ATSR Reprocessing for Climate Lake Surface Temperature: ARC-Lake

Validation Report – v1.0

Title: ATSR Reprocessing for Climate Lake Surface Temperature: ARC-Lake: Validation Report – v1.0

Document Number: ARC-Lake-Validation-Report-v1.0

Issue: 1

Revision: 1.0

Date: 8 October 2010

Signature Table

	Name	Function	Company	Signature	Date
Prepared	S MacCallum C Merchant	Researcher Reader	University of Edinburgh		8 Oct 2010
Approved					
Released					

Document Change Record

Version	Date	Author	Description
1.0	8 Oct 2010	SM, CM	

Table of Contents

List of Figures

1 INTRODUCTION

1.1 Acronyms and Abbreviations

AATSR	*Advanced ATSR*
ARC	*ATSR Reprocessing for Climate*
ATSR	*Along-Track Scanning Radiometer*
BT	*Brightness Temperature*
LIC	*Lake Ice Concentration*
Lake ST	*Lake Surface Temperature*
MD	*Match-up Dataset*
MDB	*Match-up DataBase*
NEΔT	*Noise Equivalent Differential Temperature*
NWP	*Numerical Weather Prediction*
OE	*Optimal Estimation*
RMSD	*Root-Mean-Square Deviation*
RT	*Radiative Transfer*
RTM	*Radiative Transfer Model*
RTTOV	*Radiative Transfer for TOVs (a fast RTM)*
SD	*Standard Deviation*
TCWV	*Total Column Water Vapour*
ToA	*Top of Atmosphere*

1.2 Purpose and Scope

This document is a Validation Report for the Lake Surface Temperature (Lake ST) and Lake Ice Concentration (LIC) products, generated from Along-Track Scanning Radiometer (ATSR) imagery, for the ARC-Lake project.

In terms of scope, this Validation Report covers version 1.0 products, the first public product release, for ATSR-2 and Advanced ATSR (AATSR).

1.3 Validation Report Overview

The Validation Report provides the following:

- an assessment of the performance of the Lake ST product, in quantitative terms, relative to *in situ* observations
- qualitative illustrations of the Lake ST retrievals from case study analysis at instrument resolution
- a quantitative assessment of the LIC product relative to ice charts from mixed sources (*in situ*, aircraft, and satellite)
- qualitative illustrations of the performance of the ice detection algorithm from case study analysis at instrument resolution

2 LAKE SURFACE TEMPERATURE (LAKE ST) PRODUCT

2.1 Introduction

Two methods of assessment of the Lake ST retrieval algorithm are employed: analysis of performance for case study images at full AATSR resolution and point comparisons with *in situ* observations. These twin approaches are adopted to provide qualitative visual assessments of algorithm performance across spatial domains and to provide a quantitative measure of the overall performance relative to in situ observations. Validation against *in situ* observations is described in this section while the case study analysis is presented in §3. In all cases, results are presented for ARC-Lake v1.0 retrievals.

2.2 Data

A match-up dataset (MD) was constructed from the in situ temperature data currently available to the ARC-Lake project. This consists of 52 observation locations covering 16 of the Phase One lakes. Details of the in situ data are given in Table 1. As the in situ data are from a variety of sources, with different formats, considerable effort has been put in to consolidate this data to a standard format for use in ARC-Lake, and to apply quality control measures.

Source	Lake Names (number of observation locations)
National Data Buoy Center (NDBC)	Superior (3), Huron (4), Michigan (2), Erie (1), Ontario (1)
Fisheries and Oceans Canada (FOC)	Superior (1), Huron (2), Great Slave (2), Erie (2), Winnipeg (3), Ontario (3), Woods (1), Saint Clair (1), Nipissing (1), Simcoe (1)
Swedish University of Agricultural Sciences (SLU)	Vanern (5), Vattern (2), Malaren (13)
GLobal Lake Ecological Observatory Network (GLEON)	Balaton (1)
King's College London (KCL)	Nyasa (3)

Table 1. Details of in situ data consolidated into the ARC-Lake MD.

2.3 Methods

Clear-sky Lake ST retrievals are averaged over a 5x5 pixel box, equivalent to the resolution of the ARC-Lake "per-lake" (PL) and "daily global" (DG) products (MacCallum & Merchant, 2010b), centred on the buoy location. Matching against *in situ* observations is performed spatially (within 1 km) and temporally (within 3 hours) to create Match-up Databases (MDBs) for both ATSR-2 and AATSR. In total there are ~15500 match-ups for ATSR-2 and ~17500 for AATSR. These totals are over two orders of magnitude less than the number of buoy-satellite match-ups available over the oceans. The locations of the match-ups for each instrument are shown in Figure 1. Note that for AATSR, we have as yet no validation data outside of midlatitude regions.

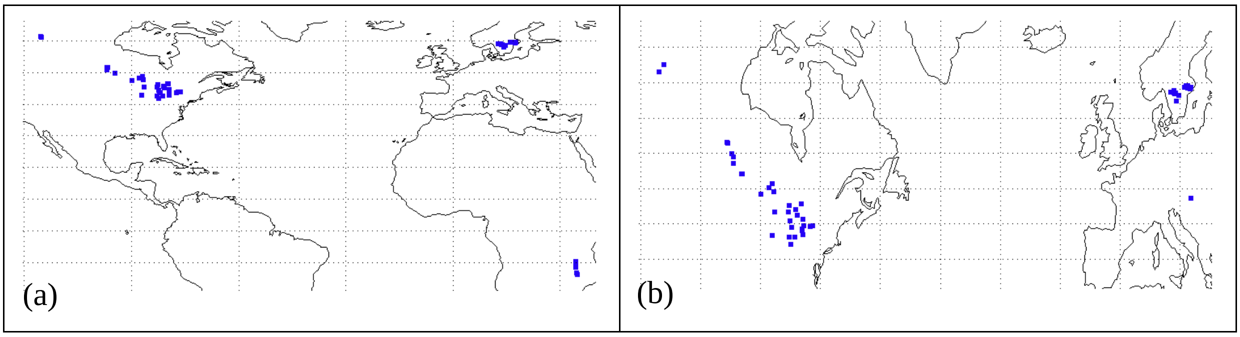


Figure 1. Locations of *in situ* observations with match-ups to (a) ATSR-2 and (b) AATSR.

Lake STs are compared to the *in situ* observations for the various cloud masks and retrieval schemes. Day time and night time retrievals are considered separately for a number of different channel/view combinations: nadir-view 2-channel (N2), nadir-view 3-channel (N3), dual-view 2-channel (D2), dual-view 3-channel (D3). A summary of the results are presented in §2.4 for all match-ups with at least one clear-sky pixel.

2.4 Results

Basic statistics from the comparisons with *in situ* observations are presented in Table 2 for operational retrievals using the SADIST cloud mask and Table 3 for the ARC-Lake OE retrievals using Bayesian (maximum channel set) cloud screening.

The percentage of the total number of match-ups where there is at least one clear-sky observation ranges from ~5-14% across the different cloud masks for ATSR-2 and from ~8-15% for AATSR. For AATSR the number of match-ups is always greater when the Bayesian rather than the SADIST cloud mask is used, with the Bayesian mask returning ~67-77% more match-ups. A similar but less pronounced result is observed for ATSR-2, where the Bayesian mask returns ~40-45% more match-ups. Observations made in the case study analysis (§3) support this result, where over-masking of clear-sky areas is seen to be more prevalent in the SADIST mask.

Mean satellite-in-situ differences for each channel combination are an indicator of 'retrieval bias' for the different 'algorithms' (although the surface skin effect and near-surface stratification can also cause some mean differences). For ideal retrievals, we would expect mean differences relative to in situ of the order of -0.2 K for night (due to skin effect) and closer to zero or slightly positive for day. In the day time case, there will be a combination of skin effect and average stratification between measurement depth and surface reflected in the mean difference; but we don't really have a good insight at present into the degree of near-surface/diurnal stratification to be expected in different lakes.

Retrieval biases (relative to in situ measurements) range from 0.12 K to 0.88 K (day) and -0.52 K to 0.18 K (night) for operational AATSR retrievals using the SADIST cloud mask, with RSDs ranging from 0.52 K to 0.65 K. The range of biases is reduced to 0.23 K to 0.37 K (day) and -0.48 K to -0.28 K (night) when the ARC-Lake OE retrieval and Bayesian cloud mask are used. For ideal retrievals, we would expect mean differences relative to in situ of the order of -0.2 K for night (due to skin effect) and closer to zero or slightly positive for day (combination of skin effect and average stratification between measurement depth and surface). RSDs from ARC-Lake are equal to or lower than those from the operational retrieval for all retrieval types, with the RSD for the N2 case ~0.2 K lower. SDs for all but the N2 retrieval are also lower, and in the case of dual-view night-time retrievals by more than 0.4 K.

These results demonstrate the advantages of the ARC-Lake OE retrieval and Bayesian cloud screening over operational equivalents. The first key advantage is the increased number of observations. This offers potentially greatly improved coverage of the lakes, thus yielding a more spatially and temporally complete data record. Secondly, there is a much greater degree of self consistency across the different channel/view combinations (i.e. a significantly smaller range of biases across the different retrievals). Thirdly, there is comparable or slightly reduced noise in the retrievals, demonstrated by comparable or lower RSDs. The SDs are generally reduced by more than the RSDs, indicating a reduction in outliers associated with cloud or ice detection failures. The consistency of biases and RSDs across retrieval schemes is of particular importance for extending the ARC-Lake project to include ATSR-1, due to the failure of the 3.7 µm channel on this instrument.

Day / Night	View / Channels	ATSR-2				AATSR			
		N	Bias	SD	RSD	N	Bias	SD	RSD
Day	N2	813	1.02	1.09	0.62	1519	0.88	1.03	0.65
Day	D2	812	0.24	1.09	0.56	1520	0.12	1.03	0.55
Night	N2	1528	0.35	0.81	0.68	1499	0.18	1.18	0.63
Night	N3	1529	0.12	0.71	0.51	1498	-0.32	1.07	0.52
Night	D2	1524	-0.49	0.78	0.61	1495	-0.52	1.89	0.60
Night	D3	1529	-0.32	0.73	0.53	1496	-0.41	1.20	0.53

Table 2. Validation statistics for Lake ST retrievals performed using SADIST cloud masking and operational Lake ST retrieval scheme. View/channels indicates the views (N = nadir, D = dual) and the number of channels used in the retrieval (e.g. N2 is a nadir-view, twin-channel retrieval).

Day / Night	View / Channels	ATSR-2				AATSR			
		N	Bias	SD	RSD	N	Bias	SD	RSD
Day	N2	1146	0.51	1.36	0.53	2535	0.23	0.97	0.47
Day	D2	1149	0.56	1.05	0.55	2539	0.37	0.89	0.51
Night	N2	2212	-0.25	1.30	0.59	2646	-0.48	1.55	0.59
Night	N3	2226	-0.17	0.78	0.52	2653	-0.31	0.77	0.51
Night	D2	2225	-0.21	0.87	0.59	2653	-0.38	0.84	0.60
Night	D3	2225	-0.14	0.79	0.54	2653	-0.28	0.78	0.53

Table 3. Validation statistics for Lake ST retrievals performed using ARC-Lake Bayesian (maximum channel set) cloud masking and ARC-Lake OE Lake ST retrieval scheme. View/channels indicates the views (N = nadir, D = dual) and the number of channels used in the retrieval (e.g. N2 is a nadir-view, twin-channel retrieval).

