking a sufficiently high number of higher distanced regions. Additionally other crop types could be evaluated for their suitability for this approach as it is clearly shown, that the identifiability of crop types using the depicted approach, varies strongly between different crop types.

Studying the correlations between class imbalance and reached accuracies also can be of interest to increase understanding on this topic.

The primary findings of this study can be summarized as follows:

a) For some crop types (e.g. ‘Maize’) a lack of training data from the classification region might be overcome by using data from remotely located regions. However the applicability of this approach seems mainly to depend on spectral characteristics of the crop types themselves and is not generally suitable for all crop types.

b) Some regions perform very well as training location in a wide range of region combinations. This indicates that it might be possible to use these to create sets of representative training data for specific suitable (as indicated in finding “a”) crop types which can be successfully utilized for classifications independent of the availability of on-site training data.

References

Bannari, Abdou, Daniel Morin, F Bonn, and AjRsr Huete. 1995. “A Review of Vegetation Indices.” *Remote Sensing Reviews* 13 (1-2): 95–120. https://doi.org/10.1080/02757259509532298.

Benos, Lefteris, Aristotelis C Tagarakis, Georgios Dolias, Remigio Berruto, Dimitrios Kateris, and Dionysis Bochtis. 2021. “Machine Learning in Agriculture: A Comprehensive Updated Review.” *Sensors* 21 (11): 3758. https://doi.org/10.3390/s21113758.

Blickensdörfer, Lukas, Marcel Schwieder, Dirk Pflugmacher, Claas Nendel, Stefan Erasmi, and Patrick Hostert. 2022. “Mapping of Crop Types and Crop Sequences with Combined Time Series of Sentinel-1, Sentinel-2 and Landsat 8 Data for Germany.” *Remote Sensing of Environment* 269: 112831. https://doi.org/10.1016/j.rse.2021.1128310.

Danielson, JJ, and DB Gesch. 2011. “Global Mid-Resolution Terrain Elevation Data 2010 (GMTED 2010).” *United States Geological Survey.*

ESA, European Space Agency. 2023. “Technical Guide - Sentinel-2 MultiSpectral Instrument (MSI) Overview.” https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument (access date: 2023-06-04).

European Commission, and the European Soil Bureau Network. 2004. “The European Soil Database Distribution Version 2.0.” *CD-ROM, EUR 19945 EN, 2004.* https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data#tabs-0-description=0.

European Union, European Environment Agency, Copernicus Land Monitoring Service. 2018. “Corine Land Cover - 100 Meter.” https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=download.

Eurostat. 2018a. “Land Use and Cover Area Frame Survey 2018,” https://ec.europa.eu/eurostat/web/lucas/data/primary-data/2018.

Eurostat. 2018b. “Technical Reference Document c-3: Classification,” https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C3-Classification.pdf.

Eurostat. 2018c. “LUCAS 2018 (Land Use / Cover Area Frame Survey) - Technical Reference Document C2 Field Form and Ground Document (Template),” https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C2-FieldForm-GD-Template.pdf.

FAO. 1998. “World Reference Base for Soil Resources, by ISSS-ISRIC-FAO.” *World Soil Resources Reports No. 84.*

Fowler, Jared, François Waldner, and Zvi Hochman. 2020. “All Pixels Are Useful, but Some Are More Useful: Efficient in Situ Data Collection for Crop-Type Mapping Using Sequential Exploration Methods.” *International Journal of Applied Earth Observation and Geoinformation* 91: 102114. https://doi.org/10.1016/j.jag.2020.102114.

Geiger, Rudolf. 1961. “Überarbeitete Neuausgabe von Geiger, r.” *Köppen-Geiger/Klima Der Erde.(Wandkarte 1: 16 Mill.)–Klett-Perthes, Gotha.*

Grira, Nizar, Michel Crucianu, and Nozha Boujemaa. 2004. “Unsupervised and Semi-Supervised Clustering: A Brief Survey.” *A Review of Machine Learning Techniques for Processing Multimedia Content* 1 (2004): 9–16.

Heupel, Katharina, Daniel Spengler, and Sibylle Itzerott. 2018. “A Progressive Crop-Type Classification Using Multitemporal Remote Sensing Data and Phenological Information.” *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 86: 53–69. https://doi.org/10.1007/s41064-018-0050-7.

Kells, Lyman Morse, Willis Frederick Kern, and James R Bland. 1940. *Spherical Trigonometry.* McGraw-Hill book Company, Incorporated.

Kinyanjui, Mwangi J. 2011. "NDVI-Based Vegetation Monitoring in Mau Forest Complex, Kenya." *African Journal of Ecology* 49 (2): 165–74. https://doi.org/10.1111/j.1365-2028.2010.01251.x.

Köppen, Wladimir. 1900. "Versuch Einer Klassifikation Der Klimate, Vorzugsweise Nach Ihren Beziehungen Zur Pflanzenwelt." *Geographische Zeitschrift* 6 (11. H): 593–611.

Ma, Zhe, Zhe Liu, Yuanyuan Zhao, Lin Zhang, Diyou Liu, Tianwei Ren, Xiaodong Zhang, and Shaoming Li. 2020. "An Unsupervised Crop Classification Method Based on Principal Components Isometric Binning." *ISPRS International Journal of Geo-Information* 9 (11): 648. https://doi.org/10.3390/ijgi9110648.

Mellor, Andrew, Samia Boukir, Andrew Haywood, and Simon Jones. 2015. "Exploring Issues of Training Data Imbalance and Mislabelling on Random Forest Performance for Large Area Land Cover Classification Using the Ensemble Margin." *ISPRS Journal of Photogrammetry and Remote Sensing* 105: 155–68. https://doi.org/10.1016/j.isprsjprs.2015.03.014.

Mueller, Lothar, Uwe Schindler, T Graham Shepherd, Bruce C Ball, Elena Smolentseva, Konstantin Pachikin, Chunsheng Hu, et al. 2014. "The Muencheberg Soil Quality Rating for Assessing the Quality of Global Farmland." *Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia*, 235–48. https://doi.org/10.1007/978-3-319-01017-5_13.

Orynbaikyzy, Aiym, Ursula Gessner, and Christopher Conrad. 2019. "Crop Type Classification Using a Combination of Optical and Radar Remote Sensing Data: A Review." *International Journal of Remote Sensing* 40 (17): 6553–95. https://doi.org/710.1080/01431161.2019.1569791.

Orynbaikyzy, Aiym, Ursula Gessner, and Christopher Conrad. 2022. "Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2." *Remote Sensing* 14 (6): 1493. https://doi.org/10.3390/rs14061493.

Pal, Mahesh, and Paul M Mather. 2003. "An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Classification." *Remote Sensing of Environment* 86 (4): 554–65. https://doi.org/10.1016/S0034-4257(03)00132-9.

Perumal, K, and R Bhaskaran. 2010. "Supervised Classification Performance of Multispectral Images." *CoRR* abs/1002.4046. https://doi.org/10.48550/arXiv.1002.4046.

Piedallu, Christian, Véronique Cheret, Jean-Philippe Denux, Vincent Perez, Jaime Sebastian Azcona, Ingrid Seynave, and Jean-Claude Gégout. 2019. "Soil and Climate Differently Impact NDVI Patterns According to the Season and the Stand Type." *Science of the Total Environment* 651: 2874–85. https://doi.org/10.1016/j.scitotenv.2018.10.052.

Pluto-Kossakowska, Joanna. 2021. "Review on Multitemporal Classification Methods of Satellite Images for Crop and Arable Land Recognition." *Agriculture* 11 (10): 999. https://doi.org/10.3390/agriculture11100999.

Prudnikova, Elena, Igor Savin, Gretelerika Vindeker, Praskovia Grubina, Ekaterina Shishkonakova, and David Sharychev. 2019. "Influence of Soil Background on Spectral Reflectance of Winter Wheat Crop Canopy." *Remote Sensing* 11 (16). https://doi.org/10.3390/rs11161932.

Ringrose, S., Wilma Matheson, C. Matlala, Toni O'Neill, and Patricia Werner. 1994. "Vegetation Spectral Reflectance Along a North-South Vegetation Gradient in Northern Australia." *Journal of Biogeography* 21 (January): 33. https://doi.org/10.2307/2845602.

Rodriguez-Galiano, Victor Francisco, Bardan Ghimire, John Rogan, Mario Chica-Olmo, and Juan Pedro Rigol-Sanchez. 2012. "An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93–104. https:

//doi.org/710.1016/j.isprsjprs.2011.11.002.

Rouse, JW. 1974. "Monitoring the Vernal Advancement of Retrogradation of Natural Vegetation." *NASA/GSFC, Type III, Final Report, Greenbelt, MD* 371.

Stevens, A. 2010. "Introduction to the Basic Drivers of Climate." *Nature Education Knowledge* 3. https://www.nature.com/scitable/knowledge/library/introduction-to-the-basic-drivers-of-climate-13368032/.

Tobler, Waldo R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46 (sup1): 234–40. https://doi.org/10.2307/143141.

Wang, Sherrie, George Azzari, and David B. Lobell. 2019. "Crop Type Mapping Without Field-Level Labels: Random Forest Transfer and Unsupervised Clustering Techniques." *Remote Sensing of Environment* 222: 303–17. https://doi.org/10.1016/j.rse.2018.12.026.

Zhang, Hongyan, Jinzhong Kang, Xiong Xu, and Liangpei Zhang. 2020. "Accessing the Temporal and Spectral Features in Crop Type Mapping Using Multi-Temporal Sentinel-2 Imagery: A Case Study of Yi'an County, Heilongjiang Province, China." *Computers and Electronics in Agriculture* 176: 105618. https://doi.org/10.1016/j.compag.2020.105618.

Appendix

Appendix A: Google Earth Engine Code

Appendix B: Generated and derived data from classifications

Appendix C: Classification accuracies multi-class classifications

Appendix D: Region combination matrices

Appendix E: Spectral profiles of ‘Maize’, ‘Common Wheat’ and ‘Other’

Appendix A-1:

Google Earth Engine Code ‘multi-class’

The following data sets are imported to the Google Earth Engine:

- var regions: regions.shp geometry of the individual study regions
- var geometry: geometry.shp used to pre-define the extent
- var training: train_data.shp training/test data

The following script was used for the multi-class classifications. The following script was used to classify pixels utilizing a classifier, trained in region "LVha-1". This script is representative for all multi-class classifications and can be modified by (un-)commenting to perform all other multi-class classifications analogously.

```
// Define the regions of interest (ROIs) as a feature collection
var roi = regions;

// Define a list of region names
var regionNames = ['LVha-1', 'LVha-2', 'CMeu-3', 'CMdy-1',  'CMdy-2', 'CMeu-1',
                   'CMca-1','CMca-2','CMdy-3', 'PZha-2','CMeu-2','PZha-1','CMeu-4',
                   'LVha-3'];

// add Sentinel-2 imageCollection and filter by date and geometry
var S2 = ee.ImageCollection('COPERNICUS/S2_SR')
  .filterDate('2018-04-01', '2018-10-31')
  .filterBounds(geometry);


//function to calculate NDVI
function addNDVI(image) {
  var ndvi = image.normalizedDifference(['B8','B4'])
  return image.addBands(ndvi.rename('NDVI'));
}

// map function over imageCollection, keeping only bandy NDVI 01-14 and SCL 01-14
var NDVI_S2 = S2.map(addNDVI).select(['NDVI', 'SCL']);



//subset imageCollection by date range and composite greenest pixel values
//(max NDVIs) to one image

var composite = NDVI_S2.filterDate('2018-04-01','2018-04-15').qualityMosaic('NDVI')
    .rename('01', 'CLD01')
  .addBands(NDVI_S2.filterDate('2018-04-16','2018-04-30').qualityMosaic('NDVI')
    .rename('02', 'CLD02'))
  .addBands(NDVI_S2.filterDate('2018-05-01','2018-05-15').qualityMosaic('NDVI')
    .rename('03', 'CLD03'))
  .addBands(NDVI_S2.filterDate('2018-05-16','2018-05-23').qualityMosaic('NDVI')
    .rename('04', 'CLD04'))
```

```
  .addBands(NDVI_S2.filterDate('2018-06-01','2018-06-15').qualityMosaic('NDVI')
    .rename('05', 'CLD05'))
  .addBands(NDVI_S2.filterDate('2018-06-16','2018-06-30').qualityMosaic('NDVI')
    .rename('06', 'CLD06'))
  .addBands(NDVI_S2.filterDate('2018-07-01','2018-07-15').qualityMosaic('NDVI')
    .rename('07', 'CLD07'))
  .addBands(NDVI_S2.filterDate('2018-07-16','2018-07-31').qualityMosaic('NDVI')
    .rename('08', 'CLD08'))
  .addBands(NDVI_S2.filterDate('2018-08-01','2018-08-15').qualityMosaic('NDVI')
    .rename('09', 'CLD09'))
  .addBands(NDVI_S2.filterDate('2018-08-16','2018-08-31').qualityMosaic('NDVI')
    .rename('10', 'CLD10'))
  .addBands(NDVI_S2.filterDate('2018-09-01','2018-09-15').qualityMosaic('NDVI')
    .rename('11', 'CLD11'))
  .addBands(NDVI_S2.filterDate('2018-09-16','2018-09-30').qualityMosaic('NDVI')
    .rename('12', 'CLD12'))
  .addBands(NDVI_S2.filterDate('2018-10-01','2018-10-15').qualityMosaic('NDVI')
    .rename('13', 'CLD13'))
  .addBands(NDVI_S2.filterDate('2018-10-16','2018-10-31').qualityMosaic('NDVI')
    .rename('14', 'CLD14'))

// clip 28-band composite to the individual study regions and assign a variable each
var LVha1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'LVha-1')))
   .set({region:"LVha-1"});
var LVha2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'LVha-2')))
   .set({region:"LVha-2"});
var CMeu3 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMeu-3')))
   .set({region:"CMeu-3"});
var CMdy1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMdy-1')))
   .set({region:"CMdy-1"});
var CMdy2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMdy-2')))
   .set({region:"CMdy-2"});
var CMeu1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMeu-1')))
   .set({region:"CMeu-1"});
var CMca1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMca-1')))
   .set({region:"CMca-1"});
var CMca2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMca-2')))
   .set({region:"CMca-2"});
var CMdy3 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMdy-3')))
   .set({region:"CMdy-3"});
var PZha2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'PZha-2')))
   .set({region:"PZha-2"});
var CMeu2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMeu-2')))
```

```
   .set({region:"CMeu-2"});
var PZha1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'PZha-1')))
   .set({region:"PZha-1"});
var CMeu4 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMeu-4')))
   .set({region:"CMeu-4"});
var LVha3 = composite.clip(roi.filter(ee.Filter.eq('Name', 'LVha-3')))
   .set({region:"LVha-3"});


//======================================== classifications
// define variables for classifier
var label = "class";
var class_bands = ['01','02','03','04','05','06','07','08','09', '10',
                   '11','12','13','14']

// Training data

// create list of all image subsets (individual regions)
var imageList = ee.List([LVha1,LVha2,CMeu3,CMdy1,CMdy2,CMeu1,CMca1,CMca2,
                         CMdy3,PZha2,CMeu2,PZha1,CMeu4,LVha3]);

// function to access training data from featureCollection for each region
// separately
function pixels(img){
  return img.sampleRegions({
    collection: training.filter(ee.Filter.eq("region",img.get("region"))),
    properties:["class", "region"],
    scale:10
  })
}

// Mapping the function to all "greenest pixel" images
// (Source: Oliver Lopez Stackexchange)
var N = imageList.size();
var s = ee.List.sequence(0,N.subtract(1));

var results = ee.FeatureCollection(s.map(function(n){
 return pixels(ee.Image(imageList.get(n)));
}
)).flatten();



// Apply cloud-filter to FeatureCollection, masking out cloudy features and
```

```
// no-data features
var featureList = results.filter(ee.Filter.and(ee.Filter.neq('CLD01', 0),
                                                ee.Filter.neq('CLD02', 0),
                                                ee.Filter.neq('CLD03', 0),
                                                ee.Filter.neq('CLD04', 0),
                                                ee.Filter.neq('CLD05', 0),
                                                ee.Filter.neq('CLD06', 0),
                                                ee.Filter.neq('CLD07', 0),
                                                ee.Filter.neq('CLD08', 0),
                                                ee.Filter.neq('CLD09', 0),
                                                ee.Filter.neq('CLD10', 0),
                                                ee.Filter.neq('CLD11', 0),
                                                ee.Filter.neq('CLD12', 0),
                                                ee.Filter.neq('CLD13', 0),
                                                ee.Filter.neq('CLD14', 0),
                                                ee.Filter.neq('CLD01', 8),
                                                ee.Filter.neq('CLD02', 8),
                                                ee.Filter.neq('CLD03', 8),
                                                ee.Filter.neq('CLD04', 8),
                                                ee.Filter.neq('CLD05', 8),
                                                ee.Filter.neq('CLD06', 8),
                                                ee.Filter.neq('CLD07', 8),
                                                ee.Filter.neq('CLD08', 8),
                                                ee.Filter.neq('CLD09', 8),
                                                ee.Filter.neq('CLD10', 8),
                                                ee.Filter.neq('CLD11', 8),
                                                ee.Filter.neq('CLD12', 8),
                                                ee.Filter.neq('CLD13', 8),
                                                ee.Filter.neq('CLD14', 8),
                                                ee.Filter.neq('CLD01', 9),
                                                ee.Filter.neq('CLD02', 9),
                                                ee.Filter.neq('CLD03', 9),
                                                ee.Filter.neq('CLD04', 9),
                                                ee.Filter.neq('CLD05', 9),
                                                ee.Filter.neq('CLD06', 9),
                                                ee.Filter.neq('CLD07', 9),
                                                ee.Filter.neq('CLD08', 9),
                                                ee.Filter.neq('CLD09', 9),
                                                ee.Filter.neq('CLD10', 9),
                                                ee.Filter.neq('CLD11', 9),
                                                ee.Filter.neq('CLD12', 9),
                                                ee.Filter.neq('CLD13', 9),
                                                ee.Filter.neq('CLD14', 9)));
```

```
// function to assign training data
function trainingData(regionName) {
  var region_assigned = featureList.filterMetadata("region", "equals", regionName)
  return region_assigned;
}

// execute function (here for region LVha)
var LVha1_train = trainingData('LVha-1')
//var LVha2_train = trainingData('LVha-2')
//var CMeu3_train = trainingData('CMeu-3')
//var CMdy1_train = trainingData('CMdy-1')
//var CMdy2_train = trainingData('CMdy-2')
//var CMeu1_train = trainingData('CMeu-1')
//var CMca1_train = trainingData('CMca-1')
//var CMca2_train = trainingData('CMca-2')
//var CMdy3_train = trainingData('CMdy-3')
//var CMeu5_train = trainingData('CMeu-5')
//var PZha2_train = trainingData('PZha-2')
//var PZha3_train = trainingData('PZha-3')
//var CMeu2_train = trainingData('CMeu-2')
//var PZha1_train = trainingData('PZha-1')
//var CMeu4_train = trainingData('CMeu-4')
//var CMca3_train = trainingData('CMca-3')
//var LVha3_train = trainingData('LVha-3')


// assigning test data for each except the training region (here: LVha1)
//var LVha1_test = trainingData('LVha-1')
var LVha2_test = trainingData('LVha-2')
var CMeu3_test = trainingData('CMeu-3')
var CMdy1_test = trainingData('CMdy-1')
var CMdy2_test = trainingData('CMdy-2')
var CMeu1_test = trainingData('CMeu-1')
var CMca1_test = trainingData('CMca-1')
var CMca2_test = trainingData('CMca-2')
var CMdy3_test = trainingData('CMdy-3')
var PZha2_test = trainingData('PZha-2')
var CMeu2_test = trainingData('CMeu-2')
var PZha1_test = trainingData('PZha-1')
var CMeu4_test = trainingData('CMeu-4')
var LVha3_test = trainingData('LVha-3')
```

```
// create list of test data (without training region LVha1)
var testList = LVha2_test
//  .merge(LVha1_test)
    .merge(CMeu3_test)
    .merge(CMdy1_test)
    .merge(CMdy2_test)
    .merge(CMeu1_test)

//  .merge(CMca1_test)
//  .merge(CMca2_test)
//  .merge(CMdy3_test)
//  .merge(PZha2_test)

//  .merge(CMeu2_test)
//  .merge(PZha1_test)
//  .merge(CMeu4_test)
//  .merge(LVha3_test);


// Make a Random Forest classifier and train it (here on LVha1 training data).
ar classifier = ee.Classifier.smileRandomForest(100)
    .train({
      features: LVha1_train,
      classProperty: label,
      inputProperties: class_bands
    });


// classify the images of all regions except training region (here: LVha1)
// this step is split up (commented) to keep the calculation cost per run manageable

//var classified_2to01 = LVha1.select(class_bands).classify(classifier);
var classified_2to02 = LVha2.select(class_bands).classify(classifier);
var classified_2to03 = CMeu3.select(class_bands).classify(classifier);
var classified_2to04 = CMdy1.select(class_bands).classify(classifier);
var classified_2to05 = CMdy2.select(class_bands).classify(classifier);
var classified_2to06 = CMeu1.select(class_bands).classify(classifier);

//var classified_2to07 = CMca1.select(class_bands).classify(classifier);
//var classified_2to08 = CMca2.select(class_bands).classify(classifier);
//var classified_2to09 = CMdy3.select(class_bands).classify(classifier);
```

```
//var classified_2to11 = PZha2.select(class_bands).classify(classifier);

//var classified_1to13 = CMeu2.select(class_bands).classify(classifier);
//var classified_1to14 = PZha1.select(class_bands).classify(classifier);
//var classified_1to15 = CMeu4.select(class_bands).classify(classifier);
//var classified_1to17 = LVha3.select(class_bands).classify(classifier);

// Accuracy Assessment
function confusionMatrix(regionName) {
  var testdat_region = testList.filterMetadata("region", "equals", regionName)
  return ee.ConfusionMatrix(testdat_region.classify(classifier).errorMatrix({
  actual: "class",
  predicted: "classification"
}));
}
//
//var OA_2to01 = ee.Feature(confusionMatrix('LVha-1').accuracy());
var OA_2to02 = ee.Feature(confusionMatrix('LVha-2').accuracy());
var OA_2to03 = ee.Feature(confusionMatrix('CMeu-3').accuracy());
var OA_2to04 = ee.Feature(confusionMatrix('CMdy-1').accuracy());
var OA_2to05 = ee.Feature(confusionMatrix('CMdy-2').accuracy());
var OA_2to06 = ee.Feature(confusionMatrix('CMeu-1').accuracy());

//var OA_2to07 = ee.Feature(confusionMatrix('CMca-1').accuracy());
//var OA_2to08 = ee.Feature(confusionMatrix('CMca-2').accuracy());
//var OA_2to09 = ee.Feature(confusionMatrix('CMdy-3').accuracy());
//var OA_2to11 = ee.Feature(confusionMatrix('PZha-2').accuracy());

//var OA_2to13 = ee.Feature(confusionMatrix('CMeu-2').accuracy());
//var OA_2to14 = ee.Feature(confusionMatrix('PZha-1').accuracy());
//var OA_2to15 = ee.Feature(confusionMatrix('CMeu-4').accuracy());
//var OA_2to17 = ee.Feature(confusionMatrix('LVha-3').accuracy());

// print Overall Accuracy (OA)
//print(OA_2to01)
print(OA_2to02)
print(OA_2to03)
print(OA_2to04)
print(OA_2to05)
print(OA_2to06)

//print(OA_2to07)
//print(OA_2to08)
```

```
//print(OA_2to09)
//print(OA_2to11)

//print(OA_2to13)
//print(OA_2to14)
//print(OA_2to15)
//print(OA_2to17)


// print training data
print(LVha1_train, 'LVha1 train'

// print training data separated by class
print(LVha1_train.filter(ee.Filter.eq('class', 0)), 'LVha-1 train, class 0')
print(LVha1_train.filter(ee.Filter.eq('class', 1)), 'LVha-1 train, class 1')
print(LVha1_train.filter(ee.Filter.eq('class', 2)), 'LVha-1 train, class 2')
print(LVha1_train.filter(ee.Filter.eq('class', 3)), 'LVha-1 train, class 3')

// print classification info
print(classifier.explain());

// print confuson matrices
//print(confusionMatrix('LVha-1'))
print(confusionMatrix('LVha-2'))
print(confusionMatrix('CMeu-3'))
print(confusionMatrix('CMdy-1'))
print(confusionMatrix('CMdy-2'))
print(confusionMatrix('CMeu-1'))

// map classification results (representatively for the first 5 classifications)
var palette = [
  'grey',
  'yellow',
  'brown',
  'orange'
];

Map.addLayer(classified_2to02, {
  min: 0,
  max: palette.length - 1,
  palette: palette
})
Map.addLayer(classified_2to03,{
```

```
  min: 0,
  max: palette.length - 1,
  palette: palette
})
Map.addLayer(classified_2to04,{
  min: 0,
  max: palette.length - 1,
  palette: palette
})
Map.addLayer(classified_2to05,{
  min: 0,
  max: palette.length - 1,
  palette: palette
})
Map.addLayer(classified_2to06,{
  min: 0,
  max: palette.length - 1,
  palette: palette
})
```

