Lakes_CCI+ - Phase 2

D3.1. System Specification Document (SSD)

Reference: CCI-LAKES2-0016-SSD

Issue 3.1 – 17/10/2023

Contract number : 4000125030/18/I-NB –Lakes_cci

11 rue Hermès, Parc Technologique du Canal – 31520 Ramonville Saint-Agne, France
Tel +33 (0)5 61 39 47 00 Fax +33(0)5 61 75 10 14
www.cls.fr

CHRONOLOGY ISSUES

Issue	Date	Object	Written by
0.1	05 May 2019	Draft version	B. Calmettes
1.0	24 March 2020	Version 1.0	B. Calmettes, S Simis, C Merchant, C Duguay, H Yésou, J-F Crétaux
1.1	24 April 2020	Revision following ESA review	B. Calmettes, S Simis, C Merchant, C Duguay, H Yésou, J-F Crétaux
1.2	15 Feb. 2021	Update following the new algorithms used for LIC date generation	C. Duguay
1.3	29 April 2021	Typo corrections following ESA review	B. Coulon
2.0	18 April 2022	Accompanying CRDPV2.0.1	
3.0	6 Sept. 2023	Linked with CRDP V2.1	B.Calmettes and S. Simis with contributions of Lakes_CCI project.
3.1	17/10/23	Revision following ESA review	B.Calmettes and S. Simis with contributions of Lakes_CCI project.

Checked by	Stefan Simis – PML J-F Cretaux - LEGOS	Stefan Simis Jean-francois Cretaux
Approved by	Alice Andral - CLS	A. Andral
Authorized by	Clément Albergel - ESA	Clement Albergel

DISTRIBUTION

Company	Names	Email
ESA	Clément Albergel	clement.albergel@esa.int
BC	Carsten Brockman	carsten.brockmann@brockmann-consult.de
BC	Dagmar Müller	dagmar.mueller@brockmann-consult.de
BC	Jorrit Scholze	jorrit.scholze@brockmann-consult.de
BC	Kerstin Stelzer	kerstin.stelzer@brockmann-consult.de
BC	Martin Boettcher	martin.boettcher@brockmann-consult.de
BC	Olaf Danne	olaf.danne@brockmann-consult.de
CLS	Alice Andral	aandral@groupcls.com
CLS	Anna Mangilli	amangilli@groupcls.com
CLS	Beatriz Calmettes	bcalmettes@groupcls.com
CLS	Nicolas Taburet	ntaburet@groupcls.com
CLS	Christophe Fatras	cfatras@groupcls.com
CLS	Pierre Thibault	pthibaut@groupcls.com
CNR	Claudia Giardino	giardino.c@irea.cnr.it
CNR	Mariano Bresciani	bresciani.m@irea.cnr.it
CNR	Monica Pinardi	pinardi.m@irea.cnr.it
CNR	Marina Amadori	amadori.m@irea.cnr.it
CNR	Rossana Caroni	caroni.r@irea.cnr.it
CNR	Giulio Tellina	tellina.g@irea.cnr.it
H2O Geo	Claude Duguay	claude.duguay@h2ogeomatics.com
H2O Geo	Yuhao Wu	mark.wu@h2ogeomatics.com
H2O Geo	Jaya Sree Mugunthan	jayasree.mugunthan@h2ogeomatics.com
H2O Geo	Justin Murfitt	justin.murfitt@h2ogeomatics.com
HYGEOS	François Steinmetz	fs@hygeos.com
LEGOS	Jean-François Cretaux	jean-francois.cretaux@cnes.fr
LEGOS	Paul-Gérard Gbektom	paul.gerard.gbetkom@legos.obs-mip.fr
PML	Stefan Simis	stsi@pml.ac.uk
PML	Xiaohan Liu	liux@pml.ac.uk
PML	Nick Selmes	nse@pml.ac.uk
PML	Mark Warren	mark1@pml.ac.uk
Sertit	Hervé Yésou	Herve.yesou@unistra.fr
Sertit Sertit	Jérôme Maxant Sabrina Amsil	maxant@unistra.fr s.amzil@unistra.fr
Sertit	Rémi Braun	remi.braun@unistra.fr
UoR	Chris Merchant	c.j.merchant@reading.ac.uk
Bangor	Iestyn Woolway	iestyn.woolway@bangor.ac.uk
UoR	Laura Carrea	l.carrea@reading.ac.uk
UoS	Dalin Jiang	dalin.jiang@stir.ac.uk
UoS	Evangelos Spyrakos	evangelos.spyrakos@stir.ac.uk
UoS	Ian Jones	ian.jones@stir.ac.uk

LIST OF CONTENTS

LIST OF TABLES AND FIGURES

REFERENCE DOCUMENTS

RD-1	Algorithm Theoretical Basis Document	CCI-LAKES2-0012-ATBD
RD-2	Product Specification Document	CCI-LAKES2-0003-PSD
RD-3	User requirement Document	CCI-LAKES2-0005-URD
RD-4	Product User Guide	CCI-LAKES2-0021-PUG
RD-5	Product Validation and Intercomparison Report	CCI-LAKES2-0022-PVIR
RD-6	Data Access Requirement Document	CCI-LAKES-0017-DARD
RD-7	Theia/Hydroweb Project	http://hydroweb.theia-land.fr/

1 Introduction

The SSD describes the data processing system, including its design, performance, operating environment, and target hardware. The SSD also describes any supporting software developed, such as validation or visualisation tools.

The objective of the SSD is to describe:

- the functionality and architectural design of the processing system
- the engineering methodologies adopted, and the reasons for adopting these
- the main requirement specifications, functions and components of the processing system
- how the features of the processing system will solve certain problems, including any trade-off criteria
- the system life cycle design, how it is implemented and its performance
- source code version control and issue tracking
- error handling and reporting
- cost effectiveness of the chosen implementation (e.g. cloud vs. stand-alone)
- the maintenance costs and risks (in view of transferring to an operational system)
- any licensing constraints

To ensure this document is accessible, a user-friendly summary giving an overview of the general characteristics of the system is included.

Data generation in Lakes_cci is a system of systems, with one processing centre per ECV product. To make a consistent Lakes_cci system and to be able to generate products that meet user needs, existing processing chains are updated to provide seamless interaction with operations that generate the user-facing dataset.

This document contains the description of the existing systems. Each system is already implemented following user community requirements, and validation tests are performed.

Section 2 defines the main development objectives for the Lakes_cci from a system development point of view and its distinction as a system of systems. It also describes the work performed towards the harmonisation of the final product to deliver a consistent data set.

Subsequently, each contributing system is described in Section 3. This includes harmonisation efforts of each individual system to generate the Lakes ECV dataset following the requirements identified in the User Requirements Document [RD-3] These requirements are based on user needs, GCOS recommendations, team experience and literature review.

Section 4 provides a compliance overview regarding the system requirements from the SOW and the URD.

2 Lakes_cci processing system objectives

2.1 Objectives

The objective of the Lakes_cci project is to produce and validate a consistent data set of the variables grouped under the Lakes Essential Climate Variable. This involves both R&D activity and the application of rules to meet CCI data Standards. At the system level, the objective is to define the process and rules to generate a consistent data set so that the user can easily manipulate five products in the Lakes ECV:

- Lake Water Level (LWL)
- Lake Water Extent (LWE)
- Lake Surface Water Temperature (LSWT)
- Lake Ice Cover (LIC)
- Lake Water Leaving Reflectance (LWLR)

The sixth product, Lake Ice Thickness (LIT), is included separately in Lakes_cci V2.1.

In the Lakes_cci context, the specification of the production system will be useful for related activities in the Copernicus Climate Change service (C3S), which already implements some of the Lakes ECV products (LWL, LSWT) with periodic releases, and the Copernicus Land Monitoring Service which includes operational products (LSWT, LWLR, XX,...).

2.2 General overview

The function of the system is the production of a consistent, long-term, multi-mission and multi-sensor global data set.

Two teams are cooperating within the organizational framework to fulfil the mission objective of the system:

- **The Science Team** oversees the physical reliability and validity of the products generated by the system. Consequently, the function of the Science Team is to design scientific algorithms capable to produce long-term consistent, multi-mission global products.
- **The System Development Team** develops the scientific algorithms based on the design of the science team and integrate them in the processing baselines.

In the Lakes_cci context, contributing products (LWL, LWE, LSWT, LIC, LWLR, LIT) are generated separately (Figure 1), with similar temporal and spatial resolutions defined in the Product Specification Document, [RD-2] to be integrated into the final Lakes ECV product (Figure 2).

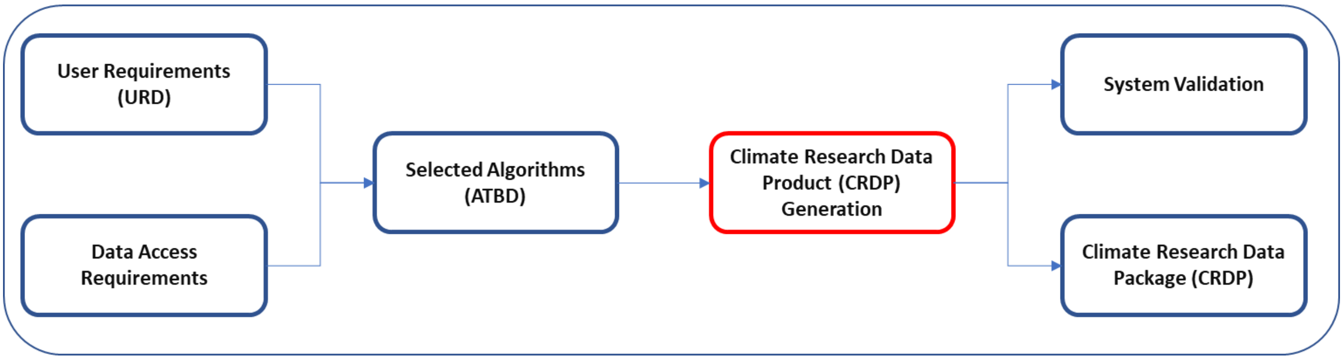

Figure 1. Structure for the product generation.

