Consortium Members

D2.5: Product Validation Plan (PVP)

Reference:	CCI-LAKES-0030-PVP
Issue:	1.2
Date:	24 April 2020

Version history:			
Issue:	Date:	Reason for change:	Author
1.0	06/12/19	Initial Version	S. Simis et al.
1.1	19/03/20	Revision following ESA review	S. Simis et al.
1.2	24/04/20	Revision following ESA review	S. Simis et al.

People involved in this issue:			Signature
Authors:	Stefan Simis, Xiaohan Liu	Plymouth Marine Laboratory	
	Jean-François Crétaux	LEGOS	
	Hervé Yésou	SERTIT	
	Erik Malnes, Hannah Vickers	NORCE	
	Pablo Blanco	TRE Altamira	
	Chris Merchant, Laura Carrea	University of Reading	
	Claude Duguay	H20 Geomatics	
Internal review:	S. Simis, 12 March 2020	Plymouth Marine Laboratory	Stefan Simis Digitally signed by Stefan Simis Date: 2020.05.07 22:43:31 +01'00'
Approved by:	B. Coulon	CLS	Bruno Coulon Signature numérique de Bruno Coulon Date : 2020.05.05 08:36:05 +02'00'
Authorisation:	C. Albergel	ESA	Clement Albergel Digitally signed by Clement Albergel Date: 2020.05.11 18:23:41 +02'00'

Distribution:		
Company	**Names**	**Contact Details**
ESA	A.M. Trofaier C. Albergel P. Cipollini	Anna.Maria.Trofaier@esa.int Clement.Albergel@esa.int Paolo.Cipollini@esa.int
BC	K. Stelzer	kerstin.stelzer@brockmann-consult.de
CLS	B. Coulon B. Calmettes	bcoulon@groupcls.com bcalmettes@groupcls.com
CNR	C. Giardino	giardino.c@irea.cnr.it

Distribution:		
Company	**Names**	**Contact Details**
H2OG	C. Duguay	claude.duguay@h2ogeomatics.com
LEGOS	J.F. Créteaux	jean-francois.cretaux@legos.obs-mip.fr
PML	S. Simis	stsi@pml.ac.uk
UoR	C. Merchant L. Carrea	c.j.merchant@reading.ac.uk l.carrea@reading.ac.uk
UoS	A. Tyler E. Spyrakos	a.n.tyler@stir.ac.uk evangelos.spyrakos@stir.ac.uk
TRE-Altamira	P. Blanco	pablo.blanco@tre-altamira.com
NORCE	E. Malnes	eima@norceresearch.no
SERTIT	H. Yesou	herve.yesou@unistra.fr
GeoEcoMar	A. Scrieciu	albert.scrieciu@geoecomar.ro

Table of Contents

1. Overview

The Lakes_cci project has three stages of product validation. The first stage is the validation of individual thematic variables by their respective teams, which often coincides with the characterisation of product uncertainties as detailed in the E3UB. Second, the consistency between these variables is investigated through a number of pre-defined use cases, each focussing on multiple variables. Finally, when the climate data records are released for external use, project scientists will be collecting feedback from data users and improving the data production chains accordingly.

This document presents, for each of the thematic variables included in the Lakes_cci, the nature of validation activities that are planned to take place. These may include in situ measurement campaigns through international collaborations, or ingestion and compilation of existing datasets with subsequent analysis of product accuracy. Gaps and opportunities for product validation are thus identified.

The feedback expected from the planned use case studies is summarized in the last chapter of this document.

2. Lake Water Level (LWL) Validation

2.1. LWL Validation activities

Two types of validation are performed for LWL.

1. General comparison of LWL products against in situ observations
2. Dedicated field work in the framework of satellite altimetry cal/val programmes.

The comparison with in situ measurements is done in coordination with the SHI (State Hydrological Institute in St Petersburg) which is an external partner to the Lakes_cci and which further collaborates in the framework of the Hydrolare lake database.

2.2. Schedule for LWL validation

Table 1 Overview of LWL validation activities

Validation activity	*Time frame for analysis*	*Implementation*
Method 1: direct comparison between LWL products against in situ data	For most of the lakes the first data were acquired in the 1990s and the last are obtained in recent years (2015 to 2019). Some in situ data are constantly acquired (in North America in particular) and regular updates of the database are planned.	For method 1 and 2 the methodology is already implemented, and results are updated yearly for both approaches.
Method 2: field work planned at specific dates with corresponding satellites coverage over Lake Issykkul, where all correction steps in LWL processing are checked against acquired in situ measurements (wet and dry troposphere, instrumental biases).	Yearly field work over Lake Issykkul has been performed since 2003. In 2020, a first campaign for LWL data will also be organised over Lake Baikal.	

2.3. Inputs and methods for LWL validation

Data from Hydrolare over the Russian territory, and some countries in Central Asia are used. RMS of differences, absolute bias, and potential drifts are produced lake by lake. A validation set of 10 lakes is included in this study (Onega, Ladoga, Baikal, Bratske, Illmen, Kuybyshev, Caspian, Khanka, Rybinsk and Issykkul).

We also extract water level from existing in situ database in North America (source: NOAA and USGS in USA, Canadian water office in Canada) and South America (water national agency in Brazil, Chile and Argentina)

Interpolation of LWL product to the dates of in situ measurements are first of all performed, then the mean bias between in situ and satellite time-series is calculated. A bias is always detected since satellite time series and in situ measurements are never given using the same geodetic reference frame. Some results of these comparison are given in Cretaux et al. (2016) and Ričko et al. (2012).

Drift can subsequently be adjusted if it is observed. Root-mean-square differences are calculated, and in case of multi-satellite data the RMS will be derived for each individual mission.

We can see the result for Lake Baikal in Figure 1 below.

