

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



D13.2

"D43.1b - Prototype Report: Improved Permanent Grassland"

Issue/Rev.: 2.0

Date Issued: 05.12.2019

submitted by:



in collaboration with the consortium partners:



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:

The contents of this document are the copyright of GAF AG and Partners. It is released by GAF AG on the condition that it will not be copied in whole, in section or otherwise reproduced (whether by photographic, reprographic or any other method) and that the contents thereof shall not be divulged to any other person other than of the addressed (save to the other authorised officers of their organisation having a need to know such contents, for the purpose of which disclosure is made by GAF AG) without prior consent of GAF AG. Nevertheless, single or common copyrights of the contributing partners remain unconditionally unaffected.

DOCUMENT RELEASE SHEET

	NAME, FUNCTION	DATE	SIGNATURE
Author(s):	Heinz Gallaun (JR) Petra Miletich (JR) Mathias Schardt (JR) Martin Puhm (JR) Annekatrin Metz-Marconcini (DLR) Benjamin Leutner (DLR) Christophe Sannier (SIRS) Sophie Villerot (SIRS) Alice Lherould (SIRS)	28.11.2019	
Review:	Petra Miletich (JR) Klaus Granica (JR)	02.12.2019	
Approval:	Eva Sevillano Marco (GAF)	05.12.2019	
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	17.07.2018	63	First issue of WP43 Deliverable
2.0	05.12.2019	92	Second issue of WP43 Deliverable. Minor updates in Sections 1-3 introducing demosites Central and South-East and major updates in all sections from phase 2 implementation in all three demosites (status change prototypes).

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	D6.1: D31.1a - Methods Compendium: Sentinel-1/2/3 Integration Strategies, (Issue 1), Issued: March 2018
AD06	D7.2: D32.1b - Methods Compendium: Time Series Preparation, (Issue 2), Issued: 15.05.2019
AD07	D8.2: D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification (Issue 2), Issued: December 2019.
AD08	D9.2: D34.1b - Methods Compendium: Time Series Analysis for Change Detection (Issue 2), Issued: 04.12.2019.
AD09	D3.2: D21.1b - Service Evolution Requirements Report (Issue 2), Issued: December 2019.
AD10	D2.3: D12.1c - DWH use for 2017/2018/2019 (Issue 3), Issued: 18.10.2019.
AD11	D16.3: D51.1d - Stakeholder Consultation Report (Issue 4), Issued: 31.10.2019.
AD12	D10.2: D35.1b - Methods Compendium: Time Series Consistency for HRL Product (Incremental) Updates (Issue 2), Issued: December 2019.
AD13	D1.5: D11.3b - Interim Progress Report (Issue 2), Issued: 09.07.2019.
AD14	D14.2: D44.1b - Prototype Report: Crop Area and Crop Status/Parameters (Issue 2), 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 optical and Sentinel-1 Synthetic Aperture Radar (SAR) data, complemented by Medium-Resolution (MR) Sentinel-3 optical and Landsat-8 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 report constitutes a prototype report of the work package (WP) 43 “Improved Permanent Grassland Identification”, directly building on the methods and processing lines developed by Task 3, especially WP32 “Time Series Preparation”, WP33 “Time Series Analysis for Thematic Classification” and WP34 “Time Series Analysis for Change Detection”. It demonstrates a grassland prototype for future potential Copernicus Land Monitoring Service (CLMS) products of the pan-European Component, which is improved compared to previous approaches such as e.g. the Copernicus HRL Grassland 2015 and has tackled improvement of the current HRL Grassland 2018.

The objective of this WP is to develop a framework for an improved identification of grassland areas using Sentinel time series with the aim to develop a prototype of a European HR Grassland Layer with high thematic accuracy, optical and SAR data integration, increased spatial resolution and high automation level as well as (mid-term) increased thematic content. Together with the outcomes of the other WPs of ECoLaSS Task 3 (Automated High Data Volume Processing Lines) and Task 4 (Thematic Proof-Of-Concept), it constitutes a basis for the investigation activities of Task 5 (Operationalisation Framework).

Chapter 1 of the document explains the purpose and objectives of WP43. Chapter 2 presents the background of Copernicus Grassland monitoring needs and the summary of related requirements. Chapter 3 gives a short summary and description of the demonstration sites. In Chapter 4, an overview of the applied time series analysis methods is given based on the results and recommendations of WP33 “Time Series Analysis for Thematic Classification” and WP34 “Time Series Analysis for Change Detection”. It reviews specifically the multi-sensor data integration methods, the usage of multi-temporal time series metrics, the random forest classification approach and the validation analysis procedure. Based on these reviews, Chapter 5 details the prototype implementation approaches, including a description of the input data, pre-processing lines, the experimental setup, and the final validation of the prototype.

The achievements compared to the existing HRL grassland 2015 can be summarized by:

- **improved level of automation** to allow a faster production and shorter monitoring intervals (e.g. continuous monitoring with yearly updates)
- Improve the **thematic classification accuracy**
- Fully exploit **optical Sentinel-2 and SAR Sentinel-1 time series** instead of using pre-selected, best-suited optical EO data scenes

- Design a fully **integrated SAR/optic** time series data analysis to benefit from the multi-sensor characteristics
- Provide a **seamless, wall-to-wall product** (e.g. no cloud cover gaps)
- Provide information on **mowing intensity**, which can be used as valuable input in downstream applications such as for example related to the assessment of grassland management intensity
- Improve the status layer's detail from 20m **spatial resolution to 10m**
- Provide a **change detection** approach to detect grassland increase and decrease

The successful application demonstrates the potential for a future large area rollout of the developed methods.

The ECoLaSS project follows a two-phased approach of two times 18 months duration. The first issue of this deliverable presented preliminary results up to month 18. In the second 18-month project cycle, this second issue is published, containing all final results.

Table of Contents

1	INTRODUCTION	1
2	BACKGROUND AND SUMMARY OF REQUIREMENTS	2
3	DEMONSTRATION SITES.....	3
3.1	ECoLaSS DEMONSTRATION SITES	3
3.2	DEMONSTRATION SITE WEST FOR GRASSLAND	5
3.3	DEMONSTRATION SITE CENTRAL FOR GRASSLAND	6
3.4	DEMONSTRATION SITE SOUTH-EAST FOR GRASSLAND	7
4	OVERVIEW OF APPLIED METHODS.....	8
4.1	MULTI-SENSOR DATA INTEGRATION.....	8
4.2	MULTI-TEMPORAL SAR AND OPTICAL METRICS	9
4.2.1	Spectral Optical Indices	9
4.2.2	Multi-seasonal Statistical SAR and Optical Features	10
4.3	RANDOM FOREST CLASSIFICATION APPROACH	10
4.4	VALIDATION ANALYSIS PROCEDURE.....	11
4.4.1	Sampling Point Generation for Grassland	11
4.4.2	Response Design.....	16
4.4.3	Validation Analysis Procedure	16
5	PROTOTYPES IMPLEMENTATION	19
5.1	DATA AND PROCESSING SETUP	21
5.1.1	Input Data and Data Integration.....	22
5.1.2	Pre-Processing	24
5.1.3	Demonstration Site West	25
5.1.4	Demonstration Site Central	28
5.1.5	Demonstration Site South-East	35
5.2	PROTOTYPE VALIDATION RESULTS.....	40
5.2.2	HRL2015 Comparison with an Independent Data Source	61
5.3	PROTOTYPE SPECIFICATIONS	66
CONCLUSION AND OUTLOOK.....		75
REFERENCES.....		77
ANNEXE 1		81

List of Figures and Tables

Figure 3-1: European Demonstration Sites (modified from: EEA, 2015).	4
Figure 3-2: Overview of Demonstration site West.....	5
Figure 3-3: Overview of Demonstration site Central.	7
Figure 3-4: Overview of Demonstration site South-East.....	8
Figure 4-1: Simple random (left) and random systematic (right) sampling designs.	12
Figure 4-2: LUCAS points located on a regular grid.....	13
Figure 4-3: Number of sample points as a function of the expected error rate	14
Figure 4-4: Replicates and sub-replicates used on LUCAS grid.	14
Figure 4-5: Validation samples over the Demonstration site GRA_2017_010m_BE_03035_V0_1.	15
Figure 4-6: Confusion Matrix for Accuracy Assessment of thematic map product.	17
Figure 5-1: General workflow for grassland status mapping.	20
Figure 5-2: General workflow for grassland change detection.	21
Figure 5-3: OBB error in relation the number of S1 and S2 input features.....	26
Figure 5-4: Top 40 S1 and S2 features for the grassland discrimination.....	26
Figure 5-5: LPIS Grassland polygons (2016) based on S2 image	27
Figure 5-6: LPIS Grassland polygons (2016) based on S1 image	27
Figure 5-7: Grassland mask 2018 Central demonstration site detail,.....	30
Figure 5-8: Map-to-map change detection in Central demo site for the reference years 2015 (HRL2015 GRA)-2018 (ECoLaSS grassland mask 2018).	31
Figure 5-9: Boundary filter steps applied to the raw change detection recoded layer (b) in the map-to-map change detection (reference years 2015-2018) in the Central demonstration site.	31
Figure 5-10: Example of the combination of the raw change layer (a) with the boundary-cleaned layer (c) in Grassland change detection (2015-2018).	32
Figure 5-11: Example of the combination of the NDVI minimum threshold recoded layer (d) with the cleaned change layer (c) in Grassland change detection (2017-2018).	32
Figure 5-12: Example of the combination of the combined layer (e) with the raw change layer (a) in Grassland change detection (2017-2018).	33
Figure 5-13: Summary of the change detection approach for the Grassland reference years 2017-2018 in Central demonstration site.	33
Figure 5-14: Intensive/Extensive grassland use layer workflow.	34
Figure 5-15: Grassland mowing intensity in Central 2018.	34
Figure 5-16: Detailed view of the grassland mowing intensity layer.	35
Figure 5-17: Development of out-of-bag grassland F1 score during Gini-based backward feature selection for the predictor sets: Sentinel-1 (S1), Sentinel-2 (S2) and combined features (S1/S2).....	36
Figure 5-18: SAR + OPT grassland classification with random forest and selected features for 2017	40
Figure 5-19: LGP grassland areas 2016.....	40
Figure 5-20: World View 1 image from the 14.03.2017.	43
Figure 5-21: GRA_West_2017 (grassland in green).	43
Figure 5-22: Grassland Probability Layer 2017.....	43
Figure 5-23: Grassland Probability Layer 2018.....	43
Figure 5-24: Sample of the grassland mowing intensity map.	44
Figure 5-25: Number of valid observations per pixel for the S2 image stack used as input to the mowing event detection algorithm (beginning of week 10/2018 to end of week 44/2018) on the left. The right image shows the grassland use intensity layer 2018 for the demonstration site West.	44
Figure 5-26: World View 1 image from the 2017.05.09.....	45
Figure 5-27: World View 1 image from the 2018.10.05.	45
Figure 5-28: CVA approach. Loss (yellow), Gain (Green).	45
Figure 5-29: Grassland masks Central 2017 (up) and 2018 (down).	47
Figure 5-30: Detailed view from the Grassland identification 2018 and 2017 showing the effects of 2018 drought.	48

Figure 5-31: Classification result detail grassland mask 2018 in Central demonstration site.....	49
Figure 5-32: Grassland mask 2018 Probability layer.....	50
Figure 5-33: Consistency between the Grassland Status Layer 2018 and the Crop Mask 2018.....	51
Figure 5-34: Grassland change layers 2015-2018 and 2017-2018	53
Figure 5-35: Detailed view of the map-to-map change detection and the NDVI time series thresholding approach applied in the Central demonstration site	54
Figure 5-36: Grassland status layers of 2017 and 2018 for the demonstration site South-East.	56
Figure 5-37: Random forest grassland probability 2017 and 2018.....	57
Figure 5-38: Detected non-continuous grassland patches, due to interdispersion with rocks, shrubs and trees.....	59
Figure 5-39: 2017 status layer and change layer 2017 – 2018.....	61
Figure 5-40: Unfiltered comparison of the HRL GRA 2015 and the ECoLaSS prototype 2017.....	63
Figure 5-41: Filtered comparison of the HRL GRA 2015 and the ECoLaSS prototype 2017	64
Figure 5-42: Overview of the MEAN spectral signature of the index Clgreen for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	81
Figure 5-43: Overview of the MEAN spectral signature of the index CIRedEdge for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.	82
Figure 5-44: Overview of the MEAN spectral signature of the index EVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	83
Figure 5-45: Overview of the MEAN spectral signature of the index MCARI/OSAVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	84
Figure 5-46: Overview of the MEAN spectral signature of the index MTCI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.	84
Figure 5-47: Overview of the MEAN spectral signature of the index NBR for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.	85
Figure 5-48: Overview of the MEAN spectral signature of the index NDMI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.	86
Figure 5-49: Overview of the MEAN spectral signature of the index NDRE1 for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	87
Figure 5-50: Overview of the MEAN spectral signature of the index NDRE2 for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.	87
Figure 5-51: Overview of the MEAN spectral signature of the index NDVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.	88

Figure 5-52: Overview of the MEAN spectral signature of the index REP for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	89
Figure 5-53: Overview of the MEAN spectral signature of the index SAVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	90
Figure 5-54: Overview of the MEAN spectral signature of the index TCB for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	90
Figure 5-55: Overview of the MEAN spectral signature of the index TCG for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	91
Figure 5-56: Overview of the MEAN spectral signature of the index TCW for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.....	92

Table 3-1: Description of the selected Demonstration Sites.....	4
Table 4-1: Number of PSU.....	15
Table 5-1: ADDITIONAL VHR1 data used for the GRA prototypes over the project years 2017, 2018 and 2019 sorted by type.....	23
Table 5-2: LUCAS 2018 inclusion rules.....	24
Table 5-3: Final variables used for Grassland status layer production in demonstration site South-East .	38
Table 5-4: Internal validation results for the GRA_West_2017 product (area-weighted plausibility approach)	41
Table 5-5: Internal validation results for the GRA_West_2017 product (area-weighted blind approach). 41	41
Table 5-6: Internal validation results for the GRA_West_2018 product (area-weighted plausibility approach).	42
Table 5-7: Internal validation results for the GRA_West_2018 product (area-weighted blind approach). 42	42
Table 5-8: Internal validation results for the GRA_West_Change_2017_18 product (area-weighted plausibility approach).	45
Table 5-9: Internal validation results for the GRA_Central_2017 product (area-weighted plausibility approach).	49
Table 5-10: Internal validation results for the GRA_Central_2018 product (area-weighted plausibility approach).	50
Table 5-11: Internal validation results for the GRA_Central_Change_2017_18 product (area-weighted plausibility approach).	55
Table 5-12: Internal validation results for the GRA_South_East_2018 product (area-weighted plausibility approach).	57
Table 5-13: Internal validation results for the GRA_South_East_2018 product (area-weighted blind approach).	58
Table 5-14: Internal validation results for the GRA_South_East_2017 product (area-weighted plausibility approach).	58
Table 5-15: Internal validation results for the GRA_South_East_2017 product (area-weighted blind approach).	59
Table 5-16: Internal validation results for the GRA_South_East_Change_2017_18 product (area-weighted plausibility approach).	60
Table 5-17: Comparison between the HRL Grassland layer 2015 and the adapted VIRP 2017.....	62

Table 5-18: Accuracies for HRL Grassland using the adapted VIRP 2017 as Reference.....	62
Table 5-19: Comparative Grassland statistic (HRL 2015, ECoLaSS 2017) for the demosite West.....	62
Table 5-20: Types of omission errors for HRL GRA 2015 and the ECoLaSS prototype 2017.....	64
Table 5-21: Types of commission errors for HRL GRA 2015 and the ECoLaSS prototype 2017	65
Table 5-22: Detailed specifications for the Improved Primary Status Grassland Layers.	68
Table 5-23: Colour palette for the probability layers.....	69
Table 5-24: Detailed specifications for the New Status Layer Grassland Change.....	69
Table 5-25: Detailed specifications for the New Status Layer Grassland Change.	70
Table 5-26: Detailed specifications for the New Status Layer Grassland Mowing intensity.....	71
Table 5-27: Detailed specifications for the New Status Layer Grassland Mowing Events.....	73

Abbreviations

AGRI	Agriculture
AVHRR	Advanced Very High Resolution Radiometer
CE	Central
CI_green	Green Chlorophyll Index
CI_red_edge	Red Edge Chlorophyll Index
CLMS	Copernicus Land Monitoring Services
CORINE	Coordination of Information on the Environment
CoV	Coefficient of Variation
CRM	Crop Mask
CRT	Crop Type
DIFF	Difference Feature
DLT	Dominant Leave Type
DWH	Data Warehouse
ECoLaSS	Evolution of Copernicus Land Services based on Sentinel data
EEA	European Environment Agency
EEA39	39 European countries
EEA-MSGI	EEA Metadata Standard for Geographic Information
EEE	Entrusted European Entities
EO	Earth Observation
ESA	European Space Agency
EVI	Enhanced Vegetation Index
GC	Ground cover
GIS	Geographic Information System
GNDVI	Green Normalized Difference Vegetation Index
GRA	Grassland
GRD	Ground Range Detected
H2020	Horizon 2020
HR	High Resolution
HRL	High Resolution Layer
IACS	InVeKoS in Austria
IMC	Imperviousness degree change
IMD	Imperviousness degree
IMP	Imperviousness
INSPIRE	Infrastructure for Spatial Information in the European Community
IRECI	Inverted Red-Edge Chlorophyll Index
ISO	International Organization for Standardization
IW	Interferometric Wide Swath Mode
JECAM	Joint Experiment for Crop Assessment and Monitoring network
JRC	Joint Research Centre
LAEA	Lambert azimuthal equal-area projection
LC	Land cover
LGP	Grassland Reference Polygons (Landbouwgebruikspercelen ALV)
LPIS	Land Parcel Identification System
LU	Land use
LUCAS	Land Use/Cover Area frame statistical Survey
LZW	Lempel–Ziv–Welch
MA	Mali
MAX	Maximum
MCARI	Modified Chlorophyll Absorption Ratio Index
MEAN	Mean
MIN	Minimum
MMU	Minimum Mapping Unit

MODIS	Moderate Resolution Imaging Spectroradiometer
MSAVI	Modified Soil-Adjusted Vegetation Index
MSAVI2	Modified Soil-Adjusted Vegetation Index 2
MSWIR	Mean Short-Wave Infrared Reflectance
MSGI	Metadata Standard for Geographic Information
MTCI	MERIS Terrestrial Chlorophyll Index
MTV2	Modified Triangular Vegetation Index (2)
NBR	Normalized Burn Ratio
NDII	Normalized Difference Infrared Index
NDMI	Normalized Difference Moisture Index
NDRE	Normalized Difference Red Edge Index
NDRE1	Normalized Difference Red Edge Index (1)
NDRE2	Normalized Difference Red Edge Index (2)
NDSVI	Normalized Difference Senescence Vegetation Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference water index
NIR	Near-InfraRed, Sentinel-2 – Near-InfraRed – B8
NNIR	Sentinel-2 – Short Wavelength Infrared – B8a
NO	North
OA	Overall Accuracy
OPT	Optical
OSAVI	Optimized Soil-Adjusted Vegetation Index
PA	Producer Accuracy
PSRI	Plant Senescence Reflectance Index
REP	Red-Edge Position
RF	Random Forest
RGR	Red-Green Ratio
S1	Sentinel-1
S2	Sentinel-2
S-3	Sentinel-3
SA	South Africa
SAR	Synthetic Aperture Radar
SAVI	Soil-Adjusted Vegetation Index
SCL	Scene classification layer
SE	Southeast
Sen2Cor	Sentinel2 Correction
SNAP	Sentinel Application Platform
STD	Standard deviation
SW	SouthWest
SWIR	Short Wavelength Infrared
SWIR1	Sentinel-2 – Short Wavelength Infrared – B11
SWIR2	Sentinel-2 - Short – B12
TCARI	Transformed Chlorophyll Absorption Ratio Index
TCB	Tasseled Cap Brightness
TCC	Tree Cover Change
TCG	Tasseled Cap Greenness
TCW	Tasseled Cap Brightness
UA	User Accuracy
USA	United States of America
VH	Vertical transmit/Horizontal receive (polarization)
VHR	Very High Resolution
VI	Vegetation Indices
VIRP	Visually interpreted reference points

VRE1	Sentinel-2 - Vegetation Red Edge – B5
VRE2	Sentinel-2 - Vegetation Red Edge – B6
VRE3	Sentinel-2 - Vegetation Red Edge – B7
VV	Vertical transmit/Vertical receive (polarization)
WE	West
WI	Wetness Index
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 for future next-generation operational Copernicus Land Monitoring Service (CLMS) products of the pan-European and Global Components. This will contribute to demonstrating operational readiness of the finally selected products, and shall allow the key CLMS stakeholders (i.e. mainly the Entrusted European Entities (EEE) EEA and JRC) to take informed decisions on potential procurement of the next generation of Copernicus Land services from 2020 onwards.

To achieve this goal, ECoLaSS makes full use of dense time series of optical as well as Synthetic Aperture Radar (SAR) satellite data. Rapidly evolving user requirements are analysed in support of a future pan-European roll-out of new/improved CLMS products, and the transfer to global applications.

This Deliverable **D13.2 - D43.1b - Prototype Report: Improved Permanent Grassland (Issue 2)** comprises a description of the provided prototype datasets of Improved Permanent Grassland products (linked to Deliverable P43.2). It provides a detailed description of the objectives together with an explanation of the methodology, results and conclusions, as derived by WP43. It addresses the prototype methodologies for preparation of in-situ reference and validation data sets, application of different grassland classification algorithms (as described in WP33) in the defined demonstration sites West, Central and South-East. Further accuracy assessment for the grassland product prototypes, and optimisation of the algorithms described in WP33 and WP34 based on the assessment results are conducted. As such, it is part of Task 4: “Thematic Proof-of-Concept/Prototype on Continental/Global Scale”, which aims at exploring and setting up a robust classification approach for an improved identification of permanent grasslands 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 **D13.4 - P43.2b - Data Sets of HRL Permanent Grassland 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 being prototypically implemented in a selection of these sites in the frame of the Task 4 WPs. In ECoLaSS, proofs-of-concept / prototype demonstrations are carried out with respect to five topics of relevance: (i) Time series derived indicators and variables WP41, (ii) Incremental Updates of HR Layers WP42, (iii) Improved permanent grassland identification WP43, (iv) Crop area and crop status / parameters monitoring WP44, and (v) New LC/LU products WP45. This deliverable focusses on the prototype GRASSLAND as part of WP43.

This report constitutes the follow-up of the first issue, in which preliminary results up to month 18 were presented. In the second 18-month project cycle, this second issue of this deliverable is published, containing all final results. It is comprised of six Chapters. Chapter 1 of the document is this introduction explaining the purpose and objectives of WP43 as well as the document structure. Chapter 2 presents the background of CLMS grassland monitoring needs and the related summary of requirements. Chapter 3 gives a short summary and description of the demonstration sites. Chapter 4 gives an overview of the time series analysis methods which have been chosen based on the results and recommendations of WP33 “Time Series Analysis for Thematic Classification” and WP34 “Time Series Analysis for Change Detection”. It reviews the multi-sensor data integration methods, the use of multi-temporal time series metrics, the random forest classification approach and the validation analysis procedure. Based on these reviews, Chapter 5 details the prototype implementation approach including a description of the input data, pre-processing lines, the experimental setup, and the final validation of the prototype.

2 Background and Summary of Requirements

After methods testing within Task 3, the demonstration activities of Task 4 show the implementation of the developed advanced processing lines for prototyping. The prototype for “improved permanent grassland” delivers new methodologies using dense Sentinel time series with the aim to develop a prototype of a next-generation European HR Grassland Layer with high thematic accuracy. The improvement of permanent grassland identification targets to enhance the specifications and quality of the current generation of HRL Grassland products, such as implemented by the EEA for the reference time step 2015 and similar for the reference time step 2018 (as confirmed when the Call for Tender was published by the EEA in 2018). The development is based on the experiences of the HRL Grassland 2015 production being the baseline for improvement, as well as high-priority user requirements that are regularly updated and documented in WP21.

The HRL Grassland 2015 is the first layer of its kind, as the previous HRL Natural Grassland with the reference year 2012 suffered from technical constraints and accuracy limitations leading to a quite restricted definition of *natural* grassland to be detected. The HRL 2015 Grassland layer, however, shows a completely new product, comprising natural, semi-natural and managed grasslands of the EEA39 countries at 20m spatial resolution and with a minimum mapping unit of 1ha. The layer represents the full range of grassland types and covers all typical grassland landscapes of Europe.

1. The HRL Grassland 2015 production used a multi-temporal and multi-sensor approach. The layer was produced by using a combined optical/SAR data analysis approach based on data from the reference period 2015 +/- 1 year. Image segments derived from multi-temporal, best-suited optical EO data were utilized to classify the multi-temporal data base of both optical and SAR input data. Automatically and manually derived training samples of the main land use classes were selected and applied in a supervised multi-temporal classification approach to compute grassland maps from both sensor types. A subsequent rule-based evaluation finally defined the optimum grassland mask. Additionally, recent and historic bare soil masks helped to identify grassland areas that show a ploughing indication and therefore were excluded from the mask. The result was a pan-European grassland/non-grassland mask showing an overall thematic accuracy of 94.32% and for the grassland class a user accuracy of 85.9% and a producer accuracy of 77.8%. These are internal validation results provided by the HRL 2015 consortium. It was based on open source VHR images in combination with the multi-temporal, multi-seasonal EO data base (e.g. Sentinel-2, Landsat) used for production.

Despite the high quality of the present pan-European HRL Grassland 2015 layer, there is potential for enhancement and improvement of the methodological approach in order to further accelerate production speed, optimise the overall and regional reliability of the classification, to reduce data gaps, increase the spatial resolution and to include further user needs. Main requirements from users for the prototype developments have been collected in WP21 [AD09] and 51 [AD11], as well as from current developments in Copernicus, and can be summarized as follows:

- Design a refined and sound workflow with an **improved level of automation** to allow a faster production and shorter monitoring intervals (e.g. yearly updates)
- Improve the **thematic classification accuracy**
- Fully exploit **optical Sentinel-2 time series** instead of using pre-selected, best-suited EO data scenes
- Design a fully **integrated SAR/optic** time series data analysis to benefit from the multi-sensor characteristics
- Provide a **seamless, wall-to-wall product** (maximally reduced data gaps due to cloud cover)

- Improve the status layer's detail from 20m **spatial resolution to 10m**. Product definitions consequently might have to be adapted, such as e.g. the Minimum Mapping Unit
- Investigate a **future change detection** approach to detect grassland increase and decrease
- Include **more seasonal information** with respect to the grassland's phenological/vitality behaviour that can be further exploited in order to support further **grassland discrimination** between e.g. intensively managed (frequently cut grassland) and extensively managed (more natural, extensively used or grazed grassland)

The methodologies tested for the grassland prototype are described in the following chapters. They address the above mentioned user demands and try to provide the best trade-off between grassland classification accuracy, level of automation and time-series integration. At present, and as proved in the implementation phase 2 in ECoLaSS, a bottleneck towards larger scale production relates to validation constraints. Besides the timeliness aspect, the visual inspection and the expert knowledge required in products verification is a drawback, and other limitations come from the lack of homogeneous reference data with inclusion probabilities for the area based accuracy assessment. More specifically, the lack of in situ data affects in particular production of the grassland mowing intensity. In terms of operational production, the HRL 2018 Grassland has recently started. Improvements include the spatial resolution of 10m, as in the ECoLaSS prototypes. The accuracy standard in HRL 2015 was 85% overall accuracy for the whole layer (i.e., EU extent), whereas in the HRL2018 the quality standards specifications aim at 85% overall accuracy within each biogeographical region. Regarding minimum mapping units (MMU), in the previous HRL 2015 MMU is defined as 1ha, whilst in the current HRL 2018, considering the product is raster based at 10m spatial resolution pixel, several options are still under discussion (e.g., filtering by 1-2 pixels is the most likely option at present). In this sense, the prototyping phase in ECoLaSS has tackled several options and follows these current HRL 2018 specifications (e.g., 10m spatial resolution of grassland status layers, 85% overall accuracy target), that in turn were gathered in the requirements analysis tasks (Task 2 and Task 5 outcomes). Expert knowledge and region-wise decisions have been applied in the ECoLaSS testing and prototyping. In turn, such decisions in the change layer (MMU 0.5ha) must consider the references layers under comparison in order to produce coherent and meaningful change layers. For further methodological details, the methodological compendium reports from Task 3 can be consulted.

3 Demonstration Sites

All prototypes are implemented in selected representative demonstration sites, which cover various biogeographic regions and biomes. The Improved Primary Status Layer for Grassland is first demonstrated in the West (Belgium) Demonstration site. In phase 2 the developed processing line on grassland identification has been implemented on the demonstration sites West, South-East and Central.

3.1 ECoLaSS Demonstration Sites

The selected larger demonstration sites (60,000/90,000km² per demonstration site) contain the 5 test sites from Task 3. These demonstration 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 West covered the Atlantic and Continental 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 ECoLaSS demonstration 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 have been prototypically implemented in one or more demonstration sites in project phase 1, and in three demonstration sites in phase 2 (including the sites of phase 1).

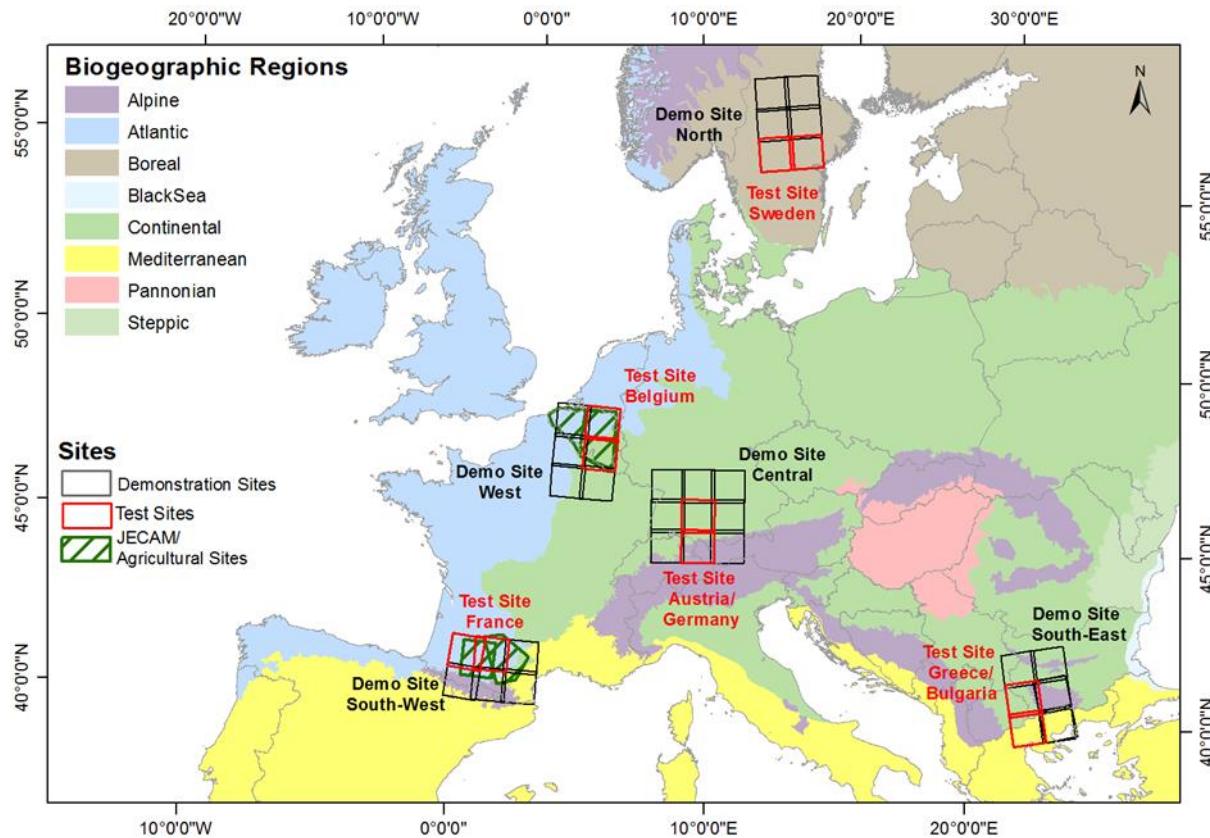


Figure 3-1: European Demonstration Sites (modified from: EEA, 2015).

A short description of the different demonstration 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%

3.2 Demonstration Site West for Grassland

The demonstration site “West” serves for demonstrating the proposed candidates for Copernicus Land Service evolution in terms of roll-out to a larger scale (Task 4) in ECoLaSS for the thematic topics: Grasslands, Indicators/Variables, Agriculture and New Land Cover products (including the CLC approach and the consistency of the combined HRL layers approach).

Within this the demonstration site West, with an area of approximately 65,019.86km², 63.08% are agriculturally used while the rest is covered by water bodies (1.61%), forests and semi-natural areas (63.08%), artificial areas (13.48%), as well as wetlands (0.40%). In this statistic, the various grassland types appear under different land cover classes. Natural grasslands (CLC code 321) with 0.10% are generalised under forests and semi-natural areas, whereas pastures (CLC code 231) with 11.7% are generalised under agricultural areas. Further, green urban areas (CLC code 141) with 0.104% and sport and leisure facilities (CLC code 142) with 0.433% are generalised under artificial areas. This shows that the demonstration site West is mainly comprised by natural grasslands and agricultural grasslands. A map of the selected demonstration site West for the HRL Grassland (GRA) prototype is provided in Figure 3-2 below.

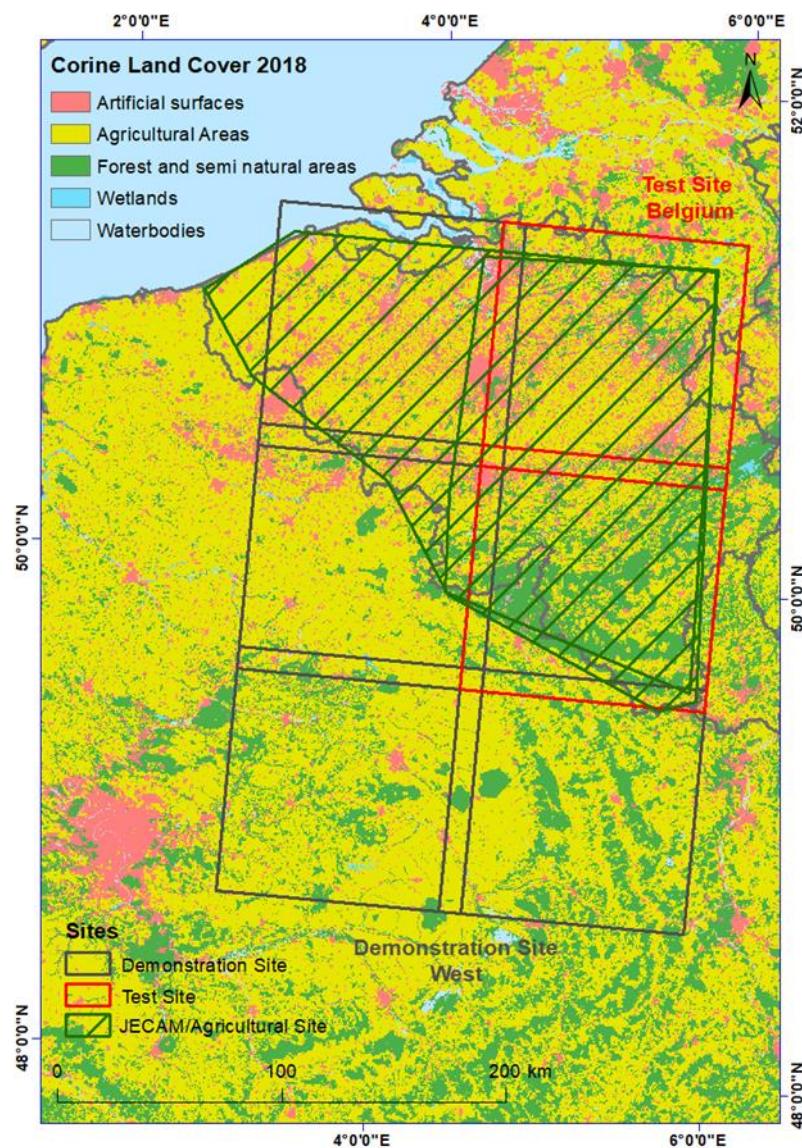


Figure 3-2: Overview of Demonstration site West.

3.3 Demonstration Site Central for Grassland

The demonstration site “Central” incorporates the areas of the test sites and serves for demonstrating the proposed candidates for Copernicus Land Service evolution in terms of roll-out to a larger scale (Task 4) in ECoLaSS for all thematic topics addressed: Grasslands, Forest, Imperviousness, Indicators/Variables, Agriculture and New Land Cover products (including the CLC approach and the consistency of the combined HRL layers approach). The selected prototype sites cover the most important environmental zones (Source: EEA) of Europe and the member states of the EEA-39. In particular, Central site is located in the Continental and Alpine Central European region (Germany, Austria, Switzerland, Italy). During phase 1, the sites were reviewed as was explained in the Periodic Technical Report and second issue of the Interim Progress Report [AD13]. Accordingly, the Central site remains a large site with 9 Sentinel-2 granules (approx. 90.000km²).

The demonstration site is dominated in the North by cropland areas, mixed with grassland (pastures). The Southern part, covering the Bavarian Alpine Foreland, is dominated by forest cover and grassland, including extensively used grassland and wetland areas. The test site covers the “Wetterstein mountain range” as part of the Alps with mountain-specific vegetation zones and stretches South down to the Inn valley. The whole demo is very challenging regarding grasslands and agriculture discrimination as it is heterogeneous in terms of climate/microclimate conditions as well as in terms of altitude leading to heterogeneous farming management systems (time windows for sowing and harvesting) as well as to shifted vegetation periods. On the North/South gradient, there is a high variety of crop types especially in the more favourable areas of north and central tiles), with more fodder crops/grasslands and pastures towards the southern areas closer to the Alps. In the alpine region agricultural areas are restricted mainly to valleys whereas grassland and pastures are dominant in higher altitudes. In the western part, towards the north, the River Rhine valley, with very mild climate even in wintertime, favours the presence of vineyards and agriculture uses, whereas towards the south the Black forest domains and the agriculture land use is present up to moderate altitudes. Croplands are also present along the Danube riverside and lake Constanza. According to CLC, within the demonstration site Central, with an area of approximately 96,013.29 km², 44.55% are agriculturally used while the rest is covered by water bodies (1.55%), forests and semi-natural areas (44.64%), artificial areas (9.03%), as well as wetlands (0.22%). In this statistic the various grassland types appear under different land cover classes. Natural grasslands (CLC code 321) with 3.76% are generalised under forests and semi-natural areas, whereas pastures (CLC code 231) with 16.12% are generalised under agricultural areas. Further, green urban areas (CLC code 141) with 0.06% and sport and leisure facilities (CLC code 142) with 0.42% are generalised under artificial areas. This shows that the demonstration site Central is mainly comprised by forests and semi-natural grasslands and agricultural grasslands. A map of the selected demonstration site Central for the HRL Grassland (GRA) prototype is provided in Figure 3-3 below.

