Product Validation and Algorithm Selection Report (PVASR)

Reference: CCI-LAKES2-0011-PVASR

Issue 3.0.0 – 10/12/2025

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

For internal use

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 OF ISSUES

Issue	Legacy Issues	Date	Objective	Written by
1.0	1.0	19/03/2020	Initial Version	S. Simis, JF Crétaux, H Yésou, E Malnes, H Vickers, P Blanco, C Merchant, L Carrea, C Duguay, Y Wu
	1.1	24/04/2020	Revision following ESA review	See above
	2.0	30/07/2020	LSWT MODIS data processing	C Merchant
	2.1	29/09/2020	Revision following ESA review	JF Crétaux, C Merchant, H Yésou, E Malnes
2.x	2.2	02/05/2022	Updates accompanying CRDP v2.0 and v2.0.1	X Liu, JF Crétaux, L Carrea, C Duguay, S Simis
	2.3	06/07/2022	Revision following ESA review	B. Calmettes
2.1	-	06/09/2023	Revision and updates accompanying CRDP v2.1	S. Simis, B Calmettes, JF Crétaux, X Liu, A Mangili, H Yésou, L Carrea, C Duguay, Y Wu
2.1.1	-	17/10/2023	Revision following ESA review	S. Simis, B Calmettes, JF Crétaux, X Liu, A Mangili, H Yésou, L Carrea, C Duguay, Y Wu
3.0.0	-	10/12/2025	Documentation accompanying CRDP v3.0.0	S. Simis, B Calmettes, JF Crétaux, C. Fatras, X Liu, A Mangili, H Yésou, L Carrea, C Duguay, Y Wu, C.Fatras

Note: prior to document issue 2.1, the PVASR and ADP were provided in a combined document. From v2.1 onwards the two documents are provided separately and versioning is kept consistent with the CRDP release cycle.

BIBLIOGRAPHIC REFERENCE

This document is only intended for internal use.

DOCUMENT REVIEW

Checked by	Stefan Simis – PML	S. Simis
Approved by	Philippe Mourot - CLS	Pmourot
Authorized by	Clément Albergel - ESA	clement.albergel

LIST OF CONTENTS

1 Introduction

The Product Validation and Algorithm Selection Report (PVASR) reports the results of algorithm round-robin intercomparison as the basis for algorithm selection and further development during the Lakes_cci project. The PVASR is an internal document.

The purpose of the PVASR is to keep track of algorithm validation and the protocols used to carry out algorithm selection, particularly where these deviate from the Product Validation Plan (PVP).

The following elements for algorithm validation and selection are taken into consideration for each of the Lakes Essential Climate Variable products:

- How round-robin comparison was performed, including the satellite EO, ancillary and validation data sets employed.
- Specification of selection criteria to rank the results, and discussion of whether and how the selection criteria can be combined to provide an overall ranking.
- Specification of the included algorithms.
- Details of any harmonisation of the algorithms performed in advance to ensure unbiased results (e.g. common ancillary data sets).
- Results of the performance assessment of each individual algorithm.
- Ranking of the algorithms according to the results of the assessment. Different rankings may need to be derived corresponding to the different selection criteria or sets of associated selection criteria.
- Reasoning and justification of the overall ranking, including a discussion of user priorities considered when making any trade-offs in ranking.

The following sections detail these procedures for each respective Lakes ECV Product.

2 Lake Water level – LWL

2.1 Candidate algorithms

The algorithms for LWL calculation were developed at LEGOS and are detailed in the Algorithm Theoretical Basis Document (ATBD). Two algorithms are in use, based on the state-of-the-art in calculating LWL from satellite altimetry. The first algorithm uses the value of the range provided in the IGDR files (generated by agencies) while the Lake Physical Processing (LPP) algorithm estimates the range based on simulation of the waveform (Crétaux et al 2009, Boy et al 2022). Since each altimeter provides a distinct Global Data Record, an initial phase of organising the data and the geophysical corrections is required to produce a coherent climate data record. Moreover, each satellite mission presents a specific altimeter bias which requires correcting (based on published results), in order to arrive at a consistent long-term multi-satellite LWL time series.

The processing software is named *Hysope*. It can be used operationally and is based on Intermediate Geophysical Data Records (IGDRs) (delivery delay 1 to 2 days). It is operated at CLS in the framework of the Hydroweb database. A version for non-operational use also runs at LEGOS and is based on the same equations but using Geophysical Data Records (GDRs) instead of IGDRs (Crétaux et al. 2016).

The procedure is run against data within a priori defined polygons of lake contours (using the common dataset of maximum water extent outlines created for Lakes_cci) which are then processed using the Hysope software which is classically using the following equation:

$$LWL = Alt - Rcorr - TE \qquad [2.1]$$

Where LWL is considered with respect to a geoid, Rcorr is the measured range between the satellite and the lake surface, Alt is the altitude of the satellite above an ellipsoid and TE is the combination of all correction factors to take into account atmospheric refraction (propagation in the ionosphere and the troposphere), tidal effects (solid Earth, lake and polar), and geoid height above the ellipsoid. For more detailed information a full discussion of the computation of LWL is found in Crétaux et al. (2009).

All corrections are released in the Geophysical Data Records (GDRs) (delivery delay = 90 days) or the IGDRs. The range is chosen from different retracking considering that generally the OCOG retracking is the most suitable for continental surface (see E3UB v2.1.1 document). The geoid correction is calculated using the repeat track technique (see E3UB v2.1.1 and Crétaux et al. 2009, 2016).

2.2 Validation results

The general algorithm used to calculate water level over lakes is well known and established in scientific literature. To address the issues that are listed in the following sections, we need to analyse lakes where reference in-situ data is available. Examples of these procedures are given in Ricko et al. (2012) and Arsen et al. (2015), comparing different databases which include lake water level observations.

Lakes_cci cooperates with the State Hydrological Institute of St Petersburg, which provides in-situ data of LWL for a set of Russian and central Asian lakes. We also use existing databases on the web to increase the number of lakes that can be used for this purpose. These external sources are indicated in Table 1.

Table 1. In situ datasets used in the validation of the LWL product

Source	Description
U.S. Army Corps of Engineers	The U.S. Army Corps of Engineer provides in-situ data on Great Lakes. All levels are referenced to the International Great Lakes Datum of 1985 (IGLD 85). Water levels have been coordinated with Canada for 1918-2018.

Source	Description
Hidricos Argentina[4]	The database base of Hidricos Argentina provides in-situ data on national rivers and lakes.
U.S. Geological Survey[5]	The USGS investigates the occurrence, quantity, quality, distribution, and movement of surface and underground waters, and disseminates the data to the public. It provides in-situ data on U.S. lakes.
Water Office of Canada[6]	The Water Office of Canada provides historical water level collected over thousands of hydrometric stations across Canada.
FOEN[7]	The Swiss Federal Office for the Environment provides hydrological data, and in particular the water levels of lakes in Switzerland.
ANA[8]	The Brasilian "Agencia Nacional de Aguas e Saneamiento Basico" (ANA) provides in-situ data on reservoirs in Brazil.

The comparative analysis allows the statistically best performing retracking algorithm to be selected, as has been widely demonstrated for lakes as well as rivers.

Additional metrics to validate the LWL products include comparison of individual LWL retrieval to the long-term LWL variability, to detect outliers. The impact of removing outliers is traced as part of this process.

2.3 Identified issues

There are two main issues currently under investigation for processing of altimetry data over lakes. The first is related to the onboard tracking system, and the second is related to the processing of altimetry over small lakes.

We have identified solutions to address onboard tracking issues based on new a priori information. For retrieval of LWL over small lakes we identify solutions based in new algorithms for SAR data. Both approaches are detailed in the Algorithm Development Plan.

Another separate consideration of retrieval performance is the calculation of relative biases when several satellites of different types of orbits are used over a given lake. When we use a series of satellites such as Topex / Poseidon, Jason-1/2/3, Sentinel-6A, we collect data from the same orbit, so that the relative bias between each mission is well described and calibrated (see Cretaux et al. 2009, 2011, 2013, 2018, Bonnefond et al., 2018). When observations from different orbit are used, however, such as with Jason and Envisat or Jason and Sentinel-3, another bias is added. The instrumental biases are known, but since the tracks do not cover the same position over the lake, an additional bias due to geoid error has to be considered. A very simple method has been developed at LEGOS to correct for this additional bias. The LWL is calculated independently using each track, over the whole period of time, and during the overlapping period we interpolate the point measurement from each pass and calculate the average difference between all interpolate points. It then corresponds to the additional bias due to geoid errors.

The observed limitations are not negligible and need further improvements over the next two years. The most notable expected improvement is a dedicated DEM uploaded to current ESA altimetric missions, to improve the radar tracking over small lakes particularly in regions with surrounding relief, like mountain areas. New processing of past missions, mainly using data generated as part of the FDR4ALT project contained in the TDP- IW (Thematic Data Product - Inland Waters), is also carried out in order to provide longer and more accurate time series for small lakes. Finally, we are implementing new methods for SAR processing on sentinel-3A and sentinel-3B satellite known as full SAR processing, which will allow sub meter resolutions along the track of the satellite.

2.4 Sentinel-3A/B reprocessing

ESA has developed dedicated processing chains for L2 products over different surfaces including inland waters with the objective of improving the performances of the current products. Concerning inland waters, hamming window and waveform zero-padding have been implemented. The reprocessing of the Sentinel-3 A/B covers from the beginning of the mission until September 2023 when those Land Thematic products are operational produced. The full Sentinel-3A and Sentinel-3B missions started reprocessing with the Thematic Instruments Processing Facilities (T-IPFs) at the end of 2022 and the full Sentinel-3 Altimetry mission Thematic products dataset is available since November 2023. With those products, the precision is strongly improved over inland waters, on particular over small targets.

The timeseries for the Lakes_cci targets monitored ty Sentinel3 A/B mission have been estimated using data from the new thematic products and the results compared to the previous timeseries obtained with the Payload Data Ground Segment (PDGS) processing.

2.4.1 Comparison of new time series to external in situ data

Eight time series of lakes being observed only by Sentinel-3A/3B missions, obtained with thematic products, were compared to the available time series from 3 in-situ datasets described in Table 1: ANA, FOEN, and Hídricos Argentina. Table 2 contains comparison statistics for time series estimated with PDGS and T-IPF data compared with in-situ data: the URMSE, Pearson Coefficient and the number of observations used to compute those values.

Table 2. Statistics between PDGS/TIPF altimetry datasets compared to in-situ measurements

		URMSE(cm)		Pearson		No. observations	
source	lake name	PDGS	T-IPF	PDGS	T-IPF	PDGS	T-IPF
ANA	Sobradino	12,438	7,119	0,99861	0,99949	160	331
ANA	Tres Marias	13,625	7,597	0,99871	0,99939	40	53
FOEN	Bodensee	4,24	3,237	0,99228	0,99626	54	54
FOEN	Leman	3,668	3,057	0,977799	0,98282	139	143
Hidricos Argentinos	Cochrane	14,018	14,266	0,49936	0,48753	32	32
Hidricos Argentinos	San Martin	34,514	33,377	0,91792	0,93594	96	106
Hidricos Argentinos	Viedma	5,237	5,743	0,99458	0,99304	192	207

For most of the comparisons, the URMSE value is lower for the time series estimated with T-IPF data. For Lake Cochrane, this is due to a shift on the in-situ measurements during the period 2020-2021 (Figure 1). This shift also explains why the value of the Pearson coefficient is very low for this series (near 0.5). For the rest of the lakes, this coefficient is greater than 0.9, indicating an excellent correlation between the altimetry time series and in-situ time series, which improves with the use of T-IPF data.

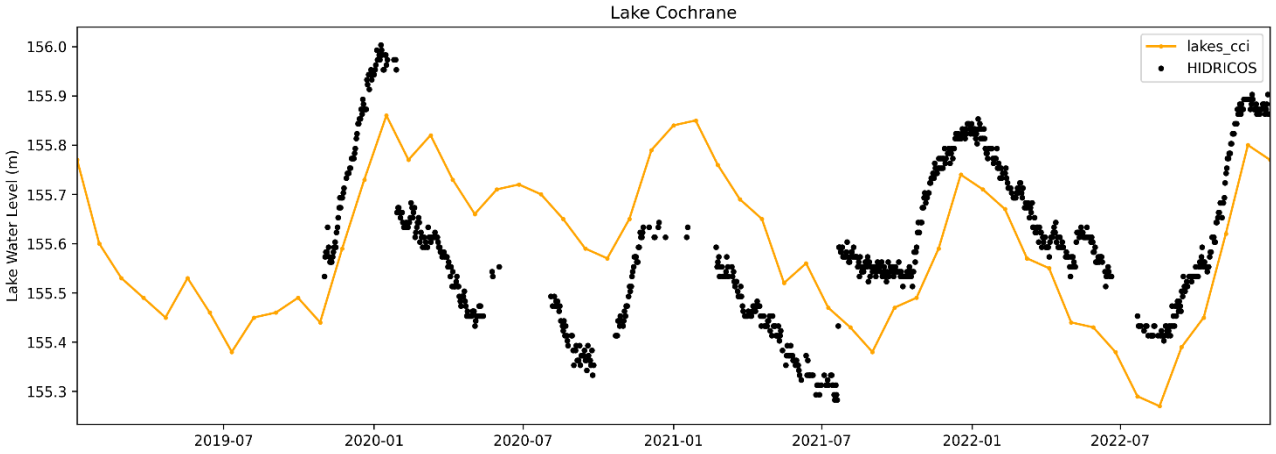


Figure 1. Lake Cochrane: Comparison T-IPF time series to Hídricos Argentina in-situ time series

2.4.2 Uncertainty improvements

One of the main advantages of T-IPF data is the reduction at level 2 of the tracking noise, which translates into an improvement in the uncertainty of water level values estimated at level 3. Figure 2 shows how the number of lakes with low uncertainty values increases with time series estimated with T-IPF data. For example, the number of lakes with a mean uncertainty lower than 5 cm using PDGS data is 126. The number of lakes using T-IPF data increases to 170.

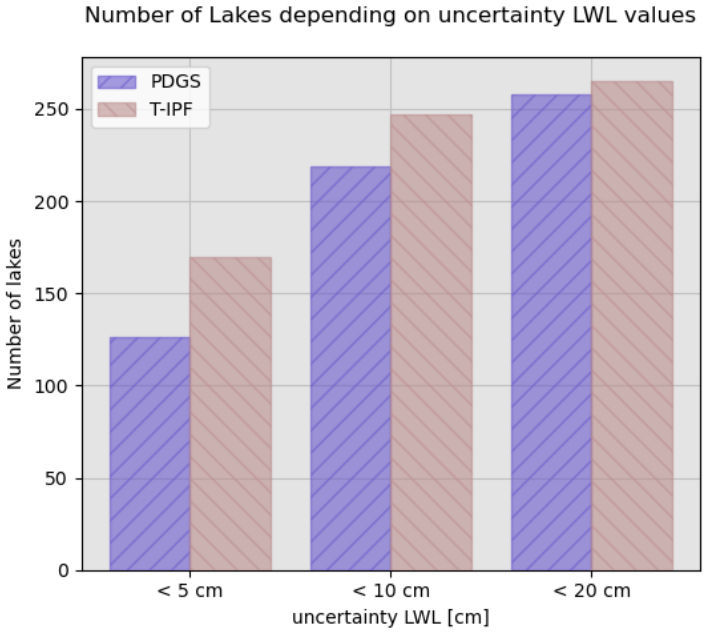


Figure 2. Number of lakes as a function of the uncertainty threshold

However, considering that the L2 data processing described in the ATBD also reduces retracking noise, the improvement in T-IPF data is more significant for small transects. Figure 3 shows that for small transects (less than 40 km), the number of lakes complying with the GCOS recommendation (uncertainty less than 10 cm) is higher with T-IPF data. For transects larger than 40 km, there is no improvement in uncertainty.

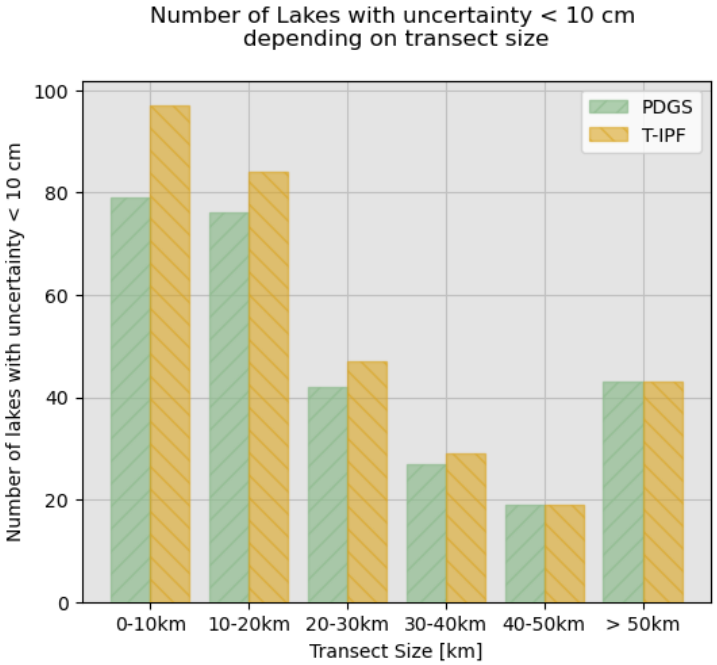


Figure 3. Number of lakes with uncertainty lower than 10 cm (GCOS recommendation) based on the transect size.

The timeseries estimated with the T-IPF data were also compared with Lakes_cci V2.1 time series. The time series of lake water level are very similar, which confirms that the algorithm used for this estimation Is very robust. The main difference concerns the uncertainty value. Figure 4 and Figure 5 show, respectively, the results obtained for lakes Cumberland and Atlin which demonstrate the improvement in the uncertainty value. In particular, the high uncertainty values correspond to outliers in the water level time series.

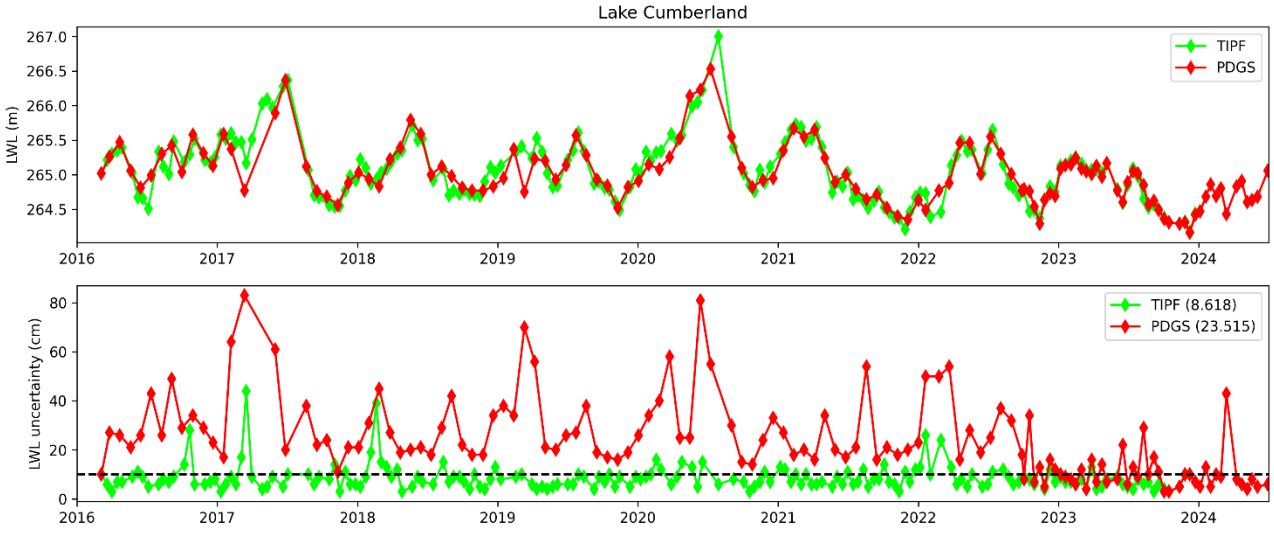


Figure 4. Lake Cumberland: Lake water level time series and uncertainty time series estimated with PDGS and T-IPF data.

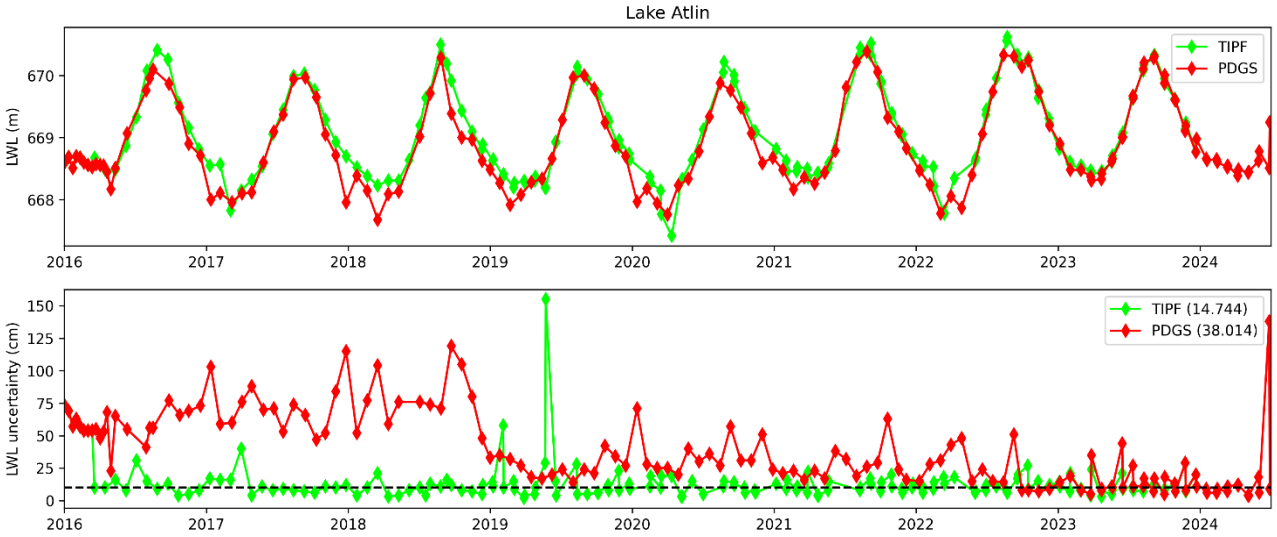


Figure 5 Lake Atlin: Lake water level time series and uncertainty time series estimated with PDGS and T-IPF data.

2.5 References

Arsen, A., JF. Cretaux, and R. Abarca-Del-Rio. (2015). Use of SARAL/AltiKa over mountainous lakes, intercomparison with Envisat mission J. of Adv. Space Res. The Saral/ALtiKa satellite Altimetry Mission, 38, 534-548. doi: 10. 1080/01490419. 2014. 1002590

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

Boy F., Cretaux J-F., Boussaroque M., & Tison C. (2022). Improving Sentinel-3 SAR mode processing over lake using numerical simulations, IEEE Transactions on Geoscience and Remote Sensing, 60, 1-18, Art no. 5220518, Doi: 10.1109/TGRS.2021.3137034

Crétaux, J. F., S. Calmant, V. Romanovski, et al. (2009). An absolute calibration site for radar altimeters in the continental domain: lake Issykkul in Central Asia, Journal of Geodesy 83 (8) 723-735. doi: 10.1007/s00190-008-0289-7

Crétaux, J. F., S. Calmant, V. Romanovski, et al. (2011). Absolute Calibration of Jason radar altimeters from GPS kinematic campaigns over Lake Issykkul, Marine Geodesy, 34 : 3-4, 291-318. doi: 10.1080/01490419. 2011. 585110

Crétaux J-F., Bergé-Nguyen M., Calmant S., Romanovski V. V., Meyssignac B., Perosanz F., Tashbaeva S., Arsen A., Fund F., Martignago N., Bonnefond P., Laurain O., Morrow R., Maisongrande P., (2013). Calibration of envisat radar altimeter over Lake Issykkul, J. Adv. Space Res., Vol 51, 8, 1523-1541. doi: 10. 1016/j. asr. 2012. 06. 039

Cretaux J-F, M. Bergé-Nguyen, S. Calmant, N. Jamangulova, R. Satylkanov, F. Lyard, F. Perosanz, J. Verron, A. S. Montazem, G. Leguilcher, D. Leroux, J. Barrie, P. Maisongrande and P. Bonnefond. (2018). Absolute calibration / validation of the altimeters on Sentinel-3A and Jason-3 over the lake Issykkul, Remote sensing, 10, 1679. doi:10. 3390/rs10111679

Egido A., and W. H. F Smith. (2017). Fully Focused SAR altimetry: Theory and Applications, IEEE, 55, 1

Ričko M., C. M. Birkett, J. A. Carton, and J-F. Cretaux. (2012). Intercomparison and validation of continental water level products derived from satellite radar altimetry, J. of Applied Rem. Sensing, Volume 6, Art N°: 061710. doi: 10. 1117/1. JRS. 6. 061710

3 Lake water extent - LWE

3.1 Candidate algorithms

As detailed in the Algorithm Theoretical Basis Document, based on results obtained in previous work an approach based only on optical HR imagery was adopted for LWE. Water surfaces are extracted from the images based on the exploitation of an in-house processing chain, named ExtractEO (Maxant et al, 2022).

3.2 Validation results

Validation of the LWE product already starts during several preprocessing and processing steps. The validation is done individually over each lake for the following steps:

- Data selection process
- Hypsometric curves computation
- Water extraction

3.2.1 Date selection for candidate images based on altimetric data.

The aim of this step is to select set of pertinent images acquired at representative stages of a given lake, i.e., lower, higher and intermediate level of water/filling of the lake. For this, an analysis of a LWL from Hydroweb is done, based on quantile, a buffer is applied to enlarge the epochs.

In the example of Bagré reservoir (Burkina Faso), 56 dates were identified to characterize lake water level stages. Applying a temporal window to each date, the image search then extended to 336 individual days. From these dates, 110 Sentinel-2 MSI were found and investigated. After running ExctractEO to calculate cloud cover percentage, 32 images were kept with < 5 % cloud cover over the lake. The representativeness of the resulting set of images was then checked to ensure that the full range of water levels was still included (Figure 6). A hypsometric relationship is then computed based on LWL from Hydroweb and LWE derived from the Sentinel-2 MSI imagery.

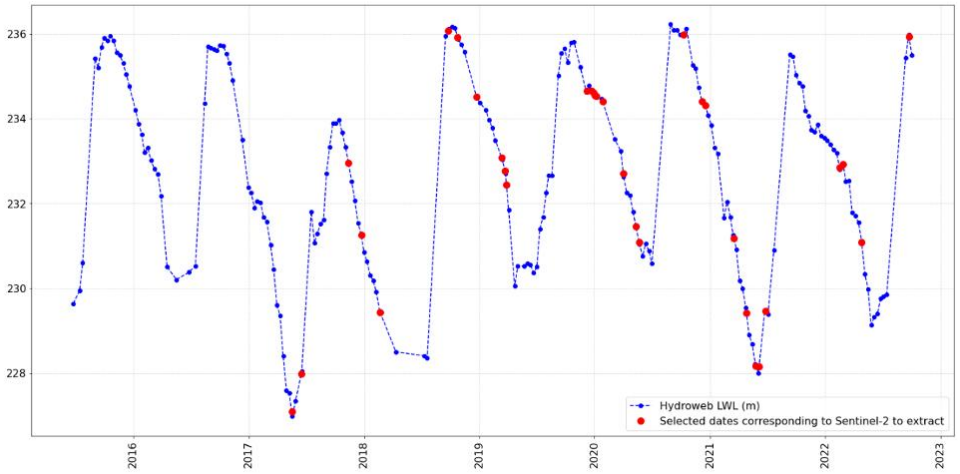


Figure 6: Date selection based on Hydroweb altimetric time series. The selected dates well cover all the stages of the lake filling, from Min to Max plus intermediate levels

3.2.2 Analysis of the hypsometric curves

The obtained hypsometric curve and repartition of the LWE/LWE pairs are analysed to observe the tendencies of the curves and identify possible anomalies, to remove outliers and to understand sources of potential omission or commission during the processing of the data (Figure 7).

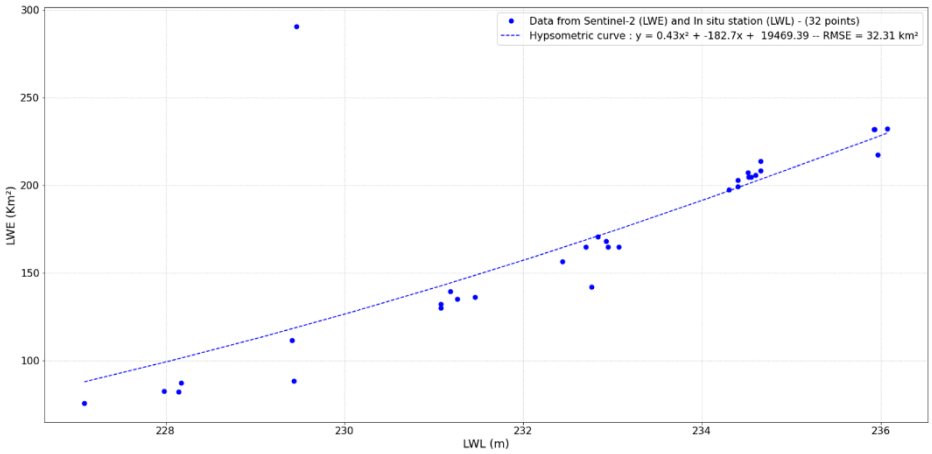


Figure 7: First estimate of the hypsometric relationship based on 32 images of Bagré Reservoir.

In this first hypsometric curve estimate, based on 32 LWL/LWE pairs, a few points are marked as outliers. It can further be noted that the curve does not integrate the highest values LWE/LWL. The resulting RMSE is relatively high at 32.31 km² (i.e., 13% of the largest observed lake surface).

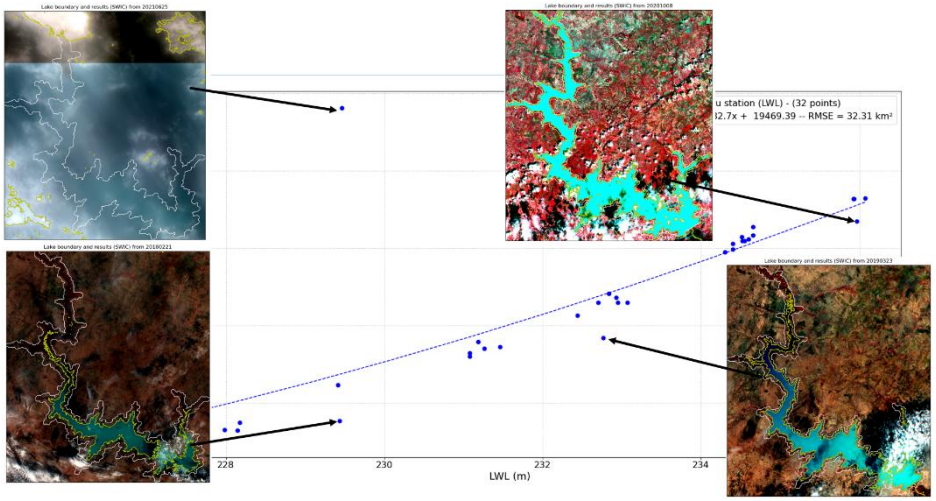


Figure 8: Analysis of problematic LWE/LWL pairs, usually related to haze, smoke or cloud.

Based on visual analysis (Figure 8), four pairs were rejected where clouds, haze and sun glint are observable.

A new hypsometric relationship is then calculated (Figure 9), where the curve extends over the full observed range and RMSE reduced from 32.3 km² (i.e. 13%) to 3.4 km² (1.4 %).

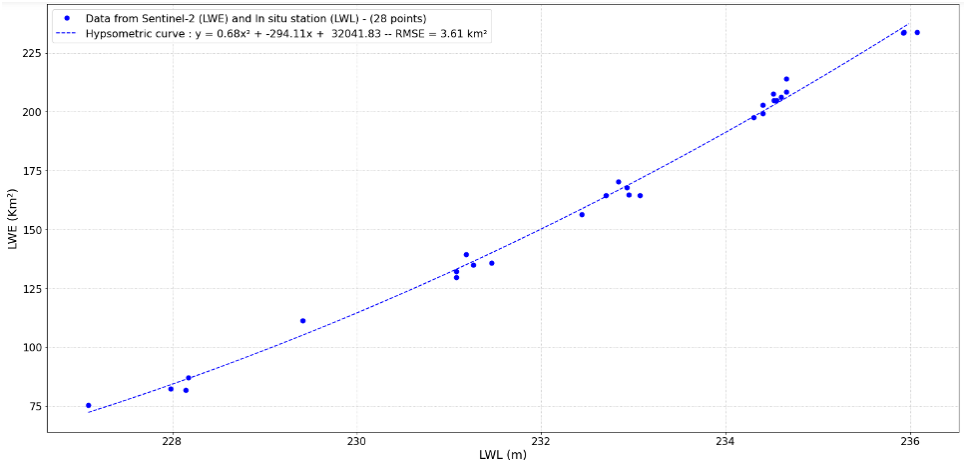


Figure 9: Final hypsometric curve for Bagré reservoir based on 28 pairs passing quality checks.

3.2.3 Water extraction process validation

Lake by lake water extents' validation is not suitable when applied to many lakes, however, it is possible, necessary and important to validate the procedure which generates lakes' area and lakes' 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 3) was obtained thanks to CNES facilities within the SWOT Cal/Val program and through Pleaides NEO promotional use in the Dinamis programme. Results of the investigation are detailed for each of the regions in the following sections.

Table 3: Pairs of VHR and HR data available for the validation of the processing chain

Site	Province Country	Size in km²	VHR dataset	Dates of acquisition VHR	Dates of acquisition HR
Fitri lake	Chad	200-1250	Pleiades HR: 0,7m	2021	
Fern Ridge Reservoir	Oregon	36	Pleiades HR: 0,7m	25072022	26072022
	USA			29072022	29072022
			SPOT 6-7: 1,5 m	11082022	13082022
Lac de Der	Champagne France	10-40	Pleiades HR: 0,70m	30122019	31122019
				06012020	06012020
Lac d'Orient	Champagne France	8-23	Pleiades NEO: 0,30m	15122022	13122022
					16122022
Lac d'Auzon-Temple	Champagne France	7-20	Pleiades NEO: 0,30m	15122022	13122022
					16122022
Lac d Amance	Champagne France	0-18	Pleiades NEO: 0,30m	15122022	13122022
					16122022

Shenjing lake	Anhui Prov. PR China	33	Pleiades NEO: 0,30m	10122021	07122021

Lakes in Champagne

In this area, several lakes are linked to the flood protection system of Paris by the Grands lacs de Seine establishment (Figure 10). They were covered by the recently launched Pleiades NEO sensors as well by Pleaides imagery.

Figure 10: Lakes in Champagne

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 Sentinel2 LWE compared to the straighter limits of VHR shown in Temple Lake (Figure 11). 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 12). 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 11: 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.

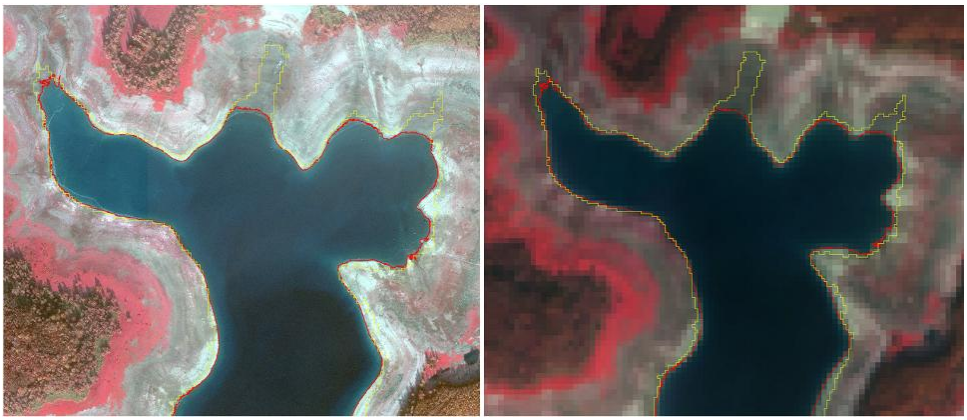


Figure 12: Comparison of VHR/HR in Orient Lake. Left: Pleaides NEO. Right: Sentinel-2. Red lines are the LWE limits derived from Pleaides NEO, yellow lines the Sentinel LWE limits.

Table 4: Metrics obtained over the Champagne lakes

All			
Metric	FREQUENCY	Area m²	Percentage
reference	72	21693449,96	100,00
database	23	21800055,36	100,49
omission	72	570486,23	2,63
commission	23	677091,66	3,11
taux_detection	17	21122963,70	97,37
taux_justesse	1	999999,00	96,89
BREAK	1	1,00	1,00
precision	1	1,00	0,97
recall	1	1,00	0,97
Fscore	1	1,00	0,97
CSI	1	1,00	0,94

Amance			
Metric	FREQUENCY	Area m²	Percentage
reference	20	1819683,51	100,00
database	6	1903600,00	104,61
omission	20	56409,20	3,10
commission	6	140325,68	7,37
taux_detectio	3	1763274,32	96,90
taux_justesse	1	999999,00	92,63
BREAK	1	1,00	1,00
precision	1	1,00	0,93
recall	1	1,00	0,97
Fscore	1	1,00	0,95
CSI	1	1,00	0,90

Auzon Temple			
Metric	FREQUENCY	Area m²	Percentage
reference	25	8711286,40	100,00
database	8	8473200,00	97,27
omission	25	338141,24	3,88
commission	8	100054,90	1,18
taux_detectio	6	8373145,10	96,12
taux_justesse	1	999999,00	98,82
BREAK	1	1,00	1,00
precision	1	1,00	0,99
recall	1	1,00	0,96
Fscore	1	1,00	0,97
CSI	1	1,00	0,95

Orient			
Metric	FREQUENCY	Area m²	Percentage
reference	13	10829060,28	100,00
database	3	11147755,36	102,94
omission	13	96436,30	0,89
commission	3	415131,38	3,72
taux_detectio	2	10732623,98	99,11
taux_justesse	1	999999,00	96,28
BREAK	1	1,00	1,00
precision	1	1,00	0,96
recall	1	1,00	0,99
Fscore	1	1,00	0,98
CSI	1	1,00	0,95

The metrics obtained on these reservoirs 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 the very good recognition and extraction of water bodies using Sentinel2 data via ExtractEO, qualifying the proposed approach for generating LWEs (Table3).

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 km² and a popular site for boating, fishing, and birdwatching. Surrounding portions of the reservoir is the Fern Ridge Wildlife Area, a wildlife management area providing a unique habitat for a variety of species including wintering waterfowl populations.

By comparing the LWE extracted from the VHR and HR images, Sentinel-2 detected more water in the eastern part of the reservoir (Figure 13)..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 Senitnel2 data. This confusion between open water and wetland stems from the use of the Sentinel2 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.

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 ?.

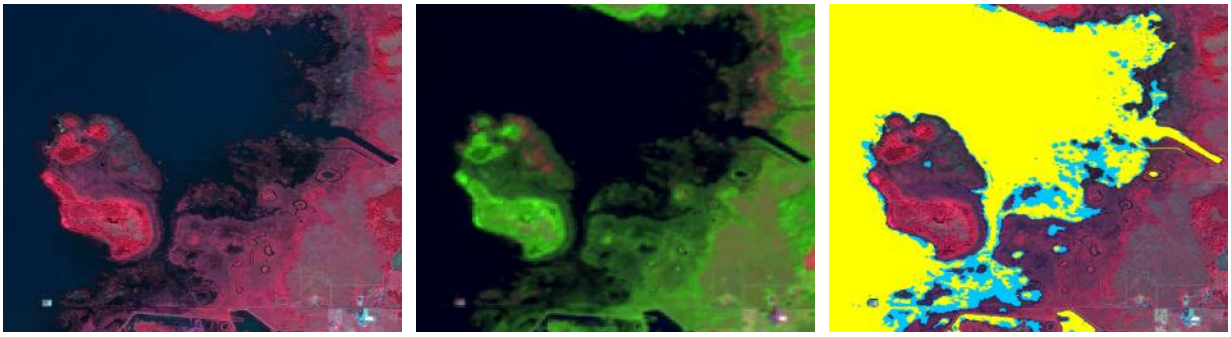


Figure 13: Fern Ridge reservoir observed with Pleiades HR (25-07-2022), Sentinel-2 SWIR false colour composite (26-07-2022). A comparison of LWE derived from Pleiades in yellow and Sentinel-2 in blue is shown to the right.

In terms of metrics, despite this relative omission related to wetlands, correspondence was good in terms of accuracy and F Score (Table 5).

Table 5: 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 Pleiades and Sentinel-2 acquired on 29/07/2022 (right).

	Pleaides on 626/07/2022, S-2 on 25/07/2022				Pleaides and S-2 on 729/07/2022		
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

Fitri Lake (Chad)

Fitri Lake (12°50'N, 17°30'E), is the largest lake inside Chad. It plays a major role in this Sahel region for human activities and as a refuge for biodiversity. It occupies an endoreic watershed with a surface area of 80,000 km² located in the middle of the Sahel to the east of Lake Chad. Fitri Lake's dynamic is directly linked to the West African monsoon regime, whose seasonal to multi-annual variability directly impacts variations in its level and its extent. As a result, Lake Fitri is a sensitive and unique record of monsoon variability and its impact on landscapes and ecosystems in the past.

Lake Fitri is an endoreic lake occupying a basin with a very shallow topography. 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 km² respectively, compared with the 194 km² observed in 2017.

Over Fitri Lake (Chad), two approaches of processing Sentinel-2 MSI were compared, one based on 10 m spatial resolution and the other at 20 m of spatial resolution.

Visual analysis highlighted the gain related to the highest spatial resolution of Pleaides. This is particularly well documented on the fringes of the main water body but also on smaller flooded branches of the lake (Figure 14 and Figure 15).

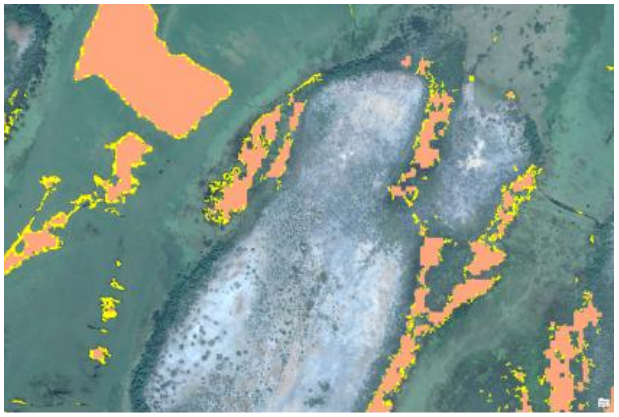


Figure 14: Fitri lake illustration of gain in water surface when exploiting VHR Pleiades imagery (yellow) compared with the Water mask derived from Sentinel-2 (orange).

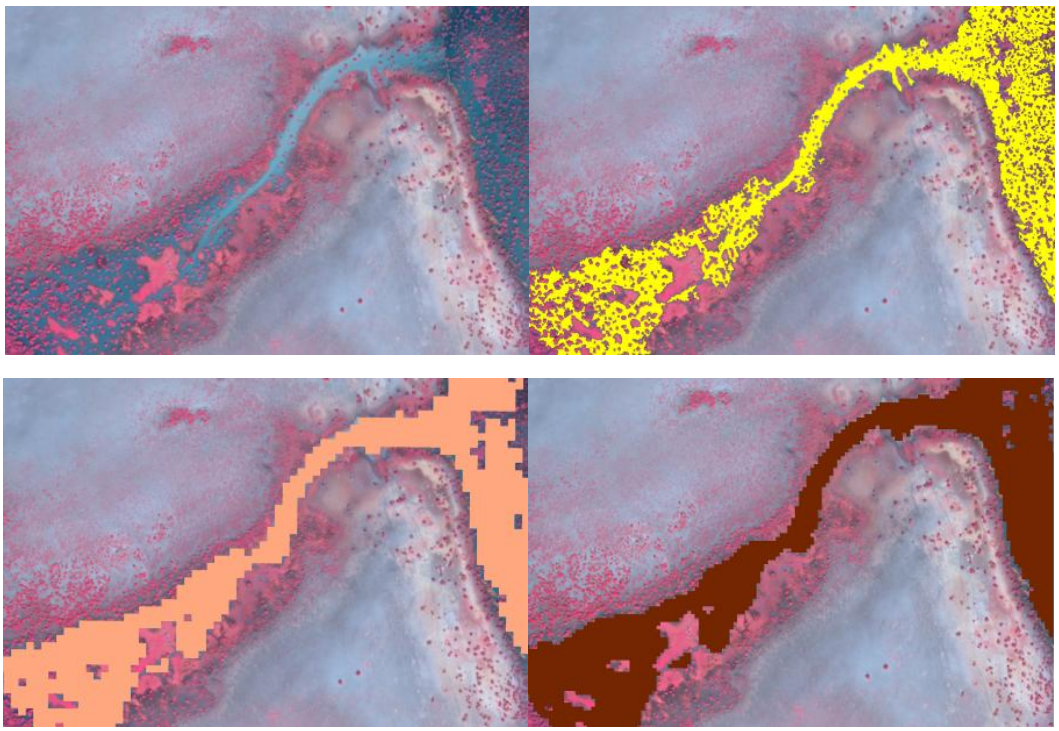


Figure 15:Comparison of LWE derived from Pleaides image and Sentinel 2 with UL, Pleiades colour composite of a flood branch of the Fitri lake, upper right, LWE derived from Pleaides, Bottom left, LWE derived from a Sentinel2 image at 20 m of spatial resolution and bottom right, LWE from A Sentinel2 processed at 10m.