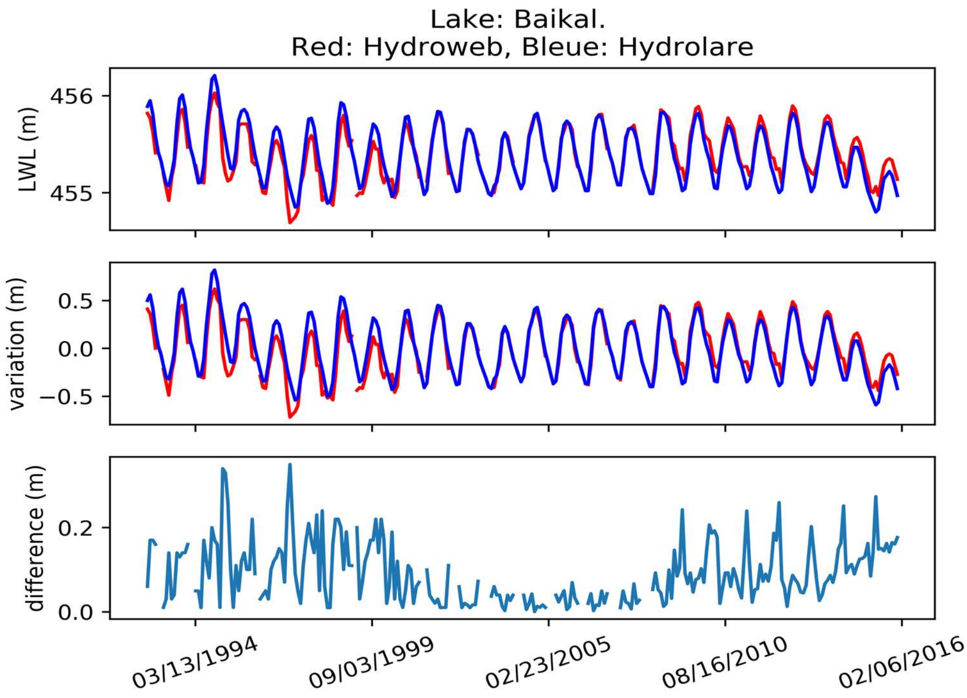

Figure 1 Time series of Lake Baikal

The second approach is based on 15 years of field work experiments over Lake Issykkul in Central Asia. This large lake (6000 km^2) was selected in 2004 to serve as a dedicated calibration / validation site for satellite altimetry over lakes. It has the advantage of overpasses by all past, present and future altimetry missions. The instrumental concept for the field work is widely described in several publications (Cretaux et al. 2009, 2011, 2013, 2018, Bonnefond et al. 2018). In brief, the field work is organised yearly or bi-yearly after consulting the ephemerides of the satellites. GPS levelling of the lake surface is performed along the satellite tracks using a GPS system. In situ fixed instrumentation allows to assess the stability of the LWL product, and also to validate the atmospheric and geodetic corrections. The main purpose is to perform full error budget analysis including the range measurements using different retracking algorithms (so called ice-1, Ice-2, ocean) and also the different corrections (ionosphere, troposphere, geoid). In 2020 in the framework of calibration / validation of Sentinel-3A/B missions, a similar experiment will start in Lake Baikal in Russia.

Figure 2 shows an example of LWL altimetry measurements against LWL from GPS levelling along two tracks (666 and 707) of sentinel-3A.

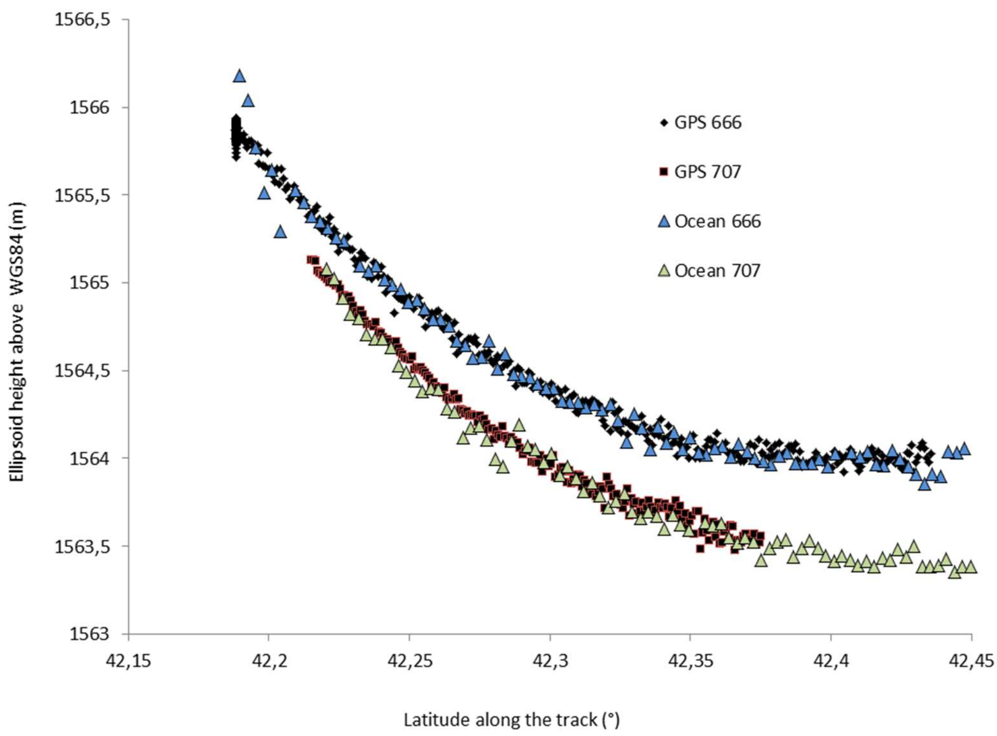

Figure 2 LWL altimetry measurements compared with GPS levelling along two S3 tracks.

2.4. Known constraints for LWL validation

The main difficulty for LWL validation using the first method is that in situ measurements are difficult to obtain, so there are few datasets. Validation against one lake may not be relevant for other lakes since we can observe large differences depending on the morphology of the lake, the geographical and climatic conditions. So using the existing list of in situ data is not representative of the general accuracy of satellite altimetry for LWL measurement.

Regarding the second approach, in situ data are collected during field work, so the comparisons are valid only for few dates (once or twice per year) and for one or two lakes lake. The objective is not general validation of LWL but error budget in a well know case study. The methodology of calibration / validation based on in situ GPS levelling is moreover quite complex to perform, and also quite expensive. It is currently funded by CNES in the framework program of calibration / validation of satellite altimeters for the Issykkul site. For Baikal, it is part of another project funded by ESA (S3MPC) for validation of sentinel-3A/B altimetry measurements.

The uncertainty given within the products are not directly relying on the validation process described above. Indeed, validation only allows providing general overview of errors budget (for in site Calibration / Validation experiment) and comparisons against in situ data give another overview, but of accuracy in an as relevant as possible context (size, morphology of lakes, environmental conditions). Uncertainties given in the products for LWL are simply the quantification of the dispersion of the individual measurements along the track of the satellites after all corrections have been performed and the average LWL has been calculated. The values of uncertainties are the therefore simply the RMS of the differences of these individual measurements against the average LWL calculated. Results of validation process described here are published and allow users to rely on a degree of confidence of the product, but this cannot be considered as direct uncertainty since the validation is evidently limited to a small number of lakes.

Moreover uncertainties being only statistical, eventual biases or long term drifts are not directly visible in these numbers. That is the reason why regular external validations are done. Instrumental biases are extracted from the in site experiments, while drifts or seasonal errors due to changing climate (presence of ice for example, or of aquatic vegetation) can be seen with comparison to in situ historical data.

2.5. References

Bonnefond, P.; Verron, J.; Aublanc, J.; Babu, K.N.; Berge-Nguyen, M.; Cancet, M.; Chaudhary, A.; Cretaux, J-F.; Frappart, F.; Haines, BJ. , Laurain, O.; Ollivier, A.; Poisson, JC.; Prandi, P.; Sharma, R.; Thibaut, P.; Watson, C. The benefits of the Ka-Band as evidenced from the SARAL/AltiKa Altimetric mission: quality assessment and unique characteristics of AltiKa data, Remote Sensing. 2018, 10(1), 83, doi:1039/rs/10010083

J-F Cretaux, M. Bergé-Nguyen, S. Calmant, N. Jamangulova, R. Satylkanov, F. Lyard, F. Perosanz, J. Verron, A.S. Montazem, G. Leguilcher, D. Leroux, J. Barrie, P. Maisongrande and P. Bonnefond, 2018, Absolute calibration / validation of the altimeters on Sentinel-3A and Jason-3 over the lake Issykkul, Remote sensing, 10, 1679,; doi:10.3390/rs10111679

Cretaux J-F, Abarca Del Rio R, Berge-Nguyen M, Arsen A, Drolon V, Clos G, Maisongrande P, Lake volume monitoring from Space, Survey in geophysics, 37: 269-305, doi 10.1007/s10712-016-9362-6 , 2016

Crétaux J-F., Bergé-Nguyen M., Calmant S., Romanovski V.V., Meyssignac B., Perosanz F., Tashbaeva S., Arsen A., Fund F., Martignago N., Bonnefond P., Laurain O., Morrow R., Maisongrande P., Calibration of envisat radar altimeter over Lake Issykkul, J. Adv. Space Res., Vol 51, 8, 1523-1541, doi: 10.1016/j.asr.2012.06.039, 2013

