Lakes_CCI+ - Phase 6 - CCN6 - Lake Storage Change D 2.1. Benchmark technical note

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

CHRONOLOGY ISSUES

Issue	Date	Object	Written by
1.0	21/09/23	Initial version	C.Fatras, I.Lucas
1.1	16/10/23	ESA comments included	C. Fatras, I.Lucas

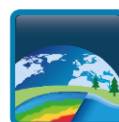|		
Checked by	Stefan Simis – PML J-F Cretaux - LEGOS	*Stefan Simis* *Jean-francois Cretaux*
Approved by	Alice Andral - CLS	*A. Andral*
Authorized by	Clément Albergel - ESA	*Clement Albergel*

DISTRIBUTION

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

LIST OF CONTENTS

LIST OF TABLES AND FIGURES

REFERENCE DOCUMENTS

Reference	Name	Access
[R1] CCI-Lakes2-0010-TN	Lake Storage Change Option. State of the art / User requirement	https://climate.esa.int/en/projects/lakes/key-documents-lakes
[R2] CCI-LAKES-0024-ATBD	Algorithm Theoretical Basis Document	https://climate.esa.int/en/projects/lakes/key-documents-lakes/

1 Purpose of this document

The objective of this document is to benchmark the different methodologies to retrieve storage change of lakes and reservoirs (LSC) globally and with a temporal depth sufficient for climate studies and climate modelling. The different methodologies to retrieve Lake Storage Change have been described in [R1]. The following diagram recalls the different steps to estimate LSC.
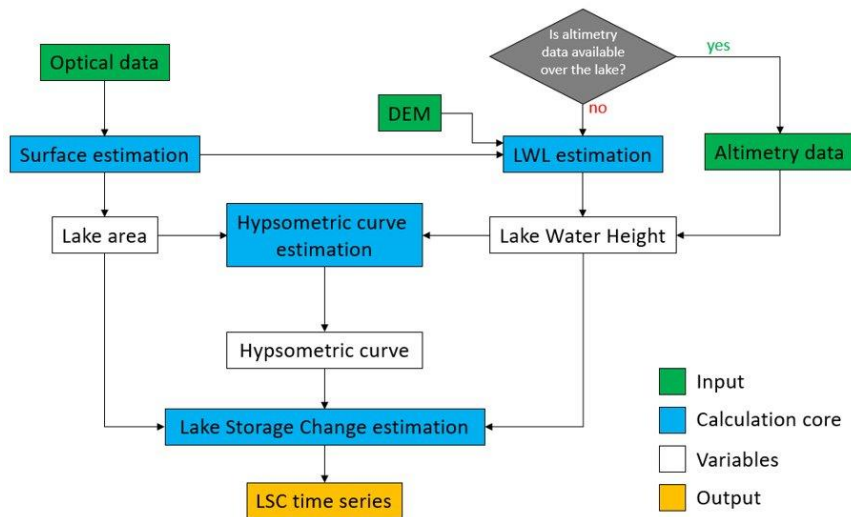
Figure 1 – Pipeline proposed in the frame of the ESA CCI LSC option

First, an evaluation of the different methodologies to retrieve the hypsometric curve mode will be presented, then we will focus on the different LSC estimation methods. The two last parts will benchmark the impact on LSC of the use of several available data sources for water level and surfaces. This analysis will allow to assess if and how the different dataset can be crossed and their impact on LSC errors.
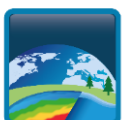
2 Validation dataset

2.1 Datasets

20 lakes and reservoirs have been selected to conduct the feasibility of LSC retrieval [R1]. Figure 2 gives the locations of the test sites and Table 1 the characteristics of these lakes and reservoirs, including all data available (from Earth observation and in-situ). The sources of information, the formats as well as the methods of acquisition of these in-situ data are site-specific and therefore will require some harmonization steps before any validation and quality assessment.
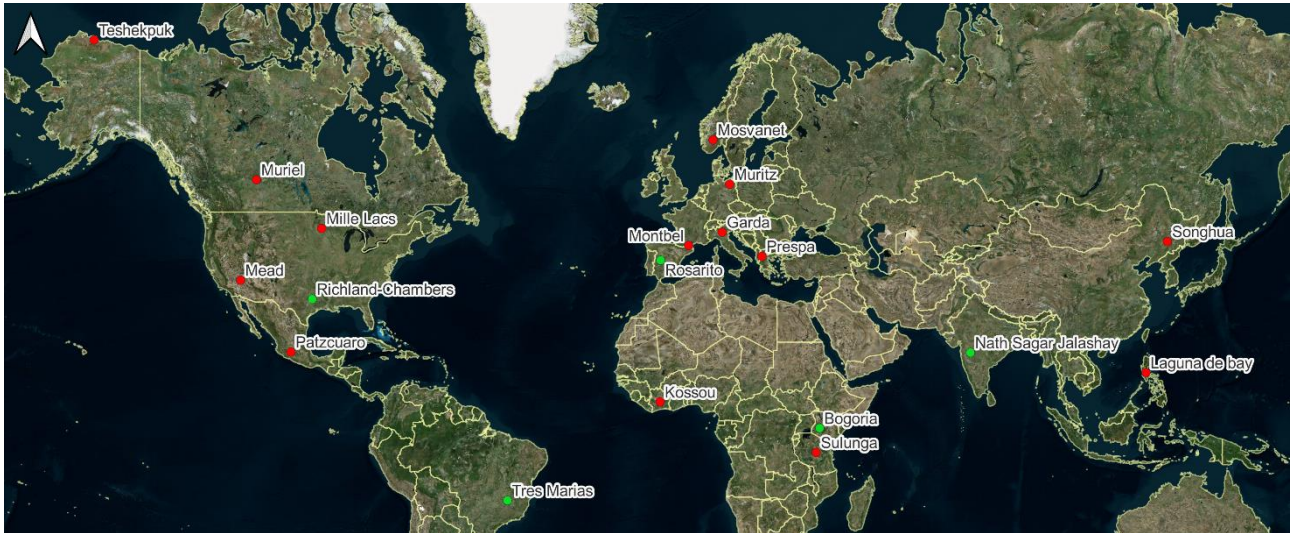
Figure 2 - Location of the twenty study lakes around the world

On all those lakes, only some will be considered in this benchmark (see the list in Table 1).

Table 1 – Lakes used in the benchmarks, In-situ data sources and descriptions (A: area, H: height, V: volume).

Name	Country	Benchmark Hypsometry	LSC	Heights	Surface	Notes	In-situ data source
Tres Marias	Brazil	X	X	X	X	In situ data (A-H-V curves, mathematical modelization. Unknown Height referential.)	Sistema Nacional de Informação de Recursos Hídricos (Brazil)
Muriel	Canada			X	X	No in-situ, no altimetry, used as an illustration	-
Nath Sagar Jalashay	India	X	X	X	X	In situ data (A-H-V, Digitalized data, from spatial determination. Unknown height referential.)	Remote Sensing Directorate, Central Water Commission (India)
Garda	Italy			X		A-H relation from algorithm on lake bathymetry. Unknown height referential.	-
Kossou	Ivory Coast		X	X		No in-situ	-
Bogoria	Kenya	X	X	X		A-H relation from algorithm on lake bathymetry. Unknown height referential.	Hickley et al., 2003
Rosarito	Spain		X	X	X	Volume time series. Unknown height referential	Ministerio para la transición ecológica y el reto demográfico (Spain)
Sulunga	Tanzania			X		Used to retrieve heights from DEM	-
Mead	USA	X	X	X		Analysis of different models from DAHITI platform	-
Mille Lacs	USA		X			Analysis of an unvarying lake	-
Richland-Chambers	USA	X	X		X	In situ data (A-H-V)	Texas Water Development Board (TWBD, USA)
Songhua	China			X		Analysis of height time series	-

In situ availability

2.2 Metrics

In order to compare methods and data, we will use metrics as indicators. Materialized by the formulas as below, metrics permit to quantify the error in a method or quantify the error according to in-situ data. In this benchmark, we will use metrics to compare models but also to compare results in relation with some in-situ data presented earlier.

Table 2. List of metrics used in the benchmark

Caption	Formula	Description
Equation 1- RMSE formula where ŷ is the estimated data an y the raw data, n the length of the times series.	$$RMSE = \sqrt{\frac{\sum_{i=0}^{n}(\hat{y}_i - y^{ref}_i)^2}{n}}$$	Root-mean-square error: measure of the differences between an estimation and an observation. Can assess a prediction model, or difference between in-situ values and estimated values.
Equation 2- Mean absolute difference formula where ŷ is the estimated data, n the length of the times series	$$mean\ offset = \frac{\sum_{i=0}^{n}(y^{ref}_i - \hat{y}_i)^2}{n}$$	Mean absolute difference or mean offset; Gives an idea of the delta between in-situ data and estimated data (in the given unit).
Equation 3- Relative Error formula	$$relative\ error = \frac{RMSE}{Scale} \times 100$$	Error in percentage; Facilitating comparisons.
Equation 4- Cosine Similarity Index formula where A and B, two time series to compare	$$CSI = \frac{A \cdot B}{\parallel A \parallel\ \parallel B \parallel}$$	Measuring similarity of two vector, the CSI permits to compare our two times series and particularly the correlation of their trends. The closer the index is to 1, the closer the time series are. The implementation is made from this formula:

3 Benchmark of hypsometry curve estimation

The common method to determine lake volume variations is to base the calculation on the hypsometric curve of the lake. This curve is defined by height and surface couples. To get the volume variation between two states of the lake whatever the height values, it is necessary to fit the curve with the adequate model.

In this section, we will first analyse which model gives the most precise curve between the power law, the polynomial, and models from geomorphic processes as the modified Strahler approach [R1], then study which computing method fits best to the model, the Gauss-Helmert compensation and /or the classical parametrical least square approach. Finally, we will explore the different methods to remove outliers when estimating the hypsometric curve (RANSAC algorithm, supervision of data sample and from water mask analysis)).

Four test lakes with in-situ hypsometric curves (either direct A-H-V relations or through provided bathymetry) are analysed: Richland-Chambers (USA), Bogoria (Kenya), Tres Marias (Brazil) and Nath Sagar Jayakwadi (India).

3.1 Curve model

The estimation of a hypsometric curve enables us to estimate lake storage changes. Whatever model is chosen, the result will approximate the actual topography. Although not considered as a test lake, the Garda Lake in Italy clearly illustrates this point by showing the very precise hypsometric curve deduced from the bathymetry made by the Italian Navy and the curve produced with a mathematical estimation (order 2), see Figure 3.
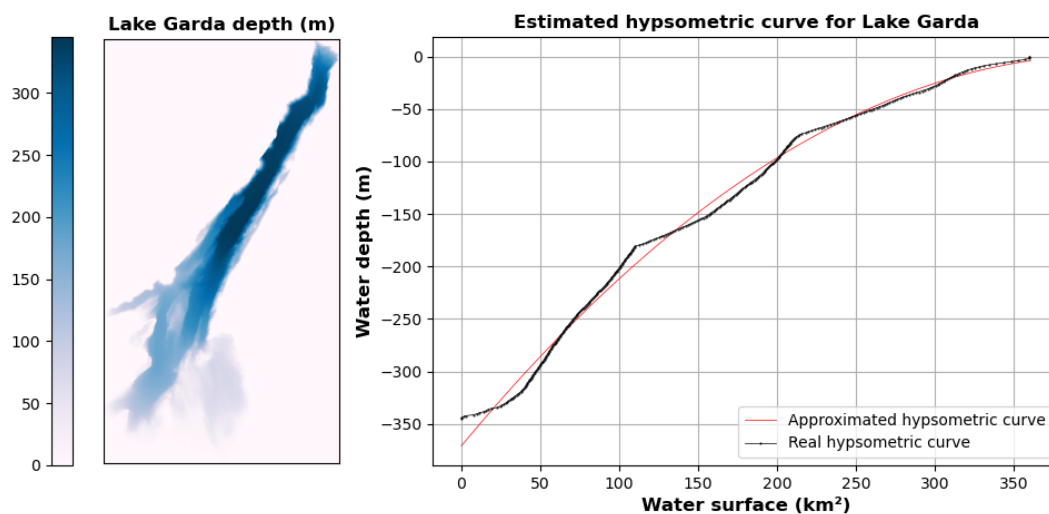
Figure 3- (Left) Lake Garda depth (bathymetry) from the Italian Navy. (Right) In-situ and approximated hypsometric curves

As we can see above, even if we approximate the curve with a mathematical model (in red), it will always be difficult to represent ground unevenness. Thus, we are only talking in terms of estimation of hypsometric curves in this benchmark because the reality is too complex to be modelled mathematically easily with high precision.

In the state of the art [R1], multiple possible models were presented: the modified Strahler approach, polynomial model and power law.
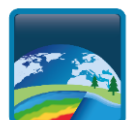
The modified Strahler approach, as used in the DAHITI platform (Schwatke & al., 2020).

