

**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



### D9.2

## “D34.1b - Methods Compendium: Time Series Analysis for Change Detection”

Issue/Rev.: 2.0

Date Issued: 04.12.2019

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](mailto:copernicus@gaf.de) – Internet: [www.gaf.de](http://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 and Partners. 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) Benjamin Leutner (DLR) David Herrmann (GAF) Pierre Defourny (UCL) Alexandre Pennec (SIRS)	21.09.2019	
Review:	Petra Miletich (JR) Klaus Granica (JR) Benjamin Leutner (DLR) Sophie Villerot (SIRS)	28.11.2019	
Approval:	Eva Sevillano (GAF)	04.12.2019	
Acceptance:	Massimo Ciscato (REA)		
Distribution:	Public		

### DISSEMINATION LEVEL

DISSEMINATION LEVEL		
PU	Public	<b>X</b>
CO	Confidential: only for members of the consortium (including the Commission Services)	

### DOCUMENT STATUS SHEET

ISSUE/REV	DATE	PAGE(S)	DESCRIPTION / CHANGES
1.0	29.03.2018	70	First issue of the Methods Compendium: Time Analysis for Change Detection
2.0	04.12.2019	110	Second issue of the Methods Compendium: Time Analysis for Change Detection. Major changes in all sections with final results from phase 2

### 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	D21.1b - Service Evolution Requirements Report, Issue 2.0, Issued: December 2019
AD06	D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification, Issue 2.0, Issued: December 2019
AD07	D32.1b- Methods Compendium: Time Series Preparation, Issue 2.0, Issued: 15.05.2019
AD08	D35.1b- Methods Compendium: HRL Time Series Consistency for HRL Product Updates, Issue 2.0, Issued: December 2019

## EXECUTIVE SUMMARY

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

Within the Copernicus Land Monitoring Services, methodological approaches for change detection and tracking are increasingly being demanded to update existing LC/LU maps and to shorten the timespan between image acquisition and final LC/LU products (e.g. for updating of CORINE LC, Urban Atlas, HRLs, for identification of permanent grasslands, for crop status monitoring, monitoring Natura2000 sites and further applications).

Estimating change from remotely sensed data is not a straightforward approach since time series contain a combination of seasonal, gradual and abrupt changes in addition to noise which originates e.g. from remnant geometric misregistration or atmospheric effects. For monitoring land cover changes on pan-European level, the availability of optical EO time series is a bottleneck because of data gaps and non-optimal acquisition time frames due to frequent cloud and snow cover conditions and/or low solar incidence angles, especially in the North of Europe and in Alpine areas. Therefore, for pan-European or global applications, a main requirement is the usage of optical as well as SAR time series to allow a homogeneous wall-to-wall coverage for change monitoring. Depending on the biogeographic region, approaches which are only based on optical data streams (e.g. in Mediterranean areas), and approaches which combine optical and SAR data streams (e.g. in areas with more frequent cloud coverage) can be applied. Therefore, optical and SAR based approaches are benchmarked separately in WP34. In order to achieve these goals ECoLaSS makes full use of dense time series from High-Resolution (HR) Sentinel-2 optical and/or Sentinel-1 Synthetic Aperture Radar (SAR) data.

Rapidly evolving scientific developments as well as user requirements were 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. In Chapter one, the current state of the art is therefore described in detail.

As the land cover dynamics heavily depend on the specific land cover category, there is no single change detection approach, which can be recommended for application in the thematic fields, which are covered within ECoLaSS. For example, the temporal characteristics of the reflectance trajectory are different for forest, cropland, grassland, water and settlement. Therefore, the testing of the change detection approach and its benchmarking is described detailly in thematically separated chapters.

A major requirement for future monitoring of land-cover changes at the pan-European or global level is the reduction of processing time between image acquisition and availability of finally derived information products. Therefore, the focus within WP34 is put on fully automated change detection approaches, which are based on dense time series of Sentinel-2 and/or dense time series of Sentinel-1. A major focus is put on the development of transferrable methods in order to achieve operational feasibility for large scale applications and achieving cost-efficiency in terms of automated approaches. The related tests have proved to be successful, providing satisfactory accuracies and the methodologies have been finally selected for generation of the incremental updates in Task 4.

## TABLE OF CONTENT

<b>1</b>	<b>STATE-OF-THE-ART</b>	<b>1</b>
1.1	TIME SERIES METHODS	1
1.1.1	<i>Signal anomaly detection based on optical Time Series</i>	1
	Image-to-image change detection:	3
	Time series based signature anomaly detection	3
1.1.2	<i>Signal anomaly detection based on SAR Time Series</i>	9
1.1.3	<i>Change detection by combining multi-temporal classifications and multi-temporal anomalies</i>	11
1.2	FOREST CHANGE/LOSS	12
1.2.1	<i>Forest loss</i>	12
1.2.2	<i>Forest regrowth</i>	16
1.3	IMPERVIOUSNESS CHANGE	18
1.4	GRASSLAND CHANGE	20
1.5	AGRICULTURAL DYNAMICS	21
<b>2</b>	<b>TESTING AND BENCHMARKING OF CANDIDATE METHODS</b>	<b>24</b>
2.1	FOREST CHANGE	24
2.1.1	<i>Description of candidate methods</i>	24
2.1.2	<i>Benchmarking criteria</i>	29
2.1.3	<i>Implementation and results of benchmarking</i>	29
2.1.4	<i>Summary and Conclusions</i>	42
2.2	IMPERVIOUSNESS CHANGE	43
2.2.1	<i>Description of candidate methods</i>	47
2.2.2	<i>Benchmarking criteria</i>	53
2.2.3	<i>Implementation and results of benchmarking</i>	55
2.2.4	<i>Summary and conclusions</i>	59
2.3	GRASSLAND CHANGE	60
2.3.1	<i>Description of candidate methods</i>	61
	Signal anomaly detection based on dense time series	61
	Map-to-Map comparison	66
	Map-to-Map comparison combined with signature anomaly detection	66
2.3.2	<i>Benchmarking criteria</i>	66
2.3.3	<i>Implementation and results of benchmarking</i>	67
	Signal anomaly detection based on dense time series	67
	Map-to-Map Comparison	77
	Combination of signal anomalies with Map-to-Map comparison	79
2.3.4	<i>Summary and conclusions</i>	85
2.4	AGRICULTURE CHANGE	86
2.4.1	<i>VIs and hue time series as candidate data sources</i>	87
2.4.2	<i>Candidate detection methods</i>	88
2.4.3	<i>Benchmarking criteria</i>	91
2.4.4	<i>Implementation and benchmarking results</i>	92
<b>3</b>	<b>CONCLUSIONS AND OUTLOOK</b>	<b>97</b>

## LIST OF FIGURES

FIGURE 1-1: COMPONENTS OF TIME SERIES .....	4
FIGURE 1-2: VARIABLES COMMONLY DERIVED IN THE CONTEXT OF REMOTE SENSING BASED TIME SERIES ANALYSIS .....	4
FIGURE 1-3: SCHEMATIC ILLUSTRATION OF DIFFERENT TIME SERIES ANALYSIS METHODS .....	5
FIGURE 1-4: GENERAL PRINCIPLES OF THE OBJECT-BASED ANOMALIES DETECTION FOR CHANGE DETECTION METHOD .....	12
FIGURE 1-5: EXAMPLE OF THE NDWI TRAJECTORY FOR A FOREST LOSS PIXEL .....	15
FIGURE 1-6: TYPICAL NDVI RESPONSE FOLLOWING A DISTURBANCE.....	17
FIGURE 2-1: PHASE 1 MAP-TO-MAP CHANGE DETECTION WITH HRL2015 TCM ( $T_0$ ) AND NEWLY CLASSIFIED 2017 TCM ( $T_1$ ).....	25
FIGURE 2-2: PHASE2 MAP-TO-MAP CHANGE DETECTION WITH TCM 2017( $T_0$ ) AND TCM 2018 ( $T_1$ ) AND NDVI PLAUSIBILITY STEP .....	26
FIGURE 2-3: WORKFLOW OF THE KALMAN FILTERING PROCEDURE INCLUDING TUNING PARAMETERS .....	29
FIGURE 2-4: 2017 CLOUD COVERAGE OF S-2 TILE VVF (LEFT) AND VVF TILE (RIGHT). BLUE: SCENES WITH < 50% CLOUD COVER.....	30
FIGURE 2-5: SENTINEL-2 DATA SCORE (NUMBER OF CLOUD-FREE OBSERVATIONS) FOR THE ECOLASS TEST SITE NORTH .....	30
FIGURE 2-6: SENTINEL-2 DATA SCORE LAYER 2017 (LEFT) AND 2018 (RIGHT) FOR THE CENTRAL DEMONSTRATION SITE .....	32
FIGURE 2-7: CHANGE DETECTION FILTER RESULTS.....	33
FIGURE 2-8: CHANGE POLYGONS IN AN AREA CLOSE TO NAISLACH (CENTRAL TEST SITE). .....	35
FIGURE 2-9: COMPARISON OF THE CHANGE AREA VALIDATION ACCURACY FOR ALL INPUT DATA SCENARIOS .....	37
FIGURE 2-10: DETAILED VIEW OF DETECTED NEGATIVE FOREST CHANGE PATCHES OF DIFFERENT DETECTION PRECISION FOR SENTINEL-1 (S1), SENTINEL-2 (S2) AND COMBINED SENTINEL-1 AND SENTINEL-2 (S1/S2) .....	38
FIGURE 2-11: EXAMPLE OF DETECTED FOREST LOSS IN TILE 32TPT– TEST SITE AUSTRIA/GERMANY. ....	41
FIGURE 2-12: OVERALL WORKFLOW OF THE IMPERVIOUSNESS IMPLEMENTATION .....	44
FIGURE 2-13: THE REFERENCE LAYER ( $T_0$ ): THE HRL IMPERVIOUSNESS 2015 BUILT-UP MASK.....	45
FIGURE 2-14: NEW STATUS LAYER ( $T_N$ ): THE HRL IMPERVIOUSNESS 2017 BUILT-UP MASK.....	45
FIGURE 2-15: NEW STATUS LAYER ( $T_N$ ): THE HRL IMPERVIOUSNESS 2018 BUILT-UP MASK.....	46
FIGURE 2-16: NEW HRL IMPERVIOUSNESS 2017 BUILT-UP MASK POST-PROCESSED WITH SPATIAL FILTERING .....	47
FIGURE 2-17: CALCULATION OF THE BUILT-UP PROBABILITY MAP FROM THE 2015 BUILT-UP MASK FOR PHASE 1 SW TEST SITE .....	48
FIGURE 2-18: CALCULATION OF THE BUILT-UP PROBABILITY MAP FROM THE 2015 BUILT-UP MASK FOR THE PHASE 2 TEST SITES .....	49
FIGURE 2-19: NEW HRL IMPERVIOUSNESS 2017 BUILT-UP MASK POST-PROCESSED WITH THE CONTEXTUAL ANALYSIS .....	49
FIGURE 2-20: NEW HRL IMPERVIOUSNESS 2018 BUILT-UP MASK POST-PROCESSED WITH THE CONTEXTUAL ANALYSIS .....	51
FIGURE 2-21: FINAL CHANGE LAYER FOR THE TEMPORAL LAPSE 2015-2017 .....	51
FIGURE 2-22: FINAL CHANGE LAYER FOR THE TEMPORAL LAPSE 2015-2018 Ce AND SE TEST SITES AND 2017-2018 SW TEST SITE .....	52
FIGURE 2-23: REFERENCE CALIBRATION DATASET OVERLAID ON THE CHANGE LAYER .....	54
FIGURE 2-24: 2018 REFERENCE CALIBRATION DATASETS OVERLAID ON THE 2015-2018 CHANGE LAYER FOR THE TEST SITES.....	55
FIGURE 2-25: EXAMPLE OF CALIBRATION POINT FOR THE NEWLY DETECTED BUILT-UP IN 2017 .....	56
FIGURE 2-26: EXAMPLE OF OMISSION ERRORS 2015 (UNDETECTED BUILT-UP IN 2015) .....	57
FIGURE 2-27: EXAMPLE OF COMMISSION ERRORS FOR 2017, FALSE BUILT-UP IN 2017 .....	57
FIGURE 2-28: EXAMPLE OF CALIBRATION POINT FOR THE NEWLY DETECTED BUILT-UP IN 2018.....	58
FIGURE 2-29: EXAMPLE OF OMISSION ERRORS 2015 (UNDETECTED BUILT-UP IN 2015) .....	59
FIGURE 2-30: EXAMPLE OF COMMISSION ERRORS FOR 2018, FALSE BUILT-UP IN 2018 .....	59
FIGURE 2-31: OVERVIEW FOR THE GRASSLAND CHANGE MONITORING WORKFLOW .....	61
FIGURE 2-32: MAD COMPONENTS .....	62
FIGURE 2-33: CVA COMPONENTS.....	63
FIGURE 2-34: COST FUNCTIONS .....	65
FIGURE 2-35: VALIDATION SAMPLES OVER THE DEMONSTRATION SITE WEST .....	67
FIGURE 2-36: NUMBER OF UNMASKED OBSERVATIONS PER PIXEL, YEAR 2015 .....	68
FIGURE 2-37: NUMBER OF UNMASKED OBSERVATIONS PER PIXEL, YEAR 2016 .....	68
FIGURE 2-38: TIME SERIES OF A GRASSLAND PIXEL LOCATED IN THE CENTRE OF THE RED RECTANGLE (TRUE COLOUR) .....	70
FIGURE 2-39: TIME SERIES OF A GRASSLAND PIXEL LOCATED IN THE CENTRE OF THE RED RECTANGLE (FALSE COLOUR) .....	71
FIGURE 2-40: TIME SERIES OF SPECTRAL BANDS .....	72
FIGURE 2-41: TIME SERIES OF SPECTRAL INDICES .....	73
FIGURE 2-42: EXAMPLE OF A TASSELED CAP BRIGHTNESS TIME SERIES OF A GRASSLAND PIXEL AND FITTED REGRESSION MODELS USING OLS AND IRLS.....	74
FIGURE 2-43: COMPUTED OBSERVATION WEIGHTS OF THE IRLS FIT .....	74
FIGURE 2-44: EXAMPLE OF GREENNESS TIME SERIES OF A GRASSLAND PIXEL AND FITTED REGRESSION MODELS USING OLS .....	75
FIGURE 2-45: COMPARISON OF AN AGRICULTURAL PIXEL TO A GRASSLAND PIXEL .....	75
FIGURE 2-46: GREENNESS TREND AND AMPLITUDES OF AN AGRICULTURE AND GRASSLAND PIXEL.....	76
FIGURE 2-47: GREENNESS TIME SERIES OF A GRASSLAND PIXEL IN CONSECUTIVE YEARS .....	76
FIGURE 2-48: GREENNESS TREND AND AMPLITUDES OF A GRASSLAND PIXEL IN CONSECUTIVE YEARS .....	77



## LIST OF TABLES

TABLE 1-1: STUDIES USING BFAST OR BFAST MONITOR WITH LANDSAT TIME SERIES .....	6
TABLE 1-2: STUDIES USING CCDC WITH LANDSAT TIME SERIES .....	7
TABLE 2-1: SAMPLE DISTRIBUTION OF TRAINING DATASET FOR THE TEST SITE NORTH .....	31
TABLE 2-2: COMPOSITION OF THE SAMPLE BASE USED FOR ASSESSING THE MAP ACCURACIES. ....	33
TABLE 2-3: INTERNAL VALIDATION RESULTS FOR THE KALMAN FILTER APPROACH .....	34
TABLE 2-4: ACCURACY ASSESSMENT OF NEGATIVE FOREST CHANGE DERIVED FROM SENTINEL-1 TIME FEATURES .....	36
TABLE 2-5: ACCURACY ASSESSMENT OF NEGATIVE FOREST CHANGE DERIVED FROM SENTINEL-2 TIME FEATURES .....	36
TABLE 2-6: ACCURACY ASSESSMENT OF NEGATIVE FOREST CHANGE DERIVED FROM COMBINED SENTINEL-1/-2 TIME FEATURES .....	37
TABLE 2-7: PERCENTAGE OF CHANGE AREA COMPARED TO THE 2015 FOREST AREA .....	38
TABLE 2-8: ACCURACY ASSESSMENT OF NEGATIVE FOREST CHANGE DERIVED FROM COMBINED SENTINEL-1/-2 TIME FEATURES WITHOUT NDVI PLAUSIBILITY OF CHANGES – TEST SITE NORTH .....	39
TABLE 2-9: ACCURACY ASSESSMENT OF NEGATIVE FOREST CHANGE DERIVED FROM COMBINED SENTINEL-1/-2 TIME FEATURES WITHOUT NDVI PLAUSIBILITY OF CHANGES – TEST SITE CENTRAL .....	39
TABLE 2-10: ACCURACY ASSESSMENT OF NEGATIVE FOREST CHANGE DERIVED FROM COMBINED SENTINEL-1/-2 TIME FEATURES WITHOUT NDVI PLAUSIBILITY OF CHANGES – TEST SITE SOUTH-EAST .....	40
TABLE 2-11: BENCHMARKING CRITERIA, CHANCES, AND ISSUES OF THE DIFFERENT INPUT DATA SCENARIOS .....	42
TABLE 2-12: CLASSIFICATION RESULTS FOR THE REFERENCE YEAR 2017 .....	56
TABLE 2-13: CLASSIFICATION RESULTS FOR THE REFERENCE YEAR 2017 .....	58
TABLE 2-14: ANOMALOUS IMAGES.....	69
TABLE 2-15: MOWING EVENTS.....	69
TABLE 2-16: MAP-TO-MAP CHANGES WITH DIFFERENT PROBABILITY THRESHOLDS .....	77
TABLE 2-17: MAD ANOMALIES COMBINED WITH RF PROBABILITIES ON POLYGON LEVEL.....	82
TABLE 2-18: CVA ANOMALIES (2 SIGMA) COMBINED WITH RF PROBABILITIES ON PIXEL LEVEL .....	82
TABLE 2-19: MAP-TO-MAP APPROACH .....	83
TABLE 2-20: METHODS WITHOUT PARAMETERIZATION. ....	93
TABLE 2-21: METHODS WITH PARAMETERIZATION. ....	93
TABLE 2-22: COMPARISON OF THE CALIBRATION AND MONTHLY SAMPLES FOR ALL METHODS FOR MSAVI .....	95

## ABBREVIATIONS

AGDD	Accumulated Growing Degree Days
BAP	Best-Available-Pixel
BBCH	Biologische Bundesanstalt, Bundessortenamt Und Chemische Industrie
BFAST	Breaks For Additive Seasons And Trend
BRDF	Bidirectional Reflectance Distribution Function
C2C	Composite2Change
CBERS	China–Brazil Earth Resources Satellite
CCA	Cross-Correlation Analysis
CCDC	Continuous Change Detection And Classification
CLMS	Copernicus Land Monitoring Services
CMFDA	Continuous Monitoring Of Forest Disturbance Algorithm
CORINE	Coordination Of Information On The Environment
CVA	Change Vector Analysis
DLT	Dominant Leaf Type
DT	Decision Tree
ECoLaSS	Evolution Of Copernicus Land Services Based On Sentinel Data
EEA	European Environment Agency
EEE	Entrusted European Entities
EO	Earth Observation
ESA	European Space Agency
ETM+	Enhanced Thematic Mapper Plus
EU	European Union
FAO	Food And Agriculture Organization
FAPAR	Fraction Of Absorbed Photosynthetically Active Radiation
FCOVER	Fractional Cover
FEWS	Famine Early Warning Systems
Fmask	Function Of Mask
FRA	Global Forest Resources Assessment
GEOGLAM	Group On Earth Observations Global Agricultural Monitoring
GDD	Growing Degree Days
GIEWS	Global Information And Early Warning System
GRD	Ground Range Detected
ha	Hectare
HJ-1	Huan Jing-1
HLS	Harmonized Landsat-Sentinel-2
HR	High Resolution
HRL	High Resolution Layer
ID	Identifier
IFZ	Integrated Forest Z-Score
IRECI	Inverted Red-Edge Chlorophyll Index
IRLS	Iteratively Reweighted Least Squares
IRS	Indian Remote-Sensing Satellite
ISA	Impervious Surface Area
JRC	Joint Research Centre
KI	Kittler-Illingworth
KT	Kauth-Thomas
LAI	Leaf Area Index
LandTrendr	Landsat-Based Detection Of Trends In Disturbance And Recovery
LC	Land Cover
LCMAP	Land Change Monitoring, Assessment, And Projection
LIDAR	Light Detection And Ranging

LPIS	Land Parcel Identification Systems
LSR	Least Squares Regression
LSMM	Linear Spectral Mixture Model
LTS	Landsat Time Series
LU	Land Use
LUCAS	Land Use/Cover Area Frame Statistical Survey
MAD	Multivariate Alteration Detector Components
MARS	Merrick Advanced Remote Sensing
MCP	multiscale change profiles
MetOp	Meteorological Operational Satellite
MLC	Maximum Likelihood Classification
MMU	Minimum Mapping Unit
MODIS	Moderate Resolution Imaging Spectroradiometer
MSS	Multispectral Scanner System
NASA	National Aeronautics And Space Administration
NBR	Normalized Burn Ratio
NLCD	National Land Cover Database
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Vegetation Index
NIR	Near-Infrared
OA	Overall Accuracy
OLS	Ordinary Least Squares
OPT	Optical
OPT/SAR	Optical And SAR Data Combined
PA	Producer Accuracy
PCC	Post-Classification Comparison
PROBA-V	Project For On-Board Autonomy – Vegetation
PSP	Phenological Sequence Patterns
PUG	Unpalatable Grasses
RF	Random Forest
RGB	Red-Green-Blue
RI	Recovery Indicator
RRI	Relative Recovery Indicator
SAR	Synthetic Aperture Radar
SAVI	A Soil-Adjusted Vegetation Index
SCD	Spectral Change Difference
Sen2Cor	Sentinel2 Correction
SMA	Spectral Mixture Analysis
sNDVI	spatially normalized NDVI
SPOT	Satellite Pour l'Observation De La Terre/Satellite For Observation Of Earth
SV	Stochastic Volatility
SVM	Support Vector Machine
SWIR	Short Wavelength Infrared
TCT	Tasseled Cap Transformation
TM	Thematic Mapper
TWDTW	Time Weighted Dynamic Time Wrapping
UA	User Accuracy
USA	United States Of America
USTC	Uncertainty-Based Spatio-Temporal Consistency
VCT	Vegetation Change Tracker
VH	Vertical Transmit/Horizontal Receive (Polarization)
VHR	Very High Resolution

VV Vertical Transmit/Vertical Receive (Polarization)  
VVF Valutazione Della Vocazionalità Faunistica  
VWF Variable Window Filters



Zhu (2017) reviewed change detection studies based on Landsat time series. Although the number of studies is rather high, the author observed that “*most of the time series studies were only interested in producing annual or biannual change maps*” and therefore “*tended to select multiple images (partly cloudy images) acquired in the same season and the same year to produce cloud-free composite images as their inputs. In this way, seasonal differences caused by solar angle differences and vegetation phenological changes were minimized, and the data volume was reduced substantially*” (Zhu, 2017). The author also noted that more recent studies use a higher amount of images per year acquired in different seasons.

The conclusions from both reviews as well as an independently conducted literature search show that there are few main algorithms for change detection that have been applied to satellite image time series with a high temporal resolution. Many older, established methods like the vegetation change tracker (VCT) proposed by (Huang et al., 2010) or LandTrendr (Kennedy et al., 2010) circumvent the problems outlined in the previous sections by using annual time series of cloud-free composites.

The optimal approach to apply depends on the time series density and on the specific thematic or geometric requirements for the respective application field such as for forest monitoring, monitoring changes in grasslands, croplands etc.

#### Seasonality:

Seasonal patterns in remotely sensed time series are primarily caused by the annual variations of temperature and rainfall that influence plant phenology. Effects of the Bidirectional Reflectance Distribution Function (BRDF) can also add to the seasonality (Zhu et al., 2012). An analysis of different seasonal patterns for a number of land-cover/land-use classes based on MODIS NDVI time series is given by Geerken (2009). Fourier analysis was applied to the data in order to find the dominant frequencies in the annual cycle. The results indicate that most of the seasonal variation of many land cover types can be modelled by using 3 harmonics corresponding to a frequency of 1, 2, and 3 periods per year, without overfitting the noise. Dynamic land cover types like cropland show larger amplitudes at the higher frequencies whereas forest types have the highest amplitude at a frequency of 1 period per year. Hence, time series with increased density and well distributed observations covering all seasons are required to monitor more dynamic land cover types. An algorithm for change detection has to be able to distinguish between normal seasonal changes and abnormal behaviour.

#### Unequal temporal interval between observations:

Satellites with a regular nadir acquisition scheme usually have a constant revisit cycle, for example, 16 days for Landsat-8 and 5 days with Sentinel-2 (two satellites). However, with increasing latitude adjacent paths overlap and the density of the time series in these areas can be increased, at the cost of irregularly spaced observations. The revisit time could also be decreased by integrating data from Landsat and Sentinel-2 (Wulder et al., 2016). Because both satellites have different repeat cycles, irregularly spaced observations can be expected. However, many methods used in time series analysis require a constant sampling interval, in other words, a constant interval between the observations. In order to use these methods, an additional pre-processing step like filtering or interpolation becomes necessary. A review and comparison of some existing methods is given by (Kandasamy et al., 2013). This pre-processing may introduce other undesired effects like overly smoothing the data or introducing another set of tuning parameters into the workflow. At spatial resolutions of Sentinel or higher, the additional computational effort might also be significant. Depending on the application, the use of more sophisticated time series analysis methods which can handle irregularly spaced observations can be preferable.

#### Missing Observations:

Clouds and cloud shadows, as well as snow, greatly influence the reflectance of different spectral bands (Dozier, 1989; Irish et al., 2006). They can be considered as noise in the data and must be masked out, resulting in missing observations in the time series. Not all methods used in the time series analysis are designed to account for missing values. Analogous to the reasons pointed out in the previous Chapter, a

decision between additional pre-processing to fill the gaps or more powerful time series analysis methods capable of handling missing observations must be made. Another aspect of the problem is concerned with the amount of cloud and snow cover that has to be expected for certain areas on Earth. In the tropical climate zone, for example, the cloud cover probability is very high (Wylie and Menzel, 1999) and it is, therefore, difficult to obtain dense time series.

Presence of outliers:

Although there are some powerful algorithms for cloud, cloud shadow and snow screening, none of them can fully fulfil the requirements for a highly accurate product. As a consequence, the presence of outliers in the time series must be considered as a given fact. Thus, a reliable change detection method has to incorporate a mechanism to avoid confusion between outliers and real changes.

This chapter presents a review of signal anomaly detection methods based on per-pixel time series analysis, which have been published in the recent years. The signal anomaly detection is based on optical Sentinel-2 data streams resulting from WP32 [AD07]. Therefore, the chapter focuses especially on algorithms capable of processing dense time series at spatial resolutions of optical sensors like the Landsat or Sentinel-2 series. In this context, time series are considered dense if several valid observations are available within one year. Before discussing some of the existing methods, a summary of the challenges, which a change detection algorithm has to address, is given.

The state-of-the-art signal anomaly detection methods can be divided into two categories. The first one is called image-to-image change detection methods and the second one is based on time series analysis (Hirschmugl et al., 2017). The availability of high frequency observation data increased over the years and with both Sentinel-2 satellites in orbit dense time series with high resolution data are available. Therefore, this chapter focuses on time-series analysis.

### **Image-to-image change detection:**

At least one image acquired before and one after a change event is required, of which the first must be from the start of a monitoring period and the second in the end. This approach is often also referred to as ‘bi temporal change detection’ [...]. Such an approach can be also applied two or more times, each time comparing two images or classifications. One major drawback of this method is the influence of cloud and cloud shadows which leads to remaining data gaps. Another drawback is the sensitivity to geometric and radiometric inconsistencies caused by short term changes in reflectance or backscattering characteristics. Because of these but also other drawbacks, the focus in EcoLaSS is on developing methods based on dense time series of satellite images, which is described in more detail in the next Chapter.

### **Time series based signature anomaly detection**

This approach requires a series of images taken continuously over a period of time. Thus, there is a need for substantially more and regular image acquisitions over the area of interest. This approach is often referred to simply as ‘time series analysis’. (Hirschmugl et al., 2017)

Time series analysis based change detection can also be done feature-to-feature based on creating temporal features over a certain time period and comparing the features with each other. Furthermore, classification results based on temporal features can be compared with each other to classify changes.

The time series based change detection methods highly depend on dense time series. Typically there are three components (see Figure 1-1), which can be exploited depending on the change detection focus. The first component is the long-term directional trend observed over the years, which represent the stable conditions with a long-term decrease or increase in the evaluated features. The second component is the seasonal component, which represents the spectral change due to climatic conditions, for example, dry seasons. The last component is the residual component, which represents short term

changes due to short term occurrences through local weather events or other short term disturbances like fires (Kuenzer et al., 2015).

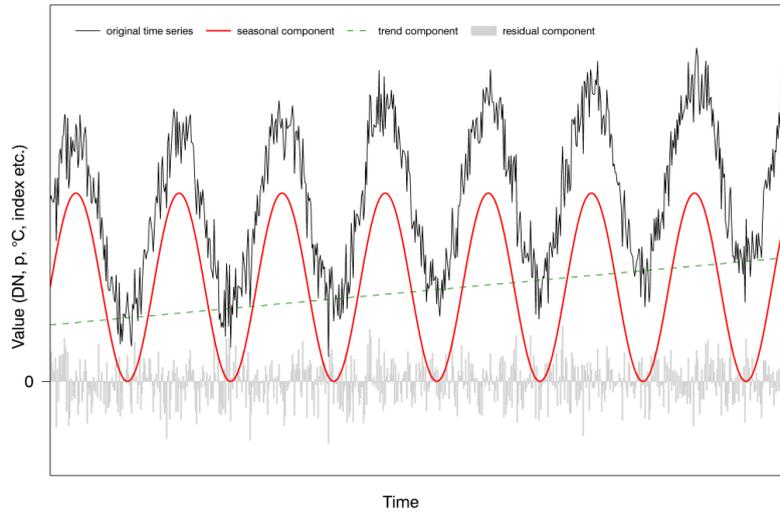


Figure 1-1: Components of time series (from Kuenzer et al., 2015)

Kuenzer et al. (2015) reviewed time series data, approaches, opportunities and challenges based on optical and radar data sources. Different features can be derived from several data sources like geophysical variables, index variables, thematic variables, topographic variables and texture variables. Those variables are described in Figure 1-2 (Kuenzer et al., 2015).

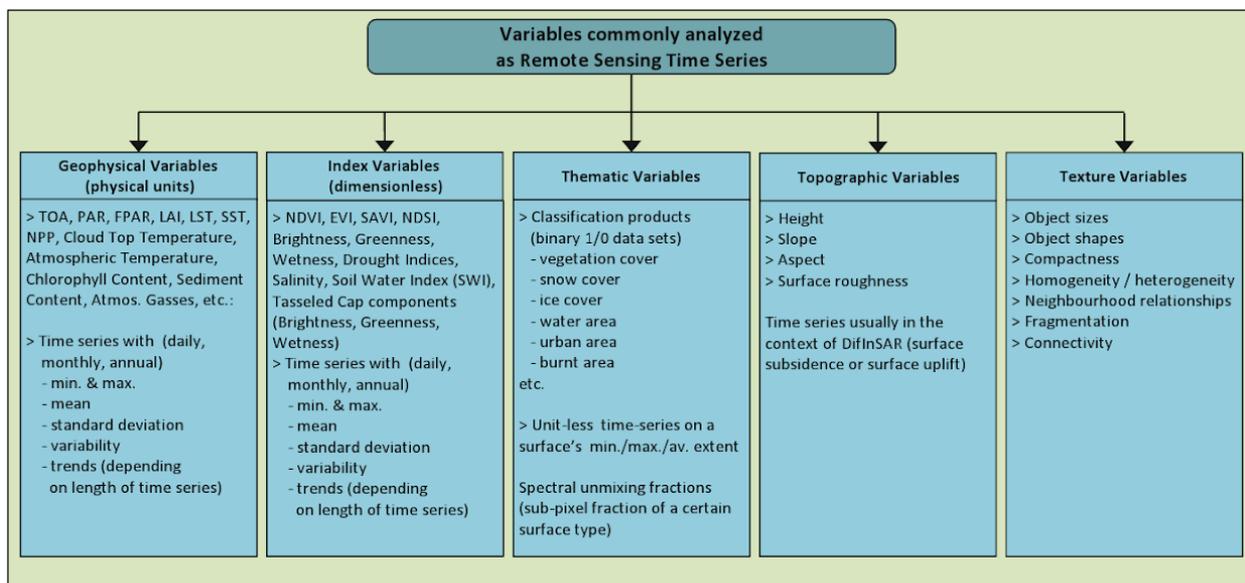
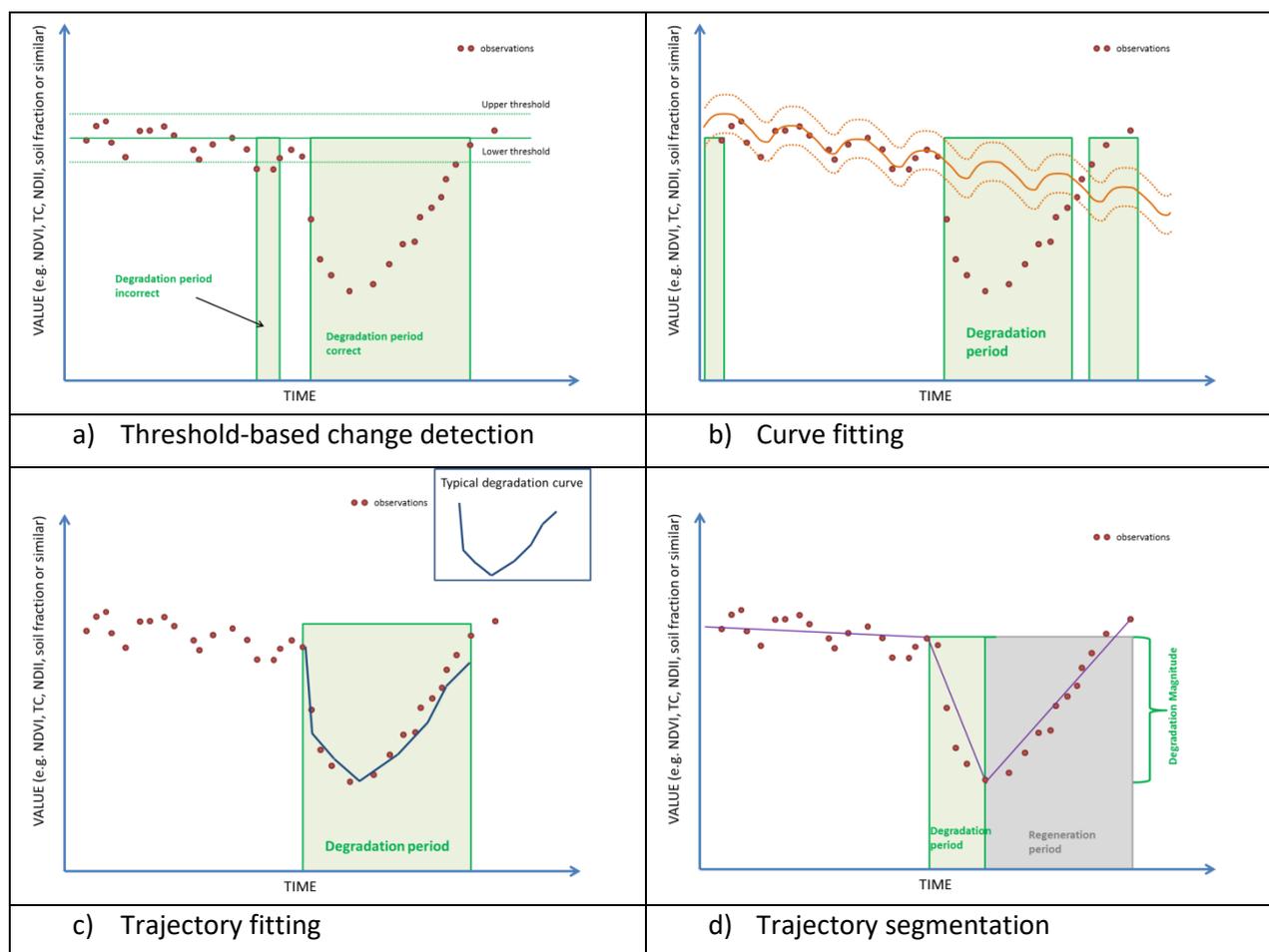


Figure 1-2: Variables commonly derived in the context of remote sensing based time series analysis (from Kuenzer et al., 2015)

Banskota et al. (2014) defines four categories for trajectory-based time series analysis methods:

- threshold based change detection
- curve fitting
- trajectory fitting
- trajectory segmentation

Figure 1-3 provides a schematic comparison of the four mentioned time series analysis methods.



**Figure 1-3: Schematic illustration of different time series analysis methods (from Hirschmugl et al., 2017)**

a) Threshold-based signature anomaly detection

The threshold based change detection methods use empirically defined thresholds to separate deviating observations from stable trend observations within the time series analysis. The main drawback of threshold-based methods is that the empirically derived thresholds are only suitable for local areas cannot be transferred to other areas, which are characterized by different vegetation types (Banskota et al., 2014). Therefore, they are in general not suited for a pan-European or global approach.

b) Curve fitting

Often, a pixel-wise trend function based upon least squares regression (LSR) is fitted between a spectral variable and time (Banskota et al., 2014). One drawback is that the underlying statistical assumptions (data normality, equal variance) are generally difficult to meet and violation of these assumptions results in an inadequate representation of the data by the fitted function (Hirschmugl et al., 2017).

BFAST - the Breaks for Additive Season and Trend algorithm was originally designed to detect the trend and seasonal changes in MODIS 16-day NDVI composites (Verbesselt et al., 2010a). It is based on the decomposition of a time series into a trend, seasonal, and remainder component. Change detection is performed by finding breakpoints in either of the estimated components. Through this approach, the number, timing, and type of changes in historic time series can be determined. A second version employed a different seasonal model based on trigonometric functions which the authors deemed to be “more suitable and robust for phenological change detection with satellite image time series” (Verbesselt et al., 2010b). The next stage in development was BFAST Monitor, “a multi-purpose time-series-based disturbance detection approach that identifies and models stable historical variation to enable change

*detection within newly acquired data”* (Verbesselt et al., 2012). Initially, it was also used with MODIS data, but a variety of more recent studies listed in demonstrate that BFAST and BFAST Monitor can also be applied to Landsat time series to detect both abrupt and gradual change. It has to be noted that all of the studies in Table 1-1 investigate forest applications. Nevertheless, it may be possible to adapt some elements of the BFAST approach to other land cover types.

**Table 1-1: Studies using BFAST or BFAST Monitor with Landsat time series**

Goal and observables	Test sites	Reference
Monitoring of forest cover loss, NDVI, combination with MODIS or rainfall data	Bolivia, tropical forest	(Dutrieux et al., 2015)
Forest monitoring in regions with persistent cloud cover, fusion of Landsat NDVI and SAR	Fiji, tropical forest	(Reiche et al., 2015)
Tracking of disturbance-regrowth dynamics using all available Landsat data, NDMI	Southern Peru, tropical forest	(DeVries et al., 2015a)
Monitoring of small-scale forest disturbances, NDVI	Southern Ethiopia	(DeVries et al., 2015b)
Deforestation mapping, sNDVI (spatially normalized NDVI, reduced seasonality by using spatial context)	Humid/dry forest, Brazil/Bolivia	(Hamunyela et al., 2016)
Mapping of deforestation and degradation, various spectral bands and indices, focus on correct change classification	Southern Ethiopia	(DeVries et al., 2016)
Evaluation of how specific effects of site and radiometric correction affect the accuracy of deforestation monitoring when using BFAST Monitor	Brazil, Ethiopia, Vietnam	(Schultz et al., 2016)
Forest disturbance detection and change agent attribution (windthrow, cleared windthrow, bark beetles, and other harvest), tasseled cap wetness	Bohemian forest, Kalkalpen, Tatra	(Oeser et al., 2017)
Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs	Multiple examples of arable lands at large scale	(Atzberger C., 2013)
Analysis of Agricultural Land Use Changes in Jombang Regency, East Java, Indonesia Using BFAST Method	Jombang agricultural lands, East Java	(Permatasari, P., A, et al, 2016)

Another tool is TIMESAT (Eklundh and Jönsson, 2015), which provides several options for separating components of changes and smoothing the time series with a variety of applications. In the TIMESAT tool, a number of processing steps for transforming noisy signals into smooth trajectories are included, such as e.g. the Savitzky-Golay filter.

CCDC - Originally, Zhu et al. (2012) developed the so-called Continuous Monitoring of Forest Disturbance Algorithm (CMFDA). *“Using all the available Landsat ETM+ images in two years, models using sines and cosines are fit for each pixel and each spectral band. These models can predict Landsat images at any date assuming there is not any land cover change.”* Change detection is performed by differencing the predicted and a newly acquired image. If the calculated difference of a change index sensitive to forest disturbance crosses a certain threshold for 3 consecutive times, a pixel is flagged as changed. The method also incorporates two-stage cloud screening where previously unmasked clouds, shadows, and snow can be detected by using multi-temporal data. This approach was later refined and published separately as *“a new algorithm ... for automated masking of cloud, cloud shadow, and snow for multi-temporal Landsat images”* (Zhu and Woodcock, 2014a).

The change detection part of CMFDA has been further developed to the Continuous Change Detection and Classification (CCDC) algorithm where the concept is extended to include more types of land cover beside forest as well as a classification framework (Zhu and Woodcock, 2014b). From the beginning, CCDC was designed to work with dense Landsat time series and can, therefore, handle seasonality, missing or irregularly spaced observations, and outliers to some extent. Both abrupt and gradual changes can be detected. Some further updates to the algorithm are described in Zhu et al. (2015). They include a mechanism to automatically adjust the complexity of the time series model based on the number of available clear observations as well as a different method to estimate the model parameters that reduce overfitting. Some recently published studies employing the CCDC algorithm are listed in Table 1-2.