Crétaux J-F, Calmant S, Romanovski V, Perosanz F, Tashbaeva S, Bonnefond P, Moreira D, Shum C.K, Nino F, Bergé-Nguyen M, Fleury S, Gegout P, Abarca Del Rio R, and Maisongrande P, Absolute Calibration of Jason radar altimeters from GPS kinematic campaigns over Lake Issykkul, Marine Geodesy, 34 : 3-4,291-318, DOI: 10.1080/01490419.2011.585110, 2011

Cretaux J.F., Calmant S., Romanovski V., Shabunin A., Lyard F., Berge-Nguyen M., Cazenave A. Hernandez F., and F Perosanz An absolute calibration site for radar altimeters in the continental domain: lake Issykkul in Central Asia, Journal of Geodesy, Vol 83, 8, 723-735, DOI: 10.1007/s00190-008-0289-7, 2009

Ričko M., C.M. Birkett, J.A. Carton, and J-F. Cretaux, Intercomparison and validation of continental water level products derived from satellite radar altimetry, J. of Applied Rem. Sensing, Volume 6, Art N°: 061710, DOI: 10.1117/1.JRS.6.061710, 2012

3. Lake Water Extent (LWE) parameter

3.1. LWE Validation activities

Lake Water Extent is an additional thematic Lakes_cci variable for which methods are under development during the first year of the project. This chapter details the methodology and plans for validation of candidate methods. The final methodology is expected to be a combination of candidate procedures, as detailed in the ATBD, and will be subject to its own validation, following the same (where applicable) principles as set out below.

Three methods are identified to validate the Lake Water Extent products:

- Cross-validation of coincident SAR and optical measurements of water extent
- Validation via the hypsometric method against in situ lake water level measurements
- Validation via the hypsometric method against altimeter measurements of the lake water level measurements

The first method gives a pixel-wise comparison of the two products. Visual inspection is sufficient to determine which of these results (SAR or optical) has highest confidence. For example, a cloud-free optical image is easy to classify and likely to be accurate, while SAR may have challenges with noise due to wind. Inversely, thin cloud or shadows impede the optical detection at times when SAR may be more accurate. The result of validation in this context will be a pixel-wise accuracy estimate, as well as overall aggregated accuracy estimates.

Method 2 and 3 are closely related, but the second is assumed to have higher accuracy since in situ measurements provide higher accuracy than satellite altimeters. On the other hand, in situ measurements are relatively scarce and therefore limit the scope of the validation. The methods give an overall error estimate of the area per classified image, provided that the hypsometric curve can be estimated with high accuracy. We can also estimate the overall area uncertainty.

3.2. Schedule for LWE validation

Validation activity	*Time frame for analysis*	*Implementation*
1) Direct comparison between S1 and S2 over selected lakes	2017-2019, for selected data from same days.	CDR v1 & v2
2) Comparison against in situ data for Altevatn, and potentially other lakes where in situ LWL measurements area available	S1: Carried out from 2002-2019 using a complete set of Envisat, Radarsat-2 and S1 data	CDR v1 & v2
3) Comparison against altimeter measurements of LWL over all lakes	Carried out over the set of vectors (LWL, LWE) used for comparison which is different from one lake to another. Range of level variability is the criteria for selection of vector (LWL, LWE). It can cover a long time frame in case of monotonic LWL time series, or it can be reduced in case of seasonal mode of variability.	CDR v1 & v2

3.3. Inputs and methods for LWE validation

Method 1:

An example of cross-comparison between S1 and S2 is shown in Figure 3.

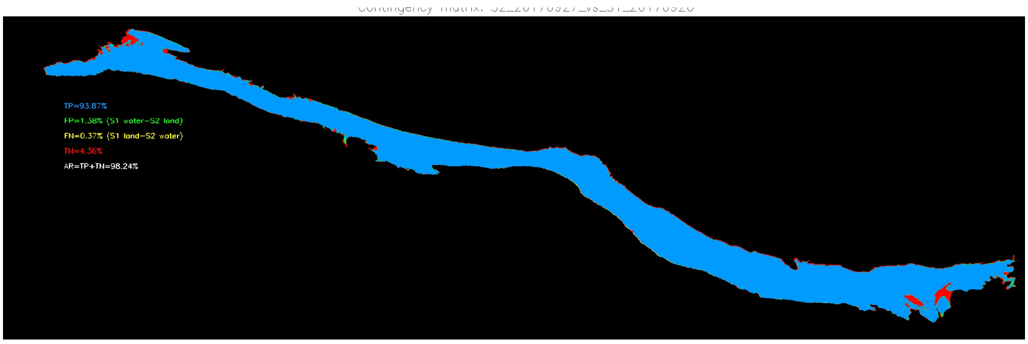

Figure 3 Comparison between S1 and S2 estimates over Altevatn for 20190927 Blue and red indicates True positives (S1 and S2 both indicate water) and true negatives (S1 and S2 both indicate land). A small fraction of pixels is classified different (yellow and green) by S1 and S2. The overall accuracy is 98%.

By combining many such pairs of S1 and S2 observations we can make an overall comparison of how good the two sensors are capable of estimating the water area. For Altevatn in the period 2016-2018 we found 12 products that could be compared.

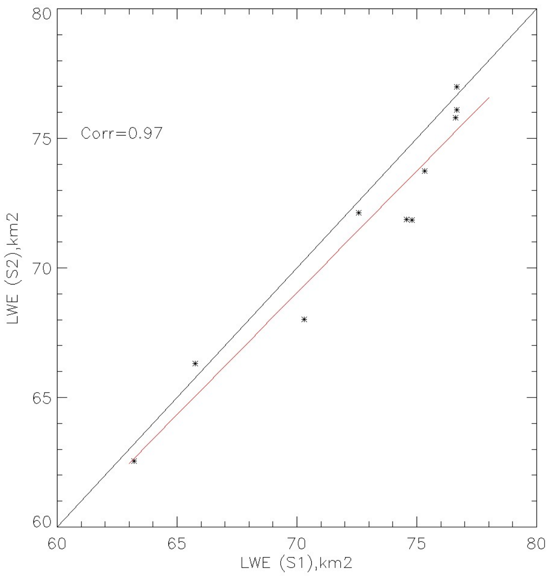

Figure 4 Comparison between S1 and S2 estimates of LWE. The overall correlation is 0.97.

Method 2:

In this method we use time-series of S1 LWE estimates and align them with accurate in situ LWL measurements. This allows us to calculate the regression between LWL and LWE (hypsometry) and the resulting accuracy of the model follows.

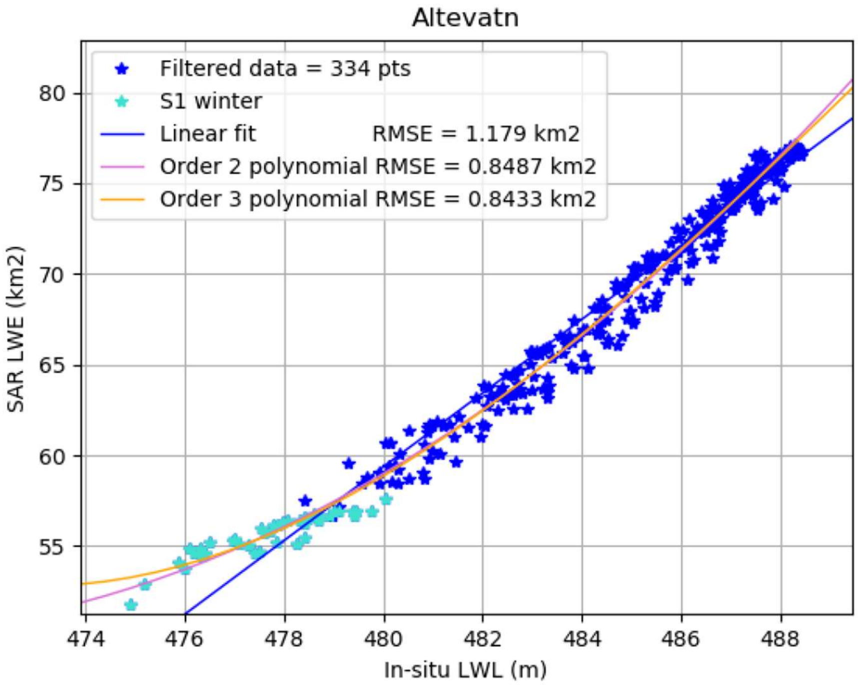

Figure 5 LWE-LWL regression for Altevatn. A polynomial fit of the thrid order gives an overall RMSE=0.84 km^2 or 1.3% error in area estimates.

