
Climate Change Initiative Extension (CCI+) Phase 2
New Essential Climate Variables (NEW ECVS)
High Resolution Land Cover ECV (HR_LandCover_cci)

Algorithm Development Plan

(ADP)

Prepared by:

Università degli Studi di Trento
Fondazione Bruno Kessler
Università degli Studi di Pavia
Università degli Studi di Genova
Université Catholique de Louvain
Politecnico di Milano
LSCE
CREAF
University of Exeter
e-GEOS s.p.a.
Planetek Italia



UNIVERSITY
OF TRENTO



FONDAZIONE
BRUNO KESSLER



UNIVERSITÀ DEGLI STUDI
DI GENOVA

UCLouvain
Earth and Life Institute - Geomatics



POLITECNICO
MILANO 1863



LSCE



CREAF



University
of Exeter



e-geos
AN ASI / TELESPAZIO COMPANY



planetek
italia

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	1	

Changelog

Issue	Changes	Date
1.0	First issue	02/08/2024
1.1	Updated version according to CCI_HR LC_Ph2_Milestone1_RID-ESA.xlsx	16/09/2024

Detailed Change Record

Issue	RID	Description of discrepancy	Sections	Change
1.0	ESA-02	The sentence "for future product versions as part of our ongoing strategy" should be replaced by "in Phase 2"	Section 1.1	The sentence has been updated.
1.0	ESA-03	The sentence "for best implementation of the" should be replaced by "for the best implementation / for best implementing the ..."	Section 1.2	The sentence has been updated.
1.0	ESA-04	The sentence "These maps were obtained by classifying time series data from Sentinel-2 L2A for the 2019 HRLC10 and HRLC30 maps" should be replaced by "These maps were obtained by classifying time series data from Sentinel-2 L2A for the 2019 HRLC10 map"	Section 2.1	The sentence has been changed for clarity as it was ambiguous.
1.0	ESA-05	The text of the caption should be modified as follow "Workflow for optical data processing chain to produce the 2019 HRLC10 and HRLC30 maps obtained by classifying the time series of Sentinel-2 L2A data, and the HRLC30 historical maps obtained by classifying the time series of Landsat L2 data for 1990, 1995, 2000, 2005, 2010 and 2015."	Section 2.1 – Figure 1	The sentence has been changed for clarity as it was ambiguous.
1.0	ESA-06	In the considered implementation of the processing chain, the team generates monthly, seasonal and annual composites to harmonize the data and <u>to remove the cloud coverage</u>	Section 2.1.1	"to remove cloud contamination" is correct. The sentence has been updated.

Contents

2	Introduction.....	4
2.1	Executive summary	4
2.2	Purpose and scope	4
2.3	Reference documents	4
2.4	Acronyms and abbreviations.....	4
3	Algorithm development Plan for Optical Data	6
3.1	Current Status of Optical data classification	6
3.1.1	Optical Data Pre-processing	6
3.1.2	Optical Feature Extraction.....	6
3.1.3	Optical Land Cover Classification	6
3.2	Development Plan for Optical Data classification	8
3.2.1	Optical Data Pre-processing	8
3.2.2	Optical Feature Extraction.....	9
3.2.3	Optical Land Cover Classification	9
4	Algorithm development plan for SAR data	9
4.1	Current status of SAR data classification	9
4.1.1	Training set extraction.....	10
4.1.3	Multitemporal despeckling filter	11
4.1.4	Features extraction	11
4.1.5	SAR Land Cover Classification.....	11
4.2	Development plan for SAR classification	12
4.2.1	SAR Land Cover Classification.....	12
4.2.2	SAR water extraction	12
5	Algorithm development plan for decision fusion	13
5.1	Current status of decision fusion	13
5.1.1	Multi-sensor geolocation	13
5.1.2	Optical-SAR multi-sensor and multi-temporal fusion	14
5.1.3	Spatial harmonisation	14
5.2	Development plan for decision fusion	14
5.2.1	Optical-SAR multi-sensor and multi-temporal fusion	15
5.2.2	Spatial harmonisation	15
6	Algorithm development plan for multitemporal change detection	15
6.1	Current status of multitemporal change detection	15
6.1.1	Time Series reconstruction.....	16
6.1.2	Abrupt change detection	16
6.2	Development plan for multitemporal change detection	17
6.2.1	Feature space design.....	17
6.2.2	Time series reconstruction and break point detector	17
7	References.....	19

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	3	

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	4	

1 Introduction

1.1 Executive summary

Building on the activities and insights from Phase 1, the algorithms will be refined and enhanced, or new algorithms and methodologies will be selected to ensure high-quality outputs. This document outlines the activities to identify the most reliable methodologies and the planned improvements in Phase 2.

1.2 Purpose and scope

This Algorithm Development Plan (ADP v1.0) provides the details on the expected evolutions to the current processing chain following the first production in phase 1. It includes planned developments to:

- The algorithms themselves.
- The necessary auxiliary data for the best implementation of the algorithms.
- The training strategy and implementation.

The algorithms under development in the next processing cycles will be those selected from Consortium inter-comparison exercise as selected from internal benchmarking/development activities. The evolutions outlined in this document will be implemented in the end-to-end system to generate the updated/new HRLandCover_cci climate data records in the next Cycle. It is also important to note that this document will be regularly updated according to the schedule, with possible intermediate notes in case of significant achievements are obtained in between.

1.3 Reference documents

Ref. Title, Issue/Rev, Date, ID

- [RD1] CCI_HRLC_Ph1-D2.1_PVASR, latest version
 [RD2] CCI_HRLC_Ph1-D2.2_ATDB, latest version
 [RD3] CCI_HRLC_Ph1-D1.1_URD, latest version
 [RD4] CCI_HRLC_Ph1-Modelers_TN

1.4 Acronyms and abbreviations

ASM	Angular Second Moment
ATDB	Algorithm Theoretical Basis Document
BOCPD	Bayesian Online Change Point Detection
BFAST	Breaks For Additive Season and Trend
CCI	Climate Change Initiative
CD	Change Detection
CPDNN	Change Point Detection based on Neural Networks
CNN	Convolutional Neural Network
DEM	Digital Elevation Model
DL	Deep Learning
ETM	Enhanced Thematic Mapper
FCN	Fully Convolutional Network
GLCM	Gray Level Co-occurrence Matrix
GRNN	General Regression Neural Network
GRU	Gated Recurrent Unit
HMM	Hidden Markov Models
HR	High Resolution

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	5	

HRLC10	High Resolution Land Cover product at 10m resolution
HRLC30	High Resolution Land Cover product at 30m resolution
HRLCC30	High Resolution Land Cover Change product at 30m resolution
ICA	Independent Component Analysis
LC	Land Cover
LCC	Land Cover Change
LPP	Locality Preserving Projections
LSTM	Long-Short-Term Memory
MOLCA	Map Of LC Agreement
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-Infrared
PCA	Principal Component Analysis
PCC	Post Classification Comparison
PCHIP	Piecewise Cubic Interpolation
PGM	Probabilistic Graphical Models
PVASR	Product Validation and Algorithm Selection Report
RBF	Radial Basis Function
RNN	Recurrent Neural Networks
RF	Random Forest
R-G-B	Reed-Green-Blue
SAR	Synthetic Aperture Radar
SITS	Satellite Image Time Series
SLC	Single Look Complex
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
TS	Time Series
UEXT	Urban EXTraction
URD	User Requirement Document
UMAP	Uniform Manifold Approximation and Projection

2 Algorithm development Plan for Optical Data

Based on the results from the three study areas of Phase 1, the team aims to explore potential improvements for each step of the optical processing chain. The following sections provide details on both the status of the optical processing chain and the algorithm development plan for Phase 2.

2.1 Current Status of Optical data classification

Figure 1 illustrates the optical data processing chain used to produce the Phase 1 HRLC maps. These maps were obtained by classifying time series data from Sentinel-2 L2A for 2019 LC maps, and from Landsat L2 for 1990, 1995, 2000, 2005, 2010, and 2015 LC maps. The processing chain involves three main steps:

1. Optical Data Pre-processing, which prepares the optical data for subsequent steps.
2. Optical Feature Extraction, which characterizes the contextual information of the pixels.
3. Optical SVM Classification, which produces the optical class-posterior probability maps.

