GLOBAL LAKE SURFACE WATER TEMPERATURES FOR BIODIVERSITY, LIMNOLOGY, METEOROLOGY AND CLIMATE

Christopher J Merchant(1), Stuart N MacCallum(2), Aisling Layden(3) and Philippe Goryl(4)

(1) University of Reading, Harry Pitt Building, Reading, RG6 7BE, UK. Email: c.j.merchant@reading.ac.uk
(2) University of Edinburgh, Crew Building, King’s Buildings, Edinburgh EH9 3JN, UK. Email: s.maccallum@ed.ac.uk
(3) University of Edinburgh, Crew Building, King’s Buildings, Edinburgh EH9 3JN, UK. Email: a.layden@sms.ed.ac.uk
(4) ESA/ESRIN, Frascati, Italy. Email: philippe.goryl@esa.int

ABSTRACT

A reprocessing of Along Track Scanning Radiometer (ATSR) archives has been undertaken to create global lake surface water temperature (LSWT) datasets for a range of applications. The project, funded by the European Space Agency and called “ARC-Lake”, applies optimal estimation (OE) retrievals and probabilistic cloud screening methods to provide LSWT estimates for approximately 1000 lakes, globally, from 1991 to 2012. This methodology is generic (i.e. applicable to all lakes) as variations in physical properties such as elevation, salinity, and atmospheric conditions are accounted for through the forward modelling of observed radiances. The publicly available data products from ARC-Lake have been used, and have further potential for application, in a variety of fields, including meteorology, climate, and ecology. We will provide an overview of ARC-Lake from methodology through to applications of the LSWT datasets.

1. INTRODUCTION

Lakes are a vital component of Earth’s fresh water resources, and are of fundamental importance for terrestrial life. Lake water temperature is one of the key parameters determining ecological conditions within a lake, influencing chemical and biological processes. In addition to the impact on lake ecology, lake water temperatures determine air-water heat and moisture exchanges, and are therefore vital for understanding the hydrological cycle. Lake surface water temperatures (LSWT) and lake ice cover (LIC) observations therefore have potential environmental and meteorological applications for inland water management, lake modelling and numerical weather prediction.

The series of Along Track Scanning Radiometers (ATSRs) are visible and infrared (IR) imagers that provide excellent radiometric qualities and dual-view scanning capability. Because of this, the ATSRs have previously been exploited for sea surface temperature observations in the ATSR Reprocessing for Climate (ARC) project [1] with very accurate results [2].

LSWT retrieval is a distinct problem from SST or land surface temperature (LST) estimation in many ways, yet previous studies have applied or adapted SST or LST schemes to LSWT retrieval, using a number of different instruments: (A)ATSR [3], MODIS [4, 5], AVHRR [6], Landsat [7], ASTER [8], and SEVIRI [9].

The European Space Agency (ESA) established the ARC-Lake project (www.geos.ed.ac.uk/arclake) to adapt SST techniques for cloud and ice detection and for surface temperature retrieval to the problem of lakes. In the ARC-Lake project, the specific challenges of LSWT retrievals (including lake-boundary definition, and variations in lake elevation and salinity) are addressed systematically for lakes worldwide. The first two phases of this project (v1 and v2) focused on large lakes (surface area > 500 km²). The third and final phase of this project (v3), applies techniques developed in the initial phases to smaller lakes. This paper outlines the LSWT retrieval methods developed in the initial phases and highlights changes required for application to smaller lakes. Applications of the LSWT data and an overview of available data products are also presented.

2. METHODS

The methodology for LSWT retrievals, used in ARC-Lake v1 and v2, is described in detail in [10]. These techniques are also applied in ARC-Lake v3, with the exception of lake definition methods. This methodology and changes to it are outlined in the following sections.