Method 3: In the third method we use the hypsometry analytical function (in general a fit of either LWE = a.LWL (linear), or LWE= a.LWL^2+b.LWL + c (first-degree polynomial)) to calculate the LWE variable once a measurement of LWL has been performed and produced. To invert the polynomial coefficient we need to use a set of vectors at different dates (LWL, LWE) measured using the satellite altimetry for LWL and the satellite imagery for LWE. We generally limit the number of vectors (LWL,LWE) to 10 to 15 per lake, which in general is largely enough to cover the whole range of level and extent variabilities. Another way of validating the method is to calculate the RMS of the differences between the measured values of LWE (from satellite imagery processing) minus the theoretical value of LWE (calculated using hypsometry function).

This is done for all lakes.

In Figure 6 an example for Lake Dagze-co (located over the Tibetan Plateau) is given. The RMS is expressed in km^2 which in this case represents less than 1% of the total surface of the lake. We have done this test on a list of 10 lakes over the Tibetan plateau, and the results have converged to a value of 1 to 2% of error compared to the total extent of each of these lakes. In Figure 6, the green dots correspond to vectors (LWL, LWE) calculated by external collaborator Guoking Zhang for intercomparison.

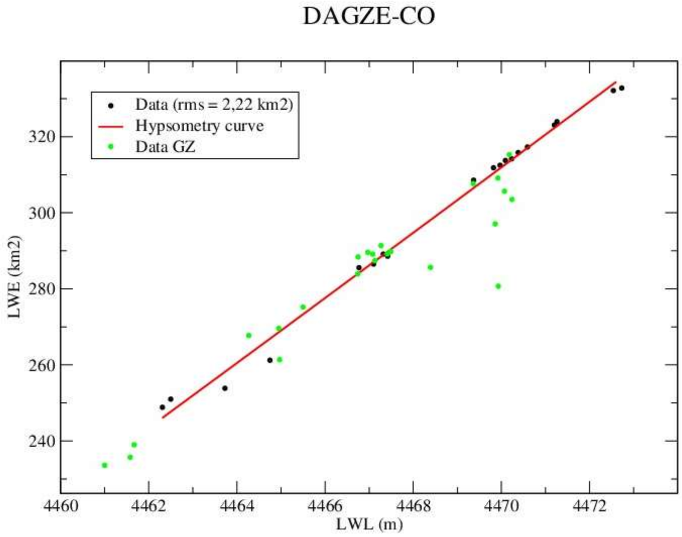

Figure 6 LWE vs LWL in Lake Dagze-co

3.4. Known constraints for LWE validation

Only method 3 will deliver validation results for all selected lakes. Method 2 could also be processed extensively for all simultaneous S1/S2 datasets. Method 1 hinges on the creation of new datasets.

An obvious drawback with method 1 is that we compare two EO datasets, where it can be subjective to determine which is most correct. There is a need for a thorough visual inspection of the two products. This could partially be overcome by using a third dataset (e.g. high resolution optical and/or SAR sensor data), but in practice it is very challenging and costly to gain access to such data over lakes where we have simultaneous measurements of S1 and S2.

The main drawback for method 2 is that in situ measurements of LWL are difficult to obtain, so there are few datasets. Validation against one lake may not be relevant for other lakes since we can observe large differences between the algorithms for different lakes/different conditions.

The main drawback using altimeters is that the accuracy of altimeter data is variable, and expectedly poorer than in situ data. Also, the fact that altimeters are not synchronized in time with S1/S2, add to the uncertainty.

At this stage of the project it is not possible to provide an uncertainty associated with the product. As mentioned previously, the first obstacle it is the validation of the product by itself. What will be done it is to assess the difference in term of water surface depending of the exploited imagery, i.e. a relative uncertainty depending of the exploited source. It is expected that, through product validation, a range of product bias estimates can be derived per sensor and resolution.

4. Lake Surface Water Temperature (LSWT) parameter

4.1. LSWT Validation activities

LSWT validation consists of comparing retrieved water-surface (skin) temperature with in situ temperature measurements, which are obtained at a point within the surface area to which the retrieved temperature applies. This means that LSWT is validated at full satellite resolution (L2 internal products, per pixel) and at L3 (on the gridded 1/120 degree LSWT product).

In situ temperature measurements in lakes are not common, and the key activity for this purpose is obtaining the maximum number of in situ data points for comparison. There is no international system for data sharing of in situ LSWT measurements, and over many years (within the projects ARC Lake, GloboLakes and EUSTACE) the team has developed a network of professional connections who share data for our validation use. The collection happens once a year, towards the end of the calendar year. Only data that have at least daily temporal resolution are considered. A significant effort is required to re-format the data received into a common format. Data were obtained in the last effort for 113 observation locations covering 43 lakes that can be remotely sensed (see Figure 7), but this should slowly increase year by year.

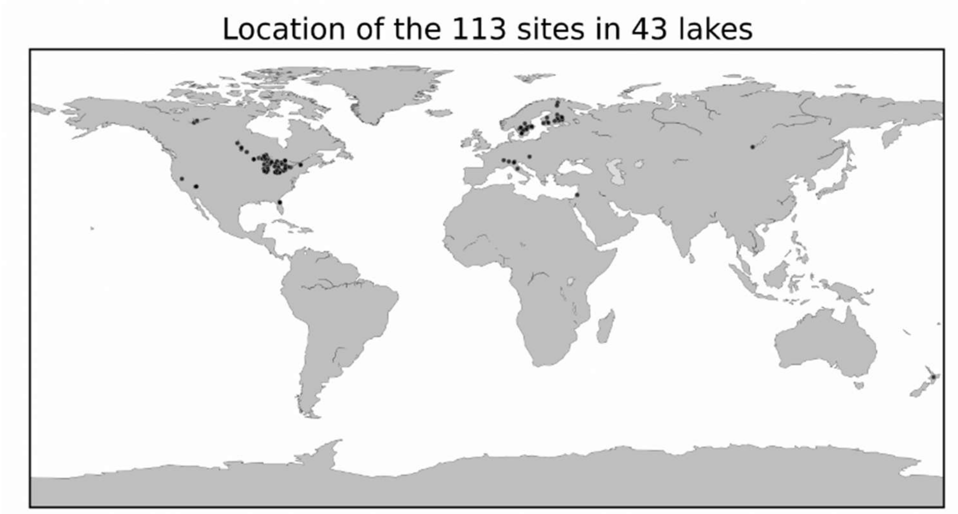

Figure 7 Locations of 113 in 43 lakes that can be remotely sensed for LSWT using 1 km infrared imagers such as SLSTR.

The validation of the L2 product will also give de facto the validation performance for the merged product, since these are at effectively the same spatial resolution.