Appendix A-2:

Google Earth Engine Code ‘Maize binary’

‘Maize’ classifications were calculated, using the same input data as the multi-class classfications. The following script was used to identify ‘Maize’ pixels utilizing a classifier, trained in region “LVha-1”. This script is representative for all ‘Maize’ classifications and can be modified by (un-)commenting to perform all other ‘Maize’ classifications analogously.

```
// Define the regions of interest (ROIs) as a feature collection
var roi = regions;

// Define a list of region names
var regionNames = ['LVha-1', 'LVha-2', 'CMeu-3', 'CMdy-1',  'CMdy-2', 'CMeu-1','PZha-1',
                   'LVha-3'];

// add Sentinel-2 imageCollection and filter by date and geometry
var S2 = ee.ImageCollection('COPERNICUS/S2_SR')
  .filterDate('2018-04-01', '2018-10-31')
  .filterBounds(geometry);

//function to calculate NDVI
function addNDVI(image) {
  var ndvi = image.normalizedDifference(['B8','B4'])
  return image.addBands(ndvi.rename('NDVI'));
}

// map function over imageCollection, keeping only bands NDVI 01-14 and SCL 01-14
var NDVI_S2 = S2.map(addNDVI).select(['NDVI', 'SCL']);


//subset imageCollection by date range and composite greenest pixel values
//(max NDVIs) to one image

var composite = NDVI_S2.filterDate('2018-04-01','2018-04-15').qualityMosaic('NDVI')
      .rename('01', 'CLD01')
  .addBands(NDVI_S2.filterDate('2018-04-16','2018-04-30').qualityMosaic('NDVI')
      .rename('02', 'CLD02'))
  .addBands(NDVI_S2.filterDate('2018-05-01','2018-05-15').qualityMosaic('NDVI')
      .rename('03', 'CLD03'))
  .addBands(NDVI_S2.filterDate('2018-05-16','2018-05-23').qualityMosaic('NDVI')
      .rename('04', 'CLD04'))
  .addBands(NDVI_S2.filterDate('2018-06-01','2018-06-15').qualityMosaic('NDVI')
      .rename('05', 'CLD05'))
  .addBands(NDVI_S2.filterDate('2018-06-16','2018-06-30').qualityMosaic('NDVI')
      .rename('06', 'CLD06'))
  .addBands(NDVI_S2.filterDate('2018-07-01','2018-07-15').qualityMosaic('NDVI')
      .rename('07', 'CLD07'))
```

```
  .addBands(NDVI_S2.filterDate('2018-07-16','2018-07-31').qualityMosaic('NDVI')
      .rename('08', 'CLD08'))
  .addBands(NDVI_S2.filterDate('2018-08-01','2018-08-15').qualityMosaic('NDVI')
      .rename('09', 'CLD09'))
  .addBands(NDVI_S2.filterDate('2018-08-16','2018-08-31').qualityMosaic('NDVI')
      .rename('10', 'CLD10'))
  .addBands(NDVI_S2.filterDate('2018-09-01','2018-09-15').qualityMosaic('NDVI')
      .rename('11', 'CLD11'))
  .addBands(NDVI_S2.filterDate('2018-09-16','2018-09-30').qualityMosaic('NDVI')
      .rename('12', 'CLD12'))
  .addBands(NDVI_S2.filterDate('2018-10-01','2018-10-15').qualityMosaic('NDVI')
      .rename('13', 'CLD13'))
  .addBands(NDVI_S2.filterDate('2018-10-16','2018-10-31').qualityMosaic('NDVI')
      .rename('14', 'CLD14'));


// clip 28-band composite to the individual study regions and assign a variable each
var LVha1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'LVha-1')))
  .set({region:'LVha-1'});
var LVha2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'LVha-2')))
  .set({region:'LVha-2'});
var CMeu3 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMeu-3')))
  .set({region:'CMeu-3'});
var CMdy1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMdy-1')))
  .set({region:'CMdy-1'});
var CMdy2 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMdy-2')))
  .set({region:'CMdy-2'});
var CMeu1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'CMeu-1')))
  .set({region:'CMeu-1'});
var PZha1 = composite.clip(roi.filter(ee.Filter.eq('Name', 'PZha-1')))
  .set({region:'PZha-1'});
var LVha3 = composite.clip(roi.filter(ee.Filter.eq('Name', 'LVha-3')))
  .set({region:'LVha-3'});


//======================================= classifications

// define variables for classifier
var label = "class_maiz";
var class_bands = ['01','02','03','04','05','06','07','08','09', '10', '11','12',
                   '13','14']

// Training data
```

```
// create list of all image subsets (individual regions)
var imageList = ee.List([LVha1,LVha2,CMeu3,CMdy1,CMdy2,CMeu1,PZha1,LVha3]);


// function to access training data from featureCollection for each region
// separately
function pixels(img){
  return img.sampleRegions({
    collection: training.filter(ee.Filter.eq("region",img.get("region"))),
    properties:["class_maiz", "region"],
    scale:10
  })
}

// Mapping the function to all "greenest pixel" images
// (Source: Oliver Lopez Stackexchange)
var N = imageList.size();
var s = ee.List.sequence(0,N.subtract(1));

var results = ee.FeatureCollection(s.map(function(n){
 return pixels(ee.Image(imageList.get(n)));
}
)).flatten();

// Apply cloud-filter to FeatureCollection, masking out cloudy features and
// no-data features
var featureList = results.filter(ee.Filter.and(ee.Filter.neq('CLD01', 0),
                                                ee.Filter.neq('CLD02', 0),
                                                ee.Filter.neq('CLD03', 0),
                                                ee.Filter.neq('CLD04', 0),
                                                ee.Filter.neq('CLD05', 0),
                                                ee.Filter.neq('CLD06', 0),
                                                ee.Filter.neq('CLD07', 0),
                                                ee.Filter.neq('CLD08', 0),
                                                ee.Filter.neq('CLD09', 0),
                                                ee.Filter.neq('CLD10', 0),
                                                ee.Filter.neq('CLD11', 0),
                                                ee.Filter.neq('CLD12', 0),
                                                ee.Filter.neq('CLD13', 0),
                                                ee.Filter.neq('CLD14', 0),
                                                ee.Filter.neq('CLD01', 8),
                                                ee.Filter.neq('CLD02', 8),
```

```
                                                  ee.Filter.neq('CLD03', 8),
                                                  ee.Filter.neq('CLD04', 8),
                                                  ee.Filter.neq('CLD05', 8),
                                                  ee.Filter.neq('CLD06', 8),
                                                  ee.Filter.neq('CLD07', 8),
                                                  ee.Filter.neq('CLD08', 8),
                                                  ee.Filter.neq('CLD09', 8),
                                                  ee.Filter.neq('CLD10', 8),
                                                  ee.Filter.neq('CLD11', 8),
                                                  ee.Filter.neq('CLD12', 8),
                                                  ee.Filter.neq('CLD13', 8),
                                                  ee.Filter.neq('CLD14', 8),
                                                  ee.Filter.neq('CLD01', 9),
                                                  ee.Filter.neq('CLD02', 9),
                                                  ee.Filter.neq('CLD03', 9),
                                                  ee.Filter.neq('CLD04', 9),
                                                  ee.Filter.neq('CLD05', 9),
                                                  ee.Filter.neq('CLD06', 9),
                                                  ee.Filter.neq('CLD07', 9),
                                                  ee.Filter.neq('CLD08', 9),
                                                  ee.Filter.neq('CLD09', 9),
                                                  ee.Filter.neq('CLD10', 9),
                                                  ee.Filter.neq('CLD11', 9),
                                                  ee.Filter.neq('CLD12', 9),
                                                  ee.Filter.neq('CLD13', 9),
                                                  ee.Filter.neq('CLD14', 9)));

//assign training data

function trainingData(regionName) {
  var region_assigned = featureList.filterMetadata("region", "equals", regionName)
  return region_assigned;
}

// execute function (here for region LVha)
var LVha1_train = trainingData('LVha-1')
//var LVha2_train = trainingData('LVha-2')
//var CMeu3_train = trainingData('CMeu-3')
//var CMdy1_train = trainingData('CMdy-1')
//var CMdy2_train = trainingData('CMdy-2')
//var CMeu1_train = trainingData('CMeu-1')
//var PZha1_train = trainingData('PZha-1')
//var LVha3_train = trainingData('LVha-3')
```

```
// assigning test data for each except the training region (here: LVha1)
//var LVha1_test = trainingData('LVha-1')
var LVha2_test = trainingData('LVha-2')
var CMeu3_test = trainingData('CMeu-3')
var CMdy1_test = trainingData('CMdy-1')
var CMdy2_test = trainingData('CMdy-2')
var CMeu1_test = trainingData('CMeu-1')
var PZha1_test = trainingData('PZha-1')
var LVha3_test = trainingData('LVha-3')


// create list of test data (without training region LVha1)
var testList = LVha2_test
//  .merge(LVha1_test)
    .merge(CMeu3_test)
    .merge(CMdy1_test)
    .merge(CMdy2_test)
    .merge(CMeu1_test)
    .merge(PZha1_test)
    .merge(LVha3_test)

// Make a Random Forest classifier and train it (here on LVha1 training data).
var classifier = ee.Classifier.smileRandomForest(100)
    .train({
      features: LVha1_train,
      classProperty: label,
      inputProperties: class_bands
    });


// classify the images of all regions except training region (here: LVha1)
//var classified_1to01 = LVha1.select(bands).classify(classifier);
var classified_1to07 = LVha2.select(class_bands).classify(classifier);
var classified_1to08 = CMeu3.select(class_bands).classify(classifier);
var classified_1to09 = CMdy1.select(class_bands).classify(classifier);
var classified_1to10 = CMdy2.select(class_bands).classify(classifier);
var classified_1to11 = CMeu1.select(class_bands).classify(classifier);
var classified_1to12 = PZha1.select(class_bands).classify(classifier);
var classified_1to13 = LVha3.select(class_bands).classify(classifier);
```

```
// Accuracy Assessment
function confusionMatrix(regionName) {
  var testdat_region = testList.filterMetadata("region", "equals", regionName)
  return ee.ConfusionMatrix(testdat_region.classify(classifier).errorMatrix({
  actual: "class_maiz",
  predicted: "classification"
}));
}

//var OA_1to01 = ee.Feature(confusionMatrix('LVha-1').accuracy());
var OA_1to07 = ee.Feature(confusionMatrix('LVha-2').accuracy());
var OA_1to08 = ee.Feature(confusionMatrix('CMeu-3').accuracy());
var OA_1to09 = ee.Feature(confusionMatrix('CMdy-1').accuracy());
var OA_1to10 = ee.Feature(confusionMatrix('CMdy-2').accuracy());
var OA_1to11 = ee.Feature(confusionMatrix('CMeu-1').accuracy());
var OA_1to12 = ee.Feature(confusionMatrix('PZha-1').accuracy());
var OA_1to13 = ee.Feature(confusionMatrix('LVha-3').accuracy());


// print Overall Accuracy (OA)
// print(OA_1to01,'LVha-1')
print(OA_1to07,'LVha-2')
print(OA_1to08,'CMeu-3')
print(OA_1to09,'CMdy-1')
print(OA_1to10,'CMdy-2')
print(OA_1to11,'CMeu-1')
print(OA_1to12,'PZha-1')
print(OA_1to13,'LVha-3')


// print training data
print(LVha1_train, 'LVha1 train')

// print training data separated by class
print(LVha1_train.filter(ee.Filter.eq('class_maiz', 0)), 'LVha-1 train, _maiz 0')
print(LVha1_train.filter(ee.Filter.eq('class_maiz', 1)), 'LVha-1 train, _maiz 1')

// print classification info
print(classifier.explain())

// print confusion matrices
//print(confusionMatrix('LVha-1'))
print(confusionMatrix('LVha-2'))
```

```
print(confusionMatrix('CMeu-3'))
print(confusionMatrix('CMdy-1'))
print(confusionMatrix('CMdy-2'))
print(confusionMatrix('CMeu-1'))
print(confusionMatrix('PZha-1'))
print(confusionMatrix('LVha-3'))

// map classification results

var palette = [
'grey',
'orange'
];


//Map.addLayer(classified_1to01, {
//min: 0,
//max: palette.length - 1,
//palette: palette
//})

Map.addLayer(classified_1to07, {
min: 0,
max: palette.length - 1,
palette: palette
}

Map.addLayer(classified_1to08, {
min: 0,
max: palette.length - 1,
palette: palette
})

Map.addLayer(classified_1to09, {
min: 0,
max: palette.length - 1,
palette: palette
})

Map.addLayer(classified_1to10, {
min: 0,
max: palette.length - 1,
palette: palette
```

```
})

Map.addLayer(classified_1to11, {
min: 0,
max: palette.length - 1,
palette: palette
})

Map.addLayer(classified_1to12, {
min: 0,
max: palette.length - 1,
palette: palette
})

Map.addLayer(classified_1to13, {
min: 0,
max: palette.length - 1,
palette: palette
})
```

Appendix B-1:

Generated and derived data from multi-class classifications

Regions		Imbalance multi-class										Imbalance 'common wheat' binary				
training_region	classified_region	number of classes (training region)	number of training pixels class 0 (other)	number of training pixels class 1 (common wheat)	number of training pixels class 2 (barley)	number of training pixels class 3 (maize)	class balance class 0	class balance class 1	class balance class 2	class balance class 3	ImbalanceIndex multi-class	number of training pixels class 0 (other)	number of training pixels class 1 (common wheat)	class balance class 0	class balance class 1	ImbalanceIndex 'common wheat' binary
CMca1	CMca2	3	1728	2262	2147	0	0,84471	1,10575	1,04954	0	0,10	3875	2262	1,26283	0,73717	0,26
CMca1	CMdy3	3	1728	2262	2147	0	0,84471	1,10575	1,04954	0	0,10	3875	2262	1,26283	0,73717	0,26
CMca1	CMeu2	3	1728	2262	2147	0	0,84471	1,10575	1,04954	0	0,10	3875	2262	1,26283	0,73717	0,26
CMca1	CMeu4	3	1728	2262	2147	0	0,84471	1,10575	1,04954	0	0,10	3875	2262	1,26283	0,73717	0,26
CMca1	PZha2	3	1728	2262	2147	0	0,84471	1,10575	1,04954	0	0,10	3875	2262	1,26283	0,73717	0,26
CMca2	CMca1	3	440	325	325	0	1,21101	0,8945	0,8945	0	0,14	765	325	1,40367	0,59633	0,40
CMca2	CMdy3	3	440	325	325	0	1,21101	0,8945	0,8945	0	0,14	765	325	1,40367	0,59633	0,40
CMca2	CMeu2	3	440	325	325	0	1,21101	0,8945	0,8945	0	0,14	765	325	1,40367	0,59633	0,40
CMca2	CMeu4	3	440	325	325	0	1,21101	0,8945	0,8945	0	0,14	765	325	1,40367	0,59633	0,40
CMca2	PZha2	3	440	325	325	0	1,21101	0,8945	0,8945	0	0,14	765	325	1,40367	0,59633	0,40
CMdy1	CMca1	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMca2	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMdy2	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMdy3	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMeu1	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMeu2	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMeu3	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	CMeu4	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	LVha1	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	LVha2	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	LVha3	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	PZha1	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy1	PZha2	4	297	122	1076	316	0,65599	0,26946	2,37659	0,69795693	0,69	1689	122	1,86527	0,13473	0,87
CMdy2	CMca1	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMca2	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMdy1	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMdy3	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMeu1	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMeu2	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMeu3	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	CMeu4	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	LVha1	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	LVha2	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	LVha3	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	PZha1	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy2	PZha2	3	828	817	0	542	1,5144	1,49428	0	0,9913123	0,34	1370	817	1,25286	0,74714	0,25
CMdy3	CMca1	3	1989	1975	405	0	1,36576	1,35615	0,2781	0	0,48	2394	1975	1,0959	0,9041	0,10
CMdy3	CMca2	3	1989	1975	405	0	1,36576	1,35615	0,2781	0	0,48	2394	1975	1,0959	0,9041	0,10
CMdy3	CMeu2	3	1989	1975	405	0	1,36576	1,35615	0,2781	0	0,48	2394	1975	1,0959	0,9041	0,10
CMdy3	CMeu4	3	1989	1975	405	0	1,36576	1,35615	0,2781	0	0,48	2394	1975	1,0959	0,9041	0,10
CMdy3	PZha2	3	1989	1975	405	0	1,36576	1,35615	0,2781	0	0,48	2394	1975	1,0959	0,9041	0,10
CMeu1	CMeu4	3	534	996		334	0,85944	1,603	0	0,53755365	0,40	868	996	0,93133	1,06867	0,07
CMeu1	LVha3	3	534	996		334	0,85944	1,603	0	0,53755365	0,40	868	996	0,93133	1,06867	0,07
CMeu2	CMca1	3	1626	500	1084	0	1,51963	0,46729	1,01308	0	0,36	2710	500	1,68847	0,31153	0,69
CMeu2	CMca2	3	1626	500	1084	0	1,51963	0,46729	1,01308	0	0,36	2710	500	1,68847	0,31153	0,69
CMeu2	CMdy3	3	1626	500	1084	0	1,51963	0,46729	1,01308	0	0,36	2710	500	1,68847	0,31153	0,69
CMeu2	CMeu4	3	1626	500	1084	0	1,51963	0,46729	1,01308	0	0,36	2710	500	1,68847	0,31153	0,69
CMeu2	PZha2	3	1626	500	1084	0	1,51963	0,46729	1,01308	0	0,36	2710	500	1,68847	0,31153	0,69
CMeu3	CMca1	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	CMca2	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	CMdy1	4	204	944	0	370	0,53755	2,48748	0	0,97496706	0,49	574	944	0,75626	1,24374	0,24
CMeu3	CMdy2	4	204	944	0	370	0,53755	2,48748	0	0,97496706	0,49	574	944	0,75626	1,24374	0,24
CMeu3	CMdy3	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	CMeu1	4	204	944	0	370	0,53755	2,48748	0	0,97496706	0,49	574	944	0,75626	1,24374	0,24
CMeu3	CMeu2	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	CMeu4	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	LVha1	4	204	944	0	370	0,53755	2,48748	0	0,97496706	0,49	574	944	0,75626	1,24374	0,24
CMeu3	LVha2	4	204	944	0	370	0,53755	2,48748	0	0,97496706	0,49	574	944	0,75626	1,24374	0,24
CMeu3	LVha3	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	PZha1	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
CMeu3	PZha2	4	102	472	0	185	0,53755	2,48748	0	0,97496706	0,49	287	472	0,75626	1,24374	0,24
LVha1	CMca1	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMca2	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMdy1	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMdy2	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMdy3	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMeu1	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMeu2	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMeu3	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	CMeu4	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	LVha2	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	LVha3	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	PZha1	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha1	PZha2	4	451	1264	1290	211	0,56095	1,57214	1,60448	0,26243781	0,59	1952	1264	1,21393	0,78607	0,21
LVha2	CMca1	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMca2	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMdy1	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMdy2	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMdy3	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMeu1	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMeu2	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMeu3	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	CMeu4	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	LVha1	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	LVha3	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	PZha1	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
LVha2	PZha2	4	1727	1651	2766	913	0,97889	0,93581	1,56781	0,51750035	0,28	5406	1651	1,5321	0,4679	0,53
PZha2	CMca1	3	1072	511	365	0	1,65092	0,78696	0,56211	0	0,43	1437	511	1,47536	0,52464	0,48
PZha2	CMca2	3	1072	511	365	0	1,65092	0,78696	0,56211	0	0,43	1437	511	1,47536	0,52464	0,48
PZha2	CMdy3	3	1072	511	365	0	1,65092	0,78696	0,56211	0	0,43	1437	511	1,47536	0,52464	0,48
PZha2	CMeu2	3	1072	511	365	0	1,65092	0,78696	0,56211	0	0,43	1437	511	1,47536	0,52464	0,48
PZha2	CMeu4	3	1072	511	365	0	1,65092	0,78696	0,56211	0	0,43	1437	511	1,47536	0,52464	0,48
Average																

