Consortium Members

D4.1: Product Validation and Intercomparison Report

Reference: CCI-LAKES-0009-PVIR

Issue: 1.2

Date: Jul. 14, 20

esa
European Space Agency

lakes cci

Version history:			
Issue:	Date:	Reason for change:	Author
1.0	30 March 2020	Initial Version	PML: Xiaohan Liu, Stefan Simis CLS: Beatriz Calmettes H2O: Claude Duguay SERTIT: Hervé Yésou
1.1	30 April 2020	Adding Chapter LSWT and minor corrections	UoR: Chris Merchant CLS: B Calmettes / B Coulon
1.2	14 July 2020	Revision following ESA Review Adding results and minor corrections in LWE part	ICUBE-SERTIT : H. Yésou NORCE : E. Malmes TRE-Altamira : P.Blanco Legos : JF Cretaux CLS: B Calmettes / B Coulon

People involved in this issue:			**Signature**
Authors:	Stefan Simis, Xiaohan Liu	Plymouth Marine Laboratory	
	B. Calmettes	CLS	
	Hervé Yésou	SERTIT	
	Chris Merchant,	University of Reading	
	Claude Duguay	H20 Geomatics	
Internal review:	Stefan Simis B. Calmettes	Plymouth Marine Laboratory CLS	Stefan Simis Digitally signed by Stefan Simis Date: 2020.10.05 12:52:11 +01'00'
Approved by:	B. Coulon	CLS	Bruno Coulon Signature numérique de Bruno Coulon Date : 2020.10.05 16:51:17 +02'00'
Authorized by:	C. Albergel	ESA	Clement Albergel Digitally signed by Clement Albergel Date: 2020.10.05 21:34:42 +01'00'

Distribution:		
Company	**Names**	**Contact Details**
ESA	C. Albergel P. Cipollini	clement.albergel@esa.int Paolo.Cipollini@esa.int
BC	K. Stelzer	kerstin.stelzer@brockmann-consult.de
CLS	B. Coulon B. Calmettes	bcoulon@groupcls.com bcalmettes@groupcls.com
CNR	C. Giardino	giardino.c@irea.cnr.it
GeoEcoMar	A. Scrieciu	albert.scrieciu@geoecomar.ro
H2OG	C. Duguay	claude.duguay@h2ogeomatics.com
LEGOS	J.F. Crétaux	jean-francois.cretaux@legos.obs-mip.fr
NORCE	E. Malnes	eima@norceresearch.no
PML	S. Simis	stsi@pml.ac.uk
SERTIT	H. Yésou	herve.yesou@unsitra.fr
TRE-ALTAMIRA	P. Blanco	pablo.blanco@tre-altamira.com
UoR	C. Merchant L. Carrea	c.j.merchant@reading.ac.uk l.carrea@reading.ac.uk
UoS	A. Tyler E. Spyrakos	a.n.tyler@stir.ac.uk evangelos.spyrakos@stir.ac.uk

List of Contents

Reference documents

RD- 1: Algorithm Theoretical Basis Document
CCI-LAKES-0024-ATBD_v1.2

RD- 2: Product Validation Plan
CCI-LAKES-0030-PVP_V1.1

RD- 3: Product User Guide
CCI-LAKES-0029-PUG_V1.0

1. Overview

This document contains product validation results for the first version of the Lake_cci dataset. An update will be issued with the next version. The Lakes_cci project has three stages of product validation:

- Validation of individual thematic variables based on direct comparison between remote sensing products and in situ data or other remoting sensing datasets
- Consistency between these variables through five use cases
- Feedback from users of the data set

The purpose of this document is to summarize the results of the different thematic products (LWL, LWE, LSWT, LIC and LWLR) according to the first activities described in the Product Validation Plan (RD- 2).

2. Lake Water Level (LWL) parameter

2.1. Data description

Lake Water Level is the measure of the absolute height of the reflecting water surface beneath the satellite with respect to a vertical datum (geoid) and expressed in metres. The time series has been computed from multiple altimetry satellites since late 1992 to 2018 inclusive. The time periods used for each satellite/instrument are provided in Table 1 but may vary from one lake to the other, depending on the orbits of the satellites with respect to the location of the lake All missions has been used when the data were available and valid, however during tandem overlapping phase (TOPEX/Poseidon/Jason1, Jason1/jason2, Sentinel3A/Sentinel3B) we always test what is the most precise solution before choosing the data used in the time series. Therefore, from one lake to another one, the month of the year used for a given mission may change.

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

Satellite	Instrument	Time Period
TOPEX/Poseidon (T/P)	Poseidon-1	08/1992 – 01/2002
Jason-1	Poseidon-2	12/2001 – 07/2013
Jason-2	Poseidon-3	06/2008 – 10/2016
Jason-3	Poseidon-3B	01/2016 – present
Envisat	Radar Altimeter (RA-2)	Mission: 03/2002 – 10/2010
Cryosat-2	SAR interferometric Radar Altimeter (Siral)	04/2010-2015
SARAL	AltiKa	02/2013 – 07/2016
Sentinel-3a	SRAL	02/2016 – present

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

2.2. Comparison methods

2.2.1. Comparison with in situ data

External in situ data are useful to assess the quality of the LWL products. The comparison with these products, using different datums and different dates, is not straight forward. However, it provides information on the product precision and accuracy. The list of datasets used is provided in Table 2

Table 2. In situ databases for LWL validation

Dataset name	Description
Hydrolare	The International Data Centre on Hydrology of Lakes And Reservoirs provides data on mean monthly water level of nearly 1200 water bodies. The Centre operates under the auspices of WMO and a detailed protocol developed by the International Steering Committee of the Centre and agreed by WMO.

Dataset name	Description
Hidricos Argentina	The database base of Hidricos Argentina provides in-situ data on national rivers and lakes.
USGS	The database of US Geological Survey provides in situ data
US Army	The 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 Office Canada	This database contains in situ historical hydrometric data in Canada.

2.2.2. Dedicated field work

The second comparison method is the base on 15 years of the dedicated field work in the framework of satellite altimetry cal/val programmes over lake Issykkul.

2.3. Description of work

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

Drift can subsequently be adjusted if it is observed. Root-mean-square differences of unbiased time series are calculated, for the complete timeseries and for the Jason 3 and Sentinel 3A missions.

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

2.4. Result analysis

2.4.1. Comparison with Hydrolare

Thanks to the collaboration with the International Data Centre on Hydrology of Lakes and Reservoirs, Hydrolare, the information on ten lakes was provided in a monthly time step. Two indicators were estimated: Bias and RMS (Table 3). These values were estimated for each lake for the full time series of nearly 30 years including data from all missions. Since 2016, the indicators were separately evaluated for Jason 3 and Sentinel 3A missions. In some cases, the bias is high (up to 2.3m) and suggest that, even if the precision is good, the accuracy of the LWL products should be considered carefully by users. Appendix A contains the figures corresponding to each lake.

Table 3. Hydrolare LWL comparison

Lake Name	Multi satellite			Jason 3 (since2016)		Sentinel 3A (since 2016)	
	Time period	Bias (cm)	RMS(m)	Bias (cm)	RMS(m)	Bias (cm)	RMS (m)
Baikal	1992/09 - 2015/12	0,975	0,093	-	-	-	-
Bratskoye	1992/09 - 2015/12	74,302	0,275	-	-	-	-
Caspian	1992/09 - 2016/12	30.845	0.062	36.020	0.052	34.700	0.050
Issykkul	1992/09 - 2017/12	-232.984	0.045	-231.098	0.017	-231.252	0.017
Khanka	2000/01 - 2018/12	101.989	0.19	-	-	-	-
Kuybyshevskoye	1992/09 - 2018/12	26.550	0.228	29.802	0.082	23.132	0.207
Ladoga	1992/09 - 2018/12	-3.727	0.053	-6.217	0.026	-6.387	0.025
Onega	1992/10 - 2018/12	39.7111	0.059	34.536	0.038	34.837	0.042
Rybinskoye	1992/09 - 2014/12	10.122	0.178	-	-	-	-
Superior	1992/09 - 2017/12	-59.093	0.042	-62.369	0.011	-62.495	0.013

The rms value for both missions Jason 3 and Sentinel 3B is very similar. This value is lower for the last two missions than for the overall period, indicating a better estimation of the LWL for the current missions. Concerning the bias, it's a point of system reference for the measurement. We are therefore more interested in the variability (Appendix A) and the correlation indicated by the Pearson coefficient.

2.4.2. Comparison to Hidricos Argentina

The information concerning the variation on the Water Lever for lake Argentino and lake General Carrera were obtained online from the Base de datos Hidrologica integrada (BDHI): bdhi.hidricosargentina.gob.ar. For those lakes two indicators were evaluated (Table 4): the RMS of the variations and Pearson coefficient, indicating the correlation between time series.

Table 4. Hidricos Argentina LWL Comparison

Lake Name	Time period	RMS(m)	Pearson
Argentino	1992/10 - 2019/12	0.159	0.969
General Carrera	2008/09 - 2014/10	0.387	0.582

Since for General Carrea the comparison period ends before 2016, analysis of performance in current missions is not possible. Concerning Argentino lake, the Pearson coefficient is 0.995 and 0.991 for Jason 3 and Sentinel 3A missions respectively showing a very strong correlation between the time series. This is also indicated by the low value of the RMS for missions: 0.063m and 0.079m for Jason 3 and Sentinel 3A

Figure 1 and Figure 2 show the comparison of the LWL variation from CCI lakes and Hidricos Argentina for lakes Argentino and General Carrera respectively. For Argentino lake, there is a very good correlation between time series, also indicated by a Pearson coefficient near to 1. However, for General Carrera lake, this correlation is less strong. The accuracy of the LWL estimation depends on multiple factors: size and shape of the lake, location of the satellite track over the lake, mean depth of the lake. The General Carrea shows a high variability.

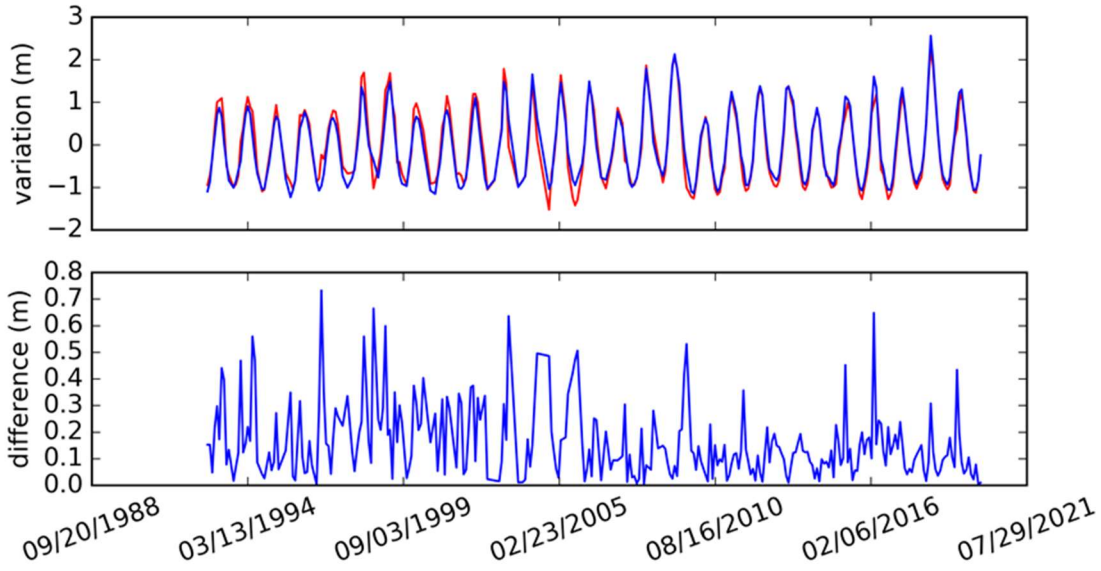


Figure 1. Hidricos Argentina comparison for the lake Argentino (red: Lakes_cci, blue: Hidricos Argentina)

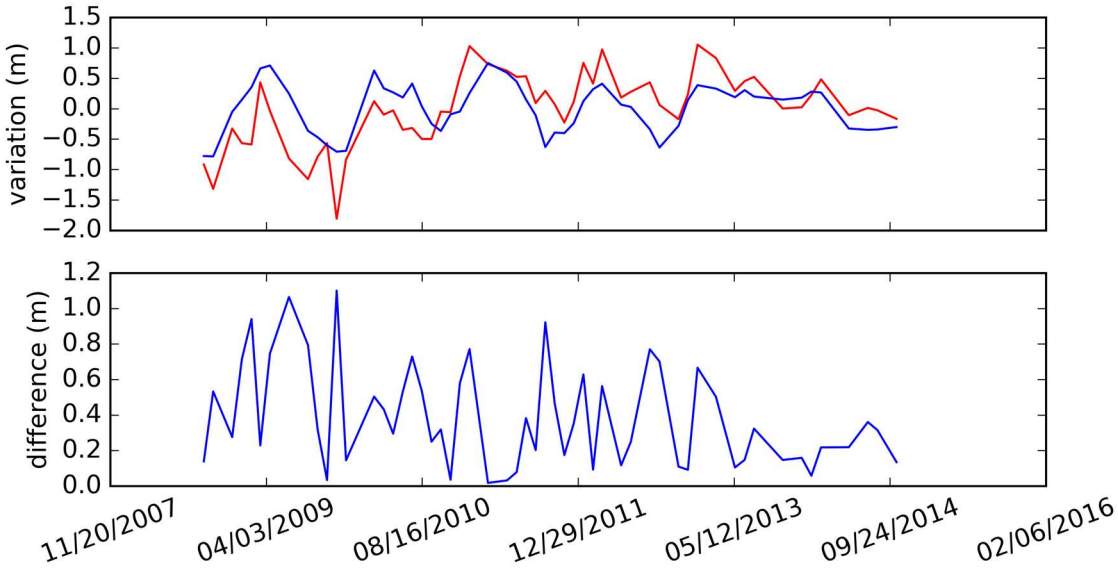


Figure 2. Hidricos Argentina comparison for the lake General Carrera ((red: Lakes_cci, blue: Hidricos Argentina))

2.4.3. Comparison to USGS

The information concerning the variation on the Water Lever for lakes Michigan and Des Bois were obtained online from the US Geological Survey database (USGS). For those lakes, as for the previous comparisons the RMS of the variations and Pearson coefficient were evaluated (Table 5). Figure 3 and Figure 4 show the comparison of the LWL variation from CCI lakes and UGS.

Table 5. USGS LWL Comparison

Lake Name	Time period	RMS(m)	Pearson
Des Bois	1992/09 - 2019/02	0.217	0.668
Michigan	1997/09 - 2019/10	0.084	0.965

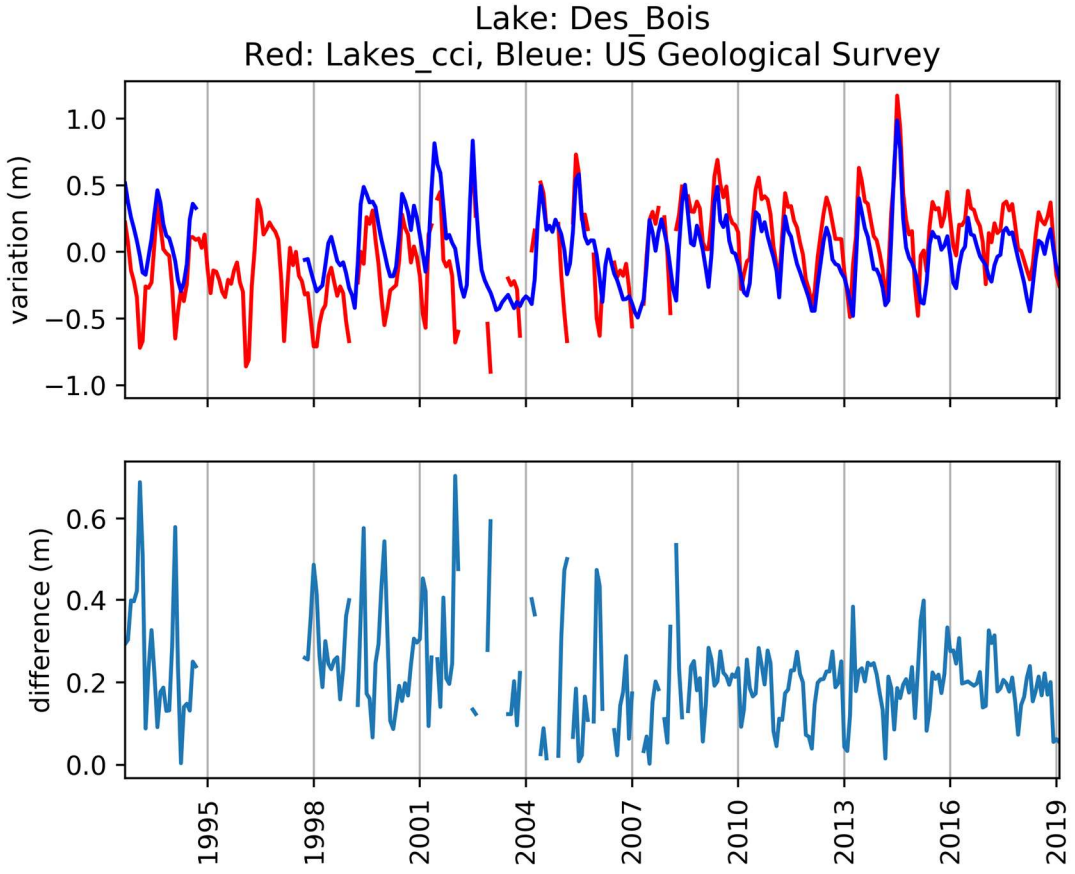


Figure 3. USGS comparison for the lake Des Bois

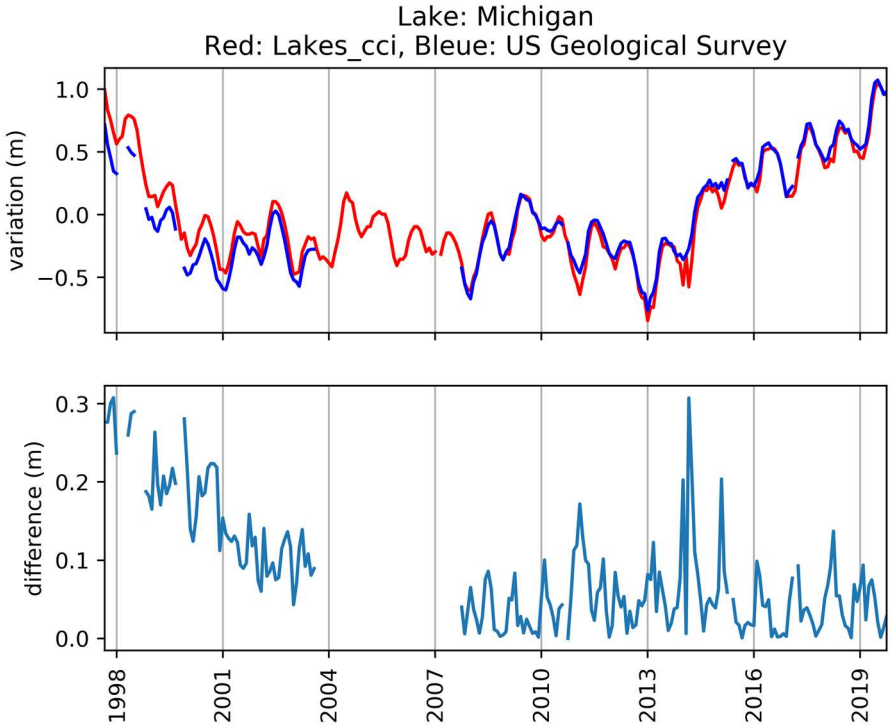


Figure 4. USGS comparison for the lake Michigan

It should be noticed that although in situ data are generally considered as the truth, which is valid most of the time, they may also present severe limitations. Some periods of time are not covered at all with in situ data. Some human errors in the data collection are also happening sometimes as we see it with the lake Onega. In other case like for the lake Michigan or the Caspian Sea, it exists several in situ instrumentations that provide different values of LWL. Sometimes this can be easily explained by local effect at high frequency (like the Seiche effect) sometimes it is less understandable. For example, we see that with two sources for the lake Michigan, the US army corps (Figure 5), and the USGS, the LWL present drifts and systematic disagreements. We can see with the USGS data between 1998 and 2004 when compared to the data of US army. In the first case the comparison with the altimetry shows big disagreements while in the second case the correlation and the RMS are much better.

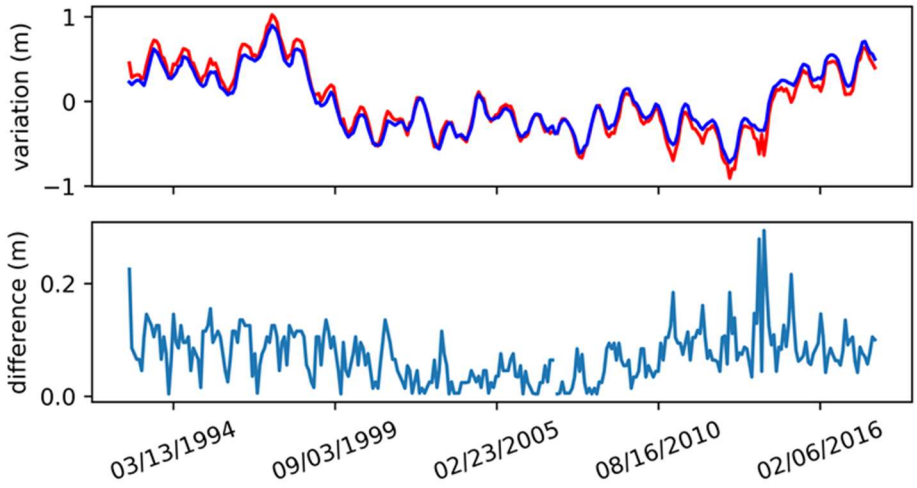


Figure 5. US Army comparison for the lake Michigan (red: Lakes_cci, blue: Hidricos Argentina)

2.4.4. Comparison to Canadian Water office

The Water Office of Canada contains information on the water levels and flood for multiple Canadian lakes and rivers at several time resolutions. Monthly historical data for 19 lakes included in first version of the CCI lakes dataset were obtained on-line: https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html.

The three indicators used with the previous in situ datasets: Bias, RMSE and Pearson Coefficient were evaluated. Figure 6 and Figure 7 show respectively the Pearson coefficient and RMS value for the 19 lakes compared. Most of them have a high Pearson coefficient showing a good time series correlation. Appendix B contains the figures of timeseries, variation and unbiased absolute difference for each lake. For the lakes with low value of Pearson coefficient, there a variety of reasons for this:

- In some cases, as for the lakes Aylmer or Caribou, there is small amount of in situ data
- In other case, as for the lakes Great Slave or Williston, the altimetric level value couldn't be estimated
- Some outliers, as for the Ontario lake, will affect the correlation between timeseries

In these cases, the low Person coefficient does not actually represent a poor correlation between time series.

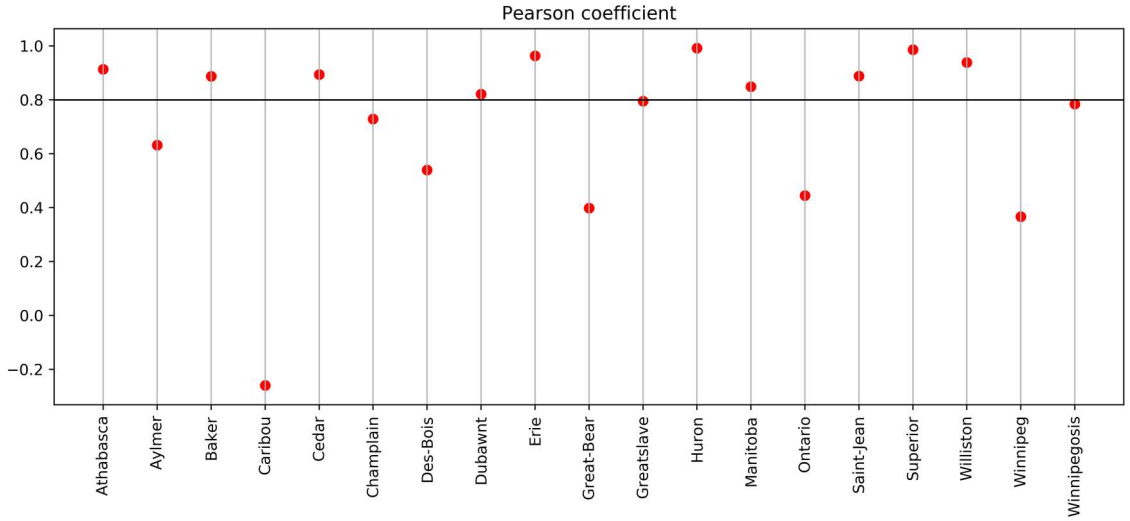


Figure 6. Pearson coefficient CCI Lakes - Water Office Canada.

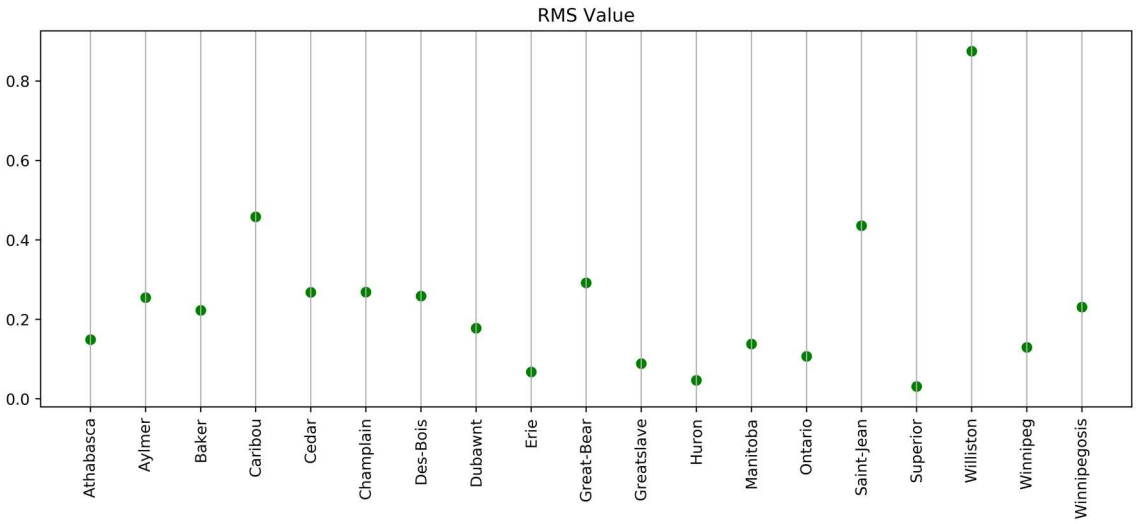


Figure 7. RMS Value. CCI Lakes – Water Office of Canada

2.4.5. Field work experiments

Figure 8 shows an example of LWL altimetry measurements with Sentinel 3 along two tracks (666 and 707) against in-situ measurements over Issykkul lake. There is an excellent correlation between both series (99%) and a low value of RMS (Figure 9)

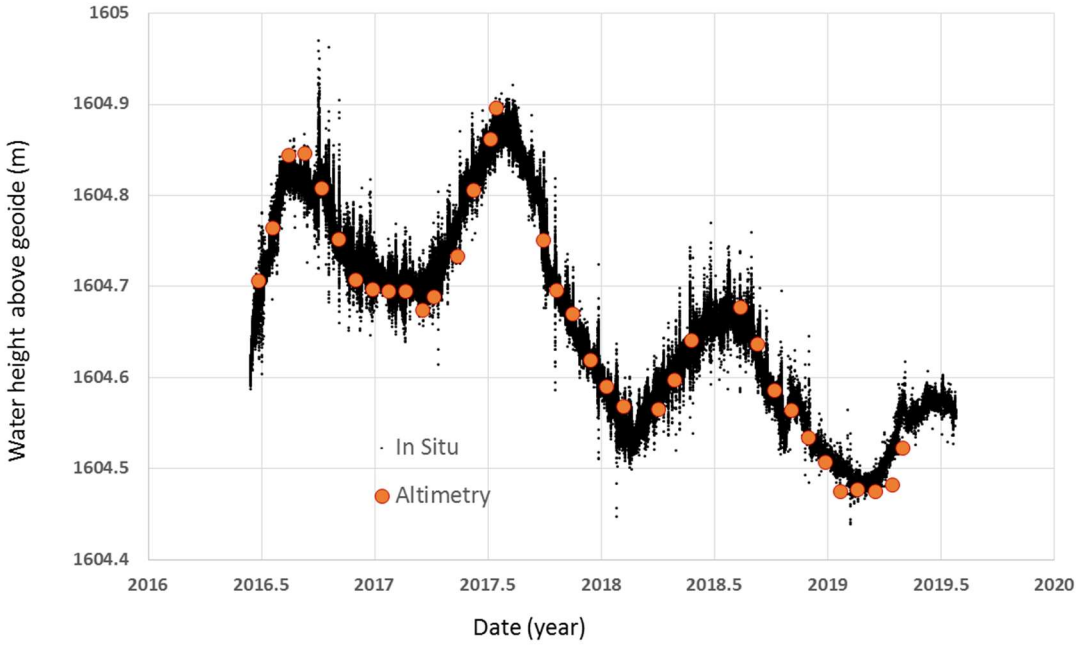


Figure 8. Sentinel 3 vs In-situ measurements for Issykkul Lake

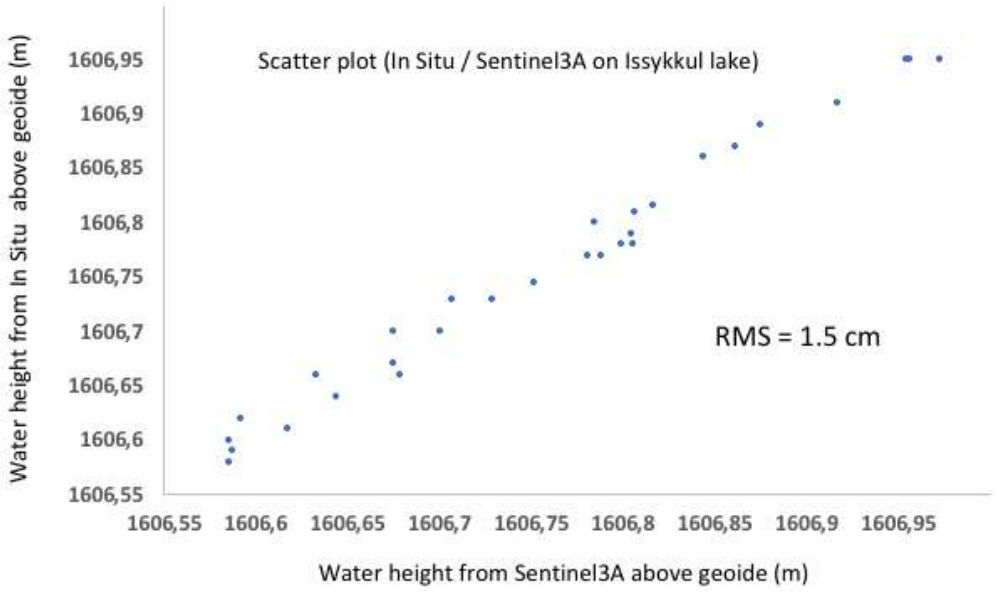


Figure 9. Correlation between Sentinel 3 vs In-situ measurements for Issykkul Lake

2.5. Conclusions and recommendations

We have validated the Lake Water Level, one of the lakes__cci products by comparing the timeseries to multiple external datasets of in situ measurements. Globally, there is a very good correlation with external datasets from different organisations with data in different regions of the world.

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

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

We would like to thank Prof. Valery Vuglinskiy (State Hydrological Institute, 23, 2nd Line, Vassilievsky Island, 199053, St Petersburg, Russia) for having provided us with in-situ data from Hydrolare lakes.

3. Lake Water Extent (LWE) parameter

3.1. Data description

3.1.1. Optical Data

Landsat 5 and 8 images and Sentinel2 time series have been exploited to derived LWE over a set of test site's lakes.

It has to be noticed that the rules for accessing Sentinel2 imagery have changed during the project. At the beginning most of the imagery was accessible directly on-line. At the present time only the recent images are accessible whereas the historical ones, i.e. the 4 last years are off-line and must be requested. When exploiting the ESA Sc Hub, this procedure is very time consuming, for example to access and download 23 products over the Argentino lakes, it took more than 5 days.

3.1.1.1. Landsat 5 TM TOA Reflectance:

Landsat 5 TM calibrated top-of-atmosphere (TOA) reflectance were exploited. Calibration coefficients are extracted from the image metadata. See Chander et al. (2009) for details on the TOA computation.(Earth Engine Data Catalog).

Table 6: Landsat 5 wavelengths and resolution

Bands

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

3.1.1.2. Landsat 8 TOA Reflectance

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

Table 7: Landsat 8 wavelength and resolution

Bands

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

3.1.1.3. Sentinel-2 L1C/L2A

3.1.1.3.1. Sentinel 2 Radiometric Resolutions

Table 8: Landsat 8 wavelength and resolution

Bands

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

More details can be obtained at ESA: Radiometric.

3.1.1.3.2. Level-1C

The Level 2 Sentinel 2 images is not systematically produced all over the world. By the way it is Level 1C data that have been proceed. Level-1C product provides orthorectified Top-Of-Atmosphere (TOA) reflectance with sub-pixel multispectral registration. Cloud and land/water masks are included in the product.

More details can be obtained at ESA: Level-1C Processing for details.

3.1.1.3.3. Level-2A

Level-2A product provides orthorectified Bottom-Of-Atmosphere (BOA) reflectance with sub-pixel multispectral registration. A Scene Classification map (cloud. cloud shadows. vegetation. soils/deserts. water. snow. etc.) is included in the product.

More details can be obtained at ESA: Level-2A product for details

3.1.2. SAR data

The SAR data used to calculate the LWE are Sentinel-1 images acquired in the Level-1 Interferometric Wide Swath (IWS) mode. NORCE has employed these data in Ground Range Detected (GRD) format while TRE-Altamira has employed Single Look Complex (SLC). GRD images contain the detected amplitude and multi-looked to reduce the impact of speckle. SLC images preserve phase information and are processed at the natural pixel spacing. IW mode is a dual-pol acquisition mode. In this case, images are acquired in both VH and VV polarization. Except for some particular cases, the SRTM DEM has been used for geocoding purposes.

Table 9: Sentinel-1 employed data spatial resolution

Mode	Resolution rg x az	Pixel spacing rg x az	Number of looks	ENL
GRD IW	20x22 m	10x10 m	5x1	4.4
SLC IW	2.7x22 m to 3.5x22 m	2.3x14.1 m	1x1	1

A database of Envisat ASAR WSM data is available which has been also used for some lakes. ASAR WSM data has in general coarser spatial resolution, and only one polarization so the quality is in general poorer than for S1 data.

3.1.3. Exogeneous database exploited as inputs

3.1.3.1. Global Surface Water database

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

3.1.3.2. Lakes contours database

The analysis is done at a given lake scale. To that matter a precise contour of the lake shore is requested, allowing to decrease processing time but more important to limit artefacts related to the lake's environment.

Whereas exploitation of Medium or low resolution satellite imagery for LWST or LWSR, is based on the analysis of "pure" water bodies, considering that parameters retrieval is done based on the lake AOIs plus a kilometric buffer, for the lake water extent, the investigation is focused on the much precise as possible shore line. This is not a simple limit; for lots of areas, related to water level increase, the water surface's expansion is observed on shoreline, but also on bordering wetlands.

So, for MR and LR satellite imagery exploitation, a relative rough definition of the Area of interest is sufficient, and the buffer application would correct some potential mistake. Then for exploitation of HR satellite imagery, such as Sentinel2, is requiring a precise definition of the AOIs, this can be done exploiting the CCI lakes AOIs

database, or the Hydrolakes database derived from the SRTM mission and containing 1,4 millions of lakes larger than 10 ha, (https://www.hydrosheds.org/pages/hydrolakes) but more often in the WP6, the AOIs were at least validated based on Sentinel 2 imagery acquired at different hydrological period, and when requested modified.

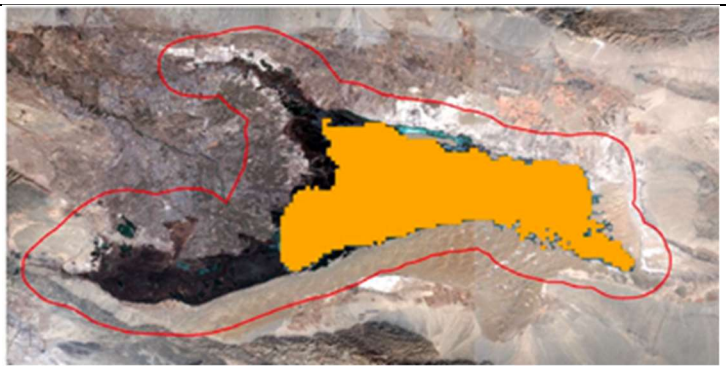
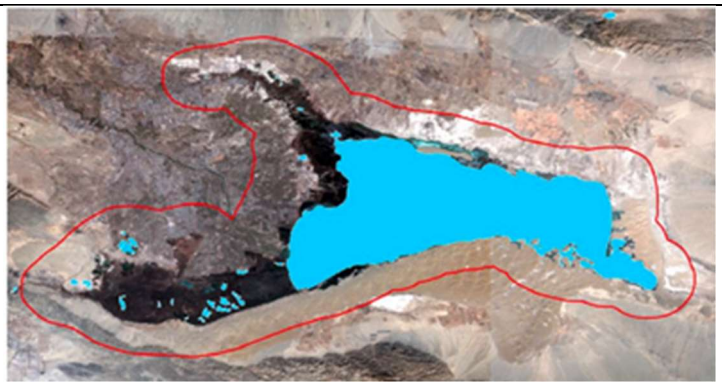
	Orange: CCI lake AOI
	Blue: Hydrolake AOI
	Sentinel 2 image acquired on the 2019-05-14 Presenting water inundating the North Western and South Western branch
	Red: AOIs defined and exploited in order to include the two "'arms" that correspond to wetlands more of less inundated due to water level increase.

Figure 10. Comparison of AOIs contours: case of Bosten lake (PR China).

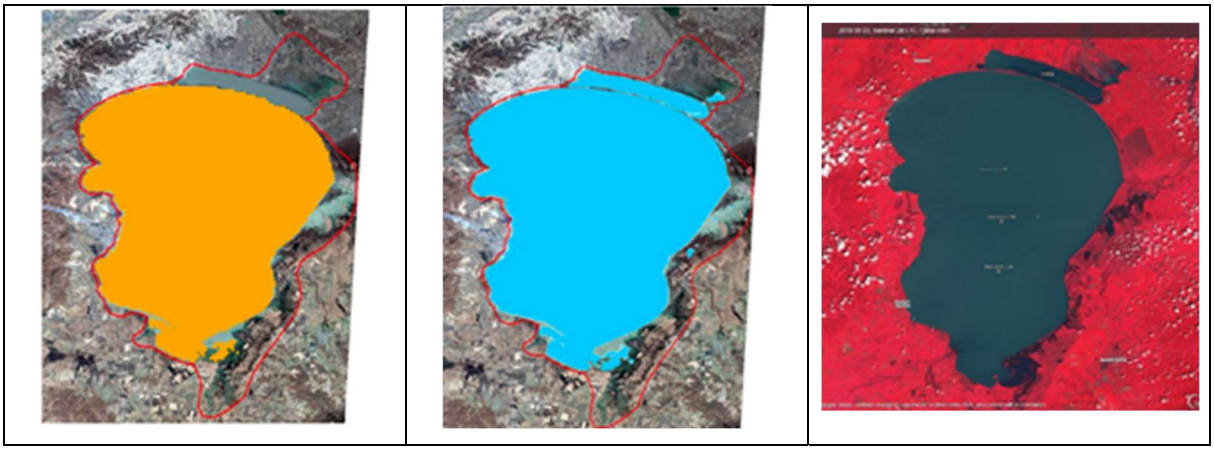


Figure 11. Comparison of AOIs contours: case of Khanka Lake (PR China-Russia) with in orange CCI Lakes AOIs, in blue, Hydrolake and in red exploited AOI for Sentinel2 exploitation.

When comparing, with a Sentinel2, 2019-08-03, it is well noticeable that the AOIs proposed by both Hydrolakes and CCI lakes database are too restrictive, as excluding the wetlands in the North and South East parts of the region.

3.2. Comparison of methods

3.2.1. Comparison of methods for optical sensors (VIS_NIR_SWIR)

The aim of this part is to compare lakes area and lakes vector extracted from sentinel-2 and Landsat images with different classifiers. Initially two none supervised approach: OTSU and K-Means and two supervised ones. SVM and Random forest have been applied on the data set over the lakes' test. Finally, for this comparison of the results from the optical processing approaches only OTSU SVM and Random forest are discussed.

3.2.1.1. Description of work

A selection of lakes with various characteristic, i.e. spectral behaviours of shallow waters, environment more or less arid, relief, presence of ice and/or snow etc.. (Table 10 and Table 11) have been done. The results obtained over these lakes based on different processing approaches have been compared.

More precisely, this analysis was done in three steps. In a first one, a relative long time series of images was selected for a first set of lakes. In a second time, a specific inter comparison was done on a short dense time series over Chad lake covering a field period survey. At least a third time, for consolidation of the approach, as well as investigated a better parametrization of the process, such as the Random Forest (RF) approaches, an additional set of lakes, with a relative low amount of images, 6 to 8, was selected in order to present different levels' state of theses lakes.

Table 10: Major characteristics of the Chad lake case study and the second set of the analysed lakes

	Altevatnet Norway	Colhue Argentina	Namtso China	Alakol Kazakhstan	Sassykkol Kazakhstan	Chilwa Malawi	Al Hamar Irak
Shadow water		XXX				X	XXX
ice	XXX		XXX	XXX	XXX		

	Altevatnet Norway	Colhue Argentina	Namtso China	Alakol Kazakhstan	Sassykkol Kazakhstan	Chilwa Malawi	Al Hamar Irak
snow	XXX		XXX	XXX	XXX		
Topographic position	XXX						
Local environment	XXX	XXX	XXX	XX	XX	XXX	XX
sunglint			XXX				
Floating vegetation						XXX	XXX
Lake dynamic	X	XXX	X	X	X	XXX	XXX

Table 11**:** Major characteristics of the Chad lake case study and the second set of the analysed lakes

	Chad	**Argentino Argentina**	**Bosten China**	**Khanka China Russia**	**Illmen Russia**	**Sary kamysh Turkménistan - Ouzbékistan**
Shadow water	X		X	X	X	
ice		XXX	XX	XXX	XX	
snow		XXX		XXX		
Topographic position		XXX				
Local environment	XX			X		X
sunglint						
Floating vegetation	XX		X	XX	X	
Lake dynamic	X		XX		XX	X

3.2.2. Validation of LWE derived from HR optical sensors based on VHR sensors

Validation of water extent is pure and great challenge by itself. Few methods can be investigated:

- Comparison of LWE with databases. There are lot of limitations, genes of the database, such as the resolution of the input data, the date/period, a LWE can change greatly from date to date, a year to another.
- Field comparison by surveying the water bodies limits walking along the shore with a GPS tracking, or using a boat or kayak to follow the shore. It is not always possible to walk around lake, all there are the question of the accessibility of the lake. In case of shallow water, what is the exact distance between the boat and the shore, a tens of meters (or more) that represents two to five/ten Sentinel pixel.
- Extract LWE from satellite of high and very high resolution.

The last approach is according to us the most promising. Therefore, it is not so convenient to be implemented, it requested pair of HR and VHR images acquired within a very short time. The data have to cover if possible,

the targeted lake as a whole, englobing the surrounding areas. Most of the time, and it is particularly the case with CCI lakes that are large lakes. So on in most of case only a part of the lake is covered by the two sensors and so what is the representativity of the covered area. And, of course, when VHR data have to be ordered it can be a costly approach. For this reason, an agreement with the CNES, French Space Agency was initiated in order to order a low coast VHR SPOT 6-7 or Pleaides imagery and share these data with the WP6 team. Therefore, the analysis of the catalogues was not so successful and this approach was abandoned. Hopefully in parallel, CNES was able to order VHR Pleiades images over large reservoirs in France, the Der and Orient lakes, and this during a dynamic period of infilling.

3.2.2.1. Location of the test areas

The Der and Orient lakes, located in the East of Paris, within the Champagne area region, are part of the Seine River flood management systems.

The Der lake is the largest artificial reservoir in France, with a surface around 48km2, for a maximal depth of 18m, the Orient water surface is 22 km2 for a maximal depth of 22 m. The functioning of the reservoir is the following:

Water is taken from the rivers, i.e. Marne River for Der lake, la Seine River for Orient lake, from November / December to June, thus filling the reservoir. From July to October, water is released to support the flow of rivers. As a result, water surfaces change considerably during the year, for the Der lake going from around 40 square kilometers during the high season, to less than ten square kilometers during the very low water period.

Figure 12. Location of the Der lake and Orient Lake.

3.2.2.2. Exploited data

The lakes are located within an Overlapping part of Sentinel 2 tracks, allowing up to 14 acquisitions by month. So, it was an ideal case to order VHR Pleaides imagery, knowing that the acquisition will by the way have at maximal one day of delay between the VHR and HR data.

Finally, two pairs of Pleiades HR data, 70 cm of spatial resolution, a panchromatic channel and 4 visible ones from blue to near Infrared channels, were acquired on the 30 of December, with a delay of one day with Sentinel2 and 6 of January 2020, same date as a Sentinel2 acquisition.

Figure 13. The Der Lake; as viewed by Pleiades on the 30 of December 2019, and on the 1 of January 2020

From the Pleiades and Sentinel2 data, LWE were extracted for each date and each reservoir based on a SVM approach.

3.2.2.3. Results

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

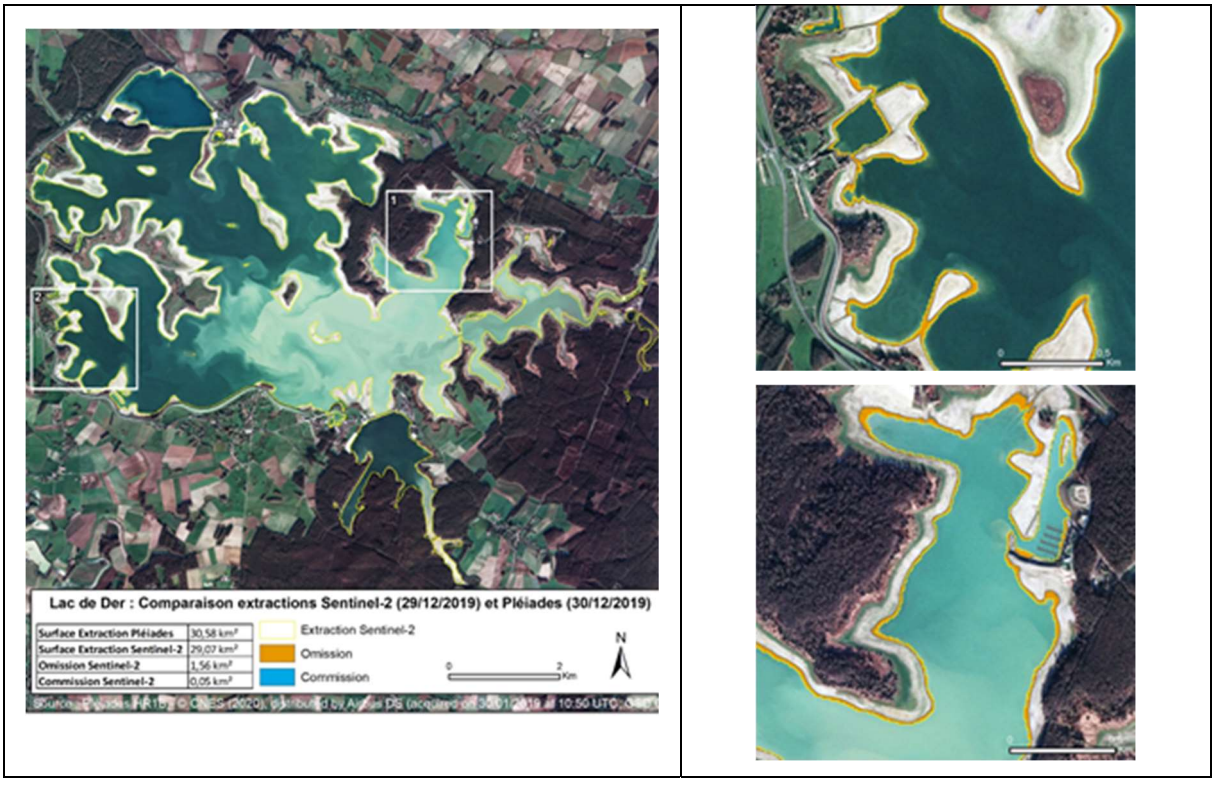


Figure 14. Omission and commission over lake de Der when exploiting VHR and HR images acquired with a 24h of delay

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

This case of study allows also to evaluate, in the context of infilling reservoir, the part of the 24h of delay between the two acquisitions.

When data are acquired the same day, the space occupied along the shore of the omission is very narrow, in fact the shoreline corresponds to a staircase, of swatooth's effects, alternating omission and commission pixel. An effect that is related in fact to the difference of spatial resolution.

Where, the LWE represent two stages of infilling, we observe a large omission belt around the lake shore. This belt in fact corresponds to the increase of the surface of water within one day. So, of course, what is seen as water on the Pleiades image, cannot be described as water on the Sentinel2 image acquired a day before.

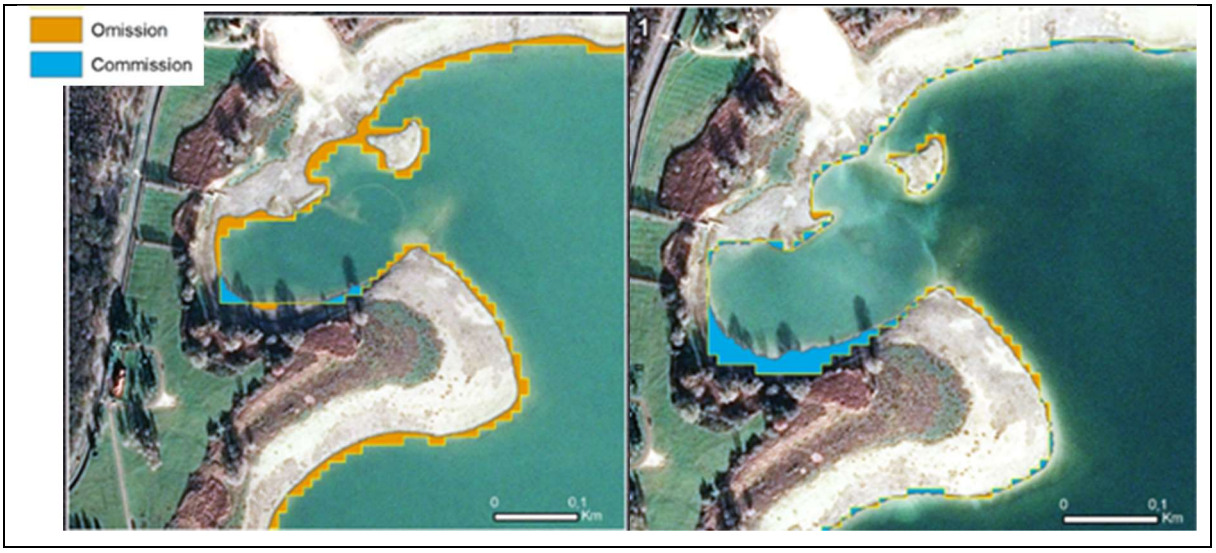


Figure 15. Comparison of LWE accuracy derived with one day of delay between the Sentinel2 and Pleaides acquisitions and the same day (right).

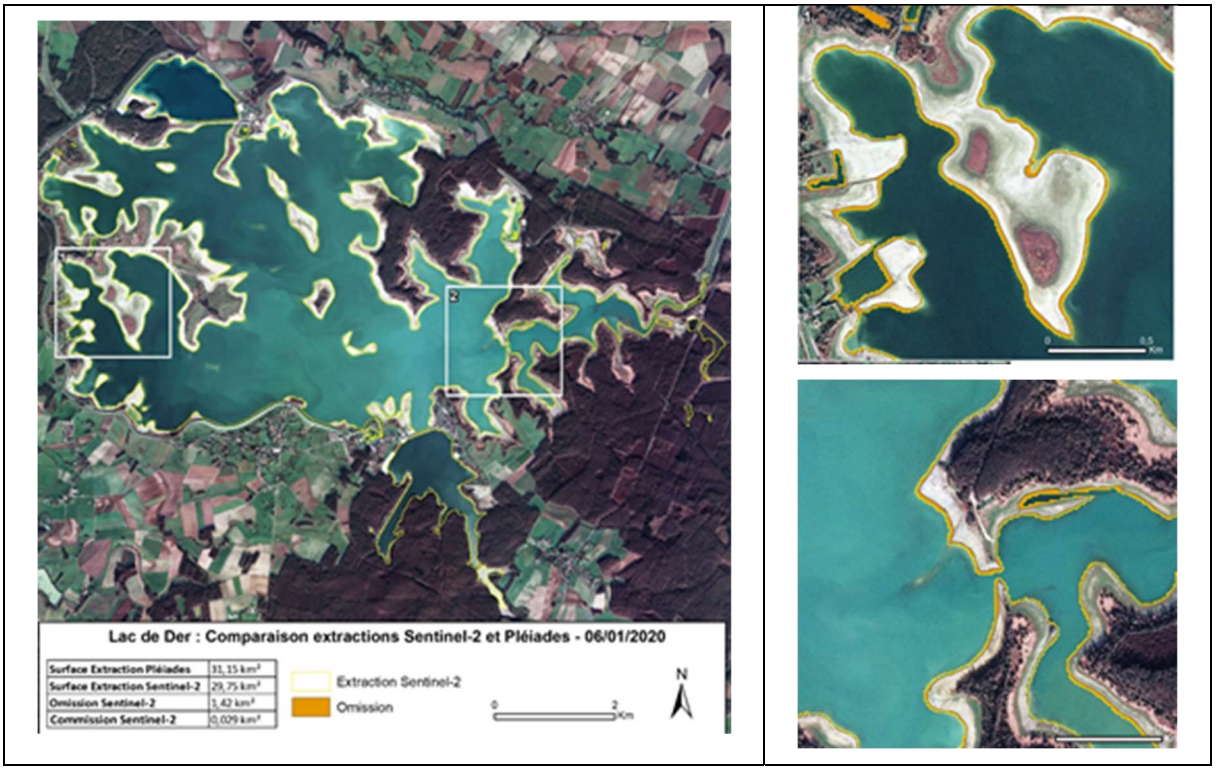


Figure 16. Omission and commission over lake de Der when exploiting VHR and HR images acquired quasi simultaneously.

3.3. CCI test sites result analysis

3.3.1. Results analysis for optical sensors-based approaches

As indicated, a first analysis was conducted lake by lake. An analysis of each important difference/gap from an approach to another one was done and commented.

3.3.1.1. Altevatnet Lake

Altevanet Lake is a narrow long lake, i.e. 2*38 km2, within an incisive valley is located rather north of Norway: 68°N. This Northern location induces the presence of snow/ice on the lake shores as well as relatively low solar position. The analysis of the observed gaps can be related to these characteristics.

