Round-robin protocol

Authors	Bontemps Sophie (UCL), Brockmann Carsten (BC), Defourny Pierre (UCL), Kirches Grit (BC), Boettcher, Martin (BC)
Distribution	ESA : Arino Olivier, Kalogirou Vasileios

Table of recorded changes

Issue Record Table

ISSUE	DATE	REASON
1.0	April 1, 2011	Completion of the first version
1.1	April 11, 2011	Update of the first version, based on RIDs received from ESA
1.2	April 21, 2011	Update to include the link between compositing period and classification algorithms in the RRob

Detailed Change Record Table

ISSUE	RID No.	SECTION	CHANGE
1.1	RRP/1	2.2.2.	Text has been updated according to suggestions (the word “algorithm” has been changed) and figures 3 and 4 have been modified accordingly
1.1	RRP/2	All documents	Tables 1, 2, 10, 11 and 20 have been completed with a size estimation for the different input and validation data that are distributed
1.1	RRP/3	4.1	Study areas common with the pre-processing have been highlighted in Table 13
1.1	RRP/6	5.4	Section about the “best” algorithm selection criteria has been added in the document
1.1	Progress	2.2.2	Figure 3 has been updated to highlight the fact that multi-year dataset have to be used in the classification algorithm
1.1.	Progress	3.2	Included the correct web address
1.1	Progress	3.2.3	New flag included with regard to retrieval accuracy
1.1	Progress	3.3	Rephrased with respect to the understanding
1.1	Progress	5.2	Clarify access to CEOS land sites
1.1	Progress	7, 8	Forms that participants are asked to filled in have been added, next to the Data Policy
1.2	Progress	2.2.2	Section has bee modified to include the test on the compositing period (update of text and figure 3)
1.3	RRP-1.2/7	2.2.2 + 5.4.2	The test of the compositing period impact is clearly stated in the section 2.2.2 (as one classification component to be tested for the classification chain – see figure 3). In addition, assessment methodology and selection criteria for this particular test have been clearly specified in the section 5.4. 2
1.3	RRP-1.2/9	5.4.2	Section has been modified to provide a more detail description of the algorithms selection criteria and methods

Symbols and acronyms

AERONET : AErosol RObotic NETwork
AP : Algorithm Provider
ASAR : Advanced Synthetic Aperture Radar
ATBD : Algorithm Theoretical Basis Document
BC : Brockmann Consult
BRDF : Bidirectional Reflectance Distribution Function
CCI : Climate Change Initiative
CCI-LC : Climate Change Initiative – Land Cover
CEOS : Committee on Earth Observation Satellites
CMC : Climate Modelling Community
DP : Data Provider
DUE : Data User Element
ENVISAT : ESA Environmental Satellite
EO : Earth Observation
ESA : European Space Agency
FR : Full Resolution
FRS : Full Resolution Full Swath
GlobCover : ESA DUE project (http://ionia1.esrin.esa.int/)
MERIS : Medium Resolution Imaging Spectrometer
PSD : Product Specification Document
RRob : Round-Robin
RRP : Round-Robin Protocol
RRDP : Round-Robin Data Package
SDR :Surface Directional Reflectance
SPOT : Satellite Pour l'Observation de la Terre
SPOT-VGT : SPOT-Vegetation
SRTM : Shuttle Radar Topography Mission
SWBD : SRTM Water Bodies Dataset
UCL : Université catholique de Louvain

Reference documents

Applicable document

ID	Title
[AP-1]	ESA Climate Change Initiative Phase I - Scientific User Consultation and Detailed Specification Statement of Work (SoW), version 1.4, 09.11.2009
[AP-2]	ESA Climate Change Initiative Phase 1 - Project Guidelines V1
[AP-3]	Defourny, P. and CCI-LC project partners, ECV Land Cover Part 3 - Technical proposal for the European Space Agency , 05/03/2010
[AP-4]	Herold M., van Groenestijn A., Kooistra L., Kalogirou V. and Arino O., CCI Land Cover project – User Requirements Document (URD), version 2.2 (23/02/2011)
[AP-5]	CMUG, 2010, Requirement Baseline Document MOHC, MPI-M, ECMWF, MétéoFrance.
[AP-6]	Brockmann C., Krüger O., Böttcher M., Kirches G., Defourny P., Bontemps S., Achard F., Mayaux P., Schmullius C., Traut K., Michaelis R., Santoro M., Kalogirou, V. and Arino O., CCI Land Cover project – Data Access Requirement Document (DARD), version 1.6 (02/03/2011)
[AP-7]	Defourny P., Bontemps S., Böttcher M., Brockmann C., Fomferra N., Kirches G., Krüger O., CCI Land Cover project – Products Specification Document (PSD), version 1.1 (15/03/2011)
[AP-8]	Kirches G., Brockmann C., Krüger O., Dane O., Böttcher M., CCI Land Cover project – Algorithm Theoretical Basis Document (ATBD) for pre-processing, version 0 (01/02/2011)
[AP-9]	Bontemps S., Defourny P., Schmullius C., Santoro M., Traut K., Michaelis R., Kalogirou V. and Arino O., CCI Land Cover project – Algorithm Theoretical Basis Document (ATBD) for classification, version 0 (01/02/2011)
[AP-10]	Achard F., Bontemps S., Defourny P., Herold M., Mayaux P., Kalogirou V. and Arino O., CCI Land Cover project – Product Validation Plan, version 1.0 (31/01/2011)

Reference document

ID	Title
[RD-1]	Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X., 2010, MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sensing of Environment, 114, 168-182.
[RD-2]	Defourny P. et al. 2010, ESA Climate change initiative: land_cover_cci, ESA Living Planet Symposium, 1st July 2010.
[RD-3]	GlobCover 2005 validation report v. 2.1. (GLOBCOVER_MAN_GVR_EUM3_I2.1), 05/11/2008
[RD-4]	GlobCover 2005 technical specification v. 4.0. (GLOBCOVER_MAN_TS_EUM3_I4.0), 13/10/2008
[RD-5]	NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.5, 25 October, 2010
[RD-6]	OpenGIS Geography Markup Language (GML) Application Schema for Earth Observation Products, Implementation Standard, OGC 06-080r4, 1.0, 2010-02-25
[RD-7]	Committee on Earth Observation Satellites, 2008, Working Group on Information Systems and Services - Interoperability Handbook, February 2008 Issue 1.1 Available at : http://wiki.ieee-earth.org/@api/deki/files/7/=Handbook_0802.pdf

Table of content

1. Objective of the round-robin activity

The objective of the round-robin (RRob) activity is the selection of the "best" algorithm(s) or combination of algorithms [AP-1]. The RRob aims at improving the understanding of algorithms (and/or sensors) performance in enabling the inter-comparison of varying algorithms. This is an open activity in which external research groups are invited to develop and propose their own methodologies.

The "best" algorithm selection is achieved following algorithm inter-comparison and product validation exercises and considering the user-focus as most important. To this end, the algorithms performance and products accuracy are evaluated towards the targeted criteria defined in the Product Specification Document (PSD) [AP-7].

The entire effort is organized as transparent as possible. All participants evolve within a standardized environment: they have the same information about context and objective and they work with the same dataset. The scope of the exercise, the input dataset, the expected output, the validation dataset and the evaluation methodology are defined in advance and described in this Round-Robin Protocol (RRP) document.

2. General framework for the round-robin activity

2.1. Project framework and algorithms

The European Space Agency (ESA) Climate Change Initiative (CCI) will deliver the next generation of satellite derived geophysical parameters that should meet requirements from the Global Climate Observing System (GCOS) for Essential Climate Variables (ECV) and the Climate Modelling Community (CMC).

In particular, the CCI Land Cover (CCI-LC) project aims at delivering, in a consistent way over years and from various Earth Observation (EO) instruments, global land cover products matching the needs of key users belonging to the CMC. To this end, a prototype system will be developed, based on a set of pre-processing and classification algorithms (Figure 1).

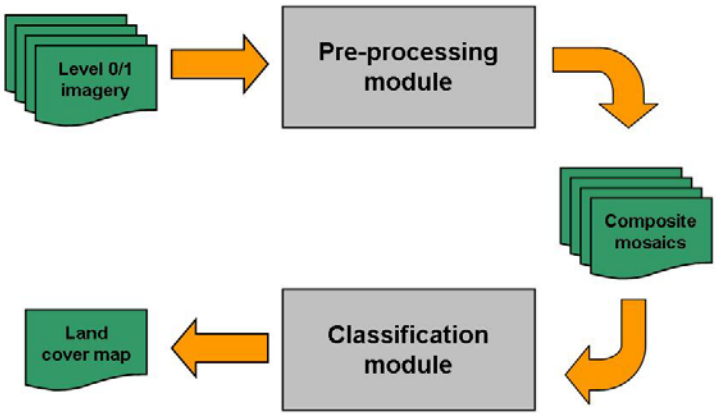

Figure 1. Schematic illustration of the pre-processing and classification modules articulation

The classification module is articulated to the pre-processing one. The pre-processing module starts from L1B imagery and produces global composited mosaics of land surface reflectance values. These mosaics serve as input to the classification module, which interprets them into land cover classes.

2.2. Algorithms involved in the round-robin activity

The RRob aims at improving the understanding of algorithms (and/or sensors) performance in enabling the inter-comparison of varying algorithms. Any inter-comparison exercise can be viewed as composed of:

- an input dataset composed of *R* events ;
- a set of *S* alternative processing algorithms ;
- *T* different output products ;
- a methodology for comparison (e.g. graphical visualisation methods and/or statistical analyses), along with appropriate validation dataset.

The CCI-LC RRob activity concentrates on specific algorithmic steps of the pre-processing and classification modules (with for each step, specific input and output dataset). The steps involved in the RRob activity are presented in the two following sections.

2.2.1. Pre-processing algorithms

The pre-processing module produces global composited mosaics of land surface reflectance values, in performing (i) radiometric and geometric correction, (ii) pixel identification, (iii) atmospheric correction with aerosol retrieval and (iv) compositing and mosaicing procedures. The produced global composited mosaics are the input to the classification module.

Figure 2 illustrates the whole pre-processing chain in highlighting in red the steps (and the associated results) that are tested in RRob activity.