This is a six parameters model which can be hard to initialise. Details of the methodology is given in [R1]. It considers the area and height scales of a lake to model the lake morphology as accurately as possible. On Figure 4, the hypsometric curves computed by the DAHITI team are plotted with different models for two study lakes (Mead and Richland Chambers in the USA) and their corresponding RMS. In the validity area of the data, polynomial and linear (polynomial of order 1) models can tend to give better results (smaller RMS). Whether we have more information on the lake characteristics (depth, in-situ data…), this approach could be interesting to use. In most cases, data are valid only in a certain window, thus, a polynomial could be enough.
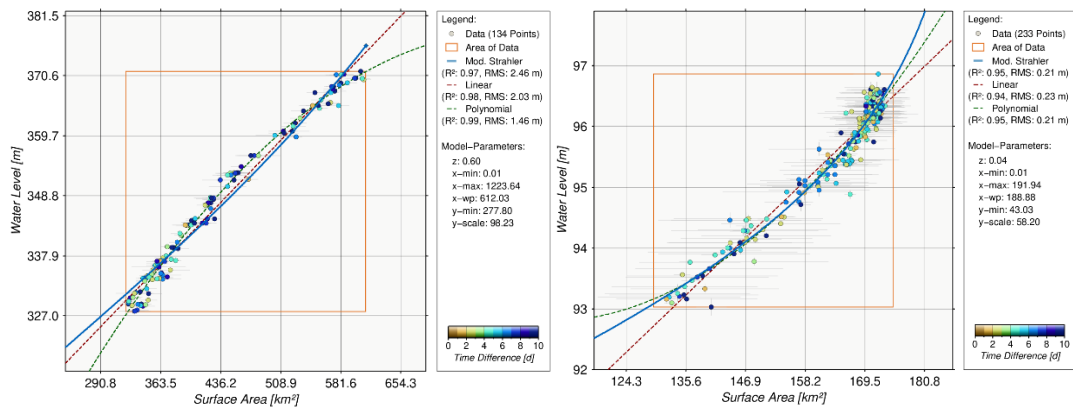
Figure 4 - Hypsometry on Mead Lake (left) and Richland Chambers Lake (right) with Modified Strahler model (figure from DAHITI website : https://dahiti.dgfi.tum.de/en/)

Polynomial model

It is the most common approach in literature like in Hydroweb and Lakes_cci project (Cretaux et al., 2016). Unlike the modified Strahler approach, it does not require any other *a priori* data than the heights and surfaces to be processed. On the different test lakes, when analysing the RMSE values, one would tend to keep the polynomials of order 1, 2 or 3 (see Table 3).

Power law

It is the easiest approach to initiate and run. Regarding all test lakes, it works only for the Tres Marias Lake in Brazil (see Figure 5) and the Bogoria Lake in Kenya. For other lakes, we encountered numerical instability, difficulty in parameters' initialization and poor fitting. Therefore, this model lacks genericity. As we can see on Figure 5, when a result is given, it is quite like a polynomial, which is more robust and quicker to initialize for our use.
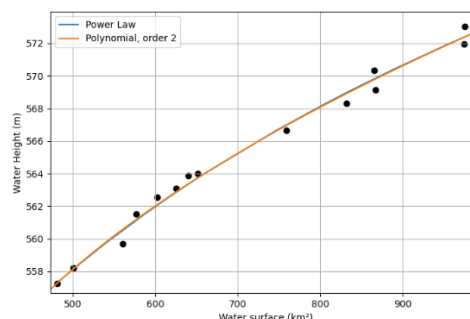
Figure 5 -Hypsometric curve from power law (blue) and order 2 polynomials (orange) on lake Tres Marias
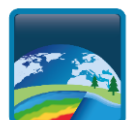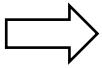
> **The most adequate model in terms of robustness, complexity, and result in most of the cases is the polynomial approach. The estimated hypsometric curve will be only available on the validity domain: where there are height /area observation couples.**

3.2 Computing method

To fit a model from raw data, two approaches are possible: the parametric estimation, also known as a least squares method and the Gauss-Helmert method (see [R1] for more details). For the parametric estimation, regression robustness is increased using the Huber loss function during iterations. For Gauss-Helmert compensation, as a total least squares compensation, both height and areas are considered in hypsometric curve estimation.

The calculation of the hypsometric curve of each test lake gives quite similar results with the parametric and Gauss-Helmert compensation when looking at RMSE and Mean offset (Table 3).

Table 3 - Statistics on the estimation of the hypsometric curve according to the calculation method

|		Richland-Chambers		Bogoria		Tres Marias		Nath-Sagar		
	Order	RMSE	Mean offset (km²)	RMSE	Mean offset (km²)	RMSE	Mean offset (km²)	RMSE	Mean offset (km²)	MEAN
Parametric	1	2,004	3,919	0,121	**0,216**	22,241	56,145	8,106	**15,768**	**13,565**
	2	1,171	3,449	0,1175	0,253	16,438	59,263	5,194	17,986	**12,984**
	3	**0,859**	**3,397**	**0,116**	0,282	14,519	58,633	2,115	19,031	**12,369**
Gauss-Helmert	1	2,002	3,895	0,121	0,217	21,991	56,46	7,666	16,357	13,589
	2	1,399	3,742	0,1179	0,236	16,511	59,777	5,186	18,077	13,131
	3	1,526	4,609	**0,116**	0,282	**14,47**	58,825	**2,082**	19,27	12,648

For this benchmark, we chose the parametric method. The results are slightly better with this method, but this choice is rather because the Gauss-Helmert method is optimal when using homogeneous data pairs (height, surface) from the same source or of equivalent accuracy. However, altimetric data is more accurate than surface data. Therefore, including these surfaces in the estimate may distort the result, whereas the parametric method ensures an estimate that matches the accuracy of the input height data. This Gauss-Helmert method will therefore be the best one to use when dealing with heights and areas from homogeneous data source which might be the case with future data from the SWOT mission.
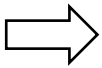
> **To be as robust as possible, the use of parametric compensation is preferred for the time being.**

Moreover, the hypsometric curve must be a steady or increasing curve. Thus, a condition on the curve growth shall be applied by testing if the model derivative has a decreasing slope.

3.3 Outlier Management

Water surface and water level time series may contain outliers (cloud cover, lake features, satellite optical defect, etc.). These errors can lead to biases on the hypsometric curves and so on the lake storage change determination. Therefore, the outlier management consists in choosing, during the LWE processing, the surfaces that gives best results (no clouds…). The selection of data should be made according to height time series to ensure the good repartition of surfaces along the curve to estimate. This is also the methodology applied for the Lakes_cci LWE (see ATBD [R2]).

Here, three approaches are compared to eliminate outliers: supervision of data sample, elimination of corrupted surface data by water mask analysis, and RANSAC algorithm.
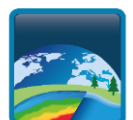
The first solution, supervision of data sample, is done manually before any processing -as in [R2]. This leads to a constrained hypsometric curve with good scores, see example in Figure 6. The main flaw of this method is the necessity of manual supervision that impacts any automatic processing.
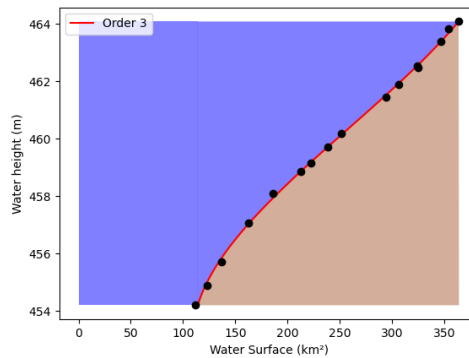
Figure 6- Curve fitting for sampled Nath-Sagar data

Another solution is to extract corrupted water mask before estimating the hypsometric curve. One idea is to estimate the proportion of erroneous/missing pixels from the permanent areas of the considered water body, and then compare it with the water body generated for one month. This can be done by using Pekel occurrences (Pekel & al., 2016) with a potential cloud mask and removing all data in the permanent water mask where the percentage of clouds/no data is high.

Another solution is to extract corrupted water mask before estimating the hypsometric curve. One idea is to estimate the proportion of erroneous/missing pixels from the permanent areas of the considered water body, and then to compare it with the water body generated for one month.

For instance, on the Tres Marias Lake (see Figure 7), we estimated a "permanent" water mask from the Pekel GSWO dataset (Pekel et al., 2016) by selecting water pixels with occurrences of at least 95%. Then, we compare this "permanent" mask with an actual water area from the monthly GSW dataset to estimate a percentage of wrong detection (no water where it should be). In the case shown below, the yellow area is correctly observed but is definitively insufficient compared to the whole permanent area displayed in red. That indicates a high proportion of missing data, so this water mask is discarded. This way, most outliers are removed from monthly time series (see Figure 7 below). Pekel occurrence maps can present errors (see section 6.3.2), and the choice of 95% occurrence for the permanent water mask can be irrelevant depending on the lake. This is still a way to manage to standardise this process, that could further be refined.
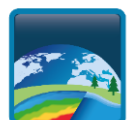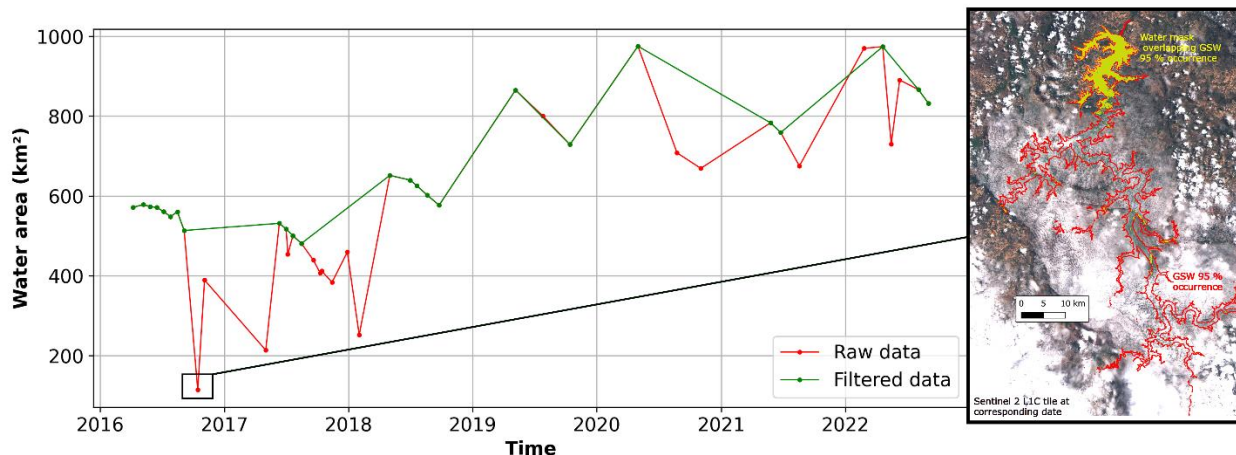
Figure 7 - Data filtering from GSWO dataset on Tres Marias Lake

Finally, the RANSAC method is a quasi-fully automated option, particularly relevant in case of datasets with strong outliers. For instance, on the Tres Marias Lake, some outliers shift the model by 35km² in area. It increases the error by 5% compared to in-situ data. The use of RANSAC removes all outliers and therefore improves the model significatively (see Figure 8),
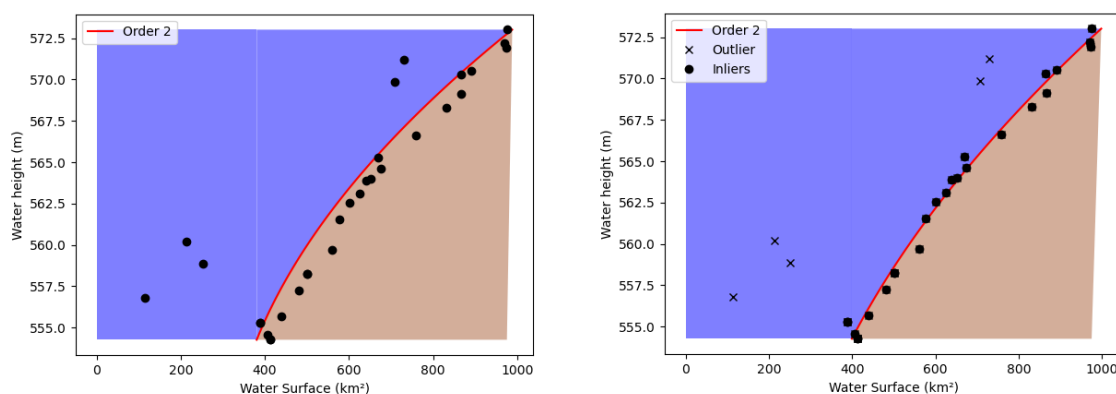
Figure 8- Fitting on the Tres Marias Lake without (left) and with RANSAC (right)

Yet, this RANSAC method does not automatically yield to the same result when random samples of the dataset are used to fit the model. Indeed, in some cases, it stops because the maximum number of iterations is reached without minimizing enough the error. This phenomenon depends on the density and the dispersion of the dataset. For example, Figure 9 shows two results obtained by running the same process several times on Bogoria lake which has dispersed data (presence of outliers): its gives two different hypsometric curves.
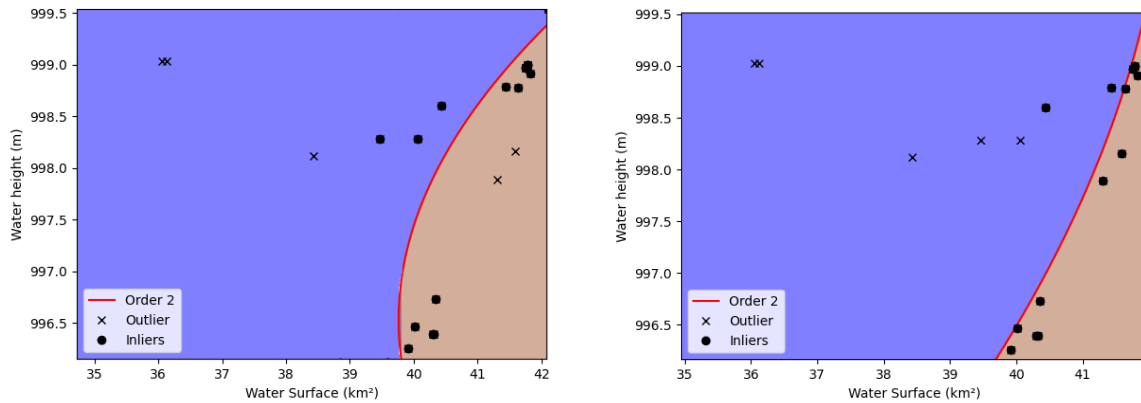
Figure 9- Two Fitting iterations with RANSAC on the Bogoria Lake

Thus, letting RANSAC choose a model of lake hypsometric curve can generate multiple bias or errors on the final volume variations when data sample contains scattered outliers. Adapting RANSAC parameters for data processing can help to minimize this problem, but adding parameters is not what is wanted for a generic and quasi-automated process.
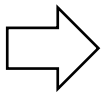
> **The estimation of the curve with outliers discarded can be partly automatized, but still needs human validation.**
>
> **A quasi-automatic removal of outliers for LSC estimation is possible by combining the automated selection of outliers (on water masks) and the application of RANSAC algorithm.**

4 Benchmark of Volume variation estimation

For lake storage change estimation, we need first to assess if the lake is stable, which is often linked to an unvarying surface, or varies in height and surface with time. Therefore, it involves two different processes for volume variation estimation that are described and benchmarked in the two following sub-sections. Note that the classification of lakes regarding the surface extent variation is specifically analysed in Section 6.1.