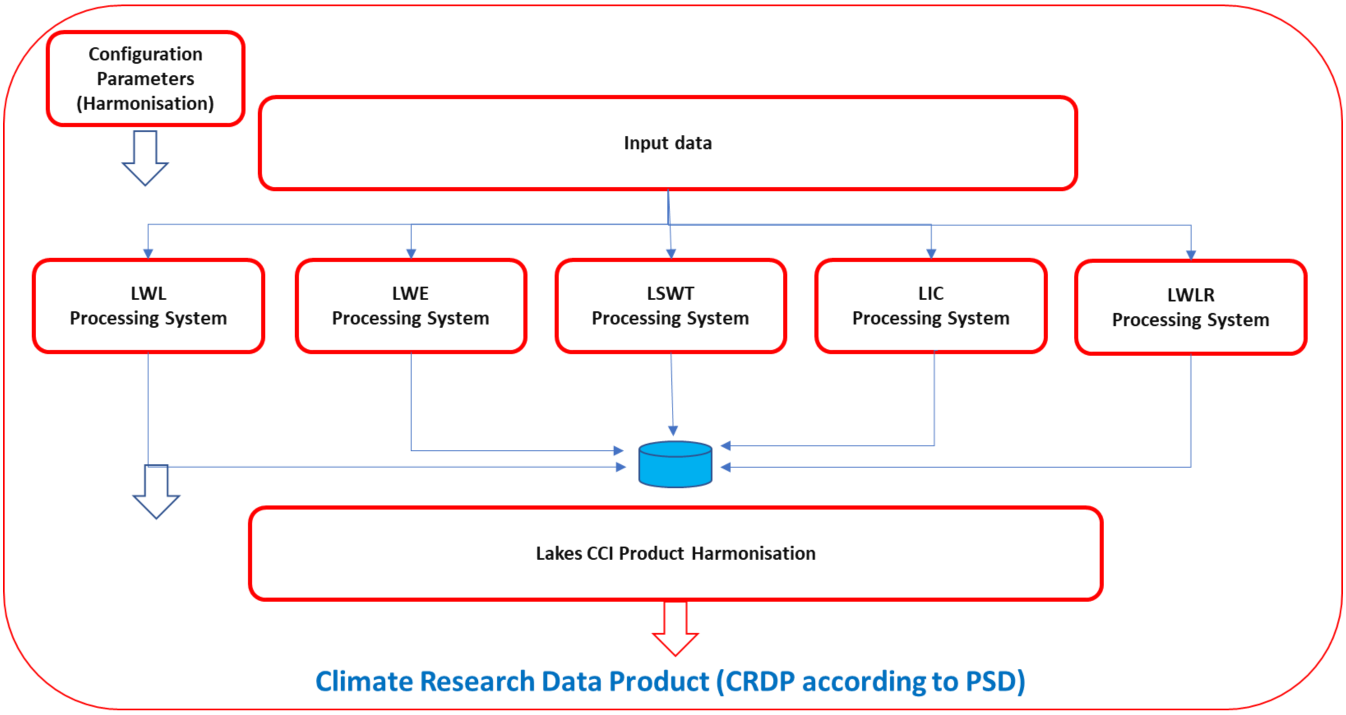

Figure 2. General Principle for harmonisation of Lakes ECV product

Each processing system generates daily products projected onto a common global grid. These are then inputs into the CRDP generation system. The harmonisation process starts by applying a set of harmonisation parameters and ends by merging all outputs from the different systems (see §2.3).

For Lakes_cci CRDP v2.1 the LIT products are delivered separately. The LIT has been added to Lakes_cci only recently and the LIT timeseries are not yet produced globally. Data for a single target, Great Slave Lake (Canada), will be added to the CRDP v2.1 on a demonstration basis. LIT products consist of 20+ years of observations at ten-day temporal resolution over three Regions of Interest (RoIs) of the target lake. The LIT products are generated according to the Lakes_cci product standards, like the other ECV products.

2.3 Harmonization process

The consistency of the product is met if each product follows the rules below:

- Identical product format and time step
- Grid common format
- Common water body definitions: geographic region of interest and unique identifier

The harmonisation process is a three-step process:

1. **Data aggregation, projection, masking, and identification**

Agreed format specifications are applied to the production sequences of all production teams:

- Product format and time step: there is one product per day, containing all parameters. If a parameter is missing the field will be filed with a default value.
- Grid common format: the product provides dada over a grid with spatial resolution of 1/120 degrees. For parameter like LWL and LWE which provide one value per lake, this value is set for each grid point according to the land mask.

- Common land mask: a common land mask is defined to limit observations to specified water body definitions, which forms the basis of processing systems for LSWT, LIC and LWLR. For LWL and LWE it is only used during the merging process. Each processing line uses the base mask according to its specific requirements.
- Common lake identifiers: each lake uses an identifier linked to existing databases or specifically created for the Lakes_cci.

2. Thematic variables generation

For each thematic variable (LWL, LWE, LSWT, LWLR, LIC), the existing processing system is adapted to generate the thematic files in the common output format in order to facilitate the merging process. LIT dataset is provided in a separate file in CRDP V2.1 and it is planned to fully integrate it in the common format in CRDP V3.0

3. Merge

Once each system has generated its own products, a merge process combines all thematic variables into a daily global 'L3S' product. In case of missing observation data, a default (fill) value is used.

There is no value transformation such as interpolation or gap filling. The merging step does not implement any algorithm that could change the values generated and checked by the processing system in charge of the respective variable.

The merging process verifies on the input side that the files for each product are available, and the corresponding variables are included, and on the output side that the corresponding file, with the correct size and the output variables were generated. These output files are compliant with CCI Data Standards V2.3 released on 26/07/2021 (see here : https://climate.esa.int/en/explore/esa-cci-data-standards/) .

The output products are fully described in the Product Specification Document (RD-3) and the Product User Guide (RD-4).

3 Processing system

3.1 Lake Water level - LWL

3.1.1 General Description

This section provides an overview of the LWL system with its system context, its main function and processing chain, and its architecture (Figure 3).

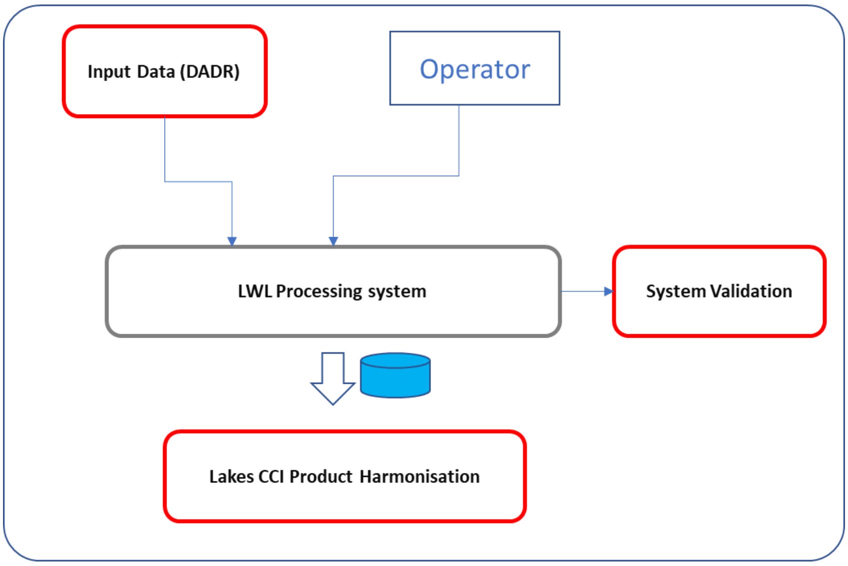

Figure 3: LWL Processing system: General Diagram

3.1.2 Main functionalities

The processing framework of Lake Water Level in the CCI Lakes project is based on the THEIA/Hydroweb project (RD-)

The LWL product is measured using satellite radar altimetry. Radar altimeters send an electromagnetic pulse to the satellite nadir and record the propagation time to and from the emitted wave and its echo from the surface. Algorithm Theorical Basis Document (RD-1) includes the complete description of the algorithm used to estimate the Lake Water Level.

Pre-processing needed to get the input data streams into the correct format to be handled by the Hysope processor. A post process is also needed to adapt the data to the Lakes_cci specifications and it's added around the Hysope processor core.

3.1.3 Architecture

Two types of processing are used to generate the LWL dataset. The first one, used for historical data is performed by LEGOS using GDR data from past missions (Topex/Poseidon, Jason 1 and 2, Envisat, Cryosat-2). The second uses NRT data from current missions (Jason 3, Sentinel 3A, Sentinel-3B and Sentinel-6A) in operational mode, performed by CLS in an operation model with close cooperation with LEGOS.

The Technical Platform and Operations Framework is composed of a set of hardware and software components interacting with each other (Figure 4).

The Lakes_cci LWL Technical Platform is realized at CLS and CNES providing two major hardware components. These hardware components are:

- The CLS cluster: 9 batch servers / 216 cores (432 threads) 128GB memory / server
- The CNES cluster (HAL): 300Tflops, 380 batch servers / 8400cores, 128GB memory / server, 6,2 PB GPFS / 200TB burst buffer/ 100GBs bandwidth Infiniband, low latency network, 4 GPGPU Nvidia Volta V100

The HAL cluster is used to gather and enhance the historical altimetry data thanks to its significant storage capability and computing power. The historical altimetry databases (L2E-HR) on the server are maintained and operated with a CLS software and database proprietary format directly compatible with LWL system.

The CLS cluster is used to copy and store only the relevant input data from the altimetry L2E-HR databases on the HAL cluster. This data is stored temporarily on a dedicated partition (netapp4-L2P-HYDRO, 3To) for the operation team use. The partition also stores the relevant missing ancillary data and houses the operational C3S LWL Processing Framework.

The Operations Framework is a based on a source code library on Github[1] and web-based repository manager GitLab[2], holding all code and configurations fragments needed to run the operational service. Write access to the GitLab code repository is restricted to the members of the Lake Water Level development team. Specific version of code packages can be cloned or downloaded to the Technical Platform. This process is automated by making use of a CLS overlayer software to the Git tools. The documentation platform of all packages is hosted on a dedicated Microsoft SharePoint repository with restricted access to the Science Team, the System Development Team, and the Operations Team.

