

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



D7.2

D32.1b – Methods Compendium: Time Series Preparation

Issue/Rev.:2.0

Date Issued: 15.05.2019

submitted by:



in collaboration with the consortium partners:



submitted to:



European Commission – Research Executive Agency

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

CONSORTIUM PARTNERS

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

CONTACT:

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

DISCLAIMER:

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

DOCUMENT RELEASE SHEET

	NAME, FUNCTION	DATE	SIGNATURE
Author(s):	Heinz Gallau (JR) Mathias Schardt (JR) Petra Miletich (JR) Janik Deutscher (JR) Klaus Granica (JR) Martin Puhm (JR) Larissa Posch (JR) Annekatrin Metz-Marconcini (DLR) Benjamin Leutner (DLR) Igor Klein (DLR) Carolin Sommer (GAF) David Herrmann (GAF) Sophie Villerot (SIRS) Élie Khalil (UCL)	29/04/2019	
Review:	Petra Miletich (JR) Katharina Schwab (GAF)	30/04/2019	
Approval:	Eva Sevillano Marco (GAF)	15/05/2019	
Acceptance:	Massimo Ciscato		
Distribution:			

DISSEMINATION LEVEL

DISSEMINATION LEVEL		
PU	Public	<input checked="" type="checkbox"/> X
CO	Confidential: only for members of the consortium (including the Commission Services)	

DOCUMENT STATUS SHEET

ISSUE/REV	DATE	PAGE(S)	DESCRIPTION / CHANGES
1.0	28.02.2018	65	First delivery of the Methods Compendium: Time Series Preparation
2.0	14.05.2019	104	Second issue of the Methods Compendium: Time Series Preparation with major updates of chapters 2.1, 2.2, 3.2, 3.3, 4 and 5.

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	D21.1a - Service Evolution Requirements Report, Issue 1.0, Issued: 09.08.2017
AD06	D22.1b - EO and other Data Requirements Report, Issue 2.0, Issued: 21.02.2019
AD07	D23.1b - Service Infrastructure/Architecture Requirements Report, Issue 2.0, Issued: 18.02.2019
AD08	D31.1 - Methods Compendium: Sentinel-1/2/3 Integration Strategies
AD09	D33.1 - Methods Compendium: Time Series Analysis for Thematic Classification
AD10	D34.1 - Methods Compendium: Time Series Analysis for Change Detection
AD11	D35.1 - Methods Compendium: HRL Time Series Consistency for HRL Product (incremental) Updates
AD12	D41.1 - Prototype Report: Time Series derived Indicators and Variables
AD13	D42.1 - Prototype Report: Consistent HR Layer Time Series/Incremental Updates
AD14	D43.1 - Prototype Report: Improved Permanent Grassland
AD15	D44.1 - Prototype Report: Crop Area and Crop Status/Parameters
AD16	D45.1 - Prototype Report: New LC/LU Products
AD17	Technical Note – Proposed substitutes for Sentinel-3 data. Issue 1.0, Issued: 04.12.2018

EXECUTIVE SUMMARY

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

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

The Deliverable “D32.1b – Methods Compendium: Time Series Preparation” analyses methods for fully automated pre-processing of Sentinel time series data streams, which are a prerequisite for all further processing tasks. The quality of the pre-processing procedure determines the thematic quality, content, and accuracy that can be achieved in any subsequent information extraction processes. The work of WP32 is focused on time series preparation of Sentinel-1 and Sentinel-2 data for the defined test sites. The focus of WP32 are comprehensive tests of optical as well as SAR pre-processing methods and the implementation of pre-operational processing lines which allow pre-processing of Sentinel-1 and Sentinel-2 data streams following the requirements as requested in AD06.

Following an Introduction in chapter 1, chapter 2 presents a comprehensive State-of-the-Art review of pre-processing methods for optical and SAR time series data streams. Each Sentinel mission brings new opportunities and challenges for the time series data processing lines. The quality of the pre-processing procedure determines the thematic quality, content, and accuracy that can be achieved in any subsequent information extraction processes. Therefore, pre-processing approaches for high volume data processing based on the high resolution satellite constellations are addressed.

Chapter 3 provides an overview of the test and demonstration sites defined followed by the description of testing and benchmarking applied. The thematic focus of the testing and benchmarking is on existing pre-processing methods for a) atmospheric correction, b) cloud, shadow and snow masking, c) topographic normalisation and d) the geometric consistency of the products with regards to Sentinel-2 data. Furthermore, testing and benchmarking of existing pre-processing methods for Sentinel-1 data is performed focusing on a) radiometric calibration, b) speckle noise reduction and d) interferometric coherence estimation.

Chapter 4 gives a description of the generic pre-processing lines for Sentinel-1 and Sentinel-2 data streams which can be adapted according to the user needs. The main challenge requested as research need is the adaptation and up-scaling of current processing lines from local level to a continental level. Accordingly, a large volume of data should be pre-processed, to be able to produce data composites and derive calibrated and validated variables. Further detailed descriptions of the pre-processing lines applied in the different test and demonstration sites in respect to the site and the thematic applications are given.

A conclusion and outlook is finally presented in chapter 5, which summarizes the findings of the testing and benchmarking applied and gives recommendations for all pre-processing steps.

Table of Contents

1	INTRODUCTION.....	1
2	STATE-OF-THE-ART.....	3
2.1	PROCESSING METHODS FOR OPTICAL TIME SERIES	3
2.1.1	<i>Atmospheric correction.....</i>	3
2.1.2	<i>Cloud, Cloud Shadow and Snow Masking.....</i>	6
2.1.3	<i>Topographic normalisation.....</i>	9
2.1.4	<i>Geometric Consistency</i>	10
2.1.5	<i>Noise reduction and gap filling</i>	12
2.1.6	<i>Sentinel-2/Sentinel-3 time series fusion</i>	13
2.2	PROCESSING METHODS FOR SAR TIME SERIES.....	14
2.2.1	<i>Geometric pre-processing.....</i>	14
2.2.2	<i>Radiometric pre-processing</i>	15
2.2.3	<i>Speckle noise reduction</i>	17
2.2.4	<i>Interferometric coherence estimation</i>	18
2.2.5	<i>Available SAR processing software tools</i>	20
3	TESTING AND BENCHMARKING OF CANDIDATE METHODS	23
3.1	TEST AND DEMONSTRATION SITES	23
3.2	BENCHMARKING OF METHODS FOR PRE-PROCESSING OF OPTICAL TIME SERIES.....	25
3.2.1	<i>Benchmarking of candidate approaches for atmospheric correction.....</i>	25
3.2.1.1	Description of Candidate Methods	26
3.2.1.2	Benchmarking Criteria.....	26
3.2.1.3	Implementation and Results of Benchmarking	27
3.2.1.4	Summary and Conclusions	31
3.2.2	<i>Benchmarking of candidate approaches for cloud, cloud shadow and snow masking</i>	32
3.2.2.1	Description of Candidate Methods	32
3.2.2.2	Benchmarking Criteria.....	33
3.2.2.3	Implementation and Results of Benchmarking	34
3.2.2.4	Summary and Conclusions	41
3.2.3	<i>Benchmarking of candidate approaches for topographic normalisation</i>	41
3.2.3.1	Description of Candidate Methods	41
3.2.3.2	Benchmarking Criteria.....	42
3.2.3.3	Implementation and Results of Benchmarking	42
3.2.3.4	Summary and Conclusions	49
3.2.4	<i>Benchmarking of candidate approaches for Geometric consistency</i>	49
3.2.4.1	Description of Candidate Methods	49
3.2.4.2	Benchmarking Criteria.....	50
3.2.4.3	Implementation and Results of Benchmarking	50
3.2.4.4	Summary and Conclusions	52
3.2.5	<i>Benchmarking of candidate approaches for noise reduction and gap filling</i>	53
3.2.5.1	Description of Candidate Methods	53
3.2.5.2	Benchmarking Criteria.....	54
3.2.5.3	Implementation and Results of Benchmarking	55
3.2.5.4	Summary and Conclusions	57
3.3	BENCHMARKING OF METHODS FOR PRE-PROCESSING OF SAR TIME SERIES.....	57
3.3.1	<i>Benchmarking of candidate approaches for radiometric pre-processing.....</i>	58
3.3.1.1	Description of Candidate Methods	58
3.3.1.2	Benchmarking Criteria.....	58
3.3.1.3	Implementation and Results of Benchmarking	58
3.3.1.4	Summary and Conclusions	59
3.3.2	<i>Benchmarking of candidate approaches for speckle noise reduction.....</i>	59
3.3.2.1	Description of Candidate Methods	59
3.3.2.2	Benchmarking Criteria.....	60
3.3.2.3	Implementation and Results of Benchmarking	60
3.3.2.4	Summary and Conclusions	64
3.3.3	<i>Benchmarking of candidate approaches for interferometric coherence estimation</i>	64
3.3.3.1	Description of Candidate Methods	64

3.3.3.2	Benchmarking Criteria.....	65
3.3.3.3	Implementation and Results of Benchmarking	65
3.3.3.4	Summary and Conclusions	72
4	PRE-PROCESSING OF ECOLASS TEST SITES	73
4.1	PRE-PROCESSING OF OPTICAL DATA	73
4.1.1	<i>Pre-processing of test sites/demo sites</i>	74
4.1.1.1	Central (Austria/ Germany).....	74
4.1.1.2	South-East (Greece/Bulgaria).....	76
4.1.1.3	West (Belgium).....	77
4.1.1.4	South-West (France)	77
4.1.1.5	Northern (Sweden).....	78
4.2	PRE-PROCESSING OF SAR DATA	80
4.2.1	<i>Pre-processing of test sites/demo sites</i>	82
4.2.1.1	Central (Austria/ Germany).....	82
4.2.1.2	South-East (Greece/Bulgaria).....	84
4.2.1.3	West (Belgium).....	84
4.2.1.4	South-West (France)	84
4.2.1.5	Northern (Sweden).....	85
5	CONCLUSIONS AND OUTLOOK	86
REFERENCES	88

List of Figures and Tables

FIGURE 2-1: THEIA LEVEL-2 DATA PRODUCTION EXTENT. (14.05.2019) SOURCE: HTTP://WWW.CESBIO.UPS-Tlse.FR/MULTITEMP/?PAGE_ID=7501	9
FIGURE 2-2: "GEOLOCATION PERFORMANCE FOR S2A (LEFT) AND S2B (RIGHT). THE PERFORMANCE IS COMPUTED OVER A SLIDING WINDOW OF 30 DAYS (S2A, LEFT) OR 10 DAYS (S2B, RIGHT). NUMBER OF PRODUCTS USED IN THE SLIDING WINDOW INDICATED FOR EACH POINT. (ACT= ACROSS TRACK, ALT= ALONG TRACK). A VERTICAL LINE INDICATES A GEOMETRIC CALIBRATION EVENT." (ESA, 2019)	12
FIGURE 2-3: LAND COVER CLASSIFICATION ACCURACIES FOR DB AND DN AND DIFFERENT PRE-PROCESSING SCENARIOS	16
FIGURE 2-4: DIFFERENT SAR RADIOMETRIC PROCESSING LEVELS: LEFT: BETA NAUGHT (SLANT RANGE GEOMETRY), MIDDLE: SIGMA NAUGHT (CORRECTED FOR LOCAL INCIDENCE ANGLE), RIGHT: FLATTENED GAMMA NAUGHT (TERRAIN CORRECTIONS ACCORDING TO SMALL, 2011)	17
FIGURE 2-5: EXAMPLE OF TWO SAR SPECKLE FILTERS APPLIED AT AN URBAN/RURAL TEST SITE	17
FIGURE 2-6: COMPARISON BETWEEN THREE DIFFERENT FULL RESOLUTION COHERENCE ESTIMATES OF THE SAME INTERFEROGRAM, (20090618-20090629): (A) A 5×5 WINDOW BOXCAR, (B) 11×11 WINDOW BOXCAR, AND (C) SIBLING-BASED (25 TO 100 SIBLINGS PER POINT) 41×41 SEARCH WINDOW). THE BOXCAR COHERENCE ESTIMATES MANY FALSE HIGH COHERENCE POINTS, SUCH AS IN THE FIELD OUTLINED IN RED. SOURCE: SPAANS AND HOOPER, 2016.	19
FIGURE 3-1: DEMONSTRATION SITES WHICH ARE PROCESSED IN TASK 4 ARE DRAWN IN BLACK, TEST SITES WHICH ARE PROCESSED IN TASK 3 ARE DRAWN IN RED.	24
FIGURE 3-2: ECOLASS TEST SITES IN AFRICA DEMONSTRATION SITES WHICH ARE PROCESSED IN TASK 4 ARE DRAWN IN YELLOW.....	25
FIGURE 3-3: "ACCURACY (A), PRECISION (P), AND UNCERTAINTY (U) SCORES PER BAND FOR THE S-2 SR PRODUCTS OF EVERY PROCESSOR AND OVER ALL TEST SITES. THE NUMBER OF POINTS (NBP) INVOLVED IN THE APU ANALYSIS VARIES DUE TO THE DIFFERENT NUMBER OF S-2 SCENES PROCESSED AND SUBMITTED BY EVERY PROCESSOR." (DOXANI ET AL., 2018).....	28
FIGURE 3-4: "COMPARISON BETWEEN THE MSI, MODIS AND OLI SENSORS. BASED ON: U.S. GEOLOGICAL SURVEY (2016), SUHET (2015), AND NASA OFFICIAL." (MILETICH, 2017)	29
FIGURE 3-5: "MEAN REFLECTANCE - RED BAND." (MILETICH, 2017)	30
FIGURE 3-6: "MEAN REFLECTANCE - NIR BAND." (MILETICH, 2017).....	30
FIGURE 3-7: "MEAN REFLECTANCE - SWIR BANDS." (MILETICH, 2017)	31
FIGURE 3-8: SEN2COR CLOUD MASK: COMPARISON BETWEEN LEVEL-1C AND LEVEL-2A PRODUCT	35
FIGURE 3-9: DIFFERENCES BETWEEN THE SENTINEL IMAGE LEVEL-2A WITHOUT CLOUD MASK (UPPER LEFT) AND THE SEN2COR CLOUD MASK (UPPER RIGHT), THE FMASK CLOUD MASK (LOWER RIGHT) AND THE MAJA CLOUD MASK (LOWER LEFT).	36
FIGURE 3-10: EXAMPLE 1: CLOUD MASK COMPARISON BETWEEN MAJA, FMASK AND SEN2COR. SENTINEL-2_20170725T105031_31UER; 49°41'8.69"N; 4°16'55.80"E.....	37
FIGURE 3-11: EXAMPLE 2: CLOUD MASKS COMPARISON BETWEEN MAJA, FMASK AND SEN2COR. SENTINEL-2_20170705T105031_31UER; 49°41'8.69"N; 4°16'55.80"E.....	38
FIGURE 3-12: SHADOW MASK: COMPARISON BETWEEN LEVEL-2A TERRAIN AND CLOUD SHADOWS.....	39
FIGURE 3-13: SHADOW MASK: COMPARISON BETWEEN TERRAIN AND CLOUD SHADOWS.....	40
FIGURE 3-14: "EXAMPLE SHOWING THE VARIOUS IMAGES BEFORE AND AFTER TOPOGRAPHIC NORMALISATION (SEE ALSO TABLE 1)." (POSCH, 2017).....	44
FIGURE 3-15: "EXAMPLE SHOWING THE REGRESSION LINE BEFORE AND AFTER TOPOGRAPHIC NORMALISATION (ON TOP 2 BEFORE AND BELOW DIAGRAM AFTER CORRECTION)." (POSCH, 2017)	45
FIGURE 3-16: "REGRESSION LINES AFTER TOPOGRAPHIC NORMALISATION USING DIFFERENT SOFTWARE PACKAGES AND DEMs." (POSCH, 2017)	46
FIGURE 3-17: "BAR DIAGRAMS SHOWING THE SLOPE (BETA1) OF THE REGRESSION LINE." (POSCH, 2017)	46
FIGURE 3-18: "BAR DIAGRAMS SHOWING THE CORRELATION BETWEEN COS(i) AND REFLECTANCE." (POSCH, 2017).....	47
FIGURE 3-19: IMAGE EXAMPLES SHOWING THE RESULTS OF THE TOPOGRAPHIC NORMALISATION WITH LIDAR DSM (LEFT) AND EUROMAPS DSM (RIGHT).	48
FIGURE 3-20: IMAGE EXAMPLES SHOWING DETAILS FROM THE RESULTS OF THE TOPOGRAPHIC NORMALISATION WITH LIDAR DSM (LEFT) AND EUROMAPS DSM (RIGHT).	48
FIGURE 3-21: GEOMETRIC ACCURACY ANALYSIS FOR SENTINEL-2 DATA AT A FLAT TERRAIN TEST SITE.	50
FIGURE 3-22: GEOMETRIC ACCURACY ANALYSIS FOR SENTINEL-2 DATA AT A MOUNTAINOUS TEST SITE.....	51
FIGURE 3-23: TWO SENTINEL-2 IMAGES FROM THE MOUNTAINOUS TEST AREA AND DISPARITY IMAGES IN EAST AND NORTH BETWEEN BOTH IMAGES.....	52
FIGURE 3-24: WEIGHTING FUNCTIONS USED BY THE ROBUST IRLS FITTING METHOD.	54
FIGURE 3-25: DISTRIBUTION OF SAMPLING POINTS OVER TILE 32UMU.....	55
FIGURE 3-26: NUMBER OF AVAILABLE OBSERVATIONS PER PIXEL FROM 2016-01-26 TO 2017-12-26.....	56

FIGURE 3-27: DIFFERENT SAR RADIOMETRIC PROCESSING LEVELS: LEFT: BETA NAUGHT (SLANT RANGE GEOMETRY), MIDDLE: SIGMA NAUGHT (CORRECTED FOR LOCAL INCIDENCE ANGLE), RIGHT: FLATTENED GAMMA NAUGHT (TERRAIN CORRECTIONS ACCORDING TO SMALL, 2011).....	59
FIGURE 3-28: EXAMPLES OF DIFFERENT ADAPTIVE SAR SPECKLE FILTERS.....	60
FIGURE 3-29: EXAMPLES OF DIFFERENT ADAPTIVE SAR SPECKLE FILTERS.....	61
FIGURE 3-30: FROST FILTER WITH DIFFERENT FILTER WINDOWS.....	61
FIGURE 3-31: LAND COVER CLASSIFICATION ACCURACIES ACHIEVED WITH DIFFERENT ADAPTIVE SAR SPECKLE FILTERS.....	62
FIGURE 3-32: CROSS-VALIDATION ACCURACIES FOR LAND COVER CLASSIFICATION ACHIEVED WITH DIFFERENT ADAPTIVE SAR SPECKLE FILTERS AND WINDOW SIZES.....	63
FIGURE 3-33: MEAN CROSS-VALIDATION ACCURACIES ACHIEVED WITH DIFFERENT SAR SPECKLE FILTERS, WINDOW SIZES, PROCESSING UNITS AND SPATIAL RESOLUTIONS.....	63
FIGURE 3-34: LEFT: SINGLE SENTINEL-1 SCENE, UNFILTERED; RIGHT: SINGLE SENTINEL-1 SCENE FILTERED WITH MTF, 19 IMAGES, 7x7 WINDOW.....	64
FIGURE 3-35: RSG COHERENCE PROCESSING MODULE.....	66
FIGURE 3-36: CALCULATION OF ORTHOGONAL BASELINES FOR 6-DAY COHERENCE PAIRS (LEFT) AND FOR ALL COMMON-MASTER PAIRS (RIGHT) FOR BELGIUM TEST SITE.....	67
FIGURE 3-37: OVERVIEW OF COHERENCE PRODUCT EXAMPLES FOR THE BELGIUM TEST SITE.....	68
FIGURE 3-38: COHERENCE PRODUCT EXAMPLES FOR POLARISATION WITHIN THE BELGIUM TEST SITE (BRIGHT: HIGH COHERENCE, DARK: LOW COHERENCE).....	68
FIGURE 3-39: COHERENCE PRODUCT EXAMPLES FOR TIME INTERVALS WITHIN THE BELGIUM TEST SITE (BRIGHT: HIGH COHERENCE, DARK: LOW COHERENCE).....	69
FIGURE 3-40: SELECTED EXAMPLES FROM COHERENCE PROCESSING OVER THE VEGETATION SEASON 2018 WITHIN THE BELGIUM TEST SITE (BRIGHT: HIGH COHERENCE, DARK: LOW COHERENCE).....	70
FIGURE 3-41: COHERENCE PRODUCT EXAMPLES FOR TARGETED SPATIAL RESOLUTIONS.....	71
FIGURE 3-42: COHERENCE PRODUCT EXAMPLES FOR POLARISATION WITHIN THE BELGIUM TEST SITE (BRIGHT: HIGH COHERENCE, DARK: LOW COHERENCE).....	71
FIGURE 3-43: COMPARISON OF SENTINEL-1 COHERENCE PRODUCTS WITHIN THE BELGIUM TEST SITE.....	72
FIGURE 3-44: COMPARISON OF SENTINEL-2 IMAGERY WITH SENTINEL-1 PRODUCTS WITHIN THE BELGIUM TEST SITE (LEFT: SENTINEL-2; MIDDLE: SENTINEL-1 GRD BACKSCATTER, RIGHT: SENTINEL-1 SLC COHERENCE).....	72
FIGURE 4-1: PRE-PROCESSING OF OPTICAL TIME SERIES – WORKFLOW.....	73
FIGURE 4-2: SEN2Cor SCL DERIVED CLOUD MASKS (IN BLUE) AND IMPROVED CLOUD MASKS DERIVED FROM THE MODIFIED SCL.....	76
FIGURE 4-3: SENTINEL-1 PRE-PROCESSING WORKFLOW IN EClaSS.....	81
FIGURE 4-4: SNAP WORKFLOW FOR COHERENCE CALCULATION.....	84
 TABLE 2-1: “GEOMETRIC IMAGE QUALITY REQUIREMENTS”. (LANGVILLE ET AL., 2015)	11
TABLE 3-1: TEST SITES AND CORRESPONDING DEMO SITES IN PHASE 2.....	23
TABLE 3-2: “CORRESPONDING BANDS REGARDING EACH SPECTRAL REGION ACROSS MSI, OLI AND MODIS.” (MILETICH, 2017)	29
TABLE 3-3: “OVERVIEW OF THE USED COMBINATIONS.” (POSCH, 2017)	43
TABLE 3-4: “OVERVIEW OF THE USED CLASSES AND NUMBER OF SAMPLES/PIXELS.” (POSCH, 2017)	43
TABLE 3-5: OVERVIEW OF THE USED COMBINATIONS.....	47
TABLE 3-6: RESULTS OF THE HARMONIC REGRESSION BENCHMARKING. N = NUMBER.....	57
TABLE 4-1: FILE SUFFIX FOR PROCESSED PRODUCTS.....	77
TABLE 4-2 - FILE NAME FILED CONVENTION FOR THE MAJA DATASETS	78
TABLE 4-3: DEMO SITES – RESULTING SENTINEL-2 TIME SERIES IMAGERY PHASE 1.....	79
TABLE 4-4: DEMO SITES - RESULTING SENTINEL-2 TIME SERIES IMAGERY PHASE 2.....	79
TABLE 4-5: TEST AND DEMO SITES – RESULTING SENTINEL-1 TIME SERIES IMAGERY IN PHASE 2.....	82
TABLE 4-6: TEST AND DEMO SITES – RESULTING SENTINEL-1 TIME SERIES IMAGERY IN PHASE 2.....	82
TABLE 4-7 - FILE NAME FILED CONVENTION FOR THE MAJA DATASETS	85

Abbreviations

ACCA	Automated-Cloud-Cover Assessment
ACIX	Atmospheric Correction Inter-comparison eXercise
AERONET	AErosol RObotic NETwork
AG	Adaptive Gaussian
AG	Asymmetric Gaussian
ALE	Absolute Location Error
ALOS	Advanced Land Observing Satellite
AOD	Aerosol Optical Depth
AOI	Area Of Interest
AOT	Aerosol Optical Thickness
APDA	Atmospheric Pre-corrected Differential Absorption
ASG	Adaptive Savitzky-Golay
ATCOR	ATmospheric CORrection
Bmerge	Burst merge
BOA	Bottom-of-Atmosphere
BT	Brightness Temperature
CACAO	Consistent Adjustment of the Climatology to Actual Observations
CADS	Calibration Annotation Data Set
CDI	Cloud Displacement Index
CEOS	Committee on Earth Observation Satellites
CESBIO	Centre d'Etudes Spatiales de la BIOsphère
CGIAR-CSI	Consortium for Spatial Information
CLMS	Copernicus Land Monitoring Services
CNES	Centre national d'études spatiales
CombSD	Combine Sentinel data
CPU	Central Processing Unit
db	Decibel
DDV	Dark Dense Vegetation
DEM	Digital Elevation Model
DLOG	Double Logistic
DLR	German Aerospace Center
DN	Digital Number
DORIS	Delft Object-Oriented Radar Interferometric software
DOs	Dark Objects
DSM	Digital Surface Model
DU	Dobson Unit
ECoLaSS	Evolution of Copernicus Land Services based on Sentinel data
EEA	European Environment Agency
EEE	Entrusted European Entities
EO	Earth Observation
ERS	European Remote-Sensing Satellite
ESA	European Space Agency
ESD	Enhanced Spectral Diversity
ETM+	Enhanced Thematic Mapper Plus
EW	Extra-Wide Swath Mode
Fmask	Function of mask
FORCE	Framework for Operational Radiometric Correction for Environmental Monitoring
FR	Full Resolution
FRE	Flat Reflectance

GAF	GAF AG
GCP	Ground Control Point
GeoTIFF	Georeferenced Tagged Image File Format
GIO	GMES Initial Operations
GMES	Global Monitoring for Environment and Security
GMT(s)	Generic Mapping Tool(s)
GNU	GNU's Not Unix
GPT	GUID Partition Table
GRD	Ground Range Detected
GRDH	Ground Range Detected High Resolution Product
GUI	Graphical User Interface
H2020	Horizon 2020
HGT	“height” format
HH	Horizontal transmit/Horizontal receive (polarization)
HPC	High Performance Computing
HR	High Resolution
HRL	High Resolution Layer
HV	Horizontal transmit/Vertical receive (polarization)
IMPACT	Image Processing and Classification Toolkit
InSAR	Interferometric Synthetic Aperture Radar
INSPIRE	Infrastructure for Spatial Information in Europe
IRECI	Inverted Red-Edge Chlorophyll Index
IW	Interferometric Wide Swath Mode
JM	Jeffries Matusita
JR	JOANNEUM RESEARCH Forschungsgesellschaft mbH
JRC	Joint Research Centre
KML	Keyhole Markup Language
LaRSC	Land Surface Reflectance Code
LC	Land Cover
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LibRadtran	Library For Radiative Transfer
LIDAR	Light Detection And Ranging
LOESS	locally weighted scatterplot smoothing
LUTs	Look Up Tables
MACCS	Multi-sensor Atmospheric Correction and Cloud Screening
MAJA	Maccs-Atcor Joint Algorithm
MMU	Minimum Mapping Unit
MODIS	Moderate Resolution Imaging Spectroradiometer
MODTRAN	MODerate resolution atmospheric TRANSmision
MR	Medium Resolution
MSI	MultiSpectral Instrument
MTF	Modulation Transfer Function
NASA	National Aeronautics and Space Administration
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-InfraRed
NLC	New Land Cover
OLI	Operational Land Imager
OLS	Ordinary Least Squares
OTB	Orfeo Toolbox
PROBA-V	Project for On-Board Autonomy – Vegetation

QI	Quality Indicators
RADAR	Radio Detection and Ranging
RMSE	Root Mean Square Error
RSG	Remote Sensing Software Graz
S2A	Sentinel-2A
S2B	Sentinel-2B
SAR	Synthetic Aperture Radar
SatRead	Read Satellite Data
SCM	Scene Classification Map
Sen2Cor	Sentinel-2 Atmospheric Correction
SIRS	Systèmes d'Information à Référence Spatiale SAS
SLC	Single Look Complex
SM	Strip Map Mode
SMAC	Simplified Method for Atmospheric Correction
SNAP	Sentinel Application Platform
SRE	Surface Reflectance
SRTM	Shuttle Radar Topography Mission
SSA	Single Scattering Albedo
SSP	Smoothing Spline
SWIR	Short Wavelength Infrared
TIR	Thermal Infrared
TM	Thematic Mapper
TOA	Top-of-Atmosphere
TOPSAR	Terrain Observation with Progressive Scans SAR
UCL	Université catholique de Louvain, Earth and Life Institute (ELI)
URL	Uniform Resource Locator
UTM	Universal Transverse Mercator
VH	Vertical transmit/Horizontal receive (polarization)
VHR	Very High Resolution
VNIR	Visible and Near InfraRed
VV	Vertical transmit/Vertical receive (polarization)
WGS	World Geodetic System
WP	Work Package
WV	Wave

1 Introduction

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

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

This present Deliverable D7.2 “**D32.1b – Methods Compendium Time Series Preparation Report (Issue 2)**” is the second issue of the “WP 32 - Time Series Preparation” main Deliverable, and takes into account the experiences made in the first and second phase of ECoLaSS in terms of pre-processing experiences from Task 3 and Task 4. It is part of Task 3: “Automated high data volume processing lines”. The main objective of this Task is to develop and benchmark methods to be capable of processing huge volumes of Sentinel data and developing and testing integration strategies to benefit from complementary multi-resolution, multi-temporal and multi-sensor characteristics, to support the generation of the prototype products for Continental and Global Land Components.

The aim of WP32 – Time Series Preparation is to find robust methods for fully automated pre-processing of Sentinel time series data streams. The assessment of pre-processing chains and methodologies as an outcome of WP32 forms part of the basis for the methodological developments in Task 3 and the prototype demonstrations in Task 4. The quality of the pre-processing procedure determines the thematic quality, content and accuracy that can be achieved in any subsequent information extraction processes. It has to be considered to process the best possible data products in order to fulfil the highest quality standards. Final conclusions and recommendations are based on literature studies, benchmarking results and experiences of consortium partners.

The generation of spatio-temporally consistent optical time series with top of atmosphere reflectance values requires substantial fully automated pre-processing including a) extraction of metadata, b) atmospheric correction, c) topographic normalisation, d) cloud, cloud shadow and snow masking, e) optional - radiometric calibration, and f) optional - spatio-temporal noise reduction and gap-filling. Considering the data size and number of scenes to be processed, an automated and rapid processing chain is essential. The research related to optical time series pre-processing is therefore focusing on testing of best suited algorithms and their application in test sites and the development of new methods as well as the implementation of prototype methods.

Time series analysis based of SAR data can be based on backscatter or coherence products or a combination of these. All radar data require substantial pre-processing including a) automated SAR image download, b) automated preparation of digital elevation data, c) automated update of orbit files (precise; restituted for near real time processing), d) thermal noise removal (for SLC) and GRD border noise removal, e) radiometric calibration, f) multi-looking and image filtering (speckle, adaptive), g)

spatio-temporal filtering (compare e.g. Quegan et al., 2000), h) radiometric terrain corrections (compare e.g. Small, 2011) and i) geo-referencing. Considering the number of scenes to be processed and the resulting data size, an automated and rapid processing chain is essential. The research related to SAR time series pre-processing is, therefore, focusing on testing of existing / currently developed algorithms and tools and implementation of spatio-temporal noise filtering tools optimized for Sentinel-1 time series, cluster processing to reduce processing time, and derivation of multi-temporal metrics.

2 State-Of-The-Art

To monitor land cover and its changes, Earth observation (EO) data with a high temporal and spatial resolution are required. However, up to recently these scientific as well as user requirements could only be satisfied by dense satellite data time series on local, continental and global scale. Based on the experiences from GMES Initial Operations (GIO) stage between 2011 and 2014, EO data availability has been identified as the major bottleneck for operational service implementation. Continuous satellite image acquisition systems are required to avoid very long range image acquisition for complete coverages and the lack of timely availability of initial contiguous image coverages [AD06]. In this context each Sentinel mission brings new opportunities and challenges for the time series data processing lines. However, robust methods for fully automated pre-processing of Sentinel time series data streams are required [AD04]. The quality of the pre-processing procedure determines the thematic quality, content, and accuracy that can be achieved in any subsequent information extraction processes. Therefore, pre-processing approaches for high volume data processing based on the high resolution satellite constellations are addressed in this chapter.

2.1 Processing methods for optical time series

The processing methods for optical time series include the generation of spatio-temporally consistent optical time series with top of atmosphere reflectance values. Therefore the following pre-processing steps are available: a) extraction of metadata, b) atmospheric correction, c) topographic normalisation, d) cloud, cloud shadow and snow masking, e) optional - radiometric calibration, f) derivation of indices and variables [AD04].

2.1.1 Atmospheric correction

In both phases of ECoLaSS, Sentinel-2 data are downloaded from the Copernicus Hubs and spatio-temporally consistent optical time-series with Bottom-of-Atmosphere reflectances are generated. In order to obtain the Bottom-of-Atmosphere (BOA) reflectances for optical data, the two main processors are Sen2Cor and MAJA.

SEN2COR

One software package for atmospheric correction is presented from ESA as Sentinel-2 Level-2 processor software named Sen2Cor, providing a set of algorithms, which are composed of several state-of-the-art techniques performing the task of correcting the reflectance values (Müller-Wilm et al., 2013).

The processor performs the atmospheric-, terrain and cirrus correction of Top-of-Atmosphere (TOA) Level-1C input data creating Bottom-of-Atmosphere (BOA) and, if required, terrain- and cirrus corrected reflectance images. Parameters like atmospheric visual transparency, aerosol type and the water vapour column are needed for the atmospheric correction process and derived through the associated bands (Müller-Wilm et al., 2013).

The Dense Dark Vegetation algorithm performs the task of the Optical Aerosol Thickness detection and aerosol type estimation, requiring areas of known reflectance behaviour. How the process works in detail is explained by Richter et al. (2006):

"The method is based on a stable relationship between surface reflectance ρ in the short-wave infrared (SWIR; 2100–2200nm) and reflectance in the blue (480 nm) and red (660 nm) wavelengths. It starts with the approximation $\rho_{SWIR} = \rho_{SWIR}^{TOA}$ where ρ^{TOA} the Top-of-Atmosphere reflectance is. This approximation is fairly accurate, because of the small path radiance and high atmospheric transmittance $\tau > 0.9$ in the SWIR region [...]. Then, the relationships $\rho_{blue} = \frac{\rho_{SWIR}}{4}$ and $\rho_{red} = \rho_{SWIR}/2$ are

employed and the differences between the TOA apparent reflectance and the reflectance from the above relationships are used to calculate the path radiance and corresponding aerosol optical depth. The spectral reflectance correlation between the red and SWIR wavelengths is sufficiently stable for an aerosol retrieval."

Louis et al. (2016b) states after a first validation process comparing with AERONET sunphotometer measurements that the algorithm only works with dark dense vegetation pixels within each granule.

Furthermore, **Water Vapour** information is required to correct reflectance values, gathered through the Atmospheric Pre-corrected Differential Absorption (APDA) algorithm, using the bands B8a and B9. Band 8a provides reflectance values in an atmospheric window region, whose values can be seen as a reference for the water vapour. Band 9 presents reflectance values in an absorption region. The offset between those two bands is a measure for absorption depth, which dominates the lower atmosphere, assuming that the surface reflectance for B9 is the same as for the B8a. In this case, the absorption depth represents water vapour column content (Müller-Wilm et al., 2013). Louis et al. (2016a) observed after a first validation process with AERONET sunphotometer measurements that the water vapour retrieval algorithm is less influenced by clouds and missing dark dense vegetation pixels than by the aerosol thickness detection results. Since the water vapour column in the upper atmosphere and the ozone absorption content cannot be estimated from image data, seasonal climatological values are included, differentiating between summer and winter atmosphere (Louis et al., 2016a).

Optionally a **cirrus correction** can be performed using band 10, which is located in the 1380 nm spectral region, where the ground reflected signal is totally absorbed receiving only scattered cirrus signal at the sensor (Richter et al., 2011). The correlation between the cirrus band (B10) and other bands in NIR and SWIR leads to the cirrus contribution. This contribution can be removed from the radiance signal (Müller-Wilm et al., 2013; Louis et al., 2016a). An established cirrus removal method exists for visible and near infra-red bands. The Sen2Cor processor uses an extended technique removing cirrus during the atmospheric reflectance calculation to avoid reflectance artefacts and improving the accuracy of surface reflectance retrievals. Richter et al. (2011) describe the standard cirrus removal method:

"A 'virtual' surface is defined consisting of the earth's surface and atmosphere beneath cirrus clouds. It includes the effects of surface reflection and atmospheric scattering and absorption processes. Above cirrus, there is the remaining part of the atmosphere with a residual water vapour column. [...] The cirrus reflectance in the VNIR (from 0.4 to 1 μ m) is linearly related to the cirrus reflectance at 1.38 μ m [and] [...] the cirrus single scattering albedo (SSA) is close to 1 in this part of the spectrum which means that scattering dominates. It does not hold in the SWIR spectrum (1.6 μ m, 2.2 μ m) where cirrus absorption is higher [...]."

This means that standard cirrus removal method only considers channels in NIR window regions, excluding the 0.94 μ m water vapour region. Hence, an enhanced cirrus removal method is introduced by Richter et al. (2011), accounting for the water vapour absorption above cirrus clouds, during the atmospheric correction process using band 10. "For instruments with similar channels as Sentinel-2, the water vapour is retrieved with a band in a window region (around 0.87 μ m) and one in the 0.94 μ m absorption region (Richter et al., 2011)." The transmittance between sun-cirrus and cirrus-satellite is initially unknown, therefore radiative transfer codes using climatological atmosphere profiles for different seasons and areas adopted from MODTRAN models are used (Richter et al., 2011).

Sen2Cor uses a combined approach for the atmospheric and topographic correction, eliminating the topographic effects during the BOA reflectance calculation. In order to the terrain correction processing, a digital elevation model can automatically be downloaded by the processor using the SRTM 3 arcs model. The 90m SRTM Digital Elevation Database required by the algorithm for rugged terrain will be downloaded from CGIAR-CSI. The topographic correction requires parameters like slope, aspect and hill

shadow. The digital elevation model will be automatically resampled and referenced to the Sentinel scene (Main-Knorn et al., 2015; Müller-Wilm et al., 2013).

"The atmospheric model of Sen2Cor (Level-2_AtmCorr) is dependent on the calculation of radiative transfer functions for different sensor and solar geometries, ground elevations, and atmospheric parameters. [...] The processor reads the parameter in form of Look Up Tables (LUTs) pertaining to this parameter space and interpolates if required. The LUTs have been generated via libRadtran, a library for the calculation of solar and thermal radiation in the Earth's atmosphere." (Müller-Wilm, 2015) There are 24 different look-up tables provided by the processor including tables regarding rural and maritime aerosols, mid-latitude summer and winter atmosphere profiles and various types of ozone concentration, which can be configured in the Level-2 GIPP.xml file (Louis et al., 2016b).

The **Aerosol-type** is selected depending on the climatology from the scene acquisition area and time. By using built-in MODTRAN aerosols, for example, rural or maritime aerosol types are differentiated. The ozone content is provided by ancillary data for each scene individually (Richter et al., 2006; Müller-Wilm et al., 2013). The ozone content concentration, measured in Dobson Units (DU), should be chosen depending on the season (summer or winter). The standard ozone content for a latitude summer atmosphere is 331 DU and in winter it is about 330 DU (Müller-Wilm, 2015).

After all required parameters for the atmospheric correction are estimated, the BOA reflectance calculation is performed using a set of the available and calculated LUTs, chosen by the user according to geographic location and climatology (Main-Knorn et al., 2015; Müller-Wilm et al., 2013). Output images are automatically resampled to a user-defined spatial resolution based on the three resolutions available in the input image. The three bands with 60m spatial resolution (Bands 1, 9 and 10) will be omitted in the level-2a output since they are not needed for land cover applications (Louis et al., 2016b; Müller-Wilm et al., 2013; Fletcher K., 2012). In a first validation process again using AERONET as reference data, Louis et al. (2016b) observed a high spectral agreement between several example spectra (e.g., forest, bright soil, etc.) and the Sentinel image data. The reference image has been calculated running the atmospheric correction on the Sentinel scene with an AOT constant measured by the AERONET instrument (Louis et al., 2016b).

Another major output of the processor is the scene classification, which detects 12 different land cover classes (Müller-Wilm et al., 2013). "The algorithm is based on a series of threshold tests that use as input Top-of-Atmosphere reflectance from the Sentinel-2 spectral bands. In addition, thresholds are applied to band ratios and indexes like the Normalized Difference Vegetation - and Snow Index (NDVI, NDSI)" (Müller-Wilm et al., 2013). In a first validation approach, the scene classification reached an overall accuracy of 78%. Random stratified sampling points are visually interpreted as a reference, with at least 50 points per class (Louis et al., 2016b).

MACCS-ATCOR JOINT ALGORITHM MAJA

MAJA is a software package for atmospheric correction presented from CNES, DLR and CESBIO. MAJA stands for MACCS-ATCOR Joint Algorithm and represents cloud screening and atmospheric correction software. The MAJA code uses multi-temporal and multi-spectral information to detect clouds, cloud shadows, water and snow. Moreover, the algorithm estimates the optical properties of the atmosphere (water vapour and aerosol content) and performs an atmospheric correction including adjacency and irradiance variations due to terrain (Rouquié et al., 2017).