4.1 LSC calculation Methodology

The following table summarizes all the methods possible to calculate volume variations (see [R1] for more details).

Table 4 – Formulas for water volume variation estimation between two states

Heron's formula	$\Delta V = \frac{1}{3}\left(A_1 + A_2 + \sqrt{A_1 * A_2}\right) * \Delta H$
Mean area	$\Delta V = \frac{1}{2}(A_1 + A_2) * \Delta H$
Basic volume (For unvarying lake areas)	$\Delta V = A * \Delta H$ where A corresponds to the unvarying surface of a lake
Integration	$\Delta V = \int_{H_1}^{H_2} f(H)\, dH \quad or \quad \int_{A_1}^{A_2} f(A)\, dA$

LSC has been computed using these 4 methods on a sample of lakes having different surface variation rates. Figure 10 gives all the results. The first results of this comparison show that all methods give consistent LSC timeseries. Besides, curves based on Integration, Heron and Mean Area methods are often overlapping (Bogoria, Rosarito, Mead, Milles Lacs). Concerning the Basic Volume formula, we observe more discrepancies in the extrema and LSC timeseries, except on the Mille Lacs Lake that has little surface variation and for which all methods are equivalent. This latter result let us assume that with a low surface variation rate, the basic volume formula, which has the advantage of using only one area value, is sufficient. This assumption will be explored in the next subsection 4.2.
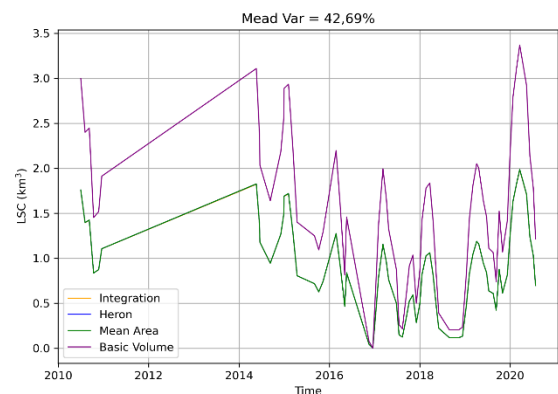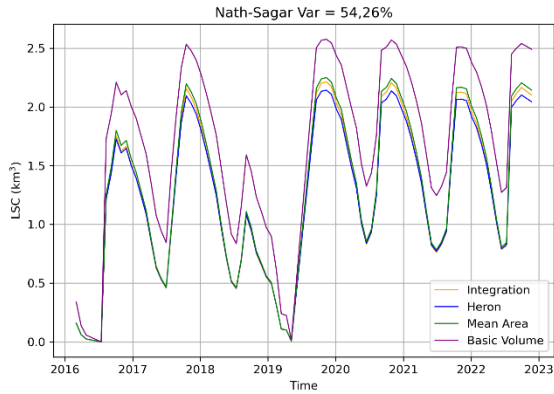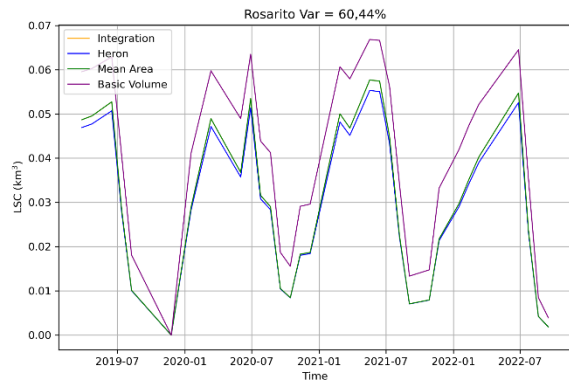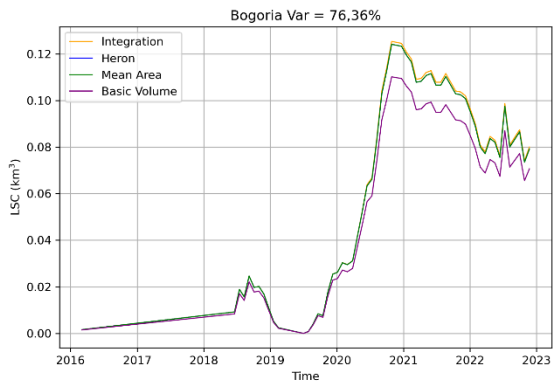
Figure 10 - LSC on 8 lakes with all methods

For medium to high surface varying lakes, further investigations are needed to choose the most adequate method.
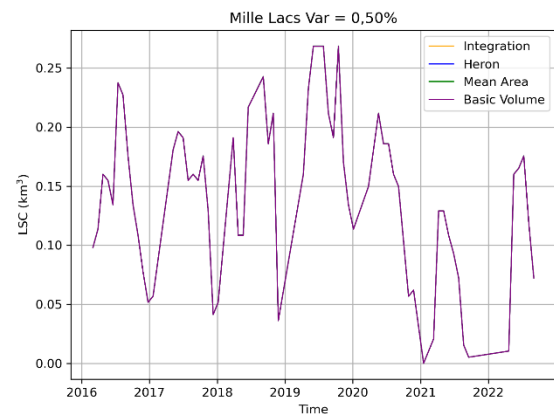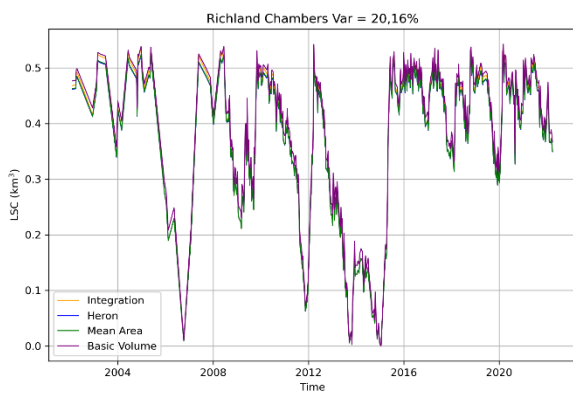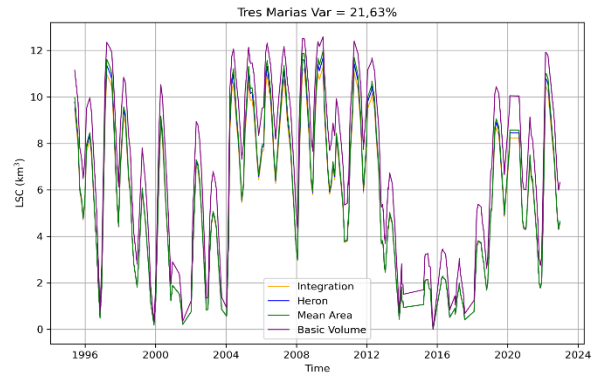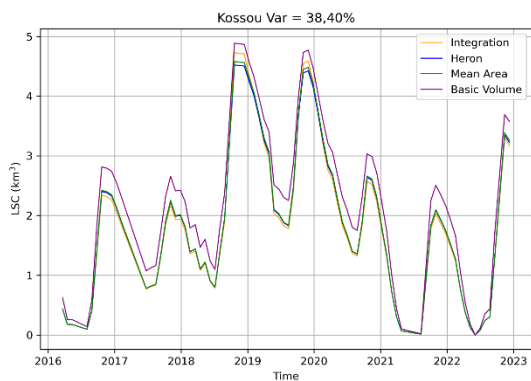
To do this, tests were made on Richland-Chambers and Tres-Marias lakes where in-situ heights, surfaces, and volume are available. For each lake, a hypsometric curve is calculated from in-situ data, then volume variations with the different methods. We also distinguish whether volume variations are computed between consecutive dates or in relation to the known minimum height. The final volume variations are compared with in-situ volume measurements. The result is summarised in the table below.

Table 5 - Statistics on LSC estimation depending on methods.

| | | Richland-Chambers LSC … | | Tres Marias LSC … | |
		…in relation to min known value	…between consecutive dates	…in relation to min known value	…between consecutive dates
Integration	Pearsons	1	1	1	1
	RMSE	0,0004	0	0,0026	0,0024
	Relative error (%)	0,0601	0,0382	0,0215	0,0199
	Mean Offset (km^3)	0,0003	0	*0,0024*	0,0017
Heron	Pearsons	1	1	1	1
	RMSE	0,0048	0,0001	0,0673	0,003
	Relative error (%)	0,8191	0,0966	0,5547	0,0251
	Mean Offset (km^3)	0,0038	0	0,0539	0,001
Mean Area	Pearsons	1	1	1	1
	RMSE	0,0042	0,0001	0,1749	0,0081
	Relative error (%)	0,7151	0,0968	1,4425	0,069
	Mean Offset (km^3)	0,0032	0	0,1306	0,0019
Basic Volume	Pearsons	1	1	1	1
	RMSE	0,0125	0,0004	1,0781	0,2338
	Relative error (%)	2,1065	0,3065	8,8921	1,982
	Mean Offset (km^3)	0,0113	0,0001	1,0275	0,1623

Results shows low relative error, particularly for Mean area, Heron and Integration methods for all lakes. However, the integration method has a better accuracy. For the basic volume formula, the chosen value for *A,* the unvarying surface, is picked for each lake in the HYDROLakes database (Messager & al., 2016). It gives poor statistics because surfaces variations are too important to be materialized by a single value and to give reliable outcomes.

> **Integration between consecutive dates gives best results to estimate lake storage change for varying lakes.**

4.2 Unvarying surface Lakes

The analysis will be based on the Mille Lacs Lake. From 2016 to 2018, its volume variation (0.75 km³) corresponds to only 0.14% of its global surface

On the Mille Lacs Lake, data are few, well sampled and not dispersed, which facilitates the computing of volume variations with the hypsometric curve. Note that, if the data height/surface couples scatter plots
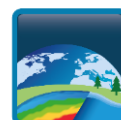
of unvarying lakes are really dispersed, height/surface couples scatter plots tend to look like a trendless shapeless point cloud which can be hard to model as a curve.

Figure 11 illustrates the weak variation of this lake. This kind of behaviour is characterized by a hypsometric curve, which can be compared to a vertical line. By zooming in on the abscissa, a polynomial curve of second order appears.
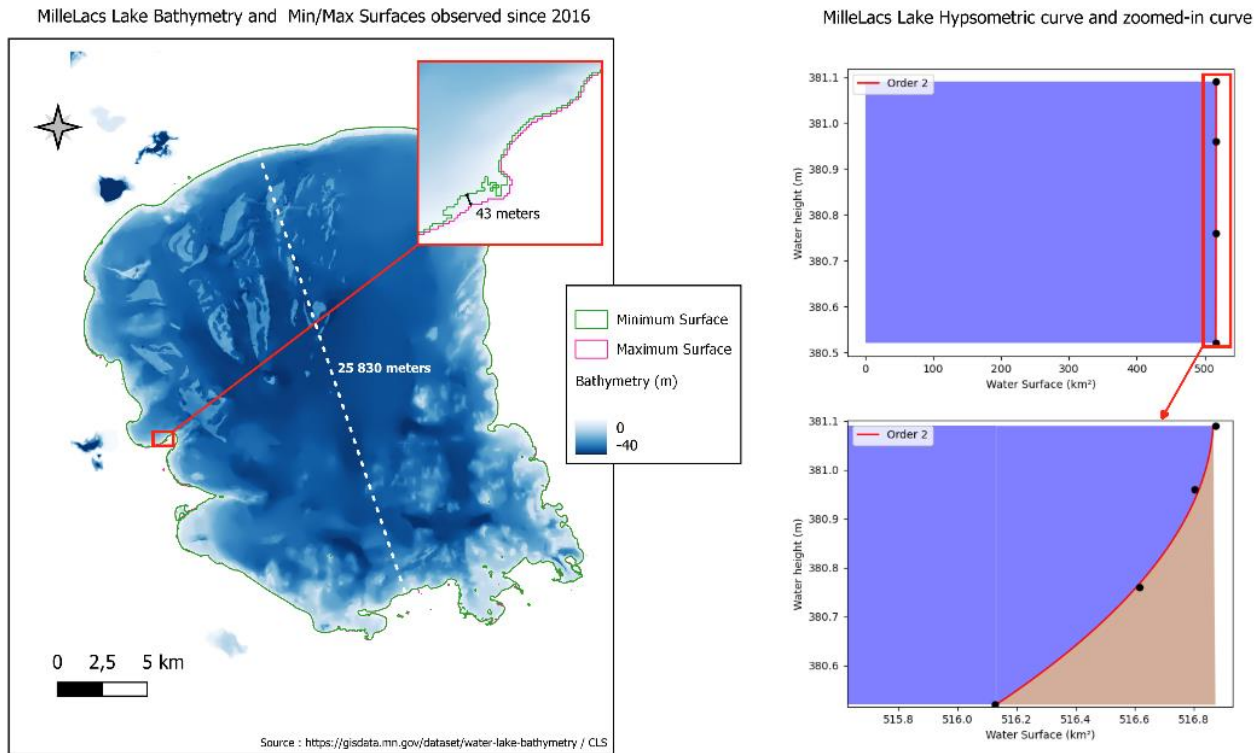
Figure 11 - Characteristics of a stable lake: Example of the Mille lacs Lake

For unvarying lakes, we compare two methods:

- LWL and water areas time series will be used to compute LSC time series with the integration of the hypsometric curve to estimate LSC (as described in the previous section)
- computation of LSC time series with the basic volume formula with LWL and a mean area derived from HydroLAKES .