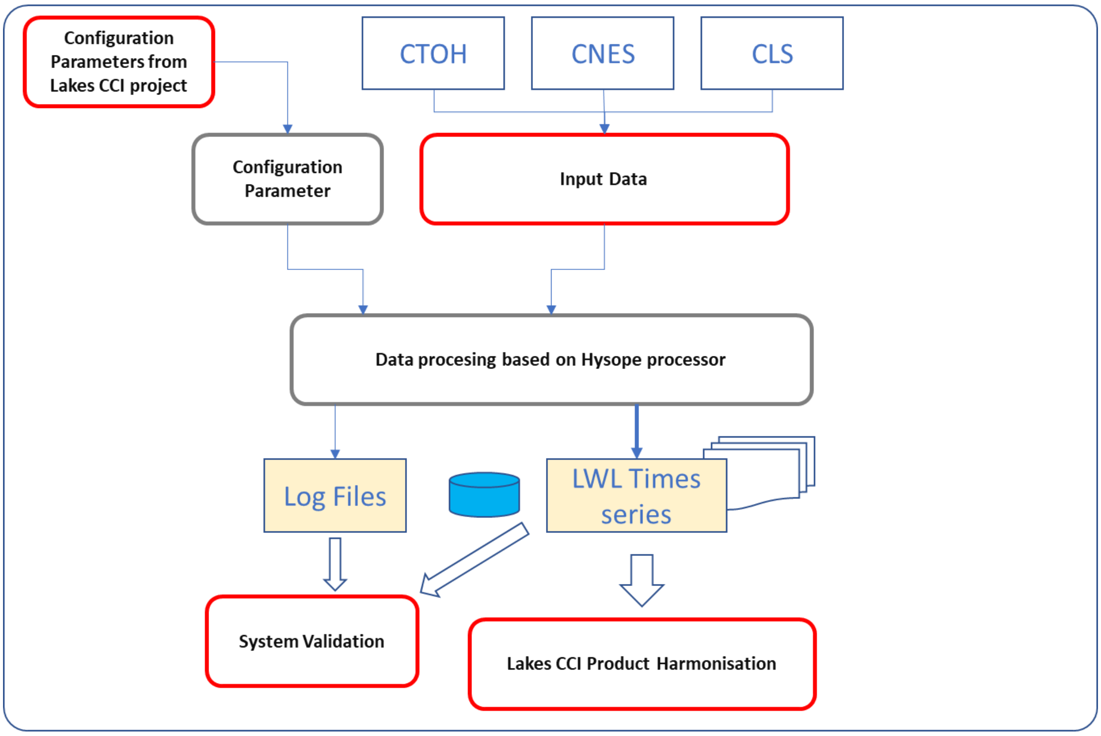

Figure 4. Overview of the elements and interfaces of the Lakes_cci LWL Processing Framework

[1] https://github.com/
[2] https://about.gitlab.com

3.1.4 Input

The original Hysope processor was designed to ingest Level-2 altimetry data. The input format can either be the NetCDF official formats delivered by the space agency (e.g. cophub) or the CLS proprietary format. This last functionality allows the Hysope processor to ingest the historical altimetry database L2E-HR housed by and maintained on the HAL cluster.

Data used in the Lakes_cci LWL can be divided into historic and NRT inputs. Historic data are level-2 altimetry products from already decommissioned satellite missions generated in one processing cycle. These datasets are reprocessed on a regular basis by space agencies to enhance them with the state-of-the art algorithms. The data archive is available on the disk storage of CLS cluster and on the CTOH (Centre for Topographic studies of the Ocean and Hydrosphere), data centre of the Science Team. The data archive and is complemented by regular backups of the data.

Active input data streams are based on observations from a series of altimeters on-board historic and current international satellite missions. Algorithms to derive lake water level from altimetry have been developed and are further research by LEGOS (Crétaux et al 2006, Crétaux et al 2011, Crétaux et al 2016).

3.1.5 Outputs

The Hysope processor can produce for each lake either a CSV or GeoJSON file. Therefore, the CCI lakes post processing block converts the format into the required NetCDF4 described in CCI Data Standards document as well as into thespatial and temporal resolution defined in the Product Specification Documents (PSD): a gridded and daily output. Given that the LWL is not a grid product, it has to be adapted according to the common lake mask defined for all the products in the Lakes_cci project.

3.1.6 Verification

Verification is the process to demonstrate that the system meets the specified requirements [RD-3] and provides data to users. Verification methods used are tests, inspection, and monitoring:

- Unit testing: it is carried out during development of the system software and before its deployment. The expected result is defined before the code execution and the actual result is compared to the expected result.
- Regression testing: it verifies that previous results are the same after a software change
- Completeness check: it ensures that the system has calculated a complete output dataset by inspecting report files and testing that the expected output files have been generated
- Visual inspection: it is the last stage of verification. It consists of opening, with appropriate software, and a looking at the results.

3.2 Lake water extent -LWE

3.2.1 General Description

The lake water extent is calculated from the hypsometric curve. The hypsometry coefficients are determined by combination of LWL and LWE. These LWE are derived from spectral contrasts in the optical reflectance of water compared to surrounding land as described in the Algorithm Theoretical Basis Document (RD-1). The curves are linear or 2d-polynomial representations of the relationship between LWL and LWE.

LWE maps used to calculate the hypsometry coefficients are determined from contrasts in the optical reflectance of water compared to surrounding land. Water surfaces are detected using a multilayer perceptron (neural network) algorithm and integrating the GSW database for sampling. This processing chain is implemented as part of a processing chain for Sentinels 1-3 used to derive several properties including water, fire, and cloud detection, in time series context with large spatial and temporal windows.

3.2.2 Main functionalities

The processing framework of Lake Water Extent in the CCI Lakes project is based on the THEIA/Hydroweb project (RD-). The LWE product is estimated as the corresponding value in the hypsometric curve for a given value of the lake water level.

As for LWL, post processing is needed to adapt the data to the Lakes_cci specifications, which is added around the Hysope processor core.

3.2.3 Architecture

Similar to LWL, the LWE data for historical missions is generated by LEGOS and data for current missions is generated by CLS, in close collaboration. The Technical Platform and Operations Framework is shown in Figure 5.

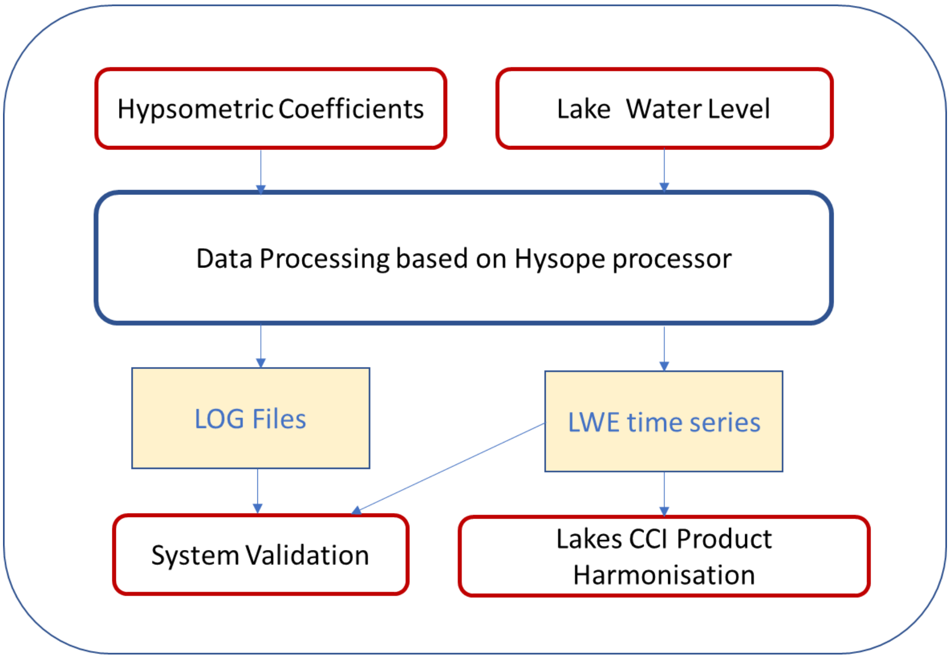

Figure 5. Overview of the elements and interfaces of the CCI Lakes LWE Processing Framework

3.2.4 Input

The satellite altimetry height (LWL) of lake water surface is used together with the hypsometry coefficients to determine the LWE variable. A pre-processing phase is applied which consists in collecting satellite Landsat (5, 7 and 8) and Sentinel-2 images, and in some cases Very High-Resolution alternatives (e.g., Pleiades) to determine lake extent for a set of dates with varying water level to cover the whole range of LWL and LWE variability. The hypsometric coefficients are then calculated through a least square adjustment using the (LWL/LWE) data sets.

3.2.5 Output

The Hysope processor generates a file in CSV format containing the LWE estimates. This file is then converted, in conjunction with the lake water level, into the netCDF4 format to comply with the requirements of the Lakes_cci products. For each lake, the daily LWE is filled into the common water/land mask.

3.2.6 Verification

Analysis of log files allows detection of possible errors during the generation of the data. Additionally, the time series of the lake water extent for each lake are visually inspected after generation. Moreover, we systematically determine the precision of each time series of LWE calculating the RMSE of the difference between LWE inferred from the polynomial function of the hypsometry of the lake, and LWE directly measured by the satellite image, over the whole dataset of images used to calculate the hypsometry coefficient.

