*Article*

# Agriculture Land Cover Identification Using One-Shot Airborne Hyperspectral Images: case study of small parcels, Poland

**Beata Hejmanowska [1] 0000-0003-0230-8386, Piotr Kramarczyk**

[1] AGH University of Science and Technology / al. Mickiewicza 30, 30-059 Krakow, Poland; galia@agh.edu.pl [1]

* Correspondence: galia@agh.edu.pl; Tel.: +48-605-061-510

Version March 18, 2023 submitted to Remote Sens.

**Abstract:** This study aimed to investigate the possibility of using one-shot hyperspectral airborne images to recognize crops for an area with many small plots. The results showed that unsupervised clustering methods could classify crops with an accuracy of 80%, which improved to 90% when restricted to only grain crops, using a single airborne hyperspectral recording. However, additional layers such as NDVI, DTM, slope, and aspect did not improve classification accuracy. For comparison, the accuracy of clustering time series Sentinel-2 images with NDVI layers and DTM-derived data yielded an accuracy of: 74%, Sentinel-2 time series 68% and single one registration before harvest 39%. The results of the random forest classification were slightly less accurate due to a lack of sufficient reference data. However, it is challenging to verify the reported accuracy of crop recognition in the literature above 90% due to differences in analysis methodologies, reference data selection, pixel/object approaches, metric choice, and calculation formulas used.

**Keywords:** airborne hyperspectral images, Sentinel-2, k-means, random forest, crop recognition

---

## 1. Introduction

Remote sensing-based land use classification in agricultural areas is not a new topic. The number of publications in this area has been steadily increasing and exceeded 1400 in 2022 Fig. 1. Machine learning techniques accounted for just over 200 of these publications. In addition, there has been a rapid growth trend in Random Forest and Deep Learning since 2015, with less growth in Support Vector Machine.

By studying this topic, one can find articles describing both various case studies and reviews. Despite the abundance of material in this area, the problem of recognizing crops or land use types in agricultural areas is not a closed, solved, or trivial problem [33]. This applies to both the choice of data and classification methods, as well as the assessment of accuracy [32].

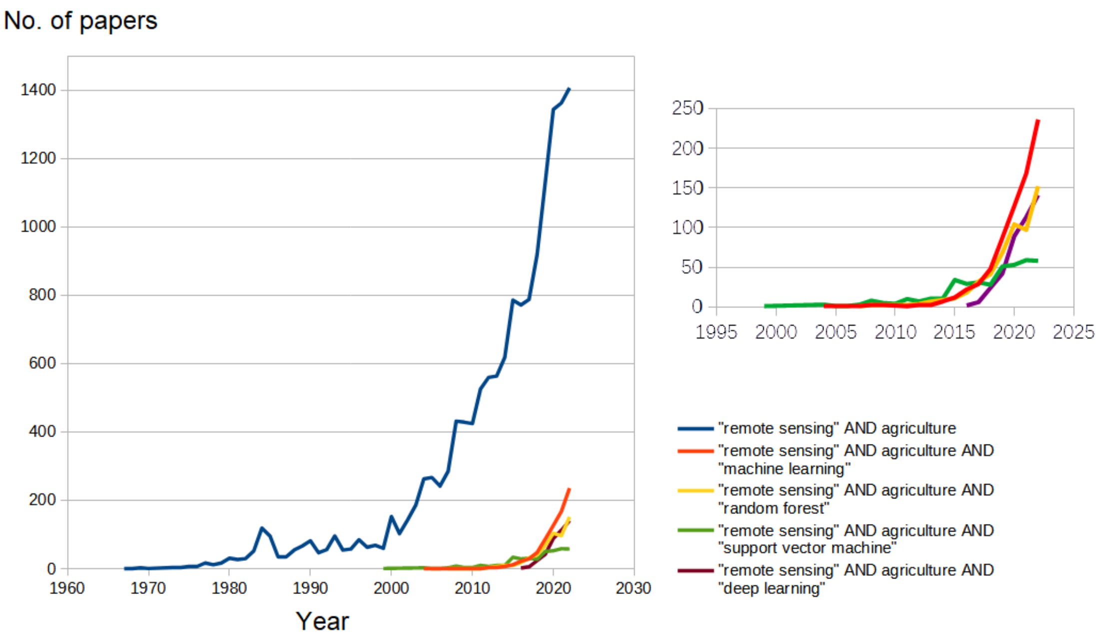


**Figure 1.** Increasing of the remote sensing applications in agriculture based on the literature, Scopus March 2023

Agricultural land cover monitoring is performed for various purposes such as yield forecasting [1–3], precision farming [4,5], and control of direct agricultural subsidies and sustainable development [6,7]. When recognizing land cover in agricultural areas, we have to deal with the delimitation of built-up areas, industrial areas and, in particular, areas of mineral extraction by the open-pit method.

The first aspect is the selection of data. In recognition of agricultural land cover, long-time series data from Sentinel-2 (S2) and Sentinel-1 (S1) covering the entire plant phenological cycle is the standard approach. Other data, such as indices calculated from the S1/S2 time series such as the Normalized Differential Vegetation Index (NDVI) [11] and radar backscattering coefficient (SIGMA) [12], as well as cadastral parcels or Digital Elevation Model (DTM), are often also included in the input data set.

The use of long time series is time-consuming and limited in many countries in the temperate climate zone, including many European countries due to cloud cover. In literature, however, this topic is rarely addressed.

Machine learning algorithms, including Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), Deep Learning (DL), and others, are used exclusively to automatically classify these large data sets. Recently, the majority of research papers on agriculture land cover recognition have focused on using either RF or SVM methods, despite the increasing popularity of DL techniques. This is due to the challenge of obtaining sufficient reference data to support deep learning models.

The selected publications that employ the RF method are presented in Table 1. The information provided includes the test region, the data utilized for classification, no of agriculture land cover (no of classes), and the Overall Accuracy (OA).

**Table 1.** RF crop classification

| Author | Test area | Features | No of classes | OA |
|---|---|---|---|---|
| Bolognesi et al. 2020 | Italy | L8, S2, NDVI | 3 | 90% |
| Hutt et al. 2020 | Germany | S1, ancillary data | 12 | 96.7%% |
| Sun et al. 2019 | China | S1, S2, L8, NDVI | 3 | 93% |
| Sun et al. 2020 | China | S1,S2 | 5 | 86.98% |
| VanTricht 2018 | Belgium | S1,S2, NDVI | 8 | 82% |

Selected studies that showcase the results of agriculture land cover classification using the SVM method are similarly summarized in 2.

**Table 2.** SVM crop classification

| Author | Test area | Features | No of classes | OA |
|---|---|---|---|---|
| Brinkhoff et al.2020 | Australia | 7 | 3 | 84.2% |
| Maponya et al. 2020 | South Africa | S2 | 5 | 82.4%% |
| Mustak et al. 2019 | India | S1, S2 | 3 | 88.94% |

In all of these studies, the input data for analysis was time series created from multiple satellite images taken at different points in time. The accuracy of the results was generally greater than 80%. It should be noted that the time series was often generated from a large number of image acquisitions, which is a prevalent trend in research using remote sensing methods in agriculture land cover recognition. This approach can be seen in such services as Sen2Agri [22] or Sen4Cap [9] dedicated to agricultural areas.

However, processing long time series is time-consuming and requires many unclouded images, which can be a challenge in temperate climates. As a result, methods based on single image acquisition are promising. Our study followed the suggestion of Maponya et.al. [19], who compared the accuracy obtained from time series versus the accuracy that could be achieved from a single image. The highest accuracy from a single image acquisition was achieved about 4 weeks before harvest, at 77.2% (compared to a maximum of 82.4% for time series). We repeated this experiment in the north of Poland, where land plots are relatively large and can be recognized in Sentinel images and obtained similar accuracy for a single registration 79% [23]. This suggests that using optical images from a single acquisition, a recognition accuracy of 80% at the plant level can be achieved for agriculture land cover. Our research was also motivated by the study presented in [24] which highlights the use of deep learning for mapping agriculture land cover during cloudy seasons with a single hyperspectral satellite image and achieved high accuracy (94%). However, the resolution of satellite images could be an issue in areas with highly fragmented agricultural structures, such as southern Poland and other regions globally. Therefore, we have established our research objective as determining the accuracy that can be obtained through the use of a single airborne hyperspectral image in recognizing land cover in areas with small plots and complex structures.

## 2. Methods and Materials

The research was carried out according to the scheme shown in Fig. 2. The activities can be divided into 3 groups: acquisition and processing of satellite images, aerial images, and acquisition of information on land cover types. Cloudless S2 images downloaded from Copernicus hub (ESA) and the SRTM numerical terrain model downloaded from the Jet Propulsion Laboratory (JPL) were used for analysis from the satellite ceiling. The result of the preparation work was an S2 stack containing processed data from ESA and JPL. From the aerial altitude, a HySpex stack was similarly prepared using hyperspectral imagery (registered by MGGP, https://www.mggpaero.com/) and Digital Tarrain Models, Digital Surface Models (DSM) from the National surveying and cartographicla service (geoportal.gov.pl). During the field work, reference data was generated, which was preprocessed and divided into training and test data. Next, image classification was performed using methods: RF and SVM with accuracy analysis based on an independent test set.

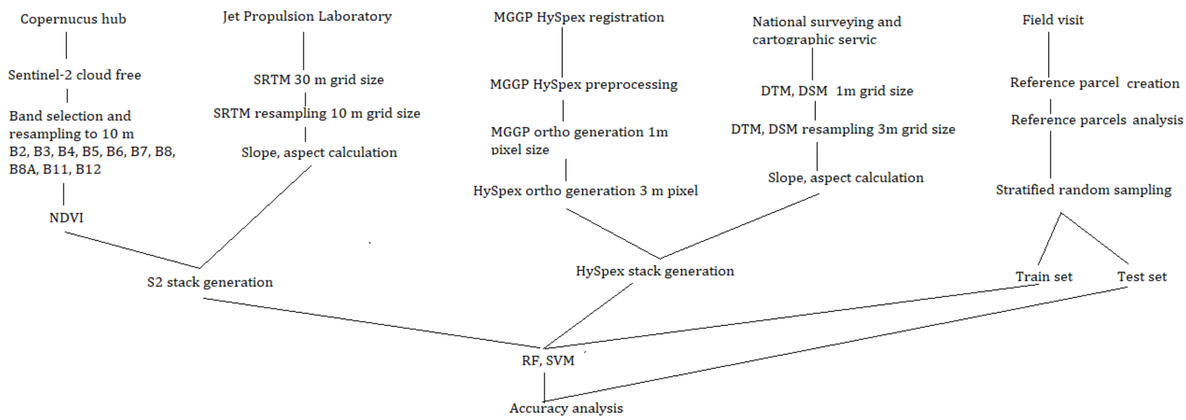


**Figure 2.** Workflow

### *2.1. Test area*

Poland is characterized by a varied agricultural landscape, with large, regularly shaped fields in the northern and central regions and small, elongated and irregular plots in the south. The use of S1/S2 imagery for crop monitoring may prove feasible for fields in the northern and central areas, but could pose difficulties for those in the south. In collaboration with the Agency for Restructuring and Modernisation of Agriculture (ARMA https://www.gov.pl/web/arimr-en) in Poland, a test area near the town of Kolbuszowa was selected as a representative sample (as shown in Figure 3). Information was obtained from ARMA on the agricultural plots that receive subsidies, with roughly 5000 such plots registered annually (as shown in Figure 4). Most of these plots are small, with 75% of the agricultural plots being less than 1 hectare in size, with a third quartile of 9499 square meters (as shown in Figure 5).

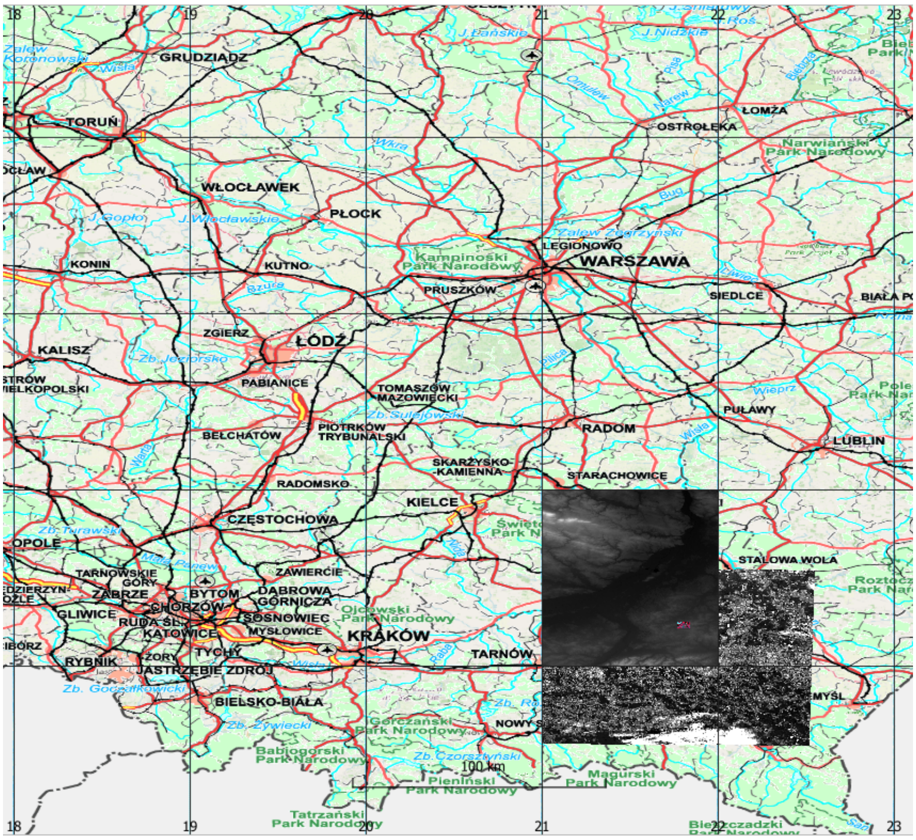


**Figure 3.** Test area location on the south of Poland, small red area in the background of S2 and SRTM, geographical coordinate system, EPSG:4326 (on topographic map https://mapy.geoportal.gov.pl/wss/service/img/guest/TOPO/MapServer/WMSServer )