The lake storage change is calculated in relation with the minimum known height.

Both methods give equivalent results (see Figure 12). The mean offset between the two methods is of 0.001 km³ which leads to a relative difference of 0.17% in relation to the mean maximum variation. With all these elements, it can be argued that for a lake with surface variations of less than 5%, the basic volume method is sufficient: less computing time, equivalent results, if not better (integration errors can introduce errors if the hypsometric curve doesn't fit well the data). Note that the question of how to define the lake surface variation rate limit is described in part 6.1.
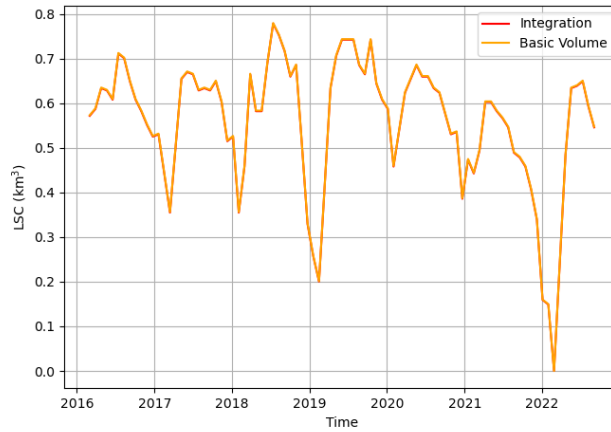
Figure 12- Overlapping curves of LSC estimation by integral and basic volume formula on Mille Lacs Lake
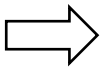
➡️ **When the lake is classified as stable (for example with surfaces varying by less than 5%), the basic volume method is preferred.**

5 Benchmark of height estimation methodologies

5.1 Height estimation from altimetry

Water levels are essential for estimating variations in lake water storage. In some global studies of lake storage change, the hypothesis is even made to have a constant surface area and therefore LSC is directly derived from water level (Cooley et al. 2021).

One common way to get lake water height is with satellite altimetry. This method allows the measurement of water levels in rivers, lakes, and floodplains (see ATBD [R2]). On this Lakes_CCI option, we will consider the historical LWL databases (HydroWEB, DAHITI and G-REALM) and analyse if these data sources are compatible or consistent with each other. To do so, we will first compare the LWL timeseries depending on the source of the data and then assess the impact of using one source compared to another on LSC estimation.

5.1.1 Assessment

There are 5 selected lakes Mead (USA), Kossou (Ivory Coast), Bogoria (Kenya), Songhua (China), and Garda (Italy) for which water level data are available in HydroWEB, DAHITI and G-REALM and 3 others with HydroWEB and DAHITI data: Tres Marias (Brazil), Rosarito (Spain) and Nath Sagar (India) (Table 6). Since volume in-situ data is available, a statistical assessment can also be done on these three lakes.

Table 6 - Data availability for studied lakes for height assessment, green means it is available, red means it is not.

	Height			In-Situ
	DAHITI	HydroWEB	G-REALM	Volume
Mead				Yearly
Kossou				
Bogoria				
Songhua				
Garda				
Tres Marias				
Rosarito				
Nath Sagar				

5.1.2 Comparison of available time series

For the 5 lakes with LWL available in HydroWeb DAHITI and G-REALM, the time series have a strong consistency together with the same tendencies (see Figure 13). However, on Bogoria Lake for example (Figure 14 e), there are offsets of 0.14m between DAHITI and HydroWEB, 0.14m also between DAHITI and G-REALM, and 0.30m between G-REALM and HydroWEB. This kind of offsets can be explained by different processes applied on the raw altimetry data and different reference geoids. Yet, these discrepancies have no significant impact on LSC estimation if we consider only one source of LWL for each lake (see part 4.1.3).
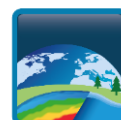
Therefore, and despite the similarities between the different database, we must stay careful when switching from one source to another and recommend using only one source of LWL for a given lake. This allows to consider more lakes for LSC estimation.

The case of Garda Lake (see Figure 13 d) shows that trends between the three data sources are the same but noisy. Nevertheless, the height-span of G-REALM (1.3m) is twice as big than the height-spans of DAHITI (0.6m) and HydroWEB (0.7m). This can be explained by the fact that heights are hard to extract from Garda Lake altimetry tracks: this lake is surrounded by a great diversity of surfaces and topography, which leads to heterogeneous backscattering. Moreover, Garda Lake does not change much in terms of height and surface compared to other water bodies, leading to scattered raw data difficult to sort out.
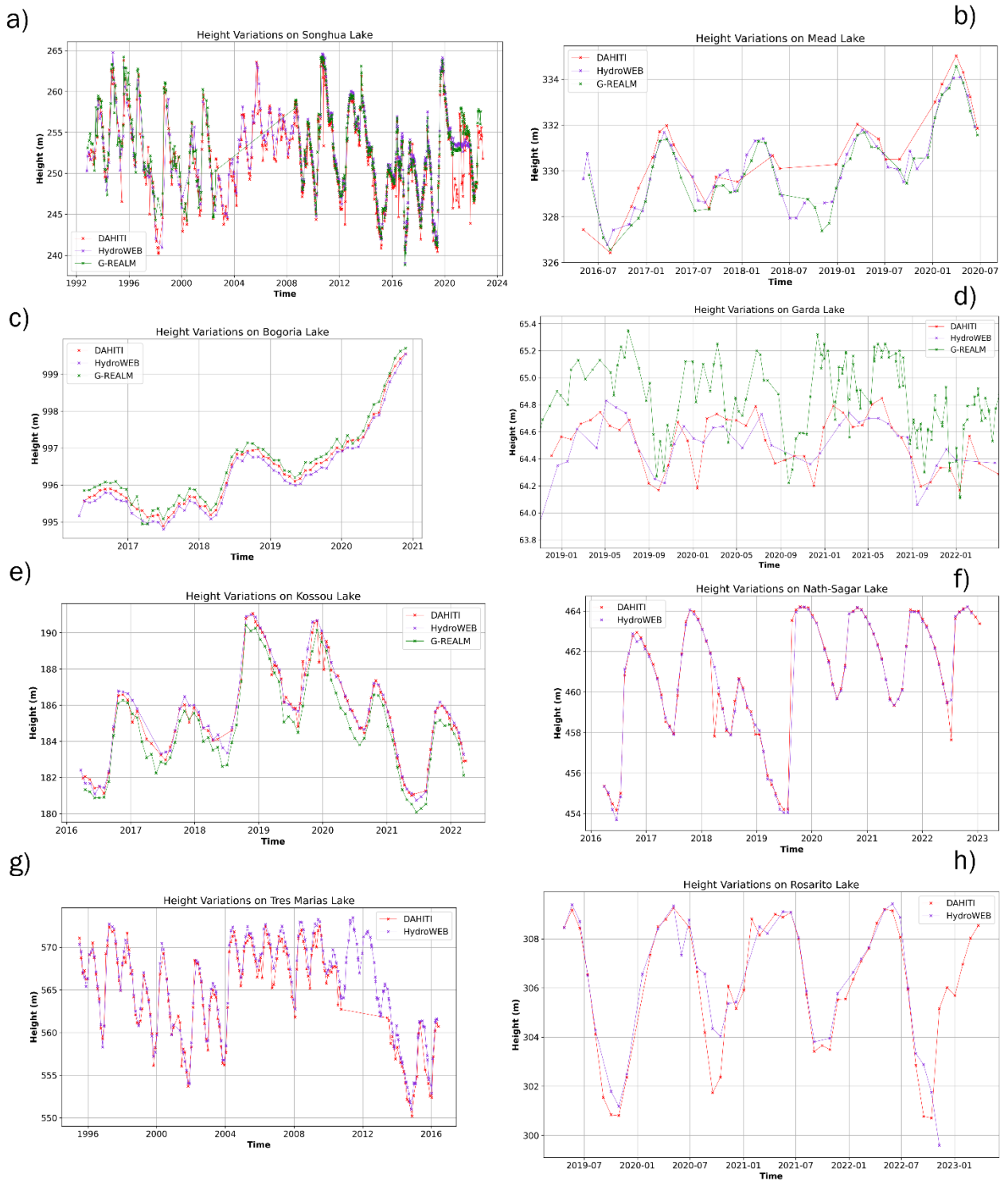
Figure 13 - Study Lakes Height Time Series

5.1.3   Impact on LSC times series

The impact of the different water height sources on LSC is evaluated in this sub-section. In-situ volume datasets were only available on three of the test lakes: the Tres Marias, the Rosarito and the Nath-Sagar lakes. The same metrics as for the assessment of the volume change calculation methods were used. and water surface area is extracted for this benchmark phase from the GSW dataset. Results are given in Table 7.

Table 7 - Statistics on volume variation depending on height data sources

Surface	Height	Metrics	Tres Marias	Rosarito	Nath Sagar
GSW	DAHITI	Pearson	0,9997	**0,9974**	0,9999
		CSI	0,9996	**0,9993**	1
		RMSE	0,5852	**0,003**	0,0576
		Relative error %	4,003	**4,1676**	2,9128
		Mean Offset (Cube km)	0,5764	**0,0023**	0,0557
	HydroWEB	Pearson	**0,9999**	0,9652	**0,9999**
		CSI	**0,9996**	0,9899	**1**
		RMSE	**0,241**	0,0145	**0,0484**
		Relative error %	**1,5909**	20,1879	**2,4293**
		Mean Offset (Cube km)	**0,2043**	0,122	**0,0455**

For the Tres Marias and the Nath Sagar lakes, the results in LSC estimation were slightly better with HydroWEB height data whereas on the Rosarito lake, results are better with DAHITI, especially regarding the relative error. The few differences in the water level time series leads to different tendencies in the hypsometric curve and then in the LSC, mainly due to the number of altimeters considered and therefore, the number of water level data (Hydroweb uses multi-mission algorithms, while DAHITI processes one altimeter). In other words, date-to-date comparisons are good between the 2 sources of LWL data but the tendency between the dots might be impacted due to the time sampling. This assessment is also observed for the Mead and Bogoria lakes studied just after.

For G-REALM data, we do not have in situ data to assess LSC with G-REALM time series. However, on the Mead and Bogoria lakes, LWL and LSC data are available in the three databases (Figure 14). The differences in LWL are linked with the algorithms used to process raw data and also with the sampling of the data (i.e., the number of altimeters considered). When looking at the LSC timeseries processed by each platform, this leads to the same differences on LSC time series. But this figure also shows that when we compute LSC based on the same methodology for each LWL timeseries (same surface extent, hypsometric curve and LSC calculation), we obtained the same LSC estimation.