A similar picture is observed for ATSR-2 retrievals. The ARC-Lake scheme returns almost 50% more match-ups with comparable (or reduced) RSD and a greatly improved self consistency for different channel combinations. The operational results have biases ranging from 0.24 K to 1.02 K (day) and -0.49 K to 0.35 K (night) for operational ATSR-2 retrievals using the SADIST cloud mask. The ARC-Lake scheme returns bias ranges of 0.51 K to 0.56 K (day) and -0.25 K to -0.14 K (day). RSDs are again typically between 0.5 K and 0.6 K.

Scatter plots for dual-view maximum channel set (i.e. D2 and D3) retrievals for day-time and night-time retrievals are shown in Figure 2, for the operational (labelled “ATS”) and ARC-Lake (labelled “OE”). The increased number of match-ups from ARC-Lake is seen to arise particularly from the lower end of the temperature range, where the SADIST threshold tests are most likely to return false positives. All the retrievals show some trend in difference against in situ temperature, quantified by the slope, m, shown on the plots. For example, ARC-Lake night match-ups using the D3 channels has m = -0.017 K K^{-1}, meaning that over the 25 K range of lake temperatures in the data, the satellite is warmer relative to in situ observations by 0.4 K for the lowest temperatures compared to the warmest temperatures. The trends for the D2 and D3 channel combinations are relatively consistent for ARC-Lake, whereas the corresponding trends in the operational case are of opposite sign.

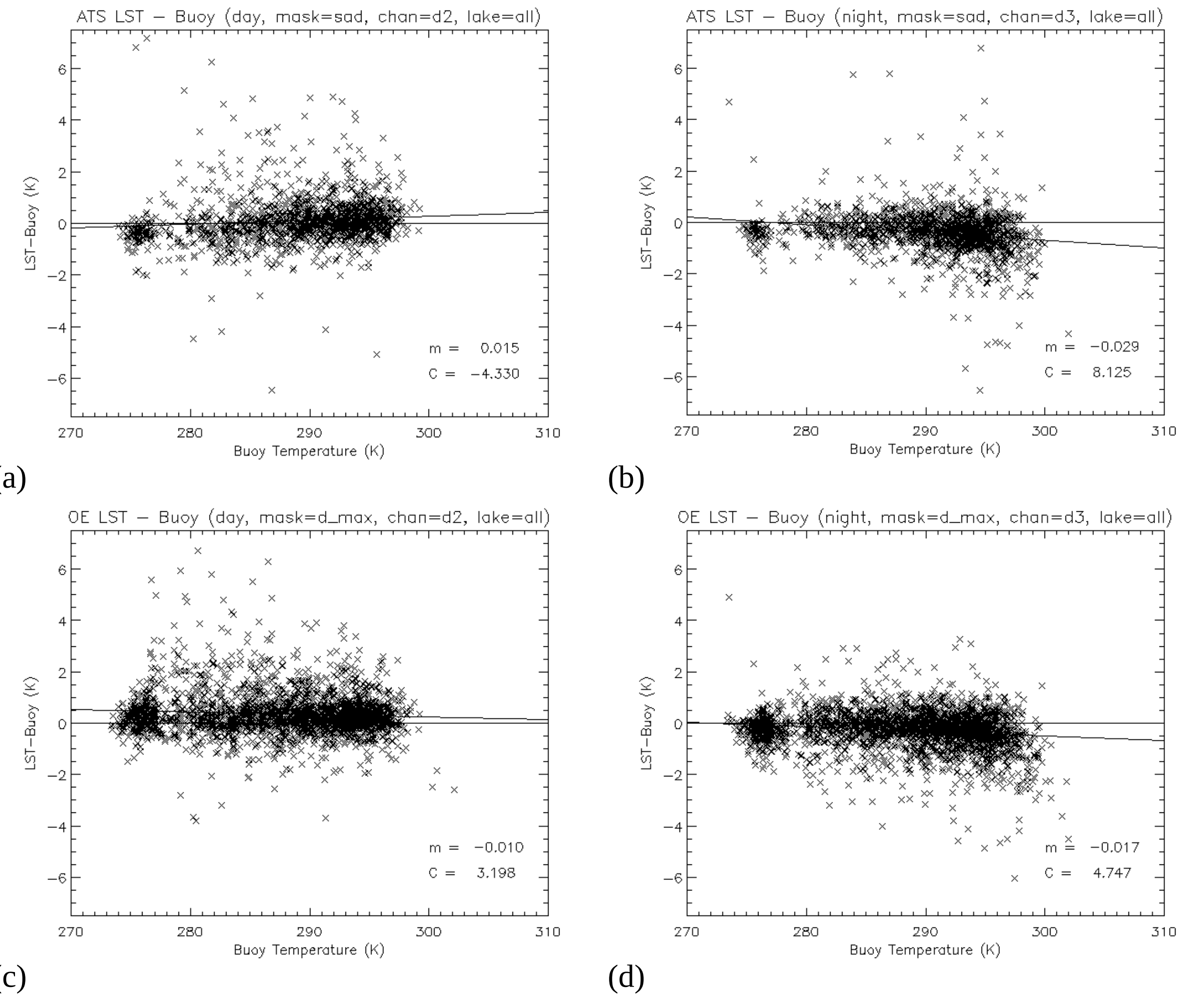


Figure 2. Lake ST-Buoy differences against buoy temperature for AATSR. (a) and (b) operational SADIST day and night, (c) and (d) ARC-Lake day and night.

Similar results are observed for ATSR-2 (Figure 3). As for AATSR, the greater number of match-ups returned by the ARC-Lake scheme is most apparent at low temperatures, and trends in bias with temperature are more consistent across retrieval schemes.

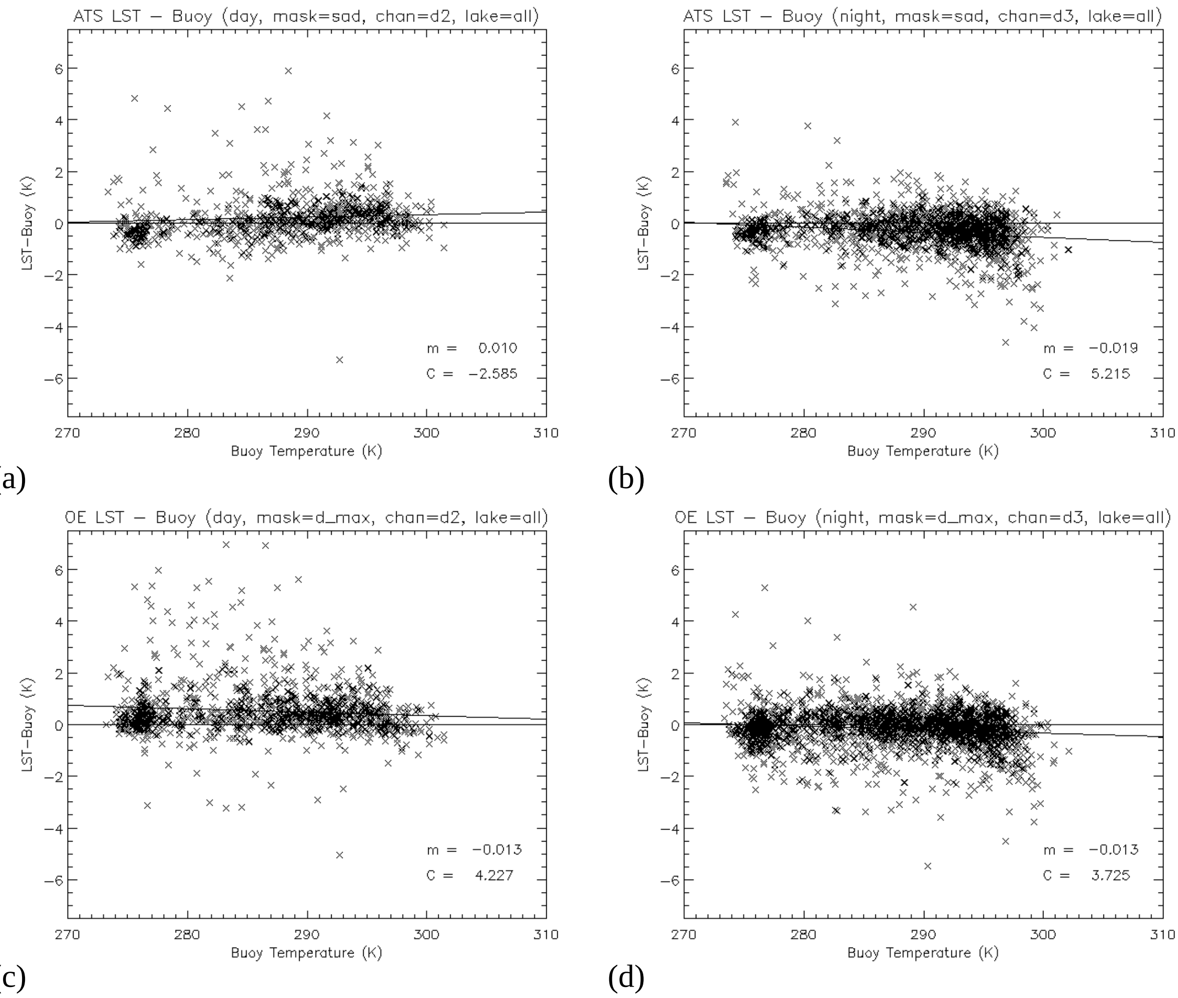
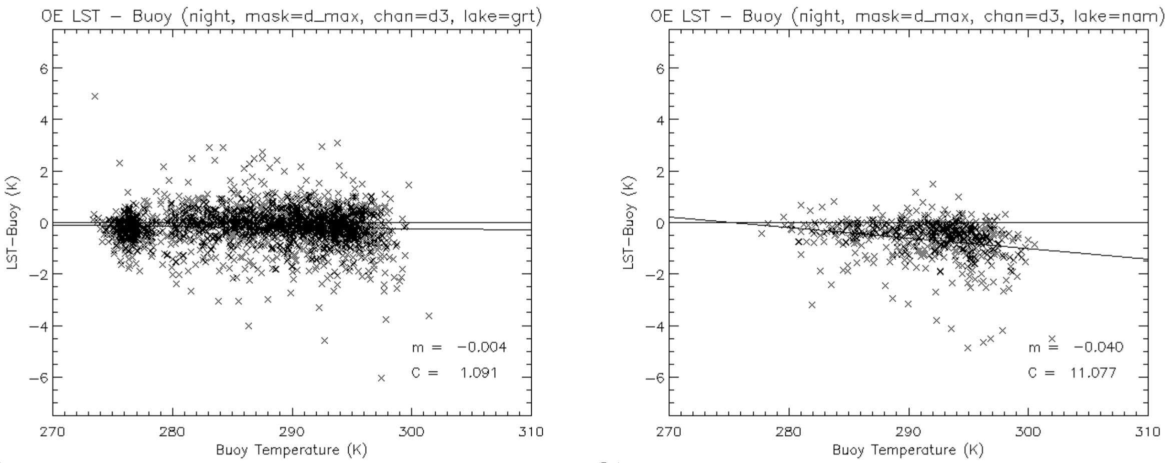


Figure 3. Lake ST-Buoy differences against buoy temperature for ATSR-2. (a) and (b) operational SADIST day and night, (c) and (d) ARC-Lake day and night.

