climate change initiative

→ LAKES NEWSLETTER

Issue n. 4 | February 2021

In this newsletter

The Lakes Climate Change Initiative project has entered the final year of its first funding phase. This brings a number of research activities to a close while we prepare to produce new and updated global datasets of the Lakes Essential Climate Variable. This newsletter highlights the work done to validate the production of the Lake Water Extent (LWE) product.

Project update

The first version of the Climate Research Data Package (CRDP V1.0) is currently used to explore five climate use case studies. In addition, we are in the process of collating version 1.1 which benefits several improvements and will be available in March of this year. V1.1 has increased coverage of Lake Water Level (LWL) and Lake Water Extent (LWE) observations measurements and implements a new Artificial Intelligence (AI) based algorithm (see doi.org/10.1016/j.rse.2020.112206) for Lake Ice Cover (LIC) detection. We plan to release version 2.0 of the CRDP in September, with increased coverage of up to 2000 inland water bodies per thematic variable.

Two new research activities started in December 2020: the first one is dedicated to the estimation of Lake Ice Thickness (LIT) and the second activity investigates the global consistency of the merged Lakes Essential Climate Variables product.

Lake Water Extent: how is it calculated?

Lake Water Extent (LWE) is typically mapped from image (radar or optical) data whereas Lake Water Level (LWL) is derived from altimetry. However, the two can be directly correlated as follows. First, historical time-series of satellite altimetry Lake Water Level (LWL) are used to determine when the lake was at low, medium or high level. High-resolution images are then collected to over times of low to high water level, and processed to produce the LWE estimate. The relationship between lake water level and extent can be established from around 10 to 15 pairs of observations by fitting a first-or second-order polynomial. This is the *hypsometry curve* of the lake. With this function, LWL from altimetry can be related to LWE from imagery or vice versa. It is beneficial to derive from LWE through hypsometry because analysing high-resolution imagery to derive LWE is a lengthy process.

The LWE information needed to derive the hypsometry curve 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 were investigated to generate LWE. Because the processing chains differ between these methodologies, candidate algorithms were tested using a set of lakes in a benchmarking exercise organised within the project.

Benchmarking the LWE estimation approaches

Four teams participated in the activity to benchmark and validate various approaches to derive LWE: ICube-SERTIT (France) and LEGOS (France) processed optical imagery from the Landsat series and Sentinel-2 A & B, whereas NORCE (Norway) and TRE-Altamira (Spain) processed radar images from Sentinel-1 A & B satellites.

The objective of the benchmarking exercise was to assess the state of the art of available methods to delineate water bodies from High Resolution optical and SAR data. The water masks produced from individual methods were then compared for a number of lakes varying in altitude, depth, surrounding landscape, and hydrology). Lakes were selected from landscapes including savanna, arid / sandy areas, agricultural areas, snow/ice, and mountainous terrain. The subsequent objective was then to establish the feasibility of transferring these methods into a large-scale processing system to produce the LWE thematic ECV for the Lakes_cci production system, and to propose a method for product validation based on coherence between all provided solutions and the precision of the hypsometry curves. The best LWE products can then be used to generate hypsometry curves.

For optical images, different candidate algorithms were tested to delineate water: thresholding methods (OTSU, OTSU-Canny), unsupervised K-Means (in two configurations), machine Learning approaches (SVM, Random Forest). These were applied to primitives/indices and raw channels and combined with a pre-selection of water samples data exploiting GSW and some test criteria. For radar imagery, principally Sentinel1 data in SLC and GRD formats, several approaches based on K Means were also suggested and tested. Details about these methods are described in the Algorithm Theoretical Base Document available through the project website.

We applied these different methods to several lakes for which long term time series of satellite altimetry were available and in different landscape context: the Bosten and Sarykamish lakes in Central Asia, Argentino in south Patagonia, the Khanka lake in oriental Siberia, the Namco lake on the Tibetan Plateau, and Lake Illmen in West Russia. Additional tests were performed on the drying Lake Chad in the African Sahel.

The validation of the calculation of the LWE is not straightforward because ground truth is difficult to obtain. Various approaches were taken:

- Field survey: GPS mapping from a boat in the Lake Chad Archipelago Area
- Comparison to VHR imagery (e.g. Pleiades versus Sentinel-2)
- Cross-comparison between LWE solutions to quantify omission/commission rates
- Validation considering the hypsometry curve, inspecting the coherence between solutions

In April 2019, a field survey was organised in the Lake Chad archipelago where the lake shore was precisely mapped using a GPS receiver. This area has a complex morphology, with sandy flat zones and both flooding and non-flooding vegetation amongst shallow turbid water areas. The impact of these complex components was mainly visible on SAR imagery which is dependent on surface roughness (omission) and sandy flat areas (commission), while displacement of floating vegetation isles under wind influence can induce a pseudo LWE signal. Figure 1 below illustrates that optical imagery in such a complex case captures the complex water mask adequately.