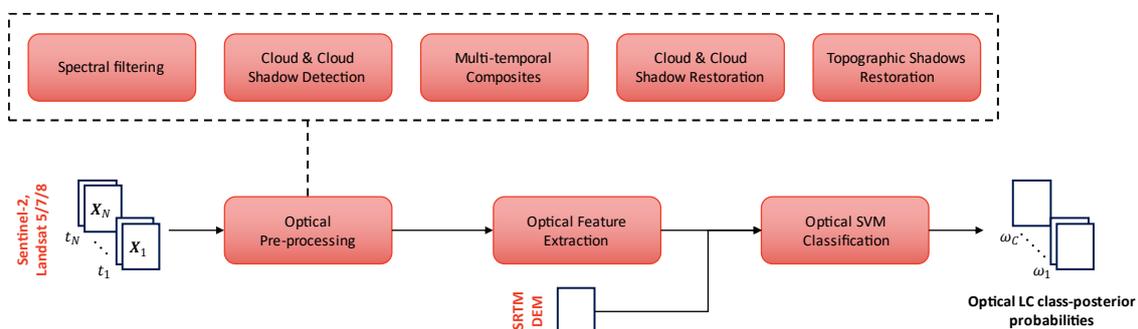


Figure 1. Workflow for optical data processing chain to produce the 2019 optical class-posterior probabilities obtained by classifying the time series of Sentinel-2 L2A data, and the 1990, 1995, 2000, 2005, 2010 and 2015 optical class-posterior probabilities obtained by classifying the time series of Landsat L2 data.

2.1.1 Optical Data Pre-processing

In the considered implementation of the processing chain, the team generates monthly, seasonal and annual composites to harmonize the data and to remove the cloud contamination, exploiting automatically generated cloud and cloud shadows masks. The considered temporal window changes depending on the data availability and quality. HRLC10 optical class-posterior probabilities are generated using 12 bimonthly Sentinel-2 composites (with overlap) to represent 2019. HRLC30 optical class-posterior probabilities are generated using seasonal composites, where Landsat L2 acquisitions of the considered season from both the target year, the previous and the subsequent years are considered for the composite generation. The Siberian area required special care due to poor data availability and quality, resulting in a single yearly composite focused on the summer season. When seasonal or bimonthly composites are considered, a standard linear temporal filling approach have been used to replace the pixels still associated to a cloud or a cloud shadow in the composites. For the seasonal composites of year 2005 and 2010 in Africa, a gap filling procedure is included to improve the spectral homogeneity of the composites when using Landsat-7 (whose scan-line corrector (SLC) failed in 2003, resulting in about 22% of the pixels per scene not being scanned) as the main source of optical acquisitions.

2.1.2 Optical Feature Extraction

In addition to the spectral bands, additional features are extracted and fed to the classifier to improve the mapping accuracy and include additional information:

- Topographic features: Altitude, extracted from a SRTM DEM raster;
- Textural features: Dissimilarity, Correlation, Contrast, Homogeneity, Energy, and Angular Second Moment, extracted from the first composite of the considered year.

2.1.3 Optical Land Cover Classification

For the automatic generation of Optical class-posterior probabilities, machine learning approaches have been adopted. The considered machine learning model is fed with the optical composites' spectral bands and the additional topographic and textural features. To train the classification model, a training database needed to be developed, such that to comprehensively represent the specific phenological patterns and spectral signatures of

the classification legend [RD2].

2.1.3.1 Training Set Production

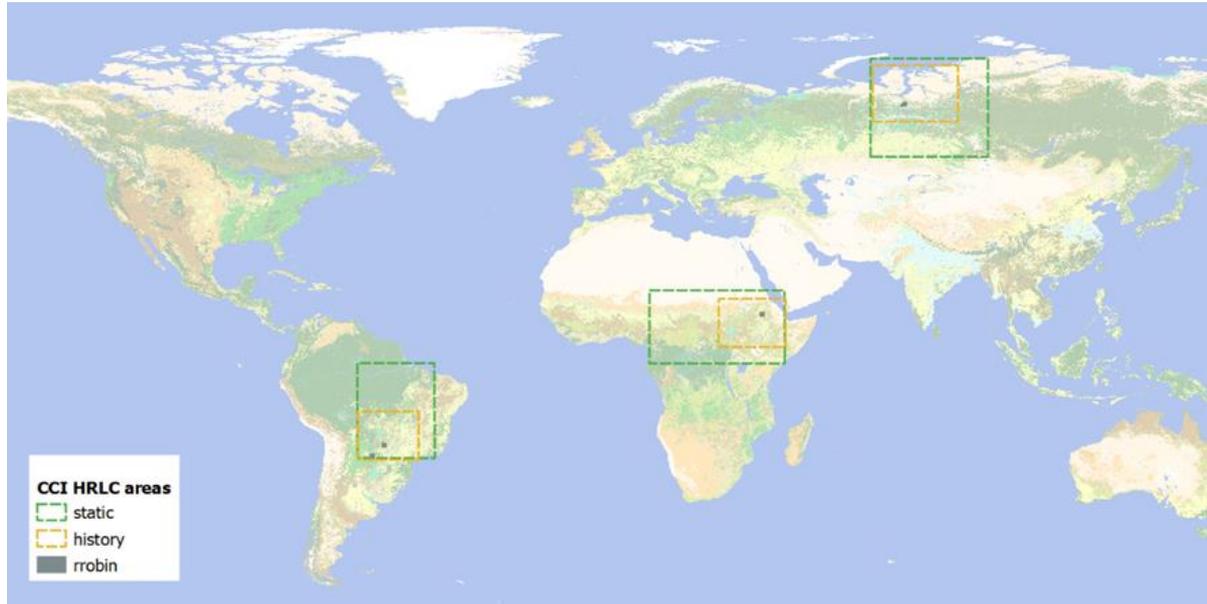


Figure 2. Considered study areas during Phase 1: Amazonia, Africa and Siberia. The HRLC10 static areas, where the training set has been extracted, are the larger green rectangles, whereas the HRLC30 historical areas are orange rectangles.

Due to the lack of available training data, the team has devoted significant effort to generating a training set for the three study areas. This effort has resulted in a high-quality photo-interpreted training set that aligns with the legend definitions and consistently represents the LC every 5 years from 1990 to 2019. Figure 3 shows a qualitative example of extraction of training set, while Figure 4 represents the number of tiles considered for the Amazonia study area.



Figure 3. Training Set Production conducted via photo-interpretation.

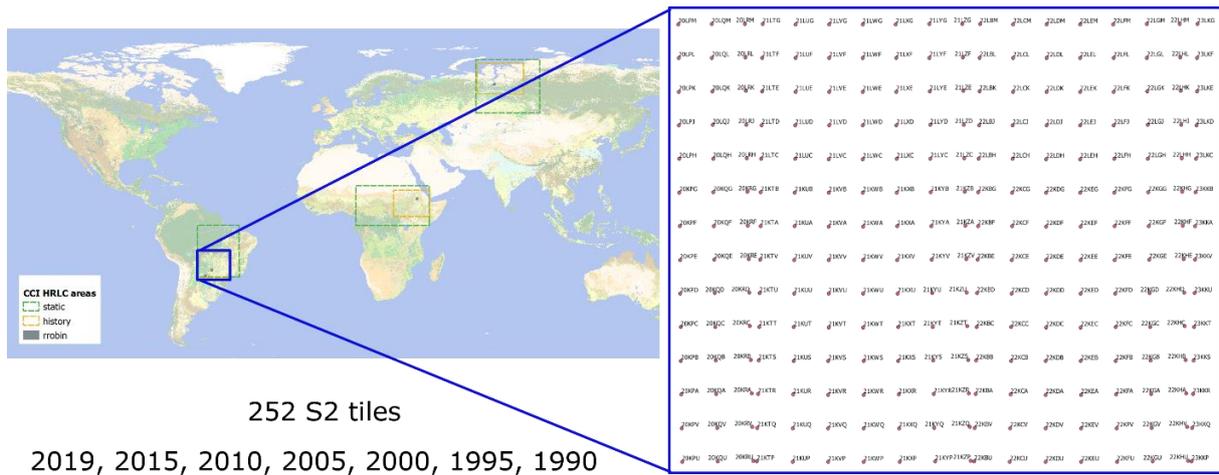


Figure 4. Example of number of tiles covered by photointerpretation in Amazonia.

While complex LC classes require reliable samples that cannot be extracted from the outdated coarse thematic products, existing thematic products represent a valid source of information for the other classes, allowing to significantly expand the training set and properly represent the whole areas to map. For this reason, only for the static map production, we integrated the training sets delivered through photointerpretation with samples extracted from the agreement of available LC products. Moreover, the increased amount of training labels unlocked the possibility of exploiting the specific properties of the local LC. This has been done by considering the global ecoregions, which are areas of water or land that contain characteristic assemblages of natural communities and species. By training a classifier for each ecoregion, we can exploit the fact that inside an ecoregion the probability of encountering different vegetation species (which may be mapped in the same class) and communities remains relatively constant. This feature is important in LC mapping as it allows to mitigate the intra-class variability, a well-known issue in remote sensing.