The validation results presented in Figure 2 and Figure 3 are dominated by the Great Lakes as they are more extensively monitored *in situ* (Table 1). This is demonstrated in Figure 4 and Table 4, where results are shown for the Great Lakes and all other North American lakes separately, for night time ARC-Lake retrievals only. There is a significant difference in bias between these subsets of data but random errors remain similar. This highlights the need for further *in situ* observations, covering a greater variety of lakes and locations, to be included in the ARC-Lake validation data set. Efforts to obtain additional *in situ* data are ongoing.



(a) (b)

Figure 4. Lake ST-Buoy differences against buoy temperature for night time AATSR. (a) The Great Lakes, (b) all other North American lakes

Lakes	Day / Night	View / Channels	AATSR			
			N	Bias	SD	RSD
Great lakes	Night	D3	2031	-0.169	0.731	0.529
Other North American	Night	D3	581	-0.627	0.754	0.492

Table 4. Lake ST-Buoy validation statistics for AATSR, corresponding to Table 4. A comparison of the D3 night-time retrievals over the Great Lakes and all other North American lakes.

3 LAKE ST CASE STUDIES

3.1 Introduction

For the qualitative analysis a set of 12 AATSR scenes were selected as case studies. These were chosen such as to provide examples over a range of different geographical locations, altitudes, lake sizes, and meteorological conditions. The case studies selected cover the following lakes: Great Slave, the Great Lakes, Titicaca, Onega, Ladoga, Vanern, Vattern, Victoria, Superior, Bay, Winnipeg, Huron, Nyasa, and a number of other smaller lakes.

Qualitative visual assessment of cloud masking and Lake ST retrievals have been carried out for all 12 AATSR case studies. Performance is variable across the case studies and across cloud masks / retrieval algorithms within case studies. Examples from the case study analysis, along with some general observations, are presented in the following sections.

Two retrieval schemes are assessed: the operational scheme (ATS), and the ARC-Lake optimal estimation method (OE). Two types of cloud mask are also assessed: the SADIST threshold based cloud mask and the probabilistic Bayesian cloud mask. The SADIST cloud mask is used in the operational retrieval scheme (ATS), while Bayesian cloud masking is used for ARC-Lake OE retrievals. Like the Lake ST retrievals themselves, Bayesian cloud masking can be performed using different view/channel set combinations. Results are presented for two of these combinations: nadir-view twin-channel retrieval with nadir only minimum channel set Bayesian cloud screening, and dual-view twin-channel retrieval with dual-view maximum channel set Bayesian cloud screening.

3.2 Case 1

The first example (Figure 5) covers part of the Great Lakes region, including lakes Huron (5), Erie (12) and Ontario (15), on 2nd April 2008. At this time of year, temperatures on these lakes are close to 0°C and ice may still be present. In the false colour image (Figure 5a) all three of the Great Lakes can be seen to be largely clear with only small patches of cloud (white), mainly across Lake Huron. Although mostly clear of cloud, there is a significant area of ice cover (yellow-brown) visible on Lake Erie. This AATSR scene is included as a case study as it provides a test of the retrieval scheme at the lowest extreme of the temperature range and also a test of the ice detection algorithm

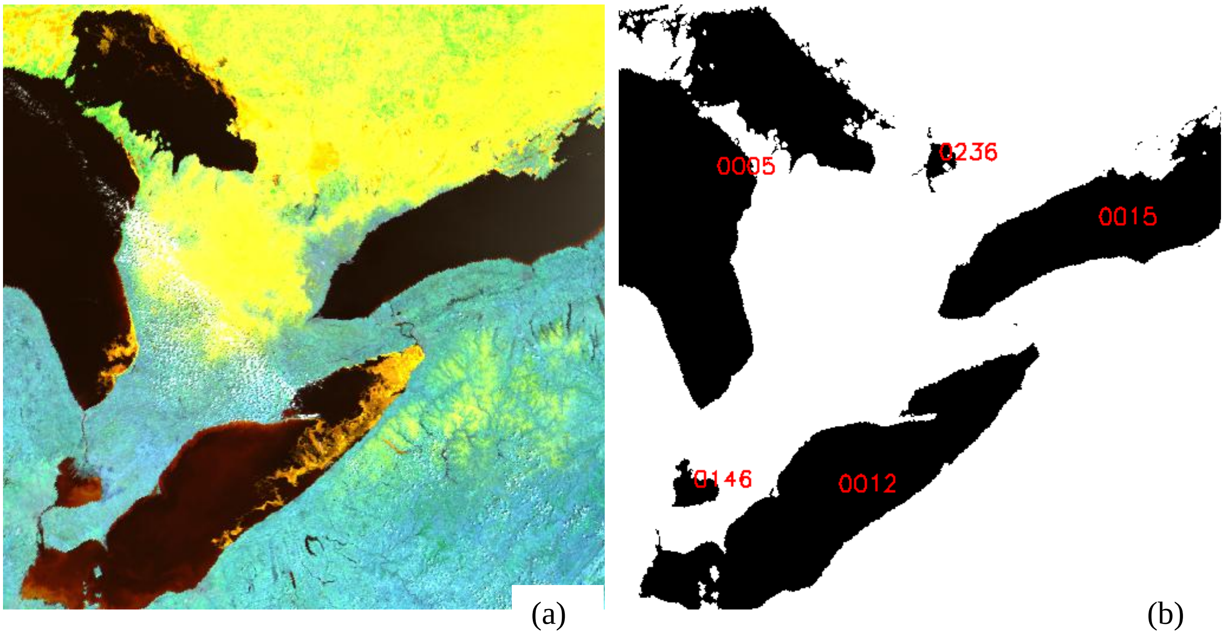


Figure 5. Case study example for the Great Lakes region on 02/04/08. (a) False colour image from AATSR reflectance channels (0.66 µm, 0.87 µm and 1.6 µm). (b) Land/water mask showing the lake locations with corresponding ARC-Lake IDs.

In Figure 6 (b) to (d) the Lake STs are shown for the various cloud detection and retrieval schemes, with the cloud mask in black. Figure 6a shows the prior Lake ST field used in the ARC-Lake OE retrievals (MacCallum & Merchant, 2010a). The SADIST cloud detection scheme (b) masks almost all of the lake surfaces. The nadir Bayesian cloud mask (c) incorrectly masks around half of the lake area. This over-masking is present in some areas of all three lakes where the prior Lake ST appears (in comparison with the OE retrievals in (d)) to be a degree or so too warm or too cool. This error in the prior Lake ST has a lesser impact on the dual-view cloud detection (d), which correctly passes the majority of lake pixels as clear sky. Some ice-affected pixels are also passed as clear sky in (d), but are flagged correctly by the ice detection test discussed below (and therefore their retrieved temperatures would not contribute to the product).

The over-masking seen in the SADIST scheme is observed some of the other case studies, while in other cases the SADIST cloud mask is comparable the Bayesian cloud masks. Both forms of the Bayesian cloud mask return fewer falsely flagged areas than the SADIST mask, but there is still significant over-masking in the nadir-view case (Figure 6c) and some over-masking in the dual-view case (Figure 6d), predominantly around the lake edges.

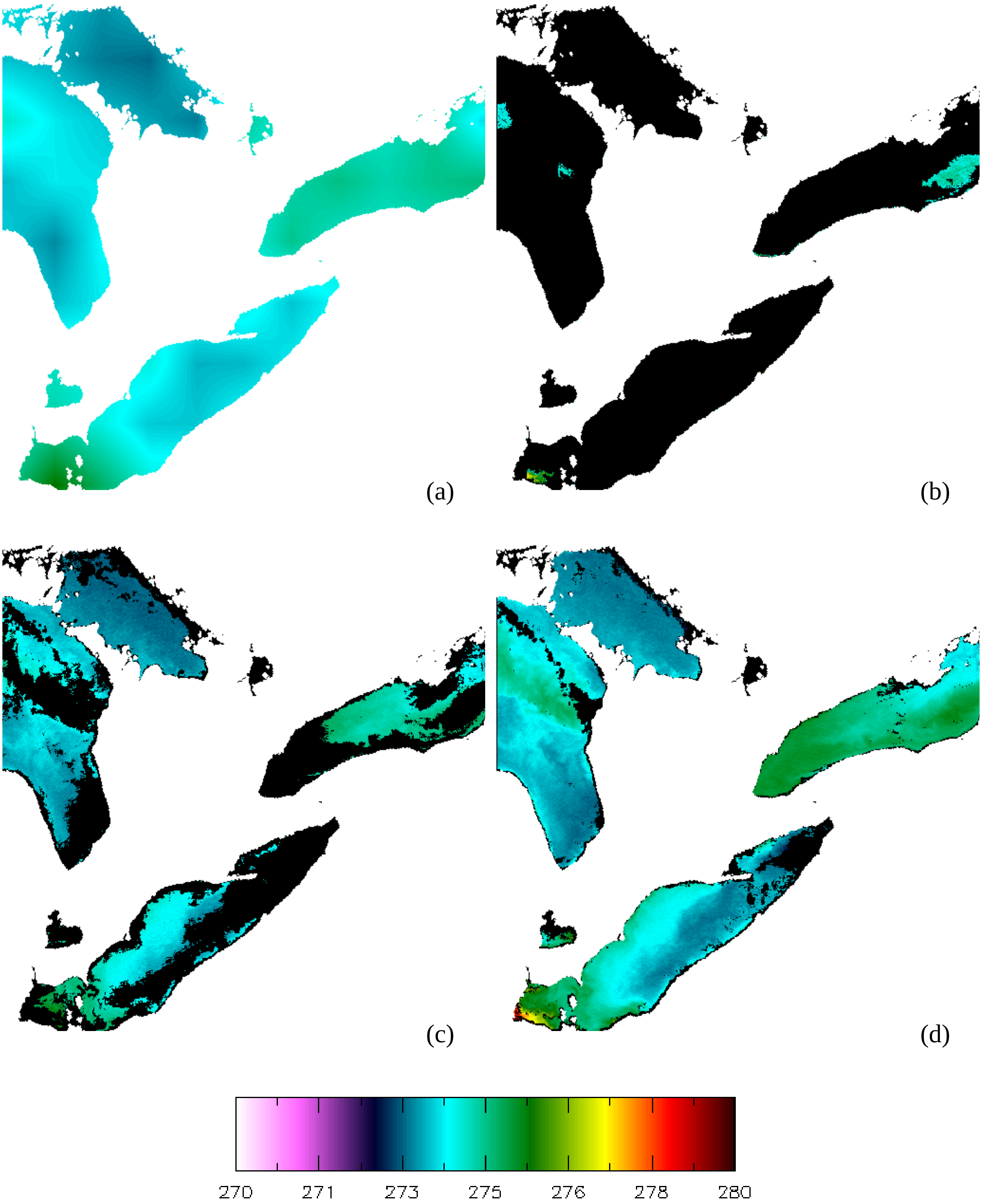


Figure 6. Results for the AATSR scene over the Great Lakes (Figure 5). (a) Prior Lake ST field. (b) Operational SST retrieval with SADIST cloud mask in black. (c) ARC-Lake OE Lake ST (nadir-view, twin-channel) with Bayesian cloud screening (nadir-view, minimum channel set). (d) OE Lake ST (dual-view, twin-channel) with

Bayesian cloud screening (dual-view, maximum channel set). The colourbar applies to all figures and in all cases the cloud mask is represented as black.

