# Interannual Variability of Lake Ice Backscatter Anomalies on Lake Neyto, Yamal, Russia

Georg Pointner[1,2] and Annett Bartsch[1,2]

[1]b.geos, Korneuburg, Austria
[2]Austrian Polar Research Institute, Vienna, Austria

## Abstract

Anomalous areas of varying shape and location characterized by low backscatter in Synthetic Aperture Radar (SAR) imagery of lake ice on lake Neyto on the Yamal Peninsula in Russia have been described qualitatively in the literature for many years. Possible suggested causes are the formation of eddies or the release of gas through the lake sediments, which could both lead to local thinning of the ice layer and alter radar backscatter. To date, the phenomenon, its cause, and its spatial and temporal dynamics are poorly understood, and studies from other geographic regions are completely absent. In order to perform first steps towards a better understanding of the phenomenon, we developed a workflow to quantitatively assess the spatial variability of the anomalies in the years 2015 to 2019 for lake Neyto. We introduce a binary image classification algorithm developed with state-of-the-art open-source image processing tools and employ metrics commonly used for describing spatial relationships of vector and raster data. This includes polygon distances, polygon intersections and cumulative pixel counts deduced from the classification results in order to quantify, for the very first time, the dynamics over a number of years. The geospatial analysis reveals large spatial variations, but also some overlap between different years. Locations of anomalies do not seem more similar between consecutive years than when they are compared over the longer period. Some of the spatial properties of the clusters of low backscatter may support the explanation of gas release as the primary cause of the observed patterns.

## Keywords:

remote sensing, Arctic, lakes, image processing, synthetic aperture radar

## 1    Introduction

Arctic lakes are important features of the hydrosphere and the cryosphere. They occupy significant parts of the Arctic tundra and play an important role in the carbon cycle (e.g. Walter Anthony et al., 2012; Wik et al., 2016).

In winter, space-borne C-band Synthetic Aperture Radar (SAR) data can be useful for monitoring lake ice phenology (e.g. Surdu et al., 2015; Duguay & Pietroniro, 2005), and especially the grounding state of lake ice (e.g. Duguay & Lafleur, 2003; Surdu et al., 2014;

Grunblatt & Atwood, 2014). Regions of floating lake ice appear bright (high backscatter) in SAR images due to the high reflection of the radar signal, which is caused by high dielectric contrast at the ice–water boundary (Duguay et al., 2002).

However, for a range of lakes on the Yamal Peninsula in northwestern Siberia, patterns of low backscatter in central parts of lakes where floating lake ice is assumed were identified by Bogoyavlensky et al. (2018). Low backscatter is usually observed from shallow shelves of lakes, where the ice is grounded (Duguay et al., 2002). The extent of these areas remains fairly constant throughout the winter, but the extent of the area with low backscatter patterns changes significantly throughout the winter. Most of the zones of low backscatter outside the shelves can be identified first in the SAR imagery in mid to late winter (usually March or April), when they start to appear mostly as circular or elongated objects. These subsequently widen until the onset of snowmelt. Possible explanations for these backscatter anomalies include the formation of eddies and the accumulation of methane released through pockmarks in the lake sediments migrating upwards in the water column under the ice layer. Both would lead to a local thinning of the ice layer and thus to lower backscatter due to increased specular reflection from the water surface.

Understanding the origin and dynamics of these patterns may be important for climate research and for understanding sub-lake permafrost dynamics in the case of methane emissions, or for hydrological research in the case of eddies.

To date, the literature has included only visual descriptions of anomalies, in only a few SAR images of lake Neyto and lakes in its vicinity. To our knowledge, there are no descriptions of similar backscatter anomalies for lakes in other geographic regions. Engram et al. (2013) demonstrated a positive statistical relationship between L-band backscatter and bubbles of methane trapped in lake ice for a range of lakes in Alaska, but they did not show consistent areas of anomalous backscatter and noted that such a relationship could not be deduced for C-band data.

In this study, we perform the first-ever quantitative analysis on these objects of varying location and shape on lake Neyto, which will contribute to understanding the phenomenon. Understanding the nature of the phenomenon may lead to new applications of Sentinel-1 data for the monitoring of gas emissions or eddies in remote Arctic locations.

An important part of the research to understand the phenomenon is the analysis of changes in the locations of anomalies in different years. This study aims to describe the variability of patterns from 2015 to 2019 based on Sentinel-1 Extra-Wide Swath (EW) data. First, a method needed to be identified which would allow the retrieval of the anomalies. In the second step, temporal patterns of object metrics were analysed. Objects of low backscatter intensity were extracted, and distance and intersection metrics were used to describe the spatial variability between the years.

