Ref. Ares(2019)7863369 - 20/12/2019

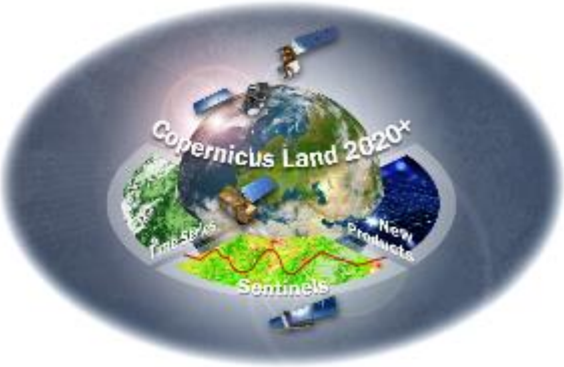
Horizon 2020

Space Call - Earth Observation: EO-3-2016: Evolution of Copernicus services

Grant Agreement No. 730008

ECoLaSS

Evolution of Copernicus Land Services based on Sentinel data



D14.1

"D44.1b – Prototype Report: Crop Area and Crop Status/Parameters (Issue 2)"

Issue/Rev.: 2.0

Date Issued: 20.12.2019

submitted by:

UCL Université catholique de Louvain

in collaboration with the consortium partners:

SIRS, JOANNEUM RESEARCH DIGITAL, GAF AG (an e-GEOS (ASI / Telespazio) Company), DLR

submitted to:

European Commission – Research Executive Agency

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 730008.

Consortium Partners

No.	Participant Organisation Name	Short name	City, Country
1	GAF AG	GAF	Munich, Germany
2	Systèmes d'Information à Référence Spatiale SAS	SIRS	Villeneuve d'Ascq, France
3	JOANNEUM RESEARCH Forschungsgesellschaft mbH	JR	Graz, Austria
4	Université catholique de Louvain, Earth and Life Institute (ELI)	UCL	Louvain-la-Neuve, Belgium
5	German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Wessling	DLR	Wessling, Germany

Contact:

GAF AG
Arnulfstr. 199 – D-80634 München – Germany
Phone: ++49 (0)89 121528 0 – FAX: ++49 (0)89 121528 79
E-mail: copernicus@gaf.de – Internet: www.gaf.de

Disclaimer:

Document Release Sheet

	Name, Function	Date	Signature
Author(s):	Pierre Defourny (UCL) Brigitte Bedoret (UCL) Nicolas Bellemans (UCL) Nicolas Deffense (UCL) Thomas De Maet (UCL) Benjamin Goffart (UCL) Diane Heymans (UCL) Marie-Julie Lambert (UCL) Philippe Malcorps (UCL) Matthieu Taymans (UCL) Kathrin Schweitzer (GAF) Christophe Sannier (SIRS) Sophie Villerot (SIRS) Antoine Masse (SIRS)	13.12.2019	
Review:	Regine Richter (GAF) Sophie Villerot (SIRS)	18.12.2019	R. Richter
Approval:	Eva Sevillano Marco (GAF)	20.12.2019	
Acceptance:	Massimo Ciscato (REA)		
Distribution:	public		

Dissemination Level

Dissemination Level		
PU	Public	X
CO	Confidential: only for members of the consortium (including the Commission Services)	

Document Status Sheet

Issue/Rev	Date	Page(s)	Description / Changes
1.0	10.08.2018	71	First issue of the Prototype Report: Crop Area and Crop Status/Parameters.
2.0	20.12.2019	112	Second issue of the Prototype Report: Crop Area and Crop Status/Parameters.

Applicable Documents

ID	Document Name / Issue Date
AD01	Horizon 2020 Work Programme 2016 – 2017, 5 iii. Leadership in Enabling and Industrial Technologies – Space. Call: EO-3-2016: Evolution of Copernicus services. Issued: 13.10.2015
AD02	Guidance Document: Research Needs Of Copernicus Operational Services. Final Version issued: 30.10.2015
AD03	Proposal: Evolution of Copernicus Land Services based on Sentinel data. Proposal acronym: ECoLaSS, Proposal number: 730008. Submitted: 03.03.2016
AD04	Grant Agreement – ECoLaSS. Grant Agreement number: 730008 – ECoLaSS – H2020-EO-2016/H2020-EO-2016. Issued: 18.10.2016
AD05	D7.2 : D32.1b - Methods Compendium: Time Series Preparation (Issue 2), Issued: 15.05.2019
AD06	D8.2: D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification, Issue 2.0. Issued: December 2019

EXECUTIVE SUMMARY

The Horizon 2020 (H2020) project, “Evolution of Copernicus Land Services based on Sentinel data” (ECoLaSS) addresses the H2020 Work Programme 5 iii. Leadership in Enabling and Industrial technologies - Space, specifically the Topic EO-3-2016: Evolution of Copernicus services. ECoLaSS is being conducted from 2017–2019 and aims at developing and prototypically demonstrating selected innovative products and methods as candidates for future next-generation operational Copernicus Land Monitoring Service (CLMS) products of the pan-European and Global Components. ECoLaSS assesses the operational readiness of such candidate products and eventually suggests some of these for implementation. This shall enable the key CLMS stakeholders (i.e. mainly the Entrusted European Entities (EEE) EEA and JRC) to take informed decisions on potential procurement as (part of) the next generation of Copernicus Land services from 2020 onwards.

To achieve this goal, ECoLaSS makes full use of dense time series of High-Resolution (HR) Sentinel-2 (S-2) optical and Sentinel-1 (S-1) Synthetic Aperture Radar (SAR) data, complemented by Medium-Resolution (MR) Sentinel-3 (S-3) optical data if needed and feasible. Rapidly evolving scientific developments as well as user requirements are continuously analysed in a close stakeholder interaction process, targeting a future pan-European roll-out of new/improved CLMS products, and assessing the potential transferability to global applications.

This Deliverable **“D.44.1b – Prototype Report: Crop Area and Crop Status Parameters”** is a deliverable of the WP 44 targeting capturing the phenology of the different crop types as well as the crop type mapping based on S-1 and/or S-2 time series.

Methods which have been developed and tested in smaller test sites in several European regions and in Africa, were now applied on a larger scale on the whole demo sites for generating Crop Mask (CRM) and Crop Type Mask (CRT) prototypes. Classification in the Demonstration sites West, Central and Mali means testing the applicability of methods and the accuracy of results under highly heterogeneous conditions:

- high diversity of agro systems: highly engineered Middle European farming management vs. African small-scale and subsistence farming
- varying Earth Observation conditions: cloudiness and drought
- differing reference data: LPIS data in the case of the European sites versus comprehensive dataset collected in the field by partners for the Mali Site

Considering the heterogeneous conditions within the demonstration sites, the intention was to harmonize the methods to a certain extent by using a similar approach, similar features and a proven classifier.

Promising results are obtained in phase 1 and confirmed in phase 2. The crop mask (CRM) products obtained from applying the Random Forest classifier on different datasets give very satisfactory performance with an overall accuracy much higher than 90% in both phases: Phase 1: OA accuracy records (i) 98 % for site West (France) from S-2 only, (ii) 97% (F1-Score 0.97) for site Central from S-1 and S-2, 97%, and (iii) (F1-Score 0.89) for site Mali from S-2. Phase 2: OA for Central showed 93,3% and Kappa index of 0,83 for the combined approach of Sentiel-1 & Sentinel-2. These high accuracies obtained for all sites including the Malian smallholder cropping systems demonstrated the maturity level of this prototype. However, important aspects should be taken into consideration including the

- optimization of the number of features (selection of the most suitable time features with grouped forwarded features selection)
- stratification to deal with uneven spatial distribution of the calibration data, and

- measures to reduce some specific confusion like that of grassland and fodder crops in Germany and that of bottoms of valleys and crops in Mali.

Also, promising results are obtained for the Crop Type Map (CRT) products. The accuracy assessments show that the overall accuracy is ranging from 64 % in Mali (20 tiles - 6 classes) and 87 % (4 tiles together - 15 classes) in France both using only S-2, to 86% in Central (22 classes) to 92 % in Belgium (for the best tile – 24 classes) both combining S-2 and S-1-derived features.

Finally, it is worth mentioning that CRM and CRT prototypes have been further developed over the West, Central and Mali Demonstration sites and, then, extended to the South France and South Africa Demonstration sites where cropland – grassland discrimination and then the year-to-year variation of the CRM prototype products are of the first priority.

Table of Contents

1 INTRODUCTION 1
2 BACKGROUND AND SUMMARY OF REQUIREMENTS 3
3 DEMONSTRATION SITES 4
3.1 ECOLASS DEMONSTRATION SITES 4
3.1.1 Demonstration site West for Crop Type Mapping 5
3.1.2 Demonstration site Central for Crop Type Mapping 7
3.1.3 Demonstration site Mali for Crop Type Mapping 8
3.1.4 Demonstration site South-Africa for Crop Type Mapping 11
4 OVERVIEW OF APPLIED METHODS...11
4.1 METHOD FOR CROP TYPE MAP – DEMO-SITE WEST 11
4.2 METHOD FOR CROP TYPE MAP – DEMO-SITE CENTRAL 14
4.3 METHOD FOR CROP TYPE MAP – DEMO-SITE MALI & SOUTH-AFRICA 17
5 PROTOTYPE IMPLEMENTATION 19
5.1 DATA AND PROCESSING SETUP 19
5.1.1 Input Data 19
5.1.1.1 Demo-site West 19
5.1.1.2 Demo-site Central 21
5.1.1.3 Demo-site Mali..27
5.1.1.4 Demo-site Western Cape Province (South-Africa)...29
5.1.2 Pre-processing 30
5.1.2.1 Demo-site West 30
5.1.2.2 Demo-site Central 32
5.1.2.3 Demo-site Mali..32
5.1.2.4 Demo-site Western Cape Province (South-Africa)...33
5.1.3 Experimental Setup 33
5.1.3.1 Demo-site West 33
5.1.3.2 Demo-site Central 36
5.1.3.3 Demo-site Mali..50
5.1.3.4 Demo-site South Africa 52
5.1.4 Post-processing...52
5.2 RESULTS AND VALIDATION 58
5.2.1 Crop type map of Demo-Site West 58
5.2.2 Crop mask map of Demo-Site Central 70
5.2.3 Crop type map of Demo-Site Central 75
5.2.4 Crop type map of Demo-Site Mali 85
5.2.5 Crop type map of Demo-Site Western Cape Province (South-Africa) 92
5.3 PROTOTYPE SPECIFICATIONS 94
6 CONCLUSION AND OUTLOOK 110
REFERENCES...111

List of Figures and Tables

Figure 3-1: European Demonstration Sites 4
Figure 3-2: Overview of the six tiles of the WEST demonstration site where the southern four tiles are over Belgium and France. 6
Figure 3-3: Overview of the demonstration-site West draped over the CORINE Land Cover classes (2018). 6
Figure 3-4: Phase 2 - Overview of the demonstration-site Central draped over the CORINE Land Cover classes (20182). In red indicated the test site where methods have been tested in phase 2 prior to be applied on the demo site (see WP33). 8
Figure 3-5: Africa Sites.. 10
Figure 4-1: Phase 2 - F1 Score for all Crop Type classes: S1, S2 and combined approach S1 & S2 and the improvement of accuracies by using both sensors 15
Figure 4-2: Phase 2 - Selected time features for the CRT-prototype classification on demo site Central .. 16
Figure 5-1: Coverage of the Demo-Site Central with Sentinel-1 data (orange) and Sentinel-2 data tiles (blue)) for the reference year 2017. 22
Figure 5-2: Data Score Layer of the Demo-Site Central depicting the amount of valid coverage per pixel for the reference year 2018. *© EuroGeographics* 23
Figure 5-3: Phase 1 - Coverage of the Demo-Site Central with LPIS data (permanent grassland excluded) for the reference year 2017 used for training and validation. 25
Figure 5-4: Coverage of the Demo-Site Central with LPIS data (permanent grassland excluded) for the reference year 2018 used for training and validation.. 27
Figure 5-5: Demonstration site (in color) covering all the Sikasso region and part of the Segou and Koulikoro administrative regions. Field data were collected around 27 villages spatially distributed over the demo site (presented in red). The natural regions of the PIRT are presented in stripped line for the Haut-Bani-Niger region and in transparent for the Koutiala region. 28
Figure 5-6: JECAM protocol to sample fields along main and secondary roads (JECAM, 2017) 28
Figure 5-7 Distribution of 2017 crop types and field sizes as provided by the Western Cape Department of Agriculture. 29
Figure 5-8 : Set of 20 Sentinel-2 tiles processed for the prototype demonstration of phase 2 in South-AfricaSet of 20 Sentinel-2 tiles processed for the prototype demonstration of phase 2 in South-Africa .. 33
Figure 5-9: Phase 1 - Number of Reference Parcels and the Mean Parcel Size per class. 38
Figure 5-10: Phase 2 - stratification of the demo site Central for the drop mask classification; 4 strata taking into account different biogeographic conditions.. 41
Figure 5-11: Phase 2 - details for the demo site Central (referring to the Sentinel-2 tile – division of the classification method), serving as basis for the stratification approach or the crop mask classification. Above: Environmental Zones after Metzger et al; below: maximum height per Sentinel-2 tile 42
Figure 5-12: Phase 2 - Average parcel size per crop type. Classes 4, 7 and 13 have been left out due to small numbers of parcels as well as small parcel size. 46
Figure 5-13: Area per crop type in ha.. 47
Figure 5-14: Phase 2 - Connection between percentage of area, maximum parcel size and F1 Score (combined approach of the two sensors) per crop type for the demo site Central 48
Figure 5-15: Phase 2 - Selected features for crop type classification for demo site Central after FFS 50
Figure 5-16: Two stratification strategies: aggregated PIRT natural zones on the left and aggregated Vintrou strata on the right based on cropping systems analysis. 51
Figure 5-17 - Raw classification for the crop types, without post-processing. 54
Figure 5-18 - Post-processed classification for the crop types, with a majority filter of 5 pixels applied. . 54
Figure 5-19: Phase 2 - Raw 3crop mask for emo site Central and 3x3 pixel filtered crop mask 56
Figure 5-20: Phase 2 - Comparison of potential and actual post-processing for the crop type prototype in demo site Central 56

Figure 5-21: Phase 2 - Comparison between LPIS reference data, crop type classification re and google map indicating the heterogeneity on parcel level....... 57
Figure 5-22: Classification results on UFR-UFS-UES tiles after majority filtering....... 65
Figure 5-23: F1-score per crop type for UFR, UFS and UES tiles. Ordered by decreasing area over the 3 tiles. 66
Figure 5-24: Performance metrics using the comprehensive typology of the 2018 Flanders crop type map produced from Sentinel-1 and Sentinel-2 time series by the Sen4CAP system....... 67
Figure 5-25: Performance metrics using the ECoLaSS paneuropean typology of the 2018 Flanders crop type map produced from Sentinel-1 and Sentinel-2 time series by the Sen4CAP system....... 67
Figure 5-26: Performance metrics using the ECoLaSS paneuropean typology of the 2018 Flanders crop type map produced from Sentinel-2 time series by the Sen2-Agri system....... 68
Figure 5-27: Performance metrics using the ECoLaSS paneuropean typology of the 2018 Wallonia crop type maps respectively produced from Sentinel-2 time series by the Sen2-Agri system and from Sentinel-1 and Sentinel-2 time series by the Sen4CAP system. 68
Figure 5-28: 10-meter crop type map for Wallonia (south of Belgium – demo-site West) for 2018 growing season, obtained at the end of the season using Sentinel-1 and Sentinel-2 time series with Sen4CAP system. 69
Figure 5-29: Overview over the resulting crop mask in the demonstration site Central....... 71
Figure 5-30: Results of the crop mask classification (middle: unfiltered, right: filtered, in grey) in NW Baden-Wurttemberg (location B) over a Spot-6 image of 2016 in combination with the reference data. 72
Figure 5-31: Results of the crop mask classification (in grey) near Innsbruck (AT, location F), draped over a Worldview-1 image of 2015 in combination with the LPIS reference data. 72
Figure 5-32: Results of the crop mask classification (in grey) in W Baden-Wurttemberg (location H) over a Deimos image of 2016 in combination with the reference data. Grassland (dark blue) is correctly masked out by the crop mask, Agricultural grassland (green lines) and Fallow (light blue lines) areas are wrongly masked out....... 73
Figure 5-33: Overview over the resulting crop mask in the demonstration site Central for the reference year 2018....... 74
Figure 5-34: Overview over the results of the crop type classification in the demonstration site Central with the locations of the examples marked in black....... 78
Figure 5-35: Results of the crop type classification (left) in central Baden-Wurttemberg (location A) with high breakties (right) and the classification errors (highlighted in white)....... 79
Figure 5-36: Results of the crop type classification on a pixel base (upper left figure) and after filtering (upper right figure) in NW Baden-Wurttemberg (location B) with high breakties (central figure) and a corresponding Spot-6 image (lower figure). 80
Figure 5-37: Results of the crop type classification (upper figure) in W Baden-Wurttemberg (location C) with low breakties (lower figure) and wrongly classified polygons (highlighted in white)....... 81
Figure 5-38: Results of the crop map classification in combination with the HRL Grassland 2015 in central (location D, upper Figures) and southern Baden-Wurttemberg (location I, below)....... 82
Figure 5-39: Phase 2 - Overview over the results of the crop type classification in the demonstration site Central for the reference year 2018....... 83
Figure 5-40: Phase 2 - Plot of the confusion matrix of the pixel-based prototype results illustrating the mixing up of cereals (winter and summer) as well as of those crop types having high percentages of grassy coverage (fodder crops+temporary grassland, wine growing, fruit trees+berries)....... 84
Figure 5-41: 10-meter cropland masks over south of Mali for 2017 (left) and 2018 (right) growing seasons, obtained at the end of the season (April-November) using Sentinel-2 and Landsat- 8 data. White and black represent cropland and non-cropland classes respectively....... 86
Figure 5-42: Overlay of the cropland mask on the 2017 VHR color composite. The upper subsets corresponds to the results with the 4 strata and the lower ones to the those without stratification. for three small areas. 88
Figure 5-43: Evolution of the accuracy metrics of the different cropland mask produced along the season over south of Mali for 2017. Four periods were considered to respectively include all the S2 observations

acquired first from April to September, then from April to October, April to November and, finally from April to December 2017. 89
Figure 5-44: 10-meter crop type map over south of Mali for 2017 growing season, obtained at the end of the season using S2 and L8 data. 90
Figure 5-45: Crop map accuracy for the 2017 main crop types over the demo-site Mali (a) Koutiala strata (on the left) (b) Haut-Bani-Niger strata (on the right). Sorghum and Millet are merged into the Dry Cereals class. 90
Figure 5-46: Close-up of the 2017 crop map for the main crop types including the cropland mask the over the demo-site Mali - region of Koutiala. 91
Figure 5-47: 10-meter crop type map over south of Mali for 2018 growing season, obtained at the end of the season using S2 and L8 data. 91
Figure 5-48: Cropland mask and crop type map performances for the 2017 winter grain season in the Western Cape Province.Cropland mask performances for the 2017 winter grain season in the Western Cape Province. 92
Figure 5-49: Crop type map of the Western Cape Province (South-Africa) for the the 2017 winter grain season as derived from Sentinel-2 and Landsat-8 time series. The initial 9 classes are mapped here while the validation results are presented after regrouping in 7 crop types. 93

Table 3-1: Description of the selected Demonstration Sites 5
Table 3-2: Crop calendar for the main crops of the site in Mali (1-2-3 corresponding to decade). 10
Table 5-1: Number of S-1 and S-2 images per tile used for the calculation of the agricultural prototype for the demonstration site Central. 21
Table 5-2: Phase 1 - Number and name of crop groups and their abbreviation for the reference year 2017. 24
Table 5-3: Phase 2 - reworked crop type nomenclature for crop type classification for the reference year 2018 comprising the main crop groups, crop classes and crop types in a Pan-European context which could be adapted to the regional context 26
Table 5-4: Description of the 2017 in situ datasets in terms of number of samples per class as provided by IER (Mali). 28
Table 5-5: Example of visual check on the samples. 31
Table 5-6: Temporal features associated with the NDVI extracted from the tile 31UEQ. The delineation represents LPIS dataset parcels 34
Table 5-7: Phase 1 - Start and end date of the 10 different scene sets and the number of S-1/S-2 images they contain including the maximum amount of cloud cover (%). 37
Table 5-8: Phase 1 - Number of polygons for each LC type in the reference data for the crop mask: All, Train set (for model training), and Test set (for validation). 37
Table 5-9: Phase -1 Number of polygons for each crop group in the reference data (All), their area, and the number of polygons used for validation (Test set). 38
Table 5-10: Phase 1 - Overview over the relevant bands/indices and the derived time features used for training and classification. 39
Table 5-11: Phase 2 - Start and end date of the 8 different scene sets and the number of S-1/S-2 images they contain including the maximum amount of cloud cover (%). 40
Table 5-12: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum Upper Left: All, Train set (for model training), and Test set (for validation). 43
Table 5-13: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum North East: All, Train set (for model training), and Test set (for validation). 43
Table 5-14: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum Central East: All, Train set (for model training), and Test set (for validation). 43
Table 5-15: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum South West: All, Train set (for model training), and Test set (for validation). 43
Table 5-16: Phase 2 - Number of parcels/sample polygons per crop type for the demo site Central 44

Table 5-17: Phase 2 - pixel support for crop type classification for demo site Central; classes 4, 7 and 13 left out due to insufficient sample number 45
Table 5-18: Phase 2 - Time features per band and index for Sentinel-1 49
Table 5-19: Phase 2 - Time features per band and index for Sentinel-2 49
Table 5-20: Cropland labels, characterizing 12 types of culture. 52
Table 5-21: Non-cropland class labels. 53
Table 5-22 - Before and after the cleaning of isolated pixels. 53
Table 5-23: Before and after majority filtering. 55
Table 5-24: Global confusion matrix for the four S-2 tiles of the French West part of the demonstration site, for phase 1. 59
Table 5-25: Confusion matrix for the derived crop mask, over the 4 tiles. 60
Table 5-26: Overall Accuracy (OA) on both products over the 4 tiles. 60
Table 5-27 - Confusion matrix for the French West tiles (for phase 2) 61
Table 5-28: Confusion matrix of the CRM prototype in the demonstration siteWest. 63
Table 5-29 - F-Scores and maximal confusion for each class over the French West tiles. 63
Table 5-30: Overall Accuracy (OA) of classification results based on specific model per tile. OA is presented before and after grouping crop classes and applying majority filter on for the 31UFR, 31UFS and 31UES tiles. 64
Table 5-31: Overall Accuracy (OA) of classification results based on a single model trained on 31UFR tile for the 31UFR, 31UFS and 31UES tiles. 65
Table 5-32: Comparison of the overall accuracies for the crop type maps produced using Sentinel-2 time series (Sen2-Agri) and Sentinel-2 and Sentinel-1 time series (Sen4CAP) for two Belgian administrative regions and two different typologies (comprehensive one and the ECoLaSS paneuropean one). 66
Table 5-33Confusion matrix for the 2018 Wallonia crop type produced from Sentinel-1 and Sentinel-2 using Sen4CAP with the ECoLaSS paneuropean crop typology. 69
Table 5-34: Selected time features used for training and classification. 70
Table 5-35: Confusion matrix of the CRM prototype in the demonstration site Central. 71
Table 5-36: Confusion matrix of the CRM prototype in the demonstration site Central for the reference year 2018. 75
Table 5-37: Accuracy results per class. 76
Table 5-38: Confusion matrix of the CRT prototype in the demonstration site Central. 77
Table 5-39: Phase 2- Confusion matrix for crop type mask prototype - pixel based validation 84
Table 5-40: Phase 2- Confusion matrix for crop type mask prototype - validation on field level 85
Table 5-41: Variability of the Mali cropland map accuracies as assessed by a fully independent dataset for the seven different training set (Leave-4-Out protocol from 27 sampled villages). 86
Table 5-42: Variability of the Mali cropland map performances as assessed by spatially independent dataset for the seven different training sets (Leave-4-Out protocol from 27 sampled villages). 87
Table 5-43: Fscore, Kappa and OA achieved by the end of the season cropland mask generated using S2 and L8 time series over south of Mali 88
Table 5-44: Confusion matrix for the 2017 crop type map derived from Sentinel-2 time series for the winter grain season. 93
Table 5-45: Product specifications for CRT_2016_010m_WE_03035_prototype_v01.tif and for CRM_2016_010m_WE_03035_prototype_v01.tif 96
Table 5-46: Color palette for CRT_2016_010m_WE_03035_prototype_v01.tif 97
Table 5-47: Color palette for CRM_2016_010m_WE_03035_prototype_v01.tif 97
Table 5-48: Product specifications for CRT_2017_010m_WE_03035_prototype_v01.tif (Belgium part) .. 98
Table 5-49: Color palette for CRT_2017_010m_WE_03035_prototype_v01.tif (Belgium part) 99
Table 5-50: Product specifications for CRT_2017_010m_WE_03035_prototype_v01.tif (France part) .. 100
Table 5-51: Color palette for CRT_2017_010m_WE_03035_prototype_v01.tif (France part) 101
Table 5-52: Color palette for CRM_2017_010m_WE_03035_prototype_v01.tif (France part) 102
Table 5-53: Product specifications for CRT_2017_010m_ML_32630_prototype_v01.tif 102
Table 5-54: Color palette for CRT_2017_010m_ML_32630_prototype_v01.tif 103

Table 5-55: Product Specifications and color palette for CRT_2017_010m_CE_03035_prototype_v01.tif 103
Table 5-56: Product Specifications and color palette for CRT_2018_010m_CE_03035_prototype_v01.tif 105
Table 5-57: Product Specifications and color palette for CRM_2017_010m_CE_03035_prototype_v01.tif 107
Table 5-58: Product Specifications and color palette for CRM_2017_010m_CE_03035_prototype_v01.tif 108
Table 5-59: Colour palette for the probability layers 109

Abbreviations

AOI	Area of Interest
BRI	Brightness Index
CLMS	Copernicus Land Monitoring Service
CMDT	Compagnie Malienne pour le Développement du Textile
Co	Cotton
CRM	Crop Mask
CRT	Cropland Type/Crop Type
CT	Classification Trees
ECoLaSS	Evolution of Copernicus Land Services based on Sentinel data
EEA	European Environmental Agency
EEE	European Entrusted Entities
EO	Earth Observation
ESA	European Space Agency
FFS	Forward Feature Selection
GIS	Geographic Information System
GRA	Grassland
GRD	Ground Range Detected
H2020	Horizon 2020
HRLs	High Resolution Layers
IRECI	Inverted Red-Edge Chlorophyll Index
ITCZ	Intertropical Convergence Zone
JRC	Joint Research Center
LC/LU	Land Cover/Land Use
LPIS	Land Parcel Identification System
Mi	Millet
MMU	Minimum Mapping Unit
Mz	Maize
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
OA	Overall Accuracy
PA	Producer's Accuracy
Rc	Paddy rice
PHENO	Phenological Product
PIRT	Projet d'Inventaire des Ressources Terrestres
RF	Random Forest
S-1	Sentinel-1
S-2	Sentinel-2
S-3	Sentinel-3
SAR	Synthetic Aperture Radar
Sg	Sorghum
UA	User's Accuracy
VHR	Very High Resolution
WBS	Work Breakdown Structure
WP	Work Package
XML	Extensible Markup Language

1 Introduction

The Horizon 2020 (H2020) project, "Evolution of Copernicus Land Services based on Sentinel data" (ECoLaSS) addresses the H2020 Work Programme 5 iii. Leadership in Enabling and Industrial technologies - Space, specifically the Topic EO-3-2016: Evolution of Copernicus services. ECoLaSS is being conducted from 2017–2019 and aims at developing and prototypically demonstrating selected innovative products and methods as candidates for future next-generation operational Copernicus Land Monitoring Service (CLMS) products of the pan-European and Global Components. ECoLaSS assesses the operational readiness of such candidate products and eventually suggests some of these for implementation. This shall enable the key CLMS stakeholders (i.e. mainly the Entrusted European Entities (EEE) EEA and JRC) to take informed decisions on potential procurement as (part of) the next generation of Copernicus Land services from 2020 onwards.

To achieve this goal, ECoLaSS makes full use of dense time series of High-Resolution (HR) Sentinel-2 optical and Sentinel-1 Synthetic Aperture Radar (SAR) data, complemented by Medium-Resolution (MR) Sentinel-3 optical data if needed and feasible. Rapidly evolving scientific developments as well as user requirements are continuously analysed in a close stakeholder interaction process, targeting a future pan-European roll-out of new/improved CLMS products, and assessing the potential transferability to global applications.

This Deliverable **D14.2-D.44.1b – Prototype Report: Crop Area and Crop Status/Parameters (Issue 2)** is devoted for demonstrating a proof-of-concept system for delivering agriculture related products on pan-European scale by taking advantage of the continuous fluxes of Sentinel-1, and -2 data. This deliverable comprises a description of the provided prototype datasets of crop area and crop status/parameters (linked to D44.2a). A detailed description of the objectives is provided together with an explanation of the methodology, results and conclusions, as derived by WP44. It addresses the developed methodologies for extending the S-2 processing chain to S-1 and to a larger number of crop types, the implementation of a classification processing chain considering multi-annual spectral signatures, the integration of Synthetic Aperture Radar (SAR) time series into optical processing lines for mapping agricultural practices, as well as explore the algorithm potential in food insecure countries. As such, it is part of WP 44 of task 4:"Thematic proof-of-Concept/Prototype on Continental/Global Scale", which aims at exploring and setting up a robust classification approach for Crop type identification based on Sentinel-2 and Sentinel-1 time series and in situ data for pan-European Land Monitoring. This report is accompanied by the Deliverable **D14.4 - P44.2b – Data Sets of Crop Area and Crop Status/Parameters Products (Issue 2)**. This report serves as documentation for the prototype dataset.

In the ECoLaSS project a prototype is defined as a prototypic / thematic proof-of-concept implementation of an improved or newly defined potential future Copernicus Land layer, building on the methods and processing lines developed in Task 3. The consortium has selected representative demonstration sites both in Europe and Africa, covering various bio-geographic regions and biomes. All prototype products and services are prototypically implemented in these sites in the frame of the Task 4 WPs. In ECoLaSS proofs-of-concept / prototype demonstration are carried out with respect to five topics of relevance: (i) Time series derived indicators and variables, (ii) Incremental Updates of HR Layers, (iii) Improved permanent grassland identification, (iv) Crop area and crop status / parameters monitoring, and (v) New LC/LU products. This deliverable focusses on the prototype Crop Type Identification as part of WP 44.

This report comprises of six sections. Section 1 of the document structure provides a comprehensive introduction on the objectives and contents of the report. Section 2 gives background information on crop type identification, prototyping phase and how they will serve the stakeholders in the framework of Copernicus HRLs and ongoing activities. Section 3 details the demonstration sites characteristics where development and validation of the prototypes take place. Then, an overview of the developed methods for crop type mapping is provided in section 4 while section 5 describes the implementation of the developed methods in a way detailing the processing chain producing crop type maps with accuracy assessment of the maps. Finally, core points of the developed methods and main results are summarized with perspectives in section 6.

2 Background and Summary of Requirements

After first methods have been tested by the Task 3 WPs (AD05, AD06, AD07, AD08, AD09) in various test sites and algorithms have been described, the demonstration activities of Task 4 have set up the developed processing lines in demonstration sites and derive final prototype versions. This comprises establishing prototypes for: (i) deriving indicators and variables both for Continental and Global Component products and services from high-volume time series data with high spatial resolution and temporal repeat frequency; (ii) improving one of the main pan-European Copernicus Land products, i.e. the current (2012, 2015) and future (2018) HRLs on Forest and Imperviousness by developing incremental update strategies and ensuring time series consistency; (iii) improved permanent grassland identification targeting the HRL Grassland 2015 improvement; (iv) crop area and crop status/parameters monitoring targeting a potential future Agricultural service; as well as (v) further novel LC/LU products, e.g. as tested in Task 3.

The project is basing all its developments on regularly updated high-priority user requirements, and assess/benchmark all operational product candidates in view of their innovation potential and technical excellence, automation level, potential for roll-out to pan-European level and/or global scale, timeliness for operational implementation, costs versus benefits, etc. (further elaborations will be performed in Task 5).

The latest production of the Copernicus High Resolution Layers contains five thematic areas: Imperviousness, Forest, Grassland, Water and Wetness, and Small Woody Features. A layer on agriculture or arable land has not yet been part of the HRLs, and is also not foreseen for the 2018 production, although it is definitely in the trending topics in the political discourse, as well as from the user perspective.
ECoLaSS targets a potential new layer on agriculture for 2020+, which is proposed to contain, on the one hand a Crop Mask (CRM), and on the other hand a Crop Type (CRT) map.

Within the user and stakeholder requirements analysis, as performed as part of WP 21, a future potential Agricultural Service was most voiced by multiple users. The Joint Research Centre (JRC) emphasised the concept note entitled “Towards Future Copernicus Service Components in support to Agriculture” which has been drafted by JRC in April 2016 (JRC, 2016). A follow-up on this concept note is expected to be issued any time soon. The ECoLaSS team is proposing a Crop Mask and different setups of Crop Type maps, which have been produced for four different prototype experiments applied onto three demonstration sites in project phase one.

3 Demonstration Sites

All prototypes are implemented in selected representative demonstration sites, which cover various bio-geographic regions and biomes. In phase 1, the crop mask and crop types 2016 are demonstrated in the demonstration site West, whereas for 2017 the demonstration sites Central, West and Mali are used. In phase 2, these masks were updated: for Central, crop mask and crop type mask based on data of the reference year 2018.

3.1 ECoLaSS Demonstration Sites

The prototype sites (60,000/90,000 km² per prototype site) contain the 5 test sites from Task 3. These prototype sites are relevant in Task 4 for demonstrating the proposed candidates for a Copernicus Land Service Evolution roll-out on a larger scale. As shown in Figure 3-1, the pre-selected demonstration sites cover the Boreal, Continental, Alpine, Atlantic and Mediterranean zones (Source EEA: http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3#tab-gis-data) of the member and associated states of the EEA-39. The selected prototype sites are located in the **North of Europe, in the Alpine/Central region, in the West, in the South-West and in the South-East of Europe**. All prototype products and services will be prototypically implemented in one or more prototype sites in project phase 1, and in three prototype sites in phase 2.

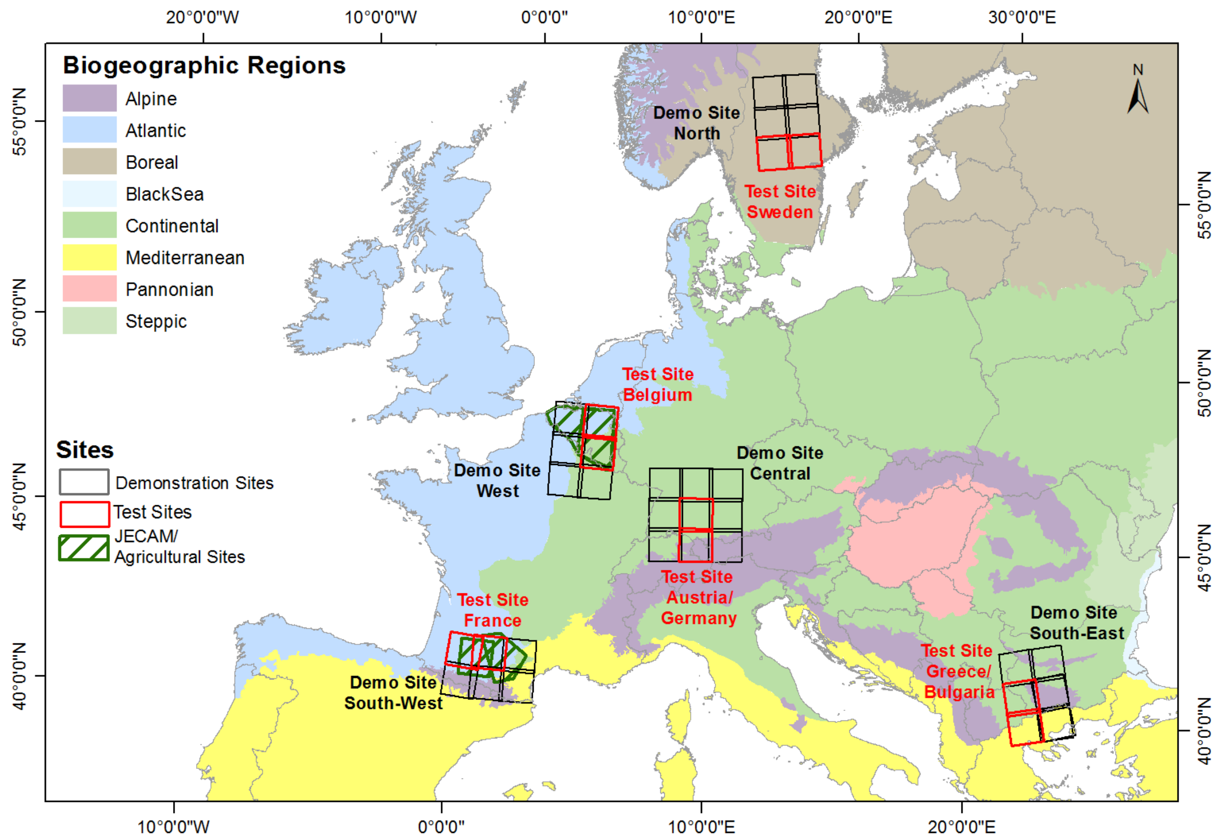


Figure 3-1: European Demonstration Sites

(Map: © European Environment Agency; administrative boundaries: ©EuroGeographics)

A short description of the different prototype sites is given in the following Table 3-1 below:

Table 3-1: Description of the selected Demonstration Sites

Location	Biogeographical region(s)	Countries	Distribution of CORINE land cover classes 2018 (Level 1) per demonstration site
Northern Europe	Boreal	Sweden, Norway	Artificial areas: 1.90%, Agricultural areas:11.87%, Forest and semi-natural areas: 69.01%, Wetlands: 3.25%, Waterbodies: 13.94%
Alpine / Central Europe	Continental, Alpine	Germany, Austria, Switzerland, Italy and Czech Republic	Artificial areas: 9.03%, Agricultural areas: 44.55%, Forest and semi-natural areas: 44.65%, Wetlands: 0.23%, Waterbodies: 1.55%
West Europe	Atlantic, Continental	Belgium, France, Luxembourg	Artificial areas: 13.47%, Agricultural areas: 63.08%, Forest and semi-natural areas: 21.43%, Wetlands: 0.39%, Waterbodies: 1.61%
South-East Europe	Mediterranean, Continental, Alpine	Serbia, Macedonia, Greece, Bulgaria and Kosovo	Artificial areas: 3.34%, Agricultural areas: 34.87%, Forest and semi-natural areas: 56.67%, Wetlands: 0.17%, Waterbodies: 4.93%
South-West Europe	Atlantic, Mediterranean, Alpine	France, Spain	Artificial areas: 3.59%, Agricultural areas: 48.05%, Forest and semi-natural areas: 47.06%, Wetlands: 0.10%, Waterbodies: 1.18%

3.1.1 Demonstration site West for Crop Type Mapping

In the frame of the Task 4 where all prototype products and services are implemented, the West demonstration site has been chosen as the primary demonstration site for the improved HRL Grassland (GRA), a potentially new phenological product (PHENO) as well as a potential future Copernicus Land High Resolution Layer on Agricultural (AGRI). The demonstration site contains the test site “Belgium” that has been studied in Task 3.

For this WP44, the site West has been processed separately because of the various constraints in in situ data availability corresponding to the different countries, i.e. Belgium and France. Regarding the French part, the area selected exhibits a strong heterogeneity of cropland, mixed with intensive grasslands, besides common areas such as urban areas, forest or water.

For SIRS demo site, all the prototype experimentation was conducted on four Sentinel-2 tiles over France and Belgium (31UER, 31UEQ, 31UFR and 31UFR), which represents a total superficies of 40.000 km² in the south part of the demo-site West, as seen in Figure 3-2.
In the rest of this report, this will be referred to as the “French West tiles” for clarity.
For the phase 1 of the UCL site, three Sentinel-2 tiles were used for conducting the prototype experimentation. In particular, the 31UES and 31UFS tiles are over Belgium while the 31UFR tile is over Belgium and France (Figure 3-2) with a total superficies of 30.000 km² in the south part of the demo-site West. For demonstration in phase 2, the entire administrative regions including these Sentinel-2 tiles have been processed. For the sake of demonstration relevance, all the tiles necessary to map completely the Flanders on one hand and Wallonia on the other hand have been considered.
In the same fashion, this will be referred as the “Belgian West tiles”.

Figure 3-2: Overview of the six tiles of the WEST demonstration site where the southern four tiles are over Belgium and France.

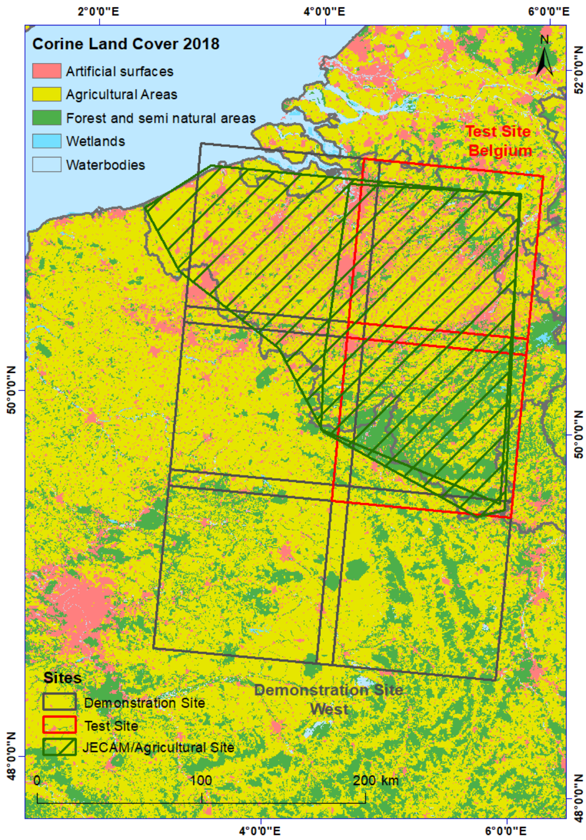


Figure 3-3: Overview of the demonstration-site West draped over the CORINE Land Cover classes (2018).

3.1.2 Demonstration site Central for Crop Type Mapping

The Demonstration Site Central covers the border region of Germany (mainly the provinces of Bayern and Baden-Wurttemberg), Austria (mainly the provinces of Tyrol and Vorarlberg), Switzerland as well as small areas of France, Italy, and Liechtenstein. The approaches for crop mapping have already been tested in the Southern part of the demo site, on a test site covering the tiles 32UNU and 32TNT. The Central demonstration site is characterized by cropland areas, mixed with permanent grassland (pastures). The adjoining part towards South, covering the foothills of the Bavarian Alps, is dominated (besides forest cover) by grasslands including specific grassland habitats and wetland types. Zones of alpine vegetation are included in the southern tiles, which contain also the Wetterstein mountain range as part of the Alps, stretching South down to the Inn valley. Lake Constance is also part of the Southern region as well as lower parts of Switzerland. Besides the Alps, the Black Forest area and the Vosges are regions at higher altitude, where forests and grassland predominate cropland.

The demo site has been chosen for crop mapping in order to test the classification approach under challenging but also realistic conditions:

- high diversity of crop categories
- heterogeneous regions (e.g. in Baden-Wurttemberg and Austria)
- strong topography (lowlands and high mountains)
- availability of Land Parcel Identification System (LPIS) in-situ data for training the crop type classification algorithm and for validations.

Figure 3-4 shows the Central Demonstration site.

