BESTMAP
Behavioural, Ecological & Socio-economic Tools for Modelling Agricultural Policy

Mapping of vegetation indices and metrics, and their utility in FSA mapping at CS scale

Deliverable D5.4

31 October 2021

Guido Riembauer, Markus Metz
mundialis Gmbh & Co KG
Guy Ziv, Arjan Gosal
University of Leeds
James Bullock, Paul Evans
UK Centre for Ecology and Hydrology
Tomáš Václavík, Fanny Langerwisch, Marek Bednář
Palacký University Olomouc
Sanja Brdar, Predrag Lugonja
BioSense Institute

BESTMAP

Behavioural, Ecological and Socio-economic Tools for Modelling Agricultural Policy

Prepared under contract from the European Commission

Grant agreement No. 817501
EU Horizon 2020 Research and Innovation action

Project acronym:	**BESTMAP**
Project full title:	**Behavioural, Ecological and Socio-economic Tools for Modelling Agricultural Policy**
Start of the project:	September 2019
Duration:	48 months
Project coordinator:	Prof. Guy Ziv School of Geography, University of Leeds, UK http://bestmap.eu/
Deliverable title:	Mapping of vegetation indices and metrics, and their utility in FSA mapping at CS scale
Deliverable n°:	D5.4
Nature of the deliverable:	Report
Dissemination level:	Public
WP responsible:	WP5
Lead beneficiary:	mundialis GmbH & Co KG
Citation:	Riembauer, G., Metz, M., Ziv, G., Gosal, A., Bullock, J., Evans, P., Václavík, T., Langerwisch, F., Brdar, S., Lugonja, P., (2021). *Mapping of vegetation indices and metrics, and their utility in FSA mapping at CS scale*. Deliverable D5.4 EU Horizon 2020 BESTMAP Project, Grant agreement No. 817501.
Due date of deliverable:	Month n°26
Actual submission date:	Month n°26

Deliverable status:

Version	Status	Date	Author(s)
1.0	Final	31 October 2021	Guido Riembauer, Markus Metz mundialis GmbH & Co KG Guy Ziv, Arjan Gosal University of Leeds James Bullock, Paul Evans UK Centre for Ecology and Hydrology Tomáš Václavík, Marek Bednář Palacký University Olomouc Sanja Brdar, Predrag Lugonja BioSense Institute

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Table of contents

_____

Summary

This deliverable provides an overview of all work conducted in the context of Activity 5.3.1 (Developing remote sensing indicators) with respect to Farming System Archetype (FSA) Mapping (Task 5.3). This work is based on the FSA definition and mapping in 'D2.2 - Conceptual Framework' and 'D3.5 - Farming System Archetypes for each CS' and investigates the potential of remote sensing methods to inform different dimensions of FSAs. Findings from this analysis will contribute to the BESTMAP roadmap (Task 5.4). Specifically, methodologies for crop type mapping, crop yield estimation, and field boundary mapping are investigated in different case study regions and their relevance for FSAs are shown.

In the context of crop type mapping, two remote sensing based crop map products are evaluated and a crop classification approach is applied using Sentinel-2 data. It is found that existing large-scale crop maps have limited accuracy on class level but can give an indication of the more abstract FSA farm specialization. A machine learning crop classification approach yields high accuracies when optimized for a specific case study region.

Emphasis is put on the assessment of yield estimation approaches that are potentially suitable to complement the economic size dimension of FSAs. A Sentinel-2 time-series is generated for an entire FADN region and the relationship between vegetation index statistics and reference yield data are analyzed. The maximum GNDVI value of the growing season is found as best correlated with reference yield. Additionally, yield mapping based on soil and weather data is assessed. Work on this topic is ongoing and the approaches are to be refined and/or merged. Finally, the need for a consistent field boundary mapping method is highlighted and a processing architecture based on Sentinel-2 data is proposed.

With respect to the Grant Agreement, the scope of Activity 5.3.1 has been slightly shifted. Originally, Sentinel-2 vegetation index time series were to be calculated for all case study regions, and Sentinel-1 backscatter time series for two case study regions additionally. However, with regard to the simplification to two FSA dimensions (see D3.5) it seemed more expedient to perform exploratory analyses on how to enrich and complement these dimensions with remote sensing methods rather than to create large time-series for all case study regions without a clear link to FSAs. Because of this, only one large time series of Sentinel-2 data was created (though scaled up to a FADN region for data availability reasons) for the assessment of vegetation indices, while Sentinel-1 data analysis as well as grassland cutting frequency calculation was omitted. In summary, the spatial scope of this activity was reduced in favor of more detailed remote sensing methodology analyses that have direct implications for the two FSA dimensions.

_____

1. Introduction: Remote Sensing for Farming System Archetype mapping

Mapping phenological metrics is a typical application of optical remote sensing in the agriculture domain. The idea of Task 5.3 being reported in this deliverable is to assess the potential of remote sensing for Farming System Archetypes (FSA) mapping. Originally, the FSA classification was envisioned to cover environmental conditions, land-use intensities and management practices as well as socio-economic factors that would provide insights in farmer's decision making. However, as outlined in D2.2, the final FSA definition applied in BESTMAP was simplified to ensure a harmonized approach across CSs and enable Europe-wide upscaling. Therefore, we made use of data that can be linked to the European Farm Accountancy Data Network (FADN) data and included only two FSA dimensions **(1) farm specialization** and **(2) economic size**. Farms were mapped according to this classification across the BESTMAP case study areas in D3.5. D3.5 also identified limitations in data availability and definition inconsistencies across the case study areas (such as definition of field) hindering the direct allocation of a farm to the two FSA dimensions.

Remote sensing data does not provide information on land ownership and its potential to predict farmers' behaviour with respect to agricultural policies is limited to providing structural information derived from land cover maps. However certain metrics and spatial information can be extracted from remote sensing data that can be valuable to inform and/or refine the two FSA dimensions. Assessing methodologies to derive this information from remote sensing and applying them to case study regions (or scaling up to FADN regions) was the main focus of Task 5.3. With regard to the simplified FSA definition, they are specifically:

- Crop type mapping: Mapping crop types from remote sensing data directly informs the farm specialization dimension of FSAs because the type of farm specialization is determined based on the extent of individual crops grown at a farm.
- Yield mapping: The estimation of field-level yield from remote sensing data can serve as a complement to the Standard Output Coefficients (SOCs) selected in D3.5 to inform the economic size dimension of FSAs. Yield is heavily dependent on yearly weather conditions etc. and as such can reflect the temporal dynamics of a farm's economic output.
- Extraction of field boundaries: Applying a spatially independent algorithm to extract field boundaries from remote sensing data reduces the dependency on member state specific IACS/LPIS and solves the inconsistency of field definitions identified in D3.5.

The following sections summarize analyses that showcase our assessment of the aforementioned remote sensing methodologies: Section 2 describes the activities with regard to crop type mapping, while section 3 gives an overview on yield estimation, providing an in-depth analysis of Sentinel-2 based vegetation index time-series. Finally, section 4 outlines a prototypical algorithm for the calculation of field boundaries. Discussions are provided at the end of each section, as the analysed remote sensing applications have different scopes and levels of maturity with regard to their contributions to FSA classification.

2. Crop Type Mapping

Mapping crop types from remote sensing data has the potential in BESTMAP for directly informing the farm specialization dimension of FSAs because the type of farm specialization is determined based on the extent of individual crops grown at a farm. Therefore, the objective of this task was to utilize new remote sensing data products (e.g. based on Sentinel-1 and Sentinel-2) to produce vegetation indicators that would allow the classification of FSAs based on these remotely sensed datasets and compare them with FSA mapping in the Case Studies (CSs) based on local field/farm data. The main existing data on farmland in each CS are IACS/LPIS (Integrated Administration and Control System / Land Parcel Identification System), currently serving as the basis for the farm specialization dimension of FSAs. Three analyses were performed to investigate to what extent remote sensing products can serve as an alternative for this purpose:

- For the Czech case study region, a state-of-the-art remote sensing product of crop classification provided by the Belarusian company OneSoil is compared to IACS/LPIS data (Section 2.1)
- For the UK case study region, the remote sensing based "Land Cover Plus: Crops" product (provided by the UK Centre for Ecology and Hydrology (UKCEH)) is used to classify the farm specialization dimension of FSAs, and a comparison to the IACS/LPIS approach is drawn (Section 2.2)
- For the Serbian case study region, where IACS/LPIS data is not available, a machine-learning approach is implemented to classify crop types based on Sentinel-2 data (Section 2.3)

2.1. Czech case study region: Comparison of OneSoil crop type data to LPIS

2.1.1. Objective

As our FSA classification uses crop type as a main input layer, we assessed the usability/accuracy of a field-level remote sensing 'crop type' product available from OneSoil (http://onesoil.ai), a Belarusian company that develops web and mobile precision farming apps that are based on satellite imagery and machine learning technologies, for the whole of the EU. We did this assessment for the Czech case study as an example, and compared the fine-scale land-use data provided by the IACS/LPIS data with the remote sensing product of OneSoil. The OneSoil data provided information on agricultural land uses for 2018 for each parcel centroid (extracted from LPIS), including 7 crop types (wheat, maize, sugar beet, barley, sunflower, rapeseed, soybeans) and 3 non-crop type categories (grass, other, not field). This is much fewer than the number of crop types identified in the IACS/LPIS database, which include over a hundred specific crops, although some of them occur only on a few fields or are grown only in some years. Thus, the comparison was conducted for the 10 crop type categories available in both datasets, using LPIS data for the same year (2018) and including only those parcels that were included in both datasets.