**Table 1-2: Studies using CCDC with Landsat time series**

Topic	Test sites	Reference
Investigation of the possibilities for monitoring gradual changes using dense Landsat time series	Several locations, USA	(Vogelmann et al., 2016)
Analysis of vegetation greenness trends considering effects due to land cover change	Guangzhou, China	(Zhu et al., 2016)
Analysis of urbanization induced land use and land cover change	Atlanta metropolitan area, USA	(Fu and Weng, 2016)
Evaluation of CCDC for use within the USGS Land Change Monitoring, Assessment, and Projection program (LCMAP)	Several locations, USA	(Pengra et al., 2016)
Mapping of forest degradation	Lam Dong Province, Vietnam	(Vogelmann et al., 2017)
Near-real-time monitoring of insect defoliation	Southern New England, USA	(Pasquarella et al., 2017)

KALMAN filtering - denotes a versatile parameter estimation technique which yields optimal estimates in a statistical sense. This technique is well established in many fields and has been applied to numerous signal processing problems (Gibbs, 2011). Harvey (1989) provided a comprehensive treatment of the Kalman filters’ application to time series analysis in econometrics. However, examples of its usage in the remote sensing context are rare. Kleynhans et al. (2011) have analyzed NDVI time series derived from MODIS images in order to map settlement growth in northern South Africa. They concluded that their Kalman filter based algorithm achieved higher accuracies compared to a “traditional annual NDVI differencing method”. Another study has been conducted by Olsson et al. (2016) in a test area located in northern Sweden, where Kalman-filtered NDVI time series have been used to detect insect induced defoliation in near real-time. MODIS 8-day composites covering the years 2000-2013 have been used as input data.

In general, the Kalman filter approach requires the definition of a dynamic model and an observation model. The observation model, on the one hand, defines the relationship of the measurements to a set of state variables which cannot be observed directly. For a time series, the state variables usually represent the series’ additive decomposition into trend, seasonal, and long-term cyclical components. The dynamic model on the other hand describes the expected temporal evolution of the state variables. By formulating the dynamic model in continuous time, the problem of gaps in the time series due to persistent cloud cover can be addressed. The Kalman filter operates recursively from one point in time to the next. Each recursion may be divided into two steps. In the time-update step, the states’ temporal evolution is predicted based on the dynamic model. It is followed by the measurement-update step, where the predicted state estimate is enhanced by incorporating newly available observations. Thus, the Kalman filter framework has some interesting properties regarding its application to remote sensing time series. Since the state variables are continuously updated, the underlying model can adapt to variations in the seasonal cycle, which may be present due to different climatic conditions compared

to the previous year. This is a key advantage over simple regression models. Moreover, the ability to make model-based forecasts including confidence intervals is an inherent part of the technique.

c) Trajectory fitting

Phenomena like land cover changes have a distinct temporal progression which can be abrupt or on long terms. Those characteristics can be represented with spectra-temporal signatures. Following trend assumptions are made to define different change types. The gradual long-term changes will be represented by a linear trend, whereas abrupt short term changes are represented by a sharp decrease or increase in the trajectory. Areas without changes show ideally a steady horizontal trend. The trajectory fitting method adopts these assumptions by fitting curve characteristics, with defined change trajectories. This supervised change detection method shows some limitations by the requirement of predefined trajectories representing all possible change types (Banskota et al., 2014).

d) Trajectory segmentation

The trajectory segmentation method is applied e.g. within the LandTrendR (Kennedy et al., 2010) approach to characterize distinct sub-trends within the trajectory. In this approach, the trajectory is decomposed into a series of straight-line segments, which capture broad features of the trajectory as well as sub-trends (Kennedy et al., 2010). *“The first phase of segmentation is the determination of the vertex years that define the end points of segments. In the second phase the best straight-line trajectory is fitted through those vertices using either point-to-point or regression lines. The result of this segmentation is a spectral trajectory which is composed of straight-line segments. The time position and spectral value of vertices of the segments provide the essential information, which is needed to produce maps.”* (Hirschmugl et al., 2017)

According to (Banskota et al., 2014), *“the advantage of this approach is that the straight-line segments allow the detection of abrupt events, such as disturbances, as well as longer-duration processes, such as regrowth”*. Another advantage is that the data determines the shape of the trajectory and therefore no typical curve of degradation is required as in the trajectory fitting method. One major drawback of the trajectory segmentation is that seasonal effects caused by phenology are not taken into account (Banskota et al., 2014).

### LIMITATIONS

All algorithms share certain basic concepts, but the individual implementation varies. They are designed to process large amounts of data in a highly automated way and, therefore, rely on data-driven statistical boundaries for detecting change, although the distinct nature and computation of these boundaries is quite different. Parts of the algorithms handle seasonality through harmonic models based on sines and cosines, but the methods used to fit them to the data differ. Regarding the limitations of time-series-analysis-based change detection algorithms, the influence of several interdependent parameters has to be considered as decisive:

(1) *Properties of the input data.* In order to detect a certain type of change, it has to be represented in the input data and spectral bands or indices have to be chosen accordingly. Cohen et al. (2017) compared forest disturbance maps produced from the output of seven different change detection algorithms. Among them were CCDC and the method proposed by Brooks et al. (2014), the remaining five methods only consider annual time series. One of the conclusions by Cohen et al. (2017) was that *“Spectral change magnitudes associated with forest disturbance are highly variable, with a population likely to be skewed towards lower-magnitude occurrences. Such disturbances are challenging to map because they are often difficult to distinguish from spectral noise common in temporal trajectories of spectral signals.”*



minimum error algorithm (Kittler and Illingworth, 1986), Otsu's algorithm (Otsu, 1979) and Huang and Wang's algorithm (Huang and Wang, 1995), whereas the KI technique was the most commonly used for unsupervised change detection.

Furthermore, Carincotte et al. (2006) applied an unsupervised change detection technique based on fuzzy hidden Markov chains. The principle of their approach is a fuzzy segmentation of the log-ratio image identifying pixels belonging to three classes: negative change, positive change and no change. All pixels which belong to two classes are considered as fuzzy class of considered mixture classes (Carincotte et al., 2006). According to Bovolo et al. (2013), change detection approaches based on ratio and log-ratio operators comparing pixel by pixel are not well suited for spatially very high resolution (VHR) images and high temporal resolution time series because of the complexity of VHR images in their electromagnetic backscattering mechanisms and possible differences in acquisition conditions of multi-temporal images (e.g. differences in view angle, moisture). In such cases, the same pixels or objects can show different backscatter behaviours without undergoing actual land cover changes, hence leading to false alarms (Bovolo et al., 2013). To overcome these challenges Bovolo et al. (2013) introduced a two way concept exploiting a multi-scale technique to preliminarily detect possible changes in backscatter and explicitly model the semantic meaning of change by using SAR image properties and prior information. Furthermore, (Gamba et al., 2006) applied a pixel as well as feature based change detection. They performed an extraction and comparison of linear features from multiple SAR images and pixel-based changes. Hachicha and Chaabane (2014) reviewed different methods for similarity measures which allow change detection in SAR images. Also (Inglada and Mercier, 2007) discussed similarity measures for automatic change detection in multi-temporal SAR images by developing local statistics between two time steps. In their approach, the local statistics are estimated by using a cumulant-based series expansion which approximates the probability density functions in the neighbourhood of each pixel (Inglada and Mercier, 2007). The *Kullback-Leibler divergence* locally measures the distance between two probability density functions and in this manner creates multi-scale change profiles (MCPs) which are then used to identify the change (Inglada and Mercier, 2007). Chatelain et al. (2008) discuss the *multivariate gamma distributions* for change detection in different SAR data sources (multi-sensoral). Moreover, wavelet based change detection methods were introduced by Bovolo and Bruzzone (2005), Celik (2010), (Gong et al., 2012), (Bovolo et al., 2013), and (Schmitt and Brisco, 2013). Other change detection approaches are based on extension change vector analysis (CVA) (Shen et al., 2007) or neural networks (Pratola et al., 2013).

The second family of SAR change detection methods aims for multi-channel SAR data e.g. using polarimetric parameters (Dierking and Skriver, 2002; Qong, 2004, Nascimento et al. 2019). Dierking and Skriver (2002) stated that intensity images from multi-temporal polarimetric SAR data provide better results for change detection compared to correlation coefficients and phase differences between co-polarized channels. They used a ratio operator to analyze temporal images and a decision mechanism with respect to the required probability of false alarm (Bazi et al., 2005). Other approaches are based on Markov random fields for multichannel SAR data (Moser et al., 2007; Moser and Serpico, 2009; Yousif and Ban, 2014), the generalized maximum likelihood test for covariance matrices (Conradsen et al., 2003), partial vectors for the suppression of the backscatter coefficient influence (Marino et al., 2013) and the Hotelling-Lawley trace statistics (Akbari et al., 2013).

Inherent to most polarimetric change detection approaches is the reliance on the complex covariance matrix, which becomes more informative with the number of dimensions. Many polarimetric approaches, therefore, rely on quad-polarized SAR data. This limits their applicability to operational land monitoring with Sentinel-1, which, although capable of acquiring quad-polarized data, acquires almost exclusively in dual polarized mode.

In addition to presented studies and methods focusing on SAR data, some promising time series analysis tools, which were developed for optical data, could be considered for their adaptation (see Chapter 1.1.1). Change Detection by Combining Multi-temporal Classifications and Multi-temporal Anomalies.

### 1.1.3 Change detection by combining multi-temporal classifications and multi-temporal anomalies

This Chapter introduces a rather original algorithm to detect forest cover change based on two or multiple satellite images of the same area acquired by the same sensor at different epochs. The principle of this algorithm has been first developed and implemented by Desclée et al. (2006) and then used over large sets of samples as reported by Duveiller et al. (2008). In the framework of the FRA-2010, FAO selected this algorithm for his global forest cover assessment (FAO, 2010). This algorithm was found very efficient to detect change in needleleaf forest, evergreen or tropical semi-deciduous forests. For forest with strong seasonality, the performances also depend on the acquisition date of the images which should be similar with regards to the foliage phenology.

This automated object-based method for change detection is a multi-step iterative process that identifies the objects showing higher reflectance differences between the two images than the other objects for the same land cover. A working assumption easily met is the fact that the majority of the objects are not affected by change between the two or more dates.

The processing chain is run by pairs of images and includes four major steps (Figure 1-4):

- (i) a multi-date multispectral image segmentation to delineate objects,
- (ii) an extraction for each object of the surface reflectance distributions corresponding to each image;
- (iii) a change detection algorithm based on a multivariate iterative trimming to detect the objects showing anomalies based on their reflectance distribution (i.e. mean and standard deviation) difference;
- (iv) mapping the objects detected as change between the two or more images.

The main advantage of this change algorithm is the fact that it does not require a strict consistency of the signal over time but assumes a spatially homogeneous impact of the atmosphere and of any issue which may affect the signal from one date to another. In other words, for the invariant target, the signal from the first to the second date can be changed or biased but this bias is assumed to be constant over the whole region of interest. The past experiences shown that this working assumption hold for an area of 20 x 20km and larger areas has to proceed by moving windows. The size of the window is a trade-off between the homogeneity of the atmospheric effects on the surface reflectance and the necessity to have enough objects which were not changed between the two dates.

Recently, this automatic multi-temporal segmentation has been adapted to Sentinel-2 image pair. Four bands from both old and recent images are used with equal weights: 3 (Red), 4 (NIR), 8 (SWIR) and 12 (SWIR). Band 12 with a spatial resolution of 20m has to be resampled at 10m resolution and for that reason, it could be more appropriate to weight a little less the band 12 for the segmentation. The product of this segmentation is a set of spectrally and spatially coherent objects and common for both images. The initial design was proposed for forest but could probably be extended to another land cover classes which are not too much affected by seasonality.

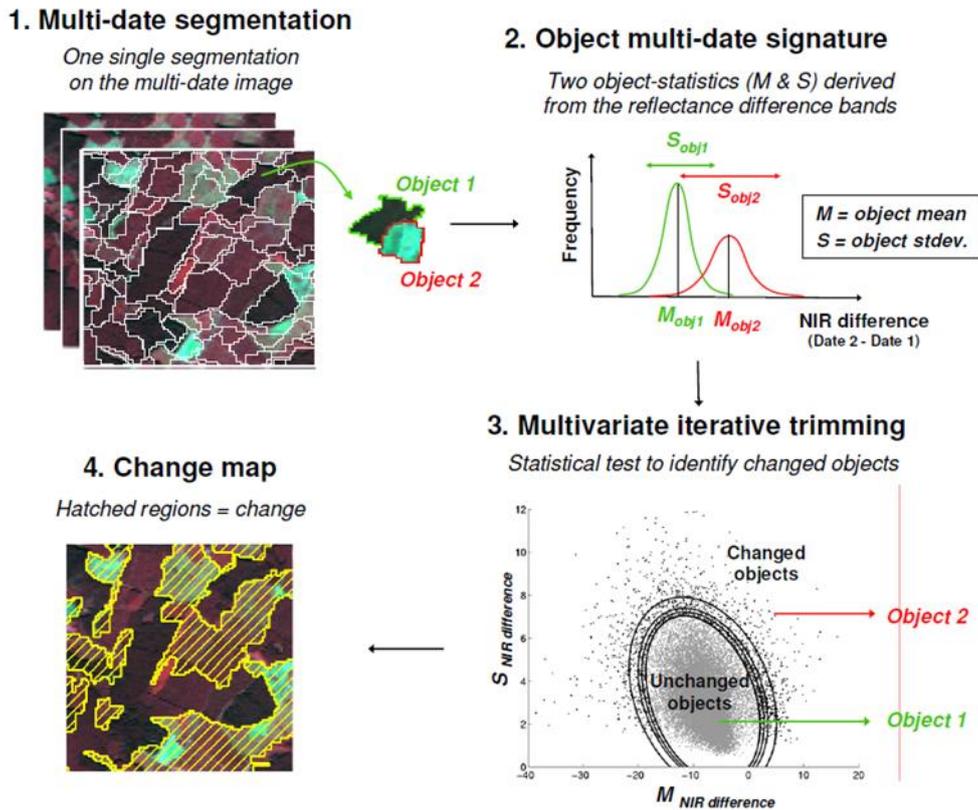


Figure 1-4: General principles of the object-based anomalies detection for change detection method (Desclée et al., 2006)

A differentiated consideration of the thematic land cover is necessary as their spectral-temporal characteristics and change patterns are fundamentally different in time as well as in space. The following sections address change detection methods specifically for Forest, Imperviousness, Grassland and Agriculture.

## 1.2 Forest Change/Loss

Currently used methods on forest change monitoring aim for accurate and timely detection of forest disturbances (Reiche et al., 2018; Hansen et al., 2016). Regarding forest recovery, most of the existing studies focus on post-disturbance regrowth rates and recovery trends using long-period Landsat time series (Bartels et al., 2016; Franks et al., 2013; Frazier et al., 2015; Pickell et al., 2016).

As aforementioned, on a medium to high resolution scale, the currently used methods based on remotely sensed data can be generally divided into two categories: a bi-temporal (image-to-image change detection) and a multi-temporal approach (time-series analysis) (Hirschmugl et al., 2017; Mitchell et al., 2017; Miettinen et al., 2014).

### 1.2.1 Forest loss

For image-to-image change detection two different approaches do exist. One approach is comparing the classification result of each classified image (post-classification) and the other approach is putting the images and their spectral bands/indices into a classification algorithm (Hirschmugl et al., 2017; Banskota et al., 2014). In context of the time-series analysis the following four different approaches can be named: threshold based change detection, curve fitting, trajectory fitting and temporal segmentation (Banskota et al., 2014; Hirschmugl et al., 2017). Commonly used algorithms to detect forest loss and regrowth are LandTrendR (Kennedy et al., 2010; Frazier et al., 2015, Grogan et al., 2015), BFAST (DeVries et al., 2015, Hamunyela et al., 2016, Reiche et al., 2015), CCDC (Zhu and Woodcock, 2014) and the

Composite2Change (C2C) approach (White et al., 2017; Hermosilla et al., 2016). The applicability of the Kalman filtering approach to Sentinel-2 time series has also been investigated.

### **LITERATURE REVIEW ON FOREST LOSS DETECTION**

There are several articles that review the state of the art in change detection through remote sensing. Many of the reviewed methods and articles are developed specifically for change detection in forests. Hirschmugl et al. (2017) review literature that is about the monitoring of forest disturbances and degradation by the use of optical remote sensing data. Most of the methods presented in their article were developed on the basis of Landsat time series. The review article by Zhu (2017) is limited to methods utilizing Landsat imagery, the vast majority of the articles reviewed referring to forest as the change target. In the review article by Banskota et al. (2014) forest monitoring studies that use Landsat time series data are reviewed.

Hirschmugl et al. (2017) discriminate between *bi-temporal* and *multi-temporal* change detection approaches. In the following section, one bi-temporal and some multi-temporal approaches for forest loss detection are described.

From the category *bi-temporal approaches* the image-to-image change detection is described. For an image-to-image change detection, two images are assessed: one before and one after a, e.g., a forest loss event. A change (e.g. a forest loss) is indicated by the difference in the spectral characteristics between the images (Hirschmugl et al., 2017).

From the category *multi-temporal approaches*, five commonly used algorithms that are applicable to forest loss are described in the following section. Four methods (Vegetation Change Tracker, LandTrendR, BFAST and the method by Hansen et al., 2013) are selected for further description because each of them is mentioned in at least two of the three review articles (i.e. Zhu, 2017; Hirschmugl et al., 2017; Banskota et al., 2014). The fifth method (CCDC) is selected because it has been chosen as the change detection algorithm for the operational land cover classification in the frame of the USGS LCMAP program for the United States (Zhu 2017).

The Vegetation Change Tracker (VCT) (Huang et al., 2010) is a thresholding method, i.e. it "employs a predefined threshold for identifying a land cover (mostly forest) in the time series, and changes are detected when there are significant deviations from the threshold" (Zhu, 2017). Huang et al. (2010) developed and tested VCT at many study sites in different states of the USA. Landsat images are transformed into a multi-spectral forest probability index called "Integrated Forest Z-score" (IFZ). An IFZ higher than a certain threshold indicates non-forest. By comparison of two images, forest loss events can be detected. Huang et al. (2010) examine disturbance patterns with a biennial interval from 1984 to 2006. In their validation, overall accuracies of around 80% are achieved. Due the design of the method, it is only suitable for abrupt changes, i.e. forest loss events, but not gradual changes like forest regrowth (Zhu, 2017).

Landsat-based detection of Trends in Disturbance and Recovery (LandTrendR, Kennedy et al., 2010) is a temporal segmentation method, i.e. an "approach that needs to have all the historical time series data available at the same time. By segmenting the time series into a series of straight line segments based on the residual-error and angle criterions, both abrupt and gradual changes can be detected based on the derived straight line segments at pixel level" (Zhu, 2017). The method was developed by Kennedy et al. (2010) for the detection of both forest loss and forest regrowth on basis of Landsat imagery. The study by Kennedy et al. (2010) relies on a very long time series for the period from 1985 to 2007.

The Breaks For Additive Season and Trend (BFAST) Monitor is a statistical boundary method. It "expects the time series to follow a statistical boundary and any significant departure from the boundary is detected as a change" (Zhu, 2017). This algorithm was originally developed for vegetation change in any land cover class on the basis of MODIS data (Verbesselt et al., 2010). Lately, BFAST Monitor has been used for change detection in forests by the use of Landsat time series (DeVries et al., 2015; Hamunyela et al., 2016; Reiche et al., 2015). BFAST allows for the detection of both abrupt and gradual changes. Therefore, the method is suitable for both, forest loss and forest regrowth monitoring. For a detailed

description of the general method see Chapter 1.1.1 Signal anomaly detection based on optical Time Series. BFAST uses a historical time series of, e.g. an index like the NDVI, to fit a curve to a series of pixel values spanning several years. In the currently available literature it is often the full Landsat time series, starting from 1984. From the three above-mentioned studies focussing on forest changes detection, DeVries et al. (2015) use a time series spanning six years for the modelling of the curve. Hamunyela et al. (2016) utilize the full Landsat time series from 1984 to 2014. The Landsat time series used by Reiche et al. (2015) spans the period from 2004 to 2012.

Another statistical boundary method is the Continuous Change Detection and Classification (CCDC; Zhu and Woodcock, 2014). CCDC evolved from the Continuous Monitoring of Forest Disturbance Algorithm (CMFDA) which was originally designed to detect forest loss (Zhu et al. 2012). CCDC, however, is designed to detect both abrupt and gradual changes, i.e. it is suitable for both forest loss and regrowth monitoring. The algorithm uses all spectral bands of all available Landsat images starting in 1984 (Zhu, 2017). For a detailed general description of the CCDC method see Chapter 1.1.1 Signal anomaly detection based on optical Time Series.

A good example of an operational application of forest loss detection is the global mapping of forest changes in the time period 2000–2018 by Hansen et al. (2013). Hansen et al. (2013) fully rely on Landsat data. The applied method is a trajectory classification method, i.e. it "first extracts information from Landsat time series for places that have undergone certain kinds of change (for training purpose), and later uses this information to further classify every Landsat time series in the image" (Zhu, 2017). The study by Hansen et al. (2013) utilizes a multi-date classification method that relies on a high frequency Landsat time series. The method allows for the detection of forest loss and gain (Zhu, 2017).

The loss layer of the Global Forest Map product by Hansen's team (Hansen et al., 2013) shows forest loss events for the time period 2000-2018 with an annual allocation of the loss events. While the details about the derivation of this layer are not documented very well, an article about land cover monitoring of the conterminous United States (Hansen et al., 2014) documents the mechanisms of such a multi-date classification method. Amongst other things, that article describes the quantification of forest cover loss in the time period 2006 to 2010. A supervised classification algorithm is used to discriminate forest change from unchanged forest on pixel level. The imagery used are weekly mosaics of Landsat imagery (spectral bands). These mosaics are used to derive image metrics (temporal-spectral features or time features; e.g. 5-year percentile of 10% per each of the Landsat bands). Expert interpretation is used to build a training set which is applied to carry out a supervised classification on basis of the time features. The outcome of the classification is a map that shows change events in the period 2006–2010. Afterwards, each change pixel in this map is analysed to identify the date of the change. The analysis is based on the Normalized Difference Water Index (NDWI) (Gao, 1996) for each of the weekly Landsat mosaics. The exact year of the change (e.g. forest loss) is deduced from the NDWI trajectory of a change pixel (see Figure 1-5).



In the frame of ECoLaSS, it is of further interest to achieve harmonization and continuity between a) historic HRL forest masks (HRL Forest Type 2012, HRL Dominant Leaf Type 2015) and b) future forest masks. An elegant method to achieve continuity is to set a historic HRL forest mask as the benchmark forest mask. Any future observation of "non-forest" within this benchmark forest mask can be assumed to be forest loss. For a detailed description of this map-to-map change detection method see Chapter 2.1.1.

One advantage of detecting forest loss using the map-to-map change detection method is that the method is fully detached from the input data used for forest map generation. A forest pixel in the HRL Dominant Leaf Type 2015 product, for example, might be derived from a combination of Sentinel-2, Landsat, IRS and SPOT data. A forest map for the year 2018 might be a combination of Sentinel-2 and Sentinel-1. The map-to-map change detection is completely independent from the input data used for the forest map generation. So, the fusion of images from different optical sensors or even the fusion of optical and SAR imagery can be solved in complete isolation from the change detection. Provided that in phase 2 in ECoLaSS the status layers for the years 2017 and 2018 (Tree Cover Masks) are generated in a comparable fashion as depicted in WP33, different change detection methods are applied to the improved 10m resolution products, on the basis of NDVI derived change indicators and Kalman filter testing.

### 1.2.2 Forest regrowth

One of the challenges in characterizing rates of forest recovery is the absence of a universal definition of what is meant by the term recovery in a forest context. Recovery involves the return of vegetation or the establishment of a secondary forest canopy following a disturbance, meaning that there are often different terminologies used such as 'reforestation', 'regrowth', or 'regeneration' to describe what happens to forests following disturbance (Bartels et al., 2016). For example, Pickell et al. (2016) defined recovery as 80% of the mean spectral value of the 2 years prior to the disturbance for each disturbed pixel in the time series. Other studies adapted this definition and used metrics that represent the number of years required for a disturbed pixel to return to 80% of the pre-disturbance spectral value (Frazier et al., 2018; White et al., 2017). However, as stated by Kennedy et al. (2012) and Griffiths et al. (2014), spectral recovery is not a direct measure of forest recovery. This means, that it should be interpreted in the context of a priori expectations of recovery, which are typically derived from ground plot measurements (White et al., 2017). As demonstrated in Bartels et al. (2016), with using field data alone, there is a notable lack of information on forest structural attributes that are important to characterizing forest recovery (e.g., canopy cover). Therefore, information from remotely sensed data can provide additional insight into relative forest regrowth rates and changes in spectral recovery. Ideally, measures made from remotely sensed data should be supplemented with ground data and an ecological understanding of the forest recovery process.

Most of the studies investigating forest regrowth using Landsat Time Series (LTS) to track forest recovery following wildfire (e.g., Bolton et al., 2015; Franks et al., 2013; Frazier et al., 2018) and less commonly, following harvest (Kennedy et al., 2012; Schroeder et al., 2007; White et al., 2017).

Numerous approaches have been used to characterize post-disturbance recovery, with vegetation indices and derivatives thereof being the most common ones (White et al., 2017; Chu and Guo, 2014). For example, the NDVI has often been used as an assessment of vegetation "greenness" to examine active photosynthetic vegetation quantity and to track forest recovery (Franks et al., 2013; Frazier et al., 2018).

Figure 1-6 shows a typical NDVI response after a disturbance event (e.g., fire) with a high pre-disturbance value sharply dropping at the time of the event and then slowly recovering several years after the event as the vegetation re-grows.



The C2C algorithm produces annual, cloud-free best-available-pixel (BAP) image composites and includes change detection and typing, spectral trend analysis of breakpoints, object-based segmentation, and Random Forest (RF) classification. The outputs of the C2C processing are (i) gap-free, surface reflectance proxy image composites, (ii) annual change detection information, and (iii) a series of descriptive change metrics that characterize the change events (i.e. magnitude, duration), as well as the pre- and post-change conditions (Hermosilla et al., 2016; White et al., 2017).

The BFAST algorithm can be used for breakpoint analysis to check whether values following a disturbance in a time series return to the previously defined stable state or not. Although a lot of studies use BFAST to monitor forest clearing (i.e., forest loss), the BFAST approach can also be adapted to include forest recovery (DeVries et al., 2015, Grogan et al., 2016).

It should be noted that post-disturbance forest recovery is a process rather than a state, meaning that data from longer time periods is required to capture a more detailed process of forest recovery. Bartels et al. (2016) found that it took 5-10 years for most forest ecosystems of Canada to attain a benchmark canopy cover of 10% after wildfire or harvest. Kennedy et al. (2012) proposed a five-year window stating that it was critical to monitor the initiation of recovery. Therefore, current HRL Forest update frequencies (i.e., every 3 years) might capture regrown forest, however, the above-mentioned studies state that larger time intervals might be necessary. Future layers with potentially more narrow intervals (e.g., an update on a yearly basis) might be well suitable for detect forest loss, but might not be able to detect forest recovery processes.

### 1.3 Imperviousness Change

There are only a limited number of studies investigating the change of Impervious Surface Area (ISA). Powell et al. (2008) used 13 dates to quantify ISA in a 34 year period between 1972 and 2006. A spectral unmixing approach was used to determine the fractional composition of Tasseled Cap components of normalized Landsat scenes. To identify the spectral endmembers, they examined digital orthophotos. Expert-based spatial and temporal rules were applied to refine the model. This resulted in a per-pixel ISA time series. To reduce the phenology based variation of the results, they applied a 3-year moving average on the ISA values to map the change. Xian and Homer 2010 and Homer et al. (2015) describe the processing of the ISA products of the National Land Cover Database (NLCD) provided by the U.S. Geological Survey. The ISA change product was processed for the time steps 2001, 2006 and 2011 and is based on normalized Landsat pairs of the respective years. The accompanying land cover products were used to determine if a change in land cover took place between two time steps for single pixels. If no change took place, the respective pixels were used as training dataset for a regression tree model to determine the imperviousness of the pixels that were identified as change pixels. Lu et al. (2011) focused in their study on the detection of imperviousness change at the complex urban-rural transition zone. They applied a two-step calibration of their Landsat-based model. The first step consists of regression model for 2008 Landsat imagery, which is based on Quick-Bird derived impervious surface. In a second step, other Landsat acquisitions were calibrated on the 2008 scene to be able to apply the existing regression model on them. The pixel-based ISA values were spatially aggregated to determine the change between the time steps. Zhou et al. (2012) applied a sub-pixel classifier on Landsat scenes for the years 1980, 1990 and 2000 in order to derive ISA and ISA change. To identify the increase and decrease of imperviousness, they applied image differencing of the ISA products of the 3 decades. Dams et al. (2013) estimate ISA by means of Landsat imagery for two time steps (1986 and 2003). First they create a binary urban mask. Afterwards, an ISA map is derived from high resolution IKONOS imagery. This map is used as training data for deriving ISA from Landsat imagery, employing a linear regression model. The change is determined by calculating the difference between the two ISA maps. Xia et al. (2011) and Du et al. (2015) used two CBERS 01/02 and one HJ-1 scene to derive the ISA percentage for the period of 2001 to 2009 and analysed the change. They applied a linear spectral mixture model (LSMM) and a nonlinear backpropagation neural network to identify ISA. The change analysis was conducted by a post-classification approach. Hao et al. (2015) analysed the relation between ISA and relative mean annual surface temperature. For this purpose, they retrieved ISA using spectral mixture analysis (SMA) by

performing a normalisation procedure (normalised SMA) to enhance the separability of the urban land cover. Next, a linear SMA was applied to derive the urban composition of each pixel. Based on the normalised spectrum a principal component analysis was conducted leading to the identification of the three endmembers vegetation, soil, and impervious surface. ISA was calculated for three time steps including 1990, 2001, and 2014. To determine the change the difference of the corresponding images was calculated. Wang et al. (2017) mapped ISA for the time steps 2000 and 2010 using Landsat data. Here, a regression tree-based approach was implemented to first derive ISA for the year 2010. In this context, the spectral bands of Landsat and four additional spectral indices were employed. Training data for regression was manually extracted by photointerpretation of very high resolution imagery. To limit the area of interest to the actual urban extent, a binary urban mask was classified using object-based texture features and the random forests classifier. Next, ISA was derived for the year 2000. For this purpose, a binary urban mask was classified for the year 2000 to identify persistent built-up areas between 2000 and 2010. Training data for the year 2000 was collected from areas where no change occurred in the built-up masks. In order to analyse the change between 2000 and 2010 a difference image was calculated. Yang and Sun (2017) applied combined pixel- and object-based methods to derive ISA change of Landsat images of 1990, 2000 and 2010 time steps. They determined the land-cover of 2000 with a pixel-based Support Vector Machine (SVM) classifier. The source image was segmented into image objects and intersected with the land-cover result. Using the IR-MAD approach (Canty and Nielsen, 2008), they identified pixels with a high probability of change between 1990 and 2000 or 2000 and 2010. The change pixels were reclassified by SVM for 1990 and 2010. They did not investigate the percentage of imperviousness per-pixel.

Besides the opportunities, which come with above mentioned studies, imperviousness change analyses have to cope with challenges and limitations on the data and processing side. Based on the geometric accuracy and the varying spatial resolution when using various sensors, detailed change detection at pixel level is difficult to conduct (Lu et al., 2011). On subpixel level the complexity may increase, due to the mixed reflectance values of the share of surface cover (Wu et al., 2015). Varying climatic seasonality between time-steps and long-term climate changes have an impact on the phenological cycle of vegetation and can confuse change monitoring analyses (Yu et al., 2017). The accuracy of the results for the different time-steps is essential for the overall accuracy of change analyses (Hao et al., 2016, Hsiao and Cheng 2016). Shi et al. (2017) used a Landsat time-series dataset from 1987 to 2016 to determine annual ISA change. They applied an SVM classification, which was revised based on the class probability with an uncertainty-based spatial-temporal consistency (USTC) model to improve the accuracy of the time-series change analysis. Their results show that the approach performs better than the SVM classifier or a model that assumes temporal consistency.

Golda et al. (2013) have tested several automated change detection methods to produce a change map over New Delhi between 1974 and 1999, using Landsat MSS and Landsat TM images. They only used a bi-temporal approach and constituted a Spectral Change Difference (SCD) dataset by simple differencing, simple rationing as well as computing the absolute and Euclidian distances and a chi-square transformation. This SCD dataset was then fed to fusion algorithms. The first one, at a feature level, used Kittler-Illingworth segmentation and thresholding as pre-processing, before using the fuzzy set theory to produce the change map. The other presented algorithms were set at decision level; majority voting, Dempster-Shafer fusion and fuzzy integral were compared. Fuzzy integral and simple rationing outperformed other methods, clocking at 81% and 80% of overall accuracy. Finally, Golda et al. (2013) applied the same panel of algorithms to the same pre-processed images, but with an intermediate step where a curvelet transformation was tested. Almost all algorithms have been improved, the fuzzy integral exhibiting an overall accuracy of 98%, closely followed by the fuzzy set fusion and the Dempster-Shafer fusion. The curvelet-based transformation method enhances the edges and provides information both in the spatial and spectral domains in the form of a multi-scale representation of curves describing the objects present in the two images. It was developed by Candes and Donoho in 1999 and has been widely tested in SAR environment for change detection, but very few studies have explored this method in the optical domain.



investigation. *“The thermal radiance values of TM/ETM 6 data, the Normalized Difference Vegetation Index (NDVI), and new variables (brightness, greenness, and wetness) generated by the Kauth–Thomas Transforms (KT) algorithms from Landsat TM/ETM data served as the feature nodes of a DT classifier and contributed to improving the classification results. It showed an overall accuracy of more than 85% and a Kappa statistic of agreement of about 0.79 in 1996 and 2001 with the exception of about 0.69 in 1988. [...] The distribution maps of land degradation in the years of 1988, 1996 and 2001 were generated respectively based on the classification results. Their change maps were created by the difference between the distribution maps from 1988 to 1996 and from 1996 to 2001 respectively.”* (Chen, S., and Rao, P. 2008)

Melville et al. (2018) identified native grassland communities found in the Tasmanian Midlands region using a random forest classification approach based on Landsat ETM+ and WorldView-2 data from 2010. To identify appropriate remote sensing datasets to derive accurate and frequently updateable maps, repeat classification and statistical significance testing has been performed. Training and validation data are based on a pre-existing field dataset. The results show that mean class accuracies lies between 54% and 87%. Higher mean accuracies were achieved with the WorldView-2 dataset.

Sun et al. (2017), identified areas of grassland degradation and restoration using the Sen + Mann-Kendall method. MODIS NDVI data from 2001 till 2012 are combined with meteorological data (temperature, ground downward shortwave radiation) and Landsat TM datasets. Trend analysis of long time-series vegetation changes is performed with regression methods. The authors state *“The proposed methods of grassland degradation and restoration monitoring and driving forces analysis were suitable for longtime-series vegetation indicators datasets at the regional scale.”* (Sun et al., 2017).

## 1.5 Agricultural Dynamics

As explained in the introduction, the type of signal anomalies measured for cropland types are usually caused by management activities and phenology therefore has to be addressed differently and sometimes to be combined with thematic information (existing LC/LU maps) or classification results from WP33 [AD06].

Time series analysis for change detection has been used in different ways in the literature as in operational systems. Most of the early warning system for agriculture monitoring such as Monitoring Agriculture by Remote Sensing (EU/JRC-MARS), Global Information for Early Warning System (FAO-GIEWS), Food Early Warning System (US-FEWS), GEOGLAM, relies on NDVI profile analysis to detect anomalies with regards to normal or averaged profiles computed from the long term archives at 250m or 1km resolution. While it is expected that the Sentinel-3 should allow capitalizing on these long term archives, they are not relevant for the Sentinel-1 and Sentinel-2 time series analysis for annual crop growth cycle monitoring at the 1km pixel or landscape level. Indeed MODIS, PROBA-V or even MetOp data are more efficient to handle for such coarse monitoring combining all parcels within a 1km pixel.

The density of the Sentinel-1 and Sentinel-2 time series open the way for application at parcel or sub-parcel level such as crop specific monitoring along the season, detection of agriculture practices like tillage, sowing or grassland mowing, or to identify land abandonment. These applications are related to change detection in the expected time series but are not standard land use or land cover change detection applications. This Chapter reviews the few recent results obtained from Sentinel-1 and Sentinel-2 as the density of these time series have opened the way to new methods.

Obviously crop specific applications first requires to know the parcel crop type either from an early crop type classification or from farmers declarations as recorded in LPIS layers developed in the EU Common Agriculture Policy framework.

Three recent methods attempts to make use of the full phenological profile for crop type mapping:

### 1. Method based on Phenological Sequence Patterns (PSP)

Bargiel (2017) introduces a novel approach to incorporate the full knowledge of the crops’ phenology using a sequence-based classification approach. Those sequences are called “frequent sequential patterns” (Julea et al., 2011). Bargiel (2017) transfers this approach to a supervised classification with the incorporation of phenological knowledge to classify crops based on phenological sequential patterns (PSP). He uses ground truth data to identify the different crop types and differentiate six phenological phases (according to the BBCH scale by Meier and Bleiholder (2006)). As input for the crop map, he uses 99 Sentinel-1A images. The probability  $p(y_s)$  for a crop-type  $y$  at a certain phenological stage  $s$  is computed for each pixel of the investigated area, using Naïve Bayes and RF. From there, a new value  $y_s$  is calculated for each class and phenological stage.

$$y_{seq} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{k} \frac{\sum_{s=1}^k p(y_s)}{c}$$

where,

- $c$  is the coefficient of variation of the sequences’ single stage probabilities,
- $i$  is the index of the pixel and its eight nearest neighbours,
- $n$  is the number of considered neighbours.

Each pixel is then labelled with the crop type maximizing  $y_s$ . The method was evaluated by comparing it to MLC and RF, without the PSP approach. The overall results obtained with the PSP approach are better than with MLC or RF for most crops. PSP strongly improved the accuracy for different types of cereals for example. Bargiel (2017) suggests several further research. The author recommends testing the PSP approach on Sentinel-1 data fused with Sentinel-2 data, evaluating the contribution of each phenological stage to the classification result and calculating the feature importance within each stage to limit the classification to the most important ones. Finally, he notes PSP based classification seems to be very sensitive to subtle anomalies in certain parcels (e. g. a potato field overgrown with weeds), but claims this establishes possibilities for detecting organic farming or drought monitoring.

### 2. Method based on Time Weighted Dynamic Time Warping (TWDTW)

Belgiu and Csillik (2018) introduce another approach for crop mapping exploiting full time series. They assessed the performance of a Time-Weighted Dynamic Time Warping (TWDTW) method, based on Sentinel-2 time series, using pixels and parcels as spatial analysis units and compare it to RF classification. The method was tested on three sites, in Romania, Italy and the USA. TWDTW involves two major steps, namely (i) creating a smoothed temporal pattern for the NDVI (8-days frequency) for the sample point; (ii) applying a logistic Dynamic Time Warping, adding a time-weight. This function takes each pixel location in the NDVI time series and analyses it in conjunction with the temporal patterns identified for training samples. The output is a raster with layers containing dissimilarity measures for each temporal pattern. The method then generates a crop type map by associating each pixel to the class with the lowest dissimilarity value. Belgiu and Csillik (2018) concluded that the TWDTW approach outperformed the Random Forest method when only small samples sets are available per crop, is robust to irregular temporal sampling and to the annual changes of vegetation phenological cycles.

### 3. Classification of Biophysical variables profiles

In order to integrate multi-sensor time series like Sentinel-2, Landsat-8 and possibly Sentinel-1 in coherent time series, Waldner et al. (2015) investigated the potential of structural biophysical variables as common features to consistently combine multi-sensor time series for land/crop cover classification. High resolution LAI, FAPAR and FCOVER were successfully retrieved from Landsat-8 and SPOT-4 and displayed consistent average temporal trajectories. However, the class variability and signal-to-noise ratio increased compared to simple NDVI. Classifications with structural biophysical variables reached

only 77% at the best end-of-season overall accuracy to be compared with 95% accuracy level achieved with the spectral bands and NDVI. When assuming that the cropland extent is known, crop type classification reaches 89% with spectral information, 87% with the NDVI and 81%–84% with biophysical variables. While this approach is elegant for multi-sensor time series fusion, the retrieval performances still do not allow to challenge classical approaches.

## 2 Testing and Benchmarking of Candidate Methods

The major challenges of with time series data sources are addressed. A major challenge are missing observations due to high frequency cloud cover and low sun incidence angles. Therefore, time series change detection using SAR as alternative or complementary image data source and the usage of multi-sensor data sets have been tested primarily in phase 2 with a subsequently applied plausibility analysis. The outcomes of this testing phase will be discussed in this chapter.

### 2.1 Forest Change

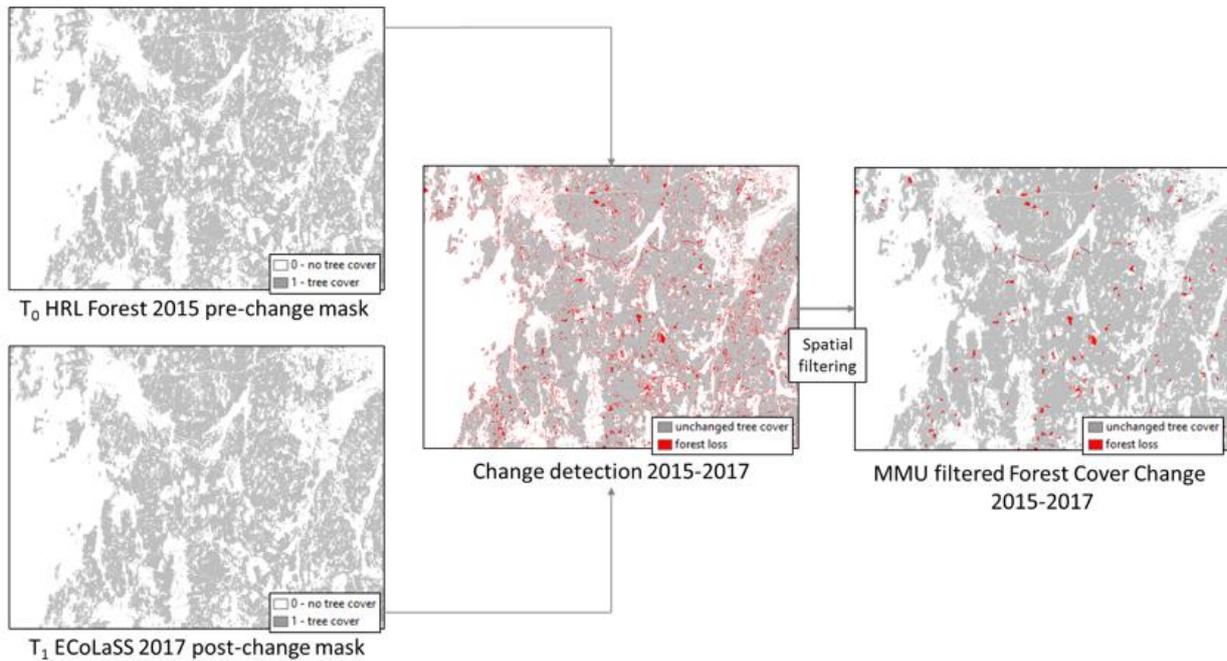
The aim of this work is the evaluation of a forest change detection approach based on temporal-spatial metrics (using Sentinel-2 and Sentinel-1 data) in the ECoLaSS North, Central, and South-East test sites. On the one hand, the image-to-image approach applied compares a pre- and post-change forest mask (phase 1: HRL2015 tree cover mask and newly classified 2017 tree cover mask; phase 2: 2017 tree cover mask and 2018 tree cover mask) to detect areas of assumed forest loss. On the other hand, time series analysis methods on the basis of time feature computation have been implemented to derive the 2017 and 2018 forest status layer Dominant Leaf Type, which is used to derive the relevant tree cover masks per reference year. Given the fact that these tree masks have been generated under a consistent workflow, the change detection in phase 2 includes a simple yet operational and accurate enough change indicator application, which is more reliable than the approach applied in phase 1. This is achieved by means of combining time features differences between 2017 and 2018 (e.g., minimum NDVI, S-2 bands statistics) and the difference image between the classifications of 2017 and 2018, plus some filtering and buffering refinements, ending in a MMU of 1ha, which is fully in line with the ongoing HRL2018 Forest production. Additionally, tests with the Kalman filter approach have been performed. While this work concentrates on negative changes, i.e. a decline in forested area, Chapter 1.2.2 of this report reviews potential methodologies for the delineation of areas of forest increase or regrowth.