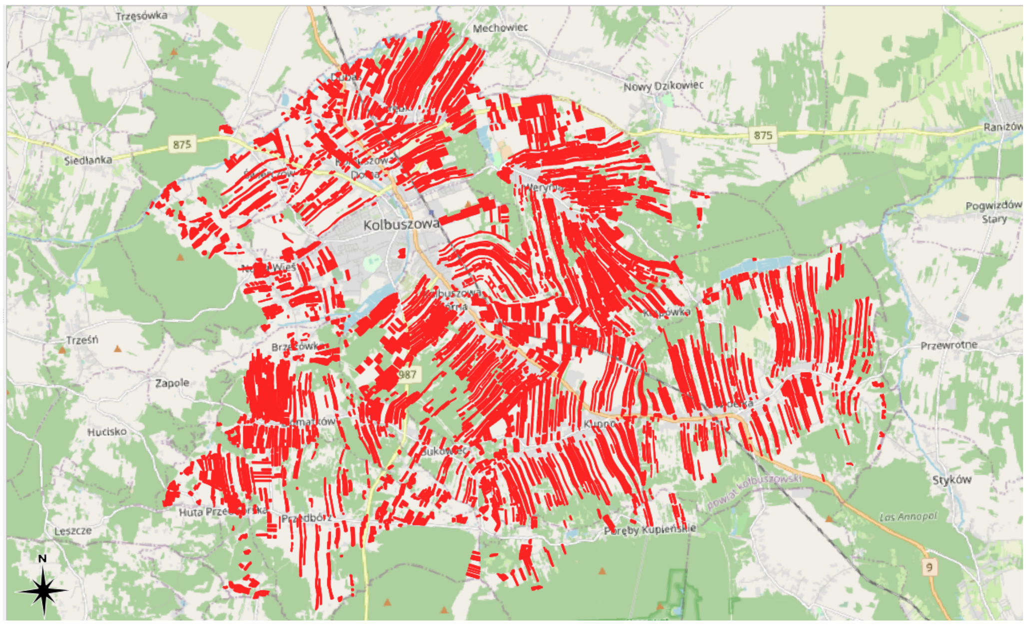


**Figure 4.** Parcels submitted for subsidies each year (thanks to https://www.gov.pl/web/arimr-en)

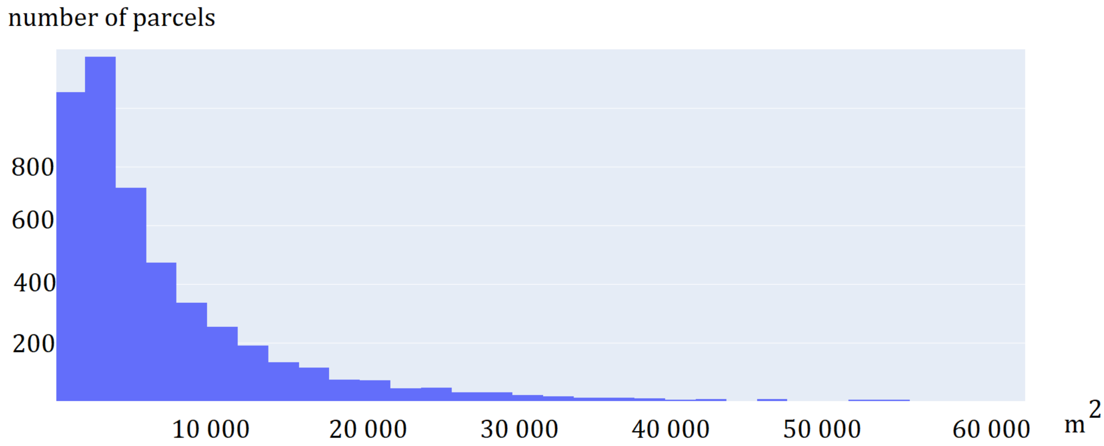


**Figure 5.** Histogram of the parcels' areas (thanks to https://www.gov.pl/web/arimr-en)

### 2.2. Data

The choice of data is dependent on its intended usage. For our research aimed at recognizing agricultural land cover types, we needed to consider the phenological stage of the crops on the agricultural plots. Based on a promising suggestion in the literature, we decided to investigate the feasibility of using data collected approximately four weeks before harvest. In Poland, the harvest season typically lasts for 2 months (July and August), starting with the small harvest of rapeseed and winter barley, followed by the large harvest of rye, spring barley, wheat, and oats [25]. To examine the potential of using a single registration for agricultural land cover recognition, a survey campaign was conducted in 2021 that collected the following data: S2 time series, one-shot hyperspectral, in-situ measurement, and topographical data. The data acquisition dates are in Table 3.

**Table 3.** Data

| Data | Date |
|---|---|
| Sentinel-2 | March-September |
| Hyperspectral | 5th of July 2021 |
| In-situ measurements | 7 July 2021 |
| SRTM (jpl.nasa.gov) | 2000 |
| TOPO (geoportal.gov.pl) | 2017 |

During the 2021 growing season, only six Sentinel-2 registration dates were cloud-free Table 4.

**Table 4.** S2_image

| Data | S2 ID |
|---|---|
| 27.03.2021 | S2B_MSIL2A_20210327T093039_N0214_R136_T34UEA_20210327T120034 |
| 11.04.2021 | S2A_MSIL2A_20210411T093031_N0300_R136_T34UEA_20210411T122810 |
| 09.05.2021 | S2B_MSIL2A_20210509T094029_N0300_R036_T34UEA_20210509T120133 |
| 25.07.2021 | S2B_MSIL2A_20210725T093039_N0301_R136_T34UEA_20210725T115620 |
| 28.07.2021 | S2B_MSIL2A_20210728T094029_N0301_R036_T34UEA_20210728T125908 |
| 06.09.2021 | S2B_MSIL2A_20210906T094029_N0301_R036_T34UEA_20210906T113414 |

Sentinel-2 images were acquired from Copernicus Open Access Hub as granules with a size of 100 per 100 km with a radiometric correction level of 2A in geographical coordinate system EPSG:4326. The images were not further corrected either geometric or radiometric. The pixel size depending on the channel is 10, 20 and 60m. A single S2 scene in SAFE (ESA) format takes approximately 1.2 gigabytes when packed (S2 range in shown in Fig. 3).

Hyperspectral data were acquired for the area ca. 5 x 4 km using HySpex VS-725 which is a very small area compared to the S2 range (in red in Fig. 3) . The registration was performed at an altitude of 867 - 882 m. The HySpex VS-725 consists of two SWIR-384 scanners and one VNIR-1800 scanner which provide 430 spectral channels (414.13 nm - 2357.43 nm). The test area was covered with 16 strips. Radiometric, geometric (PARGE), atmospheric (ATCOR4) correction was performed using the MODTRAN physical model. The final product, an orthophotomap with a pixel size of 0.5 m was registered in the UTM 34N coordinate system (EPSG:32634) and takes up about 60 gigabytes.

DTM and DSM (Digital Surface Model) were obtained from the national server: geoportal.gov.pl. Three DSM sheets (about 150 megabytes) and 15 DTM sheets (about 160 megabytes) with a pixel size of 1m.

A field visit was conducted to obtain information about the ground truth (the plant that was 2021 grown on the agriculture parcels). Information of the location was acquired using handheld GPS. In the field 10 agricultural land cover types was recognized: beet, soil, barley, maize, oats, wheat rye, wheat winter, grass, potato and rye. Soil, oats and grassland is easy to delimitation. Wheat, rye and wheat rye are important cereal crops with a rich content of starch, protein, and other nutrients (https://eos.com/products/crop-monitoring/crops/). Wheat is more commonly cultivated and used for the production of bread, pasta, and other food products, while rye is more resistant to low temperatures and used in bread production under difficult climatic conditions, as well as for animal feed and alcohol production. Wheat has awnless spikes, while rye spikes have characteristic awns, although not as long as those of barley. In the field, it is easy to distinguish wheat, rye, and barley. Triticale (wheat rye) is a hybrid of wheat and rye, obtained by crossing these two plants. Triticale has characteristics of both wheat and rye, making it more resistant to adverse weather conditions and having higher nutritional value. Triticale is mainly used as a fodder crop for animals and as a crop to produce grain for flour. In the field, it is difficult to recognize triticale from rye, and often the decision must be consulted with the farmer who sowed the grain. Some types of land cover occurred on single plots, so it was decided to omit them from further analysis. In the end, it was resolved to reduce the number of classes to as in Table 5. The spatial distribution of the reference vectors can be seen in Fig. 6 and 7.

**Table 5.** Crops

| Id | Crop |
|---|---|
| 2 | soil |
| 5 | oats |
| 6 | wheat rye |
| 7 | wheat winter |
| 8 | grass |

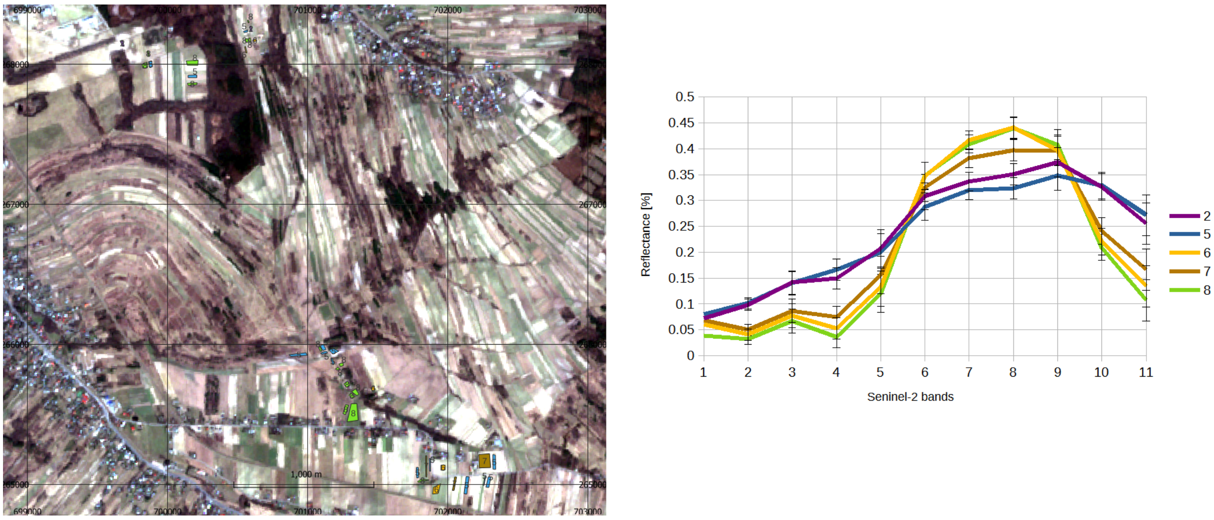


**Figure 6.** Parcels visited in the field with average spectral curves, S2 RGB - 27.03.2021 in background


**Figure 7.** Parcels visited in the field, zoom-in, left - north part, right - south part

## 2.3. Data preprocessing

Six Sentinel-2 images were subset and resampled into 10 m. From each Sentinel-2 set 10 bands were selected (B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12) for area of interest (AOI): 1295 columns x 922 rows (UL: 552250, 5567620 ; LR: 5558400, 5565200 ; EPSG:32634). The channels of all the images prepared in this way were saved in a single TIF file. In addition, NDVI was calculated for each registration date and added to the above mentioned TIF file. Image classification also uses other data that can increase the accuracy of classification, such as numerical terrain models and the slopes/aspects calculated from them. So SRTM was acquired, cropped to the AOI, and resampled to 10m. SRTM and calculated: slope and aspect were added to the TIF file as well. This means that the S2 time series stack consists of 60 Sentinel-2 channels, 6 NDVI images and 3 images containing topography information of the area.

The original Sentinel-2 images are recorded on 12 bits and stored on 16 bits as uint16 (in the metadata there is a size by which DN should be divided to calculate the reflection coefficient: 0-1 as a float32 number, which is 10,000). The values of NDVI coefficients change from -1 to 1 and were stored as float32 numbers.

The SRTM layer is of float32 type and includes for AOI values in the range of 230-310m, slope and aspect are also float 32 type and include values in the range of: 0 to 26 degrees, and 0 to 360 degrees. For the purposes of machine learning, all layers were scaled to a range: 0.0 - 1.0.

Due to the large size of the hyperspectral image, it was cropped of the area where the field visit was conducted and resampled to a pixel size of 1m and 3m. Numerical terrain models were merged and clipped to the extent of the hyperspectral image. In addition, slopes and aspects were calculated from the DTM and NDVI from hiperspectral channels. All rasters were merged into a single TIFF file (5340 cols x 6840 rows, UL: 557062, 5566510 ; LR: 559732, 5563090, 430 bands, NDVI, DTM, DSM, slope and aspect).

The hyperspectral mosaic (9484 rows x 7478 cols, UL: 554995, 5566821 ; LR: 559737, 5563082) made from the processed hyperspectral images has a spatial resolution of 0.5 m, consists of 430 spectral channels (414.13 nm - 2357.43 nm) is registered in the UTM 34N coordinate system (EPSG:32634) and takes up about 60 gigabytes

### 2.4. Methods

The image data, numerical terrain models and their derivatives were merged using own code in Python as a stack and saved as a single tif file. Separately, one file from the Sentinel-2 time series, at 10m resolution, and one file with hyperspectral data at 3m resolution (the original HySpex 0.5 data was resampled to 3m). There are 68 layers (bands) in the Sentinel-2 time series stack file, Table 6. From 1-60 Sentinel-2 channels, 61 to 66 NDVI for each date (the channels used for calculation are also given), 67-68 DTM, aspect and slope. There are 435 layers (bands) in the Hyperspectral stack file Table 7. From 1-430 HySpex channels, 431-434 DTM, DSM aspect and slope, 435 NDVI (the channels used for calculations are also given).

**Table 6.** Sentinel-2 time series stack

| Band number | Details |
|---|---|
| 1-10 | B0327 |
| 11-20 | B0411 |
| 21-30 | B0509 |
| 31-40 | B0725 |
| 41-50 | B0728 |
| 51-60 | B0906 |
| 61 | 0327_7/3 |
| 62 | 0411_17/13 |
| 63 | 0509_47/43 |
| 64 | 0725_57/53 |
| 65 | 0728_37/33 |
| 66 | 0906_27/23 |
| 67 | DTM |
| 68 | aspect |
| 68 | slope |

**Table 7.** Hyperspectral stack

| Band number | Details |
|---|---|
| 1-430 | HySpex bands |
| 431 | DTM |
| 432 | DSM |
| 433 | slope |
| 434 | aspect |
| 435 | ndvi_142_80 |

Image processing was carried out using custom scripts in Python and free plug-ins for QGIS (EnMAP-Box and GRASS). Automatic classification was performed in an unsupervised (K-means method) and supervised (Random Forest method) manner. RF classification accuracy analysis was analyzed by k-fold cross validation in EnMAP-Box. Analysis of the accuracy of the final classification result was done on independent test fields in GRASS.

Clustering in EnMAP proceeds in two stages: FitKMeans and Predict Clustering

FitKMeans in EnMAP-Box is executed using the following script:

$from - sklearn.pipeline - import - make\_pipeline$

$from - sklearn.preprocessing - import - StandardScaler$

$from - sklearn.cluster - import - KMeans$

$clusterer = KMeans()$

$KMeans(n_clusters = 8, *, init =' k - means + +', n_init = 10, max_iter = 300,$

$tol = 0.0001, verbose = 0, random_state = None, copy_x = True, algorithm =' lloyd')$

$estimator = make\_pipeline(StandardScaler(), clusterer)$

$outEstimator.pkl$

Argument *init* of class *sklearn.cluster.KMeans* - "'k-means++' : selects initial cluster centroids using sampling based on an empirical probability distribution of the points' contribution to the overall inertia" (scikit-learn 1.1.2). Number of clusters *n_clusters* can be modified (default = 8).