In terms of the temperatures retrieved, good consistency is observed across channel/view combinations in the ARC-Lake OE scheme, with lake STs within 0.15 K. The operational retrievals are less consistent with each other with temperature differences of ~1.0 K observed. As well as being more consistent across retrievals, the ARC-Lake OE scheme also provides a spatially smoother temperature product, enabling thermal features to be distinguished more easily. Again this result is observed across the case studies.

Comparison of retrieved Lake STs against *in situ* observations is also possible for this case study, with buoy measurements available at four locations on across Lake Erie and Lake Ontario. At all four of these locations the SADIST cloud mask flags the immediate area around the buoy (5x5 pixel box) as cloud. The Bayesian masks however, flag three of these areas as clear-sky, and comparisons are made between OE Lake ST and the *in situ* observations. These comparisons agree with the results for the full match-up dataset Table 3, with a mean (satellite-buoy) bias of 0.42 K across the 3 match-ups and ~0.1 K difference between nadir and dual view retrievals.

One of the reasons for selecting this case study was the clear presence of ice in the visible imagery (Figure 5a). Ice can be seen to the north east of Lake Erie and in small areas in the south and north of Lake Huron. The result of the ARC-Lake ice detection scheme for this scene is presented in Figure 7. In this case study the ice detection algorithm performs reasonably well, correctly masking the major ice visible on Lakes Erie and Huron. Further analysis of the ARC-Lake ice product is presented in §4 and §5.

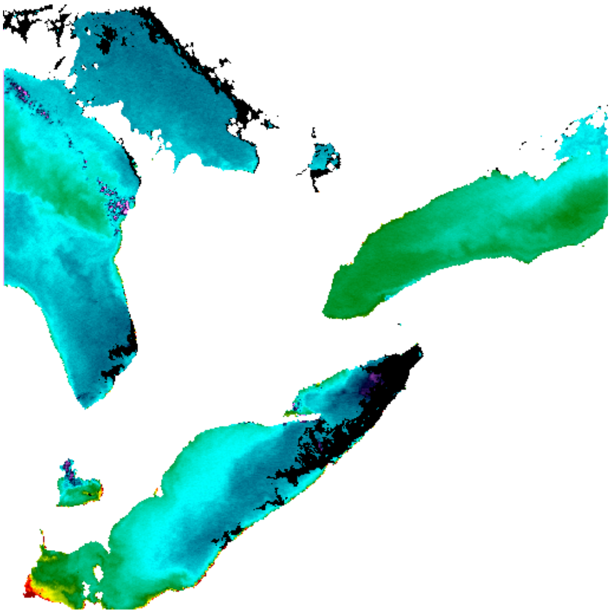
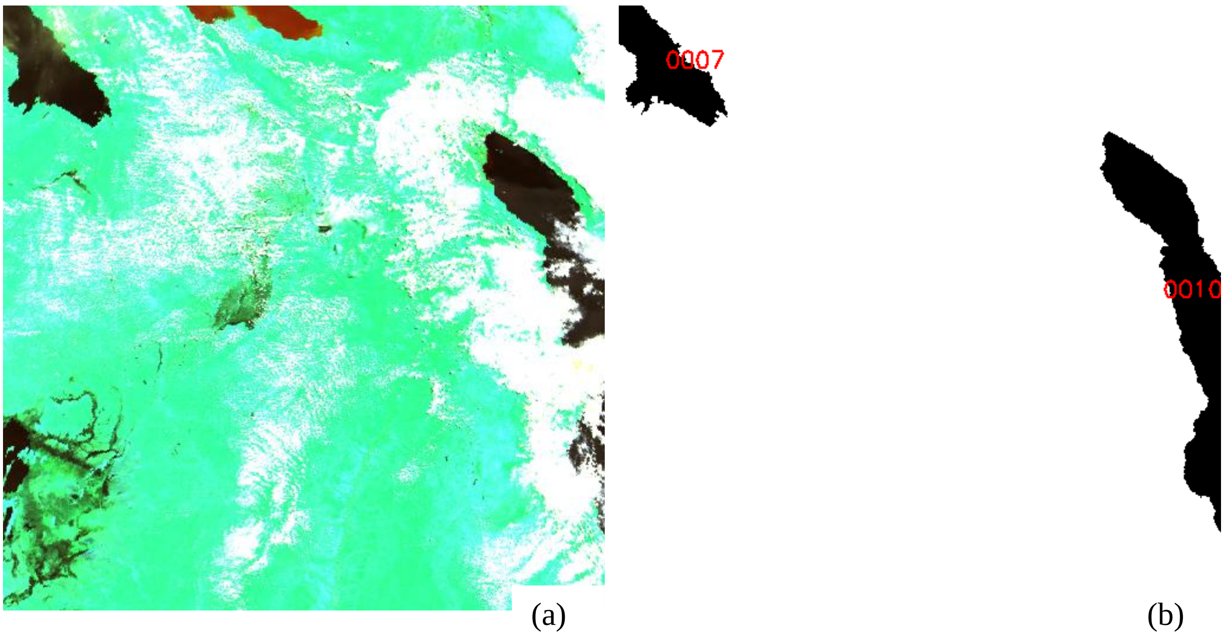


Figure 7. Results for the AATSR scene over the Great Lakes (Figure 5). OE Lake ST (dual-view, twin-channel) with ARC-Lake ice screening (MacCallum & Merchant, 2010a). Note that no cloud screening has been applied, and cloud related Lake ST biases are evident. The colourbar given in Figure 6 applies and the ice mask is represented as black.

3.3 Case 2

The second example (Figure 8) covers part of Lakes Nyasa (10) and Tanganyika (7), on 2nd April 2008. There is relatively little seasonal variation in surface temperature over these lakes, with temperatures typically only varying by ~4 K over the year with an annual mean of ~298 K. In the false colour image (Figure 8a) there is a mixture of cumulus and thin cirrus across the lakes, and some regions of clear sky. Thin cirrus that is only just discernible in the image covers most of the area of Tangayika, This AATSR scene is included as a case study as it provides a test of the retrieval scheme at the higher end of the temperature range and provides a more challenging test of the cloud detection scheme.



(a) (b)

Figure 8. Case study example for Lake Nyasa. (a) False colour image from AATSR reflectance channels (0.66 μm, 0.87 μm and 1.6 μm). (b) Land/water mask showing the lake locations with corresponding ARC-Lake IDs.

As for the first case study, the Lake STs for the various retrieval schemes are shown in Figure 9 (b) to (d) the Lake STs are shown for the various retrieval schemes, along with the appropriate cloud mask. Figure 9 (a) shows the prior Lake ST field used in the ARC-Lake OE retrievals (MacCallum & Merchant 2010a). Again there are significant differences between the SADIST and Bayesian cloud masks. In this case study, both Bayesian methods (nadir minimum channel set and dual maximum channel set) return similar cloud masks, and a reasonable representation of the cloud cover visible in the reflectance imagery.

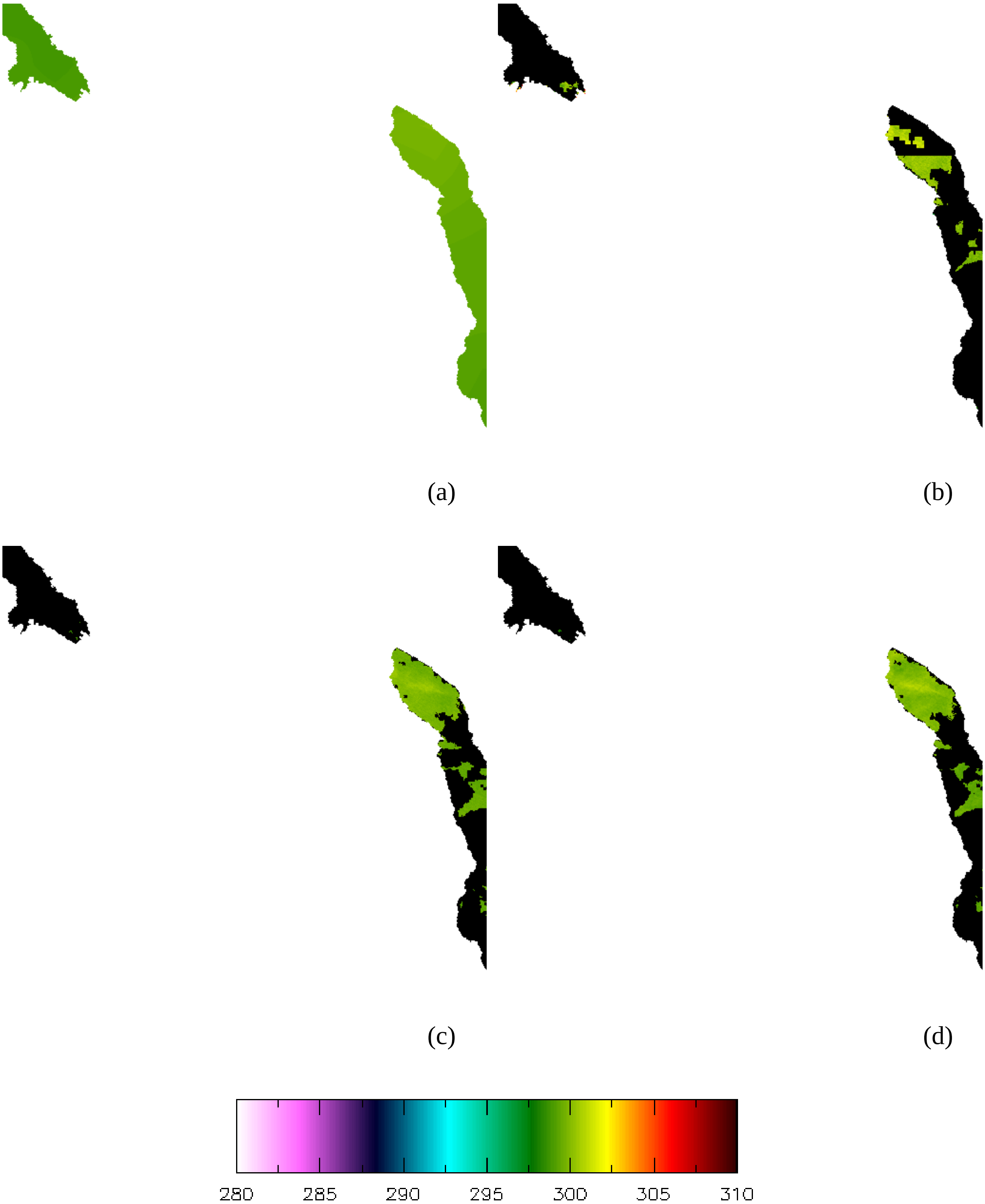


Figure 9. Results for the AATSR scene over Lake Nyasa (Figure 8). (a) Prior Lake ST field. (b) Operational SST retrieval with SADIST cloud mask. (c) ARC-Lake OE Lake ST (nadir-view, twin-channel) with Bayesian cloud screening (nadir-view, minimum channel set). (d) OE Lake ST (dual-view, twin-channel) with Bayesian

cloud screening (dual-view, maximum channel set). The colourbar applies to all figures and in all cases the cloud mask is represented as black.

