D4.1. Product Validation and Intercomparison Report (PVIR)

Reference: CCI-LAKES2-0022-PVIR

Version: 3.0.0 - Issue: 1

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 VERSION/ISSUE

Version	Issue	Date	Object	Written by
1.0	1.0	30 March 2020	Initial Version	PML: Xiaohan Liu, Stefan Simis CLS: Beatriz Calmettes H2O: Claude Duguay SERTIT: Hervé Yésou
	1.1	30 April 2020	Adding Chapter LSWT and minor corrections	UoR: Chris Merchant CLS: B Calmettes / B Coulon
	1.2	30 Sept. 2020	Revision following ESA Review Adding results and minor corrections in LWE part	ICUBE-SERTIT : H. Yésou NORCE : E. Malmes TRE-Altamira : P.Blanco Legos : JF Crétaux CLS: B Calmettes / B Coulon
	1.3	4 Oct. 2021	Accompanying CRDP v1.1 (direct product validation) and evaluating CRDP v1.0 and v1.1	C. Duguay
	1.4	11 Nov. 2021	Revision following ESA Review	B. Calmettes
2.0	2.0	22 Apr. 2022	Accompanying CRDP v2.0 (direct product validation) and evaluating CRDP v1.0 and v1.1 (user feedback, use cases)	PML: Xiaohan Liu, Stefan Simis CLS: Beatriz Calmettes H2O: Claude Duguay UoR: Chris Merchant/Laura Carrea Legos: JF Crétaux
	2.1	6 July 2022	Revision following ESA Review	B. Calmettes
2.1	0	29 January 2024	Accompanying CRDP v2.1 (direct product validation)	PML: Xiaohan Liu, Stefan Simis CLS: Beatriz Calmettes, Anna, Mangilli H2O: Claude Duguay UoR: Laura Carrea, Chris Merchant Legos: Jean-Francois Crétaux ICUBE SERTIT: Hervé Yésou BC: Dagmar Müller
2.1	1	4 April 2024	Revision following ESA Review	B . Calmettes
3.0.0	1	10 December 2025	Accompanying CRDP v3.0.x	PML: Xiaohan Liu, Stefan Simis CLS: Beatriz Calmettes, Anna, Mangilli H2O: Claude Duguay UoR: Laura Carrea, Chris Merchant Legos: Jean-Francois Crétaux ICUBE SERTIT: Hervé Yésou BC: Dagmar Müller USTIR: Dalin Jiang, Evangelos Spyrakos

* Prior to version 2.1 only the Issue was used. The version number henceforth tracks the versioning of the relevant version of the CRDP.

Checked by	Stefan Simis – PML	
Approved by	Philippe Mourot - CLS	
Authorized by	Clément Albergel - ESA	

LIST OF CONTENTS

1 Introduction

The Product Validation and Intercomparison Report (PVIR) details the results of validation and intercomparison activities carried out across the Lakes_cci. It describes the datasets used to validate and intercompare the products contained in the Climate Research Data Package (CRDP) and discusses the results of the validation and intercomparison activities.

Whereas the focus of the PVIR is to inform users of validation results, the associated activities are separately described in the Product Validation Plan (PVP) which is disseminated at the start of these activities. The validation strategy is nevertheless described in the PVIR to provide a self-contained account of all validation and intercomparison work.

Specifically, the PVIR is designed to:

- List the data sets (ground-based, radiosonde, in situ, airborne, ship based, etc.) used as independent reference measurements for validation.
- List any EO and model-based products that are used for intercomparison.
- Fully describe results of the validation of the CRDP, including the prognostic uncertainty estimates, taking into account the reliability and accuracy of the independent reference measurements.
- Discuss what relevance the validation applied to one CRDP product level (e.g. Level-2) has for usage of another product level (e.g. Level-3).
- Validate stability of the CRDP, if it is feasible to do so.
- Compare the CRDP with other commonly used EO and model-based products for the same ECV.
- Compare the stability of the CRDP with different EO and model-based products, paying particular attention to breakpoints where different satellites start or stop contributing to the time series.
- Identify and discuss the impacts of any weaknesses in the validation and intercomparison work (e.g. retrieved quantities that cannot be confidently validated due to lack of reference measurements, or due to large uncertainties or poor geographic sampling of the reference measurements)
- Quantitatively compare the quality of the CRDP with the GCOS accuracy requirements, and also with user accuracy requirements reported in the URD (if different from GCOS).
- Identify limitations in the CRDP, e.g. anomalies, or conditions under which the product quality is lower than expected.
- Provide recommendations to the EO developers on aspects of CRDP quality that are priorities for future improvement
- Provide overall conclusions to users on the CRDP quality, and recommendations for use regarding any limitations identified.

The Lakes_cci contains several observation disciplines to arrive at its constituent Lakes ECV Products. Validation opportunities, particularly the availability of in situ reference datasets, vary widely between these disciplines and products. The following sections describe, in sequence, the validation activities and results achieved to date.

2 Lake Water Level - LWL

2.1 Data description

2.1.1 Satellite data

Lake Water Level is the measure of the absolute height of the reflecting water surface beneath the satellite with respect to a vertical datum (geoid), expressed in metres. Per-lake time series have been computed from multiple altimetry satellites from late 1992 to 2022. The time periods used for each satellite/instrument are provided in Table 1 but may vary from one lake to the other, depending on the orbits of the satellites with respect to the location of the lake. All missions for which data were available and considered valid are included, however during tandem overlapping phases (TOPEX/Poseidon/Jason1, Jason1/jason2, Sentinel3A/Sentinel3B) the most precise solution is selected for the time series. Therefore, from one lake to another one, the mission used to generate specific periods in the time series may differ.

Table 1: Time periods for the satellite/instrument used to generate the lake product

Satellite	Instrument	Period
TOPEX/Poseidon (T/P)	Poseidon-1	08/1992 – 01/2002
ERS-2	RA	04/1995 – 06/2003
Jason-1	Poseidon-2	12/2001 – 07/2013
Jason-2	Poseidon-3	06/2008 – 10/2016
Jason-3	Poseidon-3B	01/2016 – 04/2022
Sentinel-6A	Poseidon-4	09/2021 – 12/2023
Envisat	Radar Altimeter (RA-2)	03/2002 – 10/2010
Cryosat-2	SAR interferometric Radar Altimeter (Siral)	04/2010 – 2015
SARAL	AltiKa	02/2013 – 07/2016
Sentinel-3a	SRAL	02/2016 – 12/2023
Sentinel-3b	SRAL	01/2019 – 12/2023

A detailed description of the product generation is provided in the Algorithm Theoretical Basis Document (ATBD) with further information on the product given in the Product User Guide (PUG)

2.1.2 In situ data

External in situ data are useful to assess the quality of the LWL products, providing insight into product precision and accuracy. The comparison with in situ data, using different datums and observation times is not straightforward. A list of datasets used to validate LWL in CRDP v3.0 is provided in Table 2.

Table 2: In situ data sources for LWL validation

Dataset name	Description
ANA	The "Agencia Nacional de Aguas e Saneamiento Basico" (ANA) provides in situ data on reservoirs in Brazil
Hidricos Argentina	The Hidricos Argentina database provides in situ data on national rivers and lakes.
USGS	The database of US Geological Survey provides in situ data
Water Office Canada	This database contains historical in situ hydrometric data in Canada.
FOEN	The Federal Office for the Environment (FOEN) provides hydrological data, in particular for water levels of lakes in Switzerland.

In addition to the external datasets listed, our team have been involved with 18 years of dedicated field work in the framework of satellite altimetry Cal/Val programmes, at Issyk Kul in Kyrgyzstan.

2.2 Comparison methods

The validation of the Lake Water Level involves comparing the cci data with in situ data from different datasets. As the temporal measures in the Lakes_cci and in situ datasets are not the same, only days with observations in both datasets are taken into account.

Additionally, a bias between the satellite time series and in situ measurements is always expected, since they rarely use the same geodetic reference frame. Some results of these comparisons are given in Cretaux et al. (2016) and Ričko et al. (2012). To avoid the problems associated with the geodetic reference, the comparison with in situ measurements was based on the variation in water level rather than the absolute value comparison.

The differences between the LWL and in situ observations are performed by estimating, for each available lake, the URMSE (Unbiased Root Mean Square) between time series as well as the Pearson coefficient indicating the correlation between the time series. The number of measurements used for the estimation of this metrics is also provided, which gives an indication of the validity of the analysis.

2.3 Description of the work

If in such analysis the in situ data are considered as ground truth, there are still difficulties to directly compare them because the observations represent different scales of observation. Satellite measurements are collected along the track of the satellite with a footprint of several square kilometres, while the in situ measurements are generally done on the lakeshore. The reference system of each of the measurements also differ, and uncertainty on the geoid over the lake adds a source of bias between the two types of measurements. Further uncertainty stems from any technical issues with gauges, data gaps, human error in data collection, which combined may attribute errors up to several centimetres. Finally, interpolation of the satellite-derived estimate is required when the in situ measurements are provided as temporal (typically monthly) averages. Uncertainties stemming from water level variations within the aggregation period should then be expected.

The uncertainties of satellite altimetry over lakes depend on several factors. Depending on lake morphology, the altimetry telemetry waveform can be complex compared to open water surfaces such as the ocean or large lakes. The footprint over narrow reservoirs, for example, combines the lake surface with non-water surfaces such as vegetation, bare soil, sandbanks, or ice. This explains how a variation from a few centimetres to several decimetres can occur in altimeter performance over a large set of lakes. Moreover, evolution of the altimetry technique - from the Low-Resolution-Mode altimeters on the

Jason/OSTM series to the SAR altimetry used with the Sentinel-3A/B and Sentinel-6A series – improves the accuracy from the oldest time series to the latest. For example, over Issyk Kul or the Illmen lakes, accuracy has improved by a factor of two since the introduction of Sentinel-3 data.

Sensor drift can be adjusted if it is observed. Root-mean-square differences of unbiased time series are calculated both for the complete time series and for the Jason 3/Sentinel-6A and Sentinel 3A missions.

Lake Issyk Kul in Central Asia was chosen in 2004 as a dedicated calibration / validation site for satellite altimetry over lakes, to circumvent many of the issues with in situ reference observations listed above. Over 20 years of experiments over the large lake (6000 km^2) provide overpasses by all past and 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). Recent campaigns (in 2021, 2022 and 2023) have been organised over the Issyk Kul and have shown that for such a large lake, the accuracy of satellite altimetry of last generation with SAR measurements is better than 2 cm (see Figure 1).

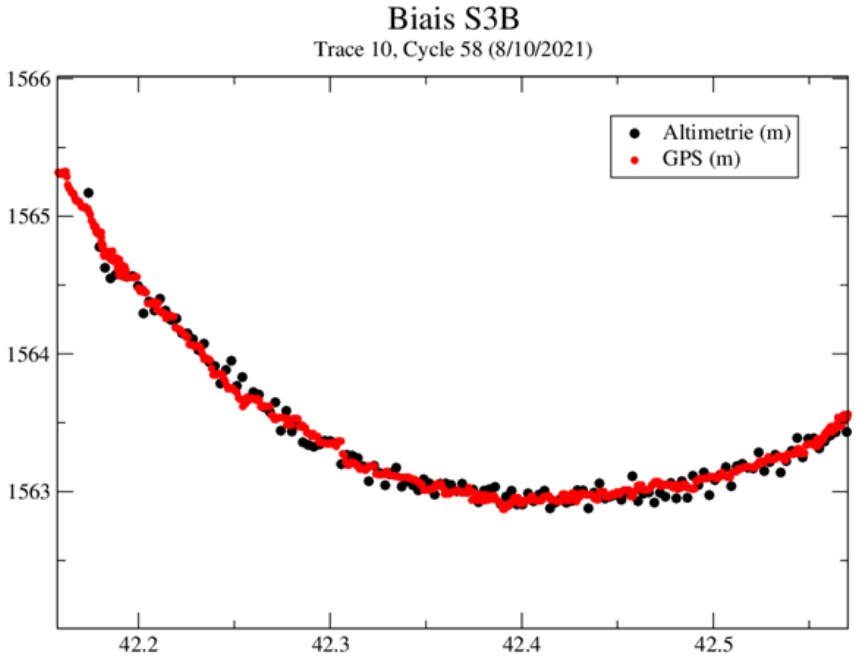


Figure 1: GNSS water height of the lake from field work in 2021, along a track of sentyinel-3B compared to water height measured by sentinel-3B along the track N°10. Absolute bias and standard deviation of the differences are below 2 cm.

2.4 Validation results

2.4.1 Validation at ANA sites

The National Water Agency of Brazil provides information of in situ water level measurements for reservoirs. The Pearson coefficient achieved for water level in reservoirs Sobradinho and Tres Marias, indicated in Table 3, is greater than 0.9 while the unbiased root-mean-squared (URMSE) is less than 40 cm. Figure 1 shows the comparison of the water level variation timeseries from Lakes_cci dataset for the Tres Marias reservoir. Figure 2 contains the histogram of the difference between the time series showing that, as expected, it has an average value centred on zero.

Table 3: ANA LWL comparison

Lake Name	Time period	URMS (m)	Pearson	Number of measurements
Sobradinho	1995/01 – 2022/12	0.397	0.967	462
Tres Marias	1996/02 - 2022/12	0.298	0.985	300

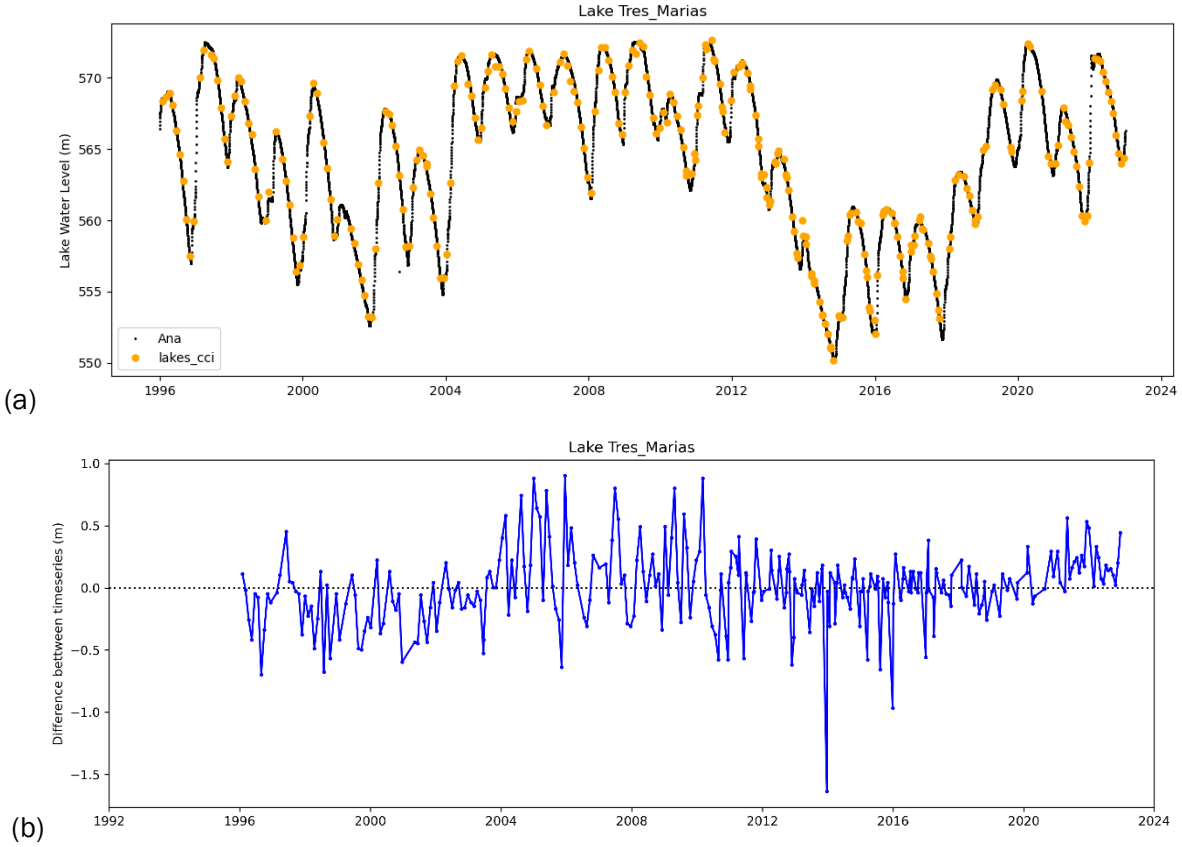


Figure 1: Reservoir Tres Marias - Comparison to ANA in situ dataset (a) time series and (b) difference between common measurements

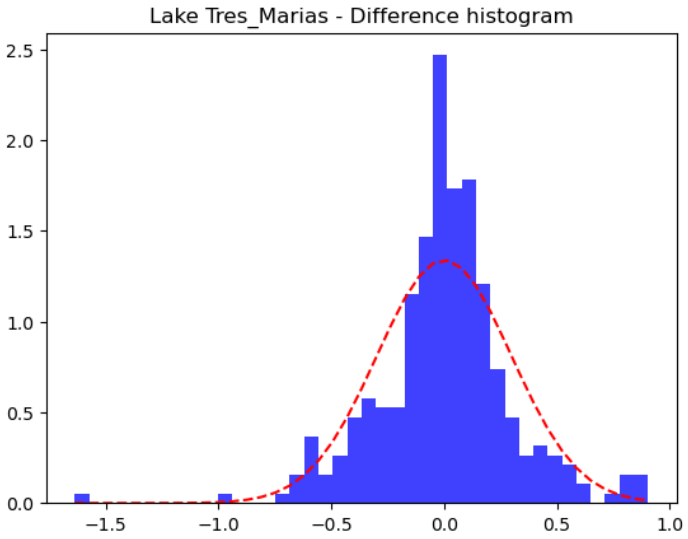


Figure 2: Lake Tres Marias. Histogram of the difference.

2.4.2 Validation at Hidricos Argentina sites

The information concerning the variation on the Water Lever for five lakes in Argentina was obtained from the *Base de Datos Hidrologica integrada* (BDHI)[1]. For those lakes, the URMS and Pearson coefficient were estimated (Table 4). As described in the Comparison methods section (Section 2.2), the water level is previously unbiased to avoid reference issues.

Table 4: Hidricos Argentina LWL Comparison

Lake Name	Time period	URMS(cm)	Pearson	Number of measurements
Argentino	1993/01 - 2022/10	16.76	0.98	946
Buenos Aires (General Carrera)	2008/09 - 2022/09	49.64	0.59	164
Cochrane	2010/03 - 2022/09	13.19	0.87	216
San Martin	2016/03 - 2022/09	67.93	0.85	196
Viedma	2013/03 - 2022/09	14.38	0.98	265

The lake with the lowest Pearson correlation coefficient value is Buenos Aires Lake, also known as General Carrera Lake. A larger difference is observed prior to 2016 when positive bias is observed (Figure 3) and the measurements with old missions are much noisier. A better correlation (Pearson coefficient 0.683 estimated with 91 measurements) and lower URMS (37.29 cm) is observed after 2016. Outliers in the Hidricos Argentina dataset in 2021 also contribute to the lower correlation coefficient.

[1] bdhi.hidricosargentina.gob.ar

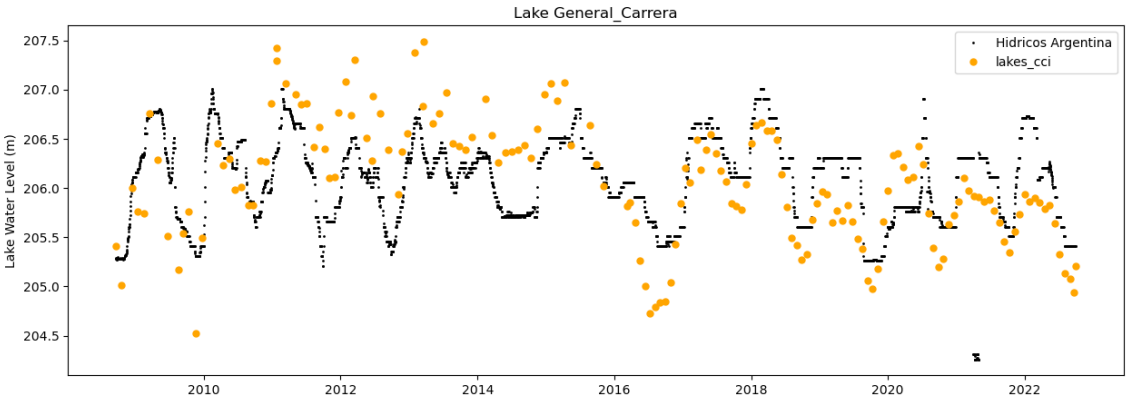


Figure 3: Lake Buenos Aires (General Carrea) - Comparison to Hidricos Argentina in situ dataset

Concerning the high value of the URMS for lake San Martin, it is probably due to some issues in the in−situ station during the period 2018-2019 as shown in Figure 4.

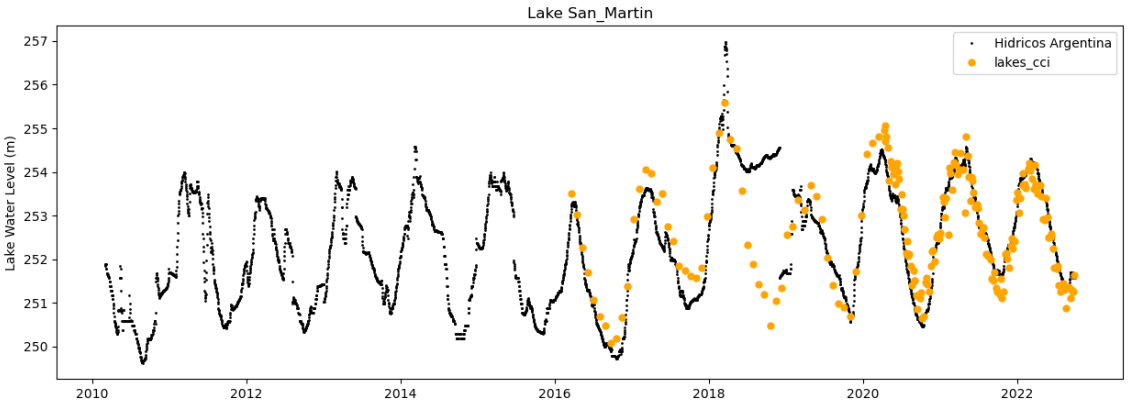


Figure 4. Lake San Martin - Comparison to Hidricos Argentina in situ dataset

2.4.3 Validation at USGS sites

The United States Geological Survey (USGS) provides information on selected water resources across the US. The USGS investigates the occurrence, quantity, quality, distribution, and movement of surface and underground waters and disseminates these to the public, state and local governments, public and private utilities, and other Federal agencies involved with managing water resources. Information concerning monthly lake water level evolution was obtained for four lakes. The Pearson coefficient values and URMS are indicated in Table 5. The correlation between time series for all lakes is very high with Pearson coefficients close to 1. Figure 5 shows the comparison between the unbiased variation of the water level and the observed difference for Lake Woods (Des Bois). The marked rise in the lake water level in 2015 and 2022 has been well measured by altimetry and the distribution of the difference is well centred in zero. Nevertheless, the error increases when the water level is low. In this case, some measures may be contaminated on shores of the lake.

Table 5: USGS LWL comparison

Lake Name	Time period	URMS(cm)	Pearson	Number of measurements
Des Bois (Woods)	2007/10 - 2022/12	11.66	0.99	646
Great_Salt	2008/10 - 2019/10	16.16	0.99	79
Michigan	2007/10 - 2022/12	6.8	0.98	1252
Walker	2013/04 - 2022/12	22.75	0.97	103

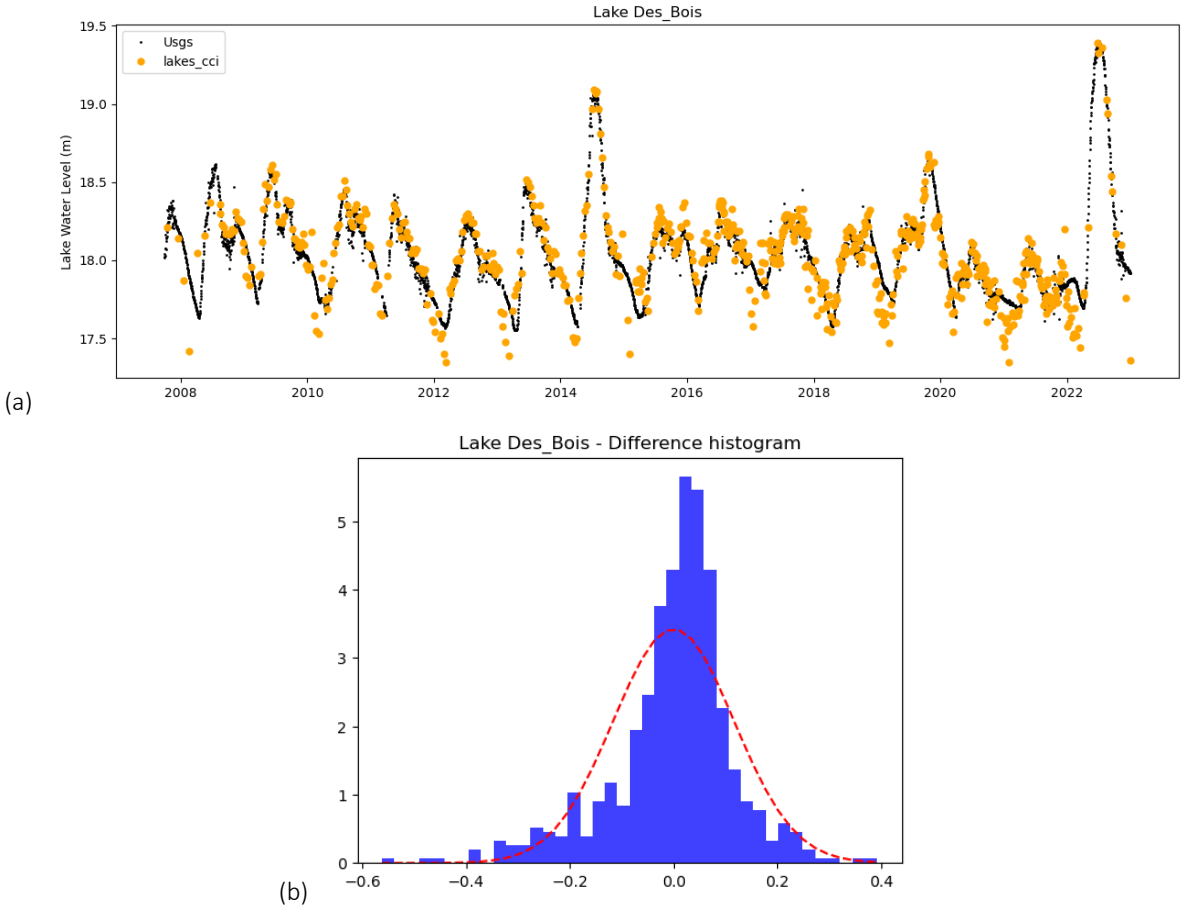


Figure 5: Lake Woods (Des Bois) - Comparison to USGS in situ dataset (a) time series and (b) Histogram of the difference

2.4.4 Validation at Canadian Water office

The Water Office of Canada[2] obtains information on water levels and floods for multiple Canadian lakes and rivers at several time resolutions. Multiple in situ stations provide information on a single lake and the mean daily value is used for the comparison to Lakes_cci data. Nevertheless, as described above, in some cases there may be anomalies in the in situ data. The case of Lake Athabasca is a good example, where data from 3 in situ stations are available. As shown in Figure 6, the data from the Fort Chipewyan

[2] https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html

station changed reference in 2014. Nevertheless, the comparison of the unbiased water level with Lakes_cci data before and after 2014, estimated separately, is very good as shown in Figure 7. The timeseries from Water Office of Canada are unbiased to be compared to Lakes_cci timeseries.

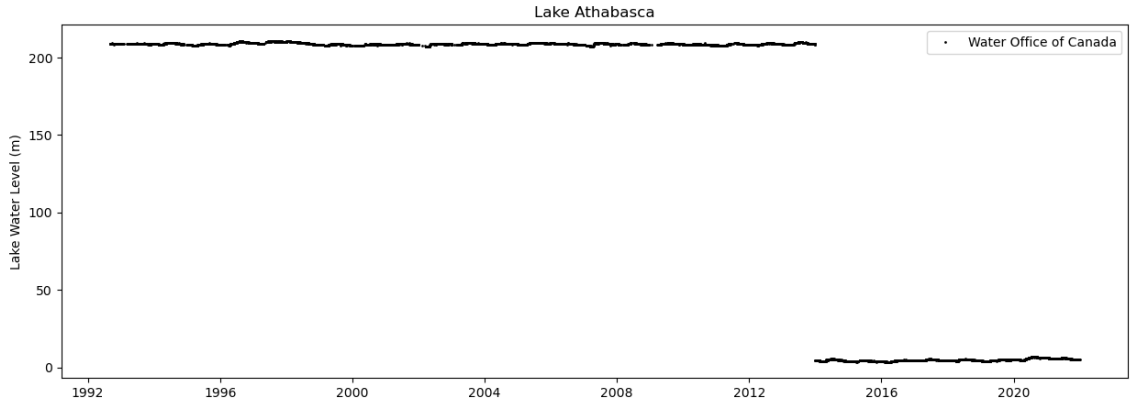


Figure 6: Lake Athabasca. In situ data from Water Office of Canada at Fort Chipewyan

Figure 7: Lake Athabasca. Comparison to Lakes_cci dataset (a) before 2014 (b) after 2014

The metrics used with the previous in situ datasets: URMSE and Pearson coefficient were evaluated. The Pearson coefficient (Figure 8) show a high correlation (Pearson coefficient > 0.8 for most of the lakes).

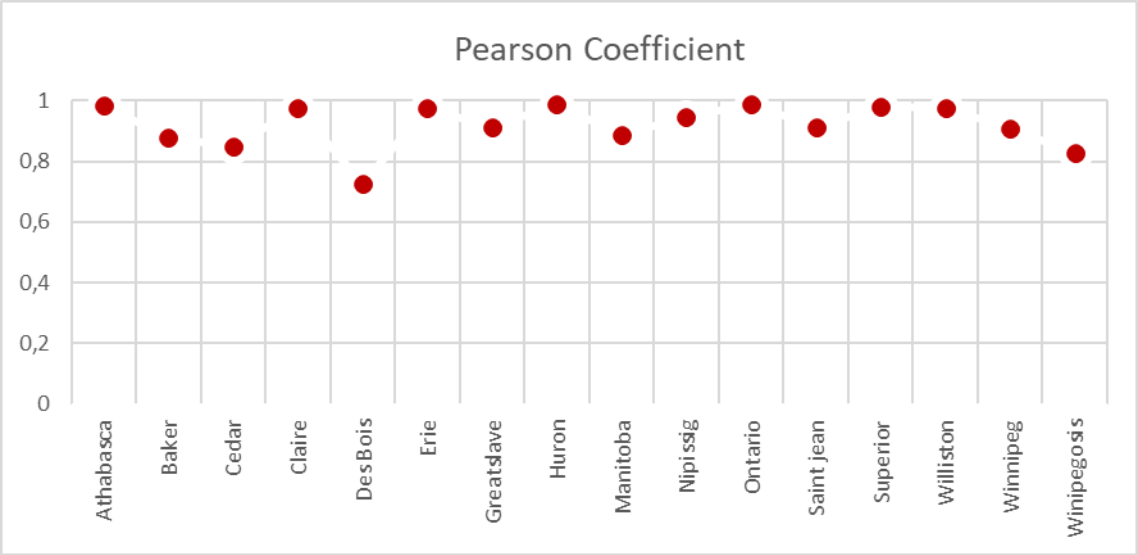


Figure 8: Pearson coefficient Lakes_cci- Water Office Canada.

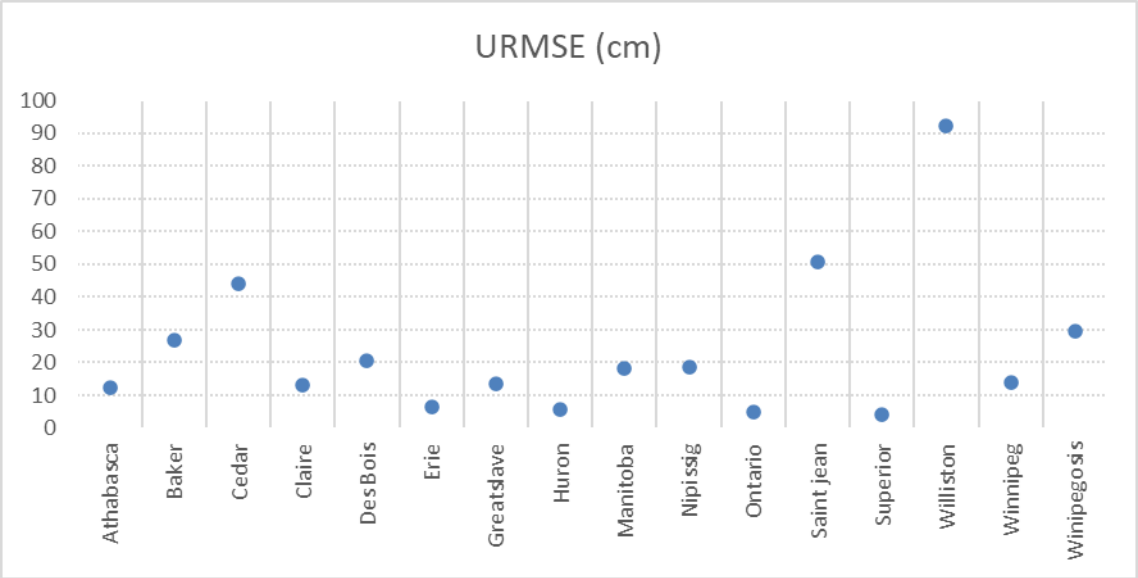


Figure 9: URMSE Value. Lakes_cci – Water Office of Canada

Figure 9 shows the URMSE values for the lakes analysed.

Concerning Lake Des Bois (Woods), the lake with the lowest Pearson coefficient, the first part of the series shows a poorer result of the earlier missions (Figure 10) and there is between the first and second part of the series. If only the period after 2008 is compared, the Pearson coefficient increases to 0.8951, indicating the very good correlation with more recent missions.

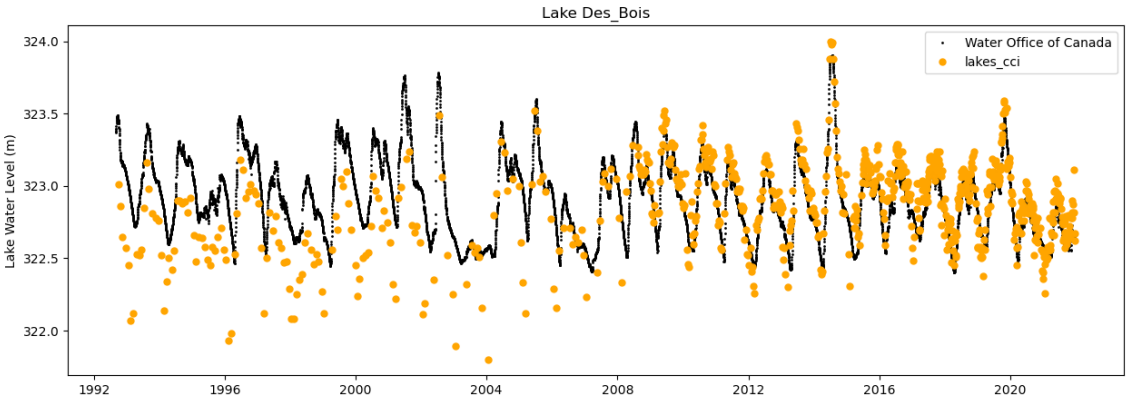


Figure 10: Lake Des Bois (Woods) - Comparison to Water Office of Canada in situ dataset

Lake Williston exhibits the highest URMSE value. This is due, as shown in Figure 11, to some very strong outliers. Nevertheless, the distribution of the differences between Lakes_cci and Water Office of Canada time series (Figure 12) is centred on zero. One of the characteristics of this lake (and the lakes with higher URMSE values) is the very pronounced variation in water level.

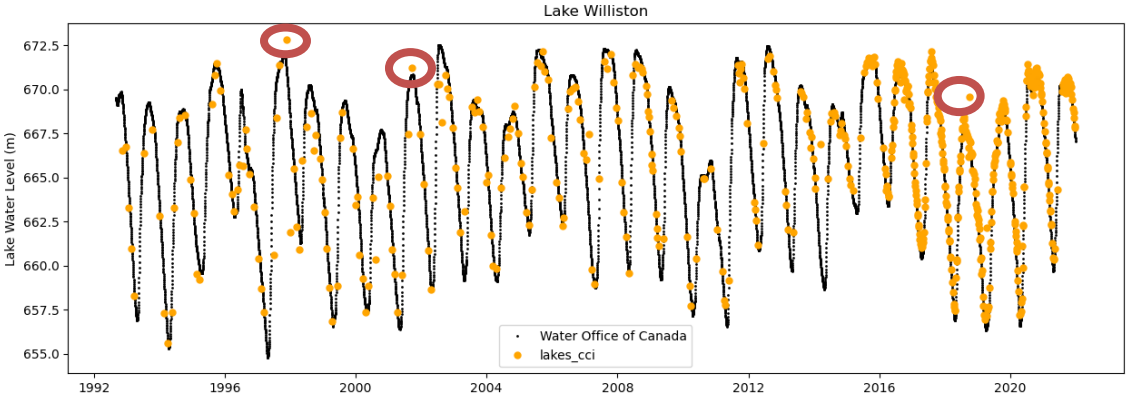


Figure 11: Lake Williston - Comparison to Water Office of Canada in situ dataset

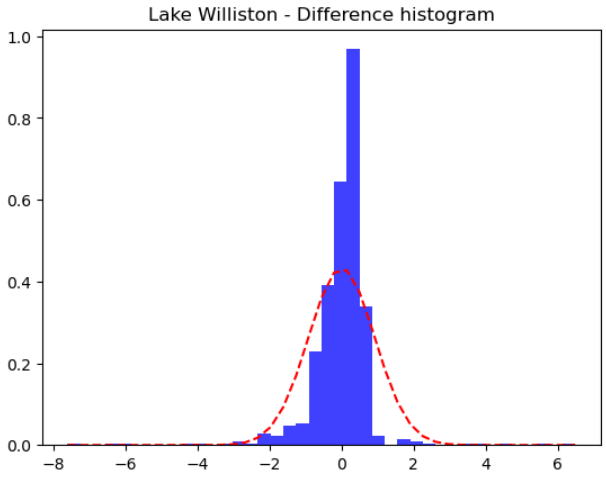


Figure 12: Lake Williston - Histogram of the difference between Lakes_cci and Water Office of Canada time series

2.4.5 Validation at Office Fédéral de l'Environnement sites

The Swiss FOEN defines environmental monitoring programs and maintains various measurement networks. It operates and coordinates several water related monitoring networks. Moreover, it monitors the level of water in Swiss rivers and lakes. A series of nearly 7 years (2016-2022) of in situ LWL over a set of mountain lakes has been acquired[3]. The results are summarized in Table 6. It shows that using only sentinel-3A and sentinel-3B in this period, the LWL product is highly accurate with URMSE < 12 cm except for the Bodensee where RMSE is nearer 20 cm.

Table 6: FOEN LWL Comparison

Lake Name	Time period	URMS(m)	Pearson
Geneva (Leman)	2016/06 - 2022/12	0.061	0.963
Neuchatel	2016/07 - 2022/12	0.116	0.797
Bodensee	2016/03 – 2022-12	0.201	0.931
Luzern	2016/07 - 2020/08	0.08	0.810
Zürich	2016/07 -2020/08	0.076	0.781

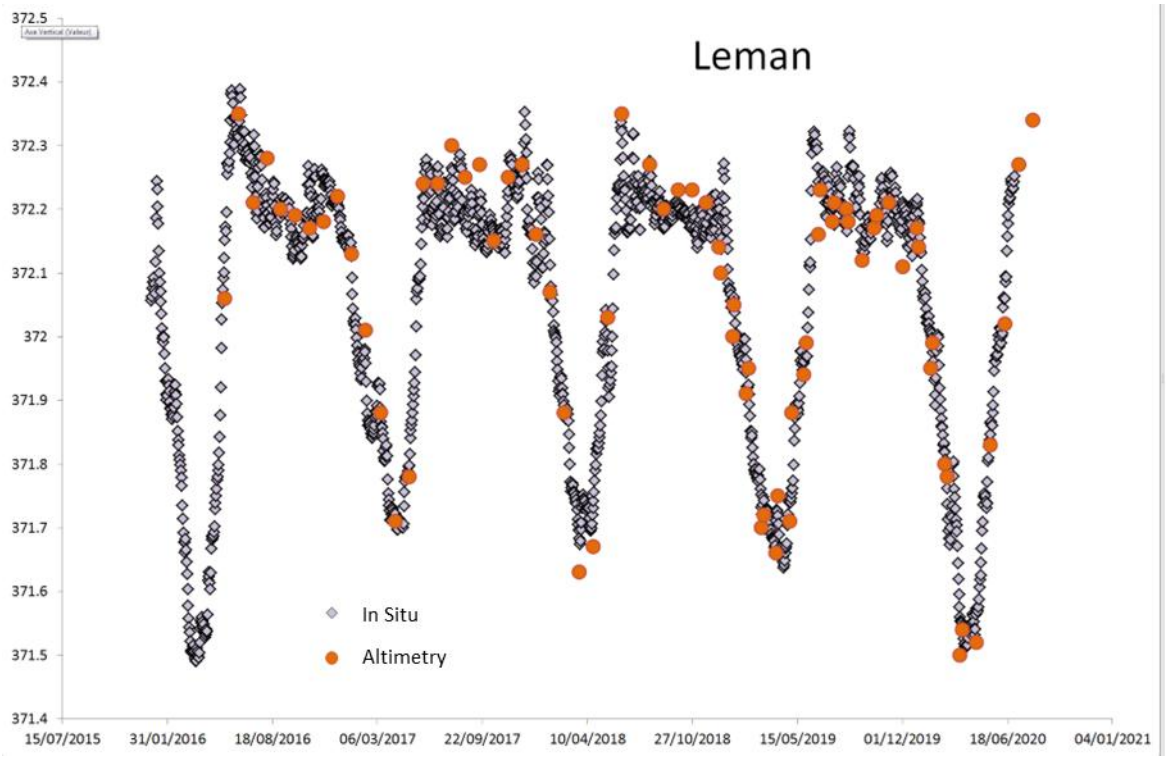


Figure 13: FOEN comparison of LWL (in m) for Lake Geneva (orange: Lakes_cci, grey: FOEN)

[3] https://www.hydrodaten.admin.ch/

2.4.6 Long-term validation at Issyk Kul

Figure 14 shows an example of LWL altimetry measurements with Sentinel 3 along two tracks (666 and 707) against in situ measurements over Issyk Kul. There is an excellent correlation between both series (99%) and a low value of URMS (Figure 15).

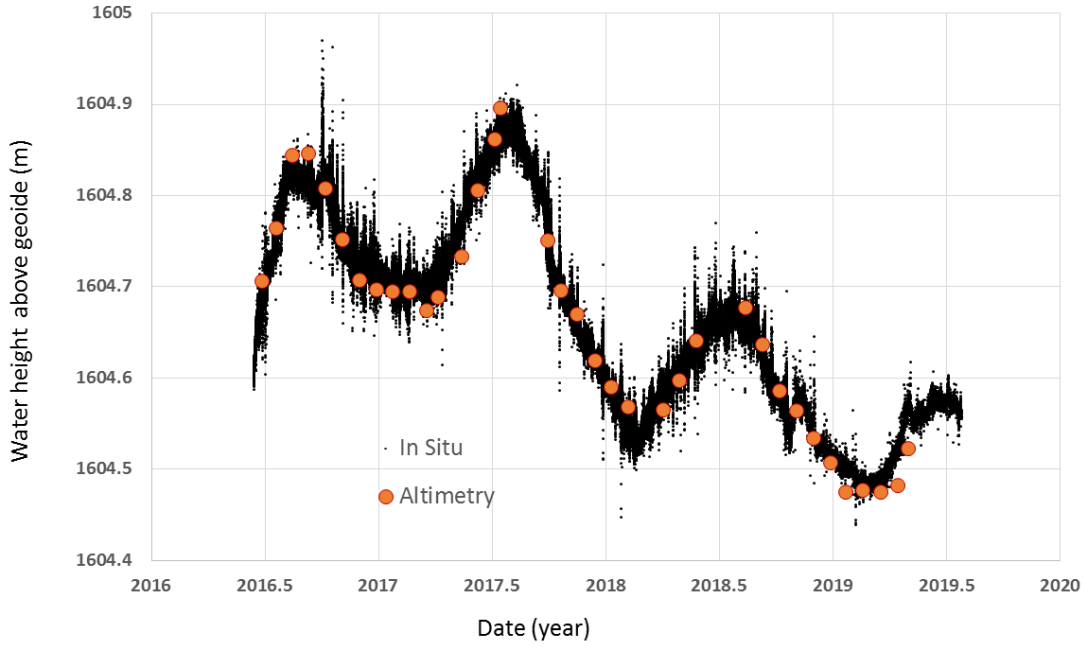


Figure 14: Sentinel 3 vs in situ measurements for Lake Issyk Kul

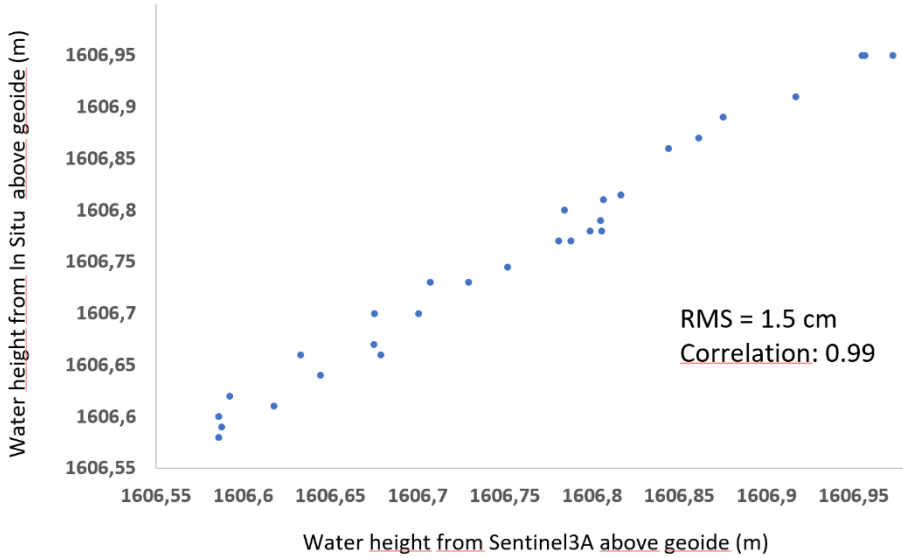


Figure 15: Correlation between Sentinel 3 vs in situ measurements for Issyk Kul

2.5 Conclusions and recommendations

We have validated the Lake Water Level product by comparing interpolated LWL time series to multiple external datasets of in situ measurements. Globally, there is a good correlation with external datasets from different organisations around the world.

Validation of the Lake water level highlights that, given that LWL is a multi-mission product, the quality of the results may differ over time. In this context a reprocessing of past missions can be very beneficial and is recommended in future generation of datasets.

The different comparisons will help us to determine where improvements in the data processing are still needed: better identification of outliers, re-analysis of past missions like Topex / Poseidon, ERS2 or Envisat, and better calculation of some inter-satellite biases.

It should be noted that, although in situ data are generally considered as truthful reference, severe limitations may be present due to a number of reasons:

- Some time periods are not covered at all with in situ data.
- Some observation error in the data collection should be expected.
- In situ measurements are always done using ground instruments that are located on the shoreline, while altimetry height provided in the CCI products, result from averaging water level measurements along the tracks of satellites, and are not collocated to the in situ gauges. It may generate also some additional discrepancies.
- The in situ measurement are in some cases provided by several instruments, that provide different values of LWL.
- Seiches effects on large lakes also cause an increase in the difference between water height measured by in situ gauges and the water height measured and averaged along the satellite track.

The decadal stability of the water height measured using satellite altimetry depends on several factors. For some lakes like Lake Des Bois or Lake Williston (Figure 10, Figure 11) the difference with in situ data is not constant over time with some trends during the 30 years of data. For other lakes like Lake Issykkul or Lake Michigan this remains constant. We still must investigate what are the reasons for this.

It has been demonstrated in many papers (over lakes, rivers, coastal areas, deltas) that the SAR altimetry has improved the accuracy of satellite altimetry for the measurements of water height of such water bodies. The SAR mode has the major interest of better focus of the effective measurement, which no longer covers (as did the previous altimeters) a circular ground footprint of several tens of km^2, but a 300 m (along-track) slice in the ground footprint. For the tracking of small targets (rivers, lakes and narrow reservoirs) this represents a breakthrough because the measurement is less altered by surrounding reflective sources, especially when the target object is oriented perpendicular to the satellite ground track. Consequently, the results for small lakes that are observed by sentinel-3A/B and sentinel6A (so after 2016) are all better than lakes that were observed by Topex/Poseidon, Envisat, Jason-1, Jason-2 and Jason-3. For example (not shown in this document), for lakes that are observed by the Topex/Poseidon until sentinel-6A, we observed less dispersion, less outliers, better stability when Sentinel-6A (which carries a SAR altimeter) is used.

Finally, the accuracy of water height not only depends on the type of instrument, or on the size of the lake, but also on the number of measurements and their position over the lake surface. It varies from one lake to another ones, depending on the position of the lake below the orbit of the satellite and its shape. In SAR altimetry and new data processing techniques such as the Lake Physical Processing (LPP), the expected accuracy, even in the worst- case scenario (small lakes, small number of measurements,

mountain area), will be significantly improve and will remain below 10 cm of accuracy for almost all the lakes.

3 Lake Water Extent - LWE

LWE can be expressed as the outline of a water body (a vector or a grid representation) or as the total areal extent of a waterbody (a single number). It is practically very challenging to process the high spatial resolution satellite imagery required to generate maps of water presence for hundreds, if not thousands, of lakes at the accuracy and temporal frequency required for climate studies. For this reason, the strategy adopted is twofold. First, for each lake, we collect a set of HR optical satellite imagery spread out over a long period. To choose the images we use the water level time series (calculated using satellite altimetry) to determine when the lake was at low, medium and high level. A relationship (a first or second order polynomial) can then be established using a set of 10 to 15 data combinations of LWL and LWE and least square adjustment. Knowing this relationship, we can then estimate LWE from LWL with each successful altimetry.

3.1 Data description

To generate LWE, only optical satellite data are exploited. In addition, a prior for maximum lake extent is taken from the Lakes_cci static database, to limit the search extent. For methodological validation, the LWE extraction from high resolution imagery is validated against very-high resolution imagery over a set of test sites.

Landsat 5 and 8/9 images and Sentinel-2 time series were selected to extract the LWE, whilst Pleaides and Pleiades NEO have been exploited for validation purposes.

3.1.1 Landsat 5 TM TOA Reflectance

Landsat 5 TM calibrated top-of-atmosphere (TOA) reflectance were used. Calibration coefficients are extracted from the image metadata (Table 7). See Chander et al. (2009) for details on the TOA computation.