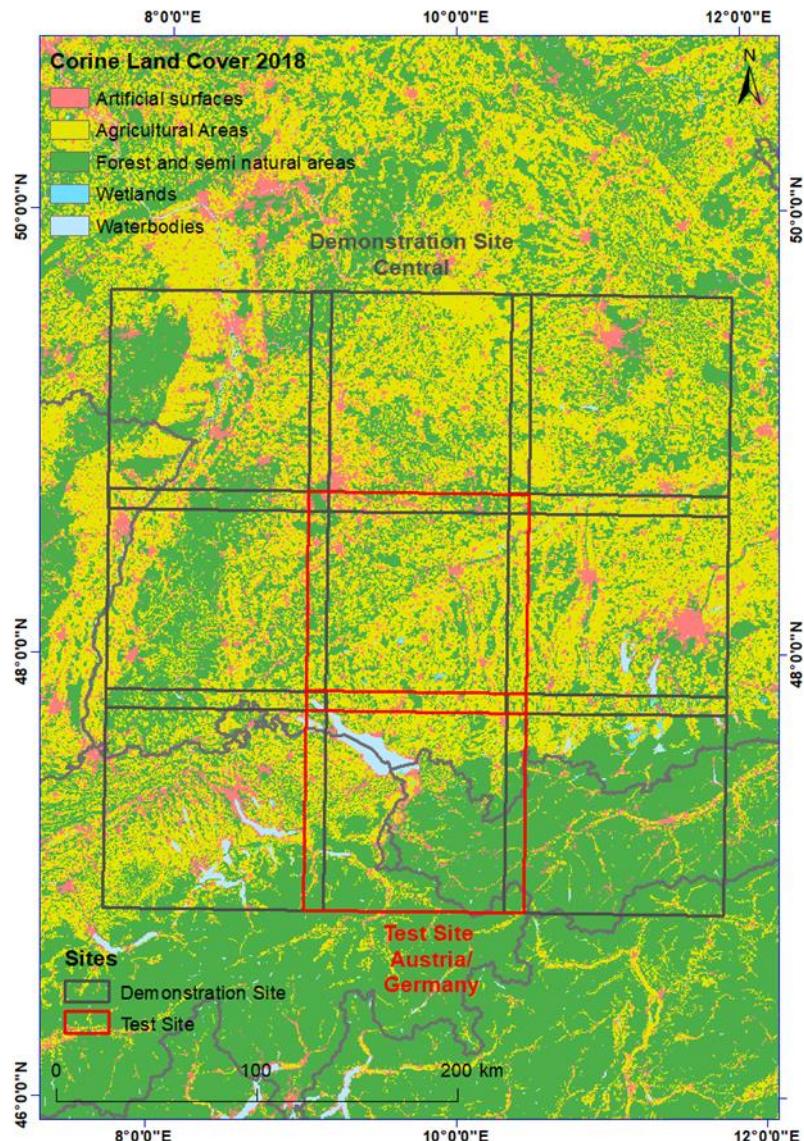


Figure 3-3: Overview of Demonstration site Central.

3.4 Demonstration Site South-East for Grassland

The demonstration site “South-East” incorporates the areas of the test sites, and serve for demonstrating the proposed candidates for Copernicus Land Service evolution in terms of roll-out to a larger scale (Task 4) in ECoLaSS for the following thematic topics addressed: Grasslands, Forest and Imperviousness. The selected prototype sites cover the most important environmental zones (Source: EEA) of Europe and the member states of the EEA-39. In particular, South-East site is located in the Mediterranean and Steppic European region (Greece/Bulgaria), including continental and alpine areas. Within the demonstration site South-East, with an area of approximately 65,545.32 km², 34.87% are agriculturally used while the rest is covered by water bodies (4.93%), forests and semi-natural areas (34.87%), artificial areas (3.34%), as well as wetlands (0.17%). In this statistic the various grassland types appear under different land cover classes. Natural grasslands (CLC code 321) with 5.73% are generalised under forests and semi-natural areas, whereas pastures (CLC code 231) with 2.56% are generalised under agricultural areas. Further, green urban areas (CLC code 141) with 0.04% and sport and leisure facilities (CLC code 142) with 0.08% are generalised under artificial areas. A map of the selected demonstration site South-East for the HRL Grassland (GRA) prototype is provided in Figure 3-4 below.

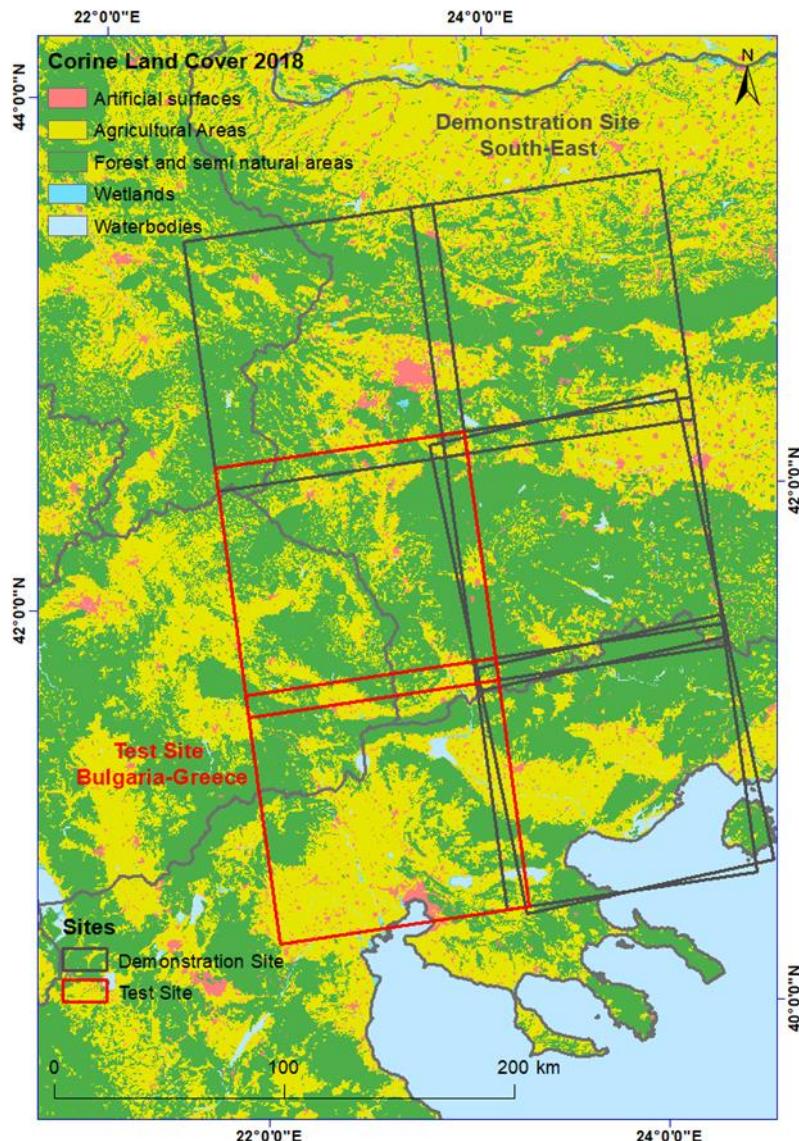


Figure 3-4: Overview of Demonstration site South-East.

4 Overview of Applied Methods

Grasslands are dynamic throughout the time and within its growing period, it is characterised by changing canopy density, chlorophyll status and ground cover (Zillmann et al., 2014). Since grasslands and crops show significant variations throughout their growing cycle, time series representing information about phenological dynamics are required. One aim of the prototype implementation is to identify the most suitable phenological seasons and time series indicators from optical and SAR time series to get the best-trade-off between minimum numbers of individual time features needed and classification accuracy.

4.1 Multi-Sensor Data Integration

The usage of only SAR time series as alternative data, and the combination and integration of S1 and S2 data, closing data gaps due to cloud cover, or as complementary information to increase the thematic accuracy has been addressed in WP31 [AD05] and Task 3 tasks ([AD07], [AD08], [AD12]). The advantages of the synergetic combination of different sensors at high spectral and spatial resolution are (i) to

achieve a denser temporal resolution by filling the missing observations of S2 data due to cloud cover with SAR data, and (ii) to create complementary information layers as input for the classification approach. Several state-of-the-art fusion methods were benchmarked and tested in WP31 [AD05] including the pixel level fusion, feature level fusion and the decision level fusion.

As described in WP31 [AD05], the image fusion on pixel level is the computationally most intensive approach because the merging is applied directly on the pixels obtained at the source images. This approach can be used to take all information into account when the sensors have similar spatial characteristics (Samadzadegan, 2004). The feature level fusion uses different extracted features derived from source data or ancillary data, which are also used for classification purposes. The feature level fusion should be used if the features can be appropriately associated (Samadzadegan, 2004). Decision fusion methods combine results obtained through separate classifications of optical and SAR features. Indices and statistics for the classification are calculated for each data set separately and the results of the classification can be combined using logical AND functions, expert systems or probabilistic techniques. This approach should be used when the sensors are very different. Additionally, it has the advantage of being less computationally intensive (Samadzadegan, 2004).

As shown in the benchmarking applications in WP31 and WP33, the integration of S1/S2 allows benefiting from the multi-sensor characteristics, using the information gained from both sensors in parallel for the classification process as they record complementary characteristics of the land surface. Therefore, integrated S1/S2 datasets were used for the prototype production.

4.2 Multi-temporal SAR and Optical Metrics

Multi-temporal features that aim at quantifying the vegetation cycle based on the images acquired by fused S2 and S1 images are used in the prototype production. Their statistical properties are computed in order to select the most relevant features to study grassland cover. Several metrics can be constructed to focus on grassland, derived from both, optical and SAR time series. As already described in WP31 [AD05] temporal metrics can be summarized to different types. The first category includes statistical metrics of spectral values calculated over one or several periods such as the average, the maximum, the minimum – especially relevant to follow seasonality and phenology. The second category includes change metrics, usually derived from a temporal trajectory which represents the magnitude and the duration of the change, or the slope of the modelled trend. Third, stationary or non-stationary shape variables derived by temporal trajectories can be seen as a function of time and provide periodic patterns or transitions patterns. Final, trend metrics can be used to find simple linear trends, seasonal trends and breakpoints to describe landscape processes. For further detailed information, see WP31 [AD05].

In the prototype implementation the above mentioned statistical metrics of spectral values calculated over one or several periods are exploited. Different multi-sensor temporal statistical metrics are derived from S1 only, S2 only and the synergetic use of optical and SAR data to enhance the accuracy of the classification result in comparison to optical or SAR only features (described in detail and benchmarked in WP33 [AD07]), including seasonal metrics and vegetation indices.

4.2.1 Spectral Optical Indices

Many spectral indices have been defined in the past. Some have been and are still widely used, such as the NDVI, while others have only been proposed as alternatives in the recent years. In particular, the novel red-edge bands of Sentinel-2 MSI provide improved opportunities for the derivation of spectral indices addressing vegetation state and photosynthetic activity. In the WP31 [AD05], main spectral indices are listed focusing on vegetation and built-up discrimination in order to explore their potential contribution to the evolution of Copernicus Land Services products. Moreover, the assessment of vegetation chlorophyll status at the canopy level, using remote sensing data is important for providing information about the health status of grassland ecosystem, for example, the physiological status, productivity, or phenology (Tong and He, 2017). The paragraphs below complement the description of

indices in WP31 with focus on those indices which were used in the feature selection process to determine the optimal set of features for grassland identification.

Zillmann et al. (2014) concluded that the usage of seasonal statistics of various vegetation indices from multi-seasonal images in conjunction with the spectral reflectances of the visible, near-infrared and short-wave infrared spectrum are useful for grassland discrimination. The most relevant seasonal vegetation indices features are the Normalized Difference Senescent Vegetation Index (NDSVI) average and maximum, due to the sensitivity of the index to drying and senescence events. Other useful seasonal indices used within the study are NDVI, ground cover (GC), Plant Senescence Reflectance Index (PSRI), Normalized Difference Infrared Index (NDII), NDSVI, Wetness Index (WI), and Brightness as they describe phenological differences of grasslands and croplands (Zillmann et al., 2014). Another study conducted from Yang et al. (2017) stated that the Normalized Difference Vegetation Index (NDVI), Red-Green Ratio (RGR), Enhanced Vegetation Index (EVI), Normalized Difference Infrared Index (NDII), Modified Triangular Vegetation Index II (MTV2), SWIR Reflectance and PSRI were important for distinguishing grassland and cropland. Furthermore, Tong and He (2017) tested the performance of 144 published broad- and narrowband vegetation indices for retrieving chlorophyll content for a semi-arid mixed-grass prairie ecosystem. In general, narrowband indices utilizing data from a wavelength from the red to the red-edge region (690–750nm) performed best. Broadband indices are found to be as effective as narrowband indices for chlorophyll content estimation at both leaf and canopy scales. Based on the results of the benchmarking applications in WP31 and WP33 and on findings described in above papers, the most promising indices were selected for the prototype implementation. Please note, that only a subset of these indices were finally used in the classification, based on the automated derived feature importance measures, as described in next chapters.

4.2.2 Multi-seasonal Statistical SAR and Optical Features

Multi-seasonal features can be generated from different seasons concerning those periods of the year where grassland could be identified the best, taking into account agricultural management schemes as well as grassland mowing cycles. To characterize different temporal behaviours of grasslands and other land cover classes, high-frequency acquisitions over the growing seasons are essential. Nevertheless, the temporal behaviour may vary considering different mowing practices. Schmidt et al. (2014) concluded from their study with RapidEye data from 2009 to 2011 that the NDVI composites from the early summer seasons are most important for the grassland discrimination, whereas spring, late summer and midsummer seasons also played an important role. This is in line with observations made during feature-selection in the ECoLaSS grassland prototype production. Likewise, Keil et al. (2015) stated that March, May and August (using NDVI composites) are important seasons in central Europe to discriminate crops and grasslands. After identifying the optimal acquisition period, it still can be impossible to find good quality images or image composites with no cloud gaps for the entire area (Zillmann et al., 2013). Wang et al. (2010), extracted temporal trajectories of the normalized difference vegetation index (NDVI) and the normalized difference moisture index (NDMI) to examine the temporal variation of warm-season grass and cool-season grass grasslands in a growth cycle. It was found that the spring–summer period revealed maximal spectral differences between these two grass types. The NDVI is stated to be more useful than NDMI in summer–fall. The NDVI trends in this period varied with both phenology and grassland treatments such as haying and grazing. As pointed out in WP33 [AD07] these seasons vary for different regions with changing climate conditions and differing management systems (Zillmann et al., 2013). Although the most important seasons are described in literature, they vary depending on climatological conditions (e.g. drought period in summer 2018) and the application areas (e.g. mountainous regions).

4.3 Random Forest Classification Approach

The amount of images used in multi-temporal classification studies has greatly increased along with enhanced temporal sensor capacities. As pointed out in WP33 [AD07], the availability of dense time series calls for the use of machine learning algorithms over the traditional statistical classification

approaches due to the increasing computational capabilities necessary to process big amounts of data (Waske and van der Linden, 2008).

The Random Forest (RF) classifier first proposed by Breiman (2001) belongs along with other boosting and bagging methods as well as classification trees in general to the ensemble learning methods, which generates many classifiers and aggregates their results to calculate their response (Liaw and Wiener, 2002; Horning, 2010). The random forest algorithm generates multiple decision trees with randomly drawn subsets, instead of using all variables from the available data. The subsets are drawn with replacement, meaning that one sample can be selected several times, while others may not be selected at all (Belgiu and Dragut, 2016; Ali et al., 2012). Regarding each random sample, a classification or regression tree is grown to the largest possible extent without pruning. At each node, a random sample of a predictor variable is extracted; among those, the best split is chosen. To predict new data, the prediction among all trees are aggregated using majority votes. The class with the maximum vote overall decision trees is the one selected for the output product (Liaw and Wiener, 2002; Ali et al., 2012). One advantage of the classifier is the calculation of the variable feature importance.

In this context, the relative importance of variables is calculated for each feature available for both optical and SAR data. Within the forest generation, every node in the decision trees is a condition on a single feature to split the dataset. The Mean Decrease Impurity (also known as Gini importance) measure, calculates the sum of the total impurity reductions at all tree nodes where the variable appears (Breiman, 2001). Therefore, each feature importance represents the sum over the number of splits across all trees that include the feature; proportionally to the number of samples it splits (Louppe et al., 2013). One drawback of this method is that the mean decrease impurity measure is biased towards preferring variables with more categories. Another drawback is when the dataset is composed of correlated features, which can be assumed to have the same importance. Nevertheless, the first feature analysed reduces the importance of other correlated features (Louppe et al., 2013). As described in the final issue of WP33 [AD07] Further the grouped forward feature selection method was also applied. The goal of feature selection is two-fold: improve the computational efficiency and reduce the generalization error of the model by removing irrelevant features or noise. This wrapper removes or adds one feature at the time based on the classifier performance until a feature subset of the desired size is reached. The difference with other methods is that the forward feature selection eliminates or adds features based on a user-defined classifier/regression performance metric. The algorithm finally yields a combination of the features with the highest accuracy. The subset of features is subsequently used for the classification process.

4.4 Validation Analysis Procedure

4.4.1 Sampling Point Generation for Grassland

The approach applied is based on the approach developed for the validation of CLMS Pan-European thematic products.

4.4.1.1 Stratification and Sampling Design

The following sub-chapters provide a description of the procedure of a scientifically and statistically sound sampling scheme for assessing the thematic quality. This comprises descriptions of the customized stratification approach, the sample size calculation procedure and the strategy for ensuring representative sample distribution and sufficient regional spread, whilst considering practical feasibility and cost-effectiveness within an operational setup.

4.4.1.2 Overview

The stratification and the sampling design consist primarily in selecting an appropriate sampling frame and sampling unit. The sampling units can either be “defined on a cartographic representation of the surveyed territory” (Gallego and JRC-IES 2004), in which case it is an area frame, or on a list of the

features. According to this study, area frames give a better representation of the population as the spatial dimension is kept.

In an area frame, sample units can be points, lines (often referred to as transects) or areas (often referred to as segments, described by (Gallego 1995). The first step is to define the geographical area for which the accuracy assessment is to be reported and the type of sample units. For the majority of cases, point samples will be used, but areas or segments may be used in specific cases such as when not only thematic accuracy need to be reported, but also the geometry of mapped objects. Points are considered as the most appropriate unit for our purpose. Polygons have also the drawback of being specific to a single map. In case of changes, the sample may not be adapted anymore.

Sampling design refers to the protocol whereby the samples are selected. A probability sampling design is preferred for its objectivity. “Simple random, stratified random, clustered random and systematic designs are all examples of probability sampling designs” (Stehman and Czaplewski 1998). Even though a simple random design is easy to implement, its main drawback is that some portions of the population may not be adequately sampled (see Figure 4-1:). Cluster sampling is often used to reduce the costs of the collection of reference data, but does not resolve geographic distribution problems. A systematic approach would solve this problem, yet it is not appropriate if the map contains cyclic patterns. A stratified approach consists in allocating a pre-defined number of samples per land-cover class. As explained in Stehman and Czaplewski (1998), stratification ensures that each class is represented.

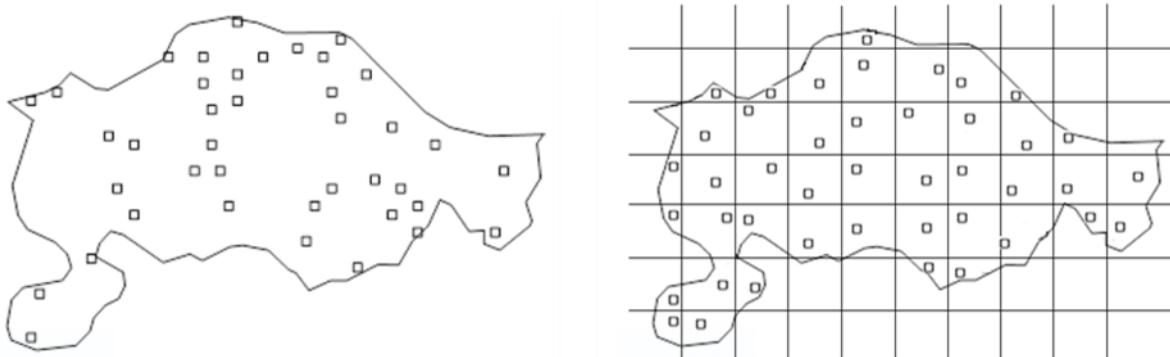


Figure 4-1: Simple random (left) and random systematic (right) sampling designs.

4.4.1.3 Selected Sampling Design

The selected sample design for thematic accuracy assessment for this Demonstration site combines a systematic and stratified approaches and benefits from the advantages of both of them. It is based on LUCAS (Land Use/Cover Area frame statistical Survey) sampling approach. LUCAS corresponds to a grid of approximatively 1,100,000 points throughout the European Union where land cover or land use type is observed. Using LUCAS points ensures traceability and coherence between the different layers.

LUCAS points are located every 2 km on a regular grid, as illustrated in Figure 4-2: . A set of 81 points located on an 18x18 km square constitutes a group in which every point is associated with a number comprised between 1 and 81 (note that the numbers do not follow each other spatially consecutively). The same pattern with the same numbers allocation is repeated all over the grid. Within this context, a replicate refers to the points with the same number selected over the whole LUCAS grid.

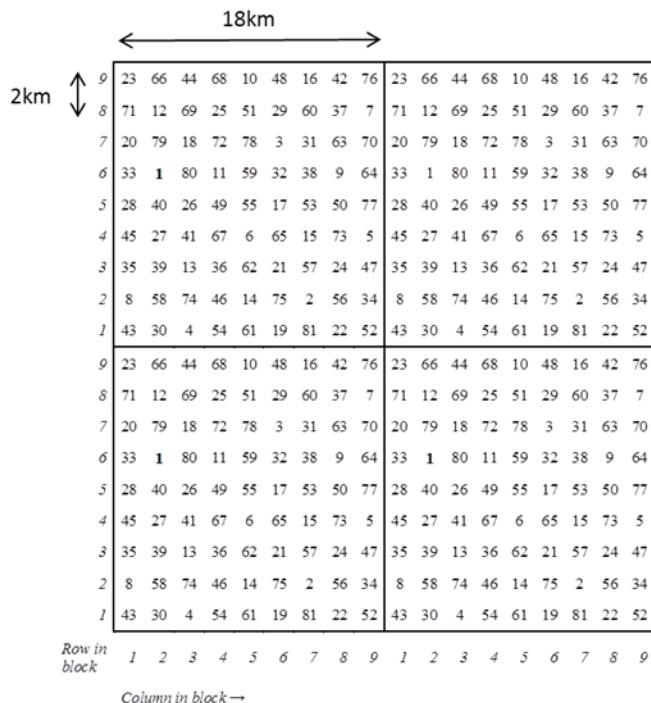


Figure 4-2: LUCAS points located on a regular grid.

At first, the number of samples to allocate to each stratum (i.e. thematic class) is calculated as a function of its areal proportion in the investigated reference area. In this manner, the sampling design is not only systematic but also stratified.

The determination of the number of sample units considers the number of thematic classes. It is possible to estimate a suitable sample size for each stratum (i.e. thematic class) based on the expected acceptable error rate. The standard error of the error rate can be calculated as follows:

$$\sigma_h = \sqrt{\frac{p_h(1-p_h)}{n_h}} \text{ where } n_h \text{ is the sample size for stratum } h \text{ and } p_h \text{ is the expected error rate. This}$$

can be reworked to express sample size n_h as a function of p_h and desired standard error σ_h :

$$n_h = \frac{p_h(1-p_h)}{\sigma_h^2}.$$

From Figure 4-3, it can be seen that for an expected 50% error rate, within a stratum, 100 sample units are required for a standard error of 5%, whereas the number of samples would need to be increased by a factor of four if the accepted standard error is divided only by a factor of 2 (i.e. to 2.5%). On the other hand, if the expected error rate is 15%, only 51 samples would be necessary with a 5% standard error. A similar approach had been adopted to determine the sample size for assessing the accuracy of CLC2006 and CLC2000-2006 changes (Büttner et al. 2012). Equal allocation of sample units per stratum works particularly well to assess commission errors, whereas the definition of an appropriate number of sample units for omission errors is more difficult because it depends on the expected spatial extent of the theme to be mapped. Therefore, to ensure that both commission and omission errors are accounted for, the approach suggested by Olofsson et al. (2014) is applied combining the allocation of a minimum number of samples for the smallest strata with proportional allocation to larger strata.

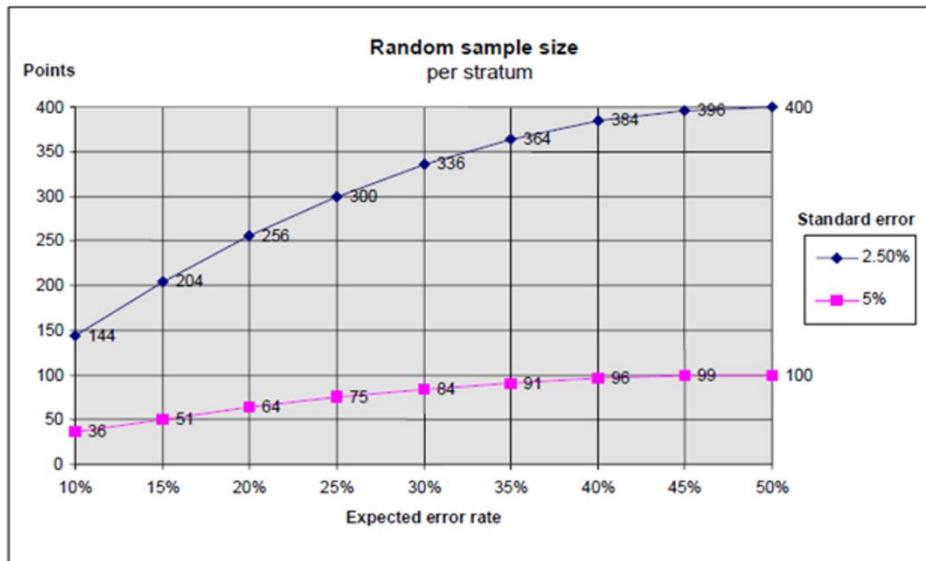


Figure 4-3: Number of sample points as a function of the expected error rate for two accepted standard error values (after Wack et al. 2012).

A minimum number of sample units per stratum is defined to ensure that even small strata are properly represented in the sample. As the error rate is expected to be less around 10%, a minimum of 50 sample units per reporting stratum should be sufficient. The number of replicates to be selected for a stratum depends on its areal cover percentage and the number of LUCAS points intersecting the stratum.

For thematic classes covering a large proportion of the study area, 1 replicate may already exceed the defined number of samples for this class. To solve this problem, replicates are split into four sub-replicates, as illustrated by the blue numbers in Figure 4-4: .

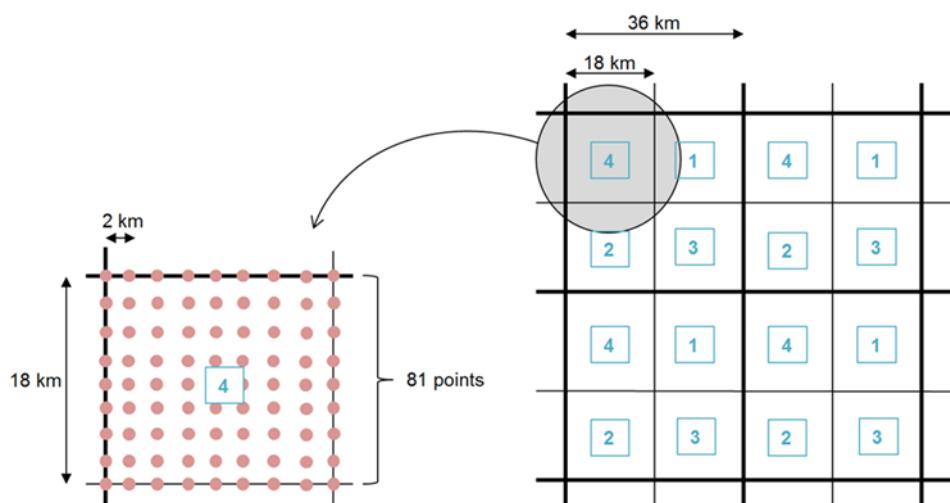


Figure 4-4: Replicates and sub-replicates used on LUCAS grid.

4.4.1.4 Stratification Approach

For the selected demonstration site over Belgium/France, a stratification is applied based on an omission and commission strata:

- Commission: Grassland class
- Omission: Non-Grassland areas

The number of sample units called Primary Sampling Units (PSUs) per stratum based on LUCAS and densified LUCAS grid should be such to ensure a sufficient level of precision. The minimum number of PSUs per stratum will be set at 50. The total number of PSU is presented in the Table below:

Table 4-1: Number of PSU.

Strata	Area of each strata (ha)	Percentage	Number of PSUs
Omission Strata	777 434	86.95%	598
Commission Strata	5 181 871	13.05%	102
Total	5 959 305	100%	700

These 700 samples units have been selected and distributed as shown in the Figure 4-5:

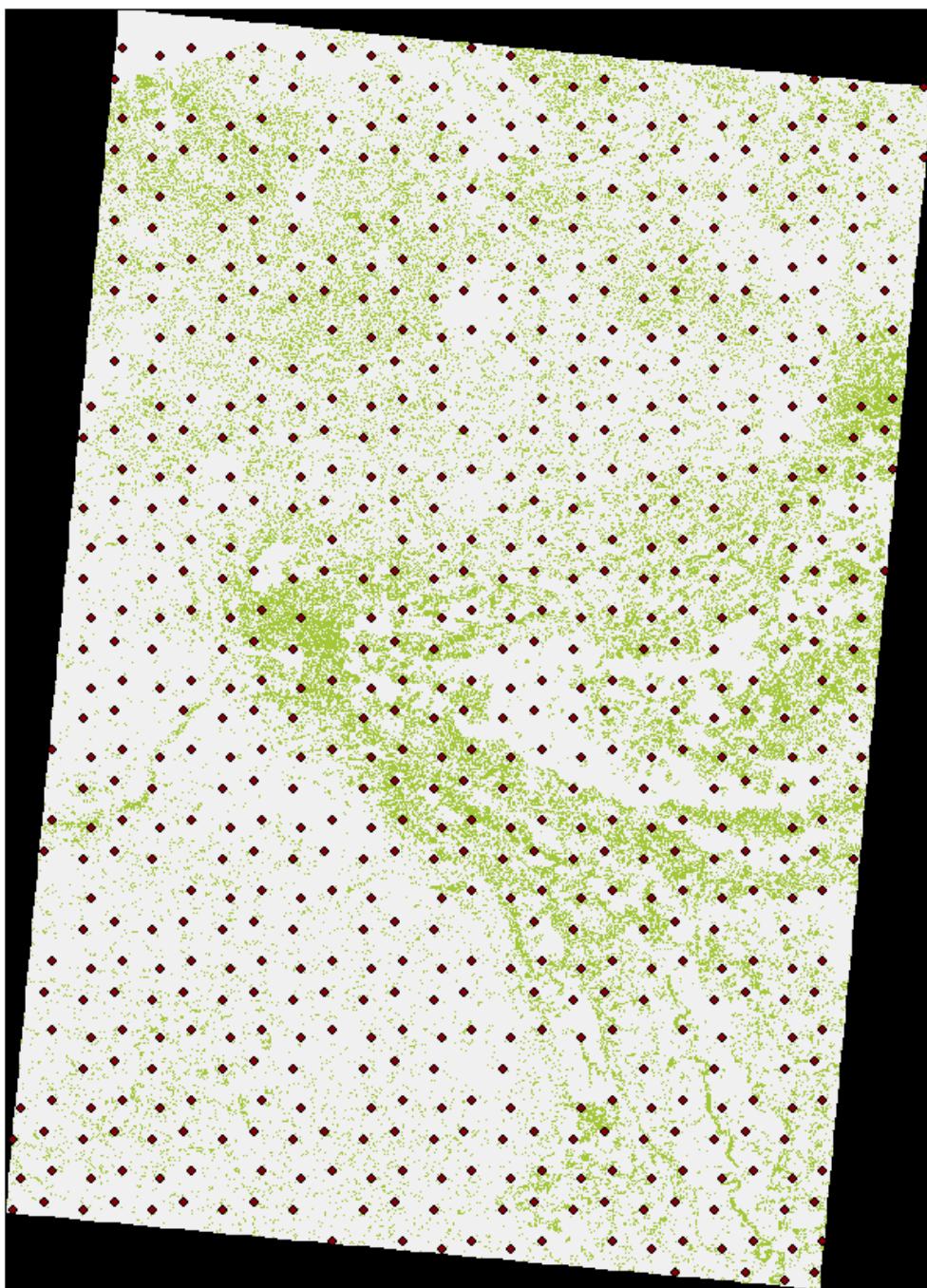


Figure 4-5: Validation samples over the Demonstration site GRA_2017_010m_BE_03035_V0_1.

4.4.2 Response Design

For the assessment of the Grassland layer, the selection of PSU will define a point within an area that should be larger than the MMU (i.e. 0.5ha) for the assessment of the product. In addition, PSUs should be surrounded by eight pixels of the same value within the Grassland product to avoid any boundary effect. The interpretation of each PSU at a particular point is based on the independent assessment (grassland or non-grassland area) at the PSU level. The response design “is the protocol for determining the reference land cover classification of a sampling unit” (Stehman & Czaplewski, 1998).

At the moment, no reference dataset has been identified at pan-European scale to proceed to the validation of an intensive/extensive status for the grassland patches. Depending on the country, some LPIS datasets include such information, that either could be partial or based on a different definition, and it should be noted that their accuracy and reliability have not been established.

The datasets against which the interpretation is performed are divided in two main groups: guiding data and reference data. The guiding data used in the production of the classifications are the re-processed HR Sentinel-2 data. The reference data provide more spatial detail and strong landscape context to the assessment. The available reference data are:

- Bing maps image / cartography layer
- Google Earth image / cartography data

The interpretation process is controlled and the results recorded for the selected sample locations in a point vector layer. PSUs are represented by centre point at the selected location. Each PSU point contained a set of attributes to both define the characteristics of the sample point and the results of the assessment. At the PSU level the guiding data is screened for clouds and cloud shadows and a note is made in the comment attributes when no useable data are available for a reference year. Either the presence of any spatial shifts between the guiding datasets or against reference datasets are also noted.

Then, using the guiding data in the context of the reference data, an assessment is made whether the surface at each PSU is a grassland and non-grassland areas. If necessary, the UNCERTAIN and COMMENT attributes will be used to record points where the coding are unclear or certain issues occurred.

The thematic accuracy assessment was conducted in a two-stage process: an initial blind interpretation in which the validation team did not have knowledge of the product's thematic classes and a plausibility analysis performed on all sample units in disagreement with the production. This chapter will provide the **plausibility** results production methodology. When possible, blind results have been added as complements – this step was indeed skipped for the change layers validation, due to the very limited spatial area affected by those changes.

4.4.3 Validation Analysis Procedure

Thematic accuracy is presented in the form of an error matrix made out of the results of the samples blind and plausibility interpretation. As explained in (Selkowitz and Stehman, 2011) unequal sampling intensity resulting from the stratified systematic sampling approach should be accounted for by applying a weight factor (p) to each sample unit based on the ration between the number of samples and the size of the stratum considered:

$$\hat{p}_{ij} = \left(\frac{1}{N} \right) \sum_{x \in (i,j)} \frac{1}{\pi_{uh}^*}$$

Where i and j are the columns and rows in the matrix, N is the total number of possible units (population) and π is the sampling intensity for a given stratum.

This is because the samples from the smaller strata potentially exhibit a higher sampling intensity than those from the larger strata. Therefore, a correction for the sampling intensity will be applied to the error matrices produced following the procedure described (Selkowitz and Stehman, 2011) and applied (Olofsson et al., 2013) leading to a weighting factor inversely proportional to the inclusion probability of samples from a given stratum. Not applying this correction could result in underestimating or overestimating map accuracies.

Thematic accuracy is usually assessed based on the construction of confusion or error matrix which can be described as illustrated in Figure 4-6 for 5 thematic classes.

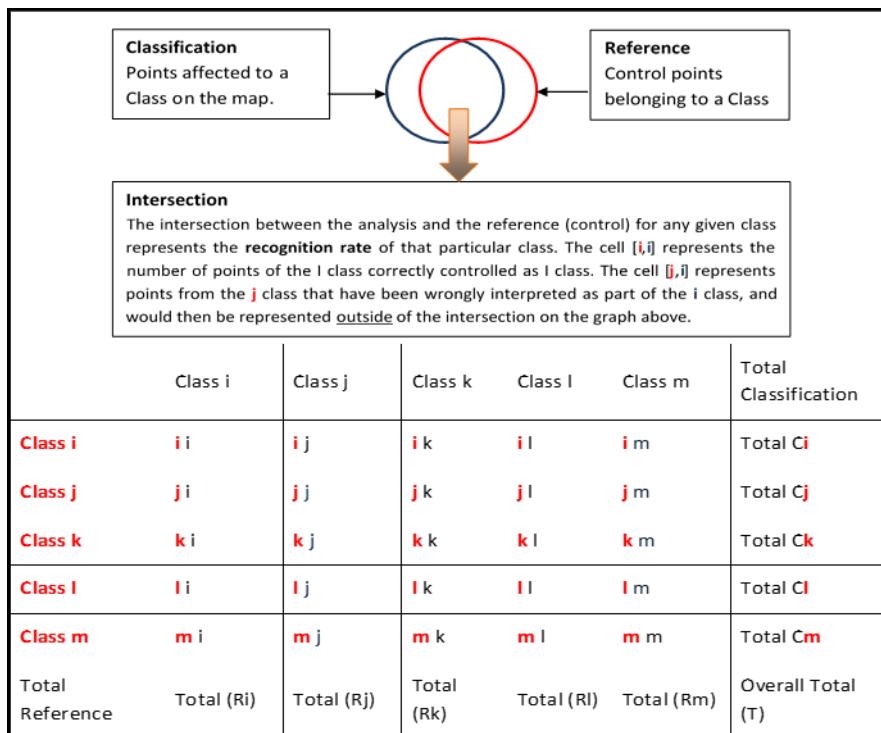


Figure 4-6: Confusion Matrix for Accuracy Assessment of thematic map product.

Let us assume α represents any given class of $[i, j, k, l, m]$, the following accuracy metrics can be calculated:

- The **Overall Accuracy** or Recognition Rate is measured by the sum of the diagonal of the Confusion Matrix divided by the total number of controlled points: $OA \text{ or } R_r = \sum_{\alpha=i}^m (\alpha\alpha)/T$. The Recognition Rate or Overall Accuracy assesses the overall agreement between the classified and reference data set. However, for single class themes such as in the HRL, it does not necessarily provide a realistic assessment of the quality of the map produced because the area covered by the theme considered (Grassland in this case) can be relatively small in comparison to the rest of the area.
- Therefore, the row and column totals and the diagonal of the Matrix are used to assess two types of accuracy, the User and Producer Accuracy:
 - **Producer Accuracy** for the α class = $\alpha\alpha/C_\alpha$ is a measure of **omission error**. For instance, an observation has been identified as grassland during the validation, but has been classified as another class: it has been omitted from the grassland class. The number of omission errors should be less than 15% for the Grassland Mask.
 - **User Accuracy** for the α class = $\alpha\alpha/R_\alpha$ is a measure of the **commission error** (or contamination risk): errors due to the wrong allocation of an observation to a class. For instance, an observation is classified as grassland, but identified as belonging to another class during the validation process: this observation has contaminated another class. The number of commission errors should be less than 15% for the Grassland Mask.