4.2. Schedule for LSWT validation

Validation activity	*Time frame for analysis*	*Implementation*
Collection of in situ data from professional network	November every year	All inputs reformatted to a standard specification
Matching between in situ data and L2 and L3 LSWT products	3 months after product generation	Matches are added to a match-up data base (netCDF)
L2 comparisons and L3 comparisons	5 months after product generation	Statistics and plots stratified by quality level are generated
Document	PVR deadlines	Present results in Product Validation Report

4.3. Inputs and methods for LSWT validation

The collected in situ data contain as a minimum daily resolution LSWT (preferably sub-daily) with time, longitude and latitude. Metadata on depth is preferable, but data are accepted if we are confident they are surface data even if depth metadata are absent. We have no control on the variety of in situ instrumentation, but essentially all are thermistor measurements. We do not generally have specific uncertainty estimates for the in situ measurements or their location, but typically thermistors have uncertainty <0.25 K and locations are specified to well within the 1 km pixel of the highest resolution satellite observations.

In situ data are quality controlled by ensuring they do not stray too far from a climatology (based on the v3 LSWT product) or fluctuate unphysically quickly (see Figure 8).

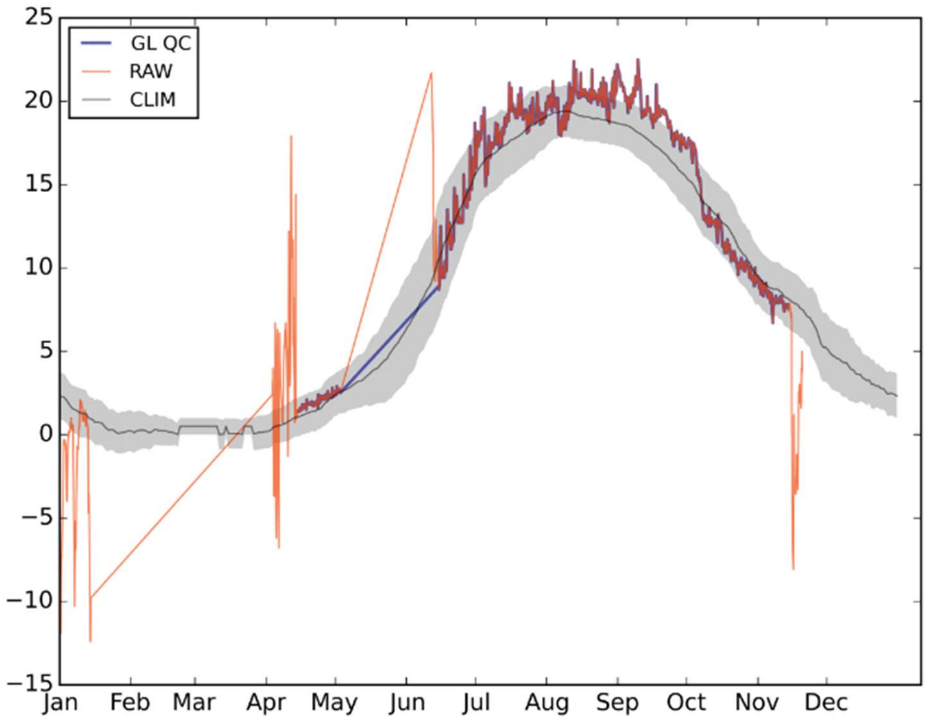

Figure 8 Illustration of quality control results, in this case for Lake Huron in year 2002. The red line is the received in situ data: when thin, this indicates periods failing quality control, and when thick, data that pass quality control. The black line and grey band show the climatological annual cycle for this location and its inter-annual variability (one sigma).

Quality controlled in situ data are matched to L2 and L3 data (ie., valid, clear-sky LSWTs), using the criterion that the in situ data fall within the area of the satellite observation and within 1 day time separation (or are daily mean data). The products are validated by quality level and by minimum quality level, using two standard tables shown below (with purely illustrative numbers):

Table 2 Validation statistics by quality level

	Median	RSD	Mean	SD	N
ql=5	-0.300	0.430	-0.389	0.935	2749
ql=4	-0.440	0.652	-0.590	1.242	1827
ql=3	-0.745	1.082	-0.946	1.620	900
ql=2	-1.445	1.749	-1.636	1.945	308
ql=1	-3.310	4.018	-4.110	4.342	1552

Table 3 Validation by minimum quality level

	Median	RSD	Mean	SD	N
ql>3	-0.340	0.504	-0.469	1.073	4576
ql>2	-0.370	0.578	-0.548	1.193	5476
ql>1	-0.390	0.608	-0.606	1.268	5784
ql>0	-0.510	0.860	-1.347	2.703	7336

These two forms relate to how users may use quality level information. For example, a user may wish (as we recommend) to use all data with quality level 4 and 5.

The statistics are: median difference; robust standard deviation of difference, calculated as 1.27 times the median absolute deviation from the median; mean; and standard deviation. The median and RSD are less influenced by outliers.

Note that the expected median difference is in the range –0.15 to –0.25 K, because of the water skin effect whereby the thin (~0.1 mm) surface layer of the water body is cooler than water below because (typically) of heat loss from the lake to air.

The methods for L3 validation are essentially the same, with the gridded data substituted.

Optimal estimation is used for LSWT retrieval, and this method returns an estimated retrieval uncertainty according to standard equations. These results are then used to provide internal uncertainty estimates for the gridded products. The uncertainty estimates assume Gaussian statistics and in principle could be validated against the robust validation statistics (RSD in the tables).

The RSD of satellite-in-situ difference is explained by three terms (to be added in quadrature): the retrieval uncertainty, the in situ uncertainty and the consequences of true geophysical variability between the satellite and the match (match uncertainty). Quantitative validation of the retrieval uncertainty from the RSD difference requires additionally precise knowledge of the other two terms. For the case of LSWT validation data, neither of these terms is well quantified, because the validation data are collected from an informal network of contacts and are collected using an unknown variety of sensors and sample-location methodologies. A lower limit on the combined in situ measurement and match uncertainties comes from analogy with sea surface temperature matches to drifting buoy matches in the open ocean, which have been intensively studied: the lower limit is 0.2 K accounts for in situ and match uncertainties combined. This in turn places an upper limit on the retrieval uncertainty of LSWT of 0.38 K (for ql = 5). The provided internal uncertainties vary around 0.35 K, so this, as far as it goes, is consistent.

4.4. Known constraints for LSWT validation

The principal constraint is that the number of remotely-sensible lakes with in situ data, which is limited. So, for example, the total number of matches for the whole AATSR mission (10 years) for all lakes is around 7300.

5. Lake Ice Cover (LIC) parameter

5.1. LIC Validation activities

In Lakes_cci, the lake ice cover (LIC) is a level 3 (L3) product generated from MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath, Collection 6 on a 250-m grid (at the Equator) then temporally (daily) and spatially aggregated onto the harmonized grid (spatial resolution of 1/120 degrees). Assessment will be performed on the 250-m product following three methods: 1) comparison against ground-based nearshore observations; 2) comparison with LIC products generated by other algorithms/groups such as NASA; and 3) validation against pixels extracted for a selection of lakes from visual interpretation of original imagery on input product used as input into the LIC retrieval algorithm. Assessment is also planned for the harmonized daily LIC product (near 1 km at the equator) against the Interactive Multisensor Snow and Ice Mapping System (IMS) 1 km product (available since December 2014) and ground-based nearshore ice/open water observations. A description of the methods and related inputs are described in Section 5.3 and foreseen challenges in Section 5.4.

