MAPPING OF LAKE ICE IN NORTHERN EUROPE USING DUAL-POLARIZATION RADARSAT-2 DATA

Hindberg, Heidi and Malnes, Eirik

Northern Research Institute (Norut), PO Box 6434 Tromsø Science Park, N9291 Tromsø, Email: heidi.hindberg@norut.no, eirik.malnes@norut.no

ABSTRACT

In this paper, we investigate the potential of including cross-polarization data in an unsupervised classification method based on SAR data to determine ice extent over lakes in Northern Europe. By introducing cross-pol data we can increase the separability between open water and ice, and we can decrease misclassifications where open water with waves is classified as ice. Cross-pol data also helps with labelling of the classes. However, cross-pol data can decrease the separability between the classes if the ice on the lake is very thin.

1. INTRODUCTION

Measurements of ice extent on inland lakes are useful for the modelling communities in numerical weather prediction, regional climate models and hydrology. These communities have indicated a general lack of observations of lake ice, and classifications of lake surfaces into open water and ice-covered water based on earth observation are highly desirable [1]. Recent investigations into the importance of lake ice cover for these fields include [2, 3]. Previously, the use of earth observation data in lake ice classification have been limited to visual/manual inspections of optical satellite data. A review of methods for monitoring lake (and river) ice can be found in [4]. Synthetic aperture radar (SAR) satellite data offers a useful alternative to optical satellite data for lakes at high latitudes, in particular during the polar night period and during cloud cover when optical sensors are useless. Thus, we can potentially get observations more often using SAR. Recently there has been an increasing interest in using SAR data and automatic processing chains for determining lake ice cover (see e.g., [5,6]).

In a previous work [7], we have developed processing chains for lake ice retrieval based on co-polarized (co-pol) Envisat ASAR wide-swath data. The aim of that work was to create a fully automatic processing chain. However, we observed that the SAR data had a variable separability between open water and ice covered water. Generally, the backscatter from water in co-pol SAR images is very low, and the backscatter from ice on the lake surface has higher values. For co-pol data, the presence of waves on the water can increase the backscatter from water areas, even to a point where the backscatter from water is higher than the backscatter from ice. This will drastically decrease the classification accuracy of the method, and it will make it difficult to find a generic rule for assigning labels to the classes found in the image. This was previously solved by using auxiliary data. Furthermore, in [7] we had to perform a manual selection to remove scenes where the water surface had wind effects. If we had measurements of wind strengths in the lake area we could include this in an automated scene selection, however this requires a weather station close to the lake in question and the wind to be constant over the lake.

Upcoming SAR sensors such as Sentinel-1 will make more dual-polarization data available. Generally, the backscatter from water in cross-polarized (cross-pol) SAR data is very low, and the backscatter from ice on the lake surface has higher values. It is well known that cross-pol SAR data is less sensitive to waves on the water. Thus, in cases where the co-pol data have areas on the lake surface with a sea state such that the backscatter values are overlapping with the backscatter from ice, we can utilize cross-pol data to increase the separability between the two classes. Since the backscatter from water will not be higher than the backscatter from ice in the cross-pol data, we can base the labelling of the classes on which class that has the lowest cross-pol values. We also note that cross-pol backscatter from ice and open water is less dependent on the incidence angle that co-pol backscatter [8]. In this study, we use Radarsat-2 dual-polarization data over lakes in Norway and Sweden to investigate how dual-polarization data can improve the lake ice classification. Radarsat-2 SCNA data also has better resolution than Envisat ASAR wide swath data, allowing us to apply the method to smaller lakes (and potentially even large rivers). Dual-pol Radarsat-2 SCNA data is comparable to Sentinel-1 EWS data, and the results of this study hence provide insight into operational use of Sentinel-1 for lake ice classification.

2. STUDY SITES AND DATA

Our source of satellite data is Radarsat-2 ScanSAR narrow (SCNA) dual-polarization data. We have considered two Norwegian lakes, Mjøsa at 60°40′N 11°00′E and Femunden at 62°12′N 11°52′E, and one Swedish lake, Torneträsk at 68°22′N 019°06′E. Note that Torneträsk is north of the polar circle, thus optical satellites will be useless for this lake during the polar night. We have selected some scenes that illustrate the

Proc. 'ESA Living Planet Symposium 2013', Edinburgh, UK 9–13 September 2013 (ESA SP-722, December 2013)

potential and the potential pitfalls of including cross-pol data in the classification. We consider two scenes from each of the lakes, where five are from the break-up period of the lakes and one is from the freeze up period, as seen in Tab. 1.