Table 7: Landsat 5 wavebands and resolution

Bands

Name	Resolution	Wavelength	Description
B1	30 meters	0.45 - 0.52 µm	Blue
B2	30 meters	0.52 - 0.60 µm	Green
B3	30 meters	0.63 - 0.69 µm	Red
B4	30 meters	0.76 - 0.90 µm	Near infrared
B5	30 meters	1.55 - 1.75 µm	Shortwave infrared 1
B6	30 meters	10.40 - 12.50 µm	Thermal Infrared 1. Resampled from 60m to 30m.
B7	30 meters	2.08 - 2.35 µm	Shortwave infrared 2

3.1.2 Landsat 8 TOA Reflectance

Calibration coefficients for Landsat 8 calibrated top-of-atmosphere (TOA) reflectance are extracted from the image metadata (Table 8). See Chander et al. (2009) for details on the TOA computation. (Earth Engine Data Catalogue).

Table 8: Landsat 8 wavebands and resolution

Bands

Name	Resolution	Wavelength	Description
B1	30 meters	0.43 - 0.45 µm	Coastal aerosol
B2	30 meters	0.45 - 0.51 µm	Blue
B3	30 meters	0.53 - 0.59 µm	Green
B4	30 meters	0.64 - 0.67 µm	Red
B5	30 meters	0.85 - 0.88 µm	Near infrared
B6	30 meters	1.57 - 1.65 µm	Shortwave infrared 1
B7	30 meters	2.11 - 2.29 µm	Shortwave infrared 2
B8	15 meters	0.52 - 0.90 µm	Band 8 Panchromatic
B9	15 meters	1.36 - 1.38 µm	Cirrus
B10	30 meters	10.60 - 11.19 µm	Thermal infrared 1, resampled from 100m to 30m
B11	30 meters	11.50 - 12.51 µm	Thermal infrared 2, resampled from 100m to 30m

3.1.3 Sentinel-2 L1C/L2A

Sentinel-2 MultiSpectral images were not systematically produced to L2 with global coverage in the first years of operation. L1C and L2 data have both been used to generate the Lakes_cci LWE product. The L1C product provides orthorectified Top-Of-Atmosphere (TOA) reflectance with sub-pixel multispectral registration. Cloud and land/water masks are included in the product. The waveband configuration and per-band resolution are provided in Table 9. More details can be obtained at ESA: Level-1C Processing for details.

Table 9: Sentinel-2 wavebands and resolution

Bands

Name	Units	Min	Max	Scale	Resolution	Wavelength	Description
B1				0.0001	60 meters	443.9nm (S2A) / 442.3nm (S2B)	Aerosols
B2				0.0001	10 meters	496.6nm (S2A) / 492.1nm (S2B)	Blue
B3				0.0001	10 meters	560nm (S2A) / 559nm (S2B)	Green
B4				0.0001	10 meters	664.5nm (S2A) / 665nm (S2B)	Red
B5				0.0001	20 meters	703.9nm (S2A) / 703.8nm (S2B)	Red Edge 1
B6				0.0001	20 meters	740.2nm (S2A) / 739.1nm (S2B)	Red Edge 2
B7				0.0001	20 meters	782.5nm (S2A) / 779.7nm (S2B)	Red Edge 3
B8				0.0001	10 meters	835.1nm (S2A) / 833nm (S2B)	NIR
B8A				0.0001	20 meters	864.8nm (S2A) / 864nm (S2B)	Red Edge 4
B9				0.0001	60 meters	945nm (S2A) / 943.2nm (S2B)	Water vapor
B11				0.0001	20 meters	1613.7nm (S2A) / 1610.4nm (S2B)	SWIR 1
B12				0.0001	20 meters	2202.4nm (S2A) / 2185.7nm (S2B)	SWIR 2

More details can be obtained at ESA: Radiometric.

L2A products provide orthorectified Bottom-Of-Atmosphere (BOA) reflectance with sub-pixel multispectral registration. A Scene Classification map (cloud, cloud shadows, vegetation, soils/deserts, water, snow) is included in the product.

For a long time, the Level 2 Sentinel-2 images is not systematically produced all over the world. By the way it is Level 1C data that have been proceed in the Lakes_cci LWE product.

3.1.4 Pléiades HR

Pléiades is an environment-focused constellation (consisting of two satellites, referred to as 1A and 1B) from CNES (France), launched on 17 December 2011 (Pléiades 1A) and 2 December 2012 (Pléiades 1B). The constellation remains operational. Sensor properties are listed in Table 10.

Table 10: Pleiades wavelengths and resolution

Name	Resolution	Waveband	Description
Pan	0.5	480-820	
B0	2	450-530	Blue
B1	2	510-590	Green
B2	2	620-700	Red
B3	2	775-915	NIR

3.1.5 Pleiades NEO

The Pléiades Neo Imager carried on board the satellites deliver very-high optical (30 cm resolution) imagery with a swath of 14 km. The wavebands configuration is provided in Table 11.

Table 11: Pleiades NEO wavelengths and resolution

Name	Resolution	Waveband	Description
Pan	0.3	450-800	
1	1.2	400-450	Deep Blue
2	1.2	450-520	Blue
3	1.2	530-590	Green
4	1.2	620-690	Red
5	1.2	700-750	Red edge
6	1.2	770-880	NIR

3.1.6 Reference classification data

3.1.6.1 Global Surface Water database

The European Commission Joint Research Centre developed the Global Surface Water Database in the framework of the Copernicus Programme. It maps the location and temporal distribution of water surfaces at the global scale over the past 32 years and provides statistics on the extent and change of those water surfaces. The dataset produced from Landsat imagery (courtesy USGS and NASA) will support applications including water resource management, climate modelling, biodiversity conservation and food security.

3.1.6.2 Active Learning for Cloud Detection (ALCD)

To validate water masks for large scenes in different seasons, a reference dataset has been developed using Active Learning for Cloud Detection (ALCD) software exploiting Sentinel-2 images. The motivation behind the use of ALCD for validating results rather than a hydrological database comes from the fact that some water bodies are highly dynamic, depending on season, meteorological conditions, management policies, etc. Fourteen test sites, or more precisely, 'tile scale sites', were identified in France (Pena Luque et la., 2021) as shown in Figure 16. This ALCD reference database has been exploited to set up and to tune the parametrization of the Extract EO chain.

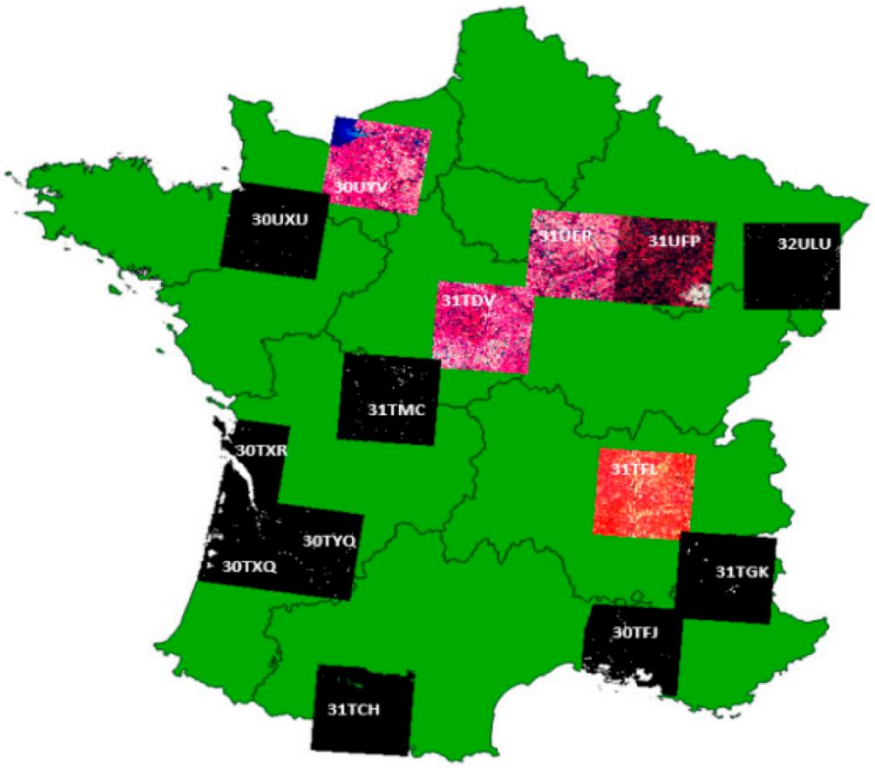


Figure 16: ALCD sites exploited to validate algorithmic approaches and parameters settings (Pena Luque et al., 2021)

The reference water masks are generated using an iterative active learning procedure which enables the creation of an accurate reference mask in less than 3–5 h for a 110 × 110 km scene. For this, it is necessary to manually create reference points on the image, train the model (based on Random Forest of OTB and predict with ALCD, then add new reference points for the most problematic areas, repeat training/predictions cycles as many times as necessary (usually 5–7 iterations). The final step consists of a manual correction of persisting errors. For the training process, the following Sentinel-2 bands were used, B2: Blue, B3: Green, B4: Red, B8: NIR, B11: SWIR1, B12: SWIR2. In addition, derived indices such as MNDWI and slope information derived from SRTM were also exploited.

The generated validation dataset is open access (Pena Luque 2019). The 14 selected sites cover different eco-climatic zones (oceanic, Mediterranean, mountainous, continental). Then, 26 scenes have been completely labelled on these sites on different seasons and conditions (snow, flood event, turbid waters, wetlands, urban areas, dry scenes).

3.1.6.3 Lake contours database

The initial search for water within the satellite scene is geographically bounded to decrease processing time. Within these bounds (and a 1-kilometre buffer), the algorithm determines the shoreline as precisely as possible. This is not a simple limit as, for many areas, water level dynamics cause surface expansion into bordering wetlands (Figure 17).

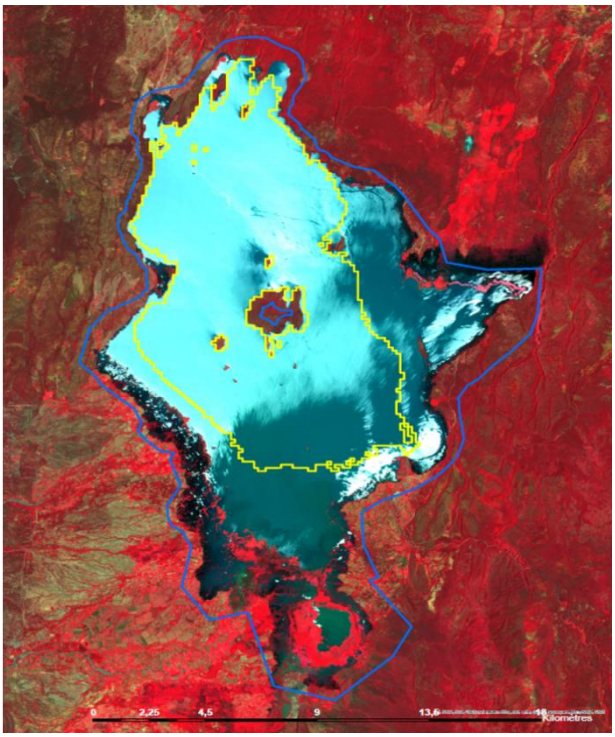


Figure 17: Example of Lake AoI correction for Lake Barringo (Kenya) with in yellow the Lakes_cci AoI and in blue the corrected ones. large part of the main open water surface was missing on the Lakes_cci AOI as well as the bordering wetlands.

For medium and lower resolution imagery a relative rough definition of the Area of interest (AoI) is sufficient, and the buffer application should correct some variations. For exploitation of (very) high resolution imagery, such as Sentinel-2, a more precise definition of the AoI is required. This is done against the Lakes_cci maximum water extent mask (from the Land Cover CCI), or alternatively the Hydrolakes database[4] derived from the SRTM mission which contains 1,4 million lake polygons larger than 10 ha. These initial bounds were validated based on Sentinel-2 imagery acquired at different hydrological period and corrected as needed (Figure 18).

[4] https://www.hydrosheds.org/pages/hydrolakes

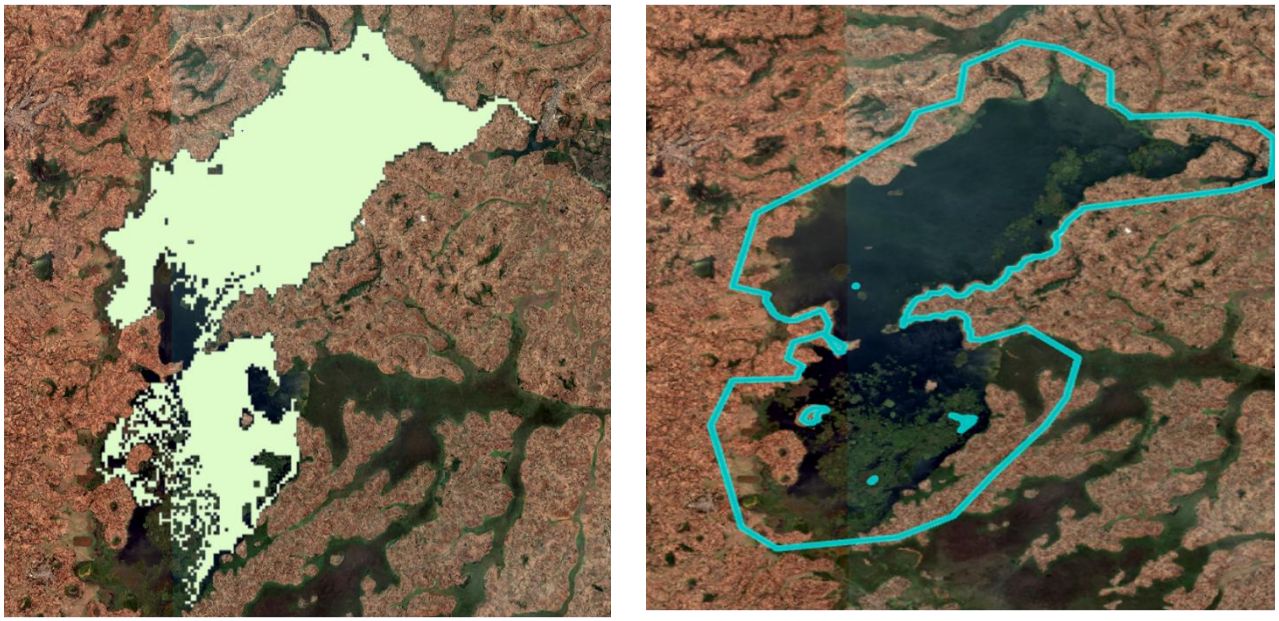


Figure 18: AoI correction: case of the Lake Finchaa (Ethiopia). Part of the bordering and recurrently flooded wetlands were missing in the Lakes_cci AoI, the limits have been corrected. In this case the precise lake limit is difficult to define, and thresholds are set subjectively.

3.2 Comparison methods

The validation of the Lake Surface Water Extent consists of the comparison with independent database, LWE derived from higher resolute satellite imagery. A part of the validation on intercomparison of results obtained from different levels of preprocessing or based on different parametrisations.

3.3 Description of the work

3.3.1 Data set selection

Cross-validation of Sentinel-2 MSI images processed to L1C and L2A was done considering small water bodies, ponds, and fish farms observable within two tiles over the Champagne region (France) for which ADLC reference tiles were also accessible (Figure 19).

Images acquired in summer and winter with under normal and flooding hydrological conditions were selected and processed. The derived LWE from L1C and L2A images were then compared with each other and with those from the ALCD reference tiles.

Metrics exploited for comparison are the classical F Score, Precision, Recall and MCC (Mathews Correlation Coefficient) a less exploited metric despite the fact that MCC produces a more informative and truthful score in evaluating binary classifications.

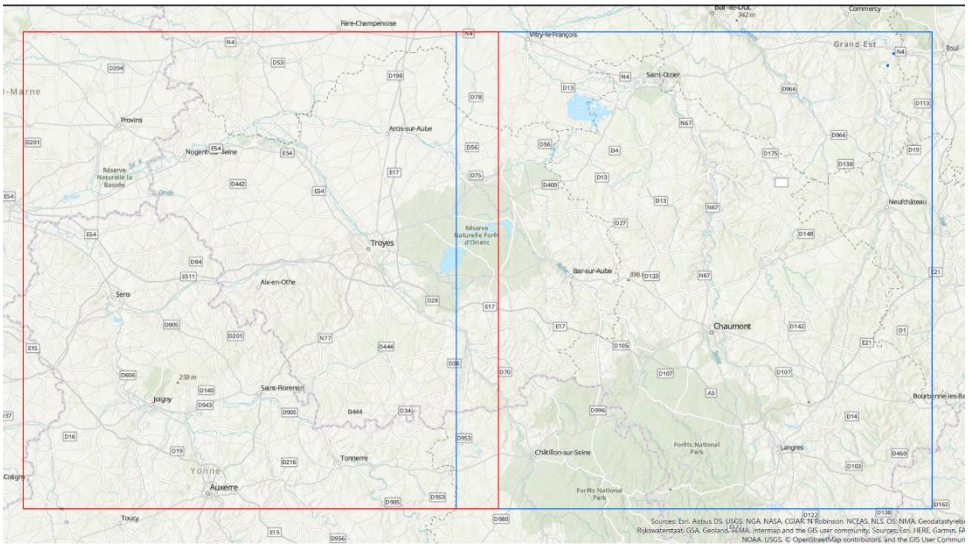


Figure 19: Lakes in Champagne (France), covered by two Sentinel-2 tiles Tile T31UEP /UFP

https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c-processing

3.3.2 Tuning of ExtractEO parameters

Water surfaces are extracted from the imagery using an in-house processing chain named ExtractEO (Maxant et al, 2022).

ExtractEO preprocessing steps correspond to:

- Area of Interest (or Region of Interest ROI) is defined including the target lake and must be as precise as possible (see above).

Selection of set of images representative of the different water levels of the target, i.e. based on altimetric water level curves from Hydroweb.

The processing then follows the following steps:

- Automatic water sample generation from Global Surface Water. Water indices are computed to remove outliers and filter the training samples to the hydrological reality of the image (water extent, resolution)
- Training using the Multi-Layer Perceptron classifier
- Slope and hillshade thresholds derived from HR DEMs are applied to refine the water extraction (post-processing)
- Minimum mapping unit (MMU) sieving to remove small features (0,1 hectares in this case)
- Water extent (in km^2) is subsequently calculated using the sum of individual pixel classified as water pixel within the ROI
- Generation of a maximum water extent mask

3.3.3 Validation of water surfaces extraction based on VHR optical imagery

Lake-by-lake water extent validation is not feasible at the scale of the Lakes_cci. However, it is possible, necessary and important to validate the procedure which generates lake area and lake extent vectors from Sentinel-2 MSI and Landsat images.

To validate the ExtractEO processing chain, cross-validation of coincident high resolution (HR) and very high resolution (VHR) optical measurements of water extent were carried out over selected lakes representing different lake environments and morphological complexity.

The result of this validation is a pixel-wise accuracy estimate, as well as overall aggregated accuracy estimates. Thus, by combining pairs of HR and VHR observations we can make an overall comparison of how good the process to generate water extent from HR optical imagery. The VHR data are either Pleiades HR or Pleiades NEO data with spatial resolution of 70 and 30 cm, respectively. This dataset (Table 12) was obtained thanks to CNES facilities within the SWOT Cal/Val programme and through Pleaides NEO promotional use in the CNES Dinamis programme. Results of the investigation are detailed for each of the regions in the following sections.

Table 12: Pairs of VHR and HR data available for the validation of the processing chain

Site	Province Country	Size in km^2	VHR dataset	Dates of acquisition VHR	Dates of acquisition HR
Lake Fitri	Chad	200-1250	Pleiades HR: 0,7m	2021	
Fern Ridge Reservoir	Oregon USA	36	Pleiades HR: 0,7m	25/07/2022	26/07/2022
				29/07/2022	29/07/2022
			SPOT 6-7: 1,5 m	11/08/2022	13/08/2022
Lac de Der	Champagne France	10-40	Pleiades HR: 0,70m	30/12/2019	31/12/2019
				06/01/2020	06/01/2020
Lac d'Orient	Champagne France	8-23	Pleiades NEO: 0,30m	15/12/2022	13/12/2022
					16/12/2022
Lac d'Auzon-Temple	Champagne France	7-20	Pleiades NEO: 0,30m	15/12/2022	13/12/2022
					16/12/2022
Lac d'Amance	Champagne France	0-18	Pleiades NEO: 0,30m	15/12/2022	13/12/2022
					16/12/2022
Lake Shenjing	Anhui Prov. PR China	33	Pleiades NEO: 0,30m	10/12/2021	07/12/2021

3.4 Validation results

3.4.1 Data set selection and parameter tuning

A set of three images over two Sentinel-2 MSI tiles were selected to represent different hydrological conditions (dry, wet and flooding) and associated variability in lake and reservoir volume. Based on LWE extracted compared with the ALCD reference, metrics were produced lake-by-lake, and at tile level (Table 13).allowing a ranking of the performance of the Level1C and 2A as well as some tuning parameters.

Table 13: Metrics derived from the preprocessing of Level 1C and 1A data set.

		UFP									
Date	Product	parameters	resolution	pekel	NDWI thr.	hillshade thr.	sieve thr.	F1 score	Precision	Recall	MCC
2019.07.17	L2A	best	10	20	0.1	45	20	0.922	0.978	0.872	0.923
		common	10	20	[-0.1 ; 0.3]	45	20	0.921	0.984	0.865	0.922
	L1C	best	10	20	0.1	45	20	0.933	0.960	0.908	0.933
		common	10	20	[0.0 ; 0.35]	45	20	0.915	0.984	0.854	0.916
2019.12.04	L2A	best	10	20	0.25	45	20	0.941	0.929	0.953	0.941
		common	10	20	[-0.1 ; 0.3]	45	20	0.927	0.886	0.972	0.928
	L1C	best	10	20	[0.0 ; 0.35]	45	20	0.937	0.914	0.962	0.937
		common									
2019.12.29	L2A	best	10	20	0	45	10	0.906	0.930	0.883	0.905
		common	10	20	[-0.1 ; 0.3]	45	20	0.893	0.977	0.822	0.896
	L1C	best	20	20	[0.0 ; 0.35]	45	10	0.877	0.890	0.863	0.875
		common	10	20	[0.0 ; 0.35]	45	20	0.869	0.848	0.891	0.868

Comparing the preprocessing level, our results (Figure 20) highlight the following that for images acquired in July, similar results for LWE extracted either from LC1 and L2A data. Therefore, when the acquisitions conditions (azimuth, sun position etc.) are less favourable such as in December.

LC1 presents higher level of commission on shadows areas as well as over mud banks and/or muddy shallows than L2A. It is therefore highly recommended to use level 2 data. In terms of remote sensing image processing, this may seem obvious. However, it should be remembered that for many years, Level 2 data was only available on a limited scale.

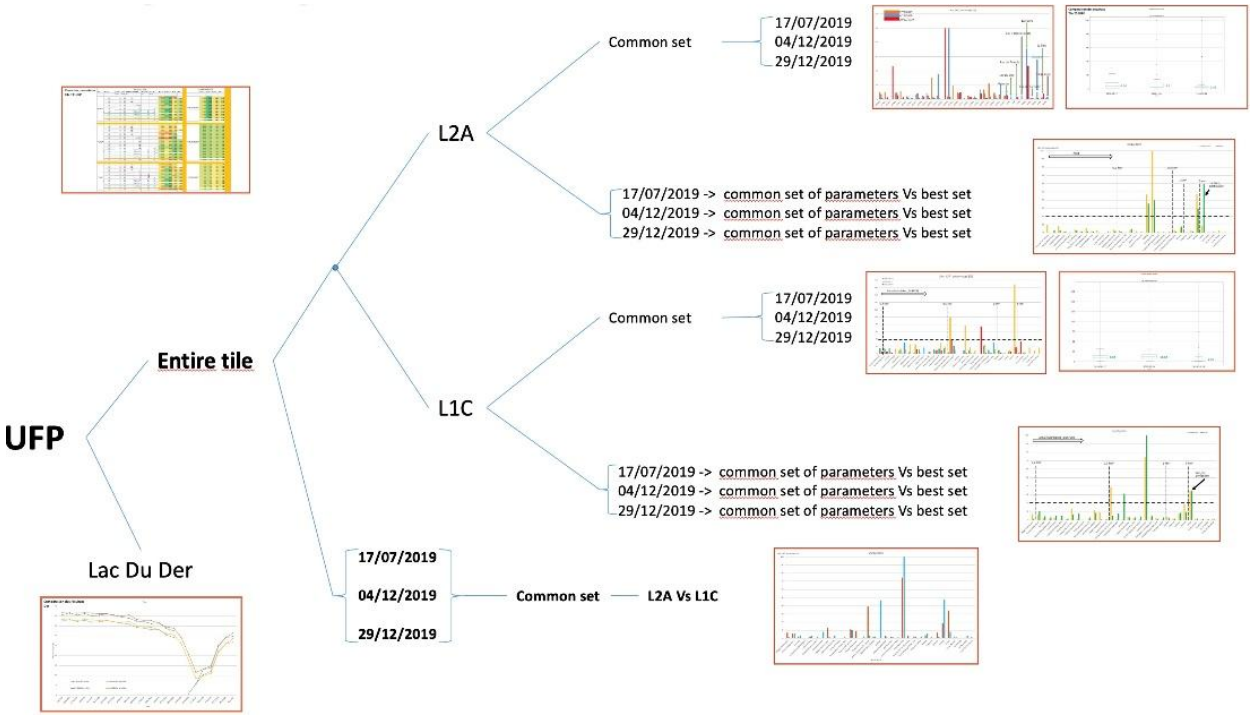


Figure 20: Scheme of testing and validating the exploited image processing level (L1C vs L2A), and ExtractEO parameterisation.

In addition to the levels of images, few parameters that can be tuned were analysed, i.e. the threshold values on NDWI indices for the training samples selected via Pekel's GSW database (to insure it is really open water and not a drying area) as well as the hillshade and sieve threshold values. Then the threshold values were identified, the final comparison with ALCD reference gives relatively small divergence (Figure 21).

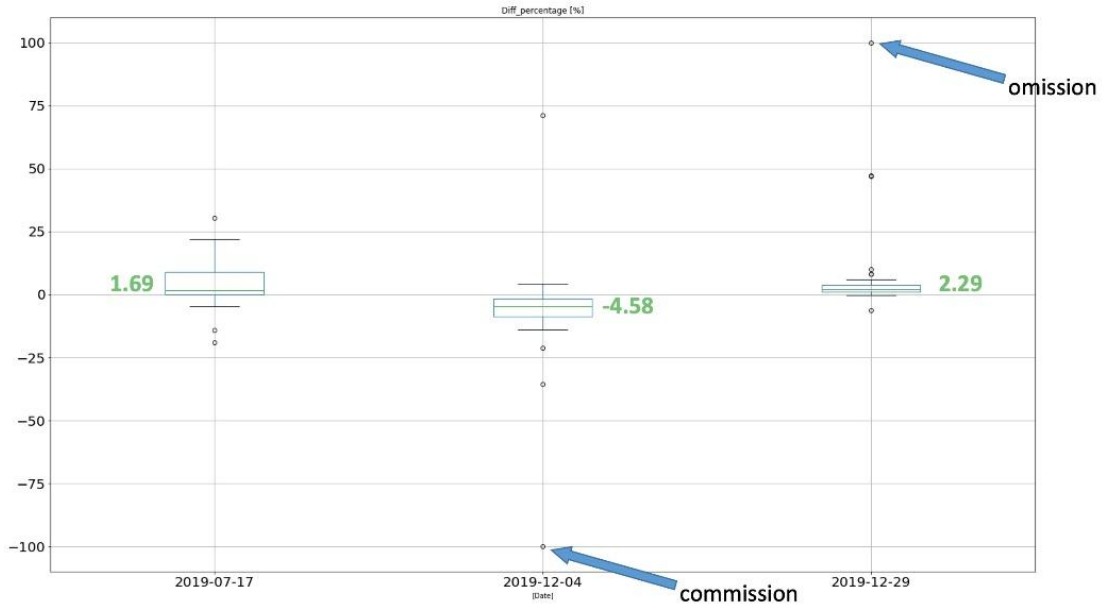


Figure 21: Difference in percentage between LWE derived from ExtractEO and the ALCD reference for the Sentinel-2 tile T31 UFP.

When analysing water body by water body and working with absolute differences, the values obtained can become very large when reaching 100 of omission for example (Figure 22). However, we need to put these results into perspective because in many cases the values relate to very small entities, just a few pixels in size; lot of these water bodies are unnamed being just identified with an ID in a database (Figure 23)

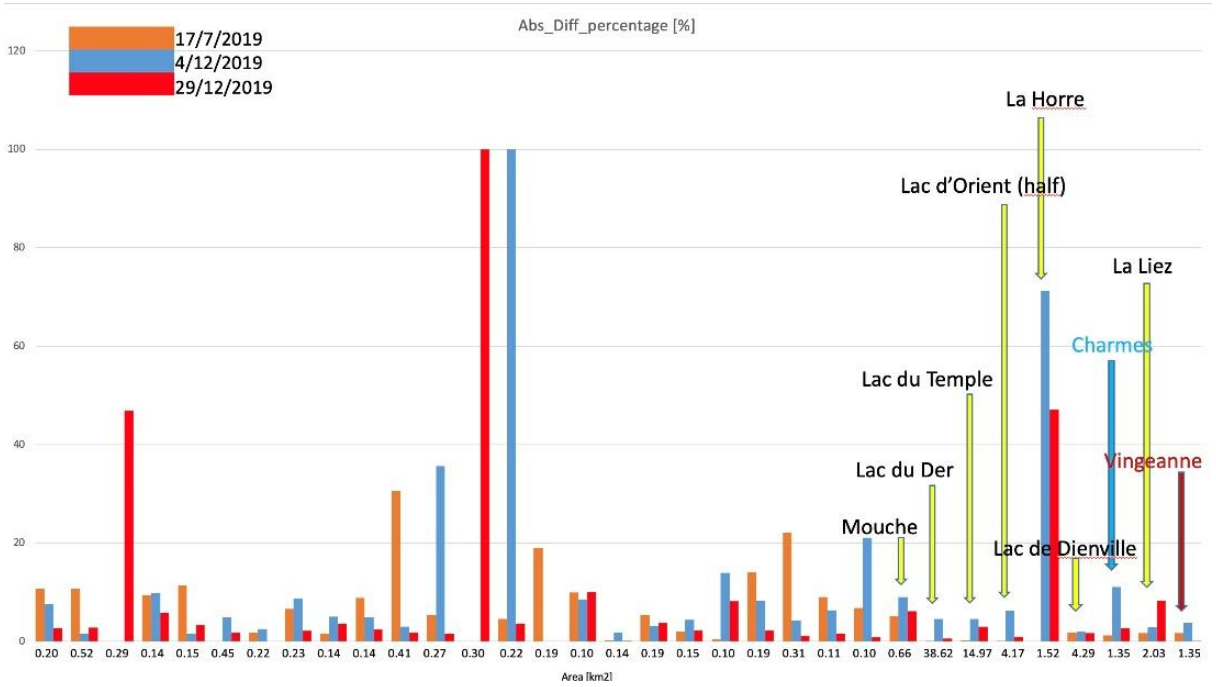


Figure 22: Absolute difference expressed in percentage between LWE derived from ExtractEO and the ALCD reference ; notice that in abscise the associated involved surfaces in km2 are indicated.

The very high value absolute difference, in the order of 90% or more, correspond to dry lakes where only a small water fraction remains. This fraction can be partly recognized and extracted by ExtractEO, whereas it was not noted in the ALCD reference (Figure 23).

Figure 23: Example of a water body presenting a large difference with the ALCD reference.

3.4.2 Comparison of Landsat versus Sentinel-2 over Lake Chad

The Chad case study presents an opportunity to compare results acquired from Sentinel-2 and Landsat 8 acquired on the same day, albeit with different spatial resolution (10 and 30 m, respectively). It further presents an opportunity to compare the satellite-derived lake extent with field tracks recorded from a boat.

Lake Chad is a historically large, shallow, endorheic lake in Africa, which has varied in size over the centuries. According to the Global Resource Information Database of the United Nations Environment Programme, it shrank by as much as 95% from around 1963 to 1998.

Th field survey was carried out in the archipelago of Lake Chad, more precisely within the Bol-Reria area; a disconnected part of interdune water bodies.

3.4.2.1 Context of Lake Chad

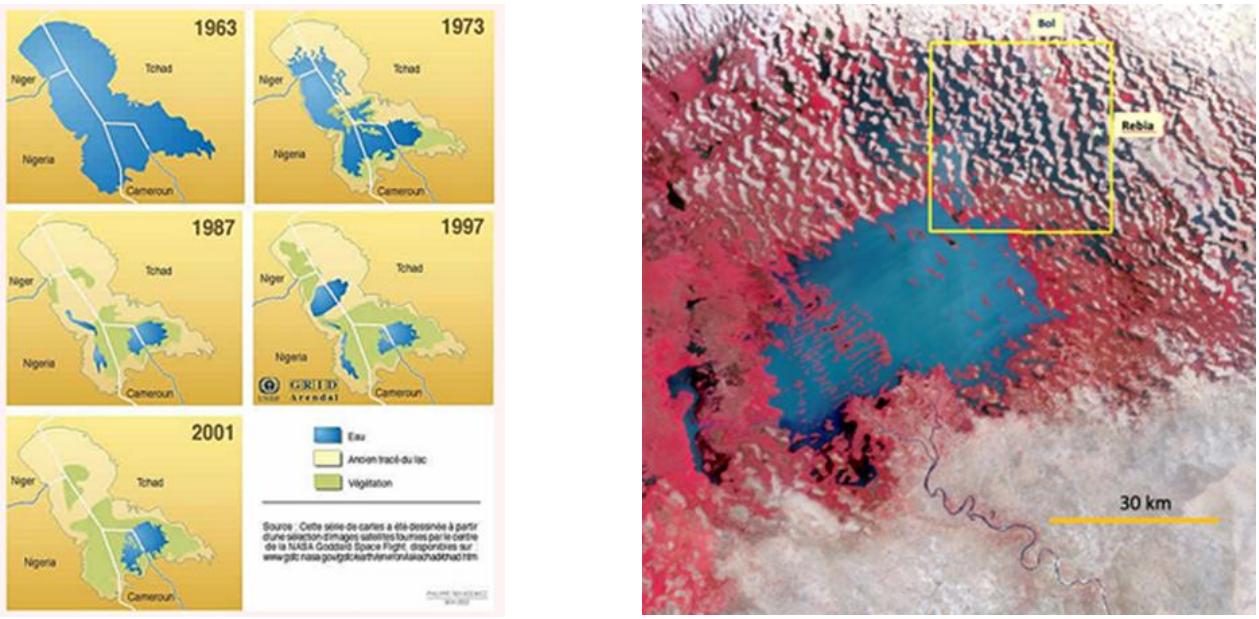


Figure 24: Lake Chad evolution over 40years (1963-2001) and location of the Bol- Rebia area (yellow rectangle) within the Archipelago region of Lake Chad

In the Archipelago area (Figure 25), the islands and peninsulas are summits of a sand dune network flooded by Lake Chad water. The dune sonnet area is part of the island where there is a small group of huts and a few millet fields when the island is inhabited. The vegetation corresponds to trees and shrubs. On the shores of the islands, fringe of macrophytes can be observed. One characteristic of this part of Lake Chad is the presence of reed islands that are anchored in shallow waters. Part of them separate from the main islands and form circular floating islands of vegetation (*kirtas*), formed from by *Papyrus* and *Phragmites*. Their sizes vary from a few meters to several hundred meters. At the time of reversal of dominants winds, during June and October generally in the rainy season when the winds shifts are frequent, these islands move back and forth, modifying the aspect of the reeds islands and closing the channels of the Archipelago (Leonard, 1974; Ittis and Lemoalle, 1983).

Land tongues with sparse vegetation

Floating and anchored vegetation

Bare sandy tongues

Figure 25: Landscapes of the Archipelago area.

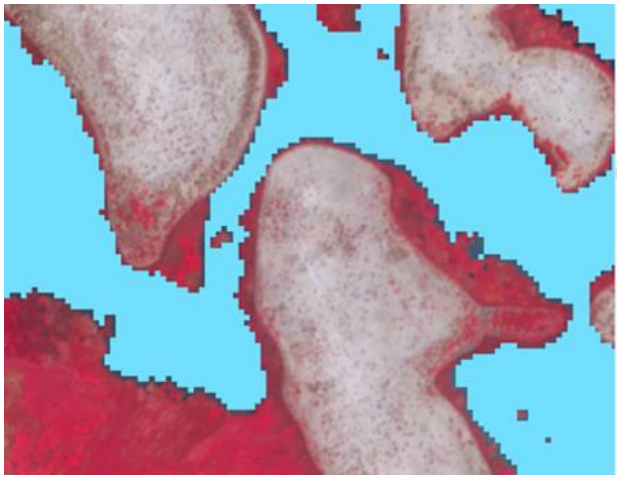
3.4.2.2 Exploited data

Landsat and Sentinel-2 present similarity but also differences and not only in terms of spatial resolution. Whereas SWIR1 and SWIR2 bands have similar properties, NIR bands are noticeably different between Sentinel-2 band B8, and Landsat-8 band 5. The narrow band B8A of Sentinel-2 is similar to band 4 of Landsat 8, with respectively a 20 and 30 meters spatial resolution (Table 14).

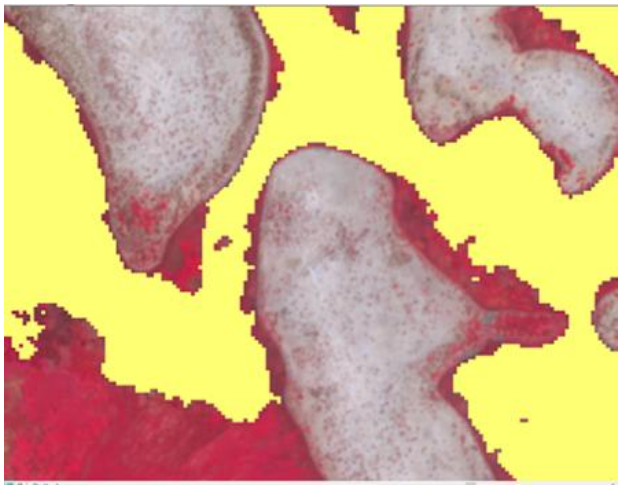
Table 14: Comparison of the Sentinel-2 and Landsat 8 bands.

	Sentinel-2		Landsat 8
NIR	0,779-0,885	B8 10m	0.85-0.88 um
NIR	0,853-0,875	B82 20m	0.85-0.88
SWIR1	1,568 – 1,659		1.57-1.65
SWIR 2	2,114- 2,289		2.11-2.29

The comparison of LWE extraction from Sentinel-2 and Landsat 8 shows that extraction based on Landsat 8 is more coherent/continuous, and less algorithm-dependent compared to Sentinel-2. LWE derived from Sentinel-2 shows that the limits of water/non water boundaries are more accurate, but some overestimation is observed in vegetation areas. On the LWE obtained from Landsat 8 image, an underestimation of water on the shore of the lake is evident due to the lower resolution (Figure 26).



LWE derived from Landsat-8 extraction on the shoreline (background Sentinel-2)



LWE derived from Sentinel-2 extraction with Sentinel-2 image (background)

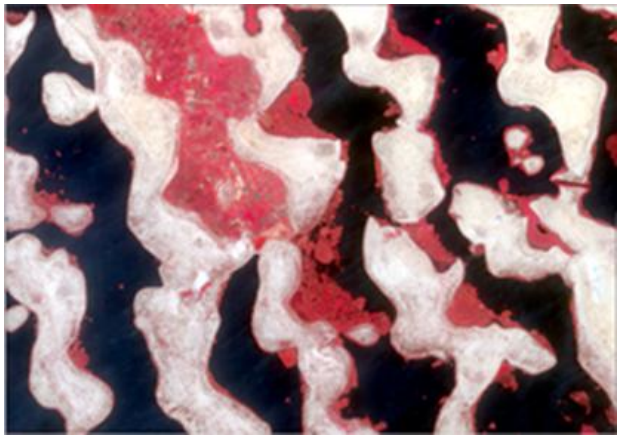
Figure 26: Comparison of Sentinel-2 and Landsat 8 LWE: influence of the spatial resolution.

A multitemporal analysis was carried out over a period of one month using five Sentinel-2 images obtained on 9, 14, 24 and 29th April and 4th May 2019. The water surface was determined in each image. Each result was then compared to the mean surface over the observed period. This showed < 2% difference between observations taken over the period of one month (Table 15).

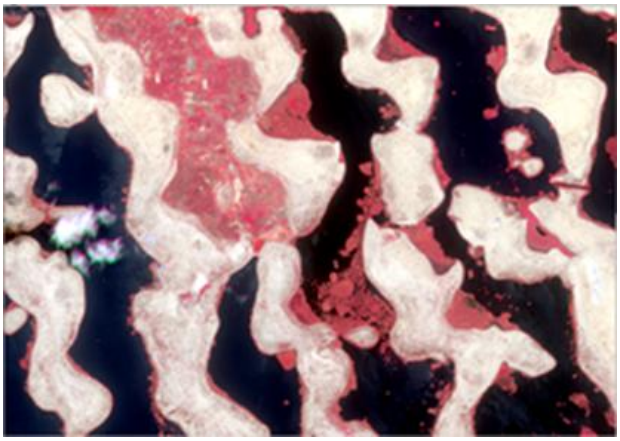
Table 15: LWE derived from Sentinel-2

Date	LWE	Difference / Mean
09/04/2019	378.7875 km²	0,7%
14/04/2019	371.5667 km²	1,4%
24/04/2019	372.1424 km²	-1,03%
29/04/2019	379.2515 km²	0,85%
04/05/2019	381.6009 km²	1,48%

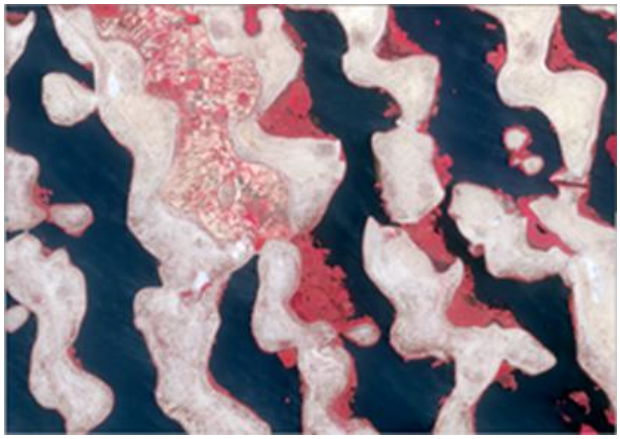
In this sequence, the apparent LWE dynamics are related to the movement of the *kirtas*, i.e. floating vegetation, under wind direction/speed dependence rather than LWL changes (Figure 27, Figure 28). Small variations further stem from small clouds on one date.



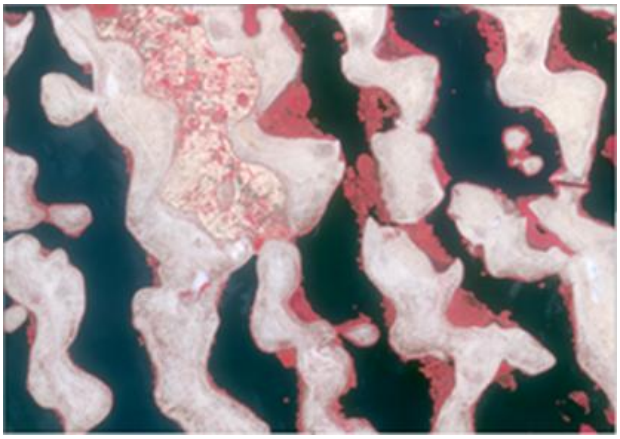
Sentinel-2 acquired on 9 April 2019 (NIR, R, G)



Sentinel-2 acquired on 14 April 2019



29 April 2019



4 May 2019

Figure 27: Displacement over one month of the k*irtas*.

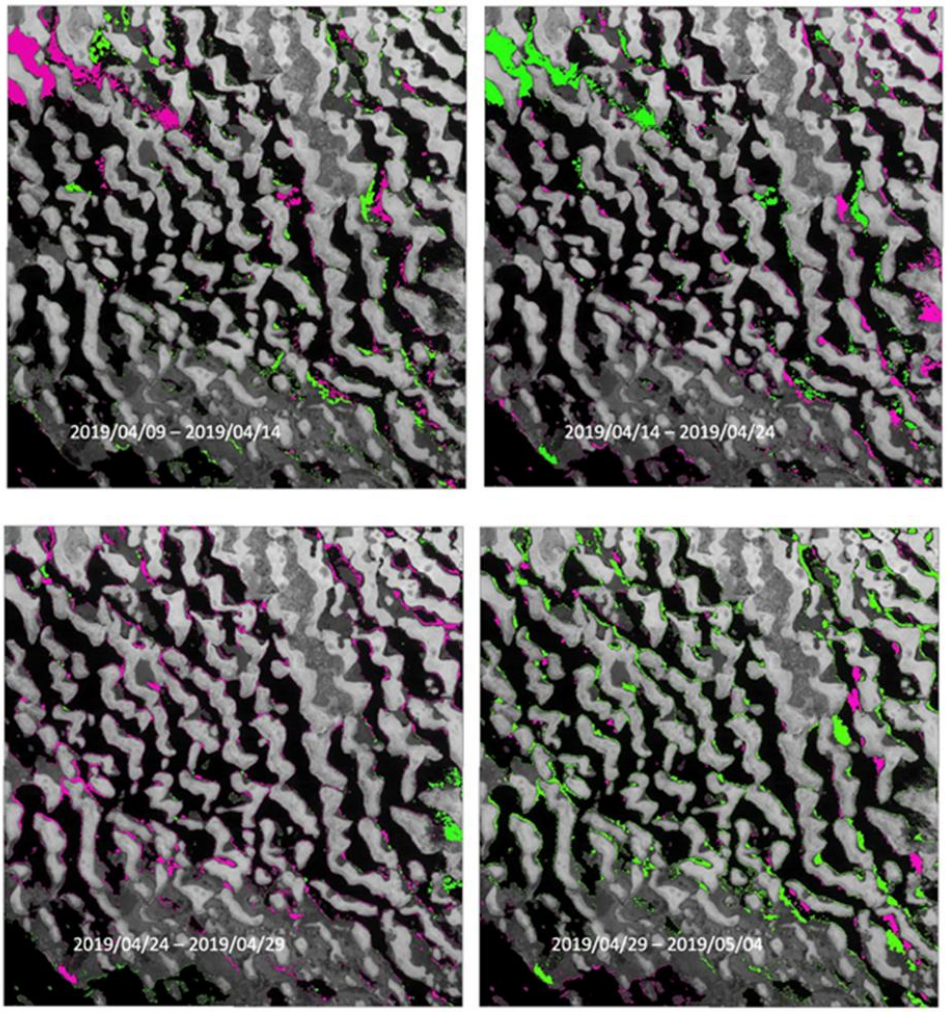


Figure 28: Displacement of vegetated islands from date to date. Green: No water in first image, water in last image. Lilac: Water in first image, no water in last image

3.4.2.3 Validation with boat track

A field trip has been organized by LEGOS on parallel of the Lakes_cci work. Boat tracks on Landsat, using a GPS receiver carefully mapped water and vegetation. The tracks also follow through vegetated islands, confirming the water aspect of these 'islands'. All changes in the nature of the terrain crossed by the boat were noted to map the lake contour along the boat route.

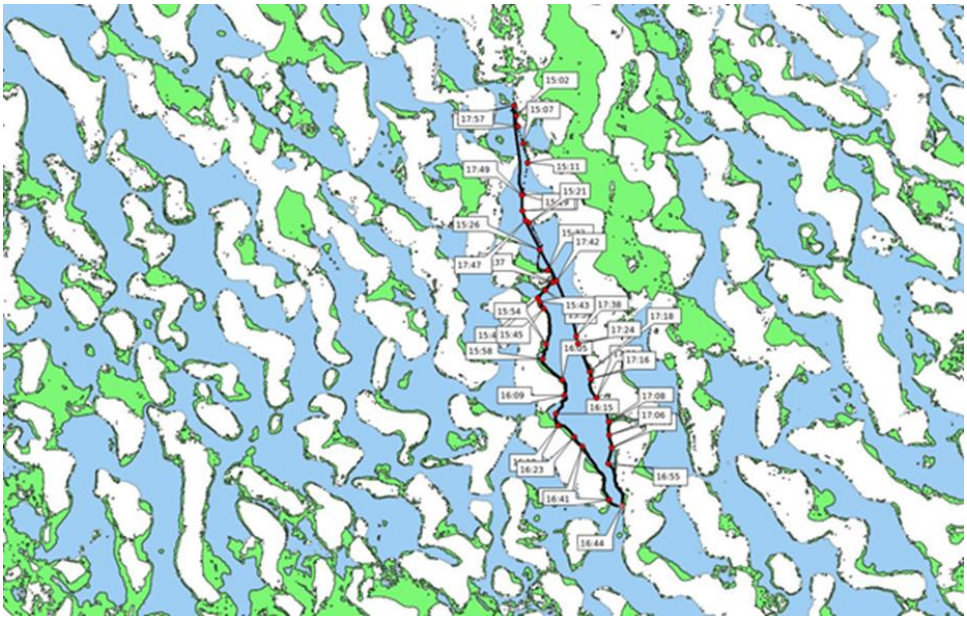


Figure 29: Tracks of boat survey, 10th April 2019, Lake Chad Archipelago (Courtesy of Legos)

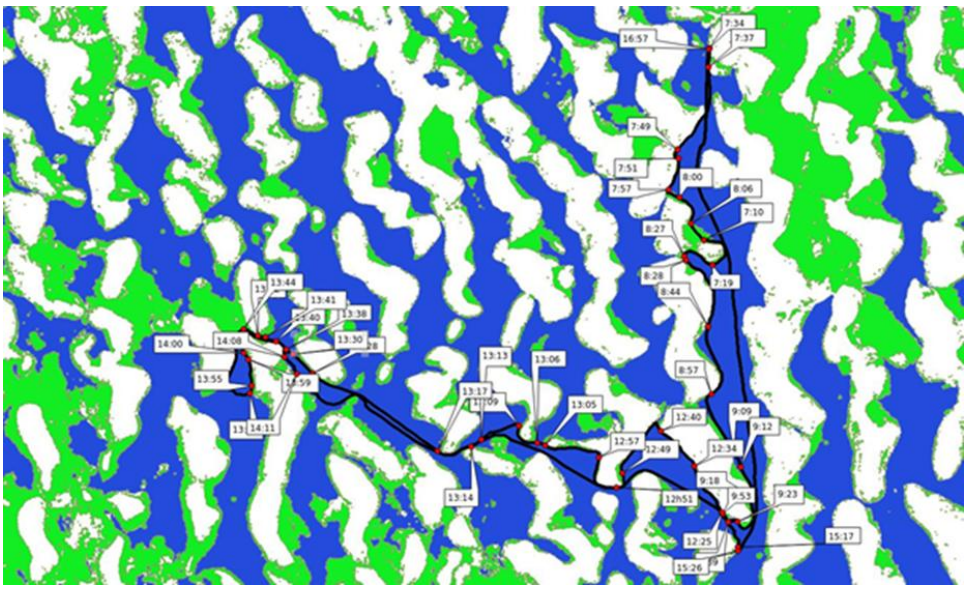


Figure 30: Tracks of boat survey, 11th April 2019, Lake Chad Archipelago (Courtesy of Legos)

3.4.3 Validation of HR extracted water extent based on VHR imagery

As previously mentioned, validation was carried out over selected lakes representing different lake environments and morphological complexity. These sites are located in Europe, Africa, Northern America. Example results are provided in the following sub-sections.