As a first step, the algorithm detects areas covered with clouds, cloud shadows, snow and water and generates corresponding masks. In these areas, the atmospheric corrections won't provide reliable results. The mask is generated based on multitemporal data sets to improve the results. Second, two components are computed for the atmospheric correction, the atmospheric absorption and the atmospheric scattering. In the MAJA package, the atmospheric absorption correction is based on the

SMAC model (Simplified Method for Atmospheric Correction) and additionally it is based on the Water Vapour Absorption knowledge derived from the corresponding band. Third, the aerosol optical thickness is computed; therefore three different methods are available within the MACCS software. The first method is based on a multi-temporal dataset assuming that the aerosol concentration varies quickly over time but slowly in space. The second method is a spectral method assuming that the vegetation is characterized by a constant ratio of reflectance for some spectral bands. The last method represents a spectro-temporal method which combines the previous two methods. The last contribution, which will be corrected by the algorithm is the blurring due to the landscape surrounding the site imaged, i.e. the adjacency effects (Lonjou et al., 2016; Rouquié et al., 2017)

“The algorithm implemented gets its reliability from the use of the temporal dimension to improve the knowledge of the area that has been imaged and make the distinction between what is slowly changing, i.e. the landscape itself, and elements quickly varying such as clouds, clouds shadows and aerosols. In addition to the BOA reflectance for all the bands, a set of composite images at low resolution is also appended into the Level-2 product; they contain the most recent good quality surface reflectance or cloud free TOA reflectance and they are enriched date after date.”(Lonjou et al., 2016)

MAJA is available as executable code for Linux RedHat/CentOS (see https://github.com/olivierhagolle/Start_maja) since October 2nd, 2017. In the latest release, version 3.2.2, MAJA has been improved further by adding the DLR thin cirrus detection and correction method (Richter et al., 2011, Donadieu and L’Helguen, 2016). Estimation of cirrus reflectance, thereby, is based on the Sentinel-2 MSI cirrus band (1.38 µm). Furthermore, MAJA now normalizes the reflectance of each pixel based on its viewing angle. As of now, this is based on a constant normalization function; however, it may be improved by introducing a land cover-dependent correction. Moreover, aerosol correction now makes use of the Copernicus Atmosphere Monitoring Service AOT forecast to improve AOT estimation.

MAJA is freely available in a binary version for non-commercial use. The software may not be used for commercial purposes, where the following points are included under “commercial use”: Sale of the software, the sale of products using the software and sale of services that have been integrated in the software (CNES, 2018).

FORCE

FORCE infers the surface reflectance from Tanré's formulation of the radiative transfer, including adjacency effect correction. “A precompiled MODIS water vapor database provides daily or climatological fallback estimates. Aerosol optical depth (AOD) is estimated over dark objects (DOs) that are identified in a combined database and image-based approach, where information on their temporal persistency is utilized. AOD is inferred with consideration of the actual target reflectance and background contamination effect. In case of absent DOs in bright scenes, a fallback approach with a modeled AOD climatology is used instead.” (Frantz et al., 2016)

2.1.2 Cloud, Cloud Shadow and Snow Masking

As a result of the correction software packages, mentioned in the last chapter, additionally, to the corrected image data, cloud, cloud shadow and snow masks can be derived. Cloud and cloud shadow detection is one of the most important pre-classification activities when dealing with optical remote sensing data as it is of high interest for most applications to discriminate between clear sky pixels and pixels covered by clouds or cloud shadow. Therefore, it is among the first processing steps after pre-processing the raw data. The identification and exclusion of clouds and cloud shadows are necessary due to their influence on spectral reflectance values and characteristics of the thematic classes. Two main cloud types are differentiated: (a) Opaque clouds, which block almost all information from the surface, and (b) cirrus clouds, which are partially transparent and share spectral signatures with the land cover beneath. Since Sentinel-2 has no bands in the thermal spectrum, the cirrus band B10 (1375 nm) is

introduced to the cloud detection, which is most suitable for detecting thin cirrus. Cloud pixels are characterized by lower temperature and higher reflectance than land cover pixels. Therefore, many cloud detection algorithms depend on the thermal band (Zhu et al., 2015; Hagolle et al., 2010). Zhu et al. (2015) demonstrate that the cirrus band is even more important than the thermal band for cloud detection processes, comparing LANDSAT 4-7, LANDSAT 8 and simulated Sentinel-2 images with the Fmask cloud detection algorithm.

There is a range of existing approaches which can be summarized in following groups: 1) local or aware of spatial context, 2) self-contained or dependent on external data, 3) being probabilistic or decision-based (Hollstein et al., 2016). Hollstein et al. (2016) tested several ready-to-use classification algorithms on a database of manually labelled Sentinel-2 data. They presented a detection scheme based on classical Bayesian probability estimation which provides higher accuracies compared to decision trees computed from simple band math formulas. In the following, most common cloud detection approaches for Sentinel-2 data are introduced.

SEN2COR APPROACH

The Level-1C products embed a cloud mask specifying dense and cirrus clouds at 60m spatial resolution based on band thresholding. Dense clouds, which are characterized by high reflectance in the blue spectral region, are classified by using a threshold for B2. Additionally, a threshold in B11 and B12 (SWIR) is used to avoid misclassification with snow. Cloud reflectance is high in the SWIR, whereas snow presents a low reflectance. “At B10, there is a high atmospheric absorption band and only high altitude clouds are detected. However, this last criterion is only applied after the first detection of cloud pixel in the blue band where cirrus is transparent.” After detecting dense clouds, an algorithm for cirrus cloud discrimination is being applied based on two criteria. First, B10 corresponds to a high atmospheric absorption band. Therefore, only high altitude clouds can be detected. Secondly, cirrus cloud being semi-transparent cannot be detected in the B2 blue band. Furthermore, morphology-based operations are applied to reduce false detections. The filter performs erosion to remove isolated pixels and dilatation to fill the gaps. Using the Level-1C products, the user can use the Sentinel-2 Toolbox to generate Level-2 data including cloud probabilities. More details can be found in the ESA Sentinel Technical Guides (2017).

The Level-2 processing includes scene classification and an atmospheric correction resulting in Bottom-of-Atmosphere (BOA) reflectances, Scene Classification Map (SCM) together with Quality Indicators (QI) for cloud and snow probabilities at 60m resolution (ESA. 2017). At Level-2, the scene classification which consists of four different classes for clouds (including cirrus), together with six different classifications for shadows, cloud shadows, vegetation, soils/deserts, water and snow. The scene classification algorithm is based on threshold tests on top-of-atmosphere reflectance and spectral indices. Initially a snow and ice probability map as well as a binary mask is derived based on the normalized snow index (NDSI). On the remaining pixels, cloud probability is estimated again by means of threshold tests, followed by cloud shadow estimation.

SEN2-AGRI APPROACH

The Level-2 processor of the Sen2-Agri project is based on Multi-sensor Atmospheric Correction and Cloud Screening (MACCS) chain and ATCOR Joint Algorithm (MAJA), which perform atmospheric correction and provides Bottom-of-Atmosphere reflectance values with snow, water, cloud and cloud shadow masks (Udroiu et al., 2017). The MACCS is a further development of the CNES Orfeo Toolbox Library (B Petracci et al., 2015) and ATCOR is a DLR atmospheric and topographic correction software (Louis et al., 2016b; Richter and Schlapfer, 2016).

FMASK APPROACH

One of the most popular approaches has been developed by Zhu and Woodcock, 2012a for Landsat 4-7. The so-called Fmask (Function of mask) approach uses Top of Atmosphere (TOA) reflectance and Brightness Temperature (BT) to detect clouds, cloud shadows and snow. This approach has been improved and a new version has been provided for Landsat-8 and Sentinel-2 data (Zhu et al., 2015). The major input for cloud detection for Landsat data was a thermal band which does not exist for Sentinel-2. Nevertheless, the new cirrus band of Sentinel-2 has been found to be useful for cloud detection especially for thin cirrus clouds (Zhu et al., 2015).

FORCE APPROACH

The Fmask approach has been improved further for the Sentinel 2 MSI sensor by Frantz et al. (2018). By exploiting sensor specific parallax effects between different NIR band viewing geometries, they could completely make up for the missing thermal band. The authors demonstrate a much improved separability of clouds from bright ground targets, i.e. the most common false positives and consequently an improvement in overall accuracy over Fmask of +0.29 to +0.39 over urban areas. The algorithm is open-source and available as a sub-module in the FORCE processing system (<https://www.uni-trier.de/index.php?id=63673>).

ACCA ALGORITHM

The Automated Cloud-Cover Assessment (ACCA) Algorithm is an unsupervised classifier for clouds, which takes advantage of known spectral properties of thick clouds, snow, bright soil, vegetation, and water but no assessment of cirrus clouds and cloud shadows. The Landsat Processing System applies ACCA on every image prior to archiving to save the cloud cover percentages for quarter scenes and full scenes within the metadata records. The scene depended approach uses the reflective and thermal signatures to detect clouds. Problems occur in regions with snow cover at high illumination angles (Irish et al., 2006a).

MACCS-ATCOR JOINT ALGORITHM MAJA

The MAJA cloud detection method is based on a number of threshold tests including the cirrus band (B10). Additionally, multi-temporal tests are carried out to detect clouds by measuring a steep increase of the blue surface reflectance. Finally, the correlation of the pixel neighbourhood with previous images is calculated to avoid over detections based on the assumption that two different clouds at the same location on successive dates will not have the same shape. If a large correlation is observed, the pixel is excluded from the cloud mask (Lonjou et al., 2016).

The CNES MUSCATE production centre produces near-real time the Sentinel-2 Level-2 data, corrected for atmospheric effects using the MAJA software. The products are available for download at <https://theia.cnes.fr/atdistrib/rocket/#/home>. The data is acquired in large areas shown in Figure 2-1.

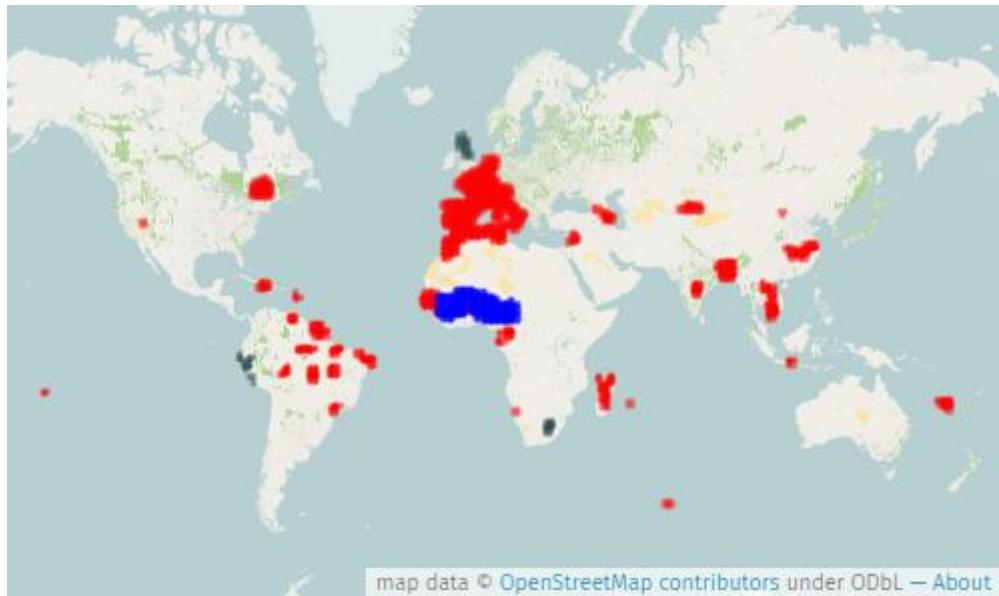


Figure 2-1: Theia Level-2 data production extent. (14.05.2019) Source: http://www.cesbio.ups-tlse.fr/multitemp/?page_id=7501

2.1.3 Topographic normalisation

A topographic correction is necessary if the test sites are characterized by mountainous terrain. The topography can significantly influence the radiometric properties of the signal received from the satellite (see Wulder and Franklin, 2012). This effect is caused by the different lighting angles resulting from the topography (Gallaun, Schardt and Linser, 2007). The aim of a topographical correction is to compensate for the differences in reflectance intensity between the areas with varying slope, exposure and inclination and to obtain the radiation values that the sensor would have in the case of a flat surface. A large number of approaches have been developed to correct the topographic influence. In these algorithms, which are used for topographic normalization, the illumination conditions for each pixel are usually estimated by calculating the cosine of the sun incidence angle based on the solar position, slope and exposure of the pixel. The following formula is used to calculate the cosine of the angle of incidence (see Formula 5.1) (Sola, González-Audícan und Álvarez-Mozos, 2016):

$$\cos \gamma_i = \cos \beta \cos \theta_s + \sin \beta \sin \theta_s \cos(\varphi_n - \varphi_s)$$

$\cos \beta$ is the angle of inclination, $\cos \theta_s$ is the solar zenith angle, φ_n is the aspect angle and φ_s is the solar azimuth angle, where $\cos \beta$ and φ_n are pixel-based values derived from the DEM.

According to Balthazar, Vanacker and Lambin (2012), the various topographical correction techniques developed can be divided into three types based on their complexity and requirements. The first group of procedures are simple empirical methods, such as ratio algorithms. If it is assumed that the lighting effect caused by the topography is proportional to the different bands, then the ratio of two bands can eliminate topographic effects. However, these algorithms do not produce satisfactory results because the radiometric variations caused by the topography are dependent on the wavelength and the differences between the bands are not equally proportional increasing or decreasing. This method also cuts out important image information. Since these approaches are not very successful, methods using digital elevation models (DEMs) that provide terrain elevation, slope inclination, exposure and other information to describe surface geomorphology are often preferred (see Timothy A. Warner, 2009; Balthazar, Vanacker and Lambin, 2012).

Methods using such digital height information belong to the second group of methods, the so-called semi-empirical approaches. These include the Cosine-correction, C-correction and Minnaert models (

Balthazar, Vanacker and Lambin, 2012). The Cosine-correction, which Park et al., 2017 in its categorization of topographic methods does not yet rank among the semi-empirical approaches but rather among the Lambertian approaches, assumes for all wavelengths (spectral bands) that the bidirectional reflection factor is constant and independent of angle of incidence and angle of failure (Park et al., 2017; Vincini and Frazzi, 2003). This type of correction is defined as follows (see formula 5.2):

$$L_n = L \cdot \frac{\cos \theta}{\cos i}$$

In the formula for Cosine-correction, L_n stands for the topographically normalized radiation, L stands for the observed radiation in uneven terrain, θ stands for the sun's angle of incidence and i for the sun's angle of incidence (Vincini and Frazzi, 2003; Wulder and Franklin, 2012). The C-correction method is based on the observed empirically linear correlation between radiation L and cosine of the angle of incidence $\cos i$. It estimates the correction factor c for each band λ by forming the ratio of the regression offset $a_{1\lambda}$ to the regression slope $b_{1\lambda}$ and then applying the correction factor to the Cosine-correction (Vincini and Frazzi, 2003; Sola et al., 2016 cited after Teillet, Guindon and Goodenough, 1982)

$$L_\lambda = a_{1\lambda} + b_{1\lambda} \cdot \cos i$$

$$L_{\lambda n} = L_\lambda \cdot \frac{\cos \theta + c_\lambda}{\cos i + c_\lambda}$$

The Minnaert correction, on the other hand, changes the Cosine-correction by a constant k , which is obtained by a regression that determines how much the observed surface behaves like a diffuse Lambert lamp (see formula: 5.5). Minnaert models derive such a Minnaert constant k for each spectral band. If it has a value between 0 and 1, this means that the surface is a perfect Lambert lamp (Kane et al., 2008; Hantson and Chuvieco, 2011; Vincini and Frazzi, 2003). The Minnaert correction was originally proposed by Minnaert (1941) and is defined as:

$$\log(L_\lambda \cdot \cos e) = \log L_{\lambda n} + l_\lambda \cdot \log(\cos i \cdot \cos e)$$

$$L_{\lambda n} = L_\lambda \cdot \frac{\cos^{1-k_\lambda} e}{\cos^{k_\lambda} i} \cdot \cos^k \theta$$

$$L_{\lambda n} = L_\lambda \left[\frac{\cos \theta}{\cos i} \right]^k$$

In this formula, e is the angle of precipitation angle or slope angle derived from the elevation model (Kane et al., 2008; Colby, 1991). As you can see, in practice the Minnaert constant is estimated by a linear equation (Hantson and Chuvieco, 2011). There are also some methods that have changed or extended the Minnaert Model (see also: Kane et al., 2008; Sola et al., 2016).

2.1.4 Geometric Consistency

Pre-processing of EO data used in time series analysis applications is one of the most crucial issues to guarantee reliable and meaningful results, especially so, when data from different sensors are used. In ECoLaSS the imperviousness, grassland, forest and agriculture applications may use various combinations of Sentinel-2, Sentinel-1 and Sentinel-3 data. Therefore, it is essential to analyse the overall data quality of the input data and the initial geometric shifts and radiometric differences between image data from the same sensor (e.g. from neighbouring orbits) and from different sensors. Ground Control Points (GCPs) are quite often used for testing geometric accuracy of satellite images. They can be derived either by using GPS or from aerial photography or by overlaying the orthophoto onto another dataset which

features better accuracy. Differences in feature locations between datasets are observed and quantified, and later used for determination of the accuracy (Pandžić et al., 2016).

Referring to Sentinel-2 optical observation data the ESA processing chain already implements a geometric refinement step which aims at improving the repetitiveness of the image geolocation. This step is necessary in order to reach the multi-temporal geolocation requirement. The main geometric quality requirements are summarized in Table 2-1.

Table 2-1: “Geometric image quality requirements”. (Langville et al., 2015)

Geometric image quality performance	Ground processing hypothesis
A priori accuracy of image location: 2km max (3σ)	No processing
Geolocation Performances Accuracy of image location: 20m (3σ)	After image processing without control points
Geolocation Performances Accuracy of image location: 12.5m (3σ)	After image processing with control points
Multitemporal registration: 3m (2σ) for 10m bands 6m (2σ) for 20m bands 18m (2σ) for 60m bands	After image processing with control points
Multispectral registration for any couple of spectral bands: 3m (3σ) for 10m bands 6m (3σ) for 20m bands 18m (3σ) for 60m bands	After image processing with control points

To meet the requirements for multi-temporal registration and the absolute geo-location, a refinement of a geometric model is based on a set of Level-1B images covering the whole globe where a spatiotriangulation algorithm obtains the geolocation information using reference GCPs (ESA, 2019; Gascon et al., 2017). Furthermore, ancillary data provided by the satellite, and models calibrated in-flight or before launch are used. The geolocation results from a combination of board and ground contributors (Langville et al., 2015; Gascon et al., 2017). “Geometric calibration activities allow the determination of all the GIPPs of the geometric calibration model which aims at better ensuring geometry compliance to the requirements for Sentinel-2 images (orientation of the viewing frames, lines of sight of the detectors of the different focal planes). These parameters have been estimated before launch and the purpose of the geometric calibration activities is to take into account any update of these parameter values that might occur.” (Gascon et al., 2017)

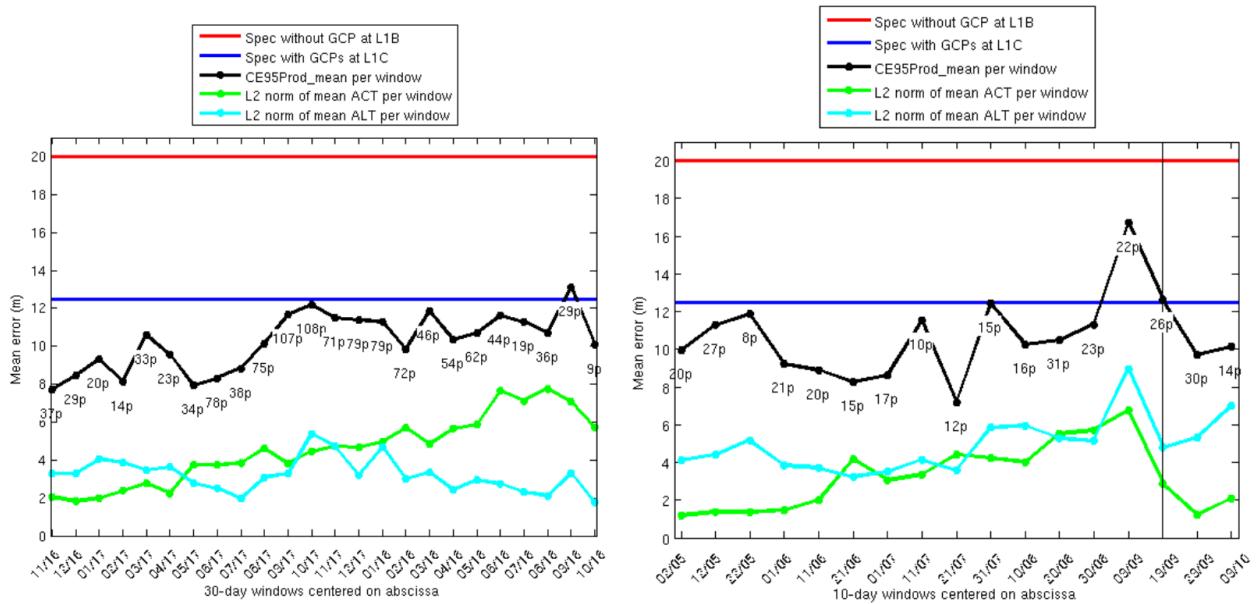


Figure 2-2: “Geolocation performance for S2A (left) and S2B (right). The performance is computed over a sliding window of 30 days (S2A, left) or 10 days (S2B, right). Number of products used in the sliding window indicated for each point. (ACT= across track, ALT= along track). A vertical line indicates a geometric calibration event.”

(ESA, 2019)

“The absolute geolocation is constantly monitored for S2A and S2B. The long-term performance is close to 11m at 95% for both satellites. The performance for S2B reported below shows the improvement carried out with the calibration of September 2018. An update of the geometric calibration for S2A and S2B is in preparation.” (ESA, 2019)

Furthermore, Pandžić et al., (2016) performed a quality analysis based on comparison of Sentinel-2 imagery with orthofotos as reference data. They used a statistical parameter to describe the geometrical relation between different images represented as the distance between corresponding curves in two datasets. “The experiment was conducted for two different test sites, Austria and Serbia. From 21 lines with a total length of ca. 120km the average offset of 6.031m (0.60 pixel of Sentinel-2) was obtained for Austria, whereas for Serbia the average offset of 12.720m (1.27 pixel of Sentinel-2) was obtained out of 10 lines with a total length of ca. 38km.” (Pandžić et al., 2016)

2.1.5 Noise reduction and gap filling

From the currently produced HRLs 2018, specifically the Grassland and the Water and Wetness products require dense times series over the reference year [AD06]. Furthermore, time series of remotely sensed data are an important data source for understanding land cover dynamics. However, time series data are not continuous and consistent in space and time and suffer from noise due atmospheric conditions (the presence of clouds, a high aerosol concentration, effects due to anisotropic reflectance behaviour of surfaces) and instrument issues (sensor problems or algorithm failures) (Moreno et al., 2014). Algorithms for time series filtering are required to reduce the impact of noise, since these disturbances negatively affect the usability for accurate land cover monitoring and change detection. Regarding the time series noise reduction spatial interpolation methods have the advantage of minimizing local anomalies so that large-scale trends in regional to global vegetation phonologies can be better identified (Chen et al., 2011; Moreno et al., 2014).

The first methods for noise reduction and gap filling are compositing methods, which operate over a local temporal window, to minimize cloud and snow cover artefacts and atmospheric effects. Nevertheless, several investigations have been focussing on the improvement of time series data using

mathematical filters to get temporally smoothed and spatially continuous products. The combination of temporal and spatial methods like geostatistical or regression methods, can overcome the limitation of spatial filters, by using historical data and temporal curves from neighbour pixels. The spatial filtering has the advantage of minimizing local anomalies but they may fail in complex landscapes, with a high proportion of mixed pixels. Temporal filtering methods are differentiated between local or global methods. “Local methods include: rank based (e.g., median, maximum or minimum filters) and linear/polynomial (e.g., moving average, adaptive Savitzky-Golay (ASG), locally weighted scatterplot smoothing (LOESS), Whittaker smoother, CACAO method, and smoothing spline (SSP) methods. The advantage of the local methods is that they make no assumptions about the underlying nature of the processes responsible for the variability in the time series while taking advantage of the neighbouring observation to generate an estimate. Nevertheless, for noisy and incomplete time series these methods are more limited and it becomes necessary to apply restrictions. Global approaches consider model fits such as polynomial, Double Logistic (DLOG), asymmetric Gaussian (AG), decomposition techniques based on Fourier analysis or wavelet transforms. Global methods are more suited to derive terrestrial biophysical parameters and extract seasonality information for phenological studies. Nevertheless, these methods also suffer several drawbacks that limit their use. [...] Global methods assume an a-priori phenological shape or magnitude that could make them not flexible enough to model the temporal response of irregular or asymmetric time series (e.g., high inter-annual variability).” (Moreno et al., 2014)

Statistical methods to smooth the time series and provide forecast information, including Fourier transforms or wavelet decomposition, are based on a decomposition of the time series into noise, seasonal variability, and trend and therefore require relatively long time series of observations (Kandasamy et al., 2013). Furthermore, the effectiveness of Savitzky-Golay and the asymmetric Gaussian function-fitting filter have been tested by several studies. Hird and McDermid (2009) showed that Savitzky-Golay and Gaussian function-fitting method filter provide a balanced performance to derive the optimal time series metrics. Both methods are outperformed by the Fourier-based fitting method. The Savitzky-Golay filter presents a method of data smoothing based on local-least-squares polynomial approximation. The method is reducing noise while maintaining the shape and height of the waveform peaks. In comparison to the asymmetric Gaussian function-fitting method it shows the more accurate derivation of the start of season and end of season measures (Hird and McDermid, 2009). Moreno et al., (2014) show that the adaptive Savitzky-Golay and the adapted local regression filter (LOESS) can robustly reconstruct the trajectories, reducing the noise up to 80%. The adapted LOESS is particularly resistant to outliers. Jönsson and Eklundh (2004) developed the TIMESAT program, which uses an adaptive Savitzky-Golay filter and optionally harmonic and asymmetric Gaussian model function methods to process time series. The results are smoothed curves which can be used for extracting seasonal parameters like the beginning and the end of the vegetation season. The Savitzky-Golay and the asymmetric Gaussian can be classified as local filtering methods. The harmonic function uses the data for the full year and can be classified as global method. Jönsson and Eklundh (2004) stated that the Savitzky-Golay filtering is relatively unaffected by noise and it is able to follow complex behaviours. For noisy time-series it is necessary to apply restrictions to achieve a fixed functional form using the asymmetric Gaussian or the combined harmonic and polynomial basis.

2.1.6 Sentinel-2/Sentinel-3 time series fusion

Fusion of Sentinel-2 / Sentinel-3 time series is described in detail in the deliverable report D6.1_D31.1, as this is mainly covered by WP31. In phase 1, no fusion of Sentinel-2 / Sentinel-3 time series was performed, as the temporal frequency of Sentinel-3 is limited, because operational data are available from the Sentinel-3A satellite only up to now. Sentinel-3B satellite was launched recently successfully on 25th April 2019, but operational data is not expected to be available in the next months from Sentinel-3B. As lined out in the technical note on proposed substitutes for Sentinel-3 data [AD06], so far the data do not satisfy vital data requirements for land applications. Being primarily geared towards marine applications, the data pose still unresolved challenges for land applications such as unsuitable viewing

angles and missing spectral bands in the SWIR or TIR. Indispensable processing steps such as atmospheric correction and cloud masking are not yet available. Given the observed high quality of Sentinel-2 time-series with up-to five day revisit frequencies; the project has found the need to incorporate MR sensor data to be negligible in order to achieve the project aims of regarding HRL prototype production.

Nevertheless, in the absence of already usable high-density Sentinel-3 time series, the complementary use of alternative MR optical satellite data is being explored by ECoLaSS. In particular, the project makes use of PROBA-V data acquired during two years of corresponding to PROBA-V and Sentinel-2 acquisitions. An algorithm to downscale the spatial resolution of Sentinel-2 to the spatial resolution of PROBA-V 1km will be developed based on the Point Spread Function of the latter instrument. Sentinel-2 derived images will be visually and quantitatively compared to PROBA-V 1km images for the same date to insure a best possible radiometric and geometric correspondence. In a second step the production of a PROBA-V-like NDVI time series at 1km from Sentinel-2 data for the years 2017 and 2018 is performed which will be compatible with the long term SPOT Vegetation / PROBA-V time series. Lastly, NDVI anomalies obtained from the original SPOT Vegetation / PROBA-V time-series are compared to NDVI anomalies obtained from the PROBA-V-like NDVI time-series from Sentinel-2 data. The corresponding anomalies detection will be then be further investigated at 10m resolution to highlight the upscaling capacity of Sentinel-2 time series to document the detected anomaly. This comparison will be implemented for some tiles of Sentinel-2 covering part of ECoLaSS demo sites [AD17].

2.2 Processing methods for SAR time series

Processing methods for SAR time series include following steps: a) automated SAR image download, b) automated preparation of digital elevation data, c) automated update of orbit files (precise; restituted for near real time processing), d) thermal noise removal (for SLC) and GRD border noise removal, e) radiometric calibration, f) multi-looking and image filtering (speckle, adaptive), g) spatio-temporal speckle filtering (compare e.g. Quegan et al., 2000), h) radiometric terrain corrections (compare e.g. Small, 2011) and i) geo-referencing [AD04]. The research related to SAR time series pre-processing is, therefore, focusing on a) benchmarking of existing / currently developed algorithms and tools, b) implementation of spatio-temporal noise filtering tools optimized for Sentinel-1 time series, c) enhanced automated detection and correction of atmospheric influences, d) cluster processing to reduce processing time, e) and derivation of multi-temporal metrics [AD04].

2.2.1 Geometric pre-processing

In general, Sentinel-1 image data are accurate in terms of geometry and no additional adjustments seem to be required. A very extensive geometric location error analysis of Sentinel-1A and Sentinel-1B data has recently been published by (Schubert et al., 2017) for all Sentinel-1 data types. The mean absolute location error (ALE) for 133 Sentinel-1 GRD products, the data type which is primarily used in the ECoLaSS workflows, was found to be very small with a slant range error of 0.20m +/- 0.24, and an azimuth error of 0.31m +/- 1.31. The currently achieved geometric location accuracies for Sentinel-1, therefore, do not require any further improvements in the ECoLaSS project.

However, for Sentinel-1 GRD data processing it is essential to download the precise orbit files from ESA which are made available about 2 weeks after image acquisition at https://qc.sentinel1.eo.esa.int/aux_poeorb/. The delay of 2 weeks before precise orbits become available is a limiting factor for near-real-time applications. The ECoLaSS Sentinel-1 pre-processing workflow includes an automated co-registration of images in image geometry based on these precise orbit files. Creating a registered multi-temporal image stack is essential before applying any multi-temporal filtering. Based on the precise orbit information, all images of the same orbit are co-registered to a master scene before being filtered and orthorectified. Mean matching errors are below 1m standard deviation. Orthorectification is primarily based on the SRTM4 model which is available as a consistent layer for most of Europe. Since the SRTM model is of low spatial resolution, DEM errors related to topography are likely to occur in

mountainous regions during orthorectification. However, due to unfavourable local incidence angles, most SAR information in mountain regions is in general of low accuracy and should be masked out before classification. Flat areas do not show any significant mislocation errors related to orthorectification and/or orbit information.

2.2.2 Radiometric pre-processing

Sentinel-1 Level-1 data are the generally available products intended for most data users. Level-1 products are produced as Single Look Complex (SLC) and Ground Range Detected (GRD). Most image products are available in dual-polarization mode (HH+HV or VV+VH), except older data for the African test sites which are only available as single polarization data.

Level-1 SLC products consist of focused SAR data geo-referenced using orbit and attitude data from the satellite and provided in zero-Doppler slant-range geometry. The products include a single look in each dimension using the full TX signal bandwidth and consist of complex samples preserving the phase information.

Level-1 GRD products consist of focused SAR data that has been detected, multi-looked and projected to the ground range using an Earth ellipsoid model. Phase information is lost. The resulting product has approximately square resolution pixels and square pixel spacing with reduced speckle at the cost of reduced geometric resolution. The resolution is dependent upon the amount of multi-looking performed. Level-1 GRD products are available in MR and HR for IW and EW modes, MR for WV mode and MR, HR and FR for SM mode. For the ECoLaSS project, we are mainly using Level-1 GRD products in HR and IW mode (10x10m; 5x1 look) which is the standard product. For coherence estimation, SLC data is required. The spatial resolution of SLC data is 2.7x22m to 3.5x22m and the resulting pixel spacing (the distance between adjacent pixels in an image) is 2.3x14.1m.

The standard radiometric unit of Sentinel-1 GRD data is in db. Data in db can be converted to Digital Number (DN) which is the square root of the backscatter coefficient. This reduces the file size by 50% and facilitates multi-temporal metric calculation from SAR image stacks. Previous studies have shown DN based land cover classification results to be slightly more accurate than db based classification results (Figure 2-3). For projects dealing with several terabytes of SAR data, also the file size reduction of 50% is a critical issue.

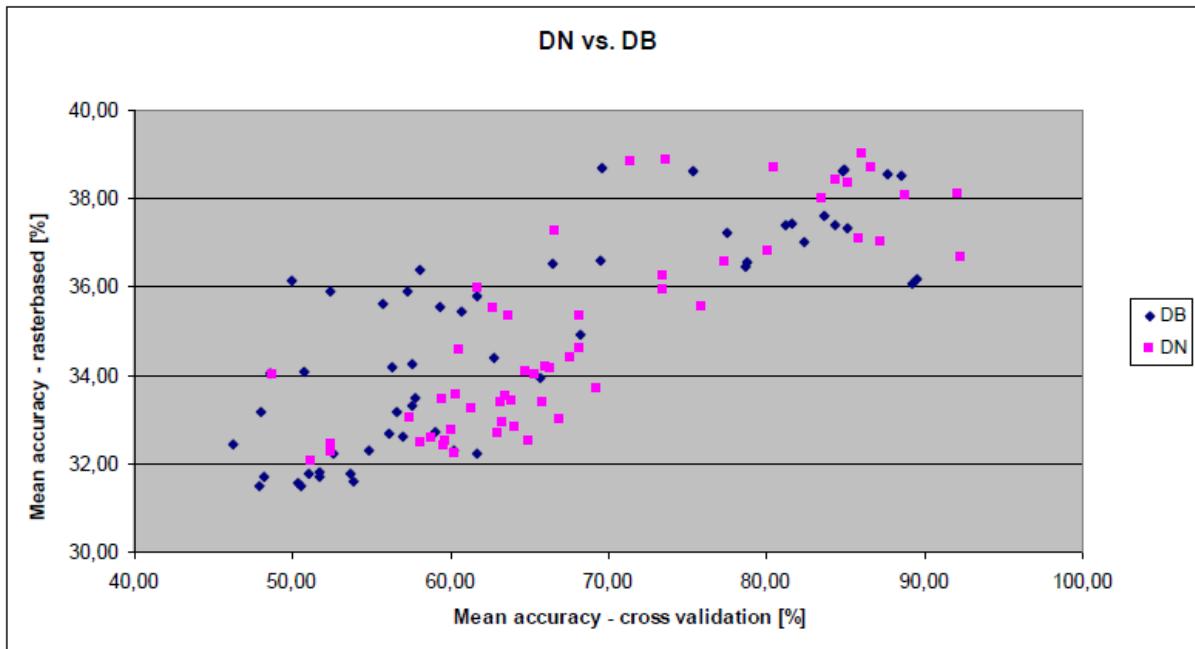


Figure 2-3: land cover classification accuracies for db and DN and different pre-processing scenarios

Radiometric calibration for Sentinel-1 data can be done simply by using the calibration Look-up Tables (LUTs) in the Calibration Annotation Data Set (CADS) delivered by ESA as part of the image metadata. The Sentinel-1 Level 1 GRD product includes several interpolated Look-Up Tables (LUTs) to convert intensity values into beta, sigma or gamma naught values. These LUTs are used by most SAR software applications to quickly perform radiometric calibration. The calibration values can however also be calculated directly. A detailed description of the calibration equations and available metadata can be found at:

<https://sentinel.esa.int/documents/247904/685163/Sentinel-1-Radiometric-Calibration-V1.0.pdf>

In addition, radiometric terrain corrections can be applied to sigma or gamma naught data in order to minimize the terrain effects in the SAR image. RSG and SNAP Sentinel-1 Toolbox use a method for gamma naught flattening described by (Small, 2011). Ellipsoid-based or sigma naught (σ^0) based incident angle approximations fail to reproduce the effect of topographic variation in their sensor model. The gamma naught (γ^0) backscatter is converted directly from beta naught (β^0) to a newly introduced terrain-flattened γ^0 normalization convention. The interpretability of imagery treated in this manner is improved in comparison to processing based on a conventional ellipsoid or local incident angle based σ^0 normalization, but the accuracies depend on the quality of the DEM used. An example of different processing levels and a terrain flattened gamma naught image are given in Figure 2-4 below.

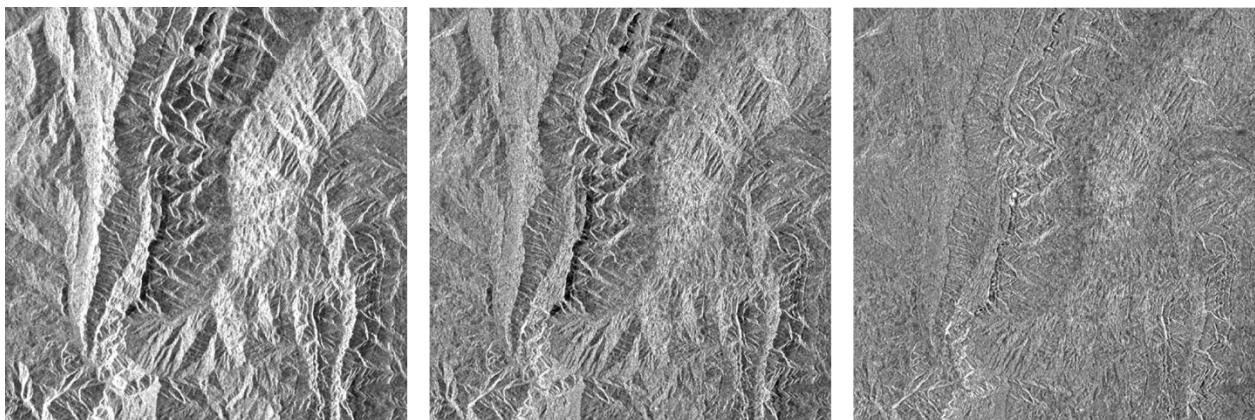


Figure 2-4: Different SAR radiometric processing levels: left: beta naught (slant range geometry), middle: sigma naught (corrected for local incidence angle), right: flattened gamma naught (terrain corrections according to Small, 2011).

2.2.3 Speckle noise reduction

SAR speckle filtering and multi-looking can be applied to reduce speckle noise in SAR imagery. Common speckle filters include Lee filter, Frost filter, Kuan filter, GammaMap filter and multiresolution speckle filters (Lee et al., 1994). Figure 2-5 shows image examples filtered with the Modified-Frost and the Multiresolution filter. However, for time series of SAR data, multitemporal SAR filtering is the most widely used procedure for noise reduction (Quegan et al., 2000). Input data to multitemporal filtering should not be filtered, multi-looked or orthorectified apriori.

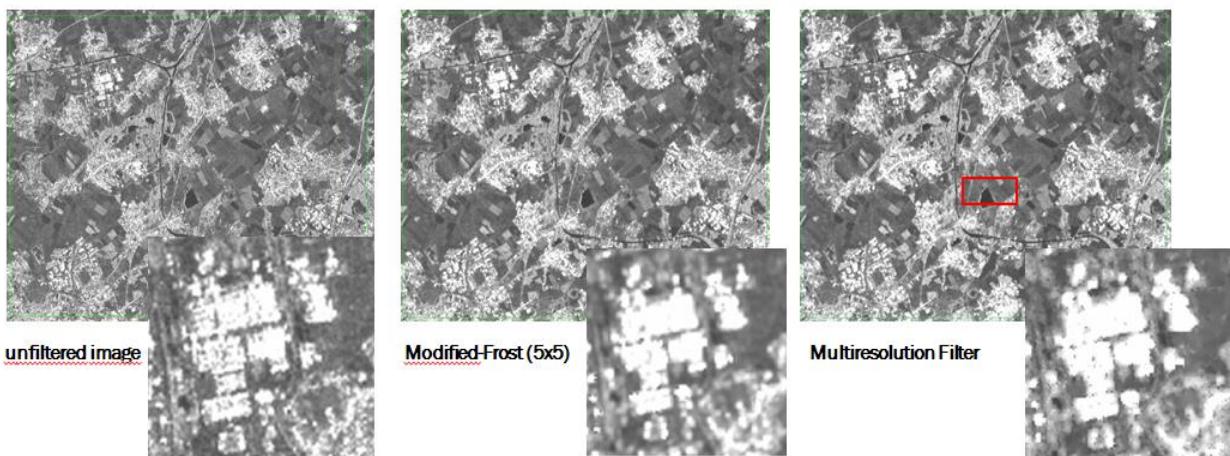


Figure 2-5: Example of two SAR speckle filters applied at an urban/rural test site

Speckle arises from random variations in the strength of the backscattered waves from objects and is seen mostly in RADAR images and is one of the most critical disturbances that regarding the quality of Synthetic Aperture Radar (SAR) coherent images (Ahmed et al., 2010). “Noise is any undesired information that contaminates an image. Speckle or Multiplicative noise, is a signal-dependent form of noise, whose magnitude is related to the value of the original pixel. This tends to reduce the image resolution and contrast.” (Gopinathan and Poornima, 2015) Noise reduction is an important pre-processing step for SAR imagery based land cover classification (Gopinathan and Poornima, 2015). Several adaptive and non-adaptive filters are widely used for despeckling in SAR images. In addition to multi-looking, which is already applied to GRD images, SAR speckle filtering can be used to reduce speckle noise in SAR imagery. Common SAR speckle filters include Lee filter, Frost filter, Kuan filter, Gamma-Map filter and multiresolution speckle filters. An overview of proved standard SAR speckle filters

is provided in Lee, 1994. However, for time series of SAR data, multitemporal SAR filtering is the most widely used procedure for noise reduction (Quegan, 2000).