2.1.3.2 Classification Algorithms

Several well-known classifiers widely employed to generate LC maps have been tested and compared during Phase 1. From this analysis, the standard SVM classifier and the deep recurrent neural network LSTM classifier has shown the most promising performance. From further analysis provided by the validation team, the SVM classifier has shown to allow the production of more robust classification results. The composites' spectral bands and the topographic and textural features are stacked into a single feature vector and classified by means of an SVM using an RBF kernel. The generated class scores are then used to estimate class-posterior probabilities using the Isotonic Regression calibration strategy.

2.2 Development Plan for Optical Data classification

2.2.1 Optical Data Pre-processing

The pre-processing of Sentinel-2 and Landsat-5/7/8 involves several steps: i) cloud/cloud shadow detection, ii) atmospheric correction, iii) spectral filtering, and iv) composites generation.

In Phase 1, cloud and cloud shadow detection relied on the operational algorithms of FMASK [1] and Sen2Cor [2] for Landsat-5/7/8 and Sentinel-2, respectively, which were then improved with custom algorithms in the case of Sentinel-2. In Phase 2, additional algorithms [3] will be considered to further enhance the quality of the generated cloud mask, such as FORCE [4], MAJA [5] and s2cloudless [6]. For Sentinel-2, attention will be given to algorithms that exploit the parallax effect due to the time delay between the acquisition of different bands to improve cloud detection, especially for reducing the false detection of clouds on bright surfaces [7]. The most recent FMASK 4.0 [8] and Sen2Cor v2.11 already include the parallax test and will be the baseline for assessing possible improvements in this direction.

Regarding atmospheric correction, the activities in Phase 2 will be related to the analysis of alternative algorithms [9] and their topographic correction capabilities. Sentinel-2 Level 2 Surface Reflectance products by default include the topographic correction, whereas the corresponding Level 2 Landsat Surface Reflectance operational products lack this processing step, which was addressed in Phase 1 with the definition of an ad-hoc topographic correction algorithm for topographic shadow correction. The activities in Phase 2 will include the analysis of possible alternatives. Also, indirect approaches will be considered, where, instead of using topographic correction, slope and sun view angle will be provided to the classifier as features, thus allowing the classifier to

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	9	

implicitly be invariant to these illumination conditions. However, note that this requires for the stratification in the training set extraction to consider both illuminated and topographic shadow areas as strata.

The activities on spectral filtering for Phase 2 will be mainly focused on enhancing the performance of gap-filling operations in Landsat-7 ETM+ SLC off images [10].

Regarding the composite generation step, Phase 1 showed that it was the most demanding step in the optical pre-processing step. Therefore, the focus will be code optimization and efficiency, and alternative algorithms will also be considered to better handle the variable number of valid observations used to compute the composite, especially when the observations available are few.

2.2.2 Optical Feature Extraction

During Phase 1, additional features were considered for the training of the SVM classifiers: GLCM-based features such as contrast, dissimilarity, homogeneity, energy, correlation and ASM, and altitude. However, GLCM features computation have shown to be time demanding. Hence, the activities in Phase 2 will focus on improving the feature extraction efficiency and compare with alternative spatial features. In addition, other features will be considered, such as the slope and aspect derived from the digital elevation models and spectral indices.

2.2.3 Optical Land Cover Classification

For the classification step in the optical processing chain, the main challenges in Phase 1 were defined by i) the scarcity of available photo-interpreted data able to properly characterize the large areas that need to be mapped, ii) the considered input features, and iii) the optimization and efficiency of the considered classification algorithm. Given the complexity of the considered classification problem, the training of the classifiers can be performed in a completely supervised, a partially supervised (or semi-supervised) and an unsupervised framework. In Phase 2 of the project, attention will be given to weakly supervised learning [11], which stands in between complete supervision and partial supervision, and is based on the use of unreliable sources of training labels. In the context of the project, weak supervision can be used to leverage obsolete maps as an additional source of labels [12], [13]. In Phase 1, the training set was augmented using part of the maps intercomparison activities, which provided weak training labels where the available land cover maps agreed. While this was shown to be helpful, there is still room for improvement. Indeed, labels produced in this way tend to be biased towards “easy” samples, thus providing little help in points where existing maps disagree. Instead, weak supervision provides a framework where all the available labels (not only the map agreement) can be exploited, and the uncertainty of the label can be considered during training to guide its effect on the learning process.

Comprehensive analysis of different classification algorithms will be performed, addressing the new state of the art in the classification of satellite image time series led by Transformer deep learning models [14]. Focus will be given to strategies for properly handling intra-annual time series of composites, but also to multi-year classification for temporally consistent classifications [15]. The considered models will be compared both in terms of performance and inference time. Indeed, focus will be given to the optimization and efficiency of the model inference step, in order to allow faster generation of optical land cover maps.

3 Algorithm development plan for SAR data

This section describes the planned improvements to SAR land cover classification based on the results obtained and defined in Phase 1 of the CCI+ project in the three identified test areas, i.e. Amazon, Africa and Siberia.

3.1 Current status of SAR data classification

The CCI+ HRLC Phase 1 initiative evaluated and compared the performance of candidate SAR (Synthetic Aperture Radar) data classification algorithms. The main objective was to identify the most promising methods for implementation in the SAR image processing chain shown in Figure 5.

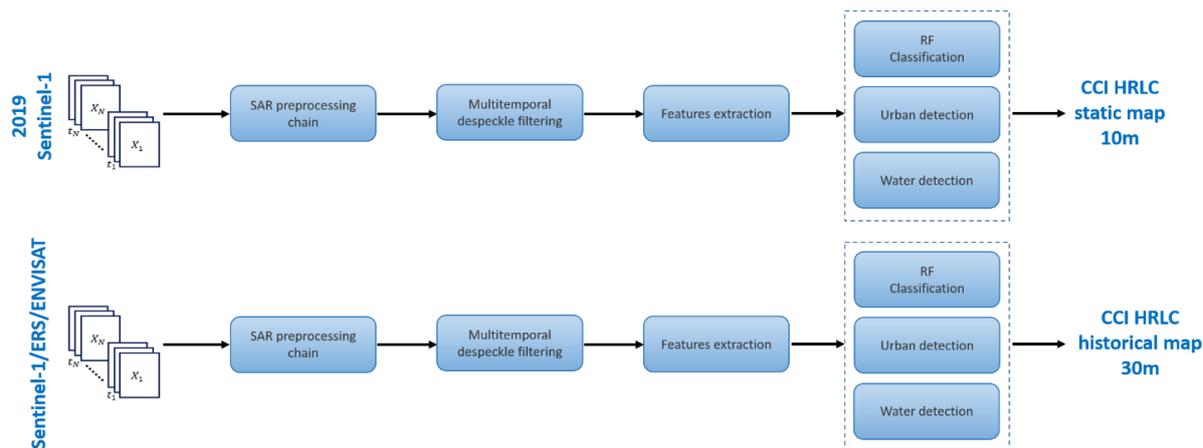


Figure 5. Workflow for the SAR data processing chain to produce the 2019 HRLC10 map using time series of Sentinel-1 data and the historical HRLC30 maps using time series of ERS or ENVISAT data for 1990, 1995, 2000, 2005, 2010 and 2015.

Initially, in the absence of a consistent and much more accurate set of reference samples, the first experiments were carried out on the four benchmark areas of the Round Robin using a training set extracted from the medium resolution products. Subsequently, in the second year of the CCI HRLC Phase 1 project, high-resolution training data sets were extracted. The high-resolution reference data are much more powerful than those extracted from the thematic products on a global scale. The method used to extract these new training points is based on a stratified approach and photo-interpretation of very high resolution imagery, as described in more detail in the PVASR [RD1] document. The main goal of the last year of the Phase 1 was to improve and release the final version of the processing chain designed for the high resolution and global land cover mapping. The novelty lies in the improvement of the classification task, to make the best use of the backscattered SAR signal to distinguish built-up settlements and water bodies. For this purpose, two dedicated routines in the Python language have been properly developed and tested. Accordingly, the Random Forest (RF) classifier was trained with training samples collected by the team using the hierarchical approach (see ATDB [RD2]), excluding the points belonging to the "built-up" and "water (permanent and seasonal)" classes.

The algorithms chosen for each step of chain in Figure 5 were:

- *Training set Selection*
 - Photointerpreted and hierarchically extracted reference data
- *Speckle Filtering*
 - Multitemporal despeckling filter [16]
- *Feature Extraction*
 - Single band analysis:
 - *Mean filter*
 - *Median filter.*
 - *Lee filter.*
 - *Minimum (maximum).*
 - Dual band analysis:
 - *Ratio, VV/VH;*
 - *Sum, VV+VH;*
 - *Mean, (VV+VH) / 2;*
 - *Difference, VV-VH.*
- *Classification*
 - RF classifier
 - UEXT algorithm (for built-up extent extraction)
 - Water permanent and water seasonal detection

These algorithms will be further improved over the Phase 2 cycles. Their current status is available in the ATBD deliverable of the Phase 1 [RD2].

