4.A Overview of the "Climate Change Initiative Extension - High Resolution Land Cover" Project

4.A.1 Relation with the CCI MR Project

The "Climate Change Initiative Extension" (CCI+) program is the natural follow up of the "Climate Change Initiative" (CCI). Its focus is to monitor the Essential Climate Variables (ECVs) identified by the Global Climate Observing System (GCOS) in support of the United Nations Framework Convention on Climate Change (UNFCC). The main purpose of the CCI+ is to continue the successful achievements made under the CCI programme focusing the attention on (among other goals):

- developing new ECVs not yet included in CCI
- new research and development studies on existing ECVs already started in CCI
- interaction and collaboration between EO science community and climate science community
- prototype product generation and system definition

These needs emerge because long-term global Earth Observation archives are now available to either the scientific and user communities from past/current and planned ESA EO missions for climate purposes (e.g., historical ERS and Envisat missions, relevant ESA managed archives of Third-Party Mission data and the Copernicus Space Component). Thus, CCI+ aims to define and validate innovative approaches for continuously generating and updating a comprehensive set of ECV products in the long term by focusing on the consistency and quality analysis, from a climate model perspective.

Among the various ECV products, the focus of this Chapter is on the "High Resolution (HR) Land Cover ECV". It addresses a new variable with respect to CCI and focuses on the impact of land cover (LC) and LC changes (LCC) on climate changes, with the primary objective of examining in detail the role of the spatial resolution of the detected changes to support climate modelling research. Although it cannot be directly used as an input to climate models, land cover is an essential climate variable to quantify surface energy and water fluxes and the sources of greenhouse gasses, to monitor variation in land use and land surface and to characterize the impacts of extreme events (e.g., floods, heatwaves, droughts or hurricanes). LC change is in fact both a cause and a consequence of climate change either when the change is human-induced or generated from natural events. This has been widely recognized by the scientific community and demonstrated by the previous CCI program, focused on the generation of Moderate Resolution (MR) Land Cover ECV maps at global scale.

The MRLC CCI provided annual LC maps at 300m resolution covering the period 1992-2015 (the 2016 annual map is expected) and, as side activity, a first preliminary 20m product based on Sentinel-2 has been generated for the African continent for 2015-16. CCI MR introduced a new concept of global LC map with respect to the past, by considering a multi-sensor approach and generating time series of interoperable and consistent annual global products. The goal of the CCI+ project is to extend and improve this new concept (multi-year and multi-sensor LC and LCC, while keeping the retro-compatibility with other products) by increasing the spatial resolution of the generated products.

Within CCI+ HRLC, the usage of high resolution remote sensing data to produce LC maps and detect LC changes opens the door to an unprecedented wide range of possibilities for climate change analysis (including improved quantification of energy, water and carbon budgets of terrestrial ecosystems). However, increasing the spatial resolution reframes the perspective with respect to the MR project both from the theoretical and the operational view point. Although HR increases the potential of analysing in detail dynamic phenomena of the LC, many challenges are introduced with respect to the MR case. The temporal availability of HR data in the past/current archives is much lower than the MR one and it strongly varies across the years. Differently from the medium resolution case (e.g., SPOT-Vegetation), no daily acquisitions are available and only in the very recent years it was possible to get a dense temporal sampling thanks to the Sentinel and Landsat-8 missions. Prior to them, the amount of yearly-based images available in archives dramatically reduces (being Landsat-5 the most relevant source), yielding to a much less uniform framework for the development of HRLC products.

This scenario leads to a more complex process for the production of historical time series of products. In particular, it requires a shift in the processing paradigm that moves from the analysis of many images per year