This demonstrates that if a height time series is missing in a data source, we can use another source of LWL without affecting LSC estimation. But due to different altitude referential or other algorithm and altimeter considerations used in each database, using several sources for one lake involves often an offset in the LWL data which has a direct impact on LSC absolute values. We must therefore retain for each lake with one data source though.
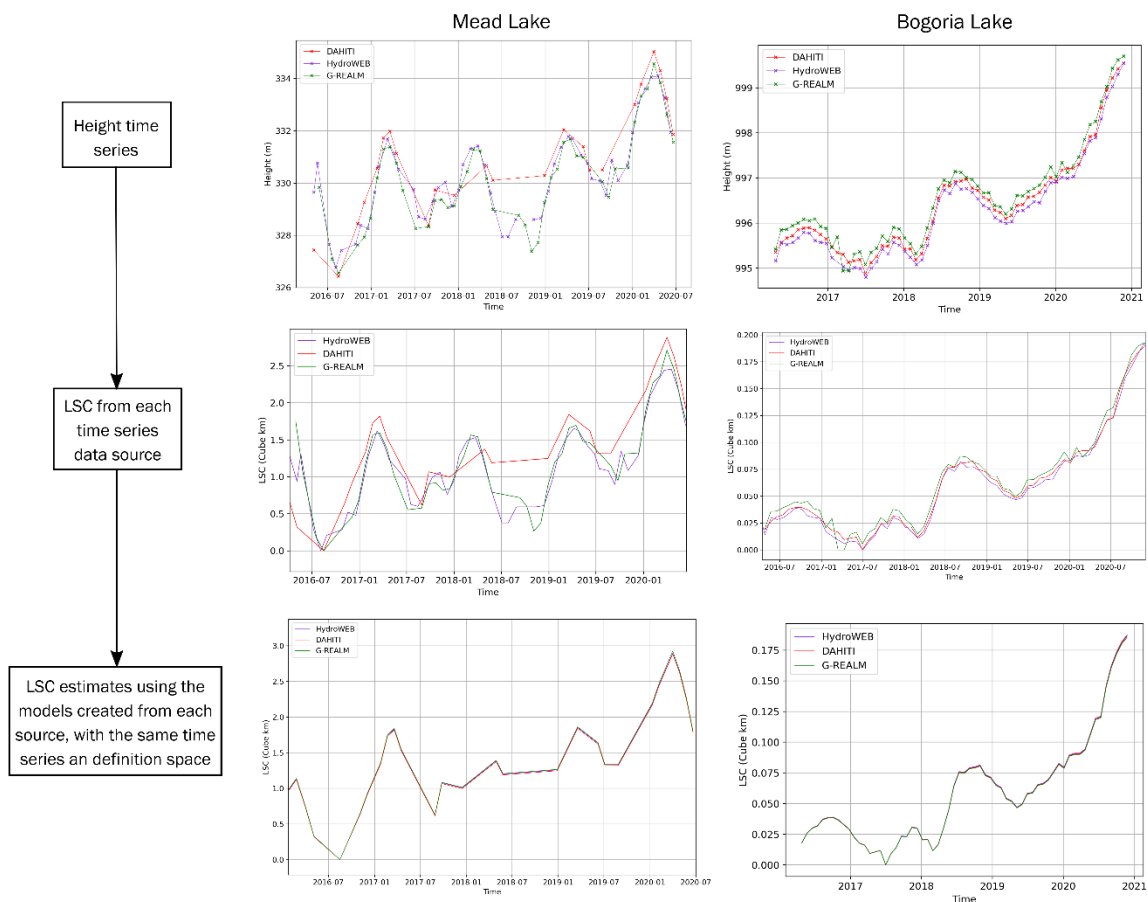
Figure 14 - Height times series to LSC depending on height data source

5.1.4 Synthesis
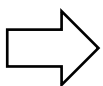
In the LSC option, ESA Lakes_cci LWL time series, available also in the Hydroweb platform, will be exploited when available. However, when LWL data are not available on Hydroweb but do exist in DAHITI and/or GREALM, we can use those data as there is a very limited impact on LSC estimation. In That case, only one source of water level must be used for a given lake.

Also, at this moment, only ~15% of world lakes (according to HydroLAKES database) are covered by radar satellite altimeters. If SWOT (launched in December 2022) is a game changer with potentially 97% or water bodies covered worldwide, the data are not yet accessible to the public, and will only cover lakes from 2023 onwards. Thus, a complementary solution can be explored by using DEM in the case no height information is possible with altimetry, notably to go back as far as possible timewise, which is of critical interest in the frame of the CCI.

5.2 Height estimation from DEM

5.2.1   General idea

In this section, the use of precise water surfaces is the starting point to estimate the water level from their contours projected on DEM. The feasibility of this technique and a critical analysis of the results observed on three different lakes is proposed. This is an explorative approach to go towards long time-series, even without past altimetry data, but also to use datasets that cover only partly an interest lake.

5.2.2   Methodology

We used two DEMs: COPDEM GLO-30 (data acquires through Tandem-X between 2011 and 2015) and MERIT (data acquired through SRTM mission in February 2000 and other optical dataset afterwards) DEMs. Water surfaces were extracted from CLS own algorithm on Sentinel-2 and Landsat images. Spatial altimetry height time series (Hydroweb, DAHITI, GREALM) were used as validation data. We have to keep in mind that ideally, and for LSC estimation, the acquisition of a DEM should be at the lowest level of the water body which is rarely the case. Hence, if we observe a lowering trend of lake water height with time, recent DEM will be more suitable than older ones.

The methodology explored is as follows:
1. Read the Digital Elevation Model geo information.
2. Open each water mask extracted from the optical sensors, and project it to precisely warp the DEM projection in extent and resolution: the nearest neighbour was used to ensure a faithful alignment with the water contours obtained from the acquisitions. The DEMs were used as spatial references, ensuring that all the projections were correctly performed towards a consistent spatial target.
3. Extract the contours from all the water zones: It should be noted here that certain water contours, sometimes less accurate due to significant cloud cover over a substantial portion of the main water body or noise along the edges caused by dense vegetation, were nonetheless retained. Corrections and filtering were necessary.
4. Extract all the DEM elevation values overlapped by the contours: After projecting the water area onto the DEM, the contour retrieval is done by tracing a curve that selects the last water pixel within the water zone. Selecting this inner edge provided the best accuracy, as the outer contour (composed of the first land pixels after the water area) resulted in overestimated heights and average did not show improvement.
5. Select a representative height value: from the extracted DEM elevation values, a histogram is used to extract the most representative height value.

5.2.3   Results and discussion

This study has been tested over Renaissance (Ethiopia), Sulunga (Tanzania) and Muriel (Canada) lakes. Renaissance lake began being filled up in 2020, which allowed the complete bathymetry to be known thanks to all available DEM. Sulunga lake was empty until the end of 2019 before being filled up thanks to heavy rainfall in Eastern Africa since 2020. Muriel lake has known a continuous decreasing trend since the 1970's but has been covered by altimetry since 2016. In this latest case, this dearth in information might be overcome with this methodology, which would be a very interesting approach in a frame of a changing climate.

Renaissance lake

The bathymetry of the Renaissance lake is perfectly know thanks to the DEM produced while it was still empty. This is a unique opportunity to get the full volume, but also to test the methdology proposed for LSC estimation on a large lake. As we can see on Figure 15 where the correlation between GREALM height series and heights derived from COPDEM GLO-30, is 0.97, the accuracy reached is very good. In the majority of the observed data in the case of the Renaissance lake, the agreement between both datasets is lower than 1m wich is the expected result as the DEM has a 1m vertical resolution. However, the last point of the curve is problematic. There is a strong offset of around 14m which is due to the fact that the observation window of the altimeter is out of the interest area of the backscattered altimetry waveform, hence inducing an underestimation of the water height.
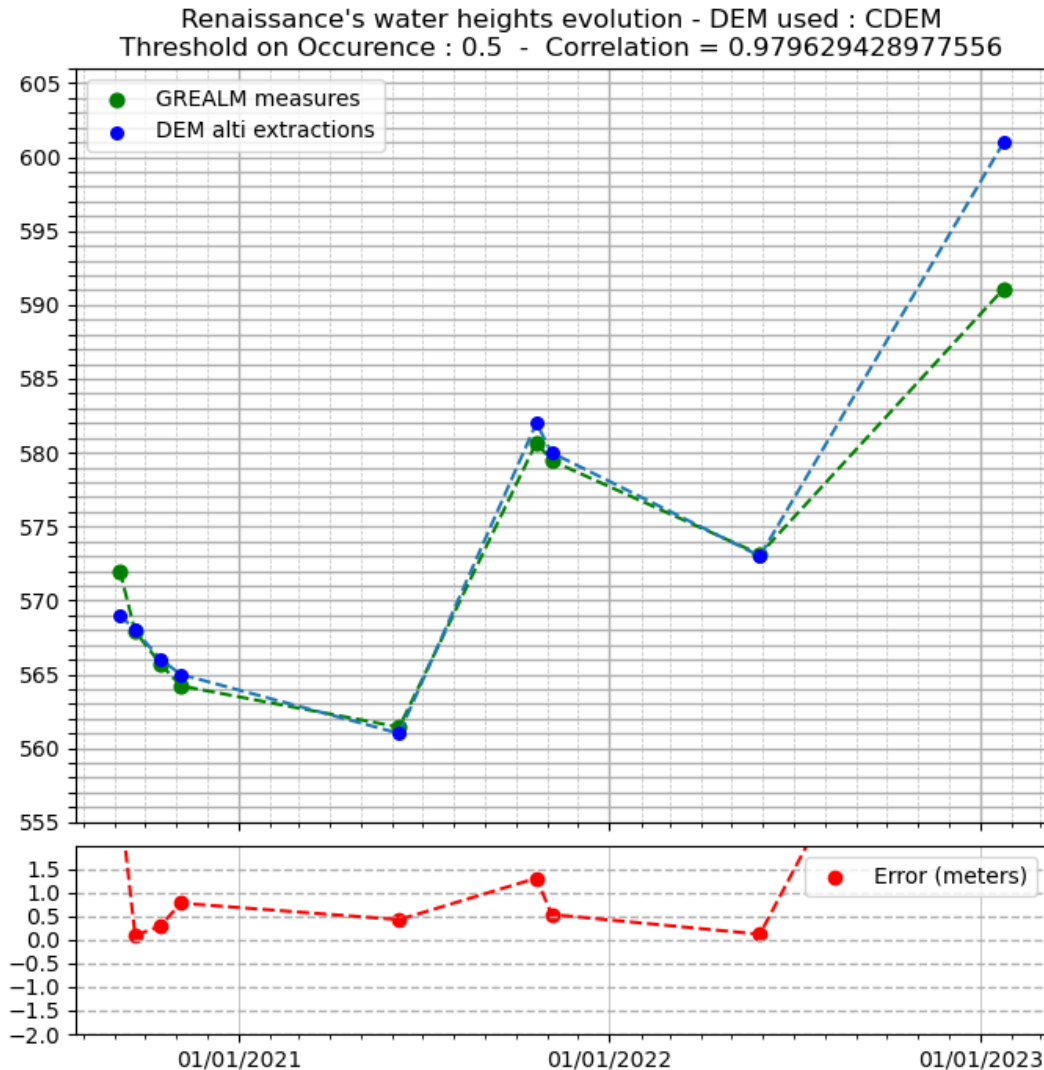

Figure 15 - Water Heights evolution on the Renaissance Lake with COPDEM GLO-30

If we compare the hypsometric curves obtained directly from COPDEM GLO-30, from the surface's contours projected on the same DSM, and the height/surfaces obtained from Sentinel-2 and altimetry,

we get the curves in Figure 16. We observe a very good agreement between the curves, showing the potential of this methodology to retrieve the hypsometric curve without altimetry information.
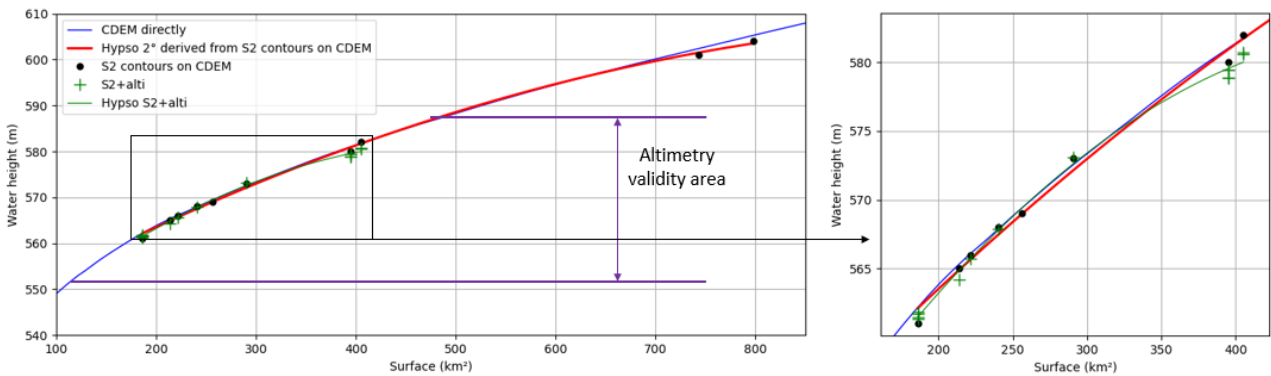
Figure 16 – Hypsometric curves over Renaissance Lake obtained from empty Copernicus DEM (blue), S2 contour projection on Copernicus DEM (red), and surface/altimetry couples (green). The current altimetry validity area is indicated in purple.

Sulunga lake

The case of the Sulunga lake (see Figure 17) in Tanzania is slightly different, as this water body was almost empty from 2014 to 2020. Once again, the differences between the two height datasets (altimetry and derived from contours projected on COPDEM GLO-30) remains in general inferior to 1m. This might be to link with the 1m resolution in height with the COPDEM GLO-30 dataset. We must mention that the surface determination is a harsh work in this case due to cloud cover and floating vegetation, what might introduce slight errors.

