Consortium Members

D2.2: Algorithm Theoretical Basis Document (ATBD)

Reference:	CCI-LAKES-0024-ATBD
Issue:	2. 1
Date:	29 Sept. 2020

esa
European Space Agency

lakes cci

Version history:			
Issue:	Date:	Reason for change:	Author
1. 0	9 Dec. 2019	Initial Version	S. Simis et al.
1. 1	24 March 2020	Revision following ESA review	S. Simis et al.
1. 2	24 April 2020	Revision following 2nd ESA review	S. Simis et al.
2. 0	30 July 30 2020	Update of the LSWT chapter to take into account MODIS data. Update of the 4. 2: Optical LWE Estimates	Chris Merchant Herve Yesou
2. 1	29 Sept. 2020	Revision following ESA Review	J-F Crétaux, Hervé Yésou, Chris Merchant

People involved in this issue:			Signature
Authors:	Stefan Simis, Xiaohan Liu	Plymouth Marine Laboratory	
	Jean-François Crétaux	LEGOS	
	Hervé Yésou	SERTIT	
	Erik Malnes, Hannah Vickers	NORCE	
	Pablo Blanco	TRE Altamira	
	Chris Merchant, Laura Carrea	University of Reading	
	Claude Duguay	H20 Geomatics	Claude Duguay Digitally signed by Claude Duguay Date: 2020.10.05 09:51:11 -04'00'
Internal review:	Stefan Simis,	Plymouth Marine Laboratory	Stefan Simis Digitally signed by Stefan Simis Date: 2020.10.05 15:44:43 +01'00'
Approved by:	B. Coulon	CLS	Bruno Coulon Signature numérique de Bruno Coulon Date : 2020.10.05 16:50:02 +02'00'
Authorized by:	C. Albergel	ESA	Clement Albergel Digitally signed by Clement Albergel Date: 2020.10.05 21:32:59 +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	bcoulon@groupcls. com

Distribution:		
Company	**Names**	**Contact Details**
	B. Calmettes	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@unistra. 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

1. Summary

This document specifies the theoretical basis for the algorithms used to produce the five thematic parameters of the Lakes Essential Climate Variable (ECV) under the Lakes_cci. The document presents, for each algorithm contributing to product generation, the detail required for users to gain an informed understanding of the technical and scientific considerations underlying these products, ranging from the scientific description to functional (inputs, outputs) and mathematical definitions of the algorithms as well as references to literature underpinning the performance of these algorithms. Readers interested in a comparative analysis of candidate algorithms are referred to the Product Validation and Intercomparison Report (PVIR). The scope of this document is the entire product generation chain.

2. Introduction

The lakes cci benefits from pre-existing, mature processing systems that are used in operational context. These are joined in a modular, distributed fashion to generate the Lakes_cci climate data records. The main challenges for the lakes cci are to produce an internally consistent ECV consisting of five domains:

- Lake water level (LWL) and lake water extent (LWE)
- Lake surface water temperature (LSWT)
- Lake Ice Cover (LIC)
- Lake Water-Leaving Reflectance (LWLR)

These are distinct domains with respect to their algorithm basis, with the exception of LWL and LWE which may be linked through hypsometry (see below). Alternative approaches to determine LWE are also discussed in this document whereas the final LWE algorithm will be issued in a future update of this document. All outputs from WP6 devoted to New Methodologies for LWE are included in this document.

The algorithms for each of the domains listed above are described in detail over the following chapters.

3. Lake Water Level (LWL) algorithms

3.1. Description

Altimetry was originally designed for oceanography in the 1970s. It was used to study favourable continental surfaces, especially in hydrology and glaciology, two themes for which monitoring the height of water or ice surfaces is crucial. The principles of measurement have not changed over time but the interpretation of measurements becomes more complex with increasing heterogeneity of the target, or the presence of slopes. Altimetry is not a singular measurement - one speaks of altimetry missions because multiple sensors on board a satellite contribute to the quality of the measurement. Accurate orbit positioning sensors are crucial, as well as radiometers to determine the influence of atmospheric moisture on the signal. Dual-frequency altimeter systems for correcting the ionospheric delay are combined onto the same platform to gain the required precision in height measurement. However, these auxiliary sensors do not work for all continental surfaces as they do on ocean surfaces. This chapter, therefore, discusses the proposed solution for the particular challenge of altimetry of lake surfaces.

Satellite altimeters are designed to measure the two-way travel time of short radar (or laser) pulses reflected from the Earth's surface which gives the distance between the satellite and the reflected surface, called "range". The shape of the reflected signal, known as the "waveform", represents the power distribution of accumulated echoes as the radar pulse hits the surface. The so-called onboard tracking system is the software which attemps to keep the reflected radar echo within the receiver observation window. The resulting waveforms are called 'tracked waveforms'. The travel time is calculated using a predefined analytic function, which fits the time distribution of the reflected energy. The fitting process of the acquired waveform is called re-tracking. The first altimetry missions were designed for the ocean domain and the corresponding algorithm, the so-called Brown model (Brown 1977) was fitted to classic ocean surfaces. There it is considered that thermal noise is followed by a rapid rise of the returned power called 'leading edge', and a gentle end sloping plateau known as 'trailing edge'. However, over the continents the waveforms are generally contaminated by noise resulting from multiple land returns such as vegetation, bare sands, or steep shorelines. Consequently, the shape of the echoes reflected by continental waters is often very different from that reflected by the ocean surface. It can thus become difficult, if not impossible, to calculate the water level of a river or small lake using the classic Brown analytic function. One way of working around this is to use alternative and more suitable re-tracking functions of the waveforms. Moreover, several corrections that are commonly well measured over ocean are also degraded over continental surface and lead to use specific models.

3.2. LWL algorithm definition

Here, LWL is measured using satellite radar altimetry (alternatively, Lidar altimetry, for example on ICESat-1 missions, can also be considered).

Radar altimeters send an electromagnetic pulse to the satellite nadir and record the propagation time to and from the emitted wave and its echo from the surface. The electromagnetic bands of interest are the Ku and Ka bands, with are reflected perfectly – without penetration – by water (which is not the case for snow and ice). Multiplied by the speed of light c, half the time it takes for the transmission Δt gives length R (the range) between the satellite and the reflective surface:

$$R = c \frac{\Delta t}{2} \qquad [3.\ 1]$$

The height H of the reflective surface is given by the following equation:

$$H = a - (R + \Sigma C_p + \Sigma C_g) \qquad [3.\ 2]$$

where *a* is the orbital altitude of the satellite with respect to the ellipsoid. Corrections must be made for propagation in the atmosphere (C_p) and also vertical movements of the Earth's crust (C_g).

The ellipsoidal height is then converted into elevation *h*, taking the local undulation of geoid N into account:

$$h = H - N \qquad [3.3]$$

The ΣC_p and ΣC_g terms in Equation 3. 2 correspond to sets of corrections that must be subtracted to arrive at an accurate estimation of H.

There are two types of corrections:

- propagation corrections (Cp) needed because the radar pulse propagates through the atmosphere at a speed below the speed of light c used in Eq. 4. 1
- geophysical corrections (Cg) linked to the vertical movements of the Earth surface (tides, for example) and for which we want to correct the measurement in order to apply it to a fixed geodetic datum in the terrestrial reference frame.

Finally, we can express the height of a lake by the full following equation:

$$h = a - R - DTC - WTC - IC - ET - PT - LT - SSB - N \qquad [3.4]$$

where DTC is the dry tropospheric correction, WTC the wet tropospheric correction, IC the ionospheric correction, ET the Earth tide, PT the polar tide, LT the lake tide and SSB the instrumental so-called sea state bias. The means to calculate these terms to determine mean lake height are given in the following section.

3.2.1. LWL input data and corrections

The altitude of the satellite (a) is calculated by the ground segment of each mission and is usually provided by the so called GDRs (Geophysical Data Records) that are made available by ESA, CNES and NASA on freely web access. We use the AVISO web site service to extract the altitude of the satellite as well as other parameters.

Over continental waters, the radar echo which is assembled into so-called waveforms in the LRM mode usually does not fit with the Brown model, for which the on-board tracking algorithm was adapted. Hence, alternate algorithms that are run during post-processing, the so-called retracking algorithms, are used to calculate the range (*R*). One of these algorithms is the Ice-1 algorithm, based on an Offset Center Of Gravity (OCOG) scheme, which outperforms alternatives and is considered most suitable for lakes. This retracker determines the shape of the waveform in terms of its amplitude, width, and centre of gravity, allowing the leading-edge position (LEP) of the waveform, which is directly linked to the range, to be calculated. The Ice-1 has proven to be robust for non-Brownian waveforms, such as those generally registered over lakes and rivers. It is, therefore, often used in the application of satellite altimetry for hydrology studies.

The DTC is directly proportional to the atmospheric pressure and it is given by the GDRs. Altitude of the lake is taken into account for atmospheric pressure used in the calculation of the DTC since the launch of Jason-2.

WTC is related to the water vapor contained in the air column that the electromagnetic wave intersects. This correction can be estimated in two ways: either with an onboard bi- or tri-frequency radiometer or from a global meteorological model, as used for dry tropospheric correction. It has been pointed out that the correction using a radiometer is highly erroneous over continental water or coastal regions, due to land contamination, up to a distance of 20 to 30 km from the coastline. Therefore, apart from very large lakes, the WTC used in continental water comes directly from a model based on climate gridded data sets of multi-layer water vapour and temperature fields based on the ECMWF reanalysis. WTC varies geographically and seasonally and can amount to several decimetres. The WTC model also takes lake altitude into account.

The IC correction is related to the interaction of the electromagnetic wave with free electrons in the upper atmosphere. It is proportional to the Total Electronic Content (TEC) in this layer of the

atmosphere and inversely proportional to the square of the pulse frequency. It has been shown that over lakes, this correction could be erroneous due to land contamination if the measurement is taken close to the shoreline. Therefore, it may be preferable to use the IC derived from the Global Ionospheric Maps (GIM), inferred from the GNSS worldwide network. This correction is provided also in the GDRs.

The Earth tide (ET) and Pole tide (PT) are estimated using models and are provided within the GDRs. PT is related to changes in centrifugal forces, and thus the flattened shape of the Earth, by variations created through fluctuations in the rotational axis of the Earth. The vertical movements of the surface of the Earth associated with this tide are at the centimetre level and are well modelled. ET is linked to astronomical gravitational forces surrounding the Earth, essentially variations in lunar and solar attraction based on their position in the sky. The vertical movements of the surface of the Earth related to the ET are around twenty centimetres. They are relatively well modelled.

LT and SSB are not considered in our calculation but are minor compared to other corrections. For large lakes the measurements may also be biased due to seiche effects that can amount to decimetres in the worst cases. If strong seiches occur at the time of the pass of the satellite, then the data are simply eliminated by filtering process of the outliers.

In addition, some altimeter biases must be taken into account. They depend on several factors, including instrumental electronic bias and error due to geoid bias between several satellite tracks over a given lake. They are calculated prior to the estimation of lake height and are provided for each of the tracks on each of the lakes in a so-called directory file. This methodology is detailed in Cretaux et al. (2011, 2018).

The term *N* in Eqs. 3. 3 and 3. 4 corresponds to the geoid correction that must be applied to each altimetry measurement and is also provided in the GDRs. However, current corrections in the GDRs are not accurate enough at short wavelengths for lakes. Work by Birkett (1995) and Cretaux and Birkett (2006) shows that a specific computation must be performed to correctly account for the slope of the geoid over a distance of several hundred metres, which is much finer than the current geoid model resolution. The “repeat track technique” is used to solve this problem. The geoid slope is recalculated for each of the satellite tracks of the satellite and averaged using all cycles. The result of this calculation is a mean vertical profile along the track which serves as geoid correction (Cretaux et al. 2016).

3.3. LWL quality assessment

Quality assessment is carried out by comparing the retrieved LWL with independent in situ measurements. In situ measurements are available for some lakes on national hydrological services, for example in Canada or Brazil, for selected lakes and reservoirs, or they are released to the project team by the State Hydrological Institute of St Petersburg within the framework of the Hydrolare data centre. A set of approximately 20 lakes regularly serve as validation data sets for altimetry products. The accuracy of lake height measurement depends on several factors: range, orbit and correction errors. Range errors result from surface roughness and quality of the retracking of the altimeter waveform. It is also important to emphasize that the altimeter measurement is an average over the footprint which intrinsically differs from a single point measurement of a ground gauge, and which is furthermore generally done along the coast line.

Performing comparisons over a set of several lakes and reservoirs of varying morphology and from different regions addresses the recurring question of accuracy of altimetry for lakes and reservoirs and its dependency on the size of the water bodies. It is not clear whether a minimum size threshold exists below which the altimeter does not provide valid water levels. Past studies (see Cretaux et al. 2016) have shown that the accuracy of LWL is largely sub-decimetre for large lakes and that lake size influences the quality of the measurement. However, these results also show that accuracy is dependent on the lake environment: mountain lakes or those with ice and snow in winter, as well as large but narrow reservoirs have degraded accuracy. For Lake Onega for example, the RMS accuracy was twice as good when winter months were excluded compared to the year as a whole. Past studies further show that accuracy of LWL products ranges from a few centimetres for very large lakes to a few decimetres for small or narrow lakes (Ričko et al. 2012, Cretaux et al. 2016). However, new missions like the Sentinel-3 constellation present new technological developments from which

improved accuracy may be expected, owing to acquisition in SAR mode. This will have little effect for large lakes (Cretaux et al. 2018) but prompts new assessment to gauge the extent of improvements over smaller lakes.

3.4. LWL references

Birkett C. M. (1995). Contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes, Journal of Geophysical Research, 100, C12, 25, 179-25, 204

Brown G. S. (1977). The Average Impulse Response of a rough surface and its applications, IEEE Trans. Antennas Propag, Vol. 25, pp. 67-74. 10. 1109/TAP. 1977. 1141536

Crétaux J-F. and Birkett C. M. (2006). Lake studies from satellite altimetry, C R Geoscience, 10. 1016/J. crte. 2006. 08. 002

Crétaux J-F. , Calmant S. , Romanovski V. , Perosanz F. , Tashbaeva S. , Bonnefond P. , Moreira D. , Shum C. K. , Nino F. , Bergé-Nguyen M. , Fleury S. , Gegout P. , Abarca Del Rio R. , Maisongrande P. (2011). Absolute Calibration of Jason radar altimeters from GPS kinematic campaigns over Lake Issykkul, Marine Geodesy, 34: 3-4, 291-318, 10. 1080/01490419. 2011. 585110

Cretaux J-F, Abarca Del Rio R. , Berge-Nguyen M. , Arsen A. , Drolon V. , Clos G. , Maisongrande P. 2016. Lake volume monitoring from Space, Survey in geophysics, 37: 269-305. 10. 1007/s10712-016-9362-6

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

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

4. Lake Water Extent (LWE) algorithms

LWE can be expressed as the presence of water (on a map), or as the total areal extent of a waterbody (a single number). It is practically very challenging, if not impossible, to process the high spatial resolution satellite imagery required to generate maps of water presence for hundreds, if not thousands, of lakes every few days. For this reason, the strategy adopted is twofold. First, for each lake, we collect a set of images (SAR and/or optical) spread out over a long period. To choose the images we use the water level time series (calculated using satellite altimetry) to determine when the lake was at low, medium and high level. A relationship (a first or second order polynomial) can then be establish using a set of about 10 to 15 couples of (LWL, LWE) and a simple least square adjustment. Knowing the function LWE=f(LWL) we can thus relate LWL from altimetry to LWE using the hypsometry process. This allows us to achieve a high temporal resolution without overwhelming image processing requirements.

To determine LWE maps to feed into the hypsometry curve, it is determined from contrasts in the optical and/or radar reflectance of water compared to surrounding land. In Lakes_cci, both optical and synthetic aperture radar (SAR) approaches are being investigated to generate LWE. Because the processing chains differ between these methodologies, candidate algorithms are described separately for SAR and optical techniques, in the following sections.

4.1. SAR LWE estimates

4.1.1. Description

The SAR approach is divided into three stages: pre-processing, lake water classification and labelling and results generation. The lake water classification-labelling stage is considered the core of the methodology. Figure 1 shows the processing chain as a block diagram. Some of the processing stages currently have two branches, where alternative and complementary methods are being evaluated to ultimately optimize the retrieval in a unified scheme. More information on the candidate algorithms and how selection takes place is provided in the PVASR-ADP document, whereas the PVP details the validation approach to either methodology. For the time being, and to illustrate the complementarity of the methodologies, we describe the processing chain as one entity, to help identify common steps that can be shared or interchanged in future.

The steps corresponding to each block are described in the following sections.