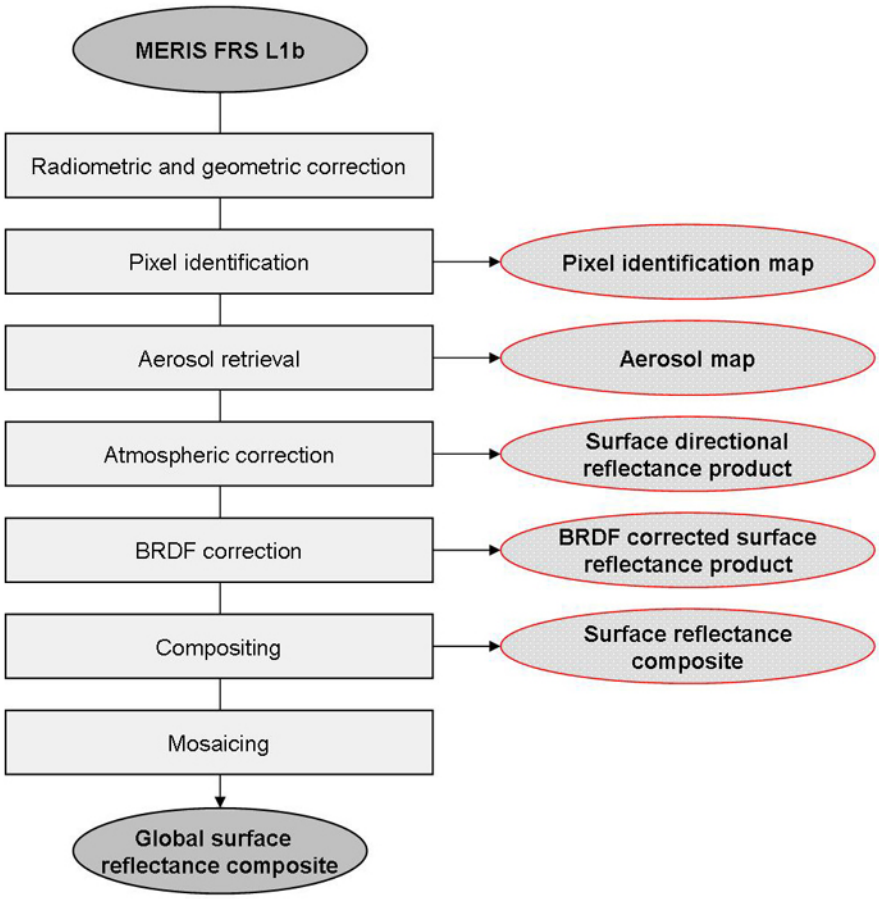

Figure 2. Schematic illustration of the pre-processing chain in the framework of the RRob activity

The five pre-processing steps challenged through the RRob activity are:

- the pixel identification, which should result in the "best" identification of pixel status (i.e. land, water, snow and ice, cloud, cloud shadow and invalid);
- the aerosol retrieval, which should result in the "best" aerosol properties estimation;
- the atmospheric correction, which should result in the "best" surface directional reflectance (SDR) products (i.e. in surface reflectance products which are not BRDF-corrected);

- the BRDF correction, which should result in the "best" surface reflectance products;
- the compositing algorithm, which should minimize the compositing length while maintain the composite quality in terms of space and time consistency.

In order to allow testing these 5 steps, two input dataset are needed:

- MERIS FRS L1b imagery, which is of use for the 3 first steps;
- SDR products, which are related to the 2 last steps.

Finally, the different outputs that can be expected from this pre-processing RRob activity are:

- the pixel identification map;
- the aerosol map;
- the SDR product, accounting for the geometric and radiometric correction, the pixel identification, the aerosol retrieval and the atmospheric corrections (i.e. all the pre-processing steps before the BRDF correction);
- the BRDF corrected surface reflectance product;
- the surface reflectance composite, accounting for the BRDF correction and the compositing algorithm.

It should be noted that the two last steps (BRDF correction and compositing algorithm) could be handled together (i.e. developing a compositing method which corrects BRDF). In this case, a single output can be proposed for these two steps.

2.2.2. Classification algorithms

With regard to the classification module, the RRob is articulated in two independent activities. The first one concerns a global and generic classification methodology using optical time series (Figure 3) while the second one is related to class-specific discrimination algorithms and makes use of both optical and SAR dataset (Figure 4).

Figure 3 and Figure 4 both present the input and output dataset and the algorithmic steps that are challenged through the RRob activities.

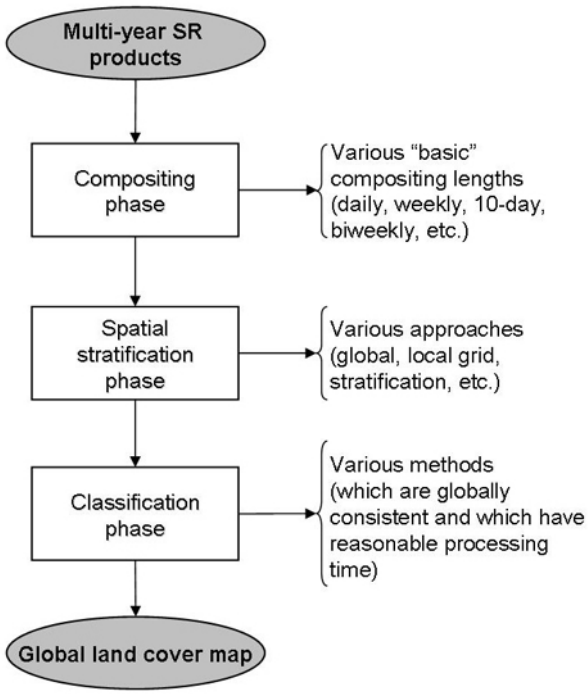

Figure 3. Schematic illustration of a generic and global classification chain: composite mosaics produced by the pre-processing are analyzed according to a specific spatial approach and a determined classification methodology to produce a global land cover map

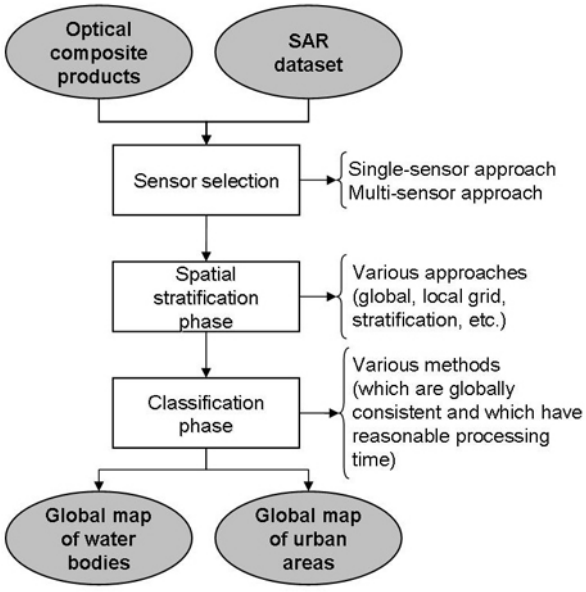

Figure 4. Schematic illustration of the two targeted class-specific classification chain: starting from optical and/or SAR dataset, two independent chains have to be developed to identify urban areas and water bodies

With regard to the first RRob activity (Figure 3), different algorithms are tested in three particular phases:

- the compositing phase, where the impact of the "basic" compositing period on the classification algorithm(s) performance should be assessed;

- the spatial stratification phase, where the alternative approaches could be either a global approach (i.e. classifying the entire world as a single dataset), the local grid approach (i.e. gridding the entire world in regular rectangular small cells) and a stratified approach (i.e. stratifying the entire world in regions homogeneous with respect to specific criteria);
- the classification phase, where any alternative can be foreseen as far as the proposed approach is consistent at the global scale, can handle with multi-year dataset to provide a unique land cover map which is consistent over time (i.e. not sensitive to annual conditions) and has a reasonable processing time.

The associated input and output dataset are:

- a multi-year dataset of global daily SDR products and of global SR composites generated by the pre-processing module (input);
- a single global land cover map with a definite legend (output).

As already noted, the second RRob activity (Figure 4) integrates the possibility to use SAR imagery. It involves three distinct phases in which varying algorithms can be tested:

- the input data selection, considering that the SAR dataset ***has to be used***. Consequently, two alternatives can be foreseen, which are a single-sensor SAR-based approach or an approach combining both dataset;
- the spatial stratification phase, where the alternative approaches could be either a global approach (i.e. classifying the entire world as a single dataset), the local grid approach (i.e. gridding the entire world in regular rectangular small cells) and a stratified approach (i.e. stratifying the entire world in regions homogeneous with respect to specific criteria);
- the classification phase, where any alternative can be foreseen as far as the proposed approach is consistent at the global scale and has a reasonable processing time.

Starting from the same set of inputs, two independent products are expected, thus leading to the necessity to develop two independent classification chains. RRob participants don't have to focus on these two chains. The sets of inputs and outputs are:

- the two input dataset, which are the surface reflectance composite products generated by the pre-processing module on the one hand and a SAR dataset over the same period on the other hand;
- the two independent outputs, which are a global map for the water bodies and the urban areas.

3. Round-robin data package for pre-processing

3.1. Input dataset

Based on the considerations of section 2.2.1, the input data belonging to the Round-Robin Data Package (RRDP) for the pre-processing include both satellite imagery (Table 1) and auxiliary dataset related to the different pre-processing steps to be tested (Table 2).