2.1.2. Results

The results showed that the overall agreement between LPIS and the state-of-the-art remote sensing product from OneSoil is relatively poor. Only about 40% of the farmland area - covered with wheat, maize or sugar beet - was classified mostly correctly. These areas were covered with the highest crop cover per parcel (dominating crop). In contrast, about 20% of the area - mainly covered with barley, sunflower, rapeseed or soy - was classified largely incorrectly.

Table 1: Overview of classification accuracy/overlap between IACS/LPIS data and the remote sensing product from OneSoil.

Crop Type	parcels in LPIS (n = 12373)	parcels in OneSoil (n = 12372)	parcels overlay in both (rel. to LPIS%, rel. to OneSoil%)
Match of more than 40%:			
Wheat	2127 (17%)	2189 (18%)	1387 (65%, 63%)
Maize	875 (7%)	1746 (14%)	782 (89%, 45%)
Sugar Beet	22 (<1%)	22 (<1%)	15 (68%, 68%)
Match of less than 40%:			
Barley	575 (5%)	552 (4%)	184 (32%,33%)
Sunflower	175 (1%)	723 (6%)	1 (0.6%, 0.1%)
Rapeseed	673 (5%)	183 (1%)	2 (0.3%, 1%)
Soybeans	73 (1%)	46 (<1%)	23 (32%, 50%)
non-crop cover			
Not Field	6097 (49%)	896 (7%)	679 (11%, 76%)
Grass	408 (3%)	3827 (31%)	175 (43%, 6%)
Other	1348 (11%)	2188 (18%)	258 (19%, 12%)

The main mis-classification occurred for wheat instead of barley, rapeseed instead of sunflower and vice versa, and maize instead of soy. This is mainly due to comparable spectral signals of the flowering plants (rapeseed / sunflower) or a comparable seasonality (e.g. wheat / barley). Overall we conclude that the use of these remotely sensed crop type data available at EU scale (in this case the OneSoil product) is not recommended, at least at this stage and quality for the classification of FSAs and therefore also not recommended for the upscaling of FSAs to the European level.

2.2. UK case study region: FSA mapping based on UKCEH Land Cover Plus: Crops data

2.2.1. Objective

Farm System Archetypes (FSA) have been calculated for farms in the Humber region of the United Kingdom in BESTMAP Deliverable 3.5, using Land Parcel Identification System (LPIS) data, which includes crop information, under licence from the Rural Payments Agency (RPA). An alternative data source from the UK Centre for Ecology and Hydrology (UKCEH) is available, who produce "Land Cover Plus: Crops", annual maps of the major crops each autumn for all fields in Great Britain from Copernicus Sentinel-1 C-band SAR (Synthetic

Aperture Radar) and Sentinel-2 optical data. UKCEH Land Cover plus Crops is based on the UKCEH Land Cover Map parcel framework (https://www.ceh.ac.uk/ukceh-land-cover-maps). Every parcel which is larger than 2 ha and categorised as agricultural land is coded with crop type information. Here, we investigate whether this data can be used as an alternative to RPA data and how this affects the FSA classification.

The UKCEH data has 13 crop types; beet (sugar beet / fodder beet), field beans, grass, maize, oilseed rape, peas, potatoes, spring barley, winter barley, spring wheat, winter wheat, winter oats and other crops. This is much fewer than the crop types that are available in the RPA (LPIS) data. A total of 123 crop types were present for the Humber in the RPA data, though after exclusions (e.g. ponds, roads, etc.), a total of 91 crops remained. As grassland is not stipulated as permanent (P4/grazing livestock and forage) or temporary (P1/general cropping) in the UKCEH data, two approaches were used to match the grassland type to either P1 or P4. It is assumed that permanent grassland would be necessary for livestock farms.

2.2.2. Methodology

UKCEH Land Cover Plus: Crops 2019 data, hereafter referred to as UKCEH data (UK Centre for Ecology and Hydrology, n.d.), was used to extract the crop types in field parcels for the Humber region in the UK. To allow comparison of the FSAs calculated based on RPA data and UKCEH data, the LPIS parcels were used as a constant delineator of field parcel boundaries. The full methodology for determining the FSA from crop types is fully explained in BESTMAP Deliverable 3.5, therefore not repeated here. The total area of the RPA parcels (2019) was aggregated by anonymised farm business to give total farm area, which was used to calculate percentage of area covered by any particular specialization (e.g. P1) from which the FSA was derived (see BESTMAP Deliverable 3.5) from the UKCEH data. For example, if UKCEH data indicated a P1 crop covering 50% of all parcels from a farm business, it would not meet the 66% threshold to be categorised as a P1 FSA and would therefore be categorised as Mixed.

A vital step was matching crops with specialization types; (P1) general cropping, (P2) horticulture, (P3) permanent crops, and (P4) grazing and livestock. This was relatively simple with most UKCEH named crops which matched to general cropping, though for the "grass" and "other crops" categories this was more difficult. As the UKCEH grass category includes temporary and permanent grassland in addition to numerous other crop types (**Redhead 2018**), it could be allocated to multiple specialization types. Previous FSA calculations from the RPA data in BESTMAP Deliverable 3.5 show that most grassland is permanent in the study region. The UKCEH "other" category can include, but is not limited to, other cereals and field vegetables (**Redhead 2018**), therefore was classified as P1. Three scenarios of analysis were therefore conducted, differing in matching the UKCEH "grass" category:

A. Multiple years (2017-2020) UKCEH data, with consistent grass classified as P4 and inconsistent grass classified as P1. For an LPIS parcel to be classified as P4, at least 50% of the parcel had to be classified as "grass" in the UKCEH data for each of the four years.

B. Single year (2019) UKCEH data, with all grass classified as P4.
C. Single year (2019) UKCEH data, with all grass classified as P1.

2.2.3. Results

A total of 3515 out of 3528 farms within the region had at least partial crop information available in the UKCEH data that were used to classify the farms into FSA: (P1) General cropping, (P2) Horticulture, (P3) Permanent crops, (P4) Grazing livestock and forage, and (Mixed) mixed farms with no clear specialization (see Deliverable 3.5 for details on how farms are classified into FSA). A total of 13 farms had no parcels with crop information (10 of these fell into the P4 RPA FSA category, and 3 into the P1 and so were not classified under any of the three analysis scenarios). Figure 1 shows the results for the different scenarios, using the RPA derived FSAs as reference. Over 75% of P1 farms were correctly classified in all scenarios with no P3 classified correctly. Approximately 50% of the P4 FSA were correctly identified (according the to the RPA derived FSA) using both multi year data and where grass was classified as P4 for a single year (Figure 1 a/b), whereas classifying grass as P1 meant no farms were correctly identified as P4 (Figure 1 c).
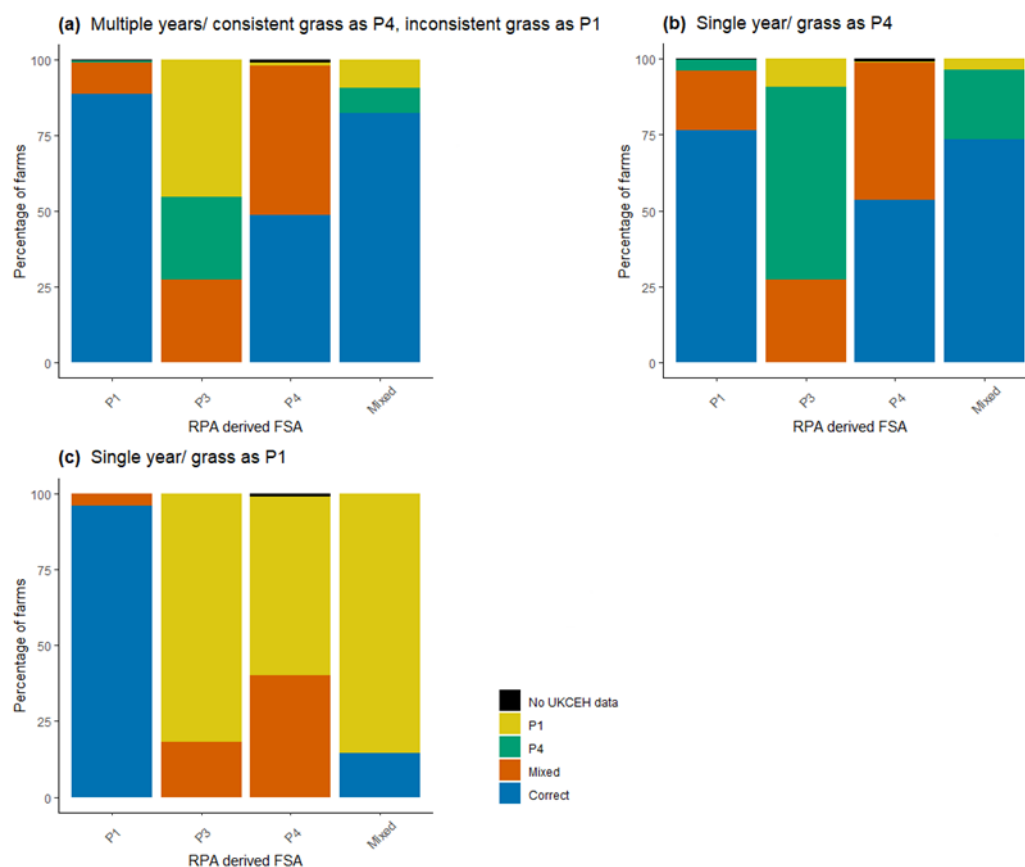

Figure 1: Plot showing the percentage of matched and mismatched UKCEH derived FSA classifications with reference to RPA derived FSA classifications; (a) multiyear data used to determine "grass" as either P1 (temporary) or P4 (permanent), (b) single-year "grass" classified as P4, and (c) single-year "grass" classified as P1. Number of samples n of RPA derived FSA are as follows: P1 = 2242, P3 = 11, P4 = 912, and Mixed = 363. The "correct"

category refers to farms that are identified as the same FSA as in the RPA derived FSA. "No UKCEH data" refers to farms where no farm parcels with UKCEH data present.