# 2  Data and Methods

## 2.1  Study area selection

The study site is one of the largest lakes on the Yamal Peninsula, lake Neyto, which is among the lakes with the largest clusters of pixels of low backscatter in central parts of the lake in late winter. The visual appearance of patterns in different years have already been described by Bogoyavlensky et al. (2018), but no quantitative analyses to characterize spatial and temporal properties had so far been carried out. Similarities to optical data were also identified (an example is shown in Figure 1). Due to these characteristics, we chose this lake as our primary study area. Because of its size and large clusters of low backscatter outside the shelf, it can be studied using relatively coarse Sentinel-1 Extra-Wide Swath (EW) data at 40-metre pixel spacing. This is crucial as data over central Yamal are mostly acquired in this mode by Sentinel-1.
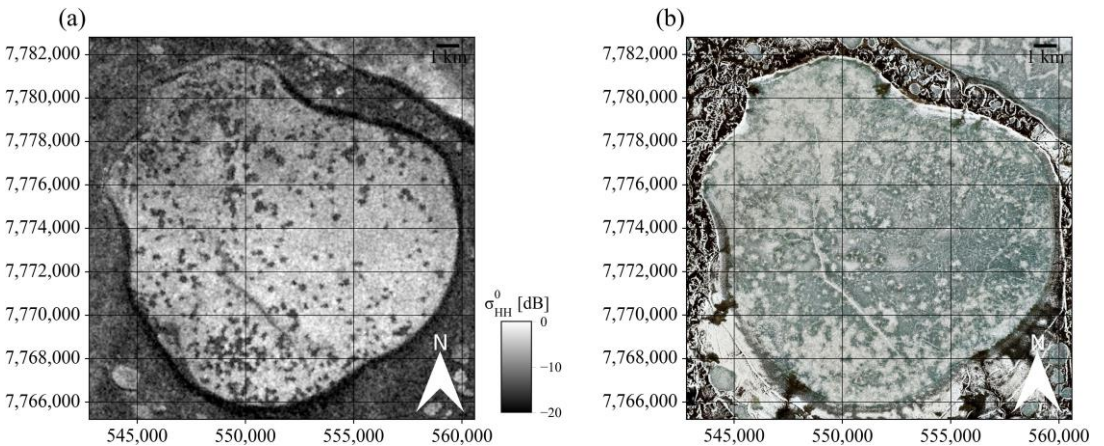


**Figure 1:** Example for clusters of low backscatter in SAR imagery and similarities to optical imagery: (a) Sentinel-1 EW HH-polarized acquisition from 27 May 2017; (b) Sentinel-2 true-colour composite from 8 June 2017.

## 2.2  Data

### Sentinel-1 SAR Data

The primary data used in this study come from the two polar-orbiting Sentinel-1 satellites, which are part of the EU's Copernicus programme. Sentinel-1A and Sentinel-1B were launched in April 2014 and April 2016, respectively. The main scientific instrument on the two satellites is an identical SAR instrument, called "C-SAR", which can be operated using different spatial resolutions and swath widths (ESA, 2013). Additionally, data can be acquired using various polarization modes: single co-polarized acquisitions, or dual co-polarized plus cross-polarized acquisitions are possible.

The default operating mode over land is the Interferometric Wide Swath mode (IW) with vertical-vertical (VV) and vertical-horizontal (VH) dual-polarization. However, data over lake Neyto are much more frequently acquired in Extra Wide Swath mode (EW) with horizontal-horizontal (HH) and horizontal-vertical (HV) dual polarization. The number of EW acquisitions over lake Neyto is more than seven times greater than the number of IW acquisitions, and no IW data are available for 2016. Therefore, we only consider EW data in this study. The main differences between the two modes are the larger swath width and coarser spatial resolution in EW mode compared to IW mode. Common pixel spacing used after pre-processing is 40 metres for EW data and 10 metres for IW data.

Sentinel-1 EW data with both HH and HV polarization channels were used to classify clusters of low backscatter in the central part of lake Neyto. The locations of mapped clusters were compared to each other between the years 2015 to 2019.

In order to assess changes in the locations between years, we used two Sentinel-1 EW acquisitions per year, hence ten images in total. For a time series analysis of the evolution of the area of backscatter anomalies, we used all available later-winter Sentinel-1 EW images, a total of 395 images (60 in 2015, 113 in 2016, 111 in 2017, 62 in 2018, 49 in 2019).

**Sentinel-2 Optical Data**

The two Sentinel-2 satellites are also part of the EU's Copernicus programme and were launched in June 2015 and March 2017. Sentinel-2A and Sentinel-2B each carry an identical multispectral imager, the 'MultiSpectral Instrument' (MSI), which acquires data in the optical and near-infrared regions of the electromagnetic spectrum in 12 spectral bands (Drusch et al., 2012). The spatial resolution depends on the band and ranges from 10 to 60 metres.

In this study, one Sentinel-2 image was used to visually highlight similarities of patterns between SAR and optical imagery for lake Neyto.

**Global Historical Climatology Network (GHCN) – Daily Data**

GHCN-Daily (Menne, Durre, Korzeniewski, et al., 2012) is a database that provides daily records of temperature, precipitation and snow over land areas worldwide (Menne, Durre, Vose, et al., 2012). In this study, we used daily air temperature records from the Seyaha station, the station closest to lake Neyto and located on the east coast of the Yamal Peninsula at a distance of approximately 80 kilometres, to assess the influence of weather conditions on ice properties in relation to observed backscatter.