Table 1. Input satellite dataset useful for the pre-processing and included in the CCI-LC RRDP

Product	Specification	Format and size	Tested algorithms
ENVISAT MERIS FRS_L1P products	300 m spatial resolution 15 spectral channels 6 spatial subsets all over the world From 2005 to 2009	ENVISAT ~2TB per year and all subsets	Pixel identification algorithm Aerosol retrieval Atmospheric correction
ENVISAT MERIS FRS surface directional reflectance products	Surface directional reflectance (not BRDF-corrected) products 300 m spatial resolution 15 spectral channels 3 spatial subsets all over the world	NetCDF ~2TB per year and all subsets	BRDF correction Compositing algorithm

Table 2. Input auxiliary dataset useful for the pre-processing and included in the CCI-LC RRDP

Product	Specification	Format and size	Tested algorithms
ENVISAT Attitude and Orbit	Restituted orbit state vectors and pointing Global coverage From 2005 to 2009	ENVISAT ~4GBper year	Pixel identification algorithm Aerosol retrieval Atmospheric correction
SRTM Water Bodies Dataset (SWBD)	Land/water mask (version 2.1) Tiles of 1°*1° over 6 spatial subsets all over the world Version 2.1 from 2000 Available at: http://dds.cr.usgs.gov/srtm/version2_1/SWBD/	Vector dataset ~1-2050 KB per tile	Pixel identification algorithm
Monthly Aerosol Optical Depth	Dataset of pre-computed aerosol optical depth derived from MERIS data global From 2005 to 2009	Beam Dimap 260 MB per year	Atmospheric correction

Product	Specification	Format and size	Tested algorithms
Getasse 3.0 Digital Elevation Model	Elevation Model (version 3.0) for use with AMORGOS Tiles over 6 spatial subsets all over the world Version 3.0 Available at: http://earth.esa.int/services/amorgos/download/getasse/	Raster dataset, ~12657 KB per Tile	Pixel identification algorithm BRDF correction Compositing algorithm

As mentioned in the tables, the data are provided over 6 spatial subsets which define the RRob study areas for the pre-processing. These spatial subsets were selected both to be strategic from the pre-processing point of view and to include validation (AERONET and/or CEOS) sites. The location of these subsets is presented in Figure 5 and their geographic extent is detailed in Table 3.

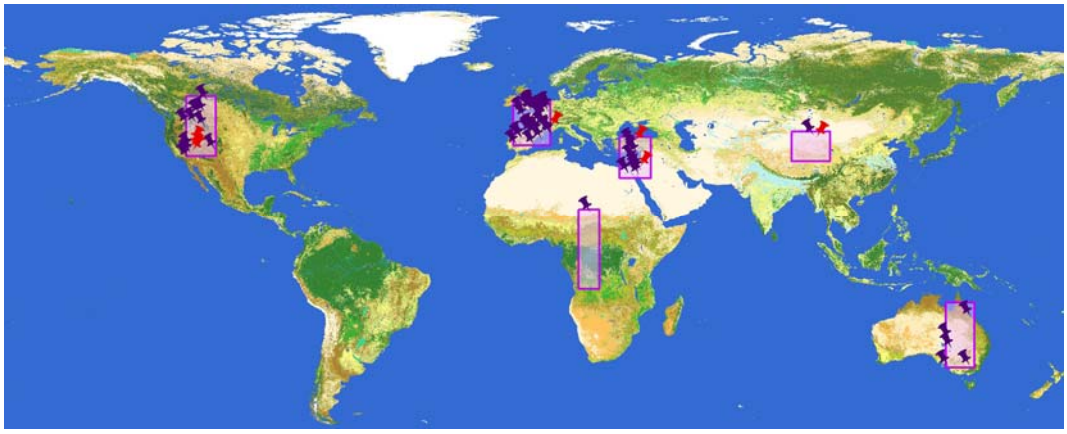

Figure 5. Localization of the 6 study areas used for the pre-processing RRob activity. Red marks indicate CEOS validation sites while mauve ones represent AERONET stations.

Table 3. Name and extent of the study areas used to test pre-processing algorithms in the RRob activity

Subset ID	Subset name	Upper left corner	Lower right corner	Size (km²)
1	North America	117°W, 55°N	107°W, 35°N	2 473 088
2	Western Europe	7°W, 54°N	5.5°E, 38.5°N	2 395 249
3	Africa	15°E, 17°N	22°E, 10°S	2 336 757
4	Central Asia	86.6°E, 43.2°N	99.7°E, 33.3°N	1 260 685
5	Northwest Asia	28.6°E, 41.4°N	27.3°E, 39.3°N	1 494 169
6	Australia	138.3°E, 14.6°S	148.2°E, 36.2°S	2 294 462

It should be noted that RRob participants are requested to apply their algorithms to **minimum 4 spatial subsets**. Indeed, one important requirement for the CCI-LC project is the need to be globally consistent and rather automated (global products over several years being to be produced within a 12-months period).

3.2. Output dataset

For each algorithm tested in the RRob activity (see section 2.2.1), a specific output is foreseen:

- For the pixel identification algorithm, a "pixel identification" map is expected, where the status of each pixel is identified (land, water, snow and ice, cloud, cloud shadow and invalid);
- For the aerosol retrieval algorithm, an aerosol map has to be provided, including aerosol optical depths and angstrom exponent;
- For the pre-processing chain including all steps until the atmospheric correction, the expected output is a surface directional reflectance (SDR) product;
- For the BRDF correction, a surface reflectance product needs to be delivered;
- For the compositing algorithm, a surface reflectance composite representing the best trade-off between the compositing length and the spatial and temporal quality is to be presented.

These five outputs are summarized in Table 4, along with their format. The format of the pre-processing RRob outputs is NetCDF-3, i.e. the "classic model" of NetCDF since most tools can handle them. The file specification follows CF conventions [RD-5].

Table 4. Output dataset expected from the pre-processing RRob

Product	**Tested algorithm**	**Number of products**	**Format**
Pixel identification map	Pixel identification algorithm	Minimum 4	NetCDF
Aerosol map	Aerosol retrieval algorithm	Minimum 4	NetCDF
Surface directional reflectance product	Pre-processing chain, including radiometric and geometric correction, pixel identification, aerosol retrieval and atmospheric correction	Minimum 4	NetCDF
Surface reflectance product	BRDF correction	Minimum 4	NetCDF
Surface reflectance composites	Compositing algorithm	Minimum 4	NetCDF

In addition, since the RRob aims at enabling the inter-comparison of varying algorithms, participating algorithms must be well-documented. Each algorithm should be provided with a report (word or pdf format) describing the developed methodologies and the products. In the particular case BRDF correction and compositing algorithms were considered together, the way BRDF correction is integrated in the compositing algorithm should be clearly documented.

The technical specifications with which the pre-processing RRob outputs have to be in line are described in detail in the following sections.

3.2.1. Product identification and filename convention

The file names of the pre-processing RRob outputs should follow the following naming convention:

<filename> ::= <id>-<subset>-<centre>-<proc time>.nc

Products are logically identified by their type, their time, and their product revision:

<id> ::= <type>-<time>-a<algorithm>

The product type is composed of the project acronym and work phase, the sensor (satellite source), the product code and level, the resolution, and the aggregation period:

<type> ::= <project>-<sensor>-<code>-<level>-<resolution>-<period>

The name components are defined in Table 5..

The dash "-" is the separator between name components. The filename convention obeys NetCDF CF by using the postfix ".nc".

Table 5. Filename convention of the pre-processing RRob outputs

Name component	Format	Description
project	"cci-lc-rrob" (constant)	Project Acronym and work phase
sensor	"meris" (constant)	Mission, platform and sensor identifier
code	e.g "pim"	Product code identifier pim - Pixel identification map aem - Aerosol map sdr - surface directional reflectance product sr - surface reflectance product csr - surface reflectance composites
level	e.g. "L2"	Product level [RD-7] Level 2 − Retrieved environmental variables (e.g., ocean wave height, soil moisture, ice concentration) at the same resolution and location as the level 1 source data. Level 3 − Data or retrieved environmental variables which have been spatially and/or temporally resampled (i.e., derived from level 1 or 2 products). Such resampling may include averaging and compositing
resolution	e.g. "300m"	Spatial resolution
period	e.g. "10d"	Aggregation period (no aggregation "00d")
time	"yyyyMMddHHmmss"	Time of measurement or start time of aggregation period
algorithm	e.g. "Idepix"	shortcut of algorithm
subset	Study areas	"vv" where: "01" is North America "02" is Western Europe "03" is Africa "04" is Central Asia "05" is Northwest Asia "06" is Australia

Name component	Format	Description
centre	Name of the processing center [a-zA-Z0-9._]*	Processing centre, generator of the product
processing time	"yyyyMMddHHmmssnnn"	Processing time: "yyyy" is the processing year "MM" is the processing month "dd" is the processing day "HH" is the processing hour "mm" is the processing minute "ss" is the processing second "nnn" (added to make file name unique if necessary)

3.2.2. Global attributes

The NetCDF file of pre-processing RRob outputs shall have the following global attributes (Table 6).

Table 6. Global attributes of the pre-processing RRob outputs, according to the structure of the NetCDF files

Attribute Name	Format	Description
title		Short description of the file content
institution		Where the original data was produced
source		Method of production of the original data
history		List of applications that have modified the original data, with time stamp, processor and parameters
references		References that describe the data or methods used to produce it.
comment		miscellaneous information about the data or method used to produce it
Conventions		Name of the conventions followed
type		Product type
id		Product identifier, see section 3.2.1
subset		Study areas
algorithm	e.g. "Idepix"	shortcut of algorithm
resolution		Resolution of the product in meters
creation_date	yyyyMMddHHmmss (UTC)	Creation time of product

3.2.3. Variables

The NetCDF file of the pre-processing RRob outputs has two dimensions that define the spatial raster, as described in Table 7.

Table 7. Dimension of the pre-processing RRob outputs, according to the structure of the NetCDF files

Dimension	Value	Description
y		Dimension that distinguishes different lines
x		Dimension that distinguishes different columns
t	2	Start and stop time of aggregation interval (closed interval at start, open interval at stop: [start;stop[) Note: no aggregation interval is defined by the same star and stop time

The variables and variable attributes that shall be contained in the NetCDF file of the pre-processing RRob outputs are described in Table 8.