When analysing the obtained LWE metrics (Table 8), 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 8: 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

Shenjin Lake

Shengjin Lake, 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. Dominated by lakes and freshwater marshes and boasting good water quality, it is one of the best-preserved inland freshwater lake systems along these stretches of the Yangtze River. The site plays a role in regulating floods along the Yangtze and contributes to water purification and regional climate regulation. It also provides an important stopover and wintering ground for migrating birds.

The case of the Shenjin Lake illustrates the difficulties to extract water extent in case of shallow waters, when the delimitation of water and wet sediments can be very delicate. This is particularly sensitive during the step of sampling selection process.

For Sentinel-2, two sets of parameters were tested (Table 9), to reduce the weight of the SWIR channels that can induce commission over wetland and mud/sandy banks such as previously shown.

Table 9: Parameters exploited for sampling selection

	Param 1	Param 2
MNDWI Threshold	0.45	0.1
NDWI Threshold	[0.1, 0.4]	[0.01,0.3]
Number of samples	10000	10000
Pekel Threshold	20	20
Refining method	SWIC	SWIC
SWIC Threshold	[0.3, 0.9]	0

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 set parameterization introduce a buffer zone along the water limits, including very light sandy areas (Figure 16).

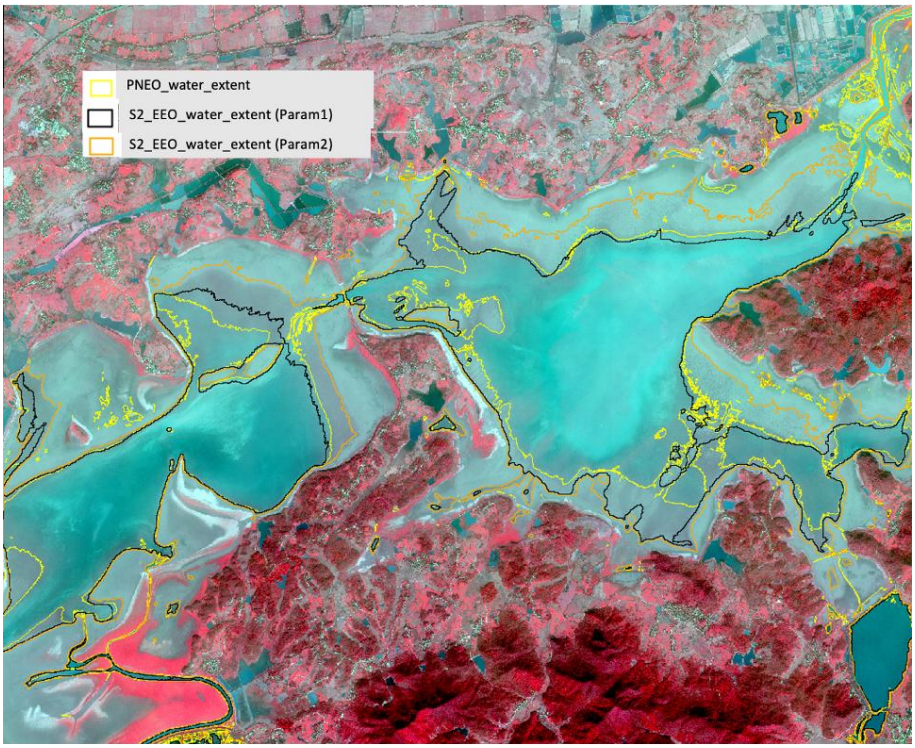


Figure 16. Comparison of the LWE limits derived from Pleiades NEO and Sentinel2

When analysing the LWE metrics (Table 10), results obtained from parameter set 2 present high values for commission and relative low accuracy values. 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 10: 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	Percent	Frequency	Sum Area	Percent
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

3.3 Identified issues

In terms of algorithm few issues have been identified:

- The presence of sun glint on water surfaces can disrupt the process of recognising and extracting water bodies. The appearance of this phenomenon depends not only on the position of the sun, but also on the location of the target in the swath. It would be beneficial to develop a sunglint flag to avoid issues with sunglint propagating through to final results. This would also allow automatic adjustment of related coefficients and thresholds in the ExtractEO processing chain.
- Refinement to the existing cloud detector in the ExtractEO chain should lead to better cloud detection in the S-2 L2A product.

3.4 References

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

4 Lake surface water temperature – LSWT

4.1 Candidate algorithms

Surface temperatures from infrared observations are obtained by coefficient-based methods or optimal estimation (OE, Merchant and Embury 2014). Because of the varied altitudes of lakes and the large differences in atmospheric absorption associated with continentality, optimal estimation is the appropriate approach for LSWT estimation (MacCallum and Merchant, 2012).

OE also provides comprehensive equations for uncertainty evaluation, on which basis uncertainty estimates are provided in LSWT products per datum.

As well as retrieval, classification of which pixels are filled with water under clear skies is a necessary part of the LSWT processing. This is done by a "fuzzy logic" style approach in which a number of metrics with fuzzy thresholds are combined into a "water detection score" that contributes to the definition of the quality level attributed to the pixel. Bayesian cloud detection, as used for sea surface temperature, was also considered to identify clear-sky pixels but is heavily compromised in its current implementation for small lakes, where the spatial coherence of the temperature of the scene is not a good indicator of cloud (unlike in the centre of large lakes and over open ocean). Because of the user requirement to increase the number of measured lakes, the latter scheme is therefore currently inapplicable for the identification of clear-sky only water pixels.

4.2 Validation results

Validation results (for the LSWT 4.5 included in Lakes_cci CRDP v2.x) are summarised in the product validation report. The validation undertaken is a comparison of satellite to matched in situ temperatures. These comparisons are limited by the non-representative sample-of-opportunity (in-situ measurements being unfortunately hard to obtain) and by variable and often unknown in-situ uncertainty characteristics and quality control. Validation results are too ambivalent in this case to be used as a discriminant between alternative algorithmic approaches, and for LSWT are not used in this way. As mentioned in the previous section, the retrieval algorithm is established by considerations rooted in physics and inverse theory.

Nonetheless, the validation exercise is sufficient to establish that for quality level (QL) 5 (best quality) LSWT data, the data have low bias (<0.1K) and uncertainty estimates are reasonable.

4.3 Identified issues

LSWT retrieval

1. Optimal estimation uses an observation-simulation error covariance matrix. This matrix is presently a simple diagonal estimate that doesn't account for the likelihood of cross-channel correlations in the simulation errors.
2. Optimal estimation uses a prior error covariance matrix. This matrix is a simple diagonal estimate based on experience of ERA-interim, and ideally should be updated for use with ERA-5.

LSWT retrieval uncertainty

(NB, this aspect is also addressed by the issues identified for LSWT retrieval.)

1. The decomposition of the optimal estimation uncertainty into different correlation length scales is an approximation; a more complete solution needs to be coded (The decomposition is relevant when creating gridded data).

Water detection

The water detection to select water-only pixels relies on day-time (reflectance) channels and therefore cannot be applied at night. The alternative that can work at night is based on Bayesian cloud screening (as used in SST CCI), and for small lakes require considerable research, development and modification.

Quality level determination

1. QL determination is based mainly on water detection results, on the sensitivity to the prior, on retrieval chi-square results (which measure the plausibility of the solution given the prior and observations). However, the chi-square results differed more than expected between Metop-A and Metop-B AVHRRs, affecting the QL attribution adversely in the case of Metop-B. This is yet to be clearly understood, but the root of the reasons very likely lies in the L1b data.

Increasing coverage processing night-time

1. By far the most impactful issue for users for LSWT is to increase the density of coverage, which is limited by sensor-orbit coverage and by cloud cover. Moreover, night-time observations are considered to provide better accuracy than daytime data when compared with bulk temperatures due to, among other things, the absence of solar heating (Hook et al., 2003).

Updating the prior LSWT

Currently, for the prior LSWT for OE and for Bayesian cloud detection is the LSWT climatology derived from the previous version of the Lakes_cci dataset. However, once the work to construct a gap-filled complementary dataset is completed, it will be worth to update the prior with the gap-filled product which will be a better a priori “guess” of the LSWT.

A work-around for the Metop-B issue with chi-square is implemented in LSWT v4.1 and LSWT v4.5, but a more fundamental resolution will be sought.

4.4 References

Merchant, C.J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., Good, S., McLaren, A., Rayner, N., Morak-Bozzo, S. and Donlon, C. (2014) Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1 (2). pp. 179-191.

MacCallum, S.N. and Merchant, C.J. (2012) Surface water temperature observations of large lakes by optimal estimation. Canadian Journal of Remote Sensing, 38 (1). pp. 25-45.

Hook, S. J., Prata, F. J., Alley, R. E., Abtahi, A., Richards, R. C., Schladow, S. G., Palmarsson, S. (2003). Retrieval of lake bulk and skin temperatures using Along-Track Scanning Radiometer (ATSR-2) data: A case study using Lake Tahoe, California. Journal of Atmospheric and Oceanic Technology, 20(4), 534–548.

5 Lake water leaving reflectance – LWLR

5.1 Candidate algorithms

Algorithms in the LWLR processing chain *Calimnos* fall into three categories:

- Pre-processing including pixel identification as water, land, cloud or ice
- Atmospheric correction yielding LWLR
- Water-column physical and biogeochemical properties derived from LWLR

The algorithm validation procedures and results associated with CRDP v3.0.0 have been updated with respect to previous releases. Both the atmospheric correction yielding the LWLR variable and the calibration of algorithms to derive biogeochemical properties have been updated. Additional candidate algorithms for chlorophyll-a were already described in the Algorithm Development Plan. A section has been added to evaluate the performance of algorithms for the light absorption by Coloured Dissolved Organic Matter (CDOM) following an earlier technical report (available through the project website) and its inclusion in CRDP v3.0.0. Work is ongoing to include solutions for the vertical diffuse attenuation coefficient (K_d), which will be reported in the user documentation.

For the **pre-processing** category processing relies on the Idepix multi-sensor processor in SNAP. This algorithm combines information from static sources (such as water extent) and dynamic pixel identification based on a neural network trained for each of the optical sensors. Depending on the capabilities of the sensor, the processing chain will rely on combinations of these processes. The algorithm is not part of validation of the Lakes_cci LWLR but its performance is taken into consideration with regard to consistency in water/land masking between the Lakes_cci thematic ECV Products.

Validation of algorithms for **atmospheric correction** requires near-coincident in situ observations of water-leaving reflectance. Due to relatively scarce in situ data from lakes, the window for 'coincidence' may be up to several days from satellite observation. Longer time windows allow more data points to be included, which is suitable to determine the best-performing algorithm but less suitable to determine product uncertainties which may then be exaggerated. For CRDP v3.0.0 we consider several new data sources whereas previous assessments were only based on LIMNADES, a community-owned repository for lake bio-optical measurements. Previously, the assessment was limited to MERIS and MODIS. The most prominent data collection used to assess CRDP v3.0.0 is now the GLORIA database (Lehmann et al. 2023). In addition, several high-frequency autonomous platforms can be considered, providing a far greater density of observations for a small number of lake stations and ship transects. These include the WISPstation observations using instrumentation by Water Insight, and So-Rad shipborne observations with automation developed at PML.

Periodically, studies are carried out to assess the comparative performance of atmospheric correction algorithms (e.g. in the ACIX-Aqua framework). Each algorithm is scored for statistical performance metrics as well as the number of (valid) results returned. The round-robin procedure is not repeated for CRDP v3.0.0 which, like previous versions, is based on the Polymer algorithm albeit an upgraded version. The performance of the selected atmospheric correction algorithm is validated against in situ radiometric observations on a per-sensor basis, characterising LWLR uncertainty in terms of random and systematic components.

A strategy for including SeaWiFS and VIIRS observations will be defined in future, based on the results of MODIS, but is not within the scope of currently planned outputs.

Finally, **algorithms for the derived water-column properties**, notably chlorophyll-a concentration, total suspended matter dry weight, and light absorption by Coloured Dissolved Organic Matter, are evaluated in a sensor-dependent manner similar to the procedure given for atmospheric correction algorithms.

These algorithms are first evaluated against the available in-situ data archives (see above) to assess their application range. The algorithms are then tuned for optimal performance as a function of their membership to a set of **Optical Water Types (OWTs),** allowing them to be mapped to satellite imagery using a weighted averaging 'blending' method. Details on the optical water type methodology are provided in the ATBD.

Several of the algorithms that may be considered for atmospheric correction are coupled atmosphere-water models yielding water-column properties including chlorophyll-a and suspended matter concentrations. These algorithms have not all been thoroughly evaluated in peer-reviewed literature and are included primarily for reference. Where they outperform better-understood alternatives they are given further consideration.

The candidate algorithms are listed per sensor in the following tables (Table 11 - Table 17). As a rule of thumb, algorithms with a transparent and published methodology are considered, and algorithms with a theoretical basis suggesting specificity and sensitivity to the target substance are strongly preferred. This may nevertheless include empirical algorithm forms as well as machine learning models. Where algorithms have been previously validated in a specific region, their validated range is given. The MERIS/OLCI algorithms used in CRDP v1.0 to v2.1.0 (listed in Table 11 and Table 12) are described in comparative detail in Neil et al. (2019).

For the MODIS/VIIRS set of candidate algorithms (Table 13 for chlorophyll-a, Table 14 and Table 15 for suspended matter and turbidity) this comparison was carried out for the first time in the Lakes_cci and used from CRDP v2.0. Details on the waveband combinations that are shown in these tables suggest there may be substantial overlap, in which case highly similar algorithms are ultimately collapsed into algorithm types prior to calibration.

Table 16 and Table 17 list candidate chlorophyll-a and TSM/Turbidity algorithms for OLCI (and in many cases, MERIS) which were added to the evaluation for CRDP v3.0.

Table 11: Candidate algorithms tested for MERIS (and OLCI by proxy) yielding chlorophyll-a concentration

Type	Model	Reference
(Semi-)empirical NIR-red BR	MERIS 2-Band 708/665	Gilerson *et al.* 2010 Gurlin *et al.* 2011 Gons *et al.* 2005
	MERIS 2-Band 753/665	Gilerson *et al.* 2010 Gitelson *et al.* 2011 Moses *et al.* 2009.
	MERIS 3-Band	Gitelson *et al.* 2008 Gitelson *et al.* 2011 Gurlin *et al.* 2011 Moses *et al.* 2009
	MERIS NDCI	Mishra *et al.* 2012
Empirical OC	MERIS OC2E MERIS OC3E MERIS OC4E	O'Reilly *et al.* 2000
Neural Network	NN_Chl NN_IOP FUB CoastColour C2RLakes(EUT/BOR)	Ioannou *et al.* 2013

Type	Model	Reference
Analytical	MERIS QAA [Turbid]	Mishra *et al.* 2013
	MERIS GSM	Maritorena *et al.* 2002
	MERIS Matrix Inversion	Boss and Roesler 2006
Peak Height Method	MPH	Matthews *et al.* 2012

Table 12: Candidate algorithms tested for MERIS (and OLCI by proxy) yielding total suspended matter.

Type	Algorithm name	Reference
Empirical	Binding red Zhang 708 Vantrepotte 665 POWERS 560	Binding *et al.* 2005 Zhang *et al.* 2010 Vantrepotte *et al.* 2011 Eleveld *et al.* 2008
	D'Sa 665/560 Dekker 490, 560 Dekker 560, 665	D'Sa *et al.* 2007 Dekker *et al.* 2002
	Loisel 3-Band	Loisel *et al.* 2014
(Semi-) Analytical	Binding A Nechad 665 Nechad 681 Nechad 708 Nechad 753	Binding *et al.* 2010 Nechad *et al.* 2010

Table 13: Summary of candidate Chla algorithms for MODIS/VIIRS

Code	Name/form	Type	Bands*	Calibration or validated range	Reference
A	OC3M	blue-green band ratio	min[R_{rs}(443), R_{rs}(488)], R_{rs}(547)	0. 2 ~ 90 mg m^{-3}	O'Reilly and Maritorena 2000
B	OC2M	blue-green band ratio	R_{rs}(488), R_{rs}(547)	0. 2 ~ 90 mg m^{-3}	O'Reilly and Maritorena 2000
C	OC2M-HI (500 m)	blue-green band ratio	R_{rs}(469), R_{rs}(555)	0. 2 ~ 90 mg m^{-3}	O'Reilly and Maritorena 2000
D	FLH	peak height	R_{rs}(665), R_{rs}(677), R_{rs}(746)	1 ~ 10 mg m^{-3}	Letelier 1996
E	linear	NIR-red band ratio	R_{rs}(748), [R_{rs}(667) or R_{rs}(678)]	4 ~ 240 mg m^{-3}	Gitelson 1992; Dall'Olmo et al. 2005; Gitelson et al. 2007, 2008; Gurlin et al. 2011
F	linear	blue-green band ratio	R_{rs}(551), R_{rs}(443)	8 ~ 17 mg m^{-3}	Ha et al. 2013

Code	Name/form	Type	Bands*	Calibration or validated range	Reference
G	linear	spectral index	R_{rc}(645), R_{rc}(859)	6. 6 ~ 113. 7 mg m^{-3}	Shi et al. 2017
H	APPEL model	empirical	R(645), R(859), R(469)	2. 5 ~ 91. 0 mg m^{-3}	El-Alem et al. 2012
I	GSM	semi-analytical	(not reproduced)	0 ~ 100 mg m^{-3}	Maritorena et al. 2002
J	QAA_v6	semi-analytical	(not reproduced)	0. 02 ~70. 21 mg m^{-3}	Lee et al. 2002
K	QAA_Tur	semi-analytical	(not reproduced)	59 ~1376 mg m^{-3}	Mishra et al. 2013, 2014
L	MODIS SA	semi-analytical	(not reproduced)	0 ~ 2 mg m^{-3}	Carder et al. 2004

*Reflectance bands are as used in the original definition, taking the following forms:
R_{rs} is above-surface remote-sensing reflectance
R_{rc} is the atmospherically Rayleigh-corrected reflectance.

Table 14: Summary of candidate Turbidity algorithms for MODIS/VIIRS

Code	Name/form	Type	Bands*	Calibration or validated range	Reference
A	polynomial	single red band	R_{rs}(645)	TSM <30 mg L^{-1}	Petus et al. 2010
B	linear	single red band	R_{rs}(645)	Turb 0 ~ 15 NTU	Moreno-Madrinan et al. 2010
C	exponential	single red band	R_{rs}(645)	Turb 1. 8 ~ 160 FTU	Constantin et al. 2017
D	power law	single red band	R_{rs}(645)	Turb 0. 9 ~ 8 NTU	Chen et al. 2007
E	polynomial	single NIR band	nL_w(869)	Turb 1~300 NTU	Wang et al. 2012
F	power law	NIR-red ratio	R_{rs}(859)/R_{rs}(645)	Turb 50 ~ 1000 NTU	Robert et al. 2016
G	exponential	NIR-red ratio	R(859)/R(645)	Turb 77. 4 ~2193 NTU TSM 77 ~ 2182 mg L^{-1}	Doxaran et al. 2009
H	semi-empirical	red or NIR	ρ_w(645), ρ_w(859)	Turb 1. 8 ~ 988 FNU	Dogliotti et al. 2015

*Reflectance bands are as used in the original definition, taking the following forms:
R_{rs} is above-surface remote-sensing reflectance
nL_w is the normalized water-leaving radiance.
R is the 'surface reflectance' of the MODIS land product.
ρ_w is the water reflectance, which is defined as πL_w (λ) / $E^0{}_d{}^+$ (λ), where L_w is the water-leaving radiance and $E^0{}_d{}^+$ is the above-water downwelling irradiance.

Table 15 Summary of candidate TSM algorithms tested for MODIS/VIIRS

Code	Name/form	Type	Bands*	Calibration or validated range	Reference
I	linear	single red band	R(645)	0 ~ 55 mg L^{-1}	Miller and McKee 2004; Sipelgas et al. 2006
J	polynomial	single red band	R_{rs}(645)	0 ~ 30 mg L^{-1}	Petus et al. 2010

K	exponential	single red band	$R_{rs}(645)$	0 ~ 300 mg L^{-1}	Zhao et al. 2011; Shi et al. 2015
L	polynomial	single red band	$nL_w(645)$	0 ~ 16 mg L^{-1}	Ondrusek et al. 2012
M	exponential	NIR-red ratio	$R(859)/R(645)$	77 ~ 2182 mg L^{-1}	Doxaran et al. 2009
N	power law	NIR-red ratio	$R_{rs}(859)/R_{rs}(645)$	18 ~ 927 mg L^{-1}	Robert et al. 2016
O	polynomial	NIR-red ratio	$log[R_{rs}(859)]/log[R_{rs}(645)]$	5. 8 ~ 577. 2 mg L^{-1}	Chen et al. 2015
P	linear	NIR-red ratio	$log[R_{rs}(859)]/log[R_{rs}(645)]$	1 ~ 64 mg L^{-1}	Wang et al. 2010a
Q	exponential	red and NIR	$R_t(645)$-$R_t(859)$	0 ~ 12 mg L^{-1}	Hu et al. 2004
R	linear	red and NIR	$R_{rs}(645)$-$R_{rs}(859)$	0. 3 ~ 20 mg L^{-1}	Tarrant et al. 2010
S	linear	two NIR bands	$\rho_w(859)$-$\rho_w(1240)$	74 ~ 881 mg L^{-1}	Wang et al. 2010b
T	exponential	three bands	$R_{rs}(488)$, $R_{rs}(555)$, $R_{rs}(645)$	1~ 300 mg L^{-1}	Zhang et al. 2010
U	semi-analytical	red	nL_w (748)	0. 18 ~ 28. 3 mg L^{-1}	Binding et al. 2010
V	generic single-band	red or NIR	$\rho_w(645)$ or $\rho_w(859)$	1 ~100 mg L^{-1}	Nechad et al. 2010; Polito et al. 2016

*Reflectance bands are as used in the original definition, taking the following forms:
R is the 'surface reflectance' of the MODIS land product.
R_{rs} is above-surface remote-sensing reflectance
R_t is the total radiance observed by MODIS (F_t) divided by the annual mean extraterrestrial solar irradiance F_0.
nL_w is the normalized water-leaving radiance.
ρ_w is the water reflectance, which is defined as $\pi L_w(\lambda) / E^0{}_d{}^+(\lambda)$, where L_w is the water-leaving radiance and $E^0{}_d{}^+$ is the above-water downwelling irradiance.

Table 16 Summary of candidate new Chla algorithms for OLCI

Algorithm	Architectural approach	Formular	Original training (mg.m-3)	reference
OC4_OLCI	Blue-green ratios	MBR=Rrs(443>490>510)/Rrs560	0.01 to 78	(O'Reilly and Werdell 2019)
OC5_OLCI	Blue-green ratios	MBR=Rrs(413>443>490>510)/Rrs560	0.01 to 78	(O'Reilly and Werdell 2019)
OC6_OLCI	Blue-green ratios	MBR=Rrs(413>443>490>510)/M(560&665)	0.01 to 78	(O'Reilly and Werdell 2019)
OC4_MERIS	Blue-green ratios	MBR=Rrs(442>490>510)/Rrs560	0.01 to 78	(O'Reilly and Werdell 2019)
OC5_MERIS	Blue-green ratios	MBR=Rrs(412>442>490>510)/Rrs560	0.01 to 78	(O'Reilly and Werdell 2019)
OC6_MERIS	Blue-green ratios	MBR=Rrs(412>442>490>510)/M(560&665)	0.01 to 78	(O'Reilly and Werdell 2019)
Optimized QAA for OLCI	Semi-analytical	1. Modified reference band (λ_0) of 709 or 754 nm in Step 3: $MCI = R_{rs}(709) - R_{rs}(665) - \left(R_{rs}(754) - R_{rs}(665)\right) * \frac{709-665}{754-665}$	5 to 100	(Liu, Li et al. 2020)

		If MCI≤0.0016, λ_0 =709 nm, else, λ_0 =754 nm 2. Modified $a_{ph}(665)$ equation and η value in step 7 $a_{ph}(665) = \eta\, a_{nw}(560) + (1-\eta)\, a_{nw}(665)$		
MDN	Machine learning	MDN	0.2 to 1209	(Pahlevan, Smith et al. 2020, Pahlevan, Smith et al. 2021, Smith, Pahlevan et al. 2021)
Bayesian	Bayesian probabilistic neural networks	BNN	0.05 to 68	(Werther, Odermatt et al. 2022)
Smith18	Switched blending (G2B and OCI)	$Chl_{blend}(mg\ m^{-3}) = \alpha_1 Chl_{G2B} + \alpha_2 Chl_{OCI}$ Where $\alpha_1 = (Ø - 0.75)/(1.15 - 0.75)$; $\alpha_2 = (1.15 - Ø)/(1.15 - 0.75)$ and Ø=$R_{rs}(708)/R_{rs}(665)$ G2B algorithm refers to Gilerson, Gitelson et al. (2010); $Chla\ [mg\ m^{-3}] = \left(A \times \left(\frac{R(709)}{R(665)}\right) + B\right)^C$ OCI algorithm refers to combined CI (Chl<0.25) and OC4E (Chl>0.25) algorithms: $CI = R_{rs}(560) - [R_{rs}(443) + (560-443)(665-443) * (R_{rs}(665) - R_{rs}(443))]$ $Chl_{CI}(mg\ m^{-3}) = 10^{(a1+b1*CI)}$ $Chl_{OC4}(mg\ m^{-3}) = 10^{(a+b\times X+c\times X^2+d\times X^3+e\times X^4)}$	0.43 to 309	(Smith, Lain et al. 2018)

Table 17 Summary of candidate new Turbidity/TSM algorithms for OLCI

Algorithm	Architectural approach	Formular	Original training (g.m-3)	reference
SOLID20	MDN-based bbp inversion	Classification based	0.1 to 2626.8	(Balasubramanian, Pahlevan et al. 2020)
Jiang21	Semi-analytical	Classification based	0.09 to 2627	(Jiang, Matsushita et al. 2021)
Novoa21G	Switch blending	Linear-Green (tsm<10): $a \times R_w(560)$ Linear-Red (tsm 10~50): $b\times R_w(665)$ Poly-NIR (tsm>50): $c \times R_w(865)^2 + d \times R_w(865)$	2.6 to 1579.1	(Novoa, Doxaran et al. 2017)

Novoa21B	Switch blending	Linear-Green (tsm<10): $a \times R_w(560)$ Nechad et al. (2010) NIR (tsm 10~50): $\frac{b \times R_w(665)}{1-R_w(665)/c}$ Nechad et al. (2010) NIR (tsm>50): $\frac{d \times R_w(865)}{1-R_w(865)/e}$	17.8 to 340.6	(Novoa, Doxaran et al. 2017)
Uudeberg20-clear	Band ratios	$logTSM = a \times R_w560 + b \times R_w665 + c \times \frac{R_w490}{R_w560} + d$	0.5 to 215.2	(Uudeberg, Aavaste et al. 2020)
Uudeberg2020-moderate	Band ratios	$TSM = a \times \left(R_w865 - \frac{R_w778.75 + R_w865}{2}\right) + b$	0.5 to 215.2	(Uudeberg, Aavaste et al. 2020)
Uudeberg20-Turbid	Band ratios	$TSM = a \times \left(R_w865 - \frac{R_w778.75 + R_w865}{2}\right) + b$	0.5 to 215.2	(Uudeberg, Aavaste et al. 2020)
Uudeberg20-VeryTurbid	Band ratios	$logTSM = a \times logR_w560 + b \times R_w665 + c \times log\frac{R_w490}{R_w560} + d$	0.5 to 215.2	(Uudeberg, Aavaste et al. 2020)
Uudeberg20-Brown	Band ratios	$lnTSM = a \times \left(R_w708.75 - \frac{R_w778.75 + R_w665}{2}\right) + b$	0.5 to 215.2	(Uudeberg, Aavaste et al. 2020)
ANTA21 (Turbidity) (based on Nechad 2009, tunned for OLCI)		T(red) was used if RW(red) < 0.05, and T(NIR) if RW(red) > 0.07, with a linear blending in the transition. Red=665 nm NIR=865 nm $T(\lambda) = \frac{A(\lambda) * R_w(\lambda)}{\frac{1 - R_w(\lambda)}{C(\lambda)}}$	0.83 to 176 FNU	(Nechad, Ruddick et al. 2009, Dogliotti, Ruddick et al. 2015, Klein, Lantuit et al. 2021)

5.2 Validation results

Prior to and in parallel to the Lakes_cci (e.g. ACIX-Aqua), extensive validation exercises reported on the performance of satellite-derived LWLR against in-situ remote-sensing reflectance (predominantly from above-water measurements) as well as on the retrieval of chlorophyll-a from atmospherically corrected LWLR. The majority of these analysis have focussed on MERIS and OLCI sensors.

5.2.1 Atmospheric correction algorithms

5.2.1.1 Comparative algorithm selection

For LWLR, six algorithms for MERIS were initially compared: MEGS8.1 (MERIS default), FUB, CoastColour, Case2Regional, SCAPE-M and POLYMER. From these results (Figure 17 to Figure 19 give examples of MEGS and the best performing algorithms), POLYMER was selected based on its superior linearity and correlation with in situ data despite a significant negative bias, which appeared associated with overestimation of the atmospheric radiance component rather than the water model, which is not yet well understood. Linearity in the response nevertheless suggested that algorithms for the retrieval of chlorophyll-a, total suspended matter or turbidity could be tuned to reproduce in situ observations. This procedure is described in more detail in the E3UB document.

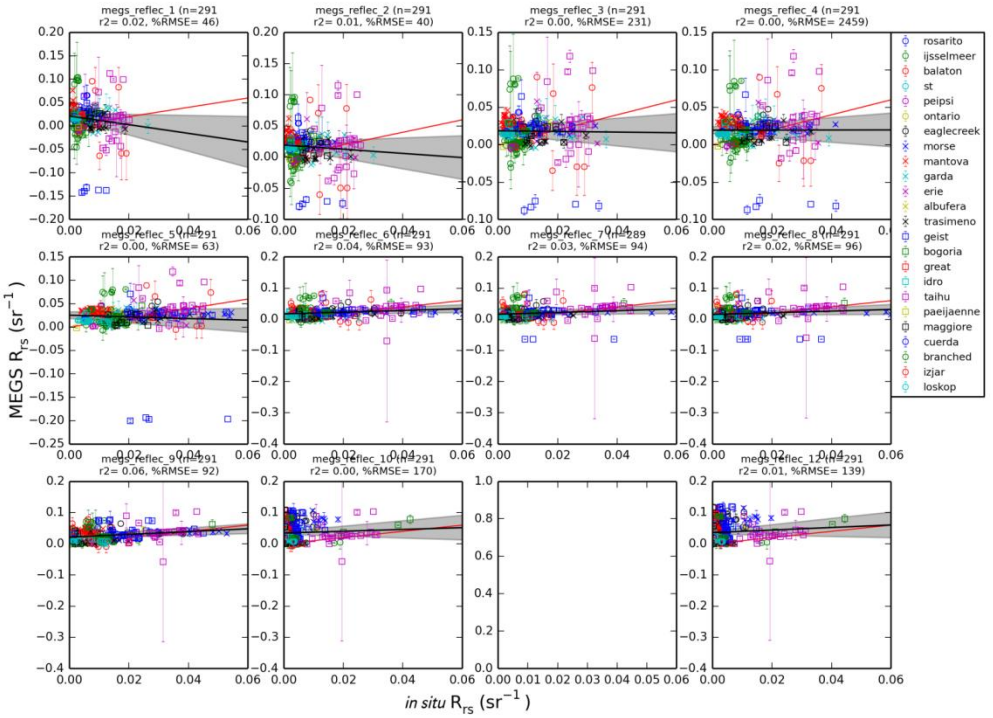


Figure 17 LWLR retrieval per MERIS waveband using the default MEGS algorithm. Matchups are for a ±7-day matchup window and 3x3 pixel extraction window and include results of 23 lakes.

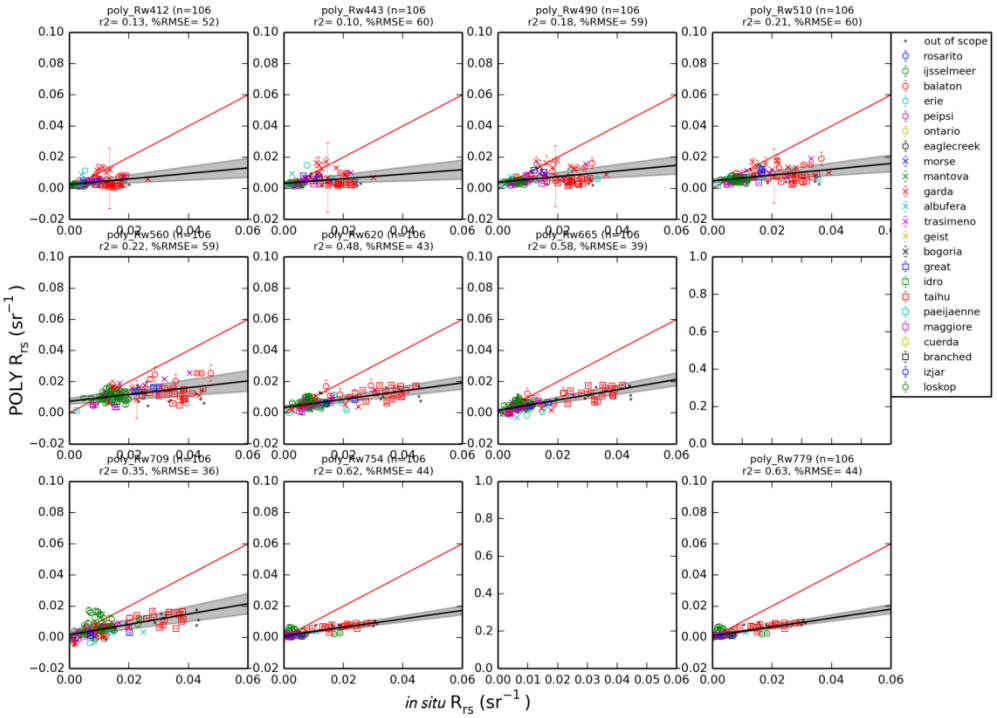


Figure 18 As previous but for the Polymer algorithm.

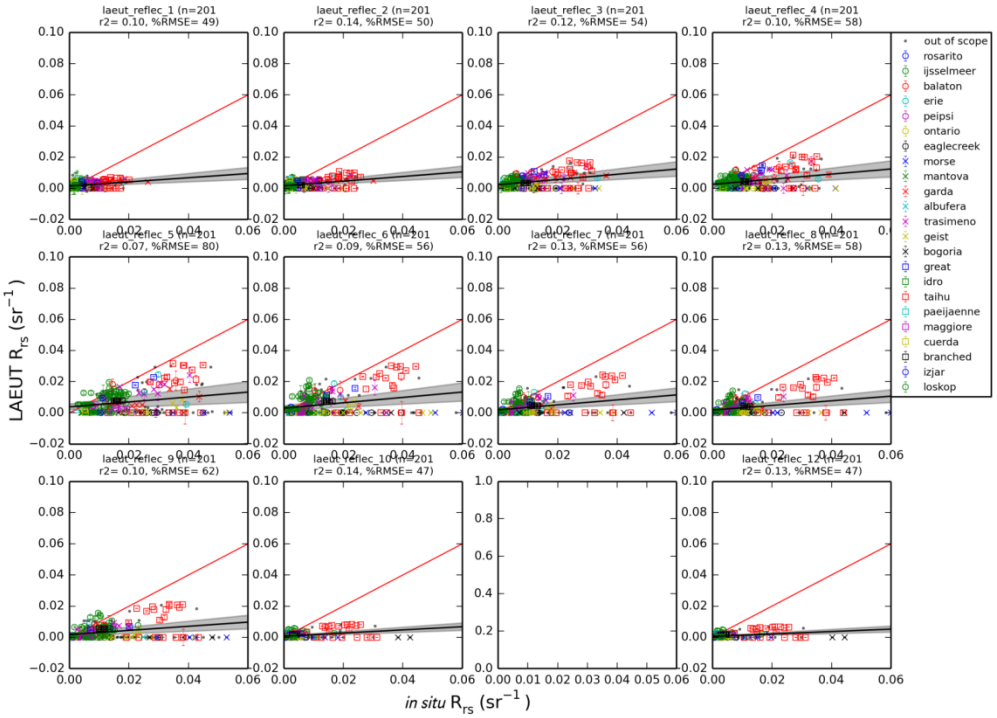


Figure 19 As previous but for the Lakes C2R algorithm.

The results illustrated above gave reason to select Polymer as the most performant algorithm for MERIS and, by extension, OLCI. For MODIS, two potentially suitable algorithms were identified and compared (l2gen and Polymer). The result for Polymer v4.12 is shown in Figure 20. Results were highly consistent between the two algorithms (Figure 21). Polymer was selected for MODIS because it generates approximately 40% more results. Full validation results are provided in the PVIR document.

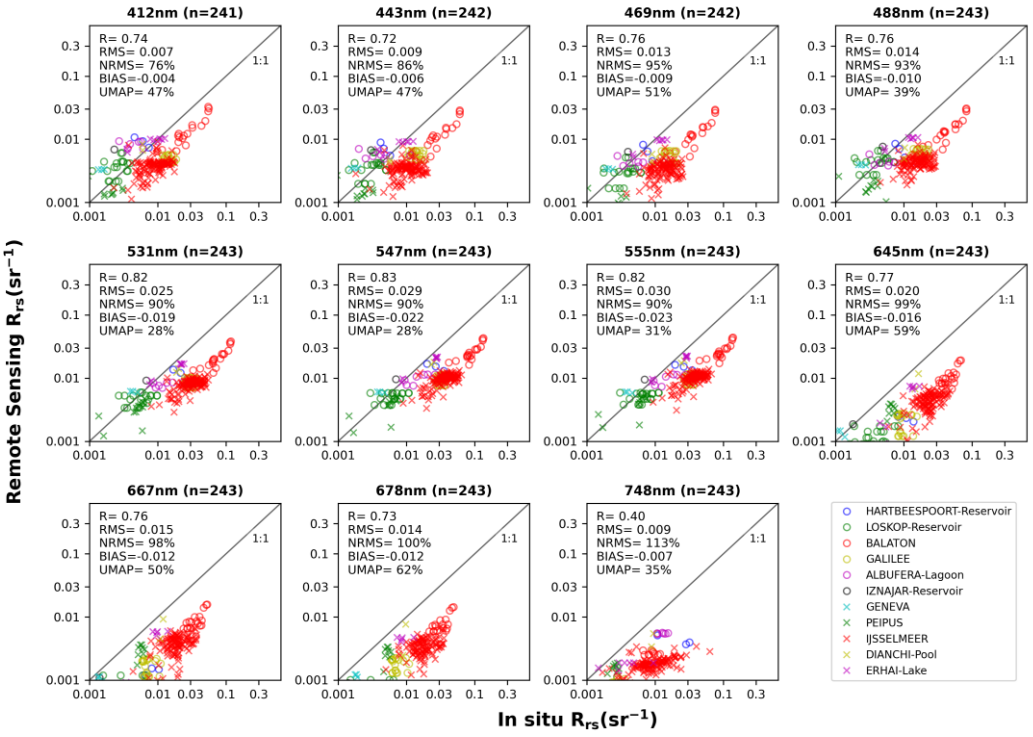


Figure 20 POLYMER v4.12 LWLR from MODIS versus LIMNADES in situ R_{rs} (±3 day matchup and 3×3 macropixel).

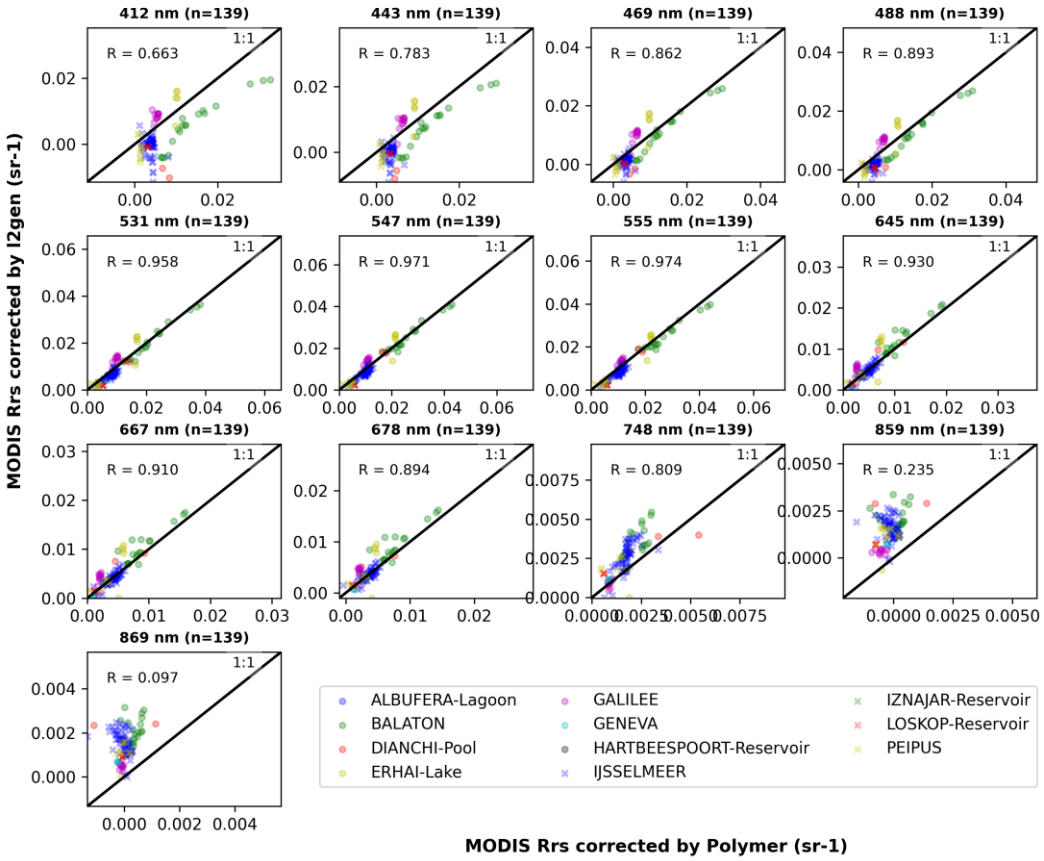


Figure 21 Comparison between POLYMER and l2gen corrected reflectance

5.2.1.2 Selected algorithm configuration

The Polymer algorithm can be configured to use specific bio-optical models and assumptions on the albedo of the water body. In essence, the algorithm then iteratively optimises atmospheric, glint and water-leaving radiance components within set bounds. These bounds are determined by the theoretical (wide) scattering and absorption estimates for inland waters. Due to the optically diverse nature of the global lakes population, such as wide optimisation space is considered necessary. However, it also introduces a risk of ambiguous solutions, which may present in the resulting product as spatio-temporal discontinuities. These can present as speckle (individual pixels) or as contrasting features, and require masking on the basis of consistency between Lakes ECV products, climatology, or acceptable value ranges.

For CRDP v3.0.0, a wider initial parameter space is introduced in the algorithm, which results in a wider output value range of the LWLR product. This is desirable as it is better capable of resolving highly-turbid waters and expected to address systematic underestimations shown in the previous section. On the other hand, this gives rise to larger uncertainties in areas where water-leaving reflectance is poorly resolved against influences from nearby land or the lake bottom. Whilst this may represent a physically more consistent result, post-hoc masking procedures will require additional attention. Additional evidence of improved performance stems from the spectral shape of the LWLR product, as detailed in the PVIR.

A comparison of the new functionality (introduced in Polymer v4.15) is shown in Figure 22 and shows closer correspondence between OLCI derived remote-sensing reflectance (R_{rs}) at 560 nm and the GLORIA in situ database. New outliers are also visible (highlighted in red box) which will need to be flagged in quality control procedures.