Ahmed et al., (2010) proposed an adaptive mathematical morphological filter to reduce the speckle noise in SAR images. The results show, that the morphological filter gives promising results for significantly suppressing speckle noise and preserving the potential targets. “It can be easily implemented where erosion and dilation operators can be represented as minimum and maximum filters respectively instead of the exponential damping function that is dependent on speckle data model that should be known before applying filters.” (Ahmed et al., 2010).

Bouchemakh et al., (2006) tested several methods for speckle noise reduction including on filtering by Minimum Coefficient of Variation based on mathematical morphology, which uses local measurements of the noise in the image to process a low pass filtering operation over regions where the original signal is estimated to be homogeneous. The second method tested consists on empirical modal decomposition, which decomposes an image in four recursive images and reconstructing the image by detecting the edge direction. The last method tested is the Daubechies wavelets decomposition, which determines the position of real edges in the image so that false edges and noise can be removed. “Comparison of these three filtering techniques with existing methods show us that they were a good compromise between reducing speckle noise, texture conservation, and edge preserving. We have done a comparative study based on statistical criteria (mean, Standard-Deviation, Speckle Index) and Visual aspect. The results obtained by these three methods tested on SAR images of Algiers city, show that we have a good reduction of speckle noise with preserving edges” (Bouchemakh et al., 2006).

2.2.4 Interferometric coherence estimation

In the case of interferometric SAR, coherence is the normalized complex cross-correlation between two complex signals (two SAR images separated by a baseline) s_1 and s_2 and is defined as:

$$\gamma = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}}, 0 \leq |\gamma| \leq 1$$

where $(..)$ denotes an average over the ensemble of pixels, usually selected by a sliding window of size (azimuth x range) in a single look complex image. Interferometric coherence is essentially a complex variable, combining both the coherence magnitude and interferometric phase. In general, the measured coherence γ can be described as a product of the following factors:

$$\gamma = \gamma_{System} \gamma_{SNR} \gamma_{Temp} \gamma_{Vol}$$

where γ_{System} combines decorrelation caused by measurement system quantization, ambiguities, relative shift of the Doppler spectra and the baseline, γ_{SNR} describes the coherence decrease caused by the finite sensitivity of the system (the signal to noise ratio), γ_{Temp} accounts for changes in the target over time and γ_{Vol} describes decorrelation caused by volume scattering over vegetated areas where several scatterers at different heights contribute to scattering. The two last terms depend on target properties and have the largest dynamics when measuring natural targets.

For land cover mapping, interferometric coherence can be of great value since regions of vegetation suffer from temporal decorrelation (Zebker and Villasenor, 1992). The standard coherence estimation is

based on a local complex cross-correlation and is known to over-estimate the real coherence value, especially in areas of low coherence (Touzi et al., 1999). In general, a larger window within cross-correlation provides a better, i.e., less biased, coherence estimate. Until recently, the standard procedure was to estimate the coherence over the same window used for multi looking. However, with high resolution SAR imagery, the coherence is highly over-estimated resulting in a noisy coherence image. Therefore, a decoupling of the window size of multi looking and cross-correlation was introduced. The resulting coherence estimate uses a variable correlation window and is a trade-off between a rather unbiased coherence estimation and a locally well-defined unblurred estimate. The specific size of the correlation window needs to be empirically determined.

A new method for coherence estimation has recently been published by Spaans and Hooper, 2016. Their new algorithm calculates the coherence for each point based on an ensemble of points with similar amplitude behaviour throughout the data set. The resulting coherence estimate is superior in resolution and noise level to the commonly used boxcar method (Figure 2-6), which has three main drawbacks. First, it suffers from a resolution problem since points with high amplitude (e.g., buildings or natural ridges) dominate the coherence estimation but leads to a smearing out of these features. Second, it tends to overestimate the coherence of a large amount of fully decorrelated points because of the random signal being similar for neighbouring points by chance. Third, as the method essentially measures the variability of phase within the window, any non-constant signal biases the coherence estimation, with high-phase gradients leading to low coherence estimates. The new method is also faster to calculate and, in contrast to existing methods, can be used in near real-time applications.



Figure 2-6: Comparison between three different full resolution coherence estimates of the same interferogram, (20090618-20090629): (a) A 5×5 window boxcar, (b) 11×11 window boxcar, and (c) sibling-based (25 to 100 siblings per point) 41×41 search window). The boxcar coherence estimates many false high coherence points, such as in the field outlined in red. Source: Spaans and Hooper, 2016.

For land cover classification and change detection, both VV and VH coherence can be used and can have advantages depending on the change detection that is required. Some recent papers analysing Sentinel-1 12-day coherence changes for agricultural applications and grassland mowing detection describe slightly higher coherence changes for VH than for VV (Zalite et al., 2014; Tamm et al, 2016). VH coherence seems to show slightly higher sensitivity for low vegetation height/density changes (e.g. grass height changes from mowing) while VV seems to be more sensitive regarding vegetation to bare soil changes. But overall differences are quite small. Some studies for built-up area detection present best results when combining VH and VV coherence. Which polarization is better suited therefore seems to be related to the application and the type of change we are looking for in ECoLaSS. It is not possible to make a general statement. Most research studies also include soil moisture products and/or local and regional precipitation data for their coherence analysis to remove unreliable SAR images affected by soil moisture. This is beyond the scope of the ECoLaSS project as in practice this is only applicable at local/regional scale. On the pan-European level, auxiliary data would need to be harmonized which is beyond the scope of the project.

2.2.5 Available SAR processing software tools

Methods and algorithms for SAR pre-processing and coherence estimation are implemented in different software packages, of which several are shortly discussed in the following paragraphs.

ORFEO TOOLBOX

Orfeo Toolbox includes several SAR processing tools but some functionalities are limited. There are a number of pre-defined recipes for SAR processing - information can be found in the OTB Cook Book current version 6.00 (<https://www.orfeo-toolbox.org/CookBook>). These tools can handle Sentinel-1 data both at SLC level and GRD processing level. It includes different radiometric processing levels (beta, sigma, gamma) but values are only interpolated from the metadata look-up tables delivered by ESA. A speckle filter can be applied. Frost, Lee, Gamma-MAP and Kuan filters are available. Polarimetric decompositions are possible, but this tool requires full polarimetric data with at least 3 polarizations (HH, HV, VH or VV). This is not the case for Sentinel-1. A multi-temporal SAR filter is not included. There are two kinds of decompositions: coherent ones and incoherent ones. In the coherent case, only the Pauli decomposition is available. In the incoherent case, there are following decompositions available: Huynen, Barnes, and H-alpha-A. The user must provide three one-band complex images HH, HV or VH, and VV.

For SLC data, the toolbox includes a SAR debursting tool required for Sentinel-1 processing. InSAR applications include modulus and phase calculations. Coherence processing is not included in the CookBook but generally seems supported by Orfeo + Monteverdi toolboxes. No information could be found regarding the specific calculation of short-term coherence products. Parallel execution is supported. BandMath tools can be used for different types of (statistical) image analysis.

SNAP SENTINEL-1 TOOLBOX

The Sentinel-1 toolbox is a very powerful tool for processing both Sentinel-1 GRD and SLC data. It supports command line processing (GPT tools) and easy GUI based workflow generation.

A large number of SAR processing tutorials for the Sentinel-1 Toolbox can be found online at ESA:

<http://step.esa.int/main/doc/tutorials/sentinel-1-toolbox-tutorials/>

The software functionalities include:

- Data import (GRD and SLC)
- Automated orbit information update
- Specific Sentinel-1 TOPS processing tools
- Radiometric calibration (based on metadata look-up tables only)
- Radiometric terrain correction
- Multi-looking
- Speckle filtering (several filters)
- Stacking tool
- Image co-registration tools (with and without DEM)
- Multi-temporal filtering
- Linear to dB conversion
- SAR mosaicking tool
- Orthorectification
- Interferometric tools include:
 - o **Coherence estimation** (no time window selection – a priori stacking of relevant images needed)
 - o Interferogram formation

- Phase unwrapping
- InSAR filtering tools

Coherence estimation with SNAP Sentinel-1 toolbox:

This operator computes the coherence image, with or without subtraction of the reference phase. The reference phase is subtracted if there is a 2d-polynomial computed as a result of the "Compute Interferogram" operator. It is not subtracted if this information is not included within the metadata, or if the number of polynomial coefficients in "Compute Interferogram" operator is set to 0. This is a "general" coherence estimation operator and not exclusive only for InSAR applications. It can be utilized to estimate the coherence information from any stack of co-registered complex images. In order to reduce the noise, as the post-processing step, you can perform multi-looking (with Multi-look Operator). So far, there is no specific tool available for short-term coherence estimation from Sentinel-1.

DORIS

The Delft Institute of Earth Observation and Space Systems of Delft University of Technology has developed an Interferometric Synthetic Aperture Radar (InSAR) processor named Doris (Delft object-oriented radar interferometric software). The Doris software can be downloaded freely, but only for non-commercial applications (conditions). The software includes a tool for coherence estimation. More information on <http://doris.tudelft.nl/>

GMTSAR

GMTSAR is an open source (GNU General Public License) InSAR processing system designed for users familiar with Generic Mapping Tools (GMTs). The code is written in C and will compile on any computer where GMT and NETCDF are installed. The system has three main components:

1. a pre-processor for each satellite data type (ERS-1/2, Envisat, ALOS-1, TerraSAR-X, COSMOS-SkyMed, Radarsat-2, Sentinel-1A/B, and ALOSentinel-2) to convert the native format and orbital information into a generic format;
2. an InSAR processor to focus and align stacks of images, map topography into phase, and form the complex interferogram;
3. a postprocessor, mostly based on GMT, to filter the interferogram and construct interferometric products of phase, coherence, phase gradient, and line-of-sight displacement in both radar and geographic coordinates;

GMT is used to display all the products as postscript files and KML images for Google Earth. A set of C-shell scripts has been developed for standard 2-pass processing as well as image alignment for stacking and time series.

RSG- REMOTE SENSING SOFTWARE PACKAGE GRAZ

The project partner JOANNEUM RESEARCH developed a comprehensive software package for SAR data processing. The Remote sensing Software package Graz (RSG), which is designed for geometric processing options for almost any kind of remote sensing image data, acquired from airborne as well as space-borne platforms and comprising the geometric aspects of optical line scanners, SAR systems, perspective images and rational polynomial modelling. RSG covers nearly all of the geometric and radiometric processing options being specified for the processor components of the processor suite.

The basic elements of the RSG software package are: Geometric Modelling tool box, being implemented for any kind of remote sensing image data and being based on ground control points and least squares adjustment techniques. This is also includes Block Adjustment, being applicable to entire blocks of

remote sensing image data; Geocoding tool box, including various other value-adding options like high-precision geocoding integrating digital elevation models, polynomial-based image registration, matching-based image registration, semi-automatic geocoding utilizing control point chips and sensor model based image-to-image or map-to-image registration; Stereo Mapping tool box, including geometric modelling, a variety of image filtering and image matching tools, determination of ground coordinates from matching results and subsequent generation of digital elevation models, optionally based on triangulation procedures; SAR Interferometry (InSAR) tool box, comprising registration of single look complex (SLC) SAR image pairs, generation of an InSAR interferogram and multi-look products, phase unwrapping, determination of ground coordinates and generation of digital elevation models, differential InSAR processing methods and generation of innovative products like topograms, slope maps or velograms from InSAR data sets. The software is available commercially.

3 Testing and benchmarking of candidate methods

In ECoLaSS, the processing of high volume data processing lines mainly based on the high resolution satellite constellations is addressed. This topic is important for the pan-European component, where a large volume of satellite earth observation data at high resolution has to be processed on a continental level. The main challenges are the combination of current processing modules into high volume data processing lines, as well as upscaling from local or country level to pan-European or potentially global level in the future. Therefore, a large volume of data needs to be efficiently processed to produce data composites and derive calibrated and validated variables.

3.1 Test and Demonstration sites

In order to conduct an in-depth analysis five test sites within different major European environmental zones and two test sites in Africa have been chosen. Pre-processing of optical and SAR time series for all test sites, as mentioned in Table 3-1 has been performed to test and identify the best suited algorithms, to develop new methods and implement prototype processing software.

Table 3-1: Test sites and corresponding demo sites in phase 2.

Test sites	Defined test-site granules	Partner responsible for pre-processing	Demo sites	Defined demonstration site granules	Partner responsible for pre-processing	Biogeographic regions
European						
Sweden	33VVF 33VWF	DLR	North	33VVH, 33VWH, 33VVG, 33VWG, 33VWF	DLR	Boreal
Austria/ Germany	32UNU 32TNT	GAF	Central	32UMV, 32UNV, 32UPV, 32UMU, 32UNU, 32UPU, 32TMT, 32TNT, 32TPT	GAF	Alpine/ Continental
Belgium	31UFR 31UFS 31UES 31UER	JR	West	31UFR, 31UFS, 31UES, 31UER, 31UFQ, 31UEQ	JR	Atlantic /Continental
France	30TYP 31TCJ	SIRS	South-West	30TYP, 30TYN, 31TCJ, 31TCJ, 31TDJ, 31TDH	SIRS	Atlantic
Greece/Bulgaria	34TFM 35TFL	UCL	South-East	34TFN, 34TGN, 34TFM, 34TGM 35TKG, 34TFL, 34TGL, 34TKF	UCL	Mediterranean /Alpine/ Continental
African						
Mali	29PRP 29PTU	UCL	Mali	29PPQ, 29PNQ, 29PMP, 29PRM, 29PRP, 29PRQ, 29PPP, 29PLP, 29PMN, 30PUA, 29PRR, 30PTT, 29PPM, 29PQN, 30PTA, 30PTS, 29PNS, 30PTU, 30PUV, 29PMQ, 29PLQ, 29PMR, 30PTV, 29PQR, 29PLR, 29PQM, 30PVA, 29PPR, 29PMS, 30PUU, 29PNN, 29PQQ, 29PPN, 29PNR, 29PNP, 29PNM, 29PRN, 29PQP, 30PVV	UCL	semi-arid regions
South Africa	35JMJ 35JNJ	UCL	South Africa	34JBL, 33HYC, 34HDH, 34JCL, 34HFG, 34HFH, 33HYB, 34HFJ, 34HGJ, 33JYG, 34HBG, 34HBK, 34HGH, 33HYD, 34HDG, 34HCG, 34HDJ, 34HEG,	UCL	natural landscapes combined with subsistence agriculture

				34HBJ, 34JBM, 34HBH, 34HDK, 34HEH, 34HCK, 33HYE, 35HKC, 33JYF, 34HCJ, 34HCH, 34JCM, 34HEJ		
--	--	--	--	---	--	--

Table 3-1 shows the five European test sites with their corresponding demonstration sites. Each demo site contains a representative test site of two granules for software and algorithm testing and development. The location of these test sites within the demonstration sites is highlighted in red colour.

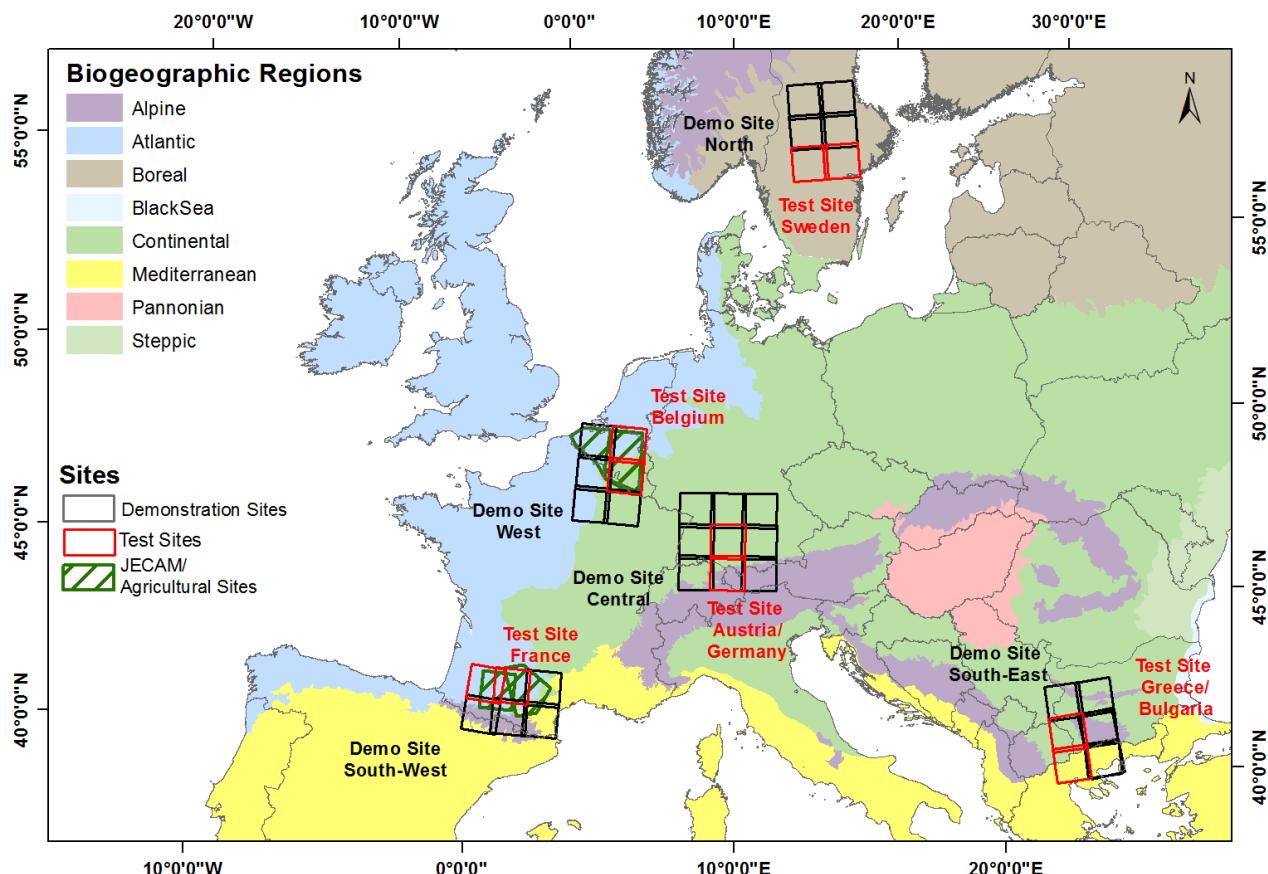


Figure 3-1: Demonstration sites which are processed in Task 4 are drawn in black, test sites which are processed in Task 3 are drawn in red.

The test sites in South Africa as well as in Mali are shown in Figure 3-2.

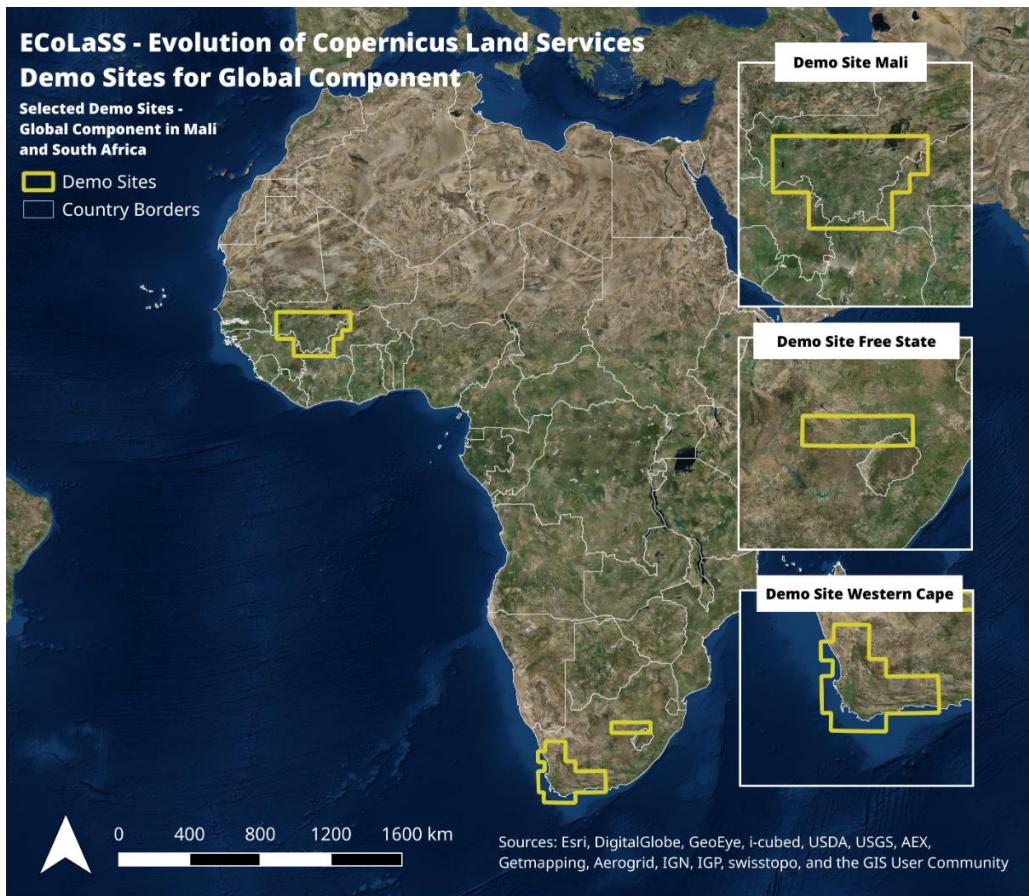


Figure 3-2: ECoLaSS test sites in Africa Demonstration sites which are processed in Task 4 are drawn in yellow.

3.2 Benchmarking of methods for pre-processing of optical time series

The processing methods for optical time series include the generation of spatio-temporally consistent optical time series with Bottom-of-Atmosphere reflectance values. Considering the data size and number of scenes within the time series, an automated and rapid processing chain, including all necessary pre-processing steps, is essential. The following sections provide methods and benchmarking thereof, focussing on the topics of atmospheric correction, cloud, cloud shadow and snow masking, topographic normalization and geometric consistency of optical time series data. The candidate methods for pre-processing optical time series, benchmarking criteria, the implementation and results of benchmarking and resulting conclusions are described in the sub-sections of this chapter.

3.2.1 Benchmarking of candidate approaches for atmospheric correction

Applying an atmospheric correction is particularly important when dealing with multitemporal data as it cannot be assumed that illumination and atmospheric conditions are the same for all images. Differences between the images caused by atmospheric influences would influence the accuracy of the classification result. Cloud masking and land cover classification tasks require an atmospheric correction beforehand.

Sentinel-2 data sets are downloaded for the ECoLaSS test and demonstration sites, from which spatio-temporally consistent optical time-series with Bottom-of-Atmosphere reflectances are being generated. In order to obtain the Bottom-of-Atmosphere reflectances for optical data, several processors like Sen2Cor (processor for Sentinel-2 Level-2A product generation) and MAJA (Maccs-Atcor Joint Algorithm)

are available. Within the CEOS context, ESA and NASA are evaluating those algorithms among others in an inter-comparison exercise project named ACIX (Atmospheric Correction Inter-comparison eXercise) (<http://calvalportal.ceos.org/projects/acix>) for understanding the different uncertainty contributors and improving the available atmospheric correction processors. The first, exploratory round of ACIX has been completed in 2018 and is currently being updated with a more rigorous testing regime in ACIX II.

The official method for the atmospheric correction of Sentinel-2 data used in the Sentinel toolbox provided by ESA is Sen2Cor.

3.2.1.1 Description of Candidate Methods

SEN2COR BY ESA

Sen2Cor presented from ESA as Sentinel-2 Level-2A processor provides a set of algorithms, which is composed of several state-of-the-art techniques performing the task of correcting atmospheric-, terrain and cirrus effects based of Top-of-Atmosphere Level-1C input data. As a result Bottom-of-Atmosphere products and if required terrain- and cirrus corrected reflectance images are created. Furthermore, parameters like the atmospheric visual transparency, the aerosol type and the water vapour column are needed for the atmospheric correction process and derived through the associated bands (Müller-Wilm et al., 2013).

MAJA BY CNES-DLR

The MAJA processor uses multi-temporal information to detect clouds, cloud shadows and to estimate the optical properties of the atmosphere (Rouquié et al., 2017). The algorithm detects areas covered with clouds, cloud shadows, snow and water based on multi-temporal data sets and generates corresponding masks. Further, two components are computed for the atmospheric correction, the atmospheric absorption and the atmospheric scattering. Another contribution which will be corrected by the algorithm is the blurring due to the landscape surrounding the site imaged, i.e. the adjacency effects (Lonjou et al., 2016; Rouquié et al., 2017).

3.2.1.2 Benchmarking Criteria

A comparative quality assessment of atmospheric correction methods can be based on following criteria, which are currently considered as standard in literature. There are three main approaches for evaluating surface reflectance products.

COMPARISON WITH GROUND TRUTH DATA

Gascon et al., 2017 and Pflug et al., 2016 used the AERONET sunphotometer measurements to validate the Sentinel-2 Bottom-of-Atmosphere product. The validation datasets have been selected for radiometric validation. They cover different land cover types, different atmospheric conditions and include different latitudes in order to cover various solar angles and seasons. The objective is to estimate the uncertainty of BOA product resulting from processing Sentinel-2 data using Level-2A processor Sen2Cor. Further Doxani et al., 2018 compared the Surface Reflectance products from several processors based on Landsat-8 and Sentinel-2 input data, to reference data sets computed by 6S radiative transfer code. The AOT, aerosol model and the column water vapour were derived from the AERONET sunphotometer measurements.

INTER-ORBIT SPECTRAL STABILITY

Another approach is to check the surface reflectance stability by comparing two images of the same sensor in areas from overlapping orbits. Therefore, a region in Germany has been selected, comparing

two overlapping images with clear skies and observed from two different orbits on two different dates in summer 2015, which are close together.

CONSISTENCY WITH OTHER SENSORS

A third approach checks for the consistency of the Top-of-Atmosphere products derived from comparable sensors. The consistency between the different atmospheric correction approaches is needed, to better combine various products for analysis. To check the consistency of the Sentinel-2 Level-2A products derived from the Sen2Cor processor (Version 2.2.1), Sentinel-2 MSI is compared with two other sensors with similar spectral resolution (LANDSAT 8 OLI and TERRA/AQUA MODIS).

3.2.1.3 Implementation and Results of Benchmarking

COMPARISON WITH GROUND TRUTH DATA

Comparison with ground truth data was performed e.g. by Gascon et al. (2017). The validation of the surface reflectance products was thereby performed comparing Sen2Cor outputs with surface reflectance reference data generated with sunphotometer measurements by the AERONET instrument. Therefore, a 9km x 9km subset around the sunphotometer location was created and a reference BOA reflectance is generated by running the Sen2Cor atmospheric correction with a constant AOT value equal to the AOT value measured. The results are considered as “truth” regarding the surface reflectance since the greatest uncertainty in atmospheric correction comes from the aerosol characterization (Gascon et al., 2017). The comparison shows small differences in surface reflectance between both products.

Pflug et al., (2016) summarize: “*The Level-2-product is clearly improved in comparison with the Level-1C-product. [...] DDV-pixels are distributed over the full image. The granule contains 3.1% DDV-pixels, which is sufficient for an accurate estimation of aerosol amount. [...] Both AOT and WV averages across the subset are in good agreement with the reference value obtained from AERONET. Sen2Cor provides an AOT value at 550 nm of 0.20 with reference value from AERONET of 0.23. WV is estimated from Sen2Cor with a value of 2.46 cm with a reference value from AERONET of 2.63 cm. Several example spectra were extracted in the reference image and in the image to validate. Example spectra for dark and bright soil, for forest and for different other vegetated locations show the expected spectral dependency and agree between reference and the spectra to be validated. A difference up to 0.04 in surface reflectance between Sen2Cor-processing and reference had been found, leading to a Normalized Density Vegetation Index (NDVI) uncertainty up to 0.06*”.

The ACIX comparison protocol also draws heavily on AERONET sunphotometer references. The first round of ACIX has been completed in 12/2016 and is summarized in Doxani et al. (2018). In total eleven processors capable of performing atmospheric correction for Sentinel-2 MSI have been compared. Much of the comparison revolved around the plausible derivation of the intermediate layers “water vapour” and AOT, both of which can be directly validated against sunphotometer measurements, as well as surface reflectance, which was validated against simulations by the 6S radiative transfer model based on Level-1C data and the sunphotometer measurements. For end-users, like ECoLaSS, reliable surface reflectance retrieval is of main interest.

In general most algorithms offered a comparable performance regarding the surface reflectance retrieval. The best estimates were found by FORCE, LaSRC, MACCS (MAJA precursor) and Sen2Cor. The former three candidate approaches succeeded in all test cases, while Sen2Cor did exhibit difficulties for some sites in correcting infra-red bands (Doxani et al. 2018). In general, arid scenes without dark dense vegetation pixels posed challenges to all algorithms with image-based AOT retrieval.

MSI Band	CorA	FORCE	iCOR	LaSRC	MACCS	S2-AC2020	GFZ-AC	Sen2Cor	
	nbp	23873202	29568870	23808647	36863274	12538144	34243490	34159390	30335882
1	A	-0.006	-0.002	-0.010	-0.010	-	-0.006	0.026	-0.003
	P	0.096	0.009	0.024	0.010	-	0.017	0.014	0.011
	U	0.096	0.009	0.026	0.014	-	0.018	0.029	0.011
2	A	0.000	-0.004	0.000	-0.007	-0.008	-0.004	0.023	-0.001
	P	0.021	0.007	0.028	0.008	0.010	0.021	0.016	0.009
	U	0.021	0.008	0.028	0.011	0.013	0.022	0.029	0.009
3	A	0.003	-0.012	0.013	-0.005	-0.008	0.000	0.031	0.004
	P	0.024	0.006	0.034	0.006	0.008	0.023	0.023	0.010
	U	0.025	0.014	0.036	0.008	0.012	0.023	0.039	0.011
4	A	0.002	-0.007	0.018	-0.003	-0.007	0.002	0.022	0.006
	P	0.027	0.005	0.036	0.006	0.007	0.025	0.020	0.012
	U	0.027	0.009	0.040	0.007	0.010	0.026	0.030	0.013
5	A	0.008	-0.008	0.027	-0.002	-0.005	0.007	0.031	0.020
	P	0.029	0.005	0.038	0.006	0.006	0.012	0.022	0.018
	U	0.030	0.009	0.046	0.006	0.008	0.014	0.038	0.027
6	A	0.005	0.001	0.024	-0.001	-0.003	0.004	0.024	0.017
	P	0.032	0.005	0.033	0.005	0.006	0.010	0.042	0.011
	U	0.032	0.005	0.041	0.005	0.007	0.011	0.049	0.021
7	A	0.006	-0.002	0.025	-0.003	-0.007	0.005	0.020	0.014
	P	0.033	0.005	0.031	0.005	0.005	0.009	0.047	0.010
	U	0.034	0.006	0.040	0.005	0.008	0.010	0.051	0.017
8	A	0.008	0.017	0.032	0.001	-0.001	0.011	0.025	0.022
	P	0.033	0.010	0.034	0.005	0.005	0.026	0.047	0.014
	U	0.034	0.019	0.047	0.005	0.006	0.028	0.053	0.026
8a	A	-0.008	0.000	0.023	-0.002	-0.008	0.003	0.016	0.013
	P	0.033	0.005	0.028	0.004	0.005	0.011	0.049	0.008
	U	0.034	0.005	0.036	0.005	0.009	0.011	0.051	0.015
11	A	0.021	-0.010	0.018	0.002	-0.003	0.009	0.017	0.020
	P	0.035	0.005	0.019	0.003	0.003	0.007	0.011	0.009
	U	0.041	0.011	0.026	0.003	0.004	0.011	0.020	0.022
12	A	0.020	0.004	0.013	0.004	0.000	0.008	0.014	0.025
	P	0.030	0.006	0.013	0.003	0.002	0.006	0.019	0.014
	U	0.036	0.007	0.018	0.005	0.003	0.010	0.024	0.028

Figure 3-3: “Accuracy (A), precision (P), and uncertainty (U) scores per band for the S-2 SR products of every processor and over all test sites. The number of points (nbp) involved in the APU analysis varies due to the different number of S-2 scenes processed and submitted by every processor.” (Doxani et al., 2018)

It is worth noting, however, that the tests took place in a very early phase of Sentinel-2 availability (2016) and do not represent the most up-to date capabilities of each processor. Moreover, the comparison framework was not strictly standardized. And multi-temporal approaches, such as MAJA, could not yet exploit their full potential. Results should therefore not be interpreted as definitive reference. The currently ongoing ACIX II efforts will address these issues.

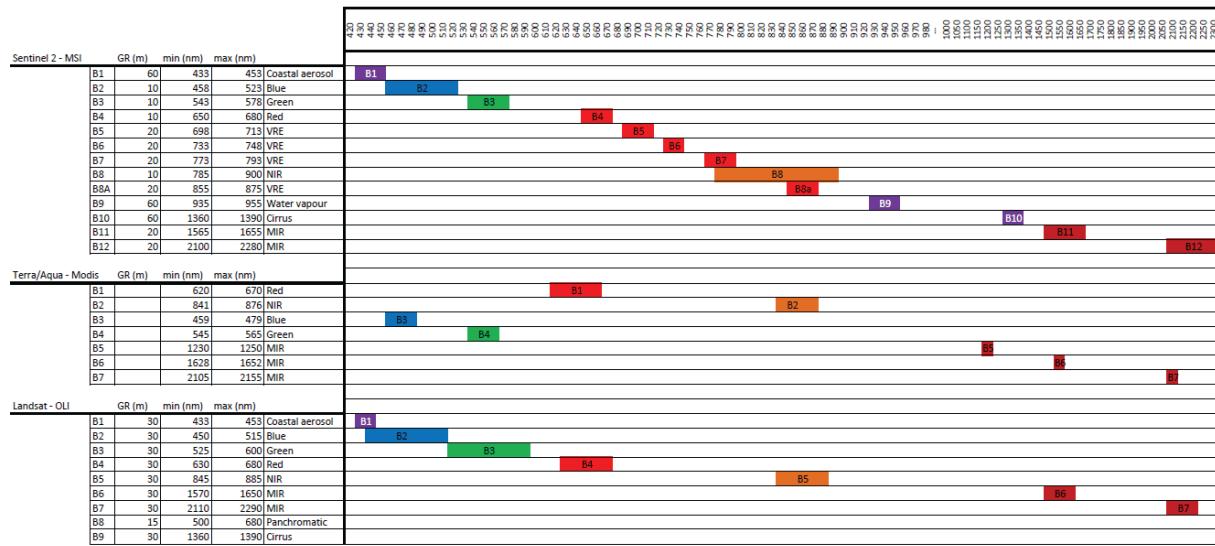
INTER-ORBIT SPECTRAL STABILITY AND CONSISTENCY WITH OTHER SENSORS:

To check the consistency of the Sentinel Level-2A products derived from the Sen2Cor processor (Version 2.2.1), we compared Sentinel-2 MSI with two other sensors with similar spectral resolution (LANDSAT-8 OLI and TERRA/AQUA MODIS). Figure 3-4 highlights the different bands for all three sensors and also the differences in bandwidths for corresponding spectral bands of each sensor. It also shows which bands are comparable with each other over all three sensors. With regards to OLI, the spatial resolution of MSI is finer except regarding the bands used for the atmospheric correction. The major spectral difference between OLI and MSI is the presence of red-edge bands in MSI and the thermal bands in OLI. Regarding MODIS the spatial resolution of MSI and OLI is much finer. In order to compare the sensors, following bands are chosen and resampled to 920m, representing almost twice the MODIS resolution (see Table

3-2). MODIS and OLI sensor products adopted for this analysis are downloaded already containing reflectance values. The Sentinel-2A scenes are processed with the Sen2Cor processor.

Table 3-2: “Corresponding bands regarding each spectral region across MSI, OLI and MODIS.” (Miletich, 2017)

	MSI	MODIS	OLI
RED	B4	B1	B4
NIR	B8a	B2	B5
SWIR	B11	B6	B6
Resolution	20m	463m	30m



MSI = multi-spectral instrument; OLI = Operational Land Imager; Modis = moderate-resolution imaging spectroradiometer; GR = groundresolution; NIR = Near Infrared; SWIR = Shortwaved Infrared; VRE = Vegetation Red Edge;

Figure 3-4: “Comparison between the MSI, MODIS and OLI sensors. Based on: U.S. Geological Survey (2016), Suhet (2015), and NASA Official.” (Miletich, 2017)

First, regions of interest are defined for the thematic classes water bodies, broad-leaved forest, coniferous forest and urban area in the overlapping area of all images to be compared including Sentinel-2A MSI of 26.08.2015, Sentinel-2A MSI of 29.08.2015, TERRA/AQUA MODIS of 29.08.2015, and LANDSAT 8 OLI of 30.08.2015. One region of interest for each class is defined containing homogeneous pixels in the resampled products and centred in the land cover feature to avoid adjacent land cover pixels influencing the comparison. The chosen classes can be considered as steady and therefore changes during short time intervals can be neglected.

Figure 3-5 displays the differences between the mean reflectance values of each region in the RED channel, presenting similar results over all three sensor systems for each example area. Especially between MODIS and OLI only small differences are observed. In relation to the other sensors, the reflectance values for OLI within the water body area are twice as high. Those differences can be explained by the presence of thin cirrus clouds. Nevertheless, it should be considered that in relation to the other classes or in absolute terms the differences are minor. Regarding broadleaved and coniferous forests the MSI sensor shows slightly lower reflectance values in both scenes, indicating that the Sen2Cor processor tends to under-correct the reflectance values. However, as shown in the statistics the reflectance differences in the red region are less than one percent.

Figure 3-6 represents the differences between the mean reflectance values of each region in the NIR channel, again showing similar results across all three sensor systems with one exception. In the case of the broadleaved region, the MSI sensor (29.08.2015) displays extreme low reflectance values in relation to the other sensors and the MSI scene acquired on 26.08.2015. A visual inspection of the images illustrates a difference between the processed granules within the Sentinel scenes. The test sites are located in two different granules. The reflectance values of the upper granule, including the broadleaved and coniferous forest test sites explain that Sen2Cor tends to under-correct due to haze within the image. Since broadleaved and coniferous forest is located in the same granule, it would be expected that the coniferous forest reflectance values would also be underestimated. Nevertheless, the reflectance values of coniferous forest are at least 20% lower than the broadleaved reflectance values; accordingly the differences are smaller.

In Figure 3-7 the differences between the mean reflectance values for the SWIR channel are represented, illustrating more and higher differences over all three sensor systems. In comparison with the RED and NIR channels, the Sentinel scenes also seem over-corrected in relation to the other two sensors. Regarding all example spectra, the Sentinel MSI sensors shows that Sen2Cor tends to over-correct expect for the broadleaved forest area, where the LANDSAT scene seems slightly under-corrected. Nevertheless, it should be noted that the absolute reflectance differences are less than 2%.