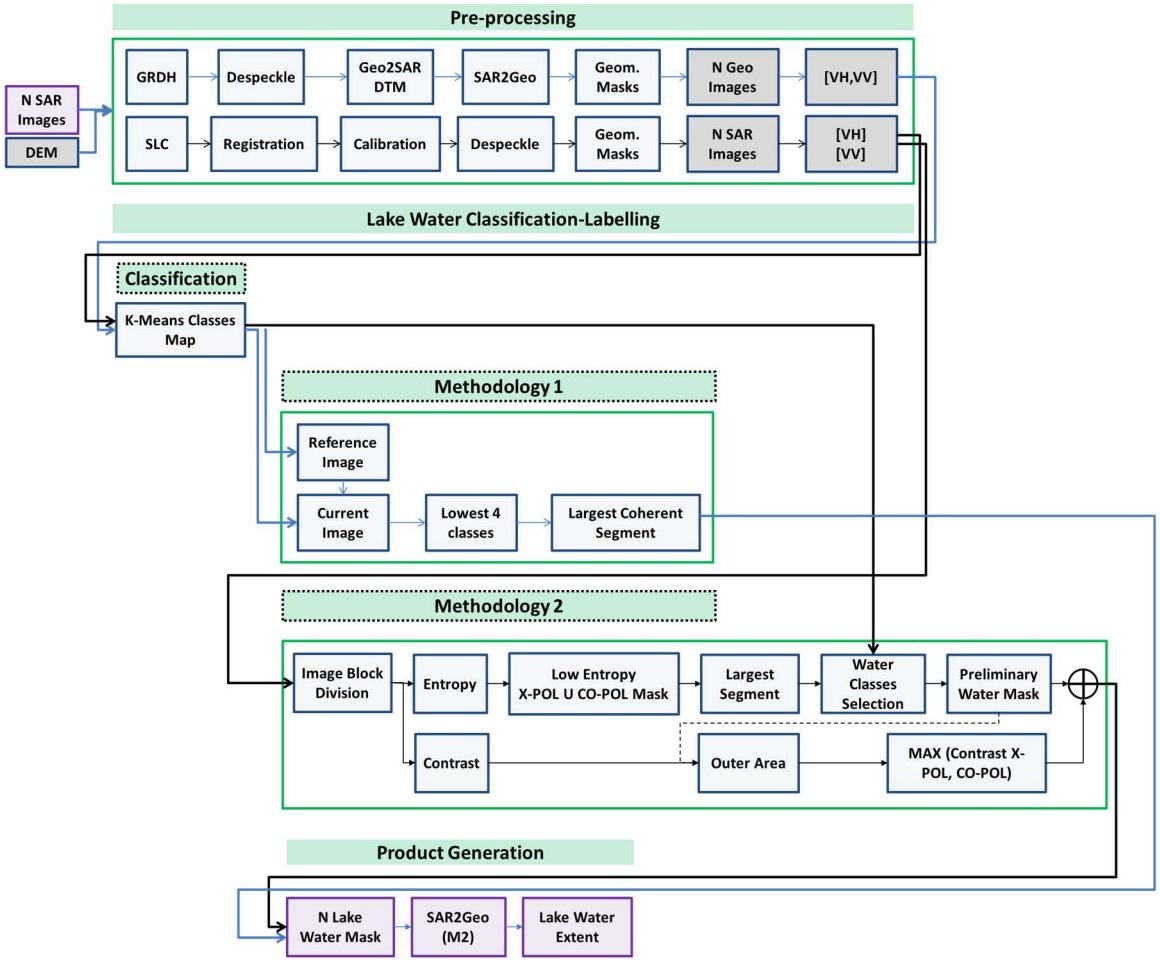

Figure 1: Block diagram of the methodology employed to obtain the Lake Water Extent using SAR images.

4.1.1.1. Pre-processing

The pre-processing accounts for manipulations done on the original input image stack to ease the lake water extent calculation. The output of this chain of processes forms the input for Lake Water classification/labelling. Two SAR images formats are accepted, GRDH (Norce) and SLC (Tre-Altamira), with the following order of processes, respectively:

4.1.1.1.1. Ground Range Detected High resolution (GRDH) input

The pre-processing part of GRDH input imagery in the NORCE algorithm involves performing a precision SAR geocoding on all SAR images using the GSAR (Generic Synthetic Aperture Radar) software (Larsen et al. 2005). We use the calibrated Sentinel-1 IW GRDH product, which is sampled with 10 m pixel spacing on the ellipsoid. The geocoding steps performed are as follows:

- Calibration using an annotated lookup table (from the GRDH product)
- Multilooking, 2 × 2 pixels averaging to suppress speckle
- Mapping the selected Universal Transverse Mercator (UTM)-projected output grid to the radar coordinates using the best available digital elevation model from a variety of sources (e. g. SRTM, http://kartverket. no) and precision orbit vectors available from the European Space Agency (ESA) (https://qc. sentinel1. eo. esa. int)
- Projection of the backscatter product to the output grid, using radar coordinate mapping (from previous step) and cubic interpolation.

The output radar backscatter amplitude images in VV and VH polarization are stored as a GeoTIFF image using the corresponding UTM zone, WGS-84 projection.

4.1.1.1.2. Single Look Complex (SLC) input

The pre-processing part of the TRE-Altamira is a stack of SAR SLC images of the lake area. The default input images are Sentinel-1 IW, but the methodology allows images from any existing sensor to be used. In the case of Sentinel-1, we consider the bursts overlapping the area of interest plus a spatial buffer zone. Under this default mode the following processing steps are applied to both cross and co-polarization images. The images are first co-registered to a common grid in SAR coordinates. The image amplitudes are then inter-calibrated and a Lee despeckling filter is applied. Using the precise orbits and a DEM, the corresponding shadow and layover masks are generated. These are then applied to the images to avoid false positives during the lake area classification/labelling stage. The amplitude images are transformed to decibels and scaled to a 0-255 range based on the significant values histogram. The initially selected area is then further cropped to the final area-of-interest, i. e. the lake and immediate surroundings.

4.1.1.1.3. Lake surface classification and labelling

These steps are in place to identify lake water pixels from the amplitude images obtained after pre-processing. An initial classification takes place, which is common to the two processing lines. Following classification, complementary processing chains are again evaluated, here denoted NORCE (Methodology 1) and TRE-Altamira (Methodology 2). It is important to stress that both methodologies use the same input and yield equivalent results: a binary lake water map that is employed to calculate the water extent. The separate methodologies are in place to define the best combination of processing steps, i. e. yielding the most accurate results.

4.1.1.1.4. Classification

The classification algorithm applies the K-means unsupervised clustering method to segment the image into a defined number of clusters. The method therefore allows for variable cluster (class) centers to which a pixel is assigned, which is preferable since the absolute backscatter values associated with water can vary from acquisition to acquisition due to for example meteorological factors. In these situations, a fixed threshold would not be as effective. The clustering is currently performed using seven classes. By over-segmenting the lake there is the possibility to merge classes to represent the water area as part of the following steps.

The LWE classification algorithm takes the SAR backscatter amplitude images in cross- and co-polarization as inputs together in the NORCE approach whereas the TRE-Altamira approach uses them separately, both with a radar layover/shadow mask and a lake maximum area mask. The purpose of the masks is to reduce the image area where water is to be detected and areas where the radar cannot image the surface, therefore reducing computational requirements.

4.1.1.1.5. Methodology 1 (NORCE)

In the first approach, a multi-temporal correction is applied to the set of classified images in order to replace unclassified pixels, for example those that have missing data, or radar layover/shadow overlapping the lake area. This is done by selecting the latest preceding image in the time series acquired at a different viewing geometry, which represents a reference image. The classification of the reference image is used to assign pixel classes to the pixels in the current image that were unclassified. Once this correction is performed, the water area is mapped by extracting the pixels in the lowest 4 classes. Only the largest coherent water region is retained in the final geocoded lake area. All images in the dataset are manually accepted or rejected based on a visual comparison of the final lake map with the original backscatter images.

4.1.1.1.6. Methodology 2 (Tre-Altamira)

In the second approach, water class selection is achieved using entropy (Zhihui et al. 2012) and coefficient of variation/contrast (Martinis et al. 2009 and 2011, Landuyt et al. , 2019) images. The images are divided into blocks of 100 x 100 pixels. Different values can be tested choosing the optimal in terms of entropy and contrast histogram separation.

First, low entropy pixels are selected by applying the Kernel Density Estimator (Gramaki 2018) to the entropy histogram since water has low entropy and these pixels have the highest probability of representing water. A binary map is generated using these pixels and this is done separately for the cross and co-polarization channels, which are binary added. From this combined map, the largest water surface segment is extracted, discarding outer water bodies. By superimposing the water entropy mask on the classification map, the classes belonging to the largest water surface segment are determined. Using the classes contained within the water surface segment (derived from entropy images) the classes are selected from the original classification map to create a final binary map representing water. Again, only the largest coherent water segment is retained for the final lake region.

Lake borders may not be clearly defined in the amplitude image as humidity or low vegetation on the lake surroundings may present similarly low backscatter to the lake water, resulting in low contrast between surroundings and water. Complementary to the entropy analysis and the lake mask, a lake "margin" area is calculated. Among the blocks containing this border area, the lake pixels of the corresponding cross or co-polarization map presenting the maximum contrast between both are selected and added to the existing inner map obtaining the final lake binary map in SAR coordinates.

4.1.1.2. Product Generation

Up to this point, the lake water maps for each of the processed images are available. Those corresponding to Methodology 1 are already geocoded. In the case of Methodology 2 they are still in SAR coordinates so a geocoding step takes place. From these geocoded lake water maps (in GeoTiff format), the lake area extents are retrieved and exported to a csv file.

4.1.2. Algorithm definition

4.1.2.1. Input data

SAR processing for LWE relies on the following input data sources:

- GRDH and SLC SAR images
- Maximum lake area mask (GeoTIFF)
- Radar layover/shadow mask (GeoTIFF or SAR)

The default images are Sentinel-1 IW (VH, VV) but the algorithm is valid for the most common SAR missions: ERS, ENVISAT-ASAR, Radarsat-1, Radarsat-2, ALOS-PALSAR, PALSAR-2, TerraSAR-X, PAZ and Cosmo-Skymed.

4.1.2.2. Output data

The SAR processing procedures result in the following outputs:

- Binary lake/water map. Water is assigned a value of 1, 0 elsewhere (GeoTIFF)
- LWE time series (one CSV file containing the lake water extent in square kilometres and the corresponding acquisition date for each of the employed images).

4.1.2.3. Mathematical statement

K-means clustering (Jin et al. , 2011) is a variant of the Expectation-Maximization algorithm, and as an unsupervised clustering algorithm requires no manual intervention. It works by separating the data samples into a specified number of clusters of equal variance by minimizing the sum of the square difference between the data samples and cluster mean (often called centroids), given by

$$\Sigma(x_i - \mu_i)^2 \qquad [4.1]$$

The first stage of the algorithm selects initial values for each cluster centroid, for example by taking actual sample values from the dataset. A sample is assigned to the cluster whose centroid is closest to the sample value and the mean of the samples in each cluster is calculated, which becomes the new centroid. The process is iterated by re-assigning samples to the nearest cluster based on the updated value of the cluster centroid and the mean of the samples in each cluster is calculated again. The process eventually converges when the difference between the previous cluster centroid and the updated value (mean of the samples in the cluster) is less than some threshold value (tolerance) or the maximum number of iterations has been reached.

Other mathematical operations that are implemented within the data processing steps include the Lee filter (Lee 1986) and the Kernel Density Estimator (Gramaki, 2018). Two mathematical operators are also used. The first is the coefficient of variation or contrast *Cv*:

$$Cv = \mu x / \sigma x \qquad [4.2]$$

where μ_x is the mean value of the amplitude pixels *x* over an image block and σ_x the standard deviation of the amplitude pixels *x* over the same image block. The second is the entropy *H*:

$$H = -\Sigma k\, pk\, log2(pk) \qquad [4.3]$$

where *k* is the number of quantified amplitude levels and p_k is the probability associated with quantified level *k*.

4.1.2.4. Quality assessment

A pixel accuracy figure is not provided as part of the result. This is mainly due to the use of the K-Means classification method and the entropy and contrast operators. The nature of these methods does not allow for the error of water pixel misclassification to be calculated. Some general indicators could be calculated by assessing for example the amplitude, entropy and contrast histograms. These reliability indicators could be generated for the images present in the data stack. At the present time this remains an open issue under discussion. In order to calculate the accuracy of the lake water maps, ground truth data are required.

4.2. Optical LWE estimates

For optical LWE estimation we investigate two unsupervised approaches, Otsu's method (Otsu 1979) and KMeans, and two supervised methods, SVM and Random Forest, applied to High resolution sensors with a SWIR channel, i. e. Landsat 5, Landsat 8, Sentinel-2. The Kmeans approach is identical to its use on SAR imagery and is not repeated here. The following sections detail the OTSU, SVM and Random Forest approaches.

4.2.1. Algorithm 1: Adaptive sampling for SVM classification

4.2.1.1. Description

An accurate and precise sampling is essential for training an efficient supervised classification. One of the challenges regarding water surfaces classification is the dynamic nature of such systems; seasonality and meteorological conditions have great impacts on the open water surface. As a result, if the classification algorithm is to be trained on each image, training samples cannot be static and have to be adapted to each image from the time series. On the other hand, in the context of a large scale water surface extraction, creating a high quality sampling would require either a very time consuming photo-interpretation work, or an already high quality algorithm capable of distinguishing water and non-water pixels. The use of an *a priori* water surface database could help by reducing the processing area to where water is most likely present. The method used here for creating these automatic samples is a combination of the said a priori database and the radiometric information obtained from each image of the time series.

The most relevant a priori water surface database would be the Global Surface Water (GSW) produced by the Joint Research Centre (Pekel et al. , 2016). Apart from being a state-of-the-art global dataset, the accuracies presented in the associated paper are high and the dataset has been used and cited in high quality scientific work (Weiss et al. 2018, Yamazaki et al. 2017). These factors motivated the use of GSW as a priori inland water surfaces dataset.

As illustrated by Figure 2, each satellite image of the time series is processed individually, extracting “water” and “non-water” samples. A machine learning algorithm then exploits both the image and the training dataset to perform a supervised classification of water as LWE. The Support Vector Machine (SVM) is widely used and considered one of the most suitable machine learning algorithms for water detection, showing high and stable performance (Martinis et al. , 2017).

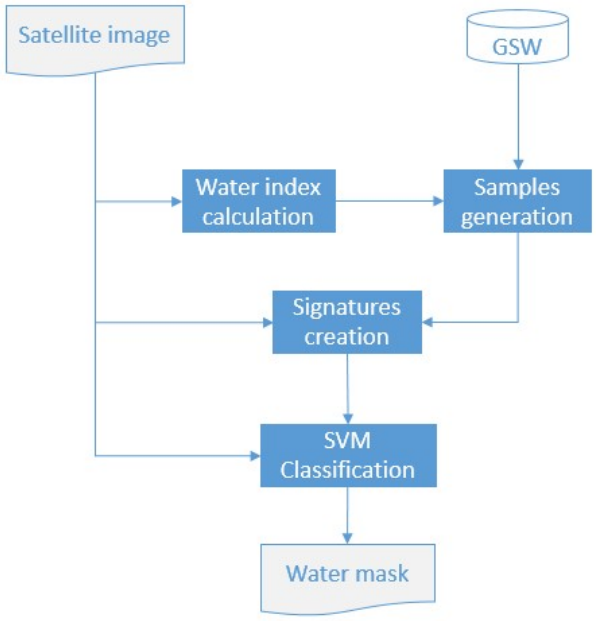

Figure 2: Algorithm 1 workflow

Water masks derived from the time series are then cropped with the lake area of interest. Water areas are calculated and reported in a table in order to visualise LWE curve in time

4.2.1.2. SVM Algorithm definition

4.2.1.2.1. Input Data

The optical LWE estimation relies on the following input data:

- Sentinel-2 L1C/L2A or Landsat-8 time series
- Global Surface Water database
- Lake area of interest in vector format

4.2.1.2.2. Output Data

The output data from optical LWE estimates are analogous to the SAR outputs, noting that instead of a raster map a vector is produced:

- Lake Water Extent derived from each EO image, in vector format
- Spreadsheet (Excel format) summarizing for a given lake, all water surfaces by acquisition date

4.2.1.2.3. Mathematical statement

The LibSVM library, implemented in the Orfeo Toolbox, is used with a linear kernel type. SVM is a powerful machine learning procedure created by Vapnik (1995) and is useful because of applicability in a large range of applications (Vapnik 1998, Cherkassky and Mulier 1998). The power of SVMs relies on the kernel function which can adapt dimension function compared to the mapping input (Cortes and Vapnik 1995). In addition, the kernel can adapt itself into a matrix kernel instead of high dimensional space if the classes are not separable in the input space.

4.2.1.3. SVM Accuracy

Validation of the algorithm has been performed in the ESA-ASAPTERRA project (Advancing SAR and Optical Methods for Rapid Mapping). The overall accuracy of LWE produced by the same algorithm over the Poyang Lake test site was 98. 65%. A reference classification produced by visual interpretation was used as ground truth in this evaluation.

4.2.2. Algorithm 2: MNDWI thresholding using the OTSU method

4.2.2.1. Description

Thresholding is one the most basic classification methods. From a greyscale image, a value is computed as the limit between two or more classes. The OTSU method selects an optimal threshold by reducing the within-class variance, or by maximizing the between-class variance. This method is easy to implement and computationally efficient. There is an alternative method associating to the OTSU threshold an edge detection operator, such as a Canny Filter.

It is anticipated that applying the OTSU algorithm to the area corresponding to water following a so-called ‘water-index’ will increase the performance of the process compared to single raw band thresholding. The Modified Normalized Difference Water Index, MNDWI has been demonstrated to be one of the most efficient indices when detecting water for satellite images.

Furthermore, focusing the study on an *a priori* known region where the water class is expected makes ideal conditions for an OTSU application for water identification. For this reason, it is important to use good reference polygons, delimiting the areas of interest, here referred to as the Lake Area of Interest. This area delineates where the largest observed water extent is expected to occur and may include some of the surrounding land area.

The raw water masks derived from the satellite time series are then cropped to the lake area of interest. Water areal extent is calculated and reported in a table to synthesize LWE dynamics over time.

4.2.2.2. Algorithm definition

4.2.2.2.1. Input Data

The input data for the Optical thresholding approach include:

- Sentinel-2 L1C/L2A, Landsat- 8 as well as Landsat-5 time series obtained from ESA Science Hub, USGS portal or Google Earth Engine.
- Lake area of interest in vector format: the maximum water extent masks derived from ESA Landcover cci.

4.2.2.2.2. Output Data

Lake Water Extents derived using the optical thresholding approach will be provided in a vector format (shapefiles). In addition, a spreadsheet summarizes the water extent per water surface within a given time series, separated by acquisition date.