Table 8. Identical variables and variables' attributes of the pre-processing RRob outputs , according to the structure of the NetCDF files

Variable	Attribute	Format	Value	Description
crs		int	0	Coordinate reference system attribute container
	grid_mapping_name		prefered - sinusoidal	
	semi_major_axis		6378137.0	
	inverse_flattening		298.257223563	
	earth_radius		6371007.181	
	false_easting		0.0	
	false_northing		0.0	
	longitude_of_central_meridian		0.0	
	scale_factor_at_central_meridian		1.0	
time		double(t)		Start and stop time of the aggregation interval
	long_name		aggregation time interval	
	units		days since 1981-01-01 00:00:00	
	calendar		julian	
y		double(y)		y coordinate of projection
	standard_name		projection_y_coordinate	

Variable	Attribute	Format	Value	Description
	long_name		y coordinate of projection	
	units		m	
x		double(x)		x coordinate of projection
	standard_name		projection_x_coordinate	
	long_name		x coordinate of projection	
	units		m	
lon		double (y,x)	-180.0 .. 180.0	Longitude coordinate of pixel
	standard_name		longitude	
	long_name		longitude coordinate	
	units		degrees east	
	valid_min		-180.0	
	valid_max		180.0	
lat		double (y,x)	-90.0 .. 90.0	Latitude coordinate of pixel
	standard_name		latitude	
	long_name		latitude coordinate	
	units		degrees north	
	valid_min		-90.0	
	valid_max		90.0	

The definition of the other variables and variable attributes depends on the product type of the pre-processing RRob outputs. They are described in the following Table 9. Following variables attributes in the NetCDF file of the output product should be defined by the RRob participants if possible:

- long_name,
- valid_min,
- valid_max
- _FillValue.

Table 9. Variables of the pre-processing RRob outputs, according to the structure of the NetCDF files

Product	Variable	Description
Pixel identification map	Clear Land	Pixel is classified as land and not covered by cloud/cloud shadow and snow/ice
	Clear Water	Pixel is classified as water and not covered by cloud/cloud shadow and snow/ice

Product	Variable	Description
Pixel identification map	Cloud	Pixel is classified covered by cloud
	Cloud shadow	Pixel is classified covered by cloud shadow
	Snow/and Ice	Pixel is classified as snow/and ice covered
	Invalid	Status of the pixel is invalid
Aerosol map	AOD	Aerosol optical depth by 550 nm
	AAC	Aerosol Angstrom Coefficient
	flag	Quality Flag
Surface directional reflectance product	sdr_1	surface directional reflectances of MERIS channel 1 (412.5 nm)
	sdr_2	surface directional reflectances of MERIS channel 2 (442.5 nm)
	sdr_3	surface directional reflectances of MERIS channel 3 (490 nm)
	sdr_4	surface directional reflectances of MERIS channel 4 (510 nm)
	sdr_5	surface directional reflectances of MERIS channel 5 (560 nm)
	sdr_6	surface directional reflectances of MERIS channel 6 (620 nm)
	sdr_7	surface directional reflectances of MERIS channel 7 (655 nm)
	sdr_8	surface directional reflectances of MERIS channel 8 (681.25 nm)
	sdr_9	surface directional reflectances of MERIS channel 9 (708.75 nm)
	sdr_10	surface directional reflectances of MERIS channel 10 (753.75 nm)
	sdr_11	surface directional reflectances of MERIS channel 11 (760.625 nm)
	sdr_12	surface directional reflectances of MERIS channel 12 (778.75 nm)
	sdr_13	surface directional reflectances of MERIS channel 13 (865 nm)
	sdr_14	surface directional reflectances of MERIS channel 14 (885 nm)
	sdr_15	surface directional reflectances of MERIS channel 15 (900 nm)
	NDVI	Normalized Vegetation Index
	Flag_L1b	Flags of the MERIS FRS L1b Product
	Flag_1 … Flag_n	Set of Flags - Quality, Error estimation etc. description of Flags included in the product specification
Surface reflectance product	sr_1	surface reflectances of MERIS channel 1 (412.5 nm)
	sr_2	surface reflectances of MERIS channel 2 (442.5 nm)
	sr_3	surface reflectances of MERIS channel 3 (490 nm)
	sr_4	surface reflectances of MERIS channel 4 (510 nm)
	sr_5	surface reflectances of MERIS channel 5 (560 nm)
	sr_6	surface reflectances of MERIS channel 6 (620 nm)
	sr_7	surface reflectances of MERIS channel 7 (655 nm)
	sr_8	surface reflectances of MERIS channel 8 (681.25 nm)
	sr_9	surface reflectances of MERIS channel 9 (708.75 nm)
	sr_10	surface reflectances of MERIS channel 10 (753.75 nm)

Product	Variable	Description
Surface reflectance product	sr_11	surface reflectances of MERIS channel 11 (760.625 nm)
	sr_12	surface reflectances of MERIS channel 12 (778.75 nm)
	sr_13	surface reflectances of MERIS channel 13 (865 nm)
	sr_14	surface reflectances of MERIS channel 14 (885 nm)
	sr_15	surface reflectances of MERIS channel 15 (900 nm)
	NDVI	Normalized Vegetation Index
	Flag_L1b	Flags of the MERIS FRS L1b Product
	Flag_1 … Flag_n	Set of Flags - Quality, Error estimation etc. description of Flags included in the product specification
Surface reflectance composite	sr_1_mean	Mean of surface reflectances of MERIS channel 1 (412.5 nm)
	sr_1_sigma	Sigma of surface reflectances of MERIS channel 1 (412.5 nm)
	sr_2_mean	Mean of surface reflectances of MERIS channel 2 (442.5 nm)
	sr_2_sigma	Sigma of surface reflectances of MERIS channel 2 (442.5 nm)
	sr_3_mean	Mean of surface reflectances of MERIS channel 3 (490 nm)
	sr_3_sigma	Sigma of surface reflectances of MERIS channel 3 (490 nm)
	sr_4_mean	Mean of surface reflectances of MERIS channel 4 (510 nm)
	sr_4_sigma	Sigma of surface reflectances of MERIS channel 4 (510 nm)
	sr_5_mean	Mean of surface reflectances of MERIS channel 5 (560 nm)
	sr_5_sigma	Sigma of surface reflectances of MERIS channel 5 (560 nm)
	sr_6_mean	Mean of surface reflectances of MERIS channel 6 (620 nm)
	sr_6_sigma	Sigma of surface reflectances of MERIS channel 6 (620 nm)
	sr_7_mean	Mean of surface reflectances of MERIS channel 7 (655 nm)
	sr_7_sigma	Sigma of surface reflectances of MERIS channel 7 (655 nm)
	sr_8_mean	Mean of surface reflectances of MERIS channel 8 (681.25 nm)
	sr_8_sigma	Sigma of surface reflectances of MERIS channel 8 (681.25 nm)
	sr_9_mean	Mean of surface reflectances of MERIS channel 9 (708.75 nm)
	sr_9_sigma	Sigma of surface reflectances of MERIS channel 9 (708.75 nm)
	sr_10_mean	Mean of surface reflectances of MERIS channel 10 (753.75 nm)
	sr_10_sigma	Sigma of surface reflectances of MERIS channel 10 (753.75 nm)
	sr_11_mean	Mean of surface reflectances of MERIS channel 11 (760.625 nm)
	sr_11_sigma	Sigma of surface reflectances of MERIS channel 11 (760.625 nm)
	sr_12_mean	Mean of surface reflectances of MERIS channel 12 (778.75 nm)
	sr_12_sigma	Mean of surface reflectances of MERIS channel 12 (778.75 nm)
	sr_13_mean	Sigma of surface reflectances of MERIS channel 13 (865 nm)
	sr_13_sigma	Mean of surface reflectances of MERIS channel 13 (865 nm)
	sr_14_mean	Sigma of surface reflectances of MERIS channel 14 (885 nm)

Product	Variable	Description
	sr_14_sigma	Mean of surface reflectances of MERIS channel 14 (885 nm)
Surface reflectance composite	sr_15_mean	Sigma of surface reflectances of MERIS channel 15 (900 nm)
	sr_15_sigma	Mean of surface reflectances of MERIS channel 15 (900 nm)
	NDVI_mean	Mean of normalised vegetation index
	NDVI_sigma	Sigma of normalised vegetation index
	status_flag	Status of the pixel, where surface reflectance is retrieved
	clear_ sky_count	Number of contributing observations over clear sky land in aggregation period
	snow_ice_count	Number of observations with snow and ice coverage in aggregation period
	cloud_count	Number of non-contributing observations with cloud coverage in aggregation period
	cl_shadow_count	Number of non-contributing observations with cloud shadow coverage in aggregation period
	valid_count	Number of valid observations in aggregation period
	invalid_count	Number of invalid observations in aggregation period
	risk_flag	Description of the risk of failed BRDF-correction

3.2.4. Metadata

Metadata of the pre-processing RRob outputs should be provided as a GML file. The corresponding structure definition can be found in the OpenGIS GML Application Schema for Earth Observation Products [RD-6]. GML also provides XML schema definitions for XML validation.

3.3. Validation dataset

The comparison between the different pre-processing algorithms proposed through the RRob activity is made thanks to a set of validation data, which are presented in Table 10.

Table 10. Validation dataset to use to evaluate the performance of the pre-processing algorithms

Product to validate	Validation dataset	Specification	Format and size
Pixel identi-fication map Aerosol map SDR product	ENVISAT MERIS FRS surface directional reflectance products (generated by the pre-processing chain)	SDR (not BRDF-corrected) products 300 m spatial resolution 15 spectral channels 3 spatial subsets all over the world	NetCDF ~2TB per year and all subsets
Surface reflectance composite	ENVISAT MERIS FRS 10-day surface reflectance composites	Surface reflectance (BRDF-corrected) composites 300 m spatial resolution 15 spectral channels	NetCDF ~500GB per year and all

	(generated by the pre-processing chain)	3 spatial subsets all over the world	subsets
Pixel identi-fication map	In-situ data	Database of manually selected and classified pixels based on MERIS data Representative points of the year 2005 - 2009 distributed over the 6 pre-processing study areas **NOT DISTRIBUTED**	Database
Aerosol map	In-situ data AERONET stations	Aerosol Optical Depth (AOD) – Level 2 representative points of the year 2005 - 2009 distributed over the 6 pre-processing study areas Available at: http://aeronet.gsfc.nasa.gov/cgi bin/combined_data_ access_new **NOT DISTRIBUTED**	ASCII tables (compressed zip file)
SDR product	In-situ data CEOS land sites	8 CEOS LANDNET sites: Standard reference sites for the post-launch calibration of space-based optical imaging sensors. Representative points of the year 2005 - 2009 distributed over the 4 pre-processing study areas Available at: http://calvalportal.ceos.org/ **NOT DISTRIBUTED**	ASCII tables
Surface reflectance (BRDF corrected) product	No external dataset – Use of the generated surface reflectance products	Surface reflectance (BRDF-corrected) products generated by the developed algorithms are validated by using intrinsic quality control methods: statistics computation including distributions, temporal evaluation of the mean reflectance values of the variance amplitudes 3 spatial subsets all over the world 2005-2009	Report pdf or word-document
Surface reflectance (BRDF corrected) composite	No external dataset – Use of the generated surface reflectance composites	Surface reflectance (BRDF-corrected) composites generated by the developed algorithms are validated by using intrinsic quality control methods: statistics computation including distributions, temporal evaluation of the mean reflectance values of the variance amplitudes 3 spatial subsets all over the world 2005-2009	Report pdf or word-document

As indicated in Table 10, only the ENVISAT MERIS FRS surface directional reflectance products and surface reflectance composites (generated by the pre-processing chain) are made available to RRob participants, who are invited to provide inter-comparison results (with these maps and/or other dataset) in a validation report. If other dataset are used, they shall be fully documented in the report.