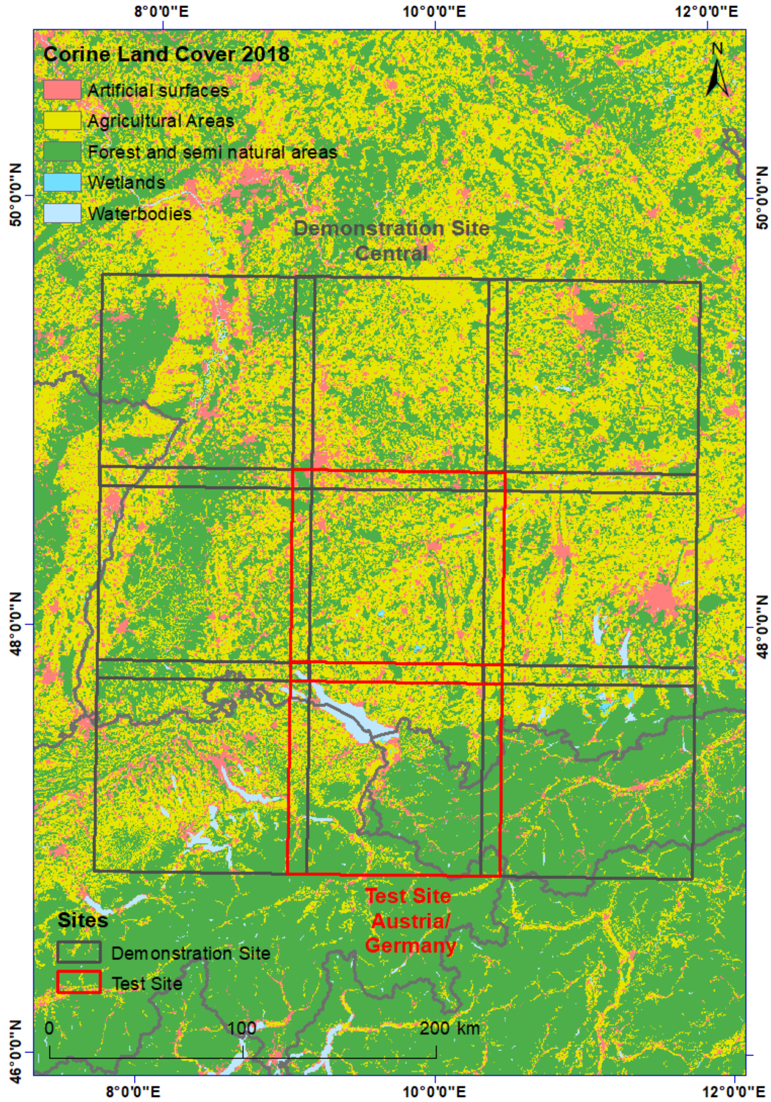


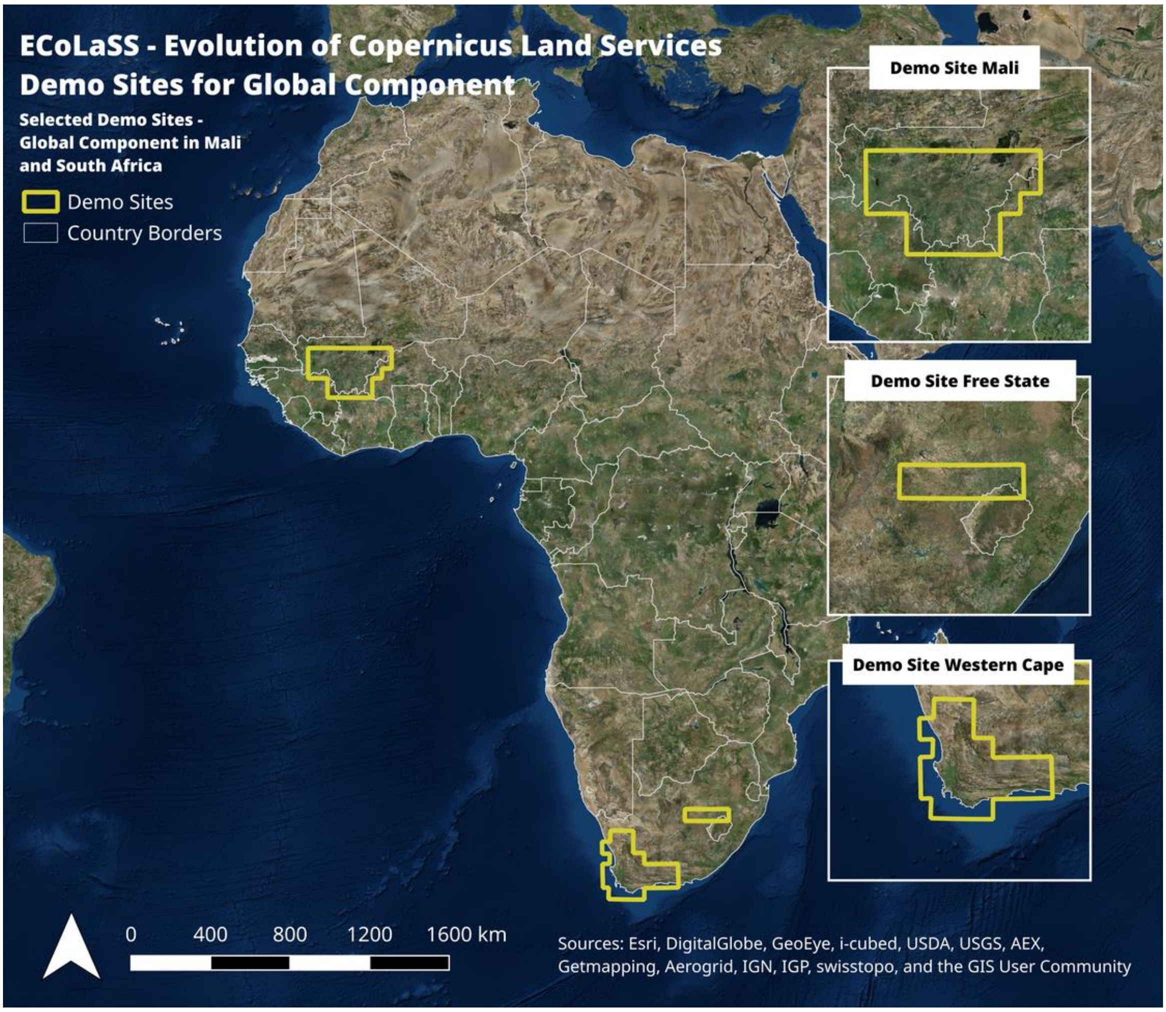
Figure 3-4: Phase 2 - Overview of the demonstration-site Central draped over the CORINE Land Cover classes (20182). In red indicated the test site where methods have been tested in phase 2 prior to be applied on the demo site (see WP33).

© European Union, Copernicus Land Monitoring Service 2012, European Environment Agency (EEA).

3.1.3 Demonstration site Mali for Crop Type Mapping

The Mali demo-site corresponds to a small holder farming system typically monitored by Early Warning Systems in the context of food unsecure countries. The outstanding challenges faced by all Early Warning Systems in assessing local food insecurity, include data accuracy, timeliness and disaggregation. Those challenges are even wider in heterogeneous smallholder farming systems, such as Southern Mali, due to small field size, heterogeneity in management practices, the resulting landscape fragmentation, and the widespread presence of trees within the fields. This Malian demo-site is representative more specifically of the Sudanian region covering a large part of the food unsecure region of West Africa. This region has a strong north-south gradient linked to the movements of the Intertropical Convergence Zone (ITCZ). The climate is tropical with two distinct dry and rainy seasons. In the Sudanian region (isohyets from 500 to 1,000 mm) the growing period is longer (more than 150 days) and spans from May to November.

The study site has been extended to largely correspond to the cotton's belt and covers about 135,500 km². From an administrative perspective, this includes the entire Sikasso's region and part of the Koulikoro (Koulikoro, Kati, Kangaba and Dioila circles) and Ségou's (Bla, San, Tominiam, Baroueli circles) regions with a North limit around 14° in latitude following the circles limits. It is covering most of the Compagnie Malienne pour le Développement du Textile (CMDT) region. The study areas includes two main agro-climatic zones: the Sudanese and the Sudano-Guinean zones that contrast mainly in the annual amount of rainfall. Two major natural regions from the Projet d'Inventaire des Ressources Terrestres (PIRT) are observed in the study regions: Haut Bani-Niger and Koutiala. Pearl millet (Pennisetum glaucum (L.) R. Br.), sorghum (Sorghum bicolor (L.) Moench), maize (Zea mays L. ssp.), paddy rice and cotton (Gossypium sp.) are the major crops in the study area covering respectively 29%, 20%, 13%, 11% and 9% of the cultivated land. A map of the selected prototype site of Mali for the Crop type map is provided in Figure 3-5:



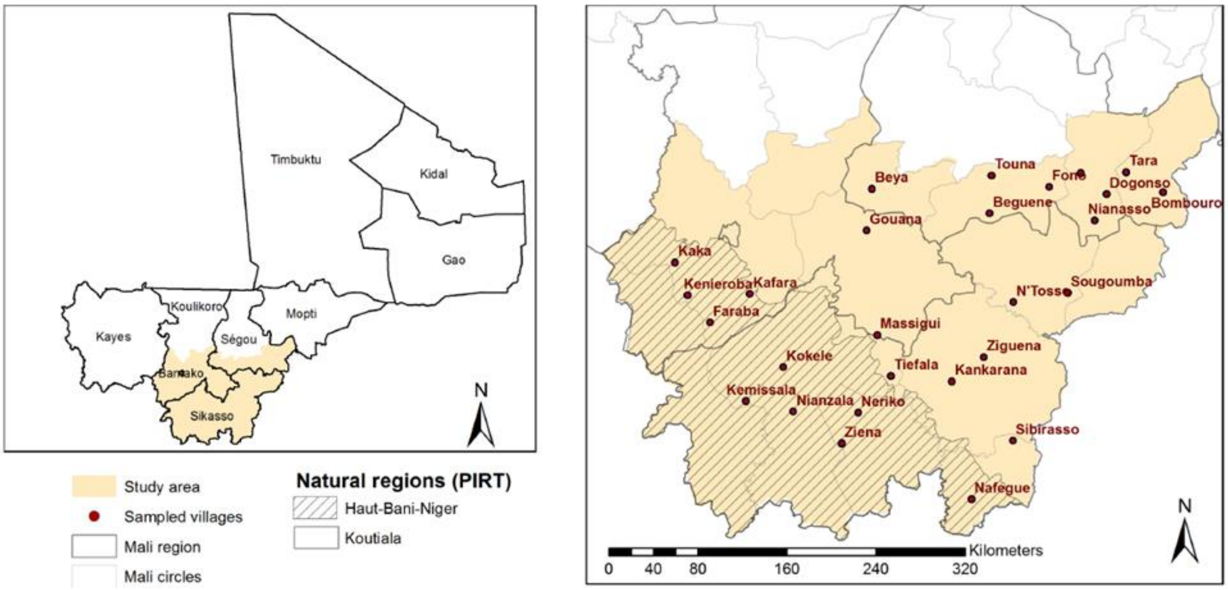


Figure 3-5: Africa Sites

The main crops are the millet (29%), sorghum (20%), maize (13%), paddy rice (11%), and cotton (9%) with a field size typically ranging from 1 to 5 ha. Table 3-2 reports the crop calendar (1-2-3 corresponding to decades) of the main crops.

Table 3-2: Crop calendar for the main crops of the site in Mali (1-2-3 corresponding to decade).

	May			June			July			August			September			October			November		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Co			Sowing			Growing							Harvest								
Mz			Sowing				Growing						Harvest								
Mi			Sowing						Growing										Harvest		
Sg			Sowing						Growing						Harvest						
Rc					Sowing					Growing							Harvest				
Gr					Sowing				Growing				Harvest								

3.1.4 Demonstration site South-Africa for Crop Type Mapping

As shown at the Figure 3, the Western Cape Province was proposed as the location of the demo-site South-Africa for crop type mapping. The Western Cape Department of Agriculture was indeed very interested by the demonstration of Sentinel-2 derived agriculture products as planned in the ECoLaSS project. The total area of the Western Cape province is about 129,500 km² but one quarter does not correspond to any arable lands."

4 Overview of applied methods

Since the duration and characteristics of the growing period strongly depends on location and regional biogeographic conditions such as temperature, precipitation, altitude, etc., it is necessary to work with time series analysis to discern the phenological dynamics associated. Moreover, since grassland and cropland can show similitudes over time, it seems important to work on time series indicators on which the phenological distinction should be optimized. Thus, this chapter will present the spectral indicators and the time features associated used to generate this prototype, the pre-classification steps, the classification algorithm employed in the processing chains, and finally, the validation analysis procedure.

4.1 Method for crop type map – Demo-site West

The chapter 4.1 contains information about feature computation and selection, the pre-classification steps, the classification algorithm and the validation procedure.

FEATURES COMPUTATION AND SELECTION

Many spectral indices have already been defined in the WP31 [AD05]. The ones used for the cropland type prototype generation will be specified in this paragraph.

NDVI – Normalized Difference Vegetation Index:

The "Normalized Difference Vegetation Index" (Rouse Jr. et al., 1974; Tucker, 1979) is used as an indicator to monitor vegetation health, and can be used as a proxy for photosynthetic activity and primary production from vegetation biomass. It is calculated as the difference in the reflectance between the NIR region and the red region – a rapid change in the spectral response of vegetation known as the "red edge" – then normalized by the sum of the reflectance in both channels:

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{Red}}}$$

NDWI – Normalized Difference Water Index:

The “Normalized Difference Water Index” (Gao, 1996) is defined as the ratio

$$\mathrm{NDWI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{SWIR}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{SWIR}}}$$

where ρ is the reflectance of each spectral channel. Both wavelengths are localized in the part of the spectrum reflected by vegetation canopies. The NIR channel is linked to a negligible absorption of light by the water content present in the vegetal, while the SWIR channel present a weak liquid absorption. The NDWI is therefore sensitive to slight changes in the liquid water absorbed by vegetation canopies, giving an indication on the vegetation water stress.

BRI – Brightness Index:

The brightness Index (BI) can expressed as (Mathieu et al., 1998):

$$\mathrm{BI} = \sqrt{\frac{\rho_{\mathrm{Red}} + \rho_{\mathrm{Green}} + \rho_{\mathrm{Blue}}}{3}}$$

which is a measure of the average reflectance magnitude in the visible bands, used to quantify the soil color effect.

IRECI - Inverted Red Edge Chlorophyll Index:

The Inverted Red Edge Chlorophyll Index (IRECI) was introduced by (Frampton, Dash, Watmough, & Milton, 2013) in order to estimate the canopy chlorophyll content, and can be written as:

$$\frac{\rho_{Red\ Edge\ 3} - \rho_{Red}}{\rho_{Red\ Edge\ 1} / \rho_{Red\ Edge\ 2}}$$

The strong correlation between the IRECI and the biophysical Leaf Area Index (LAI) parameter makes it an essential optical index to compute.

Temporal features:

Furthermore, temporal statistics on those derived indices have been applied, based on the seasonal time intervals. They are especially useful for determining vegetated classes (Esch et al., 2018). On the prototype generation only nine of them have been generated, for each index, on each considered period, at pixel-level:

- maximum value,
- mean value,
- minimal value,
- standard deviation,
- 10^{th}, 25^{th}, 50^{th} (i.e. the median), 75^{th} and 90^{th} percentiles.

For the French West tiles, this has been computed for both phases on the optical time series from 2016 and 2017.

For the Belgian West tiles, the coefficient of variation has been computed in addition to the previously mentioned temporal features on Sentinel 1 data.

PRE-CLASSIFICATION STEPS

The Pre-classification steps sub-chapter addresses the image generation and furthermore LPIS data as reference data.

Image generation

For the French West tiles, in the first phase, a composite of 81 bands was produced by SIRS. The cropland types prototype generation was established with three different time-windows: January to March, April to June and July to September, in order to find a balance between the amount of valid data and the temporal constraints to create a meaningful time series for the phenology, thanks to spectral features as the NDVI, NDWI and brightness, which highlight different characteristics of the vegetation.

For the second phase, "artificially" images are produced, as described in the section 3.1.2, in particular the Weighted Average Compositing (WAC). This method is used here over the 2017 time series of S-2A and B: cloud masks are applied on the images, and an interpolation at fixed dates is computed in order to fill gaps due mainly to cloud presence, but also to saturated pixels or anomalies on the sensor. Indices and temporal features are then extracted from those images and fed to the model.

For the Belgian West tiles, two time series of bands were used. The fist was composed of 374 bands representing Sentinel-2 data (b3:b8 and b11:b12) in addition to the calculated spectral features (NDVI, NDWI and BRI). Sentinel-2 data were cloud masked and, then, a gap-filling algorithm based on linear interpolation (every 10 days) was applied from January to December. The second time series was composed of 288 bands of Sentinel-1 data representing the backscattering (VV, VH, VV/VH) based calculated features such as mean, standard deviation, coefficient of variations, 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles. Each temporal feature has been computed on a 2 months period from January to December.

LPIS as reference data

Two LPIS were used to generate the cropland type product (one on France and another on Belgium). These datasets will be used as a reference for the year 2016 for the first phase over the French West tiles. The polygons differentiate several agricultural areas such as croplands and grasslands.

These reference data will be used for three different steps of the processing chain:

1. Samples selection on the most represented cropland types (as the classifications will be distributed on two countries, the classes need to be in common)
2. Homogenization of the results at the parcel levels (attribution of the majority number of pixels to the fields, by statistics zonal computation)
3. Qualitative comparison (i.e. look and feel) for the assessment of the cropland type product

For the second phase over the French West tiles, only the French LPIS from 2017 has been used.

The Belgian LPIS for the year 2017 was used by UCL for the training and validation of the classification.

CLASSIFICATION ALGORITHM

As explained in previous WPs of Task 3, Random Forest (RF) classification combines many decision trees to obtain better predictive performance. Each decision tree is calibrated on a selection of random subset. Algorithms such as RF have recently received increasing interest (Rodriguez-Galiano, Ghimire, Rogan, Chica-Olmo, & Rigol-Sanchez, 2012) because they have proved to be more accurate and more robust to noise than single classifiers (Shang & Breiman, 1996). Ensemble classifier, like the multitudes of decision trees in RF, are known to perform better than an individual classifier can. Breiman (Breiman, 2001) introduced RF in 2001 which presents many advantages for its application in remote sensing:

- Efficiency on large data bases;
- Thousands of input variables without variable deletion;
- Estimation of which variables are important in the classification;
- Relative robustness to outliers and noise;
- Computational lightness compared to other tree ensemble methods (e.g. Boosting);
- Much less sensitivity to overtraining or over fitting.

A RF consists of a combination of classifiers where each one of those contributes with a single vote to the assignation of the most frequent class detected for the input vector. This grant RF special characteristics which make it substantially different to traditional classification trees (CT). In fact, a RF increases the diversity of the trees by making them grow from different training data subsets created through the process.

VALIDATION PROCEDURE

Thematic accuracy is presented in the form of an error matrix made from the results of the samples interpretation, with 50% of them used for the calibration and the other 50% for the validation.

The thematic accuracy is also defined by several quality indices:
- The Overall Accuracy (also called Recognition Rate) is measured by the sum of the diagonal of the Confusion Matrix divided by the total number of controlled points. It assesses the overall agreement between the classified and reference data set.
- The User Accuracy is measured by dividing the diagonal number of pixels by the row total. It assesses the commission error (or contamination risk), i.e. the errors due to the wrong allocation of an observation to a class.

Producer Accuracy is measured by dividing the diagonal number of pixels by the column total. It assesses the omission error.

The F-score was utilized as a measure of accuracy where the precision and recall were considered to calculate the score. Specifically, the precision is the number of correct positive results divided by the number of all positive results returned by the classifier while the recall is the number of correct positive results divided by the number of all samples identified as positive.

For the phase 2 of the Belgian part of the demo site West, the open source Sen2Agri v. 2.0. is a pixel-based approach dealing only with Sentinel-2 and Landsat-8 time series. At the opposite, the Sen4CAP system relies on an object-based approach allowing to filter out the results per parcel. Both systems calibrate a RF model per stratum but no stratification is used in Belgium. However two separate models are established separately to deliver map outputs corresponding to the full extent of both administrative regions, i.e. Flanders and Wallonia. The details of the methods are fully described at the section 4.3, in Defourny et al. (2019) and in the Sen4CAP user manual (esa-sen4cap.org).

4.2 Method for crop type map – Demo-Site Central

During the testing and benchmarking performed in the frame of WP33 (AD06) one of the main results was that for the crop mask as well as the crop types, the accuracies of the classifications based on S-2 are significantly higher than those based on S-1. The benefit of using the combined approach of both, S-1 and S-2, varied for the products: whereas the accuracies of the crop mask increased only marginally in the central test site, those for the crop type mask has been strongly improved, varying for several crop type classes between F1 Score values from +1 to +10% (see figure below). It was decided to use both Sentinel-1 and Sentinel-2 data for the agricultural prototype, in the light of implementation and comparison on the larger demonstration site, a future pan-European roll out into areas of high cloud cover or short growing season, and further testing of the advantages of optical/SAR sensor integration.

After the data selection, several pre-processing steps were necessary to prepare the EO data (see section 5.1.2 and AD05).

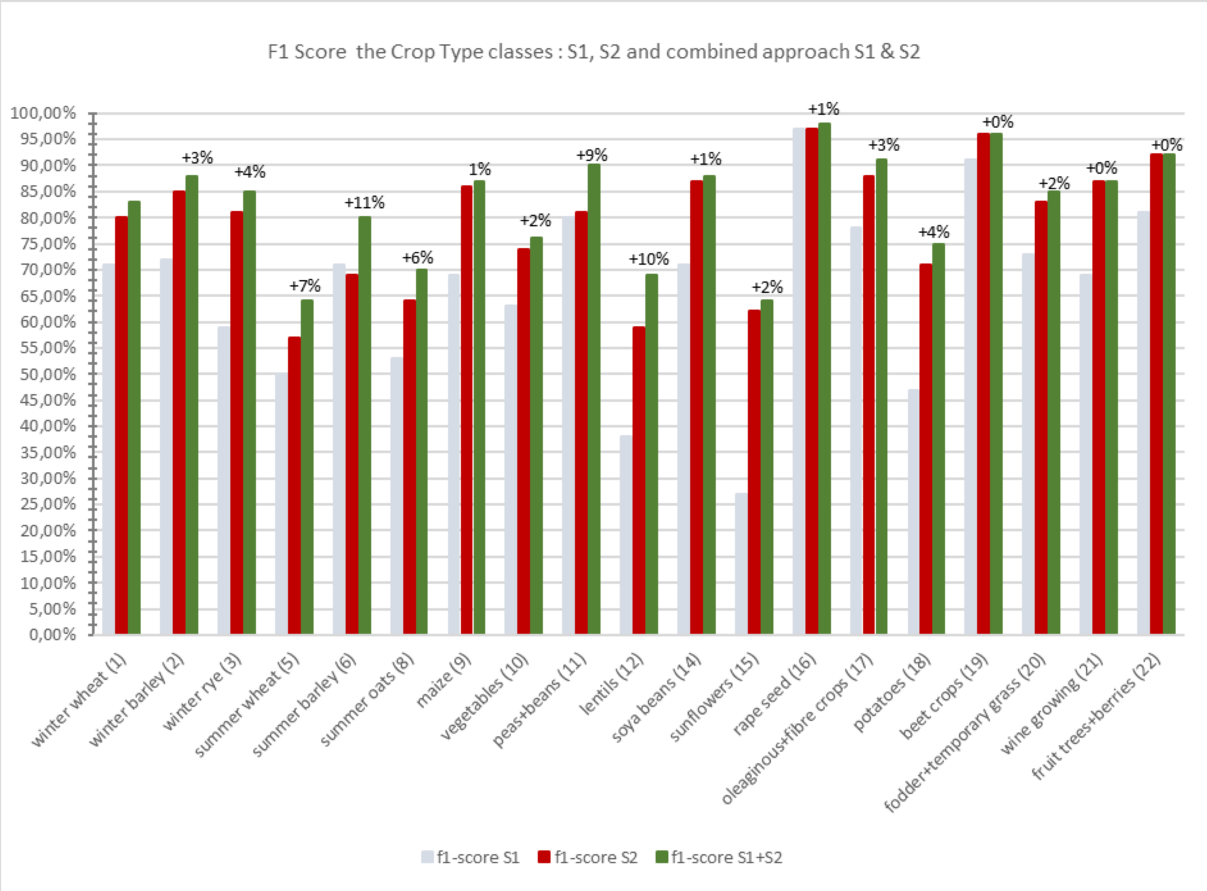


Figure 4-1: Phase 2 - F1 Score for all Crop Type classes: S1, S2 and combined approach S1 & S2 and the improvement of accuracies by using both sensors

FEATURES COMPUTATION AND SELECTION

As preceding tests in the frame of Task 3 (AD06) of this project have shown that the time features approach is an appropriate method to prepare the data for crop area and crop type classification this method is used for the calculation of the agricultural prototype in the demonstration site Central. The feature extraction step consists of computing the most discriminant variables from the time series for a specified time window to be used as input for the classification algorithm. Time features take several aspects into account: (1) spectral characteristics, such as the multispectral reflectance from S-2 or the backscatter coefficient from S-1, as well as derived indices, such as the NDVI or any other vegetation, chlorophyll or soil index, as well as SAR band rations; (2) temporal, such as the minimum, maximum or amplitude of a variable over a given time period; and (3) textural, such as the local contrast, entropy or any other variable derived from the co-occurrence matrix. More details on the time-features and their background are given in the AD06 (pp. 74).

PRE-CLASSIFICATION STEPS (FORWARD FEATURE SELECTION)

It turned out during the testing that not all time features have high impact on the classification result and its accuracies. In order to achieve maximum accuracy with less computational efforts and in a resource effective manner, only the most important features have been selected by method of grouped forward features selection. The forward feature selection is an iterative feature selection approach where the most informative features are selected This method guarantees that only relevant information goes into the classification and reduced the actual number of features from initially 676 to 221 for the final classification.

As already described in detail in WP33, the FFS is particularly suitable for taking the most out of a large time window. The number of selected features reflects not only the importance of specific features referring to the classification but also those of the most suitable time window (see Figure 4-2). The FFS for the CRT prototype for example selected predominantly S-2 features and, additionally, predominantly in the time window from Mid-July to Mid-Oct. However, even in the other time windows and also considering S-1 plays a role for the case of agricultural classification. Therefore, it is not recommended to leave out S-1 or to further constrain the time window.

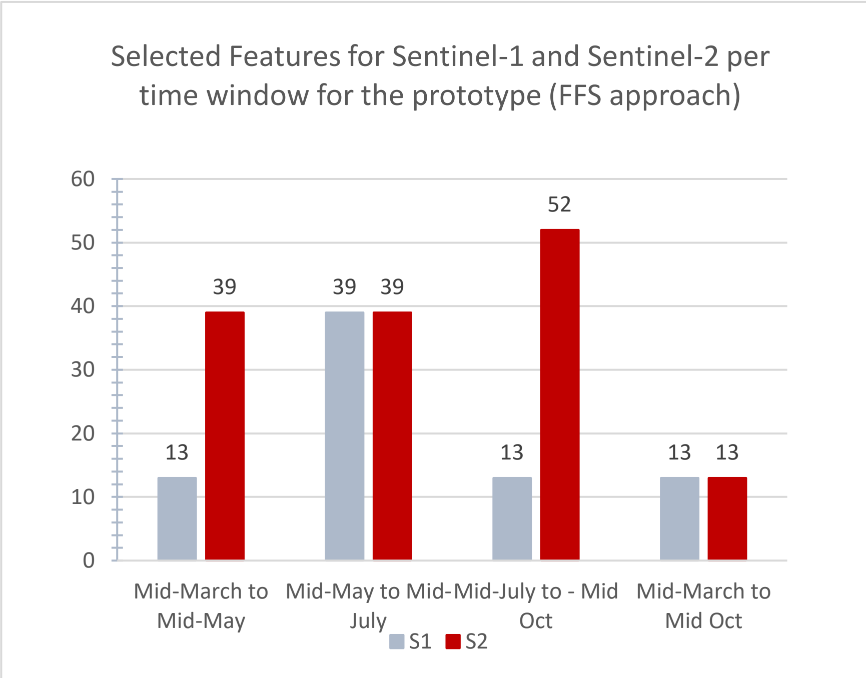


Figure 4-2: Phase 2 - Selected time features for the CRT-prototype classification on demo site Central

The finally selected features were then calculated for the whole raster data and used to train the final classification model. An independent accuracy assessment was then performed based on the test polygons, which have not been used at any point during model training (including FFS).

CLASSIFICATION ALGORITHM (RANDOM FOREST CLASSIFIER)

The classifications (crop type, crop mask, FFS) themselves consist of one or many numerical processes to finally allocate every pixel (in terms of a pixel-based classification) to one of the classes of a defined land cover typology. The applied non-parametric random forests approach was selected for the agricultural prototype. The algorithm is robust to data reduction and produces harmonized results. For further details on the classification method please see AD06 (pp. 82). The outcome of the classification are different layers depending on the number of classes: one layer for the predictions (classes), the class-probabilities (one layer per class), and three reliability layers (max. probability, breaking ties, entropy), which are combined to one single reliability layer.

For the validation of the classification result a confusion matrix was created which compared the final classification with a reference data set via cross-correlation. In this context the Producer's (PA) and User's

Accuracies (UA) were analyzed as well as the Overall Accuracy (OA) and the F-Score. These quality metrics are part of the overarching accuracy assessment as described in [AD06].

POST-PROCESSING

The final result is a pixel based product representing the most probable crop type per pixel. As users of such data are more familiar with vector based products, where each vector patch represents a parcel with a certain crop type, several tests were conducted to enhance the overall look-and-feel classification result but also correspondence to the needs of a user. The most common method to harmonize such pixel based products is a filtering of neighbouring pixels and an elimination of pixel cluster below a specified size, the minimum mapping unit.

For both, crop mask and crop type mask, the filtering approach applied a MMU of 4 pixel. This has a quite low impact on the initial classification result but highly enhances the products look-and-feel. In a kernel of 3x3 pixels, the class is determined to define the centre pixel value. With that method noise, consisting of single pixels and small lines of pixels, often representing parcel borders (due to spectral mixture with other land use and land cover like tracks between fields or the land use of neighbouring areas), can be removed. In a second step, a majority filter in combination with filtering along the longest border has been applied. As a pixel-based classification does not refer to geometries, or, more precisely, to the shape of a crop parcel, this filtering method gives the opportunity to converge the result towards what will be of "real" situation on field level. This final product appears much more harmonized but still fulfils the request of a pixel based product.

4.3 Method for crop type map – Demo-site Mali & South-Africa

The proposed method for the demo-site Mali is based on the ESA Sen2-Agri system which has been supported by ESA and recently developed as a generic free and open-source operational tool for crop mapping. It was demonstrated in several countries but never for crop type mapping over a large area in a smallholder cropping system. This standalone processing system made of modules generates a series of four agricultural products from Sentinel-2 (S2) and Landsat 8 (L8) time series acquired along the growing season of interest. It allows handling large volume of EO data in near real time and easily scale up to cover countrywide products. Two modules were tested over the demo-site Mali: (i) the dynamic cropland mask and (ii) the cultivated crop type map for the main crop types.

FEATURES COMPUTATION AND SELECTION

As described by Valero et al. (2016) and Inglada et al. (2015), the Sen2_Agri system (Inglada et al., 2015; Defourny et al., 2019) uses as input features for the classifier, 10-day gap-filled time series of S2 and L8 surface reflectance values. This means that a S2-interpolated image is computed every 10-day at 10 and 20 m resolution according to the S2 bands by a weighted linear interpolation. The no-data values were determined thanks to available missing value masks provided by the data preprocessing step, detecting pixels affected by clouds, cloud shadows or saturation effects. The weights used for the linear interpolation were computed by measuring the temporal linear distance between the interpolated date and the valid observations. In addition to the 10 spectral bands of S2, other features such as NDVI, Normalized Difference Water Index (NDWI), and brightness are computed for each of these S2-interpolated image. Unlike for the other demo-sites, the Brightness index is defined by Sen2-Agri as follows:

$$\mathrm{Sen2Agri} = \sqrt{\rho_{\mathrm{Green}}{}^2 + \rho_{\mathrm{Red}}{}^2 + \rho_{\mathrm{NIR}}{}^2 + \rho_{\mathrm{SWIR}}{}^2}$$

Therefore 13 features compiled for the 25 dates of the time series corresponding to the growing period are available for the classification process. All are used in the crop type classification process. For the cropland extent mapping, 17 phenological metrics derived from the NDVI profile (Valero et al., 2016) and 5 statistical metrics respectively computed from the NDWI and the Brightness time series (Valero et al., 2016) are compiled as input data.

PRE-CLASSIFICATION STEP (CROPLAND MASK)

As a preliminary step for the crop type mapping, the detection of the cropland extent is required and completed using a supervised Random Forest (RF) classifier exploiting a set of 27 temporal and statistical features specifically designed to depict the growing cycle. This cropland mask is mandatory as no LPIS exists in most of the food unsecured countries. The training data are composed of a set of cropland and non-cropland spatial polygons spread over the demo-site. The methodology applied to map the cropland class relies on the "cropland" definition proposed by the JECAM network in the framework of GEOGLAM. According to this definition, the "annual cropland from a remote sensing perspective is a piece of land of a minimum 0.25 ha (minimum width of 30 m) that is sowed/planted and harvested at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation" (JECAM, 2017). In this definition, perennial crops and fallows are excluded from the cropland class.

CLASSIFICATION ALGORITHM (RANDOM FOREST CLASSIFIER)

The crop type map is a map of the main crop types at 10 m resolution. Similarly to the cropland mask, the processing chain for the generation of the crop type map follows a supervised RF classification approach built on field data. The identification of individual crop types is conducted within the previously created cropland mask. All the features, i.e. the surface reflectance for the 10 S2 spectral bands and the 3 spectral indices are considered as input to establish the RF model at the stratum level. Indeed, to cope with the agro-ecological gradients and the diversity of cropping systems, a stratification layer is defined to reduce the heterogeneity observed for each main crop of interest. As the demo-site Mali covers two main agro-climatic zones: the Sudanese and the Sudano-Guinean zones that contrast mainly in the annual amount of rainfall. The two major natural regions depicted by the Projet d'Inventaire des Ressources Terrestres (PIRT) are used as stratification: Haut Bani-Niger and Koutiala.

ACCURACY ASSESSMENT

For the prototype of the demo-site Mali, the in situ information were split randomly between training samples (75%) used to feed the classification process and validation samples (25 %) used to assess the product accuracy through the computation of several well-known criteria (Overall Accuracy, and F1-Score of individual crops). In addition, a spatially independent validation data set has been prepared to further assess the quality of the cropland mask.

For the prototype of the demo-site South-Africa, the very comprehensive in situ information were split randomly between training samples (5%) for calibration and reference samples (95%) for validation.

5 Prototype Implementation

This chapter shows the prototypical implementation of the agricultural prototype, consisting of a crop mask (CRM) and a crop type (CRT) map. Firstly, the integrated EO and ancillary data is described (section 5.1), followed by the demonstration of the results of the actual prototype in the demonstration site as well as the validation (section 0), and lastly, the description of the dataset properties and its metadata, referring to **D14.4: "D44.2b – Data Sets of Crop Area and Crop Status/Parameters Products"**. For a description see section 0.

5.1 Data and Processing Setup

The following subchapters will give further details on the creation of a prototype for a potential future HRL Agriculture, implemented on the West, Central and Mali demonstration site. The following sections include the input data and their integration (section 5.1.1) as well as the several pre-processing steps (section 5.1.2) and the experimental setup for preparing and performing the classification (section 0).

5.1.1 Input Data

In the following sections, the input data of the Demonstration Sites West, Central and Mali are described. This contains the image data, reference and field data as well as Copernicus HRL 2015 data.

5.1.1.1 Demo-site West

For the French West tiles, in the first phase, the LPIS dataset for the reference year 2016 only was available over France, and the next version for the year 2017 should become available this summer 2018, hence the use of images dated from the year 2016 only. Since no SAR image covering this time window has been provided, and a complete re-processing of the raw archive of S-1 images was deemed too time-consuming, the process only used optical data from 2016. Regarding these optical data, it has been decided to split them into three trimesters – January to Mars, April to June and July to September. The aim of this approach was to establish a spectral signature over the study year (2 January of 2016 to 28 September of 2016) for the different classes, through several indices such as NDVI, NDWI or BRI (the description of the indices and their time features are available in the methodology section in chapter 4.

For all those reasons, the dataset used was restricted to 91 S-2 images, after applying the cloud mask to put aside the unexploitable captures. Those images are composed of 10 bands at 10m and 20m and are atmospherically corrected. In order to increase the accuracy of the results for further classification, it should be useful to integrate S-1 in the next iteration of the task 4.

For the second phase, only the LPIS from 2017 was available. In order to put to the test the classifier, in particular over the tile 31UFR, decision has been made not to add the Belgian LPIS from the same year.

Only optical time series from S2 A and B has been used – with a further selection of the best scenes only, in order to avoid the potential contamination of the time series by undetected clouds, which remains a recurring issue.

The numerous SAR temporal features, despite their clear improvement of the classification results, are yet to be fully understood in terms of their respective contribution to this improvement. The amount of computation required not only for the SAR pre-processing but also for the temporal features would be staggering at a pan-European scale. The focus of this second phase was to strongly upgrade the quality of the optical classification (even on a rather cloudy region such as the North of France and Belgium). With an OA up to 10 points compared to the results of the first phase, slightly exceeding the 85% mark, there is a greater confidence that agricultural products could be produced yearly over the whole Europe, with a restrained use of SAR datasets over only very cloudy regions. Notwithstanding, provided the analysis ready data developments and availability will evolve rapidly in the near future, the implementation of SAR features and multisensory integration will be part of the classification workflows and big data analysis on a normal basis at the pan-European and global scales land cover productions. Within ECoLaSS, the tests and prototypes are designed under such a frame, to set proof-of-concept products and workflows.

In contrast to SIRS, UCL used the LPIS datasets dated from the year 2017 which was concurrent to sufficient observations acquired by Sentinel-1 and Sentinel-2 satellites. For Sentinel-1, scenes of the co-polarized (VV), and cross-polarized (VH) backscattering acquired between January and December 2017 were used for calculating the backscattering ratio (VV/VH) which, then, used together (VV, VH, VV/VH) as added sources of information (to optical data acquired by Sentinel-2) for increasing the ability of crop identification. For Sentinel-2 data, observations acquired between January 3rd and December 2nd, 2017 and composed of bands 3:8 and bands 11:12 in addition to NDVI, NDWI and BRI were used after being atmospherically corrected and cloud masked. Hence, the crop type identifications was applied on features fundamentally originated from optical and Synthetic Aperture Radar (SAR) datasets after being undergone to various steps of pre-processing chain detailed in section 5.1.2

Two LPIS datasets were necessary to cover the study area: one over Belgium, called LPIS_OPENDATA_PARC_AGRI_2016 and a second over France referred to as LPIS_Registre_Parcellaire_Graphique_2016. As previously mentioned, the French LPIS for reference year 2016 was used since France has not provided the 2017 version yet at the time of the production of this layer. It is also worth underlining that this LPIS does not contain vineyard parcels, so this class will not show up in the results.

Two of the four Sentinel-2 tiles also covered part of Luxembourg but the LPIS was not complete on this zone. Only the geographical information could be exploited but no information appears in the attribute table.

Besides, the use of the LPIS fulfills two main objectives:

- To select training samples of different cropland types;
- To homogenize the results on the parcels.

The Copernicus 2015 High-Resolution layers (HRL) provide information for the whole of the EEA-39 area on 5 specific land covers: the level of sealed soil (imperviousness), tree cover density and forest type, grasslands, wetness and water, and finally, small woody features. What is more, these products were automatically computed from time series satellite imagery from 2015, before being manually enhanced and finally validated using very high-resolution (VHR) imaging. This is the reason why they are used here to generate 4 layers of non-cropland samples: urban areas, forest, water and grassland. In the final product, these non-cropland classes will be merged in a unique class to establish an overall accuracy of the prototype as detailed in section 5.2.1.

For the phase 2, the production of the 2018 prototypes considered separately the Belgian administrative regions, i.e. Flanders in the north and Wallonia in the south of Belgium. The respective 2018 LPIS layers were collected, recoded and used for each production. The initial typology of Flanders and Wallonia respectively includes 318 and 271 different LPIS labels including permanent crops and distinction of crop types according to the destination (industrial versus consumption, etc.). These labels have been recorded into 117 and 116 crop types respectively for Flanders and Wallonia. However many of these crops are very rare and have been automatically filtered out by the classification process.

As the version 2.0 of the open source Sen2Agri system was released in May 2019 and the version 1.0 of the open source Sen4CAP system using Sentinel-1 and Sentinel-2 time series was just released in the fall 2019, both systems have been run and their respective results compared. The Sen4CAP calibration dataset is more restrictive and discard any crop type class with less than 30 parcels containing at least 10 Sentinel-2 pixels.

5.1.1.2 Demo-site Central

Phase 1:
The 2017 prototype calculation for the demo-site Central is based on both, S-1 and S-2 data, covering a time period from 2017/15/03 to 2017/14/11. In case of the S-2 images, the cloud cover was constrained to a maximum of 90%. Pre-processing is described in [AD D32]. Table 5-1 shows the resulting distribution of images per tile and satellite.

Table 5-1: Number of S-1 and S-2 images per tile used for the calculation of the agricultural prototype for the demonstration site Central.

					PHASE 1					
	32TMT	32TNT	32TPT	32UMU	32UMV	32UNU	32UNV	32UPU	32UPV	SUM
S-1	38	105	66	54	37	115	103	92	103	**713**
S-2	46	47	45	47	45	49	44	42	44	**409**
SUM	**84**	**152**	**111**	**101**	**82**	**164**	**147**	**134**	**147**	**1122**
					PHASE 2					
	32TMT	32TNT	32TPT	32UMU	32UMV	32UNU	32UNV	32UPU	32UPV	SUM
S-1	47	87	75	43	48	104	92	90	44	**630**
S-2	64	29	55	65	34	63	59	55	27	**451**
SUM	**111**	**116**	**130**	**108**	**82**	**167**	**151**	**145**	**71**	**1081**

Phase 2:

As for the 2018 prototype the classification again uses the combination of optical and radar satellite imagery of S-1 and S-2, this time applying two limiting factors in order to avoid artefacts within the time features: maximum cloud cover and minimum tile cover.

The cloud cover was restricted to 90% for S-2, as in phase 1. Due to generally high cloud cover during the year 2018, some S-2 tiles show very limited number of viable imagery (see Table 5-1). Another issue was the highly fractioned imagery for the demo site Central. As it turned out in the testing that strong variations in coverage would lead to artefacts in the time feature calculation, a tile coverage parameter has been integrated. Optical data haven been used covering 80% of a S-2 tile and more, radar data covering at least 20% of the same area.

The analysed time period stretched from 2018/15/03 to 2018/10/14 and four time windows focusing on the beginning of the vegetation period (Mid-March to Mid-May), on the main growing period (Mid-May to Mid-July), the withering and harvesting period (Mid-July to Mid-Oct) plus a time window comprising the whole vegetation period from Mid-March to Mid-Oct. The amount of available imagery for phase 2 per tile and sensor is also given in Table 5-1.

The differences in number of scenes per tile are caused by the time window, the cloud cover, the tile cover and the location of the demonstration site. Regarding the time perspective: since data acquisition starts in mid-March, where the cloud cover in the Central demo-site is typically very high, there are a lot of S-2 scenes which do not fulfil the requirement of less than 90% cloud cover. In terms of location, the differences in S-1 coverage can be explained by the overlapping swaths (see Figure 5-1). The 32UNU and the 32TNT tile for example have the highest number of scenes and as the Figure 5-1 and Figure 5-2 show, the two swaths have the biggest overlap there, so there are more images available that mostly cover these tiles.

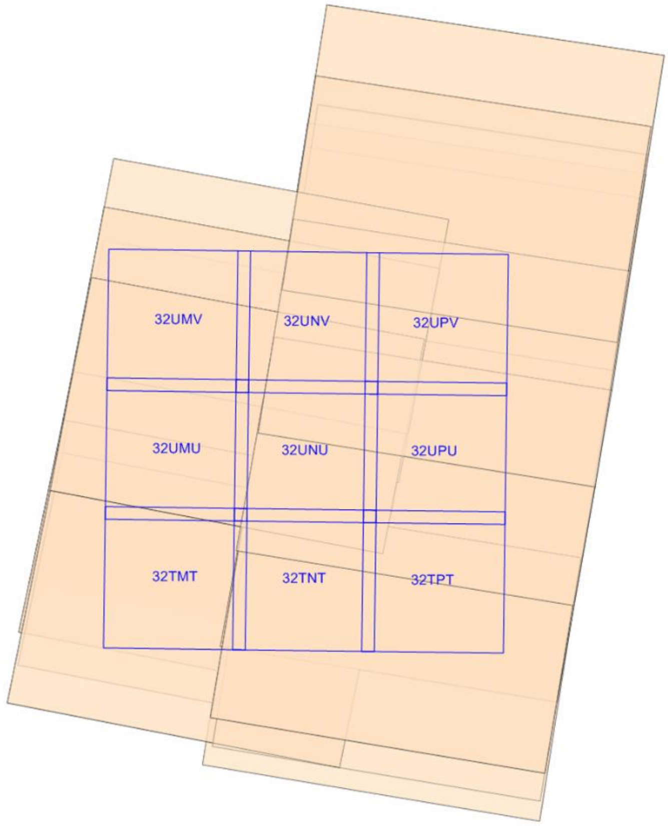


Figure 5-1: Coverage of the Demo-Site Central with Sentinel-1 data (orange) and Sentinel-2 data tiles (blue)) for the reference year 2017.