4.2.2.2.3. Mathematical statement

The OTSU method was first expressed by Otsu (1979). If we consider that an image is composed of grey level values. Each pixel n_i in the image has a grey level value i. The total number of pixels is:

$$N = \sum_{i=0}^{L} n_i \qquad [4.4]$$

The grey-level probability distribution is expressed as:

$$p_i = \frac{n_i}{N}, \text{ with } p_i \geq 0, \text{ and } \sum_{i=0}^{L} p_i = 1 \qquad [4.5]$$

Let's suppose that the image is divided in two classes using the grey values: class A with grey levels from *1* to *t*, and class B from *t+1* to *L*. A and B probabilities are:

$$\omega_A = \sum_{i=0}^{t} p_i,\ \omega_B = \sum_{i=t+1}^{L} p_i = 1 - p_A \qquad [4.6]$$

Mean grey levels in each class are:

$$\mu_A = \sum_{i=0}^{t} \frac{ip_i}{\omega_A},\ \mu_B = \sum_{i=t+1}^{L} \frac{ip_i}{\omega_B} \qquad [4.7]$$

And the total mean grey value is:

$$\mu_T = \sum_{i=0}^{L} ip_i \qquad [4.8]$$

Class variances are given by:

$$\sigma_A = \sum_{i=0}^{t} (i - \mu_A)^2 p_i / \omega_A \qquad [4.9]$$

$$\sigma_B = \sum_{i=t+1}^{L} (i - \mu_B)^2 p_i / \omega_B \qquad [4.10]$$

The Otsu method tends to maximize the between-class variance to make the two classes as distinct as possible.

$$\sigma_{BC}^2 = \omega_A \omega_B (\mu_A - \mu_B)^2 \qquad [4.11]$$

The calculated threshold is then the value that maximizes this between-class variance.

Although many previous studies work globally by exploiting a single threshold value, it is important to note that for a given lake, the threshold value must be calculated for each image. In consequence, the values of the threshold vary according to the site and the date (see examples in Figure 3, Figure 4 and Figure 5).

Dates	Threshold
30/06/2016	0,065029736
30/07/2016	0,065108026
16/08/2016	0,104964603
22/10/2016	0,325102834
30/06/2017	0,084940876
07/07/2017	0,074881105
25/07/2017	0,095057017
05/09/2017	0,074863375
07/09/2017	0,135050455
27/09/2017	0,125069731
30/09/2017	0,144615386
08/10/2017	0,255138901
13/10/2017	0,244853664
23/10/2017	0,584918515
17/06/2018	0,03478363
02/07/2018	0,064810589
27/07/2018	0,065254784
28/08/2018	0,114812733
05/09/2018	0,105352232
17/09/2018	0,155139037

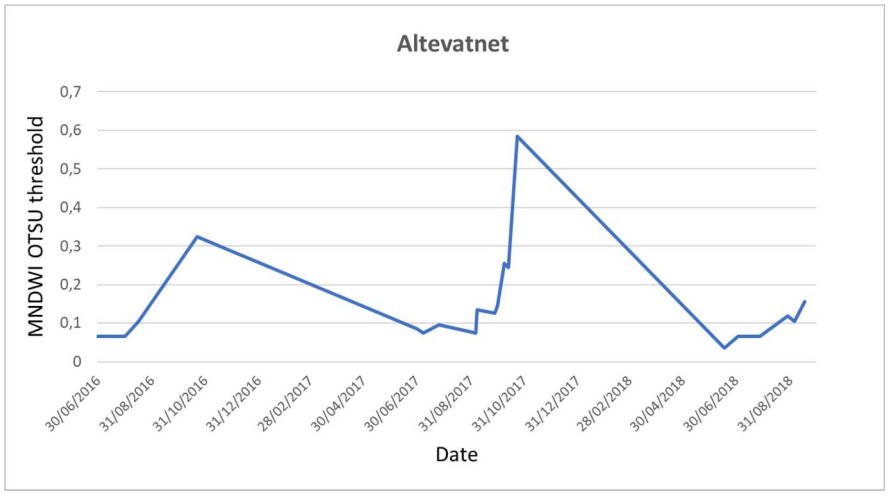

Figure 3: Variations of the threshold values over the S2 timeseries for the Altevatnet lake

Colhue	
Date	Threshold
09/01/2016	0,205033669
08/04/2016	0,04490473
02/09/2016	0,095112755
04/12/2016	0,045224937
21/03/2017	0,015560646
03/04/2017	0,344966733
07/07/2017	0,365063266
04/11/2017	0,31503865
11/11/2017	0,305017148
06/12/2017	0,315100093

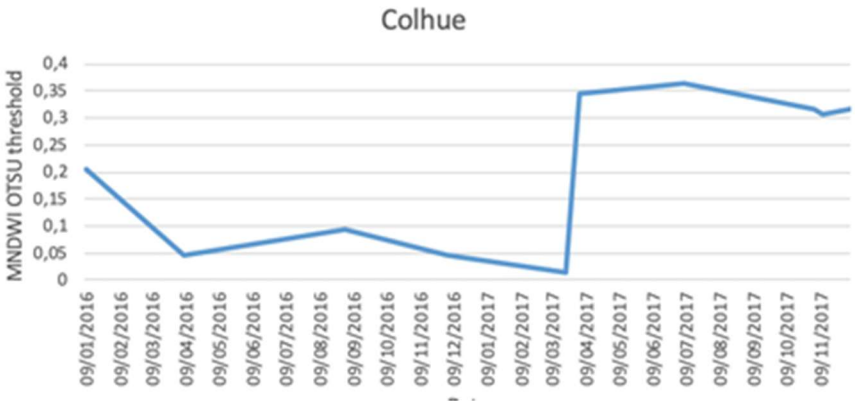

Figure 4: Variations of the threshold values over the S2 timeseries for the Colhue lake

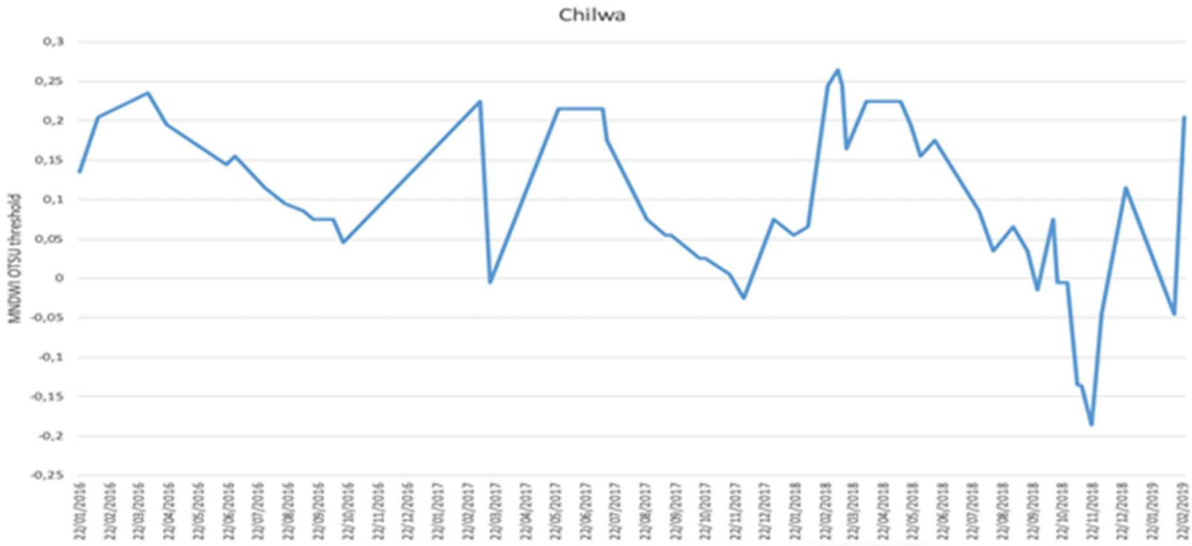

Figure 5: Variations of the threshold values over the S2 timeseries for the Chilwa lake

4.2.2.3. Alternative: CANNY OTSU approach

As already indicated there is an alternative of the OTSU threshold, it is the CANNY OTSU (Donchyts et al., 2016; Yang et al., 2020) for deriving water surface limits.

The Canny algorithm is an operator widely used for edge detection in images (Fang et al, 2009). It is here introduced as a way to focus the OTSU calculation on regions with strongly defined interfaces. In the case of water surface detection, the interface between a water surface and the surrounding land acts as such a high contrast region. This procedure has the benefit to reduce drastically the amount of data to be processed by the OTSU algorithm and focuses the statistics on water discrimination, especially when water is under-represented in the scene. The workflow is defined as follow:

1. Calculate the MNDWI.
2. Calibrate and apply the Canny filter on the MNDWI layer. Highly contrasted interfaces area will be highlighted.
3. Apply a buffer on the detected interfaces to include surrounding areas (both water and land).
4. Apply the OTSU algorithm on the MNDWI layer, within the regions defined by the buffered regions.
5. Threshold the MNDWI layer using the calculated OTSU value.

4.2.3. Algorithm 3: Random Forest (RF)

4.2.3.1. Description

This approach is based on machine learning pipeline using Random Forest. Training samples are derived from the GSW database.

4.2.3.1.1. Pre-processing

For each image in the time series a deep stack is produced. This stack is built with a selection of image bands and relevant indices (Figure 6). The bands and indices were defined based on previous studies on water bodies mapping led by ICube-Sertit.

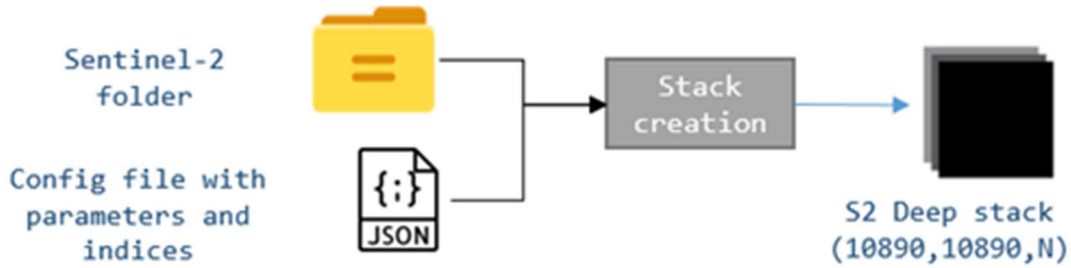

Figure 6: Illustration of the Sentinel-2 stack creation

4.2.3.1.2. Training

The Random Forest is proven to be a robust algorithm for classifications task. In order to build an efficient machine learning model, a collection of Sentinel-2 tiles from the time series with the correspondent labels from the GSW database is selected for the training (Figure 7).

The model will learn from multiple dates representing different seasons and Lake Surface dynamics which enables for model generalization (temporal and spatial) nor from each image individually (adaptative sampling by tile instead of generic samples derived from the multi-temporal stack).

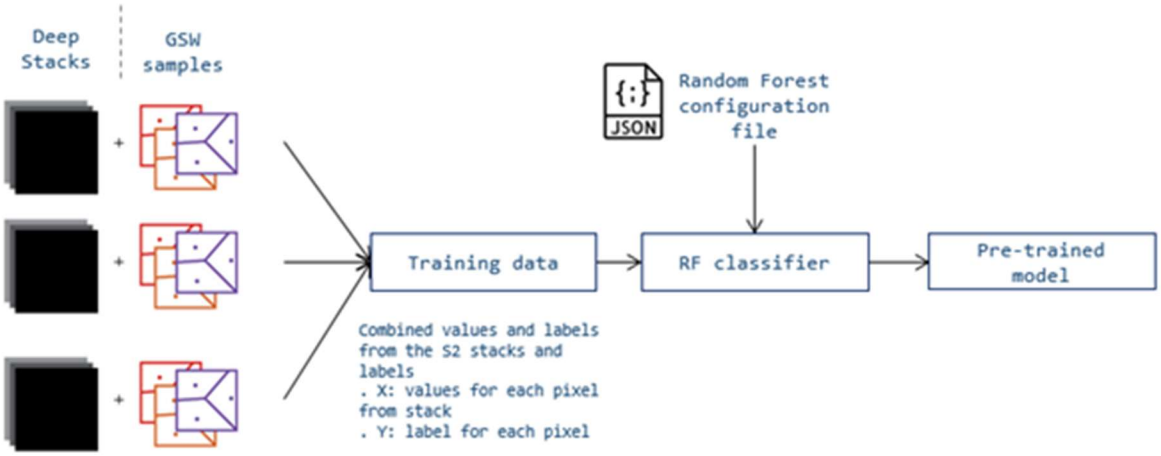

Figure 7: Illustration of the RF pipeline

4.2.3.1.3. Prediction

The trained model pipeline is saved and applied to the complete time series (Figure 8).

Figure 8: Illustration of the prediction from pre-trained models on the entire Sentinel-2 time serie

4.2.3.2. Algorithm definition

4.2.3.2.1. Input Data

Input data to the Random Forest method for optical LWE estimation include:

- Sentinel-2 data L1C/L2A
- Lake area of interest in vector format
- GSW samples

4.2.3.2.2. Output Data

The Lake Water Extents derived using the Random Forest method for optical LWE estimates will be provided in a vector format (shapefiles).

4.2.3.2.3. Mathematical statement

The random forest algorithm is part of the ensemble classifiers family. This technique is popular within the remote sensing community due to the accuracy of its classifications.

It was first introduced by Tin Kam Ho (1995). The method was extended by Breiman (2001) highlighting important concepts like bagging (bootstrap aggregating) and Out-of-bag error (OBB).

The bootstrap sampling allows for the decorrelation of the trees and therefore improve the results and the robustness of the model. The OBB error can be used to determine the importance of features used for the training.

An important concept for random forest is bagging or bootstrap aggregation. For a given vector *x* to predict, *T* number of trees and *m* the number of variables for each tree:

1. Select a random sub-sample of size *m* with replacement from the data

2. Train the tree *Ti* on the bootstrapped sample

The final estimate of the prediction for is an average of estimation from each tree *Ti*

$$f(x) = \frac{1}{T}\sum_{k=1}^{T} T_i(x) \qquad [4.12]$$

The RF model uses bagging to form an ensemble of decision trees. It only uses a sub-sample of the training data and a selection of features for each tree.

The random forest algorithm adds an extra step to the bagging algorithm, which is feature selection (feature bootstrapping)

For a given vector *x* to predict, *T* number of trees and *m* the number of variables for each tree:

1. Select a random sub-sample of size *m* with replacement from the data

2. Train the tree *Ti* on the bootstrapped sample with a selection of *n* random features from the dataset (main difference between bagging and RF)

The prediction of *x* is similar to the bagging process, an average (or vote) over all the trees.

The values of *T*, *m* and *n* are part of the hyper parameters that need to be fine-tuned in order to optimize the RF results.

4.2.3.3. Quality assessment

The data used to train the random forest model for the lake extent extraction is divided into three sub-datasets (60, 20 and 20):

- Training: data used for training only (approx. 60%)
- Development: samples used to fine-tune the model and estimate good hyper parameters (approx. 20%)
- Validation: samples used for the model training validation (approx. 20%)

This approach allows the estimation of the model performance. However, these indicators cannot evaluate how well the model will perform on the entire time series. A more thorough inspection of the models results, compared to others, is needed to evaluate the final accuracy of this method.

4.2.4. Algorithm 4: simple threshold approach based on MNDWI index

4.2.4.1. Description

The final optical water detection algorithm is based on threshold definition using the Normalized Difference Water Index (NDWI) which was first introduced by McFeeters (1996) and expressed as the band ratio:

$$NDWI = (Green-NIR)/(Green+NIR) \quad [4.\ 13]$$

The modified Normalized Difference Water Index is:

$$MNDWI = (Green-SWIR)/(Green+SWIR) \quad [4.\ 14]$$

The water extent (in km^2) is subsequently calculated using the sum of individual pixel classified as water pixel within the ROI.

Processing steps are as follows. Starting with a number of pre-processing steps:

- For each image a Region Of Interest (ROI) is defined including the target lake.
- All pixels that are already detected as cloud pixel in the input product are excluded.
- A geolocalised grid at native resolution is produced with the NDWI index.

Images are subsequently individually processed. A cloud index is provided with sentinel-2 L2 images sourced from the THEIA database (https://labo. obs-mip. fr/multitemp/quantitative-comparison-of-cloud-masks-from-maccsmaja-sen2cor-and-geosys-hand-made/) and corresponding pixels are excluded. However, cloud pixels located within the permanent lake interior (at a distance farther than 10 km from the shoreline) are converted to water pixels. Only images with cloud cover lower than 5% of the total image are selected.

For Sentinel-2: a pixel is classified as water when NDWI is higher than 0. 1, or NDWI equal to 1 or NDWI equal to -1 and when at the same time Band 4 reflectance is lower than 400.

For Landsat 5 a pixel is classified as water when NDWI is higher than 0. 02 and for landsat 8 when NDWI is higher than 0. 1.

4.2.4.1.1. Input Data

Input data to the MNDWI method for optical LWE estimation include:

- Sentinel-2 data L1C/L2A, landsat 5, and 8
- Polygon of the Region of interest
- Maximum Water Extent polygons from an external database (GLWD)

4.2.4.1.2. Output Data

A file of Lake Water Extents derived from each EO image, with associated lake water level is produced (in text format) for a set of 10-15 different dates, serving as input file for the inversion of the hypsometry of each lake. In order to calculate the hypsometry coefficient for each of the lake in the Database, we need to produce a set of vector (LWL, LWE) for about 10 – 15 images collected and processed at different dates corresponding to different water level. These value of (LWL, LWE) is therefore then used as input for the hypsometry inversion.

4.2.4.2. Quality assessment

The LWE is calculated by a combination of hypsometry and water height measured from satellite altimetry. The hypsometry is represented by a linear or a quadratic polynomial, depending on lake morphology, and is estimated using a set of 10-15 vectors (LWL, LWE) for a selection of dates at different levels. A way to estimate the associated uncertainty is to calculate the RMS of the differences between the theoretical (calculated from hypsometry's coefficient) and the measured LWE (directly from satellite imagery) for the dates that have been chosen to build the hypsometry. A first test done over a set of 10 lakes have shown that the estimated RMS of LWE was lower than 2% of the total extent for each of the 10 lakes

4.2.5. Final determination of the LWE: approach based on hypsometry

4.2.5.1. Input data

To calculate the LWE we use the vector (LWL, LWE) calculated previously (see above) to estimate, by simple least square adjustment, the coefficients representing the 2d-polynomial subsequently used as hypsometry curve.

The satellite altimetry height (LWL) of lake water surface is then used together with the hypsometry coefficient to determine the LWE variable.