3.4.3.1 Lakes in Champagne (France)

In this area, several lakes are linked to the flood protection system of Paris by the Grands lacs de Seine establishment. They were covered by the recently launched Pleiades NEO sensors as well by Pleaides imagery.

In a first step, two pairs of Pleiades HR data, 50 cm of spatial resolution, a panchromatic channel and 4 visible ones from blue to near Infrared channels, were acquired on the 30th of December, with a delay of one day with Sentinel-2 and at 6 of January 2020, same date as a Sentinel-2 acquisition.

Figure 31: The Der Lake; as acquired by Pleiades on the 30th of December 2019, and Sentinel-2 on the 1st of January 2020

When comparing the LWE derived with Pleaides and Sentinel-2 data acquired within 24 hours, the difference in term of surface is very low, i.e. one 29.07 km2 for Sentinel-2, and 30.58 km2 for Pleiades. 95% of Pleiades Water is recognized by Sentinel-2. There is a very low level of commission, 0.05 km^2 (Figure 32).

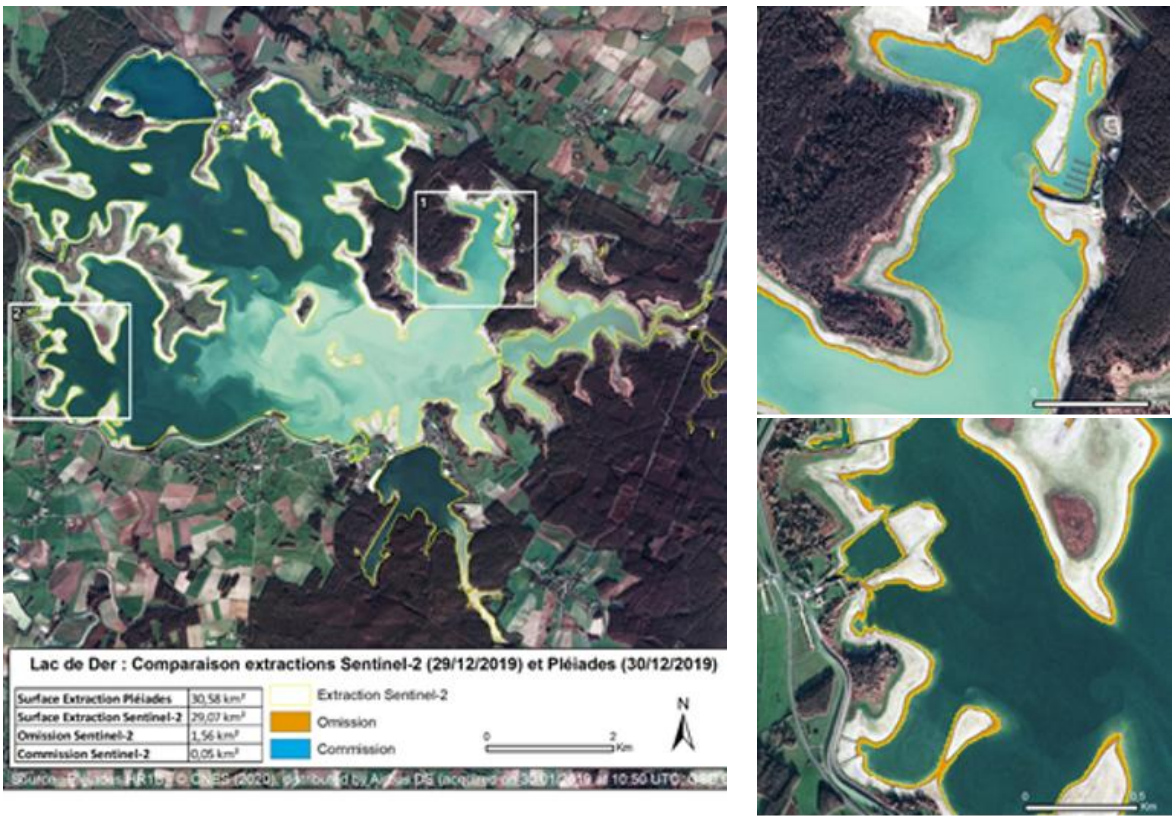


Figure 32: Omission and commission over Lake de Der when exploiting VHR and HR images acquired with a 24h of delay

When the acquisition of the VHR and HR images was the same day, 95,5% of Pleiades Water is recognized by Sentinel-2. There is an effect of resolution i.e. 10 m versus 0.70 m. The shoreline is much finer on the VHR derived LWE. Therefore, the omission is relatively low, and the commission very low.

This study allows us to evaluate, in the context of an infilling reservoir, the effect of a 24h delay between the two acquisitions.

When data are acquired the same day, the space occupied along the shore of the omission is very narrow. The shorelines show a sawtooth effect, alternating omission and commission pixels, which is due to differences in image resolution.

Where the LWEs represent two stages of infilling, we observe a large omission belt around the lake shore, corresponding to the increase of the surface of water within one day.

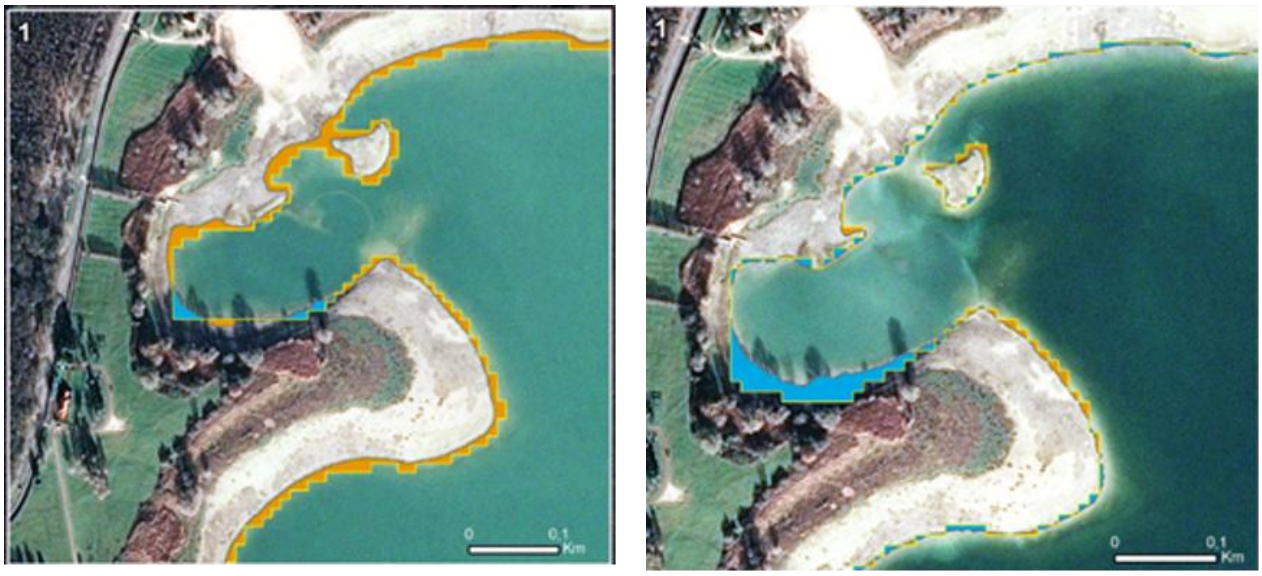


Figure 33: Comparison of LWE accuracy derived with one day of delay between the Sentinel-2 and Pleaides acquisitions and the same day (right).

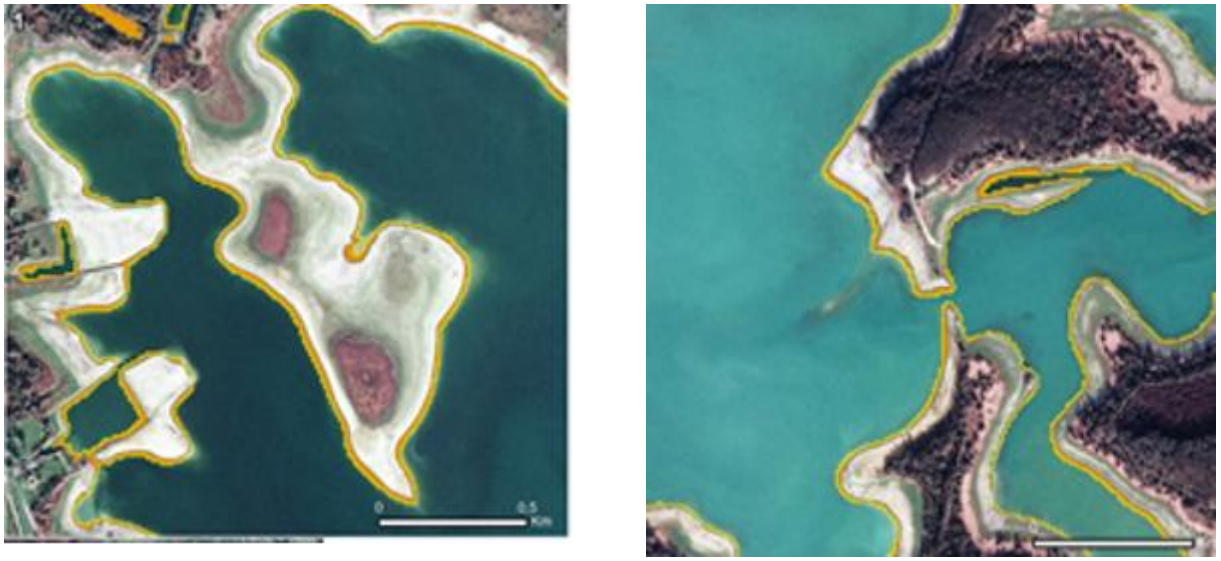


Figure 34: Omission and commission over Lake de Der when exploiting VHR and HR images acquired quasi simultaneously.

For the next example, Pleiades NEO was acquired on 15 December 2022 whereas Sentinel-2 MSI data were acquired on 13 and 16 December 2022. Due to presence of snow, the comparison was carried out for the VHR/HR pair of 15 and 13 December 2022. This winter period corresponds to low water level, when open water is surrounded by sandy-muddy banks. This offers a good opportunity to illustrate the classical effects related to the different spatial resolution, with a sawtooth/staircase limit on Sentinel-2 LWE compared to the straighter limits of VHR shown in Temple Lake (Figure 35). This comparison also highlights the apparent commission related to the exploitation of the SWIR band from Sentinel-2, shown in detail for Orient Lake (Figure 36). At this low water level, SWIR bands are not so effective at differentiating water from the surrounding environment, and do not separate wet sandy-muddy banks from open water.

Figure 35: Effect of resolution in Temple Lake. The red line marks LWE from Pleiades NEO (30cm); the yellow line shows the saw tooth limits derived from Sentinel-2 MSI

Figure 36: Comparison of VHR/HR in Orient Lake Red lines are the LWE limits derived from Pleaides NEO, yellow lines the Sentinel LWE limits

Table 16: Metrics obtained over the Champagne lakes

Metric	Frequency	Area (m^2)	Percentage
Reference	72	21693449.96	100.0
Database	23	2180055.36	100.5
Omission	72	570486.23	2.6
Commission	23	677091.66	3.1
Precision	1	1.00	97
Recall	1	1.00	97
F-score	1	1.00	97

The metrics obtained on these reservoirs (Table 16) show very high values both for the classic precision and accuracy rates and for the more interesting parameters such as the F-Score and Recall. This reflects excellent recognition and extraction of water bodies using Sentinel-2 data via ExtractEO, qualifying the proposed approach for generating LWEs.

3.4.3.2 Fern Ridge Reservoir (Oregon, USA)

Fern Ridge Reservoir (or Fern Ridge Lake) is a reservoir on the Long Tom River in Oregon (USA). The reservoir is a U.S. Army Corps of Engineers flood control project encompassing near 49 km^2. Surrounding portions of the reservoir is the Fern Ridge Wildlife Area, a wildlife management area of wetlands. This case also illustrates the problem of the definition of a lake, what are the limits of a water body of the zones of open water only or zones of open water plus the surrounding flooded wetlands?

By comparing the LWE extracted from the VHR and HR images, Sentinel-2 detected more water in the eastern part of the reservoir (Figure 37). These pixels, corresponding to wet areas of mixed waterlogged vegetation located in the Ridge Fire zone, were extracted and recognised as water areas from the Sentinel-2 data. This confusion between open water and wetland stems from the use of the Sentinel-2 SWIR bands, spectral bands which are highly discriminating for open water, but which make it difficult to differentiate these wetlands from waterlogged areas, just as in other cases from sand or mud banks. In the case of the Fern Ridge site, this relative commission represents around 4% of the water surface.

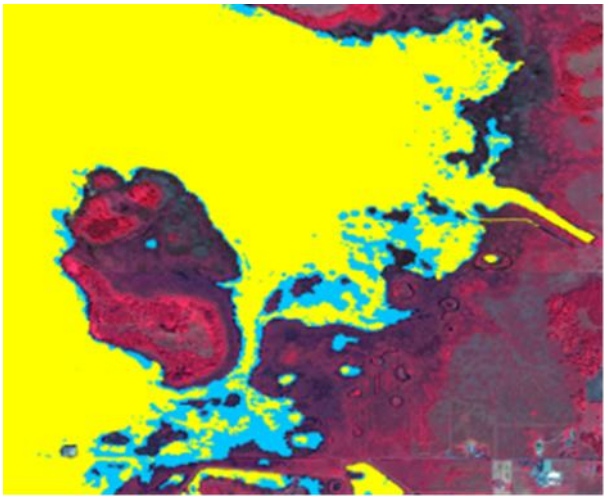


Figure 37: Comparison of LWE derived from Pleiades in yellow and Sentinel-2 in blue over Fern Ridge

In terms of metrics, despite this relative omission related to wetlands, correspondence was good in terms of accuracy and F Score (Table 17).

Table 17: Metrics obtained over Fern Ridge reservoir for the VHR/HR pairs of the Pleiades and Sentinel-2 on 26/07/2022 and 25/07/2022 (left), and for Pleaides and Sentinel-2 acquired on 29/07/2022 (right).

	Pleiades on 26/07/2022, S-2 on 25/07/2022				Pleiades and S-2 on 29/07/2022		
Type	Frequency	Sum Area	Percent		Frequency	Sum Area	Percent
Reference	91	27242769.9	100		102	27205236.4	100
Database	42	28411100.	104.29		51	28188500.	103.61
Omission	84	347057.1	1.27		94	350991.15	1.29
Commission	42	1515387.2	5.53		50	1334254.74	4.73
Detection rate	31	26895712.8	98.73		36	26854245.2	98.71
Accuracy rate	1	999999.0	94.67		1	999999.0 1.0	95.27
Break	1	1.0	1		1	1.0	1
Precision	1	1.0	0.95		1	1.0	0.95
Recall	1	1.0	0.99		1	1.0	0.99
F_score	1	1.0	0.97		1	1.0	0.97
CSI	1	1.0	0.94		1	1.0	0.94

3.4.3.3 Lake Fitri (Chad)

Lake Fitri is the largest lake inside Chad, playing a major role in this Sahel region for human activities and as a refuge for biodiversity. Lake Fitri occupies an endoreic watershed with a surface area of 80,000 km^2 located in the middle of the Sahel to the east of Lake Chad. Lake Fitri's dynamic is directly linked to the West African monsoon regime, who's seasonal to multi-annual variability directly impacts variations in its level and its extent.

The depth of the lake is less than 2.5 m (low-water period). From 1995 onwards, based on Landsat data, a gradual resurgence of the lake and its wetland was observed. Observations in recent years (2017-2022) confirm the trend towards increased extension of the lake during the rainy season. These analyses were carried out using Sentinel-2 satellite images with a resolution of 10 m (Maxant et al., 2023). They clearly highlight the surplus hydrological years of 2020 and 2022, with water surfaces of 1249 and 1180 km^2 respectively, compared with the 194 km^2 observed in 2017.

Two approaches of processing Sentinel-2 MSI were compared for Lake Fitri, one based on 10 m spatial resolution and the other at 20 m of spatial resolution.

When analysing the obtained LWE metrics (Table 18), values are very similar between the two image resolutions, with slightly higher omission at 20 m and slightly more commission at 10 m. At both Sentinel-2 resampled resolutions, the recall and precision are relatively good for this complex hydro- system. Of course, part of the omission is related to the difference in spatial resolution between the VHR and HR images.

Table 18: Metrics obtained from the comparison of LWE derived from Pleiades and Sentinel-2 MSI at 20 m (left), and 10 m (right).

	S-2 20m versus Pleiades				S-2 10m versus Pleiades		
Type	Frequency	Sum Area	Percent		Frequency	Sum Area	Percent
Reference	63797	370516696.8	100		63797	370516696.8	100
Database	7629	353542800	94.42		8193	361739810.3	97.6
Omission	62591	48843863.9	13.18		94	46444991.5	12.53
Commission	7459	31869966.8	9.01		50	37668104.8	10.41
Detection rate	7353	321672833.1	86.82		36	32417105.3	87.46
Accuracy rate	1	999999.0	90.99		1	999999.0	89.59
Break	1	1.0	1		1	1.0	1
Precision	1	1.0	0.91		1	1.0	0.90
Recall	1	1.0	0.87		1	1.0	0.87
F_score	1	1.0	0.89		1	1.0	0.86
CSI	1	1.0	0.80		1	1.0	0.79

3.4.3.4 Lake Shenjin

Lake Shengjin, located on the south bank of the middle and lower reaches of the Yangtse River (Anhui province, PR China), is a large shallow freshwater lake with a shoreline extending to 156 km at 11 m above sea level. The case of Lake Shenjin illustrates the difficulties to extract water extent in case of shallow waters, when the delimitation of water and wet sediments can be very delicate, and the bottom of the lake can play an important role in terms of returning signal. This is particularly sensitive during the step of sampling selection process. Two sets of parameters were tested to reduce the weight of the SWIR channels that can induce commission over wetland and mud/sandy banks such.

When analysing the LWE limits derived from Pleaides NEO and the two LWE results from Sentinel-2, the limits of the Pleaides NEO and Sentinel-2 with parametrization 1 are generally very similar, whereas the LWE obtained with the second parameterization set introduces a buffer zone along the water limits, including very light sandy areas (Figure 38).

Figure 38: Comparison of the LWE limits derived from Pleiades NEO and Sentinel-2

When analysing the LWE metrics (Table 19), results obtained from parameter set 2 present high values for commission and relative low accuracy values as it can be observed on the Figure 38. The commission rate decreases with the first parametrization while omission increases. These omissions are mostly related to very fine water paths observed at 0.30 cm and not at 10 m resolution. Over this complex shallow water system, the different scores are relatively good with recall or F Score values about 0.85-0.87.

Table 19: Metrics obtained from the comparison of the LWE derived from Pleaides Neo and two parametrizations of the sampling step for water recognition from Sentinel-2.

	Parametrization 1			Parametrization 2		
Type	Frequency	Sum Area	Pourcent	Frequency	Sum Area	Pourcent
Reference	513	32691145.8	100	513	32691145.8	100
Database	41	33698468.7	103.8	102	46942058	143.59
Omission	442	4099507.1	12.54	333	1155269.43	5.53
Commission	30	5106830	15.15	100	15406181.6	32.82
Detection rate	36	28591638.7	87.46	69	31535876.3	96.47
Accuracy rate	1	999999.0	84.85	1	999999.0	67.18
Break	1	1.0	1	1	1.0	1
Precision	1	1.0	0.85	1	1.0	0.67
Recall	1	1.0	0.87	1	1.0	0.96
F_score	1	1.0	0.86	1	1.0	0.79
CSI	1	1.0	0.75	1	1.0	0.66

3.5 Conclusion and recommendations

We analysed the behaviour of LWE with respect to source product processing level, the resolution of Landsat/Sentinel-2 and on lakes with varying characteristics to provide a representative overview of the observation challenges and quality of LWE estimates from high-resolution optical imagery. In addition, validation tasks were carried out in a consistent manner, exploiting VHR imagery, and higher resolution images such Pleiades NEO data at 30 cm. These analyses have been carried out for different environment conditions and lake types. Obtained results confirm the accuracy of the high resolution water extent extraction and support the chosen approach. Weaknesses and strengths of the followed approach to generate LWE within the Lakes_cci can be analysed on the basis of these results. Specific challenges such as cloud cover and associated shadows, local environment, the presence of vegetation within the water bodies, floating or submerged ones, are recognised. Many of the characteristics of water conditions are related to the colour of the water in shallow water.

A persistent problem is the definition of LWE, i.e. whether we consider open free water or open free water including ice cover, and whether we speak of continuous open water when the periphery contains inundated wetlands or connected water basins. Specific cases, such as floating islands, would logically be included in the definition of a ‘water body’ but not necessarily in ‘open water extent’. In an automated processing workflow, these are special cases without clear definition. Depending on the followed classification or algorithmic approach, or selection of training samples, omission or commission is possible.

For the operational application, i.e. monitoring LWE based on LWL, generating precise hypsometric curves, each curve requires prior inspection. More work may be needed to select, access and process appropriate images, exploiting systems such as Google Earth Engine and ESA Science Hub, and also some

developments of tools are requested such as corrections tools for missing data corrections, exploiting exogeneous database such as the Global Surface Water Explorer.

3.6 References

Iltis A., Lemoalle J., 1983. The aquatic vegetation of Lake Chad p 125 - 143 in Lake Chad, Carmouze J.P., Durand J.R. & Lévèque C. (eds), Monogr. Biologicae 53, Junk, 575 p

Leonard J., 1974: Aperçu de la végétation de la partie Est du lac Tchad. ORSTOM

Maxant J, Braun R, Caspard M, Clandillon S. (2022). ExtractEO, a Pipeline for Disaster Extent Mapping in the Context of Emergency Management. Remote Sensing. .14(20):5253. https://doi.org/10.3390/rs14205253

Peña-Luque, S.; Ferrant, S.; Cordeiro, M.C.R.; Ledauphin, T.; Maxant, J.; Martinez, J.-M. Sentinel-1&2 Multitemporal Water Surface Detection Accuracies, Evaluated at Regional and Reservoirs Level. Remote Sens. 2021, 13, 3279. https://doi.org/10.3390/rs13163279

Peña-Luque Santiago, 2019. CNES ALCD Open Water Masks (Version 1.1): Zenodo. Available online: https://zenodo.org/record/3675333#.YR4dmd8RWUk

4 Lake Surface Water Temperature – LSWT

The Lake Surface Water Temperature LSWT is defined as the temperature of the water at the surface of the water body (surface skin temperature). The CCI LSWT dataset consists of a long-term climate data record (CDR). The validation and comparison of the LSWT is based on matchups between in situ and satellite measurements. In situ measurements are collated once per year from limnologists worldwide who are willing to share in situ data of lake surface temperature measurements on a personal, ad-hoc basis supplementing the partial collections of some agencies. The annual collation is undertaken towards the end of the calendar year and includes quality control of the in situ observations, adding measurements that are suitable for satellite validation to the validation database.

4.1 Data description

4.1.1 Satellite data

LSWT time series have been computed from sensors on multiple satellites and lake-specific consistency adjustments between sensors have been applied using the MetOpA AVHRR instrument as a reference (see ATBD). In the current processor version, LSWT v4.5, MetOpA AVHRR has the best combination of length of record and data density for this purpose. A quality level is assigned to each pixel-observation as described in the ATBD. For the LSWT derived from MODIS, only the highest quality levels (4 and 5) have been used for the dataset and a constant bias correction (0.11 for quality level 5, 0.19 for quality level 4) has been applied across all lakes after a comparison with LSWT from MetOpA AVHRR. The same form of algorithms has been used to retrieve the LSWT from all sensors to obtain consistent time series for each of the lakes in the CRDP.

The time periods used for each satellite/instrument are provided in Table 20. Not all lakes include LSWT from all sensors in the series because of differing density and geometry of observation. The temperatures in the dataset are only available for cloud and ice-free observations during (in this version) day time, so gaps in time and space are common for all the lakes due to cloud cover and swath geometry of the instruments. The LSWT in the CRDP v3.0 includes observations until the end of 2023 and includes reprocessed SLSTR data so that only the latest NTC (Non-Time Critical) L1b data are used.

Table 20: Time periods for the satellite/instrument used to generate the LSWT product

Satellite	Instrument	Time Period
ERS-2	ATSR2	Jun1995 – Jun2003
Envisat	AASTR	May2002 – Apr2012
MetOpA	AVHRR	Mar2007 – Aug2019
MetOpB	AVHRR	Dec2012 – Aug2019
Terra	MODIS	Feb2000 – Dec2022
Sentinel3A	SLSTR	Jun2016 – Dec2023
Sentinel3B	SLSTR	Aug2018 – Dec2023

A detailed description of the product generation is provided in the ATBD with further information on the product given in the PUG.

4.1.2 In situ data

The in situ dataset currently used for validation has been constructed from the in situ temperature data collected through the ARCLake project, the GloboLakes project, the EU Surface Temperature for All Corners of Earth (EUSTACE) project and the Copernicus Climate Change Service (C3S). At present, this dataset consists of 465 observation locations covering 177 lakes. The overlap with lakes included in the Lakes_cci consists of 134 lakes with a total of 384 sites. Details of the in situ observation locations with their sources are given in Table 21 which reports all locations for the target lakes where there are matches. The geographical distribution of the sites is reported in Figure 39 which shows that most of the globe is covered, with a large proportion of the sites located in North America and Europe.

Table 21: List of the sources of the in situ data

Source	Lake name (number of observation sites)
NDBC – National Data Buoy Centre (USA)	Superior (3) Huron (2) Michigan (2) Erie (1) Ontario (1)
FOC – Fisheries and Oceans Canada (Canada)	Superior (1) Huron (4) Great Slave (2) Erie (2) Winnipeg (3) Ontario (4) Woods (1) Saint Claire (1) Nipissing (1) Simcoe (1)
Michigan Technological University (USA)	Superior (2) Michigan (1)
University of Minnesota (USA)	Superior (2)
Northern University of Michigan (USA)	Superior (3)
Superior Watershed Partnership (USA)	Superior (1)
U.S. Army Corps of Engineers (USA)	Superior (1)
Technical University of Kenya (Kenya)	Victoria (1)
GLERL – Great Lakes Environmental Research Lab (USA)	Huron (3) Michigan (2)
University of Wisconsin-Milwaukee (USA)	Michigan (2)
Northwestern Michigan College (USA)	Michigan (1)
University of Michigan CIGLR (USA)	Michigan (2)
Limno Tech (USA)	Michigan (3) Erie (4)
Illinois-Indiana Sea Grant and Purdue Civil Engineering (USA)	Michigan (2)
Leibniz Institute for Freshwater Ecology and Inland Fisheries (Germany)	Tanganyika (1)
Pierre Denis Plisnier	Tanganyika (4)
GLEON	Tanganyika (3)
Irkutsk State University (Russia)	Baikal (1)
Regional Science Consortium (USA)	Erie (1)
UGLOS – Upper Great Lakes Observing System (USA)	Erie (2) Douglas (1)

Source	Lake name (number of observation sites)
Xavier Lazzaro	Titicaca (1)
LEGOS – Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (France)	Issykkul (1)
SLU – Swedish University of Agricultural Science (Sweden)	Vanern (6) Vattern (2) Malaren (9) Hjalmaren (1) Siljan (1) Bolmen (2) Ekoln (1) Roxen (2)
Uppsala University (Sweden)	Vanern (1) Erken (1)
GEMStat – The Global Freshwater Quality Database – United Nation Environment Programme	Laguna Mirim (3) Peipsi (1) Laguna Madre (12) Rincon del Bonete (4) Salto Grande (2) Laguna Tamiahua (5) Biwa (2) Bodensee (1) Miguel Aleman (2) Lago Gatun (6) Nezahualcoyotl (5) Neagh (1) Great Prespa (2) Corrib (14) Derg (17) Muritz (2) Trichonida (1) Laguna de Castillos (4) Chiemsee (1) Ijsselmeer (2) Lekshm (1) Oestra Ringsjoen (1) Selenter (1) Kummerower (1) Schwerin (1) Steinhuder (1) Ammersee (1) Starnberger (1) Walchensee (1) Mueggelsee (1) Kerkini (1)
Estonian University of Life Sciences (Estonia)	Peipsi (18) Vorstjarv (4)
Sao Paulo State University (Brazil)	Tucurui (1) Itaipu (1) Tres Marias (1) Serra da Mesa (1) Furnas (1) Itumbiara (1)
Junsheng Li (China)	Taihu (1)
KU Leven (Belgium)	Kivu (1)
ENABEL for Agence Belge de Developpement at the Universite de Namur and CARRTEL through GEISHA	Kivu (1)
SYKE – Finnish Environment Institute (Finland)	Inarinjarvi (1) Päijänne (3) Pielinen (4) Oulujärvi (1) Keitele (1) Näsijärvi (1) Lokan (1) Onkivesi (2) Puulavesi (1) Höytiäinen (1) Koitere (1) Vanajavesi (1) Pyhäjärvi (1) Lappajärvi (1) Mallasvesi (2) Vuohijärvi (1) Lentua (1) Miekojärvi (1) Pyhäjärvi (1)
Vermont EPSCOR – Established Program to Stimulate Competitive Research (USA)	Champlain (1)
SUNY Plattsburgh Center for Earth and Environmental Science (USA)	Champlain (1)
Nipissing University (Canada)	Nipissing (2)
National Park Service (USA)	Mead (3) Mohave (2)
NIWA (New Zealand)	Taupo (3) Rotorua (1)
GLEON – Global Lake Ecological Observatory Network	Tanganyika (3) Balaton (1)
BLI – Balaton Limnological Institute (Hungary)	Balaton (6)

Source	Lake name (number of observation sites)
KDKVI – Central Transdanubian (Regional) Inspectorate for Environmental Protection, Nature Conservation and Water Management (Hungary)	Balaton (3)
Hungarian Met Service (Hungary)	Balaton (1)
UMR CARRTEL – Centre Alpin de Recerche sur le Réseaux Trophique des Ecosystèmes Limniques (France)	Geneva (1)
EAWAG - Swiss Federal Institute of Aquatic Science and Technology (Switzerland)	Geneva (1) Neuchatel (1) Maggiore (1) Thun (1)
Environmental Data Initiative (USA)	Winnebago (1) Diefenbaker (1) Marion (1) Moultrie (1) Truman (1) Last Mountain (1) Narach (1) George (1)
Flathead Lake Bio Station (USA)	Flathead (5)
UC-Davis Tahoe Environmental Research Center (USA)	Tahoe (1)
Utrecht University (Netherlands)	Garda (1)
CNR Italian National Research Council (Italy)	Garda (8) Trasimeno (2) Maggiore (2) Como Mezzola (5) Bolsena (1) Iseo (2) Bracciano (1)
ARPA VENETO (Italy)	Garda (2)
NOAA National Ocean Service Water Level Observation Network (USA)	St John River (3)
USGS U.S. Geological Survey (USA)	Upper Klamath (11) Calcasieu (2) Salvador (4) George (2) Walker (2) Mattamuskeet (2) Abert (1) Mohave (1)
Cornell University Biological Field Station (USA)	Oneida (1)
CNR IRSA Institute for Water Research (Italy)	Maggiore (1)
Environmental Protection Agency (Ireland)	Corrib (2) Derg (1)
Martin Dokulil (Austria)	Neusiedl (1)
Israel Oceanographic and Limnological Research (Israel)	Sea of Galilee (2)
GEISHA project	Sea of Galilee (1)
National Institute for Environmental Studies (Japan)	Kasumigaura (5)
Universidad del Valle de Guatemala (Guatemala)	Atilian (1)
Portland Water District (USA)	Sebago (1)
Universitá degli Studi di Perugia (Italy)	Trasimeno (1)
University of Waikato and the Bay of Plenty Regional Council (New Zealand)	Rotorua (1)

Source	Lake name (number of observation sites)
Centre for Ecology and Hydrology - Edinburgh (UK)	Lomond (1) Leven (1)
Freie Universitat Berlin/Fondazione Edmund Mach (Germany/Italy)	Iseo (1)
University of Latvia and Latvian Environmental Geology and Meteorology Centre (Latvia)	Razna (1)
University of Wisconsin-Madison (USA)	Mendota (1)
NTL LTER – North Temperate Lakes Long-Term Ecological Research (USA)	Mendota (1) Trout (1)
Uppsala University (Sweden)	Erken (1)
The Ohio State University (USA)	Douglas (1)
ARPA LOMBARDIA (Italy)	Idro (1) Como (5) Iseo (1)

Figure 39: Geographical distribution of sites used for LSWT validation.

Table 22 lists the 134 lakes with their maximum distance from land (Carrea et al. 2015), which is an indication of lake size that is meaningful for LSWT remote sensing. The distance to land for lake Iseo in Italy is shown in Figure 40. The best resolution of the instruments used for the retrieval of the LSWT is 1 km. If the lake has a maximum distance to land of 1.7 km, such as in lake Iseo, the LSWT retrieval is very

likely to be available only for that part of the lake and only for a limited proportion of overpasses (clear sky and observations relatively central within the swath). In particular, a combination of factors has to occur: 1) the satellite image locations line up so that some pixels are nominally fully water pixels, which requires the satellite view zenith angle (which affects the on-the-ground resolution) to be such that the half-pixel size is smaller than the distance to coast; 2) these pixels are cloud free; 3) image geolocation errors (which can be of order 1 pixel uncertainty) are small enough that the nominally water-filled pixels are truly water-filled meaning that the water detection tests are passed.

Table 22: Lakes_cci water bodies with in situ data matches

Lake id	Name	Country	N sites	Max distance to land (km)
2	Superior	Canada/USA	13	73.5
3	Victoria	Tanzania	1	84.1
5	Huron	Canada/USA	9	73.3
6	Michigan	USA	15	63.8
7	Tanganyika	Tanzania	8	34.1
8	Baikal	Russia	5	33.7
10	Malawi	Malawi	1	34.8
11	Great Slave	Canada	2	44.6
12	Erie	Canada	10	45.6
13	Winnipeg	Canada	3	40.1
15	Ontario	Canada	5	36.1
20	Titicaca	Bolivia/Peru	1	23.4
25	Issykkul	Kyrgyzstan	1	26.9
29	Vanern	Sweden	7	20.3
44	Woods	Canada	1	11.8
46	Lagoa Mirim	Brazil/Uruguay	3	18
50	Peipsi	Estonia/Russia	19	20.6
52	Tucurui	Brazil	1	6.4
65	Itaipu	Paraguay	1	3.8
66	Taihu	China	1	16
67	Kivu	Zaire	2	13
95	Vattern	Sweden	2	9.9
96	Serra da Mesa	Brazil	1	4
101	Laguna Madre	Mexico	12	9.6
132	Rincon del Bonete	Uruguay	4	6.5
139	Furnas	Brazil	1	2
144	Inarinjarvi	Finland	1	4.1
146	Saint Claire	Canada	1	13
157	Paijanne	Finland	3	3.8
161	Salto Grande	Uruguay/Argentina	2	3.4
163	Malaren	Sweden	9	2.7
165	Champlain	USA	2	5.8
188	Tres Marias	Brazil	1	2.3
195	Pielinen	Finland	4	4.1

Lake id	Name	Country	N sites	Max distance to land (km)
198	Nipissing	Canada	3	9
202	Oulujarvi	Finland	1	6
235	Laguna Tamiahua	Mexico	5	7.3
236	Simcoe	Canada	1	8.4
238	Itumbiara	Brazil	1	3
268	Biwa	Japan	2	8.2
278	Mead	USA	4	3.8
295	Taupo	New Zealand	3	9.6
310	Balaton	Hungary	11	6
327	Geneva	Switzerland	2	6.2
340	Winnebago	USA	1	7.8
346	Keitele	Finland	1	2.2
352	Bodensee	Austria/Switzerlan /Germany	1	6
369	Nasijarvi	Finland	1	2.9
375	Flathead	USA	5	5.9
376	Lokan	Finland	1	4.4
380	Tahoe	USA	1	8.2
387	Hjalmaren	Sweden	1	4.7
402	Diefenbaker	USA	1	2.1
422	Onkivesi	Finland	2	2.4
452	Miguel Aleman	Mexico	2	3.1
457	Lago Gatun	Panama	6	2.6
477	Nezahualcoyotl	Mexico	5	2.6
481	Neagh	UK	1	6.6
505	Garda	Italy	9	5.2
507	St John River	USA	1	2.4
530	Puulavesi	Finland	1	2.1
556	Marion	USA	1	4.1
559	Upper Klamath	USA	11	4
654	Siljan	Finland	1	5.4
657	Hoytiainen	Finland	1	3.2
679	Vorstjarv	Estonia	4	6.2
690	Great Prespa	Greece	2	4.8
748	Calcasieu	USA	2	5
824	Moultrie	USA	1	5.6
836	Salvador	USA	4	56
893	Nuechatel	Switzerland	1	3.6
905	Oneida	USA	1	4.2
929	George	USA	2	5.1
948	Maggiore	Italy	2	2.4

Lake id	Name	Country	N sites	Max distance to land (km)
1028	Bolmen	Sweden	2	2.7
1057	Corrib	Ireland	14	2.6
1075	Last Mountain	USA	1	2.5
1115	Neusiedl	Austria	1	3.6
1128	Walker	USA	2	4.3
1148	Mattamuskeet	USA	2	4.4
1154	Abert	USA	2	4.4
1196	Sea of Galilee	Israel	2	5.6
1201	Vanajavesi	Finland	1	3.4
1204	Kasumigaura	Japan	5	3.7
1240	Pyhajarvi	Finland	1	3.9
1246	Lappajarvi	Finland	1	3.5
1479	Atilian	Guatemala	1	4
1498	Sebago	USA	1	4.2
1519	Derg	Ireland	18	1.9
1529	Trasimeno	Italy	3	4.3
1596	Bolsena	Italy	1	5.2
1649	Muritz	Germany	2	3.5
1679	Ekoln	Sweden	1	1.6
1893	Roxen	Sweden	2	2.9
1913	Trichonida	Greece	1	2.9
2054	Vuohijarvi	Finland	1	1.9
2092	Laguna de Castillos	Uruguay	4	3.6
2317	Chiemsee	Germany	1	3.2
2425	Narach	Belarus	1	3.7
2516	Lomond	UK	1	1.5
3307	Bracciano	Italy	1	3.9
3379	Razna	Latvia	1	3.2
3493	Crater	USA	1	3.5
4503	Mendota	USA	2	2.5
6785	Erken	Sweden	1	1.5
12262	Leven	UK	1	1.5
12471	Trout	USA	1	1.4
13377	Douglas	USA	2	1.5
100000012	Ijsselmeer	Netherlands	1	10.7
100000016	Ijsselmeer	Netherlands	1	10.7
300000767	George	USA	1	1.8
300001112	Koitere	Finland	2	2.2
300001141	Mallasvesi	Finland	2	2.1
300001274	Como Mezzola	Italy	5	2
300009360	Mohave	USA	3	2.8

Lake id	Name	Country	N sites	Max distance to land (km)
300011716	Myekojarvi	Finland	1	1.5
300012023	Lentua	Finand	1	2.1
300012614	Pyhajarvi	Finalnd	1	1.7
300012491	Lekshm	Russia	1	2.9
300013182	Oestra Ringsjoen	Sweden	1	2
300013383	Selenter	Germany	1	1.7
300013431	Kummerower	Germany	1	1.8
300013454	Schwerin	Germany	1	2.2
300013601	Steinhuder	Germany	1	2.2
300013962	Ammersee	Germany	1	2.1
300013966	Starnberger	Germany	1	2.1
300013992	Walchensee	Germany	1	1.5
300014052	Thun	Switzerland	1	1.6
300014185	Iseo	Italy	3	1.7
300014478	Kerkini	Greece	1	2.8
300016649	Rotorua	New Zealand	2	3.1
300165102	Mueggelsee	Germany	1	1.2

A good portion of the lakes used for the validation are small, for which, given previous discussion, the LSWT retrieval is most challenging.

Moreover, some of the locations of in situ measurements are situated close to the coast even for large lakes, which means that the nearest water-filled pixels may not overlap the in situ measurement, thus increasing the uncertainty in the comparison from spatial representativity.

As the in situ data are from a variety of sources, with different formats, considerable effort has been put in to consolidate each new source of data to a standard format for use in validation. A quality control procedure for checking the in situ data is also necessary, since they are not always plausible. This is partly automated and partly by manual inspection. The quality control procedure was initiated within the ARCLake project and updated within GloboLakes and C3S.

The in situ data have a range of characteristics:

- Measurements have been taken at different depths up to 1 m.
- Temporal sampling of the measurements ranges from 15 minutes to few times a year.
- Temporal availability of the in situ measurements varies from few months up to covering all the satellite period.
- For some locations the measurements are averages while for others they have been taken instantaneously at the reported time.
- None of the in situ measurements which have been collected are provided with an uncertainty estimate.

While part of the data is available online, the majority have been collected through personal communications. In certain cases, we are not licensed to redistribute the data because of provider data policies.

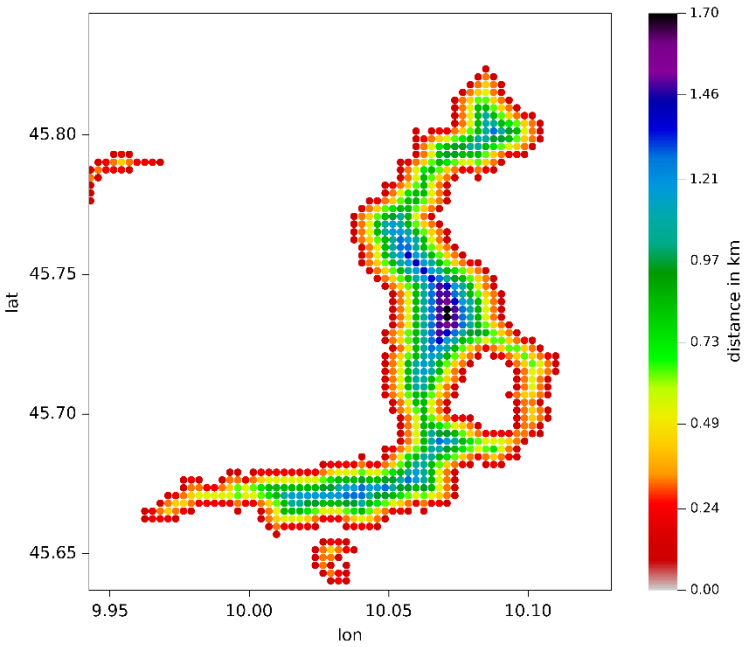


Figure 40: Distance to the nearest land for each pixel on water for lake Iseo in Italy at approximately 300m resolution.

4.2 Comparison methods

The validation of the Lake Surface Water Temperatures consists of the comparison with independent in situ data. The satellite-to-in situ-matches are created at the original satellite coordinates, at L2. The Lakes_cci products are not on satellite coordinates but are gridded in a regular grid at 1/120° resolution and "supercollated" (combined across the available sensors) making a gridded product technically referred to as "L3S". The LSWT of the L3S grid cell of the combined and regridded temperature are therefore directly validated to assess the products as seen by users. The validation of the LSWT is performed using conventional and robust statistics, the latter being less sensitive to outliers and more descriptive of the majority of data.

4.2.1 Generation of the L2 matchups

A per-sensor matchup is created at L2 and it contains satellite and in situ data as nearly as possible co-incident and space and time. The match defines the reference temperature and time from the in situ location and the associated LSWTs, quality level and uncertainty from the L2 LSWT product. The matchup is created for satellite observations based on the following criteria:

- **Spatially** within 2 km from the location of the in situ measurement and
- **Temporally** within 3 hours for the in situ measurements where the measurement time was available. For some of the lakes only daily mean temperature was recorded or the time-of-day of the measurements was not reported, and in these cases the time criterion was to match the day of observation.

4.2.2 Validation of the L3S CCI LSWT v4.5

The differences between the L3S LSWT and in situ data are analysed using both standard and robust statistics. Robust statistics is less influenced by outliers in the distribution of differences, which can be also caused by poor in situ measurements. Quality control measures have been applied to in situ measurements, but many different instruments have been used to take measurements and the operating

methods of the instruments and the reporting vary strongly between sites. Time series of the absolute temperatures together with their difference are generated differentiating the quality levels. Violin plots, where the distribution of the difference is shown, are produced for each quality level. The robust statistics are also investigated per quality level for each year and for each lake.

4.2.3 Validation of the LSWT uncertainty

The validation of the L3S Lakes_cci LSWT v4.5 (CRDP v3.0) is carried out comparing the satellite minus in situ temperature difference with the combination of the satellite uncertainty (present in the products) and an estimate of the in situ uncertainty (which is relatively poorly known). In an ideal case, the standard deviation of the differences between the satellite LSWT and a reference LSWT would equal the combined measurement uncertainty plus the uncertainty attributable to representativity effects.

4.2.4 Number of lakes in the CCI with LSWT

An assessment of the lakes with no retrieved LSWT is reported. Most of the lakes where issues are observed are too small in comparison with the satellite resolution, while others may be included in the next version.

4.3 Description of the work

The matchup is carried out per sensor over the 384 locations on 144 lakes. The total number of matches is 148,100 for any quality level and 121,679 excluding satellite LSWT of quality level equal to 1, which are “bad data”. The number of matches varies per year and since the AVHRR and the MODIS sensors have a larger swath than the ATSR sensors (ATSRs swath is 500 km, AVHRRs swath is ~2900km, MODIS swath is 2330 km and SLSTR swath is ~1400 km), after the year 2000 the number of matches clearly increases as shown in Figure 41. We can notice two other clear increase in 2007 when the AVHRR on MetOpA is introduced and in 2013 when the AVHRR on MetOpA is used together with the AVHRR on MetOpB (Figure 41) data, since a different number of locations is available every year, as shown in Figure 42. The number of locations where in situ measurements have been taken has almost tripled since 1995. However, a portion of the measurement temporal frequency is daily.

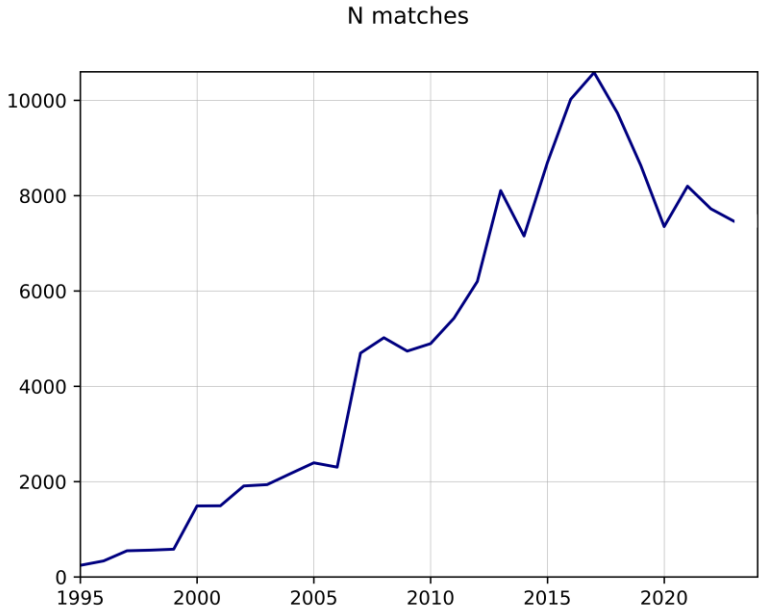


Figure 41: Number of matches for the Lakes_cci at L3 per year.

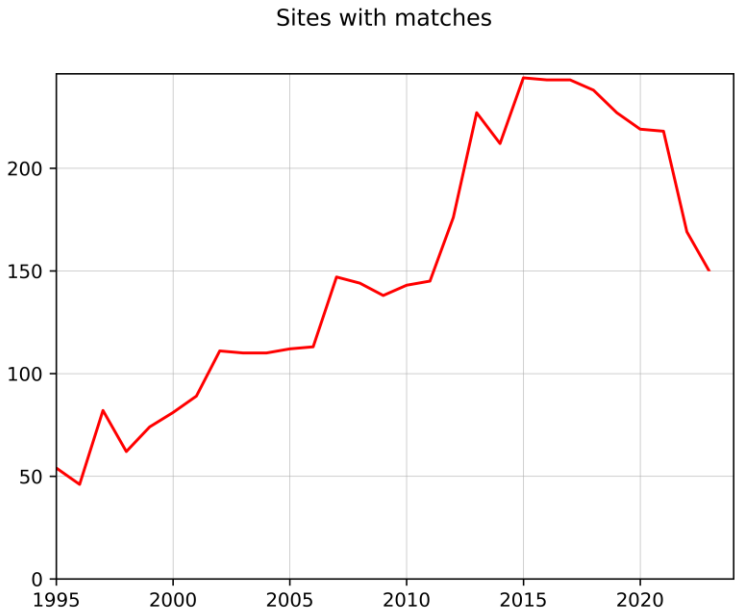


Figure 42: Number of sites with matches for the Lakes_cci at L3 per year.

4.4 Validation results

Robust and traditional statistics per quality level are reported in Table 23 for the matches across all locations where in situ measurements were available.

Table 23: Robust and traditional statistics of the LSWT satellite minus in situ difference.

QL	N	Median	RSD	Mean	SD
5	55175	-0.120	0.563	-0.040	0.975
4	32211	-0.170	0.815	-0.112	1.260
3	22166	-0.220	1.082	-0.286	1.559
2	12127	-0.390	1.571	-0.562	1.891
1	26421	-1.810	3.751	-3.056	4.746

The number of matches per quality level are listed together with the median and the robust standard deviation of the satellite minus in situ temperature difference and the traditional metrics, the mean and the standard deviation. The difference between the median and the mean is between 0.57 and 0.80 K for quality level 3,4,5 and increases as the quality levels get lower (suggesting, as expected, a higher incidence of cold-biased observations for low quality levels). The agreement between satellite and in situ measurements varies according to the quality levels in a way that is expected.

The best agreement is for quality levels 4 and 5, which are the levels that reflect a higher degree of confidence in the validity of the satellite estimate. Our recommendation to users is to use the quality level 4 and 5 for lake-climate applications in general, although lower quality level data may be relevant to users where they have specifically verified their fitness for a given lake for their application. Quality level 3 data comparison with the in situ data shows an agreement that may be acceptable to some users; however,

they have to be used with care. Quality level 1 data should never be used and they are classified as “bad data”.

A contribution to the difference on average is the expected skin effect. Infrared radiometers are sensitive to radiation emitted between the air-surface interface and 20 mm below the interface while the in situ measurements considered here are taken at a distance up to 1 m from the air-surface interface. During the night, the surface of the water is generally cooler than the subsurface by ~0.2 K [Saunders, 1967], [Embury et al, 2012]. However, during the day, if the wind speed is low enough, thermal stratification due to solar heating contributes a positive offset to the difference in temperature between the radiometric lake surface and the in situ measurement depth (up to 1 m). The positive thermal stratification would be expected to be in the range <<1 K for most observations and but occasionally of order a few kelvins (Hondzo et al, 2022). The degree of near-surface stratification to be expected in different lakes depends on fetch, weather conditions (radiative balance and wind speed), the depth of in situ measurement, and any local vertical mixing perturbations introduced by the presence of the in situ measurement system. The aggregate effect of these factors is not currently well quantified. Overall, it is plausible that for day time LSWT observations the mean stratification effect is of order one or a few tenths, as has been determined over the oceans. In summary, a geophysical contribution to the satellite minus in situ temperature difference is the expected skin effect of -0.2 K, but other positive geophysical offsets are similar in magnitude and difficult to quantify precisely. In this context, a mean agreement of the physics-based retrievals and validation within ±0.2 K is a convincing result. In terms of scatter, as well as the retrieval uncertainty and variability in the vertical stratification effects, the scatter includes in situ uncertainty and horizontal variability. Again, quantitative understanding of the scatter from these effects is not yet mature, and for this reason full uncertainty budget validation remains a research aspiration (see also E3UB v2.1).