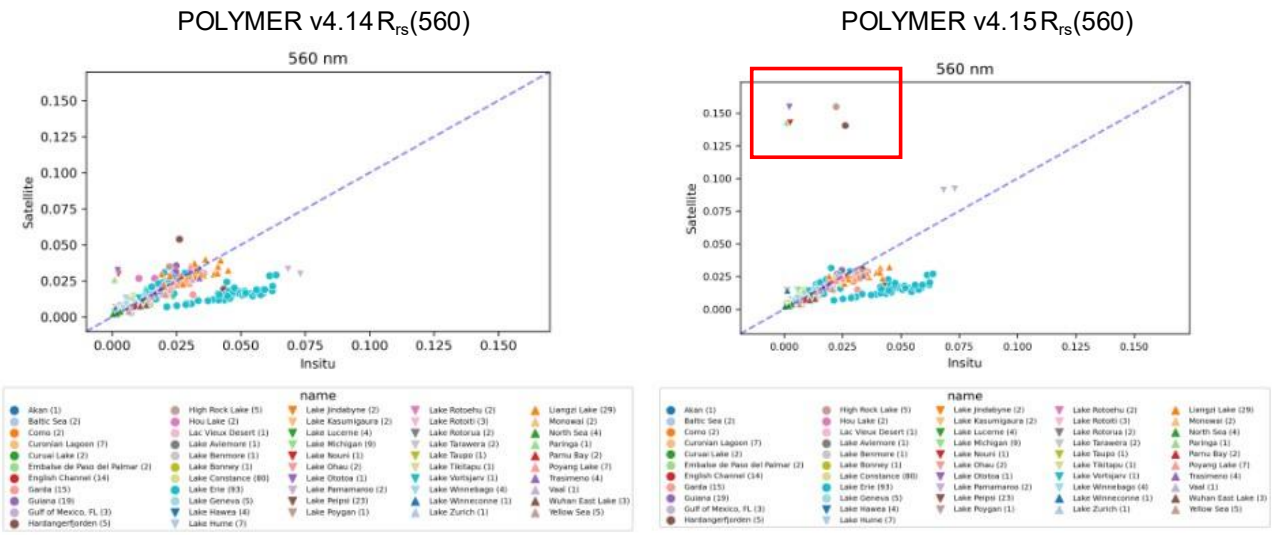
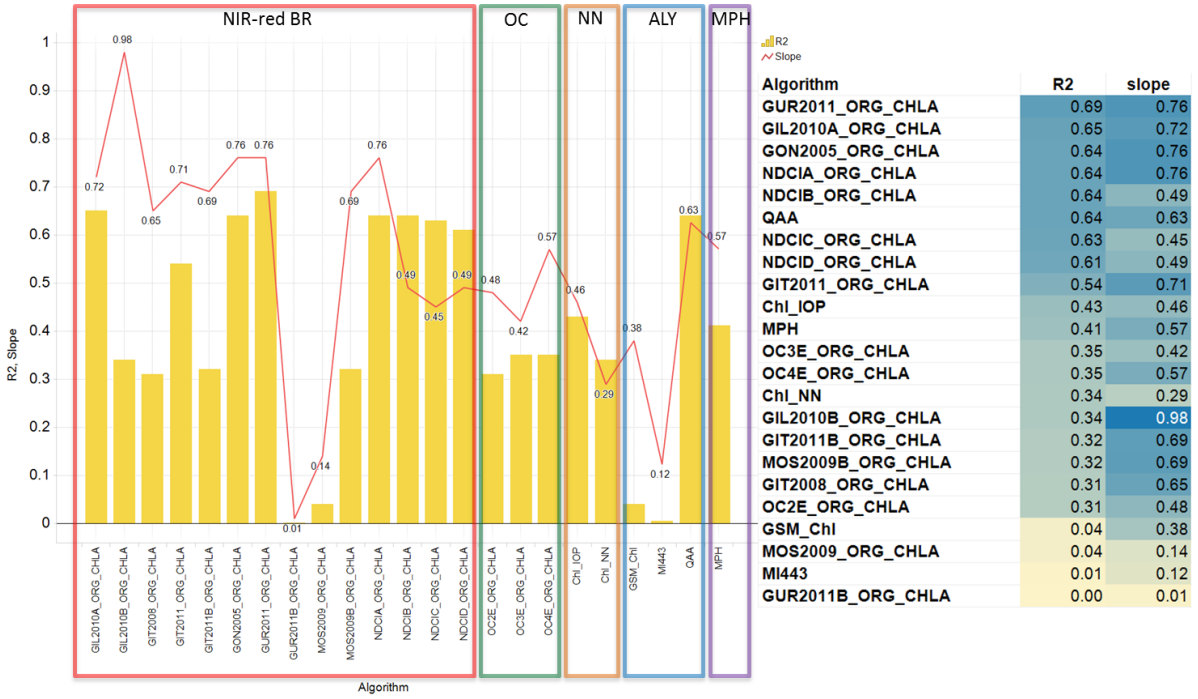


Figure 22 Comparison of Polymer v4.14 and v4.15, the latter including a wider initial optimisation range. In situ data from GLORIA (395 points over 52 water bodies, 3 × 3 macropixel and ±12 hours matchup window).

5.2.2 Chlorophyll-a algorithms

Despite higher uncertainty in the validation of LWLR due to scarce in situ reference data, chlorophyll-a and suspended matter and/or turbidity algorithms may be evaluated and subsequently tuned based on a larger number of matchups with MERIS data in the LIMNADES data set for these measurands. Figure 23 shows results of the round-robin comparison of algorithms for chlorophyll-a in the previous versions of Lakes_CCI (v2.x and before), ultimately resulting in the selection of OC2, a near infra-red (NIR) over red band algorithm based on Gilerson et al. (2010), the semi-analytical NIR-red ratio algorithm of Gons et al. (2005) and a modified Quasi-Analytical Algorithm (QAA) following Mishra et al. (2013). A separate algorithm tuning exercise was also carried out using exclusively in-situ (reflectance and concentration) data, as reported by Neil et al. (2019).

In the algorithm validation towards CRDP v3.0.0, the previous selection of four chlorophyll-a algorithms underwent a rigorous new round-robin comparison including new candidate algorithms developed in recent years, as detailed in Table 16.



Algorithm	R2	slope
GUR2011_ORG_CHLA	0.69	0.76
GIL2010A_ORG_CHLA	0.65	0.72
GON2005_ORG_CHLA	0.64	0.76
NDCIA_ORG_CHLA	0.64	0.76
NDCIB_ORG_CHLA	0.64	0.49
QAA	0.64	0.63
NDCIC_ORG_CHLA	0.63	0.45
NDCID_ORG_CHLA	0.61	0.49
GIT2011_ORG_CHLA	0.54	0.71
Chl_IOP	0.43	0.46
MPH	0.41	0.57
OC3E_ORG_CHLA	0.35	0.42
OC4E_ORG_CHLA	0.35	0.57
Chl_NN	0.34	0.29
GIL2010B_ORG_CHLA	0.34	0.98
GIT2011B_ORG_CHLA	0.32	0.69
MOS2009B_ORG_CHLA	0.32	0.69
GIT2008_ORG_CHLA	0.31	0.65
OC2E_ORG_CHLA	0.31	0.48
GSM_Chl	0.04	0.38
MOS2009_ORG_CHLA	0.04	0.14
MI443	0.01	0.12
GUR2011B_ORG_CHLA	0.00	0.01

Figure 23 Round-robin comparison of chlorophyll-a retrieval algorithms, including NIR-red band ratio algorithms, ocean colour blue-green ratio algorithms, neural networks, analytical (multi-band) inversion algorithms and the maximum peak height algorithm.

5.2.2.1 Chorophyll-a algorithms for MERIS

For MERIS, the evaluation results of thirteen chlorophyll-a algorithms, using their original parameterizations, were assessed against in situ measurements and are presented in Figure 24. Data points in the scatter plots are colour-coded based on the dominant OWT, as indicated by the colour bar on the right side of the figure. Among the evaluated algorithms, the OCX, Smith-Switch, and QAA algorithms exhibited the best performance prior to tuning, with scatter plots closely aligning with the 1:1 line and achieving superior statistical metrics compared to the other algorithms.

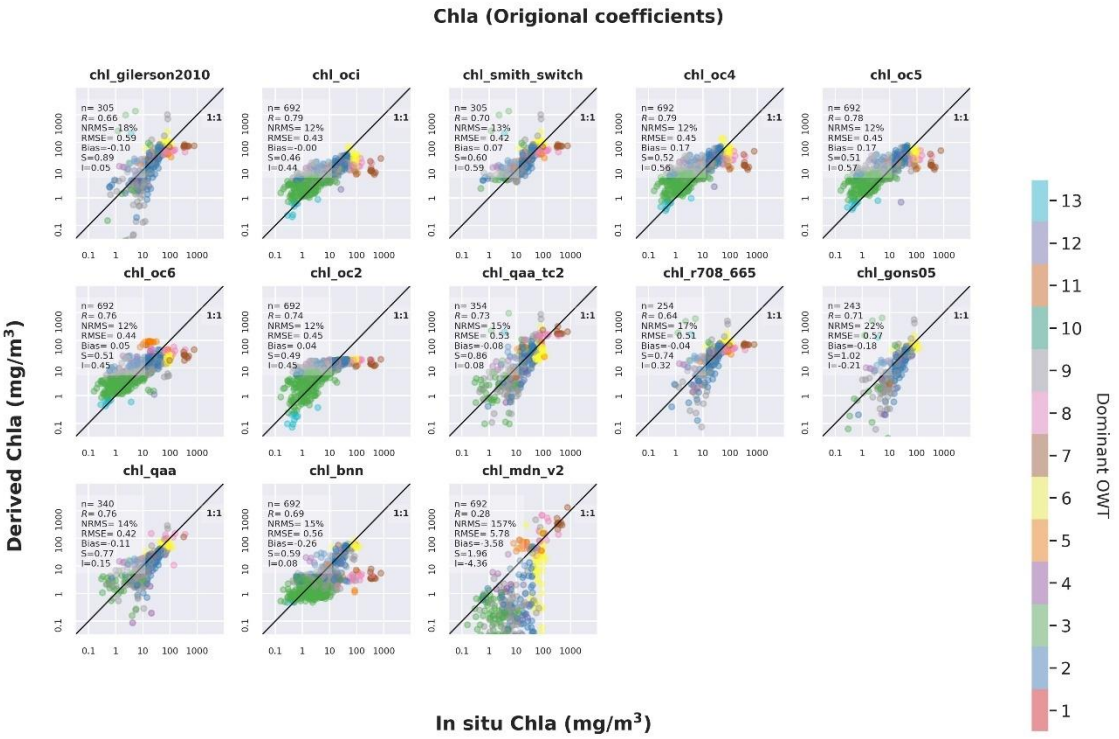


Figure 24 Round-robin comparison of chlorophyll-a retrieval algorithms for MERIS, including the four algorithms used in previous Lakes_cci versions (v2.1 and earlier) for chlorophyll-a product generation (OC2, r708_665, Gons05, and QAA), along with nine newly proposed candidate algorithms.

Algorithm selection and calibration were further refined within an OWT classification framework. Each algorithm was evaluated and tuned based on previously published parameterizations. Using a scoring system to assess algorithm performance (Liu et al., 2024), combined with visual inspection, an optimal algorithm with tuned parameters was assigned to each OWT (Figure 25 and Figure 26). Specifically, OWTs 1, 7, 8, and 10 were assigned to QAA_TC2; OWTs 2, 3, 6, 9, 11, and 13 to OCI; OWTs 4 and 12 to Smith_Switch; and OWT 5 to OC4.

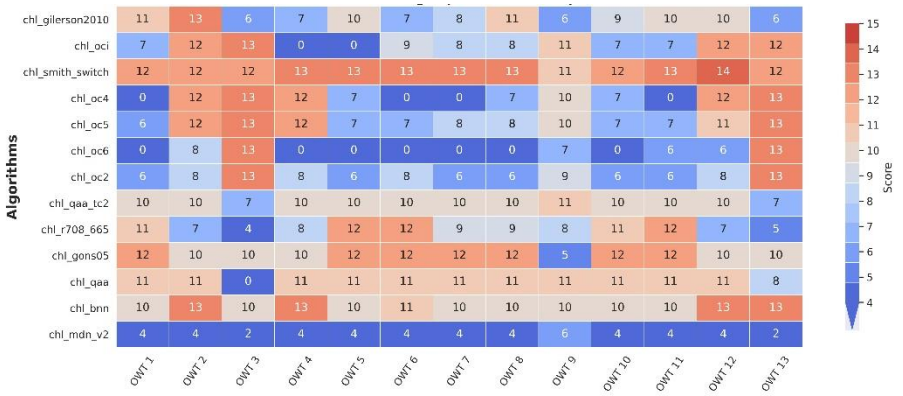


Figure 25 Performance scores of each chlorophyll-a algorithm evaluated for each OWT, derived from MERIS

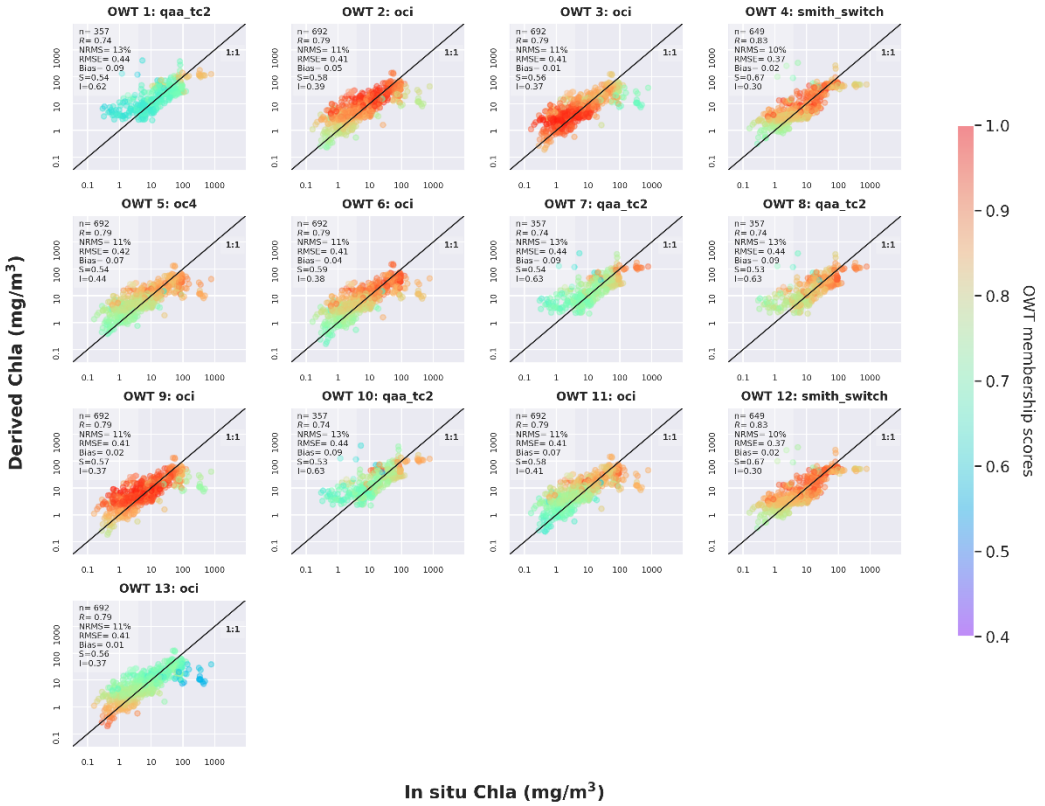


Figure 26 Chlorophyll-a algorithm assignment of each OWT for MERIS

Chlorophyll-a concentrations retrieved by satellite from each assigned algorithm were mapped and blended based on per-pixel OWT membership scores. Specifically, the algorithm outputs corresponding to the top three OWTs were averaged using their membership scores as weighting factors, after normalizing the scores between zero and one, where one represents the highest score and zero represents the score of the fourth-ranked OWT.

The comparison between in situ and blended chlorophyll-a concentrations from MERIS showed (as expected) improved performance, confirming the effectiveness of the tuning process and the subsequent assignment of each algorithm to a specific subset of the concentration range and individual OWTs (Figure 27Figure 30). The regression line closely approaches unity, with an R-value of 0.86, both measured in log space. A Leave-One-Lake-Out Validation (LOLOV) was employed to further evaluate the agreement between in situ measurements and OLCI-derived chlorophyll-a concentrations. LOLOV involves evaluating calibrated algorithms by sequentially excluding data from one lake at a time, repeating this process across all lakes to provide a comprehensive assessment of model performance. The LOLOV results indicated Normalized Root Mean Square Error (NRMS), Root Mean Square Error (RMSE), and bias values ranging from 8.2% to 10.1%, 0.30 to 0.37 mg/m³, and 0.03 to 0.08 mg/m³, respectively (Figure 27), indicating the robust performance of the recalibrated algorithms across diverse inland waters with varying optical properties.

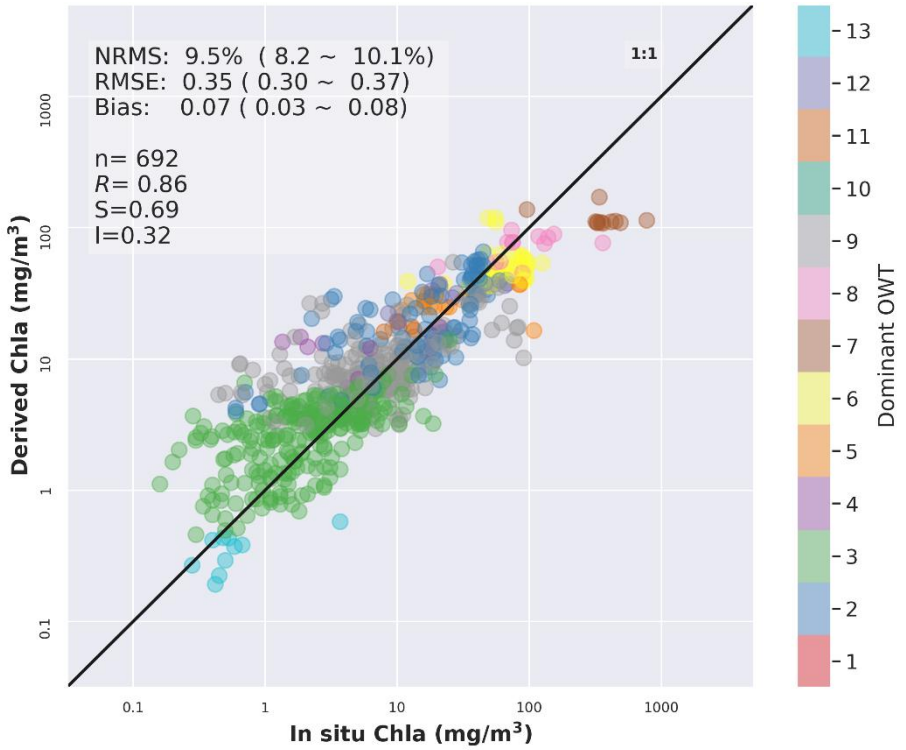


Figure 27 Comparison between in situ and weighted blending chlorophyll-a from the top-3 optical water types, derived from MERIS.

5.2.2.2 Chlorophyll-a algorithms for OLCI

The evaluation results for thirteen chlorophyll-a algorithms for OLCI, using their original parameters, against in situ measurements are presented in Figure 28. The data points are coloured according to the dominant OWT. Overall, the OCX, BNN, and MDN algorithms demonstrated superior performance prior to tuning, with scatter plots close to the 1:1 line and superior statistical metrics compared to other algorithms (Figure 28).

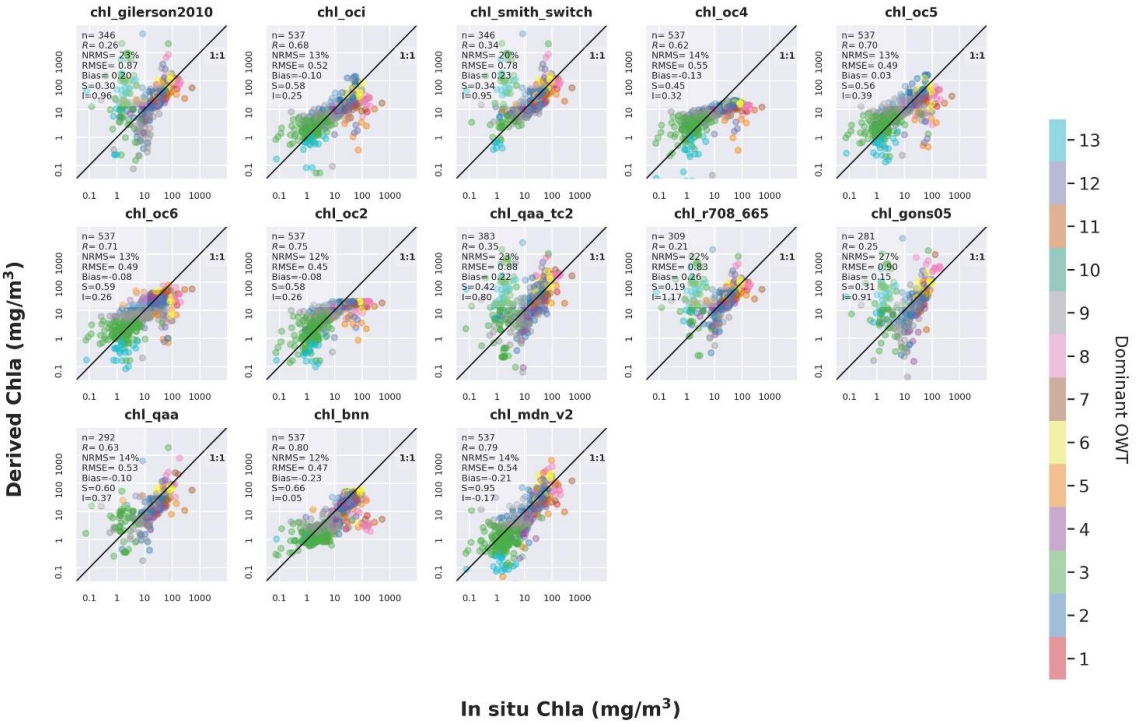


Figure 28 Round-robin comparison (with originally published algorithm coefficients) of chlorophyll-a retrieval algorithms for OLCI, including the four algorithms used in previous Lakes_cci versions (v2.1 and earlier) for chlorophyll-a product generation (OC2, r708_665, Gons05, and QAA), along with nine newly proposed candidate algorithms.

Further algorithm selection and calibration for OLCI chlorophyll-a retrieving were conducted within an OWT classification framework. Algorithms were evaluated and tuned based on previously published parameterizations, based on matchups with at least 80% similarity to a specific OWT. Using a scoring system that assessed algorithm performance (Liu et al. 2024), along with visual inspection, each OWT was assigned an optimal algorithm with tuned parameters (Figure 29 and Figure 30). OWTs 1, 8 and 9 were assigned to MDN, OWT2 is assigned to BNN, OWTs 3 and 13 were assigned to OC2, OWTs 4, 5, 6, 11, and 12 were assigned to Smith_switch, OWTs 7 and 10 were assigned to r708_665.

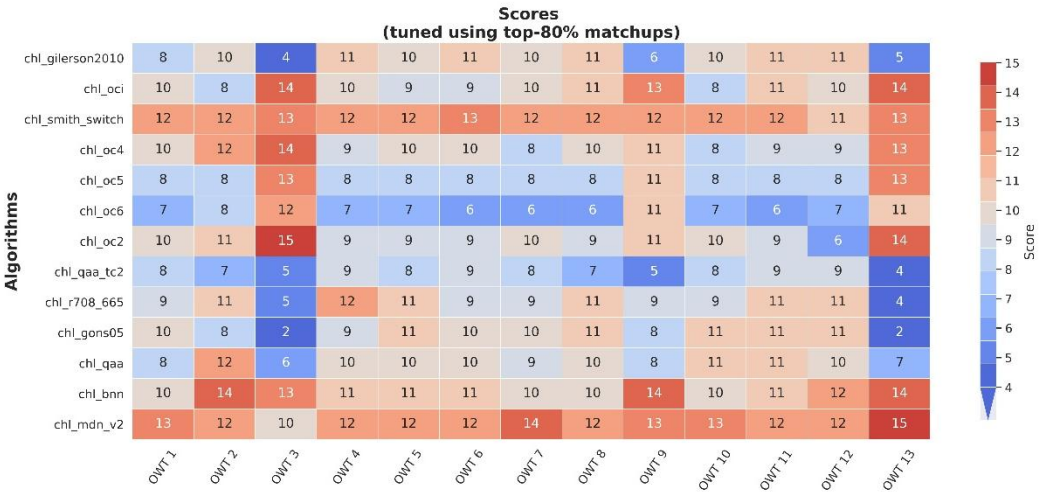


Figure 29 Performance scores of each chlorophyll-a algorithm evaluated for each OWT, derived from OLCI, with tuned algorithm coefficients.

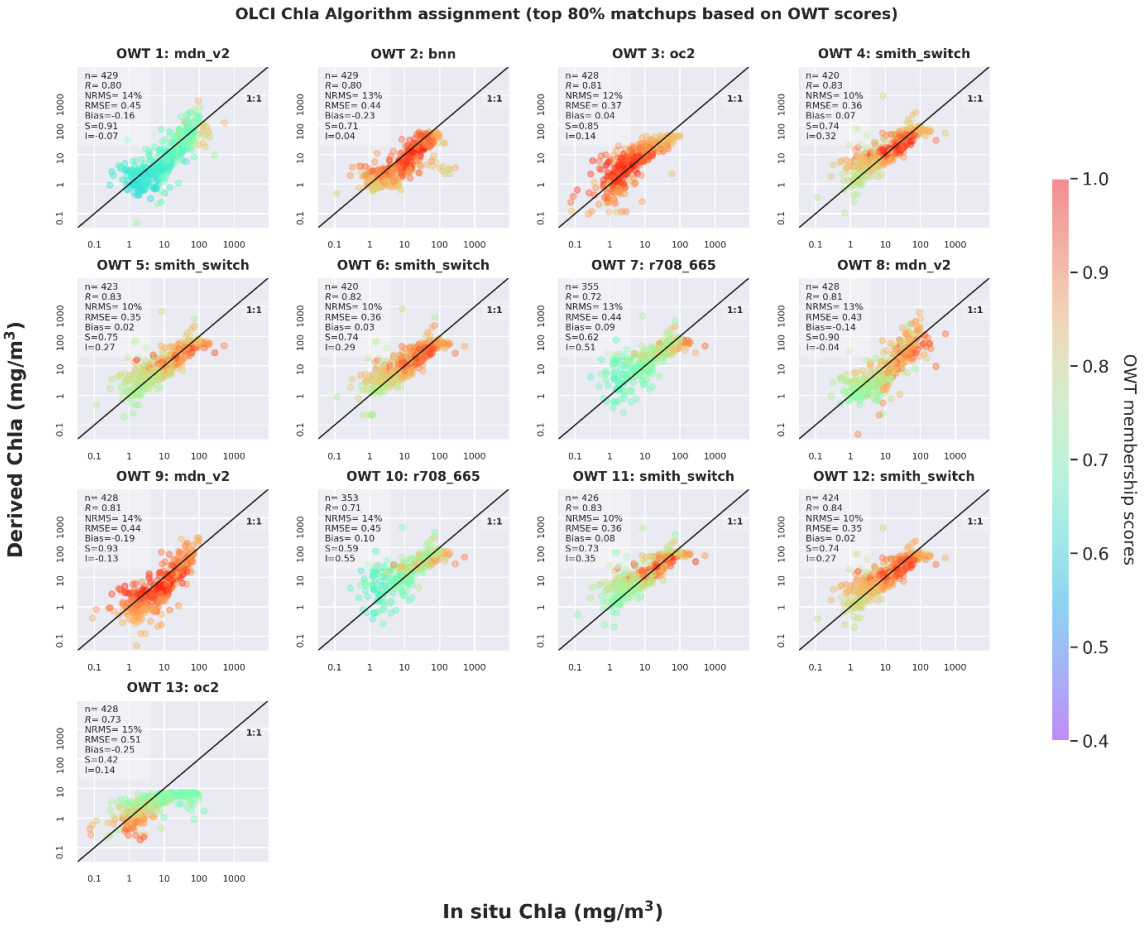


Figure 30 Chlorophyll-a algorithm assignment (following tuning) of each OWT for OLCI

For OLCI, the comparison between in situ and blended chlorophyll-a concentrations showed (as expected) improved performance, confirming the effectiveness of the tuning process and the subsequent assignment of each algorithm to a specific subset of the concentration range and individual OWTs (Figure 28 and Figure 31). The regression line closely approaches unity, with an R-value of 0.88, both measured in log space. The LOLOV results indicated NRMS, RMSE, and bias values ranging from 8.4% to 9.7%, 0.32 to 0.35 mg/m³, and -0.08 to -0.02 mg/m³, respectively (Figure 31).

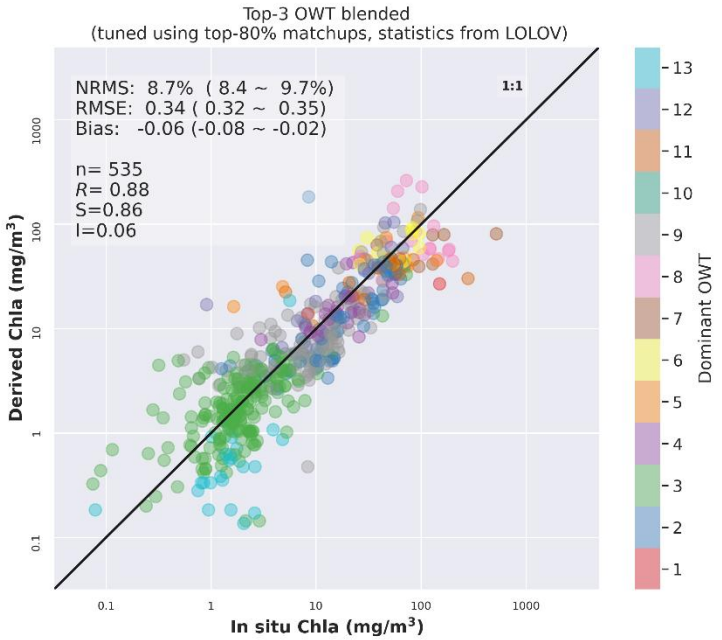


Figure 31 Comparison between in situ and weighted blending chlorophyll-a from the top-3 optical water types, derived from OLCI.

5.2.2.3 Chlorophyll-a algorithms for MODIS

A Round-robin comparison was performed to assess the candidate chlorophyll-a algorithms for MODIS, including four algorithm which were used in CRDP v2.0 (OC3, OC2, r748_667 and OC2m_hi), as well as another eight candidate algorithms. Therefore, there are a total of twelve algorithms compared, including seven blue-green band ratio algorithms, two NIR-red band ratio algorithms, two semi-analytical algorithms and one switching algorithm. Figure 32 presents the performance of these twelve algorithms before tuning, using their original coefficients as provided in the respective publications.

The OCX algorithm, which relies on blue-green band ratios, demonstrated generally strong correlations between satellite-derived and in situ chlorophyll-a concentrations. In contrast, algorithms based on NIR-red band ratios exhibited weaker relationships between satellite-derived and in situ chlorophyll-a concentrations. This is likely due to the lack of a NIR band at 709 nm and use of longer wavebands with lower signal-to-noise ratio, impacting the accuracy of NIR-red-based algorithms. Despite lower specificity to chlorophyll-a across Optical Water Types, deriving the chlorophyll-a signal from blue/green waveband ratios thus provides a more consistent result on average, but with larger uncertainties (including lake-specific bias) compared to MERIS and OLCI algorithm solutions.

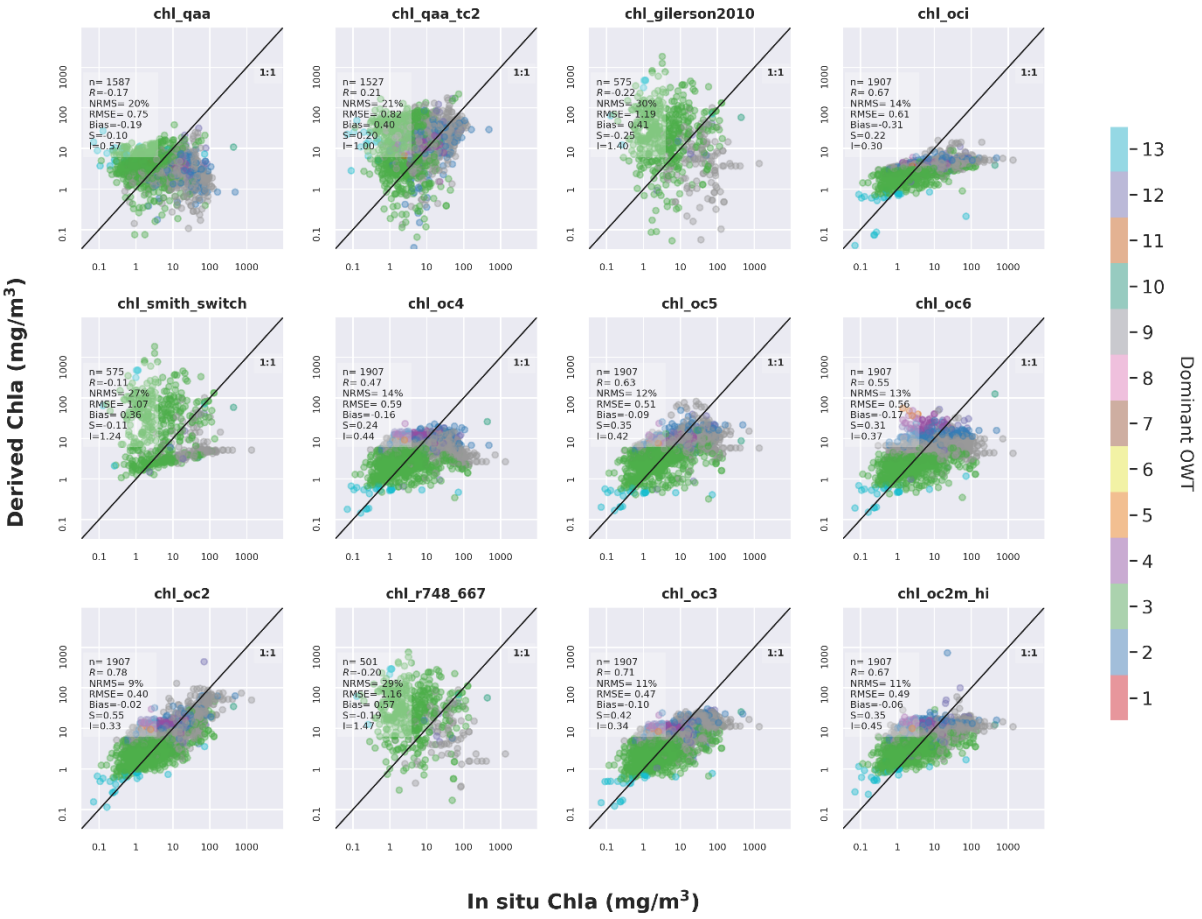


Figure 32 Round-robin comparison of chlorophyll-a retrieval algorithms for MODIS using originally published coefficients, including the four algorithms used for MODIS in previous Lakes_cci versions (v2.1 and earlier) for chlorophyll-a product generation (OC2, r748_667, OC3, and OC2m_hi), along with another eight proposed candidate algorithms.

Further algorithm selection and calibration were performed within an OWT classification framework. For matchups with at least 80% similarity to a specific OWT, chlorophyll-a algorithms were further tuned and evaluated for MODIS. Figure 33 presents a heatmap of the performance scores for each Chl-*a* algorithm after per-OWT tuning, using a scoring system that evaluates algorithm performance (Liu et al., 2024).

Consistent with previous findings, the OCX algorithms exhibit superior performance across nearly all OWTs, surpassing both the NIR-red band ratio algorithms and the two multi-band semi-analytical QAA-based algorithms (Figure 33). Based on a combination of performance scores and visual inspection, the optimal algorithm with tuned parameters was assigned to each OWT. Specifically, OWTs 4, 8, and 11 were assigned to OC3, while all remaining OWTs were assigned to OC2 (Figure 34).

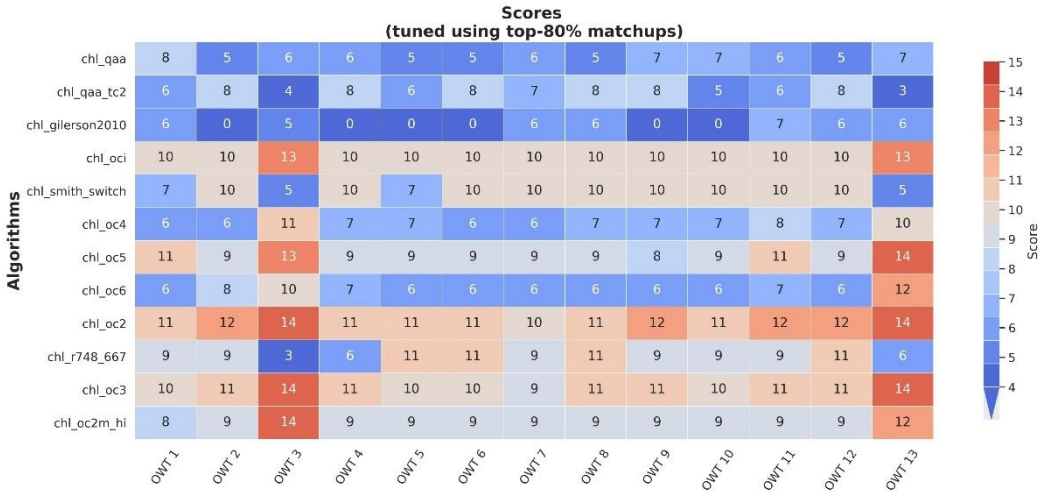


Figure 33 Performance scores of each chlorophyll-a algorithm evaluated after tunning for each OWT, derived from MODIS

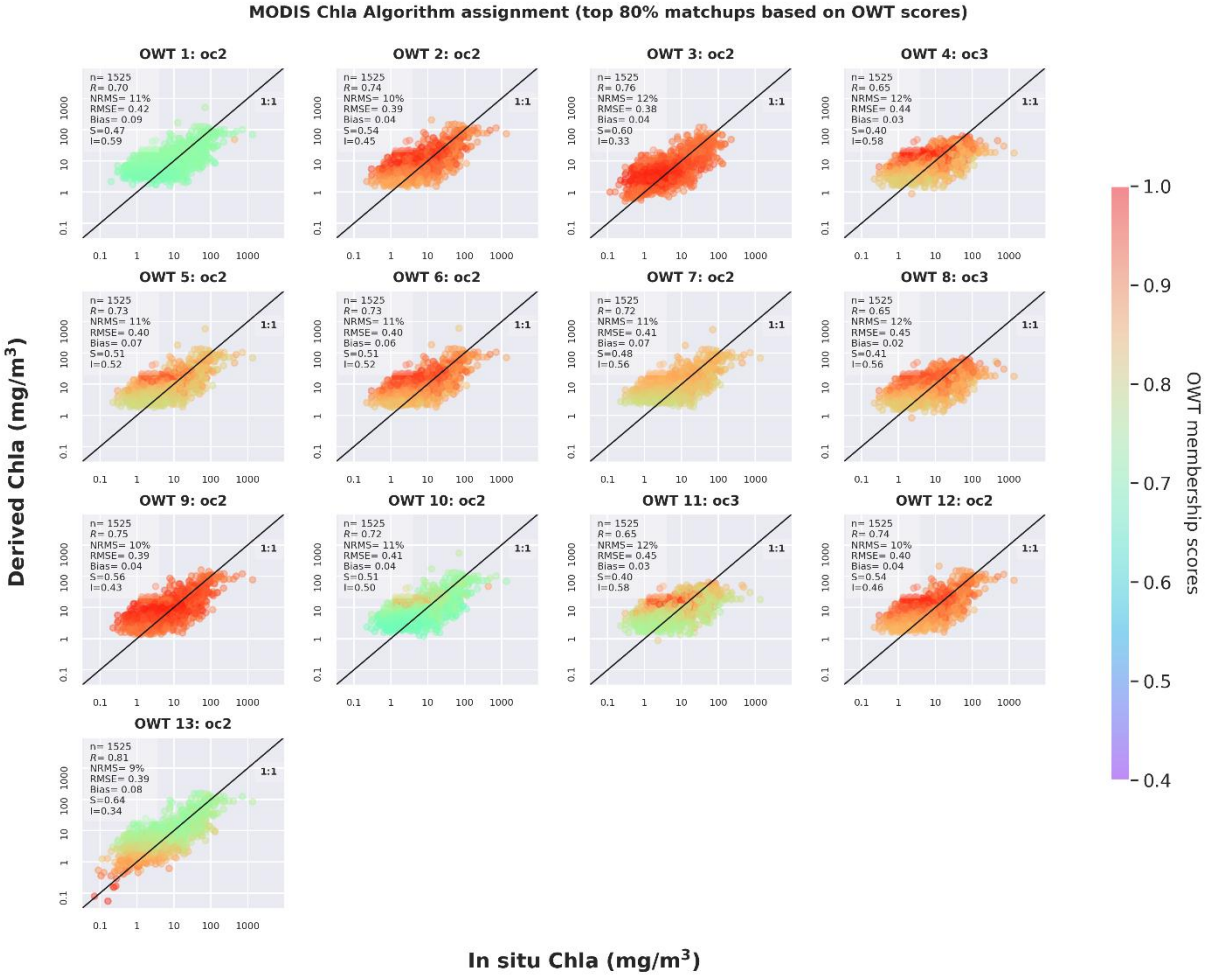


Figure 34 Chlorophyll-a algorithm assiment of each OWT for MODIS

The satellite-derived weighted blending chlorophyll-a from the top-3 OWTs show good aggrement with in situ chlorophyll-a for MODIS (Figure 35). The regression line closely approaches unity, with an R-value of

0.79, both measured in log space. The LOLOV results indicated NRMS and RMSE values ranging from 8.9% to 10.5% and 0.38 to 0.41 mg/m³, respectively (Figure 35). Despite relying exclusively on OCX algorithms with MODIS, per-OWT tuning and subsequent blending provides linearity across the matchup dataset. It is worth noting that OWT membership is also compressed (observations are relatively more similar to multiple OWTs) compared to MERIS and OLCI evaluations, which leads to a narrower diversity of algorithms driving the estimates.

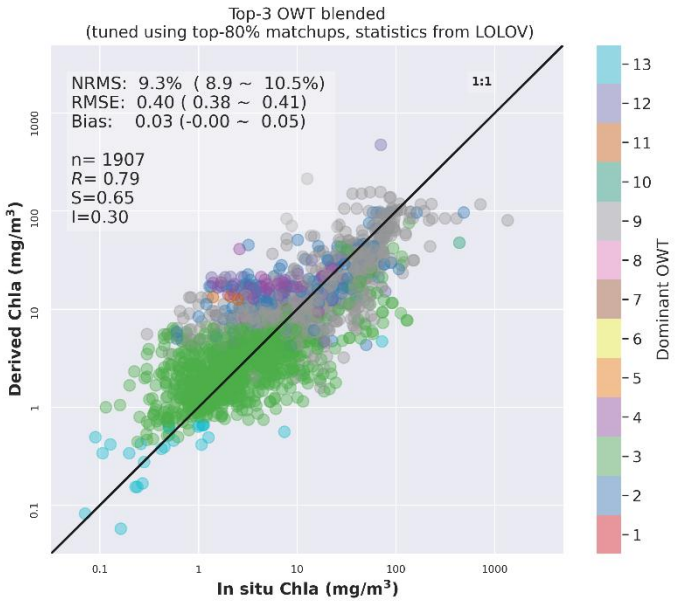


Figure 35 Comparison between in situ and weighted blending chlorophyll-a from the top-3 optical water types, derived from MODIS.

5.2.3 TSM and Turbidity algorithms

5.2.3.1 TSM and Turbidity algorithms for MERIS and OLCI

In Lakes_cci v2.X and earlier versions, an initial algorithm selection was performed during the GloboLakes project using TSM matchups from MERIS. The algorithms proposed by Zhang et al. (2010), Binding et al. (2010), and Vantrepotte et al. (2011) were chosen perviously based on their superior performance across different concentration ranges and optical water types. These selected algorithms were subsequently converted to turbidity values using the coefficients provided by Nechad et al. (2010).

For CRDP v3.0.0, the previously selected three TSM/turbidity algorithms underwent a rigorous round-robin comparison together with candidate algorithms developed in recent years, as detailed in Table 17. A slight change was introduced for v3.0.0, where TSM is provided instead of turbidity. This adjustment was made because the number of in situ TSM measurements is more than three times greater than that of turbidity and the majority of algorithms in this group are originally calibrated against TSM.

5.2.3.1.1 TSM and Turbidity algorithms for MERIS

Figure 36 illustrates the evaluation results of TSM/turbidity algorithms for MERIS, using their original coefficients, against in situ TSM measurements. Data points in the scatter plots are colour-coded based on the OWT, as indicated by the colour bar on the right side of the figure. Among the evaluated algorithms, the Nechad, Klein2021, Jiang2021, Zhang, and Binding algorithms exhibited superior performance prior to tuning, displaying scatter plots more closely aligned with the 1:1 line and achieving better statistical metrics compared to other algorithms.

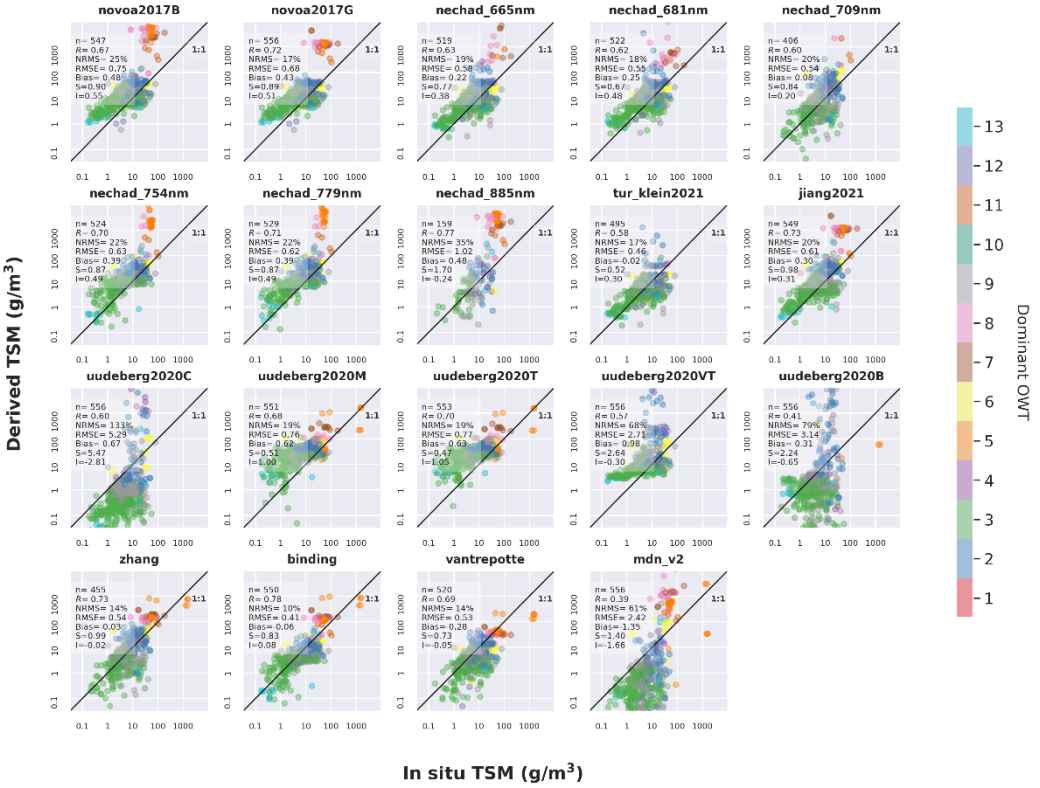


Figure 36 Round-robin comparison of TSM retrieval algorithms for MERIS, including the three algorithms used in previous Lakes_cci versions (v2.1 and earlier) for Turbidity product generation (Zhang, Binding and Vantrepotte), along with 18 newly proposed candidate algorithms.