The standard error of the error rate estimate can be calculated as follows: $\sigma_h = \sqrt{\frac{p_h(1-p_h)}{n_h}}$ where n_h is the sample size for stratum h and p_h is the expected error rate. The 95% Confidence Interval is +/- 1.96. σ_h .

5 Prototypes Implementation

This chapter shows the implementation of the Improved Primary Status Grassland Layers (2017 and 2018) and the Grassland Change (2015-2018 and 2017-2018) prototypes in the three demonstration sites: South-East, Central and West. The status layers prototypes covering the demonstration site are provided with 10m spatial resolution and a minimum mapping unit of 0.05ha. The change detection products are produced with 10m spatial resolution and a minimum mapping unit of 0.5ha, except for the change product 2015-2018 that needs to be resampled to the HRL2015 product at 20m. Additionally, Grassland Mowing Intensity products are provided within the grasslands in the status layer 2018. The mowing intensity products represent intensively/extensively mowed grasslands as binary product. The intensive mowed grasslands are defined by three or more detected mowing events and the extensive mowed grasslands by less than three detected mowing events. For the grassland prototype, different technical approaches and the usability of different temporal and seasonal features are tested, as described in WP33 [AD07]. The grasslands within the demonstration site comprise natural, semi-natural and managed grasslands. Depending on the definition, natural grasslands could be available as proxy products, for instance being derived from the intermediate product of the number of mowing events detected. This is only possible under the assumption that natural grasslands occur where no mowing events have been detected. Nevertheless, both the context and the definition need to be carefully considered as well (e.g., leaving out grasslands within urban perimeters or pastures, where also no mowing events are detected). It has to be emphasized that for the identification of natural grasslands longer time series must be employed and further ancillary data is likely required for example with respect to the fertilization issue. Figure 5-1 provides an overview of the general workflow in the prototypes implementation, as a reminder of the methodology described in the final issue of WP33 [AD07].

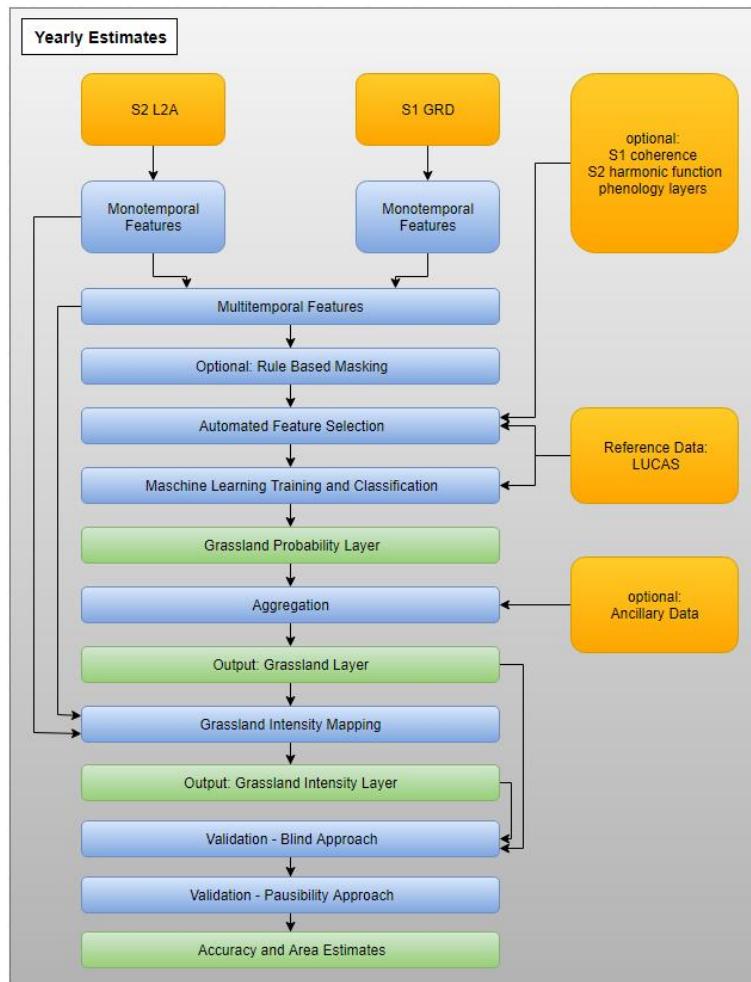


Figure 5-1: General workflow for grassland status mapping. Sentinel-2 (S2); Sentinel-1 (S1); Land Use and Land Cover Survey (LUCAS).

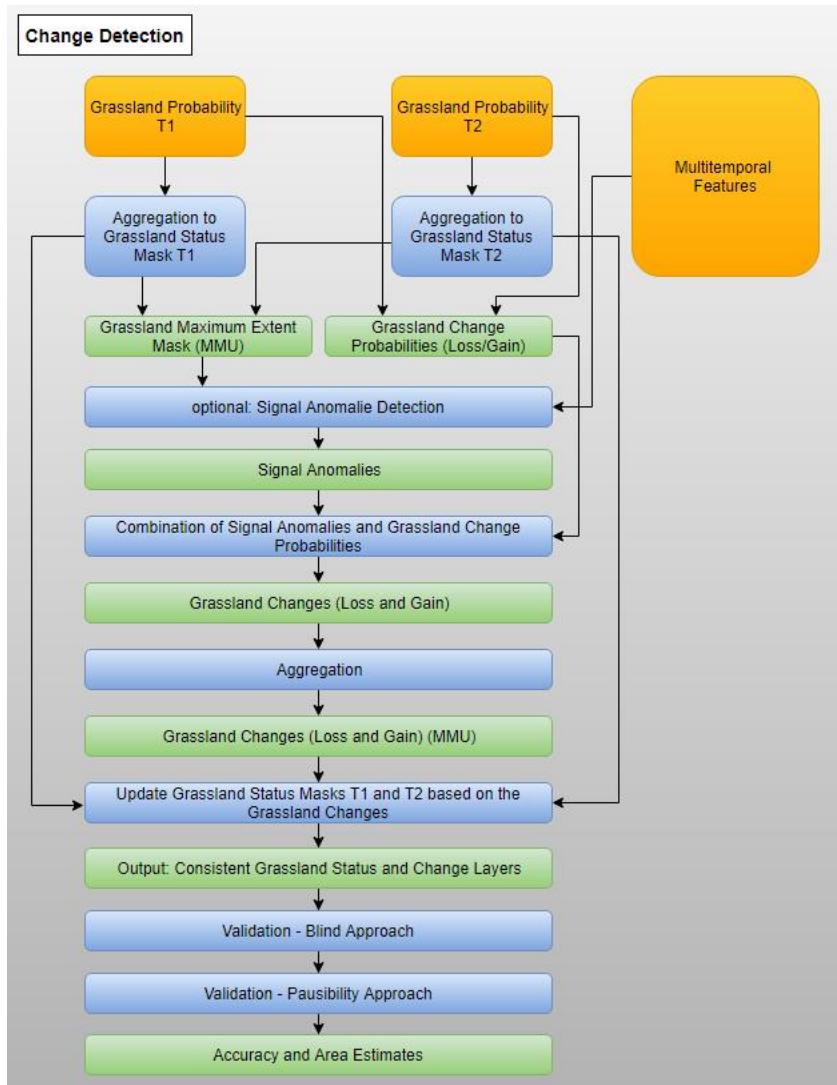


Figure 5-2: General workflow for grassland change detection.

Figure 5-1 and Figure 5-2 show the overview of the grassland status and change monitoring workflows. Respectively. The first major step, which has been conducted, is the processing setup for the prototype production, described in detail for the three different sites in Chapter 5.1. However, the occurrence of clouds specifically in Northern, Western and mountainous parts of Europe significantly reduces the density of the time series. Another drawback of using S2 data only is that there are similarities of spectral characteristics of different land cover types, which may lead to uncertainties in the resulting product. Secondly, Chapter 5.2 presents the results and the validation of the grassland prototype and lastly the description of the dataset properties and its metadata, referring to P43.2 – Data Sets of HRL Permanent Grassland Products is provided in Chapter 5.3.

5.1 Data and Processing Setup

Firstly, the pre-processing steps (Chapter 5.1.2), the integrated EO and ancillary data (Chapter 5.1.1) are described followed by the description of the prototype implementation in the different demonstration sites (5.1.3, 5.1.4, 5.1.5).

The developed processing chain is able to process a large amount of time series input data within a reasonable amount of time to provide the multi-temporal input variables for the classification process. The achieved level of automation ensures the effective application of the process for mapping at the pan-European level. The optical data proved to be more suitable to characterize the vegetation status and to exclude water covered objects and urban objects, while SAR imagery contributes e.g. the information to separate the grassland from cropland areas and supplement the S2 data in case of cloud

and cloud-shadow gaps as analysed in [AD07]. The data processing chain is operational for application at the continental level.

5.1.1 Input Data and Data Integration

SAR DATA

SAR backscatter characteristics are largely independent from weather and daytime conditions and provide complementary information to optical data for discrimination of vegetation types due to their diverse response to different polarisation signals. Compared to the optical data, microwaves are less affected by the physical-chemical characteristics of the surface, but rather by its structure such as geometry and roughness. Active energy scattered by vegetation is dependent on size, density, orientation and dielectric properties of elements compared to the size of radar wavelength helping to differ between roughness and moisture content (Rüetschi et al., 2017). SAR time series products used for the grassland classification were based on Level-1 products in Interferometric Wide swath (IW) mode provided as Level-1 Ground Range Detected (GRD) data. The IW mode is considered the main acquisition mode of Sentinel-1 over land and satisfies the majority of service requirements. Sentinel-1 data are processed by orbit because consistent acquisition geometries are fundamental to a number of pre-processing steps such as the multi-temporal speckle filtering, which can only be applied to images of the same orbit. Additionally, temporal image stack statistics have been calculated which are used as input data for the time series classification processing chains.

OPTICAL DATA

The Sentinel-2 sensor system has an overall number of 12 bands from 10m to 60m spatial resolution. For the ECoLaSS processing, only the 10m and 20m bands are used, which are in total 10 bands.

S1/S2 DATA INTEGRATION

The integration of S1/S2 allows benefiting from the multi-sensor characteristics, using the information gained from both sensors in parallel for the classification process as they record complementary characteristics of the land surface. The fusion on pixel level has been applied by stacking different S1/S2 features into one dataset, which is used for as input for the classification approach. A high geometric accuracy at sub-pixel level is required for this approach to avoid artificial errors being introduced to the fused data set, which means the image data set need to be resampled to common pixel spacing.

VHR DATA – DATA WAREHOUSE

Copernicus Contributing Mission ADDITIONAL datasets were requested via the Data Warehouse mechanism managed by ESA and an overall quota of 31,600km² has been granted to the ECoLaSS project for 2017, 2018, and 2019 [AD10]. The demonstration site West is represented by an overall extension of 1,608km² of VHR coverage, of which 909km² are archive data and 699 km² are new acquisitions. For the Central site overall 700km² of archive data and 596km² of new acquisitions were ordered, whereas for South-East 1,012km² of archive data were used (see Table 5-1). The ECoLaSS prototype generation is focussing on methods and prototypical developments based on multi-temporal Sentinel data and time series analysis. The VHR data was applied for the training as well as for validation. The VHR data sets provide additional information on grassland location, mowing dynamics and phenology in order to increase the separability of grasslands from agricultural fields and broadleaf forests.

Table 5-1: ADDITIONAL VHR1 data used for the GRA prototypes over the project years 2017, 2018 and 2019 sorted by type.

Category	West	Central	South-East
Archive_standard_Oprical_VHR1	909	700	1,012
New acquisition_standard_Optical_VHR1	699	596	0

The ordered VHR data partly cover the geographic areas where reference data is available. In order to cover the full phenology and dynamics of grasslands (approximate date and number of mowing events) VHR data for the vegetation season are needed. This is important since grasslands are much more dynamic than forests or urban areas, especially when crop rotation is applied in agricultural areas.

LANDBOUWGEbruikSPERCELEN ALV, 2016 (LGP) – REFERENCE POLYGONS:

The reference data set “Landbouwgebruikspelen ALV, 2016”(LGP) provided by the Departement Landbouw en Visserij is used for visual comparisons with the classification results. The dataset presents a polygon-wise assessment for the year 2016, differentiating between several agricultural areas including cultivation crops and grasslands. Since this reference data set was composed for agricultural purposes this reference data set does not include following grassland types:

- Grasslands in urban areas: parks, urban green spaces in residential and industrial areas, sport fields, golf courses
- Natural grasslands on military sites, airports
- Grasslands on land without use
- Semi-arid steppes with scattered Artemisia scrub
- Coastal grasslands, such as grey dunes and salt meadows located in intertidal flat areas with at least 30% graminoid species of vegetation cover

The LGP reference polygons were therefore only used for qualitative comparison with the grassland prototype.

IACS – INTEGRATED ADMINISTRATION AND CONTROL SYSTEM

As reference database for the production of the Mowing intensity Layer in the Central demonstration site the IACS data are essential for the production. The IACS data, also known as InVeKoS in Austria, contains a key attribute for the production of a mowing intensity layer. The number of mowing events can partly be derived from the attribute “ SNAR-BEZEI”. In the database the number of mowing events for all patches is not given, but different categories can be classified as extensive or extensive mowed and some categories include the number of mowing events. Unfortunately, this kind of information, is only available in the Austrian dataset that is covered by the determined test- and demonstration site. This leads to a restricted validation of the respective prototype in other sites.

LUCAS 2018

LUCAS 2018 data is used within the Grassland benchmarking and prototype generation. It is recommended to use only homogenous LUCAS points where the distance from the interpreter to the point location is less than 100m. Following attributes from the LUCAS 2018 Table can be used for the selection of homogenous points.

Table 5-2: LUCAS 2018 inclusion rules.

Inclusion Rules	Comments
"OBS_TYPE" = '1' OR "OBS_TYPE" = '3'	1: In situ < 100 meter 3: In situ PI
"CPRN_LC1N" >= 20 and "CPRNC_LC1E" >= 20 and "CPRNC_LC1S" >= 20 and "CPRNC_LC1W" >= 20	If the point is a Copernicus point it will be excluded if the homogenous LC is smaller than the circle with 20m radius.
"PARCEL_AREA_HA" >= 2	2: 0.1ha <= area < 1ha
"LC1_PERC" = 100	Land cover percentage

5.1.2 Pre-Processing

The ECoLaSS **West demonstration site** in Belgium is comprised of the footprints of six adjacent Sentinel-2 tiles (32UES, 32UER, 32UFS, 32UFR, 32UFQ and 32UEQ) for which Sentinel-2 and Sentinel-1 data were processed according to the outcome of WP32. The Sentinel-2 imagery has been atmospherically corrected and topographically normalized using the ESA Sen2Cor software. Furthermore, images with more than 90% cloud cover are excluded and an enhanced cloud mask using FMASK is applied to the remaining images.

For each Sentinel-1 orbit, the pre-processing is calculated separately as multi-temporal filtering can only be applied to images of the same orbit. The S1 time-series images are radiometrically calibrated, radiometric terrain corrected, resampled to 10m and orthorectified by applying the Remote Sensing Software Package Graz (RSG), which is developed by Joanneum Research. Based on pre-processed time series of VV and VH images the temporal image stack statistics are calculated. Next, the SAR images are temporally filtered by a multi-temporal speckle filtering approach reducing the speckle noise (3x3 kernels).

The large amount of scenes with strong cloud cover in the time series reinforces the need for the use of image composite-like time features. Different multi-temporal features based on the S1 and S2 time series from the reference year are generated. The annual features have the advantage of having included more observations to reduce biases from missing observations, whereas the seasonal features are useful to discriminate the phenological differences between vegetation types. All available images are reduced to temporal statistics including the metrics explained in the following Chapter 4.2.

In the **Central (Austria/Germany)** site the Sen2Cor package is used for pre-processing of optical data (i.e., Sentinel-2 scenes) including following steps atmospheric correction, topographic normalisation, cloud and cloud shadow detection and masking using an adapted Sen2Cor SCL mask, setting the cloud filter to below 90% cloud cover, as described in WP32 [AD06]. Another major output of the processor is the scene classification layer (SCL), which detects 12 different land cover classes. The SCL is used for cloud and cloud shadow masking during the classification process. The cloud cover metric used for analysis and classification does not rely on the official metadata cloud value provided by the original Sentinel-2A product, but is derived from the Scene Classification produced by Sen2Cor. The resulting cloud-mask layers can later be used to optimize the Sentinel-2 input data for the feature generation step and will ensure optimised input data for subsequent thematic classification steps. The affected areas in the specific satellite image are masked out so that the respective pixel values are not considered in the calculation of (time-) features.

As temporal features were calculated out of the time series (e.g. max, mean, min, etc.) and gaps were filled with scenes from the time series stack automatically, no gap-filling and noise reduction has been applied. The Sen2Cor output bands are automatically resampled to 10m. In particular, all 20m bands are re-sampled using the cubic method whereas SCL is resampled using nearest neighbourhood. The three bands with 60m spatial resolution (Bands 1, 9 and 10) are omitted in the Level-2A output since they are not needed for land cover applications. From the surface reflectance products, spectral indices are derived (e.g., NDVI, NDWI, Brightness, IRECI).

This approach proved to be sufficient in order to reduce the amount of artefacts in the time feature computation by enlarging cloud and shadow areas. It has been implemented by Python scripting and convinces by its high performance (in spite of the limitations reported in [AD06]).

In turn, the pre-processing of SAR data (i.e., Sentinel-1) is performed with ESA SNAP, as described in [AD06]. SNAP is a common architecture for all Sentinel products. The processing steps applied for the test site comprise the orbit update (includes automatic precise orbits download), thermal noise removal, radiometric calibration generating a beta band, terrain flattening to gamma naught based on SRTM 1sec HGT, terrain correction using the same DEM generating a 10m resolution product and the export of the scene in DN units. Then up to 10 scenes from the same orbit/slice are co-registered/stacked in WGS UTM 32N projection. After the orthorectification of the images, the multitemporal speckle filtering has been applied using a Frost 5x5 window. Python scripts are used in indices derivation: VV (Gamma naught), VH (Gamma naught), Normalized Difference VV/VH, Ratio VV/VH. A geometric consistency monitoring step is applied to automatically check for shifts within the co-registered scenes.

5.1.3 Demonstration Site West

In the demonstration site West, the grassland status layers are generated for the reference year 2017 and 2018. Further, a grassland change layer is produced between 2017 and 2018 for detecting areas of grassland loss and grassland gain. Additionally, the grassland mowing intensity layer is derived within the grassland status mask 2018, where the mowing events for managed grasslands and semi-natural grasslands are detected.

5.1.3.1 Grassland Identification

The methods and workflows tested and described in the final issue of WP33 D33 “Time Series Analyses for Thematic Classification” [AD07] have been implemented for the demonstration site West (Belgium/France) in phase 2. Training samples were generated based on the LUCAS 2018 samples where homogenous samples are selected through database sections queries as described in Chapter 5.1.1. In the West demonstration site, a total of 7136 LUCAS samples were available. Through selection for homogenous points 3189 are left and can be used for training and internal quality tests of the classification models the interpreted points were randomly split into training and validation sets at a ratio of 65% training to 35% validation. With aggregated classes (grassland and all other land cover) the random forest model has been trained using the best suited temporal and spectral variables selected by a Random Forest feature selection Gini Measure, as described in WP33 [AD07]. The number of trees has been set to 1000 and the number of variables to the square root of the total number of input variables.

5.1.3.1.1 Predictor Candidates

The input data finally used for the prototype generation are 32 Sentinel-1 and 320 Sentinel-2 time series derived statistical features over the vegetation period from April until September. In the West demonstration site the vegetation period could be further extended to March and October, due to the mild climatic conditions. Nevertheless, it is not recommended to extend the monitoring period to March and October, since March can be biased by remaining snow areas and in October, larger shadows occur due to the lower sun incidence angle. Time features computed for the annual features are percentiles (10, 90), median, coefficient of variation and the standard deviation. On the basis of phase 1 outcomes and tests for grassland, the seasonal time windows included are the spring period (01.04-31.05), summer (01.06-31.07) and the late summer and autumn (01.08-30.09) for 2017 and 2018. Only the median statistic is used to calculate the seasonal features, since they are robust in terms of data availability. Choosing longer time periods would benefit the data quality since more observations can be included, but for the Grassland thematic, longer periods would not capture the different phenological differences. In the case of the West Grassland prototype, the added value of incorporating complex time features like S2 harmonic regression parameter or S1 coherence features was insignificant. The data availability is

essential and in fact, the performance and reliability of the approach depend on the availability of dense time series.

5.1.3.1.2 Feature Selection

Using a random forest-based classification approach with more than 400 features, a feature selection is performed to identify the most important input features for classification, in relation to product quality as well as a cost and time efficient processing. The above described features served as input for the features selection analysis. In a first step all available features are analysed with regards to the model performance. Figure 5-3 shows the variability of the OOB error in comparison to the number of input features provided to the RF classifier. Around 40 to 50 features should be used, using less features drops the model performance significantly.

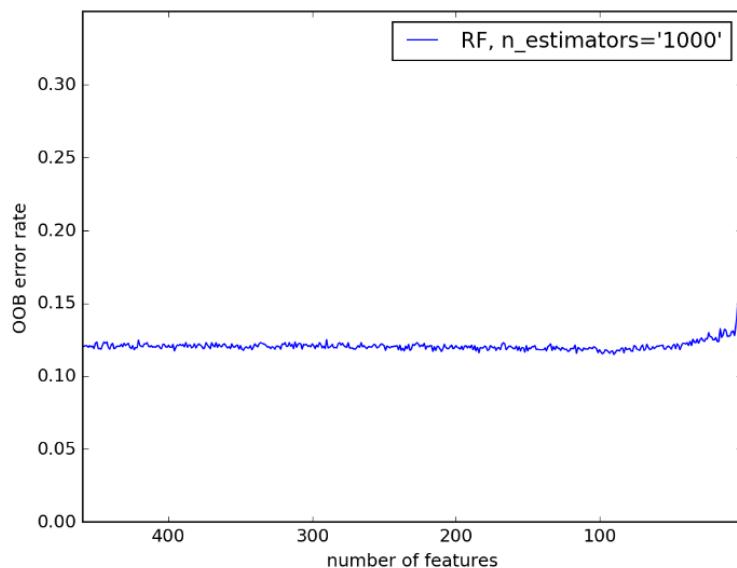


Figure 5-3: OOB error in relation the number of S1 and S2 input features.

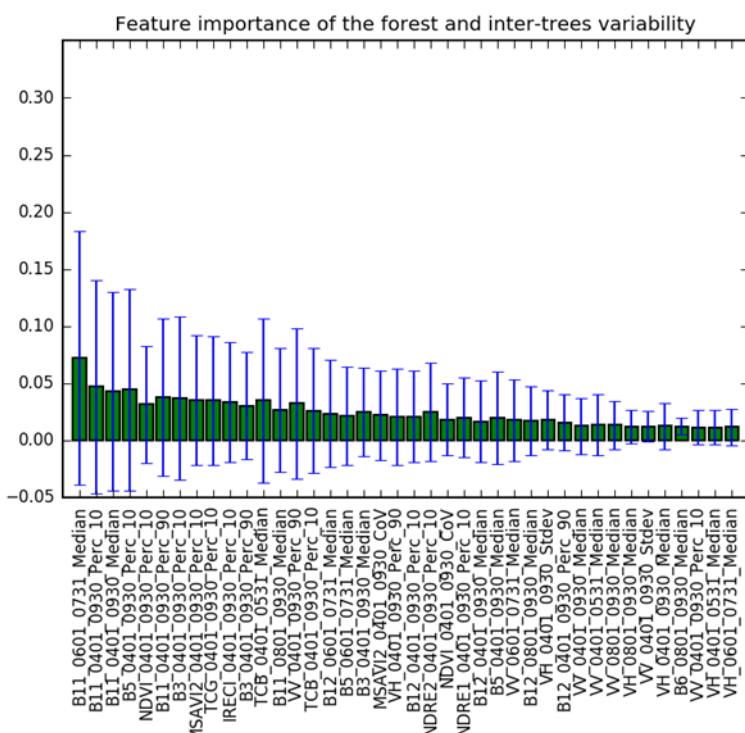


Figure 5-4: Top 40 S1 and S2 features for the grassland discrimination.

Figure 5-4 presents the 40 included time features used in the production including several annual S1 and S2 features. Overall more S2 features are needed than S1 features. The reason could be the geometric problems at feature borders and the remaining noise within homogenous regions (see Figure 5-18 and Figure 5-19).



Figure 5-5: LPIS Grassland polygons (2016) based on S2 image (R: B11_0601_0731_Median G: B11_0401_0930_Perc_10 B: B5_0401_0930_Perc_10)



Figure 5-6: LPIS Grassland polygons (2016) based on S1 image (R: VV_0401_0930_Perc_90 G: VH_0401_0930_Perc_90 B: VH_0401_0930_Stdev)

The feature selection tests automatically executed by the Random Forest Classifier have shown that, in the West demonstration site, best performing input indices are NDVI, MSAVI2, IRECI and TCB. Regarding SAR data generally, both polarisations (VV and VH) as annual and seasonal coefficients have shown the best performance. It is important to define time windows where grassland and cropland are best separable.

5.1.3.1.3 Classification and Post-Processing

Several input feature combinations and their respective feasibility, effort and accuracy: Sentinel-2, Sentinel-1, and the combination of Sentinel-1 and Sentinel-2 temporal features have been tested and the results are described in WP33.

A final aggregation of the pixel-based classification results has been applied. All patches smaller than the MMU of 0.05ha (5 pixels) were removed.

In addition to the prototype status layers, probability layers are provided as a pixel-based quality indicator / expert product. The additional 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.

5.1.3.2 Grassland Change Monitoring

For the grassland change monitoring a signal anomalies detection approach has been combined with a map-to-map change detection approach. For the signal anomaly detection the change vector analysis approach has been adopted as described in detail in WP34 [AD08]. The change vector is the difference between vectors related to selected normalised features at different years. The selection of temporal features was thereby based on a qualitative assessment by a remote sensing expert. As grasslands are

highly dynamic ecosystems, where the characteristics of the temporal trajectories depend on the management in the respective year (e.g. varying number and timing of mowing events or fertilisation), only features were selected which are “integrated” over the whole vegetation season. Based on a qualitative assessment of temporal trajectories, following features were selected for the change detection: (i) median of NDVI, (ii) median of reflectance in the red band, (iii) median of the reflectance in the green band, all calculated over the whole vegetation season. The change vectors were normalised, and the lengths of the change vectors measured as Euclidean distance are used as anomaly indicators. As the temporal density of the S2 time-series varies between 2017 (Satellite Sentinel-2B data only available from July 2017 onwards) and 2018, features based on median calculation over the vegetation period were selected, as these “robust” features are less sensitive to temporal density differences. To retain the high spatial resolution, only temporal features derived from bands with 10m spatial resolution were selected. The pixel-based change detection results were finally aggregated according to the nomenclature specifications to provide a consistent set of grassland status layer 2017, 2018 and changes. As specified, the Minimum mapping units for grassland status maps is 0.05ha and for the change layer 0.5ha.

5.1.3.3 Grassland Mowing Intensity Mapping

The grassland mowing intensity has been derived by estimating the number of mowing events from Sentinel-2 time series. A Kalman filter approach is used to track the signal levels of the Tasseled Cap Components Brightness, Greenness, and Wetness through time on pixel-level. The method assumes that the removal of healthy biomass after a mowing event causes an abrupt drop of the Greenness signal on the one hand, but also a drop of the Wetness level because of the increased soil reflectance. Thus, the implemented algorithm signals a mowing event if a statistically significant change vector in the two-dimensional feature space created by Greenness and Wetness is detected and its direction corresponds to a drop of both variables. The statistical significance is influenced not only by the vectors' magnitude, but also by the length of the time gap between consecutive observations. Large gaps in the time series will result in a lower sensitivity of the detection method, because the algorithm has not enough information to distinguish between abrupt and gradual signal changes. A detailed description of the method is given in WP33 [AD07]. Since no mowing event reference data was available for the demonstration site West, the prototype is to be considered merely a proof-of-concept. For operational production, local expert knowledge and/or in-situ training data, like InVeKoS data in Austria is required. A sample of the result is given in Figure 5-24. A 3-by-3 majority filter has been applied to the raw pixel based results. The map features distinct borders between extensively and intensively managed patches.

5.1.4 Demonstration Site Central

In the following chapters the data and processing setup is explained in detail for the Grassland Identification (5.1.4.1) including the Predictor Candidates (5.1.4.1.1), the Feature Selection (5.1.4.1.2), and the Classification and post-processing (5.1.4.1.3). Further, the setup for both Grassland Change Monitoring (5.1.4.2) and the Grassland Mowing Intensity Mapping (5.1.4.3) is elaborated.

5.1.4.1 Grassland Identification

The methods and workflows tested and described in the final issue of WP33 “Time Series Analyses for Thematic Classification” [AD07] have been implemented for the demonstration site Central (Austria/Germany) in phase 2. Training samples were generated based on the LUCAS samples through an outlier detection and an independent visual interpretation with Sentinel-2 time series data and additional VHR data if available [AD07]. Thus, in the Central demonstration site, a total of 3664 samples were used, derived from the HRLs 2015 reference maps and LUCAS 2018 data filtered by the Observation Type attribute, in order to be sure the points had been interpreted from a reasonable distance limit. Furthermore, some manual samples were added during the iterative classification process. In automatic sampling, outlier detection is a key step, as described in Task 3 [AD07]. Accordingly, an outlier detection based on spectral signals was carried out by means of applying zonal statistics for 30x30m samples. For

training and test of the classification model the interpreted points were randomly split into training and validation sets at a ratio of 50% training to 50% validation. With aggregated classes (grassland and all other land cover) the random forest model has been trained using the best suited temporal and spectral variables selected by a grouped forward feature selection, as described in WP33 [AD07], and below in subsection Feature Selection. The classifier parameter number of trees was set to 250.

5.1.4.1.1 Predictor Candidates

The input data finally used for the time feature computation (input for grassland demo production) were Sentinel-2 bands 3, 11 and 8A, as well as derived indices NDVI, NDWI, IRECI and Brightness and Sentinel-1 VV and VH bands plus normalized difference of VV and VH (NDVVVH) and the ratio of VV and VH (RATIOVVH). Based on phase 1 outcomes and tests for grassland, the time windows testing included the spring period (01.03-31.05), late winter to spring (01.01-01.06) and a larger period (01.01-30.11) for 2017 and 2018 time series. Time feature computed for the input data were percentiles (10, 25, 50, 75, 90, difference 90-10 and difference 75-25), minimum, maximum, mean and standard deviation. In the case of the Central Grassland demo, the benefit of incorporating complex time features was insignificant. Throughout the tests, the most relevant features were the percentiles.

In 2018, the data situation is not optimal due to long lasting snow cover in high altitude areas and a high cloud cover throughout the year. Additionally, imperfect cloud mask caused artefacts in time feature extraction, wherefore corrections to the cloud masks were applied. Therefore, data availability is essential and in fact, the performance and reliability of the approach depend on the availability of dense time series.

5.1.4.1.2 Feature Selection

The generation of suitable time features, especially considering upscaling and in particular the pan-European or global roll-outs, is challenging and requires large computational capacities. Indeed, many reasonable combinations of time series metrics, sensor bands, indices and suitable temporal windows are conceivable, leading to a potentially quite large number of features. The testing experiences in Task 3 for grasslands in Central contributed to the definition of the classification parameters and proved the temporal windows and features that were performing best. In this regard, the Random Forest classification algorithm provides information about feature importance.

In parallel to the new feature calculation and analysis, the grouped forward feature selection method is applied. The grouped forward feature selection method adapted and embedded in the Random Forest classification process is based on the sequential feature selector integrated in the machine learning package (python module scikit-learn in the machine learning extension MLxtend).

This sampling method is used to reduce an initial d-dimensional feature space to a k-dimensional feature subspace where $k < d$. The goal of feature selection is two-fold: improve the computational efficiency and reduce the generalization error of the model by removing irrelevant features or noise. This wrapper removes or adds one feature at the time based on the classifier performance until a feature subset of the desired size k is reached. The difference with other methods like the Recursive Feature Elimination, is that the latter is computationally less complex using the feature weight coefficients (e.g., linear models) or feature importance (tree-based algorithms) to eliminate features recursively whereas the forward feature selection eliminates or adds features based on a user-defined classifier/regression performance metric. The algorithm finally yields a combination of the features with the highest accuracy. This subset of features is used for the classification process.

5.1.4.1.3 Classification and Post-Processing

Classification was applied using different combinations of sensor data and time periods to benchmark their respective feasibility, effort and accuracy: Sentinel-2, Sentinel-1, and the combination of Sentinel-1 and Sentinel-2 temporal features.

Using a random forest-based classification approach on several hundred temporal feature descriptors, a feature selection is performed to identify the most meaningful input features for classification, ensuring

a high resultant product quality as well as a cost and time efficient processing with high accuracy. For the Central demo, 156 time features were finally selected. A probability layer indicating the percentage of reliability constitutes a pixel-based quality indicator by-product.

A final post-classification filtering and case-wise elevation thresholds (e.g., in the agriculture crop mask prototype, based on expert knowledge in the region, an altitude threshold of 1700m was applied to leave out highland grasslands, as no crops are found in this regions above that altitudes) were applied to the grassland classification. In particular, in the case of the grasslands identification, expert knowledge and regional tuning was required: applying a threshold of 2800m altitude contributed to avoid commission errors detected in the areas permanently covered by ice (e.g., glaciers). All patches smaller than five pixels, i.e. 5 pixels minimum mapping unit (MMU), were removed to close holes in grassland patches and remove very small patches of grassland. The final MMU for the status layers for 2017 and 2018 grasslands masks is therefore 0.05ha, improving the tests and demos conducted in phase 1 regarding accuracy, as well as regarding look and feel of the products.

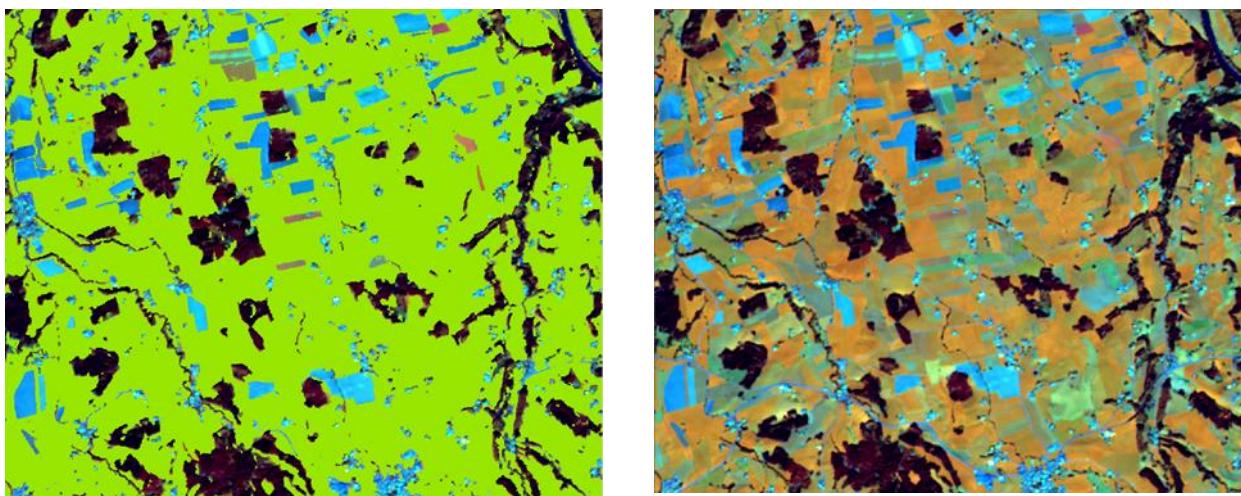


Figure 5-7: Grassland mask 2018 Central demonstration site detail,

Left: S2 band combination B08-B04-B11 overlaid with classification result for 2018 (grasslands in green colour).

Right: S2 band combination B08-B04-B11 (bluish areas are ploughed agricultural fields, grasslands in orange colour).

The status layers and corresponding accuracy metrics are detailed in the error matrices provided in Chapter 5.2.1.2. The Figure above shows a detailed view of the grassland mask 2018, reflecting how well the separation between grassland and cropland worked. On the left, grasslands appear in green and ploughed agricultural fields appear bluish. On the right, grassland appears in orange, forest in very dark red-brown, cropland appears reddish and bare soil appears bluish in the S2 band combination B08-B04-B11 (i.e., NIR-SWIR-Red).

The consistency between the ECoLaSS prototypes produced in Central for the forest, grasslands and agriculture land covers was further verified. An overlap between the forest, agriculture and grasslands prototypes in Central is shown in Chapter 5.2.1.2.

5.1.4.2 Grassland Change Monitoring

In phase 2, two different approaches have been applied in demonstration site Central for the grassland change detection for the reference years 2015 (HRL2015 grassland) and 2018 (ECoLaSS grassland status layer 2018) on the one hand, and on the other hand for the reference years 2017 and 2018 (ECoLaSS grassland status layer 2017 and 2018 respectively) [AD07].

In the first case, the approach is rather simple map-to-map comparison because no pre-processed S2 data was available for 2015 to derive further information on possible changes. This is likely to be the case in other circumstances when comparing already produced land cover maps with new classification

products. Except the application of a boundary filter that reduced the effect resulting from the different resolutions of the status layers (2015: 20m; 2018: 10m) and the geometric inaccuracies of the EO data, no further post-processing was applied to this change layer.

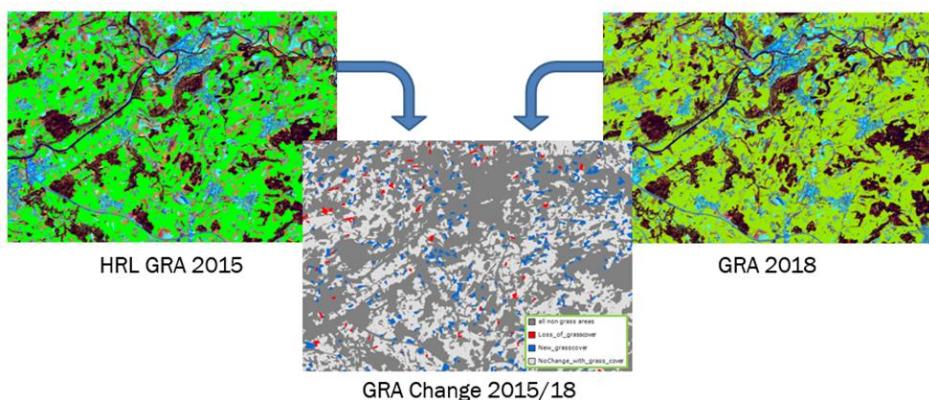


Figure 5-8: Map-to-map change detection in Central demonstration site for the reference years 2015 (HRL2015 GRA)-2018 (ECoLaSS grassland mask 2018).