4.2.5.2. Method

We first use the water level time series inferred from satellite altimetry and released as LWL products, in order to determine some key periods when the lake was at extreme heights and some intermediate values. Then, satellite images are selected at the same dates for the water mask detection using approaches described above. It is not realistic to determine water extent of a lake for each measurement of its water level, especially when a lake is too large and is not covered by only one image. We select between 10 and 15 images at different dates and calculate the hypsometry relationship (dLWE/dLWL) which is then applied to determine surface extent of the lakes each time a water level is calculated using satellite altimetry. The hypsometry is expressed as a polynomial of degree 1, 2 or 3 depending on the linearity of the couples of water level and surface extent of the lake. In such processing, we do not need to process a large amount of satellite images and this is a practical way to produce lake surface extent together with lake water level. We don't extrapolate the water surface for water height outside of the maximum and minimum values used to determine polynomial coefficients. Therefore we try to collect satellite images from different sensors (radar, optical) as close as possible to the maximum and minimum level observed from satellite altimetry.

The method is applied and described in Cretaux et al. , (2015, 2016). A final estimation of accuracy is calculated as expressed in 4. 2. 4. 2 with all vectors used to calculate the hypsometry coefficients.

4.2.5.3. Output data

LWE time series are produced for the time span corresponding to the LWL products in the form of text files.

4.2.5.4. References

Breiman, L. , 2001. Random forests. Machine learning, 45(1), pp. 5-32.

Chang, C-C. , Lin, C-J. (2011). LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27.

Cherkassky, V. , Mulier, F. M. (1998). Learning from Data: Concepts, Theory, and Method. John Wiley & Sons.

Cortes, C. , Vapnik, V. (1995). Support-vector networks. Machine Learning 20, 273-297.

Cretaux J-F. , Biancamaria S. , Arsen A. , Bergé-Nguyen M. , and Becker M. , Global surveys of reservoirs and lakes from satellites and regional application to the Syrdarya river basin, *Environmental Research Letter*, 10, 1, AN: 015002, 2015, **DOI:** 10. 1088/1748-9326/10/1/015002

Cretaux J-F, Abarca Del Rio R, Berge-Nguyen M, Arsen A, Drolon V, Clos G, Maisongrande P, Lake volume monitoring from Space, *Survey in geophysics*, 37: 269-305, doi 10. 1007/s10712-016-9362-6 , 2016

Donchyts, G. , Schellekens, J. , Winsemius, H. , Eisemann, E. , van de Giesen, N. , 2016. A 30 m Resolution Surface Water Mask Including Estimation of Positional and Thematic Differences Using Landsat 8, SRTM and OpenStreetMap: A Case Study in the Murray-Darling Basin, Australia. Remote Sens. 8, 386. doi:10. 3390/rs8050386

Fang Mei, YUE GuangXue and YU QuingCang, 2009 : The study of an applicaiotn of Otsu method in a Canny operator. Proceedings Int. Symp. On information Porcessing, Huangshan, PR China, 21-23 August, 2009, 109-112. http://citeseerx. ist. psu. edu/viewdoc/download?doi=10. 1. 1. 402. 5899&rep=rep1&type=pdf

Feyisa, G. , Meilby, H. , Fensholt, R. , & Proud, S. (2013). Automated Water Extraction Index: A new technique for surface water using Landsat imagery. ELSEVIER, Remote Sensing of Environment.

Fisher, A. , Flood, N. , & Danaher, T. (2015). Comparing Landsat water index methods for automated water classification in eastern australia. ELSEVIER, Remote Sensing of Environment.

Gramacki, A. (2018). Nonparametric Kernel Density Estimation and Its Computational Aspects, Studies in Big Data Vol 37, Springer.

Haouet, S. (2014). Développement d'un outil d'extraction de cibles "simples". Application à la cartographie des corps en eau et des séries temporelles d'images satellite HR et THR.

Ho, T. K. , 1995, August. Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition (Vol. 1, pp. 278-282). IEEE.

Jin X. , Han J. (2011). K-Means Clustering, Encyclopedia of Machine Learning, Springer.

Landuyt, L. , Van Wesemael, A. , Schumann, G. J. , Hostache, R. , Verhoest, N. E. C. and Van Coillie, F. M. B. (2019). Flood Mapping Based on Synthetic Aperture Radar: An Assessment of Established Approaches, IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 2, 722-739.

Larsen, Y. ; Engen, G. ; Lauknes, T. R. ; Malnes, E. ; Høgda, K. A. (2005). A generic differential interferometric SAR processing system, with applications to land subsidence and snow-water equivalent retrieval. Fringe ATSR Workshop 2005, Frascati, Italy, 28 November – December 6.

Lee, J. S. (1986). Speckle suppression and analysis for synthetic aperture radar images, Opt. Eng. 25(5), 636-643.

Li, Z. , Chen, X. , Luo, P. and Tian, Y. (2012). Water area segmentation of the Yangcheng Lake with SAR data based on improved 2D maximum entropy and genetic algorithm, 2012 Second International Workshop on Earth Observation and Remote Sensing Applications, Shanghai, China, 263-267.

Martinis, S. , Twele, A. and Voigt, S. (2009). Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data, Nat. Hazards Earth Syst. Sci. , 9, 303–314.

Martinis, S. and R. Christoph, R. (2015). Backscatter Analysis Using Multi-Temporal and Multi-Frequency SAR Data in the Context of Flood Mapping at River Saale, Germany, Remote Sens. 7(6), 7732-7752.

Martinis, S. , Haouet, S. (2017). ASAPTERRA – Advancing SAR and Optical Methods for Rapid Mapping

McFeeters S. K. (1996) The use of Normalized Difference Water Index (NDWI) in the delineation of open water features, International Journal of Remote Sensing, 17(7):1425–1432

Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 9(1), 62-66.

Pekel, J. F. , Cottam, A. , Gorelick, N. , Belward, A. (2016). High-resolution mapping of global surface water and its long-term changes. Nature 540, 418-422. doi:10. 1038/nature20584

Vala, H. J. , & Baxi, A. (2013). A review on Otsu image segmentation algorithm. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 2(2), 387-389.

Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. Springer-Verlag.

Vapnik, V. N. (1998). Statistical Learning Theory. Wiley.

Weiss, D. J. , Nelson, A. , Gibson, H. S. , Temperley, W. , Peedell, S. , Lieber, A. , . . . & Mappin, B. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, *553*(7688), 333.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing, 27(14), 3025-3033.

Yamazaki, D. , Ikeshima, D. , Tawatari, R. , Yamaguchi, T. , O'Loughlin, F. , Neal, J. C. , . . . & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. Geophysical Research Letters, 44(11), 5844-5853.

Yang Xiucheng, Qiming Qi, Hervé Yésou, Thomas Ledauphin, Mathieu Koehl, Pierre Grussenmeyer, 2020: Monthly estimation of the surface water extent in metropolitan France at a 10-m resolution using Sentinel-2 data. Rem Sens Enviroment, 244, https://doi. org/10. 1016/j. rse. 2020. 111803.

Yang, Y. , Liu, Y. , Zhou, M. , Zhang, S. , Zhan, W. , Sun, C. , & Duan, Y. (2015). Landsat 8 OLI image based terrestrial water extraction from heterogeneous backgrounds using a reflectance homogenization approach. ELSEVIER, Remote Sensing Environment.

5. Lake Surface Water Temperature (LSWT) algorithms

5.1. Description

This section describes the algorithm for Lake Surface Water Temperature (LSWT) production chain. For consistency with data produced in heritage projects, the version (v5. 0) described here refers to the scientific versioning of the climate data record in a manner already familiar to users of data from these projects, with GloboLakes offering the product generation as v4. 0 (Carrea et al. 2019) and the first version from Lake CCI being termed v4. 1.

The algorithm described here is consistent with the v4. 1 algorithm described in ATBD v1. 0, with the exception of the following advances:

- V5. 0 will include data from the Moderate-resolution Imaging Spectroradiometer (MODIS).
- Prior information for the retrieval process is updated, using ERA-5 (the most recent numerical weather prediction, NWP, re-analysis from ECMWF).

The inclusion of the MODIS and ERA-5 datastream will significant change the LSWT record, making v5. 0 a major version increment over v4. 0/4. 1, because use of new re-analysis data necessitates reprocessing of the complete time series.

The LSWT climate data record (CDR) is based on Along-Track Scanning Radiometer (ATSR) AVHRR series and MODIS Terra instruments. The scope of this algorithm description applies to the following steps:

(1) identification of water-only pixels for valid retrieval,

(2) the LSWT retrieval itself,

(3) estimating the daily average LSWT from the instantaneous skin observation,

(4) assigning a pixel quality level,

(5) remapping the data to a regular global grid,

(6) cross-sensor LSWT harmonization.

5.2. Algorithm definition

The LSWT processing sequence is described in Figure 9. The individual processing steps are described in turn below.

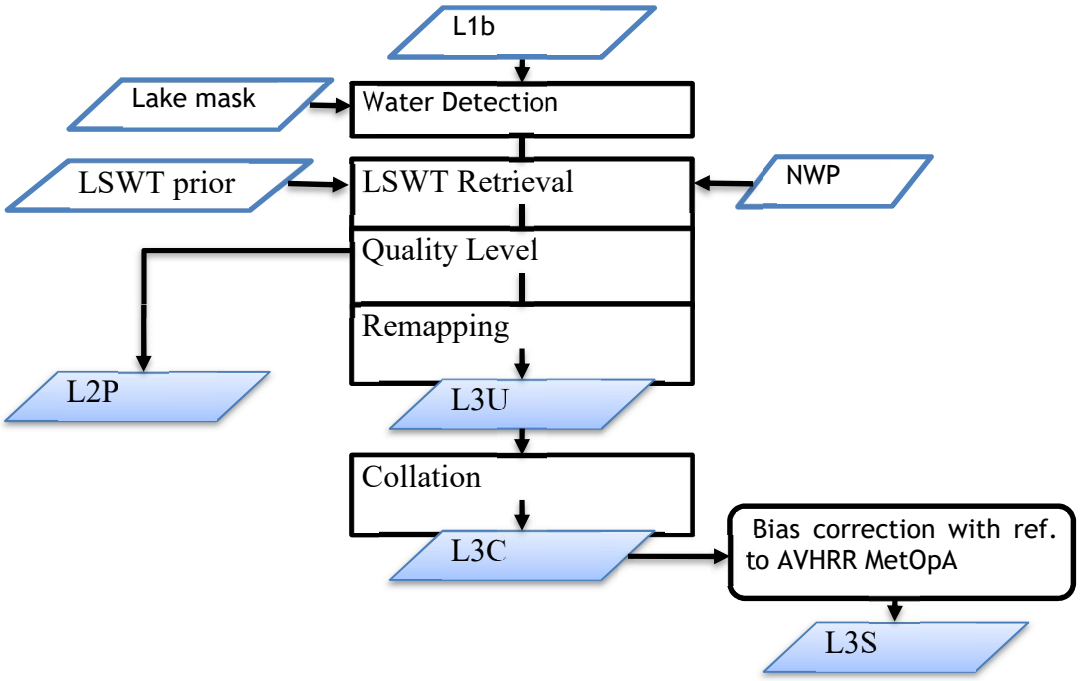

Figure 9: Major steps in LSWT processing

5.2.1. Water detection

Water detection is applied to potential inland water pixels. It operates by calculating a score against several metrics, derived from reflectance channels available. For this reason, LSWT products are in this version obtained only from daytime scenes.

The score for a given metric is defined as:

$$s = \begin{cases} 0 & if X \le t_0 \\ \frac{X - t_0}{t_1 - t_0} & if t_0 < X < t_1 \\ 1 & if X \ge t_1 \end{cases} \qquad [5.1]$$

The score is a linear ramp between 0 and 1, similar to well-known concepts of "fuzzy logic" (the scores are like probabilities).

The values of the terms in Eq. 5. 1 are given in Table 1. The first three metrics based on red, near infra-red and shortwave infra-red bands use the expectation that reflection from a cloud-affected pixel exceeds that from a clear view of a lake, with values appropriate to different wavelengths.

The MNDWI is the modified normalised difference water index. The NDVI is a normalised difference vegetation index. The setting of the thresholds was done (within GloboLakes) using AATSR imagery tuned to a probability of cloud image derived from MERIS 300 m imagery. The tuning of thresholds was done one-at-time across metrics, maximising the posterior probability that a certain pixel is cloudy or cloud free.

Table 1: Thresholds test for water detection

X, metric	Score definition	Thresholds
R_{670}	$s = \begin{cases} 0 & \text{if } X \ge t_0 \\ \frac{X - t_0}{t_1 - t_0} & \text{if } t_1 < X < t_0 \\ 1 & \text{if } X \le t_1 \end{cases}$	t_0 = 0. 132 t_1 = 0. 032

$\boldsymbol{X}$, metric	Score definition	Thresholds
R_{870}	$\boldsymbol{s} = \begin{cases} 0 & \text{if } X \geq t_0 \\ \dfrac{X - t_0}{t_1 - t_0} & \text{if } t_1 < X < t_0 \\ 1 & \text{if } X \leq t_1 \end{cases}$	t_0 = 0. 097 t_1 = 0. 022
R_{1600}	$\boldsymbol{s} = \begin{cases} 0 & \text{if } X \geq t_0 \\ \dfrac{X - t_0}{t_1 - t_0} & \text{if } t_1 < X < t_0 \\ 1 & \text{if } X \leq t_1 \end{cases}$	t_0 = 0. 048 t_1 = 0. 012
MNDWI	$\boldsymbol{s} = \begin{cases} 0 & \text{if } X \leq t_0 \\ \dfrac{X - t_0}{t_1 - t_0} & \text{if } t_0 < X < t_1 \\ 1 & \text{if } X \geq t_1 \end{cases}$	t_0 = 0. 295 t_1 = 0. 515
NDVI	$\boldsymbol{s} = \begin{cases} 0 & \text{if } X \geq t_0 \\ \dfrac{X - t_0}{t_1 - t_0} & \text{if } t_1 < X < t_0 \\ 1 & \text{if } X \leq t_1 \end{cases}$	t_0 = -0. 085 t_1 = -0. 245
MNDWI-NDVI	$\boldsymbol{s} = \begin{cases} 0 & \text{if } X \leq t_0 \\ \dfrac{X - t_0}{t_1 - t_0} & \text{if } t_0 < X < t_1 \\ 1 & \text{if } X \geq t_1 \end{cases}$	t_0 = 0. 375 t_1 = 0. 685

The first three metrics use the expectation that reflection from a cloud-affected pixel exceeds that from a clear view of a lake, with values appropriate to different wavelengths. The MNDWI is the modified normalised difference water index. The NDVI is a normalised difference vegetation index. The setting of the thresholds was done (within GloboLakes) using AATSR imagery tuned to a probability of cloud image derived from MERIS 300m imagery. The tuning of thresholds was done one-at-time across metrics, maximising the posterior probability that a certain pixel is cloudy or cloud free.

The same method will apply for AATSR, AVHRR and MODIS, although the parameter values will be re-evaluated for MODIS. Additionally, to maximise consistency across the Lake CCI project, we expect to use full resolution information from the MODIS LIC processing chain to exclude ice-covered pixels consistently.

5.2.2. LSWT retrieval

The retrieval scheme is the optimal estimation (OE) scheme (MacCallum et al. 2012) for all sensors, the equation for which is:

$$\hat{\boldsymbol{x}} = \boldsymbol{x_a} + \boldsymbol{G}\big(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x_a})\big) \qquad [5.2]$$

The retrieved state is the prior state plus an increment of $\mathbf{G}\big(\boldsymbol{y} - F(\boldsymbol{x}_a)\big)$. F is the forward model, i. e. the radiative transfer model (RTM) run for the prior re-analysis data and prior LSWT. The matrix $\mathbf{K}$ expresses how the observations change for departures from the prior state x_a, i. e. , it is a matrix where a given row contains the partial derivatives of the BT in a particular channel with respect to each element of the state vector in turn. The partial derivatives are the tangent linear outputs from the forward model $\mathbf{F}$. $\mathbf{S}_\varepsilon$ is the error covariance of the differences between the model and observed BTs. This error covariance matrix is the sum of the radiometric error covariance in the observations ($\mathbf{S}_o$) and estimated error covariance of the forward model ($\mathbf{S}_m$). $\mathbf{S}_a$ is the error covariance matrix for the prior state variables.

Standard OE theory also enables estimation of the retrieval uncertainty (to be output), a diagnostic of fit (the χ^2 of the retrieval fit) and the sensitivity of the retrieval to the true LSWT ("averaging kernel" in retrieval theory). The latter two outputs are used within quality level attribution (see below).

5.2.3. Quality Level

Quality level is treated as a concept that is distinct from uncertainty: a highly uncertain LSWT can have the highest quality level if all the conditions for giving a valid LSWT and valid LSWT uncertainty are met: the quality level reflects the degree of confidence in the validity of the uncertainty estimate and not the magnitude of data uncertainty.

The quality level assigned to a pixel will be the lowest level (row of table) which matches any of the conditions shown in Table 2. The assignments are compatible with GHRSST conventions: i. e. , a particular level is given if none of the conditions higher up any column of the table are met. In the table d is the distance to land.

Table 2: Quality level criteria

Level	Meaning	Water detection score (0. 5<d<=1. 5)	Water detection score (d>=1. 5)	Sensitivity	χ^2	Other
0	No data	<0	<0			No data; non-lakes pixel
1	Bad data	<0. 5	<0	<0. 1	>3	LSWT < 273. 15
2	Worst quality	<2	<0. 5	<0. 5	>2	Limb (θ_{sat} > 55)
3	Low quality	<3. 5	<2	<0. 9	>1	
4	Acceptable quality	<4. 5	<3. 5		>0. 35	

For instance, any pixel where s is unavailable (value is less than zero), required input BTs are unavailable, or which is over land will be assigned quality level of 0. Next, any pixels close to land which have s < 0. 5, calculated LSWT sensitivity < 0. 1 etc. will be assigned quality level of 1 and so on.

- Quality level 0 pixels should contain no other data
- Quality level 2-5 pixels should always contain valid data
- Quality level 1 pixels contain data but the data is not suitable for use (bad_data). For instance, the LSWT retrieval may have been attempted, but rejected as bad_data due to low sensitivity etc

We recommend using quality level 4 and 5, and consideration of use of quality level 3 with caution, depending on the user's application.