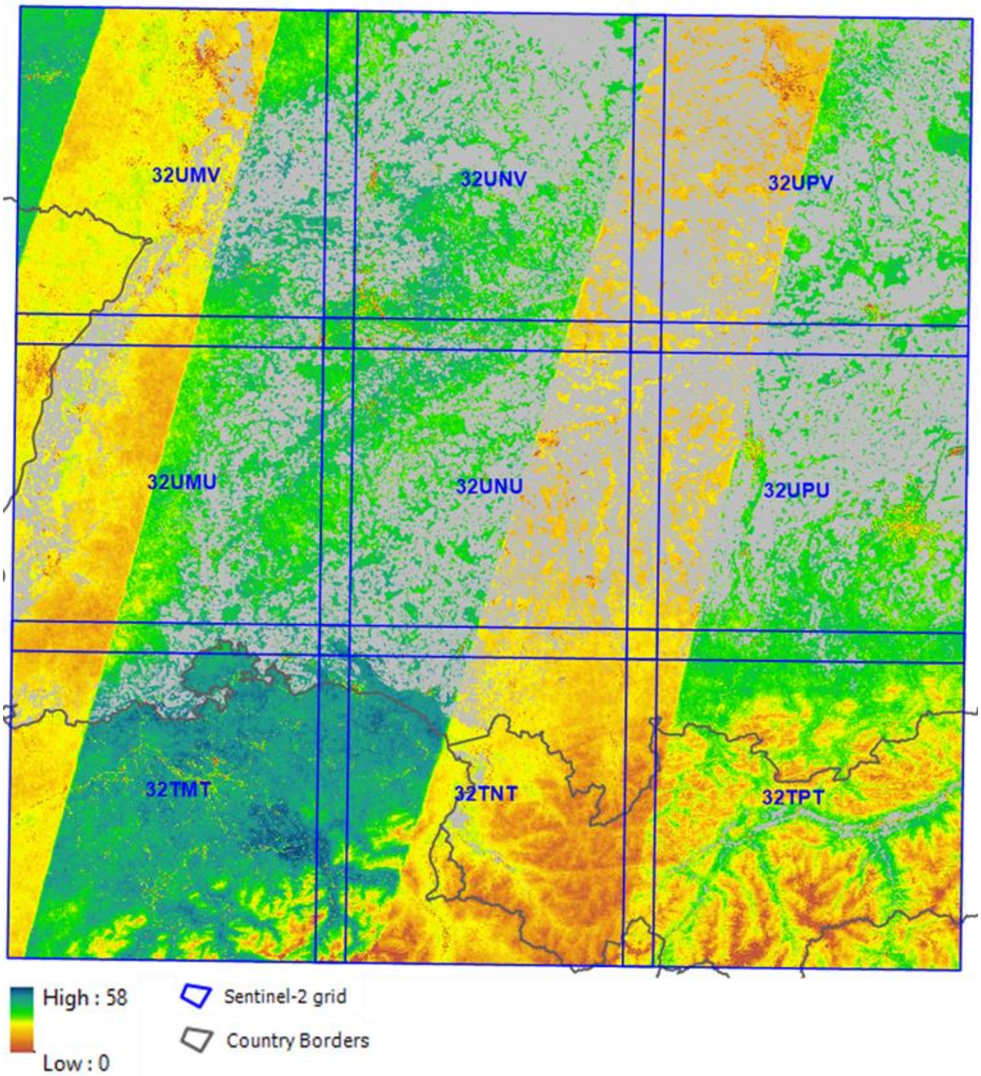


Figure 5-2: Data Score Layer of the Demo-Site Central depicting the amount of valid coverage per pixel for the reference year 2018. *© EuroGeographics*

Besides the EO data for the classification, reference data were necessary for the calibration of the classification model as well as for the validation of the classification results. A necessary precondition is that the reference data base gives detailed and up-to-date information on crop types.

Phase 1:

Phase 1 used LPIS data base for both, crop mask and crop type classification. Underrepresented crop types and not meaningful classes considering the differentiation between crop types (e.g. sweet potatoes, field margins, and agricultural roads) were excluded from the LPIS dataset. The result of the selected LPIS polygons for the reference year 2017 is visualized in Figure 5-3. This remaining LPIS data was grouped into meaningful crop groups (see section 0).

A further step in the processing chain of the Agricultural prototype is to explore the utilised reference data. The area covered by the central demonstration site includes parts of two different countries (Southwest-Germany and West-Austria) with different crop type nomenclatures, which led to several compatibility challenges when merging the data into joint classes and, thereafter, groups of classes. After some adjustments 16 classes (15 crop groups plus one rest group for “others”; see **Table 5-2**) could be identified for the reference year 2017, based on logical considerations of similar crop types, and their spectral characteristics and temporal appearance. The derived crop group classes served as input to the prediction in a further step.

Table 5-2: Phase 1 - Number and name of crop groups and their abbreviation for the reference year 2017.

ClassID	Crop Group	Crop Group Abbreviation
1	Agrarian Grass	AgrGrass
2	Fallow	Fallow
3	Fruit Trees	FruitTrees
4	Legume	Legume
5	Maize	Maize
6	Others	Others
7	Potatoes	Potatoes
8	Strawberries	Strawberries
9	Sugar Beets	SugarBeets
10	Summer Crop	SummerCrop
11	Summer Rape	SummerRape
12	Sunflowers/Topinambour	SunflTopinamb
13	Vegetables	Vegetables
14	Winegrowing	Winegrowing
15	Winter Crop	WinterCrop
16	Winter Rape	WinterRape

As can be seen in Figure 5-3, the majority of the tiles of the demo-site is covered well by the reference data. Due to LPIS data of Bavaria and for Switzerland not being available to the consortium, the prototype was developed for all nine tiles with training data distributed within the area of Baden-Wurttemberg and western Austria.

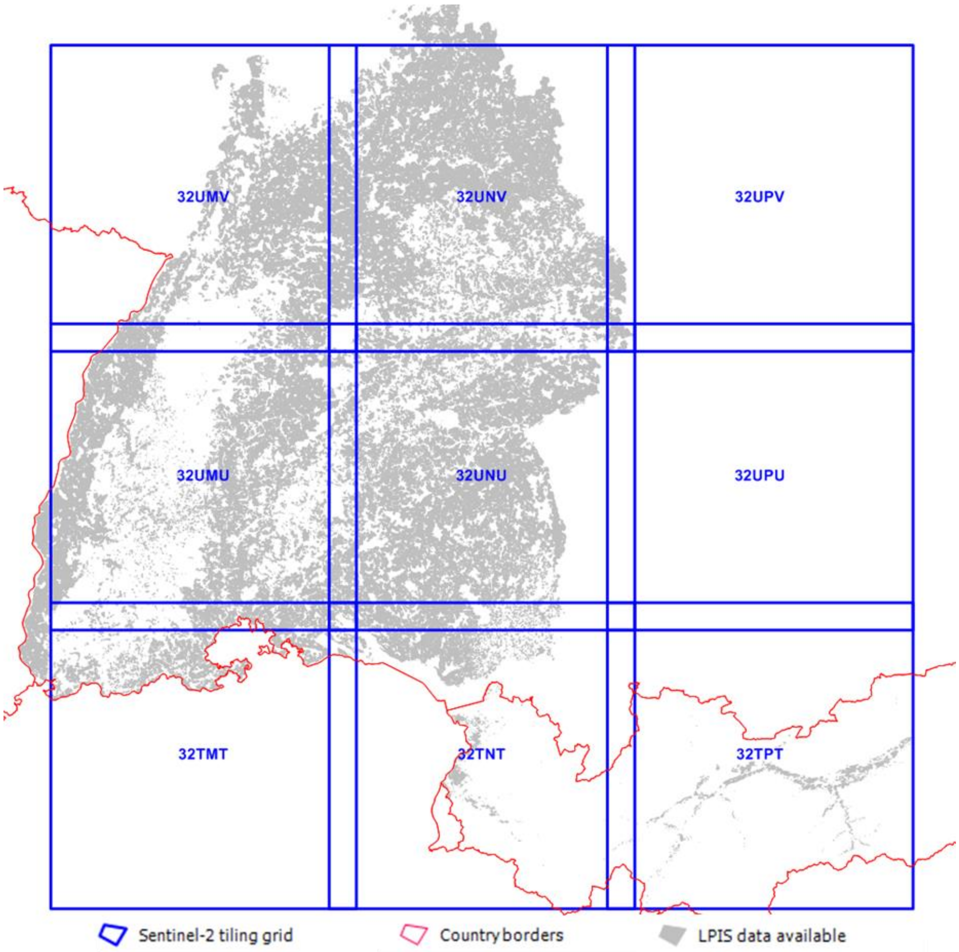


Figure 5-3: Phase 1 - Coverage of the Demo-Site Central with LPIS data (permanent grassland excluded) for the reference year 2017 used for training and validation.

Phase 2:

In contrast to phase 1, 2 different reference data sources haven been used in phase 2: The crop mask bases on LUCAS 2018 data, complemented by validated samples derived from the HRL2015 and by manual sampling. The aim was to test the potential of LUCAS data as source of reference providing large-scale coverage for a future Pan-European crop mask. An overview of the LUCAS classes used for the crop mask classification can be found in WP33.

As for the crop type mask, data of the LPIS 2018 were used, where available. In phase 2, these were LPIS data from Baden-Wurttemberg, Bavaria and Austria.

The main focus of the second project phase was to create an extended crop type nomenclature which is suitable for crop type differentiation and at the same time suitable for a Pan-European roll-out. The proven crop type nomenclature is given in the table below:

Land cover mask		Crop Group		ECoLaSS Paneuropean Crop class	Crop Type mask: ECoLaSS Crop type classes for CTM	
Crop mask	1	Cereals	11	Winter cereals	1	winter wheat
					2	winter barley
					3	winter rye
					4	winter oats
			12	Spring/summer cereals	5	summer wheat
					6	summer barley
					7	summer rye
					8	summer oats
			13	Maize	9	Maize
			14	Rice	999	Rice
	2	Vegetables, dry pulses, berries	21	Vegetables	10	vegetables
			22	dry pulses/legumes	11	peas and beans
					12	lentils
					13	legumes
	4	Industrial crops	41	Soybeans	14	soya beans
			42	Sunflower	15	sunflowers
			43	Rapeseed	16	rape seed
			44	Other oleaginous, fibre, biofuel, and beverage crops	17	oleaginous+fibre crops
	5	Root/ tuber crops	51	Potatoes	18	potatoes
			52	beet crops	19	beet crops
	6	Fodder crops	61	Temporary grasslands (<5 y.) + fodder/agrarian grass	20	temorary grassland
	7	Permanent crops	71	Grape vines	21	wine growing
			72	Olives groves	999	Olives groves
			73	Other permanent crops		fruit trees/orchards
					22	shrub fruits

Table 5-3: Phase 2 - reworked crop type nomenclature for crop type classification for the reference year 2018 comprising the main crop groups, crop classes and crop types in a Pan-European context which could be adapted to the regional context.

Permanent grassland cover has been left out, since a highly accurate High Resolution Layer focusing on those specific grassland areas (HRL2015 published, HRL2018 still in production) already exists. Another reason is that the definition of crop area differs in the Pan-European context: some countries include (all or even some specific types of) permanent grassland in the crop area, some exclude. It serves transparency to leave this decision up to the user, if permanent grassland should later on be re-integrated by using the HRL2018 data.

The approach of phase 2 clearly focused on the Pan-European perspective and its necessary preconditions. Considering crop classification at this scale, these preconditions are

- reference data sources which are easily available, up-to-date and accurate and
- a crop type nomenclature that suits both, local as well as Pan-European conditions

The Crop Mask prototype has been produced with LUCAS data as a potential source of nearly Pan-European reference basis.

The crop type classification, as already mentioned, orientates towards the LUCAS crop hierarchy but comes with a complete re-thinking of a hierarchical crop classification nomenclature.

Thanks to the approval for using the LPIS data for Bavaria by the Ministry of Food, Agriculture and Forestry, the coverage with reference data of the Central site for the second project phase (reference year 2018) could be significantly improved. Most parts of the demonstration site could be covered with reference data, except for Switzerland, Liechtenstein, the French part in the West and the German parts outside Baden-Wurttemberg (Rheinland-Pfalz, tile 32UMV).

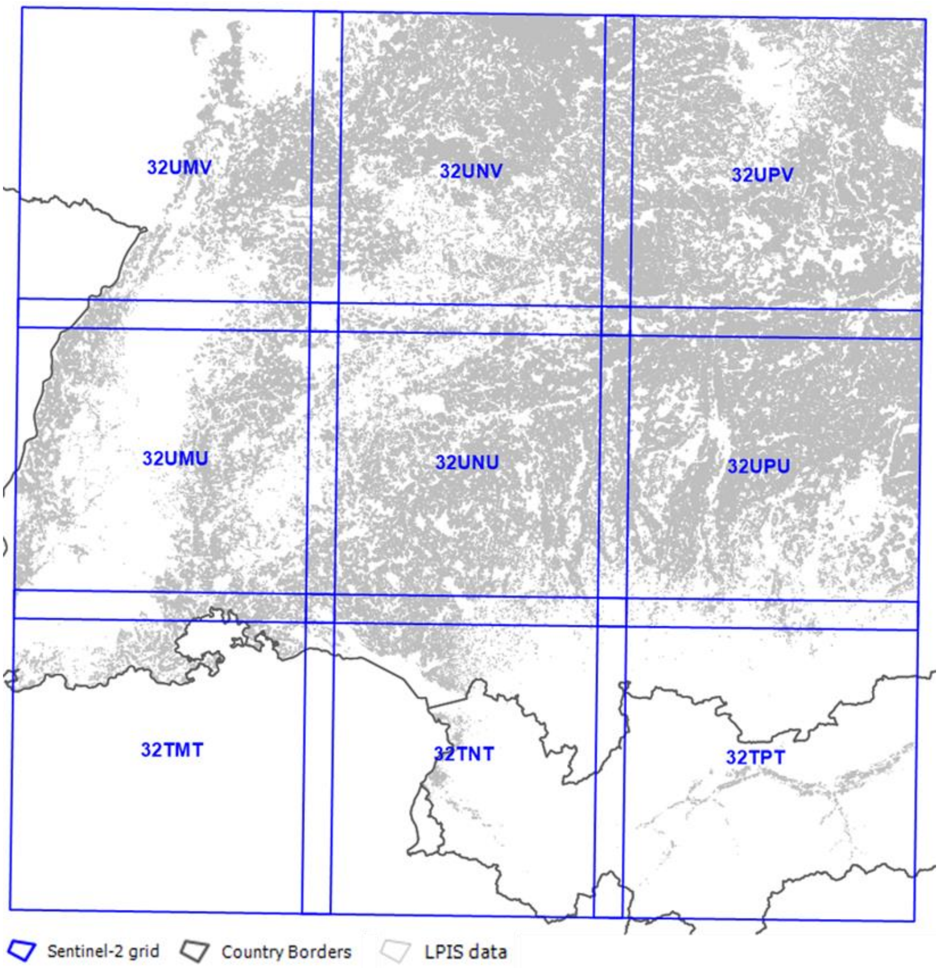


Figure 5-4: Coverage of the Demo-Site Central with LPIS data (permanent grassland excluded) for the reference year 2018 used for training and validation.

© EuroGeographics; © Agrarmarkt Austria © MLR BW © StMELF

5.1.1.3 Demo-site Mali

As proposed in the Sen2Agri system, all S2 and L8 imageries acquired over the demo site during the 2017 and 2018 growing seasons starting on the 1st April and ending on the 15 December were automatically downloaded from the ESA's Scientific data Hub (https://scihub.copernicus.eu/) and United States Geological Survey (USGS) (https://landsat.usgs.gov/) data provider platforms. For 2017, the total number of images acquired by S2 and L8 were 1811 and 273 images respectively with less than 90% cloud cover while no Sentinel-1 data have been used over Mali.

Crop type in situ observation were made available to ECOLASS by the Institut d'Economie Rurale (IER, Bamako) through UCLouvain as they were collected during the 2017 and 2018 growing seasons in the context of the IER improvement of the Sen2Agri campaign. These comprehensive field campaigns were carried out in 27 villages spread covering the demo-site Mali allowing-to capture most of the diversity of the main crop types (Figure 5-5). For 2017, the crop type field data set is composed of 6591 cultivated fields including both major and minor crop types (Table 5-3). The data were collected along the roads associated with crop type information following the JECAM guidelines (JECAM, 2017).

Table 5-4: Description of the 2017 in situ datasets in terms of number of samples per class as provided by IER (Mali).

Stratum	Cropland samples / ha	Non-cropland samples / ha
Koutiala	4205	322
Haut-Bani Niger	2386	698

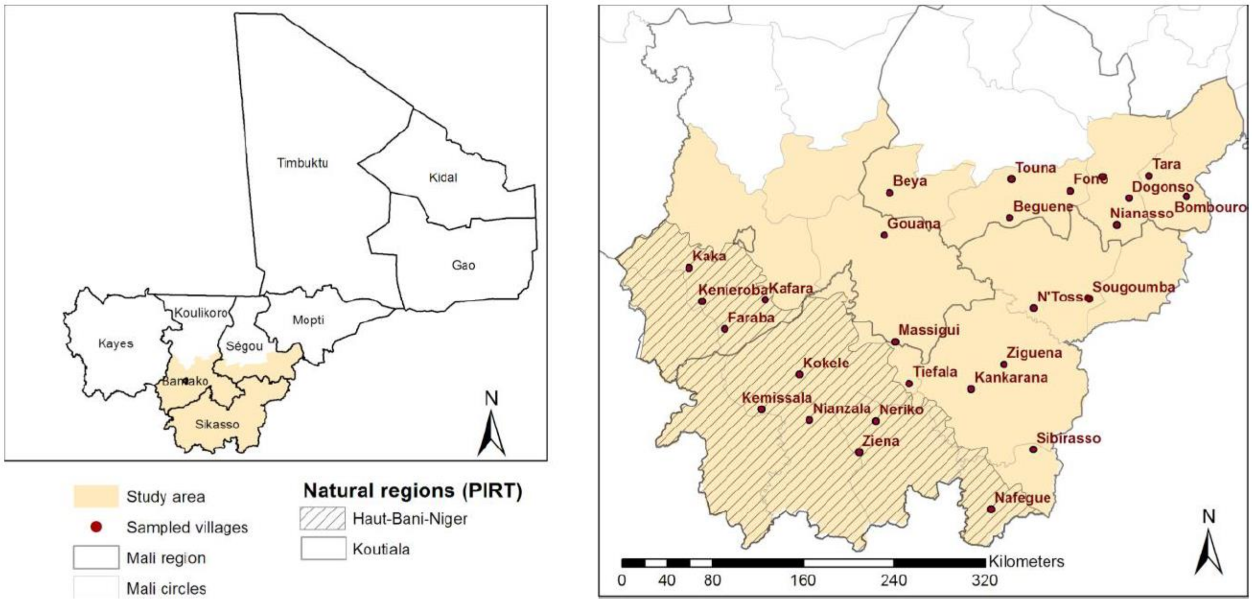


Figure 5-5: Demonstration site (in color) covering all the Sikasso region and part of the Segou and Koulikoro administrative regions. Field data were collected around 27 villages spatially distributed over the demo site (presented in red). The natural regions of the PIRT are presented in stripped line for the Haut-Bani-Niger region and in transparent for the Koutiala region.

In each sampled village, a team of field operators collected geotrace (polylines) on the border of the fields alongside vehicule roads and secondary roads (Figure 5-6), following the JECAM guidelines. Field operators traveled on motorbikes to be able to take the small paths. When fields were too small to record geotrace, operators collected georeferenced points in the middle of the parcels. All visible fields along the major roads were sampled while a minimum of 30% of the fields were collected along secondary roads or paths. For each collected sample, the field operators were asked to identify and encode the crop type as well as the relative position of the field with regards to the geotrace, i.e. left or right. The data collection was performed using the GeoODK data collection free app (http://geoodk.com/) on smartphones with GPS. On a daily basis, all collected data were sent to a web platform for near real time quality control.

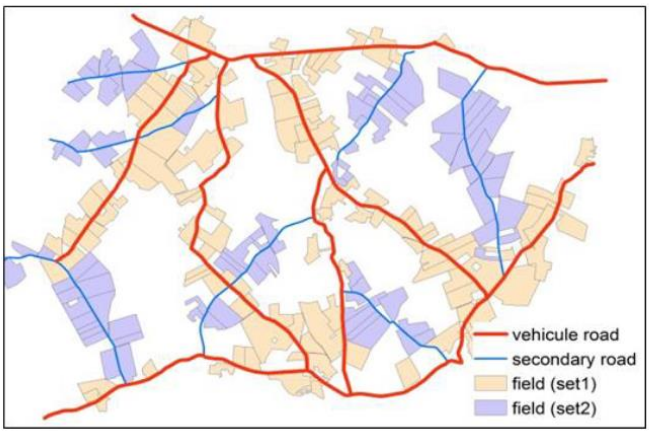


Figure 5-6: JECAM protocol to sample fields along main and secondary roads (JECAM, 2017)

In order to transform geotraces into polygons delineating individual fields, visual digitalization on recent Sentinel-2 data was completed. When no field limits were visible on the Sentinel-2 image, the operators were asked to draw a polygon of 20m width following the geotrace to limit the errors. The same methodology was applied to field points. Nevertheless, as field points were often collected in area with very small fields (e.g. bottom of valley, rice fields, peanut fields), the operators were asked to draw a small polygon of 20 X 20 m (400^2 m) around the field point.

In order to complement this very comprehensive in situ crop type survey, a non cropland dataset including +2,100 polygons corresponding to all land cover types except cropland, was built by visual interpretation of Sentinel-2, the ECoLaSS VHR images complemented by sometimes outdated VHR imagery from Google map or Bing. Those polygons were pre-selected based on a stratified random sampling over strata derived from a land cover map (GlobeLand30). GlobeLand 30 provides a generic land cover map of 10 classes at 30-m derived from Landsat images for 2010 (Chen et al, 2015). The stratified random sampling guarantees the representativeness and the good spatial distribution of all classes in the non-cropland dataset (including grassland, forest, water, urban, wetland, bare soil).

The availability of these valuable and comprehensive in situ data really justified the acquisition of VHR imagery of the Mali. The VHR images allowed the delineation of the Farming Management Unit observed from the ground and were found instrumental for the quality control of the in situ data collected during the growing seasons.

5.1.1.4 Demo-site Western Cape Province (South-Africa)

As agreed at the review meeting, the Western Cape province of South-Africa was also planned as a demo-site extending very significantly beyond the initial Free State demo-site (South-Africa). The Western Cape Department of Agriculture was indeed very interested by the demonstration of Sentinel-2 derived agriculture products as planned in the ECoLaSS project. This Department mapped in 2017 all the agricultural parcels and the associated crop types for the entire province. They made available for this ECoLaSS demonstration the 262.985 polygons distributed in 16 crop types over an area of 90.608 km² (Figure 5-7). The total area of the Western Cape Province is about 129,500 km² but one quarter does not correspond to any arable lands.

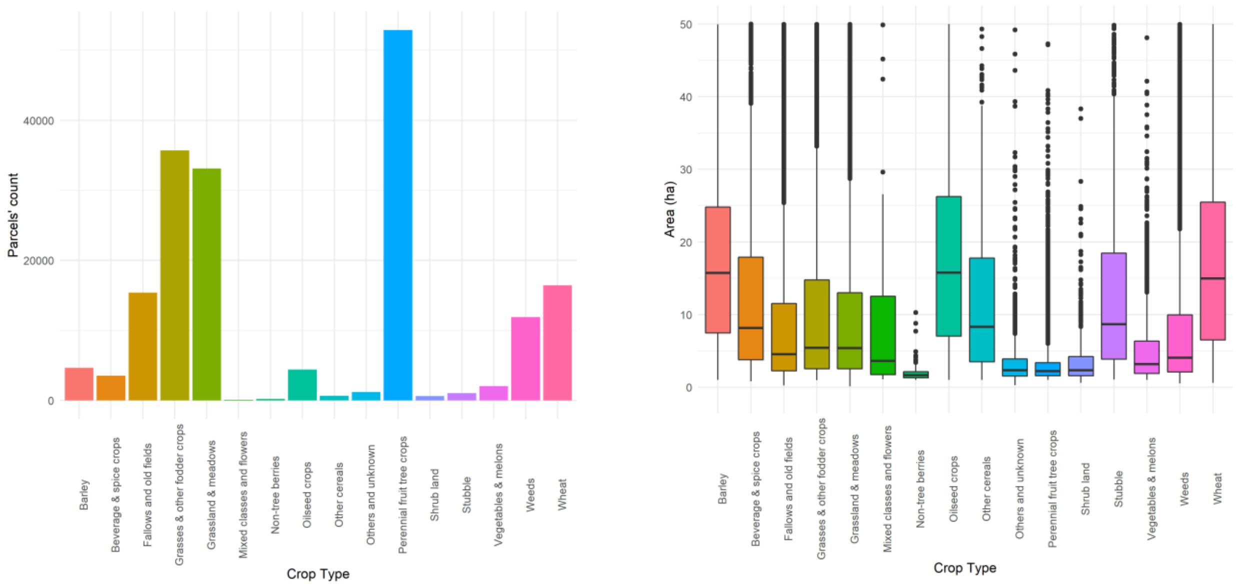


Figure 5-7 Distribution of 2017 crop types and field sizes as provided by the Western Cape Department of Agriculture.

5.1.2 Pre-processing

The Pre-processing chapter provides an overview of the pre-processing steps for the Demo-site West, Central as well as Mali. Therefore information about the Image-data and the HRLs 2015 are given.

5.1.2.1 Demo-site West

For the first phase, the WEST demonstration site data, over the 6 S-2 tiles (31UES, 31UER, 31UFS, 31UFR, 31UFQ and 31UEQ) has been pre-processed by the Joanneum Research for the year 2016, as described in the report of WP32. Due to the strong cover in the time series, only 91 images could be deemed exploitable after cloud mask application over the French West tiles. It is worth mentioning that the pre-processing of the backscattering was constrained to Sentinel-1 data acquired within the Belgian borders. Thus, the input features originated from Sentinel-2 data and, hence, the classification results were constraints to the same extent LPIS (SIRS)

For the phase 1 of UCL production, the level-2A processor named as MACCS; Multi-sensor Atmospheric Correction and Cloud Screening; was used in processing Sentinel-2 data (tiles 31UES, 31UFS and 31UFR) for detecting clouds and their shadows, estimating aerosol optical thickness (AOT), water vapor and correcting for the atmospheric effects. MACCS was jointly developed by CESBIO and CNES, as described in the report of WP32. A gap-filling algorithm employing a linear interpolation was applied every 10 days in order to obtain time series of bands 3:8, bands 11:12 having the resolution of 10 and 20 m. Bands of 20 m resolution were resampled to 10 m resolution and, then, NDVI, NDWI and BRI were calculated.
Also, the Sentinel-1 backscattering datasets (VV, VH, VV/VH) were undergone to various steps of pre-processing chain including calibration, radiometric normalization, terrain correction and speckle filtering. Then, pixel based statistics were calculated a time interval of two months from the backscattering time series resulting in input features of mean, standard deviation, coefficient of variation and the 10th, 25th, 50th, 75th and 90th quantiles.

For the French West tiles, for phase 1, as the LPIS dataset comes from farmers' declarations and from different production sites, the French and the Belgium datasets required a harmonization step. Indeed, nomenclatures from both zones do not match due to divergence in the agricultural policy of each country. This leads to the impossibility of combining certain different classes issued from those incompatible LPIS together.
Decision has been made to take into account only the most representative cropland types to circumvent this difficulty. The classification is therefore not made on all the declared classes, which minimizes the confusion on exceptionally rare classes. For future production deployment at larger scale, this step will have to be done for each kind of landscaping set.

Once this harmonization is made, several criteria for the set of samples per class have been drawn:

- Containing approximately 50 elements;
- Geographically well-defined over the all area of interest (AOI), using a regular fishnet;
- As pure as possible in order to restrain the mixed pixel effect, ensuring a more precise classification.

After the generation of points for each part of the fishnet, a buffer is made on those points and the zonal statistics are computed – this step requires the clearest image, without atmospheric veil that could have been missed by the cloud masking.
It is worth noting that LPIS is based on declaration only, resulting in potentially mixed or flawed class assignation, making the third and last criterium a crucial step in the sampling design process. What is more, only one type of cultures if declared, even though other crops are cultivated in addition during the given year, such as catch crops – this is a clear issue for the success of the classifier.

Thus, for each set of samples per class, the mean and the standard deviation is calculated. Only samples within one standard deviation of the mean are selected, and assumed to be of the "purest" form. This step is closely intertwined with a visual check of those samples, to assess the effective selection inside a parcel, or the absence of any remaining cloudy veil, for each temporal window. The Table 5-5 below shows two examples of samples which can't be keep because of calibration or validation issues. Finally, the samples are randomly selected in order to keep one sample per class in each part of the fishnet.

Table 5-5: Example of visual check on the samples.

	The sample collected is not pure enough, statistically and visually: it is then remove of the selection.
 	The sample should normally be keep for the set, if the decision was based only on statistics and on the look and feel. However, on the right image, the sample obscured by the shadow of a cloud; this will alter the image statistics in the time series during classification. It is then removed from the selection.

In the second phase, selection of samples has been directly made from the LPIS parcels. This allows "imperfection" of fields (such as slight borders contamination, or shadows of trees divided cultures for example) to be integrated into the classifier model and therefore learnt correctly. The smallest classes (with a number of fields below 50) have been excluded from the sampling, and the smallest fields for other classes were discarded.

The unified pan-European nomenclature, as shown in Table 5-3, was then used to merge the rest of the fields into the defined classes.

For UCL, the Belgian datasets of the tiles 31UES, 31UFS and 31UFR of the year 2017 were used with some constraints which are 1) applying a buffer of 15 meters from the polygons borders for avoiding mixels, 2) ignoring fields of small areas which are represented by few (sometimes scattered) pixels (less than 50) in a way distorting the classifier performance and 3) splitting the dataset into two equal parts (50% of the data each) for training and classification.

The sampling design for the HR layers is the same as the one mentioned in the previous section. A selection of samples from zonal statistics analysis is followed by a visual check of the sample relevance and then a random selection over the fishnet is applied. The sampling made in phase 1 was revised and reassessed in the second phase before being re-integrated in the sampling dataset for the next year.

The only divergence can be found in the thematic accuracy. Indeed, to gain processing time, the HRL samples are created for one year and then re-used for the next year while being also completed with new ones. This still implies a visual checking, to also ensure that the samples used are still valid. The prototype product for the cropland type is based on 2016 (then on the 2017) data while using HRL samples based on 2015 data.

For the prototype demonstration of the phase 2 over the Belgian part of the demo-site West, all the Sentinel-2 and Landsat-8 were downloaded and preprocessed using the MAJA processor, which is the combination of MACCS and ATCOR as compared in the WP32 report. The Sentinel-1 time series were preprocessed using the ESA SNAP toolbox as implemented in the Sen4CAP system.

5.1.2.2 Demo-site Central

The selected EO data (see chapter 5.1.1) were generally pre-processed as described in the Deliverables AD05 (pp. 18) and AD06.

Sentinel-1 data are pre-processed in five main processing steps: The Sentinel-1 Ground Range Detected (GRD) data (VV and VH polarisation) were pre-processed to Gamma0 values and a multi-temporal speckle filter was applied on the time series. Furthermore, the data are radiometrically calibrated and a terrain flattening as well as a terrain correction has been performed. The pre-processing was partially done using the ESA SNAP toolbox, partially by using python scripting. Only data of the descending orbit 66 was used for the reference year 2017 whereas for the year 2018 both ascending and descending orbits are incorporated in the analysis (15 and 168) in order to increase the data density of the time series for the selected time window of Mid-March to Mid-October and capture relevant changes in phenology, vitality of the plants as well as changes induced by agricultural management.
The major processing steps for all S-2 images are the atmospheric correction, cloud, cloud shadow and snow masking, as well as topographic normalisation and geometric accuracy tests. The software used for all these pre-processing steps is the Sen2Cor provided by ESA.

5.1.2.3 Demo-site Mali

S2 and L8 satellite data were atmospherically-corrected using the Multisensor Atmospheric Correction and Cloud-Screening (MACCS) algorithm (Hagolle et al., 2008; Hagolle et al., 2010), included in the Sen2Agri system. That method associates to each pixel one of the following status: “Land”, “Cloud”, “Cloud shadow”, “Water” or “Snow”. The consecutive classification methods are applied only over “Land” and “Water”.

These so called missing values were determined by value masks provided by the MACCS preprocessing step, detecting pixels affected by clouds, cloud shadows or saturation effects. The gap-filling method is a weighted linear interpolation on a 10-day time step for S2. These interpolated reflectance are used to extract the features and compute the indices.

For the prototype demonstration of the phase 2, the version 2.0 of the Sen2Agri system was implemented and used for all Sentinel-2 and Landsat-8 data preprocessing.

5.1.2.4 Demo-site Western Cape Province (South-Africa)

For the prototype demonstration of the phase 2, the version 2.0 of the Sen2Agri system was implemented and used for all Sentinel-2 and Landsat-8 data preprocessing. A set of 20 tiles of Sentinel-2 were preprocessed for the whole winter grain growing season as illustrated in the Figure 5-8.

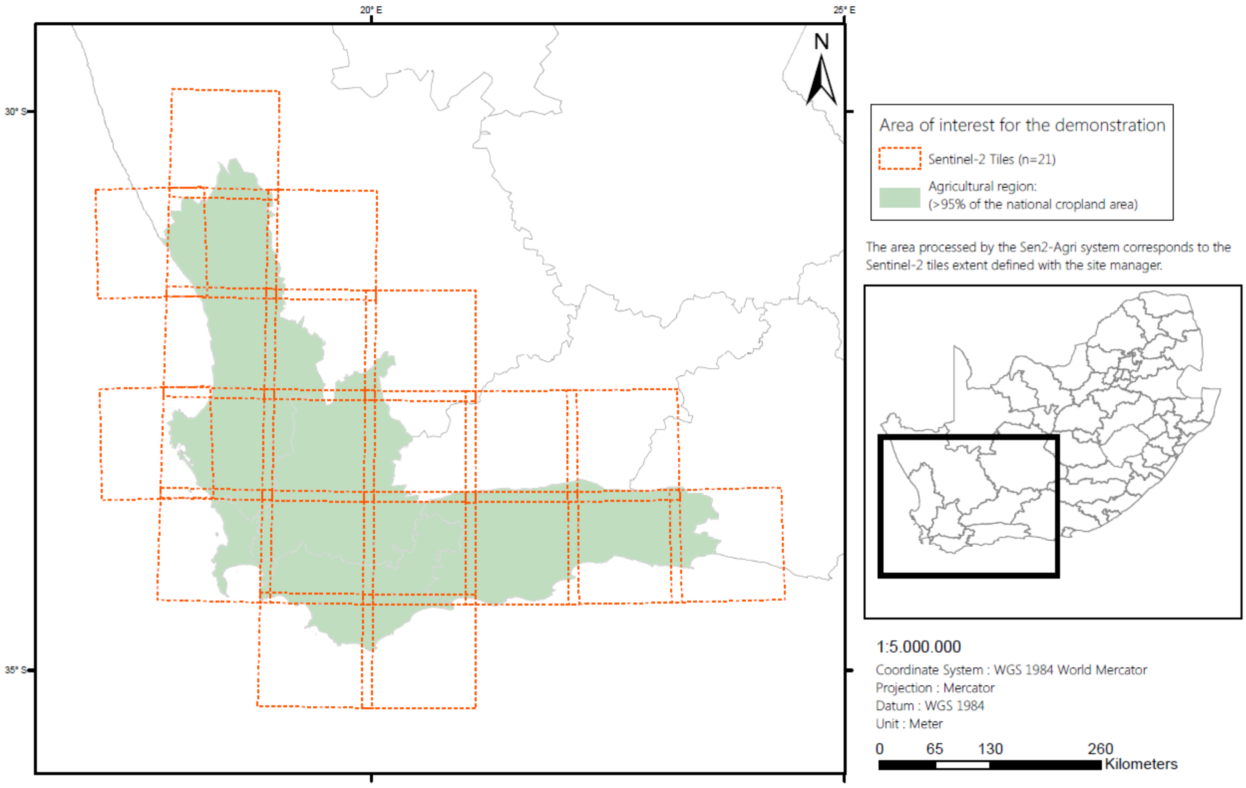


Figure 5-8 : Set of 20 Sentinel-2 tiles processed for the prototype demonstration of phase 2 in South-AfricaSet of 20 Sentinel-2 tiles processed for the prototype demonstration of phase 2 in South-Africa

5.1.3 Experimental Setup

In the following subsections the experimental setup will be described following the three different experiments carried out in the demo-site West (Belgium/France), Central (Germany/Austria), and Mali.

5.1.3.1 Demo-site West

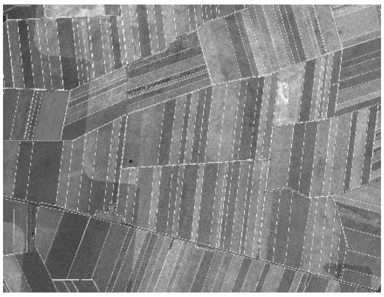
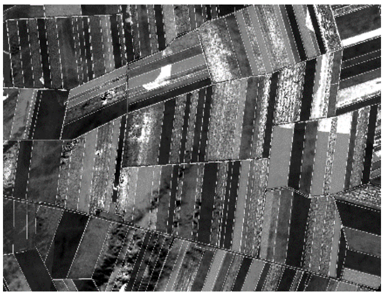
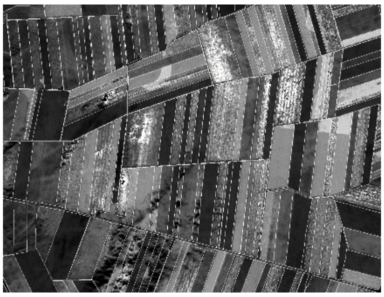
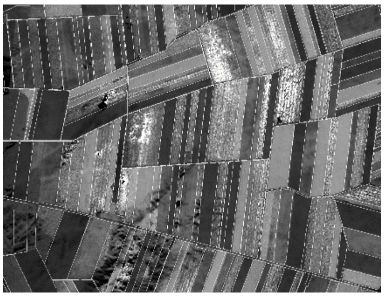
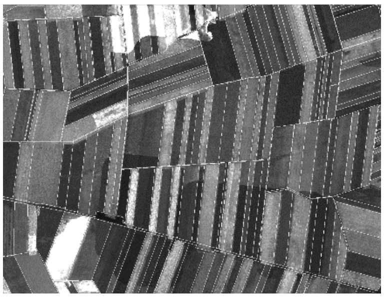
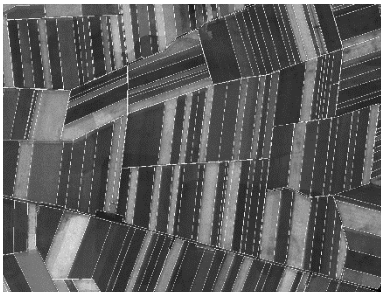
The following sections describe the time features methodology applied in the production of the cropland type prototype. The preliminary set of implemented features will be explained in the following sections, along with a description of the consecutive classification workflow implementing those time features.

PRELIMINARY SET OF IMPLEMENTED FEATURES

For the French West tiles, and the first phase, several indices have been used for the classification of the different cropland types: NDVI (Normalized Difference Vegetation Index), NDWI (Normalized difference water index) and BRI (brightness). Moreover, for each index, 9-time features have been generated: maximum, mean, minimum, percentiles (10th, 25th, 50th, 75th, 90th) and standard deviation. As the yearly time window was divided on three three-month periods, 81 indices have been therefore computed for each S-2 tile, and then stacked. Those 81-band images have then been fed as input data to the classifier.

The choice of the indices is employed to discriminate different cropland types over all the S-2 tiles, which exhibit at least one cloudless value per pixel for each the three-month period. Those three trimesters three periods are therefore used to build a spectral signature for each cropland type.
The Table 5-6 represents some examples of time features (mean, maximum and minimum) for the NDVI on the three-time periods on the working units 31UEQ.

Table 5-6: Temporal features associated with the NDVI extracted from the tile 31UEQ. The delineation represents LPIS dataset parcels.

		
(a) Period 1: maximal NDVI	(b) Period 1: mean NDVI	(c) Period 1: minimal NDVI
		
(d) Period 2: maximal NDVI	(e) Period 2: mean NDVI	(f) Period 2: minimal NDVI
		
(g) Period 3: maximal NDVI	(h) Period 3: mean NDVI	(i) Period 3: minimal NDVI

The second period, from April to June, exhibits more phenological variability than the first period from January to March. Those darker or brighter fields are a sign of a more complex and disparate range of value linked to the chlorophyll activity of the present vegetation. Due to the restrained number of available images, there is still some secondary effects linked to the cloud mask application, as can be seen, for example, on the image (c). The footprint of the cloud mask is noticeable. On the third period, the variability of the NDVI is less significant, but it still brings information on the phenology of the different crop types.

For the second phase, the vegetation parameter IRECI (Inverted Red-Edge Chlorophyll Index) has been added to the list of spectral features used over the input data, made from interpolated images of the best scenes selected over the year 2017 from S-2A and B.
One the one hand, temporal features are computed over those interpolated images (all S-2 bands stacked and resampled at 10m), while on the other hand, NDVI, NDWI, IRECI and BRI indices are also computed and stacked as input data to the classifier.

For UCL, the input features were produced from datasets of Sentinel-1 and Sentinel-2 satellites. In particular, 374 features were originated from Sentinel-2 datasets where a stuck of eleven layers (8 bands: “b3:8 b11:12”, NDVI, NDWI and BRI) was produced every 10 days over the acquisition period (January 3rd-December 2nd, 2017). In addition, 288 features were originated from the Sentinel-1 data where the mean, standard deviation, coefficient of variation and five quantiles (10th, 25th, 50th, 75th and 90th quantiles) were produced from the co-polarized (VV), the cross-polarized (VH) and polarization ratio (VV/VH) of the backscattering at a constant time interval (every two months) of the time series.

Classification workflow (Demonstration Site West)

For both phases, the main processing steps for the generation of this prototype product rely on the classification computation. The processing chain is composed of four successive steps:

1. Computing of the image statistics;
2. Generation of class features;
3. Classification with random forest algorithm;
4. Application of a majority filter to harmonize the results;
5. Calculation of an automatic confusion matrix.

During phase 1, the 81-band image is fed to the algorithm which computes various spatial statistics, such as mean and standard deviation for each band – and during phase 2, interpolated images, as well as their temporal statistics and the vegetation indices selected above were fed to the classifier. Those statistics are used to establish a classification model, according to each cropland and non-cropland class. The classes’ signatures are computed from the selected samples, whose 50% are used for calibration while the other 50% reserved for validation. The random forest algorithm, chosen after the tests conducted in WP33, can finally be applied to classify the different classes.
A majority filter is also applied to the classification. This final step presents two advantages:

- to harmonize the results, since it smooths the pixel-based classification and lets a more naturel landscape appear;
- to merge isolated pixels into larger parcels.

Once the classification generated, a first confusion matrix is automatically generated to estimate the accuracy of the results for each tile (but only the overall confusion matrix is presented in this final report for both phase). It is worth mentioning that for the phase 1, UCL applied the same classification workflow (steps 1:5 in addition to the majority filter) using 374 and 288 features/tile originated from Sentinel-2 and Sentinel-1 datasets respectively. The classification was performed in two ways where the first trained and applied the RF classifier for each tile individually while the second trained the classifier on the 31UFR tile and then applied on other tiles. The 31UFR was selected in the second case based on accuracy assessment of the first case. In addition, the classification results were constrained to the LPIS polygons within the Belgian borders, which means that all non-agricultural land covers produced by the classifier were grouped in one class named as “other land cover”. Finally, 24 classes were selected for representing the classification results in a way grouping similar classes (e.g. the summer cereals class includes mixed protein crops, summer Triticale, Quinoa, summer Rye and Sorghum).