Also to notice, the MERIT DEM derived from SRTM represents a nearly full Sulunga lake, as it might have been the case in February 2000 (measurement date of SRTM dataset). This impedes the use of MERIT to estimate the water height.
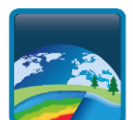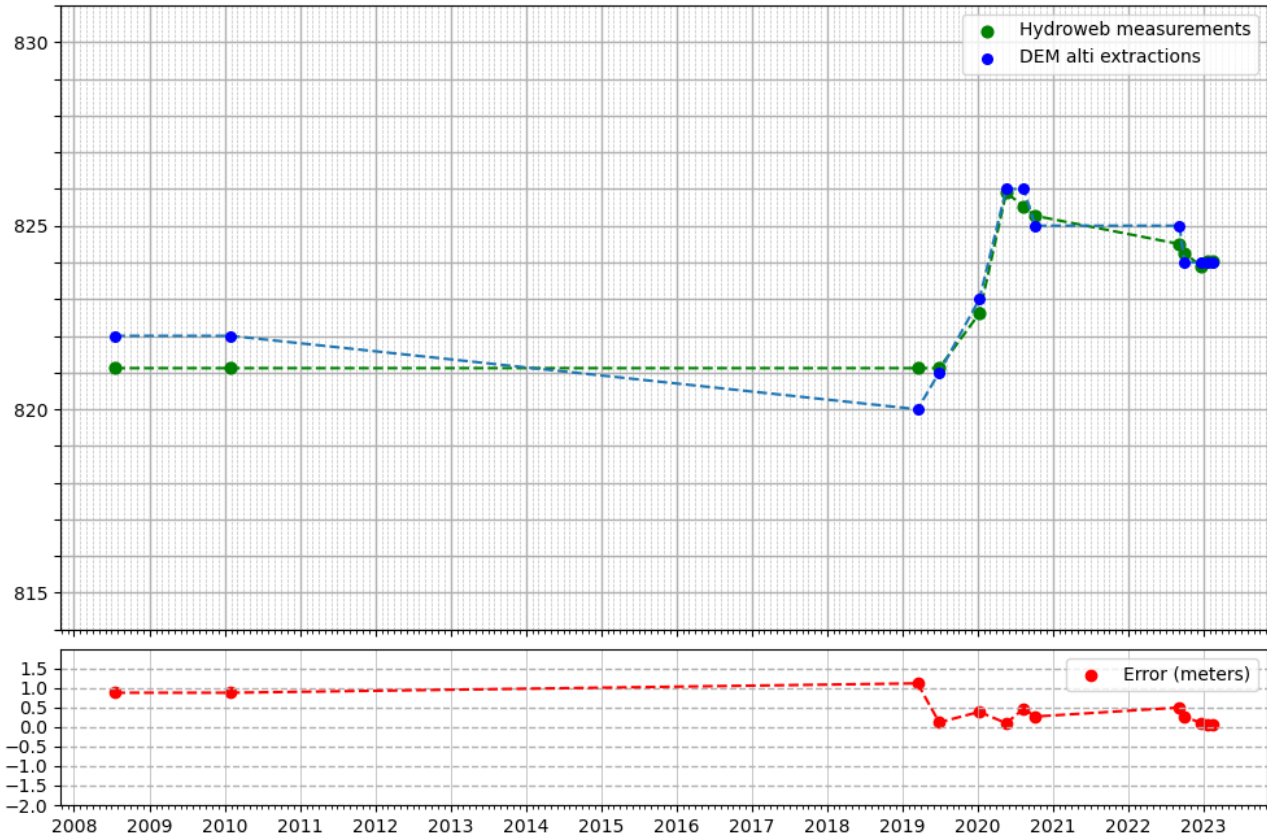
Figure 17 - Water Heights evolution on the Sulunga Lake with CDEM (COPDEM GLO-30). Differences between estimations and altimetry measurements are plotted in the bottom plot.

Muriel Lake

On the Muriel Lake (Alberta, Canada), there is no altimetry reference dataset before 2016. The lake water area has known a constant decrease from 1984 (first Landsat 4 data) to 2016, which is a marker of a decreasing water level. Hence the use of DEM can be used to retrieve such past information and to estimate the volume variation on a forty years' time span.

In Figure 18, we can see the hypsometric curve estimation from the "classic" water height from altimetry hybridized with water surface areas from Sentinel-2, but also its extension to the data based on Landsat and COPDEM GLO-30. If the behavior of the curve seems to be logical, there might still be a small water height offset issue between the two parts of the hypsometric curve, with no solution for the moment.

Figure 18 –Hypsometric curves for lake Muriel, using Sentinel-2 and Sentinel-3A information (red), Copernicus DSM directly (blue), and Landsat data projected on Copernicus DSM (black dots).

Concatenating both hypsometric curves, we can estimate the LSC variation from 1984 onwards, which is not possible considering only altimetry and optical imagery. From 2016 to 2023, there is a 100 Million Cubic Meters (MCM) variation, but thanks t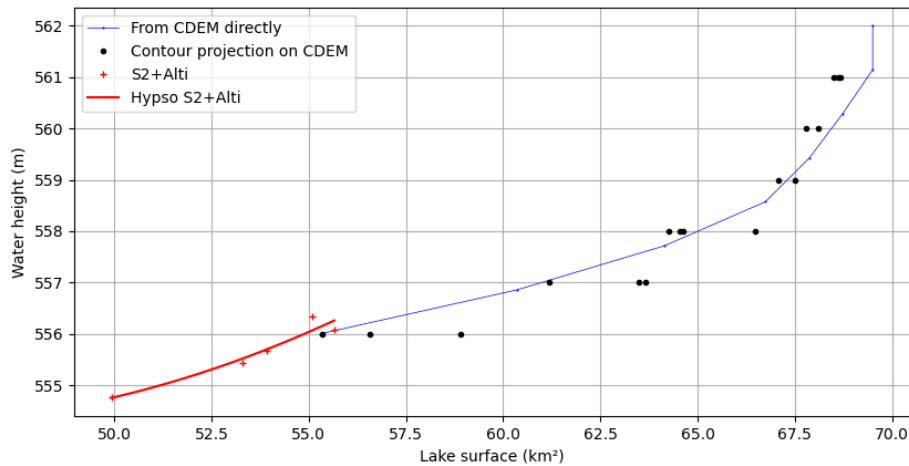o the data extracted from both Landsat and Copernicus DSM, the change between 1984 and 2016 is estimated to be around 400 MCM (see Figure 19). If this estimation is of huge interest in the frame of climate change studies, the precision of the estimation suffers directly from the low resolution of the DEM vertically and horizontally.
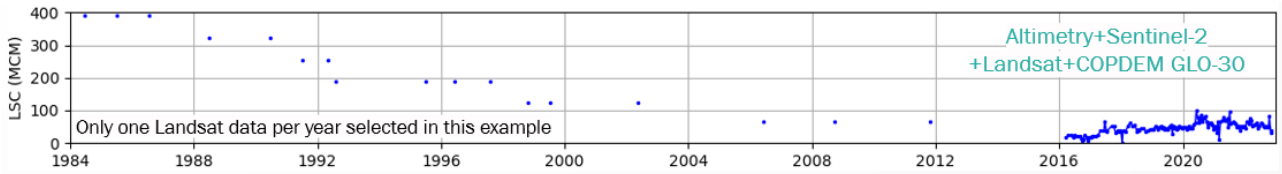

Figure 19 – LSC estimation from 1984 to 2013 for lake Muriel using Landsat, COPDEM GLO-30, Sentinel-2 and Sentinel-3A information.

5.2.4   Synthesis

This methodology allows the retrieval of height estimation from water surfaces contours projected on a DEM. We generally retrieved very good results, as shown for the Renaissance Lake in Ethiopia for instance. This may also lead to the estimation of LSC from "old" Landsat data (from 1984 onwards), what may lead to long time-series of LSC, what is of tremendous interest in the frame of the Lakes_CCI project. If we worked with complete lake images in this benchmark, this methodology could also be applied to partially covered lakes (due to cloud cover or orbit "cuts"), from which height estimation could be linked to DEM-derived hypsometric curves to retrieved past water level/surface/LSC.

The methodology is already close to an operational use, nevertheless some questions remain on the quality of the DEM (precision, resolution), but also on the number of lakes on which such time-series may be produced. If a variation between the surfaces estimated in 1984 and 2016 (beginning of Sentinel-2 time series) might be a good indicator of such feasibility, the needed AOI to encompass both water surfaces cannot be easily produced without human supervision. Also, a more recent DEM with a smaller resolution shall improve such estimations, as could be the case with the future CNES S3D2 mission. In the end, this methodology shall be applied when suitable, on a case-by-case basis, and with Landsat accurate surface estimations to be relevant.
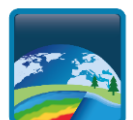
6 Benchmark of water surface estimation for LSC

To estimate lake storage change, lake surfaces as well as lake heights are necessary. Nevertheless, the lake dataset can be split into two sub-families as mentioned previously: the lakes with unvarying surfaces with height (and with time), and those that do vary with height and time. We explore first in this section a methodology to classify lakes as variant and non-variant in extent then how to estimate the surface in case of unvarying surface and benchmark the impact of all LWE methodologies presented in the state-of-the-art [R1] on LSC estimation.

6.1 Proposed methodology for lake surface variation classification

The separation between the two lake families is done using the Global Surface Water Occurrence (GSWO, see Pekel et al., 2016) dataset, which encompasses water surfaces from 1984 to 2021 using Landsat dataset. The surface difference between the surface corresponding to at least 15% of occurrence (nearly full state of the studied lake) is compared with the surface corresponding to the maximum occurrence minus 15% of occurrence (low observed state of the lake). These values have been chosen to avoid considering erroneous measurements present in the GSWO dataset or artefacts linked with border effects, or reservoir lakes created after 1984 which might propose low values of occurrence. That way, the lake surface variation is estimated in relation to its maximum extent. We chose to consider a lake as unvarying if its surface variation (ΔS) between 15% and maximum-15% of GSWO is lower than 5% of its maximum extent.

In Figure 20, we can observe the varying surface in cyan (GSWO>15%) against the permanent water in blue (GSWO>max$_{GSWO}$-15%) for three of the test lakes (Nath Sagar, Müritz and Tres Marias. Only lake Müritz can be considered as a lake unvarying surface-wise with a ΔS<5%.

Also notice that preliminary work on the area of interest (AOI) has been done to isolate the lakes and avoid considering water surfaces that do not belong to the water body of interest. The AOI precision is key for this application and must be defined carefully before any processing to make sense and be representative of water surface extrema.
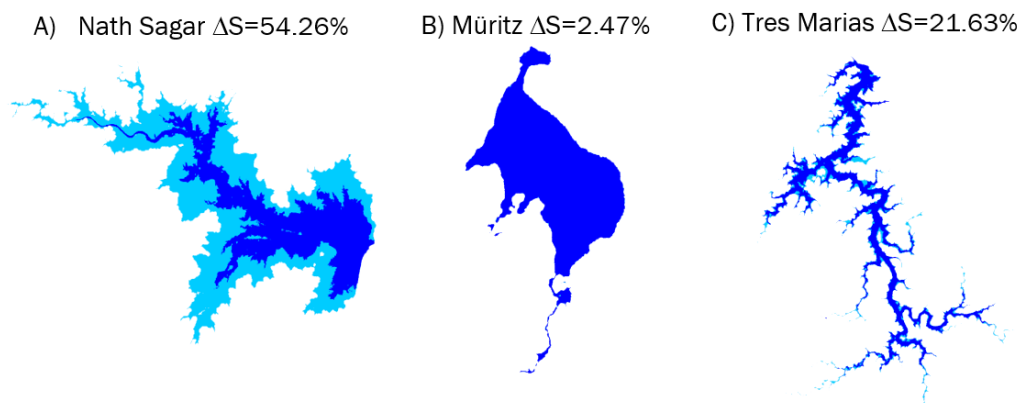
A) Nath Sagar ΔS=54.26%   B) Müritz ΔS=2.47%   C) Tres Marias ΔS=21.63%

Figure 20 – Surface variation on three lakes between 15% (cyan) and max$_{GSWO}$-15% (blue) of GSWO and their maximum extent.

To get a first glimpse of the surface variation of the lakes considered in the ESA Lakes_cci, we applied the methodology described previously and considered the AOI defined by the Lakes_cci project. The histogram obtained is displayed on Figure 21, and shows that nearly half of the lakes might be considered as unvarying if we consider a 5% threshold in (ΔS). We must remain careful though, as the AOI delimited by the Lakes_CCI shapefile does not encompass the lake's whole surface in many cases, despite a small buffer applied of 0.005°.
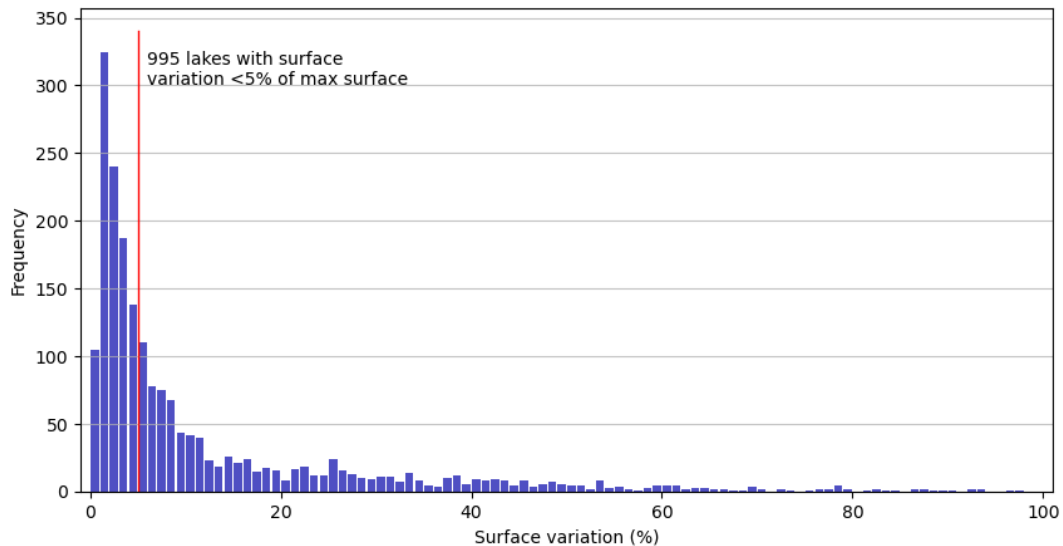

Figure 21 - Histogram of lake's surface variation between 5 and 90% occurrence on the GSWO dataset.

6.2 Surface unvarying lakes

As explained previously, the best method to estimate LSC for an unvarying lake is to use the basic approach, i.e., multiplying the constant surface with the height variation from altimetry (see Section 4.2).