HRLC CLASSES											
CODE	DN	1st LEVEL	CODE	DN	2nd LEVEL	CODE	DN	3rd LEVEL	CODE	DN	4th LEVEL
10	1	Tree cover evergreen broadleaf									
20	2	Tree cover evergreen needleleaf									
30	3	Tree cover deciduous broadleaf									
40	4	Tree cover deciduous needleleaf									
50	5	Shrub cover evergreen	51	17	Broadleaf						
			52	18	Needleleaf						
60	6	Shrub cover deciduous	61	19	Broadleaf						
			62	20	Needleleaf						
70	7	Grasslands	71	21	Natural						
			72	22	Managed						
80	8	Croplands	81	23	Winter	811	24	Rainfed			
						812	25	Irrigated	8121	26	Sparkling
									8122	27	Flooding
			82	28	Summer	821	29	Rainfed			
						822	30	Irrigated	8221	31	Sparkling
									8222	32	Flooding
			83	33	Multicropping	831	34	Rainfed			
						832	35	Irrigated	8321	36	Sparkling
									8322	37	Flooding
90	9	Woody vegetation aquatic or regularly flooded									
100	10	Grassland vegetation aquatic or regularly flooded									
110	11	Lichens and Mosses									
120	12	Bare areas	121	38	Unconsolidated	1211	39	Sands			
						1212	40	Bare soils			
			122	41	Consolidated						
130	13	Built-up	131	42	Buildings						
			132	43	Artificial Roads						
140	14	Open Water seasonal									
150	15	Open Water permanent									
160	16	Permanent snow and/or ice	161	44	Snow						
			162	45	Ice						

Figure 4.19: The legend adopted in the context of the project.

acquired at medium resolution to few (or very few for some areas and years) images characterized by high spatial resolution. A key CCI+ requirement is the production of 5 years annual LC maps with a resolution of 10-30m. This will ensure consistency in the temporal sampling among different areas and maximise the exploitation of historical data available in archives (e.g. Landsat). The higher geometrical detail requires the re-definition of the legend for the LC map to be consistent with 300m spatial resolution products. The resulting HR legend follows the specification of the Land Cover Classification System (LCCS upgraded to LCML recently) and is compatible with the already existing 300m MR one. Such newly defined legend provides a higher thematic content with respect to the MR LC map (see Figure 4.19).

4.A.2 The Three Areas of Study and the Main Products

The primary goal of the "High Resolution (HR) Land Cover ECV" is to advance the technical knowledge and capacity of the scientific, vegetation and the climate community in assessing:

- the climate impacts on observed LC change
- whether LC has changed as a function of climate either permanently or temporarily
- the energy and matter (mainly water and carbon) balances for terrestrial ecosystems, as major drivers of vegetation-climate feedbacks

Moreover, the project investigates the role of spatial-temporal resolution on the consistency of LC classification to quantify the variability of LC and LC changes detected at different scales. Because of the multitude of efforts on generating HRLC maps, the analysis is not performed at global level, but at regional and subcontinental scale. According to the MR products generated in the previous CCI phase and the preliminary requirements provided by the climate research group, three preliminary test areas have been identified: Amazon, African Sahel, and Siberia (see Figure 4.20). In such areas:

- LC change is noted in the moderate resolution products (Africa and the southern part of the Amazon forest);
- no LC change is reported, but changes in vegetation state have been observed (central part of the Amazon Forest);
- areas where the change in LC has been attributed as a contributor to

modification of climate and vice versa (Siberia, and likely west Africa).

In each of the three areas, one of the goals of the project is the generation of three main products:

- a static HRLC map at subcontinental scale produced at 10m spatial resolution mainly with Sentinel-1 and Sentinel-2 images;
- a dynamic product consisting of regional temporal products at 30m spatial resolution and with a temporal resolution of 5 years, synchronized with the MR product and going back in time till the 1990. Moreover, the historical series of LC maps is generated by using a larger number of satellite sensors than those used to generate the static map and considering, as far as possible, a data archive independent from the one used to generate the static map;
- annually long-term record of regional changes that are consistent with the dynamic products at 30m resolution.

Figure 4.20 shows the details of such products, identifying the areas of interest with colors. Indeed, the red boxes identify the regions where the static maps are generated, while the blue boxes highlight the areas for the dynamic products.

4.A.3 Overview of the Processing Chain

The processing chain proposed in the context of the project to generate the products described above is capable of handling a large volume of data (e.g., multi-temporal Sentinel-1, Sentinel-2, and the Landsat datasets, well beyond Landsat-8, and automatically ingesting remote sensing data acquired by different sensors such as SPOT, ERS, ASAR). Additionally, the processing chain has been designed to be flexible and modular, prone to easy integration of additional modules and different algorithms. Indeed, the overall processing chain is composed of:

- an optical processing chain focusing on Sentinel-2 (static map) and Landsat (dynamic product) imagery;
- a SAR processing chain focusing on Sentinel-1 (static map), ERS, and ASAR (dynamic product);
- a data fusion processing chain at the interface between the optical and the SAR processing chains.