5.2.4. Remapping (L3U)

The remapping from the L2 data in swath projection to the fixed L3 grid proceeds as follows:

- Identify L2 pixels contributing to a L3 cell
- Select highest quality level pixel(s) in the L3 cell
- Calculate average LSWT from the pixels that share the highest quality level and propagate uncertainties to the uncertainty in this average (Bulgin et al. 2015a)

When averaging from the pixel scale to L3 grid scale, the component or uncertainty from uncorrelated errors reduces (uncertainty in the mean is scaled by the familiar "$1/\sqrt{n}$"). Uncertainty in the correlated error components are not reduced by averaging, since over these small scales the degree of correlation will be very high, and is taken to be perfect ("r = 1"). The total uncertainty in the average is found by combining the propagated component uncertainties.

If the grid cell contains pixels which were not included in the averaging (e. g. due to the presence of cloud etc.), then there is an additional uncertainty due to incomplete sampling. This is calculated following Bulgin et al. (2015b) (derived for application to sea surface temperature uncertainty estimation) and is added to the uncorrelated component.

5.2.5. Daily Collation (L3C)

The polar orbiting satellite carrying the AVHRR/ATSR sensors typically complete 14-15 orbits each day resulting in the same number of L2P or L3U products. While L3U files are on a global grid, they are very sparse as the sensor will only observe a small fraction of the Earth's surface in each orbit. For ease of use the LSWT outputs are collated to produce one file for each 24-hour period, corresponding to day-time observations.

Following the GHRSST conventions [D1], when collating observations from overlapping orbits in the same day the L3C will contain the highest quality observation available in the 24-hour period. The selection of best observation is done as follows:

- Choose input cells with the highest quality level
- If multiple observations have the same quality level, then average.

5.2.6. Inter-sensor adjustment (L3S)

Inter-sensor adjustment applies small adjustment factors to reconcile typical differences between the LSWT obtained from different sensors. This is done by selecting a reference sensor and applying a per-lake adjustment to other sensors. This adjustement is applied only if:

- Enough observations where available to estimate the adjustment for the lake (more than 3 months of data for each sensor-to-reference pair).

The uncertainty of the adjustment was < 0. 049. In the previous version this was valid for 80% of the lakes.

A flag indicating whether the adjustment has been applied is present in the files and the uncertainty of the bias correction is included in the total uncertainty. For lakes where the flag is not set, the impact of changes in sensor on the long-term trends in LSWT is less well constrained, and trends should be treated with caution.

5.3. Input Data

5.3.1. AVHRR

The AVHRRs are a series of multipurpose imaging instruments carried onboard the National Oceanic and Atmospheric Administration (NOAA) Polar Operational Environmental Satellites (POES) and EUMETSAT Polar System (EPS) MetOp satellites. There are four AVHRR instruments still in operation with the final AVHRR launched onboard MetOpC in November 2018 (which is still too recent to harmonise). For the LSWT v4. 2 which are generated for the CCI Lake dataset the data from AVHRR on the MetOpA and MetOpB satellites are processed. Only the EPS AVHRRs are used because only for them do we have access to global 1. 1 km data.

The EPS AVHRR is an across-track scanning radiometer using six spectral channels, three visible and three infrared, with a spatial resolution of approximately 1. 1 km at nadir. There are 2048 pixels in each scan for a swath width of about 2800 km. MetOpA (Oct 2006 – present) and MetOpB (Jan 2013 –

present) are both in stable orbits with local equator crossing time (LECT) of 09. 30. AVHRR L1b data are available from EUMETSAT: https://navigator. eumetsat. int/

5.3.2. ATSR

The ATSR instruments were well calibrated, dual-view radiometers intended to produce long-term, consistent LSWT observations. Three ATSR instruments have flown on board ESA's two European Remote Sensing (ERS) satellites and Envisat satellite. All three satellites were in stable sun-synchronous orbits with near-constant Local Equator Crossing Times (LECTs) – the ERS-1 and ERS-2 platforms had a LECT of 10:30 and Envisat had a crossing time of 10:00 all of which were maintained within a few minutes. ATSR-2 (on ERS-2, Aug 1995 – June 2003) and Advanced ATSR (AATSR, on Envisat, July 2002 – April 2012) are used in v4. 2.

The ATSRs have a nadir resolution of 1 km, with a second, forward, view at ~55°. Because of unsolved challenges using both views over topography in relation to small lakes, only the nadir view is used in v4. 2. ATSR L1b are available from ESA: https://earth. esa. int/web/guest/data-access

5.3.3. MODIS

The MODIS instrument that will be used is the sensor on Terra, which is in a stable sun-synchronous orbit with LECT of 10. 30 am, which is relatively consistent with the ATSRs and Metop AVHRRs. It has a viewing swath width of 2, 330 km and views the entire surface of the Earth every one to two days. Its detectors measure 36 spectral bands between 0. 405 and 14. 385 μm, and it acquires data at three spatial resolutions – 250 m, 500 m, and 1, 000 m, with the channels relevant to our algorithms being those with 1, 000 m resolution.

The MODIS data used are the L1b MOD021KM and the correspondent MOD03 geolocation data (collection61) with data from the 24/02/2000. MODIS L1b data are available from NASA: https://earthdata. nasa. gov/.

5.3.4. NWP data and lake mask

ERA-5 profiles are used as an input for the atmospheric profile for the RTM required by the retrieval algorithm, mainly the temperature and humidity profiles but also surface variables (see Table 3: Variables used in RTM and retrieval). The ERA-5 surface temperature is not used for the prior LSWT, since this aspect of the re-analysis is not adequate. Therefore a climatology from v4. 0 LSWT products is used for the GloboLakes lakes of the CCI list while for the other lakes a climatology derived through a slightly different processing is utilised. The mask of the lakes to process together with a per-pixel distance to the nearest land are required as an input of LSWT processor. The mask used is from Carrea et al, (2015b) described in Carrea et al (2015a).

Table 3: Variables used in RTM and retrieval

Parameter	Type	Source	Used in
Atmospheric temperature	Analysis, profile	ERA5	RTM
Atmospheric water vapour	Analysis, profile	ERA5	RTM
Surface pressure	Analysis, surface	ERA5	RTM
Mean sea level pressure	Analysis, surface	ERA5	RTM
10m wind U-component	Analysis, surface	ERA5	RTM
10m wind V-component	Analysis, surface	ERA5	RTM
2m air temperature	Analysis, surface	ERA5	RTM

Parameter	Type	Source	Used in
2m dew point temperature	Analysis, surface	ERA5	RTM
Lake surface water temperature	Created as a climatology	LSWT V4. 0 CLIMATOLOGY	RTM, retrieval
Total Column Water Vapour	Analysis, surface	ERA5	retrieval

5.4. Output Data

The LSWT-specific output data are in netCDF4-classic format and follow the GHRSST data specifications (https://www. ghrsst. org/wp-content/uploads/2016/10/GDS20r5. pdf) which are fully compatible with CCI data standards. For the multi-variable (merged) lakes product, see the relevant product user guide. The LSWT file names have the format:

<Date><Time>-<RDAC>-<Level>-LSWT-<Dataset>-v02. 0-fv01. 0. nc

The variable components within braces: <*component*> are summarised in below and detailed in the following sections. The fixed version string "v02. 0-fv01. 0" indicates that the file is a GDS 2. 0 format file and the file version is 1. 0 (only incremented if a replacement file is later generated). The CDR version is indicated by the <*Dataset*> string.

Table 4: Filename description of LSWT products

Component	Definition	Description
<Date>	YYYYMMDD	The identifying date for this file in ISO8601 basic format
<Time>	HHMMSS	The identifying time for this file in ISO8601 basic format
<RDAC>	LakesCCI	The RDAC where the file was created
<Level>	L3S	The data processing level (gridded supercollated)
<Dataset>	v5. 0	LSWT version number

L3S LSWT-only files are supplied on a regular latitude/longitude grid and variables have dimensions:

time: 1 (defined as unlimited)

lat: Number of latitude points (3600) – i. e. 0. 05° grid

lon: Number of longitude points (7200) – i. e. 0. 05° grid

Table 5: Main output variables of LSWT products

Variable name	Description
lake_surface_water_temperature	Best estimate of $LSWT_{skin}$ as observed by the satellite, in kelvin
lswt_uncertainty	Total uncertainty in $LSWT_{skin}$, in kelvin
quality_level	Quality level of the LSWT: 0 for no data; 1 for bad data; 2 for worst usable data; 3 for low quality; 4 for good quality; 5 for best quality

obs_instr	The instruments used for the correspondent observation
flag_bias_correction	It indicates if a bias correction (to harmonise across sensors) has been applied
lake_id	Lake IDs as used in Globolakes (v4. 0) for backward compatibility described in (Carrea et al. 2015a) and available at (Carrea et al. 2015b)
cci_lake_id	Lake IDs consolidated within the Lakes CCI

5.5. Quality Assessment

The quality assessment exercise is carried out mainly comparing the retrieved LSWTs with independent in situ measurements. An in situ database has been compiled and updated towards the end of each year. Quality control checks are performed on the in situ data ranging from unrealistic values to comparison with the climatology together with its variability. The lakes where in situ data are available are distributed globally although European and especially North American lakes are the most monitored. The assessment of the differences between reference data and satellite LSWT is carried out using robust statistics which is resistant to outliers and bad data in both satellite and in situ measurements. Each LSWT is accompanied by its uncertainty which will be validated using independent reference data as well.

5.6. References

Bulgin, C. E. , Embury, O. , Corlett, G. and Merchant, C. J. (2016a) Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer instruments. Remote Sensing of Environment, 178. pp. 213-222. ISSN 0034-4257 doi:10. 1016/j. rse. 2016. 02. 022

Bulgin, C. E. , Embury, O. and Merchant, C. J. (2016b) Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data. Remote Sensing of Environment, 117. pp. 287-294. ISSN 0034-4257 doi:10. 1016/j. rse. 2016. 02. 021

Carrea, L. , Embury, O. and Merchant, C. J. (2015a) 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

Carrea, L. ; Embury, O. ; Merchant, C. J. (2015b): GloboLakes: high-resolution global limnology dataset v1. Centre for Environmental Data Analysis, 21 July 2015. doi:10. 5285/6be871bc-9572-4345-bb9a-2c42d9d85ceb

Carrea, L. ; Merchant, C. J. (2019): GloboLakes: Lake Surface Water Temperature (LSWT) v4. 0 (1995-2016). Centre for Environmental Data Analysis, 29 March 2019. doi:10. 5285/76a29c5b55204b66a40308fc2ba9cdb3.

MacCallum, S. N. and Merchant, C. J. (2012) Surface water temperature observations for large lakes by optimal estimation. Canadian Journal of Remote Sensing, 38(1). pp. 25-45. ISSN 1712-7971 doi:10. 5589/m12-010

[D1]GHRSST data specification; https://www. ghrsst. org/wp-content/uploads/2016/10/GDS20r5. pdf

6. Lake Ice Cover (LIC) algorithms

6.1. Description

Lake ice cover corresponds to the extent (or area) of a lake covered by ice. The generation of a lake ice cover (LIC) product from satellite observations requires implementation of a retrieval algorithm that can correctly label pixels as either ice (snow-free and snow-covered), open water or cloud. From such product, one can determine ice dates and ice cover duration at the pixel scale (ice-on and ice-off) and lake-wide scale (complete freeze-over (CFO) and water clear of ice (WCI)) (Duguay et al. , 2015). From a climate perspective, determination of ice onset (date of first pixel covered by ice), CFO, melt onset (date of first pixel with open water) and WCI are of most relevance to capture important ice events during the freeze-up and break-up periods. Duration of freeze-up and break-up periods and duration of ice cover over a full ice season can be determined from these dates.

The LIC product generated for Lakes cci uses MODIS (Terra and Aqua) data as to provide the most consistent and longest historical record globally to date (2000-present). The full processing chain and retrieval algorithm are described next.

6.2. Algorithm definition

An overview of the implemented processing chain is given in Figure 10. It includes three modules: data import, retrieval, and data export. Data is processed one day at a time. As part of global initialization, a water mask is loaded. Then, the data for each day is processed. One execution of the processing chain processes one day of data.

MODIS (Aqua and Terra) surface reflectance and brightness temperature bands are used for feature retrieval (i. e. for labelling as water, ice, or cloud). The surface reflectance bands are available at 250 m (QKM) and 500 m (HKM) resolutions. Brightness temperature bands are available in 1 km (1KM) resolution. Geolocation is provided at 1 km resolution and is interpolated to 250 m.

Prior to applying retrieval, pixels of interest are identified as “good” or “bad” using quality bands from the original MODIS surface reflectance products. Pixels of interest are labelled as cloud-covered or cloud-free; cloud-free pixels are then labelled as ice-covered or ice-free (water). Labelled pixels are resampled to the output grid. The processing chain has been revised for cci_lakes to generate the output grid based on specifications of the harmonized product. Aggregation is performed by taking a majority vote between ice and water, ties broken by selecting ice. If there are zero ice and water pixels, then the cell is labelled as cloud if there are non-zero cloud pixels; otherwise the output cell is labelled as “bad”.

More specifically, the processing steps presented in Figure 10 are:

Step 1: Load surface reflectance (bands 2, 3, 4, 6/7), brightness temperature (bands 20, 31, 32), geolocation (latitude and longitude), and quality bands as rasters from MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath (MOD/MYD 09) products.

Step 2: Identify lake (water) pixels of interest based on maximum water extent from ESA CCI Land Cover (v4. 0) 150-m resolution product.

Step 3: Identify pixel quality and label pixels of interest from application of retrieval algorithm for the detection of clouds, ice and open water.

Step 4: Resample labelled pixels acquired in a day from individual swaths to the output grid at 1/120 degrees resolution and perform temporal (daily) and spatial aggregation in terms of each cell in the output grid.

Step 5: Filter the output grid to discard cells (1/120 degrees resolution) which contain land pixels using maximum water extent observed in ESA CCI Land Cover (v4. 0) 150-m resolution product.

Step 6: Write and export the daily lake ice cover product in the required format (NetCDF) with metadata.

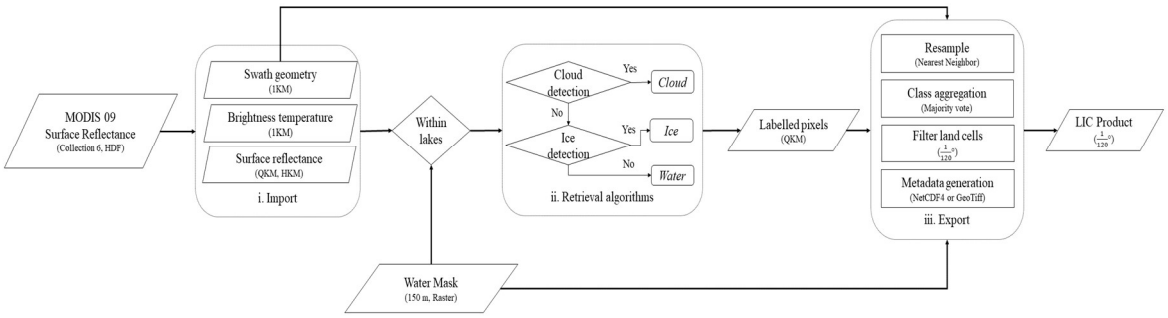

Figure 10: Overview of the processing chain for generation of MODIS LIC daily product

6.2.1. Input Data

Satellite input data:

- MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath, Collection 6 (Vermotte et al. , 2015).

Lake water boundaries:

- Maximum water extent observed in ESA CCI Land Cover (v4. 0) at 150-m resolution.

6.2.2. Output Data

The output data is produced in the harmonized grid format. The edge of each grid cell subtends 1/120 degrees latitude or longitude. The list of variables included in the LIC product are provided in Table 6.

Table 6: Output variables in LIC product

Band	Variable name	Description	Values
1	lake_ice_cover	Label assigned to grid cell	1: water, 2: ice, 3: cloud, 4: bad
2	lake_ice_cover_uncertainty	Uncertainty of the label (%); currently determined from accuracy assessment of individual 250-m products (internal evaluation)	1. 15: water, 8. 29: ice, 4. 37: cloud

6.2.3. Retrieval Algorithm

The retrieval algorithm consists of two parts, one for cloud detection (clear-sky or cloud cover) and other for ice detection (ice or open water).

To develop and validate the retrieval algorithm, 17 lakes distributed across the Northern Hemisphere were selected (Figure 11 and Table 7). 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 (15 years) to ensure algorithm stability. For each lake, one image acquired during freeze-up (FU) and one image during break-up (BU) were selected. False colour composites (R: Band 2, G: Band 2, B: Band 1) with a 250 m spatial resolution were used as reference images to manually collect areas of interest (AOI) with labels (lake ice, open 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.