2.1. Lake Definition

ARC-Lake v1 and v2 focused on the world’s “large” natural lakes, conventionally taken to be those in excess of 500 km² in surface area [11,12], with the addition, some slightly smaller lakes and reservoirs of interest. The location of these 263 initial targets is shown in Fig. 1. To enable satellite observations at a given location to be attributed to an individual lake, a mask was derived through consolidation of the NAVOCEANO gridded land/water mask (https://www.ghrsst.org/data/ghrsst-data-tools/land-water-mask) and level-1 polygons from the Global Lakes and Wetlands Database, GLWD [12]. The use of the GLWD polygons to provide unique lake IDs enables lakes with highly complex shapes to be represented, correctly associating multiple separate groups of water cells to the same ID.
Although implemented successfully in ARC-Lake v1 and v2, this lake-mask has the limitation of being static in time. Consequently, lakes known to have large seasonal or long-term variations in surface area were excluded. The static nature of the polygons, also results in some cases where polygons, defined for an earlier time period, do not accurately represent the lake area for the ATSR time period. To allow inclusion of smaller lakes and lakes with temporarily varying surface area, while minimizing the risk of land contamination, a water detection scheme was introduced to the processing for v3. Following initial application of this water detection scheme (§2.2) and matching with GLWD IDs, a set of 1628 initial targets was defined for ARC-Lake v3 (Fig. 2). In addition to smaller lakes and lakes with variable area, v3 also includes reservoirs.

### 2.2. Water Detection

The water detection scheme implemented in ARC-Lake v3 utilises the visible (VIS), near-IR (NIR), and short-wave-IR (SWIR) channels of the ATSR2 and AATSR instruments, through a combination of simple threshold tests. Two tests are based on normalized difference indices: the Modified Normalised Difference Water Index (MNDWI) [13] and the Normalised Difference Vegetation Index (NDVI) [14]. Further threshold tests are performed on the individual channel reflectances and on the 11 μm channel brightness temperature (BT).

This water detection scheme was applied to the full ATSR2/AATSR time period, recording counts of positive water detection (Ncounts) on a 1/120° grid. Using the GLWD polygons as a basis for lake location, but not a limit, GLWD IDs were assigned to each region identified as water.
Subsequent filtering based on Ncounts values was applied on a lake-by-lake basis to reduce potential land contamination from mixed land/water pixels. The resulting lake mask provides an estimate of the maximum extent of each lake over the ATSR2/AATSR time period. For day-time observations, this lake mask is used in conjunction with an active implementation of the water detection scheme, within the LSWT processing system. This allows the full extent of the lake to be represented while at the same time capturing temporal variations in surface area.

An alternative approach is required for night-time observations, as the active water detection relies on reflectance channels only available during periods of solar illumination. Annual masks of minimum lake extent are estimated from annual accumulations of water counts (Ncounts) on the 1/120° grid, using a similar Ncounts filtering method as applied to minimize potential land contamination in the maximum extent mask. For night-time observations, LSWT retrievals are only performed for pixels defined as water in the minimum extent mask for the year of observation.

Advantages of such an approach over a static mask are illustrated in Fig. 3, where an example of an annual minimum extent mask is compared to the maximum extent mask and the GLWD polygon, for the Aral Sea. Here, the GLWD polygon (red) has been defined for a time when the Aral Sea surface area was much larger than over the ATSR2/AATSR time period and therefore poorly represents the area over this period, captured by the maximum extent mask (blue). The difference between the maximum and minimum masks highlights the significant changes in the extent of the Aral Sea within, while also illustrating the ability to capture these changes through the annual estimates of minimum lake area derived from water detection counts.

2.3. Retrieval Algorithms

LSWT retrievals are performed using combinations of the three IR channels available on the ATSRs: 3.7, 10.8, and 12 μm. The ATSR IR radiometer is calibrated to high accuracy [15] and a dual-view geometry enables robust atmospheric correction [16]. Global coverage is provided every three days in the tropics, with more frequent observation possible at higher latitudes. All ATSRs have flown on platforms with stable late-morning orbits, yielding consistent overlap periods to support their application to global climate monitoring. Spatial resolution is ~1 km at the nadir, and restricts the size of lake surface features (and indeed the minimum size of lakes) that can be observed. However, this is still fine enough to enable some spatial variations in LSWT to be resolved.

The LSWT retrieval algorithm consists of cloud detection and temperature retrieval components, both of which depend upon forward modelling of clear-sky IR observations of the ATSRs. The radiative transfer model used is RTTOV [17], driven by the nearest numerical prediction (NWP) profile for the state of the atmosphere from the European Centre for Medium Range Weather Forecasting (ECMWF).

A major motivation for this study is the relatively inadequate observational information available to NWP systems on LSWT, which is becoming more important as lake dynamics are increasingly included in land-atmosphere interaction schemes. A corollary of this is that NWP is not a good source of surface temperature for forward modelling of the IR satellite observations. Following initialisation using a combination of monthly climatology from MODIS observations [18] and lake-mean temperature climatology simulations from the lake model FLake [19], empirical orthogonal function (EOF) techniques [20] are used to reconstruct a spatially complete time-series of LSWT from the sparse ATSR observations. An iteratively updated version of this is used as the prior LSWT in the forward modelling.