Further algorithm selection and calibration were conducted for MERIS TSM within an OWT classification framework. For matchups with at top 80% similarity to a specific OWT, the TSM algorithms were evaluated and tuned based on previously published parameterizations. Using a scoring system that assessed algorithm performance (Liu et al. 2024), along with visual inspection, each OWT was assigned an optimal algorithm with tuned parameters (Figure 37 and Figure 38). OWTs 1 was assigned to Uudeberg2020T, OWTs 2, 10 and 11 to Nechad at 709nm, OWTs 3 and 9 to Klein2021, OWTs 4 and 13 to Nechad at 665nm, OWTs 5 and 8 to Nechad at 681nm, OWT 6 to Binding, OWTs 7 and 12 to Zhang (Figure 38).

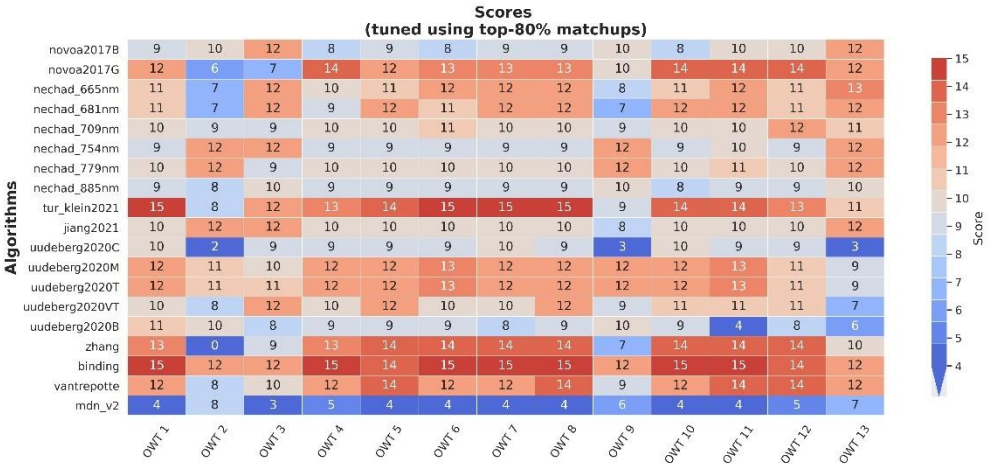


Figure 37 Performance scores of each TSM algorithm evaluated for each OWT, derived from MERIS

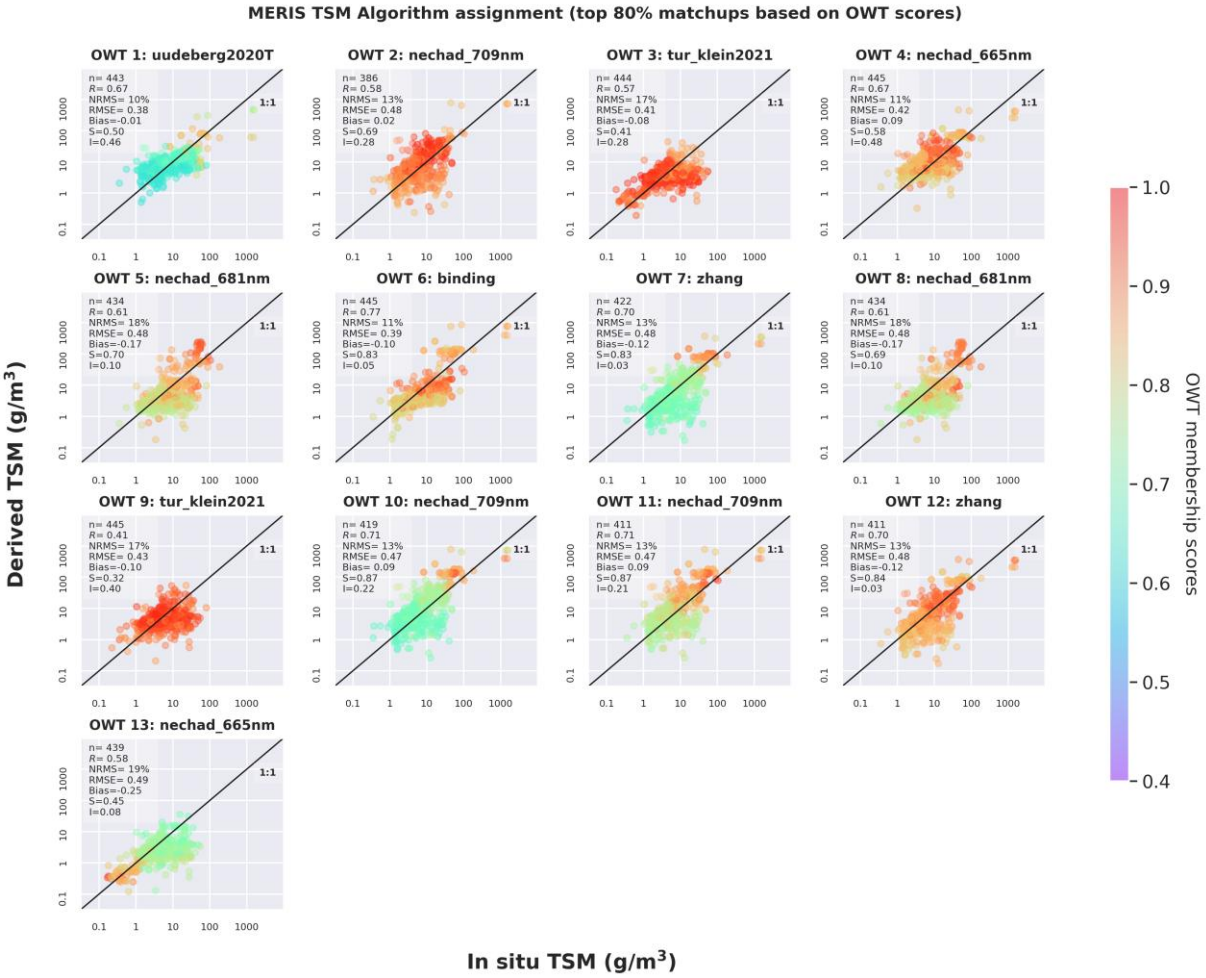


Figure 38 TSM algorithm assignment of each OWT for MERIS

TSM concentrations retrieved by MERIS from each assigned algorithm were mapped and blended based on per-pixel OWT membership scores (Figure 39). The comparison between in situ and top-3 blended TSM concentrations showed (as expected) improved performance, confirming the effectiveness of the tuning process and the subsequent assignment of each algorithm (Figure 39). The regression line closely approaches unity, with an R-value of 0.80, both measured in log space. The LOLOV results indicated NRMS, RMSE, and bias values ranging from 9.2% to 9.8%, 0.37 to 0.39 g/m³, and -0.01 to 0.08 g/m³, respectively (Figure 39Figure 43), demonstrating the robust performance of the recalibrated algorithms across diverse inland waters with varying optical properties.

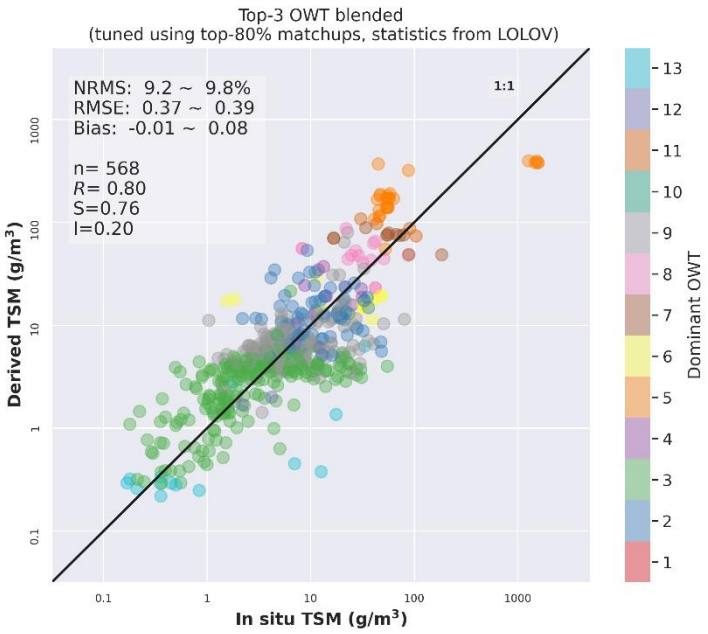


Figure 39 Weighted blending TSM result from combining the top-3 optical water types membership, derived from MERIS. as a function of in situ TSM.

5.2.3.1.2 TSM and Turbidity algorithms for OLCI

Figure 40 presents the evaluation results for TSM/turbidity algorithms for OLCI, using their original parameters, compared against in situ TSM measurements. Data points are coloured according to the OWT. Overall, the Nechad, Zhang, Vantrepotte, Bala2020, and MDN algorithms demonstrated superior performance prior to tuning, with scatter plots close to the 1:1 line and superior statistical metrics compared to other algorithms (Figure 40).

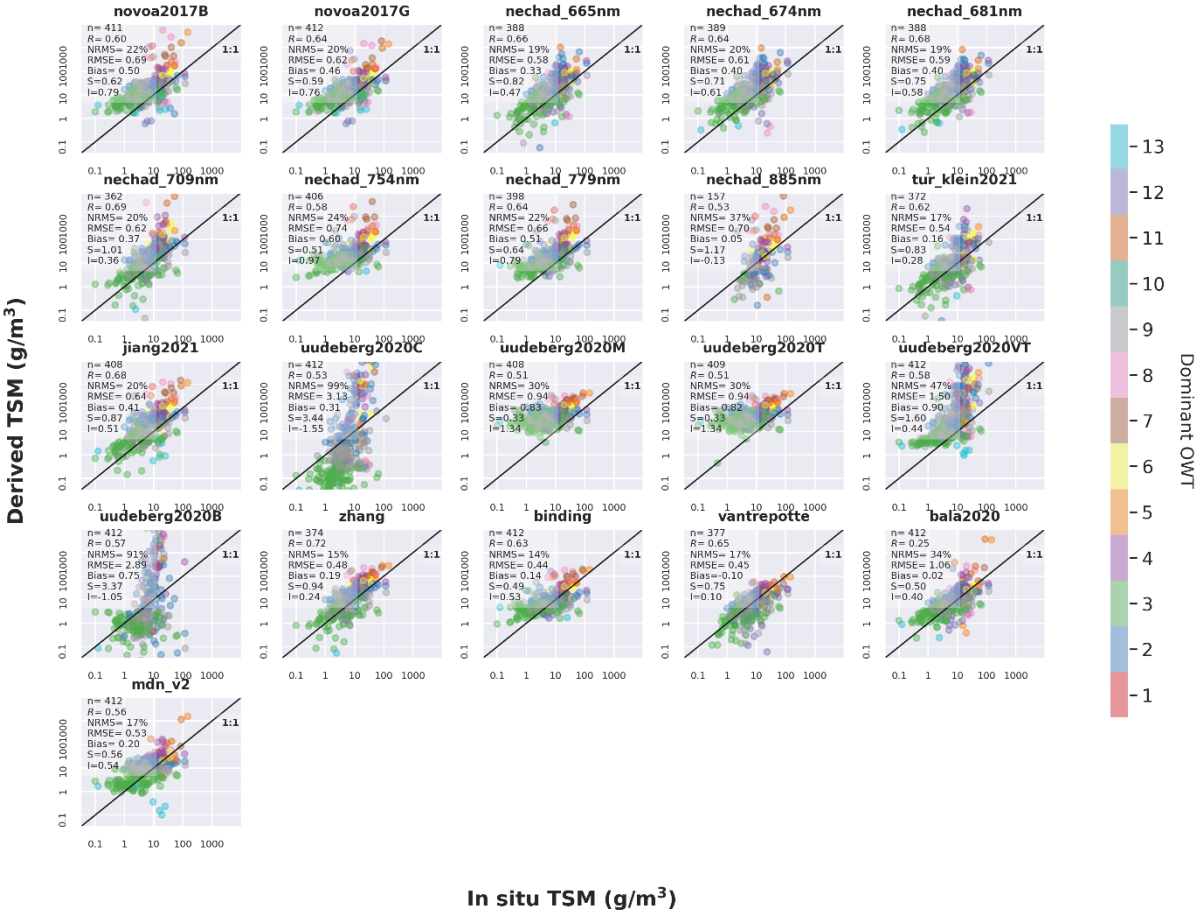


Figure 40 Round-robin comparison of TSM retrieval algorithms for OLCI, including the three algorithms used in previous Lakes_cci versions (v2.1 and earlier) for Turbidity product generation (Zhang, Binding and Vantrepotte), along with 18 newly proposed candidate algorithms.

Algorithm selection and calibration were further refined for OLCI within the OWT classification framework. TSM algorithms were evaluated and tuned, using matchups with top 80% similarity to a specific OWT, based on previously published parameterizations. Using a scoring system to assess algorithm performance (Liu et al., 2024), combined with visual inspection, an optimal algorithm with tuned parameters was assigned to each OWT (Figure 41 and Figure 42). Specifically, OWTs 1 and 5 were assigned to the Vantrepotte algorithm, OWTs 2, 7, and 10 to the Zhang algorithm, OWT 3 to the Nechad algorithm at 665 nm, OWTs 4, 9, and 12 to the Nechad algorithm at 709 nm, OWT 6 to Jiang2021, OWTs 8 and 11 to Binding, and OWT 13 to the Nechad algorithm at 681 nm.

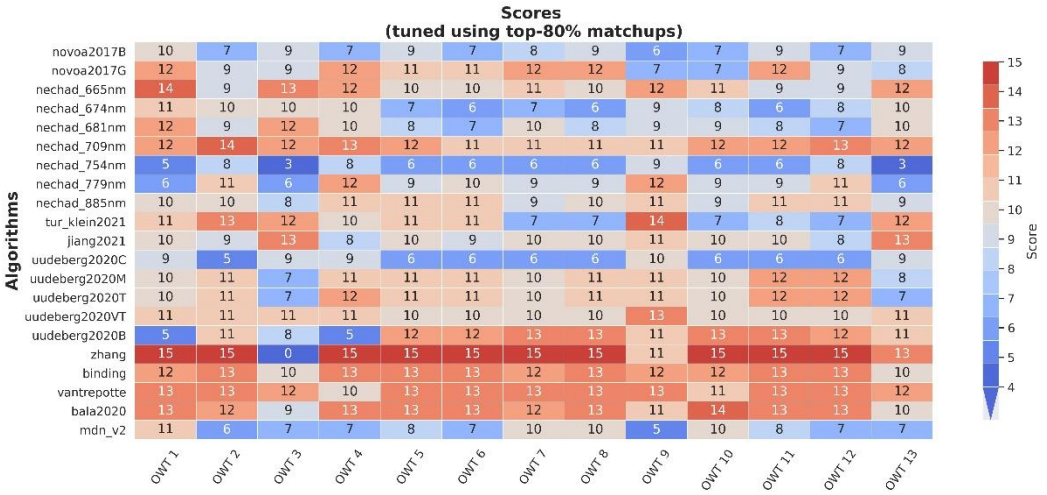


Figure 41 Performance scores of each TSM algorithm evaluated for each OWT, derived from OLCI

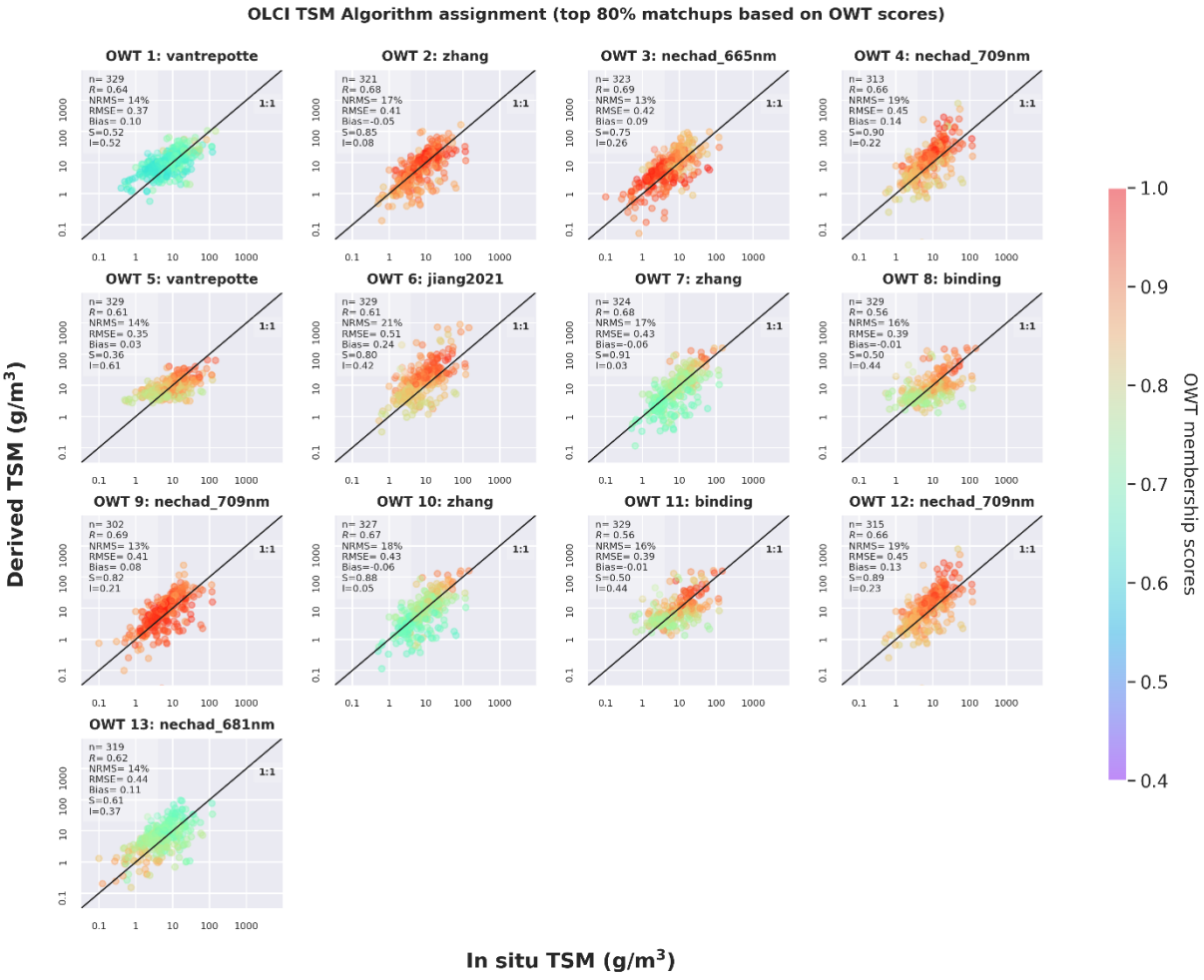


Figure 42 TSM algorithm assignment of each OWT for OLCI

The comparison between in situ and top-3 blended TSM concentrations for OLCI showed improved performance as expected, validating the effectiveness of the tuning process and the subsequent assignment of each algorithm to a specific subset of the concentration range and individual OWTs (Figure

43). The regression line closely aligns with the 1:1 line, achieving an R-value of 0.77, calculated in log space. The LOLOV results indicated NRMS, RMSE, and bias values ranging from 11.7% to 15.8%, 0.37 to 0.43 g/m³, and -0.001 to 0.11 g/m³, respectively (Figure 43). These results highlight the robust performance of the recalibrated algorithms across diverse inland waters with varying optical properties.

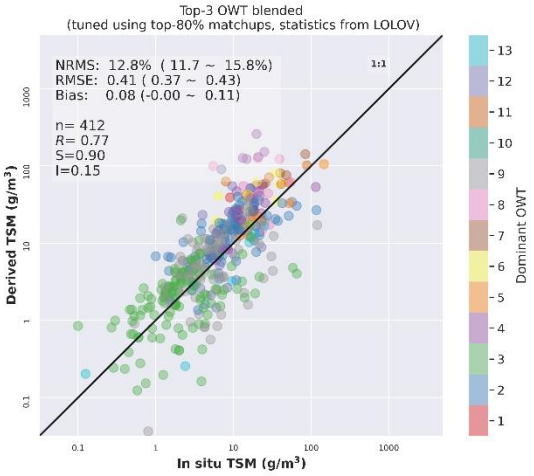


Figure 43 Comparison between in situ and weighted blending TSM from the top-3 optical water types, derived from OLCI.

5.2.3.1.3 TSM and Turbidity algorithms for MODIS

Figure 44 presents the evaluation results of TSM/turbidity algorithms for MODIS, using their original coefficients, against in situ TSM measurements. Among the assessed algorithms, Nechad, Klein, Zhang, and Vantrepotte demonstrated relatively better pre-tuning performance, with scatter plots closely aligning with the 1:1 line and achieving better statistical metrics compared to the other algorithms.

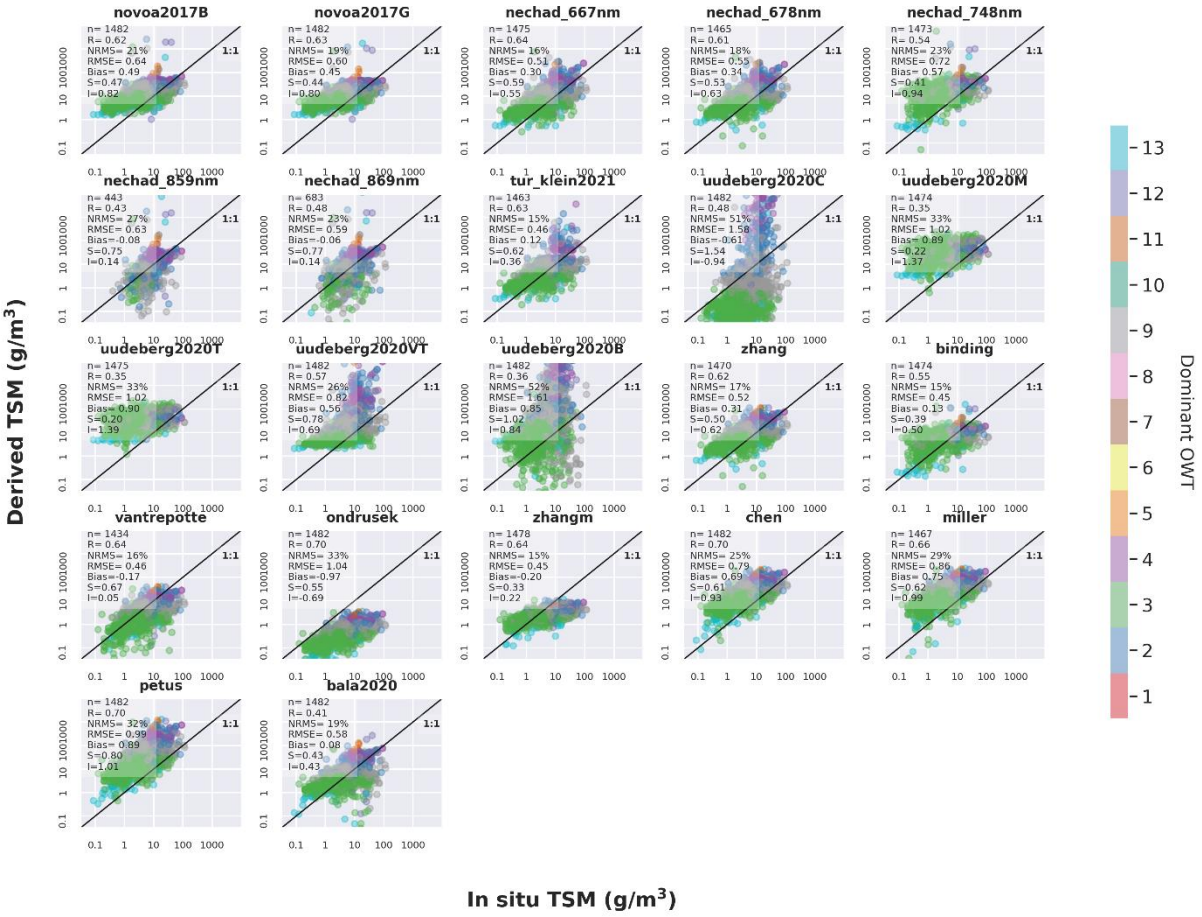


Figure 44 Round-robin comparison of TSM retrieval algorithms for MODIS, including the three algorithms used in previous Lakes_cci versions (v2.1 and earlier) for Turbidity product generation (Ondrusek, ZhangM, Chen, Miller, Petus), along with 17 newly proposed candidate algorithms.

Further algorithm selection and calibration were performed within an OWT classification framework. TSM algorithms were further tuned and evaluated for MODIS, using matchups with top-80% similarity to a specific OWT. Figure 45 presents a heatmap of the performance scores for each TSM algorithm after per-OWT tuning, using a scoring system that evaluates algorithm performance (Liu et al., 2024). Based on statistical analysis and visual checks, the OWTs 1, 5 and 10 were assigned to Nechad at 667 nm, OWTs 2, 7 and 11 assigned to Miller, OWT 3 assigned to Petus, OWTs 4, 6, 8, and 12 assigned to Klein, OWT 9 assigned to Ondrusek, and OWT 13 assigned to Chen.

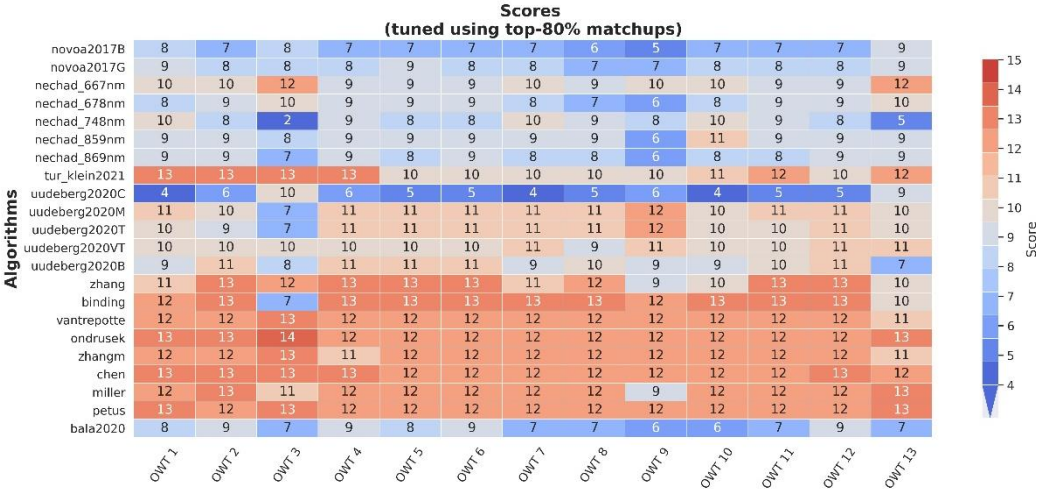


Figure 45 Performance scores of each TSM algorithm evaluated for each OWT, derived from MODIS

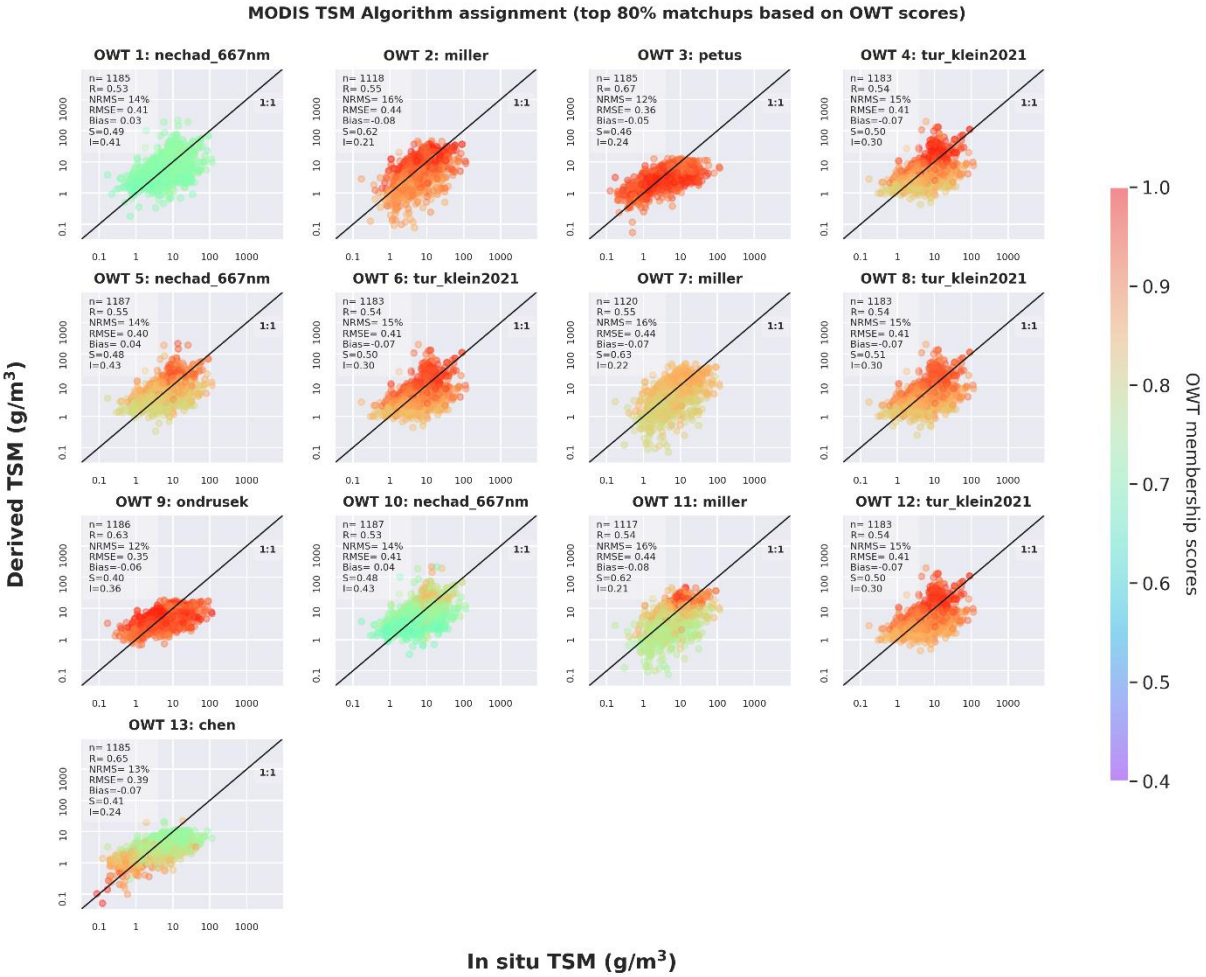


Figure 46 TSM algorithm assignment of each OWT for MODIS

Figure 47 presents a comparison between in situ measurements and the weighted blending of TSM from the top three OWTs for MODIS. The results indicate improved performance compared to individual tuned algorithms, reaffirming the effectiveness of the tuning process and the subsequent algorithm assignments. The regression line closely approaches unity, with an R-value of 0.69, both measured in log space. The LOLOV results indicated NRMS and RMSE values ranging from 12.1% to 13.6%, 0.38 to 0.41

g/m³, respectively (Figure 47Figure 43), demonstrating the robust performance of the recalibrated algorithms across diverse inland waters with varying optical properties.

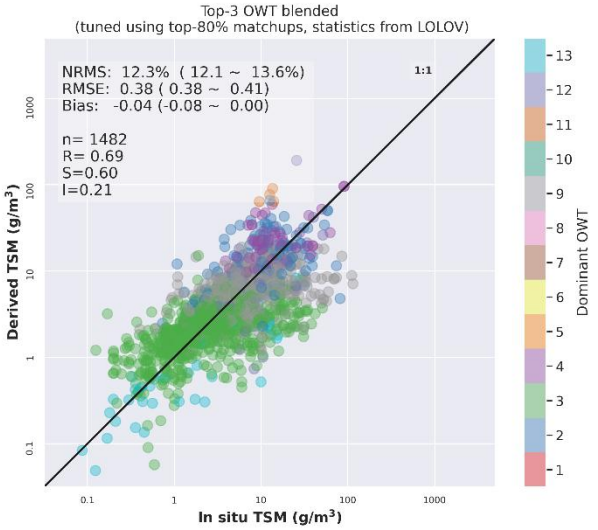


Figure 47 Comparison between in situ and weighted blending TSM from the top-3 optical water types, derived from MODIS.

5.2.4 CDOM light absorption algorithms

Eight satellite algorithms for estimating CDOM light absorption coefficients at 440 nm ($a_{CDOM}(440)$) were initially selected from 16 candidate algorithms and subsequently refined through per-OWT tuning for MERIS, OLCI, and MODIS (see CDOM Technical Note v1.2 available through the project website[1]). This process aimed to determine the most suitable algorithm for $a_{CDOM}(440)$ estimation across different optical water types.

Figure 48, Figure 49 and Figure 50 present the validation results comparing in situ and satellite-derived $a_{CDOM}(440)$ for the three sensors (MERIS, OLCI and MODIS) using the original coefficients from their respective published studies. Overall, the Fick and Tiwari-Shanmugam algorithms exhibited superior performance prior-tuning compared to other evaluated $a_{CDOM}(440)$ estimation methods.

[1] https://climate.esa.int/media/documents/CDOM_Techinical_Note_CCN-D-1_V1.2_final_ZybCWtm.pdf

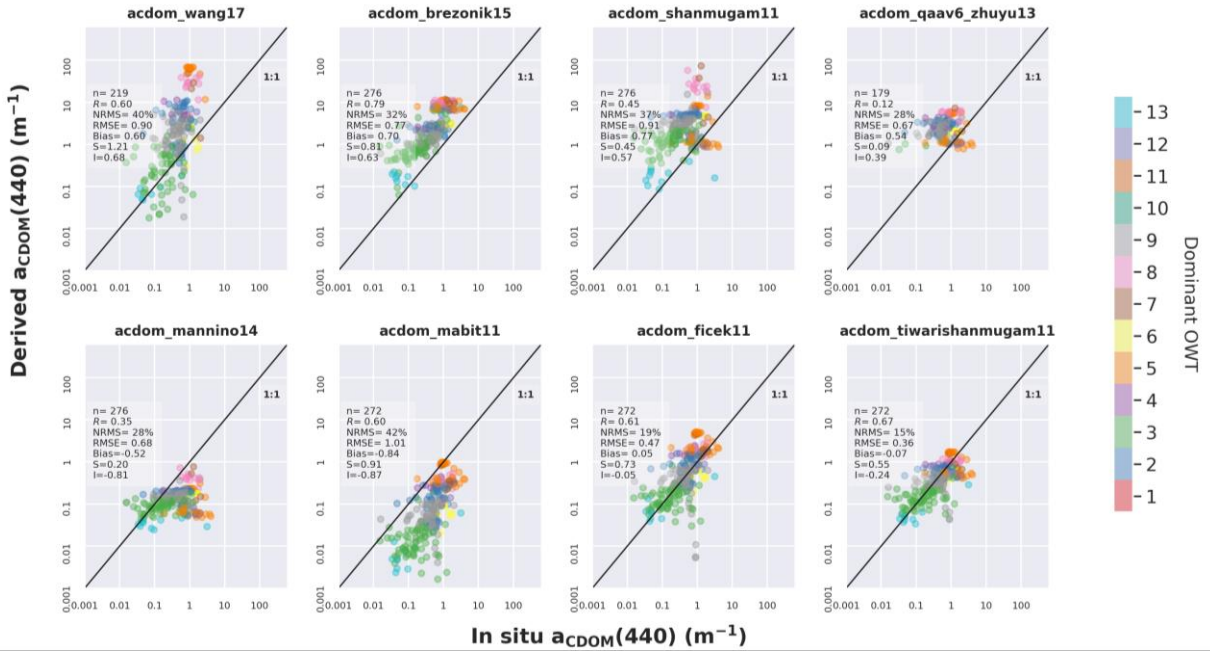


Figure 48 Round-robin comparison of a_{CDOM} retrieval algorithms for MERIS, using original coefficients.

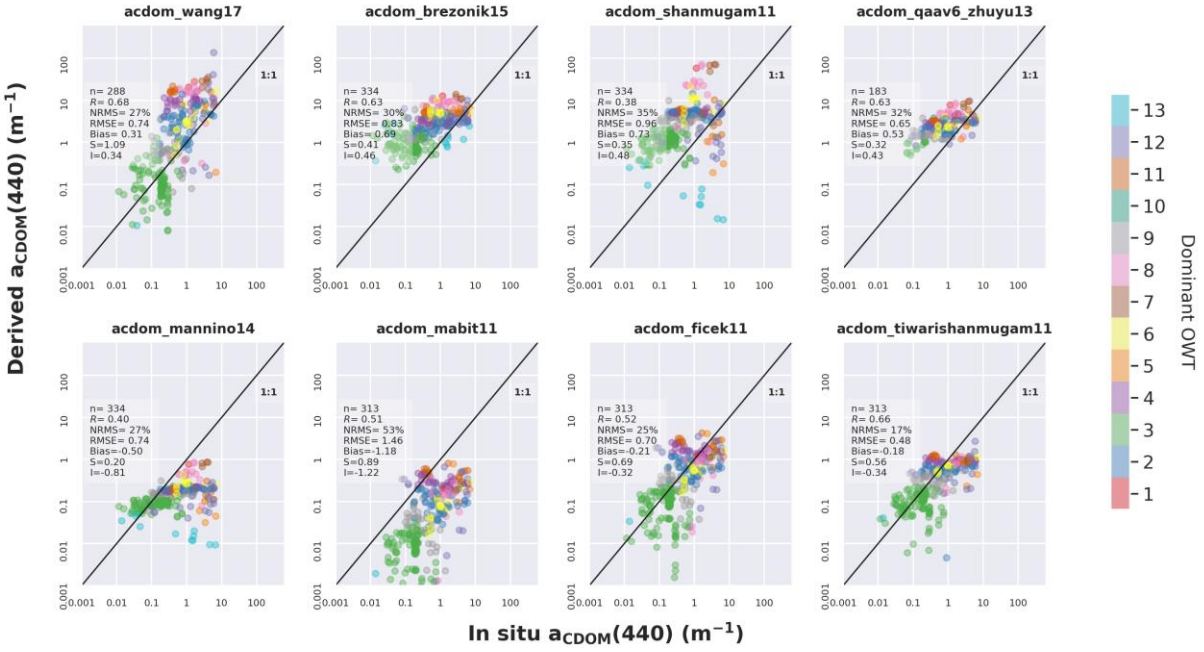


Figure 49 Round-robin comparison of a_{CDOM} retrieval algorithms for OLCI, using original coefficients.

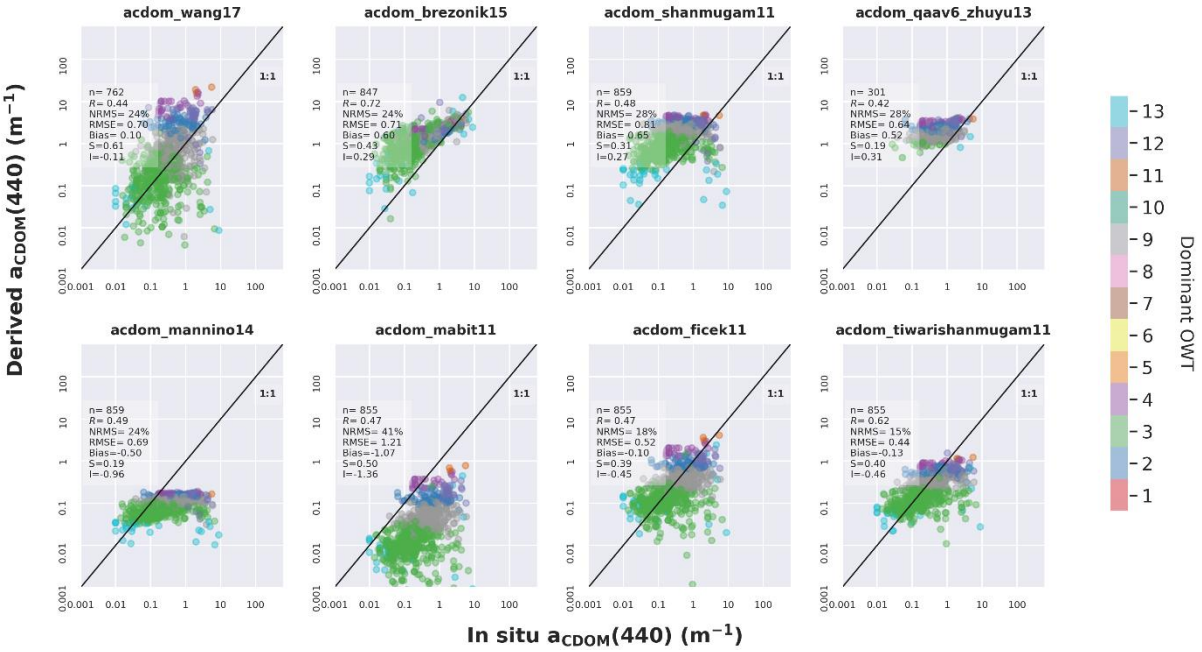


Figure 50 Round-robin comparison of $a_{CDOM}(440)$ retrieval algorithms for MODIS, using original coefficients.

Further algorithm selection and calibration were conducted within an OWT classification framework, using matchups with top-80% similarity to a specific OWT. Figure 51 presents heatmaps illustrating the performance scores of each $a_{CDOM}(440)$ algorithm following per-OWT tuning, based on a scoring system designed to evaluate algorithm effectiveness (Liu et al., 2024). The results indicate that the Brezonik, Manino, and Tiwari & Shanmugam algorithms exhibit superior performance for both MERIS and OLCI. For MODIS, the Brezonik, Manino, Mabit, and Tiwari & Shanmugam algorithms demonstrate the highest performance among all evaluated algorithms.

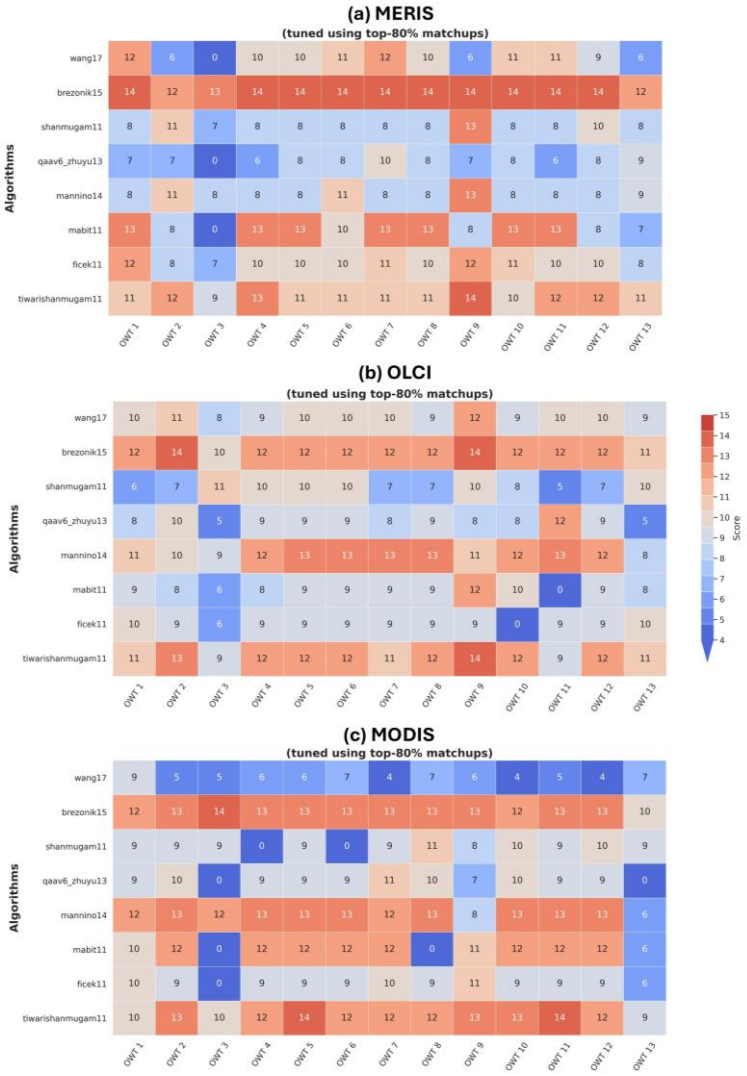


Figure 51 Performance scores of each $a_{CDOM}(440)$ algorithm evaluated for each OWT following per-OWT tuning, for (a) MERIS, (b) OLCI and (c) MODIS

Combining statistical performance scores with visual inspection, the most effective $a_{CDOM}(440)$ algorithm was selected and assigned to each OWT for MERIS, OLCI, and MODIS (Figure 52, Figure 53 and Figure 54). Specifically, for MERIS, OWTs 1, 5, 8, and 10 were assigned to Ficek; OWTs 2, 3, 4, 6, 9, 11, and 12 to Brezonik; and OWTs 7 and 13 to Tiwari & Shanmugam. For OLCI, OWTs 1, 6, 7, 8, and 11 were assigned to Mannino; OWTs 2, 4, 5, 12, and 13 to Brezonik; OWT 3 to Tiwari & Shanmugam; OWT 9 to Shanmugam; and OWT 10 to Wang. For MODIS, the majority of OWTs were assigned to Brezonik, with the exception of OWT 9, which was assigned to Ficek.