#### 2.1.1 Description of candidate methods

Two change mapping approaches have been tested in the frame of WP34. The classical image-to-image approach considering different data scenarios, and the Kalman filter.

##### IMAGE-TO-IMAGE

In phase 1 of EcoLaSS, the most common image-to-image change approach has been tested using a difference image of the benchmark HRL2015 tree cover mask and new tree cover classifications for the year 2017, derived from different time series data and indicators. The post-change mask is delineated via a random forest classifier for three different input data scenarios, using time features (AD06, Chapter 3.1.4) of Sentinel-1, Sentinel-2 and combined Sentinel-1 and Sentinel-2 data. The forest/non-forest training sample data was initially extracted from the HRL2015 DLT product and manually verified and improved. The change methodology, shown in Figure 2-1, focusses on the detection of negative forest loss only, as the temporal distance between both masks is less than two years. The difference image of both masks is limited to the intersecting forest area of the HRL2015 mask and the new 2017 classification. This comparison is made at 20m by resampling the improved 10m Tree Cover Mask 2017. Subsequently, different MMUs have been tested.



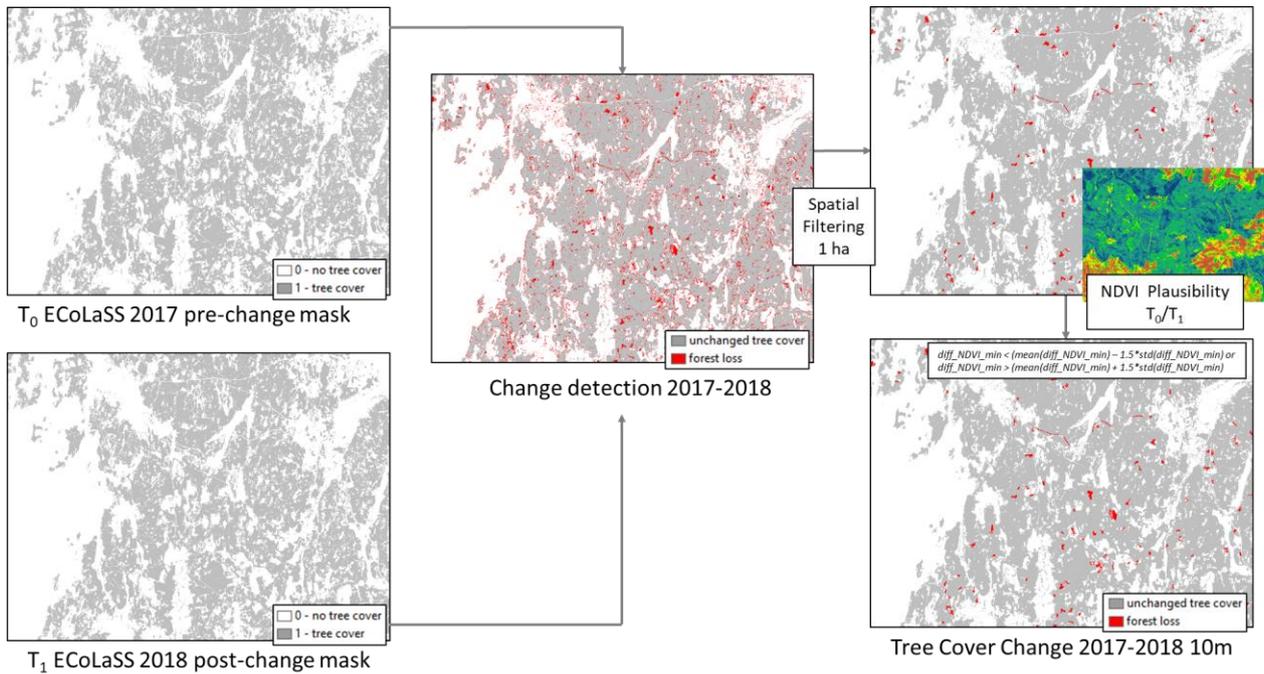
**Figure 2-1: Phase 1 map-to-map change detection with HRL2015 Tree Cover Mask ( $T_0$ ) and newly classified 2017 Tree Cover Mask ( $T_1$ )**

In phase 2 the 2017 and 2018 tree cover masks have been generated, a modified image-to-image change detection approach has been successfully tested. The aim was to develop a simple and transferrable method in order to achieve operational feasibility at larger scales and allowing for a cost-efficient automation. The tests have proved successful, providing satisfactory accuracies in all test sites, and the methodology has been finally selected for generation of the incremental updates in Task 4.

The difference between the status layers of both reference years is computed to derive the forest changes (unchanged, forest gain and forest loss). Forest gain is generally discarded due to the short span of the time series when compared to forest regrowth dynamics. As aforementioned, the change detection is targeting forest loss detection. Filtering isolated pixels and negative buffering reduce artefacts resulting from positional shifts of Sentinel data and border effects coming from mixed pixels. In turn, the change indicator is computed as the NDVI time feature difference between 2018 and 2017. The resulting time feature mask (e.g., minimum NDVI) is recoded assigning change to all areas where the mask is less than the mean minus the standard deviation and unchanged otherwise. Next, by combining the time feature mask recoded and the cleaned difference between the status layers (post filtering, buffering and recoded), pixels labelled as changes in both, remain as such. Finally, the combination of this layer providing the change is combined with binary tree cover mask. Filtering by a suitable MMU allows for erasing small differences between the two status layers (it could happen that a small patch was captured better in one or the other layer and therefore a part of the patch is highlighted as a change). In case of forest, and in view of the testing experiences in the different biogeographic regions, it is considered that one hectare is a meaningful MMU to highlight changes in the tree/forest cover. This approach was first applied to the grassland change detection tests in ECoLaSS and has been successfully tested in the Forest domain as well to verify the transferability to other domains. It is characterised by its simplicity, cost-efficiency and operational feasibility from regional to continental scales, albeit a proper thresholding might point to local/regional stratifications for quality performance.

The Figure below depicts the workflow of the change detection approach developed in phase 2 with a subsequently applied NDVI time series thresholding. This method can be applied to other time features change indicators (e.g., derived from S2 band or other indices). In phase 2, the change product is significantly improved as a consequence of the enhanced spatial resolution of the compared status layers

(tree cover mask 2017/2018 at 10m), which have been produced in a consistent manner with integration of S1 time features and a common feature selection method (k-fold or variance).



**Figure 2-2: Phase2 map-to-map change detection with Tree Cover Mask 2017(T<sub>0</sub>) and Tree Cover Mask 2018 (T<sub>1</sub>) and NDVI plausibility step**

Subsequently, the forest area change maps were validated using a stratified random sampling approach including LUCAS 2018 points. Results are presented in Chapter 2.1.3.

### KALMAN FILTER

The underlying assumption of the near real-time monitoring approach presented in this chapter is that the normal temporal trajectory of a given spectral band over the course of the year can be captured by a time-series model. For example, the phenological cycle typically encountered in forests can be approximated by a sum of trigonometric functions of different frequencies and a trend component. Consequently, newly available observations may be compared to a model-based forecast. Significant changes of the spectral signature, possibly linked to a forest disturbance, are indicated by statistically significant deviations between newly available observations and the forecast.

Thus, the presented Kalman filter approach shares many principles with the change-detection part of CCDC (Zhu and Woodcock, 2014b) and BFAST Monitor (Verbesselt et al. 2012), but instead of modelling the normal spectral trajectory as a deterministic function of time, the Kalman filter allows a stochastic evolution of the trajectory by putting more weight on recent observations.

For the Kalman filter to be applied, the time-series model has to be mathematically formulated using the discrete-time state space representation. The measurement equation linking the system state to a set of observables is

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{r}_k$$

where  $\mathbf{z}_k$  is an  $n_k$ -element vector of observations,  $\mathbf{x}_k$  is the  $p$ -element state vector,  $\mathbf{H}_k$  is an  $n_k \times p$  matrix, and  $\mathbf{r}_k$  is an  $n_k$ -element vector of serially uncorrelated observation noise with mean zero and  $n_k \times n_k$  covariance matrix  $\mathbf{R}_k$ :

$$E(\mathbf{r}_k) = \mathbf{0} \text{ and } \text{Var}(\mathbf{r}_k) = \mathbf{R}_k$$

The temporal evolution of the state vector is described by a dynamic model using the transition equation

$$\mathbf{x}_k = \Phi_k \mathbf{x}_{k-1} + \mathbf{q}_k$$

where  $\Phi_k$  denotes the  $p \times p$  transition matrix and  $\mathbf{q}_k$  is a  $p$ -element vector of serially uncorrelated process noise with mean zero and  $p \times p$  covariance matrix  $\mathbf{Q}_k$ :

$$E(\mathbf{q}_k) = \mathbf{0} \text{ and } \text{Var}(\mathbf{q}_k) = \mathbf{Q}_k$$

The subscript  $k$  indicates that a variable may vary with time. Note that no assumptions regarding the distributions of the observation and process noises are made at this point, but they are supposed to be uncorrelated with each other in all epochs. It is further assumed that the initial state  $\mathbf{x}_0$  is known with a level of uncertainty characterized by the state error covariance matrix  $\mathbf{P}_0$ . The variables  $\mathbf{H}_k$ ,  $\mathbf{R}_k$ ,  $\Phi_k$ , and  $\mathbf{Q}_k$  are known as system matrices. To clarify the notation it should be mentioned that an index  $k$  on  $\Phi$  always means that the respective transition matrix describes the temporal change of the state vector with respect to the previous epoch. For a detailed account regarding the definition of the system matrices of time-series models please refer to Harvey (1989).

Once the system matrices and the initial values  $\mathbf{x}_0$  and  $\mathbf{P}_0$  are defined, the discrete-time Kalman filter algorithm can be used to obtain estimates for the state and its error covariance matrix in subsequent epochs  $k = 1, 2, \dots, m$ . The Kalman filter processes observations recursively, one epoch at a time, and each recursion may be divided into several steps. The first one is known as time update step and yields the predicted (*a-priori*) estimates  $\tilde{\mathbf{x}}_k$  and  $\tilde{\mathbf{P}}_k$  based on the dynamic model and the previous estimates at time  $t_{k-1}$ :

$$\tilde{\mathbf{x}}_k = \Phi_k \hat{\mathbf{x}}_{k-1}$$

$$\tilde{\mathbf{P}}_k = \Phi_k \hat{\mathbf{P}}_{k-1} \Phi_k^T + \mathbf{Q}_k$$

Step two is to compute the *a-priori* measurement residual  $\mathbf{y}_k$  and its associated covariance matrix  $\mathbf{C}_k$  according to the equations given below. The residual represents the difference of the prediction to the actual measurements and is referred to as *innovation*, since it contains new information currently not present in the predicted state.

$$\mathbf{y}_k = \mathbf{z}_k - \mathbf{H}_k \tilde{\mathbf{x}}_k$$

$$\mathbf{C}_k = \mathbf{H}_k \tilde{\mathbf{P}}_k \mathbf{H}_k^T + \mathbf{R}_k$$

In the final step of each recursion, the new information is merged with the predictions to obtain improved (*a-posteriori*) estimates  $\hat{\mathbf{x}}_k$  and  $\hat{\mathbf{P}}_k$ . Therefore, it is known as measurement update step. The Kalman gain matrix  $\mathbf{K}_k$  given below determines how much the newly acquired measurements will influence the *a-posteriori* estimates of the state and its error covariance.

$$\mathbf{K}_k = \tilde{\mathbf{P}}_k \mathbf{H}_k^T \mathbf{C}_k^{-1}$$

The elements of  $\mathbf{K}_k$  are ranging from 0 to 1, thus it can be interpreted as a weighting matrix. With the update equations stated below, the basic Kalman filter recursion is complete.

$$\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k + \mathbf{K}_k \mathbf{y}_k$$

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \tilde{\mathbf{P}}_k$$

When the sequence of measurements processed by the filter contains outliers, an additional outlier detection step should be included *before* the measurement update step. The properties of the innovations can be exploited to detect anomalous measurements by means of a statistical test. Provided that the underlying model assumptions are correct and the observation noise is Gaussian, the innovations will be normally distributed with mean zero and covariance matrix  $\mathbf{C}_k$ , that is

$$\mathbf{y}_k \sim N(\mathbf{0}, \mathbf{C}_k)$$

The test statistic  $\hat{T}_k$  given in below follows the  $\chi^2$ -distribution with  $n_k$  degrees of freedom, where  $n_k$  is the number of observations in epoch  $k$ . The hypotheses to be tested on a significance level  $\alpha$  are stated below.

$$\hat{T}_k = \mathbf{y}_k^T \mathbf{C}_k^{-1} \mathbf{y}_k, \hat{T}_k \sim \chi^2(n_k)$$

$$H_0: \mathbf{y}_k = \mathbf{0} \text{ if } \hat{T}_k \leq \chi_{n_k, 1-\alpha}^2$$

$$H_1: \mathbf{y}_k \neq \mathbf{0} \text{ otherwise}$$

Considering that anomalous observations will cause large innovations, the null hypothesis will be rejected if outliers are present in the current epoch. In order to avoid a negative influence on the state estimate, the measurement update step is not carried out.

The flowchart depicted in Figure 2-3 illustrates how the methods discussed in the preceding chapters are joined together in order to create a data-driven algorithm capable of detecting abrupt changes on pixel-level. Some further explanatory comments are given below.

- (1) The initial state, as well as the observation noise, are estimated on a pixel-level using the IRLS method. Therefore, the user has to supply a stack of historic images as a training dataset.
- (2) The tuning parameters and the time-series model components specified by the user are applied globally; hence they are the same for all pixels.
- (3) New images acquired in the monitoring period are processed one at a time in a Kalman filter loop. A hypothesis test is used to identify anomalous observations showing significant deviations to the prediction.
- (4) Each pixel features an “anomaly counter”. Each time an observation is marked as anomalous, the counter is *incremented* by 1. In contrast, the counter is also *decremented* by 1 if the current observation is not statistically conspicuous, although it can never become lower than zero. This principle ensures that a certain user-defined “change threshold” of the counter can only be reached when outliers occur temporally aggregated.
- (5) Whenever an observation is marked as anomalous, the measurement update step is bypassed.

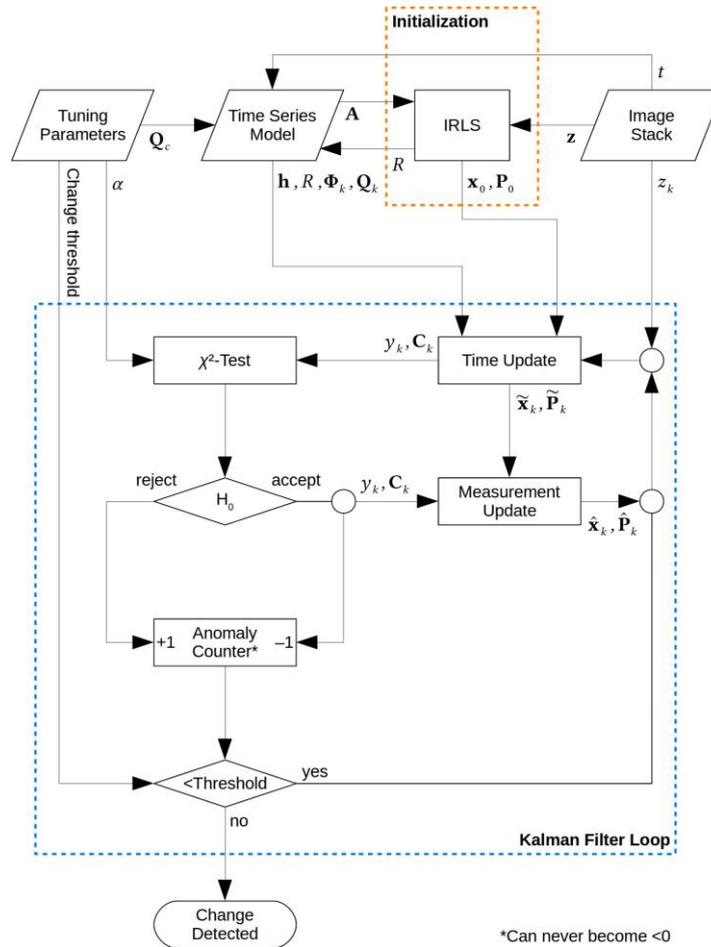


Figure 2-3: Workflow of the Kalman filtering procedure including tuning parameters

### 2.1.2 Benchmarking criteria

The quality of the delineated forest change areas is assessed in an accuracy assessment as described in Chapter 2.1.3.3. In addition, several other criteria were used to evaluate the trade-off between optimal results and suitable effort of the different experiments, which can be summarized as cost criteria. These criteria include the estimated processing time and advantages or disadvantages specific to the sensor data.

### 2.1.3 Implementation and results of benchmarking

In the following chapters, the implementation of the benchmarking process is described. An overview of the classification reference samples and predictor features (Chapter 2.1.3.1 and Chapter 2.1.3.2), the results of the forest change detection (Chapter 2.1.3.3) and the benchmarking evaluation of the different input data configurations (Chapter 2.1.3.4) are presented.

#### 2.1.3.1 Classification of input data

##### PHASE 1

In phase 1, all relevant tests have been performed in the ECoLaSS North test site in Sweden, which consists of two adjacent Sentinel-2 tile footprints (tile IDs 33VVF and 33VWF) for which Sentinel-1 and Sentinel-2 data were processed. The 2017 tree cover classification, which precedes the forest change detection, uses the same temporal-spatial metrics of Sentinel-2 as described in WP33 [AD06]. The Sentinel-2 imagery was atmospherically corrected and topographically normalized using the ESA Sen2Cor

software (Louis et al., 2016). Only scenes with cloud cover lower than 50% were used for the classification and analysis. This filter does not rely on the official metadata cloud score provided by the original Sentinel-2 Level-1C product, it is instead derived from the Level-2A product cloud mask, which was calculated as part of the data pre-processing chain using Sen2Cor. Figure 2-4 shows the Sentinel-2 scene cloud cover distribution in the test site, which is frequently affected by clouds. Figure 2-5 shows the respective data score (inverted cloud score) for each pixel in the area of interest, which is the number of available Sentinel-2 observations with average cloud cover <50% per pixel, within the full year 2017.

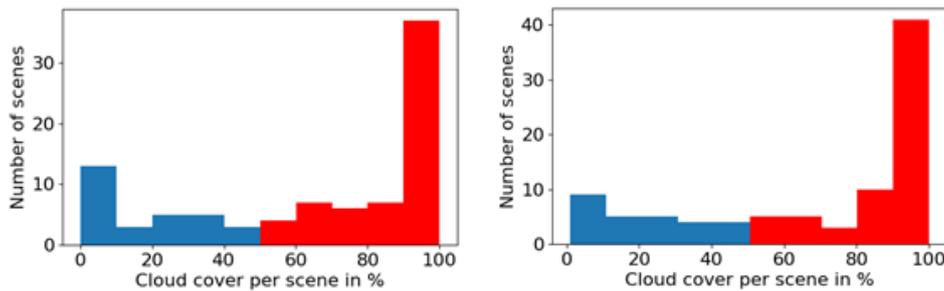


Figure 2-4: 2017 cloud coverage of S-2 tile VVF (left) and VWF tile (right). Blue: Scenes with < 50% cloud cover

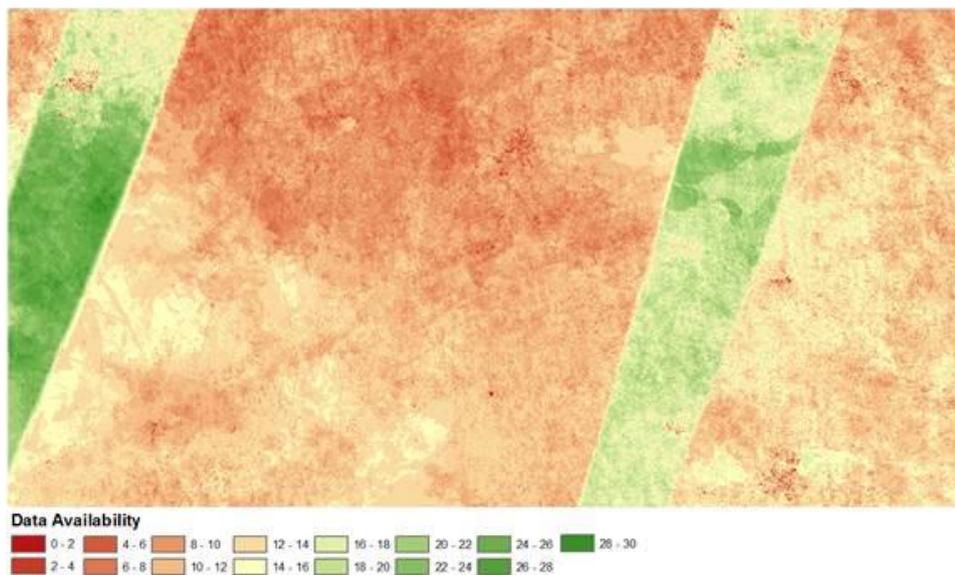


Figure 2-5: Sentinel-2 data score (number of cloud-free observations) for the ECoLaSS test site North (VVF/VWF tiles); input data: pre-filtered scenes with average cloud cover <50%, within the full year 2017

The training samples for the forest mask classification were extracted from the combined HRL2015 Dominant Leaf Type (DLT) product and the HRL2015 Grassland (GRA) product. Table 2-1 shows the distribution of sample points of the training dataset. Certain measures were undertaken to reduce the number of outliers and errors in these samples:

1. Reduction of edge effects and mixed pixels through negative buffering (60m) of the HRL DLT product classes (coniferous forest, broadleaf forest and non-forest). The remaining forest patches usually represent patches of relatively homogenous leaf type.
2. Removal of patches smaller than 1ha
3. Stratified random point sampling within the remaining coniferous/broadleaf strata
4. Removal of sampling errors through visual checks of samples
5. Iterative resampling and a visual check of samples for the broadleaf forest class to match the number of coniferous forest samples

6. Creation of rectangle polygons (corresponding to 3x3 10m pixels) from the point samples by applying a 15m buffer

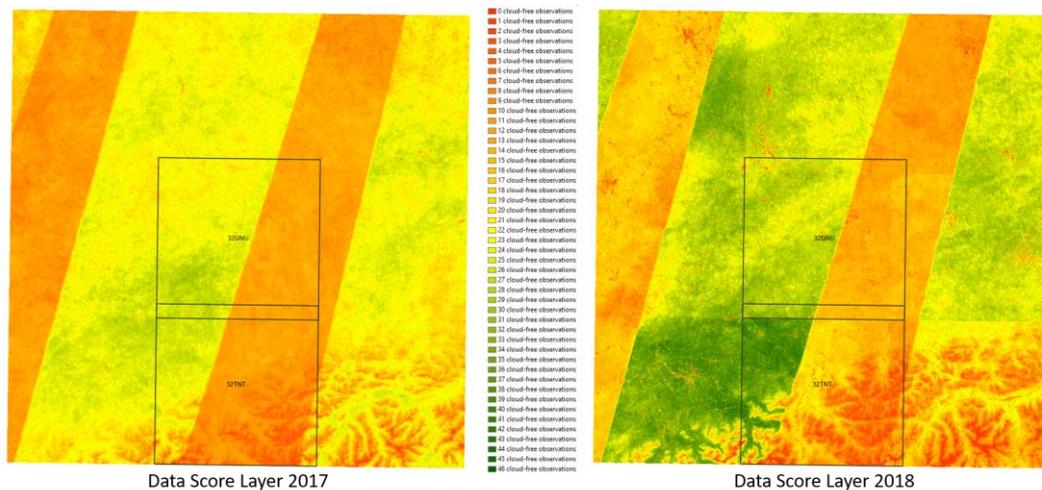
**Table 2-1: Sample distribution of training dataset for the test site North**

Class ID	Class name	Training data # polygons
0	non-forest	500
1	forest	400

## **PHASE 2**

Phase 2 could directly benefit from the experiences made in the first implementation phase as part of Task 4. The maximum allowable cloud cover threshold has been increased up to 60% and new time features have been calculated based on the Sentinel-2 spectral bands, providing additional information for the tree cover classification as input for the image-to-image change detection tests. The 2017 and 2018 tree cover masks, which precedes the forest change detection, uses the same temporal-spatial metrics of Sentinel-1 and Sentinel-2 as described in WP33 [AD06]. Tests have been performed in all three Forest test sites: The ECoLaSS North test site in Sweden (tile IDs 33VVF and 33VWF), the Central test site (tile IDs 32TNT and 32UNU) and the South-East test site (tile IDs 34TFM and 34TFL) for which Sentinel-1 and Sentinel-2 data were processed. The Sentinel-2 imagery was atmospherically corrected and topographically normalized using the ESA Sen2Cor software (Louis et al., 2016) and MACCS, as described in WP32 Deliverable [AD07]. The time features described in AD06, Chapter 3.1.4 of WP33 [AD06] were calculated for the Sentinel-2 NDVI, NDWI, Brightness and IRECI indices and all 10m/20m spectral bands. For Sentinel-1 the regular time features were calculated for gamma nought of the VV and VH polarizations and the normalized difference as well as the ratio of VV and VH.

A large amount of scenes with strong cloud cover in the time series reinforces the need for the use of image composite-like time features as described and utilized in WP33 (AD06, Chapter 3.1.4). Due to preliminary research on vegetation phenology, for both Sentinel-1 and Sentinel-2 the time features were calculated for the 2017 and 2018 complete growing season period (March 15 - to September 15), after testing the spring period (March 15 to June 15) in phase 1. This and others, in the end, are discarded as the scene availability was significantly reduced due to high and frequent cloud cover when considering shorter time windows for specific tiles, and this is likely to be the case for larger areas and cloud-prone regions, especially within Northern Europe. Figure 2-6 shows the Data Score Layer for the reference years 2017 and 2018 (vegetation period: March 15 to September 15) of the Central demonstration site, representing the largest site in ECoLaSS. As it can be clearly seen, the image situation in 2017 was less good (max. 33 cloud-free observation per pixel) than in 2018 (max. 46 cloud-free observation per pixel), which benefits from the availability of Sentinel-2B for the full observation period, whereas in 2017 Sentinel-2B was only available from the second onwards. Sentinel-2 swath borders run through the demonstration site and are well presented by the Data Score Layer. In this context, the Data Score Layer acts as an additional quality layer, which supports the sampling process by pointing to areas with generally good or worse data availability.



**Figure 2-6: Sentinel-2 Data Score Layer 2017 (left) and 2018 (right) for the Central demonstration site, overlaid by the test site**

Phase 2 tests have been performed using an improved sample base that has been generated from the HRL2015 products Dominant Leaf Type, Imperviousness, Grassland and Water at 20m spatial resolution. This so-called Sample Layer (SLA) 2015 provides five thematic classes from which training samples were automatically extracted. Similar measures as in phase 1 were undertaken to reduce the number of outliers and errors:

1. Reduction of edge effects and mixed pixels through negative buffering (20m) of the HRL2015 SLA
2. Removal of patches smaller than 1ha
3. Systematic stratified random polygon (30m x 30m )sampling within the five SLA 2015 classes, following the proportional distribution of classes within the test site, considering the general Sentinel-2 data availability by incorporation of the Sentinel-2 Data Score Layer
4. Removal of sampling errors through scatter-plot analysis based on time features
5. Iterative resampling and outlier detection

This approach pursues a generally wider representation of forest and non-forest samples for generation of the tree cover status maps as input for the change detection.

### 2.1.3.2 Change detection methodology

The fundamental properties of the change detection methodology using time features as input for an image-to-image change detection approach has not been changed in the second project phase. Consequently, the general statements from project phase 1 are still valid and presented below. For the detection of changes, results of a random forest classification based on

- exclusively Sentinel-1 time features,
- Sentinel-2 time features and
- a combination of Sentinel-1/Sentinel-2 time features

were compared to the HRL2015 Forest Mask, to derive negative changes (forest loss) between the 2015 references layer and the 2017 tree cover mask developed in ECoLaSS. For the reduction of border-effects, linear structures in the resulting difference maps were removed as well as patches (connected pixels) smaller than the HRL Forest product minimum mapping unit of 0.5ha (here: 13 pixels). Figure 2-7 shows a subset of the resulting change maps at 20m spatial resolution for Sentinel-1, Sentinel-2 and the combination of the two, where the respective unfiltered maps (before the reduction of border-effects) are shown in a lighter shade. In general, the unfiltered Sentinel-1 change map tends to show more



The above described experiences and outcomes from project phase 1 have been taken into account to improve the change detection methodology, which is depicted in Figure 2-2. The combined Sentinel-1/Sentinel-2 data scenario has been rated as the most promising one to serve the user needs in the best possible manner.

**KALMAN FILTER**

In phase 2, a test of the Kalman filter approach has been conducted using tile 32UMU in the Central area. All available Sentinel-2 images of relative orbit 108 acquired from January 2016 to the end of June 2019 have been used. Data from the overlapping second relative orbit 65 has been discarded because the different viewing angle introduces a systematic noise component in the resulting time series which should be addressed during pre-processing.

The initial state for the filter has been derived by fitting a first-order harmonic regression model to the data from 2016 to the end of June 2017. Hence all subsequent images from July 2017 onwards have been processed within the Kalman filter loop outlined in Figure 2-3. The anomaly counter threshold has been set to 3 and the level of significance of the anomaly test was 1%. For each processed spectral band, the main output of the change-detection algorithm is a map containing information about which pixels were flagged as changed and when. Analyses of the spectral change vectors corresponding to forest loss events showed that abrupt reflectance increases in the red band as well as both short wave infrared bands are reliable indicators. Hence the final change areas and per-pixel confirmation times have been aggregated by applying a 3-dimensional majority filter with 3x3x3 kernel size to the stack of red, SWIR1, and SWIR2 change maps.

**Table 2-3: Internal validation results for the Kalman Filter approach (count based approach, therefore, no confidence intervals are given in this case).**

Kalman Filter		REFERENCE			Total	User Accuracy	
		Non-Forest	No-Change	Change			
PRODUCT	Non-Forest	329	5		334	98.50	
	No Change	24	103	3	130	79.23	
	Change			41	41	100.00	
	Total	353	108	44	505		
Producer Accuracy		93.20	95.37	93.18		0.94	Overall Accuracy
						0.95	F-Score Non-Forest
						0.86	F-Score No-Change
						0.96	F-Score Change
						0.86	Kappa

An example of the derived change polygons is depicted below. The area is located west of Naislach in the northern part of the Black Forest, where pre- and post-change aerial imagery is available on Google Earth. It can be seen that some of the delineated polygons correspond to areas where complete clear cuts took place, while others show only partial removals. One of the remaining difficulties with the Kalman filter approach is the translation of simple changed/unchanged information to labels like percent-tree-cover-density-loss or similar metrics. Spectral change vectors corresponding to deforestation are highly-variable and can be influenced by many variables like

- forest type,
- biogeographic region,
- pre-change tree cover density,
- time of the change event within the season,

- undergrowth type,
- observation availability, that is the time gap between the change event and subsequent images.

As already mentioned in the state-of-the-art chapter, another source of error is given by the limited length of the available time series. We observed that this is less critical for evergreen forests, but larger errors have to be expected for deciduous woodlands where the normal dynamics of the spectral signal are more difficult to modelling approach.



Figure 2-8: Change polygons in an area close to Naislach (Central test site).

### 2.1.3.3 Forest change results

The phase 1 forest change results of all three input data scenarios (using time features of Sentinel-1, Sentinel-2 and combined Sentinel-1/Sentinel-2 data) correspond well to the forest change validation samples, but show gradual differences between the number of omission errors and miss-detections. The best overall result is offered by the combined use of Sentinel-1/Sentinel-2 time features.

In phase 2, the combined Sentinel-1/Sentinel-2 data scenario has been selected with the highest potential to address the ever-growing user requirements. The image-to-image change detection approach has been modified by integration of a subsequent NDVI plausibility step with a certain NDVI threshold applied. Based on consistently produced Tree Cover/Forest Maps at the same spatial resolution, resulting change results are provided on an annual basis with very promising accuracies.

The accuracy metrics of phase 1 are provided in the Table 2-4 to Table 2-6 below. Phase 2 metrics are presented in the Table 2-8 to Table 2-10. Figures >90% are exceeding the frequently asked accuracy threshold of 90% and are highlighted in green.

**Table 2-4: Accuracy assessment of negative forest change derived from Sentinel-1 time features**

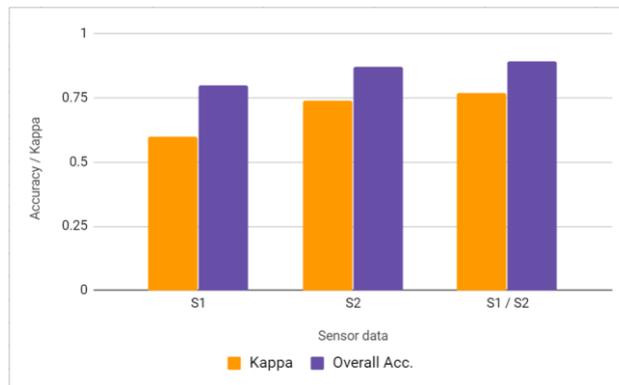
Sentinel-1 time features		REFERENCE		Total	User Accuracy	Confidence Interval
		Change	No Change			
PRODUCT	Change	63	5	68	92.65%	6.94%
	No Change	19	37	56	66.07%	13.29%
Total		82	42	124		
Producer Accuracy		76.83%	88.10%		80.65%	Overall Accuracy
Confidence Interval		9.74%	10.98%		7.36%	Confidence Interval
					0.84	F-Score Change
					0.755	F-Score No Change
					0.60	Kappa

**Table 2-5: Accuracy assessment of negative forest change derived from Sentinel-2 time features**

Sentinel-2 time features		REFERENCE		Total	User Accuracy	Confidence Interval
		Change	No Change			
PRODUCT	Change	66	0	66	100.00%	0.76%
	No Change	16	42	58	72.41%	12.36%
Total		82	42	124		
Producer Accuracy		80.49%	100.00%		87.10%	Overall Accuracy
Confidence Interval		9.19%	1.19%		6.30%	Confidence Interval
					0.89	F-Score Change
					0.84	F-Score No Change
					0.73	Kappa

**Table 2-6: Accuracy assessment of negative forest change derived from combined Sentinel-1/-2 time features**

Combined Sentinel-1 & Sentinel-2 time features		REFERENCE		Total	User Accuracy	Confidence Interval
		Change	No Change			
PRODUCT	Change	70	2	72	97.22%	4.49%
	No Change	11	39	50	78.00%	12.48%
	Total	81	41	122		
	Producer Accuracy	86.42%	95.12%		89.34%	Overall Accuracy
	Confidence Interval	8.08%	7.81%		5.89%	Confidence Interval
					0.97	F-Score Change
					0.97	F-Score No Change
					0.77	Kappa



**Figure 2-9: Comparison of the change area validation accuracy for all input data scenarios**

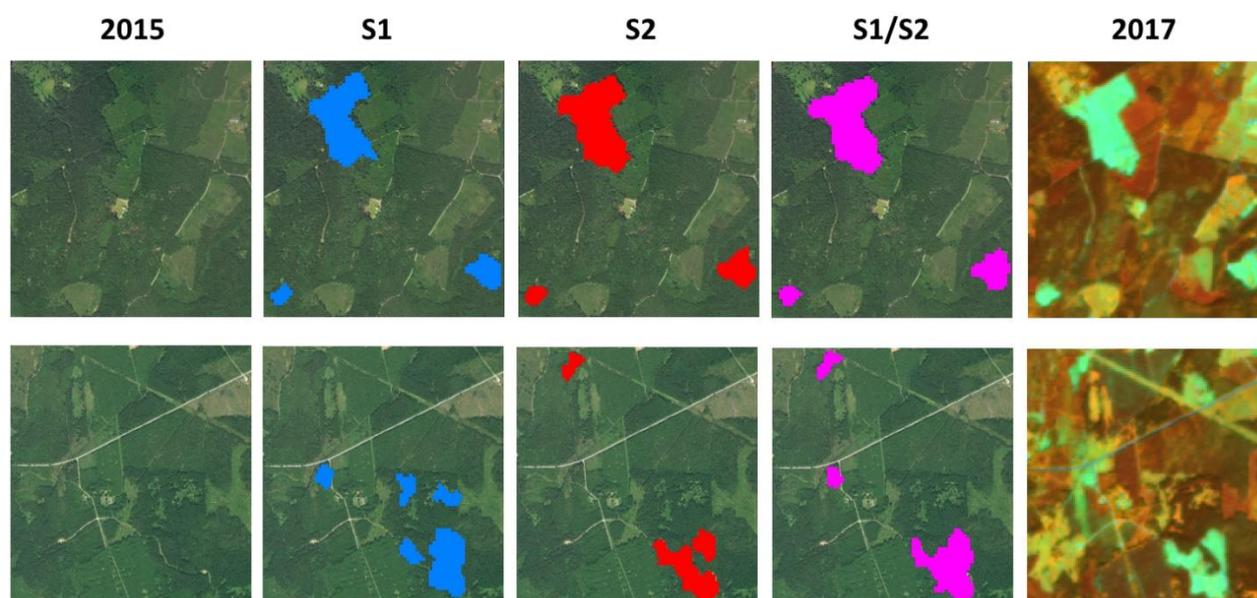
All three input data scenarios show very high user’s accuracies for the negative change areas, with 100% of the detected change patches being true positives for Sentinel-2 and 97.22% for Sentinel-1/Sentinel-2. With 92.65%, Sentinel-1 shows the lowest result, as it moderately overestimates the amount of change areas. At the same time, the Sentinel-1 method also shows the lowest producer’s accuracy for change areas, it misses roughly one out of four actual validation samples. Many of the patches delineated via Sentinel-1 classification are a bit smaller compared to the respective Sentinel-2 detections. Also, the majority of the overestimated patches are quite small. These observations are represented in **Fehler! Verweisquelle konnte nicht gefunden werden.** With 1.84%, the total percentage of detected change area within the HRL2015 forest mask is lower for Sentinel-1 compared to Sentinel-2 and combined Sentinel-1/Sentinel-2. A ~ 2% change in the forest area from 2015 to 2017 seems relatively high, but the forest in the ECoLaSS North test site is actively managed for timber production, and many typical clear-cut areas can be identified by visual interpretation. With 80.49%, the Sentinel-2 scenario shows about 4% less omission errors compared to Sentinel-1, and the best result can be attributed to the combined usage of Sentinel-1/Sentinel-2 time features, with 85.42% producer’s accuracy for the change class. The result for the combined usage of Sentinel-1/Sentinel-2 time features shows the highest overall accuracy, with little miss-detections of change, and relatively few omission errors.

This can also be observed in Figure 2-10, which shows two selected examples of change patches for the Sentinel-1 and Sentinel-2 input-data configuration. The first example (top row) shows an accurate representation of the negative change areas and has a good accordance between the different input data scenarios. The agreeing change patches in the second example (bottom row) can also be visually verified in the 2017 image. However, the additional Sentinel-1 detections, which are not detected by

Sentinel-2 only, cannot be fully verified as change areas. Sentinel-2 misses one patch that shows up in both the Sentinel-1 and Sentinel-1/Sentinel-2 change images.

**Table 2-7: Percentage of change area compared to the 2015 forest area**

	change area [%]	change area [ha]	2015 forest area [ha]
<b>Sentinel-1</b>	1.84	5,774	314,582
<b>Sentinel-2</b>	2.08	6,537	314,582
<b>Sentinel-1/Sentinel-2</b>	1.93	6,063	314,582



**Figure 2-10: Detailed view of detected negative forest change patches of different detection precision for Sentinel-1 (S1), Sentinel-2 (S2) and combined Sentinel-1 and Sentinel-2 (S1/S2) compared to a 2015 Bing Maps RGB image (left) and 2017-07-06 Sentinel-2 composite NIR/SWIR/RED (right)**

In order to further improve the change results in phase 2, the combined Sentinel-1/Sentinel-2 time features have been extended by the 10m/20m spectral bands of Sentinel-2 under consideration of a longer observation period (March to September). The thereof derived 10m tree cover masks for the reference years 2017 and 2018 have been consistently produced and already show a high overall accuracy. Error propagation of omission and commission errors could be drastically reduced, but is naturally present. With a subsequently applied NDVI plausibility step, detected changes can be further categorized into probably ‘real changes’ and ‘technical changes’. The latter one is flagged in the initial difference raster and subsequently recoded to ‘No Change’. Thus, the producer accuracy of the ‘No Change’ class, as well as the user accuracy of the ‘Change’ class, can be significantly improved. The following tables show the results for the three test sites with such a NDVI plausibility of detected changes applied.