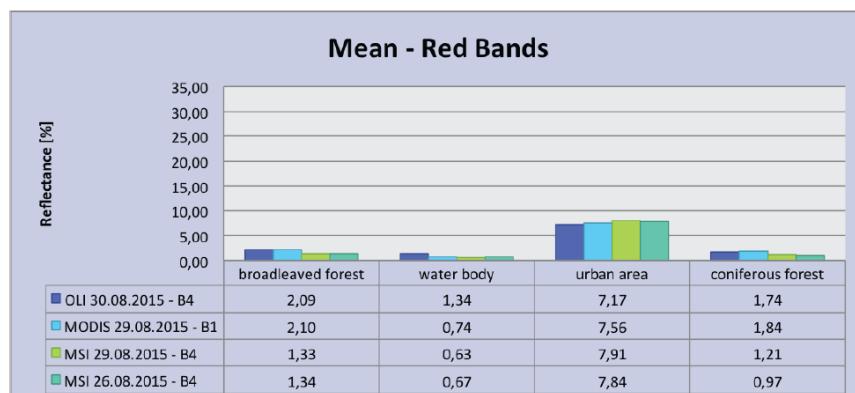


Figure 3-5: “Mean reflectance - RED band.” (Miletich, 2017)

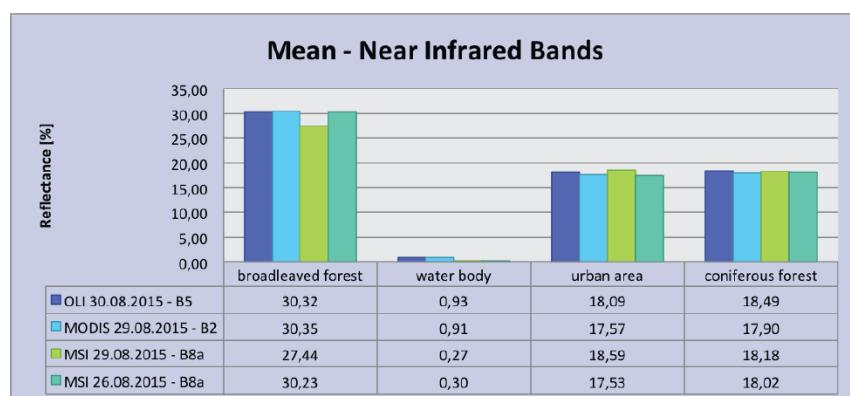


Figure 3-6: “Mean reflectance - NIR band.” (Miletich, 2017)

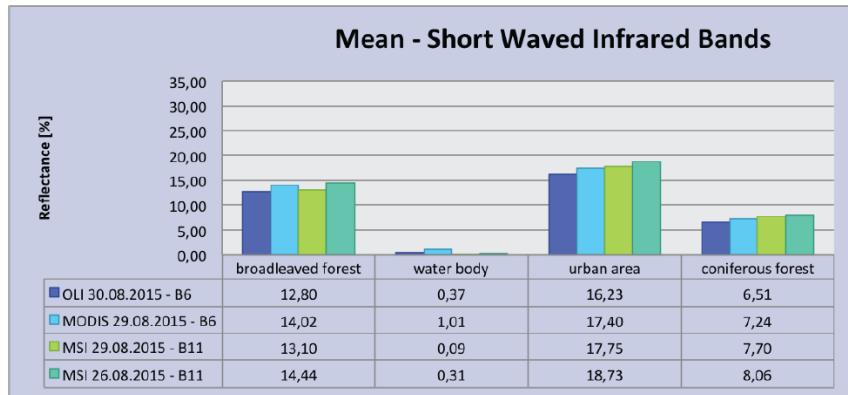


Figure 3-7: "Mean reflectance - SWIR bands." (Miletich, 2017)

These differences are probably due to directional effects as the observations from two different orbits are made with different viewing angles.

3.2.1.4 Summary and Conclusions

In general, the example spectra for coniferous forest, broad-leaved forest, water and urban area show the expected spectral agreement between reference spectra and images spectra to be compared. For the most part, atmospheric correction performed with Sen2Cor works consistent like shown in the comparison in chapter 3.2.1.3. Gascon et al. (2017) achieved in their early validation approach of the Sen2Cor processor similar results stating that the early Sentinel-2 Level-2A products show good inter-orbit reflectance comparability if there are DDV pixels present in the granule.

"The first results are very encouraging in term of image quality performance as well as in term of computing performance. Results obtained show the very good performances of the mission products both in terms of radiometry and geometry. Thanks to a robust in-flight calibration strategy, the radiometry is both accurate (<5% absolute uncertainty) and stable (<1%/year variation estimated). Cancelling seasonal effects on diffuser acquisition is the key to this performance: this involves an accurate model of the Sun-Earth distance and of the diffuser bi-directional Reflectance Function. Further progress on the latter point should lead to improved pixel response stability (i.e. Fixed Pattern Noise) in the near future." (Gascon et al., 2017).

Nevertheless, difficulties arise, when no dark dense vegetation pixels are within the granule. Small over-correction in comparison with the AERONET sunphotometer measurements is observed by Louis et al. (2016a), which may result from false aerosol parameters because Sen2Cor estimated higher aerosol optical thickness compared to the reference data. It should be noted that there is still a possibility that the reference data obtained from AERONET sunphotometer measurements can be under-corrected. The major problem is the granule wise processing of Sen2Cor, which results in granule borders within the scenes in cases of missing dark dense vegetation pixels.

Comparing the two main atmospheric correction software packages both show advantages and drawbacks and are not performing sufficiently enough to achieve a high thematic quality, content and accuracy in any subsequent information extraction process. Concerning the atmospheric correction the inter-orbit spectral stability and the consistency with other sensors show that Sen2Cor is able to achieve consistent results only if the data sets fulfil certain characteristics, e.g. presence of DDV pixels. MAJA atmospheric corrected products are not evaluated due to the fact that MAJA cannot be used in commercial projects such as HRL production, considering the licence and implementation issues.

3.2.2 Benchmarking of candidate approaches for cloud, cloud shadow and snow masking

Reliable identification of clouds and cloud shadows are necessary for any type of optical remote sensing image analysis, especially in operational and fully automatic setups. A significant improvement of current methods for the classification/masking of clouds, cloud shadows and snow is required for enabling a fully automated generation of highly accurate masks in near-real time. Whereas current methods, such as implemented in the Sen2Cor (Louis et al., 2016b) or LEDAPS software (Home et al., 2013), classify each scene separately, improved methods are currently being developed exploiting the additional information provided by time series.

3.2.2.1 Description of Candidate Methods

SEN2COR

Sentinel-2A products are provided with different already available cloud masks depending on the product level. Within the Level-2 product, the scene classification algorithms generate a classification map based on spectral threshold tests applied to the cirrus band and band ratios like NDVI and NDSI. The scene classification map provides four different classes of clouds (thin cirrus clouds, high, medium and low probability clouds), four land cover classes (vegetation, soil/deserts, water and snow) and two extra classes for shadows (differentiating between cloud shadows and terrain shadows) (Fletcher K., 2012; Müller-Wilm et al., 2013). Furthermore, Level-1C products are provided with spatially filtered cloud masks differentiating between opaque and cirrus clouds based on the scene classification map with 60m spatial resolution.

Apart from the scene classification map a probabilistic cloud map and a snow probability map, are produced during the processing with Sen2Cor. Within several iterations with different thresholds, the cloud probability map is updated during the Sen2Cor processing. The refined threshold values for the cloud detection algorithm are derived from LANDSAT 7 Automatic Cloud Cover Assessment. The opaque cloud detection algorithm uses the water vapour band (B1) and blue band (B2) in combination with two SWIR bands (B11, B12) to differentiate between snow and clouds. The algorithm first defines potential cloudy pixels within the red part of the spectrum through a certain threshold. Further, those pixels undergo a filtering process based on spectra thresholds, band ratios and the indices normalized difference snow index (NDSI), normalized difference vegetation index (NDVI). Pixels with values lower -0.1 in the NDSI are representing snow and are excluded as potential clouds. Vegetation pixels are excluded from the cloud map deriving information from NDVI and a NIR/Green ratio. Bare soil and water are also excluded from threshold with a blue/NIR ratio. A median filter (3x3 or 5x5) is applied on the three cloud classes to reduce false classifications occurring at border region with high contrast like river contours or shorelines (Louis et al., 2010).

Further, the cirrus band (B10) is used to detect ice high-altitude clouds, which are represented by high reflectance values in B10 and low reflectance values in B1 and B2 (Fletcher K., 2012; Müller-Wilm et al., 2013). The strength of the cloud band lies in the strong water vapour absorption within the mentioned spectral region. Thresholds are used for separating clear sky pixels from cloud pixels (Louis et al., 2010).

If a correlation of the cirrus signal in the cirrus band and the other wavelengths in the NIR region can be found, the cirrus contribution can be removed. Rudolf Richter et al. (2011) introduced an enhanced cirrus removal method for multispectral images including channels in the NIR and SWIR region as well as a water vapour band and a cirrus band. During the water vapour and surface reflectance calculation within the atmospheric correction process water vapour above cirrus clouds derived from the cirrus channel is considered, leading to more accurate assessments of the water vapour map. Hence, cirrus artefacts on the SWIR region can be avoided. As mentioned in chapter 6.1.1 radiative transfer look-up

tables are used for the water vapour and reflectance retrieval. Regarding the cirrus correction further look-up tables containing cirrus transmittances for different solar and viewing conditions are used (Rudolf Richter et al., 2011).

For the snow detection, four consecutive filters are applied on B2, B3, B8 and B11 reflectance values. The first filter is based on the assumption that cloud and snow reflectances are similar in band 3. On the contrary considering clouds B11 reflectance values are very high while for snow the reflectance values are low in B11. Therefore, the NDSI is calculated and thresholds are applied. The second filter eliminates pixels from the snow mask that are characterised by high NDSI values and low reflectance values in B8. The third filter removes pixels from the mask that have high NDSI values and low reflectance values in B2. The last filter eliminates water pixels from the snow mask by excluding pixels with a high NDSI value and a low ratio between B2 and B4 (Louis et al., 2010; Fletcher K., 2012; Müller-Wilm et al., 2013).

MAJA

The CNES MUSCATE production centre and the DLR produce Sentinel-2 surface reflectance data using the MAJA processor. The CNES MUSCATE production centre product generations covers areas shown in Figure 2-1 (Donadieu and L'Helguen, 2016).

The cloud mask product computed by the MACCS software is available in 10m or 20m (CLM_R1.tif or CLM_R2.tif) resolution and made out of one band coded over 8 useful bits explained below:"

- *bit 0 (1): all clouds except the thinnest and all shadows*
- *bit 1 (2): all clouds (except the thinnest)*
- *bit 2 (4): clouds detected via mono-temporal thresholds*
- *bit 3 (8): clouds detected via multi-temporal thresholds*
- *bit 4 (16): thinnest clouds*
- *bit 5 (32): cloud shadows cast by a detected cloud*
- *bit 6 (64): cloud shadows cast by a cloud outside image*
- *bit 7 (128): high clouds detected by 1.38 µm" (Donadieu and L'Helguen, 2016)*

The MAJA algorithm performs atmospheric correction and cloud detection in Sentinel-2 images using time series. It is based on a number of threshold tests including the cirrus band (B10), multi-temporal tests to detect clouds and the correlation of the pixel neighbourhood with previous images to avoid over detections (Donadieu and L'Helguen, 2016).

FMASK

One of the most popular approaches has been developed by Zhu and Woodcock, 2012 for Landsat 4-7. The so-called Fmask (Function of mask) approach uses Top of Atmosphere (TOA) reflectance and Brightness Temperature (BT) to detect clouds and cloud shadows. This approach has been improved and a new version has been provided for Landsat-8 and Sentinel-2 data (Zhu et al., 2015). The major input for cloud detection for Landsat data is a thermal band which does not exist for Sentinel-2. Nevertheless, the new cirrus band of Sentinel-2 has been found to be useful for cloud detection especially for thin cirrus clouds (Zhu et al., 2015).

3.2.2.2 Benchmarking Criteria

VISUAL ASSESSMENT

The assessment criterion is a visual interpretation of the resulting cloud masks. In the following, some image samples show the differences between the processors MAJA, FORCE and Sen2Cor as well as the

difference between the product levels available regarding Sentinel-2 data. The assessment of the cloud, cloud shadow and snow masks are done by visual interpretation of the different products.

QUANTITATIVE ASSESSMENT

A quantitative assessment was recently published by Baetens et al. (2019), who validate Copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and Fmask processors using reference cloud masks generated with a supervised active learning procedure. The main benchmarking criterion is the overall accuracy.

3.2.2.3 Implementation and Results of Benchmarking

Figure 3-8 shows the different cloud classes provided by the Level-1C and Level-2 masks. The mask quality is examined by visual inspection. Level-1C cloud mask comprises cirrus and opaque not covering the full cloud extent (omission error). Level-2 scene classification presents three different classes of opaque clouds based on probabilities (high, medium, or low). Misclassifications are observed in urban areas and dry mountainous terrain, concerning all three probability classes. The high probability cloud class is calculated with a high threshold from the cloud probability mask. It is missing some thin cirrus clouds but on the contrary, misclassifications are reduced. With higher thresholds, only the cloud centres are detected, missing cloudy pixels at the borders. The lower the probability cloud class calculated with a low threshold suffers from false detection for bright surfaces in dry environments in urban areas or on mountainous terrain. In addition, Level-1C and Level-2 products contribute thin cirrus cloud classes taking advantage of the new cirrus band B10. In the Level-1C product, the mask is already further processed to reduce misclassification and fill gaps (Müller-Wilm et al., 2013). Both masks are screened reasonably well, but some thin clouds (mostly thinner plane contrails) are missed by the detection in Sen2Cor. Although Sen2Cor cloud mask tends to over detect clouds, high, medium and low probability clouds and the cirrus are adopted and further processed, due to the fact the low probabilities cloud mask are best at detecting the actual clouds.

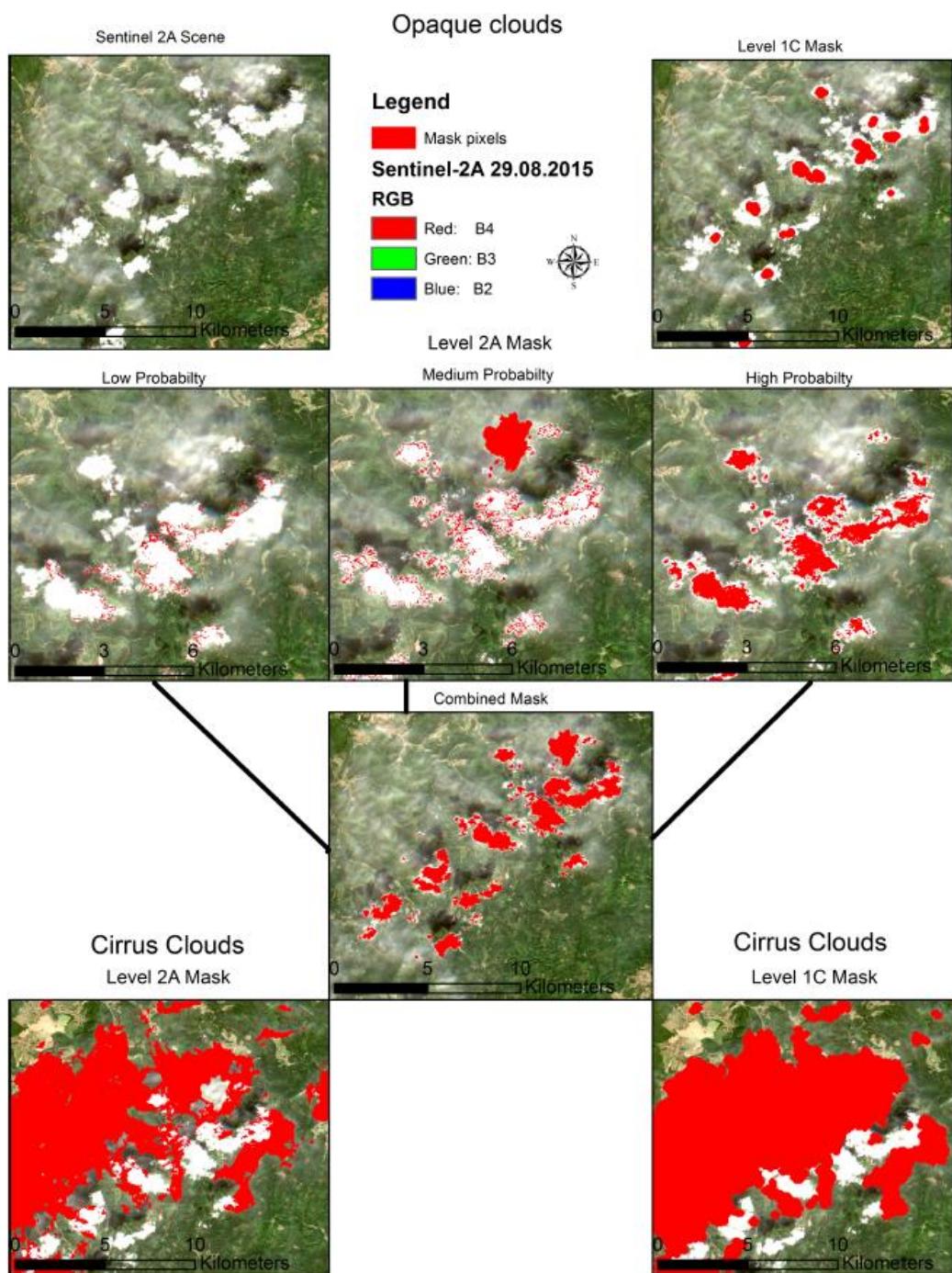


Figure 3-8: Sen2Cor cloud mask: Comparison between Level-1C and Level-2A product

VISUAL ASSESSMENT COMPARISON MAJA, SEN2COR AND FMASK

Sen2Cor cloud masks are based on single scenes, and provide 3 levels of cloud probability: low, medium, high, and also the cirrus mask (Louis et al., 2016b). In contrast to Sen2Cor, the MAJA algorithm uses a time-series based approach and compares pixels from previous images to assess the cloud probability (Lonjou et al., 2016).

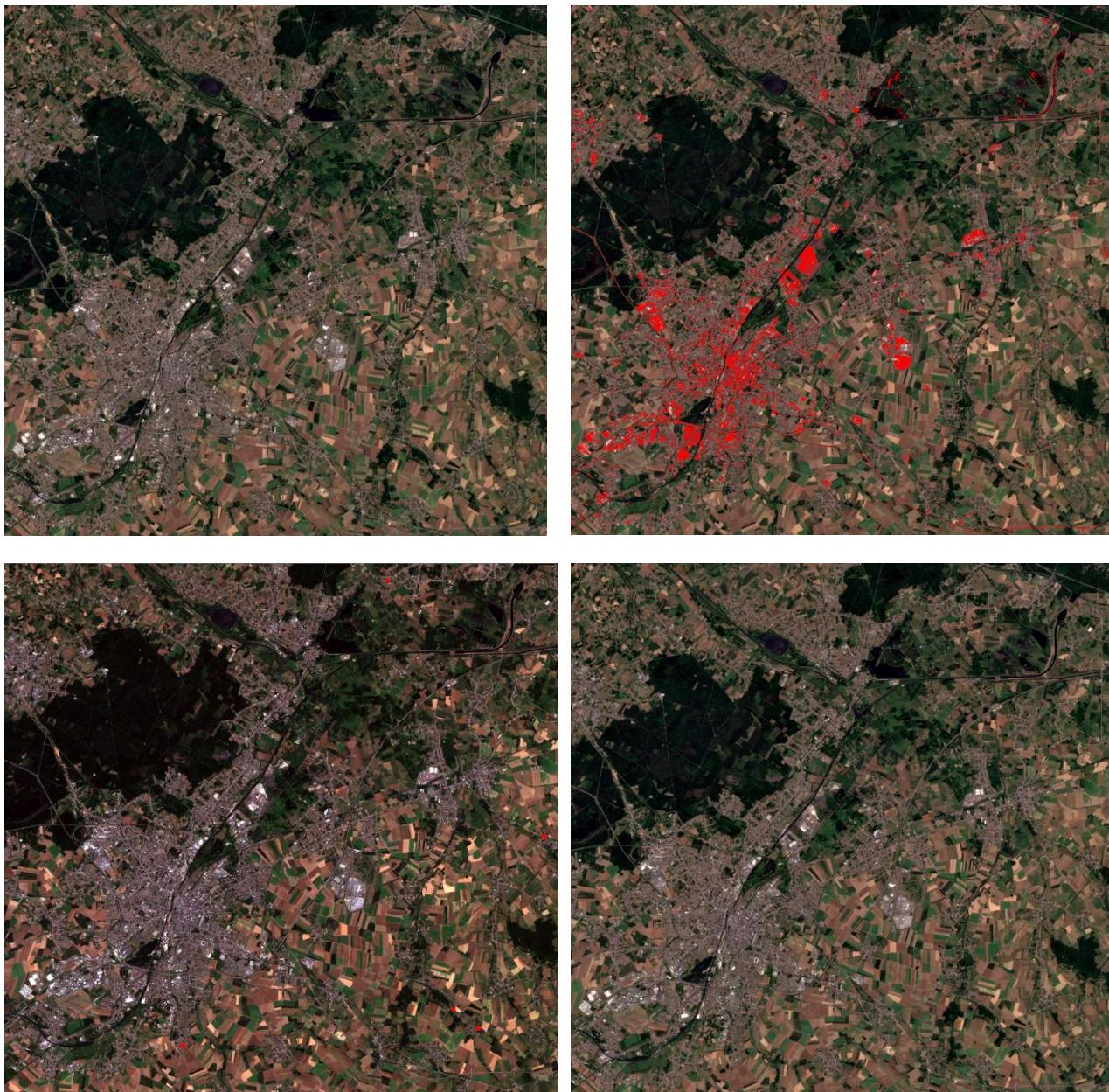


Figure 3-9: Differences between the Sentinel image Level-2A without cloud mask (upper left) and the Sen2Cor cloud mask (upper right), the Fmask cloud mask (lower right) and the MAJA cloud mask (lower left).

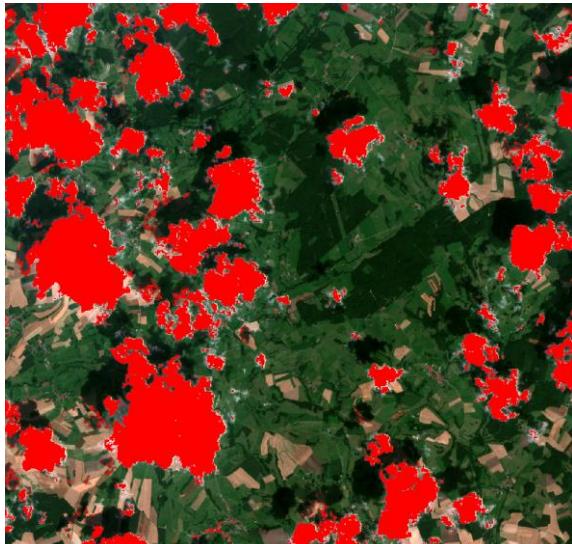
Notably, Sen2Cor tends to misclassify bright areas for clouds as shown in Figure 3-9. The example comes from a completely cloud-free image obtained above France at the beginning of May (Sentinel-2_20170705T105031_31UER). No commission errors can be seen within the MAJA cloud mask, while Sen2Cor classifies, urban areas, bare areas, agricultural areas and water bodies as clouds (low probability clouds). Fmask shows only a few commission errors. Therefore, it has been decided to use only the medium and high probability clouds from the scene classification within a cloud mask for further analysis and visual comparisons.



Level-2A image (Band combination:4-,B3-,B2)



MAJA cloud mask according to bit1 (= all clouds except the thinnest) (cloud mask in red).



Sen2Cor (Level-2A) scene classification (high, medium and low probability clouds)



Fmask cloud mask (cloud mask in red)

Figure 3-10: Example 1: Cloud mask comparison between MAJA, Fmask and Sen2Cor. Sentinel-2_20170725T105031_31UER; 49°41'8.69"N; 4°16'55.80"E



Level-2A image (Band combination:4-,B3-,B2)



MAJA cloud mask according to bit1 (= all clouds except the thinnest) (cloud mask in red).



Sen2Cor (Level-2A) scene classification (high, medium and low probability clouds)



Fmask cloud mask (cloud mask in red)

Figure 3-11: Example 2: Cloud masks comparison between MAJA, Fmask and Sen2Cor. Sentinel-2_20170705T105031_31UER; 49°41'8.69"N; 4°16'55.80"E

Figure 3-10 and Figure 3-11 show a visual comparison between the MAJA, Fmask and Sen2Cor Level-2 cloud mask outputs based on a true colour image (the cloud mask are represented in red). Both methods detect large clouds but show omission and commission errors. In particular, Sen2Cor shows large commission errors over bright land cover objects when including low probability clouds. On the other hand, a significant number of omission errors are introduced at the edges if low probability cloud pixels are excluded. On the contrary, MAJA shows omission errors because it tends to underestimate small clouds and cloud shadows. Furthermore, MAJA shows commission error over dark land cover objects (e.g. coniferous forest). Both, MAJA and Sen2Cor omit cirrus clouds. Concluding it can be stated that both cloud masking approaches show room for improvements. Fmask tends to detect the clouds quite well with few omission and commission errors.

QUANTITATIVE ASSESSMENT COMPARISON MAJA, SEN2COR AND FMASK

The quantitative assessment by Baetens et al. (2019), who validate Copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and Fmask processors using reference cloud masks generated with a supervised active learning procedure conclude: “The results show that MAJA and FMask perform similarly, with an overall accuracy around 90% (91% for MAJA, 90% for FMask), while Sen2Cor’s overall accuracy is 84%” (Baetens et al., 2019). These results correspond with above visual assessment and with the experience from the consortium partners of ECoLaSS which were gained in various other projects.

LEVEL-2A CLOUD SHADOW AND TERRAIN SHADOW MASKS:

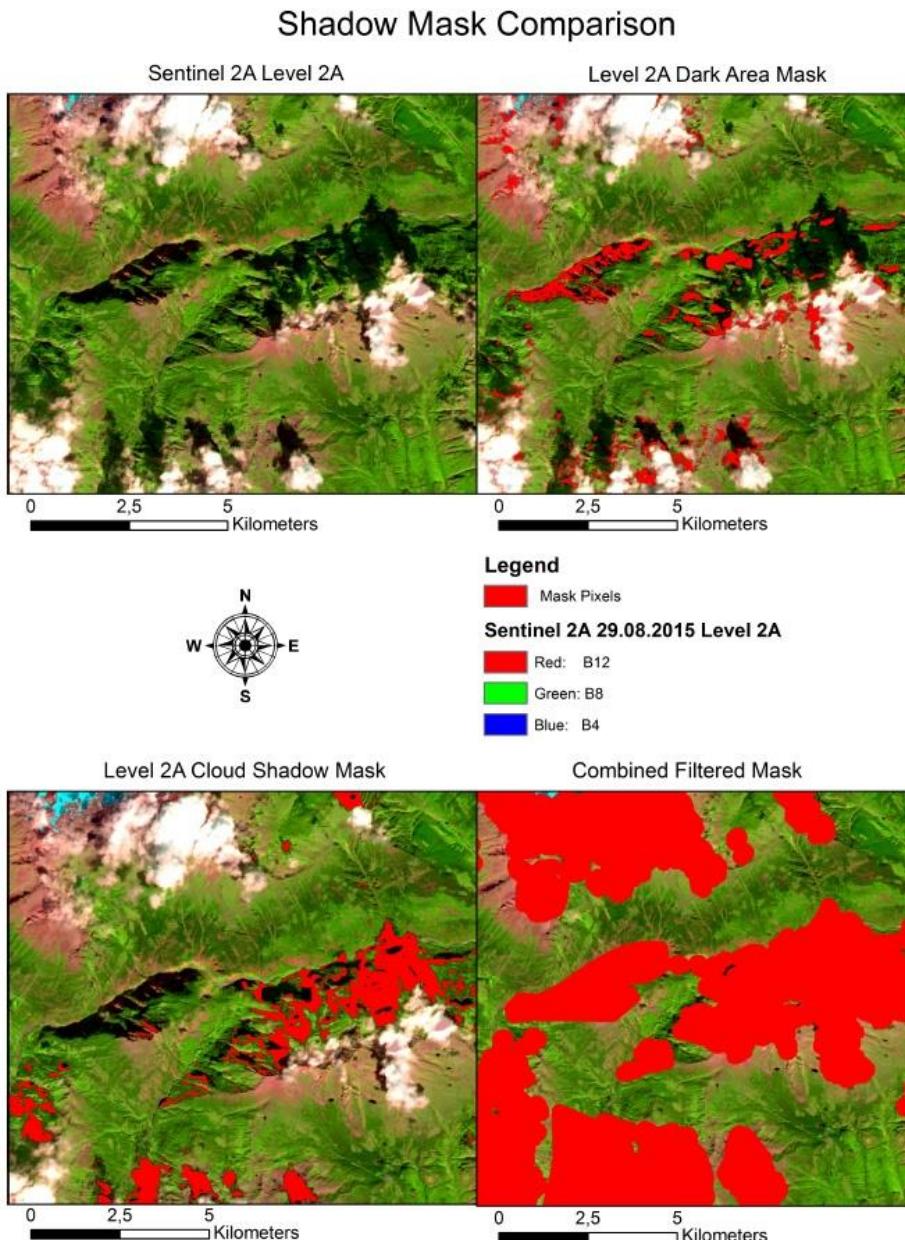


Figure 3-12: Shadow mask: Comparison between Level-2A terrain and cloud shadows.

The classes cloud and terrain shadows are only provided by the Level-2 product represented in Figure 3-12. Cloud shadows are often misclassified because they share similar spectral signatures with

topographic shadows, water and wetlands (Zhu et al., 2015). The cloud and mountainous shadows in the scene classification are classified too small; therefore a buffer of 200m is applied to enlarge the shadow areas. Some water pixels with low reflectance area are masked as cloud shadow pixels due to similar reflectance values in the NIR and SWIR bands. To reduce the misclassification in the shadow mask patches with a size less than 400m (corresponding to four 10 x 10m pixels) are eliminated automatically.

In most cases, cloud and cloud shadows are located next to each other depending on the sun location and image acquisition time and the shadow size should be nearly the same size as the cloud. To reduce the cloud shadow misclassification geometric relationships between clouds and cloud shadows could be used (Jin et al., 2013).



Level-2A image (Band combination:4-,B3-,B2)



MAJA shadow mask according to bit5 (=cloud shadows cast by a detected cloud) (shadow mask in red).



Sen2Cor (Level-2A) scene classification (cloud shadows in red)



Fmask shadow mask (shadow mask in red)

Figure 3-13: Shadow mask: Comparison between terrain and cloud shadows.

Figure 3-13 presents a comparison of the shadow detection between MAJA, Sen2Cor and Fmask. The resulting shadow masks are highlighted in red. Due to a generous buffer, MAJA produces a usable cloud-shadow mask, however, at the cost of overestimating shadow areas. Moreover, shadows between closely neighbouring clouds are often not detected as such. However, in most cases these areas fall into

the buffer around the cloud itself, and can therefore also be considered as masked. Cloud shadow detection in Sen2Cor performs relatively poor with lots of omissions, whereas Fmask performs the best by detecting cloud shadows very accurately, which in most cases, can be directly used further, even without additional buffering.

3.2.2.4 Summary and Conclusions

Sen2Cor presents generally reasonable detection of clouds and cloud shadows excluding the low probability cloud pixels. The result can be improved by applying an adequate buffer. Still, high confusion with bright surfaces is present. Additionally, some small clouds and parts of shadows are omitted. In general, cloud omissions should be avoided, meaning bright objects should not be confused with clouds and cloud shadows should not be confused with dark objects. To avoid commission errors, the multitemporal approach used with MAJA promises better results. Regarding the usability of the cloud mask products omission errors at the edges as shown in the Sen2Cor product are preferred over misclassifications due to an easier applicable cloud mask enhancement processing chain. For example, enhancements can be achieved by applying a buffer to the estimated cloud areas.

Sen2Cor performs reasonably with the given information, the software implementation and usage is user-friendly and open for public usage. Considering multi-temporal information, MAJA produces better cloud masks. Nevertheless, due to the licence and implementation issues MAJA fits better for a systematic production at the ground segments, whereas Sen2Cor is easier to implement and run operationally on the user side. Both methods still need to be improved. Sen2Cor still detects too many clouds and confuses snow and clouds as well as water and shadows. MAJA also needs some improvement due to the omission errors.

3.2.3 Benchmarking of candidate approaches for topographic normalisation

A topographic correction is necessary if the test sites are characterized by mountainous terrain. The topography can significantly influence the radiometric properties of the signal received from the satellite (Wulder and Franklin, 2012). This effect is caused by the different lighting angles resulting from the topography (Gallaun et al., 2007). The aim of a topographical correction is to compensate for the differences in reflectance intensity between the areas with varying slope, exposure and inclination and to obtain the radiation values that the sensor would have in the case of a flat surface. A large number of approaches have been developed to correct the topographic influence. In these algorithms, which are used for topographic normalization, the illumination conditions for each pixel are usually estimated by calculating the cosine of the sun incidence angle based on the solar position, slope and exposure of the pixel.

3.2.3.1 Description of Candidate Methods

Semi-empirical approaches are using Digital Elevation Models (DEMs) that provide terrain elevation, slope inclination, exposure and other information to describe surface geomorphology (Balthazar, Vanacker and Lambin, 2012). These include the Cosine correction, C-correction and Minnaert models (Balthazar, Vanacker and Lambin, 2012).

TOPOGRAPHIC CORRECTION WITHIN SEN2COR SOFTWARE

The algorithm for rugged terrain integrated within the Sen2Cor package requires the existence of an appropriate Digital Elevation Model (DEM). The DEM is used to derive products like slope, aspect and terrain shadow maps. Within the software package, only two pre-existing DEMs can be used. The first one is the 90m SRTM Digital Elevation Database from CGIAR-CSI. The second supported format is the commercial 90m DTED-1 Format PlanetDEM from PlanetObserver, which must be purchased (Louis et al., 2016b).

SELF-CALIBRATING TOPONORM FROM JR-IMPACT SOFTWARE

In this approach, the Minnaert correction is iteratively applied, where the Minnaert constant k is derived land-cover specific, by a regression (see formula: 5.5). Minnaert models derive such a Minnaert constant k for each spectral band. If it has a value of 1, this means that the surface is a perfect Lambert reflector (Kane et al., 2008; Hantson and Chuvieco, 2011; Vincini and Frazzi, 2003). In practice, the Minnaert constant is estimated by a linear equation (Hantson and Chuvieco, 2011). There are also some methods that have changed or extended the Minnaert Model (see also: Kane et al., 2008; Sola et al., 2016). The method implemented by JR extended this model by applying an iterative land-cover specific correction (Gallaun et al., 2007).

3.2.3.2 Benchmarking Criteria

The quality of a topographic correction depends heavily on the type of land surface. Therefore, the most suitable method can change depending on the surface type (Park et al., 2017). The accuracy and resolution of the elevation model used for normalization also play an important role. If the selected resolution is too coarse, artefacts may appear in the image after the correction (see Richter, Louis and Müller-Wilm, 2012). Sola et al., 2016 have each identified different evaluation methods for evaluating the quality of a topographical correction in several studies.

VISUAL ASSESSMENT

The first indicator of the quality of the corrections is the visual evaluation of the images before and after the topographic correction (Sola et al., 2016). In-situ measurements are also possible. These are measurements on site, for a few selected points, with an optical spectroradiometer (Singh et al., 2015). However, these are associated with high time expenditure and high costs. Moreover, the performance cannot be differentiated according to different spectral bands (Park et al., 2017).

LINEAR REGRESSION AND CORRELATION COEFFICIENT

The most commonly used quantitative evaluation method is to consider the dependence between the cosine of the solar angle of incidence and the reflection values of each band. This can be measured by forming the correlation coefficient of all pixels in the image and/or by looking at the slope of linear regression. A lower dependency between the angle of incidence and reflectance would mean that the correction was successful. However, this assumption is not valid for areas where slope and exposure influence the distribution of land cover. In such areas, a certain residual correlation between the angle of incidence and reflection must be expected (Sola et al., 2016). Instead, however, it is also possible to view the relationship between the cosine of the angle of incidence and the reflection values of each band only for the pixels of certain land cover classes and not for the pixels of the entire image. The extent to which the variability of the reflection values within a land cover class has decreased can be calculated using the standard deviation and the coefficient of variation of the reflection. After the topographic correction, a higher homogeneity of the land cover is to be expected, as the radiometric variations caused by the topography are minimized within a class by normalization. This evaluation method is one of the most objective evaluation methods.

3.2.3.3 Implementation and Results of Benchmarking

Two methods have been compared and analysed, i.e. self-calibrating topographic normalisation from JR-IMPACT software and topographic correction within Sen2Cor software. The analysis was based on reference land cover type data on tree species and different Digital Surface Models (DSMs) as input, i.e. SRTM 90m, ALOS DSM 30m and LiDAR 10m. First, the results are evaluated with visual interpretation and second statistical parameters are derived and analysed.

As input image, a Sentinel-2A atmospheric corrected image with Sen2Cor from 13.08.2015 Level-1C covering the granules 32TPT/32TQT was selected, which covers different elevation zones. The differentiation of the tree species types is analysed using separability distance measures, i.e. Jeffries Matusita (JM). Following combinations are investigated in order to assess the topographic methods and the used DEMs:

Table 3-3: “Overview of the used combinations.” (Posch, 2017)

Image	Atm. Correction	Topo. Correction	Used DSM
o.A.	No atm. correction	No topographic correction	-
o.T.	Sen2Cor	No topographic correction	-
Impact 0.9	Sen2Cor	Impact	10m LiDAR DSM
Sen2Cor	Sen2Cor	Sen2Cor	90m SRTM
Imp.SRTM	Sen2Cor	Impact	90m SRTM
Impact 1.0	Sen2Cor	Impact	10m LiDAR DEM
ALOS DSM	Sen2Cor	Impact	30m ALOS DSM

For the comparison of the various results, it was decided to compare cosine of the sun angle with the reflectance values of each band. The correlation coefficient is calculated and the slope of the linear regression line evaluated. It is assumed that the correction is successfully applied if the interdependence of the two variables is low. However, it has to be noticed that the two parameters slope and exposition could bias the results, showing a higher correlation than expected. For the assessment of the land cover variability, the statistical measures standard deviation and coefficient of variation are used. It is assumed that the reflectance will become more homogeneous after the normalisation due to the reduction of the radiometric variation.

Preconditions are to be set before the investigations have been performed. Not all forest samples from the reference can be used for the investigations due to following reasons. Forest sites with crown coverage of less than 60% show much influence from other surface classes, e.g. bushes, grass or bare soil, and thus excluded. Secondly the classes mixed forest, fir and pine are not used due to their high spectral variability which could bias the results in the calculation. Another factor influencing the reflectance values is the age of the trees. In order to avoid a mixture of the age classes, it is necessary to split them into two separate classes: old and young.

REFERENCE AREAS

For the analysis, the reference areas are sorted out according to the above-described requirements. Altogether 138 samples were used for the analysis. Table 3-4 illustrates the number of samples for each forest type and the corresponding number of pixels.

Table 3-4: “Overview of the used classes and number of samples/pixels.” (Posch, 2017)

Species Type	Number of Samples	Number of Pixels
Spruce old	34	1680
Spruce young	27	1476
Broadleaf old	32	1768
Broadleaf young	24	1101
Larch old	21	1033

VISUAL ASSESSMENT

In mountainous areas, the topography causes different illumination on slopes depending on their exposition. In general, reflectance values for land cover classes on north oriented slopes are lower than reflectance values on south oriented slopes. This rationale creates a 3-dimensional effect based on the dark and bright slopes, which gives a distinct spatial impression to the observer (see Figure 3-14: images o.A. and o.T. products are explained in Table 3-3: “Overview of the used combinations.”). Thus it is assumed that the loss of the 3D effect and an equalized illumination within a scene is an indicator for a successful topographic correction (see Figure 3-14).

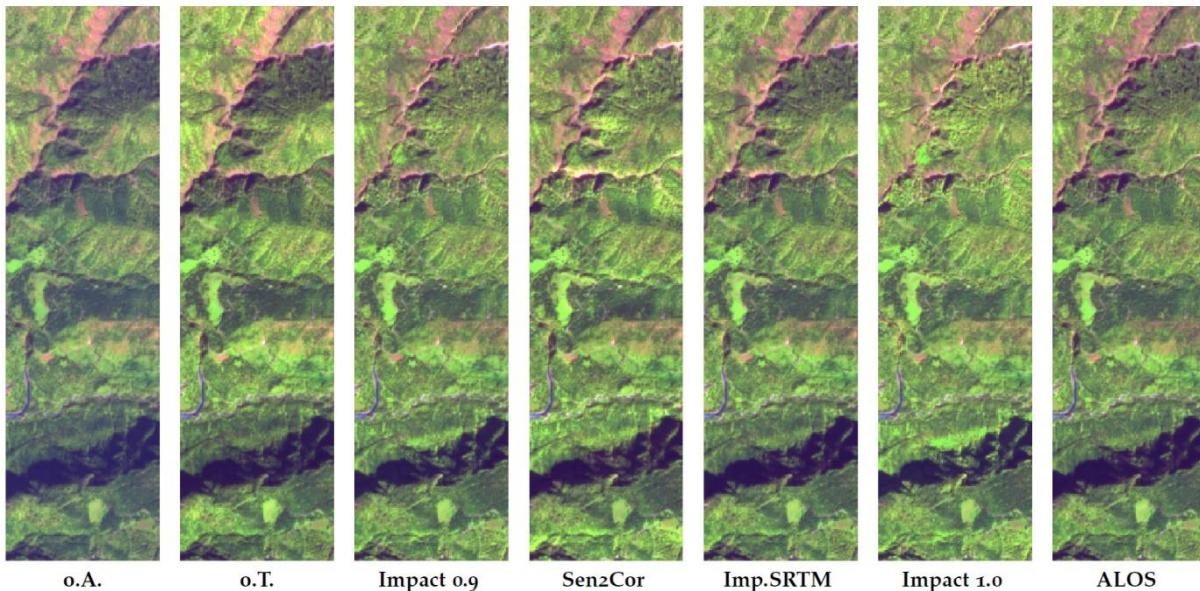


Figure 3-14: “Example showing the various images before and after topographic normalisation (see also Table 1).” (Posch, 2017)

The relationship between the cosine of the sun angle and the image reflectance values is significant before the topographic correction. In order to analyse this relationship, it is necessary to calculate the variables such as regression line, regression coefficient and correlation coefficient.

LINEAR REGRESSION

In a first approach, the linear regression is calculated and compared within the data sets, i.e. origin data sets without topographic correction versus the in Table 3-3 listed topographic corrected images. The rationale behind this approach is based on the assumption that after the topographic correction the slope of the regression line equals 0. The relief of the terrain should not influence anymore the reflectance values within the scene. This ideal case is not fully realistically to be achieved owing to natural variation properties of forest stands on one hand and the selection of reference areas on the other hand. That means in a certain class the reference areas to be used for the calculation could be more heterogeneous than in another class. It can be stated, that after a successful topographic correction the slope has to become flat. An example is displayed in Figure 3-15.