Classification results (accuracies and confusion matrices)

training_region	classified_region	OA	class 0 classified as 0	class 0 classified as 1	class 0 classified as 2	class 0 classified as 3	Summe Class 0	User's Accuracy, class0	Producers Accuracy, class0	OA_other	class 1 classified as 0	class 1 classified as 1
CMca1	CMca2	0,33809	140	266	34		440	0,27	0,32	0,30	69	239
CMca1	CMdy3	0,73358	1758	226	5		1989	0,70	0,88	0,79	505	1365
CMca1	CMeu2	0,68660	1280	307	39		1626	0,68	0,79	0,73	172	290
CMca1	CMeu4	0,84256	1425	359	112		1896	1,00	0,75	0,88	0	1497
CMca1	PZha2	0,62577	962	110	0		1072	0,92	0,90	0,91	27	234
CMca2	CMca1	0,46163	1685	8	35		1728	0,45	0,98	0,71	1055	756
CMca2	CMdy3	0,46670	1859	17	113		1989	0,54	0,93	0,74	1539	123
CMca2	CMeu2	0,62928	1391	73	162		1626	0,81	0,86	0,83	296	56
CMca2	CMeu4	0,55277	1850	15	31		1896	0,65	0,98	0,81	978	67
CMca2	PZha2	0,58830	1060	10	2		1072	0,71	0,99	0,85	426	85
CMdy1	CMca1	0,37983	1477	12	62	177	1728	0,55	0,85	0,70	662	854
CMdy1	CMca2	0,52899	242	91	107		440	0,55	0,55	0,55	32	223
CMdy1	CMdy2	0,48505	633	52	127	16	828	0,60	0,76	0,68	164	21
CMdy1	CMdy3	0,46143	1931	0	37	21	1989	0,52	0,97	0,74	1815	85
CMdy1	CMeu1	0,42006	531	0	0	3	534	0,90	0,99	0,95	58	2
CMdy1	CMeu2	0,64299	1484	100	11	31	1626	0,85	0,91	0,88	238	206
CMdy1	CMeu3	0,37813	204	0	0	0	204	0,24	1,00	0,62	630	0
CMdy1	CMeu4	0,60121	1652	0	136	108	1896	0,69	0,87	0,78	733	433
CMdy1	LVha1	0,53794	219	0	195	37	451	0,90	0,49	0,70	20	37
CMdy1	LVha2	0,31926	726	246	755	0	1727	0,63	0,42	0,53	420	565
CMdy1	LVha3	0,71392	451	0	0	44	495	0,54	0,91	0,73	381	0
CMdy1	PZha1	0,56798	1737	46	128	0	1911	0,59	0,91	0,75	1127	225
CMdy1	PZha2	0,48819	951	0	0	121	1072	0,76	0,89	0,82	299	0
CMdy2	CMca1	0,64429	1694	31	3	0	1728	0,70	0,98	0,84	736	1023
CMdy2	CMca2	0,54059	235	193	11	1	440	0,49	0,53	0,51	32	290
CMdy2	CMdy1	0,39867	271	20	0	6	297	0,49	0,91	0,70	57	65
CMdy2	CMdy3	0,45754	1918	0	46	25	1989	0,55	0,96	0,76	1544	81
CMdy2	CMeu1	0,43509	529	0	0	5	534	0,39	0,99	0,69	812	120
CMdy2	CMeu2	0,66573	1422	13	175	16	1626	0,90	0,87	0,89	80	196
CMdy2	CMeu3	0,13439	102	0	0	0	102	0,18	1,00	0,59	472	0
CMdy2	CMeu4	0,51499	1749	110	7	30	1896	0,64	0,92	0,78	971	37
CMdy2	LVha1	0,44372	300	88	59	4	451	0,43	0,67	0,55	241	1018
CMdy2	LVha2	0,44736	1263	280	184	0	1727	0,70	0,73	0,71	252	1250
CMdy2	LVha3	0,46101	475	0	0	20	495	0,47	0,96	0,72	527	32
CMdy2	PZha1	0,59010	1797	79	26	9	1911	0,60	0,94	0,77	1217	496
CMdy2	PZha2	0,54363	1059	0	5	8	1072	0,82	0,99	0,90	237	0
CMdy3	CMca1	0,51915	1679	46	3		1728	0,39	0,97	0,68	826	1423
CMdy3	CMca2	0,41481	168	266	6		440	0,39	0,38	0,39	27	297
CMdy3	CMeu2	0,51495	1151	442	33		1626	0,52	0,71	0,61	4	496
CMdy3	CMeu4	0,73155	1489	360	47		1896	0,74	0,79	0,77	510	1048
CMdy3	PZha2	0,57752	1022	29	21		1072	0,75	0,95	0,85	328	54
CMeu1	CMeu4	0,62572	1369	271	0	256	1896	0,66	0,72	0,69	696	801
CMeu1	LVha3	0,86990	455	0	0	40	495	0,59	0,92	0,76	314	335
CMeu2	CMca1	0,56982	1724	2	2		1728	0,66	1,00	0,83	876	673
CMeu2	CMca2	0,49509	262	178	0		440	0,44	0,60	0,52	32	293
CMeu2	CMdy3	0,51614	1954	0	35		1989	0,57	0,98	0,77	1504	297
CMeu2	CMeu4	0,75231	1791	98	7		1896	0,87	0,94	0,91	259	781
CMeu2	PZha2	0,55031	1072	0	0		1072	0,68	1,00	0,84	511	0
CMeu3	CMca1	0,48134	1362	209	0	157	1728	0,34	0,79	0,57	643	1592
CMeu3	CMca2	0,35950	88	352	0		440	0,23	0,20	0,22	10	315
CMeu3	CMdy1	0,38377	266	30	0	1	297	0,31	0,90	0,60	3	119
CMeu3	CMdy2	0,48400	581	238	0	9	828	0,48	0,70	0,59	65	752
CMeu3	CMdy3	0,75051	1769	215	0	5	1989	0,75	0,89	0,82	465	1510
CMeu3	CMeu1	0,50429	530	4	0	0	534	0,41	0,99	0,70	753	207
CMeu3	CMeu2	0,44517	1115	511	0		1626	0,47	0,69	0,58	186	314
CMeu3	CMeu4	0,83304	1566	321	0	9	1896	0,86	0,83	0,84	249	1323
CMeu3	LVha1	0,53731	293	158	0	0	451	0,18	0,65	0,42	10	1254
CMeu3	LVha2	0,35964	514	1213	0	0	1727	0,19	0,30	0,24	207	1444
Average		0,55						0,60	0,79	0,69		

Appendix B-2:

Generated and derived data from ‘Common Wheat’ classifications

Regions													Classification results (accuracies and confusion matrices)												
training_region	latitude training region	longitude training region (radian measure)	latitude training region (radian measure)	classified_region	latitude classification region	longitude classification region (radian measure)	latitude classification region (radian measure)	total distance	lateral distance	altitude training region	altitude classification region	altitude differ	number of training pixels	number of test pixels	OA	class 0 classified as 0	class 0 classified as 1	User's Accuracy, class 0	Producers Accuracy, class 0	class0, avg acc	class 1 classified as 0	class 1 classified as 1	User's Accuracy, common wheat	Producers Accuracy, common wheat	class1_avg
CMca1	55,595428	0,23117975	0,970323268	PZha2	62,425483	0,440450435	1,089530215	1020,41649	758,136105	35,5836578	814,382365	778,798707	6137	1948	0,75553	962	110	0,78	0,90	0,84	277	234	0,68	0,46	0,57
CMca1	55,595428	0,23117975	0,970323268	CMeu4	58,463351	0,271415489	1,020377967	347,97208	318,339453	35,5836578	59,8075383	24,2238805	6137	3468	0,87486	1537	359	0,95	0,81	0,88	75	1497	0,81	0,95	0,88
CMca1	55,595428	0,23117975	0,970323268	CMeu2	55,54431	0,223182215	0,96943109	29,3591974	5,674098	35,5836578	36,9023908	1,31873301	6137	3210	0,75682	1319	307	0,86	0,81	0,84	210	290	0,49	0,58	0,53
CMca1	55,595428	0,23117975	0,970323268	CMdy3	58,989195	0,266648837	1,029555676	396,523742	376,708137	35,5836578	74,6200558	39,036398	6137	4369	0,78910	1763	226	0,74	0,89	0,81	610	1365	0,86	0,69	0,77
CMca1	55,595428	0,23117975	0,970323268	CMca2	41,810499	-0,081378978	0,729730869	2006,63478	1530,12712	35,5836578	88,8776568	53,293999	6137	1121	0,53987	174	266	0,67	0,40	0,53	86	239	0,47	0,74	0,60
CMca2	41,810499	-0,081378978	0,729730869	PZha2	62,425483	0,440450435	1,089530215	3015,21901	2288,26322	814,382365	35,5836578	778,798707	1090	1948	0,72457	1062	10	0,71	0,99	0,85	426	85	0,89	0,17	0,53
CMca2	41,810499	-0,081378978	0,729730869	CMeu4	58,463351	0,271415489	1,020377967	2326,4029	1848,46657	814,382365	59,8075383	754,574827	1090	3468	0,56171	1881	15	0,56	0,99	0,77	1505	67	0,82	0,04	0,43
CMca2	41,810499	-0,081378978	0,729730869	CMeu2	55,54431	0,223182215	0,96943109	1981,61032	1524,45302	814,382365	36,9023908	777,479974	1090	3210	0,75682	1553	73	0,78	0,96	0,87	444	56	0,43	0,11	0,27
CMca2	41,810499	-0,081378978	0,729730869	CMdy3	58,989195	0,266648837	1,029555676	2356,21925	1906,83526	814,382365	74,6200558	739,762309	1090	4369	0,52851	1972	17	0,52	0,99	0,75	1852	123	0,88	0,06	0,47
CMca2	41,810499	-0,081378978	0,729730869	CMca1	55,595428	0,23117975	0,970323268	2006,63478	1530,12712	814,382365	88,8776568	725,504708	1090	6137	0,62055	1720	8	0,53	1,00	0,76	1506	756	0,99	0,33	0,66
CMdy1	47,370444	-0,029513012	0,826770216	PZha2	62,425483	0,440450435	1,089530215	2368,13163	1671,10933	105,625458	35,5836578	70,0418006	1811	1948	0,67720	1072	0	0,68	1,00	0,84	511	0	0,00	0,00	0,00
CMdy1	47,370444	-0,029513012	0,826770216	PZha1	55,742682	0,159806727	0,972893335	1192,39215	929,318418	105,625458	814,382365	708,756906	1811	5516	0,57560	1865	46	0,56	0,98	0,77	1495	225	0,83	0,13	0,48
CMdy1	47,370444	-0,029513012	0,826770216	LVha3	45,347315	0,172629155	0,791459954	915,716238	224,567319	105,625458	481,69786	376,072402	1811	2744	0,43269	495	0	0,43	1,00	0,72	649	0	0,00	0,00	0,00
CMdy1	47,370444	-0,029513012	0,826770216	LVha2	54,912528	0,180695491	0,958404414	1183,8786	837,171324	105,625458	59,8075383	45,8179201	1811	7057	0,60568	1481	246	0,58	0,86	0,72	1086	565	0,70	0,34	0,52
CMdy1	47,370444	-0,029513012	0,826770216	LVha1	50,306353	0,058790718	0,878011495	493,366307	325,885899	105,625458	43,6912925	61,934166	1811	3216	0,28455	451	0	0,27	1,00	0,63	1227	37	1,00	0,03	0,51
CMdy1	47,370444	-0,029513012	0,826770216	CMeu4	58,463351	0,271415489	1,020377967	1680,60881	1231,31268	105,625458	36,9023908	68,7230676	1811	3468	0,67157	1896	0	0,62	1,00	0,81	1139	433	1,00	0,28	0,64
CMdy1	47,370444	-0,029513012	0,826770216	CMeu3	48,149395	-0,048235106	0,840365476	118,011443	86,463561	105,625458	124,359561	18,7341023	1811	1518	0,17770	204	0	0,18	1,00	0,59	944	0	0,00	0,00	0,00
CMdy1	47,370444	-0,029513012	0,826770216	CMeu2	55,54431	0,223182215	0,96943109	1348,60823	907,299126	105,625458	74,6200558	31,0054027	1811	3210	0,81468	1526	100	0,84	0,94	0,89	294	206	0,67	0,41	0,54
CMdy1	47,370444	-0,029513012	0,826770216	CMeu1	45,007306	0,192076888	0,785525677	1010,58732	262,308318	105,625458	113,279752	7,65429351	1811	1864	0,35033	534	0	0,35	1,00	0,67	994	2	1,00	0,00	0,50
CMdy1	47,370444	-0,029513012	0,826770216	CMdy3	58,989195	0,266648837	1,029555676	1706,77045	1289,68136	105,625458	41,5558315	64,069627	1811	4369	0,52321	1989	0	0,51	1,00	0,76	1890	85	1,00	0,04	0,52
CMdy1	47,370444	-0,029513012	0,826770216	CMdy2	49,746377	0,25063593	0,86823807	1207,76727	263,728563	105,625458	83,4594681	22,1659903	1811	2843	0,48450	776	52	0,49	0,94	0,72	796	21	0,29	0,03	0,16
CMdy1	47,370444	-0,029513012	0,826770216	CMca2	41,810499	-0,081378978	0,729730869	661,270925	617,153895	105,625458	33,9919698	71,6334887	1811	1121	0,74771	349	91	0,77	0,79	0,78	102	223	0,71	0,69	0,70
CMdy1	47,370444	-0,029513012	0,826770216	CMca1	55,595428	0,23117975	0,970323268	1375,27962	912,973224	105,625458	88,8776568	16,7478016	1811	6137	0,64411	1716	12	0,55	0,99	0,77	1408	854	0,99	0,38	0,68
CMdy2	49,746377	0,25063593	0,86823807	PZha2	62,425483	0,440450435	1,089530215	1557,8899	1407,38077	481,69786	35,5836578	446,114202	2187	1948	0,67720	1072	0	0,68	1,00	0,84	511	0	0,00	0,00	0,00
CMdy2	49,746377	0,25063593	0,86823807	PZha1	55,742682	0,159806727	0,972893335	752,569258	665,589855	481,69786	814,382365	332,684505	2187	5516	0,64115	1832	79	0,60	0,96	0,78	1224	496	0,86	0,29	0,58
CMdy2	49,746377	0,25063593	0,86823807	LVha3	45,347315	0,172629155	0,791459954	592,795721	488,295882	481,69786	105,625458	376,072402	2187	2744	0,46066	495	0	0,45	1,00	0,72	617	32	1,00	0,05	0,52
CMdy2	49,746377	0,25063593	0,86823807	LVha2	54,912528	0,180695491	0,958404414	635,370791	573,442761	481,69786	59,8075383	421,890322	2187	7057	0,79840	1447	280	0,78	0,84	0,81	401	1250	0,82	0,76	0,79
CMdy2	49,746377	0,25063593	0,86823807	LVha1	50,306353	0,058790718	0,878011495	786,832752	62,157336	481,69786	43,6912925	438,006568	2187	3216	0,80525	363	88	0,60	0,80	0,70	246	1018	0,92	0,81	0,86
CMdy2	49,746377	0,25063593	0,86823807	CMeu4	58,463351	0,271415489	1,020377967	972,193152	967,584114	481,69786	36,9023908	444,795469	2187	3468	0,52566	1786	110	0,54	0,94	0,74	1535	37	0,25	0,02	0,14
CMdy2	49,746377	0,25063593	0,86823807	CMeu3	48,149395	-0,048235106	0,840365476	1260,03322	177,265002	481,69786	124,359561	357,3383	2187	759	0,17770	102	0	0,18	1,00	0,59	472	0	0,00	0,00	0,00
CMdy2	49,746377	0,25063593	0,86823807	CMeu2	55,54431	0,223182215	0,96943109	653,228835	643,570563	481,69786	74,6200558	407,077804	2187	3210	0,85089	1613	13	0,84	0,99	0,92	304	196	0,94	0,39	0,66
CMdy2	49,746377	0,25063593	0,86823807	CMeu1	45,007306	0,192076888	0,785525677	584,150615	526,036881	481,69786	113,279752	368,418108	2187	1864	0,42745	534	0	0,38	1,00	0,69	876	120	1,00	0,12	0,56
CMdy2	49,746377	0,25063593	0,86823807	CMdy3	58,989195	0,266648837	1,029555676	1029,28443	1025,9528	481,69786	41,5558315	440,142029	2187	4369	0,52220	1989	0	0,51	1,00	0,76	1894	81	1,00	0,04	0,52
CMdy2	49,746377	0,25063593	0,86823807	CMdy1	47,370444	-0,029513012	0,826770216	1207,76727	263,728563	481,69786	83,4594681	398,238392	2187	1811	0,81623	277	20	0,83	0,93	0,88	57	65	0,76	0,53	0,65