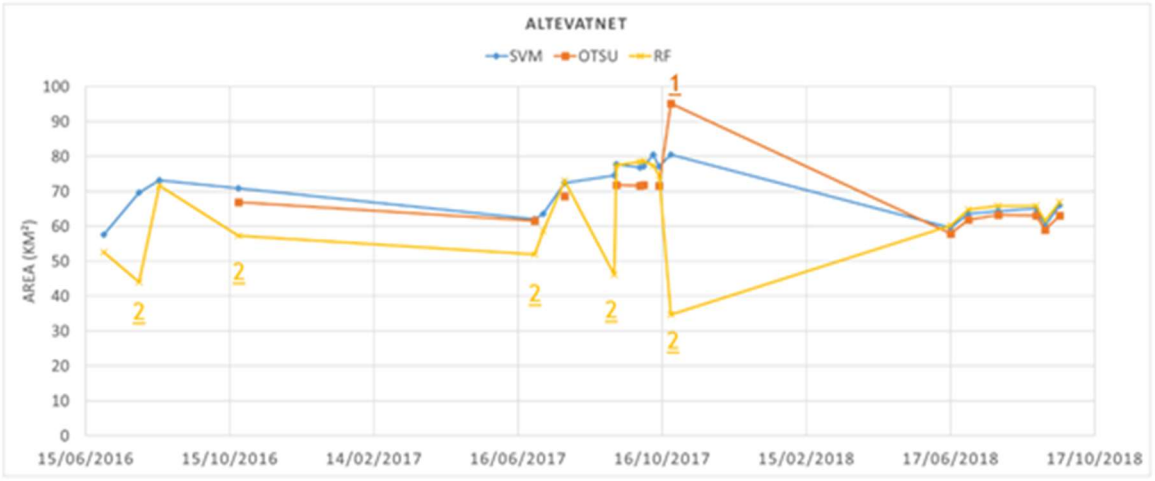
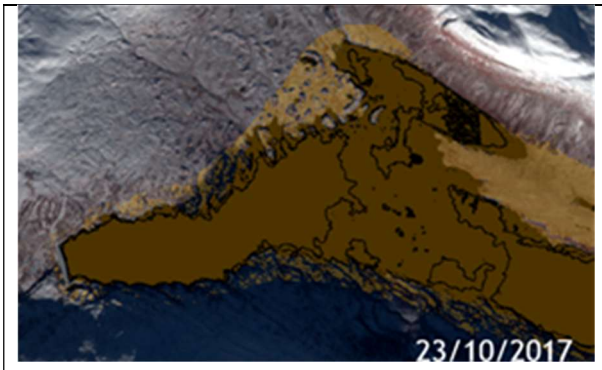


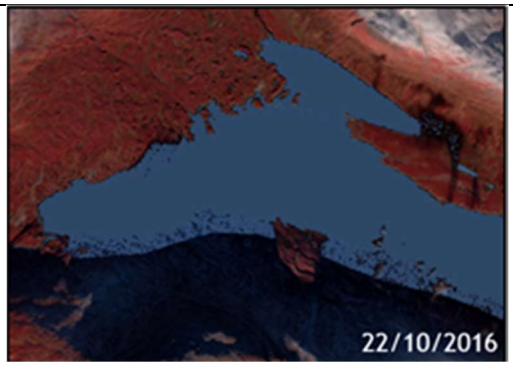
Figure 17. LWE derived from optical imagery based on different approaches over Altevatnet lake.

Observations based on the graphic and vector extractions analyse are the following:

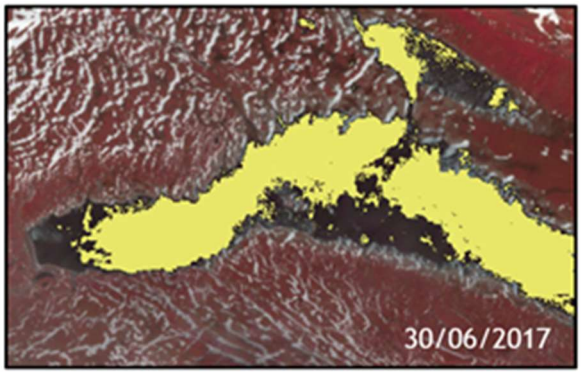
- On the S2 image of the 23/10/2017. OTSU's overestimation caused by snow.
- On the images of the 30/07/2016. 22/10/2016. 30/06/2017. 05/09/2017 & 23/10/2017 RF's generates an underestimation caused by low reflectance (lower than training image) and wisp of cloud.
- In some case, SVM classifies shadows into water's class. Same problem can also be observed when applying RF.

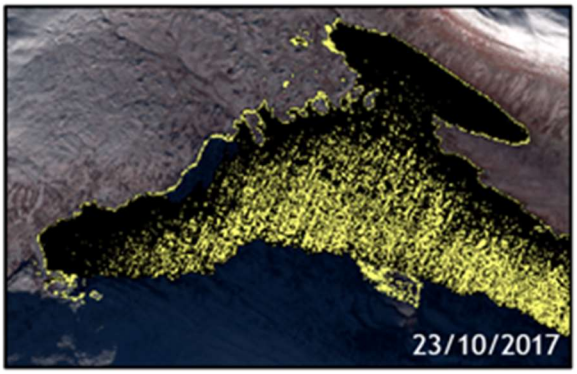


OTSU: commission snow/water



SVM: shadows classified as water





RF's generates an underestimation caused by low reflectance (lower than training image) and wisp of cloud

Figure 18. Omission and commission observed over Altevatnet lake

In addition, over the Altevatnet site, Otsu and Otsu Canny approaches were tested. It is appearing that the OTSU Canny is more restrictive on water, therefore it is also OTSU Canny that is more strongly influenced by shadows, snow, clouds.

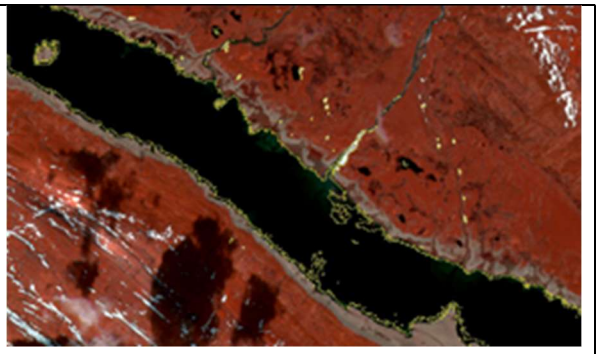
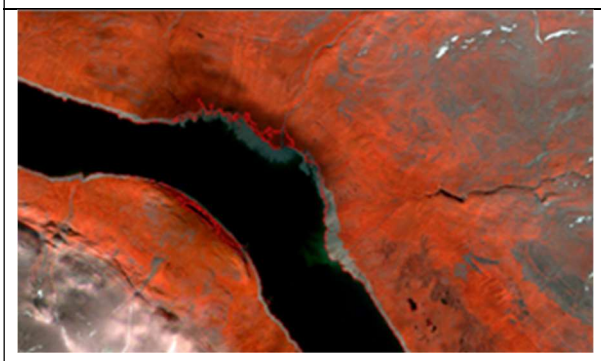
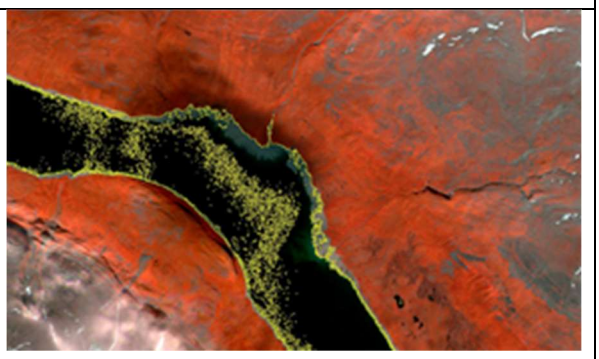
	
OTSU applied over a Sentinel2 image acquired the 17-06-2018 over Altevatnet lake	Canny-OTSU applied over a Sentinel2 image acquired the 17-06-2018 over Altevatnet lake
	
OTSU applied over a Sentinel2 image acquired the 27-07-2018 over Altevatnet lake	Canny-OTSU applied over a Sentinel2 image acquired the 27-07-2018 over Altevatnet lake

Figure 19. Comparison of Otsu and Canny-Otsu results.

3.3.1.2. Colhue Lake

Colhue Lake. It is located in a tectonic depression reshaped depression by wind and fluvial activities on the Western margin of the Patagonian plains and it is protected from Western rains by a North South a ridge of 600m. It is a shallow lake that knows fluctuations in terms of precipitations and rivers discharges resulting on LWE variations at least from 1998 to 2015. The analysis of the Sentinel 2 time series indicates that for the bear period, i.e. 2015-2017, the lake known a period of total dryness and of rapid infilling.

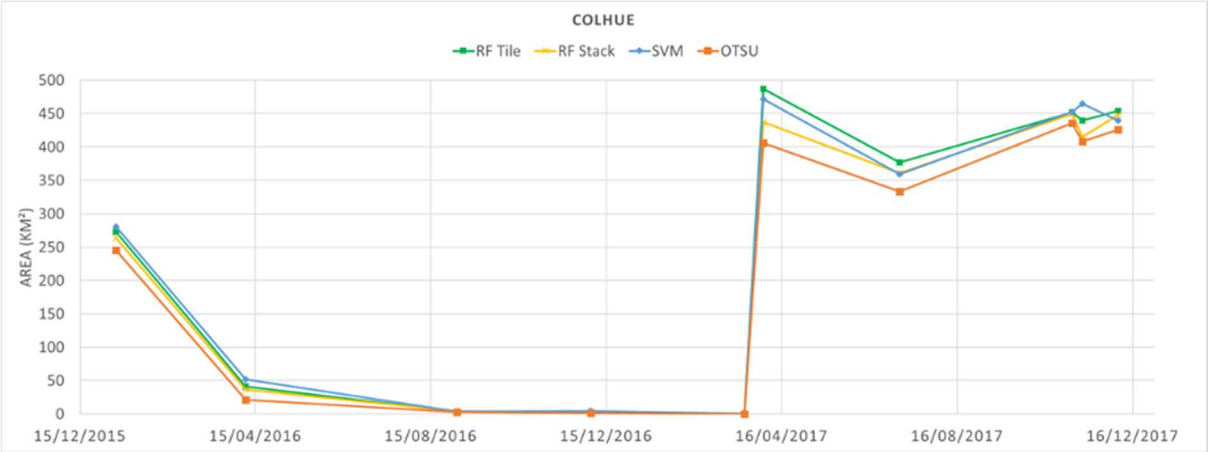
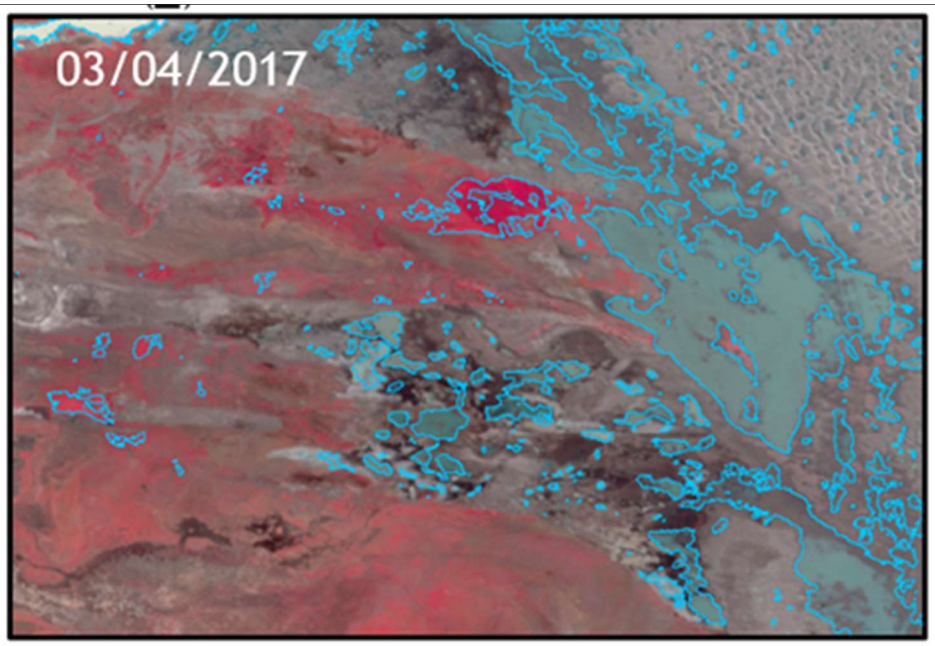


Figure 20. LWE derived from optical imagery based on different approaches over Colhue Lake.

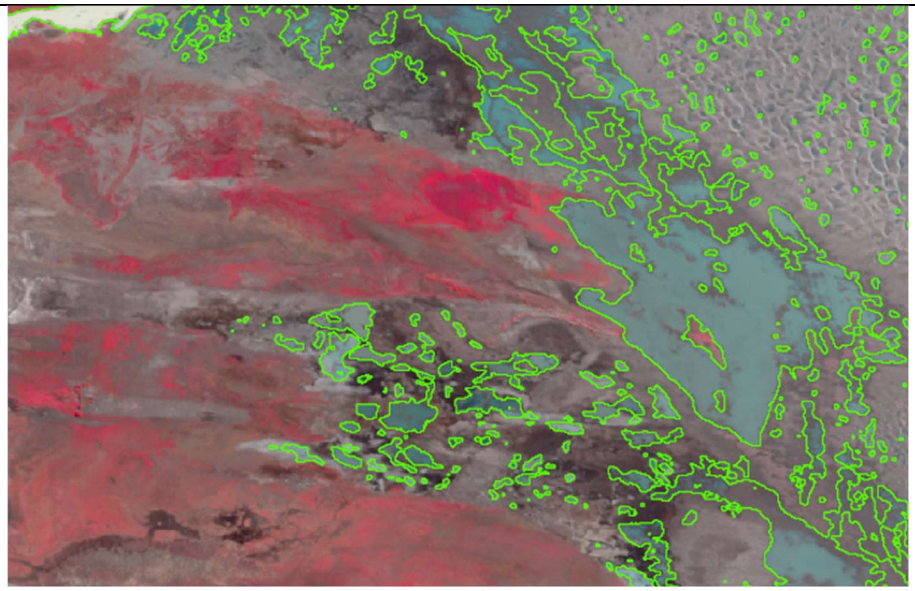
Over Colhue Lake, the difference processing's approaches provided very similar results

Observations based on the graphic and vector extractions are the following:

- Main differences between methods are caused by the misidentification of the water/non-water limits. SVM and RF classify muddy part of the lake/trickles of water.
- SVM classify also some vegetation as water surface. Not detected by RF.



Commission between vegetation and water, when applying SVM on Sentinel2 image



Vegetation not detected as water, when applying RF Tile on Sentinel2 image

Figure 21. Analysis and comparison between SVM and RF Tile approaches over vegetated areas bordering the Colhue Lake.

In addition, over this test site, Otsu and Otsu Canny approaches were tested. The obtained results are for this case very similar.

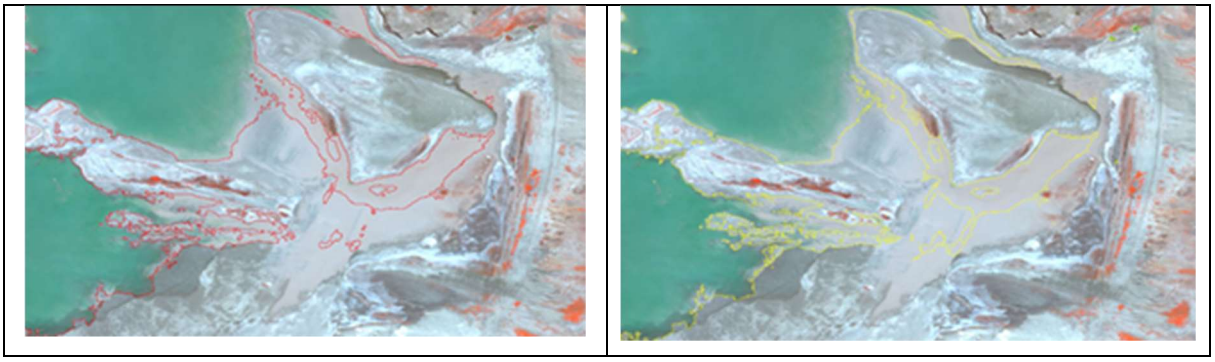


Figure 22. OTSU (left) and Canny-OTSU (rigth) applied on a Sentinel2 image acquired the 27/02/2019

3.3.1.3. Namtso Lake

Namtso Lake is a large water bodies of 1900 km2 located on the Tibetan Plateau. This lake knows a relative increase of water height and water extent (+600km2) from 1994 to 2000 and would have been relatively stable since then. It is located at a very high elevation, ie 4720 m, by the way the lake is covered by ice a long part of the year, from November to April, and the atmospheric effect are important.

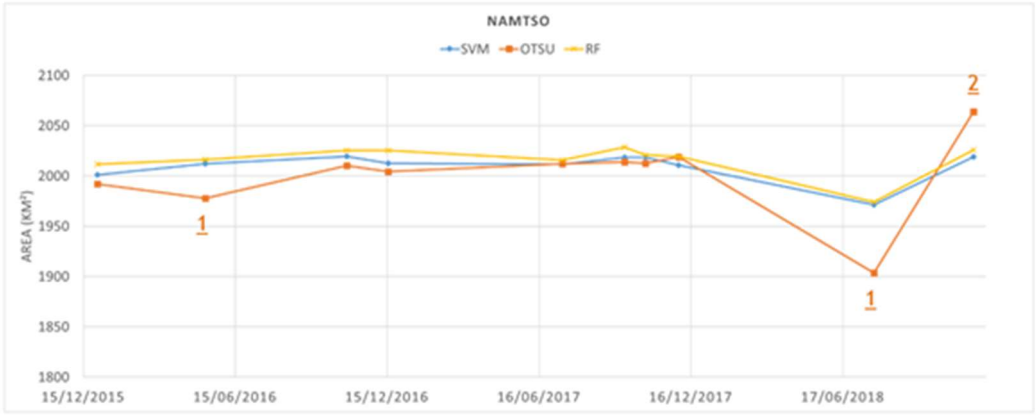


Figure 23. LWE derived from optical imagery based on different approaches over Namsto Lake.

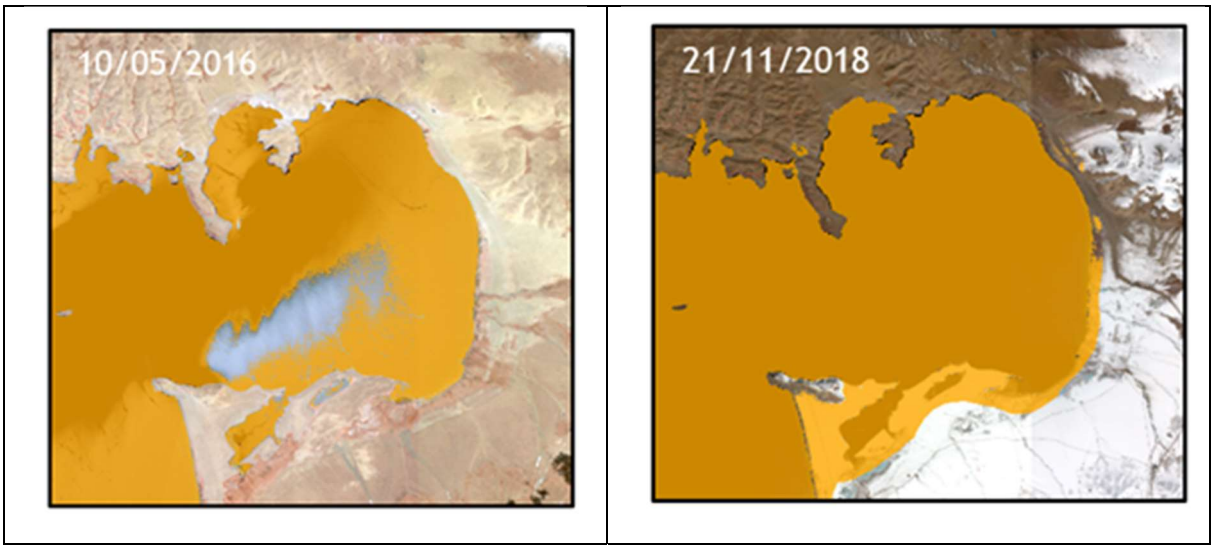


Figure 24. OTSU processing: under estimation du to sunglint and over estimation du to snow.

Observations based on the graphic and vector extractions analyse:

- On the images of the 10/05/2016 and of the 24/07/2018 OTSU's underestimation caused by sunlight.
- On the S2 image of the 21/11/2018, there are OTSU's overestimation caused by snow.
- SVM and RF have pretty much the same good results. Differences between methods are lower than 5%.

3.3.1.4. Sassykol Lake

Sassykol Lake is located at an altitude of 350 m in the southeast of the eastern province of Kazakhstan in the Balkhash-Alakol lowland between mountain systems of Zhetysu Alatau in the south. Tarbagatai in the north, and Barlyk in the east. It is the part of a complex hydro system formed by three successive lakes: Sassykol on the upstream part, the Kosharkol and in downstream the Alakol Lake. Sassykol lake communicates with the Alakol Lake through an extensive wetland surrounding the intermediate lake named the Kosharko Lake Sassykol is a shallow fresh-water lake which shores are gently sloping and densely covered with reeds. From November to the end of March-April, the lakes are covered by ice.

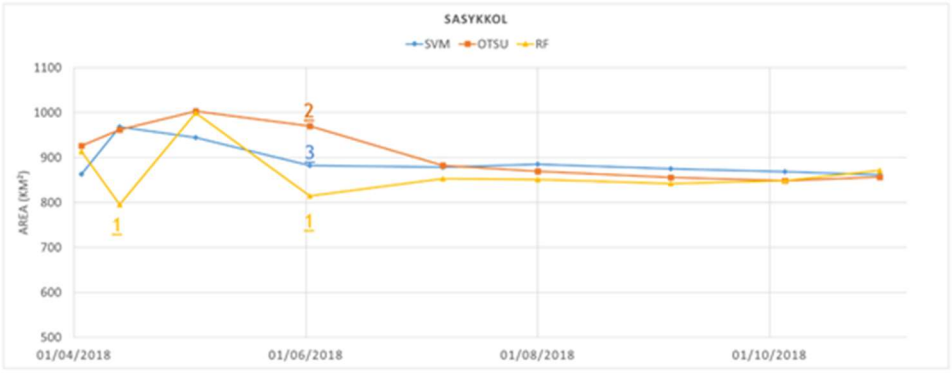


Figure 25. LWE derived from optical imagery based on different approaches over Sassykol lake.

1. Random Forest approaches (RF) involve an underestimation that could be caused by low reflectance (lower than training image) and wisp of clouds.
2. OTSU's approach generates an overestimation: it classifies vegetation and wisp of the cloud.
3. SVM: it shows a good classification. The boundaries of the water bodies are well marked, the vegetation is well distinguished.

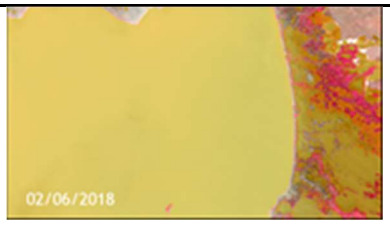
		
RF: underestimation related to sunglint	Otsu: over estimation; integrating as water, part of vegetation	SVM: good classification between water, vegetation

Figure 26. Artefacts observed over Sassykol lake exploiting different processing approaches.

3.3.1.5. Alakol Lake

Alakol Lake is an endorheic salted lake relatively deep, 45m, and large, more than near 3000 km2. The shores of Alakol Lake are rugged and an unstable coastal zone, with also large islands. As already indicated, from November to end of March-April, the lake can be covered by ice.

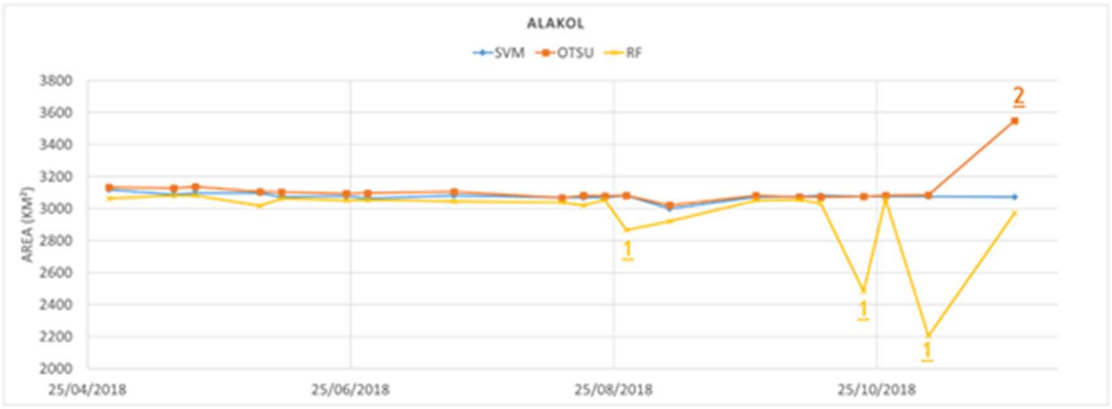


Figure 27. LWE derived from optical imagery based on different approaches over Alakol lake.

Observations based on the graphic and vector extractions analysis are the following:

- RF-underestimation caused by a wisp of cloud for the images from 28/08/2018, 22/10/2018, 06/11/2018 & 24/07/2018, indicated with (1) on the Figure 28.
- OTSU's approach presents an overestimation caused by snow, on the Sentinel2 image acquired on 26/11/2018, indicated with (2) on the Figure 28.

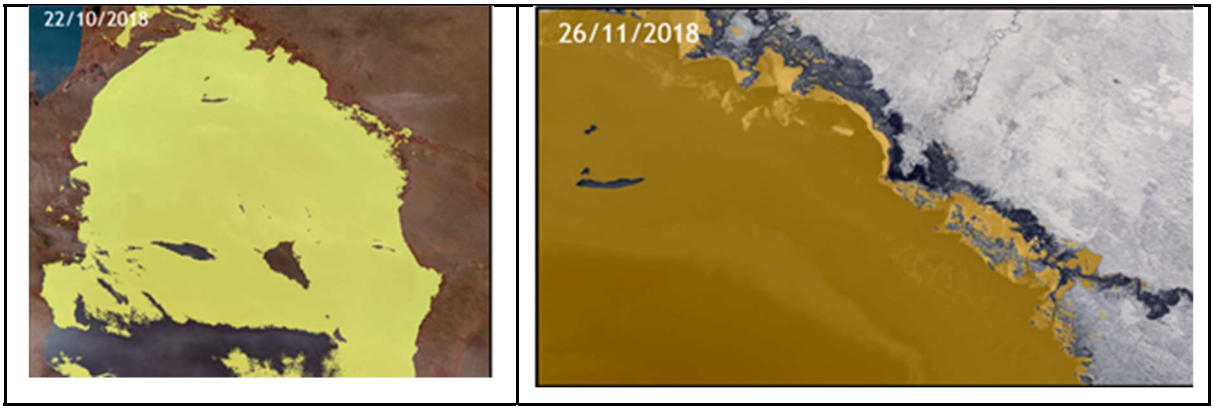


Figure 28. Artefacts observed over Alakol Lake when applying, left; RF: underestimation related to clouds right, OTSU; commission with snow.

3.3.1.6. Chilwa Lake

Lake Chilwa is the second-largest lake in Malawi after Lake Malawi. It is in eastern Zomba District, near the border with Mozambique. Lake Chilwa is a shallow, i.e. 3m, enclosed saline lake located along the East African Rift Valley in southern Malawi near its border with Mozambique.

Approximately 60 km long and 40 km wide. The lake is surrounded by extensive wetlands. There is a large island in the middle of the lake called Chisi Island. The lake has no outlet and the level of water is greatly affected by seasonal rains and summer evaporation. In recent years, Lake Chilwa has been shrinking.

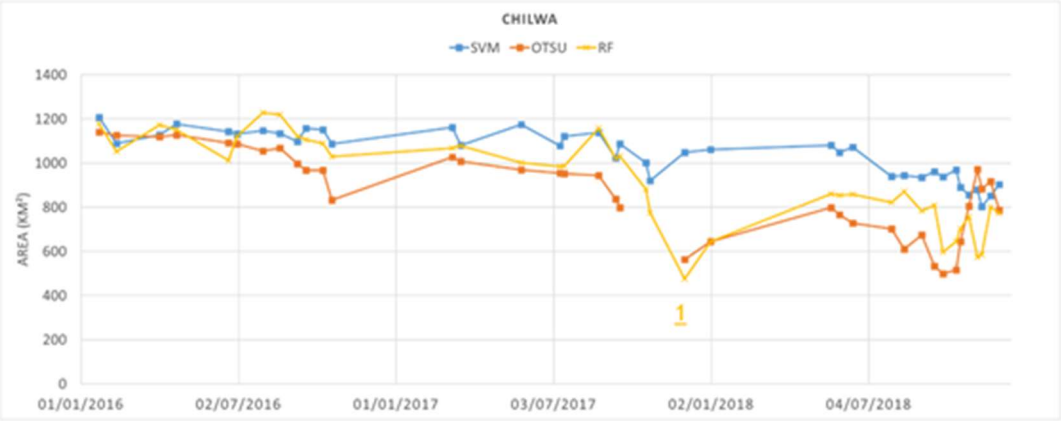
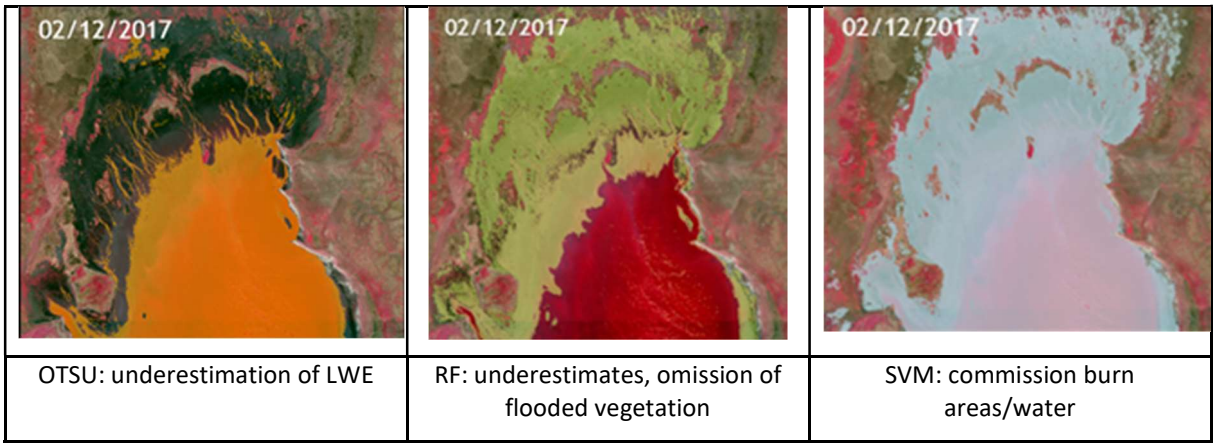


Figure 29. LWE derived from optical imagery based on different approaches over Chilwa Lake

Observations based on the graphic and vector extractions analysis are the following:

- OTSU: underestimates the water surface indicate by (1) on Figure 30.
- RF: underestimates. don't classify water with vegetation
- SVM: overestimates LWE as classifying burned areas as water surfaces



OTSU: underestimation of LWE	RF: underestimates, omission of flooded vegetation	SVM: commission burn areas/water

Figure 30. Artefacts observed over Chilwa Lake when applying OTSU, RF Tile and SWM on Sentinel2 imagery.

3.3.1.7. Al Hamar Wetland

Al Hammar wetlands are a large wetland complex in South Eastern Iraq that are part of the Mesopotamian Marshes in the Tigris–Euphrates river system. Historically, the Hammar Marshes extended up to 4,500 km2 during seasonal floods. They were destroyed during the 1990s by large-scale drainage, dam and dike construction projects. Since 2003, they are recovering following reflooding and destruction of dams. The water spatial distribution is still very controlled by inherited structures.

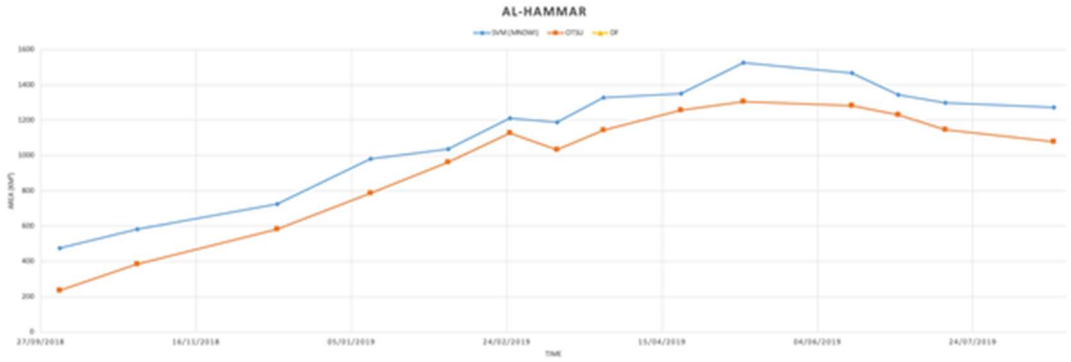


Figure 31. LWE derived from optical imagery based on different approaches over Al Hamar wetlands.

The LWE dynamic derived from satellite image, thanks to OTSU or SVM approaches are similar. Therefore, the values derived from SVM algorithm are higher than the OTSU ones. This is related to the fact that OTSU is more restrictive, taking into account free open water surface, rather than SVM which includes also a part of flooded vegetation.

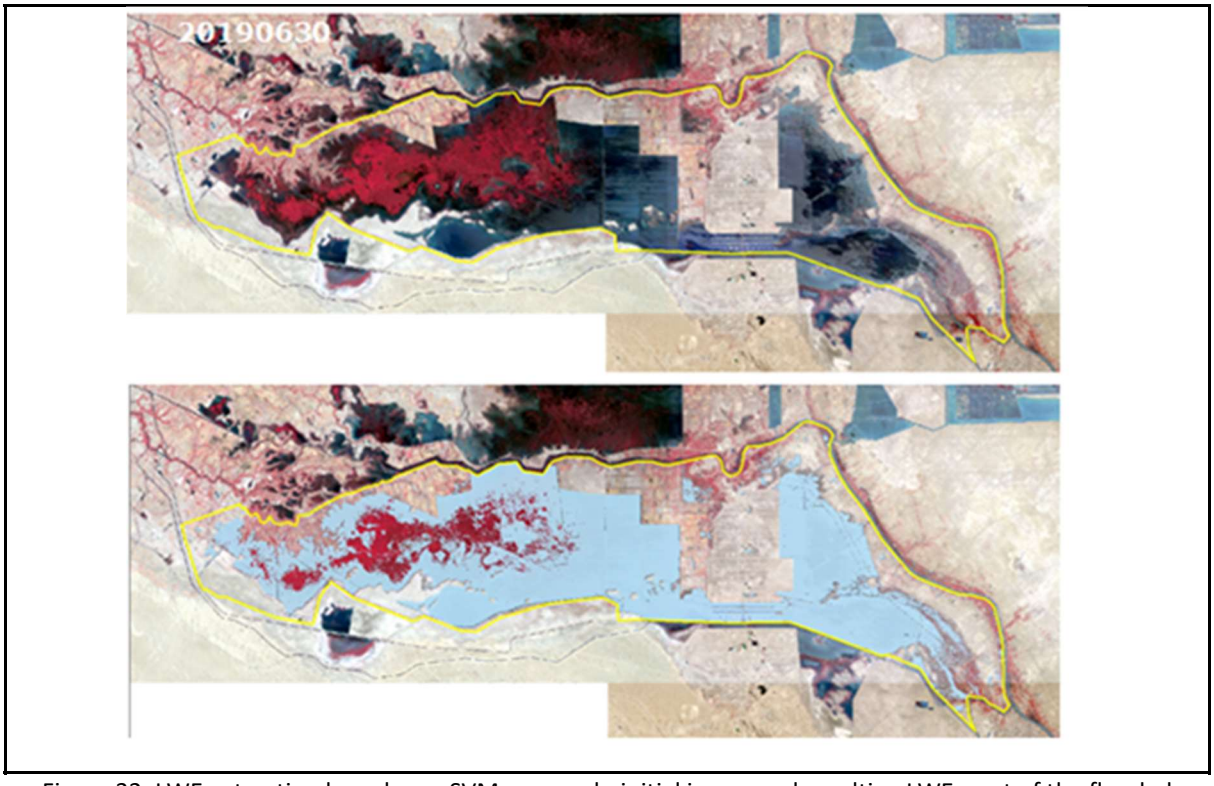


Figure 32. LWE extraction based on a SVM approach, initial image and resulting LWE: part of the flooded vegetation is integrated within the LWE.

3.3.1.8. Chad case study

The Chad case of study is particular in regard to the others tested lakes cases. It was an opportunity to:

- Test 5 different algorithms for LWE extraction based on optical imagery
- Compare results acquired from Sentinel2 and landsat8 data acquired the same day, but of course having different spatial resolution, ie 10 and 30m respectively.
- Compare LWE derived both from Optical and SAR imagery

- Compare the EO derived map with the with field tracks recorded from a boat.

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

This field survey was carried out in the Archipelago of the Chad lake, the Bol Reria area a disconnected part of interdune water bodies.

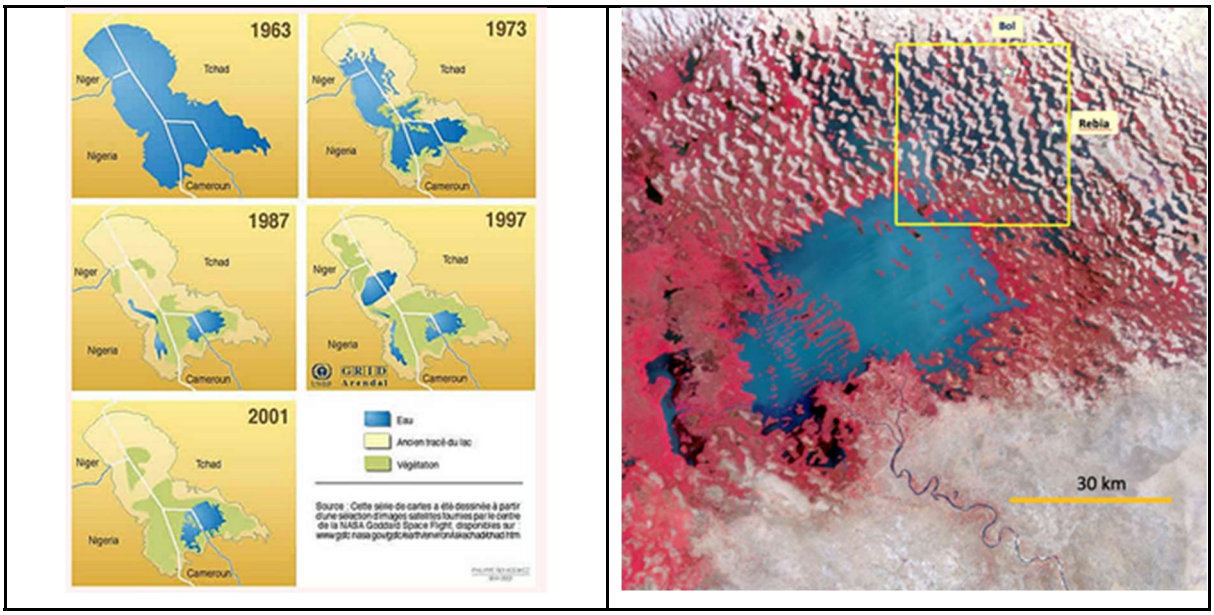


Figure 33. Chad lake evolution over 40years (1963-2001) and location of the Bol area within the Archipelago region of Chad Lake

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

Land tongues with sparse vegetation

Floating and anchored vegetation

Bare sandy tongues

Figure 34. Landscapes of the Archipelago area.

Five different methods were tested, OTSU, Cany-OTSU, Kmean, SVM and RF. Obtained LWE are presented in the Figure 35.

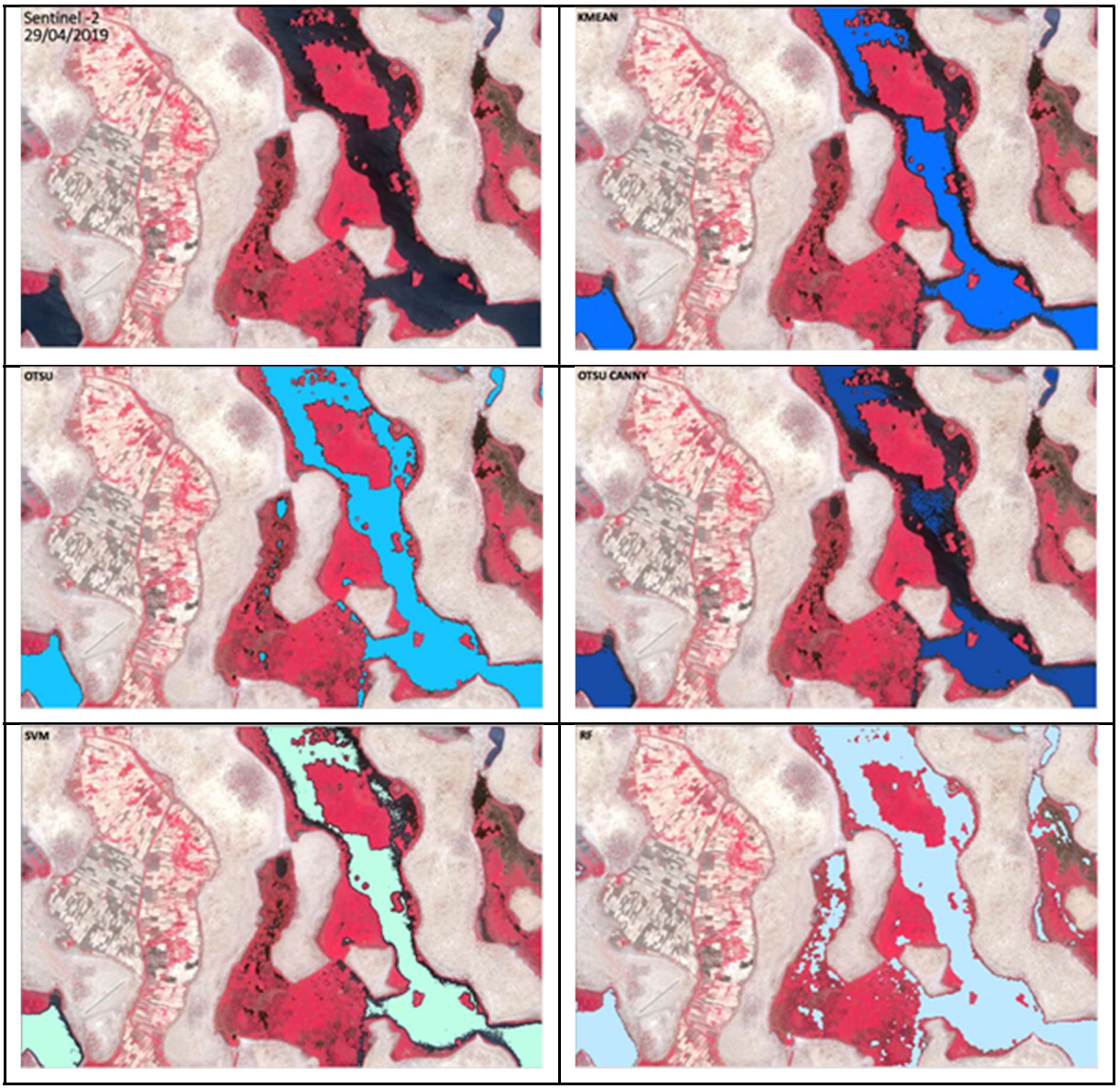


Figure 35. LWLs extracted exploiting different approaches from the Sentinel2 image acquired on the 29 of April 2019.

In this complex environment, the analysis and comparison of the derived LWE from Sentinel2 imagery are the following:

- Canny OTSU and Kmeans, a none negligible underestimate the LWE
- SWM a small under estimation
- OTSU, the most realistic over this landscape
- RF a small overestimation as integrating part of the wet/floating vegetation

All methods retrieve about 70 % of the potential observed water, all real open water surface. The differences are noticeable on very shallow water and on the immediate environment or inside floating islands.

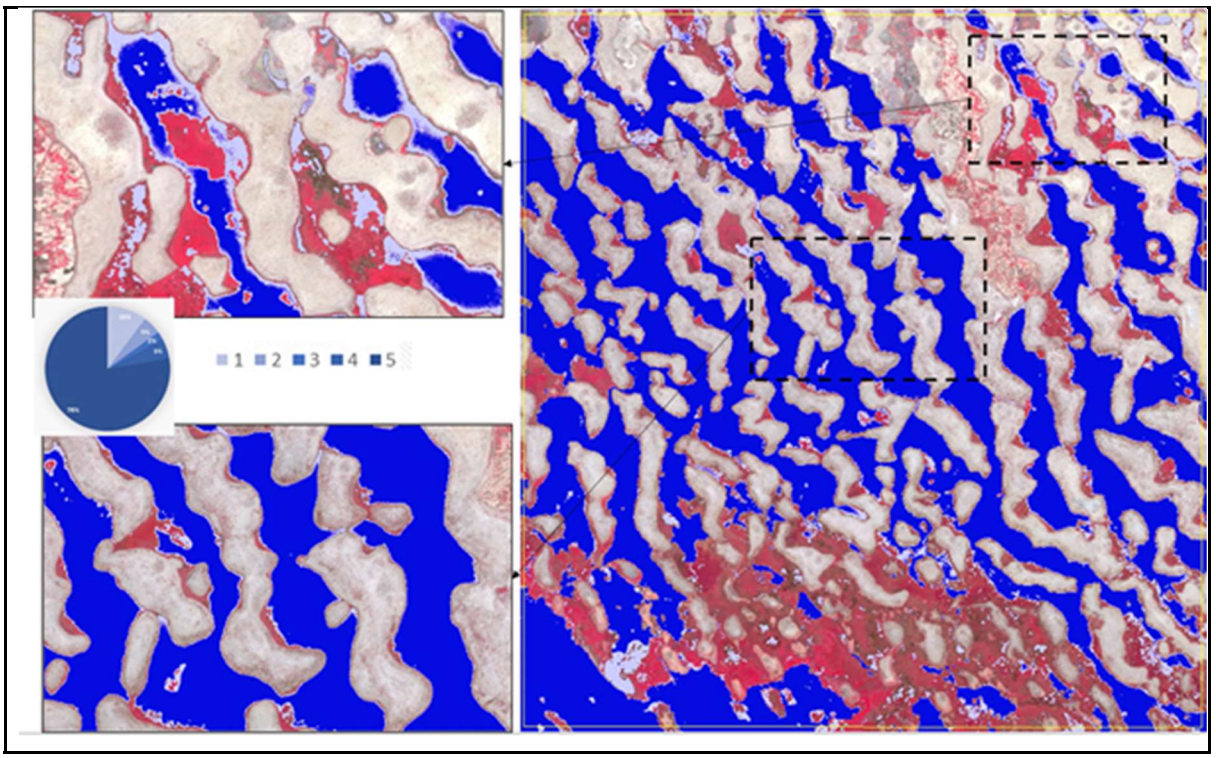


Figure 36. Occurrence map of observed water based on Sentinel2 imagery.

The same five approaches were also followed to retrieve LWE from the Landsat imagery. In this complex environment, the analysis and comparison of the derived LWE from Landsat imagery are the following:

- Kmeans, a none negligible underestimate the LWE
- Canny-OTSU presents a small under estimation
- SVM and OTSU, are very similar in terms of obtained results and present the most realistic over this landscape
- RF a small overestimation as integrating part of the wet/floating vegetation

Based on Landsat 8 imagery, i.e. with a pixel integrating more surface, all methods present in common more than 80 % of the potential observed water, all real open water surface. The slight differences are noticeable on very shallow water and on the immediate environment or inside floating islands.

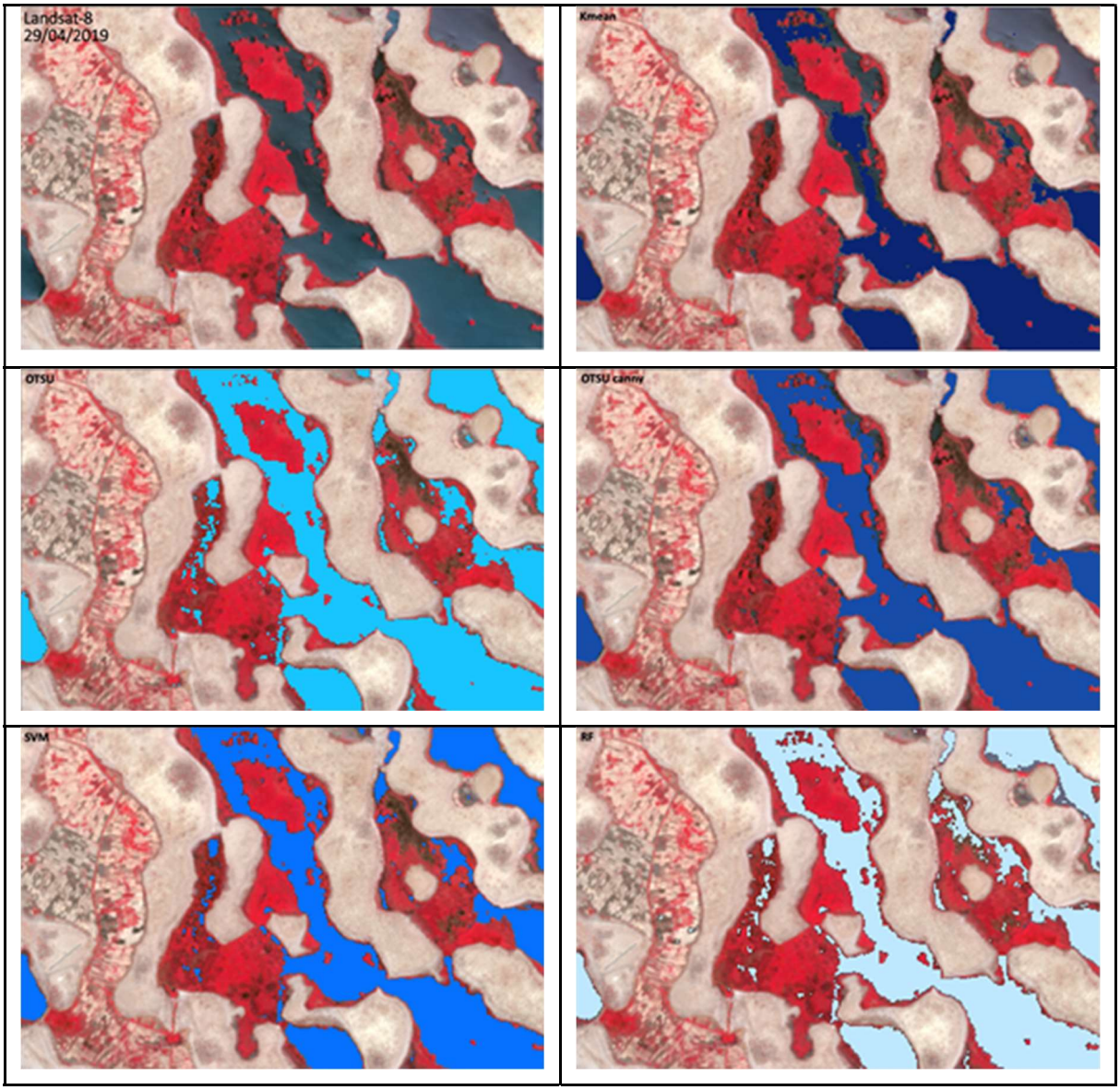


Figure 37. Occurrence map of observed water based on Landsat 8 imagery.

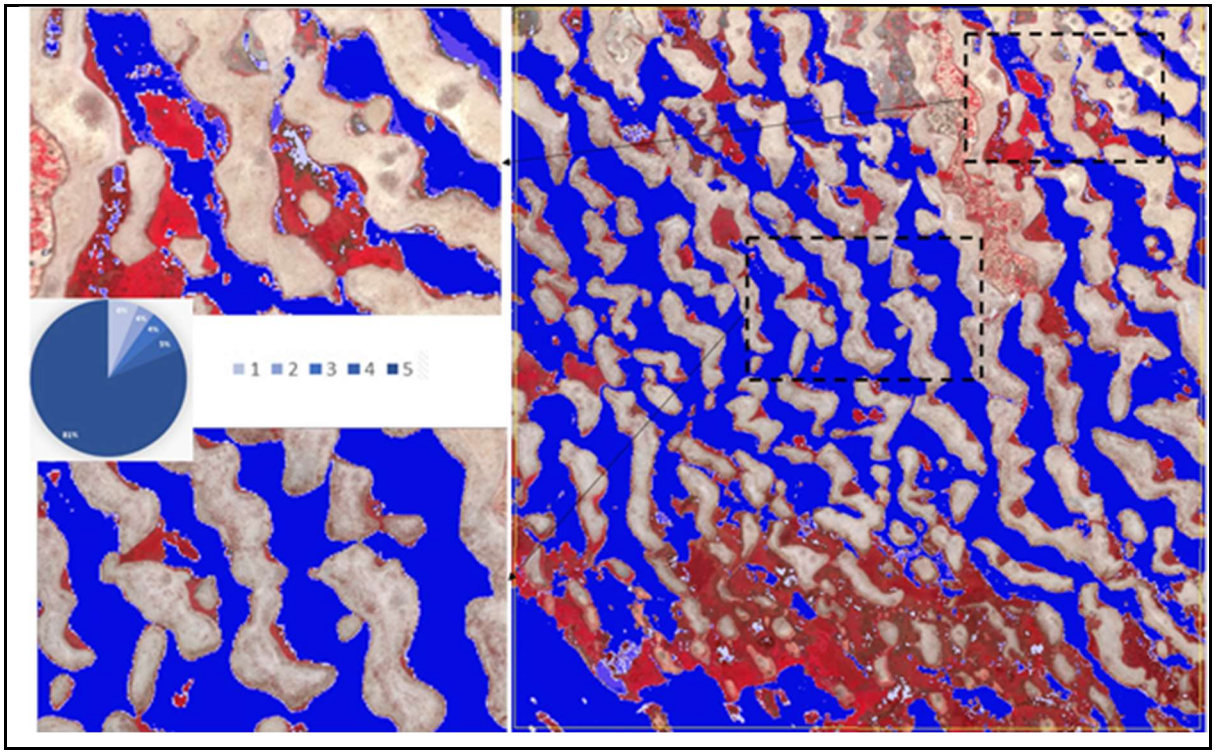


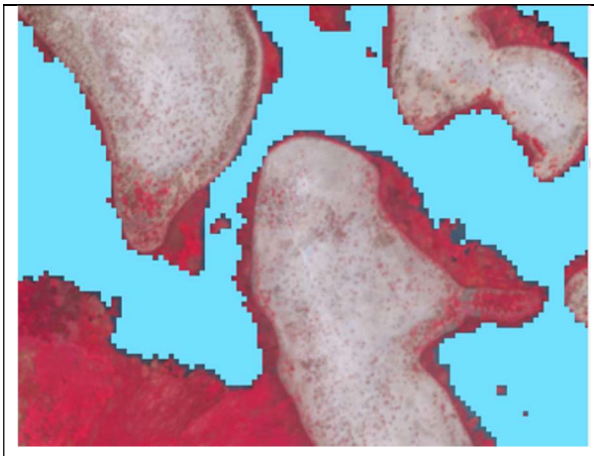
Figure 38. Map of occurrence of observed water based on Landsat8 imagery.

The comparison of LWE extraction from Sentinel2 and Landsat 8 is interesting, the extraction based on Landsat 8 is more coherent/continuous, with less difference between the results derived from the various methods.

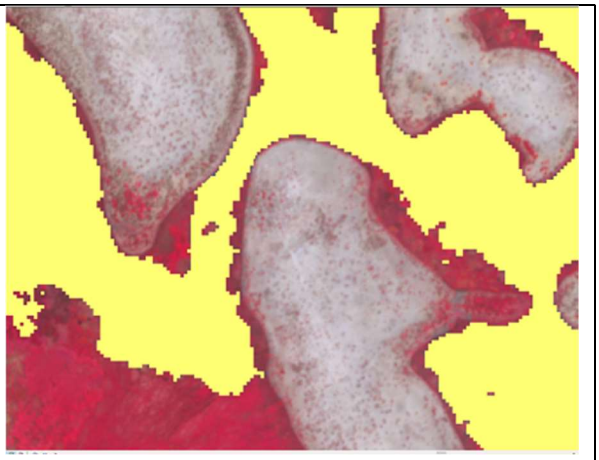
Table 12. Comparison of the Sentinel2 and Landsat 8 bands.

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

If in the SWIR1 and SWIR , the bands are located similarly and have same width, on the NIR there are noticeable difference between Sentinel2, band B8 10m, and landsat8, band 5 whereas the narrow band B8A of Sentinel2 is similar to the NIR band, band 4, of Landsat 8, with respectively a 20 and 30 meters of spatial resolution. Whereas, on the LWE derived from Sentinel2, the limits water/non water are more accurate, but some overestimation is observed with vegetation areas classified. On the LWE obtained from a RF approach over a Landsat 8 image, it can be observed an underestimation of water on the shore of the lake due to the lower resolution.



LWE derived from Landsat-8 extraction on the shoreline (background Sentinel2)



RF Sentinel-2 extraction with Sentinel-2 image (background)

Figure 39. Comparison of Sentinel 2 and Landsat 8 LWE: influence of the spatial resolution.

In addition, a multitemporal analysis was done over a period of one month, exploiting five Sentinel 2 images. These data were acquired at the following dates.