The developed processing chain is able to treat a large amount of input data within a reasonable temporal window to provide two cropland-related layers, without manual enhancement. The achieved level of automation ensures the effective application of the process to map different cropland types at pan-European level.

The workflow for the production of the crop types prototypes over the French West tiles is listed hereafter:

For phase 1:

1. Layer stacking of the bands from the data pre-processing by Joanneum Research
2. Application of the cloud mask
3. Regrouping of data per time-windows to balance cloud presence and phenological variability
4. Computation of spectral indices, such as NDVI, NDWI and BRI
5. Computation of time features, such as maximum, mean, minimum, median and percentiles (10th, 25th, 75th, 90th) and standard deviation
6. Concatenation of all the previously generated data

For phase 2:

1. Pre-processing using MAJA
2. Interpolation of the time series of images at fixed dates
3. Resample of each 20m band at 10m
4. Stacking of all the spectral bands for each "artificial" image
5. Computation of spectral indices on those stacks
6. Computation of temporal statistics on those stacks

For both phases:

7. Classification – which could be divided into the following subtasks:
    a. Computation of bands spatial statistics
    b. Generation of class features with 50% of the samples used for calibration and the remaining 50% kept for validation
    c. Classification with RF algorithm
    d. Application of a majority filter to smooth the results
    e. Computation of an automatic confusion matrix for validation, using the other half of the samples

For phase 1 only:

8. Aggregation of the results at field parcel level according to the LPIS shapefile
9. Mosaicking of the results obtained over each S-2 tile

Finally, for both phases:

10. Merge of isolated pixels into larger ensembles (especially in the corridors between two parcels of the LPIS)
11. Formatting of the metadata

5.1.3.2 Demo-site Central

Phase 1:

For the production of the phase 1 prototype, the used S-1 and S-2 images were grouped into 10 different scene sets based on different time steps, five for S-2 and five for S-1 (see Table 5-7). The results of the tests regarding the respective start and end date further described in WP33 (AD06) were taken into account here. There are 5 scene sets per sensor on which the feature extraction (see section 4.2) was performed: one of them covering the whole growing season from mid-March until mid-November and four others each covering a two-month period starting from mid-March until mid-September.

Table 5-7: Phase 1 - Start and end date of the 10 different scene sets and the number of S-1/S-2 images they contain including the maximum amount of cloud cover (%).

Sensor	Sceneset Name	Start Date	End Date	Number of Scenes per Tile	Max. Cloud Cover (%)
S-1	EL_S1_2017-03-15TO2017-05-14_CC00	2017-03-15	2017-05-14	9-30	0
	EL_S1_2017-05-15TO2017-07-14_CC00	2017-05-15	2017-07-14	8-26	0
	EL_S1_2017-07-15TO2017-09-14_CC00	2017-07-15	2017-09-14	10-31	0
	EL_S1_2017-09-15TO2017-11-14_CC00	2017-09-15	2017-11-14	9-29	0
	EL_S1_2017-03-15TO2017-11-14_CC00	2017-03-15	2017-11-14	37-115	0
S-2	EL_S2_2017-03-15TO2017-05-14_CC90	2017-03-15	2017-05-14	5-7	90
	EL_S2_2017-05-15TO2017-07-14_CC90	2017-05-15	2017-07-14	10-12	90
	EL_S2_2017-07-15TO2017-09-14_CC90	2017-07-15	2017-09-14	13-15	90
	EL_S2_2017-09-15TO2017-11-14_CC90	2017-09-15	2017-11-14	10-16	90
	EL_S2_2017-03-15TO2017-11-14_CC90	2017-03-15	2017-11-14	42-49	90

For the crop mask classification, reference samples covering the basic LC types (Forest, Cropland, Grassland, Urban Areas, and Water bodies) were available from the HRL 2015 layer production. LPIS data have not been used for reference sampling, as these are not available in all European countries.

Splitting the dataset into a training- and test set was performed for the crop mask classification, meaning that approx. 50% (~400–1300 polygons for each class) were used for training and the remaining 50% for testing. This led to a total of ~7300 sample polygons for all classes amounting to more than 1 million pixels. In order to reduce the amount of samples for training, 50,000 random pixels were extracted out of the full train set. For the urban areas the amount of pixels in the full train set was <50,000. Therefore, all available pixels of the urban areas (~31,000) were used (see Table 5-8).

Table 5-8: Phase 1 - Number of polygons for each LC type in the reference data for the crop mask: All, Train set (for model training), and Test set (for validation).

LC Type	All (polygons)	Trainset (pixels)	Testset (polygons)
Forest	2672	50000	1336
Cropland	1129	50000	565
Grassland	1538	50000	769
Urban areas	1142	31300	571
Waterbodies	846	50000	423
SUM	**7327**	**231300**	**3664**

Regarding the sampling dataset for the crop type classification the complete LPIS reference data was randomly split in training (30%) and test (70%) polygons. For the model training, in order to derive a balanced and reduced training/calibration data set, a random sample of 10,000 pixels was drawn for each crop group (the train set).

The following Figure 5-9 and the Table 5-9 shows the referring classes and the number of LPIS polygons they contain (number of total LPIS polygons and their respective size (km²), as well as a number of 70% of polygons per class: the test set). Winter Crop has by far the most polygons in the demonstration site Central whereas Summer Rape or Sunflowers/Topinambour cover the least area. Both large in polygons as well as field size are Maize, Summer Crop and Agrarian Grassland, whereas fruit trees and vegetables cover small areas in relation to their amount of polygons. Regarding the mean parcel size Sugar Beets and Winter Rape have remarkable big parcel sizes compared to their amount of reference parcels. Winter Crop and Maize are again both large in number of reference parcels and mean parcel size. The smallest mean parcel size of all classes has winegrowing.

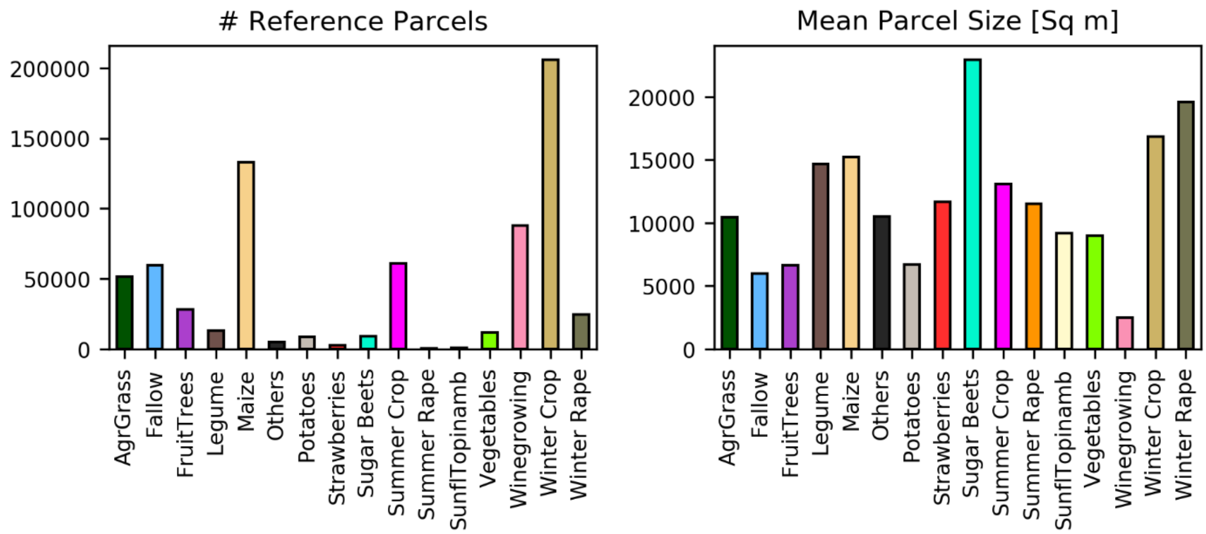


Figure 5-9: Phase 1 - Number of Reference Parcels and the Mean Parcel Size per class.

Table 5-9: Phase -1 Number of polygons for each crop group in the reference data (All), their area, and the number of polygons used for validation (Test set).

Crop Group	All (polygons)	Area of All (km²)	Trainset (Pixels)	Testset (polygons)
Winter Crop	205861	3471,1	10000	144103
Maize	132789	2018,83	10000	92952
Winegrowing	87831	219,68	10000	61482
Summer Crop	61150	801,66	10000	42805
Fallow	59729	355,97	10000	41810
Agrarian Grass	51493	538,86	10000	36045
Fruit Trees	24932	169,59	10000	17452
Winter Rape	24484	479,42	10000	17139
Legume	13217	193,85	10000	9252
Vegetables	9701	79,23	10000	6791
Sugar Beets	9053	207,67	10000	6337
Potatoes	8708	58,33	10000	6096
Strawberries	2715	31,63	10000	1901
Others	1844	22,89	10000	1291
Sunflowers/Topinambour	767	7,03	10000	537
Summer Rape	256	2,94	10000	179
SUM	**694530**	**8658,68**	**160000**	**486172**

Another important step is the calculation of spatially contiguous features from the pre-processed and stacked input data. A feature is a statistical metric which is derived for each band/index and pixel from a set of valid (i.e. cloud/cloud shadow free) observations over a specific time range, e.g. the median of all valid NDVI values of the defined period from 2017-03-17 to 2017-05-14. Table 5-10 lists the bands/indices and the derived time features from them. For more information on the features please refer to AD06. In phase 1 of ECoLaSS, these features were calculated for all periods listed in Table 5-7 which led to a total of 1246 features. But as mentioned in section 4.2 it is more efficient to reduce the number of features for the following reasons: (i) to keep the computational effort as low as possible, and (ii) to avoid redundant information and reduce noise. For the selection of the most relevant features the FFS was applied. For the feature selection the 10,000 training samples were split in a 50% training and 50% validation set for classifying and evaluating the feature performance with the Random Forest (RF). The final model with the selected features was then trained with the full training samples.

Table 5-10: Phase 1 - Overview over the relevant bands/indices and the derived time features used for training and classification.

Sensor	Bands/Indices	Time Features Groups
Sentinel-1a/b	• VV (Gamma0) • VH (Gamma0) • Norm. Difference VV/VH (NDVVVH) • Ratio VV/VH (RATIOVVVH)	activity (only phase 1), cov, dif, max, maxmean, mean, median, min, percentiles (p010, p025, p050, p075, p090), pdiff 075025, pdiff 090010, std, trend
Sentinel-2a/b	• Brightness (derived through summation of the values of the bands Green, Red, NIR and SWIR1) • IRECI (Inverted Red Edge Chlorophyll Index) • NDVI (Normalized Difference Vegetation Index) • NDWI (Normalized Difference Water Index, based on SWIR and NIR) • B03, B04, B08, B11, B12	

Both, for the crop mask and crop type classification, the RF classifier was used as classification approach for the final model training as well as for the classification approach used for the FFS and the classification of the final products. The parameters were the same during FFS, crop mask classification and the crop type classification (number of trees: 500, 'gini' criterion for measuring the split quality, number of features to consider when looking for the best split: square root of the number of features, no constrain at the depth of the trees; for further information, see AD06 (pp. 82).

The crop mask was validated with the sample dataset as described in section 5.1.1, taking all pixels of the test set into account. For validating the results of the crop type map, a buffer was applied on the selected 70% of the LPIS dataset (the reference data) where all pixels with a distance ≤2 pixels from the polygon border (which represent boarders between agricultural fields) were not considered for validation.

Phase 2:

An analysis of the growing cycles of the main crop plants in Europe induced a restructuring of the time window used for production in phase 2. Chosen time windows are Mid-March to Mid-May (spring period– sowing and sprouting of summer crops as well as growing of winter crops), Mid-May to Mid-July (summer period: growing of summer crops, vegetation peak for winter crops, planting of additional summer crops like vegetables) and Mid-July to Mid-Oct (autumn period: vegetation peak for further crops, period of plants' maturing, withering, harvesting, first tilling events). In order to grasp meaningful information about plant growth and farming management's impact, a 2-3 months period turned out to be minimum for creating sound and reliable time features. The additional time window covering the whole period is used to collect further information of the whole life cycle of all plants. This approach led to the following table of time feature sets:

Table 5-11: Phase 2 - Start and end date of the 8 different scene sets and the number of S-1/S-2 images they contain including the maximum amount of cloud cover (%).

Sensor	Sceneset Name	Start Date	End date	Number of scenes (demosite)	Max. Cloud Cover (%)	Min. tile Coverage
S-1	EL_S1_2018-03-15TO2018-05-14_CC00_TC02	15.03.2018	14.05.2018	163	n.a.	20%
S-1	EL_S1_2018-05-15TO2018-07-14_CC00_TC02	15.05.2018	14.07.2018	178	n.a.	20%
S-1	EL_S1_2018-07-15TO2018-10-14_CC00_TC02	15.07.2018	14.10.2018	276	n.a.	20%
S-1	EL_S1_2018-03-15TO2018-10-14_CC00_TC02	15.03.2018	14.10.2018	630	n.a.	20%
S-2	EL_S2_2018-03-15TO2018-05-14_CC90_TC08	15.03.2018	14.05.2018	106	90%	80%
S-2	EL_S2_2018-05-15TO2018-07-14_CC90_TC08	15.05.2018	14.07.2018	122	90%	80%
S-2	EL_S2_2018-07-15TO2018-10-14_CC90_TC08	15.07.2018	14.10.2018	222	90%	80%
S-2	EL_S2_2018-03-15TO2018-10-14_CC90_TC08	15.03.2018	14.10.2018	451	90%	80%

Since the accuracy of the crop mask is decisive for the quality of the crop type mask, phase 2 strongly focused on methods for enhancing the crop/non-crop differentiation. Due to the heterogeneous biogeographic conditions on the demo site, agricultural farming routines differs regionally. Sowing and harvesting dates are shifted in areas of higher altitudes (such as the Southern tiles in the Alpine region) and in regions with mild and temperate conditions (such as the tiles upper left of the demo site). The results of the test site classification in phase 2 led to the assumption that a stratification approach would support the differentiation between crop types in general, and between winter and summer crops in particular. Four different strata have been identified for the classification of the crop mask in demo site Central:

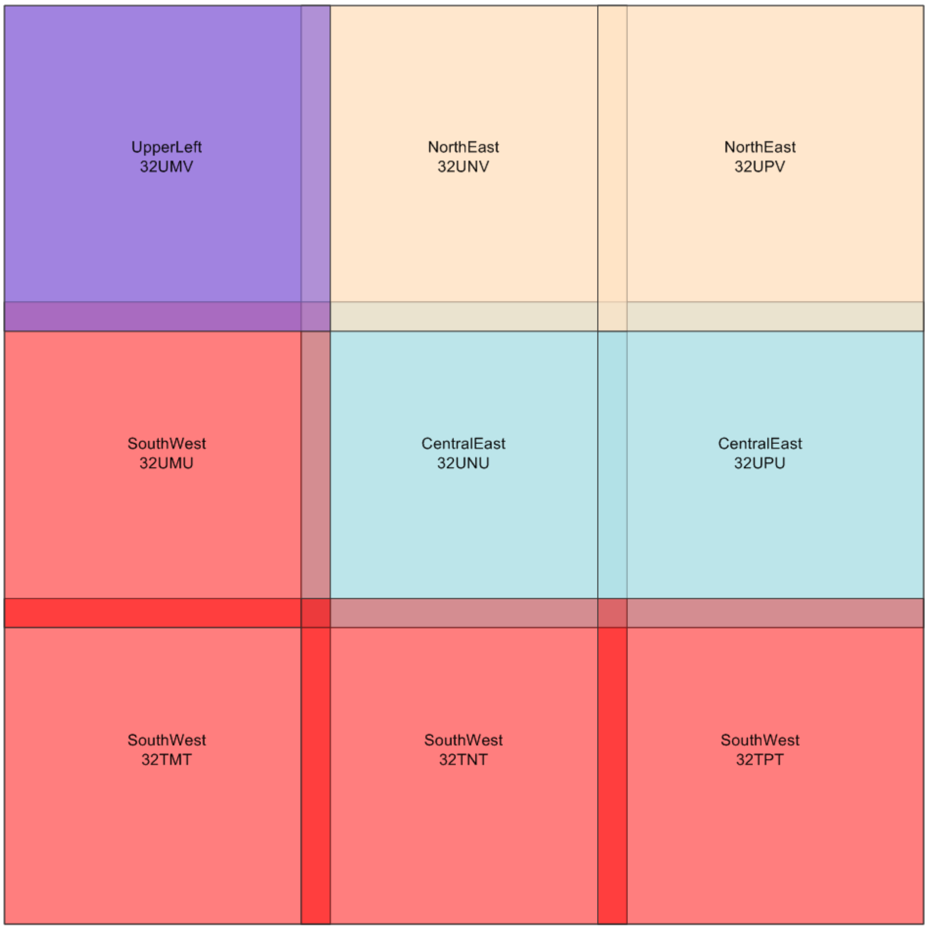


Figure 5-10: Phase 2 - stratification of the demo site Central for the drop mask classification; 4 strata taking into account different biogeographic conditions

The stratification approach takes several aspects into account: the mean temperature, the mean precipitation, the altitude and the predominant biogeographic region (following the approach of Metzger et al. for a global stratification in environmental zones; taking biophysical aspects into account such as temperature, precipitation, soil characteristics, altitude and more)

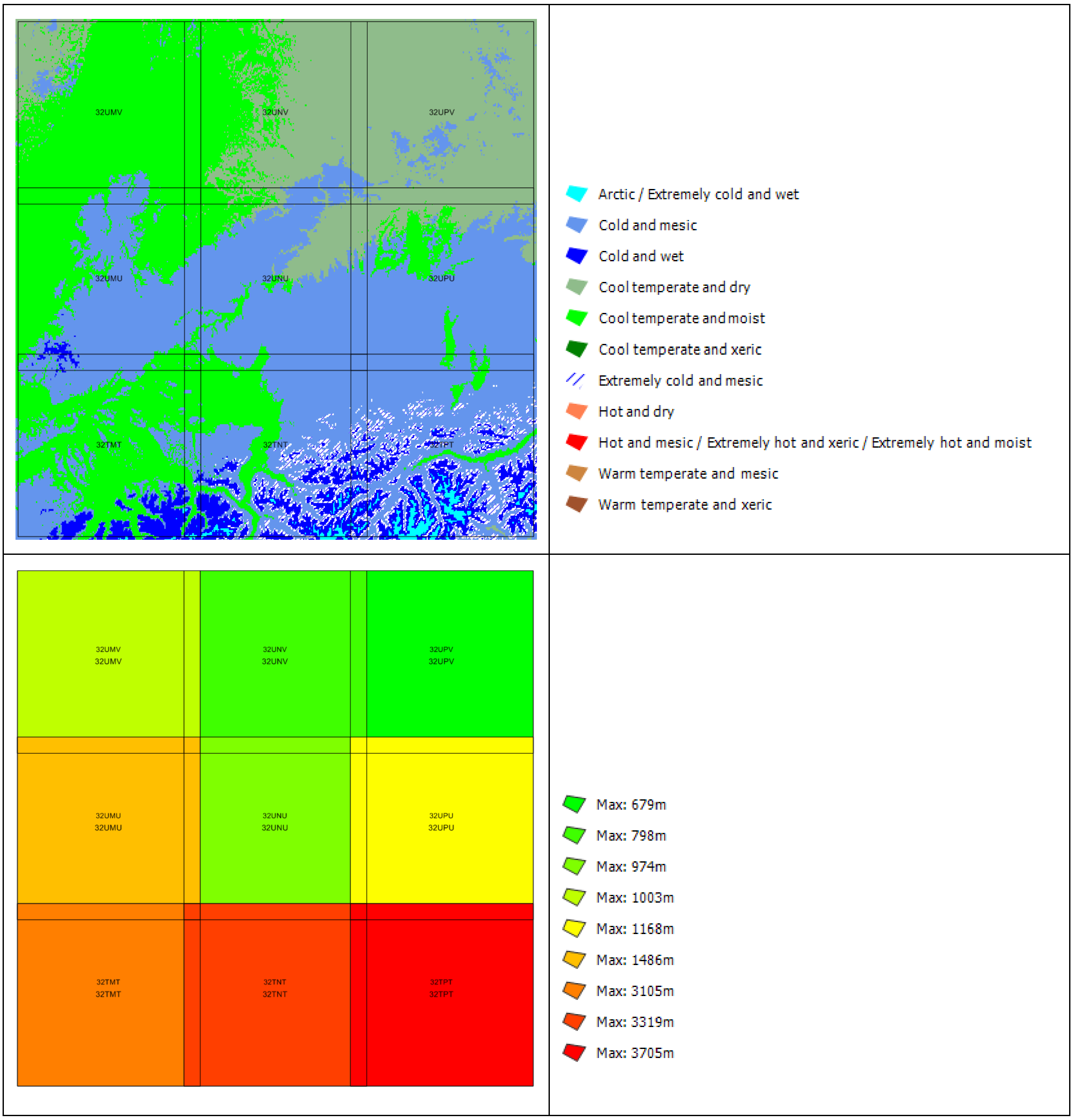


Figure 5-11: Phase 2 - details for the demo site Central (referring to the Sentinel-2 tile – division of the classification method), serving as basis for the stratification approach or the crop mask classification. Above: Environmental Zones after Metzger et al; below: maximum height per Sentinel-2 tile

The classification for the crop mask has been generated for the respective stratum with adapted sample files. As in previous tests, the train/test-split was 50% training samples, 50% test samples. The sample files cover the basic LC types (forest, cropland, grassland, urban areas, and water bodies) derived from the LUCAS2018 data, complemented by samples of HRL2015 and manual samples for orchards and vineyards. The number of pixels per class varies due to heterogeneous land cover per tile (see table below).

Table 5-12: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum Upper Left: All, Train set (for model training), and Test set (for validation).

	number of pixel		
class code	reference set	train set	test set
1 - broadleaved trees	7754	3875	3879
2 - coniferous trees	6136	3068	3068
3 - imperviousness	2553	1276	1277
4 - weater	126	63	63
5 - grassland	3037	1517	1520
6 - cropland	3204	1602	1602

Table 5-13: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum North East: All, Train set (for model training), and Test set (for validation).

	number of pixel		
class code	reference set	train set	test set
1 - broadleaved trees	11868	5928	5940
2 - coniferous trees	9582	4789	4793
3 - imperviousness	3356	1673	1683
4 - weater	154	82	72
5 - grassland	7504	3752	3752
6 - cropland	6097	3049	3048

Table 5-14: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum Central East: All, Train set (for model training), and Test set (for validation).

	number of pixel		
class code	reference set	train set	test set
1 - broadleaved trees	6936	3545	3391
2 - coniferous trees	6128	3068	3060
3 - imperviousness	3817	1935	1882
4 - weater	477	234	243
5 - grassland	7921	4114	3807
6 - cropland	5878	3060	2818

Table 5-15: Phase 2 - Number of polygons for each LC type in the reference data for the crop mask for stratum South West: All, Train set (for model training), and Test set (for validation).

	number of pixel		
class code	reference set	train set	test set
1 - broadleaved trees	15168	7579	7589
2 - coniferous trees	22574	11287	11287
3 - imperviousness	4250	2133	2117
4 - weater	2419	1215	1204
5 - grassland	18778	9398	9380
6 - cropland	10539	5261	5278

As for the crop type mask, LPIS data from 2018 considering Baden-Wurttemberg, Bavaria and Austria have been used and adapted to the reworked crop type nomenclature for phase 2. Testing showed, that a sufficient number of samples is essential for a distinct differentiation of crop types within the classification. In order to avoid low numbers of samples for specific crop types, the classification for the demo site has been performed without stratification. However, stratification could well be an approach in a larger context, such as one larger biogeographic region that covers several Sentionel-2 tiles. Precondition will be a sufficient number of samples per crop type.

Concerning the test and train split, the crop type samples have been split by 70% for the model training and 30% for the classification because this percentage lead to best results. Indeed, model building for those many crop classes needs more information input and subsequent training than the model building for a binary crop mask. These were reduced to those crop type polygons which correspond to the crop type nomenclature and filtered by a minimum parcel size of >1ha. The samples file showed the following statistics:

Table 5-16: Phase 2 - Number of parcels/sample polygons per crop type for the demo site Central

class code	CropType	number of parcels
1	winter wheat	124389
2	winter barley	66189
3	winter rye	7562
5	summer wheat	1736
6	summer barley	30512
8	summer oats	9708
9	maize	149882
10	vegetables	3894
11	peas+beans	5566
12	lentils	439
14	soya beans	3650
15	sunflowers	360
16	rape seed	27119
17	oleaginous + fibre crops	6477
18	potatoes	9301
19	beet crops	14329
20	temporary grassland + fodder crops	31850
21	wine growing	8147
22	fruit trees + berries	6048

Aiming at a balanced and reduced training/calibration data set, a random sample of 10,000 pixels per polygon was drawn for each crop group (the training set) out of the sample polygons.

Table 5-17: Phase 2 - pixel support for crop type classification for demo site Central; classes 4, 7 and 13 left out due to insufficient sample number

class	PA	UA	F1	support
winter wheat (1)	0,91	0,83	0,87	63.859
winter barley (2)	0,82	0,81	0,82	57.674
winter rye (3)	0,83	0,81	0,82	59.130
summer wheat (5)	0,56	0,63	0,59	49.688
summer barley (6)	0,74	0,72	0,73	57.941
summer oats (8)	0,58	0,66	0,61	45.206
maize (9)	0,90	0,83	0,86	56.774
vegetables (10)	0,77	0,67	0,71	43.076
peas+beans (11)	0,78	0,80	0,79	55.834
lentils (12)	0,75	0,69	0,72	23.190
soya beans (14)	0,82	0,81	0,81	60.630
sunflowers (15)	0,71	0,90	0,79	41.819
rape seed (16)	0,99	0,98	0,98	65.016
oleaginous+fibre crops (17)	0,94	0,92	0,93	53.491
potatoes (18)	0,81	0,82	0,82	56.124
beet crops (19)	0,92	0,94	0,93	74.155
fodder+temporary grass (20)	0,78	0,69	0,73	48.042
wine growing (21)	0,60	0,83	0,69	37.672
fruit trees+berries (22)	0,88	0,84	0,86	43.827

Several aspects concerning sample characteristics have impact on the classification results: the number of samples polygons/parcels, the average size of the parcels and the number of large parcels. Table 5-17 gives an overview of the percentage of number of parcels versus parcels size. As already discussed in WP33, specific spectral characteristics can compensate potential limitations on parcel number and/or size, as it can be seen in the example of sunflower: Even though sunflowers are grown on a limited area, the average parcel size is large enough to be well detected with 10m resolution imagery. Additionally, sunflowers offer a very distinct texture and unique spectral characteristics to be well detectable and to end up with high accuracies.

All crop types showing small parcel sizes risk to be less detectable. Winter crops tend to be grown on large parcels and on large areas as well and show high accuracies – except the winter oats, giving small parcels as well as little number of parcels. According this reasoning, lentils and vineyards should also have been left out respectively risk to be less reliably detected. Indeed, the accuracies of these crop types are lower

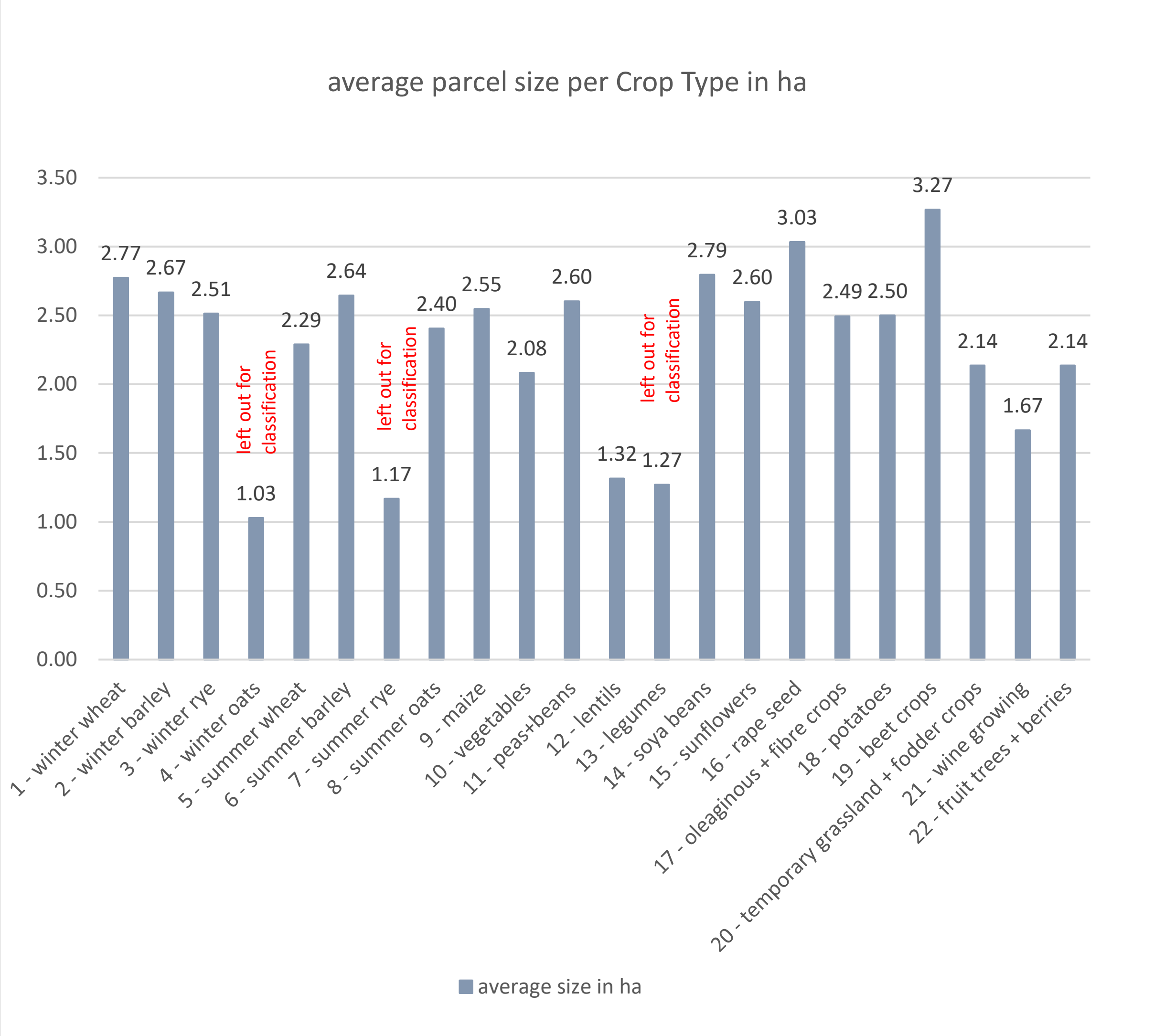


Figure 5-12: Phase 2 - Average parcel size per crop type. Classes 4, 7 and 13 have been left out due to small numbers of parcels as well as small parcel size.

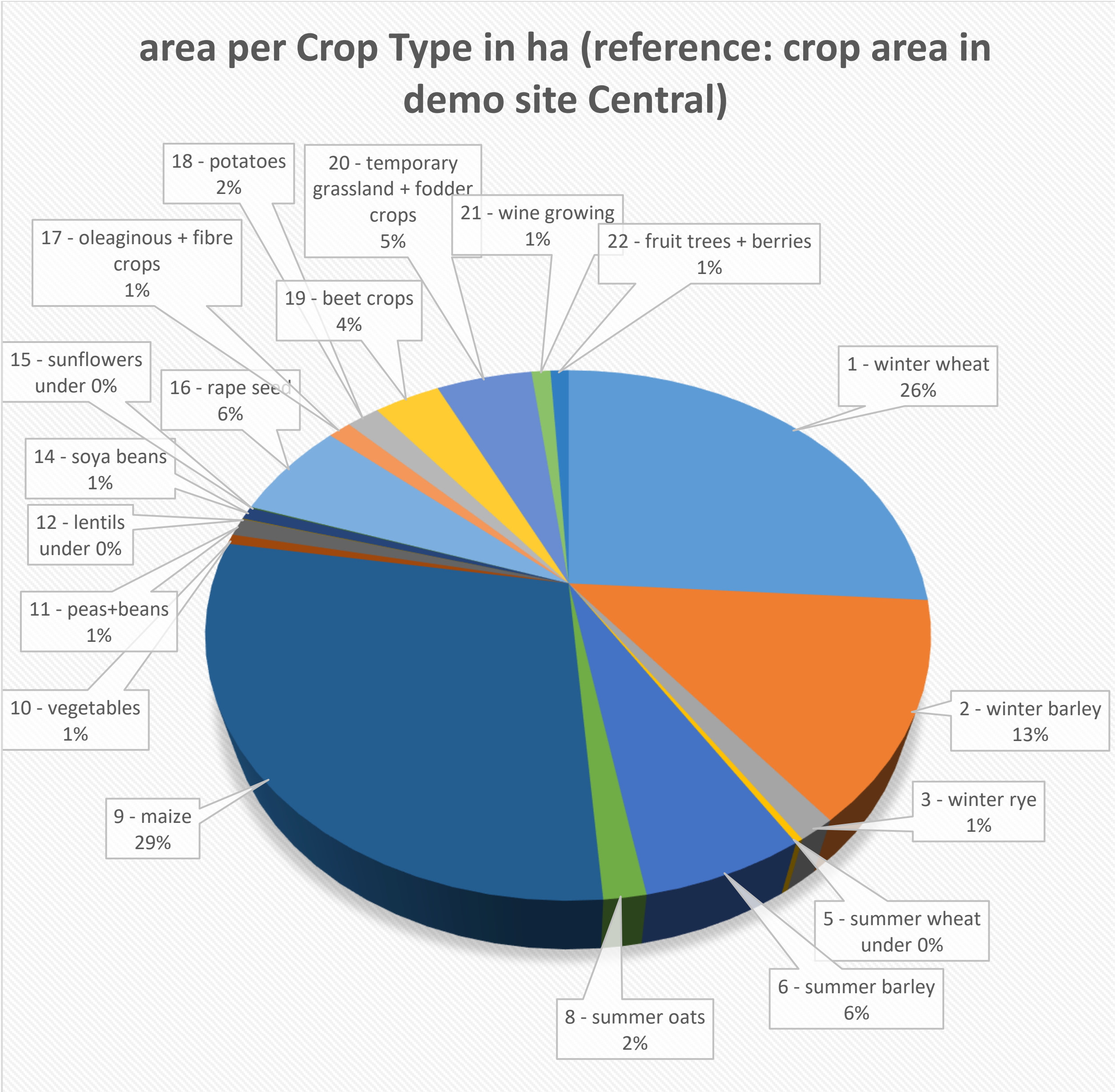


Figure 5-13: Area per crop type in ha

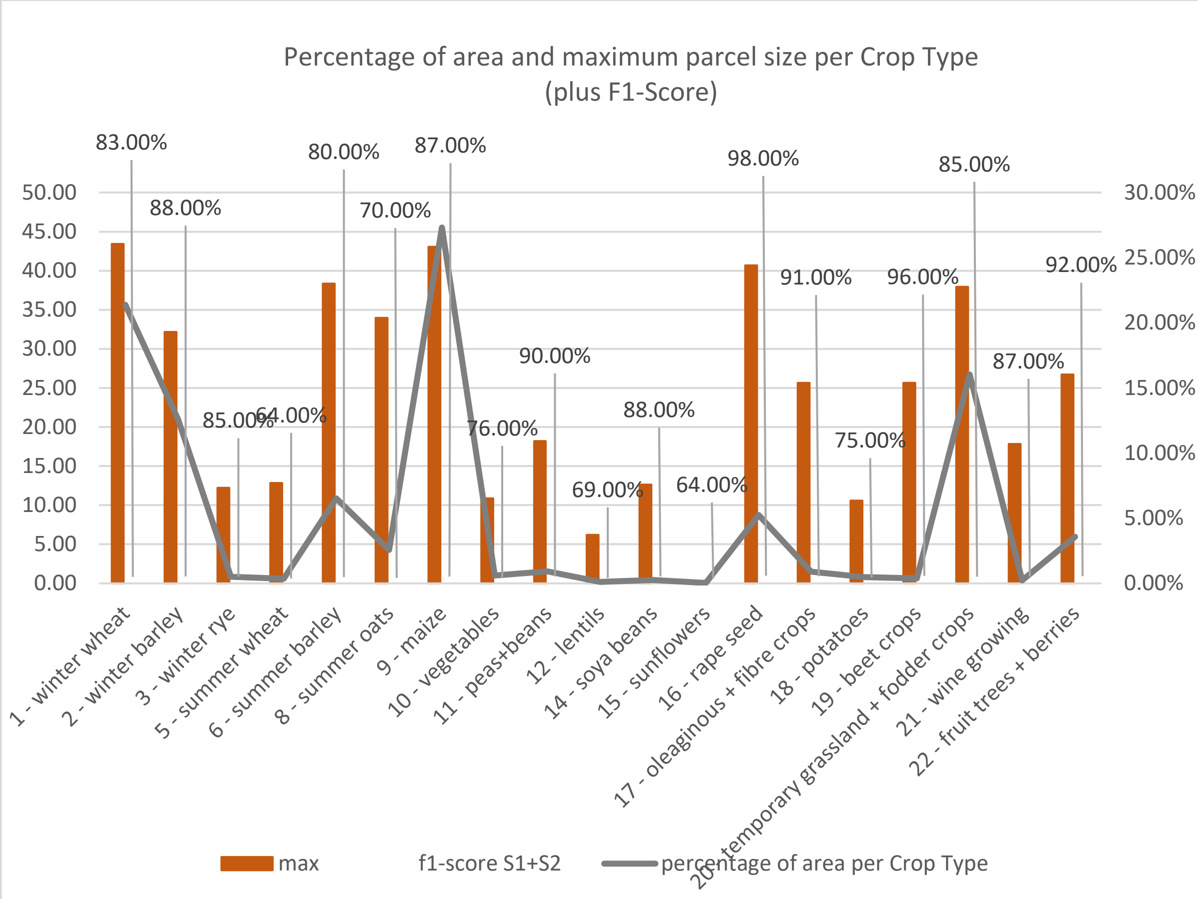


Figure 5-14: Phase 2 - Connection between percentage of area, maximum parcel size and F1 Score (combined approach of the two sensors) per crop type for the demo site Central

<u>Time feature calculation</u>

The basic set of statistical metrics for the demo site is the same as for the test site as this set proved to guarantee spatially contiguous information for the chosen time windows. As statistical metrics for each band and index and per pixel from a dense time series of valid EO data, both S-1 and S-2, they make out most of the vegetation periods' information.

Table 5-18: Phase 2 - Time features per band and index for Sentinel-1

sensor		band + indices		features
S-1	→	VV	→	covariance (cov)
		VH		standard deviation (sdt)
		NDVVVH		mean
		RATIOVVVH		min
				max
				median
				p010
				p025
				p050
				p075
				p090
				pdiff075025
				pdiff090010

Table 5-19: Phase 2 - Time features per band and index for Sentinel-2

sensor		band + indices		features
S-2	→	band 03	→	covariance (cov)
		band 04		standard deviation (sdt)
		band 08		mean
		band 11		min
		band 12		max
		NDVI		median
		NDWI		p010
		IRECI		p025
		BRIGHTNESS		p050
				p075
				p090
				pdiff075025
				pdiff090010

The initial set of time features consists of 676 features altogether. For the sake of efficiency, the method of grouped Forward Feature Selection has been used to reduce the number of features to reduce the computational effort, redundant information and noise. After the FFS, the reduced number of 221 time features has been used as input for the classification of crop mask and crop type.

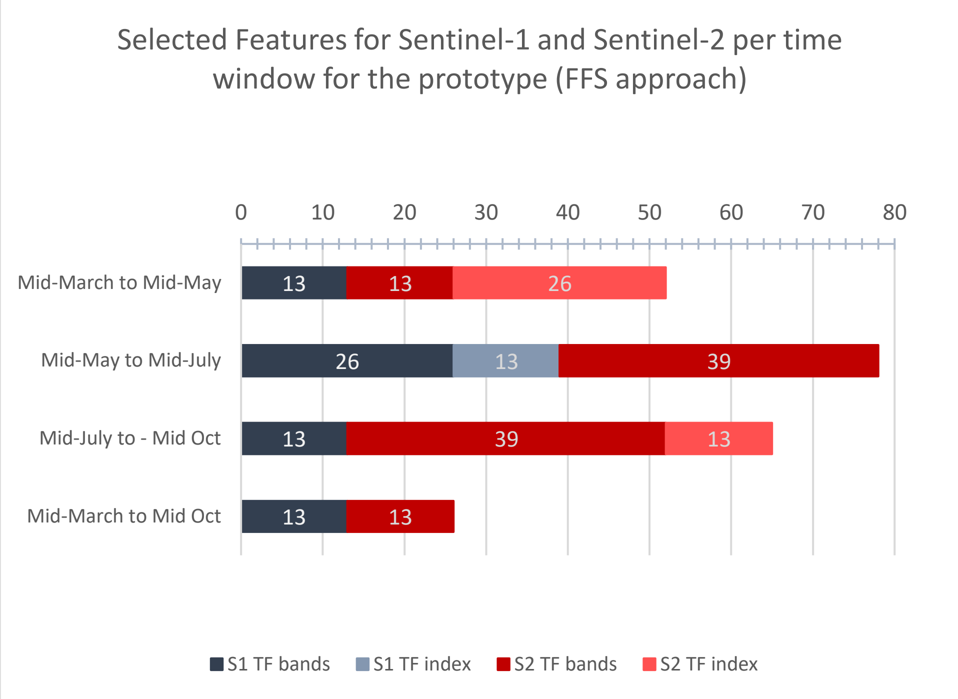


Figure 5-15: Phase 2 - Selected features for crop type classification for demo site Central after FFS

Classification:

The Random Forest classifier has proven to be the most suitable one for agricultural land cover classification. Therefore, for both, crop mask and crop type mask, this classifier has been used. The parameters were the same as for the test site, such as 500 as number of trees and maximum depth of 10 n-estimators.

Validation:

The crop mask has been validated with LUCAS 2018 points that have not been used for training nor for tests during classification. The crop type mask could only be validated by using LPIS data since LUCAS or any other reference data could not provide the specific information on the crop types. As far as possible, LPIS data that have not gone into the classification workflow have been used. As for the crop types with only a small number of samples (where all polygons/parcel were needed as input), only the testing set was used for validation. Please find the results of the validation in chapter 5.2.

5.1.3.3 Demo-site Mali

The Sen2-Agri products were processed and specifically assessed using overall accuracy and F-score figures to address four operational questions to be answered at the demo-site level: the stability of a RF classification process from one run to another and from one year to another, the impact of the independence of the validation dataset, the impact of stratification, and the accuracy evolution according to the delivery time along the season.

For the 2017 year, the Sen2-Agri system v.1.7 was run first to assess different validation setups, then according three different stratification strategy (none, 2 PIRT strata, aggregated strata of Vintrou et al. (2012), and for different periods assessing the accuracy evolution along the season to deliver cropland masks as early as possible. Indeed this cropland mask can be used to focus on the crop monitoring activities or the early warning system based on daily MODIS or PROBA-V NDVI time series. Therefore, the accuracy evolution of the cropland mask along the accumulation of Sentinel-2 observation is assessed for four different periods all starting in April to September, October, November and, December.

Splitting randomly the in situ dataset into training and validation subset is quite valid when the entire region of interest is systematically covered. While the comprehensive field data campaign insured a widely distributed sampling, only 27 villages and their surroundings where samples for 135 000 km². Therefore two different strategies have been applied.

First, an independent randomly stratified crop - non crop dataset was collected by photointerpretation in order to perform spatially independent validation. The random selection of points was stratified with GlobeLand 30 land cover to guarantee a certain quantity of samples in each land cover class. A set of 500 cropland samples have been photo-interpreted on screen by a field expert using Google maps and Bing.

Second, the impact of the spatial dependency between the calibration and validation datasets on the validation performances was assessed for cases where in situ data are not covering systematically the whole area of interest, like in most cases of in situ data collection. Two methods for the preparation of the calibration / validation datasets were compared: (i) a classical scheme where the in situ data are randomly split in two subsets, resulting in spatially correlated samples and (ii) a leave-4 out (L4O) with a spatially independent classification. Indeed, most classification approaches simply use a percentage of the ground data for training and the remaining data for the validation, i.e. 75 % for training and 25 % for validation in the Sen2-Agri system. Such a method results in some spatial dependency of the validation dataset versus the calibration one as both sets of polygons come from all the 28 sampled villages (Figure 5-5). Next to that, a more spatially independent strategy was designed as a Leave-4-Out (L40) classification was run by leaving 4 villages out and then using them to validate the map. As 28 villages were sampled for field data collection, 7 classifications were run to leave each village once out. The number of repeat is limited due to high computing time.