CMdy2	49,746377	0,25063593	0,86823807	CMca2	41,810499	-0,081378978	0,729730869	1711,53453	880,882458	481,69786	33,9919698	447,70589	2187	1121	0,70196	247	193	0,88	0,56	0,72	35	290	0,60	0,89	0,75
CMdy2	49,746377	0,25063593	0,86823807	CMca1	55,595428	0,23117975	0,970323268	654,587923	649,244661	481,69786	88,8776568	392,820203	2187	6137	0,68170	1697	31	0,58	0,98	0,78	1239	1023	0,97	0,45	0,71
CMdy3	58,989195	0,266648837	1,029555676	PZha2	62,425483	0,440450435	1,089530215	661,726962	381,427968	59,8075383	35,5836578	24,2238805	4369	1948	0,69299	1043	29	0,70	0,97	0,83	457	54	0,65	0,11	0,38
CMdy3	58,989195	0,266648837	1,029555676	CMeu4	58,463351	0,271415489	1,020377967	60,5495515	58,368684	59,8075383	814,382365	754,574827	4369	3468	0,74510	1536	360	0,75	0,81	0,78	524	1048	0,74	0,67	0,71
CMdy3	58,989195	0,266648837	1,029555676	CMeu2	55,54431	0,223182215	0,96943109	411,146396	382,382235	59,8075383	36,9023908	22,9051475	4369	3210	0,79022	1184	442	1,00	0,73	0,86	4	496	0,53	0,99	0,76
CMdy3	58,989195	0,266648837	1,029555676	CMca2	41,810499	-0,081378978	0,729730869	2356,21925	1906,83526	59,8075383	74,6200558	14,8125174	4369	1121	0,61569	174	266	0,86	0,40	0,63	28	297	0,53	0,91	0,72
CMdy3	58,989195	0,266648837	1,029555676	CMca1	55,595428	0,23117975	0,970323268	396,523742	376,708137	59,8075383	88,8776568	29,0701185	4369	6137	0,77820	1682	46	0,67	0,97	0,82	839	1423	0,97	0,63	0,80
CMeu1	45,007306	0,192076888	0,785525677	LVha3	45,347315	0,172629155	0,791459954	95,1557647	37,740999	43,6912925	74,6200558	30,9287633	1864	2744	0,72552	495	0	0,61	1,00	0,81	314	335	1,00	0,52	0,76
CMeu1	45,007306	0,192076888	0,785525677	CMeu4	58,463351	0,271415489	1,020377967	1527,52701	1493,621	43,6912925	83,4594681	39,7681757	1864	3468	0,69954	1625	271	0,68	0,86	0,77	771	801	0,75	0,51	0,63
CMeu2	55,54431	0,223182215	0,96943109	PZha2	62,425483	0,440450435	1,089530215	1042,43497	763,810203	36,9023908	35,5836578	1,31873301	4369	1948	0,67720	1072	0	0,68	1,00	0,84	511	0	0,00	0,00	0,00
CMeu2	55,54431	0,223182215	0,96943109	CMeu4	58,463351	0,271415489	1,020377967	365,053531	324,013551	36,9023908	814,382365	777,479974	4369	3468	0,74366	1798	98	0,69	0,95	0,82	791	781	0,89	0,50	0,69
CMeu2	55,54431	0,223182215	0,96943109	CMdy3	58,989195	0,266648837	1,029555676	411,146396	382,382235	36,9023908	59,8075383	22,9051475	4369	4369	0,57669	1989	0	0,54	1,00	0,77	1678	297	1,00	0,15	0,58
CMeu2	55,54431	0,223182215	0,96943109	CMca2	41,810499	-0,081378978	0,729730869	1981,61032	1524,45302	36,9023908	74,6200558	37,717665	4369	1121	0,72549	262	178	0,89	0,60	0,74	32	293	0,62	0,90	0,76
CMeu2	55,54431	0,223182215	0,96943109	CMca1	55,595428	0,23117975	0,970323268	29,3591974	5,674098	36,9023908	88,8776568	51,975266	4369	5147	0,60125	1726	2	0,52	1,00	0,76	1589	673	1,00	0,30	0,65
CMeu3	48,149395	-0,048235106	0,840365476	PZha2	62,425483	0,440450435	1,089530215	2345,57478	1584,64577	124,359561	35,5836578	88,7759029	759	1948	0,67467	1068	4	0,68	1,00	0,84	511	0	0,00	0,00	0,00
CMeu3	48,149395	-0,048235106	0,840365476	PZha1	55,742682	0,159806727	0,972893335	1171,62636	842,854857	124,359561	814,382365	690,022804	759	5516	0,89810	1544	367	1,00	0,81	0,90	3	1717	0,82	1,00	0,91
CMeu3	48,149395	-0,048235106	0,840365476	LVha3	45,347315	0,172629155	0,791459954	1011,7109	311,03088	124,359561	105,625458	18,7341023	759	2744	0,47290	480	15	0,45	0,97	0,71	588	61	0,80	0,09	0,45
CMeu3	48,149395	-0,048235106	0,840365476	LVha2	54,912528	0,180695491	0,958404414	1175,01339	750,707763	124,359561	481,69786	357,3383	1518	7057	0,57963	514	1213	0,71	0,30	0,51	207	1444	0,54	0,87	0,71
CMeu3	48,149395	-0,048235106	0,840365476	LVha1	50,306353	0,058790718	0,878011495	505,472186	239,422338	124,359561	59,8075383	64,5520224	1518	3216	0,90204	293	158	0,97	0,65	0,81	10	1254	0,89	0,99	0,94
CMeu3	48,149395	-0,048235106	0,840365476	CMeu4	58,463351	0,271415489	1,020377967	1661,77831	1144,84912	124,359561	43,6912925	80,6682682	759	3468	0,83564	1575	321	0,86	0,83	0,85	249	1323	0,80	0,84	0,82
CMeu3	48,149395	-0,048235106	0,840365476	CMeu2	55,54431	0,223182215	0,96943109	1342,79341	820,835565	124,359561	36,9023908	87,4571699	759	3210	0,67215	1115	511	0,86	0,69	0,77	186	314	0,38	0,63	0,50
CMeu3	48,149395	-0,048235106	0,840365476	CMeu1	45,007306	0,192076888	0,785525677	1106,8626	348,771879	124,359561	74,6200558	49,7395049	1518	1864	0,48170	530	4	0,40	0,99	0,70	789	207	0,98	0,21	0,59
CMeu3	48,149395	-0,048235106	0,840365476	CMdy3	58,989195	0,266648837	1,029555676	1684,06568	1203,2178	124,359561	113,279752	11,0798087	759	4369	0,82846	1774	215	0,79	0,89	0,84	465	1510	0,88	0,76	0,82
CMeu3	48,149395	-0,048235106	0,840365476	CMdy2	49,746377	0,25063593	0,86823807	1260,03322	177,265002	124,359561	41,5558315	82,8037292	1518	2843	0,81581	590	238	0,90	0,71	0,81	65	752	0,76	0,92	0,84
CMeu3	48,149395	-0,048235106	0,840365476	CMdy1	47,370444	-0,029513012	0,826770216	118,011443	86,463561	124,359561	83,4594681	40,9000926	1518	1811	0,92124	267	30	0,99	0,90	0,94	3	119	0,80	0,98	0,89
CMeu3	48,149395	-0,048235106	0,840365476	CMca2	41,810499	-0,081378978	0,729730869	720,328195	703,617456	124,359561	33,9919698	90,3675909	759	1121	0,52680	88	352	0,90	0,20	0,55	10	315	0,47	0,97	0,72
CMeu3	48,149395	-0,048235106	0,840365476	CMca1	55,595428	0,23117975	0,970323268	1370,43443	826,509663	124,359561	88,8776568	35,4819039	759	6137	0,77970	1519	209	0,69	0,88	0,79	670	1592	0,88	0,70	0,79
LVha1	50,306353	0,058790718	0,878011495	PZha2	62,425483	0,440450435	1,089530215	1887,05291	1345,22343	113,279752	35,5836578	77,6960941	3216	1948	0,79975	1052	20	0,78	0,98	0,88	297	214	0,91	0,42	0,67
LVha1	50,306353	0,058790718	0,878011495	PZha1	55,742682	0,159806727	0,972893335	717,147497	603,432519	113,279752	814,382365	701,102613	3216	5516	0,94464	1713	198	1,00	0,90	0,95	3	1717	0,90	1,00	0,95
LVha1	50,306353	0,058790718	0,878011495	LVha3	45,347315	0,172629155	0,791459954	734,954963	550,453218	113,279752	105,625458	7,65429351	3216	2744	0,43531	495	0	0,43	1,00	0,72	646	3	1,00	0,00	0,50
LVha1	50,306353	0,058790718	0,878011495	LVha2	54,912528	0,180695491	0,958404414	695,48276	511,285425	113,279752	481,69786	368,418108	3216	7057	0,74571	1078	649	0,84	0,62	0,73	210	1441	0,69	0,87	0,78
LVha1	50,306353	0,058790718	0,878011495	CMeu4	58,463351	0,271415489	1,020377967	1198,23698	905,426778	113,279752	59,8075383	53,4722136	3216	3468	0,84141	1508	388	0,90	0,80	0,85	162	1410	0,78	0,90	0,84
LVha1	50,306353	0,058790718	0,878011495	CMeu3	48,149395	-0,048235106	0,840365476	505,472186	239,422338	113,279752	43,6912925	69,5884595	3216	759	0,74564	99	3	0,41	0,97	0,69	143	329	0,99	0,70	0,84
LVha1	50,306353	0,058790718	0,878011495	CMeu2	55,54431	0,223182215	0,96943109	857,524499	581,413227	113,279752	36,9023908	76,3773611	3216	3210	0,74506	1284	342	0,87	0,79	0,83	200	300	0,47	0,60	0,53
LVha1	50,306353	0,058790718	0,878011495	CMeu1	45,007306	0,192076888	0,785525677	820,237344	588,194217	113,279752	124,359561	11,0798087	3216	1864	0,34902	534	0	0,35	1,00	0,67	996	0	0,00	0,00	0,00
LVha1	50,306353	0,058790718	0,878011495	CMdy3	58,989195	0,266648837	1,029555676	1228,64014	963,795462	113,279752	74,6200558	38,6596962	3216	4369	0,80323	1803	186	0,75	0,91	0,83	594	1381	0,88	0,70	0,79
LVha1	50,306353	0,058790718	0,878011495	CMdy2	49,746377	0,25063593	0,86823807	786,832752	62,157336	113,279752	41,5558315	71,7239205	3216	2843	0,75805	658	170	0,74	0,79	0,77	228	589	0,78	0,72	0,75
LVha1	50,306353	0,058790718	0,878011495	CMdy1	47,370444	-0,029513012	0,826770216	493,366307	325,885899	113,279752	83,4594681	29,8202838	3216	1811	0,81384	265	32	0,85	0,89	0,87	46	76	0,70	0,62	0,66
LVha1	50,306353	0,058790718	0,878011495	CMca2	41,810499	-0,081378978	0,729730869	1128,16622	943,039794	113,279752	33,9919698	79,2877822	3216	1121	0,52288	279	161	0,58	0,63	0,61	204	121	0,43	0,37	0,40
LVha1	50,306353	0,058790718	0,878011495	CMca1	55,595428	0,23117975	0,970323268	883,669526	587,087325	113,279752	88,8776568	24,4020952	3216	6137	0,75714	1688	40	0,65	0,98	0,81	929	1333	0,97	0,59	0,78
LVha2	54,912528	0,180695491	0,958404414	PZha2	62,425483	0,440450435	1,089530215	1193,84936	833,938005	41,5558315	35,5836578	5,97217366	7057	1948	0,65761	1040	32	0,67	0,97	0,82	510	1	0,03	0,00	0,02
LVha2	54,912528	0,180695491	0,958404414	PZha1	55,742682	0,159806727	0,972893335	119,362224	92,147094	41,5558315	814,382365	772,826533	7057	5516	0,50427	1681	230	0,52	0,88	0,70	1570	150	0,39	0,09	0,24
LVha2	54,912528	0,180695491	0,958404414	LVha3	45,347315	0,172629155	0,791459954	1063,93992	1061,73864	41,5558315	105,625458	64,069627	7057	2744	0,76399	474	21	0,66	0,96	0,81	249	400	0,95	0,62	0,78
LVha2	54,912528	0,180695491	0,958404414	LVha1	50,306353	0,058790718	0,878011495	695,48276	511,285425	41,5558315	481,69786	440,142029	7057	3216	0,74869	419	32	0,51	0,93	0,72	399	865	0,96	0,68	0,82
LVha2	54,912528	0,180695491	0,958404414	CMeu4	58,463351	0,271415489	1,020377967	506,227906	394,141353	41,5558315	59,8075383	18,2517069	7057	3468	0,91494	1782	114	0,91	0,94	0,92	181	1391	0,92	0,88	0,90
LVha2	54,912528	0,180695491	0,958404414	CMeu3	48,149395	-0,048235106	0,840365476	1175,01339	750,707763	41,5558315	43,6912925	2,135461	7057	759	0,17770	102	0	0,18	1,00	0,59	472	0	0,00	0,00	0,00
LVha2	54,912528	0,180695491	0,958404414	CMeu2	55,54431	0,223182215	0,96943109	169,566405	70,127802	41,5558315	36,9023908	4,65344065	7057	3210	0,84431	1443	183	0,91	0,89	0,90	148	352	0,66	0,70	0,68
LVha2	54,912528	0,180695491	0,958404414	CMeu1	45,007306	0,192076888	0,785525677	1102,21186	1099,47964	41,5558315	124,359561	82,8037292	7057	1864	0,35294	532	2	0,35	1,00	0,67	988	8	0,80	0,01	0,40
LVha2	54,912528	0,180695491	0,958404414	CMdy3	58,989195	0,266648837	1,029555676	542,436566	452,510037	41,5558315	74,6200558	33,0642243	7057	4369	0,69854	1795	194	0,64	0,90	0,77	1001	974	0,83	0,49	0,66
LVha2	54,912528	0,180695491	0,958404414	CMdy2	49,746377	0,25063593	0,86823807	635,370791	573,442761	41,5558315	113,279752	71,7239205	7057	2843	0,77082	775	53	0,71	0,94	0,82	324	493	0,90	0,60	0,75
LVha2	54,912528	0,180695491	0,958404414	CMdy1	47,370444	-0,029513012	0,826770216	1183,8786	837,171324	41,5558315	83,4594681	41,9036367	7057	1811	0,68496	277	20	0,71	0,93	0,82	112	10	0,33	0,08	0,21
LVha2	54,912528	0,180695491	0,958404414	CMca2	41,810499	-0,081378978	0,729730869	1822,76	1454,32522	41,5558315	33,9919698	7,5638617	7057	1121	0,43007	261	179	0,50	0,59	0,55	257	68	0,28	0,21	0,24
LVha2	54,912528	0,180695491	0,958404414	CMca1	55,595428	0,23117975	0,970323268	198,367372	75,8019	41,5558315	88,8776568	47,3218253	7057	6137	0,58296	1683	45	0,51	0,97	0,74	1619	643	0,93	0,28	0,61
PZha2	62,425483	0,440450435	1,089530215	CMeu4	58,463351	0,271415489	1,020377967	688,816241	439,796652	88,8776568	35,5836578	53,293999	1948	3468	0,78143	1771	125	0,74	0,93	0,84	633	939	0,88	0,60	0,74
PZha2	62,425483	0,440450435	1,089530215	CMeu2	55,54431	0,223182215	0,96943109	1042,43497	763,810203	88,8776568	814,382365	725,504708	1948	3210	0,89605	1507	119	0,94	0,93	0,93	102	398	0,77	0,80	0,78
PZha2	62,425483	0,440450435	1,089530215	CMdy3	58,989195	0,266648837	1,029555676	661,726962	381,427968	88,8776568	59,8075383	29,0701185	1948	4369	0,59662	1934	55	0,56	0,97	0,76	1544	431	0,89	0,22	0,55
PZha2	62,425483	0,440450435	1,089530215	CMca2	41,810499	-0,081378978	0,729730869	3015,21901	2288,26322	88,8776568	36,9023908	51,975266	1948	1121	0,47582	39	401	1,00	0,09	0,54	0	325	0,45	1,00	0,72
PZha2	62,425483	0,440450435	1,089530215	CMca1	55,595428	0,23117975	0,970323268	1020,41649	758,136105	88,8776568	74,6200558	14,257601	1948	6137	0,59323	1704	24	0,52	0,99	0,75	1599	663	0,97	0,29	0,63
average															0,66								0,69	0,45	0,57