Table 1 Date of acquisitions

Lake	Scene 1	Scene 2
Torneträsk	20121205	20130529
Mjøsa	20130424	20130502
Femunden	20130516	20130518

3. METHOD

3.1. Pre-processing

We create backscatter images from the SAR data using precision geocoding software [9], projecting the images in UTM zone 33 north coordinates (WGS-84 ellipsoid) with a spatial resolution of 50 meters. A Lee filter of size 3x3 is applied to each image to reduce speckle noise while preserving edges in the images. Since we are only interested in the lake surface, we then apply a land mask to remove land pixels. We also mask out the five pixels closest to the shoreline, to avoid using pixels that contain both water and land in the classification scheme. Since we use a statistical method requiring several input pixels, we do not consider images that cover less than 30% of the lake's surface. We also remove images that, based on the date of the acquisition, should contain only water or only ice.

3.2. Classification

Since we have dual-polarization data, we have two data values for each pixel in an image. We have three possible input data sets to a classification method: (i) use only co-pol data, (ii) use only cross-pol data, and (iii) use both polarizations simultaneously. Thus, the feature vectors are 1D for (i) and (ii) and 2D for (iii). We model the feature vectors using a Gaussian Mixture Model (GMM), i.e., we assume that the probability density of the feature vectors is a sum of normal distributions. Each normal distribution represents one class in the image. The mean, covariance and prior probability of each normal distribution is estimated from a subset of the data using an Expectation Maximization (EM) algorithm. We use a Bayes classifier, which assigns a class to an input feature vector by selecting the class whose probability density in the GMM gives the highest value give the input feature vector. The K-means classifier was successfully used in [5] to classify lake ice, we note that K-means can be considered as a special case of the Bayes classifier assuming a GMM.

The classification method requires us to select the number of classes we want to segment the image into. We should have two classes in the final result, but for some scenes we need to segment the image into three (or potentially more) classes and merge some of the classes to obtain the final result. The number of classes used for each scene is selected by visual inspection of the backscatter images and the marginal and 2D histograms. If we use more than two classes, we need to merge the classes so we are left with two classes. For case (iii), we can have situations where we need one class for calm open water areas that are dark in both polarizations, one class for open water areas affected by wind which are dark in cross-pol and brighter in co-pol, and one class for ice covered areas that are bright in both polarizations. Based on this consideration, we choose to merge the two classes with lowest cross-pol values and label the resulting class as water. If we use more than two classes in case (i) or (ii) it is more difficult to propose a general merging rule.

Using the Bayes classifier we obtain a segmented image where all pixels on the lake surface are assigned a class number. To get a final product we need to label the classes, i.e., specify which classes represent water and which classes represent ice. Since the strength of the backscatter from water relative to the backscatter from ice varies with the presence of wind in co-pol data, we need information about wind conditions to label the classes correctly. For cross-pol data, the backscatter from water should not be higher than the backscatter from ice regardless of wind conditions, thus the class with the lowest mean value should be labelled water and the others as ice. Likewise, we set the class with the lowest cross-pol mean in case (iii) as water.

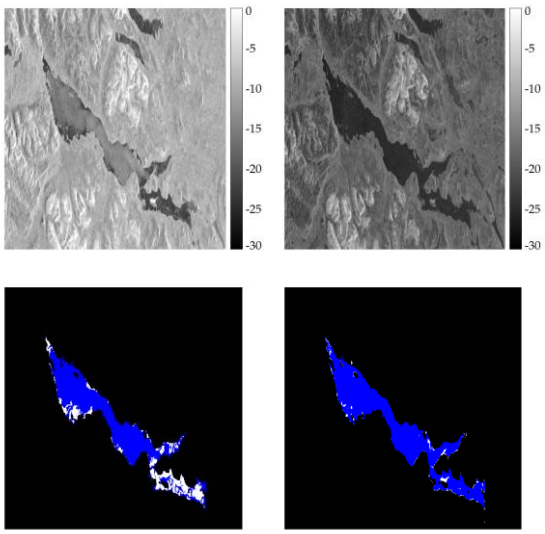

Figure 1. Upper: Backscatter images for Torneträsk on 20121205, co-pol (left) and cross-pol (right). Lower: Classification results from co-pol only (left) and from combined dual-polarization data (right).