Both Sen2-Agri versions v.1.7 and v.2.0 relies on a supervised Random Forest (RF) classifier built on a set of features was computed from a temporally resampled and gap-filled reflectances time series. The pre-processing module however varies from MACCS to MAJA when moving from the v.1.7 used for the 2017 maps to the v.2.0 used for the 2018 maps production. Furthermore, the 2018 products have been extended to a larger extent to include Bamako.

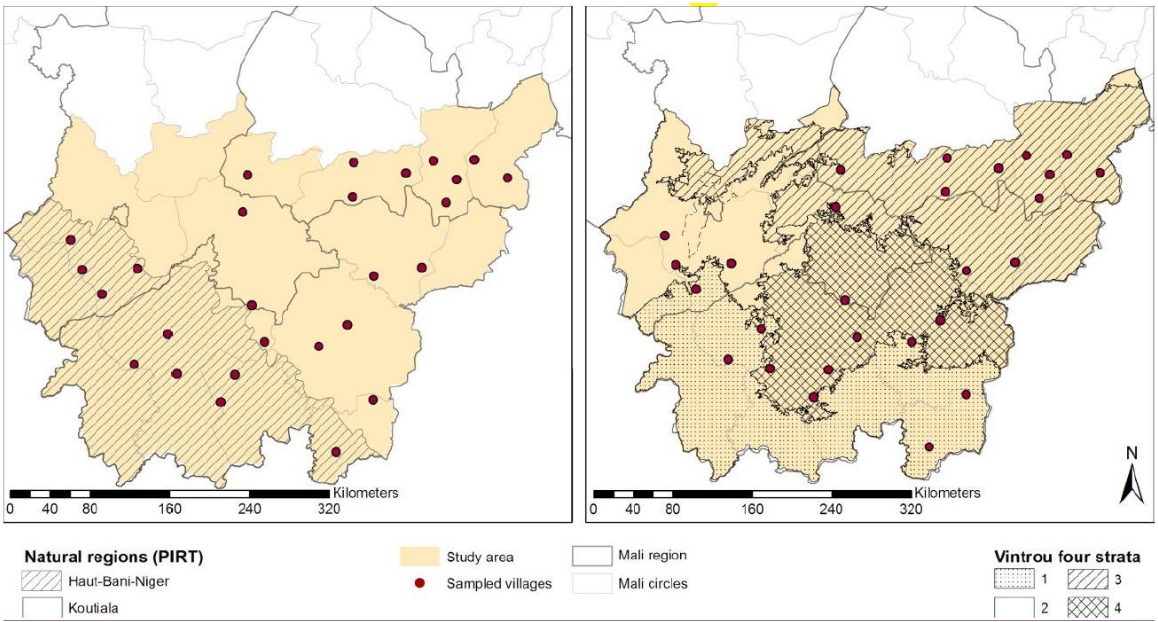


Figure 5-16: **Two stratification strategies: aggregated PIRT natural zones on the left and aggregated Vintrou strata on the right based on cropping systems analysis.**

As described here above and in Defourny et al. (2019), the cropland classification processor of the Sen2-Agri system relies on a RF model built on temporal features derived from the NDVI time series and on statistical features extracted from Normalized Difference Water Index (NDWI), brightness and S2 red-edge indices (Red edge NDVI, S2 Red Edge Position, Plant Senescence Reflectance Index, Chlorophyll Red−Edge). The Sen2-Agri crop type classification processor relies also on a RF model but built on the gap-filled Sentinel-2 reflectance time series and the corresponding NDVI, NDWI, and brightness.

5.1.3.4 Demo-site South Africa

The cropland and the crop type maps for Western Cape Province were produced using the Sen2-Agri v.2.0 system in order to assess the accuracy of agricultural products covering entire administrative unit. No stratification was considered for this large study area. Unlike for Mali, 5 % of the available in situ data were used for calibration and the remaining 95 % served for validation. After an inner buffer of 15 m and the removal of parcels smaller than 0,5ha, 184.046 parcels out of 262.985 were considered as input data for the demonstration. The calibration dataset is made of 5 % of this input dataset plus 2,293 non crop polygons (built up, water, forest, …) delineated based on on-line available VHR imagery (n calib.= 11,490).

The initial crop typology included 200 classes (including non crop classes) which were regrouped in 9 crop type classes.

5.1.4 Post-processing

DEMO-SITE WEST (FRENCH PART)

The look and feel gives a first qualitative impression on the prototype product, based on its proximity with the LPIS dataset. The confusion matrix automatically generated for the filter classification leads to a first quantitative assessment of the results.
Using those two combined techniques, the classification has been launched on several occasions - for each iteration, the classes have been re-arranged differently (through fusion, suppression or addition) until a final nomenclature of 23 classes (19 cropland classes and 4 non-cropland classes) has been selected, whose details are shown in the Table 5-20 as well as in Table 5-21.

Table 5-20: Cropland labels, characterizing 12 types of culture.

Class ID	1	2	3	4	5	6	7	8	9	10	11	12
Category	Cropland type											
Class	Maize	Summer Cereals	Winter Cereals	Rice	Sugarbeets	Potatoes	Root crops	Vegetables	Flowers	Soybeans	Fruits Trees	Grapes Vines

Class ID	13	14	15	16	17	18	19
Category	Cropland type						
Class	Olives Groves	Tree Crops	Sunflower	Oleaginous	Rapeseed	Grassland	Fodder

Table 5-21: Non-cropland class labels.

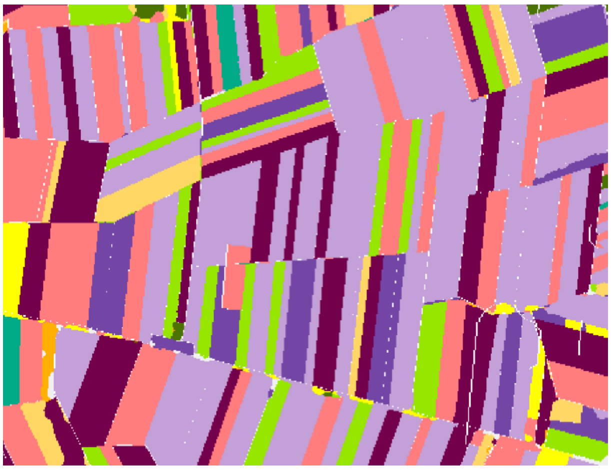
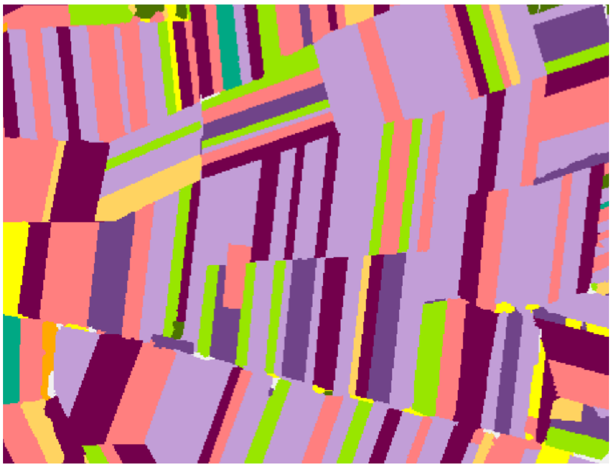
Class ID	20	21	22	23
Category	Other Land Cover (Non-Cropland)			
Class	Grassland	Urban areas	Forest	Water

There may still be room for improvement by refining the quality of the samples for example, or by integrating S-1 data, as demonstrated in the WP43.

However, the overall final prototype product reached an even more satisfying level, with this configuration of 19 classes than in phase 1.

During phase 1, the following step in the post-processing lies in the aggregation of pixels at field-level according to the LPIS shapefile. This leads to a harmonization of the processed layer for all the cropland classes, resulting from a majority occurrence re-classification of pixel at field level.
Each S-2 tile classification has then been mosaicked to obtain one unique layer over the whole demonstration site. In the overlapping area between two adjacent tiles (a corridor of 10km by 100km), a majority vote has been applied, based on the number maximal number of observations leading to a particular label. For example, a considered pixel identified as belonging to a particular class, as the result of a classifier using only 3 images will be overruled by a classifier result made from 6 optical images identifying the pixel as containing another kind of crop.
Finally, the last part of the post-processing merges isolated pixels into larger ensembles, especially on the corridors between two adjacent field shapes, as taken from the LPIS. A comparative oversight is displayed in the Table 5-22 below.

Table 5-22 - Before and after the cleaning of isolated pixels.

	
Results before the cleaning of the isolated pixels between two contiguous parcels of the LPIS	Results after the cleaning of the isolated pixels between two contiguous parcels of the LPIS

In phase 2, only a majority filter of 5 pixels is applied, which removes most of those border artefacts, as seen in Figure 5-17 and Figure 5-18.

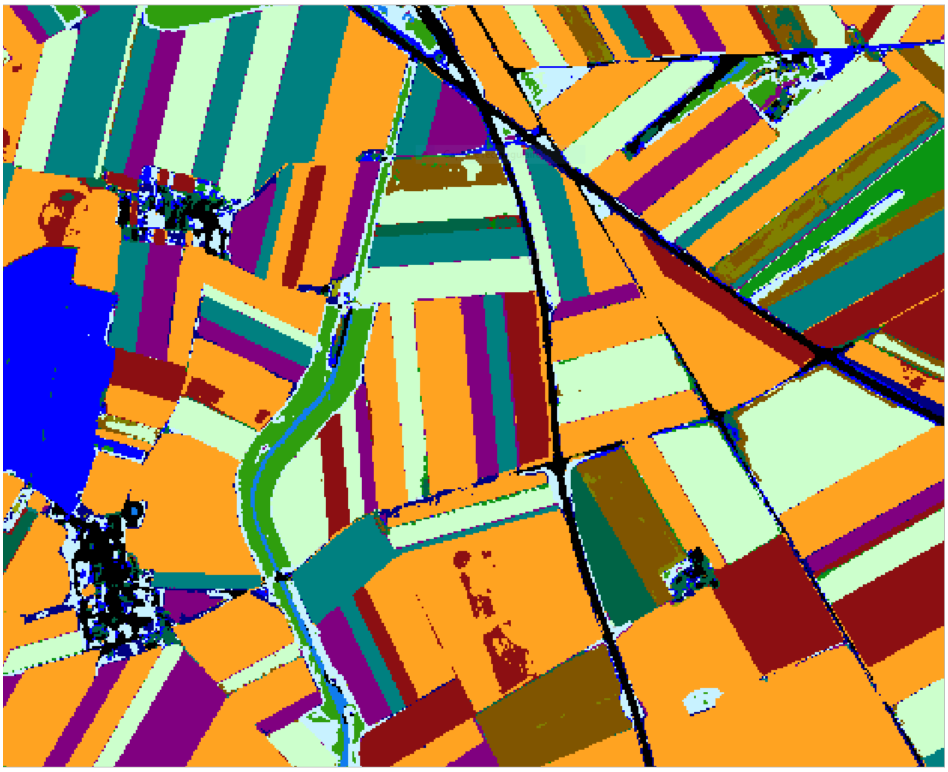


Figure 5-17 - Raw classification for the crop types, without post-processing.

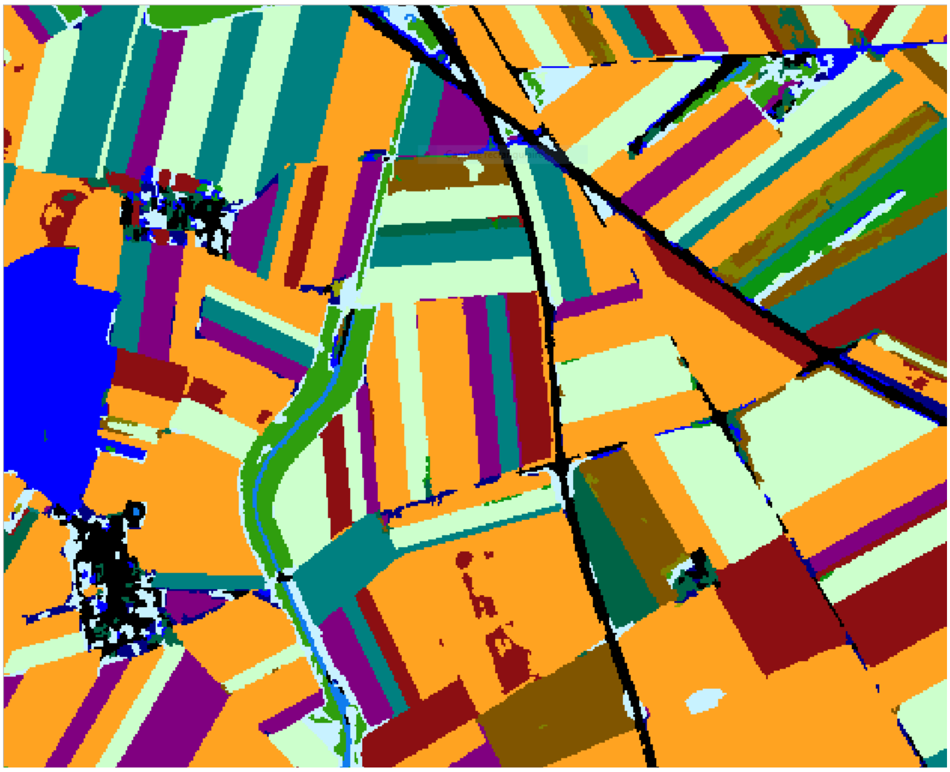


Figure 5-18 - Post-processed classification for the crop types, with a majority filter of 5 pixels applied.

The layer names are set following naming convention: layer_year_resolution_demosite_EPSGC_version.tif, as discussed between partners.

In phase 1, two products, named respectively CRT_2016_010m_WE_03035_prototype_v01.tif for the detailed cropland type prototype layer and CRM_2016_010m_WE_03035_prototype_v01.tif for the derived binary mask between cropland and non-cropland pixels are conformed to the following specifications inTable 5-45.

In phase 2, the two products are then: CRT_2017_010m_WE_03035_prototype_v01.tif and CRM_2017_010m_WE_03035_prototype_v01.tif.

Demo-site West (Belgium Part)

As a first post-processing step, the LPIS polygons was used as a cropland mask to constrain the classification results to the actual agricultural parcels.
A majority filter based on the LPIS polygons was then applied as a second post-processing step to produce a coherent result in the form of one crop type class per field parcel (Table 5-23).

Table 5-23: Before and after majority filtering.

	
Results before applying the majority filter	Results after applying the majority filter

Demo-site Central (phase 2)

Post-processing of the CRM comprised a MMU filtering to enhance the look-and-feel of the classification result.

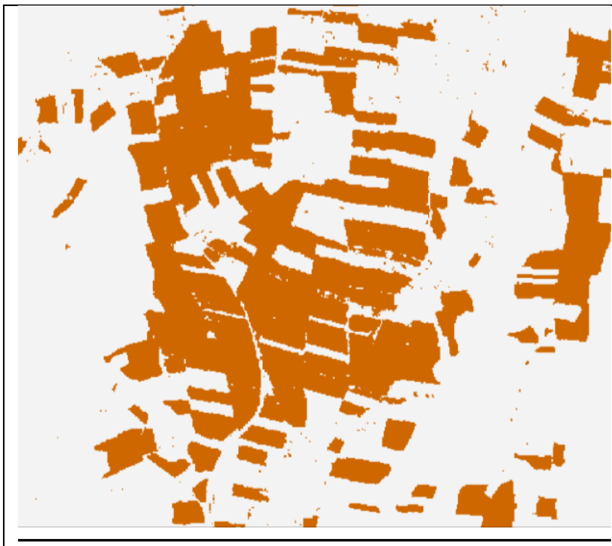
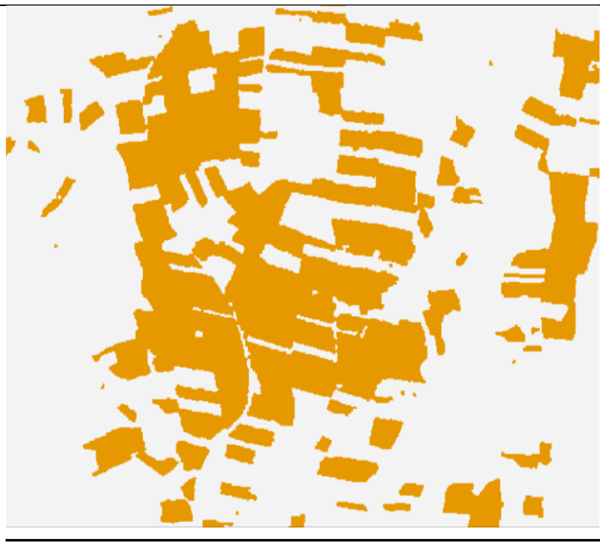
	
Raw crop mask for demo site central	3x3 pixel filtered crop mask for demo site Central

Figure 5-19: Phase 2 - Raw 3crop mask for emo site Central and 3x3 pixel filtered crop mask

The post processing of the CRT strongly depends on the availability of reference data. If LPIS reference data are available, the LPIS geometries can be used for calculating the majority per parcel. This approach highly corresponds to reality where predominantly one crop grows per parcel. In case, no reference parcels are available, a morphological filter approach can be applied to provide homogeneous results. Figure 5-20 displays those two possibilities.

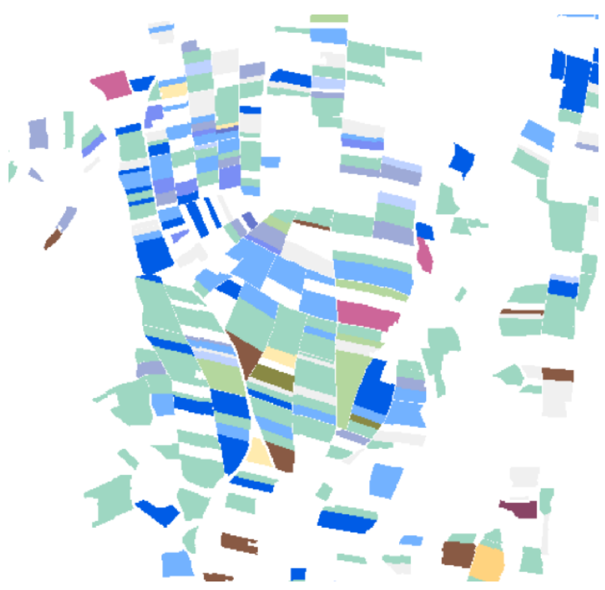
	
Crop type mask with morphology filtering + filter at longest border (Filter method applied for the final prototype)	Crop type mask with majority filtering using LPIS geometries

Figure 5-20: Phase 2 - Comparison of potential and actual post-processing for the crop type prototype in demo site Central

Figure 5-21 illustrates the situation on parcel level: apparently most of the parcel area is classified as winter barley, which is the right crop type. As the growing of plants is often uneven due to local soil characteristics, unequal humidity and therefore patchy vegetation cover during the growing period, the classifier detects several crop types. Using a majority filtering basing on actual parcel borders can limit this technical issue and lead to a more accurate product in a thematic perspective.

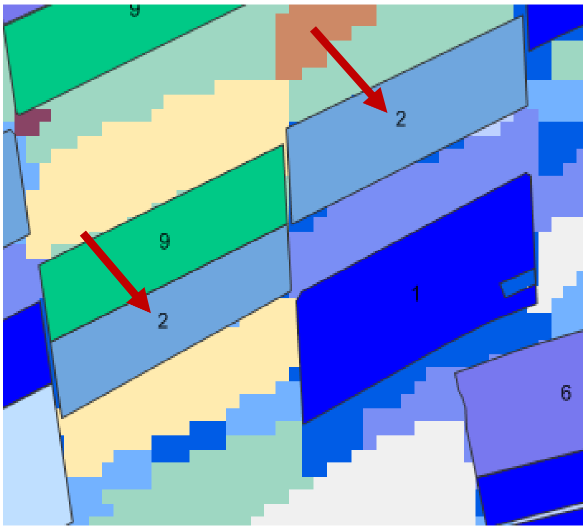
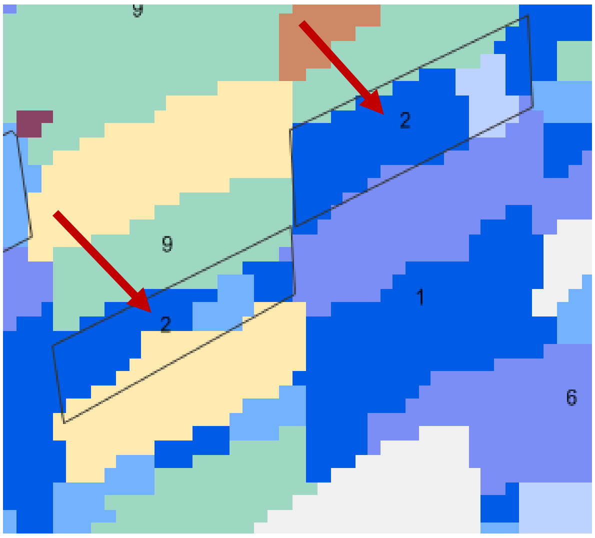
	
LPIS: 2 parcels of winter barley	
	
Classification: no clear majority of crop type within one parcel due to lacking parcel borders/geometries	Uneven growing status ending up in heterogeneity of phenology and thus in mixed crop type for classification

Figure 5-21: Phase 2 - Comparison between LPIS reference data, crop type classification re and google map indicating the heterogeneity on parcel level.

Unfortunately, reference data fully covering large areas are still the limiting factor. In a Pan-European perspective, the approach of using geometries in order to match the field level reality and offering the user a parcel-based crop type mask is not yet possible.

5.2 Results and Validation

This Chapter contains the main results. It contains information about the Crop mask (CRM) and Crop Type (CRT) maps of the Demonstration-Sites West, Central as wells as Mali.

5.2.1 Crop type map of Demo-Site West

During the phase 1, several tests were conducted over each of the 4 Sentinel-2 tiles of the French part of the demonstration site until the handling of the number of crop type classes gave satisfying results. After various addition, suppression or fusion of classes.

These results were quite promising, especially for the tile 31UFR, where the overall accuracy is near 85%. However, the tile 31UEQ overall accuracy remains below the 70% threshold, despite various attempts to improve the classification results. The final confusion matrix for the four demonstrations tiles is shown in Table 5-24 and reaches 77% as overall accuracy. According to this matrix, some classes are well identified, such as beets, while others, in particular potatoes and linen, still exhibit strong confusion with other classes. The rest of the classes are identified at an acceptable producer accuracy (between 63% and 77%) and a user accuracy between 60% and 78%. The overall confusion matrix can be seen in Table 5-24.

- .

Table 5-24: Global confusion matrix for the four S-2 tiles of the French West part of the demonstration site, for phase 1.

DEMO-SITE WEST (4 WUs) GLOBAL ASSESSMENT		REFERENCE LABELS													User Accuracy	Commission
		Other Land Cover	Winter Wheat	Winter Barley	Spring Cereals	Peas	Winter Rape	Maize	Agrarian Grassland	Beets	Potatoes	Fallow	Linen	Chicory		
PRODUCED LABELS	Other Land Cover	14796	79	0	32	0	28	200	113	4	26	22	36	1	96.47%	3.53%
	Winter Wheat	17	4175	426	699	46	194	0	48	0	0	48	3	0	73.82%	26.18%
	Winter Barley	17	542	3686	229	65	173	0	94	0	2	13	27	0	76.03%	23.97%
	Spring Cereals	221	605	72	5171	358	630	150	80	0	55	234	324	103	64.61%	35.39%
	Peas	140	35	142	197	1856	168	0	79	0	10	183	0	0	66.05%	33.95%
	Winter Rape	37	269	203	218	12	3347	0	1	0	17	5	24	0	80.98%	19.02%
	Maize	201	12	71	232	3	5	4437	412	75	112	98	1	203	75.69%	24.31%
	Agrarian Grassland	41	172	174	336	3	243	308	5706	0	34	250	5	0	78.47%	21.53%
	Beets	115	0	0	0	0	0	143	5	4554	79	0	40	2	92.22%	7.78%
	Potatoes	48	86	0	219	6	54	149	124	91	1075	114	56	35	52.26%	47.74%
	Fallow	212	22	0	41	3	355	324	1099	0	112	2125	4	0	49.45%	50.55%
	Linen	50	122	0	288	62	50	19	121	5	16	143	3033	184	74.10%	25.90%
	Chicory	0	0	32	91	0	18	183	0	0	783	147	0	1869	59.85%	40.15%
Producer Accuracy		93.09%	68.23%	76.70%	66.70%	76.88%	63.57%	75.04%	72.39%	96.30%	46.32%	62.83%	85.36%	77.97%	77.08%	
Omission		6.91%	31.77%	23.30%	33.30%	23.12%	36.43%	24.96%	27.61%	3.70%	53.68%	37.17%	14.64%	22.03%		

The derived product from this first prototype, the cropland mask, can be also used for validation. This secondary product is a reclassification of the first one, but a different quantitative approach for validation was conducted: a set of stratified random points was generated over the 4 tiles, and the following confusion matrix was extracted, see Table 5-25.

Table 5-25: Confusion matrix for the derived crop mask, over the 4 tiles.

CROPLAND MASK (plausibility)		REFERENCE LABELS			
		Non-Cropland	Cropland	User Accuracy	Commission
PRODUCT LABELS	Non-Cropland	245.19	15.92	93.90%	6.10%
	Cropland	4.51	134.38	96.75%	3.25%
	Producer Accuracy	98.19%	89.41%	94.89%	
	Omission	1.81%	10.59%		

Table 5-26: Overall Accuracy (OA) on both products over the 4 tiles.

	31UEQ	31UER	31UFQ	31UFR	Cropland Type	Cropland Mask	Remarks
OA	69.03%	78.30%	78.26%	84.30%	77.08%	98.89%	Could be improved with: - a modification of the classes, - better samples - use of textural indices - use of S-1 data - reclassification of the cropland types on the cropland mask only

For phase 2, the classes have been altered to fit the unified nomenclature set between partners. The sampling process has been entirely revised – and even though the SAR dataset has not been used for the classification step, mainly due to the issue addressed in section 5.1.1.1, the overall accuracy of the crop type product reaches now 87%, a gain of 10 points, in comparison with the first phase. The results for the cropmask and the crop types products are displayed in Table 5-27 and Table 5-28. Those results are automatically generated by the algorithm, based on the calibration dataset based on the samples.

Table 5-27 - Confusion matrix for the French West tiles (for phase 2)

French West 4-tile site		REFERENCE																				
		maize	Su mer cereals	winter_cereals	sugarbeet	potatoes	Root crops	vegetables	soybeans	Fruit trees	Grape vines	sunflower	oleaginous	rapeseed	grassland	fodder	grasshrl	urban	forest	water	Total	UA
PRODUCT	maize	5893524	14341	31350	21102	88611	36172	140858	50680	78017	134526	76106	28085	13516	87195	9752	1192	114255	30207	4271	6853760	0.86
	Summer cereals	37770	6546597	320851	51664	7164	26369	128115	2582	72085	51862	34581	389728	54702	59624	12527	1824	62104	17124	9319	7886592	0.83
	Winter cereals	40827	1231569	34190486	62702	15065	49564	153288	3334	489843	42071	26769	417355	306360	859070	104728	57053	195709	49153	34933	38329879	0.89
	sugarbeet	78555	42745	31089	8172387	286274	162434	68675	14513	30232	30890	122175	14793	13497	2850	15024	49	3662	6476	771	9097091	0.90
	potatoes	31860	5655	5583	59240	1456867	61359	183578	28729	4630	12016	77623	27559	2597	1479	3212	24	5211	713	154	1968089	0.74
	root_crops	1913	3997	508	4377	14314	196318	10391	829	735	1748	735	1169	56	325	303	0	674	36	20	238448	0.82
	vegetables	9490	26236	4463	10759	29763	18263	470833	1523	3316	7229	10704	34169	1774	1482	785	63	10398	768	361	642379	0.73
	soybeans	1231	89	511	374	4741	542	823	21301	416	782	3287	61	20	730	413	13	170	146	15	35665	0.60
	fruit_trees	465	194	1585	296	16	658	3283	41	80631	3535	76	146	207	14418	646	2725	1668	3639	98	114327	0.71

grape_vines	437	137	177	201	55	381	1600	7	3965	140436	180	64	10	1275	37	6	1420	69	42	150499	0.93		
sunflower	5085	958	1271	1989	16320	850	4928	5905	2075	4459	221486	2526	883	845	2143	4	3016	512	327	275582	0.80		
oleaginous	3572	31764	5810	3683	13451	3340	30307	7611	2645	6809	7202	401140	4561	1160	1220	1	6309	510	105	531200	0.76		
rapeseed	13636	83311	178689	13325	4296	1266	38641	1342	57193	2809	22179	28916	8677070	64653	42263	2416	4870	7124	3488	9247487	0.94		
grassland	10099	9700	30564	1911	2692	14428	3316	3109	128713	19373	4442	4096	6054	575609	56459	85696	13720	13622	1101	984704	0.58		
fodder	20404	60925	77252	61340	26679	39820	43120	15987	102381	35555	11583	28442	29171	434903	2269588	40512	8700	9412	854	3316628	0.68		
grasshrl	0	0	30	0	0	0	0	0	355	154	0	0	0	2115	169	1425	28	28	0	4304	0.33		
urban	0	0	3	0	2	1	234	0	58	69	12	1	0	26	0	0	4884	0	2	5292	0.92		
forest	2	17	0	0	0	0	2	0	378	30	0	0	0	18	0	2	143	5371	30	5993	0.90		
water	45	2	74	4	34	9	0	30	110	17	0	0	4	67	0	24	280	165	3301	4166	0.80		
Total	6148915	8058237	34880296	8465354	1966344	611774	1281992	157523	1057778	494370	619140	1378250	9110482	2107844	2519269	193029	437221	145075	59192	79692085			
PA	0.96	0.81	0.98	0.97	0.74	0.32	0.37	0.14	0.08	0.28	0.36	0.29	0.95	0.27	0.90	0.01	0.01	0.04	0.06			0.870	**OA**
																						0.825	**Kappa**

Table 5-28: Confusion matrix of the CRM prototype in the demonstration siteWest.

		Reference Data		
		0	1	
Class Name	**CRM 2017**	Non-cropland	Cropland	Totals
Non-cropland	0	**14981**	4774	19755
Cropland	1	10358057	**69314273**	79672330
		10373038	69319047	79692085

The F-Scores, as well as each major confusion between crops and non-cropland classes, can be found in Table 5-24.

Table 5-29 - F-Scores and maximal confusion for each class over the French West tiles.

Classes	F-Scores	Maximal confusion with:
Maize	0.907	Vegetables, Grape vines
Summer Cereals	0.821	Oleaginous, Winter cereals
Winter Cereals	0.934	Summer cereals, Cultivated Grasslands
Sugarbeet	0.931	Potatoes, Root crops
Potatoes	0.741	Vegetables, Sunflower
Root Crops	0.462	Potatoes, Vegetables
Vegetables	0.489	Oleaginous, Potatoes
Soybeans	0.221	Potatoes, Sunflower
Fruit Trees	0.138	Cultivated Grasslands, Forest
Grape Vines	0.436	Fruit Trees, Vegetables
Sunflower	0.495	Potatoes, Soybeans
Oleaginous	0.420	Summer cereals, Vegetables
Rapeseed	0.945	Winter cereals, Summer Cereals
Cultivated grasslands	0.372	Fruit Trees, Natural Grasslands
Fodder	0.778	Cultivated Grassland, Fruit Trees
Natural grasslands	0.014	Cultivated Grasslands, Fruit Trees
Urban	0.022	Vegetables, Grape vines
Forest	0.071	Fruit Trees, Urban
Water	0.104	Urban, Forest

For UCL, as detailed in section 5.1.3.1, the RF classifier was applied in two ways. The first is by training and applying RF on each tile individually while the second is by training RF on the 31UFR only.

For the first case, the overall accuracy was computed for each tile before and after grouping the crop type classes as well as after applying the majority filtering (Table 5-30). The 31UFR tile recorded the highest overall accuracy (88.81%) followed by the 31UFS (84.13%) and 31UES (82.99%) in case of not grouping similar classes. The overall accuracy has improved after grouping similar classes with an increase between 3% and 8% while the highest overall accuracies for all tiles (> 90%) are obtained after applying the majority filter.

Table 5-30: Overall Accuracy (OA) of classification results based on specific model per tile. OA is presented before and after grouping crop classes and applying majority filter on for the 31UFR, 31UFS and 31UES tiles.

Tile	OA before grouping classes (up to 159 different classes)	OA for classes grouped (24 classes)	OA after majority filtering (24 classes)
31UFR	88.81 %	92.35 %	93.92 %
31UFS	84.13 %	88.55 %	90.76 %
31UES	82.99 %	91.13 %	93.55 %

For the second case, the accuracy assessment detailed in Table 5-30 show that the classifier has lower performance on the tiles not used for the training (which is expected). Thus, the final product of the tiles is kept separated.

Based on the abovementioned accuracy assessment, the final products are considered as the maps produced from training and applying the RF classifier along with applying the post processing steps including the majority filter (Figure 5-8).

Table 5-31: Overall Accuracy (OA) of classification results based on a single model trained on 31UFR tile for the 31UFR, 31UFS and 31UES tiles.

Tile	OA for classes grouped (24 classes)
31UFR	92.35 %
31UFS	68.59 %
31UES	59.71 %

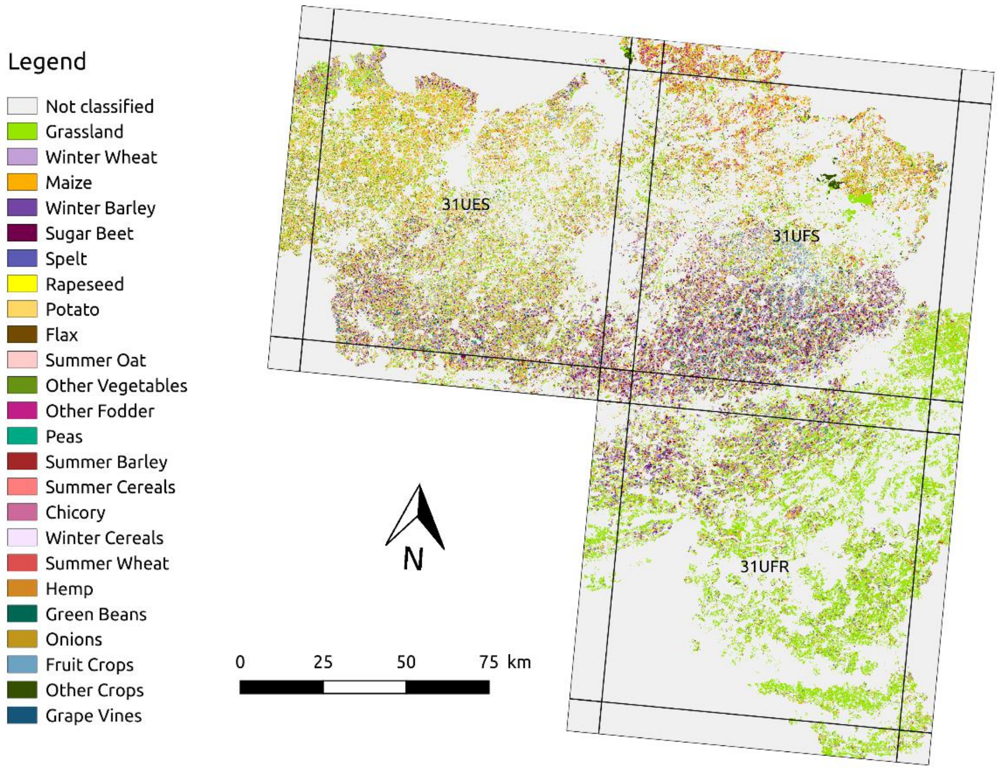


Figure 5-22: Classification results on UFR-UFS-UES tiles after majority filtering.

Figure 5-23 depicts the F1-score where the classes are ordered based on decreasing area coverage. Overall the classifier was able to differentiate several of the minority classes. The differences in accuracy between the tiles can be explained by the occurrence of these classes in the training dataset between tiles (for example the absence of Grape Vines in the 31UES tile).

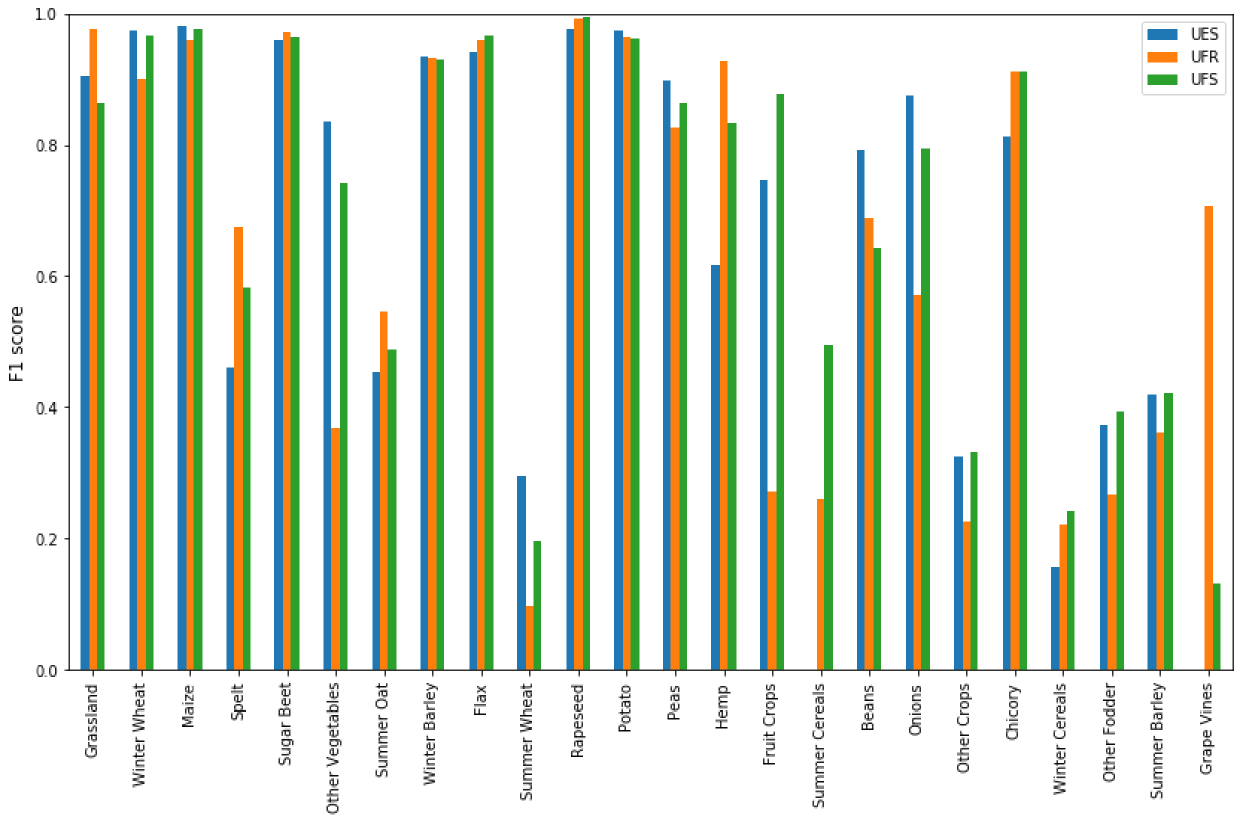


Figure 5-23: F1-score per crop type for UFR, UFS and UES tiles. Ordered by decreasing area over the 3 tiles.

The phase 2 of the demonstration for demo-site West is based on the most recent availability of the Sen2-Agri v.2.0 and Sen4CAP systems.

The Sen2-Agri based on Sentinel-2 only and the Sen4CAP outputs are reported at the Table 5-32 for a comprehensive typology, including respectively 75 and 41 crop types for Flanders and Wallonia and for the ECoLaSS paneuropean typology (22 crop types). This results synthesis highlight the high performance of the crop type mapping based on LPIS. The comparison first concludes to a slight improvement (around 3 % in OA) by adding the Sentinel-1 time series to the Sentinel-2 observation. The comparison also highlights the interest of the ECoLaSS paneuropean typology which provides figures higher than 95 % with 22 crop types. It is worth mentioning that these metrics always include the grasslands which contributes for more than one third of the area.

Table 5-32: Comparison of the overall accuracies for the crop type maps produced using Sentinel-2 time series (Sen2-Agri) and Sentinel-2 and Sentinel-1 time series (Sen4CAP) for two Belgian administrative regions and two different typologies (comprehensive one and the ECoLaSS paneuropean one).

Overall accuracy	Sen2-Agri		Sen4CAP	
Crop type typology	Original	Ecolass	Original	Ecolass
Wallonia	93.05%	96.41%	95.08%	97.67%
Flanders			92.09%	95.91%

A more systematic analysis of the Fscore distribution according to the crop types is also quite informative. All the results below report the Fscore ranked in descending order according to the area covered by the crop type. The dotted black line is the cumulative curve of the total area. It could be used as a weight to compute a overall accuracy weighted by area, which is not the case here. The Figure 5-24 and Figure 5-25 indicate that 5 main crops are mapped with a Fscore higher than 0,9 and 4 extra ones higher than 0,8. They also show that grouping in the ECoLaSS typology increases the overall accuracy from 0,92 to 0,96 but do not change significantly the FScore performances for the best mapped classes. The rare classes were filtered out by the classification process as no SMOTE algorithm was applied; the resulting typology are made of 16 and 17 crop types for Flanders and Wallonia respectively.

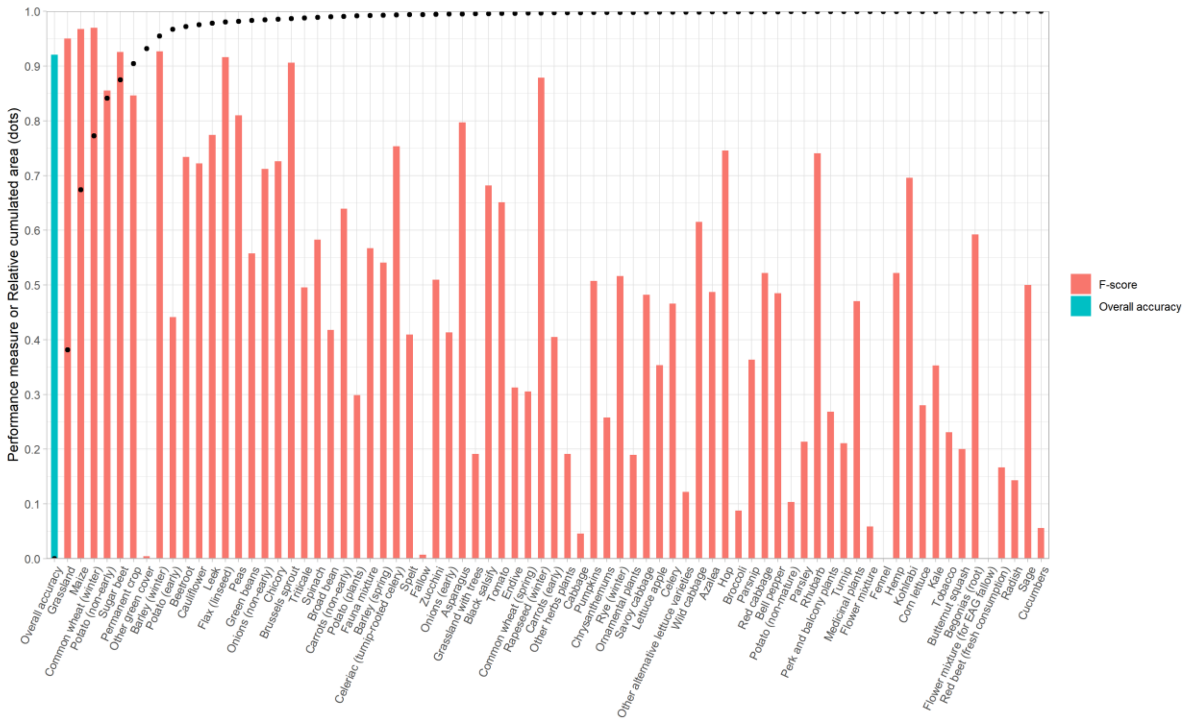


Figure 5-24: Performance metrics using the comprehensive typology of the 2018 Flanders crop type map produced from Sentinel-1 and Sentinel-2 time series by the Sen4CAP system.