Figure 2 shows the spatial effect of this classification for the Humber CS region. Scenario A (assigning "grass" to FSA P1 or P4 depending on its temporal consistency) clearly shows the best agreement with the reference RPA derived FSA map. This scenario shows the best overall agreement with 77.5% (see Table 2). An issue with the RPA derived FSA calculations (BESTMAP Deliverable 3.5) was the lack of information whether a field was used for horticulture (P2), and this issue is also prevalent in the UKCEH derived FSA, in addition to having no clearly defined permanent crop (P3) information in the UKCEH data. Therefore both derived FSA might suffer from misclassified P2 specialised crops.
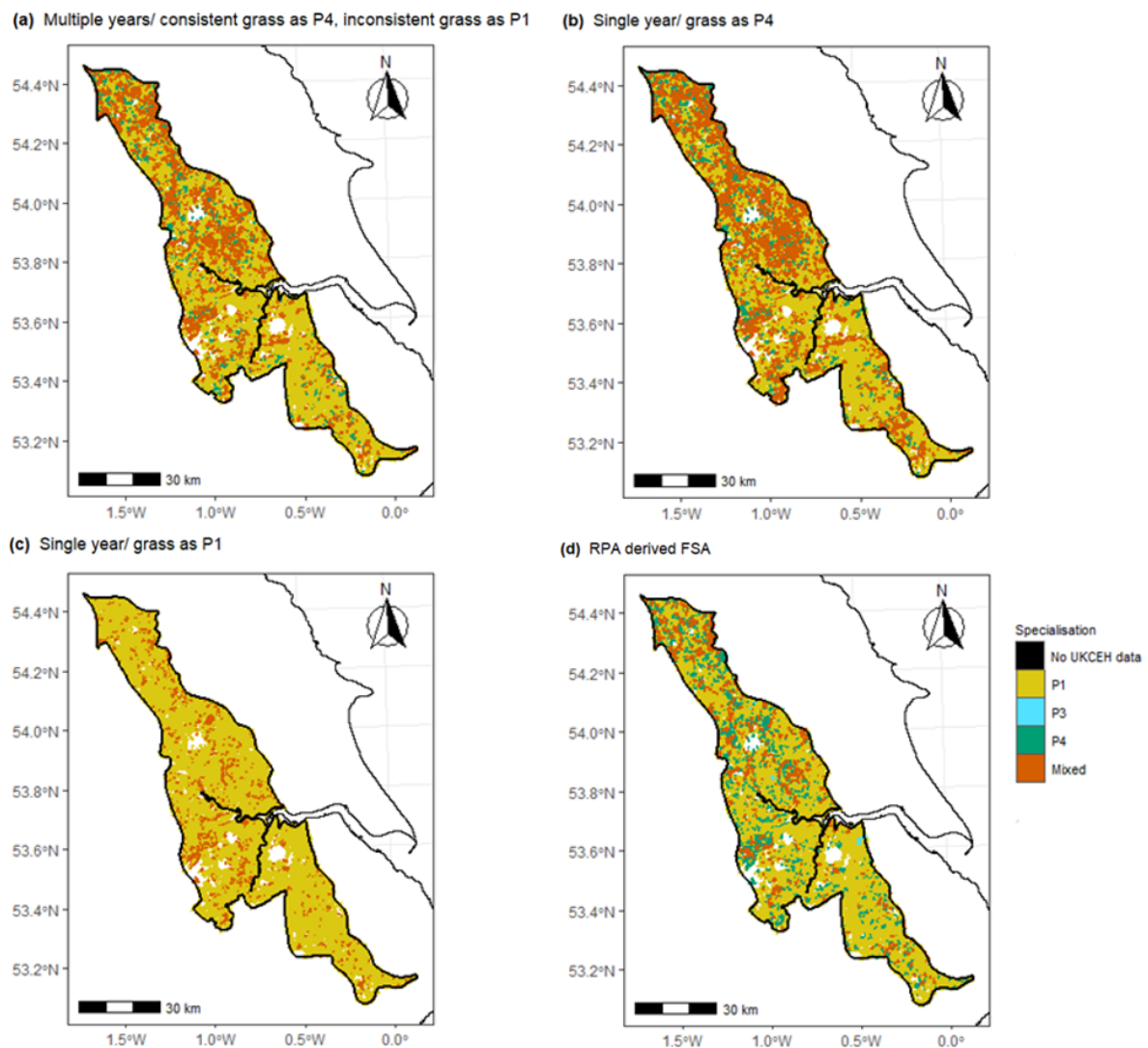

Figure 2: Maps showing the UKCEH derived FSA for scenarios on how crop types are matched to specializations; (a) multiyear data used to determine "grass" as either P1 (temporary) or P4 (permanent), (b) single-year "grass" classified as P4, and (c) single-year "grass" classified as P1, and (e) RPA derived FSA. Data source: UKCEH Land Cover® Plus: Crops © 2019 UKCEH, © RSAC, © Crown Copyright 2007, Licence number 100017572.

Table 2: The number of farms (out of 3528) that were correctly classified using three scenarios of analysis

UKCEH derived FSA scenario	FSA classified farms compared to RPA derived FSA	
	Correct	**Misclassified**
a) Multiple years/ consistent grass as P4, inconsistent grass as P1	2733 (77.5%)	795 (22.5%)
b) Single year/ grass as P4	2466 (70.0%)	1062 (30.0%)
c) Single year/ grass as P1	2204 (62.5%)	1324 (37.5%)

2.3. Serbian case study region: Crop type mapping from Sentinel-2 data

2.3.1. Objective

The Serbian case study region heavily relies on remote sensing products since LPIS is still not implemented to support agricultural policies. Instead of LPIS, data coming from AgroSens, a voluntary platform of digital agriculture in Serbia (https://agrosens.rs/#/app-h/welcome) is utilised. The platform enables monitoring of crops by combining processed Sentinel-2 and meteorological data (historical data and forecasts) and other ground information obtained through various sensors and devices. The main limitation of the platform comes from its coverage, which currently encompasses only one quarter of the farmers. It is expected that the number of farmers using AgroSens will grow as new free services are released. To overcome the limitations, remote sensing products are used to upscale available information. In this exercise, a pixel-based crop classification algorithm based on Sentinel-2 data is developed, validated, and applied to the case study area.

2.3.2. Methodology and results

Crop classification maps were generated for the whole case study area and are derived using a Random Forest classifier utilizing satellite data. Starting from 2017, maps were generated based on Sentinel-2 data. The classifier is pixel based, where pixels align with Sentinel-2 10x10 m pixels. All available cloud free images during the growing season were used as input data to train a classifier on collected ground truth data (~1000 parcels per season). Model is trained through a 10-fold cross validation procedure. Splits into training and validation are additionally constrained by taking into account that pixels corresponding to the same parcel are either placed into training or validation part to avoid overfitting. Furthermore splits into the training and validation areas are balanced in terms of distribution of parcel sizes and overall number of pixels.

Overall accuracy in different seasons is ~90% for the classification of 5 most common crops (maize, wheat, soybean, sugarbeet, sunflower) in the region of Vojvodina. Table 2 provides

obtained overall accuracy and precision for each crop across observed years. Accuracy depends on the number of available images (dates with cloud-free images) and their distribution during growing season and quality/variability of collected ground truth data.