3.1.1 Training set extraction

The training set extraction followed the hierarchical approach outlined in the ATBD document to achieve a consistent and accurate training set through photointerpretation of very high-resolution imagery. The final

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	11	

version of the land cover legend used is detailed in the URD document [RD3]. Specifically, the team focused on collecting training samples in Siberia, targeting both static and historical data production.

3.1.2 SAR pre-processing chain

SAR scenes go through a standard pre-processing chain using a workflow implemented in a graph of the SNAP toolbox [17]. These steps are: 1) the application of the orbit file to correct the satellite position and add velocity information; 2) border noise removal, to remove invalid data at the edges of the image; 3) radiometric calibration, to adjust the grey level of the SAR image according to the backscatter signals from the objects; 4) geometric terrain correction, to remove the distortion caused by the topographic variations and is performed by using a Digital Elevation Mode (DEM).

3.1.3 Multitemporal despeckling filter

The multitemporal denoising filter used to reduce speckle in SAR images is the one described in [16], which was applied to the four "seasonal" SAR clusters. The multitemporal approach seems to give better results than a spatial filter applied independently on each SAR image, thanks to the exploitation of the temporal sequence in favour of a better preservation of the spatial resolution. The filter is ratio-based and computes an image, called the *super image*, by exploiting the SAR time series. In fact, the temporal averaging of the SAR time series produces the super image, where speckle is reduced, and spatial resolution is preserved. The filtered image is thus recovered by exploiting the statistical properties associated with the original super image. Basically, the method consists of three steps: a) calculation of the super image by arithmetic averaging of SAR time series images; b) denoising of the ratio image; c) calculation of the final image given by the multiplication between the denoised ratio and the super image.

3.1.4 Features extraction

Radar images are a valuable source of spatial and temporal information, as they make it possible to study the nature of the scene observed and captured by the sensor, and to study in detail the physical and morphological characteristics of the terrain, also known as *features*. In this context, feature extractors allow the extrapolation of structural descriptors to improve the radar image classification task and optimise the class recognition. Spatio-temporal feature extraction is performed using the polarimetric information derived from the SAR intensities. Rather than considering complex spatial features such as shape and size, which would require unsupervised segmentation of the image, a set of texture features is computed for each of the four seasonal composites: *Lee filter*, *spatial median*, *mean*, *maximum*, *minimum* and *range* (maximum minus minimum).

3.1.5 SAR Land Cover Classification

SAR land cover mapping is performed using the seasonal set of features as input to the RF algorithm. However, the classification results obtained in the first two years of the Phase 1 revealed problems due to outliers in the detection of built-up and water (both permanent and seasonal) land cover classes. Based on a careful preliminary analysis and a study of the state of the art in the literature, two modules dedicated exclusively to the extraction of the two imputed classes, i.e. built-up and water, were developed. The maps obtained were then integrated with the resulting RF classification map. The built-up and water classifiers have been optimized to work mainly with radar sequences acquired by the Copernicus Sentinel-1 sensor, which provides high-resolution images at 10m per pixel. The built-up area recognition module uses the average image, called the super image, which is given by the time average of the SAR temporal series. From the super image, the built-up 'seeds' (i.e. source positions) and the cost of each pixel are then extracted to evaluate a cumulative cost map from which the built-up area recognition information is extracted. For water detection, this module employs features that track changes in the backscattering coefficient over time. Specifically, the temporal mean, minimum, maximum, and variance are computed from monthly SAR time series data. These metrics are then fed into the unsupervised *K-means* classifier to differentiate between water and non-water pixels. Next, morphological operations (closing and opening, which involve erosion and dilation) are applied to remove any outliers misclassified as water or built-up areas. This process produces monthly water masks (where water pixels are assigned a value of 1 and non-water pixels a value of 0) for the year of interest. To create the final water map, distinguishing between seasonal and permanent water classes, an additional step is performed. This step involves pixel-based analysis of the calculated monthly water masks: a pixel is labeled as 'permanent' if it remains water (value 1) for more than 6 months; otherwise, it is labeled as 'seasonal.' Finally, the maps given by the RF classifier, built-up and water detectors are merged to produce the SAR HRLC10 map.

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	12	

3.2 Development plan for SAR classification

Phase 1 successfully established a Synthetic Aperture Radar (SAR) processing framework, and the current phase seeks to enhance this chain using Deep Learning (DL) techniques on multitemporal SAR sequences for the classification task. This section describes the proposed improvements in the Phase 2 of the project for the SAR processing system, focusing on:

- SAR land cover classification.
- SAR Water extraction.

3.2.1 SAR Land Cover Classification

The current plan for SAR LC classification involves using a classification pipeline that employs a DL network applied to multitemporal SAR sequences. The proposed approach uses a DL architecture to classify SAR sequences by dividing them into seasonal subsequences and extracting spatial features, as done in Phase 1. These features are then processed using DL. The methodology, designed to work on spatial subsets, allows for comprehensive geographical coverage. It has been tested on Sentinel-1 SAR datasets from Amazon, Africa, and Siberia to evaluate its effectiveness across diverse environmental and climatic conditions.

A considered Sentinel-1 SAR sequence consists of multiple images from the same orbit, beam, and polarization. These images undergo several correction and refinement steps, as outlined in the pre-processing procedures used in Phase 1 of the CCI+ project, including orbit file application, thermal noise removal, border noise removal, radiometric calibration, and geometric terrain correction. After pre-processing, spatio-temporal SAR features are extracted for use in DL models. The SAR sequence is divided into four seasonal clusters to represent different LC types. Due to speckle distortion inherent in SAR imaging, a multitemporal despeckle filter is applied to mitigate this effect while preserving spatial resolution. Subsequently, 7 textural features (Lee filter, spatial median, mean, maximum, minimum and range, i.e. maximum minus minimum) are computed from each seasonal composite using a 5×5 pixel kernel, resulting in a total of 28 features for DL classification.

The LC information from the Map Of LC Agreement (MOLCA) [18] will be used to build the training set for the DL approach. MOLCA was generated using existing global High-Resolution LC (HRLC) maps, retaining only areas where all datasets agree on LC classes and discarding areas of disagreement (these pixels are labeled 'no-data'). The MOLCA images, arranged according to the tiling grid of the Sentinel-2 Level-1C product and distributed in GeoTIFF format, cover about 19 million square kilometers in three regions (Amazonia, Africa, and Siberia) with approximately 117 billion pixels at 10m resolution.

Three DL-based systems (*Attention Unet* [19], *Swin-Unet* [20], or 3D-FCN [21]) are currently under evaluation, and their performance will be compared to identify the most effective method for LC classification. These approaches use DL systems that depend solely on radar data, employing temporal and spatial synthetic SAR image features derived from annual series organized into seasonal clusters. Instead of using dense temporal image sequences, these synthetic features are input into the DL network. This approach not only provides spatial information about the scenes but also integrates multitemporal data through seasonal partitioning.

The proposed methodology achieves high LC classification performance even with Sentinel-2 tiles that have limited images per season, proving effective in situations with sparse temporal sequences. Preliminary tests suggest that the system performs well for global LC classifications, delivering robust results by efficiently summarizing spatial features from less dense temporal data and adapting to the climatic conditions of various regions at a 10m resolution. Future work includes enhancing the classification capabilities of the water extractor, a specialized module designed to analyze SAR time series for the classification of water bodies.

3.2.2 SAR water extraction

The water improvement method involves fusing optical data and an index-based approach to enhance water extent extraction accuracy, particularly in areas with limited SAR data or challenging environmental conditions. As outlined in [22], the previous project processing chain used unsupervised *k-means* clustering on SAR sequences to extract water features. However, this approach is less effective on smooth surfaces like desert sand. To address this, the classification process is improved by integrating data from SAR sequences with multitemporal indices from multispectral data sequences. For instance, the NDVI (Normalized Difference Vegetation Index) and NDWI (Normalized Difference Water Index) can be calculated. NDVI provides information on vegetation health and quantity by comparing Near-Infrared (NIR) and Red (R) reflectance, while NDWI evaluates the presence of water or soil moisture by comparing NIR and green (G) reflectance. The time-averaged values of these indices, combined with features extracted from the Red-Green-Blue (R-G-B) bands (such as maximum, minimum, and variance values over the time period of interest), can be merged with features from SAR sequences and fed into the *k-means* algorithm. Preliminary tests in the La Picasa Laguna area in Argentina indicate that this method yields a more accurate water mask, particularly on sandy or smooth surfaces.