With regard to the in-situ data points of the LC CCI pre-processing validation database, they are not distributed since they will be used by the CCI-LC consortium to perform the final comparison between all algorithms proposed in the RRob activity. Furthermore, its non-distribution allows being sure that

none algorithm will be tuned towards this dataset and thus, ensures a completely independent validation.

4. Round-robin data package for classification

4.1. Input dataset

Based on the considerations of section 2.2.2, the input data belonging to the classification RRDP include both optical and SAR imagery, as illustrated in Table 11. Table 12 presents the auxiliary information provided along with the satellite data.

With regard the optical imagery, two kinds of inputs are foreseen. The first one corresponds to the MERIS time series as they will be produced by the CCI-LC pre-processing chain. For organization reasons, it cannot be released at the start of the RRP but it will be made available after one month. In order to be able to start developing classification algorithms, the MERIS dataset which was used as input in the GlobCover project is provided at the start of the RRP.

Furthermore, a 5-year SPOT-VEGETATION (SPOT-VGT) dataset is also included in the RRDP in order to complement the MERIS time series.

Table 11. Input satellite dataset useful for the classification and included in the CCI-LC RRDP

Product	**Specification**	**Format and size**
ENVISAT MERIS FRS surface directional reflectance products (generated by the pre-processing chain)	Surface directional reflectance (not BRDF-corrected) products 300 m spatial resolution 15 spectral channels 3 spatial subsets all over the world	NetCDF ~2TB per year and all subsets
ENVISAT MERIS FRS 10-day surface reflectance composites (generated by the pre-processing chain)	Surface reflectance (BRDF-corrected) composites 300 m spatial resolution 15 spectral channels 3 spatial subsets all over the world	NetCDF ~500GB per year and all subsets
ENVISAT MERIS FRS 15-day surface reflectance composites (input dataset of the GlobCover 2005 project)	Surface reflectance (BRDF-corrected) composites 300m spatial resolution 3 spectral channels (+ status map) 7 spatial subsets all over the world	GeoTiff ~10 GB for all subsets (zip files)
SPOT-VGT 1-km daily (S1) products	Surface reflectance daily images 1-km spatial resolution 4 spectral channels (+ flags) 7 spatial subsets all over the world	GeoTiff ~60 GB/year for all subsets (zip files)
ENVISAT ASAR ASA_WSM_1P	Georeferenced calibrated intensity products Full swath 75-m spatial resolution C band 7 spatial subsets all over the world	ENVISAT

Table 12. Input auxiliary dataset useful for the classification and included in the CCI-LC RRDP

Product	Specification	Format
CCI global land cover map legend	Legend expressed in LCCS and in PFT expected to depict the produced land cover map	This document
Water bodies definition	Technical specification of the water bodies product	This document
Urban areas definition	Technical specification of the water bodies product	This document

As mentioned in the tables, the data are provided over 7 spatial subsets which define the RRob study areas for the classification. These spatial subsets were selected to be representative of the variety and the most critical ecological and climatic conditions. The size of the study areas for the optical (MERIS and SPOT-VGT) time series is equivalent to 1500 km *1500 km while smaller ones (500 km * 500 km) have been defined, inside the larger ones, for the SAR dataset. They are represented in Figure 6 and their geographic extent is detailed in Table 13. Like for the pre-processing, participants are requested to work **over minimum 4 spatial subsets** in order to demonstrate the spatial consistency and the generalization potential of their methods.

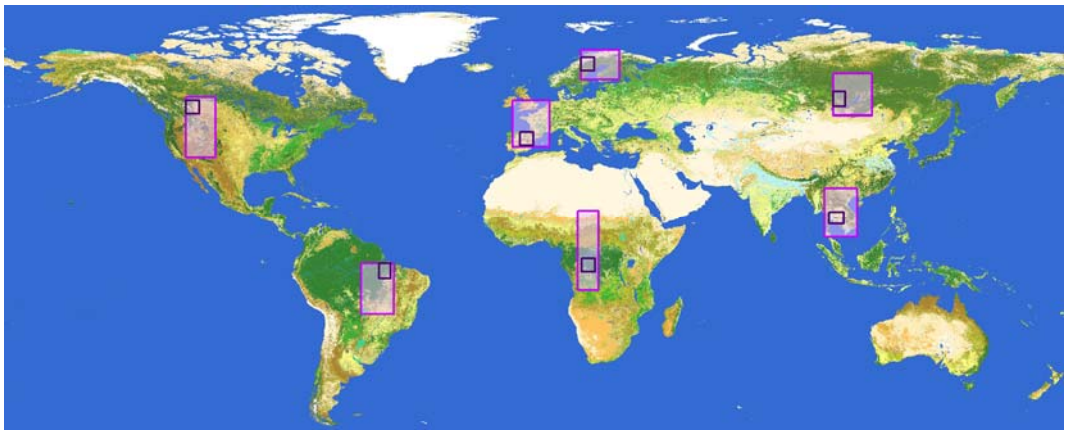

Figure 6. Localization of the 7 study areas used in the RRob for the classification algorithms. Pale areas are valid for the optical dataset while the darker rectangles, including in the pale ones, stand for the SAR dataset

Table 13. Name and extent of the study areas used to test classification algorithms in the RRob, using both optical and SAR satellite dataset. Study areas for optical dataset which are common with the pre-processing ones are highlighted in bold

Subset ID	Subset name	Dataset	Upper left corner	Lower right corner	Size (km²)
1	**North America**	**Optical**	**117°W, 55°N**	**107°W, 35°N**	**2 473 088**
		SAR	117°W, 54°N	112.5°W, 49.5°N	249 500
2	South America	Optical	58°W,1°S	47°W,18°S	2 311 470
		SAR	52°W, 1°S	48°W, 6°S	246 420
3	**Western Europe**	**Optical**	**7°W, 54°N**	**5.5°E, 38.5°N**	**2 395 249**
		SAR	4.5°W, 43.5°N	0°E, 39°N	249 500
4	Scandinavia	Optical	16°E, 71°N	29°E, 61°N	1 607 952

		SAR	16°E, 68.5°N	20.5°E, 64°N	249 500
5	Eastern Europe	Optical	101°E, 63°N	114°E, 49°N	2 251 422
		SAR	101°E, 57°N	105°E, 52°N	246 420
6	**Africa**	**Optical**	**15°E, 17°N**	**22°E, 10°S**	**2 336 757**
		SAR	16.3°E, 0.7°N	20.8°E, 3.8°S	238 411
7	Southeast Asia	Optical	98°E, 24.5°N	109°E, 8°N	2 271 111
		SAR	99.5°E, 16.5°N	104.5°E, 12.5°N	246 420

Over the 7 classification study areas related to optical satellite imagery, three are common with the pre-processing ones (Western Europe, North America and Africa). The input dataset made of MERIS time series generated by the pre-processing chain will only be provided over these 3 subsets. Therefore, **one of the four mandatory subsets to be tested should be one of these**.

4.2. Output dataset

For each algorithm included in the RRob activity (see section 2.2.2), a specific output is foreseen:

- For the generic and global classification algorithm, a land cover map is expected described the legend provided as auxiliary data;
- For the class-specific classification algorithms, two kind of outputs can be produced (not necessarily together):
    - a map of water bodies, meeting the specifications provided as auxiliary data;
    - a map of urban areas, meeting the specifications provided as auxiliary data.

These outputs are summarized in Table 14, along with their format. In addition, since the RRob aims at enabling the inter-comparison of varying algorithms, participating algorithms must be well-documented. Each algorithm should be provided with a report (word or pdf format) describing the developed methodologies and the products.

Table 14. Output dataset expected from the classification RRob

Product	Tested algorithm	Number of products	Format
Land cover map	Generic and global classification algorithm, based on optical (MERIS and/or SPOT-VGT) time series	Minimum 4 (including at least Europe, North-America or Africa)	GeoTiff
Water bodies map	Water bodies detection algorithm, using SAR dataset (with or without optical time series)	Minimum 4	GeoTiff
Urban areas map	Urban areas detection algorithm, using SAR dataset (with or without optical time series)	Minimum 4	GeoTiff

The technical specifications with which the classification RRob outputs have to be in line are described in detail in the following sections.

4.2.1. Product identification and filename convention

The file names of the classification RRob outputs should follow the following naming convention:

File name = <type>-<subset>-<centre>-<proc time>.tiff

where <type> = <project>-<code>-<level>-<resolution>-<epoch>

The dash "-" is the separator between name components. The different name components are defined in Table 15.