- 09/04/2019
- 14/04/2019
- 24/04/2019
- 29/04/2019
- 04/05/2019

For each image, a water surface based on an SVM approach was derived. From these a mean surface was calculated and then each surface compared to this mean surface.

Table 13.

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

Very stable water bodies, with less than 2% of difference between the observations within a month. In addition, the LWE apparent dynamic is more related to the movement of floating vegetation (under wind direction/speed dependence) rather than LWL changes. Some also apparent changes are related to some artefact linked with small clouds on one date.

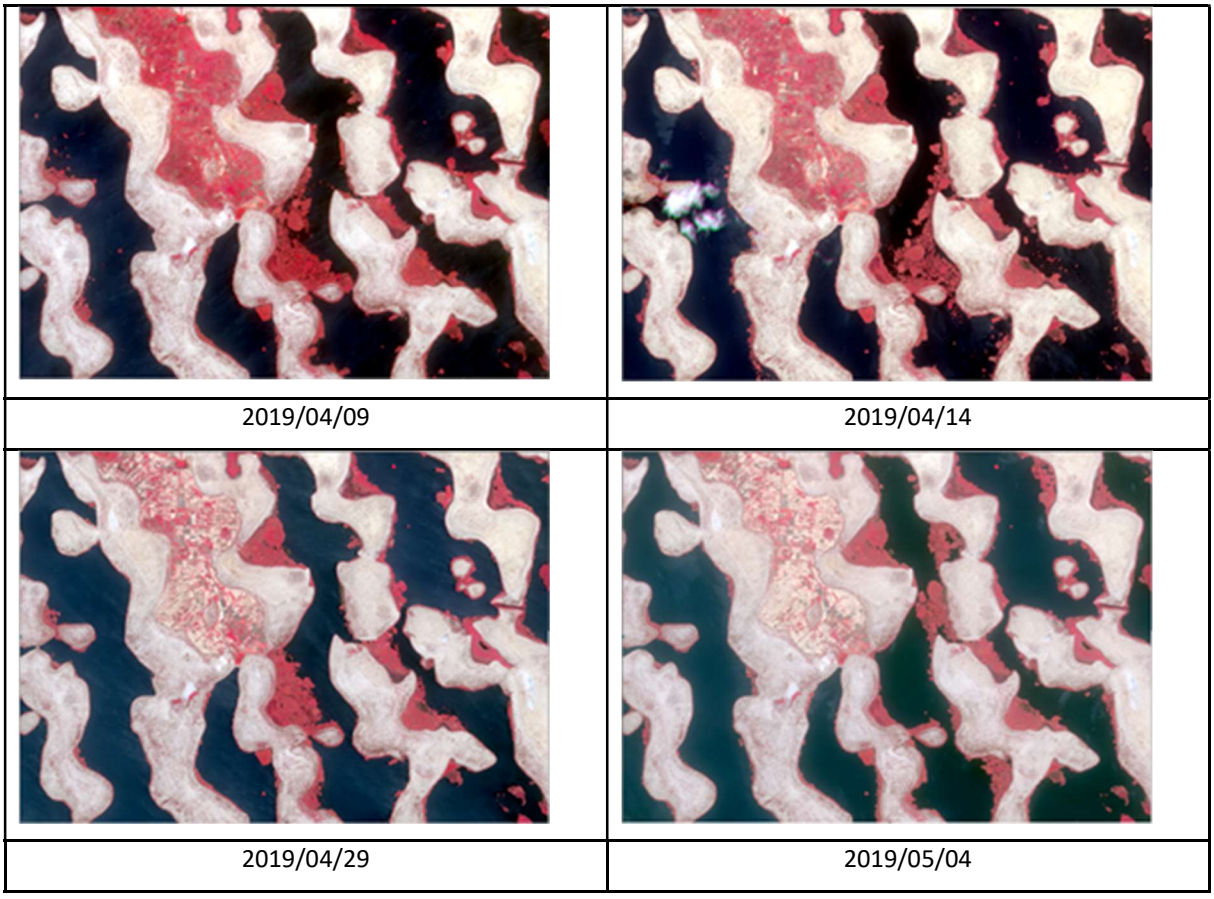
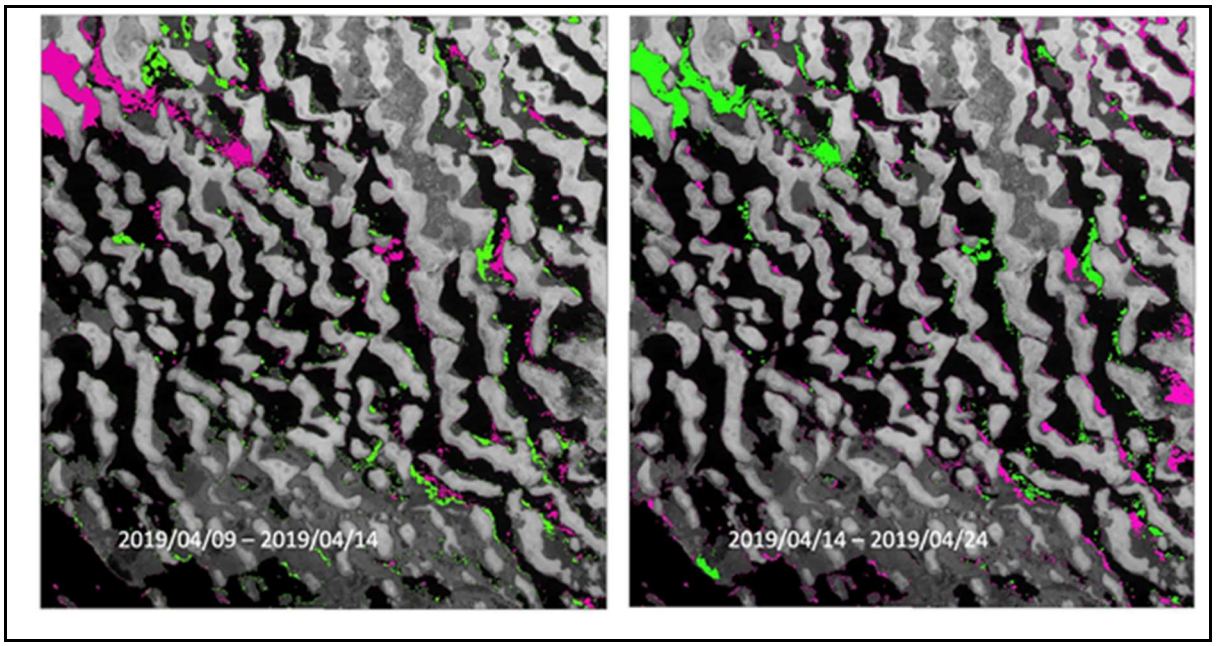


Figure 40. Displacement over one month of the Kirtas.



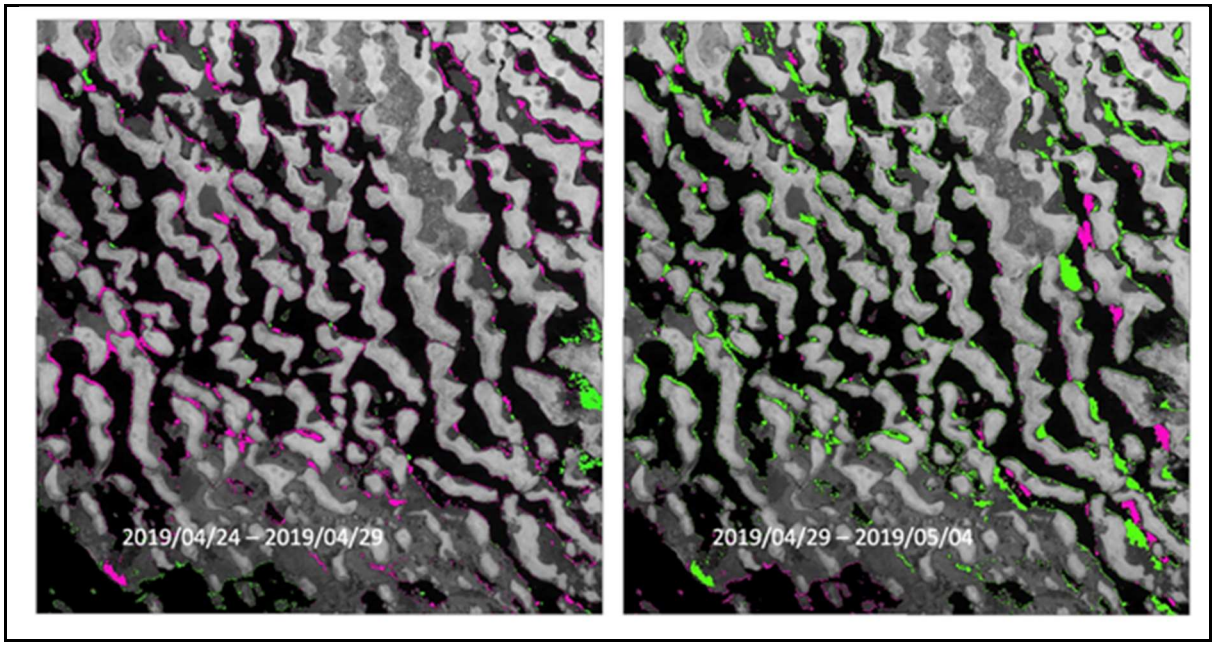


Figure 41. Displacement of vegetated islands from date to date. Green colour: No water in first image, water in last image. Lily colour: Water in first image, no water in last image.

3.3.1.8.1. Validation with boat track.

A field trip has been organized by LEGOS on parallel of the CCI lake work. Boat tracks on Landsat , using a GPS, receiver allowed to carefully map the water and the vegetation. It is allowed locally to validate the water limits, another part the boat is in the middle of bays. An interesting thing is also that the tracks cross within vegetated island, confirming well the “water” aspect of these island. . All changes in the nature of the terrain crossed by the boat were noted allowing to precisely map the lake countour along the boat’s route.

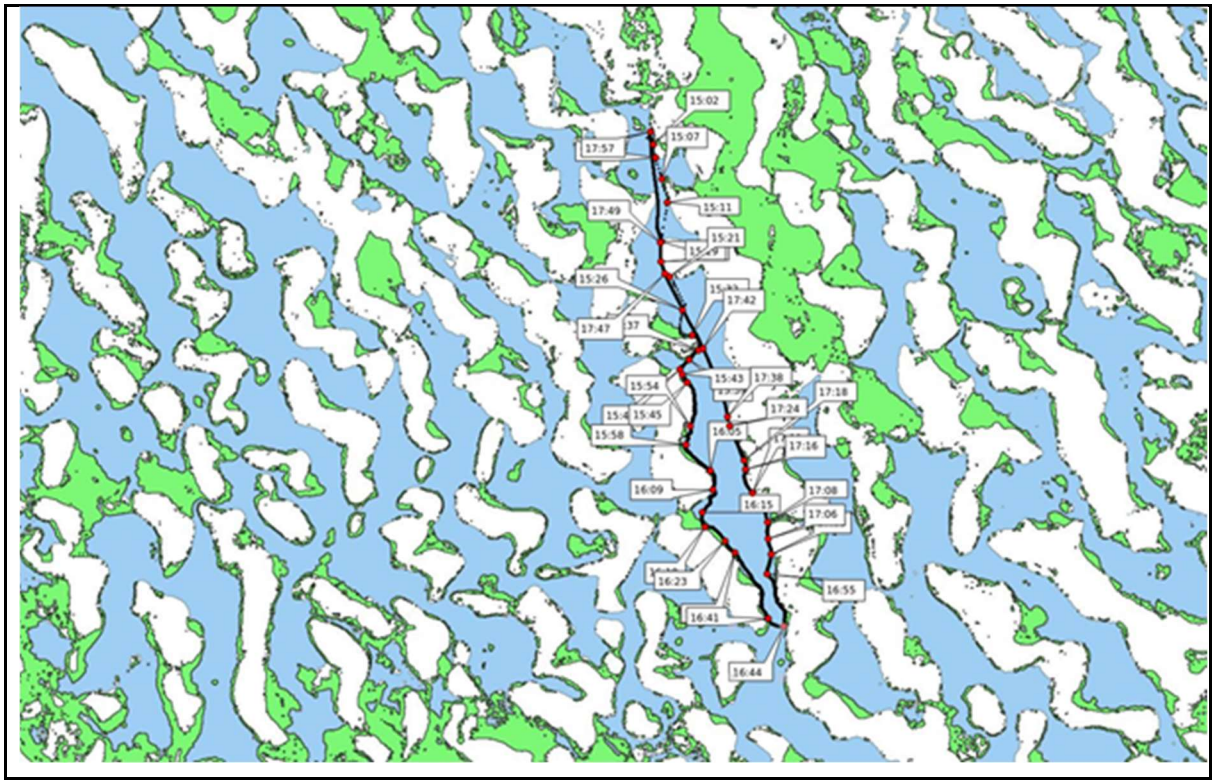


Figure 42. Tracks of boat survey, 10 of April 2019 within the Chad lake Archipelago (Courtesy of Legos).

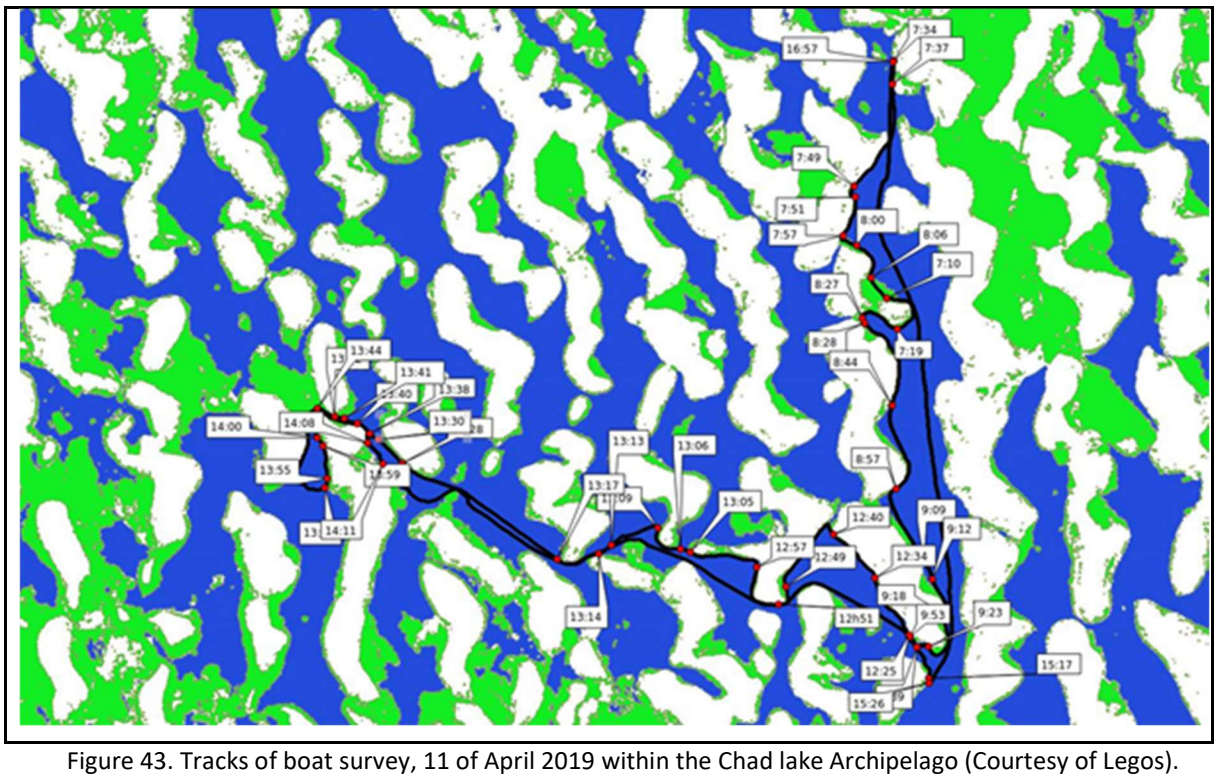
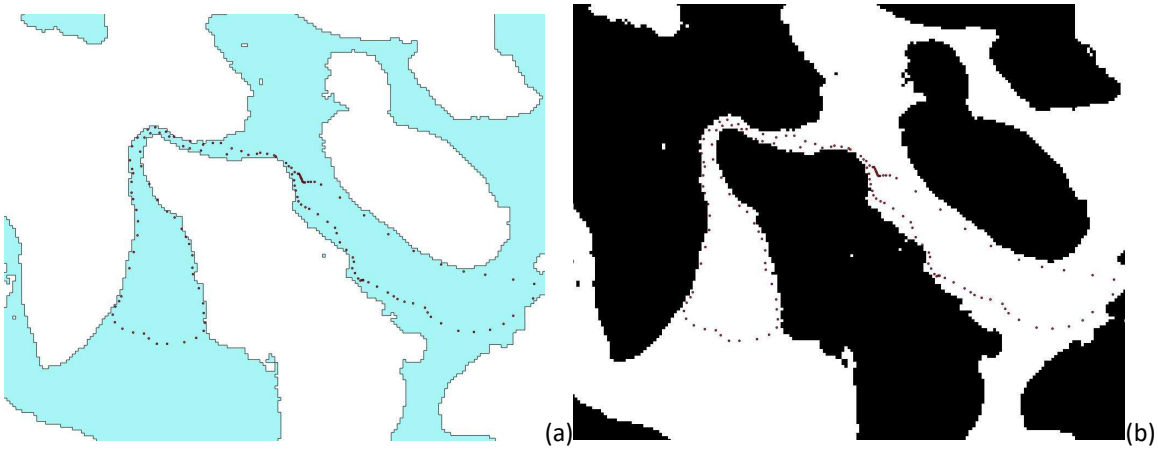


Figure 43. Tracks of boat survey, 11 of April 2019 within the Chad lake Archipelago (Courtesy of Legos).



(a) (b)

Figure 44 (a) Processing using normalized MNDWI index with threshold of -0.1 (b) using OTSU method. Both processing of the landsat 8 images allow to correctly map the transition between land and water in this complex archipelago system.

Reference:

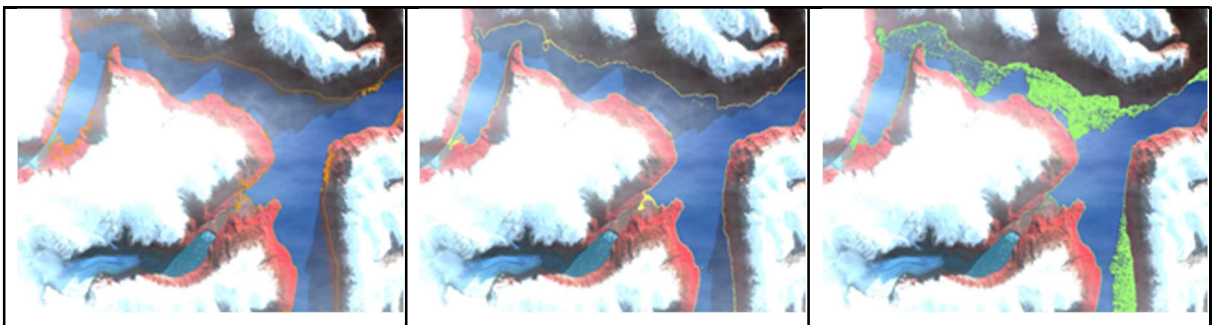
Leonard J., 1974: Aperçu de la végétation de la partie Est du lac Tchad. ORSTOM

Iltis André and Jacques Lemoalle, 1983: the aquatic vegetation of lake Chad, in lake Chad, Ecology and productivity of a shallow tropical ecosystem. JP Carmouze, Durand JR and C Leveque edts. Junk Publishers.

3.3.1.9. Argentino lake

Lago Argentino, also name El Calafate, is a lake in the Patagonian province of Santa Cruz, Argentina. It is the biggest freshwater lake in Argentina, with a surface area of 1,415 km2. It has an average depth of 150 m, and a maximum depth of 500 m. The lake lies within the Los Glaciares National Park in a landscape with numerous glaciers and is fed by the glacial meltwater of several rivers, the water from Lake Viedma brought by the La Leona River, and many mountain streams

The argentine lake case is interesting, due to the potential impact of environment on the retrieval of LWE. Two major elements have to be taken into account, the topographic position of the lake surrounding by mountains, and the presence of glaciers feeding the lakes and of snow covering the borders. By the way it is challenging case for testing the different methods of LWE based on optical imagery.



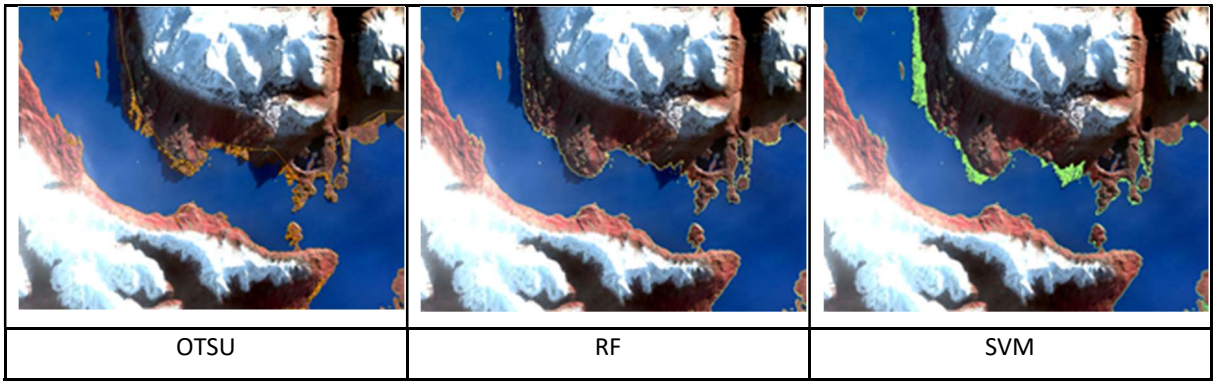
OTSU	RF	SVM

Figure 45. Artefacts observed over Argentino Lake when applying OTSU, RF Tile and SWM on Sentinel2 imagery acquired the 02-06-2019.

In this second set of tests for each processed image, SVM and RF was trained based on Pekel water mask occurrence.

It is appearing that:

- OTSU and SVM are sensitive to shadows on water surfaces.
- RF is less sensible to shadows effect and by the way presents a more realistic shoreline.



OTSU	RF	SVM

Figure 46**. Artefacts observed over Argentino Lake when applying OTSU, RF Tile and SWM on Sentinel2 imagery acquired the 02-06-2019.**

When analyzing the temporal evolution of the LWE, a relative coherence of the LWE evolution over the time obtained by the different process. Therefore, the OTSU tends to overestimate the water extent, whereas the SVM, underestimates it. This explains the differences between the methods at 02/06/2017. Large shadows are present at the limit between the lake and the land. RF Tile is more consistent than SVM which don't classify shadows on water and OTSU which classify shadows on land.

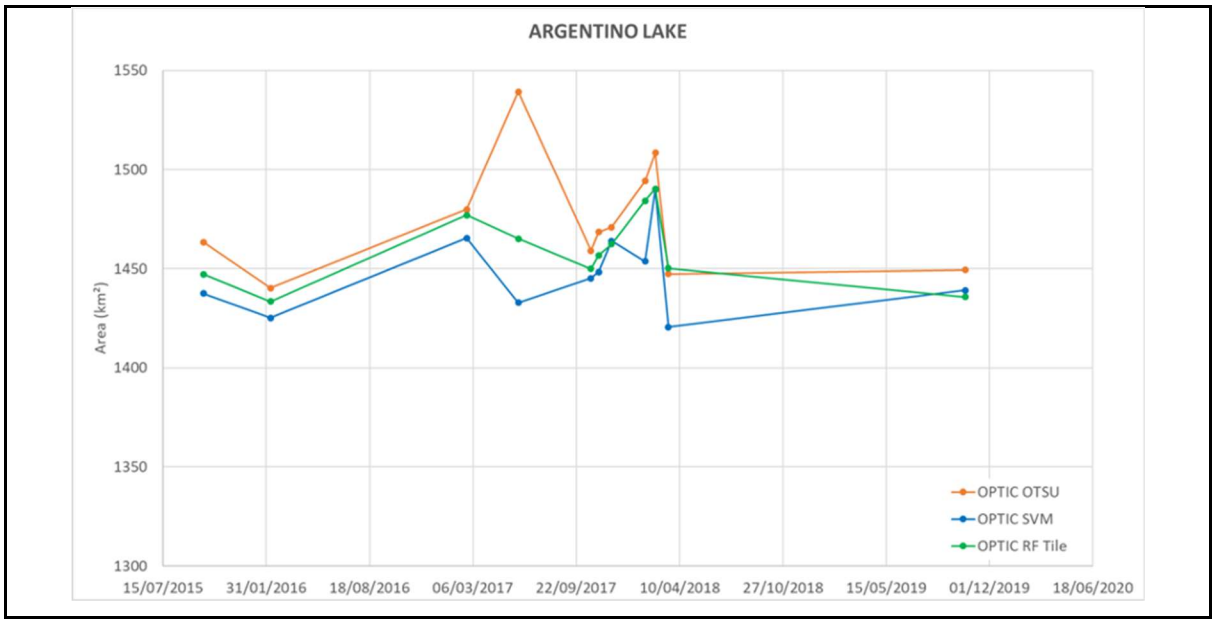


Figure 47. LWE derived from optical imagery based on different approaches over Argentino Lake

3.3.1.10. Bosten lake

The Bosten lake, also known as Bagrash lake, is a large freshwater lake on the North Eastern rim of the Tarim Basin, about 20 km East of Yanqi, Xinjiang, in the Bayingholin Mongol Autonomous Prefecture, an extremely arid region in the North West of China. Covering an area of about 1000 km2, it is the largest lake in Xinjiang. The mean water depth is 8,2m with a maximum depth of 17m. The lake is frozen during the winter (up to March). In the Western part: there is a huge wetlands complex that is totally dry in beginning spring. In addition, the environment of the lake varies a lot around the year why lot of vegetation in summer (as well on the nearly agricultural fields that could induce some potential confusion with the wetlands vegetation. In fall, the vegetation is relatively dry.

In term of data accessibility and quality, it has to be noticed that a part of the analysis is spoiled by the fact that a relative important amount, i.e. 7 dates, of Landsat 7 images covering an interesting hydrological period where not exploitable due to some instrument artefact.

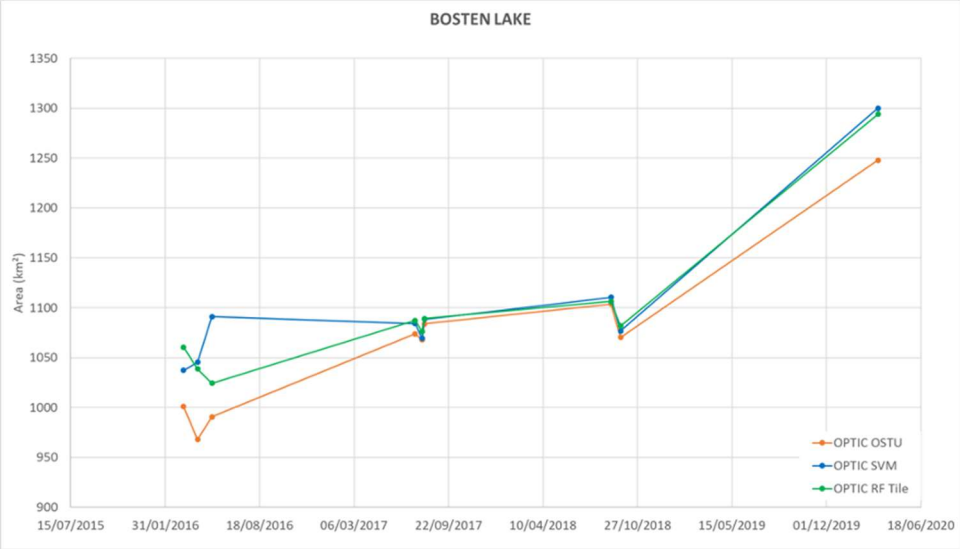
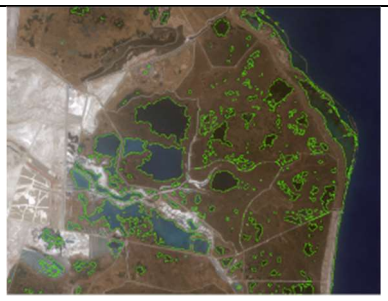
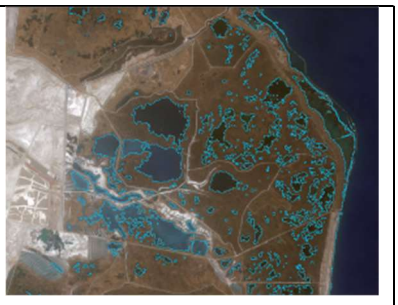
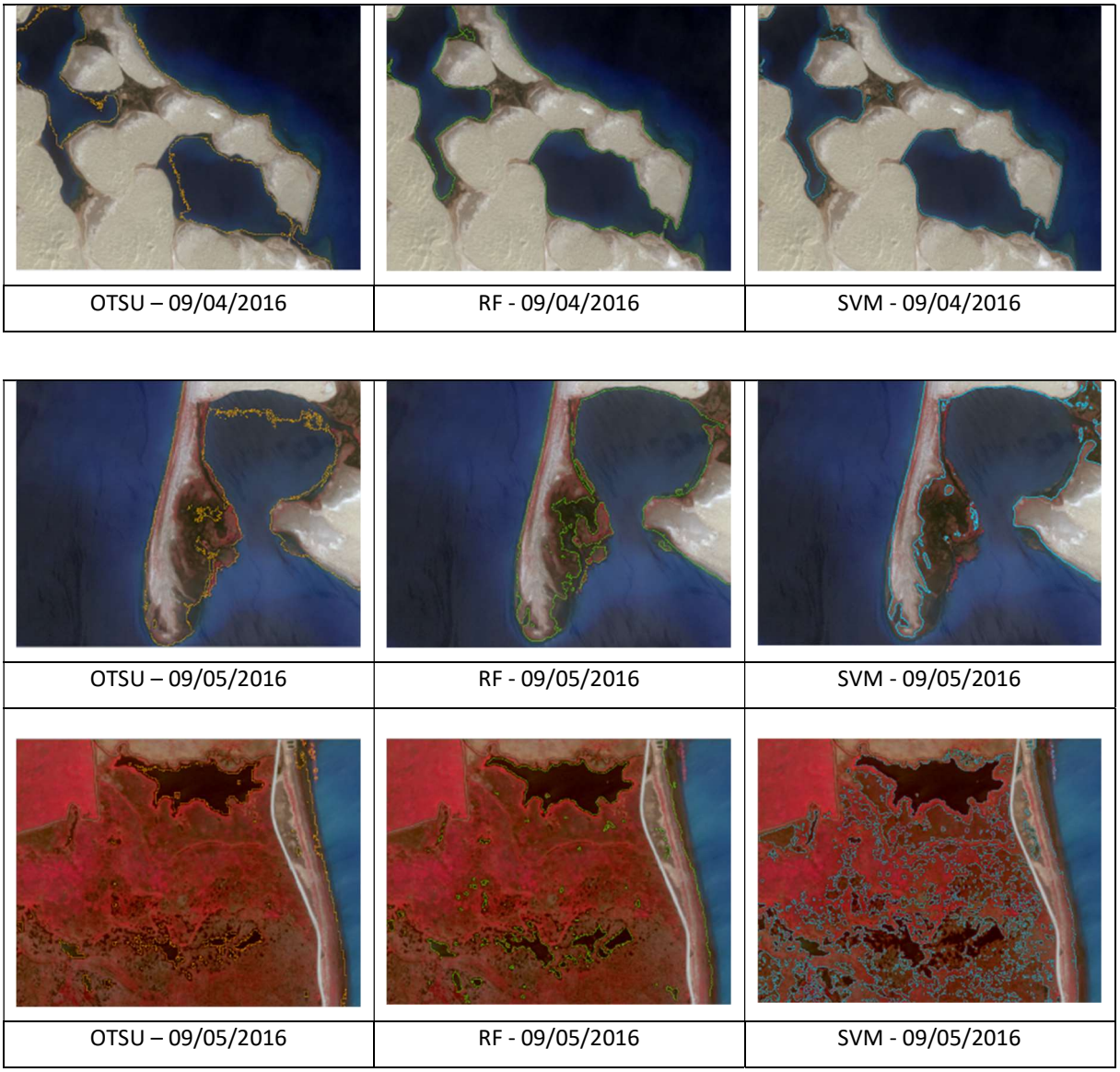


Figure 48. LWE derived from optical imagery based on different approaches over Bosten lake.

		
OTSU – 09/04/2016	RF - 09/04/2016	SVM - 09/04/2016



OTSU – 09/04/2016	RF - 09/04/2016	SVM - 09/04/2016
OTSU – 09/05/2016	RF - 09/05/2016	SVM - 09/05/2016
OTSU – 09/05/2016	RF - 09/05/2016	SVM - 09/05/2016

Figure 49. Comparison of LWE limits obtained from OTSU, Random Forest (RF) et SWM over Bosten wetlands.

Observations based on the graphic and vector extractions analysis are the following:

- OTSU: underestimates some shallow water
- RF: The most realistic over this landscape.
- SVM: overestimates water surface, especially in dry wetland
- The RF approach is the best technical solution to derive accurate LWEs over the Bosten Lake.

As illustrated above with water masks derived from the images Sentinel-2 acquired the 09/04/2016 and the 09/05/2016, OTSU is overestimating and SVM is underestimating on this data.

3.3.1.11. Khanka

Lake Khanka is located on the border of the People's Republic of China (China) and the Russian Federation (Russia). It is the largest lake in Northeast Asia, as well as a transboundary waterbody between China and Russia. The lake is called Lake Xingkai in China and Khanka Lake in Russia. The water plane of the lake varies between 4,000-4,400km2. The average lake depth is 4.5m and maximum lake depth is 10.6m.

There are 23 inflowing rivers to the lake, 8 draining from China and 15 draining from Russia. The Song'acha River is the only outflow river from the lake and is subsequently connected with the Wusuli/Ussuri River and the Heilong/Amur River system. The drainage basin of Lake Xingkai/Khanka is a habitat for important animal and plant species of both countries, particularly the wetlands surrounding the lake. The Russian Federation designated the lake as a Ramsar Convention wetland site, on the basis of its importance for migratory bird species.

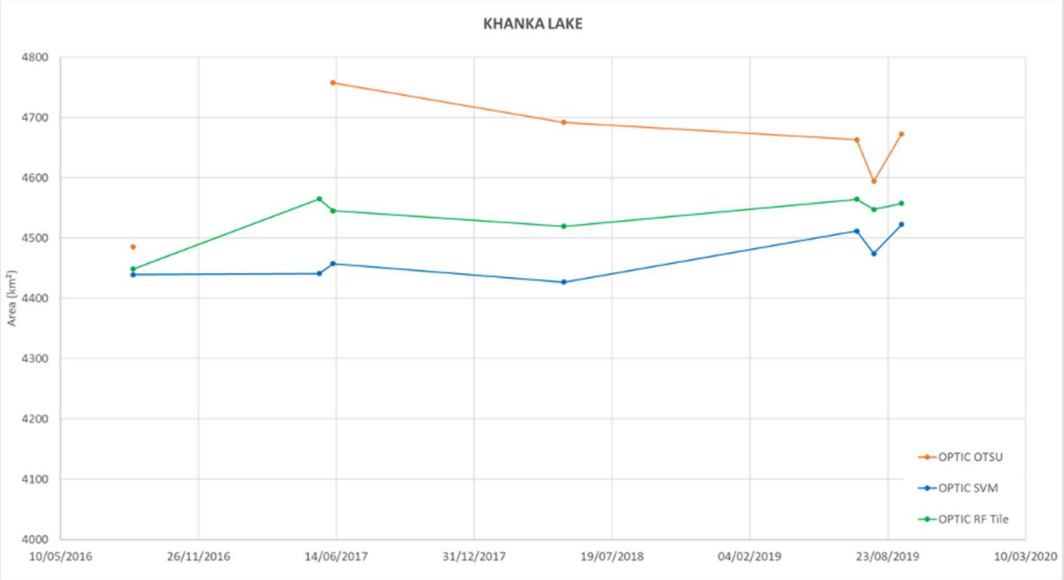
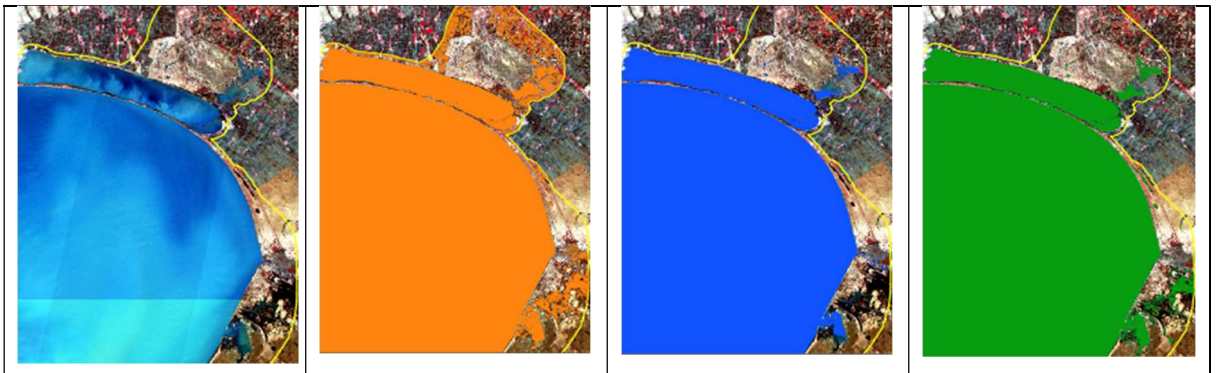


Figure 50. LWE derived from optical imagery based on different approaches over Khanka Lake.

The comparison of the LWE obtained based on OTSU, SV and RF indicates:

- OTSU, Overestimation observed in dry wetlands and in crops that could extend far beyond the AOI boundaries.
- SVM, Underestimation in small water bodies.
- RF is Closest from the ground truth.



Sentinel-2 image 10/05/2018	In orange OTSU LWE	In blue SVM LWE	In green RF LWE

Figure 51. LWE obtained with OTSU, SV and RF.

Presented at local scale in the above figure, these trends are observed on the whole S2 image. Conclusion over Khanka case of study is that RF approach allows to derive accurate LWEs.

3.3.1.12. Ilmen lake

The Ilmen lake is a large lake in the Novgorod Oblast of Russia. The average surface area is 982 Km2, therefore it may vary between 733 km2 and 2090 square km2 depending on water level. The lake is fed by 52 inflowing rivers, the four main ones being the Msta, the Pola, the Lovat, and the Shelon. It is drained through a single outlet, the Volkhov, into Lake Ladoga.

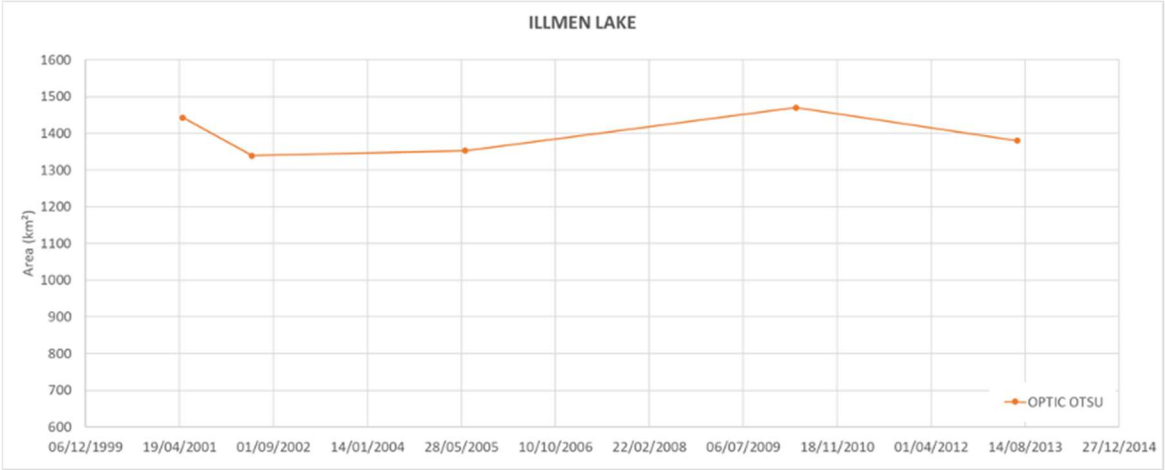


Figure 52. LWE derived from optical imagery based on different approaches over Illmen Lake.

Part of the processing was done exploiting OTSU approach from GEE. Therefore, not all the images selected from USGS sites, are accessible from GEE, by hence, the optical time series is not so large.

3.3.1.13. Sary kamysh Lake

The Sarykamysh or Sary-Kamysh, also known as Sarygamysh is situated in Central Asia approximately midway between the Caspian Sea and the Aral Sea. The lake sits in an oval depression of tectonic origin, which was later affected by aeolian erosion during successive glaciations from 2 million years ago, to 10 000BC. The Northern quarter of the lake belongs to the country of Uzbekistan, while the rest belongs to Turkmenistan. The Sarykamish lake was formed in 1971 as a result of flooding of a set of small lakes located within the depression. Now Sarykamish lake is a large drainage water body which has been used as a discharge collector of salty irrigation water from the fields. Since 1992 the lake has been progressively increased in size a, reaching maximal levels in the beginning of 2000, an increasing still on going with some recession phases. The mean water depth is 8 m with a maximum depth of 4m; its surface is about 4000km2.

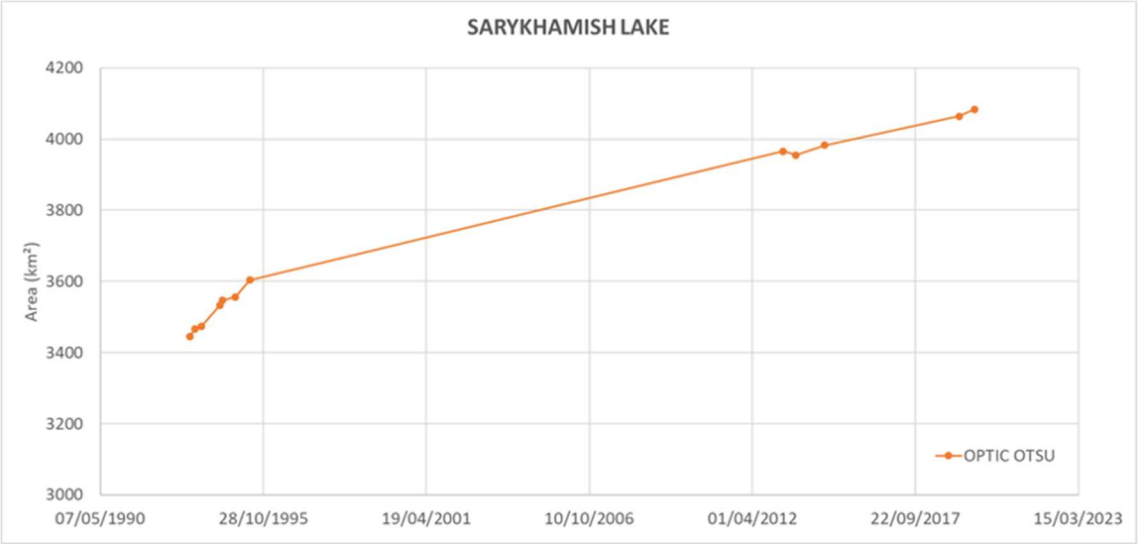


Figure 53.: LWEs derived from optical imagery based on different approaches over Sary Khamish Lake.

The idea was to have a selection of images covering the different phases of lake inflow. The selected data started from the beginning of the 90' up the recent years. Unfortunately, a relative long period, in the beginning of the 2000' was not covered due to the instrumental problem on Landsat 7.

3.3.2. Results analysis for SAR sensors-based approaches

As commented in Section 3.1.2, series of Sentinel-1 images have been exploited to derive their Lake Water Mask and their corresponding Lake Water Extent. The LWE measurements will be used along with the Lake Water Level in order to generate the corresponding hypsometric curves.

As described in the ATBD document, NORCE and TRE-Altamira have employed different approaches (even though sharing some common steps) and have used GRD and SLC images respectively. A main limitation with S1 data is that this constrains the time-series to the period after October 2014. S1 data are geocoded using the SRTM DEM and a most precise one when available (as for Lake Altevatnet).

The Lake Water Masks generated by NORCE and TRE-Altamira over the same images' dates (or presenting a low temporal difference) are used in order to compare both methodologies' relative performances. In order to maximize the contribution of the SAR LWE to the hypsometric curve calculation, NORCE and TRE-Altamira employed different S1 datasets (including different orbits) but also considering some common dates for comparison purposes. In the same way, images presenting a time difference of 12 days have also been considered. This can be questionable sometimes when the LWE changes rapidly, but in order to obtain sufficient amount of comparisons we have used this criterion. This analysis has been carried out over the following lakes: Chad (section 3.3.2.1), Illmen (section 3.3.2.2), Bosten (section 3.3.2.3), Argentino (section 3.3.2.4) and Khanka (section 3.3.2.5).

The comparison statistics are derived from the contingency matrix. For its calculation an analysis polygon covering the lake is selected and one of the two results is taken as the reference. Then the following parameters are calculated in %:

- True positive (TP): both results detecting water.
- False negative (FN): reference result detecting water and the other land.
- False Positive (FP): reference result detecting land and the other water.
- True negative (TN): both results detecting land.
- Accuracy rate (AR): TP + TN.

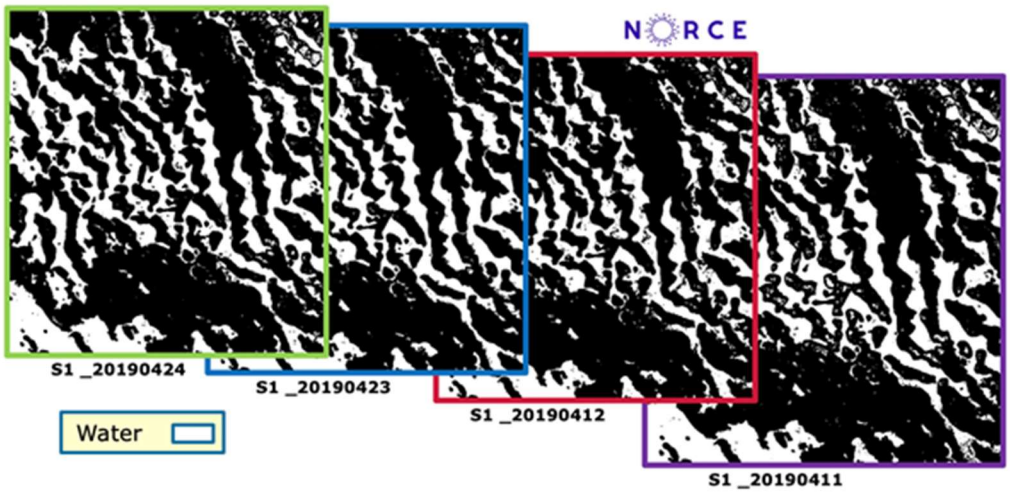
It is worth saying that the naming of the statistical parameters (TP, FN, FP and TN) is also somewhat arbitrary since it is not always clear which of the datasets that is closest to the ground truth. For the sake of comparison,

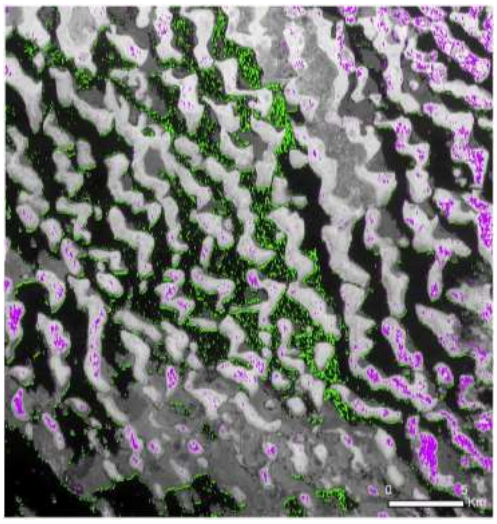
we always set the first dataset as truth. In order to illustrate the comparison, the contingency matrix as well as the temporal LWE series are shown.

In section 3.3.2.6 a summary with the tabular contingency matrices results for all the lakes that are inter-comparable between the two SAR algorithms are provided.

It is also important to mention that an Envisat ASAR WSM database is available. For some of the lakes we have found interesting data from the period 2002-2012 which are useful for some of the lakes that have had significant changes in the LWE in the period after 2000. ASAR WSM data has in general coarser spatial resolution, and only one polarization so the quality is in general poorer than for S1 data.

3.3.2.1. Chad





	LWE (km2)	
Date	NORCE	TRE-ALTAMIRA
11/04/2019	354.127	324.364
12/04/2019	339.114	299.972

	LWE (km2)	
Date	NORCE	TRE-ALTAMIRA
23/04/2019	354.471	308.309
24/04/2019	346.339	335.068

3.3.2.2. Illmen

NORCE's and TRE-ALTA's lake water mask and their corresponding contingency matrices for a couple of coincident dates (20191024 and 20200116) which present a significant extent variation among them, are depicted in Figure 54. The overall agreement for those images are 97.73% and 96.19% respectively. Most of the differences between the two SAR approaches is related to the integration of inundated wetlands.

For Lake Illmen, the contingency matrix values for all NORCE's and TRE-ALTA's results corresponding to images presenting a temporal difference lower or equal to 12 days are can be found in Table 14.

A representation of the temporal evolution of all NORCE's and TRE-ALTA's LWE is depicted in Figure 55. The same representation but just focusing on the 12 days difference images is depicted in Figure 56.

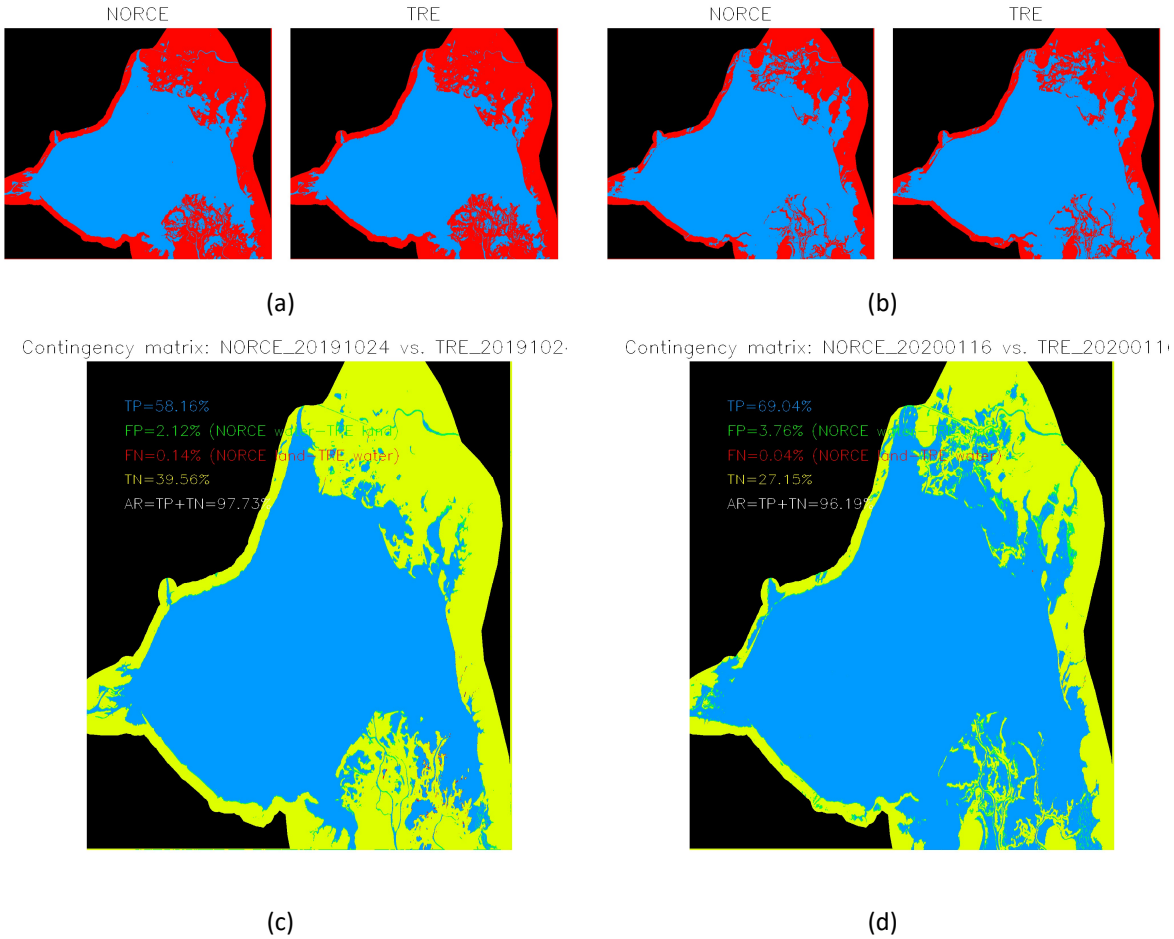


Figure 54. (a) NORCE and TRE-ALTA 20191024 lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b) NORCE and TRE-ALTA 20200116 lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

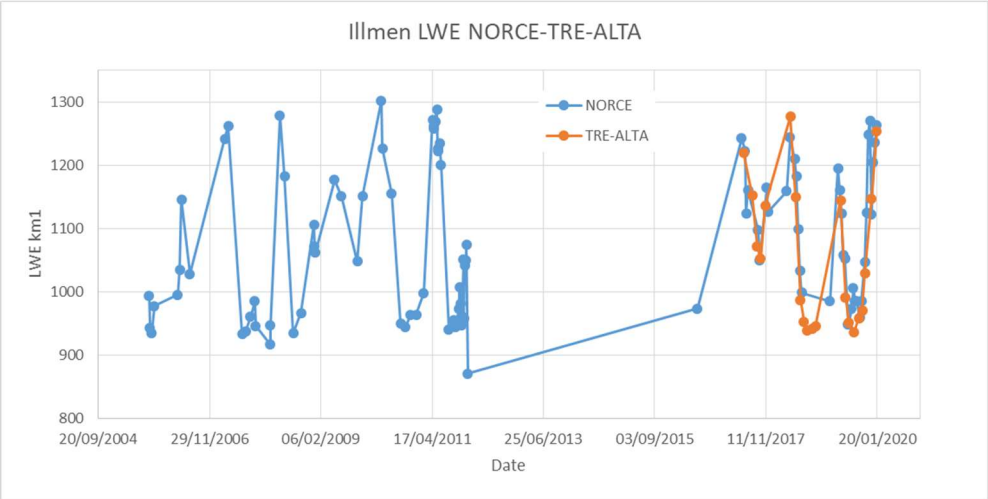


Figure 55. NORCE and TRE-ALTA LWE time series for Lake Illmen.

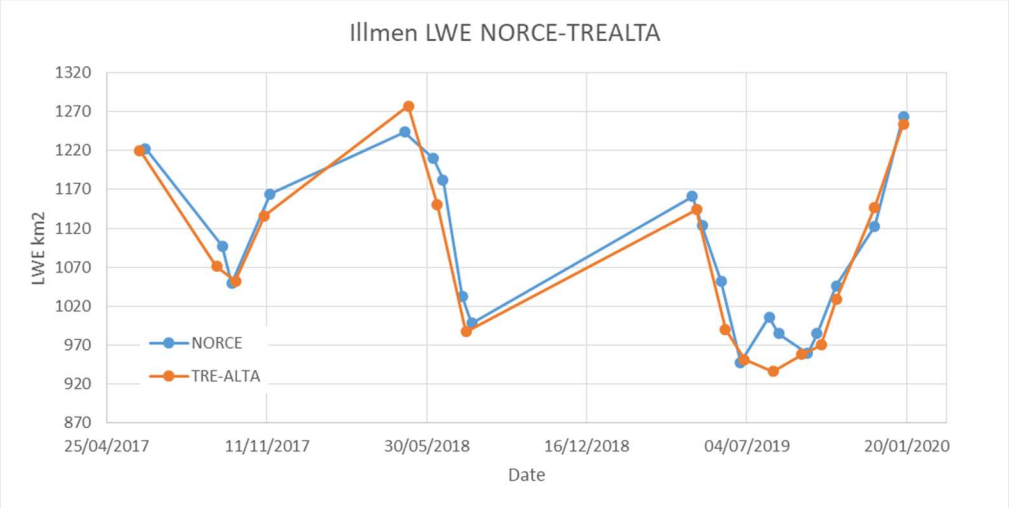


Figure 56. NORCE and TRE-ALTA 12 days maximum difference LWE time series for Lake Illmen.