Table 2: Crop classification 10-fold cross validation

Year	No. of dates	Accuracy of crop classification	Maize precision	Wheat precision	Soybean precision	Sugar Beet precision	Sunflower precision
2017	22	88.19 %	85.01 %	97.09 %	71.41 %	92.13 %	86.54 %
2018	23	92.18 %	92.35 %	94.21 %	91.59 %	96.65 %	79.88 %
2019	16	91.69 %	87.49 %	92.18 %	82.80 %	98.19 %	89.20 %
2020	36	97.45 %	98.56 %	97.95 %	97.69 %	98.61 %	91.65 %
2021	30	94.79 %	93.57 %	98.36 %	87.53 %	96.64 %	93.80 %

Figure 3 (a, b) shows a part of the pixel level crop classification map along with the corresponding post-processed map. It can be observed that the pixel based classification map has some errors due to within fields vegetation variability. These errors can be corrected through post-processing (e.g. majority voting of classified crop across pixels in field) if accurate production field boundaries are available (see also **Serra & Pons 2016)**. Figure 4 shows the whole region for which crop classification was performed and includes an additional class, oilseed rape, that was included in monitoring in the last season.
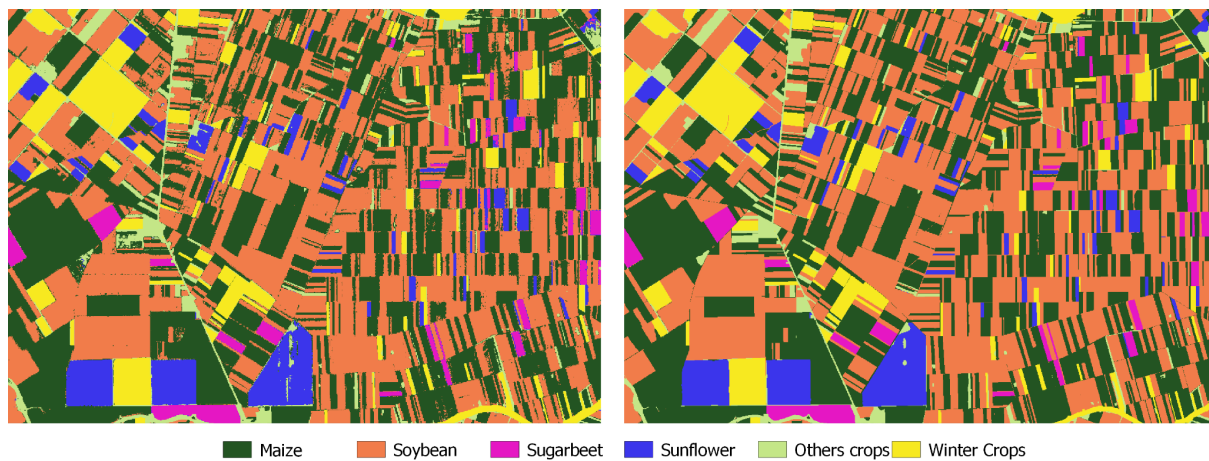

Maize   Soybean   Sugarbeet   Sunflower   Others crops   Winter Crops

Figure 3: (a) Pixel based classification (b) filtering based on fields boundaries from cadastral data
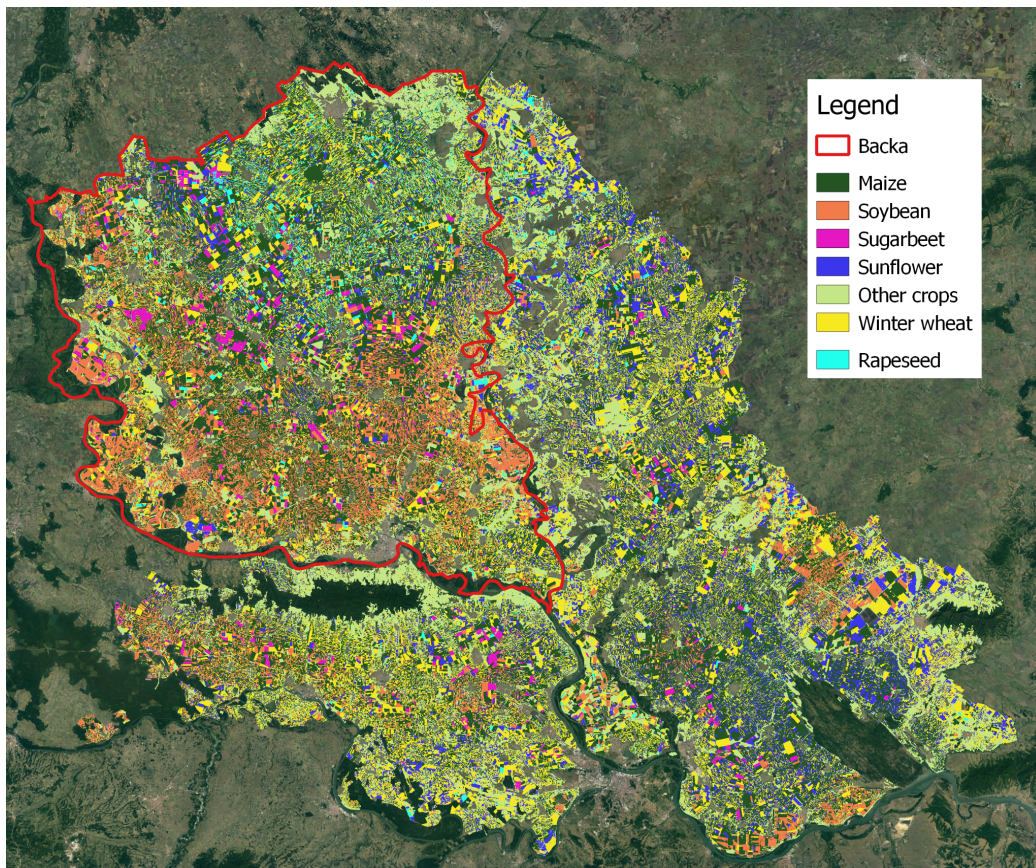
Figure 4: case study region crop classification map

2.4. Discussion and Outlook

We compared crop classification products based on remote sensing data to reference data from IACS/LPIS and assessed their suitability for the classification of FSA farm specialization. For the Czech case study region, the OneSoil crop classification was directly compared to crop type data from IACS/LPIS, resulting in poor agreement. For the UK case study region, comparison was performed on the FSA level by classifying FSA farm specialization based on a remote sensing based classification product (UKCEH Land Cover plus: Crops) and using the IACS/LPIS based FSA classification from D3.5 as reference. Here, agreement of up to 77.5% could be achieved when classification products from different years were combined to distinguish permanent from temporary grassland to assign farms to FSA specializations P1 (general cropping) or P4 (livestock farming). However it should be noted that no horticulture (P2) was present in the reference FSAs of the study area and only 11 farms represented permanent crops (P3), none of which were classified correctly using UKCEH data.

From this we conclude that remote sensing based crop type classifications can be used to give an indication of FSA farm specialization in a case study region, provided the classification product is optimized for the region of interest: The Europe-wide OneSoil product showed insufficient accuracy for this purpose. The analysis of the Serbian case study region confirms this assumption as high crop classification accuracies (~90%) can be achieved using Sentinel-2 time-series data in a machine learning approach within a limited

area. Further research is needed into the applicability of crop classification data for FSA derivation on larger scales than case study areas. As more large-scale crop classification datasets become available (e.g. **d'Andrimont et al. 2021**), effort should be put in analysing their quality and suitability for the BESTMAP project. Finally, accurate boundaries of parcels and how these parcels aggregate into farms will still need to be known to allow FSAs to be derived, regardless of the crop data used (see also Section 4).

3. Crop Yield Mapping

Crop yield is an essential agricultural variable that can be estimated using remote sensing data because it is related to biophysical parameters (i.e. above-ground-biomass (AGB)). Crop yield is relevant to BESTMAP as a measure of the economic performance of a farm, and provides information on the economic dimension for mapping FSAs.

Two analyses were performed focusing on different regions of interest and yield mapping approaches:

- FADN region 412 (East England): Sentinel-2 time-series data was analysed to assess the suitability of remote sensing indicators for yield estimation. This analysis was performed on an entire FADN region instead of a BESTMAP case study area because of better availability of reference yield data. (Section 3.1)
- For the Serbian case study region, an experimental yield mapping product is generated based on soil and weather data. (Section 3.2)

3.1. Remotely sensed phenological indicators for yield estimation

3.1.1. Objective

The goal of this activity is to identify remote sensing indicators from a time-series of Sentinel-2 data that can be used to give an estimate of yield distribution at a FADN region scale. Specifically, this analysis focuses on FADN region 412 (East England) as an example, as yield reference data is available for this region. However, the general availability of yield reference data at a field level cannot be assumed for a region, crop type, or time of interest. Thus, this exercise investigates whether a general relationship between remote sensing indicators and yield can be established. The goal is not to get accurate yield estimates for individual fields (even though the analysis is performed on field level), but to find patterns of high/low yield on landscape scale by identifying the best correlated proxy variables.

Using a Sentinel-2 time-series, growing season vegetation index metrics can be calculated and assigned to individual fields. As reference yield data is available on field level, the most suitable predictor metric can be identified. Field-average yield data for an area of England covered by FADN region 412 and held by UKCEH is used for this exercise.

3.1.2. State of the art agricultural yield mapping from optical remote sensing data

Existing approaches to estimate yield from optical remote sensing data typically employ a metric derived from a vegetation index or a set of indices. Most vegetation indices make use

of wavelengths where absorption by chlorophyll is greatest (i.e. visible red light), vegetation "greenness" can be assessed (i.e. visible green light), or the leaf cell structure leads to strong reflection (i.e. near infrared) (**Zhao et al. 2020**). Apart from the well-known normalised difference vegetation index (NDVI, e.g. used by **Skakun et al. 2018)**, further indices such as green normalised difference vegetation index (GNDVI, **Rahman & Robson 2020**), green chlorophyll vegetation index (GCVI, **Jain et al. 2016**), optimised soil-adjusted vegetation index (OSAVI, **Zhao et al. 2020**), or enhanced vegetation index (EVI, **Shammi & Meng 2021**) are applied in the literature for yield mapping. Table 3 summarizes these most popular indices. The leaf area index (LAI) is also widely used for yield estimation (e.g. **Baez-Gonzalez et al. 2005**), but it represents a biophysical variable itself and as such cannot be directly derived from spectral bands without modelling. Existing algorithms for Sentinel-2 based LAI estimation appear to be prone to errors on field level (**Kganyago et al. 2020**), so the LAI is not included in this analysis.