**Table 2-8: Accuracy assessment of negative forest change derived from combined Sentinel-1/-2 time features without NDVI plausibility of changes – test site North**

Extended Sentinel-1 & Sentinel-2 time features with NDVI plausibility		REFERENCE			Total	User Accuracy	Confidence Interval
		Non-Forest	No-Change	Change			
PRODUCT	Non-Forest	476	0	0	476	100.00%	0.11%
	No Change	25	161	3	189	85.19%	5.33%
	Change	2	0	49	51	100.00%	6.31%
	Total	503	161	52	716		
	Producer Accuracy	94.63%	100.00%	94.23%		95.81%	Overall Accuracy
	Confidence Interval	2.07%	0.31%	7.30%		1.54%	Confidence Interval
					0.972	F-Score no forest	
					0.951	F-Score Change	
					0.92	F-Score No Change	
					0.91	Kappa	

**Table 2-9: Accuracy assessment of negative forest change derived from combined Sentinel-1/-2 time features without NDVI plausibility of changes – test site Central**

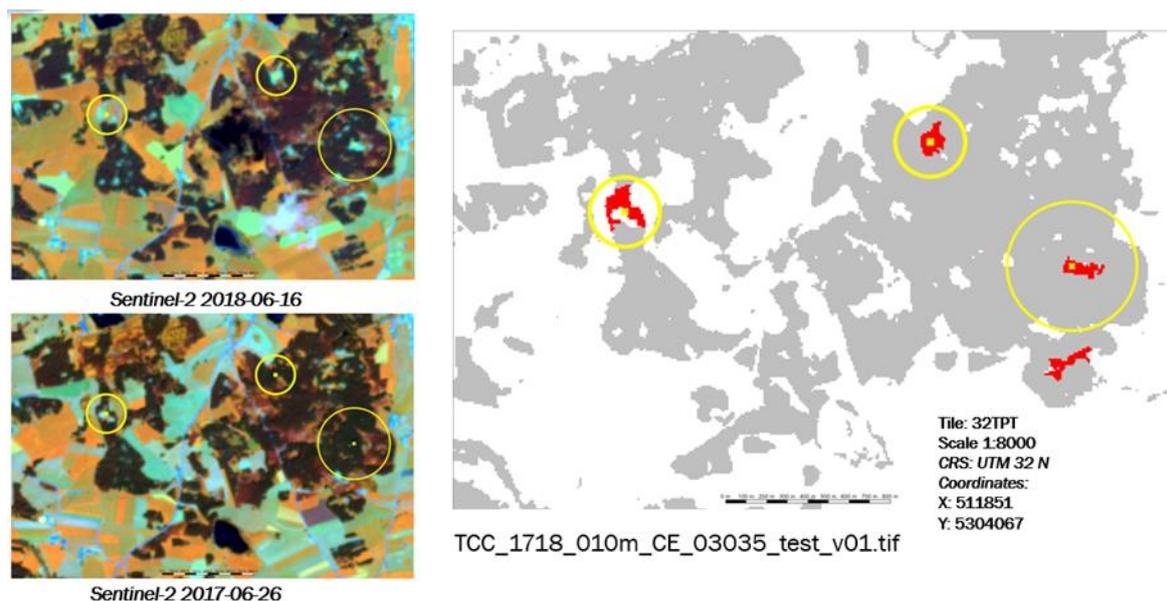
Extended Sentinel-1 & Sentinel-2 time features with NDVI plausibility		REFERENCE			Total	User Accuracy	Confidence Interval
		Non-Forest	No-Change	Change			
PRODUCT	Non-Forest	646	0	0	646	100.00%	0.08%
	No Change	29	148	1	178	83.15%	5.78%
	Change	2	1	50	53	94.34%	7.16%
	Total	677	149	51	877		
	Producer Accuracy	95.42%	99.33%	98.04%		96.24%	Overall Accuracy
	Confidence Interval	1.65%	1.65%	4.79%		1.32%	Confidence Interval
					0.976	F-Score no forest	
					0.961	F-Score Change	
					0.905	F-Score No Change	
					0.90	Kappa	

**Table 2-10: Accuracy assessment of negative forest change derived from combined Sentinel-1/-2 time features without NDVI plausibility of changes – test site South-East**

		REFERENCE			Total	User Accuracy	Confidence Interval
		Non-Forest	No-Change	Change			
<b>Extended Sentinel-1 &amp; Sentinel-2 time features with NDVI plausibility</b>	<b>PRODUCT</b>						
	<b>Non-Forest</b>	527	0	0	527	100.00%	0.109%
	<b>No Change</b>	8	155	0	163	95.09%	3.62%
	<b>Change</b>	13	4	66	83	79.52%	9.28%
	<b>Total</b>	548	159	66	773		
	<b>Producer Accuracy</b>	96.17%	97.48%	100.00%		96.77%	<b>Overall Accuracy</b>
	<b>Confidence Interval</b>	1.70%	2.75%	0.76%		1.31%	<b>Confidence Interval</b>
					0.980	<b>F-Score no forest</b>	
					0.885	<b>F-Score Change</b>	
					0.962	<b>F-Score No Change</b>	
					0.93	<b>Kappa</b>	

Validation results in all three test sites show very high overall accuracies being greater than 95% with a Kappa coefficient of 0.90 or better. In any single case, the producer accuracy is exceeding the 90%. In view of the user accuracies, some differences can be observed between the test sites. Even though the user accuracy for the 'Non-Forest' class is always at 100%, the accuracies for the 'No Change' class range between 83.15% (Central) and 85.19% (North) and 95.09% (South-East). Whereas the 'Change' class exceeds the 90% in the test sites North and Central, South-East shows the lowest user accuracy with 79.52%.

With respect to the test sites North and Central, the relatively low user accuracies of the 'No Change' class can be explained by the lower thematic accuracy of the 2017 input data mask. Here, the overestimation of tree cover has an adverse effect on the change result. Figure 2-11 shows exemplarily an area of detected forest loss in the Central test site.



**Figure 2-11: Example of detected Forest loss in Tile 32TPT– Test Site Austria/Germany.**

Regarding the South-East test site, the inadequate cloud-masking performed by the MACCS processor has been identified for the primary source of errors resulting in the relatively low user accuracy of 79.52% for the ‘Change’ class. This is due to a very high proportion of artefacts in the calculated time features representing the input for the thematic classification as well as for the subsequent plausibility. These artefacts hamper the plausibility approach, as the 2018 NDVI time features show largely spread artefacts with clear edge effects and even nodata gaps in every single tile.

#### 2.1.3.4 Benchmarking results

Table 2-11 gives a summary of how classification results (accuracy) relate to processing costs. Furthermore, scenario-specific chances and issues are listed. This type of benchmarking has been firstly made in phase 1 within Task 3 (compare Issue 1 of the document at hand) but had to be revised after the experiences and lessons learned from the first prototypic demonstration in Task 4, as the processing costs have been generally underestimated. Therefore, the benchmark criteria have been extended to the item storage cost. Regarding the overall processing costs, the following principles are valid:

- a) The longer the observation period the more data is needed and the higher the processing costs.
- b) The more time features are calculated, the higher the storage costs.

Consequently, the overall processing costs are dependent on the length of the time series (observation period) and the overall number of time features (no. time features) to be calculated. The latter significantly increases the storage costs, being part of the overall processing costs.

The Sentinel-2 scenario and combined Sentinel-1/Sentinel-2 scenario both reach rather high accuracy values with a Kappa coefficient of 0.74 and 0.77, respectively, whereas the Sentinel-1 scenario reaches a considerably lower Kappa coefficient of 0.60.

Originating from the time feature set described in [AD06], the processing costs for Sentinel-1 are generally lower as for Sentinel-2. From this perspective, the Sentinel-1 scenario might appear favourable when minimizing processing costs is an important objective, or if cloud cover is frequently very high over the region of interest.

The Sentinel-2 scenario with its high accuracy offers a good cost-benefit relation. In case of the test site North in Sweden, the problem of clouds and cloud shadows in the imagery is solved by the use of time features for a 3-month period (spring period): for 100% of the forested areas in the test site a classification decision could be made. This, however, is always dependent on cloud cover situations

specific to a region. This could be directly shown in the implementation phase of the forest prototypes in the demonstration site North: the spring period was sufficient for the southern two S-2 tiles (representing the test site), but was not in the four northern tiles, where a much more frequent cloud cover could be observed. A similar situation was observed in the test site South-East. Whereas 2017 could be covered fully cloud-free (even without time features) the situation in 2018 was extremely difficult and led to a variety of artefacts in the derived time features, which have been notably already extended to a 6-month period.

Consequently, the addition of Sentinel-1 data to the data scenario is recommended when cloud cover is a big issue in the area of interest. This solution also provides an additional benefit: The combination of Sentinel-1 and Sentinel-2 input data further increases the accuracy of the change detection results. On the other hand, increasing costs are a direct consequence.

No issues in the transferability of the image-to-image approach to the other Forest test sites, representing different biogeographical regions could be observed.

**Table 2-11: Benchmarking criteria, chances, and issues of the different input data scenarios**

	Accuracy (Kappa)	Processing cost	Storage cost	Chances	Issues
<b>Sentinel-1</b>	0.60	+	+	Independent from cloud cover	SAR inherent properties (foreshortening, layover in strong relief, speckle)
<b>Sentinel-2</b>	0.74	++	+++	Dependent on cloud cover, but time features mitigate problematic areas	Clouds/cloud shadows, artefacts, nodata gaps
<b>Sentinel-1 Sentinel-2</b>	0.77	+++	++++	Partially dependent on cloud cover, but SAR and Sentinel-2 time features mitigate problematic areas	Clouds/cloud shadows, artefacts, SAR inherent properties (foreshortening, layover in strong relief, speckle)
<b>Kalman</b>	0.86	++	++	Once the system is initialized, historic images are no longer required and new images are processed one-at-a-time. Despite only using optical data, the principle of change confirmation reduces the number of false alarms due to errors of the cloud mask.	Long time series required in the beginning to set up a meaningful time series model and initialize the Kalman filter state. This might fail in areas with extremely low number of clear-sky observations and/or persistent snow cover. At least 3 post-change observations are required to detect and confirm a change – this can lead to cases where changes occurring at the end of the year cannot be confirmed in the same year.

### 2.1.4 Summary and Conclusions

This work investigates the potential of Sentinel-1 and Sentinel-2 data for automated forest change detection in a frequently cloudy test site in Sweden. Method testing is done in preparation for a future improved HRL Forest change product. The delineation of the negative forest change is based on the comparison of a pre- and post-change forest mask product. The classification of the post-change mask uses spatio-temporal input features. The methodology can incorporate both Sentinel-1 and Sentinel-2 data without significant adjustments. This makes it flexible to be used in areas with frequent cloud coverage, which does not affect SAR imagery.

The detected forest change patches of all three input data scenarios (Sentinel-1, Sentinel-2 and combined Sentinel-1/Sentinel-2 data) reveal similar forest-loss areas and show high validation accuracies. The combined usage of Sentinel-1/Sentinel-2 delivers the best detection result, but naturally has the highest benchmarking cost, while the Sentinel-2 scenario gives slightly worse results with about half the cost. The limitation on Sentinel-1 leads to an overestimation of detected change patches resulting in lower user's accuracy, but overall, can still deliver a useable forest change mask. This is especially important considering the limited availability of optical imagery in areas of frequent cloud cover.

The methodology enables the delineation of a forest loss layer independent from the type of sensor input data used for both the reference and post-change forest masks. However, the success of the approach is very reliant on the quality of the reference forest mask, as over- or underestimation of forest cover can lead to respective over- or underestimation of negative forest change. Also, the current methodology is not able to determine the exact date of the change due to the use of time features. In the case of forest degradation or loss occurring during the period of the investigated time series (2017), time feature layers provide ambiguous information - especially those that deal with basic statistics of the entire period. For the detection of these specific changes, the combination of the approach with methods that focus on feature dynamics of forest change events within the respective time period may be considered, which is as such not yet part of Copernicus Land products, but is a recommended issue for future investigations.

Further research was carried out to prove the transferability of the methodology to areas of different geographic conditions. However, with the geographic limitations of the ECoLaSS sites, the transferability still needs to be validated at larger scale. Besides that, the influence of "technical" non-changes that can be caused by SAR speckle etc. could be further investigated. This is also valid for the integration of SAR coherence which could further improve the change detection accuracy.

The accuracy has been improved for all forest layers in phase 2. The method transferability has been proved feasible and from the tests in WP33 and WP35 and the corresponding applications to the larger demonstration sites, it should be stressed that the change product quality relies in turn on the respective status layers accuracy. Therefore, although for forest loss detection a one year update cycle for operational roll-out using S1/S2 features is achievable, the limitations described in WP32 and WP33 apply: multitemporal co-registration of S2 needs to be improved, a better DEM for improved topographic normalization is strongly recommended and the improvement of cloud masks coming from various processors (e.g. Sen2Cor) is nowadays an asset for Copernicus Sentinel-1/-2 applications.

## 2.2 Imperviousness Change

The aim of the WP34 is to explore and set up a methodological approach for automated change detection based on Sentinel time series. The selected approach, described in the workflow of Figure 2-12, is based on a new iterative classification approach. Indeed, the resulting IMD status layer for 2018 obtained in the WP33 is combined with the built-up mask from the HRL Imperviousness 2015 (or 2017). This step does not only reveal 2015/17-2018 built-up changes, with a more precise confidence interval, but it also detects potential omission errors in the built-up mask 2015 as well as potential commission errors of the 2018 built-up area. Thus, the procedure aims to both minimise the occurrence of change "false positives", and detecting omission/commission errors in this previous 2015 dataset and aims to differentiate between „actual changes“ and „technical changes“. For this purpose, the WP34 relies on densification and optimization reference dataset to determine the relative proportion of actual change versus all the error components.

### THE INPUT DATA:

The imperviousness change detection relies on two input data:

- The reference layer ( $t_0$ ): the HRL Imperviousness 2015, depicted in Figure 2-13;

- The new status layer ( $t_n$ ): the HRL Imperviousness 2017, deduced from the full dataset of Sentinel-2 pre-processed images, cloud-free, with all the spectral bands available, for the year 2016, as depicted in Figure 2-14.

The selection of the HRL Imperviousness 2017 follows the outcome of the WP33. Thus, the new status layer results from the classification based on the following set-up methods:

- A mono-temporal classification;
- The Support Vector Machine classifier;
- The data and classification fusion based on the Dempster-Shafer Theory, which will give the overall accuracy metric.

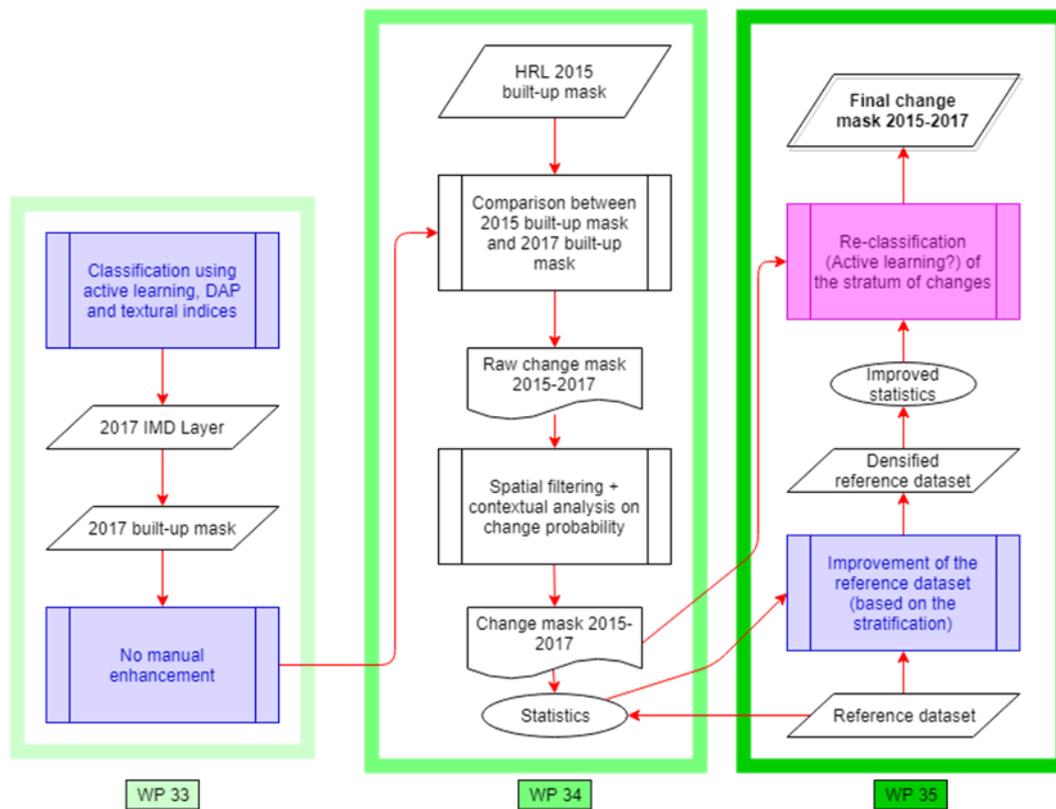


Figure 2-12: Overall workflow of the imperviousness implementation

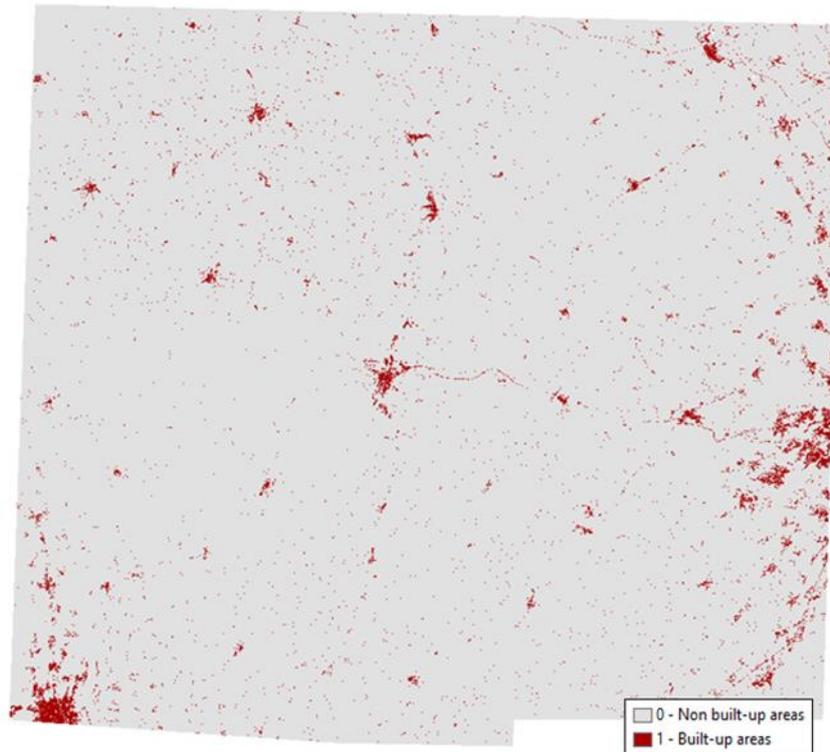


Figure 2-13: The reference layer ( $t_0$ ): the HRL Imperviousness 2015 built-up mask

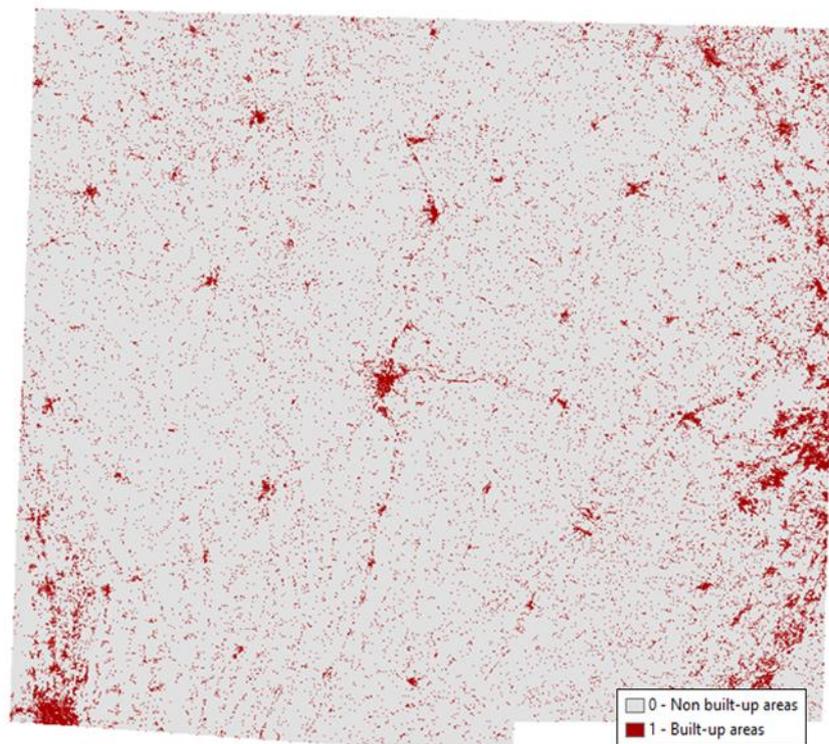


Figure 2-14: New status layer ( $t_n$ ): the HRL Imperviousness 2017 built-up mask

For the phase 2, the imperviousness change detection relies on two input data:

- The reference layer ( $t_0$ ): the HRL Imperviousness 2015,
- The new status layer ( $t_n$ ): the HRL Imperviousness 2018, deduced from the full dataset of Sentinel-1+2 pre-processed images, cloud-free, with the 10m spectral bands available, for the year 2018, as depicted in Figure 2-15.

During phase 2 implementation, the HRL Imperviousness 2018 followed the outcomes of phase 1 where a classification procedure was adopted but with some improvements in the processing chain based on additional tests. Thus, the new status layer results from the classification are based on the following set-up methods:

- A Sentinel-2 mono-temporal classification;
- A Sentinel-1 time series features classification;
- Supervised classification algorithm based on active learning and textural indices
- The data and classification fusion based on the Dempster-Shafer Theory, which will give the overall accuracy metric.

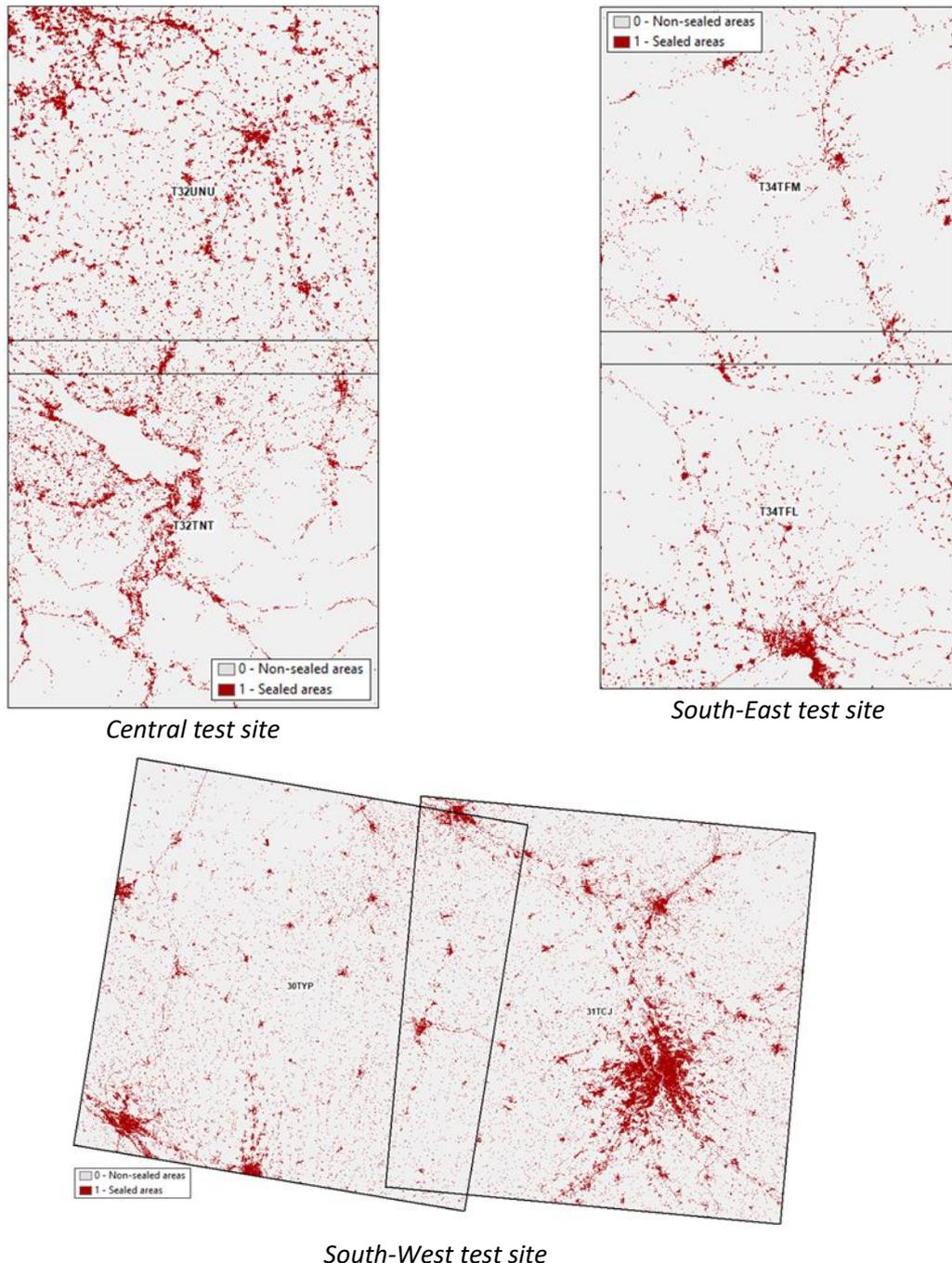


Figure 2-15: New status layer (tn): the HRL Imperviousness 2018 built-up mask

### 2.2.1 Description of candidate methods

Two main methods are mainly used to detect changes:

- Directly detect and classify the changes from a pair of images as described in (Hansen et al., 2008; Bruzzone and Serpico, 1997; Coopin et al., 2004).
- Or perform mono-temporal classifications and then compare both results to classify the changes (Lefebvre et al., 2016).

The first approach presents the advantage of having one single process and so reduces the errors accumulated through multiple classifications.

However, in the context of the updating of urban areas where classified regions in the first image are already available, the comparison with a new classification issued from the new image appears more rational. This is the process chosen for the tests.

Under the second procedure, two approaches can be presented:

- An Image to image comparison - this is the simplest way to detect changes and it is based on the simple difference between two raw classifications. However, this approach is too limited because of known problems related to differences in terms of acquisition, geometry, phenology between the two periods. That is why, studies such as (Lefebvre et al., 2016) regarding the monitoring of urban areas prefer the following method.
- The Post-Classification Comparison (PCC) – this method implies post-processing of the layers to attain:
  - o A spatial and temporal consistency based on the reference data ( $t_0$ ) – the purpose is to ensure the consistency and comparability between the different dates, so as to prevent problems mentioned in the “image-to-image approach”, related to the possible divergences in terms of acquisition, and/or of geometry between the two periods.
  - o A post-processing filtering (shown in Figure 2-16) usually used in global urban map, for example in the works of (Kemper et al., 2015) – indeed, there is a significant portion of noise due to single pixels or isolated pixels (small aggregated group of pixels), which are most likely misclassifications. Such noises should be reduced/removed with post-classification filtering approaches.

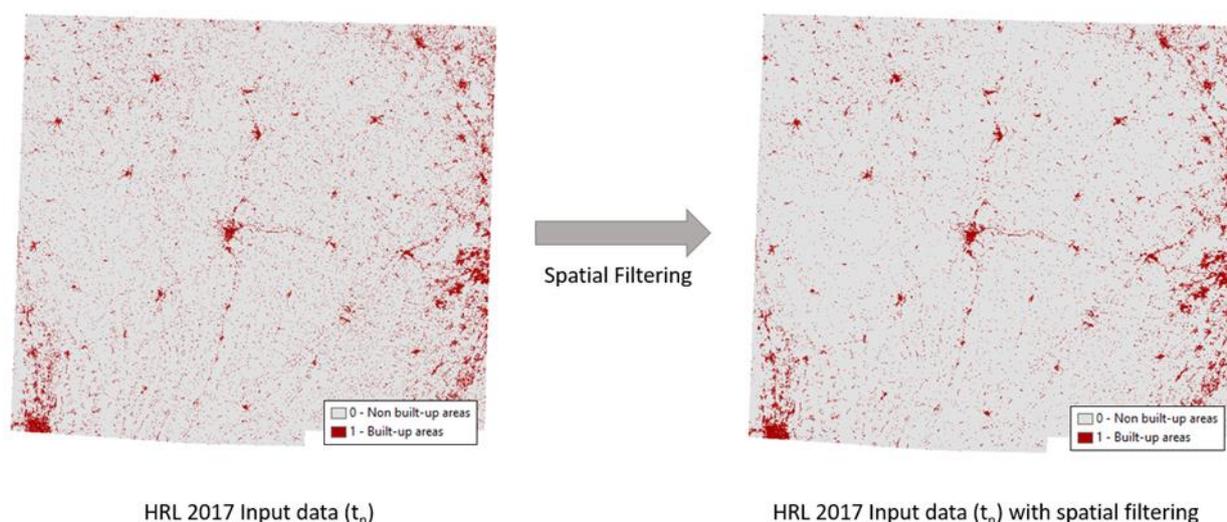
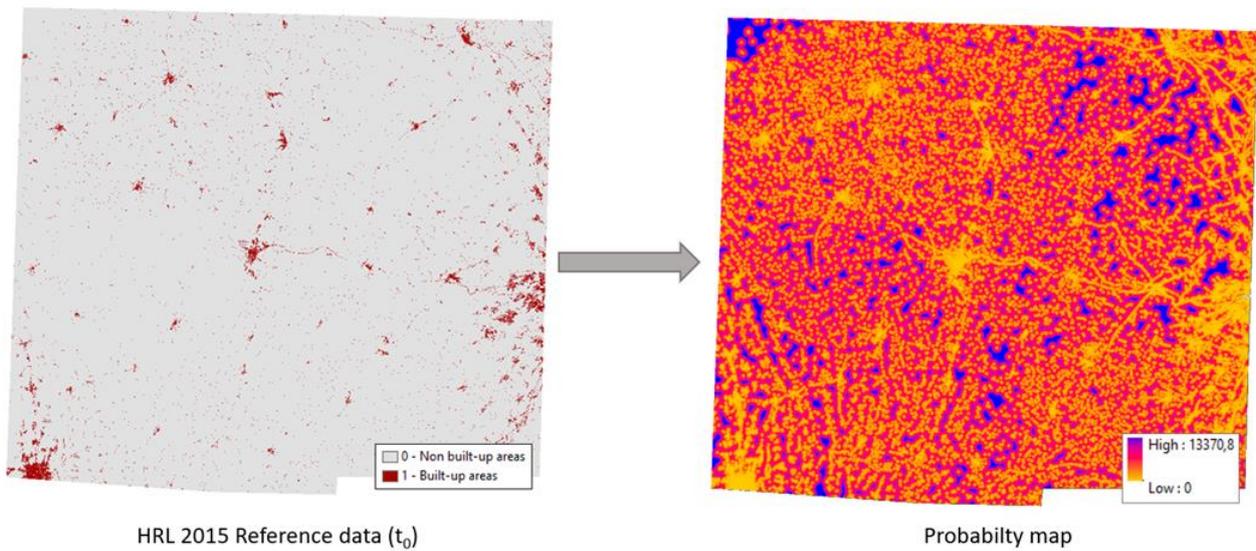


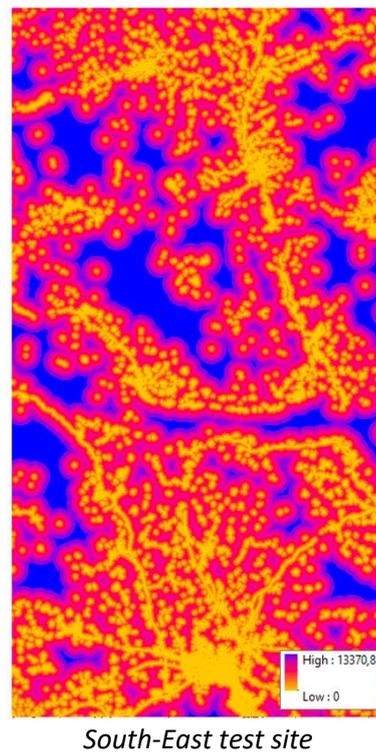
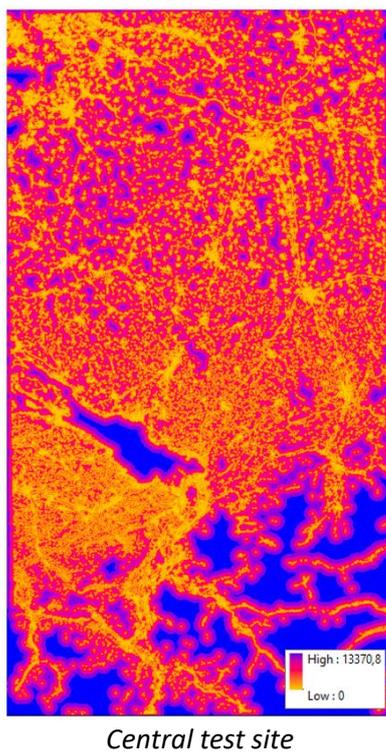
Figure 2-16: New HRL Imperviousness 2017 built-up mask post-processed with spatial filtering

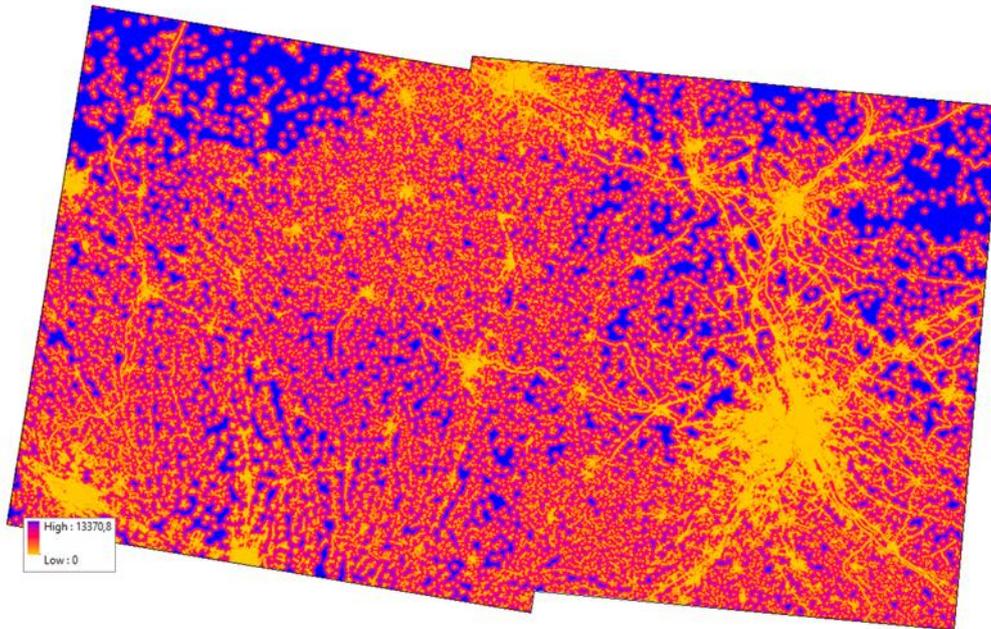
- o A contextual analysis based on change probability, such as discussed in (Lefebvre et al., 2016) and depicted in Figure 2-16 and Figure 2-17– the aim is to take into account the built-up pixels in the 2015 built-up mask in order to establish a probability map of

changes. The analysis describes each cell's relationship or membership to a source, or a set of sources based on probabilities.



**Figure 2-17: Calculation of the built-up probability map from the 2015 built-up mask for the phase 1 South-West test site**





South-West test site

Figure 2-18: Calculation of the built-up probability map from the 2015 built-up mask for the phase 2 test sites

The assumption made using the contextual analysis is that urbanized areas spread more than they appear randomly in the landscape, leading to the post-processed result displayed in Figure 2-19 and Figure 2-20.

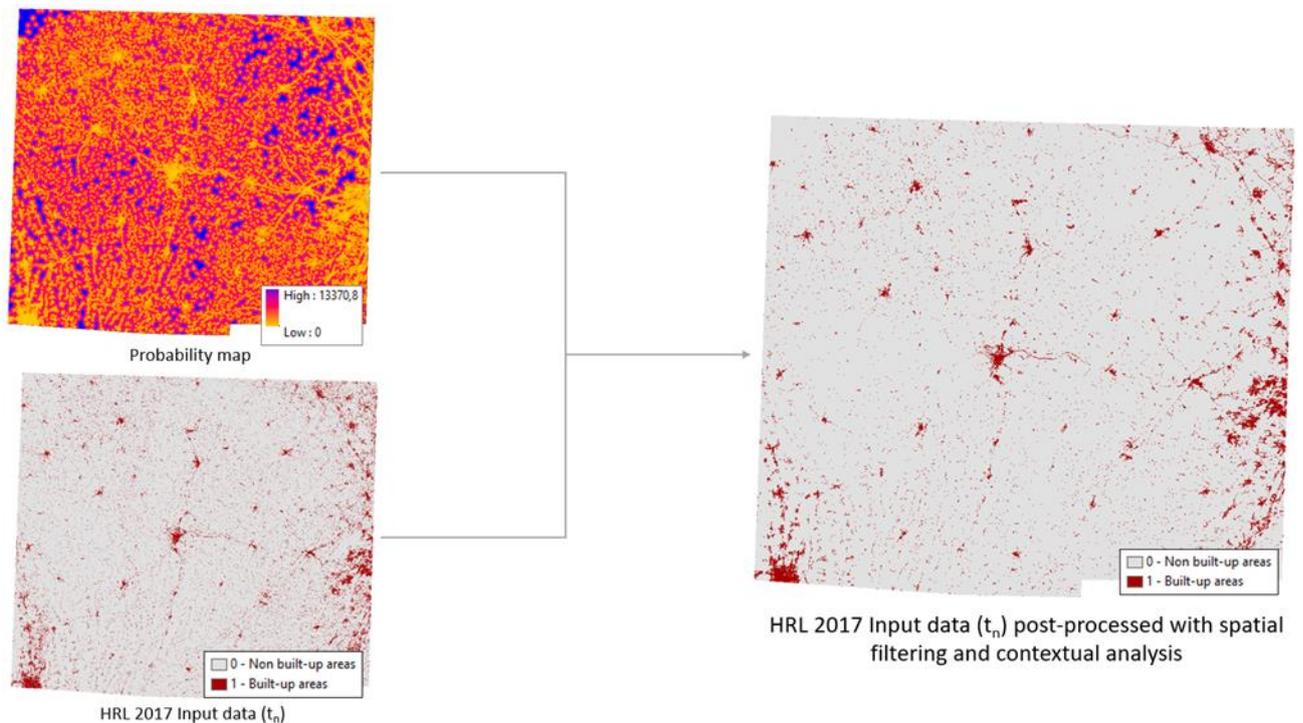
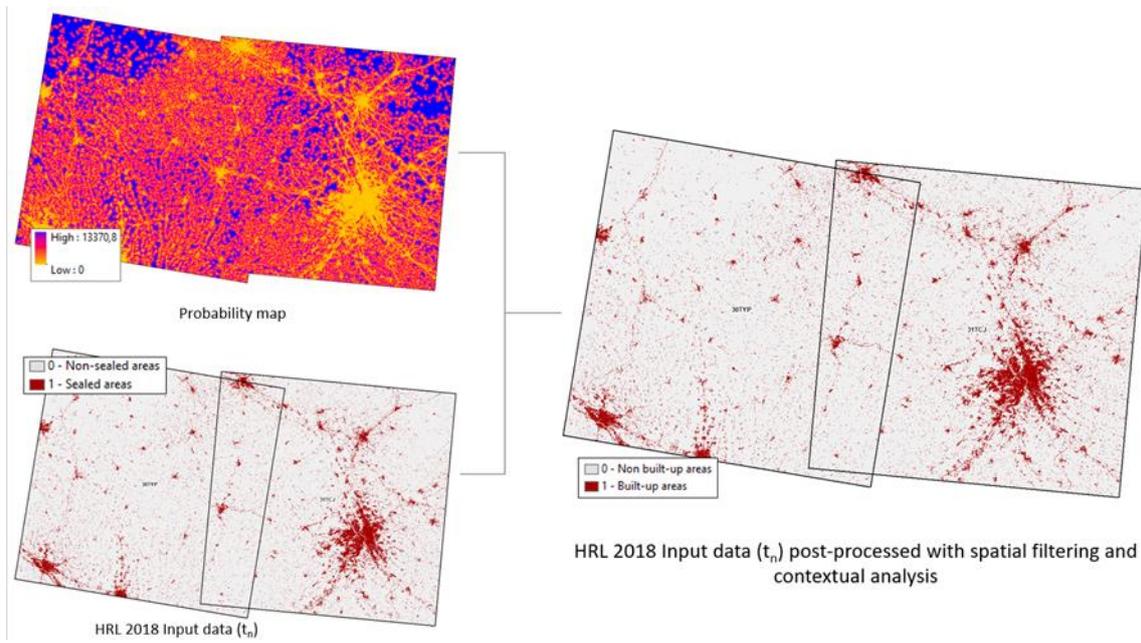
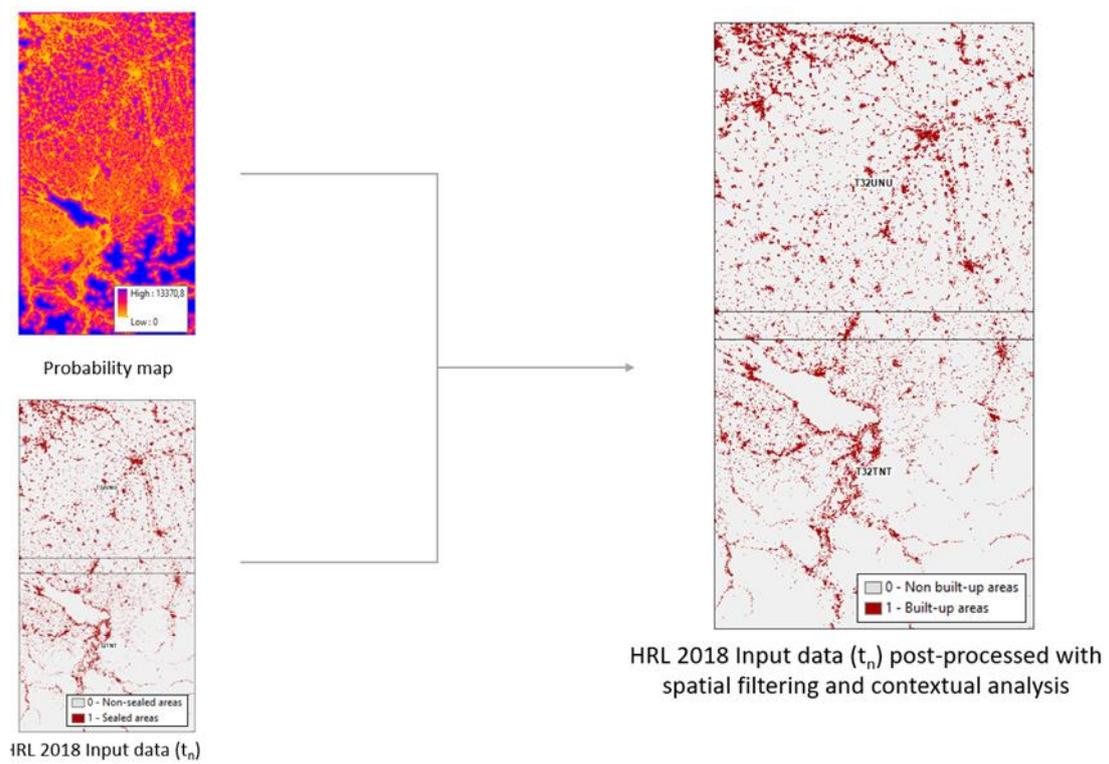


Figure 2-19: New HRL Imperviousness 2017 built-up mask post-processed with the contextual analysis

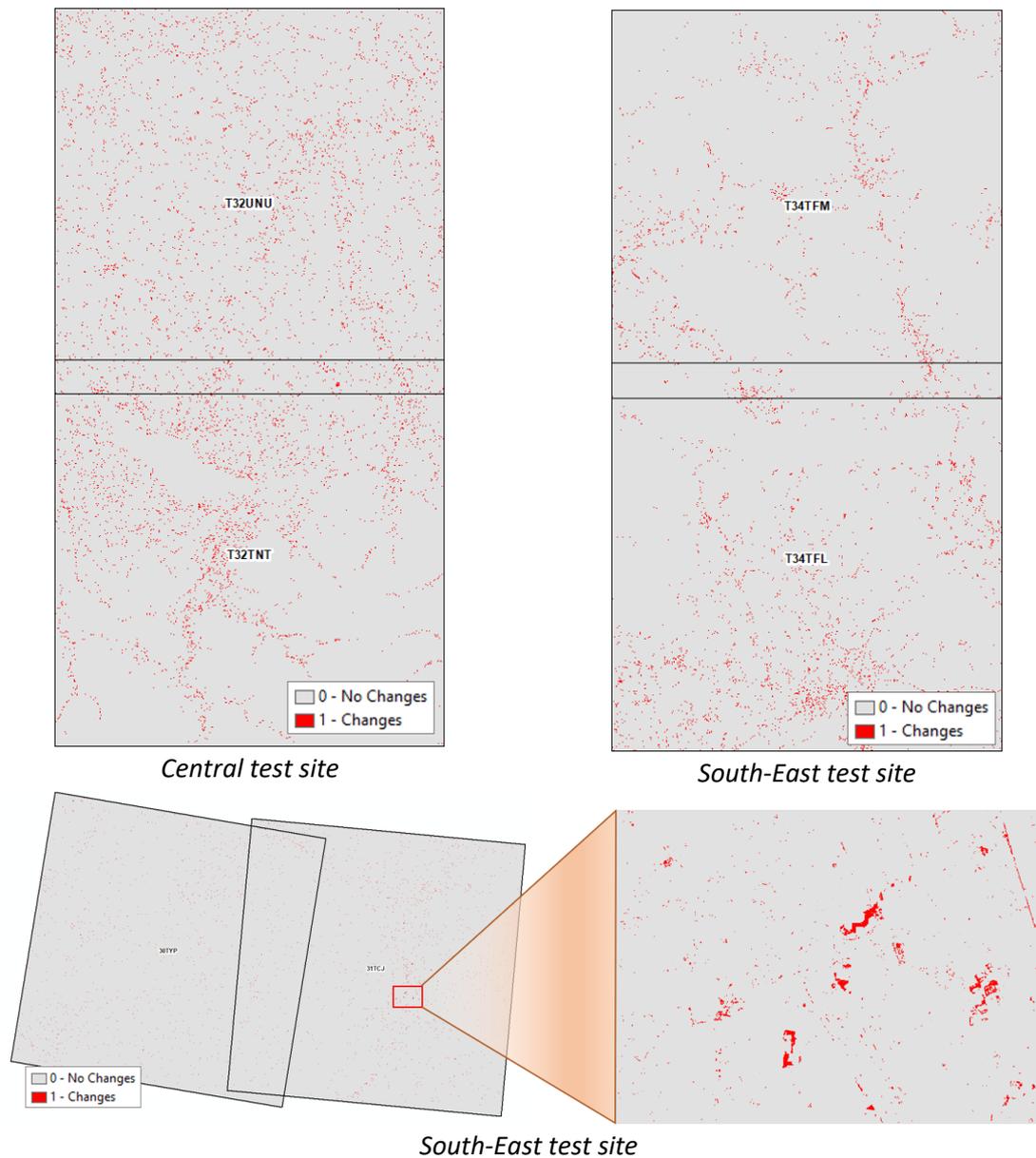


*South-West test site*



*Central test site*





**Figure 2-22: Final change layer for the temporal lapse 2015-2018 the Central and South-East test sites and 2017-2018 for the South-West test site**

Although some errors could be present in the reference layer 2015, new errors could have appeared in the detection of change between two time periods. The change layers' errors can be due to the following factors:

- Some omissions of change – new urban areas that appear between 2015 and 2017 (or 2018) were not detected;
- Some technical changes due to:
  - o Commission errors added for the new period – areas falsely flagged as new urban zones;
  - o Omission errors detected for the previous period – urban areas, already present in 2015, that were not then detected as such, but have now been flagged as urban areas in the 2017 and 2018 layers.