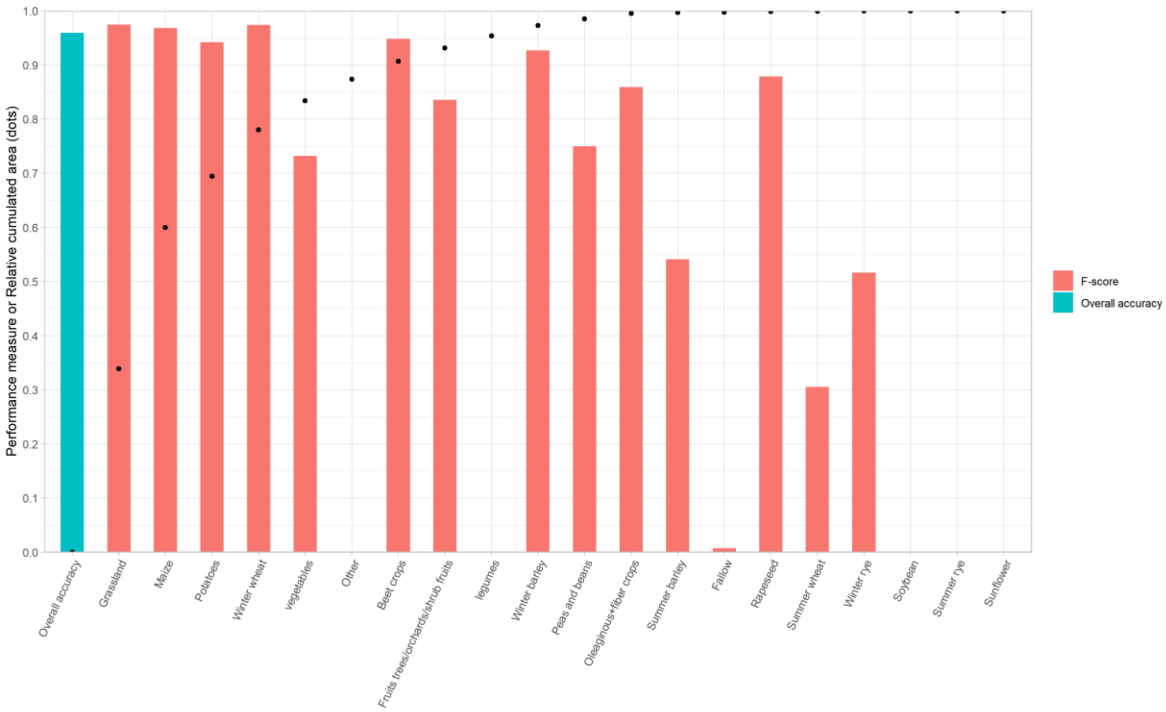


Figure 5-25: Performance metrics using the ECoLaSS paneuropean typology of the 2018 Flanders crop type map produced from Sentinel-1 and Sentinel-2 time series by the Sen4CAP system.

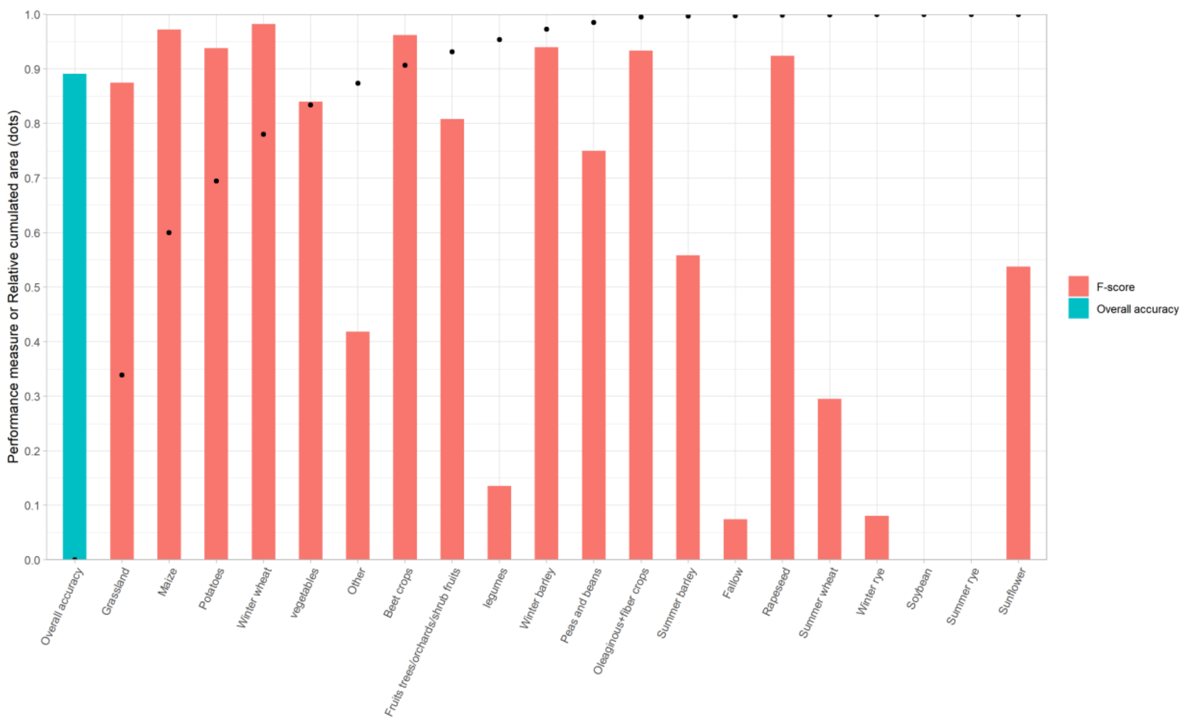


Figure 5-26: Performance metrics using the ECoLaSS paneuropean typology of the 2018 Flanders crop type map produced from Sentinel-2 time series by the Sen2-Agri system.

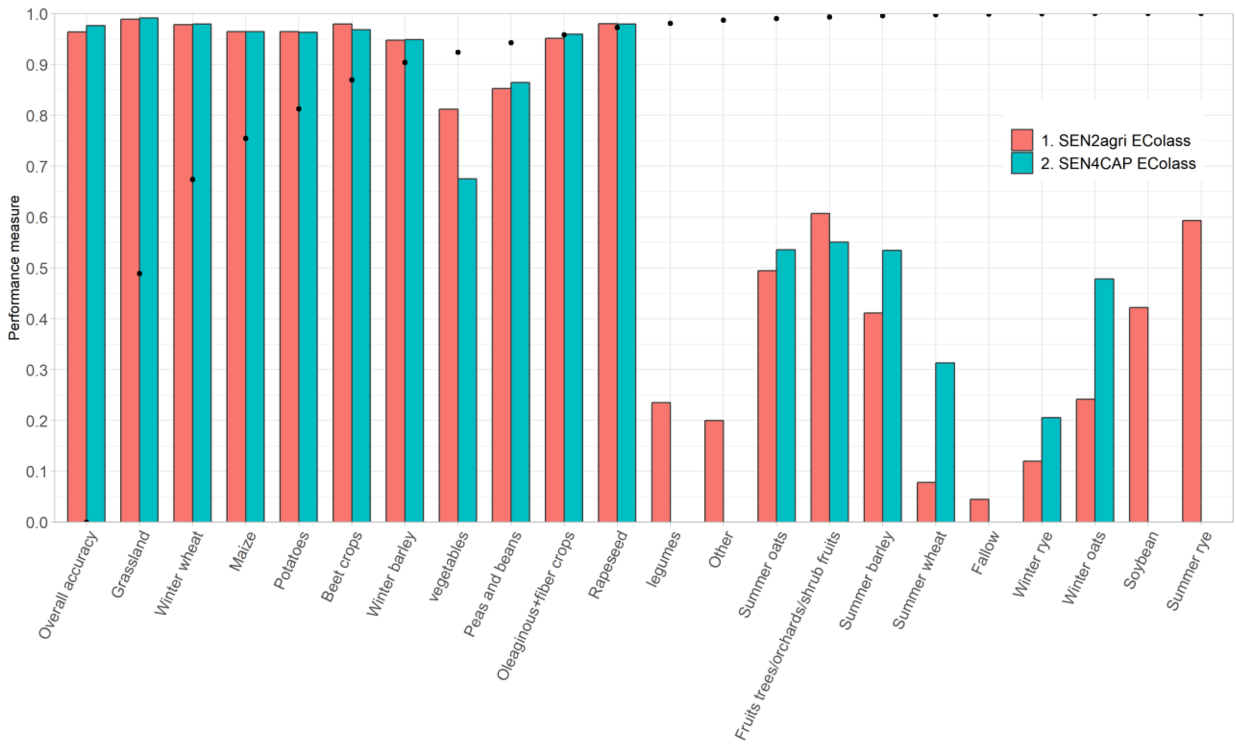


Figure 5-27: Performance metrics using the ECoLaSS paneuropean typology of the 2018 Wallonia crop type maps respectively produced from Sentinel-2 time series by the Sen2-Agri system and from Sentinel-1 and Sentinel-2 time series by the Sen4CAP system.

Both Sen2-Agri and Sen4CAP used 25 % of the in situ data as calibration and in spite of differences in the calibration dataset preparation, they provide rather similar results. It is clear that the Sentinel-1 processing cost is more justified by the enhanced reliability of the observing system than by the slight thematic improvement. The Figure 5-27 reports both the Sen2-Agri and Sen4CAP results allowing a direct and systematic comparison of both results highlighting the contribution of the Sentinel-1 time series. For most crops the use of Sentinel-1 very slight increases the Fscore values except for vegetables and to a lesser extent for beet crops.

Overall the metrics values obtained by the accuracy assessment of the 8 different crop type maps shown that the results for Wallonia are very similar to the one for Flanders. The crop type map and the corresponding confusion matrix based on the ECoLaSS paneuropean crop typology are reported here below as illustration of the prototypes.

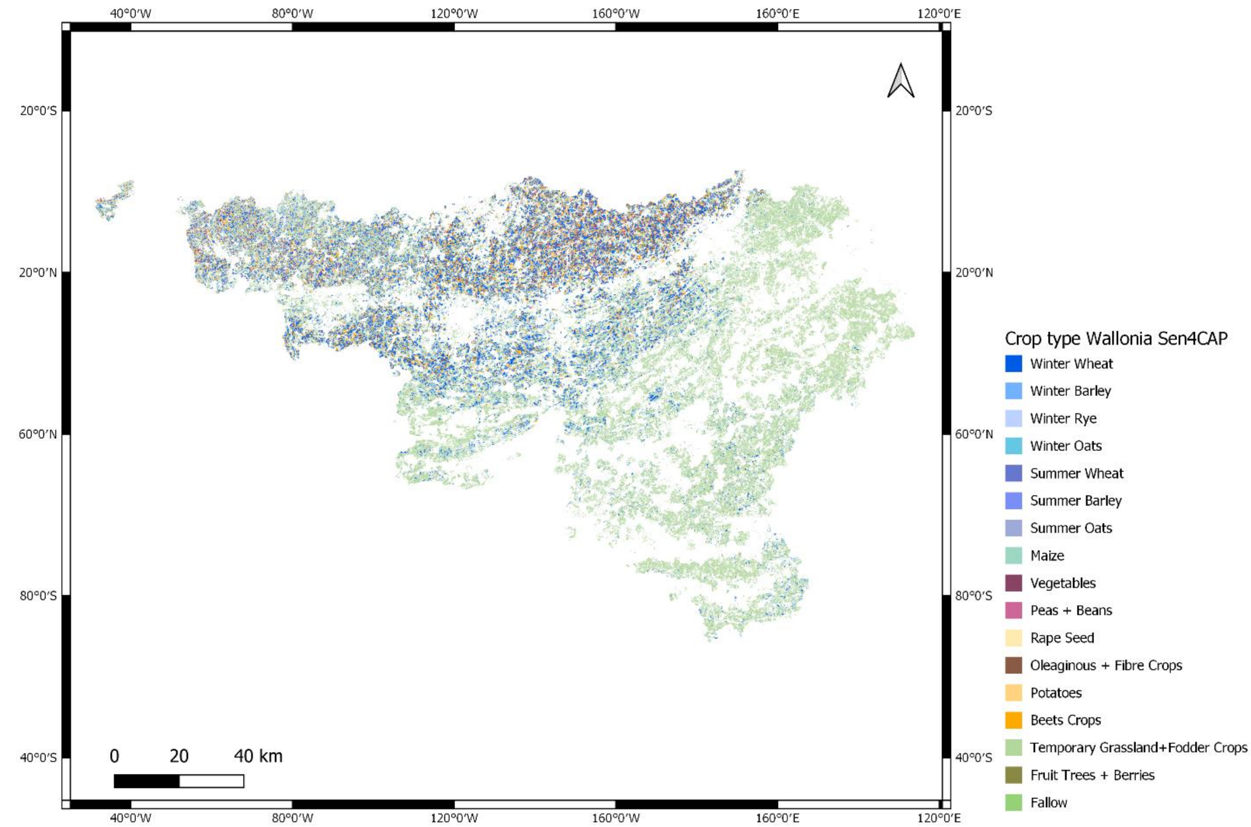


Figure 5-28: 10-meter crop type map for Wallonia (south of Belgium – demo-site West) for 2018 growing season, obtained at the end of the season using Sentinel-1 and Sentinel-2 time series with Sen4CAP system.

Table 5-33Confusion matrix for the 2018 Wallonia crop type produced from Sentinel-1 and Sentinel-2 using Sen4CAP with the ECoLaSS paneuropean crop typology.

	Reference	1	2	3	4	5	6	8	9	10	11	16	17	18	19	20	22	23		**User's**
Sen4CAP	LPIS 2018	Winter Wh	Summer Ba	Winter Rye	Winter Oa	Summer W	Winter Bar	Summer O	Maize	Vegetables	Peas + Bea	Rape Seed	Oleaginous	Potatoes	Beet crops	Grassland	Fruit Trees	Fallow	Totals	**accuracy**
Winter Wheat	1	25367	171	6	15	18	4	7	33	1	2	6	1	4	4	45	0	1	25685	98,8
Summer Barley	2	84	4799	4	1	0	14	2	9	1	0	1	2	2	0	12	0	2	4933	97,3
Winter Rye	3	134	35	26	0	0	0	0	0	0	0	2	0	0	0	3	0	1	201	12,9
Winter Oats	4	1	0	0	11	0	0	1	0	0	0	0	0	0	0	0	0	0	13	84,6
Summer Wheat	5	95	4	0	1	39	3	3	2	0	0	0	4	0	0	7	0	0	158	24,7
Winter Barley	6	46	6	1	0	3	104	6	3	0	2	0	1	2	0	41	0	1	216	48,1
Summer Oats	8	68	4	0	4	19	15	143	8	0	1	0	17	3	1	39	0	1	323	44,3
Maize	9	49	7	1	1	2	10	12	18598	37	13	4	7	200	54	306	5	14	19320	96,3
Vegetables	10	12	3	0	0	1	0	4	119	897	21	0	4	72	180	196	33	20	1562	57,4
Peas + Beans	11	17	7	0	0	2	2	0	29	37	941	0	6	39	5	57	5	2	1149	81,9
Rape Seed	16	15	4	0	0	0	0	0	2	0	2	1509	0	0	0	1	0	0	1533	98,4
Oleaginous + Fiber cr	17	22	3	0	0	1	2	5	3	0	0	0	1145	1	1	7	0	0	1190	96,2
Potatoes	18	10	5	0	0	0	0	0	55	12	14	0	7	7505	72	16	0	2	7698	97,5
Beet crops	19	10	0	0	0	0	0	0	20	79	3	0	1	33	7238	5	1	1	7391	97,9
Grassland	20	177	134	14	0	6	19	28	368	32	29	26	1	23	5	108469	107	204	109642	98,9
Fruit Trees + Berries	22	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	95	0	99	96,0
Fallow	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0,0
	Totals	26107	5182	52	33	91	173	211	19250	1096	1028	1548	1196	7884	7560	109208	246	249	181114	
Producer's accuracy		97,2	92,6	50,0	33,3	42,9	60,1	67,8	96,6	81,8	91,5	97,5	95,7	95,2	95,7	99,3	38,6	0,0		97,7

5.2.2 Crop mask map of Demo-Site Central

This chapter focusses on the results of the crop mask and crop type classification in the Central demonstration site for both reference years (2017 and 2018). The results of the feature selection are described, followed by the crop mask and crop type results and accuracies. Finally, a comparison with the HRL Grassland 2015 is provided.

During the group-based FFS performed in Phase 1 for the analysis of 2017, the accuracy reached its plateau with 13 selected feature groups, which comprise 228 features out of a total of 1246 features. After the first FFS 172 of the 228 selected features had their origin in the optical data. 56 features were derived from the radar data. In the second FFS, the 228 were further reduced to 28 features. After the two runs of FFS the selected 28 features, which are listed below in Table 5-34, were used for the crop mask and the crop type classification for the demonstration site Central. Most of these features could be derived from the optical data (20 out of 28), 8 from SAR data (4 derived from indices and 4 derived from VV and VH bands). The 20 optical feature consisted of 17 features derived from indices (only NDVI and NDWI whereas BRIGHTNESS- and IRECI-based features were not selected), and only 3 features were derived from spectral bands.

Table 5-34: Selected time features used for training and classification.

Feature No.	Scene Set (Start Date TO End Date)	Sensor	Band/Index	Featuretype
1	2017-03-15 TO 2017-11-14	S1	NDVVVH	std
2	2017-03-15 TO 2017-05-14	S2	NDVI	p090
3	2017-07-15 TO 2017-09-14	S2	NDWI	p050
4	2017-05-15 TO 2017-07-14	S1	VV	median
5	2017-03-15 TO 2017-11-14	S2	B08	p090
6	2017-05-15 TO 2017-07-14	S2	NDWI	p075
7	2017-07-15 TO 2017-09-14	S2	B03	p050
8	2017-05-15 TO 2017-07-14	S2	NDWI	actn0t42
9	2017-09-15 TO 2017-11-14	S2	NDWI	mean
10	2017-03-15 TO 2017-05-14	S2	NDVI	cov
11	2017-05-15 TO 2017-07-14	S2	NDVI	p025
12	2017-03-15 TO 2017-11-14	S1	VH	std
13	2017-03-15 TO 2017-05-14	S1	RATIOVVVH	mean
14	2017-03-15 TO 2017-11-14	S2	NDVI	max
15	2017-07-15 TO 2017-09-14	S2	NDWI	pdiff075025
16	2017-03-15 TO 2017-11-14	S2	NDVI	median
17	2017-09-15 TO 2017-11-14	S2	NDWI	std
18	2017-05-15 TO 2017-07-14	S2	NDWI	min
19	2017-03-15 TO 2017-11-14	S1	NDVVVH	pdiff090010
20	2017-05-15 TO 2017-07-14	S2	NDVI	difmax3mean
21	2017-07-15 TO 2017-09-14	S2	NDWI	actn0t42
22	2017-05-15 TO 2017-07-14	S2	NDVI	min
23	2017-03-15 TO 2017-11-14	S1	NDVVVH	mean
24	2017-03-15 TO 2017-11-14	S1	VH	mean
25	2017-03-15 TO 2017-11-14	S2	B08	std
26	2017-05-15 TO 2017-07-14	S1	VV	cov
27	2017-05-15 TO 2017-07-14	S2	NDVI	p075
28	2017-05-15 TO 2017-07-14	S2	NDWI	difdif3mean

The prototype product of phase 1 was created pixel-based and no MMU filter has been applied. This means that single classified pixels were kept in the product. Such MMU filter, as it is applied for some of the current HRLs, was applied in phase 2 and will probably be subject of future definitions of a potential agricultural HRL.

The cross-validation of the crop mask yielded an Overall Accuracy of 97% (F1-Score 0.97). As can be seen in the results shown in Table 5-35 the classification model for the distinction between cropland and non-cropland works well in general. Producer's and User's Accuracies of the non-crop class are very high (between 96% and 98%) as well as the Producer's Accuracy for the crop class (93%) whereas the User's Accuracy of the crop class is lower (85%). The latter value signalizes that 15% of the cropland is wrongly classified as non-crop.

Table 5-35: Confusion matrix of the CRM prototype in the demonstration site Central.

		Reference Data		
		0	1	
Class Name	**CRM 2017**	Non-cropland	Cropland	Totals
Non-cropland	0	**445929**	15625	461554,00
Cropland	1	6224	**94850**	101074,00
		452153,00	110475,00	562628,00

Figure 5-29 shows the whole 2017 crop mask for the demonstration site Central with the location of the examples described hereafter. It is visible that in the higher regions the cropland gets less compared to the lower parts, e.g. in the Black Forest or in the Alps in Austria and Switzerland.

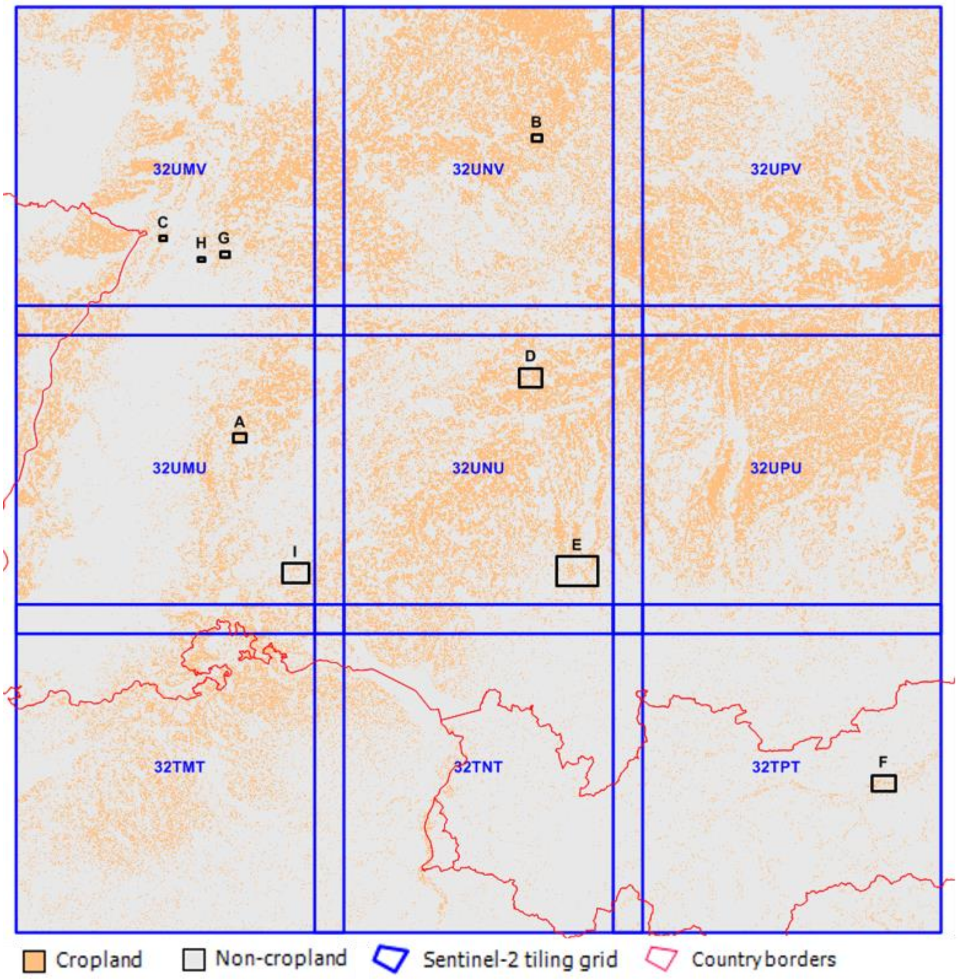


Figure 5-29: Overview over the resulting crop mask in the demonstration site Central.

A closer look on the crop mask is given in Figure 5-30. Forest areas as well as waterbodies, grassland and urban areas are masked very precisely. Only some parts of gardens and green areas in urban areas are wrongly classified as cropland and therefore not masked out. In order to address this problem, the whole pixel-based product was filtered as it is described in chapter 4.2. The resulting crop mask fixes the problem of wrongly unmasked urban areas for the most part.

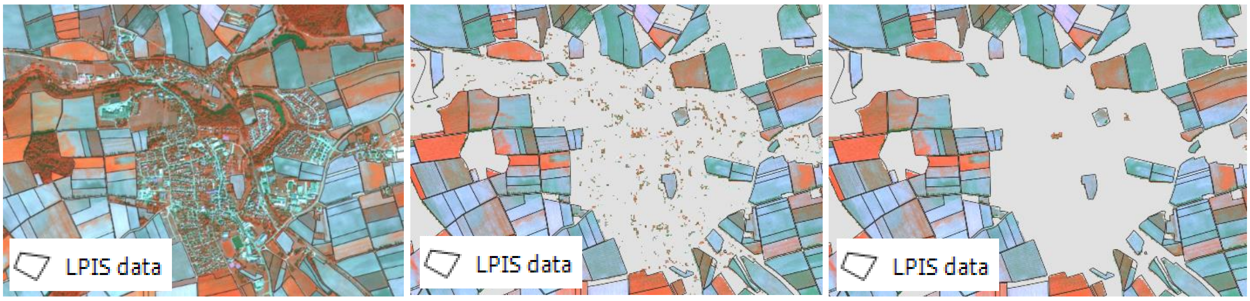


Figure 5-30: Results of the crop mask classification (middle: unfiltered, right: filtered, in grey) in NW Baden-Wurttemberg (location B) over a Spot-6 image of 2016 in combination with the reference data.

However, predominantly in the Austrian part of the demonstration-site the crop map does not only include some green areas within cities but also a lot of grassland in rural and mountainous areas (see Figure 5-31). Therefore, in particular grassland areas are sometimes recognized as cropland in the mask and not masked out, some potential for improvement is identified regarding this issue. The example below from Austria shows that forests and water are mostly cut out by the mask but some grassland areas that should have been masked out are not covered by the mask.

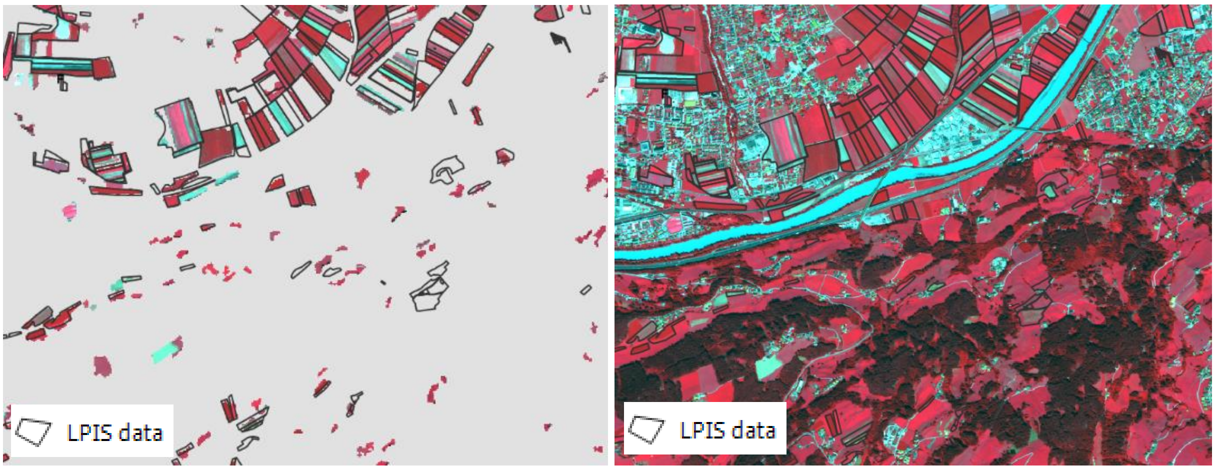


Figure 5-31: Results of the crop mask classification (in grey) near Innsbruck (AT, location F), draped over a Worldview-1 image of 2015 in combination with the LPIS reference data.

In contrary to the previous case of over-estimation of cropland, another fact that should be addressed in the future is that some fields which are present in the reference data get cut out by the crop mask, which leads to under-estimation in some areas (see Figure 5-32). This, however, primarily happens with the class Agricultural Grassland and Fallow, and especially when these fields are surrounded by a lot of Grassland, which was excluded from the prototype classification and thereby is correctly masked out in the figure below. This effect has not been largely observed for other crop types.

Figure 5-32: Results of the crop mask classification (in grey) in W Baden-Wurttemberg (location H) over a Deimos image of 2016 in combination with the reference data. Grassland (dark blue) is correctly masked out by the crop mask, Agricultural grassland (green lines) and Fallow (light blue lines) areas are wrongly masked out.

In Phase 2, the FFS reduced the number of highly relevant time features from 676 to 221 time features. Compared to phase one this number of feature is higher and related to the face that more features per time windows have been used. The optical data usually predominate for both, crop mask and crop type classification. The high accuracies for the S-2 only experimental setup for the crop mask led to the assumption that focusing on S-2 features only would lead to a sufficient accuracy (see WP33). However, using the combined approach is the better option considering the then denser time series even in regions with high cloud cover. The high percentage of S-1 features in the FFS for the crop mask 2018 supports this impression.

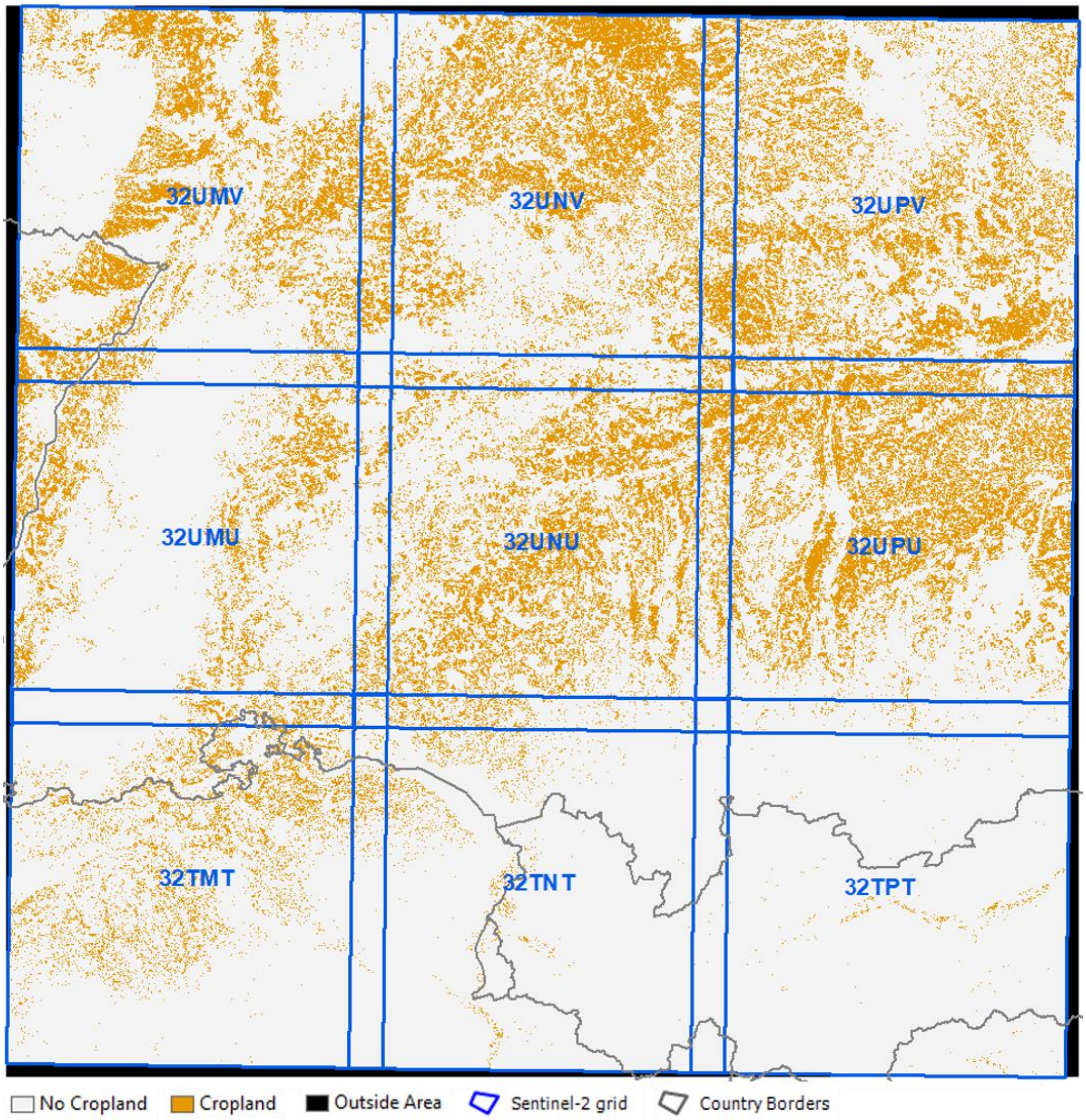


Figure 5-33: Overview over the resulting crop mask in the demonstration site Central for the reference year 2018.

The Crop mask, created by using the selected time features and samples derived from LUCAS 2018 points and complemented by validated samples from HRL2015 and manual samples, shows high accuracies: OA of the CRM 2018 is 93,28% with high producer and user accuracies. Compared to phase 1, the OA is a bit lower which is probably linked to the general data quality (cloud cover), effects of the summer drought 2018. User´s accuracies for the crop class could be improved while the PA for crops and the UA for non-crops provide comparable values for both phases.

Table 5-36: Confusion matrix of the CRM prototype in the demonstration site Central for the reference year 2018.

CRM_2018_10m_CE_03035_prototype_v01		REFERENCE					
		No Cropland	Cropland	Total	User Accuracy	Confidence Interval	F1-Score
PRODUCT	No Cropland	2766	155	2921	94,69%	0,83%	95,28%
	Cropland	119	1035	1154	89,69%	1,80%	88,31%
	Total	2885	1190	4075			
	Producer Accuracy	95,88%	86,97%		93,28%	Overall Accuracy	
	Confidence Interval	0,74%	1,95%		92,49%	Confidence Interval	
					82,72%	Kappa	

The crop mask prototype illustrates that the differentiation between cropland/fodder crops and grassland types needs high attention. Further, the detection of crop areas which show mixed information by definition, such as orchards (heterogeneous trees cover mixed with grassland patches) or vineyards (small rows of vine combined with grassland in between) are difficult to map in a comprehensive manner. A thorough selection of appropriate training samples is necessary to support good CRM results.

5.2.3 Crop type map of Demo-Site Central

The results of the phase 1 crop type classification are very promising. The overall accuracy of the 2017 classification result is at 89%. A closer look into the several classes and their accuracy results reveals some differences in performance. Table 5-37 shows that many classes (e.g. Maize, SugarBeets, WinterCrop, WinterRape) are very well distinguished with Producer's and User's accuracies between 91% and 98%. Several other classes, e.g. FruitTrees, Legume and SummerCrop still show acceptable User's and Producer's accuracies (around 71 – 86%). Classes such as strawberries and vegetables show high Producer's accuracies (around 65 – 80%) but low Producer's accuracies (around 40 – 50%) which is explainable by the fact that most reference pixels are correctly classified but many other classes are confused with these classes. Only some predominantly smaller classes, e.g. SummerRape and SunflTopinamb could not be classified very precisely, characterized by low Producer's accuracies. The class others, which provides a mix of all remaining under-represented crop types is by definition very inhomogeneous and amounts to accuracies above 50%.

Table 5-37: Accuracy results per class.

Crop Group	PA (%)	UA (%)
AgrGrass	58	74
Fallow	49	71
FruitTrees	75	71
Legume	76	75
Maize	94	95
Others	53	73
Potatoes	51	59
Strawberries	79	41
SugarBeets	95	91
SummerCrop	86	82
SummerRape	46	21
SunflTopinamb	64	28
Vegetables	68	48
Winegrowing	96	95
WinterCrop	94	94
WinterRape	98	93

The following Table 5-38 shows the confusion matrix of the crop type classification on a polygon level, with 70% of the LPIS dataset used as reference pixels (see Table 5-9 for comparison), and the reference dataset being represented in the columns. One has to take into account, however, that the total number of polygons strongly varies among the classes In general, the accuracy levels can be closely related to the size of the referring classes: High-accuracy classes such as WinterCrop, WinterRape, SummerCrop and Maize occur more often in the demonstration site, whereas SummerRape, SunflTopinamb, Strawberries and Others are underrepresented and are therefore of lower importance for the crop type discrimination in the Central demonstration site.

The confusion matrix shows that some classes, e.g. Maize, WinterCrop, WinterRape and SugarBeets could be classified very well. Their User's Accuracies are between 91% and 95%, whereas the Producer's Accuracies are between 94% and 98%. In case of SummerCrop which has a PA of 86% and a UA of 82% it is noticeable that nearly all of the scattering goes towards WinterCrop. Reasons for this confusions will be further investigated, potentially they are in the timing of the crop growth phenology of particular summer crops.

It is also demonstrated that for example 22% of the polygons classified as SunflTopinamb are actually Fallow and 16% are Maize in the reference data. Furthermore the SummerRape, which was less successfully classified has many omission errors. Nearly 80% of the SummerRape in the reference data was wrongly classified as Fallow (21%), SummerCrop (16%), or WinterRape (12%). This leads to a low User's Accuracy of 28% (SunflTopinamb) and 21% (SummerRape) respectively. While the SunflowerTopinamb class has an acceptable Producer's Accuracy of 64% the SummerRape is also problematic with a PA of 46%. One reason for that is the rather small size of the classes (see Table 5-9 and Figure 5-9).

Table 5-38: Confusion matrix of the CRT prototype in the demonstration site Central.

		Reference Data																
Class Name		1	3	5	8	9	10	11	13	14	15	16	17	18	19	20	21	
	CRT 2017	AgrGra	Fallow	Fruit Trees	Legume	Maize	Others	Pota-toes	Straw-berries	Sugar Beets	Summer Crop	Summer Rape	Sunfl/ Topinamb	Vege-tables	Wine-growing	Winter Crop	Winter Rape	**Totals**
AgrGra	1	**4997**	240	18	70	232	47	84	5	4	220	5	0	51	2	757	6	**6738**
Fallow	3	390	**5892**	44	188	628	55	161	19	16	363	5	13	79	21	412	12	**8298**
FruitTrees	5	52	119	**3235**	3	50	12	6	12	0	15	0	0	31	949	56	3	**4543**
Legume	8	198	416	12	**5921**	236	104	108	11	10	507	15	2	78	1	272	28	**7919**
Maize	9	528	849	99	216	**75207**	105	362	22	35	516	5	12	234	12	1170	41	**79413**
Others	10	12	51	11	21	156	**1046**	5	4	0	21	0	4	55	1	46	0	**1433**
Potatoes	11	177	312	12	165	191	51	**1985**	14	33	63	4	13	306	1	35	3	**3365**
Strawberries	13	115	182	51	73	106	48	64	**933**	5	78	1	3	219	59	326	20	**2283**
SugarBeets	14	41	172	12	9	81	8	32	2	**5785**	53	1	1	100	4	76	3	**6380**
SummerCrop	15	331	460	22	399	427	83	311	7	35	**30689**	4	7	105	4	4488	21	**37393**
SummerRape	16	2	54	0	26	9	0	5	0	1	40	**51**	0	3	0	27	29	**247**
Sunfl/Topinamb	17	21	136	1	22	99	17	41	1	30	32	1	**174**	18	0	17	2	**612**
Vegetables	18	398	557	120	205	861	225	275	109	52	221	7	24	**3260**	222	186	5	**6727**
Winegrowing	19	114	570	499	10	125	39	17	20	2	25	0	2	77	**28302**	40	11	**29853**
WinterCrop	20	1102	1643	187	433	1342	126	416	23	69	2680	5	11	193	34	**122950**	187	**131401**
WinterRape	21	114	275	9	38	153	16	61	2	6	160	7	4	23	4	273	**15396**	**16541**
Totals		**8592**	**11928**	**4332**	**7799**	**79903**	**1982**	**3933**	**1184**	**6083**	**35683**	**111**	**270**	**4832**	**29616**	**131131**	**15767**	**343146**

The following Figure 5-34 gives an overview over the crop type classification of phase 1 in the demonstration site Central. Marked in black and numerated from A-I are the examples, which will be presented afterwards.

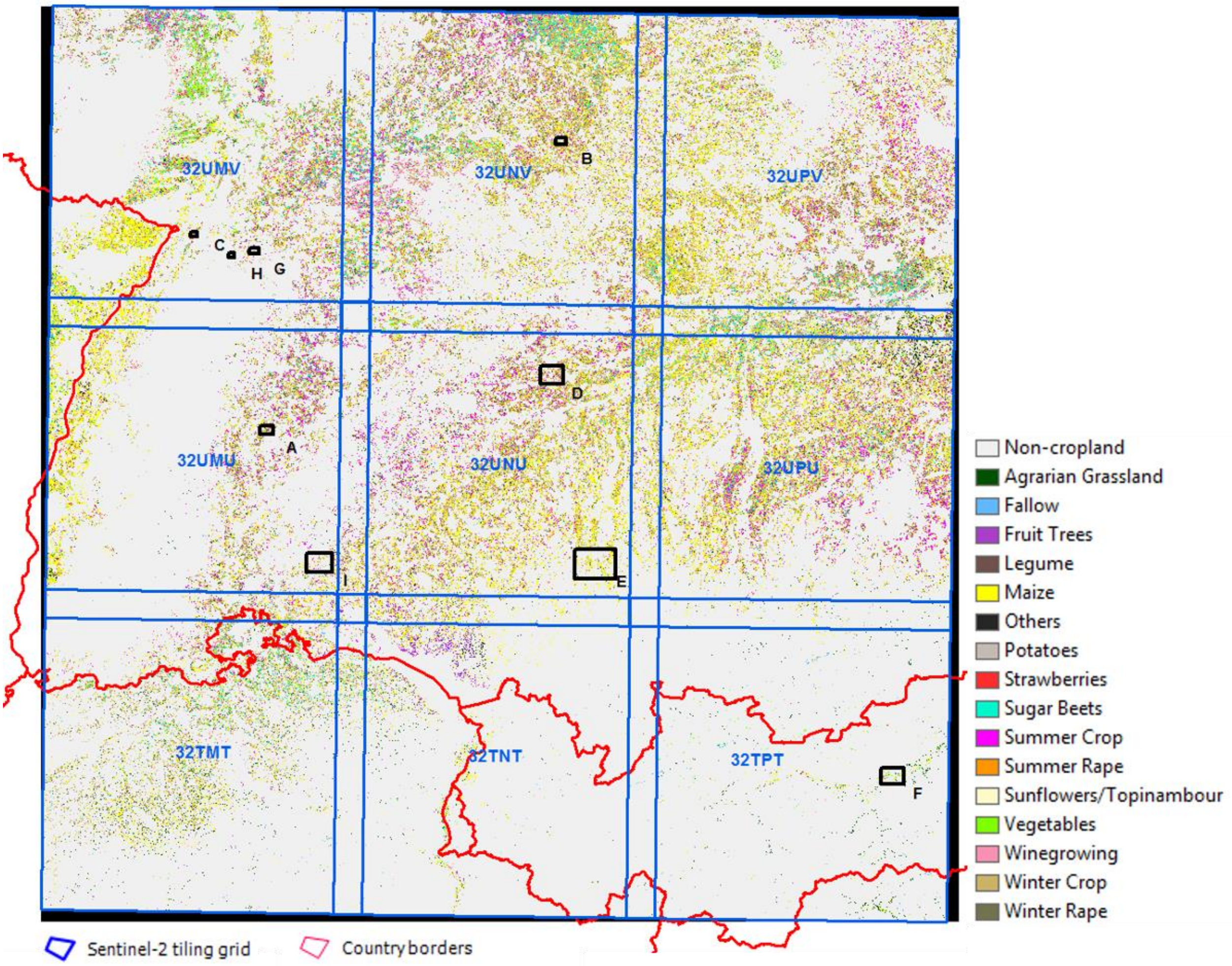
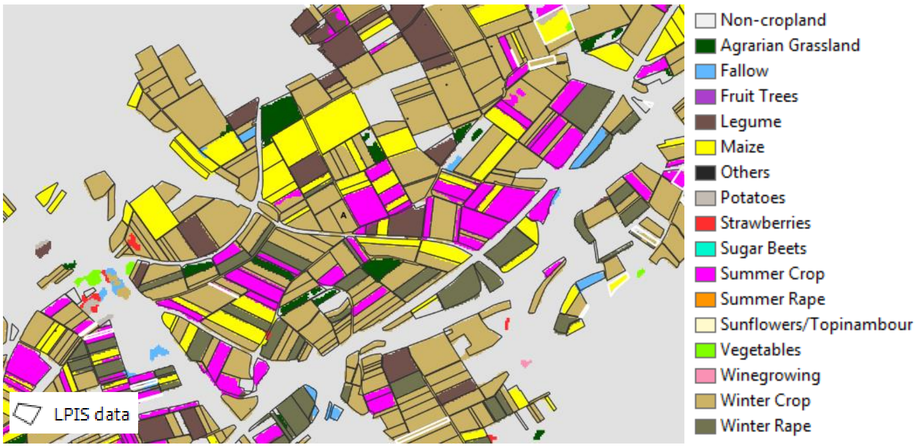


Figure 5-34: Overview over the results of the crop type classification in the demonstration site Central with the locations of the examples marked in black.