Table 15. Components that make the name of the LC maps delivered by the CCI-LC project

Field	Signification	Value
project	Project Acronym and work phase	CCI-LC-RRob (constant)
code	Product code identifier	With the following convention: "LCMap", for the output of the generic and global classification algorithm generic; "WBmap", for the output of the water bodies classification algorithm; "UAmap", for the output of the urban areas classification algorithm.
level	Processing level	L4 (constant)
resolution	Spatial resolution	300m for the LCMap and the UAMap (constant) 150m for the WB map (constant)
epoch	Epoch of the product	2005 to 2009
subset	Two numbers identifying the study areas	"vvhh" where: "01" is North America; "02" is South America; "03" is Western Europe; "04" is Scandinavia; "05" is Eastern Europe; "06" is Africa; "07" is Southeast Asia.
centre	Name of the processing centre, generator of the product	[a-zA-Z0-9]
proc time	Processing time, nnn to make file name unique if necessary	"yyyyMMddHHmmssnnn" where: "yyyy" is the processing year "MM" is the processing month "dd" is the processing day "HH" is the processing hour "mm" is the processing minute "ss" is the processing second "nnn" is added to make file name unique if necessary

4.2.2. Thematic content and legend

With regard to the land cover map (output of the generic and global classification algorithm generic), it has to be made with a dominance of MERIS time series and could be complemented by the SPOT-VGT dataset, either to compensate for a possible lack of MERIS acquisitions and/or to increase the MERIS spectral resolution by providing a SWIR channel.

The classification algorithm has to transform the surface reflectance mosaics into products which are meaningful for the climate modellers' point of view. Two criteria are of importance: the accuracy and the thematic content.

The accuracy is evaluated both through overall accuracy and temporal stability. The targeted overall accuracy is set to 80%. The accuracy is also assessed in terms of "temporal stability", i.e. the land cover is expected to be mapped in a consistent way over time, independently of inter-annual variability.

Two different versions of the product are expected: the first one based on the LCCS-based CCI-LC legend (Table 16) and the second one expressed in Plant Functional Types (PFTs) (Table 17).

Table 16. Legend of the global LC maps based on LCCS

Class	Description
1	Cropland, rainfed
2	Cropland, irrigated or post-flooding
3	Mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%)
4	Mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%) / cropland (< 50%)
5	Tree cover, broadleaved, evergreen, closed to open (>15%)
6	Tree cover, broadleaved, deciduous, closed (>40%)
7	Tree cover, broadleaved, deciduous, open (15-40%)
8	Tree cover, needleleaved, evergreen, closed (>40%)
9	Tree cover, needleleaved, evergreen, open (15-40%)
10	Tree cover, needleleaved, deciduous, closed (>40%)
11	Tree cover, needleleaved, deciduous, open (15-40%)
12	Tree cover, mixed leaf type (broadleaved and needleleaved)
13	Mosaic tree and shrub (>50%) / herbaceous cover (< 50%)
14	Mosaic herbaceous cover (>50%) / tree and shrub (<50%)
15	Shrubland
16	Grassland
17	Sparse vegetation (tree, shrub, herbaceous cover)
18	Tree cover, flooded, fresh or brakish water
19	Tree cover, flooded, saline water
20	Shrub or herbaceous cover, flooded, fresh-saline or brakish water
21	Urban
22	Bare areas

Class	Description
23	Water
24	Snow and ice
25	No data

Table 17. Legend of the global LC maps expressed in PFTs

Class	Description
1	Broadleaved, evergreen
2	Broadleaved, deciduous
3	Needleleaved, evergreen
4	Needleleaved, deciduous
5	Shrubs
6	Grassland
7	Cropland, irrigated
8	Cropland, non-irrigated
9	Wetland
10	Barren land or sparse vegetation
11	Urban
12	Water
13	Snow & Ice

With regard to the water bodies' map (output of the class-specific discrimination algorithm using SAR data), it should contain the following thematic information:

- A pixel is water or land;
- In the case of a pixel identified as "water", the water is permanent or seasonal;
- In the case of a pixel identified as "permanent water", the pixel corresponds to an open water body or a permanent wetland;
- In the case of pixel identified as "seasonal water", the pixel corresponds to an open water body, an irrigated area, a flooding or a seasonal wetland.

An example of legend is provided in Table 18.

Table 18. Legend of the water bodies' map

	1st level	2nd level
10	Permanent water bodies	
11		Open water bodies
12		Permanent wetland
20	Seasonal (non-permanent) water bodies	
21		Open water bodies

	1st level	2nd level
22		Flooding
23		Irrigated area
24		Seasonal wetland

With regard to the difference between permanent and seasonal water bodies, the critical information to provide is the date of detection (start and end dates of detection).

As for the identification of irrigated area, the detection date is highly valuable additional information.

As for the wetlands identification, recognizing that defining wetlands is a really complicated task, it is proposed to RRob participants to determine their wetlands typology according to the results of their classification. However, two criteria should be considered as important, which are:

- the distinction between inland and coastal wetlands;
- the life form of the wetland vegetation (e.g. trees or herbaceous vegetation).

With regard to the urban areas' map (output of the class-specific discrimination algorithm using SAR data), it should contain the following thematic information (Table 19):

- A pixel is urban or not;
- In the case of a pixel identified as "urban", the urban area is dense or scattered.

Recognizing that urban detection at global scale using medium spatial resolution is a challenging task, the focus should be placed on the first level of information (urban or not). The second one can be considered as a refinement.

Table 19. Legend of the urban areas' map

	1st level	2nd level
10	Urban areas	
11		Dense urban areas
12		Scattered urban areas

4.2.3. Quality flags

The quality of the three different maps needs to be documented through quality flags. At least three kinds of flags should be provided:

- a flag indicating if the pixel has been processed or not;
- if the pixel has been processed, a flag indicating the number of observations used for the classification;
- if the pixel has been processed, a flag informing about the confidence in the classification diagnostic, which should be related to the developed algorithm (e.g. distance metric, probability value, etc.).

4.2.4. Metadata

Metadata of the classification RRob outputs should be provided as a GML file. The corresponding structure definition can be found in the OpenGIS GML Application Schema for Earth Observation Products [RD-6]. GML also provides XML schema definitions for XML validation.

4.3. Validation dataset

The comparison between the different algorithms proposed through the RRP is made thanks to a set of validation data, which are presented in Table 20.

Table 20. Validation dataset to use to evaluate the performance of the classification algorithms

Product to validate	**Validation dataset**	**Specification**	**Format and size**
Generic land cover map, urban areas map	GlobCover 2005 land cover map (version 2.2)	Land cover map at 300m spatial resolution, representative of the period from December 2004 to June 2006 Over the 7 classification study areas Provided along with a legend translation to be compatible with the CCI-LC legend Available at: http://ionia1.esrin.esa.int/	GeoTiff ~160 MB for all subsets (zip files)
Generic land cover map, urban areas map, water bodies map	National or continental land cover maps	Corine Land Cover maps from 2000 and 2006 (map derived from photo-interpretation of high spatial resolution imagery – source: EEA). Available at: http://www.eea.europa.eu/data-and-maps National land cover database of United States from 2001 (map derived from Landsat imagery – source: USGS). Available at: http://landcover.usgs.gov/landcoverdata.php Africover maps from 2000 (map derived from photo-interpretation of high spatial resolution imagery – source: FAO). Available at: http://www.africover.org/ GLC2000 regional maps (maps derived from SPOT-VGT imagery – source: JRC). Available at: http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php	ESRI Grid, GeoTiff
Water bodies map	SWBD	Representative of the year 2000 Water mask over the 7 classification study areas Available at: http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/Shuttle_Radar_Topography_Mission_Water_Body_Dataset	Shapefiles
Generic land cover map, urban areas map	Points of the GlobCover validation database	400 points representative of the year 2005 and distributed over the 7 classification study areas Points with LCCS description that can be easily translated in the CCI-LC legend **NOT DISTRIBUTED**	Access file

As indicated in Table 20, only maps are made available to RRob participants, who are invited to provide inter-comparison results (with these maps and/or other dataset) in a validation report. If other dataset are used, they shall be fully documented in the report.

With regard to the points of the GlobCover validation database, they are not distributed since they will be used by the CCI-LC consortium to perform the final comparison between all algorithms proposed in the RRob activity. Furthermore, its non-distribution allows being sure that none algorithm will be tuned towards this dataset and thus, ensures a completely independent validation.

5. Protocol

5.1. Participation

The round-robin is an activity open to everyone.

Participants to the RRP can be either Algorithm Providers (AP) or Data Providers (DP).

- Algorithm Providers are scientists who develop algorithms for processing the input datasets. One contact scientist name is required for each AP.
- Data Providers are scientists who contribute in data which can be used either as ancillary or reference datasets. A contact will be identified for each provided data, which will be protected via the CCI-LC Data Policy (see section 0).

To enable their integration in the RRP, AP are expected to:

- Submit, to the CCI-LC consortium, a proposal which presents research objectives, principal investigator and co-investigators (if any), required input dataset, methodology, expected results and deliverables. The access to the Round-Robin Data Package is conditioned by the approval of this proposal.
- Commit to actively participate to the whole RRob activity. This involves:
    - Downloading the requested data;
    - Developing and applying their algorithms to minimum 4 study areas (in order to demonstrate the global consistency and generalization potential of the proposed algorithms);
    - Delivering the CCI-LC consortium outputs which are compliant with requested format, with the RRob activity timeline and with the requirements in terms of deliverables. With regard to this latter concern, 3 deliverables need to be provided:
        - the products (in the requested format);
        - an Algorithm Theoretical Baseline Document containing a description of the developed algorithm and mentioning any previous applications of the algorithm. Any aspects that may affect independence of the tests (e.g. previous tuning/validation of the algorithm with part of the RRP) should also be mentioned to allow a correct scientific analysis. This documentation can make reference to published document.
        - a Validation Report describing the algorithm performance and the products accuracy assessed using the validation dataset which those provided in the Round-Robin Data Package or their own ones. In this latter case, the validation dataset should be fully documented.
- Accept and sign the CCI-LC Data Policy (see section 0).
- Acknowledge that no funding will be provided by the CCI-LC consortium for this research activity.