3.3.2.3. Bosten

NORCE's and TRE-ALTA's lake water mask and their corresponding contingency matrices for a couple of coincident dates (20170729, 20190414 are depicted in Figure 57. The overall agreements for those images are 98.89% and 98.75% respectively.

For lake Bosten, the contingency matrix values for all NORCE's and TRE-ALTA's results corresponding to images presenting a temporal difference lower or equal to 12 days are represented in Table 16.

A representation of the temporal evolution of all NORCE's and TRE-ALTA's LWE is depicted in Figure 58. The same representation but just focusing on the 12 days difference images is depicted in Figure 59.

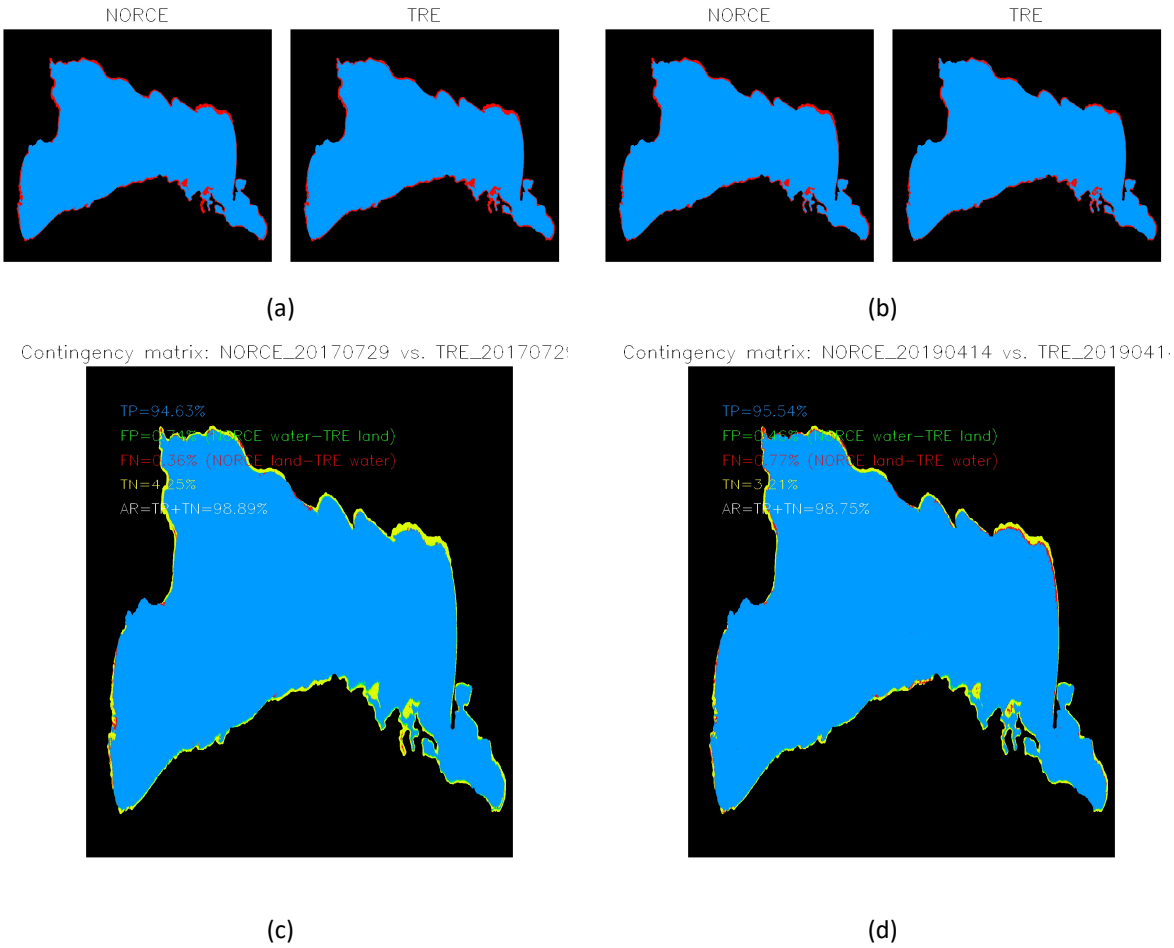


Figure 57. (a) NORCE and TRE-ALTA 20170729 lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b) NORCE and TRE-ALTA 20190414 lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

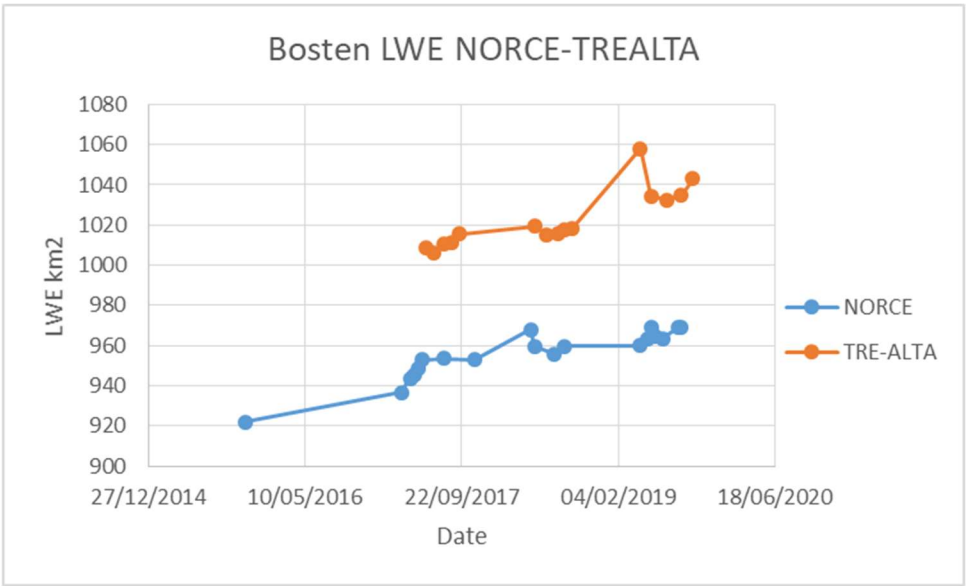


Figure 58. NORCE and TRE-ALTA LWE time series for lake Bosten.

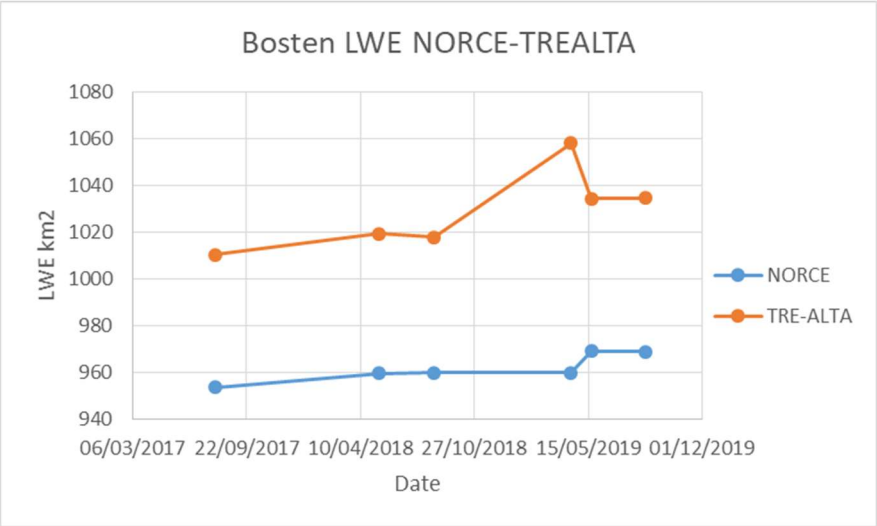


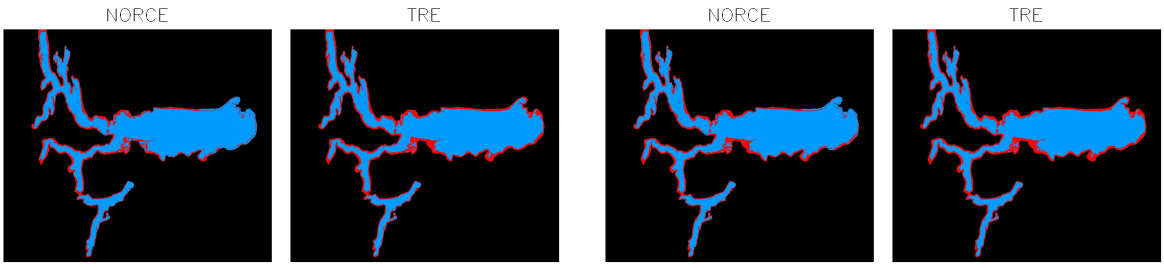
Figure 59. NORCE and TRE-ALTA 12 days maximum difference LWE time series for lake Bosten.

3.3.2.4. Argentino

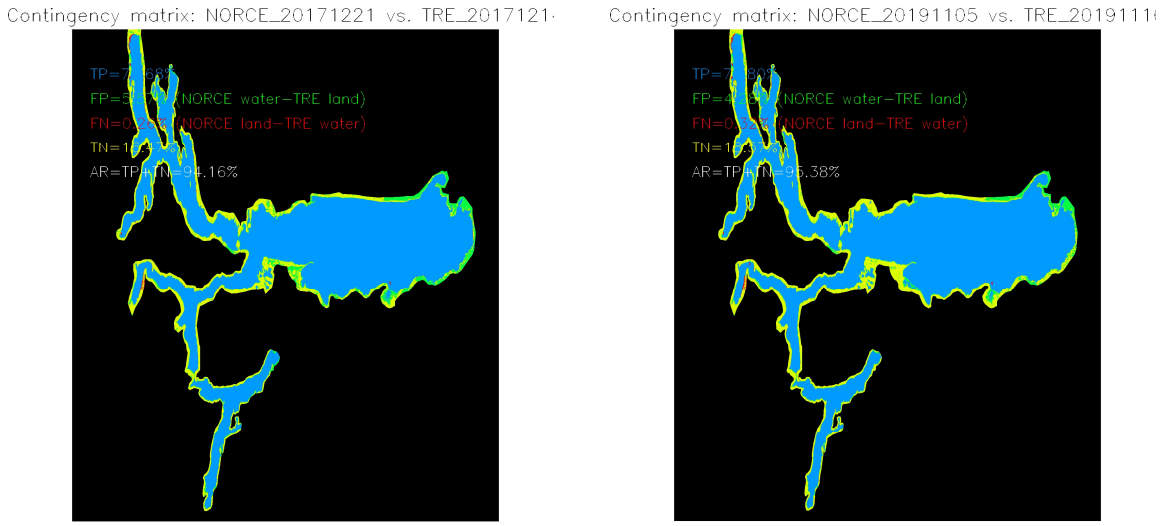
NORCE's and TRE-ALTA's lake water mask and their corresponding contingency matrices for a 7 days difference couple (20171221-20171214) and a 5 days difference one (20191105-20191110) are depicted in Figure 60. The overall agreement for those images are 94.16% and 95.38% respectively.

For lake Argentino, the contingency matrix values for all NORCE's and TRE-ALTA's results corresponding to images presenting a temporal differences lower or equal to 12 days are represented in Table 18.

A representation of the temporal evolution of all NORCE's and TRE-ALTA's LWE is depicted in Figure 61. The same representation but just focusing on the 12 days difference images is depicted in Figure 62.



(a) (b)



(c) (d)

Figure 60. (a) NORCE and TRE-ALTA (20171221-20171214) lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b) NORCE and TRE-ALTA (20191105-20191110) lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

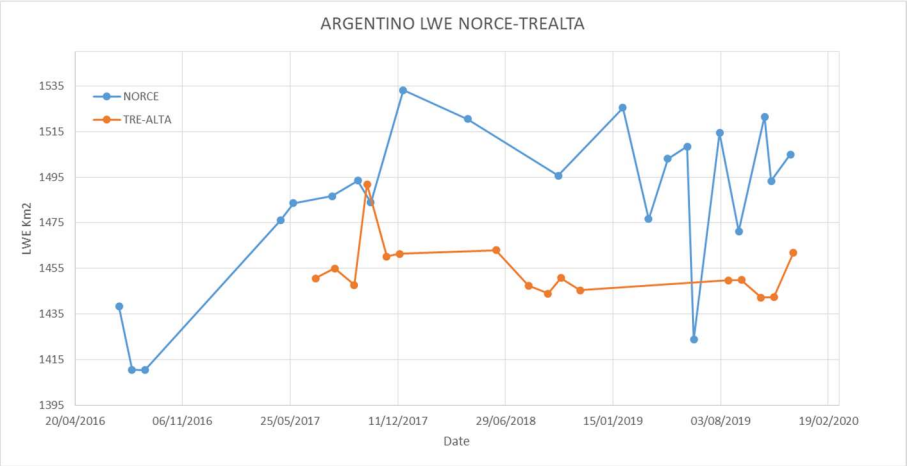


Figure 61. NORCE and TRE-ALTA LWE time series for lake Argentino.

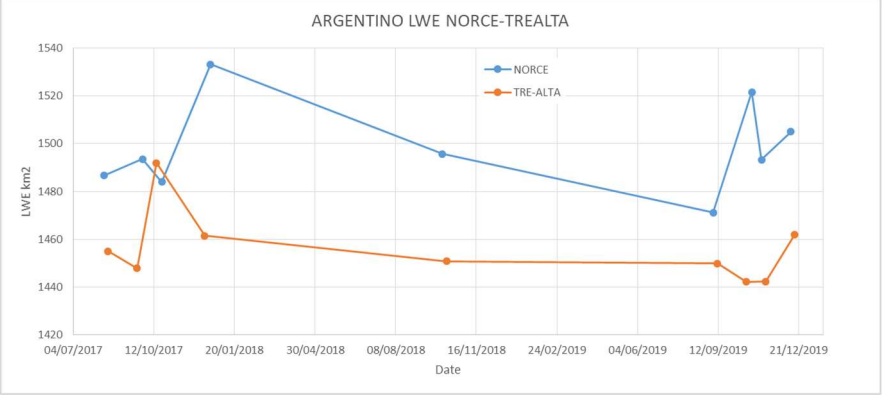


Figure 62. NORCE and TRE-ALTA 12 days maximum difference LWE time series for lake Argentino.

3.3.2.5. Khanka

NORCE's and TRE-ALTA's lake water mask and their corresponding contingency matrices for a couple of coincident dates (20180515, 20190510 are depicted in Figure 63. The overall agreements for those images are 98.83% and 97.62% respectively.

For lake Khanka, the contingency matrix values for all NORCE's and TRE-ALTA's results corresponding to images presenting a temporal difference lower or equal to 12 days are represented in Table 20.

A representation of the temporal evolution of all NORCE's and TRE-ALTA's LWE is depicted in Figure 64. The same representation but just focusing on the 12 days difference images is depicted in Figure 65.

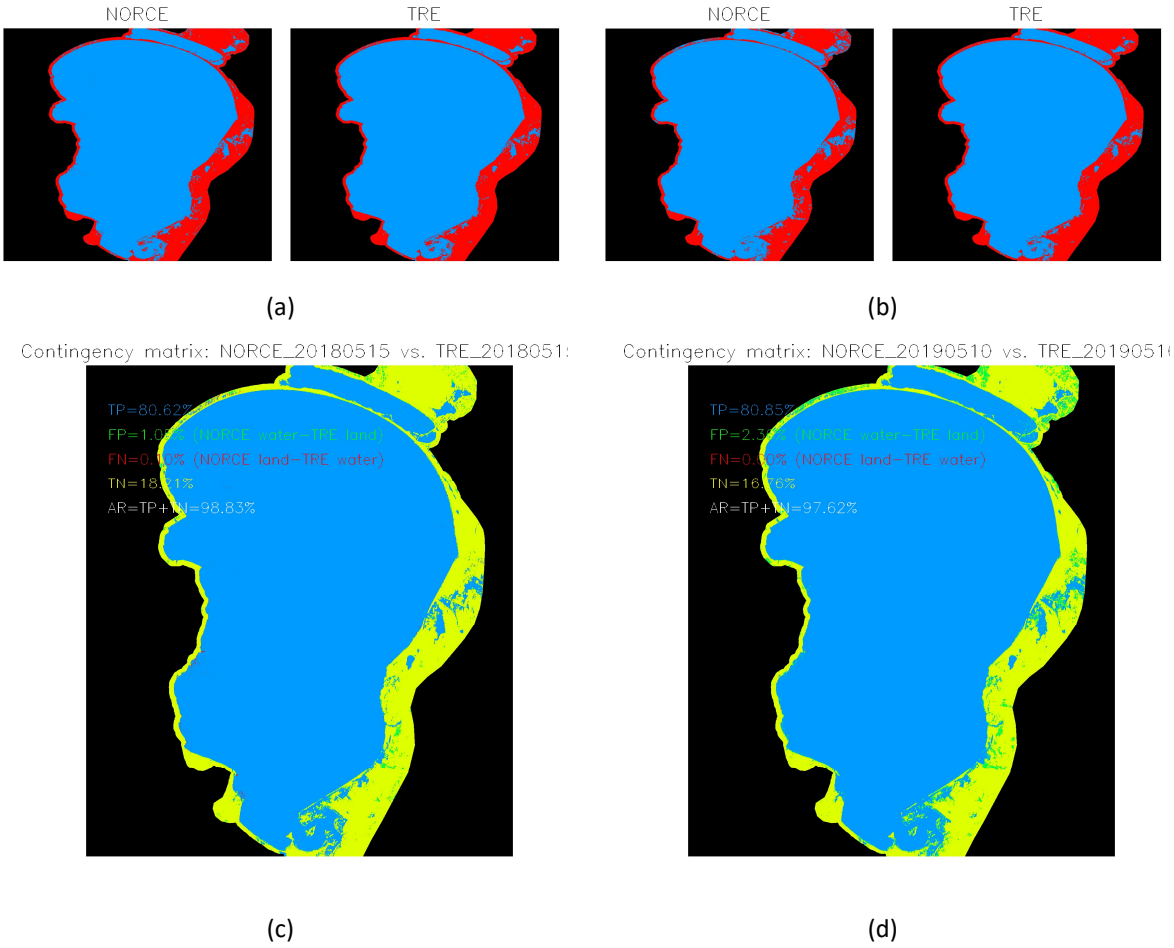


Figure 63. (a) NORCE and TRE-ALTA 20180515 lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b) NORCE and TRE-ALTA 20190510 lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

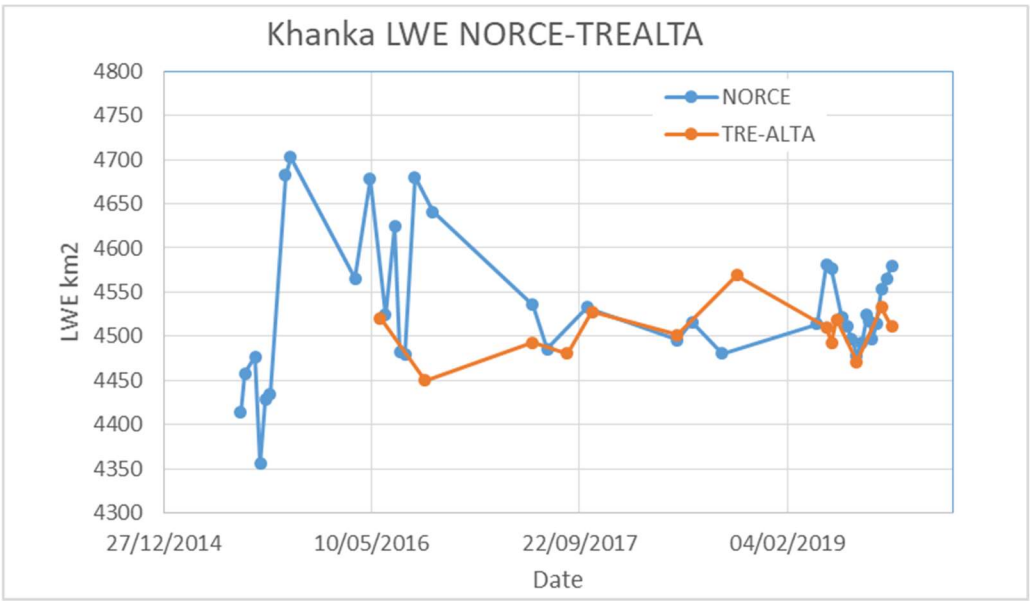


Figure 64. NORCE and TRE-ALTA LWE time series for lake Khanka.

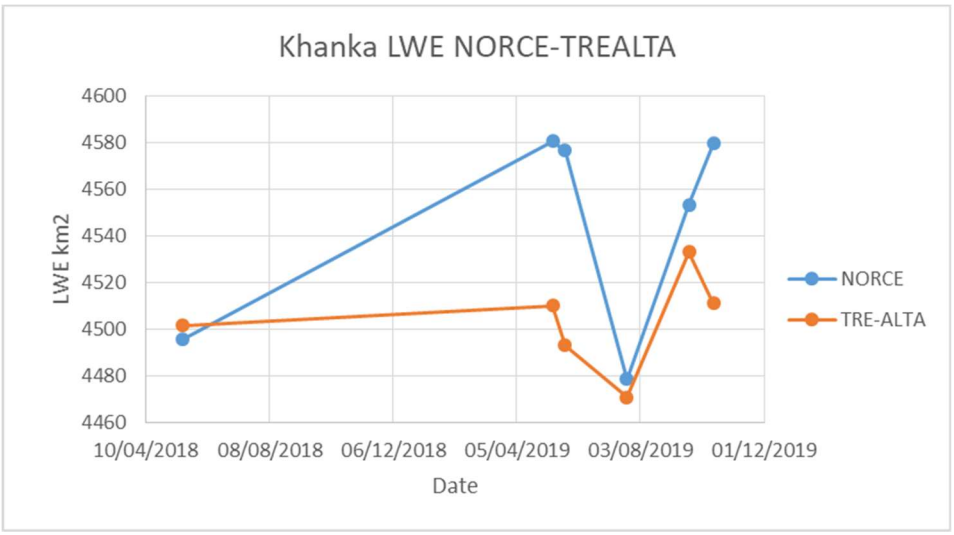


Figure 65. NORCE and TRE-ALTA 12 days maximum difference LWE time series for lake Khanka.

3.3.2.6. Overall comparisons between the two SAR algorithms

In this section we provide tabular contingency matrices for all the lakes that are inter-comparable between the two SAR algorithms. Due to different selections of dates for the two SAR classifiers (NORCE and TRE-ALTA) different dates will be inter-compared. The main criterion for inter-comparison is that the two images (NORCE's and TRE-ALTA's) are close in time. For that, a maximum time difference to 12 days is set.

The following tables have been generated:

- Lake Illmen: contingency matrix in Table 14 and corresponding Area difference mean values in Table 15.
- Lake Bosten: contingency matrix in Table 16 and corresponding Area difference mean values in Table 17.
- Lake Argentino: contingency matrix in Table 18 and corresponding Area difference mean values in Table 19.

- Lake Khanka: contingency matrix in Table 20 and corresponding Area difference mean values in Table 21.
- Mean Values of all previous: Table 22

Tables' acronyms stand for: TP (True Positive in %), FP (False Positive in %), FN (False Negative in %), TN (True Negative in %), AR (Accuracy Rate in %), Td (time difference in days), Ad (Area LWE difference in %) and Ard (Area LWE difference in square kilometers).

After the comparison results, it can be stated the performance for the two SAR results is in general very similar, although differences are observed due to different dates/lakes and slight differences in methodology.

Table 14. NORCE and TRE-ALTA contingency matrix values summary for Lake Illmen

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Illmen	20170612	20170605	67.38	3.03	0.11	29.47	96.85	7	4.15	2.27
Illmen	20170916	20170909	59.83	3.37	0.25	36.52	96.36	7	4.92	25.97
Illmen	20170928	20171003	58.61	1.82	0.63	38.92	97.54	5	1.95	3.25
Illmen	20171115	20171108	63.48	3.58	0.04	32.88	96.36	7	5.28	28.01
Illmen	20180502	20180507	69.55	2.10	0.31	28.03	97.58	5	2.49	33.16
Illmen	20180607	20180612	64.11	5.58	0.00	30.29	94.40	5	7.99	59.6
Illmen	20180619	20180612	64.03	4.08	0.08	31.79	95.82	7	5.87	32.31
Illmen	20180713	20180718	55.73	3.77	0.16	40.33	96.06	5	6.06	45.68
Illmen	20180725	20180718	55.69	1.82	0.19	42.27	97.97	7	2.83	11.32
Illmen	20190427	20190502	64.06	2.82	0.13	32.97	97.03	7	4.03	16.86
Illmen	20190509	20190502	63.16	1.55	1.03	34.25	97.41	7	0.80	20.91
Illmen	20190602	20190607	56.13	4.49	0.01	39.35	95.49	5	7.39	61.91
Illmen	20190626	20190701	53.59	1.01	0.39	44.99	98.58	5	1.14	3.38
Illmen	20190801	20190806	53.16	4.78	0.00	42.04	95.21	5	8.25	69.85
Illmen	20190813	20190806	53.16	3.57	0.00	43.24	96.41	7	6.29	48.84
Illmen	20190918	20190911	54.02	1.22	0.36	44.37	98.40	7	1.55	0.82
Illmen	20190930	20191005	54.58	2.15	0.41	42.83	97.42	7	3.06	14.52
Illmen	20191024	20191024	58.16	2.12	0.14	39.56	97.73	0	3.27	17.92
Illmen	20191211	20191211	61.77	2.90	1.09	34.21	95.99	0	2.80	24.6
Illmen	20200116	20200116	69.04	3.76	0.04	27.15	96.19	0	5.10	9.67

Table 15. NORCE and TRE-ALTA mean and stdv LWE difference for Lake Illmen

	Mean(Ar)	Stdv(Ar)	Mean(Ard)	Stdv(Ard)
Illmen	96,74	1,09	26,54	20,97

Table 16. NORCE and TRE-ALTA contingency matrix values summary for Lake Bosten

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Bosten	20170729	20170729	94.63	0.74	0.36	4.25	98.89	0	0.40	56.70
Bosten	20180513	20180513	95.12	0.85	0.32	3.69	98.81	0	0.55	59.68
Bosten	20180817	20180817	95.25	0.73	0.39	3.61	98.87	0	0.35	58.06
Bosten	20190414	20190414	95.54	0.46	0.77	3.21	98.75	0	0.32	98.22
Bosten	20190520	20190520	95.83	1.08	0.26	2.81	98.64	0	0.84	65.30
Bosten	20190824	20190824	95.95	0.94	0.20	2.90	98.85	0	0.76	65.78

Table 17. NORCE and TRE-ALTA mean and stdv LWE difference for Lake Illmen

	Mean(Ar)	Stdv(Ar)	Mean(Ard)	Stdv(Ard)
Bosten	98.80	0.09	67.29	15.61

Table 18. NORCE and TRE-ALTA contingency matrix values summary for Lake Argentino

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Argentino	20170811	20170816	76.49	4.24	1.11	18.14	94.63	5	3.86	31.7
Argentino	20170928	20170921	76.62	4.47	0.74	18.15	94.78	7	4.60	45.78
Argentino	20171022	20171015	76.99	3.59	2.31	17.09	94.09	7	1.58	7.94
Argentino	20171221	20171214	77.68	5.57	0.26	16.47	94.16	7	6.37	71.85
Argentino	20181005	20181010	76.56	4.65	0.83	17.94	94.51	5	4.70	44.94
Argentino	20190906	20190911	76.29	3.59	1.16	18.93	95.23	5	3.03	21.32
Argentino	20191024	20191017	76.76	5.86	0.32	17.04	93.81	7	6.71	79.42
Argentino	20191105	20191110	76.80	4.28	0.32	18.57	95.38	5	4.88	51.08
Argentino	20191211	20191216	77.78	3.94	0.46	17.80	95.58	5	4.25	43.09

Table 19. NORCE and TRE-ALTA mean and stdv LWE difference for Lake Argentino

	Mean(Ar)	Stdv(Ar)	Mean(Ard)	Stdv(Ard)
Argentino	94.68	0.61	44.12	22.50

Table 20. NORCE and TRE-ALTA contingency matrix values summary for Lake Khanka

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Khanka	20170601	20170601	80.58	1.83	0.02	17.56	98.14	0.00	2.19	6.01
Khanka	20180515	20180515	80.62	1.05	0.10	18.21	98.83	0.00	1.15	70.53
Khanka	20190510	20190510	80.85	2.36	0.00	16.76	97.62	0.00	2.83	83.5
Khanka	20190522	20190522	80.59	2.56	0.00	16.84	97.43	0.00	3.07	7.72
Khanka	20190721	20190721	80.21	1.15	0.01	18.61	98.83	0.00	1.40	20.23

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Khanka	20190919	20190919	81.24	1.48	0.04	17.22	98.46	0.00	1.73	68.48
Khanka	20191013	20191013	80.88	2.32	0.00	16.78	97.66	0.00	2.78	6.01

Table 21. NORCE and TRE-ALTA mean and stdv LWE difference for Lake Khanka

	Mean(Ar)	Stdv(Ar)	Mean(Ard)	Stdv(Ard)
Khanka	98.13	0.58	42.74	35.15

Table 22. Mean Value for all previous contingency matrix AR, Td and Ad values

	AR	Ard	Td	Ad
Average	96.87	37.95	4.8	3.75

3.3.3. Comparison of optical and SAR sensors-based approaches

Similarly to what has been done in the previous section with the two SAR based methodologies, a comparison between the SAR and the optical LWE approaches is now done. On the SAR side Sentinel-1 and Envisat-SAR have been employed while on the optical side Sentinel2 and Landsat series have been used.

Temporal LWE series are generated and compared and for some of the lakes the contingency matrix is calculated whenever having images with a temporal difference equal or lower than 12 days.

3.3.3.1. Altevatnet

Comparison between SAR and optical LWEs have been carried out over more than 10 pairs of acquisitions, thanks to the high latitudes, the revisit is very high.

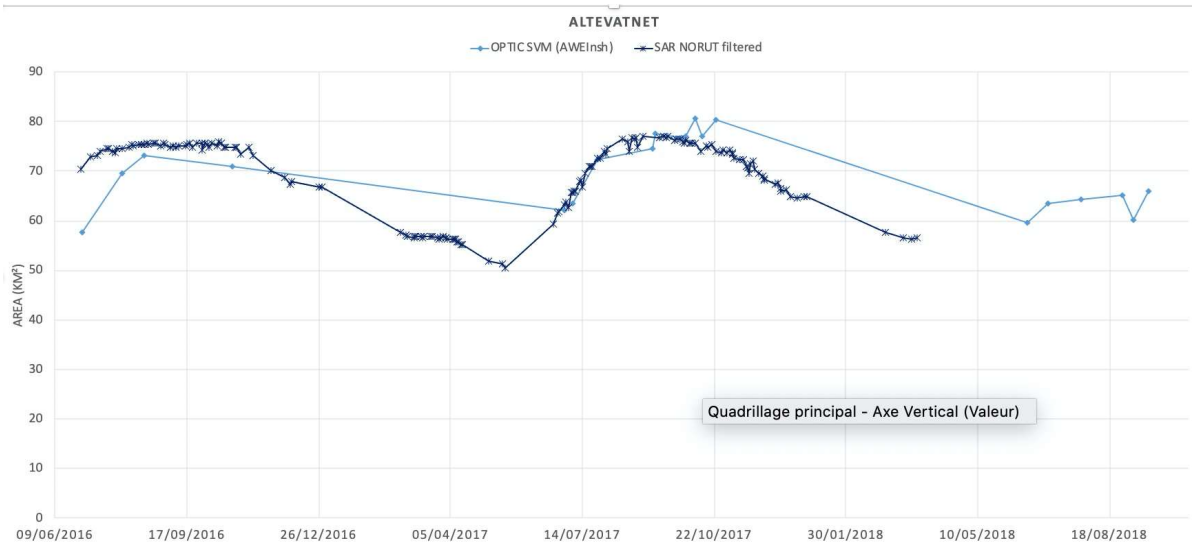


Figure 66. Comparison SAR and optical LWEs derived over Altevatnet

The value of the derived LWEs from optical and SAR are very coherent between them, with a correlation of 0.97

	S2 LWE product	
S1, LWE product	S2, water	S2.land
S1, water	82.96 % (TP)	2.35% (FP)
S1, land	1.43 % (FN)	13.24% (TN)

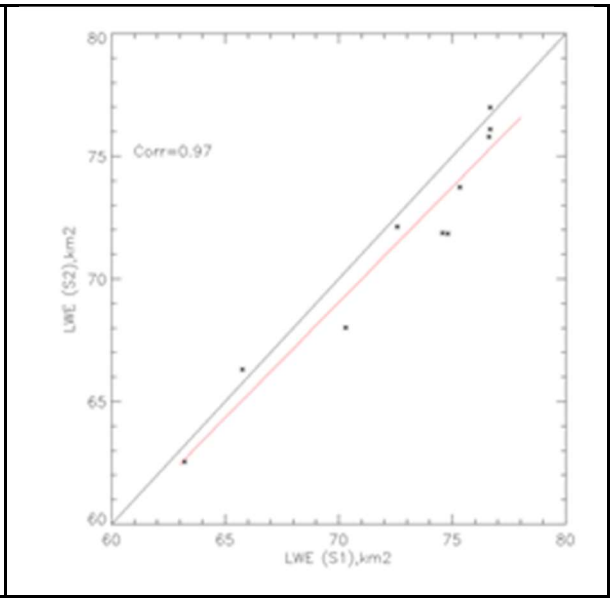


Figure 67. Analysis over 12 pairs of Sentinel1 and Sentinel2 of the commission and omission rates.

Of course, over this Nordic landscape, the observed commission and omission are the expected ones and are related to:

- Presence of cloud on optical imagery.
- Sensitivity to windy conditions and ice presence in SAR.

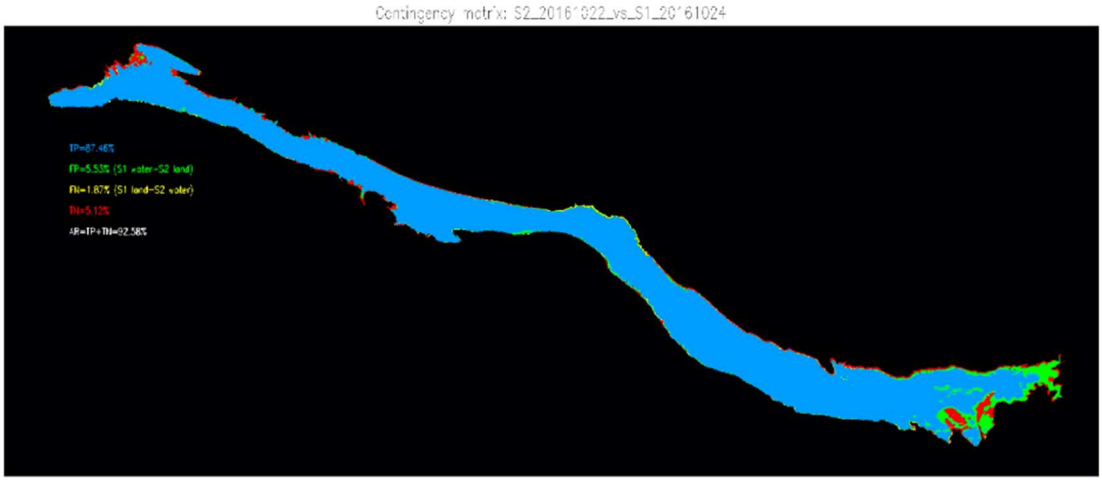


Figure 68. Analysis of the omission and commission between SAR and optical approaches exploiting Sentintel1 and Sentinel2.

3.3.3.2. Colhue

The period with dual observation is relatively limited, the trends are similar but with apparently an overestimation of water surface by SAR. This could be related to the nearby environment, i.e. desert and also sandy shore.

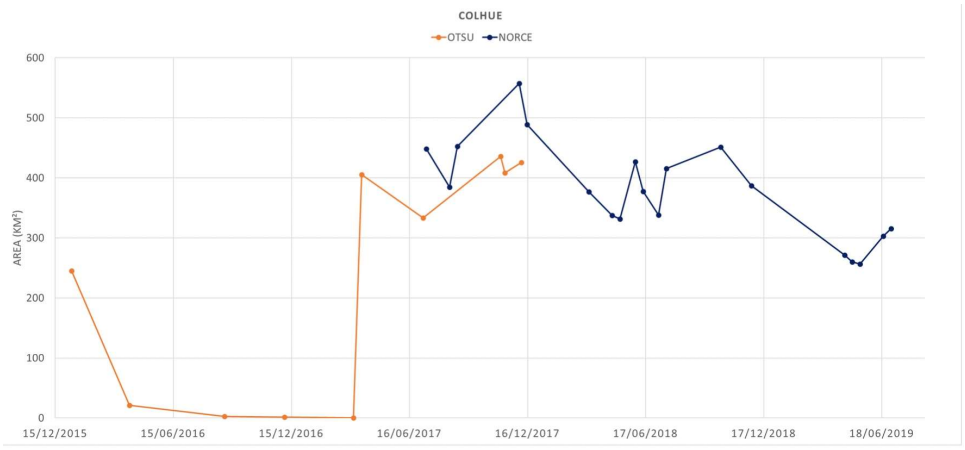


Figure 69. Comparison SAR and optical LWEs derived over Colhue Lake.

3.3.3.3. Al Hamar wetland

There are important differences between SAR and optical observations over the complex area of Al Hamar wetlands. Whereas optical imagery allows us to monitor the increase of water surfaces, SAR LWEs presents a drop at the beginning of the series, and then stay at a lower level than optical. This is due to the facts:

- Optical LWE integrates a part of flooded vegetation whereas SAR retrieves open water surface.
- SAR is more sensitive to the environment, i.e. sandy flat area, than optical sensors.

Figure 70. Comparison SAR and optical LWEs derived over Al-Hammar.

3.3.3.4. Sassykol and Alakol

This case of study is particularly interesting, as it is two neighbouring lakes being connected through a wetland complex. In addition, Sassykhol is a relatively shallow lake, with surrounding wetland, whereas Alakol is a deep lake with a more classical shape.

Over Alakol lake, the difference of LWE between SAR derived LWE and Optical ones, is about 3% that is quite satisfactory.

Therefore, over Sasykholol, the difference is about 14%, this is related to the fact large part of the increase of water surface is related to the flooding of wetland, which is not observed by SAR.

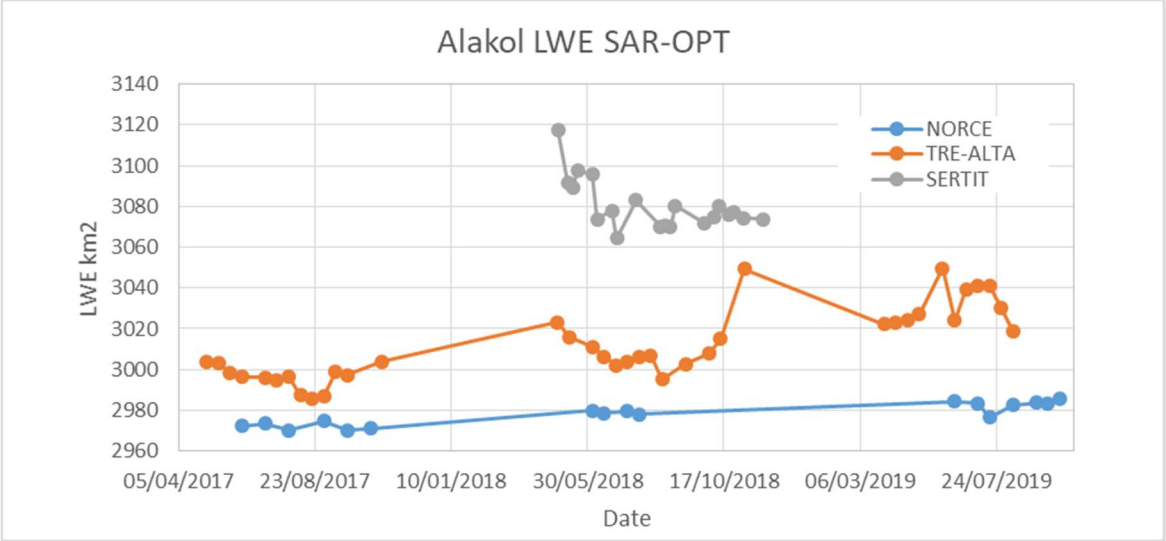
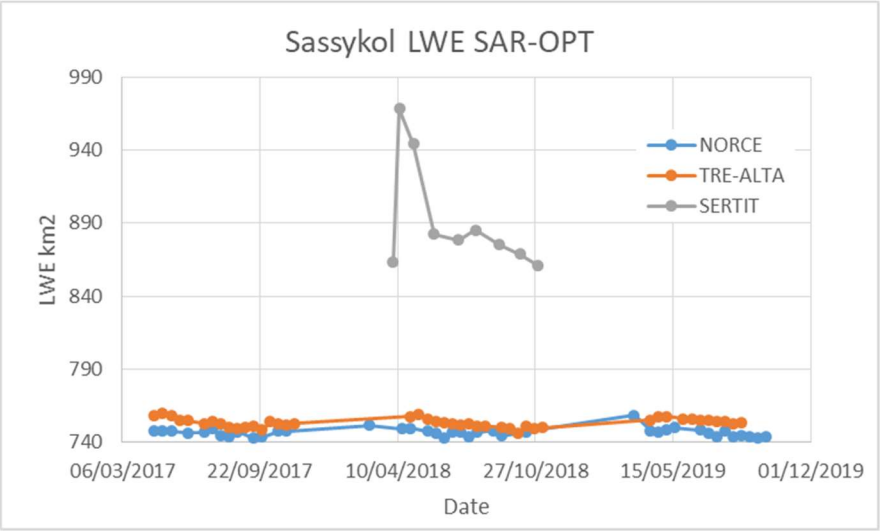


Figure 71. Comparison SAR and optical LWEs derived over Sassykol and Alakol lakes.



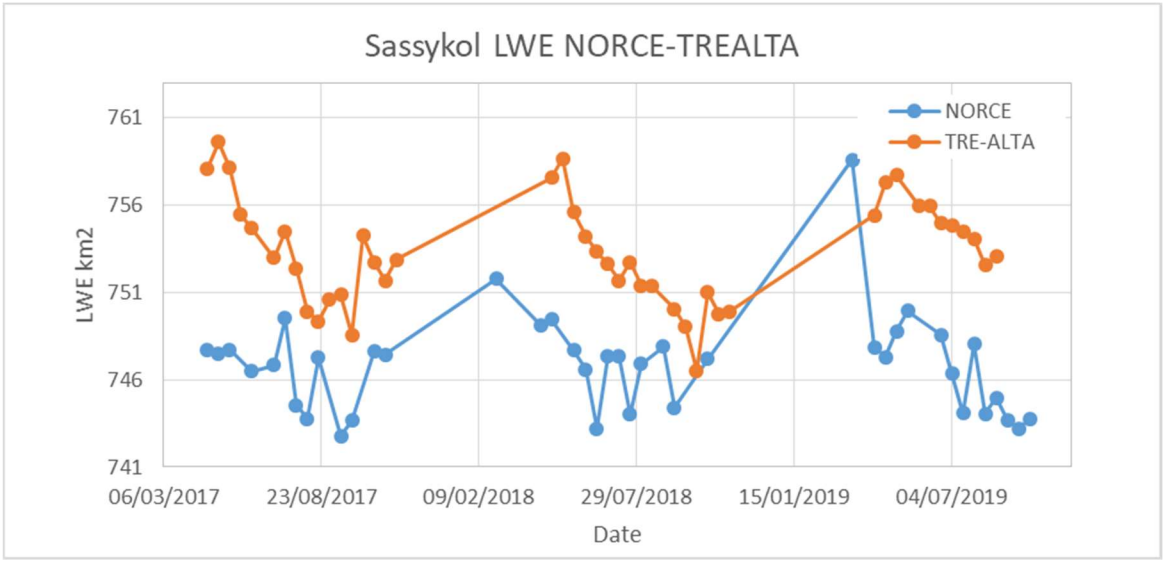


Figure 72. Comparison SAR and optical LWEs derived over Sassykol and Alakol lakes.

3.3.3.5. Chad

There were near simultaneous acquisition of Sentniel1 SAR image and Sentinel2 optical images on the 24th of April 2019. This gives the opportunity to compare the capabilities and limitations of the two system.

Within the common AOI defined there were 14.68 km² detected as water in Sentinel 1 and none water in Sentinel 2. These areas correspond partially to very small and located clouds that have intercepted the optical signal. In this case SAR, and not the optical sensor, were closest to the truth. Therefore, most of the "SAR water alone" correspond to commission with bare soil on the NE part of the study area.

At the opposite, 40.75 km² detected as water on Sentinel 2 and no water on Sentinel 1. These areas correspond mostly to shoreline of the lake, plus some floating islands.

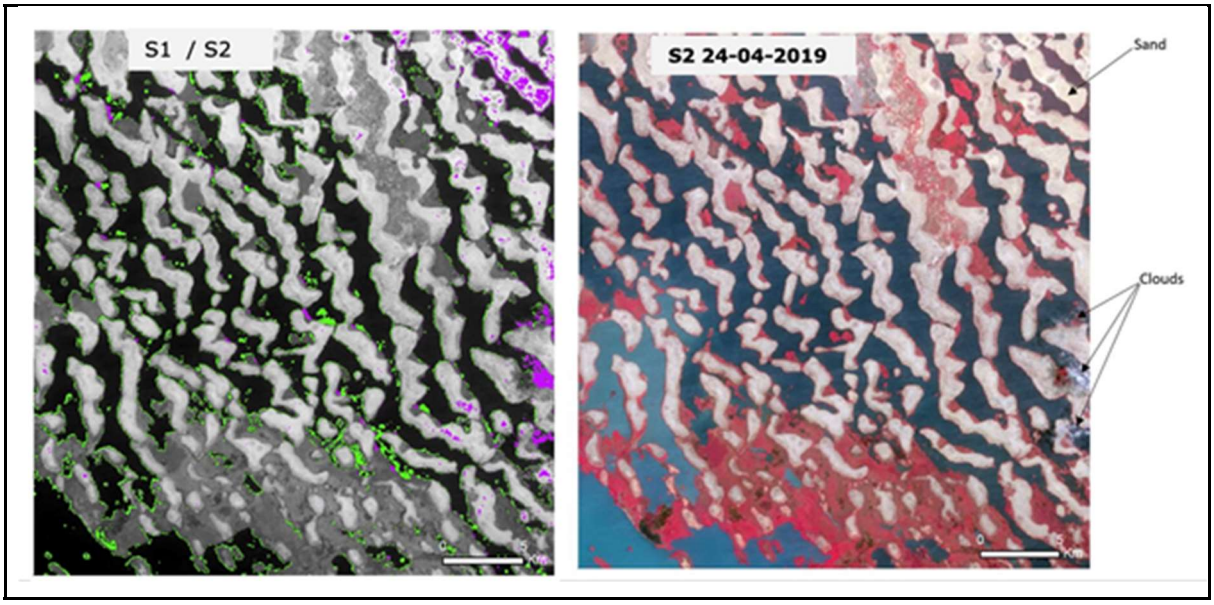


Figure 73. Comparison of the LWE derived from SAR and optical imagery on the 24 of April 2019: in Purple, detected as water on S1 and none water based on S2; in Green detected as water on S2 and none water on S1.

3.3.3.6. Chilwa lake

The difference between optical LWE and SAR classification is about 20%. This is related to the fact that large parts of the increase in water surface is related to the flooding of wetland, which is not observed by SAR. In this specific complex context, the SAR extraction seems underestimated over wetlands.

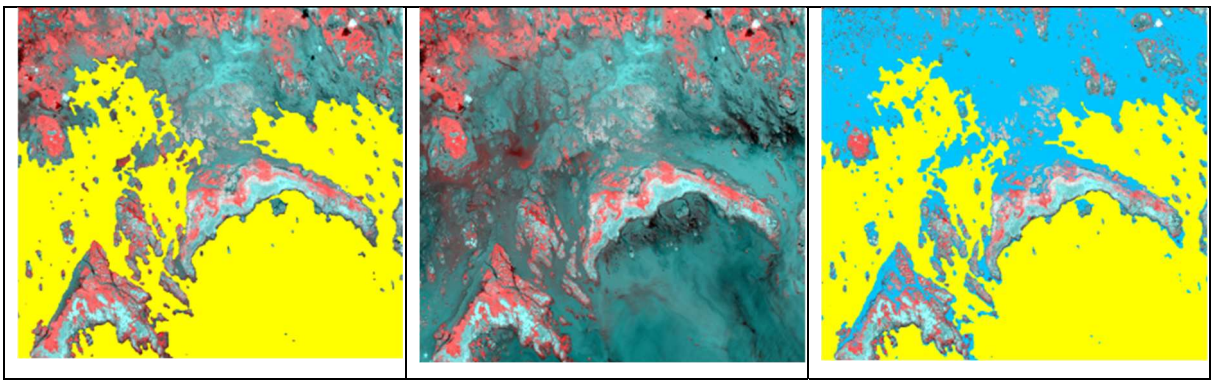


Figure 74. Comparison of SAR and optical LWEs; in yellow LWE derived from SAR with in background the Sentinel2, False colour Sentintel2 acquired the 24/04/2018; in blue the additional water surfaces derived from a Sentinel-2 image with the SVM algorithm.

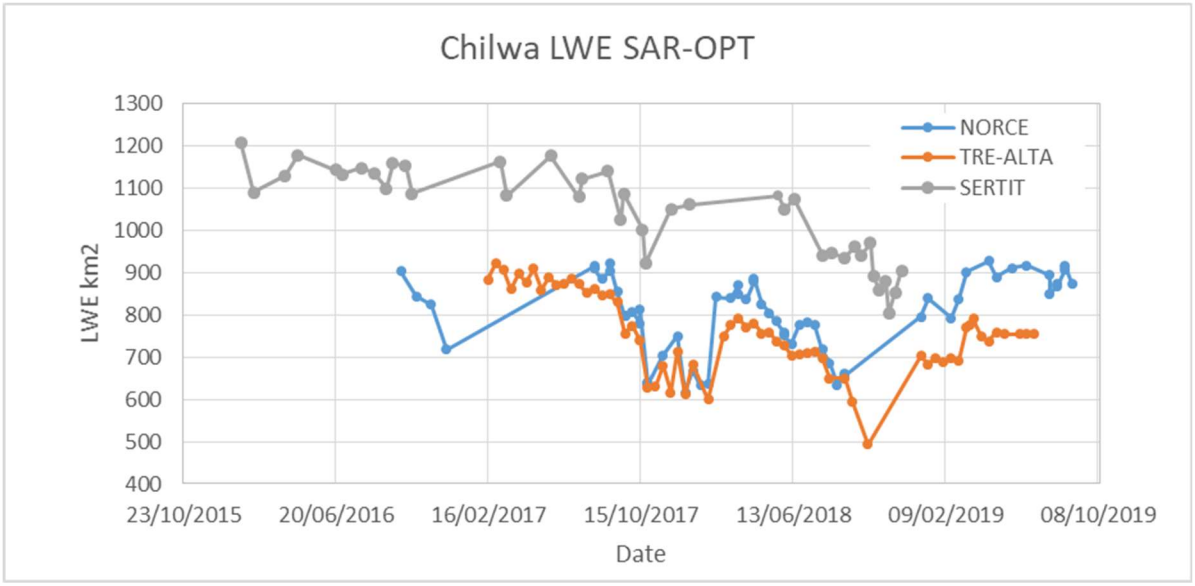


Figure 75. Comparison SAR and optical LWEs derived over Chilwa lake.

3.3.3.7. Argentino

NORCE-SERTIT and TRE-ALTA-SERTIT lake water mask and their corresponding contingency matrices for a couple of couple of images presenting a time difference smaller than 12 days (20171022-20171015 and 2071015-20171020 respectively) are depicted in Figure 76. The overall agreements for those images are 96.66% and 95.82% respectively.

For Lake Argentino, the contingency matrix values for all NORCE and SERTIT LWE and TRE-ALTA and SERTIT corresponding to images with a temporal difference lower or equal to 12 days are represented in Table 24 and Table 29 respectively.

A representation of the temporal evolution of all NORCE, TRE-ALTA and SERTIT LWE is depicted in Figure 77. The same representation but just focusing on the 12 days difference images is depicted in Figure 78.

Argentino is a lake surrounded by steeply mountains. The shape of the lake is defined by narrow valleys. Due to the altitude, the lake is often roughened by wind, and this is a challenge for the SAR retrieval of the LWE. When comparing LWEs derived from Sentinel1 and Sentinel2 acquired close in time, the differences are:

- Commission errors for SAR on lakeshore, particularly the Eastern parts of the lake, where the sandy, muddy shores are relative flat and soft surfaces that has a radar signature similar to water.
- Some SAR omission-errors that could be related to windy water surface, but also in the north branch of the lake, to some commission errors from the glacier terminus or lake ice.

Depending on the applied processing, the difference in terms of detected water surface can be significant (more than 5%) or relatively small (less than 1% - 3.3%).

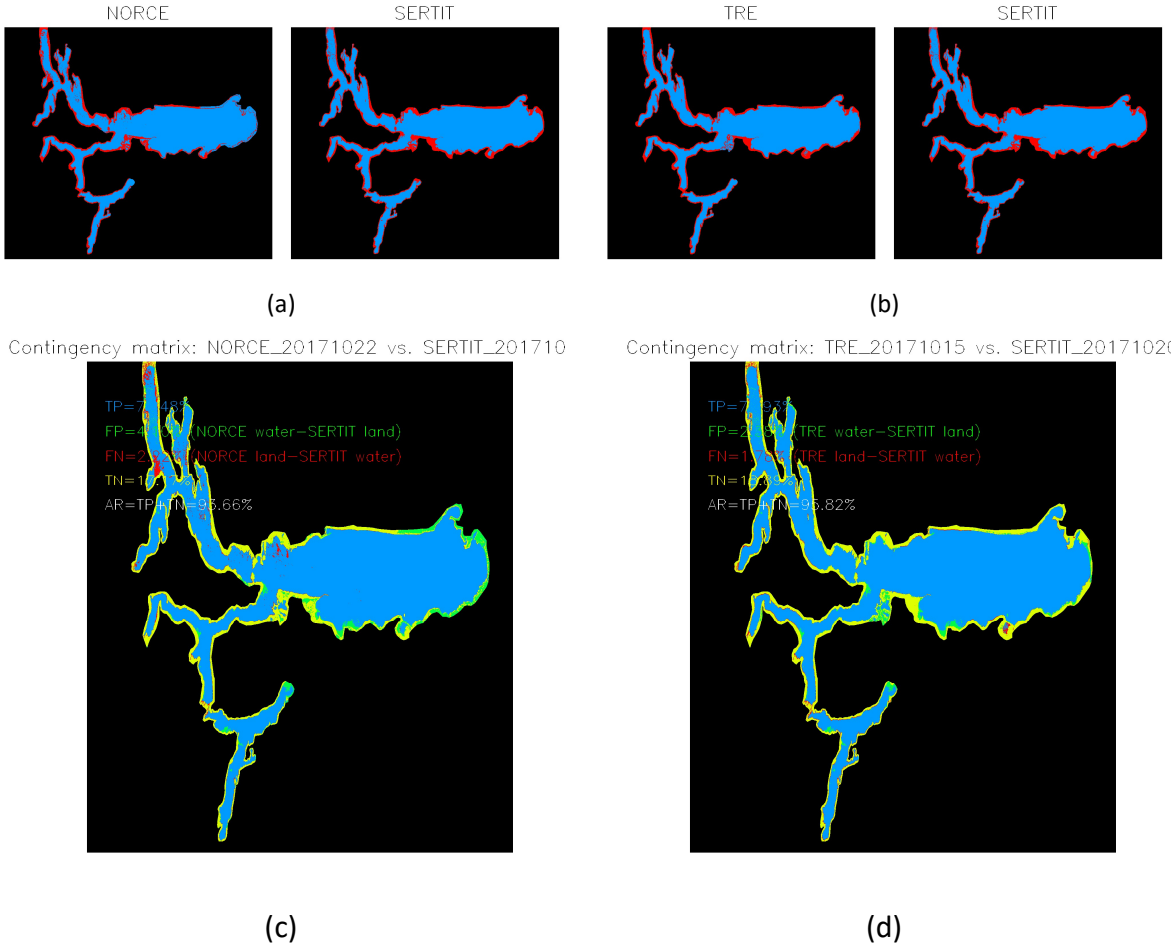


Figure 76. (a) NORCE-SERTIT (20171022-20171015) lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b) TRE-ALTA-SERTIT (2071015-20171020) lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

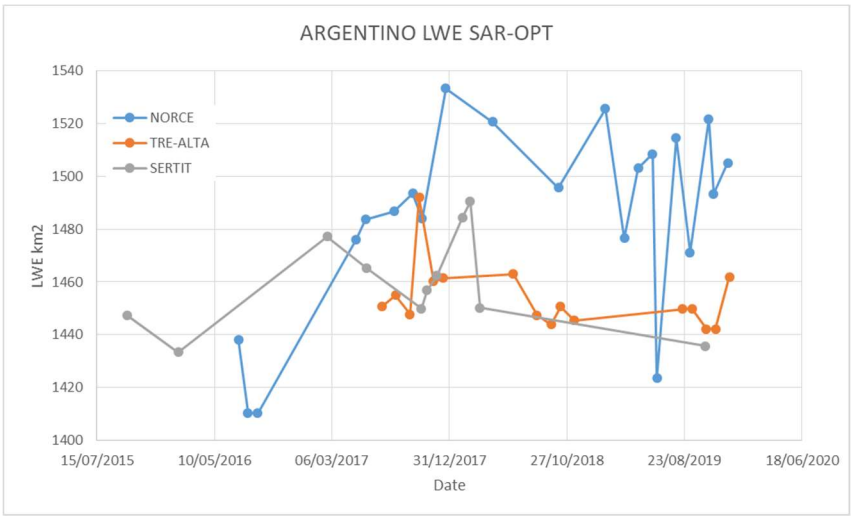


Figure 77. NORCE SAR, TRE-ALTA SAR and SERTIT LWE time series for lake Argentino.