The distributions of the satellite minus in situ temperature differences per quality level 2,3,4,5 are reported in Figure 43 as box plots. The distributions become more stretched and less symmetric with longer tail towards negative differences as the quality levels decreases.

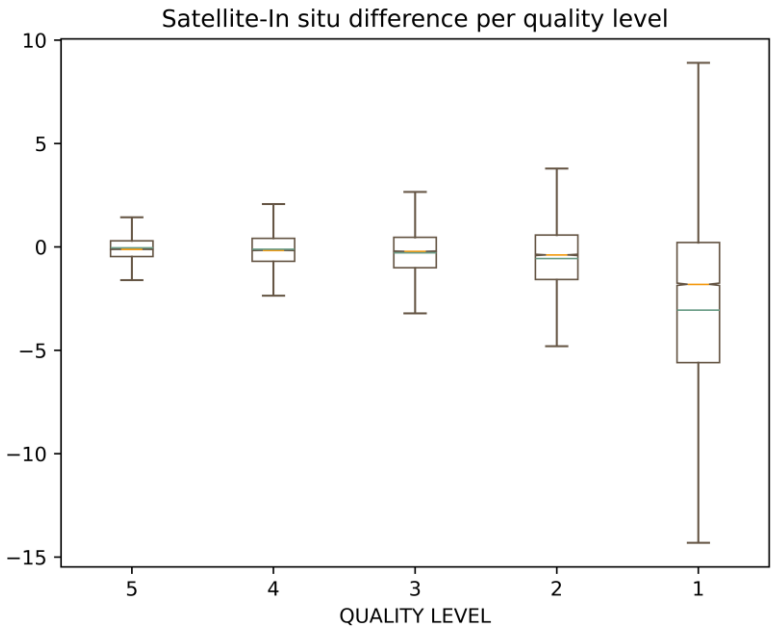


Figure 43: Distributions of the satellite minus in situ temperature difference per quality level as box plots.

The median and the robust standard deviation per quality level per year for all the lakes is shown in Figure 44 and Figure 45 together with the number of matches. For high quality levels the median and the robust standard deviation of the satellite minus in situ differences are consistently small throughout the years

when different instruments have been adopted and a different number of matches is available. They deteriorate as the quality goes lower. The number of matches for quality level 5 is consistently the highest.

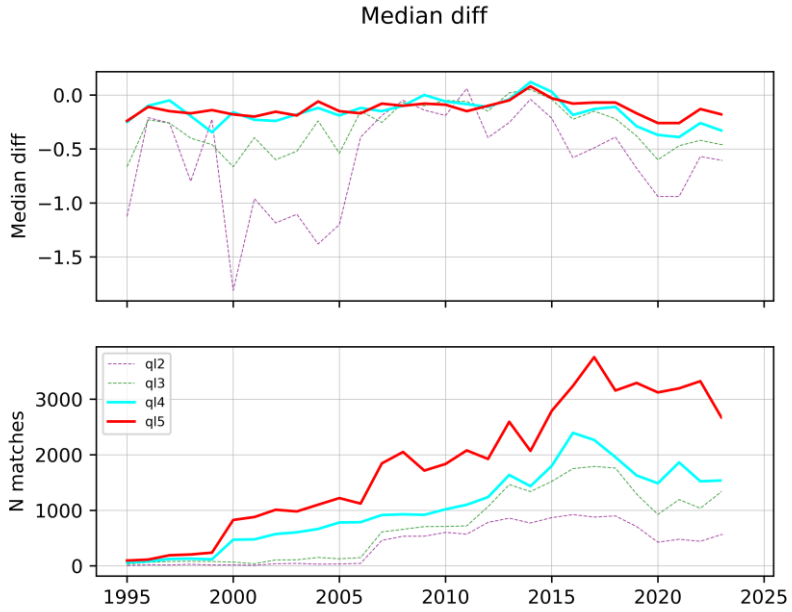


Figure 44: Satellite minus in situ temperature difference median per year (upper plot) and number of matches (lower plot) per quality level.

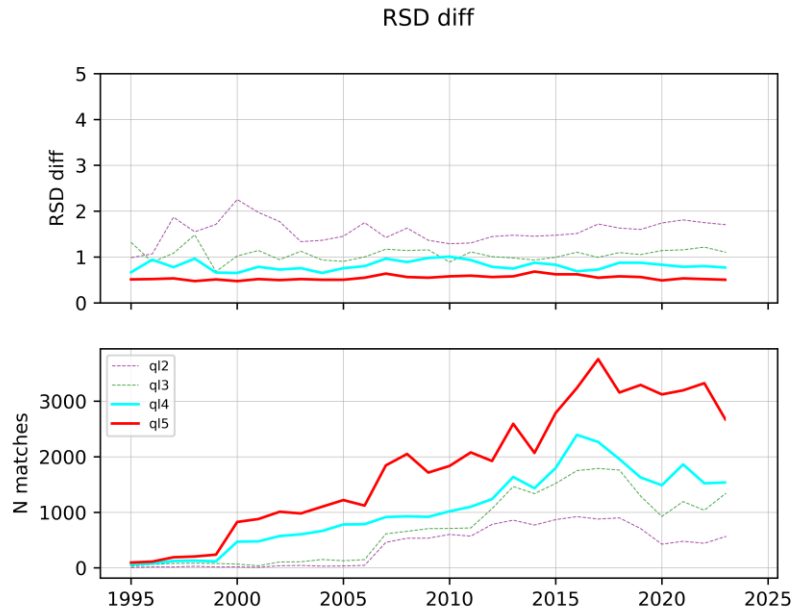


Figure 45: Satellite minus in situ temperature difference robust standard deviation per year (upper plot) and number of matches (lower plot) per quality level.

The median and robust standard deviation have been inspected also for each lake. Figure 46 and Figure 47 show the plots together with the correspondent number of matches. Higher numbers of matches are for lakes where data were available for longer periods but also where hourly/subhourly measurements were available and for sites far from the coast. The median and robust standard deviation are consistently

better for quality level 5 throughout the lakes. For both quality level 5 and 4 the instances of greater variation are related to lower numbers of matches. However, for some lakes/quality-level combinations the in situ measurements are very few: for example, for Lake Taihu (lake ID 66), only one match with LSWT of quality level 4 is available. For Lake Paijanne (lake ID 157), the median difference and the robust standard deviation are unusually large. The in situ data for this lake come from two different originators for three sites as shown in Figure 48. Figure 49 shows the satellite LSWTs, the in situ values and the climatology (for reference) for the three sites on Lake Paijanne in 2015. The majority of the measurements have been taken at the site 1 which is very close to the coast where the satellite minus in situ difference shows cooler satellite LSWT in the first part of the year and warmer in the second part. This behaviour is consistently throughout the years and is consistent with an effect of shallow-water energy balance on the in situ measurement that differences from the satellite location of the matches that are at a spatial distance up to 2 km; for comparing to measurements close to the coast in shallow water, this is a significant distance. For the other two sites the in situ and satellite measurements have a good agreement but the in situ data have a lower temporal resolution.

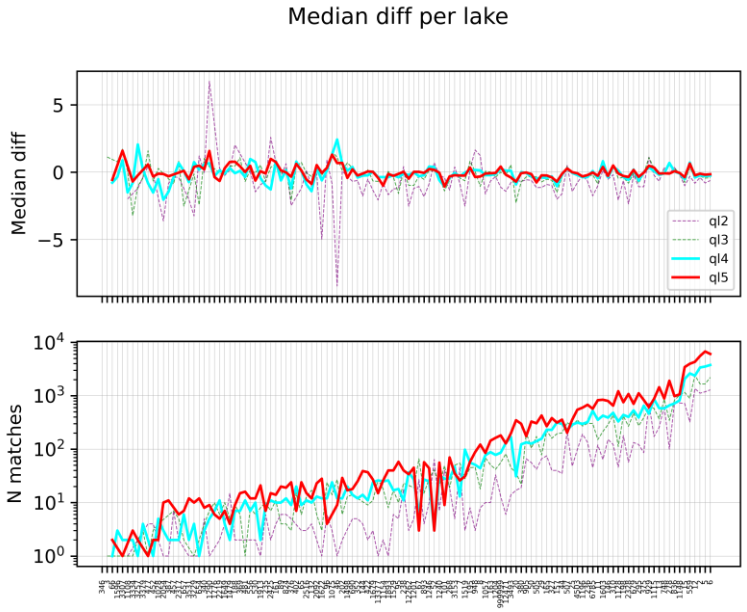


Figure 46: Satellite minus in situ temperature difference median per lake (upper plot) and number of matches (lower plot) per quality level.

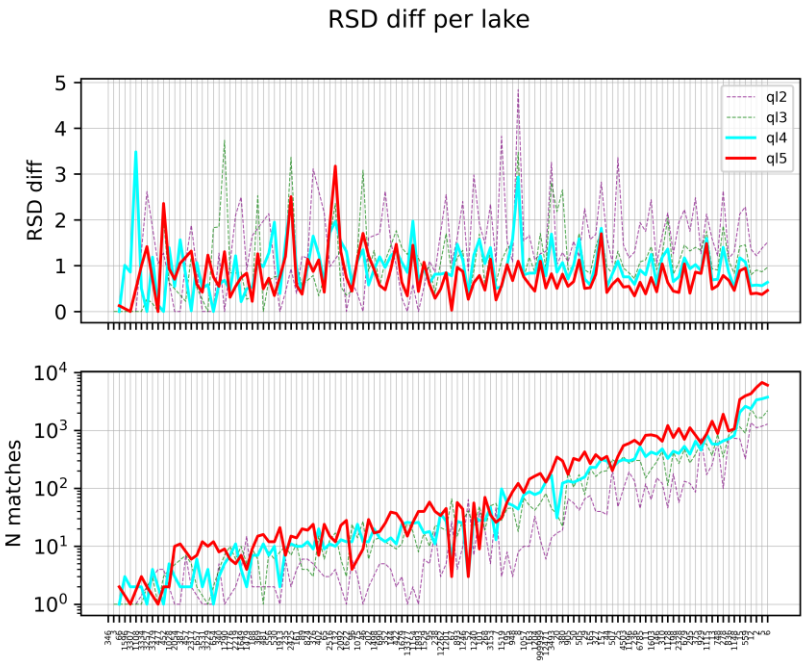


Figure 47: Satellite minus in situ temperature difference robust standard deviation per lake (upper plot) and number of matches (lower plot) per quality level.

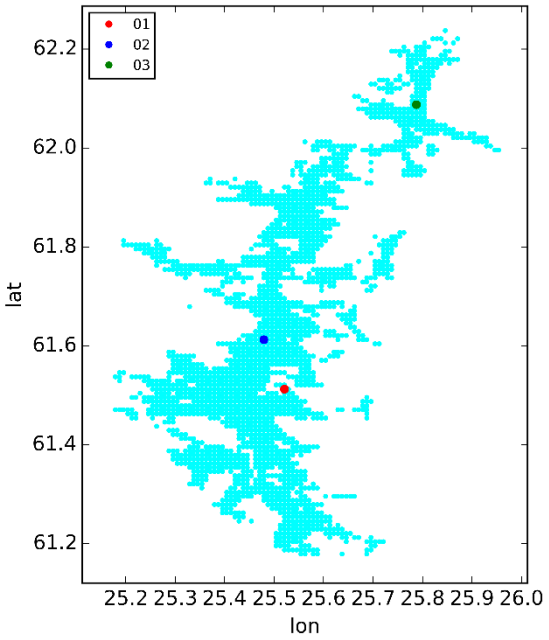


Figure 48: Locations of in situ measurements for Lake Paijanne in Finland. Each dot represents a 1/120°x1/120° cell.

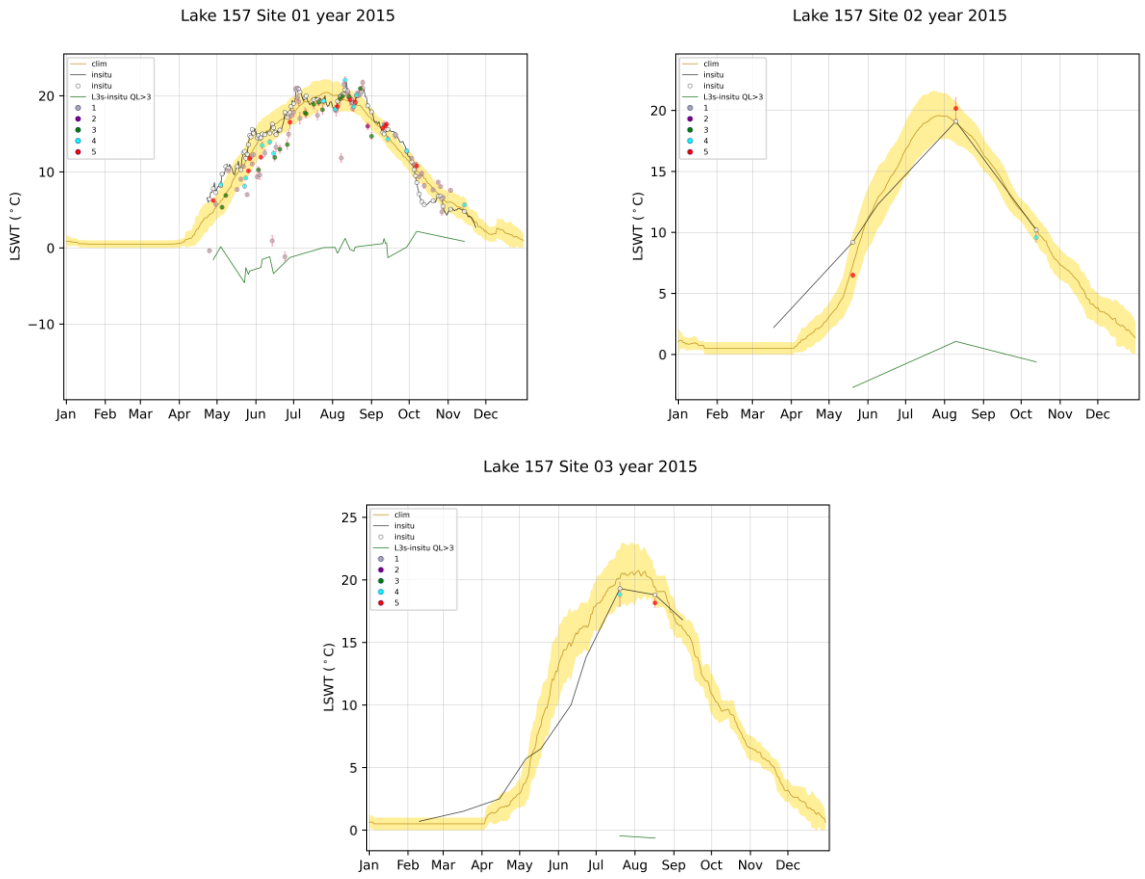


Figure 49: Validation plots for the three sites on Lake Paijanne (Finland) for 2015. T yellow plot is the climatology, the black line and white dots are the in situ measurements, the coloured dots are the satellite LSWTs per quality level and the green line is the satellite minus in situ difference.

As another example, consider Lake Superior where many sites are available. The robust statistics of the difference for all the matches of quality level 3,4,5 have been plotted per sites in Figure 50 showing consistency for sites near each other, and a higher variability of the differences for sites close to the coast than those far into the lake, as expected. Figure 51 and Figure 52 show the LSWT and the in situ measurements in 2014 respectively for three sites on the Lake Superior. In 2014 a sharp increase in temperature can be observed in the beginning of August consistently for the three sites and consistently for the satellite and in situ measurements. The timing of temperature increase is consistent at the three very different locations, one being closer to the coast and the other two more offshore. Figure 52 (right hand side) shows the position of the in situ measurement sites on Lake Superior.

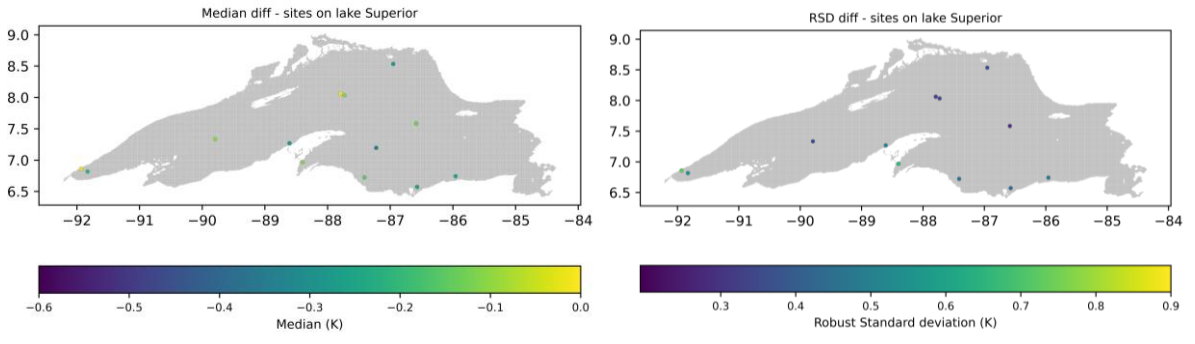


Figure 50: Satellite minus in situ temperature difference median and robust standard deviation for all the sites on Lake Superior for quality level 3,4,5.

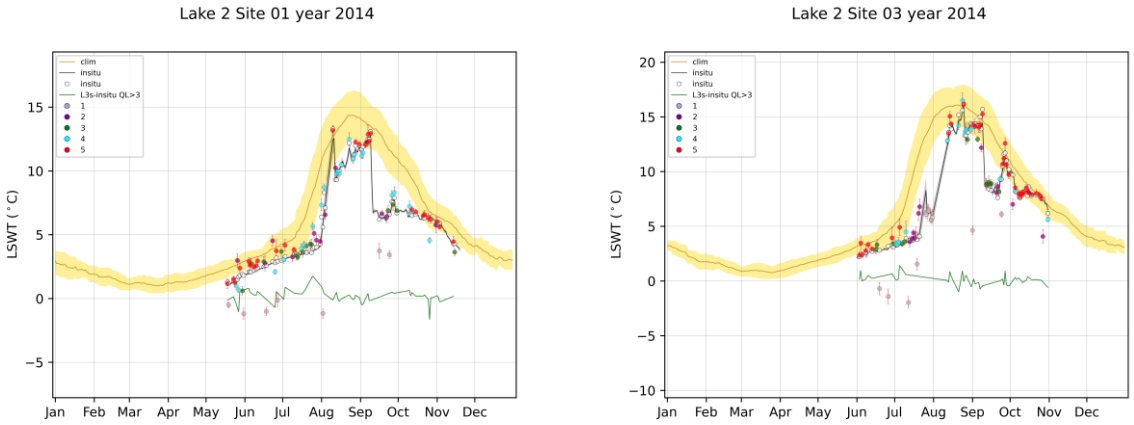


Figure 51: Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for Lake Superior in 2014 at site 01 (on the left) and site 03 (on the right).

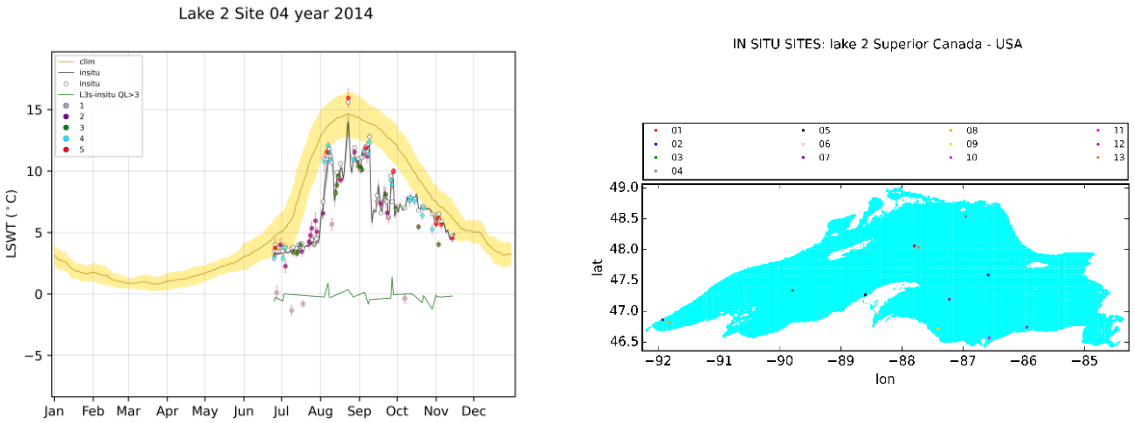


Figure 52: Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for Lake Superior in 2014 at site 04 (on the left) and location of the in situ measurement sites on Lake Superior on a 1/120° grid (on the right).

The time series of the satellite and the in situ temperature together with their difference have been inspected and they are reported here for two “difficult” validation cases as examples. The first is a small lake (Lake Erken in Sweden). The second is Lake Kasumigaura in Japan, a lake with low-temporal-frequency data.

The location where the in situ measurements have been collected on Lake Erken in Sweden is shown in Figure 53 (red dot). Figure 54 shows the satellite observations and the in situ measurements in 1995 when only ATSR2 was utilised and in 2014 when observations from AVHRR-A, AVHRR-B and MODIS were used. For both the years the satellite observations follow remarkably well the in situ measurements, which were very high frequency measurements. The peak in the difference (green line) in 1995 is very likely due to a slight temporal mismatch. Despite the peculiar behaviour of the temperatures through the year in both cases with LSWT being well beyond one standard deviation from the climatological reference, the satellite and in situ are mimicking each other remarkably well. Note that in this case the measurements site is close to the shore but matching within 2 km does not have a strong influence in this case because the lake is small and more consistent in LSWT across its area.

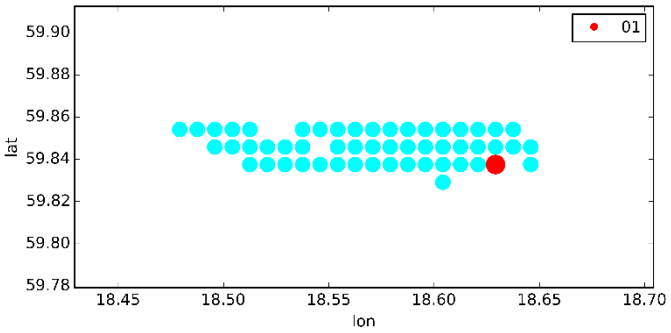


Figure 53: Location of in situ measurements for Lake Erken in Sweden. Each dot represents a 1/120°x1/120° cell.

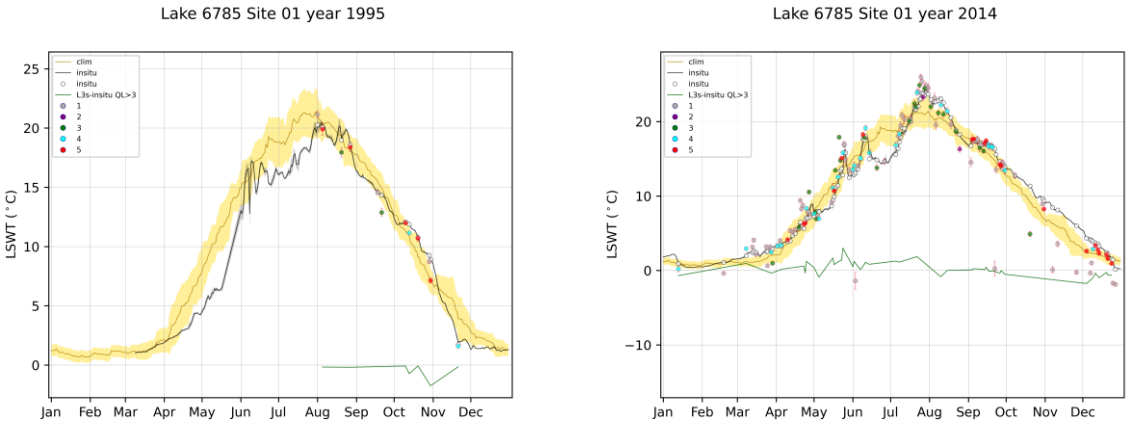


Figure 54: Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for Lake Erken in Sweden in 1995 (left) and 2014 (right).

Figure 55, Figure 56, Figure 57 present Lake Kasumigaura in Japan at the four sites in year 2016 and Figure 58 presents the site 05 in 2016 and 2020. Sites 01 and 03 are very close to the coast, so the matches are lower in number than for the other sites. However, a consistency between in situ and satellite LSWT can be observed for all the sites.

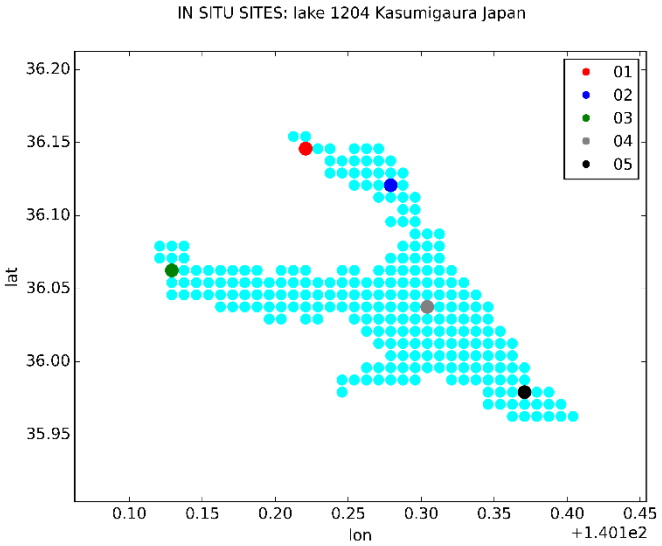


Figure 55: Location of in situ measurements for Lake Kasumigaura in Japan. Each dot represents a 1/120°x1/120° cell.

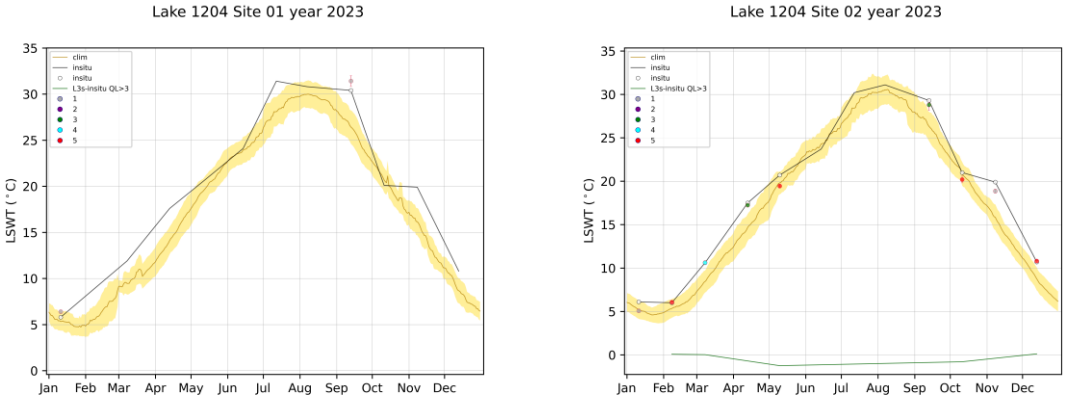


Figure 56: Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for Lake Kasumigaura in Japan in 2023 for site 01 (left), site 02 (right).

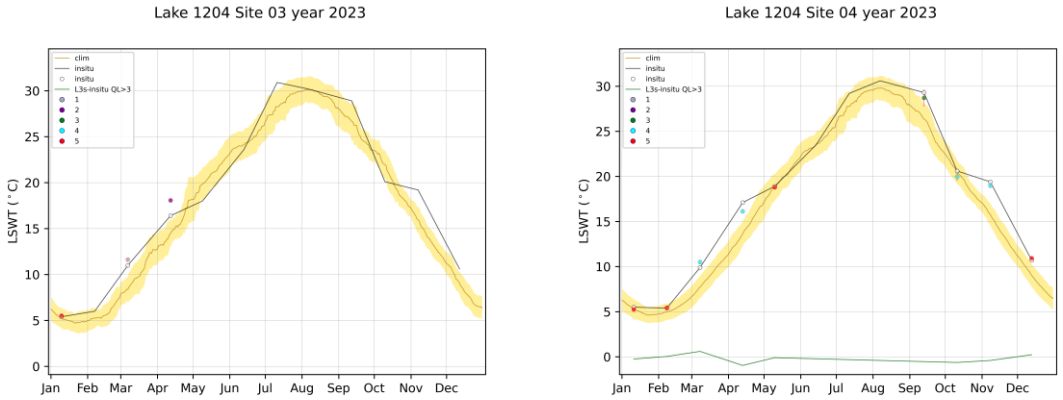


Figure 57: Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for Lake Kasumigaura in Japan in 2023 for site 03 (left), site 04 (right).

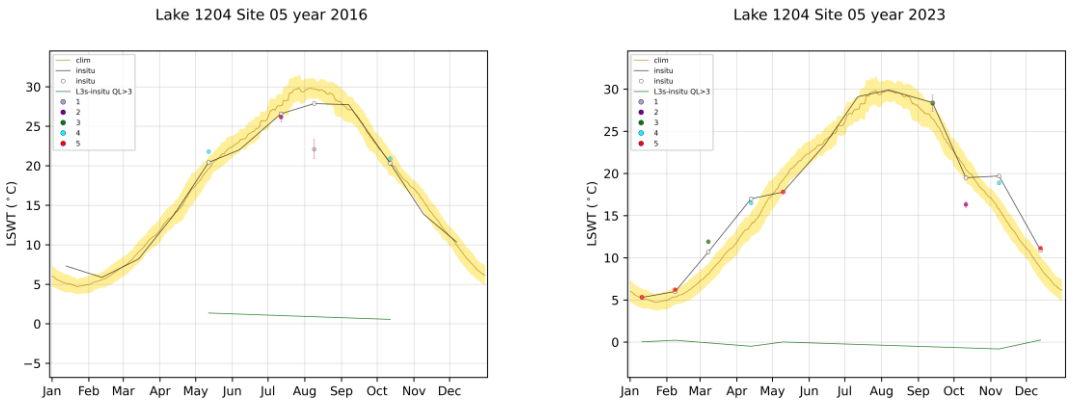


Figure 58. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for Lake Kasumigaura in Japan in 2016 and 2023 for site 05.

4.4.1 Validation of the uncertainty LSWT v4.5

The LSWT uncertainty estimate has been validated comparing the difference satellite minus in situ temperatures and the correspondent LSWT and in situ uncertainties. The following quantity is calculated for each match:

$$\Delta = \frac{T_{LSWT} - T_{INSITU}}{\sqrt{\sigma_{LSWT}^2 + \sigma_{INSITU}^2 + \sigma_{repr}^2}}$$

where *T* indicates temperature, for LSWT and in situ as indicated in the subscripts. σ means the standard deviation from measurement uncertainty (for LSWT and in situ) and from real differences because of point-to-pixel representativity effects.

The in situ measurements uncertainty is not known for the data we have. We explore two assumptions: σ_{INSITU}=0.2 K, a value based on deployment of similar measurement technologies to the ocean, and σ_{INSITU}=0.5 K which would be at the upper end of our expectations for in situ uncertainty. The representativity effect is presently unquantified and we set it to 0 K for the present; neglecting representativity has the tendency to make the LSWT uncertainty look underestimated. σ_{LSWT}^2 is context sensitive and varies from match to match, which is why the validation approach involves the calculation of the above metric: the distribution of Δ should be a Gaussian distribution with mean equal to 0 and standard deviation equal to 1 when all standard deviations are well estimated and the retrieval is unbiased relative to the in situ and any mean geophysical effect. Figure 59 shows the histograms of the uncertainties per quality level where also the fitted Gaussian and the target Gaussian distributions are shown. In Figure 59 an in situ uncertainty of 0.2 K has been assumed.

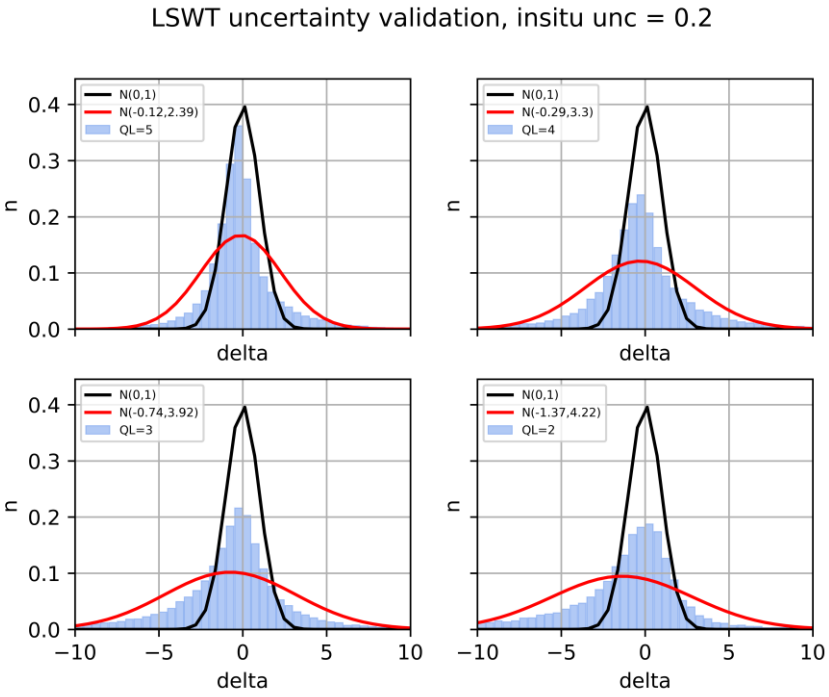


Figure 59: LSWT uncertainty validation (in situ uncertainty = 0.2K) per quality level (indicated in legend): histograms of Δ

For quality level 5, the Gaussian fit has width 2.39, which means that observed differences are more different than expected from the quoted uncertainties. This may be partly because the product uncertainties are underestimated but could also arise to the degree that lake in situ data (being more diverse) have larger uncertainty than the assumed value (based on experience of ocean observations), and because representativity is neglected. Interpretation of this outcome is therefore currently ambiguous, and research is needed to better understand the in situ uncertainty and representativity effects.

We used an in situ uncertainty of 0.5 K to explore the level of in situ uncertainty that would better fit the Gaussian. Figure 60 shows the uncertainty validation for this value. The width of the Gaussian fit for quality level 5 is much closer to one (1.5), and thus an in situ uncertainty of 0.5 K may be closer to the combined effect of in situ and representativity uncertainty. However, more investigations of in situ uncertainty and representativity need to be carried out to understand this better.

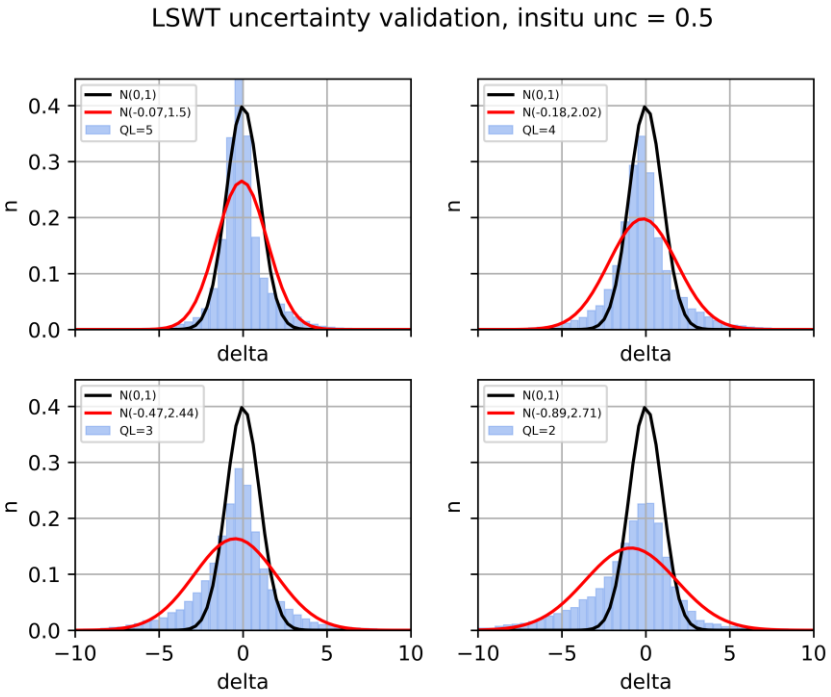


Figure 60: LSWT uncertainty validation (in situ uncertainty = 0.5K) per quality level (indicated in legend): histograms of Δ

4.4.2 Lakes without LSWT results

For 13 target lakes out of the 2025 lakes on the Lakes_cci target list, LSWT has been obtained with less than 100 days of observations of quality level>1, largely due to the fact that they are too small. The lakes are listed in Table 24 together with the estimated maximum distance to land. The majority of the lakes are not feasible because of their small size, except for lakes CGL200000071 and CGL20000072 where in the mask used during the processing they were labelled as sea rather than lakes.

Table 24. Lakes without LSWT observations

Lakes_cci ID	Name	Country	Max distance to land (km)	N days of observation	Comments
GLWD00018089	Macnean	United Kingdom	0.6	0	Too small
GLWD00164651	Portmore	United Kingdom	0.5	3	Too small
GLWD00208840	Mantua	Italy	0.5	0	Too small
GLWD00215215	Morse	USA	0.5	0	Too small
GLWD00215311	Geist	USA	0.5	0	Too small
GLWD00215339	Eagle Creek	USA	0.6	0	Too small
GLBL00000004	Mourve	United Kingdom	0.2	0	Too small
CGL200000071	Patos Lagoon	Brazil	25.2	0	Labelled as sea
CGL200000072	Maracaibo	Venezuela	46.5	0	Labelled as sea
HYLA00134644	Rihpojavri	Norway	0.7	0	Too small
GLWD00016662	Jijila	Romania	1.3	3	Too small
HYLA00016021	Ambussel	Tanzania	1.8	23	Too small
GLWD00000799	Hawizeh marshes	Iraq/Iran	1	34	Too small/dry

4.5 Conclusion and recommendations

The validation of the LSWT shows very good mean agreement, comfortably within ±0.2 K between satellite LSWT and independent in situ temperature measurements.

The uncertainty validation principles are clear, but the results are not mature enough to make strong validation statements regarding the uncertainty information since in situ uncertainty and representativity uncertainty are poorly known for lakes: experience from the ocean suggests in situ/representativity uncertainty of order 0.2 K, but for lakes this may be an underestimate. An initial analysis suggests 0.5 K is closer to the case.

The LSWT and uncertainty validation show that the quality level accompanying the LSWT are a very important variable for the proper use of the data. Quality levels provide the confidence on the LSWT retrieval. We recommend using quality levels 4 and 5 for lake-climate applications.

Manual inspection of all products for more than 2000 water bodies is impossible and, in most cases, requires local knowledge. The validation of the products will always be based on a small sample of well-studied areas. Users of these products are therefore advised to inspect the results for their area of interest before generating derivative products and any feedback to Lakes_cci would be most useful.

4.6 Acknowledgement

We would like to thank Iestyn R. Woolway for having established the contacts to set up the in situ database and all the institutions listed in Table 2 that have provided us with in situ data and in particular: Enner Alcantara (Saõ Paulo State University, Saõ Paulo, Brazil), Gil Bohrer (The Ohio State University, Columbus, USA), Jean-Francois Cretaux (LEGOS, Toulouse, France), Xavier Lazzaro, (Universidad Mayor de San Andres, Bolivia), Margaret Dix (Universidad del Valle de Guatemala, Guatemala), Hilary Dugan (Center for Limnology, University of Wisconsin-Madison, USA), Martin Dokulil (Mondsee, Austria), Gideon Gal (Yigal Allon Kinneret Limnological Laboratory, Israel Oceanographic and Limnological Research, Migdal, Israel), Claudia Giardino (Istituto per il Rilavemento Elettromagnetico dell'Ambiente, National Research Council of Italy, Italy), April James (Nipissing University, Canada), Ilga Kokorite (University of Latvia and Latvian Environmental Geology and Meteorology Centre, Latvia), Johanna Korhonen (SYKE, Helsinki, Finland), Ben Kraemer (Leibniz Institute for Freshwater Ecology and Inland Fisheries, Berlin, Germany), Alo Laas (Estonian University of Life Sciences, Tartu, Estonia), Eric Leibensperger (Center for Earth and Environmental Science, SUNY Plattsburgh, USA), Junsheng Li (Institute of Remote Sensing and Digital Earth, Chinese Academy of Science, China), Alessandro Ludovisi (Dipartimento di Biologia Cellulare e Ambientale, Universita degli Studi di Perugia, Italy), Chris MacBride (University of Waikato and the Bay of Plenty Regional Council, New Zealand), Shin-ichiro Matsuzaki (National Institute for Environmental Studies, Japan), Linda May (Centre for Ecology and Hydrology, Edinburgh, UK), Ghislaine Monet (UMR CARRTEL, Thonon le Bains, France), Tiina Nogesand Peeter Noges (Estonian University of Life Sciences, Tartu, Estonia), Sajid Pareeth (Freie Universitat Berlin/Fondazione Edmund Mach, Germany/Italy), Sebastiano Piccolroaz (Institute for Marine and Atmospheric Research, Department of Physics, Utrecht University), Don Pierson (Uppsala University, Sweden), Pierre-Denis Plisnier, Merja Pulkkanen (SYKE, Helsinki, Finland), Antti Raike (SYKE, Helsinki, Finland), Alon Rimmer (Yigal Allon Kinneret Limnological Laboratory, Israel Oceanographic and Limnological Research, Migdal, Israel), Michela Rogora (CNR Institute for Water Research (CNR IRSA), Italy), Geoffrey Schladow (UC-Davis Tahoe Environmental Research Center, USA), Eugene Silow (Irkutsk State University, Russia), Lewis Sitoki (Department of Earth Environmental Science and Technology, Technical University of Kenya, Nairobi), Evangelos Spyrakos (Biological and Environmental Science, University of Stirling, Scotland, UK), Wim Thiery (Department of Earth and Environmental Sciences, KU Leuven, Belgium), Piet Verburg (NIWA, New Zealand), Gesa Weyhenmeyer (Department of Ecology and Genetics, Uppsala University, Sweden), Caroline Wynne (Environmental Protection Agency, Ireland).

5 Lake Water Leaving Reflectance – LWLR

Lake Water-Leaving Reflectance (LWLR), also referred to as water colour, is the measurement of the quantity of sunlight reaching the remote detector after interaction with the water column. The validation and comparison of the LWLR products, including LWLR, chlorophyll-a (Chl-a) and total suspended matter is based on matchups between in situ and satellite measurements. The same dataset is used to characterise the final product uncertainty after modelling it to known data properties, such as the similarity to specific Optical Water Types.

Lake Water-Leaving Reflectance (LWLR) is the result of atmospheric correction of top-of-atmosphere radiance over water pixels. This correction is the result of model optimisation and subject to the possibility of ambiguous solutions. The main effects that introduce uncertainty are mixing of reflectance from water and nearby land in the atmosphere, bottom effects, water-column bio-optical model ambiguities and limited sensor band configurations to bound the mentioned numerical optimisation. Limited in situ reference observations and a bias favouring turbid, productive and large lakes in historical datasets, further hamper uncertainty characterisation over the full range of water types and lake geophysical and geospatial characteristics.

The problem of lacking in situ data is somewhat overcome when the end-to-end validation of biogeochemical products derived from LWLR is concerned. Here, a spatio-temporal sampling bias still exists, with most of the available in situ data having been collected since the launch of MERIS. Systematic error in the LWLR retrieval can be compensated in algorithm calibration for biogeochemical products.

MERIS and OLCI are the best-available sensors for inland optical water quality products. To fill the observation gap from April 2012 to April 2016 between MERIS and OLCI, MODIS-Aqua observations are analysed. Only a limited number of lakes show consistent LWLR retrieval between all three sensors during overlapping periods, the remainder presenting optical conditions that are beyond the MODIS-Aqua sensor and subsequent atmospheric capabilities to resolve.

CRDP v3.0.0 represents a fully revised set of algorithms used to derive LWLR and its derived variables. The new release includes additional identification of suspect observations either via masking when data should not be considered by users, or as a series of quality flags indicating potential data quality risks.

5.1 Data description

5.1.1 In situ data

The validation dataset used in this study is GLORIA (Lehmann et al. 2021) which comprises a global collection of hyperspectral lake water-leaving reflectance measurements with associated downstream biogeochemical water quality variables. In addition, data were separately collated from various community contributions including through GLEON, to validate CDOM and K_d products included in CRDP v3.0.0. For the latter, the LIMNADES repository was also used, despite it being no longer maintained.

5.1.2 Satellite data

CRDP v3.0.0 features the 4th reprocessing of ESA MERIS full resolution L1B data, NASA MODIS-Aqua L1A, and ESA/EUMETSAT OLCI Non-time-critical L1B from Sentinel-3 A and B. These were processed to LWLR by applying Polymer (v4.17b; Steinmetz et al. 2011) MODIS data were obtained from the National Aeronautics and Space Administration (NASA) for the period 2011-2018, providing overlap with MERIS and OLCI to evaluate inter-sensor bias. All L1 data were masked using a combination of L1 and Idepix flags (see ATBD for details). The masks applied were Cosmetic, duplicated, glint risk, suspect, land/ocean, bright, coastline, and invalid from the L1 product and invalid, cloud, cloud ambiguous, cloud sure, cloud buffer, cloud shadow, snow ice, bright, white, coastline, land and glint risk from Idepix.

5.2 Comparison methods

Product validation of the LWLR and derived products (Chl-a, TSM, CDOM, K_d) is based on comparison against in situ observations.

The satellite matchups were extracted from 3×3 pixel windows unless otherwise indicated, with a temporal window of ±1 day. The mean value was calculated from the macro-pixel which most closely matched the time of the observations. If the same satellite pass was found to be a valid match-up for multiple in situ observations from the same location, then the nearest in situ value in time was selected for the match-up.

Depending on which product was being validated, statistical measures of performance included the coefficient of determination (R^2), Root-Mean-Square difference (RMS), Normalized RMS difference (NRMS), Mean Absolute Percentage difference (MAP) and bias.

5.3 Description of the work

5.3.1 Lake Water-Leaving Reflectance

For water-leaving reflectance, comparisons between the in situ and satellite measurements were performed for each waveband. In this report, the atmospherically corrected fully normalised water-leaving reflectance (R_w) were converted to remote sensing reflectance R_{rs} (assuming $R_{rs} = R_w/\pi$ [sr^{-1}]) to facilitate the comparison between in situ and satellite measurements.

5.3.2 LWLR-derived products

After extensive algorithm intercomparison and tuning, algorithm sets were selected to cover all Optical Water Types specified by Spyrakos et al. (2018). The assignment of algorithms to each OWT is detailed in the ATBD, whereas the overall blending procedure and uncertainty propagation is detailed in Liu et al (2021) and the E3UB document.

5.3.3 Inter-sensor consistency evaluations

For the 1000 largest lakes, initial consistency evaluations were performed by intercomparing MODIS with MERIS for the overlapping year 2011 and with OLCI for the overlapping year 2018.

5.3.2.1 Data extraction

Data extraction of MERIS, OLCI and MODIS was based on L2 data. For each sensor on each day, to mitigate land adjacency effects, one pixel along the lake margins in MODIS and two pixels in MERIS and OLCI were excluded, reflecting the finer spatial resolution of MERIS and OLCI relative to MODIS in level 2 data (300 m vs. 1 km). Percentiles of 20, 50, and 80 were then calculated and saved for further consistency evaluations.

5.3.2.2 Statistical analysis

Three variables were selected for the evaluation including two water quality parameter (Chl-a) and one water-leaving reflectance band (polymer corrected Rw at 488/490 nm). It is noted that 488 nm is a MODIS band, and 490 nm is the corresponding MERIS or OLCI band. Based on previous inter-sensor analyses on LWLR, R_w(488-490) showed the best consistency between MODIS and MERIS or OLCI. For brevity, R_w(488) is used to refer this band for the three sensors hereafter.

For all of the large lakes, two statistic analyses were conducted on extracted daily median values of the three variables (Chl-a, TSM and R_w(488)).

- Two-sided t-test on related samples: This is a dependent t-test for the null hypothesis that 2 related or repeated samples have identical average (expected) values. If the calculated *p*-value is below the threshold chosen for statistical significance (0.001 was chosen for this study), then the null hypothesis is rejected in favour of the alternative hypothesis.
- Linear regression analysis was also performed on the daily extracted median values between sensors for each lake, correlation coefficient (R) was calculated as an indicator of the strength and direction of the linear relationship between the observations of two sensors.

Following the above analysis, lakes with $P > 0.001$ (from t-test) or $R > 0.5$ for all of the three evaluated variables were selected for a further visual check, which finally resulted in 73 lakes that show sufficient consistency. It is also noted that Turbidity is not included from MODIS due to insufficient performance of all algorithm candidates.

5.3.2.3 Climatologic filtering

A climatologic filtering procedure is used to detect and eliminate potential inconsistencies among the ECV Products . Specifically, it focused on identifying potentially unphysical combinations of LWLR, LSWT, and/or LIC.

The underlying theory guiding this filtering process was that extreme values found in Chl-a and TSM may be caused by partial or thin ice cover observed in low LSWT conditions (e.g., below freezing point) or when ice is already detected. In these situations, LWLR products showing extreme values should be considered inconsistent. As detailed in the Algorithm Theoretical Basis Document (ATBD), this climatological filtering process led to the introduction of two critical consistency-related indicators within the 'lwlr_quality_flags' band. These are: 'lwlr_poor_consistency' and 'lwlr_low_consistency'.

To evaluate the efficacy of these climatological flags, we conducted an analysis using 10th and 90th percentile data extractions. We used Level 3 daily data from both MERIS for the period 2002-2012, and OLCI for the period 2016-2022. This analysis was performed on the Great Slave Lake and Amadjuak, both in Canada, which are known to have ice cover. The data were examined both prior to and following the application of the consistency flags, to assess their efficacy.

5.4 Validation results

5.4.1 Per-sensor results

5.4.1.1 LWLR validation

Previous round-robin comparisons of atmospheric correction algorithms for MERIS carried out in GloboLakes showed that Polymer yielded the statistically most robust retrieval of reflectances, although a systematic negative bias was observed. This led to Polymer being adopted as the state-of-the-art atmospheric correction processor for Lakes_cci to provide LWLR. The Rrs validations for MERIS, OLCI, and MODIS shown below are all based on the GLORIA in situ dataset.

For MERIS, validation against in situ Rrs across 14 bands (412–885 nm) showed generally good agreement in the visible domain, with correlations peaking at 620 nm ($R = 0.75$) and remaining high between 560 and 709 nm ($R > 0.65$) (Figure 61). Performance was weaker in the blue (<490 nm) and near-infrared (>760 nm) bands, where correlations dropped below 0.4 and errors increased. Importantly, no consistent systematic bias was observed, confirming the improved reliability of POLYMER-corrected MERIS Rrs compared with previous versions.