Retrieved Lake ST values from the ARC-Lake OE scheme are again consistent across channel/view combinations, within 0.1 K. The operational retrievals are again less consistent with each other with temperature differences of ~0.8 K observed. As in case study 1, the ARC-Lake ST product is also more consistent spatially.

Comparison of retrieved Lake STs against *in situ* observations is not possible for this case study, as *in situ* observations are extremely sparse and limited to ATSR-2.

4 LAKE ICE CONCENTRATION (LIC) PRODUCT

4.1 Introduction

As for the Lake ST products, two methods of assessment of the LIC retrieval algorithm are employed: qualitative analysis of performance for case study images at full AATSR resolution and quantitative comparisons with *in situ* observations. Quantitative validation against *in situ* observations is described in this section while the qualitative case study analysis is presented in §5. In all cases, results are presented for ARC-Lake v1.0 retrievals

4.2 Data

Quantitative assessment of the LIC product is conducted using ice observations obtained from the NOAA Great Lakes Ice Atlas (Assel, 2003) and the National Ice Center (http://www.natice.noaa.gov). Both of these sources provide ice charts for the Great Lakes: Superior, Huron, Michigan, Erie, and Ontario. These ice charts, described by Assel (1983)**,** are a blend of observations from different data sources (ship, shore, aircraft, and satellite) and cover the full lifetime of the ATSR series of instruments. Ice concentration data are provided as the fraction of a unit of lake surface area that is completely covered with ice, where each grid cell has a nominal resolution of 2.5 km x 2.5 km (Assel *et al*, 2002). Ice charts are provided for each winter season (Dec 1st to April 30th approx.) for the full lifetime of the ATSR instruments. This data is used to provide a quantitative indicator of the performance of the ice detection algorithm, under clear-sky conditions. The ARC-Lake ice detection algorithm is based on the Normalised Difference Snow Index (NDSI) of Hall *et al* (1995) and is described in MacCallum & Merchant (2010a). This is applied on a pixel-by-pixel basis and a count of the number of ice pixels in each 0.05°x0.05° cell is stored in the ARC-Lake v1.0 products.

Note that, since the test uses reflectance channels, it is only available for day time scenes (under clear skies). For the Great Lakes validated here, day time imagery is available throughout the year, but for extreme northern lakes, there may be periods where no ice detection can be done.

4.3 Methods

For quantitative comparison with the ARC-Lake LIC product the digitised ice chart data from the Great Lakes Ice Atlas and the National Ice Center are averaged to the same 0.05°x0.05° as the v1.0 ARC-Lake LIC product. Ice chart data were compared with the ARC-Lake LIC product only for days where the ice chart data was taken from observations (i.e. interpolated data was excluded), and only for days where there was at least one clear-sky ATSR observation of the lake available (either open-water or ice). As the ice charts are only provided during the period where the lakes are likely to be (partially) frozen, the inclusion of days where no ice is present should not unduly bias the results of this comparison towards successful detection of open-water.

Four surface categories are defined for this analysis: open-water, ice-covered, mixed-water, and mixed-ice. The definitions of these are given in Table 5. Each grid cell is classified as one of these categories in both the ice charts and the LIC product. When calculating the percentage ice-cover in the LIC product, only clear-sky observations are included (i.e. ice cover = N ice pixels / (N ice pixels + N water pixels), and cloudy pixels are not considered). Where the number of non-cloudy pixels in a cell is low, the sampling-related error in the LIC for the cell can be large. For each day where ice chart and LIC products have matched, grid cells containing at least one clear-sky observation were considered in the analysis.

Category Name	Short Name	Percentage Ice-Cover
Open-water	OW	0
Mixed-water	MW	1-15
Mixed-ice	MI	15-85
Ice-covered	IC	>85

Table 5. Categories of ice-cover used in analysis of ARC-Lake LIC product.

4.4 Results

Percentage ice-cover values are compared between ice charts and the LIC product for each of the five Great Lakes, over all observations for ATSR-2 and AATSR independently. The results of this analysis, considered over all the Great Lakes, are presented in Table 6 and Table 7 for AATSR and ATSR-2 respectively. These tables show the percentage of cells where each pair of surface categories (Table 5) is observed between the ARC-Lake LIC product and the ice charts (e.g for AATSR observations over all the Great Lakes (Table 6), 2.52 % of cells are classed as open-water in the ARC-Lake LIC product and as ice-covered in the ice charts).

Ice Charts \ ARC-Lake	0 %	1-15 %	15-85 %	>85 %
0 %	63.72	0.81	0.67	0.19
1-15 %	8.75	0.61	0.65	0.66
15-85 %	2.43	0.91	2.29	2.63
>85 %	2.52	1.05	3.31	8.80

Table 6. Results of comparison of ARC-Lake LIC product from AATSR with ice charts over all the Great Lakes. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 152487 grid cells.

Ice Charts \ ARC-Lake	0 %	1-15 %	15-85 %	>85 %
0 %	65.77	1.08	1.15	0.42
1-15 %	5.86	0.26	0.36	0.24
15-85 %	3.17	1.07	2.21	1.70
>85 %	3.55	0.96	3.27	8.95

Table 7. Results of comparison of ARC-Lake LIC product from ATSR-2 with ice charts over all the Great Lakes. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 188359 grid cells.

The results of the analysis, in Table 6 and Table 7, demonstrate reasonable levels of agreement between the ARC-Lake LIC product and the ice charts. Ideally the diagonal elements of these tables should be large, particularly so for the both OW and both IC cases where there should be less ambiguity about the surface type. By summing equivalent elements of these tables, a summary of the agreement between the two ice products can be obtained. Table 8 provides this summary, where four levels of agreement have been defined, corresponding to the number of surface categories by which the two ice products disagree (e.g. level 0 indicates both products class the cell as the same category, level 3 indicates one product classes the cell as OW while the other classes it as IC).

Level of Disagreement	AATSR	ATSR-2
0 (Agree)	75.43	77.19
1	17.05	13.34
2	4.81	5.51
3 (Disagree)	2.71	3.96

Table 8. Summary of the level of disagreement between ARC-Lake LIC and ice charts. The level number indicates the number of surface categories by which the two ice products disagree (e.g. level 0 indicates both products class the cell as the same category, level 3 indicates one product classes the cell as OW while the other classes it as IC).

Reasonable levels of agreement are observed between the ARC-Lake LIC product and the ice charts (Table 8). For both ATSR instruments, the LIC product classifies the surface in the same category as the ice-chart in over 75% of the cells assessed, mainly due to very reliable identification of the 0% ice class. The percentage of cells which agree to within one class exceeds 90 % for both sensors.

Closer assessment of Table 6 and Table 7 reveals that the ARC-Lake LIC product underestimates the amount of ice-cover, relative to the ice charts. This can be seen by considering the elements of the tables on either side of the main diagonal: elements above represent cells where more ice is observed in the LIC product than the ice chart, while the opposite is true for elements below the diagonal. Assessing the results in this way reveals that the ARC-Lake LIC product may fail to detect ice coverage accurately in ~18-24% of cells, while at the same time falsely flagging open-water as ice in ~5% of cells, for both ATSR-2 and AATSR.

The full breakdown of results for each of the Great Lakes is given in the Table 10 to Table 19 in the appendix (§8).

5 LIC CASE STUDIES

5.1 Introduction

For the qualitative analysis the ARC-Lake ice detection algorithm is assessed visually at the instrument pixel resolution (~1 km at nadir). The set of case study images defined in §3 contains three scenes where ice is visible in the reflectance imagery, covering the Great Lakes, Lake Winnipeg, and Lakes Onega and Ladoga. In addition to these scenes, the ice cover product is assessed at the pixel resolution for days, identified in the analysis with ice charts (§4), where significant ice-cover was present. Examples from this analysis, along with some general observations, are presented in the following sections.

5.2 Example from Case Studies

An example of the ARC-Lake LIC product at the pixel resolution has been presented in Figure 7, above. In that case study, the ARC-Lake product is seen to provide a reasonable representation of the ice cover over the Great Lakes, visible in the reflectance channel image (Figure 5).

A further example from the case studies is shown in Figure 10. As in Figure 7, the ice detection algorithm appears to work quite effectively in this scene, successfully masking the ice visible in the north of Lake Onega while correctly identifying the small ice-free region in the south of the lake as open water. In the third of the case studies (Lake Winnipeg on 01/01/08) where ice is visible in the reflectance imagery (not shown), the ice mask fails to detect any ice in the cloud free areas of the scene. In this case the BTs in the scene are such that they have triggered a gross cloud test that, failure of which bypasses all further processing for the pixel (Bayesian cloud detection, ice detection module and retrieval). This is not appropriate ARC-Lake processing, and will be changed prior to future releases of the ARC-Lake LIC products.