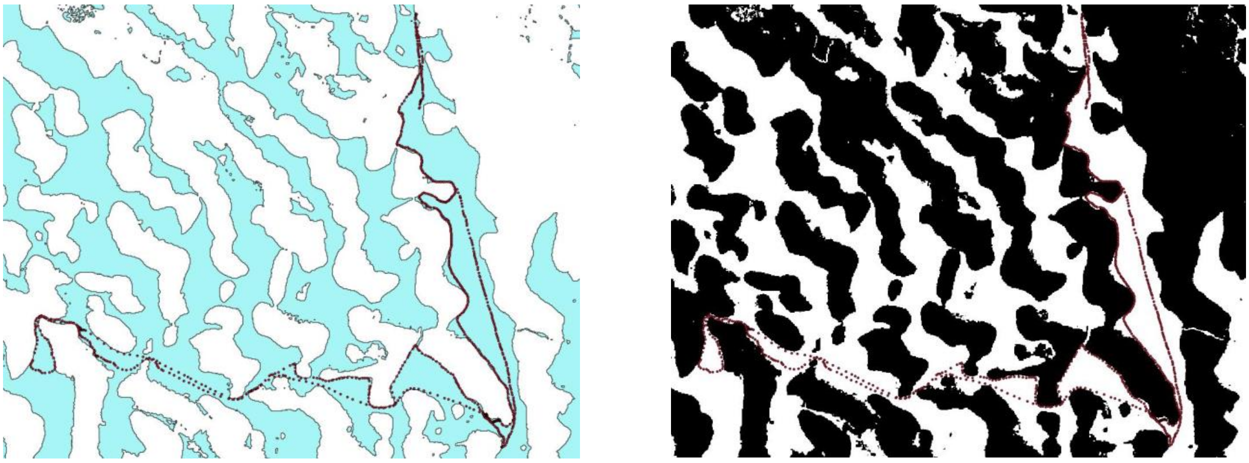

Figure 1: Lake Chad archipelago boat survey (red dots) and classification with simple NDWI threshold method on NDWI (left) and based on OTSU approach (right).

It is common to observe irregularities between the various LWE results. Figure 2 illustrates some of the differences found between SAR solutions and the optically derived water extent.