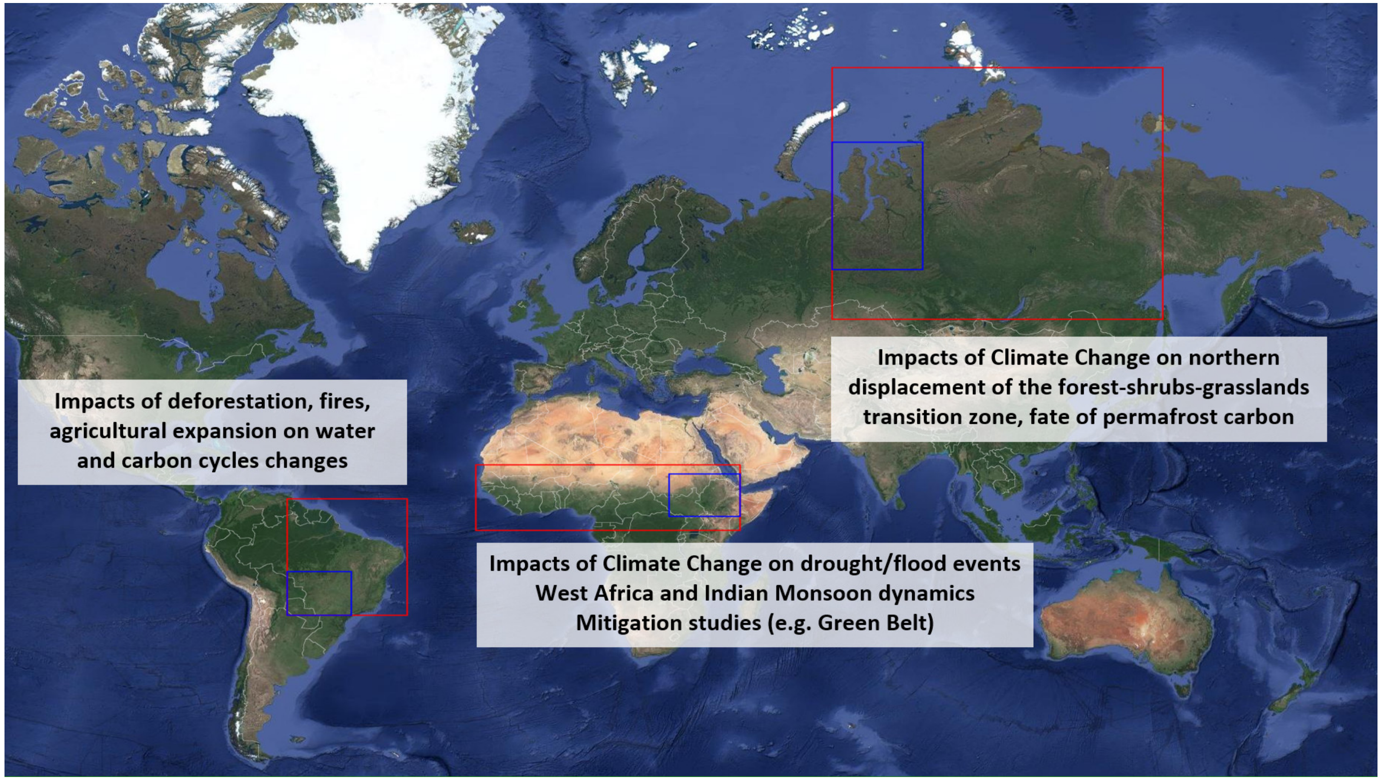

Figure 4.20: The three sub-continental areas the CCI HRLC project focuses on: Amazon, African Sahel, and Siberia. Details on the reason why such areas have been chosen are also reported in the text boxes. The red boxes identify the regions where the static maps are generated, while the blue boxes highlight the areas for the dynamic products.