## 2.2.2 Benchmarking criteria

Due to the semi-automated nature of the HRL production workflow, it is not possible to guarantee that all errors can be removed from the change layer.

However, the relative magnitude of actual change versus the errors contained in the change layer for each time interval should be known to provide a basis for improving the temporal consistency between each layer – however, that work was done as part of the WP35.

Therefore, there is a strong need to develop a reference calibration dataset that will be used to determine the relative proportion of actual change versus all the error components described in the previous paragraph. To be valid, this calibration dataset has been selected based on a probability sampling approach similar, but independent, to that of the validation dataset as implemented in the report of the WP33.

The stratification and the sampling design primarily consist of selecting an appropriate sampling frame and a sampling unit. The sampling units can either be “defined on a cartographic representation of the surveyed territory” (Gallego, 2004), in which case it is an area frame, or on a list of the features. According to the author, 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, as 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 most of the cases, point samples will be used, but areas or segments may be used in specific cases such as when not only thematic accuracy needs to be reported, but also the geometry of mapped objects. Polygons have also the drawback of being specific to a single map. In case of changes, the sample may not be adapted anymore. Points are considered as the most appropriate unit for our tests.

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” as explained in (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. 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 predefined number of samples per land-cover class. As explained in (Stehman and Czaplewski, 1998), stratification ensures that each class is represented.

The validation approach chosen combines random and stratified approaches and benefits from the advantages of both of them.

For the purpose of the current tests, a stratification is applied to the change areas strata. Since the focus is on change, the approach will help to assess the new built-up areas for the year 2017 and the omission errors from 2015 (the undetected built-up pixels of 2015) as well as the commission errors from 2017 (the pixels falsely flagged as built-up in 2017).

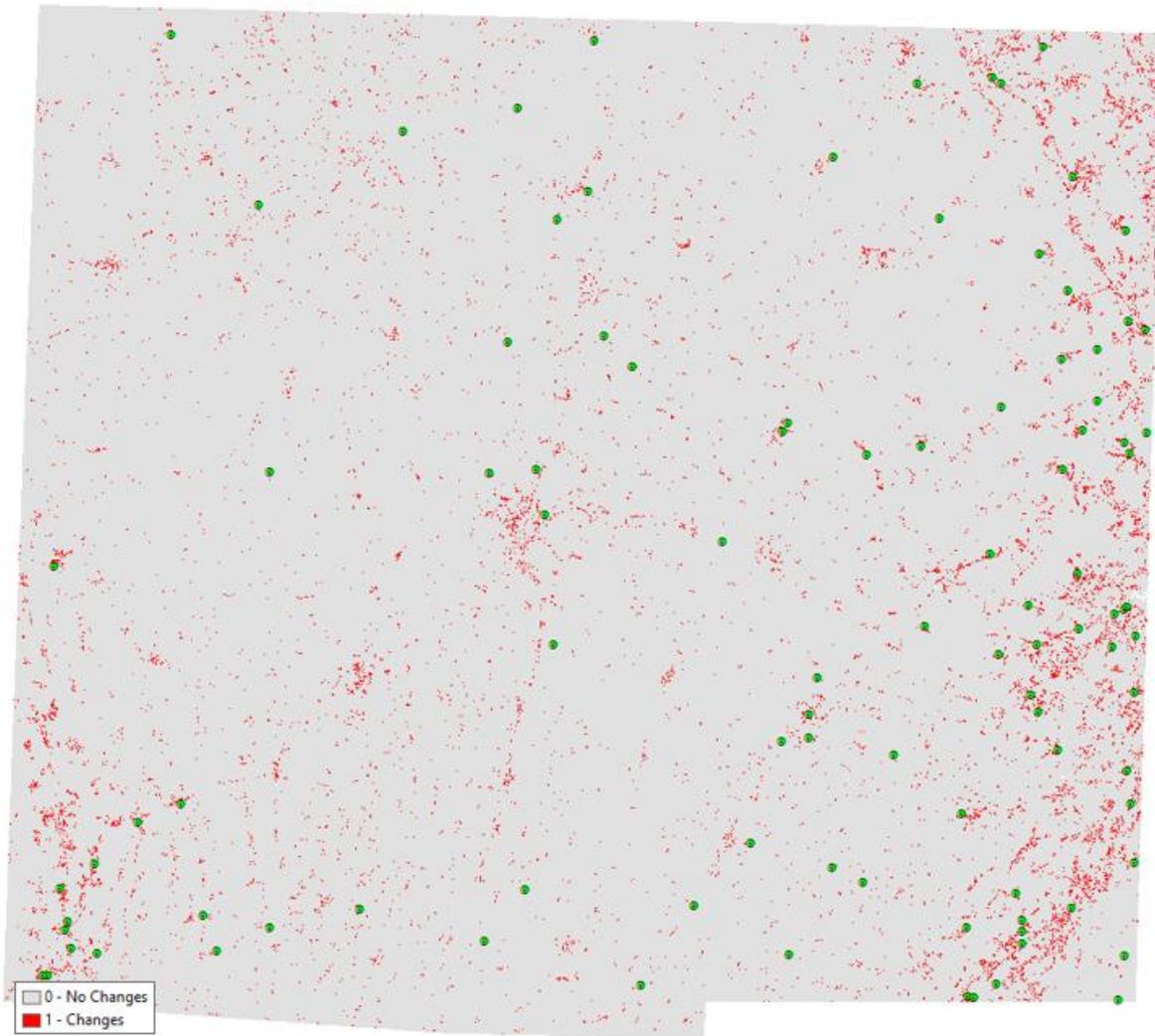


Figure 2-23: Reference calibration dataset overlaid on the change layer



Figure 2-24: 2018 Reference calibration datasets overlaid on the 2015-2018 change layer for the test sites

### 2.2.3 Implementation and results of benchmarking

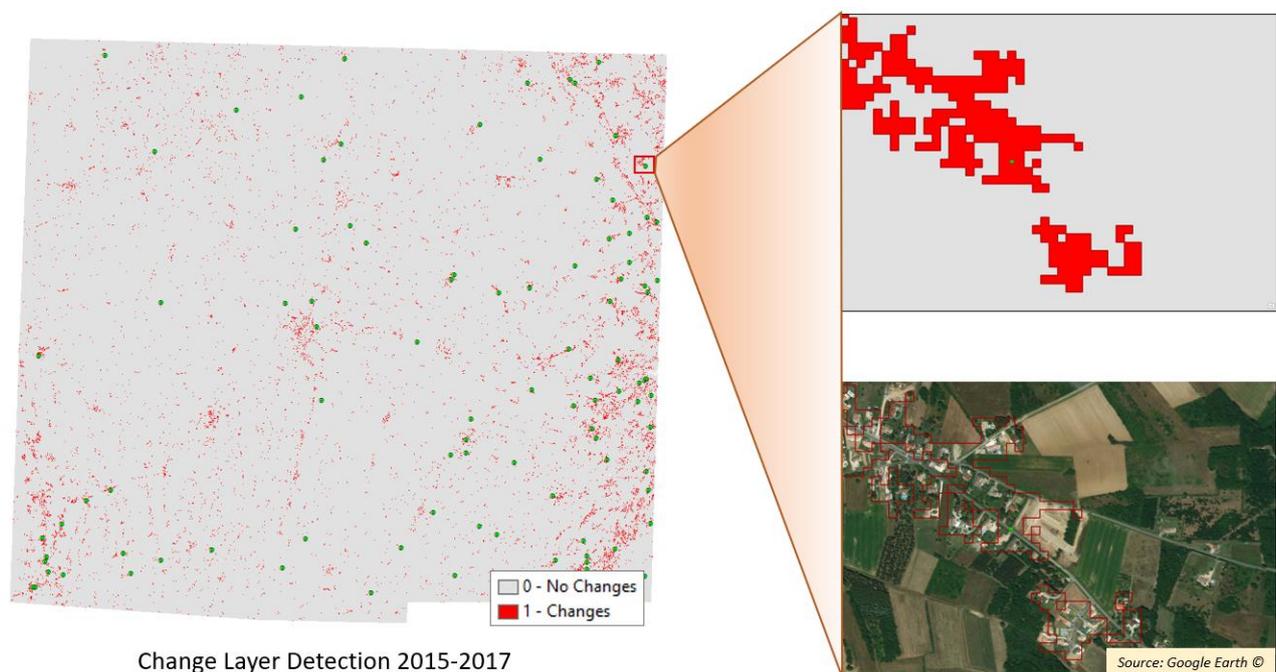
As described before, the benchmarking is only done on the best status classification for 2017 based on Sentinel-2 cloud-free images (high resolution spectral and spatial). The implementation has been done

on the test site in South-West site of France, over the tiles 30TYP and 31TCJ. The results obtained from the reference calibration dataset are presented in Table 2-12 .

**Table 2-12: Classification results for the reference year 2017**

	% of total change areas
<b>New built-up 2017</b>	9%
<b>Omission errors 2015 (undetected built-up 2015):</b>	58%
<b>Commission errors 2017 (false built-up 2017)</b>	33%

Based on the calibration dataset, the relative magnitude of actual change is estimated to 9% of the total change areas (in Figure 2-28). Thus, the errors concern the remaining 91% of the change areas detected. There is a high amount of omission errors (in Figure 2-29) coming from the reference data impacting more than half of the change areas detected.



**Figure 2-25: Example of calibration point for the newly detected built-up in 2017**

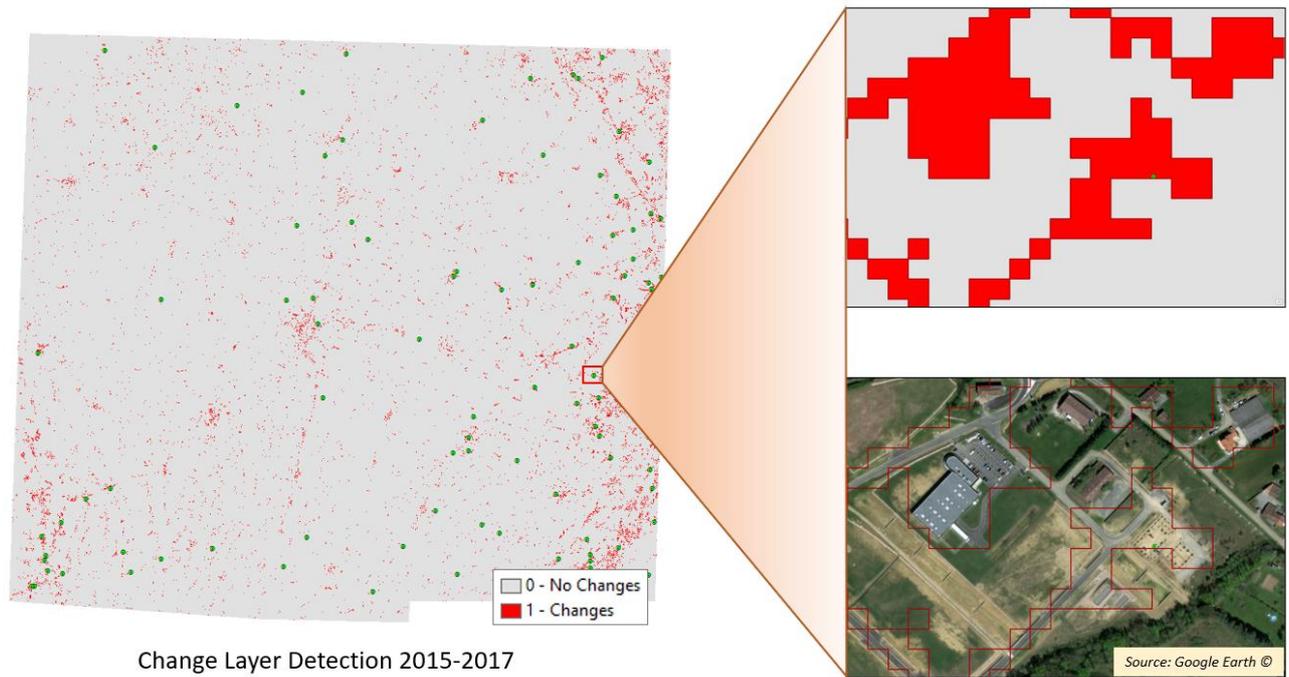


Figure 2-26: Example of omission errors 2015 (undetected built-up in 2015)

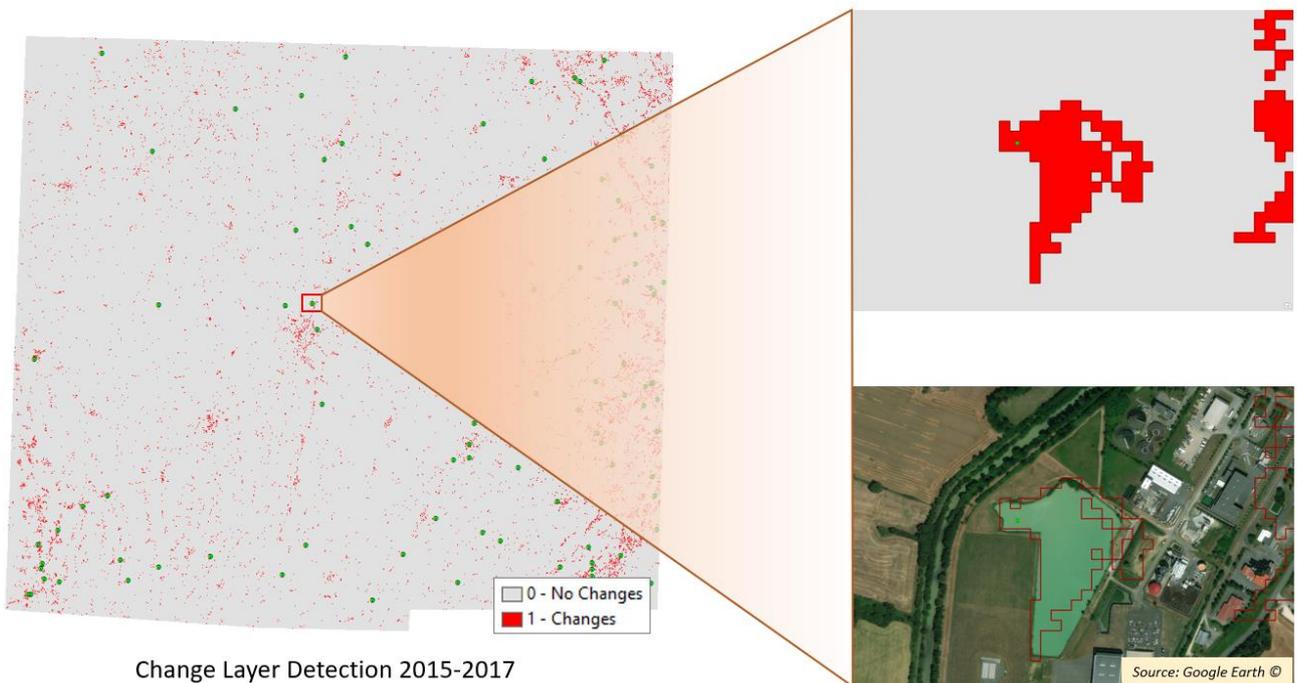


Figure 2-27: Example of commission errors for 2017, false built-up in 2017

In phase 2, the implementation of the change detection has been done on the 3 test sites South-West (tiles 30TYP and 31TCJ), Central (tiles 32UNU and 32TNT) and South-East (34TFM and 34 TFL).

The results obtained from the reference calibration dataset are presented in Table 2-13.

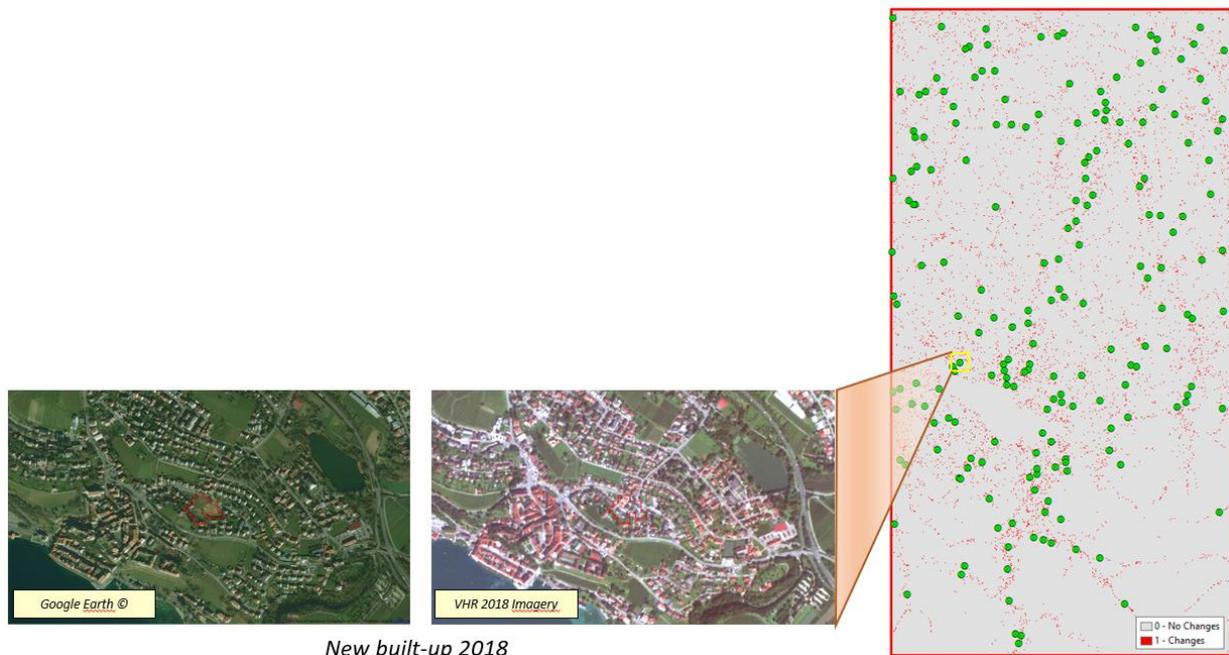
**Table 2-13: Classification results for the reference year 2017**

% of total change areas				
	South-West		Central	South-East
<b>New built-up 2018</b>	19.00%	<b>New built-up 2018</b>	14.00%	7.00%
<b>Omission errors 2017* (undetected built-up 2017):</b>	48.00%	<b>Omission errors 2015* (undetected built-up 2015):</b>	68.50%	69.50%
<b>Commission errors 2018* (false built-up 2018)</b>	33.00%	<b>Commission errors 2015* (false built-up 2015)</b>	17.50%	23.50%
<b>Total Change Area (km<sup>2</sup>)</b>	75.70	<b>Total Change Area (km<sup>2</sup>)</b>	410.25	280.85

It has to be noticed that the change detection for the South-West test site is done on the 2017-2018 time period since the HRL IMD 2017 has been created as part of phase 1; explaining the relative proportionality between the omission and commission errors (48% vs 33%) and the greater new built-up statistic (19%). The low total of change area (75.70km<sup>2</sup>) in comparison with the 2015-2017 change layer also shows the relevance of the change detection approach.

Regarding the Central and South-East test sites, the calibration dataset shows the same tendency. Thus, the relative magnitude of actual change is low (respectively 14% and 7% of the total change areas for the Central and South-East test sites). Thus, the errors concern more than the remaining 85% of the change areas detected. There is a high amount of omission errors coming from the reference data impacting almost 70% of the change areas detected.

Most of the omission errors concern small and isolated built-up features as shown in Figure 2-29. Regarding the commissions, we find usual errors like arable lands without crops, construction sites, quarries (e.g. bare soils) and permanent water bodies as shown in Figure 2-30.



**Figure 2-28: Example of calibration point for the newly detected built-up in 2018**



Omission 2015

Figure 2-29: Example of omission errors 2015 (undetected built-up in 2015)



Commission 2018

Figure 2-30: Example of commission errors for 2018, false built-up in 2018

## 2.2.4 Summary and conclusions

The results show that the magnitude of real change is around 7-14% of the change areas detected for the time period 2015-2017/18. The re-processing of the reference layer – the Imperviousness layer of 2015 – and the new status layer for 2017/2018 needs to be done as part of the WP35 [AD08] to ensure that there are no spatial inconsistencies between the layers of the different epochs. There is still a need to ensure the temporal consistency and comparability between the different time intervals. The main reasons for temporal inconsistencies can be listed as the following:

- Impervious areas typically represent less than 5% of the total area and even though the level of omission and commission are still below the set threshold of 10%. The area represented in those

10% errors in the delineation of the built-up mask can still be greater than the actual area of change between the two periods.

- Although some errors are present in the reference layer, new errors can also appear in the detection of change between two time periods. The change layers' errors can be due, as mentioned before, to the following factors:
  - o Omission of change;
  - o Commission errors added for the new period;
  - o Omission errors detected for the previous period.

In phase 1, the mono-source approach, based on one HR sensor, Sentinel-2, doesn't seem sufficient. Indeed, the times series is not dense enough to take advantage of the phenology of inter-yearly and of the intra-yearly seasonal dynamics. Further investigation has been done in the second phase, including a multi-sensor approach. The phase 2 outcomes of the 2017-2018 change detection show the relevance of the procedure with a reduction of the change areas over the time period. The matter of fact that the implementation of the change detection included this multi-sensor approach with a combination of Sentinel-1 and 2.

## 2.3 Grassland Change

To detect anomalies, which can be used to derive grassland changes dense time series are a prerequisite.

Another approach undertaken in phase 2 exclusively using Sentinel time series is of further interest to achieve harmonization and continuity between historic HRL grassland masks (i.e., HRL Grassland status layer 2015) and the improved and future grassland masks. In this sense, as explained in the case of the forest change detection methods applied, to achieve continuity a historic HRL grassland mask is set as the benchmark grassland mask. Any future observation of "non-grassland" and "grassland" within this benchmark grassland mask can be assumed to be respectively grassland loss or grassland gains (i.e., grassland change). For a detailed description of this map-to-map change detection method see Chapter 2.1.1.

One advantage of detecting changes using the map-to-map change detection method is that the method is fully detached from the input data used for map generation. A grassland pixel in the HRL Grassland status 2015 product, for example, might be derived from a combination of Sentinel-2, Landsat, IRS and SPOT data. A grassland map for the year 2018 might be a combination of Sentinel-2 and Sentinel-1. The map-to-map change detection is completely independent from the input data used for the map generation. So, the fusion of images from different optical sensors or even the fusion of optical and SAR imagery can be solved in complete isolation from the change detection. In addition, in phase 2 the status layers for the years 2017 and 2018 (grassland masks) are generated in a comparable fashion as depicted in WP33, analogously to the forest status layers, 2017-2018 change detection implementation, a different change detection method is applied to the improved 10m resolution products. In this case e.g. on the basis of a minimum NDVI derived change indicator.

An overview of the change monitoring is given in Figure 2-31. As shown in the testing within WP34, change detection approaches depend on the specific thematic application. For example, different characteristics of the temporal trajectory are to be defined for grasslands (e.g., depending on mowing dates) and forests. In the Grassland test in ECoLaSS, the main change detection approaches implemented are:

- a) Signature anomaly detection
- b) Map-to-Map comparison
- c) Map-to-Map comparison combined with signature anomaly detection



### **MULTIVARIATE ALTERATION DETECTOR COMPONENTS (MAD)**

This chapter introduces the multivariate alteration detection (MAD) transformation which is based on the established canonical correlations analysis by Nielsen in 1998. “The second canonical correlation and canonical variates are determined subject to the condition that they are orthogonal to the first ones, and this process goes on for the higher-order canonical correlations and variates. Performing differences between these pairs of variates allow then to consider a change detection analysis based on linear combinations of the original variables, ordered by correlation (similarity) between pairs” (Zanchetta & Bitelli, 2017). For N bands MAD give N output difference images, which are called MAD components. These components bear the maximum variance between the two sets of variables that are positively correlated. Therefore, each MAD represents different type of change, where similar types of spatial change will be grouped into one component as shown in Figure 2-32. One drawback is that the MAD fails to consider the overall spatial coherence of the change, is based on single-pixel analysis. In order to get the significant change a threshold can be applied to the MAD since the variates are approximately normally distributed about zero and un-correlated. Therefore a good candidate for thresholding when no ground truth is available is the standard deviation. In this sense all the pixel whose intensities are within +/- 1 sigma of zero can be considered as unchanged. Depending on the application, also the threshold of +/- 2 sigma is used to avoid overdetection of anomalies. The method provides an almost automated, time effective and reproducible technique. Therefore, it can be used as complementary procedure or as a first interpretation step in change detection studies (Zanchetta & Bitelli, 2017).

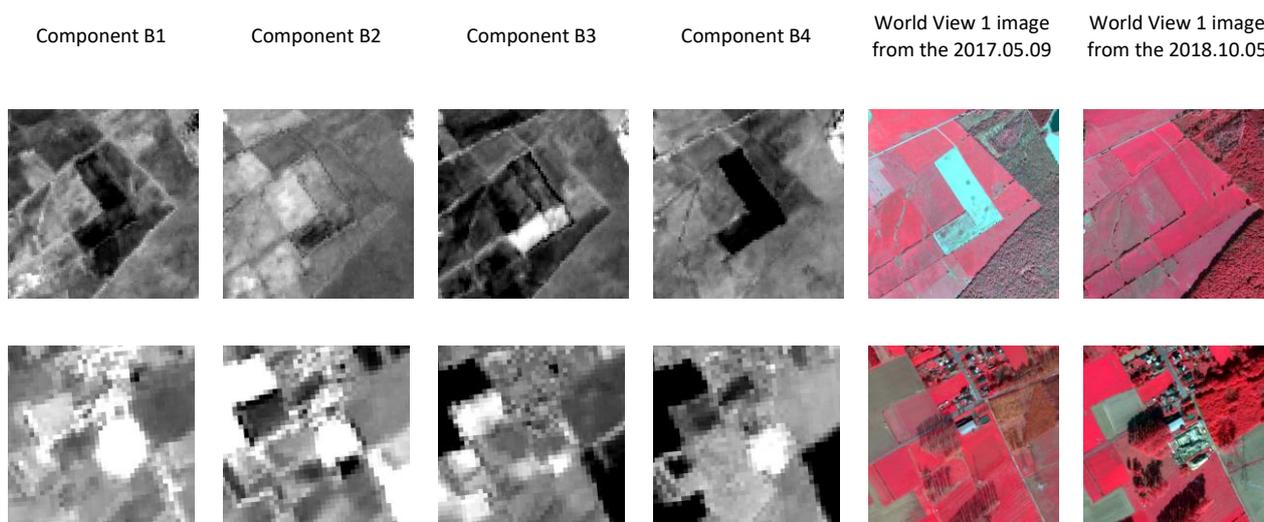


Figure 2-32: MAD components

### **CHANGE VECTOR ANALYSIS (CVA)**

The change vector is the difference between vectors at different years, with the change vector length indicative of the magnitude of change (see e.g. Bovolo & Bruzzone, 2007). As a measure of the change magnitude, the Euclidian distance measure can be applied. This approach was mainly applied for forest monitoring by bi-temporal image change detection in the past. However, for detecting changes in grasslands, an “integration” of the reflectance trajectory over the whole vegetation period is a prerequisite, as in general the mowing events differ significantly from year to year. In addition to this “integration” this leads to a reduction of image noise, which allows a pixel-based comparison of the derived annual features. The selection of temporal features was based on a qualitative assessment by a remote sensing expert. 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. Based on these criteria and the qualitative analysis, the median of NDVI, the reflectance of the red band, and the reflectance of the green band over the vegetation period were selected as features for the change vector analysis. Figure 2-33 shows the CVA results with different thresholds for gain and loss areas.

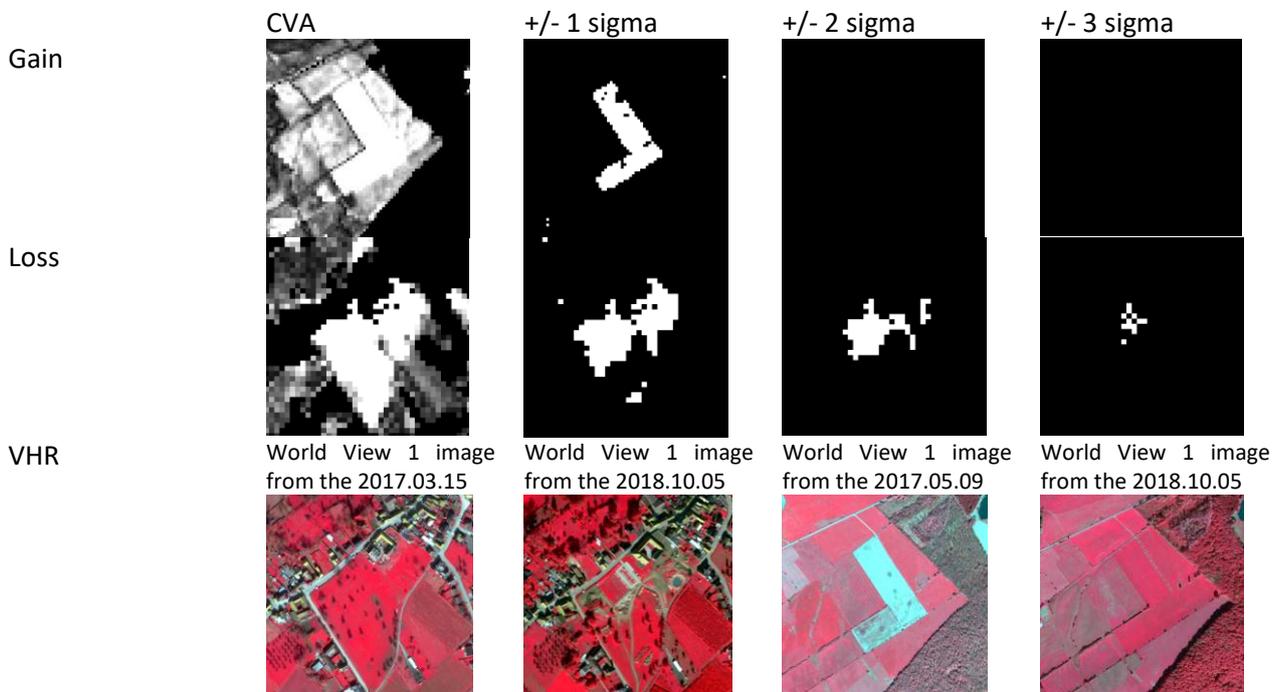


Figure 2-33: CVA components

### TIME SERIES MODELS

The underlying regression models are based on the assumption that a given time series can be described by an additive model. Each observation  $z_k$  is considered as the realization of a random variable  $Z_k$ , which is itself a sum of four components:

$$Z_k = T_k + C_k + S_k + R_k, k = 1, 2, \dots, m$$

The first two terms on the right-hand side of the equation above are called *trend* and *cycle* respectively. Both reflect non-random long-term movements in the series and while the trend may take a variety of shapes, the cycle component is explicitly defined as a recurring pattern of growth and decline. Furthermore,  $S_k$  is called the *seasonal* component and reflects non-random short term cyclical patterns which repeat themselves more or less every year. The remaining non-systematic deviations left in the series are captured by the *irregular* component  $R_k$ , which is a random variable used to model all kinds of random influences (Falk et al., 2012). Additive decomposition models *describe* the series in terms of components which cannot be observed directly. One of the advantages of this strategy is that each component of interest can be analyzed and interpreted separately. Additionally, the complexity of the model can be scaled simply by adding or removing components in order to find a model which is consistent with the data, without being unnecessarily over-specified. Once an appropriate model for a given time series has been found, it can be used to *predict* future observations.

In a regression model, the non-stochastic time series components are represented by deterministic functions of time. It is also assumed that these functions depend on specific sets of parameters which can be estimated using least squares methods. The trend component is formulated as a polynomial of degree  $n$ ,

$$T(t) = \sum_{i=0}^n c_i t^i$$

A linear trend is therefore determined by the parameters  $c_0$  and  $c_1$  representing the intercept and the slope. The seasonal component is constructed by a sum of  $P$ -periodic cosine waves, similar to the principle of a Fourier series, with  $P$  denoting the fundamental duration of the seasonal cycle. Therefore, the set  $\Omega$  containing the seasonal (angular) frequencies is

$$\Omega = \{\omega_1, \omega_2, \dots\} = \left\{ \frac{2\pi j}{P} \mid j = 1, 2, \dots, \infty \right\}$$

with  $j$  as the corresponding number of periods per cycle. Considering the nature of the time series investigated, it is appropriate to measure the time between epochs in days and thus set  $P$  to 365.25. An element  $\gamma_j$  of the seasonal component may be formulated as a cosine wave of frequency  $\omega_j$  modulated by the parameters amplitude and phase, denoted by  $A_j$  and  $\phi_j$  respectively:

$$\gamma_j(t) = A_j \cos(\omega_j t - \phi_j)$$

In this form, the phase parameter is an argument to the cosine function, resulting in a non-linear regression model. A trigonometric identity can be exploited to obtain the more desirable form

$$\gamma_j(t) = \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)$$

where the original amplitude and phase parameters are replaced by  $\alpha_j$  and  $\beta_j$ . The relationships of the new parameters to  $A_j$  and  $\phi_j$  are given below.

$$A = \sqrt{\alpha^2 + \beta^2}$$

$$\phi = \cos^{-1}\left(\frac{\alpha}{A}\right) = \sin^{-1}\left(\frac{\beta}{A}\right)$$

In conclusion, a seasonal component incorporating an arbitrary subset of frequencies taken from  $\Omega$  is given by

$$S(t) = \sum_{j \in \mathbb{N}_{>0}} \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)$$

Note that the expressions for the trend and seasonal components are formulated in continuous time, hence unequal time intervals between observations do not pose a problem.

### **ROBUST PARAMETER ESTIMATION**

The measurement equation of a linear regression model links an  $m$ -element observation vector  $\mathbf{z}$  to the  $p$ -element parameter vector  $\mathbf{x}$  containing the various coefficients  $c_i$ ,  $\alpha_j$ , and  $\beta_j$  of the time series model. Using a vector/matrix notation that is

$$\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{r}$$

where  $\mathbf{A}$  is the  $m \times p$  design matrix and  $\mathbf{r}$  is an  $m$ -element vector capturing the deviations from the model, also called residuals. Assuming  $m > p$  and normally distributed residuals with mean zero and constant variance, the parameter vector can be estimated using the method of ordinary least squares (OLS). Considering the presence of outliers in the data, this assumption may be violated and a robust parameter estimation approach following the implementation of Heiberger and Becker (1992) is

employed. The technique is known as iteratively reweighted least squares (IRLS) and belongs to the class of  $M$ -estimators (Huber, 1964). The parameter estimate  $\hat{\mathbf{x}}$  is computed based on the cost function

$$\Phi = (\mathbf{z} - \mathbf{A}\hat{\mathbf{x}})^T \mathbf{W}(\mathbf{z} - \mathbf{A}\hat{\mathbf{x}}) \rightarrow \min.$$

where  $\mathbf{W}$  is an  $m \times m$  diagonal matrix of observation weights ranging from 0 to 1. Consequently,  $\mathbf{W}$  can also be expressed in terms of an  $m$ -element vector  $\mathbf{w}$ , that is

$$\mathbf{W} = \text{diag}(\mathbf{w})$$

The parameter vector minimizing  $\Phi$  is given by

$$\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{z}$$

The initial solution of the IRLS procedure is obtained by setting  $\mathbf{W}$  equal to an identity matrix, thus it is equal to the OLS solution. In subsequent iterations, the weights are calculated based on the residuals  $r_1, r_2, \dots, r_m$ . Observations with large residual values are down-weighted. Each residual is divided by the scale  $s$  to obtain the normalized residual  $u_i$ :

$$u_i = r_i/s \text{ using } s = \text{median}(|r_1|, |r_2|, \dots)/0.6745$$

The normalized residuals are then used as input to a cost function which determines the weight of the corresponding observation. Heiberger and Becker (1992) suggest a two-stage procedure employing the Huber cost function (Huber, 1964) until convergence followed by (up to) two additional iterations applying the Bisquare cost function (Beaton and Tukey, 1974). Convergence is achieved when the difference of the Euclidean norms of  $\hat{\mathbf{x}}$  between consecutive iterations is below a given threshold. The respective elements of  $\mathbf{w}$  are determined by the equations given below. An illustration of the cost functions is given in Figure 2-34.

$$w_{i,Huber} = \begin{cases} 1 & |u_i| \leq c \\ \frac{c}{|u_i|} & |u_i| > c \end{cases} \text{ default } c = 1.345$$

$$w_{i,Bisquare} = \begin{cases} \left[1 - \left(\frac{|u_i|}{c}\right)^2\right]^2 & 0 < \frac{|u_i|}{c} \leq 1 \\ 0 & \frac{|u_i|}{c} > 1 \end{cases} \text{ default } c = 4.685$$

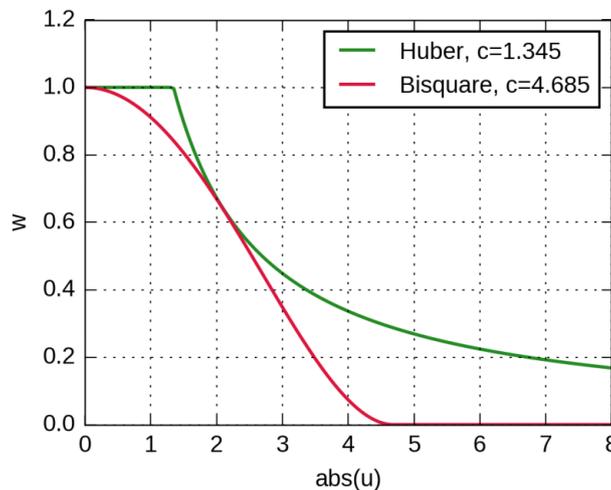


Figure 2-34: Cost functions

## Map-to-Map comparison

There are two common Image-to-Image change detection methods which can be applied to detect changes based on multi-temporal integrated image data.

The first approach uses a pair of integrated data sets based on multi-temporal features which are calculated over a certain period. The advantage of this approach is reducing the accumulated errors of the classification products.

The second approach performs classifications based on those integrated data sets including multi-temporal features and compares both results to detect the changes. Using the random forest classifier a grassland probability map for each reference year can be produced. Instead of comparing the classifications with each other, the probabilities can be combined resulting in a probability difference map differentiating between increase and decrease. To evaluate the changes, thresholds can be applied on the probability difference map derived through iterative accuracy evaluation.