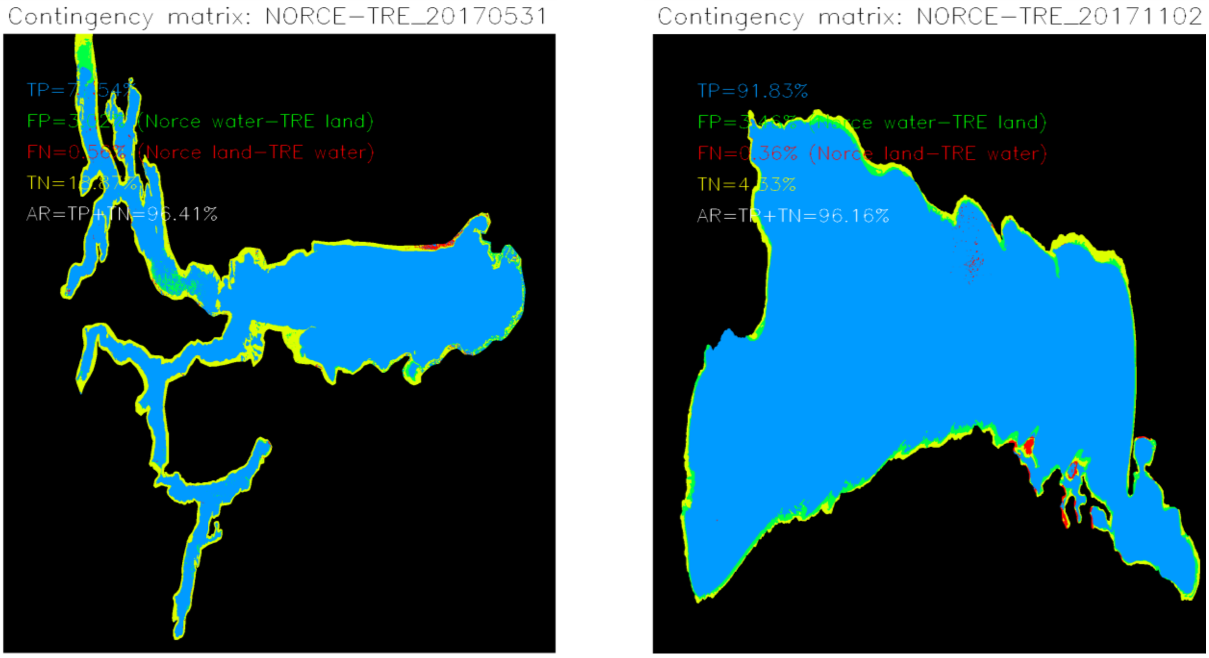

Figure 2: Lake Argentino (left) and Lake Bosten (right) show similar water masks between the SAR solutions, with wind induced surface roughness causing some observed differences. Lake Bosten shows differences concentrated in shallow areas at the mud/sandy wet transition from land to water.

In a further example, the very shallow Sarykamish Lake located in a sandy desert shows SAR images tended to overestimate the LWE compared to optical imagery.

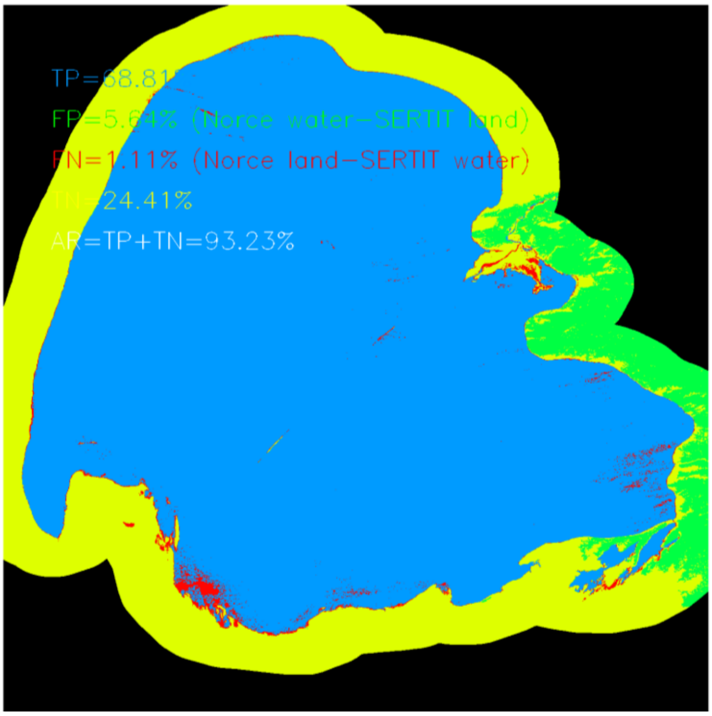

Figure 3: Sarykamish Lake is surrounded by a sandy flat area which is interpreted as water by SAR (green area). Wind or boats can also induce some omission, and rough water surfaces are interpreted as land by SAR (red areas).

Some examples of hypsometry curves are shown below. Figure 4 shows both an example of a coherent curve (Lake Bosten) and an example where methods diverge (Lake Namco). In the optimal cases where the four solutions agree, the final RMS is approximately 1% of the total extent of the lake, and the whole range of variations of level and extent is well covered.

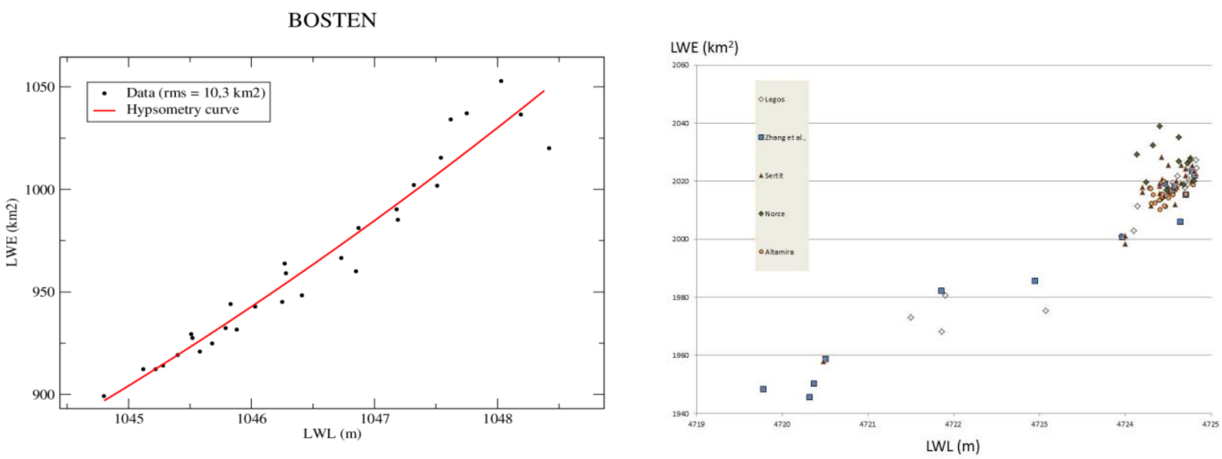

Figure 4: Hypsometry curves for Lake Bosten (left) using all LWE solutions (optical and radar) and for Lake Namco (right) showing a high degree of divergence.