© EuroGeographics



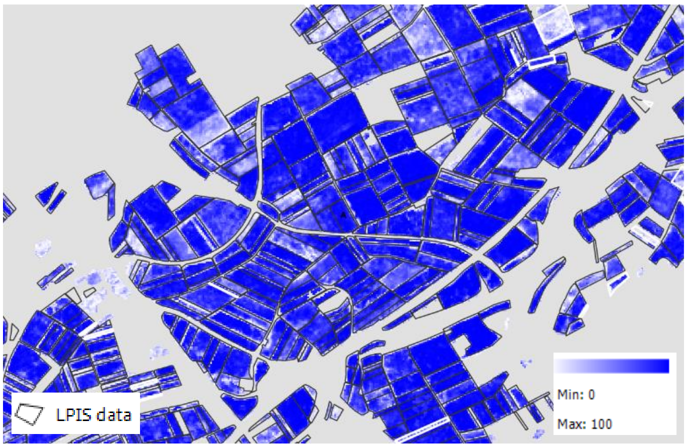


Figure 5-35: Results of the crop type classification (left) in central Baden-Wurttemberg (location A) with high breakties (right) and the classification errors (highlighted in white).

© MLR BW

Figure 5-35 depicts an example for a successful crop type classification. The upper figure shows the classification results and the one below shows the breakties of the results. Darker blue areas and higher values stand for higher breakties and thereby reliability of a crop type belonging to a certain class in the reference dataset. Additionally, field polygons where classification errors occurred are highlighted in white. Most of the fields in this area are classified correctly and only a few errors are visible.

The following example also shows good results regarding the crop type classification. Besides that, the crop map problems of the unfiltered prototype within urban areas are obvious as there are a lot of pixels defined as cropland which actually aren't cropland. As already described previously in this section, a filter was applied to the pixel-based product. The results are shown in the upper right Figure (see Figure 5-36).

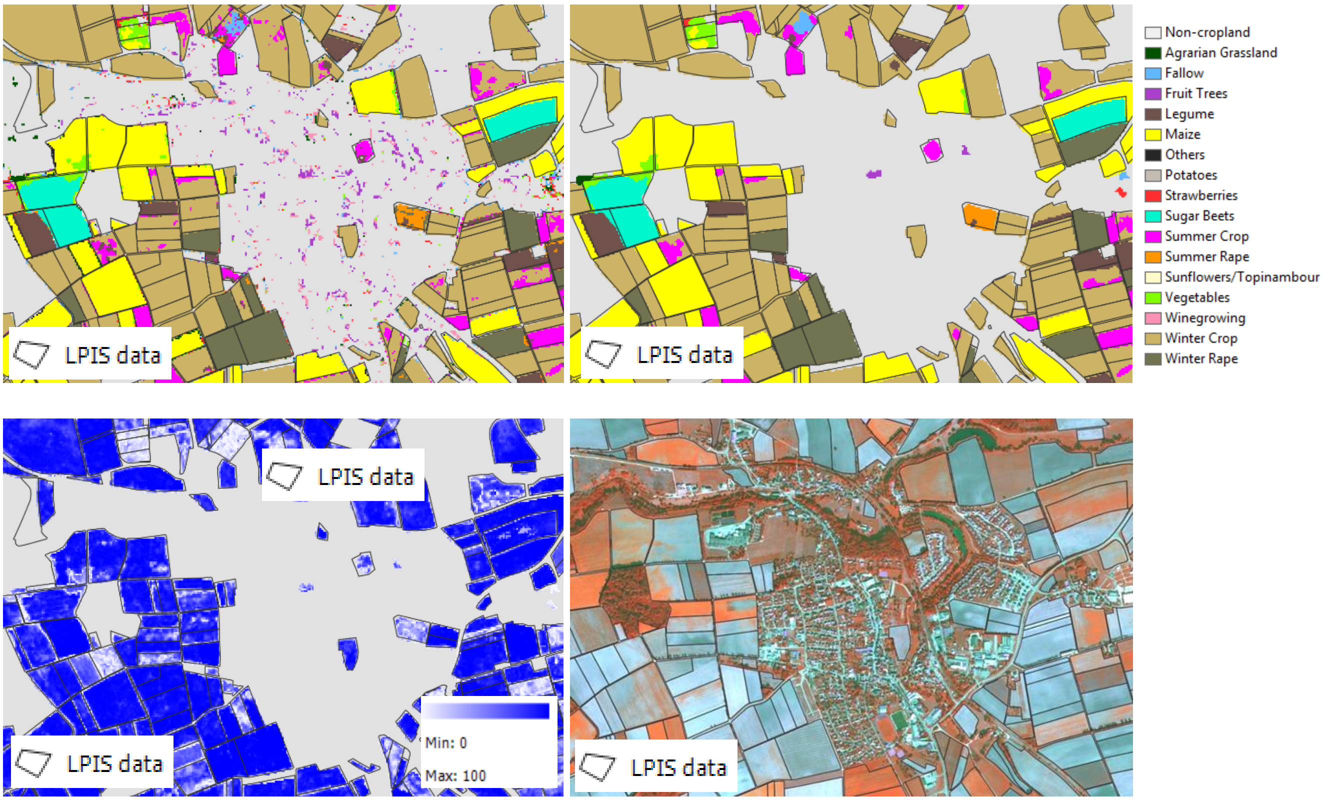


Figure 5-36: Results of the crop type classification on a pixel base (upper left figure) and after filtering (upper right figure) in NW Baden-Wurttemberg (location B) with high breakties (central figure) and a corresponding Spot-6 image (lower figure).

In some other cases the classification didn't yield as good results as in the previous examples. As the following Figure 5-37 clearly shows, the model works better for larger crop fields which was confirmed in phase 2 as well. If the field parcels are very narrow confusions with other classes are more likely and therefore crop type classification errors. Lower breakties in terms of crop types in this area of narrower fields can be depicted from the lower part of Figure 5-37. In the eastern part of the image some omission errors of the crop mask are visible.

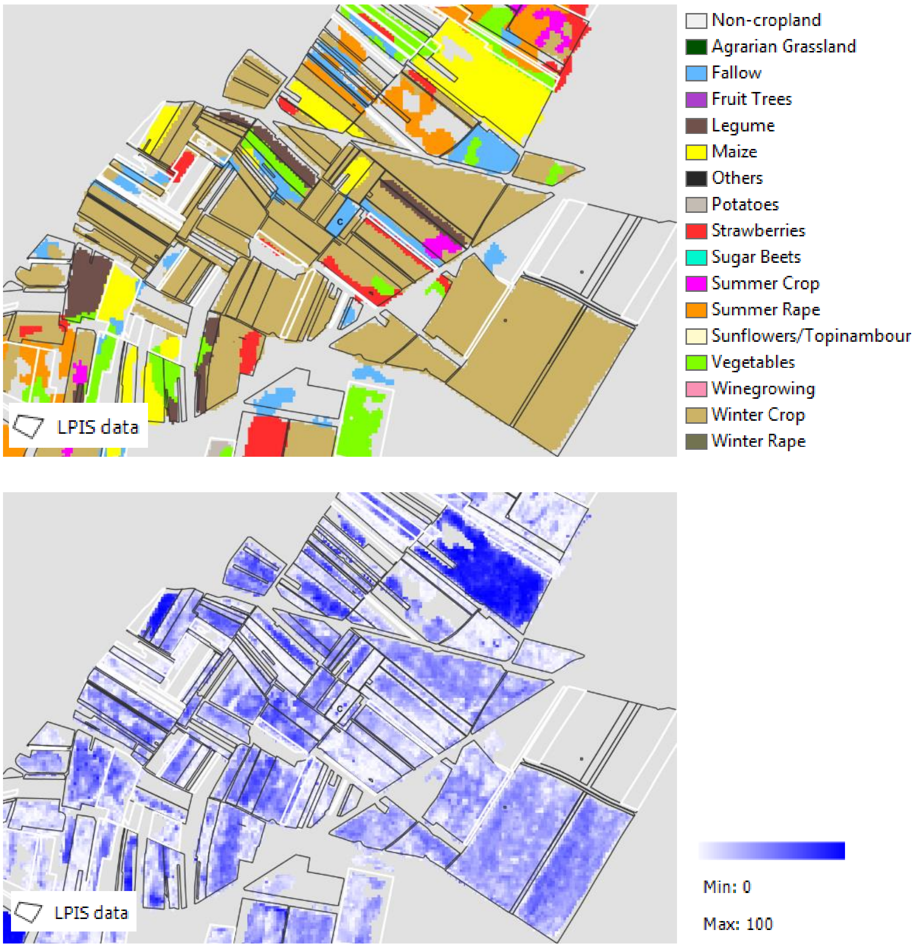


Figure 5-37: Results of the crop type classification (upper figure) in W Baden-Wurttemberg (location C) with low breakties (lower figure) and wrongly classified polygons (highlighted in white).

© MLR BW

The following figure shows the crop mask and the results of the crop type classification in combination with the HRL Grassland 2015 in southern Baden-Wurttemberg. A comparison of the crop type map with the HRL Grassland 2015 shows that the two datasets are complementary to a high degree, with a low degree of overlap. Some of the possible inaccuracies are due to the fact that all HRL 2015 products have a pixel size of 20m whereas the pixel size of the ECoLaSS prototypes is 10m.

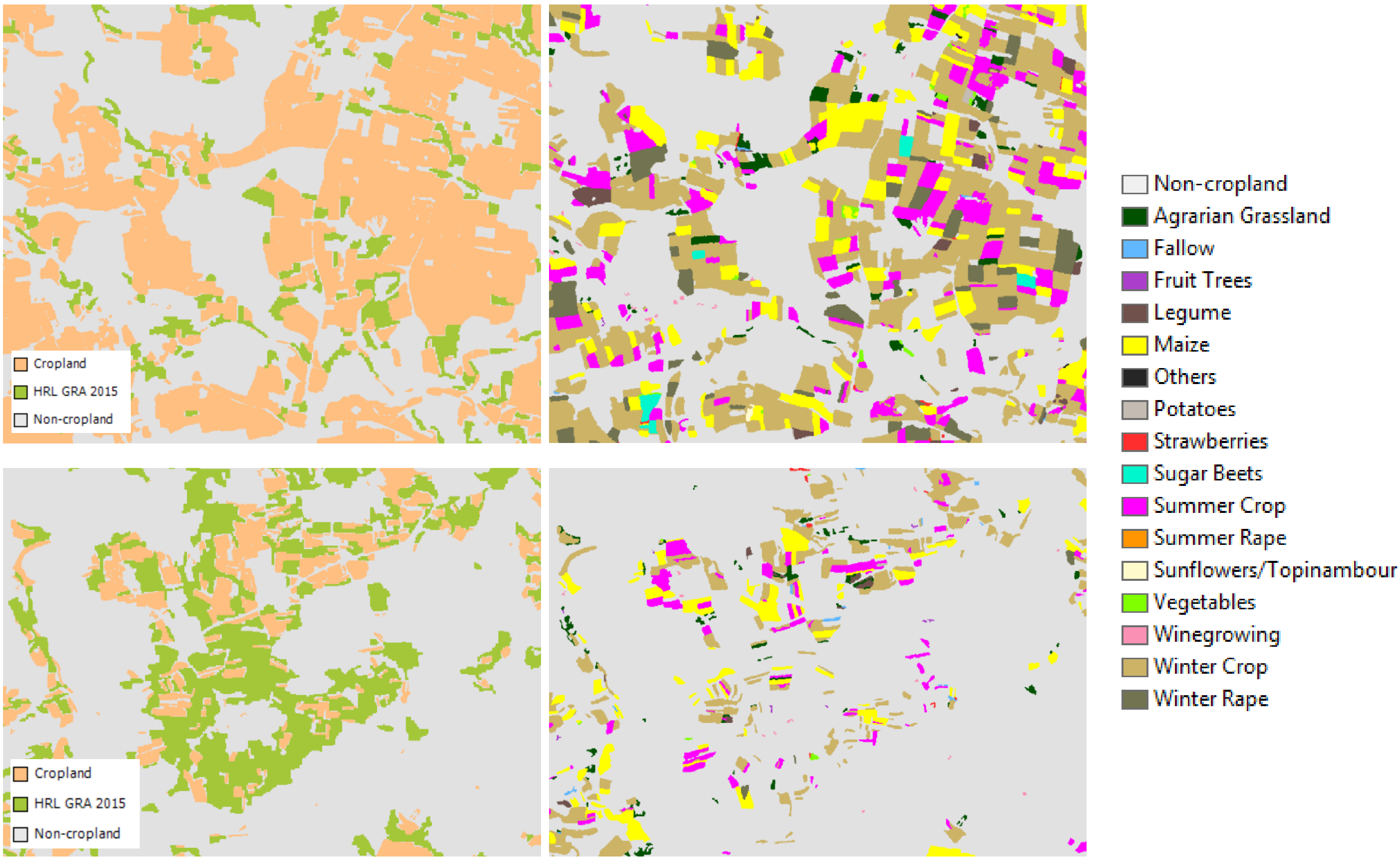


Figure 5-38: Results of the crop map classification in combination with the HRL Grassland 2015 in central (location D, upper Figures) and southern Baden-Wurttemberg (location I, below).

The crop type classification for 2018, Phase 2, uses a crop type nomenclature of 22 crop covering the most relevant pan-European crops. 18 crop types have been classified in Central with an OA of 81,83%. The resulting map is presented in Figure 5-39.

Compared to the results of phase 1, the OA of phase 2 is a bit lower. This can be explained by the following findings:

- The classification is based on a new crop nomenclature covering the most relevant European crops. Some crops, however, are difficult to differentiate and produce confusions (e.g. different winter and summer cereals).
- Demo site Central covers a variety of different landscapes why a stratification approach could be useful and support classification accuracies
- Limitations of the crop mask have an impact on the crop type; omission of crop area concerning orchards and vineyards in the crop mask lead to omissions within the crop types of fruit trees + berries and vineyards
- Overlaps between temporary grassland/fodder crops and grassland lead to omissions as well as commissions
- Crop types with more heterogeneity in the sample base show more commission errors and are mixing up with other crop types
- Similar crop types show omissions and commissions. Especially cereals are mixing up. However, the confusion matrix shows that a distinction between the various cereal types is indeed possible; a compromise could be to decide for one class winter cereals and one class summer cereals

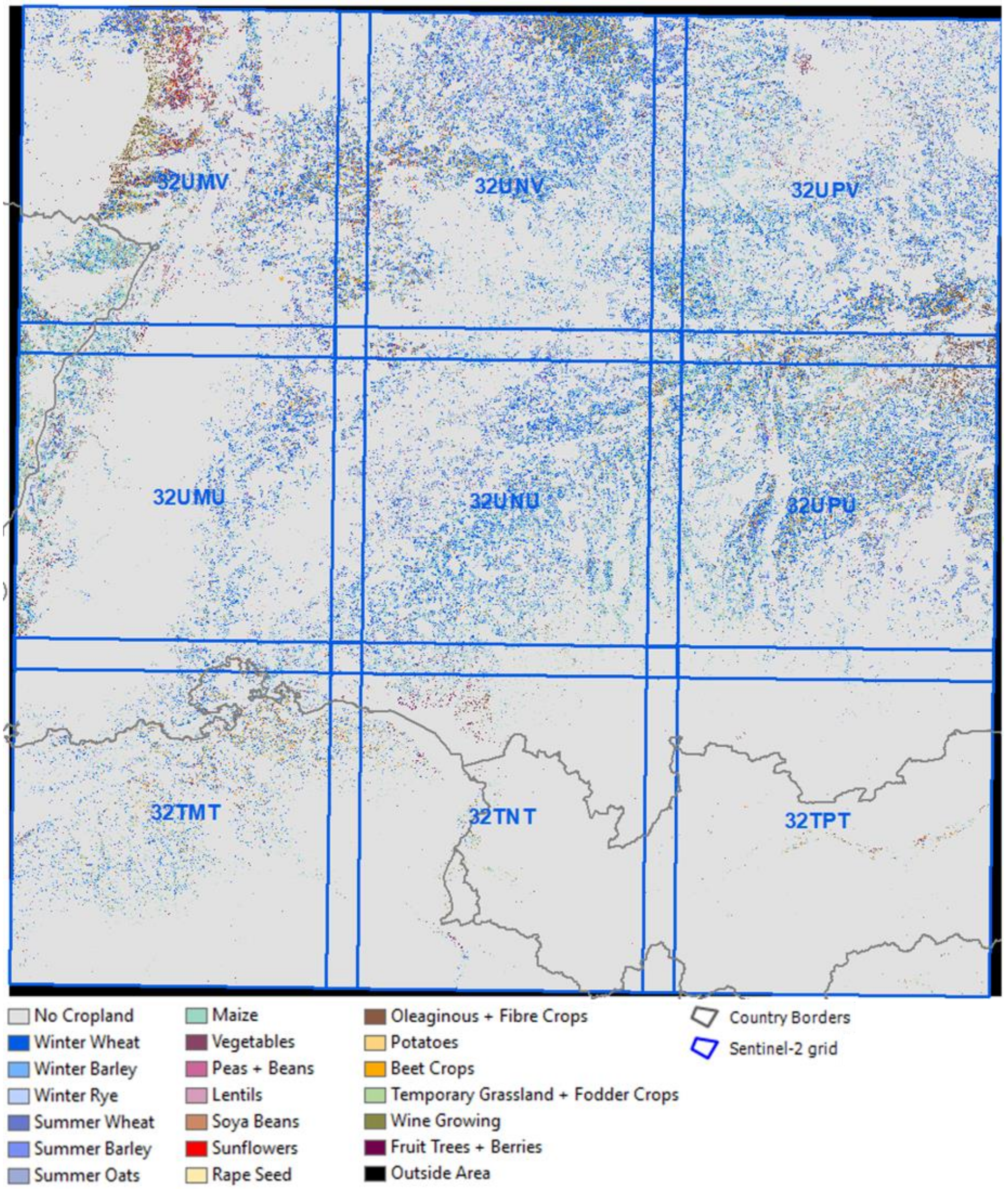


Figure 5-39: Phase 2 - Overview over the results of the crop type classification in the demonstration site Central for the reference year 2018.

Being a pixel-based product, the crop type mask first of all has been validated per pixel, using validation samples of the LPIS testing data within the crop area. In order to assess the suitability of illustrating also field level realities, a second validation focused on field level accuracies by using the LPIS geometries. The accuracies slightly differ, the OA for the validation on pixel level is 81, 84% whereas the OA on field level is 83,0%.

Table 5-39: Phase 2- Confusion matrix for crop type mask prototype - pixel based validation

		Reference Data																					
Class No		1	2	3	5	6	8	9	10	11	12	14	15	16	17	18	19	20	21	22		
	MAP (Classif	winter wheat	winter barley	winter rye	summer wheat	summer barley	summer oats	maize	vegetables	peas+beans	lentils	soya beans	sunflowers	rape seed	oleaginous + fibre crops	potatoes	beet crops	temporary grassland + fodder crops	wine growing	fruit trees + berries	total	UA
1	winter wheat	**216**	13	3	18	0	1	0	0	3	0	0	0	0	0	1	1	0	1	0	257	0,84
2	winter barley	1	**205**	27	1	1	1	2	4	1	0	1	5	0	1	1	0	1	2	0	254	0,81
3	winter rye	10	21	**194**	4	0	1	1	2	1	1	0	1	1	0	0	0	0	0	0	237	0,82
5	summer wheat	9	0	1	**131**	23	32	0	1	3	2	0	0	0	0	0	0	1	0	0	203	0,65
6	summer barley	3	0	0	22	**187**	16	0	1	4	12	0	0	0	0	0	0	0	1	0	246	0,76
8	summer oats	4	2	1	50	28	**167**	3	1	6	7	0	0	0	0	0	0	1	2	0	272	0,61
9	maize	1	1	4	1	0	8	**215**	3	1	9	7	2	0	1	0	1	3	2	0	259	0,83
10	vegetables	0	0	1	1	0	1	4	**156**	4	1	8	8	0	6	11	2	1	3	3	210	0,74
11	peas+beans	0	0	0	1	2	2	0	4	**208**	4	1	4	3	2	4	0	0	0	0	235	0,89
12	lentils	1	0	0	2	4	6	2	0	4	**127**	0	1	0	0	2	0	2	0	0	151	0,84
14	soya beans	1	0	1	0	0	1	11	0	2	1	**215**	4	0	1	16	1	3	1	0	258	0,83
15	sunflowers	0	0	0	0	0	0	0	2	1	0	0	**125**	0	0	9	0	0	0	0	137	0,91
16	rape seed	0	0	0	0	0	0	0	0	0	1	0	0	**238**	0	0	0	0	0	0	239	1,00
17	oleaginous + fibre crops	0	0	1	0	0	0	0	3	1	0	0	0	0	**190**	0	0	2	3	1	201	0,95
18	potatoes	0	0	0	0	0	0	1	12	0	0	13	15	0	1	**178**	2	0	0	0	222	0,80
19	beet crops	0	0	0	0	0	0	0	12	3	0	2	3	0	1	5	**238**	0	1	0	265	0,90
20	temporary grassland + fodder crops	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	**35**	13	0	50	0,70
21	wine growing	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	4	**13**	0	19	0,68
22	fruit trees + berries	0	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	**36**	41	0,88
	total	246	242	234	231	245	236	240	204	242	165	248	169	242	203	228	245	53	43	40	3756	
	PA	0,88	0,85	0,83	0,57	0,76	0,71	0,90	0,76	0,86	0,77	0,87	0,74	0,98	0,94	0,78	0,97	0,66	0,30	0,90		

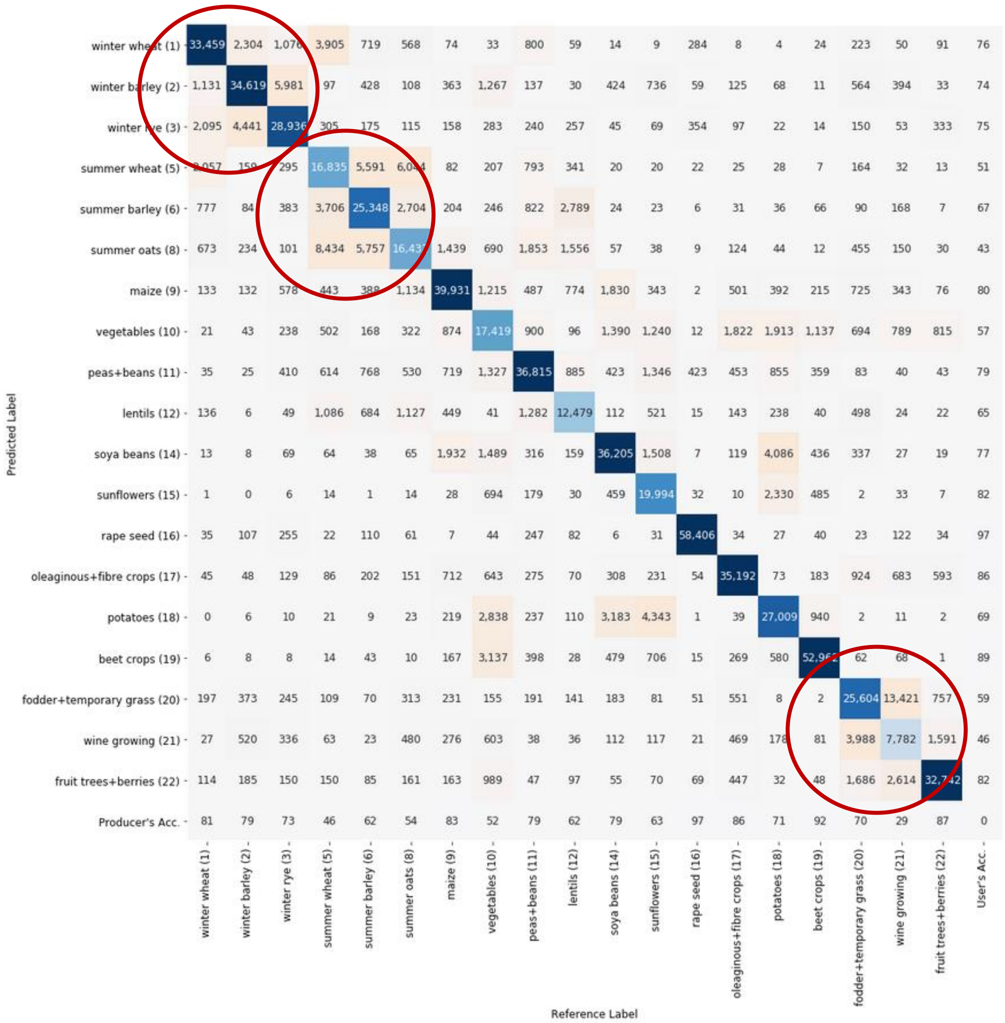


Figure 5-40: Phase 2 - Plot of the confusion matrix of the pixel-based prototype results illustrating the mixing up of cereals (winter and summer) as well as of those crop types having high percentages of grassy coverage (fodder crops+temporary grassland, wine growing, fruit trees+berries)

Table 5-40: Phase 2- Confusion matrix for crop type mask prototype - validation on field level

			LPIS reference parcels																			
			1	2	3	5	6	8	9	10	11	12	14	15	16	17	18	19	20	21	22		
			winter wheat	winter barley	winter rye	summer wheat	summer barley	summer oats	maize	vegetables	peas+beans	lentils	soya beans	sunflowers	rape seed	oleaginous + fibre crops	potatoes	beet crops	temporary grassland + fodder crops	wine growing	fruit trees + berries	total	UA
classification	1	winter wheat	41	5	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	8	5,13
	2	winter barley	2	40	9	0	0	1	2	0	0	0	0	1	0	0	0	0	0	0	0	55	0,73
	3	winter rye	1	1	35	0	1	0	0	1	0	2	0	0	0	0	0	0	0	0	0	41	0,85
	5	summer wheat	2	0	0	26	8	5	0	0	3	1	0	0	0	0	0	0	1	0	0	46	0,57
	6	summer barley	2	0	0	7	33	2	0	1	1	1	0	0	0	0	0	0	0	0	0	47	0,70
	8	summer oats	0	0	0	8	1	33	1	0	0	3	0	0	0	0	0	0	1	1	0	48	0,69
	9	maize	1	1	0	0	0	1	43	2	2	6	1	1	0	1	1	0	1	0	0	61	0,70
	10	vegetables	0	0	0	0	0	1	0	25	0	0	1	3	0	1	0	1	1	0	0	33	0,76
	11	peas+beans	0	0	0	0	0	1	0	3	42	1	0	2	1	0	1	0	2	0	0	53	0,79
	12	lentils	0	0	0	0	4	4	0	0	0	24	0	1	1	0	1	0	0	0	0	35	0,69
	14	soya beans	0	0	0	0	0	1	0	2	1	1	45	4	0	0	6	1	0	0	0	61	0,74
	15	sunflowers	0	0	0	0	0	0	0	0	0	0	0	31	0	0	1	0	0	0	0	32	0,97
	16	rape seed	0	0	0	0	0	0	0	0	0	1	0	0	48	0	0	0	0	0	0	49	0,98
	17	oleaginous + fibre crops	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	1	1	0	42	0,95
	18	potatoes	0	0	0	0	0	0	0	2	0	0	2	4	0	0	35	0	0	0	0	43	0,81
	19	beet crops	0	0	0	0	0	0	0	4	0	0	0	2	0	0	0	47	0	0	0	53	0,89
	20	temporary grassland + fodder crops	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	9	3	7	21	0,43
	21	wine growing	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	17	0	18	0,94
	22	fruit trees + berries	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	19	1,00
		total	49	47	45	42	47	49	48	41	49	40	49	50	50	42	45	49	16	22	26	765	
		PA	0,84	0,85	0,78	0,62	0,70	0,67	0,90	0,61	0,86	0,60	0,92	0,62	0,96	0,95	0,78	0,96	0,56	0,77	0,73		

5.2.4 Crop type map of Demo-Site Mali

This chapter contains information about the cropland masks, the crop type map as well as the prototype specifications.

Four operational questions were adressed by the prototypes production using the Sen2-Agri system, i.e. the stability of a RF classification process from one run to another and from one year to another, the impact of the independence of the validation dataset, the impact of stratification, and the accuracy evolution according to the delivery time along the season.

5.2.3.1 Cropland masks

Firstly, the overall accuracy of the 2017 cropland mask produced for the whole growing season has been estimated as high as 97% with a F1-Score of respectively 98% and 89% for non cropland and cropland (Table 5-41). The Figure 5-41 presents the overview of the 10-m cropland mask obtained over the south of Mali at the end of the season (November 2017). This result largely outperforms the product obtained during the demonstration phase of the Sen2-Agri mainly because of the sampling strategy and the quality of the in situ data campaign. The first Sen2-Agri map for Mali obtained in the ESA project was based on the field survey completed regularly by officers for the agriculture statistics. While the 2017 cropland mask shows an orbit artefact visible in the central part, the 2018 cropland mask obtained without stratification from April-November observation does not include any of these.

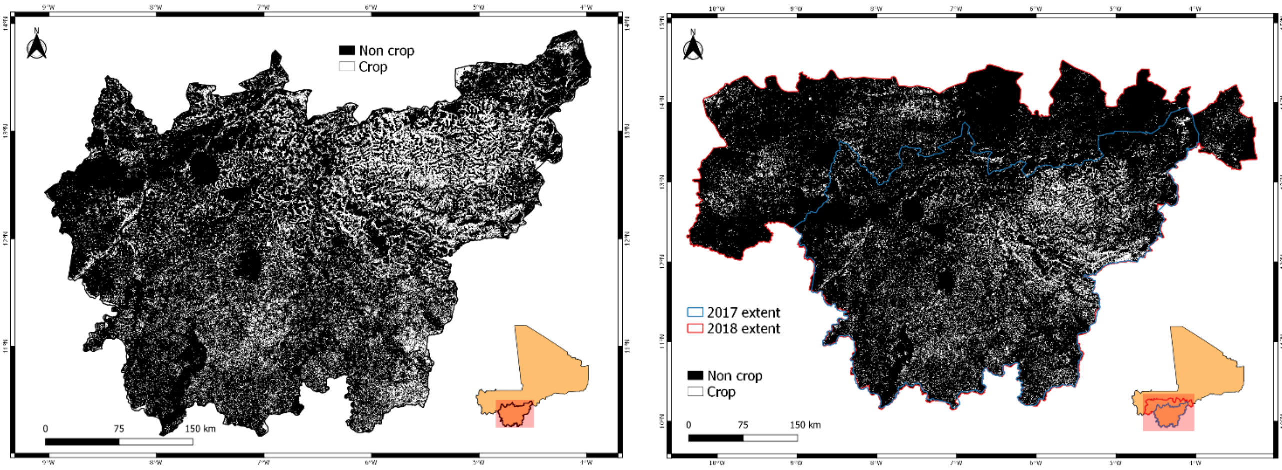


Figure 5-41: 10-meter cropland masks over south of Mali for 2017 (left) and 2018 (right) growing seasons, obtained at the end of the season (April-November) using Sentinel-2 and Landsat- 8 data. White and black represent cropland and non-cropland classes respectively.

Cropland mapping performances might vary according to the available set of in situ data and with the random process included in the RF algorithm (i.e. features and calibration data selection), seven repetitions of the RF classification based on the L4O protocol were validated thanks to the fully independent validation dataset derived from the on-screen photointerpretation. As reported by the Table 5-41, the variability induced by different location of in situ dataset and by the RF algorithm is quite limited in terms of overall accuracy (range of 0,7%) and Fscore.

It is however interesting to highlight that the assessment based on 25 % of the in situ dataset (75% for calibration) reaches an overall accuracy of 97% with a Fscore of respectively 98% and 89% for non cropland and cropland while a fully independent dataset provides an overall accuracy of 94% with a Fscore of respectively 97% and 78% for non cropland and cropland. This large Fscore difference for the cropland seems to be related to the discrepancy between on-line cropland interpretation and in situ identification as the 89% Fscore is obtained from a smaller calibration dataset (3/4 rather than 6/7 of the in situ dataset).

Table 5-41: Variability of the Mali cropland map accuracies as assessed by a fully independent dataset for the seven different training set (Leave-4-Out protocol from 27 sampled villages).

Village			Non-Cropland			Cropland		
partitions	OA	Kappa	Precision	Recall	Fscore	Precision	Recall	Fscore
1	93.9%	73.3%	0.963	0.967	0.965	0.778	0.758	0.768
2	94.0%	73.0%	0.96	0.972	0.956	0.801	0.732	0.765
3	93.8%	79.4%	0.961	0.967	0.964	0.775	0.748	0.76
4	94.3%	74.9%	0.964	0.971	0.968	0.801	0.763	0.781
5	94.5%	74.9%	0.960	0.977	0.969	0.830	0.737	0.781
6	94.0%	73.8%	0.963	0.969	0.966	0.791	0.755	0.772
7	93.9%	73.3%	0.963	0.967	0.965	0.778	0.758	0.768
Average	94.1%	74.7%	0.962	0.970	0.965	0.793	0.750	0.771
Standard deviation	0.00	0.02	0.00	0.00	0.00	0.02	0.01	0.01
Majority map	94.0%	75.0%	0.964	0.970	0.967	0.796	0.763	0.779

Similarly, the impact of validation dataset and of its spatial distribution is assessed by the L4O protocol leaving 4 villages out of the calibration but being the only one used for the validation. The reduction of the validation dataset (1/7 rather than 1/4) and the spatial independence of the validation dataset induced a much larger sensitivity (up to 12,8% in OA) to the respective sets of calibration and validation (Table 5-42).

In particular some sets, i.e. 1 and 5, produced poor or very poor results while the others are more similar. These results call for a particular attention in terms of size and spatial distribution of validation dataset.

Table 5-42: Variability of the Mali cropland map performances as assessed by spatially independent dataset for the seven different training sets (Leave-4-Out protocol from 27 sampled villages).

Village partitions			Non-Cropland			Cropland		
	OA	Kappa	Precision	Recall	Fscore	Precision	Recall	Fscore
1	88.5%	63.6%	0.893	0.967	0.929	0.843	0.606	0.705
2	94.1%	96.2%	0.943	0.972	0.957	0.936	0.875	0.905
3	91.9%	80.2%	0.923	0.966	0.943	0.91	0.812	0.858
4	95.0%	88.8%	0.955	0.971	0.963	0.941	0.909	0.925
5	82.3%	48.0%	0.815	0.977	0.888	0.878	0.428	0.576
6	91.5%	77.8%	0.917	0.969	0.943	0.907	0.773	0.835
7	95.1%	85.4%	0.97	0.967	0.969	0.88	0.892	0.886
Average	91.2%	77.1%	0.92	0.97	0.94	0.90	0.76	0.81
Standard deviation	0.05	0.16	0.05	0.00	0.03	0.03	0.18	0.13

The Table 5-43 reports the significant impact of the stratification on the quality of the cropland mask. The OA, kappa and F-score are computed for each stratum of the three different stratification strategies. The Fscore for the cropland is below 0.9 without stratification while four strata maintained a Fscore higher than 0.94 for each stratum. Some strata even showed Fscore for cropland as high as for non cropland showing that a more local RF model performed significantly better than a generic one. It seems clear that the PIRT-based stratification has a negative impact on the cropland class detection. It is worth mentioning that the OA is not really reflecting the accuracy for binary maps as it also depends on the proportion of the binary classes.

The classification performed with 4 strata outperformed the classification without stratification according to Table 5-43. However, a visual analysis balanced this conclusion and shows a larger under-estimation with the 4 strata. This was observed only on 3 small areas where VHR images were acquired at the end of the 2017 season.

Table 5-43: Fscore, Kappa and OA achieved by the end of the season cropland mask generated using S2 and L8 time series over south of Mali

	Overall accuracy (%)	Kappa (%)	Fscore	
			Non cropland	Cropland
CM no strata	97	88	0.98	0.9
CM 2 (PIRT) Strata 0	95	75	0.97	0.78
CM 2 (PIRT) Strata 2	94	72	0.97	0.76
CM 4 strata (Vintrou)	96.7	92.9	0.973	0.956
Strata 1	96.5	91.1	0.975	0.944
Strata 2	99.1	97.0	0.995	0.975
Strata 3	95.6	91.0	0.959	0.951
Strata 4	96.8	93.7	0.969	0.967

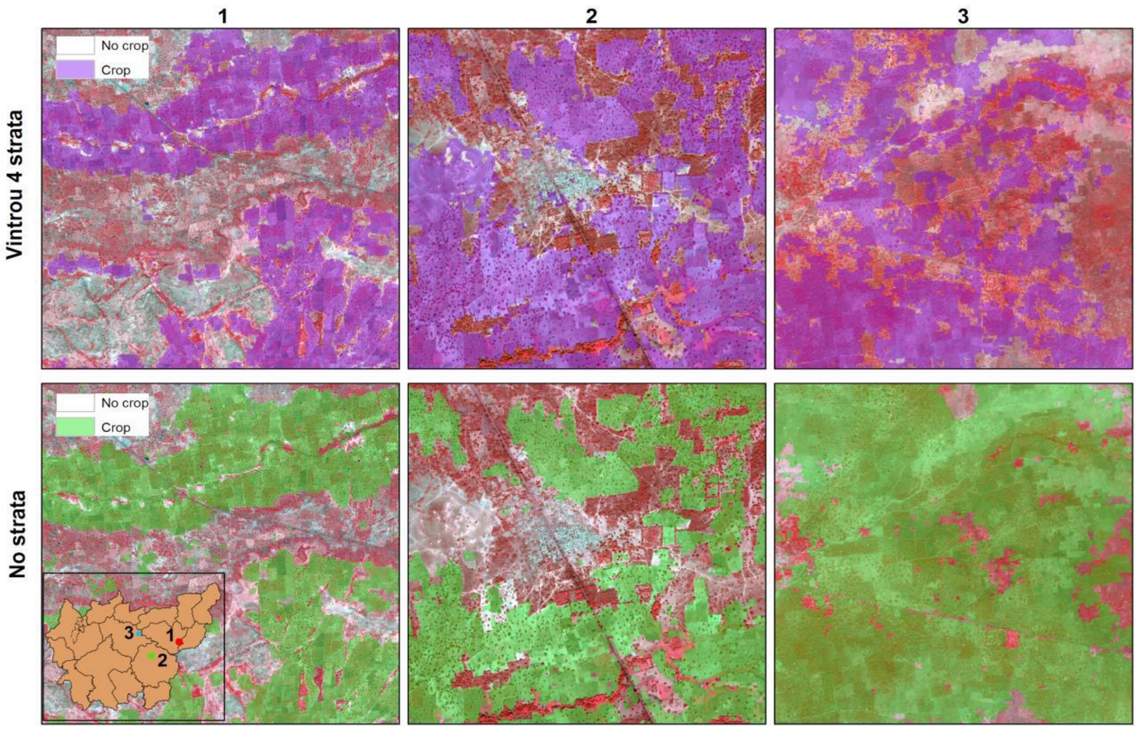


Figure 5-42: Overlay of the cropland mask on the 2017 VHR color composite. The upper subsets corresponds to the results with the 4 strata and the lower ones to the those without stratification. for three small areas.

The evolution of the mask performances along the season as illustrated in the Figure 5-43 highlights the best delivery time to top the map accuracy of the cropland mask in 2017. Among the four periods respectively corresponding to April to September, then April to October, April to November and, finally April to December 2017, the first one is probably the best trade-off between timeliness and accuracy. Last but not least, December observation clearly introduced more confusion than information probably due to the burned areas and the on going fire activities.

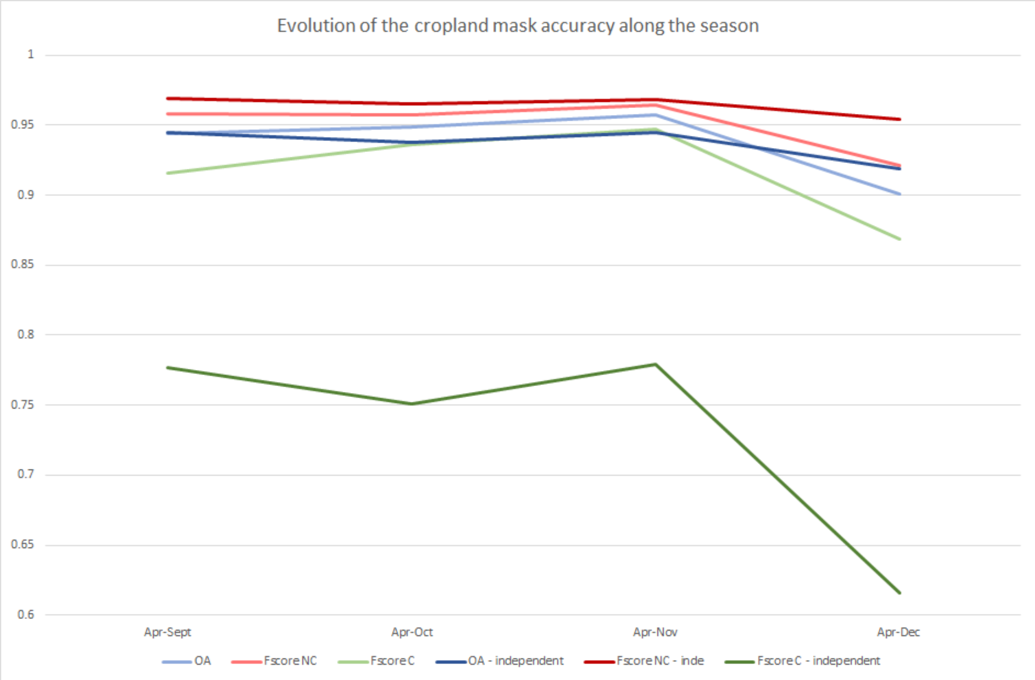


Figure 5-43: Evolution of the accuracy metrics of the different cropland mask produced along the season over south of Mali for 2017. Four periods were considered to respectively include all the S2 observations acquired first from April to September, then from April to October, April to November and, finally from April to December 2017.

5.2.3.2 Crop type maps

The accuracy of the crop type maps produced using the two strata is presented in the Figure 5-44. The crop type map reaches an OA of 63% in the northern strata (Figure 5-45(a)) and 54.6% OA in the southern strata (Figure 5-45(b)). In the northern strata, cotton crop yields to higher accuracy with a Fscore of 80% and is followed by millet (71%), sorghum (64%), maize (62%), rice (59%), other crops (43%) and groundnut (22%). In the southern strata, rice presents the higher Fscore of 89% and is followed by maize, cotton, sorghum, groundnut, other crops and millet with a Fscore of respectively 69%, 59%, 56%, 41%, 30%.

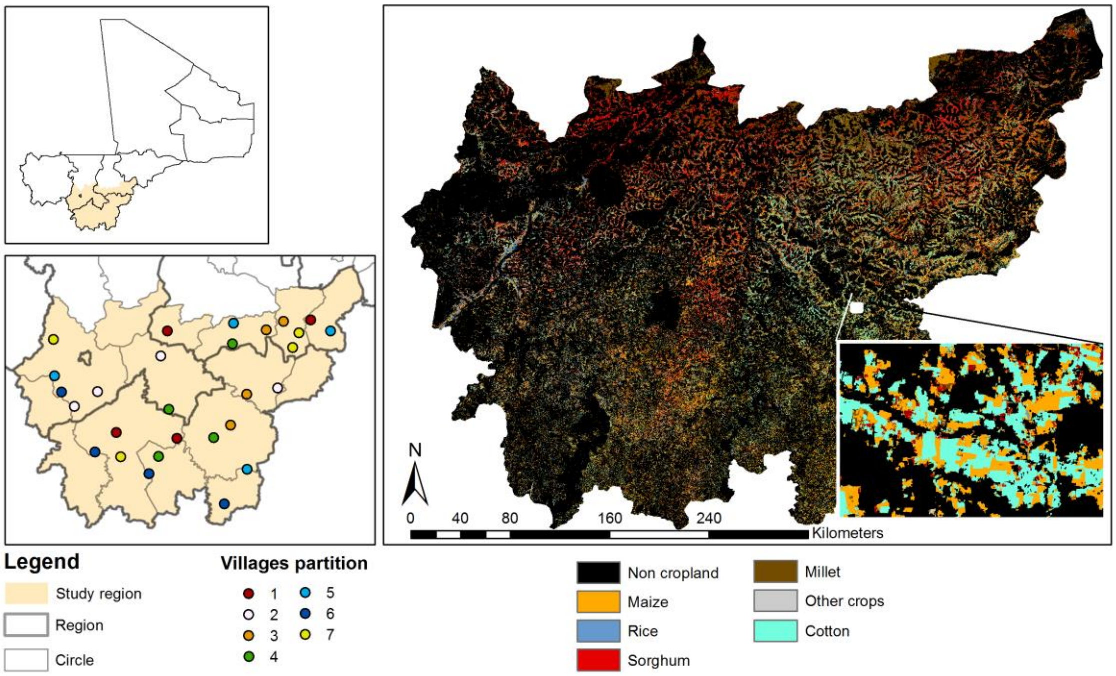


Figure 5-44: 10-meter crop type map over south of Mali for 2017 growing season, obtained at the end of the season using S2 and L8 data.