3.3 Lake water surface temperature - LWST

3.3.1 General Description

The overview of the LSWT processing system is shown in Figure 6

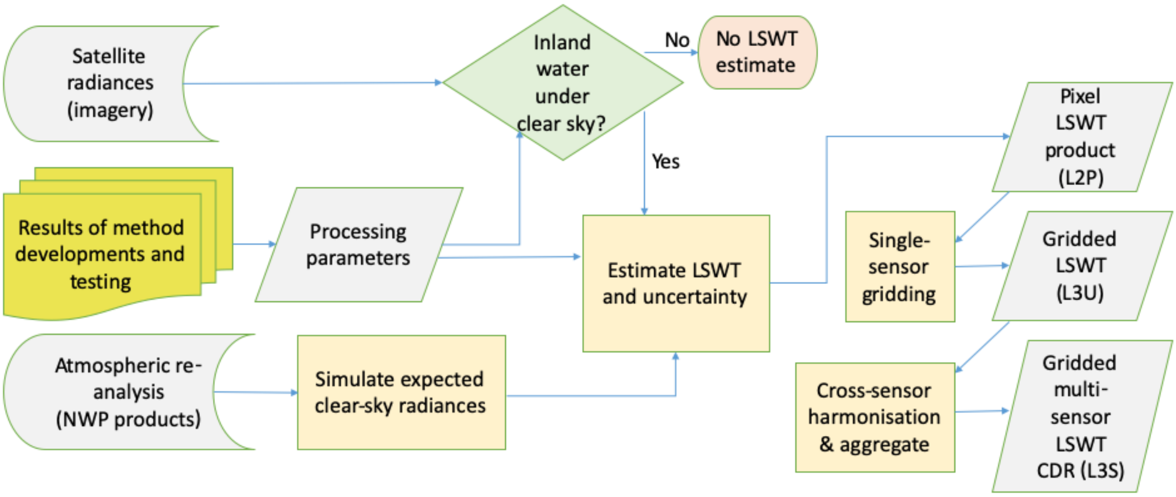

Figure 6. LSWT System overview

3.3.2 Main functionalities

The LSWT processor undertakes the following functions:

- Satellite L1 data read (all required input data streams)
- Extraction of matching prior information from numerical weather prediction (NWP) fields
- Identification of candidate satellite pixels filled with inland water (static mask)
- Dynamic water detection on candidate satellite pixels and cloud screening
- Radiative transfer modelling of satellite radiances and Jacobians given NWP (calling external software)

- Optimal estimation of LSWT, total column water vapour (discarded) and uncertainties
- Output of L2P (full-resolution swath LSWT) products
- Averaging/regridding to regular grids creating single-sensor gridded uncollated (L3U) products
- Harmonisation of mean LSWT per lake using overlapping data across sensors

Final output of multi-sensor gridded ("super-collated") climate data records (L3S)

3.3.3 Architecture

The LSWT processor is hosted at the Centre for Environmental Data Analysis (CEDA), in a Linux environment enabled to see multiple TB (Terabytes) of input files on "spinning disk" access. The processor consists of several core modules in Fortran augmented by Python 3. Trivially parallel processing is available, as each input file may be processed independently of all others to the point of generation of L3U. Once all L3U are generated, a separate process is initiated to combine to the L3S climate data record.

3.3.4 Input

The Data Access Requirement Document (RD-6) contains the characteristics and full references of the input data used to estimate the LSWT from the missions listed below, the ancillary data (inland water mask and distance to land raster file, LSWT climatology raster file plus NWP fields) and in-situ data (for validation purposes).

The missions presently processed are:

- Along-Track Scanning Radiometer 2 (ATSR2)
- Advanced Along-Track Scanning Radiometer (AATSR)
- Moderate Resolution Imaging Spectroradiometer (MODIS) Terra
- Advanced Very High-Resolution Radiometer (AVHRR) MetOp-A
- Advanced Very High-Resolution Radiometer (AVHRR) MetOp-B
- Sea and Land Surface Temperature Radiometer (SLSTR) Sentinel-3A
- Sea and Land Surface Temperature Radiometer (SLSTR) Sentinel-3B

The ancillary NWP have been ERA-Interim analysis for ATRS2, AATSR and the AVHRRs. ERA5 NWP has been utilised for MODIS and the ECMWF operational NWP in-file data has been utilised for the SLSTRs. The system development for switching to ERA-5 is complete but more investigations are needed on the output for future reprocessing, and the future system will also add a data stream from new sensors such as SLSTR on Sentinel3C. Adding a further sensor with similar channels is a direct evolution from the system point of view: the elements that need to be modified are: the reader for ingesting satellite radiance imagery; the processing parameters (from results of scientific algorithm development); and the simulation of expected clear-sky radiances (configuration of simulation for new sensor). Other elements in the processing chain operate as for other sensors.

3.3.5 Outputs

Outputs are netCDF files compliant with the recommendations of the CCI Data Standards Working Groups.

The distributed outputs are the L3S outputs which are generated on two grids: the legacy grid of 0.05 degrees; and the Lakes CCI common grid (1/120th degree). The latter data are distributed only within the all-variables Lakes CCI product.

3.3.6 Verification

Verification comprises:

- Visual inspection of plots from random samples of each run
- Validation of extracted data against matched in situ LSWT observations, using the quality levels and water detection generated by the processor

The results of validation of extracted data are described in detail in the Product Validation and Inter-Comparison Report (RD-5).

3.4 Lake water leaving reflectance - LWLR

3.4.1 General Description

The *Calimnos* processing chain combines data discovery, subsetting by target area (individual water bodies), radiometric and atmospheric corrections, pixel identification (land/cloud/water/ice), optical water type classification, individual algorithms (per parameter and water type), algorithm blending, conversion and aggregation into a single processing chain.

A schematic overview of *Calimnos* is given in Figure 7. The main processing stages and their corresponding algorithms are listed in the next section and are detailed in the ATBD (RD-1).

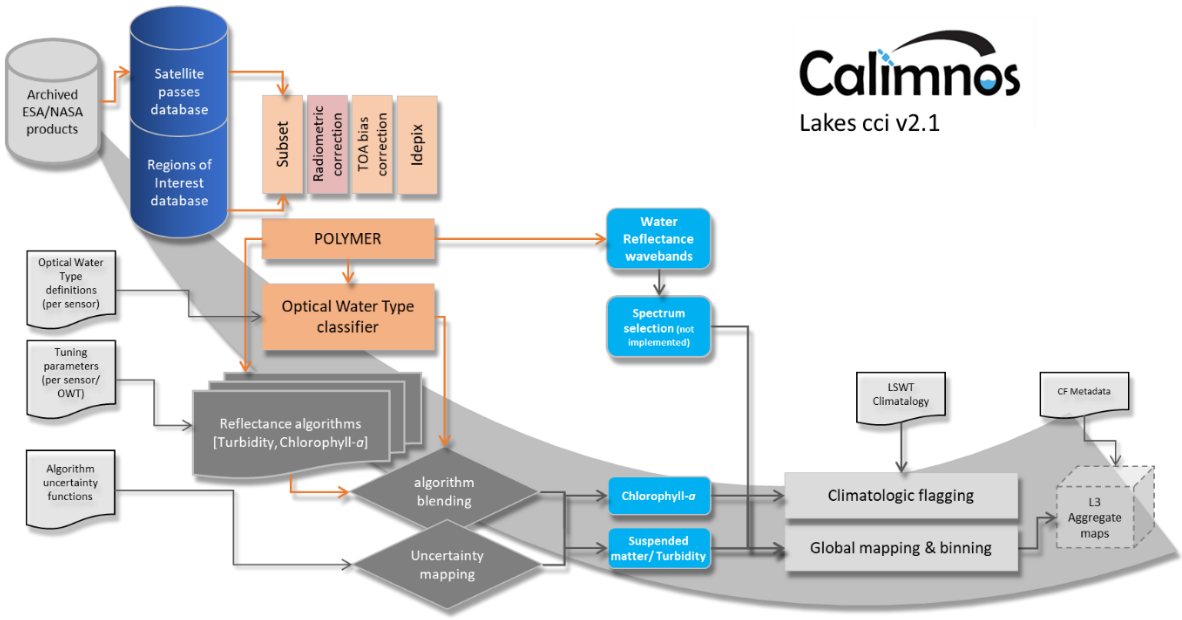

Figure 7: Schematic overview of the Calimnos processing chain for LWLR, Chlorophyll-*a* and turbidity or suspended matter.

3.4.2 Main functionalities

L2 processing steps

To produce Lake Water-Leaving Reflectance:

- Data discovery
- Subsetting around the lake areas of interest
- Radiometric and/or sensor Bias corrections
- Pixel identification as water/land/ice/cloud/cloud-shadow
- Atmospheric correction

To produce derived water-column properties:

- Optical water type classification
- Algorithm mapping and blending

For uncertainty characterization:

- Uncertainty mapping per algorithm and per optical water type

L3 processing steps

For merged lakes ECV product format consistency:

- Aggregation to 1-day intervals and Reprojection to a common planetary lat/lon grid
- Masking inconsistencies
- Mosaicking into a single global product

3.4.3 Architecture

The processing system is built around the concept of workspaces which are individually submitted to a high-performance computing environment (SGE to date, migrating to SLURM).

L2 workspaces are created to contain symbolic links to all required input data for each given combination of target area and sensor overpass. L3 processing workspaces similarly combine all L2 inputs for a given aggregation period. Workspace creation and processing is done using in-house software written for Python 3, submitting each workspace to a set of processing stages (described in RD-1) which are individually monitored and timed. Log files are created for bulk workspace creation and individual workspace monitoring. For operational processing, a job monitoring database (postgreSQL) is used to follow the timing of completion of L2 jobs, re-start jobs that are in error, and launch L3 processing when all upstream jobs are completed. This process is fully automated.

For a typical lake area, the number of L2 processing workspaces for MERIS and OLCI observations is in the order of 3000-5000 observations per lake or approximately 1 million individual processing jobs for the 250 lakes included in CDRP V1.0 and in the order of 9-10 million for CRDP v2.0, with over 1 million processing units added for each additional year seeing operation of two OLCI sensors. For MODIS-Aqua we consider a period of 9 years spanning the gap between MERIS and OLCI which yields approximately 10,000 observations per lake. The overlapping years are only used to determine that the MODIS outputs are consistent with the other sensors, which is only the case for a subset of the investigated lakes.

Workspaces are submitted to a Sun Grid Engine (or increasingly SLURM) processing queue and controlled using resource allocations. Memory (RAM) and temporary local storage criteria are set at the individual job level depending on lake size to ensure that processing nodes with different capabilities are optimally used. Each grid node has access to a shared software repository and storage media containing the input data, through a Gigabit network interface. The processing grid is normally configured to provide up to 850 individual processing slots, which are used in parallel. A containerised implementation of the processing chain has been developed and was tested on a separate massive parallel processing environment using SLURM. The containerisation minimises any risks of software version deviations going unnoticed on older or newly introduced hardware.

When an instance of a workspace is started on a processing node it is copied to local storage and a supervisor script then executes each stage of the processing chain, logging both standard and error output streams to disk. The processing stages are coded into an xml file where each stage can be switched on or off depending on requirements. A separate configuration file stores common algorithm parameterisation settings and instructions for individual processing stages, which allows a common processing stage configuration to be used with project-specific parameterisation.