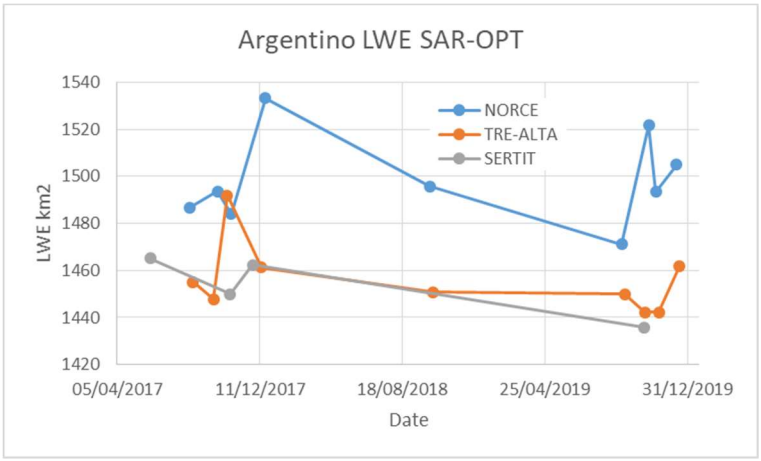


Figure 78. NORCE SAR, TRE-ALTA SAR and SERTIT Optical LWE time series for lake Argentino.

3.3.3.8. Illmen

In the Illmen lake case no temporally close images were processed among the three groups. In Figure 79, the three LWE series are depicted. From it, it is hard to evaluate SAR and Optical relative performances.

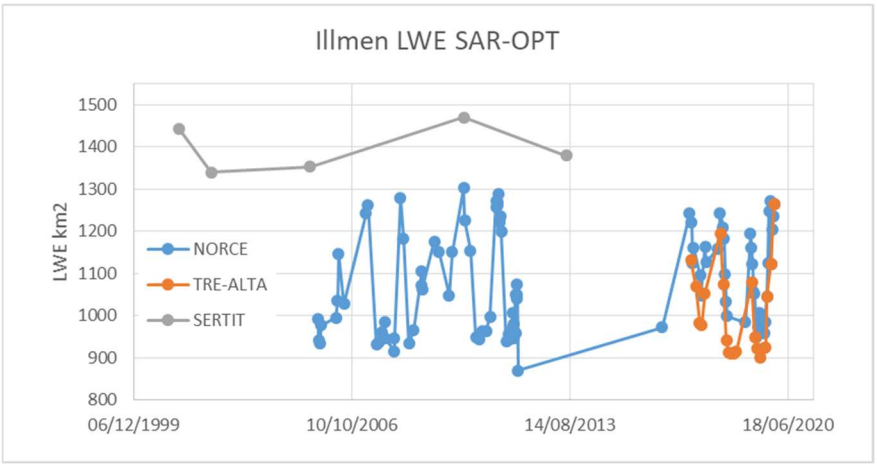


Figure 79. NORCE SAR, TRE-ALTA SAR and SERTIT Optical LWE time series for lake Illmen.

3.3.3.9. Bosten

NORCE-SERTIT and TRE-ALTA-SERTIT lake water mask and their corresponding contingency matrices for a couple of couple of images with a time difference smaller than 12 days (20170729-20170728 for both cases) are depicted in Figure 80. The overall agreement for those images are 98.02% and 97.65%, respectively.

The systematic bias between optical and SAR LWE-estimates is 4% to 8% depending on the SAR approach. This is related to the fact that large part of the increase in water extent occurs in wetlands east of the main lake, and SAR has limited capabilities in resolving the water in such wetlands. The bias between the NORCE and the TRE ALTAMIRA approach for Bosten is related to the use of different masks.

For Lake Bosten, the contingency matrix values for all NORCE and SERTIT LWE and TRE-ALTA and SERTIT corresponding to images with a temporal difference lower or equal to 12 days are represented in Table 23 and Table 28 respectively.

The temporal evolution of all NORCE, TRE-ALTA and SERTIT LWE-estimates are depicted in Figure 81. The same representation but just focusing on the 12 days difference images is depicted in Figure 82.

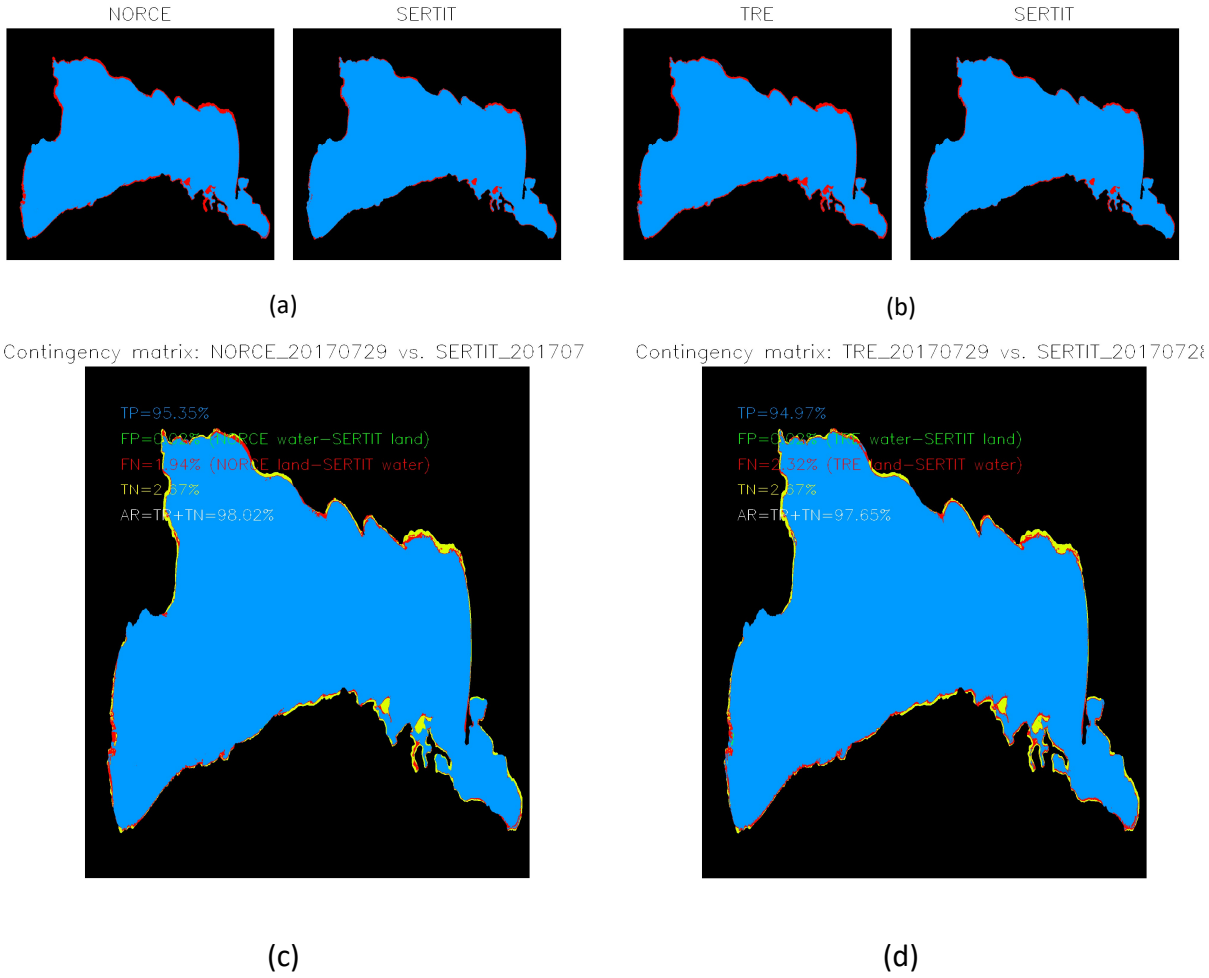


Figure 80. (a) NORCE and SERTIT (20170729-20170728) lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b) TRE-ALTA and SERTIT (20170729-20170728) lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

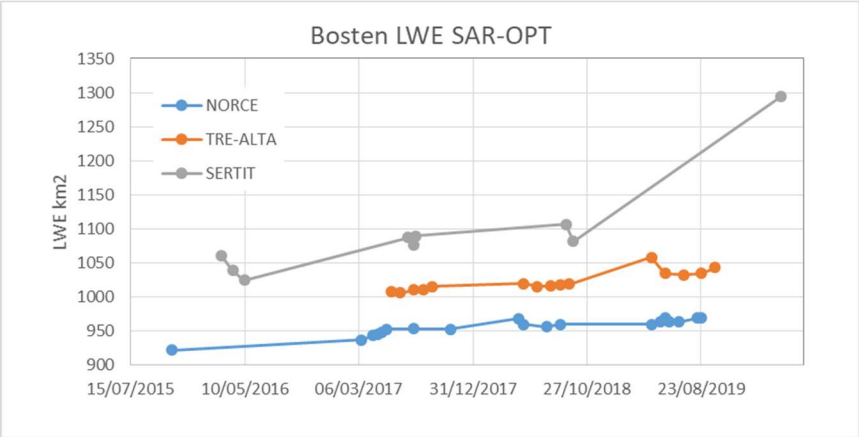


Figure 81. NORCE SAR, TRE-ALTA SAR and SERTIT LWE time series for lake Bosten.

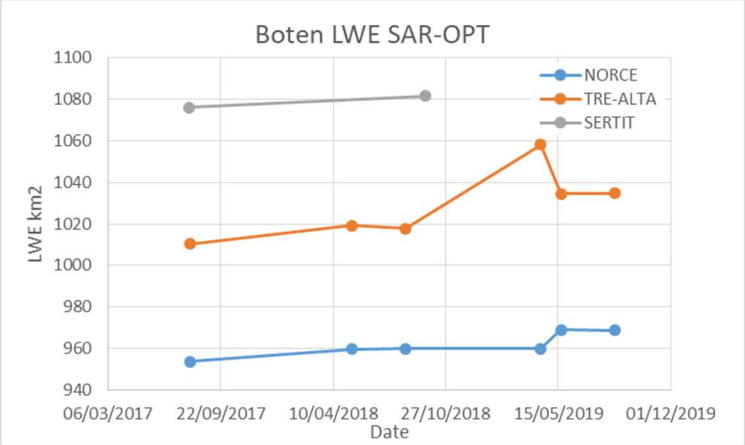


Figure 82. NORCE SAR, TRE-ALTA SAR and SERTIT Optical LWE time series for lake Bosten.

3.3.3.10. Lake Khanka

NORCE-SERTIT and TRE-ALTA-SERTIT lake water mask and their corresponding contingency matrices for a couple of couple of images with a time difference smaller than 12 days (20180515-20180510, 20180515-20180510 respectively) are depicted in Figure 83. The overall agreement for those images are 98.18% and 98.34%, respectively.

For Lake Khanka, the contingency matrix values for all NORCE and SERTIT LWE and TRE-ALTA and SERTIT corresponding to images with a temporal difference lower or equal to 12 days are represented in Table 25 and Table 30Table 28 respectively.

The temporal evolution of all NORCE, TRE-ALTA and SERTIT LWE-estimates are depicted in Figure 84. The same representation but just focusing on the 12 days difference images is depicted in Figure 85.

The difference between SAR and optical LWE is about 50 km^2, which represents about 1.1% of the water lake extent. It is interesting to compare pair by pair the evolution of commission and omission between the LWEs derived from Sentinel1 and from Sentinel2. Water detection errors in optical imagery is often related to poor cloud discrimination. In these examples, the omission and commission errors are very low. SAR omission, in red in the figure corresponds to vegetated areas along the lake shore, areas than can be inundated. In the pair of the end of July 2019, effects of the presence of a clouds is noticeable, inducing both omission as water is not recognized and commission over the land surface. This example highlights the sensitivity and limitations of exploiting optical imagery for the recognition of water surface due to the presence of clouds.

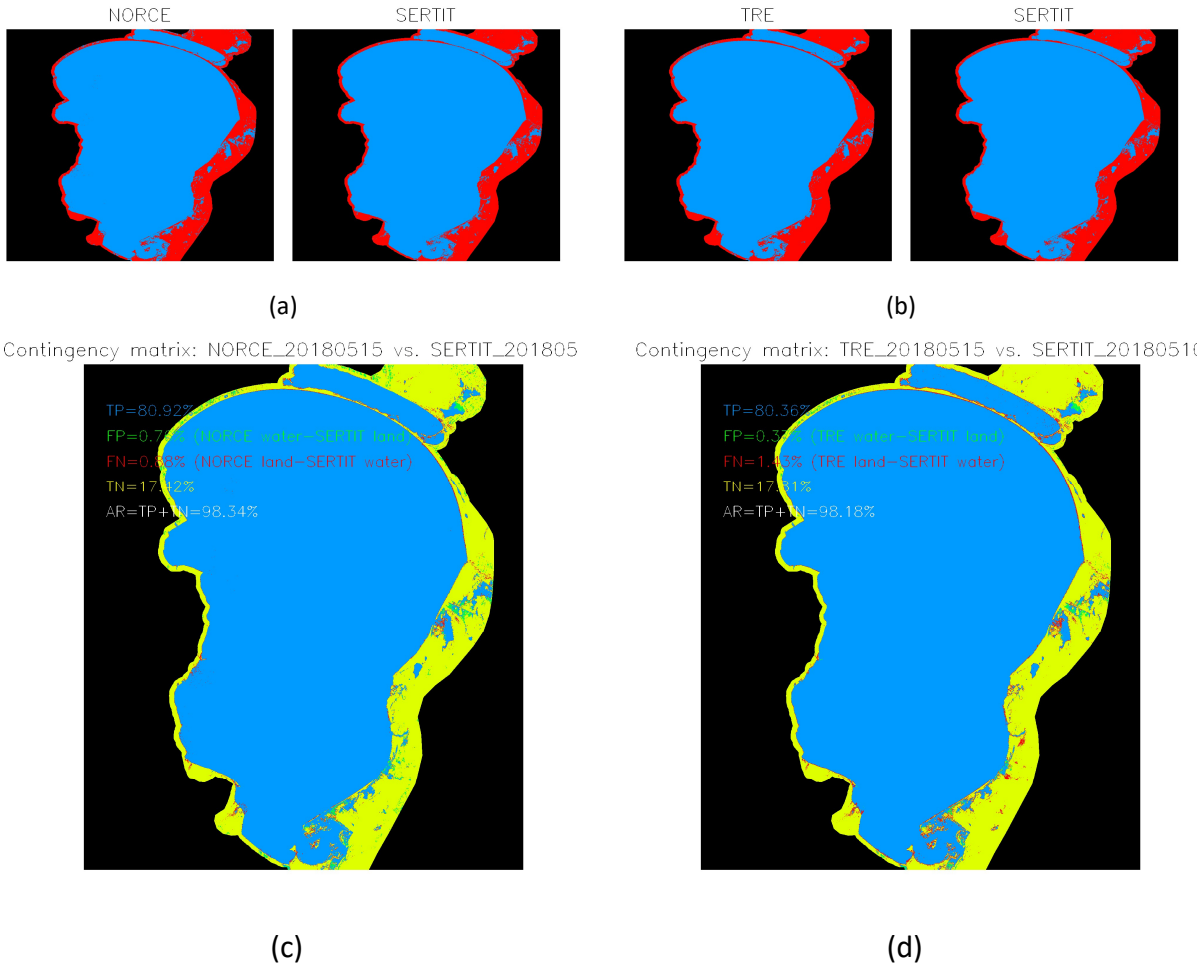


Figure 83. (a) NORCE and SERTIT (20180515-20180510) lake water masks (blue) over a selected analysis polygon (red) and (c) its corresponding contingency matrix. (b)TRE-ALTA- SERTIT (20180515-20180510) lake water masks (blue) over a selected analysis polygon (red) and (d) its corresponding contingency matrix.

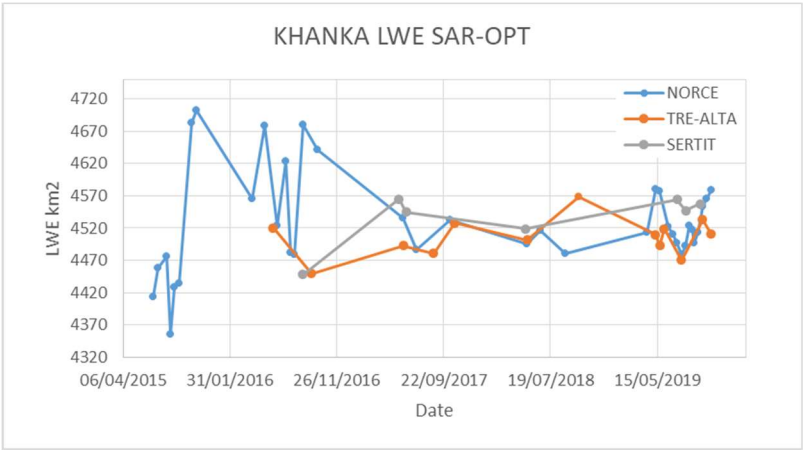


Figure 84. NORCE SAR, TRE-ALTA SAR and SERTIT LWE time series for lake Khanka.

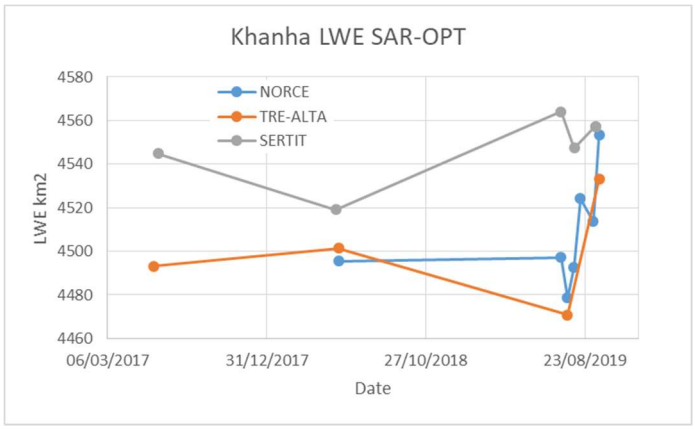


Figure 85. NORCE SAR, TRE-ALTA SAR and SERTIT Optical LWE time series for lake Khanka.

3.3.3.11. Sarykamish Lake

As already indicated, SaryKamish is located in a desert zone with very limited vegetation. The landscape consists mostly in sandy and muddy bare soil.

In the SAR LWE there is an important commission error between land and water of 5.3 %. This is related to the nature/texture/structure of the soil surrounding the lake. The area viewed as water in the SAR image, corresponds to a bare crust having similar backscatter as nearby water body.

The omissions are relatively scarce and corresponds to rough water surface and inundated vegetation within a delta-zone.

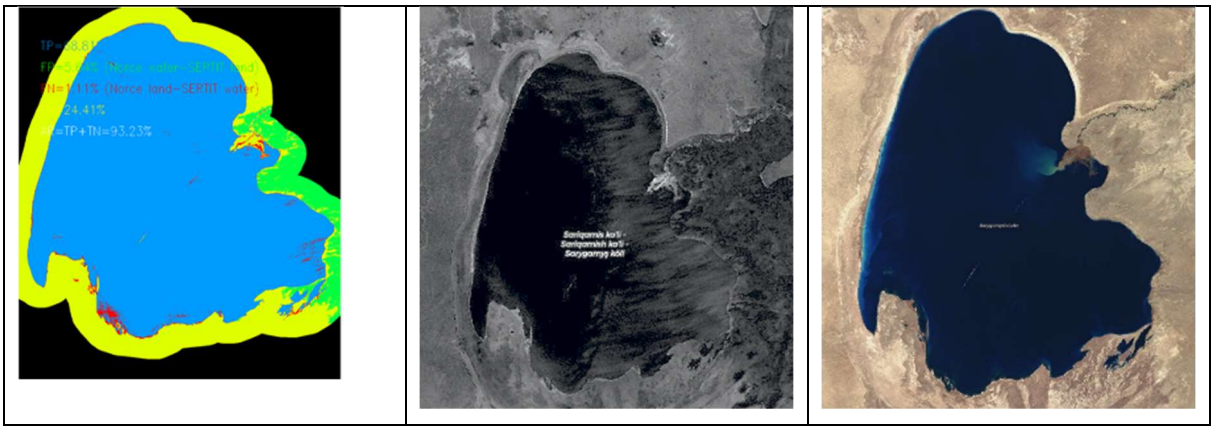


Figure 86: Comparison of omission and commission between pairs of Sentinel1, and Sentinel 2 acquired respectively the 2019-03-07, and 2019-03-11 over the Sary Kamish lake.

3.3.3.12. Summary of all inter-comparisons between optical and SAR classification

In this section we provide tabular contingency matrices for all the lakes that are inter-comparable between the two SAR and the Optical algorithms. Due to different selections of dates for the three approaches different dates will be inter-compared. The main criterion for inter-comparison is that the pair of images to compare

(either coming from any of NORCE's,TRE-ALTA's and SERTIT's) are close in time. For that, a maximum time difference to 12 days is set.

The following tables have been generated:

- Lake Bosten: NORCE-SERTIT Table 23 and TRE-ALTA –SERTIT Table 28.
- Lake Argentino: NORCE-SERTIT Table 24 and TRE-ALTA –SERTIT Table 29.
- Lake Khanka: NORCE-SERTIT Table 25 and TRE-ALTA –SERTIT Table 30.
- Mean Values of all previous: NORCE-SERTIT Table 27 and TRE-ALTA Table 31.

Tables' acronyms stand for: TP (True Positive in %), FP (False Positive in %), FN (False Negative in %), TN (True Negative in %), AR (Accuracy Rate in %), Td (time difference in days), Ad (Area LWE difference in %) and Ard (Area LWE difference in square kilometres).

Table 23. NORCE and SERTIT contingency matrix values summary for Lake Bosten

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Bosten	20170729	20170728	95.35	0.02	1.94	2.67	98.02	1	2.01	122.45

Table 24. NORCE and SERTIT contingency matrix values summary for Lake Argentino

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Argentino	20170531	20170602	78.35	2.21	1.20	18.22	96.58	3	1.26	18.39
Argentino	20171022	20171020	76.48	4.10	2.22	17.17	93.66	2	2.34	34.02
Argentino	20191024	20191015	77.12	5.51	0.85	16.50	93.63	9	5.64	85.87

Table 25. NORCE and SERTIT contingency matrix values summary for Lake Khanka

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Khanka	20180515	20180510	80.92	0.76	0.88	17.42	98.34	5	0.14	23.68

Table 26. NORCE and SERTIT contingency matrix values summary for Lake Sarykamish

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Sarykamish	20190307	20190311	68.81	5.64	1.11	24.41	93.23	4	6.07	217,78

Table 27. Mean Value for all previous NORCE and SERTIT contingency matrix AR, Td and Ad values

	AR	Td	Ad	Ard
Average	94.69	4,1	4.01	83,69

Table 28. TRE-ALTA and SERTIT contingency matrix values summary for Lake Bosten

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Bosten	20170729	20170728	94.97	0.02	2.32	2.67	97.65	1	2.42	65.74
Bosten	20180910	20180901	95.64	0.03	1.89	2.42	98.07	9	1.94	87.72

Table 29. TRE-ALTA and SERTIT contingency matrix values summary for Lake Argentino

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Argentino	20171015	20171020	76.93	2.38	1.78	18.89	95.82	5	0.76	41.96
Argentino	20171120	20171129	77.41	0.71	1.99	19.87	97.29	9	1.63	2.24
Argentino	20191017	20191015	75.86	1.22	2.10	20.79	96.66	8	1.14	6.45

Table 30. TRE-ALTA and SERTIT contingency matrix values summary for Lake Khanka

	Date 1	Date 2	TP	FP	FN	TN	AR	Td	Ad	Ard
Khanka	20170601	20170609	80.40	0.21	1.91	17.47	97.87	8.00	2.10	9,10
Khanka	20180515	20180510	80.36	0.37	1.43	17.81	98.18	5.00	1.31	23.68
Khanka	20190919	20190912	80.91	0.38	1.50	17.19	98.11	7.00	1.38	4.01

Table 31. Mean Value for all previous TRE-ALTA and SERTIT contingency matrix AR, Td and Ad values

	AR	Td	Ad	Ard
Average	94.82	5.5	5.41	30,11

3.3.4. Results analysis for hypsometric approach

The objectives are to compare the results of lake area from different methods and type of images in term of building resulting hypsometry curve. This approach was tested over the second set of lakes, Khanka (China_Russia), Illmen (Russia), Sary Kamysh lake (Turkménistan – Ouzbékistan), Bosten lake (China), and Argentino, (Argentina).

3.3.4.1. Hypsometric approach for the Argentino Lake

Over Argentino Lake, three series were compared:one from optical imagery and two from SAR imagery. Following remarks can be dressed

we see that SERTIT and NORCE solutions have the same mean value of lake water extent, since it is much lower for TRE-Altamira. The optical solution (SERTIT) presents non monotonic trend but this is likely due to two evident outliers as seen on the figure. The NORCE solution looks very disturbed with high dispersion of the hypsometry data. The TRE-Altamira looks quite coherent, with low RMS and monotonic trend but over a lower range of water level observations (no data for high water level above 180 m as for the 2 other solutions).

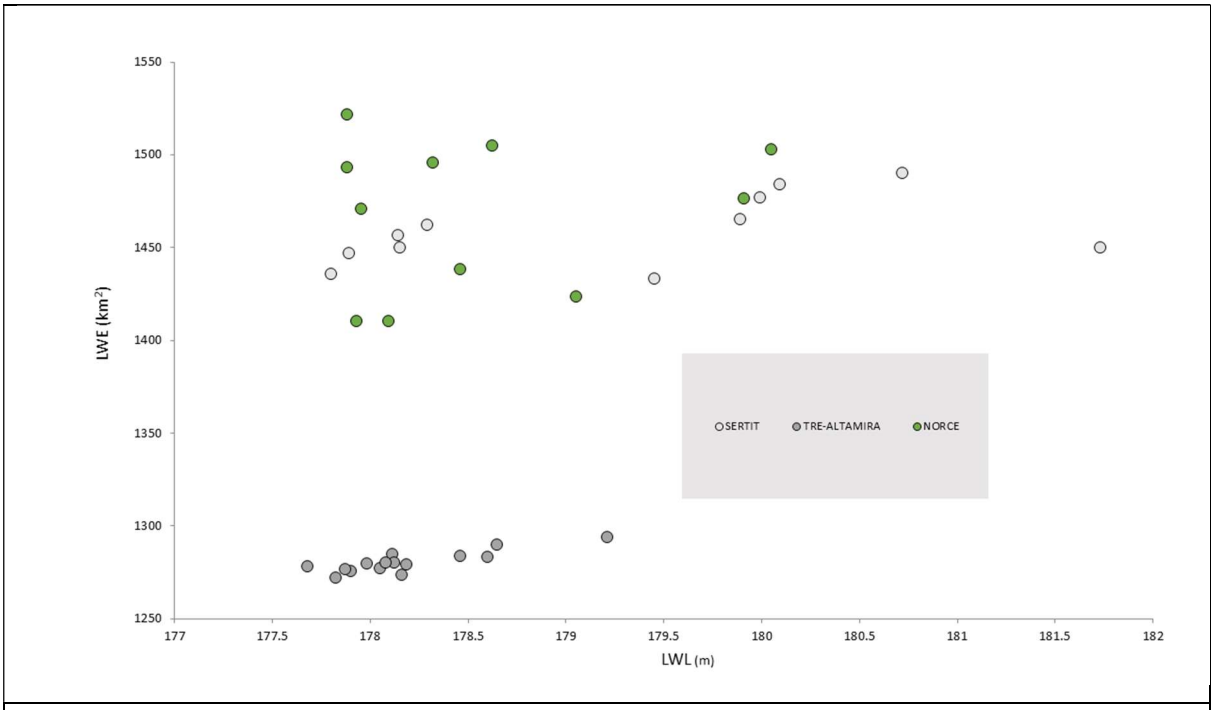


Figure 87 . Hypsometric approach over Argentino lake

The hypsometry coefficient can also be calculated using each of the solution with also the estimation of uncertainty, which here is characterized by the RMS of the difference between the theorical hypsometry and the data used to calculate it.

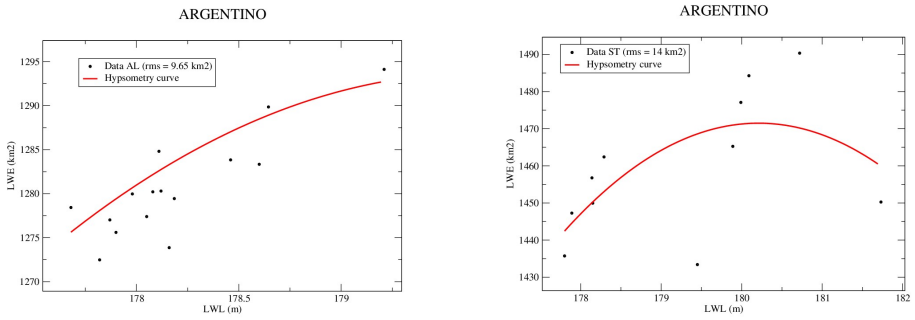


Figure 88. Hypsometry coefficient for Argentino lake

We see from this figure that the sensitivity of hypsometry to outliers (like for the optical solution: SERTIT) can be very high.

3.3.4.2. Hypsometric approach for the Bosten Lake

Over Bosten Lake, four series were compared, two from SAR imagery and the two from optical data. Following remarks can be dressed:

- The 4 solutions look quite coherent
- It remains some outliers

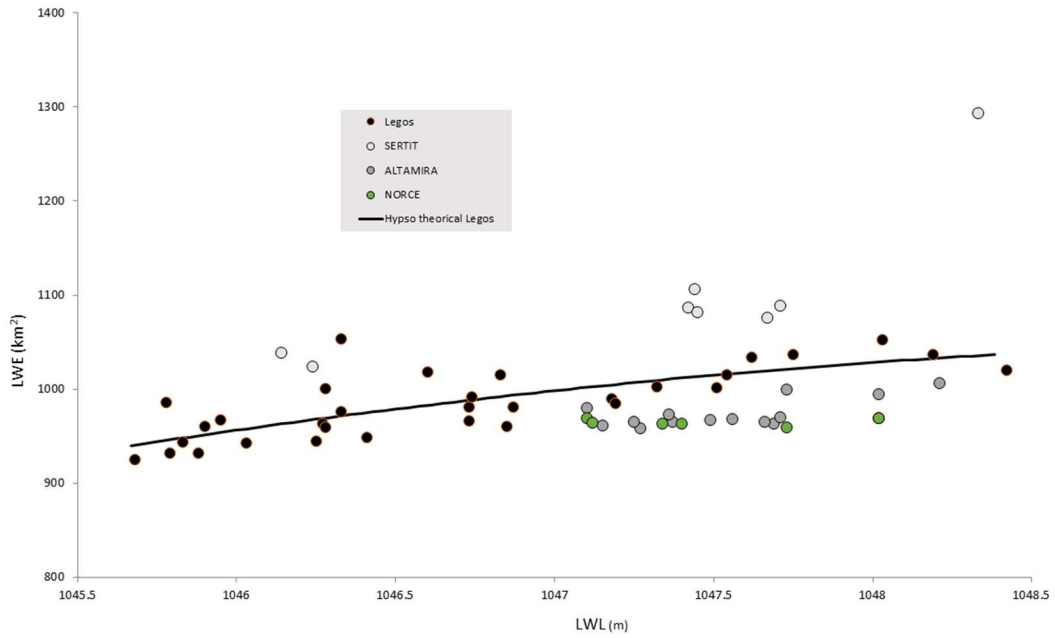


Figure 89. Hypsometric approach over Bosten Lake

It highlights the need to extract water extent of the lake over the largest range of water level variations. This minimizes the impact of dispersion of the solution on the calculation of the hypsometry coefficients, although it may increase the RMS.

The uncertainty using the theorical hypsometry curve has been calculated only with LEGOS solution. The RMS of LEGOS Solution is of 5%

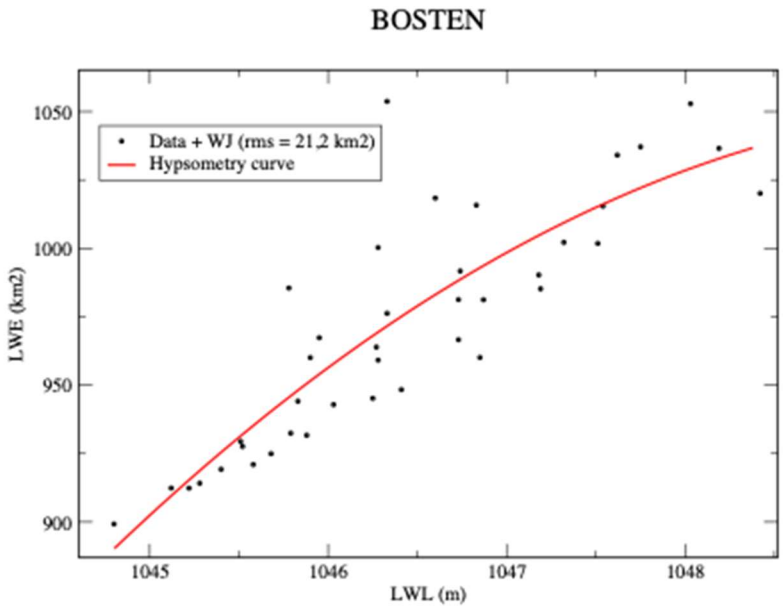


Figure 90. Uncertainty approach based on hypsometric over Bosten Lake

3.3.4.3. Hypsometric approach for the Illmen lake

Over Illmen Lake, three series were compared, two from SAR imagery and one from optical data. Following remarks can be dressed:

•Optical imageries look (Sertit) overestimates the LWE when compared to the 2 radar solutions, but not enough images have been processed

We observed from the two radar imagery datasets a small bias between both solutions: NORCE slightly overestimates the extent compared to TRE-Altamira. Both series looks however consistent, with monotonic trend and a good coverage of the water level changes over time

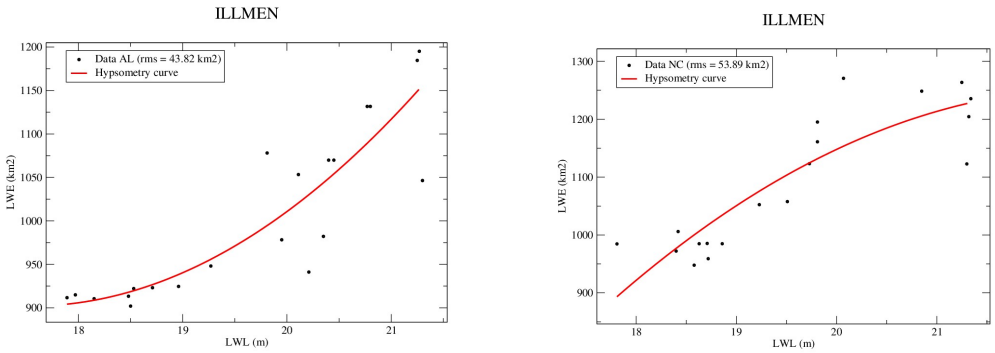


Figure 91. Lake Illmen. Theorical hypsometry for the 2 radar solutions

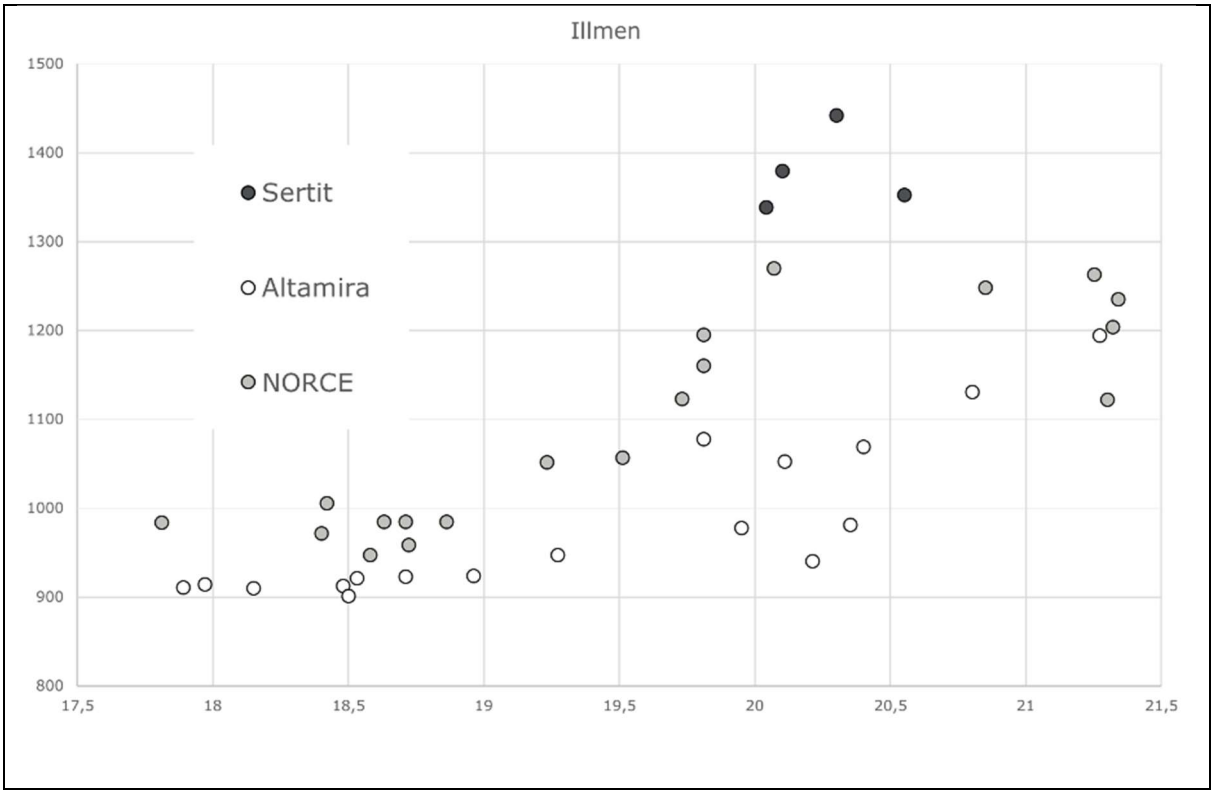


Figure 92. Hypsometric approach over Illmen Lake

3.3.4.4. Hypsometric approach for the Khanka Lake

Over Khanka Lake, four series were compared, two from SAR imagery and the two from optical data. Following remarks can be dressed:

•Very complicate case: 2 groups of solutions (optical Sertit/SAR NORCE & optical LEGOS/SAR Altamira) with strong bias

The LEGOS solution has a strong RMS, due to high dispersion, but has a monotonic trend which is more realistic. The NORCE solution cumulate monotonic trend and relatively low RMS, but over a much shorter range of observations.

•The 2 radar solutions don't have variation of LWE w.r.t variations of LWL (no trends).

They both doesn't have enough images processed. This is a typical case where hypsometry analysis does not really help building realistic relationship between LWE and LWL.

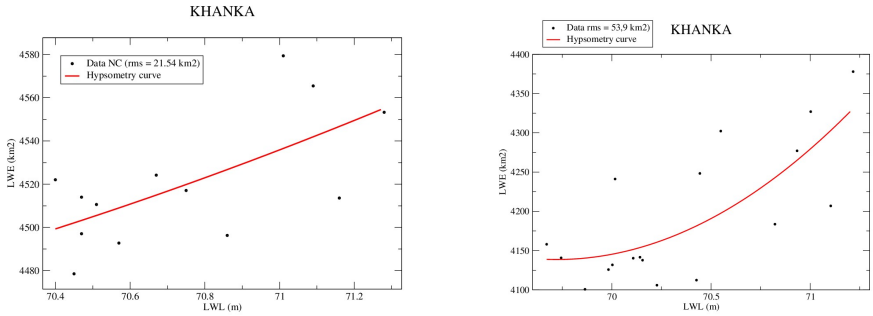


Figure 93. Theorical hypsometry from NORCE and LEGOS solutions

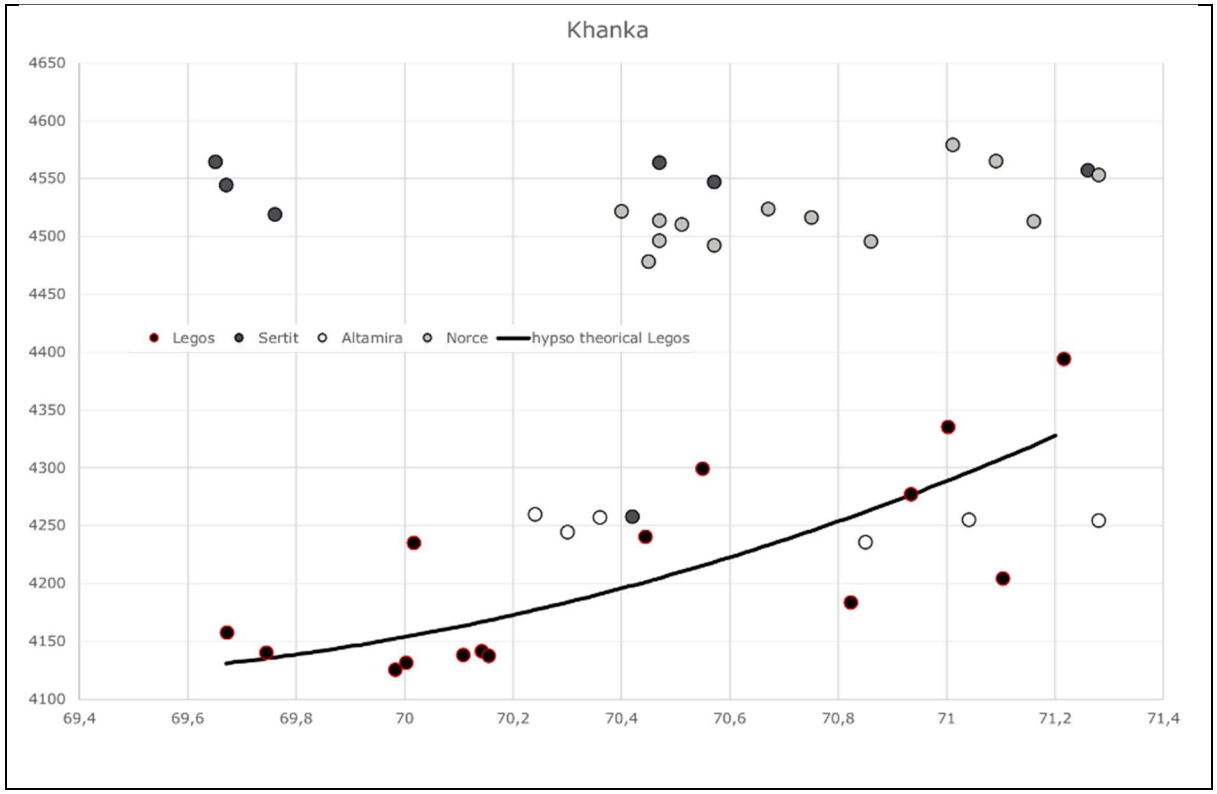


Figure 94. hypsometric approach over Khanka Lake

The uncertainty using the theorical hypsometry curve has been calculated only with LEGOS solution. The RMS of LEGOS Solution is of 1,3%

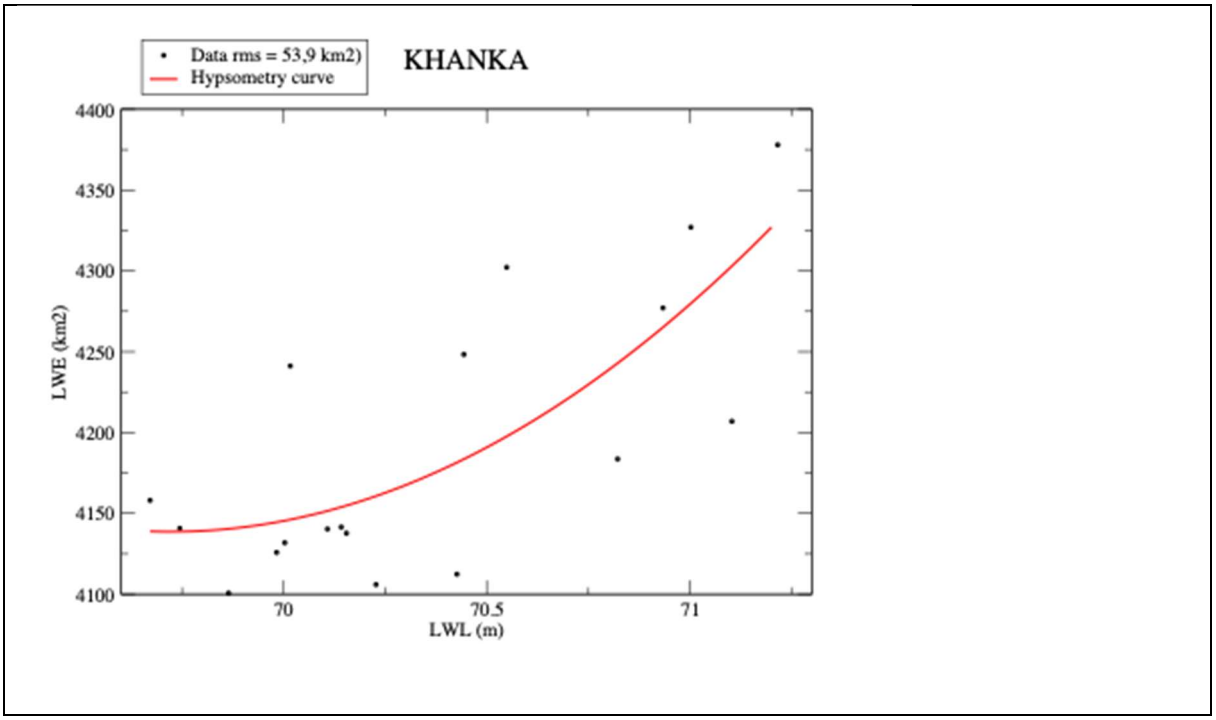
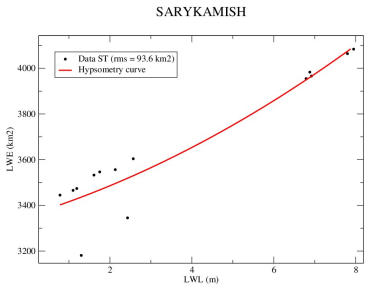


Figure 95. Uncertainty approach based on hypsometric over Khanka Lake

3.3.4.5. Hypsometric approach for the Sary -Kamisk Lake

Over Sary-Kamish Lake, four series were compared, two rom SAR imagery and wo from optical data. Following remarks can be dressed:

- The tow optical solutions are very coherent together (remains 2 outliers in the Sertit solution)
- Due to high LWL variation in time, LEGOS solution is uncompleted for small LWL
- It remains some outliers in the SERTIT solution
- the two solutions based on radar imagery are limited to a too low range of variations, since the water level change of this lake has strongly increased over the last 20 years. Radar imageries has been obtained only over the last 3-4 years when the lake was already at high level, as we may see from the X axis for these two solutions (level around 8 meters). This explains that the shape of the hypsometry depends too much on the dispersion of the LWL/LWE data and in such case hypsometry cannot be calculated using only these data.
- The uncertainty of the hypsometry using the two optical imagery's solution is lower than 1% of the total extent of the lake



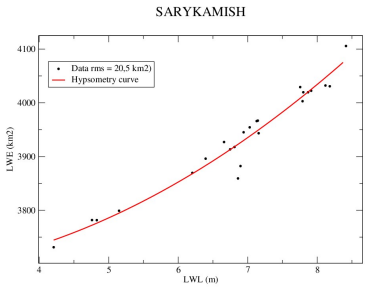


Figure 96 Theorical hypsometry from SERTIT and LEGOS solutions

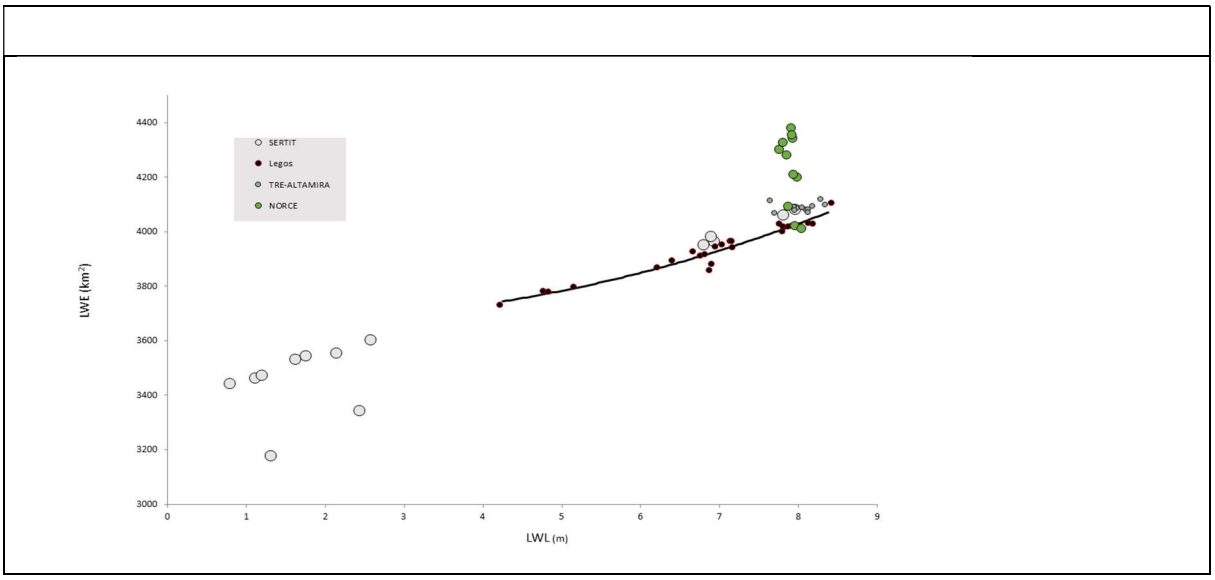


Figure 97. Hypsometric approach over SaryKamish Lake

The conclusion here is that for the Sarykamish lake, we may simply remove the two outliers (seen on the Sertit solution) and recalculate the hypsometry coefficient using the Legos, the TRE-Altamira and the Sertit solutions. It will allow to extend the domain of validity of the hypsometry from very low to very high lake water level.

The uncertainty using the theorical hypsometry curve has been calculated only with LEGOS solution. The RMS of LEGOS Solution is of 1,3%

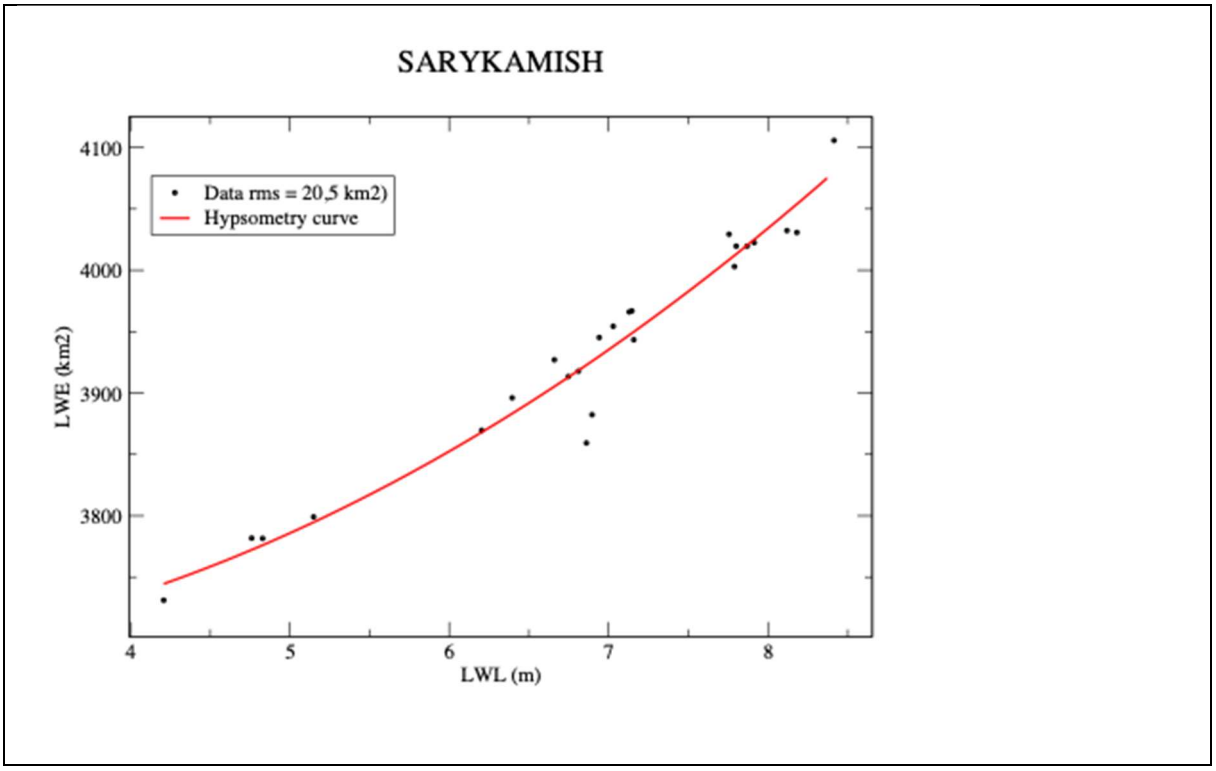


Figure 98. Uncertainty approach based on hypsometric over SaryKamish Lake

3.4. Conclusions and recommendations

We analysed the behaviour of two supervised and one unsupervised approaches on five lakes with various characteristics that provided a relatively good overview of what can be observed when estimating LWE from high-resolution optical imagery.

Each method has its weaknesses and strengths. Of course, we had to deal with the expected classical traps which are related to external condition such as clouds cover and associated shadows, local environment, the presence of vegetation within the water bodies, floating or submerged ones. Many of the characteristics of water conditions are related to the colour of the water in shallow water.

One main problem is also the definition of LWE. Do we have to consider open free water or open free water plus ice cover? In addition, are we speaking about open water, in case of water bodies surrounding or imbricated with wetlands? The delimitation between these entities is not so obvious. Do we have to consider a floating island of vegetation as water or not? Depending of the followed approach and or the exploited indices nor training samples selection process, omission or commission would /could happen.

The conclusion and recommendations at this stage, after two sets of benchmarking of methods, are the following:

- For optical data exploitation, the best method is a random forest one with a testing procedure on the training set based on Pekel GWS database.
- For SAR processing, KMeans approaches are satisfactory, but in some challenging environments, such as deserts or lakes with very fragmented or vegetated shoreline, it can result in large errors. Estimating the LWE during ice cover is also a challenging task, although it can be possible during the cold winter period (Vickers et al., 2019).

For an operational application, i.e. monitor LWE based on LWL, this require to generate de precise hypsometric curves, a work to be done lake after lake. More works are needed in order to select the

appropriate images, to access these images as there are some restriction in term of data access exploiting either GEE , nor ESA SciHub.

4. Lake Surface Water Temperature (LSWT) parameter

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

4.1. Data description

4.1.1. Satellite data

LSWT time series have been computed from sensors on multiple satellites and lake-specific consistency adjustments between sensors have been applied using the MetOpA AVHRR instrument as a reference (see RD- 1). In the current version, LSWT v4.1, MetOpA AVHRR has the best combination of length of record and data density for this purpose. The same form of algorithms has been used to retrieve the LSWT from all sensors in order to obtain consistent time series for each of the CCI target lakes. The target list was defined within the first phase of the Lakes_cci project can be found at RD- 3.