One commonly used dataset giving several lakes parameters is the Hydrolakes database (see [R1]). Nevertheless, errors have been observed in certain cases, as can be seen on Figure 22 on Garda (Italy), Iseo (Italy) and Richland-Chambers (USA) lakes. The blue parts correspond to water pixels from GSWO which are not considered as so by Hydrolakes (orange polygons). This shows that Hydrolakes might not be reliable enough to precisely determine the lake's surface, as an error on the constant surface would introduce a constant error in the LSC estimation.
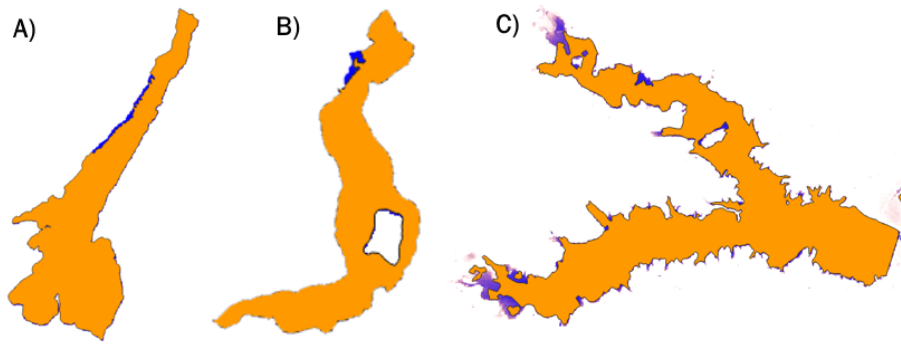
Figure 22 – Lakes as observed by GSWO (in blue) and by Hydrolakes (in orange). A) Garda (Italy). B) Iseo (Italy). C) Richland-Chambers (USA)

Therefore, the best solution is to calculate the mean surface area between the surfaces measured at highest and lowest water levels.

6.3 Impact of LWE on LSC for varying lake

In the case where lake vary surface-wise in relation with the water height, the establishment of a hypsometric curve is needed to estimate properly the water volume changes. In the state of the art [R1], some water area data sources were introduced. In this section, we evaluate the impact of the use of these surface products and algorithms on the LSC estimation. To that end, a comparative analysis was made on four lakes where in-situ data of volume are available: Richland-Chambers (USA), Rosarito (Spain), Tres Marias (Brazil) and Nath Sagar (India).

6.3.1   Assessment method

The data sources for water surface areas to benchmark are (see [R1]):

-    Lakes_cci LWE : water surfaces used to compute the hypsometric curve. [R1]
-    DAHITI (Database for Hydrological Time Series of Inland Waters), see Schwatke & al. 2020
-    GSW (Global Surface Water), see Pekel & al. 2016
-    GWW (Global Water Watch), see Donchyts & al. 2022
-    SurfWater developed by CNES, see Peña-Luque & al. 2021
-    WIZIR developed by CLS

For each source, a hypsometric curve is estimated and the LSC computed to analyse the impact of the different area datasets on LSC estimation. Complementary metrics were calculated: metric of correlation to be assured that trends are the same or equivalent and metrics to evaluate the relative and absolute difference between time series. These two aspects of validation are necessary because trends can be right even if differences appear due to a height offset for instance. For correlation measurement, the cosines similarity index (CSI) and Pearson correlation were chosen. The implementation of CSI is made from Equation 4 in Table 2, Pearson correlation is implemented in the NumPy library. For relative and absolute difference, as for LSC method choice, the relative error and the mean offset will be used (see Table 2).

The time series based on GSW and CLS, were filtered to remove obvious outliers (corrupted images...).

6.3.2   Results

Table 8 synthetized all numerical results and Figure 26 shows the LSC timeseries. Whatever the water area source chosen, results are good, and all correlations are above 95%. For each lake, the water extent that gives the best results compared to the in-situ data varies and despite different estimated hypsometric curves, the LSC time series are consistent with each other and with in-situ data.

The first reason of this result is related to the fact that all water area time series are coherent with each other (see Figure 23), except for some offsets which can be linked to the AOI of the lake, the spatial and temporal resolution, or the method used to determinates these areas (see further discussion in next sub-section).
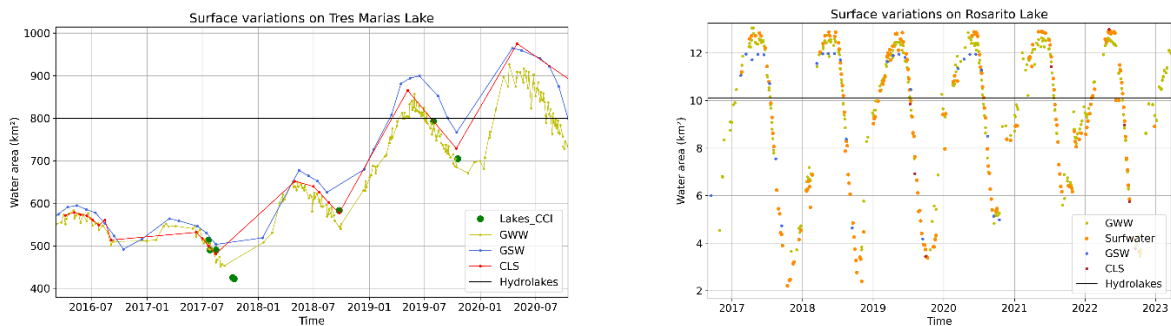

Figure 23- Surface time series of Tres Marias and Rosarito Lakes for different water surface areas sources

The water surface used to provide the **Lakes_cci LWE** are available on Mead, Kossou, Prespa, Tres Marias and Nath-Sagar lakes. Results are very good and quite coherent with other data sources. Best results are obtained on Nath-Sagar LSC. Most offsets are linked to differences in estimation method. A more specific analysis of the importance of the AOI is given in Section 6.3.3.

For the Tres Marias and the Nath Sagar lakes, statistics would recommend using the Global Surface Water **GSW data** covering more than 30 years of Landsat images (Pekel et al. 2016). The main advantage of this dataset is its temporal coverage from 1984 to 2021, plus the fact that water extents are already processed. Having the longest timeseries possible of water extent is of prime importance for climate studies to ensure to have extrema. One relevant example is the Muriel Lake in Canada. Figure 24 shows the water extent timeseries from 1984 to 2021 that underlines an important extent decrease of the lake since 1984 that cannot be seen if only Sentinel 2 data are considered.
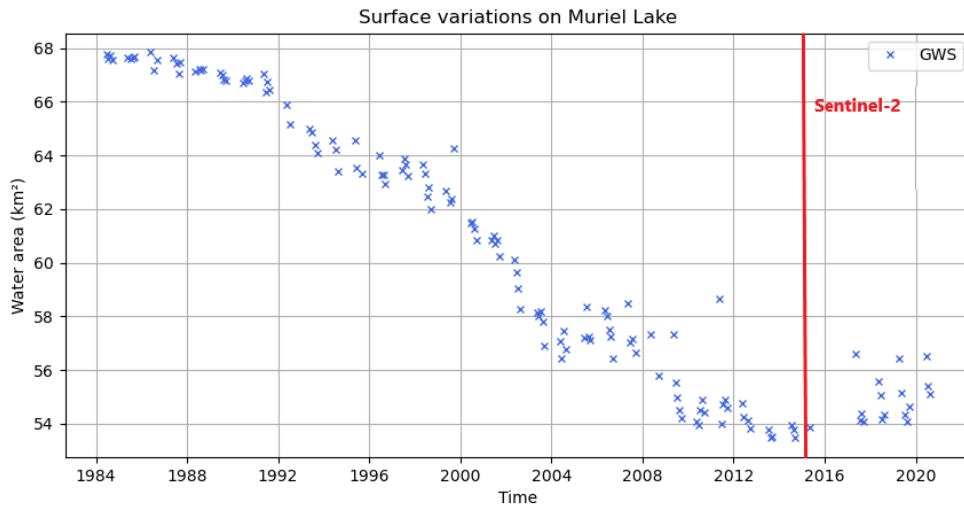
Figure 24- GSW monthly water area time series of Muriel Lake -1984 to 2021 with the materialization of the availability of Sentinel 2 data

However, GSW data have been pre- and post-processed carefully to obtain such results. Indeed, each GSW tile covers about 10°x10°: pre-processing is necessary to extract the lake extent. Sometimes tiles are empty or corrupted by some artefacts: on the Figure 25 below, water has been classified as land for Bogoria. The linear artefacts are symptomatic of Landsat-7 damaged sensors. With a resolution of 30m per pixel, this kind of error can make a surface product unusable as is.
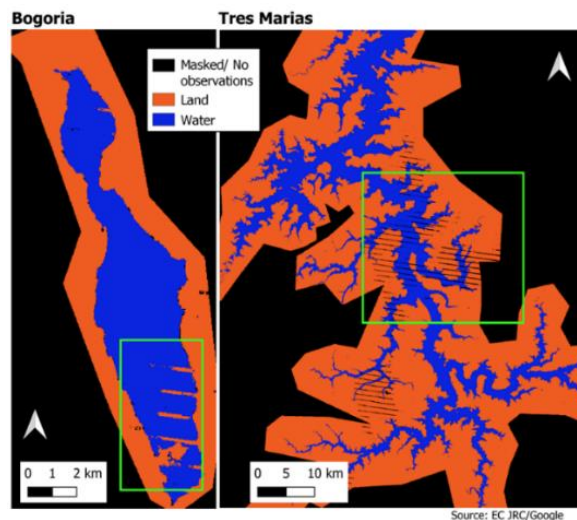

Figure 25 - Characteristic defects in Landsat images on GSW surfaces

Thus, even if we gain some time by not processing raw satellite data to extract water extent, some pre and post processes on GSW tiles are mandatory.
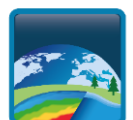
Table 8 - Statistics on volume variation estimation depending on water area data sources.

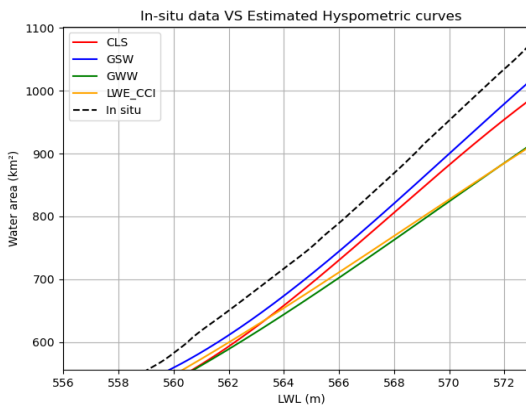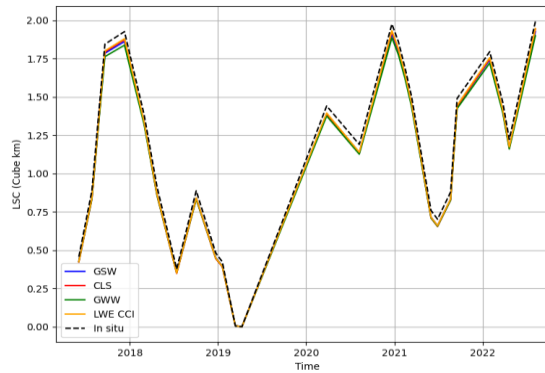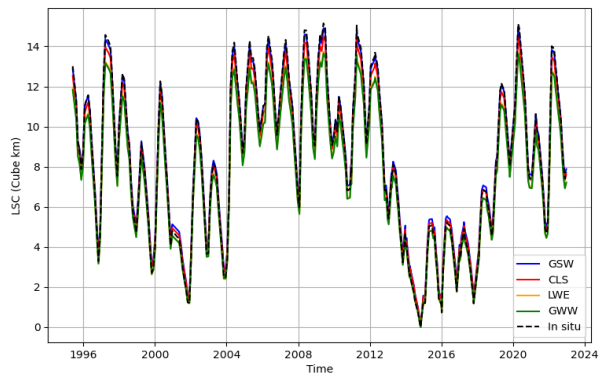|		Tres Marias	Nath-Sagar	Richland-Chambers	Rosarito
LWE CCI LWL : HydroWEB	R2 Pearson	0,9998	0,9999	-	-
	CSI	0,9999	0,9999	-	-
	RMSE	0,6782	0,0417	-	-
	Relative error (%)	4,4766	2,0936	-	-
	Offset (Cube km)	0,5551	0,0398	-	-
GSW LWL : HydroWEB	R2 Pearson	0,9999	0,9999	0,9888	0,9652
	CSI	0,9999	1	0,9983	0,9899
	RMSE	0,3228	0,0484	0,0312	0,0145
	Relative error (%)	2.7501	2,4293	5,4698	20,1879
	Offset (Cube km)	0.2640	0,0455	0,0244	0,0122
SurfWater LWL : HydroWEB	R2 Pearson	-	-	-	0,9703
	CSI	-	-	-	0,9922
	RMSE	-	-	-	0,0104
	Relative error (%)	-	-	-	14,3807
	Offset (Cube km)	-	-	-	0,0083
DAHITI LWL : DAHITI	R2 Pearson	-	-	0,9895	-
	CSI	-	-	0,9983	-
	RMSE	-	-	0,0213	-
	Relative error (%)	-	-	3,7341	-
	Offset (Cube km)	-	-	0,016	-
GWW LWL : HydroWEB	R2 Pearson	0,9999	0,9999	0,9862	0,9643
	CSI	0,9999	1	0,9981	0,9894
	RMSE	0,7301	0,0615	0,023	0,015
	Relative error (%)	5,3094	3,0863	4,0274	20,8866
	Offset (Cube km)	0,6008	0,0567	0,017	0,0127
CLS LWL : HydroWEB	R2 Pearson	1	0,9998	0,989	0,9696
	CSI	1	0,9999	0,9983	0,9919
	RMSE	0,5154	0,0481	0,0224	0,0095
	Relative error (%)	4.2509	2,4168	3,929	13,1781
	Offset (Cube km)	0,4437	0,0458	0,0166	0,0074
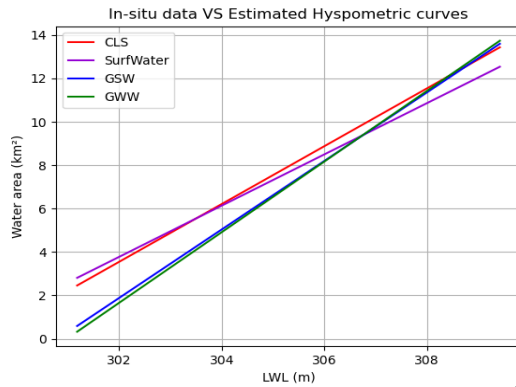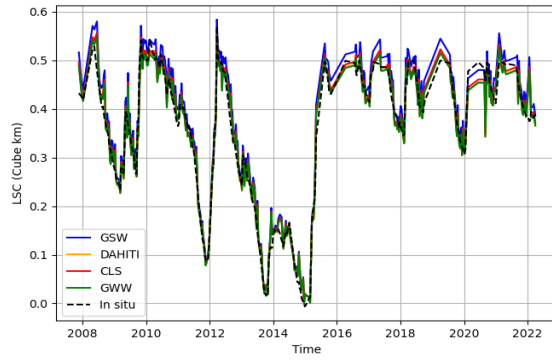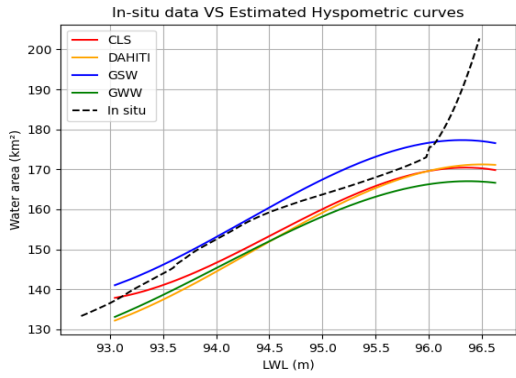
Figure 26- Hypsometric curves and LSC estimation on test lakes