In the most extreme cases it cannot be determined which solution performs best and therefore which LWE data to select for the final production of LWE time series. Lake Khanka, shown below, is an example where the hypsometry curves using SAR or optical data are fully separate, without any coherence even between the two SAR solutions and between the two optical solutions.

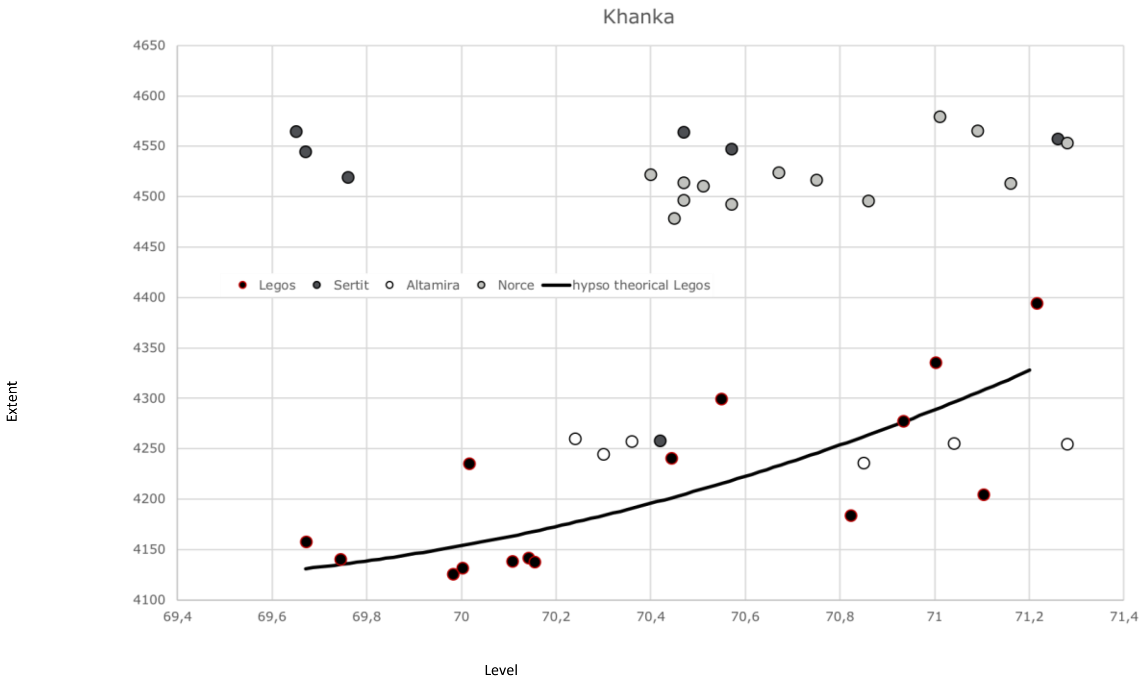

Figure 5: attempt to calculate hypsometry for Lake Khanka. We observe two groups of solutions, (Sertit/NORCE vs LEGOS/Altamira) with a strong bias. In the radar-based solutions LWE did not increase with LWL.

Conclusions

There is already evidence to suggest that a single solution to derive LWE may not be adequate. Further test cases will be useful to identify the context in which optical vs radar approaches are most useful to produce water delineations. Hypsometry nevertheless provides a suitable method to examine the coherence in LWE estimates calculated using satellite imagery, particularly in the absence of ground truth. This does not solve all incoherence but validates the use of the hypsometry curve to derive LWE from LWL, describing the level of uncertainty around the estimate at the same time.

Water delineation is a challenging task depending on multiple factors. Many observational and computational advances have led to a growing catalogue of hypsometry curves for the Lakes_cci project, but at the same time there remains ample room for future improvements in terms of processing, assessment and validation. The LWE product in the CRDP remain currently based on optical imagery analysis using a simple thresholding of the Normalized Difference Water Index. In future we would aim to incorporate additional methodologies described in this work.