Upon completion of a workspace, any files that need to be archived are copied to their archiving destination. Workspace logs are normally kept up to 1 month following production. A copy of the configuration settings is kept with the output data.

External factors such as server disk space and internet connections are monitored for the processing site using Nagios.

3.4.4 Input

Input satellite data consist of MERIS Reduced Resolution L1B from the 3rd reprocessing (for CRDP v3.0 the 4th reprocessing will be used), OLCI full resolution L1B data from the Sentinel-3A and 3B platforms, and MODIS-Aqua. All downloaded products (a full archive is kept) are referenced in a postgreSQL database.

A polygonised and manually corrected version of the ESA CCI Land Cover (v4.0) maximum water extent at 150-m resolution is used as initial lake mask. The polygons are available at https://github.com/pmlrsg/lake-polygons-PML. They are accessible to *Calimnos* from a postgreSQL database.

Configuration files are used as input to each L2 or L3 workspace and comprise:

- Processing stages in xml format
- Common algorithm parameterization in ascii format compatible with Python configparser.
- Optical water type spectra in comma separated ascii format
- Processing environment requirements in ascii format, generated from a common template at workspace creation for each lake/sensor combination.

LSWT climatologies are used to mask inconsistent results during periods with possible lake ice.

3.4.5 Outputs

The outputs comprise:

- L2 products, at the native sensor resolution, 1 file per combination of satellite overpass and region of interest. These contain most intermediary products and can be used for match-up validation and further algorithm development. They are not disseminated.
- L3 products, at a common grid, 1 file per combination of region of interest and aggregation time frame.
- Mosaicked L3 product, 1 file per time step containing all regions for which input data was available.
- Log files of each processing level, separated by processing stage.

The L3 product only super-collated during the Sentinel-3 A/B joint observation period. MERIS inputs are used exclusively for the period April 2002 - April 2012, and MODIS-Aqua only for the period between MERIS and OLCI availability.

3.4.6 Verification

The Calimnos codebase is versioned using Git (private Gitlab repository) and tagged by software version number. The software version number for the Lakes_cci CDRPv1.0 through 2.0.2 is v1.4.

Integration tests are completed as part of code review, for each compatible sensor, against a template configuration which includes all processing stages.

Selected lakes and time periods are produced and inspected prior to large-scale processing. Log files are screened for common errors (such as network glitches) while any remaining errors are manually inspected and resolved. Consistency in file size is checked and selected files are visually inspected.

Implementation of new versions of software dependencies normally results in a minor version increase and will be subject to the above tests.

3.5 Lake Ice cover – LIC

3.5.1 General Description

An overview of the LIC processing system is given in Figure 8. The goal of the LIC processing system is to determine the state of lake surfaces. The state can be assigned as either ice or water, or cloud in which case no observation is able to effectively observe the lake surface. The surface of the Earth (and therefore the lakes) is gridded by latitude and longitude into "squares" whose edges subtend 1/120th of a degree (approx. 1 km at the equator.) The cells of this grid which are of interest conform to the specifications of ESA.

The input data is provided at multiple resolutions with subpixels that are centered at locations that are perturbed from an ideal grid. Pixel correlation is done by nearest neighbour resampling. Data from each pixel is then fed to the LIC retrieval algorithm which produces a label for the pixel.

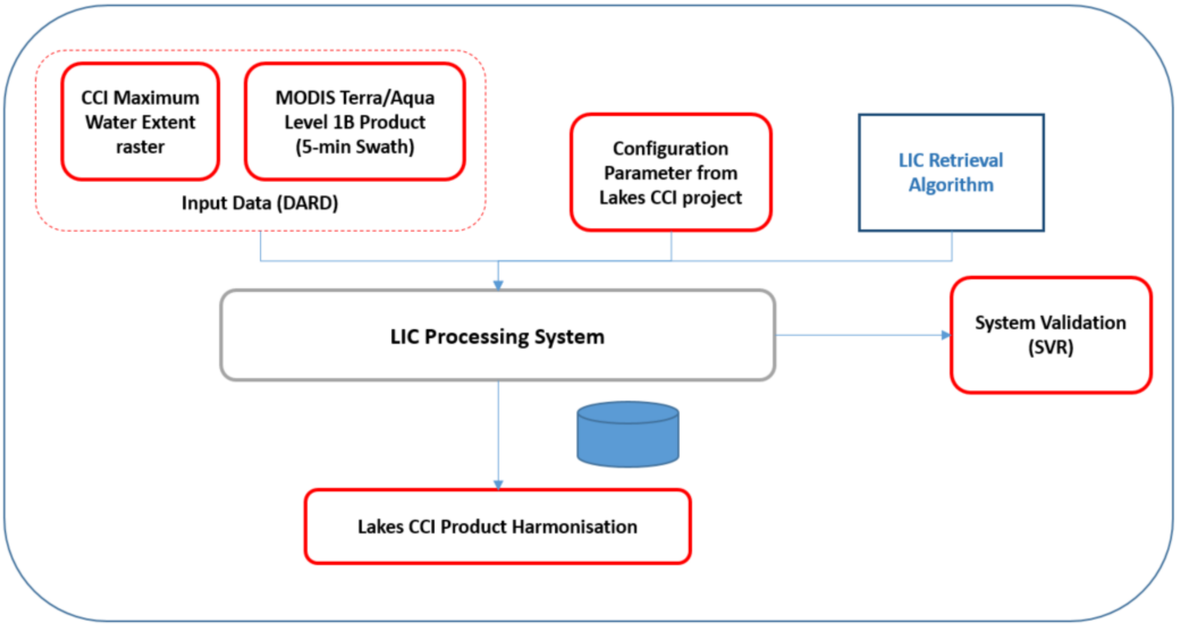

Figure 8. LIC Processing System: Context Diagram

3.5.2 Main functionalities

The LIC product (v2.1) is generated from a random forest algorithm using MODIS Terra/Aqua Level 1B data, which records the percentage of light reflected by the top of the atmosphere.

The retrieval algorithm using a well-trained random forest classifier turns the satellite observations into labels as either lake ice, open water or cloud cover. Moreover, a global water mask is employed as auxiliary data to filter land and ocean pixels. Details of the random forest algorithm and its assessment for lake ice cover mapping from MODIS Level 1B imagery are described in Wu et al. (2021).

Prior to generating the LIC product, the required MODIS product is downloaded from NASA's server and loaded into the processing chain. Subsequent to algorithm retrieval, the LIC product is written out following the Lakes_cci specifications.

3.5.3 Architecture

The LIC processing system is written in C++11 with support modules written in Python 3.6. The software is mostly hardware independent, though single precision floating point calculations are used. It should be the case that all IEEE 754 compliant hardware and compilers produce effectively identical results. The software requires a POSIX environment. Figure 9 shows the external interfaces of the data processing software.

The program requires approximately 7 GB of RAM to run one instance (this depends on the total area of the lakes of interest). Instances are fully independent. Each instance requires one process with two threads. The threads are only for pipelining – close to optimal throughput can be achieved with only one physical core.

One day of input (one instance) is approximately 160 GB. Total input size is currently about 1.3 PB for 22 years of MODIS level 1B data Terra (2000-2022) and Aqua (2002-2022).

The software is deployed on a high-performance computing (HPC) environment (Table 1). The cluster is attached to multiple storage systems which total to more than 95 PB of storage. A low-latency high-bandwidth InfiniBand fabric connects all nodes and scratch storage. Nodes each have multiple CPUs. The CPUs vary among the Xeon E5 v4, Xeon E7 v4, and Xeon Gold product lines.

Table 1. Overview of HPC environment

Operating System	CentOS 7
Number of nodes	1185
Number of cores per node	32 ~ 64
Memory per server	128 ~ 3070 GB

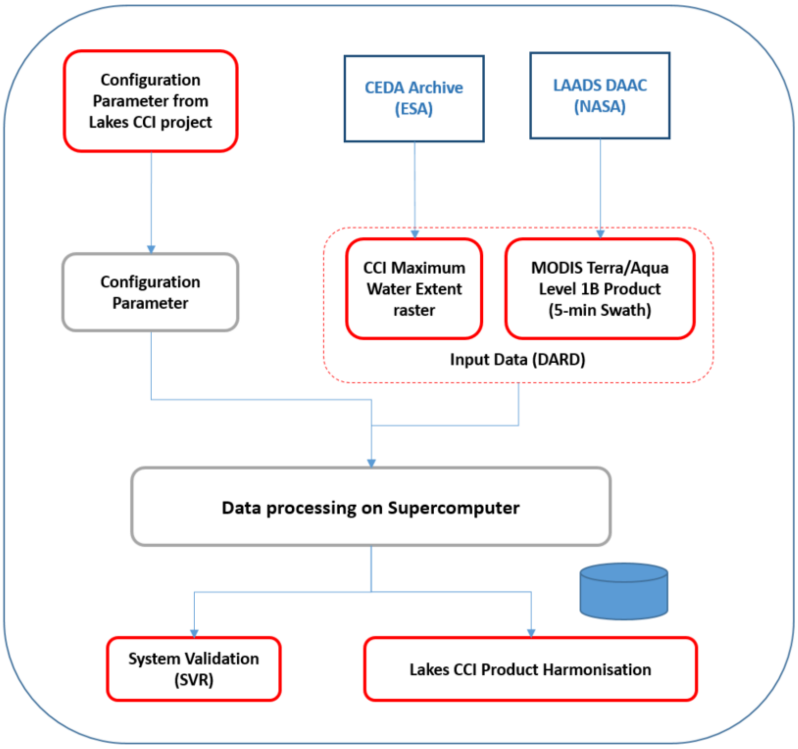

Figure 9. Overview of the elements and interfaces of the CCI lakes

LAADS DAAC (NASA) refers to the Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (DAAC) serving the global NASA Terra, and Aqua MODIS products. The CEDA Archive used by ESA refers to the UK Natural Environment Research Council's Data Repository hosting the ESA CCI Water Extent product.

3.5.4 Input

The LIC producing program requires two input data sources, MODIS Terra/Aqua Level 1B (MOD02/MYD02) product and water extent raster.

The MODIS instruments onboard Terra and Aqua have been delivering data since 2000 and 2002, respectively. The MODIS Terra/Aqua Level 1B product records Earth observations (top-of-the-atmosphere reflectance) in 5-min orbital swath format without projection. The MODIS product format is the Hierarchical Data Format (HDF5), which contains multidimensional arrays to store data from multiple channels and associated metadata. In addition, each HDF file provides longitude and latitude bands in WGS 84 as georeferenced data.