The time periods used for each satellite/instrument are provided in Table 32. Not all lakes include LSWT from all sensors in the series because of differing density and geometry of observation. The temperatures in the dataset are only available for cloud and ice-free observations during (in this version) day time, so gaps in time and space are common for all the lakes due to cloud cover and swath geometry of the instruments.

Table 32 - Time periods for the satellite/instrument used to generate the LSWTs

Satellite	Instrument	Time Period
ERS-2	ATSR2	Jun1995 – Jun2003
Envisat	AASTR	May2002 – Apr2012
MetOpA	AVHRR	Mar2007 – Aug2019
MetOpB	AVHRR	Dec2012 – Aug2019

A detailed description of the product generation is provided in the Algorithm Theoretical Basis Document (ATBD RD- 1) with further information on the product given in the Product User Guide (PUG RD- 3).

4.1.2. In situ data

The in situ dataset currently used for validation has been constructed from the in situ temperature data collected through the ARCLake project, the GloboLakes project, the EU Surface Temperature for

All Corners of Earth (EUSTACE) project and the Copernicus Climate Change Service (C3S) product. At present, this dataset consists of 272 observation locations covering 132 lakes. However, the number of lakes which are also present in the CCI lake list of the first phase are 44 with a total of 159 sites. Details of the in situ observation locations with their sources are given in Table 33 which reports all locations for the target lakes where there are matches. The geographical distribution of the sites is reported in Figure 99 which shows that most of the globe is covered with a big proportion of the sites located in North America and Europe.

Table 33- List of the sources of the in situ data

Satellite	Instrument
NDBC – National Data Buoy Centre (USA)	Superior (3) Huron (2) Michigan (2) Erie (1) Ontario (1)
FOC – Fisheries and Oceans Canada (Canada)	Superior (1) Huron (4) Great Slave (2) Erie (2) Winnipeg (3) Ontario (4) Woods (1) Saint Claire (1) Nipissing (1) Simcoe (1)
Michigan Technological University (USA)	Superior (2) Michigan (1)
University of Minnesota (USA)	Superior (2)
Northern University of Michigan (USA)	Superior (3)
Superior Watershed Partnership (USA)	Superior (1)
U.S. Army Corps of Engineers (USA)	Superior (1)
Technical University of Kenya (Kenya)	Victoria (1)
GLERL – Great Lakes Environmental Research Lab (USA)	Huron (3) Michigan (2)
University of Wisconsin-Milwaukee (USA)	Michigan (2)
Northwestern Michigan College (USA)	Michigan (1)
University of Michigan CIGLR (USA)	Michigan (2)
Limno Tech (USA)	Michigan (3) Erie (4)
Illinois-Indiana Sea Grant and Purdue Civil Engineering (USA)	Michigan (2)
Leibniz Institute for Freshwater Ecology and Inland Fisheries (Germany)	Tanganyika (1)
Pierre Denis Plisnier	Tanganyika (4)
Irkutsk State University (Russia)	Baikal (1)
Regional Science Consortium (USA)	Erie (1)
UGLOS – Upper Great Lakes Observing System (USA)	Erie (2) Douglas (1)
LEGOS – Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (France)	Issykkul (1)
SLU – Swedish University of Agricultural Science (Sweden)	Vanern (6) Vattern (2) Malaren (9) Bolmen (2)
Uppsala University (Sweden)	Vanern (1) Erken(1)
Sao Paulo State University (Brazil)	Tucurui (1) Itaipu (1) Tres Marias (1)
Junsheng Li (China)	Taihu (1)
KU Leven (Belgium)	Kivu (1)
SYKE – Finnish Environment Institute (Finland)	Paijanne (3) Oulujarvi (1) Pyhajarvi(1)
Vermont EPSCOR – Established Program to Stimulate	Champlain (1)

Satellite	Instrument
Competitive Research (USA)	
SUNY Plattsburgh Center for Earth and Environmental Science (USA)	Champlain (1)
Nipissing University (Canada)	Nipissing (2)
National Park Service (USA)	Mead (3) Mohave (2)
GLEON – *Global Lake Ecological Observatory Network*	Tanganyika (3) Balaton (1)
BLI – Balaton Limnological Institute (Hungary)	Balaton (6)
KDKVI – Central Transdanubian (Regional) Inspectorate for Environmental Protection, Nature Conservation and Water Management (Hungary)	Balaton (3)
UMR CARRTEL – Centre Alpin de Recerche sur le Réseaux Trophique des Ecosystèmes Limniques (France)	Geneva (1)
UC-Davis Tahoe Environmental Research Center (USA)	Tahoe (1)
Utrecht University (Nederlands)	Garda (1)
Italian National Research Council (Italy)	Garda (8) Trasimeno (2) Maggiore (2) Iseo (2)
NOAA National Ocean Service Water Level Observation Network (USA)	St John River (3)
Estonian University of Life Sciences (Estonia)	Vorstjarv (4)
Israel Oceanographic and Limnological Research (Israel)	Sea of Galilee (2)
National Institute for Environmental Studies (Japan)	Kasumigaura (5)
Universitá degli Studi di Perugia (Italy)	Trasimeno (1)
Freie Universitat Berlin/Fondazione Edmund Mach (Germany/Italy)	Iseo (1)
University of Wisconsin-Madison (USA)	Mendota (1)
NTL LTER – North Temperate Lakes Long-Term Ecological Research (USA)	Mendota (1)
The Ohio State University (USA)	Douglas (1)

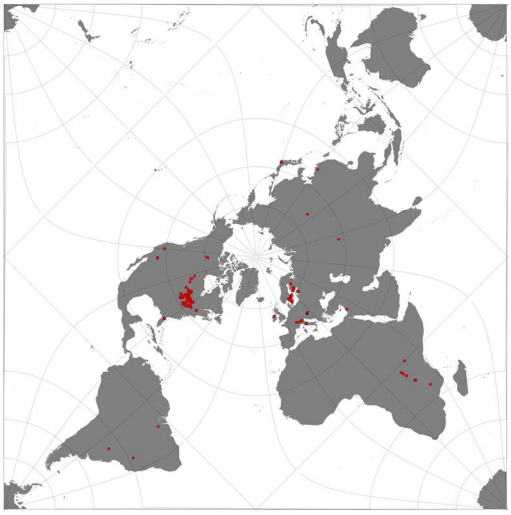


Figure 99 – Geographical distribution of sites used for LSWT validation.

Table 34. Lakes_cci lakes with in situ data. lists the 44 lakes together with their maximum distance from land [Carrea et al. 2015], which is an indication of each lakes' size that is meaningful for LSWT remote sensing. The distance to land for lake Iseo in Italy is shown in Figure 100. The best resolution of the instruments used for the retrieval of the LSWT is 1 km. If the lake has a maximum distance to land of 1.7 km such as lake Iseo, the LSWT retrieval is very likely to be available only for that part of the lake and only for a limited proportion of overpasses (clear sky and observations relatively central within the swath). In particular, a combination of factors has to occur: 1) the satellite image locations line up so that some pixels are nominally fully water pixels, which requires the satellite view zenith angle (which affects the on-the-ground resolution) to be such that the half-pixel size is smaller than the distance to coast; 2) these pixels are cloud free; 3) image geolocation errors (which can be of order 1 pixel uncertainty) are small enough that the nominally water-filled pixels are truly water-filled meaning that the water detection tests are passed.

Table 34. Lakes_cci lakes with in situ data.

Lake id	Name	Country	N sites	Max distance to land (km)
2	Superior	Canada/USA	13	73.5
3	Victoria	Tanzania	1	84.1
5	Huron	Canada/USA	9	73.3
6	Michigan	USA	15	63.8
7	Tanganyika	Tanzania	8	34.1
8	Baikal	Russia	1	33.7
11	Great Slave	Canada	2	44.6
12	Erie	Canada	10	45.6
13	Winnipeg	Canada	3	40.1
15	Ontario	Canada	5	36.1
25	Issykkul	Kyrgyzstan	1	26.9
29	Vanern	Sweden	7	20.3
44	Woods	Canada	1	11.8
52	Tucurui	Brazil	1	6.4
65	Itaipu	Paraguay	1	3.8
66	Taihu	China	1	16
67	Kivu	Zaire	1	13
95	Vattern	Sweden	2	9.9
146	Saint Claire	Canada	1	13
157	Paijanne	Finland	3	3.8
163	Malaren	Sweden	9	2.7
165	Champlain	USA	2	5.8
188	Tres Marias	Brazil	1	2.3
198	Nipissing	Canada	3	9
202	Oulujarvi	Finland	1	6
236	Simcoe	Canada	1	8.4
278	Mead	USA	3	3.8

Lake id	Name	Country	N sites	Max distance to land (km)
310	Balaton	Hungary	10	6
327	Geneva	Switzerland	1	6.2
380	Tahoe	USA	1	8.2
505	Garda	Italy	9	5.2
507	St John River	USA	1	2.4
679	Vorstjarv	Estonia	4	6.2
948	Maggiore	Italy	2	2.4
1028	Bolmen	Sweden	2	2.7
1196	Sea of Galilee	Israel	2	5.6
1204	Kasumigaura	Japan	5	3.7
1240	Pyhajarvi	Finland	1	3.9
1529	Trasimeno	Italy	3	4.3
1603	Mohave	USA	2	2.8
3153	Iseo	Italy	3	1.7
4503	Mendota	USA	2	2.5
6785	Erken	Sweden	1	1.5
13377	Douglas	USA	2	1.5

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

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

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

The in situ data have a range of characteristics:

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

While part of the data are available online, the majority has been collected through personal communications and in a proportion of cases we are not licensed to redistribute the data because of the provider's data policy.

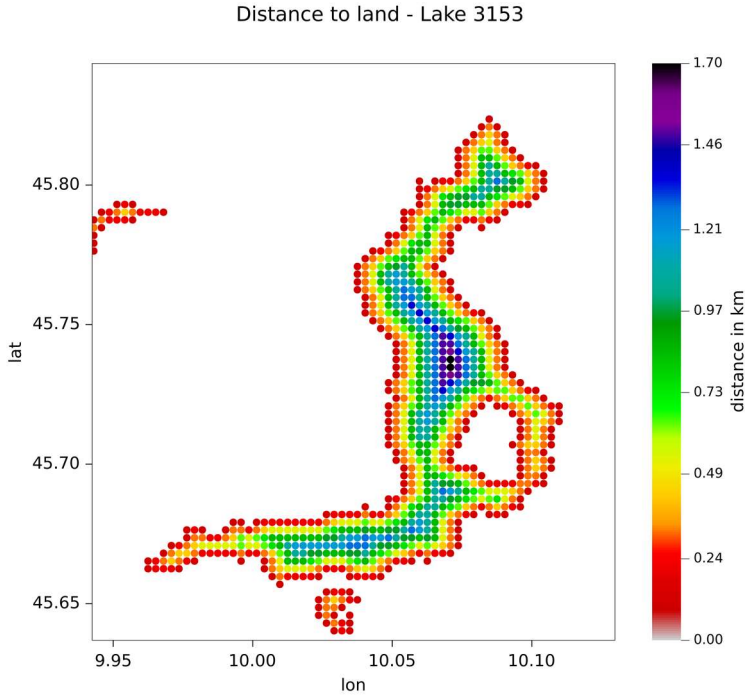


Figure 100. Distance to the nearest land for each pixel on water for lake Iseo in Italy at about 300m resolution.

4.2. Comparison methods

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

4.2.1. Generation of the L2 matchups

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

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

4.2.2. Validation of the L3S CCI LSWT v4.1

The differences between the L3S LSWT and in situ data are analysed using both standard and robust statistics. Robust statistics is less influenced by outliers in the distribution of differences, which can be also caused by poor in situ measurements. Quality control measures have been applied to in situ measurements but many different instruments have been used to take measurements and the operating methods of the instruments and the reporting vary strongly between sites. Time series of the absolute temperatures together with their difference are generated differentiating the quality levels. “Violin” plots where the distribution of the difference is shown are produced for each quality level. The robust statistics is also investigated per quality level for each year and for each lake.

4.2.3. Validation of the LSWT uncertainty

The validation of the L3S Lakes_CCI LSWT v4.1 is carried out comparing the satellite minus in situ temperature difference with the combination of the satellite uncertainty (present in the products) and an estimate of the in situ uncertainty (which is relatively poorly known). In an ideal case, the standard deviation of the differences between the satellite LSWT and a reference LSWT would equal the combined measurement uncertainty plus the uncertainty attributable to representativity effects.

4.2.4. Number of CCI lakes with LSWT

An assessment of the lakes with no retrieved LSWT is reported. Most of the lakes are too small in comparison with the satellite resolution and other can be included in the next version.

4.3. Description of work

The matchup is carried out per sensor over the 159 locations on 44 lakes. The total number of matches is 114,487 for any quality level and 98,119 excluding satellite LSWT of quality level equal to 1. The number of matches varies per year and since the AVHRR sensors have a larger swath than the ATSR sensors (ATSRs swath is 500 km and AVHRRs swath is ~2900km), after 2007 the number of matches clearly increases as it is shown in Figure 101. Number of matches for the CCI lakes at L3 per year.. We can notice another clear increase in 2013 when the AVHRR on MetOpA is used together with the AVHRR on MetOpB. In 2019 the number of matches is lower than the previous year because the LSWT time series end at the end of August 2019. The number of matches depends also on the availability of the in situ since a different number of locations is available every year as shown in Figure 101Figure 102. Number of sites with matches for the CCI lakes at L3 per year. The number of locations where in situ measurements have been taken has almost tripled since 1995; however, a portion of the measurement temporal frequency is daily.

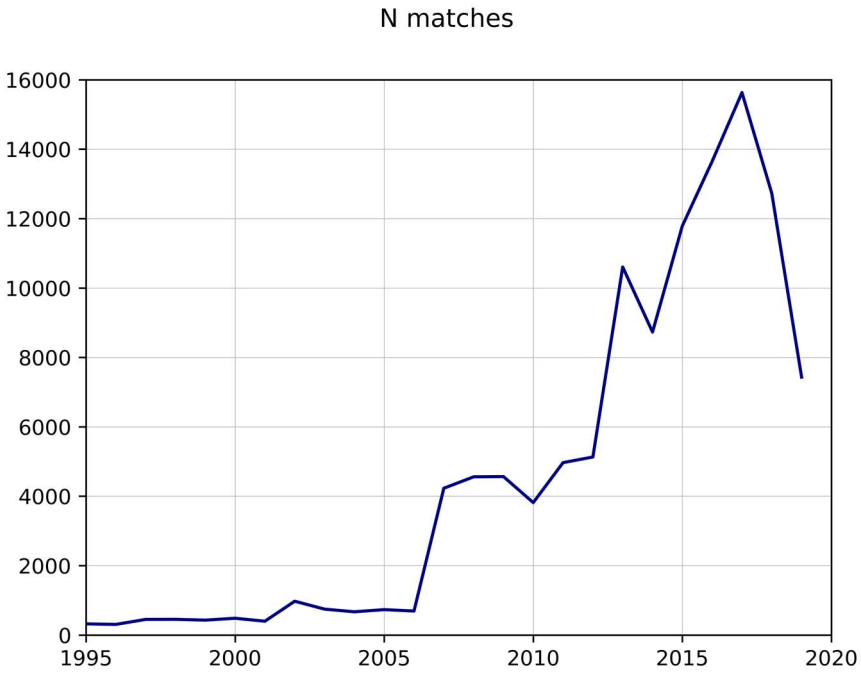


Figure 101. Number of matches for the CCI lakes at L3 per year.

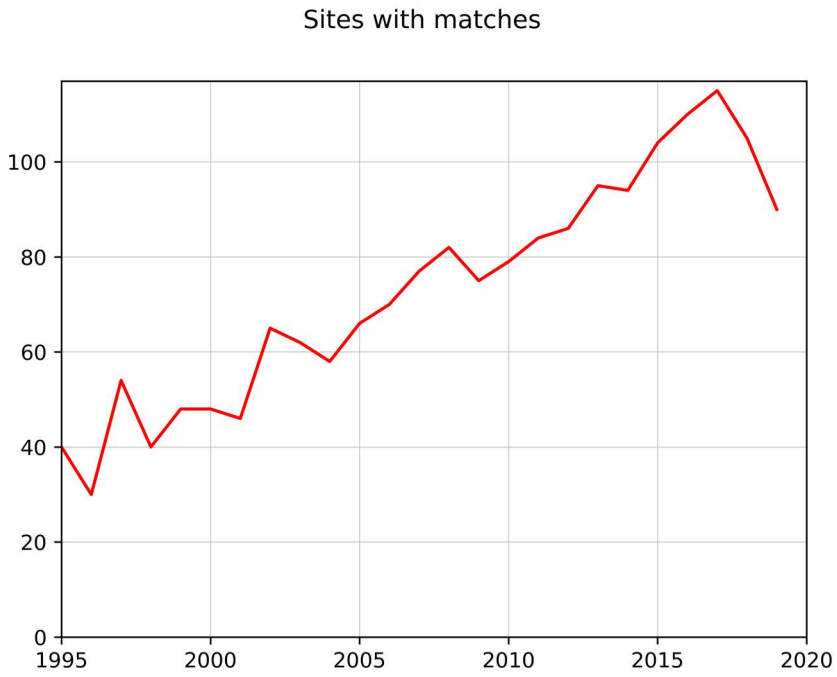


Figure 102. Number of sites with matches for the CCI lakes at L3 per year.

4.4. Result analysis

4.4.1. Validation of LSWTs

Robust statistics and the traditional statistics per quality level are reported in Table 35. List of the CCI lakes with in situ data. for the matches across all the locations where in situ measurements were available as reported in Table 34.

Table 35. List of the CCI lakes with in situ data.

QL	N	Median	RSD	Mean	SD
5	**34681**	**-0.060**	**0.504**	-0.056	0.932
4	**26757**	**-0.160**	**0.682**	-0.183	1.135
3	**23559**	**-0.240**	**0.934**	-0.336	1.356
2	**13122**	**-0.570**	**1.394**	-0.740	1.718
1	**16368**	**-2.730**	**4.240**	-3.879	4.913

In Table 35 the number of matches per quality levels are listed together with the median and the robust standard deviation of the satellite minus in situ temperature difference and the traditional metrics, the mean and the standard deviation. The difference between the median and the mean is almost negligible for quality level 5 and it increases as the quality levels get lower (suggesting, as expected, a higher incidence of cold-biased observations for low quality levels). The agreement between satellite and in situ measurements varies according to the quality levels in a way that is expected.

The best agreement is for quality levels 4 and 5, which are the levels that reflect a higher degree of confidence in the validity of the satellite estimate. Our recommendation to users is to use the quality level 4 and 5 for lake-climate applications in general, although lower quality level data may be relevant to users where they have specifically verified their fitness for a given lake for their application. Quality level 3 data comparison with the in situ data shows an agreement that may be acceptable to some users; however, they have to be used with care. Quality level 1 data should never be used and they are classified as “bad data”.

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

variability. Again, quantitative understanding of the scatter from these effects is not yet mature, and for this reason full uncertainty budget validation remains a research aspiration (see also E3UB).

The distributions of the satellite minus in situ temperature differences per quality level 2,3,4,5 are reported in Figure 103 as "violin" plots. The distributions become more stretched and less symmetric with longer tail towards negative differences as the quality levels decreases.

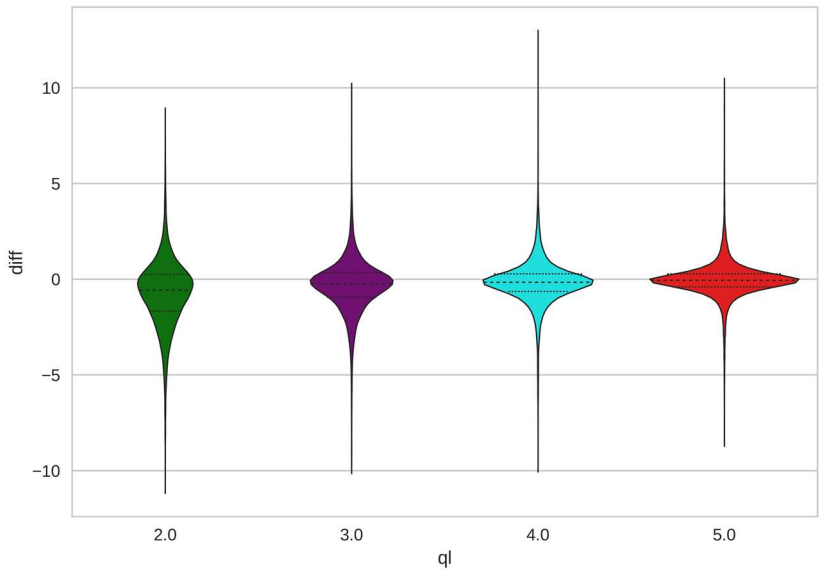


Figure 103. Distributions of the satellite minus in situ temperature difference per quality level as "violin" plots where the widths indicate the density of data for a given difference.

The median and the robust standard deviation per quality level per year for all the lakes is shown in Figure 104 and Figure 105 together with the number of matches. For high quality levels the median and the robust standard deviation are consistently low throughout the years when different instruments have been adopted and a different number of matches is available. They deteriorate as the quality goes lower especially for the ATSRs sensors which are relied on exclusively until 2007. The number of matches for quality level 5 is consistently the highest.

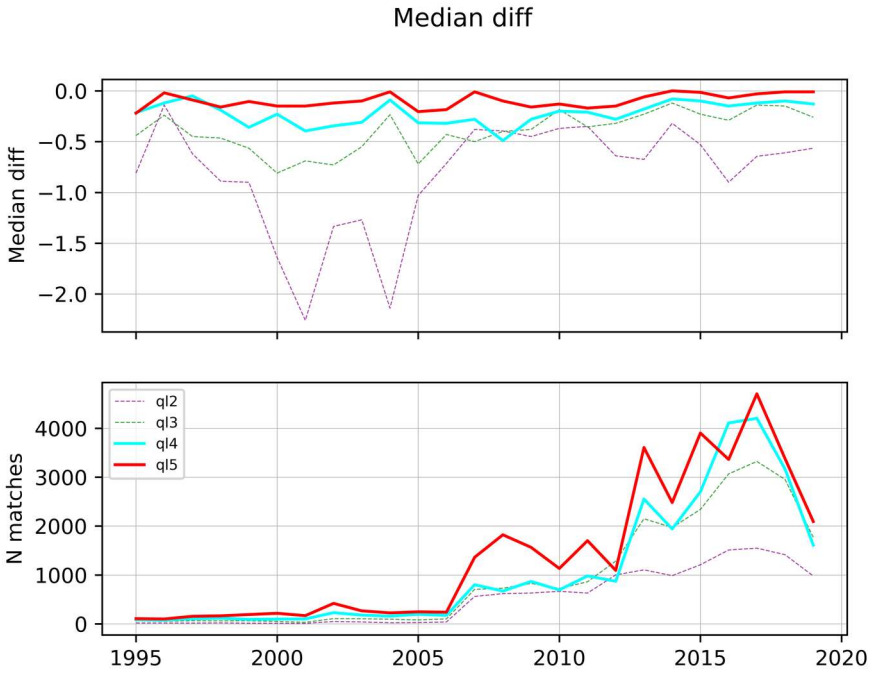


Figure 104. Satellite minus in situ temperature difference median per year (upper plot) and number of matches (lower plot) per quality level.

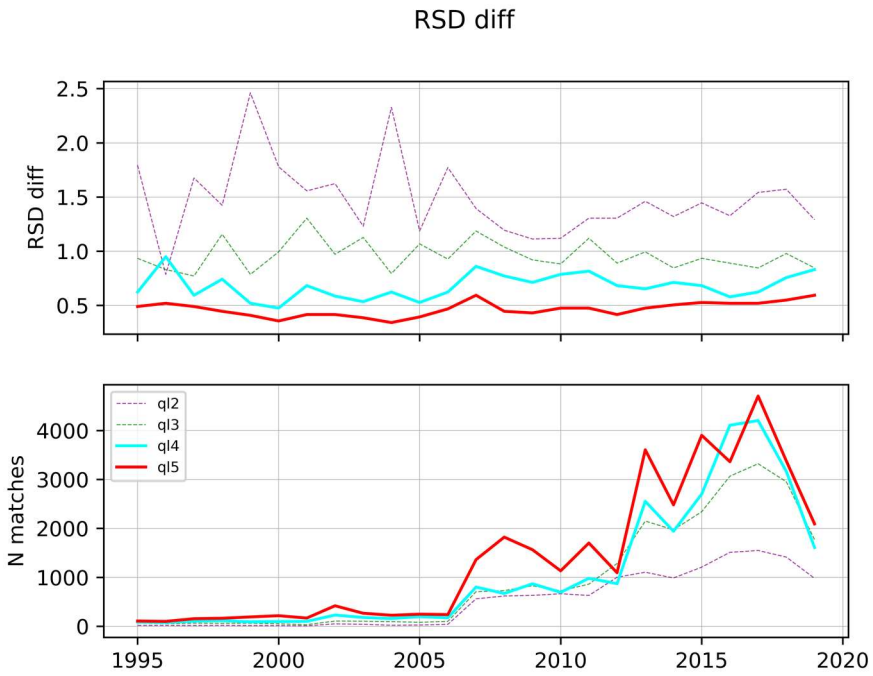


Figure 105. Satellite minus in situ temperature difference robust standard deviation per year (upper plot) and number of matches (lower plot) per quality level.

The median and robust standard deviation have been inspected also for each lake. Figure 106 and Figure 107 show the plots together with the correspondent number of matches. Higher numbers of matches are for lakes where data were available for longer periods but also where hourly/subhourly measurements were available and for sites far from the coast. The median and robust standard deviation are consistently better for quality level 5 throughout the lakes, while for quality level 4 the instances of greater variation are related to lower numbers of matches. However, for some lakes/quality-level combinations the in situ measurements are very few: for example, for lake Taihu (lake ID 66), only one match with LSWT of quality level 4 is available. For lake Paijanne (lake ID 157), the median difference and the robust standard deviation are unusually large. The in situ data for this lake come from two different originators for three sites as shown in Figure 108. Figure 109 shows the satellite LSWTs, the in situ values and the climatology (for reference) for the three sites on lake Paijanne in 2015. The majority of the measurements have been taken at the site 1 which is very close to the coast where the satellite minus in situ difference show cooler satellite LSWT in the first part of the year and warmer in the second part. This behaviour is consistently throughout the years and is consistent with an effect of shallow-water energy balance on the in situ measurement that differences from the satellite location of the matches that are at a spatial distance up to 3 km; for comparing to measurements close to the coast in shallow water, this is a significant distance. For the other two sites the in situ and satellite measurements have a good agreement but the in situ data have a lower temporal resolution.

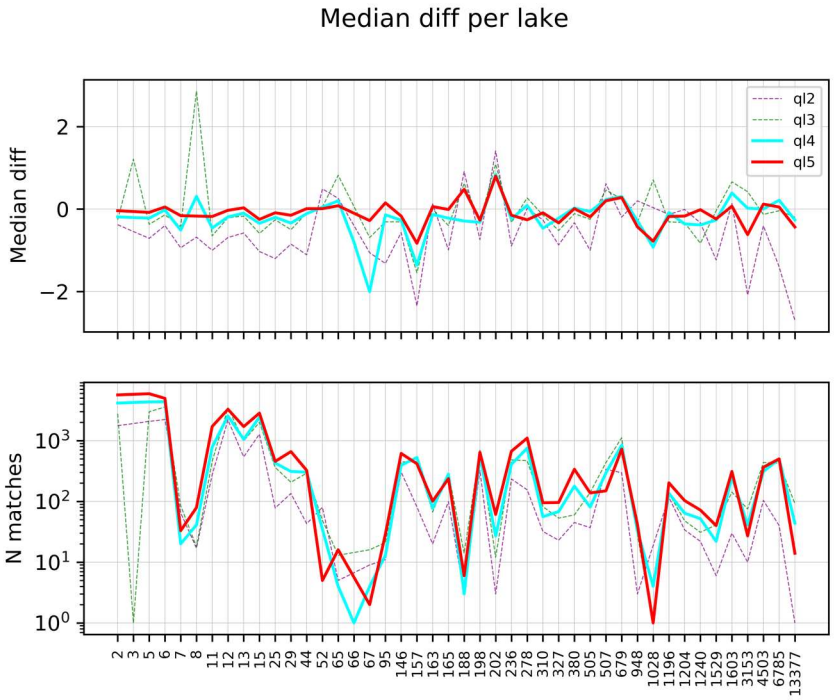


Figure 106. Satellite minus in situ temperature difference median per lake (upper plot) and number of matches (lower plot) per quality level.

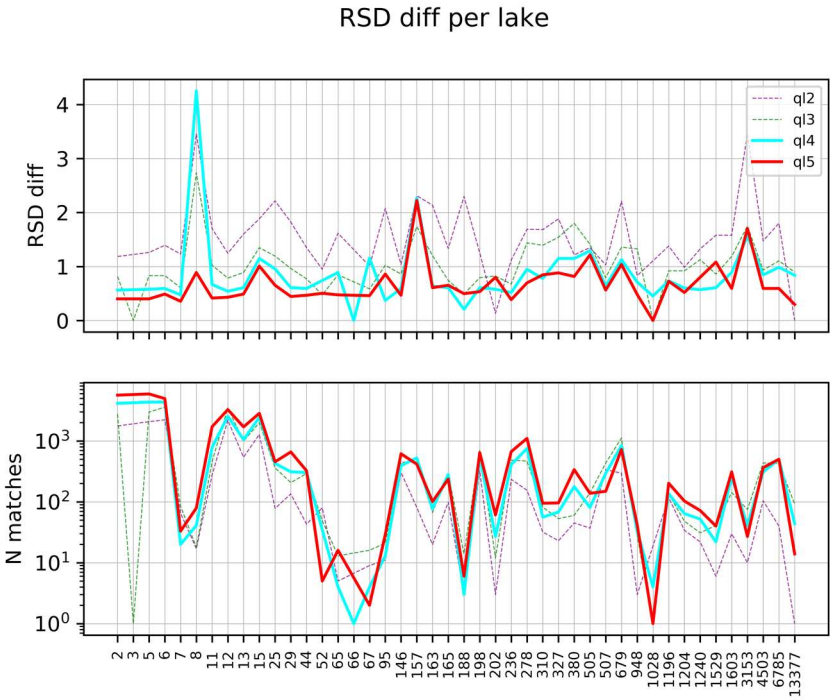


Figure 107. Satellite minus in situ temperature difference robust standard deviation per lake (upper plot) and number of matches (lower plot) per quality level.

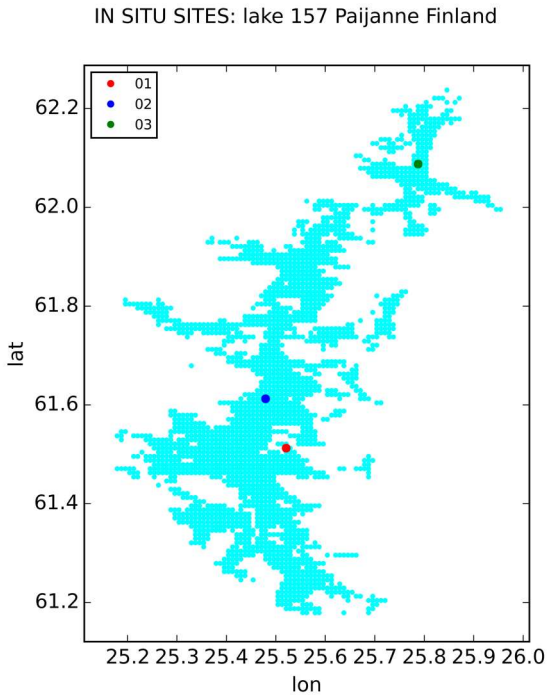


Figure 108. Locations of in situ measurements for lake Paijanne in Finland. Each dot represents a 1/120°x1/120° cell.

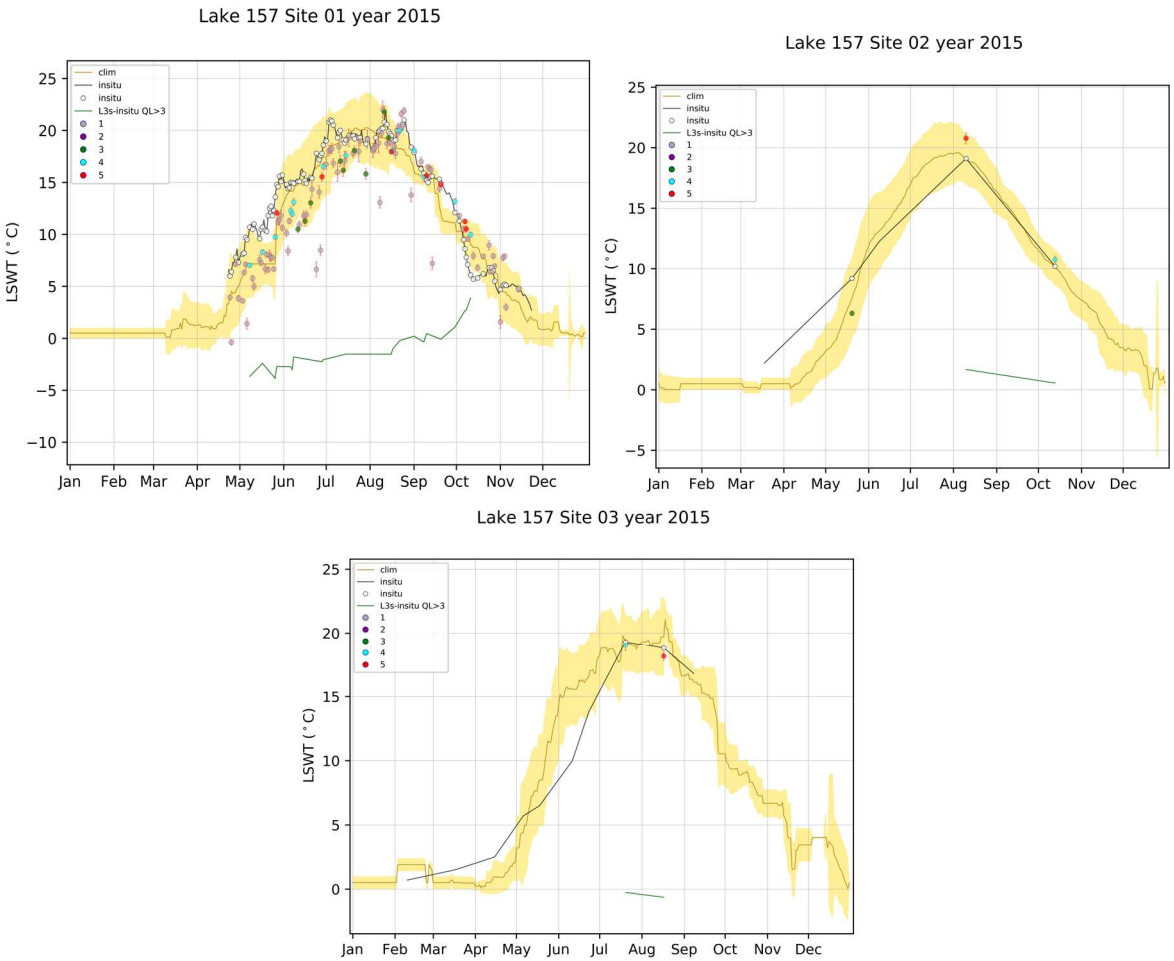


Figure 109. Validation plots for the three sites on lake Paijanne (Finland) for year 2015 where the yellow plot is the climatology, the black line and the white dots are the in situ measurements, the coloured dots are the satellite LSWTs per quality level and the green line is the satellite minus in situ difference.

As another example, consider lake Superior where many sites are available. The robust statistics of the difference for all the matches of quality level 3,4,5 have been plotted per sites in Figure 110 showing consistency for sites near each other, and a higher variability of the differences for sites close to the coast than those far into the lake, as expected. Figure 111 and Figure 112 show the LSWT and the in situ measurements in 2014 respectively for three sites on the lake Superior. In 2014 a sharp increase in temperature can be observed in the beginning of August consistently for the three sites and consistently for the satellite and in situ measurements. The timing of temperature increase is consistent at the three very different locations, one being closer to the coast and the other two more offshore. Figure 112 (right hand side) shows the position of the in situ measurement sites on lake Superior.

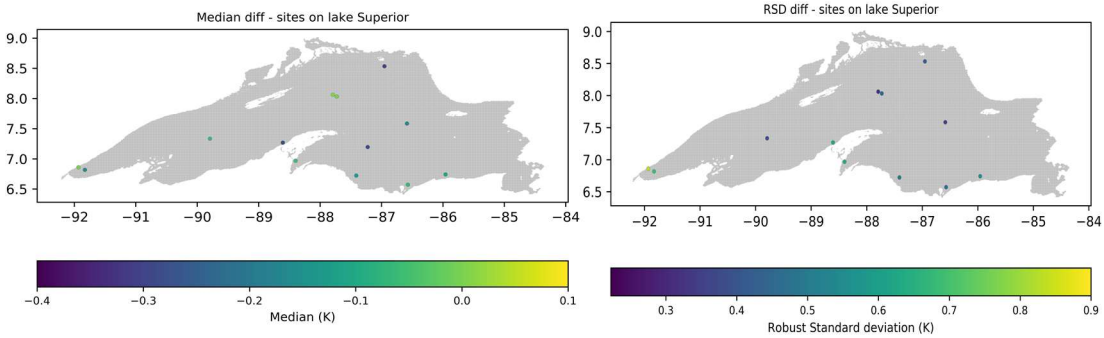


Figure 110. Satellite minus in situ temperature difference median and robust standard deviation for all the sites on lake Superior for quality level 3,4,5.

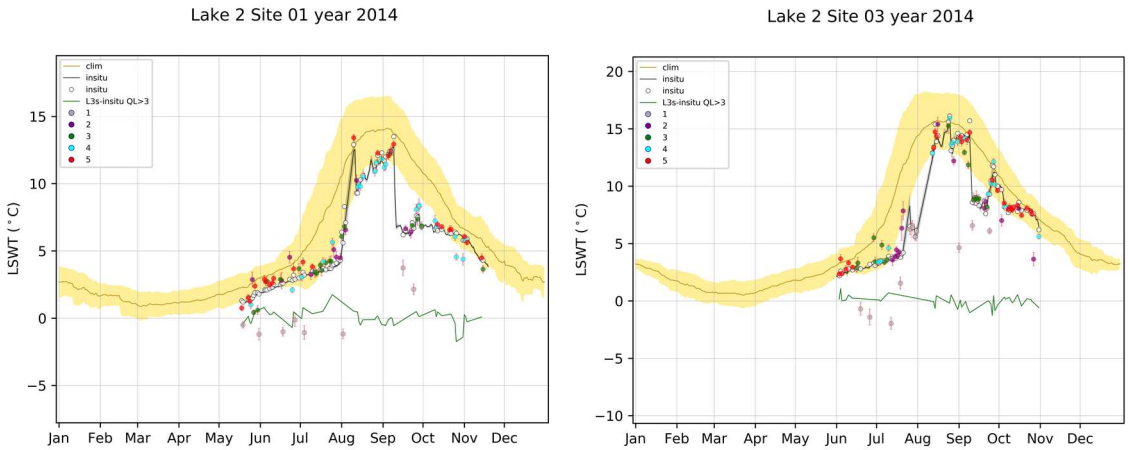


Figure 111. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for lake Superior in 2014 at site 01 (on the left) and site 03 (on the right).

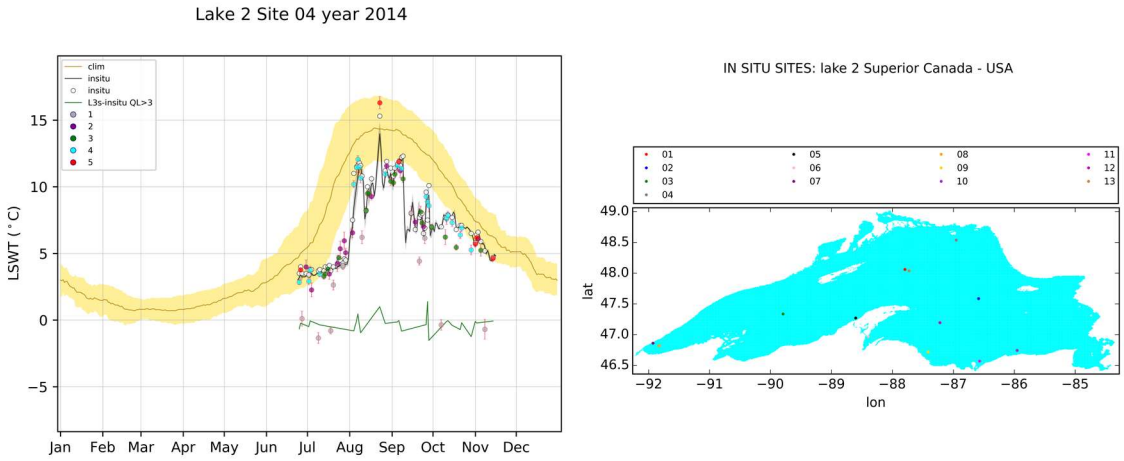


Figure 112. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for lake Superior in 2014 at site 04 (on the left) and location of the in situ measurement sites on lake Superior on a 1/120° grid (on the right).

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

The location where the in situ measurements have been collected on Lake Erken in Sweden is shown in Figure 113 (red dot). Figure F16 and Figure F17 show the satellite observations and the in situ measurements in 1997 when only ATSR2 was utilised and in 2014 when observations from AVHRR-A and AVHRR-B were used. For both the years the satellite observations follow remarkably well the in situ measurements, which were very high frequency measurements. The peak in the difference (green line) in 1995 is very likely due to a slight temporal mismatch. Despite the peculiar behaviour of the temperatures through the year in both cases the satellite and in situ are mimicking each other remarkably well. Note that in this case the measurements site is close to the shore but matching within 3 km does not have a strong influence in this case because the lake is small and more consistent in LSWT across its area.

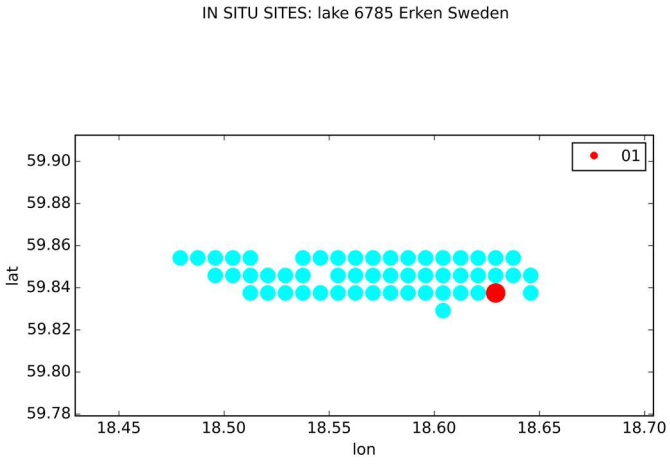


Figure 113. Location of in situ measurements for lake Erken in Sweden. Each dot represents a 1/120°x1/120° cell.

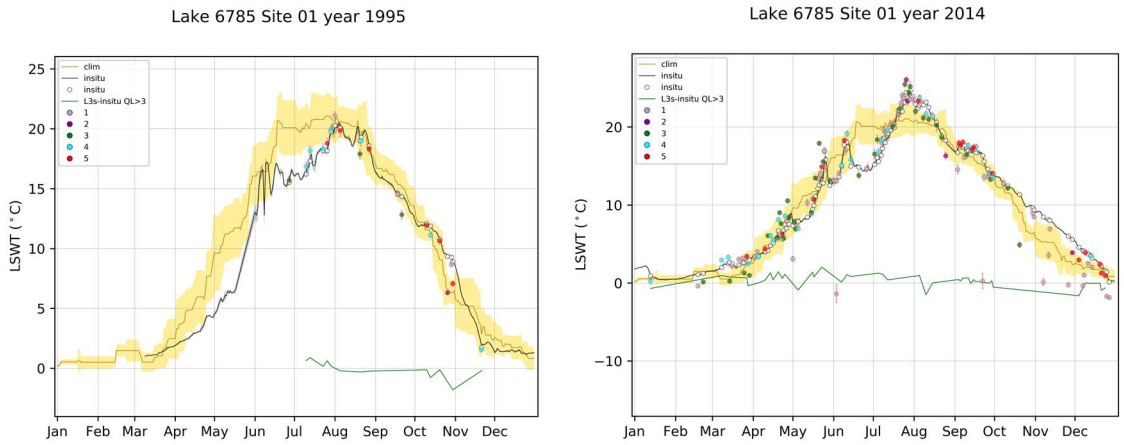


Figure 114. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for lake Erken in Sweden in 1995 (on the left) and 2014 (on the right).

Figure 115, Figure 116, Figure 117 and Figure 118 present lake Kasumigaura in Japan at the five sites in year 2016. The sites 01 and 03 are very close to the coast, so the matches are lower in number than for the other sites. However, a consistency between in situ and satellite LSWT can be observed for all the sites.

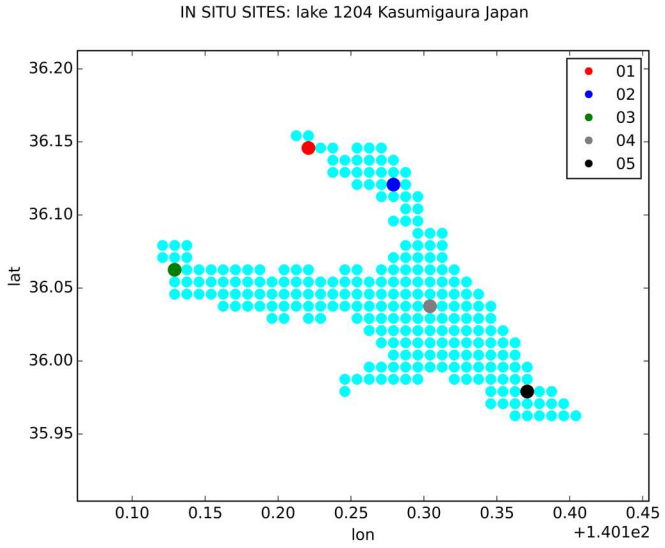


Figure 115. Location of in situ measurements for lake Kasumigaura in Japan. Each dot represents a 1/120°x1/120° cell.

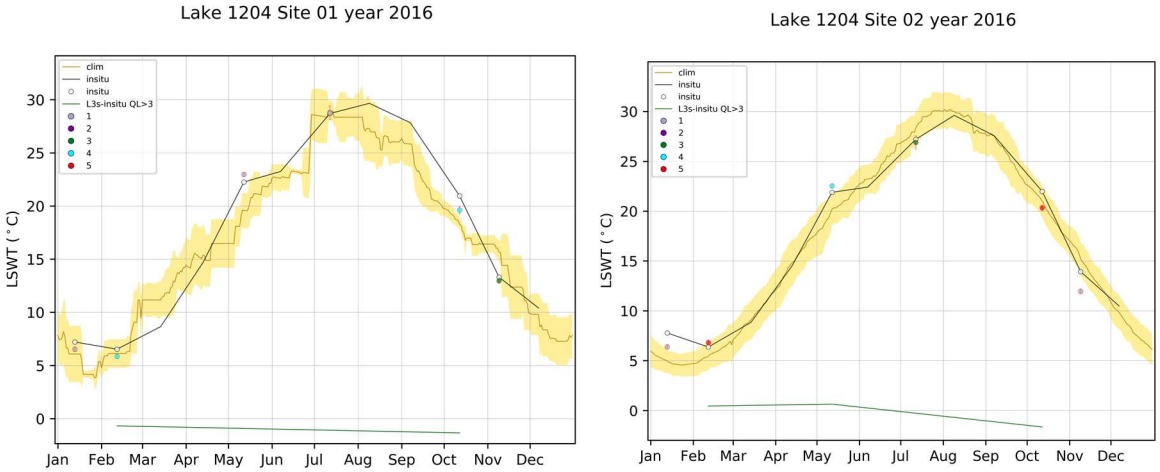


Figure 116. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for lake Kasumigaura in Japan in 2016 for site 01 (on the left), site 02 (on the right).

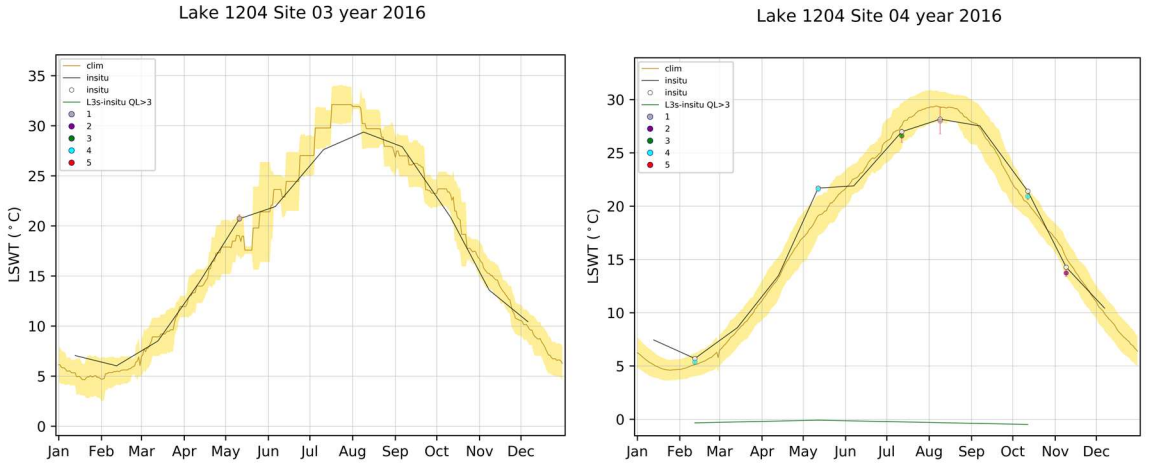


Figure 117. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for lake Kasumigaura in Japan in 2016 for site 03 (on the left), site 04 (on the right).

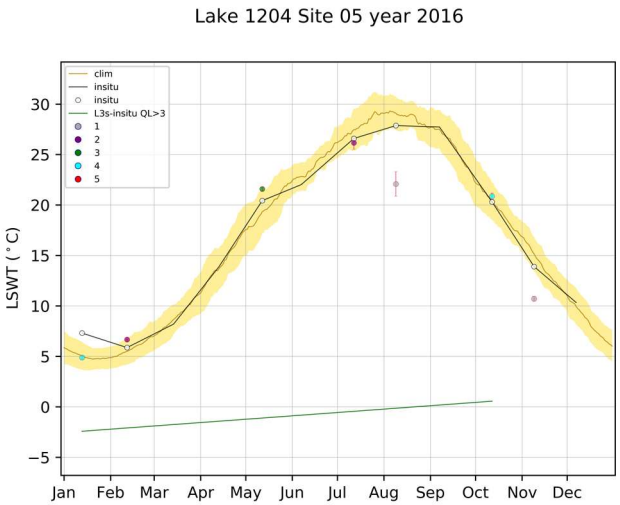


Figure 118. Satellite observations (dots), in situ matches (white dots), in situ measurements (black line), satellite minus in situ T difference for quality levels 4,5 (green line) and climatology (golden line with climatological variability as the yellow band) for lake Kasumigaura in Japan in 2016 for site 05.

4.4.2. Validation of the uncertainty LSWT v4.1

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

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

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

The in situ measurements uncertainty is not known for the data we have. We explore two assumptions: σ_{INSITU}=0.2 K, a value based from deployment of similar measurement technologies to the ocean, and σ_{INSITU}=0.5 K which would be at the upper end of our expectations for in situ uncertainty. The representativity effect is presently unquantified and we set it to 0 K for the present; neglecting representativity has the tendency to make the LSWT uncertainty look underestimated.

Lakes_cci products, σ_{LSWT}^2 is context sensitive and varies from match to match, which is why the validation approach involves the calculation of the above metric: the distribution of Δ should be a Gaussian distribution with mean equal to 0 and standard deviation equal to 1 when all standard deviations are well estimated and the retrieval is unbiased relative to the in situ and any mean geophysical effect. Figure 119 shows the histograms of the uncertainties per quality level where also the fitted Gaussian and the target Gaussian distributions are shown. In Figure 119 an in situ uncertainty of 0.2 K has been assumed.

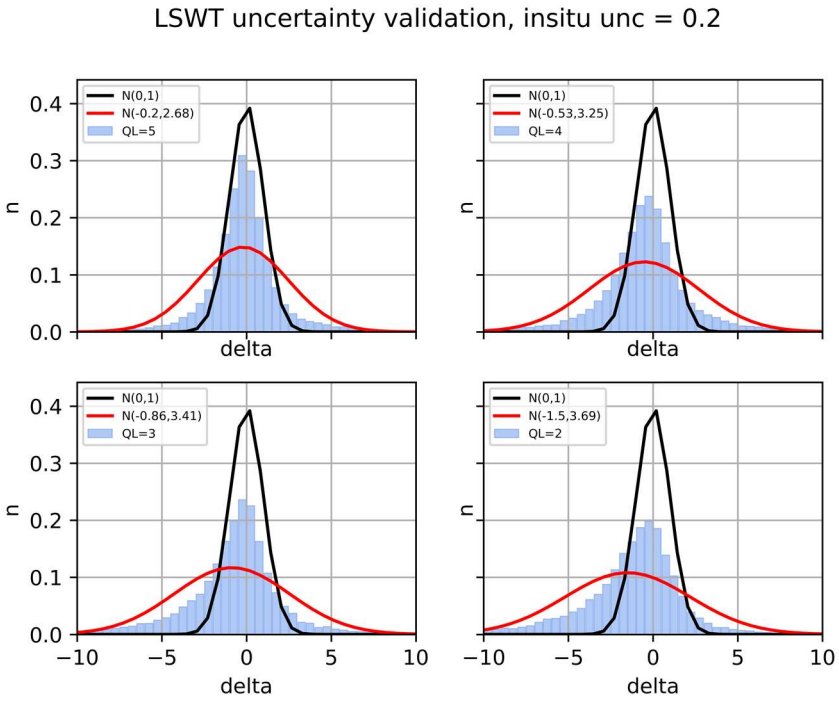


Figure 119. LSWT uncertainty validation (in situ uncertainty = 0.2K) per quality level (indicated in legend): histograms of D

For quality level 5, the Gaussian fit has width 2.68, which means that observed differences are more different than expected from the quoted uncertainties. This may be partly because the product uncertainties are underestimated, but could also arise to the degree that lake in situ data (being more diverse) have larger uncertainty than the assumed value (based on experience of ocean observations), and because representativity is neglected. Interpretation of this outcome is therefore currently ambiguous, and research is needed to better understand the in situ uncertainty and representativity effects.

We used an in situ uncertainty of 0.5 K to explore the level of in situ uncertainty that would better fit the Gaussian. Figure 120 shows the uncertainty validation for this value. The width of the Gaussian fit for quality level 5 is closer to one, and thus 0.5 K may be closer to the combined effect of in situ and representativity uncertainty. However, more investigations of in situ uncertainty and representativity need to be carried out to understand this better.

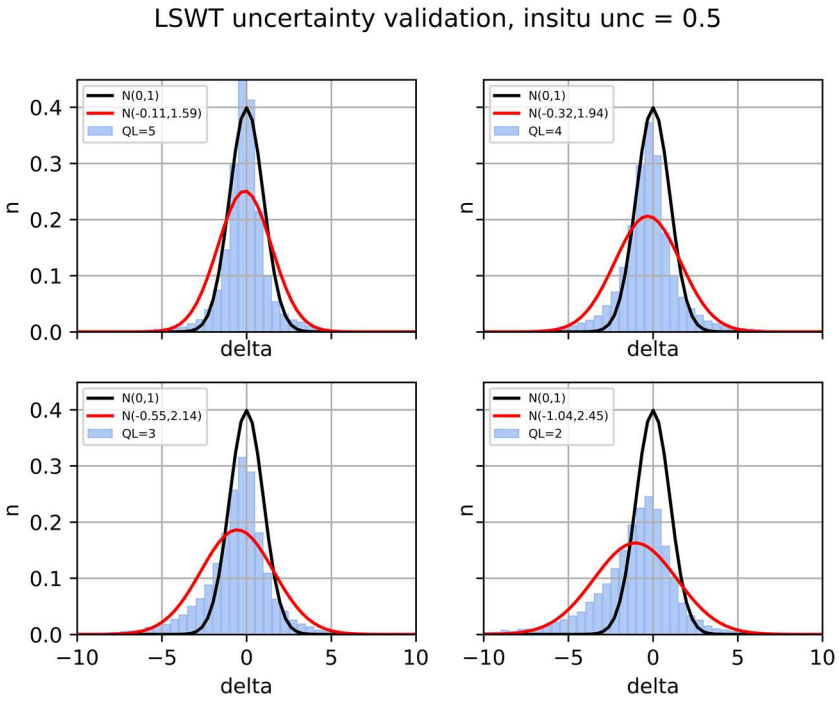


Figure 120. LSWT uncertainty validation (in situ uncertainty = 0.5K) per quality level (indicated in legend): histograms of D

4.4.3. Lakes with no LSWT

For 13 target lakes, no LSWT has been obtained, largely due to the fact that they are too small. The lakes are listed in Table 36 together with the estimated maximum distance to land. The majority of the lakes is not feasible because of their small size, except for the lake 1099 in Greenland, which was not successful due to missing climatology, and lakes 200000071 and 20000072 where in the mask used during the processing they were labelled as sea rather than lakes. For the next version of the Lakes_cci LSWT these lakes will be included.