5.2. Schedule for LIC validation

Validation activity	*Time frame for analysis*	*Implementation*
LIC algorithm “internal” validation by H2O on 250-m product, improvement and uncertainty characterisation for MODIS	02/2019 – 11/2019	MODIS algorithm and uncertainty in CDR v1
Comparison of LIC v1 (250-m and harmonized products) against ground-based and other satellite-derived LIC products. Round-robin evaluation by user groups/collaborators (ECMWF, CIS, NOAA)	11/2019 – 06/2020	MODIS evolution of algorithm and processing chain for CDR v2
LIC algorithm internal evaluation and uncertainty characterisation for VIIRS and OLCI/SLSTR	03/2020 – 10/2020	VIIRS and OLCI/SLSTR algorithms and uncertainty in CDR v2
“Internal” validation by H2O and round-robin evaluation of LIC product v2 by collaborators (ECMWF, CIS, NOAA) and other user groups	11/2020 – 06/2021	

5.3. Inputs and methods for LIC validation

Inputs for validation of the LIC product depend on the method selected. Each method and planned datasets for validation are described below.

Method 1: Comparison against ground-based nearshore observations

This method involves a comparison between satellite-derived ice/open water and ground-based observations of ice presence/absence. Ground-based ice observations are made along lake shores, sometimes in bays, such that they can only be compared to one or a few satellite pixels located close to shore but yet uncontaminated by land. Lake sites where ice observations are recorded coincident with satellite derived LIC product dates are likely to be small in number since there has been a drastic decline in such sites globally over the last three decades. When ground-based observations are available, they will be compared to MODIS LIC (250-m and harmonized daily at 1/120 degrees) from pixels located within 500-1,000 m from shore as to avoid land contamination (to be tested).

Method 2: Comparison with LIC products generated from other algorithms/groups

The Lakes_cci LIC 250-m product will be compared to ice products generated by other groups where and when possible (i.e. same lakes and overlapping time periods). At the global scale, the LIC product will be compared to NASA's Terra and Aqua (MOD10/MYD10) 500-m Snow Products (Collection 6) which identify ice, open water and clouds over lakes. For northern Europe, the product will also be compared to the lake ice extent (LIE) 250-m product derived from MODIS on Terra (not available for Aqua and only starting in March 2017) as part of the Copernicus Global Land Service (CGLS). For a common set of lakes, pixels will be randomly selected to determine overall agreement/difference between products in the identification of ice, open water and clouds. For the harmonized daily LIC product comparison is planned against the IMS 1 km product (starting December 2014).

Method 3: Validation against sets of pixels extracted for a selection of lakes from visual interpretation of MODIS reflectance imagery

Most importantly, the Lakes_cci LIC product will be validated against ice, open water and clouds determined from the visual interpretation of MODIS Terra/Aqua 250-m Atmospherically Corrected Surface Reflectance 5-Min L2 Swath data – the primary product used as input in the LIC retrieval algorithm – during both the freeze-up and break-up periods. Sets of pixels will be randomly chosen from a selection of lakes distributed globally to assess the accuracy of the LIC product (class uncertainty computed from confusion matrix). The goal is to achieve the required measurement uncertainty of 10% set by GCOS.

5.4. Known constraints for LIC validation

There are no foreseen constraints for the validation/cross-comparison of the 250-m LIC product which is generated at an intermediate step before creation of the harmonized product (daily and 1/120 degrees). Greater challenges are expected for assessment of the harmonized product. First, the IMS 1 km product is a "cloud-free" product generated from ice/snow analysts who assign all grid cells in a lake as either ice or open water. Therefore, comparison will be made only for ice/open water matchups of the harmonized product and IMS (focusing on clear days). In addition, the IMS product on a 1 km grid has only been produced since December 2014. Hence, the number of matchups may be low. Finally, the use of ground-based observations may be of limited use (TBD) for evaluation of the harmonized product since observations are made close to shore while the harmonized product provides ice/open water presence over a larger distance away from shore.

Validation and product uncertainties

As described in section 5.3, product validation activities are to include comparison against ground-based nearshore ice/open water observations and LIC products generated from other algorithms/groups, and validation against sets of pixels extracted from a selection of lakes from visual interpretation of MODIS RGB reflectance imagery. The latter approach has been used internally to quantify total uncertainty of the LIC product through computation of a confusion matrix built on an independent statistical validation. Classification error (%) derived from the confusion matrix is the metric used to report total uncertainty for each class (ice, water, cloud). Hence, in the case of LIC, there is a strong relation between validation activities and product uncertainty characterization.

The main sources of uncertainty of the LIC product are described in detail in E3UB. They have not been quantified independently. The identified sources include MODIS Aqua/Terra detectors noise/sensors degradation, observation noise (optical thickness of the atmosphere and view zenith angle/solar zenith angle), the threshold-based retrieval algorithm used, and misclassification. The round-robin evaluation by user groups (see schedule provided under section 5.2) will permit to identify in space (different lakes and lake sections) and in time (ice and open water seasons) where and when one or more of these sources of uncertainty are affecting the quality of the LIC product.

6. Lake Water-Leaving Reflectance (LWLR) parameter

6.1. LWLR Validation activities

Validation of LWLR concentrates on the atmospherically corrected LWLR and derived optical-biogeochemical water column properties, including (at least for V1 of the CDR) chlorophyll-*a* and turbidity.

The processing chain for the variables derived from LWLR includes a dynamic mapping of algorithms depending on lake Optical Water Type (OWT). The benefit of this approach is that in situ data belonging to the same optical water type can be pooled together from geographically different sources to firstly calibrate and then validate the system for that OWT. This also means that the algorithm is then expected to perform equally well over waterbodies exhibiting the same OWT, but for which no in situ data are available, which constitute the vast majority of waterbodies.

Validation activities in Lakes_cci take the form of round-robin comparisons where algorithm selection is not yet established as a result of prior research, typically followed by algorithm optimisation resulting in per-sensor and per-OWT algorithm definitions and associated uncertainty models.

Ultimately, the procedure followed per sensor, per variable of interest and per OWT depends on the availability of in situ matchup data, which is typically scarce. The following considerations are important when selecting data for validation of LWLR:

- The parameter of interest, either LWLR from in situ radiometry or a biogeochemical or physico-chemical component of the water column
- The sampling depth, and whether it can be assumed to represent the water column that is visible from the remote sensor
- The (expected) accuracy of the in situ measurement
- The sampling location, particularly whether close to shore or on open water
- The time window allowed for in situ and satellite comparison, which depends on whether absolute (narrow window) or relative performance between algorithms (wide window) is evaluated. Typically, the time window will vary from ±1 to ±7 days.

6.2. Schedule for LWLR validation

Due to scarcity of recent in situ data in common archives we do not currently plan to have a specific algorithm calibration, validation and uncertainty characterisation for the OLCI observation period, which will instead be based on statistical sensor inter-bias correction and propagation of MERIS algorithms. It is, however, expected that the LIMNADES database (see Section 6.3) will grow to include OLCI observations once the database is re-launched in 2019/2020, at which point this plan may be revised.

For the MERIS and MODIS sensors LIMNADES contains sufficient observation data for initial algorithm calibration and uncertainty characterisation, based on initial estimates. Further investigation of the optical variability in the data set is required to establish whether end-to-end validation is possible for each optical water type used to select algorithms during processing, and whether this can be done in addition to algorithm calibration. This further depends on the license attached to each data set, and permissions obtained from contributors, which is an ongoing activity in 2019. Table 4 lists the priorities in terms of calibration, validation and uncertainty characterisation with their intended time frame. Top priority is to include the mechanism for uncertainty characterisation to the MERIS (and propagated to OLCI) in the first lakes climate data record of Lakes_cci, followed by several activities to evolve the processing chain towards version 2.

Table 4 Validation activities per sensor and release of the Lakes ECV Climate Data Record.