The map-to-map change layer 2015-2018 implies first computing the geometric difference (layer a) between the HRL2015 grassland status layer and the ECoLaSS Grassland 2018 demonstration site produced in the Central region (the latter resampled to 20m). The different combinations of detected change in this raw change layer (e.g., grassland to crop, crop to grassland, etc.) were reclassified to a binary change/no change layer (layer b) in step 1. Subsequently, a boundary filter is applied to reduce the artefacts resulting from the different spatial resolutions and potential positional shifts of the Sentinel data: no change areas are applied a 1 pixel buffer, then a filter of a single pixel reduces potential noise (step 2). A buffer of 1 pixel is then applied to the remaining change areas (step 3). This last step brings meaningful change areas back to the original size. The Figure below shows this sequence:



Figure 5-9: Boundary filter steps applied to the raw change detection recoded layer (b) in the map-to-map change detection (reference years 2015-2018) in the Central demonstration site.

Finally, the result of these steps (layer c) is combined with the raw change raster (layer a). The pixels are recoded to gain and loss where the according classes from the raw change layer overlap with the valid change areas in layer c, to unchanged grassland where both years show grassland and to non-grassland areas where both classifications did not show grassland. An example of this process is depicted in Figure 5-10.

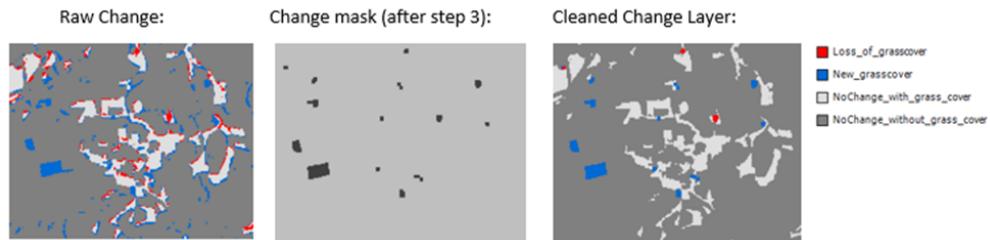


Figure 5-10: Example of the combination of the raw change layer (a) with the boundary-cleaned layer (c) in Grassland change detection (2015-2018).

A final filter to delete all areas below 1ha is applied because it must be taken into account that the HRL2015 MMU is 1ha. Smaller filters would not make sense from a technical point of view, as the result would be the “creation” of change areas where HRL2015 by definition could not show grassland or non-grassland.

In turn, the change layer 2017-2018 computation starts from the difference between the corresponding status layers (layer a), subsequently followed by the same recoding to generate a binary change/no change mask (layer b) and the same boundary cleaning as explained for the 2015-2018 change detection layer.

To improve the accuracy and to make the change layer more meaningful in terms of real changes, an additional change indicator is computed from the NDVI time series of both reference years. Therefore, the NDVI minimum temporal feature is computed for 2017 and 2018 respectively, and the difference between them is derived. The resulting difference raster (called diff_NDVI_min in this context) is then recoded according to a threshold, shown in this formula:

Change is indicated where:

$$\text{diff_NDVI_min} < (\text{mean}(\text{diff_NDVI_min}) - 1.5 * \text{std}(\text{diff_NDVI_min})) \text{ or} \\ \text{diff_NDVI_min} > (\text{mean}(\text{diff_NDVI_min}) + 1.5 * \text{std}(\text{diff_NDVI_min}))$$

All other pixels are assigned as no-change. This layer (d) implies change detection is located in areas where there is a significant difference in the minimum NDVI of both reference years.

Then, layer d is combined with the boundary-cleaned raw change mask (c), and change areas are identified where both layer d and layer c are indicating a change. All other areas are classified as no-change.

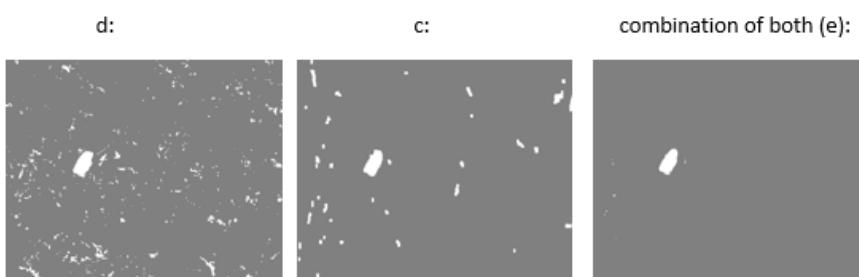


Figure 5-11: Example of the combination of the NDVI minimum threshold recoded layer (d) with the cleaned change layer (c) in Grassland change detection (2017-2018).

In the next step, the grassland raw change layer is re-classified using the combined change mask. All unchanged areas (grassland as well as non-grassland) remain untouched, whereas loss and gain areas are only adopted, when the combined change mask indicates change. This procedure is shown in Figure 5-12.

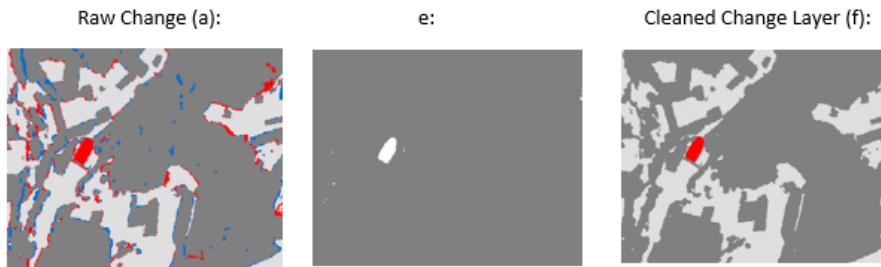


Figure 5-12: Example of the combination of the combined layer (e) with the raw change layer (a) in Grassland change detection (2017-2018).

Last, a final filtering of 0.5ha is applied. This parameter has been defined on the basis of the tests carried out in Task 3. In our experience, areas below this MMU are often due to small differences between the two status layers (where one maybe was more accurate than the other). Instead, above half a hectare changes have proved to be more meaningful. Figure 5-13 shows a summary of the whole approach again to get a better overview.

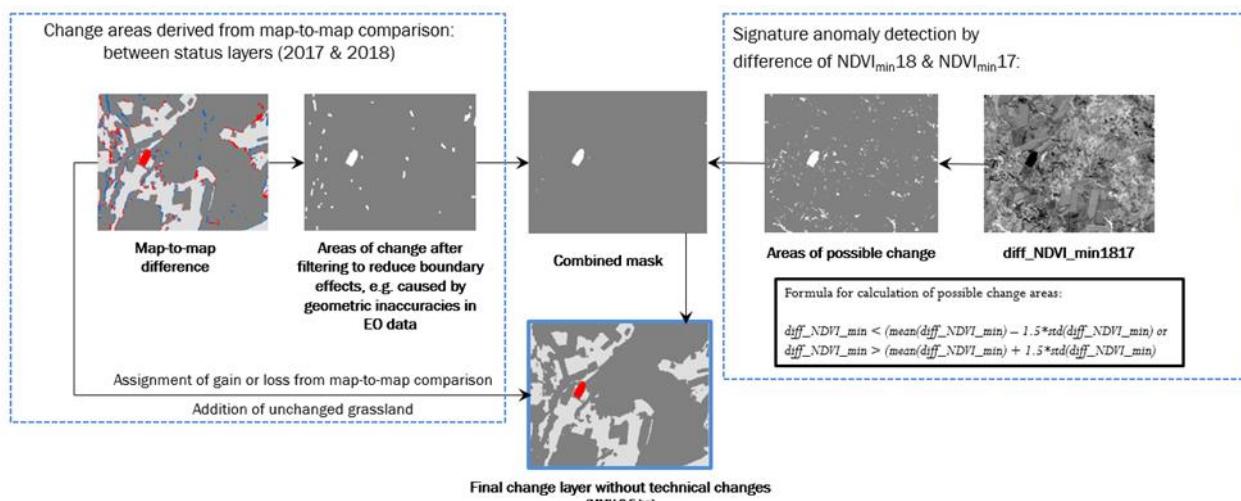


Figure 5-13: Summary of the change detection approach for the Grassland reference years 2017-2018 in Central demonstration site.

5.1.4.3 Grassland Mowing Intensity Mapping

In Task 3, tests on a new product were carried out which should provide more details for the detected grassland areas, regarding the mowing intensity. The resulting mowing intensity layer at 10m spatial resolution is based on the number of mowing events detected. For this purpose, NDVI time series are used to derive a layer showing the number of mowing events. This layer is then clipped to the grassland mask and re-classified by defining the extensive use category when less than or equal to two mowing events are detected during the year.

The intermediate product provided is the number of mowing events. From this layer, other products can be subsequently derived. For instance, a proxy of natural grasslands can be approached and estimated by considering areas where no mowing events have been detected (i.e., zero mowing events detected). As described in WP33 [AD07], at first the coherence features from SAR data were tested, although according to the accuracy versus performance benchmarking (e.g., computation costs, product quality and timeliness, etc.) and taking into account the upscaling of the products to larger scales in a cost-efficient manner, another approach was selected instead for implementation in the demonstration site. In addition, it was concluded that coherences are highly sensible to changes, even on micro-level, and therefore, events like heavy rainfall are likely to make coherence images unusable for intensity analysis. This is highly risky, besides the expense of the processing of SAR coherences, when considering

automation and large scale products. Consequently, in Central the approach based on NDVI time series has been applied.

NDVIs were computed for all scenes available in 2018 to detect mowing events all throughout the year by subtracting consecutive NDVI acquisitions and a rule based classification, defining that all pixels with ≥ 3 mowing events are intensively used and 0-2 mowing events means extensively used, in terms of mowing events. The following Figure sketches the workflow applied in the mowing intensity layer generation:

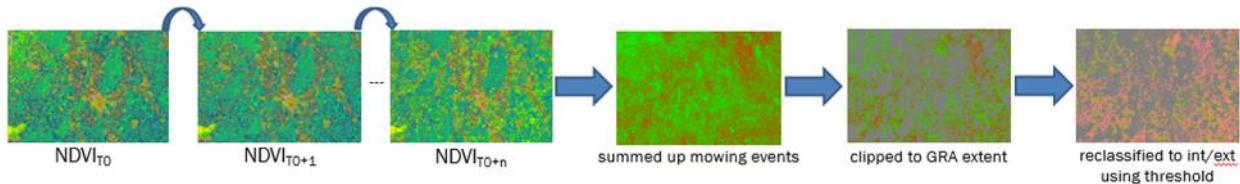


Figure 5-14: Intensive/Extensive grassland use layer workflow.

As for the other grassland layers, filtering improves significantly the look and feel by reducing noise. For the mowing intensity layer, a filter of 4 pixels in size was applied. All areas within the grassland mask where filtered, so that there is no patch for one of the two intensity classes smaller than 5 pixels in the end. Within small grassland patches, it might happen that e.g. 3 pixels are classified as extensive and 2 are classified as intensive. In such cases, the filter would cause the class values to jump between classes with each filter iteration without getting a patch of 5 unique values. If so, it was filtered in favour of intensive use because most of the areas are used intensively in the demonstration site. In any case, the number of mowing events layer, which is the previous step to the binary extensive/intensive use decision, is available for consultation. This layer is also useful to check for natural grasslands if it is assumed that the latter are present when no mowing events are detected at all. For this assumption to be more reliable, a longer time series (e.g., several years) should be considered.

The Figure below shows the test in Central for the grasslands mowing intensity in 2018.

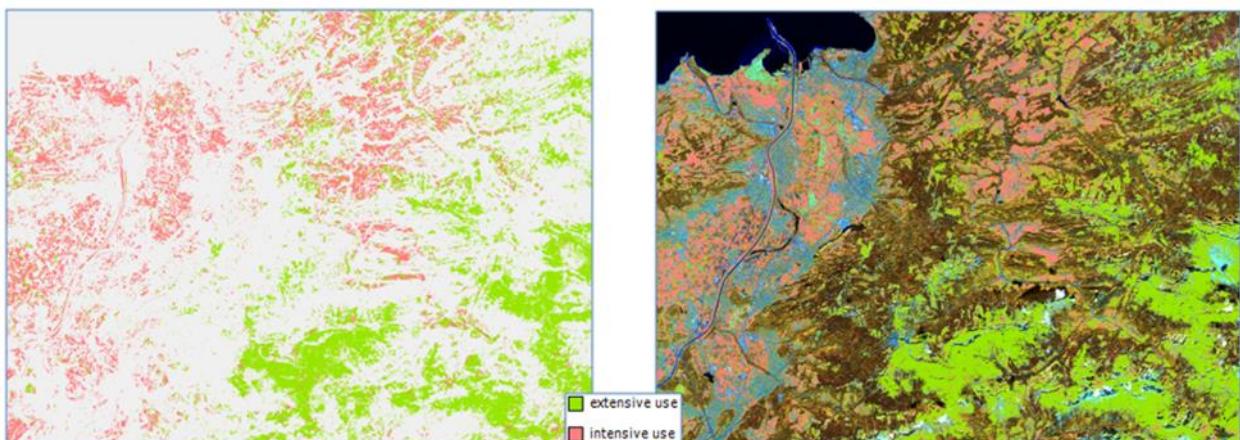


Figure 5-15: Grassland mowing intensity in Central 2018.

In Figure above, and in Figure 5-16 it can be observed that grasslands are extensively managed in Alpine regions whereas more intensively in valleys around settlements.

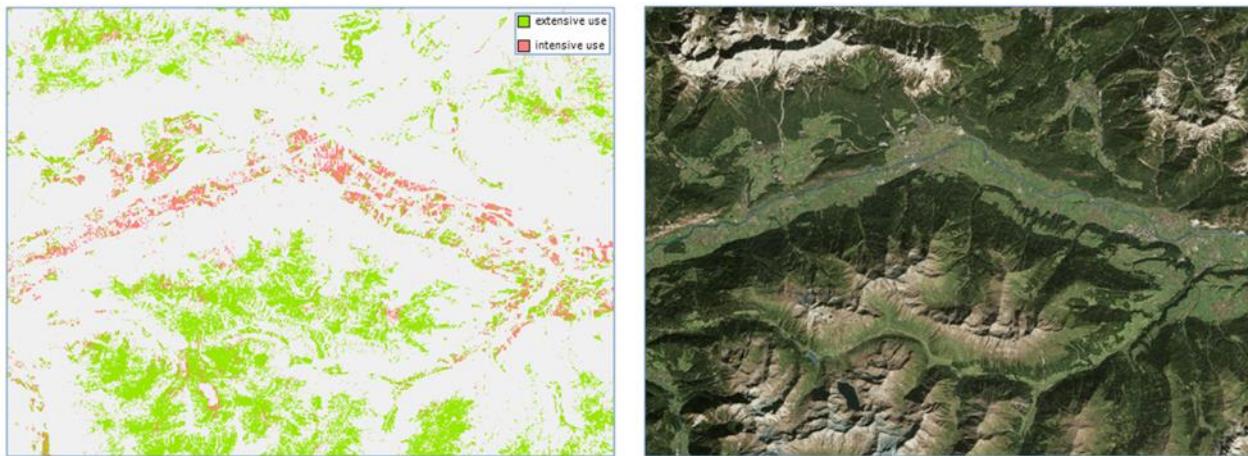


Figure 5-16: Detailed view of the grassland mowing intensity layer compared to Bing Maps Aerial of the same region, showing more intense management is concentrated in valleys around human settlements.

5.1.5 Demonstration Site South-East

5.1.5.1 Grassland Identification

Grassland identification in the South-East demonstration site was accomplished as described in Chapter 4. LUCAS 2018 points were used as training data and for internal validation. Predictor sets were a) optical-only (S2); b) SAR-only (S1); as well as the combination of both optical and SAR based predictors (S1/S2).

In total 3871 LUCAS samples were available in the demonstration site, of which 743 belonged to the grassland classes. Given that not all countries in the demonstration site are members of the EU and hence are not covered by the 2018 LUCAS sampling, i.e. Serbia and North-Macedonia, training sample size was comparatively small and available only for parts of the prototype demonstration area. LUCAS data were filtered by the following criteria:

- Observation type = In situ (LUCAS variable: OBS_TYPE)
- Homogenous landcover within 20m radius (LUCAS_variable: CPRN_LC1*)
- Dominant landcover percentage = 100% (LUCAS variable: LC1_PERC)
- Parcel area > 0.1ha (LUCAS variable: PARCEL_AREA)

After filtering, 2168 LUCAS samples remained (482 grassland samples), of which 25% were set-aside for internal validation. The LUCAS land-cover classes were then converted into binary form, that is, “grassland” and “non-grassland”. Multi-class modelling, that is, mapping the classes “grassland”, “cropland”, “forest”, and “other” separately was conducted as well, yet, presumably due to sample size limitations, this led to worse performance considering the class “grassland” than in the binary mode, hence this avenue was not pursued further.

Modelling was accomplished with the random forest algorithm first proposed by Breiman (2001). The number of trees was set to 1000, although performance did already stabilize at around 500 trees in most cases. The number of predictors to evaluate per split was set to the square root of the total number of predictors. Initial attempts at tuning this parameter did not improve the results. Apart from the binary grassland – non-grassland map, a probability layer was produced based on the fraction of “grassland” votes of each tree in the random forest.

5.1.5.1.1 Predictor Candidates

For the south-east demonstration site on average 83 Sentinel-2 acquisition dates throughout the entire year 2018 were available per granule with cloud cover less than 80%. Sentinel-2 data were processed

from Level 1C to Level-2A using Sen2Cor. Cloud and cloud-shadow masking was accomplished with the FMASK software as described in [AD06], which was found to perform superior in the study area compared to the SCL mask provided by Sen2Cor.

From the S2 Level-2A data the following spectral indices (see AD06 for formulae) were calculated for each acquisition date: NDVI, GNDVI, NDWI, NDRE1, NDRE2, MSAVI2, mean SWIR; IRECI, CI_red_edge, PSRI, REP and MCARI, Tasselled Cap Brightness, Tasselled Cap Wetness and Tasselled Cap Greenness. S1 data were pre-processed with ESA SNAP from the GRDH product to gamma naught for VH and VV polarisations, as described in Chapter 5.1.2, including multi-temporal speckle filtering.

Based on the single-date spectral bands, spectral indices and backscatter coefficients a suite of multi-temporal metrics were derived for relevant time-periods throughout the vegetation period. The statistics derived were the 10%, 50% (median) and 90% percentiles, the mean, the standard-deviation as well as the coefficient of variation. The aggregation time-periods which were chosen were:

- March – October for capturing the yearly characteristics during the vegetation period;
- Bi-monthly intervals: March – April, May – June, July – August and September – October;
- Tri-monthly intervals March – May, June – August and September – October.

Overall, 1200 Sentinel-2 and 66 Sentinel-1-based time-series predictors were derived, resulting in a total number of 1266 potential predictors.

5.1.5.1.2 Feature Selection

Based on this pool of 1266 potential predictors the random forest Gini-based feature importance was used in a backward, element-wise feature selection procedure, that is, by successively removing the least important feature. During feature selection the data-set was randomly subsampled in order not to over fit to the training data-set. Eventually the feature combination resulting in the highest out-of-bag F1 score for the grassland class was chosen for the final model fitting. For S1 the optimal set comprised 24 predictors, for S2 optimal performance was achieved with 39 predictors. The performance of the combined S1/S2 model was superior throughout and peaked at 36 predictors.

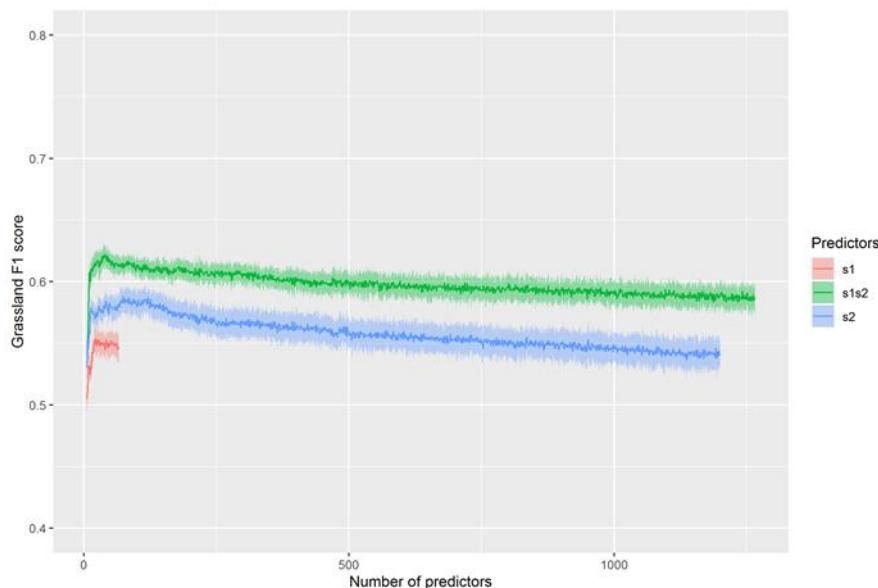


Figure 5-17: Development of out-of-bag grassland F1 score during Gini-based backward feature selection for the predictor sets: Sentinel-1 (S1), Sentinel-2 (S2) and combined features (S1/S2).

The observation that model performance increased with a decreasing number of predictors (Figure 5-17) is an indication of a sample size limitation. Even though the random forest algorithm is to a large degree robust to very high-dimensional feature spaces, this behaviour shows that it is likely that an improvement could have been possible had more training samples been available.

While attributing an absolute importance to each predictor is not straight forward due to their high collinearity, the ability to reduce the number of features which need to be computed beforehand without compromising mapping performance is the most important outcome. From the perspective of an operational implementation, a valuable observation is that the number of input features may remain moderate around at maximum of 50 predictors. The final set of predictors used for the demonstration site South-East production is listed in (Table 5-3).

Table 5-3: Final variables used for Grassland status layer production in demonstration site South-East. Temporal statistics: q10 = 10% percentile, q50 = 50% percentile (median), q90 = 90% percentile, sd = standard deviation, cv = coefficient of variation. TC green = tasseled cap greenness component, TC wet = tasseled cap wetness component.

Type	Variable	Time period (months)	Statistic
S2, spectral band	Band 3	03-04	q50
S2, spectral band	Band 3	03-10	q10
S2, spectral band	Band 5	03-10	q10
S2, spectral band	Band 4	07-08	q10
S2, spectral band	Band 11	09-10	q10
S2, spectral band	Band 11	09-10	q50
S2, spectral index	CI_red_edge	03-10	cv
S2, spectral index	GNDVI	03-10	cv
S2, spectral index	MSAVI2	03-10	q10
S2, spectral index	MSWIR	09-10	q50
S2, spectral index	MSWIR	09-11	q50
S2, spectral index	NDRE	03-10	cv
S2, spectral index	NDWI	03-10	cv
S2, spectral index	NDWI	05-06	cv
S2, spectral index	REP	03-10	q90
S2, spectral index	REP	05-06	q90
S2, spectral index	REP	05-06	sd
S2, spectral index	REP	06-08	sd
S2, spectral index	REP	07-08	sd
S2, spectral index	TC green	03-10	q10
S2, spectral index	TC green	05-06	cv
S2, spectral index	TC wet	09-10	q50
S2, spectral index	TC wet	09-11	q50
S2, spectral index	TC wet	09-11	mean
S1, gamma naught	VH	03-04	q90
S1, gamma naught	VH	03-10	q90
S1, gamma naught	VH	06-08	q90
S1, gamma naught	VV	03-05	q90
S1, gamma naught	VV	03-05	mean
S1, gamma naught	VV	03-10	q50
S1, gamma naught	VV	03-10	q90
S1, gamma naught	VV	03-10	mean
S1, gamma naught	VV	06-08	q50
S1, gamma naught	VV	06-08	q90
S1, gamma naught	VV	07-08	q50
S1, gamma naught	VV	07-08	mean

5.1.5.1.3 Classification, Prediction and Post-processing

After the feature selection was complete, all available training samples were used to build the final random forest model on the selected predictor sets, which was then used to spatially predict a) the probability that a pixel belongs to the grassland class, b) a binary “grassland”—“non-grassland” layer, where the probability exceeded 50% and c) a quality layer indicating possible uncertainty due to

temporally interpolated missing values. Subsequently the binary classification was subjected to filtering of a MMU of 5 pixels in four by four neighbourhood relationships, i.e. an MMU of 0.05ha.

5.1.5.1.4 Change Detection

After producing maps for both 2018 and 2017 a change map was calculated, for the classes a) no-change, grassland, b) no-change, not-grassland, c) grassland gain and d) grassland loss. No-change areas, i.e., both a) and b), were buffered with a dilatation filter of pixel size 1, a followed by a MMU filter set to 0.5ha.

5.1.5.1.5 Grassland Intensity

Grassland mowing intensity was calculated from NDVI time-series change events by counting signal drops larger than 0.2 NDVI from a pre-cutting NDVI greater than 0.5 as described in AD06. While this method is rapid in production, it requires careful tuning of the two thresholds. Given that no ancillary information was available for the study site South-East, the prototype is to be considered merely a proof-of-concept. For actual production it would require local expert knowledge or in-situ training data, like it is available with InVeKoS data in Austria for a meaningful rule-based calibration of the algorithm.

5.2 Prototype Validation Results

In this chapter, the validation results for all three demonstration site areas are presented. The validation procedure applied are described in detail in Chapter 4.4.3.

5.2.1.1 Demonstration Site West

With the availability of dense optical and SAR time series from Sentinel-1 and Sentinel-2, grassland mapping can profit from the increased information content provided by temporal measurements of the reflectance of grassland areas over the year. Based on selected LUCAS samples a supervised classification approach using the Random Forest classifier has been successfully applied. An example of the status layers 2017 is shown in Figure 5-18 compared to the LGP polygons from 2016 (see Figure 5-19).

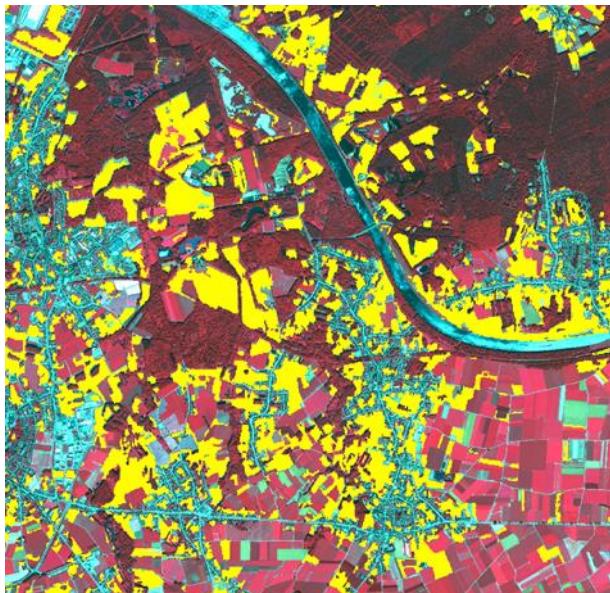


Figure 5-18: SAR + OPT grassland classification with random forest and selected features for 2017 ($p>50\%$) (grassland in yellow).

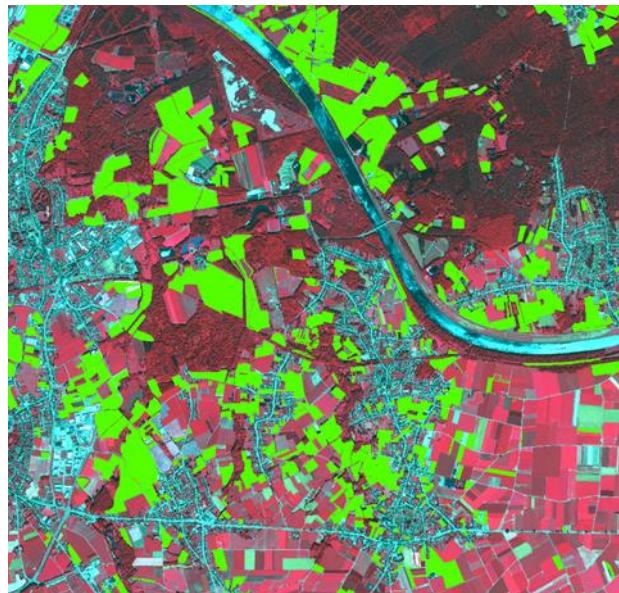


Figure 5-19: LGP grassland areas 2016 in green mapped on the World View 1 image from the 27. 06. 2018.

The S1 and S2 combined approach shows more homogenous patches than using SAR data only and diminishing confusion of grassland with other classes such like tree cover and plantations, that otherwise cannot be excluded when using only S2.

The final validation has been carried out by the consortium partner SIRS for the most promising result benchmarked in WP33 [AD07]. The result of the internal validation for the final West demonstration site (Status layer 2017: “GRA_2017_010m_WE_03035_prototype_v01”) is presented in the below confusion matrix (OA: 97.91) and in the following matrix from the blind approach (OA: 97.39).

Table 5-4: Internal validation results for the GRA_West_2017 product (area-weighted plausibility approach).

GRA_2017_010m_WE_03035_prototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval
		Non-Grassland	Grassland	Total		
PRODUCT	Non-Grassland	386.00	4.11	390.11	98.95%	0.83%
	Grassland	6.34	103.70	110.04	94.24%	0.81%
	Total	392.34	107.81	500		
	Producer Accuracy	98.38%	96.19%		97.91%	Overall Accuracy
		2.65%	1.56%		0.83%	Confidence Interval
					0.94	Kappa
					0.95	F1-score
					0.99	F0-score

Table 5-5: Internal validation results for the GRA_West_2017 product (area-weighted blind approach).

GRA_2017_010m_WE_03035_prototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval
		Non-Grassland	Grassland	Total		
PRODUCT	Non-Grassland	383.76	6.35	390.11	98.37%	1.03%
	Grassland	6.71	103.34	110.04	93.91%	0.84%
	Total	390.47	109.68	500		
	Producer Accuracy	98.28%	94.21%		97.39%	Overall Accuracy
		2.51%	1.91%		1.00%	Confidence Interval
					0.92	Kappa
					0.94	F1-score
					0.98	F0-score

The result of the internal validation for the final West demonstration site (Status layer 2018: “GRA_2018_010m_WE_03035_prototype_v01”) is presented in the below confusion matrix (OA: 98.93) and in the following matrix from the blind approach (OA: 98.01).

Table 5-6: Internal validation results for the GRA_West_2018 product (area-weighted plausibility approach).

GRA_2018_010m_WE_03035_prototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval
		Non-Grassland	Grassland	Total		
PRODUCT	Non-Grassland	386.69	4.11	390.80	98.95%	0.83%
	Grassland	3.96	105.38	109.35	96.38%	0.66%
	Total	390.66	109.49	500		
	Producer Accuracy	98.99%	96.25%		98.39%	Overall Accuracy
		2.74%	1.55%		0.80%	Confidence Interval
					0.95	Kappa
					0.96	F1-score
					0.99	F0-score

Table 5-7: Internal validation results for the GRA_West_2018 product (area-weighted blind approach).

GRA_2018_010m_WE_03035_prototype_v01 Blind Approach		REFERENCE			User Accuracy	Confidence Interval
		Non-Grassland	Grassland	Total		
PRODUCT	Non-Grassland	385.18	5.62	390.80	98.56%	0.97%
	Grassland	4.32	105.02	109.35	96.05%	0.69%
	Total	389.50	110.65	500		
	Producer Accuracy	98.89%	94.92%		98.01%	Overall Accuracy
		2.47%	1.79%		0.92%	Confidence Interval
					0.94	Kappa
					0.95	F1-score
					0.99	F0-score

Table 5-4 and Table 5-6 show the plausibility results for the grassland status mapping 2017 and 2018 respectively. Regarding the blind and plausibility approach, an overall accuracy over 97% could be achieved for both. The differences between the blind and plausibility are not significant. The user and producer accuracy is particularly high for both the grassland and the non-grassland class in the blind and plausibility approach.

Misclassifications occur between crop areas and grass areas. Cropland can be excluded from the layer with a higher reliability if the ploughing event is captured in the time series. If that is not the case, the crop areas share the similar spectral signatures with grass areas and cannot be excluded. The ploughing event is the main character to discriminate the grassland areas from crop areas. The grassland over detection due to missing ploughing events in the time series is the primary issue regarding the classification for 2017, where the time series is a lot sparser. Changing or abnormal environmental conditions between seasons also have an effect on the classification result. The drought period in summer 2018 resulted in a under detection of grassland areas as their spectral signature has been similar to some crop areas. Further, some misclassifications with fruit orchards remain as the grass cover between the orchards influences the spectral signature and the algorithm cannot separate them. Nevertheless, most of the orchards can be excluded using S1 and S2 depending on the tree size and management practices (see Figure 5-20 and Figure 5-21).



Figure 5-20: World View 1 image from the 14.03.2017.



Figure 5-21: GRA_West_2017 (grassland in green).

Additional to the grassland status layers 2017 and 2018 the grassland probability layers are generated as pixel-based quality measure shown in Figure 5-22 and Figure 5-23 for 2017 and 2018, respectively.

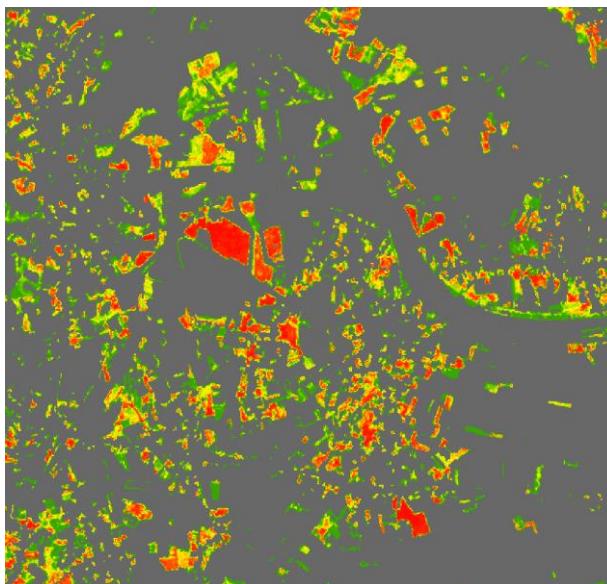


Figure 5-22: Grassland Probability Layer 2017.

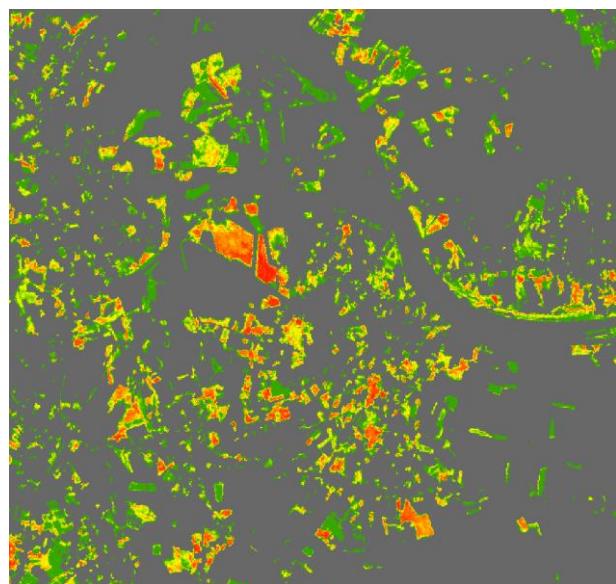


Figure 5-23: Grassland Probability Layer 2018.

In addition to the status layers, a grassland intensity layer is produced in 2018. Due to missing ground reference data a quantitative validation of the results could not be performed for the WEST demonstration site. A qualitative assessment with respect to the plausibility has been conducted by a remote sensing expert.

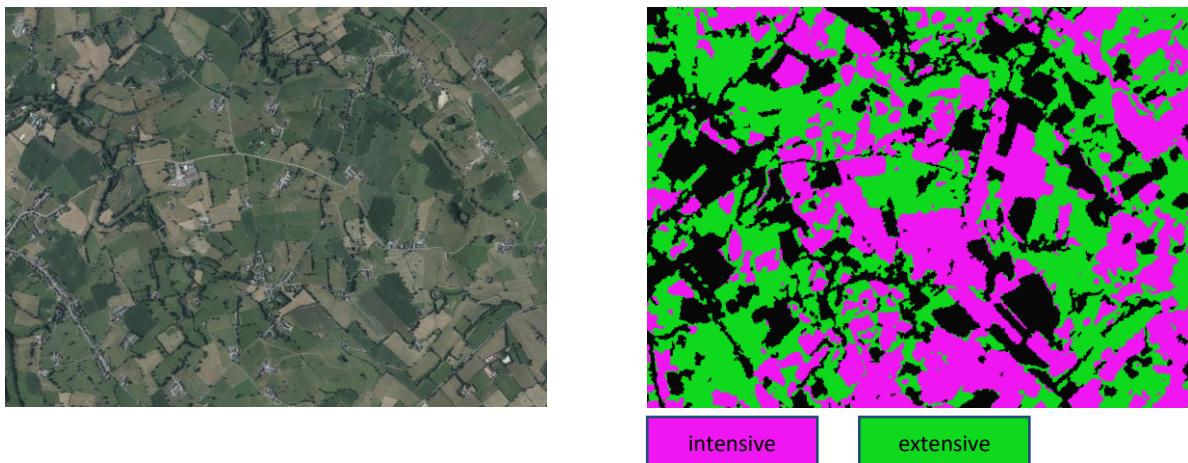


Figure 5-24: Sample of the grassland mowing intensity map.

Thereby, a systematic influence caused by the locally varying number of available observations could be identified. Figure 5-25 illustrates this issue. If data from overlapping orbits is not available, the number of clear-sky observations can become critically low and a reliable detection of mowing events is unrealistic. As a consequence, the number of patches labelled as “intensive” is lower in areas with less than 20 available observations.

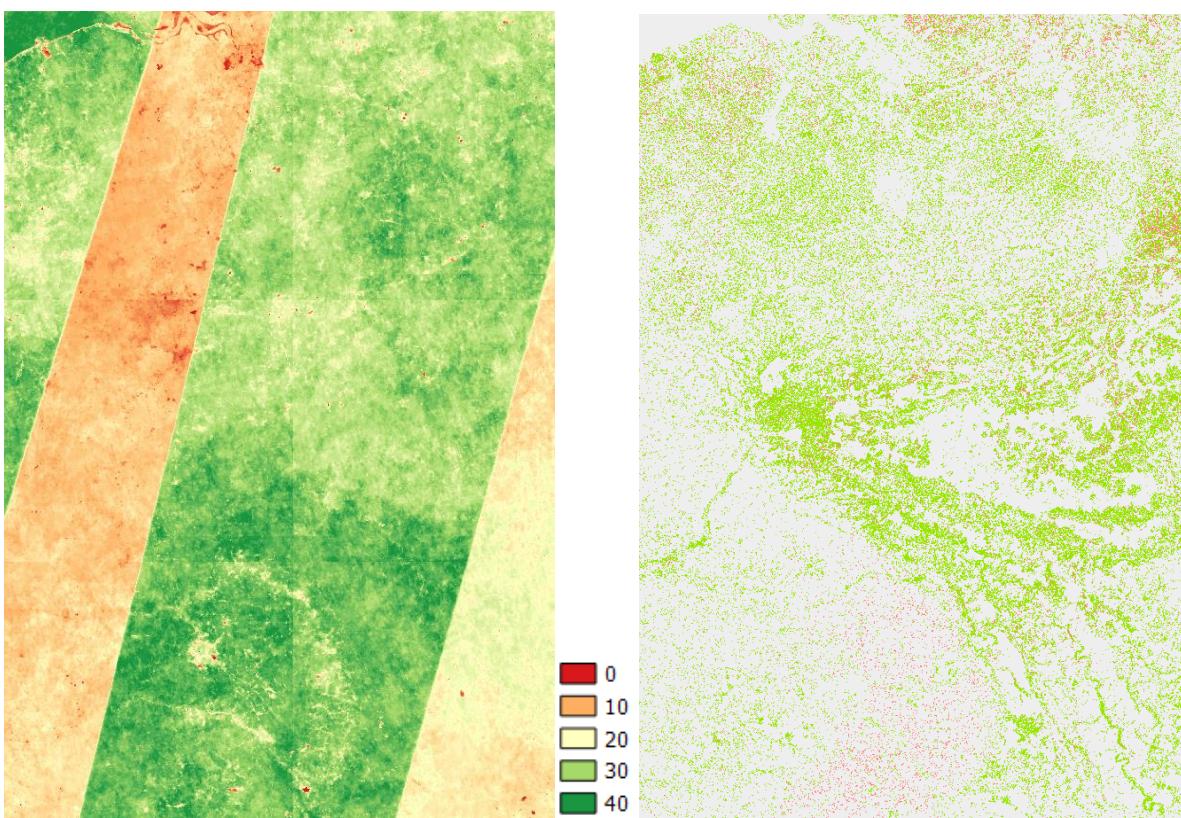


Figure 5-25: Number of valid observations per pixel for the S2 image stack used as input to the mowing event detection algorithm (beginning of week 10/2018 to end of week 44/2018) on the left. The right image shows the grassland use intensity layer 2018 for the demonstration site West.