Another issue is standardize features by removing the mean and scaling to unit variance. The standard score of a sample x is calculated as: z = (x - u) / s where u is the mean of the training samples or zero if with_mean=False, and s is the standard deviation of the training samples or one if with_std=False (scikit-learn 1.1.2).

The second step performs *PredictClustering*, which applies a *clusterer* to a raster.

The resulting clusters were mapped to crop type and analyzed in the context of reference fields to evaluate the effectiveness of the method.

Random Forest in EnMAP in classic approach proceeds using the following script:

$fromsklearn.ensembleimportRandomForestClassifierestimator = RandomForestClassifier(n\_estimators = 100, oob\_score = True)$

The set of reference parcels was divided into two separate sets using stratified random sampling. The learning process was based on the training set. The model's accuracy, known as validation accuracy, was determined using k-fold cross-validation (defaulting to 3 folds). To select the optimal hyperparameters for our Random Forest (RF) model, we utilized the widely used grid search method through scikit-learn's Grid Search CV class. Our training phase employed three evaluation metrics - accuracy, balanced accuracy (mean recall), and f1-weighted (weighted average of precision and recall) - to measure the model's performance. After testing various configurations using 10-fold cross-validation, we ultimately settled on the following classification settings.

- classification:scikit-learnlibrary, sklearn.ensemblemodule, RandomForestClassifier,
- number of trees: 100,
- min_samples_split': 2,
- min_samples_lef: 2,
- bootstrap: True,
- max_depth: None,
- max_features: None.

In the next stage, we conducted an accuracy analysis based on a test set that was not involved in the learning process. Although accuracy analysis on independent test fields is available in EnMAP-Box, the *r.kappa* function from the GRASS plugin was used for technical reasons. Function "*r.kappa* tabulates the error matrix of classification result by crossing classified map layer with respect to reference map layer. Both overall kappa (accompanied by its variance) and conditional kappa values are calculated. This analysis program respects the current geographic region and mask settings" (https://ibiblio.org/pub/packages/gis/grass/grass63/manuals/html63_user/r.kappa.html). In this manner, pixel accuracy was determined. In addition, accuracy was analyzed at the plot level by performing an automatic majority class extraction for each polygon (QGIS-Processing tools-Zonal statistic).

Even though researchers have investigated different metrics over the years [26–31]), the most commonly recommended metric remains Overall Accuracy (OA) [32], which is defined as OA = TP / (TN + FP + FN) [28]. Therefore, in our study, we limited our analysis to OA.

The study aimed to compare the accuracy of classification results from Sentinel-2 and hyperspectral images captured at significantly different altitudes (10 m for satellites and 1-3 m for aerial images). Processing of single registrations from both types of images as well as time series of Sentinel-2 was conducted to achieve this goal. The complete dataset (stack) was used for classification, and the accuracy was compared to the classification results obtained by excluding NDVI, DEM, and their processing from the remote sensing data.

## 3. Results

Before presenting the classification results, we would like to showcase the crop structure in the test area, as seen in Figure A1. The plots have a complex and intricate cultivation pattern, with small, elongated, and often irregular shapes. A clearer view can be seen in the False Color Composite (FCC = B8A, B4, B3) of the image captured on July 28th, 2021, just before the first harvest, as shown in Figure A2. The image clearly displays green vegetation (represented in red), buildings and bare soil (in cyan), fields of mature crops (in dark green), forests, and water. The stacks were created according to the table, and in the attachment there is an example: composition in false colors (FCC) of the first S-2 image, NDVI, DTM, Slope and Aspect (Figures A3 to A6). HySpex_stack data is similarly provided.

Further in the attachment are the results of clustering and classification, with details on the calculation of accuracy.

The first part of the results presents a mask showing the areas excluded from the analyses. The second part contains the results obtained using the K-means method, and the third part presents the results from RF method.

### *3.1. Mask preparation*

The best clustering results were obtained from the S2 stack, as shown in Figure A9. The results from clustering other datasets can be found in Appendix A (Figures A7 and A8). Figure A7 shows correctly classified green vegetation (green) and crops in some fields (yellow). However, a phenomenon of indistinguishability between built-up areas and bare soil (white and gray) can always be noted for a single registration date.

The clustering of the time series, both in the S2 stack (Figure A9) and the S2 time series (Figure A8), allows for the separation of built-up areas (gray), industrial areas (violet), and bare soils (brown). Additionally, mature cereals can be differentiated from green vegetation in the fields (yellow and green). Clustering the time series of remote sensing data alone, however, provides better differentiation within the green vegetation (represented in light green and cyan in Figure A9).

From the clustering results, we extracted classes to mask the areas not analyzed further for crop recognition.

- residential and industrial areas (Fig. A9 class 4 and 3),
- forests (Fig. A9, class 1 and 6)
- bare soils mixed with industrial areas (Fig. A7, Class 8)

In creating the mask, we analyzed the effectiveness of separating industrial areas, including open-pit mineral extraction from developed areas and bare soils. Just across the road to the south of the area shown in Fig. 11 et seq. is the Wienerberger Kupno open pit mineral extraction plant. Only the class covering industrial areas is shown in Fig. 8, the rest of the classes are transparent. The red color shows the separations made from one registration S20728. The separation on the S2 stack subsection covers the S20728 separation and slightly extends outside. To compare industrial areas on entire image see in Fig. A7, A8 and A9) in violet color).


**Figure 8.** Open-pit mineral extraction (geoportal.gov.pl)

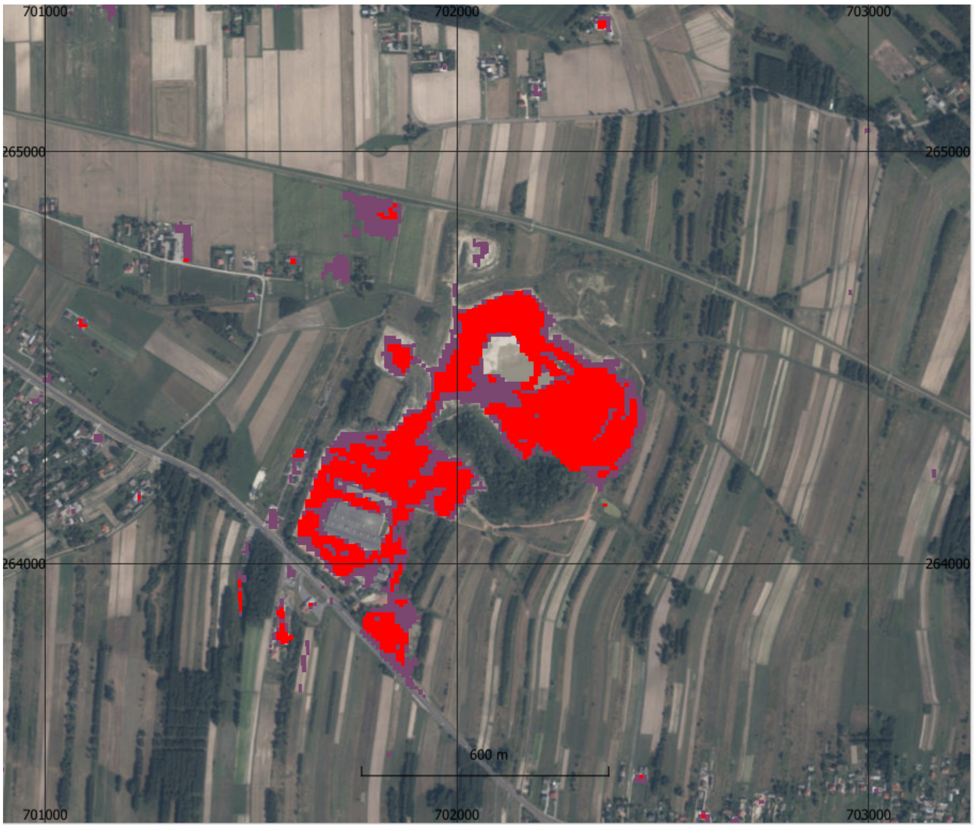


**Figure 9.** Open-pit mineral extraction - distinguished form urban and bare soils, clustering comparison S2 stack and one shot before harvest S2 0728

### 3.2. Clustering

Clustering numerates classes in automatic way, so it is necessary to map them to the ground classes. There is interpretation of the classes in the description of each figure, however the palettes are unified. Zoom of the best result using the K-means method for S2 is with the mask presented in Fig. 10 (entire image - in Fig. A10). Clustering S2 resulted in the separation of mature cereals (two classes: 3, 4 - yellow, olive) and grass (2 - grass). The bare soil in the area shown in the figure Fig. 10 was generalized within the cereal crop. Total data on separated classes in different S2 variants can be found in Table A1, (best score for S2 stack OA=0.74). Grain classes: 5 - oats as 4, 6 - wheat rye as 3 and 8 - grass were separated well. Errors in classification appeared for class: 2 - bare soil and 7 - wheat winter.

For hyperspectral data, the best results were obtained when the spatial resolution was reduced to 3m, as shown in Figures A18 and 14. The classification resulted in a separation within the cereal class: wheat rye and wheat winter. The bare soil within the crop classes was also distinguished. This is due to the significantly higher spatial resolution (3 m instead of 10 m). Total data on separated classes in different HySex variants can be found in Table A3, (best score for HySpex stack OA=0.81). Grain classes: 2 - soil as 2, 5 - oats as 1 and 8 - grass as 3 were separated well. Unfortunately class 6 - wheat rye was mixed with 7 - wheat winter as 6.

The results of the clustering accuracy analysis are presented in Figure 12, Detailed results of the accuracy analysis are provided in Appendix A, including Tables A1, A2 for Sentinel-2 data, and Tables A3, A4 for HySpex data, along with Figures A20 and A21.

The highest clustering accuracy was obtained for HySpex 3m with a score of 0.81, and adding additional data did not increase accuracy (Figure 12 blue bars). However, the accuracy of HySpex 1m clustering was low, at 0.55. The maximum accuracy from Sentinel-2 was obtained for the S2 stack, which scored 0.74, for the S2 time series 68% and for single registration S2 0728 39%. For grain crops excluding bare soil and corn (Figure 12 red bars), the accuracy was above 0.90, with scores of 0.93 for the S2 stack and 0.97 for HySpex 3m. The accuracy of clustering was determined for all reference objects/parcels as no teaching data are required in this case.

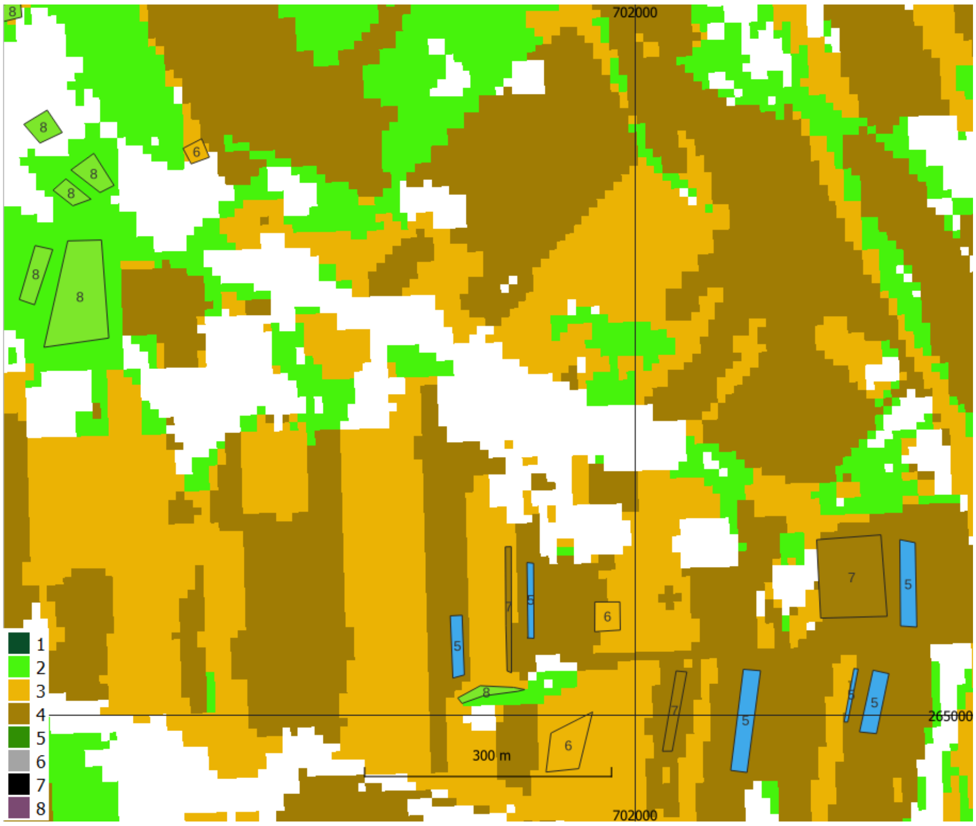


**Figure 10.** Cluster S2 stack masked; 1-conifer forest (masked), 2-grass, 3-maturing/mature cereals, 4-maturing/mature cereals, 5-deciduous forest (masked), 6-urban (masked), 7-residual class (no visible), 8-industrial (masked)