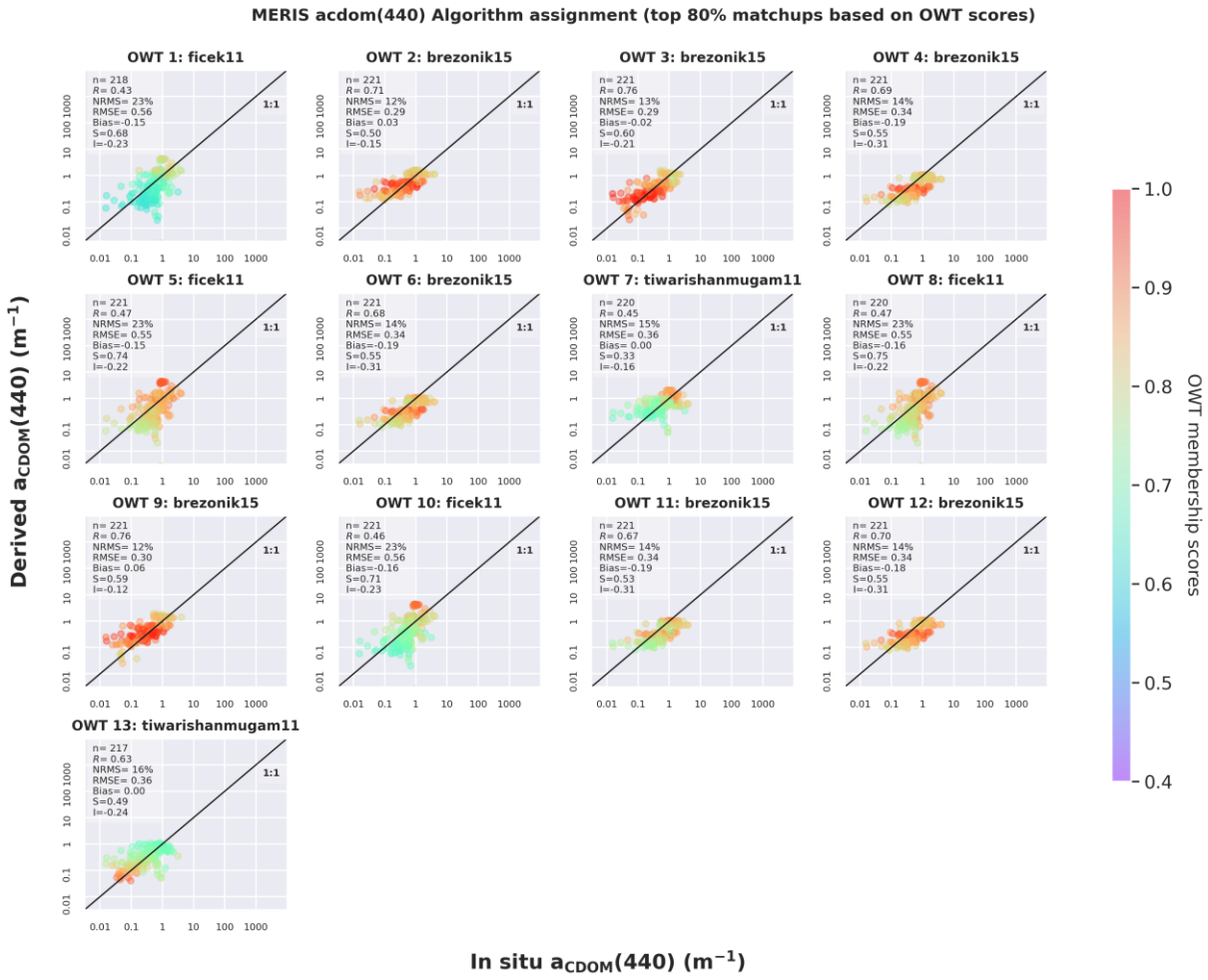


Figure 52 $a_{CDOM}(440)$ algorithm assiment of each OWT for MERIS.

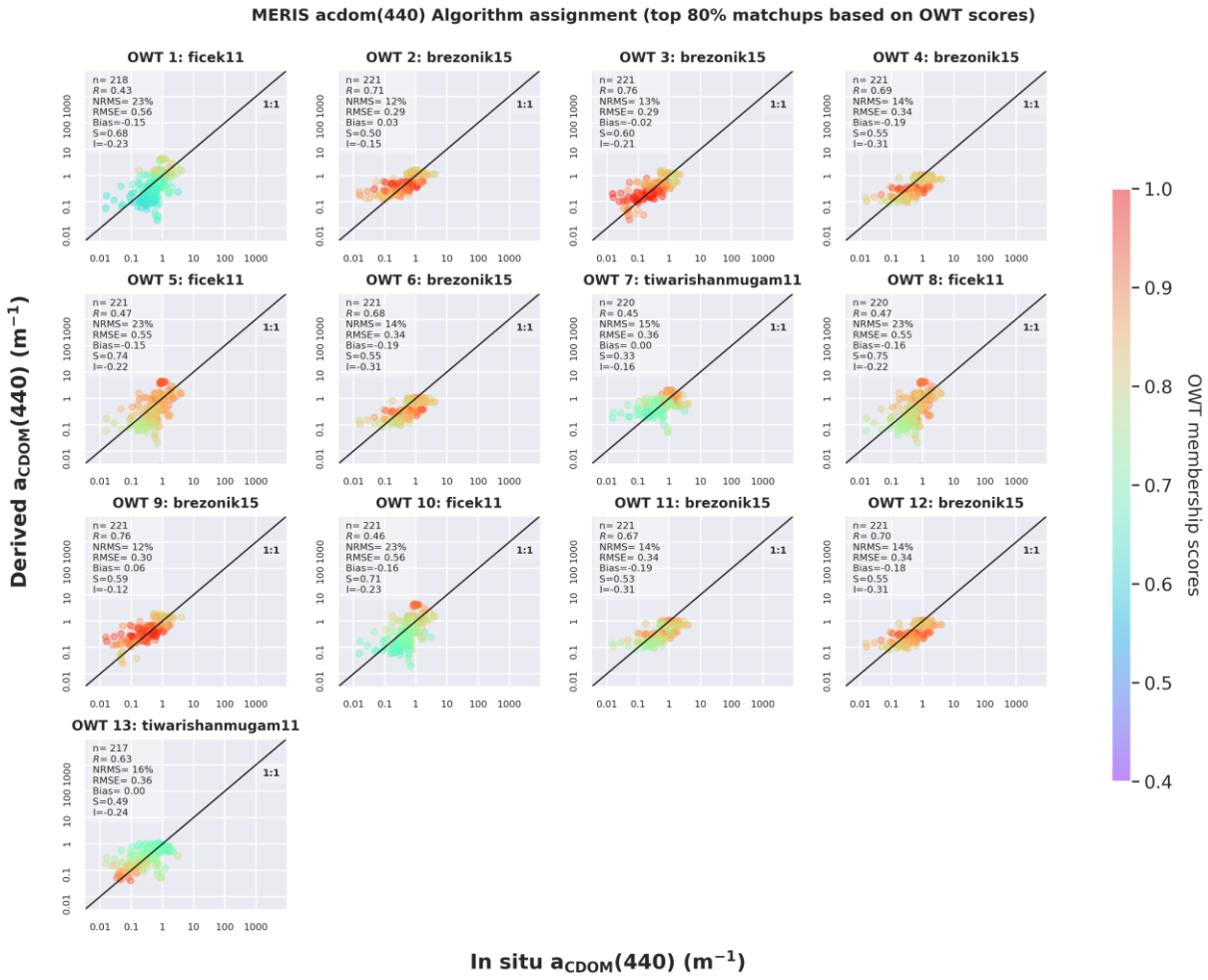


Figure 53 $a_{CDOM}(440)$ algorithm assiment of each OWT for OLCI.

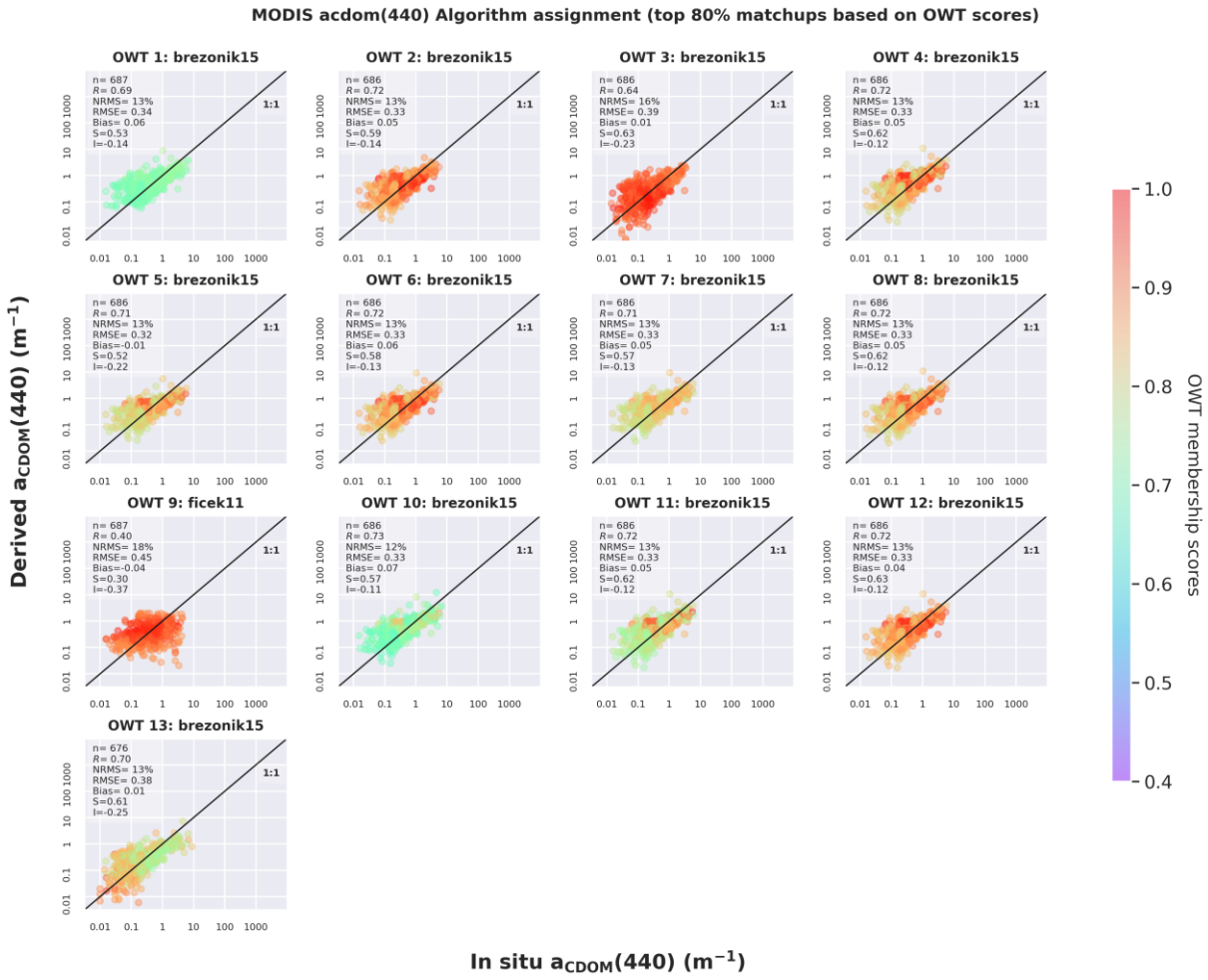


Figure 54 $a_{CDOM}(440)$ algorithm assiment of each OWT for MODIS.

Figure 55 presents the comparison between in situ measurements and the weighted blending of aCDOM(440) from the top three OWTs for MERIS, OLCI, and MODIS. The regression line closely aligns with the 1:1 line, with R-values of 0.73, 0.71, and 0.77 for MERIS, OLCI, and MODIS, respectively, all measured in log space. The average NRMS, RMSE, and Bias for MERIS, OLCI, and MODIS are 13.4%, 0.32, and 0.02; 15.1%, 0.42, and -0.08; and 11.8%, 0.35, and 0.04, respectively.

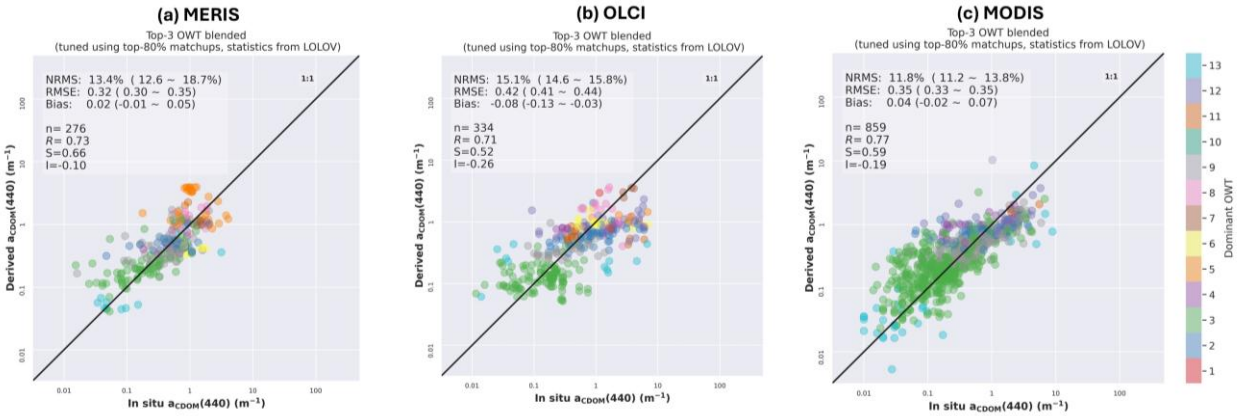


Figure 55 Comparison between in situ and weighted blending $a_{CDOM}(440)$ from the top-3 optical water types, for (a) MERIS, (b) OLCI and (c) MODIS.

5.2.5 K_d algorithms

Semi-analytically retrieval of K_d from R_{rs} is highly depending on the step of accurately retrieving absorption and backscattering coefficients (a, b_b) from R_{rs}. The Quasi-analytical algorithm (QAA) has been widely used for this purpose. Different types of QAA have been published for applications in waters with different bio-optical characteristics. In Lakes_cci, totally 16 published QAAs were selected and evaluated for MERIS and OLCI (Table 18), and nine of them were assessed for MODIS (indicated with * in Table 18) because some of the required bands (e.g., 709 nm, 779 nm) for QAAs (e.g., QAA_Mishra) are not available from MODIS.

MERIS, OLCI and MODIS satellite matchups containing satellite L_w extracted from CRDP v3.0 and in situ K_d were compiled. L_w were converted to R_{rs} by dividing by pi for QAA inputs. Candidate QAAs were then assessed using the matchup dataset, and the retrieved K_d were compared with in situ K_d for each OWT separately, to select the best-performing QAA for each OWT.

Table 18. Summary of candidate QAA algorithms evaluated for estimating K_d.

	Acronym	Algorithm name	Input band (nm)	Reference band (nm)	Reference
1*	v5	QAA_v5	443, 490, 560, 665	560	Lee et al., 2009
2*	v6	QAA_v6	443, 490, 560, 665	665	IOCCG, 2014
3*	P21	QAA_RGB	490, 560, 665	560	Pitarch et al., 2021
4	M14	QAA_Mishra	443,560,620,709	709	Mishra et al., 2014
5	C20	QAA_ Curtarelli	560, 665,704	560	Pedroso Curtarelli et al., 2020
6	L19	QAA_Liu	443, 490, 560, 665, 780	560	Liu et al., 2019
7	Y13	QAA_Yang	750, 780	750	Yang et al., 2013
8*	O16	QAA_Ogashawara	412, 443, 560, 665	560	Ogashawara, et al., 2016
9*	R17	QAA_Rodrigues	443, 490,560,665	560	Rodrigues, 2017
10	W16	QAA_Watanabe	443, 560, 620, 665, 709	709	Watanabe et al., 2016
11*	W17	QAA_Wang	490, 680	680	Wang et al., 2017
12*	C15	QAA_Chen	560, 665	560	Chen & Zhang, 2015
13*	S19	QAA_Shi	750, 865	750 or 865	Shi & Wang, 2019
14*	J18	QAA_Joshi	560, 671	560	Joshi & D'Sa, 2018
15	L09	QAA_Le	560, 709, 750	709	Le et al., 2009
16	G05	QAA_Gons	443, 560, 779	779	Gons et al., 2005

*Candidate algorithms for MODIS

5.2.5.1 K_d algorithms for MERIS and OLCI

Due to the lack of enough matchups for each OWT in the MERIS matchup dataset, algorithm assessments were carried out for MERIS and OLCI collectively because of their similarity in bands. Original coefficients and equations of QAAs were used in the algorithm evaluation, and algorithm performance in deriving K_d at 490 nm, 560 nm, and 665 nm were assessed.

The performances of QAAs were evaluated based on the top 20% of the membership score data of each OWT. Metrics used for quantifying errors include regression slope, R^2, RMSE, MAPE, Bias and Number of valid results, and these metrics were ranked from 1 (indicate better performance) to 16 (indicating poorer performance) between the results of all QAAs. The heatmaps of the performance of QAAs for estimating $K_d(490)$, $K_d(560)$, and $K_d(665)$ are shown in Figure 56, Figure 57, and Figure 58, respectively.

Figure 56. Comparison of QAAs in deriving $K_d(490)$ for MERIS and OLCI matchups.

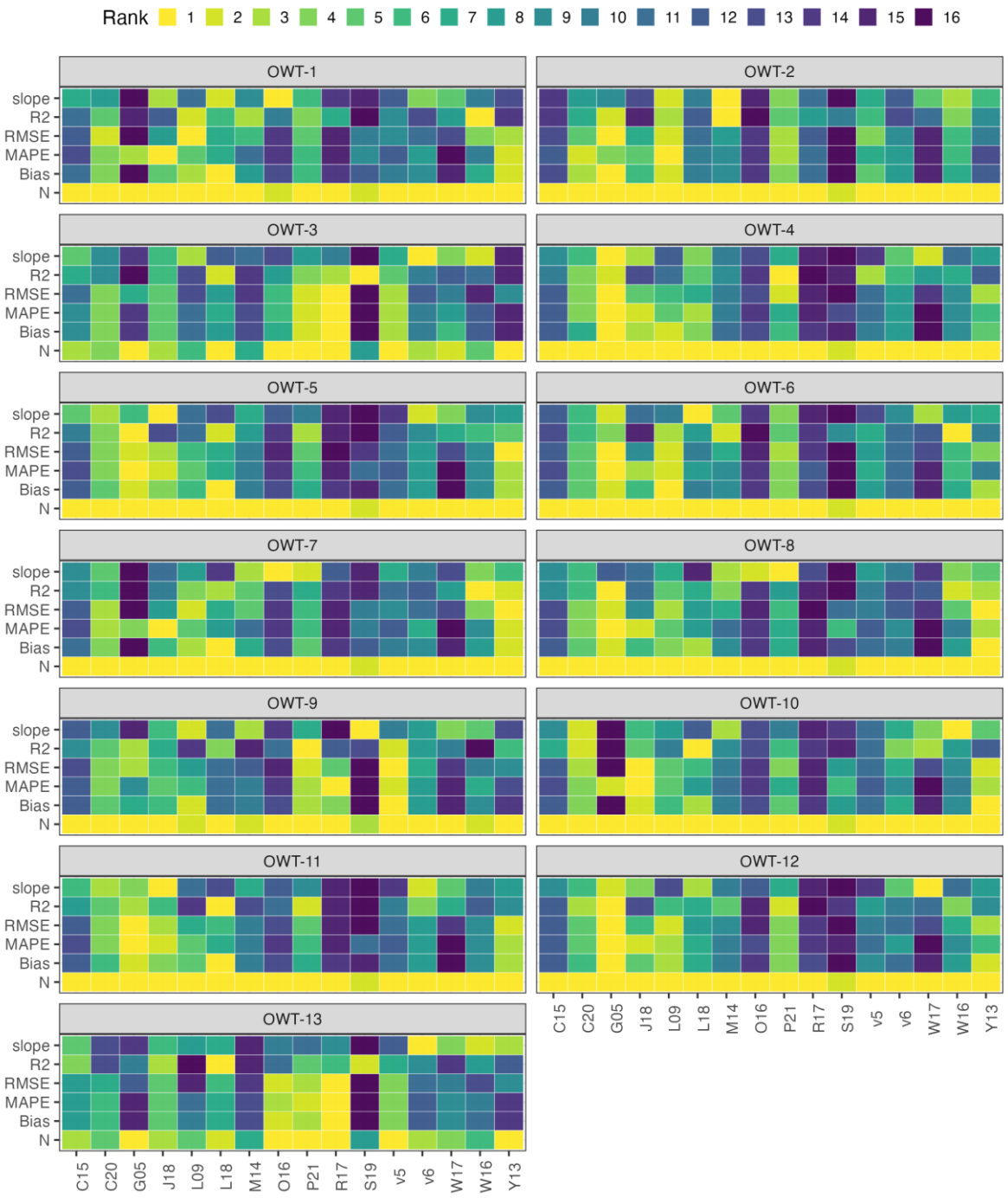


Figure 57. Comparison of QAAs in deriving $K_d(560)$ for MERIS and OLCI matchups.

Figure 58. Comparison of QAAs in deriving $K_d(665)$ for MERIS and OLCI matchups.

Visual inspections for the retrieval results were also carried out in addition to checking the error metrics. Results showed that the following algorithms performed better than others in:

- K_d(490) retrieval: Pedroso Curtarelli et al. (2020) for OWTs 1, 7; Gons et al. (2005) for OWTs 2, 5, 6, 8, 11, 12; Rodrigues et al. (2017) for OWTs 3, 13; Liu et al. (2019) for OWT 4; Pitarch et al. (2021) for OWT 9; Yang et al. (2013) for OWT 10;
- K_d(560) retrieval: Le et al. (2009) for OWTs 1, 2, 6; Rodrigues et al. (2017) for OWTs 3, 13; Gons et al. (2005) for OWTs 4, 11, 12; Yang et al. (2013) for OWTs 5, 7, 8, 10; QAA v5 for OWT 9;
- K_d(665) retrieval: Pitarch et al. (2021) for OWTs 1, 4, 5, 6, 7, 8, 10, 12; Le et al. (2009) for OWT 2; QAA v6 for OWTs 3, 13; Joshi et al. (2018) for OWTs 9, 11;

Scatterplots of the best-performing QAA in deriving K_d(490), K_d(560), and K_d(665) for each OWT are shown in Figure 59, Figure 60, and Figure 61, respectively.

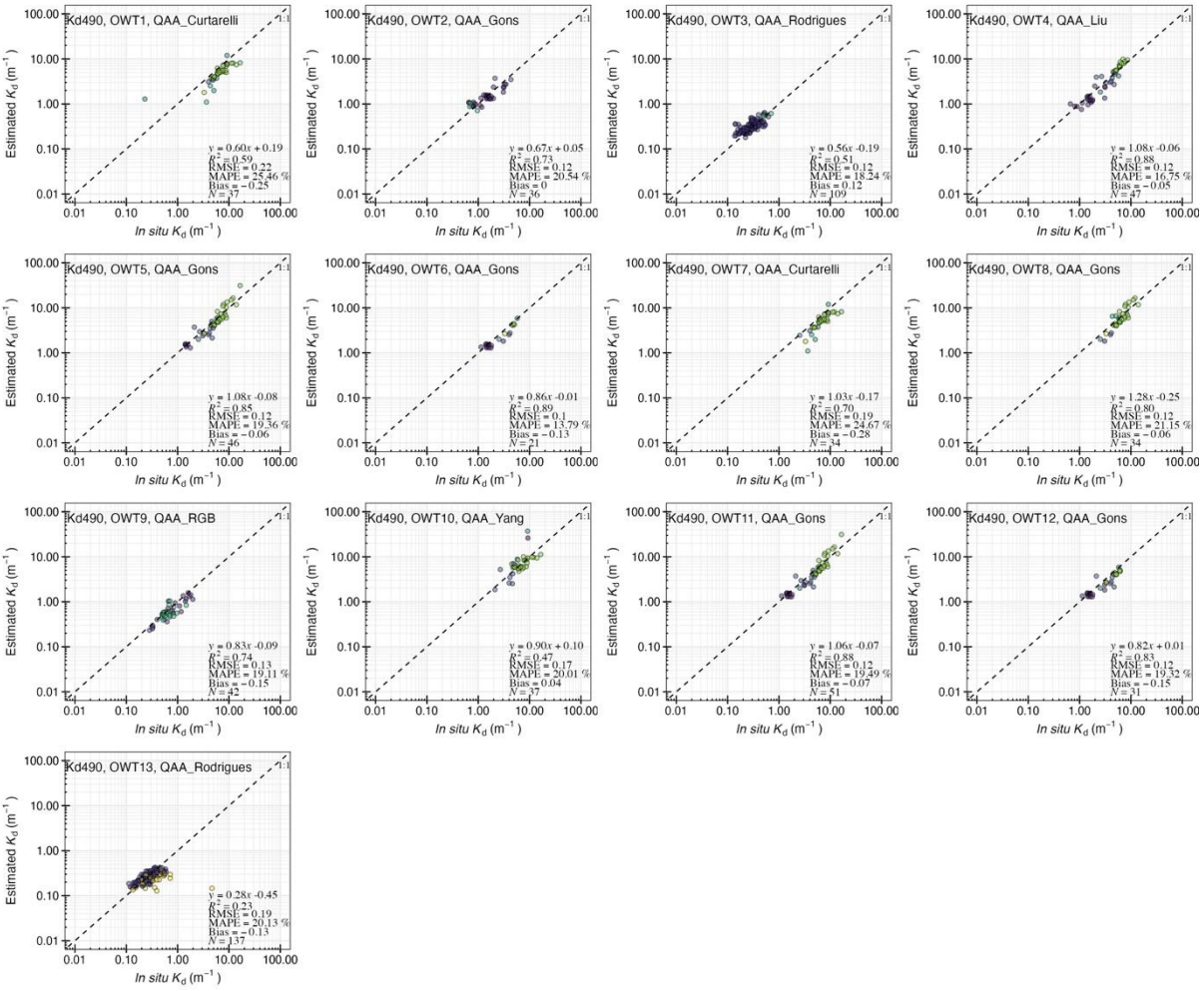


Figure 59. Best-performing QAA algorithm in deriving K_d(490) for each OWT. Data points of top 20% OWT membership score are shown for each OWT.

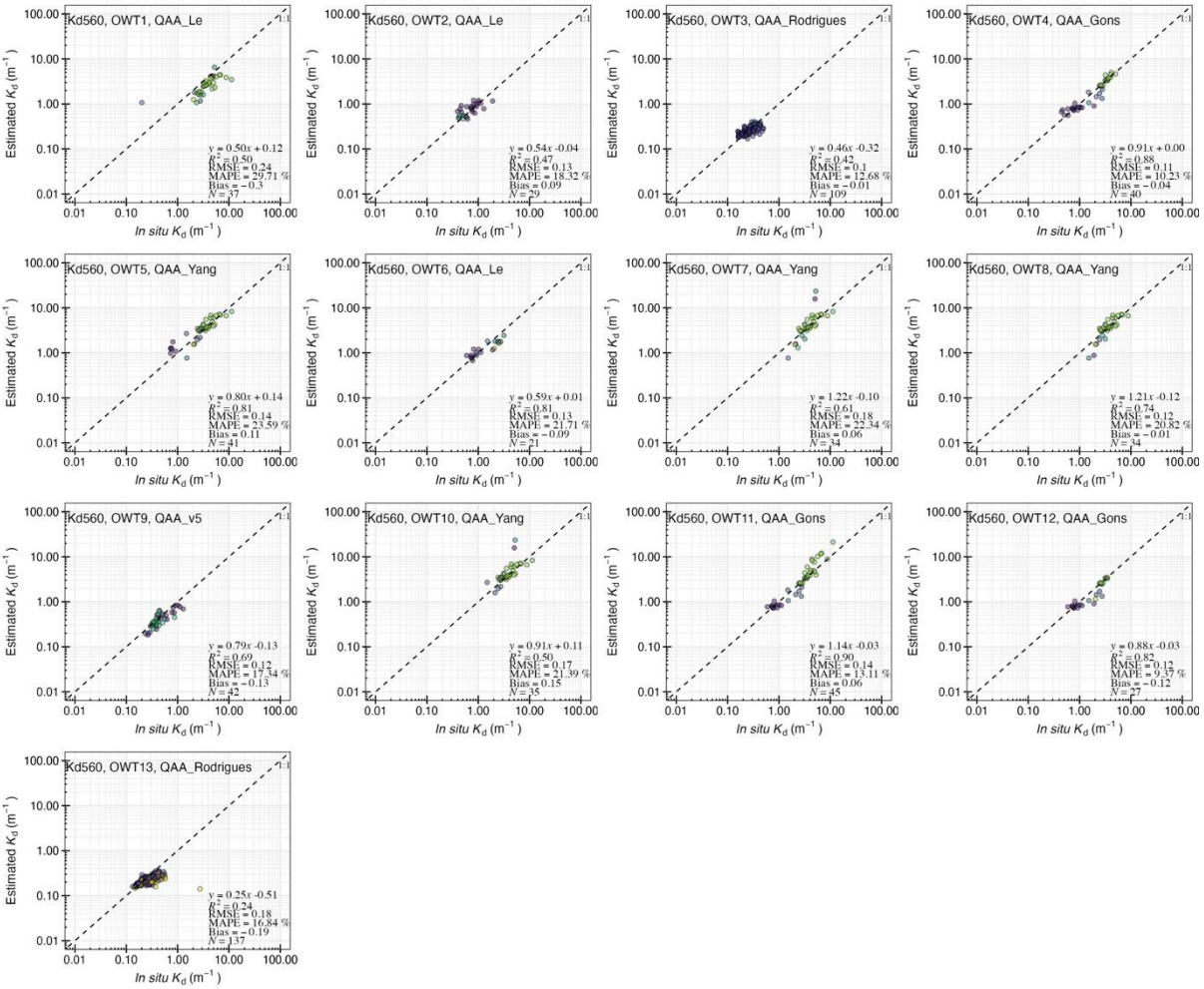


Figure 60 Best-performing QAA algorithm in deriving $K_d(560)$ for each OWT. Data points of top 20% OWT membership score are shown for each OWT.

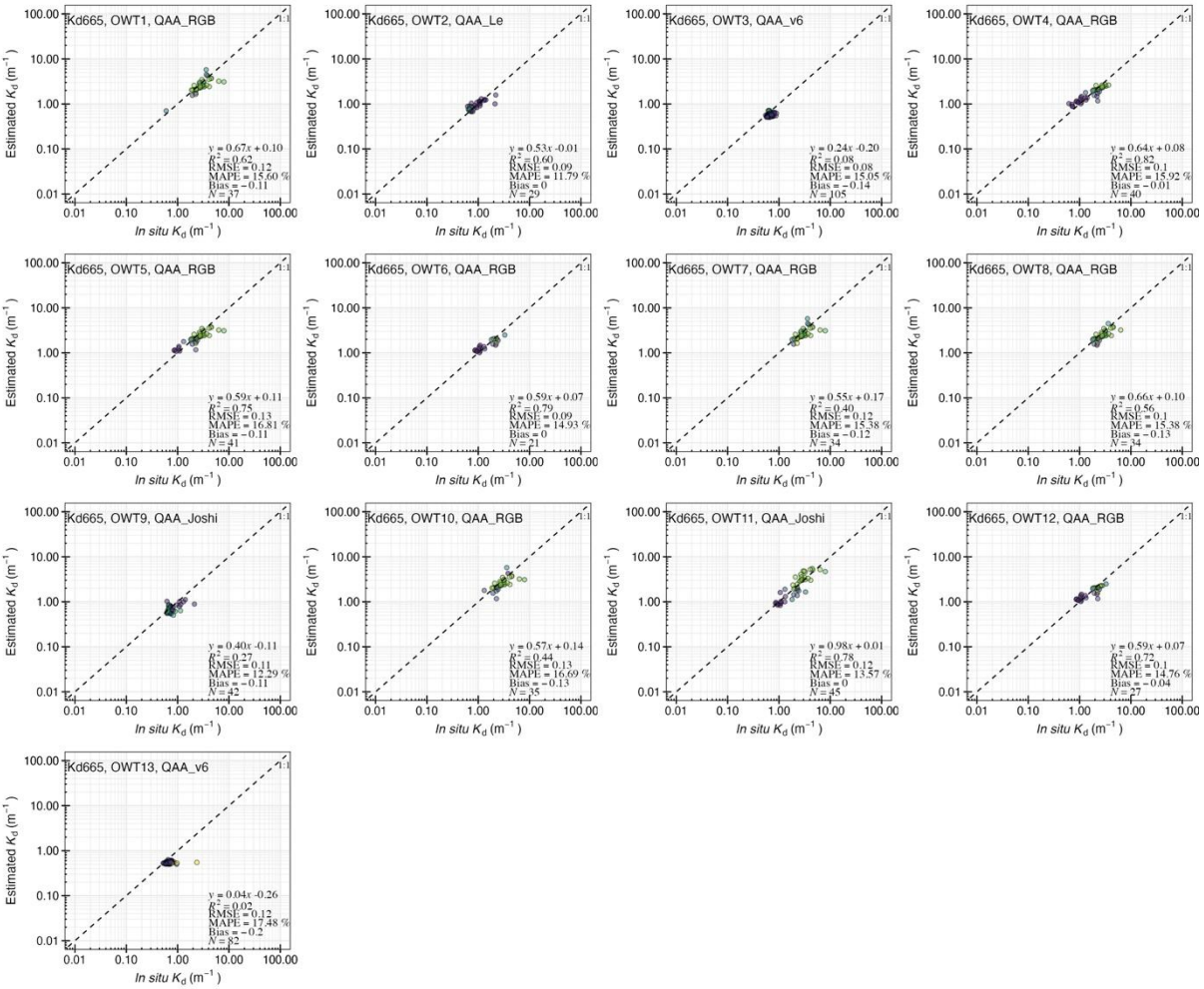


Figure 61. Best-performing QAA algorithm in deriving $K_d(665)$ for each OWT. Data points of top 20% OWT membership score are shown for each OWT.

The estimated K_d after blending the best-performing QAA algorithms are shown in Figure 62. Results showed that the blended algorithms can retrieve accurate $K_d(490)$, $K_d(560)$ and $K_d(665)$ from clear to turbid waters. A few outliers were observed in the K_d results estimated using the blended algorithms, but these are mainly associated with the quality of the reflectance after atmospheric correction. A test was carried out using the QWIP reflectance quality control method (Dierssen et al., 2022), applying a threshold of 0.2 (i.e., abs(QWIP) < 0.2 for good reflectance) can remove outliers in the results in Figure 62.

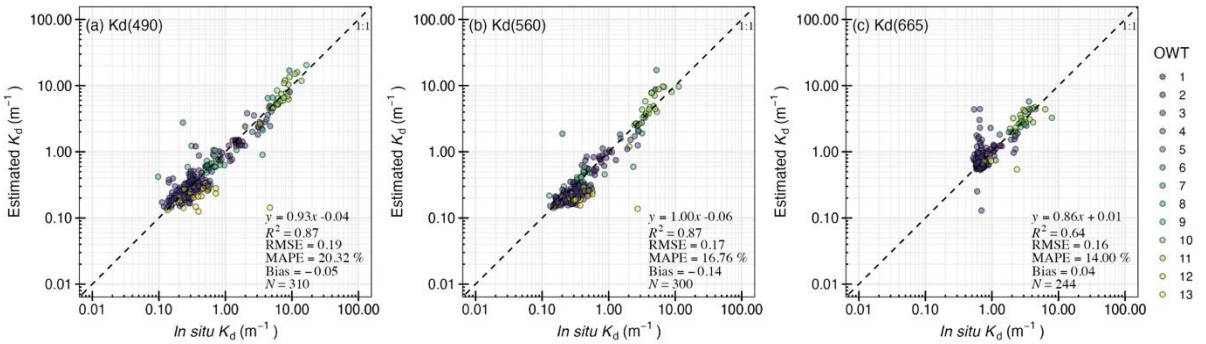


Figure 62. Estimated $K_d(490)$, $K_d(560)$ and $K_d(665)$ from MERIS and OLCI matchups using the blended algorithm.

5.2.5.2 K_d algorithms for MODIS

Nine QAAs (Table 18) were evaluated for estimating absorption and backscattering coefficients, and consequently for estimating K_d at 488 nm, 555 nm, and 667 nm from MODIS reflectance. The same approach was used for algorithm evaluation, comparison and selection as for MERIS and OLCI. The performance and comparison of different QAAs in deriving $K_d(488)$, $K_d(555)$, and $K_d(667)$ are shown in Figure 63, Figure 64, and Figure 65, respectively.

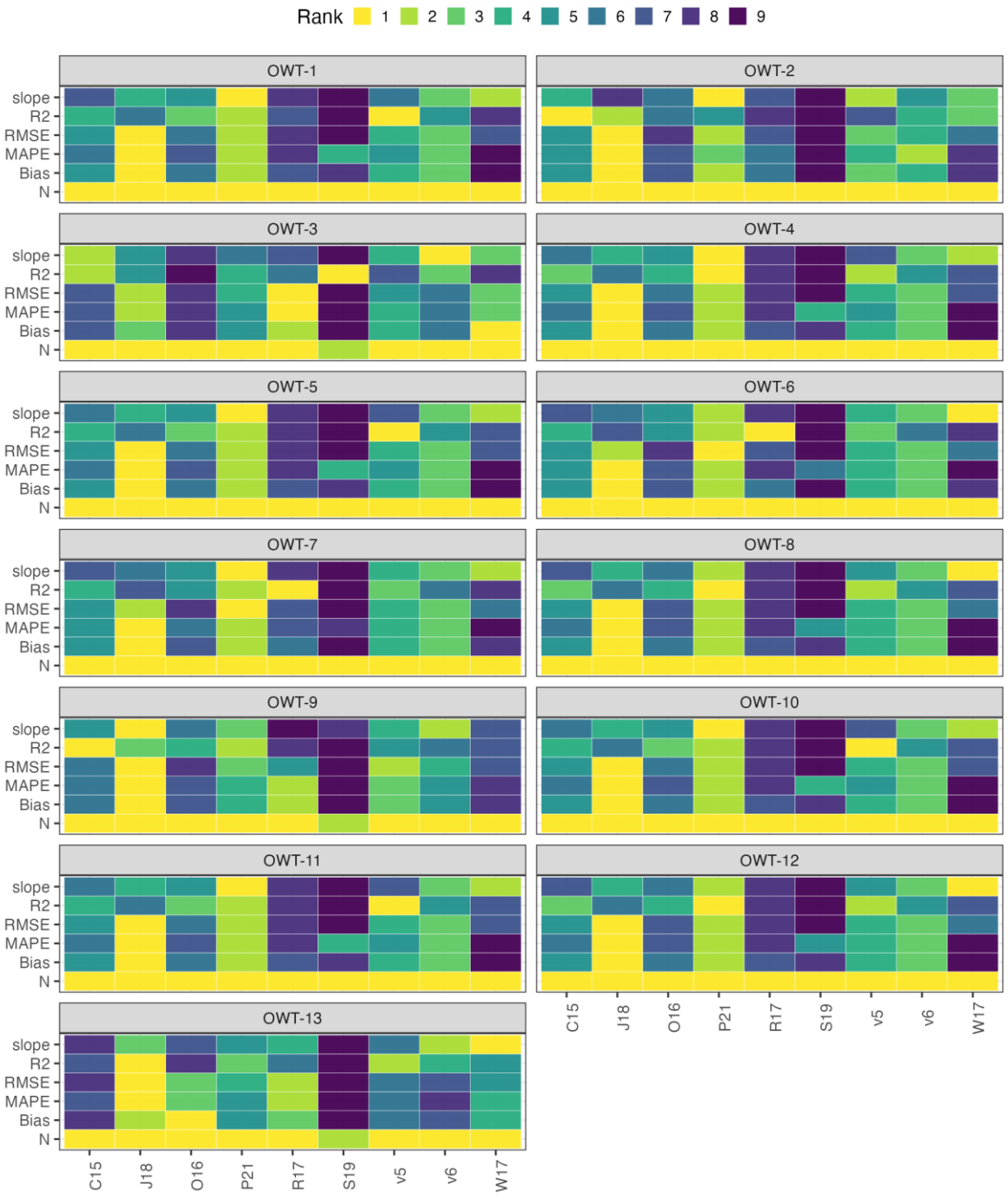


Figure 63. Comparison of different QAAs in deriving $K_d(488)$ from MODIS matchups.

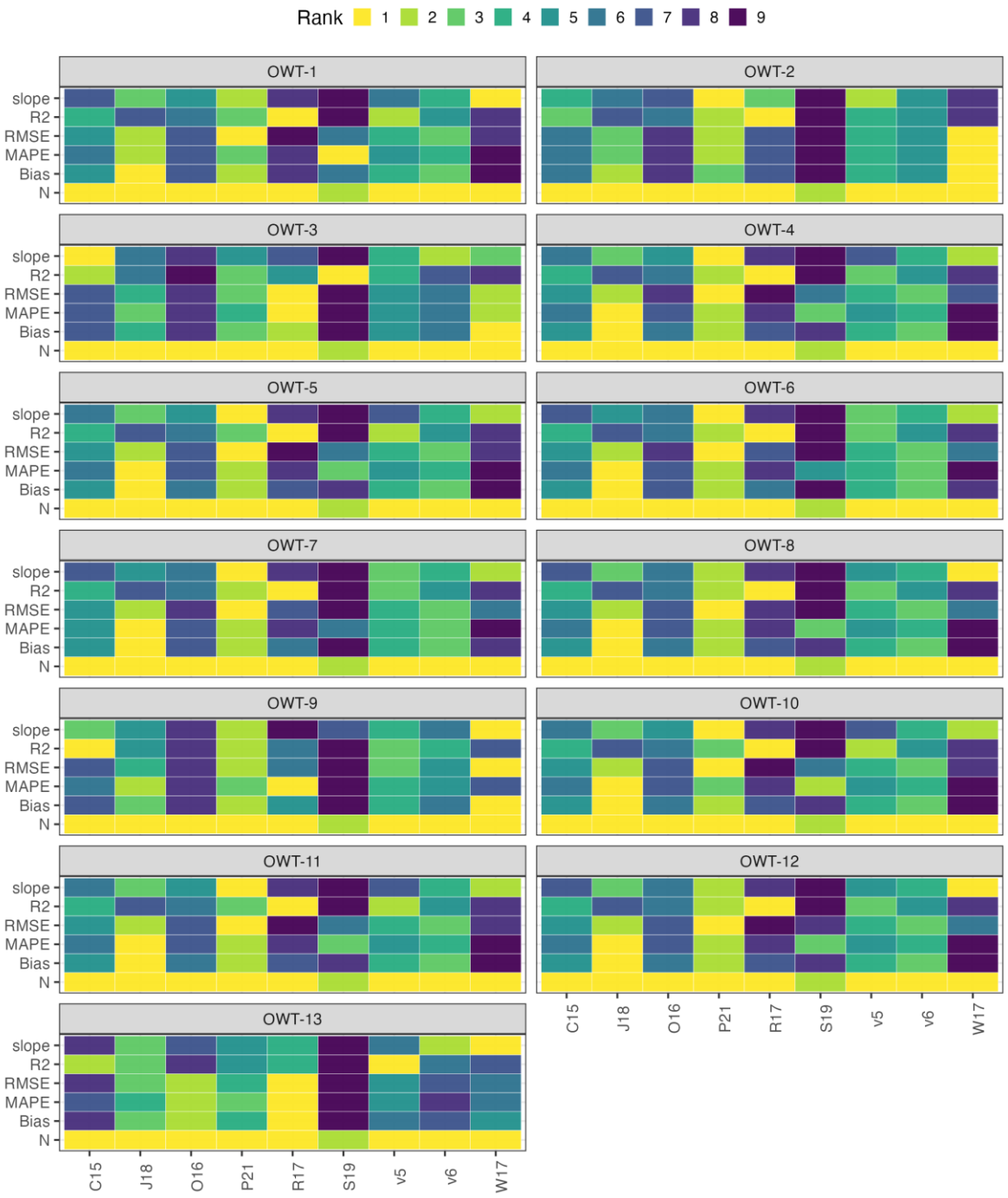


Figure 64. Comparison of different QAAs in deriving $K_d(555)$ from MODIS matchups.