BTs seen for lakes by imagers such as the ATSRs are generally less than the true surface temperature, due to net absorption of IR radiance by the atmosphere and by the surface emissivity being less than unity. BT-LSWT relationships therefore depend on the altitude (affecting atmospheric impact) and salinity (affecting emissivity) of the lakes, both of which are highly variable across the target lakes. These variables are captured in the forward model, supporting the case for using forward modelling-based cloud detection and LSWT retrieval.
Inadequacies in cloud detection are linked to significant uncertainties. Typical threshold based cloud detection schemes for SST [Error! Bookmark not defined.] distinguish clear and cloudy skies through predefined tests on ranges for BTs and inter-channel BT differences. Ideally, these tests should be dependent on parameters such as surface temperature, atmospheric profiles and satellite zenith angle. Pre-specifying thresholds that are successful across a wide set of circumstances is challenging, particularly so for lakes, where the range of circumstances is greater than for ocean surfaces. Spatial coherence information is also used to distinguish clouds and water, and similar comments apply to determining these thresholds also.

Applying cloud detection for SST to lake bodies gives a useful result in some cases – particularly those with characteristics (e.g. size, altitude, and salinity) most similar to oceans (e.g. the Great Lakes and Caspian Sea). However, smaller lakes can display greater spatial variability than is typical for the ocean, which can trip spatial coherence tests, leading to false detection of cloud. BT-LSWT relationships are also changed by high spatial coherence tests, leading to false detection of variability than is typical for the ocean, which can trip the probablity above which L SWT retrievals are likely to return false detection of clouds.

For cloud detection in this study, we use a Bayesian approach [21,22,23] informed by the forward modelling discussed above. Comparing the expected (modelled) and observed BTs, the probability the observation being clear-sky is calculated (in the context of various relevant uncertainties). The only threshold in the scheme is that of the probability above which LSWT retrievals are made. Although the Bayesian approach adapts to the atmospheric conditions automatically (to the degree represented in NWP), the spatial coherence statistics used are still those developed for ARC SST [1].

Earlier work [24] established that LSWT retrieval using standard ATSR SST retrieval coefficients is prone, for some lakes, to retrieval biases of 0.5 K. This contrasts with a level of SST retrieval bias for ATSR that is generally <0.2 K. A possible solution of specifying lake-specific retrieval coefficients is not practical when considering the large number (1628) of lakes in v3.

The LSWT retrieval is therefore done by optimal estimation (OE). We use a simplified formulation of the inverse problem originally developed for SST observations from the Advanced Very High Resolution Radiometer (AVHRR) [25]. This formulation includes only LSWT and total column water vapour as retrieved (state) variables (all though full profile forward modelling is of course used). No radiance bias correction is yet derived for ATSR BTs, so the RTTOV simulated BTs are used “as is”.

2.4. Data Product Generation

A key aim of the ARC-Lake project is to provide spatially and temporally resolved data products of LSWT and LIC. Clear-sky LSWT retrievals and LIC estimates are averaged over longitude/latitude grid cells (1°20' x 1°20') for each day/night of observations. These averaged observations are stored (along with ancillary information) in NetCDF files on a per-lake basis (all years of observations for a single lake) and a global basis (all lakes for a single day).

Further spatial and temporal averaging is also applied, to generate spatially-resolved (1°20' x 1°20') and lake-mean climatology and time-series over a range of averaging intervals. Equivalent data products are also generated from the spatially complete EOF-based LSWT reconstructions. A summary of the possible variants of data products from ARC-Lake v2 is given in Tab. 1. Data products are freely available via the ARC-Lake project website (www.geos.ed.ac.uk/arclake).

Table 1. Overview ARC-Lake data products.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Observations / Reconstructions</td>
</tr>
<tr>
<td>Coverage</td>
<td>Per-lake / Global</td>
</tr>
<tr>
<td>Time</td>
<td>Day / Night</td>
</tr>
<tr>
<td>Spatial Resolution</td>
<td>0.05° grid / Lake-mean</td>
</tr>
<tr>
<td>Temporal Averaging Type</td>
<td>Climatology / Time-series</td>
</tr>
<tr>
<td>Temporal Averaging Period</td>
<td>Seasonal / Monthly / Twice-monthly / Daily</td>
</tr>
</tbody>
</table>

3. LSWT VALIDATION

Results of a validation study, where OE LSWT retrievals were compared with in situ observations and operational retrievals (designed principally for SST), are presented in Tab. 2, and Tab. 3 for AATSR. In situ observations are from buoys, located at 54 sites across 18 lakes (and comprise those data that were freely available during ARC-Lake v2). All results are for direct comparison of satellite and buoy observations: no adjustment is made for the skin-bulk effect.