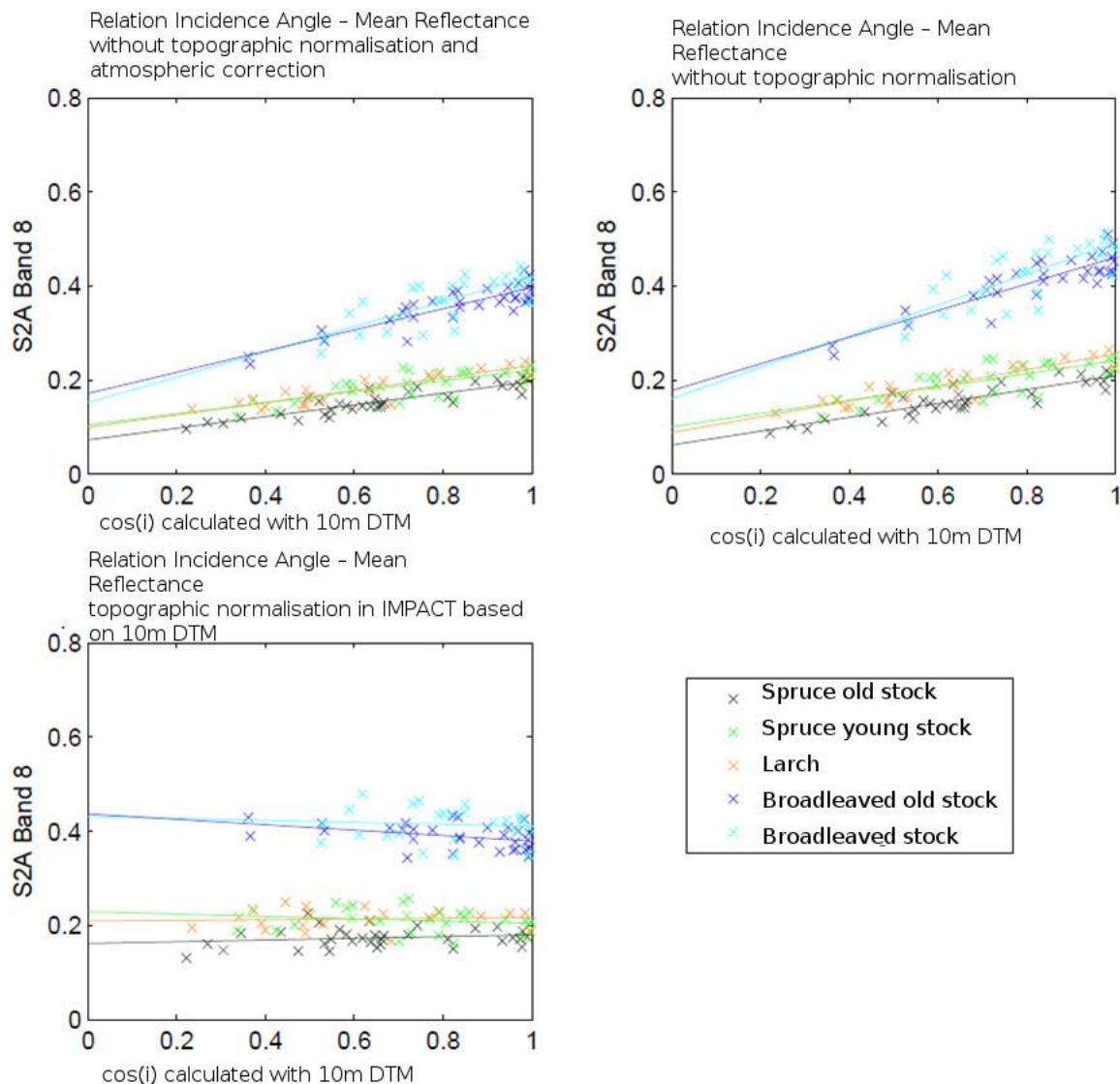


Figure 3-15: “Example showing the regression line before and after topographic normalisation (on top 2 before and below diagram after correction).” (Posch, 2017)

In Figure 3-15 the diagrams on the top show the regression lines representing the two input scenes, without any atmospheric or topographic correction (left) and with atmospheric correction (right). The diagram on the bottom displays the regression lines after the topographic correction using IMPACT with 10m DSM. It can be observed that the reflectance of broadleaf classes or the younger tree types is higher than the coniferous classes or older stands. Two effects can also be observed after the topographic correction. The classes can be over-corrected (negative line) or under-corrected (positive line) after the topographic normalisation. In Figure 3-16 a comparison of the five image derivatives can be seen.

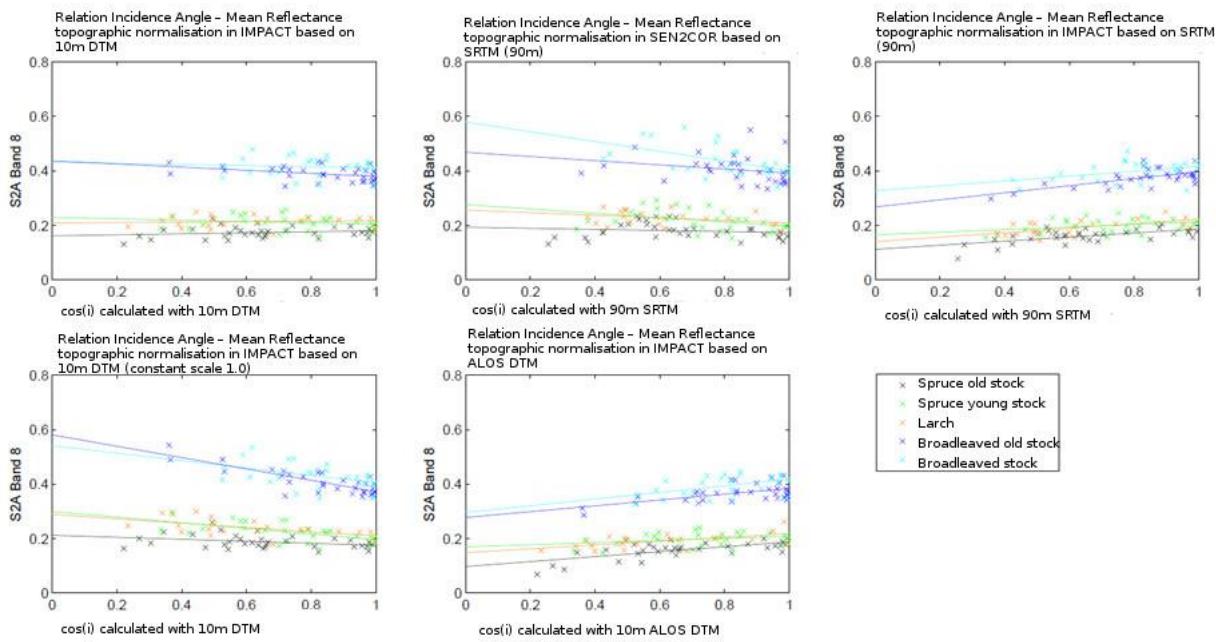


Figure 3-16: “Regression lines after topographic normalisation using different software packages and DEMs.”
(Posch, 2017)

In Figure 3-17 the slopes (β_{1}) of the regression line are illustrated with bar diagrams. It can be observed that for the Impact0.9 results the slopes in all bands are low. The Sen2Cor image shows higher values for the young stands and the slope is tending towards negative values. For the SRTM and ALOS images similar results with higher values can be observed and show an over-correction. This fact indicates a strong influence of the DSM. Generally, the broadleaf classes show higher negative slopes than the other classes.

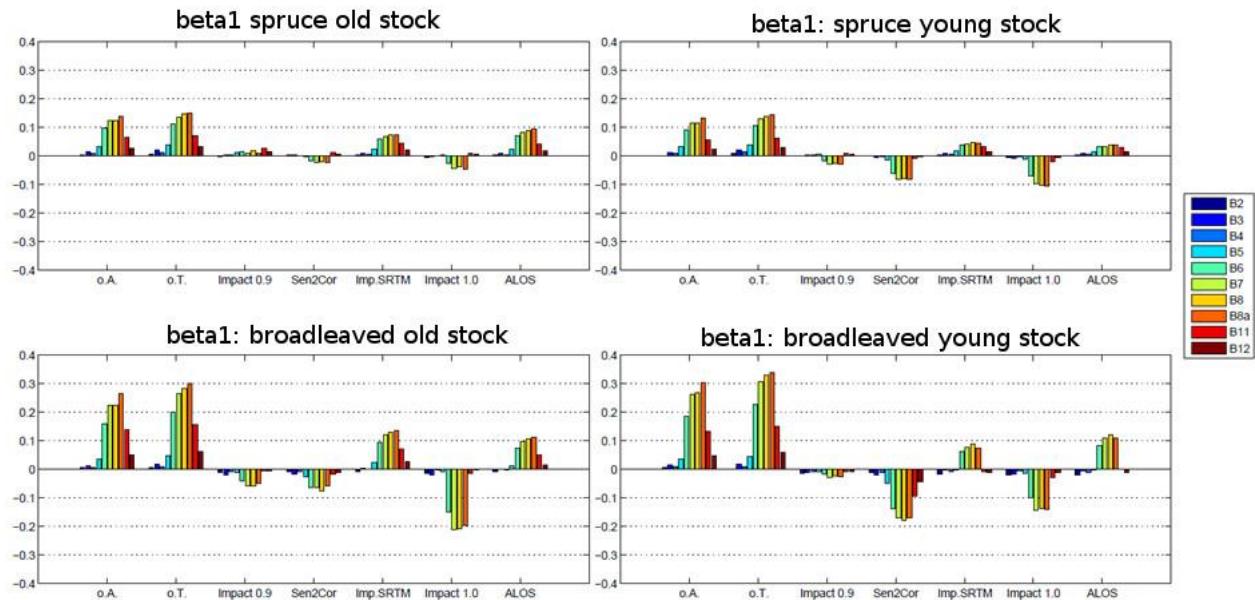


Figure 3-17: “Bar Diagrams showing the slope (Beta1) of the regression line.”
(Posch, 2017)

Another statistical value for the assessment of the analysis is the p-value which is to be used to confirm the null hypothesis or reject it. The null hypothesis is defined as the correlation between $\cos(i)$ and *reflectance values*. That means, after a successful topographic correction the null hypothesis should be rejected (p-value advances to 0, the probability of error = 5%), meaning that no topographic effect from the terrain can be observed. For the Impact0.9 image, the null hypothesis can be confirmed. For the Sen2Cor image, the null hypothesis is confirmed except for the class Broadleaf young. For this statistical parameter, the topographic correction for both images can be considered successful. For the images Imp.SRTM, ALOS and Impact1.0 the null hypothesis is rejected, meaning that the topographic correction cannot be considered as successful.

However for this investigation is also important to analyse the results using the correlation coefficient (between 0 and 1) of $\cos(i)$ and *reflectance values*. Looking at the two input images in Figure 3-18 it can be seen that both show a high correlation coefficient, which is expected because the topography is not corrected. Also for the images Imp.SRTM, ALOS and Impact1 the correlation coefficient is rather high. For the two other images Impact0.9 and Sen2Cor, the correlation coefficient is rather low, confirming again a successful correction.

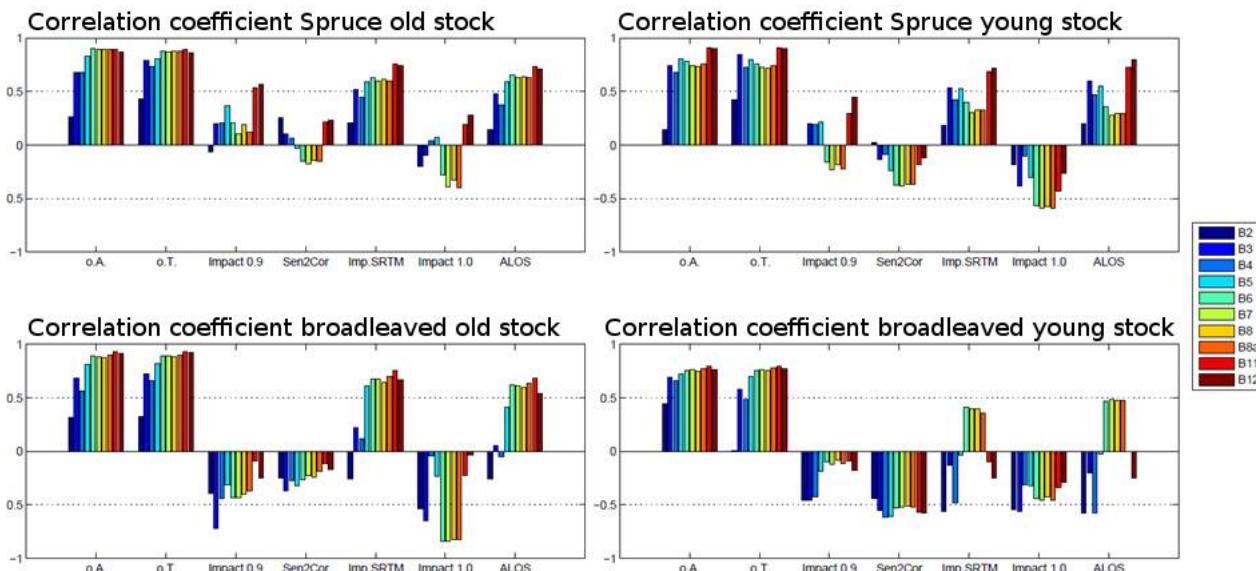


Figure 3-18: “Bar Diagrams showing the correlation between $\cos(i)$ and *reflectance*.” (Posch, 2017)

In a later stage of this investigation an additional benchmarking was applied using the 5m DEM from Euromaps for the topographic normalisation. The outcome was then compared by visual interpretation with the topographic normalized image using the LiDAR DSM.

Table 3-5: Overview of the used combinations.

Image	Atm. Correction	Topo. Correction	Used DSM
o.T.	Sen2Cor	No topographic correction	-
Impact 0.9	Sen2Cor	Impact	1m LiDAR DSM
Euromaps DSM	Sen2Cor	Impact	5m Euromaps DSM

In the analysis of the various DSMs (see also Table 3-3) it has to be stated that the spatial resolution of the global available DSMs was limited to a best resolution of 30m. The results above have shown that a highly accurate DSM from LiDAR with 1m spatial resolution gives the most sufficient correction results. Additionally, the new Euromaps DSM with 5m spatial resolution offers an alternative to the rather expensive product of LiDAR DSM.

For this purpose, a topographic correction was applied with the Euromaps DSM using the same parameters as for the correction with the LiDAR DSM.

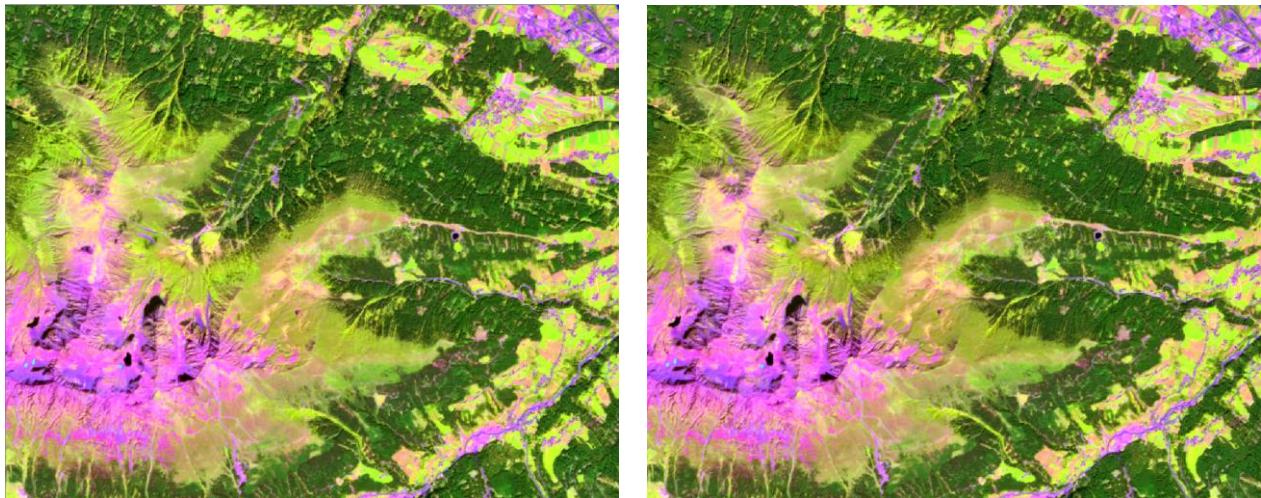


Figure 3-19: Image examples showing the results of the topographic normalisation with LiDAR DSM (left) and Euromaps DSM (right).

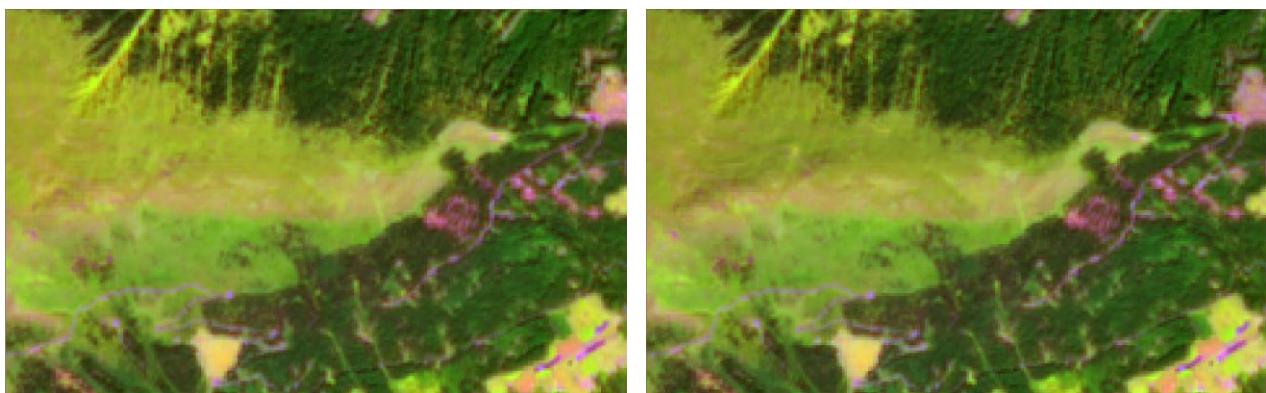


Figure 3-20: Image examples showing details from the results of the topographic normalisation with LiDAR DSM (left) and Euromaps DSM (right).

A visual interpretation on the results (see Figure 3-19 and Figure 3-20) show that the use of a Euromaps DSM gives good correction results. In comparison with the LiDAR DSM only minor differences can be recognized, which occur only in extreme terrain locations, e.g. ridges or deep valleys. The overall picture for the Euromaps DSM image shows a harmonized and equal correction, and especially the 3D effect disappeared. The latter is the strongest indicator for a successful correction or not.

Thus, it can be stated that the Euromaps DSM is a good alternative for a proper topographic correction, and can highly be recommended in case of availability.

3.2.3.4 Summary and Conclusions

The comparison of the topographical correction of the Sen2Cor processor with the one of the IMPACT software (each using the 30m SRTM) shows different results depending on evaluation method. The visual comparison shows that after the normalisation in Sen2Cor over-corrected images are obtained, while after the normalisation in Impact a slightly under-corrected picture is obtained. The same result is reached by considering the correlation between the cosine of the incidence angle and the reflectance. The decrease of variation within a class is higher after the normalisation in IMPACT than after the correction in Sen2Cor.

Furthermore, corrections with elevation models of different spatial resolutions have been performed. With the 1m LiDAR elevation model, an improvement of the correction because of the higher resolution has been reached compared to the one performed with the 90m SRTM. No improvement can be detected if the 30m ALOS DSM is applied instead of the 90m SRTM. The evaluation results with the 30m ALOS and the 90m SRTM elevation model are very similar. Looking at the two corrections performed in IMPACT, each with a different scaling factor of the Minnaert constant it can be seen that setting the scaling factor to "1" leads to an over-correction. With setting the factor to "0.9" better results are achieved.

One limitation is the quality of the DEM, the current Level-1C DEM (SRTM 90m). The main issue with Sen2Cor implementation of the topographic normalisation is the over-correction of south facing slopes. From the analysis, we conclude that topographic normalization should be applied in hilly to mountainous terrain for the ECoLaSS demonstration site production. The DEM with a spatial resolution in the range of the imagery which is corrected should be used. Therefore, we would recommend using topographic normalisation tools which allow the user to define their own DEM, if only the SRTM 90m model is available Sen2Cor topographic normalisation also provides acceptable results. In general, DEMs with higher resolution are already available freely, such as recent SRTM model or JAXA's global ALOS 3D world model, both have a resolution of 1 arc second. The use of the Euromaps DSM is an alternative for topographic correction as the outcome shows similar correction results as with the other higher resolution DSM.

3.2.4 Benchmarking of candidate approaches for Geometric consistency

Pre-processing of EO data used in time series analysis applications is one of the most crucial issues to guarantee reliable and meaningful results, especially so, when data from different sensors are used. In ECoLaSS the imperviousness, grassland, forest and agriculture applications may use various combinations of Sentinel-2, Sentinel-1 and Sentinel-3 data. Therefore, it is essential to analyse the overall data quality of the input data and the initial geometric shifts and radiometric differences between image data from the same sensor (e.g. from neighbouring orbits) and from different sensors. The results from this analysis form the basis for the pre-processing improvements applied in the ECoLaSS pre-processing workflows.

3.2.4.1 Description of Candidate Methods

SENTINEL-2 GEOMETRIC ACCURACY TESTS

Currently, Sentinel-2 data is delivered as orthorectified images only. Orthorectification at ESA is performed based on a medium resolution PlanetDEM from PlanetObserver. In mountainous areas, geometric accuracy strongly depends on the quality of the DEM and the satellite viewing angle. Since the accuracy of ESA's DEM at the ECoLaSS test sites is not known to the consortium, some geometric accuracy tests were performed in order to quantify the DEM related location errors in alpine settings. In addition to DEM related errors, we also expected to find linear location errors between Sentinel-2 images for adjacent orbits. In addition to DEM related errors, a slight yaw drift was detected for Sentinel-2A and corrected end of May 2016. Images processed before June 2016, therefore, have slightly higher

geometric mislocation errors. A list of known quality issues with Sentinel-2 data is published in the Sentinel-2 data quality report (<https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/data-quality-report>). As of summer 2017, when the first pre-processing chains were defined in the ECoLaSS project, Sentinel-2 data from neighbouring orbits were not co-registered by tie-points. In near future, images will be matched to a Sentinel-2 Global Reference image (Dechoz et al., 2015). This will significantly reduce the geometric location errors between images of overlapping orbits.

3.2.4.2 Benchmarking Criteria

SENTINEL-2 LOCATION ERRORS

Until the Sentinel-2 global reference image is available, it could be necessary for many change detection algorithms to register all Sentinel-2 images to a master reference scene in order to guarantee sub-pixel positioning accuracies. This is especially true when data from different relative orbits is used.

In order to estimate the effect of Sentinel-2 geometry errors on ECoLaSS products and services, we investigated the location errors in flat and mountainous terrain at one of the ECoLaSS test sites.

TEST SITE SELECTION FOR GEOMETRIC ANALYSIS

The Central test site, covering parts of Austria and Germany, was chosen to analyse the geometric quality of the input EO data. It is characterized by both topographically flat and mountainous terrain and it covers all land cover classes to be mapped in ECoLaSS – i.e. imperviousness, grassland, forest, agriculture. A mountainous test was seen as being important for the accuracy analysis since Sentinel-2 data is delivered as orthorectified image data and non-linear geolocation errors related to the underlying DEM used by ESA during orthorectification also need to be assessed.

3.2.4.3 Implementation and Results of Benchmarking

GEOMETRIC QUALITY TESTS

In order to separate the different error components (DEM related, orbit related – non-linear, linear), geometric tests were performed at the Central test site. In a first analysis, 9 Sentinel-2 granules with mostly flat terrain were matched to a Sentinel-2 master scene using a linear matching approach based on an automated point matching using areal mutual information as for cost function. The current relative offsets at the test site are on the order of 0.1-1.3 pixels, with a mean offset of 1.8m in x and 8.9m in y-direction between all Sentinel-2 images illustrated in Figure 3-21. However, maximum mean shifts between overlapping relative orbits are up to 22m in y-direction.

name	orbit	X [m]	Y [m]	#points
2015-07-31-S2	1	-0.384	9.474	844
2015-08-07-S2	2	0.087	8.087	1100
2016-02-06-S2	1	2.666	8.728	1087
2016-03-14-S2	2	-1.430	-7.938	1107
2016-03-17-S2	1	4.478	8.674	102
2016-03-27-S2	1	3.713	12.991	1026
2016-04-13-S2	2	-2.060	-9.667	1134
2016-06-22-S2	2	-2.071	5.239	987
Ref 2016-06-25-S2	1	-0.067	0.180	1134
Abs. mean		1.773	8.875	1049
Max. mean shifts		5.784	22.465	
std		2.474	8.040	93

Figure 3-21: Geometric accuracy analysis for Sentinel-2 data at a flat terrain test site.

These values are similar to those expected and reported by ESA (Languille et al., 2015); “1st Sentinel-2 Validation Team Meeting” at ESA in Nov. 2016) and to findings of other studies. A recent study has analysed the absolute location error of Sentinel-2 in Austria by matching Sentinel-2 and the Geoland Basemap Orthophoto of Austria (Pandžić et al., 2016). This reference imagery was acquired with 20cm resolution and orthorectified with a 10m LiDAR DSM. The authors report a mean shift of 0.6 pixels measured by estimating the offset of line segments at a test site in Austria. At a test site in Serbia a larger mean registration error of 1.27 pixels was found (Pandžić et al., 2016).

In a second analysis, 19 Sentinel-2 granules with mostly mountainous terrain were matched to a Sentinel-2 master scene using again a linear matching approach based on an automated point matching using areal mutual information as for cost function. Results are presented in Figure 3-22. While the mean errors of 4m in x and 5.2m in y are similar to those found in topographically flat areas, maximum local non-linear shifts of more than 30m were observed in the mountainous areas. This is similar to mislocation errors described for alpine areas of Switzerland and Norway (Käab et al., 2016). The authors report significant DEM errors that propagate into locally varying lateral offsets in the images of up to several pixels with respect to other georeferenced data or between Sentinel-2 data from different orbits. Such a shift of up to 30m or 3 pixels makes pixel-based change detection techniques very difficult in mountainous environments. The main reason is assumed to be DEM errors in the PlanetDEM used by ESA during orthorectification. Ressl and Pfeiffer have found that the PlanetDEM has an irregular tile structure with systematic height errors in the Alps of up to 50m (1st Sentinel-2 Validation Team Meeting”: [last accessed Sep 28, 2017]). The modelled geometric accuracy of Sentinel-2 data is still reported to be very high, with 95% of all errors between -3.5 and +3.5m, but significantly larger errors are expected for isolated mountain regions.

name	orbit	X [m]	Y [m]	#points
20160304	2	-3.801	9.443	745
20160307	1	-6.765	3.511	1564
20160308	2	-3.801	9.443	745
20160317	2	1.498	2.552	832
20160811	1	-6.947	8.304	1274
20160912	1	-2.521	9.090	958
20160215	2	1.670	10.882	780
20151218	1	-0.386	-2.998	1259
20151221	2	1.421	11.573	619
reference 20151228	1	0.002	0.001	1616
20160320	2	2.822	10.078	808
20160804	1	-3.265	2.269	533
20160824	1	-4.935	-2.277	809
20160827	2	-6.580	3.625	943
20160903	1	-5.204	0.212	744
20160923	1	-7.039	0.761	1255
20161013	1	-5.829	3.445	1243
20161016	2	-5.121	3.257	321
20161023	1	-7.803	6.359	1229
20161115	2	-0.509	4.327	850
Abs. mean		3.935	5.249	956.350
Max mean shifts		10.625	14.571	
std		3.416	4.454	338.377

Figure 3-22: Geometric accuracy analysis for Sentinel-2 data at a mountainous test site.

It is expected that in 2019, ESA will release its Sentinel-2 global reference image (Dechoz et al., 2015) which in principle is similar to the Landsat GLC and that all new Sentinel-2 will then be co-registered to this reference image. Most linear shifts between individual Sentinel-2 images should then be eliminated.

For analysing the non-linear errors that result from DEM errors during the orthorectification process we applied a full disparity matching at an alpine test site based on the same 20 Sentinel-2 scenes used for the linear matching tests. The full disparity matching uses a cross-correlation as cost function in image matching to derived image disparities. These disparities are then used for non-linear co-registration. Figure 3-23 shows two examples of the disparity maps. We found that in mountainous terrain, the linear

component alone is not sufficient to correct the observed mislocations as both the linear and the non-linear component have approximately the same magnitude. A linear workflow will therefore not be able to account for the non-linear DEM related errors.

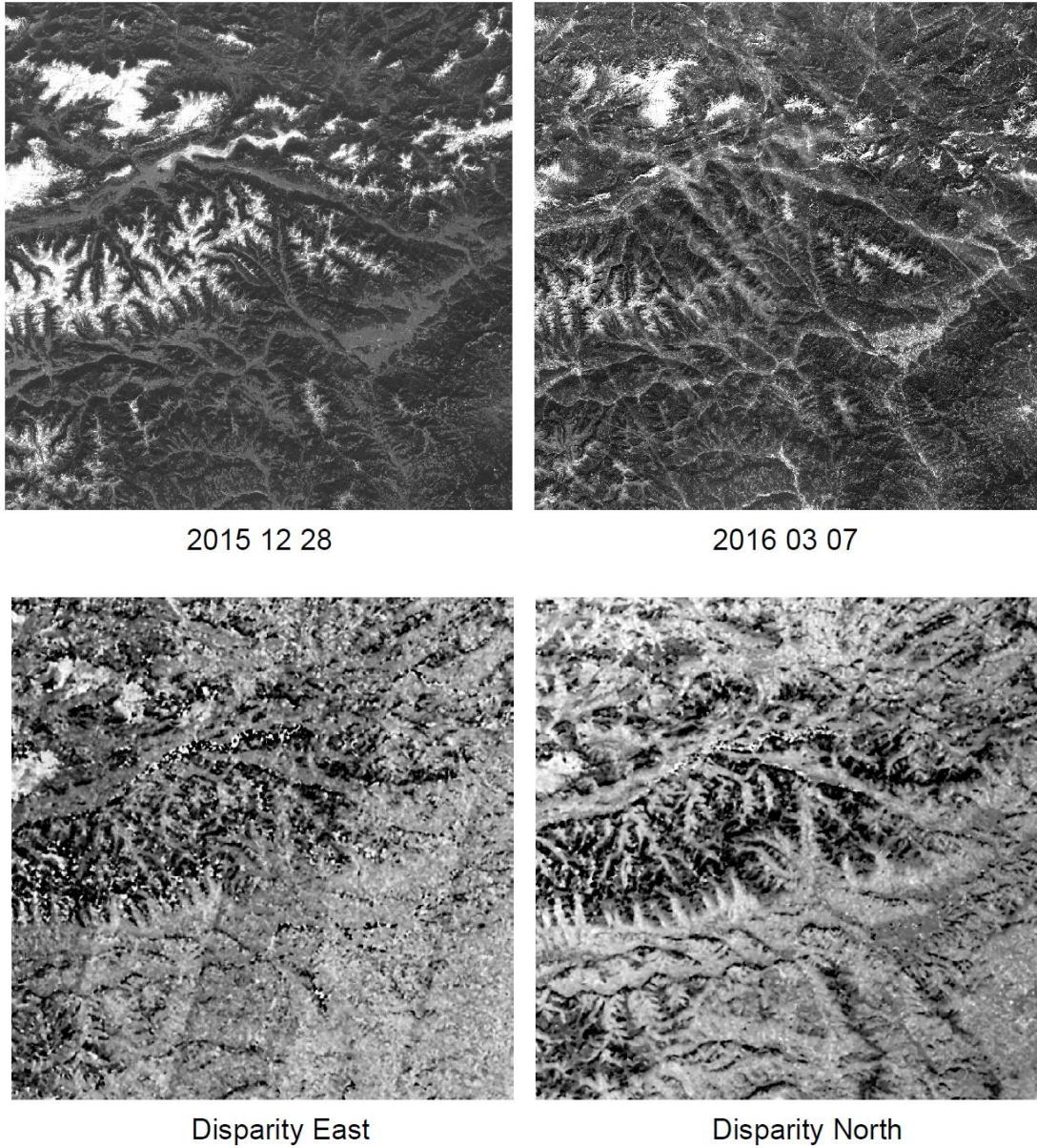


Figure 3-23: Two Sentinel-2 images from the mountainous test area and disparity images in East and North between both images.

3.2.4.4 Summary and Conclusions

While full disparity matching algorithms can account for the non-linear shifts, they consume a lot of computational resources and processing time. In a pan-European approach, this would be a strong limiting factor for any update services. The final co-registered images will also show radiometric distortions in areas of large geometry shifts which can make a classification or change detection impossible in these adjusted areas. Hence, the overall benefit of geometrically adjusting Sentinel-2 imagery in mountain areas can be questioned. The best solution would be, to use a better DEM during orthorectification at ESA. Since non-orthorectified Level-1B data is not made available to the scientific

community by ESA, the EClaSS consortium has to work with the available orthorectified Level-1C datasets. With a growing number of DEM related error reports, there are plans at ESA for integrating better DEMs during the orthorectification process, with Norway being a first test region. During the test phase used for algorithm development, the observed linear shifts are not so critical to require a co-registration procedure.

3.2.5 Benchmarking of candidate approaches for noise reduction and gap filling

This chapter investigates the applicability of the harmonic regression method for gap filling in forest areas. The underlying assumption of the approach is that the temporal trajectory of a given spectral band over the course of the year can be captured by a model featuring a sum of trigonometric functions of different frequencies and a trend component. Therefore the bottom-of-atmosphere (BOA) reflectance is modelled independently for each pixel as a continuous function of time. Consequently, synthetic images can be produced for any point in time in order to fill gaps in the time series. The concept of harmonic regression has been successfully applied by Verbesselt et al. (2012), Zhu et al. (2014), or Brooks et al. (2014).

3.2.5.1 Description of Candidate Methods

Regression models together with robust least-squares techniques to estimate their parameters are widely used. A general linear regression model links an m -element vector of observations \mathbf{z} to a p -element parameter vector \mathbf{x} through the measurement equation

$$\mathbf{z} = \mathbf{Ax} + \mathbf{r}$$

where \mathbf{A} is the $m \times p$ design matrix of the model and \mathbf{r} is an m -element vector of normally distributed observation noise with mean zero and covariance matrix \mathbf{R} . The observation noise represents the differences between the data and the model and therefore \mathbf{r} is also known as vector of residuals. Assuming $m > p$ and normally distributed and uncorrelated observation noise with mean zero and constant variance σ^2 , the parameter vector can be estimated using the method of ordinary least squares (OLS). Considering the presence of outliers in the data, for example un-masked clouds, this assumption is 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

$$F = (\mathbf{z} - \mathbf{Ax})^T \mathbf{W} (\mathbf{z} - \mathbf{Ax}) \rightarrow \min.$$

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 estimate minimizing F 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 recalculated 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 :

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

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{x} between consecutive iterations is below an appropriate threshold. The respective elements of w are determined by one of the weighting functions given below, where c is a tuning parameter. An illustration of the cost functions is given in Figure 3-24. The Huber cost function is more conservative regarding the down-weighting of suspicious observations.

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

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

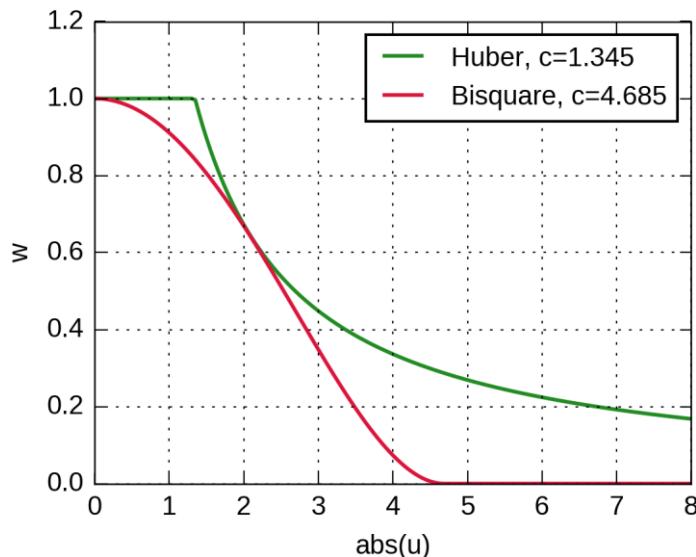


Figure 3-24: Weighting functions used by the robust IRLS fitting method.

The harmonic regression function of order n modelling the reflectance values z as a function of time t is

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

where c_0 , α_j , and β_j represent the parameters to be estimated and the ω_j 's are taken from the set of applicable angular frequencies given as

$$\Omega = \{\omega_1, \omega_2, \omega_3, \dots\} = \left\{ \frac{2\pi j}{365.25} \mid j \in \mathbb{N} \right\}$$

if time is measured in days.

3.2.5.2 Benchmarking Criteria

Within the benchmarking, the spectral differences between a synthetic image generated for a given target date and actual observations made at the same date are investigated with respect to different

model orders up to $n=3$. Spectral differences are calculated for 5000 sampling points randomly distributed in forest areas, see Figure 3-25. The square root of the mean of the squared differences represents a scalar metric used for the comparisons.

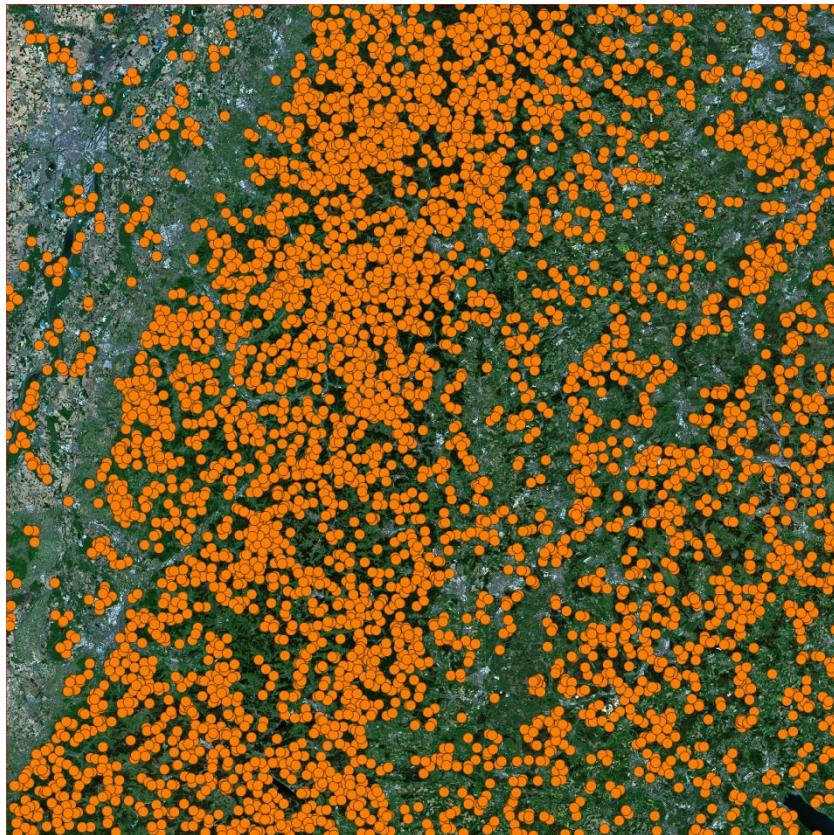


Figure 3-25: Distribution of sampling points over tile 32UMU.

3.2.5.3 Implementation and Results of Benchmarking

The benchmarking has been implemented in the central demo site, specifically for tile 32UMU. A stack of 108 Sentinel-2 images covering the time from 2016-01-26 to 2018-12-26 has been used as input data set. Two restrictions for image selection have been applied:

- Only images of relative orbit number 108 (full coverage of the tile)
- Only images with a nominal cloud cover of less than 85% according to the metadata information of the Level-1C images

After conversion of the Level-1C data to BOA reflectance using Sen2Cor, snow, clouds, and cloud shadows have been masked out based on the results of the Fmask algorithm (Zhu et al., 2015). For each sampling point and band, the harmonic regression is computed using all available data acquired in the years 2016 and 2017. Then, for each available image of 2018, reflectance values are predicted for the respective acquisition date based on the fitted model and differences can be calculated. This procedure ensures that the predicted values are independent from the reference values which they are compared to.

Please note that there are two restrictions which limit the number of sampling points available for evaluation for each 2018 image. On the one hand, only sampling points outside of masked areas can be taken into account. On the other hand the harmonic regression is only computed if a certain minimum number of observations are available in the training period 2016-2017. Since the highest model order within the benchmarking is 3, there are at most 7 parameters to estimate. For the robust regression to

yield reliable results, tests have shown that the degree of over-determination of the system should be at least 3, which means that at least 21 observations should be available to estimate the 7 parameters of the 3rd-order harmonic model. Hence only sampling points where this condition is met have been taken into account. Figure 3-26 illustrates the number of available un-masked observations per pixel within the training period.