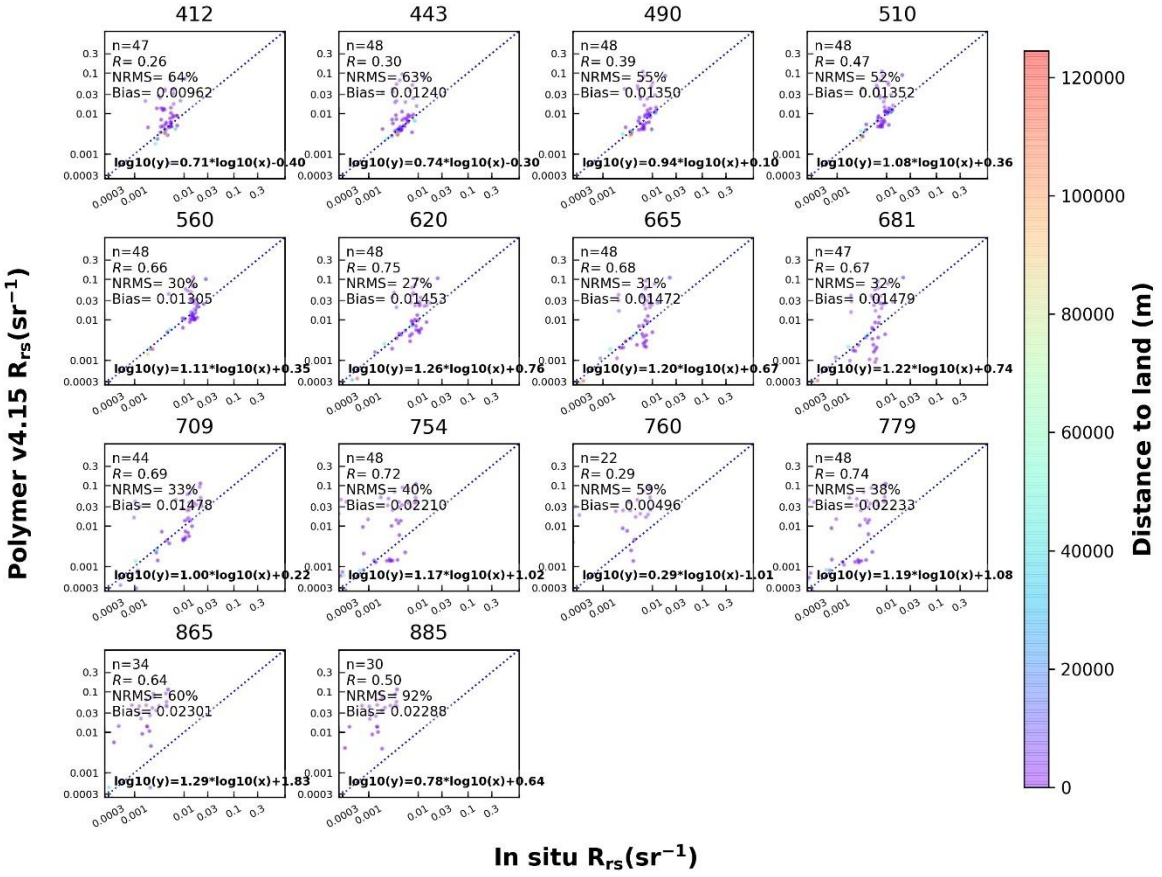


Figure 61: Comparison between in situ and MERIS R_{rs} in each band

The performance of POLYMER-corrected OLCI Rrs was evaluated against in situ measurements at 13 wavebands from 400 to 865 nm, with the number of matchups ranging from 53 at 865 nm to 72 at 779 nm (Figure 62). Overall, significant but moderate linear relationships were found across the visible spectrum. The highest agreement observed in the red band around 620 nm (R = 0.56, NRMSE = 33%). Shortwave (400–443 nm) and near-infrared bands (>754 nm) show weaker performance, with lower correlations and higher NRMSE values. POLYMER-corrected Rrs retrievals show robust performance in the visible domain, particularly between 510 and 709 nm, while performance degrades in the blue and near-infrared regions.

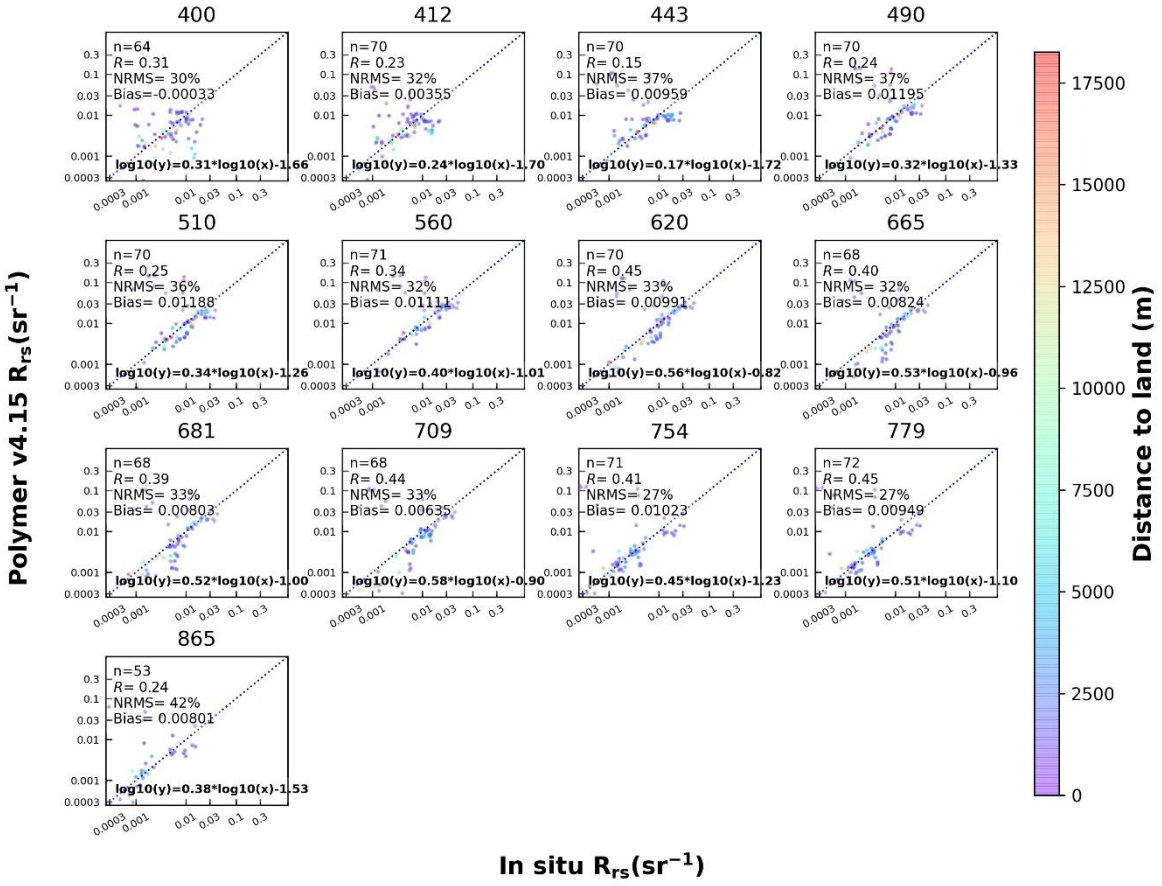


Figure 62 Comparison between in situ and OLCI R_{rs} in each band

For MODIS, validation against in situ Rrs across 12 spectral bands (412–869 nm) showed strong consistency in the visible range, with the highest correlation at 531 nm (R = 0.75) and similarly high values at 547, 555, and 667 nm (R ≈ 0.70–0.74) (Figure 63). Correlations in the blue (412–469 nm) and red (645–678 nm) bands were moderate, while performance degraded in the near-infrared (>748 nm). The number of matchups was substantially higher than for MERIS and OLCI (>1000 for most bands), reflecting MODIS's longer operation period and higher revisit frequency. Biases across all bands were generally negligible, and NRMSE values remained relatively low (12–20%) in the visible domain, further supporting the robustness of POLYMER-corrected MODIS Rrs.

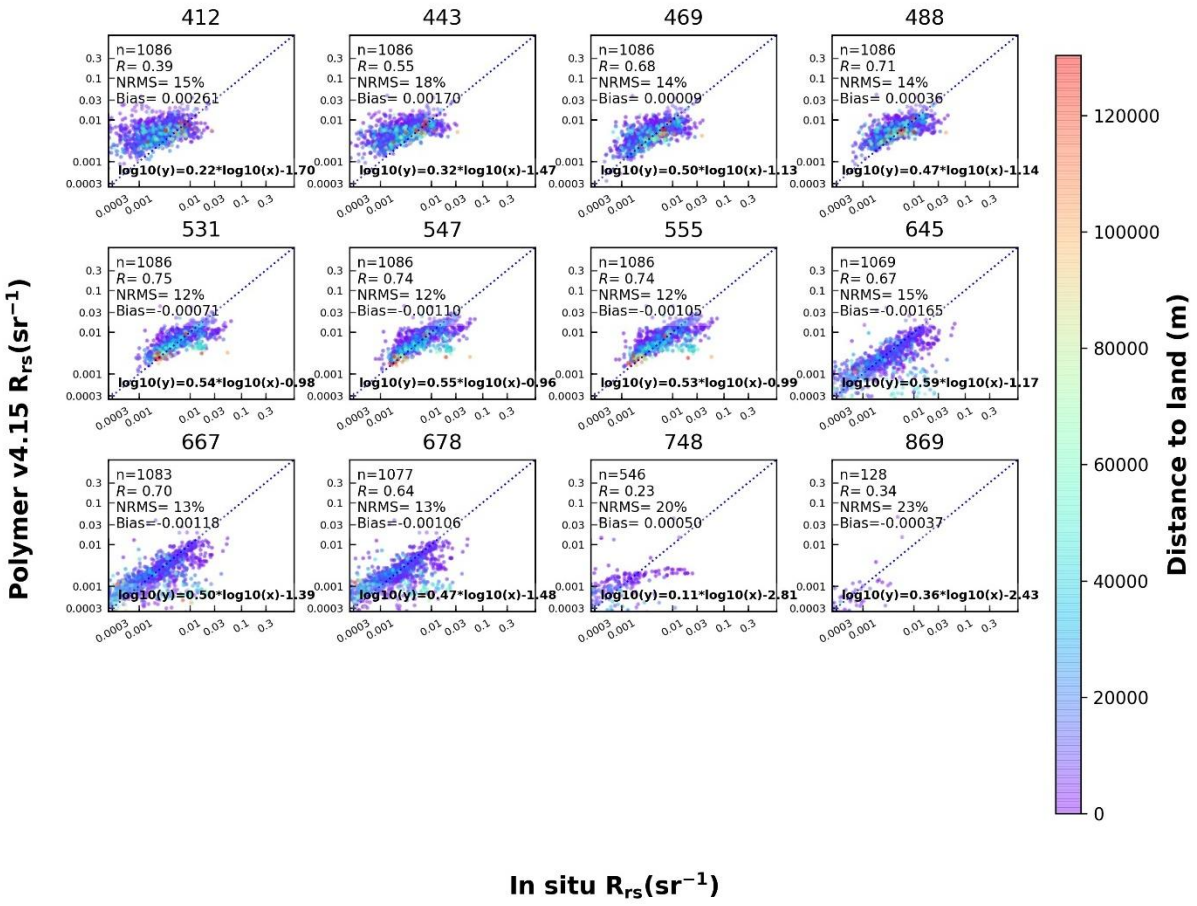


Figure 63 Comparison between in situ and MODIS R_{rs} in each band

Figure 64 presents spectral comparisons of Rrs between in situ observations and POLYMER-corrected satellite retrievals for MERIS, OLCI, and MODIS. Each panel shows the median spectra along with the 30th–70th and 20th–80th percentiles.

For all three sensors, the satellite-derived Rrs generally follow the shape and magnitude of the in situ spectra across the visible to near-infrared range. The agreement is strong in the visible domain (approximately 443–709 nm), particularly around the green (~560 nm) and red (~620–665 nm) bands, where the central tendencies of the satellite and in situ spectra overlap closely. MERIS and OLCI show relatively consistent alignment, though MERIS exhibits more variability at longer wavelengths (>700 nm). MODIS displays good overall correspondence, with the median spectra from both sources nearly coinciding, though variability is lower than for the other two sensors.

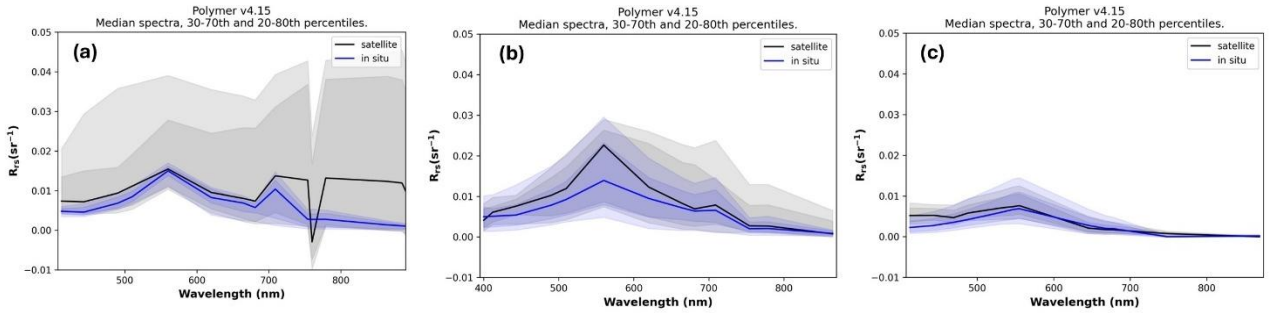


Figure 64 Spectral comparison of in situ and satellite observed R_{rs} for (a) MERIS, (b) OLCI and (c) MODIS

5.4.1.2 Chlorophyll-a validation

Figure 65 illustrates the validation of Ch-la retrievals derived from the top-3 OWT blending approach for MERIS, OLCI, and MODIS. Across the three sensors, the blended algorithm showed strong agreement with in situ Chl-a, with correlation coefficients of R = 0.86 for MERIS, 0.87 for OLCI, and 0.79 for MODIS. Errors were relatively low, with NRMSE values of 9.5%, 9.1%, and 9.3% for MERIS, OLCI, and MODIS, respectively. MERIS and OLCI exhibited the strongest performance, characterized by low bias and slopes close to unity, while MODIS, despite slightly higher errors, benefited from a substantially larger matchup dataset (n = 1907). These results demonstrate that the OWT-blended approach provides consistent and reliable Chl-a estimates across multiple sensors.

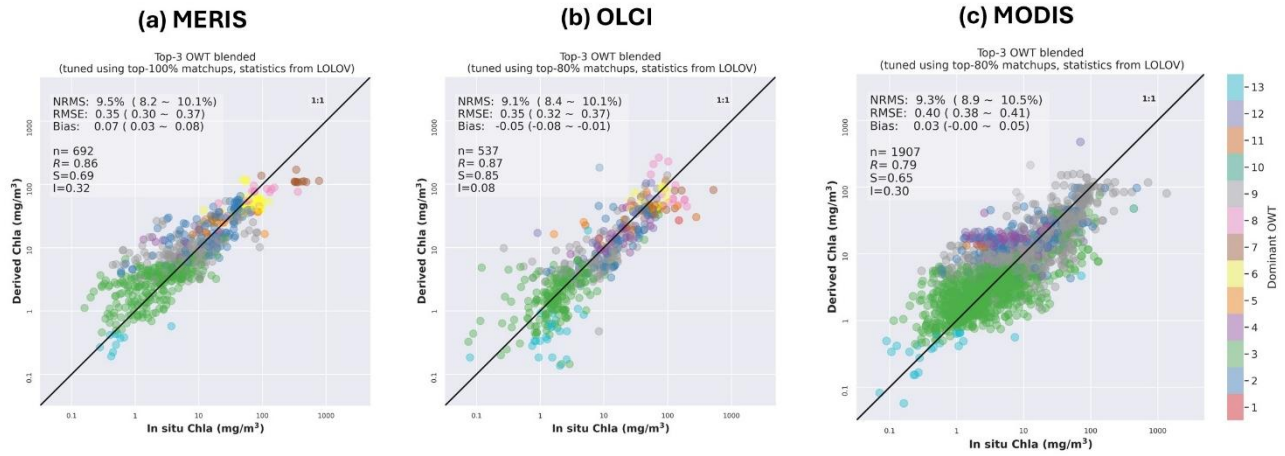


Figure 65 Comparison between in situ and top-3 blended chlorophyll-a derived from (a) MERIS, (b) OLCI and (c) MODIS.

5.4.1.3 Total Suspended Matter validation

Figure 66 presents the validation of TSM derived from the top-3 OWT blending approach for MERIS, OLCI, and MODIS. For MERIS, 568 matchups were obtained, yielding strong agreement with in situ TSM (R = 0.77) and low errors (NRMSE = 10.5%, RMSE ≈ 0.42 g m^{-3}). OLCI showed moderate performance, with 417 matchups, a correlation of R = 0.77, and slightly higher errors (NRMSE = 12.8%). MODIS, supported by the largest matchup dataset (n = 1482), achieved R = 0.69 with an NRMSE of 12.3%, reflecting larger

variability but negligible systematic bias. Collectively, these results indicate that the blended OWT approach provides reliable TSM retrievals across all three sensors,

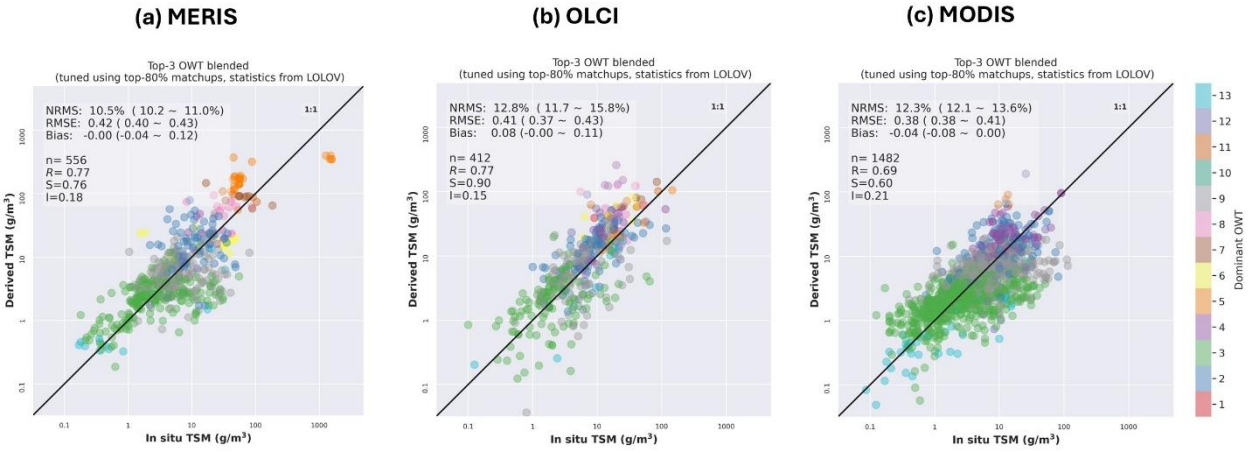


Figure 66 Comparison between in situ and top-3 blended Total Suspended Matter derived from (a) MERIS, (b) OLCI and (c) MODIS.

5.4.1.4 CDOM absorption coefficients at 440nm validation

Figure 67 shows the validation of CDOM absorption coefficients at 440 nm ($a_{CDOM}(440)$) retrieved using the top-3 OWT blending approach for MERIS, OLCI, and MODIS. MERIS yielded strong agreement with in situ observations (R = 0.73) with relatively low errors (RMSE = 0.32 m^{-1}; NRMSE = 13.4%) and minimal bias. OLCI produced comparable correlations (R = 0.71), but with larger uncertainties (RMSE = 0.42 m^{-1}; NRMSE = 15.1%) and a tendency to overestimate $a_{CDOM}(440)$ at lower values. MODIS demonstrated the most robust performance (R = 0.77), combining low error metrics (RMSE = 0.35 m^{-1}; NRMSE = 11.8%) with negligible bias.

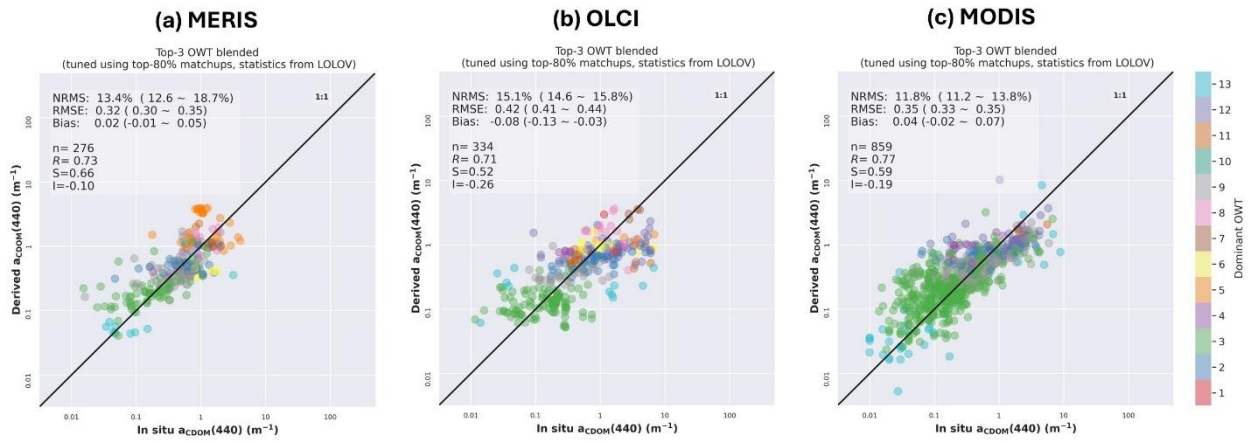


Figure 67 Comparison between in situ and top-3 blended CDOM absorption coefficients at 440nm derived from (a) MERIS, (b) OLCI and (c) MODIS.

5.4.1.5 Diffuse attenuation coefficient of downwelling irradiance validation

Figure 68 shows the validation of the diffuse attenuation coefficient of downwelling irradiance (K_d) at 490 nm, 560 nm, and 665 nm from MERIS and OLCI. We conducted the validation for MERIS and OLCI collectively because the similarity in bands between the two sensors and the lack of in situ data for each

OWT. Overall, the calculated K_d from MERIS and OLCI using the blending approach agrees well with in situ data at all three bands across clear to turbid waters, with a median absolute percentage error (MAPE) of 20.32%, 16.76% and 14.0%, respectively. A few outliers were observed in the K_d results estimated using the blended algorithms, but we found these are mainly associated with the quality of the reflectance after atmospheric correction.

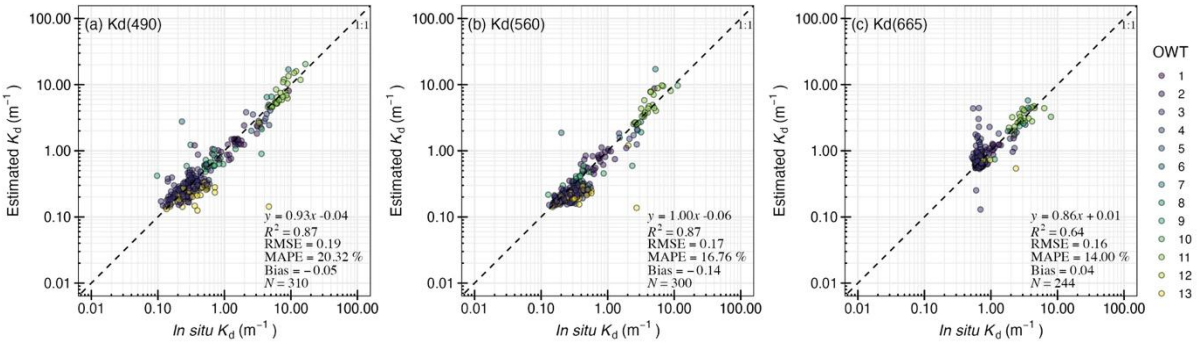


Figure 68 Comparison between in situ and top-3 blended diffuse attenuation coefficients of downwelling irradiance at (a) 490 nm, (b) 560 nm and (c) 665 nm from MERIS and OLCI.

Figure 69shows the validation of calculated K_d at 488 nm, 555 nm, and 667 nm from MODIS using the blending approach. The estimated K_d at all three bands agree well with in situ data with MAPEs of 23.19%, 20.82%, 16.13%, respectively. Compared to the validation for MERIS and OLCI, the results from MODIS shows slightly higher uncertainties. Particularly, some underestimations were observed for K_d at 555 nm although the R^2 is 0.88 and slope is 0.92 indicating the trend from clear to turbid waters is well captured.

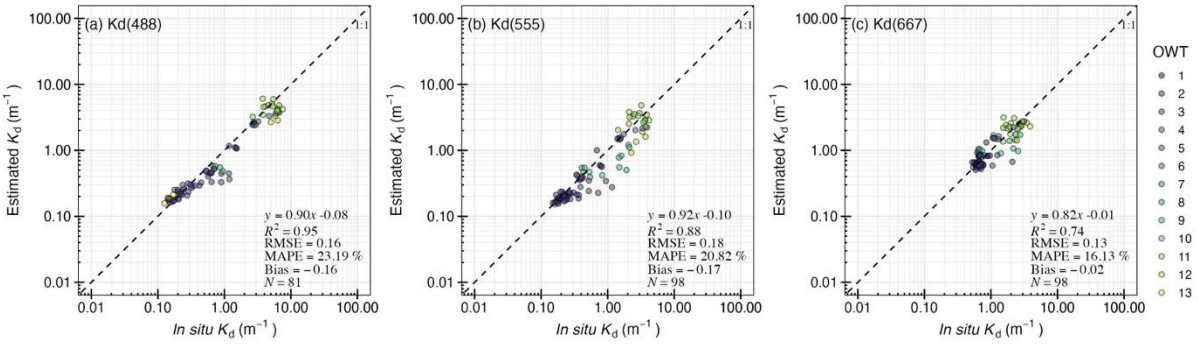


Figure 69 Comparison between in situ and top-3 blended diffuse attenuation coefficients of downwelling irradiance at (a) 488 nm, (b) 555 nm and (c) 667 nm from MODIS.

5.4.2 Inter-sensor consistency evaluation

Following statistical analysis of Chl-a, TSM and LWLR (at 488nm), 73 lakes out of the largest 1000 lakes in the Lakes_cci were identified as showing sufficient consistency between MERIS/OLCI and MODIS (see previous section 5.3.2.2). Although provided in the CRDP v3.0.0, acdom(440) was not included in the inter-sensor consistency evaluation since rare lakes show adequent inter-sensor consistency across the 3 sensors. The lakes included in the CRDP v3.0.0 for MODIS to fill the gap between MERIS (April 2012) and OLCI (April 2016) are listed in Table 25.

Table 25 List of lakes provided with MODIS observations (2012-2016)

Region name	Lake name	Continents	Area(km²)	Latitude	Longitude
GLWD00000012	Erie	North America	25938	42.144	-81.238
GLWD00000015	Ontario	North America	19842	43.826	-77.635
GLWD00000020	Titicaca	South America	7753	-15.916	-69.303
GLWD00000040	Zaysan	Asia	4334	48.717	83.432
GLWD00000041	Qinghai	Asia	4250	36.891	100.047
GLWD00000051	Van	Asia	3594	38.656	42.807
GLWD00000053	Uvs	Asia;Europe	3630	50.338	92.762
GLWD00000058	Alakol	Asia	2970	46.119	81.685
GLWD00000064	Boeng Tonle Chhma	Asia	2526	12.830	104.070
GLWD00000103	Williston	North America	1675	56.049	-124.075
GLWD00000121	Khyargas	Asia	1400	49.145	93.449
GLWD00000131	Qapshaghay Bogeni	Asia	1279	43.830	77.635
GLWD00000135	Sevan	Asia	1246	40.394	45.349
GLWD00000143		North America	1354	67.873	-97.719
GLWD00000146	Saint Clair	North America	1235	42.494	-82.703
GLWD00000179	Zhari Namco	Asia	985	30.908	85.607
GLWD00000181	Bosten	Asia	1162	41.983	87.064
GLWD00000215	Tangra	Asia	837	31.060	86.576
GLWD00000239	Ulungar	Asia	865	47.219	87.212
GLWD00000241	Sarykamyshskoye	Asia	3853	41.961	57.360
GLWD00000244	Dead Sea	Asia	644	31.534	35.478
GLWD00000249	Hirakud	Asia	616	21.658	83.734
GLWD00000250	Manych oudilo	Europe	753	46.274	42.890
GLWD00000253	Guillaume-Delisle	North America	700	56.323	-76.303
GLWD00000296	Chardarinskoye	Asia	387	41.138	68.122
GLWD00000310	Balaton	Europe	584	46.883	17.846
GLWD00000312	Ayakkum	Asia	738	37.543	89.399
GLWD00000327	Geneva	Europe	581	46.362	6.394
GLWD00000352	Bodensee	Europe	476	47.648	9.344
GLWD00000375	Flathead-Lake	North America	498	47.884	-114.125
GLWD00000382	Qyaring-Lake	Asia	481	31.135	88.355
GLWD00000383	Taro-Lake	Asia	479	31.132	84.122
GLWD00000390	Egirdir	Asia	463	38.060	30.894
GLWD00000396	Mcalpine-Lane	North America	475	66.515	-102.646
GLWD00000403	Markakol	Asia	456	48.741	85.735
GLWD00000409	Primrose	North America	435	54.892	-109.744
GLWD00000414	Cienaga-Grande-De-Santa-Marta	South America	741	10.863	-74.462
GLWD00000437	Mica	North America	401	52.147	-118.435
GLWD00000447	Kluane	North America	415	61.315	-138.783
GLWD00000469	Netsilik	North America	413	69.248	-93.116
GLWD00000488	Scutari	Europe	364	42.193	19.296
GLWD00000511	Aqqikkol	Asia	374	37.069	88.428
GLWD00000520	Chibzhang-Lake	Asia	485	33.490	90.362
GLWD00000536	Bangong	Asia	354	33.722	79.341
GLWD00000545	Cormorant	North America	349	54.222	-100.910

GLWD00000546	Pend-Oreille	North America	380	48.134	-116.375
GLWD00000549	Dorgon	Asia	346	47.747	93.408
GLWD00000554	Clark	North America	338	60.233	-154.299
GLWD00000609	Hoh Xil	Asia	306	35.583	91.092
GLWD00000649	Puma-Yumco	Asia	291	28.560	90.401
GLWD00000690	Prespa	Europe	265	40.885	21.035
GLWD00000718	Gozha	Asia	259	35.033	81.087
GLWD00000744	Ta-tse	Asia	249	31.901	87.540
GLWD00000812	Lixi'Oidain	Asia	230	35.753	90.174
GLWD00000817	Great Bitter	Africa	218	30.297	32.392
GLWD00000892	Fuxian	Asia	215	24.497	102.886
GLWD00000893	Neuchatel	Europe	216	46.898	6.843
GLWD00000913	Hsu-ju	Asia	209	30.277	86.409
GLWD00000982	Sangiyn Dalay	Asia	184	49.222	99.108
GLWD00000983	Rinqin-Xub-Lake	Asia	187	31.276	83.467
GLWD00001066	Bam	Asia	236	31.258	90.587
GLWD00001076	Dukan	Asia	116	36.103	44.925
GLWD00001079	Cuona	Asia	192	32.026	91.474
GLWD00001110	Hirfanli	Asia	220	39.145	33.673
GLWD00001128	Walker	North America	142	38.698	-118.717
GLWD00001130	Toson	Asia	137	37.145	96.942
GLWD00001192	Kyebxang-Lake	Asia	155	32.459	89.985
GLWD00001196	Sea of Galilee	Asia	164	32.799	35.589
GLWD00001354	Memar	Asia	139	34.210	82.315
GLWD00001529	Trasimeno	Europe	121	43.137	12.103
GLWD00001534	Dawa	Asia	118	31.237	84.965
GLWD00001627	Azuei	North America	118	18.572	-72.001
HYLA00001704	Hawea	Oceania	148	-44.445	169.318

5.4.2.1 Inter-sensor consistency in value distributions

5.4.2.1.1 Chlorophyll-a distributions

An inter-sensor comparison of same-day Chl-a distributions in Lake Titicaca, the largest freshwater lake in South America, from MODIS, MERIS, and OLCI was shown in Figure 70. On 29 September 2005, MODIS and MERIS both captured similar spatial patterns of Chl-a, although the median concentration from MODIS (1.46 mg m^{-3}) was higher than that from MERIS (0.92 mg m^{-3}). On 9 November 2018, MODIS and OLCI also showed consistent spatial distribution features, but with different magnitudes, as indicated by median values of 0.76 and 0.34 mg m^{-3}, respectively.

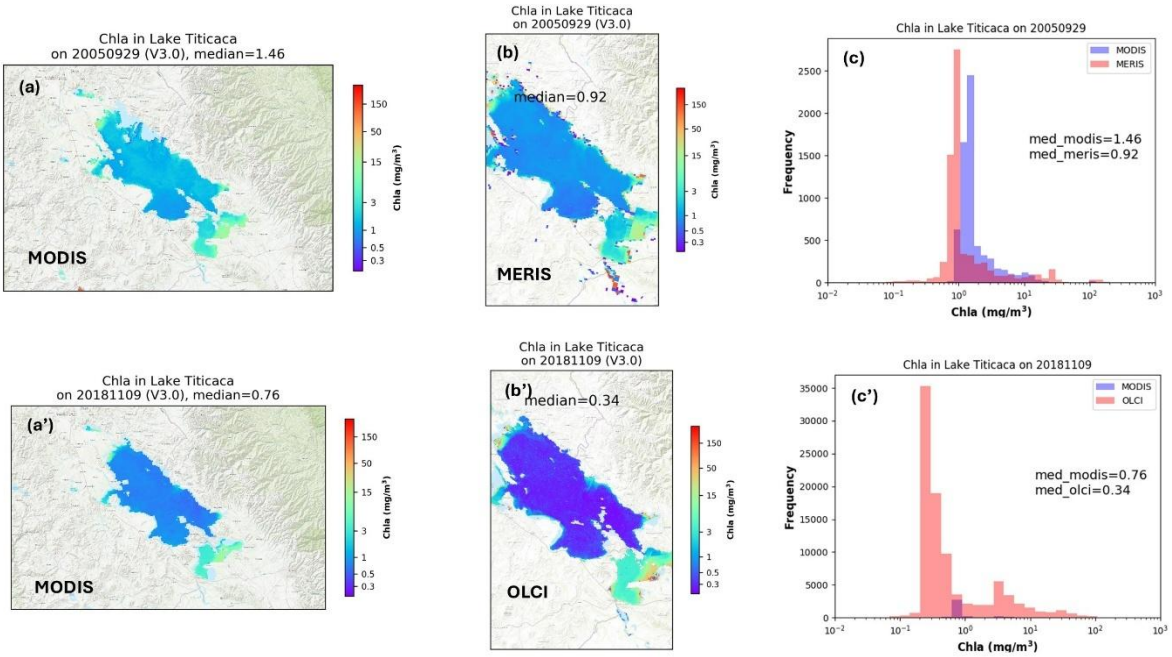


Figure 70 Inter-sensor comparison of same-day chlorophyll-a distributions in Lake Titicaca: (a–c) MODIS versus MERIS on 29 September 2005, and (a′–c′) MODIS versus OLCI on 9 November 2018.

Figure 71 presents an inter-sensor comparison of same-day Chl-a distributions in Lake Vänern from MODIS, MERIS, and OLCI. On 23 July 2008, MODIS and MERIS both captured highly similar spatial patterns of Chl-a, with median concentrations of 5.79 mg m^{-3} and 6.87 mg m^{-3}, respectively. On 21 July 2018, MODIS and OLCI likewise showed consistent spatial distribution features, although the median concentration from MODIS (4.82 mg m^{-3}) was higher than that from OLCI (3.19 mg m^{-3}). Despite these differences in magnitude, the overall spatial distribution patterns were comparable across sensors, demonstrating good inter-sensor consistency in mapping lake-wide Chl-a.

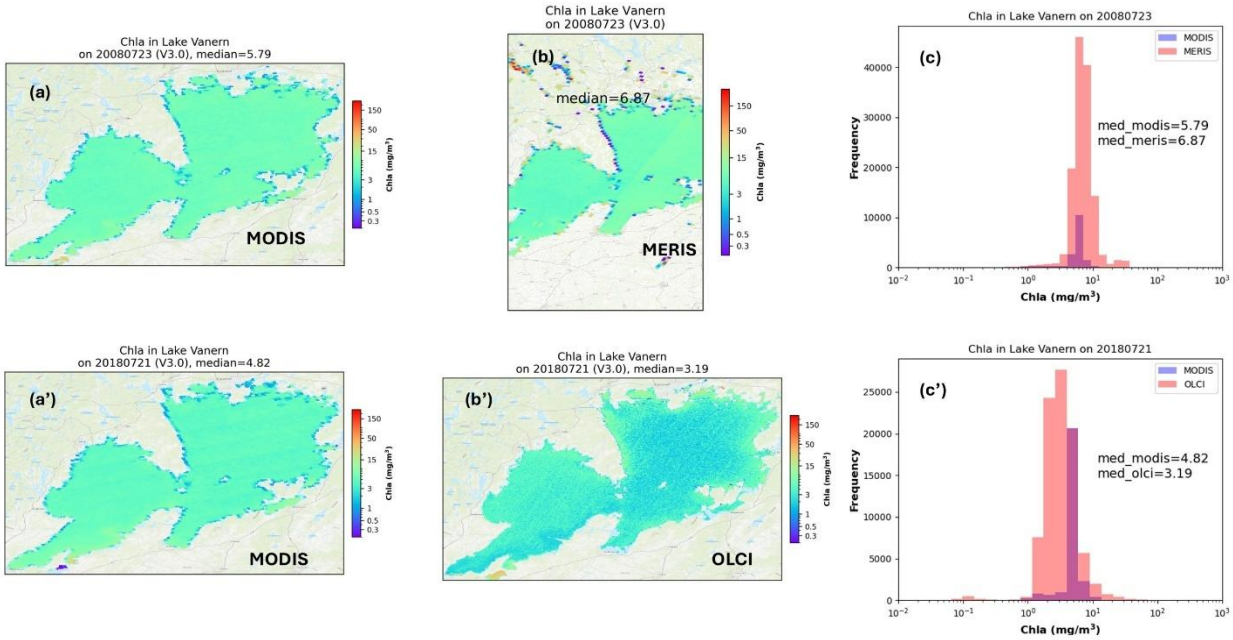


Figure 71 Inter-sensor comparison of same-day chlorophyll-a distributions in Lake Vanern: (a–c) MODIS versus MERIS on 23 July 2008, and (a′–c′) MODIS versus OLCI on 21 July 2018.

5.4.2.1.2 Total Suspended Matter distributions

Figure 72 shows same-day TSM retrievals for Lake Titicaca, enabling inter-sensor comparison between MODIS, MERIS, and OLCI. On 29 September 2005, MODIS and MERIS both showed nearly identical spatial patterns of TSM, with closely matching median concentrations of 0.28 g m^{-3} and 0.25 g m^{-3}, respectively. On 9 November 2018, MODIS and OLCI likewise produced very similar distribution patterns, again accompanied by comparable median values of 0.20 g m^{-3} and 0.25 g m^{-3}. These results demonstrate that both the spatial distribution and the central tendency of TSM retrievals are consistent across sensors, confirming strong inter-sensor agreement in lake-wide TSM estimates.

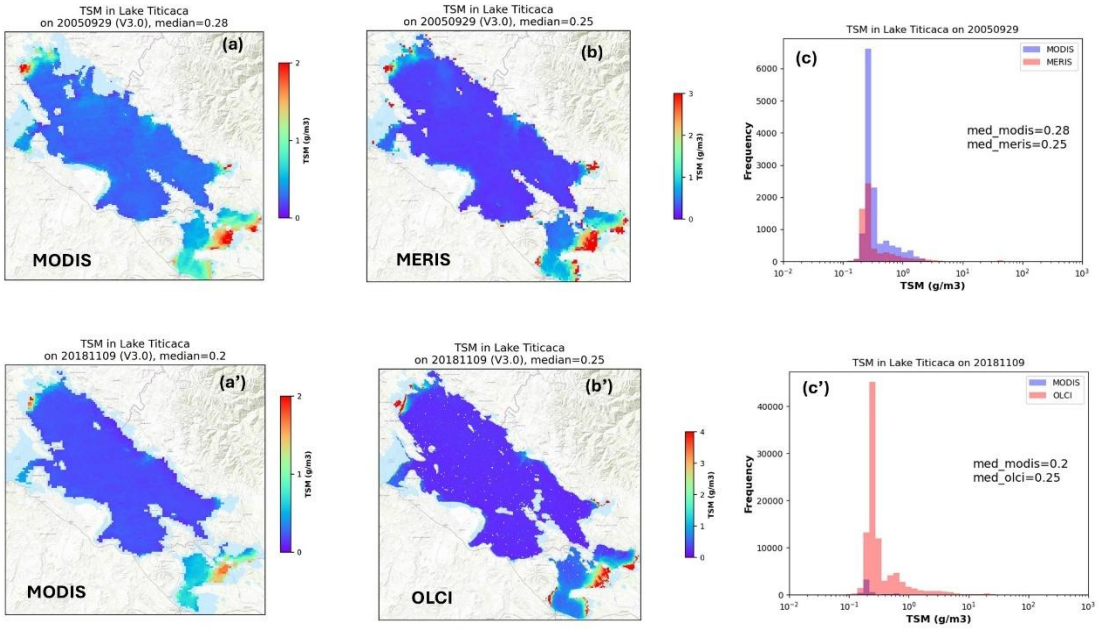


Figure 72 Inter-sensor comparison of same-day Total Suspended Matter distributions in Lake Titicaca: (a–c) MODIS versus MERIS on 29 September 2005, and (a′–c′) MODIS versus OLCI on 9 November 2018.

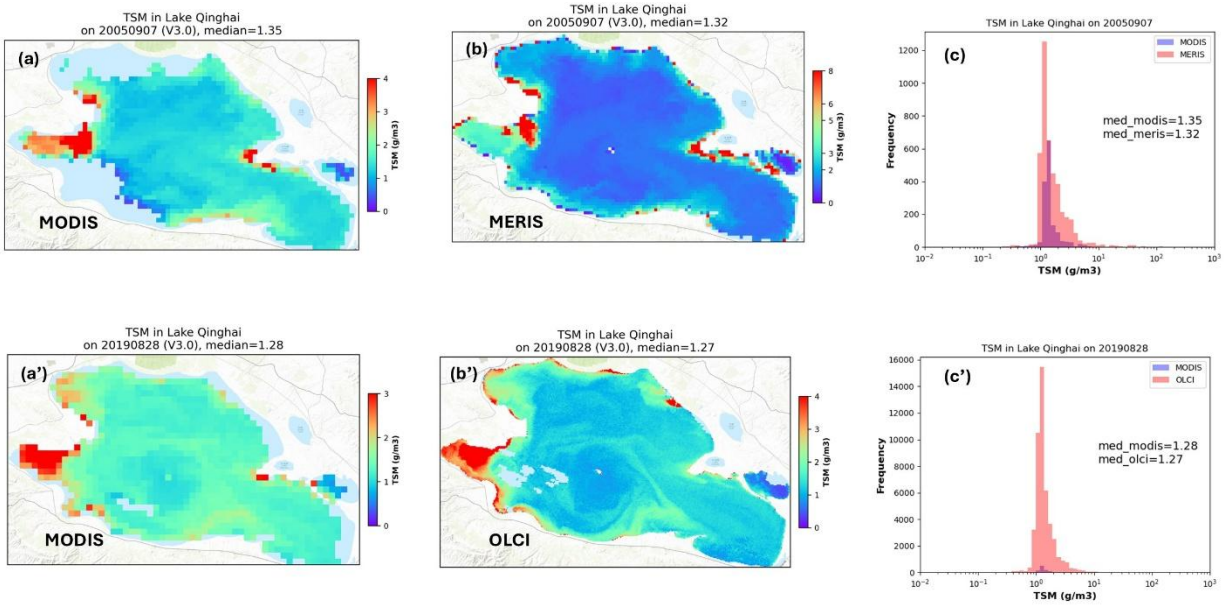


Figure 73 Inter-sensor comparison of same-day Total Suspended Matter distributions in Lake Qinghai: (a–c) MODIS versus MERIS on 7 September 2005, and (a′–c′) MODIS versus OLCI on 28 August 2018.

Figure 73 shows the inter-sensor comparison of same-day TSM distributions in Lake Qinghai derived from MODIS, MERIS, and OLCI. On 7 September 2005, the spatial patterns retrieved by MODIS and MERIS were highly consistent, with nearly identical median concentrations of 1.35 g m^{-3} and 1.32 g m^{-3}. On 28 August 2019, MODIS and OLCI likewise exhibited very similar spatial distribution features, accompanied by almost the same median values of 1.28 g m^{-3} and 1.27 g m^{-3}. The close agreement in both spatial patterns and median values highlights the strong inter-sensor consistency of TSM retrievals across different sensors.

5.4.2.1.3 CDOM absorption coefficients at 440nm distributions

Figure 74 presents an inter-sensor comparison of same-day $a_{CDOM}(440)$ distributions in Lake Razelm from MODIS, MERIS, and OLCI. On 25 July 2010, MODIS and MERIS exhibited broadly consistent spatial patterns, although the median value from MODIS (0.67 m^{-1}) was slightly higher than that from MERIS (0.47 m^{-1}). On 14 July 2016, MODIS and OLCI also revealed very similar spatial distributions, with close agreement in median concentrations (0.59 m^{-1} for MODIS and 0.50 m^{-1} for OLCI). These results indicate that, despite minor differences in magnitude, both the spatial distribution and central tendency of $a_{CDOM}(440)$ are comparable across sensors, demonstrating strong inter-sensor consistency.

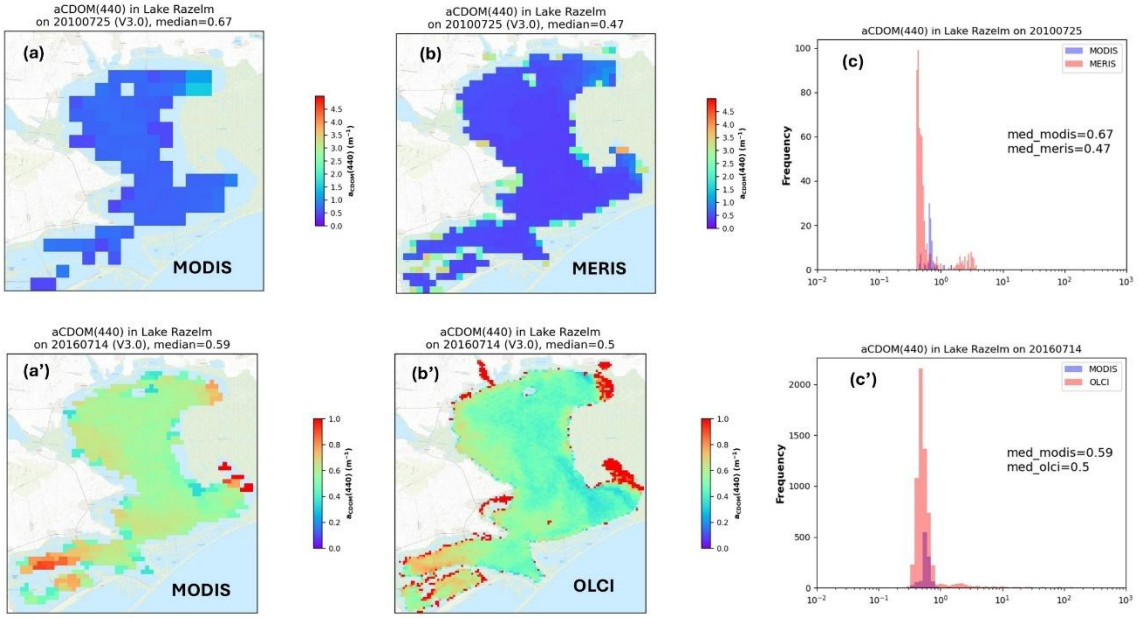


Figure 74 Inter-sensor comparison of same-day $a_{CDOM}(440)$ distributions in Lake Razelm: (a–c) MODIS versus MERIS on 25 July 2010, and (a′–c′) MODIS versus OLCI on 14 July 2016.

Figure 75 presents an inter-sensor comparison of same-day $a_{CDOM}(440)$ distributions in Lake Titicaca from MODIS, MERIS, and OLCI. On 29 September 2005, MODIS and MERIS exhibited nearly identical spatial patterns with closely aligned median values (0.11 m^{-1} and 0.12 m^{-1}, respectively), indicating strong inter-sensor consistency. On 9 November 2018, MODIS and OLCI also captured similar spatial distribution features; however, OLCI reported a substantially higher median concentration (0.40 m^{-1}) compared to MODIS (0.12 m^{-1}). This notable overestimation by OLCI suggests a systematic bias in $a_{CDOM}(440)$ retrievals for this sensor, especially in low observation range (Figure 67), despite overall agreement in spatial patterns.

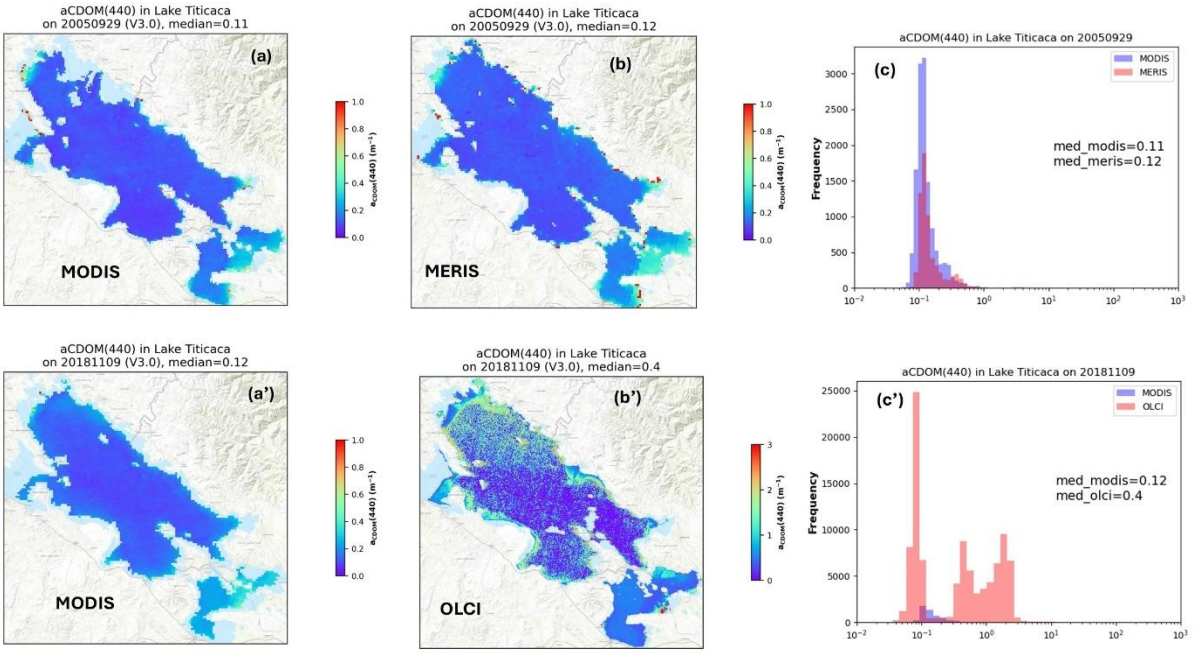


Figure 75 Inter-sensor comparison of same-day $a_{CDOM}(440)$ distributions in Lake Titicaca: (a–c) MODIS versus MERIS on 29 September 2005, and (a′–c′) MODIS versus OLCI on 9 November 2018.