In the Central site, the change detection was applied, on the one hand, between the 2017 status layer (re-sampled to 20m) and the reference year HRL2015 grassland. On the other hand, for the 2017 and 2018 status layers change detection is computed by combining a map-to-map comparison with change indicators derived from Sentinel-2 time series for 2017 and 2018.

## Map-to-Map comparison combined with signature anomaly detection

As depicted in other sections, change detection approaches can also use integrated optical and SAR images to detect signal anomalies. Furthermore, multi-temporal classifications can be derived from those integrated stacks as described in WP33 [AD06]. In this chapter those multi-temporal classifications derived from the pre-processed time-series data from WP32 [AD07] serve as input for the change detection.

### 2.3.2 Benchmarking criteria

One aspect of benchmarking is to select the spectral bands and indices whose multi-temporal signatures are represented most accurately by a time series model of sines and cosines. The ability to separate different land-use/land-cover classes based on multi-temporal signatures is also investigated in this regard. The number of parameters used in the model is a variable suitable for benchmarking. Another important criterion to assess is the influence of outliers caused by unmasked clouds, cloud shadows, and snow to the signature of a given band.

### AREA BASED THEMATIC ACCURACY

The thematic accuracy assessment can be performed by comparing the detected changes with a reference data set if available. The main purposes of the accuracy assessment and error analysis are to permit quantitative comparisons between several methods (Congalton, 1991). The thematic accuracy of the classification results can be assessed with an error matrix and following accuracy metrics:

- Overall Accuracy and Error
- User's accuracy
- Producer's accuracy
- Kappa Coefficient
- Confidence Intervals

For an internal validation of the benchmarking results a stratified random sampling approach has been used to generate validation samples within the first derived change product for the demonstration site

West. For the selected demonstration site over Belgium/France, a stratification is applied based on three strata:

- Grassland Gain
- Grassland Loss
- Grassland unchanged
- Non-Grassland unchanged

1075 samples are generated over the four strata and visually checked using combined Sentinel-2 and Landsat-8 time series and available VHR data (Google, Bing).

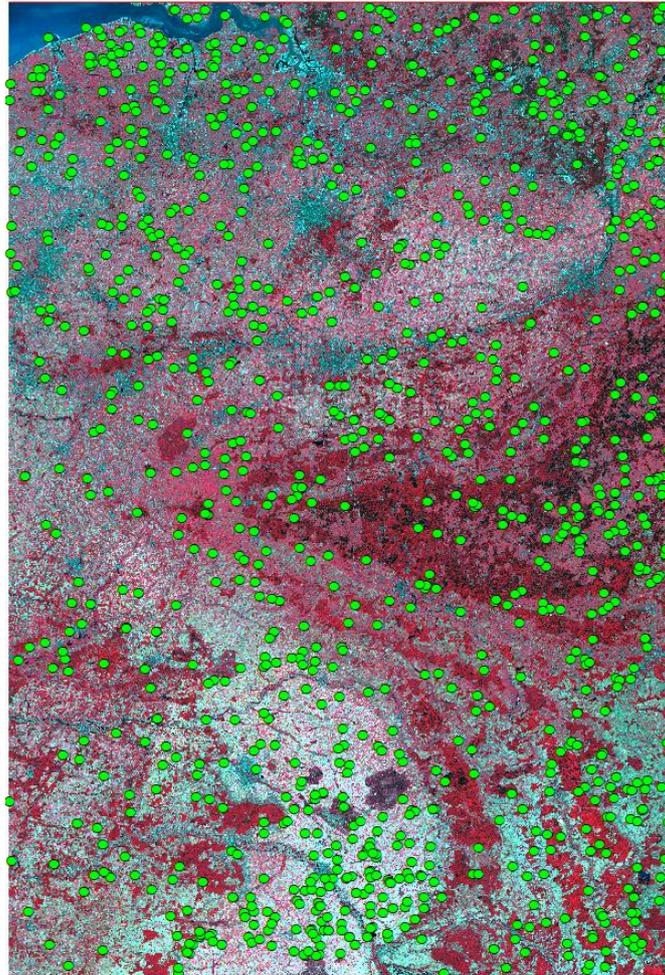


Figure 2-35: Validation samples over the demonstration site WEST

## **QUALITATIVE ASSESSMENT**

The assessment criterion is a visual interpretation of the resulting change map. In the following, some image samples show parts of the change maps compared with high-resolution data like the ArcGIS Basemap or the multi-temporal mosaics for each reference year.

### **2.3.3 Implementation and results of benchmarking**

#### **Signal anomaly detection based on dense time series**

To investigate the applicability of the regression method, a test site covering 100km<sup>2</sup> in the north of Belgium has been selected. The nature of the approach requires a large number of observations in order to capture the intra-annual variations. As Sentinel-2B data is available only from July 2017 onwards we

used data supplied by the Harmonized Landsat-Sentinel-2 (HLS) project. This project initiated by NASA combines Landsat-8 and Sentinel-2 surface reflectance data. *“In this context a “harmonized” reflectance product means that necessary radiometric, spectral, geometric, and spatial corrections have been applied to create a seamless time series, such that it is transparent to the user which sensor contributed any particular observation. It is also desirable that the harmonized products share a common gridding & tiling system, such that they are “stackable”.*” (HLS user guide v1.3, page 3) Sentinel-2 data is re-sampled to the 30m resolution of Landsat-8. Additional to the increased time series density, the HLS dataset contains cloud masks produced by the established Fmask algorithm (Zhu and Woodcock, 2012). Although a robust parameter estimation method is used, high-quality cloud masks are needed to reduce the overall number of outliers in the first place.

The temporal coverage of the test data stack is limited to the years 2015 and 2016, because the HLS project currently does not provide any data past July 2017. Illustrations of the available number of unmasked observations per pixel are given in Figure 2-36 and Figure 2-37 for the years 2015 and 2016 respectively. Due to the fact that Sentinel-2A has not been launched until the middle of 2015, the average number of observations is significantly higher in 2016. Regions with persistently few observations correspond to bright urban areas which trigger false positives in the cloud screening process.

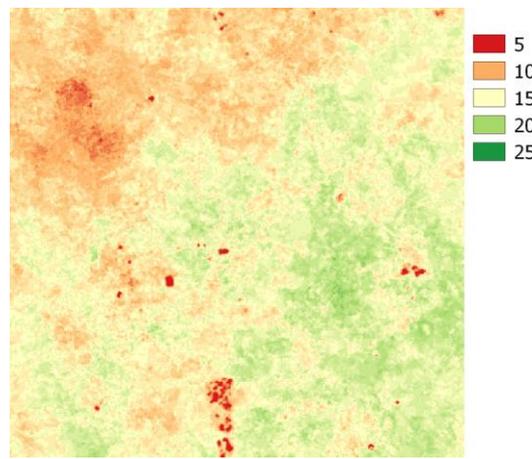


Figure 2-36: Number of unmasked observations per pixel, year 2015

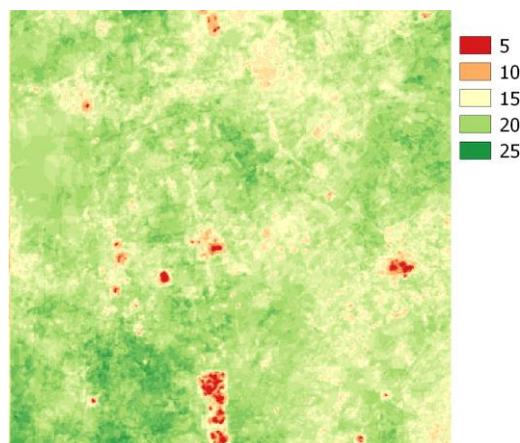


Figure 2-37: Number of unmasked observations per pixel, year 2016

**MULTITEMPORAL SPECTRAL SIGNATURE ANALYSIS**

The first step is to assess the multi-temporal signature of a grassland pixel in several spectral bands and indices. In Figure 2-38 a time series covering the year 2016 is illustrated by image chips in a true-colour band combination. Figure 2-39 shows the same series in a false colour band combination (R: SWIR1; G: NIR; B: red). The investigated pixel is located in the centre of the red rectangle displayed in each image. All images where the respective pixel is masked according to the Fmask algorithm are excluded. Nonetheless, there are still images left in the series, which are affected by anomalous conditions, see Table 2-14. Additionally, two mowing events can be identified, see Table 2-15. Note that on two occasions, namely 2016-05-01 and 2016-07-20, there are two images available for the same day.

**Table 2-14: Anomalous images**

Acquisition date	Condition
2016-01-26	Cloud shadow
2016-05-01	Cloud shadow
2016-10-08	Cloud
2016-12-02	Cloud shadow

**Table 2-15: Mowing events**

Image before	Image after
2016-05-01	2016-05-08
2016-06-07	2016-06-09

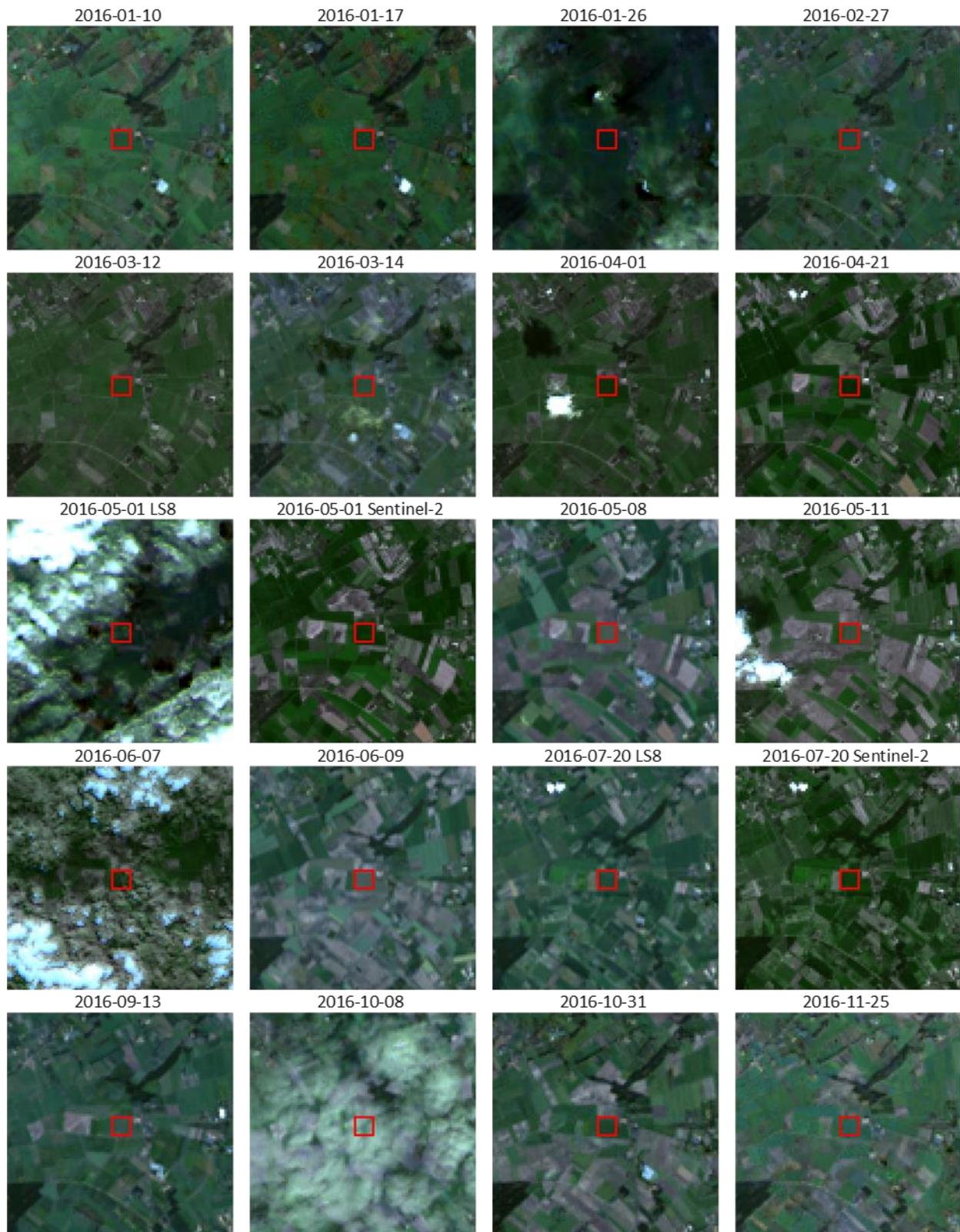


Figure 2-38: Time series of a grassland pixel located in the centre of the red rectangle (True colour)

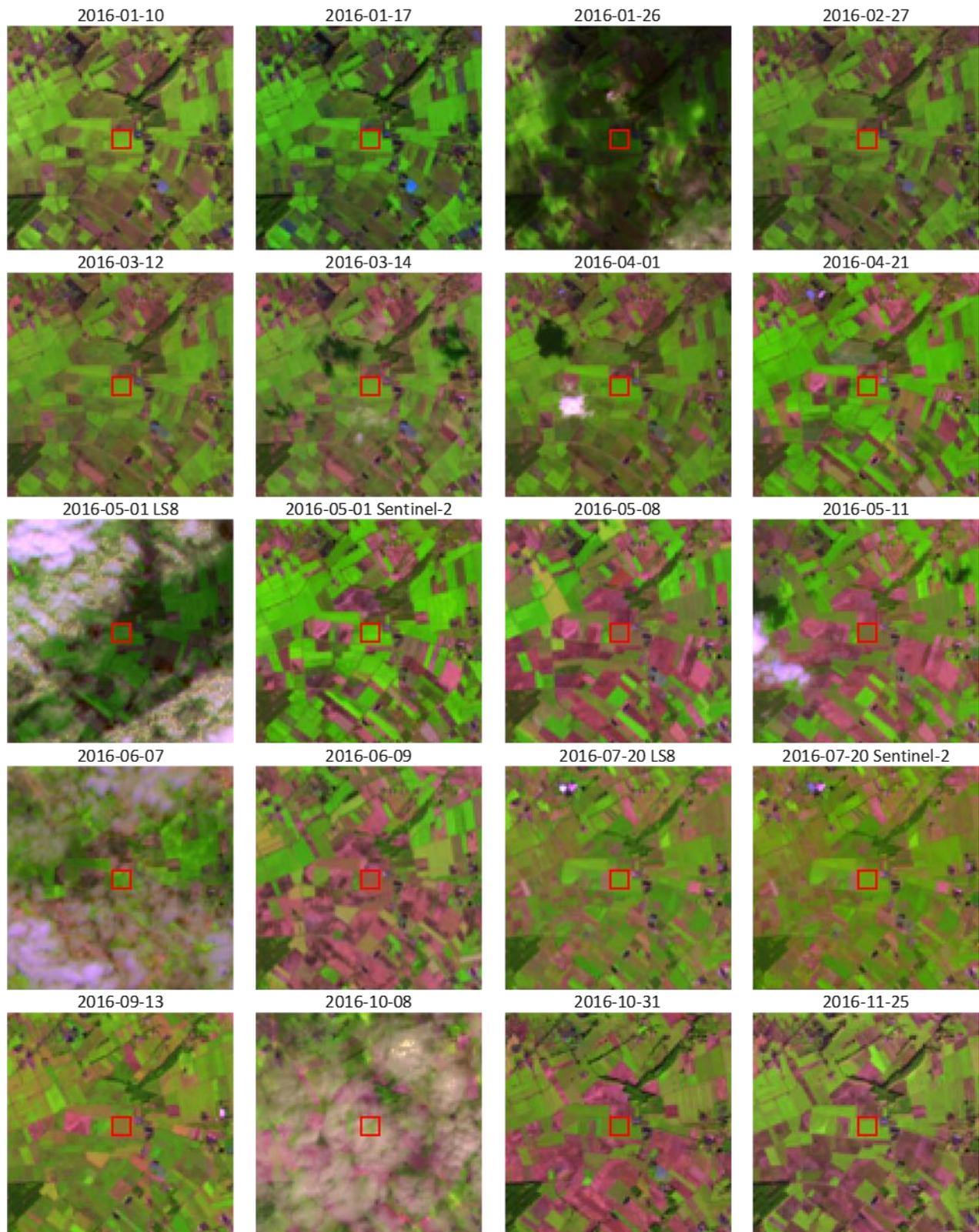


Figure 2-39: Time series of a grassland pixel located in the centre of the red rectangle (False colour)

The time series of 6 spectral bands common to Landsat-8 and Sentinel-2 plotted in Figure 2-40. Additionally, the time series of the NDVI and the Tasseled Cap components Brightness, Greenness and Wetness (Crist, 1985) are plotted in Figure 2-41. Each time series corresponds to the investigated grassland pixels with a resolution of 30x30 meters. Several observations can be made by inspecting these Figures:

- There is often little separation between outliers and neighbouring valid observations.
- Abrupt jumps caused by the mowing events are visible in the time series of all 6 spectral bands and 3 out of 4 indices.
- The only index quite insensitive to the mowing events, but sensitive with respect to the outliers, is Tasseled Cap Brightness.

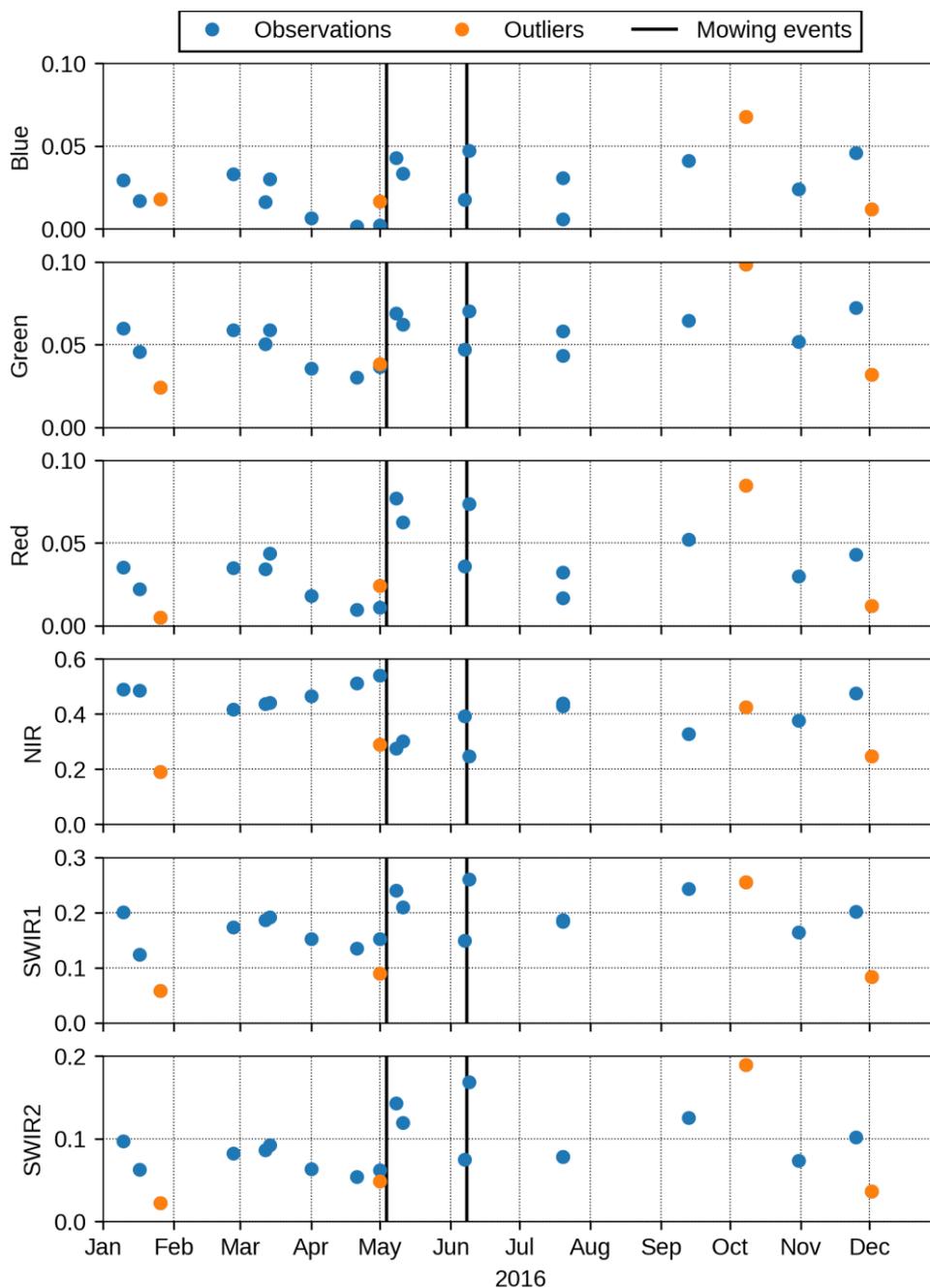


Figure 2-40: Time series of spectral bands

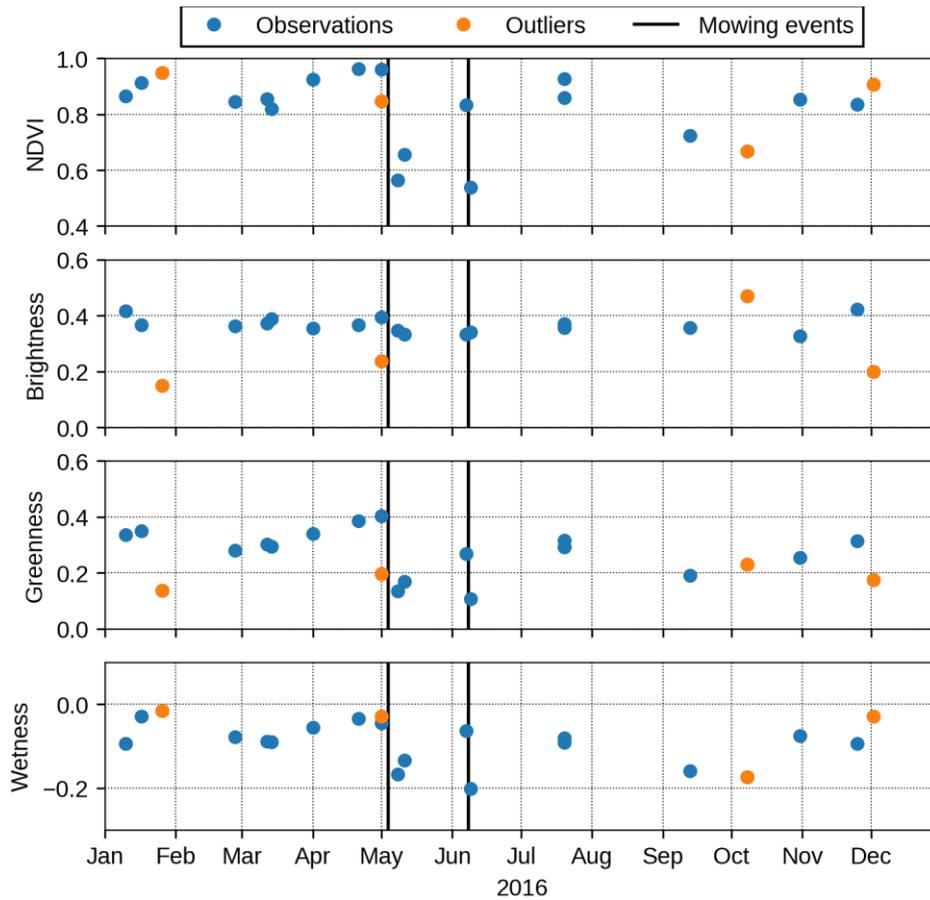


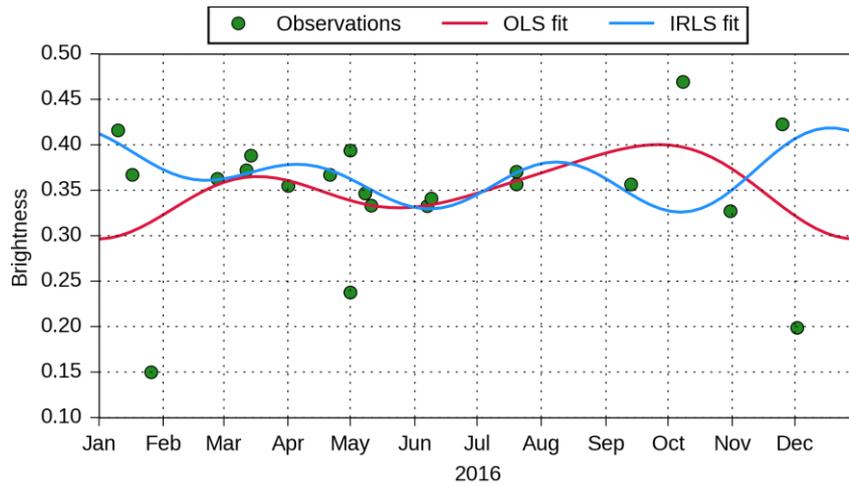
Figure 2-41: Time series of spectral indices

**APPLICATION OF THE CONCEPT**

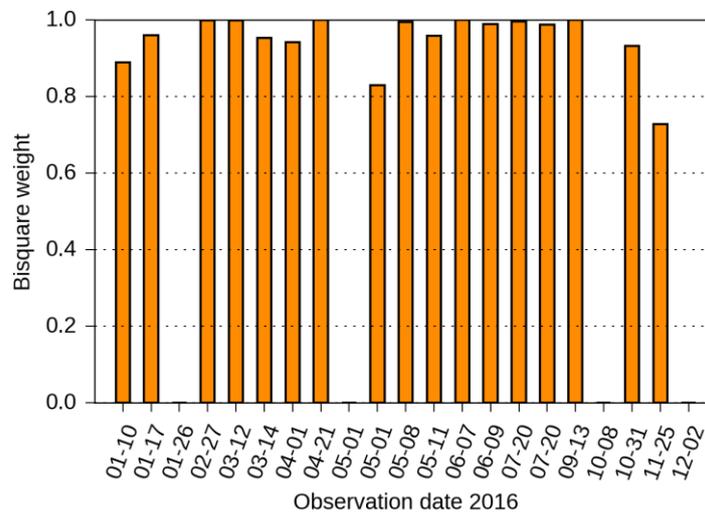
The regression model, which has been fitted to the time series of each pixel of the test site, is given below. It features a constant trend as well as a seasonal component of 3 frequencies, thus there are 7 parameters to estimate. To introduce a certain degree of over-determination, the minimum number of available unmasked observations required to carry out the IRLS procedure is set to 14.

$$z(t) = c_0 + \sum_{j=1}^3 \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)$$

The first step of the implemented test set-up is to eliminate remaining gross outliers in the data set caused by unmasked clouds, snow, and cloud shadows. It is assumed that these outliers are visible more prominently in the time series of Tasseled Cap Brightness, subsequently referred to simply as Brightness. The coefficients for the Tasseled Cap transformation were taken from (Crist, 1985). Clouds and snow are expected to cause unusually high Brightness values, whereas cloud shadows should correspond to low magnitudes of Brightness. An example of a Brightness time series for a single grassland pixel is given in Figure 2-42. Additional to the observations, the results of an OLS fit on the one hand and an IRLS fit on the other are plotted. The OLS solution is influenced by the outlying values in the series, whereas the robust IRLS fit is not. The weights assigned to each observation by the IRLS procedure are plotted in Figure 2-43 and it can be seen that 4 weights are zero. Using Figure 2-42, it can be verified that weights of zero in Figure 2-43 correspond to anomalous observations. Therefore, all observations with an assigned weight of zero are excluded from further processing steps.



**Figure 2-42: Example of a Tasseled Cap Brightness time series of a grassland pixel and fitted regression models using OLS and IRLS**



**Figure 2-43: Computed observation weights of the IRLS fit**

The second step of the implemented test set-up is to fit the regression model to the time series of Tasseled Cap Greenness, subsequently referred to simply as Greenness, which is assumed to be an appropriate spectral index to capture vegetation dynamics. Since most outliers should have been eliminated in the previous step, OLS is used to estimate the parameters. Taking the same grassland pixel as discussed above, the associated Greenness time series is plotted in Figure 2-44.

Inspecting Figure 2-44 several observations can be made:

- Two mowing events designated by abrupt jumps in the Greenness level can be clearly identified, with a possible third one indicated somewhere at the end of August.
- The time series model cannot capture the high temporal dynamics of the mowing events since it does neither account for abrupt jumps nor short growing periods of only one month.
- The Greenness level of the outliers identified as cloud shadows is similar to the level after a mowing event, thus a confusion of the two conditions is possible. This emphasizes the necessity of outlier detection.

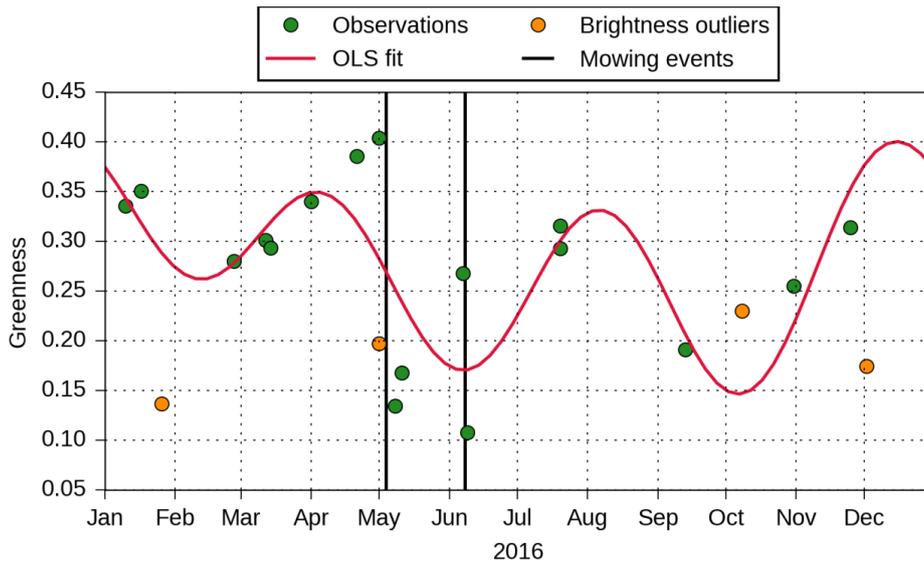


Figure 2-44: Example of Greenness time series of a grassland pixel and fitted regression models using OLS

Although the time series model cannot capture abrupt jumps, the set of estimated parameters may still be interpreted as a compressed form of the multi-temporal information contained in the image stack. While the parameters do not reflect all the variation in a time series, they might hold enough information to separate different land cover/land use classes. The idea is illustrated in Figure 2-45, where the Greenness time series of an agriculture pixel is contrasted with the time series of the previously used grassland pixel. The same workflow has been applied to obtain the OLS fit. Several differences between the fitted curves can be observed, for example:

- the value of the trend parameter  $c_0$ ,
- the composition of the seasonal pattern with respect to the amplitudes of the different frequencies, and
- the minimal value and range (difference max, min) of the fitted curve.

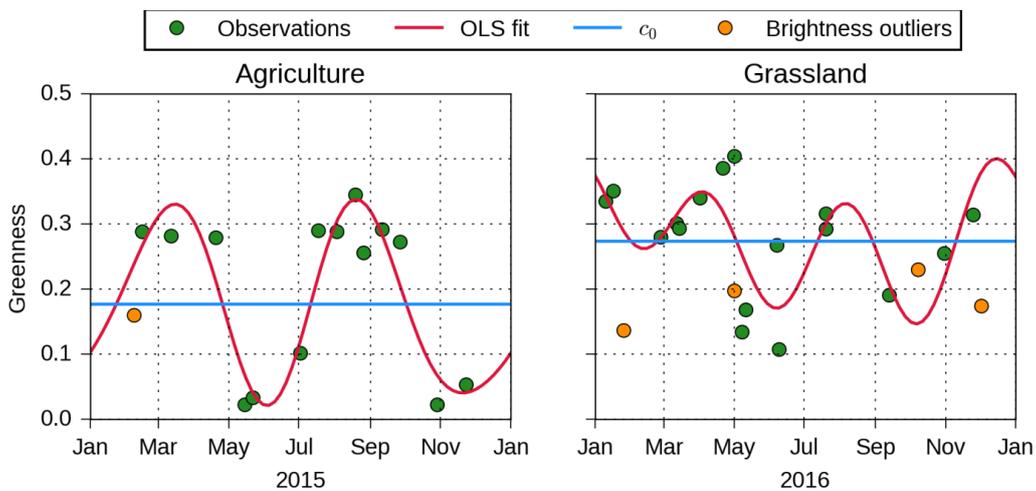


Figure 2-45: Comparison of an agricultural pixel to a grassland pixel

The parameters  $\alpha_j$  and  $\beta_j$  can be converted to amplitude values  $A_j$  using

$$A_j = \sqrt{\alpha_j^2 + \beta_j^2}$$

Additional to the parameter values, the OLS method also yields the covariance matrix of the estimates, which can be further used to derive the uncertainty of the amplitude values. Figure 2-46 contrasts the trend and amplitude parameters of the agriculture pixel with the corresponding values of the grassland pixel. The 90% confidence interval of each estimate is also illustrated, suggesting that there is a statistically significant difference in the  $c_0$  and  $A_2$  parameters. A possible approach to detect anomalies and indicate change is to carry out hypotheses tests to determine if one or more parameters have significantly changed from one year to another.

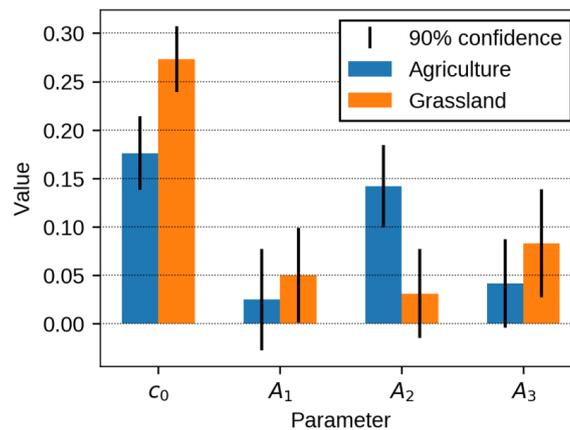


Figure 2-46: Greenness trend and amplitudes of an agriculture and grassland pixel

A different example illustrates the behaviour of a grassland pixel in consecutive years. Figure 2-47 shows the time series of Greenness and Figure 2-48 the corresponding values of the trend parameter as well as the amplitudes of the seasonal components. While the fitted curves look very different, the overall trend and the amplitudes stay roughly at the same level when their confidence intervals are taken into account. The change in the appearance of the fitted curves can be explained by phase shifts in the seasonal components.

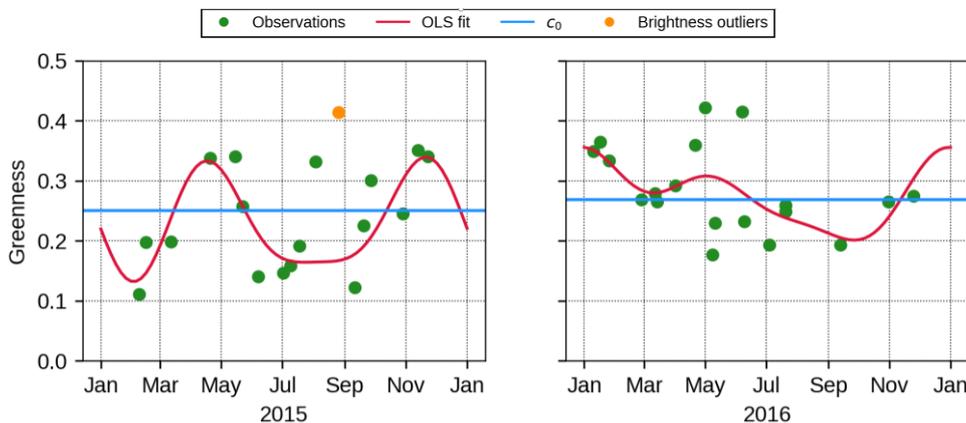


Figure 2-47: Greenness time series of a grassland pixel in consecutive years





**Figure 2-52: Changes with grassland change probability above 50%. Basis layer: ArcGIS Basemap**



**Figure 2-53: Changes with grassland change probability above 60%. Basis layer: ArcGIS Basemap**



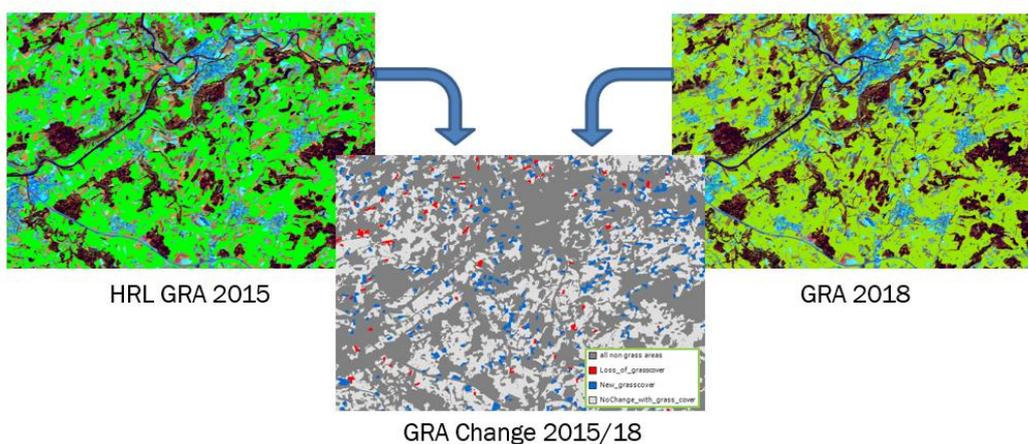
**Figure 2-54: Changes with grassland change probability above 70%. Basis layer: ArcGIS Basemap**

The map-to-map approach is highly dependent on the accuracy of the status layers. With an expected change area of 1-2% [AD08] the change rates are in general smaller than the error of the status layers, which makes it very difficult to detect changes by using only the status layers. Overall threshold based products, an overall high number of changes is observed which does not correspond to the expected change rate. The main issue of wrongly detected changes is the confusion between grassland and croplands, which strongly depend on the number of observations over the time period.

Hence, this approach strongly depends on the availability of dense time series. The annual features used as input for the grassland probability derivation require enough observations to distinguish between cropland and grassland. The methodology aggregates as change the detection of both negative and positive grassland change in this map-to-map approach.

**DEMONSTRATION SITE CENTRAL**

Figure 2-55 shows the map to map comparison between the reference year 2015 and the grassland product 2018.



**Figure 2-55: Map-to-map grassland change detection (2018/HRL 2015 reference)**

In this case, as there was no pre-processed S2 data available for 2015, which might be the case when comparing other previously existing reference maps with classifications, only map-to-map change detection is possible. Except for a boundary filter that reduces the effects from diverse resolutions of the status layers compared and geometric inaccuracies of the EO data, no further post-processing was applied to the change layer.



In the first step, the features are selected which are used to derive the change vectors. 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: (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 were used as anomaly indicators.

Figure 2-60, Figure 2-61 and Figure 2-62 show the results of the three approaches tested (MAD, CVA, map-to-map) for an area with a gain of grassland, whereas Figure 2-66, Figure 2-67 and Figure 2-68 show the results for an area with loss of grassland.

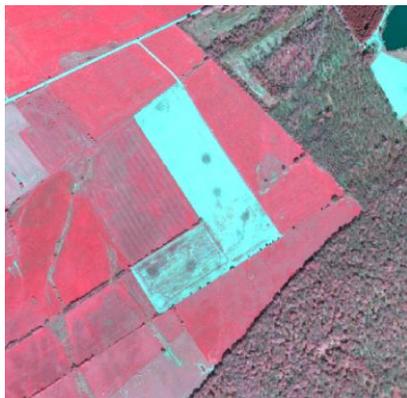


Figure 2-57: World View 1 image from the 2017.05.09



Figure 2-58: World View 1 image from the 2018.10.05

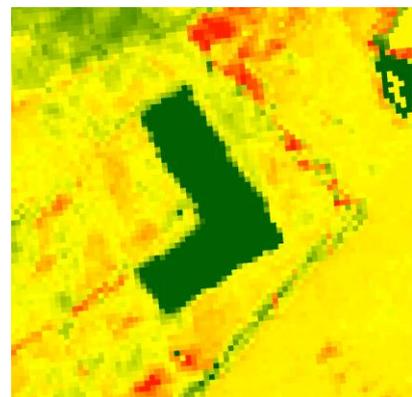


Figure 2-59: Grassland change probability map

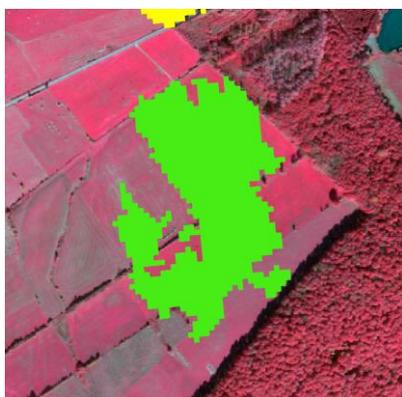


Figure 2-60: MAD approach. Loss (yellow), Gain (Green)



Figure 2-61: CVA approach. Loss (yellow), Gain (Green)

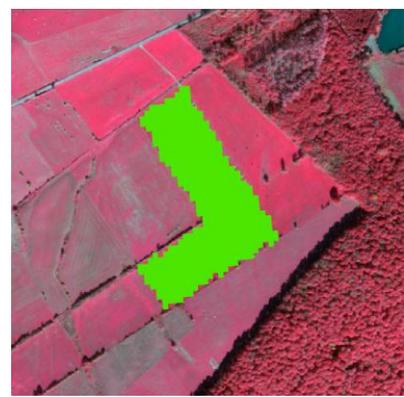


Figure 2-62: map-to-map approach. Loss (yellow), Gain (Green)



Figure 2-63: World View 1 image from the 2017.05.09

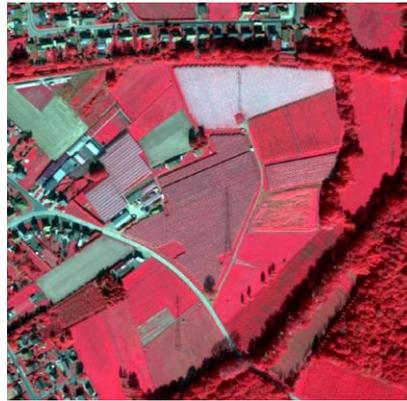


Figure 2-64: World View 1 image from the 2018.10.05

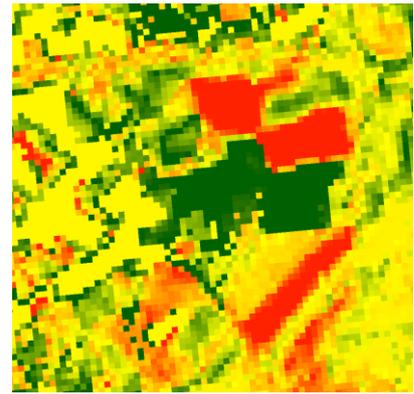


Figure 2-65: Grassland change probability map



Figure 2-66: MAD approach. Loss (yellow), Gain (Green)



Figure 2-67: CVA approach. Loss (yellow), Gain (Green)



Figure 2-68: map-to-map approach. Loss (yellow), Gain (Green)

The qualitative analysis revealed, that the polygon-based MAD approach performed worse due to overestimated anomaly areas and patch artefacts. The map-to-map approach also shows an overestimation of changes, which however can be reduced by applying the CVA approach. 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.