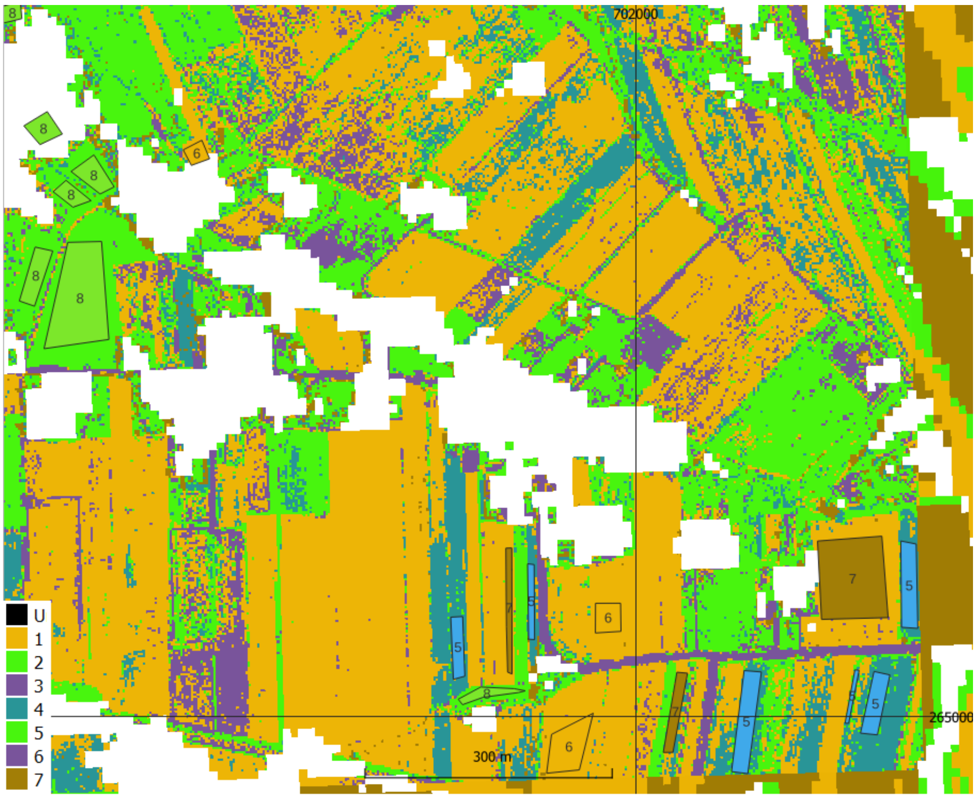


**Figure 11.** HySex 3m K-means, 1/7 - wheat rye, 2/5 - grass, 3/6 - soil, 4 - wheat winter

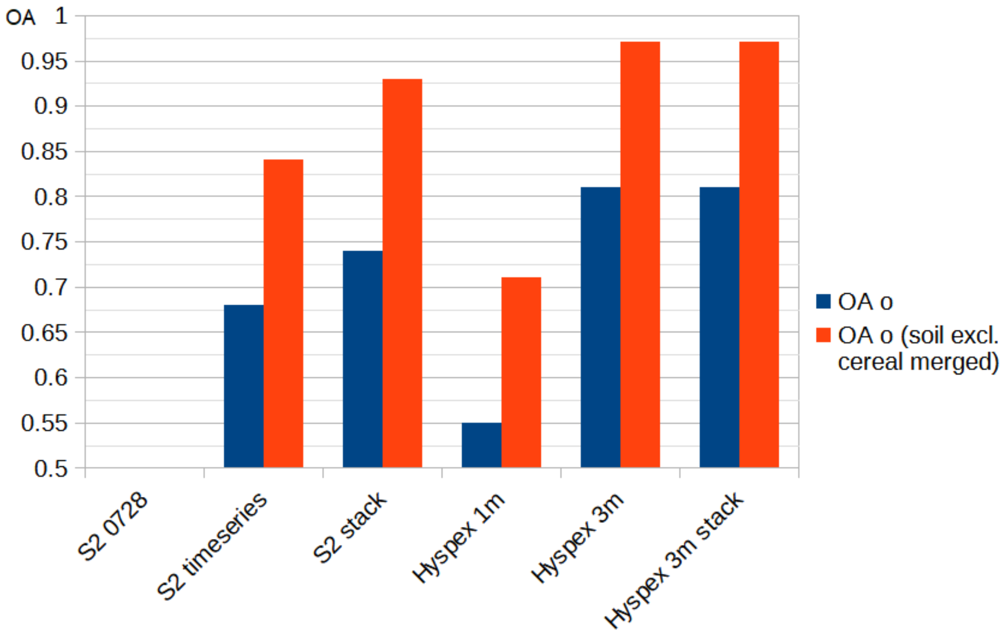


**Figure 12.** Comparison of clustering accuracy - object level, all reference data

### 3.3. Random Forest

Details of RF classification results can be found in Attachment. In Fig. 13 and Fig. 14 are enlargements of the best results obtained from the satellite and airborne ceiling. As a result of supervised classification, the class numbers are consistent with the identifiers used for learning (Tab. 5). In addition, the same palette was adopted for visualization. Visually, the classification results obtained by the RF method appear to be better than the K-means results. Due to the spatial resolution, more detail can be observed in the aerial images. At the same time, generalization is observed on satellite results. The generalization of exposed soils that occurred on the clustering results was repeated. However, analyzing the aggregate results in the Tables, it can be seen that the RF method yielded worse results than the clustering method. In the case of the RF method, unlike the clustering method, it is possible to analyze the results of method validation, from training data and the accuracy of the method on test data.

The RF method had a greater difference in accuracy between the validation and testing for Sentinel-2 images than for hyperspectral data (Figure 15). The validation accuracy on the S2 stack was 0.97 and the accuracy on the test pixels was 0.70, while the validation accuracy for HySpex 3m was 0.85 and the accuracy on the test pixels was 0.73. The object accuracy was slightly less than the pixel accuracy for the S2 stack (0.69), but greater than the pixel accuracy for HySpex 3m (0.75). The effectiveness of the RF method in correctly classifying grain crops excluding bare soil and cereal merged 100%.

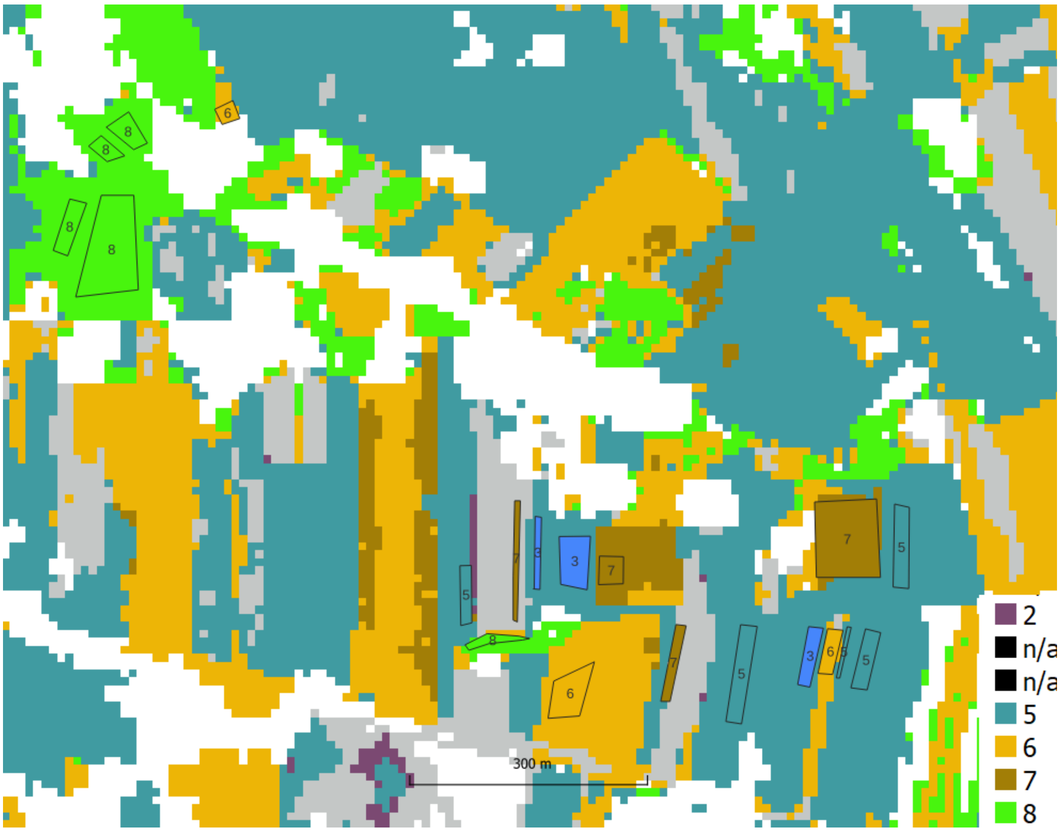


**Figure 13.** RF S2 stack masked (zoom); legend and polygon labels - according Table 5, 2-soil, 5-oats, 6-winter rye, 7- wheat winter, 8-grass

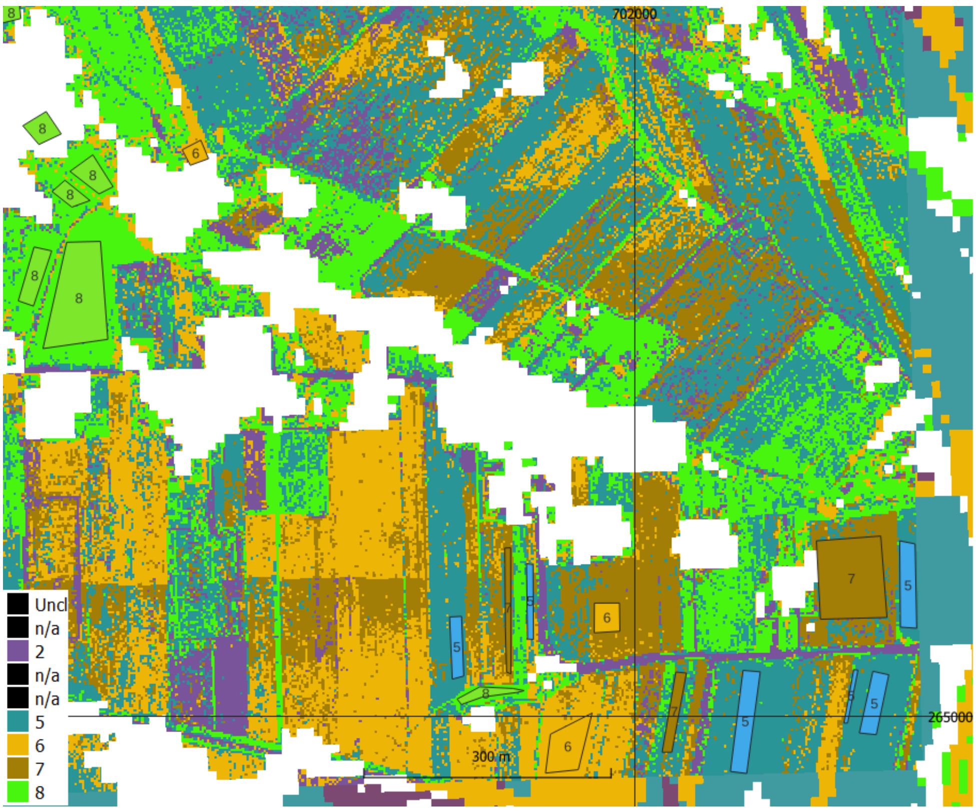


**Figure 14.** RF HySpex 3m

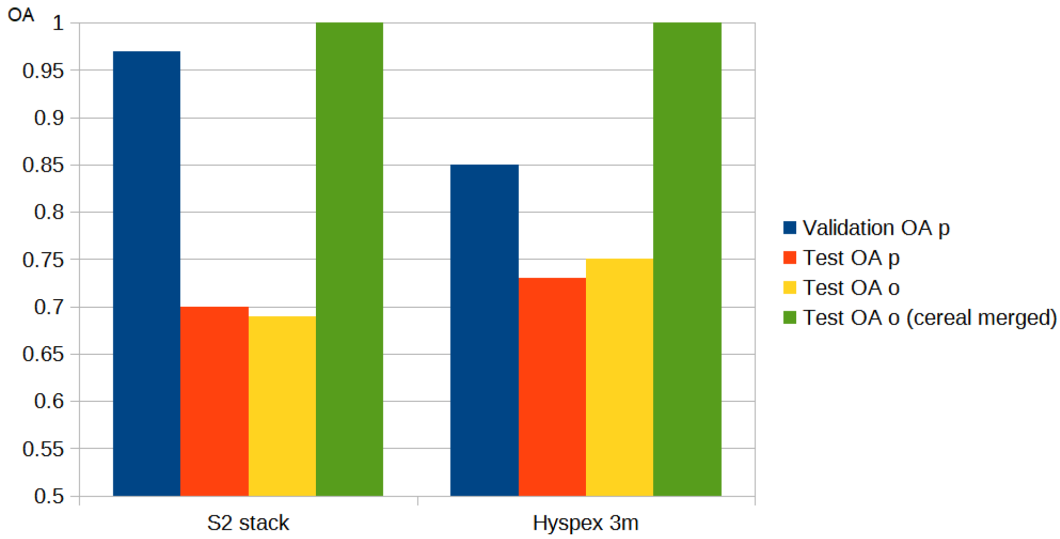


**Figure 15.** Comparison of overall accuracy obtained in RF classification; pixel (OA p ); object (OA o)

In summary of the above described results, it can be concluded that:

- With a single airborne hyperspectral recording, it was possible to classify crops at 80% accuracy using unsupervised clustering methods. When restricted to only grain crops, accuracy improved to 90%.
- Additional layers (NDVI, DTM, slope and aspect) did not increase the classification accuracy of aerial hyperspectral images.
- The accuracy of S2 stack (S2 Sentinel-2 time series plus NDVI, DTM, slope, and aspect) clustering was relatively high - 74%, especially for an area with a large number of small plots; in comparison, the accuracy of clustering only S2 time series was 68%.
- The accuracy of a single Sentinel-2 recording was surprisingly low, at less than 50%. The reason for this discrepancy is unclear, but it may be related to differences in crop structure in northern Poland where a one-shot FR S2 test had an accuracy of approximately 80% ([23]).
- The results of the random forest classification were slightly less accurate due to a lack of enough reference data. Clustering methods did not require training data, while random forest methods required dividing reference data into learning and test sets.
- The accuracy of crop recognition reported in the literature above 90% is difficult to verify due to differences in accuracy analysis methodologies, reference data selection, pixel/object approaches, metric choice, and calculation formula used.

## 4. Discussion

When comparing classification results, three aspects should be taken into account:

- which metric was chosen to assess accuracy
- whether the results concern method validation or testing on independent test data not used for learning
- whether the reference data was divided in a way that prevents data correlation

### *4.1. Metrics*

Often, articles present the learning process and report only the validation accuracy, which is always higher than the accuracy obtained on independent test data. K-fold cross validation is used to analyze learning effectiveness and results in higher validation accuracy compared to independent test data [33]. Reference data is also often divided in a way that falsely increases accuracy, such as selecting pixels from the same plot for the training and testing sets, which are correlated with each other. To avoid this issue in agriculture land cover recognition, reference set separation should be made at the plot level, not the pixel level. Therefore, for testing, plots that have never been seen during learning should be selected. Finally, it is sometimes reported as OA values that are actually ACC values. Machine learning typically uses four classification results types (TP, TN, FP, FN) and metrics like sensitivity, specificity, and accuracy (ACC=(TP+TN)/(FP+FN+TP+TN) [34]), which is only equal to OA in one-class classification. In multi-class classification, OA is calculated as TP/(TN+FP+FN) [28], while ACC results in much higher accuracy estimates than OA. Examples can be found in journals of both proper and improper use of ACC for classifying features [35,36].

Our article was influenced by a publication [24] that explored the use of one-shot hyperspectral satellite imagery compared to multispectral time series for crop recognition. The authors reported an accuracy of 94%. However, it is important to keep in mind that the accuracy was calculated using OA=(TP+TN)/(TP+TN+FP+FN), which is the de facto ACC accuracy, and only two crops, winter wheat and rapeseed, were tested. This calculation method gives a higher value for ACC compared to OA for many classes OA=TP/(TP+TN+FP+FN) as it takes into account both TPs and TNs. If we compiled from Table A4 confusion matrix we could calculate, in addition to OA (0.75), producer accuracy (PA) and user accuracy (UA). It is also possible to calculate the ACC for each class, it is the same in the ACC column and ACC row.

It is worth noting that all ACC values are high (greater then 0.80), even if PA(7 - wheat winter)=0.00, ACC=0.81. This is due to the fact that there are a large number of TN=13 (only 3 FN cases).