Table 5-8: Internal validation results for the GRA_West_Change_2017_18 product (area-weighted plausibility approach).

GRC_1718_010m_WE_03035_prototype_v01 Plausibility Approach		REFERENCE					User Accuracy
		No change without grass cover	New grass cover	Loss of grass cover	No change with grass cover	Total	
PRODUCT	No change without grass cover	378.42	1.08	0.35	2.66	382.53	98.93%
	New grass cover	1.80	4.68		1.08	7.57	61.90%
	Loss of grass cover	2.87		5.03	0.35	8.26	60.87%
	No change with grass cover	1.07		0.35	100.33	101.77	98.59%
	Total	384.18	5.77	5.75	104.44	500	
	Producer Accuracy	98.50%	81.25%	87.50%	96.06%		97.67% Overall Accuracy
						0.94 Kappa	
						0.99 F-score No change without grass cover	
						0.70 F-score New grass cover	
						0.72 F-score Loss of grass cover	
						0.97 F-score No change with grass cover	

The Change layer 2017-2018 has also been validated. Due to the complexity of interpretation of changes between 2 years the analysis was directly done as a plausibility analysis. The error matrix for the change product is shown Table 5-8. The change layer from the 2017 to 2018 status layers is displayed in Figure 5-28. Despite filtering for an MMU of 0.5ha a lot of change has been detected. Regarding the plausibility approach, an overall accuracy over 97% could be achieved. The user accuracy is particularly low for both the new grass cover and the loss of grass cover class.



Figure 5-26: World View 1 image from the 2017.05.09.



Figure 5-27: World View 1 image from the 2018.10.05.



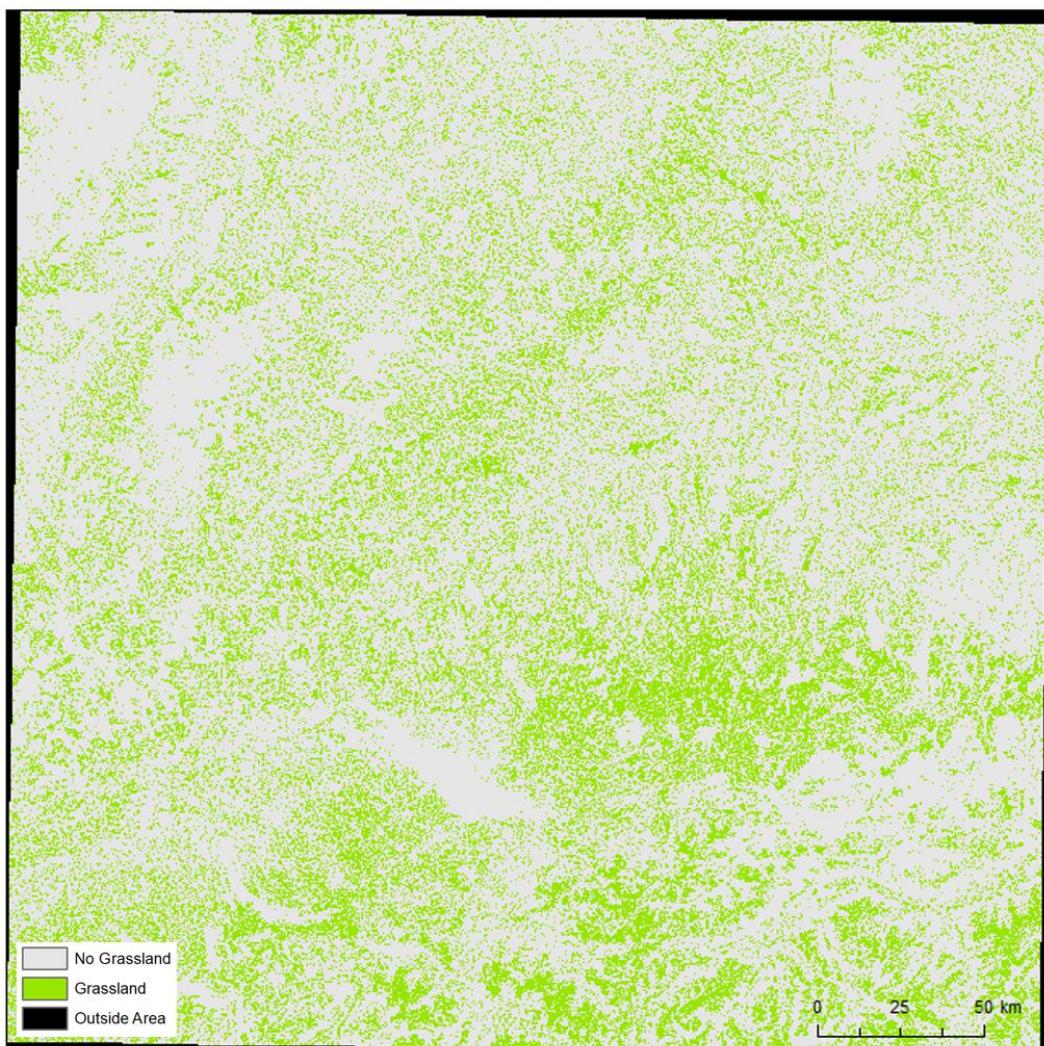
Figure 5-28: CVA approach. Loss (yellow), Gain (Green).

In general, too many changes appear, due to the misclassifications in the status maps/probabilities, which occur especially in areas where the number of observations is too low. Specifically, agricultural fields are misclassified as grassland if the ploughing event is not captured in the time series.

It must be taken into account that no historical time series were included in the grassland classification process to detect ploughing events and exclude grassland younger than 5 years, therefore agricultural grassland is included in the status layer. However, many changes actually result from such grassland, which may be present in one year while tilled in another year. Higher details in subsequent status products could be applied to improve the change detection in this regard. To reduce misclassified change caused by agricultural grasslands, which is present in one year, but not in the other, the usage of a historical time series is required to detect ploughing events. In this manner, areas that were ploughed e.g. within the last 5 years can be excluded.

5.2.1.2 Demonstration Site Central

With the availability of dense optical and SAR time series from Sentinel-1 and Sentinel-2, grassland mapping can profit from the increased information content provided by temporal measurements of the reflectance of grassland areas over the year. Based on re-interpreted LUCAS samples a supervised classification approach using the Random Forest classifier has been successfully applied. The status layers 2017 and 2018 are shown in Figure 5-29.



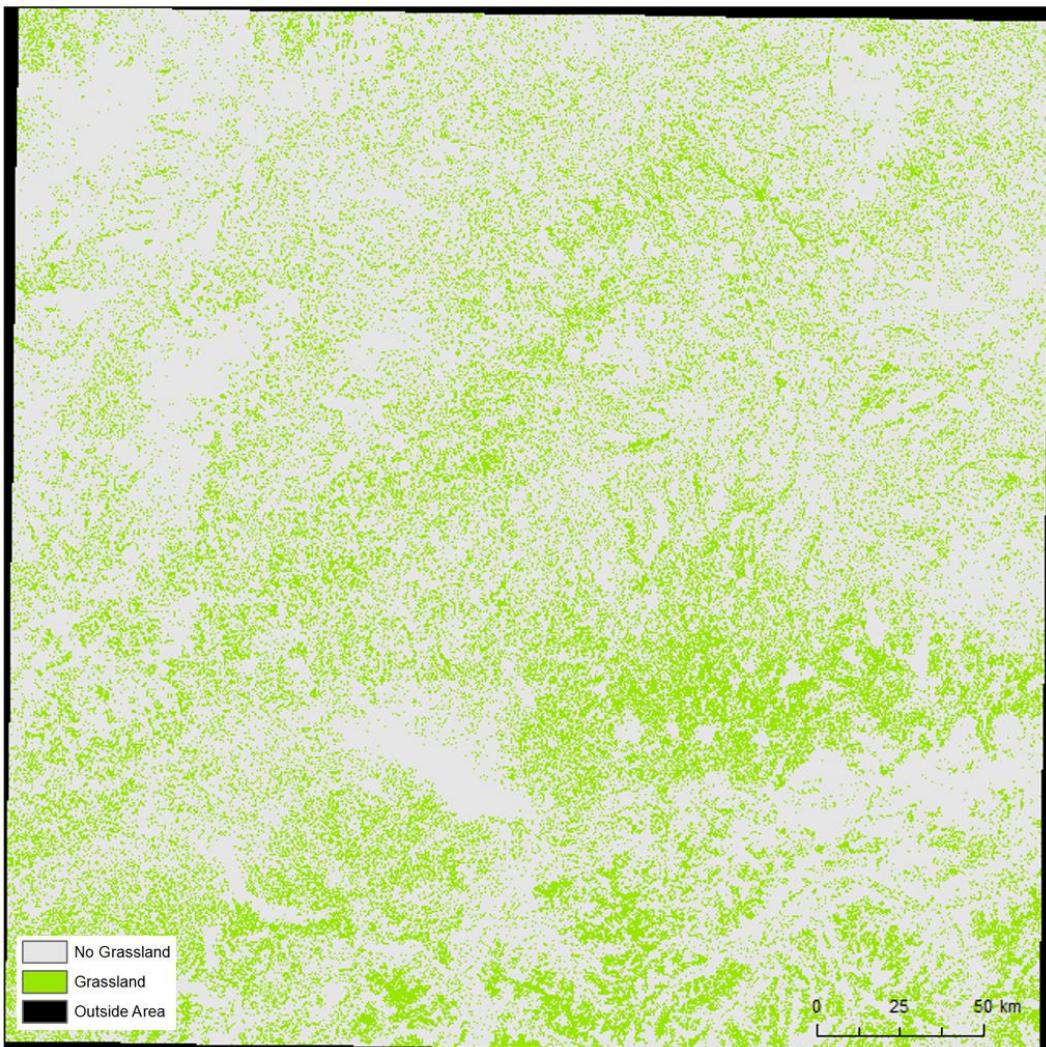


Figure 5-29: Grassland masks Central 2017 (up) and 2018 (down).

The feature selection tests automatically executed by the Random Forest Classifier have shown that, in the Central demonstration site, best performing input indices are NDVI, NDWI, IRECI and Brightness. Regarding SAR data generally the following annual coefficients have shown the best performance: VH polarisation, the VV polarisation the NDVVVH and the RATIOVVH. The later winter to spring (January to June) time window in the end did not bring an added value for classification. In different biogeographical regions the short time window can differ (which is also the case in the forest and agriculture products). To better match the local conditions, it is important to define a time window where grassland and cropland (class generating most misclassifications) are best separable (e.g., define the period when grassland is already greening whilst cropland is not). In Mediterranean regions, for instance, this window may be shifted more towards winter (e.g. Dec-Mar). Considering data availability and the not homogeneous data situation across Europe, another consideration towards larger area and constant production, is that the time period should not be too short, as otherwise results might not be meaningful due to the likeliness of limited number of scenes in some regions. Changing or abnormal environmental conditions between seasons also have an effect on the classification result. The Figure below shows the influence of the 2018 drought in the grassland product in Central. The change product as explained in D34 [AD08] is particularly affected by such circumstances.

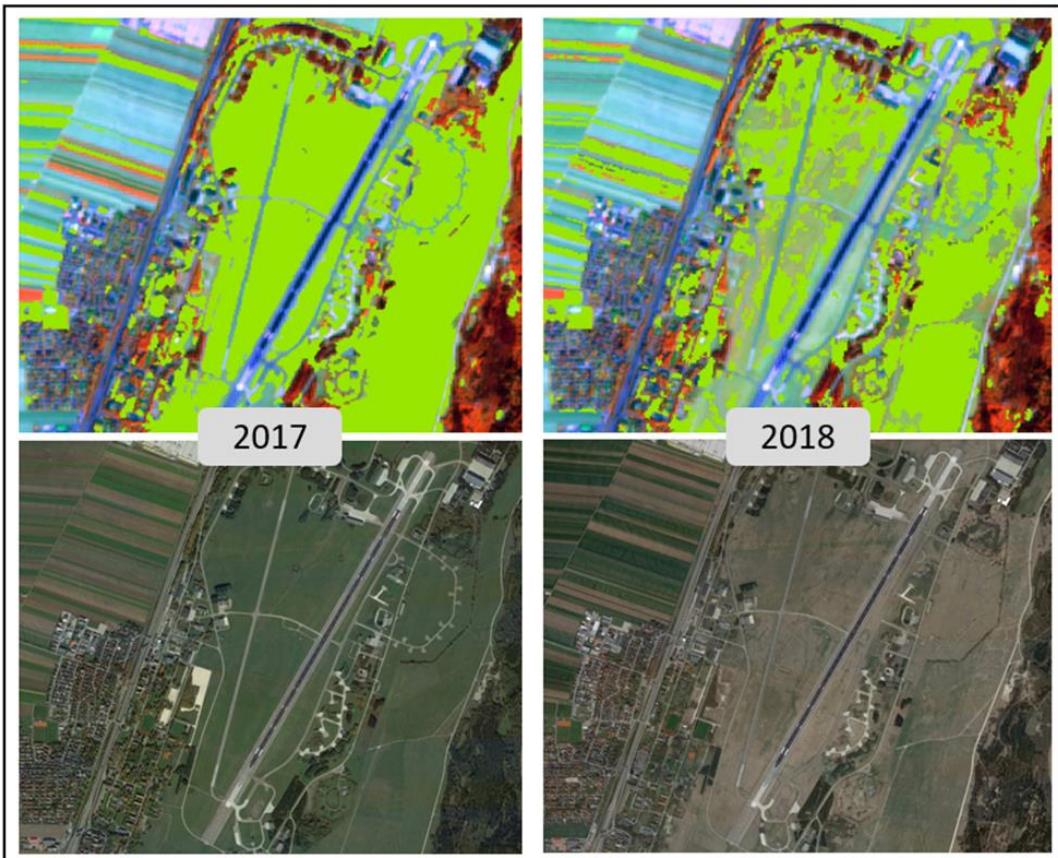


Figure 5-30: Detailed view from the Grassland identification 2018 and 2017 showing the effects of 2018 drought.

With this in view, the visual inspection as part of the validation procedure, as depicted in the accuracy guidelines in D33 [AD07] is required to guarantee reliability of the status layers and facilitate subsequent correct change detection (WP34 [AD08]) and incremental updates (WP35 [AD12]).

After the visual interpretation of all classifications, it can be observed that using optical data only, more confusion between grasslands and cropland are present, whereas using SAR data only, more misclassifications between grassland and roads are present. The combined approach shows more homogenous patches than using SAR data only and diminishing confusion of grassland with other classes suchlike tree cover and plantations, that otherwise cannot be excluded when using only S2. This is relevant as in fact the overall accuracy values obtained are misleading while the visual inspection demonstrates the combined approach performs best. A main requirement however is the precise pre-processing of the dense time series including a topographic normalisation for hilly to mountainous terrain. For SAR time-series the application of multi-temporal filtering on gamma naught corrected imagery is recommended.

For the Central site Grassland mask 2018, the overall accuracy is 96.6% (0.93% Confidence Interval). Analogously, for 2017 the overall accuracy is 96.4% (1.1% Confidence Interval). Producer and user accuracies are shown in the corresponding confusion matrices below.

Table 5-9: Internal validation results for the GRA_Central_2017 product (area-weighted plausibility approach).

GRA_2017_010m_CE_03035_p prototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval	
		Non- Grassland	Grassland	Total			
PRODUCT	Non-Grassland	590.01	16.54	606.55	97.27%	1.10%	
	Grassland	10.76	132.69	143.45	92.50%	0.70%	
		600.77	149.23	750			
		98.21%	88.92%		96.36%	Overall Accuracy	
		4.16%	1.33%		1.04%	Confidence Interval	
					0.88	Kappa	
					0.91	F1-score	
					0.98	F0-score	

The Figure below shows a detailed view of the 32UNU tile compared to LPIS grasslands polygons outlined in yellow. The base image is a S2 scene from 4th of October 2018 with the colour combination of R: NIR. G: SWIR and B: Red.

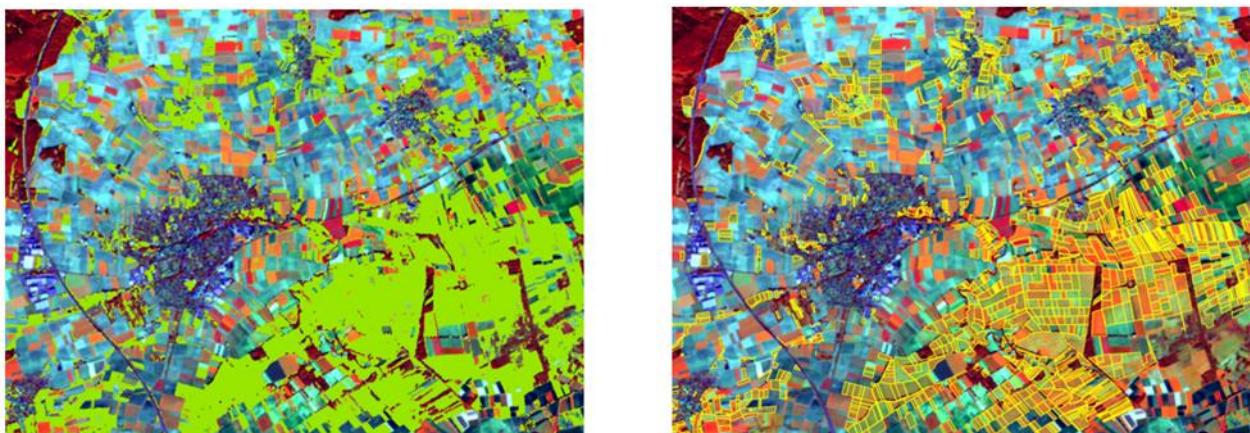


Figure 5-31: Classification result detail grassland (green) mask 2018 in Central demonstration site.

The corresponding probability layer depicts the high degree of confidence achieved in the classification result in a particular complex area where the distinction between grasslands and croplands is challenging.

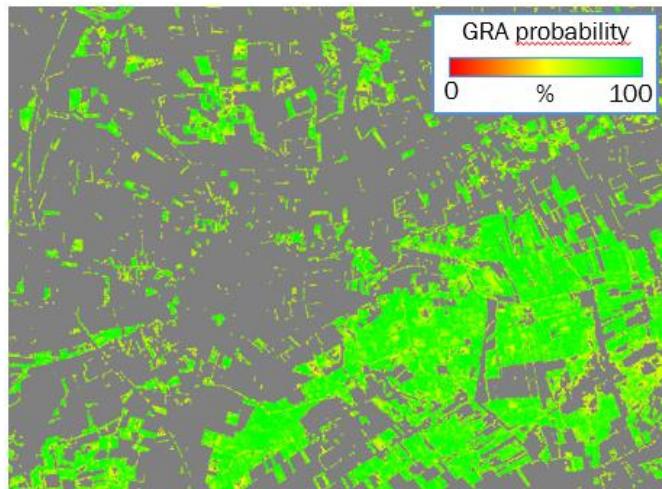


Figure 5-32: Grassland mask 2018 Probability layer.

For the 2017 status layer, the User Accuracy is particularly high, and the results show very low level of omission errors with less than 1.8%. Commissions errors are more numerous (2.73%) but remain in line with the product specifications.

Table 5-10: Internal validation results for the GRA_Central_2018 product (area-weighted plausibility approach).

GRA_2018_010m_CE_03035_prototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval
		Non-Grassland	Grassland	Total		
PRODUCT	Non-Grassland	600.29	12.27	612.56	98.00%	0.96%
	Grassland	12.99	124.44	137.43	90.55%	0.74%
	Total	613.28	136.71	750		
	Producer Accuracy	97.88%	91.02%		96.63%	Overall Accuracy
	Confidence Interval	3.40%	1.40%		0.93%	Confidence Interval
					0.89	Kappa
					0.91	F1-score
					0.98	F0-score

For the 2018 product, the User Accuracy is particularly high, and the results show very low level of commission errors with less than 2% and a confidence interval at 0.96%. Omissions errors are also few (2.12%) so the overall result is in line with high quality product specifications approaching 97% accuracy. The reference data for the validation was independent points interpreted by CAPI (750 points for the whole demonstration site). The results of the internal validation for the Status Layer 2018 are quite similar to the results of the Status layer 2017 even if omission errors are a bit higher but still in line with product specifications.

The consistency between the status layers in Central developed in ECoLaSS for Agriculture, Grasslands and Forest can be seen in the Figure below, showing the crop mask 2018 [AD07, AD14].

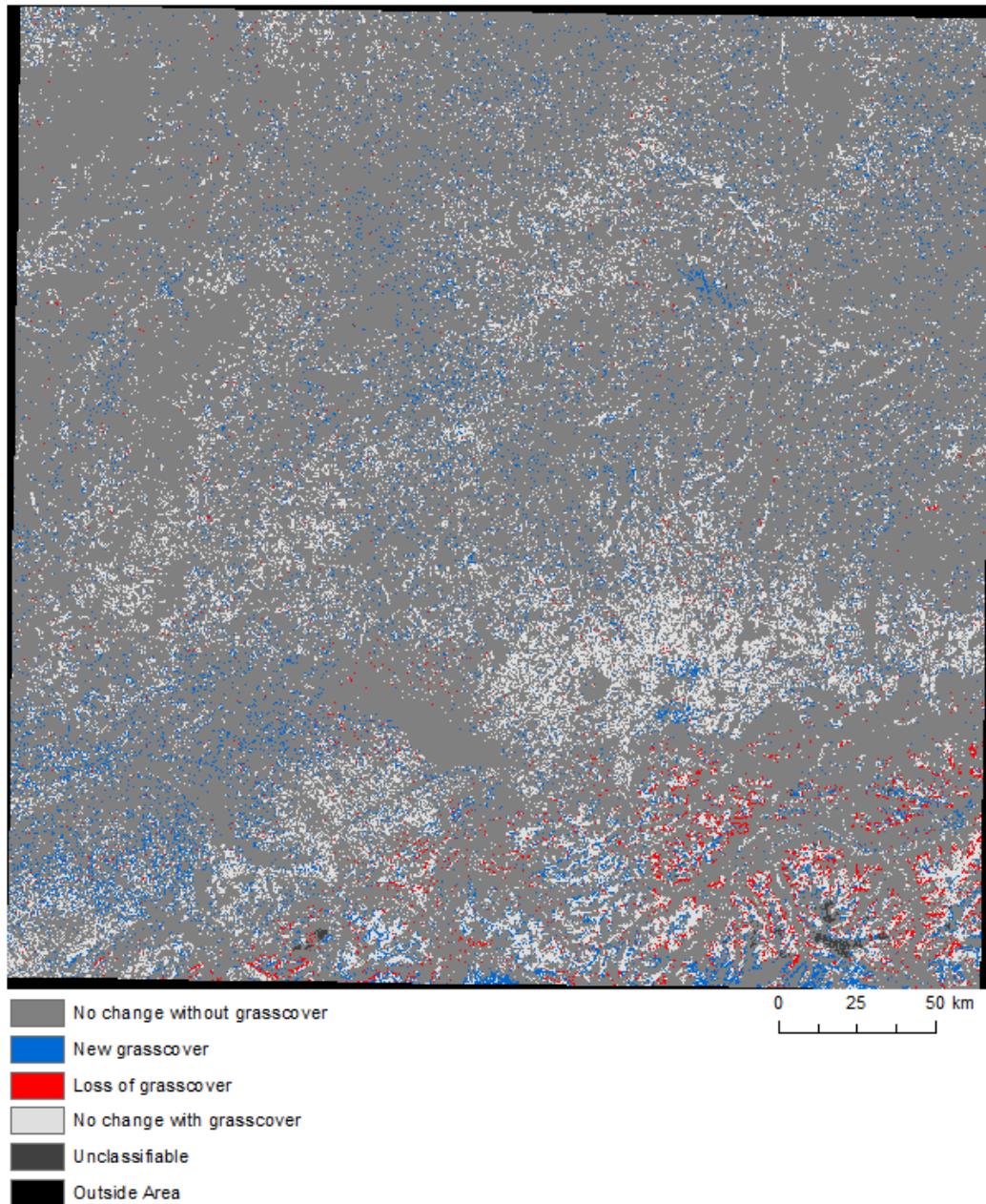


Figure 5-33: Consistency between the Grassland Status Layer 2018 and the Crop Mask 2018.

Regarding the change layers, as only the geometrical change areas are shown and no further exclusion of non-change by the usage of e.g. NDVI minimum or other change indicators is applied to the 2018-2015 change layer, there are many more changes visible compared to the 2017/18 change layer. Most of them in truth are no real change, instead resulting from weaknesses in the 2015 layer.

In areas of higher elevation and where snow cover is found for long periods of the year in the South of the demonstration site, the classification is not that accurate for both years (2015 & 2018) which in turn leads to greater differences and, therefore, change detection which is basically no change. An elevation threshold has been applied to reduce such errors, which significantly improved the respective grassland masks in the most problematic regions. Filtering significantly contributes to keep meaningful changes while removing small areas, likely classification errors. All areas below 1ha (25 pixels, 20m resolution) were filtered.

The screenshots below show respectively the change layers 2015-2018 and 2017-2018 for the Central demonstration site. Loss of grassland cover is depicted in red, gain of grassland is depicted in blue whereas no change with grasslands is shown in light grey and no change detected in non-grasslands areas is shown in dark grey.



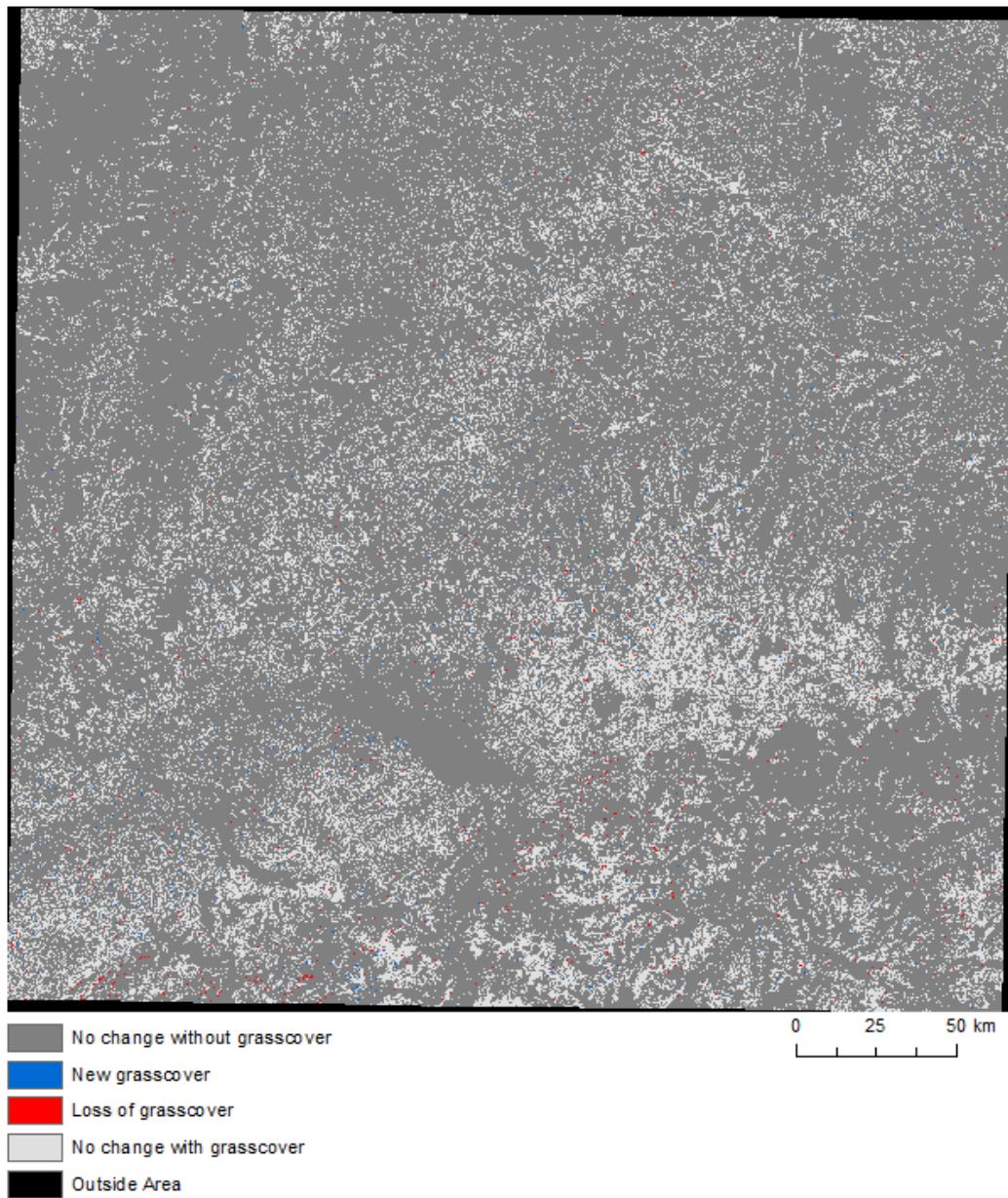


Figure 5-34: Grassland change layers 2015-2018 (up) and 2017-2018 (down).

Due to the difference in the methodological approaches between these two layers as explained in the methods description in Task 3 and Task 4 reports, the change detection for the 2015-2018 cannot be validated in a comparable way. However, the approach is still interesting because it allows for change detection between pre-existing maps and new classifications. The result of these two change detection approaches can be compared in Figure 5-35.

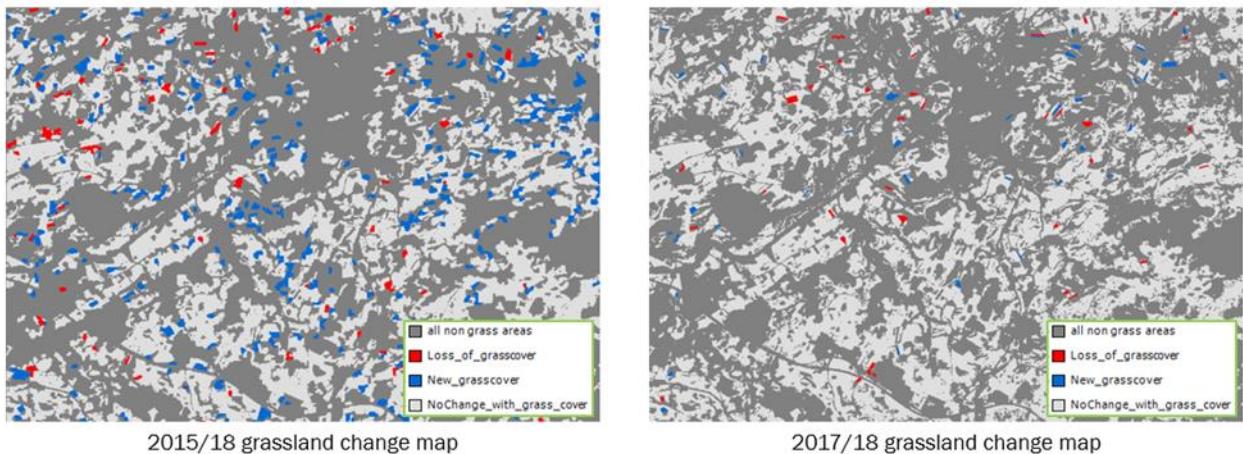


Figure 5-35: Detailed view of the map-to-map change detection and the NDVI time series thresholding approach applied in the Central demonstration site (reference years 2015-2018 and 2017-2018 respectively).

In the map-to-map comparison it has been verified that many changes are a result of the different resolutions of the respective status products (i.e., HRL2015 20m versus 2018 prototype 10m), even having resampled the 2018 product to the match the 2015 product. Additionally, erroneous changes are derived from the methodological limitations, as the application of this simple approach does not allow for the elimination of technical changes. As a consequence, in this product, the weaknesses of the HRL2015 appear as grasslands loss/gain. In contrast, the 2017-2018 change product provides a more consistent change detection. Less changes are identified as a result of the more homogeneous quality of both status layers (e.g., both reference years have the same native resolution: 10m), and the more sophisticated approach implemented that leads to avoiding the inclusion of misclassified technical changes.

The Change layer 2017-2018 has also been validated. Due to the complexity of interpretation of changes between 2 years, the analysis was directly done as a plausibility analysis. The error matrix for the change product is shown below.

Table 5-11: Internal validation results for the GRA_Central_Change_2017_18 product (area-weighted plausibility approach).

PRODUCT		REFERENCE					User Accuracy
		No change without grass cover	New grass cover	Loss of grass cover	No change with grass cover	Total	
GRC_1718_010m_CE_03035_prototype_v01 Plausibility Approach	No change without grass cover	639.56			14.65	654.21	97.76%
	New grass cover	0.09	0.41		0.07	0.57	72.00%
	Loss of grass cover	0.08		1.43	0.1	1.61	88.62%
	No change with grass cover	6.36			87.25	93.61	93.20%
	Total	646.09	0.41	1.43	102.07	750	
	Producer Accuracy	98.99%	100.00%	100.00%	85.48%		97.15% Overall Accuracy
						0.88	Kappa
						0.98	F-score No change without grass cover
						0.84	F-score New grass cover
						0.94	F-score Loss of grass cover

As reported in WP33 [AD07] the EO data availability can clearly impose constraints in this mowing intensity detection methodology. The current temporal density of optical time-series from Sentinel-2 for the years 2017 and 2016 restricted the applicability of time series methods, as reflectance trajectories depend on the grassland dynamics over the vegetation period such as e.g. mowing events. It is clear that a high cloud coverage can limit the method of comparing NDVIs of consecutive acquisition dates (as was proved in the test sites in 2018). This implies that mowing events may not be detected in areas covered by clouds for a long period of time.

Regarding the reference data employed for the validation of the Grassland mowing intensity layer, the new grassland attributes available for some points of the LUCAS 2018 data in the demonstration site, could be very helpful if the points density was sufficient in the target area. Alternatively, the validation samples have been extracted from IACS data, containing information about the mowing frequencies. The datasets used for validation for the mowing intensity grassland product are the IACS, generating an overall accuracy of 81.5%. Unfortunately, the information on number of mowing events is only included in the dataset for Austria and therefore only parts of the demonstration site are covered: two out of nine tiles in the case of Central. Validation in this case is computed by cross-checking the number of mowing events in the IACS samples (SNAR-BEZEI attribute) with the number of mowing events detected in the intermediate product of the mowing intensity layer (i.e., layer with the number of mowing events). Consequently, the mowing intensity layer cannot be fully validated, nor at the demo scale nor for a potential Pan-European/global roll-out. A qualitative inspection was also implemented.

All in all, the grassland intensity 2018 layer is consistent as derived from the areas defined as grassland within the grassland status layer, and based on a simple although effective NDVI time series analysis, with a good trade-off between computer resources and pre-processing timing (i.e., impact in updates production timeliness) and achieved thematic accuracy.

5.2.1.3 Demonstration Site South-East

The production of the grassland status layer for the demonstration site South-East underlines the complementarity of SAR and optical data for grassland mapping. While mapping grasslands was possible with either one, optimal mapping performance could only be achieved when combining both Sentinel-1 and Sentinel-2 data.

The final grassland status layers of the years 2017 and 2018 are shown in Figure 5-36. In 2017 a total 11.43% of the area was detected as grassland. In 2018 detected grasslands covered 11.06% of the area. For most areas a successful production with the selected features was possible, due to a sufficient number of cloud-free observations. Some areas remained, where the optical data were incomplete, which was more of a problem in 2018 than in 2017. Regarding model uncertainty the 2018 model distinguished more clearly between grassland and non-grassland (Figure 5-37). This was expected, given that the LUCAS training data were from the year 2018.

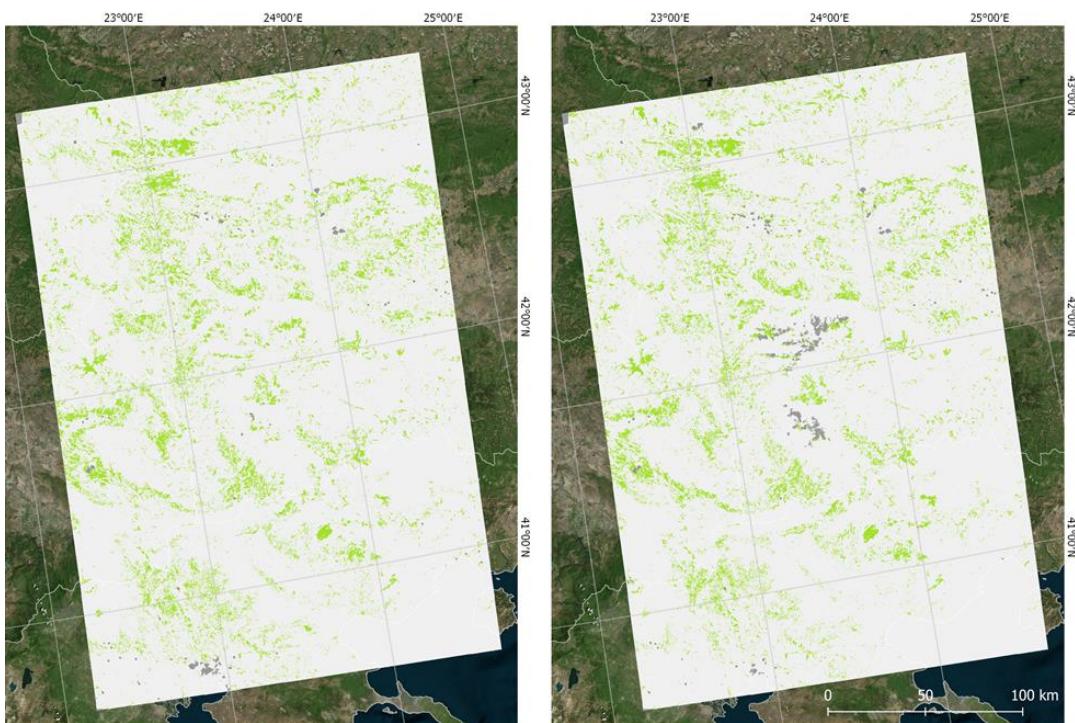


Figure 5-36: Grassland status layers of 2017 (left) and 2018 (right) for the demonstration site South-East. Dark grey areas denote incomplete data due to actual clouds and misclassified clouds due to very bright targets data.