Table 3: Vegetation indices used for yield mapping. R,G,B,NIR refer to the red, green, blue, and near-infrared bands respectively.

Index	Formula	Reference
Normalized difference vegetation index	$NDVI = (NIR - R)/(NIR + R)$	Rouse et al. (1973)
Green normalised difference vegetation index	$GNDVI = (NIR - G)/(NIR + G)$	Gittelson et al. (1996)
Green chlorophyll vegetation index	$GCVI = (NIR/G) - 1$	Gittelson et al. (2003)
Optimized soil-adjusted vegetation index	$OSAVI = 1.16 \times (NIR - R)/(NIR + R + 0.16)$	Rondeaux & Steven (1995)
Enhanced vegetation index	$EVI = 2.5 \times (NIR - R)/(NIR + 6 \times R - 7.5 \times B + 1)$	Huete et al. (2002)

Indices from a single satellite acquisition may contain information related to yield, but this requires the acquisition to be cloud-free and taken in the peak of the respective growing season. It is more beneficial and established practice to analyse a time-series of indices covering the entire growing season (**Hunt et al. 2019**). The relatively high temporal (~5 days revisit time in central Europe) and spatial (10 m for visible and NIR bands) resolution of Copernicus Sentinel-2 data makes this satellite system a preferable choice for yield mapping at field (or even within-field) level (**Drusch et al. 2012**). The high temporal resolution is especially important in areas where cloud coverage is typical in the summer months (such as central and northern Europe), as even a 5-day revisit time may yield only one cloud-free acquisition per month.

Different metrics are applied in existing studies, but most approaches find the maximum value of an index time-series to be best correlated with reference yield data: **Zhao et al.** (2020) establish a linear relationship between Sentinel-2 based peak NDVI and OSAVI values and yield at field scale for wheat in Eastern Australia. **Shammi & Meng** (2021) analyse a set of crop growth metrics extracted from MODIS NDVI and EVI time-series and their correlation with soybean yield at a county level in Mississippi, USA. **Rahman &**

Robson (2020) use linear regression to create sugarcane yield prediction models at block level in Eastern Australia based on combined Landsat-8 and Sentinel-2 GNDVI time-series. They find very high correlation between the maximum GNDVI and yield when the time-series are temporally densely sampled and exactly aligned to the growing period. **Jain et al.** (2016) apply the maximum GCVI for wheat yield mapping at field level in India from SkySat data, while **Skakun et al.** (2017) use the maximum NDVI from a Landsat-8/Sentinel-2 time-series to estimate wheat yield at regional scale in Ukraine. Several studies also enrich the modelling with environmental data to account for plant stress, which improves the yield estimation, but the availability of such data at the required spatial and temporal resolution is usually limited (**Hunt et al. 2019, Zhao et al. 2020**).

All such studies rely on the availability of reference yield data (either on field or within-field level) that is collected during harvest and/or reported by the farms. A subset of reference yield data is used to parameterize a model (e.g. fit a regression model or train a machine learning algorithm), while the rest is used for cross-validation.

3.1.3.   Generation of a Sentinel-2 based indicator time series

For FADN region 412, we downloaded all available Sentinel-2 data from 2017 to 2019 with a cloud coverage < 60%. As datasets older than 12 months are moved to the long-term archive on ESA's Copernicus Open Access Hub, all data were instead downloaded from Google Cloud Storage ([Link to Sentinel-2 resource](#)). Preferably, bottom-of-atmosphere Level-2A reflectance data were downloaded. For acquisition dates before April 2018 no Level-2A data were available, and top-of-atmosphere Level-1C reflectance data were retrieved instead. All Level-1C datasets were then atmospherically corrected using ESA's Sen2Cor software suite, in order to be consistent with the rest of the data. All following data processing was performed using the free and open source GIS software GRASS GIS (**Neteler et al. 2012**). A multitemporal datastack (Space-Time-Raster Dataset) was generated from all individual Sentinel-2 scenes. Cloud detection was then applied using the methodology from **Parmes et al.** (2017) (see also the [relevant GRASS GIS implementation](#)) - all clouds were removed from the individual time-steps. As they are found to be best correlated with yield in literature, the indices NDVI and GNDVI are calculated and cloud free index time-series are generated.

Depending on the crop type, we assumed different growing seasons. For winter crops planted in autumn or early winter (such as winter wheat, winter barley, oilseed rape) we defined the growing season between December and July, while for spring crops (e.g. spring wheat, spring barley) it was set between March and August. The [USDA crop calendars](#) as well as spatially relevant yield mapping literature (e.g. **Hunt et al. 2019**) were used as sources to define the growing seasons for East England.

Cloud-free pixel availability varied significantly across FADN region 412 as Figure 5 a) shows. This was due to different cloud occurrence patterns as well as the Sentinel-2 acquisition geometry (the latter explaining the large diagonal strip of low coverage and tile overlaps).
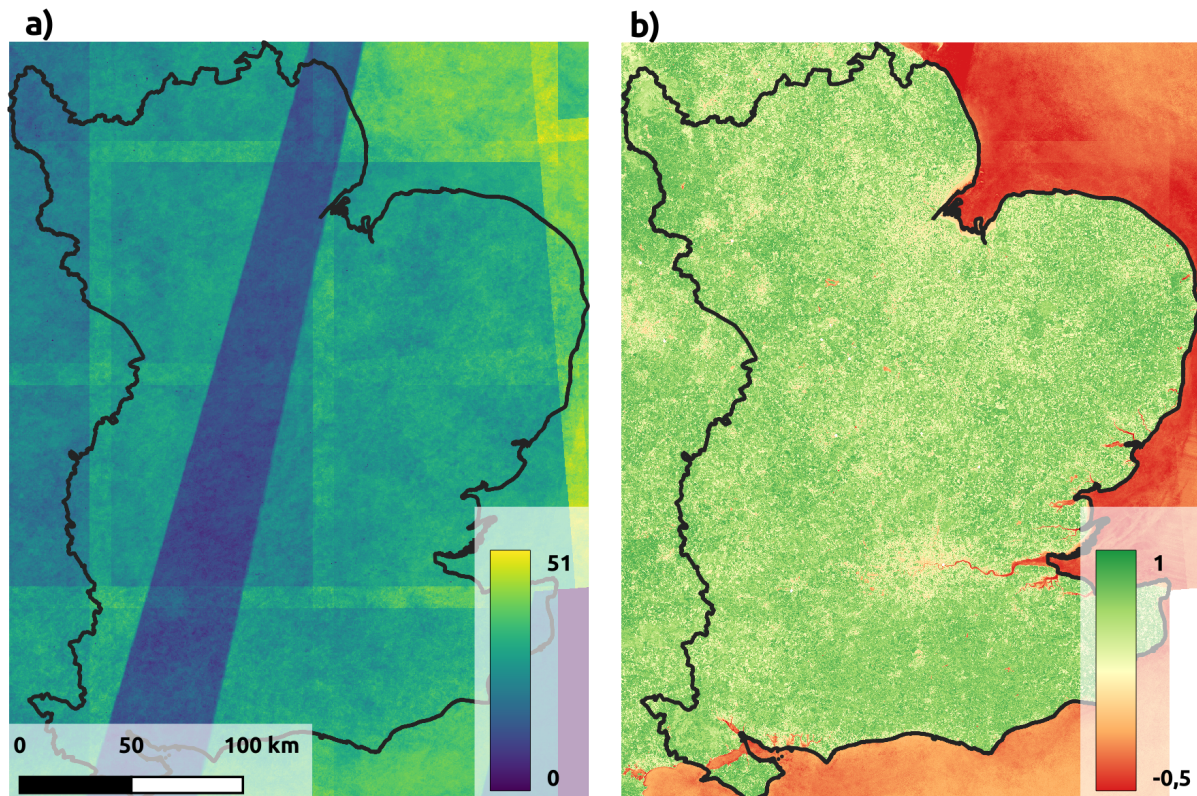
Figure 5: Statistics of growing season 12/2018 to 07/2019 for Southeast England (FADN region 412). **a)** Number of cloud-free pixels, **b)** median value of NDVI

We used the index time series to calculate a number of metrics for the different growing seasons ([maximum, median, third quartile, integral, average of individual months] of both NDVI and GNDVI time-series during growing season). These individual statistics were then averaged at a field level using UKCEH Land Cover plus Crops data of the respective year (see e.g. https://www.ceh.ac.uk/services/ceh-land-cover-plus-crops-2015).