Appendix B-3:

Generated and derived data from 'Maize' classifications

Regions													Imbalance 'maize' binary					Classification results (accuracies and confusion matrices)													Comparison post-/pre processing approach			altitude		
training_region	latitude training region	longitude training region (radian measure)	latitude training region (radian measure)	classified_region	latitude classification region	longitude classification region (radian measure)	latitude classification region (radian measure)	total distance	lateral distance	altitude training region	altitude classification region	altitude difference	pixel class 0	pixel class 1	class balance class 0	class balance class 1	Imbalance Index	number of training pixels	number of test pixels	OA	class 0 classified as 0	class 0 classified as 1	User's Accuracy, class0	Producers Accuracy, class0	class0, avg acc	class 1 classified as 0	class 1 classified as 1	User's Accuracy, maize	Producer's Accuracy, maize	class1_avg	OA maize (post-processing)	|OA post - OA pre|	OA post - OA pre	altitude_train	altitude_class	altitude_diff
CMdy1	47,370444	-0,029513012	0,826770216	CMdy2	49,746377	0,25063593	0,86823807	1207,767268	263,728563	105,6254584	481,6978603	376,0724018	1495	316	1,651021535	0,348978465	0,65	1811	2843	0,91	2268	33	0,90756303	0,98565841	0,94661072	231	311	0,90406977	0,57380074	0,73893525	0,91	0,00	0,00	105,6254584	481,6978603	376,0724018
CMdy1	47,370444	-0,029513012	0,826770216	CMeu1	45,007306	0,19207689	0,78552568	1010,587325	262,308318	105,6254584	43,69129247	61,93416597	1495	316	1,651021535	0,348978465	0,65	1811	1864	0,90	1525	5	0,89285714	0,99673203	0,94479458	183	151	0,96794872	0,45209581	0,71002226	0,95	0,06	-0,06	105,6254584	43,69129247	61,93416597
CMdy1	47,370444	-0,029513012	0,826770216	CMeu3	48,149395	-0,04823511	0,84036548	118,0114434	86,463561	105,6254584	124,3595607	18,73410225	1495	316	1,651021535	0,348978465	0,65	1811	759	1,00	574	0	1	1	1	0	185	1	1	1	1,00	0,00	0,00	105,6254584	124,3595607	18,73410225
CMdy1	47,370444	-0,029513012	0,826770216	LVha1	50,306353	0,05879072	0,87801149	493,3663074	325,885899	105,6254584	113,279752	7,654293514	1495	316	1,651021535	0,348978465	0,65	1811	3216	0,97	2931	74	0,99592253	0,97537438	0,98564845	12	199	0,72893773	0,94312796	0,83603285	1,00	0,02	-0,02	105,6254584	113,279752	7,654293514
CMdy1	47,370444	-0,029513012	0,826770216	LVha2	54,912528	0,18069549	0,95840441	1183,878597	837,171324	105,6254584	41,55583147	64,06962697	1495	316	1,651021535	0,348978465	0,65	1811	7057	0,96	6144	0	0,95925059	1	0,97962529	261	652	1	0,71412924	0,85706462	0,96	0,00	0,00	105,6254584	41,55583147	64,06962697
CMdy1	47,370444	-0,029513012	0,826770216	LVha3	45,347315	0,17262915	0,79145995	915,7162385	224,567319	105,6254584	83,45946813	22,16599031	1495	316	1,651021535	0,348978465	0,65	1811	2744	0,83	1119	25	0,72193548	0,97814685	0,85004117	431	1169	0,97906198	0,730625	0,85484349	0,97	0,13	-0,13	105,6254584	83,45946813	22,16599031
CMdy1	47,370444	-0,029513012	0,826770216	PZha1	55,742682	0,15980673	0,97289333	1192,392152	929,318418	105,6254584	33,99196977	71,63348867	1495	316	1,651021535	0,348978465	0,65	1811	5516	0,96	4127	21	0,95709647	0,99493732	0,9760169	185	1183	0,98255814	0,86476608	0,92366211	0,96	0,00	0,00	105,6254584	33,99196977	71,63348867
CMdy2	49,746377	0,25063593	0,86823807	CMdy1	47,370444	-0,02951301	0,82677022	1207,767268	263,728563	481,6978603	105,6254584	376,0724018	2301	542	1,618712628	0,381287372	0,62	2843	1811	0,96	1473	22	0,96844181	0,98528428	0,97686305	48	268	0,92413793	0,84810127	0,8861196	0,93	0,03	0,03	481,6978603	105,6254584	376,0724018
CMdy2	49,746377	0,25063593	0,86823807	CMeu1	45,007306	0,19207689	0,78552568	584,1506152	526,036881	481,6978603	43,69129247	438,0065678	2301	542	1,618712628	0,381287372	0,62	2843	1864	0,90	1526	4	0,89711934	0,99738562	0,94725248	175	159	0,97546012	0,4760479	0,72575401	0,90	0,01	0,01	481,6978603	43,69129247	438,0065678
CMdy2	49,746377	0,25063593	0,86823807	CMeu3	48,149395	-0,04823511	0,84036548	1260,033222	177,265002	481,6978603	124,3595607	357,3382996	2301	542	1,618712628	0,381287372	0,62	2843	759	0,76	574	0	0,7602649	1	0,88013245	181	4	1	0,02162162	0,51081081	0,76	0,01	0,01	481,6978603	124,3595607	357,3382996
CMdy2	49,746377	0,25063593	0,86823807	LVha1	50,306353	0,05879072	0,87801149	786,8327521	62,157336	481,6978603	113,279752	368,4181083	2301	542	1,618712628	0,381287372	0,62	2843	3216	0,91	2780	225	0,97818438	0,92512479	0,95165458	62	149	0,39839572	0,70616114	0,55227843	0,94	0,02	-0,02	481,6978603	113,279752	368,4181083
CMdy2	49,746377	0,25063593	0,86823807	LVha2	54,912528	0,18069549	0,95840441	635,3707908	573,442761	481,6978603	41,55583147	440,1420288	2301	542	1,618712628	0,381287372	0,62	2843	7057	0,96	6142	2	0,95878864	0,99967448	0,97923156	264	649	0,9969278	0,71084337	0,85388559	0,96	0,00	0,00	481,6978603	41,55583147	440,1420288
CMdy2	49,746377	0,25063593	0,86823807	LVha3	45,347315	0,17262915	0,79145995	592,7957214	488,295882	481,6978603	83,45946813	398,2383921	2301	542	1,618712628	0,381287372	0,62	2843	2744	0,68	1126	18	0,56983806	0,98426573	0,7770519	850	750	0,9765625	0,46875	0,72265625	0,70	0,01	-0,01	481,6978603	83,45946813	398,2383921
CMdy2	49,746377	0,25063593	0,86823807	PZha1	55,742682	0,15980673	0,97289333	752,5692582	665,589855	481,6978603	33,99196977	447,7058905	2301	542	1,618712628	0,381287372	0,62	2843	5516	0,93	4131	17	0,91983968	0,99590164	0,95787066	360	1008	0,98341463	0,73684211	0,86012837	0,92	0,02	0,02	481,6978603	33,99196977	447,7058905
CMeu1	45,007306	0,192076888	0,785525677	CMdy1	47,370444	-0,02951301	0,82677022	1010,587325	262,308318	43,69129247	105,6254584	61,93416597	1530	334	1,641630901	0,358369099	0,64	1864	1811	0,99	1493	2	0,98678123	0,99866221	0,99272172	20	296	0,99328859	0,93670886	0,96499873				43,69129247	105,6254584	61,93416597
CMeu1	45,007306	0,192076888	0,785525677	CMdy2	49,746377	0,25063593	0,86823807	584,1506152	526,036881	43,69129247	481,6978603	438,0065678	1530	334	1,641630901	0,358369099	0,64	1864	2843	0,92	2271	30	0,91831783	0,98696219	0,95264001	202	340	0,91891892	0,62730627	0,7731126				43,69129247	481,6978603	438,0065678
CMeu1	45,007306	0,192076888	0,785525677	CMeu3	48,149395	-0,04823511	0,84036548	1106,862601	348,771879	43,69129247	124,3595607	80,66826823	1530	334	1,641630901	0,358369099	0,64	1864	759	0,92	574	0	0,90251572	1	0,95125786	62	123	1	0,66486486	0,83243243				43,69129247	124,3595607	80,66826823
CMeu1	45,007306	0,192076888	0,785525677	LVha1	50,306353	0,05879072	0,87801149	820,2373444	588,194217	43,69129247	113,279752	69,58845949	1530	334	1,641630901	0,358369099	0,64	1864	3216	0,96	2933	72	0,98290885	0,97603993	0,97947439	51	160	0,68965517	0,75829384	0,72397451				43,69129247	113,279752	69,58845949
CMeu1	45,007306	0,192076888	0,785525677	LVha2	54,912528	0,18069549	0,95840441	1102,211862	1099,479642	43,69129247	41,55583147	2,135460996	1530	334	1,641630901	0,358369099	0,64	1864	7057	0,95	6120	24	0,94927873	0,99609375	0,97268624	327	586	0,96065574	0,64184009	0,80124791				43,69129247	41,55583147	2,135460996
CMeu1	45,007306	0,192076888	0,785525677	LVha3	45,347315	0,17262915	0,79145995	95,15576468	37,740999	43,69129247	83,45946813	39,76817566	1530	334	1,641630901	0,358369099	0,64	1864	2744	0,98	1106	38	0,99371069	0,96678322	0,98024695	7	1593	0,97670141	0,995625	0,98616321	1,00	0,02	-0,02	43,69129247	83,45946813	39,76817566
CMeu1	45,007306	0,192076888	0,785525677	PZha1	55,742682	0,15980673	0,97289333	1200,597172	1191,626736	43,69129247	33,99196977	9,699322694	1530	334	1,641630901	0,358369099	0,64	1864	5516	0,92	4091	57	0,91032488	0,98625844	0,94829166	403	965	0,94422701	0,70540936	0,82481818				43,69129247	33,99196977	9,699322694
CMeu3	48,149395	-0,048235106	0,840365476	CMdy1	47,370444	-0,02951301	0,82677022	118,0114434	86,463561	124,3595607	105,6254584	18,73410225	574	185	1,512516469	0,487483531	0,51	759	1811	0,99	1488	7	0,99798793	0,99531773	0,99665283	3	313	0,978125	0,99050633	0,98431566	1,00	0,00	0,00	124,3595607	105,6254584	18,73410225
CMeu3	48,149395	-0,048235106	0,840365476	CMdy2	49,746377	0,25063593	0,86823807	1260,033222	177,265002	124,3595607	481,6978603	357,3382996	574	185	1,512516469	0,487483531	0,51	759	2843	0,89	2270	31	0,89405278	0,9865276	0,94029019	269	273	0,89802632	0,50369004	0,70085818	0,82	0,07	0,07	124,3595607	481,6978603	357,3382996
CMeu3	48,149395	-0,048235106	0,840365476	CMeu1	45,007306	0,19207689	0,78552568	1106,862601	348,771879	124,3595607	43,69129247	80,66826823	574	185	1,512516469	0,487483531	0,51	759	1864	0,92	1466	64	0,94519665	0,95816993	0,95168329	85	249	0,79552716	0,74550898	0,77051807	0,93	0,01	-0,01	124,3595607	43,69129247	80,66826823
CMeu3	48,149395	-0,048235106	0,840365476	LVha1	50,306353	0,05879072	0,87801149	505,4721859	239,422338	124,3595607	113,279752	11,07980874	574	185	1,512516469	0,487483531	0,51	759	3216	0,99	2996	9	0,99106848	0,99700499	0,99403673	27	184	0,95336788	0,87203791	0,9127029	0,99	0,00	0,00	124,3595607	113,279752	11,07980874
CMeu3	48,149395	-0,048235106	0,840365476	LVha2	54,912528	0,18069549	0,95840441	1175,013391	750,707763	124,3595607	41,55583147	82,80372922	574	185	1,512516469	0,487483531	0,51	759	7057	0,96	6144	0	0,95925059	1	0,97962529	261	652	1	0,71412924	0,85706462	0,95	0,01	0,01	124,3595607	41,55583147	82,80372922
CMeu3	48,149395	-0,048235106	0,840365476	LVha3	45,347315	0,17262915	0,79145995	1011,710896	311,03088	124,3595607	83,45946813	40,90009256	574	185	1,512516469	0,487483531	0,51	759	2744	0,87	1073	71	0,796585	0,93793706	0,86726103	274	1326	0,94917681	0,82875	0,8889634	0,77	0,10	0,10	124,3595607	83,45946813	40,90009256
CMeu3	48,149395	-0,048235106	0,840365476	PZha1	55,742682	0,15980673	0,97289333	1171,626365	842,854857	124,3595607	33,99196977	90,36759092	574	185	1,512516469	0,487483531	0,51	759	5516	0,95	4144	4	0,93713252	0,99903568	0,9680841	278	1090	0,99634369	0,79678363	0,89656366	0,92	0,03	0,03	124,3595607	33,99196977	90,36759092
LVha1	50,306353	0,058790718	0,878011495	CMdy1	47,370444	-0,02951301	0,82677022	493,3663074	325,885899	113,279752	105,6254584	7,654293514	3005	211	1,868781095	0,131218905	0,87	3216	1811	0,99	1481	14	1	0,99063545	0,99531773	0	316	0,95757576	1	0,97878788	1,00	0,01	-0,01	113,279752	105,6254584	7,654293514
LVha1	50,306353	0,058790718	0,878011495	CMdy2	49,746377	0,25063593	0,86823807	786,8327521	62,157336	113,279752	481,6978603	368,4181083	3005	211	1,868781095	0,131218905	0,87	3216	2843	0,87	2251	50	0,87758285	0,97827032	0,92792658	314	228	0,82014388	0,42066421	0,62040405	0,92	0,05	-0,05	113,279752	481,6978603	368,4181083
LVha1	50,306353	0,058790718	0,878011495	CMeu1	45,007306	0,19207689	0,78552568	820,2373444	588,194217	113,279752	43,69129247	69,58845949	3005	211	1,868781095	0,131218905	0,87	3216	1864	0,95	1487	43	0,96936115	0,97189542	0,97062829	47	287	0,86969697	0,85928144	0,8644892	0,95	0,00	0,00	113,279752	43,69129247	69,58845949
LVha1	50,306353	0,058790718	0,878011495	CMeu3	48,149395	-0,04823511	0,84036548	505,4721859	239,422338	113,279752	124,3595607	11,07980874	3005	211	1,868781095	0,131218905	0,87	3216	759	0,99	567	7	1	0,98780488	0,99390244	0	185	0,96354167	1	0,98177083				113,279752	124,3595607	11,07980874
LVha1	50,306353	0,058790718	0,878011495	LVha2	54,912528	0,18069549	0,95840441	695,4827604	511,285425	113,279752	41,55583147	71,72392049	3005	211	1,868781095	0,131218905	0,87	3216	7057	0,96	6137	7	0,9590561	0,99886068	0,97895839	262	651	0,9893617	0,71303395	0,85119783	0,96	0,00	0,00	113,279752	41,55583147	71,72392049
LVha1	50,306353	0,058790718	0,878011495	LVha3	45,347315	0,17262915	0,79145995	734,954963	550,453218	113,279752	83,45946813	29,82028383	3005	211	1,868781095	0,131218905	0,87	3216	2744	0,85	809	335	0,92351598	0,70716783	0,81534191	67	1533	0,82066381	0,958125	0,88939441	0,92	0,06	-0,06	113,279752	83,45946813	29,82028383
LVha1	50,306353	0,058790718	0,878011495	PZha1	55,742682	0,15980673	0,97289333	717,1474966	603,432519	113,279752	33,99196977	79,28778218	3005	211	1,868781095	0,131218905	0,87	3216	5516	0,92	4130	18	0,91149857	0,99566056	0,95357956	401	967	0,98172589	0,70687135	0,84429862	0,92	0,00	0,00	113,279752	33,99196977	79,28778218
LVha2	54,912528	0,180695491	0,958404414	CMdy1	47,370444	-0,02951301	0,82677022	1183,878597	837,171324	41,55583147	105,6254584	64,06962697	5000	913	1,691188906	0,308811094	0,69	5913	1811	1,00	1494	1	0,99732977	0,9993311	0,99833044	4	312	0,99680511	0,98734177	0,99207344	1,00	0,00	0,00	41,55583147	105,6254584	64,06962697
LVha2	54,912528	0,180695491	0,958404414	CMdy2	49,746377	0,25063593	0,86823807	635,3707908	573,442761	41,55583147	481,6978603	440,1420288	5000	913	1,691188906	0,308811094	0,69	5913	2843	0,89	2299	2	0,88457099	0,99913081	0,9418509	300	242	0,99180328	0,44649446	0,71914887	0,93	0,03	-0,03	41,55583147	481,6978603	440,1420288
LVha2	54,912528	0,180695491	0,958404414	CMeu1	45,007306	0,19207689	0,78552568	1102,211862	1099,479642	41,55583147	43,69129247	2,135460996	5000	913	1,691188906	0,308811094	0,69	5913	1864	0,89	1525	5	0,88303416	0,99673203	0,93988309	202	132	0,96350365	0,39520958	0,67935662	0,90	0,01	-0,01	41,55583147	43,69129247	2,135460996
LVha2	54,912528	0,180695491	0,958404414	CMeu3	48,149395	-0,04823511	0,84036548	1175,013391	750,707763	41,55583147	124,3595607	82,80372922	5000	913	1,691188906	0,308811094	0,69	5913	759	1,00	574	0	1	1	1	0	185	1	1	1	1,00	0,00	0,00	41,55583147	124,3595607	82,80372922
LVha2	54,912528	0,180695491	0,958404414	LVha1	50,306353	0,05879072	0,87801149	695,4827604	511,285425	41,55583147	113,279752	71,72392049	5000	913	1,691188906	0,308811094	0,69	5913	3252	0,98	2944	61	0,99762792	0,9797005	0,98866421	7	240	0,79734219	0,97165992	0,88450106	1,00	0,02	-0,02	41,55583147	113,279752	71,72392049
LVha2	54,912528	0,180695491	0,958404414	LVha3	45,347315	0,17262915	0,79145995	1063,93992	1061,738643	41,55583147	83,45946813	41,90363666	5000	913	1,691188906	0,308811094	0,69	5913	2744	0,71	1126	18	0,58737611	0,98426573	0,78582092	791	809	0,97823458	0,505625	0,74192979	0,92	0,21	-0,21	41,55583147	83,45946813	41,90363666
LVha2	54,912528	0,180695491	0,958404414	PZha1	55,742682	0,15980673	0,97289333	119,3622238	92,147094	41,55583147	33,99196977	7,563861699	5000	913	1,691188906	0,308811094	0,69	5913	5516	0,95	4141	7	0,93730195	0,99831244	0,96780719	277	1091	0,99362477	0,79751462	0,8955697	0,96	0,01	-0,01	41,55583147	33,99196977	7,563861699
LVha3	45,347315	0,172629155	0,791459954	CMdy1	47,370444	-0,02951301	0,82677022	915,7162385	224,567319	83,45946813	105,6254584	22,16599031	1144	1600	0,833819242	1,166180758	0,17	2744	1811	0,97	1490	5	0,96502591	0,99665552	0,98084071	54	262	0,98127341	0,82911392	0,90519367				83,45946813	105,6254584	22,16599031
LVha3	45,347315	0,172629155	0,791459954	CMdy2	49,746377	0,25063593	0,86823807	592,7957214	488,295882	83,45946813	481,6978603	398,2383921	1144	1600	0,833819242	1,166180758	0,17	2744	2843	0,82	2301	0	0,81915272	1	0,90957636	508	34	1	0,06273063	0,53136531				83,45946813	481,6978603	398,2383921
LVha3	45,347315	0,172629155	0,791459954	CMeu1	45,007306	0,19207689	0,78552568	95,15576468	37,740999	83,45946813	43,69129247	39,76817566	1144	1600	0,833819242	1,166180758	0,17	2744	1864	0,93	1463	67	0,95746073	0,95620915	0,95683494	65	269	0,80059524	0,80538922	0,80299223				83,45946813	43,69129247	39,76817566
LVha3	45,347315	0,172629155	0,791459954	CMeu3	48,149395	-0,04823511	0,84036548	1011,710896	311,03088	83,45946813	124,3595607	40,90009256	1144	1600	0,833819242	1,166180758	0,17	2744	759	1,00	574	0	1	1	1	0	185	1	1	1	1,00	0,00	0,00	83,45946813	124,3595607	40,90009256
LVha3	45,347315	0,172629155	0,791459954	LVha1	50,306353	0,05879072	0,87801149	734,954963	550,453218	83,45946813	113,279752	29,82028383	1144	1600	0,833819242	1,166180758	0,17	2744	3216	0,95	3005	0	0,94645669	1	0,97322835	170	41	1	0,1943128	0,5971564				83,45946813	113,279752	29,82028383
LVha3	45,347315	0,172629155	0,791459954	LVha2	54,912528	0,18069549	0,95840441	1063,93992	1061,738643	83,45946813	41,55583147	41,90363666	1144	1600	0,833819242	1,166180758	0,17	2744	7057	0,89	6144	0	0,88798959	1	0,9439948	775	138	1	0,15115005	0,57557503				83,45946813	41,55583147	41,90363666
LVha3	45,347315	0,172629155	0,791459954	PZha1	55,742682	0,15980673	0,97289333	1156,878196	1153,885737	83,45946813	33,99196977	49,46749836	1144	1600	0,833819242	1,166180758	0,17	2744	5516	0,81	4120	28	0,80015537	0,99324976	0,89670256	1029	339	0,92370572	0,24780702	0,58575637				83,45946813	33,99196977	49,46749836
PZha1	55,742682	0,159806727	0,972893335	CMdy1	47,370444	-0,02951301	0,82677022	1192,392152	929,318418	33,99196977	105,6254584	71,63348867	4148	1368	1,503988397	0,496011603	0,50	5516	1811	0,98	1461	34	1	0,97725753	0,98862876	0	316	0,90285714	1	0,95142857				33,99196977	105,6254584	71,63348867
PZha1	55,742682	0,159806727	0,972893335	CMdy2	49,746377	0,25063593	0,86823807	752,5692582	665,589855	33,99196977	481,6978603	447,7058905	4148	1368	1,503988397	0,496011603	0,50	5516	2843	0,95	2247	54	0,95739242	0,97653194	0,96696218	100	442	0,89112903	0,81549815	0,85331359				33,99196977	481,6978603	447,7058905
PZha1	55,742682	0,159806727	0,972893335	CMeu1	45,007306	0,19207689	0,78552568	1200,597172	1191,626736	33,99196977	43,69129247	9,699322694	4148	1368	1,503988397	0,496011603	0,50	5516	1864	0,99	1512	18	0,99473684	0,98823529	0,99148607	8	326	0,94767442	0,9760479	0,96186116				33,99196977	43,69129247	9,699322694
PZha1	55,742682	0,159806727	0,972893335	CMeu3	48,149395	-0,04823511	0,84036548	1171,626365	842,854857	33,99196977	124,3595607	90,36759092	4148	1368	1,503988397	0,496011603	0,50	5516	759	1,00	574	0	1	1	1	0	185	1	1	1				33,99196977	124,3595607	90,36759092
PZha1	55,742682	0,159806727	0,972893335	LVha1	50,306353	0,05879072	0,87801149	717,1474966	603,432519	33,99196977	113,279752	79,28778218	4148	1368	1,503988397	0,496011603	0,50	5516	3216	0,96	2895	110	0,99621473	0,96339434	0,97980454	11	200	0,64516129	0,9478673	0,79651429				33,99196977	113,279752	79,28778218
PZha1	55,742682	0,159806727	0,972893335	LVha2	54,912528	0,18069549	0,95840441	119,3622238	92,147094	33,99196977	41,55583147	7,563861699	4148	1368	1,503988397	0,496011603	0,50	5516	7057	0,96	6141	3	0,95923149	0,99951172	0,9793716	261	652	0,99541985	0,71412924	0,85477455				33,99196977	41,55583147	7,563861699
PZha1	55,742682	0,159806727	0,972893335	LVha3	45,347315	0,17262915	0,79145995	1156,878196	1153,885737	33,99196977	83,45946813	49,46749836	4148	1368	1,503988397	0,496011603	0,50	5516	2744	0,86	959	185	0,82246998	0,83828671	0,83037835	207	1393	0,88276299	0,870625	0,876694				33,99196977	83,45946813	49,46749836
Average																				0,927								0,928	0,722			0,03	-0,01			

Appendix C:

Classification accuracies multi-class classifications

Table 5: Achieved Accuracies of all classifications, ordered from best to poorest Overall Accuracy (OA). Average accuracies are calculated as the mean of User's and Producer's Accuracy for each class and classification

No.	Training Region	Classified Region	Overall Accuracy	avg accuracy other	avg accuracy common wheat	avg accuracy barley	avg accuracy maize
1	LVha1	CMeu3	0.92	0.65	0.84	0.00	0.98
2	CMeu1	LVha3	0.87	0.76	0.76	-	0.99
3	CMca1	CMeu4	0.84	0.88	0.88	-	-
4	CMeu3	CMeu4	0.83	0.84	0.82	-	-
5	LVha2	CMeu4	0.82	0.90	0.90	-	-
6	LVha1	PZha1	0.79	0.86	0.88	0.27	0.83
7	PZha2	CMeu4	0.78	0.83	0.74	-	-
8	LVha1	CMeu4	0.78	0.81	0.84	-	-
9	CMeu3	PZha1	0.76	0.82	0.85	0.00	0.84
10	CMeu2	CMeu4	0.75	0.91	0.69	-	-
11	CMeu3	CMdy3	0.75	0.82	0.79	0.00	-
12	LVha2	LVha3	0.74	0.76	0.78	-	0.93
13	CMca1	CMdy3	0.73	0.79	0.72	0.31	-
14	CMdy3	CMeu4	0.73	0.77	0.71	-	-
15	LVha1	CMdy3	0.73	0.80	0.74	0.24	-
16	CMdy1	LVha3	0.71	0.73	0.00	-	0.96
17	LVha1	CMdy1	0.71	0.65	0.40	0.76	0.98
18	CMca1	CMeu2	0.69	0.73	0.43	0.74	-
19	LVha2	CMca1	0.67	0.77	0.61	0.84	-
20	CMdy2	CMeu2	0.67	0.89	0.53	0.52	-
21	LVha2	LVha1	0.66	0.43	0.76	0.71	0.82
22	LVha2	CMeu2	0.66	0.68	0.54	0.67	-
23	CMdy2	CMca1	0.64	0.84	0.71	0.64	-
24	CMdy1	CMeu2	0.64	0.88	0.52	0.60	-
25	LVha1	CMdy2	0.64	0.63	0.65	0.50	0.76
26	LVha1	LVha3	0.64	0.90	0.50	-	0.87
27	CMca2	CMeu2	0.63	0.83	0.22	0.59	-
28	CMca1	PZha2	0.63	0.91	0.52	0.07	-
29	CMeu1	CMeu4	0.63	0.69	0.63	-	-
30	LVha2	PZha1	0.62	0.65	0.20	0.73	0.91
31	CMdy1	CMeu4	0.60	0.78	0.64	-	-
32	PZha2	CMeu2	0.59	0.93	0.78	0.00	-
33	CMdy2	PZha1	0.59	0.77	0.57	0.33	0.83
34	CMca2	PZha2	0.59	0.85	0.49	0.40	-
35	LVha2	CMdy3	0.59	0.68	0.65	0.54	-

No.	Training Region	Classified Region	Overall Accuracy	avg accuracy other	avg accuracy common wheat	avg accuracy barley	avg accuracy maize
36	LVha1	CMca1	0.58	0.65	0.56	0.46	-
37	CMdy3	PZha2	0.58	0.85	0.32	0.19	-
38	CMeu2	CMca1	0.57	0.83	0.63	0.11	-
39	CMdy1	PZha1	0.57	0.75	0.47	0.01	0.93
40	LVha1	PZha2	0.56	0.89	0.65	0.00	-
41	CMca2	CMeu4	0.55	0.81	0.43	-	-
42	CMeu2	PZha2	0.55	0.84	0.00	0.00	-
43	CMeu3	LVha3	0.55	0.70	0.45	-	0.80
44	CMdy2	PZha2	0.54	0.90	0.00	0.00	-
45	CMeu3	PZha2	0.54	0.85	0.00	0.00	-
46	CMdy2	CMca2	0.54	0.51	0.66	0.54	-
47	CMdy1	LVha1	0.54	0.70	0.48	0.73	0.90
48	CMeu3	LVha1	0.54	0.42	0.73	0.00	0.93
49	PZha2	CMdy3	0.53	0.71	0.50	0.29	-
50	CMdy1	CMca2	0.53	0.55	0.58	0.39	-
51	CMdy3	CMca1	0.52	0.68	0.53	0.44	-
52	CMeu2	CMdy3	0.52	0.77	0.58	0.01	-
53	CMdy2	CMeu4	0.51	0.78	0.14	-	-
54	CMdy3	CMeu2	0.51	0.61	0.62	0.08	-
55	LVha2	CMdy2	0.51	0.66	0.62	0.06	0.78
56	CMeu3	CMeu1	0.50	0.70	0.59	-	0.73
57	CMeu2	CMca2	0.50	0.52	0.64	0.00	-
58	CMdy1	PZha2	0.49	0.82	0.00	0.00	-
59	CMdy1	CMdy2	0.49	0.68	0.07	0.50	0.74
60	CMeu3	CMdy2	0.48	0.59	0.68	0.00	0.45
61	CMeu3	CMca1	0.48	0.57	0.56	0.00	-
62	CMca2	CMdy3	0.47	0.74	0.35	0.22	-
63	CMca2	CMca1	0.46	0.71	0.38	0.52	-
64	CMdy1	CMdy3	0.46	0.74	0.52	0.00	-
65	CMdy2	LVha3	0.46	0.72	0.52	-	0.73
66	CMdy2	CMdy3	0.46	0.76	0.52	0.00	-
67	CMdy2	LVha2	0.45	0.71	0.72	0.02	0.84
68	CMeu3	CMeu2	0.45	0.58	0.40	0.00	-
69	CMdy2	LVha1	0.44	0.55	0.81	0.36	0.23
70	CMdy2	CMeu1	0.44	0.69	0.56	-	0.69
71	LVha1	LVha2	0.43	0.25	0.63	0.29	0.85
72	CMdy1	CMeu1	0.42	0.95	0.50	-	0.87
73	CMdy3	CMca2	0.41	0.39	0.64	0.00	-
74	LVha1	CMeu1	0.40	0.75	0.00	-	0.82
75	CMdy2	CMdy1	0.40	0.70	0.37	0.58	0.78

No.	Training Region	Classified Region	Overall Accuracy	avg accuracy other	avg accuracy common wheat	avg accuracy barley	avg accuracy maize
76	LVha1	CMeu2	0.39	0.53	0.39	0.00	-
77	PZha2	CMca1	0.39	0.73	0.50	0.10	-
78	CMeu3	CMdy1	0.38	0.60	0.57	0.00	0.99
79	CMdy1	CMca1	0.38	0.70	0.49	0.00	-
80	CMdy1	CMeu3	0.38	0.62	0.00	0.00	1.00
81	LVha2	CMeu1	0.36	0.66	0.40	-	0.59
82	CMeu3	LVha2	0.36	0.24	0.59	0.00	0.82
83	CMeu3	CMca2	0.36	0.22	0.65	0.00	-
84	LVha2	CMeu3	0.36	0.85	0.00	0.00	0.99
85	LVha1	CMca2	0.35	0.23	0.34	0.47	-
86	CMca1	CMca2	0.34	0.30	0.52	0.00	-
87	LVha2	CMdy1	0.33	0.76	0.10	0.50	0.98
88	PZha2	CMca2	0.32	0.10	0.66	0.00	-
89	CMdy1	LVha2	0.32	0.53	0.52	0.15	0.86
90	LVha2	PZha2	0.31	0.66	0.01	0.26	-
91	LVha2	CMca2	0.26	0.51	0.21	0.13	-
92	CMdy2	CMeu3	0.13	0.59	0.00	0.00	-