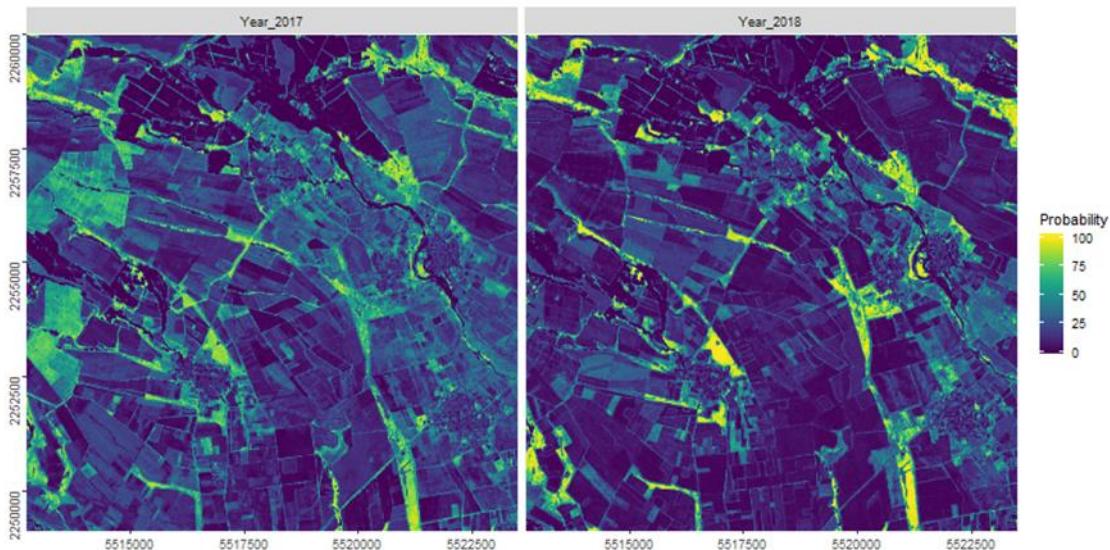


Figure 5-37: Random forest grassland probability 2017 (left) and 2018 (right).

The final internal validation has been carried out by the consortium partner SIRS for the most promising result of the feature selection. The result of the internal validation for the final South-East Demonstration site (Status layer 2018: “GRA_2018_010m_SE_03035_prototype_v01”) is presented in the below confusion matrix (Table 5-12) and in the following matrix from the blind approach (Table 5-13).

Table 5-12: Internal validation results for the GRA_South_East_2018 product (area-weighted plausibility approach).

GRA_2018_010m_SE_03035_p rototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval
		Non- Grassland	Grassland	Total		
PRODUCT	Non-Grassland	432.52	10.90	443.43	97.54%	1.41%
	Grassland	3.37	53.04	56.41	94.03%	0.45%
	Total	435.89	63.94	499.84		
	Producer Accuracy	99.23%	82.95%		97.14%	Overall Accuracy
	Confidence Interval	2.73%	3.42%		1.34%	Confidence Interval
					0.87	Kappa
					0.88	F1-score
					0.98	F0-score

Table 5-13: Internal validation results for the GRA_South_East_2018 product (area-weighted blind approach).

GRA_2018_010m_SE_03035_p prototype_v01 Blind Approach		REFERENCE			User Accuracy	Confidence Interval	
		Non- Grassland	Grassland	Total			
PRODUCT	Non-Grassland	411.93	31.50	443.43	92.90%	2.34%	
	Grassland	8.00	48.41	56.41	85.82%	0.67%	
		Total	419.93	79.91	499.84		
		Producer Accuracy	98.10%	60.58%	92.10%	Overall Accuracy	
		Confidence Interval	1.66%	4.45%	2.23%	Confidence Interval	
						Kappa	
						0.71	
						F1-score	
						F0-score	

The user accuracy is particularly high for the grassland class. The results show a low level of commission errors with less than 5.97% and are in line with the product specifications. Some commission errors have been detected in orchards and frequently occurring mixtures of grassland and shrubland. The omission errors were higher after the blind analysis (39.42%). The result after the plausibility analysis is better and does reach the product specification with a Producer Accuracy of 82.95% ($\pm 3.42\%$). Omissions were detected on edges of grassland plots.

The result of the internal validation for the final South East demonstration site (Status layer 2017: “GRA_2017_010m_SE_03035_prototype_v01”) is presented in the confusion matrix Table 5-14 and in the following matrix from the blind approach (Table 5-15).

Table 5-14: Internal validation results for the GRA_South_East_2017 product (area-weighted plausibility approach).

GRA_2017_010m_SE_03035_p prototype_v01 Plausibility Approach		REFERENCE			User Accuracy	Confidence Interval	
		Non- Grassland	Grassland	Total			
PRODUCT	Non-Grassland	427.56	17.12	444.68	96.15%	1.71%	
	Grassland	2.47	51.47	53.94	95.41%	0.43%	
		Total	430.04	68.58	498.62 *		
		Producer Accuracy	99.42%	75.04%	96.07%	Overall Accuracy	
		Confidence Interval	1.14%	3.85%	1.62%	Confidence Interval	
						Kappa	
						0.84	
						F1-score	
						F0-score	

* Please note that the total changed between 2017 and 2018 due to a validation point located in a cloudy area (code 254)

Table 5-15: Internal validation results for the GRA_South_East_2017 product (area-weighted blind approach).

GRA_2017_010m_SE_03035_p prototype_v01 Blind Approach		REFERENCE			User Accuracy	Confidence Interval	
		Non- Grassland	Grassland	Total			
PRODUCT	Non-Grassland	410.50	34.18	444.68	92.31%	2.37%	
	Grassland	8.21	45.73	53.94	84.77%	0.74%	
		Total	418.71	79.91	498.62 *		
		Producer Accuracy	98.04%	57.23%	91.50%	Overall Accuracy	
		Confidence Interval	2.08%	4.40%	2.26%	Confidence Interval	
						0.64	
						0.68	
						0.95	
						F0-score	

* Please note that the total changed between 2017 and 2018 due to a validation point located in a cloudy area (code 254)

The accuracy of the grassland product for 2017 are lower than for the grassland product 2018. However, the User Accuracy is particularly high for the grassland class and non-grassland classes and are in line with the product specifications. Some commissions are still detected in orchards.

The Omissions errors were higher after the blind analysis (42.77%). The result after the plausibility analysis is better but not enough to reach the product specification with a Producer Accuracy of 75.04% ($\pm 3.85\%$). Omissions were still detected on the edges of grassland plots but also on areas with small grassland patches. A lack of homogeneity has also been noted: some homogeneous plots of grassland were partially coded in grassland.

Particular difficulty occurring frequently within the South-East demonstration site are shrubland mosaics where grassland is interdispersed with shrubs and trees but also rocks. This leads to difficulties to identify larger continuous patches of grassland (e.g. Figure 5-38). A multi-scale analysis may be necessary to improve further on this.

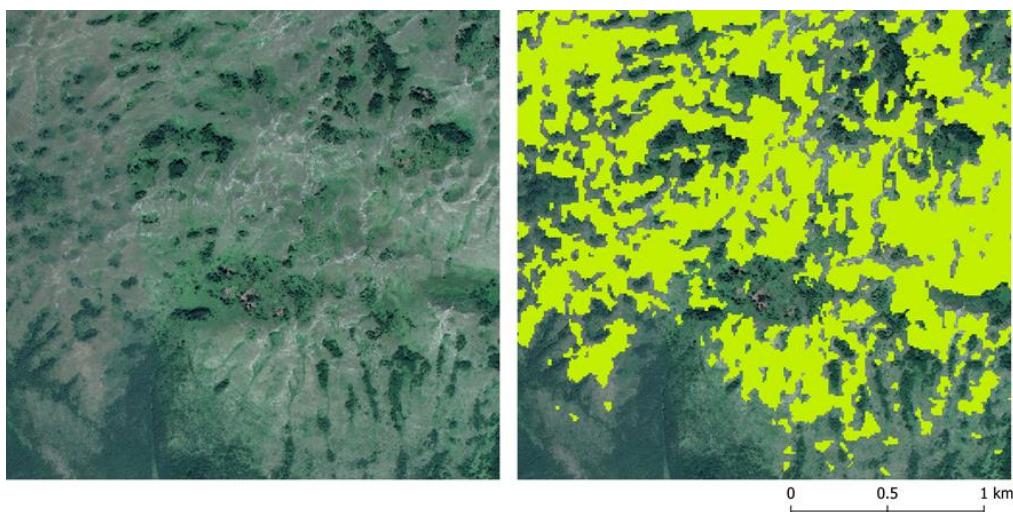


Figure 5-38: Detected non-continuous grassland patches, due to interdispersion with rocks, shrubs and trees.

The change layer from the 2017 to 2018 status layers is displayed in Figure 5-39: Despite filtering for an MMU of 0.5ha a lot of change has been detected. This is more likely attributed to model uncertainty due

to the small sample size and missing 2017 in-situ samples, than to actual changes. This is mirrored in poor User and Producer accuracies following the plausibility validation as displayed in Table 5-16.

Table 5-16: Internal validation results for the GRA_South_East_Change_2017_18 product (area-weighted plausibility approach).

GRC_1718_010m_SE_03035_prototype_v01 Plausibility Approach		REFERENCE					User Accuracy
		No change without grass cover	New grass cover	Loss of grass cover	No change with grass cover	Total	
PRODUCT	No change without grass cover	421.55	1.17	1.17	19.85	443.74	95.00%
	New grass cover	0.65	1.12		0.75	2.52	44.44%
	Loss of grass cover	1.56		1.33	0.11	3.00	44.44%
	No change with grass cover	3.10	0.77		47.23	51.10	92.42%
	Total	426.86	3.06	2.50	67.94	500	
	Producer Accuracy	98.76%	36.60%	53.31%	69.51%		94.18% Overall Accuracy
						0.75	Kappa
						0.97	F-score No change without grass cover
						0.40	F-score New grass cover
						0.48	F-score Loss of grass cover

Both omission and commission errors were numerous on the change strata. During the interpretation of the validation points, many “false” losses of grassland have been detected and mainly due to confusion with cropland plots (based on historical images and based on available LPIS) and ploughed field / bare soils in 2018. Concerning the “false” gain / new grassland in 2018, many of them are, in fact, grasslands for many years and cannot be coded as a gain of grassland. These plots are probably omissions errors of the status layer 2017.

From the presented results, it is clear that while there has been some success in mapping grasslands solely based on LUCAS training data, overall was found to be insufficient for operational purposes in the South East study area. It is likely however, that making use of all available LUCAS data from the wider geographic region surrounding the demonstration site could alleviate this shortcoming.

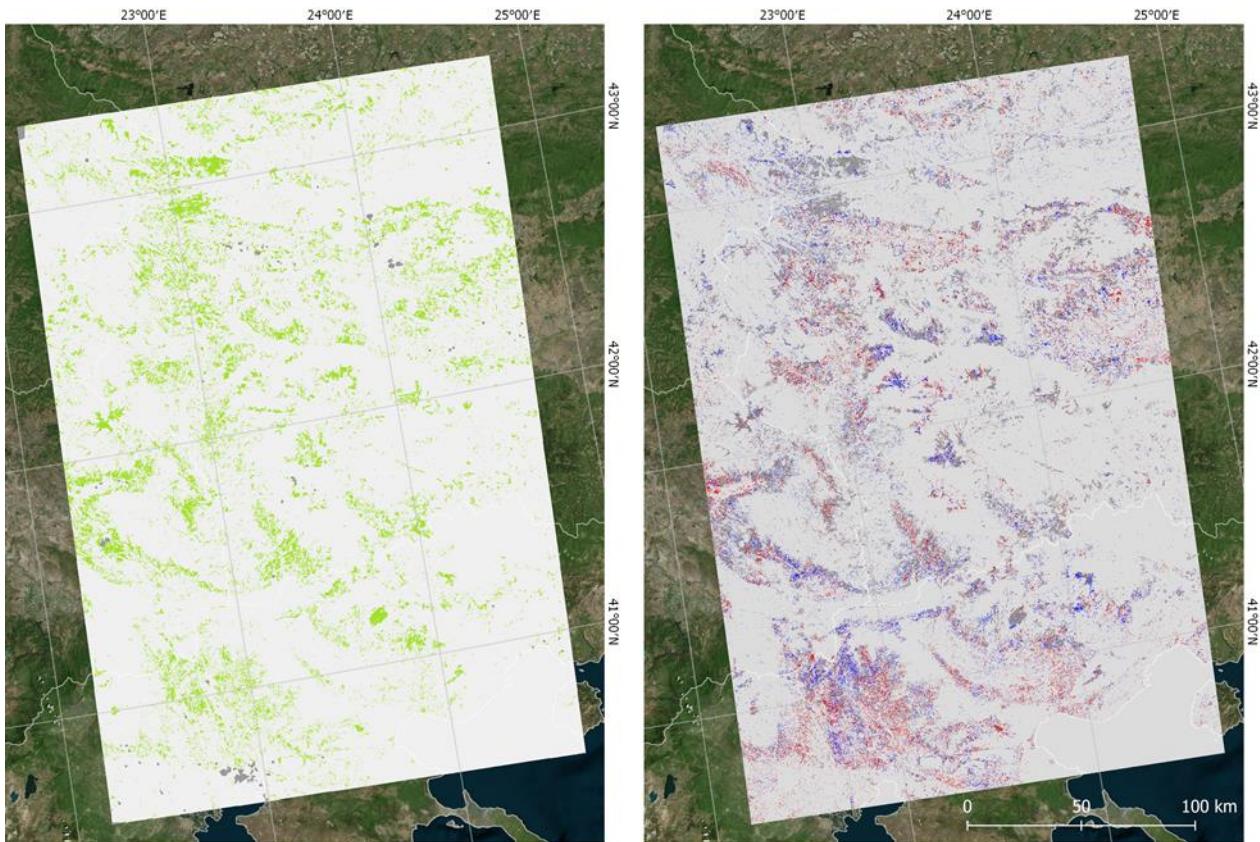


Figure 5-39: Left: 2017 status layer, right: change layer 2017 – 2018. blue = gain, red = loss, dark grey no-change, light gray, no grassland.

5.2.2 HRL2015 Comparison with an Independent Data Source

For a better understanding of the quality of the pan-European HRL Grassland 2015 layer and to be able to quantify the improvements achieved by the methodological developments of the present project, it was intended to test the HRL Grassland 2015 against the training and validation data used by the ECoLaSS project: the Visually Interpreted Reference Points (VIRP). The analysis should confirm the overall thematic accuracy of the HRL Grassland layer 2015 of more than 85% and provide information on most relevant omission and commission errors with other land cover classes as a basis for further methodological enhancements.

As described in Chapter 5.1.1 the VIRP have been extracted from the LUCAS point database and visually interpreted on Sentinel-2 time series from 2016 till 2017 and by using VHR data (like Bing maps) as additional data source. The points extracted represent main land cover classes and features with a minimum mapping unit of at least 900m². To be able to use the VIRP for analysing the quality of the HRL grassland classification, the database was initially adapted to the characteristics of the HRL data set. Each point representing grassland was reviewed and classified as grassland or non-grassland by applying the HRL grassland definition and respective minimum mapping unit of 1ha. The result was a data set with 597 grassland and 2811 non-grassland points that was subsequently combined with the classification of the HRL layer for further analysis. The result of the comparison is presented in the

Table 5-17 below. While the majority of 525 grassland points could be confirmed, 72 grassland points have been assigned to non-grassland in the HRL layer. Moreover, 100 out of 2811 non-grassland VIRP turned out to be grassland in the HRL Grassland layer.

Table 5-17: Comparison between the HRL Grassland layer 2015 and the adapted VIRP 2017.

Class	HRL Grassland 2015	VIRP 2017		Totals
		No GRA	GRA	
No GRA	0	2711	100	2811
	1	72	525	597
Totals		2783	625	3408

Table 5-18: Accuracies for HRL Grassland using the adapted VIRP 2017 as Reference.

Class Name	Producer's Accuracy	95% Confidence Interval		User's Accuracy	95% Confidence Interval	
No GRA/No GRA	97,41%	96,81%	98,02%	96,44%	95,74%	97,15%
GRA/GRA	84,00%	81,05%	86,95%	87,94%	85,24%	90,64%

Overall Accuracy: 94,95%

Based on these results, a confusion matrix (see Table 5-18) was calculated to evaluate the overall thematic accuracy and respective user's and producer's accuracies. The result confirms a high overall thematic accuracy of approx. 95% for the HRL Grassland layer. While the producer's accuracy is slightly below 85%, the user's accuracy provides satisfying values with approx. 88% thematic accuracy.

When analysing the wrong class assignments of the HRL data set, confusions between cropland and managed (mowed) grassland could be identified as the most important reason. Special attention should, therefore, be paid to this effect in the ECoLaSS project. The Grassland HRL 2015 was also tested against the first outcome of the ECoLaSS prototype. The ECoLaSS grassland prototype was created for the reference year 2017 and was trained on the VIRP (see Chapter 5.2.1). This is a first version of this comparison.

The area figures for grassland and non-grassland for both layers are quite similar; however, in general the ECoLaSS grassland prototype layer shows some more grassland than the HRL GRA 2015. Reasons for that will be discussed in the following paragraphs. An overall grassland statistic can be found in Table 5-19.

Table 5-19: Comparative Grassland statistic (HRL 2015, ECoLaSS 2017) for the demonstration site West.

Class	HRL GRA 2015	ECoLaSS grassland prototype 2017
Grassland	10,365 km ² / 83.9%	10,943 km ² / 83.2%
Non-Grassland	53,957 km ² / 16.1%	54,083 km ² / 16.8%
Clouds/Cloud-Shadows	2 km ² / 0.0%	0 km ² / 0.0%
Overall	64,324 km ²	65,023 km ²

The slight difference in the overall area is due to the estuary of Westerschelde, which was classified as non-grassland in the ECoLaSS prototype but excluded from the HRL 2015.

Figure 5-40 shows a visual comparison of both layers highlighting the differences. Patches classified as grassland in both layers are shown in green, whereas additional grassland patches in the ECoLaSS prototype are highlighted in red and additional grassland patches in the HRL 2015 are highlighted in blue.

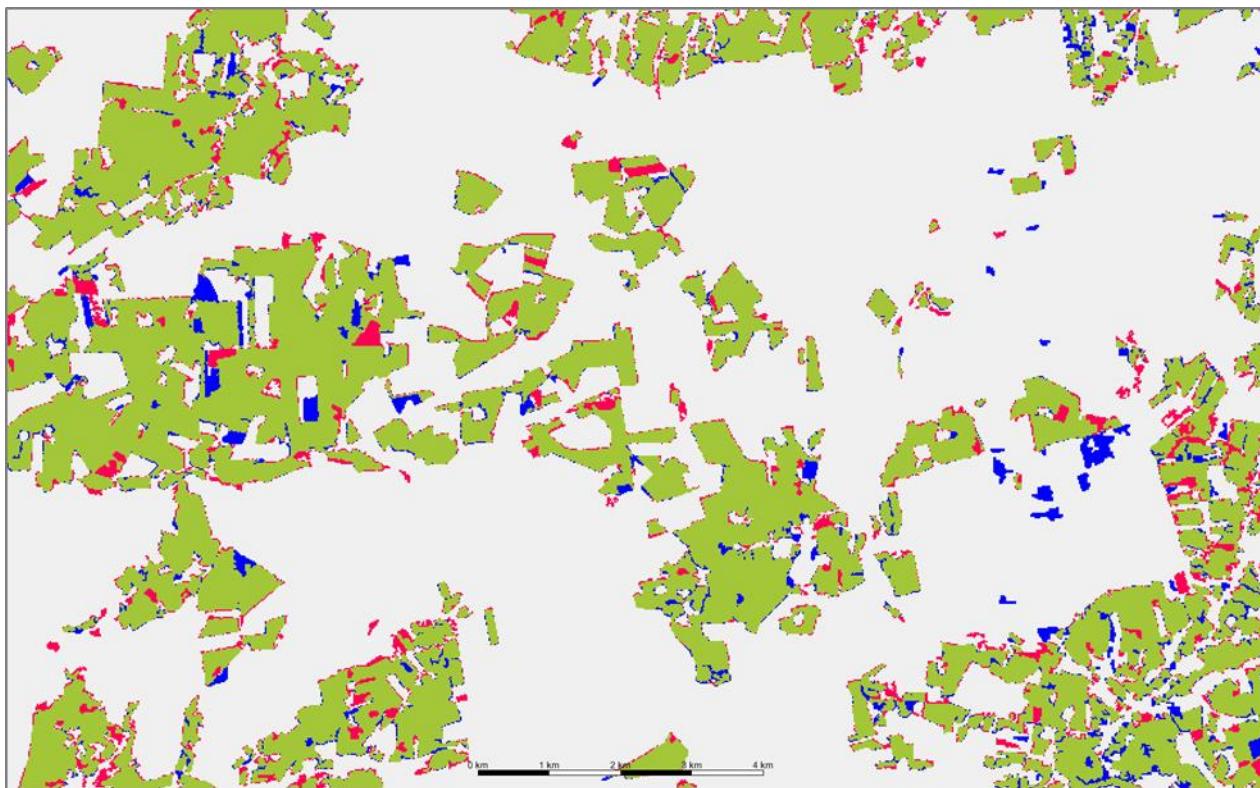


Figure 5-40: Unfiltered comparison of the HRL GRA 2015 and the ECoLaSS prototype 2017: grassland in both layers (green), grassland only in the ECoLaSS prototype (red), and only in the HRL GRA 2015 (blue).

To compare both datasets the ECoLaSS prototype in 10m spatial resolution was resampled to the 20m resolution of the HRL GRA 2015. It is obvious that differences appear continually at the border of grasslands, which is mainly due to the different geometric resolution and the resampling. Another reason could be the different input base data source, which is mainly Sentinel-1 and Sentinel-2 but acquired in different years. Slight geometric differences as well as illumination differences can lead to a different delineation of patches. Moreover, the different classification processes can provoke these deviances. The HRL GRA 2015 included an object-based approach with a segmentation of patches, while the ECoLaSS prototype is constricted to a pixel-based classification approach. Another issue is the MMU of the HRL GRA 2015 of 1ha, which was not applied in the same way to the ECoLaSS prototype. The HRL GRA 2015 allowed only pixel which share a direct border to another pixel (4 neighbouring pixels are taken into account) and groups of pixel needed to be at least 1ha of size. The ECoLaSS prototype allowed also pixel which do not share a direct border but a corner of an adjacent pixel (8 neighbouring pixels are taken into account). Therefore, the ECoLaSS prototype includes also single (diagonal) pixels, as compared to the HRL 2015 which does not.

In the first step of comparison the difference between both layers was calculated but on purpose not yet filtered, to check (i) if also very small grassland patches could have been correctly detected and (ii) if very small patches are mainly noise and misclassifications of mixed pixels in the base data source. These resulting differences, described above, are therefore not real changes, just technical issues, and are not topic of the following interpretation of the differences between both data sets.

As a second step, this difference map between the ECoLaSS prototype 2017 and the HRL GRA 2015 was also filtered to the MMU of the HRL grassland layer 2015 of 1ha. The outcome is shown in Figure 5-41. Most of the technical differences related to single pixel and lines of pixel at borders of grassland patches were removed. Resulting compact patches are a proxy for a possible grassland change product. The following analyses will show that due to different classification approaches the filtered map-to-map difference can still not be considered as a change layer, as it is mainly highlighting misclassifications in one or the other of the two layers.

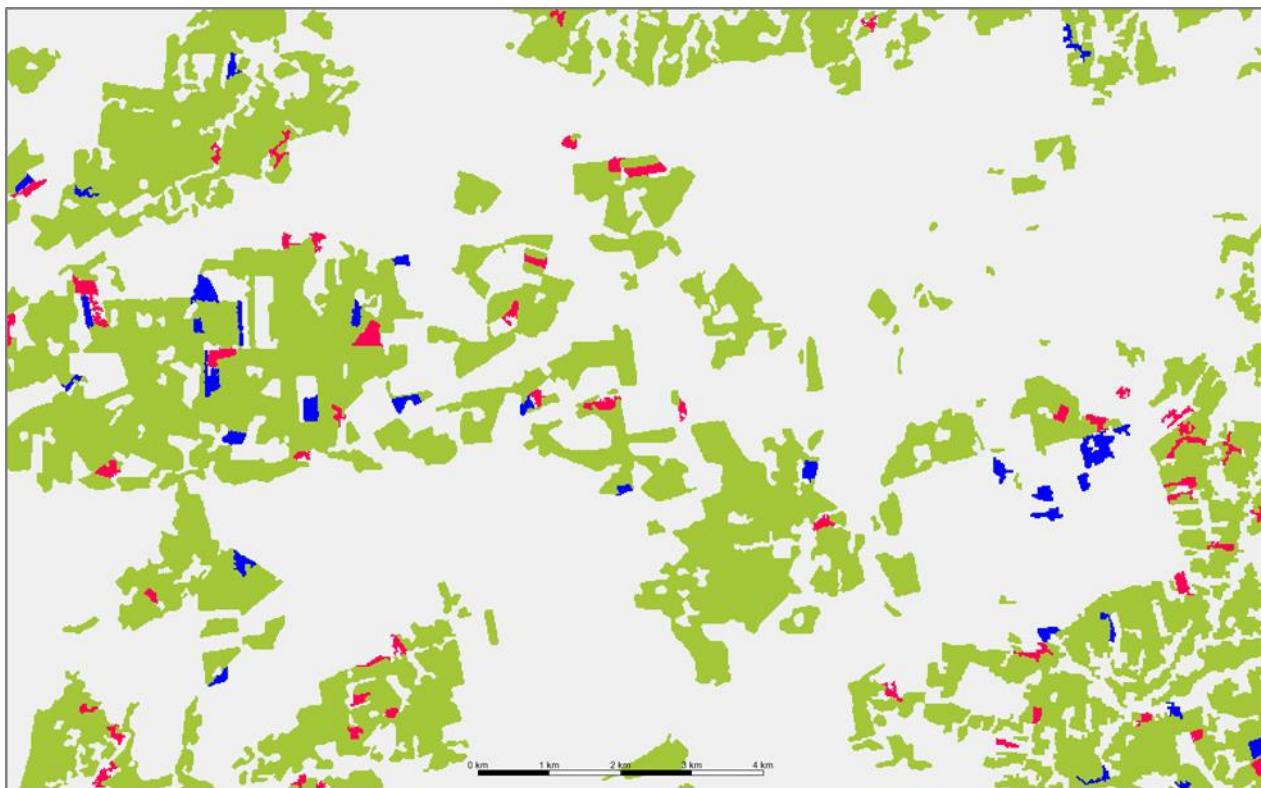


Figure 5-41: Filtered comparison of the HRL GRA 2015 and the ECoLaSS prototype 2017: grassland in both layers (green), grassland only in the ECoLaSS prototype (blue), and only in the HRL GRA 2015 (red).

The following description of misclassifications included a comparison of both datasets and a cross-check against VHR data from GoogleEarth and BingMaps.

SPECIFIC TYPES OF GRASSLAND THAT WERE MISSED OUT IN ONE OF THE LAYERS (OMISSION ERRORS)

Table 5-20 shows types of grassland, which were left out and, therefore, not detected by the classifier in one or the other layer.

Table 5-20: Types of omission errors for HRL GRA 2015 and the ECoLaSS prototype 2017.

Grassland Type	HRL GRA 2015	ECoLaSS prototype 2017
Urban green areas like sport fields, parks and residential areas	Partly missing	Partly missing
Natural Grassland on military sites	Partly missing	Missing

Reasons for these omissions in the ECoLaSS 2017 layer could be missing reference samples of these classes (in the VIRP). Reasons for omissions in the HRL GRA 2015 for urban areas are mainly related to the MMU, as urban green areas are very often mixed with trees, buildings etc. and do therefore often not fulfill the MMU of the HRL GRA 2015 of at least 1ha. It was expected that the ECoLaSS prototype will show some more urban green areas due to the better resolution. This was partly the case, however, not valid for all urban areas. Even without filtering no systematic caption of small urban green areas could be determined.

The Omissions of natural grassland on military sites are related to confusion of this land cover feature with low density shrub and forest areas, but the results for this land cover type are much more compact within the HRL GRA 2015 than in the ECoLaSS prototype 2017. Compact patches better represent the natural grassland features on the military sites and therefore it will be targeted to improve the capturing of such sites.

SPECIFIC TYPES OF NON-GRASSLAND THAT ARE WRONGLY OVERREPRESENTED IN ONE OF THE LAYERS (COMMISSION ERRORS)

Table 5-21 shows types of non-grassland by definition, which are overrepresented in the grassland mask in one or the other layer.

Table 5-21: Types of commission errors for HRL GRA 2015 and the ECoLaSS prototype 2017.

Grassland Type	HRL GRA 2015	ECoLaSS prototype 2017
Low density urban areas	Not detected	Misclassification of garden plots with certain percentage of buildings and urban infrastructures (relevant for the whole demonstration site).
Cultivated and managed areas of agricultural use	Not detected	Misclassification of e.g. Orchards and managed grassland with fruit trees (southern part of the demonstration site).
Moore and heathland	Not included in Grassland by definition, but still included in the HRL GRA 2015 at least in some areas over Europe.	Successfully excluded from the grassland mask of the ECoLaSS grassland prototype (especially northern part of the demonstration site)

Reasons for these commissions in the ECoLaSS 2017 layer could be missing reference samples of these classes (in the VIRP). Reasons for commissions in the HRL GRA 2015 for moor and heathland are mainly related to spectral confusion with wet grassland. The successful exclusion of these land cover features is a great improvement of the ECoLaSS prototype 2017.

The misclassifications present in the ECoLaSS prototype 2017 related to orchards and garden plots were successfully avoided within the HRL GRA 2015 project, in particular by the implementation of SAR EO data. As the ECoLaSS 2017 prototype layer is based on dense time series of both, optical and SAR data, the removal of these commissions should be further investigated in the second project phase, and is to be tested with a focus on SAR data.

DIFFERENCES BETWEEN THE TWO LAYERS RELATED TO REAL CHANGES BETWEEN THE YEARS 2015 AND 2017, WHICH CAN BE AN INDICATION ABOUT APPROPRIATE UPDATE CYCLES OF SUCH A PRODUCT

Based on inspection of selected change patches it can be concluded that most of the detected differences between the two products (omissions and commissions) are mainly related to technical changes. E.g., within the demonstration site no agricultural site was found where a change of the actual land use to grassland took place. Likewise, no conversion from grassland into other land covers such as urban areas was found. Therefore, the differences between the two layers are mainly caused by the classification errors described above or due to the differences of the MMU. As not every single patch was reviewed, it could still be the case that also several real changes are included in such a proxy for a possible change product. In the second phase it will be investigated how such a simple map-to-map difference layer could be separated into (i) real change classes, to derive a real change layer and (ii) technical changes, which could be used to improve the former and consequently, the grassland mask for future improved HRL products.

APPROPRIATE MMU FOR A 10M PRODUCT

Regarding the appropriate MMU for future 10m resolution grassland products, further testing shall follow in phase two of ECoLaSS. Though not all green urban areas were included in the ECoLaSS prototype 2017, some improvements, with respect to the HRL GRA 2015, have been found. These revealed that patches even smaller than the HRL GRA 2015 MMU of 1ha could have been detected. The

appropriate MMU size of a new 10m product still is under investigation and has been further explored in the second project phase.

5.3 Prototype Specifications

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. 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 will be applied to all raster prototypes and associated reference files and is documented in the Deliverables of Task 4.

The following 22 prototypes files as part of **D13.4 – P43.2b – Data Sets HRL Permanent Grassland Products (Issue 2)** for Grasslands in the demosites West, Central and South-East were submitted:

- GRA_2017_010m_WE_03035_prototype_v01.tif
- GRA_2018_010m_WE_03035_prototype_v01.tif
- GRA_2017_010m_WE_03035_probability_v01.tif
- GRA_2018_010m_WE_03035_probability_v01.tif
- GRA_2017_010m_CE_03035_prototype_v01.tif
- GRA_2018_010m_CE_03035_prototype_v01.tif
- GRA_2017_010m_CE_03035_probability_v01.tif
- GRA_2018_010m_CE_03035_probability_v01.tif
- GRA_2017_010m_SE_03035_prototype_v01.tif
- GRA_2018_010m_SE_03035_prototype_v01.tif
- GRA_2017_010m_SE_03035_probability_v01.tif
- GRA_2018_010m_SE_03035_probability_v01.tif
- GRA_2018_010m_SE_03035_mowing_v01.tif
- GRU_2018_010m_WE_03035_prototype_v01.tif
- GRA_2018_010m_CE_03035_mowing_v01.tif
- GRU_2018_010m_CE_03035_prototype_v01.tif
- GRA_2018_010m_SE_03035_mowing_v01.tif
- GRU_2018_010m_SE_03035_prototype_v01.tif
- GRC_1718_010m_WE_03035_prototype_v01.tif
- GRC_1518_020m_CE_03035_prototype_v01.tif
- GRC_1718_010m_CE_03035_prototype_v01.tif
- GRC_1718_010m_SE_03035_prototype_v01.tif

The file names generally contain the following 7 main aspects:

THEME YEAR RESOLUTION EXTENT EPSG TYPE VERSION

THEME

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

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:

"GRA_2017_010m_WE_03035_prototype_v01.tif" stands for: Grassland, 2017 reference year, 10m, Demonstration-site West, European projection (EPSG: 3035), prototype, version 01

The product specifications of the ECoLaSS grassland status prototypes 2017 and 2018 are shown in Table 5-22.

Table 5-22: Detailed specifications for the Improved Primary Status Grassland Layers.

Grassland 10m	Acronym	Product category			
	GRA	Improved Primary Status Layer			
Reference year					
2017					
Geometric resolution					
Pixel resolution 10m x 10m, fully conform with the EEA reference grid					
Coordinate Reference System					
European ETRS89 LAEA projection					
Geometric accuracy (positioning scale)					
Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.					
Thematic accuracy					
Minimum 85% overall accuracy					
Data type					
8bit unsigned raster with LZW compression					
Minimum Mapping Unit (MMU)					
Pixel-based (1ha)					
Necessary attributes					
Raster value, count, class name, area (in km ²), area percentage (taking outside area not into account)					
Raster coding (thematic pixel values)					
0: all non-grass areas					
1: Grassy and non-woody vegetation					
254: unclassifiable (no satellite image available, or clouds, or shadows)					
255: outside area					
Metadata					
XML metadata files according to INSPIRE metadata standards					
Delivery format					
GeoTIFF					
Colour table					
ArcGIS *.clr format					
Class Code	Class Name	Red	Green	Blue	
0	all non-grass areas	240	240	240	
1	Grassy and non-woody vegetation	70	158	74	
254	unclassifiable (no satellite image available, or clouds, or shadows)	153	153	153	
255	outside area	0	0	0	

In addition to the prototype layers, probability layers are provided as a pixel-based quality indicator. The additional 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 5-23.

Table 5-23: Colour palette for the probability layers.

Probability Layers					
Class Code	Class Name	Red	Green	Blue	
0	0% probability	245	0	0	
50	50% probability	245	241	0	
100	100% probability	20	245	0	
101	areas excluded by binary mask	128	128	128	
255	outside area	0	0	0	

Further grassland change prototypes for 2017 - 2018 are produced and their specifications are given in Figure 5-19.

Table 5-24: Detailed specifications for the New Status Layer Grassland Change.

Grassland Change	Acronym	Product category
Reference year	GRC	Incremental Update Layer
2017/2018		
Extent		
Demonstration sites West, Central & South-East		
Geometric resolution		
Pixel resolution 10m x 10m		
Coordinate Reference System		
European ETRS89 LAEA projection		
Geometric accuracy (positioning scale)		
Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy		
80-85% overall accuracy		
Data type		
8bit unsigned integer raster with LZW compression		
Minimum Mapping Unit (MMU)		
0,05ha for grassland and 0,5ha for changes		
Necessary attributes		
Raster value, count, class name, area (in km ²), percentage (taking outside area not into account)		
Raster coding (thematic pixel values)		
0: non-grass areas		
10: unchanged grass areas		

1: grassland gain																																										
2: grassland loss																																										
254: unclassifiable (no satellite image available, or clouds, or shadows)																																										
255: outside area																																										
Metadata																																										
XML metadata files according to INSPIRE metadata standards																																										
Delivery format																																										
GeoTIFF																																										
Colour table																																										
ArcGIS *.clr format																																										
<table border="1"> <thead> <tr> <th>Class Code</th> <th>Class Name</th> <th>Red</th> <th>Green</th> <th>Blue</th> <th></th> </tr> </thead> <tbody> <tr> <td>0</td> <td>non-grass areas</td> <td>240</td> <td>240</td> <td>240</td> <td></td> </tr> <tr> <td>10</td> <td>unchanged grass areas</td> <td>225</td> <td>225</td> <td>225</td> <td></td> </tr> <tr> <td>1</td> <td>grassland gain</td> <td>28</td> <td>72</td> <td>201</td> <td></td> </tr> <tr> <td>2</td> <td>grassland loss</td> <td>255</td> <td>0</td> <td>0</td> <td></td> </tr> <tr> <td>254</td> <td>unclassifiable (no satellite image available, or clouds, or shadows)</td> <td>153</td> <td>153</td> <td>153</td> <td></td> </tr> <tr> <td>255</td> <td>outside area</td> <td>0</td> <td>0</td> <td>0</td> <td></td> </tr> </tbody> </table>	Class Code	Class Name	Red	Green	Blue		0	non-grass areas	240	240	240		10	unchanged grass areas	225	225	225		1	grassland gain	28	72	201		2	grassland loss	255	0	0		254	unclassifiable (no satellite image available, or clouds, or shadows)	153	153	153		255	outside area	0	0	0	
Class Code	Class Name	Red	Green	Blue																																						
0	non-grass areas	240	240	240																																						
10	unchanged grass areas	225	225	225																																						
1	grassland gain	28	72	201																																						
2	grassland loss	255	0	0																																						
254	unclassifiable (no satellite image available, or clouds, or shadows)	153	153	153																																						
255	outside area	0	0	0																																						

Regarding the demonstration site Central grassland changes are derived between 2015 and 2018. The product specification are given in the Table below.

Table 5-25: Detailed specifications for the New Status Layer Grassland Change.