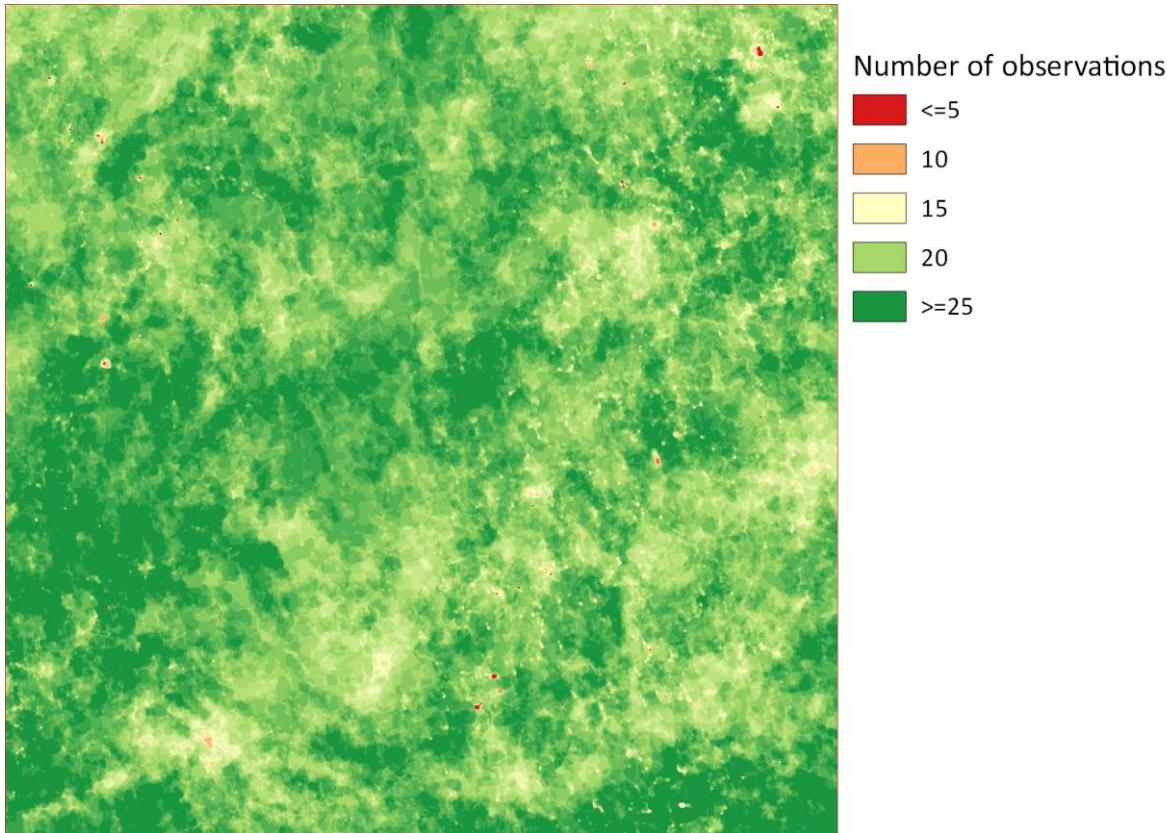


Figure 3-26: Number of available observations per pixel from 2016-01-26 to 2017-12-26.

The results of the benchmarking are presented in Table 3-6 in the form of root-mean-squared errors (RMSE) computed from the differences between model forecasts and actual images acquired in 2018. Depending on the cloud cover of each image, the number of applicable sampling points varies and results are only listed in the table if at least 1000 points were available. The table presents a quite detailed account of the forecast quality for the Sentinel-2 bands B04 (Red), B8A (NIR), and B11 (SWIR1) over the year.

Looking at the results, the first and probably surprising observation is that the RMSE values are quite similar regardless of the model order and that there is no general tendency showing that higher-order models achieve a lower RMSE. The table also shows that the highest RMSE values occur during the vegetation green-up in April and May, where also the phenological dynamics are highest. Both the steep increase of NIR reflectance as well shifts in the timing of the green-up are problematic and forecasts based on regression models can be subject to large errors. The other seasons beside spring are represented more accurately.

Table 3-6: Results of the harmonic regression benchmarking. N = Number.

Date 2018	N. of valid sampling points	RMSE over sampling points [% BOA reflectance]								
		Red band			NIR band			SWIR1 band		
		Model order			Model order			Model order		
		1	2	3	1	2	3	1	2	3
02.14.	1054	0.80	0.78	0.85	2.62	2.73	2.59	1.51	1.52	1.65
04.20.	3696	0.74	0.72	0.75	2.65	2.75	3.42	1.62	1.61	1.60
04.25.	1109	1.04	1.09	1.16	5.82	6.61	7.57	1.74	1.91	2.00
04.30.	1490	1.02	1.17	1.14	6.63	7.35	8.10	2.30	2.51	2.45
05.05	2370	0.98	1.06	1.08	7.07	7.47	7.89	1.70	1.83	1.88
05.25.	1255	0.89	0.95	0.89	6.67	6.23	5.60	1.76	1.79	1.79
05.30.	1980	0.80	0.81	0.72	5.83	5.23	4.36	1.83	1.72	1.83
06.04.	2111	0.82	0.85	0.83	5.25	4.55	3.64	1.64	1.58	1.58
07.09.	3085	0.52	0.46	0.49	2.56	2.57	2.98	1.40	1.26	1.23
08.03.	2029	0.99	1.12	1.01	2.40	2.60	2.36	1.37	1.37	1.23
08.23.	1102	1.39	1.41	1.50	3.51	3.34	3.37	1.76	1.79	1.88
09.12.	3913	0.36	0.39	0.38	2.31	2.05	1.99	1.10	1.11	1.24
09.17.	3913	0.47	0.56	0.48	2.95	2.72	2.75	1.38	1.41	1.47
09.27.	3914	0.62	0.77	0.72	2.58	2.71	2.48	1.18	1.18	1.21
10.12.	3900	0.54	0.50	0.55	2.57	2.75	2.40	1.36	1.22	1.13
10.17.	2702	0.94	0.73	0.67	2.81	2.78	2.68	1.97	1.77	1.64
10.22.	2973	0.88	0.74	0.74	3.44	3.33	3.23	2.16	1.99	1.78
11.16.	1494	1.07	1.01	0.92	2.53	2.47	2.52	1.93	1.97	1.87
12.26	2449	0.68	0.50	0.50	2.72	2.34	2.18	1.87	1.64	1.48

3.2.5.4 Summary and Conclusions

This chapter investigated the applicability of the harmonic regression method for gap filling in forest areas and different orders of the harmonic regression function have been benchmarked. A harmonic model has been trained using Sentinel-2 images from 2016 and 2017. Synthetic images based on the model have been created for several target dates in 2018 and compared to actual imagery acquired at the same date. The benchmarking showed that the forecast quality is reduced in spring. Due to the deterministic nature of the regression model, shifts in the timing of the green-up can result in larger errors and higher-order models are even more affected.

3.3 Benchmarking of methods for pre-processing of SAR time series

Processing methods for SAR time series in literature are focusing on a) benchmarking of existing / currently developed algorithms and tools, b) implementation of spatio-temporal noise filtering tools optimized for Sentinel-1 time series, c) enhanced automated detection and correction of atmospheric influences, d) cluster processing to reduce processing time, e) and derivation of multi-temporal metrics [AD04]. Therefore, the following sections provide methods and benchmarking thereof, focussing on the topics of radiometric calibration, speckle noise reduction and interferometric coherence estimation of SAR time series data.

3.3.1 Benchmarking of candidate approaches for radiometric pre-processing

As mentioned in Chapter 3.2.2, SAR intensity values can be processed to different radiometric calibration levels. These calibration levels can correct for backscatter differences related to local incidence values and thus can make data from different orbits comparable. Here, we test different processing levels for their suitability in the ECoLaSS project. Different calibration levels are calculated and tested for their performance to level backscatter values of individual land cover classes and for images of different orbits.

3.3.1.1 Description of Candidate Methods

Standard radiometric calibration levels for SAR data are beta, sigma and gamma nought. Each applies different radiometric corrections. In addition, removal of terrain effects can be applied, which is sometimes termed SAR terrain flattening.

We test the following calibration levels: Beta nought gives values per unit area in slant range geometry. Sigma nought values are per unit area in ground range, so corrected for local incidence angle. Gamma nought values are per unit area perpendicular to slant range. Flattened gamma nought is additionally corrected for local terrain effects based on a DEM (Small, 2011). Gamma nought is usually used when data from different orbits are combined, since local incidence angles vary from different orbits. We use the SRTM4 DEM as elevation source.

3.3.1.2 Benchmarking Criteria

First, only a visual analysis of output results was performed. Then similarity studies were performed among land cover types at varying incidence angles and for backscatter values in overlapping images from different orbits. The testing criteria were:

REDUCTION OF TERRAIN EFFECTS

The similarity analysis of backscatter values for an identical land cover class at different local incidence values. A number of points for different land cover classes have been derived from existing land cover classifications. The spreading of values has been estimated for the forest class, each calibration level, and ranges of local incidence values. A lower spread of backscatter values is considered better for subsequent land cover information extraction.

SIMILARITY OF IMAGE BACKSCATTER VALUES FROM OVERLAPPING ORBITS IN OVERLAPPING AREAS

Image differencing of images from different orbits is applied to image data processed regarding all above mentioned calibration levels. Remaining differences are likely to be remnants of local incidence values related to either topography or orbital imaging geometry. Lower difference values between adjacent orbits are considered better for consecutive homogeneous land cover classifications.

3.3.1.3 Implementation and Results of Benchmarking

Beta, sigma and flattened gamma nought images from two overlapping orbits were calculated for an alpine test area in the Central test site. An example is given in Figure 3-27.

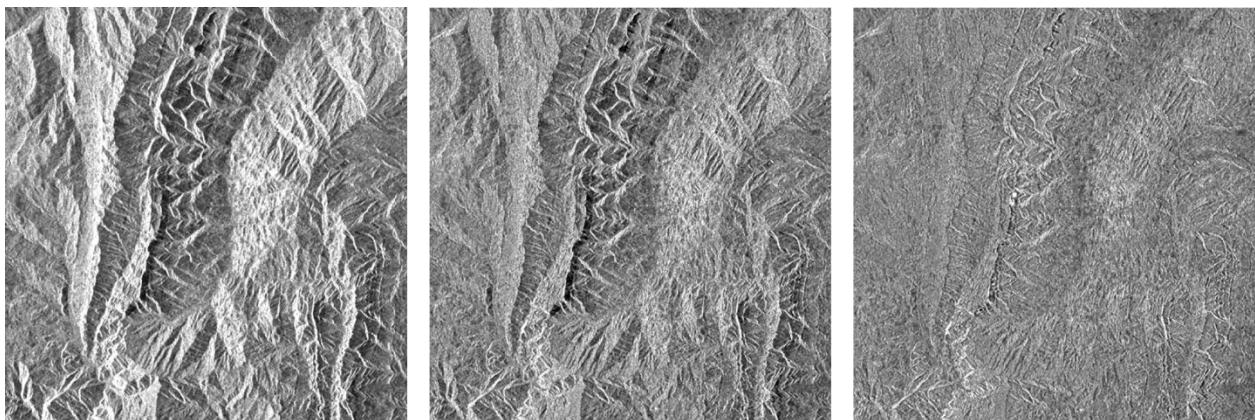


Figure 3-27: different SAR radiometric processing levels: left: beta naught (slant range geometry), middle: sigma naught (corrected for local incidence angle), right: flattened gamma naught (terrain corrections according to Small, 2011).

Difference images were calculated for overlapping areas. For grassland and forest classes which are the most common classes in alpine terrain, AOIs were selected from different local incidence angles and backscatter values were compared.

3.3.1.4 Summary and Conclusions

Beta naught images in alpine settings show strong local incidence related artefacts in backscatter values and are thus unsuitable for pan-European land cover classification. In sigma naught imagery most, but not all, terrain effects could be removed. We find the SRTM4 model is not accurate enough for a full removal of terrain effects. However, it is the only freely available DEM currently available at pan-European scale. With sigma naught processing, imagery from different orbits cannot be used jointly because backscatter values still differ strongly among the different orbits. Flattened gamma naught imagery shows best results for terrain correction and for the similarity of imagery from different orbits. When incidence angles are too steep values are still unreliable. We therefore recommend using flattened gamma naught SAR backscatter data for the ECoLaSS demo site processing if data from different orbits is combined, but to eliminate areas with large incidence angles above 70°.

3.3.2 Benchmarking of candidate approaches for speckle noise reduction

Noise reduction is an important pre-processing step for SAR imagery based land cover classification. In addition to multi-looking, which is already applied to GRD images, SAR speckle filtering can be used to reduce speckle noise in SAR imagery. Common SAR speckle filters include Lee filter, Frost filter, Kuan filter, Gamma-Map filter and multiresolution speckle filters. An overview of proved standard SAR speckle filters is provided in Lee, 1994. However, for time series of SAR data, multitemporal SAR filtering is the most widely used procedure for noise reduction (Quegan, 2000). Here, we present results from a test of different filters for their suitability for land cover classification.

3.3.2.1 Description of Candidate Methods

In a study on SAR filters different adaptive SAR, speckle filters were tested with respect to their suitability for land cover classification (Hoermann, 2011). In addition, a multitemporal SAR filter proposed by Quegan et al., 2000, was tested. Among the mono-temporal SAR filters the Frost-filter (Frost et al., 1981), Modified-Frost filter, Lee filter (Lee, 1980), Refined Lee filter (Lee, 1981), Kuan filter (Kuan, 1985), GammaMAP filter (Lopes et al., 1990), and a multiresolution filter (Meer et al., 1994) were compared. All filters were tested with different filter windows (3x3, 5x5, 7x7) except for the multiresolution filter which automatically adapts the window depending on local backscatter homogeneity. When working with multitemporal SAR data it is usually more efficient to use

multitemporal noise filters. The multitemporal filter preserves or even enhances the separability of small targets while most mono-temporal and adaptive filters do not preserve the geometric detail of the image.

3.3.2.2 Benchmarking Criteria

For estimating the quality of different SAR image filtering techniques both a visual qualitative analysis and a quantitative analysis have been performed. Land cover classification accuracies after filtering were compared both raster based on pixel level and by statistical cross-validation. Additional processing criteria were filtered window size (3×3 , 5×5 , 7×7), and processing unit (db vs. DN).

3.3.2.3 Implementation and Results of Benchmarking

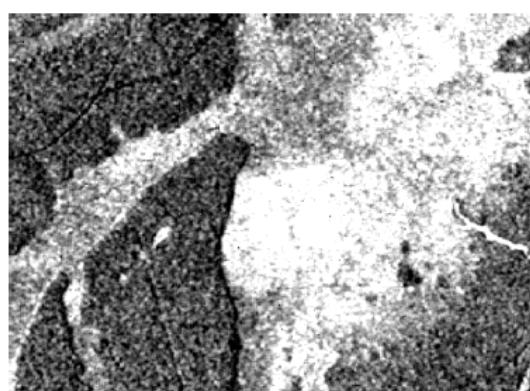
For a test area characterized by different land cover classes, images were pre-processed with different filters using JR RSG software. Most filters used here are also available in other (open-source) SAR software. The following figures (Figure 3-28; Figure 3-29; Figure 3-30) show examples of the different noise filtering results and examples of different filter windows. For larger filter windows, noise is reduced much stronger, but the image becomes blurred and does not preserve radiometric detail. Small targets are lost with large filter windows.



Ungefiltertes Bild



Frost 3x3



Modified Frost 3x3



Lee 3x3

Figure 3-28: Examples of different adaptive SAR speckle filters.

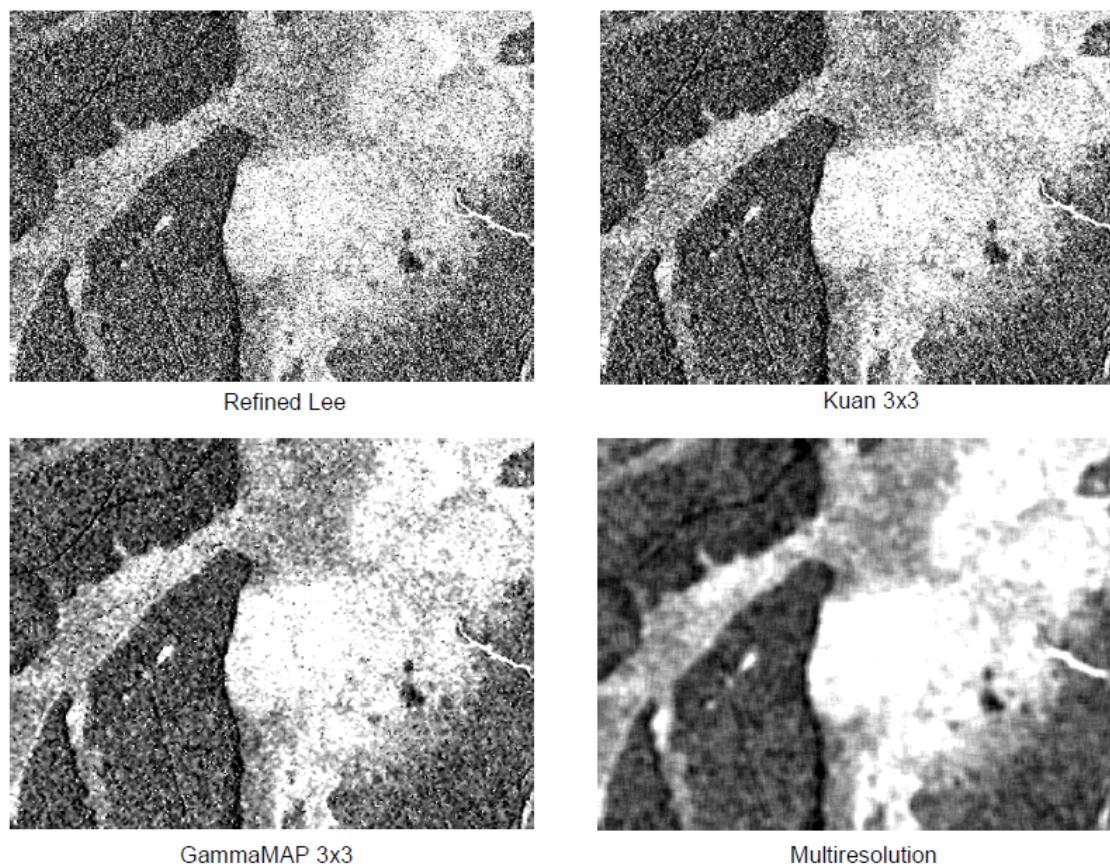


Figure 3-29: Examples of different adaptive SAR speckle filters.

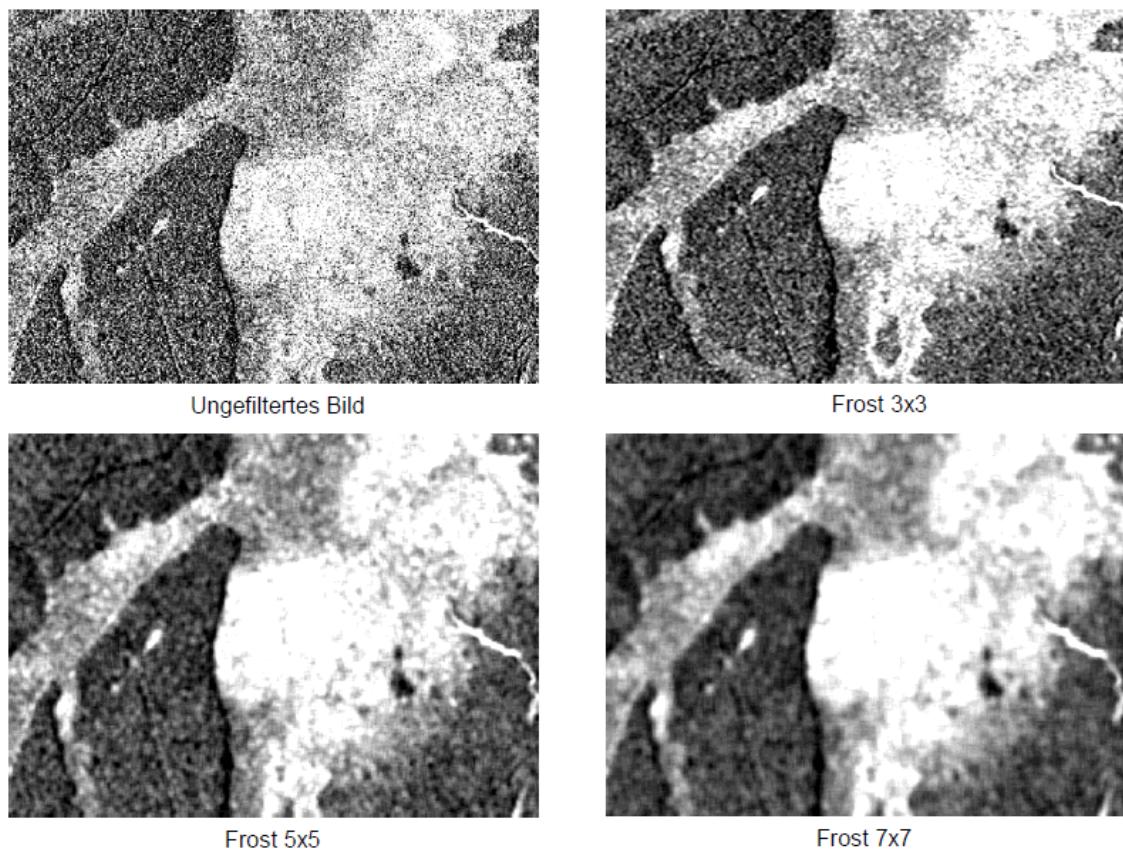


Figure 3-30: Frost filter with different filter windows.

Six land cover classes (water bodies, wetlands, urban, forest, agriculture –vegetated-, bare soil) were chosen for a simple maximum-likelihood based classification based on reference areas. Results were compared to classification results based on optical data.

The following figures (Figure 3-31, Figure 3-32, Figure 3-33) show the accuracies achieved with different filters and additional parameters (window size, a processing unit, spatial resolution). Best results were achieved with Multiresolution, Frost, and Modified-Frost filters, which have a strong smoothing effect. All filtered images show higher classification accuracies than the unfiltered image.

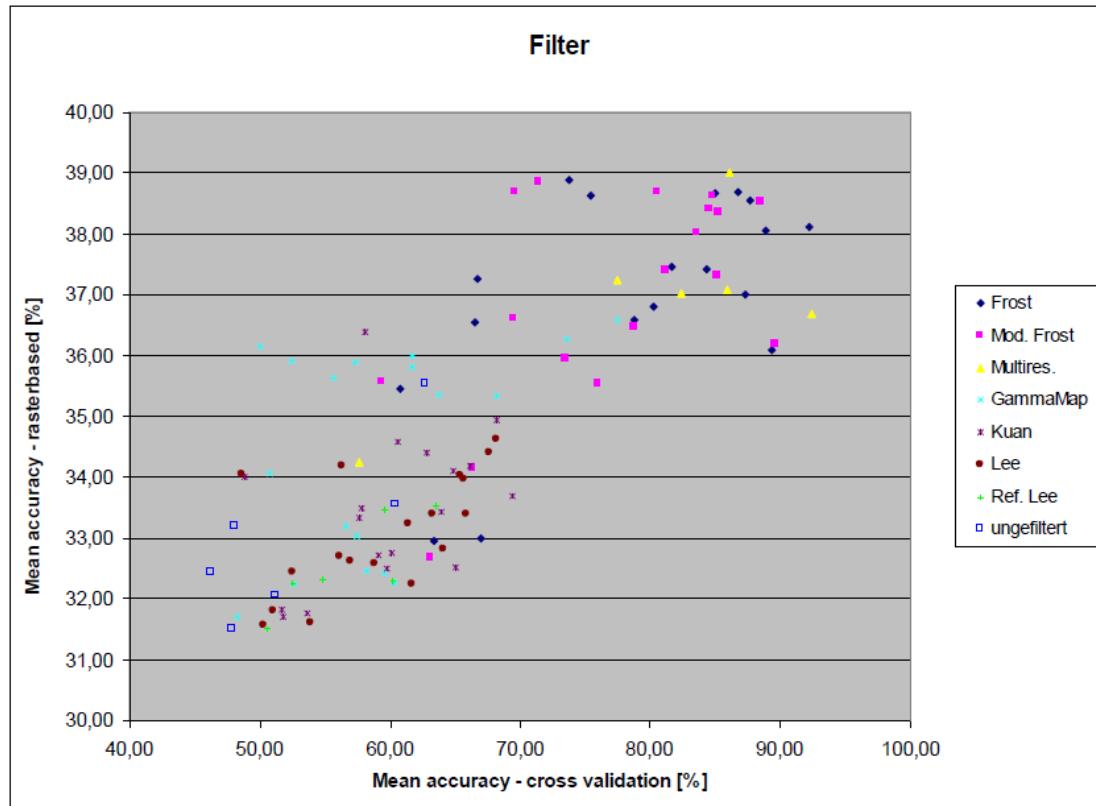


Figure 3-31: land cover classification accuracies achieved with different adaptive SAR speckle filters.

1	DN_15m_frost (3x3)	63,37
2	DN_15m_frost (5x5)	66,95
3	DN_15m_frost (7x7)	66,68
4	DN_15m_modified frost (3x3)	63,05
5	DN_15m_modified frost (5x5)	66,36
6	DN_15m_modified frost (7x7)	80,55
7	DN_15m_multiresolution	85,94
8	DN_15m_gammamap (3x3)	60,34
9	DN_15m_gammamap (5x5)	58,15
10	DN_15m_gammamap (7x7)	59,60
11	DN_15m_kuan (3x3)	65,06
12	DN_15m_kuan (5x5)	60,61
13	DN_15m_kuan (7x7)	60,12
14	DN_15m_lee (3x3)	52,53
15	DN_15m_lee (5x5)	58,87
16	DN_15m_lee (7x7)	64,13
17	DN_15m_refined lee	52,50
18	DN_15m_ungefiltert	51,25

Figure 3-32: cross-validation accuracies for land cover classification achieved with different adaptive SAR speckle filters and window sizes.

	Cross validation
Frost	78,64
Modified Frost	77,26
Multiresolution	80,32
GammaMAP	59,20
Kuan	59,95
Lee	59,34
Refined Lee	56,87
ungefiltert	52,76

Figure 3-33: Mean cross-validation accuracies achieved with different SAR speckle filters, window sizes, processing units and spatial resolutions.

Multi-temporal filtering is the most commonly used method for speckle noise reduction in multi-temporal SAR data stacks.

For a sequence of N registered multitemporal images, with intensity at position (x, y) in the kth image denoted by $I_k(x, y)$, the temporally filtered images are given by:

$$J_k(x, y) = \frac{E[I_k]}{N} \sum_{i=1}^N \frac{I_i(x, y)}{E[I_i]}$$

for $k = 1, \dots, N$, where $E[I]$ is the local mean value of pixels in a window centred at (x, y) in the image I.

(Quegan 2010; the same filter is used in ESA SNAP Sentinel-1 Toolbox)

The advantage of multi-temporal filtering is that it preserves or even enhances the level of geometric detail and increases the signal to noise ratio. Figure 3-34 shows an example of an unfiltered and a multi-temporally filtered seasonal mean image from a SAR stack covering one entire vegetation season. The

preservation of geometric detail is obvious and is beneficial for high resolution LC products and LC classes with small targets (urban built-up areas, urban vegetation).

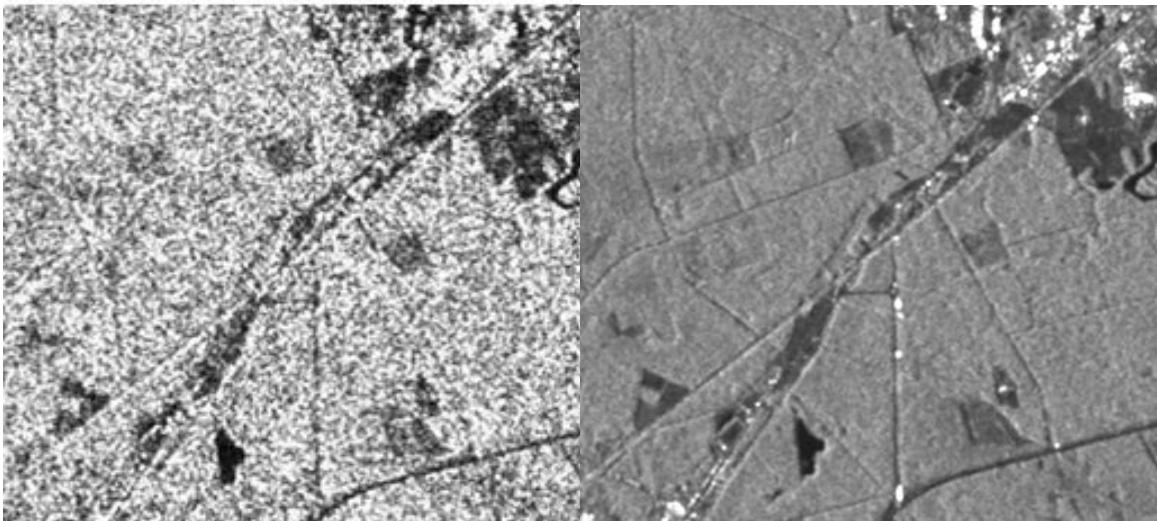


Figure 3-34: left: Single Sentinel-1 scene, unfiltered; right: Single Sentinel-1 scene filtered with MTF, 19 images, 7x7 window.

3.3.2.4 Summary and Conclusions

When working with single SAR images, adaptive speckle noise filters with strong smoothing effects, such as Frost, Modified-Frost and Multiresolution filter, usually show best results for overall land cover classification accuracies. However, geometric detail is lost due to blurring effects. When multitemporal SAR data is available multitemporal SAR filters are better suited for high resolution land cover classification with full resolution, e.g. is a 10m resolution for Sentinel-1 /Sentinel-2. Multitemporal filters, such as the one described by Quegan, preserve the radiometry while reducing noise and sharpening the image. The calculation of seasonal means further reduces noise and enhances class separability.

For ECoLaSS pre-processing, our analysis concludes that multitemporal SAR filters should be applied to pre-processed gamma nought data in DN processing unit. Multitemporal filters can be applied to the entire (or annual) image stack. Seasonal means and seasonal statistics further reduce speckle noise and should be used as input regarding SAR data classification of the HRL land cover classes.

3.3.3 Benchmarking of candidate approaches for interferometric coherence estimation

Interferometric coherence is a measure of decorrelation and can be used to separate different land cover classes. For vegetated areas decorrelation is very high and short-term coherence is required to map differences among vegetated areas. In this section, methods to calculate short-term coherence from Sentinel-1A and Sentinel-1B are tested, where different time windows are analysed by a simple evaluation of the best short-term coherence window. For the short term interferograms, coherent pixels are identified following the same strategy adopted in Pepe et. al. (2015) that is based on analysing the interferometric phase information within a boxcar averaging window.

3.3.3.1 Description of Candidate Methods

6-day, 12-day and 18-day coherence products are calculated using a standard box car coherence estimation approach. The interferometric coherence is a statistical average. It requires averaging over many samples of the same distribution. A commonly used technique for this is the boxcar filter where samples in a 3x3, 5x5 or 7x7 moving window are included in the coherence computation. The boxcar

filter has the advantage of being simple to calculate, being very effective in speckle noise reduction and in preserving the mean value. Its deficiencies are a reduction of spatial resolution due to indiscriminately averaging pixels from inhomogeneous areas and thereby blurring edges and smearing point targets.

3.3.3.2 Benchmarking Criteria

The main focus is on determining reliable short-term InSAR coherence products. Those products can be derived from Sentinel-1 SLC data. In addition, different short-term coherence windows (6-day, 12-day, 18-day) are tested and analysed regarding their potential for land cover mapping. Single Look Complex (SLC) processing focuses the data in azimuth and range to form an image. The SLC processing takes as input, the signal data and the pre-processing output including orbit information and Doppler centroid estimation polynomials. The processing then applies range processing, azimuth pre-processing, azimuth processing and azimuth post-processing. For further description see the technical guide from ESA (ref. <https://earth.esa.int/web/sentinel/level-1-slc-processing-algorithms>).

For the benchmarking tests the SLC images from 2018 were downloaded and the processing was performed according to the following four criteria:

- a. Polarisation, i.e. VV versus VH
- b. Time interval, i.e. 6days, 12 days or 18 days coherence
- c. Resolution, i.e. 20m or 40m
- d. Seasonality for different periods, i.e. 2 months, 3 months or 8 months

3.3.3.3 Implementation and Results of Benchmarking

Testing was performed with Joanneum Research RSG software, more precisely the “Coherence Processing” module (see Figure 3-35). The InSAR coherence pairs were calculated for both available polarisations (VV and VH; ref. <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation-scenario>) and in two modes: First, the InSAR coherence was calculated between a master scene and all other scenes in the stack. The ideal master scene was chosen by the program without user input. Thereafter, all short-term coherence pairs below a maximum temporal baseline (i.e. the difference between two dates) were calculated. Regarding the processing, 6-day and 12-day and 18-day coherence products were generated for the Belgium demo site based on Sentinel-1A and Sentinel-1B SLC imagery for different time periods in 2018, e.g. March/April, April/May, April/June and for the entire period from February to November 2018. The coherence estimation has also been performed for different output resolutions, i.e. 20m and 40m.

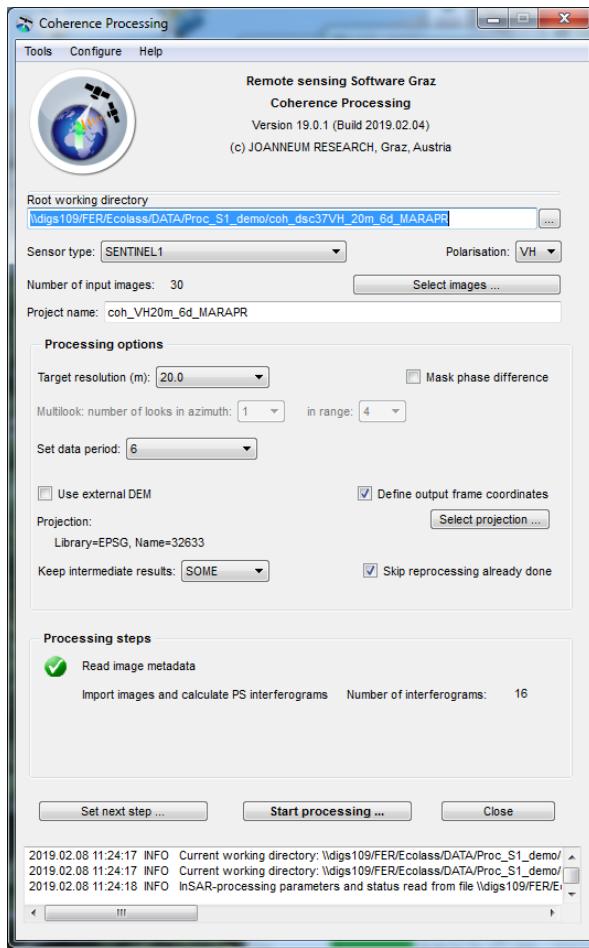


Figure 3-35: RSG Coherence Processing Module.

As can be seen in Figure 3-35 there are three main input options that can be inserted into the “*Coherence Processing*” module, i.e. polarisation, spatial resolution and data period. Additionally, the projection and the output frame - via coordinates insertion - can be defined selectively.

For the processing several steps are implemented and are running sequentially after invoking “*Start Processing*”.

- *Read image data*: reading of metadata from single SLC images
- *Import of images and calculate Interferograms*
 - SatRead: importing of precise orbit files
 - CombSD: combine Sentinel data slices of same relative Orbit and dates
 - Bmerge: merging of bursts and multi-look complex rastering
 - ESD in InSAR pair: co-registration of InSAR pairs based on enhanced spectral diversity
 - Coherence: calculation of coherence of interferograms

Figure 3-36 shows the orthogonal baselines (the temporal and spatial differences) between the recorded scenes. The left Figure shows the baselines for all six-day interval scene pairs, the right depicts the spatial and temporal distances of all scenes to the master scene. The temporal and spatial baselines indicate the scene pairs for which a coherence image was calculated. Each blue line in Figure 3-36, therefore, stands for one result image. The temporal baseline is noted in the number of days a scene is away from the master scene. For classification purposes, it is reasonable to calculate the mean of multiple consecutive coherence pairs rather than a longer-term coherence.

From the images relative to the master scene a comparison of reasonable time intervals to be used in classification can be calculated.

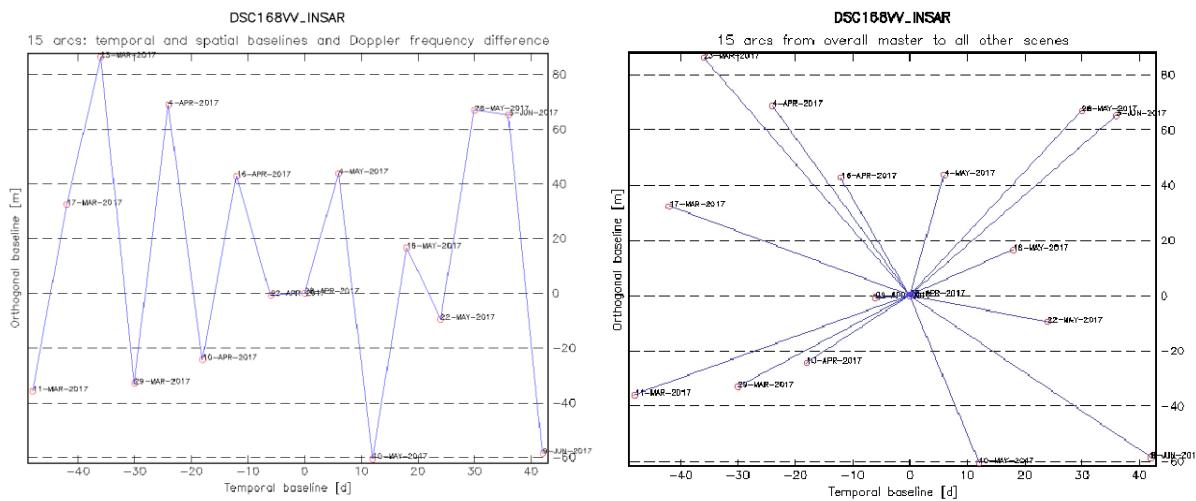
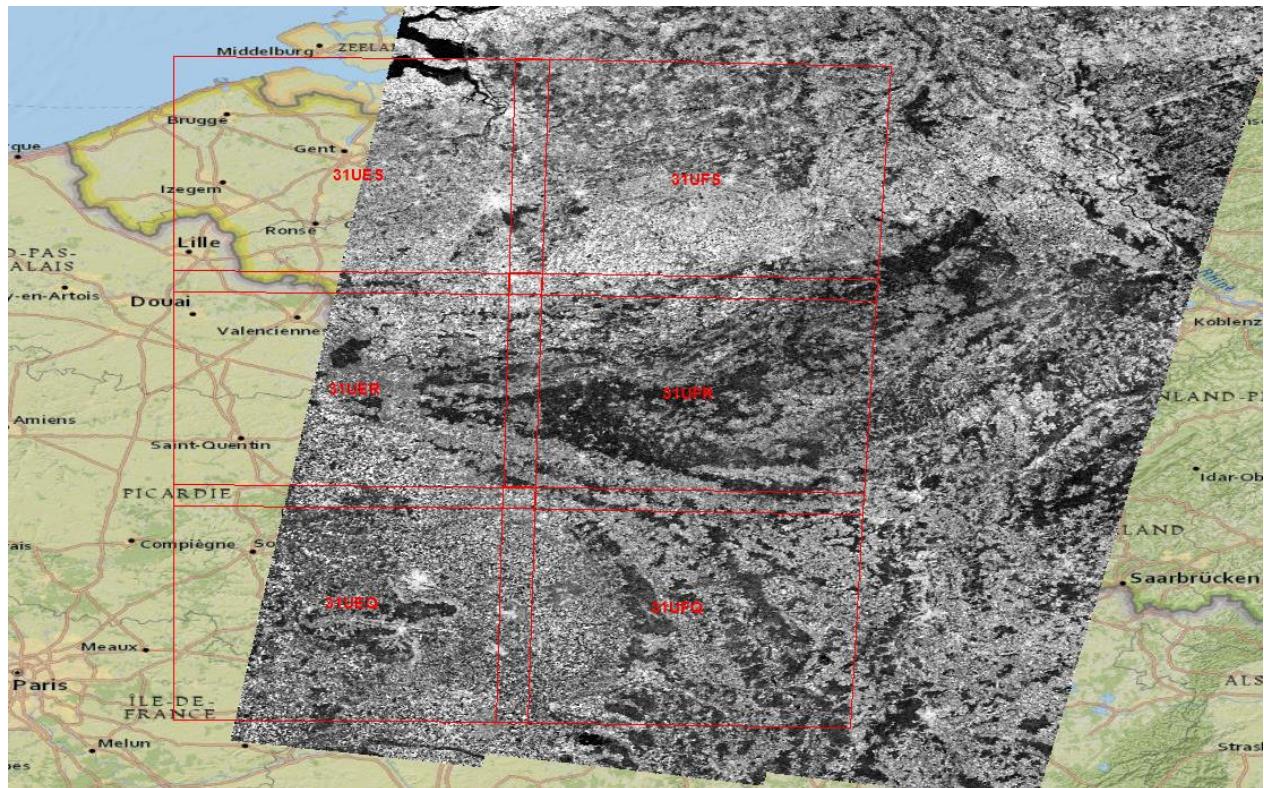


Figure 3-36: calculation of orthogonal baselines for 6-day coherence pairs (left) and for all common-master pairs (right) for Belgium test site.