4. RESULTS

We classified the six scenes with the three different input data sets. For some scenes, there was little difference between the results from cross-pol and the results from both polarizations simultaneously. In these cases, we only show the combined classification.

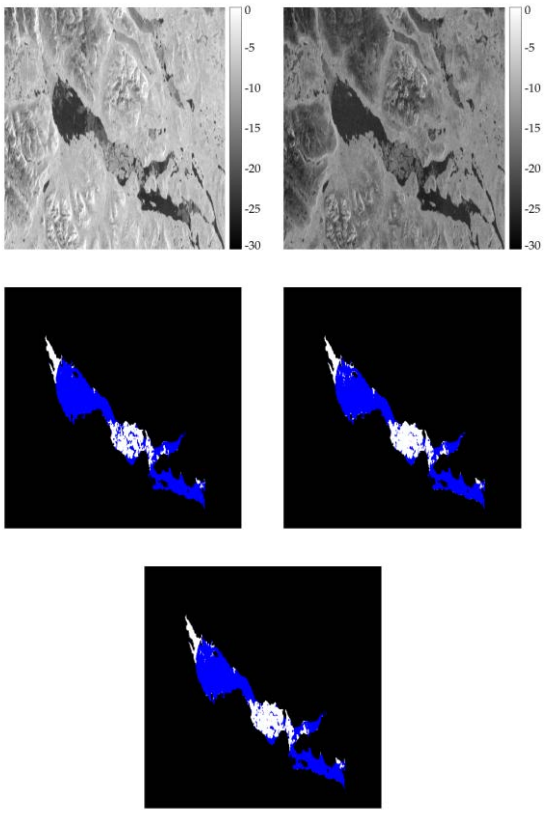

Figure 1. Upper: Backscatter images for Torneträsk on 20130529, co-pol (left) and cross-pol (right). Middle: Classification results from co-pol only (left) and from cross-pol data only (right). Lower: Classification based on both co- and cross-pol data.

4.1. Torneträsk

Fig. 1 shows the co-pol and cross-pol backscatter images for Torneträsk on 20121205, along with the classification results assuming two classes based on only the co-pol data and based on dual-pol data. The result based only on cross-pol data is similar to the dual-pol result, and is thus not shown here. There appears to be two classes present in the co-pol backscatter image, it looks like the darker areas on the lake surface are ice and the light areas are water with waves. However, these potentially ice-covered areas are not apparent in the cross-pol image. We know that the backscatter in cross-pol from water is less sensitive to waves, such that the water areas in cross-pol will all be dark regardless of the wind. The scene is from early in the winter season, such that the ice on the lake has most likely just started forming. In this phase, the ice will necessarily be quite thin and may have a smooth surface, which will give a very low backscatter from the ice in both co-pol and cross-pol. Thus, we cannot separate the ice and water in cross-pol, and the only reason we can separate them in co-pol is the waves on the water. To avoid errors in the classification, we would need the entire water surface of the lake to have waves; areas without wind would be erroneously classified as ice. We see that the classification using co-pol data segments the image into two classes, and the result seems plausible. With the classification using the dual-pol data we have segmented the image into one class that is very bright in both backscatter images and one class for the rest of the pixels. These bright areas are probably land areas that have been included because of imperfections in the lake mask. In this example, we degrade the classification result by including the cross-pol data. Since the backscatter signatures from thin smooth ice and calm water (or water in all states in cross-pol) are hard to separate, we would not recommend using cross-pol data for the freeze period. Also, we need scenes with waves on the water to be able to use the co-pol data with any confidence. Note that we selected the labels for the classes manually for this scene, as the higher backscatter in co-pol represent the water in this case.