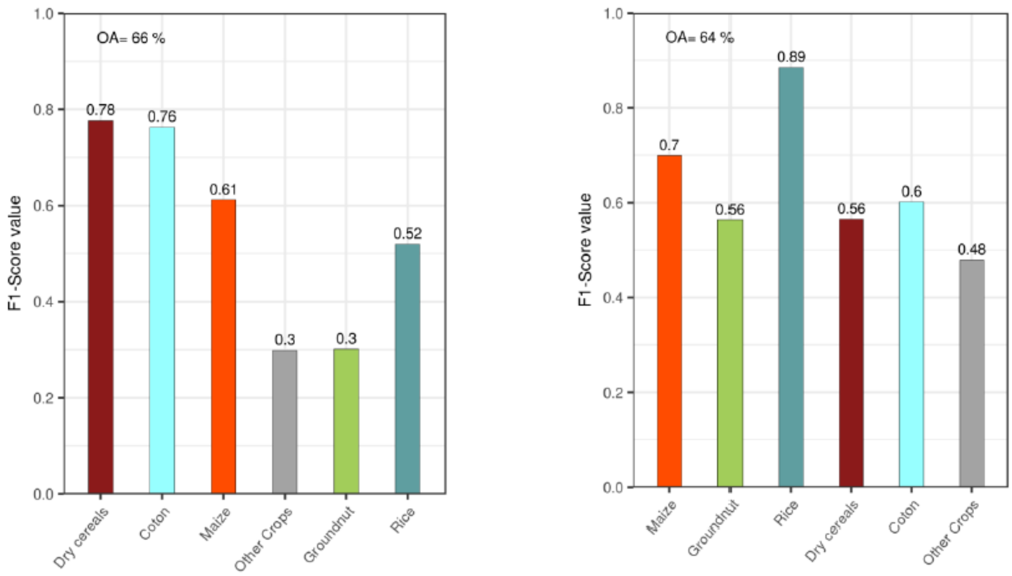


Figure 5-45: Crop map accuracy for the 2017 main crop types over the demo-site Mali (a) Koutiala strata (on the left) (b) Haut-Bani-Niger strata (on the right). Sorghum and Millet are merged into the Dry Cereals class.

A major confusion is observed between millet and sorghum which looks rather similar for a long period of the growing season and both common in the region. Combining the two crops into a dry cereals crops provided a better overall accuracy and higher Fscore. While the OA remains too low to derive useable crop area estimate, some crops are already well detected and maps of the rice, cotton and dry cereals seems becoming relevant from users perspective.

In spite of the accuracy metrics which must be improved, the spatial pattern of the crop type map resulting from a pixel-based classification is very promising as it seems to depict precisely the different fields in the complex smallholder cropping system. It is also of paramount importance to mention that these results are hardly included the contribution of Sentinel-2b satellite which should largely enhance the quantity of the satellite observation.

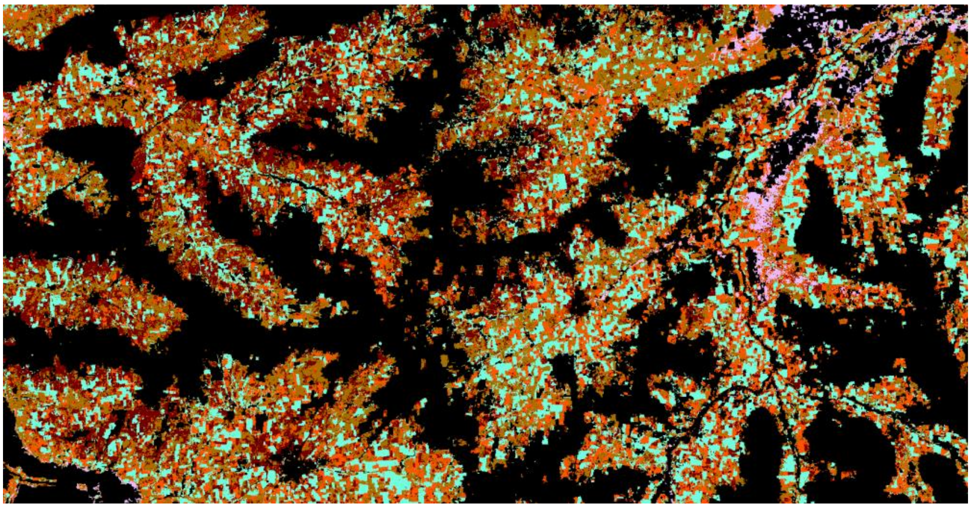


Figure 5-46: Close-up of the 2017 crop map for the main crop types including the cropland mask the over the demo-site Mali - region of Koutiala.

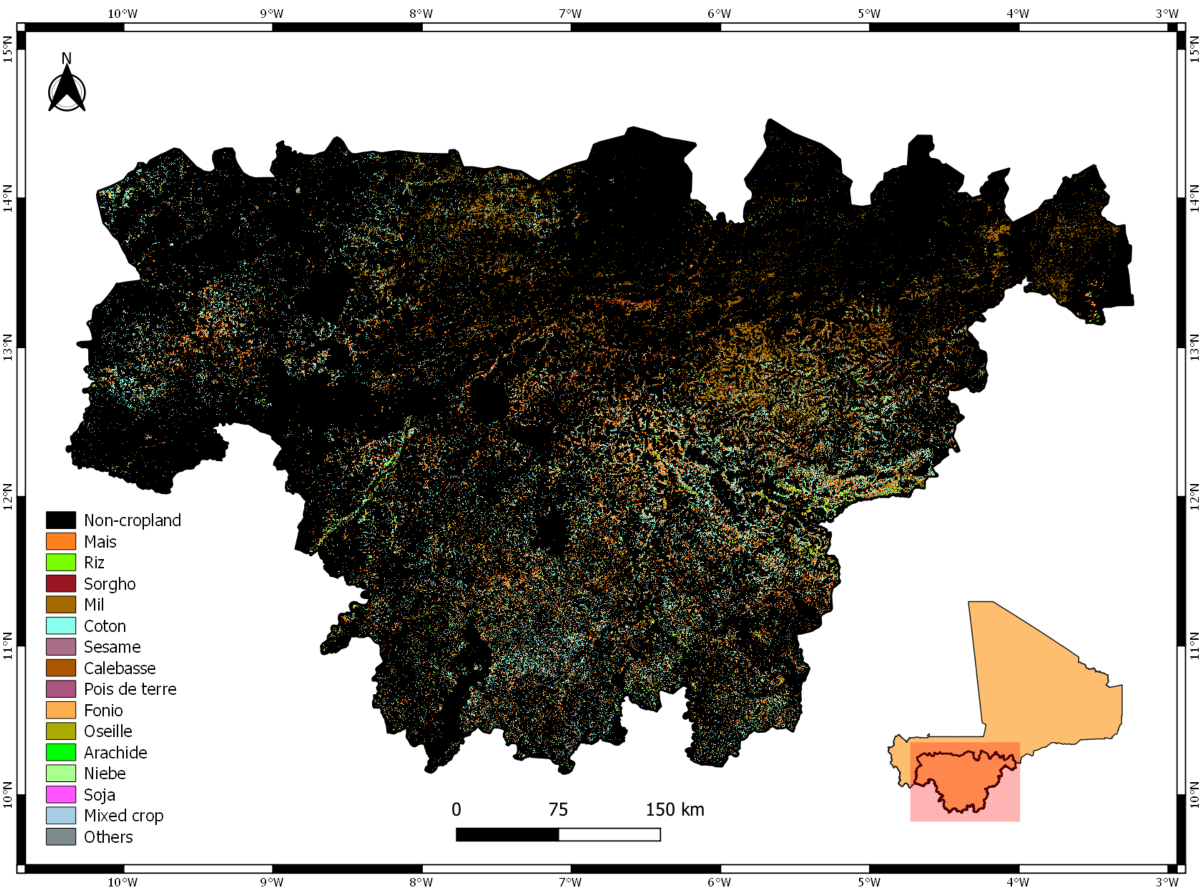


Figure 5-47: 10-meter crop type map over south of Mali for 2018 growing season, obtained at the end of the season using S2 and L8 data.

5.2.5 Crop type map of Demo-Site Western Cape Province (South-Africa)

This chapter contains information about the cropland mask and the crop type map for 2017 winter grain season.

5.2.3.1 Cropland mask

Firstly, the overall accuracy of the cropland mask has been estimated at 0,80 using 95 % of the available in situ dataset Figure 5-48. The OA is significantly lower than for the demo-site Mali but the Fscore of cropland is better here (0,88) than for Mali (0,8), probably because of the amount of in situ crop observation available. A possible option to improve these performances is to define a stratification.

Curiously, the mask accuracy is also slightly lower than the Sen2-Agri demonstration results for the same area reaching an OA of 0,85 and Fscore of 0,88 and 0,78 for cropland and cropland respectively (Defourny et al., 2019).

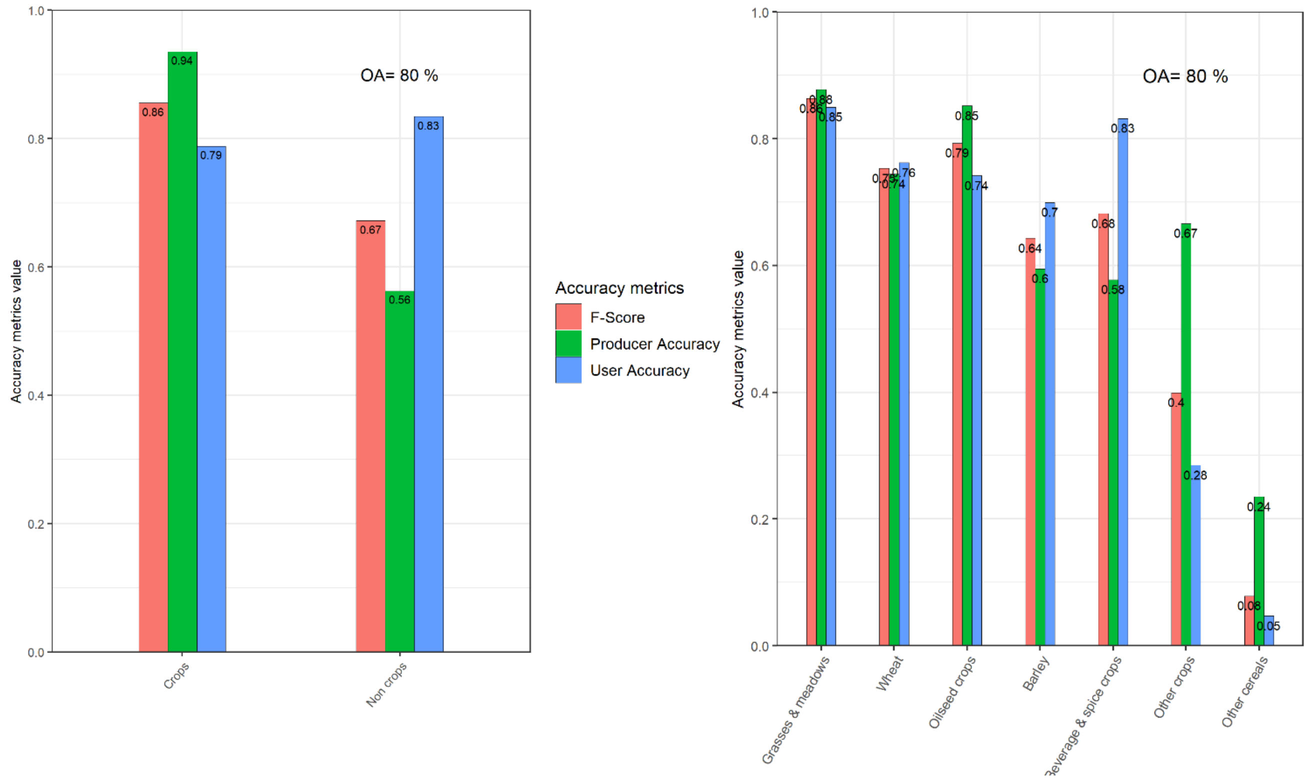


Figure 5-48: Cropland mask and crop type map performances for the 2017 winter grain season in the Western Cape Province.Cropland mask performances for the 2017 winter grain season in the Western Cape Province.

The Figure 5-49 shows the crop type map as produced by Sen2-Agri v. 2.0 system. The crop types have been further regrouped in 7 classes and provide an overall accuracy of 0,80. This is very close to the Sen2-Agri demonstration results (0,81) obtained for 4 crops only. The combination of the “other cereals” class with another one does not improve the quantitative results. The complete confusion matrix is reported at the Table 5-44. It is clear that further investigation will proceed with some stratification in order to improve the accuracy metrics and make it more useable. Last but not least these results are better than the initial one described in the WP33 report.

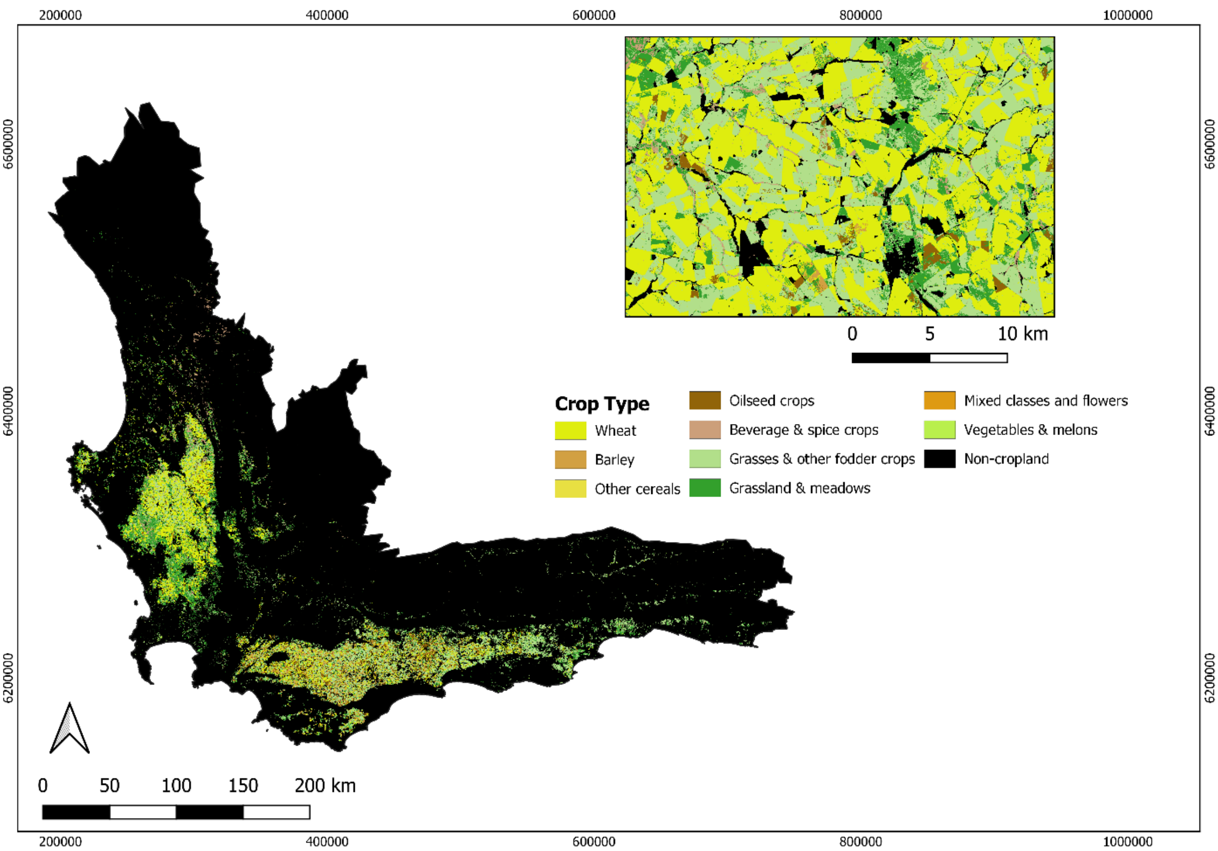


Figure 5-49: Crop type map of the Western Cape Province (South-Africa) for the the 2017 winter grain season as derived from Sentinel-2 and Landsat-8 time series. The initial 9 classes are mapped here while the validation results are presented after regrouping in 7 crop types.

Table 5-44: Confusion matrix for the 2017 crop type map derived from Sentinel-2 time series for the winter grain season.

	CRT 2018	Barley	Beverage & spice crops	Grasses & meadows	Other crops	Oilseed crops	Other cereals	Wheat	Totals	User's Accuracy
Barley	1	**4840340**	3508	563742	196	40842	872	1469432	**6918932**	70,0
Beverage & spice crops	2	281	**4068251**	779110	4998	2948	3852	29165	**4888605**	83,2
Grasses & meadows	4 + 5	1021708	2449662	**52054739**	96591	634575	76459	4917523	**61251257**	85,0
Other crops	6 + 14	10878	134632	479566	**267717**	12870	5979	28102	**939744**	28,5
Oilseed crops	8	112169	76136	1164868	7060	**5398498**	7944	510514	**7277189**	74,2
Other cereals	9	81635	64344	311633	7369	29797	**35861**	229034	**759673**	4,7
Wheat	16	2067898	246412	3948887	17888	214789	21593	**20916771**	**27434238**	76,2
Totals		**8134909**	**7042945**	**59302545**	**401819**	**6334319**	**152560**	**28100541**	**109469638**	
Producer's accuracy		59,5	57,8	87,8	66,6	85,2	23,5	74,4		**80,01**

5.3 Prototype Specifications

This section provides a description of the dataset properties and metadata for the implemented prototypes, also referring to "*P44.2a - Data Sets of Crop Area and Crop Status/Parameters Products*".

The raster products of the ECoLaSS prototypes are delivered as GeoTIFF (*.tif) with world file (*.tfw), pyramids (*.ovr), attribute table (*.dbf) and statistics (*.aux.xml), enabling an instant illustration and analysis of the products within Geographic Information System (GIS) software. Each product is accompanied with a product-specific colour table (*.clr) and INSPIRE-compliant metadata in XML format, and includes the probability layer as an additional band in the raster.

Metadata are provided together with the products as INSPIRE-compliant XML files according to the EEA Metadata Standard for Geographic Information (EEA-MSGI). EEA-MSGI has been developed by EEA to meet needs and demands for inter-operability of metadata. EEA's standard for metadata is a profile of the ISO 19115 standard for geographic metadata and contains more elements than the minimum required to comply the INSPIRE metadata regulation. Detailed conceptual specifications on EEA-MSGI and other relevant information on metadata can be found at http://www.eionet.europa.eu/gis.

The consortium has developed a standardised and harmonised product file naming convention for all prototypes produced as part of ECoLaSS based on the file naming convention of the CLMS High Resolution Layers. This file naming convention has been applied to all raster prototypes and associated reference files and is documented in the Deliverables of Task 4.

The naming convention consists of the following 7 descriptors:

THEME YEAR RESOLUTION EXTENT EPSG TYPE VERSION
as follows:

THEME

3 letter abbreviation for main products (DLT (dominant leaf type), TCC (tree cover change), GRA (grassland), IMD (imperviousness degree), IMC (imperviousness change classified), CRT (crop type), CRM (crop mask) and NLC (new land cover products).

REFERENCE YEAR

2017 in four digits; change products in four digits (e.g. 1517)

RESOLUTION

Four-digit (020m and 010m)

EXTENT

2-digit code for demonstration-sites (CE (central), NO (north), WE (west), SW (southwest), SE (southeast), SA (South Africa), ML (Mali))

EPSG

5-digit EPSG code (geodetic parameter dataset code by the European Petroleum Survey Group) "03035" for the European LAEA projection

TYPE

prototype

VERSION

3-digit code "v01"

EXAMPLE:

"CRM_2017_010m_CE_03035_prototype_v01.tif" stands for: Crop Mask, 2017 reference year, 10m, Demonstration-site Central, European projection (EPSG: 3035), prototype, version 01

"CRT_2017_010m_CE_03035_prototype_v01.tif" stands for: Crop Type, 2017 reference year, 10m, Demonstration-site Central, European projection (EPSG: 3035), prototype, version 01

The following six prototypes files as part of ***D14.4 - P44.2b – Data Sets of Crop Area and Crop Status/Parameters Products (Issue 2)*** were submitted:
The specifications of the product created by the SIRS group for the West demonstration site (French part) for the year 2016/2017 are listed in Table 5-45 to Table 5-47 and Table 5-50 to Table 5-52.

- **CRM_2016_010m_WE_03035_prototype_v01.tif**
- **CRT_2016_010m_WE_03035_prototype_v01.tif**
- **CRM_2017_010m_WE_03035_prototype_v01.tif**
- **CRT_2017_010m_WE_03035_prototype_v01.tif**

Furthermore, one product was produced by the UCL group over the West demonstration site (Belgium part) for 2017/2018. This product is conformed to the specifications detailed in Table 5-48 with associated colour palette listed in Table 5-49/Table 5-56.

- **CRT_2017_010m_WE_03035_prototype_v01.tif**
- **CRT_2018_010m_WEflan_03035_prototype_v01.tif**
- **CRT_2018_010m_WEwall_03035_prototype_v01.tif**

Product specifications for classification on MALI/SA sites are listed in Table 5-53 with associated colour palettes of Table 5-54.

- **CRT_2017_010m_ML_32630_prototype_v01.tif**
- **CRM_2018_010m_ML_32629_prototype_v01.tif**
- **CRT_2018_010m_ML_32629_prototype_v01.tif**
- **CRM_2017_010m_SA_32734_prototype_v01.tif**
- **CRT_2017_010m_SA_32734_prototype_v01.tif**

Detailed specifications for 10m CRM and CRT status layers of the demonstration site Central (produced by GAF) are listed in Table 5-55, Table 5-56, Table 5-57 and Table 5-58.

- **CRM_2017_010m_CE_03035_prototype_v01.tif**
- **CRM_2018_010m_CE_03035_prototype_v01.tif**
- **CRT_2017_010m_CE_03035_prototype_v01.tif**
- **CRT_2018_010m_CE_03035_prototype_v01.tif**

Table 5-45: Product specifications for CRT_2016_010m_WE_03035_prototype_v01.tif and for CRM_2016_010m_WE_03035_prototype_v01.tif

Products	
Cropland types 2016 – CRT_2016_10m	Cropland mask 2016 – CRT_2016_10m
Extent Demo site West	
Geometric resolution Pixel resolution 10m, grid to fully conform to the EEA Reference Grid.	
Coordinate Reference System European ETRS89 LAEA projection	
Geometric accuracy (positioning scale) Less than one pixel, according to the quality report on S-2 products	
Thematic accuracy 77% of overall accuracy	
Data type 8bit unsigned Raster, compressed with LZW	
Minimum mapping unit (MMU) One pixel (10m)	
Necessary attributes Raster value, count, class name,	
Raster coding (Thematic pixel values)	
Cropland types 2016 – CRT_2016_10m 0: Other Land cover (settlements, forest, grassland and water) 11: Winter Wheat 31: Winter Barley 40: Spring Cereals 61: Peas 71: Winter Rape 81: Maize 91: Agrarian grassland 101: Beets 131: Potatoes 141: Fallow 151: Linen 161: Chicory 254: unclassifiable (no satellite image available, or clouds, shadows, or snow) 255: outside Area	Cropland mask 2016 – CRM_2016_10m 0: Non-cropland 1: Cropland mask 254: unclassifiable (no satellite image available, or clouds, shadows, or snow) 255: outside Area
Metadata XML metadata files are to be produced according to INSPIRE metadata standards	
Delivery format GeoTIFF	

Table 5-46: Color palette for CRT_2016_010m_WE_03035_prototype_v01.tif

Cropland type 2016 – CRM_2016_10m					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0	Other Land Cover	240	240	240	
11	Winter Wheat	195	160	215	
31	Winter Barley	115	70	165	
40	Spring Cereals	255	125	125	
61	Peas	0	170	135	
71	Winter Rape	255	255	0	
81	Maize	255	175	0	
91	Agrarian Grassland	150	230	0	
101	Beets	115	0	75	
131	Potatoes	255	215	100	
141	Fallow	75	115	0	
151	Linen	115	75	0	
161	Chicory	205	105	155	
254	unclassifiable (no satellite image available, or clouds, shadows, or snow)	153	153	153	
255	outside Area	0	0	0	

Table 5-47: Color palette for CRM_2016_010m_WE_03035_prototype_v01.tif

Cropland mask 2016 – CRM_2016_10m					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0	Non- Cropland	240	240	240	
1	Cropland Mask	245	225	45	
254	unclassifiable (no satellite image available, or clouds, shadows, or snow)	153	153	153	
255	outside Area	0	0	0	

Table 5-48: Product specifications for CRT_2017_010m_WE_03035_prototype_v01.tif (Belgium part)

Products
Cropland types 2017 : CRT_2017_010m_WE_03035_prototype_v01.tif
Extent Demo site West
Geometric resolution Pixel resolution 10m, grid to fully conform to the EEA Reference Grid.
Coordinate Reference System European ETRS89 LAEA projection
Geometric accuracy (positioning scale) Less than one pixel, according to the quality report on S-2 products
Thematic accuracy More than 90% of Overall Accuracy
Data type 8bit unsigned Raster, compressed with LZW
Minimum mapping unit (MMU) One pixel (10m)
Necessary attributes Raster value, count, class name,
Raster coding (Thematic pixel values)
Cropland types 2017 – CRT_2017_10m 0: Other Land cover 1: Grassland 2: Winter Wheat 3: Maize 4: Winter Barley 5: Sugar Beet 6: Spelt 7: Rapeseed 8: Potato 9: Flax 10: Summer Oat 11: Other Vegetables 12: Other Fodder 13: Peas 14: Summer Barley 15: Summer Cereals 16: Chicory 17: Winter Cereals 18: Summer Wheat 19: Hemp 20: Green Beans 21: Onions 22: Fruit Crops 23: Other Crops 24: Grape Vines
Metadata XML metadata files are to be produced according to INSPIRE metadata standards
Delivery format GeoTIFF

Table 5-49: Color palette for CRT_2017_010m_WE_03035_prototype_v01.tif (Belgium part)

Cropland type 2016 – CRM_2016_10m					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0	Other Land Cover	240	240	240	
1	Grassland	150	230	0	
2	Winter Wheat	195	160	215	
3	Maize	255	175	0	
4	Winter Barley	115	70	165	
5	Sugar Beet	115	0	75	
6	Spelt	91	91	180	
7	Rapeseed	255	255	0	
8	Potato	255	215	100	
9	Flax	115	75	0	
10	Summer Oat	255	203	203	
11	Other Vegetables	104	147	21	
12	Other Fodder	192	29	136	
13	Peas	0	170	135	
14	Summer Barley	164	39	39	
15	Summer Cereals	255	125	125	
16	Chicory	205	105	155	
17	Winter Cereals	245	227	255	
18	Summer Wheat	220	79	79	
19	Hemp	210	135	35	
20	Green Beans	0	103	83	
21	Onions	193	150	27	
22	Fruit Crops	108	162	194	

23	Other Crops	52	80	0	
24	Grape Vines	23	86	123	

Table 5-50: Product specifications for CRT_2017_010m_WE_03035_prototype_v01.tif (France part)

Products
Cropland types 2017: CRT_2017_010m_WE_03035_prototype_v01.tif
Extent Demo site West
Geometric resolution Pixel resolution 10m, grid to fully conform to the EEA Reference Grid.
Coordinate Reference System European ETRS89 LAEA projection
Geometric accuracy (positioning scale) Less than one pixel, according to the quality report on S-2 products
Thematic accuracy More than 90% of Overall Accuracy
Data type 8bit unsigned Raster, compressed with LZW
Minimum mapping unit (MMU) One pixel (10m)
Necessary attributes Raster value, count, class name
Raster coding (Thematic pixel values)
Cropland types 2017 – CRT_2017_10m 1: maize 2: summer_cereals 3: winter_cereals 5: sugarbeet 6: potatoes 7: root_crops 8: vegetables 10: soybeans 11: fruit_trees 12: grape_vines 15: sunflower 16: oleaginous 17: rapeseed 18: grassland 19: fodder
Metadata XML metadata files are to be produced according to INSPIRE metadata standards
Delivery format GeoTIFF

Table 5-51: Color palette for CRT_2017_010m_WE_03035_prototype_v01.tif (France part)

Cropland type 2017 – CRM_2017_10m					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0	No Cropland	240	240	240	
1	Maize	158	215	194	
2	Summer Cereals	102	119	205	
3	Winter Cereals	0	92	230	
5	Sugar Beets	205	102	153	
6	Potatoes	255	215	100	
7	Root Crops	255	170	0	
8	Vegetables	255	157	188	
10	Soy Beans	205	137	102	
11	Fruit Trees	115	0	75	
12	Grape Vines	137	137	68	
15	Sunflower	255	0	0	
16	Oleaginous	137	90	68	
17	Rape Seed	115	178	255	
18	Grassland	180	215	158	
19	Fodder	240	115	95	

Table 5-52: Color palette for CRM_2017_010m_WE_03035_prototype_v01.tif (France part)

CRM_2017_010m_WE_03035_prototype_v01					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0	Non- Cropland	240	240	240	
1	Cropland Mask	245	225	45	
255	outside Area	0	0	0	

Table 5-53: Product specifications for CRT_2017_010m_ML_32630_prototype_v01.tif

Products
Cropland types 2017 – CRT_2017_10m CRT_2017_010m_ML_32630_prototype_v01.tif
Extent Demo site MALI
Geometric resolution Pixel resolution 10m, grid to fully conform to the EEA Reference Grid.
Coordinate Reference System WGS84 – UTM zone 29/30 South
Geometric accuracy (positioning scale) Less than one pixel, according to the quality report on S-2 products
Thematic accuracy >90% of overall accuracy
Data type 8bit unsigned Raster, compressed with LZW
Minimum mapping unit (MMU) One pixel (10m)
Necessary attributes Raster value, class name
Raster coding (Thematic pixel values)
0: Non-Cropland 1: Millet 2: Sorghum 3: Cotton 4: Maize 5: Other Crops 6: Groundnuts 7: Rice
Metadata XML metadata files are to be produced according to INSPIRE metadata standards
Delivery format GeoTIFF

Table 5-54: Color palette for CRT_2017_010m_ML_32630_prototype_v01.tif

Cropland type 2016 – CRM_2016_10m					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0	Non-Cropland	0	0	0	
1	Millet	164	89	20	
2	Sorghum	139	26	26	
3	Cotton	151	255	255	
4	Maize	255	76	0	
5	Other Crops	163	163	163	
6	Groundnuts	162	205	90	
7	Rice	95	158	160	

Table 5-55: Product Specifications and color palette for CRT_2017_010m_CE_03035_prototype_v01.tif

Crop Type 10m	**Acronym** CRT	**Product category** Primary Status Layer
Reference year 2017		
Extent Demonstration site Central		
Geometric resolution Pixel resolution 10m x 10m, fully conform to the EEA reference grid.		
Coordinate Reference System European ETRS89 LAEA projection		
Geometric accuracy (positioning scale) Less than one pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy 85%		
Data type 8bit unsigned Raster, with LZW compression		
Minimum mapping unit (MMU) Pixel-based (no MMU)		
Necessary attributes Raster value, count, class name, area (in km2), area percentage (taking outside area not into account)		
Raster coding (Thematic pixel values)		

0: Non-cropland
1: Agrarian Grassland
3: Fallow
5: Fruit Trees
8: Legume
9: Maize
10: Others
11: Potatoes
13: Strawberries
14: Sugar Beets
15: Summer Crop
16: Summer Rape
17: Sunflowers/Topinambour
18: Vegetables
19: Winegrowing
20: Winter Crop
21: Winter Rape
255: Outside area

Metadata
XML metadata files are to be produced according to INSPIRE metadata standards

Delivery format
GeoTIFF

Colour Table
ArcGIS *.clr format

Class Code	Class Name	Red	Green	Blue	
0	Non-cropland	240	240	240	
1	Agrarian Grassland	0	82	0	
3	Fallow	97	184	255	
5	Fruit Trees	171	63	204	
8	Legume	112	81	75	
9	Maize	255	255	0	
10	Others	38	38	38	
11	Potatoes	196	188	179	
13	Strawberries	255	46	46	
14	Sugar Beets	0	247	206	
15	Summer Crop	255	0	255	
16	Summer Rape	255	149	0	

17	Sunflowers/Topinambour	255	250	204	
18	Vegetables	128	255	0	
19	Winegrowing	250	145	180	
20	Winter Crop	204	179	102	
21	Winter Rape	114	115	80	
255	Outside Area	0	0	0	

Table 5-56: Product Specifications and color palette for CRT_2018_010m_CE_03035_prototype_v01.tif

Crop Type 10m	**Acronym** CRT	**Product category** Primary Status Layer
Reference year 2018		
Extent Demonstration site Central		
Geometric resolution Pixel resolution 10m x 10m, fully conform to the EEA reference grid.		
Coordinate Reference System European ETRS89 LAEA projection		
Geometric accuracy (positioning scale) Less than one pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy 85%		
Data type 8bit unsigned Raster, with LZW compression		
Minimum mapping unit (MMU) Pixel-based (no MMU)		
Necessary attributes Raster value, count, class name, area (in km^2)		
Raster coding (Thematic pixel values) 0: No Cropland 1: Winter Wheat 2: Winter Barley 3: Winter Rye 5: Summer Wheat 6: Summer Barley 8: Summer Oats 9: Maize		

10: Vegetables
11: Peas + Beans
12: Lentils
14: Soya Beans
15: Sunflowers
16: Rape Seed
17: Oleaginous + Fibre Crops
18: Potatoes
19: Beet Crops
20: Temporary Grassland + Fodder Crops
21: Wine Growing
22: Fruit Trees + Berries
255: Outside Area

Metadata
XML metadata files are to be produced according to INSPIRE metadata standards

Delivery format
GeoTIFF

Colour Table
ArcGIS *.clr format

Class Code	Class Name	Red	Green	Blue	
0	No Cropland	255	255	255	
1	Winter Wheat	0	92	230	
2	Winter Barley	115	178	255	
3	Winter Rye	190	210	255	
5	Summer Wheat	102	119	205	
6	Summer Barley	122	142	245	
8	Summer Oats	158	170	215	
9	Maize	158	215	194	
10	Vegetables	137	68	101	
11	Peas + Beans	205	102	153	
12	Lentils	214	157	188	
14	Soya Beans	205	137	102	
15	Sunflowers	255	0	0	
16	Rape Seed	255	235	175	
17	Oleaginous + Fibre Crops	137	90	68	

18	Potatoes	255	211	127	
19	Beet Crops	255	170	0	
20	Temporary Grassland + Fodder Crops	180	215	158	
21	Wine Growing	137	137	68	
22	Fruit Trees + Berries	115	0	76	
255	Outside Area	0	0	0	

Table 5-57: Product Specifications and color palette for CRM_2017_010m_CE_03035_prototype_v01.tif

Crop Mask 10m	**Acronym** CRM	**Product category** Primary Status Layer
Reference year 2017		
Extent Demonstration site Central		
Geometric resolution Pixel resolution 10m x 10m, fully conform to the EEA reference grid.		
Coordinate Reference System European ETRS89 LAEA projection		
Geometric accuracy (positioning scale) Less than one pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy 90%		
Data type 8bit unsigned Raster, with LZW compression		
Minimum mapping unit (MMU) Pixel-based (no MMU)		
Necessary attributes Raster value, count, class name, area (in km2), area percentage (taking outside area not into account)		
Raster coding (Thematic pixel values) 0: Non-cropland 1: Cropland 255: Outside area		
Metadata XML metadata files are to be produced according to INSPIRE metadata standards		
Delivery format GeoTIFF		
Colour Table ArcGIS *.clr format		

Class Code	Class Name	Red	Green	Blue	
0	Non-cropland	240	240	240	
1	Cropland	0	82	0	
255	Outside Area	0	0	0	

Table 5-58: Product Specifications and color palette for CRM_2017_010m_CE_03035_prototype_v01.tif

Crop Mask 10m	Acronym CRM	Product category Primary Status Layer
Reference year 2018		
Extent Demonstration site Central		
Geometric resolution Pixel resolution 10m x 10m, fully conform to the EEA reference grid.		
Coordinate Reference System European ETRS89 LAEA projection		
Geometric accuracy (positioning scale) Less than one pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy 90%		
Data type 8bit unsigned Raster, with LZW compression		
Minimum mapping unit (MMU) Pixel-based (no MMU)		
Necessary attributes Raster value, count, class name, area (in km2), area percentage (taking outside area not into account)		
Raster coding (Thematic pixel values) 0: Non-cropland 1: Cropland 255: Outside area		
Metadata XML metadata files are to be produced according to INSPIRE metadata standards		
Delivery format GeoTIFF		
Colour Table ArcGIS *.clr format		

Class Code	Class Name	Red	Green	Blue	
0	No Cropland	243	243	243	
1	Cropland	230	125	0	
255	Outside Area	0	0	0	

In addition to the prototype layers, probability layers are provided as a by-product. The additional band in the *.tiff-file serves as one of the accuracy parameters that are described in detail in the WP33 final report [AD07] and range from 0 to 100%. The higher the percentage the higher the probability is that the respective pixel belongs to the depicted class. In this manner, the probability band depicts the error map at pixel level. Further, areas that are excluded by the referring mask get the value "101". An overview of the probabilities' colour palette is given in Table 5-59.

Table 5-59: Colour palette for the probability layers

Probabilities					
Class Code	**Class Name**	**Red**	**Green**	**Blue**	
0-100	Probabilities 0-100%	-	-	-	
101	Areas excluded by binary Mask	128	128	128	
255	Outside Area	0	0	0	

6 Conclusion and Outlook

As reported in the Deliverable D3.1: "D21.1a – Service Evolution Requirements Report" of the WP 21: "WP 21 – Assessment of Service Evolution Requirements", the agriculture is one of the two thematic gaps in the current Copernicus Land Service portfolio. As clearly expressed in this report, the most frequently voiced new service was a pan-European Agricultural Service. As expected, this service are potentially interlinked with the phenology service as a key target is to capture the phenology of the different crop type also mapped by Sentinel data. While a preliminary document has been prepared by JRC and circulated to various partners, the "Concept note for a Copernicus Agricultural Service" also prepared by JRC is under examination by the Copernicus office and not available yet for the ECoLaSS partners. Therefore the targeted prototype at this stage is the crop type mapping based on S1 and S2 time series.

The Crop Mask (CRM) and the Crop Type Map (CRT) prototypes have been extensively tested in this WP 44 for three different Demonstration site (West, Central and Mali) to encompass a large diversity of agrosystems and Earth Observation conditions (cloudiness in particular). The harmonization of the protocols across the Demonstration sites concerned mainly the type of features, the Random Forest classifier and the accuracy assessment report. In addition for the European sites, the calibration and validation datasets were derived from the respective LPIS layers. At the opposite, the study in the Mali site relied on a comprehensive dataset collected on the field by Malian partner and quality controlled by UCL.

The Crop Mask (CRM) products obtained for the three Demonstration sites provided very satisfactory results with an overall accuracy far beyond 90 % (98 % for site West (France) from S-2 only, 97% (F1-Score 0.97) for site Central from S-1 and S-2, 97% (F1-Score 0.89) for site Mali from S-2). These high accuracy including in the Malian smallholder cropping systems demonstrated the maturity level of this prototype. Nevertheless, few elements should be further investigated: the optimization of the number of features (for the site Central 28 features out of the 1246 computed were found efficient), the stratification to deal with uneven spatial distribution of the calibration data, and some specific confusion like the grassland in Germany and the bottoms of valley in Mali.

The Crop Type Map (CRT) products obtained for three Demonstration sites provided promising results with the overall accuracy ranging from 64 % in Mali (20 tiles - 6 classes) and 87 % (4 tiles together – 15 crop classes and 4 other classes) in France both using only S-2 to 89 % in Central (9 tiles together – 16 classes) to 92 % in Belgium (for the best tile – 24 classes) both combining S-2 and S-1-derived features. These results could be further improved by masking using the CRM the CRT products. The F1-score values for the different crops show a very large range with some crops with F1-score beyond 0.8 (dominant crops in the agricultural landscape) but most of them below 0.8 down to 0.3 for marginal classes. Furthermore, the performance metrics for the different sites are apparently also driven by the fraction of validation corresponding to the (agrarian) grassland class. In France, the poor discrimination of this class reduces the metric value while the large fraction of grassland in the validation dataset for Belgium improves the metric value.

The development of these two essential CRM and CRT prototypes further continued over the three Demonstration sites (West, Central and Mali) and was further extended to the Demonstration site South Africa. First the focus is on the improvement of cropland – grassland discrimination and then the year-to-year variation of the CRM prototype products will be assessed for several sites. For the CRT prototype products, the crop type discrimination (F1_score) will be further improved thanks to (i) an advanced exploitation of the S-1 time series (possibly using an object-based approach), (ii) a tuning of the crop type distribution in the calibration dataset, the use of stratification to improve the tile mosaicking process, and (iii) a complementarity of the CRT product with the CRM product or the GRA HRL in the European sites.

References

Breiman, L. (2001). Random Forests. *Machine Learning, 4*(1), 5-32.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., et al.. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103(0):7 – 27. Global Land Cover Mapping and Monitoring.

Defourny P., Bontemps S., Bellemans N., et al. (2018). Sentinel-2 near real-time agriculture monitoring at parcel level thanks to the Sen2-Agri automated system: methods and demonstration from local to national scale for cropping systems around the world. Remote Sensing of Environment (under revision).

Esch, Thomas and Üreyen, Soner and Zeidler, Julian and Hirner, Andreas and Asamer, Hubert and Metz-Marconcini, Annekatrin and Tum, Markus and Böttcher, Martin and Kuchař, Štěpán and Svaton, Vaclav and Marconcini, Mattia (2018) *Exploiting Big Earth Data from Space – First Experiences with the TimeScan Processing Chain.* Big Earth Data, pp. 36-55. Taylor & Francis. DOI: 10.1080/20964471.2018.1433790 ISSN 2096-4471

Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS Journal of Photogrammetry and Remote Sensing, 82, 83-92. doi:doi.org/10.1016/j.isprsjprs.2013.04.007

Hagolle, O., Dedieu, G., Mougenot, B., Debaecker, V., Duchemin, B., & Meygret, A. (2008). Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: Application to Formosat-2 images. Remote Sensing of Environment, 112(4), 1689-1701.

Hagolle, O., Huc, M., Villa Pascual, D., Dedieu, G., 2010. A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENµS, LANDSAT and SENTINEL-2 images. Remote Sens. Environ. 114, 1747–1755.

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., Defourny, P., et al.. 2015. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. Remote Sensing, 7(9):12356–12379.

JECAM (2017). http://jecam.org/documents/

Mathieu, R., Pouget, M., Cervelle, B. and Escadafal, R. (1998). Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment. Remote Sensing of Environment 66, 17–28.

Metzger, M.J., Bunce, R.G.H., Jongman, R.G.H., Mücher, C.A., Watkins, J.W., 2005. A climatic stratification of the environment of Europe. Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2005) vol. 14, pp. 549–563.

Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photgrammetry and Remote Sensing, 67, 93-104.

Rouse Jr, J., Haas, R., and Schell, J. (1974). Monitoring vegetation systems in the great plains with erts. NASA Special Publication, 351, 309

Shang, N., & Breiman, L. (1996). Distribution based trees are more accurate. *Ionosphere, 33*(2), 351.

Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8 (2), 127–150.

Valero, S., Morin, D., Inglada, J., Sepulcre, G., Arias, M., Hagolle, O. & Koetz, B. (2016). Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. Remote Sensing, 8(1), 55.

Vintrou, E., Desbrosse, A., Bégué, A. and Traore, P. (2012). Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. In: International Journal of Applied Earth Observation and Geoinformation 14(1):83-93. DOI: 10.1016/j.jag.2011.06.010