Table 36. Lakes with no LSWT

CCI Lake id	Name	Country	Max distance to land (km)	Comments
1	Caspian	Azerbaijan/Russia/Turkmenistan/Iran	164.8	Processed as SST
1099	NN Glacial Lake	Greenland	1.8	Missing climatology
2637	Saysan	Kazakhstan	0.9	Too small
7889	Melvin	Ireland	1.3	Too small
18089	Macnean	United Kingdom	0.6	Too small
164651	Portmore	United Kingdom	0.5	Too small
208840	Mantua	Italy	0.5	Too small
215215	Morse	USA	0.5	Too small
215311	Geist	USA	0.5	Too small
215339	Eagle Creek	USA	0.6	Too small
100000004	Mourve	United Kingdom	0.2	Too small
200000071	Patos Lagoon	Brazil	25.2	Mis-abelled as sea
200000072	Maracaibo	Venezuela	46.5	Mis-abelled as sea
300134644	Rihpojavri	Norway	0.7	Too small

4.5. Conclusions and recommendations

The validation of the LSWT shows very good mean agreement (comfortably within +/-0.2 K) between satellite LSWT and independent in situ temperature measurements.

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

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

Manual inspection of all products for more than 1000 water bodies is impossible and in most cases requires local knowledge. The validation of the products is, and always will be, based on a small sample of well-studied areas. Users of these products are therefore advised to inspect the results for their area of interest before generating derivative products and any feedback to Lakes_cci would be most useful.

4.6. Acknowledgement

We would like to thank Iestyn R. Woolway for having established the contacts to set up the in situ database and all the institutions listed in Table 2 that have provided us with in-situ data and in particular: Enner Alcantara (Saõ Paulo State University, Saõ Paulo, Brazil), Gil Bohrer (The Ohio State University, Columbus, USA), Jean-Francois Cretaux (LEGOS, Toulouse, France), Hilary Dugan (Center for Limnology, University of Wisconsin-Madison, USA), Gideon Gal (Yigal Allon Kinneret Limnological Laboratory, Israel Oceanographic and Limnological Research, Migdal, Israel), Claudia Giardino (Istituto per il Rilavemento Elettromagnetico dell'Ambiente, National Research Council of Italy, Italy), April James (Nipissing University, Canada), Johanna Korhonen (SYKE, Helsinki, Finland), Ben Kraemer (Leibniz institute for freshwater ecology and inland fisheries, Berlin, Germany), Alo Laas (Estonian University of Life Sciences, Tartu, Estonia), Eric Leibensperger (Center for Earth and Environmental Science, SUNY Plattsburgh, USA), Junsheng Li (Institute of Remote Sensing and Digital Earth, Chinese Academy of Science, China), Alessandro Ludovisi (Dipartimento di Biologia Cellulare e Ambientale, Universita degli Studi di Perugia, Italy), Shin-ichiro Matsuzaki (National Institute for Environmental Studies, Japan), Ghislaine Monet (UMR CARRTEL, Thonon le Bains, France), Tiina Nogesand Peeter Noges (Estonian University of Life Sciences, Tartu, Estonia), Sajid Pareeth (Freie Universitat Berlin/Fondazione Edmund Mach, Germany/Italy), Sebastiano Piccolroaz (Institute for Marine and Atmospheric Research, Department of Physics, Utrecht University), Don Pierson (Uppsala University, Sweden), Pierre-Denis Plisnier, Antti Raike (SYKE, Helsinki, Finland), Alon Rimmer (Yigal Allon Kinneret Limnological Laboratory, Israel Oceanographic and Limnological Research, Migdal, Israel), Michela Rogora (CNR Institute for Water Research (CNR IRSA), Italy), Geoffrey Schladow (UC-Davis Tahoe Environmental Research Center, USA), Eugene Silow (Irkutsk State University, Russia), Lewis Sitoki (Department of Earth Environmental Science and Technology, Technical University of Kenya, Nairobi), Evangelos Spyrakos (Biological and Environmental Science, University of Stirling, Scotland, UK), Wim Thiery (Department of Earth and Environmental Sciences, KU Leuven, Belgium), Gesa Weyhenmeyer (Department of Ecology and Genetics, Uppsala University, Sweden).

5. Lake Ice Cover (LIC) parameter

Lake ice cover (LIC) refers to the extent (or area) of a lake covered by ice. In Lakes_cci, LIC is a daily merged level 3 (L3) product generated from MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath (MOD09/MYD09), Collection 6, data (see ATBD for details). In the LIC product, each pixel is assigned one of four possible class labels: water (value=1), ice (value=2), cloud (value=3), and bad (value=4; case where a retrieval was not possible due to poor data quality). Only pixels labelled as water, ice or cloud are considered in product validation. It is important to note that validation is performed on Terra and Aqua derived LIC individually, before merging into the daily Terra/Aqua L3 product and before aggregating into the ca. 1 km harmonized grid.

Product validation can be performed following three approaches: (1) comparison against ground-based nearshore observations; (2) comparison with LIC products generated by other algorithms/groups such as NASA's MODIS Snow Cover products from Terra and Aqua (MOD10/MYD10); and (3) validation against groups of pixels (Areas Of Interest or AOI) extracted for a selection of lakes from visual interpretation of original MODIS Terra and Aqua imagery used as input into the LIC retrieval algorithm (i.e. MOD09/MYD09). Here, validation involved a mix of approaches (2) and (3) whereby the thematic accuracy (water, ice and cloud) of both Lakes_cci LIC and NASA MODIS Snow Cover products was evaluated, and results compared. Approach (1) was not considered herein since ground-based nearshore observations were unavailable or non-existent at the time of writing of this report; however, most importantly, such observations are of a more limited value than the two other approaches for lake-wide validation of the LIC product.

5.1 Data description

5.1.1. MODIS Terra/Aqua surface reflectance product (MOD09/MYD09)

MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath, Collection 6 (Vermote et al., 2015) – MOD09/MYD09 false color composites (R: band 2, G: band 2, B: band 1) with a 250 m spatial resolution were used as reference images to manually collect AOIs with assigned labels (lake ice, open water, and cloud) to assess the accuracy of both the Lakes_cci LIC and NASA Snow Cover products. MODIS band 1 is centered at 645nm (red) and band 2 at 865 nm (near-IR).

5.1.2. MODIS Snow Cover product

MODIS Terra/Aqua Snow Cover 5-Min L2 Swath 500 m, Collection 6 (C6), daily products (MOD10/MYD10) were also validated and compared with the Lakes_cci LIC product. In MOD10/MYD10, lake ice cover is identified using the same criteria as snow over land (i.e. Normalized Difference Snow Index (NDSI) > 0 and data screens) and a 500 m lake mask. The NDSI is derived from MODIS radiance data acquired by Terra/Aqua satellites (i.e. MODIS Level 1B product (MOD02/MYD02) top-of-the-atmosphere (TOA) calibrated radiance/reflectance data). Input data to the Snowmap algorithm used in the production of MOD10/MYD10 are shown in Table 37.

Table 37. MODIS data inputs for the collection 6 (C6) snow algorithm (adapted from Riggs et al., 2016)

Input Data	Product Description
MOD02HKM	Calibrated Radiances (500 m)
MOD021KM	Calibrated Radiances (1 km)
MOD03	Geolocation Fields (1 km)
MOD35_L2	Cloud Mask and Spectral Test Results (250 m and 1 km)

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

The MODIS Snow Cover product experiences issues in differentiating between clouds and snow/ice, and both errors of commission and omission can be found in the product. Omission occurs when the MODIS cloud mask (MOD35) identifies an area of snow/ice as certain cloud, therefore excluding it from the MODIS snow product (Riggs et al., 2016). Errors of commission can occur on the periphery of cloud cover or where clouds appear similar to snow/ice and are excluded from the cloud mask and instead identified as snow/ice (Riggs et al., 2016). Cloud cover can be removed from the product by extracting the classification value from the previous and proceeding pixel to determine the possible classification of a cloud covered pixel. High solar zenith angle can also prevent accurate classification of snow/ice cover, which is important when studying freeze-up at northern high latitudes. Pixels are screened using a solar zenith angle mask of > 70° and pixels with a solar zenith angle ≥ 85° are classified as night.

5.2 Comparison methods

Validation of the Lakes_cci LIC product and NASA MODIS Snow Cover product (also used for comparison) has been performed through computation of confusion matrices built on independent statistical validation. The reference data for validation were collected for water, ice and cloud as AOIs from the visual interpretation of the false colour composite surface reflectance images (MOD09/MYD09), as described in section 5.1.1, over three ice seasons (freeze-up and break-up periods) interspersed across the 20-year MODIS record by skilled ice analysts.

5.3 Description of work

A total of 17 lakes have been selected across the Northern Hemisphere to serve for the purpose of both development and validation of the Lakes_cci LIC product (Figure 121). Samples were collected for three ice seasons (2002-2003, 2009-2010, 2016-2017) as to include MODIS data from Terra (2000-present) and Aqua (2002-present), and provide a good temporal spread over the full record to ensure algorithm stability. For each lake, one image from the freeze-up period and one image from the break-up period were selected for both Terra and Aqua images if available. False color composites (R: Band 2, G: Band 2, B: Band 1) with a 250 m spatial resolution were used as reference images to manually extract AOIs with labels (ice, water, and cloud). In total 108 images (40 from FU, 68 from BU) were selected and 552,006 pixels were sampled from the selected AOIs. Of that sample, 30% of the pixels (n = 165,697) were randomly selected for the purpose of product (Lakes_cci LIC and MODIS Snow Cover) validation and comparison (the other 70% of pixels were used for algorithm development).

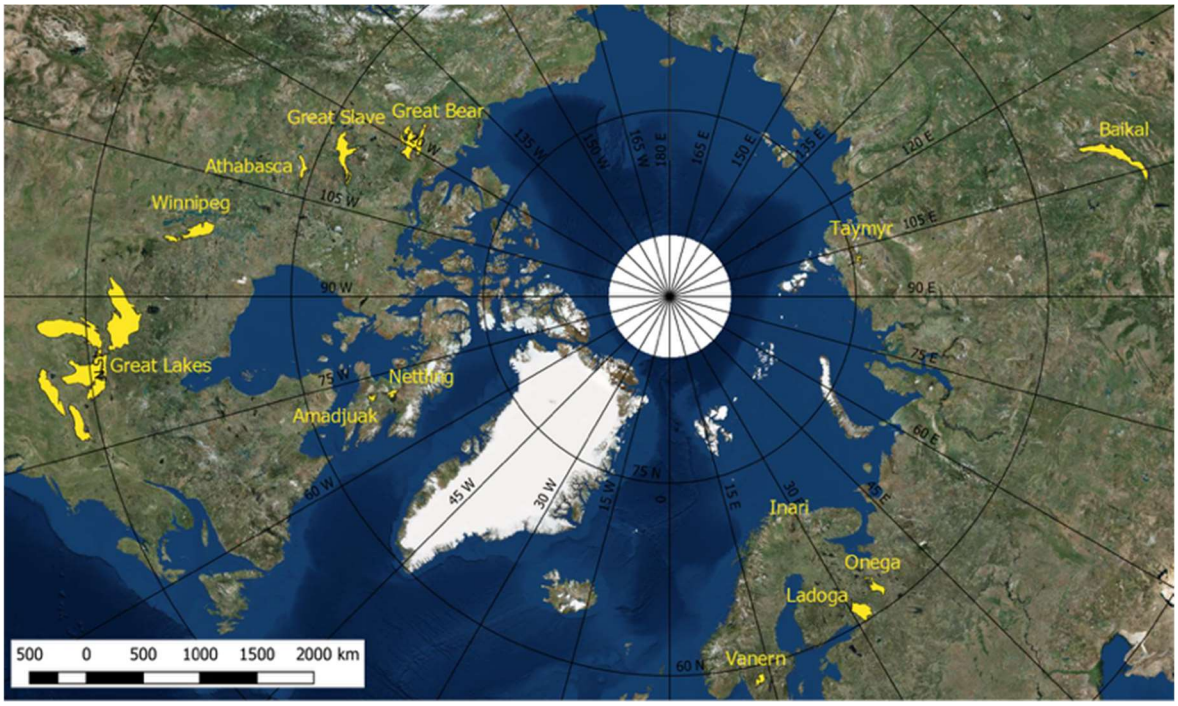


Figure 121. Geographical distribution of lakes used for Lakes_cci LIC algorithm development and validation

5.4 Result analysis

Table 38 shows the overall accuracy (OA) of the Lakes_cci LIC and NASA Snow Cover products (break-up and freeze-up periods combined). The algorithm developed for Lakes_cci provides a higher overall accuracy (95.54%), and also for individual classes (91.71% for ice cover, 98.85% for water, and 95.63% for cloud cover) compared to the existing MODIS Snow Cover products (MOD10/MYD10). The OA of MOD10/MYD10 is 87.09%. The accuracy of 74.55% for lake ice cover in MODIS Snow Cover is related to an underestimation of ice cover caused by the overestimation of cloud cover, which attains a total class accuracy of 93.32%. In addition to ice, water, cloud, MOD10/MYD10 products contain other possible labels (No Decision, Night, Missing, and Ocean).

Table 38. Confusion matrices with retrieval accuracies for Lakes_cci LIC and MODIS Snow Cover products (break-up and freeze-up periods combined)

		Retrieval Algorithm (classification)					
Lakes_cci LIC		Ice	Water	Cloud	Other		Accuracy
User-defined	Ice	46968	2737	1510	N/A		91.71%
	Water	506	57435	165	N/A		98.85%
	Cloud	2272	192	53912	N/A		95.63%
			Overall Accuracy: 95.54%				
MODIS Snow Cover		Ice	Water	Cloud	Other *		Accuracy
User-defined	Ice	38328	3911	7198	1978		74.55%
	Water	558	53178	1790	2179		92.15%
	Cloud	1546	857	52795	1379		93.32%
			Overall Accuracy: 87.09%				

Other: For MODIS Snow (MOD10/MYD10) this class includes lake pixels that were identified as “No Decision, Night, Missing or Ocean.”

Table 39 provides the accuracies reached by the Lakes_cci LIC and MODIS Snow Cover products for the break-up (BU) and freeze-up (FU) periods taken individually. There is no notable difference in the accuracy of the Lakes_cci LIC product between the BU (OA: 95.80%) and FU (OA: 95.12%) periods, and the classification accuracies are consistent across classes. Overall accuracies for the NASA's MODIS Snow Cover are lower by about 7% during BU and 10% during FU. Lower class accuracy is also achieved for the NASA product with regards to ice cover during both the BU (77.64%) and FU (69.49%) periods compared to Lakes_cci LIC (BU: 91.61%, FU: 91.87%). Differences are in the order of 14% and 22% for BU and FU, respectively, with the Lakes_cci LIC retrieval algorithm outperforming NASA's Snowmap algorithm.

Table 39. Confusion matrices with retrieval accuracies for Lakes_cci LIC and MODIS Snow Cover products (break-up and freeze-up periods individually)

		Break-Up					**Freeze-Up**				
		Retrieval Algorithm (classification)					Retrieval Algorithm (classification)				
Lakes_cci LIC		Ice	Water	Cloud	Other	Accuracy	Ice	Water	Cloud	Other	Accuracy
User-defined	Ice	29128	2535	133	N/A	91.61%	17840	202	1377	N/A	91.87%
	Water	16	39520	47	N/A	99.84%	490	17915	118	N/A	96.72%
	Cloud	1418	162	29750	N/A	94.96%	854	30	24162	N/A	96.47%
		Overall Accuracy: 95.80%					Overall Accuracy: 95.12%				
MODIS Snow Cover		Ice	Water	Cloud	Other	Accuracy	Ice	Water	Cloud	Other	Accuracy
User-defined	Ice	24760	3528	2016	1585	77.64%	13568	383	5182	393	69.49%
	Water	182	37132	414	1605	94.40%	376	16046	1376	574	87.34%
	Cloud	1009	691	28890	897	91.75%	537	166	23905	482	95.28%
		Overall Accuracy: 88.39%					Overall Accuracy: 84.97%				

Other: For MODIS Snow Cover (MOD10/MYD10) this class includes lake pixels that were identified as "No Decision, Night, Missing or Ocean."

The Lakes_cci LIC product also captures better the spatial distribution of cloud cover compared to the MODIS Snow Cover (MYD10/MOD10) (Figure 122). MOD10/MYD10 Snow Cover products overestimate cloud cover on all three dates for the lakes shown as examples.

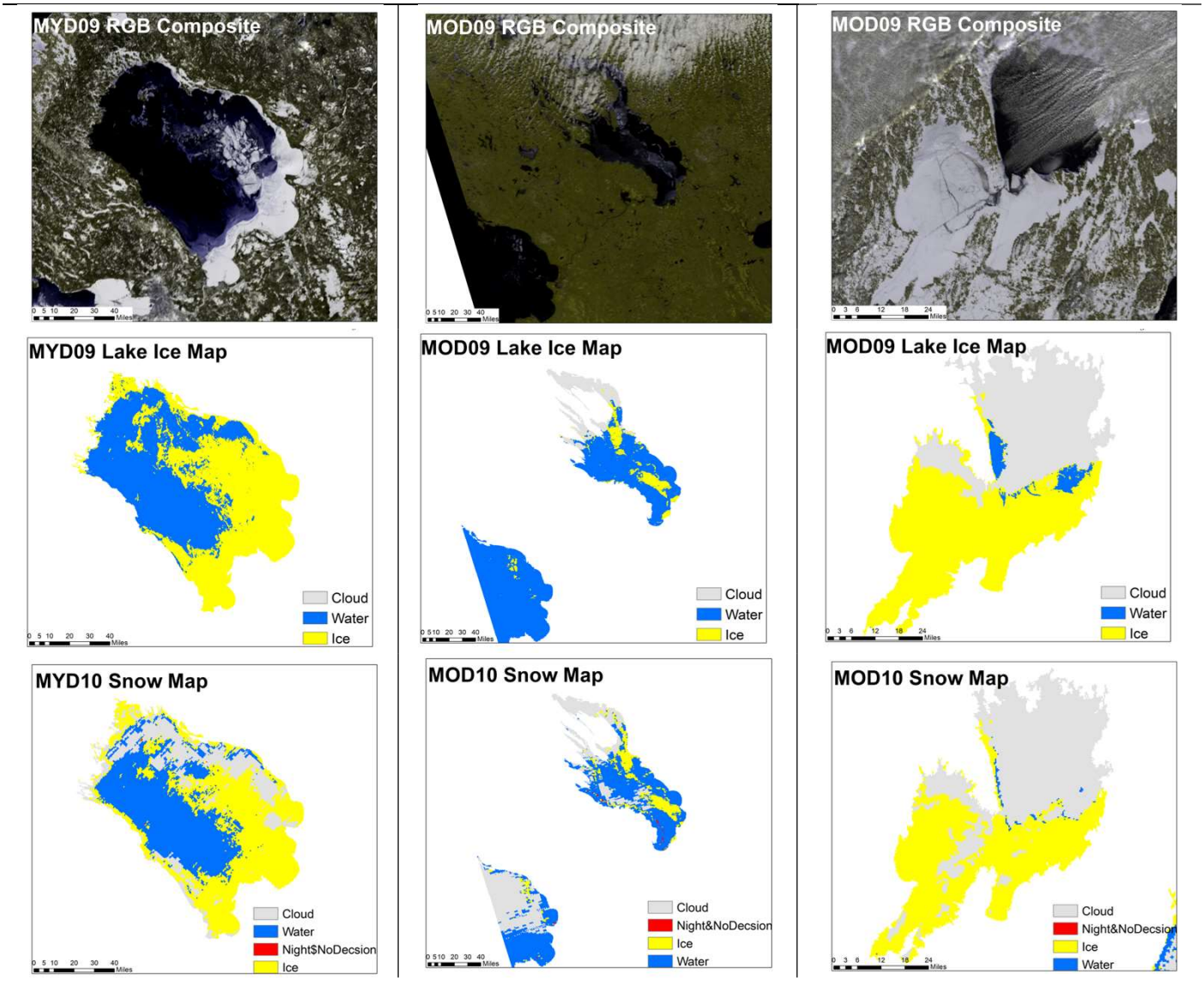


Figure 122. MYD09/MOD09 (surface reflectance) colour composites (top), MYD09/MOD09 derived Lakes_cci LIC product (middle) and MYD10/MOD10 NASA Snow Cover product (bottom) over Lake Ladoga (a), Lake Onega (b), and Lake Vanern (c)

Another source of error for the MODIS Snow Cover (MOD10/MYD10) product is the geolocation of some lakes. MOD10/MYD10 employs the MODIS Geolocation L1A product in order to mask land and water (Riggs et al., 2016). However, as shown in Figure 123, the MODIS Snowmap algorithm classifies the majority of Lake Nettilling (Canada) as ocean resulting in no classification for the ice cover on the lake. However, the more accurate lake area mask used for the generation of the Lakes_cci LIC product provides correct classification of the ice conditions on that day.

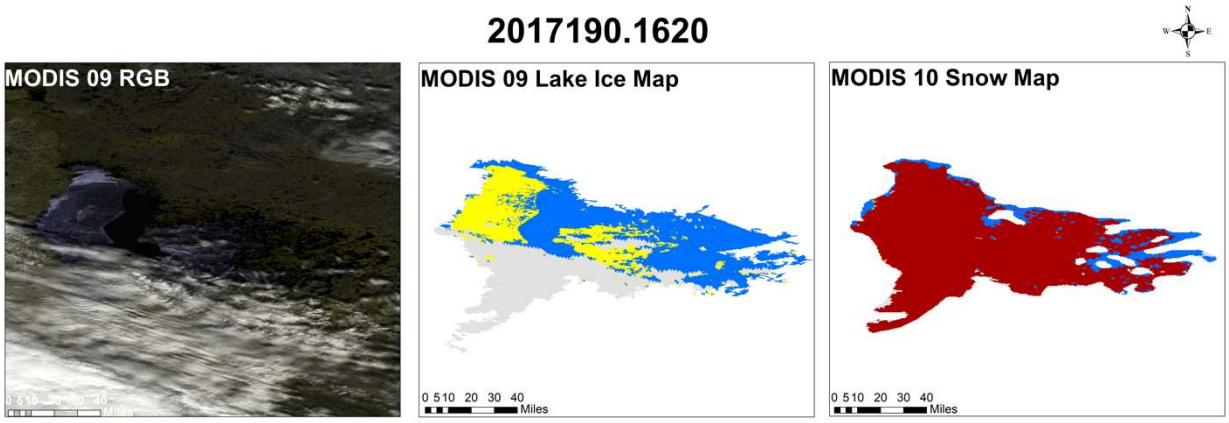


Figure 123. MODIS surface reflectance RGB colour composite (left), MODIS-derived Lakes_cci LIC (middle), and MODIS Snow Cover product (right) for Lake Nettilling (Canada) during the BU period on 9 July 2017 (UTC 16:20). The red colour corresponds to "ocean" in this MODIS Snow Cover product

5.5 Conclusions and recommendations

For Lakes_cci, LIC CDRP V1.0 is generated from a threshold-based retrieval algorithm using MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath (MOD09/MYD09), Collection 6, as primary data input. Lakes_cci LIC is a gridded product where cells are assigned one of four possible labels: water, ice, or bad (case where a retrieval was not possible due to poor data quality).

Validation of the Lakes_cci LIC product was performed via computation of confusion matrices built on independent statistical validation. Reference data for validation of class labels (water, ice, cloud) were collected from 17 lakes in Europe and North America over three ice seasons (2002-2003, 2009-2010, 2016-2017) through visual interpretation of the false color composite surface reflectance images (MOD09/MYD09) by skilled ice analysts. The reference data also served to validate lake ice/water/cloud retrievals contained in NASA's MODIS Terra/Aqua Snow Cover products (MOD10/MYD10), which were then compared to those from Lakes_cci LIC.

Results show that the retrieval algorithm implemented for Lakes_cci LIC production outperforms NASA's Snowmap algorithm, attaining an overall accuracy (FU and BU periods combined) of 95.54% compared to 87.09%. Retrieval accuracies are also found to be more consistent between classes and also ice periods for the Lakes_cci algorithm (BU: 14% and FU: 22% higher accuracy than NASA's Snowmap algorithm), in addition to better capturing the spatial distribution of cloud cover compared to the MODIS Snow Cover (MYD10/MOD10). Individual class accuracies are all above 90% (errors less than 10%) for LIC CDRP V1.0 (ice: 91.71%, water: 98.85%, cloud: 95.63%) which meet uncertainty requirements of 10% set by GCOS for LIC ECV (see E3UB document for details on determination of errors and uncertainty in LIC product).

Further assessment of the Lakes_cci LIC product and its comparison with other products is planned in the future. This includes comparison with: 1) nearshore in situ ice/open water observations if such observations are available for any of the 120 lakes forming ice cover in LIC CDRP V1.0 release (TBD); 2) NOAA's Interactive Multisensor Snow and Ice Mapping System (IMS) daily 1 km global product (available since December 2014); 3) daily ice charts of the Great Lakes produce by the North American Ice Service; and 4) weekly ice fraction product from the Canadian Ice Service. Ice cover observations from these various sources will also be valuable for assessing the quality of LIC CDRP V2.0 prior to its release.

6. Lake Water-Leaving Reflectance (LWLR) parameter

Lake Water-Leaving Reflectance (LWLR), also referred to as water colour, is the measurement of the quantity of sunlight reaching the remote detector after interaction with the water column. The validation and comparison of the LWLR products, including LWLR, chlorophyll-a (Chl-a) and total suspended matter (either as TSM or expressed as Turbidity), is based on matchups between in situ and satellite measurements.

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

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

6.1. Data description

6.1.1. In situ data

The validation dataset used in this study comprised 17 individual datasets from lakes and inland water bodies across the globe requested from the LIMNADES repository. This combined data set consisted of 1982 individual observations of remote sensing reflectance (Rrs, sr^{-1}), 28726 for Chl-a (mg/m^3) and 6955 of total suspended matter (TSM, mg/m^3). TSM measurements were more numerous than Turbidity observations. In the satellite products the two are interchangeable through a single conversion factor so results will be comparable.

6.1.2. Satellite data

Due to scarcity of recent in situ data for the OLCI observation period, the validation study described in this section is still based on MERIS. The 3rd reprocessing of MERIS full resolution L1B data was used for CDRP V1.0. These were processed to LWLR by applying radiometric calibration (SNAP toolbox, see ATBD for details) and Polymer (v4.6; Steinmetz et al. (2011)). We note that the version of Polymer used in the CDRP v1.0 is the latest (v4.12) but this does not cause differences in the handling of MERIS data. The L1B data were masked using a combination of L1B and Idepix flags (see ATBD for details). The masks applied were Cosmetic, duplicated, glint risk, suspect, land/ocean, bright, coastline, and invalid from the L1B product and invalid, cloud, cloud ambiguous, cloud sure, cloud buffer, cloud shadow, snow ice, bright, white, coastline, land and glint risk from Idepix.

6.2. Comparison methods

Product validation of the LWLR and derived products (Chl-a and TSM) is based on comparison against in situ observations. Results presented here are for in situ validation carried out against satellite observations with the MERIS sensor, for which the most in situ data are available by far.

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

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

6.3. Description of work

6.3.1. Lake Water-Leaving Reflectance

For water-leaving reflectance, comparisons between the in situ and satellite measurements were performed for each band. In this report, the atmospheric-corrected MERIS water-leaving reflectance were converted to remote sensing reflectance (Rrs = Rw/π [sr^{-1}]) to facilitate the comparison between the in situ and satellite measurements.

6.3.2. Chlorophyll-a and TSM

After extensive product validation in the GloboLakes project, the water constituent retrieval algorithms listed in Table 40 were identified as showing individual best performance against sets of Optical Water Types (Neil et al. 2019). For Chl-a, the algorithms included for validation are Gons05 (Gons et al. 2005), the NASA OC2 algorithm (O'Reilly et al. 1998), R708/R665 (Gilerson et al. 2010), and QAA (Mishra et al. 2013a; Mishra et al. 2013b). For TSM, the algorithms include those of Zhang et al. (2014), Vantrepotte et al. (2011), and Binding et al. (2010).

These algorithms are mapped to pixels depending on their similarity to a set of Optical Water Types (OWT), determined from in situ reflectance data in the GloboLakes project (Spyrakos et al. 2018). The assignment of algorithms to each OWT is shown in Table 40, which also provides the calibrated algorithm coefficients. It is noted that while the methodology of algorithm tuning is as described in Neil et al. (2019), tuned coefficients may differ since the former are derived from in situ reflectance data against LIMNADES while *Calimnos* uses coefficients optimized for satellite-derived Polymer-corrected water-leaving reflectance (Rw).

Table 40 Chlorophyll-a/TSM algorithms per optical water type and tuned parameters

Product	Algorithm	Optical Water Type number	Parameters	Tuned
Chl-a	OC2 oceancolor.gsfc.nasa.gov/cms/atbd/chlor_a	3, 9, 10, 13	a0, a1, a2, a3, a4	0.1731, -3.9630, -0.5620, 4.5008, -3.0020
	708/665 empirical band ratio based on Gilerson et al. (2010)	2, 8, 11, 12	a, b, c	79.62, 0.7393, -54.99
	Semi-analytical NIR-Red band algorithm for MERIS based on Gons et al. (2005).	1, 4, 5, 6	aph*	0.025
	Adapted QAA algorithm according to Mishra et al. (2013)	7	S_{CDOM}	0.0135
TSM	Based on Zhang et al. (2014)	1, 7, 10	a, b	2524, 1.113
	Based on Vantrepotte et al. (2011)	2, 4, 6, 8, 12	a, b, c	206.4, 20460, 0.7921
	Based on Binding et al. (2010)	3, 5, 9, 11, 13	b^*_{SPM}	0.664

The satellite retrieved water constituent products provided in the CDRP are based on a weighted blending procedure, recombining the individual algorithm results with the weighting determined by their OWT membership scores. The selected algorithms are mapped to individual satellite measurement (per pixel) from the OWTs with the top-3 classification scores. The algorithm results corresponding to those three OWTs are averaged using the membership score as weighting factor, after normalizing the scores between 0 and 1 where 1 is the highest score and 0 is the score of the 4th ranking OWT. The derived Chl-a or TSM satellite products following this procedure were compared with the in situ matchups for validation for this analysis.

In this report, the validation of the selected algorithms is performed based on the in situ TSM observations. Turbidity is obtained using a conversion factor of 1.17 NTU/g m^{-3} as formulated by (Nechad et al. 2016; Nechad et al. 2010).

6.4. Validation results

6.4.1. Validation of LWLR

Previous round-robin comparisons of atmospheric correction algorithms for MERIS carried out in GloboLakes showed that Polymer yielded the statistically most robust retrieval of reflectances, although a systematic negative bias was observed. This lead to Polymer being adopted as the state-of-the-art atmospheric correction processor for Lakes_cci to provide LWLR. Figure 124 shows R_{rs} matchup results for MERIS in 11 lakes at 11 wavebands from 412 nm to 779 nm, for which in situ data were available in LIMNADES. In general, significant linear relationships were found between the MERIS and in situ R_{rs}, with the highest R^2 of 0.67 returned in the 560 nm band and outliers in the blue bands associated with the hypereutrophic HARTBEESPOORT-Reservoir leading to the relatively worst performance. Systematic underestimation of MERIS R_{rs} is still observed. From inspecting the individual components of the atmospheric correction procedure, the underestimation is understood to be the result of overestimating the atmospheric path radiance with increasingly turbid waters, rather than a failure of the in-water bio-optical model or glint retrieval.

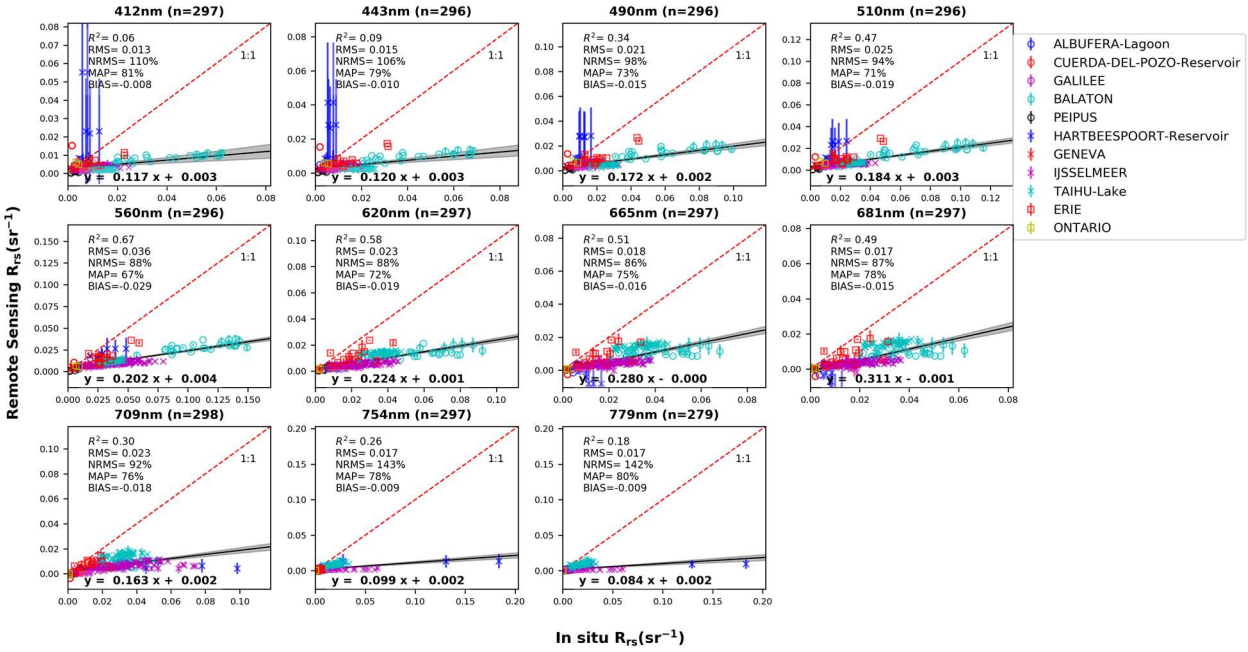


Figure 124 Comparison between in situ and MERIS R_{rs} in each band

To show the level of spectral consistency in the validation data set, Figure 125 presents averaged spectra of all R_{rs} matchups for each individual lake. There is general agreement between the in situ and satellite derived reflectance spectra, with the exception of the Cuerda-del-Pozo Reservoir.

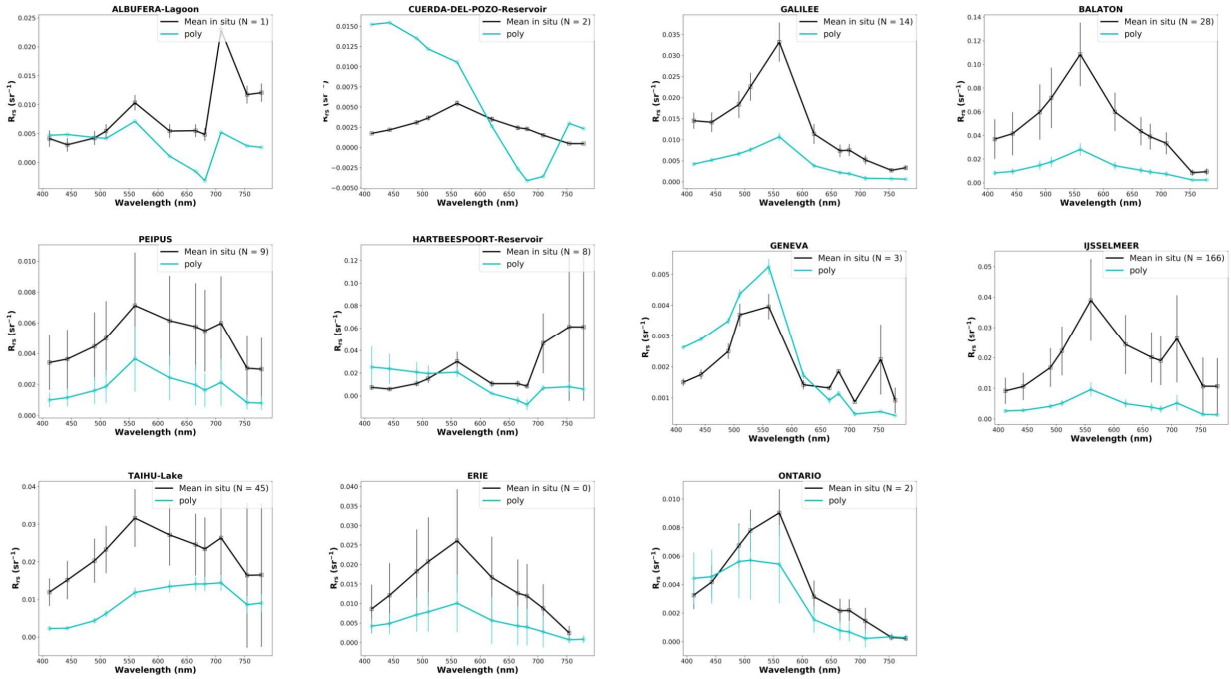


Figure 125 Spectral comparison of in situ and MERIS R_{rs}

6.4.2. Validation of chlorophyll-*a* products

Scatter plots of Chl-a validation results for each pre-tuned algorithm are shown in Figure 126, using the whole data set regardless of optical water type classification. For the OC2 algorithm, although it returned the highest Normalized Root Mean Squared Error (NRMS) of 172%, the differences are evenly distributed around unity resulting in a slope of linear regression line close to 1. Saturation of the algorithm appears, as may be expected, at concentrations > 10 mg m^{-3}. Algorithms R708_R665, Gons05, and QAA show a general overestimation at low Chl-a values, and an underestimation at high Chl-a values. The highest R^2 of 0.42 and lowest Normalized-Root-Mean-Squared-Error (NRMS) of 38% are returned by the QAA among the four algorithms.

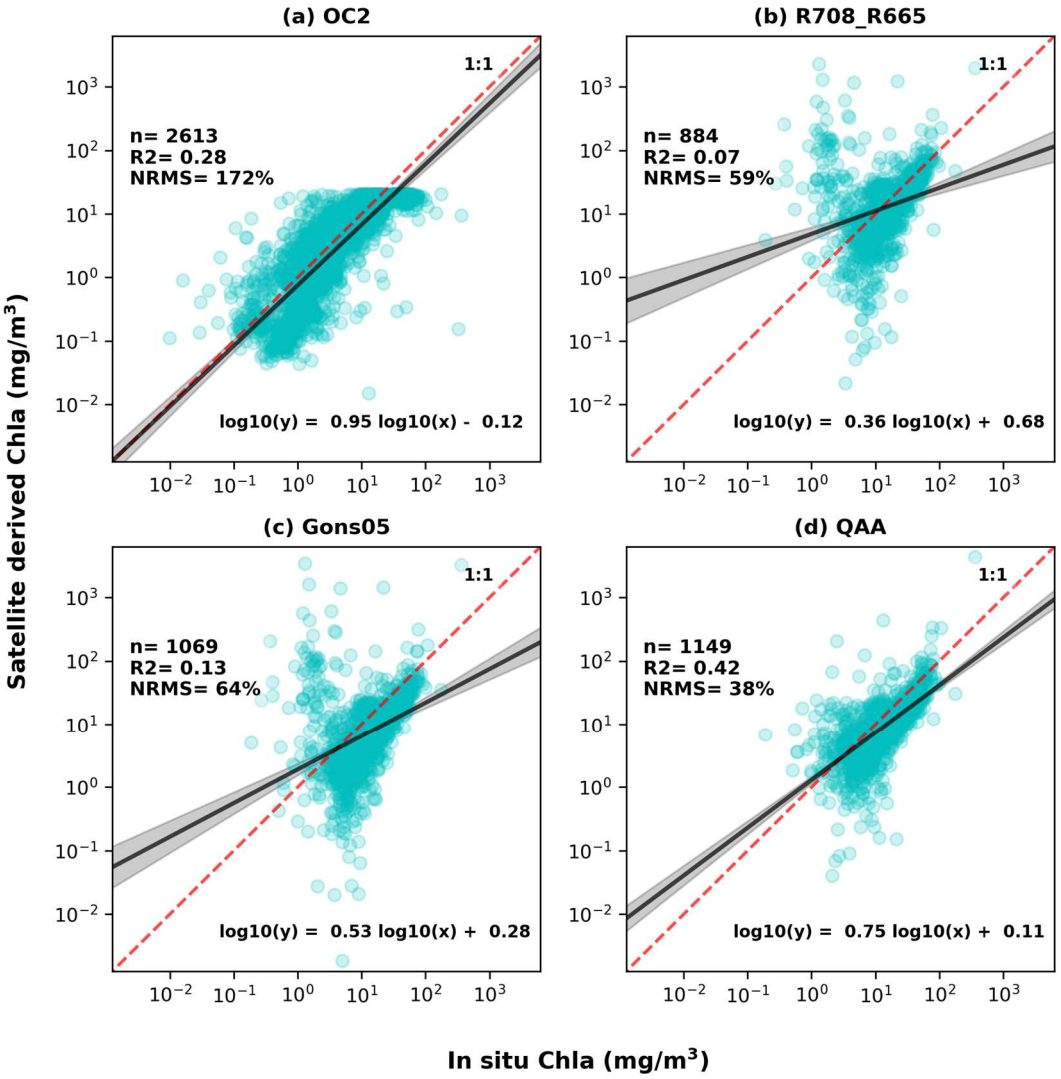


Figure 126 Comparison between in situ and MERIS-derived chlorophyll-a using the (a) OC2, (b) R708_R665, (c) Gons05, and (d) QAA algorithms.

The agreement between in situ and satellite retrieved Chl-a improves dramatically when the input algorithms are blended according to the per-pixel Optical Water Type membership (Figure 127). The regression line is close to unity, with an R^2 of 0.69 and NRMS of 78%, both measured in log space.

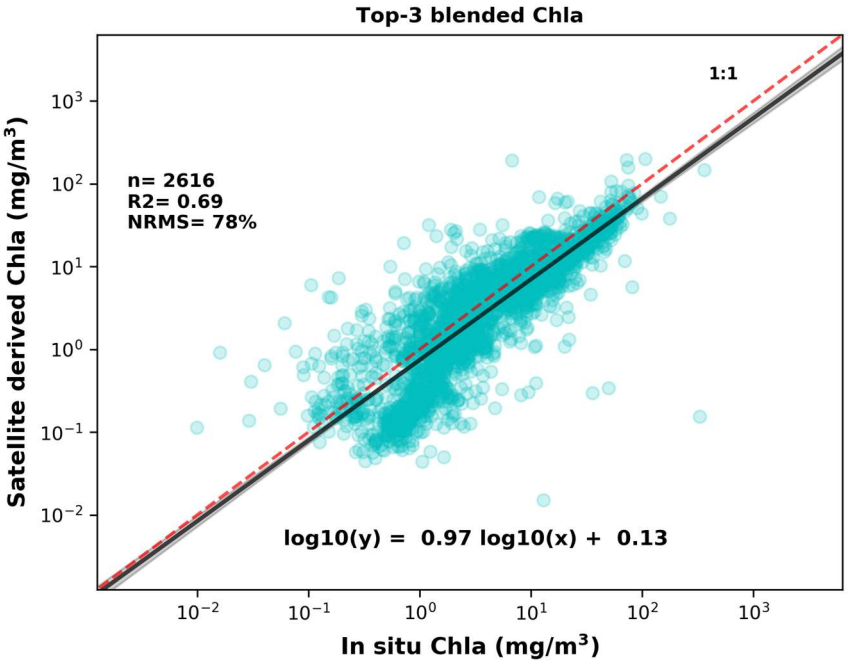


Figure 127 Comparison between in situ and top-3 blended chlorophyll-a derived from MERIS.

6.4.3 Validation of Total Suspended Matter products

The performance of individual TSM algorithms is shown in Figure 128, comparing in situ TSM measurements to those retrieved from MERIS matchup data. A general underestimation at the high TSM section was observed for all of the three algorithms. The slope of the regression line (expressed in linear scaling) for the Zhang algorithm approached unity with an intercept of -0.296 g m^{-3} suggesting systematic underestimation. The Vantrepotte algorithm showed the lowest NRMS (58%) among the three algorithms, although the lowest R^2 (0.29) was also returned by this algorithm indicating it lacks broad sensitivity. The Binding algorithm showed the highest R^2 of 0.65 and with the largest NRMS of 65%. The Binding algorithm shows no decrease in sensitivity with increasing concentrations, despite the increasing bias. These results suggest that a re-tuning of the Zhang and Binding algorithms from the original analysis carried out in the GloboLakes project would likely improve overall performance.

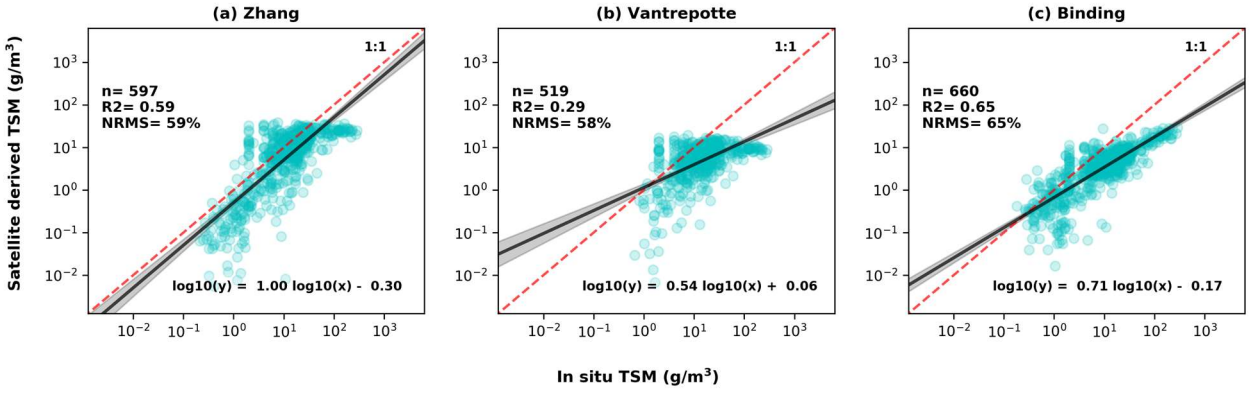


Figure 128 Comparison between in situ and retrieved TSM using the (a) Zhang, (b) Vantrepotte, and (c) Binding algorithms.

Agreement between the in situ and satellite retrieved TSM slightly improved with algorithm blending by OWT membership (Figure 129). Negative bias at the high TSM values is significant and confirms that re-turning of the algorithms, particularly the Binding algorithm, will likely improve the results. At present, the R^2 is 0.61 with NRMS = 54%.

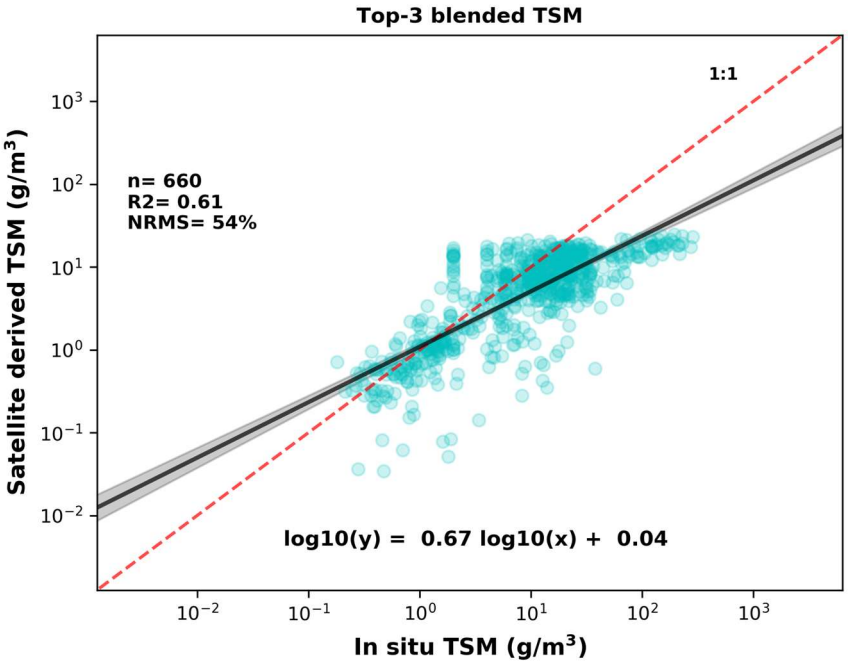


Figure 129 Comparison between in situ and top-3 blended TSM

6.5. Conclusions and recommendations

The Lakes_cci has inherited a large body of work to dynamically map algorithms to optical water types from the GloboLakes project. However, much of this work has not yet been fully published while the availability of in situ data has somewhat improved since. Thus it comes as no surprise that further algorithm optimisation is essential.

A systematic negative bias was observed in the R_{rs}, due to the challenges that are faced with accurately performing atmospheric correction in optically complex inland waters. This underestimation in R_{rs} is propagated to the derived Chl-a and TSM, but corrected again by end-to-end algorithm tuning. This is a necessary extra step because in situ matchup data of Chl-a and TSM are far more numerous than those including R_{rs}.

Validation of the Chl-a algorithms shows that the weighted-blending procedure dramatically improved the retrieval performance compared to that of the individual algorithms. For TSM, the agreement between the in situ and satellite retrieved TSM slightly improved after the blending stage compared to the individual algorithms, which show room for individual improvement using currently available data sets.

The validation presented in this report is only based on the MERIS matchups, because of the scarce in situ data availability for other sensors. Further community effort and plenty of patience are likely required to fill these data gaps. Analyses completed in the Copernicus Land Monitoring Service, which uses the same configuration of *Calimnos*, have shown that there is consistency between per-lake time series of Chl-a and Turbidity observed with MERIS (2002-2012) and OLCI (2016-present), which strongly suggests that the same algorithms may be applied to both sensors until further in situ reference data become available. However, the additional bands available on OLCI may well be able to improve on aspects of the retrieval of both LWLR and the derived

biogeochemical products. An independent validation of algorithms for MODIS is expected to accompany CDRP v2, necessitated by differences in band configuration of MODIS compared to MERIS and OLCI.

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

7. References

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

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

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

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

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

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

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

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

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

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

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

Riggs, G. A., Hall, D. K., & Román. M.O. (2016). *MODIS Snow Products Collection 6 User Guide.*

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

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

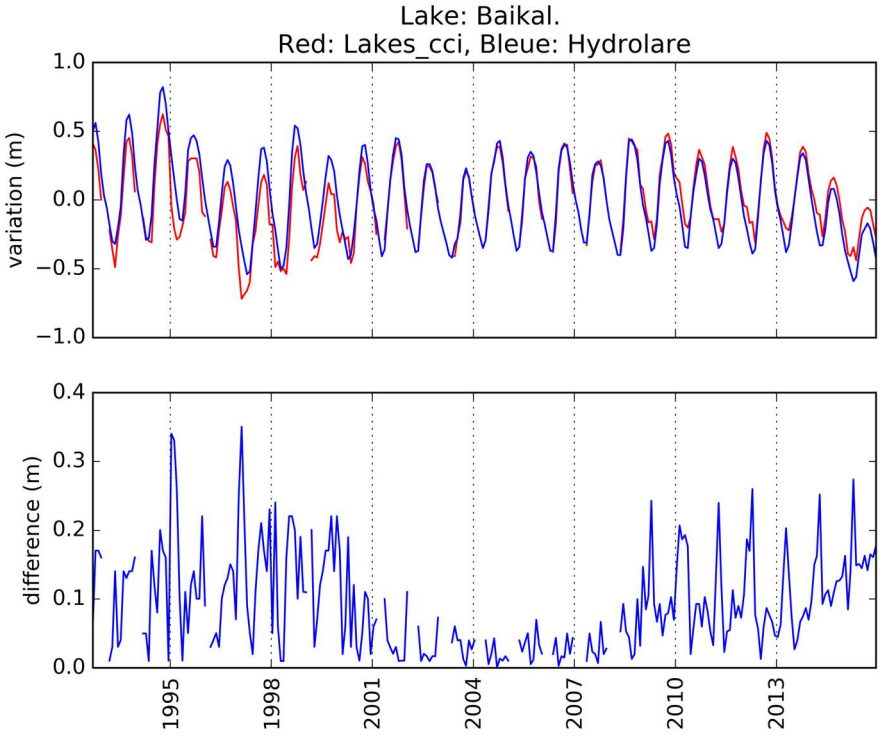
Steinmetz, F., Deschamps, P.-Y., & Ramon, D. (2011). Atmospheric correction in presence of sun glint: application to MERIS. Optics Express, 19, 9783-9800

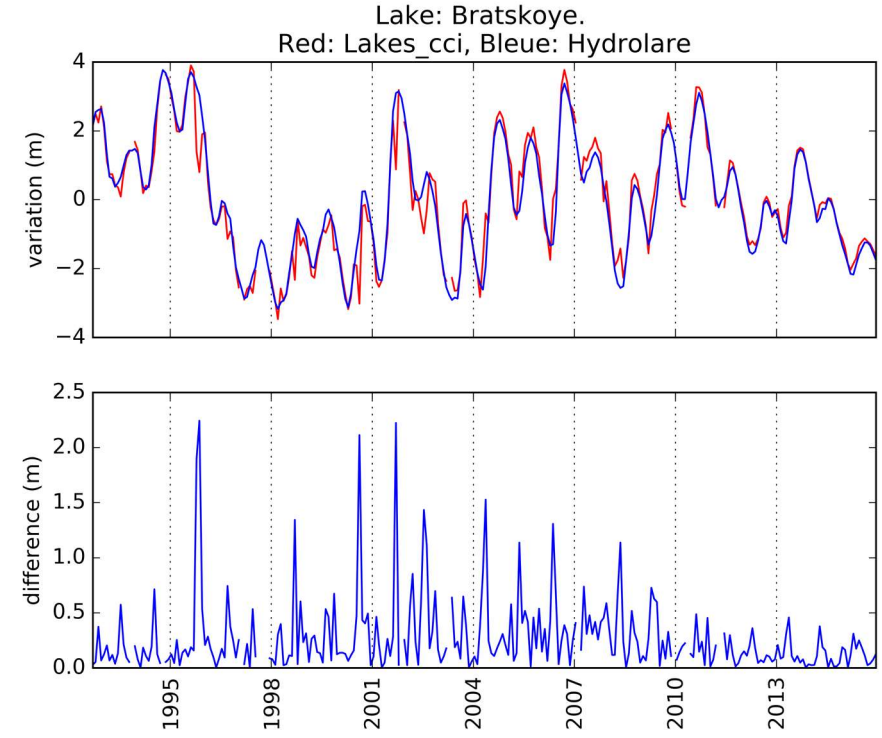
Vantrepotte, V., Loisel, H., Mériaux, X., Neukermans, G., Dessailly, D., Jamet, C., Gensac, E., & Gardel, A. (2011). Seasonal and inter-annual (2002-2010) variability of the suspended particulate matter as retrieved from satellite ocean color sensor over the French Guiana coastal waters. Journal of Coastal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

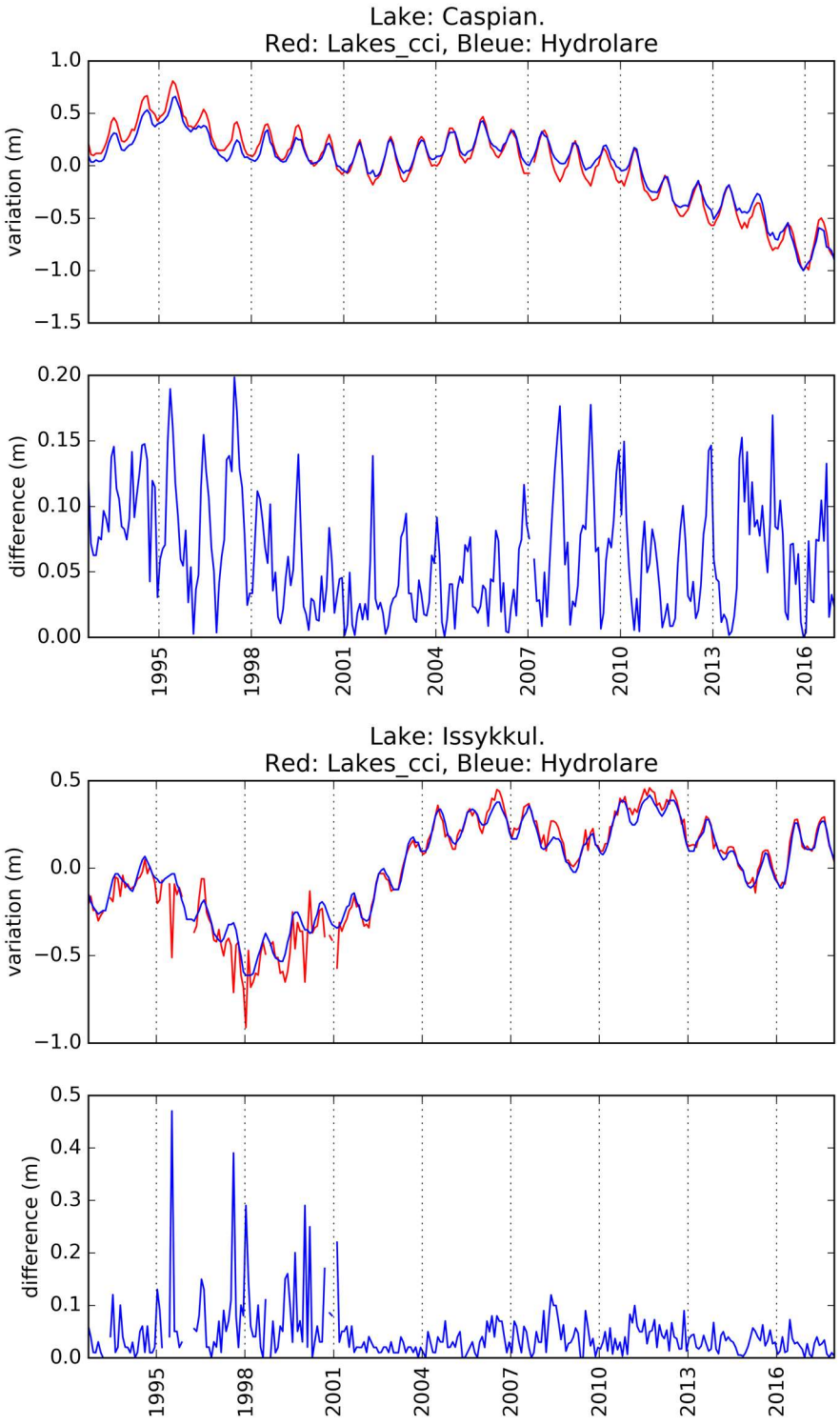
Vermote, E. F., Roger, J. C., & Ray, J. P. (2015). *MODIS Surface Reflectance User's Guide Collection 6.* Maryland: MODIS Land Surface Reflectance Science Computing Facility.