Figure 3. Segment of Lance data from 20130528 over Torneträsk

Likewise, we show the result from 20130529 in Fig. 2. The backscatter images show the ice-free water as dark areas and the ice-covered areas as bright in both co-pol and cross-pol. We also note some areas that are bright in co-pol and dark in cross-pol, which are most likely areas where wind is causing waves. Thus, we assume that the 2D input vectors consist of three classes: calm open water, open water with waves and ice-covered water. By combining the two classes with the lowest cross-pol mean values, we get a final product indicating ice covered and open water. If we use the same procedure on only the co-pol data, we also get a visually acceptable classification result. However, we needed the cross-pol data to indicate that the class with the middle mean value actually should represent water instead of ice. If we base the classification on only the cross-pol data, we can segment the image directly into two classes, and the results are not that different from the other approaches. To support our conclusion that the middle class (high co-pol backscatter, low cross-pol backscatter) is not ice, we compare it to a segment from

the MODIS Near Real-Time Images from the previous day in Fig. 3 (http://rapidfire.sci.gsfc.nasa.gov/realtime/). By visual inspection, we see that our classification results appear to match this optical image.

.

4.2. Mjøsa

The scene over Mjøsa on the 20130424 in Fig. 4 shows how the open water can be completely inseparable from the ice-covered areas in co-pol, while cross-pol gives a clear distinction between the two classes. Using only co-pol, we conclude that there is only ice in the image. The product based on cross-pol show some of the open water, but there are some "noisy" classifications of ice in the water areas. By using both polarizations at the same time, we get a cleaner and more plausible classification result.

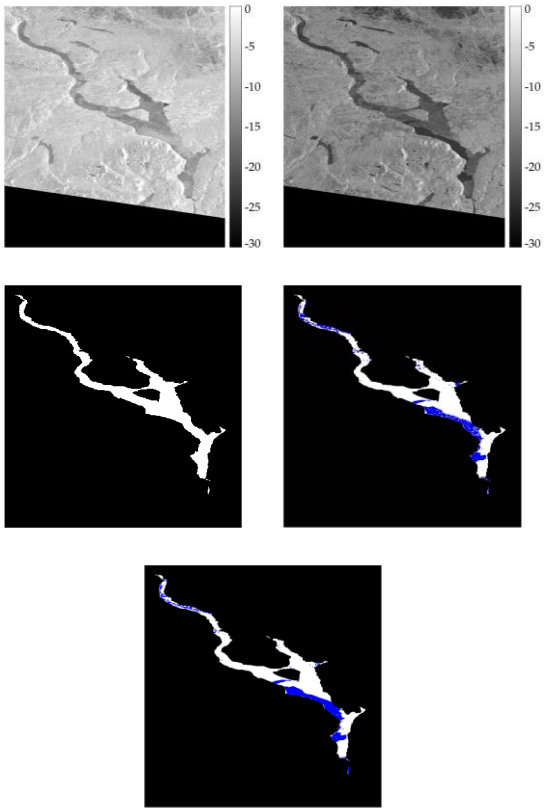

Figure 4. Upper: Backscatter images for Mjøsa on 20130424, co-pol (left) and cross-pol (right). Middle: Classification results from co-pol only (left) and from cross-pol data only (right). Lower: Classification based on both co- and cross-pol data.

The result for 20130502 are shown in Fig. 5. Here, the co-pol backscatter is somewhat confusing, with some small dark areas, brighter areas to the north and bright smooth areas at the centre of the lake. Based on the cross-pol image it looks like there is ice at the north end of the lake and open water over the rest of the lake. However, it is not clear if the darker areas in co-pol actually are calm open water or water covered by a thin ice layer that is undetectable in cross-pol. We segment the 1D input vectors into two classes, and the 2D input vector into three classes. To produce the final ice cover product, we combine the two classes with the lowest mean values in cross-pol. We note that there are some differences in the classification results based on the three different approaches. Here, either the result based on cross-pol only or the result based on both input data sets seem the most plausible.