4 Algorithm development plan for decision fusion

Based on the experimental results and the validation outcome on the three study areas of Phase 1, the team plans on exploring potential improvements within the data fusion processing chain and especially its higher semantic level components (spatial / temporal / multi-sensor decision fusion and harmonisation). The following subsections provide details on both the status of the decision fusion processing chain and the algorithm development plan for Phase 2.

4.1 Current status of decision fusion

The current pipeline of decision fusion, which corresponds to the ones built and validated during Phase 1, is illustrated in Figure 6. Figure 6.a indicates the block diagram to produce the static map while Figure 6.b shows the steps in the creation of the historical maps. The first data fusion processor is the multi-sensor geolocation, which takes inputs from both the optical pre-processing chain and the SAR pre-processing chain. In this step, the SAR data is spatially aligned to match the coordinate system of the optical data. Then, the decision fusion module, in the form of optical-SAR multi-sensor fusion, combines the classes taking into account the posterior probabilities of the processed optical and SAR data. In the historical map pipeline, the cascade multi-temporal model is additionally applied to this module to favour the temporal consistency of the product. The spatial harmonisation step is also performed to ensure a smooth spatial fusion on the overlapping part of two neighbouring Sentinel-2 granules. The final HRLC map is built by mosaicking together the spatially harmonised granules of the final fusion products.

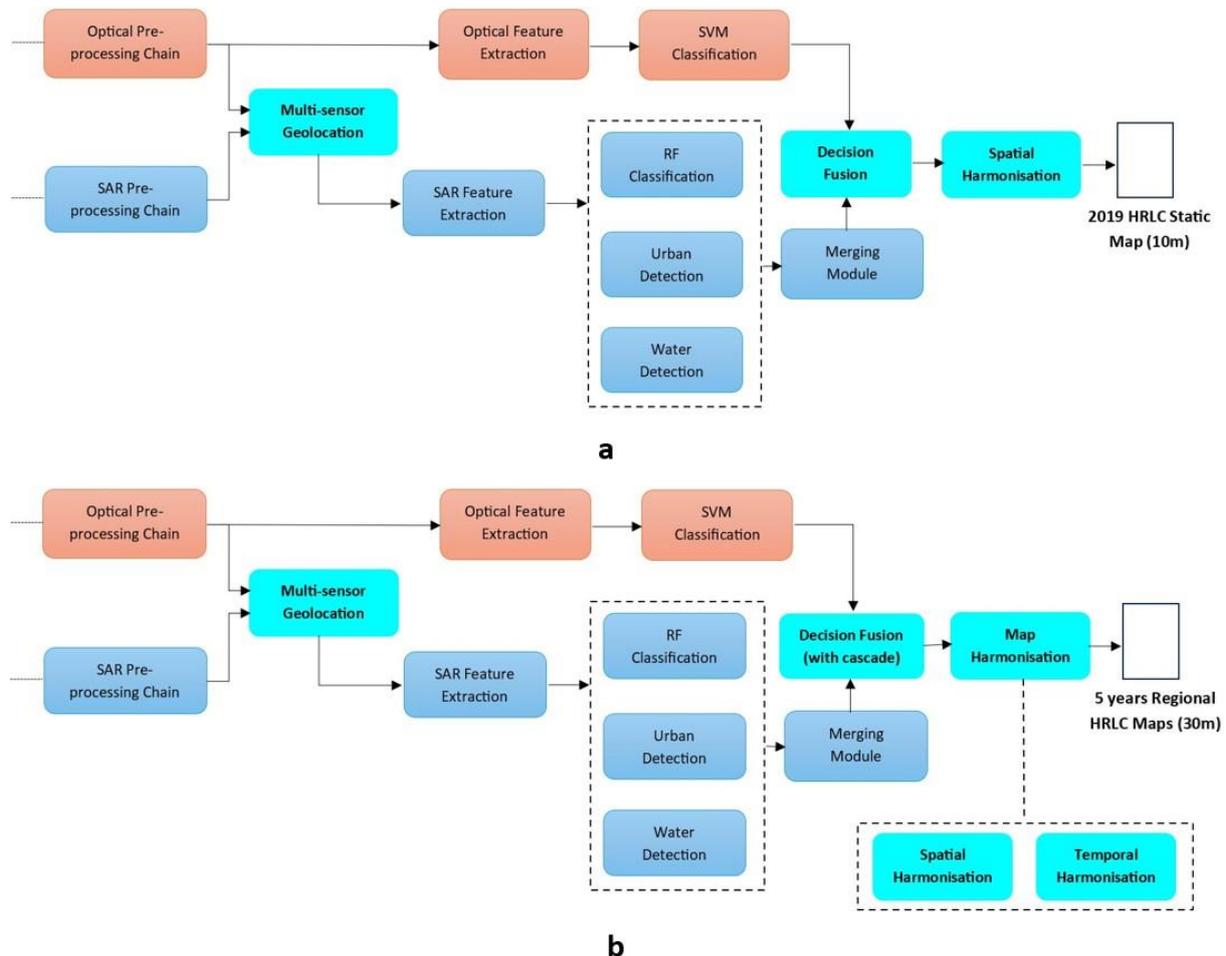


Figure 6. Decision fusion processing chain. The steps belonging to this chain are indicated in bold and light blue (cyan) background. (a) shows the workflow for the static map production while (b) indicates the pipeline for the historical maps.

4.1.1 Multi-sensor geolocation

In order to compare or integrate the data corresponding to the same scene but obtained from different measurements, image registration is necessary. Image registration is a general term to describe the process to align sets of images by referencing them into a common coordinate system [23]. In the CCI+ HRLC pipeline, image

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	14	

registration is applied to the outputs from the optical and SAR pre-processing chains to spatially align the data from both sensors, which is referred to as the multi-sensor geolocation. Multi-sensor geolocation is done to enable the joint use of data from optical and SAR sensors to eventually produce the land cover map. The optical imagery is taken as the “reference image” and the input i.e., data from SAR, is transformed to match the fixed reference image. The registration method is mainly composed of three elements: the geometric transformation, the similarity measure, and the optimization strategy. The geometric transformation consists of mapping the input image to the coordinate system of the reference image by applying affine transformations, and performing the nearest neighbour interpolation as the resampling strategy [24]. The similarity measures define to which extent the input and reference images match together. Within the pipeline of CCI+ HRLC, the similarity measure is based on the maximum mutual information [23], [25], [26] between the two images. The optimization strategy is utilized to find the optimal matching condition for the two images. This is achieved by Powell’s algorithm with barrier functions [27]. Moreover, an automatic tiling algorithm is developed in order to accommodate the possible local distortions due to the large size of the Sentinel-2 granules. Hence, the multi-sensor geolocation is applied to the patches divided from both input and reference images, instead of applying a single global transformation to the whole image [28].

In Phase 2, the multi-sensor geolocation will not be modified. This choice is because of both (1) the effectiveness of this processor, confirmed by the extensive validation conducted during Phase 1 with both optical-SAR and multi-mission SAR data [28], [29], and (2) the consistency with the general plan for Phase 2 to operate starting from the pixel-wise posteriors generated by the optical and SAR chains, whenever possible.

4.1.2 Optical-SAR multi-sensor and multi-temporal fusion

Each classification model applied to optical and SAR data not only assigns each pixel to a class, but also takes into account the uncertainty associated with each source by producing pixel-wise posterior probabilities. In particular, based on the sources of the data, there are classes that are exclusively classified using only the optical data or the SAR data, and a subset of classes which is found in both optical and SAR classification outputs. While the sensor-exclusive classes are used as they are, the common classes across the output of the two sensors are fused using optical-SAR multi-sensor fusion method. Considering that the sets of classes confidently classified from optical and SAR data are generally different, i.e., the data from the optical sensor are normally useful to discriminate all considered land cover classes while the SAR data in general are beneficial in distinguishing especially built-up areas and water bodies, a class-specific probabilistic decision fusion is utilized to address these output’s characteristics [30]. Furthermore, after this pixel-wise decision fusion step, a Markov Random Field (MRF) model [31] is applied in order to include the contextual information in the form of class interactions among the neighbouring pixels. This spatial MRF is formulated while taking into account the desired attributes of the final output of HRLC map in terms of smoothness, the degree of the salt-and-pepper classification noise, and the detail preservation, as indicated by the Climate Group. A fast implementation of the MRF energy minimization process, based on appropriate convolution operators, is integrated in the pipeline to minimize the computational burden [30].

Moreover, in the historical LC map pipeline, there is also a variability in the availability of the data due to the acquisitions that are often sparsely taken in different months of the year. To address this problem, a multi-temporal cascade model is used to propagate information towards past years in a probabilistic manner. This harmonisation step favours the temporal consistency across the historical land-cover products generated every five years.