Validation activity	*Time frame for analysis*	*Implementation*
LIMNADES algorithm calibration and uncertainty characterisation for MERIS	Jan-Dec 2019	MERIS and OLCI (propagated) algorithms and pixel uncertainties in CDR v1
LIMNADES in situ matchups with MODIS and MERIS observations. A round-robin evaluation and selection of atmospheric correction methods.	Nov 2019 – Jun 2020	MODIS evolution of the processing chain for CDR v2
LIMNADES algorithm calibration and uncertainty characterisation for MODIS and VIIRS	Mar – Oct 2020	MODIS algorithms and pixel uncertainties in CDR v2
LIMNADES algorithm calibration exercise and uncertainty characterisation for OLCI	Pending additional in situ data	OLCI algorithms and uncertainties CDR V2
LIMNADES algorithm calibration exercise and uncertainty characterisation for SeaWIFS	Pending initial in situ data analysis (geographic bias, lake types)	SeaWIFS algorithms and uncertainties CDR V2

6.3. Inputs and methods for LWLR validation

We recognise two separate sources of in situ validation data: those collected as part of national monitoring programmes and those collected by research institutes for biogeochemical or bio-optical research. Challenges with statutory monitoring data include shore-based sampling which has very limited value as reference for optical remote sensing, and common lack of accuracy in recording sampling locations (often documented as static but in practise varying). Furthermore, such data sets are invariably limited to biogeochemical and physicochemical observations, therefore only supporting end-to-end validation of biogeochemical products without validating LWLR along the way. This is a problem because the dominant source of uncertainty in the final product is the atmospheric correction which can only be established by having reference measurements of LWLR. Language may form a further accessibility barrier in accessing these data sets, and therefore geographic bias is a further issue. For these reasons, very few data sets from national monitoring programmes have been successfully incorporated in large scale satellite matchup activities. National monitoring data sets are being increasingly made open access, so their uptake may increase in future.

The challenge with existing research-quality in situ data sets is that these originate from a scattered landscape of limnological research laboratories using varying measurement protocols (not always documented) and a widely varying set of license terms and accessibility issues. The effort associated with ensuring licenses are established and then honoured when resulting in reports and publications is substantial. To overcome some of these obstacles, LIMNADES, a community-owned bio-optical data archive hosted by University of Stirling, was launched during the UK-GloboLakes project, with many research teams contributing in situ campaign data. License terms were updated in 2019 to support wider use of the data set. Because the LIMNADES initiative was not initially a funded activity and the effort of collating and harmonising the contributions is substantial, private and public access to the data has been poor since its inception. Current initiatives, notably the H2020-MONOCLE project, have invested effort in making the database contents available on demand (while respecting licenses and embargoes) through a set of web services compatible with Open Geospatial Consortium standards. LIMNADES primarily includes observations from inland water bodies and additionally includes some near-coastal observations.

The number of in situ LWLR observations in LIMNADES is limited to the MERIS/MODIS/VIIRS observation period (Figure 9), whereas earlier observations overlapping with SEAWIFS include only biogeochemical and physico-chemical observations. It should be noted that the histograms shown include data under any license and not all may be available to the project without further consultation of contributing groups, and optical-biogeochemical observations may be based on poorly calibrated sensors (e.g. fluoroprobes) which cannot be used for absolute validation. The data shown have also not been screened for the number of actual cloud-free matchups with satellite data, which typically reduces the number of observations to approximately 10% of the number shown, depending on the allowable matchup time-window, satellite overpass frequency and (seasonal) cloudiness.

6.4. Known constraints for LWLR validation

Because some OWTs are more commonly observed than others and because in situ bio-optical and remote sensing research in the last decades has focussed on waterbodies suffering the effects of eutrophication, it is not possible to produce a set of satellite matchups for each OWT that is sufficiently large to separate calibration, validation and uncertainty characterisation. This consideration is part of the research conducted within the Lakes_cci project.

Validation and product uncertainties

The E3UB document describes in detail why and how product validation and product uncertainty are necessarily linked for LSWT and derived variables. In brief, the application of fuzzy pixel classification and algorithm selection in combination with non-linear optimization techniques favours statistical uncertainty characterization on the final product over analytical error propagation. In addition, in situ reference data to inform uncertainty characterization are unlikely to cover the range observed in nature. The use of in situ data for uncertainty characterization has two implications for the validation of the products and product uncertainty itself:

- The uncertainty characterization approach must take the range of the in situ reference data into account, i.e. extrapolation of product uncertainty beyond the value range of the validation data shall not be allowed. Ultimately, uncertainty maps will thus also in a spatial sense reflect regions where product uncertainty is not yet known. This can then guide further field work (into affected and/or new optical water types) contributing to both product validation and broader uncertainty reporting.
- Uncertainty in the in situ validation data, due to the use of different (even if optimised) protocols by contributing research groups, will contribute to product uncertainty. Divergence in observation protocols is likely to show as systematic bias that is different between contributed datasets – at present this is expected to be a minor effect. Another effect stems from divergence in instrument deployment protocols, such as the frequency of maintenance or the distance from shore at which samples are taken, which could introduce more significant bias.