Appendix D:

Region combination matrices

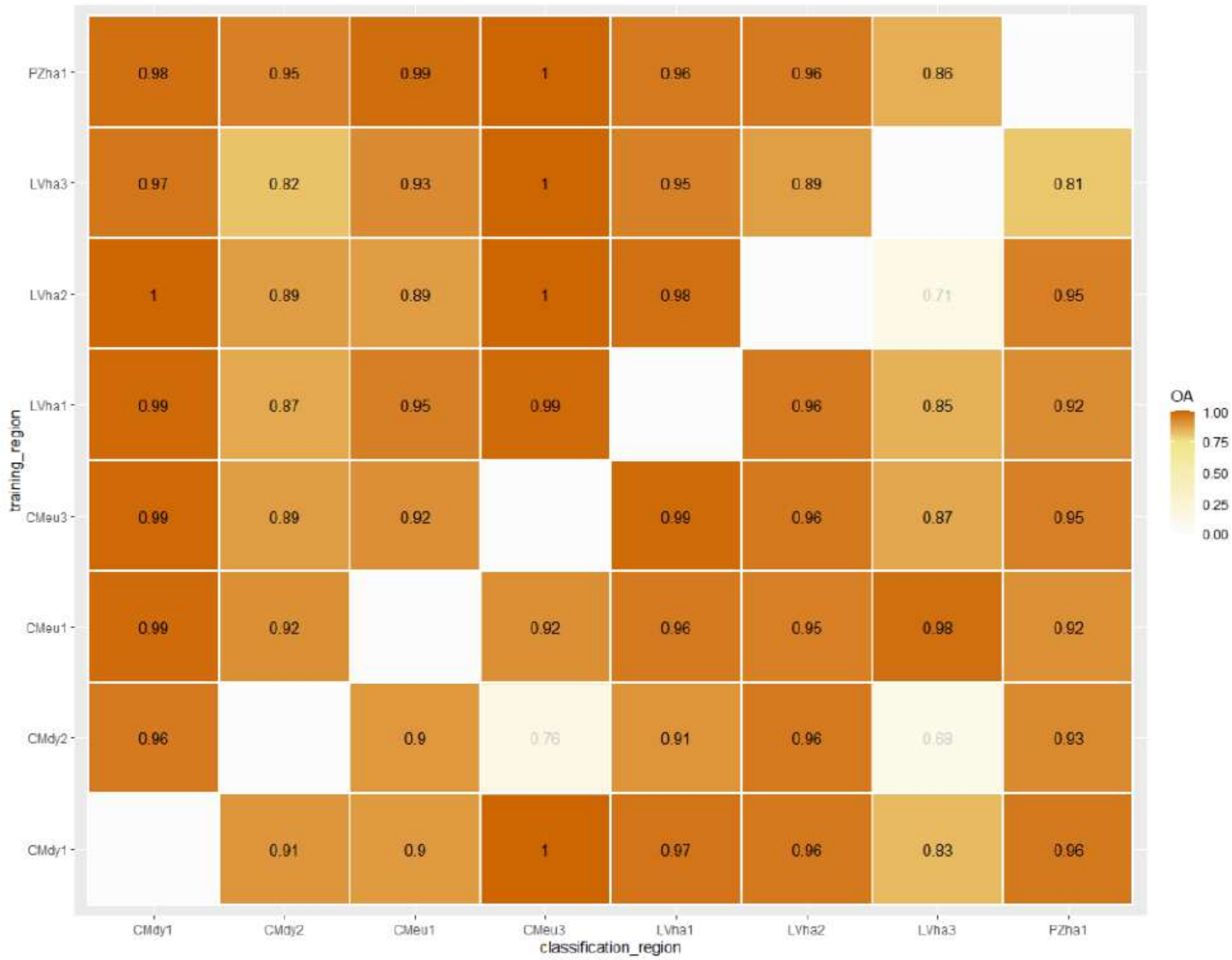

Figure 23: Overall Accuracies for classification of maize with OA > 0.8 highlighted.

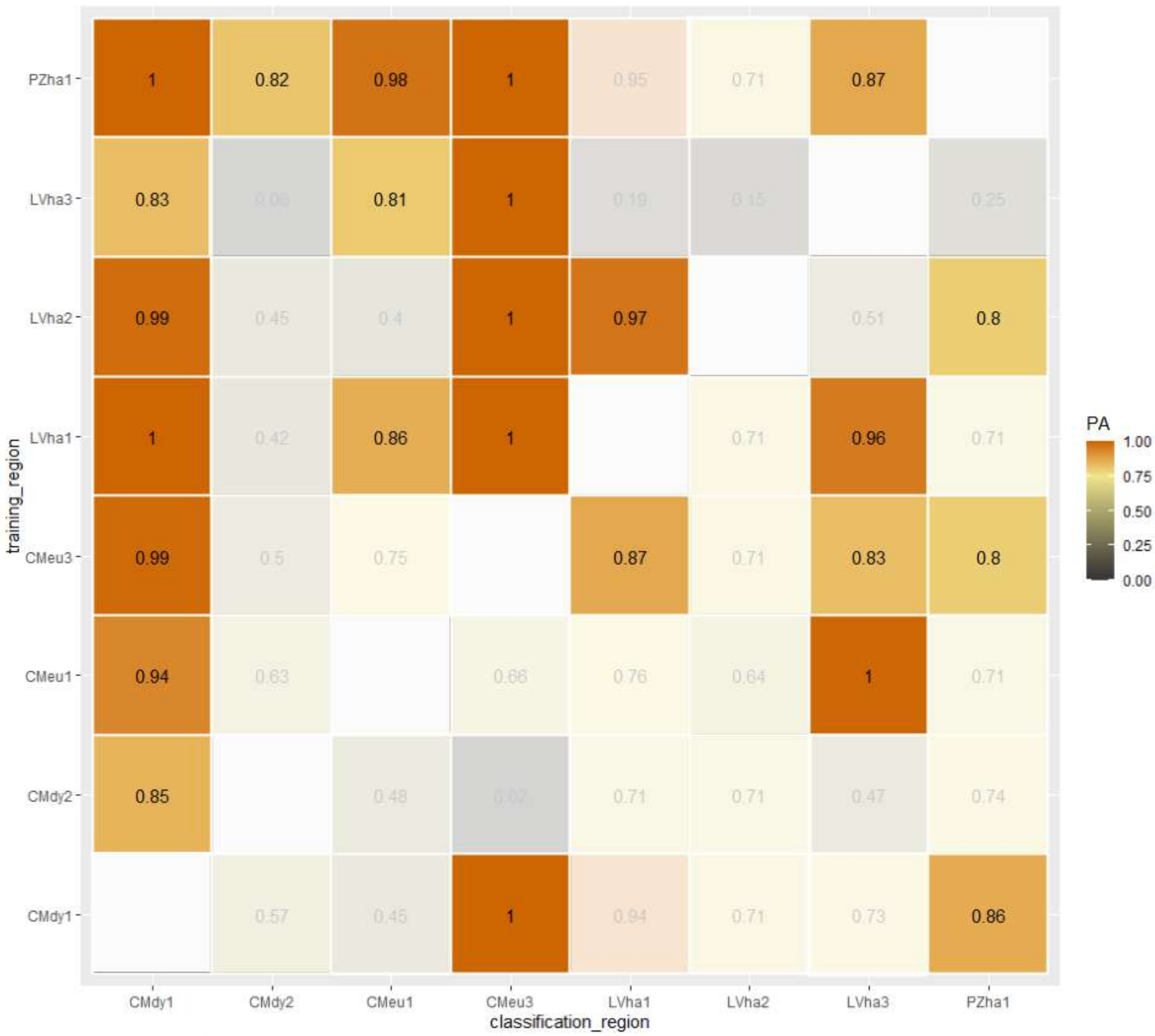

Figure 24: Producer's Accuracies for classification of maize with PA and UA simultaneously > 0.8 highlighted.

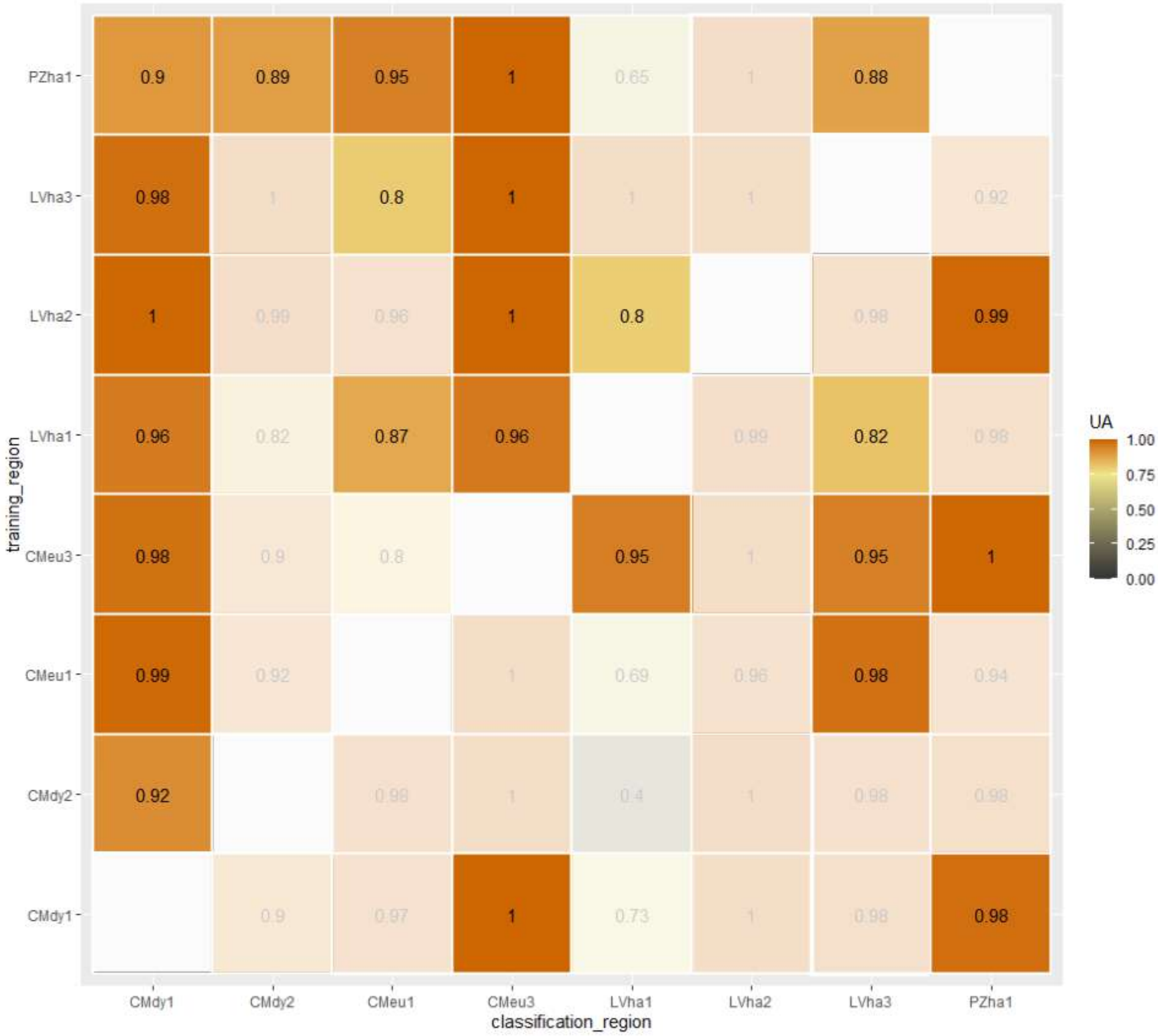

Figure 25: User's Accuracies for classification of maize with PA and UA simultaneously > 0.8 highlighted.

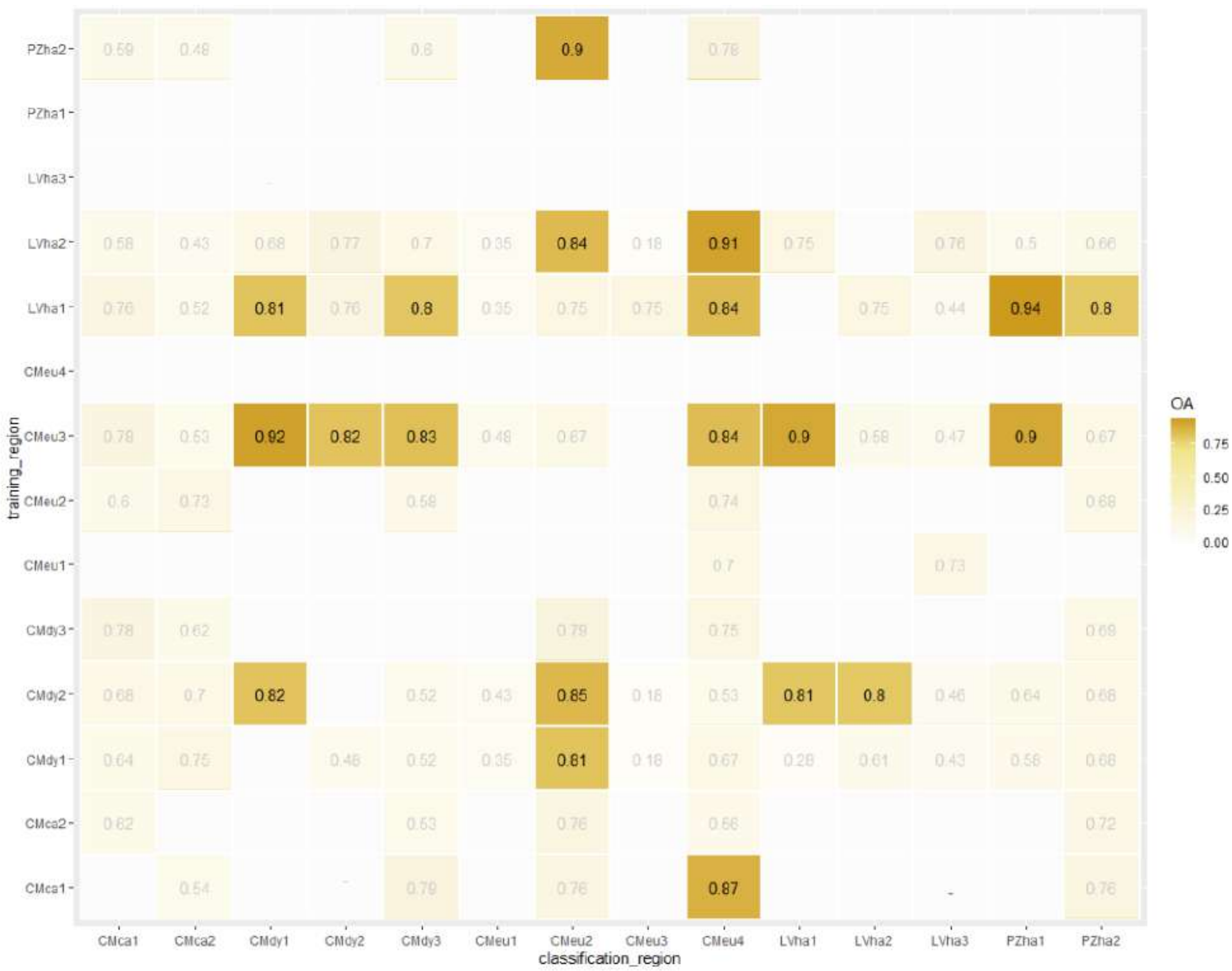

Figure 26: Overall Accuracies for classification of 'Common Wheat' with OA > 0.8 highlighted.

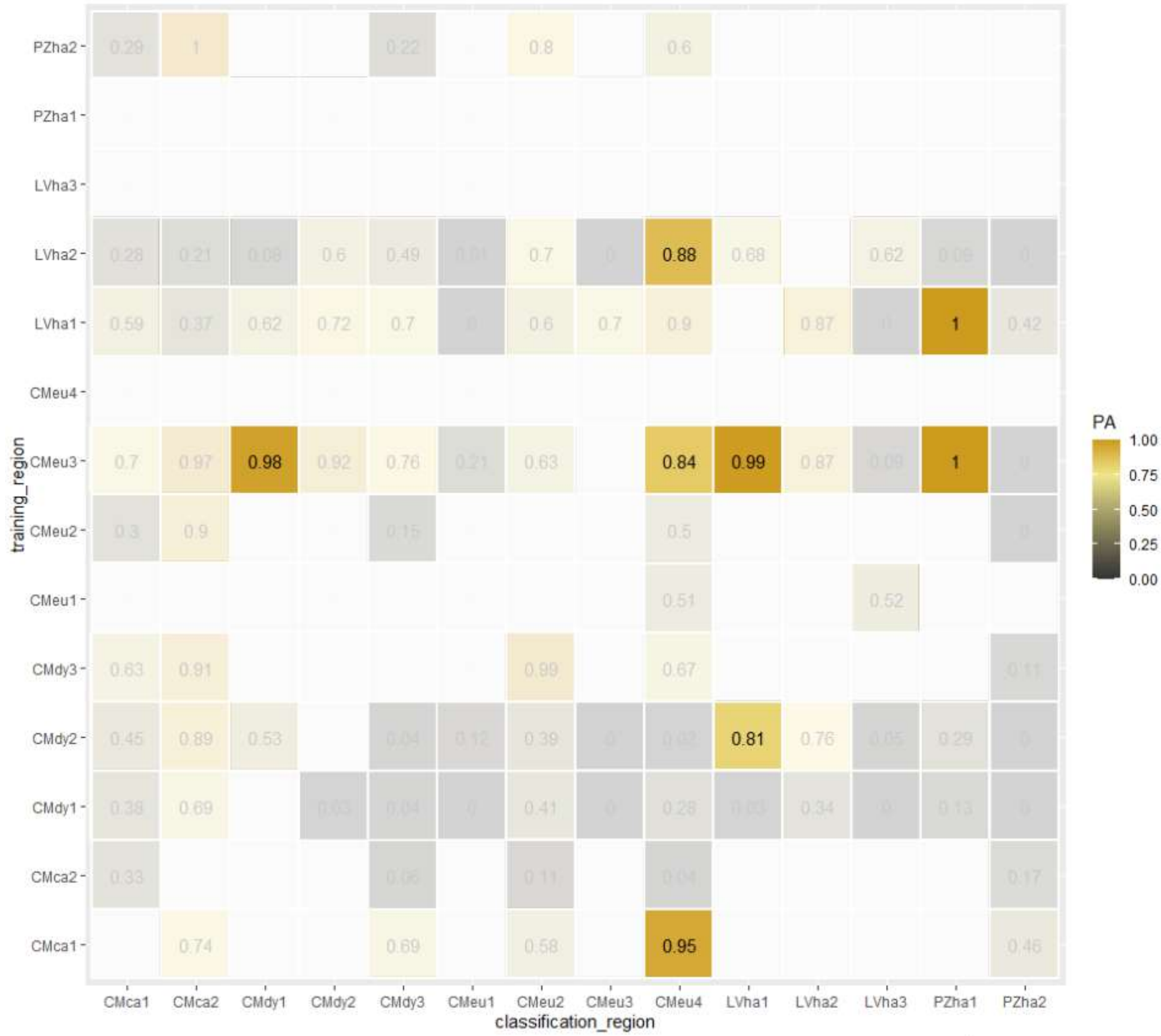

Figure 27: Producer's Accuracies for classification of 'Common Wheat' with PA and UA simultaneously > 0.8 highlighted.

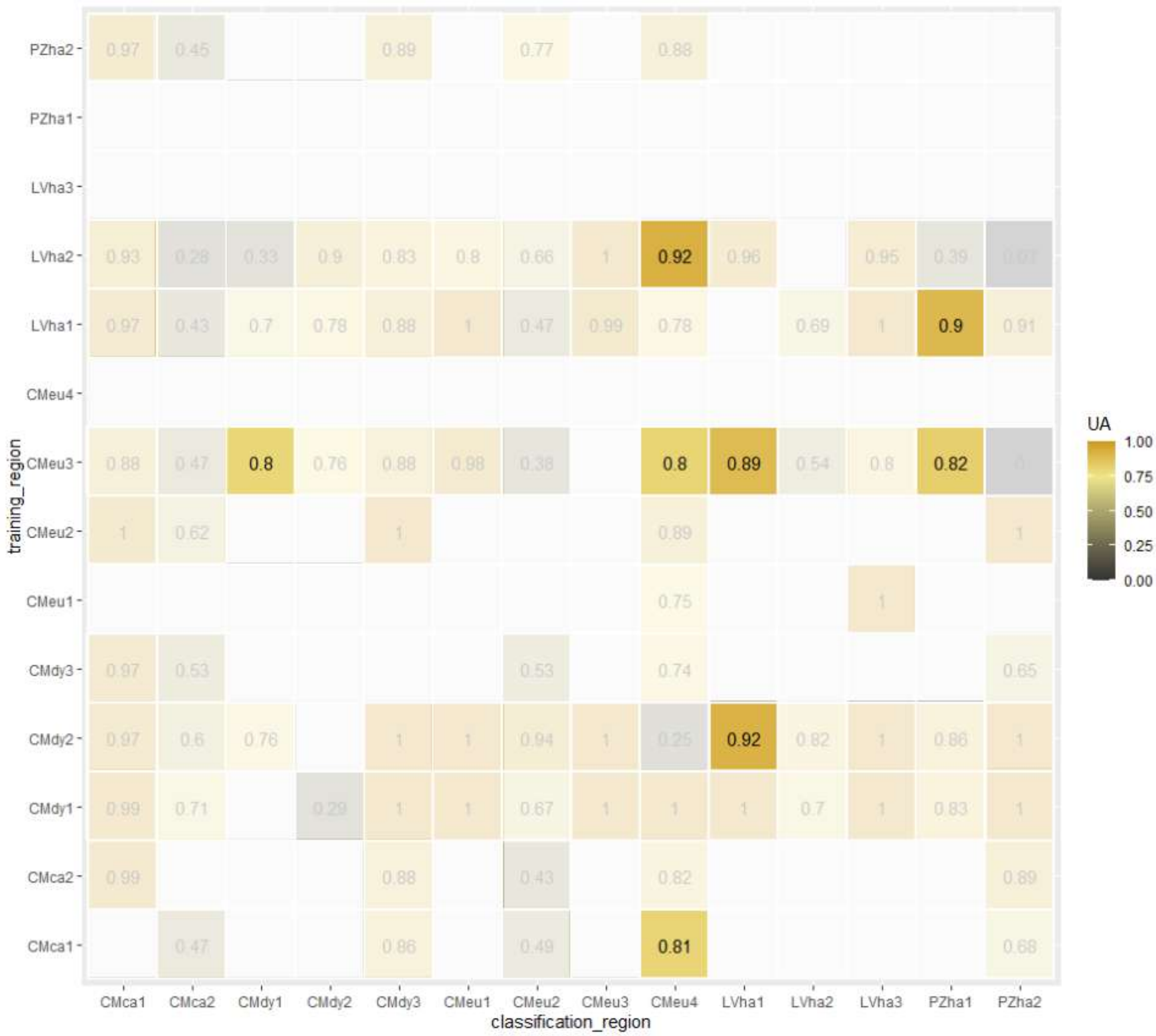

Figure 28: User's Accuracies for classification of 'Common Wheat' with PA and UA simultaneously > 0.8 highlighted.