5.2. Data access to the round-robin data package

The RRDP consist of the input (satellite and ancillary) dataset described in sections 3.1 and 4.1, of the validation dataset described in sections 3.3 and 4.3 and of this document. It is distributed from the CCI-LC website from **April 2011** with four levels of access:

1) Secured access ftp site

The following dataset are available by password-protected FTP access to RR participants:

- ENVISAT MERIS FRS 15-day surface reflectance composites (input dataset of the GlobCover 2005 project);
- ENVISAT ASAR ASA_WSM_1P products;
- Monthly Aerosol Optical Depth;
- GlobCover 2005 validation dataset;
- Round-Robin protocol (this document).

2) Copy on external disks

With regard to the large-size dataset, the CCI-LC consortium asks participants to send external disks so that copies can be made. This level of access concerns the following dataset:

- ENVISAT MERIS FRS_1P products with the auxiliary data;
- ENVISAT MERIS FRS SDR products;
- ENVISAT MERIS FRS 10-day surface reflectance composites.

3) Product whose access is regulated by an external institution

The access to the SPOT-VGT dataset – which is not the intellectual property of the CCI-LC consortium – will require introducing a request form to the SPOT-VGT programme. This form is available on the following website:

http://www.spot-vegetation.com/home/distribution/distribution03082004.htm.

Once the request approved by the SPOT-VGT programme, the dataset will be made available by the CCI-LC consortium through the above-mentioned secured access ftp site.

The access to the in-situ data of the CEOS land sites demands the registration on the CEOS CAL/VAL Portal: http://calvalportal.ceos.org/cvp/web/guest/get-general-information

4) Product which are freely and easily available on the web

This access concerns data that are open and easily available on the web. Data specification and appropriate websites are mentioned in this Round-Robin Protocol document. This level of access concerns the following dataset:

- SRTM Water Bodies Dataset (SWBD);
- Getasse 3.0 Digital Elevation Model;
- In-situ data for aerosol - AERONET stations

The CCI-LC RRDP will not be made available at once, at the start of the RRob activity. In total, **four RRDP will be distributed**, as detailed in Table 21. The timing of these deliveries is described in section 5.3.

Table 21. RRDP delivered during the RRob period

# RRDP	RRDP description
1	Pre-processing (for the pixel identification, aerosol retrieval and atmospheric correction): - input satellite dataset: ENVISAT MERIS FRS_L1B products - input auxiliary dataset - validation dataset Optical classification: - ENVISAT MERIS FRS 15-day surface reflectance composites (input dataset of the GlobCover 2005 project) - SPOT-VGT 1-km daily (S1) products - input auxiliary dataset - validation dataset
2	Pre-processing (for the BRDF correction and compositing algorithm): - ENVISAT MERIS FRS surface directional reflectance products, over Europe Optical classification: - ENVISAT MERIS FRS surface directional reflectance products (generated by the pre-processing chain), over Europe - ENVISAT MERIS FRS 10-day surface reflectance composites (generated by the pre-processing chain), over Europe
3	Pre-processing (for the BRDF correction and compositing algorithm): - ENVISAT MERIS FRS surface directional reflectance products, over North America and Africa Optical classification: - ENVISAT MERIS FRS surface directional reflectance products (generated by the pre-processing chain), over North America and Africa - ENVISAT MERIS FRS 10-day surface reflectance composites (generated by the pre-processing chain), over North America and Africa
4	SAR-based classification: - input (satellite and auxiliary) dataset - validation dataset

5.3. Timeline

The official time period for the RRob activity is between April and August 2011. Two different periods should be distinguished, the first one, from April to June, being related to the pre-processing and the optical classification algorithms and the second one, from June to August, for the SAR-based classification algorithms. An overall summary of the RRob timing is given in Table 22 and illustrated in Figure 7.

Table 22. Timing of the RRob activity

Activity	**Date**
Start of the RRP for the pre-processing and the optical classification algorithms	1st April
Release of the RRDP #1	1st April
Release of the RRDP #2	15th April
Release of the RRDP #3	15st May
Start of the RRP for the SAR-based classification algorithms	1st June
Release of the RRDP #4	1st June
End of the RRP for the pre-processing and the optical classification algorithms	30th June
Delivery of results, ATBD and validation report for the pre-processing and optical classification algorithms	30th June
End of the RRP for the pre-processing and the optical classification algorithms	15th August
Delivery of results, ATBD and validation report for the pre-processing and optical classification algorithms	15th August

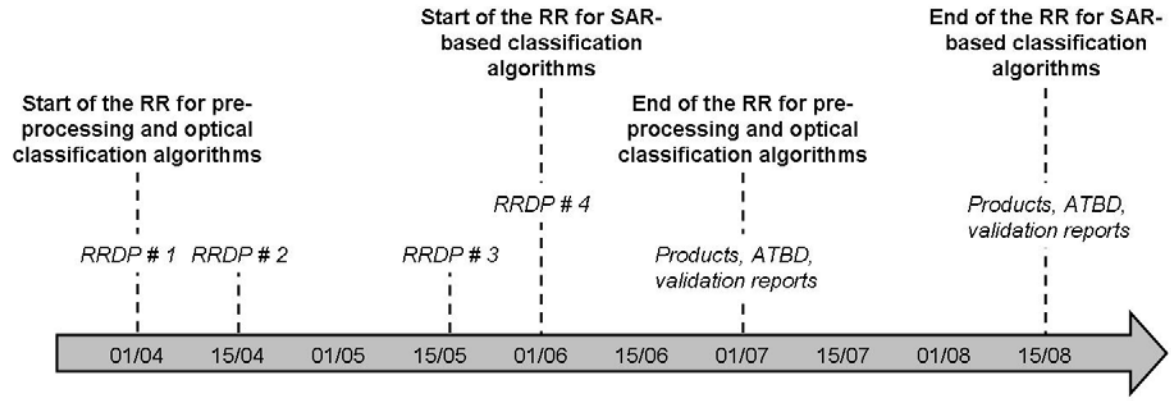

Figure 7. Timing of the CCI-LC RRob activity

5.4. Selection criteria

As already mentioned, the objective of the RRob activity is the selection of the "best" algorithm(s) or combination of algorithms. Two categories of criteria will guide this selection:

- General criteria, which are related to the scope of the CCI-LC project:
  - Proposed methodologies should be valid at the global scale, i.e. results and algorithms performance should be similar for the 4 study areas tested;
  - Proposed methodologies should allow handling large amounts of data;

- Proposed methodologies should be computing efficient, i.e. they should show reasonable processing time in order to allow delivering the CCI products within the project deadlines.

- Thematic criteria, which depend on each tested methodology and which are described in detail in the following sections.

5.4.1. Pre-processing algorithms

The different phases tested in the RRob activity (pixel identification, atmospheric correction including the aerosol retrieval, BRDF correction and the compositing algorithm) will be assessed with regard to the accuracy of the product, the error estimation and the feasibility of processing chain. This assessment will be made through two different processes, which are an independent statistical product validation using in-situ data and an analysis of the generated surface reflectance time series.

Validation of the products with in-situ data will be performed using statistical analysis to generate correlation figures and confusion matrices. This will allow the statistical evaluation of the developed algorithms performance [AP-8]. For each validation approach, the following accuracy measurements will be provided at the global level (i.e. for the minimum 4 study areas considered together): overall accuracy, user's and producer's accuracy as well as Cohen's kappa, Scott's pi and Krippendorf's alpha coefficient.

Furthermore, the generated surface reflectance time series will be investigated with regard to the temporal variation of the mean and variance over selected areas (including "reference" areas such as pixels covered by dark vegetation or desert).

The "best" algorithm will be the one that is associated with the highest overall accuracy figures at the global scale and is able to provide error estimation. As already mentioned, the feasability of the pre-processing chain with respect of the time scale and the large quantity of data will also be considered for the selection of the "best" algorithm.

5.4.2. Classification algorithms

The "best" algorithm(s) or combinations of algorithms for the different phases of the RRob activity (sensor selection, compositing period, spatial stratification approach and classification methodologies) will be identified by an assessment of the produced land cover maps in terms of (i) macroscopic errors and (ii) accuracy.

The **macroscopic errors** will be identified through a **qualitative validation** based on a visual comparison with a reference. Each land cover product generated by a particular classification algorithm will be compared to this reference and the quality of the outputs will be qualitatively documented in terms of thematic content, spatial consistency and overall quality. The reference will consist in existing land cover datasets which have been – if possible – consistently validated (i.e. for which the overall accuracy is known and the strengths and weaknesses are documented). In particular, the GLC2000 (global and/or regional) maps as well as the GlobCover products will be used.

Products **accuracy** will be evaluated through an **independent statistical product validation**, with the objective to determine the map's "fitness for use". It will be achieved by the CCI-LC consortium, using the GlobCover validation database (see Table 20). This database counts around 400 points representative of the year 2005 and distributed over the 7 classification study areas. Each point has been characterized in terms of LCCS classifiers by independent experts during the GlobCover 2005 project [RD-3]. Map's "fitness for use" will be derived by comparing the different proposed products

with this set of points and by generating confusion matrices. For each proposed approaches, the following **accuracy measurements** will be provided at the global level (i.e. for the minimum 4 study areas considered together) and by legend: overall accuracy and user's and producer's accuracy. The "best" methodology will be the one that is associated with the highest overall accuracy figures at the global scale. User's and producer's accuracies will be used as indicators to ensure that the methodologies are valid for all classes.

Table 23 shows the interest of these two distinct assessment processes for the different phases tested in the RRob activity.

Table 23. Detail of selection processes for the algorithms tested in the RRob classification activity

	Macroscopic errors assessment	**Accuracy assessment**
Sensor selection	Inform about: - the interest of the selected sensor combination in terms of landscape patterns identification and spatial consistency - possible links between classification errors and input data quality/quantity	Provide quantitative information about: - the sensor selection on the classification algorithm performance
Compositing period	Inform about: - the radiometric quality of the data - the capacity of the seasonal composites (based on the basic composites) to identify major landscape patterns - possible links between classification errors and input data quality/quantity	Provide quantitative information about: - the impact of the basic compositing period on the classification algorithm performance
Spatial stratification	Inform about: - the strata limits visibility - the spatial consistency of the classification on both sides of the strata limits (e.g. anomalous absence of a class in one stratum)	Provide quantitative information about: - the impact of the stratification approach the classification algorithm performance
Classification algorithm	Inform about: - the nature of classification errors (wrong label, missing classes, wrong position of the boundary between classes, disappearance of small patches) - evaluating the classification algorithm performance over space	Provide quantitative information about: - the classification algorithm overall performance - the classification algorithm performance by class

6. Annex A. Data policy for the round-robin activity [1]

Access to the CCI-LC Round-Robin Data Package can be accepted if the participant agrees with the conditions mentioned below. In case the proposal is submitted by a consortium, all of its members should sign this agreement.