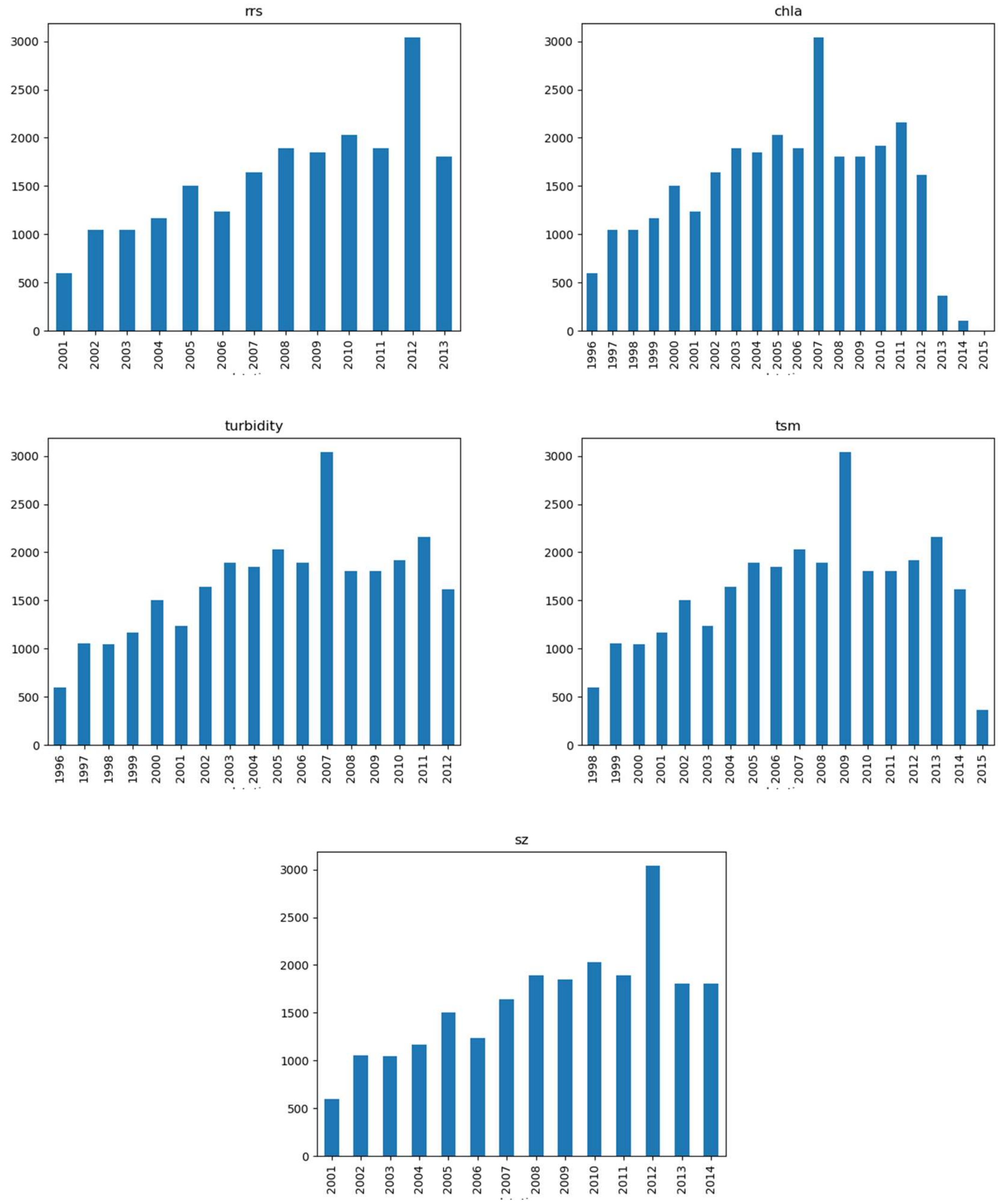

Figure 9 Availability of in situ data per year (source: LIMNADES). Rrs = Remote-sensing reflectance, chla = chlorophyll-*a* (including in vivo fluorescence), tsm = total suspended matter dry weight, sz = secchi disk transparency.

7. Lakes cci product consistency validation

Five use cases studies are undertaken in the project to address key questions in which the Lakes_cci products provide important inputs. Overall, the development of use cases depends on auxiliary in situ data, EO algorithms and climate models, while part of the activities might also include validation plans. The use cases are due to start with the release of the V1 of the Lakes_cci climate data record and will be carried out during the second project year. In this case it is expected that these activities will provide a further independent validation of the satellite derived products. Because the use cases are carried out by the Lakes_cci team, product quality feedback will be taken up during the course of these studies, allowing improvements to be researched or implemented well before V2 of the climate data record is due. The type of validation considered by each uses case is therefore described in the following paragraphs, along with the main objectives of the use cases.

Overall, the use cases might consider different type of validation (stability vs. accuracy vs. consistency). Then, in situ reference data are often considered as the main source of data to validate satellite products, while in some case also ecological or hydrodynamic modelling might provide a benchmark for validation.

The proposed validation activities for each use case, along with a summary of the goals and satellite-derived variables to be investigated, is as follows.

Use case #1 - Changes of lakes in Greenland: multi-ECV analysis.

Global increase in LSWTs have been documented widely in recent years, with summer LSWTs in some regions warming faster than ambient air temperature. In particular, lakes in Greenland, on which this use case is focusing on, are located in an area very sensitive to global warming.

LSWTs products of these lakes will be analysed to assess the impact of climate change on water stratification.

For this use case most of the validation activities will be based on evaluating the consistency with other variables. In particular, it has been suggested that reductions in LIC are related mechanistically to the concomitant increase in summer LSWT, possibly through ice-albedo feedbacks and the timing and duration of the summer stratification period.

Lake morphology can undergo rapid change in this domain. The area-of-interest will take into account the recently produced Glacier_cci high resolution water extent map as a possible source of information about varying lake extents within the context of this analysis.

Use case #2 - Analysis and interpretation of ECV for large lakes

This use case is focused on the analysis and interpretation of LWLR and LSWT for large lakes (>500 km2) in Europe, Africa, Central and South Asia and Central and South America. The aim is to determine the temporal and spatial dynamics of LWLR in relation to global warming and to a climatic gradient.

In situ data available through the LIMNADES database (https://limnades.stir.ac.uk) and aligned projects (e.g. H2020 EOMORES, H2020 MONOCLE, H2020 Danubius, NERC GloboLakes) and collaborators (GEO-Aquawatch) will be used to validate the individual variables and derived products. In situ data will include remote sensing reflectance, surface temperature but also chlorophyll-a, coloured dissolved organic matter, total suspended sediment and phycocyanin concentrations.

Feedback in terms of product (in)consistency will focus on improved product masking and attribution of problem areas.

Use case #3 - Exploiting ECV in Long Term Ecosystem Research

This use case exploits the Lakes-cci datasets to meet the needs of existing consolidated networks, which deal with the development of management options in response to global change. In particular, this use case refers to the Long-Term Ecosystem Research (LTER) network that concentrates on studies of ecological processes that play out at time scales spanning decades to centuries. Long-term data sets from the LTER program provide a context to evaluate the nature and pace of ecological change, to interpret its effects, and to forecast the range of future biological responses to variation. Long-term data are typically compiled from measurements gathered in the field and satellite observations can contribute to these series by uniquely providing timely, frequent and synoptic information from local to global scales. The main satellite data providing inputs to this use case are: LWL, LWST and LWLR and for the latter, chlorophyll-a concentration and water turbidity in particular.

Part of the activity to be developed in this use case will include the evaluation of trends on lake parameters as derived from both satellite observations (from Lake_cci) and in situ measurements (from LTER). Because there will be several independent sources of time-series data, statistical methods suitable to handle multi-temporal data are used. First, techniques will be explored to fill temporal gaps that might exist in both datasets. Then, for the time line in which most correspondence between field and satellite data will be available, consistency in terms of trends rather than absolute values will be analysed. Where possible, the degree of fitting will be analysed by considering the effects of spatial scale in terms of pixel size window or in terms of position of the radar altimetry track with respect to the sampling station.

Use case #4 - Brownification in Scandinavian lakes

This use case is meant to quantify and analyse the temporal evolution of the brownification in selected Scandinavian lakes using the satellite products timing of ice melting from LIC, LWLR and related water constituent products, in particular the absorption in the blue due to coloured dissolved organic matter (CDOM). CDOM is not provided in the CDR v1 because globally validated algorithms are lacking. This use case, therefore, analyses whether lake hue and intensity are good proxies for the effects of CDOM on lake colour.

There is a natural intra-annual cycle of the CDOM content. Maximum CDOM levels normally occur after snow/ice thawing and spring floods, and usually slightly decrease over the season. Heavy rain falls might increase the level of organic matter during summer and autumn. Organic matter is a natural degradation product from plants and boreal lakes are characterized by a high content of dissolved organic matter, which strongly influences the lake colour. The verification of such aspect will be used to validate the products in terms of consistency of different variables. Therefore, time series of the different variables will enable a good assessment of seasonal trends, consistency of adjacent measurements and correlation between different parameters. Time series will be drawn as time series plots and as heatmaps. Correlation between different parameters will be given in scatterplots. Whenever possible, field measurements will be used as an additional source of reference data.

Feedback in terms of product (in)consistency will focus on LWLR, in particular the suspected consistent underestimation of LWLR in relatively turbid and small lakes. The brownification trend is expected to be spatially relatively homogeneous. Thus, spatial inconsistencies in the retrieval of LWLR due to inaccurate atmospheric correction may be demonstrated in this study, which could then be used to improve product uncertainty models.

Use case #5 - Lakes ECV products consistency in the Danube river-lake-lagoon

This use case is aims to validate the consistency of ECV variables in the Danube river- lake-lagoon system, particularly in the highly dynamic Danube delta and connected Razelm Sinoe Lagoon System. Lagoons are not widely studied using remotely sensed data. Here, the study will use information from satellite data records on the temporal dynamics of LSWT, LIC, LWE (possibly LWL) and LWLR (mainly turbidity). In situ measurements will be also used to evaluate the products whenever there is a match-up between the two set of data. Finally, satellite data are expected to fill essential data gaps between field campaigns to corroborate (or not) trends in the in situ data.