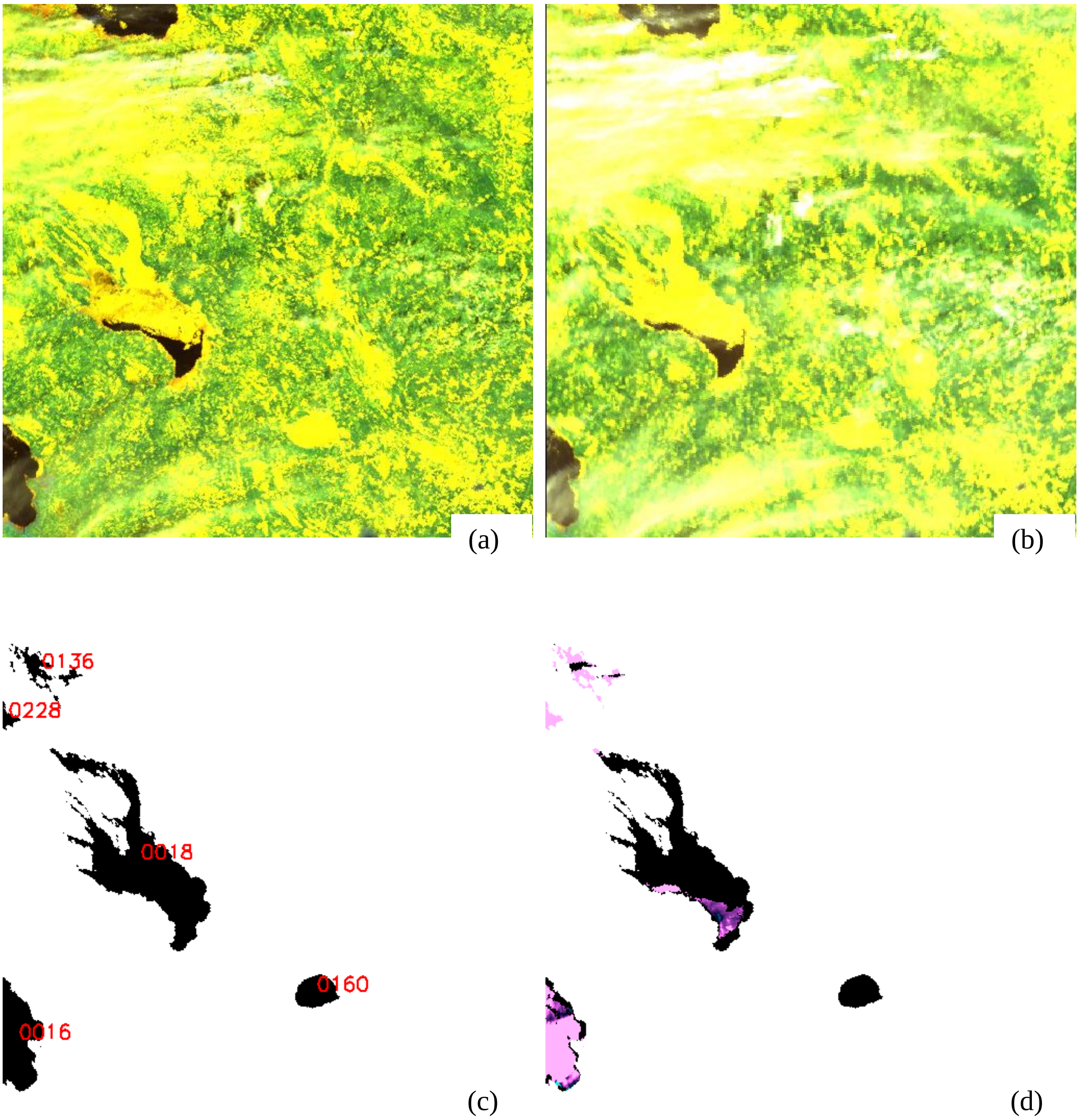


Figure 10. Ice cover case study over Lake Onega (18) on 03/04/2008. (a) False colour image from AATSR reflectance channels (0.66 µm, 0.87 µm and 1.6 µm) for nadir view. (b) As (a) but for forward view. (c) Land/water mask showing the lake locations with corresponding ARC-Lake IDs. (d) Ice-mask with pixels flagged as ice represented as black ; other colours indicate either cloud or Lake ST.

5.3 Examples from Ice Chart Analysis

The ice chart data described in §4 was used to identify days on which there was significant ice-cover over the Great Lakes. Days on which there was also good clear-sky coverage in the ATSR observations were then identified and a visual assessment of the ARC-Lake ice detection at full pixel resolution was performed. Two examples of this analysis are provided in Figures 11 to 14. The ice chart for ice chart for 21st January 1997 is shown in Figure 11. On this day, Lake Erie was judged to be largely frozen and varying degrees of ice cover were attributed around the edges of the other Great Lakes.

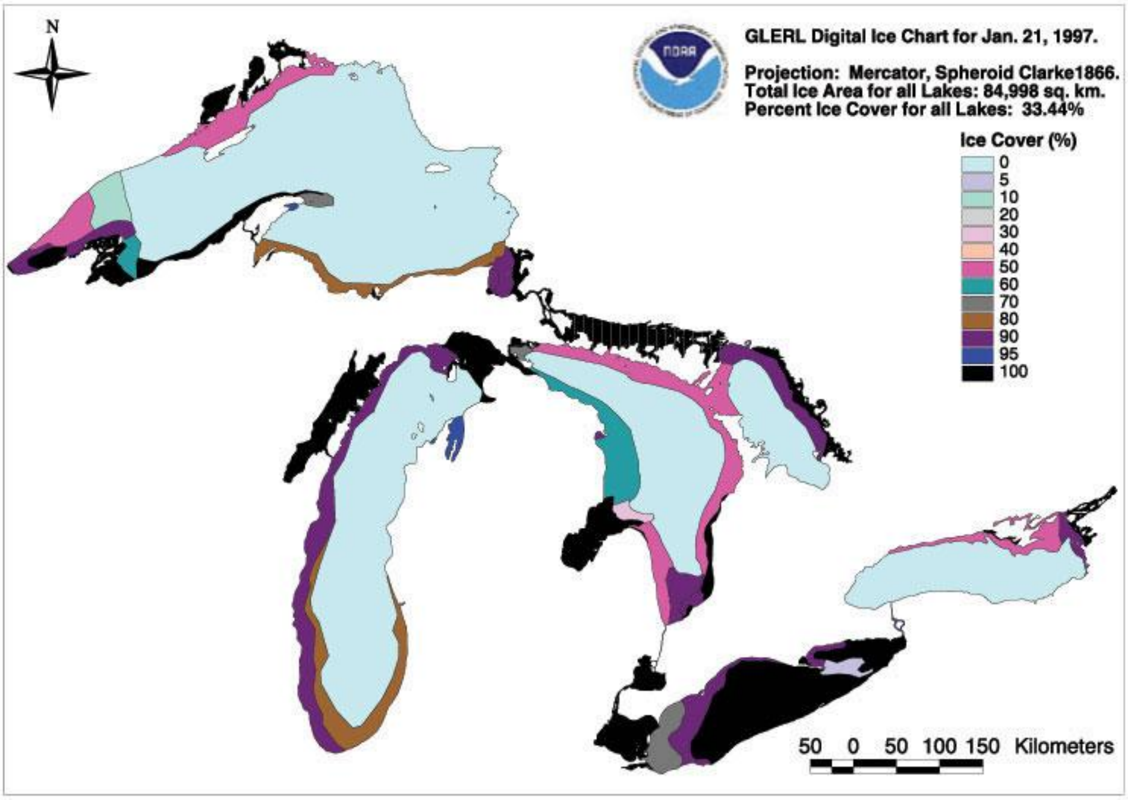
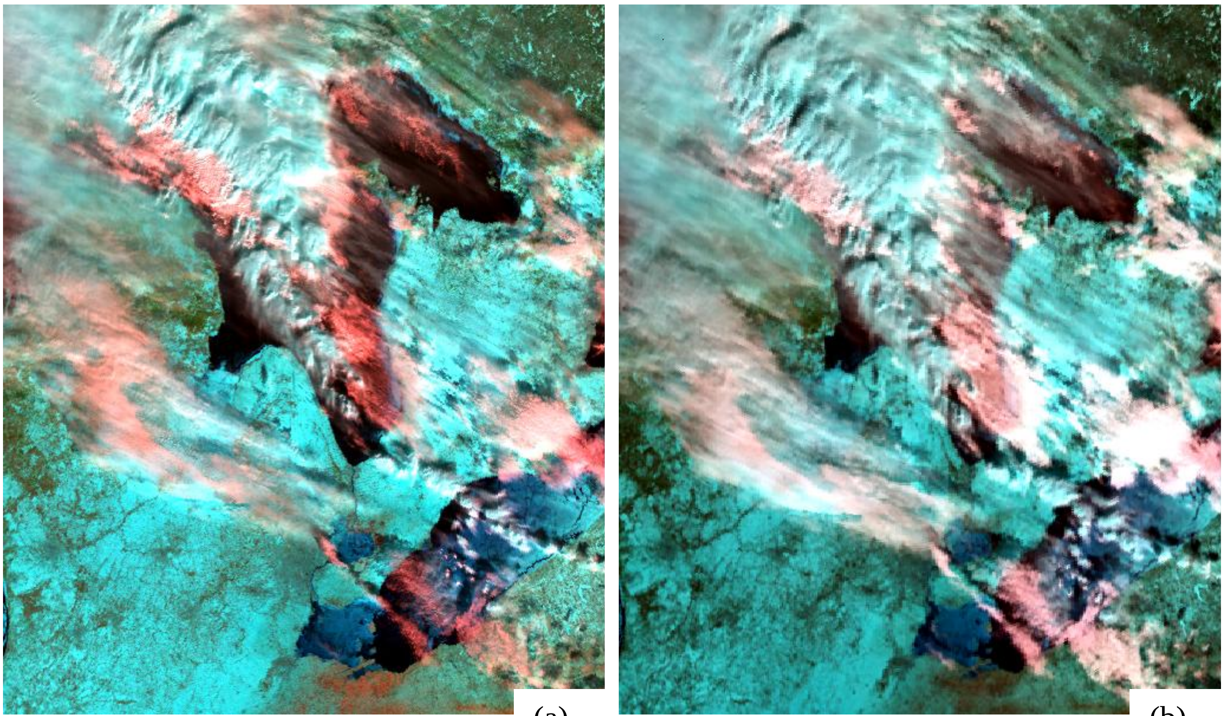


Figure 11. NOAA Great Lakes Ice Atlas ice chart for 21/01/97. Significant ice-cover is observed over Lake Erie and in southern tip of Lake Huron.

Reflectance channel imagery and the ARC-Lake ice mask are shown in Figure 12. Ice is clearly visible in the false colour images (Figure 12 a and b), as mid-blue regions (darker than the land) with adjacent black areas being open water. There is good correspondence between the ice cover visible in the ATSR-2 imagery (red areas in Figure 12c) and the black areas (100%) ice represented in the ice chart (Figure 11). This supports the use of the ice chart data for validation purposes (§4). The ice mask from ARC-Lake Figure 12c captures the visible ice cover well under clear-skies. This example also highlights one of the likely causes of the apparent underestimation of ice cover in the ARC-Lake product relative to the ice charts (§4). In the ice chart Figure 11 there are bands, roughly running north-south, of 90% and 70% ice cover towards the west of Lake Erie. While some of the corresponding area is covered in cloud, no ice is visible in the ATSR-2 imagery. There are two possibilities. It may be that not all forms of ice cover are visible in the ATSR reflectance channels and therefore may not trigger the ice detection test. Certainly, for water-logged or very thin ice, this is possible. The other possibility is that the ATSR-2 image more accurately delineates the ice than the chart, for this area. This possibility is supported by the brisk north-westerly flow (evident from the shapes of the clouds) which could tend to pile the ice of Lake Erie on the southern shore and open up a lead in the ice corresponding to the apparently open water area in the ATSR-2 image.