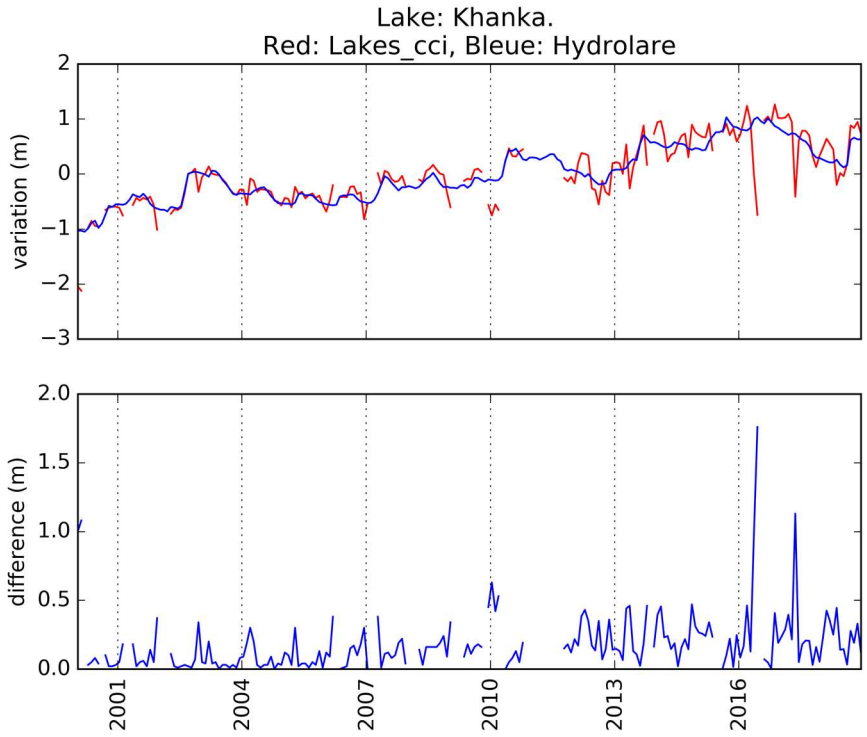
Zhang, Y., Shi, K., Liu, X., Zhou, Y., & Qin, B. (2014). Lake topography and wind waves determining seasonal-spatial dynamics of total suspended matter in turbid Lake Taihu, China: assessment using long-term high-resolution MERIS data. PLoS One, 9

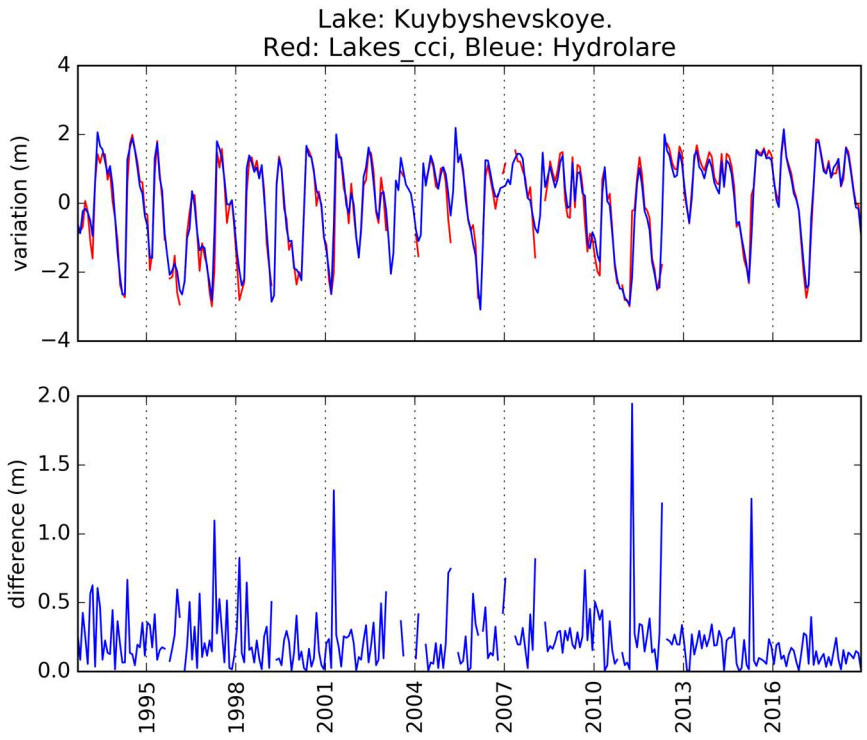
Appendix A. Hydrolare Comparison

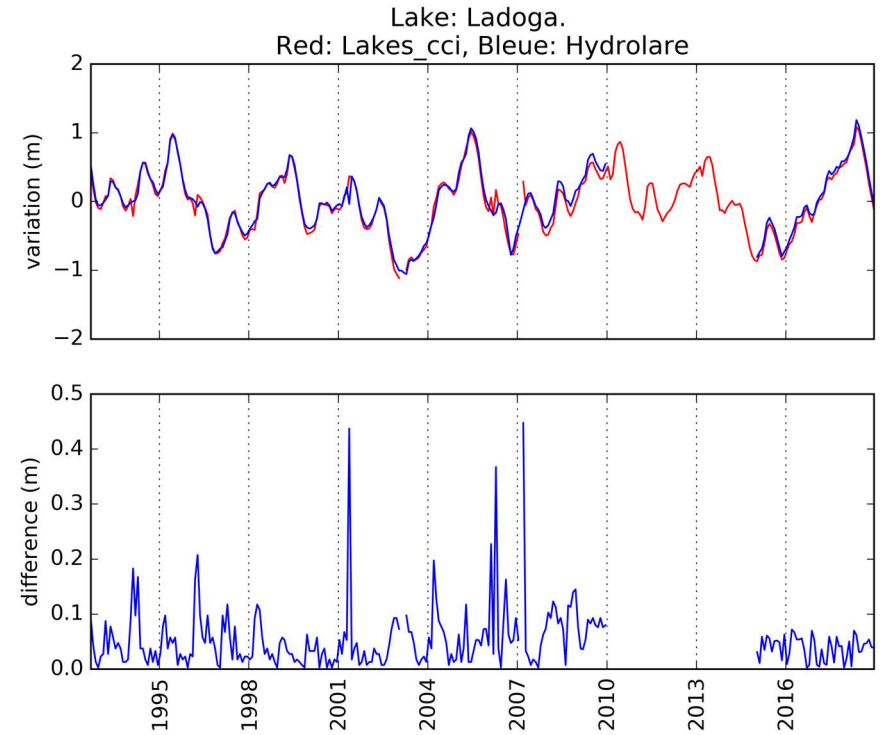


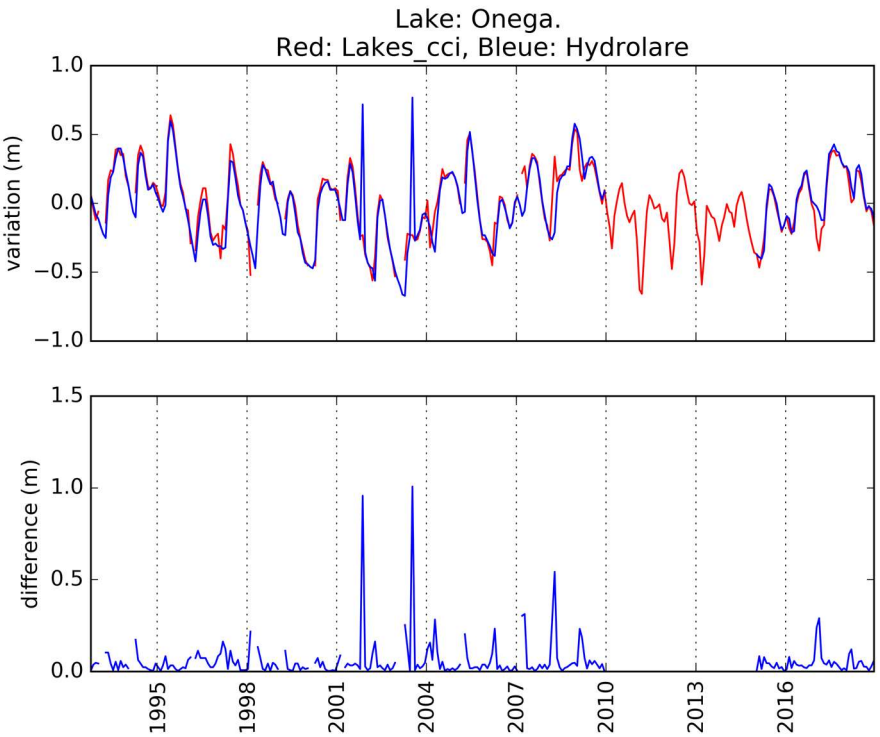


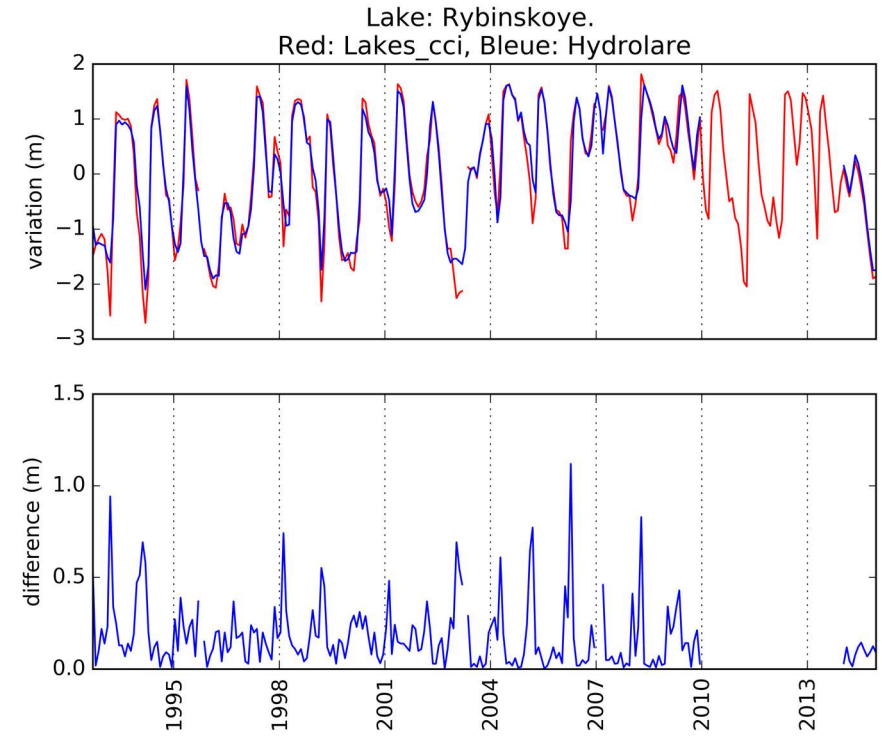

Lake: Caspian.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
1.0
0.5
0.0
−0.5
−1.0
−1.5
difference (m)
0.20
0.15
0.10
0.05
0.00
1995
1998
2001
2004
2007
2010
2013
2016
Lake: Issykkul.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
0.5
0.0
−0.5
−1.0
difference (m)
0.5
0.4
0.3
0.2
0.1
0.0
1995
1998
2001
2004
2007
2010
2013
2016

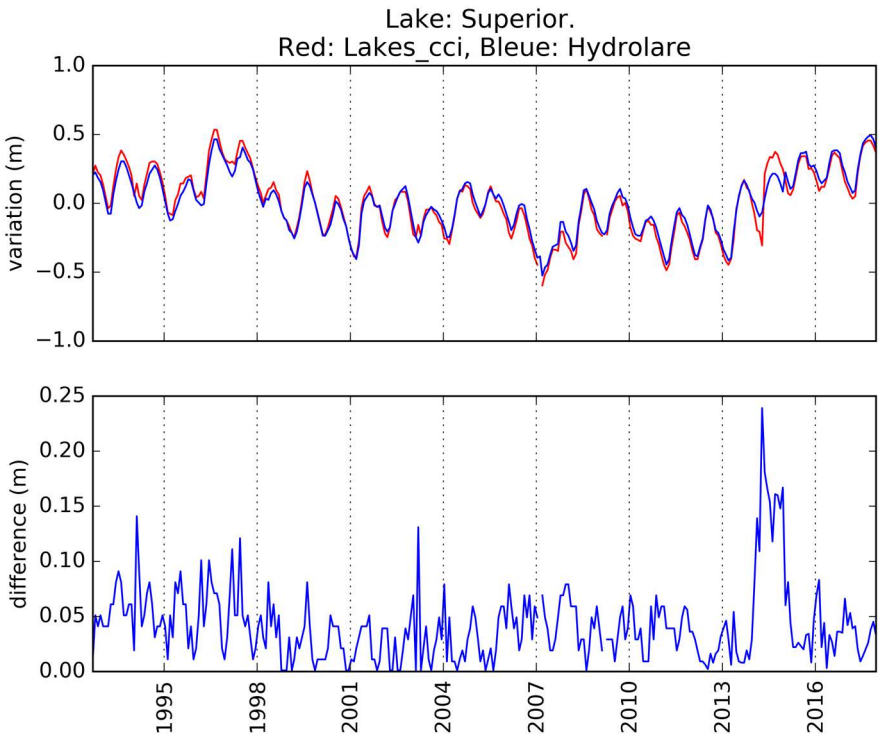

Lake: Khanka.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
2
1
0
−1
−2
−3
difference (m)
2.0
1.5
1.0
0.5
0.0
2001
2004
2007
2010
2013
2016


Lake: Kuybyshevskoye.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
4
2
0
−2
−4
difference (m)
2.0
1.5
1.0
0.5
0.0
1995
1998
2001
2004
2007
2010
2013
2016

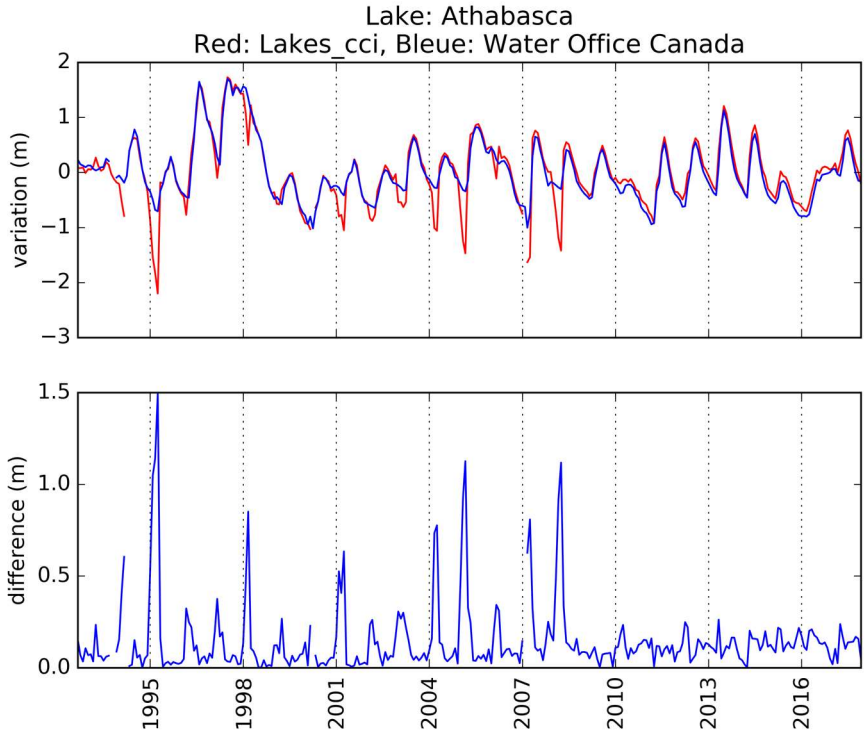

Lake: Ladoga.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
2
1
0
−1
−2
difference (m)
0.5
0.4
0.3
0.2
0.1
0.0
1995
1998
2001
2004
2007
2010
2013
2016

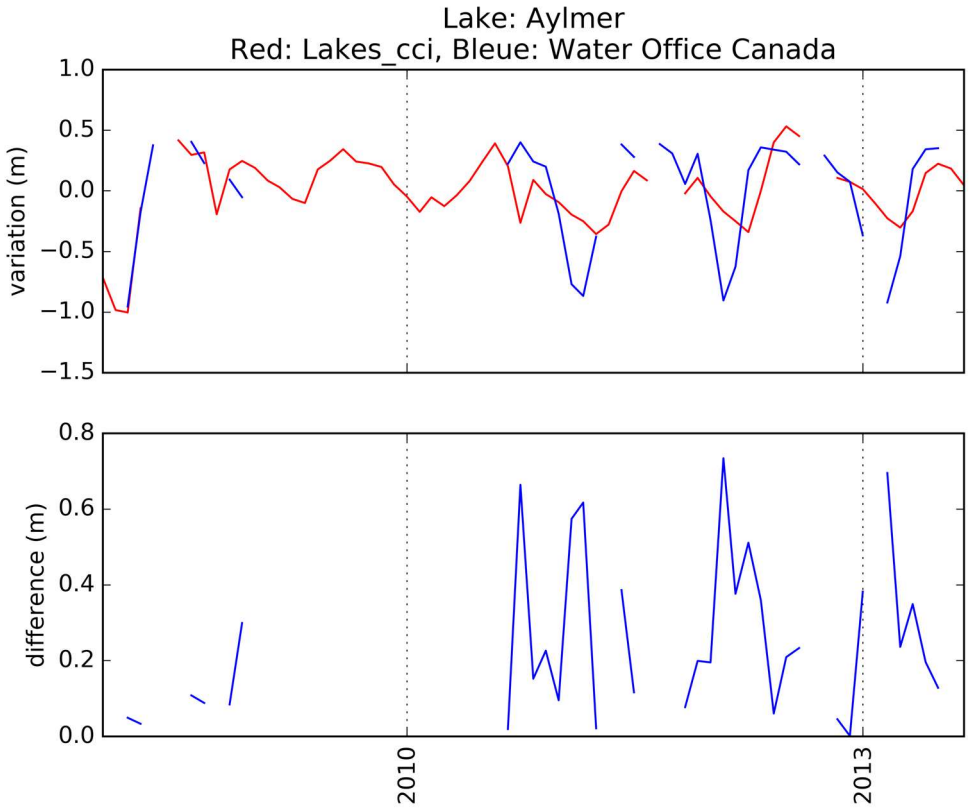

Lake: Onega.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
1.0
0.5
0.0
−0.5
−1.0
difference (m)
1.5
1.0
0.5
0.0
1995
1998
2001
2004
2007
2010
2013
2016

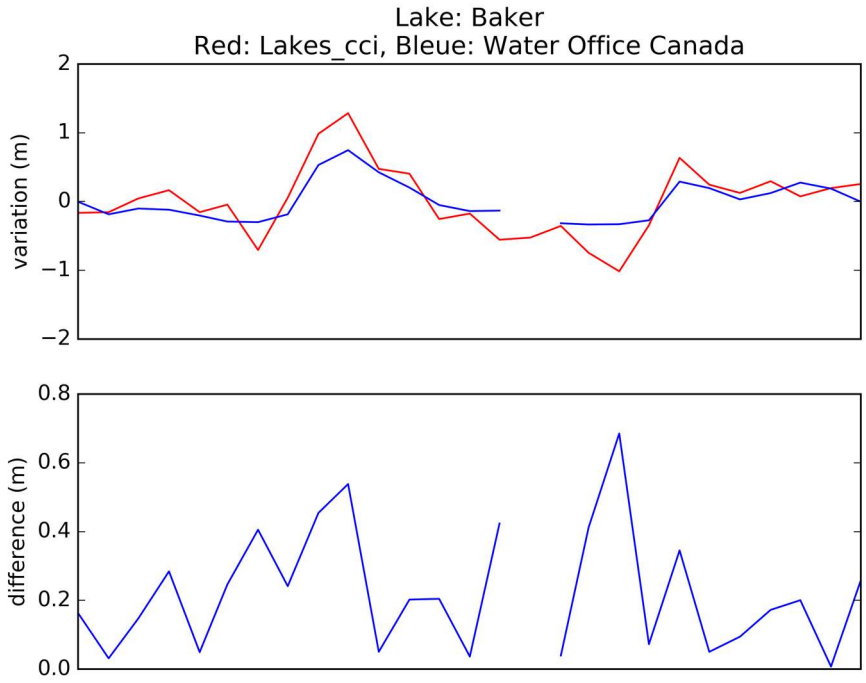

Lake: Rybinskoye.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
2
1
0
−1
−2
−3
difference (m)
1.5
1.0
0.5
0.0
1995
1998
2001
2004
2007
2010
2013

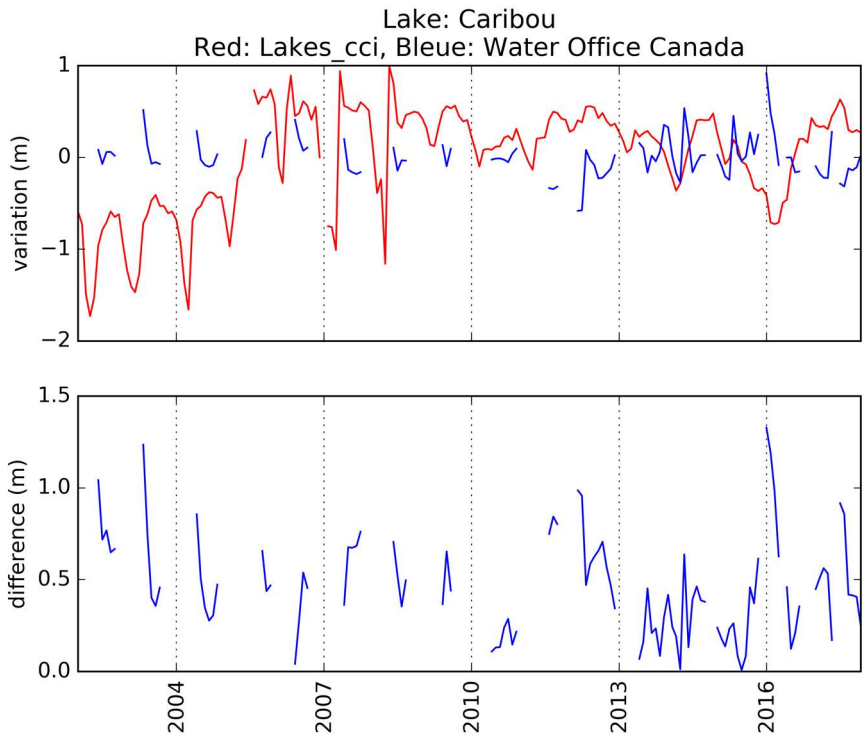

Lake: Superior.
Red: Lakes_cci, Bleue: Hydrolare
variation (m)
1.0
0.5
0.0
−0.5
−1.0
difference (m)
0.25
0.20
0.15
0.10
0.05
0.00
1995
1998
2001
2004
2007
2010
2013
2016

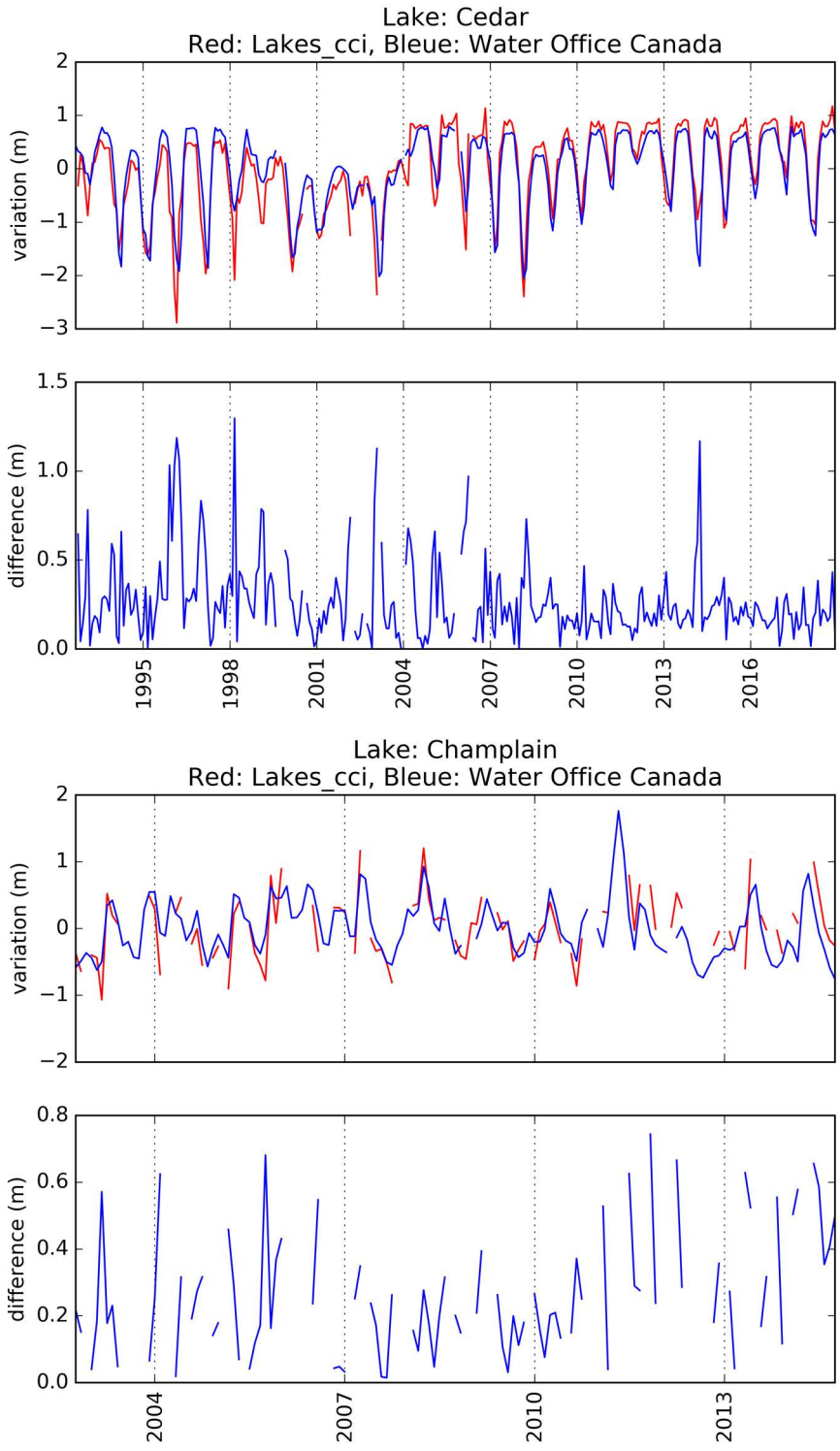
Appendix B. Water Office Canada Comparison

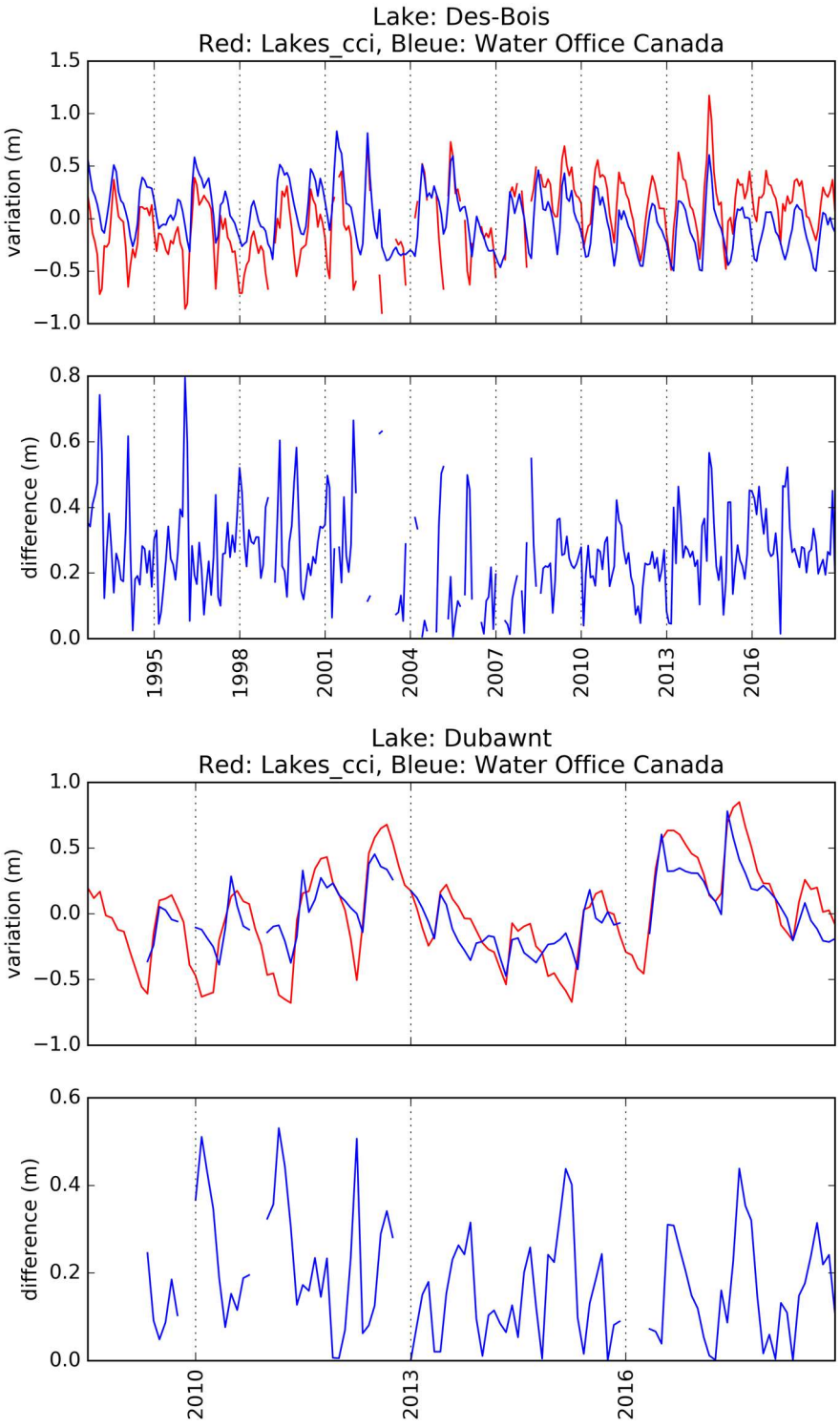


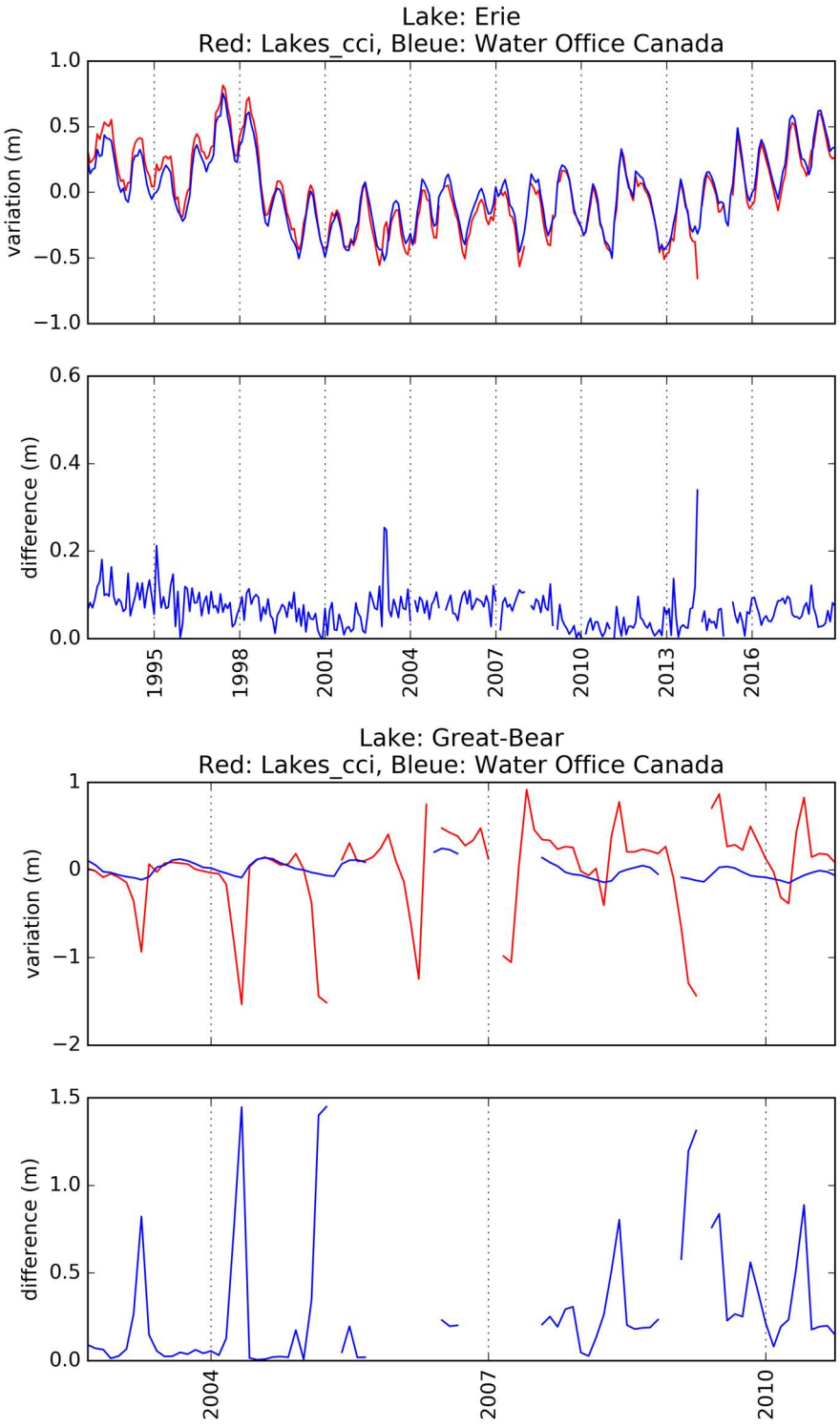


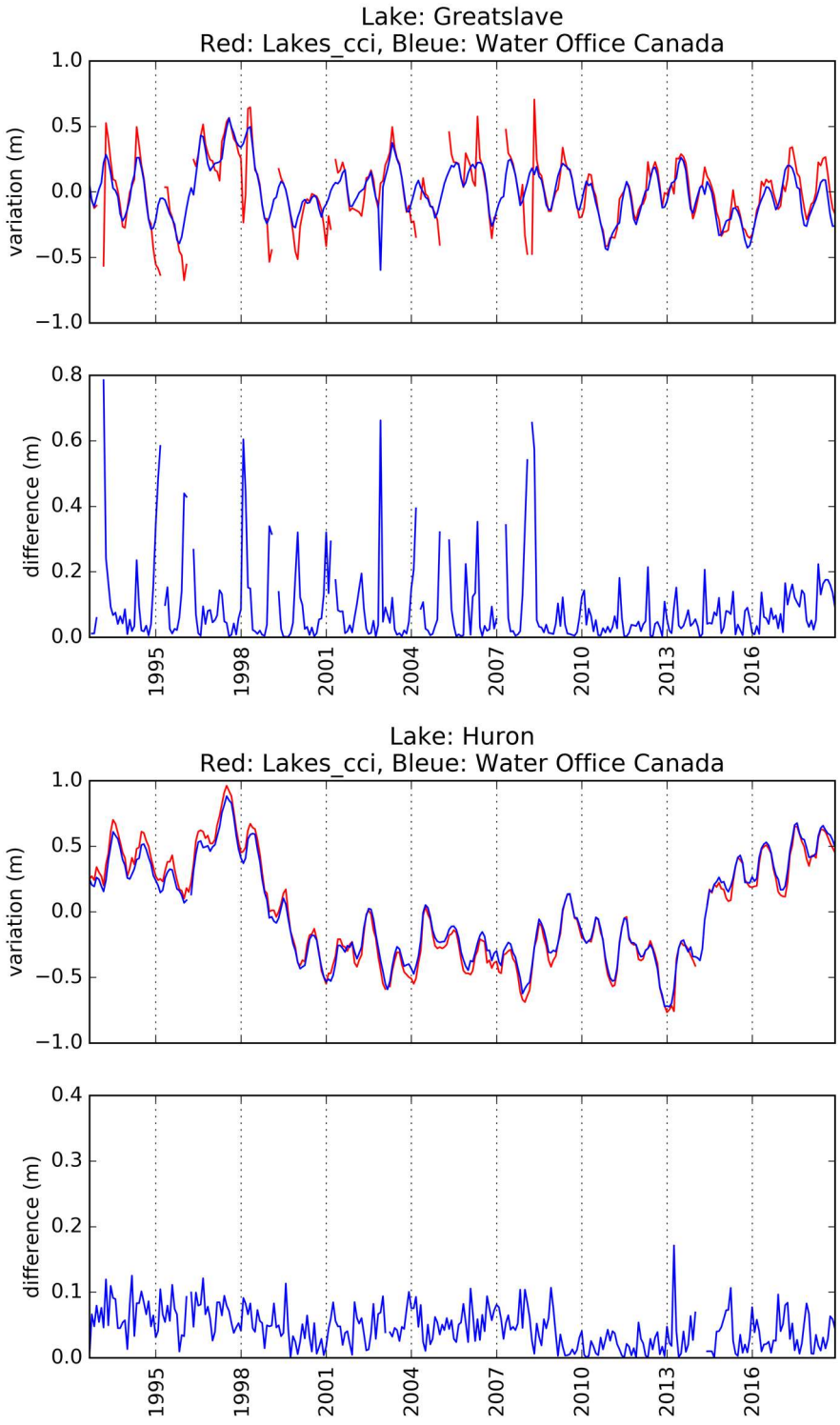

Lake: Baker
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
2
1
0
−1
−2
difference (m)
0.8
0.6
0.4
0.2
0.0

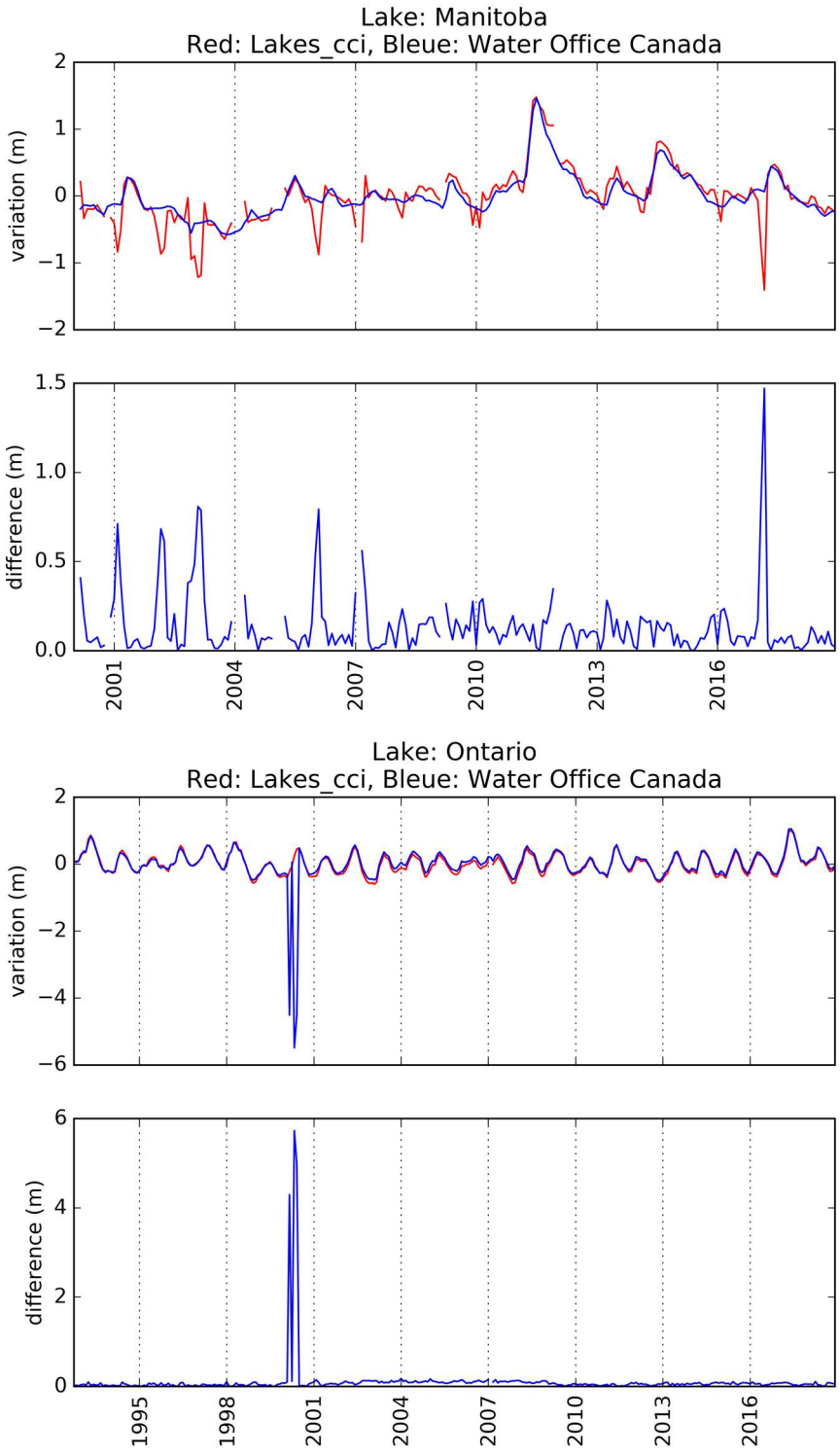

Lake: Caribou
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1
0
−1
−2
difference (m)
1.5
1.0
0.5
0.0
2004
2007
2010
2013
2016

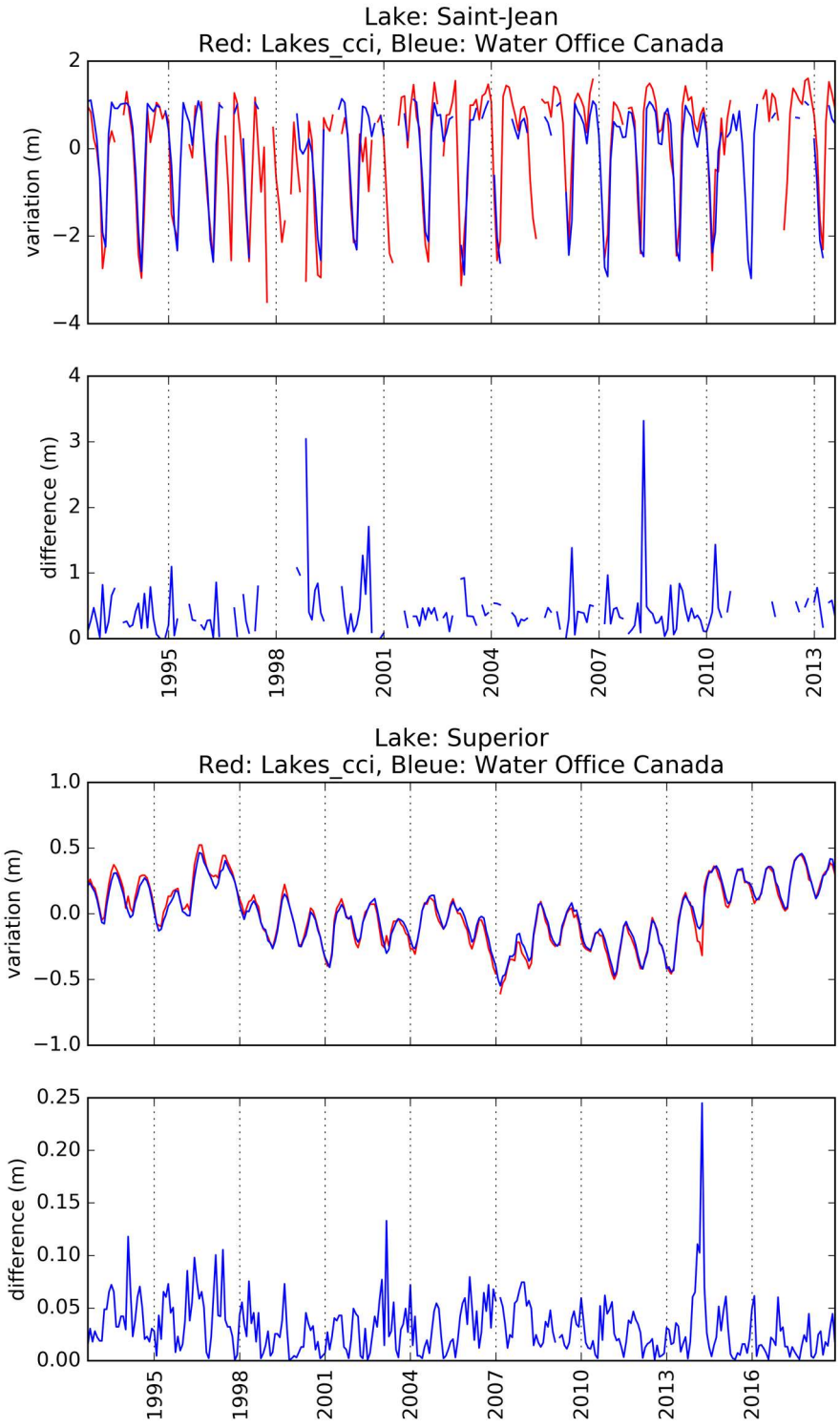

Lake: Cedar
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
2
1
0
−1
−2
−3
difference (m)
1.5
1.0
0.5
0.0
1995
1998
2001
2004
2007
2010
2013
2016
Lake: Champlain
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
2
1
0
−1
−2
difference (m)
0.8
0.6
0.4
0.2
0.0
2004
2007
2010
2013

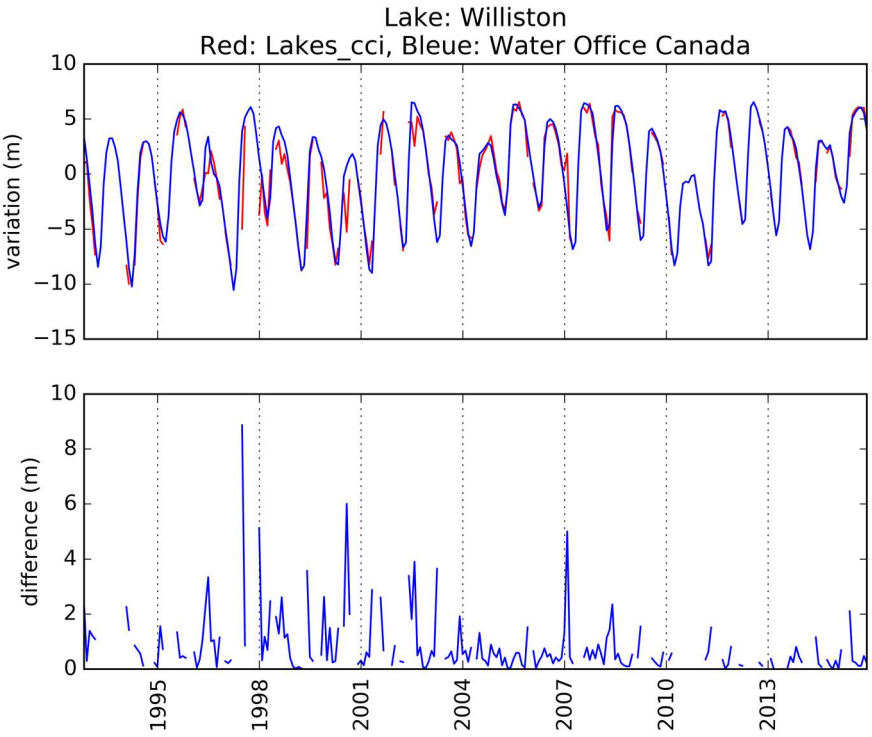

Lake: Des-Bois
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1.5
1.0
0.5
0.0
−0.5
−1.0
difference (m)
0.8
0.6
0.4
0.2
0.0
1995
1998
2001
2004
2007
2010
2013
2016
Lake: Dubawnt
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1.0
0.5
0.0
−0.5
−1.0
difference (m)
0.6
0.4
0.2
0.0
2010
2013
2016

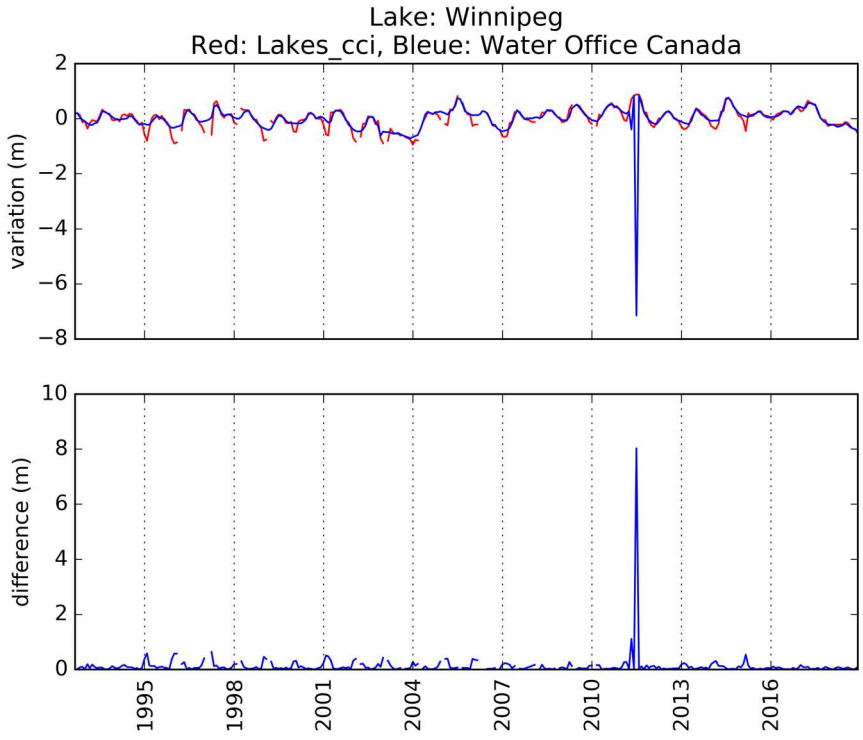

Lake: Erie
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1.0
0.5
0.0
−0.5
−1.0
difference (m)
0.6
0.4
0.2
0.0
1995
1998
2001
2004
2007
2010
2013
2016
Lake: Great-Bear
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1
0
−1
−2
difference (m)
1.5
1.0
0.5
0.0
2004
2007
2010

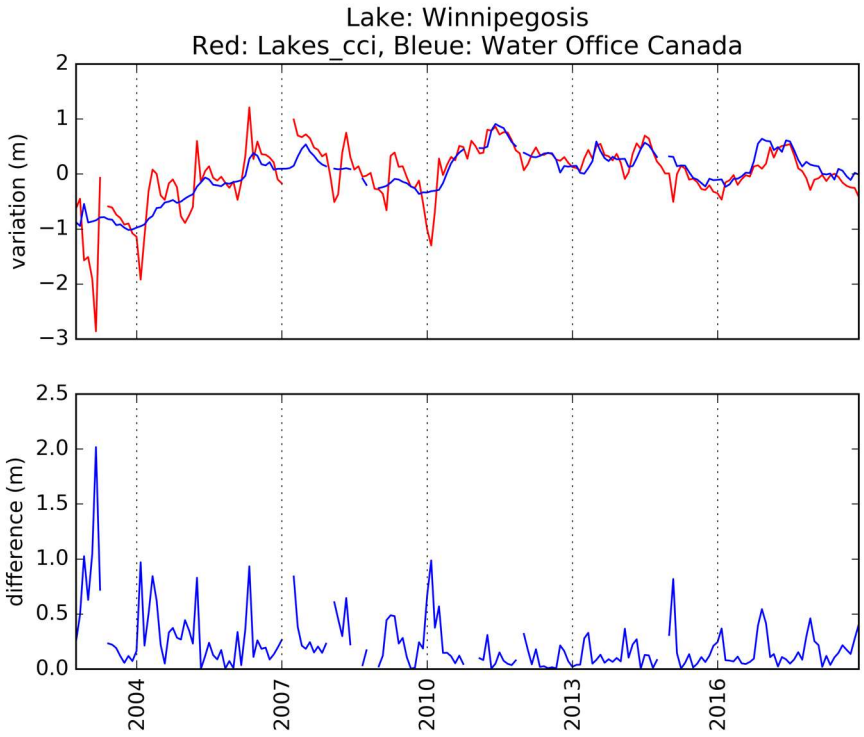

Lake: Greatslave
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1.0
0.5
0.0
−0.5
−1.0
difference (m)
0.8
0.6
0.4
0.2
0.0
1995
1998
2001
2004
2007
2010
2013
2016
Lake: Huron
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
1.0
0.5
0.0
−0.5
−1.0
difference (m)
0.4
0.3
0.2
0.1
0.0
1995
1998
2001
2004
2007
2010
2013
2016


Lake: Manitoba
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
difference (m)
Lake: Ontario
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
difference (m)


Lake: Saint-Jean
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
difference (m)
Lake: Superior
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
difference (m)


Lake: Williston
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
10
5
0
−5
−10
−15
difference (m)
10
8
6
4
2
0
1995
1998
2001
2004
2007
2010
2013


Lake: Winnipeg
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
2
0
−2
−4
−6
−8
difference (m)
10
8
6
4
2
0
1995
1998
2001
2004
2007
2010
2013
2016


Lake: Winnipegosis
Red: Lakes_cci, Bleue: Water Office Canada
variation (m)
2
1
0
−1
−2
−3
difference (m)
2.5
2.0
1.5
1.0
0.5
0.0
2004
2007
2010
2013
2016