Example of the resulting images and their extent can be seen in Figure 3-36. Adjacent Sentinel-1 SLC products (North/South) are combined during processing to cover the entire test site area. This procedure is valid for all orbits related to the area of interest.



Figure

3-37

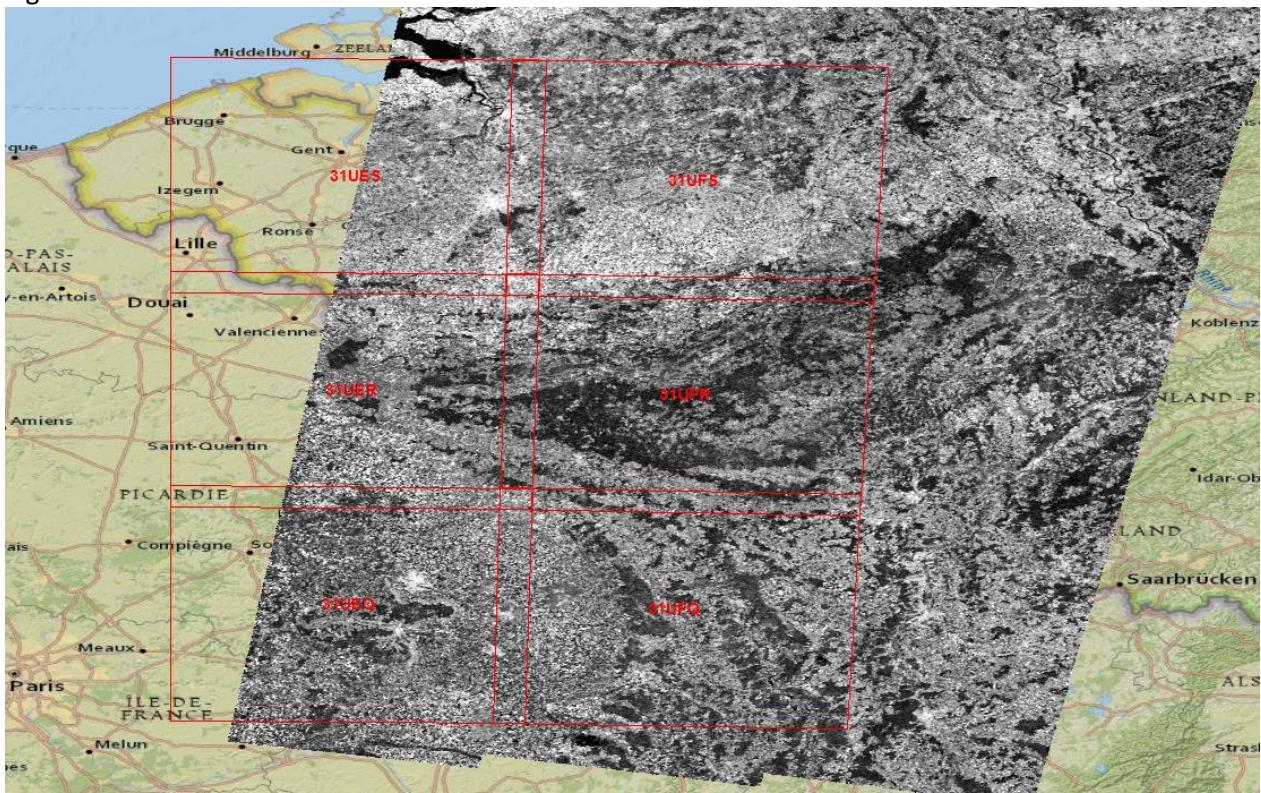
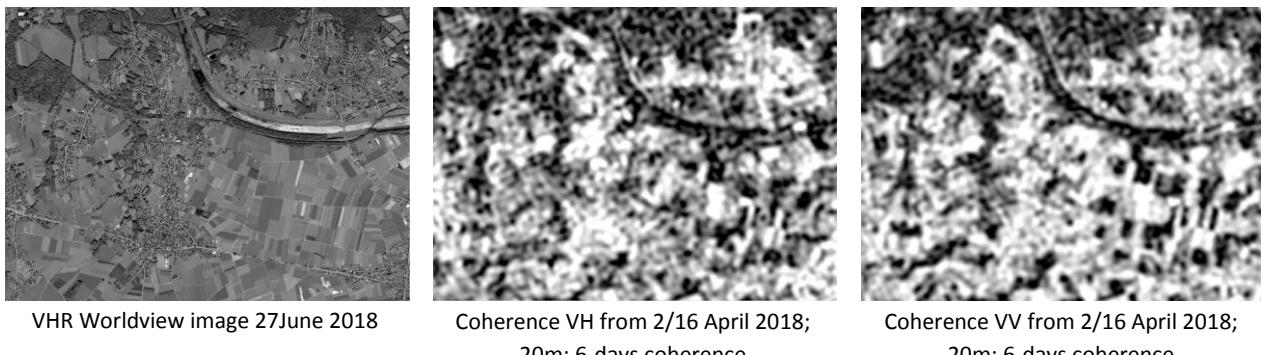


Figure 3-37: Overview of coherence product examples for the Belgium test site.

Polarisation

Sentinel-1 SLC imagery is recorded in two polarisation modes (see Figure 3-38), i.e. VV and VH, as described in the ESA observation scenario. Both polarisation show slightly different results after processing and thus can be used as complementary information.



VHR Worldview image 27June 2018

Coherence VH from 2/16 April 2018;
 20m; 6-days coherence

Coherence VV from 2/16 April 2018;
 20m; 6-days coherence

Figure 3-38: Coherence product examples for polarisation within the Belgium test site (bright: high coherence, dark: low coherence).

Time Interval

Figure 3-39 shows examples of 6-day, 12-day and 18-day Sentinel-1 coherence and its decorrelation over time at an urban area in the Central test site. The master scene was from April 28th, 2017. The decorrelation becomes visible in the enhanced noise in the 12- and 18-day coherence images (center and right) compared to the 6-day coherence (left). This occurs as a result of gradual changes in volume scattering of the vegetated area caused by e.g. plant growth and mowing events. Only entirely stable elements (urban built-up areas) on the surface show no or little decorrelation over long time periods.

The 6-day coherence image also shows high coherence values for agricultural fields and other non-vegetated areas. In the 18-day coherence image, only the built-up areas feature high coherence values while agricultural areas are no longer distinguishable. For identifying urban areas longer coherence windows might, therefore, be more useful. For agricultural applications, short term coherence of 6-days or different images of 18-day coherence and 6-day coherence might be more useful. For other land cover classes, coherence seems rather irrelevant. Forests for example show very low coherence values with a high level of noise.

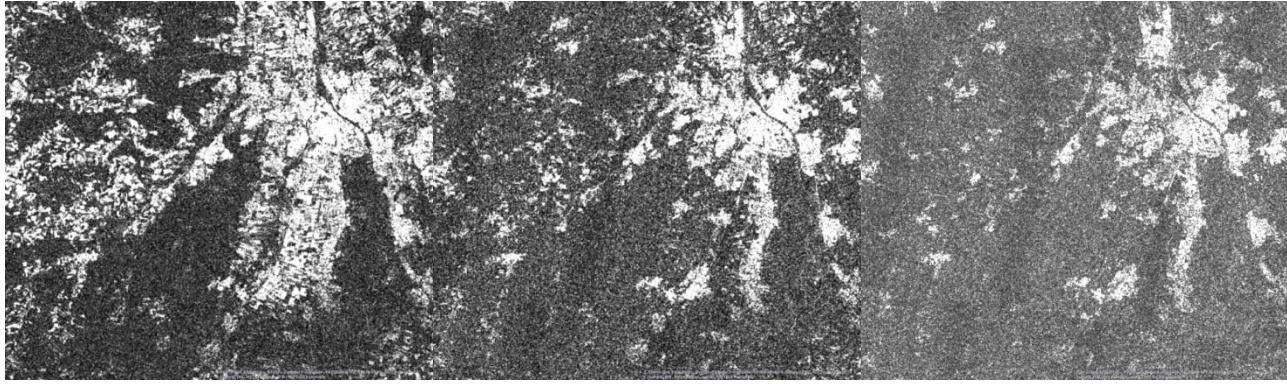


Figure 3-39: Coherence product examples for time intervals within the Belgium test site (bright: high coherence, dark: low coherence).

Seasonality

With a focus on the vegetation season a time series of coherence images was computed from beginning of March until end of October 2018. The 6-days coherence images were calculated with 20m spatial resolution. It can be observed that forest areas show low coherence and urban areas high coherence over the season. Agricultural areas show diverse coherence values, higher values in spring and lower values in summertime.

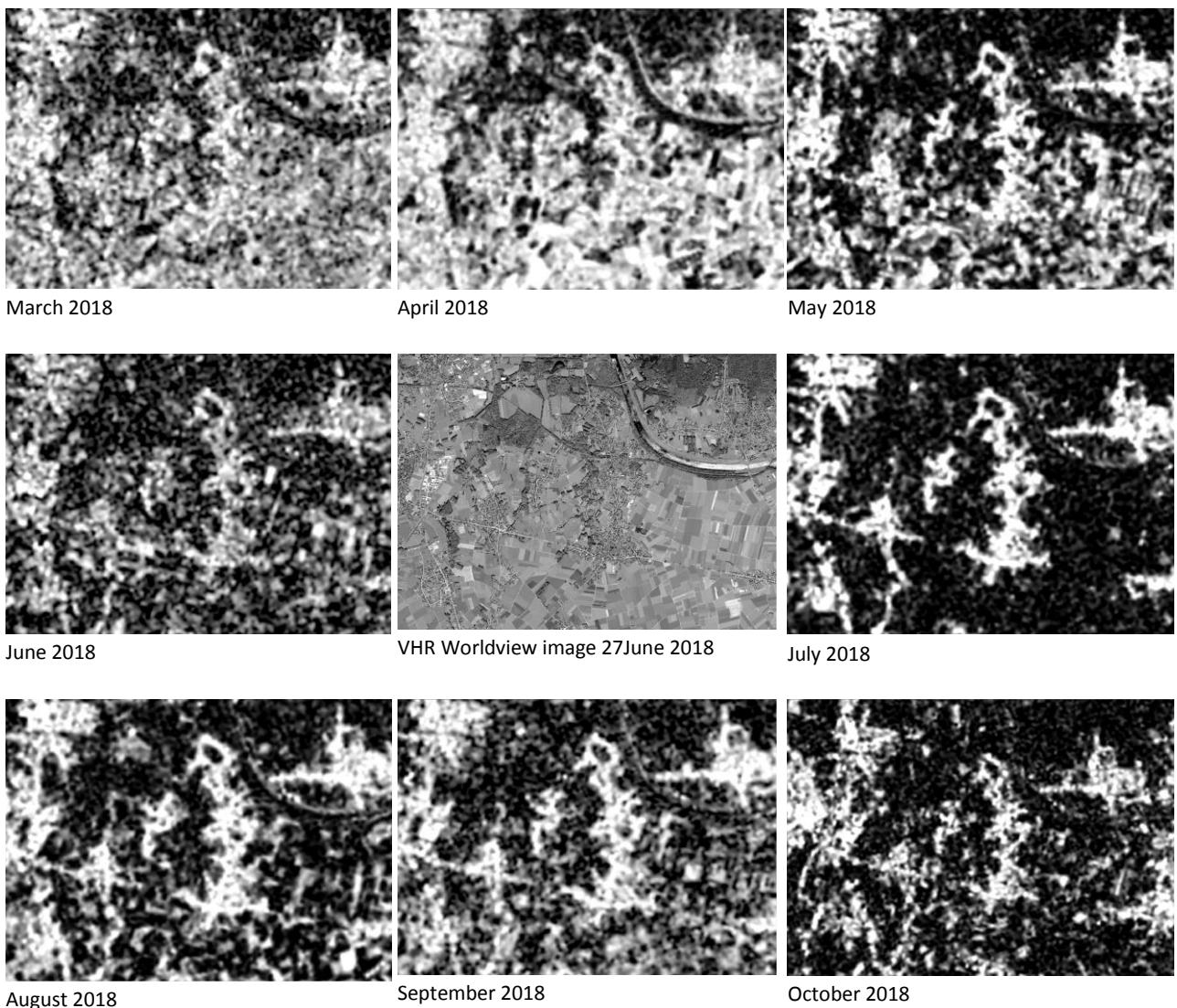


Figure 3-40: Selected examples from coherence processing over the vegetation season 2018 within the Belgium test site (bright: high coherence, dark: low coherence).

It has to be investigated how the heterogeneity in the image contents over the vegetation season can be used for grassland detection and classification on one hand, and if the rather low spatial resolution compared to the optical Sentinel-2 image content can be incorporated into a processing chain on the other hand. It has to be emphasized that the processing is time and resource consuming (one processing run can exceed more than 4TB), thus the computing can only be handled properly with high performance hard and software equipment.

Spatial resolution

The Sentinel-1 radar operates in two main modes: Interferometric Wide swath and Wave. Interferometric Wide swath mode, the default mode over land, has a swath width of 250km and a ground resolution of 5 x 20m. Therefore, the images are computed with a standard spatial resolution of 20m (and with 40m for comparison).

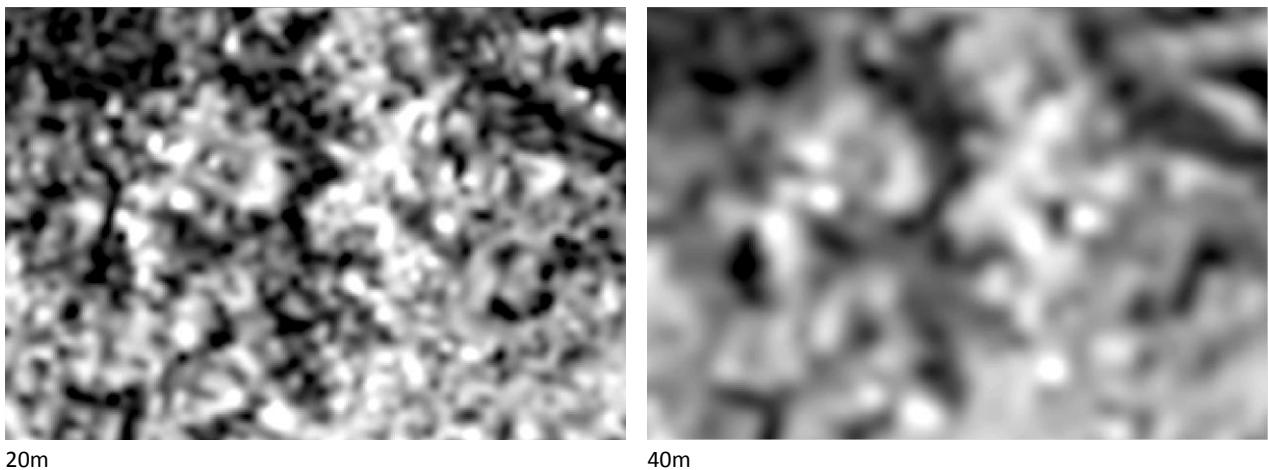


Figure 3-41: Coherence product examples for targeted spatial resolutions.

In Figure 3-41 two examples are displayed showing the outcome from processing with 20m versus 40m, respectively. It can be seen that the 20m product shows significantly better details than the 40m product. For further processing it is therefore advised to select the 20m resolution.

The final results were orthorectified in a post-processing operation to 10m spatial resolution in order to harmonize the images with the Sentinel-2 data. Tests were applied using linear, cubic or 16point sinc resampling. The results show that linear method is not advised to be used for further processing due to geometric artefacts. The other methods show only slightly differences, but the exact field structures are not displayed in a detailed manner, i.e. border lines or ratio of field sizes.

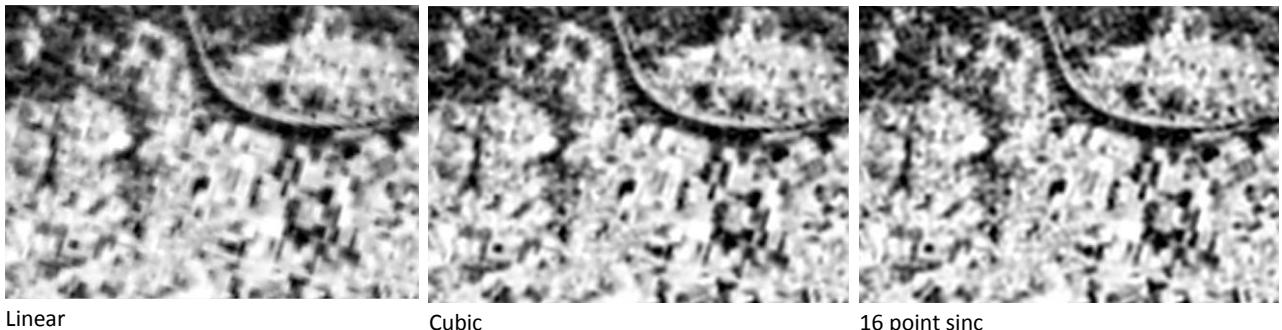
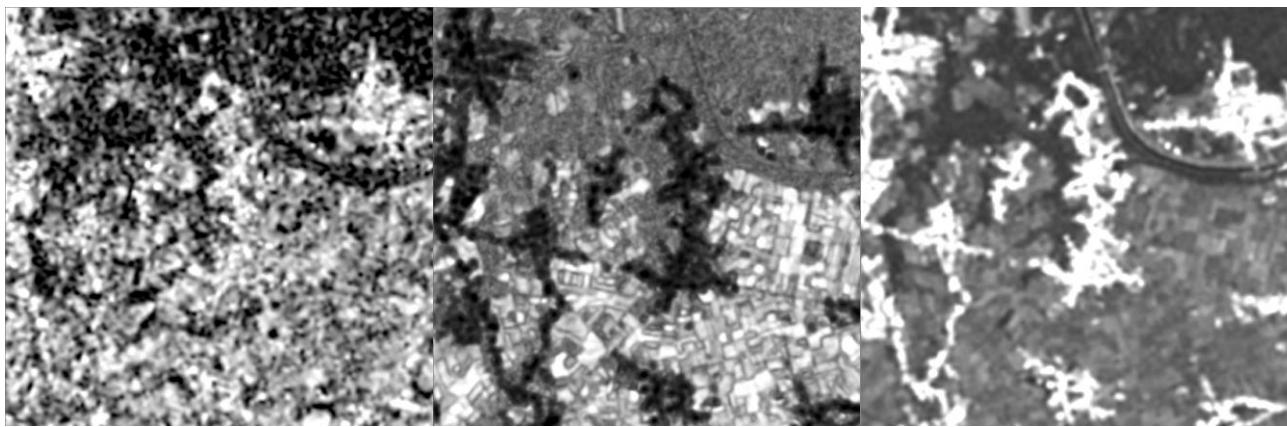


Figure 3-42: Coherence product examples for polarisation within the Belgium test site (bright: high coherence, dark: low coherence).

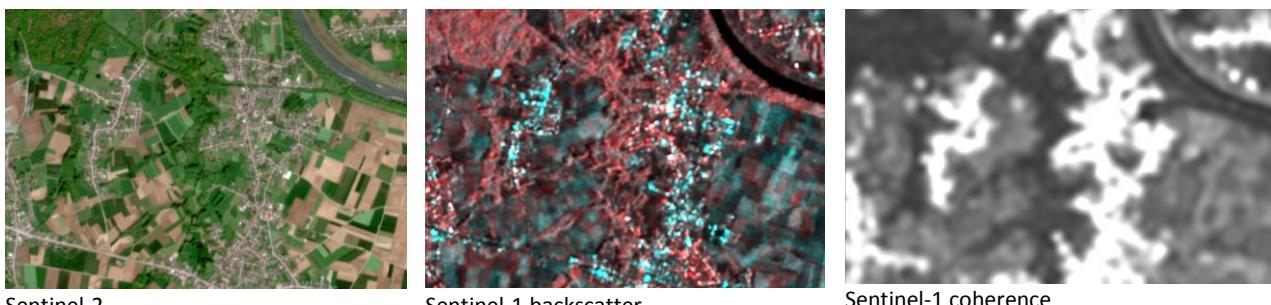
A post-processing has been applied on the time series coherence images using filtering technique of the single images. The filtering results of two periods are compared, i.e. 2 months and 8 months, with the coherence single images. Either the coefficient of variation or the mean for the eight months period were calculated. As can be seen in Figure 3-43 the “speckle” of the single coherence images is reduced due to the calculation of the mean for the 69 input images, but in the filtered images no distinct borders, e.g. for fields or urban, can be recognized. Therefore, the outcome filtered images do not give sufficient spatial detail if compared with the Sentinel-2 image.



Coherence: single image Coherence: eight months coefficient of variation filtering Coherence: eight months mean filtering

Figure 3-43: Comparison of Sentinel-1 coherence products within the Belgium test site.

The processed single coherence images are lacking for spatial details compared to the Sentinel-2 optical imagery with 10m standard spatial resolution (see Figure 3-44). The delineation of agricultural fields can nicely be observed in the Sentinel-2 imagery and also with limitations in the Sentinel-1 backscatter imagery. How the different information content of the Sentinel-1 coherence imagery can be used with other data sources will be investigated in WP33.



Sentinel-2 Sentinel-1 backscatter Sentinel-1 coherence

Figure 3-44: Comparison of Sentinel-2 imagery with Sentinel-1 products within the Belgium test site (left: Sentinel-2; middle: Sentinel-1 GRD backscatter, right: Sentinel-1 SLC coherence).

3.3.3.4 Summary and Conclusions

We have demonstrated that joining Sentinel-1A and Sentinel-1B SLC data in order to derive short-term InSAR coherence is possible using our own JR in-house software RSG, module “Coherence Processing”. The software package is easy to handle and thus can be learned very quickly. However, an analysis of other SAR software packages (see chapter 3.2.4) has shown that most do not support short-term coherence estimation directly.

Short-term coherence seems to be of interest for agricultural applications (bare field conditions) and possibly grasslands, e.g. detection of mowing events or differentiation of intensive versus extensive grassland. Longer coherence time intervals could be of interest for medium resolution detection of urban built-up areas.

Whether coherence data adds significant additional value for HR land cover classification compared to using only filtered GRD backscatter data is investigated further in the thematic classification chains in WP33. By taking into account that huge processing resources are required for large area applications, we recommend to include coherences only in case of significant improvement of the classification accuracies.

4 Pre-processing of ECoLaSS test sites

The main challenge requested as research need is the adaptation and up-scaling of current processing lines from local level to a continental level. Accordingly, a large volume of data should be pre-processed, to be able to produce data composites and derive calibrated and validated variables. This chapter describes pre-processing chains applied within the ECoLaSS project for Sentinel-1 and Sentinel-2 time series.

4.1 Pre-processing of optical data

To assess the required infrastructure, necessary to process high volume of optical Sentinel-2 data streams several software packages can be used depending on the user needs. Which package should be used depends on location area, thematic application and the available software environment. Figure 4-1 gives a general overview of the pre-processing workflow including necessary and optional pre-processing steps regarding Sentinel-2 optical time series data.

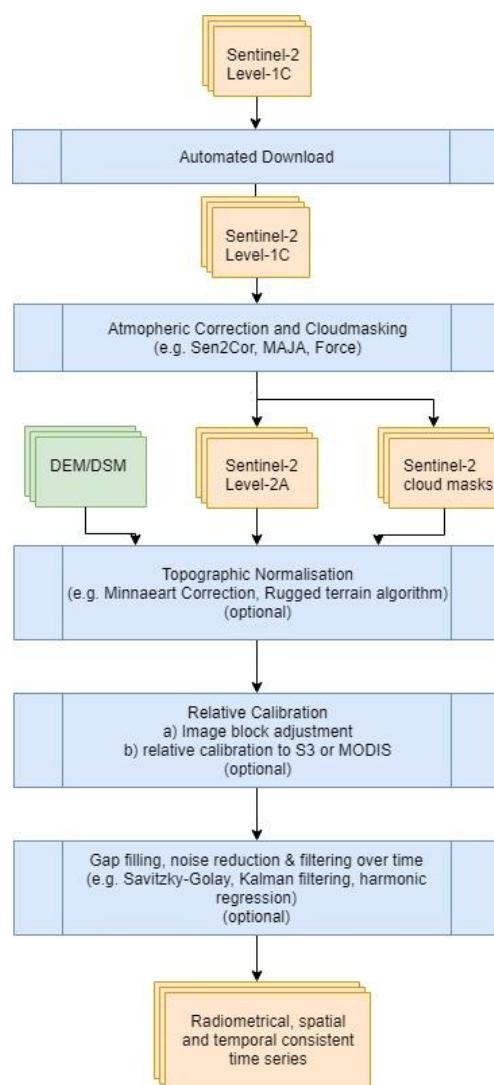


Figure 4-1: Pre-processing of optical time series – workflow.

Which steps should be applied depends on the user requirement, the area and the thematic application. For example a topographic normalisation is recommended in mountainous and hilly regions.

The high volume data processing workflow requires an automated download from available web platforms like the CopHub (<https://cophub.copernicus.eu/dhus/>). The Sentinel-2 pre-processing in WP32 and Task 4 has been performed with several pre-processing packages like Sen2Cor, MAJA, MACCS or Force depending on the site location and thematic content to be derived. Additional, optional pre-processing and post-processing steps were applied by the consortium partners depending on the test site and the demonstration site:

- cirrus removal with Sen2Cor
- terrain correction with the rugged terrain algorithm or Minneart correction
- images export in tif format
- re-projection
- solar angel map
- scene classification calculation and export
- enhanced cloud mask processing
- valid data mask generation

4.1.1 Pre-processing of test sites/demo sites

In the following chapter a detailed description of the optical pre-processing lines applied in the different test and demonstration sites in respect to the site and the thematic applications is given.

4.1.1.1 Central (Austria/ Germany)

In the **Central (Austria/ Germany)** site the Sen2Cor package is used for pre-processing including following steps atmospheric correction, scene classification and other derivates like indices. Level-1C Sentinel-2 scenes are accessed from the CopHub site and downloaded via the SentinelSat tool, setting the cloud filter to below 90% cloud cover. The ESA Sen2Cor package is used for pre-processing including following steps:

- Absolute Atmospheric Correction (Bottom-of-Atmosphere)
- Topographic Normalisation
- Cloud and cloud shadow detection and masking (using the sen2cor SCL mask)

Temporal features were calculated out of the time series (e.g. max, mean, min, etc.) and gaps were filled with scenes from the time series stack automatically. Consequently, no gap-filling and noise reduction has been applied.

The Central demo site contains flat areas and complex mountainous surfaces (parts of the Alps). The topography can significantly influence the radiometric properties of the optical data, caused by the different lighting angles resulting from the topography. A topographic normalisation has been performed in order to correct the topographic influence and to guarantee a proper consistency in the time series (e.g. shadows dramatically vary throughout the seasons). The topographic normalisation approach implemented in the Sen2Cor software uses a 90m DEM and a set of derived products like slope, aspect and terrain shadow maps.

The Sen2Cor output bands are automatically resampled to 10m. In particular, all 20m bands are resampled using the cubic method whereas SCL is resampled using nearest neighbourhood. The three bands with 60m spatial resolution (Bands 1, 9 and 10) will be omitted in the Level-2a output since they are not needed for land cover applications.

Another major output of the processor is the scene classification layer (SCL), which detects 12 different land cover classes. The SCL is used for cloud and cloud shadow masking during the classification process. The cloud cover metric used for analysis and classification does not rely on the official metadata cloud

value provided by the original Sentinel-2A product, but is derived from the Scene Classification produced by Sen2Cor. The resulting cloud-mask layers can later be used to optimize the Sentinel-2 input data for the feature generation step and will ensure optimised input data for subsequent thematic classification steps. The affected areas in the specific satellite image are masked out so that the respective pixel values are not considered in the calculation of (time-) features.

From the surface reflectance products, the following spectral indices are derived:

- NDVI (Normalized Difference Vegetation Index)
- NDWI (Normalized Difference Water Index)
- BRIGHTNESS (derived through summation of the values of the bands Green, Red, NIR and SWIR1)
- IRECI (Inverted Red-Edge Chlorophyll Index)

Together with the spectral information, these vegetation indices are the input for the subsequent thematic processing.

For processing the phase 2 test and demo tiles, it was decided to use the same pre-processing workflow as for phase 1. This decision was made on the basis of performance in phase 1, and to guarantee time series consistency throughout the project. The workflow developed for phase 1 was stable and produced reliable results. In addition, Sen2Cor is especially designed to process Sentinel-2 data and with regard to upscaling from the test site to the demo site or even to a pan-European level, it can be used to process consistent and high quality Level-2A data with minor effort.

However, issues were encountered regarding the cloud masks generated by the Sen2Cor processor. Clouds and cloud shadows were not fully and reliably detected, which led to artefacts in posterior time-features computations. Therefore, a post-processing has been applied to the original Sen2Cor Scene Classification Layer to enlarge the cloud mask. The processing consists of the following steps:

- Creation of a binary mask from the original SCL (classes 0, 1, 3, 8, 9, 10, 11)
- Computation of the object size for each object geometry
- MMU filtering (0.25ha) for class 3, 8, 9 and 10
- Buffering: 50m for class 3, 8 and 10; 100m for class 9
- Mosaicking of the results
- Substitution of the original SCL by the modified raster

This approach proved to be sufficient in order to reduce the amount of artefacts in the time feature computation by enlarging cloud and shadow areas. It has been implemented by Python scripting and convinces by its high performance. Results are shown in Figure 4-2.

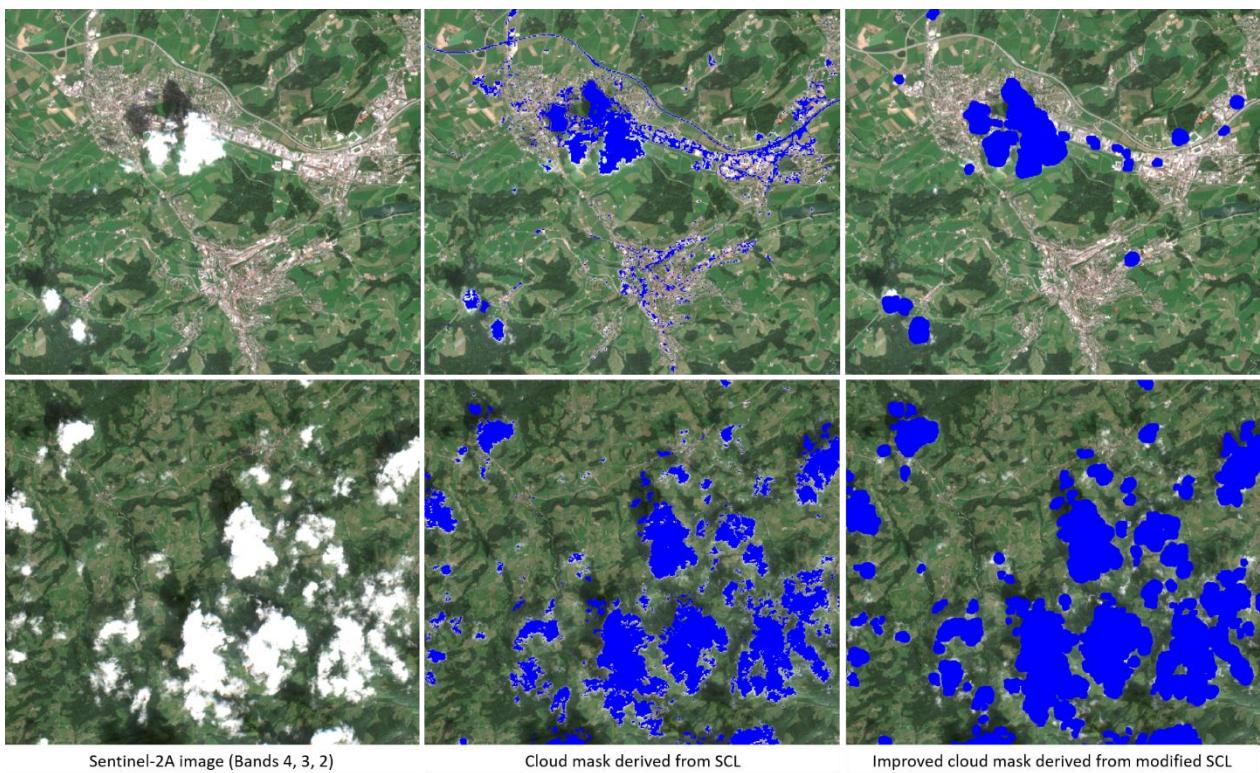


Figure 4-2: Sen2Cor SCL derived cloud masks (in blue) and improved cloud masks derived from the modified SCL.

The automation and recommendations of applying a larger buffer regarding cloud masks outlier detection have been improved in order to work in larger areas with a certain guarantee to avoid artefacts. However, this is done at the cost of reducing the data sample. In cloud prone regions this can seriously affect the time series density regarding key growing periods and phenology indicators. To complement Sen2Cor and overcome some issues, parallel processing scripting and the use of improved cloud masks is recommended

In case of the agriculture task, for the shorter periods, e.g. the time period from 15th of March to 14th of May, this results in even less valid data pixels, as usually the spring period has more scenes affected by heavy cloud cover than the summer or autumn period. Therefore, it is recommended to look at another cloud masking approach, like for example, the Fmask algorithm. Further improvements of the Fmask algorithm for Sentinel-2 images can be expected from Frantz et al. (2018) who implemented the Cloud Displacement Index (CDI) for a better separation of clouds and bright surfaces. The CDI makes use of the near-infrared parallax and fully compensates for the missing thermal band in the Sentinel-2 data. The parallax approach looks promising but was not available for testing during the current proposal.

Another promising approach is provided by the FORCE tool, which not only generates cloud masks but can be used as a Level-2 Processing System. It is intended to be an all-in-one solution for the mass-processing of medium-resolution satellite imagery to enable large area and time series applications.

Issues remain related to the co-registration between Sentinel-2A and Sentinel-2B, and the multi-temporal registration in general. In addition, some users can also incur into disk space limitations in frame of the time series analysis.

4.1.1.2 South-East (Greece/Bulgaria)

In the **South-East (Greece/Bulgaria)** site the workflow has been extended to testing MACCS and MAJA. The first step is an automatic download of the Sentinel-2 Level-1C products inside the AOI from the

Copernicus Hub or the Amazon S3 Bucket. MACCS/MAJA performs the following steps: Clouds and cloud shadows detection, cirrus detection and correction, slope correction and atmospheric detection. Each 20 meters band of the Sentinel-2 Level-2A products are resampled at 10 meters of spatial resolution. The products of interest are the Flat REflectance (FRE) products from MACCS/MAJA where corrected atmospheric and slope effects are applied. The cloud masks values are encoded using the most significant bit and the sum of those for handling overlapping classes in the same band. The snow detected values in the geophysical mask are extracted and merged into the cloud masks accordingly to the previously mentioned method adding a higher bit value.

4.1.1.3 West (Belgium)

In the **West (Belgium)** site the pre-processing chain is performed with the ‘Sentinel-2 Reflectance Data Processing’ tool implemented by JR, which presents a combined workflow to process corrected Bottom-of-Atmosphere products. The Sentinel-2 Level-1C products are automatically downloaded from the CopHub and processed including following steps atmospheric correction, scene classification and topographic normalization with the STRM 90m DEM. An additional cloud mask post-processing has been performed (applying a filling threshold of 30 pixels, a minimum cloud size of 10 pixels and a dilate operation to extend the cloud borders by 10 pixels). Further tests were applied using the Fmask algorithm for cloud masking.

The results of this processing step are comprised of atmospherically corrected imagery and supplementary data listed in Table 4-1.

Table 4-1: File suffix for processed products.

File suffix	content
.tif	GeoTIFF of Scene
.zip	output from Sen2Cor
_cld	cloud probability
_cloudmask	mask from derived clouds
_mask	scene, no-scene
_ncl	scene without clouds
_sol	solar azimuth and solar zenith in degree
_scl	scene classification

Additionally, to the previous mentioned pre-processing steps short filenames can be generated, to allow easier management and access to the time series data.

4.1.1.4 South-West (France)

In the **South-West (France)** site pre-processing has been performed using the MAJA executable (Hagolle, Huc, Desjardins, Auer, & Richter, 2017). It comprises the following processing: Estimation of H₂O content; Absorption correction using weather analysis; Cloud detection; Aerosol detection; Atmospheric correction; Composite image update and adjacency effect and slope correction; Creation of the archive. The MAJA packages provide two types of surface reflectance products: (a) Surface Reflectance products which are corrected for atmospheric effects, including adjacency effects and (b) Flat Reflectance products which are also corrected for slope effect, which consists in suppressing the apparent reflectance variations.

Several prototypes developed on this demonstration site:

- IMP layers, status layers 2017 and 2018, change layers 2015-2017 and 2017-2018

- New Land Cover layers, status layers 2017 and 2018

The NLC requires multiple thematic classes – the choice was made to use the most robust pre-processing algorithm, regarding the atmospheric correction and cloud masking (Baetens, Camille, & Hagolle, 2019) since the demo site covers 3 distinct biogeographical regions with a plurality of land covers that can turn out to be a source of confusion for classification.

A commercial licence is currently in preparation at the CNES – for the time being, the MAJA datasets are freely downloadable from the <http://www.theia-land.fr/en> website for research purposes.

The naming convention, that can be generically described by **Sentinel-2X_YYYYMMDD-HHMMSS-XXX_L2A_Txxxxx_D_V1-9_FRE/SRE_BXX/STACK.tif**, is explained in Table 4-2.

Table 4-2 - File name filed convention for the MAJA datasets

File name section	Meaning
Sentinel-2X_	From a Sentinel-2 satellite system
YYYYMMDD-HHMMSS-XXX	Datatake sensing start time and orbit number
L2A	Level of processing
V1-9	Version of pre-processing
Txxxxx	Tile number
FRE	<i>Flat REflectance</i>
SRE	<i>Surface REflectance</i>
BXX	<i>band number</i>
STACK	<i>all bands stacked</i>

4.1.1.5 Northern (Sweden)

In the **Northern (Sweden)** site pre-processing chain was built entirely around the freely available software tools provided by ESA. Sentinel-2 Level-1C data were downloaded from the Copernicus Open Access Hub and processed to Level-2A (surface reflectance) by means of Sen2Cor (v2.4.0). The scene classification performed by Sen2Cor was used for cloud masking. Sen2Cor was set-up to process at the highest possible spatial resolution of 10m.

In general, the experience with Sen2Cor was very positive. In the used version it has been found to be robust enough for operational processing on a HPC cluster environment. Nevertheless, for large-scale processing, it would be desirable if peak memory consumption was reduced further in order to scale to a higher number of CPUs per node.

RESULTING DATA

In phase 1 of ECoLaSS, an overall data amount of 3,81 scenes and a processed data volume of approx. 7.4 TB, has been processed covering the test and demonstration sites (in Task 3 and Task 4).

Table 4-3: Demo sites – resulting Sentinel-2 time series imagery phase 1.

Demo sites	S2 no. images	Size (km ²)	processed data volume	Time frame
North	577	65,056	671 GB	01.01.2017 - 30.11.2017
Central	641	96,032	1.19 TB	01.10.2016 – 14.11.2017
West	410	65,026	1.4 TB	01.01.2017 – 15.11.2017
South-West	320	62,345	523 GB	03.01.2017 – 14.11.2017
Mali	1,811	44,920	3.5 TB	01.04.2017 – 31.12.2017
South Africa	56	45,017	123 GB	07.10.2016 – 22.04.2017
SUM	3,815	378,396	7.4 TB	NA

Table 4-4: Demo sites - resulting Sentinel-2 time series imagery phase 2.

Demo sites	S2 no. images	Size (km ²)	processed data volume	Time frame
North	908	65,056	1.9 TB (Compressed)	01.01.2018 – 11.11.2018
Central	1,068	96,032	4.54 TB	01.01.2018 – 31.12.2018
West	702	65,026	2.1 TB	15.11.2017 – 15.11.2018
South-West	479	62,345	686 GB	01.01.2018 - 15.11.2018
South-East	1,681(*)	64,960	722.75 GB(*)	01.01.2017 – 31.12.2018
Mali	2,905(*)	470,093	1.61 TB(*)	01.04.2018 – 31.12.2018
South Africa	2,138(*)	373,687	1.07 TB(*)	01.04.2017 – 30.11.2017
SUM	9,881	1197,199	12.62 TB (*)	NA

(*) can change after the finalization of the pre-processing.

For the Sentinel-2 optical data an overall number of 12015 scenes and a processed data volume of about 19.3 TB, have been processed covering the test and demonstration sites (in Task 3 and Task 4) for phase 1 and 2.

Regarding the pre-processing of Sentinel-2 optical data streams several software packages have been compared (MAJA, Sen2Cor, Force). In general, all software packages have advantages and drawbacks, not performing sufficiently enough to achieve a high thematic quality, content and accuracy in any subsequent information extraction process.

Concerning the atmospheric correction most algorithms offered comparable performance. The best estimates were found by FORCE, LaSRC, MACCS (MAJA precursor) and Sen2Cor (Doxani et al. 2018). In

general, arid scenes without dark dense vegetation pixels posed challenges to all algorithms with image-based AOT retrieval.