In the case in discussion, we gave an accuracy of 0.75 (as OA), if we had counted the average ACC we would have reported 0.90.

**Table 8.** Confusion matrix based on Table A4

| | | **Test** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 2 | 5 | 6 | 7 | 8 | PA | ACC |
| Ref | 2 | 3 | 0 | 0 | 0 | 0 | 1.00 | 1.00 |
| | 5 | 0 | 5 | 1 | 2 | 0 | 0.63 | 0.81 |
| | 6 | 0 | 0 | 1 | 0 | 0 | 1.00 | 0.88 |
| | 7 | 0 | 0 | 1 | 0 | 0 | 0.00 | **0.81** |
| | 8 | 0 | 0 | 0 | 0 | 3 | 1.00 | 1.00 |
| | UA | 1.00 | 1.00 | 0.33 | 0.00 | 1.00 | **0.75** | |
| ACC | | 1.00 | 0.81 | 0.88 | **0.81** | 1.00 | | **0.90** |

### *4.2. Validation, testing sets*

[37] classified 6 classes with the correct calculation of accuracy, resulting in 95.85% for SVM and an increase in accuracy using deep learning: PCA = 8, epoch = 30 - 97.1%; PCA = 16, epoch = 30 - 98%; and PCA = 24, epoch = 30 - 98.6%. However, the reference data was divided at the pixel level, meaning

only selected test pixels were analyzed, which makes it difficult to assess the reported accuracy. As the authors mentioned, potatoes were the worst misclassified crop with a total of 32 pixels misclassified (19 too few and 13 too many). The producer accuracy was 93.04% and user accuracy was 95.15%. However, looking at Fig. 16 [37], there is a mismatch between the reference plot with potatoes and the SVM classification results, making it challenging to trust the accuracy based on pixel analysis.

In the reference set (in Fig. 16 in the middle), only one plot is covered with potatoes. As a result of the classification (in Fig. 16) right), many plots are classified as potatoes. Mainly it is the class (others), which the authors write about as not in agricultural use. In our case, this type of land cover was either masked or classified as bare soil and taken in whole for accuracy analysis.

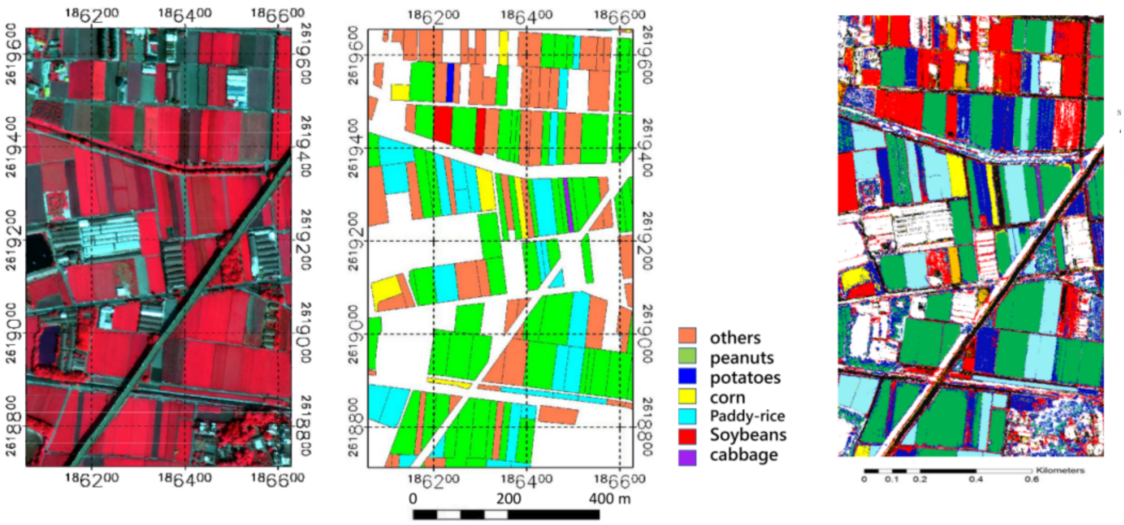


**Figure 16.** FCC from hyperspectral image, reference plots and result of SVM classification [37]

### 4.3. Reference to other comparable works

The classification accuracy of agricultural land cover types, calculated as OA using uncorrelated training and test sets, typically falls between 80-90%, sometimes even higher, especially when dealing with time series data (Table 1 and 2). While most publications focus on time series, some suggest using a data reduction approach to classification, such as a single registration [23]. The accuracy obtained in this scenario is significantly impacted by the registration deadline. Maponya et.al. suggest registering four weeks before harvest to achieve 77% accuracy, which is just 5% lower than the accuracy achieved using time series data. Despite this recommendation, we were unable to attain such accuracy using a single Sentinel-2 registration. The S2 stack accuracy reached only 74%, and a single registration resulted in an accuracy of less than 50% in our study.

This may be due to the crop structure specific to the test area, as our earlier research [23] in northern Poland, with larger plots, yielded more optimistic results that were even slightly better than reported by Maponya et al.

It also seems that the proposed registration date is appropriate, as aerial imagery allowed us to achieve an accuracy of 81%.

## 5. Conclusions

Prior to making any conclusions, it is crucial to consider the role of metrics in evaluating the precision of image classification. One commonly employed metric is accuracy, which plays a significant role in its computation. Traditionally, accuracy in remote sensing is measured using OA, while machine learning employs ACC. ACC is always higher than OA since it incorporates both true positive and true negative cases.

In our analysis, we obtained the highest accuracy 0.81 (calculated as OA) but we could report 0.96 or even 0.98 if we use ACC. Therefore, in any comparative study, it is crucial to carefully analyze the metrics used to calculate the classification accuracy. This is especially important now, as there are numerous publications presenting various machine learning models with high reported accuracies.

Regarding the primary objective of the study, which was to evaluate the accuracy of one-shot registration for agricultural land cover mapping, the research found that an accuracy of 80% can be achieved using airborne hyperspectral data. It is also recommended to perform the registration about four weeks before harvest, as confirmed by the research [19].

Undoubtedly, the future belongs to machine learning, including deep learning. The practical use of such models will be possible if a very large amount of training data is provided, which may pose a certain problem. However, it is always necessary to remember about the proper assessment of accuracy.

**Author Contributions:** Conceptualization, B.H. and P.M.; methodology, B.H. and PK; software, P.K.; validation, P.K.; investigation, P.K. and B.H.; writing—original draft preparation P.K.; writing—review and editing B.H.; visualization, P.K.; supervision, B.H.; funding acquisition BH All authors have read and agreed to the published version of the manuscript.

**Funding:** Research project supported by program Excellence initiative—research university for the AGH University of Science and Technology no. 501.696.7996. The project title: Integration of remote sensing data for control in the system of direct agricultural subsidies (IACS).

**Acknowledgments:** The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| RF | Random Forest |
| SVM | Supported Vector Machine |
| CNN | Convolution Neural Network |
| DL | Deep Learning |
| TP | True Positive |
| TN | True Negative |
| FP | False Positive |
| FN | False Negative |
| OA | Overall Accuracy=TP/(TP+TN+FP+FN) |
| ACC | Accuracy=(TP+TN)/(TP+TN+FP+FN) |