4.1.3 Spatial harmonisation

When mosaicking granules together to generate the final LC map, the spatial harmonisation step is responsible for favouring the spatial regularity across the boundary between adjacent mapping tiles. Due to the different properties of the data, and more generally due to different data availabilities, the land cover maps of neighbouring granules may have slightly different characteristics, which impact on the mosaicking products in the form of edge artifacts at the interface between the neighbouring granules. In order to mediate this possible issue, the spatial harmonisation is performed on the overlapping parts of the two neighbouring granules utilizing the linear opinion pool (which belongs to the family of probabilistic decision fusion approaches) that incorporates space-varying weights to favour a seamless spatial fusion.

4.2 Development plan for decision fusion

The current decision fusion processing chain consists of multi-sensor geolocation, optical-SAR multi-sensor and multi-temporal fusion, as well as spatial harmonisation. As mentioned before, the development of the decision fusion processing chain will focus on the other high-semantic level modules (and not on the geolocation one),

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	15	

mainly to enhance the quality of the performance of the processors. For the multi-sensor fusion, the improvement in both the combination method for the class-specific pixel-wise posteriors and the fusion that takes into account spatial information is emphasized, while the temporal and spatial harmonisations will be refined to have a smoother spatial transition and a better temporal consistency, respectively. Furthermore, all of the enhancements will take into account the degree of the pixel fragmentation which will be agreed upon with the Validation and the Climate teams to determine to which extent the pixel labelling should be regularized according to the semantic of the surrounding classes.

4.2.1 Optical-SAR multi-sensor and multi-temporal fusion

The development for the optical-SAR multi-sensor fusion will put a focal point on the improvement on the fusion of the pixel-wise posterior probabilities coming from both sensors on all tiles for which input SAR image availability is sufficient, as well as on producing a better spatial quality of the fusion map. To achieve this goal, the artefacts on the fusion map will be minimized by adopting a probabilistic standpoint that takes into account the semantics of the classes explicitly. Furthermore, in order to capture the underlying spatial structure more accurately, the broader area of Probabilistic Graphical Models (PGM) [32], [33], [34] will be considered. Focus will be given again to the feedback from the Climate team, to make sure that the spatial properties of the map (smoothness, detail-preservation, etc.) are consistent with the behaviour expected by the climate community. Regarding the multi-temporal fusion, the problem will be formulated in the broader methodological framework of satellite image time series (SITS) analysis. The probabilistic fusion models will be developed, especially from the areas of 1-D Markov models [35] and Recurrent Neural Networks (RNNs) [36]. On one hand, the family of 1-D Markov models, such as Markov chains [34], hidden Markov models (HMM) [37], [38], or generally Bayesian networks i.e., PGMs on directed graphs [32], [39], is a class of powerful and flexible methods to model the time series data while still maintaining computational efficiency. On the other hand, the RNN family, including architectures such as the Long Short-Term Memory (LSTM) and the Gated Recurrent Unit (GRU) networks [40], [41], [42], is well-known from past studies to favour a good mapping accuracy. However, in their use, there are many aspects that need to be considered, especially in terms of the requirements of training data and the processing time for training. The choice among the PGM and neural families of methods will be addressed taking carefully into account these requirements.

We emphasize that all considered data fusion methodologies either are based on or can be linked to probabilistic formulations. This property ensures that the input uncertainty information, represented by the input posterior distributions, can be used to endow the output maps with a further probabilistic uncertainty measure.

4.2.2 Spatial harmonisation

The refinement of the spatial harmonisation involves reducing the possible residual harmonisation artefacts. The identification of these artefacts will be done by focusing on the modelling of the spatial gradient on the overlapping part across the boundary of two adjacent mapping granules. Moreover, a special attention will be given to the role of each class label, possibly conditioning the spatial gradient on the estimated labels. In this case as well, the spatial harmonization will be framed as a probabilistic fusion process, thus ensuring the possibility to propagate uncertainty information from the input to the output of the processing module.

5 Algorithm development plan for multitemporal change detection

The multitemporal change detection and trend analysis utilize Satellite Image Time Series (SITS) data from optical sensors, along with five-year regional HRLC maps at 30m spatial resolution. These data are analysed to identify abrupt and permanent changes as well as trends. The change detection process considers all available yearly datasets within the processing time and detects abrupt and permanent changes between consecutive years. The method employs regional LC maps every five years to identify candidate change pixels. The time series (TS) analysis is conducted at pixel level, utilizing all available data.

5.1 Current status of multitemporal change detection

Current status of the multitemporal Land Cover Change (LCC) detection is demonstrated in Figure 7 considering the developed methodology in phase 1 [RD1]. The input dataset is the multi-annual SITS containing six years of multi-temporal acquisitions to detect the changes between the adjacent years. If insufficient data is available for adjacent years (for more details see [RD1]) the method searches for subsequent years to find a year with enough information to proceed with the change detection analysis. The process begins with the extraction of features

by considering all possible pairs of the available sensor bands to compute a set of Normalized Difference Indices. A feature reduction strategy is employed to retain only the most informative features [43]. The time series is regularized to generate a denser, uniformly sampled sequence compared to the original signal [44]. During this stage, cloud/shadow and Post Classification Comparison (PCC) masks (from regional LC maps every five years) are applied to filter out cloudy pixels, detect candidate changed pixels and to select the priority change pixels for each region, as introduced by the climate modelers [RD4]. Subsequently, a binary change detection method based on Breaks For Seasonal and Trend (BFAST) [45] is utilized to further highlight changed and remove unchanged ones. To minimize computational complexity, a feature fusion strategy is applied to combine the reduced features.

The resulting product is a four-channel image: the first channel indicates the years in which changes occurred, the second provides the probability of a particular change, the third assesses the reliability of the reported year (taking into account the gaps in data across years), and the fourth is the PCC map, which highlights high and low priority pixels. This processing chain was implemented consecutively over a five-year Landsat 5, 7, and 8 SITS dataset.

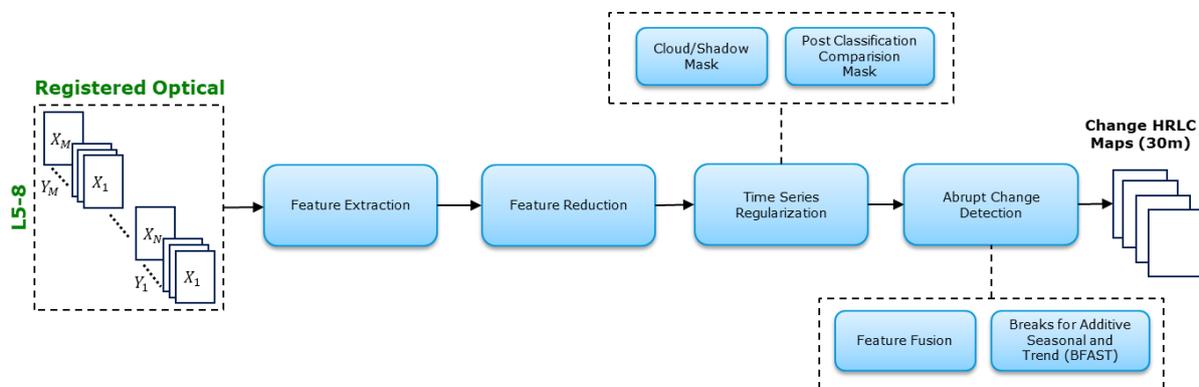


Figure 7. Change detection processing chain

5.1.1 Time Series reconstruction

Missing data and irregularities happen because of the clouds, cloud shadows, and radiometric effects. They lead to decrease data quality and hampers LCC detection. To address this, a time series reconstruction technique is applied to the extracted features from SITS, ensuring continuous and regular time series that are denser than the original signal. The choice of interpolation method is critical as it directly affects change detection accuracy and processing time. For non-vegetation profiles, upper-envelope piecewise cubic interpolation (PCHIP interpolation) is used, while for vegetation classes, an adaptive non-parametric regression using a General Regression Neural Network (GRNN) is performed [44]. Non-parametric regression captures temporal signature trends, reducing profile complexity and arithmetic dependency, thereby enhancing the overall quality and reliability of the reconstructed time series.

5.1.2 Abrupt change detection

To detect abrupt changes, the Breaks For Additive Season and Trend (BFAST) method is selected. BFAST models both linear trends and seasonal variations without requiring a reference period, setting a threshold, or defining a change trajectory, making it suitable for detecting changes in diverse areas with different types of changes. BFAST works by recursively evaluating the possibility of each time point being a breakpoint and then selecting the most optimal set of breakpoints. It requires a set of parameters to run, such as parameter h , which defines the minimal segment size between potentially detected breaks in the trend model. The magnitude of change is crucial for distinguishing real abrupt changes from other disturbances in the time series. Additionally, the probability of change for each pixel is calculated as a measure of uncertainty using the p -values from the OLS-MOSUM structural change test, the same test that BFAST employs to filter out time series with no change prior to run. This assigns probabilities of change to every time series.