DAHITI area data are only available for the Richland-Chambers. It can nevertheless be said that on the Richland-Chambers Lake, DAHITI water areas gives the best results on LSC estimation. This can be linked to the fact that water height data are also from DAHITI. The data is quite precise, and an error value is given. Also, the data is easily accessible for download. Nevertheless, we do not have access to the water extent raster, so we do not know the lake boundaries according to DAHITI which can be a limitation in the future LSC determination for data coherence between sources.

As for DAHITI, **SurfWater** data was only available for one test lake, the Rosarito Lake. The Surfwater data provides the best correlation of LSC with the in-situ data on this lake, and quite similar results with CLS. Currently Surfwater water extent maps are based on Sentinel 2 data and not available worldwide yet. The operational production is done within the THEIA Land Data centre in Europe and in certain regions of Africa and is expected to be extended at global scale soon.

The **CLS inhouse algorithm**, based on a dynamic thresholding approach for multiple indexes of Sentinel 2 and Landsat image, gives overall rather consistent results of LSC compared to the in-situ.

The **Global Water Watch** GWW platform does not provide the best results of LSC according to in-situ data, although they are satisfactory for a global product. These water area time series have the advantage to be easy to download and use. The observed differences could be explained by the boundaries of the lakes chosen to create water areas time series.

6.3.3   The importance of the AOI

We noted during the analysis of the difference water surface data source that the way the AOI is defined may have a significant impact on LSC estimation. It concerns all water surface methods and varies depending on the considered lake.

The following example on the Kossou Lake illustrates the importance of the AOI. Indeed, the strongest offsets of water surface between the Lakes_CCI data and CLS data (blue points on Figure 27) are observed on high value extents. It is directly linked to the way AOI is determined. Here, the AOI used with CLS algorithm is larger more than 100km² (Figure 27) than the AOI used for the Lakes_CCI LWE. As we can see on Figure 28, this leads to quite different hypsometric curves. However, the consequences on LSC estimation are limited generally, but can go up to 0.4 km³ (~10% of maximum LSC). In this example, we don't have reference data to consider which AIO is best but clearly and whatever the source, a vigilance on the creation of the AOI must be carried out.
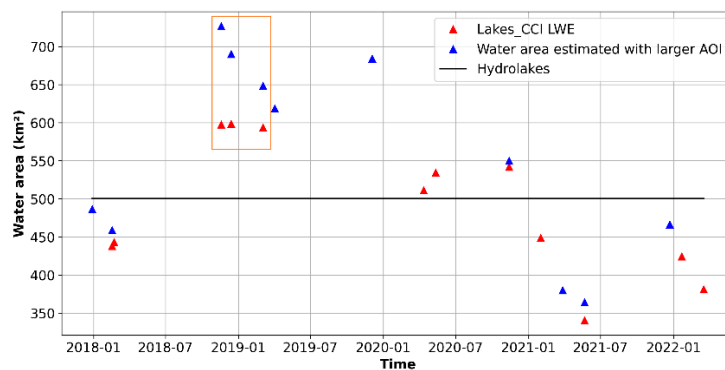

Figure 27 - CLS water areas (blue) and LWE CCI times series, estimated with different AOI
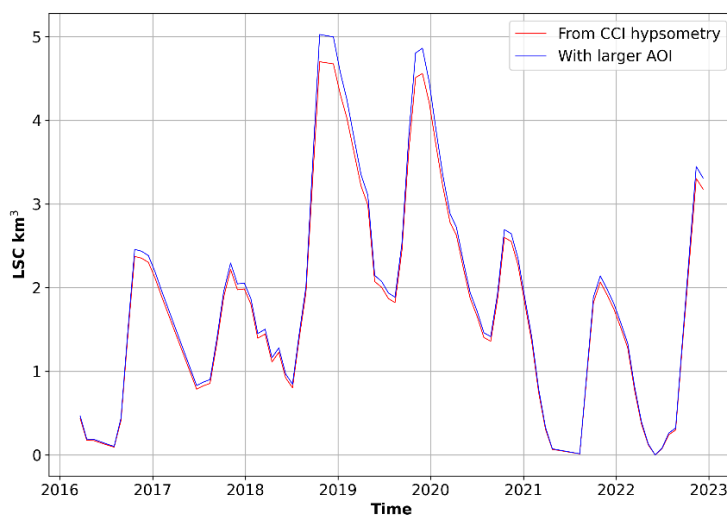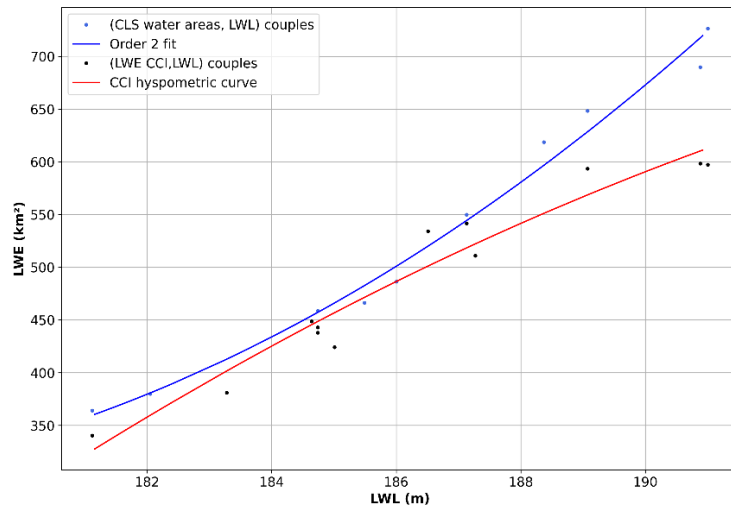
Figure 28 - CCI hypsometric and hypsometry from data with larger AOI and its impact on LSC (Kossou Lake)
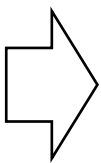
The impact of the water surface regarding the different methodologies on LSC is limited. We however recommend using water surface timeseries from the Lakes_CCI or GSW, Surfwater and CLS algorithm (DAHITI water surface maps are not available and GWW gives the poorest results). Prior to any processing, a check of the AOI and consistency of water surface data is essential to ensure to capture the extrema and have the best LSC estimation.

For unvarying lakes surface-wise, a mean surface between highest and lowest heights will be estimated.

7 Conclusions

All the methodologies of the different steps of the LSC estimation have been reviewed in this benchmark document.

The main conclusions are first to determine if the lakes surface variation is higher than 5% or not. If not, lake can be considered as constant in surface and therefore LSC will be estimated with the basic volume method and a mean water surface between the lowest and highest water level. Otherwise, the process must go through the hypsometric curve generation from water level and extent couples.

For the demonstrated lakes in this benchmark: hypsometry will be preferred for all lakes except Garda, Mille Lacs and Müritz lakes (respectively 2.56%, 0.5% and 2.47% of surface variation).

The impact of water level time series with altimetry from 3 different databases and water surface timeseries from 6 available methods have a limited impact on LSC errors. Due to the offset in water level regarding the source of the data, one must ensure to use only one source of data for a given lake. And for lake surface data, the AOI must be checked to avoid any additional errors in LSC.

Very promising results were obtained using DEM to retrieve water level. It can be very useful when no altimeter overpasses a lake and to extend the LSC timeseries for climate studies when only water level from recent altimeters is available.

The definition of the hypsometric curve is for the moment best with polynomial and parametric compensation, only in the validity domain where there are height /area observation couples. With the SWOT satellite, we expect that the Gauss Helmet method will work best. In any case, a pre-processing to remove outliers must be performed first and we would recommend a semi-automatically method by combining the automated selection of outliers (on water masks) and the application of RANSAC algorithm.

And finally, to estimate LSC from the hypsometric curve, integration between two consecutive dates gives best results.

In the next phase of conception, we will define and develop the complete pipeline: lake morphology classification, hypsometry estimation, LSC computation. To do this, more testing will be needed on the automation of lake morphology characterization, which requires the use of reliable AOI, and further studies will be done on the eventual pre-processing of outliers' detection in surfaces or/and heights data.

This phase done, we will be able to do the validation and quality assessment of the process.
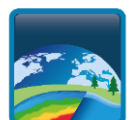
Appendix A - Test lakes

Table 9- Test lakes and their description following the selection criteria

Country	Name	L or R?	Year built	Altimetry	Elevation (m)	Max size (km²)	(A)	(B)	(C)	(D)	(E)	Surface variation variability	CloudCover (% annual)	CCI?	CCI lake LWL	CCI lake LWE	Hydroweb H	Hydroweb V	Dahiti H	Dahiti V	GRE ALM H	A-H	V-H	Bat.	TS H	TS A	TS V	Empty
Brazil	Tres Marias	R	1962	S3B	566	1040	1					Medium	45	1	1	1	1	1	0,5	0	0	X	X					
Canada	Muriel	L	n/a	S3A	555	64,1				1		Very low	56	1	0	0	0	0	1	0	0			X	X			
China	Songhua	R	1937	JA/S3A	243	550				1		Medium	52	1	1	0	1	0	1	0	1	X	X					
Ethiopia	Renaissance	R	2020	JA	450	1874	1					Very high	31	0	0	0	0	0	0,5	0	1				X			X
France	Montbel	R	1985	S3B	400	5,5			1			Very high	60	0	0,5	0	0,5	0	0	0	0				X	?	?	?
Germany	Müritz	L	n/a	JA/S3B	59	117				1		Very low	70	1	0	0	0	0	0	0	0			X				
Greece Macedonia Albania	Prespa	L	s	S3A	841	259			1			Low	50	1	1	0	1	0	1	0	0				X	X		
India	Nath Sagar Jalashay	R	1976	S3A/S3B	460	350		1				High	39	1	1	1	1	1	1	0	0	X	X					
Italy	Garda	L	n/a	ENV/JA/S3B	62	370			1			Very low	46	1	1	0	1	0	1	0	0			X				
Ivory coast	Kossou	R	1972	S3A/S3B	185	1700	1					Medium	66	1	1	0	1	0,5	1	0	1			X				
Kenya	Bogoria	L	n/a	S3A	990	34	1					High	49	1	1	1	1	1	1	0	1			X				
Mexico	Patzcuaro	L	n/a	JA/S3B	2037	260			1			Low	51	1	0	0	0	0	0	0	1			X				
Norway	Møsvatn	R	1906	JA/S3B	919	79,1					1		76	1	0	0	0	0	1	0	0							
Philippines	Laguna de bay	L	n/a	S3B	2	911	1					Medium	67	1	0	0	1	0	0,5	0	0	X	X	X				
Spain	Rosarito	R	1958	S3B	306	15			1			High	40	1	0,5	0	0,5	0	0,5	0	0						X	
Tanzania	Sulunga	L	n/a	ENV/S3B	823	974		1				Very high	38	1	0,5	0	0,5	0	1	0	0							X
USA	Teshekpuk	L	n/a	S3A/S3B	9	830					1	Very low	47	1	0,5	0	0,5	0	1	0	0			X				
USA	Mead	R	1936	JA/S3A/S3B	372	640		1				High	24	1	1	0	1	1	1	1	1				X	X	X	
USA	Mille lacs	L	n/a	S3A/S3B	377	536				1		Very low	58	1	1	0	1	0	1	0	0			X				
USA	Richland-Chambers	R	1988	JA/S3B	97	168			1			Medium	49	1	0	0	0	0	1	1	0	X	X		X	X	X	