## Appendix A. Results of classifications and accuracy analysis


**Figure A1.** Test area - Google Maps

428 *Appendix A.1. Satellite level*

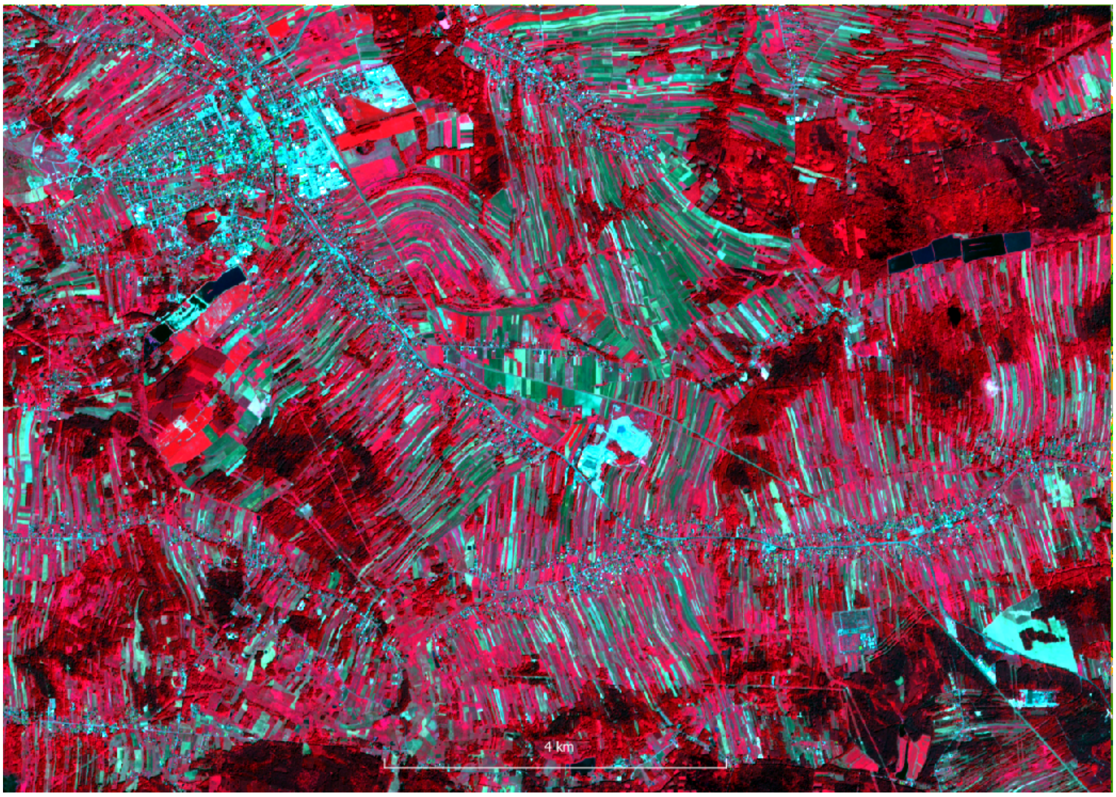


**Figure A2.** S2 FCC 0728

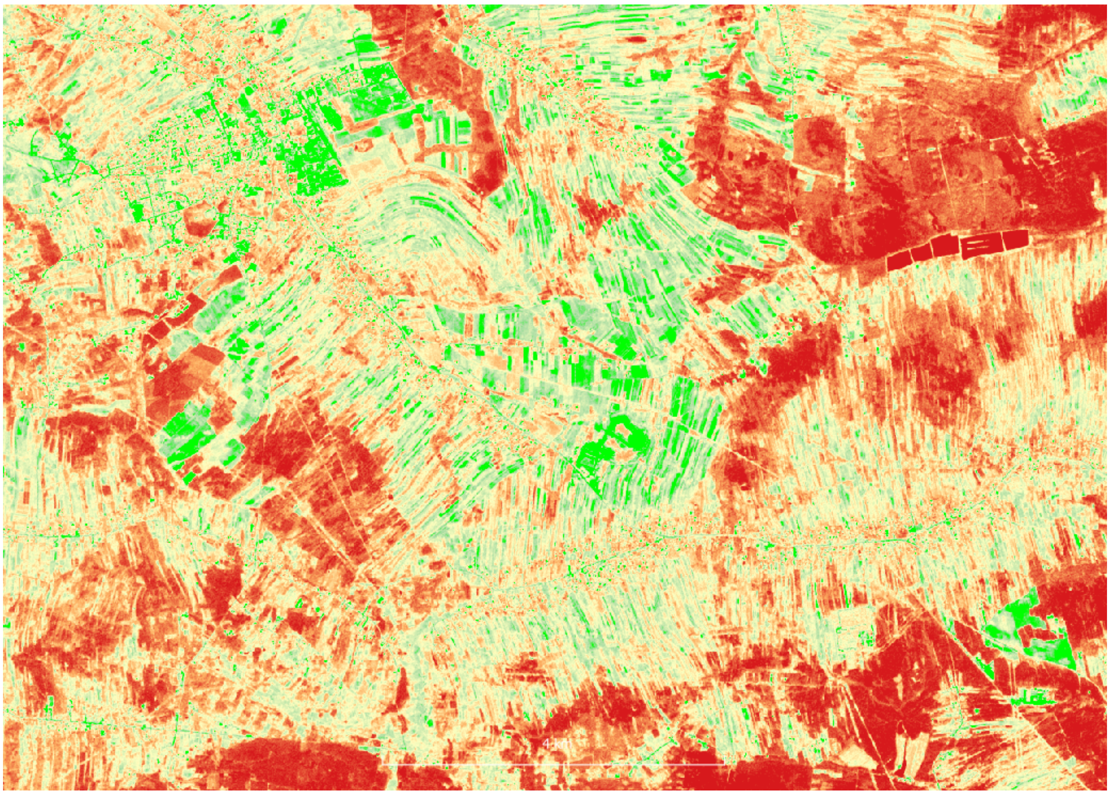

**Figure A3.** S2 NDVI 0728

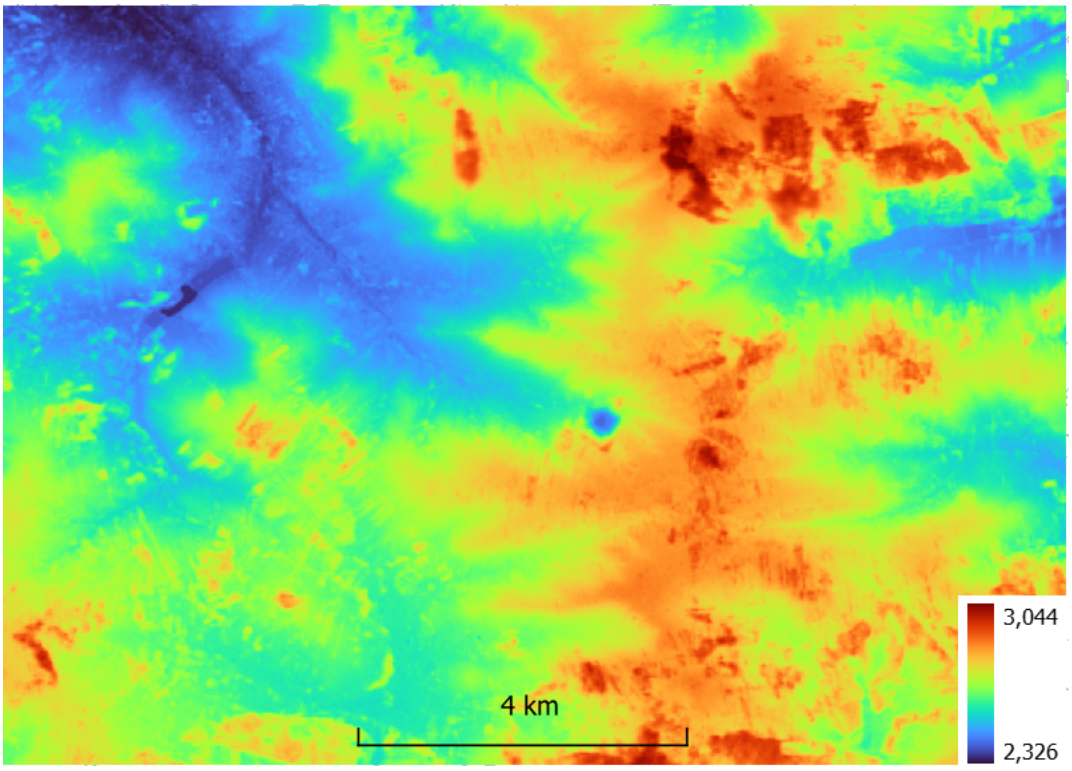


**Figure A4.** SRTM

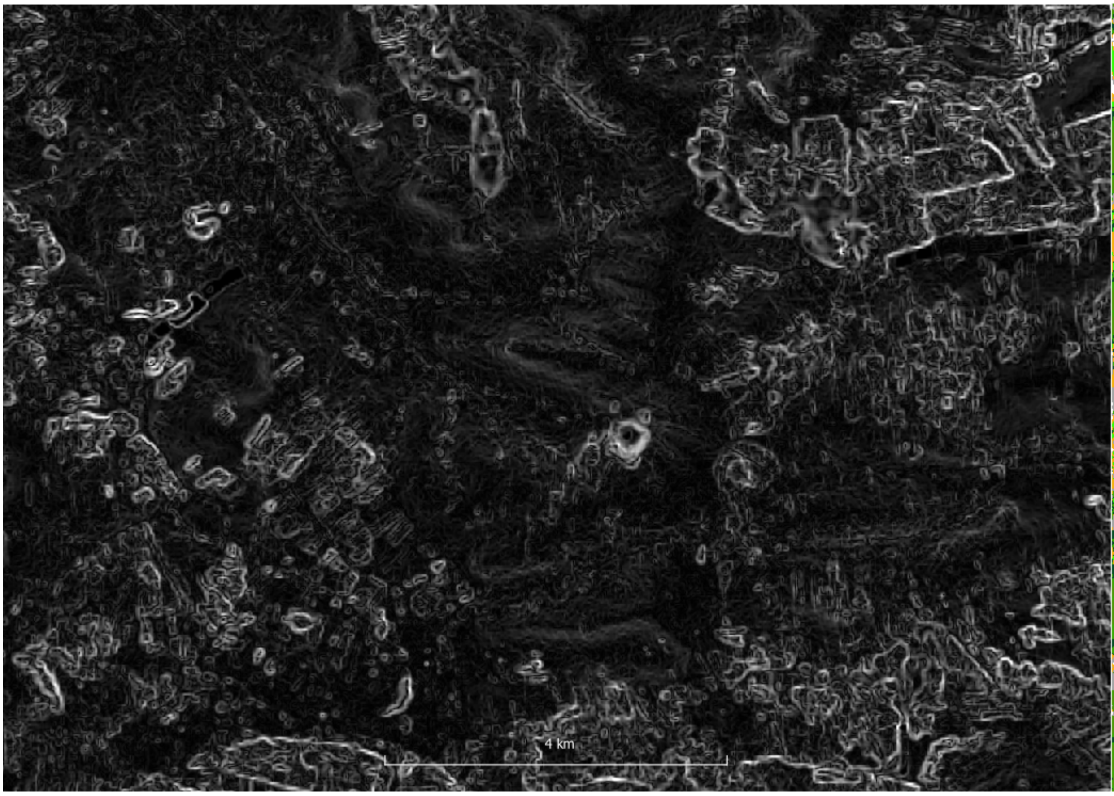


**Figure A5.** SRTM - slope

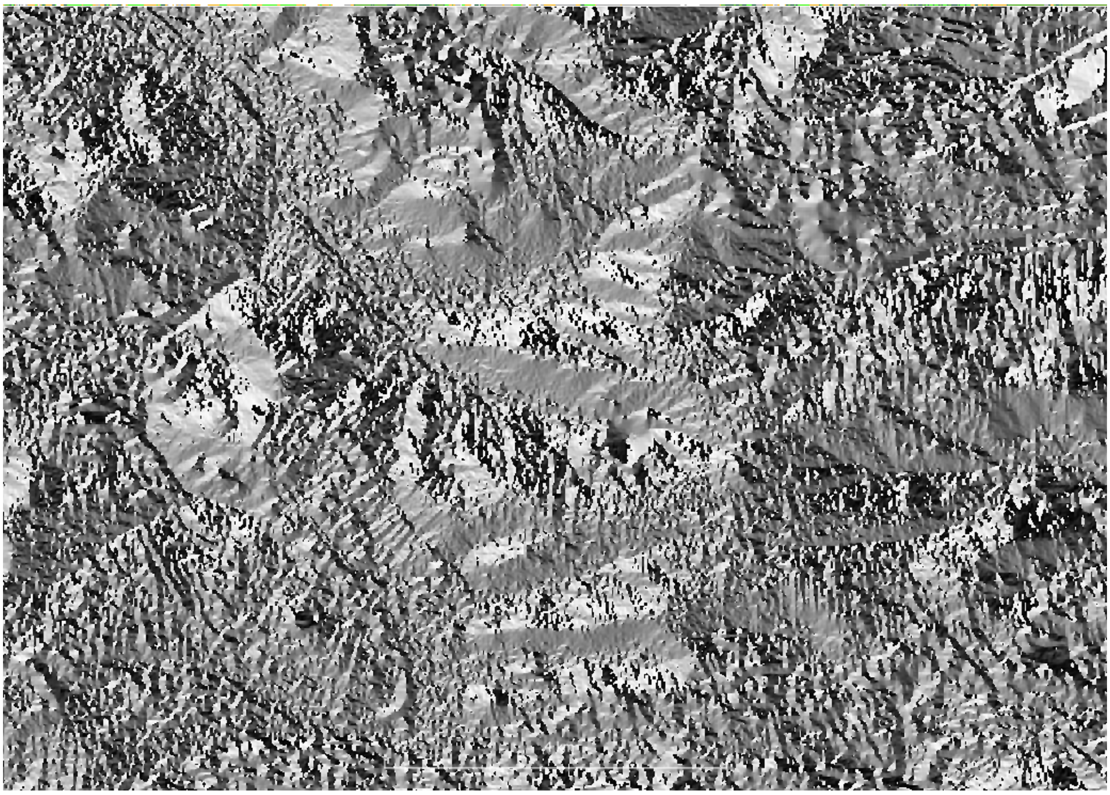

**Figure A6.** SRTM - aspect

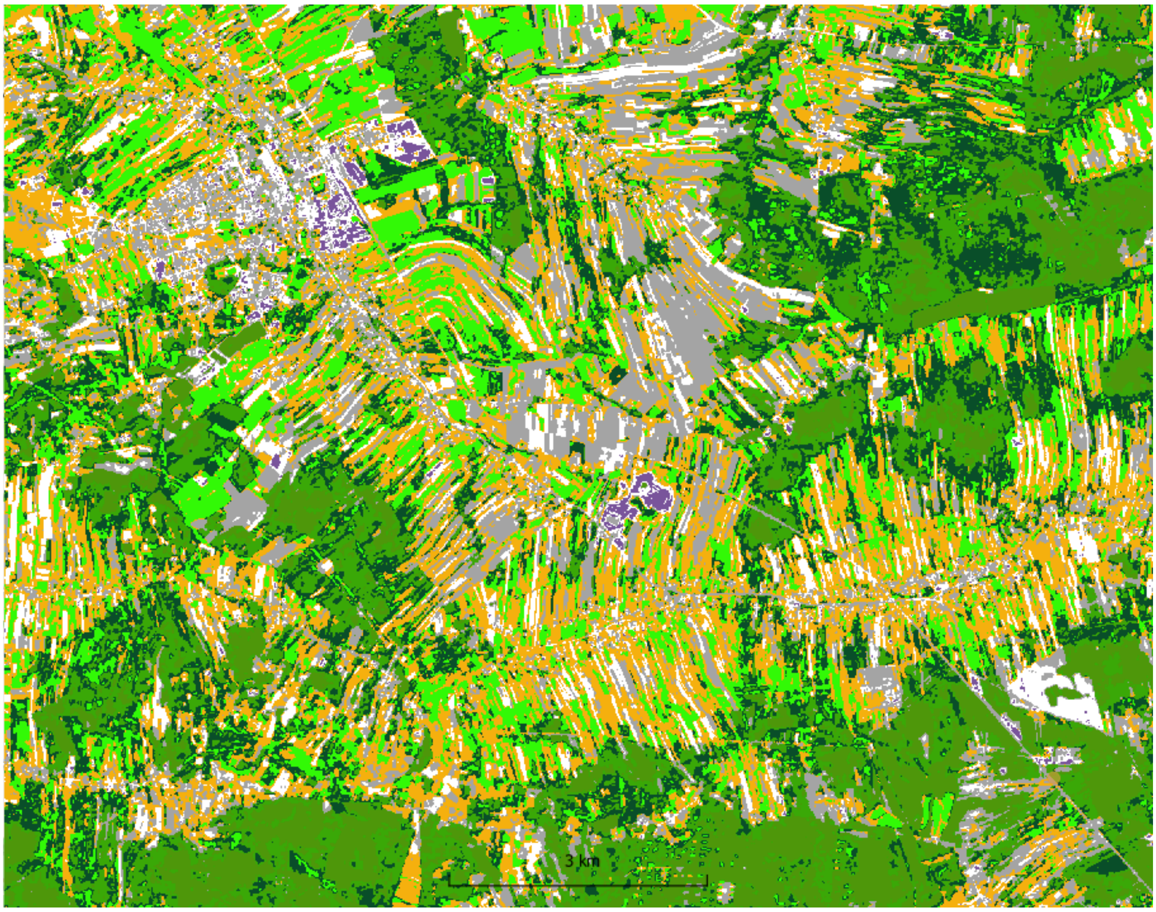


**Figure A7.** Cluster S2 0728; 1-deciduous forest, 2-maturing cereals, 3-conifer forest, 4-industrial, rocks, 5-mix forest, 6-mature cereals, urban, 7-green vegetation, 8-bare soils

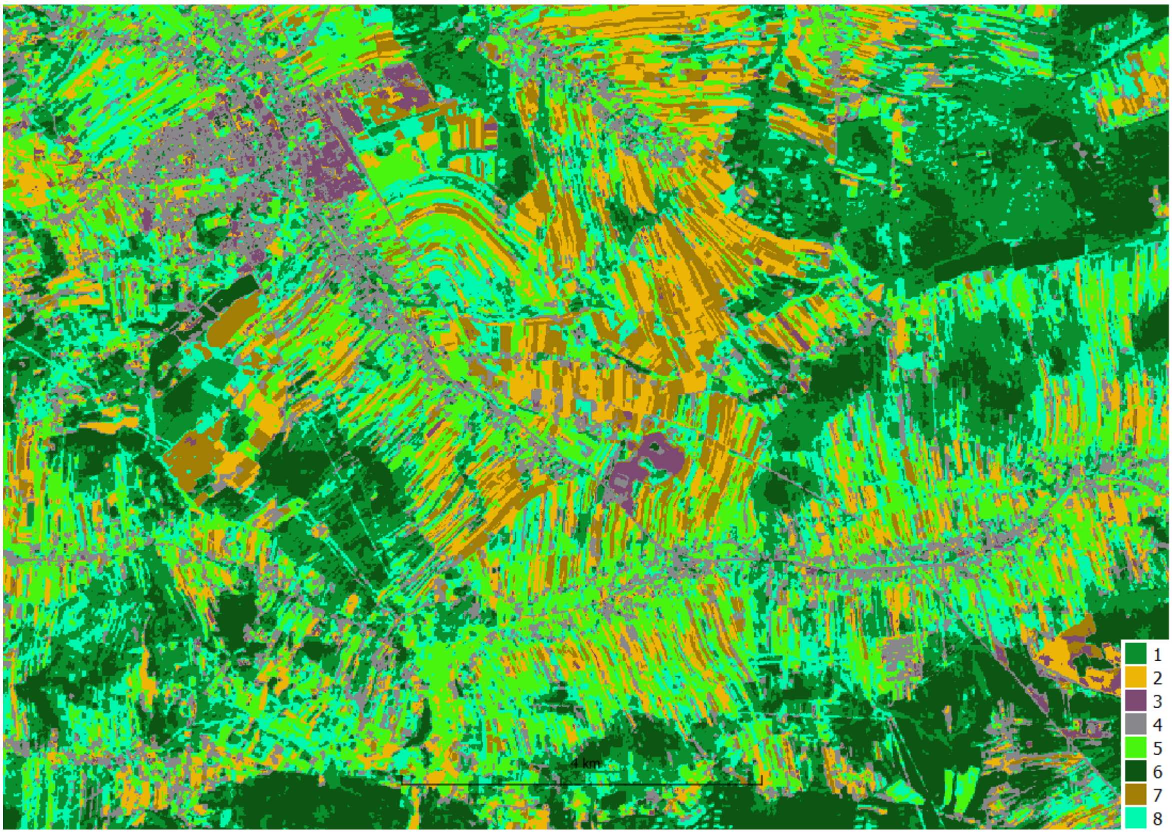


**Figure A8.** Cluster S2 time series; 1-mix forest, 2-mature cereals, 3-industrial, 4-urban, 5-maturing cereals, green vegetation, 6-conifer forest, 7-mature cereals, green vegetation, 8-grass

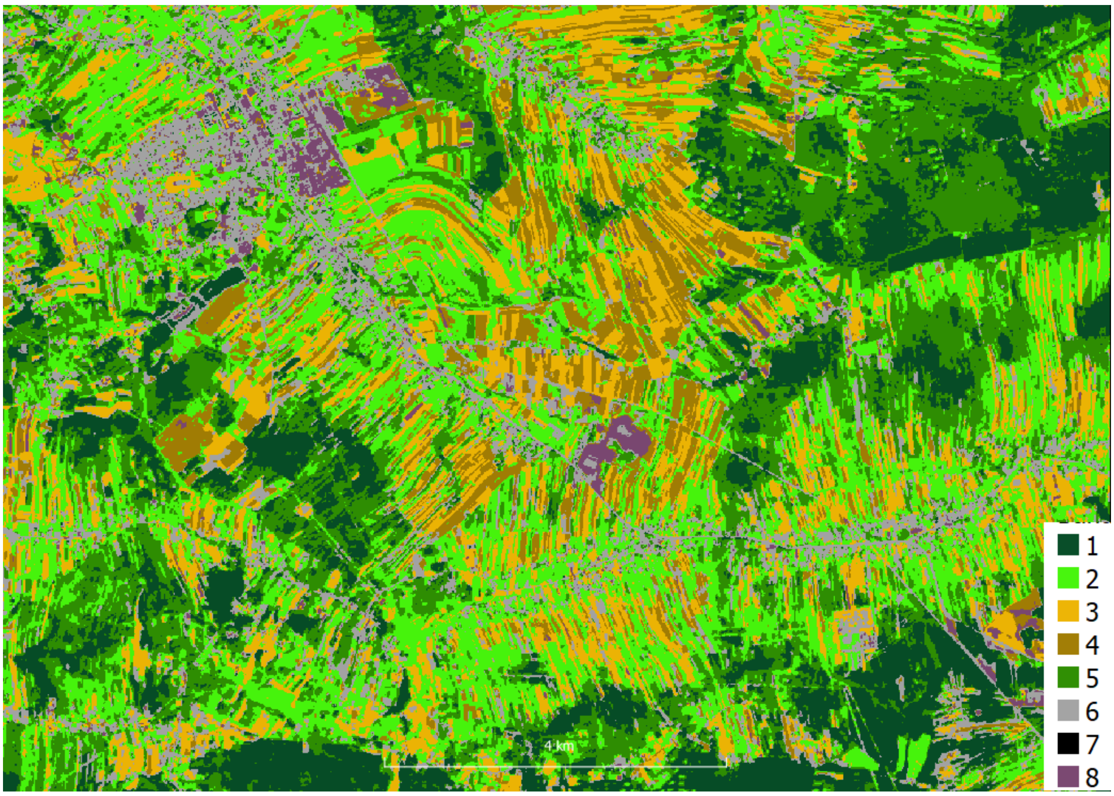


**Figure A9.** Cluster S2 stack; 1-conifer forest, 2-grass, 3-maturing/mature cereals, 4-maturing/mature cereals, 5-deciduous forest, 6-urban, 7-residual class, 8-industrial