The maximum water extent provided in ESA CCI Land Cover (v4.0) at 150-m resolution is employed as lake mask to filter land and ocean pixels from MODIS.

3.5.5 Outputs

The output data is produced in the harmonised grid format. The edge of each grid cell subtends 1/120th degrees latitude/longitude. The output variables in the LIC product include label assigned to grid cell (bands 1 and 2) and uncertainty of the label (%) (band 3). The ATBD (RD-1**Erreur ! Source du renvoi introuvable.**) describes the complete details of the output variables of LIC product.

3.5.6 Verification

In addition to the main processing chain (Map Producer), another branch of the LIC mapping system has been designed for data sampling (collect, train/optimization, test) and product validation (Figure 10). This branch can produce RGB colour composites, and a graphical user interface (GUI) developed in Python allows users to access and display RGB images to manually sample pixels with labels.

The output of labelled samples from the sampling GUI and LIC maps can be read by the validation program to conduct accuracy assessment. Moreover, the validation program also can export misclassified observations and samples with spectral data for further testing and research. The ATBD and E3UB provide details on the accuracy assessment of the LIC product.

The verification of the LIC product involves quantitative and qualitative assessments derived from the computation of confusion matrices and manual inspection, respectively. Manual inspection of the product is done via loading and display of a sample of images from various seasons, regions, and ice years in either image analysis or GIS software packages (ESA SNAP and ArcMap). Verification involves examination of metadata as well as visual inspection of pixel values for ice map classes (ice/water/cloud/bad data) against RGB colour composites for quality assurance and identification of errors to consider for future algorithm improvements leading to CDRP V3.0.

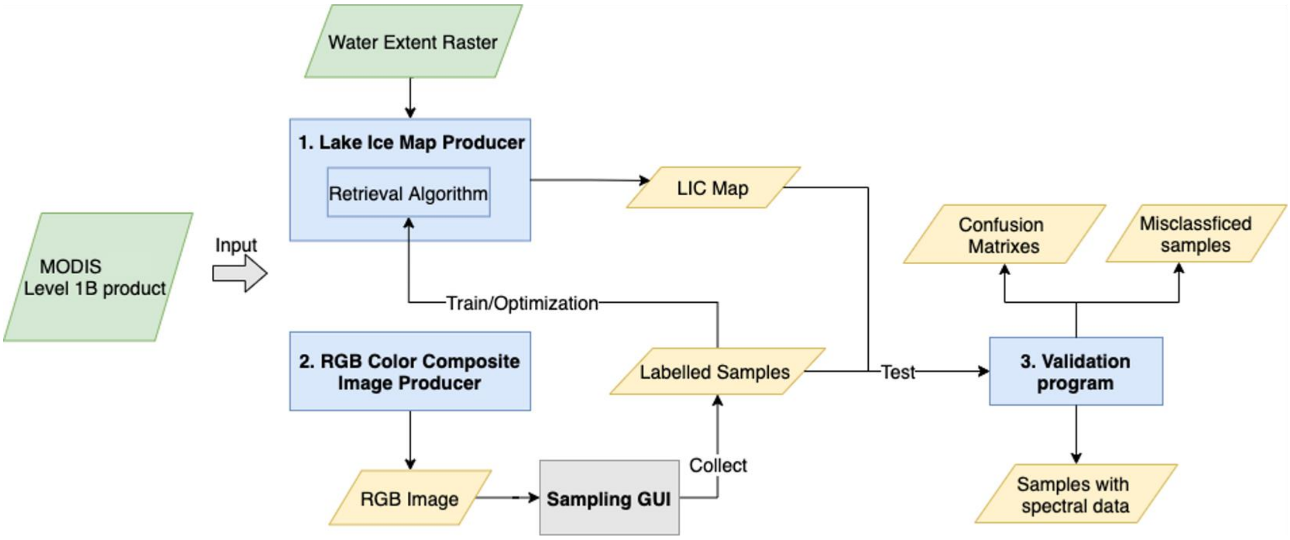

Figure 10. Overview of the LIC mapping system

3.5.7 References

Wu, Y., Duguay, C. R., and Xu, L. (2021). Assessment of machine learning classifiers for global lake ice cover mapping from MODIS TOA reflectance data. *Remote Sensing of Environment*, 253, doi:10.1016/j.rse.2020.112206.

3.6 Lake Ice Thickness- LIT

3.6.1 General Description

This section provides the description of the LIT processing system which is summarized in Figure 11. The goal of the LIT processing is to provide LIT measurements by analysing Ku-band radar altimetry waveforms in Low Resolution Mode (LRM) that contain information correlated with the seasonal evolution of ice thickness over freshwater lakes. For a target lake, the LIT analysis is performed on along track segments (RoI) chosen to avoid land contamination that could perturb the LIT signature on the radar waveforms (and thus the LIT measurements). The LIT RoI therefore depends on the target coverage of each radar altimetry mission used in the analysis (each mission can have different overpass over a lake and thus different RoI) and must contain enough waveforms to have robust and representative LIT estimations. For each target, the LIT measurements are first generated for each mission and RoI, and a data editing step is also performed to remove eventual outliers before storing the results into intermediate products, which are validated internally and then merged into the final LIT timeseries products.

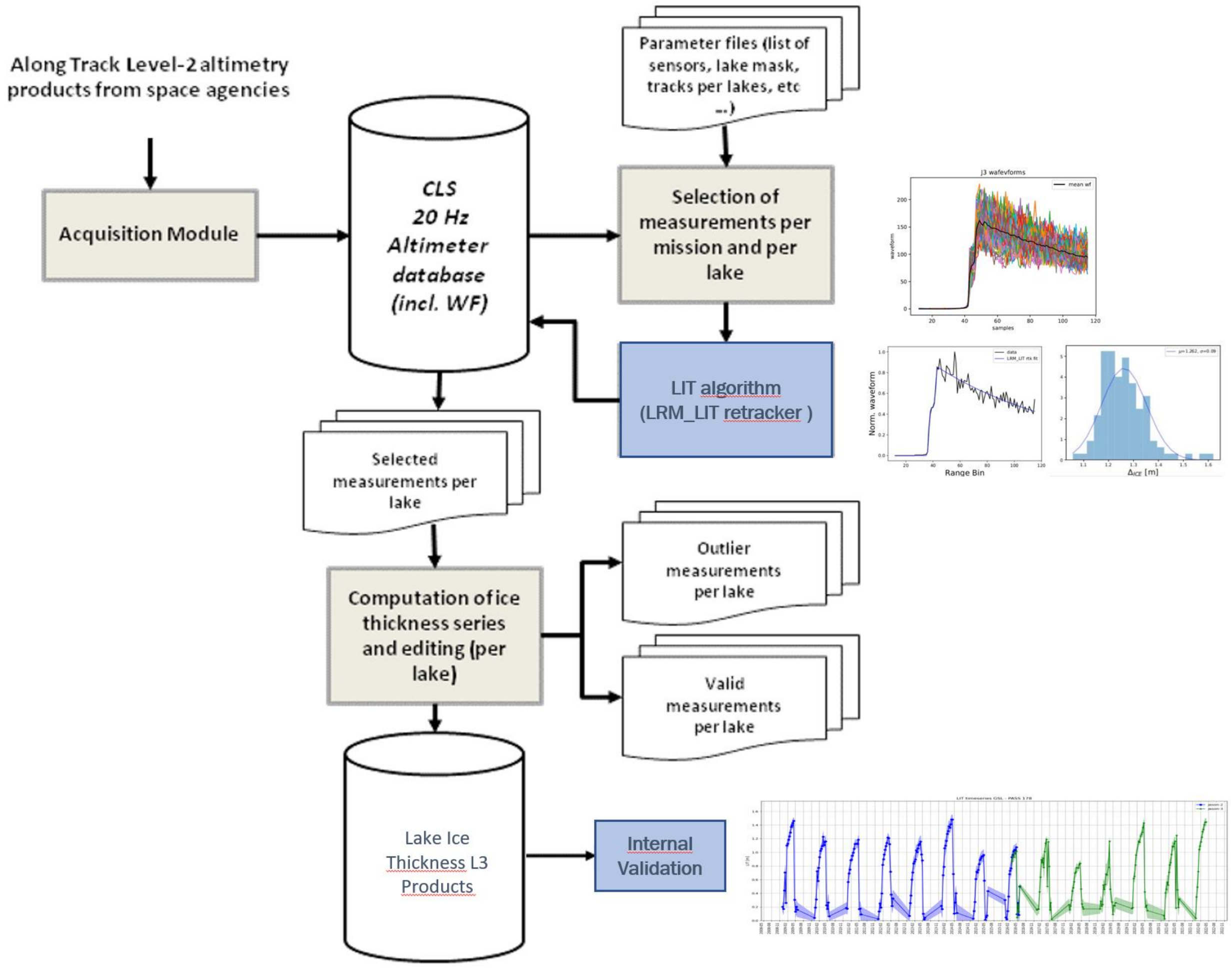

Figure 11. LIT processing system summary diagram

3.6.2 Main functionalities

The core of the LIT processing chain is the LRM_LIT retracking algorithm [Mangilli et al 2022] that generates the LIT estimates by fitting the Ku band radar waveforms with a tailored analytical model specific to LIT. The main steps integrated in the LIT processor, for each target, mission, and overpass, are:

- Input data: extraction of the radar altimetry data from the CLS database
- Data selection: selection of the waveform data in the Region of Interest (RoI)
- Retracking step: LRM_LIT retracker processing (optimisation and parameter estimation)
- Data editing: outliers' identification and removal
- LIT timeseries and intermediate product generation

Once the intermediate products for one target are generated for each mission and overpass, they are merged and validated through the following steps:

- Merging of the LIT timeseries into one final product for each target and RoI
- Internal product validation (consistency and missing data checks)

3.6.3 Architecture

The LIT processing system is hosted at CLS in a Unix environment that enables to efficiently exploit the radar altimetry data stored in the CLS database. The LIT processor is written in Python 3.9 and includes specific Python libraries for the LIT analysis and to deal with the acquisition and handling of the CLS data base.