Grassland Change	Acronym	Product category
	GRC	Incremental Update Layer
Reference year		
2015/2018		
Extent		
Central		
Geometric resolution		
Pixel resolution 20m x 20m		
Coordinate Reference System		
European ETRS89 LAEA projection		
Geometric accuracy (positioning scale)		
Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy		
80-85% overall accuracy		
Data type		
8bit unsigned integer raster with LZW compression		

Minimum Mapping Unit (MMU)																																										
0.05ha for grassland and 0.5ha for changes																																										
Necessary attributes																																										
Raster value, count, class name, area (in km2), percentage (taking outside area not into account)																																										
Raster coding (thematic pixel values)																																										
0: non-grass areas																																										
10: unchanged grass areas																																										
1: grassland gain																																										
2: grassland loss																																										
254: unclassifiable (no satellite image available, or clouds, or shadows)																																										
255: outside area																																										
Metadata																																										
XML metadata files according to INSPIRE metadata standards																																										
Delivery format																																										
GeoTIFF																																										
Colour table																																										
ArcGIS *.clr format																																										
<table border="1"> <thead> <tr> <th>Class Code</th> <th>Class Name</th> <th>Red</th> <th>Green</th> <th>Blue</th> <th></th> </tr> </thead> <tbody> <tr> <td>0</td> <td>non-grass areas</td> <td>240</td> <td>240</td> <td>240</td> <td></td> </tr> <tr> <td>10</td> <td>unchanged grass areas</td> <td>225</td> <td>225</td> <td>225</td> <td></td> </tr> <tr> <td>1</td> <td>grassland gain</td> <td>28</td> <td>72</td> <td>201</td> <td></td> </tr> <tr> <td>2</td> <td>grassland loss</td> <td>255</td> <td>0</td> <td>0</td> <td></td> </tr> <tr> <td>254</td> <td>unclassifiable (no satellite image available, or clouds, or shadows)</td> <td>153</td> <td>153</td> <td>153</td> <td></td> </tr> <tr> <td>255</td> <td>outside area</td> <td>0</td> <td>0</td> <td>0</td> <td></td> </tr> </tbody> </table>	Class Code	Class Name	Red	Green	Blue		0	non-grass areas	240	240	240		10	unchanged grass areas	225	225	225		1	grassland gain	28	72	201		2	grassland loss	255	0	0		254	unclassifiable (no satellite image available, or clouds, or shadows)	153	153	153		255	outside area	0	0	0	
Class Code	Class Name	Red	Green	Blue																																						
0	non-grass areas	240	240	240																																						
10	unchanged grass areas	225	225	225																																						
1	grassland gain	28	72	201																																						
2	grassland loss	255	0	0																																						
254	unclassifiable (no satellite image available, or clouds, or shadows)	153	153	153																																						
255	outside area	0	0	0																																						

Table 5-26: Detailed specifications for the New Status Layer Grassland Mowing intensity.

Grassland Use Intensity	Acronym	Product category
	GRU	New Primary Status Layer
Reference year		
2018		
Extent		
Demonstration sites West, Central & South-East		
Geometric resolution		
Pixel resolution 10m x 10m		
Coordinate Reference System		
European ETRS89 LAEA projection		

Geometric accuracy (positioning scale)

Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.

Thematic accuracy

Data type

8bit unsigned integer raster with LZW compression

Minimum Mapping Unit (MMU)

0,05ha for grassland

Necessary attributes

Raster value, count, class name, area (in km²), percentage (taking outside area not into account)

Raster coding (thematic pixel values)

0: non-grass areas

1: extensive mowed grassland

2: intensive mowed grassland

254: unclassifiable (no satellite image available, or clouds, or shadows)

255: outside area

Metadata

XML metadata files according to INSPIRE metadata standards

Delivery format

GeoTIFF

Colour table

ArcGIS *.clr format

Class Code	Class Name	Red	Green	Blue	
0	non-grass areas	240	240	240	
1	Extensive mowed grassland	64	178	0	
2	Intensive mowed grassland	255	0	0	
254	unclassifiable (no satellite image available, or clouds, or shadows)	153	153	153	
255	outside area	0	0	0	

Table 5-27: Detailed specifications for the New Status Layer Grassland Mowing Events.

Grassland Mowing Events	Acronym	Product category
	GRA	New Primary Status Layer
Reference year		
2018		
Extent		
Demonstration sites West, Central & South-East		
Geometric resolution		
Pixel resolution 10m x 10m		
Coordinate Reference System		
European ETRS89 LAEA projection		
Geometric accuracy (positioning scale)		
Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.		
Thematic accuracy		
Data type		
8bit unsigned integer raster with LZW compression		
Minimum Mapping Unit (MMU)		
0,05ha for grassland		
Necessary attributes		
Raster value, count, class name, area (in km ²), percentage (taking outside area not into account)		
Raster coding (thematic pixel values)		
0-100: grassland mowing events		
254: unclassifiable (no satellite image available, or clouds, or shadows)		
255: outside area		
Metadata		
XML metadata files according to INSPIRE metadata standards		
Delivery format		
GeoTIFF		
Colour table		
ArcGIS *.clr format		

Mowing Events					
Class Code	Class Name	Red	Green	Blue	
0-100	Mowing Events	-	-	-	
101	Areas excluded by binary Mask	128	128	128	
255	Outside Area	0	0	0	

Conclusion and Outlook

The objective of this WP is to develop a framework for an improved identification of grassland areas using Sentinel time series with the aim to develop a prototype of a European HR Grassland Layer. It should comprise following components, i.e. a high thematic accuracy, optical and SAR data integration, increased spatial resolution and high automation level as well as (mid-term) increased thematic content.

To demonstrate the operational applicability of the methods, they were applied on representative demonstration sites for large area mapping. The “Alpine/Central Europe” demonstration site is representative for the Continental and Alpine bio-geographic region, the “West Europe” site for the Atlantic and Continental, and the “South-East Europe” site for the Mediterranean, Continental and Alpine bio-geographic regions.

The basis of this working task relies on the results achieved in Task 3, where extensive testing and benchmarking of various approaches have been undertaken. Therefore, we applied a supervised classification approach using the Random Forest classifier, which belongs to the ensemble learning methods. LUCAS sample plot data, which are available for most EU countries at a regular basis, are the main reference data set. This facilitates future operational EU-wide rollout, as usage of already available reference data significantly reduces the required effort for large area applications.

As the results of Task 3 have shown, the synergetic use of temporal features derived from optical and SAR data streams enhances the accuracy of the classifications in comparison to either single optical or SAR only approaches. High-frequency optical and SAR acquisitions over the growing seasons were therefore used to derive temporal features over the grassland growing season as input for the classifications. The generation of suitable time features, especially considering upscaling to pan-European or global levels, is challenging and requires large computational capacities. Automated feature selection at the bio-geographic level is therefore applied, to reduce the computational effort for potential future operational large area roll-out. While attributing an absolute importance to each feature is not straight forward due to their high co-linearity, the ability to reduce the number of features which need to be computed beforehand without compromising mapping performance is the most important outcome. From the perspective of an operational implementation, a valuable observation is that the number of input features may remain moderate around at a maximum of 50 predictors. In addition to the classification results, a probability layer indicating the percentage of reliability constitutes a pixel-based quality indicator by-product from the random forest classification. We recommend that such “expert product” is also derived in case of future operational roll-out, as it can facilitate harmonising long term time series e.g. for continuous monitoring of grasslands.

Additionally, Grassland Use Intensity products are provided within the grasslands status layer 2018. The use intensity products represent intensively/extensively mowed grasslands as binary product. The intensive mowed grasslands are defined by three or more detected mowing events and the extensive mowed grasslands by less than three detected mowing events. Unfortunately, reference data regarding mowing intensity is hardly available, e.g. for the demonstration sites, only the IACS dataset for Austria provides information on the number of mowing events. Therefore, the mowing intensity prototypes implemented within the demonstration sites are merely considered a proof-of-concept.

For grassland change monitoring, approaches developed within WP34 by combining signal anomaly detection with map-to-map comparison are applied. The results are demonstrated for wall-to-wall monitoring of grassland changes in the demonstration sites for the years 2017 to 2018. In addition, 2017 status layers are also compared with already existing COPERNICUS grassland HRL 2015 by map-to-map comparison. For potential future rollout, we recommend to combine the signal-anomaly-detection approach with the map-to-map comparison approach as implemented for the demonstration change products 2017 to 2018. However, for future roll-out we recommend implementation of continuous grassland monitoring over longer time periods. To provide consistent time-series over longer time periods, above described status probability layers expert products can be used as inputs in the aggregation process.

To summarize the improvements compared to the existing HRL grassland 2015, we achieved:

- **improved level of automation** to allow a faster production and shorter monitoring intervals (e.g. continuous monitoring with yearly updates)
- Improve the **thematic classification accuracy**
- Fully exploit **optical Sentinel-2 and SAR Sentinel-1 time series** instead of using pre-selected, best-suited optical EO data scenes
- Design a fully **integrated SAR/optic** time series data analysis to benefit from the multi-sensor characteristics
- Provide a **seamless, wall-to-wall product** (e.g. no cloud cover gaps)
- Provide information on **mowing intensity**, which can be used as valuable input in downstream applications such as for example related to the assessment of grassland management intensity
- Improve the status layer's detail from 20m **spatial resolution to 10m**
- Provide a **change detection** approach to detect grassland increase and decrease

The successful application demonstrates the potential for a future large area rollout of the developed methods.

References

- Ali, Jehad, et al. (2012): "Random forests and decision trees." *International Journal of Computer Science Issues (IJCSI)* 9.5, 272.
- Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R.J., Li, H. & Moran, M.S.J. (2000, July). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In *Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, USA* (Vol. 1619).
- Belgiu, Mariana, and Lucian Drăguț (2016): "Random forest in remote sensing: A review of applications and future directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24-31.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 4(1), 5-32.
- Büttner G., Kosztra B., Maucha G., Pataki R. (2012): Implementation and achievements of CLC2006, ETC-LUSI, EEA, 65.
- Cho, M. A., & Skidmore, A. K. (2009). Hyperspectral predictors for monitoring biomass production in Mediterranean mountain grasslands: Majella National Park, Italy. *International Journal of Remote Sensing*, 30(2), 499-515.
- Clevers, J. G., & Kooistra, L. (2012). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2), 574-583.
- Clevers, J. G., & Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351.
- Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey Iii, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote sensing of Environment*, 74(2), 229-239.
- Dymond, C. C., Mladenoff, D. J., & Radeloff, V. C. (2002). Phenological differences in Tasseled Cap indices improve deciduous forest classification. *Remote Sensing of Environment*, 80(3), 460-472.
- Fernández-Manso, A., Fernández-Manso, O., & Quintano, C. (2016). Sentinel-2A red-edge spectral indices suitability for discriminating burn severity. *International journal of applied earth observation and geoinformation*, 50, 170-175.
- EEA (2015). Biogeographical regions dataset. Available: <http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3#tab-gis-data>. European Environment Agency (EEA).
- Foody, G. M., & Dash, J. (2007). Discriminating and mapping the C3 and C4 composition of grasslands in the northern Great Plains, USA. *Ecological Informatics*, 2(2), 89-93.
- Gallego, F. J. (1995). *Sampling frames of square segments*. Office for Official Publ. of the European Communities.
- Gallego, J. and JRC-IES, I. I. (2004). Area Frames for Land Cover Estimation: Improving the European LUCAS Survey. Presented at the Proceedings of the 3rd World Conference on Agricultural and Environmental Statistical Application, Cancun, Mexico, 2–4.

Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3), 257-266.

Gitelson, A., & Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology*, 143(3), 286-292.

Gitelson, A. A., Vina, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, 32(8).

Henrich, V., Krauss, G., Götze, C., Sandow, C. (2012). IDB. URL: <http://www.indexdatabase.de> (Entwicklung einer Datenbank für Fernerkundungsindizes. AK Fernerkundung, Bochum).

Horler, D. N. H., DOCKRAY, M., & Barber, J. (1983). The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, 4(2), 273-288.

Horning, Ned (2010): "Random Forests: An algorithm for image classification and generation of continuous fields data sets." Proceedings of the International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences, Osaka, Japan. Vol. 911.

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote sensing of environment*, 25(3), 295-309.

Huete, A. R., Liu, H. Q., Batchily, K., & Van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *REMOTE SENSING OF ENVIRONMENT-NEW YORK-*, 59, 440-451.

Kawamura, K., Akiyama, T., Yokota, H. O., Tsutsumi, M., Yasuda, T., Watanabe, O., & Wang, S. (2005). Comparing MODIS vegetation indices with AVHRR NDVI for monitoring the forage quantity and quality in Inner Mongolia grassland, China. *Grassland Science*, 51(1), 33-40.

Keil, M., Esch, T., Divanis, A., Marconcini, M., Metz, A., Ottinger, M., Voinov, S., Wiesner, M., Wurm, M. & Zeidler, J. (2015). Updating the Land Use and Land Cover Database CLC for the Year 2012-„Backdating“ of DLM-DE of the Reference Year 2009 to the Year 2006.

Liaw, Andy, and Matthew Wiener (2002): "Classification and regression by random Forest." *R news* 2.3 18-22.

Louppe, G., Wehenkel, L., Sutera, A., & Geurts, P. (2013): Understanding variable importances in forests of randomized trees. In *Advances in neural information processing systems*, 431-439.

Magiera, A., Feilhauer, H., Waldhardt, R., Wiesmair, M., & Otte, A. (2017). Modelling biomass of mountainous grasslands by including a species composition map. *Ecological Indicators*, 78, 8-18.

Olofsson, P., Foody, G. M., Stehman, S. V., and Woodcock, C. E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129, 122–131.

Rossini, M., Migliavacca, M., Galvagno, M., Meroni, M., Cogliati, S., Cremonese, E., Fava, F., Gitelson, A., Julitta, T., Morra di Cellia, U., Siniscalco, C. & Colombo, R. (2014). Remote estimation of grassland gross primary production during extreme meteorological seasons. *International Journal of Applied Earth Observation and Geoinformation*, 29, 1-10.

Rüetschi, M., Schaepman, M. E., & Small, D. (2017). Using Multi-temporal Sentinel-1 C-band Backscatter to Monitor Phenology and Classify Deciduous and Coniferous Forests in Northern Switzerland. *Remote Sensing*, 10(1), 55.

Samadzadegan, F. (2004). Data integration related to sensors, data and models. In *XXth ISPRS Congress. Istanbul, Turkey, ISPRS*.

Schmidt, T., Schuster, C., Kleinschmit, B., & Förster, M. (2014). Evaluating an Intra-Annual Time Series for Grassland Classification—How Many Acquisitions and What Seasonal Origin Are Optimal?. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8), 3428-3439.

Selkowitz, D. J. and Stehman, S. V. (2011). Thematic accuracy of the National Land Cover Database (NLCD) 2001 land cover for Alaska. *Remote Sensing of Environment*, 115 (6), 1401–1407.

Stehman, S. V. and Czaplewski, R. L., (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, 64 (3), 331–344.

Suhet (2015). Sentinel-2 User Handbook. E. S. Agency, Ed.,1.2 ed.

Tong, A., & He, Y. (2017): Estimating and mapping chlorophyll content for a heterogeneous grassland: Comparing prediction power of a suite of vegetation indices across scales between years. *ISPRS Journal of Photogrammetry and Remote Sensing*, 126, 146-167.

Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M. (1999). Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote sensing of environment*, 70(1), 52-68.

Vicente, L. E., LOEBMANN, D. D. S., Zillmann, E., González, A., Vicente, A. K., de OLIVEIRA, B. P., & de Araújo, L. S. Assessment of vegetational indices applied to sugarcane monitoring using Rapideye images. In *Embrapa Meio Ambiente-Artigo em anais de congresso (ALICE)*. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 18., 2017, Santos. Anais Santos: Inpe, 2017. Trabalho 59847.

Wang, C., Jamison, B. E., & Spicci, A. A. (2010). Trajectory-based warm season grassland mapping in Missouri prairies with multi-temporal ASTER imagery. *Remote sensing of environment*, 114(3), 531-539.

Waske, B., & van der Linden, S. (2008). Classifying multilevel imagery from SAR and optical sensors by decision fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5), 1457-1466.

Woodhouse, I. H. (2017). *Introduction to Microwave Remote Sensing*. CRC press.

Wu, C., Niu, Z., Tang, Q., & Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and forest meteorology*, 148(8-9), 1230-1241.

Wu, C., Niu, Z., & Gao, S. (2012). The potential of the satellite derived green chlorophyll index for estimating midday light use efficiency in maize, coniferous forest and grassland. *Ecological Indicators*, 14(1), 66-73.

Xu, D., Guo, X., Li, Z., Yang, X., & Yin, H. (2014). Measuring the dead component of mixed grassland with Landsat imagery. *Remote Sensing of Environment*, 142, 33-43.

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: a review of developments and applications. *Journal of Sensors*, 2017.

Yang, Y. H., Fang, J. Y., Pan, Y. D., & Ji, C. J. (2009). Aboveground biomass in Tibetan grasslands. *Journal of Arid Environments*, 73(1), 91-95.

Yang, X., Smith, A. M., & Hill, M. J. (2017). Updating the Grassland Vegetation Inventory Using Change Vector Analysis and Functionally-Based Vegetation Indices. *Canadian Journal of Remote Sensing*, 43(1), 62-78.

Zillmann, E., Weichelt, H., Herrero, E. M., Esch, T., Keil, M., & van Wolvelaer, J. (2013, June). Mapping of grassland using seasonal statistics derived from multi-temporal satellite images. In *Analysis of Multi-temporal Remote Sensing Images, MultiTemp 2013: 7th International Workshop on the* (pp. 1-3). IEEE.

Zillmann, E., Gonzalez, A., Herrero, E. J. M., van Wolvelaer, J., Esch, T., Keil, M., Weichelt, H. & Garzón, A. M. (2014). Pan-European grassland mapping using seasonal statistics from multisensor image time series. *IEEE Journal of selected topics in applied Earth Observations and Remote Sensing*, 7(8), 3461-3472.

Annexe 1

In the following, a detailed overview of the signal analysis of the five different grassland types present in the demonstration site WEST is given showing besides the MEAN spectral signature of the different indices for all five grassland types together also the MEAN +/- the standard deviation for the five grassland types separately.

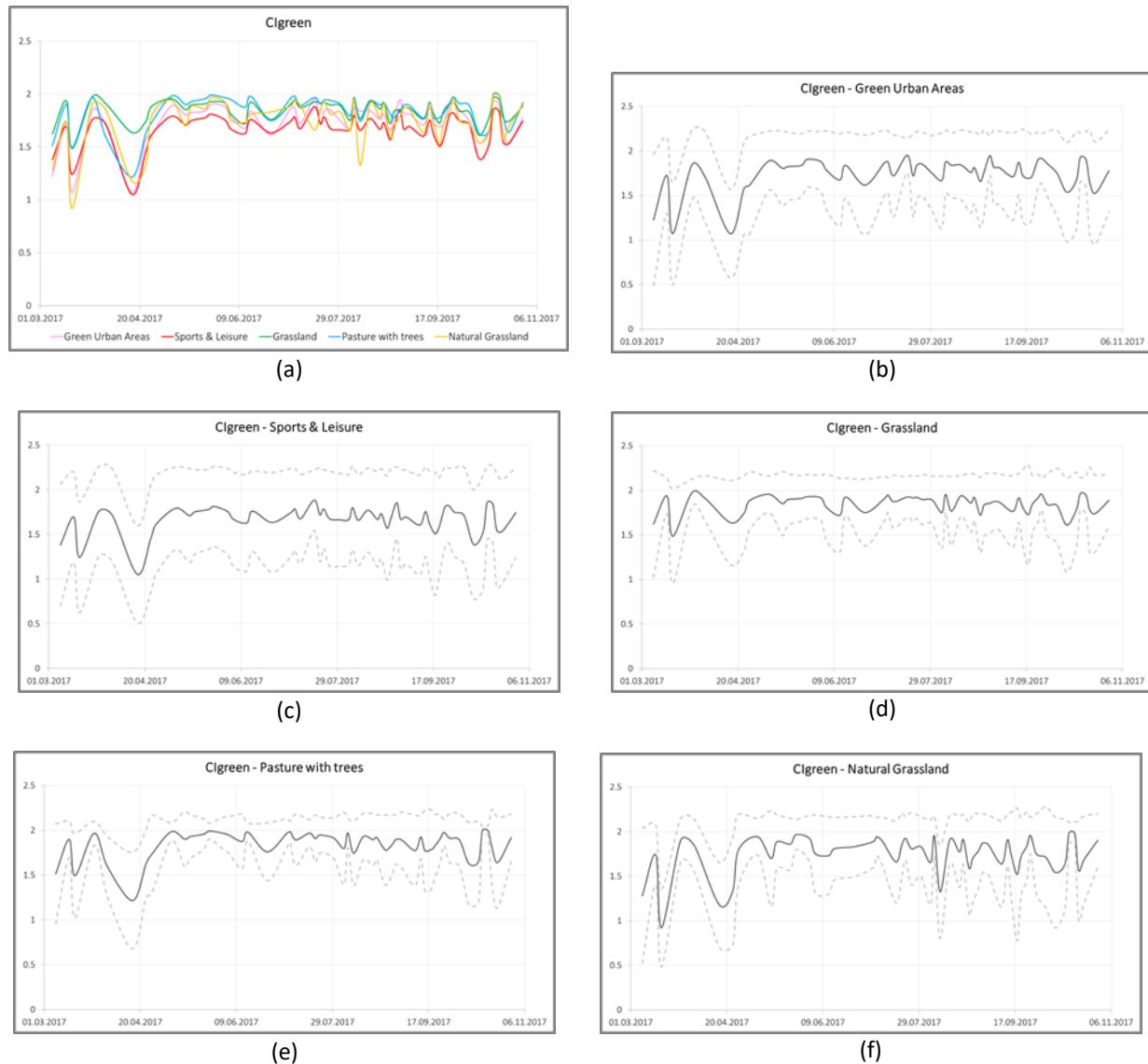


Figure 5-42: Overview of the MEAN spectral signature of the index Clgreen for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

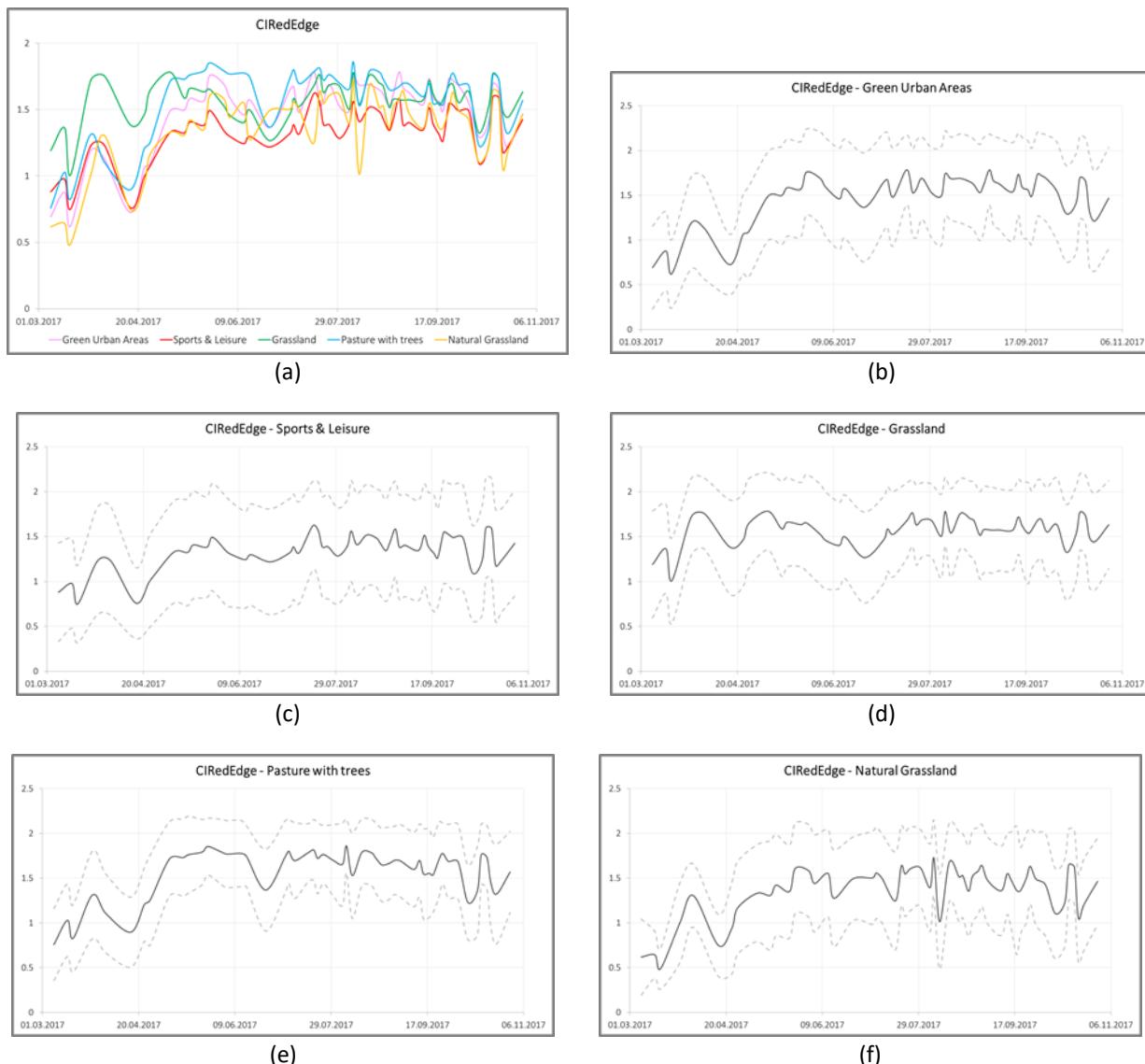
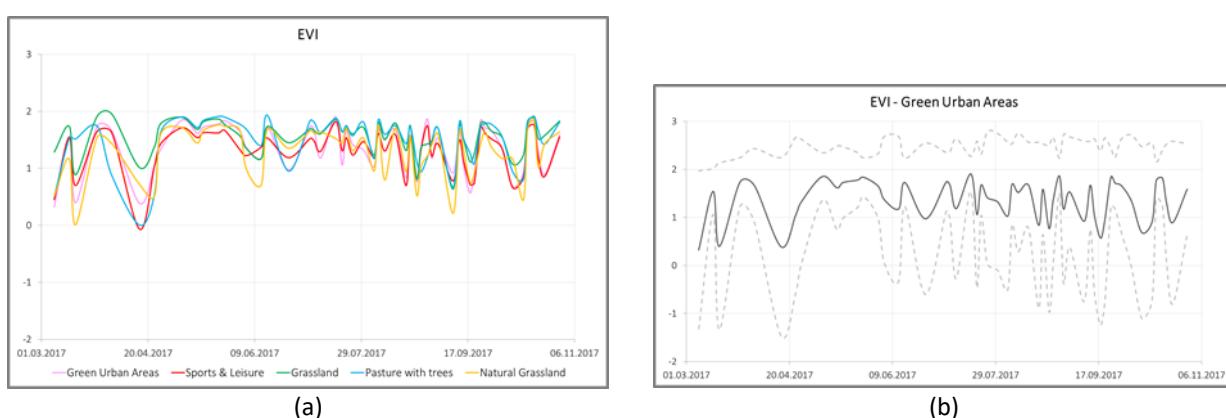


Figure 5-43: Overview of the MEAN spectral signature of the index CIRedEdge for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.



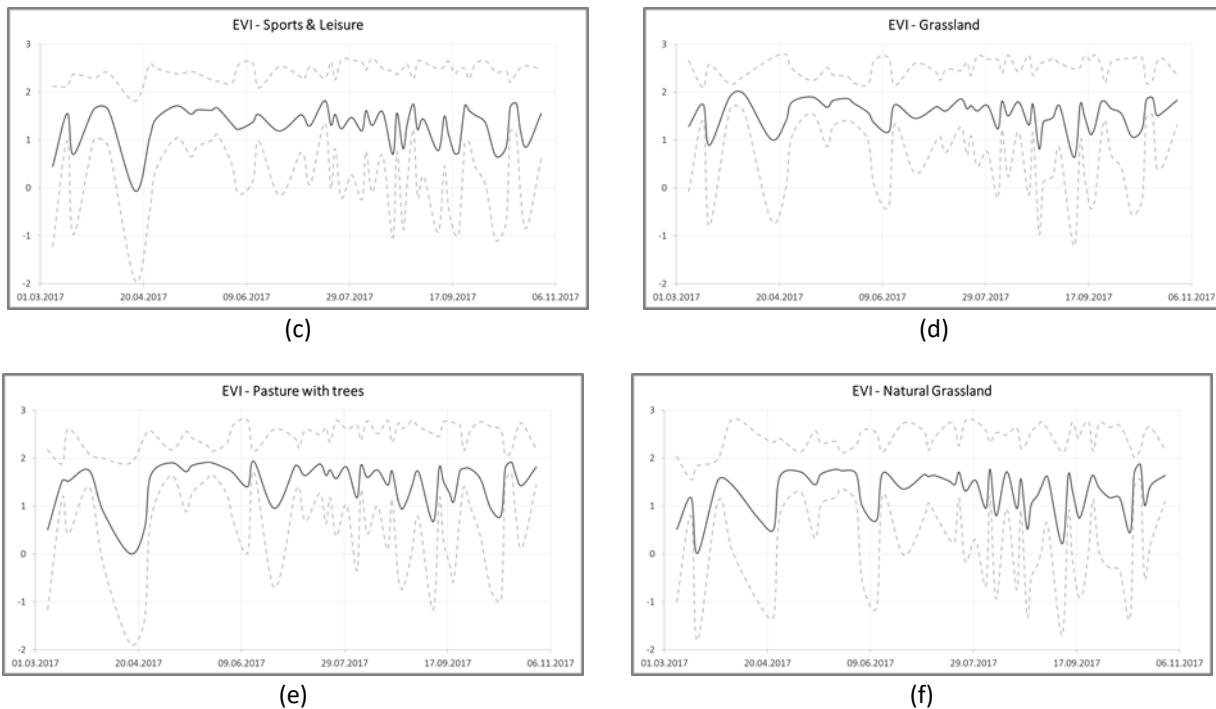
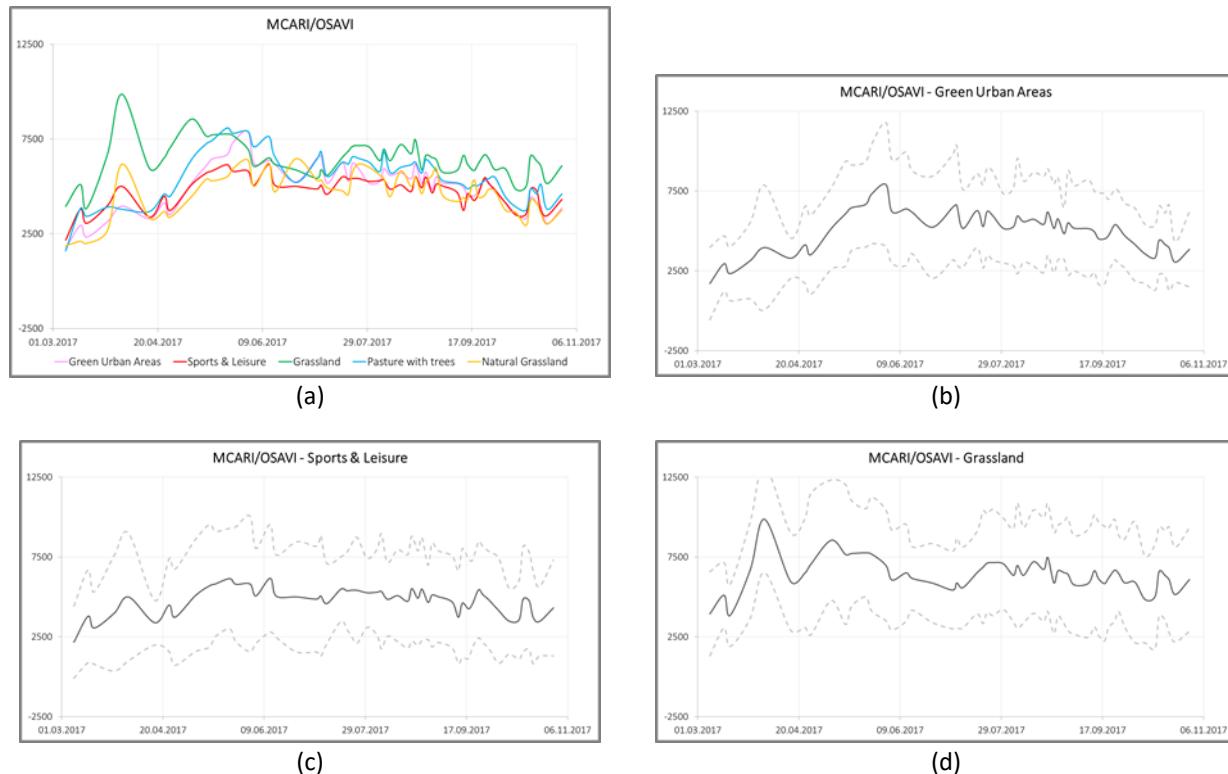


Figure 5-44: Overview of the MEAN spectral signature of the index EVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.



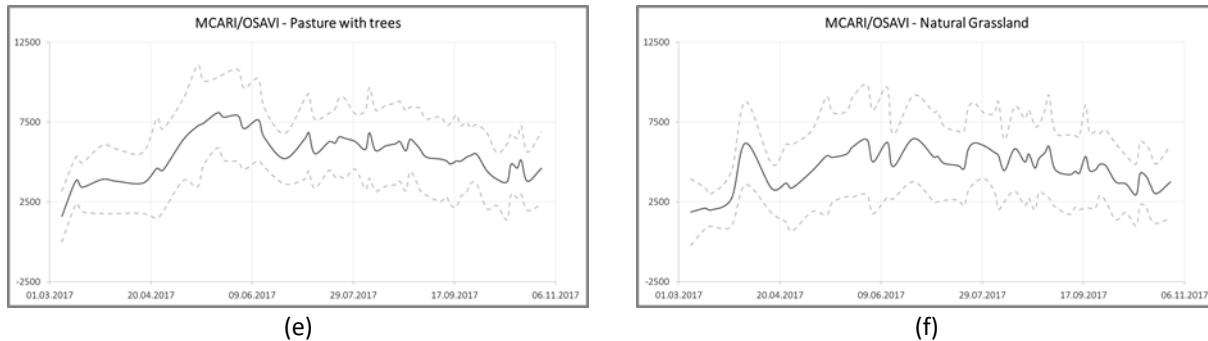


Figure 5-45: Overview of the MEAN spectral signature of the index MCARI/OSAVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

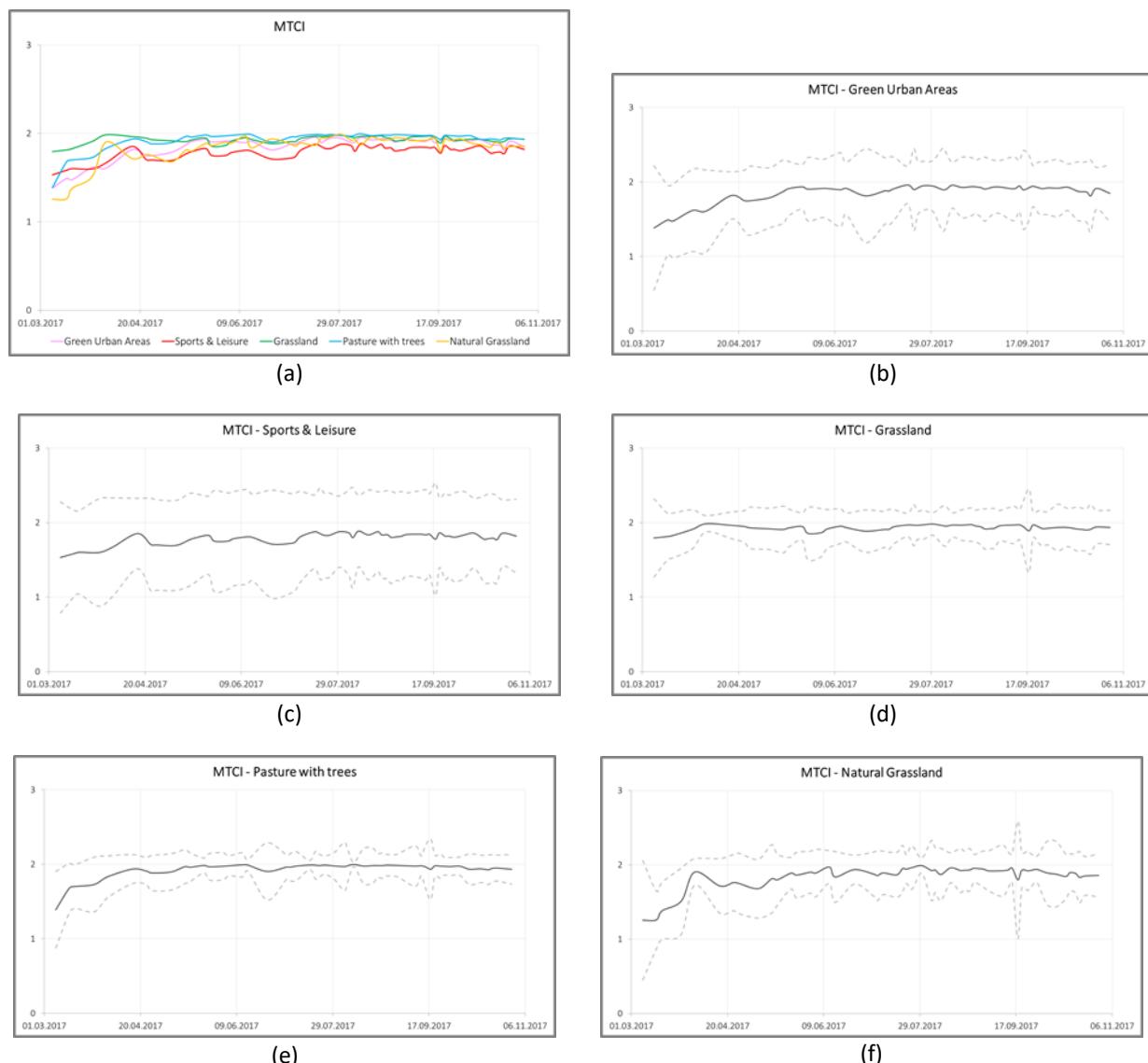


Figure 5-46: Overview of the MEAN spectral signature of the index MTCI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

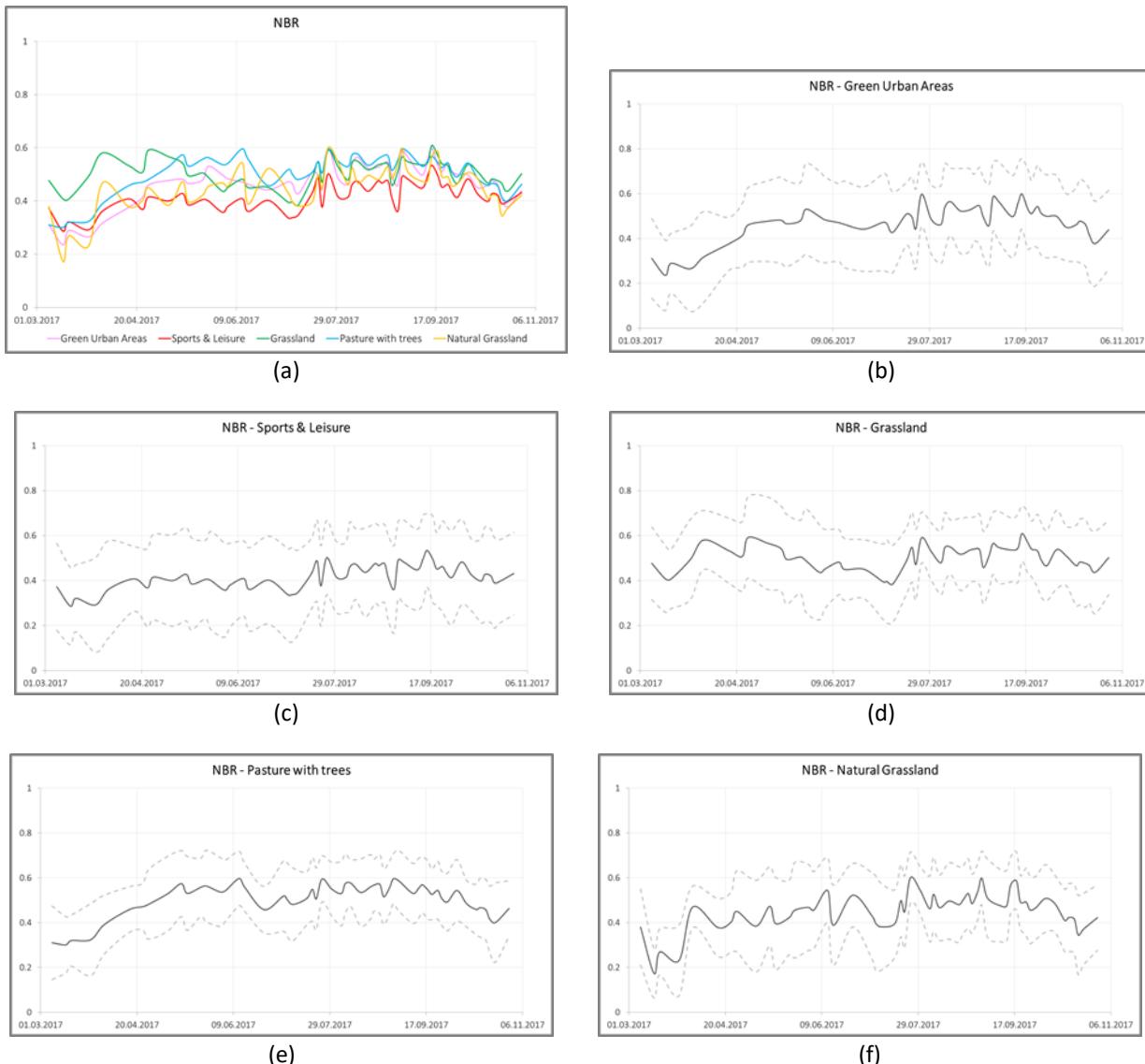
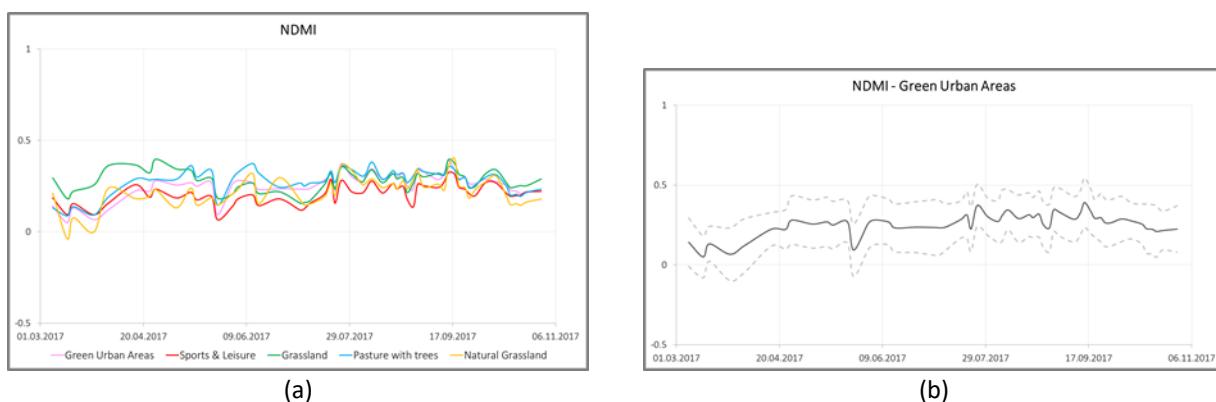


Figure 5-47: Overview of the MEAN spectral signature of the index NBR for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.