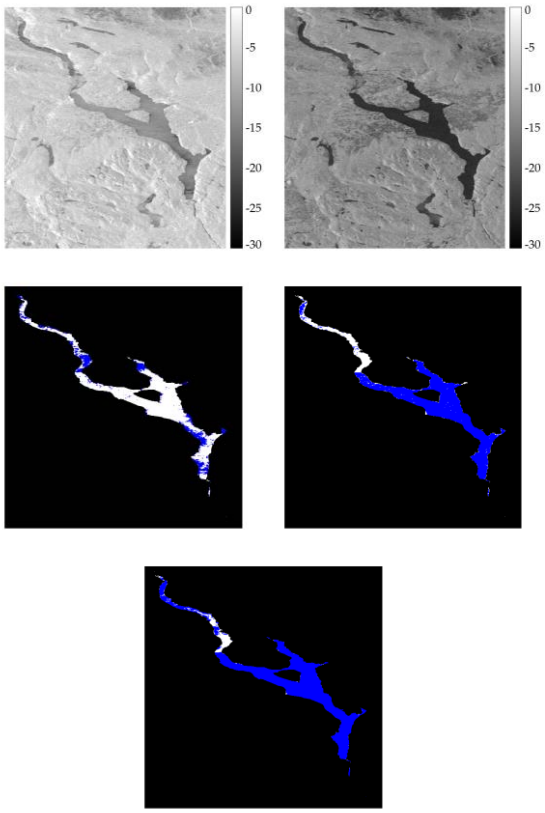

Figure 5. Upper: Backscatter images for Mjøsa on 20130502, co-pol (left) and cross-pol (right). Middle: Classification results from co-pol only (left) and from cross-pol data only (right). Lower: Classification based on both co- and cross-pol data.

4.3. Femunden

In Fig. 6, we again see a situation where the backscatter from water and ice has overlapping values in co-pol. In the cross-pol image, the values are quite different from the two classes. Here, we use two classes for all three different input vectors. The results based on cross-pol only and the results from the 2D data set were almost identical, so we do not show the cross-pol result here. These results are far better than the result based on co-pol only, which have many "noisy" classifications.

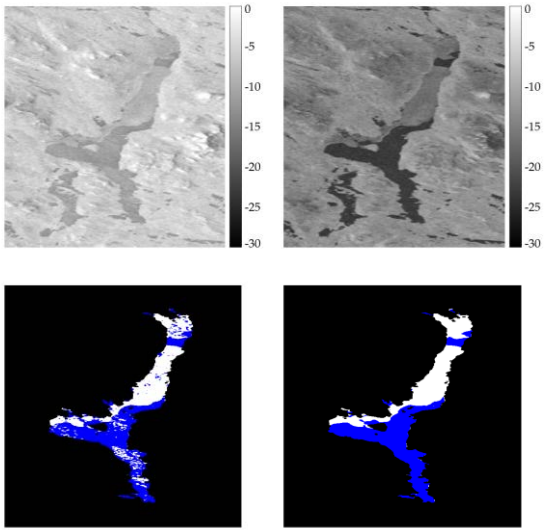

Figure 6. Upper: Backscatter images for Femunden on 20130516, co-pol (left) and cross-pol (right). Lower: Classification results from co-pol only (left) and from combined dual-polarization data (right).

Finally, Fig. 7 from 20130518 show a situation with separability in the cross-pol image and in the co-pol image, but with a small bright area in co-pol probably caused by wind. In Fig. 2 the backscatter from the wind areas were actually distinguishable from the backscatter from ice, and it was natural to classify the co-pol image into three classes. Here, the backscatter values are similar and the 1D co-pol data suggests only using two classes.

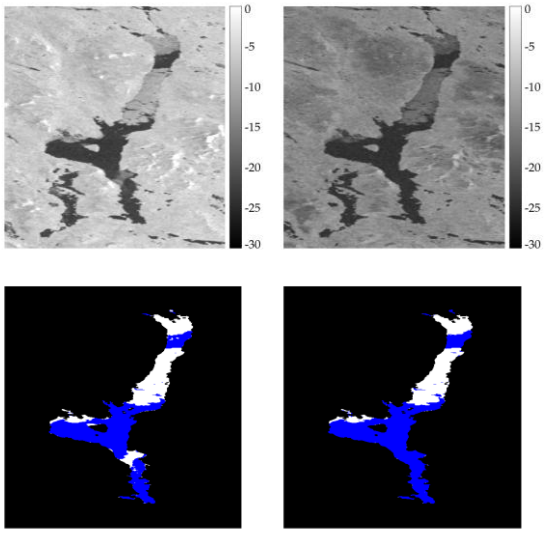

Figure 7. Upper: Backscatter images for Femunden on 20130518, co-pol (left) and cross-pol (right). Lower: Classification results from co-pol only (left) and from combined dual-polarization data (right).

Likewise, we use two classes for the 1D cross-pol data, but the full 2D data is segmented into three classes. The results from the combined data set were comparable to the result from only cross-pol, thus we omit the cross-pol results from the figure. We see that the results from co-pol only classifies the wind area as ice, while the combined result classifies these areas as water.