Concerning the cloud/snow masking Sen2Cor presents a generally reasonable detection of clouds and cloud shadows, but a high confusion with bright surfaces is present. Furthermore, confusions between clouds and fractional snow as well as cloud shadows and water could be noticed. The Sentinel Level-1C Opaque cloud and Cirrus Cloud mask detects clouds and cirrus with notable cloud commission errors. MAJA produces reasonable cloud masking results, by avoiding commission errors through bright surfaces using multitemporal data. MAJA uses a 400m buffer around each cloud, in case it has fuzzy edges, therefore here some clear pixels are lost. Fmask 4.0 cloud and shadow masks identify clouds, Cirrus and shadows reasonably well. The software has the possibility to adjust cloud probability to tune the amount of false cloud commissions/omissions. No notable confusion between snow and clouds can be noticed. Still the algorithm shows a tendency of overestimating the clouds, depending on the region and snow conditions. A cloud mask post-processing step after all processors might be still necessary to reduce noise and fill small gaps.

It is recommended that a topographic normalization should be applied in hilly to mountainous terrain. One big limitation is the quality of the DEM, the current Level-1C DEM (SRTM 90m) may not be good enough and there is no possibility to use another DEM within the Sen2Cor package. The main issue with the topographic normalisation in Sen2Cor is the over-correction of south-facing slopes. The DEM with the highest spatial resolution should be used. It is therefore recommended using topographic normalisation tools which allow the user to define their own DEM.

4.2 Pre-processing of SAR data

In order to design an optimized SAR pre-processing workflow for the ECoLaSS project and future Sentinel based Copernicus land cover mapping, a state-of-the-art literature review on SAR pre-processing, has been performed analysis of existing SAR pre-processing software and performed a number of tests and benchmarking with pre-processed Sentinel-1 data. SAR time series are based on Level-1 products in Interferometric Wide swath (IW) mode and Level-1 Ground Range Detected (GRD). The IW mode is considered the main acquisition mode over land and satisfies the majority of service requirements. In addition to GRD backscatter products, short-term coherence products were processed based on SLC data.

As a result of above testing and benchmarking an optimized SAR pre-processing workflow has been designed for the ECoLaSS project which is presented in Figure 4-3. This generic workflow can be conducted with several SAR pre-processing tools like Snap, Impact, Doris etc. The main difference between these software packages is the order of pre-processing steps. If the precise orbit files are available, the co-registration can be done before the orthorectification. It is recommended to apply the orthorectification step as last step and to be able to apply the filtering methods on the pixel in image geometry. Otherwise the distortions most notably in mountainous areas would lead to blurring effects after filtering.

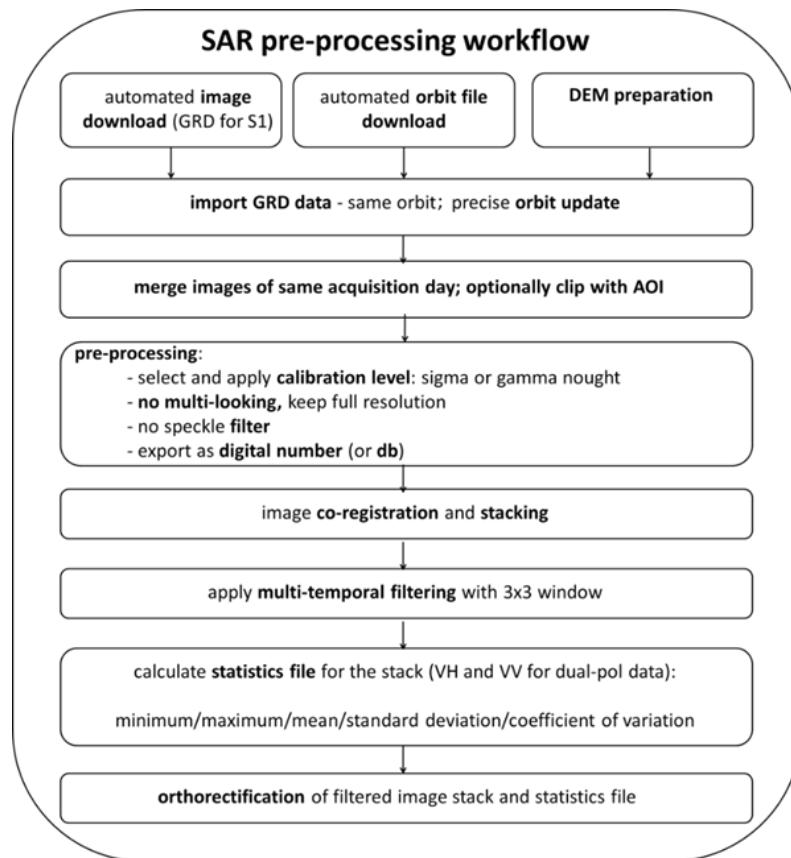


Figure 4-3: Sentinel-1 pre-processing workflow in ECoLaSS.

The pre-processing of S-1 time series is based on Level-1 products in Interferometric Wide swath (IW) mode and Level-1 Ground Range Detected (GRD). This mode has been chosen because the IW is considered to be the main acquisition mode over land. An automated processing chain for Sentinel-1 data pre-processing has then been implemented by each partner using SNAP or the Remote Sensing Software Package Graz (RSG) module “Space Suite”. It comprises the following optional processing steps:

- automated SAR image download by time window
- automated preparation of digital elevation data (SRTM-V4)
- automated download and update of precise orbit files
- merge of GRD tiles from same orbit and same acquisition day
- clip of image extent according to the sites
- calculation of local incidence angle map
- radiometric calibration
- radiometric terrain corrections,
- image registration in image geometry
- multi-temporal speckle filtering
- calculation of temporal image stack statistics (see WP33)
- orthorectification to LAEA or UTM
- coherence calculation

The orthorectification should be applied as last step due to the pixel distortion.

Table 4-5: Test and Demo sites – resulting Sentinel-1 time series imagery in phase 2.

Test and demo sites	Sentinel-1 no. images	Sentinel-1 processed data volume	Time frame	Size (km ²)
North	498	484 GB	01.01.2017 - 30.11.2018	65056
Central	243	2.24 TB	01.10.2016 – 14.11.2017	96032
West	704	2.3 TB	01.01.2017 – 11.11.2017	65026
South-West	785	1.8 TB	01.01.2017 – 15.11.2017	62345
Mali	NA	NA	NA	NA
South Africa	NA	NA	NA	NA
SUM	2824	6.5 TB	-	-

Table 4-6: Test and Demo sites – resulting Sentinel-1 time series imagery in phase 2.

Test and demo sites	Sentinel-1 no. images	Sentinel-1 processed data volume	Time frame	Size (km ²)
North	498	484 GB (compressed)	01.01.2017 - 30.11.2018	65056
Central	291	1.9TB	01.01.2018 – 31.12.2018	96032
West	733	2.5 TB	15.11.2017 – 15.11.2018	65026
South-West	665	2.6 TB	01.01.2018 - 15.11.2018	62345
South-East	1966 (*)	14.9 TB (*)	01.01.2017 – 31.12.2018	64960
Mali	318 (*)	2.24 TB (*)	01.04.2018 – 31.12.2018	470093
South Africa	266 (*)	1.93 TB (*)	01.04.2017 – 30.11.2017	373687
SUM	4737	26.56 TB	-	1197199

For the Sentinel-1 RADAR data an overall number of 7561 scenes and a processed data volume of about 33.06 TB, have been processed covering the test and demonstration sites (in Task 3 and Task 4) for phase 1 and 2. Altogether, optical and radar, the pre-processing data volume encompasses almost 52.36 TB (in Task 3 and Task 4).

4.2.1 Pre-processing of test sites/demo sites

In the following chapters a detailed description of the SAR pre-processing lines applied in the different test and demonstration sites in respect to the site and the thematic applications is given.

4.2.1.1 Central (Austria/ Germany)

In the **Central (Austria/ Germany)** site the Sentinel-1 data is downloaded from the CopHub via sentinel-sat tool. Only scenes from the ascending orbit with the relative orbit number 15 from 1st of January until 1st of November 2018 was downloaded, covering the time window required for classification purposes as a result of Phase I performances.

The pre-processing is performed with ESA SNAP. SNAP is a common architecture for all Sentinel products. It consists of a collection of SAR processing tools including calibration, speckle filtering, co-registration, orthorectification. It is easy to handle and provides a very good workflow builder (Graph Processing Tool). Also, it is regularly updated and has a vivid community of users.

The processing steps applied for the test site comprise the orbit update (includes automatic precise orbits download), thermal noise removal, radiometric calibration generating a beta band, terrain flattening to gamma naught based on SRTM 1sec HGT, terrain correction using the same DEM generating a 10m resolution product and the export of the scene in DN units. Then up to 10 scenes from the same orbit/slice are co-registered/stacked in WGS UTM 32N projection. After the orthorectification of the images, the multitemporal speckle filtering has been applied using a Frost 5x5 window. Python scripts are used in indices derivation: VV (Gamma naught), VH (Gamma naught), Normalized Difference VV/VH, Ratio VV/VH. After pre-processing, issues were encountered regarding geometric shifts between some of the scenes (we experienced this problem with 14 of 92 processed scenes for the test site). This problem occurred probably during the co-registration step, which is needed for the multitemporal filtering. However, we have not been able to identify or reproduce the exact reason of this artefact. It might have been caused within the SNAP software manipulations. We were able to re-process some of the scenes, but in the end, we had to eliminate 7 scenes from the time stack due to this geometry shift. To overcome this problem, it is recommended to implement a geometric consistency check that automatically checks for shifts within the co-registered scenes.

Other issues that came up during processing of the Sentinel-1 data were mainly related to the SNAP software. The regular updates sometimes cause processing problems (e.g. old versions are not available anymore). Another issue is that sometimes an empty output is created. The errors are not traceable, which means a lot of data has to be processed again. Other issues are geometric offsets caused by DEM problems or during co-registration. SNAP is sometimes a black box, that's why it's hard to trace down the issues, which are mostly not reproducible.

For future processing of SAR data, it is recommended to build a more stable SNAP workflow which includes automatic quality control steps. Also, SNAP provides a lot of parameter settings for the different steps. These settings could further be tested to see if they have impacts on the output and the thematic processing of the data. For consistency reasons, we used for most of the steps the same parameter settings that were already used in Phase I.

Regarding coherence, besides the tool in place for the project, it was decided to use the SNAP software under the same premises of consistency and assess the Sentinel software performance because we understand it is the tool that most users will employ when processing radar data. Nonetheless, to work in large areas, batch scripting for an automated download (e.g., precise orbits and DEMs) and data management tasks like creation of logical folder structure is strongly recommended when using the Remote Sensing Graz tool. Besides this, comparing both approaches performances (JR Remote Sensing Graz tool versus SNAP), processing time from SLC to TIF-coherence output is similar to coherence workflow with RSG-tool. The workflow for coherence calculation from a pair of SLC-images using SNAP is shown in Figure 4-4. The scripts for batch processing are executed within the SNAP Command-Line.

Coherence is used as input for time features extraction in tests related to grassland (Task 3 WP33). Although a prototype for the grassland intensive/extensive management is not envisaged in ECoLaSS, the testing is intended to explore the foundations of a high-resolution layer on the basis of a continuous/ordinal variable (e.g., use intensity percentage range).

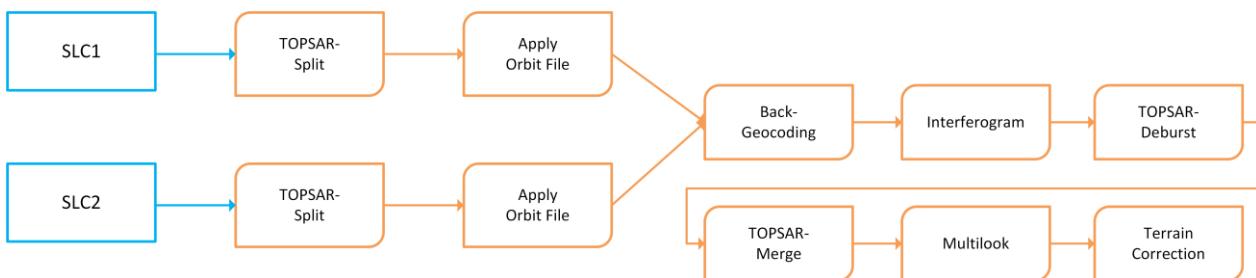


Figure 4-4: SNAP workflow for coherence calculation.

4.2.1.2 South-East (Greece/Bulgaria)

In the **South-East (Greece/Bulgaria)** site pre-processing has been performed using the SNAP toolbox and internal tools. The image ingestion part is handled by the sentinelsat utility to automate the download of Sentinel-1 Level-1 Single Look Complex (SLC) from the Copernicus Hub. The Snap toolbox processing line comprises the following steps : (1) Orbit file applied with polynomial degree 3 and Calibration (2) Removal of Thermal Noise (3) TOPSAR Deburst (4) Subsetting to the extent of the AOI (5) Multi-look of 5 range and 1 azimuth (6) Terrain Flattening and Correction based on a bicubic (Flattening) and bilinear (Correction) interpolation in the UTM projection of the AOI and using as DEM the SRTM 1Sec HGT (7) Combining polarization in one image stack (band 1: VH, band2: VV, band3: RATIOVVVH). The internal tool implementing a multi-temporal filter is applied on an annual image stack of gamma backscattered SAR data in DN processing unit.

4.2.1.3 West (Belgium)

In the **West (Belgium)** site pre-processing has been performed with the RSG module “Space Suite”. It comprises the following processing steps: (a) Image ingestion: bulk import of original images to RSG .rsx files, orbit update (precise orbits), automated combination of adjacent scenes of same acquisition date and orbit (b) Image pre-processing: definition of image frame extent (based on selected granules), full image resolution, no speckle filtering, no multitemporal filtering, radiometric terrain correction to gamma naught based on SRTM 4.1 model, combine polarizations in one image stack (band1: VH; band2: VV), (c) Resulting in 10m resolution images resampled with nearest neighbour in UTM WGS84. After the orthorectification of the images the multitemporal filtering has been applied, (d) calculation of incidence angle map.

Within the Sentinel-1 processing chain a conversion of db to DN is recommended which allows reducing the file size by 50% and facilitates multitemporal metric calculation from SAR image stacks. Furthermore, it is recommended using flattened gamma naught SAR backscatter data if different orbits are combined and to eliminate areas with large incidence angles. The SRTM4 model is not accurate enough for a full removal of terrain effects; however, it is freely available at pan-European scale. The benchmarking analysis concludes that multitemporal SAR filters should be applied to pre-processed gamma naught data in DN processing unit. Multitemporal filters can be applied to the entire (or annual) image stack. Seasonal means and seasonal statistics further reduce speckle noise and should be used as input regarding SAR data classification of the HRL land cover classes.

4.2.1.4 South-West (France)

In the **South-West (France)** the pre-processing has been performed using internal tools, based in particular on SNAP toolbox. It comprises the following processing steps: (a) Image ingestion: download from Copernicus Sci-Hub (b) Orbit file applied with polynomial degree 3 (c) Removal of GRD border noise and thermal noise, (d) Terrain correction with SRTM 1 arcsec DEM (e) Ortho-rectification with a pixel size of 10m (f) Output files encoded as int 16-bits, on 2 bands for VH and VV.

The main use of those datasets lays in the creation of a classification for the IMP status layer over this demonstration site. In order to limit the alteration of each image and thus the urban structures present on them, no further pre-processing has been applied beyond the geometric ortho-rectification to respect the geometry – hence no speckle filtering applied (Basly, Cauneau, Ranchin, & Wald, 2000).

The naming convention, which can be generically described by GAMMA_DN_SENTINEL1_YYYY-MM-DD_XXX.tif, is explained in Table 4-2.

Table 4-7 - File name filed convention for the MAJA datasets

File name section	Meaning
<i>Gamma</i>	<i>SAR backscatter processing level</i>
<i>DN or DZ</i>	<i>digital number/decibel</i>
<i>SENTINEL1</i>	<i>satellite</i>
<i>YYYY-MMM-DD</i>	<i>Date of recording</i>
<i>_XXX.tif</i>	<i>Processing number</i>

4.2.1.5 Northern (Sweden)

In the **Northern (Sweden) site** pre-processing chain was built entirely around the freely available software tools distributed by ESA. Sentinel-1 processing was performed by means of the ESA SNAP Sentinel-1 Toolbox. Sentinel-1 GRDH data were downloaded from the Copernicus Open Access Hub and processed to sigma0 backscatter coefficients. The decision to process sigma0 instead of gamma0 was taken due to the comparatively flat terrain. Processing involved the usual pre-processing steps of

- precise orbit estimation
- thermal noise suppression
- border noise suppression
- range-doppler terrain correction to 10m spatial resolution

The so processed data were then filtered by a multitemporal Lee Sigma filter, followed by a 7x7 window Frost filter. The separate processing of filtering in the multitemporal and spatial domain had become necessary due to a software bug in the Sentinel-1 toolbox.

5 Conclusions and Outlook

The main work of WP32 is focused in phase 1 of the ECoLaSS project, as it is a pre-requisite for work packages 33, 34 and 35, as well as for all work packages of Task 4. Comprehensive testing of optical as well as SAR pre-processing methods which are relevant within the ECoLaSS project are therefore undertaken in phase 1 and pre-operational processing lines which allow pre-processing of Sentinel-1 and Sentinel-2 data streams were implemented in phase 1. In phase 2, optimisation and fine-tuning of the pre-operational processing lines has been performed. Additionally, to phase 1 for Sentinel-1 six days short term coherence tests and for Sentinel-2 noise reduction and gap filling tests have been performed in phase 2.

Regarding the pre-processing of Sentinel-2 optical data streams the main software packages have been compared (MAJA, Sen2Cor, MACCS and FORCE). In general, all software packages have advantages and drawbacks, not performing sufficiently enough to achieve a high thematic quality, content and accuracy in any subsequent information extraction process.

Concerning the atmospheric correction the inter-orbit spectral stability and the consistency with other sensors show that Sen2Cor is able to achieve consistent results only if the data sets fulfil certain characteristics, e.g. presence of DDV pixels. Doxani et al., 2018 compared within the ACIX initiative the surface reflectance products from several different processors (CorA, FORCE, iCOR, LaSRC, MACCS, S2-AC2020, GFZ-AC and Sen2Cor). The overall analysis shows that FORCE, LaSRC, MACCS and Sen2Cor managed to estimate the reflectance quite well in comparison to AERONET corrected data.

Furthermore, the cloud masks have been benchmarked in chapter 3.2.2, Sen2Cor presents a generally reasonable detection of clouds and cloud shadows, but the detection is unsatisfactory due to the fact that low altitude clouds can be undetectable in the cirrus band and bright land surfaces especially build up structures are often misclassified as clouds. With the multitemporal approach, MAJA is able to avoid commission errors in urban areas from build-up areas. Frantz et al., 2018 introduced the Cloud Displacement Index (CDI) which makes use of the three highly correlated near infrared bands that are observed with different view angles. Therefore, elevated objects like clouds are observed under a parallax and can be reliably separated from bright ground objects. The CDI is integrated in the Fmask algorithm which is incorporated into the FORCE processor.

The topographic pre-processing step is recommended in hilly to mountainous terrain. The main issue with the topographic normalisation in Sen2Cor is the over-correction of south-facing slopes. The DEM with the highest spatial resolution should be used. It is therefore recommended using topographic normalisation tools which allow the user to define their own DEM. If only the SRTM 90m model is available Sen2Cor topographic normalisation also provides acceptable results.

The benchmarking of the noise reduction and gap filling showed that the RMSE values are quite similar regardless of the model order and that there is no general tendency showing that higher-order models achieve a lower RMSE. Further it can be noticed that the forecast quality is reduced in spring. Due to the deterministic nature of the regression model, shifts in the timing of the green-up can result in larger errors and higher-order models are even more affected. The highest RMSE values occur during the vegetation green-up in April and May, where also the phenological dynamics are highest. Both the steep increase of NIR reflectance as well shifts in the timing of the green-up are problematic and forecasts based on regression models can be subject to large errors. The other seasons beside spring are represented more accurately.

The Sen2Cor software implementation and handling are user-friendly with a GUI or as command line and it is open for public usage. MAJA might achieve better cloud masking results, but due to the licence and implementation issues it needs to be clarified if the software can be used for commercial purposes.

Force is a free command line software intended for mass processing. The hardware requirements are closely tied to the data volume therefore it is recommended to use multi-CPU server systems. It can be redistributed and/or modified under the terms of the GNU General Public License (Frantz, 2017).

Fusion of Sentinel-2 and PROBA-V time-series are being performed in phase 2, because Sentinel-3 satellite imagery products were not mature enough. The inclusion of PROBA-V allows for achieving the required temporal density of the optical Sentinel-2 time-series. NDVI anomalies obtained from the original SPOT Vegetation/PROBA-V time-series are compared to NDVI anomalies obtained from the PROBA-V-like NDVI time-series from Sentinel-2 data. The corresponding anomalies detection will be then be further investigated at 10m resolution to highlight the upscaling capacity of Sentinel-2 time series to document the detected anomaly. This comparison will be implemented for some tiles of Sentinel-2 covering part of ECoLaSS demo sites, in particular for the phenological layers produced in the framework of the WP41.

Concerning Sentinel-1 data sets a conversion of db to DN is recommended. The standard radiometric unit of Sentinel-1 GRD data is in db. The conversion from db to DN reduces the file size by 50% and facilitates multi-temporal metric calculation from SAR image stacks. The DN based land cover classification results are slightly more accurate than db based classification results. Furthermore, we recommend using flattened gamma nought SAR backscatter data for the ECoLaSS demo site processing if data from different orbits is combined, but to eliminate areas with large incidence angles above 70°. Flattened gamma nought imagery shows best results for terrain correction and for the similarity of imagery from different orbits. When incidence angles are too steep, values are still unreliable. With sigma nought processing, imagery from different orbits cannot be used in hilly to mountainous terrain jointly because backscatter values still differ strongly among the different orbits. The SRTM4 model is not accurate enough for a full removal of terrain effects; however, it is the only freely available DEM currently available at pan-European scale. For ECoLaSS pre-processing, our analysis concludes that multi-temporal SAR filters should be applied to pre-processed gamma nought data in DN processing unit. Multi-temporal filters can be applied to the entire (or annual) image stack. Seasonal means and seasonal statistics further reduce speckle noise and should be used as input regarding SAR data classification of the HRL land cover classes. Which seasonal statistics provide reliable features for classification depends on the thematic application and is therefore assessed in detail in WP33. Further the value of coherence data adds for HR land cover classification compared to using only filtered GRD backscatter data is investigated further in the thematic classification chains in WP33. By taking into account that huge processing resources are required for large area applications, we recommend to include coherences only in case of significant improvement of the classification accuracies.

It is expected that pre-processing software will be significantly improved in the near future, e.g. it is expected that atmospheric correction software such as Sen2Cor will be further developed and main issues such as related to inaccurate cloud masks are considerably improved. Further improvements are expected when ESA provides the Sentinel-2 time-series geometrically co-registered to the global reference image. In phase 2 of ECoLaSS we are closely following these developments and feedback-loops between WP32 and WP33, WP34, WP35 together with the demo site implementations carried out in phase 1 and 2 have contributed to guarantee continuous adaptations and improvements of the pre-processing chains to fulfil the requirements of the land cover classification and change detection workflows. In this sense different software packages are used in different demonstration sites depending on the thematic application, the site and the processing environments.

References

- Ahmed, S. M., Eldin, F. A. E., and Tarek, A. M., 2010. Speckle noise reduction in SAR images using adaptive morphological filter. In Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on (pp. 260-265). IEEE.
- Ariza, A., Ressel, C., Pfeifer, N., 2016. Evaluation of the Terrain Model Influence on the Orthorectification of Sentinel-2 Satellite Images Over Different Land Forms in Austria, in: 1st Sentinel-2 Validation Team Meeting.
- Baetens, L., Camille, D., & Hagolle, O. 2019. Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure. *Remote Sensing*, 11(433), p. 1-25.
- Balthazar, V., Vanacker, V., Lambin, E. F., 2012. Evaluation and parameterization of ATCOR3 topographic correction method for forest cover mapping in mountain areas. *International Journal of Applied Earth Observation and Geoinformation* 18, p. 436–450.
- Basly, L., Cauneau, F., Ranchin, T., & Wald, L. 2000. SAR Imagery in Urban Area. *EARSeL Symposium 1999 “remote sensing in the 21st Century: economic and environmental applications*, p. 563-568. Valladolid.
- Beaton, A.E., Tukey, J.W., 1974. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics* 16, 147–185. doi:10.1080/00401706.1974.10489171
- Bouchemakh, L., Smara, Y., Benali, M., & Cheikh, Z. B., 2006. Implementation of new methods of speckle noise reduction in SAR images. In Global Developments in Environmental Earth Observation from Space: Proceedings of the 25th Annual Symposium of the European Association of Remote Sensing Laboratories, Porto, Portugal, June 6-11, 2005 (p. 52). Ios Press Inc.
- 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 Trans. Geosci. Remote Sens.* 52, 3316–3332. doi:10.1109/TGRS.2013.2272545
- Chen, X., Mithal, V., Vangala, S. R., Brugere, I., Boriah, S., Kumar, V., 2011. A study of time series noise reduction techniques in the context of land cover change detection. Department of Computer Science and Engineering-Technical Reports, no. TR 11-016.
- Dechoz, C., Poulain, V., Massera, S., Languille, F., Greslou, D., de Lussy, F., Gaudel, A., L'Helguen, C., Picard, C., Trémas, T., 2015. Sentinel-2 global reference image. Presented at the Proc. SPIE, p. 96430A.
- Delwart, S., Bourg, L., 2004. Radiometric calibration of MERIS. Presented at the Sensors, Systems, and Next-Generation Satellites VIII, International Society for Optics and Photonics.
- Donadieu, J., L'Helguen, C., 2016. Sentinel-2A Level-2 products description (Technical Note No. 00), PSC-NT-411-0362-CNES. CNES.
- Doxani, G., Vermote, E., Roger, J.-C., Gascon, F., Adriaensen, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G.;Li, F., Louis, J., Mangin, A., Pahlevan, N., Pfugl, B., Vanhellemont, Q. 2018. Atmospheric Correction Inter-Comparison Exercise. *Remote Sensing* 10, 352.
- Fletcher, K., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services (European Spatial Agency SP-1322/2).
- Flood, N., 2013. Seasonal composite Landsat TM/ETM+ images using the medoid (a multi-dimensional median). *Remote Sensing*, 5(12), 6481-6500.
- Fougne, B., Desjardins, C., Besson, B., Bruniquel, V., Meskini, N., Nieke, J., Bouvet, M., 2016. Results from the radiometric validation of Sentinel-3 optical sensors using natural targets. Presented at the Earth Observing Systems XXI, International Society for Optics and Photonics, p. 997200.
- Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J. 2018. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effect. *Remote Sensing of Environment*, 215, 471-481.
- Frantz, D., Röder, A., Stellmes, M., & Hill, J. 2016. An operational radiometric Landsat preprocessing framework for large-area time series applications. *IEEE Transactions on Geoscience and Remote Sensing*, 54(7), 3928-3943.

- Frost, V., Stiles, J., Shanmugam, K. S., Holtzman, J., Smith, S., 1981. An adaptive filter for smoothing noisy radar images. *Proceedings of the IEEE* 69, p. 133–135.
- Gallaun, H., Schardt, M., Linser, S., 2007. Remote sensing based forest map of austria and derived environmental indicators. *Proceedings of ForestSat*.
- Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., 2017. Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing* 9, 584.
- Gopinathan, S., and Poornima, S., 2015. Enhancement of images with speckle noise reduction using different filters. *International Journal of Applied Science and Engineering Research*, 4(3), 333-352.
- Gorroño, J., Fomferra, N., Peters, M., Gascon, F., Underwood, C. I., Fox, N. P., Kirches, G., Brockmann, C., 2017. A Radiometric Uncertainty Tool for the Sentinel-2 Mission. *Remote Sensing* 9, p. 178.
- Hagolle, O., Huc, M., Pascual, D. V., Dedieu, G., 2010. A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENµS, LANDSAT and Sentinel-2 images. *Remote Sensing of Environment* 114, p. 1747–1755.
- Hagolle, O., Huc, M., Desjardins, C., Auer, S., & Richter, R. 2017. *MAJA ATBD Algorithm Theoretical Basis Document*.
- Hamilton, J. D., 1994. Time series analysis. Princeton University Press Princeton.
- Hantson, S., Chuvieco, E., 2011. Evaluation of different topographic correction methods for Landsat imagery. *International Journal of Applied Earth Observation and Geoinformation* 13, p. 691–700.
- Harvey, A. C., 1990. Forecasting, structural time series models and the Kalman filter. Cambridge -- University Press.
- Heiberger, R.M., Becker, R.A., 1992. Design of an S function for robust regression using iteratively reweighted least squares. *J. Comput. Graph. Stat.* 1, p.181–196.
doi:10.1080/10618600.1992.10474580
- Hird, J. N., & McDermid, G. J., 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment*, 113(1), 248-258.
- Hollstein, A., Segl, K., Guanter, L., Brell, M., Enesco, M., 2016. Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. *Remote Sensing* 8, p. 666.
- Home, O.D.H.N., List, C.D., NPP, N.P.P., Questions, D.P., Curation, D., Visualizer, S. M., Checker, L.-W., 2013. LEDAPS calibration, reflectance, atmospheric correction preprocessing code, version 2.
- Huber, P.J., 1964. Robust estimation of a location parameter. *Ann. Math. Stat.* 35, 73–101.
doi:10.1214/aoms/117703732
- Hurvich, C. M., Tsai, C.-L., 1989. Regression and time series model selection in small samples. *Biometrika* 76, p. 297–307.
- Inglada, J., Giros, A., 2004. On the possibility of automatic multisensor image registration. *IEEE Transactions on Geoscience and Remote Sensing* 42, p. 2104–2120.
- Irish, R. R., Barker, J. L., Goward, S.N., Arvidson, T., 2006b. Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm. *Photogrammetric Engineering & Remote Sensing* 72, 1179–1188.
- Jin, S., Homer, C., Yang, L., Xian, G., Fry, J., Danielson, P., Townsend, P. A., 2013. Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA. *International Journal of Remote Sensing* 34, 1540–1560.
- Jönsson, P., & Eklundh, L., 2004. TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30(8), 833-845.
- Kääb, A., Winsvold, S.H., Altena, B., Nuth, C., Nagler, T., Wuite, J., 2016. Glacier Remote Sensing Using Sentinel-2. Part I: Radiometric and Geometric Performance, and Application to Ice Velocity. *Remote Sensing* 8, p. 598.
- 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(6), 4055-4071.

- Kane, V. R., Gillespie, A. R., McGaughey, R., Lutz, J. A., Ceder, K., Franklin, J. F., 2008. Interpretation and topographic compensation of conifer canopy self-shadowing. *Remote Sensing of Environment* 112, p. 3820–3832.
- Karlsson, K.-G., Johansson, E., 2014. Multi-Sensor calibration studies of AVHRR-heritage channel radiances using the simultaneous nadir observation approach. *Remote Sensing* 6, p. 1845–1862.
- Kuan, D. T., Sawchuk, A. A., Strand, T. C., Chavel, P., 1985. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE transactions on pattern analysis and machine intelligence* p. 165–177.
- Languille, F., Déchoz, C., Gaudel, A., Greslou, D., De Lussy, F., Trémas, T., & Poulain, V., 2015. Sentinel-2 geometric image quality commissioning: first results. In *Image and Signal Processing for Remote Sensing XXI* (Vol. 9643, p. 964306). International Society for Optics and Photonics.
- Lee, J.-S., 1980. Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, p. 165–168.
- Lee, J.-S., 1981. Refined filtering of image noise using local statistics. *Computer graphics and image processing* 15, p. 380–389.
- Lee, J.-S., Jurkevich, L., Dewaele, P., Wambacq, P., Oosterlinck, A., 1994. Speckle filtering of synthetic aperture radar images: A review. *Remote Sensing Reviews* 8, p. 313–340.
- Li, S., Ganguly, S., Dungan, J. L., Wang, W., Nemani, R. R., 2017. Sentinel-2 MSI radiometric characterization and cross-calibration with Landsat-8 OLI. *Adv. Remote Sens* 6, p. 147.
- Lonjou, V., Desjardins, C., Hagolle, O., Petrucci, B., Tremas, T., Dejus, M., Makarau, A., Auer, S., 2016. MACCS-ATCOR joint algorithm (MAJA). Presented at the *Remote Sensing of Clouds and the Atmosphere XXI*, International Society for Optics and Photonics, p. 1000107.
- Lopes, A., Nezry, E., Touzi, R., Laur, H., 1990. Maximum a posteriori speckle filtering and first order texture models in SAR images. Presented at the *Geoscience and Remote Sensing Symposium, 1990. IGARSS'90.'Remote Sensing Science for the Nineties'*, 10th Annual International, IEEE, p. 2409–2412.
- Louis, J., Debaecker, V., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., Pflug, Bringfried, 2016a. Sentinel-2 Sen2Cor: Level-2 Processor for Users. Presented at the *Proceedings Living Planet Symposium 2016a*, Spacebooks Online, p. 1–8.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016b. Sentinel-2 Level-2 Processor Sen2Cor.
- Main-Knorn, M., Pflug, B., Debaecker, V., Louis, J., 2015. Calibration and validation plan for the L2a processor and products of the Sentinel-2 mission. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40, p.1249.
- Meer, P., Park, R.-H., Cho, K., 1994. Multiresolution adaptive image smoothing. *CVGIP: Graphical Models and Image Processing* 56, 140–148.
- Miletich, P., 2017. European large area automatic forest mapping based on random sampling points using Sentinel-2A. Master's Thesis, Graz University of Technology and JR.
- Montgomery, D. C., Jennings, C. L., Kulahci, M., 2015. *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Moreno, Á., García-Haro, F.J., Martínez, B., Gilabert, M. A., 2014. Noise reduction and gap filling of fapar time series using an adapted local regression filter. *Remote Sensing* 6, 8238–8260.
- Müller-Wilm, U., Louis, J., Richter, R., Gascon, F., Nieuwpoort, M., 2013. Sentinel-2 level 2A prototype processor: Architecture, algorithms and first results. Presented at the *Proceedings of the 2013 ESA Living Planet Symposium*, Edinburgh, UK, p. 9–13.
- Nieke, J., Mavrocordatos, C., Donlon, C., Berruti, B., Garnier, T., Riti, J., Delclaud, Y., 2015. Ocean and Land Color Imager on Sentinel-3. *Optical Payloads for Space Missions* 223.
- Pandžić, M., Mihajlović, D., Pandžić, J., Pfeifer, N., 2016. Assessment Of The Geometric Quality Of Sentinel-2 Data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 41.

- Park, S.-H., Jung, H.-S., Choi, J., Jeon, S., 2017. A quantitative method to evaluate the performance of topographic correction models used to improve land cover identification. *Advances in Space Research* 60, p. 1488–1503.
- Pepe, A., Yang, Y., Manzo, M., & Lanari, R. (2015). Improved EMCF-SBAS processing chain based on advanced techniques for the noise-filtering and selection of small baseline multi-look DInSAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 53(8), p. 4394-4417.
- Petrucci, B., Huc, M., Feuvrier, T., Ruffel, C., Hagolle, O., Lonjou, V., Desjardins, C., 2015. MACCS: Multi-Mission Atmospheric Correction and Cloud Screening tool for high-frequency revisit data processing. Presented at the Image and Signal Processing for Remote Sensing XXI, International Society for Optics and Photonics, p. 964307.
- Pflug, B., Main-Knorn, M., Bieniarz, J., Debaecker, V., Louis, J., 2016. Early Validation of Sentinel-2 Level-2 Processor and Products. Presented at the Proceedings of ‘Living Planet Symposium 2016,’ Spacebooks Online, p. 1–6.
- Posch, L., 2017. Untersuchung der Trennbarkeit alpiner Waldtypen unter Berücksichtigung der topographischen Normalisierung. Master’s Thesis, Graz University of Technology and JR.
- Quegan, S., Le Toan, T., Yu, J. J., Ribbes, F., Flouri, N., 2000. Multitemporal ERS SAR analysis applied to forest mapping. *IEEE Transactions on Geoscience and Remote Sensing* 38, p. 741–753.
- Richter, R., Schläpfer, D., 2011. Atmospheric/topographic correction for airborne imagery. ATCOR-4 user guide.
- Richter, R., Schläpfer, D., Müller, A., 2006. An automatic atmospheric correction algorithm for visible/NIR imagery. *International Journal of Remote Sensing* 27, p. 2077–2085.
- Richter, R., Wang, X., Bachmann, M., Schläpfer, D., 2011. Correction of cirrus effects in Sentinel-2 type of imagery. *International Journal of Remote Sensing* 32, p. 2931–2941.
- Rouquié, B., Hagolle, O., Bréon, F.-M., Boucher, O., Desjardins, C., Rémy, S., 2017. Using Copernicus Atmosphere Monitoring Service Products to Constrain the Aerosol Type in the Atmospheric Correction Processor MAJA. *Remote Sensing* 9, p. 1230.
- Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., Zhang, C., 2010. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, 114(1), 35-49.
- Santer, R., 2006. Calibration Validation Requirements Consolidation.
- Schubert, A., Miranda, N., Geudtner, D., Small, D., 2017. Sentinel-1A/B Combined Product Geolocation Accuracy. *Remote Sensing* 9, p. 607.
- Singh, M., Mishra, V., Thakur, N., Sharma, J. D., 2015. Expansion of empirical-statistical based topographic correction algorithm for reflectance modeling on Himalayan terrain using AWIFS and MODIS sensor. *Journal of the Indian Society of Remote Sensing* 43, p. 379–393.
- Sola, I., González-Audicana, M., Álvarez-Mozos, J., 2016. Multi-criteria evaluation of topographic correction methods. *Remote Sensing of Environment* 184, p. 247–262.
- Spaans, K., Hooper, A., 2016. InSAR processing for volcano monitoring and other near-real time applications. *Journal of Geophysical Research: Solid Earth* 121, p. 2947–2960.
- Tamm, T., Zalite, K., Voormansik, K., Talgre, L., 2016. Relating Sentinel-1 Interferometric Coherence to Mowing Events on Grasslands. *Remote Sensing* 8, p. 802.
- Teillet, P., Guindon, B., Goodenough, D., 1982. On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing* 8, p. 84–106.
- Touzi, R., Lopes, A., Bruniquel, J., Vachon, P.W., 1999. Coherence estimation for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing* 37, p. 135–149.
- Udroiu, C., Grosu, A., Popescu, R., Nicola, L., Manda, R., Savinaud, M., Bellemans, N., Bontemps, S., 2017. Sentinel-2 Agriculture Software User Manual.
- Verbesselt, J., Herold, M., Zeileis, A., 2012. Near real-time disturbance detection using satellite image time series. *Remote Sens. Environ.* 123, 98–108. doi:10.1016/j.rse.2012.02.022
- Verbesselt, J., Hyndman, R., Newham, G., Culvenor, D., 2010. Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment* 114, p. 106–115.

- Vincini, M., Frazzi, E., 2003. Multitemporal evaluation of topographic normalization methods on deciduous forest TM data. *IEEE Transactions on Geoscience and Remote Sensing* 41, p. 2586–2590.
- Warner, T. A., Foody, G.M., Nellis, M. D., 2009. *The SAGE handbook of remote sensing*. Sage Publications.
- Wulder, M., Franklin, S. E., 2012. *Remote sensing of forest environments: concepts and case studies*. Springer Science & Business Media.
- Zalite, K., Voormansik, K., Praks, J., Antropov, O., Noorma, M., 2014. Towards detecting mowing of agricultural grasslands from multi-temporal COSMO-SkyMed data. Presented at the Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International, IEEE, p. 5076–5079.
- Zebker, H.A., Villasenor, J., 1992. Decorrelation in interferometric radar echoes. *IEEE Transactions On Geoscience and Remote Sensing* 30, 950–959.
- Zhu, Z., Wang, S., Woodcock, C. E., 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel-2 images. *Remote Sensing of Environment* 159, p. 269–277.
- Zhu, Z., Woodcock, C. E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment* 118, p. 83–94.
- Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* 144, 152–171. doi:10.1016/j.rse.2014.01.011

Internet sources

- CNES (2018) Logiciels Gratuits: <https://logiciels.cnes.fr/en/license/57/144ESA>
- CNES, (2018). MAJA | Logiciels Gratuits. URL <https://logiciels.cnes.fr/en/license/57/144> (accessed 25.01.18).
- ESA, 2019. Sentinel-2 Data Quality Report. Issue 34, ESA REF: S2-PDGS-MPC-DQR, URL: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_Level-1C_Data_Quality_Report (accessed 06.02.19).
- Frantz, D, 2017. FORCE – Technical User Guide v. 1.0. beta. URL: <https://www.uni-trier.de/fileadmin/fb6/prof/FER/Downloads/Software/FORCE/FORCE-user-guide.pdf> (accessed 12.04.19).
- GEOS, 2017): <https://eocalibration.wordpress.com/>, (accessed 23.01.18).
- MAJA processor, 2017. <https://logiciels.cnes.fr/en/content/maja>, (accessed 20.01.18).
- Sen2agri Software User Manual, (2017): <http://www.esa-sen2agri.org/wp-content/uploads/resources/technical-documents/Sen2-Agri-Software-User-Manual-2.4.pdf>, (accessed 25.01.18).
- Sentinel Technical Guides, 2017. <https://sentinel.esa.int/web/sentinel/sentinel-technical-guides>, (accessed 20.01.18).
- VITO, 2012. http://calvalportal.ceos.org/c/document_library/get_file?uuid=7aa5e7dd-ffd1-4400-8c70-f7c8c68dc168&groupId=10136, (accessed 23.01.18).