BFAST is typically used on medium spatial resolution Satellite Image Time Series (SITS) and often analysed for the Normalized Difference Vegetation Index (NDVI) or other vegetation indices. However, it has been modified to handle multi-feature data using high-resolution SITS. BFAST outputs include the timing of abrupt changes and

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	17	

the probability of change. Given the substantial amount of data processed by the BFAST methodology for different regions, a feature fusion step based on calculating feature magnitude is implemented. This step alleviates the computational burden of BFAST for the various features extracted in the feature extraction phase.

5.2 Development plan for multitemporal change detection

The current processing chain for multitemporal change detection and trend analysis utilizes SITS data from optical sensors and five-year regional HRLC maps at 30m spatial resolution, analysing this data to identify abrupt and permanent changes. Considering that new areas are introduced for analysis and change detection will be performed using Sentinel-2 datasets, the methodologies need to be updated. More attention will be given to develop an integrated and unified preprocessing step for both LC classification and LCC detection, thereby minimizing the processing steps for optical data analysis and generating composites. This approach will ensure a homogeneous preprocessing, reducing the computational burden and make the preprocessing step as well as subsequent processing robust for and coherent between both LC classification and change detection analyses. The feature design strategy will be updated to extract better features related to specific LC changes for each area, and this step will be integrated with the LC classification processing chain to generate common features for both analyses in a single step, thereby reducing computational requirements. The feature extraction and feature reduction techniques will be further enhanced to leverage the higher spatial and temporal resolution of Sentinel-2. Additionally, time series reconstruction methods will be updated to handle the increased data volume and finer temporal resolution. A more robust and fast method for breakpoint detection will be implemented to efficiently process the larger datasets while maintaining or improving change detection accuracy. These updates will enable the processing chain to better handle the new data and areas, leading to more accurate and efficient change detection and trend analysis.

5.2.1 Feature space design

The development plan for the feature space design will involve creating region-specific feature spaces tailored to the unique characteristics of each region. This will be achieved using state-of-the-art methods to ensure the most effective change detection analysis [46], [47]. For instance, the feature space for change detection will incorporate spatial, temporal, and fine-grained features that capture the intricacies of each region landscape and changes over time. Pretrained deep learning architectures will be tested to extract detailed spatial and temporal features, leveraging their ability to learn complex patterns from large datasets [48], [49].

In addition, the feature reduction and fusion methods will be updated by employing advanced techniques, both deep learning and non-deep learning. For example, autoencoders and convolutional neural networks (CNNs) will be explored for deep learning-based feature reduction and fusion [50], [51]. Non-deep learning methods such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), and newer approaches like Uniform Manifold Approximation and Projection (UMAP) [52], and Locality Preserving Projections (LPP) [53] will also be evaluated for their effectiveness in this context.

The feature space design and dimension reduction will be an iterative process, involving continuous evaluation and refinement based on the results of preliminary analyses. By tailoring the feature space to each regions specific characteristics and utilizing cutting-edge techniques for feature extraction, reduction and fusion to enhance the representation of relevant changes, reducing noise, and capturing finer details that are critical for distinguishing variations in LCs. This comprehensive approach will ensure that the processing chain is robust and capable of handling the diverse and dynamic nature of the new areas introduced for analysis.

5.2.2 Time series reconstruction and break point detector

The upper envelope strategy for non-vegetation classes and a non-parametric regressor for vegetation classes developed in phase 1 shows limitations in computational time. Accordingly, experiments were conducted to understand if weekly time series reconstruction is a reliable time frequency. For Sentinel-2 data, with its frequent weekly acquisitions, this step can be omitted if sufficient data are available, thereby saving processing time and computational power. For periods with less frequent acquisitions, a different strategy will be implemented by considering the monthly, bi-monthly, and/or seasonal composites generated in the preprocessing stage (optical

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	18	

pre-processing block in Figure 1). The generated composites for LC classification will be also used for the change detection process, providing a set of input data with efficient temporal availability across the years in common with LC classification.

In phase2, other break point detectors such as the Bayesian Online Change Point Detection (BOCPD) method [54], known for its speed and efficiency, will be considered as alternatives to BFAST. Moreover, Change Point Detection based on Neural Networks (CPDNN) [55], [56], and Piecewise Linear Models with a specific parameters based on the area of the analysis will be explored to evaluate their capability and efficiency in detecting LCCs.

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	19	

6 References

- [1] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images," *Remote Sens. Environ.*, vol. 159, pp. 269–277, Mar. 2015, doi: 10.1016/j.rse.2014.12.014.
- [2] J. Louis *et al.*, "Sentinel-2 Global Surface Reflectance Level-2a Product Generated with Sen2Cor," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2019, pp. 8522–8525. doi: 10.1109/IGARSS.2019.8898540.
- [3] S. Skakun *et al.*, "Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2," *Remote Sens. Environ.*, vol. 274, p. 112990, Jun. 2022, doi: 10.1016/j.rse.2022.112990.
- [4] D. Frantz, "FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond," *Remote Sens.*, vol. 11, no. 9, Art. no. 9, Jan. 2019, doi: 10.3390/rs11091124.
- [5] O. Hagolle, M. Huc, C. Desjardins, S. Auer, and R. Richter, "MAJA Algorithm Theoretical Basis Document," Dec. 2017, doi: 10.5281/zenodo.1209633.
- [6] A. Zupanc, "Improving Cloud Detection with Machine Learning," Sentinel Hub Blog. Accessed: Jun. 27, 2024. [Online]. Available: <https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13>
- [7] D. Frantz, E. Haß, A. Uhl, J. Stoffels, and J. Hill, "Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects," *Remote Sens. Environ.*, vol. 215, pp. 471–481, Sep. 2018, doi: 10.1016/j.rse.2018.04.046.
- [8] S. Qiu, Z. Zhu, and B. He, "Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery," *Remote Sens. Environ.*, vol. 231, p. 111205, Sep. 2019, doi: 10.1016/j.rse.2019.05.024.
- [9] G. Doxani *et al.*, "Atmospheric Correction Inter-comparison eXercise, ACIX-II Land: An assessment of atmospheric correction processors for Landsat 8 and Sentinel-2 over land," *Remote Sens. Environ.*, vol. 285, p. 113412, Feb. 2023, doi: 10.1016/j.rse.2022.113412.
- [10] J. Chen, X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin, "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 1053–1064, Apr. 2011, doi: 10.1016/j.rse.2010.12.010.
- [11] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *Natl. Sci. Rev.*, vol. 5, no. 1, pp. 44–53, Jan. 2018, doi: 10.1093/nsr/nwx106.
- [12] L. Bruzzone, "Multisource Labeled Data: an Opportunity for Training Deep Learning Networks," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2019, pp. 4799–4802. doi: 10.1109/IGARSS.2019.8898311.
- [13] G. Perantoni and L. Bruzzone, "A Novel Technique for Robust Training of Deep Networks With Multisource Weak Labeled Remote Sensing Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022, doi: 10.1109/TGRS.2021.3091482.
- [14] M. Rußwurm and M. Körner, "Self-attention for raw optical Satellite Time Series Classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 169, pp. 421–435, Nov. 2020, doi: 10.1016/j.isprsjprs.2020.06.006.
- [15] R. Sedona, C. Paris, L. Tian, M. Riedel, and G. Cavallaro, "An Automatic Approach for the Production of a Time Series of Consistent Land-Cover Maps Based on Long-Short Term Memory," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2022, pp. 203–206. doi: 10.1109/IGARSS46834.2022.9883655.
- [16] W. Zhao, C.-A. Deledalle, L. Denis, H. Maître, J.-M. Nicolas, and F. Tupin, "Ratio-Based Multitemporal SAR Images Denoising: RABASAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3552–3565, Jun. 2019, doi: 10.1109/TGRS.2018.2885683.
- [17] A. Sorriso, D. Marzi, and P. Gamba, "A General Land Cover Classification Framework for Sentinel-1 SAR Data," in *2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI)*, Sep. 2021, pp. 211–216. doi: 10.1109/RTSI50628.2021.9597319.
- [18] Gorica Bratic and Maria Antonia Virelli, "Map Of Land Cover Agreement - MOLCA." Accessed: Jun. 06, 2024. [Online]. Available: <https://zenodo.org/records/8071675>
- [19] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," May 20, 2018, *arXiv: arXiv:1804.03999*. doi: 10.48550/arXiv.1804.03999.
- [20] H. Cao *et al.*, "Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation," in *Computer Vision – ECCV 2022 Workshops*, L. Karlinsky, T. Michaeli, and K. Nishino, Eds., Cham: Springer Nature Switzerland, 2023, pp. 205–218. doi: 10.1007/978-3-031-25066-8_9.
- [21] D. Marzi, J. I. S. Jara, and P. Gamba, "A 3-D Fully Convolutional Network Approach for Land Cover Mapping Using Multitemporal Sentinel-1 SAR Data," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, pp. 1–5, 2024, doi:

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	20	

- 10.1109/LGRS.2023.3332765.
- [22] D. Marzi and P. Gamba, "Inland Water Body Mapping Using Multi-temporal Sentinel-1 SAR Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. PP, pp. 1–1, Nov. 2021, doi: 10.1109/JSTARS.2021.3127748.
- [23] J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, Eds., *Image Registration for Remote Sensing*, 1st ed. Cambridge University Press, 2011. doi: 10.1017/CBO9780511777684.
- [24] A. A. Goshtasby, "Fusion of multi-exposure images," *Image Vis. Comput.*, vol. 23, no. 6, pp. 611–618, Jun. 2005, doi: 10.1016/j.imavis.2005.02.004.
- [25] A. A. Cole-Rhodes, K. L. Johnson, J. LeMoigne, and I. Zavorin, "Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1495–1511, Dec. 2003, doi: 10.1109/TIP.2003.819237.
- [26] D. Solarna, A. Gotelli, J. Le Moigne, G. Moser, and S. B. Serpico, "Crater Detection and Registration of Planetary Images Through Marked Point Processes, Multiscale Decomposition, and Region-Based Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6039–6058, Sep. 2020, doi: 10.1109/TGRS.2020.2970908.
- [27] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, no. 2, pp. 155–162, Feb. 1964, doi: 10.1093/comjnl/7.2.155.
- [28] D. Solarna, L. Maggiolo, G. Moser, and S. B. Serpico, "A Tiling-Based Strategy for Large-Scale Multisensor Optical-Sar Image Registration," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Kuala Lumpur, Malaysia: IEEE, Jul. 2022, pp. 127–130. doi: 10.1109/IGARSS46834.2022.9884048.
- [29] B. Pinel-Puysegur *et al.*, "Experimental Comparison of Registration Methods for Multisensor Sar-Optical Data," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Brussels, Belgium: IEEE, Jul. 2021, pp. 3022–3025. doi: 10.1109/IGARSS47720.2021.9553640.
- [30] L. Maggiolo, D. Solarna, G. Moser, and S. B. Serpico, "Optical-SAR Decision Fusion with Markov Random Fields for High-Resolution Large-Scale Land Cover Mapping," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Kuala Lumpur, Malaysia: IEEE, Jul. 2022, pp. 5508–5511. doi: 10.1109/IGARSS46834.2022.9884751.
- [31] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. in Advances in Pattern Recognition. London: Springer London, 2009. doi: 10.1007/978-1-84800-279-1.
- [32] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*, Nachdr. in Adaptive computation and machine learning. Cambridge, Mass.: MIT Press, 2010.
- [33] M. J. Wainwright and M. I. Jordan, "Graphical Models, Exponential Families, and Variational Inference," *Found. Trends® Mach. Learn.*, vol. 1, no. 1–2, pp. 1–305, 2007, doi: 10.1561/22000000001.
- [34] M. Pastorino, G. Moser, S. B. Serpico, and J. Zerubia, "Semantic Segmentation of Remote-Sensing Images Through Fully Convolutional Neural Networks and Hierarchical Probabilistic Graphical Models," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022, doi: 10.1109/TGRS.2022.3141996.
- [35] A. S. Willsky, "Multiresolution Markov models for signal and image processing," *Proc. IEEE*, vol. 90, no. 8, pp. 1396–1458, Aug. 2002, doi: 10.1109/JPROC.2002.800717.
- [36] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. in Adaptive computation and machine learning. Cambridge, Mass: The MIT press, 2016.
- [37] K. M. Leon-Lopez, F. Mouret, H. Arguello, and J.-Y. Tourneret, "Anomaly Detection and Classification in Multispectral Time Series Based on Hidden Markov Models," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–11, 2022, doi: 10.1109/TGRS.2021.3101127.
- [38] S. P. Abercrombie and M. A. Friedl, "Improving the Consistency of Multitemporal Land Cover Maps Using a Hidden Markov Model," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 703–713, Feb. 2016, doi: 10.1109/TGRS.2015.2463689.
- [39] J. Tao, W. Wu, and M. Xu, "Using the Bayesian Network to Map Large-Scale Cropping Intensity by Fusing Multi-Source Data," *Remote Sens.*, vol. 11, no. 2, p. 168, Jan. 2019, doi: 10.3390/rs11020168.
- [40] S. Wang, J. Cao, and P. S. Yu, "Deep Learning for Spatio-Temporal Data Mining: A Survey," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 8, pp. 3681–3700, Aug. 2022, doi: 10.1109/TKDE.2020.3025580.
- [41] M. Alameh, Y. Abbass, A. Ibrahim, G. Moser, and M. Valle, "Touch Modality Classification Using Recurrent Neural Networks," *IEEE Sens. J.*, vol. 21, no. 8, pp. 9983–9993, Apr. 2021, doi: 10.1109/JSEN.2021.3055565.
- [42] A. Trucco, A. Barla, R. Bozzano, S. Pensieri, A. Verri, and D. Solarna, "Introducing Temporal Correlation in Rainfall and Wind Prediction From Underwater Noise," *IEEE J. Ocean. Eng.*, vol. 48, no. 2, pp. 349–364, Apr. 2023, doi: 10.1109/JOE.2022.3223406.
- [43] S. Solorio-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, "A review of unsupervised feature selection methods," *Artif. Intell. Rev.*, vol. 53, pp. 907–948, 2020.
- [44] Y. T. Solano-Correa, K. Meshkini, F. Bovolo, and L. Bruzzone, "A Land-cover Driven Approach for Fitting

	Ref	D2.4 - ADP		
	Issue	Date	Page	
	1.0	02/08/2024	21	

Satellite Image Time Series in a Change Detection Context," *Image Signal Process. Remote Sens. XXVI*, vol. 11533, 2020.

- [45] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sens. Environ.*, vol. 114, pp. 106–115, 2010.
- [46] J. Seo, W. Park, and T. Kim, "Feature-Based Approach to Change Detection of Small Objects from High-Resolution Satellite Images," *Remote Sens.*, vol. 14, no. 3, p. 462, Jan. 2022, doi: 10.3390/rs14030462.
- [47] J. Li, S. Zhu, Y. Gao, G. Zhang, and Y. Xu, "Change Detection for High-Resolution Remote Sensing Images Based on a Multi-Scale Attention Siamese Network," *Remote Sens.*, vol. 14, no. 14, p. 3464, Jul. 2022, doi: 10.3390/rs14143464.
- [48] K. Meshkini, F. Bovolo, and L. Bruzzone, "An Unsupervised Change Detection Approach for Dense Satellite Image Time Series Using 3D CNN," *IEEE Int. Geosci. Remote Sens. Symp. IGARSS*, 2021.
- [49] A. M. El Amin, Q. Liu, and Y. Wang, "Convolutional neural network features based change detection in satellite images," *Proc SPIE*, vol. 10011, p. 100110W, 2016.
- [50] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006, doi: 10.1126/science.1127647.
- [51] C. Zhang *et al.*, "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 166, pp. 183–200, Aug. 2020, doi: 10.1016/j.isprsjprs.2020.06.003.
- [52] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," 2018, *arXiv*. doi: 10.48550/ARXIV.1802.03426.
- [53] A. Wang, S. Zhao, J. Liu, J. Yang, L. Liu, and G. Chen, "Locality adaptive preserving projections for linear dimensionality reduction," *Expert Syst. Appl.*, vol. 151, p. 113352, Aug. 2020, doi: 10.1016/j.eswa.2020.113352.
- [54] G. Yoshizawa, "Bayesian Online Change Point Detection for Baseline Shifts," *Stat. Optim. Inf. Comput.*, vol. 9, no. 1, pp. 1–16, Dec. 2020, doi: 10.19139/soic-2310-5070-1072.
- [55] Z. Atashgahi, D. C. Mocanu, R. Veldhuis, and M. Pechenizkiy, "Memory-free Online Change-point Detection: A Novel Neural Network Approach," 2022, *arXiv*. doi: 10.48550/ARXIV.2207.03932.
- [56] M. Hushchyn, K. Arzymatov, and D. Derkach, "Online Neural Networks for Change-Point Detection," 2020, *arXiv*. doi: 10.48550/ARXIV.2010.01388.