**Table 2-17: MAD anomalies combined with RF probabilities on polygon level**

MAD approach		REFERENCE			Grassland	Total	User Accuracy	Confidence Interval
		Non Grassland	Gain	Loss				
PRODUCT	Non Grassland	77.07	0.01	0.02	0.33	77.43	99.54	0.61
	Gain	0.04	0.22	0.00	0.23	0.48	44.91	6.59
	Loss	0.15	0.00	0.19	0.12	0.45	41.21	8.70
	Grassland	4.82	0.00	0.02	16.79	21.63	77.64	4.90
	Total	82.08	0.24	0.22	17.47	100.00		
	Producer Accuracy	93.90	92.47	85.69	96.12		94.27	Overall Accuracy
	Confidence Interval	1.21	37.14	47.01	2.51		1.16	Confidence Interval
							F-Score Non-Grassland	0.97
							F-Score Gain	0.60
							F-Score Loss	0.56
							F-Score Grassland	0.86
							Kappa	0.62

**Table 2-18: CVA anomalies (2 sigma) combined with RF probabilities on pixel level**

CVA approach		REFERENCE			Grassland	Total	User Accuracy	Confidence Interval
		Non Grassland	Gain	Loss				
PRODUCT	Non Grassland	82.10	0.04	0.14	2.21	84.48	97.18	1.28
	Gain	0.05	0.19	0.00	0.23	0.47	40.18	6.90
	Loss	0.06	0.00	0.17	0.00	0.24	72.34	7.87
	Grassland	0.11	0.00	0.00	14.70	14.81	99.26	1.43
	Total	82.32	0.23	0.31	17.14	100.00		
	Producer Accuracy	99.73	82.32	55.56	85.78		97.16	Overall Accuracy
	Confidence Interval	0.26	50.92	47.17	5.22		1.10	Confidence Interval
							F-Score Non-Grassland	0.98
							F-Score Gain	0.54
							F-Score Loss	0.63
							F-Score Grassland	0.92
							Kappa	0.60

**Table 2-19: map-to-map approach**

Map-to-Map approach		REFERENCE			Grassland	Total	User Accuracy	Confidence Interval
		Non Grassland	Gain	Loss				
PRODUCT	Non Grassland	81.10	0.00	0.01	2.33	83.45	97.19	1.80
	Gain	0.44	0.23	0.00	0.23	0.91	25.78	4.61
	Loss	0.28	0.00	0.20	0.23	0.71	28.26	5.44
	Grassland	0.11	0.00	0.00	14.82	14.93	99.26	1.43
	Total	81.93	0.24	0.22	17.61	100.00		
	Producer Accuracy	98.98	98.63	92.40	84.16		96.36	Overall Accuracy
	Confidence Interval	0.27	19.41	52.02	7.16		1.52	Confidence Interval
							F-Score Non-Grassland	0.98
							F-Score Gain	0.41
							F-Score Loss	0.43
							F-Score Grassland	0.91
							Kappa	0.50

The statistics in Table 2-17, Table 2-18 and Table 2-19 confirm the results from the qualitative assessment. The map-to-map approach significantly overestimates the changes. The MAD polygon based approach does spatially not cover the areas good enough but is able to remove some of the misclassifications and the CVA approach is able to reduce some of the misclassifications compared to the map-to-map approach but still shows an overestimation due the accuracy of the status maps. Further, it has to be noted that the grassland change area for the demonstration site is lower than the error of the status maps. This can also be reflected in the statistics when for most categories the confidence intervals are broad. For operational applications we recommend to combine a signal anomaly detection approach with the map-to-map comparison as shown above for the change vector approach.

**DEMONSTRATION SITE CENTRAL**

Additionally, once the 2017 and 2018 grassland masks have been generated a simple and transferrable method is implemented to achieve operational feasibility at larger scales and allowing for a cost-efficient automation. Actually, as the tests proved successful, the approach was further tested and applied analogously in the forest domain tests, confirming satisfactory accuracies in all test sites and domains. The geometric difference between the status layers of both reference years is computed to derive the recoded unchanged grassland and changes, embedding both new grassland cover and loss of grassland cover changes. Filtering isolated pixels and buffering reduce artefacts resulting from positional shifts of Sentinel data and border effects coming from mixed pixels. In turn, the change indicator is computed as the time feature difference between 2018 and the 2017 minimum NDVI. The resulting time feature mask (e.g., minimum NDVI) is recoded assigning change to all areas where the mask is less than the mean minus 1.5 times the standard deviation or the values greater than the mean plus 1.5 times the standard deviation and unchanged otherwise. This threshold is defined on the basis of visual inspection and expert knowledge. Next, by combining the time feature mask recoded and the cleaned difference between the status layers (post filtering, buffering and recoded), pixels labelled as changes in both, remain as such. Finally, the combination of this layer providing the change is combined with binary grassland mask. Filtering by a suitable MMU (i.e., 0.5ha) allows for erasing small differences between the two status layers (e.g., it could happen that a small patch was captured better in one or the other layer and therefore a part of the patch is highlighted as a change). In the case of grassland, in view of the testing experiences in the different biogeographic regions, it is considered that half a hectare changes are a lot more meaningful.



As reported in WP33, the production of the change layer is dependent on the status layers from which the change is derived. For instance, in areas of higher elevation where snow cover is found for long periods of the year, the classification is not that accurate. This was the case for both years (2017 & 2018), which led to greater differences and therefore change detection which is basically no change in the end. To overcome the misclassified change in higher elevated areas, a height threshold can be defined, as it is reasonable to assume that over a certain height, no tilling is applied and therefore there can be no change.

It must be taken into account that as in the grassland classification there was no historical time series included to detect ploughing events and exclude grassland younger than 5 years, 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 within the last 5 years (or whatever is given by definition) can be excluded.

In the test and demos, it has been proved that the filtering improves the look and feel of the respective layers by reducing noise. This contributes to keep meaningful changes in the change product while removing small areas, which are likely due to errors in classification. Thus, it was decided to filter all areas below 0.5ha (50 pixels).

### 2.3.4 Summary and conclusions

The benchmarking has been conducted in two different demonstration sites WEST and Central where different approaches like map-to-map, signal anomaly detections based on dense time series and the combination of both are benchmarked.

In the Central site, the change detection was applied, on the one hand, between the 2017 status layer (re-sampled to 20m) and the reference year HRL2015 grassland. On the other hand, for the 2017 and 2018 status layers change detection is computed by combining a map-to-map comparison with change indicators derived from Sentinel-2 time series for 2017 and 2018. In the 2017/2018 change detection, fewer changes were found when compared to the map-to-map comparison between the 2018 grassland product and the reference HRL2015 grasslands. Reasons explaining this are the different resolutions of the compared status layers, as the HRL2015 is produced at 20m spatial resolution, whereas the ECoLaSS Grassland 2018 is provided at an improved 10m resolution, which was re-sampled to 20m to enable a comparable change detection.

In the West site, two anomaly detection approaches are tested, with similar results. It is recommended to use the pixel-based approaches instead of the polygon-based. The qualitative analysis revealed that the polygon-based MAD approach performed worse due to overestimated anomaly areas and patch artefacts. The map-to-map approach also shows an overestimation of changes, which however can be reduced by applying the CVA approach. 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.

The signal anomaly detection approaches are also compared to the map-to-map approach where it can be concluded that the combination improves the results significantly. The map-to-map approach strongly depends on the availability of dense time series. The annual features used as input for the grassland probability derivation require enough observations to distinguish between cropland and grassland. The methodology aggregates as change the detection of both negative and positive grassland change in this map-to-map approach.

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.

Further, it has to be noted that the grassland change area for the demonstration site is lower than the error of the status maps. This is also reflected in the statistics where for most categories the confidence intervals are broad. For operational applications we recommend to combine a signal anomaly detection approach with the map-to-map comparison. In the test and demos, it has been proved that the filtering improves the look and feel of the respective layers by reducing noise. This contributes to keep meaningful changes in the change product while removing small areas, which are likely due to errors in classification. Thus, it was decided to aggregate all changes by applying a minimum mapping unit of 0.5ha (50 pixels).

In phase 2, from the tests in the thematic topics Agriculture and Grassland these conclusions are reinforced, as was expected. Data availability in key phenological moments throughout the season is a must for separability of these land covers. As explained in WP33, in the first place, the change detection is dependent on the quality of the respective status layers. For the classification of grasslands, the time period analysed should not be too short, otherwise, results are not meaningful due to the limited number of scenes. Moreover, in different biogeographical regions the definition of the most suitable time window (e.g., spring) is likely to be different. The important point is to define a time window where grassland and cropland (responsible class for most misclassifications) are best separable (e.g., capture the moment when grassland is already greening while cropland is not). When compared to the tests in Central, in Mediterranean regions this window is shifted more towards winter (e.g. Dec-Mar).

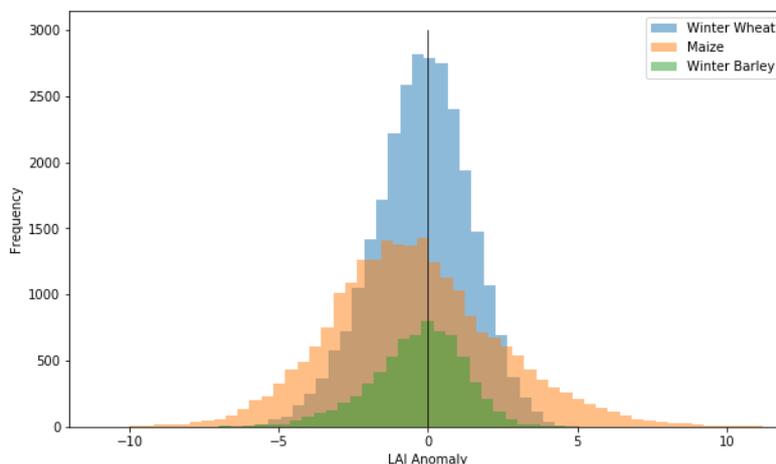
For operational applications, we recommend combining a signal anomaly detection approach with the map-to-map comparison as shown above for the NDVI anomaly detection approach and the change vector approach. The benchmarking results indicated that signal anomaly detection can contribute to successful monitoring of grassland, forest, water and imperviousness land cover changes when combined with changes detected by multi-temporal classification of land cover as described in the next chapter.

## 2.4 Agriculture change

In the beginning of the Sentinel era, the detection of agriculture change clearly refers to spatial or temporal deviation with regards to expected or most frequent behaviour and not no longer on anomalies detection with regards historical record or climatological datasets. Two different applications were considered for testing new Sentinel-2 methods: mapping the local deviation of crop growth trajectory and mapping the diversity of a crop phenological stage, i.e. the emergence date.

In order to exploit the high spatial and temporal resolution of Sentinel-2, the tested approach targets assessing crop growth condition by the trajectory of variables representing photosynthetic activities such as Green Area Index (GAI) and/or Leaf Area Index (LAI). The approach uses the BVnet algorithm employing artificial neuronal networks trained on simulated LAI and reflectance values where reflectance values are simulated by the ProSail radiative transfer model for the Sentinel-2 bands (Delloye et al., 2018). The tested method proceeds at the field level in relative terms with regards to neighbouring fields of the same crop (within a radius of 3km). This product aims to identify local marginal behaviour along the season in terms of crop growth cycle, crop development or management practices. An example of quantitative anomaly derived from the comparison between the LAI profile of each field with its surrounding ones provide a wall-to-wall map of the local growth cycle deviation.

The LAI trajectory described by the area under the LAI interpolated curve differs from one field to another and characterises the growth conditions of a crop in a given field with regards to the average LAI curve of the neighbouring fields cultivated with the same crop. As a proof of concept the Figure 2-71 depicts the cumulative LAI anomalies of three main crops of Wallonia region (part of West site) where winter wheat LAI anomaly shows bell shaped distribution while the majority of maize LAI anomaly have negative values and winter barley anomaly has skewed distributions (to the left). These differences can be partly related to the heterogeneity of agricultural practices typical for each crop.



**Figure 2-71: Distribution of the LAI anomalies for the three main crops (Winter Wheat, Maize, and Winter Barley) in 2017. Crops: The anomaly corresponds to the difference between the area under LAI curve for the field of interest and the one for the local average**

The second application aims at detecting the emergence date of maize and several methods have been benchmarked in the South Africa site. The approach is based on the analysis of Sentinel-2 time-series to identify the first phenological stages (crop emergence, first leaf development). The phenological stages are commonly defined according various classification systems such as the Biologische Bundesanstalt and Bundessortenamt und Chemische Industrie (BBCH) scale.

In particular, a specific BBCH stage is reached when at least 50% of the plants are within the definition of that stage (López-Sánchez et al., 2012). The onset date of the stages presents an important inter-annual variation. It can be explained by environmental and climatic factors as well as farm-level management decisions (crop variety, crop rotation, input availability, etc.) (Sakamoto et al., 2010). Generally, temperature and water are the main climatic factors impacting the development of the majority of species. A large number of species is also impacted by the length of the photoperiod. In temperate climates, light is generally the primary limiting growth factor. In humid climates, light and nutrients are both limiting. In tropical or dry subtropical climates, water is the main constraint but the absorption of nutrient is also reduced.

### 2.4.1 VIs and hue time series as candidate data sources

Several vegetation indices (VIs) and hue computed from surface reflectance time series are qualified candidate data source as VIs and hue works as a proxy of the Fraction of Absorbed Photosynthetically Active Radiation (fAPAR).

Vegetation Indices (VIs) can be seen as a proxy of the Fraction of Absorbed Photosynthetically Active Radiation (fAPAR) as they relate "greenness" with the measure of the absorption characteristics of the vegetation in the red and NIR spectral bands (Pinty et al., 2009). NDVI first relates to total green biomass and is sensitive to low to moderate LAI values but saturates at high values (Nguy-Robertson et al., 2014). VIs do not present a straightforward biophysical interpretation, although they are strongly correlated with biophysical variables (White et al., 1997; Eklundh et al., 2003; Pinty et al., 2009). LAI retrieval

algorithm can also be considered but the underlying assumption and the higher computing cost of such an algorithm prevent considering it at this stage.

Studies showed that the NDVI gives good estimates of the vegetation dynamics when the vegetation is photosynthetically active (Palacios-Orueta et al., 2012). Other VIs based on the Short-Wavelength Infrared (SWIR) are better for assessing low vegetation density zones. In semi-arid areas, the information contained in both the Mid-Infrared (MIR) and blue regions relates to soil properties which helps in distinguishing vegetation type classes (Hüttich et al., 2009).

Pekel et al. (2011) studied the detection of green vegetation in semi-dry and dry areas. The image color is transformed from a Red-Green-Blue (RGB) to a Hue-Saturation-Value (HSV) representation. As the MIR and NIR regions present several advantages for soil discrimination, the three MIR-NIR-Red bands are used instead of RGB. The Hue component is the only parameter conserved because it is able to discriminate land cover type where Saturation and Value fail. Marinho et al. (2014) applied the method developed by Pekel et al. (2011) and tested it for sowing date estimation which had not been investigated yet. The study aimed at estimating green-up onset dates in arid and semi-arid regions (i.e. the Sahel region) from MODIS 250m resolution images and RFE 8km resolution rainfall estimate and, then, comparing it with ground-truth data.

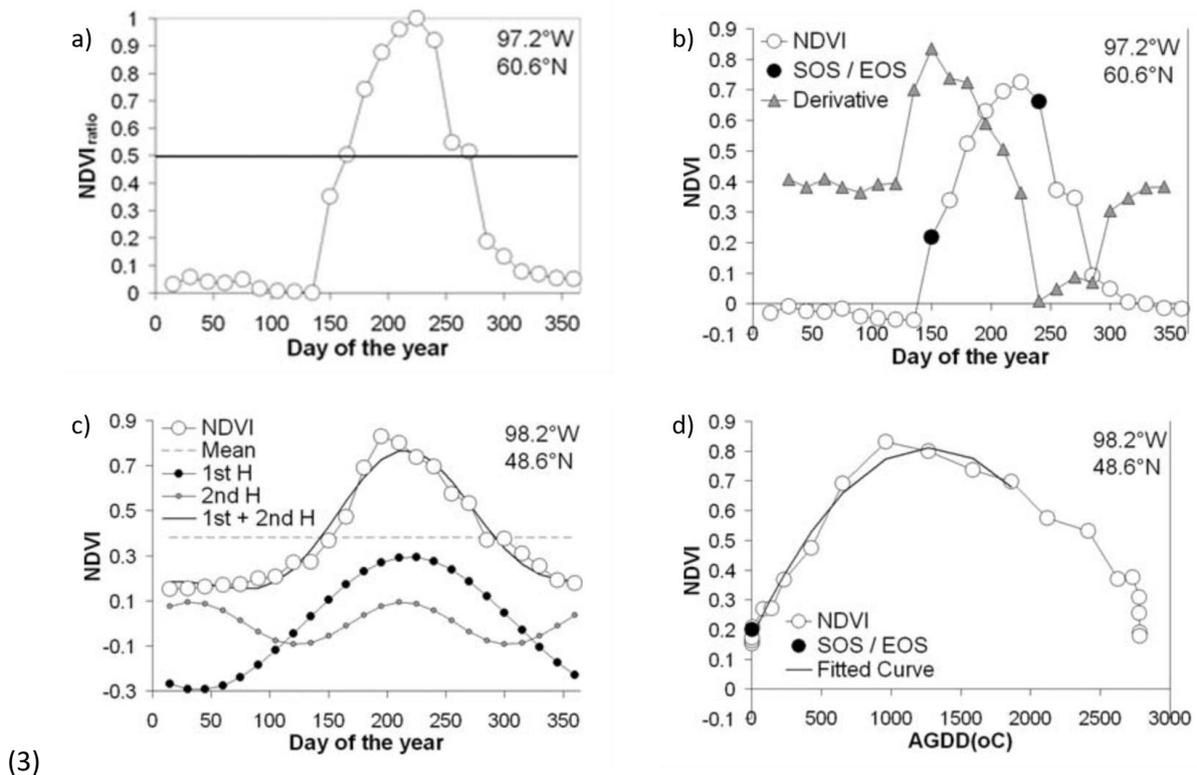
To reduce the noise and fill the gap in the time series, two interpolation methods are tested. First, a simple linear interpolation method is applied between all the successive observations, between October and the end of April. Alternatively, a logistic interpolation method commonly used to reduce noise in the vegetation profiles is also selected to interpolate the satellite observations during plant growth (i.e. between minimum and maximum index values): a four-parameter logistic function.

## 2.4.2 Candidate detection methods

The candidate methods for phenology studies are often grouped into four main categories: threshold, moving window, function fitting and model fitting methods (Zeng et al., 2016; de Beurs and Henebry, 2010).

(1) The threshold method (Figure 2-72 a) is based on linking a phenological event with the crossing of a certain value of the VI curve. The threshold can either be fixed or dynamic and varies with land cover, soil background, view and solar angle (Reed et al., 1994). However, they do not rely on an underlying biophysical meaning. For example, White et al. (1997) identified the onset and end of greenness when the NDVI ratio of the smoothed curve exceeds or falls below 0.5 respectively. Lobell et al. (2013) defined the green up phase as the date when the double-logistic fitted function exceeds 10% of the year's maximum amplitude. The main drawback of the fixed threshold is the disability for reflecting the spatial changes of larger study area and inconsistency for a wide variety of land covers (de Beurs and Henebry, 2010; Reed et al., 1994). Plethora of thresholds have been used based on the long-term VIs mean, yearly VIs, NDVI ratios, Normalized Difference Water Index (NDWI), etc. The ratio approach has the advantage of being independent from the geographic location and land cover of the area. As such, the NDWI is particularly indicated for heavily snowed areas.

(2) The moving window method can be derivative or backward-looking moving average. The derivative method (Figure 2-72 b) is founded on the assumption that the fastest green-up or greatest leaf expansion corresponds to the most ecologically relevant SOS (White et al., 1997). In other words, the maximal increase and decrease of NDVI tally with SOS and EOS (de Beurs and Henebry, 2010). Moving windows of a certain temporal extension are applied on each pixel and the slope (or derivative) is calculated. The highest positive and lowest negative derivatives are then easily extracted.



**Figure 2-72: Four methods based on the NDVI to detect start and end of the season. a) fixed threshold, b) derivative, c) Fourier transform, d) quadratic fitting based on AGDD (de Beurs and Henebry, 2010)**

Some methods retrieve the second derivative and determine the SOS as the time point combining a positive slope and a local maximum (Balzter et al., 2007). Cong et al. (2013) defined green-up onset date as the highest positive relative change of the average NDVI time-series for a 15-day moving window. Moulin et al. (1997) identified the beginning of the vegetation cycle (b\_date) on three conditions: (i) NDVI value at b\_date is close to a bare soil value, (ii) left derivative (before b\_date) should be equal to zero because NDVI is assumed constant before the growth season, (iii) right derivative (after b\_date) should be positive on two weeks' time window. de Beurs and Henebry (2010) reported that this method gives good results where the NDVI curve displays a sharp increase and a steep decrease.

The backward-looking moving average method identifies the onset of greenness as the date when the VI curve crosses the moving average function which represents a significant change in the growth trend. The moving average is built as the average of the last *i* observations. The choice of the temporal window (i.e. number of *i* observations) is crucial and arbitrary as it introduces a time lag: a large time interval is less sensitive whereas a small interval may take insignificant variations into account (Reed et al., 1994; de Beurs and Henebry, 2010).

These two first methods present the advantage of being able to retrieve multiple growing seasons (Verhegghen, 2013). However, they are not good at distinguishing the basic temporal variations of the vegetation reflectance (noise) from the relevant seasonal changes. That, the date retrievals based on local minima, maxima, or fixed thresholds can be completely shifted if observation errors contaminate the original dataset. For instance, atmospheric constituents, bi-directional reflectance distribution function, cloud coverage, and the mixed-pixel effect often influence MODIS images (Sakamoto et al., 2010).

(3) The function fitting method (Figure 2-72 c) applies a mathematical function to a given VI curve to smooth or filter the data and extract the main information. Different functions have proved to be useful: Fourier analysis, wavelet transformation, Principal Component Analysis (PCA), Canopy Structure Dynamic Model (CSDM), etc. The Fourier analysis which decomposes a complicated curve into a sum of sinusoidal

waves, is able to approximate a VI (de Beurs and Henebry, 2010). This segmentation is sensitive to systematic changes and reduces the non-systematic data noise. To interpret the new curve, the first Fourier harmonic is considered to represent the mean NDVI. The wavelet transform also decomposes the VI time-series into a set of small local waves (named wavelets) assuming the fact that the noise components have higher frequencies than the main seasonal changes (Sakamoto et al., 2010). An important aspect is that this type of frequency decomposition performs better on long time-series showing periodic changes. Consequently, the source observations should be measured at a regular time interval or require gap filling to be adequately processed (de Beurs and Henebry, 2010).

To retrieve phenological events from a fitted curve, the procedure of Sakamoto et al. (2005) can be used: the minimal or inflection point earlier than 60 days before the maximum value (defined as heading date) is selected, then, the later of the two points is identified as planting date. However, the Root Mean Square Error (RMSE) of 12.1 days for planting date estimate is not satisfactory. Another way to account for key information lies on the PCA. Through a linear combination of the original observations, the primary factors explaining the main variance of the dataset are kept. Again, the interpretation of the resulted parameters (eigenvectors) is not self-evident and does not remain consistent over the years limiting the comparative power of the method (de Beurs and Henebry, 2010). The advantage of those fitting methods is to reduce noise and adjacency pixel problem (between pixel effects) and their ability to derive phenological metrics in a consistent way (Palacios-Orueta et al., 2012).

The model fitting method fits a model to the remote sensed observations. These models can be simple (logistic models, etc.) or more complex (Gaussian Local Functions, etc.) and are previously defined or dynamically built with input data (de Beurs and Henebry, 2010; Zeng et al., 2016). The number of input parameters compared to the amount of observations available for their identification and the need of large-scale ground-truth data is crucial when assessing the scope and implementation of these model fitting (Duchemin et al., 2008).

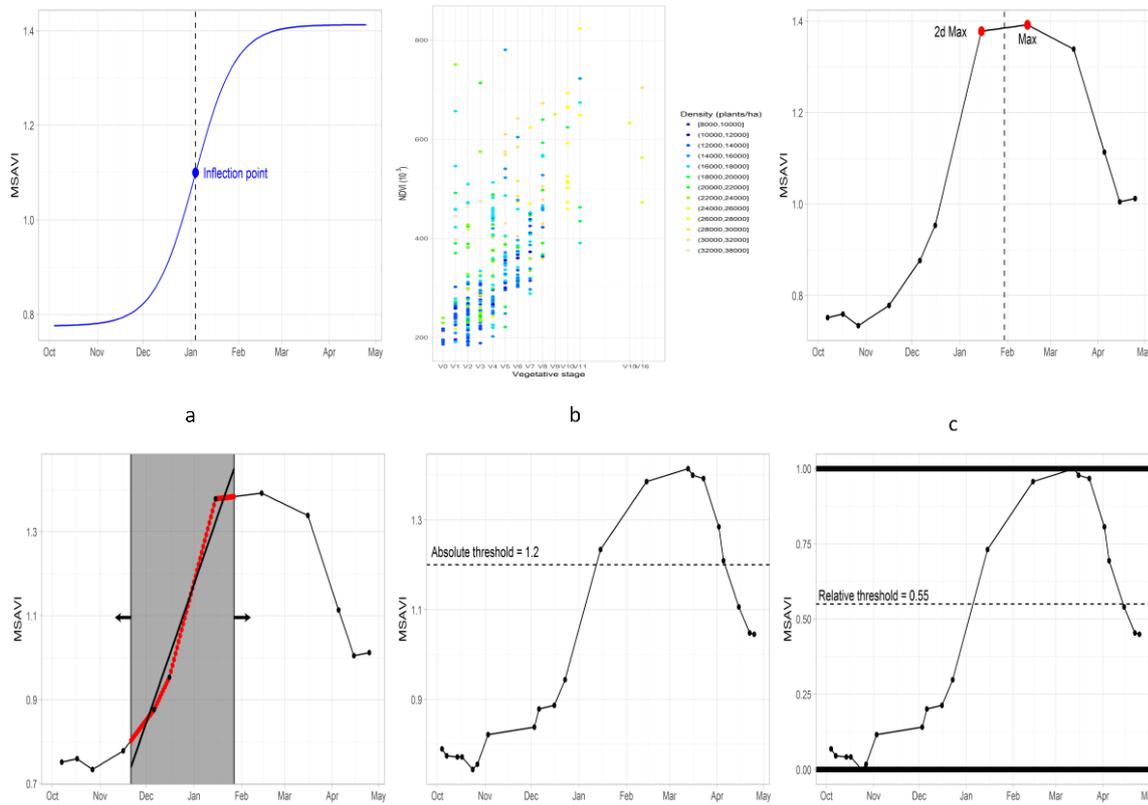
Accumulated Growing Degree Days (AGDD) can be interpreted as a measure of the accumulated heat above a specified base temperature from the beginning of the season: maize base temperature is estimated around 10 °C. Modelling vegetation growth under AGDD instead of anthropocentric calendar time (Figure 2-72 d) is more relevant especially during the first half of the growing season when day length and water stress are not the main contributors yet (de Beurs and Henebry, 2010).

These four methods can be combined. Threshold methods are generally applied on smoothed function to reduce data noise.

Then, a logistic function was applied to each identified increasing or decreasing period and key phenological dates were then retrieved from the fitted curve. For a single growth cycle, the following logistic function modelled the curve:

$$y(t) = \frac{c}{1 + e^{a+bt}} + d$$

where  $t$  is time in days,  $y(t)$  is the VI value at time  $t$ ,  $a$  and  $b$  are fitting parameters,  $c + d$  is the maximum VI value, and  $d$  is the initial background VI value.



**Figure 2-73: Overview of the six emergence estimation methods. Methods without parameterization: (a) inflection point, (b) base logistic (c) maximum value. Methods with parameterization: (d) highest slope, (e) absolute threshold, (f) relative threshold. The threshold methods are both tested on the linear and logistic interpolated observations. Presented vegetation profiles are examples from the maize calibration sample**

In the phase 2 this application has been further investigated. On one hand an alternative time scale no longer based on calendar date but rather on ‘Growing Degree Days’ providing a physiologically-based time. This is expected to provide a more generic retrieval algorithm, more robust and more scalable over large areas.

### 2.4.3 Benchmarking criteria

Furthermore, the comparison of the results of a systematic combination of detection methods and vegetation proxies is necessary for detecting the best combination. The systematic analysis of three approaches (i.e. threshold, derivative method and model fitting) and three vegetation proxies (NDVI, soil-adjusted MSAVI and Hue index (MIR-NIR-Red)) interpolated using two functions (linear, logistic), and the assessment of their relative performances against ground-truth measurements are achieved.

It is important to notice that each tested method aim to detect different time metrics (threshold intersection, highest slope, inflection point, maximum value and base logistic value) to estimate emergence date. Obviously the emergence date is then estimated by using a correction coefficient. The correction coefficient is named the time-lag which corresponds to the interval between the emergence date and the temporal metric specific for each method. Hence, the estimated emergence date from each profile corresponds to the temporal metric date minus the mean time-lag for this method. Subsequently time-lag stability (between the profiles) is the criteria to assess the performance between methods and VIs profiles.

Given that stable time-lag reducing the spread around its value is targeted, two statistical indicators are used to discriminate the different methods and VIs: Standard deviation (SD) and median absolute deviation from the median (MAD). In particular, SD is used for measuring the dispersion of the

emergence date estimations around the mean time-lag but it is very sensitive to outliers. Therefore, MAD is incorporated as addressed by Varmuza and Filzmoser (2016).

$$\text{MAD} = b \times \text{median} ( | (x_i - \text{median} (x_n)) | )$$

with,  $x_n$ , the  $n$  original observations and  $b$ , a multiplicative parameter of 1.4826, assuming the normality of the data and disregarding the abnormality induced by outliers.

Although, utilizing both SD and MAD avoid the outliers affect and treat the data as it has Gaussian distribution, significant difference are shown by the two indicators (Leys et al., 2013).

#### 2.4.4 Implementation and benchmarking results

The objective was to compare different VIs and detection methods to provide an accurate estimation of the emergence date at the field level. This benchmarking was carried out using a large in situ dataset provided by an insurance company (for more information, see deliverable D4.1). This experience demonstrated for maize that the relative threshold yields the most accurate estimate of the emergence date (about 10 days accuracy), probably due to the iterative adjustment process. More specifically, this method applied on NDVI time series provided a SD of 8.3 days and MAD of 7.4 days. Very similar results were obtained from MSAVI time series (i.e. a SD of 8.5 days and MAD of 7.4 days) with a very low coefficient of variation (2.9%). The results obtained for maize are most likely to be generalized to other crops. The sunflower results are not reported here due to the limited validation dataset.

Surprisingly, the absolute and relative thresholds do not present significantly different results, although the variations in sowing densities appear to affect the index values. This might notably be due to a relatively good homogeneity of the observed fields (high quality fields) and/or the good weather conditions for the entire region for this year.

The relative threshold method is expected to be more flexible and match more diverse situations at a larger scale. Density, cultivar, management practices are often quite different over large areas and absolute threshold might rapidly prove inaccurate. Therefore, the relative threshold is probably more indicated for scaling up this product over large regions. To apply this method, it is however necessary to wait for a crop type product (WP44) because the parameters are currently crop-specific. This crop type product is generally available at the mid-season. Later on parameter values might probably be used for multiple crops but this would require further investigation. At the field level, this method might also help in detecting anomalies or within field spatial patterns relevant to precision agriculture.

The main drawback of the relative threshold is that it will output a result regardless of the real vegetation status. A combination of an absolute threshold discriminating bare and cropped soils with a relative threshold might prove useful. Furthermore, these threshold methods are also strongly impacted by a saw-tooth pattern of observation, which was very limited in this dataset. For more erratic observations, the logistic interpolation might prove very useful to smooth the index time-series

With regards to the indices, the MSAVI, NDVI and Hue index show similar growth dynamics. MSAVI proves to yield the most accurate results with lowest SD and MAD in comparison to the other methods. A second reason to favor MSAVI at the expense of the other three indices is its remarkably low and stable coefficient of variation which is 2.9% at the extraction zone level. The objective of mitigating soil background effects for MSAVI is a reasonable explanation for this behaviour.

The Hue index in accordance with the results of Pekel et al. (2011) exhibits a sharper increase than NDVI with, however, a slightly lower start. This rapid raise does not seem to better the results. The conditional definition of the index could cause a higher variability during the season which, averaged, is less stable than the other indices.

It is also more interesting to discuss the potential and complementary uses of a set of methods. Indeed, if one is interested in near-real time results, the time-lag value or the need to wait for the maximum value might be more decisive than the overall accuracy of the method to make a choice. One might

prefer a method robust to outliers, or, on the contrary, a method that half of the time gives very accurate results. One might prefer a method that gives good relative results for a within field application and not valuing the exact emergence date estimation. The three methods without parametrization seem to present a bias to emergence date as emergence date tends to be delayed for early-planted fields and advanced for late-planted fields, particularly for the base logistic method. Also, results of the methods with parameterization presented by Table 2-20. Table 2-21 shows that the best method is the relative threshold for MSAVI and and NDVI with lowest SD and MAD.

**Table 2-20: Methods without parameterization. Results are expressed in days. Results for LAI on logistic interpolation are not presented as the number of fitted vegetation profiles fall below 15%**

		MSAVI		NDVI		Hue index		LAI	
		calib	valid	cal	valid	cal	valid	calib	valid
				ib		ib			
Inflection point	SD	19.7	19.8	17	14.1	12	37.9		
	MAD	7.4	9.6	.6	8.9	.5	7.4		
				7.		8.			
				4		9		Not applicable	
	Time-lag	30.6	31.1	35	33.3	26	27.5		
				.1		.4			
Base logistic	SD	14.1	16.7	12	15.1	13	12.2		
	MAD	13.3	14.1	.1	11.1	.5	9.6		
				10		10			
				.4		.4		Not applicable	
	Time-lag	1.3	-2.1	8.	5.5	5.	5.8		
				0		9			
Maximum value	SD	17.6	17.5	17	17.4	18	17.9	34	31.4
	MAD	19.3	20	.7	20.8	.5	20	31.1	23.7
				19		19			
				.3		.3			
	Time-lag	96.4	94.6	96	94.6	97	95.8	79.4	78.6
				.3		.9			

**Table 2-21: Methods with parameterization. Results are expressed in days. Results for LAI on logistic interpolation are not presented. The min/bias row corresponds for the calibration sample to the indicator (SD or MAD) which is minimized while bias is the difference between the zero-mean expected validation and validation emergence date estimations**

		MSAVI		NDVI		Hue index	
		calib	valid	calib	valid	calib	valid
Absolute threshold (linear interpolation)	SD						
	MAD	10.8	8.9	12.9	11.7	8.9	12.4
	Time-lag	17.0	10.4	28.5	11.9	22.7	10.4
	min/bias	SD	-0.9	MAD	0.9	SD	-2.3
Relative threshold (linear interpolation)	SD	9.8					
	MAD	7.4	11.6	10.2	7.4	11.6	11.1
	Time-lag	28.3	8.9	30.4	8.9	21.9	12
	min/bias	SD	-1.2	MAD	-0.2	SD	-1.5
Absolute threshold (log. interpolation)	SD	10.2				9.3	
	MAD	11.9	10.5	10.7	7.4	11.5	9.6
	Time-lag	25.3	8.2	22.6	10.4	24.1	7.4
	min/bias	SD	-1.5	MAD	-1.2	SD	-1.3
Relative threshold (log. interpolation)	SD	9.4		9.0		9.3	
	MAD	<b>7.4</b>	9.6	<b>7.4</b>	11.8	8.2	12.1
	Time-lag	24.4	<b>8.5</b>	26.5	<b>8.3</b>	22.6	11.1
	min/bias	SD	-0.6	SD	-1.6	MAD	-1.4
Highest slope	SD	13.6	14.6	17.6	17.4	13.6	13.9
	MAD	8.9	11.1	10.4	10.4	10.4	10.4
	Time-lag	29.4	29.0	33.0	32.8	24.9	24.1

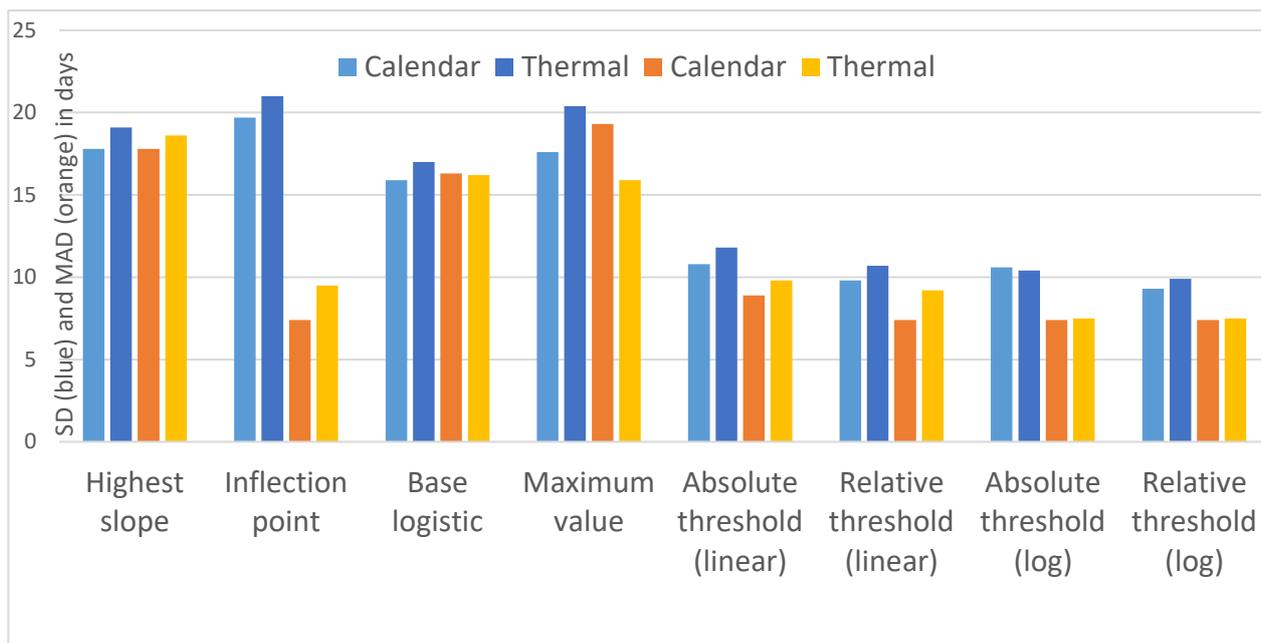
The window size for the highest slope method is optimized for each index. The optimal window is the one minimizing SD and is stable for the three VIs. Furthermore, SD is varying slowly for close window size values, it is therefore not a highly sensitive parameter.

The lack of satellite images between January and February (i.e. close to the maximal index values) might have affected the above results. The so called monthly sample was then compiled to investigate the best possible performance from a set of fields observed at least once a month and by just looking at the calibration performance as they are too few samples (n=133). Table 2-22 compares the monthly samples and calibration of all methods using the MSAVI. For the threshold methods, improvements of one day for the SD is observed for MSAVI and NDVI and half a day for the Hue index. The lowest values are a SD of 8.3 days and MAD of 7.4 days for NDVI and 8.5 days, 7.4 days for MSAVI respectively for the relative threshold method on logistic interpolation (Table 2-21). The inflection point method is greatly improved (probably due to the absence of outliers) with values close to the threshold results: SD of 9.4, 9.0, 10.3 days and MAD of 8.9, 7.4, 8.9 days, for MSAVI, NDVI and the Hue index respectively. These results highlighted the sensitivity of the results to the time series density.

**Table 2-22: Comparison of the calibration and monthly samples for all methods for MSAVI**

		MSAVI	
		calib	monthly sample
Highest slope	SD	13.6	10.7
	MAD	8.9	8.9
Inflection point	SD	19.7	9.4
	MAD	7.4	8.9
Base logistic	SD	14.1	17.5
	MAD	13.3	17.8
Maximum value	SD	17.6	13.0
	MAD	19.3	14.8
Absolute threshold (linear interpolation)	SD	10.8	9.8
	MAD	8.9	8.9
	min	SD	SD
Relative threshold (linear interpolation)	SD	9.8	9.0
	MAD	7.4	5.9
	min	SD	SD
Absolute threshold (logistic interpolation)	SD	10.2	9.1
	MAD	11.9	10.4
	min	SD	MAD
Relative threshold (logistic interpolation)	SD	9.4	8.5
	MAD	7.4	7.4
	min	SD	MAD

In the phase 2, calendar and GDD time scale have been tested using the same dataset and for the different methods. Whatever the method to retrieve the emergence, it was found that the performances are much less sensitive to the time scale than to the methods for both performance metrics as reported in the Figure 2-74.



**Figure 2-74: Methods performances for detecting the maize green onset for two time representation. Blue corresponds to the calendar days while orange to the Growing degree days. SD stands for standard deviation and MAD for Median Absolute Deviation**

Beyond the time scale, the plant density and the crop row width vary over large regions according to local agriculture practices and meteorological conditions (dry year, drought, etc.). Their respective influence on the NDVI has been assessed carefully and was shown to have a limited impact.

### 3 Conclusions and Outlook

Estimating change from remotely sensed data is not straightforward since time series contain a combination of seasonal, gradual and abrupt changes in addition to noise which originates e.g. from remnant geometric misregistration or atmospheric effects. For monitoring land cover changes on pan-European level, the availability of optical EO time series is a bottleneck because of data gaps and non-optimal acquisition time frames due to high frequency cloud and snow cover and/or low solar incidence angles especially in the North of Europe and in Alpine areas. Therefore, for pan-European or global applications, a main requirement is the usage of optical as well as SAR time series to allow a homogeneous wall-to-wall coverage for change monitoring. Depending on the biogeographic region, approaches which are only based on optical data streams (e.g. in Mediterranean areas), and approaches which combine optical and SAR data streams (e.g. in areas with more frequent cloud cover) can be applied. Therefore, optical and SAR based approaches were benchmarked separately.