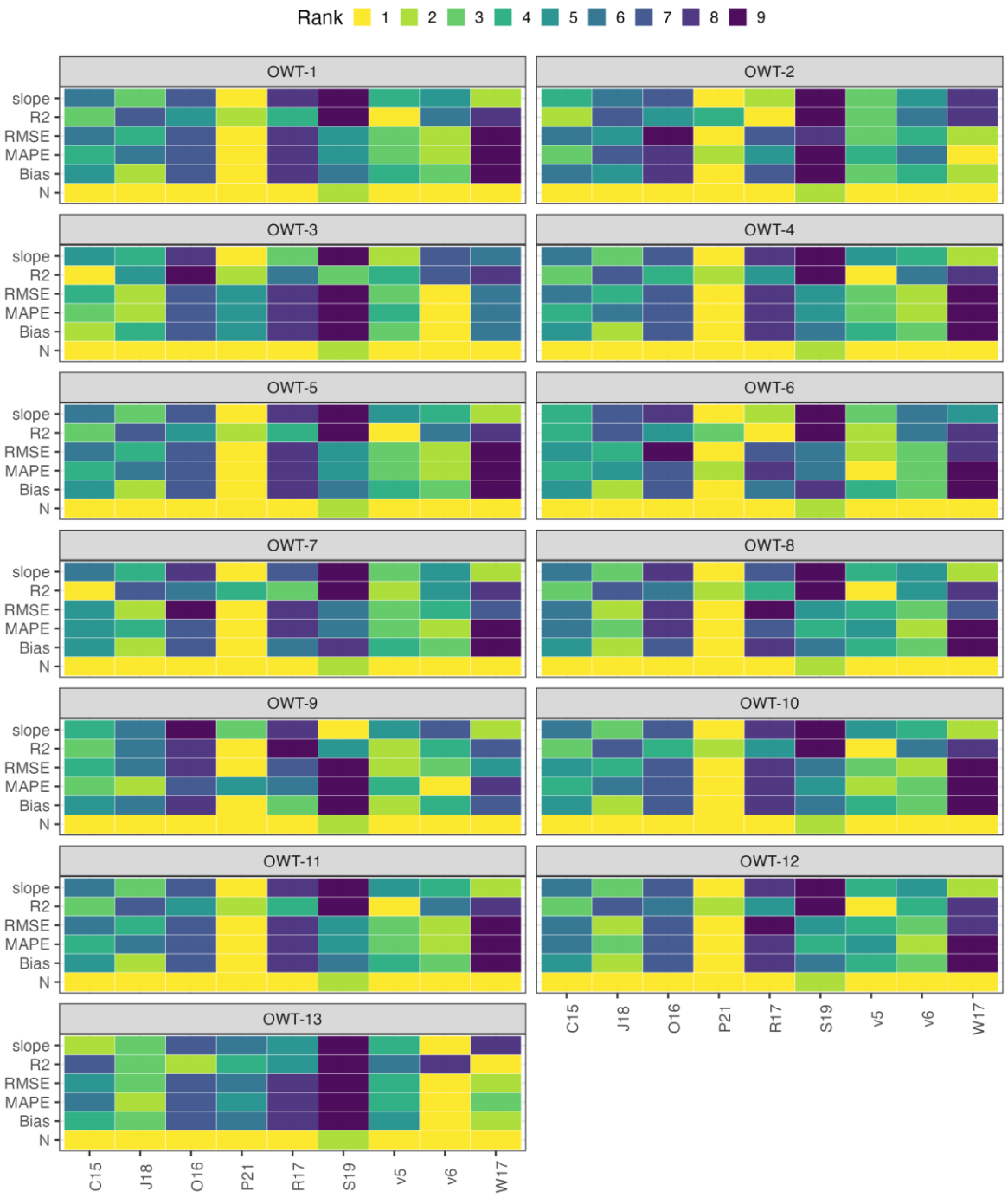


Figure 65. Comparison of different QAAs in deriving $K_d(667)$ from MODIS matchups.

By checking the error metrics and visually inspecting the scatterplots of the estimated K_d per OWT, we found that the best-performing QAAs are:

- $K_d(488)$: Joshi et al. (2018) for OWTs 1, 2, 4, 5, 6, 7, 8, 9, 10, 11; Rodrigues et al. (2017) for OWTs 3, 13; Pitarch et al. (2021) for OWT 12;
- $K_d(555)$: Joshi et al. (2018) for OWTs 1, 2, 4, 5, 6, 7, 8, 10, 11; Rodrigues et al. (2017) for OWTs 3, 13; Wang et al. (2017) for OWT 9; Pitarch et al. (2021) for OWT 12;
- $K_d(667)$: Pitarch et al. (2021) for OWTs 1, 2, 4, 5, 6, 7, 8, 10, 11, 12; QAA v6 for OWTs 3, 13; Wang et al (2017) for OWT 9.

The results of estimated K_d using the best-performing QAA for each OWT are shown in Figure 66, Figure 67, and Figure 68.

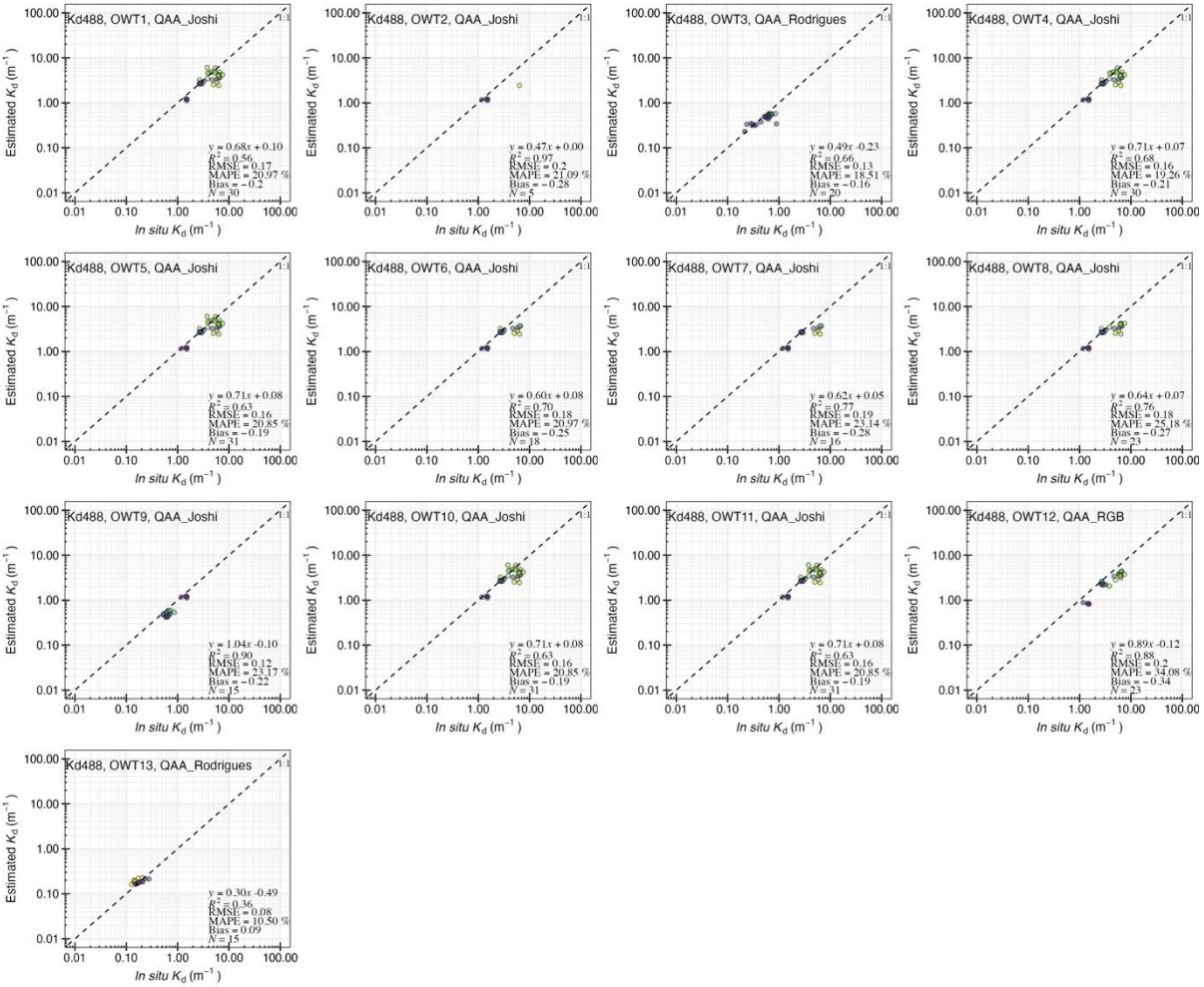


Figure 66. Best-performing QAA algorithm in deriving $K_d(488)$ for each OWT. Data points of top 20% membership score are shown for each OWT.

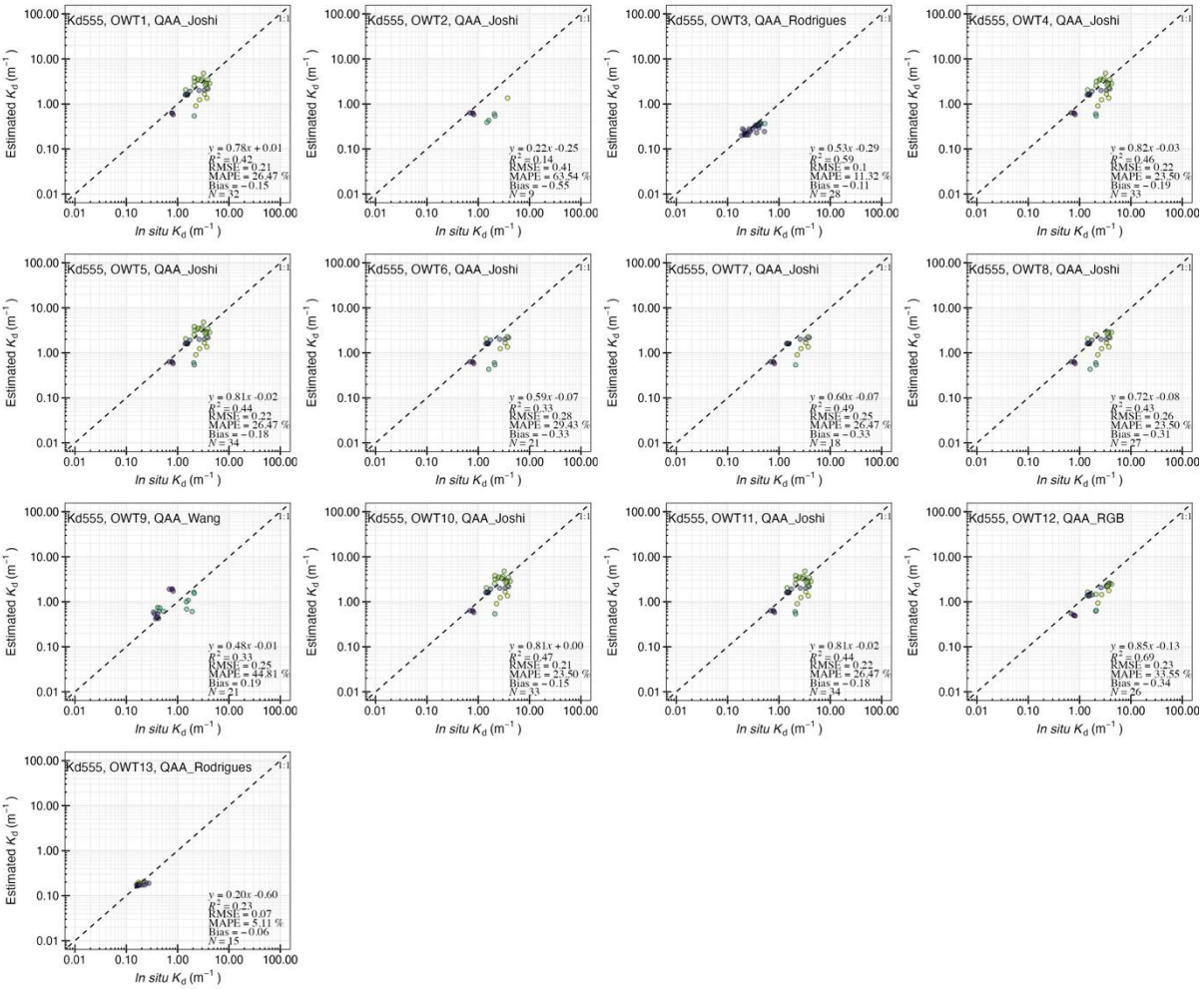


Figure 67. Best-performing QAA algorithm in deriving $K_d(555)$ for each OWT. Data points of top 20% OWT membership score are shown for each OWT.

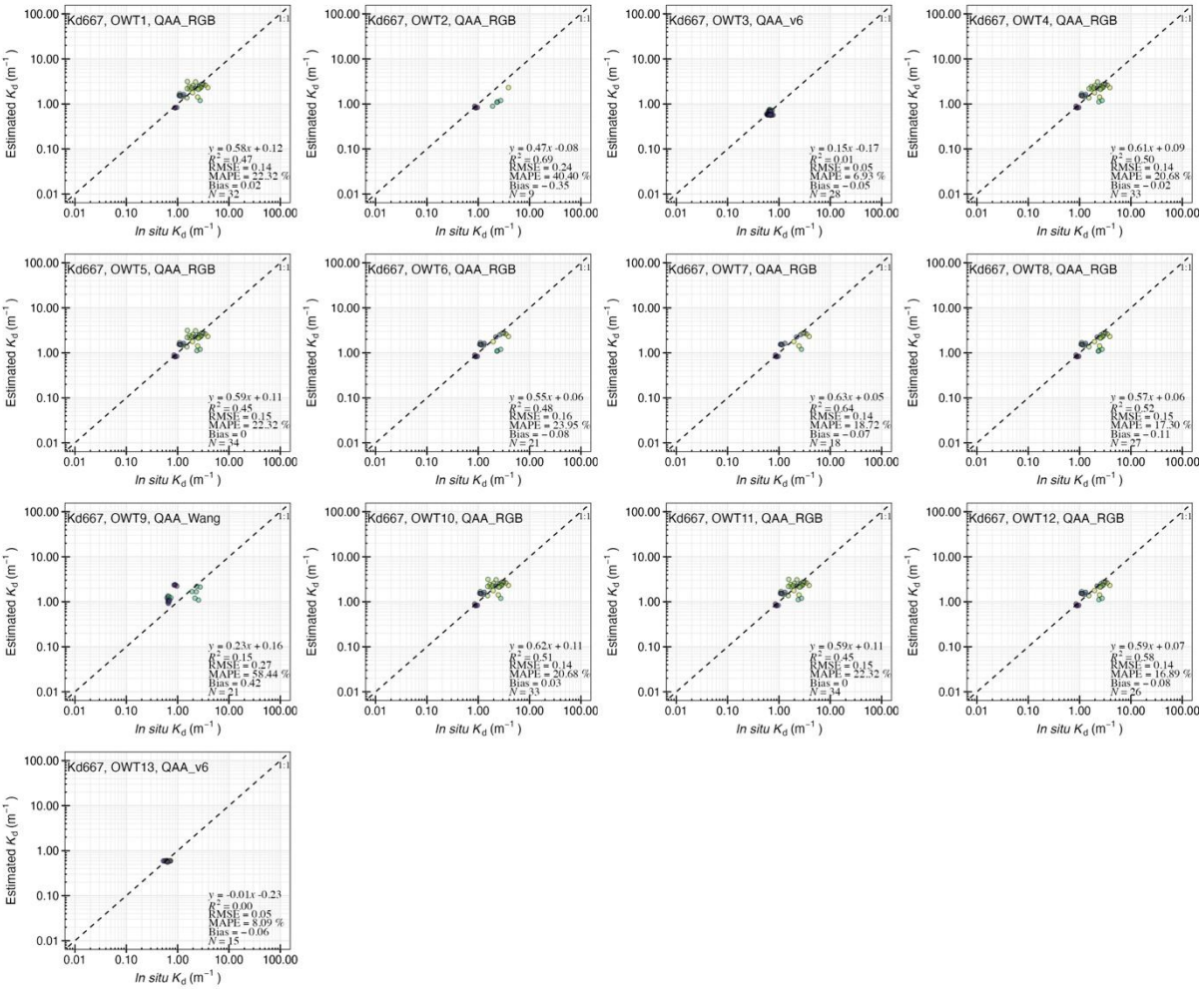


Figure 68. Best-performing QAA algorithm in deriving $K_d(667)$ for each OWT. Data points of top 20% OWT membership score are shown for each OWT.

The results of estimated K_d after blending the best-performing QAA algorithms are shown in Figure 69. Results showed that the blended algorithms can overall retrieve accurate $K_d(488)$, $K_d(555)$ and $K_d(667)$ from clear to turbid waters, some slight underestimations for $K_d(488)$ and $K_d(555)$ were observed. In addition, the number of matchups from MODIS is less than MERIS and OLCI. In the future versions, further validations with more MODIS matchups and development of new algorithms can potentially improve the K_d estimates from MODIS.

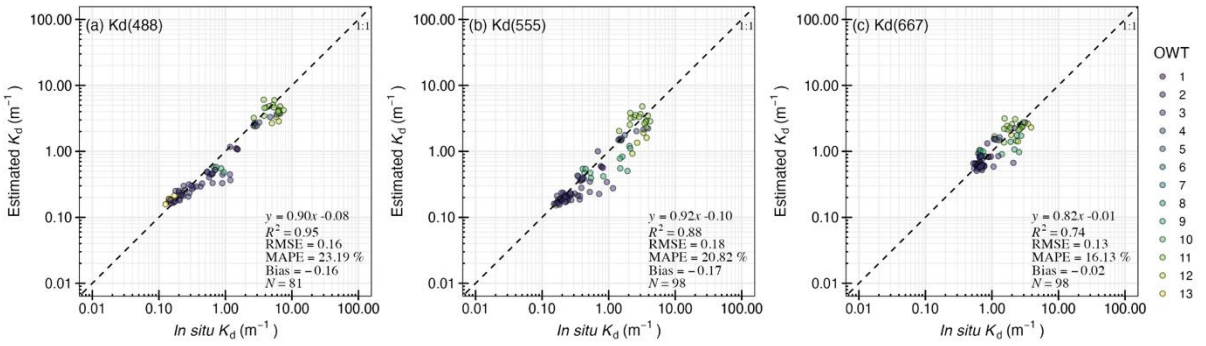


Figure 69. Estimated $K_d(488)$, $K_d(555)$ and $K_d(667)$ from MODIS matchups using the blended algorithm.

The retrieved K_d at blue, green and red bands using the blended algorithms were input to estimate K_d(PAR) for MERIS, OLCI, and MODIS matchups, and then compared to in situ K_d(PAR). It should be noted that the in situ K_d(PAR) were modelled from in situ K_d at blue, green and red bands, to maximise the number of available K_d(PAR) data and evaluate the full processing steps from R_{rs} to K_d and K_d(PAR). As shown in Figure 70, K_d(PAR) can be accurately estimated from MERIS, OLCI, and MODIS R_{rs}.

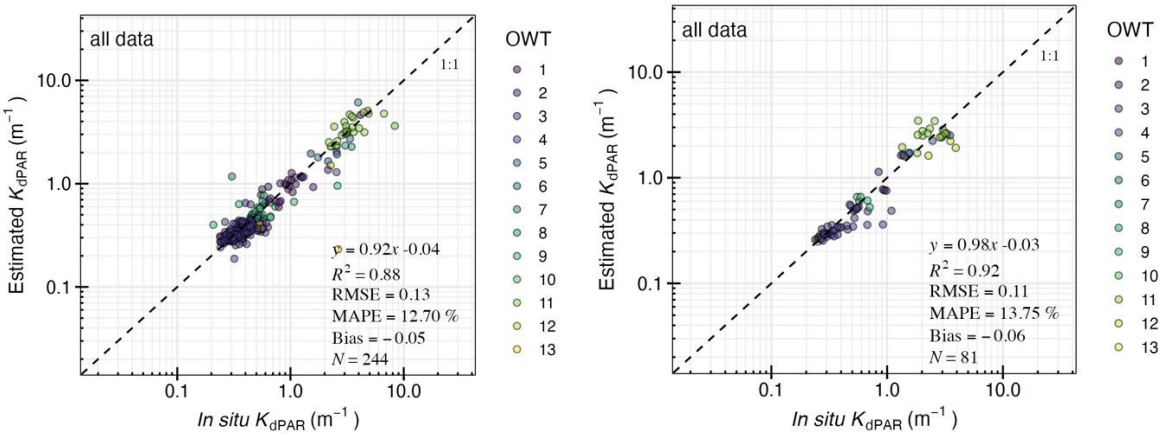


Figure 70. Estimated K_d(PAR) from (left) MERIS and OLCI matchups, and (right) MODIS matchups.

5.3 Identified issues

Prior to CRDP v3.0.0, no specific validation or tuning of individual algorithms was carried out for OLCI which instead inherited algorithms from tuning and validation for MERIS. Similarities in radiometric performance and waveband configuration between MERIS and OLCI allowed us to extrapolate the MERIS results to OLCI. Nevertheless, given the continued operation of OLCI and the emergence of new algorithms in recent years, a comprehensive analysis specifically tailored to OLCI was necessary. To address this, a thorough evaluation of newly published algorithms has now been conducted for all three sensors and incorporated into CRDP v3.0.0.

The capabilities of MERIS and OLCI, although similar, differ in several aspects, one of these being the SWIR channel of OLCI which is not present with MERIS. Overall, improved stability in OLCI and MERIS is seen in highly turbid conditions (see PVIR for full results) but some discontinuities remain observed with MERIS. Fortunately, these exceptions tend to cause extremely high LWLR, and therefore identifiable with new quality masking procedures.

It is important to acknowledge that atmospheric correction poses a significant challenge in remote estimations of water quality variables. An analysis on Chla (Pahlevan et al. 2021) suggests that atmospheric correction can lead to a loss of performance of at least 30%. Therefore, when directly applying these algorithms to satellite observations rather than in situ radiometric observations, their performance should be expected to decrease. To address this issue, it is beneficial to conduct tuning using satellite-derived reflectance. This additional tuning step aims to eliminate systematic biases introduced by the atmospheric correction procedure. By incorporating satellite-derived reflectance, the algorithms can better account for and mitigate the impact of atmospheric effects, enhancing their performance for water quality parameter estimations. This process has now been completed across all sensors and variables for CRDP v3.0.0.

An issue affecting MODIS data accuracy is band saturation in highly turbid waters. Certain detectors within the MODIS instrument, especially Red and NIR bands, have been observed to reach saturation,

particularly in highly turbid waters. This saturation results in invalid reflectance products after atmospheric-correction, which are then masked out. The saturation of several MODIS wavebands in highly turbid waters presents a specific challenge to present consistent long-term time-series without significant bias to more favourable conditions. Therefore, a cross-sensor validation between MERIS, OLCI and MODIS based on over-lapping observations is necessary, to identify lakes which are less likely to be influenced by this inherient issue by MODIS when applied to inland waters.

The following potential issues have been identified for the candidate algorithms tested for OLCI, MERIS, and MODIS:

- The MDN algorithm is based on machine learning, which relies on the dataset used for training. Therefore, the performance of this algorithm is heavily dependent on the quality and representativeness of the training data. Further tuning may not be applicable for this algorithm, and validation will not be independent from the training data set.
- Similarly, the Bayesian probabilistic neural networks (BNN) method is a machine learning-based data-driven approach. Its performance is highly influenced by the dataset used for model training. The original training range of this model is designed for oligotrophic and mesotrophic lakes, and its applicability to eutrophic and hyper-eutrophic lakes is yet to be tested. Further tuning for this algorithm might not be applicable either.
- The Smith18 algorithm is a switch-blending algorithm that combines the G2B and OCI algorithms. Its performance, especially when integrated into our weighted blending procedure, still needs to be evaluated. Tuning of this algorithm should be conducted separately for the two underlying algorithms (Table 17)
- The SOLID20 algorithm uses a classification-based approach with an MDN method for b_{bp} (particulate backscatter) retrieval. As it is based on machine learning, further tuning may not be applicable to this algorithm.
- The Novoa21G and NovoaB algorithms are both switch-blending algorithms. It is important to verify whether the outputs of these algorithms exhibit a seamless transition between the different blending modes. Additionally, each algorithm incorporated in this switch-blending approach should be individually tuned to optimize their performance.

5.4 Acknowledgments

All in situ datasets used in this research are obtained from two primary sources: the **LIMNADES** (Lake Bio-optical Measurements and Matchup Data for Remote Sensing) repository, maintained by the University of Stirling (https://limnades.stir.ac.uk, not currently accessible), and the **GLORIA** (GLObal Reflectance community dataset for Imaging and Optical Sensing of Aquatic Environments), which is openly available online (Lehmann et al, 2023). We also thank all organizations and individuals who made in situ data available through LIMNADES for *Calimnos* validation under the GloboLakes project and separate agreements: Agri-food and Biosciences Institute in Northern Ireland; General Directorate of Water Management (OVF, Hungary); Central Transdanubian Inspectorate for Environmental and Natural Protection (Közép-dunántúli Környezetvédelmi és Természetvédelmi Felügyelőség); Balaton Limnological Institute; Finnish Environment Institute; Mariano Bresciani and Claudia Giardino (Institute for Electromagnetic Sensing of the Environment, CNR-IREA, Milano, Italy); Bunkei Matsushita (Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan); Mark W. Matthews (CyanoLakes (Pty) Ltd., Cape Town, South Africa); Daniel Andrade Maciel, Victor Pedroso Curtarelli, Cláudio Clemente Faria Barbosa (National Institute for Space Research, Brazil); Daniela Gurlin (Wisconsin Department of Natural Resources, USA); Enner Alcântara, Thanan Pequeno (São Paulo State University, Brazil); Kersti Kangro, Krista Alikas (Tartu University, Estonia); Mortimer Werther, Daniel Odermatt (Eawag, Switzerland); Sachi Mishra, Deepak Mishra (University of Georgia, USA); Thi Thu Ha Nguyen (Vietnam National University, Vietnam).

5.5 References

Balasubramanian, S.V., Pahlevan, N., Smith, B., Binding, C., Schalles, J., Loisel, H., Gurlin, D., Greb, S., Alikas, K., & Randla, M. (2020). Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters. Remote Sens. Environ., 246: 111768

Binding, C., Bowers, D., Mitchelson-Jacob, E. (2005). Estimating suspended sediment concentrations from ocean colour measurements in moderately turbid waters; the impact of variable particle scattering properties. Rem Sens Env 94(3): 373-383.

Binding, C., Jerome, J., Bukata, R. and Booty, W. (2010). Suspended particulate matter in Lake Erie derived from MODIS aquatic colour imagery. Int J Remote Sens 31(19).

Boss, E. and Roesler C. (2006). Over Constrained Linear Matrix Inversion with Statistical Selection. In: Lee, ZP (Ed.), IOCCG Report nr 5. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications., pp. 126.

Carder KL, Chen FR, Cannizzaro JP, et al (2004) Performance of the MODIS semi-analytical ocean color algorithm for chlorophyll-a. Adv Space Res 33:1152–1159.

Chen S, Han L, Chen X, et al (2015) Estimating wide range Total Suspended Solids concentrations from MODIS 250-m imageries: An improved method. ISPRS J Photogramm Remote Sens 99:58–69.

Chen, S., & Zhang, T. (2015). Evaluation of a QAA-based algorithm using MODIS land bands data for retrieval of IOPs in the Eastern China Seas. Optics express, 23(11), 13953-13971.

Chen Z, Hu C, Muller-Karger F (2007) Monitoring turbidity in Tampa Bay using MODIS/Aqua 250-m imagery. Remote Sens Environ 109:207–220.

Constantin S, Constantinescu Ştefan, Doxaran D (2017) Long-term analysis of turbidity patterns in Danube Delta coastal area based on MODIS satellite data. J Marine Syst 170:10–21.

Dall'Olmo G, Gitelson AA, Rundquist DC, et al (2005) Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. Remote Sens Environ 96:176–187.

Dekker, A. Vos, R., Peters, S. (2002). Analytical algorithms for lake water TSM estimation for retrospective analyses of TM and SPOT sensor data, Int J Rem Sens 23(1):15-35.

Dierssen, H. M., Vandermeulen, R. A., Barnes, B. B., Castagna, A., Knaeps, E., & Vanhellemont, Q. (2022). QWIP: A quantitative metric for quality control of aquatic reflectance spectral shape using the apparent visible wavelength. Frontiers in Remote Sensing, 3, 869611.Dogliotti, A.I., Ruddick, K., Nechad, B., Doxaran, D., & Knaeps, E. (2015). A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarine waters. Remote Sens. Environ., 156: 157-168

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

D'Sa, E. , Miller, R., McKee, B. (2007). Suspended particulate matter dynamics in coastal waters from ocean color: Application to the northern Gulf of Mexico, Geophys Res Lett 34, L23611.

El-Alem A, Chokmani K, Laurion I, El-Adlouni SE (2012) Comparative Analysis of Four Models to Estimate Chlorophyll-a Concentration in Case-2 Waters Using MODerate Resolution Imaging Spectroradiometer (MODIS) Imagery. Remote Sens 4:2373–2400.

Eleveld, M., Pasterkamp, R., Van der Woerd, H., Pietrzak, J. (2008). Remotely sensed seasonality in the spatial distribution of sea-surface suspended particulate matter in the southern North Sea. Est Coast Shelf Sci 80(10): 103-113.

Gilerson, A., Gitelson, A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I. and Ahmed, S. (2010). Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Opt Express 18(23): 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. Int J Remote Sens 13:3367–3373

Gitelson AA, Schalles JF, Hladik CM (2007) Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sens Environ 109:464–472.

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

Gitelson, A., Gurlin, D., Moses, W., Yacobi, Y. (2011). Remote estimation of Chlorophyll-a concentration in inland, estuarine, and coastal waters. In: Weng, Q. (Ed.), Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications. CRC Press, Taylor and Francis Group, pp. 449–478 (610 p. ISBN: 9781420091755).

Gons H., Rijkeboer M., Ruddick K. (2005). Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. J Plankton Res 27(1):125-7.

Guanter L, Del Carmen González-Sanpedro M, Moreno J (2007) A method for the atmospheric correction of ENVISAT/MERIS data over land targets. Int J Remote Sens 28:709–728

Gurlin, D., Gitelson, A., Moses, W. (2011). Remote estimation of chl-a concentration in turbid productive waters — return to a simple two-band NIR-red model? Remote Sens Environ 115 (12), 3479–3490.

Ha N, Koike K, Nhuan M (2013) Improved Accuracy of Chlorophyll-a Concentration Estimates from MODIS Imagery Using a Two-Band Ratio Algorithm and Geostatistics: As Applied to the Monitoring of Eutrophication Processes over Tien Yen Bay (Northern Vietnam). Remote Sens 6:421–442.

Ioannou, I., Gilerson, A., Gross, B., Moshary, F., Ahmed, S. (2013). Deriving ocean color products using neural networks. Remote Sens Environ 134, 78e91.

IOCCG, 2014. Update of the Quasi-Analytical Algorithm (QAA_v6). Available online. http://www.ioccg.org/groups/Software_OCA/QAA_v6_2014209.pdf.

Jiang, D., Matsushita, B., Pahlevan, N., Gurlin, D., Lehmann, M.K., Fichot, C.G., Schalles, J., Loisel, H., Binding, C., & Zhang, Y. (2021). Remotely estimating total suspended solids concentration in clear to extremely turbid waters using a novel semi-analytical method. Remote Sens. Environ., 258: 112386

Joshi, I. D., & D'Sa, E. J. (2018). An estuarine-tuned quasi-analytical algorithm (QAA-V): assessment and application to satellite estimates of SPM in Galveston Bay following Hurricane Harvey. Biogeosciences, 15(13), 4065-4086.

Klein, K.P., Lantuit, H., Heim, B., Doxaran, D., Juhls, B., Nitze, I., Walch, D., Poste, A., & Søreide, J.E. (2021). The Arctic Nearshore Turbidity Algorithm (ANTA)-A multi sensor turbidity algorithm for Arctic nearshore environments. Science of Remote Sensing, 4: 100036

Le, C. F., Li, Y. M., Zha, Y., Sun, D., & Yin, B. (2009). Validation of a quasi-analytical algorithm for highly turbid eutrophic water of Meiliang Bay in Taihu Lake, China. IEEE Transactions on Geoscience and Remote Sensing, 47(8), 2492-2500.

Lee Z, Carder KL, Arnone RA (2002) Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. Appl Opt 41:5755–5772

Lee, Z., Lubac, B., Werdell, J., & Arnone, R. (2009). An update of the quasi-analytical algorithm (QAA_v5). International Ocean Color Group Software Report, 1, 1-9.

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 (2023). https://doi.org/10.1038/s41597-023-01973-y

Letelier R (1996) An analysis of chlorophyll fluorescence algorithms for the moderate resolution imaging spectrometer (MODIS). Remote Sens Environ 58:215–223.

Liu, G., Li, L., Song, K., Li, Y., Lyu, H., Wen, Z., Fang, C., Bi, S., Sun, X., & Wang, Z. (2020). An OLCI-based algorithm for semi-empirically partitioning absorption coefficient and estimating chlorophyll a concentration in various turbid case-2 waters. Remote Sens. Environ., 239: 111648

Liu, X., et al. (2024). "Quantifying decadal stability of lake reflectance and chlorophyll-a from medium-resolution ocean color sensors." Remote Sens. Environ., 306: 114120.

Liu, X., Lee, Z., Zhang, Y., Lin, J., Shi, K., Zhou, Y., ... & Sun, Z. (2019). Remote sensing of secchi depth in highly turbid lake waters and its application with MERIS data. Remote Sensing, 11(19), 2226.

Loisel, H., Mangin, A., Vantrepotte, V., Dessailly, D., Dinh, D. N., Garnesson, P. Ouillon, S., Lefebvre, J. P., Mériaux, X., Phan, T. M. (2014). Variability of suspended particulate matter concentration in coastal waters under the Mekong's influence from ocean color (MERIS) remote sensing over the last decade. Rem Sens Environ 150: 218-230.

Maritorena, S., Siegel, D., Peterson, A. (2002). Optimization of a semianalytical ocean color model for global-scale applications. Appl Opt 41, 2705–2714.

Matthews, M., Bernard, S., Robertson, L. (2012). An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. Remote Sens Environ 124, 637–652.
Miller RL, McKee BA (2004) Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. Remote Sens Environ 93:259–266.
Mishra S., Mishra D., Lee Z. (2014). Bio-optical inversion in highly turbid and cyanobacteria-dominated waters. (2014): IEEE Trans Geosci Remote Sens 52(1):375-88.
Mishra, S., Mishra, D. (2012). Normalized difference chlorophyll index: a novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. Remote Sens Environ 117, 394–406.
Moreno-Madrinan MJ, Al-Hamdan MZ, Rickman DL, Muller-Karger FE (2010) Using the Surface Reflectance MODIS Terra Product to Estimate Turbidity in Tampa Bay, Florida. Remote Sens 2:2713–2728.
Moses, W., Gitelson, A., Berdnikov, S., Povazhnyy, V. (2009). Satellite estimation of chlorophyll-a concentration using the red and NIR bands of MERIS — the Azov Sea case study. IEEE Geosci Remote Sens Lett. 6 (4), 845–849.
Nechad, B., Dogliotti, A. I., Ruddick, K. G., Doxaran, D. (2016). Particulate Backscattering and suspended matter concentration retrieval from remote-sensed turbidity in various coastal and riverine turbid waters. Proceedings of ESA Living Planet Symposium, Prague, 9-13 May 2016, ESA-SP 740.
Nechad, B., Ruddick, K., & Neukermans, G. (2009). Calibration and validation of a generic multisensor algorithm for mapping of turbidity in coastal waters. In, Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2009 (pp. 161-171): SPIE
Nechad, B., Ruddick, K., Park, Y. (2010). Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters". Remote Sens Environ 114 (2010) 854–866.
Neil, C., Spyrakos, E., Hunter, P., Tyler, A. (2019). A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. Remote Sens Environ 229: 159-178.
Novoa, S., Doxaran, D., Ody, A., Vanhellemont, Q., Lafon, V., Lubac, B., & Gernez, P. (2017). Atmospheric corrections and multi-conditional algorithm for multi-sensor remote sensing of suspended particulate matter in low-to-high turbidity levels coastal waters. Remote Sens., 9(1): 61
Ogashawara, I., Mishra, D. R., Nascimento, R. F., Alcântara, E. H., Kampel, M., & Stech, J. L. (2016). Re-parameterization of a quasi-analytical algorithm for colored dissolved organic matter dominant inland waters. International Journal of Applied Earth Observation and Geoinformation, 53, 128-145.
O'Reilly, J, Maritorena, S., O'brien, M., Siegel, D., Toole, D., Menzies, D., Smith, R. et al. (2000). SeaWiFS postlaunch calibration and validation analyses, part 3. NASA tech. memo 206892, no. 11 (2000): 3-8.
Ondrusek M, Stengel E, Kinkade CS, et al (2012) The development of a new optical total suspended matter algorithm for the Chesapeake Bay. Remote Sens Environ 119:243–254.
O'Reilly, J.E., & Werdell, P.J. (2019). Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6. Remote Sens. Environ., 229: 32-47
Pahlevan, N., Smith, B., Binding, C., Gurlin, D., Li, L., Bresciani, M., & Giardino, C. (2021). Hyperspectral retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. Remote Sens. Environ., 253: 112200
Pahlevan, N., Smith, B., Schalles, J., Binding, C., Cao, Z., Ma, R., Alikas, K., Kangro, K., Gurlin, D., & Hà, N. (2020). Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. Remote Sens. Environ., 240: 111604
Pedroso Curtarelli, V., Clemente Faria Barbosa, C., Andrade Maciel, D., Flores Júnior, R., Menino Carlos, F., de Moraes Novo, E. M. L., ... & da Silva, E. F. F. (2020). Diffuse attenuation of clear water tropical reservoir: a remote sensing semi-analytical approach. Remote Sensing, 12(17), 2828.
Petus C, Chust G, Gohin F, et al (2010) Estimating turbidity and total suspended matter in the Adour River plume (South Bay of Biscay) using MODIS 250-m imagery. Cont Shelf Res 30:379–392.

Pitarch, J., & Vanhellemont, Q. (2021). The QAA-RGB: A universal three-band absorption and backscattering retrieval algorithm for high resolution satellite sensors. Development and implementation in ACOLITE. Remote Sensing of Environment, 265, 112667.

Polito CD, Di Polito C, Ciancia E, et al (2016) On the Potential of Robust Satellite Techniques Approach for SPM Monitoring in Coastal Waters: Implementation and Application over the Basilicata Ionian Coastal Waters Using MODIS-Aqua. Remote Sens 8:922.

Robert E, Grippa M, Kergoat L, et al (2016) Monitoring water turbidity and surface suspended sediment concentration of the Bagre Reservoir (Burkina Faso) using MODIS and field reflectance data. Int J Appl Earth Obs. 52:243–251.

Rodrigues, T., Alcântara, E., Watanabe, F., & Imai, N. (2017). Retrieval of Secchi disk depth from a reservoir using a semi-analytical scheme. Remote Sensing of Environment, 198, 213-228.

Ruddick KG, Ovidio F, Rijkeboer M (2000) Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. Appl Opt 39:897–912.

Shi, W., & Wang, M. (2019). A blended inherent optical property algorithm for global satellite ocean color observations. Limnology and Oceanography: Methods, 17(7), 377-394.

Smith, B., Pahlevan, N., Schalles, J., Ruberg, S., Errera, R., Ma, R., Giardino, C., Bresciani, M., Barbosa, C., & Moore, T. (2021). A chlorophyll-a algorithm for Landsat-8 based on mixture density networks. Frontiers in Remote Sensing, 1: 623678

Smith, M.E., Lain, L.R., & Bernard, S. (2018). An optimized chlorophyll a switching algorithm for MERIS and OLCI in phytoplankton-dominated waters. Remote Sens. Environ., 215: 217-227

Spyrakos E, O'Donnell R, Hunter PD, Miller C, Scott M, Simis S, et al. (2018). Optical types of inland and coastal waters. Limnol Oceanogr. 63(2).

Uudeberg, K., Aavaste, A., Kõks, K.-L., Ansper, A., Uusõue, M., Kangro, K., Ansko, I., Ligi, M., Toming, K., & Reinart, A. (2020). Optical water type guided approach to estimate optical water quality parameters. Remote Sens., 12(6): 931

Wang, Y., Shen, F., Sokoletsky, L., & Sun, X. (2017). Validation and calibration of QAA algorithm for CDOM absorption retrieval in the Changjiang (Yangtze) estuarine and coastal waters. Remote Sensing, 9(11), 1192.

Watanabe, F., Mishra, D. R., Astuti, I., Rodrigues, T., Alcântara, E., Imai, N. N., & Barbosa, C. (2016). Parametrization and calibration of a quasi-analytical algorithm for tropical eutrophic waters. ISPRS Journal of Photogrammetry and Remote Sensing, 121, 28-47.

Werther, M., Odermatt, D., Simis, S.G., Gurlin, D., Lehmann, M.K., Kutser, T., Gupana, R., Varley, A., Hunter, P.D., & Tyler, A.N.J.R.S.o.E. (2022). A Bayesian approach for remote sensing of chlorophyll-a and associated retrieval uncertainty in oligotrophic and mesotrophic lakes, 283: 113295

6 Lake Ice cover – LIC

6.1 Candidate algorithms

Four machine learning (ML) algorithms were evaluated for their performance in classifying lake ice cover, open water and cloud cover: multinomial logistic regression (MLR), support vector machine (SVM), random forest (RF), and gradient boosting trees (GBT). The characteristics of each ML algorithm are summarized below. Full details, including the hyperparameters used by each classifier, can be found in Wu et al. (2021).

Multinomial logistic regression (MLR). MLR is an extension of logistic regression applied to multiple response variables. Logistic regression is used as an approach to develop a model of the log odds of binary class probabilities as a linear function of one or more explanatory variables (Murphy, 2013). Then, the model can inversely compute the probability of each class using the explanatory variables of a given unknown sample. Specifically, in order to tackle multi-class problem, MLR designates one of the response variables as the baseline class. In this manner, the probability of membership in the different classes is related to the probability of membership in the baseline class. The optimal values of the function parameters are computed using the training data. The MLR probability estimate for each class falls within a range from 0 to 1, resulting in a realistic probability surface. The maximum probability among the classes is the predicted class for an unknown sample.

Support vector machine (SVM). SVM's basic idea is to determine support vectors to build an optimal boundary separating the given observations in terms of classes (Burges, 1998; Vapnik, 1998; Weston & Watkins, 1999). The distance from the support vectors to a hyperplane is known as the margin. SVM, in its simplest form, is a linear binary classifier that labels a given sample using a hyperplane in the original input space. However, to solve the inseparability problem in the original space, SVM maps multidimensional data into an enlarged feature space to build a hyperplane using a kernel function (e.g., polynomial, radial basis, sigmoid). Since the radial basis function (RBF) kernel has a promising ability in non-linear classification, and as shown in several recent studies (Féret et al., 2019; Ge et al., 2018; Huang et al., 2019; Tian et al., 2020), the RBF kernel was adopted in this research over the linear kernel. We applied the one-vs-one scheme to handle the multiclass problem. SVM is sometimes called a soft margin classifier because training samples could lie on the incorrect side of the hyperplane, thereby creating a violation. The model hyperparameter, *Cost*, is a regularization constant controlling the violation degree. Another model hyperparameter, *Gamma*, is the kernel width of RBF.

Random forest (RF). As an ensemble approach, RF integrates decision trees developed by bagging samples to improve the limitations of the single-tree structure (Breiman, 2001). The bagging creates several subsets randomly from training samples with replacement (i.e. a sample can be collected several times in the same subset whereas other samples are probably not selected in this subset). Subsequently, each data subset is used to train a decision tree. For building a single tree, a random sample with a number of variables is chosen as split candidates from all variables. The number of variables available to a split is one of key RF hyperparameters, denoted as *mtry*. For the whole RF model, the number of trees (*ntree*) is defined *a priori* to develop various independent classifier outputs. The final class of each unknown sample is assigned by the majority vote of all outputs from the trees.

Gradient boosting trees (GBT). GBT is another ensemble classifier inspired by the boosting technique developed by Freund and Schapire (1996). In contrast to RF, GBT applies the entire training dataset on classification rather than resampling partial samples. The training samples are initially assigned equal weights in the first iteration to develop the first tree, and afterwards the weights are altered based on the fitting performance to the training dataset. Misclassified samples in the previous iteration are assigned a higher weight in subsequent iterations. Each tree is also given a weight based on the fitting error. The

final class of an unknown observation is assigned by computing the output of all trees multiplied by their weights. The term, gradient, is associated with iterative functional gradient descent algorithms used to optimize cost functions. Similar to RF, the hyperparameters of GBT, as a tree-based classifier, include the number of variables available to a split (*mtry*) and the number of trees (or iterations) (*ntree*). Moreover, an additional hyperparameter, learning rate (*lr*), controls overfitting in the range between 0 and 1 via shrinkage. The higher *lr* drives a faster learning process, and vice versa.

The candidate algorithm retained for the retrieval of LIC is the RF classifier. RF was found to outperform MLR and SVM, and comparable to GBT for lake ice cover, open water and cloud classification. While RF and GBT provided similar results following a comprehensive accuracy assessment (cross validation (CV): random k-fold as well as spatial and temporal CV), the former was selected for LIC product generation since it was determined to be less sensitive to the to the choice of hyperparameters necessary for classification compared to GBT, MLR and SVM. Validation results that support the selection of the RF classifier for Lakes_cci LIC product generation are provided below.

6.2 Validation results

To identify the best classifier, 17 lakes distributed across the Northern Hemisphere were selected (Figure 71 and Table 19). Training, testing and validation of the four machine learning algorithms found that RF with a combination of visible, near infrared, and mid infrared bands was the best choice for LIC product (*Figure 72*). More specifically, six MODIS (Terra/Aqua) TOA reflectance bands and the solar zenith angle (SZA) are used for feature retrieval (i. e. for labelling as water, ice, or cloud) (Wu et al., 2021). The reflectance bands are MOD02QKM/MYD02QKM at 250 m (band 1: 0.645 µm and band 2: 0.858µm) and MOD02HKM/MYD02HKM at 500 m (band 3: 0.469µm; band 4: 0.555µm; band 6: 1.640 µm; band 7: 2.130 µm) resolutions. Geolocation is provided at 1 km resolution and is interpolated to 250 m.