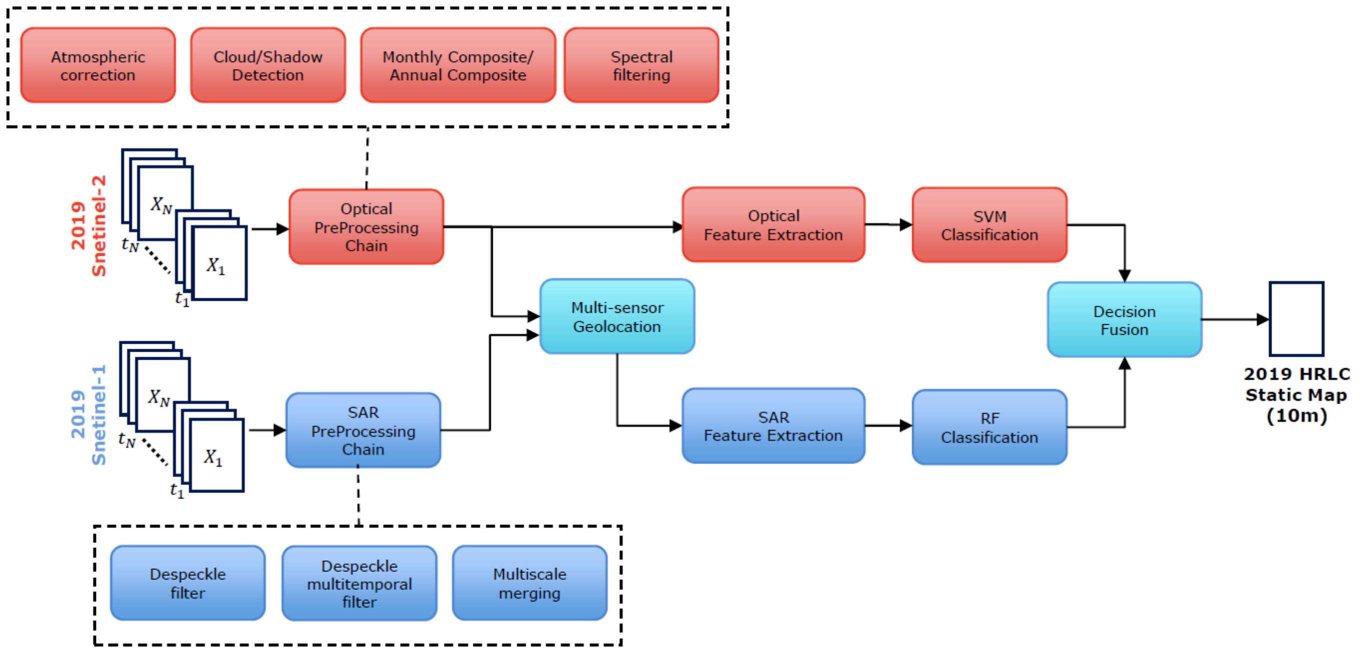

Figure 4.21: Flowchart of the processing chain designed to generate the static product. It is worth noting that the data fusion modules (i.e., the multisensor geolocation and the decision fusion blocks) are located at the interface between the optical and SAR processing chains. This ensures independence between the processing chains of the different types of data, that are combined only at specific steps in the processing chain. Indeed, the resulting processing chain is highly modular and flexible.