Retrieval biases (relative to in situ measurements) are lower for OE retrievals, relative to operational retrievals, and are of the order expected for skin-to-depth comparisons. RSDs are comparable across the retrievals but SDs from the OE scheme are significantly lower, indicating a reduction in outliers compared to the operational scheme. This is despite the Bayesian cloud screening passing a far greater number of matches as clear (~66% more). The increased number of match-ups mainly arises from the lower end of the temperature range, where the operational cloud tests seem most likely to return false detection of clouds.
4. APPLICATIONS

4.1. Climatology

Perhaps the most obvious application of the ARC-Lake data products is in improving our knowledge of basic lake climatology. Most lakes are poorly monitored in situ; therefore, ATSR observations offer a far more globally complete lake climatology, since 1991. Figs. 4 and 5 give examples of the climatological information that can be determined from the ARC-Lake data. Fig. 4 shows the lake-mean seasonal LSWT range for 258 target lakes, illustrating the global nature of the ATSR coverage. Low LSWT ranges are observed in the tropics, with peak ranges occurring at around 45° N. Moving to higher latitudes the LSWT range generally decreases again as the lakes receive less heating following the frozen period. Fig. 5 shows the lake-mean seasonal LSWT cycle for Lake Balaton (Hungary), comparing ARC-Lake climatology to that from MODIS and from the online lake model, Flake [19]. Broadly good agreement is observed here but for other cases ARC-Lake provides a more reasonable seasonal climatology than MODIS.

Table 2. Validation statistics for AATSR using operational dual-view LSWT retrievals and SADIST cloud masking.

<table>
<thead>
<tr>
<th>Day / Night</th>
<th>Channels (μm)</th>
<th>AATSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Day 11, 12</td>
<td>1969</td>
<td>-0.59</td>
</tr>
<tr>
<td>Night 3.7, 11, 12</td>
<td>1934</td>
<td>-0.23</td>
</tr>
</tbody>
</table>

Table 3. Validation statistics for AATSR using OE dual-view LSWT retrievals and Bayesian cloud masking.

<table>
<thead>
<tr>
<th>Day / Night</th>
<th>Channels (μm)</th>
<th>AATSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Day 11, 12</td>
<td>3273</td>
<td>-0.34</td>
</tr>
<tr>
<td>Night 3.7, 11, 12</td>
<td>3220</td>
<td>-0.15</td>
</tr>
</tbody>
</table>

4.2. Lake Modelling

Another potential application of the LSWT data from ARC-Lake is in the assessment and improvement of lake models. A trial study on 21 cold water lakes, comparing simulations using the full FLake model [26] (driven by ECMWF data) with lake-mean observations from ARC-Lake, revealed FLake consistently overestimates the July-August-September (JAS) LSWT by ~4 K (also shown in Fig. 6). This trial study also demonstrated that FLake underestimates the cold phase (where LSWT < 1 °C) of the lakes by approximately 4 weeks, relative to ARC-Lake observations.

A tuning scheme was developed to improve the representation of the annual LSWT cycle in FLake, based on the variation of three FLake parameters: effective depth (Z_D), albedo (α), and light extinction coefficient (κ). Through comparison of four metrics (Tab. 4) between FLake and observed LSWTs, a set of optimal run conditions (Z_D, α, and κ) were determined for 160 cold water lakes covering a wide range of lake characteristics (e.g. depth, altitude, salinity, etc.).

Table 4. Metrics used to determine optimal FLake run conditions.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Parameter(s) of influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean absolute error (MAE) in daily LSWT</td>
<td>Z_D, α, and κ</td>
</tr>
<tr>
<td>Mean JAS LSWT</td>
<td>κ</td>
</tr>
<tr>
<td>1 °C cooling day</td>
<td>Z_D</td>
</tr>
<tr>
<td>1 °C warming day</td>
<td>α</td>
</tr>
</tbody>
</table>

Following optimisation, improvements were observed across all metrics (Tab.5) and further highlighted in the time-series of LSWT for Lake Simcoe (Fig. 6). Applying this optimization to FLake provides potential for LSWTs to be modelled more accurately over a time period beyond that of the ATSR instruments.
Figure 6. Comparison of time-series of LSWT from optimized and non-optimized versions of FLake with ARC-Lake observations for Lake Simcoe.