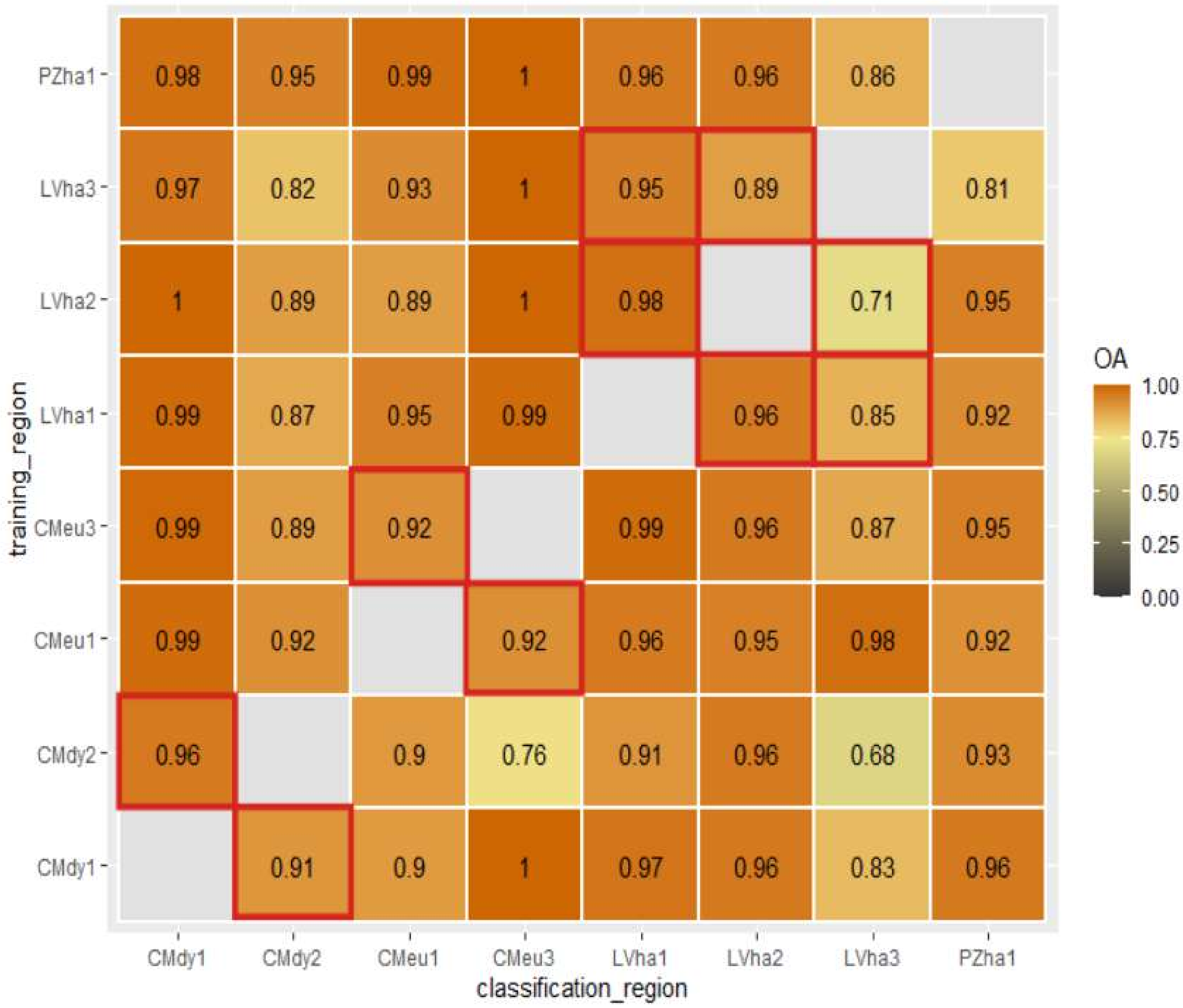

Figure 29: Overall Accuracies for classification of 'Maize' in context of soil type. Red squares indicate results from classifications, where training and classification region are of the same soil type.

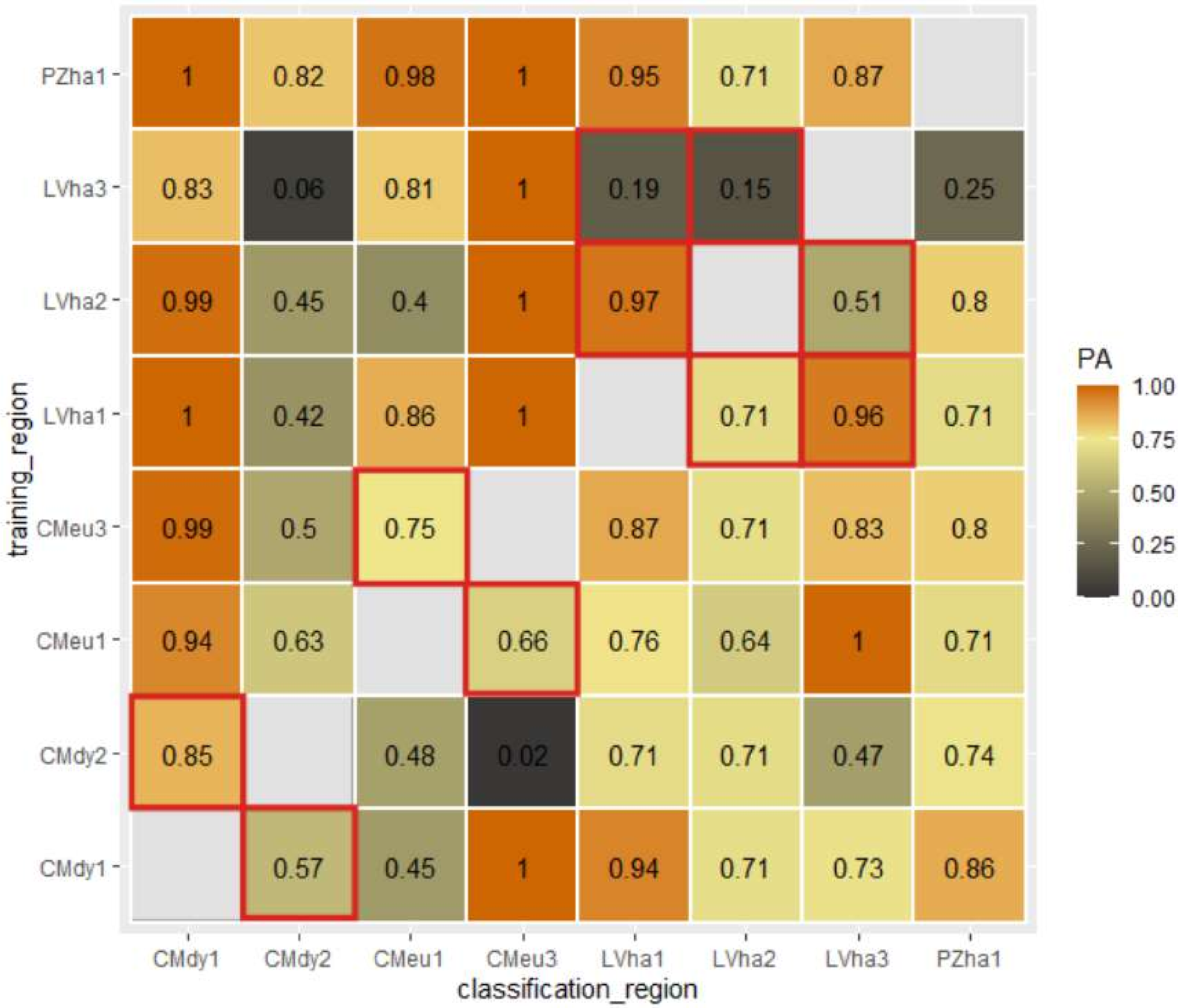

Figure 30: Producer's Accuracies for classification of 'Maize' in context of soil type. Red squares indicate results from classifications, where training and classification region are of the same soil type.

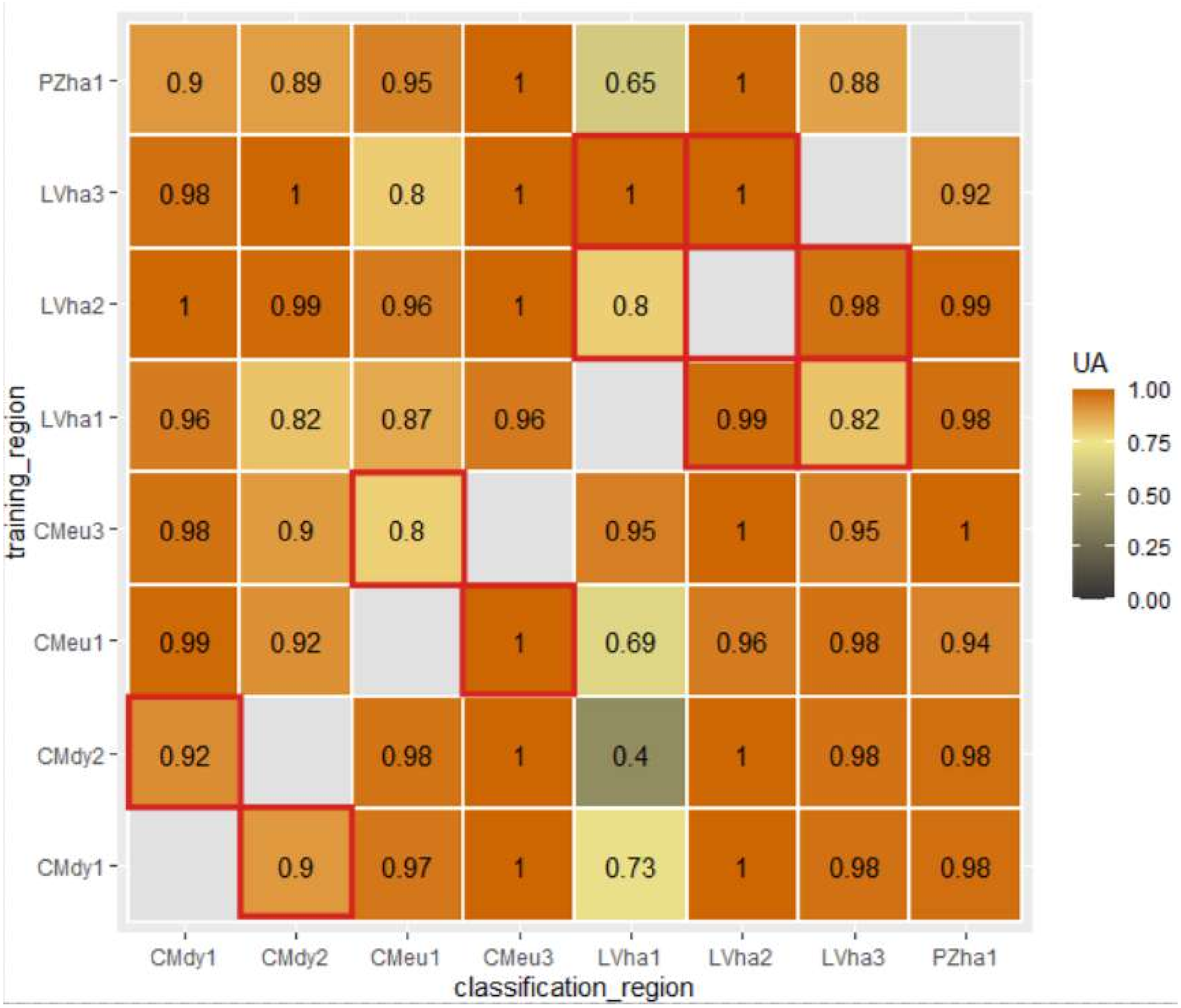

Figure 31: User's Accuracies for classification of 'Maize' in context of soil type. Red squares indicate results from classifications, where training and classification region are of the same soil type.

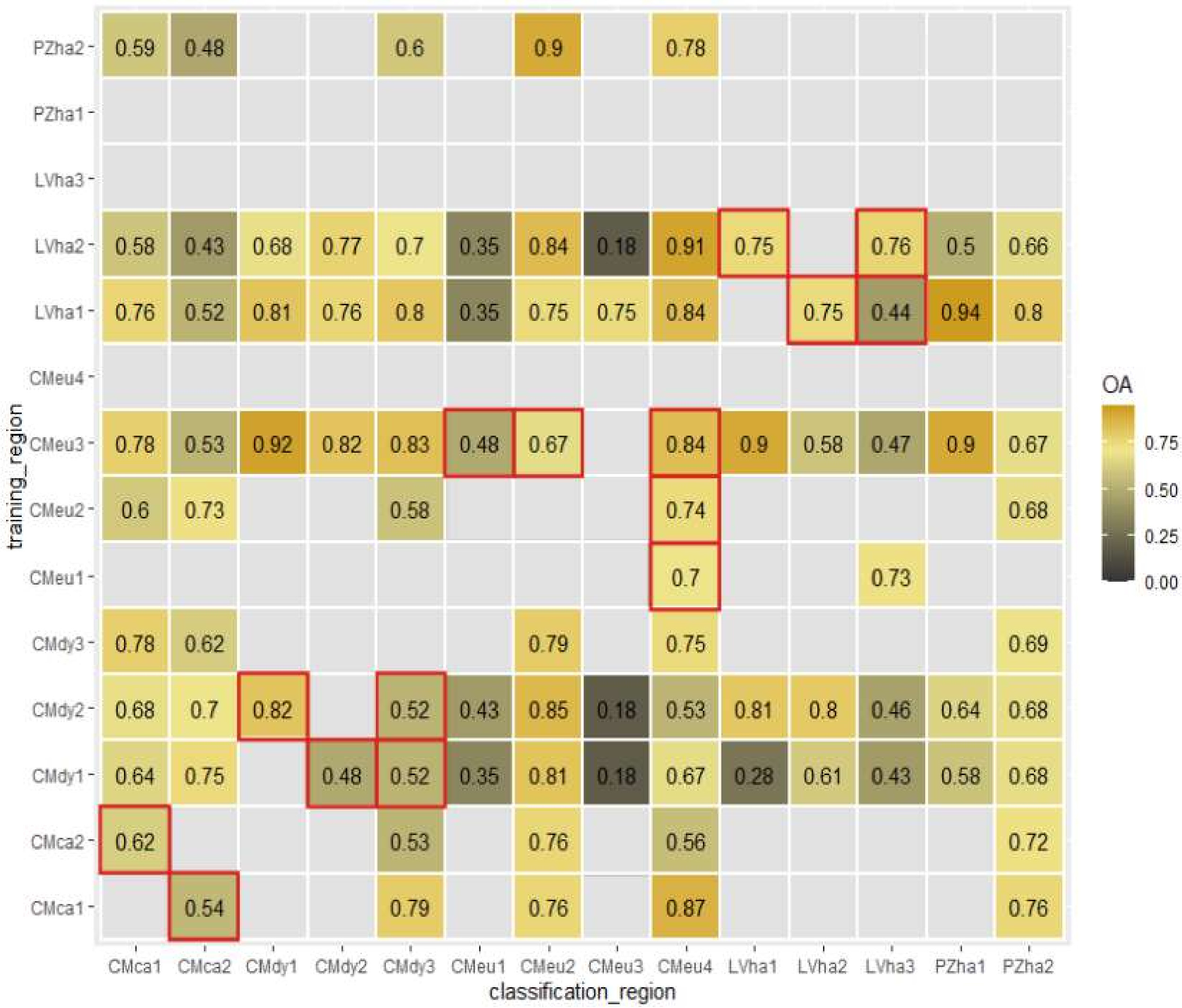

Figure 32: Overall Accuracies for classification of 'Common Wheat' in context of soil type. Red squares indicate results from classifications, where training and classification region are of the same soil type.

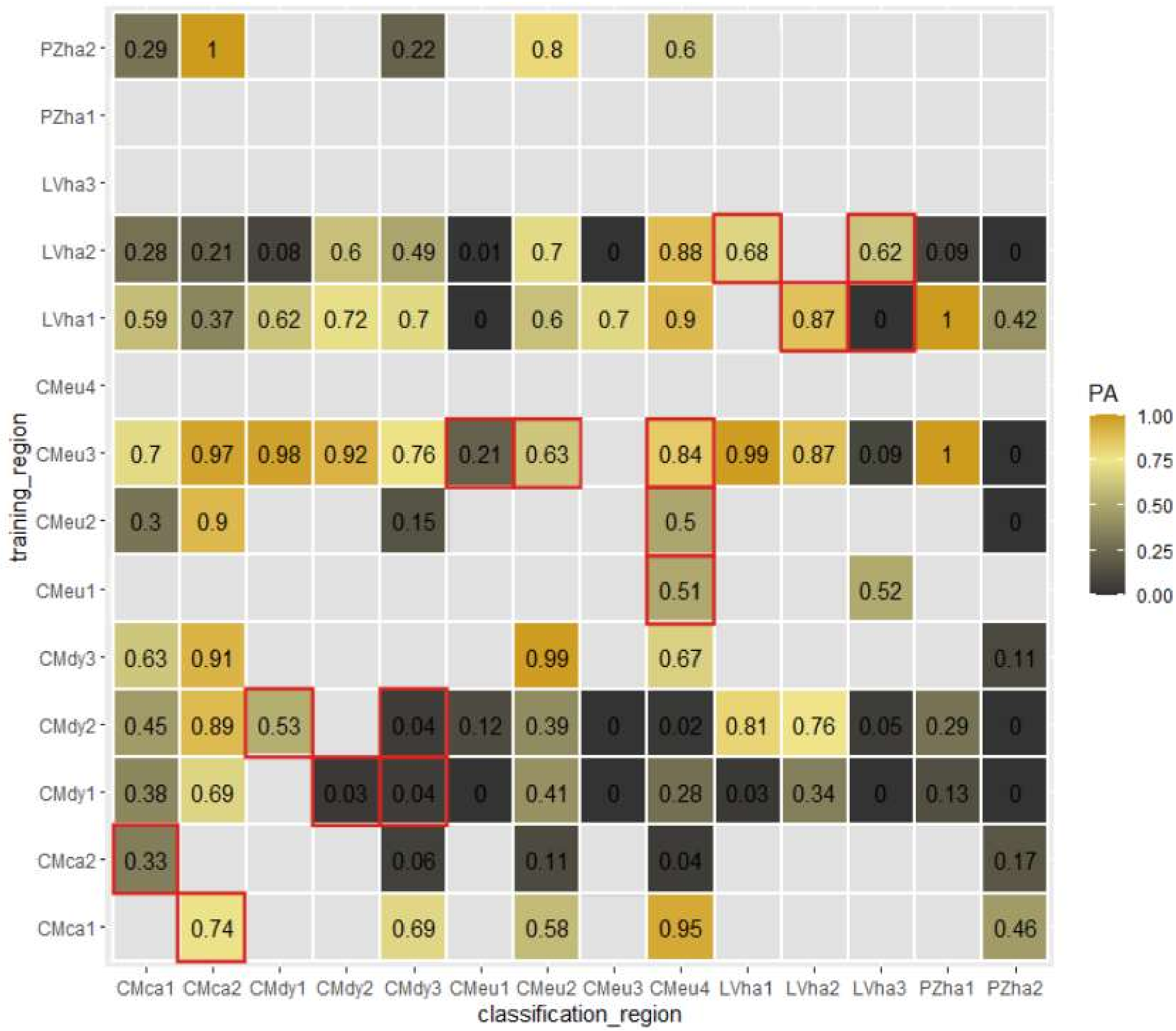

Figure 33: Producer's Accuracies for classification of 'Common Wheat' in context of soil type. Red squares indicate results from classifications, where training and classification region are of the same soil type.