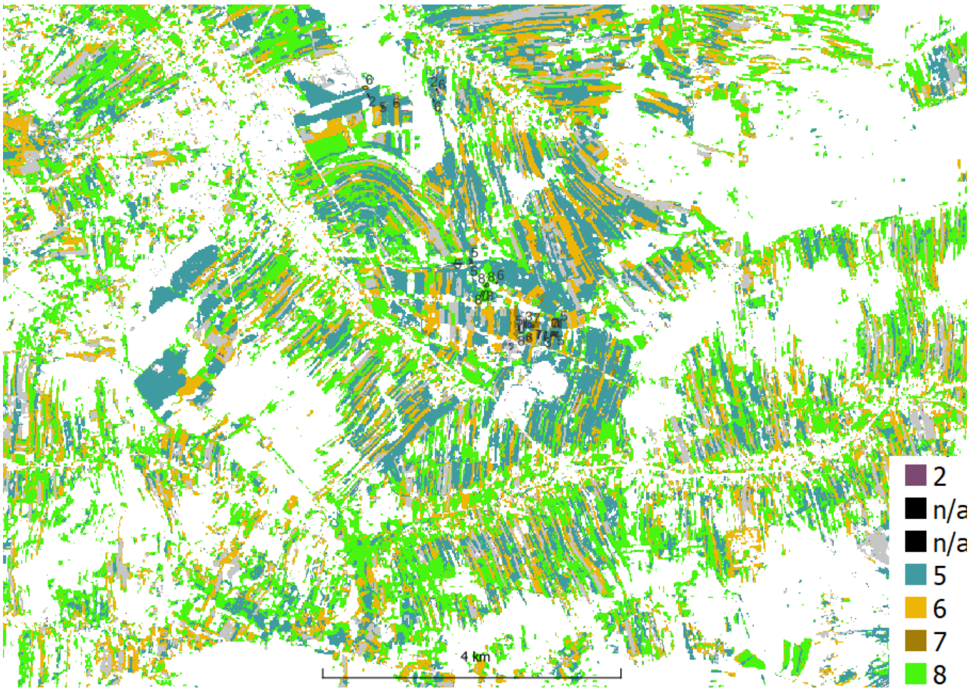


**Figure A10.** RF S2 stack masked, classes according Table 5, 2-soil, 5-oats, 6-winter rye, 7- wheat winter, 8-grass

**Table A1.** Object accuracy analysis Sentinel-2 cluster - class mapping

| Id crop | S2 stack | S2 time series | S2 0728 |
|---|---|---|---|
| 2 | 4 | 2 | 8 |
| 2 | 3 | 2 | 4 |
| 2 | 8 | 3 | 4 |
| 2 | 4 | 7 | 2 |
| 2 | 3 | 2 | 8 |
| 2 | 8 | 3 | 8 |
| 5 | 4 | 7 | 6 |
| 5 | 6 | 4 | 2 |
| 5 | 4 | 7 | 6 |
| 5 | 4 | 7 | 6 |
| 5 | 4 | 7 | 6 |
| 5 | 4 | 7 | 6 |
| 5 | 4 | 7 | 6 |
| 5 | 4 | 7 | 6 |
| 5 | 4 | 7 | 2 |
| 6 | 3 | 2 | 6 |
| 6 | 3 | 2 | 6 |
| 6 | 3 | 2 | 8 |
| 6 | 3 | 2 | 6 |
| 6 | 3 | 2 | 6 |
| 6 | 3 | 4 | 6 |
| 6 | 3 | 2 | 6 |
| 7 | 4 | 2 | 6 |
| 7 | 3 | 2 | 8 |
| 7 | 4 | 7 | 8 |
| 7 | 4 | 7 | 6 |
| 8 | 2 | 5 | 7 |
| 8 | 2 | 5 | 7 |
| 8 | 2 | 5 | 2 |
| 8 | 2 | 5 | 7 |
| 8 | 2 | 5 | 7 |
| OA | 0.74 | 0.68 | 0.39 |
| OA without bare soils | 0.92 | 0.84 | 0.48 |

| | Reference Class | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | Sum |
| (1) n/a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (2) 2 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 12 |
| (3) n/a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (4) n/a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (5) 5 | 0 | 0 | 0 | 0 | 53 | 3 | 0 | 0 | 56 |
| (6) 6 | 0 | 1 | 0 | 0 | 0 | 43 | 0 | 0 | 44 |
| (7) 7 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 9 |
| (8) 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 27 |
| Sum | 0 | 13 | 0 | 0 | 53 | 46 | 9 | 27 | 148 |

**Figure A11.** Pixel accuracy analysis, confusion matrix S2 stack RF, train, OA=0.97

| RF class | Test: 2 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|
| **2** | 3 | 0 | 0 | 3 | 0 |
| **5** | 6 | 54 | 4 | 52 | 0 |
| **6** | 24 | 0 | 33 | 2 | 0 |
| **7** | 0 | 0 | 0 | 34 | 0 |
| **8** | 0 | 0 | 1 | 0 | 89 |
| **OA** | **0.70** | | | | |

**Figure A12.** Pixel accuracy analysis, confusion matrix S2 stack RF, test

**Table A2.** Object accuracy analysis S2 stack RF, test

| Id crop | S2 stack RF |
|---|---|
| 2 | 5 |
| 2 | 6 |
| 2 | 2 |
| 5 | 5 |
| 5 | 5 |
| 5 | 5 |
| 5 | 5 |
| 5 | 5 |
| 6 | 6 |
| 6 | 6 |
| 6 | 5 |
| 7 | 5 |
| 7 | 5 |
| 8 | 8 |
| 8 | 8 |
| 8 | 8 |
| OA | 0.69 |
| OA without bare soils and corn merged | 1.00 |