Table 5. Results of FLake optimization scheme for 160 cold water lakes.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Non-optimized (21 lakes)</th>
<th>Normal (135 lakes)</th>
<th>Shallow (25 lakes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE (°C)</td>
<td>3.07</td>
<td>0.74</td>
<td>1.11</td>
</tr>
<tr>
<td>Mean JAS (°C)</td>
<td>-3.71</td>
<td>0.01</td>
<td>-0.34</td>
</tr>
<tr>
<td>1 °C cooling day</td>
<td>12</td>
<td>-1</td>
<td>-1.3</td>
</tr>
<tr>
<td>1 °C warming day</td>
<td>-27.1</td>
<td>0.5</td>
<td>-0.5</td>
</tr>
</tbody>
</table>

4.3. Numerical Weather Prediction

LSWT observations of the form available from ARC-Lake, if made operational, have potential to improve NWP, through assimilation. This is demonstrated in Fig. 7, where ARC-Lake observations are compared with NWP data and in situ observations for Lake Malawi. A climatological cycle is represented in the NWP data but the magnitude of the NWP temperatures are significantly different to those observed in ARC-Lake. There is good agreement between ARC-Lake and in situ observations, giving confidence that the ARC-Lake observations provide accurate LSWTs.

Following a validation study where good agreement was found between trends in satellite and in situ LSWTs (a root mean square discrepancy of 0.035°C yr-1 across trends for 15 locations in six lakes), time-series of average seasonal LSWT anomaly were determined for three latitude regions and are shown in Fig. 8. The largest warming trend (0.05 °C yr-1) is observed in northern hemisphere lakes with trends around zero observed for both the tropics and the southern hemisphere. This is in broad agreement with a previous study [28] as discussed in more detail in [29].

5. ARC-LAKE V3

The methodology outlined in §2 has been applied to the initial v3 target list of 1628 lakes and reservoirs, defined through analysis of water detection counts over the ATSR2/AATSR time period. As the VIS and NIR channels are not available on ATSR1, the water
detection scheme is not possible for this instrument and therefore the v3 dataset will only cover 1995 to 2012.

Comparison of Figs. 1 and 2 highlights the improved global coverage offered by ARC-Lake v3 data. This larger target list includes lakes with area of order 50 km$^2$ and Fig. 9 illustrates the potential of the v3 data to provide valuable LSWT information for lakes of this size. Intermittent lakes are also included in the v3 target list. Fig. 10 demonstrates the effectiveness of the water detection scheme in capturing periods of drought for such lakes, while still returning a reliable LSWT time-series for periods where water is present.

![Figure 9. Daily lake-mean LSWT observations from ARC-Lake v3 for lake ID=3218, Argentina (area = 55 km$^2$).](image)

![Figure 10. Daily lake-mean LSWT observations from ARC-Lake v3 for Lake Hindmarsh, Australia (area = 130 km$^2$).](image)

6. CONCLUSIONS

The success of optimal estimation (OE) LSWT temperature retrievals and Bayesian cloud screening method have been demonstrated for 258 large target lakes in ARC-Lake v1 and v2, giving improved performance over techniques designed principally for SST. No empirical tuning of retrievals has been used, so that the new satellite LSWTs are independent of in situ observations, being based on the physics of radiative transfer.

Several potential applications of the ARC-Lake LSWT products have been demonstrated: climatology, lake modelling, NWP, and trend assessment. Such potential will be enhanced with the release of ARC-Lake v3 products, which extend the dataset to smaller lakes, reservoirs, and lakes with variable surface area. ARC-Lake v2 products are available and v3 products will soon be available from [www.geos.ed.ac.uk/arclake](http://www.geos.ed.ac.uk/arclake).

Future work will see the ARC-Lake methods used to provide LSWTs in an ecological and physical lake observation system [http://www.globolakes.ac.uk/](http://www.globolakes.ac.uk/).

7. ACKNOWLEDGEMENTS

The authors gratefully acknowledge that this work is funded by the European Space Agency under contract 22184/09/1-OL. The lake processor is a modification of the processor built during the ATSR Reprocessing for Climate (ARC) project that was funded by the UK Natural Environment Research Council (NE/D001129/1) and the Department of Energy and Climate Change (CPEG 31).

8. REFERENCES