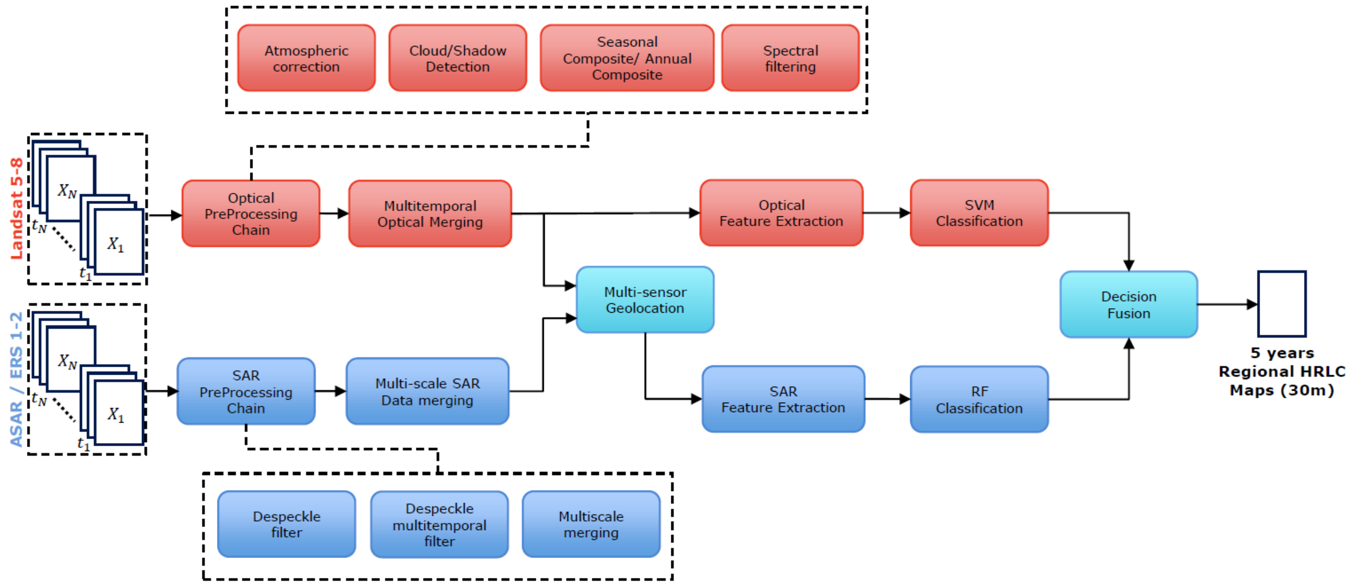

Figure 4.22: Flowchart of the processing chain designed to generate the dynamic products. As for the static product case, it is worth noting that the data fusion modules (i.e., the multisensor geolocation and the decision fusion blocks) are located at the interface between the optical and SAR processing chains. This ensures independence between the processing chains of the different types of data, that are combined only at specific steps in the processing chain. Indeed, the resulting processing chain is highly modular and flexible.

Figure 4.21 and Figure 4.22 show the flowchart of the two processing chains, the former in charge of producing the static map for the larger sub-continental areas in 2019, and the latter in charge of generating the dynamic maps every 5 years and back to the 1990s. It is worth noting the main building blocks of the two chains. Indeed, the following two sections are aimed at describing them in details.

As for the static product, Figure 4.21 shows the input data to be the Sentinel-2 and Sentinel-1 data of 2019. Then, the raw data is pre-processed by two separate chains for the two types of input data. The pre-processed images are then forwarded to the first of the two data fusion modules, which is the multisensor geolocation module. Here, the optical and the SAR images are registered in order for the resulting classification maps to be defined on the same pixel grid. Due to the fact that the Sentinel-2 images are better suited for the discrimination of the majority of the classes present in the legend (see Figure 4.19), and due to the fact that the optical imagery is able to grant more informative spatial features, the optical image is used as a master (i.e., it is not modified by the geolocation module), while the SAR data is transformed in order for it to match the spatial grid of the optical. The registered data is then forwarded to the two separate processing chains responsible for the optical and SAR classification. The output class probabilities (and not the classification maps) are then combined by the second data fusion module (i.e., the decision fusion block). The fusion step is responsible for generating the final product, which consists in the static product for the 2019, together with information about the pixelwise uncertainty in the classification.

Conversely, for the dynamic product case, Figure 4.22 shows the different input data being Landsat 5-8 for the optical chain and ASAR and ERS 1-2 for the SAR chain. The type of sensor that is being used depends on the year, as the dynamic products span a time frame that goes from 1990 to 2015, and the list of satellites operating in that period is not stable from product to product. As for the static map, there are again two separate processing chains for the optical and the SAR imagery. After the pre-processing chains, that are different from the static product case, the resulting data is passed to the multisensor geolocation module. The optical images are again used as master, and the SAR images are transformed to match the optical pixel grid. Moreover, the pixel grid is setup according to the Landsat path/row scheme at 30m resolution. After the geolocation, the classification chains produce the posterior probabilities on the 30m grid that are then fused by the decision fusion module. Finally, the output product consists of the land cover classification maps for the three smaller areas (i.e., the areas highlighted by the blue boxes in Figure 4.20) in the years ranging from 1990 to 2015 and with a time resolution of 5 years. Also in this case, the land cover maps are augmented with information related to the pixelwise uncertainty.