Using winter wheat for the growing season 12/2018 to 07/2019 as an example, we assessed the influence of the large variability of cloud-free pixel availability on the different index time-series metrics.
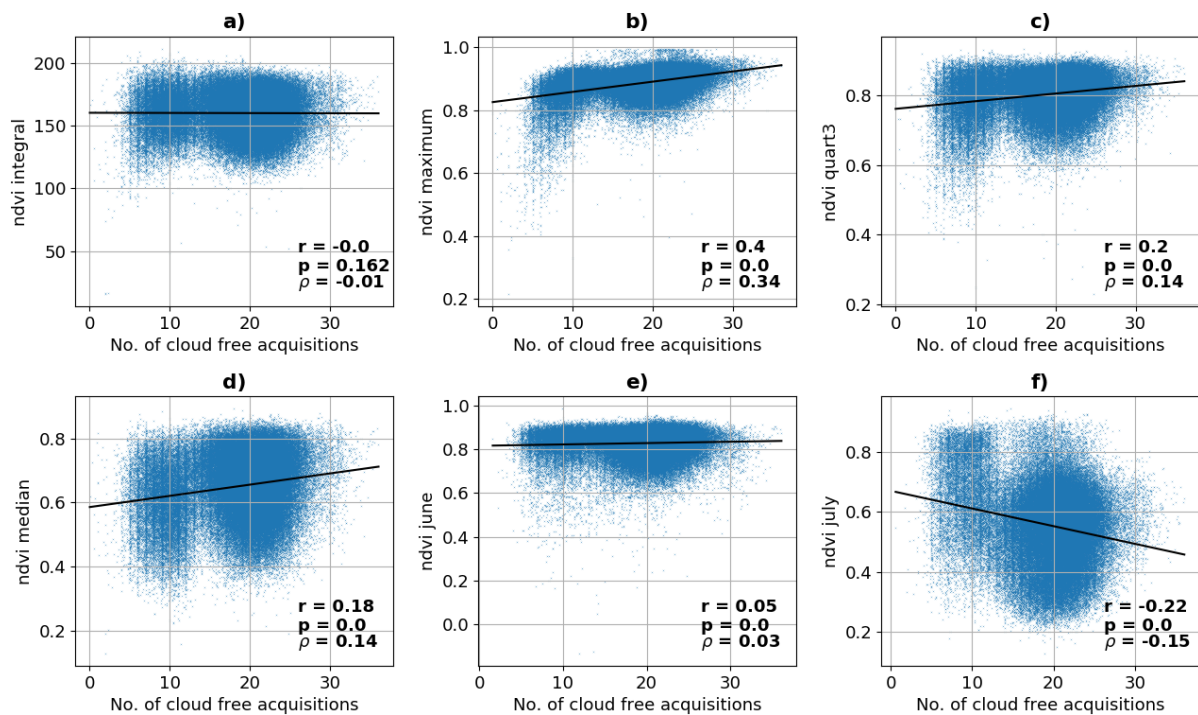
Figure 6: Relationship between the number of cloud-free acquisitions per winter wheat field in 2019 on different index time-series metrics

From Figure 6 it can be seen that this influence is rather low and no correlation exists between the number of cloud-free acquisitions and e.g. the integral of the NDVI time-series (Figure 6 a). However, an effect exists for the maximum value of the NDVI time-series (Figure 6 b). This is logical, because there is a higher chance to capture the point in time with the highest potential in-situ NDVI if the time-series has more temporal samples. As the maximum value is the most suitable predictor for yield as identified from literature, the number of available cloud-free acquisitions will likely have an influence on yield estimation and only fields with a minimum number of ten acquisitions were included in the analysis.

3.1.4. Yield correlation analysis

The index time-series metrics linked to each winter wheat field of 2019 were combined with field-average yield data held by CEH to assess the suitability of each metric as predictor of yield. These field data were gathered in the ASSIST project as described in Hunt et al. (2019, see also https://assist.ceh.ac.uk/). In brief, optical yield sensors that were attached to GPS-fitted combine harvesters measured crop yields during the crop harvest of the relevant year (i.e. 2019). Harvested grain passed through the optical beam of the sensor, enabling determination of grain volume. The sensors also recorded accurate location, grain moisture content, time, and machinery details (e.g. speed). These yield data were voluntarily submitted by farmers using the CLAAS Telematics cloud platform (www.claas-telematics.com). After submission, the raw crop yield data were cleaned following the standard cleaning procedures for precision yield in Muhammed, Milne, Marchant, Griffin, & Whitmore (2016) and Fincham et al. (*in prep*), which removed unrealistic yield data points (e.g. excessively high yields, outside field boundaries). Where more than 1 combine harvester was involved in collecting yield from the same field, yield values were

standardised to the mean of all relevant harvesters, to account for any differences in calibration. Subsampling of the data was used to ensure a consistent data resolution across all fields.

Harmonization was required between the field-level remote sensing index statistics that were calculated based on the CEH Land Cover plus Crops fields and the ASSIST dataset holding information on yield. Differences in the polygon boundaries occurred between both datasets and the percentage of overlap was calculated as an additional variable. Where polygons overlapped less than 50%, they were omitted. Table 4 summarizes the linear regression results for different remote sensing metrics and winter wheat yield for the reference year 2019. Overall, NDVI and GNDVI metrics performed very similarly, with slightly better performance of the GNDVI. Because of this, only GNDVI metrics are shown in Table 4.

Table 4: Correlation analysis of different GNDVI metrics and yield for winter wheat 2019 (n = 1233)

index metric	R^2	p-value
GNDVI median	0.0	0.44
GNDVI 3. quartile	0.077	<0.001
GNDVI integral	0.007	0.002
GNDVI maximum	0.182	<0.001
GNDVI june	0.058	<0.001
GNDVI july	0.022	<0.001

While several metrics showed significant correlation with yield, only the R^2 of the GNDVI maximum appeared promising enough for further analysis. This is consistent with current findings in literature as described in section **3.1.2**. Figure 7 shows the regression results for the GNDVI maximum. Using the number of cloud-free acquisitions as a weight to linear regression (Figure 7 B) improves the coefficient of determination (R^2) to 0.232. This might be expected as more cloud-free acquisitions would increase the chance of obtaining the true maximum value. We also investigated whether the percentage of spatial overlap between the CEH Land Cover plus Crops based remote sensing index product and the ASSIST yield dataset affects the regression but found no effect (Figure 7 C). It should also be noted that there was no indication of non-linearity in these relationships and adding a quadratic term did not improve R^2 values (analyses not shown).

A WW yield (t/ha) in 2019

$R^2 = 0.182$

B
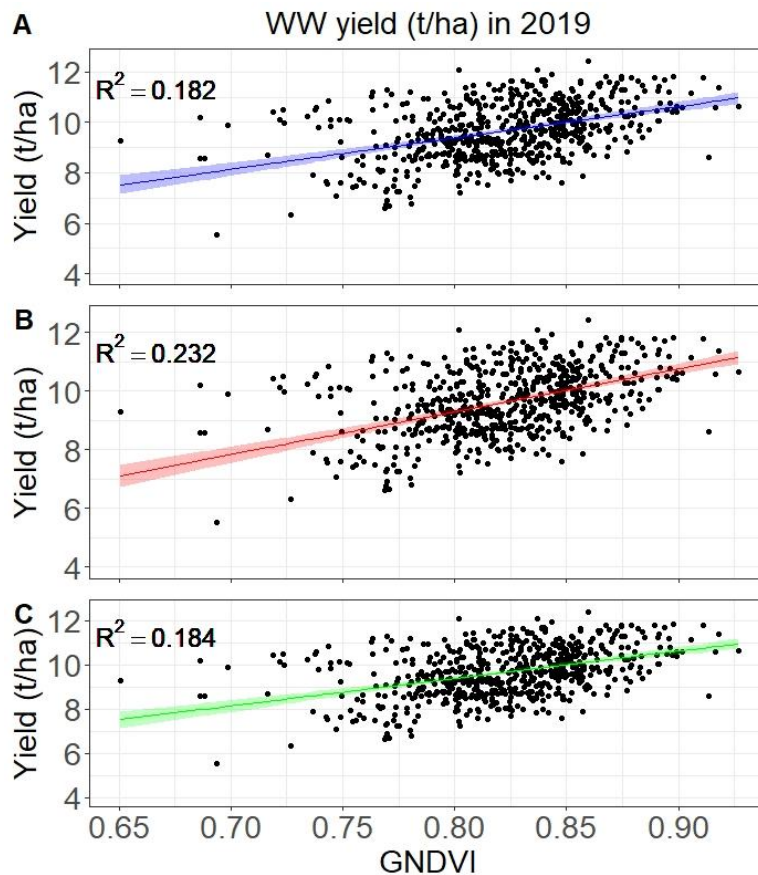
$R^2 = 0.232$

C

$R^2 = 0.184$

GNDVI

Figure 7: Linear regression between the field-level maximum GNDVI value of the 12/2018 - 07/2019 growing season of winter wheat (WW) and yield using no weighting variable (A), using the number of cloud-free acquisitions as a weighting variable (B), using the spatial overlap between fields assigned with maximum GNDVI values and ASSIST yield polygons as a weighting variable (C).

We then calculated the maximum GNDVI values for different years (2017, 2018, 2019) and different crops (winter wheat, winter barley, oilseed rape) and created linear regression models with the ASSIST yield data in order to analyse the stability of the maximum GNDVI - yield relationship. The results are summarized in Table 5.

Table 5: Correlation analysis of maximum GNDVI for different crop types and years.

Crop Type	year	R^2	R^2 (weighted by no. of acquisitions)	Spearman's ρ
winter wheat	2017	0.03	0.076	0.259
winter wheat	2018	0.315	0.31	0.557
winter wheat	2019	0.182	0.231	0.428
winter barley	2017	0.261	0.267	0.511
winter barley	2018	0.198	0.247	0.472
winter barley	2019	0.216	0.205	0.488
oilseed rape	2017	0.159	0.227	0.453
oilseed rape	2018	0.141	0.138	0.357
oilseed rape	2019	0.209	0.224	0.46

_____

Except for winter wheat in 2017, the coefficients of determination (R^2) range from 0.141 to 0.315 and tend to increase when the number of available acquisitions is used as a weighting factor. This is a weak to moderate correlation that is nonetheless significant ($p<0.05$) for all tested crop types and years. The Spearman rank correlation coefficient ρ varies around 0.5, indicating that the maximum GNDVI value might be used as a proxy variable for yield mapping at large scale, if information on the general (non-spatial) distribution and range of yield is available for a given region of interest.