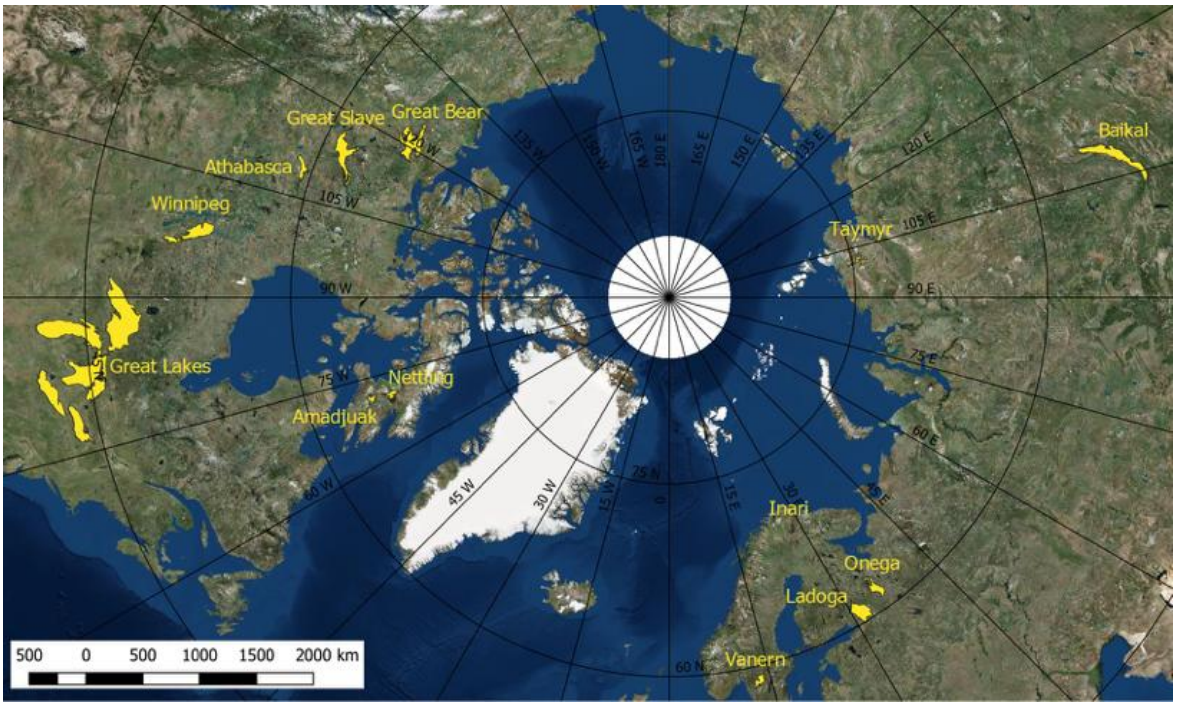


Figure 71 Geographical distribution of lakes used for LIC algorithm development and validation

Table 19 List of lakes for LIC algorithm development and (internal) validation

Lake	Country	Latitude	Longitude	Elevation (m)	Area (km²)
Amadjuak	Canada	64.925	-71.149	113	3,115
Athabasca	Canada	59.424	-109.34	213	7,900
Baikal	Russia	53.525	108.207	456	31,500
Erie	Canada/USA	42.209	-81.246	174	25,821
Great Bear	Canada	66.024	-120.61	186	31,153
Great Slave	Canada	61.579	-114.196	156	28,568
Huron	Canada/USA	44.918	-82.455	176	59,570
Inari	Finland	69.048	27.876	118	1,040
Ladoga	Russia	60.83	31.578	5	18,135
Michigan	USA	43.862	-87.093	177	58,016
Nettilling	Canada	66.42	-70.28	30	5,542
Onega	Russia	61.75	35.407	35	9,890
Ontario	Canada/USA	43.636	-77.727	75	19,009
Superior	Canada/USA	47.945	-87.32	183	82,367
Taymyr	Russia	74.538	101.639	6	4,560
Vanern	Sweden	58.88	13.22	44	5,650
Winnipeg	Canada	52.421	-97.677	217	23,750

Figure 73 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 MLR and SVM and comparable to GBT for lake ice cover, open water and cloud classification in a 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 20).

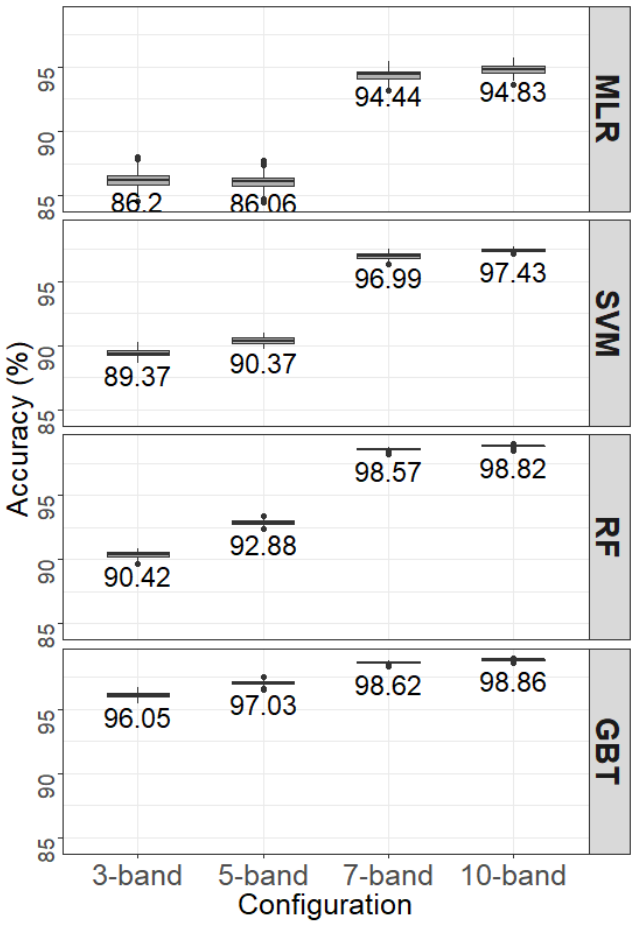


Figure 72 Comparison of classification accuracies (%) obtained with different band configurations across classifiers. The 7-band combination using RF is the one retained for generation of the LIC v2.1 product (Wu et al., 2021)

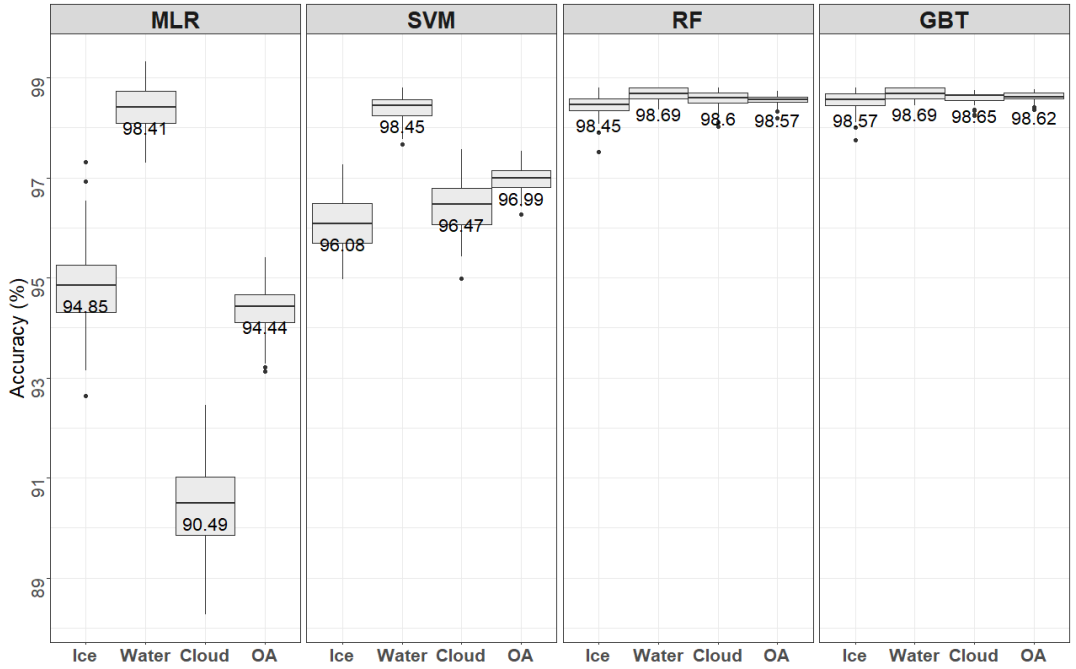


Figure 73 Comparison of accuracies (%) obtained using random 100-fold CV across classifiers for the ice, water and cloud classes individually, and overall (OA) (Wu et al., 2021)

Table 20 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%

6.3 Identified issues

The RF classifier was selected for LIC product generation since it has been shown to outperform threshold-based approaches (as determined during the early stages of Lakes_cci Phase 1) and the other machine learning classifiers as presented in the previous section. High overall accuracies (>95%) has been achieved with the RF classifier in both spatial and temporal transferability assessments (Wu et al., 2021).

As with any lake product generated from optical data, the presence of clouds as well as extensive cloud cover periods and low solar illumination angles, particularly during the fall freeze-up at high latitudes, introduce classification errors and limit the retrieval of open water and ice cover for many days of the year. In CRDP v2.1, highly turbid lakes or sections of lakes were occasionally misclassified as ice-covered during the open water season. In CRDP v3.0 this has been addressed through the addition of a temperature filter which corrects water/ice pixels depending on MODIS emissive band 31. Total uncertainty is provided per-pixel. There are also cases for a few lakes that are characterized by snow-free "blue" clear ice during spring break-up; here ice was misclassified as open water. There is ongoing work to address this issue.

One limitation of the LIC product is that no retrieval is performed when the solar zenith angle is >85 degrees; a limitation due to the use of MODIS shortwave bands that record very low surface reflectance during ice formation late fall and wintertime. Work is underway to explore the addition of new sensors to the LIC production chain (Sentinel-3) as well as explore other alternatives to fill in gaps caused by cloud cover and low-illumination conditions early in the processing chain.

6.4 References

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*(2), 121–167. https://doi.org/10.1023/A:1009715923555

Féret, J. -., le Maire, G., Jay, S., Berveiller, D., Bendoula, R., Hmimina, G., . . . Lefèvre-Fonollosa, M. -. (2019). Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: Potential and limitations of physical modeling and machine learning. *Remote Sensing of Environment*, 231, doi:10.1016/j.rse.2018.11.002

Freund, Y., & Schapire, R. E. (1996). Experiments with a new boosting algorithm. In *International Conference on Machine Learning* (pp. 148–156). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.6252

Ge, J., Meng, B., Liang, T., Feng, Q., Gao, J., Yang, S., . . . Xie, H. (2018). Modeling alpine grassland cover based on MODIS data and support vector machine regression in the headwater region of the Huanghe river, China. *Remote Sensing of Environment*, 218, 162-173. doi:10.1016/j.rse.2018.09.019

Huang, X., Liao, C., Xing, M., Ziniti, B., Wang, J., Shang, J., . . . Torbick, N. (2019). A multi-temporal binary-tree classification using polarimetric RADARSAT-2 imagery. *Remote Sensing of Environment*, 235, doi:10.1016/j.rse.2019.111478

Murphy, K. P. (2013). *Machine learning : a probabilistic perspective*. Cambridge, Mass. [u.a.]: MIT Press. Retrieved from https://www.amazon.com/Machine-Learning-Probabilistic-Perspective-Computation/dp/0262018020/ref=sr_1_2?ie=UTF8&qid=1336857747&sr=8-2

Tian, J., Wang, L., Yin, D., Li, X., Diao, C., Gong, H., . . . Liu, X. (2020). Development of spectral-phenological features for deep learning to understand spartina alterniflora invasion. *Remote Sensing of Environment*, 242, doi:10.1016/j.rse.2020.111745

Vapnik, V. (1998). The Support Vector Method of Function Estimation. In J. A. K. Suykens & J. Vandewalle (Eds.), *Nonlinear Modeling: Advanced Black-Box Techniques* (pp. 55–85). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-5703-6_3

Weston, J., & Watkins, C. (1999). Support vector machines for multi-class pattern recognition. *Proceedings of the 7th European Symposium on Artificial Neural Networks (ESANN-99)*, (April), 219–224.

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 Candidate algorithms

Few studies have investigated the potential of satellite remote sensing data for the estimation of LIT to date. Kang et al. (2010) first showed that brightness temperature (Tb) measurements from the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) at 18.7 GHz frequency (V polarization) to be highly sensitive (R2 = 0.91) to the seasonal evolution of ice thickness on Great Bear Lake (GBL) and Great Slave Lake (GSL), Canada. Based on this finding, Kang et al 2014 proposed empirical (linear regression) equations to estimate LIT for the two lakes using 18.7 GHz V-pol data (2002-2009), achieving a mean bias error (MBE) of 0.06 m and root mean square error (RMSE) of 0.19 m when compared to in-situ measurements. Surface temperature observations of snow-covered lake ice from the Moderate Resolution Imaging Spectroradiometer (MODIS) have also been assessed for the estimation of LIT. Using heat balance terms and snow depth derived from the Canadian Lake Ice Model (CLIMo, Duguay et al. 2003), Kheyrollah et al. (2017) retrieved ice thicknesses up to ~ 1.2 m from MODIS (2002-2014) with an RMSE of 0.17 m and MBE of 0.07 m when comparing LIT values from single pixels (1 km x 1 km) to those from close by near-shore field measurements collected on GSL and Baker Lake, Canada. Beckers et al. (2017) analysed waveforms from CryoSat2 (CS2) Ku-band synthetic aperture radar (SAR) altimetry for the estimation of LIT on the Great Bear Lake and Great Slave Lake. By exploiting the increasing distance between peak radar returns from the snow-ice and ice-water interfaces on the leading edge of waveforms with ice growth, the authors estimated ice thickness empirically with RMSE < 0.33 m when compared to in-situ measurements from the same near-shore location on GSL as in previous investigations. While data from CS2 show strong potential for the retrieval of LIT, the drifting orbit of the satellite makes it difficult to build a geographically precise time series of LIT measurements (i.e. repeated along the same tracks over the lifetime of the satellite) required for climate monitoring. Also, the LIT retracker algorithm developed in Beckers et al. (2017) relies on the empirical thresholding of the radar waveforms that is hard to generalize to follow the LIT evolution, in particular at the seasonal transitions, and can lead to biases and sub-optimal LIT estimates. More recent studies (e.g., Shu et al., 2020; Yang et al., 2021), have estimated LIT with radar altimetry data, more specifically from Sentinel-3 and Jason-3 missions, in the context of lake water level analysis, as the presence of lake ice has been shown to introduce a bias on winter water level measurements. These studies also used empirical methods based on already existing retrackers that are not specifically designed for the estimation of LIT. To overcome these limitations, Mangilli et al. (2022) developed a novel physically-based retracking algorithm, the LRM_LIT retracker, founded on the exploitation of the Ku band radar waveforms data in Low Resolution Mode (LRM) specifically tailored for the retrieval of LIT. The advantage of a physically-based and analytical retracker is that it does not rely on empirical or by-hand settings, allowing to derive robust and continuous LIT estimates over different target lakes and LRM radar altimetry missions, making the LRM_LIT algorithm the suitable tool to build robust and long LIT timeseries for climate monitoring. The LRM_LIT retacker is the algorithm currently being implemented in the lakes_cci LIT processor.

7.2 Validation results

The validation results of the selected LRM_LIT retracker algorithm are presented in detail in Mangilli et al. (2022) and are also described in the ATBD v2.1.1, E3UB v2.1.1 and PVP v2.1.1 documents.

The LRM_LIT retracker retained for the generation of the LIT product has been validated on simulations representative of Jason-like missions over Great Slave Lake. A summary is given in Figure 74, where the top plots refer to winter-like simulated waveforms (left panel) and the LIT histogram (right panel).

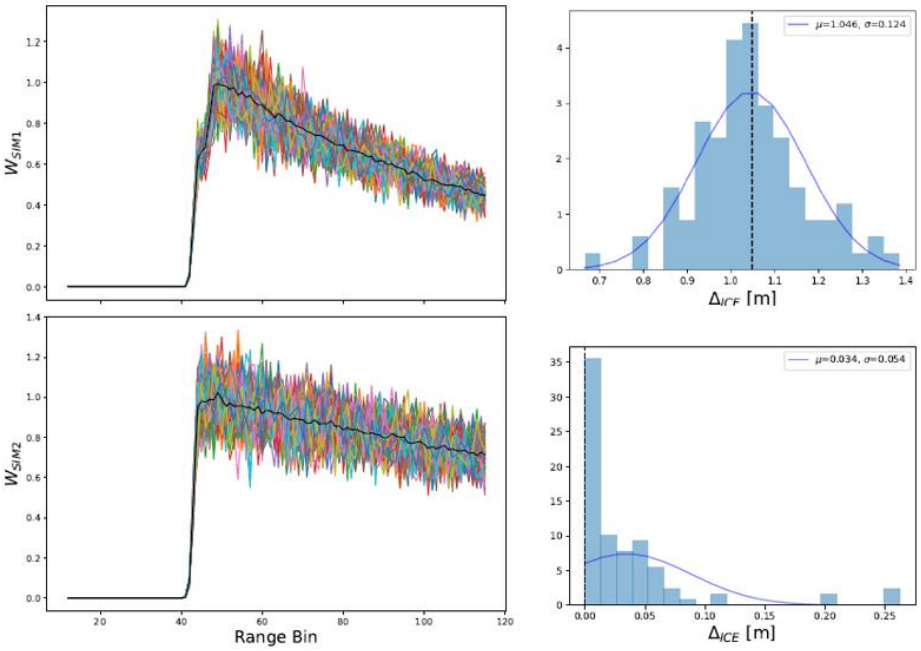


Figure 74: Validation of the LIT estimation with the LRM_LIT retracker on Jason-like simulations. In the left column are shown Jason-like waveform simulations corresponding to the winter-like SIM1 waveforms (top) and to the summer-like SIM2 waveforms. In the right column are shown LIT histograms computed for the winter-like simulations SIM1 (top panel) and for the summer-like simulations SIM2 (bottom panel). The blue lines correspond to the Gaussian fit of the histograms. The input values used to generate the simulations are also shown as dashed black lines

The input value used to generate the simulations is shown as a dashed line. The same description applies to the bottom plots, for the summer-like simulations without the ice signature. In both cases the LRM_LIT retracker gives unbiased LIT results. The uncertainty of the LIT retrieval from the winter-like simulations is ~10 cm.

LIT retrievals from satellite missions have been evaluated against LIT simulations from the thermodynamic lake ice model CLIMo (Duguay et al., 2003). A qualitative comparison with in-situ data was also performed when possible[2]. Figure 75 provides an example of the comparison of the LIT estimations obtained within a winter season over Great Slave Lake with the LRM_LIT retracker applied to Jason-2 data (blue triangles) and Jason-3 data (red stars) and LIT from CLIMo simulations with different on-ice snow depth scenarios (diamonds) and in-situ data (circles). There is an excellent agreement between Jason-2 and Jason-3 LIT estimates, which are fully compatible with the thermodynamic simulations and qualitatively in agreement with in-situ data. We note that, in general, the LIT melting phase is detected earlier with the satellite-based measurements because of snow melting that perturbs the radar echoes.

[2] It is worth noting that the comparison between LIT estimates from satellite missions and in-situ data must be taken with caution. In-situ data are typically collected near the shore, while satellite data are taken from the middle of the lake to avoid land contamination. These can represent different environments in terms of bathymetry, wind exposure, snow type and quantity. All these parameters play a key role on ice formation and thickness and they can lead to significant LIT differences.

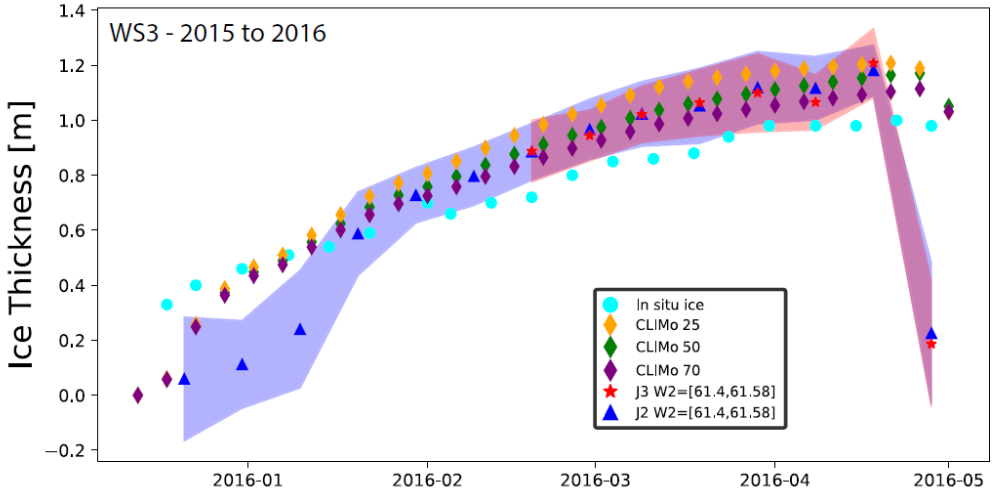


Figure 75 LIT estimates over Great Slave Lake for the 2015-16 winter season. Shown are a comparison between LIT estimates with the LRM_LIT retracker from Jason-2 (triangles) and Jason-3 (stars) data, CLIMo simulations (diamonds) and in-situ data (circles). The shaded areas correspond to the LIT estimation uncertainties computed from Jason-2 data (blue) and Jason-3 data (red). Three different realizations of CLIMo simulations are shown by varying the amount of snow on the ice. The in-situ data consist of ice thickness measurements collected in Back Bay near Yellowknife

To quantify the comparison between the LIT estimates from Jason-2, Jason-3 and CLIMo, two statistics are computed, the MBE:

$$MBE = \frac{1}{N}\sum_{i}^{N}(LIT_i^{J2} - LIT_i^{DS}) \qquad [7.1]$$

And the RMSE:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(LIT_i^{J2} - LIT_i^{DS})^2} \qquad [7.2]$$

of the LIT estimates derived from Jason-2, LIT^{J2}_i , and the other data sets, LIT^{DS}_i , for the N measurements obtained in the middle of the winter season. In the illustrative case of the 2015-16 winter season shown in Figure 75, agreement between the Jason-2 (blue triangles) and Jason-3 (red stars) LIT measurements is excellent. In the middle of the ice season, the MBE is only 0.013 m and the RMSE is 0.024 m between the two data sets. Also, the difference in the LIT mean value is only 0.02 m and that of maximum LIT is 0.025 m. Both Jason-2 and Jason-3 LIT are in strong agreement with the thermodynamics simulations with 50% of snow on ice as input (CLIMo-50 simulations), in particular in the middle of the ice season where the MBE between Jason-2 and CLIMo-50 is less than 0.01 m and the RMSE is 0.019 m. Overall, these results demonstrate that LIT estimates can be retrieved from radar altimetry data that are compatible with thermodynamic simulations and qualitatively in agreement with in-situ measurements.

Finally, the superposition of the LIT retrievals on RADARSAT-2 synthetic aperture radar (SAR) and MODIS optical images allows for a better assessment of the consistency of the LIT estimates as they provide valuable information about the state of the ice and overlaying snow cover. Figure 76 shows (left column from top to bottom) the Jason-2 LIT estimates superimposed on RADARSAT-2/MODIS images obtained within one day of Jason-2 overpasses in December, February, March and end of April. The ice thickness is colour coded and ranges from no or thin ice (0-0.32 m) in light blue to LIT in the range of 1.28- 1.60 m in pink. Plots on the right-hand side of Figure 76, show the spatial evolution of the Jason-2 LIT estimates (top plots) and the corresponding evolution of the reduced chi-squared statistics as a function of the latitude (bottom plots) for the same dates as in the left column.

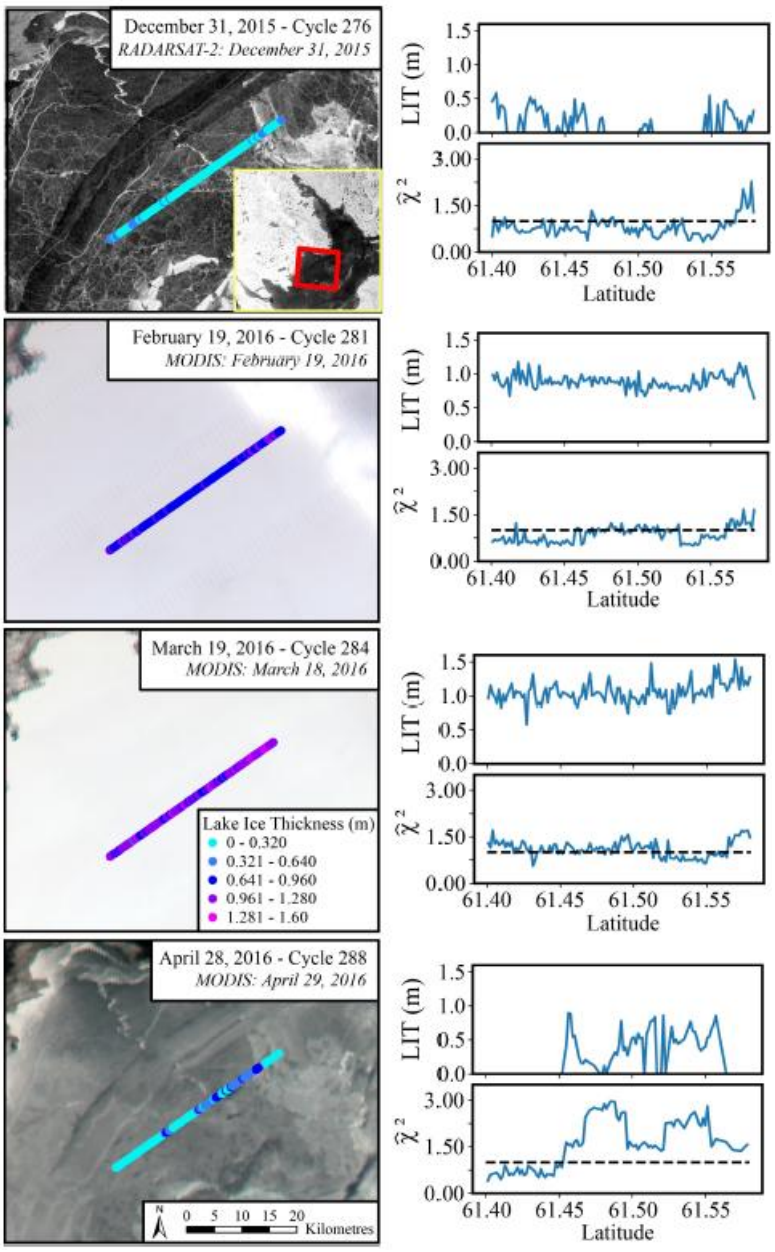


Figure 76: Evolution of Jason-2 LIT estimates over Great Slave Lake along track 45 for WS3 (winter season 2015-2016). Plots in the left column show, from top to bottom, along-track Jason-2 LIT estimates superimposed on MODIS images on the same date or within one day in December, February, March and end of April. Plots in the right column show the evolution of the Jason-2 LIT estimates as a function of latitude (top plots) and of the reduced chi-squared statistics as a function of the latitude (bottom plots) along the track

7.3 Identified issues

- The LRM_LIT retracker can capture the seasonal transitions of ice forming and melting but cannot precisely follow the ice evolution at the transitions because of the difficulty of retracking heterogeneous surfaces when the ice is too thin and when snow on the ice surface begins to melt.
- The LRM_LIT retracker works if the ice related signature, that is, the step like feature in the leading edge of the radar waveforms, is present. This signature in the LRM waveform data depends on the properties and thickness of the snowpack and the ice layer and could be erased if some conditions are not met, as for instance in the case of snow-free lake ice or melting snow on the ice surface. Therefore, there could be lakes for which the ice signature is not clearly marked or not present. In this case, these targets could not be analysed with the LRM_LIT retracker.

7.4 References

Beckers, J. F., J. Alec Casey, and C. Haas. (2017). Retrievals of lake ice thickness from Great Slave Lake and Great Bear Lake using CryoSat-2. *IEEE Transactions on Geoscience and Remote Sensing* 55(7): 3708–3720.

Kang, K.-K., C. R. Duguay, S. E. L. Howell, C. Derksen, and R. Kelly. (2010). Sensitivity of AMSR-E brightness temperatures to the seasonal evolution of lake ice thickness. *IEEE Geoscience and Remote Sensing Letters* 7(4): 751–755.

Kang, K.-K., C. Duguay, J. Lemmetyinen, and Y. Gel. (2014). Estimation of ice thickness on large northern lakes from AMSR-E brightness temperature measurements. *Remote Sensing of Environment* 150: 1–19.

Kheyrollah Pour, H., C. R. Duguay, K. A. Scott, and K.- K. Kang. (2017). Improvement of Lake Ice Thickness Retrieval From MODIS Satellite Data Using a Thermodynamic Model. *IEEE Transactions on Geoscience and Remote Sensing* 55(10): 5956–5965.

Mangilli, A., P. Thibaut, C. R. Duguay and J. Murfitt. (2022). A New Approach for the Estimation of Lake Ice Thickness From Conventional Radar Altimetry. *IEEE Transactions on Geoscience and Remote Sensing* 60: 1-15, 4305515. https://doi.org/10.1109/TGRS.2022.3197109

Shu, S., H. Liu, R. A. Beck, F. Frappart, J. Korhonen, M. Xu, B. Yang, K. M. Hinkel, Y. Huang, and B. Yu. (2020). Analysis of Sentinel-3 SAR altimetry waveform retracking algorithms for deriving temporally consistent water levels over ice-covered lakes. *Remote Sensing of Environment* 239: 111643.

Yang, Y., P. Moore, Z. Li, and F. Li. (2021). Lake level change from satellite altimetry over seasonally ice-covered lakes in the Mackenzie River Basin. *IEEE Transactions on Geoscience and Remote Sensing* 59 (10): 8143–8152.

8 Lake Storage Change – LSC

8.1 Candidate algorithms

In the proposed algorithm, the first step is to check the surface variability of the lake to distinguish between surface-area-varying and unvarying lakes. This information determines how lake storage change will be retrieved. If the lake to monitor is varying surface-wise, an estimation of the hypsometric curve (through parametric estimation with polynomial model) from height and surfaces time series is done. If the hypsometric coefficients are already available from other products (LWL, LWE), they will be used to extract LSC from the hypsometric curve. If the lake is considered as unvarying surface-wise, the LSC is directly computed from the mean area of the lake and height time series with the basic volume formula.

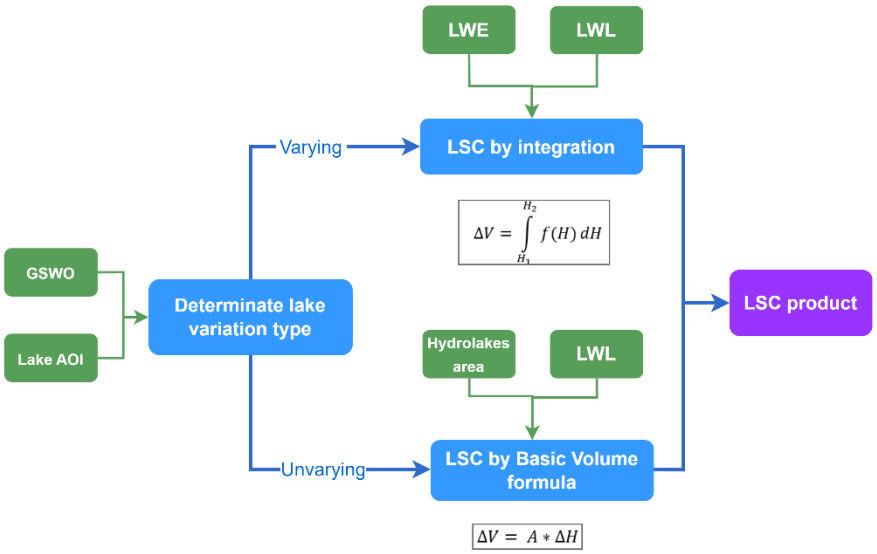


Figure 77: LSC estimation process overview

In the absence of altitude time series, a complementary method has been developed to reconstruct altitude time series. Indeed, the use of precise water surfaces is a starting point to estimate the water level from their contours projected on DEM. This way, both height and surface are retrieved at the same time, which allows the direct estimation of the hypsometric curve.

8.2 Validation results

In situ lake volume time series data were pre-processed to extract lake storage change time series, the dates of which would coincide with the data produced. From these data, the coefficient of determination (R^2), Root Mean Square Difference (RMSD), and bias were assessed using linear regression statistics.

Lake volume time-series for validation on our study dataset were only available for the Rosarito reservoir. In addition, the Richland-Chambers reservoir, the Tres Marias Reservoir and the Songhua reservoir have an area-height-volume relationship which allows us to directly relate the satellite altimetry height to the volume. Bathymetry data were accessible for Lakes Garda, Laguna de Bay and Mille Lacs.

In the end, comparison between the in-situ observations and satellite-derived LSC are as displayed in Figure 78 and Table 21.

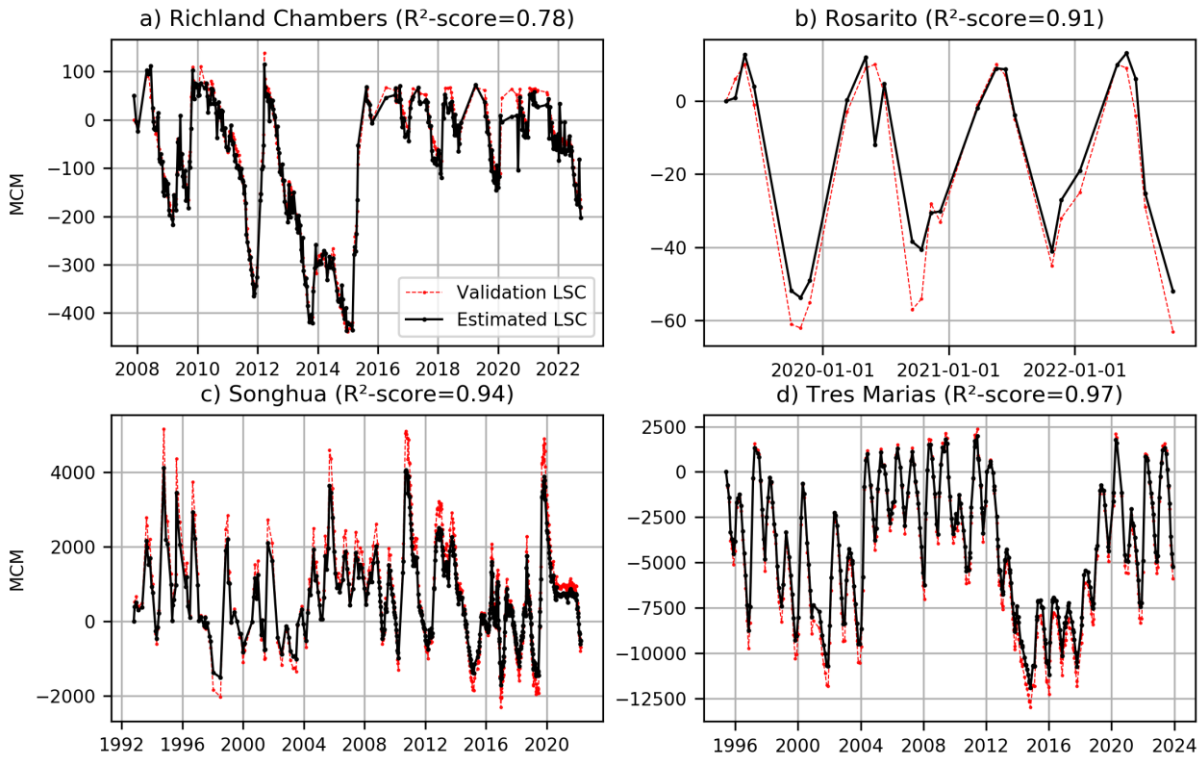


Figure 78: Validation of LSC time-series with in-situ data. R2 is included for each lake

Table 21: Statistics on LSC time-series comparison with in-situ data

Lake name	R^2	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

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.

8.3 Identified issues

Generally, few information about the in-situ data production and quality is available and may induce errors. 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.

8.4 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 Inter-product consistency

Consistency is tested between all products where a meaningful relationship is expected from physical principles. Consistency between ECV Products is always analysed on the CRDP following its publication, so that CRDP v3.0.0 is currently being prepared with input from consistency analysis on CRDP v2.1, reported here. Comparison of seasonal counts of valid pixels and ECV Product climatology already allows for interesting insights into consistency in products derived from multiple sensor systems. One such consistency test, between LIC, LSWT and LWLR products of CRDP v2.1, is introduced in this chapter. The method is based on the analysis of per-lake climatology, specifically the number of valid pixels and mean values of continuous variables. The information is provided in a set of three climatological time series figures, which can be easily analysed visually. Some automatically identified points in time like start and end of full ice cover or ice-free period are summarized in maps and are presented in the Lakes_cci dashboard.

9.1 Data description

Consistency CRDP v2.1 (daily L3 data) has been used in this study. The analysis of LIC and LSWT consistency considers all lakes individually, covering the full temporal extent of the dataset. LSWT data is always restricted to the recommended quality levels (values 3-5).

9.2 Comparison nethods

9.2.1 Consistency in spatial coverage

While deriving climatologies from lake observation data by aggregating to a single lake-average or median value per day, we may neglect differences in spatial coverage. The spatial data distribution for the different product types can vary a lot in relative coverage both spatially and temporally. Orbital coverage and different quality filtering techniques on different sensors can lead to different spatial patterns. When interpreting climatologies, users should be aware of the spatial coverage they are representing.

Two types of observation counts are gathered from the entire dataset on a pixel basis: the number of valid pixels per pixel and the number of daily observations reported in a year. Daily observations are not equivalent to satellite overpasses as there can be multiple overpasses per day in polar regions, while the L3S dataset aggregates these into a single ‘daily observation’. Thus, in the L3S dataset, an observation refers to one or more satellite overpasses. At lower latitudes, overpasses can become less frequent than daily, and it is here, that the number of observations becomes equivalent to overpasses.

Valid pixels are defined for each product separately:

- LSWT: lswt_quality_level > 3
- LWLR: lwlr_quality_flag < 2 & $R_w(560) < 1$
- LIC: lake_ice_cover_flag == 2 (forms ice)

The LIC product is counted according to the classes which are assigned to each pixel. Counts are gathered for the classes ice, water and cloud. Classification results are counted, if the LIC flag is valid (see above). The reference number of observations is based on all days with the valid expression fulfilled and the lake_ice_cover_class reporting a value greater than zero.

These numbers are aggregated over the time range restricted to years 2002 – 2022, when LWLR products are available. Maps of relative percentages of number of valid pixels to overpass counts reveal the climatological coverage of the lake.

The Great Salt Lake can showcase as an extreme example. A dam divides the lake into two distinct main water bodies which is clearly visible in LSWT product, while the dam can be seen in the LWLR product (Figure 79). The processing and quality control methods of the LSWT products are sensitive to the change in water types from the Northern to the Southern basin, but it looks like LWLR is not influenced. Looking at relative coverage per year and sensor (Figure 80) reveals a large difference between the sensors and their distribution of valid pixels. While MERIS coverage shows a similar pattern to LSWT, with the majority of years having lower coverage in the northern part of the lake, the coverage pattern changes dramatically for MODIS. Larger areas are completely uncovered, which either refers to actual changes in water level or to the severe differences in atmospheric correction and the quality control of the resulting water leaving reflectance spectra. All the shallow parts are excluded. The coverage difference between the northern and southern parts are particularly evident during 2012 and 2013. After switching to OLCI the pattern turns to higher coverage in the northern part for the years 2020-2022. On average, this will lead to the balanced distribution over all years. The MODIS LWLR coverage is well reflected in the LIC water classification, which is also based on MODIS data, though it has been independently processed.

Relative temporal coverage of LWLR products is lower than coverage of LSWT products. LWLR relative coverage shows a similar pattern as the LIC water classification, which is good sign for consistent water detection across sensors and algorithms.

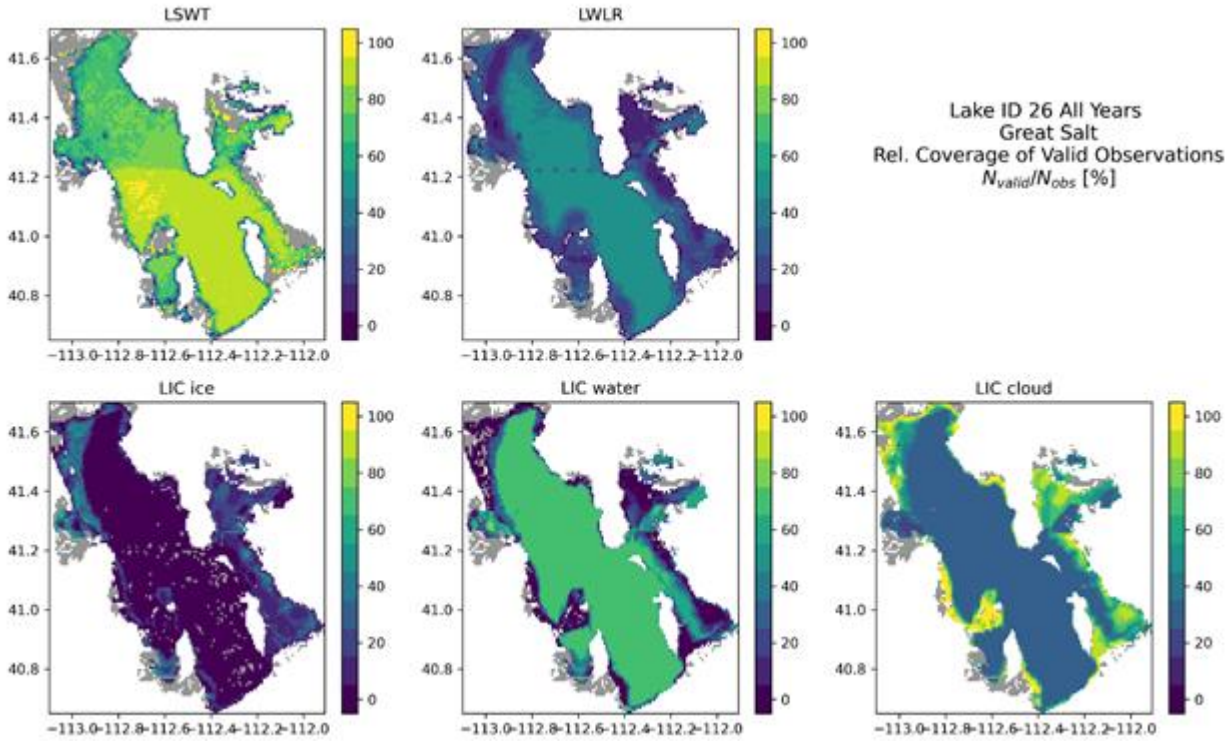


Figure 79 Percentage of coverage (valid pixels in relation to days with observations in the L3 product) for the products LSWT (quality >3) and LWLR (based on band 560nm) (top row) and relative occurrence of ice, water and clouds according to the LIC classification (bottom).Marked in grey are pixels with no valid observations, but which are still part of the static lake water mask.

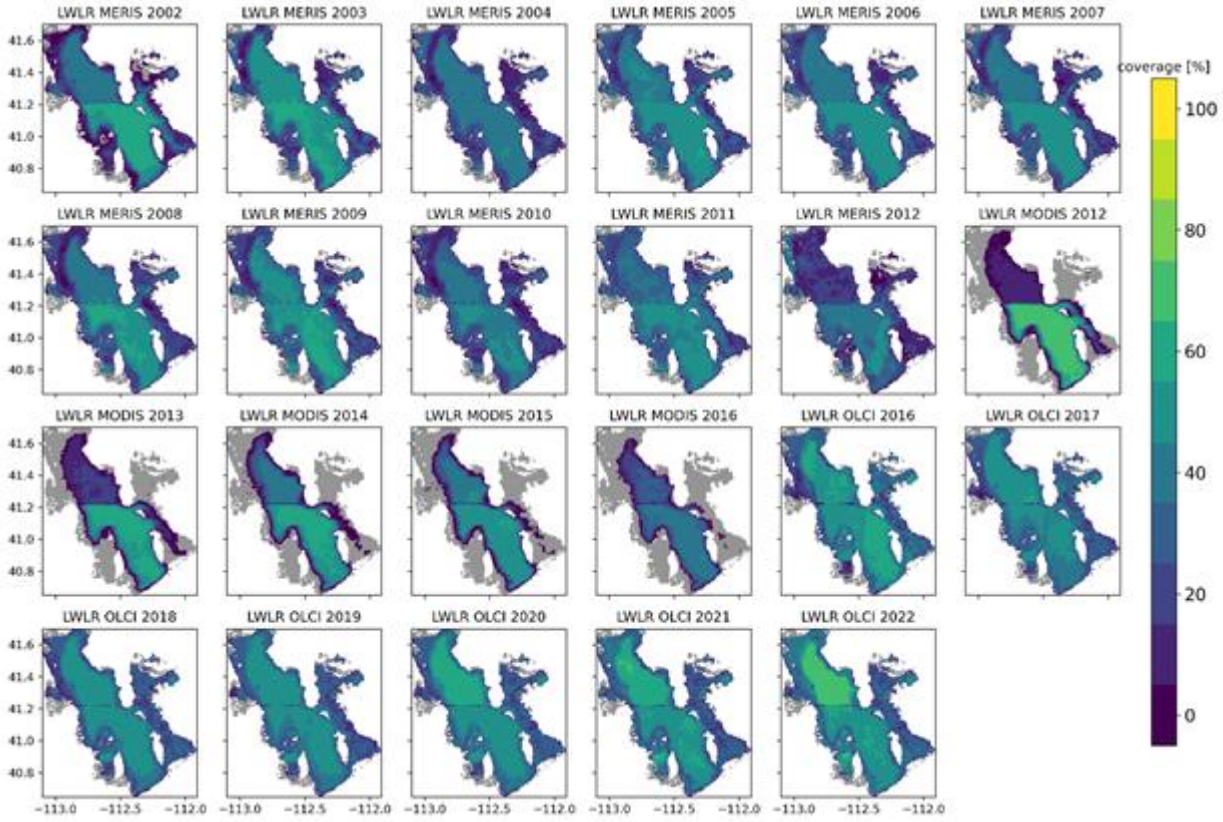


Figure 80 Percentage temporal coverage per year and per sensor for the lake water-leaving reflectance. The year 2012 is split to show MERIS (till April) and MODIS (starting in April) separately, and again in 2016 to separate coverage by MODIS and OLCI.