In phase one of ECoLaSS, a comprehensive state of the art review was performed, which together with the comprehensive know-how of the consortium partners, formed the basis of the development and implementation work within WP34 in phase 2, integrating all technical tasks outcomes. Special emphasis was drawn on signal anomaly detection for dense time series for both, optical and SAR data, as well as on multi-temporal classification.

As the land cover dynamics heavily depend on the specific land cover category, there is no single change detection approach, which can be recommended for application in all thematic fields, covered within ECoLaSS. For example, the temporal characteristics of the reflectance trajectory are different for forest, cropland types, grassland, water and settlement. Therefore, the change detection testing and benchmarking was performed separately for different thematic fields.

The **forest** change testing investigated the potential of Sentinel-1 and Sentinel-2 data for automated forest change detection to be more flexible in areas of frequent cloud cover. The combined usage of Sentinel-1/Sentinel-2 delivered the best detection result, but has the highest implementation cost, while the Sentinel-2 scenario gives slightly worse results with about half the cost. The current methodology is not able to determine the exact date of the change due to the use of time features. For the detection of these specific changes, the combination of the approach with methods that focus on feature dynamics of forest change events within the respective time period may be considered, which is recommended for future investigations. Further research is also required to validate the transferability of the methodology to areas of different geographic conditions and seasonal patterns, as the tests in different sites in phase 2 proved. Besides that, the influence of "technical" non-changes that can be caused by SAR speckle etc. could be further investigated. The integration of SAR coherence could further improve the change detection accuracy although the operational costs and timeliness might be compromised by the heavy pre-processing implied and limitations in hilly areas applications, as was confirmed by tests in the forest and grassland domains.

In the case of the **imperviousness** change studies, a mono-source approach based on one HR sensor Sentinel-2 has been tested. The results show that the mono-source approach doesn't seem to be sufficient. The main issue is that yet the times series is not dense enough to take advantage of the phenology of inter-yearly and of the intra-yearly seasonal dynamics. Further investigation has been explored in the second phase, including a multi-sensor approach, with the input data from Sentinel-1 and Sentinel-2, and in some cases medium resolution PROBA-V. There is still a need to ensure the temporal consistency and comparability between the different time intervals regarding the re-processing of the reference layer 2015 – and the new status layer for 2017.

For **grassland** monitoring, signal anomaly detection approaches are compared to the map-to-map approach where it can be concluded that the combination improves the results. The map-to-map approach strongly depends on the availability of dense time series. The annual features used as input for the grassland probability derivation require enough observations to distinguish between cropland and grassland. For the map-to-map approach, the changes are directly aggregated from the status layers. 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 e.g. ploughed within the last 5 years can be excluded.

In order to exploit the high spatial and temporal resolution of Sentinel-2, the tested **agricultural** change approach targets assessing crop growth condition by the trajectory of variables representing photosynthetic activities such as Green Area Index (GAI) and/or Leaf Area Index (LAI). Surprisingly, the absolute and relative thresholds do not present significantly different results, although the variations in sowing densities appear to affect the index values. The relative threshold method is expected to be more flexible and match more diverse situations at a larger scale. Density, cultivar, management practices are often quite different over large areas and absolute threshold might rapidly prove inaccurate. The main drawback of the relative threshold is that it will output a result regardless of the real vegetation status. With regards to the indices, the MSAVI, NDVI and Hue index show similar growth dynamics. MSAVI proves to yield the most accurate results with lowest SD and MAD in comparison to the other methods. The benchmarking results highlighted the sensitivity of the results to the time series density. Calendar and GDD time scale have been tested using the same dataset and for the different methods. Whatever the method to retrieve the emergence, it was found that the performances are much less sensitive to the time scale than to the methods for both performance metrics. Beyond the time scale, the plant density and the crop row width vary over large regions according to local agriculture practices and meteorological conditions (dry year, drought, etc.). Their respective influence on the NDVI has been assessed carefully and was shown to have a limited impact.

For future applications we thus recommend to combine a signal anomaly detection approach with the map-to-map comparison as demonstrated for the NDVI anomaly detection approach and the change vector approach. The combination of Sentinel-1 and Sentinel-2 data can be recommended for operational application for grassland and forest change detection applications. The change detection in both forest and grassland domains approaches provide operational advantages and transferability. On the one hand, map-to-map change detection is applicable evermore considering the many status products generated in the past, with different resolutions, and the growing availability of land cover maps. Limitations cannot be overcome, although properly documented, the procedure still achieves satisfactory performance. Improvements to the map-to-map approach, on the basis of probability layers of the ECoLaSS higher resolution and comparable 2017/2018 reference status layers and change indicators derived from time features like the minimum NDVI thresholding and change vector approach have been successfully tested and implemented in the demonstration sites in Task 4.

The derived products from WP34 serve as input for WP35 [AD08] where the change indicators are combined with thematic information (existing LC/LU maps), the classification results from WP33 [AD06] and the pre-processed data stacks from WP32 [AD07] to automatically generate thematic updates of the LC/LU maps.

It should be stressed, that the change product qualities rely on the respective status layers accuracy. Although a one year update cycle for operational rollout using S1/S2 features is achievable, the limitations described in WP32 and WP33 apply:

Multitemporal co-registration of S2 needs to be improved, a better DEM for improved topographic normalisation is required and the improvement of cloud masks e.g. from the Sen2Cor processor is necessary.

For future applications we recommend to apply a continuous change detection with yearly updates, however over longer time-spans such as e.g. over five or more years. A major requirement for future monitoring of land-cover changes at the pan-European or global level is the reduction of time between

image acquisition and availability of finally derived information products. Therefore, the focus within WP34 was on fully automated change detection approaches, which are based on dense time series of Sentinel-2 and/or dense time series of Sentinel-1. The aim was to develop transferrable methods in order to achieve operational feasibility for large scale applications and allowing for cost-efficient automation. The tests have proved successful, providing satisfactory accuracies and the methodologies have been finally selected for generation of the incremental updates in Task 4.

## REFERENCES

- Akbari, V., Anfinson, S. N., Dougeris, A. P., and Eltoft, T. (2013). The Hotelling-Lawley Trace Statistic for Change Detection in Polarimetric Sar Data under the Complex Wishart Distribution. 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 4162-4165 <https://doi.org/10.1109/Igarss.2013.6723750>
- Ali I, Cawkwell F, Dwyer E, Barrett B, Green S. 2016. Satellite Remote Sensing of grasslands: from observation to management –a review. *Journal of Plant Ecology* 1-23. doi:10.1093/jpe/rtw005
- Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote sensing*, 5(2), 949-981
- Balzter, H., Gerard, F., George, C., Weedon, G., Grey, W., Combal, B. & Los, S. (2007). Coupling of vegetation growing season anomalies and fire activity with hemispheric and regional-scale climate patterns in central and east Siberia. *Journal of Climate*, 20(15), 3713-3729
- Ban, Y., and Yousif, O. A. (2012). Multitemporal Spaceborne SAR Data for Urban Change Detection in China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4), 1087-1094 <https://doi.org/10.1109/JSTARS.2012.2201135>
- Banskota, A., Kayastha, N., Falkowski, M. J., Wulder, M.A., Froese, R. E. and White, J. C. (2014). Forest Monitoring Using Landsat Time Series Data: A Review, *Canadian Journal of Remote Sensing*, 40(5), 362-384, <http://dx.doi.org/10.1080/07038992.2014.987376>
- Baret, F., Guyot, G., and Major, D. (1989). TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation. *Geoscience and Remote Sensing Symposium, 1989. IGARSS'89. 12th Canadian Symposium on Remote Sensing*, 3, 1355-1358
- Bargiel D., (2017). A new method for crop classification combining time series of radar images and crop phenology information, *Remote Sensing of Environment*, 198, 369-383
- Bartels, S. F., Chen, H. Y., Wulder, M. A., and White, J. C. (2016). Trends in post-disturbance recovery rates of Canada's forests following wildfire and harvest. *Forest Ecology and Management*, 361, 194-207
- Bazi, Y., Bruzzone, L., and Melgani, F. (2005). An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. *Ieee Transactions on Geoscience and Remote Sensing*, 43(4), 874-887 <https://doi.org/10.1109/Tgrs.2004.842441>
- Belgiu M., Csillik O.(2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis, *Remote Sensing of Environment*, 204, 509-523
- Belgiu M., Drăguț L. (2016). Random forest in remote sensing: A review of applications and future directions, *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31
- Bolton, D. K., Coops, N. C., and Wulder, M. A. (2015). Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne lidar data. *Remote Sensing of Environment*, 163, 48–60
- Bovolo, F., and Bruzzone, L. (2005). A detail-preserving scale-driven approach to change detection in multitemporal SAR images. *Ieee Transactions on Geoscience and Remote Sensing*, 43(12), 2963-2972 <https://doi.org/10.1109/TGRS.2005.857987>
- Bovolo, F., Marin, C., and Bruzzone, L. (2013). A Hierarchical Approach to Change Detection in Very High Resolution SAR Images for Surveillance Applications. *Ieee Transactions on Geoscience and Remote Sensing*, 51(4), 2042-2054 <https://doi.org/10.1109/TGRS.2012.2223219>
- Brink A., Bodart C., Brodsky L., Defourny P., Ernst C., Donnay F., Lupi A., Tuckova K. (2014). Anthropogenic pressure in East Africa - monitoring 20 years of land cover changes by means of medium resolution satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 28, 60–69
- Brooks, E.B., Wynne, R.H., Thomas, V.A., Blinn, C.E., Coulston, J.W. (2014). On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 3316–3332 doi:10.1109/TGRS.2013.2272545
- Bruzzone, L., and Serpico, S. (1997). Detection of changes in remotely-sensed images by the selective use of multi-spectral information. *International Journal of Remote Sensing*, 18, 3883-3888

- Buma, B. (2012). Evaluating the utility and seasonality of NDVI values for assessing post-disturbance recovery in a subalpine forest. *Environmental Monitoring and Assessment*, 184(6), 3849–3860
- Canty, M. J., and Nielsen, A. A. (2008). Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sensing of Environment*, 112(3), 1025–1036. <http://doi.org/10.1016/j.rse.2007.07.013>
- Carincotte, C., Derrode, S., and Bourennane, S. (2006). Unsupervised change detection on SAR images using fuzzy hidden Markov chains. *IEEE Transactions on Geoscience and Remote Sensing*, 44(2), 432-441 <https://doi.org/10.1109/Tgrs.2005.861007>
- Celik, T. (2010). A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images. *Signal Processing*, 90(5), 1471-1485 <https://doi.org/10.1016/j.sigpro.2009.10.018>
- Chatelain, F., Tourneret, J. Y., and Inglada, J. (2008). Change detection in multisensor SAR images using bivariate gamma distributions. *IEEE Transactions on Image Processing*, 17(3), 249-258 <https://doi.org/10.1109/Ti2008.916047>
- Chen, S., and Rao, P. (2008). Land degradation monitoring using multi-temporal Landsat TM/ETM data in a transition zone between grassland and cropland of northeast China. *International Journal of Remote Sensing*, 29(7), 2055-2073 [doi.org/10.1080/01431160701355280](https://doi.org/10.1080/01431160701355280)
- Chen, W., Chen, W., Li, J. (2010). Comparison of surface reflectance derived by relative radiometric normalization versus atmospheric correction for generating large-scale landsat mosaics, *Remote Sensing Letters*, 1(2), 103–109
- Chen, W., Moriya, K., Sakai, T., Koyama, L., and Cao, C. (2014). Monitoring of post-fire forest recovery under different restoration modes based on time series Landsat data. *European Journal of Remote Sensing*, 47(1), 153–168
- Chu, T., and Guo, X. (2014): Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: A review. *Remote Sensing*, 6(1), 470–520
- Cohen, W. B.; Yang, Z.; Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. Timesync—Tools for calibration and validation. *Remote Sensing Environment*, 2010, 114, 2911–2924
- Cohen, W.B., Healey, S.P., Yang, Z., Stehman, S. V, Brewer, C.K., Brooks, E.B., Gorelick, N., Huang, C., Hughes, M.J., Kennedy, R.E., Loveland, T.R., Moisen, G.G., Schroeder, T.A., Vogelmann, J.E., Woodcock, C.E., Yang, L., Zhu, Z. (2017). How similar are forest disturbance maps derived from different Landsat time series algorithms? *Forests* 8. doi:10.3390/f8040098
- Cong, N., Wang, T., Nan, H., Ma, Y., Wang, X., Myneni, R. B., & Piao, S. (2013). Changes in satellite-derived spring vegetation green-up date and its linkage to climate in China from 1982 to 2010: a multimethod analysis. *Global change biology*, 19(3), 881-891
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, 37(1), 35-46
- Conradsen, K., Nielsen, A. A., and Skriver, H. (2016). Determining the Points of Change in Time Series of Polarimetric SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(5), 3007-3024 <https://doi.org/10.1109/TGRS.2015.2510160>
- Conradsen, K., Nielsen, A. A., Sehou, J., and Skriver, H. (2003). A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(1), 4-19 <https://doi.org/10.1109/Tgrs.2002.808066>
- Coopin, P., Jonckheere, I., Nackaerts, K., and Muys, B. (2004). Digital change detection methods in ecosystem monitoring: A Review. *International Journal of Remote Sensing*, 25(9), 1565-1596
- Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17(3), 301-306.
- Dams, J., Dujardin, J., Reggers, R., Bashir, I., Canters, F., and Batelaan, O. (2013). Mapping impervious surface change from remote sensing for hydrological modeling. *Journal of Hydrology*, 485, 84-95
- Dams, J., Dujardin, J., Reggers, R., Bashir, I., Canters, F., and Batelaan, O. (2013). Mapping impervious surface change from remote sensing for hydrological modeling. *Journal of Hydrology*, 485, 84-95
- de Beurs, K. M., & Henebry, G. M. (2010). Spatio-temporal statistical methods for modelling land surface phenology. In *Phenological research* (177-208). Springer, Dordrecht



- Frazier, R. J., N. C. Coops, and M. A. Wulder (2015). Boreal Shield Forest Disturbance and Recovery Trends using Landsat Time Series. *Remote Sensing of Environment*, 170, 317-327
- Frazier, R. J., N. C. Coops, M. A. Wulder., T. Hermosilla and J. C. White (2018). Analyzing spatial and temporal variability in short-term rates of post-fire vegetation return from Landsat time series, *Remote Sensing of Environment*, 205, 2018, 32-45, ISSN 0034-4257
- Furby, S., Caccetta, P., Wu, X., and Chia, J. (2008). Continental scale land cover change monitoring in Australia using Landsat imagery. *International Earth Conference: Studying, Modeling and Sense Making of Planet Earth*, (1-8). Mytilene, Lesvos, Greece.
- Gallego, J. (1995). *Sampling Frames of Square Segments*. Ispra: Office for Publications of the E.C. Luxembourg
- Gallego, J. (2004). Area Frames for Land Cover Estimation: Improving the European LUCAS Survey
- Gamba, P., Dell'Acqua, F., and Lisini, G. (2006). Change Detection of Multitemporal SAR Data in Urban Areas Combining Feature-Based and Pixel-Based Techniques. *Ieee Transactions on Geoscience and Remote Sensing*, 44(10), 2820-2827 <https://doi.org/10.1109/TGRS.2006.879498>
- Gao, B. (1996). NDWI — A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266
- Geerken, R. A., (2009). An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 422–431. doi:10.1016/j.isprsjprs.2009.03.001
- Gibbs, B. P. (2011). *Advanced Kalman filtering, least-squares and modeling: a practical handbook*. John Wiley & Sons
- Gobron, N., Knorr, W., Belward, A. S., & Pinty, B. (2009). Fraction of absorbed photosynthetically active radiation (FAPAR). *State of the Climate in*, 91(7)
- Golda, S. A. J., and Murugan, B. B. (2013). Automatic Change Detection Approach Based on Fusion of Multi-temporal Images over Urban Areas. *International Journal of Engineering Research and Applications*, 3(2), 1546-1551
- Gómez, C., White, J. C., Wulder, M. A., and Alejandro, P. (2015). Integrated Object-Based Spatiotemporal Characterization of Forest Change from an Annual Time Series of Landsat Image Composites. *Canadian Journal of Remote Sensing*, 41(4), 271–292
- Gong, M., Zhou, Z., and Ma, J. (2012). Change Detection in Synthetic Aperture Radar Images based on Image Fusion and Fuzzy Clustering. *Ieee Transactions on Image Processing*, 21(4), 2141-2151 <https://doi.org/10.1109/TIP.2011.2170702>
- Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, 151, 72–88
- Griffiths, P., Hostert, P. (2015). *Remote Sensing Time Series: Revealing Land Surface Dynamics*, chapter 15: Forest Cover Dynamics During Massive Ownership Changes - Annual Disturbance Mapping Using Annual Landsat Time-Series, 307–322. Springer International Publishing
- Grogan, K., Pflugmacher, D., Hostert, P., Kennedy, R., Fensholt, R. (2015). Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series. *Remote Sensing Environment*, 169, 438–453. [http:// dx.doi.org/10.1016/j.rse.2015.03.001](http://dx.doi.org/10.1016/j.rse.2015.03.001)
- Hachicha, S., and Chaabane, F. (2014). On the SAR change detection review and optimal decision. *International Journal of Remote Sensing*, 35(5), 1693-1714 <https://doi.org/10.1080/01431161.2014.882030>
- Hamunyela, E., Verbesselt, J., Herold, M. (2016). Using spatial context to improve early detection of deforestation from Landsat time series. *Remote Sensing Environment*, 172, 126–138 <http://dx.doi.org/10.1016/j.rse.2015.11.006>.
- Hansen, M C., Egorov, A., Potapov, P V., Stehman, S V., Tyukavina, A, Turubanova, S A., Roy, D. P., Goetz, S J., Loveland, T R., Ju, J, Kommareddy, A., Kovalskyy, V., Forsyth, C., Bents, T. (2014). Monitoring conterminous United States (CONUS) land cover change with Web-Enabled Landsat Data (WELD), *Remote Sensing of Environment*, 140, 466-484 <http://dx.doi.org/10.1016/j.rse.2013.08.014>

- Hansen, M. C., Krylov, A., Tyukavina, A., Potapov, P. V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R. (2016). Humid tropical forest disturbance alerts using Landsat data. *Environ. Res. Lett.* 11, 34008 <http://dx.doi.org/10.1088/1748-9326/11/3/034008>
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O. and Townshend, J. R. G. (2013): High-Resolution Global Maps of 21st-Century Forest Cover Change, *Science* 342 (15 November), 850–853. Data available on-line from: <http://earthenginepartners.appspot.com/science-2013-global-forest>
- Hansen, M., Roy, D. L., Adusei, B., Justice, C., and Alstatt, A. (2008). A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Conga Basin. *Remote Sensing of Environment*, 112, 2495-2513
- Hao, P., Niu, Z., Zhan, Y. and Wu, Y., Wang, L., and Liu, Y. (2015). Spatiotemporal changes of urban impervious surface area and land surface temperature in Beijing from 1990 to 2014. *GIScience & Remote Sensing*. 1-22
- Haralick, R., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *Proceedings of the IEEE*, 5(41), 786-804
- Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press
- Hermosilla, T., Wulder, M. A. , White, J. C., Coops, N. C., Hobart, G. W. and Campbell, L. B. (2016). Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring, *International Journal of Digital Earth*, DOI: 10.1080/17538947.2016.1187673
- Hirschmugl, M., Gallaun, H., Dees, M., Datta, P., Deutscher, J., Koutsias, N., Schardt, M. (2017). Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a review. *Current Forest Reports*, 3, 32–45 doi:10.1007/s40725-017-0047-2
- Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J. and Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States—representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5), 345-354
- Hu, H., and Ban, Y. (2014). Unsupervised Change Detection in Multitemporal SAR Images Over Large Urban Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8), 3248-3261 <https://doi.org/10.1109/JSTARS.2014.2344017>
- Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., Vogelmann, J. E., (2010a). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, 114, 183–198 doi:10.1016/j.rse.2009.08.017
- Huang, C., Thomas, N., Goward, S. N., Masek, J. G., Zhu, Z., Townshend, J. R. G., Vogelmann, J. E., (2010b). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images, *International Journal of Remote Sensing*, 31, 5449–5464, October 2010.
- Huang, L. K., and Wang, M. J. J. (1995). Image Thresholding by Minimizing the Measures of Fuzziness. *Pattern Recognition*, 28(1), 41-51 [https://doi.org/10.1016/0031-3203\(94\)E0043-K](https://doi.org/10.1016/0031-3203(94)E0043-K)
- Hüttich, C., Gessner, U., Herold, M., Strohbach, B., Schmidt, M., Keil, M., & Dech, S. (2009). On the suitability of MODIS time series metrics to map vegetation types in dry savanna ecosystems: A case study in the Kalahari of NE Namibia. *Remote sensing*, 1(4), 620-643
- Inglada, J., and Mercier, G. (2007). A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5), 1432-1445 <https://doi.org/10.1109/Tgrs.2007.893568>
- Inglada, J., Vincent, A., Arias, M., Tardt, B., Morin, D., and Rodes, I. (2017). Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing*, 9(9), 1-35
- Irish, R. R., Barker, J. L., Goward, S. N., Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. *Photogrammetric Engineering Remote Sensing*, 72, 1179–1188 doi:10.14358/PERS.72.10.1179
- Julea A., Méger N., Rigotti C., Trouvé E., Bolon P., Lăzărescu V. (2011). Mining pixel evolutions in satellite image time series for agricultural monitoring. In: Perner, P. (Ed.), *Advances in Data Mining*.

- Applications and Theoretical Aspects chapter Mining Pixel Evolutions in Satellite Image Time Series for Agricultural Monitoring. Springer Berlin Heidelberg, Berlin, Heidelberg, 189–203
- Kandasamy, S., Baret, F., Verger, a., Neveux, P., Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to MODIS LAI products. *Biogeosciences* 10, 4055–4071 doi:10.5194/bg-10-4055-2013
- Kaufman, Y., and Tanré, D. (1992). Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2), 261-270
- Kemper, T., Mudau, N., Mangara, P., and Pesaresi, M. (2015). Towards an automated monitoring of human settlements in South Africa using high resolution SPOT satellite imagery. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(7), 1389
- Kennedy, R. E., Yang, Z., Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. *Remote Sensing of Environment*, 114, 2897–2910 doi:10.1016/j.rse.2010.07.008
- Kennedy, R. E., Yang, Z., Cohen, W. B., Pfaff, E., Braaten, J., and Nelson, P. (2012). Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sensing of Environment*, 122, 117–133
- Kindu, M., Schneider, T., Teketay, T., Knoke, T. (2013). Land Use/Land Cover Change Analysis Using Object-Based Classification Approach in Munessa-Shashemene Landscape of the Ethiopian Highlands. *Remote Sensing*, 5(5), 2411 – 2435 <http://www.mdpi.com/2072-4292/5/5/2411/html>  
Last access: February 2018.
- Kittler, J., and Illingworth, J. (1986). Minimum Error Thresholding. *Pattern Recognition*, 19(1), 41-47 [https://doi.org/10.1016/0031-3203\(86\)90030-0](https://doi.org/10.1016/0031-3203(86)90030-0).
- Kleynhans, W. & Olivier, J.C. & Wessels, K.J. & Salmon, B.P. & Bergh, F. & Steenkamp, Karen. (2011). Detecting land cover change using an extended Kalman filter on MODIS NDVI time series data.. *Geoscience and Remote Sensing Letters, IEEE*. 8. 507 - 511 10.1109/LGRS.2010.2089495.
- Kuenzer, C., Dech, S., Wagner, W. (2015b). *Remote Sensing Time Series: Revealing Land Surface Dynamics*, ch. 1: Remote Sensing Time Series Revealing Land Surface Dynamics: Status Quo and the Pathway Ahead, 1–24. Springer International Publishing.
- Kuenzer, C., Dech, S., Wagner, W., eds., (2015a). *Remote Sensing Time Series: Revealing Land Surface Dynamics*. Springer International Publishing. Comprehensive review of different time series approaches for different data sets and applications from around the globe.
- Lefebvre, A., Sannier, C., and Corpetti, T. (2016). Monitoring Urban Areas with Sentinel-2A Data: Application to the Update of the Copernicus High Resolution Layer Imperviousness Degree. *Remote Sensing*, 8(606), 1-21
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764-766
- Liu, H., and Huete, A. (1995). Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2), 457-465
- Liu, Y., Zha, Y., Gao, J., and Ni, S. (2004). Assessment of grassland degradation near Lake Qinghai, West China, using Landsat TM and in situ reflectance spectra data. *International Journal of Remote Sensing*, 25(20), 4177-4189
- Lobell, D. B., Hammer, G. L., McLean, G., Messina, C., Roberts, M. J., & Schlenker, W. (2013). The critical role of extreme heat for maize production in the United States. *Nature Climate Change*, 3(5), 497
- López-Sánchez, J. M., Ballester-Berman, J. D., Navarro-Sanchez, V. D., & Vicente-Guijalba, F. (2012, July). Experimental validation of the interferometric coherence formulation in single-transmit mode. In 2012 IEEE International Geoscience and Remote Sensing Symposium (3114-3117). IEEE.
- Lu, D., Moran, E., and Hetrick, S. (2011). Detection of impervious surface change with multitemporal Landsat images in an urban–rural frontier. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 298-306
- Marinho, E., Vancutsem, C., Fasbender, D., Kayitakire, F., Pini, G., & Pekel, J. F. (2014). From Remotely Sensed Vegetation Onset to Sowing Dates: Aggregating Pixel-Level Detections into Village-Level Sowing Probabilities. *Remote Sensing*, 6(11), 10947-10965

- Marino, A., Cloude, S. R., and Lopez-Sanchez, J. M. (2013). A New Polarimetric Change Detector in Radar Imagery. *Ieee Transactions on Geoscience and Remote Sensing*, 51(5), 2986-3000 <https://doi.org/10.1109/Tgrs.2012.2211883>
- Marsett, R. C., Qi, J., Heilman, P., Biedenbender, S. H., Watson, M. C., Amer, S., Weltz, M., Goodrich, D., Marsett, R. (2006). Remote Sensing for Grassland Management in the Arid Southwest. *Rangeland Ecology and Management*, 59 (5), 530 – 540 [doi.org/10.2111/05-201R.1](https://doi.org/10.2111/05-201R.1)
- Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J., Lim, T.-K. (2006). A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geoscience and Remote Sensing Letters* 3, 68–72 doi:10.1109/LGRS.2005.857030
- Mathieu, R., Pouget, J., Cervelle, B., and Escadafal, R. (1998). Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment. *Remote Sensing of Environment*, 66, 17-28
- Meier U., Bleiholder H. (2006). Growth stages of mono-and dicotyledonous plants. BBCH Monograph, Federal Biological Research Centre for Agriculture and Forestry.
- Melgani, F., Moser, G., and Serpico, S. B. (2002). Unsupervised change-detection methods for remote-sensing images. *Optical Engineering*, 41, 3288 - 3297 <https://doi.org/10.1117/1.1518995>
- Melville, B., Lucieer, A., and Aryal, J. (2018). Object-based random forest classification of Landsat ETM+ and WorldView-2 satellite imagery for mapping lowland native grassland communities in Tasmania, Australia. *International Journal of Applied Earth Observation and Geoinformation*, 66, 46-55
- Moser, G., and Serpico, S. B. (2006). Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery. *Ieee Transactions on Geoscience and Remote Sensing*, 44(10), 2972-2982 <https://doi.org/10.1109/TGRS.2006.876288>
- Moser, G., and Serpico, S. B. (2009). Unsupervised Change Detection From Multichannel SAR Data by Markovian Data Fusion. *Ieee Transactions on Geoscience and Remote Sensing*, 47(7), 2114-2128 <https://doi.org/10.1109/Tgrs.2009.2012407>
- Moser, G., Melgani, F., and Serpico, S. B. (2003). Advances in unsupervised change detection. In C. H. Chen (Ed.), *Frontiers of Remote Sensing Information Processing* (403-426). Singapore: World Scientific
- Moser, G., Serpico, S., and Vernazza, G. (2007). Unsupervised change detection from multichannel SAR images. *Ieee Geoscience and Remote Sensing Letters*, 4(2), 278-282. <https://doi.org/10.1109/Lgrs.2007.890549>
- Moulin, S., Kergoat, L., Viovy, N., & Dedieu, G. (1997). Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. *Journal of Climate*, 10(6), 1154-1170.
- Nascimento, A. D., Frery, C. and Cintra, R. J. Detecting Changes in Fully Polarimetric SAR Imagery With Statistical Information Theory (2019), *IEEE Transactions on Geoscience and Remote Sensing*, 57(3), 1380-1392, doi: 10.1109/TGRS.2018.2866367
- Nguyen, H. T., Soto-Berelov, M., Jones, S. D., Haywood, A., and Hislop, S. (2017). Mapping forest disturbance and recovery for forest dynamics over large areas using Landsat time-series remote sensing. *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIX*, (November), 35.
- Nguy-Robertson, A. L., Peng, Y., Gitelson, A. A., Arkebauer, T. J., Pimstein, A., Herrmann, I. & Bonfil, D. J. (2014). Estimating green LAI in four crops: Potential of determining optimal spectral bands for a universal algorithm. *Agricultural and forest meteorology*, 192, 140-148
- Oeser, J., Pflugmacher, D., Senf, C., Heurich, M., Hostert, P. (2017). Using intra-annual Landsat time series for attributing forest disturbance agents in central Europe. *Forests* 8, 251. doi:10.3390/f8070251
- Olsson, P. O., Lindström, J., & Eklundh, L. (2016). Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote sensing of environment*, 181, 42-53
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66 <https://doi.org/10.1109/TSMC.1979.4310076>
- Palacios-Orueta, A., Huesca, M., Whiting, M. L., Litago, J., Khanna, S., Garcia, M., & Ustin, S. L. (2012). Derivation of phenological metrics by function fitting to time-series of Spectral Shape Indexes AS1 and AS2: Mapping cotton phenological stages using MODIS time series. *Remote sensing of environment*, 126, 148-159

- Pasquarella, V., Bradley, B., Woodcock, C. (2017). Near-real-time monitoring of insect defoliation using Landsat time series. *Forests* 8, 275 doi:10.3390/f8080275
- Pekel, J. F., Ceccato, P., Vancutsem, C., Cressman, K., Vanbogaert, E., & Defourny, P. (2011). Development and Application of Multi-Temporal Colorimetric Transformation to Monitor Vegetation in the Desert Locust Habitat. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(4), 318-326
- Permatasari, P. A., Fatikhunnada, A., Setiawan, Y. Nurdiana, A. (2016). Analysis of agricultural land use changes in Jombang Regency, East Java, Indonesia using BFAST method. *Procedia Environmental Sciences*, 33, 27-35
- Pesaresi, M., Gerhardinger, A., and Kayitakire, F. (2008). A Robust Built-Up Area Presence Index by Anisotropic Rotation-Invariant Textural Measure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(3), 180-192
- Petitjean, F., Inglada, J., and Gancarski, P. (2014). Satellite image time series analysis under time warping. *IEEE Trans. Geosci. Remote Sensing*, 50(8), 3081-3095
- Pflugmacher, D., Cohen, W. B., Kennedy, R. E. (2012). Using landsat-derived disturbance history (1972–2010) to predict current forest structure, *Remote Sensing of Environment*, 122, 146–165, July 2012
- Pickell, P. D., Hermosilla, T., Frazier, R. J., Coops, N. C., Wulder, M. A. (2016). Forest re-recovery trends derived from Landsat time series for North American boreal forests. *Int. J. Remote Sensing*, 37(1), 138–149
- Powell, S. L., Cohen, W. B., Yang, Z., Pierce, J. D., & Alberti, M. (2008). Quantification of impervious surface in the Snohomish water resources inventory area of western Washington from 1972–2006. *Remote Sensing of Environment*, 112(4), 1895-1908
- Pratola, C., Del Frate, F., Schiavon, G., and Solimini, D. (2013). Toward Fully Automatic Detection of Changes in Suburban Areas From VHR SAR Images by Combining Multiple Neural-Network Models. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4), 2055-2066 <https://doi.org/10.1109/Tgrs.2012.2236846>
- Qi, J., Chehbouni, A., Huete, A., Kerr, Y., and Sorooshian, S. (1994). A Modified Soil Adjusted Vegetation Index. *Remote Sensing Environment*, 48(2), 119-126
- Qong, M. (2004). Polarization State Conformation and Its Application to Change Detection in Polarimetric SAR Data. *IEEE Geoscience and Remote Sensing Letters*, 1(4), 304-308 <https://doi.org/10.1109/Lgrs.2004.835221>
- Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of vegetation science*, 5(5), 703-714.
- White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global biogeochemical cycles*, 11(2), 217-234
- Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M. (2018). Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2, *Remote Sensing of Environment*, 204, 147-161, ISSN 0034-4257
- Reiche, J., Verbesselt, J., Hoekman, D., Herold, M. (2015). Fusing Landsat and SAR time series to detect deforestation in the tropics. *Remote Sensing of Environment*, 156, 276–293 doi:10.1016/j.rse.2014.10.001
- Richardson, A., and Weigand, C. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43
- Rignot, E. J. M., and van Zyl, J. J. (1993). Change detection techniques for ERS-1 SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 31(4), 896-906 <https://doi.org/10.1109/36.239913>
- Rouse Jr, J., Haas, R., and Schell, J. (1974). Monitoring vegetation systems in the great plains with erts. *NASA Special Publication*, 351, 309
- Sakamoto, T., Wardlow, B. D., Gitelson, A. A., Verma, S. B., Suyker, A. E., & Arkebauer, T. J. (2010). A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sensing of Environment*, 114 (10), 2146-2159

- Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A crop phenology detection method using time-series MODIS data. *Remote sensing of environment*, 96(3-4), 366-374
- Schlaepfer, D., Borel, C. C., Keller, J., Itten, K. I. (2016). Atmospheric precorrected differential absorption technique to retrieve columnar water vapour, *Remote Sensing of Environment*, 65, 353–366, 1998
- Schmitt, A., and Brisco, B. (2013). Wetland Monitoring Using the Curvelet-Based Change Detection Method on Polarimetric SAR Imagery. *Water*, 5(3), 1036-1051 <https://doi.org/10.3390/w5031036>
- Schroeder, T. A., Cohen, W. B., and Yang, Z. (2007). Patterns of forest regrowth following clearcutting in western Oregon as determined from a Landsat time-series. *Forest Ecology and Management*, 243(2–3), 259–273
- Schultz, M., Verbesselt, J., Avitabile, V., Souza, C., Herold, M. (2016). Error sources in deforestation detection using BFAST Monitor on Landsat time series across three tropical sites. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 9, 3667–3679 doi:10.1109/JSTARS.2015.2477473
- Shen, G., Guo, H., and Liao, J. (2007). Change vector analysis method for inundation change detection using multi-temporal multi-polarized SAR images - art. no. 679007. *Remote Sensing and Gis Data Processing and Applications; and Innovative Multispectral Technology and Applications*, Pts 1 and 2, 6790, 79007-79007 <https://doi.org/10.1117/12.739265>
- Shi, L., Ling, F., Ge, Y., Foody, G., Li, X., Wang, L., Zhang, Y. and Du, Y. (2017). Impervious surface change mapping with an uncertainty-based spatial-temporal consistency model: A case study in wuhan city using landsat time-series datasets from 1987 to 2016. *Remote Sensing*, 9(11), 1148
- Stehman, S., and Czaplewski, R. (1998). Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sensing of Environment*, 34, 331-344
- Sun, B., Li, Z., Gao, Z., Guo, Z., Wang, B., Hu, X., and Bai, L. (2017). Grassland degradation and restoration monitoring and driving forces analysis based on long time-series remote sensing data in Xilin Gol League. *Acta Ecologica Sinica*, 37(4), 219-228
- Tarantino, C., Adamo, M., Lucas, R., and Blonda, P. (2016). Detection of changes in semi-natural grasslands by cross correlation analysis with WorldView-2 images and new Landsat 8 data. *Remote sensing of environment*, 175, 65-72
- Touzi, R., Lopes, A., and Bousquet, P. (1988). A Statistical and Geometrical Edge Detector for Sar Images. *IEEE Transactions on Geoscience and Remote Sensing*, 26(6), 764-773 <https://doi.org/10.1109/36.7708>
- Turner, D., Cohen, W., and Kennedy, R. (1999). Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing Environment*, 70, 52-68
- Varmuza, K., & Filzmoser, P. (2016). *Introduction to multivariate statistical analysis in chemometrics*. CRC press
- Verbesselt, J., Herold, M., Zeileis, A. (2012). Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, 123, 98–108 doi:10.1016/j.rse.2012.02.022
- Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D. (2010a). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114, 106–115 doi:10.1016/j.rse.2009.08.014
- Verbesselt, J., Hyndman, R., Zeileis, A., Culvenor, D. (2010b). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sensing of Environment*, 114, 2970–2980 doi:10.1016/j.rse.2010.08.003
- Verhegghen, A. (2013). *Global land surface vegetation phenology using 13 years of SPOT VEGETATION daily observations* (Doctoral dissertation, UCL-Université Catholique de Louvain)
- Vogelmann, J. E., Gallant, A. L., Shi, H., Zhu, Z. (2016). Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data. *Remote Sensing of Environment*, 185, 258–270 doi:10.1016/j.rse.2016.02.060
- Vogelmann, J. E., Khoa, P., Lan, D., Shermeyer, J., Shi, H., Wimberly, M., Duong, H., Huong, L., (2017). Assessment of forest degradation in Vietnam using Landsat time series data. *Forests* 8, 238 doi:10.3390/f8070238

- Vogelmann, J. E., Xian, G., Homer, C., Tolk, B. (2012). Monitoring gradual ecosystem change using Landsat time series analyses: Case studies in selected forest and rangeland ecosystems, *Remote Sensing of Environment*, 122, 92–105
- Waldner F., Lambert M.-J., Li W., Weiss M., Demarez V., Morin M., Marais-Sicre C., Hagolle O., Baret F. and Defourny P. (2015). Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series, *Remote Sensing* 2015, 7, 10400-10424
- Wang, P., Huang, C, Brown de Colstoun, E.C. Mapping 2000–2010 Impervious Surface Change in India Using Global Land Survey Landsat Data. *Remote Sensing*. 2017; 9(4):366
- Wang, Z., Yao, W., Tang, Q., Liu, L., Xiao, P., Kong, X., Zhang, P., Shi, F., Wang, Y. (2018) Continuous Change Detection of Forest/Grassland and Cropland in the Loess Plateau of China Using All Available Landsat Data. *Remote Sensing*, 10, 1775
- White, J. C., M. A. Wulder, T. Hermosilla, N. C. Coops, G. W. Hobart (2017). A nationwide annual characterization of 25years of forest disturbance and recovery for Canada using Landsat time series, *Remote Sensing of Environment*, 194, 303-321
- White, J. D., Running, S. W., Nemani, R., Keane, R. E., & Ryan, K. C. (1997). Measurement and remote sensing of LAI in Rocky Mountain montane ecosystems. *Canadian Journal of Forest Research*, 27(11), 1714-1727
- Wu, K., Yi, W., Niu, R., and Wei, L. (2015). Subpixel land cover change mapping with multitemporal remote-sensed images at different resolution. *Journal of Applied Remote Sensing*, 9(1), 097299.
- Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, 185, 271–283 doi:10.1016/j.rse.2015.11.032
- Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, 185, 271–283 doi:10.1016/j.rse.2015.11.032
- Wylie, D. P. and Menzel, W. P. (1999). Eight years of cloud statistics using HIRS. *Journal of Climate* 12, 170–184
- Xia, J., Du, P., Zhang, H., & Yuan, L. (2011, April). Monitoring urban impervious surface area change using CBERS and HJ-1 remote sensing images. In 2011 Joint Urban Remote Sensing Event (pp. 177-180) IEEE
- Xian, G., and Homer, C. (2010). Updating the 2001 National Land Cover Database impervious surface products to 2006 using Landsat imagery change detection methods. *Remote Sensing of Environment*, 114(8), 1676-1686
- Yang, A. and Sun, G. (2017). Landsat-Based Land Cover Change in the Beijing-Tianjin-Tangshan Urban Agglomeration in 1990, 2000 and 2010. *ISPRS Int. J. Geo-Inf.* 6(3), 59
- Yousif, O., and Ban, Y. F. (2014). Improving SAR-Based Urban Change Detection by Combining MAP-MRF Classifier and Nonlocal Means Similarity Weights. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(10), 4288-4300 <https://doi.org/10.1109/Jstars.2014.2347171>
- Yu, L., Liu, T., Bu, K., Yan, F., Yang, J., Chang, L. and Zhang, S (2017) Monitoring the long term vegetation phenology change in Northeast China from 1982 to 2015. *Scientific Reports*, 7(1), 14770 doi:10.1038/s41598-017-14918-4
- Zha, Y., Gao, J., and Ni, S. (2003). Use of Normalized Difference Built-Up Index in Automatically Mapping Urban Areas from TM Imagery. *International Journal of Remote Sensing*, 24(3), 583-594
- Zanchetta, A., & Bitelli, G. (2017). A combined change detection procedure to study desertification using opensource tools. *Open Geospatial Data, Software and Standards*, 2(1), 10
- Zeng, L., Wardlow, B. D., Wang, R., Shan, J., Tadesse, T., Hayes, M. J., & Li, D. (2016). A hybrid approach for detecting corn and soybean phenology with time-series MODIS data. *Remote sensing of environment*, 181, 237-250
- Zha, Y., and Gao, J. (2011). Quantitative detection of change in grass cover from multi-temporal TM satellite data. *International journal of remote sensing*, 32(5), 1289-1302

- Zha, Y., Gao, J., and Ni, S. (2003). Use of Normalized Difference Built-Up Index in Automatically Mapping Urban Areas from TM Imagery. *International Journal of Remote Sensing*, 24(3), 583-594
- Zhou, B., He, H. S., Nigh, T. A., and Schulz, J. H. (2012). Mapping and analyzing change of impervious surface for two decades using multi-temporal Landsat imagery in Missouri. *International Journal of Applied Earth Observation and Geoinformation*, 18, 195-206
- Zhu, Z. (2017). Change detection using landsat time series: A review of frequencies, pre-processing, algorithms, and applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 130, 370–384 doi:10.1016/j.isprsjprs.2017.06.013
- Zhu, Z., & Woodcock, C. E. (2014a). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152, 217-234
- Zhu, Z., & Woodcock, C. E. (2014b). Continuous change detection and classification of land cover using all available Landsat data. *Remote sensing of Environment*, 144, 152-171
- Zhu, Z., Fu, Y., Woodcock, C. E., Olofsson, P., Vogelmann, J.E., Holden, C., Wang, M., Dai, S., Yu, Y. (2016). Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). *Remote Sensing of Environment*, 185, 243–257 doi:10.1016/j.rse.2016.03.036
- Zhu, Z., Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83–94 doi:10.1016/j.rse.2011.10.028
- Zhu, Z., Wang, S., Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsat 4-7, 8, and Sentinel 2 images, *Remote Sensing of Environment*, 159, 269 – 277