3.2. Yield mapping based on soil and weather data

3.2.1. Objective

For the Serbian case study region Bačka, a different yield mapping approach was tested complementary to the remote sensing analysis presented in the previous section. The objective was to utilize soil and weather data to create experimental yield maps. Reference yield data is used from the AgroSens platform and additionally through data obtained from several agricultural producers across the study region (see section 2.3.1) under signed NDA.

3.2.2. Methodology and results

Yield mapping for the Serbian case study region is derived from a trained machine learning model based on a random forest regression algorithm. The model was trained on available ground truth data (~2000 yield records on field level) collected through numerous NDA agreements. We used 10-fold cross validation to select model parameters and evaluate model performance. The input data for the model were obtained from open sources dedicated to soil (SoilGrids) and weather data (Copernicus Climate Store). Feature engineering on weather data was performed before the machine learning part. Models were developed for five of the most common crops in the region (maize, wheat, soybean, sugarbeet, sunflower). Figure 8 shows the yield map of maize for the case study region with yield variability of 7.5 t/ha to 12.2 t/ha. The spatial resolution is constrained by the resolution of weather data, which is 0.25° x 0.25°. The best performing model was used to generate the map for the case study region. Perspectively, the experimental yield estimation maps will be combined with satellite based vegetation indices to derive correlations and select best indices for upscaling the study to other regions.
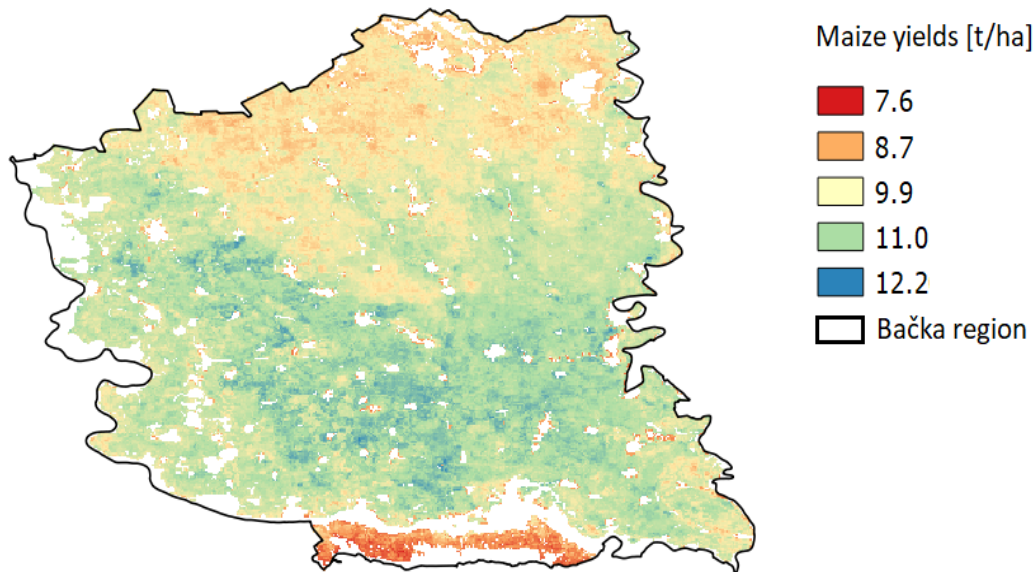
Figure 8: Yield estimates for case study Serbia (Bačka region) for Maize, where red to green colors indicate lower to higher yields

3.3. Discussion and Outlook

Two analyses with different focus were conducted with regard to yield mapping: One in-depth analysis of remotely sensed vegetation indicators and their relationship to yield and the generation of an experimental yield product based on soil and weather data. The first approach tries to overcome the need for exhaustive reference data and finds a stable, but moderate correlation of the maximum GNDVI and yield. The second approach depends heavily on reference yield data and has a limited spatial resolution, but can produce large-scale yield estimates if input data is available.

From these experimental results, we plan to investigate further one or both of two potential approaches for yield mapping summarized in the following. For both approaches, basic information on the general statistics of yield for a given crop type and region are required such that a general (non-spatial) distribution of yield can be inferred. This can be extracted from the FADN sample of farms, and the analysis will be resumed once this dataset will be made available.

- The advantages of both tested approaches can be combined: For FADN region 412, the existing regression models could be enriched with auxiliary predictors, such as climatological or topographical variables (as successfully applied to the Serbian case study region). The goal would be to find a model relationship between a variable set and yield that is stable enough to be applied with a minimum or no yield reference data. FADN data could then be used to validate this approach.
- Alternatively, if the resolution, accuracy, and/or availability of auxiliary variables prove to be insufficient for the application across case study regions, using a remote sensing indicator as single predictor could be further pursued: As the rank correlation

_____

between yield and maximum GNDVI shows, a monotonic and linear relationship exists between the two variables. As a consequence, we can apply quantile-quantile matching to map the theoretical distribution of yield (derived from FADN) and the distribution of maximum GNDVI values at a field level together, effectively assigning a yield value to each field polygon (This has the simplifying assumption that FADN yield distribution on farm level can be applied to field level). Finally we can use the ASSIST dataset to validate the resulting yield maps.

Yield estimates will further help in defining more precise FSA types that are initially derived from farm size, grown crops and average yield for the observed region. As the yield map in Figure 8 demonstrates high variability of yields across the region we can expect that spatial estimates of the yields will improve FSA categorization. Figure 9 shows the current FSA mapping for available farms in the Serbian case study region into 4 categories based on economic size. Each group can be further refined with estimates on farm productivity derived from yield maps and vegetation indices.
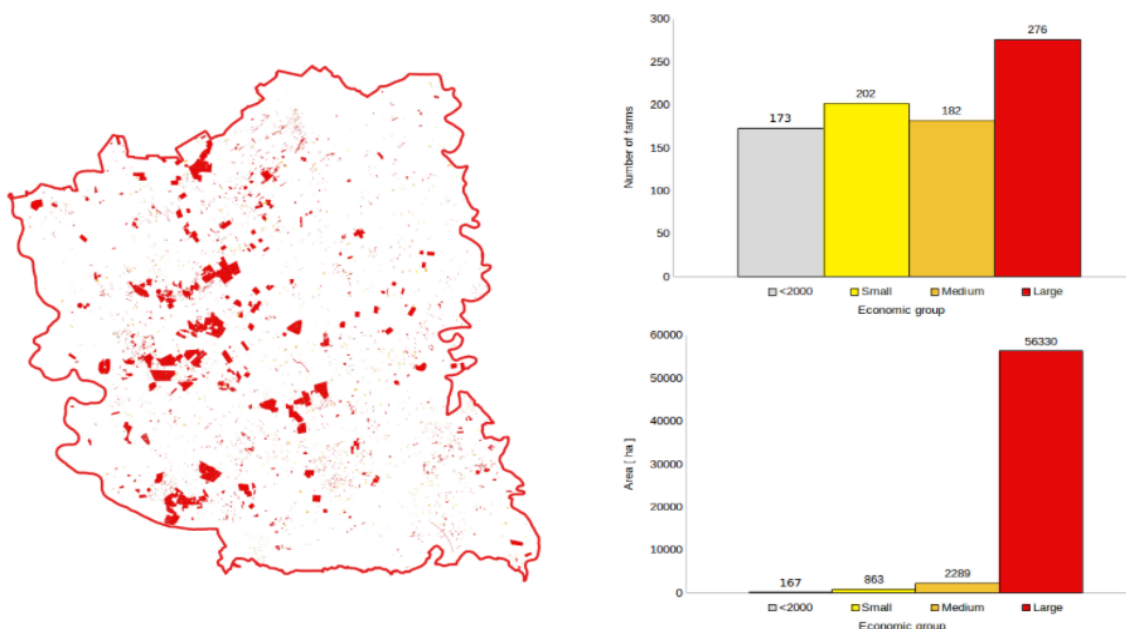

Figure 9: Available farms in case study Serbia (Bačka region) across 4 economic groups (<2000, small, medium and large), their counts and sizes

4. Field boundary mapping

4.1. Objective

As the previous analyses show, remote sensing methods for agricultural applications rely on consistent field (or parcel) boundaries in order to attribute the calculated parameters to the correct geographical unit. For most of the case study areas, LPIS data providing such information is available, but even there inconsistencies in the definition of 'fields' occur across countries or administrative units (see D3.5). In countries where LPIS is not available,

such as Serbia, field boundary data may be available from the respective administration, but accuracy, definitions, and update cycles can vary. Because of this, a remote sensing approach is proposed to delineate up-to-date field boundaries of individual years based on Sentinel-2 imagery. This has the advantage that a developed methodology could be applied to any region of interest with sufficient data coverage to produce consistent maps of production parcels. A prototypical approach is currently being drafted for the Serbian case study region.