5.4.2.2 Inter-sensor consistency in time-series variation

Figure 76 presents inter-sensor comparisons of Rrs at 443, 490, and 560 nm, Chl-a, TSM, and $a_{CDOM}(440)$ between MODIS and MERIS and between MODIS and OLCI in Lake Sevan. For MODIS–MERIS, the temporal dynamics of Rrs, Chl-a, and TSM were generally consistent, with high concordance correlation coefficients (CCC > 0.8) and strong rank correlations (Kendall's τ = 0.62–0.86). Median values were also closely aligned, though MODIS reported slightly higher Rrs and Chl-a, and lower $a_{CDOM}(440)$ compared with MERIS. For MODIS–OLCI, agreement was also strong across most variables (CCC up to 0.98), with excellent consistency in the seasonal cycles of Rrs, Chl-a, and TSM. However, OLCI tended to produce substantially higher $a_{CDOM}(440)$ than MODIS, indicating a systematic bias for this parameter, especially in the low concentration range (Figure 67). Overall, the results demonstrate robust inter-sensor consistency in spectral reflectance and biogeochemical products, while highlighting sensor-specific differences, especially for $a_{CDOM}(440)$.

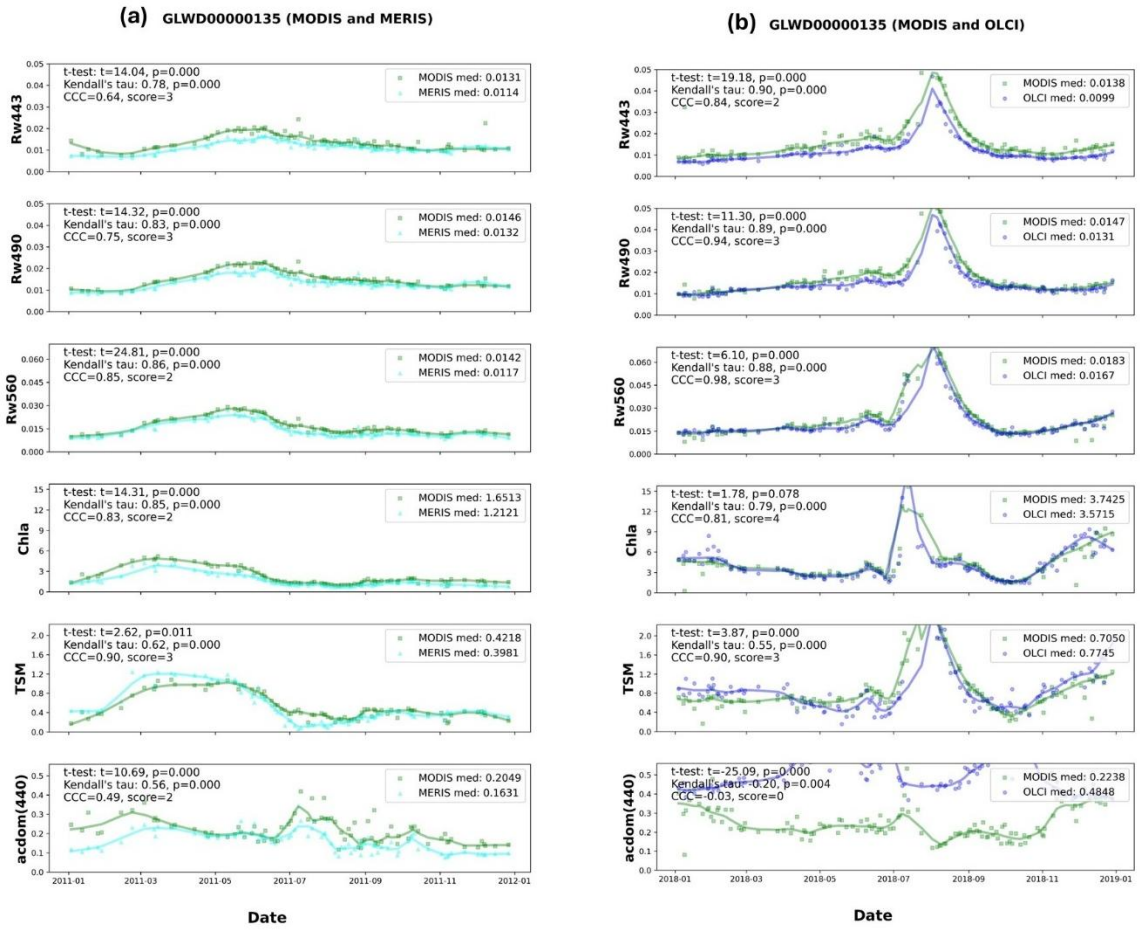


Figure 76 Inter-sensor comparison results of Rw443, Rw490, Rw560, chlorophyll-a, TSM and $a_{CDOM}(440)$ between (a) MERIS and MODIS (2011), (b) OLCI and MODIS (2018) in Lake Sevan.

5.4.3 Climatologic filtering procedure evaluation

As in CRDP v2.1, a similar climatological filtering procedure was also applied in CRDP v3.0.0. Because a systematic analysis of the effects of climatological filtering was not available at the time of the publication of the report, an illustrative example based on turbidity product in CRDP v2.1 is provided here. Figure 77 presents time-series of turbidity from Great Slave Lake and Amadjuak Lake. The figure highlights the daily 10th and 90th percentiles for each lake, indicated by the lower and upper arrows, respectively. This comparison illustrates the data both before and after implementing climatological consistency flags ('lwlr_poor_consistency' and 'lwlr_low_consistency') found in the 'lwlr_quality_flags' band of CRDP v2.1. Notably, the application of these flags effectively removes extreme outliers, particularly those observed during winter, in both lakes. This results in a more coherent and reasonable time-series representation, enhancing the reliability of our water quality data.

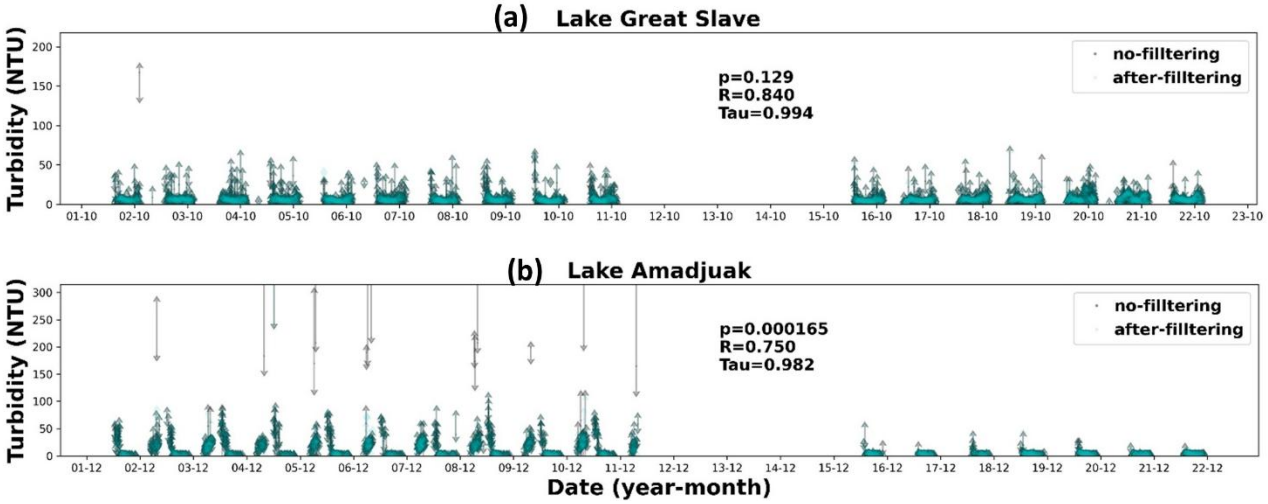


Figure 77 Turbidity Time-Series in (a) Great Slave Lake and (b) Amadjuak Lake, Before and After Applying Climatological Consistency Flags. The figure displays daily 10th and 90th percentile turbidity, indicated by vertical bars.

5.5 Conclusion and recommendations

In CRDP v3.0.0, validation of remote-sensing reflectance (Rrs) against in situ observations did not reveal systematic bias, representing a clear improvement over previous CRDP versions where a systematic negative bias was evident. This improvement largely stems from the upgraded POLYMER atmospheric correction algorithm. The new Polymer v4.17 provides a wider range of initialization conditions and makes more appropriate use of the mineral absorption component in the bio-optical model, which had previously been unstable. In addition, this version incorporates the SWIR band, improving performance under highly turbid water conditions.

Validation and calibration of Chl-a, TSM, and $a_{CDOM}(440)$ algorithms were performed for each OWT across MERIS, OLCI, and MODIS. The weighted-blending procedure further improved retrieval performance relative to the individual algorithms, with the top-3 blended retrievals showing good agreement with in situ observations for Chl-a and TSM across all sensors. While $a_{CDOM}(440)$ retrievals also showed significant agreement, performance was notably lower than for Chl-a and TSM. In particular, OLCI $a_{CDOM}(440)$ aCDOM(440) retrievals were found to be overestimated in the low range, which is expected given the known challenges of retrieving a_{CDOM} in optically complex inland waters and the reduced quality of atmospheric-corrected Rrs in the shorter spectral bands upon which $a_{CDOM}(440)$ algorithms rely.

Validation of MODIS reflectance and water quality products show the capability of using MODIS products to fill the data gap between MERIS and OLCI for specific lakes. In the previous CRDPs, Polymer was evaluated against the NASA operational atmospheric correction algorithm for MODIS (l2gen), with both showing systematic negative bias in the R_{rs} for the bands evaluated, due to the challenges for accurately performing atmospheric correction over inland optical complex waters. Polymer was found to produce almost twice the number of valid matchups with in situ measurements compared to l2gen, demonstrating its capability of handling optically complex waters better, and therefore adopted here. However, it should be noted that the validation data set does not adequately include all optical water conditions and there is a risk of unnecessarily masking, loss or misrepresentation of LWLR with Polymer, particularly in extremely turbid waters, an issue which is being investigated.

Validation of Chl-a and TSM algorithms shows that the per-OWT tuning and weighted-bending procedure dramatically improved the retrieval performance compared to that of the individual algorithms. The overall

performance of Chl-a and TSM algorithms for MODIS was relatively poor compared to that of MERIS, which is reasonable considering that MODIS was primarily designed to observe oceanic waters.

The integration of MODIS into CRDP v3.0.0 enables the generation of an uninterrupted global inland water quality dataset. At present, 73 lakes have been selected for provision with the MODIS dataset following primary inter-sensor evaluations with MERIS and OLCI, which offer more reliable observations and are less affected by land adjacency effects. The combination of upgrades to the atmospheric correction algorithm and the per-OWT algorithm validation and calibration undertaken for this version has increased the number of lakes with consistent observations from 48 (in CRDP v2.1) to 73 in CRDP v3.0.0 and, for the first time, enabled the inclusion of an uninterrupted TSM product.

In the CRDP v3.0.0, we introduced a new 'lwlr_quality_flags' band, which is pivotal for users to understand the rationale behind missing observations and to effectively utilize the dataset. This band includes several flags ('lwlr_poor_consistency' and 'lwlr_low_consistency'), among which two are particularly informative about climatological consistency, stemming from our detailed consistency investigation efforts. Our analysis comparing turbidity time-series before and after applying these consistency flags indicates good performance, demonstrating the importance and efficacy of this procedure. However, it's important to note that pixels currently flagged as 'low' and 'poor' consistency are included in CRDP v3.0.0. In the future CRDP dataset, it is considered to evaluate the potential benefits of excluding these flagged observations to enhance overall data quality.

Finally, it should be noted that, thus far, algorithm calibration, validation and uncertainty characterisation have made use of all available (under license) data for each analysis. In future, data sets of sufficient density will need to be split into calibration and validation data sets, with the latter contributing to the characterisation of product uncertainty.

5.6 References

Binding, C., Jerome, J., Bukata, R., & Booty, W. (2010). Suspended particulate matter in Lake Erie derived from MODIS aquatic colour imagery. International Journal of Remote Sensing, 31, 5239-5255

Carrea, L., Embury, O. and Merchant, C. J. (2015) Datasets related to in-land water for limnology and remote sensing applications: distance-to-land, distance-to-water, water-body identifiers and lake-centre co-ordinates. Geoscience Data Journal, 2(2). pp. 83-97. ISSN 2049-6060 doi:10.1002/gdj3.32

Chen, S., Han, L., Chen, X., Li, D., Sun, L. and Li, Y. (2015). Estimating wide range Total Suspended Solids concentrations from MODIS 250-m imageries: An improved method. ISPRS Journal of Photogrammetry Remote Sensing, 9958-69.

Chen, Z., Hu, C. and Muller-Karger, F. (2007). Monitoring turbidity in Tampa Bay using MODIS/Aqua 250-m imagery. Remote Sensing of Environment, 109(2), 207-220.

Dall'Olmo, G., Gitelson, A.A., Rundquist, D.C., Leavitt, B., Barrow, T. and Holz, J.C. (2005). Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. Remote Sensing of Environment, 96(2), 176-187.

Dogliotti, A.I., Ruddick, K., Nechad, B., Doxaran, D. and Knaeps, E. (2015). A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarin waters. Remote Sensing of Environment, 156157-168.

Doxaran, D., Froidefond, J.-M., Castaing, P. and Babin, M. (2009). Dynamics of the turbidity maximum zone in a macrotidal estuary (the Gironde, France): Observations from field and MODIS satellite data. Estuarine, Coastal Shelf Science, 81(3), 321-332.

El-Alem, A., Chokmani, K., Laurion, I. and El-Adlouni, S.E. (2012). Comparative analysis of four models to estimate chlorophyll-a concentration in case-2 waters using MODerate resolution imaging spectroradiometer (MODIS) imagery. Remote Sensing, 4(8), 2373-2400.

Embury, O., Merchant, C. J. and Corlett G.K. (2012) A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: Initial validation, accounting for skin and diurnal variability effects. Remote Sensing of Environment, 116. pp. 62-78. ISSN 0034-4257 doi:10.1016/j.rse.2011.02.028

Gilerson, A.A., Gitelson, A.A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., & Ahmed, S.A. (2010). Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Optics Express, 18, 24109-24125

Gitelson, A. (1992). The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. International Journal of Remote Sensing, 13(17), 3367-3373.

Gitelson, A.A., Dall'Olmo, G., Moses, W., Rundquist, D.C., Barrow, T., Fisher, T.R., Gurlin, D. and Holz, J. (2008). A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. Remote Sensing of Environment, 112(9), 3582-3593.

Gitelson, A.A., Schalles, J.F. and Hladik, C.M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sensing of Environment, 109(4): 464-472.

Gons, H.J., Rijkeboer, M., & Ruddick, K.G. (2005). Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. Journal of Plankton Research, 27, 125-127

Gurlin, D., Gitelson, A.A. and Moses, W.J. (2011). Remote estimation of chl-a concentration in turbid productive waters—Return to a simple two-band NIR-red model? Remote Sensing of Environment, 115(12), 3479-3490.

Hondzo, M., You, J., Taylor, J., Bartlet, G., Voller, V.R. (2022) Measurements and scaling of lake surface skin temperatures. Geophysical Reaserch Letters, 49, e2021GL093226.

Hu, C., Chen, Z., Clayton, T.D., Swarzenski, P., Brock, J.C. and Muller–Karger, F.E. (2004). Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: Initial results from Tampa Bay, FL. Remote Sensing of Environment, 93(3), 423-441.

Lehmann, M.K., Gurlin, D., Pahlevan, N. et al. (2023). GLORIA - A globally representative hyperspectral in situ dataset for optical sensing of water quality. Sci Data 10, 100.

Letelier, R.M. and Abbott, M.R. (1996). An analysis of chlorophyll fluorescence algorithms for the Moderate Resolution Imaging Spectrometer (MODIS). Remote Sensing of Environment, 58(2), 215-223.

Maritorena, S., Siegel, D.A. and Peterson, A.R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. Applied Optics, 41(15), 2705-2714.

Miller, R.L. and McKee, B.A. (2004). Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. Remote Sensing of Environment, 93(1-2), 259-266.

Mishra, S., Mishra, D.R., & Lee, Z. (2013a). Bio-optical inversion in highly turbid and cyanobacteria-dominated waters. IEEE Transactions on Geoscience Remote Sensing, 52, 375-388

Mishra, S., Mishra, D.R., Lee, Z., & Tucker, C.S. (2013b). Quantifying cyanobacterial phycocyanin concentration in turbid productive waters: A quasi-analytical approach. Remote Sensing of Environment, 133, 141-151

MODIS Characterization Support Team., “MODIS Level 1B Product User’s Guide- For Level 1B Version 6.2.2 (Terra) and Version 6.2.1 (Aqua),” (2017) https://mcst.gsfc.nasa.gov/sites/default/files/file_attachments/M1054E_PUG_2017_0901_V6.2.2_Terra_V6.2.1_Aqua.pdf

Nechad, B., Dogliotti, A., Ruddick, K., & Doxaran, D. (2016). Particulate backscattering and suspended matter concentration retrieval from remote-sensed turbidity in various coastal and riverine turbid waters. In, Proceedings of ESA living planet symposium, Prague (pp. 9-13)

Nechad, B., Ruddick, K., & Park, Y.J.R.S.o.E. (2010). Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters, 114, 854-866

Neil, C., Spyrakos, E., Hunter, P.D., & Tyler, A.N. (2019). A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. Remote Sensing of Environment, 229, 159-178

O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., & McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. Journal of Geophysical Research: Oceans, 103, 24937-24953

Ondrusek, M., Stengel, E., Kinkade, C.S., Vogel, R.L., Keegstra, P., Hunter, C. and Kim, C. (2012). The development of a new optical total suspended matter algorithm for the Chesapeake Bay. Remote Sensing of Environment, 119243-254.

Petus, C., Chust, G., Gohin, F., Doxaran, D., Froidefond, J.-M. and Sagarminaga, Y. (2010). Estimating turbidity and total suspended matter in the Adour River plume (South Bay of Biscay) using MODIS 250-m imagery. Continental Shelf Research, 30(5), 379-392.

Riggs, G. A., Hall, D. K., & Román. M.O. (2019). MODIS Snow Products Collection 6.1 User Guide.

Shi, K., Zhang, Y., Xu, H., Zhu, G., Qin, B., Huang, C., Liu, X., Zhou, Y. and Lv, H. (2015a). Long-Term Satellite Observations of Microcystin Concentrations in Lake Taihu during Cyanobacterial Bloom Periods. Environ Science Technology, http://www.ncbi.nlm.nih.gov/pubmed/25936388.

Shi, K., Zhang, Y., Zhu, G., Liu, X., Zhou, Y., Xu, H., Qin, B., Liu, G. and Li, Y. (2015b). Long-term remote monitoring of total suspended matter concentration in Lake Taihu using 250 m MODIS-Aqua data. Remote Sensing of Environment, 16443-56.

Sipelgas, L., Raudsepp, U. and Kõuts, T. (2006). Operational monitoring of suspended matter distribution using MODIS images and numerical modelling. Advances in Space Research 38(10), 2182-2188.

Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G., Neil, C., Barbosa, C.C., Binding, C.E., & Bradt, S. (2018). Optical types of inland and coastal waters. Limnology Oceanography, 63, 846-870

Saunders, P.M. (1967) The temperature at the ocean-air interface. Journal of the Atmospheric Science, 24. pp. 269-273. doi:0.1175/1520-0469(1967)024<0269:TTATOA>2.0.CO;2

Steinmetz, F., Deschamps, P.-Y., & Ramon, D. (2011). Atmospheric correction in presence of sun glint: application to MERIS. Optics Express, 19, 9783-9800

Vantrepotte, V., Loisel, H., Mériaux, X., Neukermans, G., Dessailly, D., Jamet, C., Gensac, E., & Gardel, A. (2011). Seasonal and inter-annual (2002-2010) variability of the suspended particulate matter as retrieved from satellite ocean color sensor over the French Guiana coastal waters. Journal of Coastal Research, 1750-1754

Vermote, E. F., Roger, J. C., & Ray, J. P. (2015). *MODIS Surface Reflectance User's Guide Collection 6.* Maryland: MODIS Land Surface Reflectance Science Computing Facility.

Wang, M., Nim, C.J., Son, S. and Shi, W. (2012). Characterization of turbidity in Florida's Lake Okeechobee and Caloosahatchee and St. Lucie estuaries using MODIS-Aqua measurements. Water research, 46(16), 5410-5422.

Wu, Y., Duguay, C.R. & Xu, L. (2021). Assessment of machine learning classifiers for global lake ice cover mapping from MODIS TOA reflectance data. Remote Sensing of Environment, 253, 112206, https://doi.org/10.1016/j.rse.2020.112206.

Zhang, M., Tang, J., Dong, Q., Song, Q. and Ding, J. (2010). Retrieval of total suspended matter concentration in the Yellow and East China Seas from MODIS imagery. Remote Sensing of Environment, 114(2), 392-403.

Zhang, Y., Shi, K., Liu, X., Zhou, Y., & Qin, B. (2014). Lake topography and wind waves determining seasonal-spatial dynamics of total suspended matter in turbid Lake Taihu, China: assessment using long-term high-resolution MERIS data. PLoS One, 9

Zhao, H., Chen, Q., Walker, N.D., Zheng, Q. and MacIntyre, H.L. (2011). A study of sediment transport in a shallow estuary using MODIS imagery and particle tracking simulation. International Journal of Remote Sensing, 32(21), 6653-6671.

6 Lake Ice Cover – LIC

Lake ice cover (LIC) refers to the extent of a lake covered by ice. The Lakes_cci LIC is generated from MODIS Terra/Aqua Calibrated Radiance 5-Min L1B Swath (MOD02/MYD02), Collection 6.1 data (see ATBD for details). Each pixel is assigned one of three possible class labels: water (value=1), ice (value=2), or cloud (value=3). It is important to note that validation of the Random Forest (RF model) used to generate the classification is performed individually on Terra and Aqua derived LIC, respectively.

Product validation is performed using three approaches: (1) comparison with LIC products generated by other algorithms/groups such as NASA MODIS Snow Cover products from Terra and Aqua (MOD10/MYD10); and (2) validation against groups of pixels (Areas Of Interest or AOI) extracted for a selection of lakes from visual interpretation of original MODIS Terra and Aqua imagery used as input into the LIC retrieval algorithm (i.e. MOD02/MYD02); and (3) comparison of LIC extent between the CCI product and other operational products such as the Multisensor Snow and Ice Mapping System (IMS).

6.1 Data description

6.1.1 MODIS Terra/Aqua Calibrated Radiances Level 1B product (MOD02/MYD02)

MODIS Terra/Aqua Calibrated Radiances L1B 5-Min Swath, Collection 6.1 (MODIS Characterization Support Team , 2017) – MOD02/MYD02 false color composites (R: band 2, G: band 2, B: band 1) with a 250 m spatial resolution were used as reference images to manually collect AOIs with assigned labels (lake ice, open water, and cloud) to assess the accuracy of the Lakes_cci LIC. MODIS band 1 is centred at 645nm (red) and band 2 at 865 nm (near-IR). In addition to false color composites, true color composites were also used for visual inspection (R: band 1, G: band 4, B: band 3). MODIS band 3 is centered at 469nm (blue) and band 4 at 555nm (green) with 500m.

6.1.2 MODIS Snow Cover product (MOD10/MYD10)

MODIS Terra/Aqua Snow Cover 5-Min L2 Swath 500 m, Collection 6.1 (C6.1), daily products (MOD10/MYD10) were also validated and compared with the Lakes_cci LIC product. In MOD10/MYD10, lake ice cover is identified using a similar criterion as snow over land (i.e. Normalized Difference Snow Index (NDSI) ≥ 10 and data screens) and a land/water mask provided in MOD/MYD03 products (Riggs et al., 2019). The NDSI is derived from MODIS radiance data acquired by Terra/Aqua satellites (i.e. MODIS Level 1B product (MOD02/MYD02) top-of-the-atmosphere (TOA) calibrated radiance/reflectance data). MODIS Cloud Mask products (MOD/MYD35_L2) have been also applied to filter cloud cover for MOD10/MYD10 products. Input data to the Snowmap algorithm used in the production of MOD10/MYD10 are shown in Table 26. Therefore, four labels (i.e., ice, water, cloud, other) are presented in MOD10/MYD10 snow maps to indicate ice cover over lake areas.

Table 26. MODIS data inputs for the collection 6.1 (C6.1) snow algorithm (adapted from Riggs et al., 2019)

Input Data	Product Description
MOD/MYD02HKM	Calibrated Radiances (500 m)
MOD/MYD021KM	Calibrated Radiances (1 km)
MOD/MYD03	Geolocation Fields (1 km)
MOD/MYD35_L2	Cloud Mask and Spectral Test Results (250 m and 1 km)

* Depends on platform, MOD indicates Terra, MYD indicates Aqua

The MODIS Snow Cover product experiences issues in differentiating between clouds and snow/ice, and both errors of commission and omission can be found in the product. Omission occurs when the MODIS cloud mask (MOD/MYD 35) identifies an area of snow/ice as certain cloud, therefore excluding it from the MODIS snow product (Riggs et al., 2019). Errors of commission can occur on the periphery of cloud cover or where clouds appear similar to snow/ice and are excluded from the cloud mask and instead identified as snow/ice (Riggs et al., 2019). Cloud cover can be removed from the product by extracting the classification value from the previous and proceeding pixel to determine the possible classification of a cloud covered pixel. High solar zenith angle can also prevent accurate classification of snow/ice cover, which is important when studying freeze-up at northern high latitudes. Pixels are screened using a solar zenith angle mask of > 70° and pixels with a solar zenith angle ≥ 85° are classified as night.

6.1.3 Multisensor Snow and Ice Mapping System (IMS)

The multisensor snow and ice mapping system (IMS) is a daily product generated by the National Oceanic and Atmospheric Administration (NOAA) (Ramsay, 1998). The product is available in three spatial resolutions, which are available with different temporal coverage. The 24-km grid is available 1997 to present, the 4-km grid is available 2004 to present, and the 1-km grid is available 2014 to present (NOAA, 2022). Grid cells are classified as either open water, land without snow, sea ice or lake ice, and snow covered land. IMS is gap-free as each cell has a daily classification (cloud cover is absent). When there is not enough information to make a classification decision (i.e. cloud cover or other obstruction), then the previous classification is used. Grid cells are ice-covered if >40% of the cell is ice covered, this decision is made by an analyst and is not a threshold-based algorithm (NOAA, 2022).

There are numerous data sources used to generate the IMS product. These sources include, but are not limited to, passive microwave data (AMSR-2), active microwave data (RADARSAT-2, Sentinel-1), optical imagery (MODIS, AVHRR, and GOES Imager), analysis products (Canadian Snow Analysis – Environment Climate Change Canada), and weather reports (NOAA, 2022). These data are used to make the decisions regarding the classification of each grid cell.

6.2 Comparison methods

Validation of the Lakes_cci LIC product and NASA MODIS Snow Cover product (also used for comparison) has been performed through computation of confusion matrices built on independent statistical validation. The reference data for validation were collected for water, ice and cloud as AOIs from the visual interpretation of the false colour composite surface reflectance images (MOD09/MYD09) over three ice seasons (freeze-up and break-up periods) interspersed across the 20-year MODIS record by skilled ice analysts.

In addition to statistical validation, comparison of the Lakes_cci LIC product and NASA MODIS Snow Cover product was performed through visual inspection. The reference data include false colour composite and true colour composite images.

For IMS data, lake ice extent is calculated for the LIC and IMS product. The validation activity is conducted using the 4-km product as it allows for the widest coverage of CCI target lakes. The ice extent is compared between the products to calculate mean absolute error (MAE) and root mean square errors (RMSE). When performing the comparison, observations from the LIC product where >70% of the lake is obstructed are removed to limit the impact of cloud cover on the calculated statistics. The pattern in lake ice extent is also compared visually.

6.3 Description of the work

Work on product validation and intercomparison is comprised of three parts. The first part includes algorithm development and validation for the Lakes_cci LIC product derived from MODIS Terra. Validation for the Lakes_cci LIC product generated from MODIS Aqua is presented in the second part. The last part shows intercomparison of the Lakes_cci LIC and NASA Snow cover products by visual inspection.

In the first part, 17 lakes have been selected across the Northern Hemisphere to serve for the purpose of both development and validation of the Lakes_cci LIC product derived from MODIS Terra (Figure 78). Samples were collected for three ice seasons (2002-2003, 2009-2010, 2016-2017) as to include MODIS data from Terra (2000-present) and Aqua (2002-present) and provide a good temporal spread over the full record to ensure algorithm stability. For each lake, one image from the freeze-up period and one image from the break-up period were selected for both Terra and Aqua images if available. False colour composites (R: Band 2, G: Band 2, B: Band 1) with a 250 m spatial resolution were used as reference images to manually extract AOIs with labels (ice, water, and cloud). In total 54 images (20 from FU, 34 from BU) were selected and 276,003 pixels were sampled from the selected AOIs. A 100-fold cross-validation was used to conduct model comparison and examine model transferability. In order to evaluate the performance of the algorithm for the freeze-up and break-up periods separately, additional samples from MODIS Terra were collected for a full ice season (2018 - 2019) from Great Slave Lake and Lake Ladoga. A total of 10,075,081 pixels taken from 229 swaths over the two lakes were used to present classification performance by confusion matrices.

The second part shows the validation of the Lakes_cci LIC product generated by MODIS Aqua. Similar to the work for the Lakes_cci LIC product from MODIS Terra, 1,665,188 samples were collected based on false colour composite images produced from the MYD02 Aqua product over Great Slave Lake (156 swaths), Lake Ladoga and Lake Onega (61 swaths) in 2020. A confusion matrix computed by using the samples is provided to present the classification performance.

In the last part, the Lakes_cci LIC product is compared with the NASA MODIS Snow Cover product through visual inspection. Several classification challenges for the two products, such as thin ice detection, thin cloud/fog retrieval, and confusion between ice and cloud are examined. The lake ice extent is then compared to IMS lake ice extent for 28 lakes across the Northern Hemisphere.

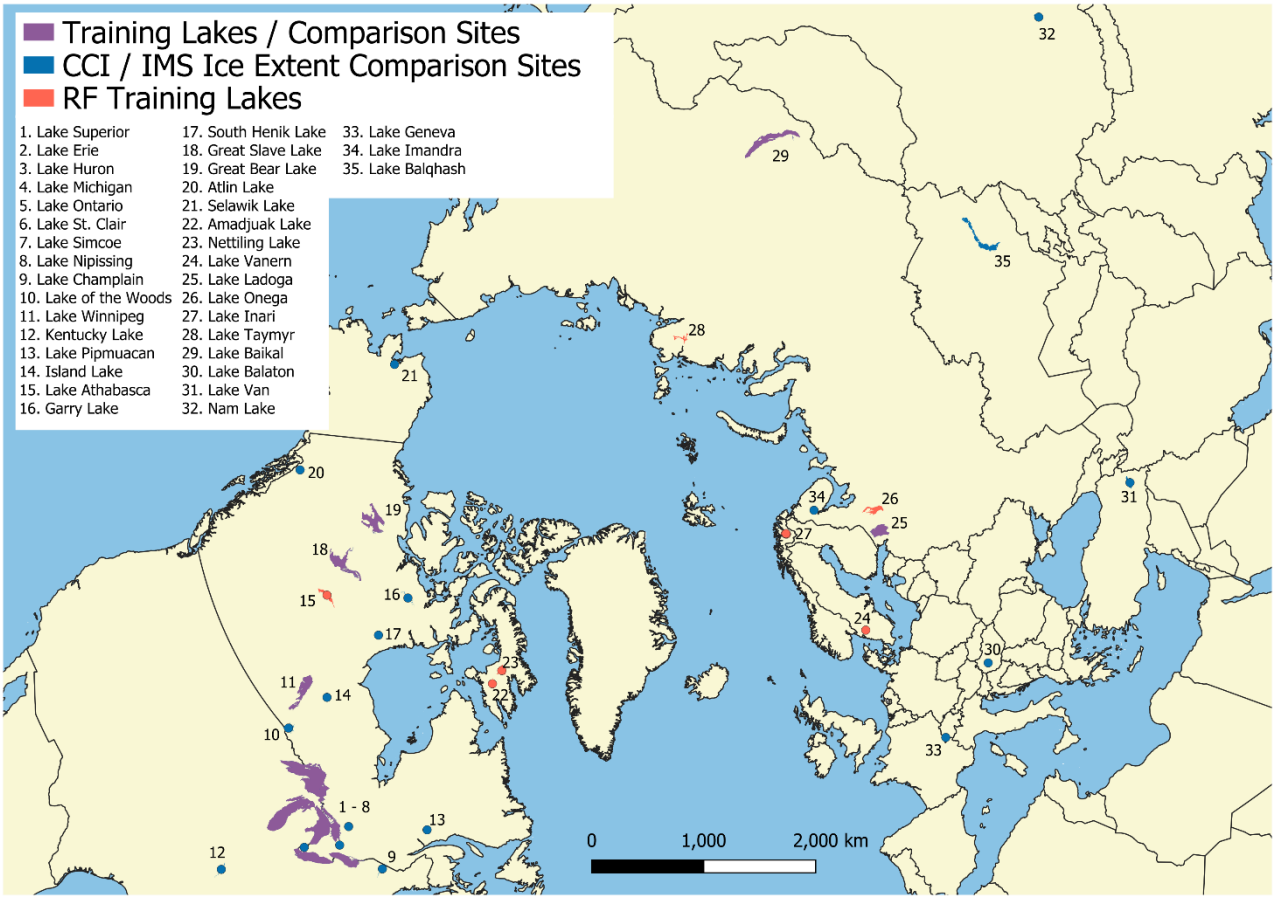


Figure 78: Geographical distribution of lakes used for Lakes_cci LIC algorithm development and validation

6.4 Validation results

Figure 79 shows the accuracies computed from a 100-fold cross-validation (CV) using the samples of MODIS Terra from the 17 lakes. Random forest (RF) was found to outperform two other machine learning algorithms (multinomial logistic regression, MLR, and support vector machine, SVM) and comparable to gradient boosting trees (GBT) for lake ice cover, open water and cloud classification in a recent paper by the developers of the current LIC product (Wu et al. 2021). Furthermore, RF provided consistent results based on a comprehensive accuracy assessment (random k-fold as well as spatial and temporal CV as shown in Table 27).

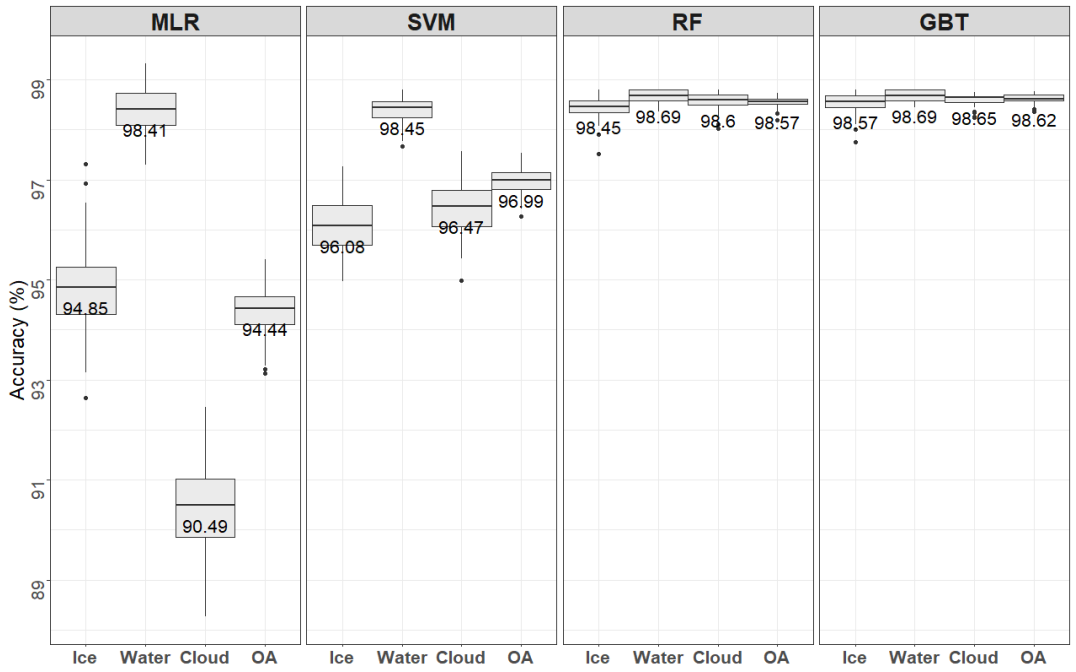


Figure 79: Comparison of accuracies (%) obtained using random 100-fold CV across classifiers for the ice, water and cloud classes individually, and overall (OA) (Wu et al., 2021)

Table 27: Accuracy assessment using temporal and spatial CV methods (adapted from Wu et al. 2021)

	MLR	SVM	RF	GBT	
Temporal CV accuracy	93.21%	83.00%	95.49%	95.15%	
Spatial CV accuracy	90.98%	79.36%	95.64%	95.26%	

Table 28 provides the accuracies reached by the Lakes_cci LIC product with input from MODIS Terra for the break-up (BU) and freeze-up (FU) periods taken individually. There is no notable difference in the accuracy of the Lakes_cci LIC product between the BU (OA: 98.14%) and FU (OA: 96.83%) periods, and the classification accuracies are consistent across classes. Table 29 presents classification results for MODIS Aqua Lakes_cci LIC, which are comparable to the performance of MODIS Terra Lakes_cci LIC shown in Table 28.

Table 28: Confusion matrices with retrieval accuracies for Lakes_cci LIC derived from MODIS Terra (break-up and freeze-up periods individually) (adapted from Wu et al., 2021)

		Freeze-Up				**Break-Up**			
		RF (classification)				RF (classification)			
Lakes_cci LIC Terra		Ice	Water	Cloud	Accuracy	Ice	Water	Cloud	Accuracy
User-defined	Ice	740,105	3,305	15,520	97.52%	774,412	1,213	14,537	98.01%
	Water	613	509,797	4,863	98.94%	927	776,296	4,350	99.32%
	Cloud	154,578	15,790	4,700,459	96.50%	45,418	6,475	2,306,423	97.80%
				Overall Accuracy: 96.83%				Overall Accuracy: 98.14%	

Table 29: Confusion matrices with retrieval accuracies for Lakes_cci LIC derived from MODIS Aqua

Lakes_cci LIC Aqua		RF (classification)			
		Ice	Water	Cloud	Accuracy
User-defined	Ice	246,032	2,487	4,751	97.14%
	Water	1,743	538,668	4,655	98.83%
	Cloud	19,738	5,324	841,790	97.11%
				Overall Accuracy:	**97.68%**

In NASA MOD10/MYD10 products, lake ice cover is estimated as snow cover with the scenario where NDSI is greater than 10 (Riggs et al., 2019). Besides NDSI, low visible reflectance screen thresholds of band 2 ≤ 0.10 or band 4 ≤ 0.11 are applied by the MODIS Snowmap algorithm in inland water bodies to detect lake ice cover (Riggs et al., 2019). Therefore, congelation ice with low visible reflectance is misclassified as open water in MOD10/MYD10 as shown in Figure 80 - Figure 82. Congelation ice usually occurs in the early ice formation or late melting period. Therefore, the performance of congelation ice detection affects the quality of lake ice phenology estimation in MOD10/MYD10. The Lakes_cci LIC product presents more accurate classification results of lake ice cover in the four cases below.

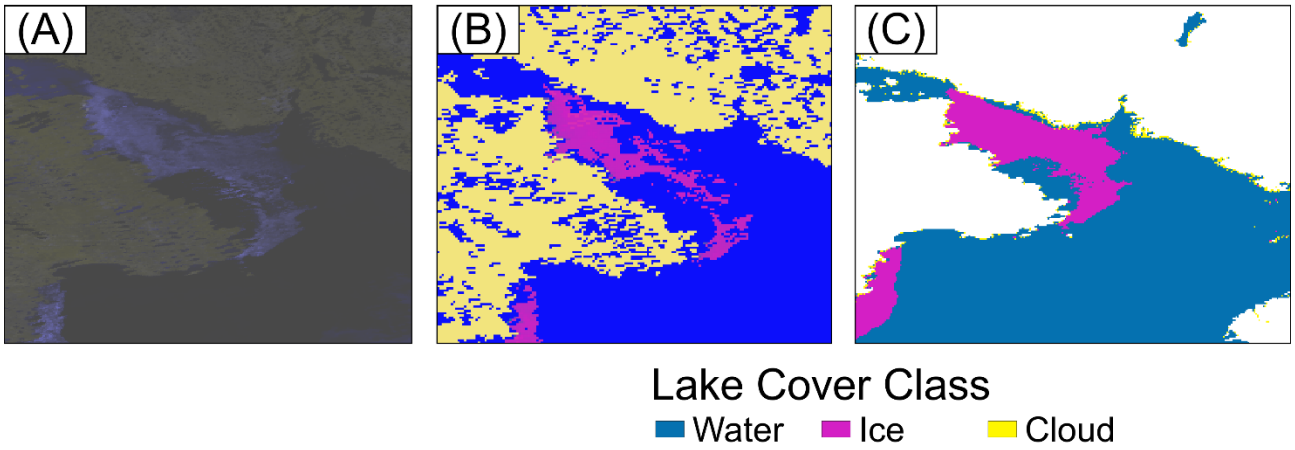


Figure 80: MODIS true RGB colour composite (a), MODIS false RGB colour composite (b), MODIS-derived Lakes_cci LIC (c), and MODIS Snow Cover product (d) for Great Slave Lake (Canada) on 19 June 2020 (UTC 19:25) by Aqua

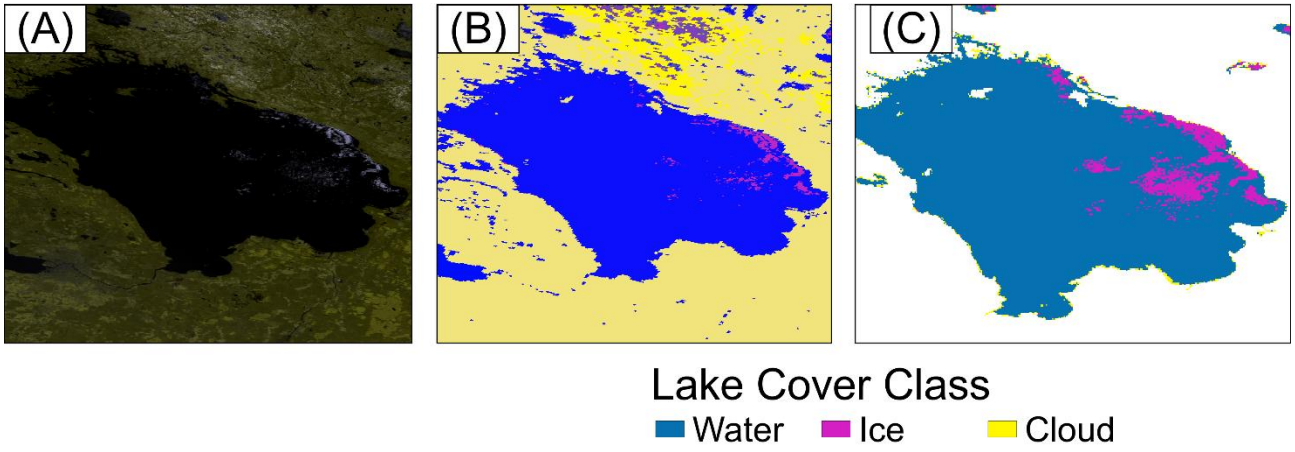


Figure 81: MODIS false RGB colour composite (A), MODIS Snow Cover product (B), and MODIS-derived Lakes_cci (C) for Lake Ladoga (Russia) on 26 April 2019 (UTC 09:45) by Terra

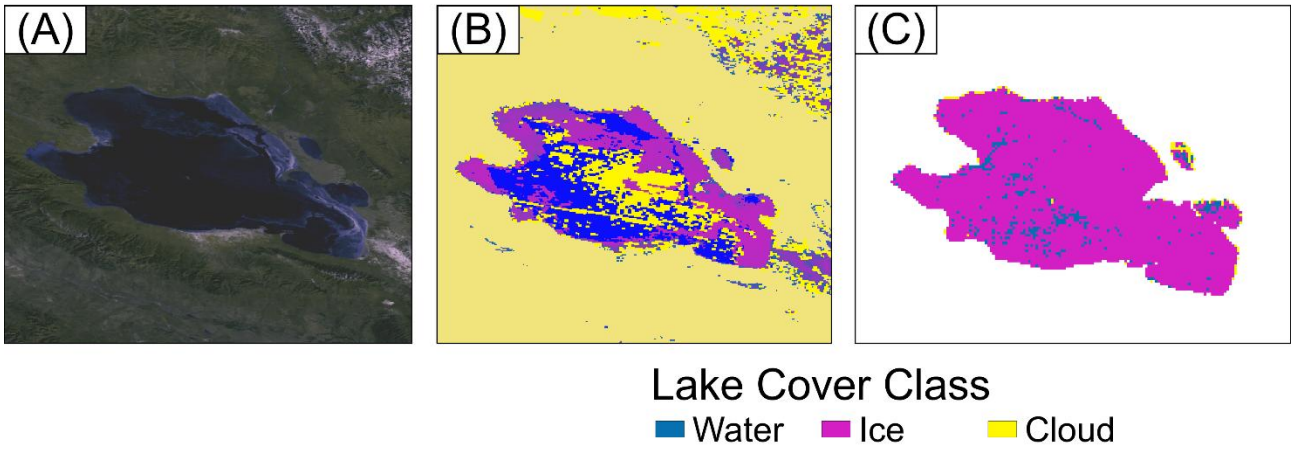


Figure 82: MODIS false RGB colour composite (A), MODIS Snow Cover product (B), and MODIS-derived Lakes_cci (C) for Qinghai Lake (China) on 4 January 2020 (UTC 04:25) by Terra

The MODIS Cloud Cover product (MOD/MYD35_L2) is used in the generation of MOD10/MYD10. Hence, the omission and commission errors of the MODIS Cloud Cover product lead to misclassification in MOD10/MYD10. Figure 83 shows misclassification of thin ice in the MODIS snow cover product due to cloud overestimation by MOD/MYD35_L2. In contrast, the Lakes_cci LIC product can detect the lake ice cover correctly. Moreover, as Figure 84 illustrates, the commission errors of cloud cover in MOD10 also

occur under high solar zenith conditions, whereas the Lakes_cci LIC product performs better in the classification of lake ice cover.

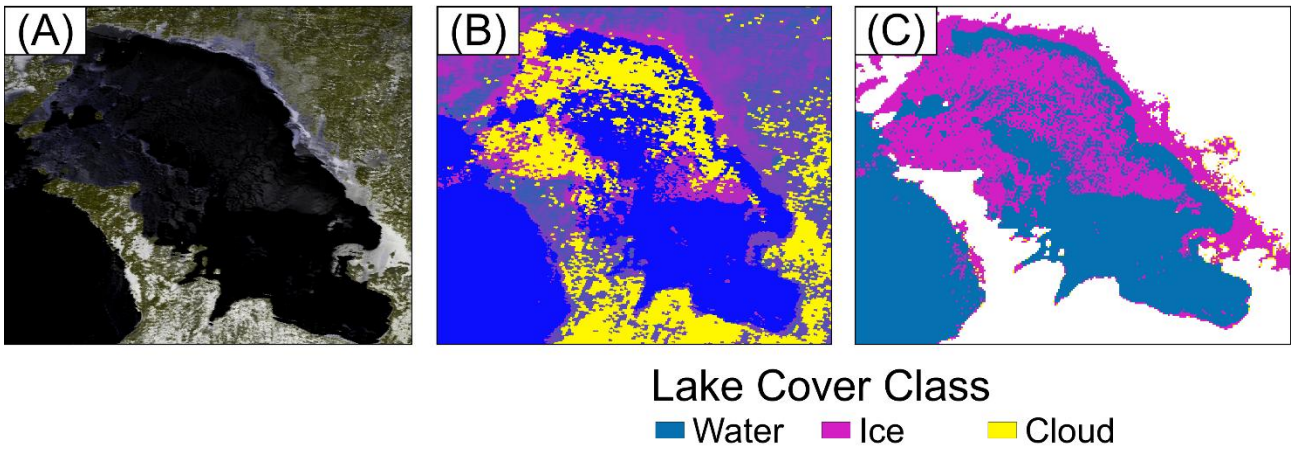


Figure 83: MODIS false RGB colour composite (A), MODIS Snow Cover product (B), and MODIS-derived Lakes_cci (C) for Lake Huron (Canada) on 10 February 2005 (UTC 18:30) by Aqua

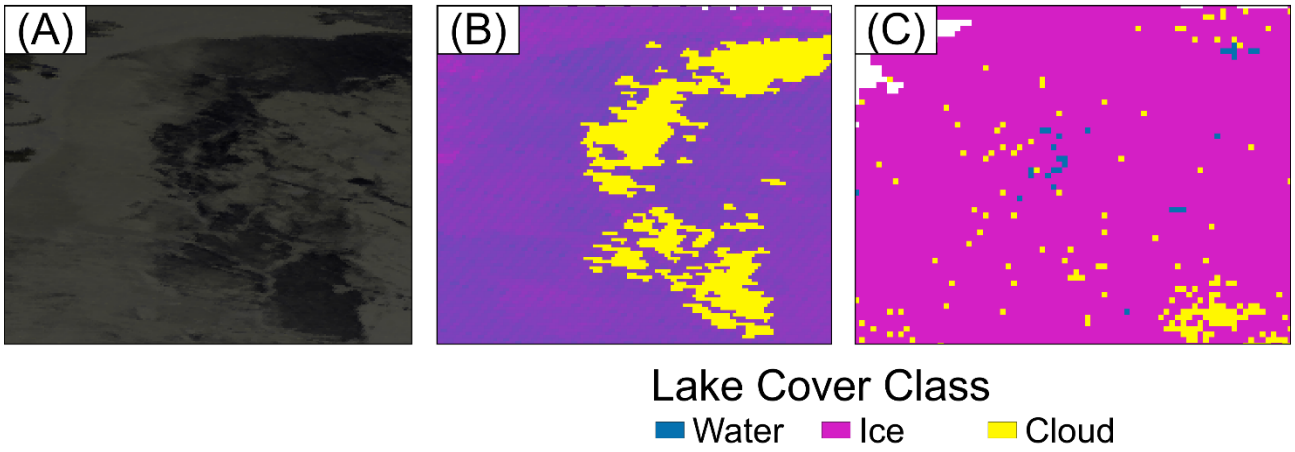


Figure 84: MODIS false RGB colour composite (A), MODIS Snow Cover product (B), and MODIS-derived Lakes_cci (C) for Great Slave Lake (Canada) on 30 December 2018 (UTC 19:20) by Terra under high solar zenith angle conditions (higher than 80 degrees)

The Lakes_cci LIC product also captures better the spatial distribution of cloud cover compared to MYD10/MOD10 **Error! Reference source not found.**(Figure 85). MOD/MYD35_L2 can mislabel low-level clouds and/or fog as "clear sky", resulting in classification errors of cloud to ice in MOD10/MYD10 as shown in Figure 85. In this example, even though the sky is clear the MOD10 product identifies low percentages of ice. However, Lakes_cci LIC is capable of correctly classifying these low-level clouds and

reporting open water. This is aided through the combination of MOD/MYD02 swaths.