Conditions

1. The request is made by a 'bona fide' research institution or consortium.
2. The research, based on these data, is carried out in the frame of the Round-Robin activity of the European Space Agency Climate Change Initiative Land Cover project.
3. Research is well-documented in terms of objectives, principal investigator, co-investigators, work plan, schedule, expected results and deliverables.
4. The products delivered in the Round-Robin Data Package and any information processed, reconstructed, derived, enhanced, or reproduced from such data may be used, reproduced, and disclosed only for the purpose of research authorised in the framework of the Round-Robin activity of the European Space Agency Climate Change Initiative Land Cover project.
5. The products delivered in the Round-Robin Data Package and any information processed, reconstructed, derived, enhanced, or reproduced from such data will not be used, reproduced, or disclosed for commercial or operational purposes. Accordingly, in the event of the breach or threatened breach of this agreement by the Round-Robin participant, the consortium of the European Space Agency Climate Change Initiative Land Cover project may exercise all applicable legal remedies.
6. The cost of the product medium, packing and expedition is borne by the user.
7. This agreement covers all the products that are part of the delivered Round-Robin Data Package.
8. By signing this agreement, the user agrees with all the terms listed above.

_____

Signature of the interested parties Date

[1] Data policy will be proposed as a form that RRob participant need to sign and send to the CCI-LC consortium. Their participation to the RRob activity is conditioned by their agreement with this data policy

7. Annex B. Declaration of interest for the round-robin activity [2]

In order to enable their integration into the CCI-LC round-robin, interested parties are expected to commit themselves to actively participate in the whole activity. This involves:

- Submitting, to the CCI-LC consortium, a proposal which briefly presents research objectives, principal investigator and co-investigators (if any), required input dataset, methodology, expected results and deliverables.
- Downloading the requested data;
- Developing and applying their algorithms to minimum 4 study areas (in order to demonstrate the global consistency and generalization potential of the proposed algorithms);
- Delivering the CCI-LC consortium outputs which are compliant with requested format, the round-robin activity timeline and the requirements in terms of deliverables. With regard to this latter concern, 3 deliverables need to be provided:
    - the products (in the requested format);
    - an Algorithm Theoretical Baseline Document containing a description of the developed algorithm and mentioning any previous applications of the algorithm. Any aspects that may affect independence of the tests (e.g. previous tuning/validation of the algorithm used in the round-robin activity) should also be mentioned to allow a correct scientific analysis. This documentation can make reference to published documents.
    - a Validation Report describing the algorithm performance and the products accuracy assessed using validation datasets (either those provided in the Round-Robin Data Package or their own ones). In this latter case, the validation dataset should be fully documented.
- Accept and sign the CCI-LC Data Policy.
- Acknowledge that no funding will be provided by the CCI-LC consortium for this research activity.

_____

Signature of the interested parties Date

_____

[2] Declaration of interest will be proposed as a form that RRob participant need to sign and send to the CCI-LC consortium. Their participation to the RRob activity is conditioned by their signature.

8. Annex C. Proposal for the round-robin activity [3]

Algorithm(s) selection

Please select the algorithm(s), which are you proposed for the participation in the round robin activity:

- □ Pre-processing algorithms
  - □ Pixel Identification
  - □ Aerosol retrieval
  - □ SDR retrieval
  - □ BRDF - Correction
  - □ Compositing
  - □ Entire pre-processing chain
- □ Classification algorithms
  - □ Generic and global classification chain (based on optical imagery)
    - □ Spatial stratification
    - □ Classification algorithm
    - □ Entire chain
  - □ Water bodies classification chain (using SAR and optical imagery)
    - □ Sensor selection
    - □ Spatial stratification
    - □ Classification algorithm
    - □ Entire chain
  - □ Urban areas classification chain (using SAR and optical imagery)
    - □ Sensor selection
    - □ Spatial stratification
    - □ Classification algorithm
    - □ Entire chain

[3] Data policy will be proposed as a form that RRob participant need to sign and send to the CCI-LC consortium. Their participation to the RRob activity is conditioned by their agreement with this data policy.

Short descriptive note about your proposed algorithm(s) or key references

(Not mandatory but we would appreciate you doing fill in this section)

Results and deliverables

Please select the results and deliverables expected from your participation in the round robin activity:

- □ Pre-processing algorithms
  - □ Pixel Identification map, ATBD and validation report
  - □ Aerosol retrieval map, ATBD and validation report
  - □ Surface directional reflectance product, ATBD and validation report
  - □ Surface reflectance product, ATBD and validation report
  - □ Surface reflectance composite - temporal resolution: _____, ATBD and validation report
- □ Classification algorithms
  - □ Global and generic map, ATBD and validation report
  - □ Global water bodies map, ATBD and validation report
  - □ Global urban map, ATBD and validation report

Data request

Please, select the data you are interested in (more information about data specification can be found in the Round-Robin Protocol). Note that results are expected for each requested dataset.

Pre-processing				
Data	**Study area**	**Date of delivery**	**Medium of delivery**	**Interest**
Input Satellite: ENVISAT MERIS FRS_L1P products	North America	1 April 2011	External disks to send to Brockmann Consult	☐
	Western Europe	1 April 2011	External disks to send to Brockmann Consult	☐
	Africa	1 April 2011	External disks to send to Brockmann Consult	☐
	Central Asia	1 April 2011	External disks to send to Brockmann Consult	☐
	Northwest Asia	1 April 2011	External disks to send to Brockmann Consult	☐
	Australia	1 April 2011	External disks to send to Brockmann Consult	☐
Input Satellite + Validation: ENVISAT MERIS FRS surface directional reflectance products	North America	15 May 2011	External disks to send to Brockmann Consult	☐
	Western Europe	15 April 2011	External disks to send to Brockmann Consult	☐
	Africa	15 May 2011	External disks to send to Brockmann Consult	☐
Input auxiliary: ENVISAT Attitude and Orbit	All subsets	1 April 2011	External disks to send to Brockmann Consult	☐
Input auxiliary: SRTM Water Bodies Dataset	All subsets	1 April 2011	Available online at: http://dds.cr.usgs.gov/srtm/version2_1/SWBD/	N/A
Input auxiliary: Monthly Aerosol Optical Depth	All subsets	1 April 2011	Secured ftp	☐
Input auxiliary: Getasse 3.0 Digital Elevation Model	All subsets	1 April 2011	Available online at: http://earth.esa.int/services/amorgos/download/getasse/	N/A
Validation : ENVISAT MERIS FRS 10-day surface	North America	15 May 2011	External disks to send to Brockmann Consult	☐
	Western Europe	15 April 2011	External disks to send to Brockmann Consult	☐

reflectance composites	Africa	15 May 2011	External disks to send to Brockmann Consult	☐

Classification				
Data	**Study area**	**Date of delivery**	**Medium of delivery**	**Interest**
Input Satellite: ENVISAT MERIS FRS 15-day surface reflectance composites (input dataset of the GlobCover 2005 project)	North America	1 April 2011	Secured ftp	☐
	South America	1 April 2011	Secured ftp	☐
	Western Europe	1 April 2011	Secured ftp	☐
	Scandinavia	1 April 2011	Secured ftp	☐
	Eastern Europe	1 April 2011	Secured ftp	☐
	Africa	1 April 2011	Secured ftp	☐
	Southeast Asia	1 April 2011	Secured ftp	☐
Input Satellite : ENVISAT MERIS FRS surface directional reflectance products (generated by the pre-processing chain)	North America	15 May 2011	External disks to send to Brockmann Consult	☐
	Western Europe	15 April 2011	External disks to send to Brockmann Consult	☐
	Africa	15 May 2011	External disks to send to Brockmann Consult	☐
Input Satellite: ENVISAT MERIS FRS 10-day surface reflectance composites (generated by the pre-processing chain)	North America	15 May 2011	External disks to send to Brockmann Consult	☐
	Western Europe	15 April 2011	External disks to send to Brockmann Consult	☐
	Africa	15 May 2011	External disks to send to Brockmann Consult	☐
Input Satellite: SPOT-VGT 1-km daily (S1) products	North America	1 April 2011	Secured ftp	☐
	South America	1 April 2011	Secured ftp	☐
	Western Europe	1 April 2011	Secured ftp	☐
	Scandinavia	1 April 2011	Secured ftp	☐
	Eastern Europe	1 April 2011	Secured ftp	☐
	Africa	1 April 2011	Secured ftp	☐

	Southeast Asia	1 April 2011	Secured ftp	☐
Input Satellite: ENVISAT ASAR ASA_WSM_1P	North America	1 June 2011	Secured ftp	☐
	South America	1 June 2011	Secured ftp	☐
	Western Europe	1 June 2011	Secured ftp	☐
	Scandinavia	1 June 2011	Secured ftp	☐
	Eastern Europe	1 June 2011	Secured ftp	☐
	Africa	1 June 2011	Secured ftp	☐
	Southeast Asia	1 June 2011	Secured ftp	☐
Validation: GlobCover 2005	All subsets	1 April 2011	Available online at: http://ionia1.esrin.esa.int/	N/A
Validation: Corine Land Cover	All subsets	1 April 2011	Available online at: http://www.eea.europa.eu/data-and-maps	N/A
Validation: National land cover database of United States	All subsets	1 April 2011	Available online at: http://landcover.usgs.gov/landcoverdata.php	N/A
Validation: Africover maps	All subsets	1 April 2011	Available online at: http://www.africover.org/	N/A
Validation: GLC2000 products	All subsets	1 April 2011	Available online at: http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php	N/A
Validation: SRTM Water Bodies Dataset	All Subsets	1 April 2011	Available online at: http://dds.cr.usgs.gov/srtm/version2_1/SWBD/	N/A

_____

Signature of the interested parties Date