3.6.4 Input

For each mission and pass over a target lake, the inputs are the Ku bands radar waveforms in Low Resolution Mode (LRM) at 20Hz resolution. The data are the AVISO L2 sensor products for each mission which are acquired and stored in the CLS database to ensure an efficient handling of the data for long time series analysis.

3.6.5 Outputs

- Main outputs (stored in the LIT products): LIT mean and standard deviation within the RoI analysis window over a target lake, and the associated quality flag
- Complementary outputs (stored in the intermediate products only): the best fit LIT estimations from the fit of each radar waveform in the LIT analysis window RoI over the target lake with the corresponding coordinates (LIT spatial evolution). This typically consists of a ~few hundreds of LIT estimations for each data cycle. This output is used for the internal validation and the consistency checks.

3.6.6 Verification

The different products generated for each target, mission and RoI, are merged in the final LIT timeseries NetCDF products. At this step, the merged products are generated to be compliant with Lakes_cci data standards.

4 Exploring lake_cci data

To access and explore the Lakes_cci data, users can access, explore, download, and analyse the Lakes_cci dataset in various ways.

4.1 CCI Open Data Portal

To access the full global datasets and lake masks with the maximum extent, Lakes_cci data can be downloaded using multiple download mechanisms and ways to extract a set of variables, or data for a specific region lakes_cci_V2.0.2.

4.2 Jupyter notebooks

A series of notebooks is freely available of charge to help new users familiarise themselves with data extraction for further analysis: lakes_cci GitHub

4.3 Interactive visualisation

Using WebGIS you can view most of the Lakes_cci variables and extract small sections of the data for download or plotting (Figure 12). This is also a useful resource for training and education, e.g. by sharing links to your visualisation with others: https://lakescci.eofrom.space/

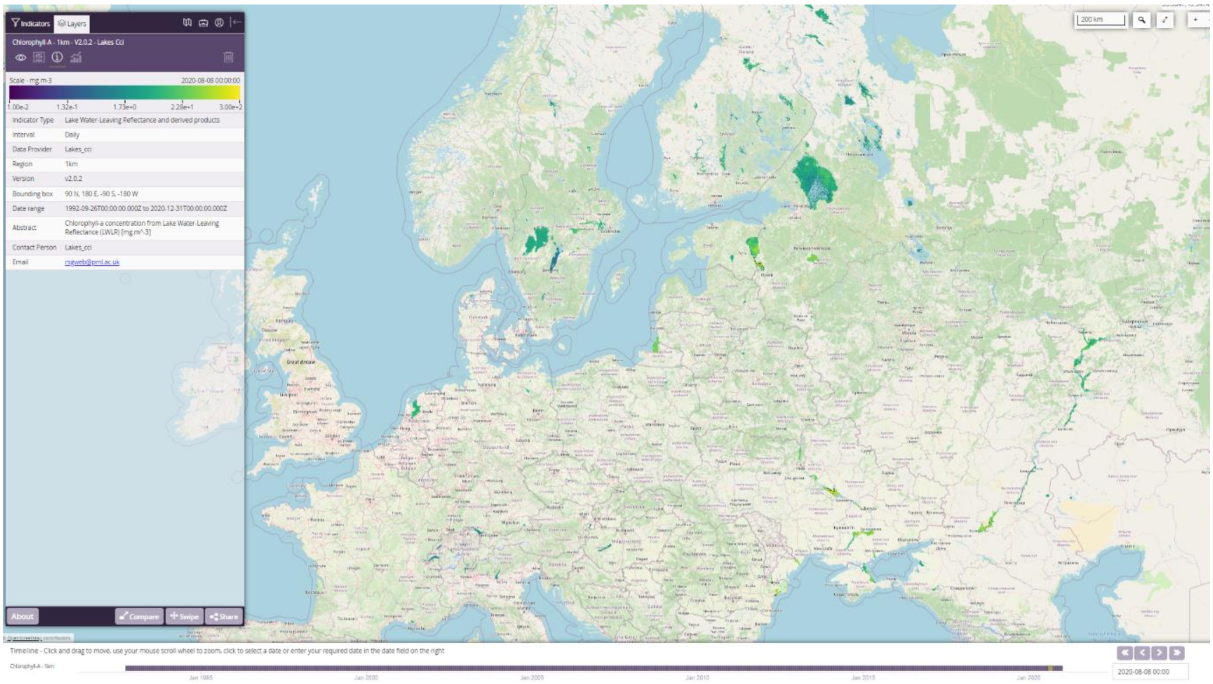

Figure 12. Lakes_cci WebGIS

4.4 Software tools

Lakes_cci data is stored in the NetCDF4 classic file format. A wide choice of software packages can be used to visualise or manipulate the NetCDF data. A list of software is provided on the Unidata web site (https://www.unidata.ucar.edu/software/netcdf/software.html). The Lakes_cci files can be visualised with the Climate Analysis Toolbox (Cate), the reference software for visualising data developed within the CCI Program funded by ESA.

5 Requirements coverage

This part indicates if the parameter/product is compliant with the Cardinal Requirements (CR) and Technical Requirements (TR) indicated in the SoW as well as the requirements from the users.

Table 2. Requirement coverage list

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
CR-1	Develop and validate algorithms to approach the GCOS ECV and meet the wider requirements of the Climate Community (i.e. long term, consistent, stable, uncertainty-characterized) global satellite data products from multi-sensor data archives.	Yes	Yes	Yes	Yes	Partial	Yes
CR-2	Produce, validate and deliver consistent time series of multi-sensor global satellite ECV data products for climate science	Yes	Yes	Yes	Yes	Yes	Yes
CR-3	Maximise the impact of European EO mission data on climate data records	Use all altimetric missions	Yes	Yes	No for v1.0. Yes for v2.0.	Yes	V2.2:Jasons mission 1/2/3)
CR-4	Generate and fully document a production system capable of processing and reprocessing the data in CR-2, with the aim of supporting transfer to operational activities outside CCI (such as C3S).	Yes	Yes	Yes	Yes	Yes	Yes
R-3	Each CCI project (the contractor) shall make significant progress towards meeting the corresponding GCOS requirements for their ECV	Yes	Yes	Yes	Yes	Yes	Yes
R-7	Each CCI project (the contractor) shall take into account the legacy of the CCI in their projects. This involves: • adopting the community consensus concept • learning from, understanding and building upon the success achieved in CCI • espousing the implicit request to contribute actively to the core elements of the CCI Programme, in particular Colocation and Integration meetings, working groups, cross-programme activities (Open Data Portal, Toolbox, Knowledge Exchange) and cross-project initiatives.	Yes	Yes	Yes	Yes	Yes	Yes

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
R-9	Each CCI project team (the contractor) shall take full account of the following key technical constraints when planning and implementing the CCI project: During Phase 1 the project consortium shall need to respond to the following technical constraints: • Need for scientific consensus on detailed ECV product and performance specifications • Availability and quality input data from EO Archives (ESA and non-ESA) • Availability and quality of associated metadata, cal/val data, and documentation • Compatibility of data from different missions and sensors • Trade-offs between cost, complexity and impact of new algorithms to be developed and validated during the project • Advance planning for data from new missions to be integrated during the project • End-to-end throughput of ECV production systems • Re-use of existing capabilities within Europe • Compliance to applicable standards • Availability of external validation data • No duplication of activities covered by other projects or programmes (e.g.H2020, Copernicus, national funding)	Yes	Yes	Yes	Yes	Yes	Yes
R-16	Each CCI project team (the contractor) shall integrate data from the Copernicus Sentinels and other key satellite missions within the relevant CCI processing systems and ECV data products	Yes	Yes	Yes (Sentinels will be V2.0)	Yes	Yes	Yes

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
R-17	Each CCI project team (the contractor) shall ensure that the system is adequately dimensioned to accommodate the growing volumes of input and output data, and the increasing computational loads needed to process, reprocess, quality control, validate, and disseminate multi-decadal, global, ECV data products, of the required climate quality, in a timely manner.	Yes	Yes	Yes	Yes	Yes	Yes
TR-3	The Contractor shall confirm and demonstrate that all work undertaken in Lakes_cci is complementary to other Lake ECV development and delivery activities being conducted by ESA and other agencies e.g. GloboLakes, HydroWeb, Copernicus Climate Change Service and Copernicus Global Land Services. It shall coordinate its activities with these efforts to ensure the information on lakes is effectively and efficiently produced for the investment from these different projects.	Yes	Yes	Yes	Yes	Yes	Yes
TR-4	The project shall be consistent and compatible with previous CCI projects, in particular, LandCover_cci in terms of its Permanent Water Bodies product	Yes	Yes	Yes	Yes	Yes	Yes
TR-7	The project shall ensure consistency across the different lake products and deliver all the products to established lakes databases (GTN-H - HYDROLARE, HYDROWEB, NSIDC Global Lake and River Ice Phenology, GLTC).	Yes	Yes	Yes	Yes	Yes	Partial (1 target for CRDP2.2)
TR-8	The project shall ensure consistency in production across multiple satellites, in particular focusing on different spatial and temporal resolutions offered consistent with the range of lake sizes that require monitoring	Yes	Yes	Yes	No for v1.0. Yes for v2.0.	Yes	Yes
TR-10	The contractor shall aim to generate the most accurate, stable and uncertainty characterised products by continuously evaluating and developing the selected algorithms and keeping abreast of new developments in the field	Yes	Yes	Yes	Yes	Yes	Yes

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
TR-11	The contractor shall ensure that all products are consistent across product streams and thus the system delivers all products for the lakes under evaluation. This shall include consideration of baseline lake identifiers, sensor geolocation, algorithm stability and accuracy, required preprocessing/sensor corrections,	Yes	Yes	Yes	Yes	Yes	
TR-14	The Contractor shall review algorithm performance and their implementation in the light of developments in the scientific literature, application to multiple sensors, robustness and accuracy on a regular basis and evaluate the need to conduct reprocessing to improve lake ECV products on a regular basis	Yes	Yes	Yes	Yes	Yes	Yes
TR-17	The contractor shall develop a prototype processing system (Lake-System) to generate all products for the Lake ECV	Yes	Yes	Yes	Yes	Supported	Yes
TR-18	The Lake-System shall provide the scientists a configurable, flexible, agile and open Lake CCI workflow for managing the evolution of the Lake ECV. It shall have the following capabilities: • Multi-sensor full mission (re-)processing to all Lake ECV products • Systematic, data-driven processing, allowing rapid ingestion of new data • Configuration management (processor versions, auxiliary data, input and output data, etc …) • Capability to integrate (existing) tools developed in different programming languages • Capability for algorithm developers to easily trial plug in, without recompilation where practical, and execute new algorithms, new versions of algorithms or new parameterisation, for testing, intercomparison and evaluation purposes • Support efficiently all the needs of the scientific algorithms and the new development	Yes					