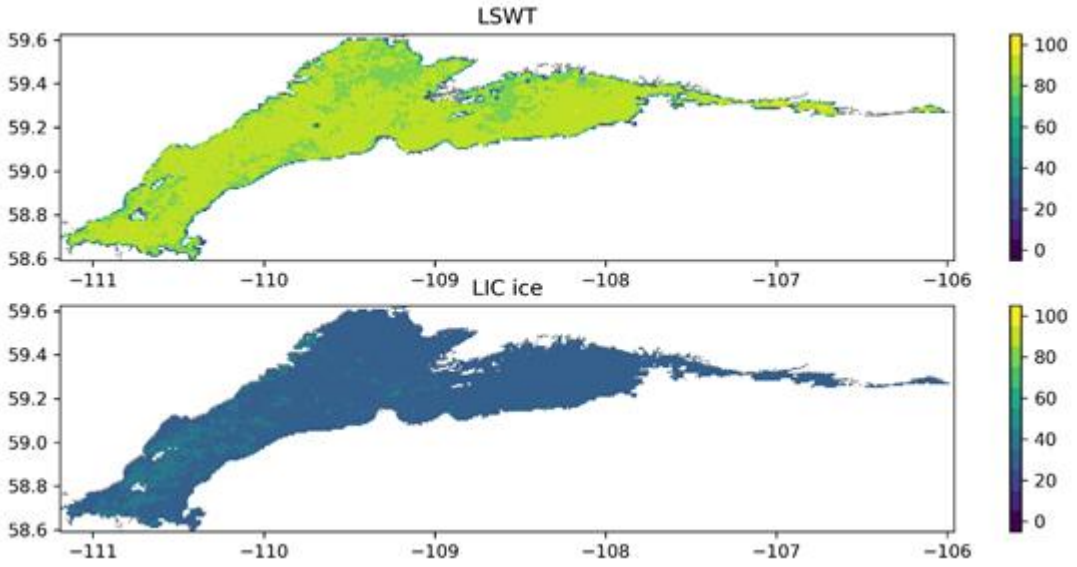


Figure 81 Percentage of coverage (valid pixels in relation to days with observations in the L3 product) for the products LSWT (quality >3) and the LIC ice product for Lake Athabasca (Canada).

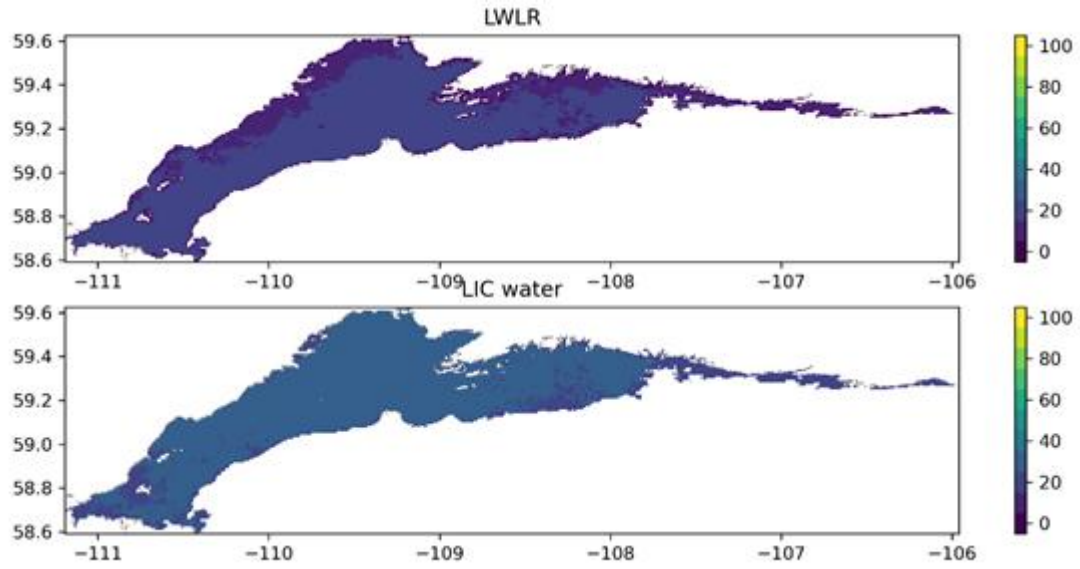


Figure 82 Percentage of coverage (valid pixels in relation to days with observations in the L3 product) for the products LWLR and the LIC water product for Lake Athabasca (Canada).

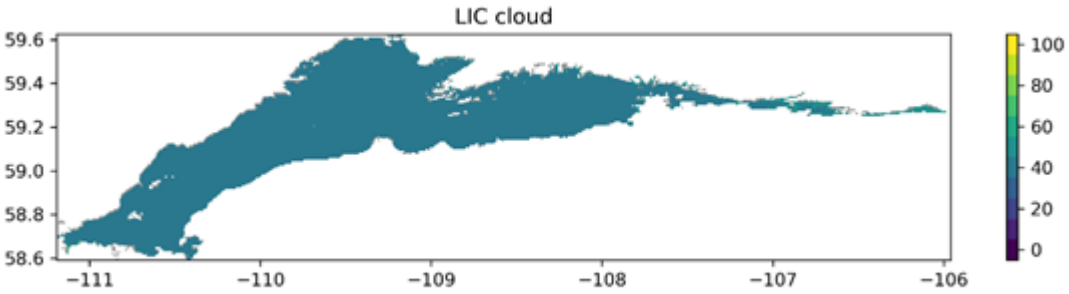


Figure 83 Relative Percentage of coverage (valid pixels in relation to days with observations in the L3 product) for the LIC cloud product for Lake Athabasca (Canada).

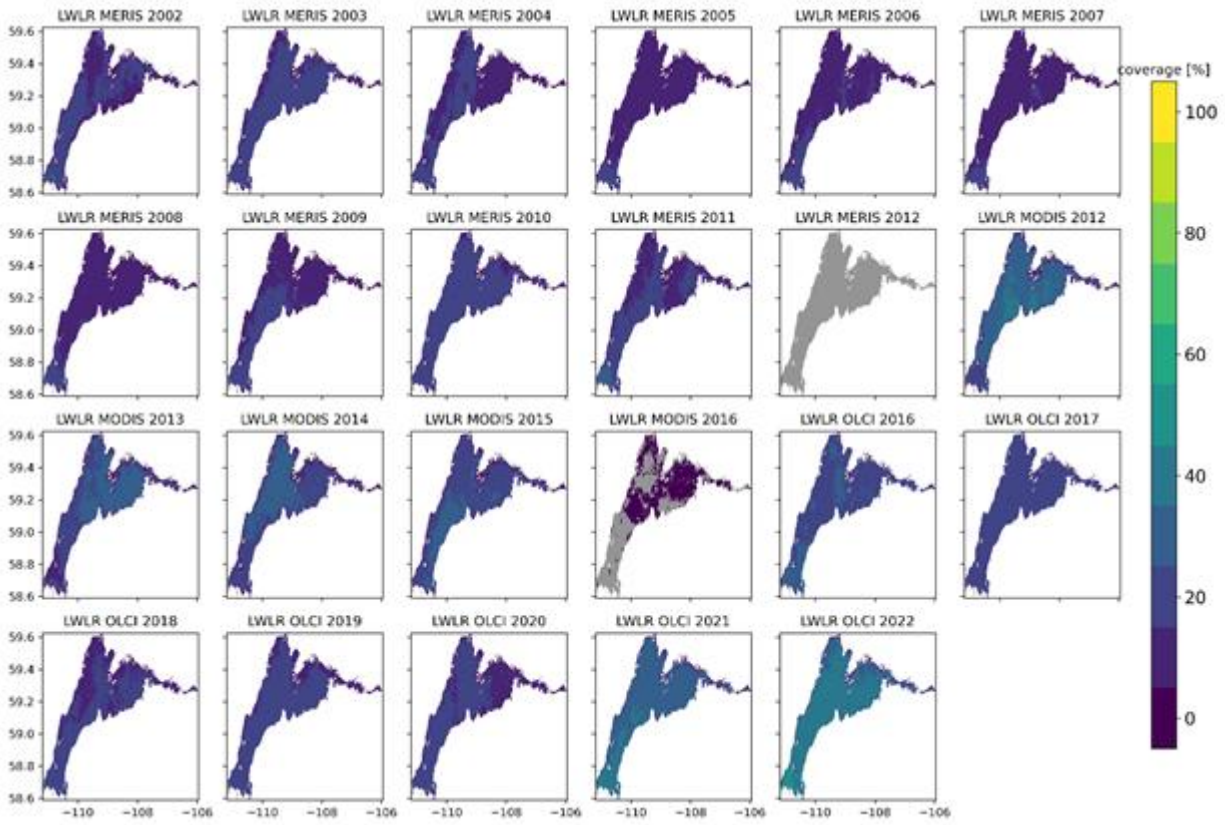


Figure 84 Percentage temporal coverage per year and per sensor for the lake water-leaving reflectance in Lake Athabasca (Canada). The year 2012 is split to show MERIS (till April) and MODIS (starting in April) data separately, and again in 2016 the coverage of MODIS and OLCI.

9.2.2 Consistency in climatologies and counts

Previous attempts in defining consistency measures got too complicated to interpret or summarize efficiently. The following simplistic approach is promising in terms of automated analysis: climatologies of continuous variables like LSWT and chl-a are compared to climatologies of counts of valid pixels of the different products, e.g. the classification of LIC or the valid water pixels of LWLR. The database comprises the L3 products so that all products are available on the same grid.

The climatologies are combined thematically in three groups per lake:

- Climatologies of the amount of water pixels for the three products LWLR (either all years or per sensor), LIC water and LSWT.
- LIC classes (water, ice, cloud, total) summarized as daily counts per lake combined with the LSWT climatology.
- LWLR climatologies of chl-a for all years, chl-a by sensor and turbidity for all years.

The analysis of these consistency measures is demonstrated on Lake Athabasca (Canada, maximum distance to land 20.6km, center point at 59.3319°N, -109.4181°E, ID 23). Figure 81 to Figure 84 give an overview of the data coverage for the different products.

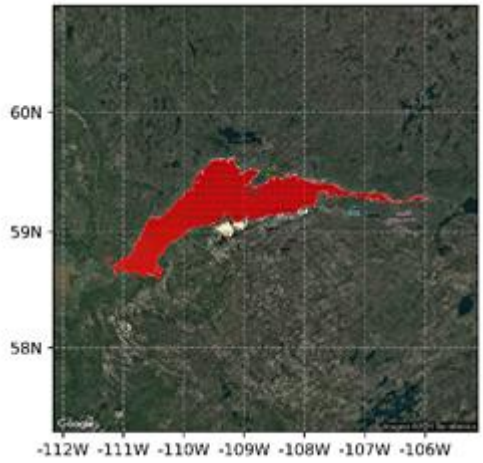


Figure 85 Lake Athabasca (Canada), ID 23.

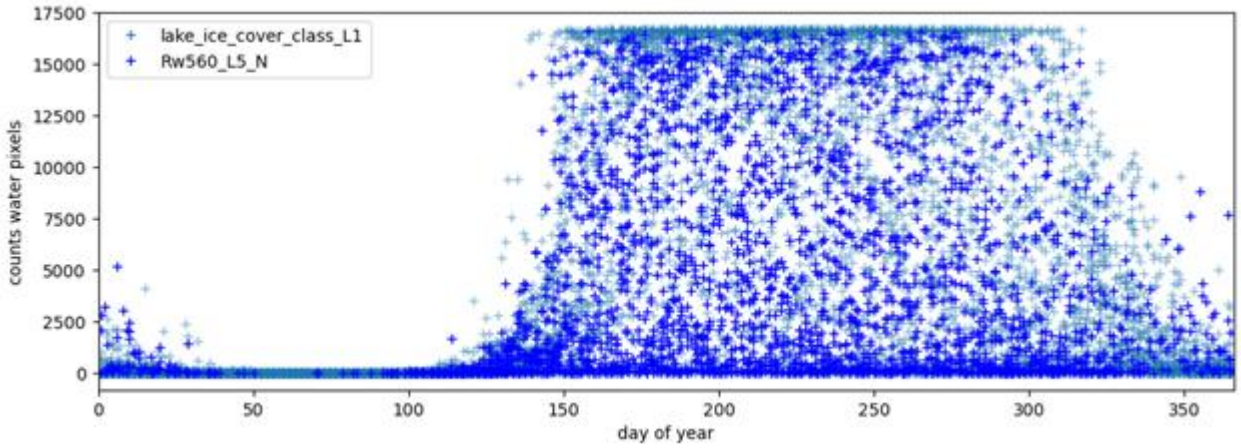


Figure 86 Timeseries count data of Lake Athabasca for LIC water class ('lake_ice_cover_class_L1', blue) and valid water pixels of LWLR ('Rw560_L5_N', orange) per day of year. Climatologies are calculated based on this dataset.

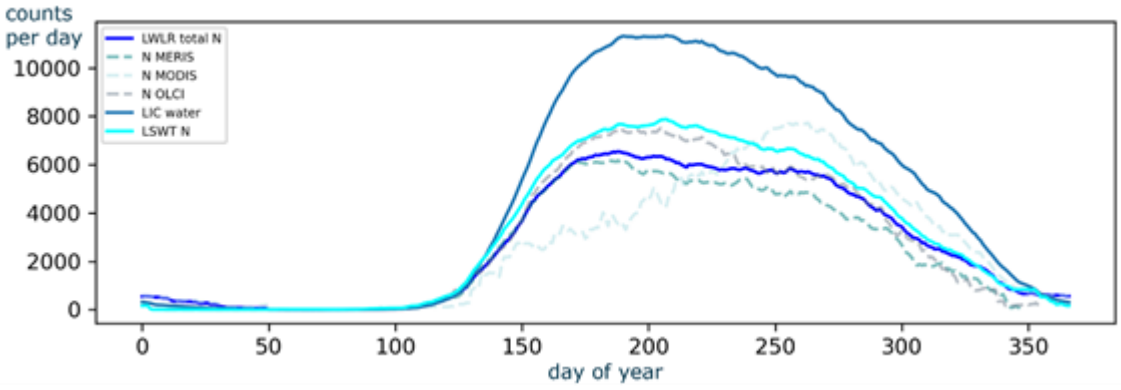


Figure 87 Climatologies of counts of water pixels in Lake Athabasca. LWLR comprises MERIS, MODIS and OLCI data, whose climatologies are also shown separately. LIC water is based on MODIS data, LSWT on several sensors.

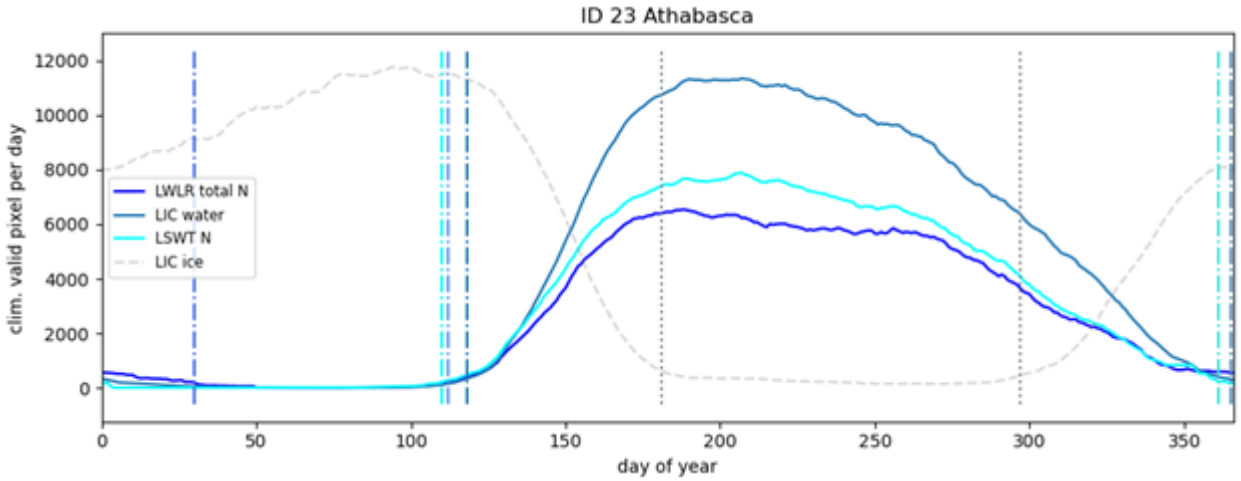


Figure 88 LIC ice classification (grey, dashed) as complement to the counts of water pixels from different products. Start and end of melting and ice formation are highlighted by vertical dotted lines: start of melting (blue-grey, dash-dotted) based on LIC water passing a threshold of 3% of the maximum range, end of melting (grey, dotted) based on LIC ice, start of ice formation (grey, dotted) based on LIC ice and end of ice formation (blue-grey, dash-dotted) based on LIC water. Other lines mark the passing of the thresholds based on LWLR valid water pixels (blue) and LSWT pixels (light blue).

9.2.2.1 Climatology of water counts

The number of valid water pixels for the three spatial products (lake water leaving reflectance LWLR, lake ice cover LIC and lake surface water temperature LSWT) are summarised as climatologies. The annual cycle of observed water pixels can be compared in shape and amplitude (Figure 87). The total number of valid observations in terms of water leaving reflectance and surface water temperature are much more similar in shape and number than the water class of the lake ice classification. The amplitude of the climatological counts represents the average spatial coverage of the data products. The LIC product covers the lake fully during summer (Figure 86), as the plateau in the number of water pixels shows. The number of valid LWLR pixels is restricted not only by clouds but also by further quality flags and the lower repetition time of MERIS and OLCI compared to MODIS.

The start and end of melting or ice formation can be detected at a 3% threshold above minimum of counts. This minimum is not necessarily zero, the values are scaled by the range of counts.

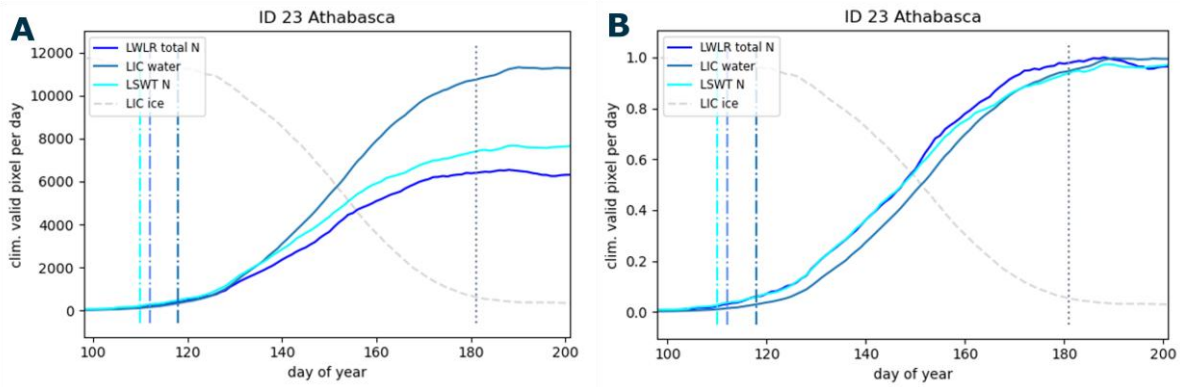


Figure 89 Climatology detail of melting season for Lake Athabasca (figure A) and normalised counts scaled by their respective maximum (figure B). The dotted line marks the end of melting using the LIC ice counts, the dashed-dotted lines indicate the start of melting based on LWLR, LWLR and LIC water counts.

The impact of melting ice starts roughly at the same time (Figure 89) in the products considered, with LSWT and LWLR preceding LIC. Normalised climatological counts reveal that the slope associated with the melting is very similar for all products (Figure 89B). The LIC classification is more conservative with respect to ice, so that the relative amount of water pixels starts to increase approximately 5 days later.

The identification of water in LSWT and LWLR products (light blue and dark blue) is very similar in the beginning of the melting season, while LSWT and LIC are more consistent towards the end, which could be associated with dark ice. The LWLR product reaches its maximum spatial extent first (day of year 188).

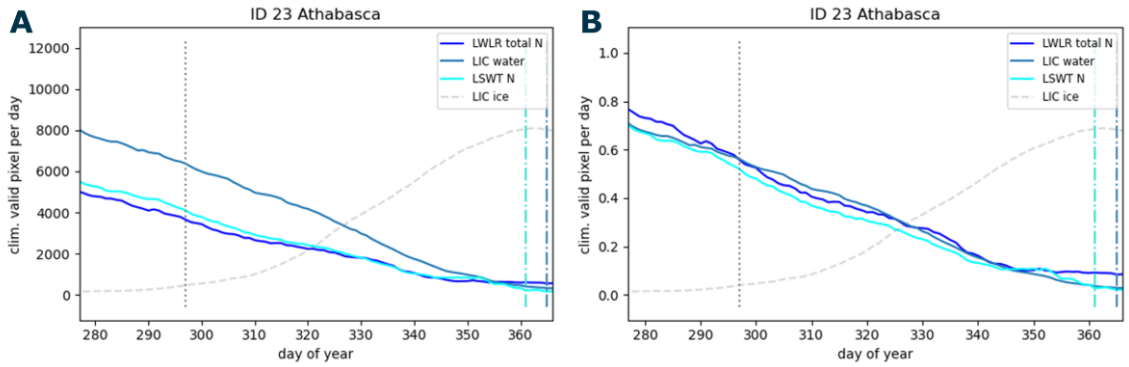


Figure 90 Climatology detail of summer and ice formation season for Lake Athabasca (figure A) and normalized counts scaled by their respective maximum (figure B). According to LIC ice class formation begins at day 297. Before, decline is due to increasing cloud coverage.

Before the beginning of ice formation (day of year 297, see Figure 90), the LWLR product still considers pixels as water where LSWT and LIC behave consistently and see less water. These differences arise from cloud masking procedures.

The MODIS LWLR coverage (Figure 87) behaves very differently from the other sensors, with significantly fewer valid pixels in summer and a later maximum spatial coverage before the start of the ice formation.

9.2.2.2 Climatology of LIC classes

In addition to the LIC water counts the other classes are also represented in their climatological form. During melting season, the behaviour of ice and water is nicely reciprocal, because cloud coverage is almost constant. The decrease in water pixel in summer and autumn is driven by the increase in cloud coverage. LSWT starts its increase before the defined melting threshold is reached in LIC water. The interannual variability of LSWT is large during the melting period.

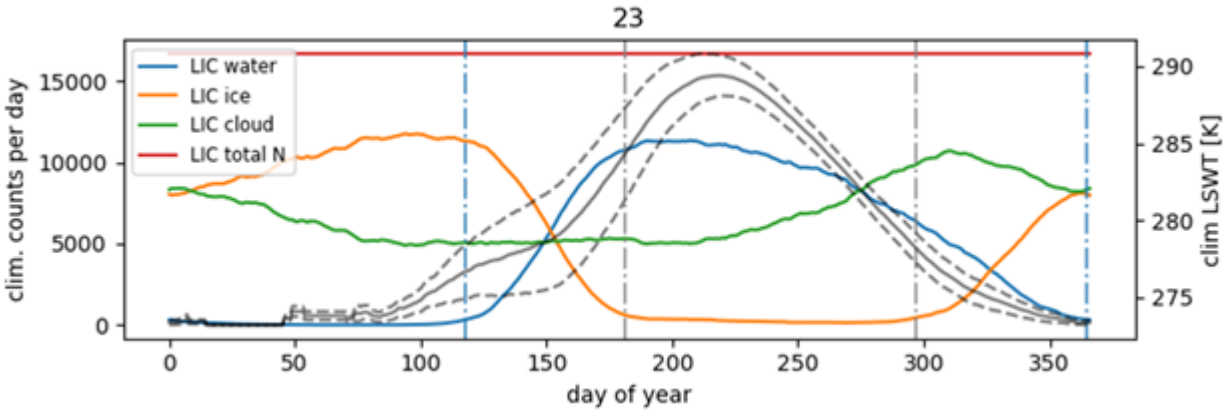


Figure 91 Climatology of LIC classes water, ice, cloud and total numbers combined with LSWT climatology plus standard deviation (dashed line). Vertical lines mark start and stop days of melting and ice formation based on the LIC ice and LIC water counts.

9.2.2.3 *Evaluating LWLR products*

The climatologies of the number of observation are supplemented by climatologies based on the products themselves (Figure 92). The LIC classification counts are combined with the lake surface water temperature climatology (Figure 92A), which allows to check the consistency of ice and water observations against water temperature.

Chlorophyll-a concentration and turbidity climatologies are calculated from respective products, which summarize either the entire observation period (chl all) or have been split into the observations of different sensors (Figure 92B). They are very stable during ice-free periods. The increase in the chlorophyll-a climatology across all sensors is driven by MODIS-derived products, while MERIS and OLCI provide almost stable values during ice formation. Sensor-dependent flagging of products can cause these differences, and particular caution should be employed to the winter months during the MODIS period. The climatology during melting and ice formation is easily influenced by data scarcity and outliers during the period of minimum open water.

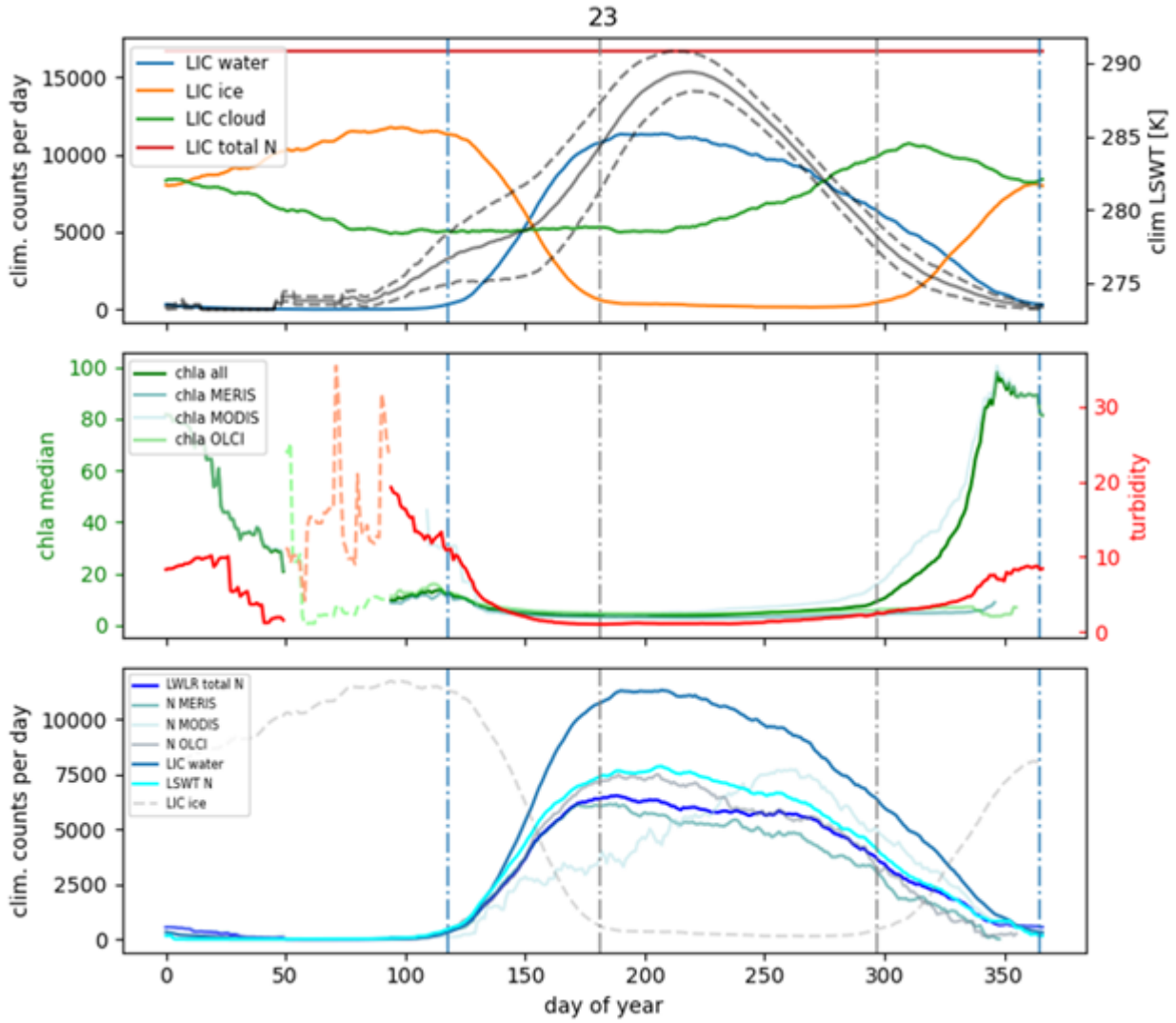


Figure 92 Climatologies of all products (counts and continuous). Vertical lines indicate ice melting and formation periods.

9.2.3 Summary of climatological analysis

While climatologies are a good approach to describe the status of the dataset (Table 22), it may be less straightforward to users how they can use this information to their benefit. If their focus lies on

biogeochemical water products like chlorophyll-a or turbidity, the sensor dependent climatologies can inform on biases introduced by either different processing or different flagging and might help to select a time range in which consistency between sensors is high.

From the automated detection of start and stop times of ice formation and melting based on the LIC ice and water counts climatology, a series of maps can be produced which show nicely the expected or the suspect behaviour of some lakes. Ice free periods are calculated from the LIC ice climatology exploiting the 3% threshold of the ice cover range. The ice-free periods increase towards the equator in general, unless the lakes are situated at high altitudes like in Himalayan lakes (Figure 93).

The climatological start of ice formation (in terms of day of the year) decreases with higher latitudes in general (Figure 94Figure 94). Some very early onsets in mid latitude regions suggest that the algorithm is not working properly here. It is known that dry-fallen areas can be misinterpreted as ice and lead to such results. The global pattern of start of ice melting is not as clear as the one for ice formation (Figure 95).

From the LIC water count climatologies the duration of full ice coverage is calculated (Figure 96). In this example of North America, the duration increases towards the North. Outliers of long ice cover times in the mid-latitude zone are associated with salt lakes or dry-fallen areas, which are not interpreted as water and not automatically detected within this algorithm. Some shorter durations in the latitude range between 65°-70°N are caused by missing observation due to low sun elevation in winter and therefore missing information for the correct determination end day of ice formation. The algorithm needs to be adjusted here, or these values flagged as unreliable due to lack of data. Here, end time of ice formation falls into the unobservable period.

For the Lakes-cci dashboard, these variables are calculated on the yearly bases. To reduce the noise (caused by outliers due to misclassification) in the count data, the temperature information is used as a filtering criterion.

Table 22 Variables from analysis of climatological count data of L3 products

Variable	Unit	Based on Data, climatology
Start ice formation	Day of year	LIC ice
End ice formation	Day of year	LIC water
Start melting	Day of year	LIC water
End melting	Day of year	LIC ice
Length melting	days	LIC ice, LIC water
Length ice formation	Days	LIC ice, LIC water
Ice free period	Days	LIC ice
Full ice cover	days	LIC water
Day of maximum LSWT	Day of year	LSWT
Relative ice cover at Maximum of ice cover	1	Maximum LIC ice / (maximum LIC water + ice)

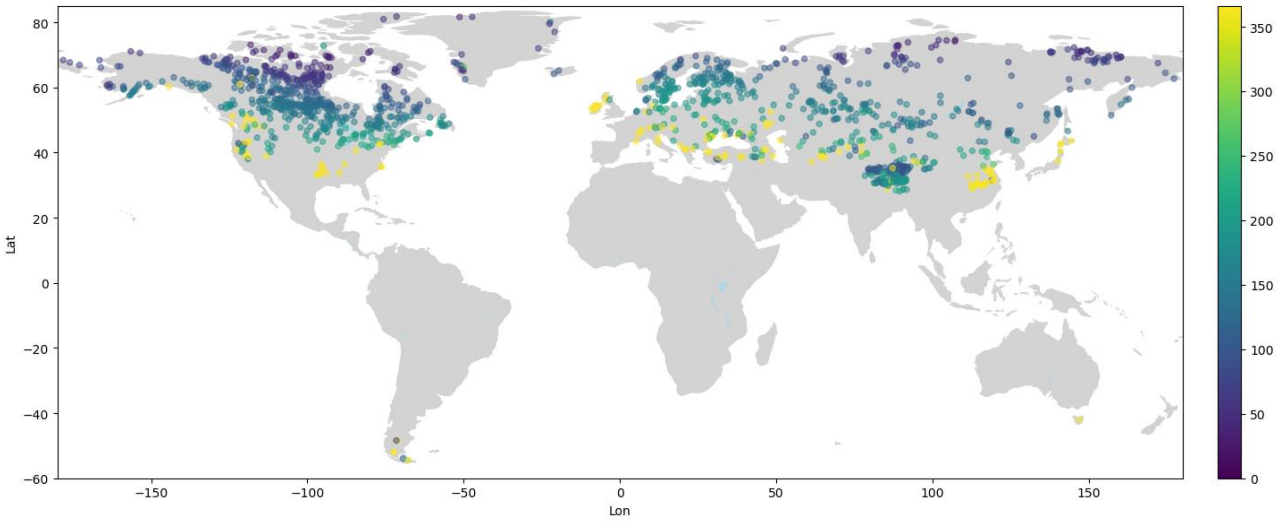


Figure 93 Duration of ice-free period in days (one dot per lake). Start of ice formation and end of melting are based on the LIC ice climatology, using the 3% threshold in the ice cover range.

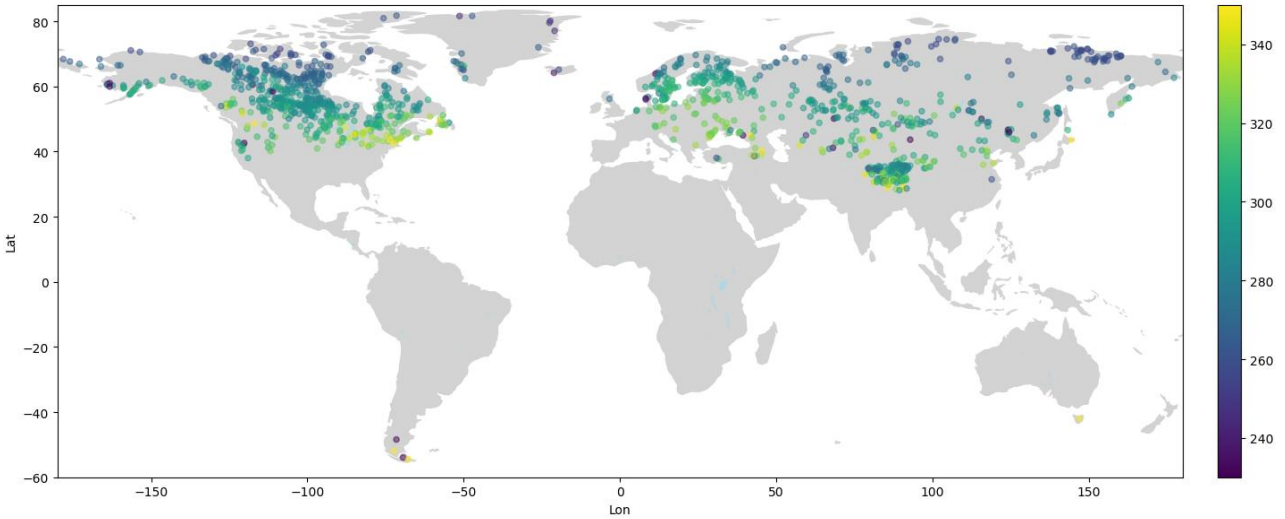


Figure 94 Start of ice formation in terms of day of the year based on the LIC ice climatology.

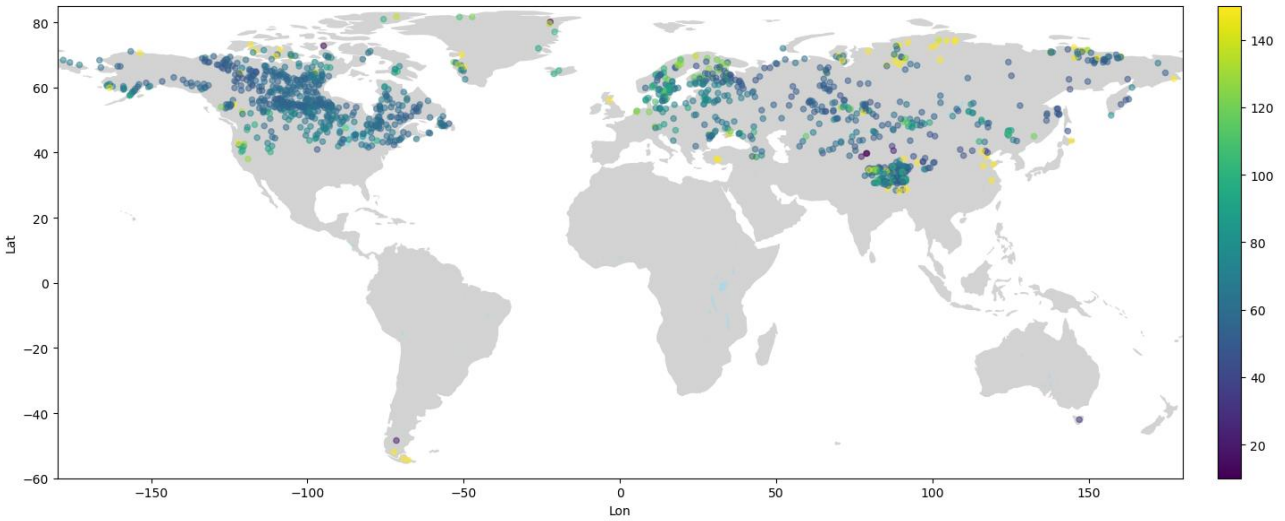


Figure 95 Start day of melting in terms of day of the year based on LIC water climatology.

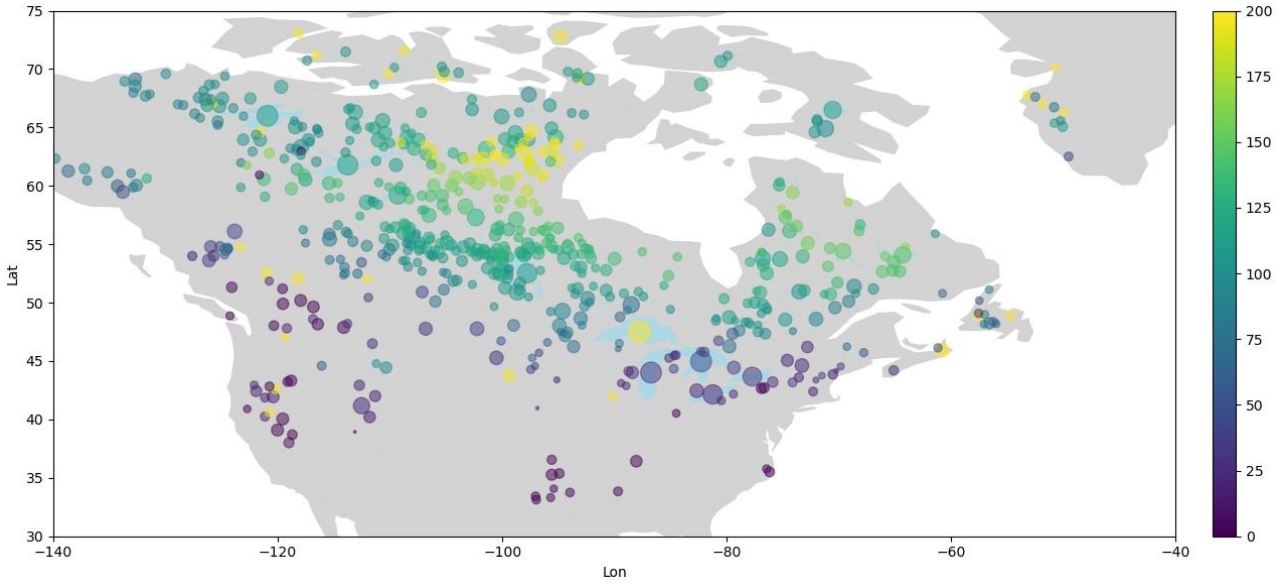


Figure 96 Duration of full ice cover in days in North America. Values are calculated from LIC water count climatology as the difference between end of ice formation and beginning of melting.

Appendix A - List of acronyms

AATSR	Advanced Along Track Scanning Radiometer
AATSR	Advanced Along Track Scanning Radiometer
AERONET-OC	Aerosol Robotic NETwork – Ocean Color
AMI	Active Microwave Instrument
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
APP	Alternating Polarization mode Precision
ASAR	Advanced Synthetic Aperture Radar
ASLO	Association for the Sciences of Limnology and Oceanography
ATBD	Algorithm Theoretical Basis Document
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced very-high-resolution radiometer
BAMS	Bulletin of the American Meteorological Society
BC	Brockman Consult
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CDR	Climate Data Record
CDOM	Coloured Dissolved Organic Matter
CEDA	Centre for Environmental Data Archival
CEMS	Centre for Environmental Monitoring from Space
CEOS	Committee on Earth Observation Satellites
CGLOPS	Copernicus Global Land Operation Service
CIS	Canadian Ice Service
CLS	Collecte Localisation Satellite
CMEMS	Copernicus Marine Environment Monitoring Service
CMUG	Climate Modelling User Group
CNES	Centre national d'études spatiales
CNR	National Research Council of Italy
CORALS	Climate Oriented Record of Altimetry and Sea-Level
CPD	Communication Plan Document
CR	Cardinal Requirement
CRG	Climate Research Group
CSWG	Climate Science Working Group
CTOH	Center for Topographic studies of the Ocean and Hydrosphere
DOC	Dissolved Organic Carbon
DUE	Data User Element
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
ELLS-IAGRL	European Large Lakes Symposium-International Association for Great Lakes Research
ENVISAT	Environmental Satellite
EO	Earth Observation
EOMORES	Earth Observation-based Services for Monitoring and Reporting of Ecological Status
ERS	European Remote-Sensing Satellite
ESA	European Space Agency
ESRIN	European Space Research Institute
ETM+	Enhanced Thematic Mapper Plus
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAQ	Frequently Asked Questions
FCDR	Fundamental Climate Data Record

FIDUCEO	Fidelity and Uncertainty in Climate data records from Earth Observations
FP7	Seventh Framework Programme
GAC	Global Area Coverage
GCOS	Global Climate Observing System
GEMS/Water	Global Environment Monitoring System for freshwater
GEO	Group on Earth Observations
GEWEX	Global Energy and Water Exchanges
GloboLakes	Global Observatory of Lake Responses to Environmental Change
GLOPS	Copernicus Global Land Service
GTN-H	Global Terrestrial Network – Hydrology
GTN-L	Global Terrestrial Network – Lakes
H2020	Horizon 2020
HYDROLARE	International Data Centre on Hydrology of Lakes and Reservoirs
ILEC	International Lake Environment Committee
INFORM	Index for Risk Management
IPCC	Intergovernmental Panel on Climate Change
ISC	International Science Council
ISO	International Organization for Standardization
ISRO	Indian Space Research Organisation
JRC	Joint Research Centre
KPI	Key Performance Indicators
LEGOS	Laboratoire d'Etudes en Géophysique et Océanographie Spatiales
LIC	Lake Ice Cover
LIT	Lake Ice Thickness
LSC	Lake Storage Change
LSWT	Lake Surface Water Temperature
LWE	Lake Water Extent
LWL	Lake Water Level
LWLR	Lake Water Leaving Reflectance
MERIS	MEdium Resolution Imaging Spectrometer
MGDR	Merged Geophysical Data Record
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	MultiSpectral Instrument
MSS	MultiSpectral Scanner
NASA	National Aeronautics and Space Administration
NERC	Natural Environment Research Council
NetCDF	Network Common Data Form
NOAA	National Oceanic and Atmospheric Administration
NSERC	Natural Sciences and Engineering Research Council
NSIDC	National Snow & Ice Data Center
NTU	Nephelometric Turbidity Unit
NWP	Numerical Weather Prediction
OLCI	Ocean and Land Colour Instrument
OLI	Operational Land Imager
OSTST	Ocean Surface Topography Science Team
PML	Plymouth Marine Laboratory
PP	Payment Plan
PRISMA	PRecursore IperSpettrale della Missione Applicativa
Proba	Project for On-Board Autonomy
QSR	Quarterly Status Report
R	Linear Correlation Coefficient
RA	Radar Altimeter
RMSE	Root Mean Square Error

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