429 Appendix A.1.1. Clustering - distinguish between industrial areas/mining pits and bare soils

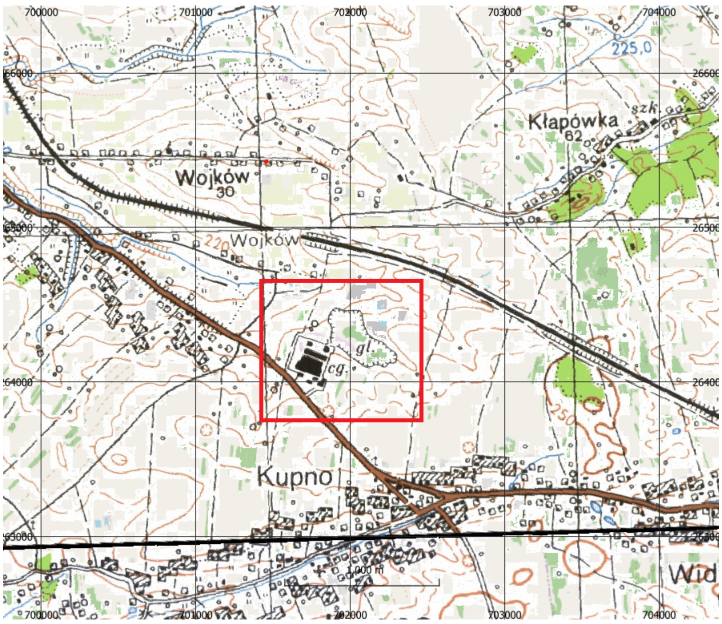


**Figure A13.** Topographical map (geoportal.gov.pl)

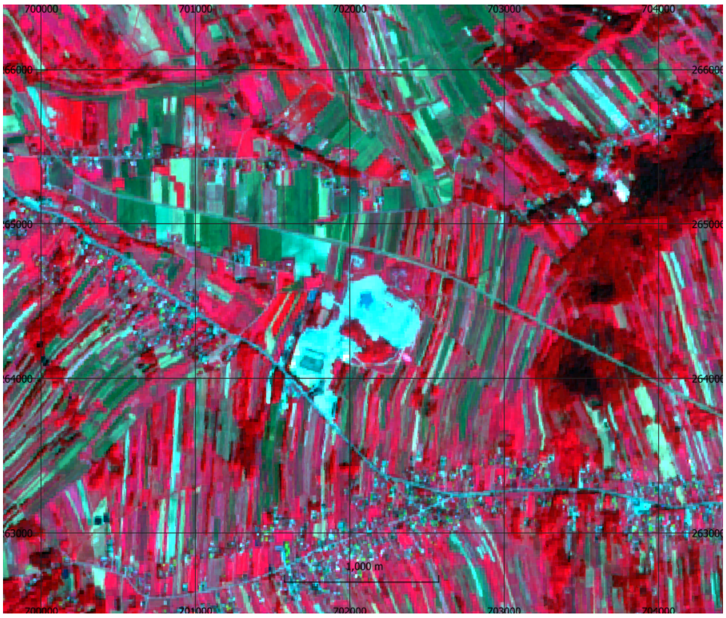


**Figure A14.** FCC S2

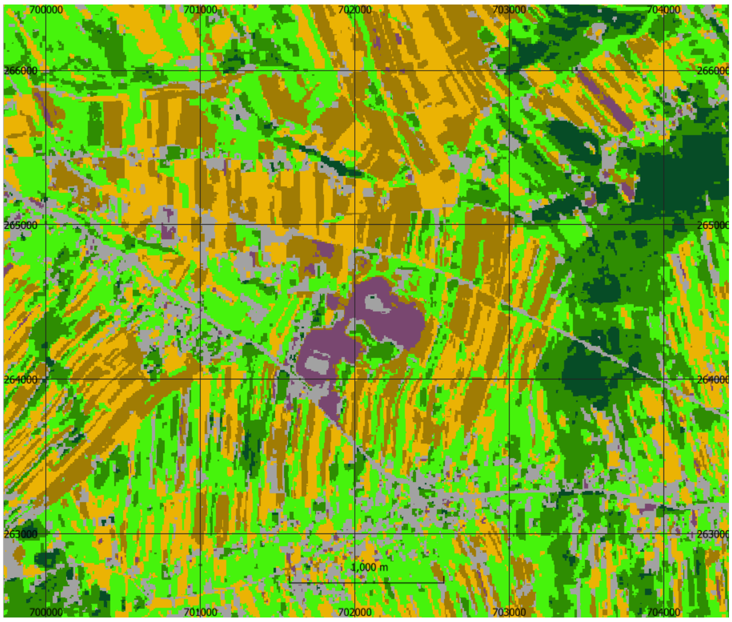


**Figure A15.** Clustering S2 stack

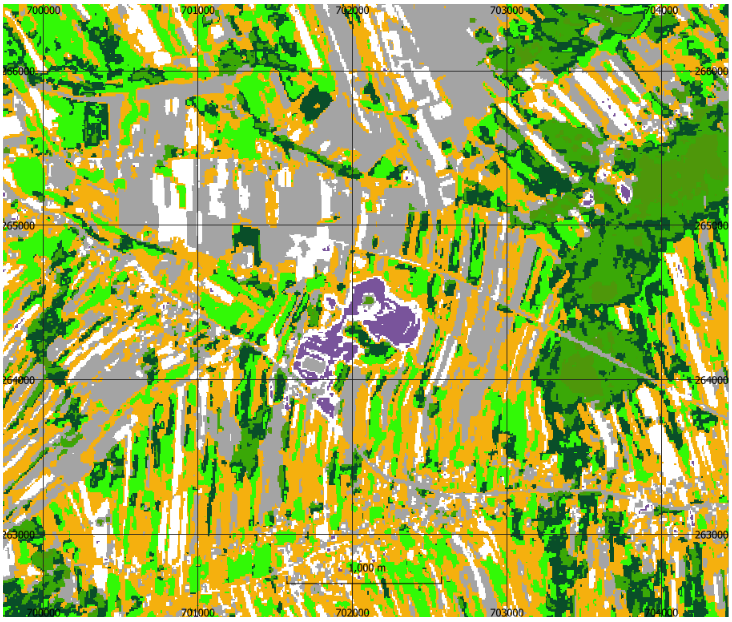


**Figure A16.** Clustering S20728

*Appendix A.2. Airborne level*

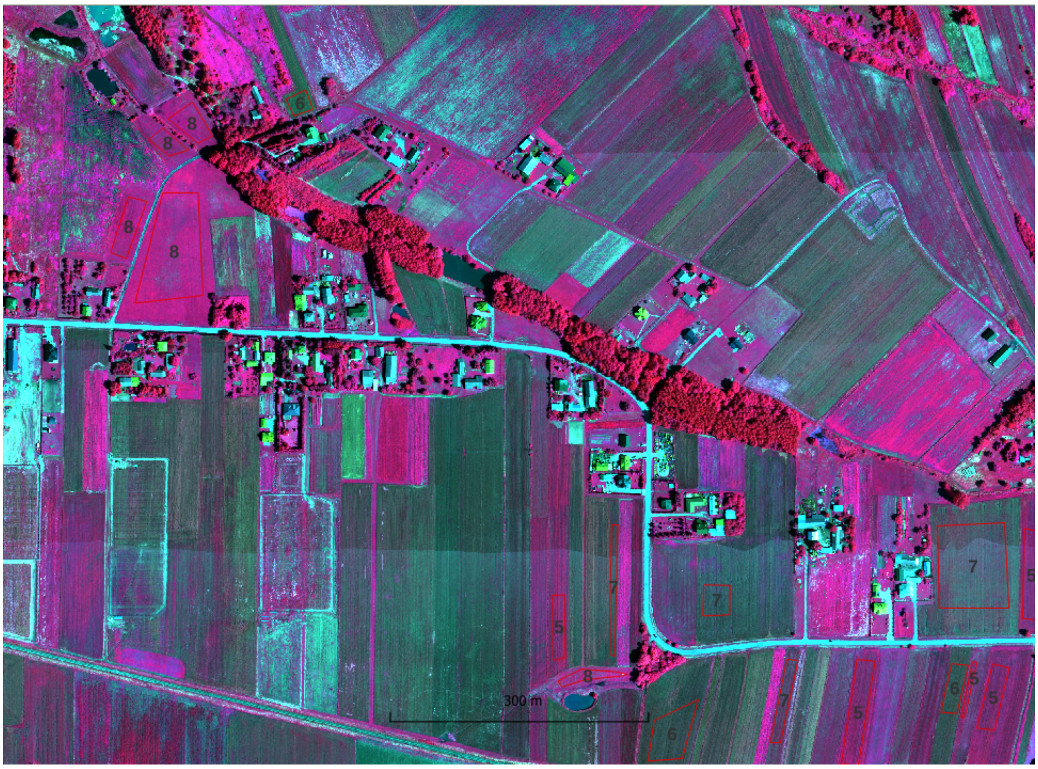


**Figure A17.** HySex FCC 1m

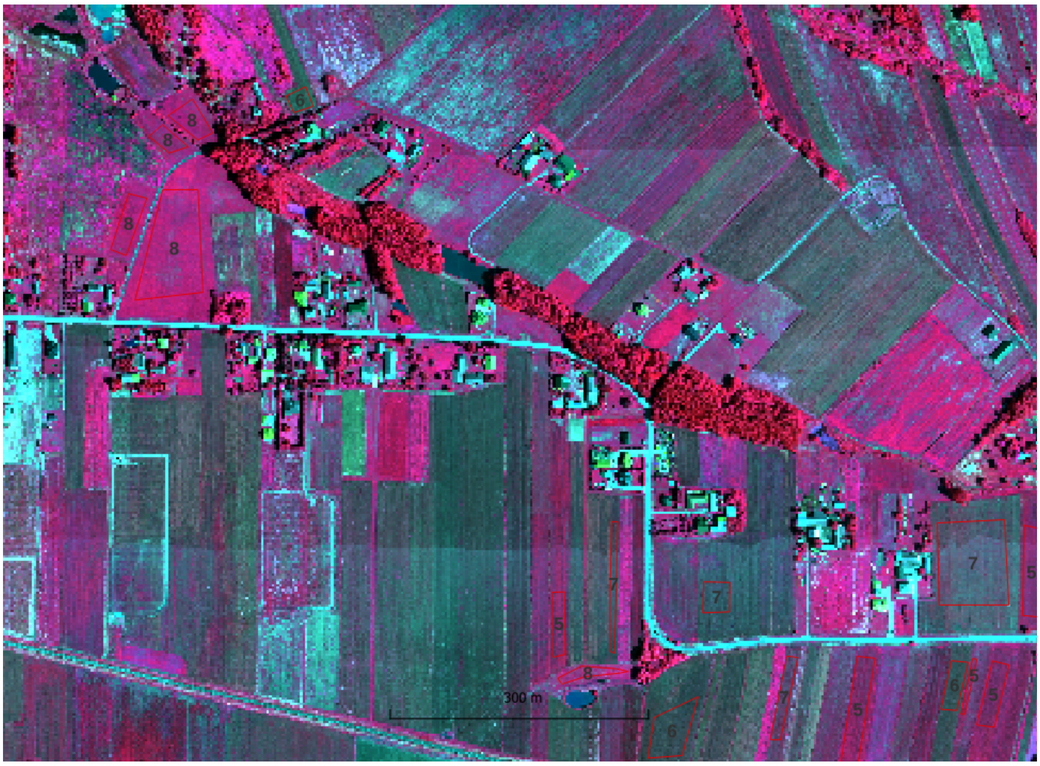


**Figure A18.** HySpex FCC 3m


**Figure A19.** HySex FCC 1m K-means

**Table A3.** Object accuracy analysis HySpex cluster - class mapping

| Id crop | HySpex 1m | HySpex 3m | HySpex stack |
|---|---|---|---|
| 2 | 6 | 6 | 4 |
| 2 | 1 | 3 | 2 |
| 2 | 7 | 3 | 2 |
| 2 | 1 | 3 | 2 |
| 2 | 1 | 3 | 2 |
| 2 | 7 | 3 | 2 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 5 | 5 | 4 | 1 |
| 6 | 6 | 1 | 6 |
| 6 | 5 | 4 | 1 |
| 6 | 3 | 1 | 6 |
| 6 | 6 | 1 | 6 |
| 6 | 6 | 1 | 6 |
| 6 | 6 | 1 | 6 |
| 6 | 6 | 1 | 6 |
| 7 | 6 | 1 | 6 |
| 7 | 6 | 1 | 6 |
| 7 | 6 | 1 | 6 |
| 7 | 6 | 1 | 6 |
| 8 | 5 | 2 | 3 |
| 8 | 5 | 2 | 3 |
| 8 | 5 | 2 | 3 |
| 8 | 5 | 2 | 3 |
| 8 | 5 | 2 | 3 |
| OA | 0.55 | 0.81 | 0.81 |
| OA corn together | 0.71 | 0.97 | 0.97 |

| | Reference Class | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | Sum |
| (1) n/a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (2) 2 | 0 | 117 | 0 | 0 | 0 | 0 | 0 | 0 | 117 |
| (3) n/a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (4) n/a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (5) 5 | 0 | 0 | 0 | 0 | 562 | 117 | 4 | 8 | 691 |
| (6) 6 | 0 | 28 | 0 | 0 | 5 | 319 | 22 | 0 | 374 |
| (7) 7 | 0 | 13 | 0 | 0 | 2 | 59 | 205 | 0 | 279 |
| (8) 8 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 279 | 289 |
| Sum | 0 | 158 | 0 | 0 | 579 | 495 | 231 | 287 | 1750 |

**Figure A20.** Pixel accuracy assessment, confusion matrix HySpex 3m stack RF, train, OA=0.85

| | | | Test | | | |
|---|---|---|---|---|---|---|
| RF | Id crop | 2 | 5 | 6 | 7 | 8 |
| c | 2 | 367 | 0 | 16 | 16 | 0 |
| l | 5 | 5 | 725 | 131 | 526 | 43 |
| a | 6 | 0 | 4 | 117 | 17 | 0 |
| s | 7 | 0 | 7 | 133 | 427 | 0 |
| s | 8 | 8 | 4 | 22 | 0 | 927 |
| OA | 0.73 | | | | | |

**Figure A21.** Pixel accuracy assessment, confusion matrix HySpex 3m stack RF, test, OA=0.73

**Table A4.** Object accuracy analysis RF HySpex 3m stack, test

| Id crop | HySpex 3m stack RF |
|---|---|
| 2 | 2 |
| 2 | 2 |
| 2 | 2 |
| 5 | 5 |
| 5 | 5 |
| 5 | 5 |
| 5 | 5 |
| 5 | 5 |
| 6 | 6 |
| 6 | 5 |
| 6 | 7 |
| 7 | 5 |
| 7 | 5 |
| 8 | 8 |
| 8 | 8 |
| 8 | 8 |
| OA | 0.75 |
| OA corn merged | 1.00 |

## References


1. Cao, J.; Wang, H.; Li, J.; Tian, Q.; Niyogi, D. Improving the Forecasting of Winter Wheat Yields in Northern China with Machine Learning;Dynamical Hybrid Subseasonal-to-Seasonal Ensemble Prediction. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14071707.
2. Bojanowski, J.S.; Sikora, S.; Musiał, J.P.; Woźniak, E.; Dąbrowska-Zielińska, K.; Slesiński, P.; Milewski, T.; Łączyński, A. Integration of Sentinel-3 and MODIS Vegetation Indices with ERA-5 Agro-Meteorological Indicators for Operational Crop Yield Forecasting. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14051238.
3. Zare, H.; Weber, T.K.D.; Ingwersen, J.; Nowak, W.; Gayler, S.; Streck, T. Combining Crop Modeling with Remote Sensing Data Using a Particle Filtering Technique to Produce Real-Time Forecasts of Winter Wheat Yields under Uncertain Boundary Conditions. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14061360.
4. Roma, E.; Catania, P. Precision Oliviculture: Research Topics, Challenges, and Opportunities;A Review. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14071668.
5. Ma, C.; Johansen, K.; McCabe, M.F. Monitoring Irrigation Events and Crop Dynamics Using Sentinel-1 and Sentinel-2 Time Series. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14051205.
6. Borowiec, N.; Marmol, U. Using LiDAR System as a Data Source for Agricultural Land Boundaries. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14041048.
7. Beriaux, E.; Jago, A.; Lucau-Danila, C.; Planchon, V.; Defourny, P. Sentinel-1 Time Series for Crop Identification in the Framework of the Future CAP Monitoring. *Remote Sensing* **2021**, *13*.
8. Commission, E. Integrated Administration and Control System (IACS), 2022.
9. consortium, S. Sentinels for Common Agricultural Policy - Sen4CAP, 2022.
10. consortium, N. New IACS Vision in Action - NIVA, 2022.

11. Rouse, J.; Haas, R.; Schell, J.; Deering, D.; Harlan, J. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. NASA/GSFC Type III Final Report, Greenbelt, MD, 371 p. . "https://ntrs.nasa.gov/api/citations/19750020419/downloads/19750020419.pdf", 1974.
12. Laur, H.; Bally, P.; Meadows, P.; Sanchez, J.; Schättler, B.; Lopinto, E.; Esteban, D. Derivation of the backscattering coefficient sigma nought in ESA ERS SAR PRI products. Technical Report ES-TN-RS-PM-HL09, ESA, September 1998. Issue 2, Rev. 5b. "https://earth.esa.int/documents/10174/13019/ers_sar_calibration_issue2_5f.pdf", 1994.
13. Bolognesi, S.; Pasolli, E.; Belfiore, O.; De Michele, C.; D'Urso, G. Harmonized landsat 8 and sentinel-2 time series data to detect irrigated areas: An application in Southern Italy. *Remote Sensing* **2020**, *12*. doi:10.3390/RS12081275.
14. Hütt, C.; Waldhoff, G.; Bareth, G. Fusion of sentinel-1 with official topographic and cadastral geodata for crop-type enriched LULC mapping using FOSS and open data. *ISPRS International Journal of Geo-Information* **2020**, *9*. doi:10.3390/ijgi9020120.
15. Sun, C.; Bian, Y.; Zhou, T.; Pan, J. Using of multi-source and multi-temporal remote sensing data improves crop-type mapping in the subtropical agriculture region. *Sensors (Switzerland)* **2019**, *19*. doi:10.3390/s19102401.
16. Sun, L.; Chen, J.; Guo, S.; Deng, X.; Han, Y. Integration of time series sentinel-1 and sentinel-2 imagery for crop type mapping over oasis agricultural areas. *Remote Sensing* **2020**, *12*. doi:10.3390/RS12010158.
17. Van Tricht, K.; Gobin, A.; Gilliams, S.; Piccard, I. Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for Belgium. *Remote Sensing* **2018**, *10*. doi:10.3390/rs10101642.
18. Brinkhoff, J.; Vardanega, J.; Robson, A. Land cover classification of nine perennial crops using sentinel-1 and -2 data. *Remote Sensing* **2020**, *12*. doi:10.3390/rs12010096.
19. Maponya, M.G.; van Niekerk, A.; Mashimbye, Z.E. Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Computers and Electronics in Agriculture* **2020**, *169*. doi:10.1016/j.compag.2019.105164.
20. Mustak, S.; Uday, G.; Ramesh, B.; Praveen, B. Evaluation of the performance of SAR and SAR-optical fused dataset for crop discrimination. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 2019, Vol. 42, pp. 563–571. doi:10.5194/isprs-archives-XLII-3-W6-563-2019.
21. d'Andrimont, R.; Verhegghen, A.; Lemoine, G.; Kempeneers, P.; Meroni, M.; van der Velde, M. From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. *Remote Sensing of Environment* **2021**, *266*, 112708. doi:https://doi.org/10.1016/j.rse.2021.112708.
22. consortium, S.A. SEN2-AGRI SYSTEM, 2022.
23. Hejmanowska, B.; Kramarczyk, P.; Głowienka, E.; Mikrut, S. Reliable Crops Classification Using Limited Number of Sentinel-2 and Sentinel-1 Images. *Remote Sensing* **2021**, *13*. doi:10.3390/rs13163176.
24. Meng, S.; Wang, X.; Hu, X.; Luo, C.; Zhong, Y. Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery. *Computers and Electronics in Agriculture* **2021**, *186*. doi:10.1016/j.compag.2021.106188.
25. Wikipedia. Żniwa, 2022.
26. Hord, M.R.; Brooner, W. Land-Use Map Accuracy Criteria. *Photogramm. Eng. Rem. S.* **1976**, *42*.
27. van Genderen, J.; Lock, B. Testing Land Use Map Accuracy. *Photogramm. Eng. Rem. S.* **1977**, *43*.
28. Russell, G.C. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. doi:https://doi.org/10.1016/0034-4257(91)90048-B.
29. Canran, L.; Paul, F.; Lalit, K. Comparative assessment of the measures of thematic classification accuracy. *Remote Sens. Environ.* **2007**, *107*, 606–616. doi:https://doi.org/10.1016/j.rse.2006.10.010.
30. Pontus, O.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. doi:https://doi.org/10.1016/j.rse.2012.10.031.
31. Foody, G.M. Impacts of Sample Design for Validation Data on the Accuracy of Feedforward Neural Network Classification. *Appl. Sci.* **2017**, *7*, 888. doi:10.3390/app7090888.
32. Stehman, S.V.; Foody, G.M. Key issues in rigorous accuracy assessment of land cover products. *Remote Sens. Environ.* **2019**, *231*, 111199. doi:https://doi.org/10.1016/j.rse.2019.05.018.

33. Morales-Barquero, L.; Lyons, M.B.; Phinn, S.R.; Roelfsema, C.M. Trends in Remote Sensing Accuracy Assessment Approaches in the Context of Natural Resources. *Remote Sens.* **2019**, *11*. doi:10.3390/rs11192305.
34. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. ROC Analysis in Pattern Recognition, doi:https://doi.org/10.1016/j.patrec.2005.10.010.
35. Mann, H.M.R.; Iosifidis, A.; Jepsen, J.U.; Welker, J.M.; Loonen, M.J.J.E.; Høye, T.T. Automatic flower detection and phenology monitoring using time-lapse cameras and deep learning. *Remote Sensing in Ecology and Conservation* **2022**, *8*, 765–777.
36. Lake, T.A.; Briscoe Runquist, R.D.; Moeller, D.A. Deep learning detects invasive plant species across complex landscapes using Worldview-2 and Planetscope satellite imagery. *Remote Sensing in Ecology and Conservation* **2022**, *8*, 875–889.
37. Wan, S.; Yeh, M.L.; Ma, H.L. An innovative intelligent system with integrated CNN and SVM: Considering various crops through hyperspectral image data. *ISPRS International Journal of Geo-Information* **2021**, *10*. doi:10.3390/ijgi10040242.