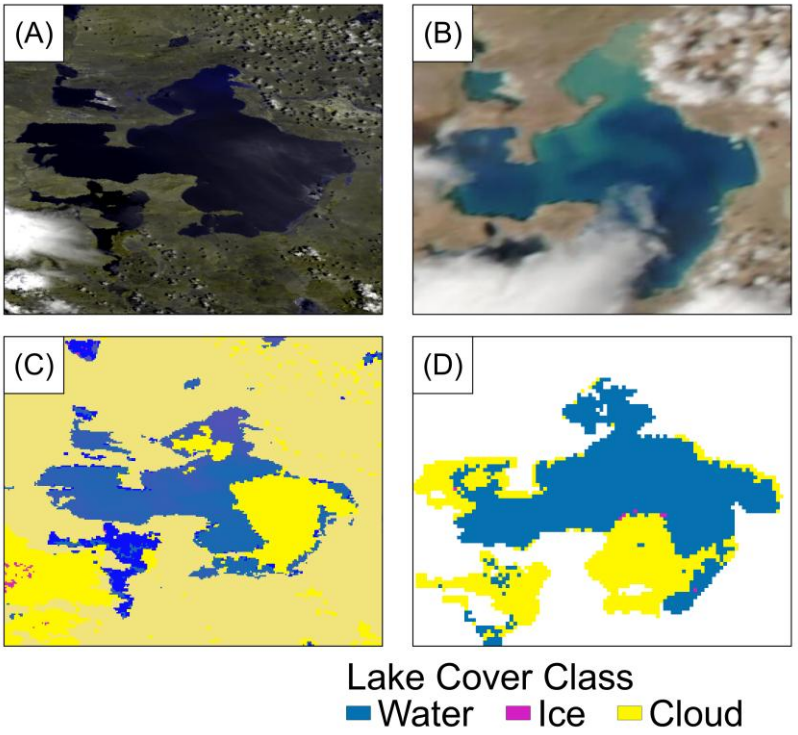


Figure 85: MODIS false RGB colour composite by Terra (A), MODIS false RGB colour composite by Aqua (B), MODIS Snow Cover product (C), and MODIS-derived Lakes_cci (D) for Siling Lake (China) on 1 July 2020.

The basic assumption where a water body is deep and clear, thus absorbing solar radiation incident upon it, is employed by the Snowmap algorithm for surface classification over inland water bodies for MOD10/MYD10 (Riggs et al., 2019). Therefore, water with high turbidity or algal blooms can be incorrectly classified as ice cover. As Figure 86 shows, the Lakes_cci LIC product retrieves the water area that is misclassified as ice by MOD10.

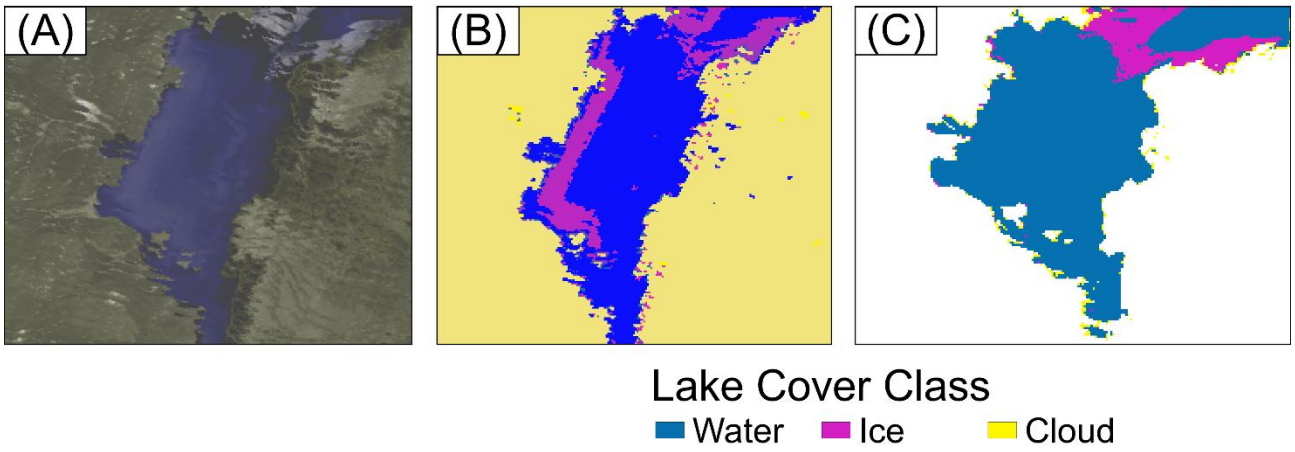


Figure 86: MODIS false RGB colour composite (A), MODIS Snow Cover product (B), and MODIS-derived Lakes_cci (C) for Lake Balkhash (Kazakhstan) on 25 March 2020 (UTC 06:45) by Terra

Another source of error for the MODIS Snow Cover (MOD10/MYD10) product is the geolocation of some lakes. MOD10/MYD10 employs the MODIS Geolocation L1A product in order to mask land and water (Riggs et al., 2019). However, as shown in Figure 87, the MODIS Snowmap algorithm classifies the

majority of Lake Nettilling (Canada) as ocean resulting in no classification for the ice cover on the lake. However, the more accurate lake area mask used for the generation of the Lakes_cci LIC product provides correct classification of the ice conditions on that day.

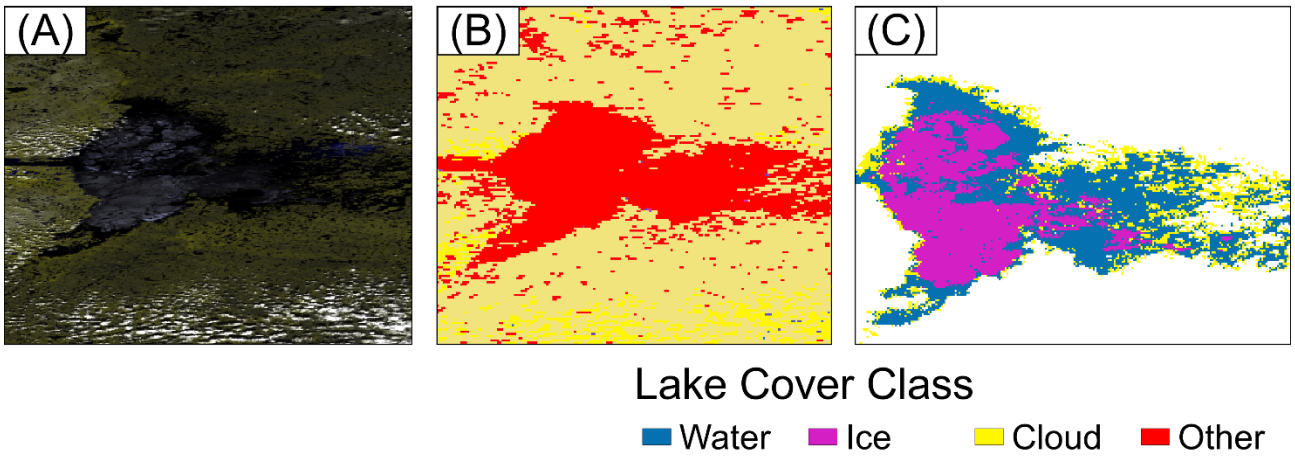


Figure 87: MODIS false RGB colour composite (A), MODIS Snow Cover product (B), and MODIS-derived Lakes_cci (C) for Lake Nettilling (Canada) on 25 July 2020 (UTC 15:45). The red colour corresponds to "ocean" in this MODIS Snow Cover product

Comparisons of ice cover extent were conducted for 28 lakes between the Lakes_cci LIC product and IMS operational product. When the average MAE for all lakes is 0.04% and the average RMSE is 0.13% indicating a strong agreement between the two products. Additionally, this is an improvement over version 2 of the Lakes_cci product which had an average MAE of 0.05% and average RMSE of 0.14%. The small improvement is likely due to the implementation of the temperature filter in v3 of the Lakes_cci product. Visual comparisons of the ice cover extent were also conducted (Figure 88, Figure 89, and Figure 90). For Lake Baikal and Great Slave Lake (Figure 88 and Figure 89) both IMS and the Lakes_cci LIC product follow an expected pattern of seasonal ice cover with an increase to 100% ice cover in the winter and then an absence of ice during the summer months. When observations where >70% of the lake is obstructed are removed, there is confirmation of strong agreement between the CCI LIC product and the IMS data. For Lake Superior (Figure 90), a pattern is observed which is more typical for the Great Lakes where ice coverage does not reach 100% every ice season. Initial Lakes_cci LIC data appears noisy, however, when obstructed dates are removed there is strong agreement between the two datasets as seen for Lake Baikal and Great Slave Lake.

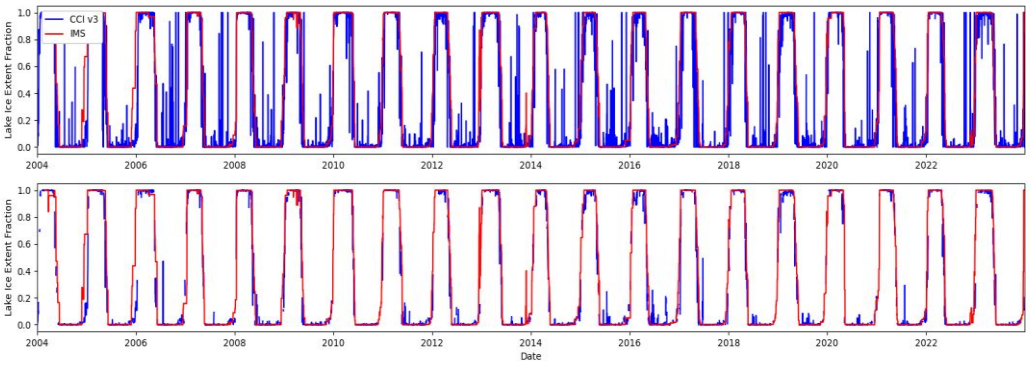


Figure 88: Comparison between Lakes_cci and IMS ice cover extent for Lake Baikal. The top plot shows a comparison between all data, and the bottom plot shows a comparison with Lakes_cci data observations are removed when the lake is >70% obstructed.

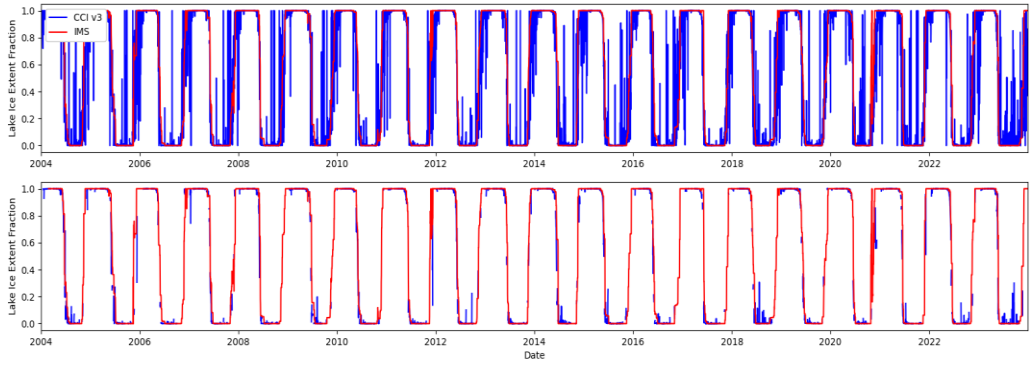


Figure 89: Comparison between Lakes_cci and IMS ice cover extent for Great Slave Lake. The top plot shows a comparison between all data, and the bottom plot shows a comparison with Lakes_cci data where observations are removed when the lake is >70% obstructed are removed.

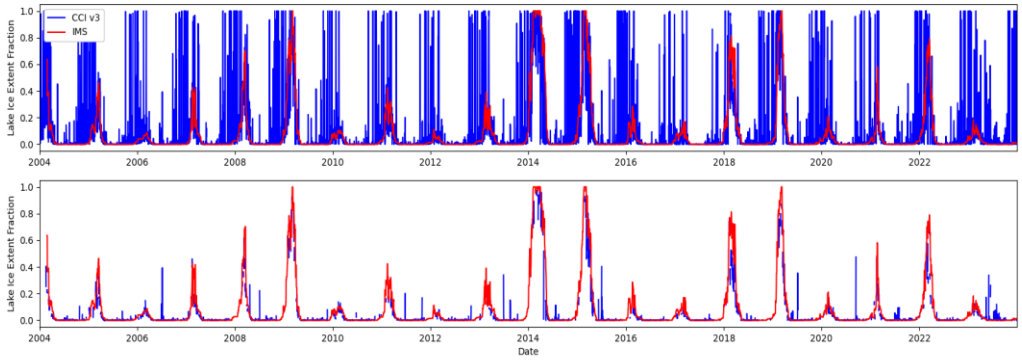


Figure 90: Comparison between Lakes_cci and IMS ice cover extent for Lake Superior. The top plot shows a comparison between all data, and the bottom plot shows a comparison with Lakes_cci data where observations are removed when the lake is >70% obstructed.

6.5 Conclusion and recommendations

For Lakes_cci CDRP V3.0, LIC is generated from a RF algorithm using MODIS Terra/Aqua Calibrated Radiance 5-Min L1B Swath (MOD02/MYD02), Collection 6.1, as primary data input. Lakes_cci LIC is a gridded product where cells are assigned one of three possible labels: water, ice, or cloud.

Validation of the Lakes_cci LIC product was performed via computation of confusion matrices built on independent statistical validation. Reference data for validation of class labels (water, ice, cloud) were collected from three dataset of samples collected through visual interpretation of the false color composite surface reflectance images (MOD02/MYD02) by skilled ice analysts. The reference data also served to validate lake ice/water/cloud retrievals contained in NASA MODIS Terra/Aqua Snow Cover products (MOD10/MYD10), which were then compared to those from Lakes_cci LIC. Additional comparisons were also made between the CCI LIC and operational IMS data.

Results show that the retrieval algorithm (RF) implemented for Lakes_cci LIC production provides robust classification of lake ice cover. Retrieval accuracies are also found to be more consistent between classes and also ice periods for the Lakes_cci algorithm. Moreover, RF also produced comparable classification results from MODIS Terra and Aqua. Individual class accuracies are all above 90% (errors less than 10%) for LIC CDRP V3.0 which meet uncertainty requirements of 10% set by GCOS for LIC ECV (see E3UB

document for details on determination of errors and uncertainty in LIC product). Additionally, according to visual inspection, compared to NASA's MODIS snow cover product (MYD10/MOD10), the Lakes_cci LIC product performs more accurately in several difficult classification cases illustrated in section 6.4. Furthermore, comparisons between CCI LIC and IMS ice cover extent showed average MAE values <0.1 days which meets uncertainty requirements set by GCOS as mentioned above. Visual comparison of the ice extent fractions demonstrates that the CCI LIC product can capture standard ice progression (full freezing and melting of the lake) as well as capturing patterns where a lake partially freezes and melts.

Continued assessment of the Lakes_cci LIC product and its comparison with ice products other than NASA's MOD10/MYD10 and IMS is planned in the future. This includes comparison with: 1) nearshore in situ ice/open water observations if such observations are available for any of the 2,024 lakes forming ice cover ; 2) other operational ice products such as the Copernicus Lake Ice Extent product; and 3) weekly ice fraction product from the Canadian Ice Service. Ice cover observations from these various sources will be valuable for further assessing the quality of future LIC CDRP versions.

6.6 References

Heinilä, K., Metsämäki, S., & Mattila, O. P. (2017). Lake ice extent–algorithm theoretica l basis document. *Copernicus Global Land Operations "Cryosphere and Water*, 22.

National Oceanic and Atmospheric Administration (2022). IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions, Version 1 – User Guide. https://nsidc.org/sites/default/files/g02156-v001-userguide_1_1.pdf

Ramsay, B. H. (1998). The interactive multisensor snow and ice mapping system. *Hydrological Processes*, *12*(10-11), 1537-1546.

Riggs, G., Hall, D.K. & Román, M.O. (2019). MODIS Snow Products Collection 6.1 User Guide, Version 1.0. 66 p.

Wu, Y., Duguay, C.R. & Xu, L. (2021). Assessment of machine learning classifiers for global lake ice cover mapping from MODIS TOA reflectance data. Remote Sensing of Environment, 253, 112206, https://doi.org/10.1016/j.rse.2020.112206.

7 Lake Ice Thickness - LIT

7.1 Data description

7.1.1 Radar altimetry data

The radar altimetry data used to generate the LIT time series are Low Resolution Mode (LRM) Ku-band waveform data at 20 Hz from Jason-1, Jason-2 and Jason-3 missions. Since the satellites follow the same orbit, they provide consistent coverage, allowing for robust LIT estimations over time.

The LIT dataset includes 14 lakes including Amadjuak Lake (passes 43 and 72), Lake Athabasca (pass 95), Baker Lake (pass 19), Dubawnt Lake (pass 19), Gauer Lake (pass 152), Great Bear Lake (passes 24, 100, and 225), Reindeer Lake (passes 76 and 145), Schultz Lake (pass 121), South Henik Lake (pass 202), Southern Indian Lake (pass 152), Tulemalu Lake (pass 24), Wharton Lake (pass 176), and Lake Winnipeg (pass 195) (Figure 91). In addition, Great Slave Lake is included, as was already the case in the previous product release.

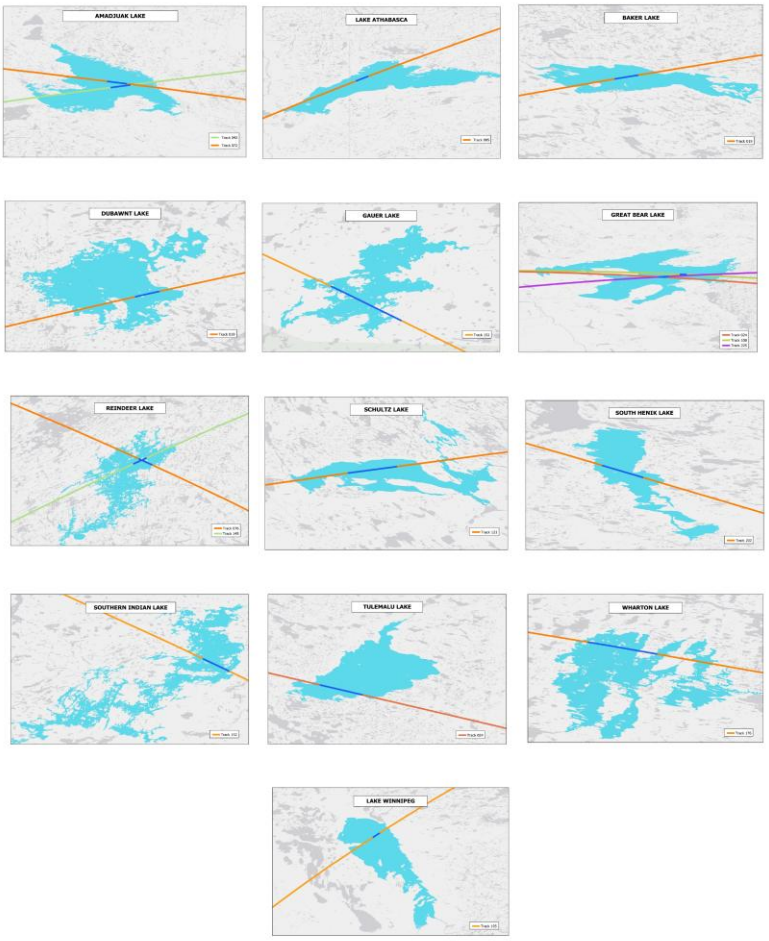


Figure 91: 13 of the lakes with corresponding Jason passes.

Each of the ROI along track segment used for the LIT analysis spans ~0.1° in latitude, including about one hundred radar waveforms, within a distance of ≃ 35 km along the altimeter track. The emplacement of each ROI has been chosen to be far enough from the coast, so that the radar echoes are not contaminated by spurious signals from land, while the length of the segment has been chosen to have

enough statistics for the LIT estimation while limiting the spatial coverage and thus the spatial evolution of the LIT within the segment in order to get consistent constraints, in particular at the seasonal transitions. The LIT time series products have been generated by analysing the radar waveform data with the LIT retracker described in Mangilli et al. (2022).

7.1.2 CLIMo simulations

The Canadian Lake Ice Model (CLIMo) is a one-dimensional thermodynamic model that can use weather station observations, atmospheric reanalysis data, or gridded products from climate models as inputs. CLIMo has been designed to simulate lake ice phenology, ice thickness, and all components of the radiation and energy balance during both the ice and open water seasons (Duguay et al., 2003). CLIMo was executed with varying percentages of snow cover on the Great Slave Lake (GSL), specifically at 25%, 50%, and 75% of the snow depth measured at the Hay River weather (land) station (southwest shore of GSL) and with a mean snow density of 300 kg m^{-3}. The smaller percentages are meant to consider the redistribution of snow on the ice surface of the large open area of the lake.

Simulations of the Lake Ice Thickness (LIT) were conducted near tracks 45, 178, and 254, respectively, covering the time span from 2002 to 2022 at daily time steps. It is important to note that two sets of LIT simulations were performed on GSL. As tracks 45 and 254 overlap, one CLIMo simulation was obtained in the overlapping region, while the second simulation was extracted from the highlighted portion within track 178.

The same CLIMo configuration and modeling framework were subsequently applied to the 13 most recently added lakes included in the extended LIT dataset. Simulations were performed using 0% and 100% snow cover conditions, along with the other snow cover scenarios tested for GSL.

7.1.3 In situ measurements

In situ LIT measurements are available for Back Bay, a small bay in Great Slave Lake located by the city of Yellowknife, for the period between November 2002 and April 2016 (no data available since then). The data were obtained from the Canadian Ice Service portal (https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/latest-conditions/archive-overview/thickness-data.html). These in situ measurements are made annually on a weekly basis, with data collection restricted to times when the ice is safe for walking. The measurements cover the ice season, beginning shortly after freeze-up and extending until the ice melts.

Similar to GSL, in situ measurements were available but only for one of the additional study lakes, Baker Lake. LIT data for Baker Lake were obtained from the YBK station and span the entire study period from 2002 to 2022, providing a continuous record for model validation.

7.1.4 Image data

MODIS/Landsat and Sentinel-1 SAR images were obtained to visualize and better understand the impact of lake surface conditions at or near the time altimetry data acquisitions. Multiple images were strategically chosen to provide a comprehensive view and ensure the availability of diverse validation data. This approach enhanced our ability to validate and interpret the LIT time series generated from the Jason series, particularly in identifying cases where retrieval errors were found to be large.

7.2 Comparison methods

The LIT time series generated from the Jason series (altimetry) is compared with the LIT values obtained from CLIMo simulations (Duguay et al. 2003). For a “qualitative” assessment, the LIT product is also compared with available in situ measurements. Note that the meteorological stations at which the measurements are collected could be located several tens of km away from the location of the altimeter

tracks where retrievals were performed. Hence, some differences in LIT are expected due to site location, where lake depth and snow regime are different (Mangilli et al., 2022). To better understand the results and to ensure consistency in retrievals, the LIT product undergoes a quality check using optical (MODIS/Landsat) and SAR (Sentinel-1) images.

As part of the validation process, the quality of the altimetry-derived LIT data was first examined through a series of checks, including verification of value ranges, assessment of quality flags, and evaluation of data reliability to ensure consistency across all observations. The temporal continuity of the LIT time series was also reviewed to identify gaps or anomalous variations. These quality control steps ensured that only reliable and representative data were used for subsequent validation against CLIMo simulations, in situ measurements, and optical/SAR imagery. LIT data with flag 0 (=good data) were used for the validation analysis. We note that flag 2 data, which indicate degraded fit performances typically resulting in increased uncertainty values beyond the typical validity range of flag 0 data ($0.35<\sigma_LIT<1$m), were not included in the validation. These represent few data points, specifically identified during ice growing/melting seasonal transitions, and excluding them has not impact on the validation results.

7.3 Description of the work

Figure 92 shows the LIT time series obtained at the GSL. The validation of the LIT product is carried out through three different methods.

The first method involves comparing the LIT time series derived from altimetry data with the LIT simulations generated using CLIMo. This validation encompasses the entire LIT time series, as data generated from CLIMo simulations are available for the full study period.

The second method consists of comparing the LIT product with in situ measurements obtained from the Canadian Ice Service (CIS). To reiterate, it is important to note that, in the case of Great Slave Lake, the Back Bay site, where the measurements are collected, is located several km away from the Region of Interest (ROI) for all three tracks. However, tracks 45 and 254 are relatively closer to Back Bay compared to track 178. Given the significant variations in depth across GSL, the comparison of in situ measurements with the LIT product should be viewed as a qualitative comparison, rather quantitative and robust validation method.

The third method involves conducting a consistency check of the LIT product by comparing it with MODIS/Sentinel-1/Landsat images on specific dates over the GSL through visual interpretation. This is done to ensure that the LIT product accurately captures the ice phenology conditions, especially the formation and melting of ice, as well as transitions, times when Jason retrievals may show larger errors.

For the 13 most recently added lakes in the LIT dataset, the same validation approach was applied, focusing primarily on comparisons between altimetry-derived LIT and CLIMo simulations. For Baker Lake, additional validation with in situ measurements were also performed due to the availability of data. Complementing these comparisons, statistical summaries were generated to evaluate the performance of the LIT product against both CLIMo simulations and in situ observations, providing a comprehensive assessment of its accuracy across these lakes.

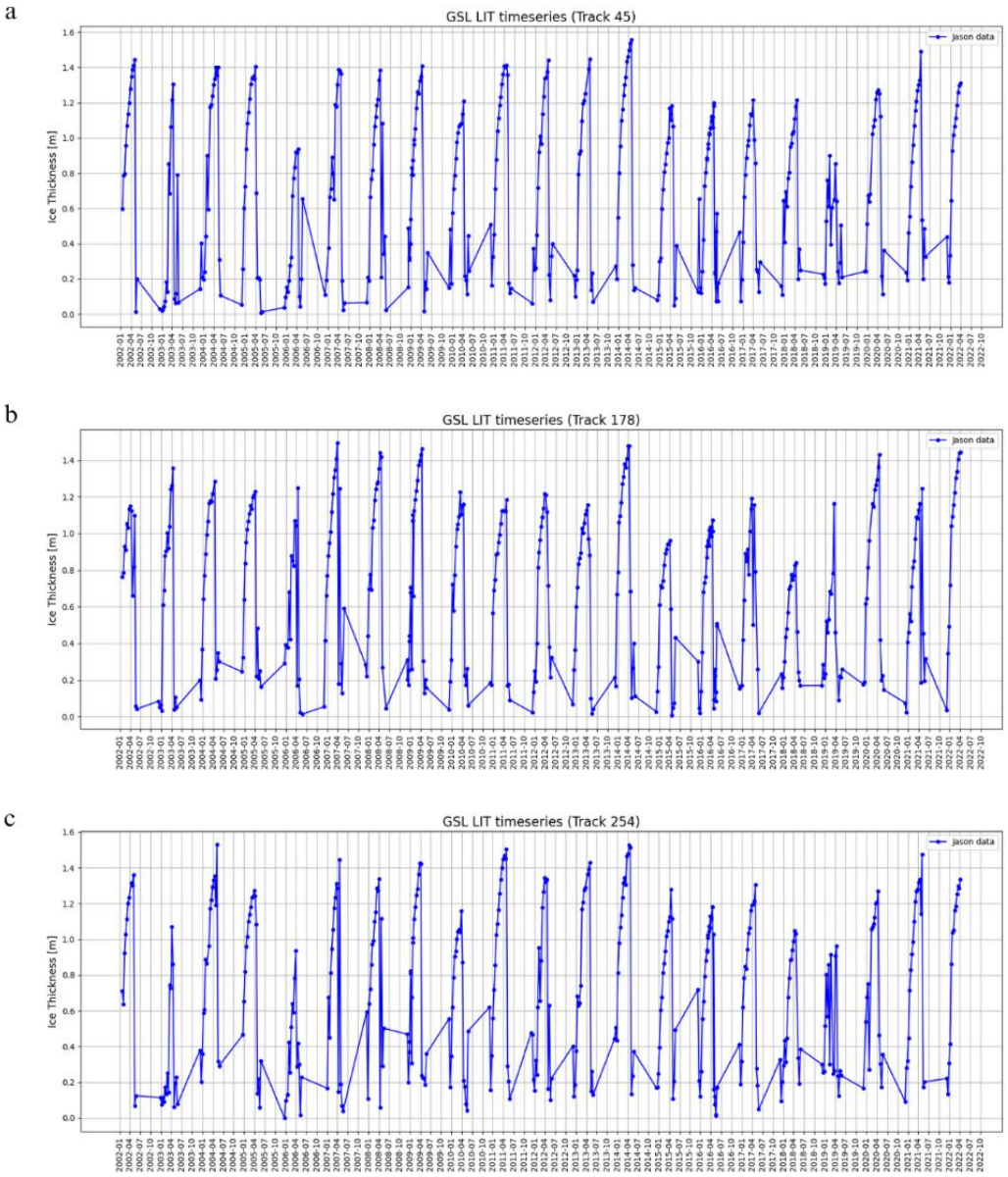


Figure 92: LIT time series on the GSL a) Track 45, b) Track 178 and c) Track 254

7.4 Validation results

7.4.1 Comparison with CLIMo

The following analysis compares LIT obtained from Jason time series data with simulations from the lake ice model (CLIMo). The scatter plots visually illustrate the match between these datasets, aiming to reveal patterns and discrepancies (Figure 93, Figure 94 and Figure 95).

Upon examination of the scatterplots and the corresponding tables that provide the statistical details (per ice year and for all years combined) for tracks 45, 178, and 254 (Table 30, Table 31 and Table 32), a strong agreement is observed between satellite altimetry data and lake ice model simulations. However, LIT product values are consistently lower than those obtained from CLIMo simulations, especially during

the early freeze and melt periods. This underestimation pattern persists across tracks and time series. However, seasonal variations in LIT are evident, highlighting the effectiveness of the satellite altimetry data in capturing temporal changes.

In all three tracks, LIT product values closely align with thicknesses obtained from CLIMo simulations with comparatively less snow cover (25%). Additionally, in a few instances, the LIT product recorded maximum thicknesses larger than CLIMo, which indicates that the percentage of snow on the ice surface may be lower than 25% in some years. A lower snow amount on the ice promotes faster/greater ice growth. However, without the availability of in situ snow depth measurements (and other snow properties such as density, layering and grain size of layers) on the ice surface along or near the Jason altimeter tracks, it is not possible to draw a definitive conclusion as to whether the retrieved maximum LIT values are at times too large or CLIMo simulated values too low.

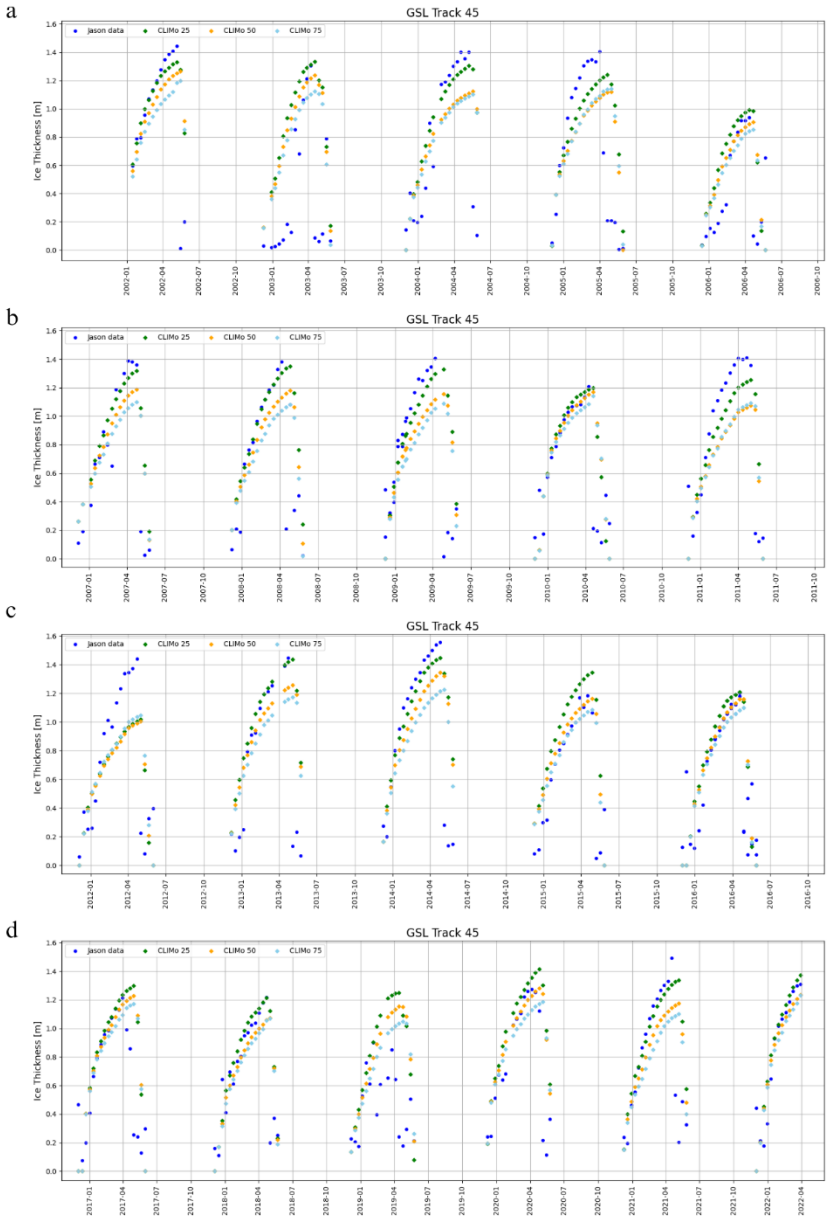


Figure 93: Comparison of LIT estimates with CLIMo simulation over GSL (Track 45) during the period a) 2002-2006, b) 2007-2011, c) 2012-2016 and d) 2017-2022. Dark blue circles represent the altimetry-based LIT product, and green, yellow and sky-blue circles represent CLIMo simulations with varying snow cover percentages on the ice surface (25%, 50%, and 75%)

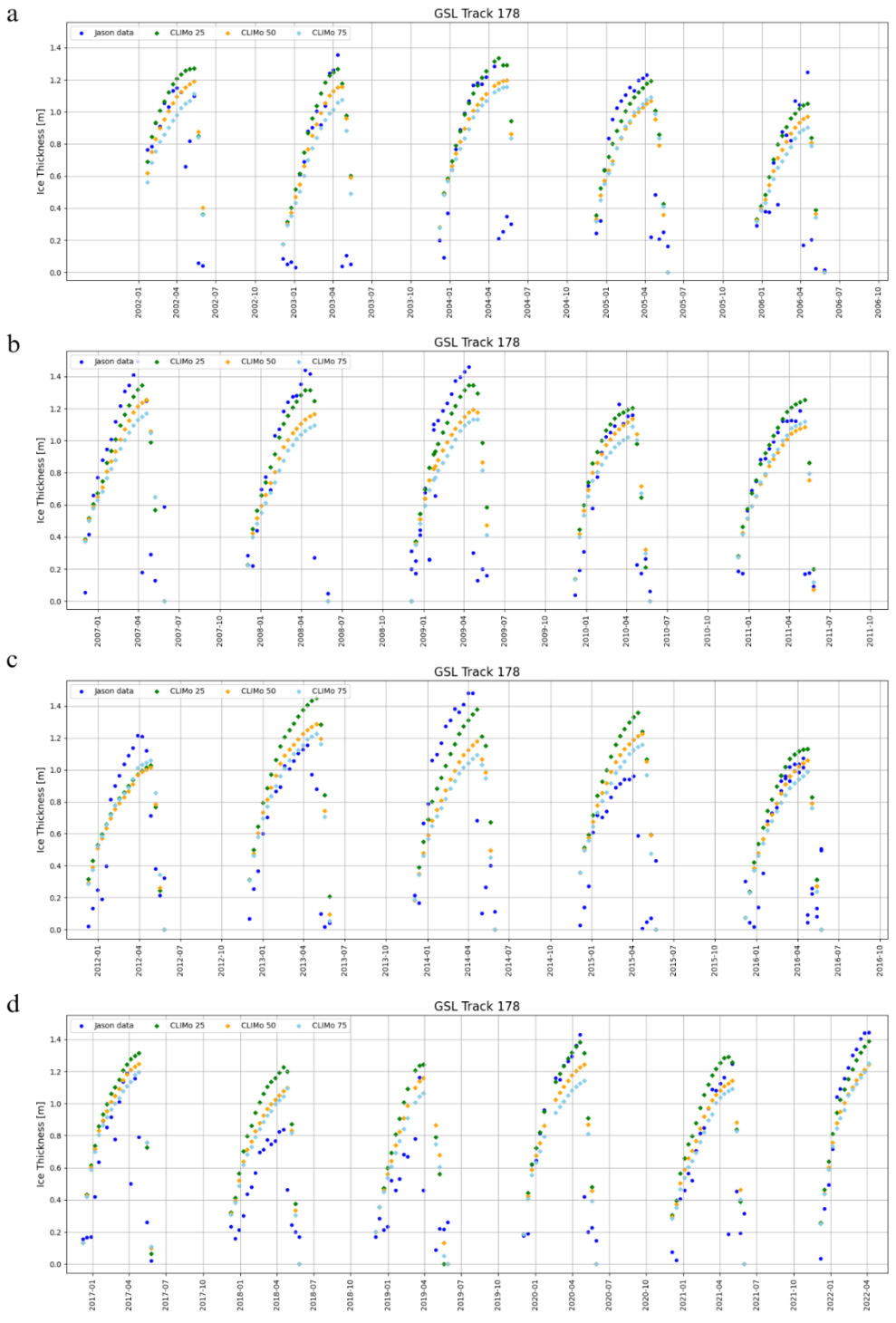


Figure 94: Comparison of LIT estimates with CLIMo simulation over GSL (Track 178) during the period a) 2002-2006, b) 2007-2011, c) 2012-2016 and d) 2017-2022. Dark blue circles represent the altimetry-based LIT product, and green, yellow and sky-blue circles represent CLIMo simulations with varying snow cover percentages on the ice surface (25%, 50%, and 75%)

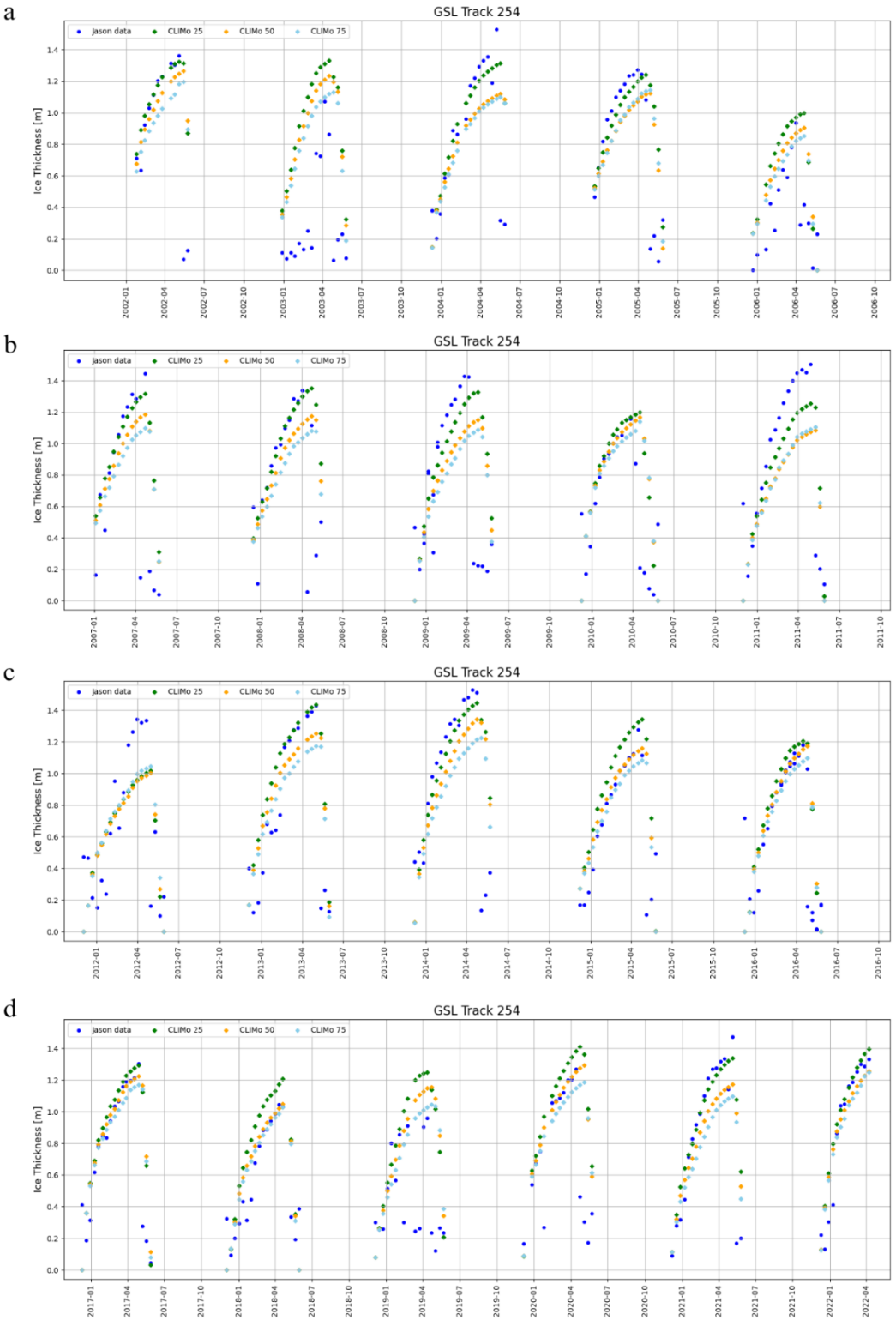


Figure 95: Comparison of LIT estimates with CLIMo simulation over GSL (Track 254) during the period a) 2002-2006, b) 2007-2011, c) 2012-2016 and d) 2017-2022. Dark blue circles represent the altimetry-based LIT product, and green, yellow and sky-blue circles represent CLIMo simulations with varying snow cover percentages on the ice surface (25%, 50%, and 75%).

Table 30: Comparison of the LIT product and the lake ice model (CLIMo) simulations for the Track 45

Season	Data	LIT Max (m)	Date of LIT Max	LIT Mean (m)	MBE (m)	RMSE (m)
2001-2002	Jason	1.443	2002-05-06	1.112	-	-
	CLIMo 25	1.329	2002-05-06	1.111	0.001	0.056
	CLIMo 50	1.268	2002-05-15	1.019	0.094	0.112
	CLIMo 75	1.198	2002-05-15	0.933	0.179	0.195
2002-2003	Jason	1.306	2003-04-08	0.776	-	-
	CLIMo 25	1.333	2003-04-18	1.164	-0.388	0.501
	CLIMo 50	1.239	2003-04-18	1.063	-0.287	0.431
	CLIMo 75	1.124	2003-04-18	0.969	-0.193	0.386
2003-2004	Jason	1.402	2004-05-08	1.139	-	-
	CLIMo 25	1.304	2004-05-08	1.127	0.012	0.163
	CLIMo 50	1.123	2004-05-18	0.968	0.17	0.248
	CLIMo 75	1.102	2004-05-18	0.94	0.198	0.261
2004-2005	Jason	1.406	2005-04-01	1.235	-	-
	CLIMo 25	1.243	2005-04-20	1.13	0.105	0.283
	CLIMo 50	1.121	2005-04-30	1.013	0.222	0.342
	CLIMo 75	1.143	2005-04-30	1.025	0.21	0.338
2005-2006	Jason	0.937	2006-04-12	0.706	-	-
	CLIMo 25	0.992	2006-04-12	0.871	-0.166	0.224
	CLIMo 50	0.908	2006-04-22	0.769	-0.063	0.167
	CLIMo 75	0.855	2006-04-22	0.714	-0.008	0.158
2006-2007	Jason	1.388	2007-04-04	1.097	-	-
	CLIMo 25	1.32	2007-04-24	1.124	-0.027	0.167
	CLIMo 50	1.188	2007-04-24	1.014	0.083	0.19
	CLIMo 75	1.101	2007-04-24	0.938	0.16	0.239
2007-2008	Jason	1.384	2008-04-05	1.06	-	-
	CLIMo 25	1.352	2008-04-25	1.177	-0.118	0.4
	CLIMo 50	1.18	2008-04-25	1.028	0.031	0.374
	CLIMo 75	1.084	2008-04-25	0.938	0.121	0.39
2008-2009	Jason	1.408	2009-04-07	1.259	-	-
	CLIMo 25	1.33	2009-04-27	1.139	0.12	0.124
	CLIMo 50	1.155	2009-04-27	0.99	0.269	0.271
	CLIMo 75	1.09	2009-04-27	0.919	0.341	0.342
2009-2010	Jason	1.208	2010-04-09	1.082	-	-
	CLIMo 25	1.198	2010-04-19	1.117	-0.034	0.044
	CLIMo 50	1.171	2010-04-19	1.066	0.016	0.027
	CLIMo 75	1.142	2010-04-19	1.01	0.072	0.076

2010-2011	Jason	1.412	2011-04-21	1.286	-	-
	CLIMo 25	1.257	2011-05-01	1.092	0.194	0.195
	CLIMo 50	1.076	2011-05-01	0.932	0.354	0.355
	CLIMo 75	1.094	2011-05-01	0.941	0.345	0.345
2011-2012	Jason	1.441	2012-04-22	1.166	-	-
	CLIMo 25	1.019	2012-05-01	0.863	0.302	0.313
	CLIMo 50	1.005	2012-05-01	0.845	0.32	0.331
	CLIMo 75	1.047	2012-05-01	0.878	0.287	0.297
2012-2013	Jason	1.448	2013-04-23	1.179	-	-
	CLIMo 25	1.438	2013-05-03	1.219	-0.04	0.059
	CLIMo 50	1.256	2013-05-03	1.078	0.101	0.117
	CLIMo 75	1.175	2013-05-03	0.991	0.188	0.196
2013-2014	Jason	1.557	2014-04-25	1.372	-	-
	CLIMo 25	1.446	2014-04-25	1.284	0.088	0.089
	CLIMo 50	1.346	2014-04-25	1.163	0.209	0.209
	CLIMo 75	1.227	2014-05-05	1.056	0.316	0.317
2014-2015	Jason	1.183	2015-04-17	0.973	-	-
	CLIMo 25	1.346	2015-04-27	1.158	-0.185	0.19
	CLIMo 50	1.162	2015-04-27	1.01	-0.037	0.062
	CLIMo 75	1.087	2015-04-27	0.936	0.038	0.062
2015-2016	Jason	1.2	2016-04-18	1.002	-	-
	CLIMo 25	1.207	2016-04-18	1.089	-0.086	0.088
	CLIMo 50	1.16	2016-04-18	1.016	-0.013	0.025
	CLIMo 75	1.101	2016-04-28	0.952	0.051	0.055
2016-2017	Jason	1.214	2017-03-31	1.07	-	-
	CLIMo 25	1.299	2017-04-30	1.135	-0.065	0.109
	CLIMo 50	1.228	2017-04-30	1.074	-0.005	0.084
	CLIMo 75	1.173	2017-04-30	1.019	0.051	0.101
2017-2018	Jason	1.216	2018-04-22	0.98	-	-
	CLIMo 25	1.212	2018-04-22	1.037	-0.057	0.066
	CLIMo 50	1.074	2018-05-02	0.901	0.08	0.09
	CLIMo 75	1.069	2018-05-02	0.865	0.115	0.122
2018-2019	Jason	0.9	2019-02-03	0.613	-	-
	CLIMo 25	1.25	2019-04-14	1.135	-0.522	0.592
	CLIMo 50	1.152	2019-04-14	1.018	-0.405	0.494
	CLIMo 75	1.047	2019-04-24	0.913	-0.3	0.405
2019-2020	Jason	1.273	2020-04-05	1.171	-	-
	CLIMo 25	1.414	2020-04-25	1.262	-0.091	0.096

	CLIMo 50	1.281	2020-04-25	1.151	0.02	0.036
	CLIMo 75	1.187	2020-05-05	1.06	0.11	0.116
2020-2021	Jason	1.492	2021-04-17	1.184	-	-
	CLIMo 25	1.338	2021-05-06	1.125	0.059	0.06
	CLIMo 50	1.175	2021-05-06	0.988	0.196	0.197
	CLIMo 75	1.101	2021-05-06	0.901	0.283	0.284
2021-2022	Jason	1.31	2022-03-30	1.206	-	-
	CLIMo 25	1.373	2022-03-30	1.249	-0.043	0.045
	CLIMo 50	1.235	2022-03-30	1.14	0.066	0.069
	CLIMo 75	1.232	2022-03-30	1.108	0.098	0.1
Overall	Jason	1.557	2014-04-25	0.71	-	-
	CLIMo 25	1.446	2014-04-25	0.853	-0.143	0.382
	CLIMo 50	1.346	2014-04-25	0.782	-0.072	0.368
	CLIMo 75	1.232	2022-03-30	0.74	-0.031	0.36

Table 31: Comparison of the LIT product and the lake ice model (CLIMo) simulations for the Track 178

Season	Data	LIT Max (m)	Date of LIT Max	LIT Mean (m)	MBE (m)	RMSE (m)
2001-2002	Jason	1.151	2002-04-01	1.015	-	-
	CLIMo 25	1.271	2002-05-11	1.073	-0.059	0.069
	CLIMo 50	1.191	2002-05-11	0.964	0.051	0.063
	CLIMo 75	1.112	2002-05-11	0.871	0.144	0.149
2002-2003	Jason	1.356	2003-04-13	1.076	-	-
	CLIMo 25	1.267	2003-04-13	1.114	-0.038	0.095
	CLIMo 50	1.155	2003-04-23	0.997	0.079	0.115
	CLIMo 75	1.075	2003-04-23	0.903	0.172	0.193
2003-2004	Jason	1.285	2004-04-14	1.155	-	-
	CLIMo 25	1.334	2004-04-24	1.157	-0.002	0.032
	CLIMo 50	1.196	2004-05-13	1.038	0.117	0.119
	CLIMo 75	1.156	2004-05-13	0.991	0.163	0.165
2004-2005	Jason	1.229	2005-04-06	1.157	-	-
	CLIMo 25	1.193	2005-04-16	1.077	0.08	0.085
	CLIMo 50	1.067	2005-04-16	0.96	0.197	0.198
	CLIMo 75	1.091	2005-04-16	0.976	0.18	0.183
2005-2006	Jason	1.248	2006-04-17	0.751	-	-
	CLIMo 25	1.052	2006-04-17	0.939	-0.187	0.365
	CLIMo 50	0.972	2006-04-17	0.85	-0.099	0.33
	CLIMo 75	0.904	2006-04-17	0.781	-0.029	0.318
2006-2007	Jason	1.495	2007-03-31	1.153	-	-