Another point worth noting is that the ice detection using the normalized index works well in cloud-shadow areas as well as directly illuminated areas (look at Lake Erie).



(a) (b)

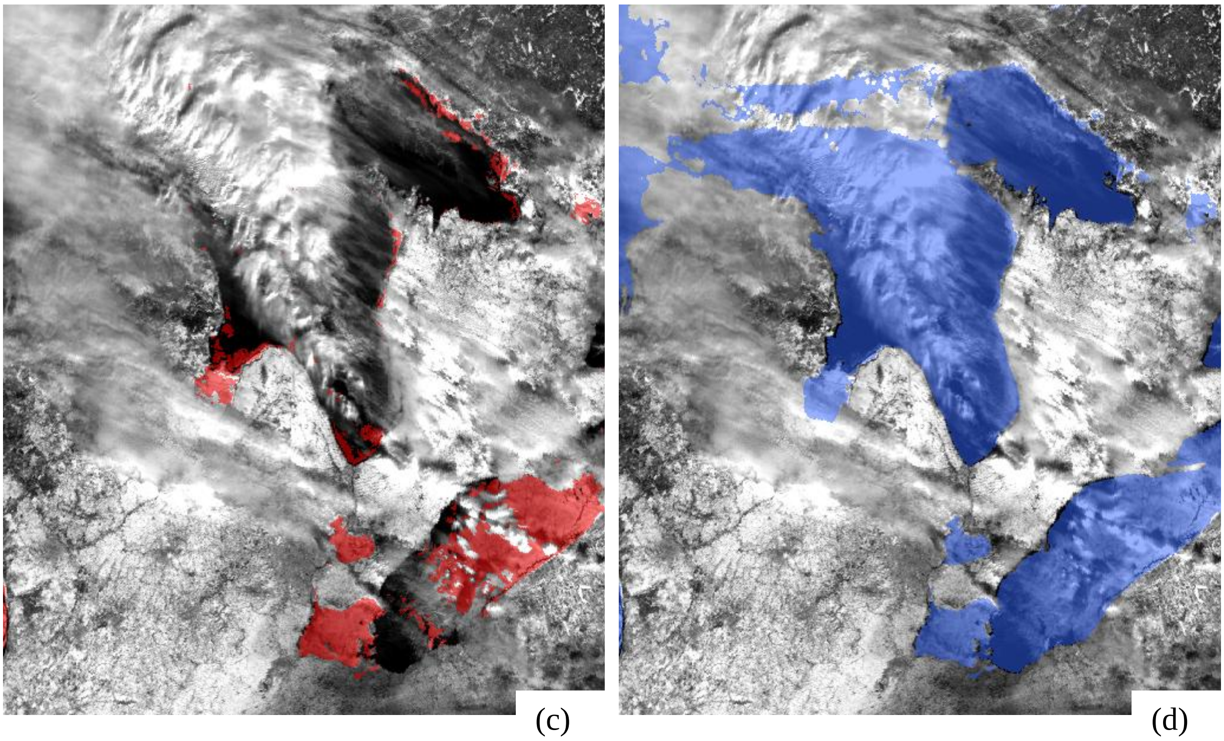


Figure 12. Analysis of ARC-Lake ice detection for Lakes Erie and Huron on for 21/01/97. (a) False colour image from ATSR-2 reflectance channels (0.66 µm, 0.87 µm and 1.6 µm) for nadir view. (b) As (a) but for forward view. (c) 0.66 µm reflectance in nadir view with ARC-Lake ice mask overlain in red. (d) 0.66 µm reflectance in nadir view with ARC-Lake land mask overlain (lake pixels are blue).

A second example of the analysis of the ARC-Lake ice cover product is presented in Figure 13 and Figure 14. The ice chart for 2nd February 1999 is shown in Figure 13. On this day, a large region in the north of Lake Erie was analysed to be > 95% ice covered. Low (~30%) ice concentrations were also analysed for the north east of Lake Ontario.

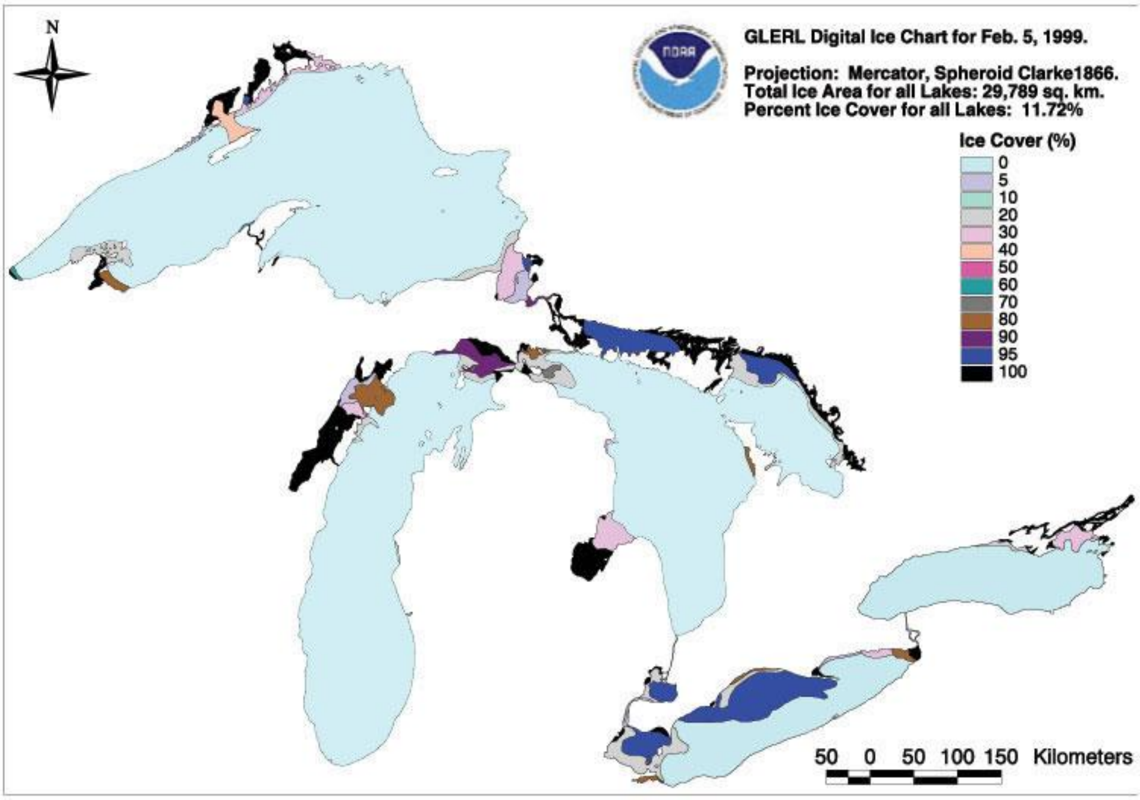


Figure 13. NOAA Great Lakes Ice Atlas ice chart for 05/02/99. Significant ice-cover is observed over the north of Lake Erie and partial ice cover is observed in the north east of Lake Ontario.

Reflectance channel imagery and the ARC-Lake ice mask are shown in Figure 14. As in the previous example, ice is clearly visible in the false colour images (Figure 14 a and b), as blue coloured regions. Again, there is good correspondence between the ice cover visible in the ATSR-2 imagery and that represented in the ice chart (Figure 13). In this case, the ice mask from ARC-Lake Figure 14c does not provide an accurate representation of the ice cover as in the previous example (Figure 12). The ice area on Lake Erie is only partly flagged, and this sort of occurrence will contribute to the underestimation of ice cover in the ARC-Lake product relative to the ice charts seen in §4. This is a result of the pre-ice detection step in the ARC-Lake processing (MacCallum and Merchant, 2010a), employed to limit misclassification of open-water as ice. Were this threshold test not implemented, approximately ¾ of the open water surface of Lake Erie would be misclassified as ice in this example. Further optimisation of this threshold test and the NDSI test are required.

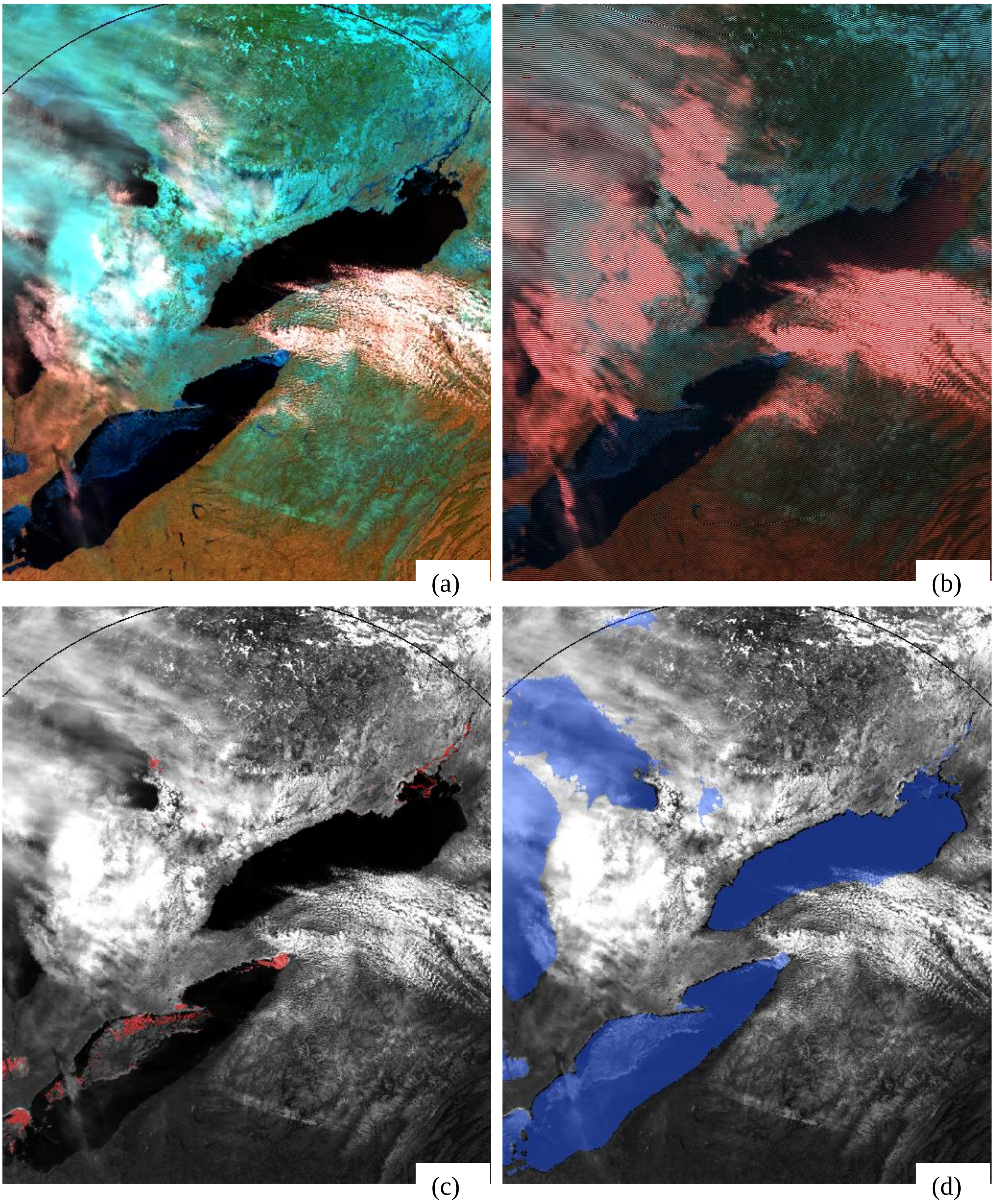


Figure 14. Analysis of ARC-Lake ice detection for Lakes Erie and Ontario on for 05/02/99. (a) False colour image from ATSR-2 reflectance channels (0.66 µm, 0.87 µm and 1.6 µm) for nadir view. (b) As (a) but for forward view. (c) 0.66 µm reflectance in nadir view with ARC-Lake ice mask overlain in red. (d) 0.66 µm reflectance in nadir view with ARC-Lake land mask overlain (lake pixels are blue).