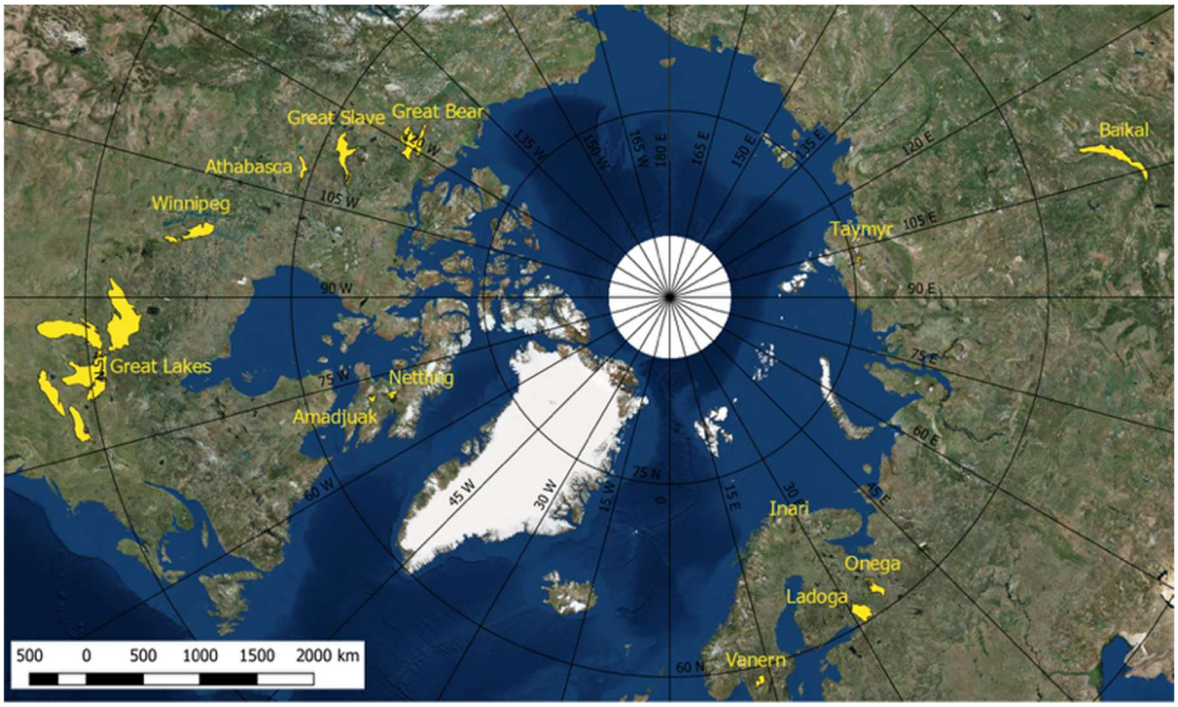

Figure 11: Geographical distribution of lakes used for LIC algorithm development and validation

Table 7: List of lakes for LIC algorithm development and (internal) validation

Lake	Country	Latitude	Longitude	Elevation (m)	Area (km^2)
Amadjuak	Canada	64. 925	-71. 149	113	3, 115
Athabasca	Canada	59. 424	-109. 34	213	7, 900
Baikal	Russia	53. 525	108. 207	456	31, 500
Erie	Canada/USA	42. 209	-81. 246	174	25, 821
Great Bear	Canada	66. 024	-120. 61	186	31, 153
Great Slave	Canada	61. 579	-114. 196	156	28, 568
Huron	Canada/USA	44. 918	-82. 455	176	59, 570
Inari	Finland	69. 048	27. 876	118	1, 040
Ladoga	Russia	60. 83	31. 578	5	18, 135
Michigan	USA	43. 862	-87. 093	177	58, 016
Nettilling	Canada	66. 42	-70. 28	30	5, 542
Onega	Russia	61. 75	35. 407	35	9, 890
Ontario	Canada/USA	43. 636	-77. 727	75	19, 009
Superior	Canada/USA	47. 945	-87. 32	183	82, 367
Taymyr	Russia	74. 538	101. 639	6	4, 560
Vanern	Sweden	58. 88	13. 22	44	5, 650

Winnipeg	Canada	52. 421	-97. 677	217	23, 750

6.2.3.1. Cloud detection

The Simple Cloud Detection Algorithm (SCDA) proposed by Metsämäki et al. (2011) has been selected and slightly modified for cloud detection in the LIC processing chain (Figure 12). The original SCDA is applied to MODIS TOA L1 while the current lakes_cci LIC processing chain uses MODIS surface reflectance products. The main approach of SCDA relies on the difference in brightness temperature (TB) between MODIS Band 31 (11 μm) and Band 20 (3. 7 μm). The difference value of cloudy observations is a large negative value due to the strong reflectance of solar radiation at 3. 7 μm. In order to cope with the difficulty of cloud detection over regions where snow is present, $Diff_{thre}$ is introduced in the algorithm. $Diff_{thre}$ is defined by the regional brightness temperature (Band 32), resulting in colder features obtaining a lower threshold. However, $Diff_{thre}$ cannot be higher than a value of -6 (determined from empirical testing). The combination of Scenario 1 (for opaque clouds) and Scenario 2 (for non-opaque clouds) determines if a pixel is obscured by cloud. The Normalized Snow Difference Index (NDSI) and TOA reflectance in Band 4 are also considered in the conditions. NDSI is calculated from the normalized difference of visible (Band 4) and shortwave infrared (Band 6) for MODIS Terra (MOD). Since some detectors of the Aqua MODIS Band 6 are either noisy or non-functional, Band 7 is used for Aqua images (MYD) instead. This algorithm has previously been tested using nine images from AATSR compared to three images from the NASA MODIS cloud product (MOD35) under different atmospheric conditions and solar zenith angles. Based on a visual evaluation, cloud masks generated by SCDA compared well to the NASA MODIS cloud product, and did provide some improvements, particularly in cold regions (Metsämäki et al. , 2011).

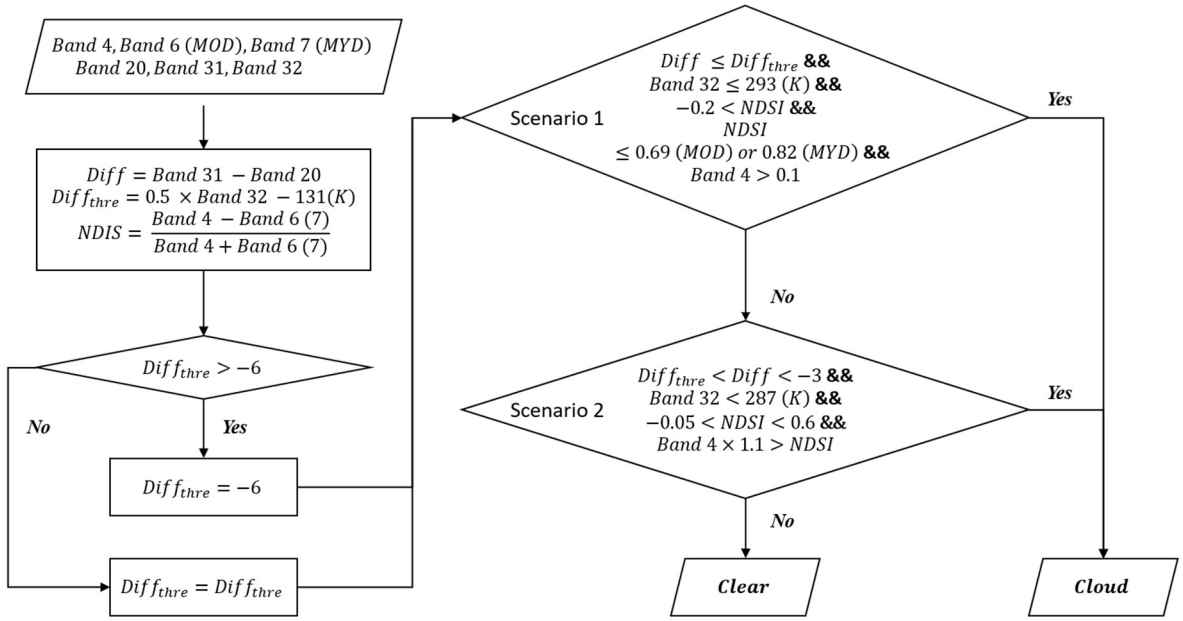

Figure 12: Cloud detection threshold-based algorithm

6.2.3.2. Ice detection

The reflectance of ice and snow is high in the visible and near-infrared portions of the spectrum (Svacina et al. , 2014). Previous studies show that the near-infrared (NIR) MODIS Band 2 is the most suitable band for discriminating ice from open water (Nonaka et al. , 2007; Jönsson and Eklundh, 2004; Šmejkalová et al. , 2016). However, the usage of Band 2 alone is not sufficient for distinguishing ice from open water in areas of lakes where sediment loads are high. As an example, as the RGB

colour composite (image on the left) of Figure 13 shows, Slave River can transport large amounts of sediment into Great Slave Lake, Canada. By using MODIS NIR Band 2 alone in the ice detection algorithm, the area of the lake with a high sediment load is misclassified as ice cover (image on the right). Adding MODIS Band 3 (blue) and Band 4 (green) in the retrieval algorithm results in correct identification of open water (image in centre).

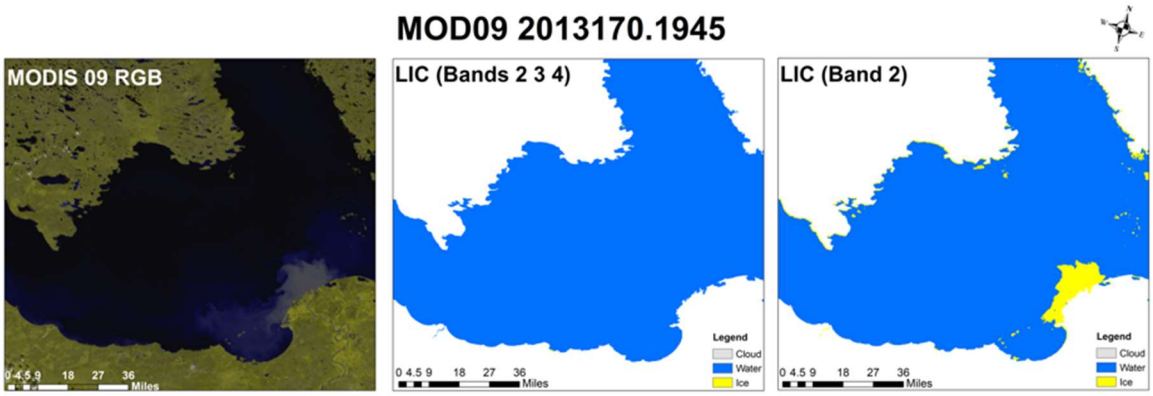

Figure 13: Example where open water is falsely identified as ice covered (image on right) in an area of Great Slave Lake (Canada) due to high sediment load (clearly seen in RGB image on left) from the Slave River entering the lake. False detection of ice occurs when only MODIS NIR band 2 is used. Open water is correctly retrieved (image in centre) when MODIS bands 3 (blue) and 4 (green) are used in addition to NIR band 2

The ice detection algorithm is a computationally low-cost threshold-based classifier that uses MODIS bands 2 (NIR), 3 (blue) and 4 (green). The thresholds for these three bands have been optimized using training data composed of a random 70% selection of sampled data (n = 386, 309). The training data were randomly and evenly separated into 100 groups. An objective function for optimization was applied based on a tradeoff equation (Eq. 6. 1).

$$tradeoff\ value = \varepsilon^2 + s^2 \qquad [6.\ 1]$$

where ε^2 is the mean classification error of the training data and s^2 is the variance of 100 errors computed from each group. Using the three MODIS bands, the algorithm iterates through each possible combination of bands for surface reflectance values between 0 and 0. 15 at an interval of 0. 001, resulting in a total of 3, 375, 000 combinations. Each combination has been tested using the tradeoff value and the combination with the lowest tradeoff value was selected as the most optimal. Figure 14 shows the optimal thresholds values obtained and implemented in the processing chain.

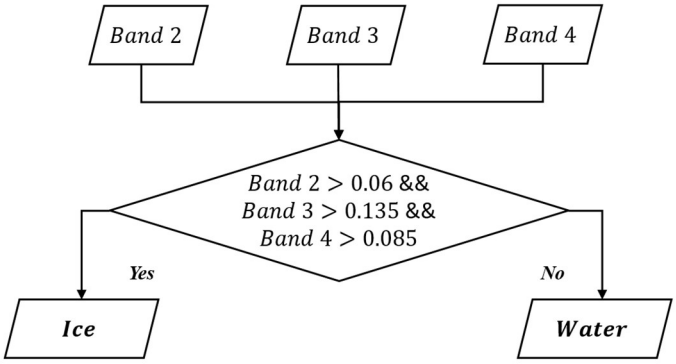

Figure 14: Ice/open water detection threshold-based algorithm

6.3. Quality Assessment

Quality assessment of the LIC product is accomplished by comparing retrieved ice, water and cloud pixels against those obtained from visual interpretation of RGB colour composite images from MODIS Terra/Aqua Atmospherically Corrected Surface Reflectance 5-Min L2 Swath. The images are selected to include several lakes across the Northern Hemisphere over a few ice seasons of MODIS Terra/Aqua record (2000-2019). Other approaches to quality assessment, such as comparison with in-situ observations (from lake shore) and satellite-based LIC products generated by other groups is envisaged (see PVP for details). Uncertainty values reported in LIC are currently derived from overall classification errors for each class (ice, water and clouds).

6.4. References

Duguay, C. R. , Bernier, M. , Gauthier, Y. , & Kouraev, A. (2015). Remote sensing of lake and river ice. In Remote Sensing of the Cryosphere, Edited by M. Tedesco. Wiley-Blackwell (Oxford, UK), 273-306.

Jönsson, P. , & Eklundh, L. (2004). TIMESAT—a program for analyzing time-series of satellite sensor data. Computers & Geosciences, 833-845.

Metsämäki, S. , Sandner, R. , Nagler, T. , Solberg, R. , Wangensteen, B. , Luojus, K. , et al. (2011). Cloud detection algorithm SCDA. GlobSnow Technical Note 2. European Space Agency.

Nonaka, T. , Matsunaga, T. , & Hoyano, A. (2007). Estimating ice breakup dates on Eurasian lakes using water temperature trends and threshold surface temperatures derived from MODIS data. International Journal of Remote Sensing, 2163-2179.

Šmejkalová, T. , Edwards, M. E. , & Dash, J. (2016). Arctic lakes show strong decadal trend in earlier spring ice-out. Scientific Reports, 1-8.

Svacina, N. A. , Duguay, C. R. , & Brown, L. C. (2014). Modelled and satellite-derived surface albedo of lake ice – Part I: evaluation of the albedo parameterization scheme of the Canadian Lake Ice Model. Hydrological Processes, 28, 4550–4561.

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.

7. Lake Water Leaving Reflectance (LWLR) algorithms

7.1. Description

LWLR processing inherits the *Calimnos* processing chain that was initially developed for the UK-based GloboLakes project, in turn based on the first global inland waterbodies processing chain using a comprehensive library of algorithms and resultant products in the ESA Diversity-2 project. These processing chains were built to process archived ENVISAT-MERIS data at full resolution (300m).

Calimnos is presently used to deliver the *Copernicus Land Monitoring Service (CLMS) – Lake water quality* products (LWLR, Turbidity, Trophic State Index) at 10-day aggregation intervals. For CLMS, archived MERIS full resolution data are presently available from *Calimnos* v1. 1, operational processing of Sentinel-3 OLCI uses *Calimnos* v1. 2 and Sentinel-2 MSI scenes are being processed for selected regions using *Calimnos* v1. 3. The incremental versions represent evolutions of the chain to handle newly introduced satellites, as well as updated dependencies such as POLYMER for atmospheric correction or Idepix for land/cloud/water masking.

Lakes cci extends *Calimnos* to include per-pixel product uncertainties, use of MODIS-Aqua, and inter-sensor bias corrections. Extension of the data record to SeaWiFs has lower priority and is not yet discussed in this document. It should be noted that *Calimnos* is a processing chain with many processing stages, each of which are described in documents referred to in the following sections as relevant. The algorithms that form the core of atmospheric correction and retrieval of water column optical properties are all based on published literature whereas algorithm-specific tuning and their assignment to specific optical water types is unique to *Calimnos*. The algorithm basis described here is equivalent to the ATBD provided for CLMS, with algorithm evolution and new elements for lakes cci specified in additional detail.

7.2. Algorithm overview

Calimnos combines data discovery, subsetting by target area (individual water bodies), radiometric and atmospheric corrections, pixel identification (land/cloud/water/ice), optical water type classification, individual algorithms (per parameter and water type), algorithm blending, conversion and aggregation into a single processing chain.

A schematic overview of *Calimnos* is given in Figure 15. The main processing stages and their corresponding algorithms are given below, with stages in current development for lakes CCI highlighted:

To produce Lake Water-Leaving Reflectance:

- Data discovery. Following download of new satellite passes at L1B these are entered into a geospatial database. Target regions are similarly specified in a geospatial database and satellite products which overlap any of the target regions are queued for processing. In the context of re-processing, any duplicate passes are removed. The procedure relies on in-house python scripts and postgres database functionality.
- Subsetting. For best processing performance, satellite passes are subset to bounding boxes around each target area. The subsetting routine is part of the SNAP toolbox, called through the Graph Processing Tool (GPT).
- Radiometric corrections. Any radiometric corrections defined following the release of the data are applied to the L1B imagery before submitting the data to atmospheric correction. For MERIS reprocessing the 3rd reprocessing and subsequently published radiometric corrections are used as part of the SNAP toolboxes using GPT. The 4th reprocessing of MERIS will be used once it becomes available. Radiometric correction for OLCI is currently pending, and not yet implemented. Radiometric corrections for MODIS will be added in this project.
    - Bias corrections: If inter-sensor bias corrections can be determined at top-of-atmosphere they will be included here to correct the input data. Whether this is indeed possible depends on the atmospheric correction procedure used (e. g. is it adjusted to use un-corrected data)

- Pixel identification. The *Idepix* neural network routine is applied for initial pixel identification as water, land, cloud/haze, or snow/ice. *Idepix* is called through SNAP using the GPT. Pixel identification masks are stored for later masking of invalid (non-water) pixels.
- Atmospheric correction. POLYMER is applied to the corrected L1B data of MERIS and OLCI sensors and yields water-leaving reflectance wavebands. The outputs are fully normalized water-leaving reflectance per waveband. POLYMER is called using a function wrapper in Python. *For MODIS data, investigations of the most reliable atmospheric correction are ongoing with POLYMER and l2gen as candidate algorithms.*

To produce derived water-column properties (turbidity and concentrations of chlorophyll-a as well as other water constituents once these reach methodological maturity):

- Optical water type classification. The optical water type (OWT) classification developed in the GloboLakes project (Spyrakos et al. 2018) is applied to each pixel to determine the similarity of the observed water-leaving reflectance spectrum to thirteen known types.
- Algorithm mapping and blending. For each OWT a best-performing algorithm (see section 4) has been selected and tuned against the global in situ reference LIMNADES data set, as part of the GloboLakes project. *Further updates of the tuning parameters for algorithms for chlorophyll-a and turbidity are under development for newly introduced sensors.* Insofar as bias correction has not already been corrected within the top-of-atmosphere product (see radiometric correction above), this step includes further corrections per algorithm, optical water type and sensor.

For uncertainty characterization:

- Uncertainty mapping. The uncertainty mapper uses results from in-situ validation, separated by optical water type, to produce bias and root-mean-square uncertainties per pixel. A predetermined set of uncertainty functions expresses, per output product (LWLR or derived water constituent concentrations), any non-linearity in product uncertainty as a function of (for the time being) optical water type membership.