4.2. Envisioned approach

Field boundaries for the Serbian case study region are available from the Republic Geodetic Authority (https://a3.geosrbija.rs/). Official cadastral data in Serbia are not updated regularly and there are discrepancies in the records with regard to the actual situation in the field. Agglomeration of the fields is usually not accompanied by changes in the field boundaries. Figure 10 demonstrates mismatches between cadastral parcels and production parcels.
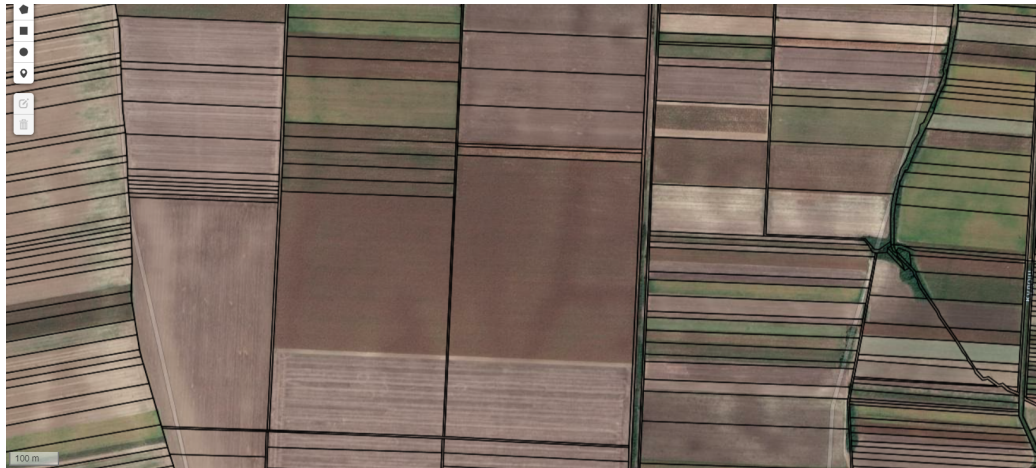

Figure 10: Mismatch between cadastral parcels and real production parcels

To improve the monitoring of agricultural production, solutions for delineating field boundaries using satellite imagery and machine learning approaches are necessary. Development of such solutions will enable reliable and up-to-date field-level information. A promising approach for the extraction of agricultural parcels and their boundaries is based on deep learning techniques by processing multispectral Sentinel-2 satellite images (**Masoud et al. 2020**). The envisioned solution for the Serbian case study area uses a multi-task based convolutional neural network (CNN) on Sentinel-2 data. Figure 11 illustrates the architecture of the solution.
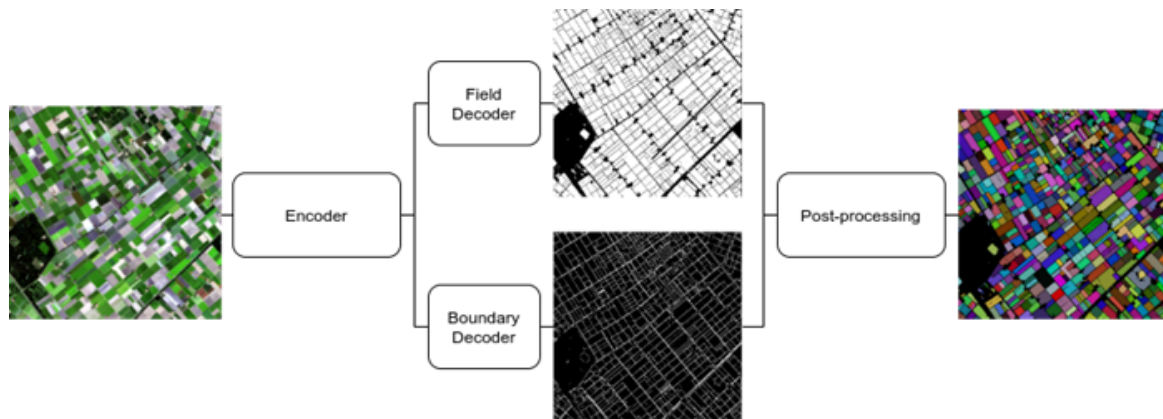
Figure 11: Multi-task convolutional neural network for detecting production parcels

The proposed architecture jointly learns two semantically similar tasks: Boundary delineation and segmentation of parcels. Both tasks have shared representations in shallower layers of the network and then the architecture splits into two branches for separate tasks. The final step performs post-processing for connecting the partial boundaries. This multi-task network outperforms a single CNN dedicated to boundary delineation only (**Masoud et al. 2020**). The precision of the obtained result in terms of detecting boundaries is 72% and in terms of spatial agreement between segmented and real parcels it is 89%. The influence of the parcel sizes on the detection accuracy will be assessed in a following step.

5. Acknowledgements

6. References

Baez-Gonzalez, A.D.; Kiniry, J.R.; Maas, S.J.; Tiscareno, M.L.; Macias C.; Mendoza, J.L.; Richardson, C.W.; Salinas G.; Manjarrez, J.R.: *Large-Area Maize Yield Forecasting Using Leaf Area Index Based Yield Model*. Agronomy Journal 97, **2005**.

d'Andrimont, R.; Verheggen, 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 (Preprint)*. Remote Sensing of Environment 266, **2021**.

Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; Meygret, A.; Spoto, F.; Sy, O.; Marchese, F.; Bargellini, P.: Sentinel-2: *ESA's optical high-resolution mission for GMES operational services*. Remote Sensing of Environment 120, **2012**.

Gittelson, A.A.; Kaufman, Y.K.; Merzlyak, M.N.: *Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS*. Remote Sensing of Environment 58, **1996**.

Gittelson A.A.; Vina, A.; Arkebaur, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B.: *Remote Estimation of Leaf Area Index and Green Leaf Biomass in Maize Canopies*. Geophysical Research Letters 30, **2003**.

Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G.: *Overview of the radiometric and biophysical performance of the MODIS vegetation indices*. Remote Sensing of the Environment 83, **2002**.

Hunt, M.L.; Blackburn, G.A.; Carrasco, L.; Redhead, J.W.; Rowland, C.S.: *High resolution wheat yield mapping using Sentinel-2*. Remote Sensing of Environment 233, **2019**.

Jain, M.; Srivastava, A.K.; Balwinder-Singh; Joon, R.K.; McDonald, A.; Royal, K.; Lisaius, M.C.; Lobell, D.B.: *Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data*. Remote Sensing 8(10), 860, **2016**.

Kganyago, M.; Mhangara, P.; Alexandridis, T.; Laneve, G.; Ovakoglou, G.; Mashiyi, N.: *Validation of Sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape*. Remote Sensing Letters 11(10), **2020**.

Land Cover plus: Crops, [FileGeoDatabase geospatial data], Scale 1:2500, Tiles: GB, Updated: 22 November 2019, CEH, Using: EDINA Environment Digimap Service, <https://digimap.edina.ac.uk>, Downloaded: 2021-04-23 16:35:55.975, **2019**.

Masoud K.M.; Persello C.; Tolpekin V.A.; *Delineation of agricultural field boundaries from Sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks*. Remote sensing. 12(1):59, **2020.**

Muhammed, S.; Milne, A.; Marchant, B.; Griffin, S.; Whitmore, A.: Exploiting Yield Maps and Soil Management Zones. AHDB Project Report No. 565, **2016.**

Neteler, M.; Bowman, M., Landa, M., Metz, M.: GRASS GIS: *A multi-purpose open source GIS*. Environmental Modelling & Software 31, **2012**.

Parmes, E.; Rauste, Y.; Molinier, M.; Andersson, K.; Seitsonen, L.: Automatic Cloud and Shadow Detection in Optical Satellite Imagery without using Thermal Bands - Application to Suomi NPP VIIRS Images over Fennoscandia. Remote Sensing 9(8), 806, **2017**.

Rahman, M.M.; Robson, A.: *Integrating Landsat-8 and Sentinel-2 Time Series Data for Yield Prediction of Sugarcane Crops at the Block Level*. Remote Sensing 12(8), 1313, **2020**.

_____

Redhead, John, *CEH Land Cover® plus Crop Map: Quality Assurance* [WWW Document]. URL https://www.ceh.ac.uk/contact-us (accessed 09.08.21), **2018**.

Rondeaux, G.; Steven, M.; Baret, F.: *Optimization of soil-adjusted Vegetation Indices*. Remote Sensing of Environment 55, **1996**.

Rouse, J.W.; Hass, R.H.; Schell, J.A.; Deering, D.W.: *Monitoring Vegetation Systems in the Great Plains with ETRS*. Proc. of Third Earth Resour. Technol. Satell. Symp. **1973**.

Serra, P.; Pons, X.: *Uncertainty visualization of remote sensing crop maps enriched at parcel scale: a contribution for a more conscious GIS dataset usage*. Journal of Maps 12:5. **2016**.

Shammi, S.A.; Meng, Q.: Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling. Ecological Indicators 121, **2021**.

Skakun, S.; Vermote, E.; Roger, J.C.; French, B.: *Combined use of Landsat-8 and Sentinel-2A images for winter crop mapping and winter wheat yield assessment at regional scale.* AIMS Geoscience 3(2), 164, **2017**.

UK Centre for Ecology and Hydrology, n.d. *UKCEH Land Cover® plus: Crops* [WWW Document]. URL https://www.ceh.ac.uk/services/ceh-land-cover-plus-crops-2015 (accessed 09.08.21), **2019**.

Zhao, Y.; Potgieter, A. B.; Zhang, M.; Wu, B.; Hammer, G.L.: *Predicting Wheat Yield at the Field Scale by Combining High-Resolution Sentinel-2 Satellite Imagery and Crop Modelling*. Remote Sensing 12(6), 1024, **2020**.