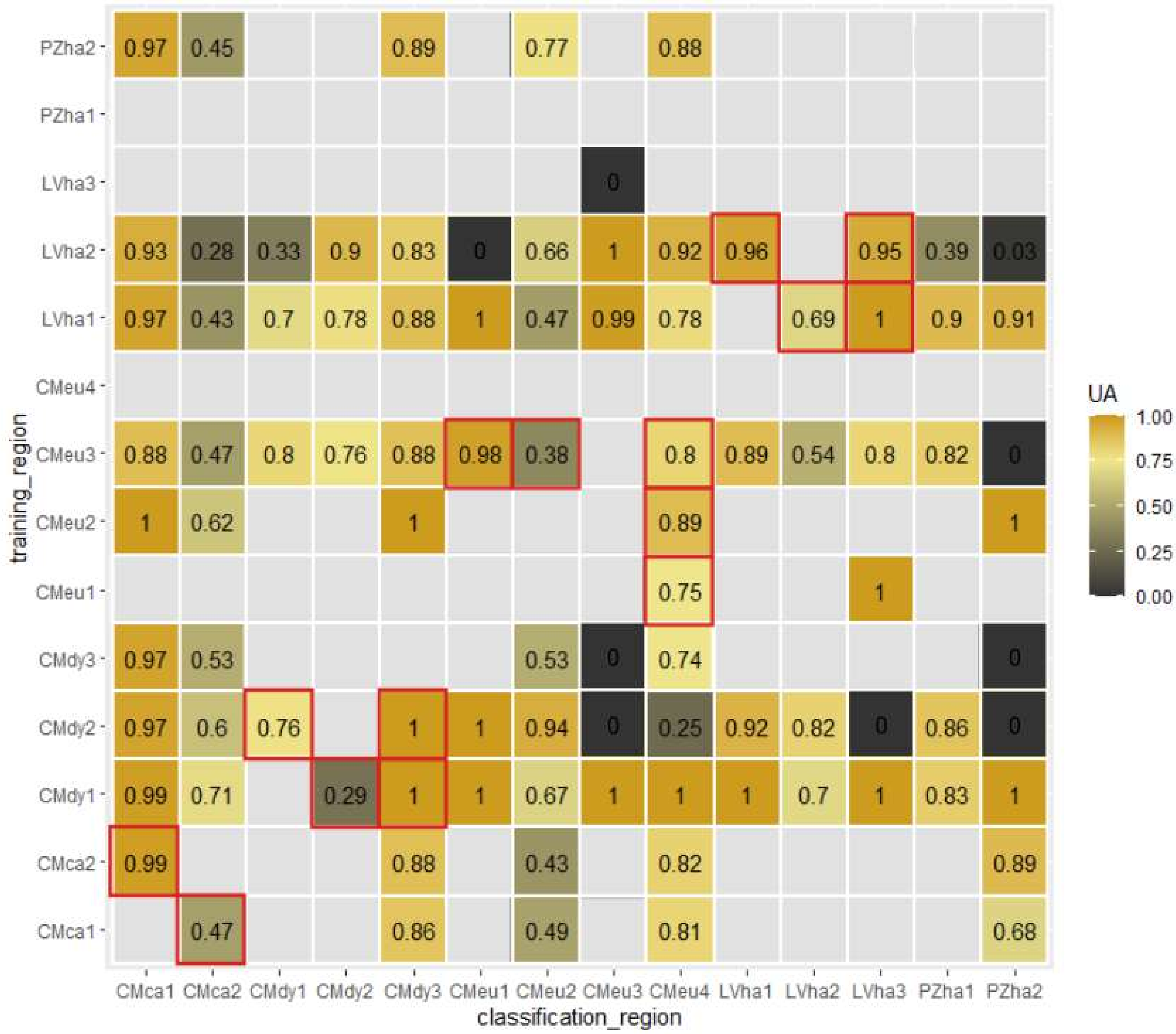

Figure 34: User's Accuracies for classification of 'Common Wheat' in context of soil type. Red squares indicate results from classifications, where training and classification region are of the same soil type.

Appendix E:

Spectral profiles of ‘Maize’, ‘Common Wheat’ and ‘Other’

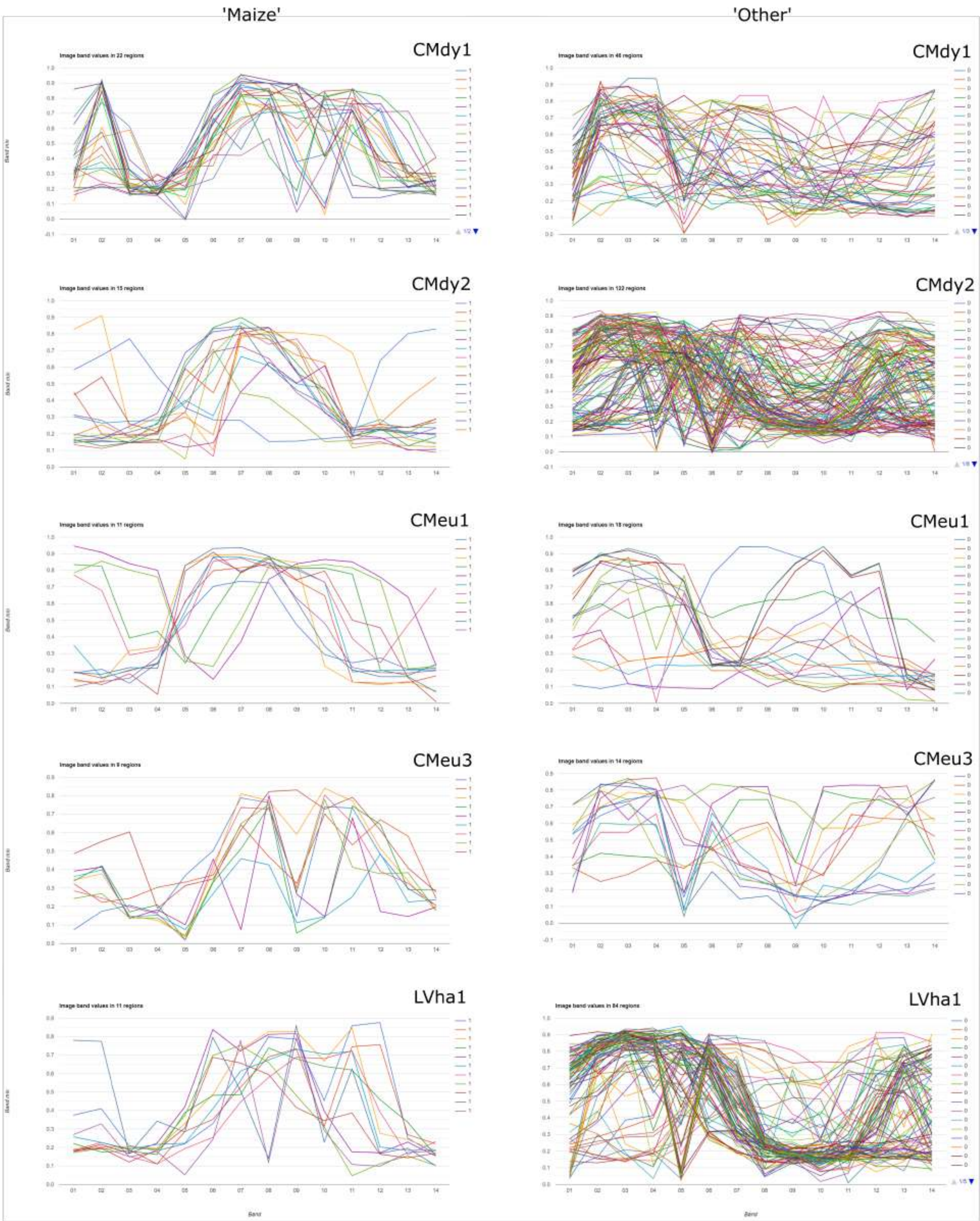

Figure 35: Part 1 - Spectral profiles of 'Maize' and 'Other' from binary classifications (NDVI 1-14)

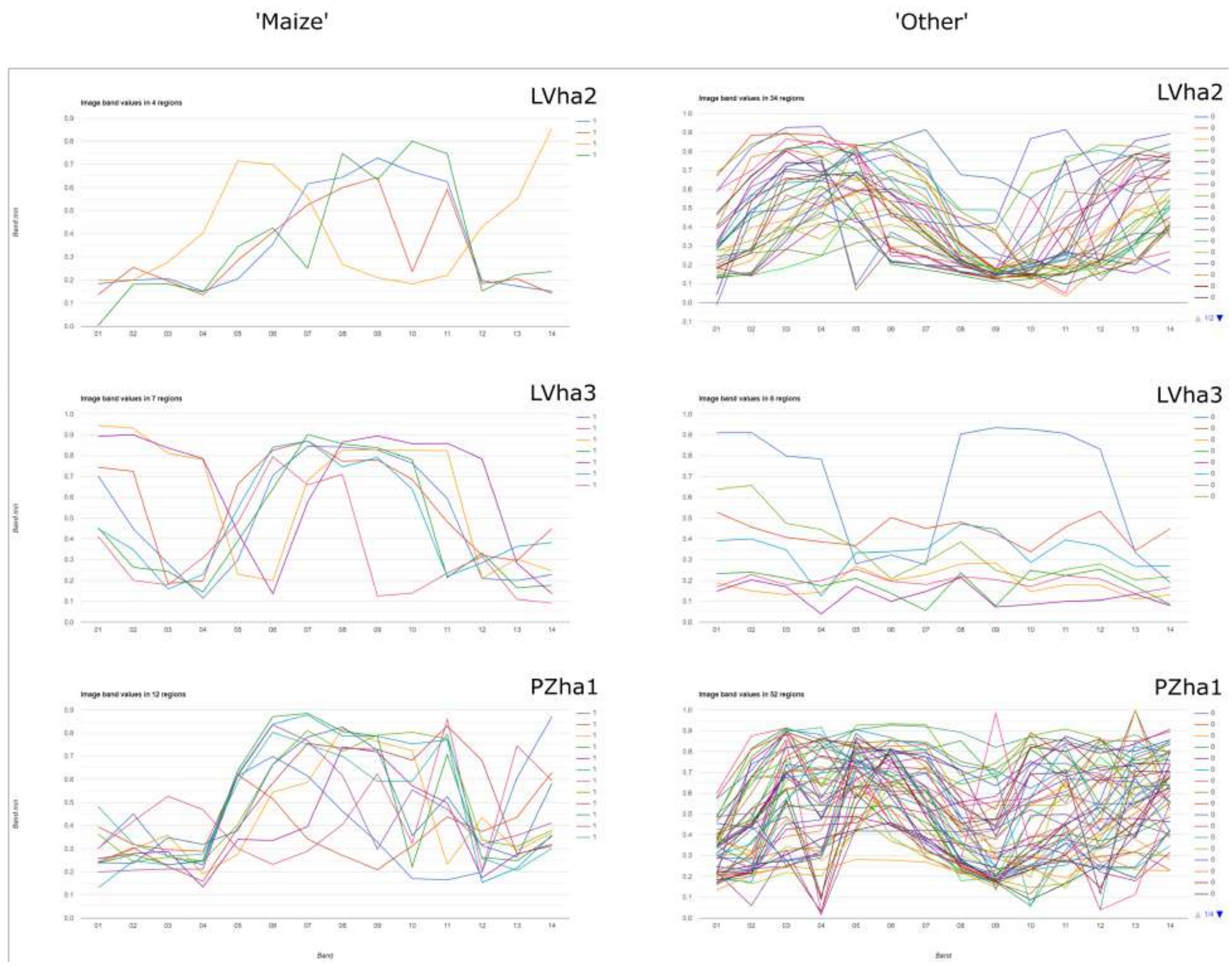

Figure 36: Part 2 - Spectral profiles of 'Maize' and 'Other' from binary classifications (NDVI 1-14)

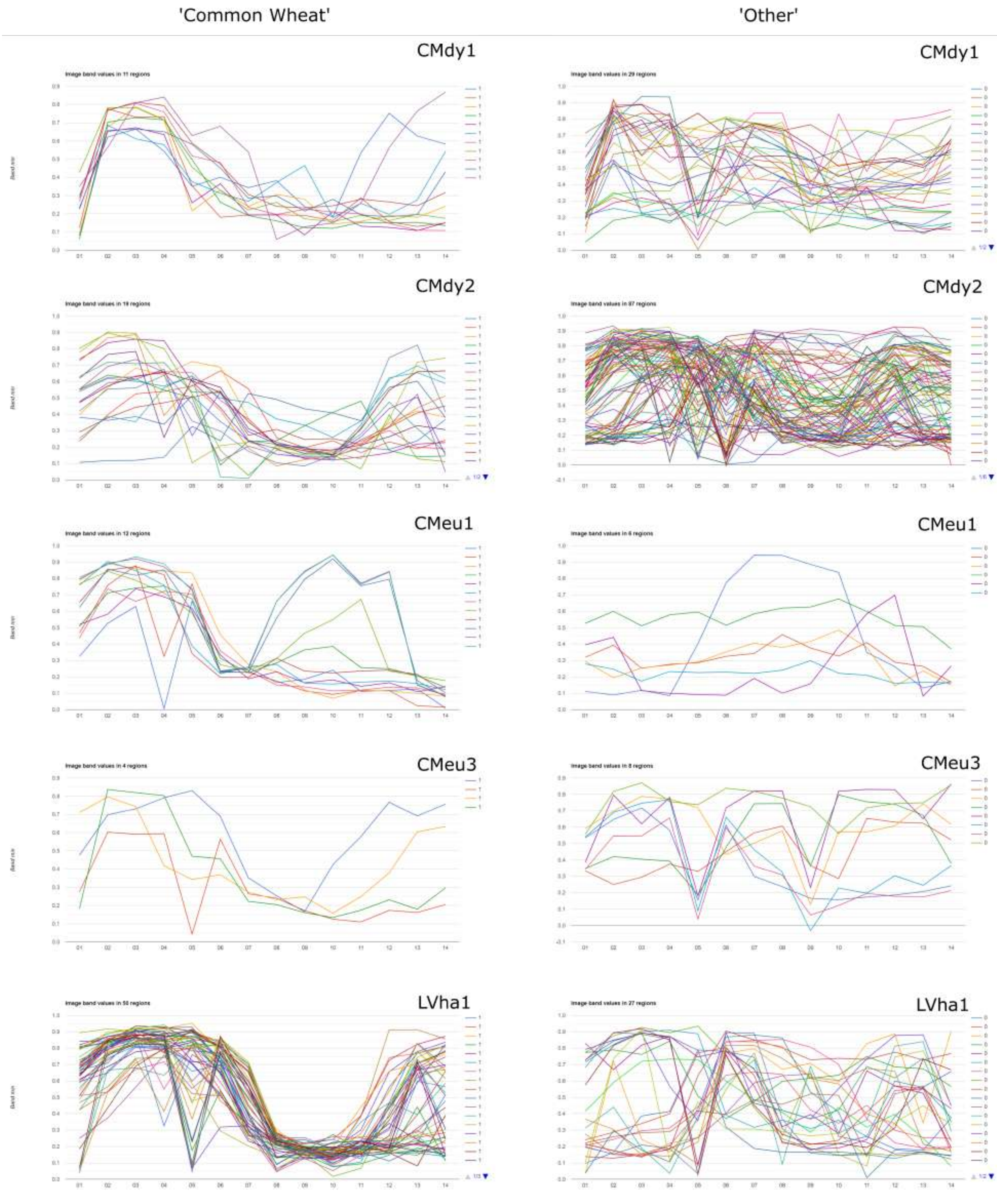

Figure 37: Part 1 - Spectral profiles of 'Common Wheat' and 'Other' from multi-class classifications (NDVI 1-14)

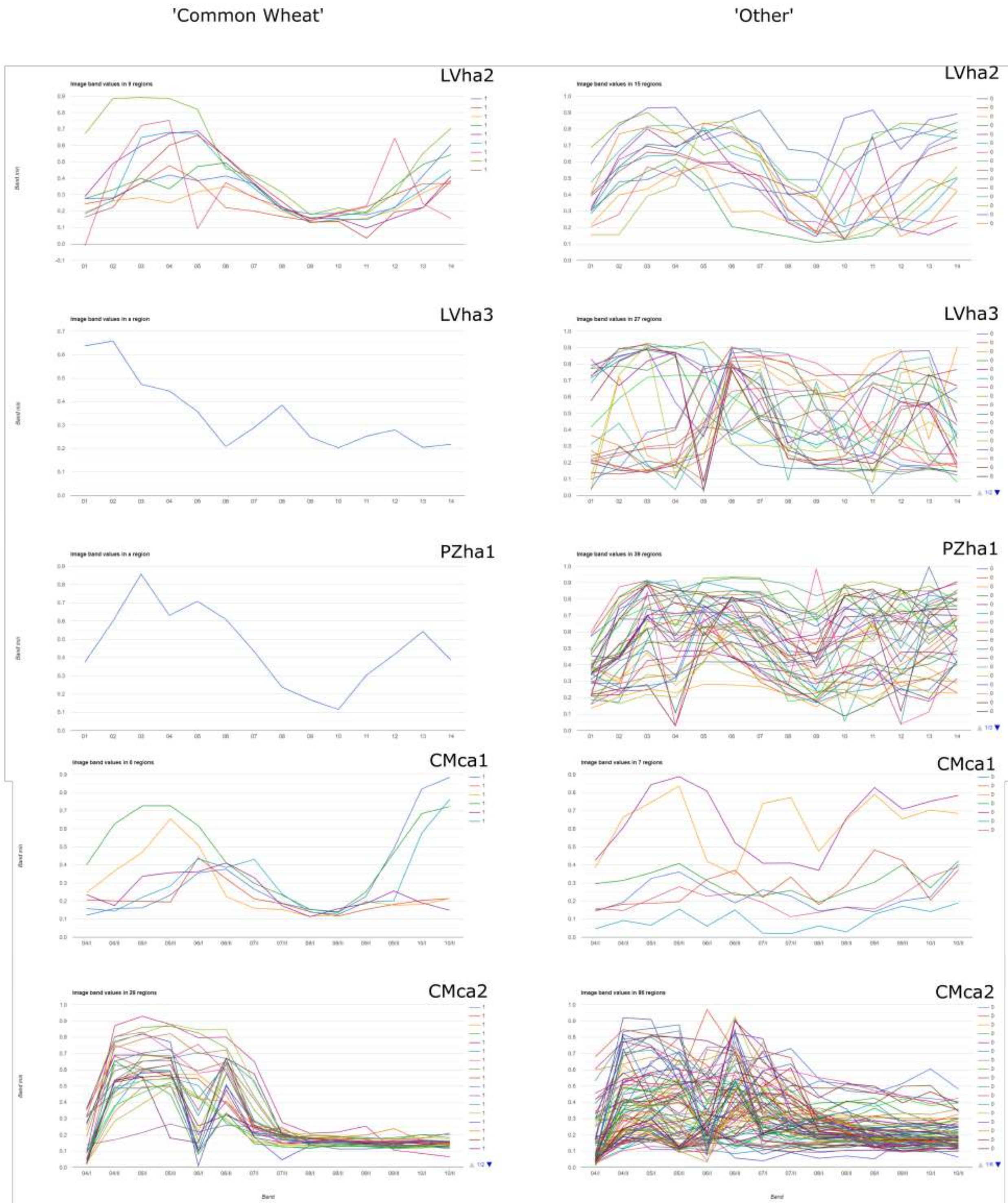

Figure 38: Part 2 - Spectral profiles of 'Common Wheat' and 'Other' from multi-class classifications (NDVI 1-14)

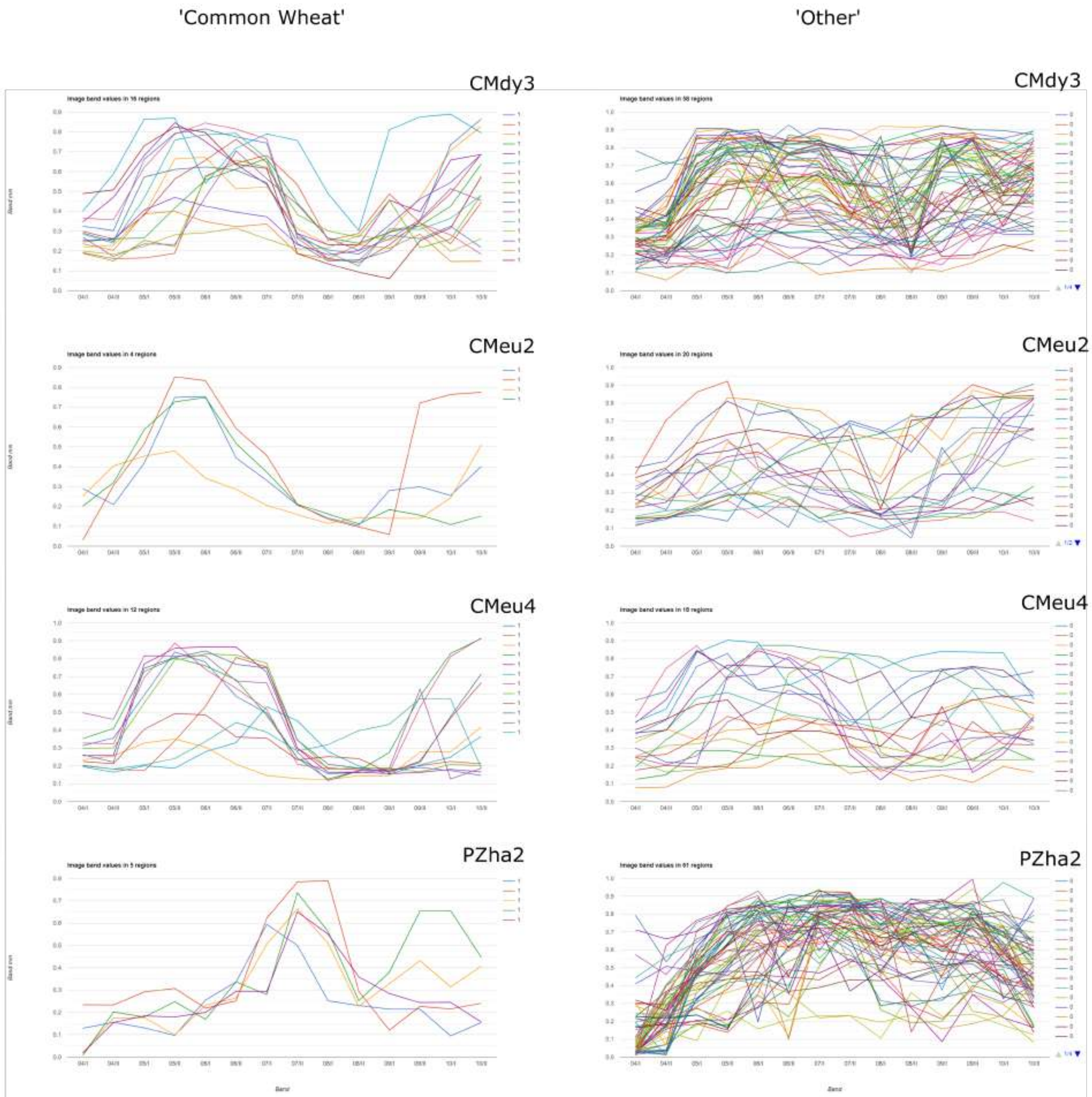

Figure 39: Part 3 -Spectral profiles of 'Common Wheat' and 'Other' from multi-class classifications (NDVI 1-14)