For merged lakes ECV product format consistency:

- Aggregation. Aggregation is done on a per-lake basis using all imagery available on a given day, after applying masks to select data for water pixels. *When multiple LWLR products are available for a given day, preference is given to the MERIS or OLCI derived cloud-free results closest to solar noon, over e. g. MODIS or averaging of LWLR spectra. Merged product from multiple sensors will be considered as an alternative strategy to enhance spatio-temporal coverage. Product uncertainty will be the primary inclusion criterion.* The aggregated products which contain the reflectance, chlorophyll-a and turbidity parameters are then mapped to a global grid according to Lakes_cci specifications.
- Mosaicking. For consistency with other thematic lakes cci variables, the products are combined on a global grid extending from -180 to 180°longitude and -90 to 90° latitude.

7.2.1. Algorithm assumptions and known limitations

The following assumptions relating to individual algorithms are core to the performance of the processing chain for LWLR and derived substance concentrations:

- *Idepix* can differentiate adequately between pixels containing water and any of the following conditions: mixed land/water, cloud, cloud shadow, ice, snow, and haze.
- *POLYMER* successfully retains the shape and amplitude of water-leaving reflectance.
- The Optical Water Type classification sufficiently captures the diversity of natural water types so that the most appropriate algorithms can be used and tuned to remove systematic bias.
- Tuning of reflectance algorithms for chlorophyll-*a* and turbidity is adequate for each water type, and based on sufficient in situ data availability to achieve statistical rigour.

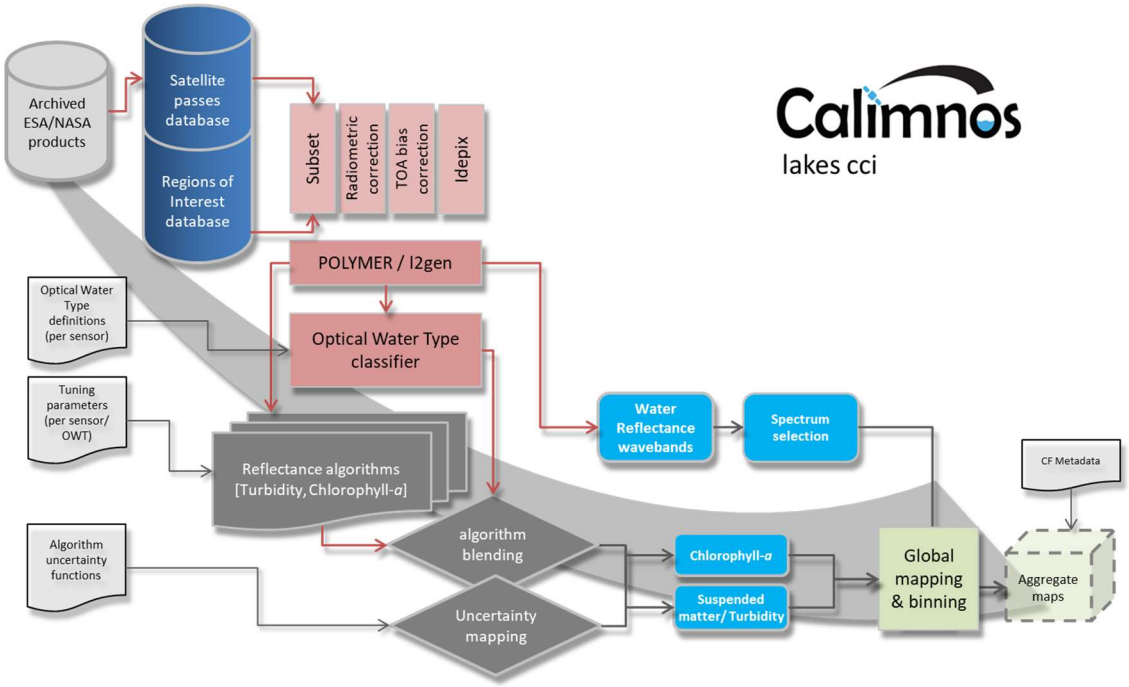

Figure 15: Schematic overview of the Calimnos processing chain for LWLR, Chlorophyll-*a* and turbidity or suspended matter.

7.2.2. Specific algorithms for LWLR

7.2.2.1. Pixel geolocation

The standard MERIS **geometric correction** needs an optimization step which is performed by the AMORGOS software developed by ACRI-ST and provided by ESA (see earth. esa. int/ services/amorgos/download/Amorgos_ICD-SUM_3. 0a. pdf and earth. esa. int/services /amorgos/download/Amorgos_STD_i3r0p1. pdf). It includes a precise orbit determination, instrument pointing and performs an ortho-rectification. The improvement in the geolocation is documented in Bicheron et al. (2011). For *Calimnos,* AMORGOS geometrically corrected lat/lon bands are archived alongside the original L1B product, used to identify the area to be processed and patched in at the end of L2 processing. The resulting accuracy is improved by Amorgos and is specified in Bicheron et al. (2011) as better than 70m for MERIS FR pixels.

For OLCI, no further processing step for improving the geolocation position is required. The accuracy is given by ESA with 0. 2 – 0. 7 pixels for processing version 2. 23 and <0. 1 pixel for processing version 2. 29.

For MODIS Aqua, a GEO file is generated based on satellite attitude and ephemeris data provided by NASA, to produce L1B data equivalent as input to further processing.

7.2.2.2. Radiometric corrections

OLCI-A will be the reference system for inter-sensor bias corrections, because it has the largest number of spectral bands and benefits from a larger vicarious calibration network than any earlier sensor. A per-band system vicarious gain correction is currently available for use with POLYMER.

For MERIS, the coherent noise equalization method reduces detector-to-detector and camera-to-camera systematic radiometric differences and results into a diminution of the vertical stripping observed on MERIS L1b products following the algorithm developed by Bouvet and Ramino (2010); coefficients for each detector line are retrieved from MERIS archive over ice. Another step for radiometric improvements is the smile correction, which corrects for the small variations of the

spectral wavelength of each pixel along the image by the estimation of the reflectance spectral slope from the measurements in two neighbouring bands. Polymer uses actual pixel wavelength, so smile correction is not used in the current processing chain.

For MERIS and MODIS, sensor-specific bias corrections, including any known system vicarious calibration gains, are applied.

Configuration of operators:

- MERIS: L1b Radiometric Correction v5. 0. 3 and sensor bias correction (TBD)
- OLCI: System Vicarious Calibration gains (2019)
- MODIS: Sensor bias correction (TBD)

7.2.2.3. Pixel identification

The cloud detection function of the Idepix algorithm developed by Brockmann Consult was used in several processing chains, e. g. those used in CoastColour L1P and Diversity II. Meanwhile many steps of the Idepix algorithm are included in the upcoming MERIS 4th reprocessing as standard algorithm by ESA. Due to the good performance of Idepix cloud screening in these applications, it is also selected for *Calimnos*. Idepix is based on a cloud probability derived from a neural net which has been trained with >60, 000 manually classified pixels and which is combined with a number of additional tests on e. g. brightness, whiteness, glint. After clouds have been identified, a buffer can be defined in order to provide for a safety margin along cloud borders. This buffer radius (in pixel) can be parameterized and is set to 2 pixels.

Validation is performed by applying the PixBox Validation, a procedure where manually selected pixels are categorized to different categories and characterized with expert knowledge, e. g. to clear land, clear water, totally cloudy, semi-transparent cloud, cloud shadow, snow/ice, etc. A set of 17k MERIS FR pixels was collected in the scope of the CoastColour project, and detailed validation results are provided in the corresponding report (Ruescas et al. , 2014).

Retrieval of water quality parameters is also strongly influenced by the occurrence of cloud shadow, which need to be identified and eliminated from further processing. Potential cloud shadow areas are identified by the geometry of the sun angle, viewing angle and the cloud height and the cloud bottom. The cloud height is gained by either the pressure or the temperature, but if this information is missing (not all sensors offer the respective bands), a maximum cloud height needs to be defined. The most difficult prediction is the height of the cloud base as it is not seen by the sensor. In Idepix it is defined as the minimum cloud height detected within the respective cloud minus an offset. Basis of good cloud shadow detection is a good cloud detection. Validation of the cloud shadow detection is done by visual inspection of different images under different conditions (cloud types and geometries).

Configuration:

- MERIS: Idepix. Envisat. MERIS v1. 0
- OLCI: Idepix. Sentinel3. Olci v1. 0
- MODIS: Idepix. TerraAqua. MODIS v2. 2.

In general, the most progressive combination of available cloud masks is selected, favouring accuracy over observation coverage.

7.2.2.4. Atmospheric correction

POLYMER v4. 12 is the latest version of an atmospheric correction processor initially designed to resolve water-leaving reflectance in clear ocean (case-1) waters including areas affected by sun glint (Steinmetz et al. 2011). The versatility of the processor to deal with bright waters has tested positively with a variety of optically complex (including inland) waters compared to alternative processors (Qin et al. 2017, Warren et al. 2019), although systematic under-estimation of reflectance in turbid and productive waters is evident. POLYMER applies a spectral optimization based on bio-optical model in conjunction with radiative transfer models to separate atmospheric (including glint) and water reflectance. The principle of the algorithm is a spectral matching method using a

polynomial to model the spectral reflectance of the atmosphere and sun glint, and a bio-optical forward reflectance model for the water part. The algorithm uses the full set of wavebands available (user-configurable) as opposed to alternative ocean-colour methods that primarily extrapolate from near infra-red bands. The output are fully normalized water-leaving reflectances.

Configuration:

- POLYMER according to Steinmetz et al. (2011), updated in Steinmetz (2016 and 2018), parameterized to use the Park and Ruddick (2005) bidirectional reflectance distribution function and operating only on pixels identified as water by the *Idepix* module (masks generated by POLYMER are not used). Starting conditions for the optimization procedure are set to chlorophyll-a = 1 mg m^{-3} and total suspended matter = 1 g m^{-3}.

7.2.2.5. Specific algorithms for derived water quality products

7.2.2.5.1. Optical water type (OWT) membership

The OWT classification module was written at PML based on the work of Moore et al. (2001) and equivalent software developed for ESA ocean colour cci. The algorithm used for lakes relies on a spectral library (spectral means) defined in the GloboLakes project by the University of Stirling (Spyrakos et al. 2018). In contrast to OWT mapping used in earlier versions of *Calimnos*, CLMS and Lakes_cci will adopt the spectral angle (Kruse et al. 1993) rather than Mahalonobis distance as metric for similarity between spectra. The spectral angle is here defined over a range of 0 to 1 where 1 implies identical spectra.

7.2.2.5.2. Water constituent algorithms

Weighted blending: Water constituent retrieval algorithms tuned to each OWT (Spyrakos et al. 2018, Neil et al. 2019) are mapped to individual pixels from the OWTs with the three highest classification scores for that pixel. 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 4^{th} ranking OWT. This procedure is used to derive maps of total suspended matter (TSM) and chlorophyll-*a* (Chla) without discontinuities at the edge of the applicable range of any single algorithm.

Uncertainty mapping: to propagate product uncertainty from the individual algorithms, the weighted OWT membership score is again used, in combination with a set of uncertainty functions resulting from in situ algorithm validation. The uncertainty functions describe product uncertainty as a function of OWT membership score and target substance concentration (where relevant). The per-OWT uncertainties are weighted according to OWT-membership to allow propagation to the final product. Where in situ data are lacking in lakes to determine product uncertainties (e. g. in the case of a specific OWT, substance concentration range or sensors), and uncertainty cannot be provided this value is set to Inf (infinity).

7.2.2.5.3. Chlorophyll-a algorithms

Table 8 below lists the mapping of algorithms to specific OWTs as given in Spyrakos et al. (2018). Each algorithm is tuned to one or more OWTs, depending on their published range of applicability and the wavebands used in their design. The algorithms set out below are those selected for MERIS after extensive product validation in the GloboLakes project (Neil et al. 2019, Steele et al. *in prep*). It is noted that while the methodology of algorithm tuning is as described in Neil et al. (2019), tuned algorithm coefficients may differ since the former are derived from calibration of in situ reflectance data against LIMNADES whereas *Calimnos* uses coefficients optimised for POLYMER-corrected normalized water-leaving reflectance. OLCI offers the same set of wavebands as MERIS, but validation of OLCI-specific algorithms is limited by scarce in situ data availability. For OLCI, therefore, we adopt the MERIS configuration of algorithms which may be further adjusted following consistency checks (preliminary analysis within CLMS suggests that no adjustment is required). For MODIS, algorithm selection is pending an evaluation of product consistency over time, for selected lakes.

Table 8: Chlorophyll-a algorithms per optical water type

Optical water type number	Algorithm source	Algorithm optimization
3, 9, 10, 13	OC2 oceancolor. gsfc. nasa. gov/cms/atbd/chlor_a	Empirical re-tuning of algorithm parameters based on GloboLakes calibration against the Limnades database, specific to each optical water type (Neil et al. 2019).
2, 8, 11, 12	708/665 empirical band ratio based on Gilerson et al. (2010)	
1, 4, 5, 6	Semi-analytical NIR-Red band algorithm for MERIS based on Gons et al. 2005.	
7	Adapted QAA algorithm according to Mishra et al. (2013)	

Selected algorithm parameters for each of the four options were tuned in GloboLakes against the LIMNADES database to accommodate for variability in retrieval accuracy over the range of inland waters encountered in the LIMNADES dataset. This procedure also accounts for any uncertainties that stem from systematic bias in the retrieval of water-leaving reflectance from POLYMER. This whole-chain validation is applied to each algorithm in *Calimnos*, recognizing the fact that the most prominent error in inland water quality retrieval is the atmospheric correction step. It is noted that validation data for relatively clear inland waters are very scarce.

The OC2 algorithm, originally formulated to retrieve chlorophyll-a concentration from relatively clear ocean waters where phytoplankton and other optically active substances covary, relies on a ratio of blue and green wavebands. The algorithm is formulated as:

$$log(Chla) = a_0 + (a_1 x) + a_2 x^2 + a_3 x^3 + a_4 x^4 \qquad [7.1]$$

where x is the reflectance band ratio:

$$x = log\ \frac{R_w(490)}{R_w(560)}. \qquad [7.2]$$

The tuned algorithm coefficients used in Calimnos are a_0 = 0. 1731, a_1 = -3. 9630, a_2 = -0. 5620, a_3 = 4. 5008 and a_4 = -3. 0020. R_w is the fully normalized water-leaving reflectance.

The remaining three algorithms are variations on empirical and semi-analytical retrieval methods that focus on interpreting the red to near-infrared part of the spectrum. These methods are the most successful approach to retrieve chlorophyll-a concentration over a range of turbidity and trophic levels, because the red part of the spectrum is least influenced by overlapping absorption signatures of dissolved substances and other phytoplankton pigments.

The Gilerson et al. (2010) algorithm is an empirically tuned ratio of bands 708 and 665 nm. The tuning of the algorithm is revised in *Calimnos*, by calibrating against the LIMNADES data set (for OWTs 2, 8, 11, and 12) which is much larger than the data set used in the original publication (Neil et al. 2019). The final configuration is a highly simplified form of the original algorithm:

$$Chla\ [mg\ m^{-3}] = A \times \left(\frac{R(709)}{R(665)}\right)^B + C \qquad [7.3$$

Where A = 76. 62, B = 0. 7393 and C = -54. 99 are tuning coefficients empirically calibrated against LIMNADES. $R(\lambda)$ is the reflectance (irrespective of whether it is expressed above or below water and normalized for viewing geometry or not) at waveband λ.

The Gons et al. (2005) algorithm (tuned to OWTs 1, 4, 5, 6) uses the same band ratio and additionally analytically retrieves the backscattering coefficient from the 778 nm band. Subsequently the absorption at 665 nm is analytically retrieved by inverting the Gordon reflectance model and attributed to chlorophyll-*a* and water. Empirical tuning is restricted to the slope of the backscattering

coefficient and a chlorophyll-*a* specific absorption coefficient determined from eutrophic inland waters. The algorithm is thus specified as follows:

$$Chla\ [mg\ m^{-3}] = \left[\left(\frac{R(709)}{R(665)}\right) \times (a_w(709) + b_b) - a_w(665) - b_b^P\right] / a^*_{chl}(665) \qquad [7.4]$$

Where $a_w(709)$=0. 84784 m^{-1} and $a_w(665)$=0. 431138 m^{-1} represent the absorption by pure water from Roettgers et al. (2011). Further, $a^*_{chl}(665)$ = 0. 025 $m^2\ mg^{-1}$ is the chlorophyll-a specific absorption coefficient following calibration against LIMNADES. The empirical constant P=1. 06 was not changed from the original formulation. The backscattering coefficient b_b is considered spectrally neutral and derived from a single near infra-red waveband:

$$b_b = \frac{0.6 \times a_w(779) \times R_w(779)}{0.082 - 0.6 \times R_w(779)} \qquad [7.5]$$

The Mishra et al. (2013) implementation of the Quasi-Analytical Approach (QAA) is similar to the analytical inversion of Gons et al. (2005) but with a slightly different set of input bands and subsequent tuning parameters. In *Calimnos* it is exclusively mapped to OWT 7. The QAA chlorophyll-a product is derived from the phytoplankton absorption at 665nm, empirically tuned against LIMNADES as follows:

$$Chla\ [mg\ m^{-3}] = A \times a_{ph}(665)^B \qquad [7.6]$$

Where A = 63. 375 and B = 0. 442. The $a_{ph}(665)$ is retrieved from a set of equations, accounting for non-phytoplankton absorption in this band through the interpretation of absorption in blue and green wavebands:

$$a_{ph}(665) = a(665) - a_w(665) - a_{ys}(665),$$

$$a_{ys}(665) = a_{ys}(442) e^{-S(665-442)},$$

$$a_{ys}(442) = \frac{(a(412) - s_1 \times a(442)) - (a_w(412) - (e_1 \times a_w(412)))}{e_1 - s_1} \qquad [7.7]$$

where the absorption by water a_w is again obtained from Roettgers et al. (2011), S = 0. 0135 nm^{-1} is the average exponential slope coefficient for yellow substances derived from LIMNADES, and e_1 and s_1 are defined as

$$s_1 = 0.74 + \frac{0.2}{0.8 + \frac{rrs(442)}{rrs(560)}},$$

$$e_1 = e^{S(442-412)} \qquad [7.8]$$

where and the absorption in bands 412, 442 and 665 nm is obtained as

$$a(\lambda) = \frac{(1.0 - u(\lambda)) \times (b_{bw}(\lambda) + b_b(\lambda))}{u(\lambda)}. \qquad [7.9]$$

Here, $b_{bw}(\lambda)$ is the backscattering coefficient of pure water obtained from Morel (1974) assuming zero salinity. In turn, $u(\lambda)$is the ratio of backscattering to the sum of backscattering and absorption, which according to the work by Gordon et al. (1988) can be obtained from below-surface remote-sensing reflectance $rrs(\lambda)$ as:

$$u(\lambda) = \frac{-g_0 + \sqrt{g^2 + (4 \times g_1) \times rrs(\lambda)}}{2 \times g_1}, \qquad [7.10]$$

With g_0 = 0. 089 and g_1 = 0. 125. The *rrs* bands are obtained from the LWLR (which corresponds to fully normalized water-leaving reflectance, $R_w(\lambda)$) using:

$$rrs(\lambda) = \frac{R_w(\lambda)}{\pi(0.52+0.54\times R_w(\lambda))} \quad [7.11]$$

7.2.2.5.4. Turbidity and suspended matter algorithms

Turbidity and total suspended matter (TSM) may be retrieved from reflectance in wavebands where phytoplankton and dissolved organic matter absorption do not significantly influence the amplitude of the reflectance. Ultimately, turbidity and suspended matter algorithms are empirically related to efficiency of light backscattering compared to absorption. The absorption of light becomes increasingly predictable with waveband due to the efficiency of absorption by pure water at longer wavelengths. Candidate algorithms to directly convert the signal to either Turbidity or suspended matter dry weight, using a conversion factor of 1. 17 NTU/g m^{-3} between suspended matter and Turbidity, have been formulated by Nechad et al. (2010, 2016). This conversion factor is currently used in *Calimnos* to obtain Turbidity from a number of suspended matter retrieval algorithms, pending validation of which algorithms perform better over specific optical water types. The suspended matter algorithms are selected per Optical Water Type as shown in Table 9.