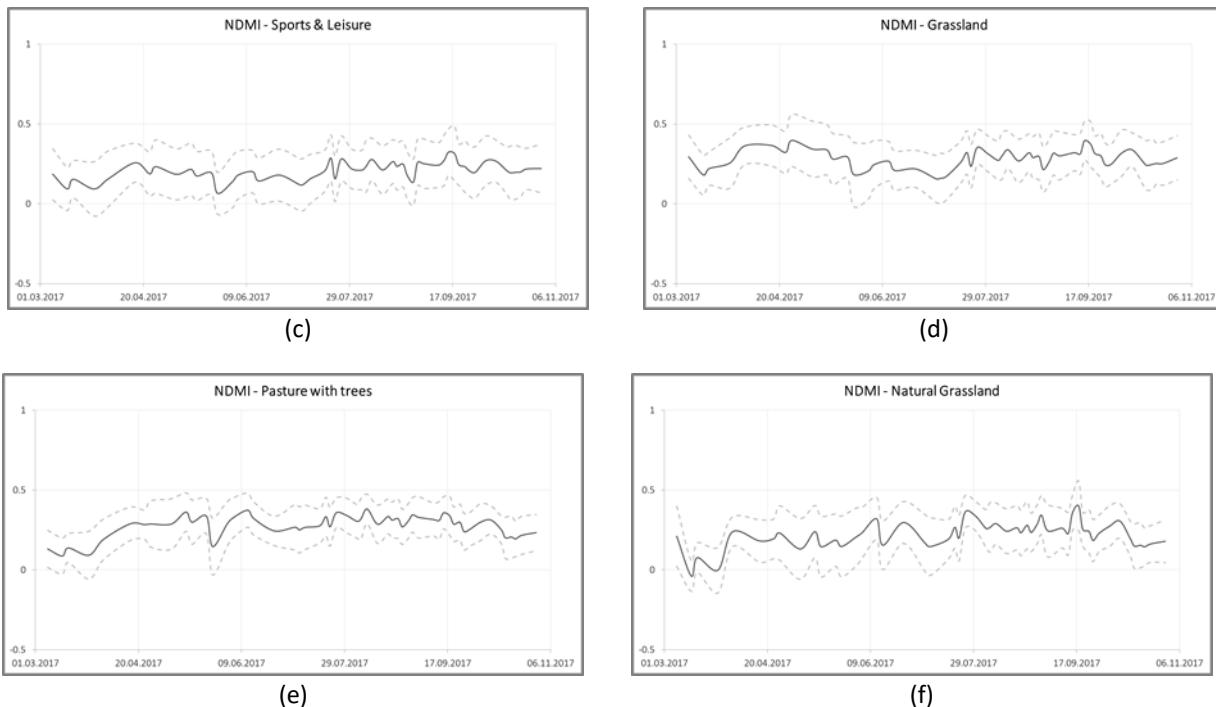
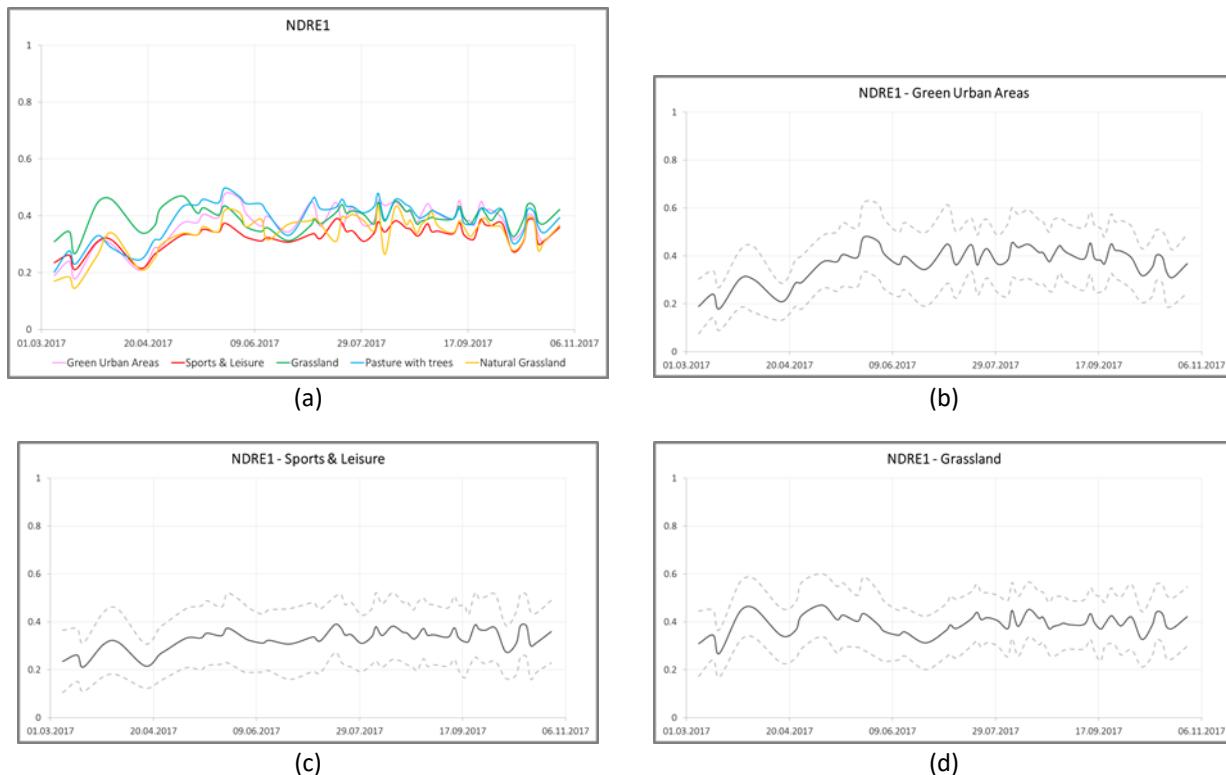


Figure 5-48: Overview of the MEAN spectral signature of the index NDMI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.



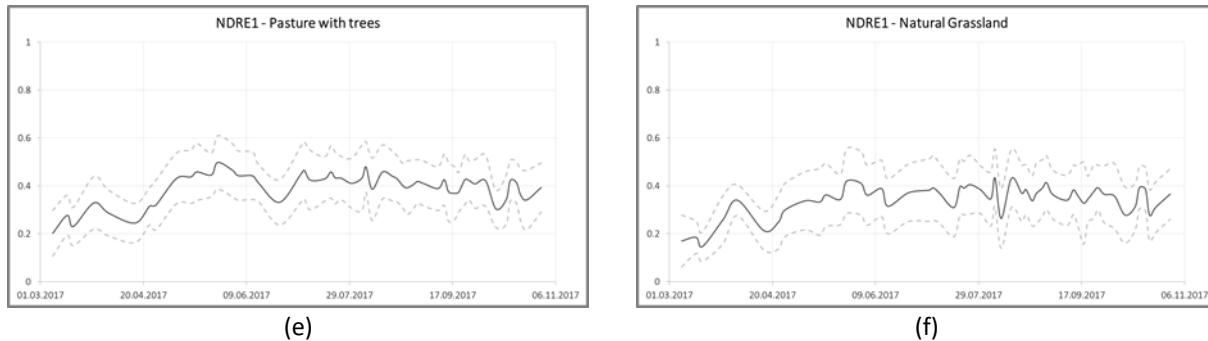


Figure 5-49: Overview of the MEAN spectral signature of the index NDRE1 for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

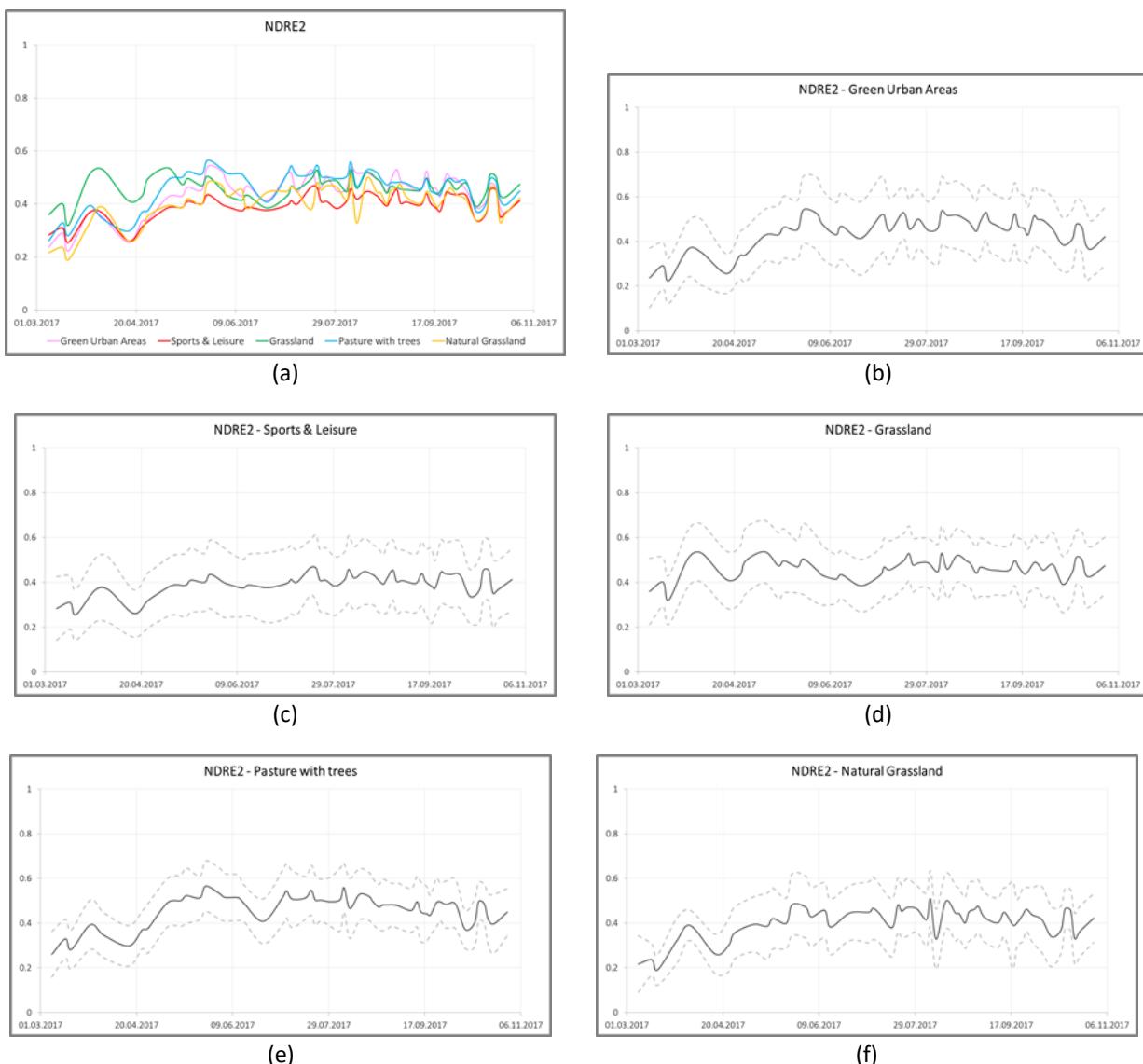


Figure 5-50: Overview of the MEAN spectral signature of the index NDRE2 for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

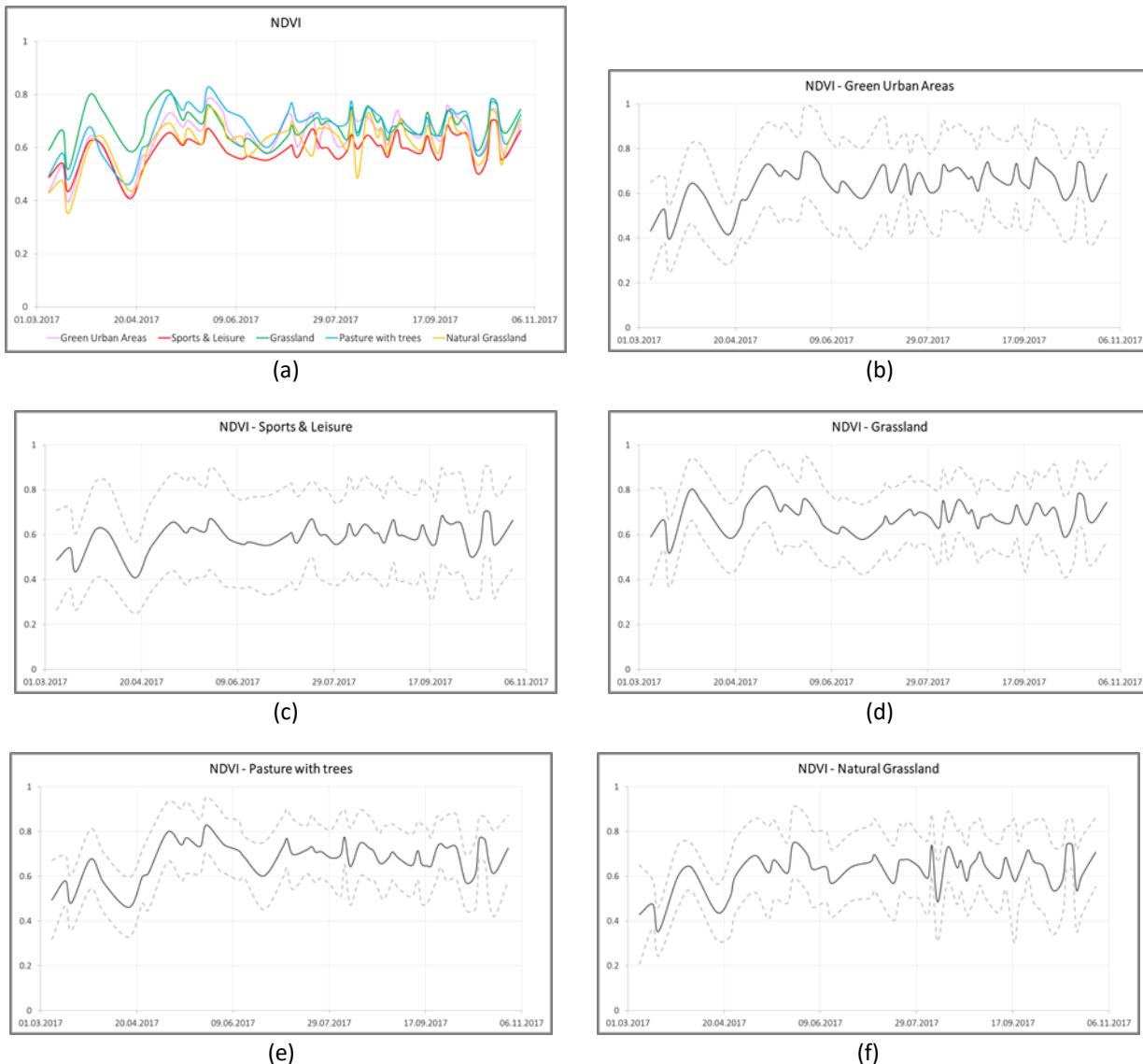
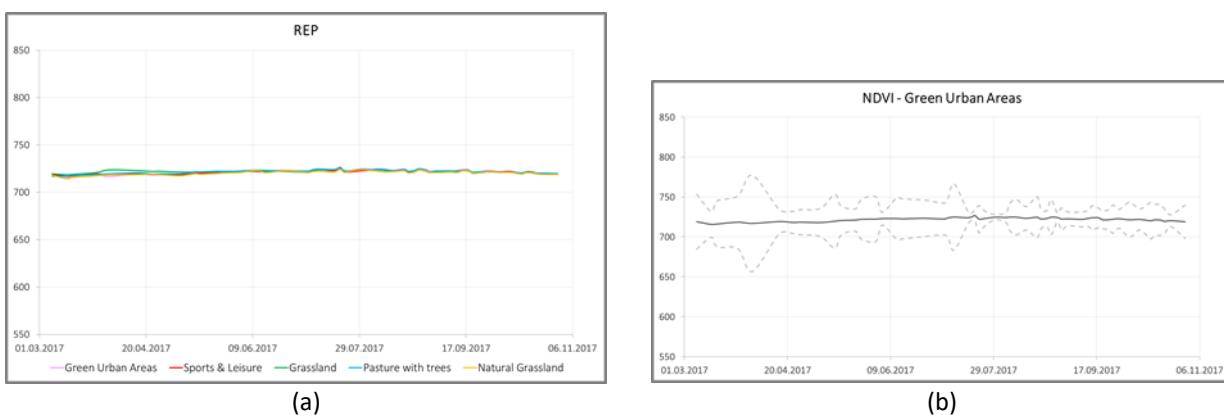


Figure 5-51: Overview of the MEAN spectral signature of the index NDVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.



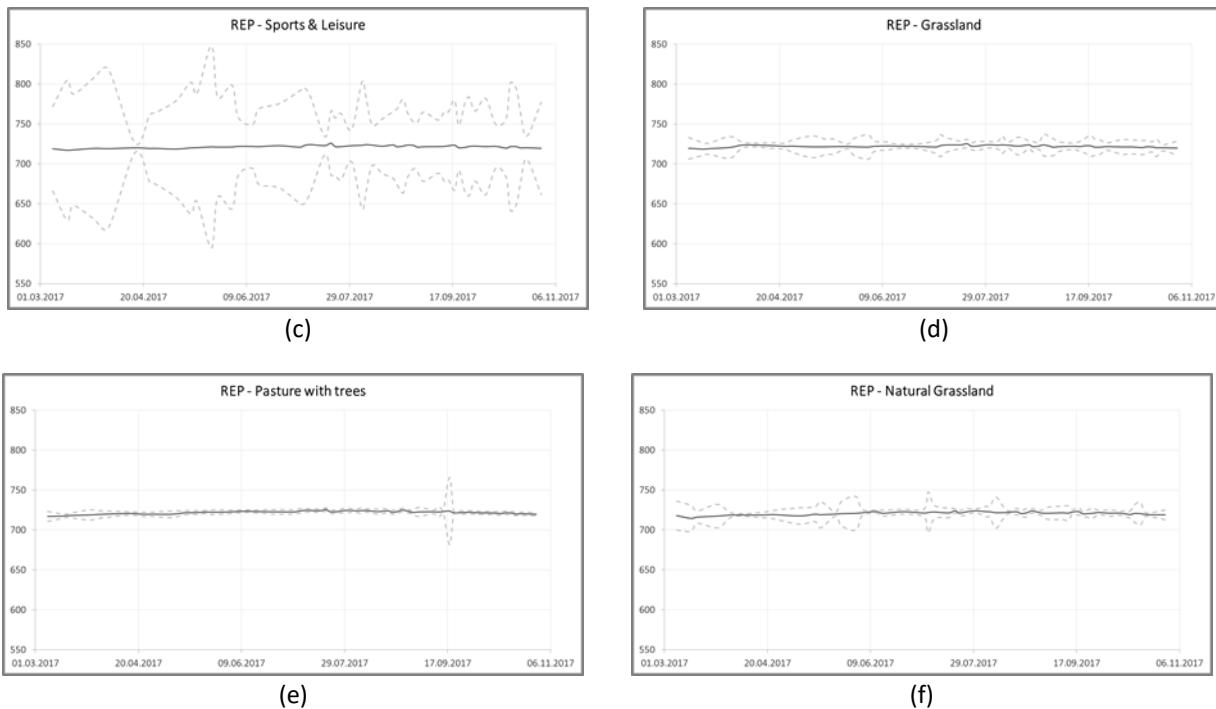
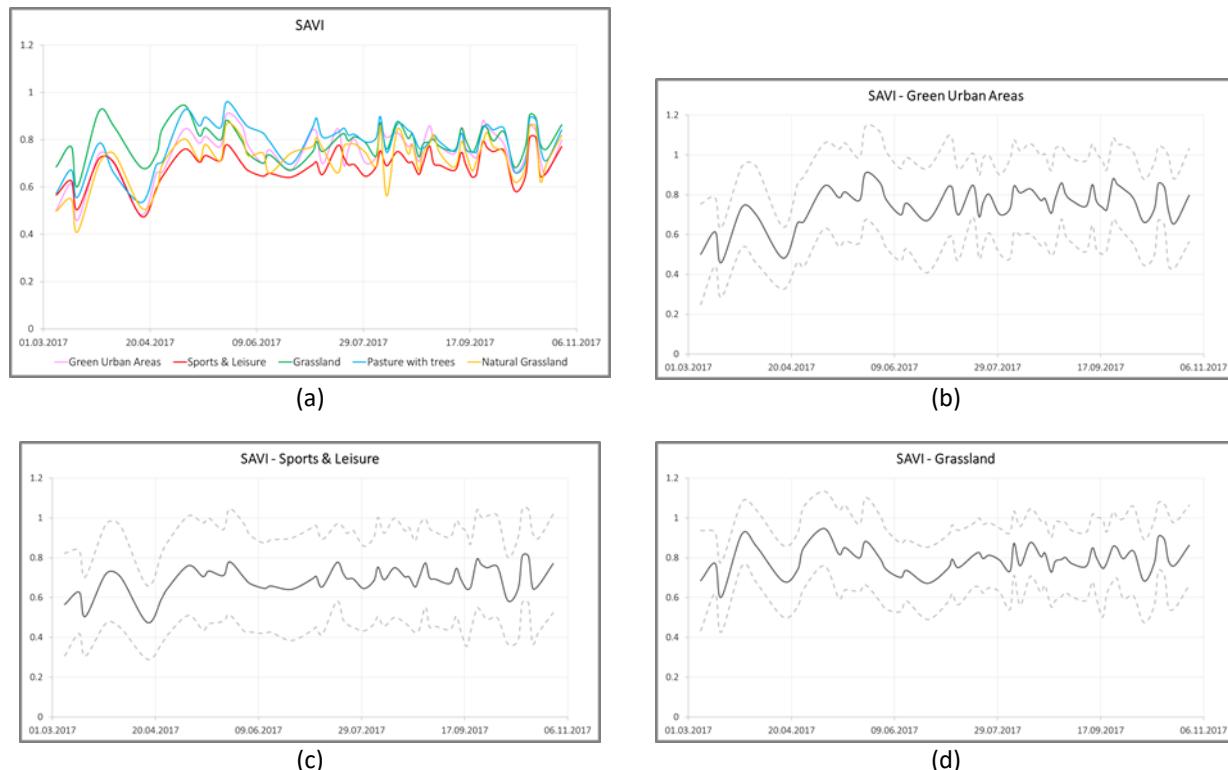


Figure 5-52: Overview of the MEAN spectral signature of the index REP for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.



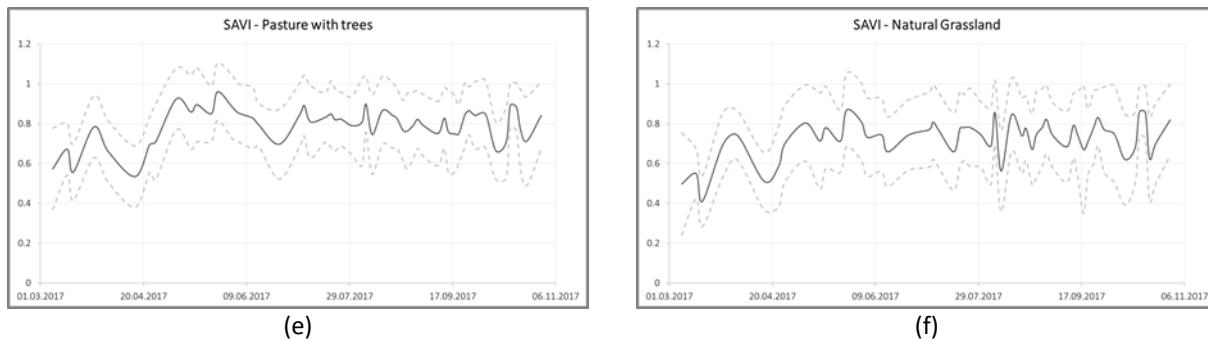


Figure 5-53: Overview of the MEAN spectral signature of the index SAVI for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

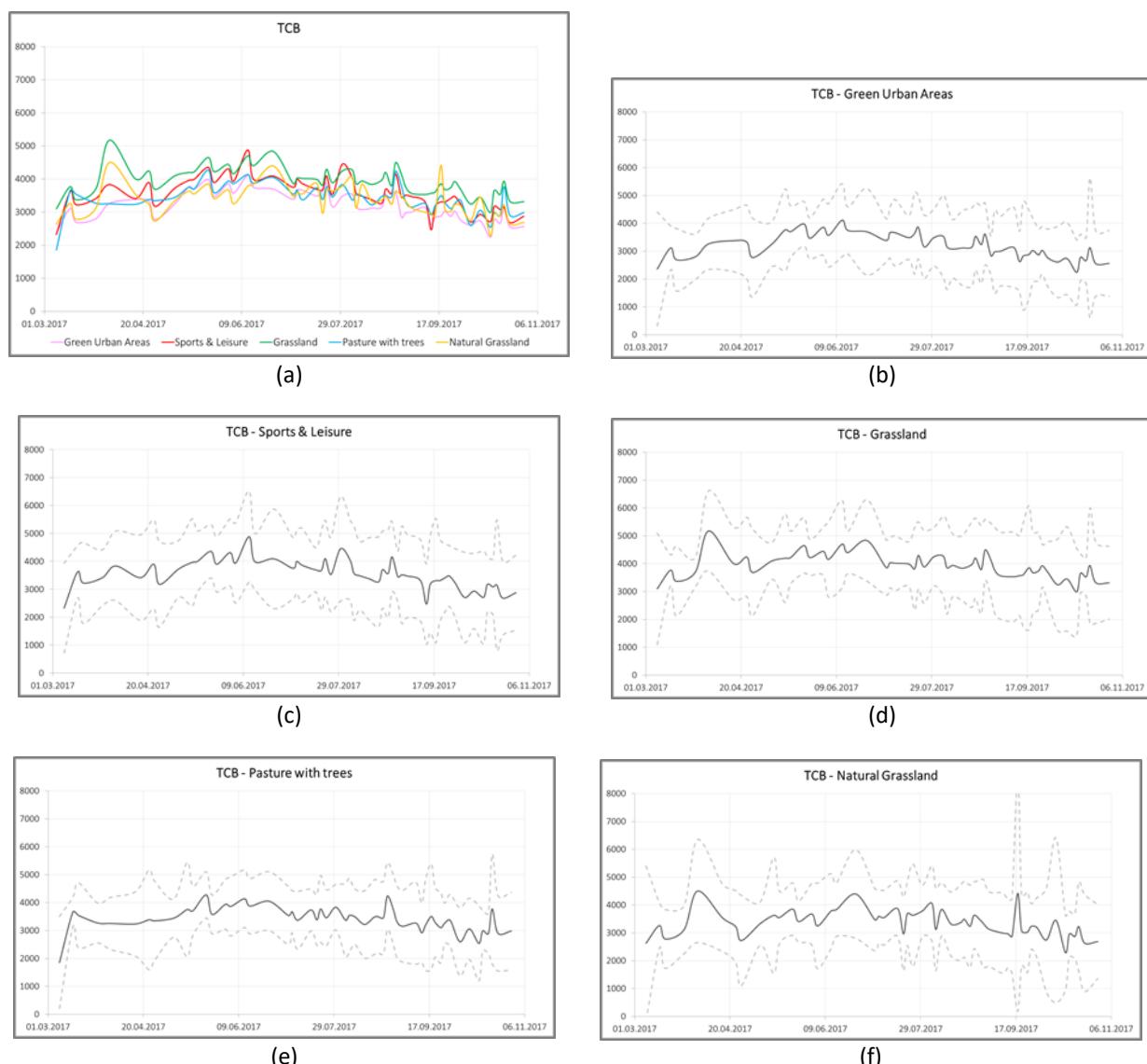


Figure 5-54: Overview of the MEAN spectral signature of the index TCB for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

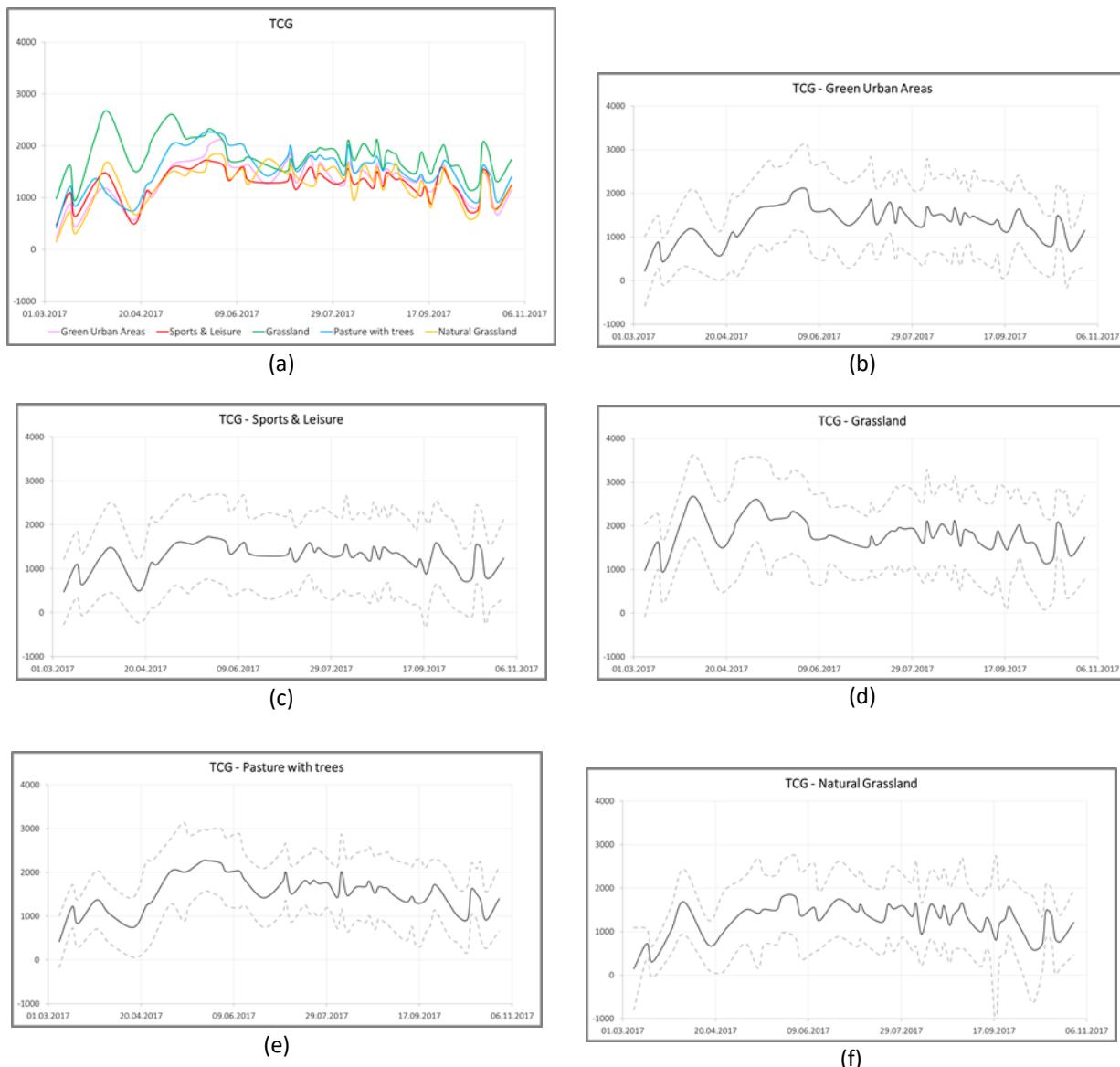
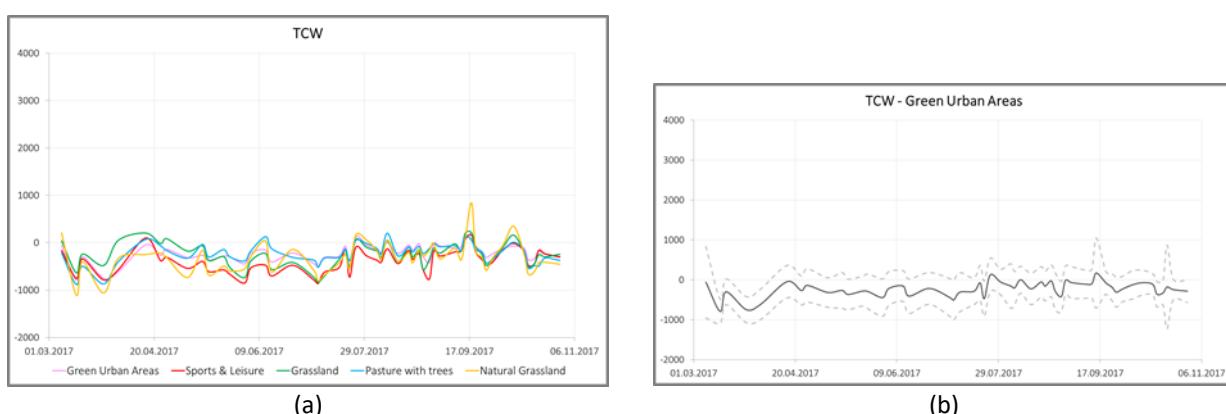
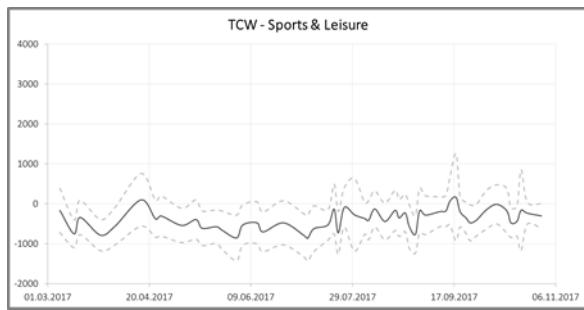
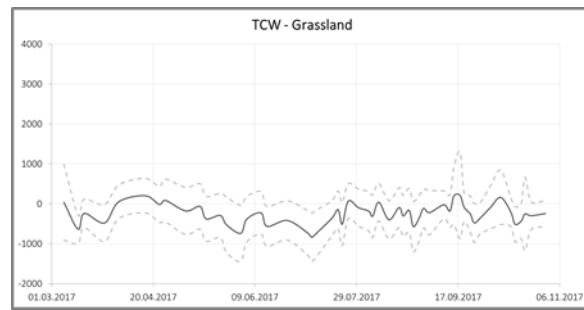


Figure 5-55: Overview of the MEAN spectral signature of the index TCG for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.

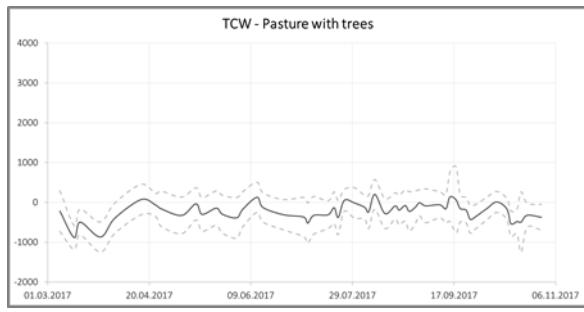




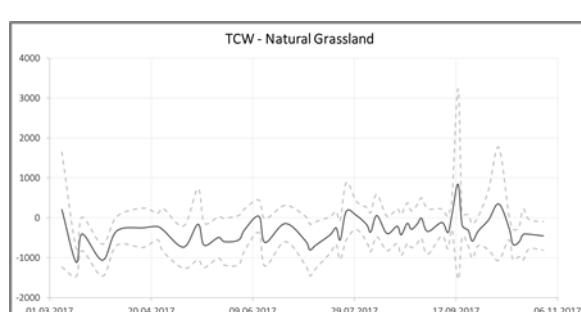
(c)



(d)



(e)



(f)

Figure 5-56: Overview of the MEAN spectral signature of the index TCW for all five grassland types (a) as well as the MEAN +/- the standard deviation for the single five grassland types: (b) green urban areas, (c) sports and leisure areas, (d) pastures with trees, (e) grassland (agriculturally used), and (f) natural grasslands.