5. DISCUSSION

We have investigated the potential advantage of combining co-pol and cross-pol data in unsupervised classification method for determining ice cover on lake surfaces in Northern Europe. While we have illustrated that cross-pol data can increase the separability of ice and open water on a lake in thaw, including the cross-pol for the freeze up can decrease the classification accuracy. Including cross-pol data has been successfully done for sea ice classification (e.g., [10]). However, it is far less likely to encounter sea ice that is as thin and smooth as to be indistinguishable from water in cross-pol, which is the main problem we had with the cross-pol data. We also observed that by only using the cross-pol data we risk a result with more spurious wrong classifications than combining the co-pol and cross-pol data. Generally, combining the two images in a 2D input vector seems to give the more plausible results. Note that even if the classification method is unsupervised, a manual step is required to remove scenes with only water or only ice, and scenes that simply do not have enough separability between ice-covered and open water.

Sentinel-1A will be launched in 2014. The sensor will be operated mainly in the interferometric wide swath mode (IW) over Europe, but also in the Extended Wide swath mode (EW) over parts of Sweden, Finland and Russia where most of the largest lakes in Europe are localised [11]. For the EW-mode with dual-polarimetric data we expect that the results reported above will be directly parallel. Similar resolutions should, but more frequent acquisitions should yield a good potential for accurate classification of lake ice. The IW mode will only be delivered with single polarimetric data (HH or VV) due to limitations in duty-cycle and downlink capacity. The IW mode will have better spatial resolution at the cost of smaller swath width and lack of cross-pol channel. The higher spatial resolution, 25m for IW mode vs. 50m for EW mode, could in some cases (e.g. for river ice classification) be an advantage, but in most cases for large scale automatic lake ice classification, we expect that it will yield a poorer lake ice extent product.

6. REFERENCES

1. Kourzeneva, E., Samuelsson P., Ganbat G., & Mironov D. (2008). Implementation of lake model FLake into HIRLAM. HIRLAM Newsletter, no. 54, 2008.

2. Eerola, K., Rontu L., Kourzeneva E., & Shcherbak E. (2010). A study on effects of lake temperature and ice cover in HIRLAM, *Boreal Env. Res.*, **15**, 130-142.

3. Martynov, A., Sushama L., & Laprise R. (2010). Simulation of temperate freezing lakes by one-dimensional lake models: performance assessment for interactive coupling with regional climate models, *Boreal Env. Res.*, **15**, 143-164.

4. Jeffries, M. O., Morris K., & Kozlenko N. (2005). Ice characteristics and processes, and remote sensing of frozen river and lakes, *Remote Sensing in Northern Hydrology*, C. R. Duguay and A. Pietroniero (eds), AGU Monograph 163, 63-90, 2005.

5. Sobiech, J., & Dierking, W. (2013). Observing lake- and river-ice decay with SAR: advantages and limitations of the unsupervised k-means classification approach. *Annals Glac.* **54**(62), 65–72.

6. Geldsetzer, T. van der Sanden, J., & Brisco, B. (2010). Monitoring lake ice during spring melt using RADARSAT-2 SAR, *Can. J. Remote Sensing,* **36**, S391-S400.

7. Hindberg H., Malnes E., Asalmi, H., & Mattila, O.-P. (2012). Validation of ENVISAT ASAR based lake ice maps on Lake Päijänne. In Proc IEEE IGARSS, 2012, July 22-27, Munich, Germany, 5238-5241.

8. Duguay, C. R., Pultz, T.J., Lafleur, P.M., & Drai, D. (2002). RADARSAT backscatter characteristics of ice growing on shallow sub-Artic lakes, Churchill, Manitoba, Canada. *Hydrological Processes,* **16**(8), 1631-1644.

9. Larsen, Y., Engen G., Lauknes T. R., Malnes E., & Høgda, K. A.(2005). A generic differential InSAR processing system, with applications to land subsidence and SWE retrieval. FRINGE, ESA ESRIN, Frascati, Italy, November 28 - December 2.

10. Dierking, W., & Pedersen, L.T. (2012). Monitoring sea ice using ENVISAT ASAR - A new era starting 10 years ago. In Proc IEEE IGARSS, 2012, July 22-27, Munich, Germany, 1852 - 1855.

11. ESA (2013), Sentinel High Level operations plan, ESA/PB-EO(2013)1, Paris, France.