Table 9 Suspended matter algorithms per optical water type

Optical water type number	Algorithm source	Algorithm optimization
1, 7, 10	Based on Zhang et al. (2014)	Empirical re-tuning of algorithm parameters based on GloboLakes calibration against the Limnades database, specific to each optical water type (analogous to Neil et al. 2019 for chlorophyll-*a*).
2, 4, 6, 8, 12	Based on Vantrepotte et al. (2011)	
3, 5, 9, 11, 13	Based on Binding et al. (2010)	

The Binding et al. (2010) algorithm as it is implemented here is based on the analytical inversion of reflectance in the 754 nm band and converting the resultant particulate backscattering signal using a mass-specific backscattering coefficient for suspended matter:

$$TSM\ [g\ m^{-3}] = \frac{a_w(754)\times R_w(754)}{f\times B\times b^*_{TSM}} \quad [7.12]$$

where the absorption by water $a_w(754)$ = 2. 8 m^{-1}, the backscattering-to-scattering ratio *B*=0. 019. The TSM-specific scattering coefficient b_{TSM}^* = 0. 664 $m^2\ g^{-1}$ and is the only coefficient that was optimized against LIMNADES.

The Vantrepotte et al. (2011) algorithm as it is implemented here is similar but uses the 665 nm band and an additional empirical factor, which is also tuned to provide the best match for the corresponding water types:

$$TSM\ [g\ m^{-3}] = \frac{A\times R_w(665)}{1-\frac{R_w(665)}{B}} + C \quad [7.13]$$

where the optimized empirical coefficients are *A* = 206. 4, *B* = 20460. 0 and *C* = -0. 7921.

The Zhang et al. (2014) algorithm as it is implemented here is an empirical relation between the 709 nm band and in situ measured suspended matter dry weight, tuned as with the algorithms above:

$$TSM\ [g\ m^{-3}] = A\times\frac{R_w(709)^B}{\pi} \quad [7.14]$$

where the optimized empirical coefficients are *A* = 2524. 0 and *B* = 1. 113.

7.3. Input products and dependencies

Satellite input data:

- Envisat MERIS L1B (3rd reprocessing, 4th once available) at reduced resolution
- Sentinel 3A/B OLCI L1B (SAFE format)
- MODIS Aqua L1A (+ GEO files)

Lake water boundaries:

- maximum water extent observed in ESA CCI Land Cover (v4. 0) at 150-m resolution
- polygons generated (including manual inspection) at PML (doi: 10. 5281/zenodo. 3349547).

POLYMER ancillary data:

- ECMWF ERA-Interim global atmospheric data set

Optical water type definitions (mean spectra standardized using Simpson's criterion)

- GloboLakes project (Spyrakos et al. 2018)

7.4. Output product

The output data (product bands) are produced as variables in a NetCDF file. Variables include a band for each reflectance band, the derived chlorophyll-a and turbidity and the associated pixel uncertainty for each of these. Intermediary products are not distributed but are generally stored for product validation and improvement purposes. These include the specific outputs from individual algorithms (prior to mapping/blending) and all processor-generated flags. A detailed overview of the output bands, their data types and attributes is provided in the Product Specification Document (PSD).

7.5. Quality Assessment

Quality assessment of the LWLR and derived products is based on validation against in situ observations. This section contains an overview of product quality assessment efforts resulting from past and present validation activities, in order to provide the user sufficient information to determine which variables are suitable for their particular use case.

Results provided here are for in situ validation carried out against observations with the MERIS sensor, for which by far the most in situ data are available.

7.5.1. Quality assessment of the atmospheric correction

The initial round-robin comparison of atmospheric correction algorithms for MERIS showed that POLYMER (v3. 5) yielded the statistically most robust retrieval of R_w particularly with respect to linearity and relative errors. However, a systematic negative bias was observed in the matchup validation (Figure 16). Given the low number of matchups in the analysis, an expanded round-robin comparison of algorithms for MERIS and other sensors is foreseen in the lakes_CCI.

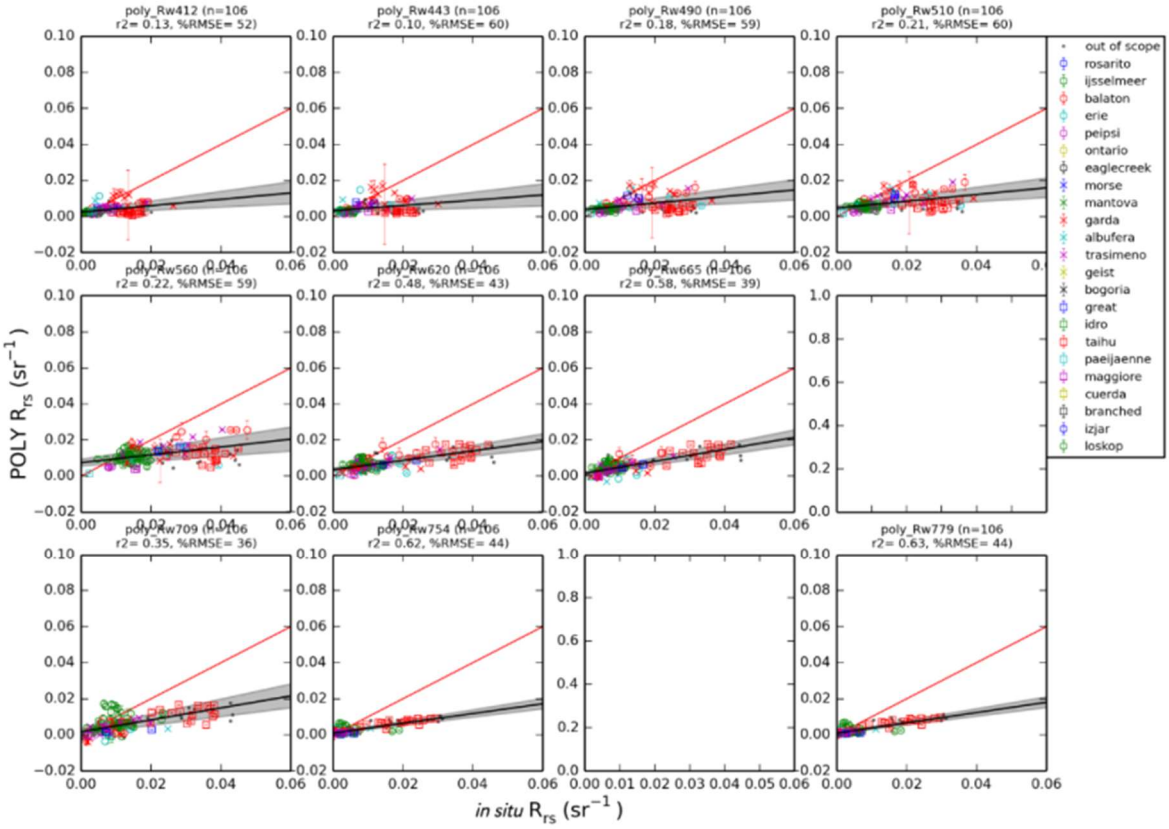

Figure 16: POLYMER v3. 5 matchups with in situ reflectance data in LIMNADES, using a ±7 day matchup window and 3x3 pixel extraction window (source: GloboLakes).

7.5.2. Quality assessment of derived water-column products

The end-to-end calibration methodology with specific tuning of algorithms corresponding to sets of optical water types is believed to counteract the systematic bias in LWLR retrieval. Furthermore, the number of matchup data points for validation of chlorophyll-a and turbidity or suspended matter is much higher.

Neil et al. (2019) presented the results of algorithm calibration using exclusively in situ radiometry and substance concentrations determined from water samples. During GloboLakes, further tuning of the algorithms was carried out to calibrate the algorithms against atmospherically corrected satellite data (using POLYMER and other candidate atmospheric correction algorithms for MERIS).

Results shown below are from a recent uncertainty characterization analysis of the chlorophyll-*a* product in the *Calimnos* v1. 04 data set, which is similar to the result expected in the first climate data record of the Lakes_cci. Further multi-sensor calibration and uncertainty characterization is part of the planned Lakes_cci work.

These results (Figure 17) clearly show that applying the weighted average of a combination of (two) tuned algorithms for each observed pixel provides a marked improvement over selecting, for each lake and each observation day, the algorithm that is most suitable for the lake-wide predominant optical water type. One clearly visible effect is the need to select algorithms that can deal with a (very) high concentration range, likely associated with patchy phytoplankton blooms surrounded by lower biomass conditions. At the scale of the whole lake, bloom-affected pixels are a minority such that mid-range algorithms would perform best for the lake as a whole. A whole-lake algorithm selection can ignore dense (but likely patchy) blooms, shown as saturation at concentrations of approximately 1000 mg m^{-3}, whereas per-pixel algorithm selection allows retrieval up to two orders

of magnitude higher. The dynamic algorithm selection approach is therefore preferred as it is better equipped to deal with optical gradients within individual lakes.

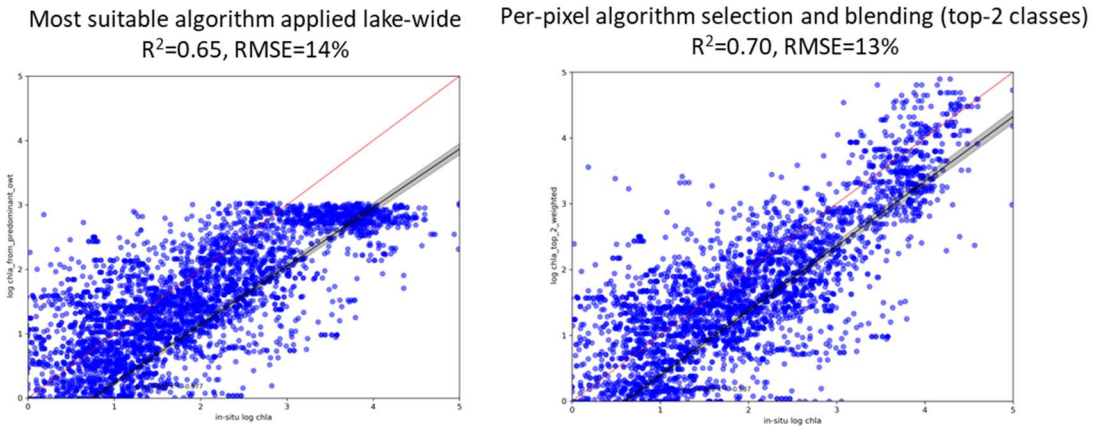

Figure 17: Comparison of the performance of (left) the best chlorophyll-a algorithm for the predominant optical water type of each lake and (right) per-pixel selection and blending of the two highest ranking algorithms based on optical water type membership scores.

7.6. LWLR References

Bicheron, P. , Amberg, V. , Bourg, L. , Petit, D. , Huc, M. , Miras, B. , Arino, O. (2011): Geolocation Assessment of MERIS GlobCover Orthorectified Products. *IEEE Trans Geosci Remote Sens* , 49(8), 2972–2982.

Binding, C. E. , J. H. Jerome, R. P. Bukata & W. G. Booty. (2010): Suspended particulate matter in Lake Erie derived from MODIS aquatic colour imagery. International Journal of Remote Sensing 31(19).

Bouvet M. , Ramoino F. (2010): Equalization of MERIS L1b products from the 2nd reprocessing, ESA TN TEC-EEP/2009. 521/MB.

Diversity-II (2015): ESA DUE DIVERSITY II- Algorithm Theoretic Baseline Document (ATBD), Version 2. 4, http://www.diversity2.info/products/documents/DEL5/DIV2_Algorithm_Theoretical_Basis_Document_v2. 4. pdf.

Gilerson, A. A. , Gitelson, A. A. Zhou, J, Gurlin, D. , Moses, W. , Ioannou, I. and Ahmed, S. A. (2010): Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Opt Express 18(23): 24109-24125.

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

Gordon, H. R. , O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark. (1988): A semianalytic radiance model of ocean color. J Geophys Res: Atmos 93: 10909-10924.

Kruse F. A. , A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, A. F. H. Goetz. (1993): The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data, Remote Sens Environ, 44(2–3):145-163.

Mishra, S. , D. R. Mishra, Z. Lee, Craig S. Tucker. (2013): Quantifying cyanobacterial phycocyanin concentration in turbid productive waters: A quasi-analytical approach, Remote Sens Environ, 133: 141-151.

Moore, TS. , JW. Campbell and Hui Feng. (2001): A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. IEEE Transactions on Geoscience and Remote Sens 39(8) 1764-1776.

Morel, A. (1974). Optical properties of pure water and pure sea water. In: Jerlov MG, Nielsen ES, editors. Optical Aspects of Oceanography. New York: Academic Press; 1974. p. 1–24.

Nechad, B. , Dogliotti, A. I. , Ruddick, K. G. , Doxaran, D. , (2016): Particulate Backscattering and suspended matter concentration retrieval from remote-sensed turbidity in various coastal and riverine turbid waters. Proceedings of ESA Living Planet Symposium, Prague, 9-13 May 2016, ESA-SP 740.

Nechad, B. , Ruddick, K. G, Park, Y. (2010): Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters". Remote Sens Environ 114 (2010) 854–866.

Neil, C. , Spyrakos, E. , Hunter, PD, Tyler, AN. (2019): A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. Remote Sens Environ 229: 159-178. 10. 1016/j. rse. 2019. 04. 027

Park Y. J. and Ruddick K. Model of remote-sensing reflectance including bidirectional effects for case 1 and case 2 waters. (2005): Appl Optics. 2005;44(7):1236-49. doi: 10. 1364/ao. 44. 001236

Roettgers, R. , Doerffer, R. , McKee, D. and Schoenfeld, W. (2011): Pure water spectral absorption, scattering, and real part of refractive index model. ESA Water radiance project. From www.brockmann-consult.de/beam-wiki/download/attachments/17563679/WOPP.zip?version=1&modificationDate=1299075673760

Ruescas, A. B. , Brockmann, C. , Stelzer, K. , Tilstone, G. H. , Beltran, J. (2014): DUE CoastColour Validation Report (p. 58). Geesthacht, Germany: Brockmann Consult. Retrieved from http://www. coastcolour. org/documents/DEL-27%20Validation%20Report_v1. pdf.

Spyrakos E, O'Donnell R, Hunter PD, Miller C, Scott M, Simis S, et al. (2018): Optical types of inland and coastal waters. Limnol Oceanogr. 63(2). doi: 10. 1002/lno. 10674.

Steinmetz F, Deschamps P-Y, Ramon D. Atmospheric correction in presence of sun glint: application to MERIS. (2011): Optics Express. 19(10):9783-800. doi: 10. 1364/oe. 19. 009783

Steinmetz, F. Ramon, D. , Deschamps, P-Y (2016): ATBD v1 – Polymer atmospheric correction algorithm, D2. 3 OC-CCI project http://www. esa-oceancolour-cci. org/?q=webfm_send/658

Steinmetz, F. (2018): ATBD v1 – Evolution of Polymer Atmospheric correction within Copernicus Global Land Service – Inland Water products. On request

Qin P, Simis SGH, Tilstone GH. (2017): Radiometric validation of atmospheric correction for MERIS in the Baltic Sea based on continuous observations from ships and AERONET-OC. Remote Sens Environ 200:263-80.

Vantrepotte, V. , H. Loisel , X. Mériaux, G. Neukermans, D. Dessailly, C. Jamet, E. Gensac, and A. Gardel, (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. J Coast Res. SI 64 (Proceedings of the 11th International Coastal Symposium), – . Szczecin, Poland, ISSN 0749-0208

Warren MA, Simis SGH, Martinez-Vicente V, Poser K, Bresciani M, Alikas K, et al. (2019): Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters. Remote Sens Environ 225:267-89.

Zhang, Y. , K. Shi, X. Liu, Y. Zhou, B. Qin. (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(5): e98055. doi:10. 1371/journal. pone. 0098055.