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
TR-19	The Lake-System shall include tools for full mission time series analysis, round robin algorithm intercomparison, quality control and other comprehensive massive product analysis (feature extraction, dependencies (e.g. from sensor, detector, geometry, latitude/longitude, etc.). The development of these tools shall be coordinated with, and transferred to, the CCI Toolbox and, where possible, compatible with existing web services (e.g. Global Surface Water Explorer).	No plans for such tools					
TR-20	The Lake ECV products shall be made available to the users through appropriate distribution mechanisms with full user support and analysis tools	Data available from: • CEDA • Interactive visualisation : https://lakescci.eofrom.space/ • Jupyter notebooks: lakes_cci GitHub					
TR-21	The Lake-System shall include data access, ingestion, product conversion tools and distribution functionality, and shall address long-term archiving of both input and output products. Both requirements shall meet the generic needs outlined in the main body of the CCI SoW.	Yes					
TR-22	The design of the Lake-system shall be based on experience where relevant from previous CCI projects and/or external processing systems	Yes					
TR-23	The Lake-System shall have data access interfaces to handle efficiently, and in a standardised fashion, the massive primary and auxiliary data streams from Sentinel 1, 2 and 3. The interfaces shall support the possibility of cross-ECV synergies.	Yes					
TR-24	The contractor shall link with the CCI portal and data analysis/visualisation tools available through the CCI Toolbox	To be done by those projects					
TR-25	The Contractor shall generate global maps of lakes at multiple resolutions derived from multi-sensor Lake FCDRs. The products shall address all GCOS Lake ECV requirements, pushing products as far as scientifically and technically possible.	All GCOS addressed					

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
TR-26	The project shall generate prototype products of: • Lake level • Lake area • Lake surface water temperature • Lake colour (and its derived properties: turbidity, chlorophyll and coloured dissolved organic matter) • Lake ice coverage (where relevant) • Lake ice thickness (where relevant)	Yes	Yes	Yes	Yes	Partial: CDOM not yet possible	YES
TR-27	The project shall generate products for at least 2000 lakes covering the range of lake sizes and types with priority given to those lakes of climate importance as a demonstration of system capability	250 lakes selected for the first version of the dataset 2024 lakes selected for late versions (since V2.0)					
TR-28	The project shall focus on addressing lakes in the priority regions identified by the IPCC. Namely: • Lake surface warming, biogeochemistry and water column stratification increases in the African Great Lakes, Lake Kariba. • Increased lake water temperatures, biogeochemistry and ice phenology in the Arctic. • Associated change in lake patterns from permafrost degradation in Siberia, Central Asia, Tibetan Plateau, Arctic (both thermakarst lake loss and new lake generation in frozen peat). • Shrinking mountain glaciers across most of Asia, Andes, western North America, Arctic and associated development and expansion on foreglacier lakes. • Lake changes in response to changes in snowpack in Australasia and western North America. • Lakes associated with changes in river discharge patterns in circumpolar rivers, Amazon River, western Andes and La Plata river.	Yes					
TR-29	The project shall aim to provide the longest time series possible based on available data products with 10 years being the minimum requirement	Yes					

From	Target	LWL	LWE	LSWT	LIC	LWLR	LIT
TR-30	The Lake ECV products shall: • Cover the time frame from 1992-end of this contract. • Be derived from all available and suitable satellite instruments listed in Tables 4-8 of the SoW • Include RMSE and bias uncertainty estimates on a per pixel basis following guidelines expressed in the main Statement of Work (see also [RD-8]). The confidence in these uncertainty estimates shall be stated	V1: 1992-2019	V1: 1992-2019	Yes	Yes	Yes V1.0: 2002-2012 and 2016-2019 MERIS & OLCI currently suitable v2.0 includes MODIS during 2012-216 for 38 lakes	V2.2 2001-2022
TR-31	The Lake ECV products shall comply with the actual version of the CCI Guidelines for Data Producers [RD-7]. Note: The CCI project shall adapt to changes in these Guidelines as these will be further developed jointly by the CCI projects.	Yes					
TR-33	The Contractor shall generate at least 2 versions of ECV products: one following the round-robin, then one incorporating the results from user feedback.	Yes	Yes	Yes	Yes	V1.0 based on prior round-robins	Yes
TR-34	Full ECV product validation shall be achieved by a combination of: a. Activities of the CCI EO Science Team b. Activities of the Climate Research Group engaged in the Lake CCI project c. Involvement in the CMUG process d. Interaction with key ecosystem modelling groups e. Involvement of key stakeholders of other international climate research projects.	Yes					

Appendix A - Project Acronyms

Acronym	
AATSR	Advanced Along Track Scanning Radiometer
AATSR	Advanced Along Track Scanning Radiometer
AERONET-OC	AErosol RObotic NETwork - Ocean Color
AMI	Active Microwave Instrument
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
APP	Alternating Polarization mode Precision
ASAR	Advanced Synthetic Aperture Radar
ASLO	Association for the Sciences of Limnology and Oceanography
ATBD	Algorithm Theoretical Basis Document
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced very-high-resolution radiometer
BAMS	Bulletin of the American Meteorological Society
BC	Brockman Consult
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CDR	Climate Data Record
CDOM	Coloured Dissolved Organic Matter
CEDA	Centre for Environmental Data Archival
CEMS	Centre for Environmental Monitoring from Space
CEOS	Commitee on Earth Observation Satellites
CF	Climate and Forecast
CGLOPS	Copernicus Global Land Operation Service
CIS	Canadian Ice Service
CLS	Collecte Localisation Satellite
CMEMS	Copernicus Marine Environment Monitoring Service
CMUG	Climate Modelling User Group
CNES	Centre national d'études spatiales
CNR	Compagnie Nationale du Rhône
CORALS	Climate Oriented Record of Altimetry and Sea-Level
CPD	Communiation Plan Document
CR	Cardinal Requirement
CRG	Climate Research Group
CSWG	Climate Science Working Group

Acronym	
CTOH	Center for Topographic studies of the Ocean and Hydrosphere
DUE	Data User Element
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
ELLS-IAGRL	European Large Lakes Symposium-International Association for Great Lakes Research
ENVISAT	Environmental Satellite
EO	Earth Observation
EOMORES	Earth Observation-based Services for Monitoring and Reporting of Ecological Status
ERS	European Remote-Sensing Satellite
ESA	European Space Agency
ESRIN	European Space Research Institute
ETM+	Enhanced Thematic Mapper Plus
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAQ	Frequently Asqued Questions
FCDR	Fundamental Climate Data Record
FIDUCEO	Fidelity and Uncertainty in Climate data records from Earth Observations
FP7	Seventh Framework Programme
GAC	Global Area Coverage
GCOS	Global Climate Observing System
GEMS/Water	Global Enfonment Monitoring System for freshwater
GEO	Group on Earth Observations
GEWEX	Global Energy and Water Exchanges
GloboLakes	Global Observatory of Lake Responses to Environmental Change
GLOPS	Copernicus Global Land Service
GTN-H	Global Terrestrial Network – Hydrology
GTN-L	Global Terrestrial Network – Lakes
H2020	Horizon 2020
HYDROLARE	International Data Centre on Hydrology of Lakes and Reservoirs
ILEC	International Lake Environment Commettee
INFORM	Index for Risk Management
IPCC	Intergovernmental Panel on Climate Change
ISC	International Science Council
ISO	International Organization for Standardization
ISRO	Indian Space Research Organisation
JRC	Joint Research Centre

Acronym	
KPI	Key Performance Indicateurs
LEGOS	Laboratoire d'Etudes en Géophysique et Océanographie Spatiales
LIC	Lake Ice Cover
LIT	Lake Ice Thickness
LSC	Lake Storage Change
LSWT	Lake Surface Water Temperature
LWE	Lake Water Extent
LWL	Lake Water Level
LWLR	Lake Water Leaving Reflectance
MERIS	MEdium Resolution Imaging Spectrometer
MGDR	Merged Geophysical Data Record
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	MultiSpectral Instrument
MSS	MultiSpectral Scanner
NASA	National Aeronautics and Space Administration
NERC	Natural Environment Research Council
NetCDF	Network Common Data Form
NOAA	National Oceanic and Atmospheric Administration
NSERC	Natural Sciences and Engineering Research Council
NSIDC	National Snow & Ice Data Center
NTU	Nephelometric Turbidity Unit
NWP	Numerical Weather Prediction
OLCI	Ocean and Land Colour Instrument
OLI	Operational Land Imager
OSTST	Ocean Surface Topography Science Team
PML	Plymouth Marine Laboratory
PP	Payment Plan
PRISMA	PRecursore IperSpettrale della Missione Applicativa
Proba	Project for On-Board Autonomy
QSR	Linear Correlation Coefficient
R	Linear Correlation Coefficient
RA	Radar Altimeter
RMSE	Root Mean Square Error
SAF	Satellite Application Facility
SAR	Synthetic Aperture Radar
SeaWIFS	Sea-viewing Wide Field-of-view Sensor

Acronym	
SIL	International Society of Limnology
SLSTR	Sea and Land Surface Temperature Radiometer
SoW	Statement of Work
SPONGE	SPaceborne Observations to Nourish the GEMS
SRD	System Requirements Document
SSD	System Specification Document
SST	Sea Surface Temperature
STSE	Support To Science Element
SWOT	Surface Water and Ocean Topography
TAPAS	Tools for Assessment and Planning of Aquaculture Sustainability
TB	Brightness Temperature
TM	Thematic Mapper
TOA	Top Of Atmosphere
TR	Technical Requirement
UNEP	United Nations Environment Programme
UoR	University of Reading
UoS	University of Stirling
US	United States
VIIRS	Visible Infrared Imaging Radiometer Suite
WCRP	World Climate Research Program
WHYCOS	World Hydrological Cycle Observing Systems
WMO	World Meteorological Organization
WP	Work Package