6 CONCLUSIONS

The ARC-Lake LIC and Lake ST products have been assessed quantitatively and qualitatively through comparisons with *in situ* observations (and blended products from *in situ* and remote sensing). Results presented in §2 and §3 demonstrate the performance of the Lake ST product in terms of the accuracy of the temperature retrieval itself, and the performance of the cloud screening methods on which the accuracy of the Lake ST product is highly dependent. Through this analysis, the Bayesian cloud screening methods (Merchant *et al*, 2005) employed in ARC-Lake are demonstrated to offer a more accurate classification of the cloud cover than the operational cloud screening, with smaller rates of falsely flagged clear-sky pixels as cloud. This is illustrated in Figures 5 to 9 and quantitatively in terms of the number of clear-sky match-ups with *in situ* observations (§2).

Validation of the Lake ST product against *in situ* observations (§2) shows good consistency across different channel/view combinations within ARC-Lake, with satellite-*in situ* biases within ~0.2 K across all combinations (considering day and night separately). Uncertainties are also consistent across retrievals, with RSDs within 0.1 K. A summary of typical performance of the Lake ST retrievals is presented in Table 9.

Day / Night	ATSR-2		AATSR	
	Bias	RSD	Bias	RSD
Day	0.53	0.54	0.3	0.49
Night	-0.19	0.56	-0.36	0.56

Table 9. Average validation statistics for Lake ST retrievals from ARC-Lake. Bias is calculated as satellite-*in situ*.

The ARC-Lake LIC product has been assessed in §4 and §5. Through comparison with ice charts (a blend of *in situ* and remote sensing observations), the LIC product is shown to provide a reasonable representation of ice cover, where cloud cover permits. There is agreement about the broad ice concentration class between the LIC product and the ice charts in over 75% of cells assessed. There is a tendency for the LIC product to underestimate the ice cover. In some cases this under-masking arises as a result of earlier coarse cloud screening, and this will be remedied in future developments of the ARC-Lake processing scheme. In other cases, there is probably a low sensitivity of the test to thin or water-logged ice, although there is also the possibility that leads are not correctly analysed in the validation data. Further development of the ice detection test will be required, and in particular it would be desirable to integrate it with the cloud detection in a single classification step.

In general, validation results from both the Lake ST and LIC ARC-Lake v1.0 products are positive, but there are identified improvements to be made during the next phase of the project, as follows:

- Version 1.1 (due by end 2010) will see

  - implementation of salinity dependent emissivity (currently, sea water emissivity is used for all lakes)
  - a further iteration of the Lake ST prior (MacCallum & Merchant, 2010a, §5 and §4 respectively

- Version 2.0 (due June 2011) will extend the Lake ST timeseries to 1991 (although LIC will not be possible with the ATSR-1 channels).
- Version 2.1 (due September 2011) will apply adjustments to bring ATSR-1, ATSR-2 and AATSR into mutual agreement using overlaps between the sensors, creating a homogenized data set

Secondary objectives for the next phase include

- possible retuning of the NSDI-based ice detection test
- possible consolidation of cloud detection and ice detection in a single Bayesian classifier
- extending the amount of validation data collected, particularly for lakes in different climate regimes than the Great Lakes

The majority of the validation work reported here has been carried out over the Great Lakes. This has arisen through availability of suitable validation data rather than by choice. As the Great Lakes are some of the largest lakes in the world and are within a relatively small geographical region, they do not represent the full diversity of lakes considered in Phase One of the ARC-Lake project. Consequently, efforts are ongoing to identify and obtain validation data for other lakes around the globe.

7 References

Assel, R. A. (1983). NOAA Data Report ERL GLERL-24: A Computerized Ice Concentration Data Base For The Great Lakes, NOAA, Great Lakes Environmental Research Laboratory.

Assel, R. A., Norton, D. C., & Cronk, K. C. (2002). NOAA Technical Memorandum GLERL-121: A Great Lakes Ice Cover Digital Data Set For Winters 1973-2000, Great Lakes Environmental Research Laboratory.

Assel, R. A. (2003). NOAA Technical Memorandum GLERL-125: Great Lakes Ice Cover, First Ice, Last Ice, and Ice Duration: Winters 1973-2002, NOAA, Great Lakes Environmental Research Laboratory.

Hall, D., Riggs, G. A., & Salomonson, V. V. (1995). Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment*, *54*(2), 127-140. doi: 10.1016/0034-4257(95)00137-P.

MacCallum, S. N., and Merchant, C. J. (2010a), ATSR Reprocessing for Climate Lake Surface Temperature: ARC-Lake - Algorithm Theoretical Basis Document – v1.0, School of GeoSciences, The University of Edinburgh.

MacCallum, S. N., and Merchant, C. J. (2010b), ARC-Lake Data Product Description – v1.0, School of GeoSciences, The University of Edinburgh.

Merchant, C. J., Harris, a. R., Maturi, E., & Maccallum, S. (2005). Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval. *Quarterly Journal of the Royal Meteorological Society*, *131*(611), 2735-2755. doi: 10.1256/qj.05.15.

8 Appendix

In this appendix, the results of the comparison of the ARC-Lake LIC product with ice chart data are presented for AATSR and ATSR-2.

8.1 LIC for AATSR

ARC-Lake / Ice Charts	0 %	1-15 %	15-85 %	>85 %
0 %	64.67	0.87	0.85	0.29
1-15 %	10.15	0.60	0.71	0.64
15-85 %	2.47	0.80	2.12	2.65
>85 %	1.65	1.38	3.72	6.42

Table 10. Results of comparison of ARC-Lake LIC product from AATSR with ice charts for Lake Superior. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 82 days of observations and 56827 grid cells.

ARC-Lake / Ice Charts	0 %	1-15 %	15-85 %	>85 %
0 %	55.87	0.89	0.64	0.14
1-15 %	10.14	0.82	0.76	1.04
15-85 %	2.62	1.22	2.83	2.01
>85 %	2.60	1.09	4.28	13.05

Table 11. Results of comparison of ARC-Lake LIC product from AATSR with ice charts for Lake Huron. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 77 days of observations and 42484 grid cells.

ARC-Lake / Ice Charts	0 %	1-15 %	15-85 %	>85 %
0 %	73.36	0.68	0.52	0.13
1-15 %	5.85	0.35	0.40	0.43
15-85 %	2.09	0.84	2.09	2.09
>85 %	2.20	0.82	2.48	5.65

Table 12. Results of comparison of ARC-Lake LIC product from AATSR with ice charts for Lake Michigan. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 55 days of observations and 28966 grid cells.

ARC-Lake / Ice Charts	0 %	1-15 %	15-85 %	>85 %
0 %	55.70	0.53	0.39	0.02
1-15 %	7.20	0.46	0.49	0.41
15-85 %	1.73	0.75	1.71	5.00
>85 %	7.74	0.74	2.40	14.75

Table 13. Results of comparison of ARC-Lake LIC product from AATSR with ice charts for Lake Erie. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 56 days of observations and 14218 grid cells.

Ice Charts \ ARC-Lake	0 %	1-15 %	15-85 %	>85 %
0 %	75.17	0.87	0.58	0.18
1-15 %	5.48	0.75	0.76	0.19
15-85 %	3.42	0.67	2.43	3.30
>85 %	0.59	0.08	0.52	4.99

Table 14. Results of comparison of ARC-Lake LIC product from AATSR with ice charts for Lake Erie. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 53 days of observations and 9992 grid cells.

8.2 LIC for ATSR-2

Ice Charts \ ARC-Lake	0 %	1-15 %	15-85 %	>85 %
0 %	65.70	1.01	0.90	0.13
1-15 %	9.59	0.32	0.30	0.06
15-85 %	2.23	1.03	2.24	1.19
>85 %	3.96	1.02	3.08	7.24

Table 15. Results of comparison of ARC-Lake LIC product from ATSR-2 with ice charts for Lake Superior. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 119 days of observations and 62379 grid cells.

Ice Charts \ ARC-Lake	0 %	1-15 %	15-85 %	>85 %
0 %	55.66	1.52	1.79	0.62
1-15 %	6.36	0.36	0.64	0.55
15-85 %	3.90	1.60	3.33	2.28
>85 %	3.10	0.99	4.54	12.77

Table 16. Results of comparison of ARC-Lake LIC product from ATSR-2 with ice charts for Lake Huron. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 106 days of observations and 53649 grid cells.

Ice Charts \ ARC-Lake	0 %	1-15 %	15-85 %	>85 %
0 %	80.17	0.71	0.65	0.17
1-15 %	1.46	0.20	0.28	0.04
15-85 %	2.62	0.70	1.49	0.89

>85 %	3.26	0.70	1.77	4.90

Table 17. Results of comparison of ARC-Lake LIC product from ATSR-2 with ice charts for Lake Michigan. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 84 days of observations and 34836 grid cells.

Ice Charts \ **ARC-Lake**	**0 %**	**1-15 %**	**15-85 %**	**>85 %**
0 %	58.88	0.67	0.70	0.19
1-15 %	4.25	0.05	0.09	0.36
15-85 %	5.03	0.90	1.42	3.14
>85 %	5.38	1.49	4.35	13.09

Table 18. Results of comparison of ARC-Lake LIC product from ATSR-2 with ice charts for Lake Erie. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 76 days of observations and 24616 grid cells.

Ice Charts \ **ARC-Lake**	**0 %**	**1-15 %**	**15-85 %**	**>85 %**
0 %	82.44	1.42	1.96	2.05
1-15 %	0.71	0.05	0.19	0.08
15-85 %	2.55	0.40	0.89	1.13
>85 %	0.67	0.18	0.92	4.36

Table 19. Results of comparison of ARC-Lake LIC product from ATSR-2 with ice charts for Lake Ontario. Values are the percentage of cells matching each surface classification pair between ARC-Lake LIC and the ice charts. These results represent 66 days of observations and 12879 grid cells.