	CLIMo 25	1.344	2007-04-09	1.204	-0.051	0.457
	CLIMo 50	1.253	2007-04-19	1.11	0.044	0.452
	CLIMo 75	1.169	2007-04-19	1.037	0.116	0.46
2007-2008	Jason	1.441	2008-04-10	1.235	-	-
	CLIMo 25	1.316	2008-04-10	1.156	0.079	0.084
	CLIMo 50	1.168	2008-04-30	1.002	0.232	0.235
	CLIMo 75	1.095	2008-04-30	0.932	0.303	0.305
2008-2009	Jason	1.461	2009-04-12	1.312	-	-
	CLIMo 25	1.345	2009-04-12	1.184	0.128	0.129
	CLIMo 50	1.194	2009-04-22	1.046	0.266	0.267
	CLIMo 75	1.133	2009-04-22	0.978	0.334	0.334
2009-2010	Jason	1.227	2010-03-15	1.093	-	-
	CLIMo 25	1.205	2010-04-14	1.13	-0.037	0.055
	CLIMo 50	1.137	2010-04-14	1.046	0.047	0.064
	CLIMo 75	1.088	2010-04-14	0.966	0.126	0.136
2010-2011	Jason	1.186	2011-04-26	1.037	-	-
	CLIMo 25	1.255	2011-05-06	1.076	-0.039	0.045
	CLIMo 50	1.087	2011-05-06	0.924	0.113	0.115
	CLIMo 75	1.118	2011-05-06	0.948	0.089	0.092
2011-2012	Jason	1.216	2012-03-28	1.079	-	-
	CLIMo 25	1.03	2012-04-27	0.894	0.185	0.189
	CLIMo 50	1.017	2012-04-27	0.872	0.207	0.21
	CLIMo 75	1.059	2012-04-27	0.903	0.176	0.178
2012-2013	Jason	1.156	2013-04-09	1.053	-	-
	CLIMo 25	1.449	2013-04-29	1.288	-0.235	0.236
	CLIMo 50	1.288	2013-04-29	1.154	-0.102	0.105
	CLIMo 75	1.225	2013-04-29	1.086	-0.033	0.042
2013-2014	Jason	1.479	2014-04-01	1.384	-	-
	CLIMo 25	1.377	2014-04-21	1.206	0.179	0.184
	CLIMo 50	1.178	2014-04-21	1.032	0.352	0.353
	CLIMo 75	1.092	2014-04-21	0.962	0.422	0.423
2014-2015	Jason	0.961	2015-04-03	0.851	-	-
	CLIMo 25	1.357	2015-04-13	1.212	-0.362	0.395
	CLIMo 50	1.226	2015-04-22	1.093	-0.242	0.283
	CLIMo 75	1.157	2015-04-22	0.025	-0.174	0.228
2015-2016	Jason	1.074	2016-04-13	0.949	-	-
	CLIMo 25	1.131	2016-04-23	1.036	-0.087	0.092
	CLIMo 50	1.059	2016-04-23	0.934	0.016	0.038

	CLIMo 75	0.987	2016-04-23	0.865	0.085	0.091
2016-2017	Jason	1.191	2017-03-26	0.942	-	-
	CLIMo 25	1.316	2017-04-25	1.168	-0.225	0.316
	CLIMo 50	1.247	2017-04-25	1.109	-0.167	0.276
	CLIMo 75	1.195	2017-04-25	1.053	-0.111	0.243
2017-2018	Jason	0.839	2018-04-17	0.728	-	-
	CLIMo 25	1.227	2018-04-17	1.087	-0.358	0.359
	CLIMo 50	1.099	2018-04-27	0.953	-0.225	0.227
	CLIMo 75	1.095	2018-04-27	0.917	-0.189	0.191
2018-2019	Jason	1.163	2019-03-20	0.751	-	-
	CLIMo 25	1.244	2019-03-30	1.157	-0.406	0.466
	CLIMo 50	1.161	2019-03-30	1.059	-0.308	0.386
	CLIMo 75	1.065	2019-03-30	0.973	-0.222	0.321
2019-2020	Jason	1.43	2020-04-20	1.245	-	-
	CLIMo 25	1.381	2020-04-20	1.252	-0.007	0.022
	CLIMo 50	1.244	2020-04-30	1.12	0.125	0.127
	CLIMo 75	1.144	2020-04-30	1.029	0.216	0.217
2020-2021	Jason	1.246	2021-05-02	0.976	-	-
	CLIMo 25	1.29	2021-04-22	1.119	-0.144	0.147
	CLIMo 50	1.143	2021-05-02	0.972	0.004	0.05
	CLIMo 75	1.092	2021-05-02	0.901	0.074	0.085
2021-2022	Jason	1.444	2022-04-04	1.329	-	-
	CLIMo 25	1.388	2022-04-04	1.255	0.074	0.075
	CLIMo 50	1.245	2022-04-04	1.135	0.194	0.196
	CLIMo 75	1.249	2022-04-04	1.119	0.21	0.212
Overall	Jason	1.495	2007-03-31	0.692	-	-
	CLIMo 25	1.449	2013-04-29	0.864	-0.172	0.36
	CLIMo 50	1.288	2013-04-29	0.788	-0.097	0.337
	CLIMo 75	1.249	2022-04-04	0.749	-0.058	0.329

Table 32: Comparison of the LIT product and the lake ice model (CLIMo) simulations for the Track 254

Season	Data	LIT Max (m)	Date of LIT Max	LIT Mean (m)	MBE (m)	RMSE (m)
2001-2002	Jason	1.362	2002-05-04	1.064	-	-
	CLIMo 25	1.326	2002-05-04	1.104	-0.04	0.1
	CLIMo 50	1.266	2002-05-14	1.013	0.051	0.111
	CLIMo 75	1.196	2002-05-14	0.928	0.136	0.176
2002-2003	Jason	1.071	2003-04-06	0.463	-	-

	CLIMo 25	1.33	2003-04-16	1.151	-0.688	0.732
	CLIMo 50	1.235	2003-04-16	1.05	-0.587	0.639
	CLIMo 75	1.132	2003-04-26	0.957	-0.494	0.56
2003-2004	Jason	1.53	2004-05-07	1.14	-	-
	CLIMo 25	1.315	2004-05-16	1.117	0.023	0.079
	CLIMo 50	1.121	2004-05-16	0.96	0.18	0.204
	CLIMo 75	1.099	2004-05-16	0.933	0.207	0.225
2004-2005	Jason	1.274	2005-03-30	1.202	-	-
	CLIMo 25	1.241	2005-04-19	1.122	0.08	0.084
	CLIMo 50	1.122	2005-04-28	1.007	0.195	0.196
	CLIMo 75	1.143	2005-04-28	1.02	0.182	0.184
2005-2006	Jason	0.937	2006-04-01	0.571	-	-
	CLIMo 25	0.997	2006-04-20	0.891	-0.32	0.379
	CLIMo 50	0.905	2006-04-20	0.788	-0.217	0.298
	CLIMo 75	0.853	2006-04-20	0.732	-0.16	0.261
2006-2007	Jason	1.445	2007-04-22	0.996	-	-
	CLIMo 25	1.317	2007-04-22	1.114	-0.118	0.409
	CLIMo 50	1.185	2007-04-22	1.005	-0.009	0.385
	CLIMo 75	1.098	2007-04-22	0.928	0.067	0.387
2007-2008	Jason	1.338	2008-04-03	1.022	-	-
	CLIMo 25	1.35	2008-04-23	1.167	-0.146	0.453
	CLIMo 50	1.176	2008-04-23	1.02	0.002	0.418
	CLIMo 75	1.08	2008-04-23	0.931	0.09	1.425
2008-2009	Jason	1.427	2009-03-26	1.16	-	-
	CLIMo 25	1.329	2009-04-25	1.153	0.006	0.412
	CLIMo 50	1.152	2009-04-25	1.002	0.158	0.43
	CLIMo 75	1.087	2009-04-25	0.931	0.229	0.462
2009-2010	Jason	1.158	2010-03-28	1.015	-	-
	CLIMo 25	1.198	2010-04-17	1.112	-0.097	0.135
	CLIMo 50	1.168	2010-04-17	1.061	-0.046	0.108
	CLIMo 75	1.139	2010-04-17	1.006	0.009	0.094
2010-2011	Jason	1.505	2011-04-29	1.309	-	-
	CLIMo 25	1.254	2011-04-29	1.083	0.226	0.228
	CLIMo 50	1.086	2011-05-09	0.925	0.384	0.386
	CLIMo 75	1.104	2011-05-09	0.934	0.374	0.376
2011-2012	Jason	1.344	2012-03-31	1.085	-	-
	CLIMo 25	1.02	2012-04-30	0.878	0.208	0.273
	CLIMo 50	1.004	2012-04-30	0.86	0.226	0.285

	CLIMo 75	1.046	2012-04-30	0.894	0.191	0.254
2012-2013	Jason	1.429	2013-05-02	1.098	-	-
	CLIMo 25	1.436	2013-05-02	1.224	-0.126	0.211
	CLIMo 50	1.254	2013-05-02	1.082	0.016	0.184
	CLIMo 75	1.173	2013-05-02	0.996	0.101	0.205
2013-2014	Jason	1.529	2014-04-14	1.351	-	-
	CLIMo 25	1.445	2014-04-24	1.274	0.078	0.089
	CLIMo 50	1.344	2014-04-24	1.152	0.199	0.203
	CLIMo 75	1.227	2014-05-03	1.047	0.304	0.307
2014-2015	Jason	1.278	2015-04-16	0.986	-	-
	CLIMo 25	1.342	2015-04-25	1.148	-0.162	0.162
	CLIMo 50	1.159	2015-04-25	1.001	-0.015	0.027
	CLIMo 75	1.084	2015-04-25	0.929	0.058	0.063
2015-2016	Jason	1.182	2016-04-16	1.01	-	-
	CLIMo 25	1.204	2016-04-16	1.09	-0.08	0.082
	CLIMo 50	1.172	2016-04-26	1.017	-0.007	0.018
	CLIMo 75	1.099	2016-04-26	0.953	0.057	0.06
2016-2017	Jason	1.306	2017-04-28	1.061	-	-
	CLIMo 25	1.296	2017-04-28	1.127	-0.066	0.074
	CLIMo 50	1.225	2017-04-28	1.068	-0.007	0.04
	CLIMo 75	1.171	2017-04-28	1.013	0.048	0.069
2017-2018	Jason	1.047	2018-04-10	0.886	-	-
	CLIMo 25	1.207	2018-04-20	1.058	-0.172	0.175
	CLIMo 50	1.05	2018-04-20	0.918	-0.032	0.051
	CLIMo 75	1.029	2014-04-20	0.882	0.004	0.041
2018-2019	Jason	0.961	2019-04-12	0.635	-	-
	CLIMo 25	1.249	2019-04-12	1.129	-0.494	0.606
	CLIMo 50	1.157	2019-04-22	1.012	-0.376	0.511
	CLIMo 75	1.045	2019-04-22	0.908	-0.273	0.435
2019-2020	Jason	1.271	2020-04-13	1.145	-	-
	CLIMo 25	1.41	2020-04-23	1.253	-0.108	0.112
	CLIMo 50	1.293	2020-05-03	1.144	0.001	0.024
	CLIMo 75	1.186	2020-05-03	1.054	0.091	0.093
2020-2021	Jason	1.474	2021-05-05	1.177	-	-
	CLIMo 25	1.338	2021-05-05	1.136	0.041	0.048
	CLIMo 50	1.174	2021-05-05	0.998	0.179	0.182
	CLIMo 75	1.099	2021-05-05	0.914	0.263	0.265
2021-2022	Jason	1.334	2022-04-07	1.225	-	-

	CLIMo 25	1.398	2022-04-07	1.261	-0.036	0.045
	CLIMo 50	1.255	2022-04-07	1.149	0.076	0.078
	CLIMo 75	1.248	2022-04-07	1.121	0.104	0.107
Overall	Jason	1.53	2004-05-07	0.717	-	-
	CLIMo 25	1.445	2014-04-24	0.881	-0.164	0.39
	CLIMo 50	1.344	2014-04-24	0.808	-0.091	0.374
	CLIMo 75	1.248	2022-04-07	0.765	-0.048	0.363

For the 13 most recently added lakes, validation was performed following the same approach as for GSL. To provide a representative overview, validation plots are shown for only Baker Lake (pass 19) and Great Bear Lake (pass 100). These plots illustrate the comparison between Jason based LIT, CLIMo simulations. ERA5 Variable LIT is also included as an additional reference, providing further context for assessing the performance of the altimetry-derived LIT product across different lake conditions (Figure 93 & 94).

Similar to the observations for GSL, GBL also shows a strong alignment between Jason-derived LIT and CLIMo simulation trends (Figure 93). The best-matching CLIMo snow cover scenario varies between ice seasons, suggesting that the lake likely experienced different amounts of snow accumulation each year. During the early freeze-up and melt periods, Jason-derived LIT values approach zero and tend to underestimate the ice thickness, a pattern consistently observed across all time periods. Nevertheless, when examining the full time series, the Jason-derived LIT closely follows the CLIMo simulations and effectively captures the seasonal evolution of lake ice thickness.

A similar trend and underestimation pattern are also observed for Baker Lake (Figure 94). Interestingly, the in situ measurements align well with the Jason-derived LIT and further support the algorithm's capability to retrieve realistic ice thickness values across different lake and climatic conditions.

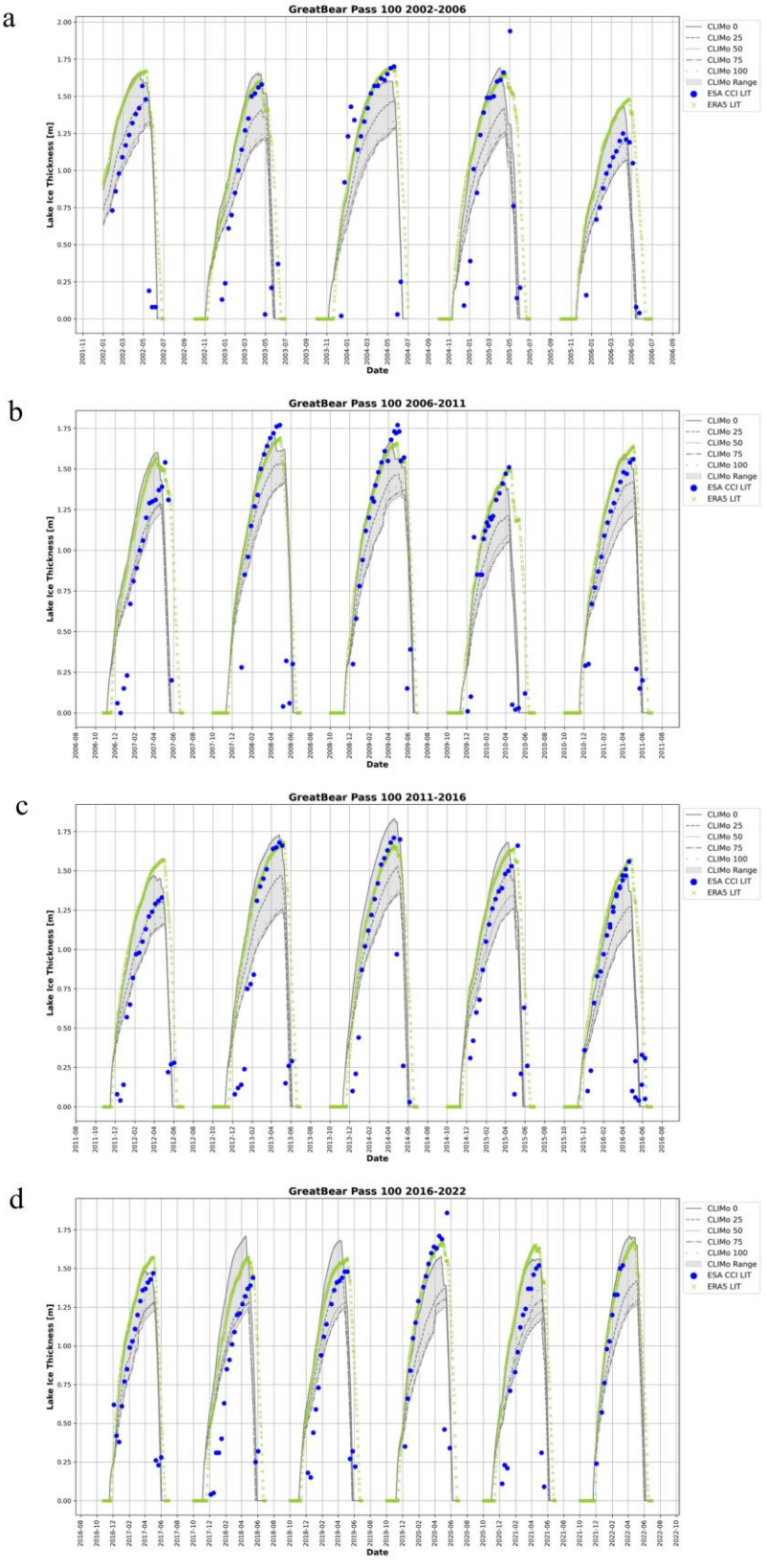


Figure 96: Comparison of Jason-derived LIT estimates with CLIMo simulations and ERA5 LIT over GBL (pass 100) for the periods: (a) 2002–2006, (b) 2007–2011, (c) 2012–2016, and (d) 2017–2022. Dark blue circles represent the altimetry-derived LIT product, while the grey-shaded area indicates the range of CLIMo simulations, with individual lines corresponding to different snow cover percentages. Green squares denote LIT values obtained from ERA5.

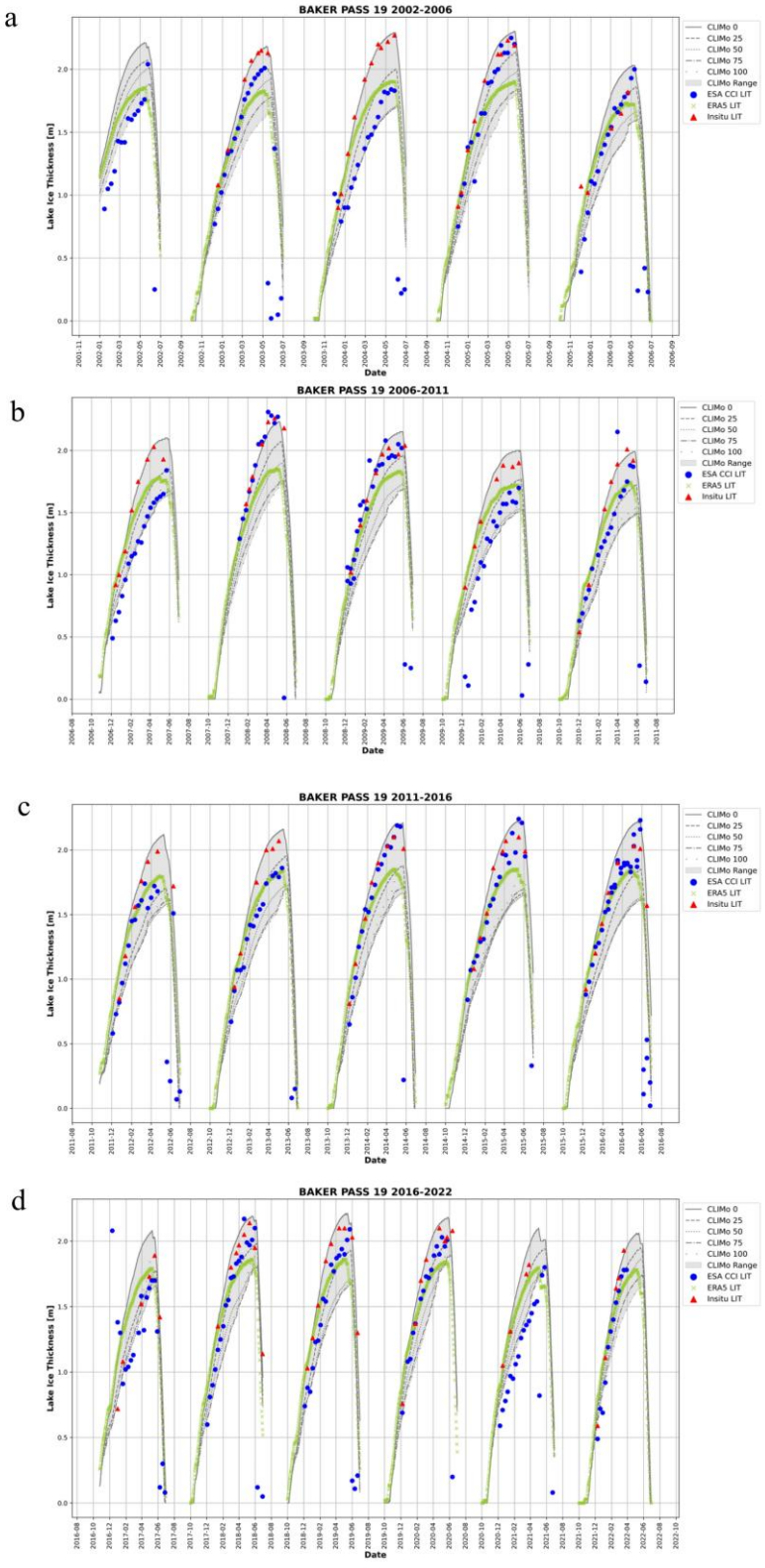


Figure 97: Comparison of Jason-derived LIT estimates with CLIMo simulations and ERA5 LIT over Baker Lake (pass 19) for the periods: (a) 2002–2006, (b) 2007–2011, (c) 2012–2016, and (d) 2017–2022. Dark blue circles represent the altimetry-derived LIT product, while the grey-shaded area indicates the range of CLIMo simulations, with individual lines corresponding to different snow cover percentages. Green squares denote LIT values obtained from ERA5, and red triangles represent in situ LIT measurements.

A single summary table (Table 37) presents overall statistics for all 13 additional lakes, including key metrics comparing altimetry-derived LIT with CLIMo simulations. This provides a concise yet comprehensive overview of the LIT algorithm's performance across the expanded dataset. Examination of the RMSE values indicates that they are generally large for all lakes, with considerable variability across the dataset. This is primarily due to underestimation of Jason-derived LIT during the early freeze-up and melt periods. To further illustrate this, Table 38 presents the comparison for Great Bear Lake and Baker Lake, considering only the fully ice-covered period and excluding early freeze-up and melt. The RMSE values show substantial improvement compared to those in Table 37. For example, the full ice-season RMSE range for Baker Lake is 0.07–0.58 m, whereas excluding the early freeze-up and melt periods reduces the range to 0.02–0.20 m. These results demonstrate that accounting for the surface state of the lake ice, such as whether it is newly formed or undergoing melt, allows the LIT algorithm to select the most reliable data and improves the accuracy of lake ice thickness retrievals across all lakes.

Table 37: Comparison of the LIT product and the lake ice model (CLIMo) simulations for the additional 13 study lakes.

Lake	Pass	MBE range (m)	RMSE range (m)	Best CLIMo Count (per season)
Amadjuak Lake	43	0.00 to 0.32	0.29 to 0.53	CLIMo 100: 7; CLIMo 25: 1; CLIMo 50: 1; CLIMo 75: 10
Amadjuak Lake	72	-0.06 to 0.25	0.04 to 0.63	CLIMo 100: 8; CLIMo 25: 2; CLIMo 50: 1; CLIMo 75: 9
Lake Athabasca	95	0.02 to 0.24	0.13 to 0.46	CLIMo 100: 13; CLIMo 25: 4; CLIMo 50: 2; CLIMo 75: 2
Baker Lake	19	-0.08 to 0.18	0.07 to 0.58	CLIMo 0: 2; CLIMo 100: 3; CLIMo 25: 10; CLIMo 50: 4; CLIMo 75: 2
Dubawnt Lake	19	-0.03 to 0.65	0.06 to 0.99	CLIMo 100: 18; CLIMo 50: 2; CLIMo 75: 1
Gauer Lake	152	0.06 to 0.47	0.10 to 0.60	CLIMo 100: 1; CLIMo 25: 1; CLIMo 50: 8; CLIMo 75: 4
Great Bear Lake	24	-0.07 to 0.43	0.21 to 0.57	CLIMo 100: 16; CLIMo 25: 1; CLIMo 50: 1; CLIMo 75: 3
Great Bear Lake	100	-0.08 to 0.13	0.13 to 0.52	CLIMo 0: 3; CLIMo 100: 5; CLIMo 25: 10; CLIMo 50: 2; CLIMo 75: 1
Great Bear Lake	225	-0.00 to 0.28	0.10 to 0.59	CLIMo 100: 13; CLIMo 25: 2; CLIMo 50: 1; CLIMo 75: 5
Reindeer Lake	76	-0.08 to 0.33	0.08 to 0.64	CLIMo 0: 1; CLIMo 100: 3; CLIMo 25: 2;

				CLIMo 50: 10; CLIMo 75: 5
Reindeer Lake	145	-0.13 to 0.33	0.25 to 0.45	CLIMo 0: 1; CLIMo 100: 4; CLIMo 25: 2; CLIMo 50: 12; CLIMo 75: 2
Schultz Lake	121	-0.26 to 0.58	0.11 to 0.65	CLIMo 0: 1; CLIMo 100: 1; CLIMo 25: 13; CLIMo 50: 5; CLIMo 75: 1
South Henik Lake	202	-0.06 to 0.08	0.07 to 0.85	CLIMo 0: 1; CLIMo 25: 7; CLIMo 50: 3; CLIMo 75: 2
Southern Indian Lake	152	-0.05 to 0.41	0.11 to 0.55	CLIMo 100: 16; CLIMo 25: 1; CLIMo 50: 2; CLIMo 75: 2
Tulemalu Lake	24	-0.20 to 0.91	0.13 to 0.98	CLIMo 100: 1; CLIMo 25: 9; CLIMo 50: 7; CLIMo 75: 2
Wharton Lake	176	-0.19 to 0.12	0.17 to 0.50	CLIMo 0: 1; CLIMo 100: 2; CLIMo 25: 10
Lake Winnipeg	195	-0.02 to 0.19	0.07 to 0.43	CLIMo 0: 3; CLIMo 100: 2; CLIMo 25: 9; CLIMo 50: 6; CLIMo 75: 1

Table 38: Comparison of the LIT product and the lake ice model (CLIMo) simulations for Baker Lak and GBL considering only the fully ice-covered period and excluding early freeze-up and melt phases.

Lake	Pass	MBE range (m)	RMSE range (m)	Best CLIMo Count (per season)
Baker Lake	19	-0.19 to 0.17	0.02 to 0.20	CLIMo 0: 4; CLIMo 100: 4; CLIMo 25: 11; CLIMo 50: 2
Great Bear Lake	24	-0.03 to 0.48	0.03 to 0.55	CLIMo 100: 7; CLIMo 25: 1; CLIMo 50: 7; CLIMo 75: 6
Great Bear Lake	100	-0.11 to 0.11	0.05 to 0.28	CLIMo 0: 7; CLIMo 100: 2; CLIMo 25: 10; CLIMo 50: 2
Great Bear Lake	225	-0.05 to 0.10	0.03 to 0.27	CLIMo 100: 7; CLIMo 25: 4; CLIMo 50: 8; CLIMo 75: 2

7.4.2 Comparison with in situ measurements

In situ measurements were available from the Canadian Ice Service only for the period 2002-2016, not for the entire LIT time series. Similar to the comparison with CLIMo simulations, the LIT product shows a strong agreement with in situ measurements (Figure 98). The temporal trajectory and magnitude of the ice growth are well captured by the LIT product.

In most ice seasons, maximum ice thickness measurements obtained in situ are less than those from the LIT product. This discrepancy may be attributed to the location of the Yellowknife station (Back Bay, GSL), from which the in situ measurements were obtained. This station is situated in a relatively shallow region (~10 m; Ménard et al., 2022) compared to the Region of Interest (ROI) of tracks 45, 178, and 254 (20-50 m). The timing of ice formation (freeze-up) and ice growth are influenced by lake depth and the snow regime, respectively, so that differences are expected. Therefore, the maximum thickness obtained by the LIT product appears reliable.

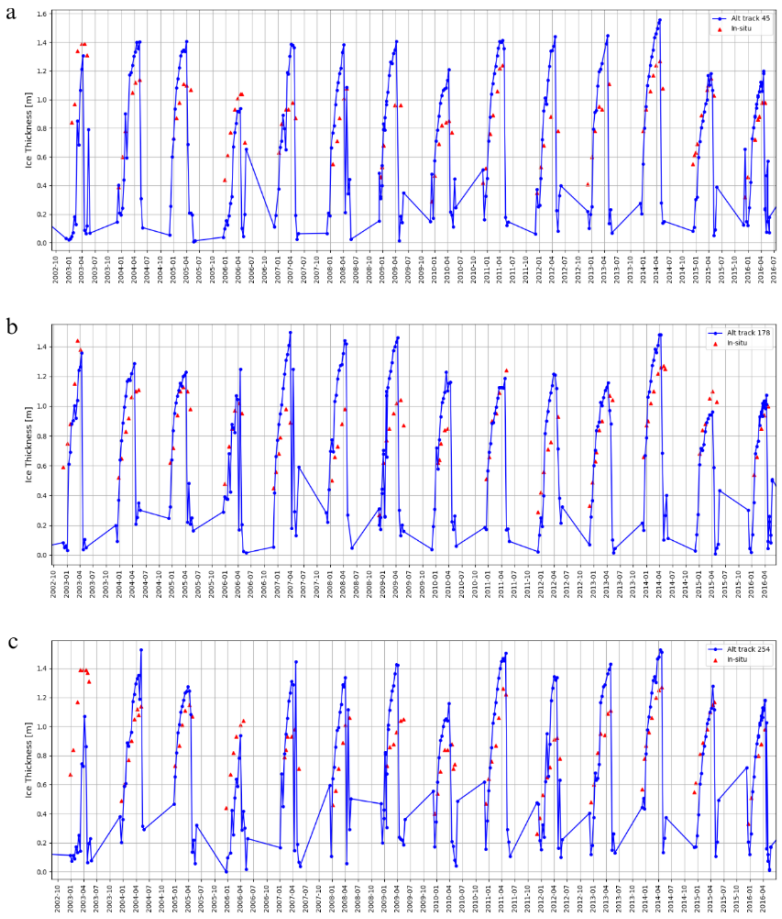


Figure 98: LIT product vs in situ measurements (GSL) Comparison with image data

From the comparison plots (Figure 93 - Figure 98) and the statistical tables (Table 30 - Table 32), it is clear that most LIT values generated from altimetry during the ice growth period are close to those from CLIMo simulations, following a similar trend of thickness evolution. However, upon closer examination, two anomalous years were identified where the altimetry-based thickness estimations were unusually low and irregular.

The first anomaly occurred during the 2018-2019 ice season and was observed in tracks 45 and 254. Interestingly, this anomaly was not observed in track 178. It is possible that tracks 45 and 254, being closer to each other, share a similarity that sets them apart from track 178, which is farther away. Further investigation revealed that CLIMo did not follow this anomaly pattern; instead, it adhered to the usual trend of increasing LIT from the start of ice formation until breakup (Figure 99). The possible reason behind the 2019 anomaly may be the absence or minimal presence of snow over the ice in that location. Visual interpretation of the optical and SAR images acquired over the 2018-2019 ice season also supports this possibility because the upper-middle part of the GSL (where tracks 45 and 254 are located) often shows no snow or thin snow. Since the LIT estimation method relies on double backscattering (from snow-ice and ice-water interfaces), such a situation with no or minimal snow leads to the underestimation of the LIT. Figure 99 displays a few examples of images illustrating the aforementioned snow cover scenarios. Although 2019 is considered an anomalous year, the LIT estimates from track 178 were close to the CLIMo simulations. From the satellite images, it is evident that there was always snow in that area (track 178). Thus, the LIT algorithm performs well on ice with the presence of snow (see Figure 99).

The second anomaly was identified during the 2002-2003 ice season. Among the three tracks, only the LIT estimates from track 254 were unusually low compared to the CLIMo simulations. Throughout the ice season, the Mean Bias Error (MBE) and Root Mean Square Error (RMSE) between the LIT product and CLIMo simulations ranged from -0.494 to -0.688 and 0.56 to 0.732, respectively (Table 32). Upon investigating the ice season using satellite images, it was revealed that a substantial lead lay directly perpendicular to track 254 and remained consistently present for several weeks during the ice season. Since altimetry signals are sensitive to open water (leads) or a mixed footprint of ice and water, the algorithm could not accurately retrieve LIT. The mentioned lead was situated somewhat farther from the location of track 45 (Region of Interest), explaining the absence of a similar anomaly in its LIT estimates.

In both cases of anomalies in the LIT time series, the LIT estimates from track 178 closely matched the CLIMo simulations, as evident in Figure 94 and Table 31. This demonstrates that LIT estimates can be generated from the altimetry product, aligning quantitatively with CLIMo simulations and qualitatively with in situ measurements as long as there is the presence of snow on the ice surface and the absence of open water areas (leads) within the altimeter footprint.

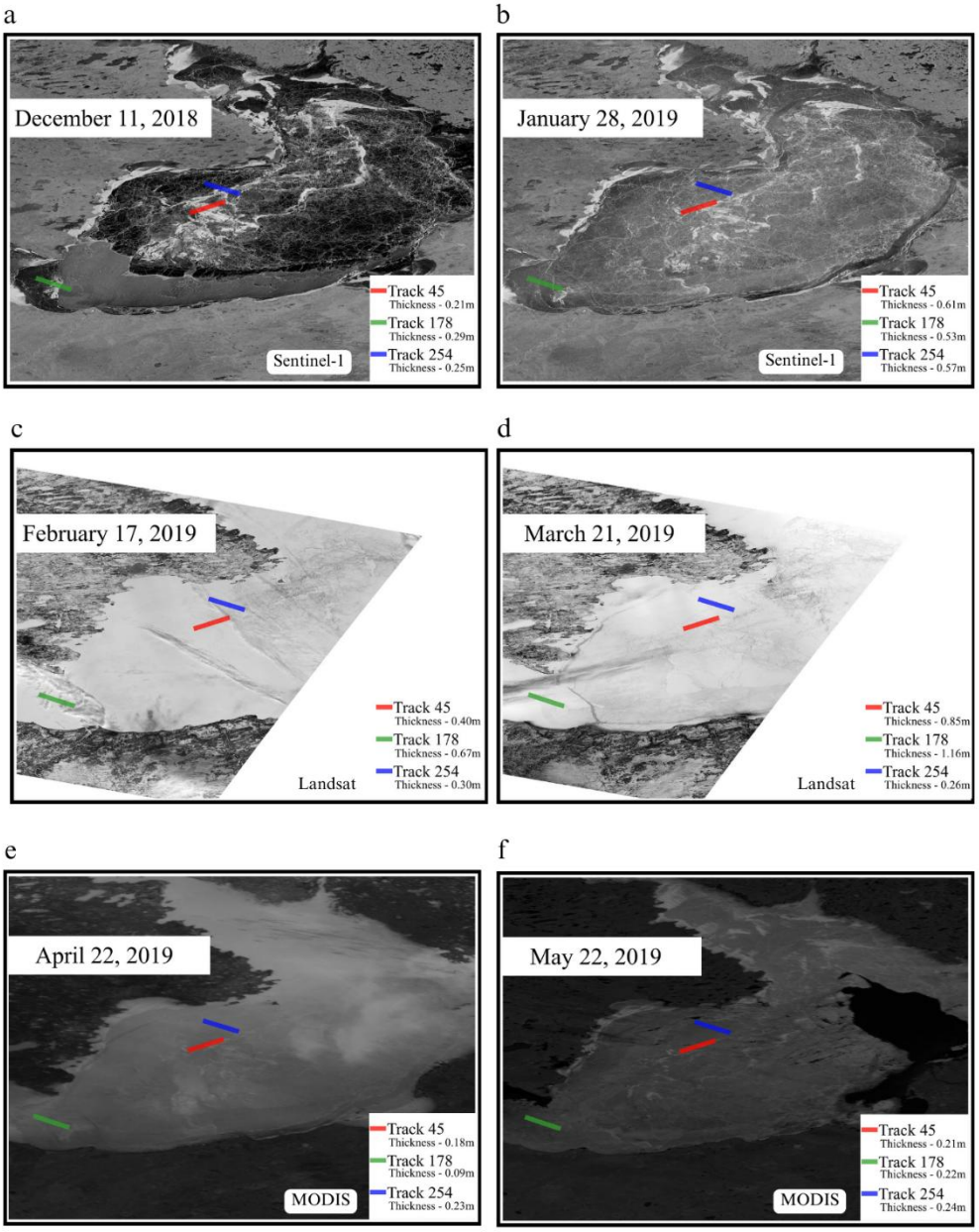


Figure 99: LIT estimates (2018 – 2019) superimposed on the corresponding satellite images

In Figure 100, we present an illustrative example of an ice season (2020-2021) where the retrieval of LIT from altimetry closely matches CLIMo simulations. This specific case serves as a demonstration of the consistent and accurate nature of the LIT estimates derived from altimetric data at Great Slave Lake.

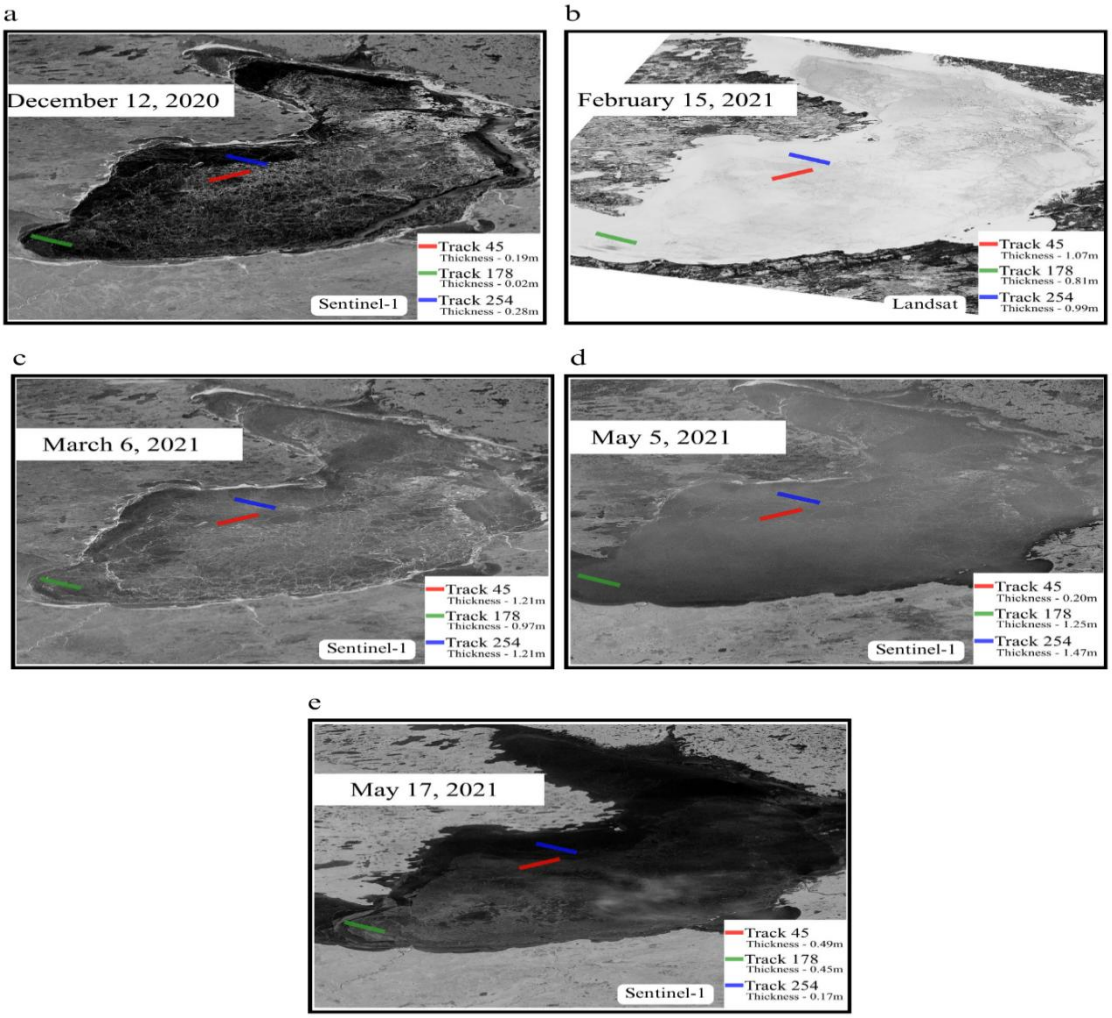


Figure 100: LIT estimates (2020 – 2021) superimposed on the corresponding satellite images

7.5 Conclusion and recommendations

The validation and intercomparison of the Lake Ice Thickness (LIT) product with lake ice model simulations and available in situ measurements demonstrate the capability of the LIT retrieval algorithm to generate reliable and consistent estimates of LIT across multiple lakes. The analysis of 13 additional lakes confirms that the algorithm performs robustly beyond the initial study site, capturing seasonal ice growth for diverse lake types and climatic conditions.

However, the results also reveal limitations in the LIT product, in particular at the seasonal transitions:

1. The LIT retracker faces challenges in precisely estimating ice thickness in the initial days of ice formation and once melt begins. This issue has been reported in Mangilli et al. (2022).
2. In the absence or minimal presence of snow on the ice, the step-like feature required to estimate LIT from the analytical approach (Mangilli et al., 2022) does not appear in the radar waveforms, leading to underestimated LIT retrievals.
3. The presence of leads or footprints with mixed open water and ice results in poor estimation of LIT values.

Despite these challenges, the LIT product successfully captures overall seasonal patterns of ice formation and melting, as supported by comparisons with CLIMo simulations, ERA5-derived LIT, and in situ

measurements at Back Bay (GSL) and YBK station (Baker Lake). However, the retracker cannot precisely follow ice evolution during the seasonal transitions due to the difficulty of retracking heterogeneous surfaces when the ice is very thin or when snow on the ice surface begins to melt. Statistical analyses show that RMSE values improve substantially when early freeze-up and melt periods are excluded, highlighting the importance of considering the ice surface state in data selection

Given that the majority of underestimation occurs at the start of freeze-up and after the onset of melt, it would be beneficial to implement a pre-processing step to eliminate error-inducing data points. Mugunthan et al. (2023) recently introduced a machine learning-based classification approach to flag lake surfaces using altimetry waveform parameters and backscatter coefficients. This classification method categorizes altimetry data points as young ice, growing ice, open water, or melting ice. In line with the needs of the LIT retrieval algorithm, the classification pre-processing step could inform the algorithm about which points to consider during the LIT retrieval procedure.

Mugunthan et al. (2023) achieved high and consistent results in their accuracy assessments for 11 lakes across the northern hemisphere, with overall classification accuracy exceeding 95% and spatial cross-validation accuracy surpassing 93.7%. It is worth noting that the mentioned work is based on SAR altimetry data, and efforts to implement the methodology with Low Resolution Mode (LRM) data are needed. Within this context, future analysis of Ku-band SAR (Unfocused SAR and Fully Focused SAR) data could further improve the LIT retrievals because of the improved spatial resolution [Mangili et al. (2024)] and should be considered to further expand and enhance the LIT timeseries and targets.

Overall, the extended validation demonstrates that the altimetry-derived LIT product provides robust estimates across multiple lakes and time periods, while targeted improvements during ice formation and melt can further enhance its accuracy and applicability for monitoring lake ice thickness at regional and global scales.

7.6 References

Duguay, C. R., Flato, G. M., Jeffries, M. O., Ménard, P., Morris, K., & Rouse, W. R. (2003). Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. Hydrological Processes, 17(17), 3465-3483.

Mangilli, A., Thibaut, P., Duguay, C.R., & Murfitt, J. (2022). A new approach for the estimation of lake ice thickness from conventional radar altimetry. IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-15, 2022, Art no. 4305515, doi: 10.1109/TGRS.2022.3186253.

Mangilli, A., Duguay, C. R., Murfitt, J., Moreau, T., Amraoui, S., Mugunthan, J. S., Thibaut, P., & Donlon, C. (2024). Improving the estimation of lake ice thickness with high-resolution radar altimetry data. Remote Sensing, 16(14), 2510. https://doi.org/10.3390/rs16142510

Ménard, P., Duguay, C.R., Flato, G.M., & Rouse, W.R. (2002). Simulation of ice phenology on Great Slave Lake, Northwest Territories, Canada. Hydrological Processes, 16(18): 3691-3706.

Mugunthan, J. S., Duguay, C. R., & Zakharova, E. (2023). Machine learning based classification of lakeice and open water from Sentinel-3 SAR altimetry waveforms. Remote Sensing of Environment, 299, 113891.

8 Lake Storage Change - LSC

8.1 Data Description

The Lake Storage Change (LSC) product is derived from two Essential Climate Variables (ECVs) used as priors: Lake Water Level (LWL) and Lake Water Extent (LWE). Consequently, the accuracy and consistency of LSC estimates are inherently dependent on the quality of these underlying datasets.

Globally, in situ LSC datasets are extremely scarce, which poses significant challenges for validating satellite-derived estimates with the level of precision required. Only a limited number of in situ observations are available, and these have been used to validate the overall methodology and processing pipeline on a lake-by-lake basis, as detailed in the following sections.

The comparison methodology consists of pre-processing in situ lake volume time-series to derive lake storage change (LSC) time-series aligned with altimetry-derived data. Statistical comparison is based on linear regression, using the coefficient of determination (R^2), root mean square error (RMSE), and bias. The approach assumes lakes with variable surface area as water level changes; for lakes with constant surface area, HydroLakes data will be validated against the GSWO dataset before producing time-series. Additionally, minima and maxima from LWL and LWE are used to define appropriate validation windows.

8.2 Description of the work

Four lakes were selected according to ins-situ data availability: Richland-Chambers (USA), Rosarito (Spain), Songhua (China), and Tres Marias (Brazil). Complete lake volume time-series were only available for Rosarito. For Richland-Chambers, Songhua, and Tres Marias, established area–height–volume relationships allow direct conversion of altimetry-derived height into volume estimates. In the case of Tres Marias, these relationships were provided by the Brazilian national water resources agency. Additional bathymetric data were available for Lakes Garda, Laguna de Bay, and Mille Lacs, supporting further validation and analysis.

8.3 Validation Results

Of the four available validation datasets, we measured good agreement between our estimates and the in situ LSC time-series, with R^2 from 0.78 to 0.97 and an excellent replication of the temporal trends. However, the number of freely accessible dataset of LSC worldwide limits validation on a larger reference dataset as mentioned previously. That is why for production purposes we consider our LSC estimation methodology valid, and rely absolutely on the quality of the LWL and LWE production teams.

In the end, comparison between the in-situ observations and satellite-derived LSC are as displayed in Figure 101 and Table 33.

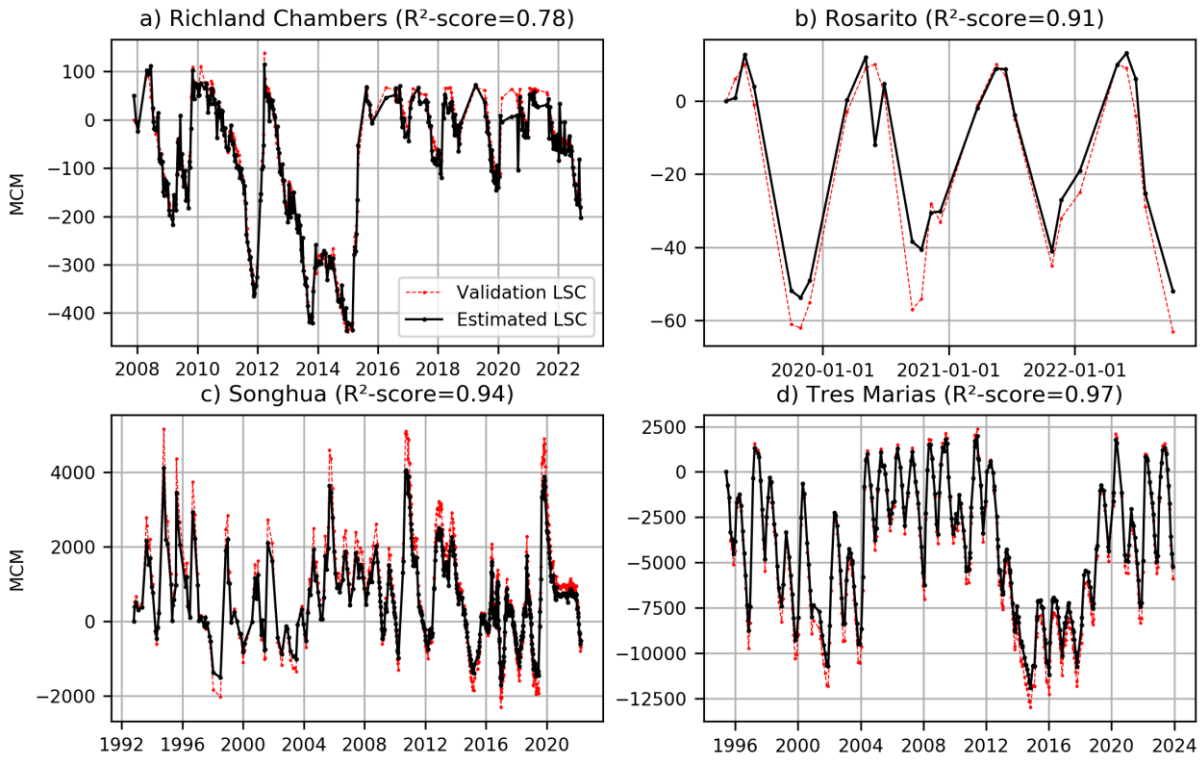


Figure 101 : Validation of LSC time-series with in-situ data. R2 is included for each lake

Table 33 : Statistics on LSC time-series comparison with in-situ data

Lake name	R²	RMSD (MCM)	Bias (MCM)
Richland-Chambers	0.78	63.17	58.36
Rosarito	0.91	7.76	-3.32
Songhua	0.94	361.67	130.84
Tres Marias	0.97	662.26	-4939

8.4 Conclusions and recommendation

The comparison process highlights limitations primarily related to the availability and quality of in situ datasets. In many cases, little information is provided about how these datasets were produced or their accuracy, which can introduce uncertainties. For example, the lack of information about the region of interest of in-situ datasets may lead to underestimation or overestimation of the LSC error. Another error source can be the differences of height system reference between computed LSC and the one of in-situ data we compare these with.

Differences are observed more often on extrema: for instance, on the Songhua Lake, for high height values the estimated LSC is lower than the one derived from in situ values. This could be linked to an offset between the two height references used for the height-area-volume curve. At the opposite, the estimated LSC is higher at low water height stages, mainly during winter season. This might be linked to errors in the water surface estimate due to ice cover and potential bias induced in water height derived from altimetry. For the Richland-Chambers water body, an offset at lower values might be linked to border effects in the water body contour from Sentinel-2. Thus, product uncertainties have various sources (height referential, ice cover, border effects…) whose impacts is mainly visible on extrema.

Based on these findings, the following recommendations are proposed. To improve the quality of this product, priority should be given to integrating numerical uncertainties from LWL and LWE into the hypsometric model for lakes with varying surface areas. For lakes with constant surface areas, a pre-analysis and correction of the mean lake area should be implemented to reduce bias.

Since LSC estimates are derived from LWL and LWE, users are strongly encouraged to examine these source datasets and their associated uncertainties to better interpret the results. LSC uncertainty is provided as a qualitative classification (low, medium, high), which should be considered when using the data for scientific or operational purposes. The consideration of quality flag enhances the reliability of LSC products and support their effective use in climate and hydrological applications.

8.5 References

Crétaux, J.-F., Abarca-del-Río, R., Bergé-Nguyen, M., Arsen, A., Drolon, V., Clos, G., & Maisongrande, P. (2016). Lake Volume Monitoring from Space. *Surveys in Geophysics*, *37*(2), 269–305. https://doi.org/10.1007/s10712-016-9362-6

Magome, J., Ishidara, H., & Takeuchi, K. (2003). Method for satellite monitoring of water storage in reservoirs for efficient regional water management. IAHS Publ. no.281, pp.303-310.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*(1), 13603. https://doi.org/10.1038/ncomms13603

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540(7633), 418–422. https://doi.org/10.1038/nature20584

9 Consistency

The most recent product inter-consistency analysis is available in PVIR v2.1, available on the project website. Several results of that analysis have since been incorporated as quality controls on CRDP v3.0. The analysis of CRDP v3.0 will be included in the next version of the PVIR.

Appendix A - List Figures and Tables