## 2.3  Pre-processing of Sentinel-1 SAR images

The pre-processing of Sentinel-1 EW data was conducted with the Sentinel Application Platform (SNAP) toolbox provided by the European Space Agency (ESA). The main steps applied were sub-setting, radiometric calibration to backscatter coefficient $\sigma^0$, thermal noise removal, terrain correction, conversion to decibels (dB), and incidence angle normalization. All these steps were performed on both polarization channels (HH and HV).

## 2.4  Binary Classification of Sentinel-1 SAR images

The classification algorithm will be briefly outlined here. We describe the main steps and provide visualizations of classification outcomes in the results section of this paper.

The inputs for the binary classification algorithm are the pre-processed Sentinel-1 images in map geometry. All steps described below were applied in identical fashion to both polarization channels. The main tool used for the classification was the Python module 'scikit-image' (van der Walt et al., 2014).

First, areas outside the lake and its shelf area where ground-fast ice is present needed to be masked from the imagery. We deduced lake masks from late-autumn Sentinel-1 EW imagery and shelf masks from earlier-winter Sentinel-1 EW imagery through binary classification. The images were rescaled to fit the pixel values from -1 to 1 required for the image-processing algorithms. The steps for the image processing included bilateral filtering to remove noise from the images, auto-levelling to balance out the unevenly distributed backscatter level across the lake, and Yen-thresholding (Yen et al., 1995) to automatically classify the images. The outputs of these steps are two classified binary images, one for the HH-channel and one for the HV-channel.

To counter the problem of the lack of in-situ data for calibration, we chose a conservative approach: for the final classification map, we kept only pixels belonging to low backscatter patterns (positive class) in the binary classification outcome of both polarization channels; otherwise the pixels were assigned to the negative class (regular floating lake ice). This corresponds to a logical AND between the classification on the HH-channel and the classification on the HV-channel.

Because Yen-thresholding determines thresholds automatically, it is only applicable if clusters of low backscatter are actually present in the image. We therefore needed to apply a mechanism to detect whether these clusters were present. Our approach tests the similarity between binary classification outcomes of the two polarization channels using Cohen's Kappa score $\varkappa$ (Cohen, 1960). Only if $\varkappa$ was above 0.2, which corresponds to 'fair agreement' following Landis & Koch (1977), was the final classification map produced as defined above. If $\varkappa$ was below 0.2, all pixels in the image were assigned to the negative class.

## 2.5  Determination of variations of locations between years

Clusters of low backscatter emerge primarily from locations of only a few pixels. Over time, these clusters widen out significantly. New clusters can form later and merge with the widening clusters, but only very rarely do pixels of low backscatter clusters revert to high backscatter within a single year. A comparison of Figures 2 and 3 in the Section 3 (Results) provides some explanation for these observations.

The basis of our analysis are two comparisons of five single binary classification results from the years 2015, 2016, 2017, 2018 and 2019. The images and their respective acquisition dates for the comparisons were chosen according to two criteria based on the observations of pattern development over time as described above. The images selected for the first comparison are the ones where the classified pattern area first exceeded 20km². The images

for the second comparison were acquired on the last date the patterns were detectable. After snowmelt sets in, the clusters of low backscatter can no longer be observed, because very low backscatter is observed from the entire lake surface.

For the two comparisons themselves, we calculated the mean minimum distance and percentages of intersecting areas between polygonized classification outcomes of the positive class (backscatter anomalies), pairwise for all years. This is the first time that study of the phenomenon has focused on single objects and the spatial relationships among them. For the most part, the Python packages Shapely (Gillies, 2007) and Fiona (Gillies, 2011), which are essential tools for geospatial programming, were used for the calculations. Since the mean minimum distance calculated from one polygon set A to another polygon set B is not equal to the mean minimum distance of polygon set B to polygon set A, the result of our calculations is a square matrix of shape 5x5 (because of 5 distinct years). Similarly, the percentage of intersecting areas is also asymmetrical, as it is calculated as the area of intersection between images of two years divided by the classified area in one year. Hence, the result is also a 5x5 matrix. Additionally, we calculated cumulative counts of positively classified pixels for the two comparisons, ranging from 0 (no occurrence in any year) to 5 (the pixel was classified positively in all five years). Further, we considered a time series of classified pattern areas for our interpretations.

## 3   Results

### 3.1  Classification results

The classification results of the positive class for the first criterion, where the pattern area first exceeded 20 km$^2$ in 2015 to 2019, are shown in Figure 2 (a)–(e). Similarly, the classification results of the positive class for the second criterion, the date (2015 to 2019) of the last available useful acquisitions in the years concerned are shown in Figure 3 (a)–(e). These images may serve for a visual assessment of the binary classification outcome and to aid understanding of the other results. The expansion of clusters of low backscatter can seen in part by comparing Figure 2 (a)–(e) with Figure 3 (a)–(e).
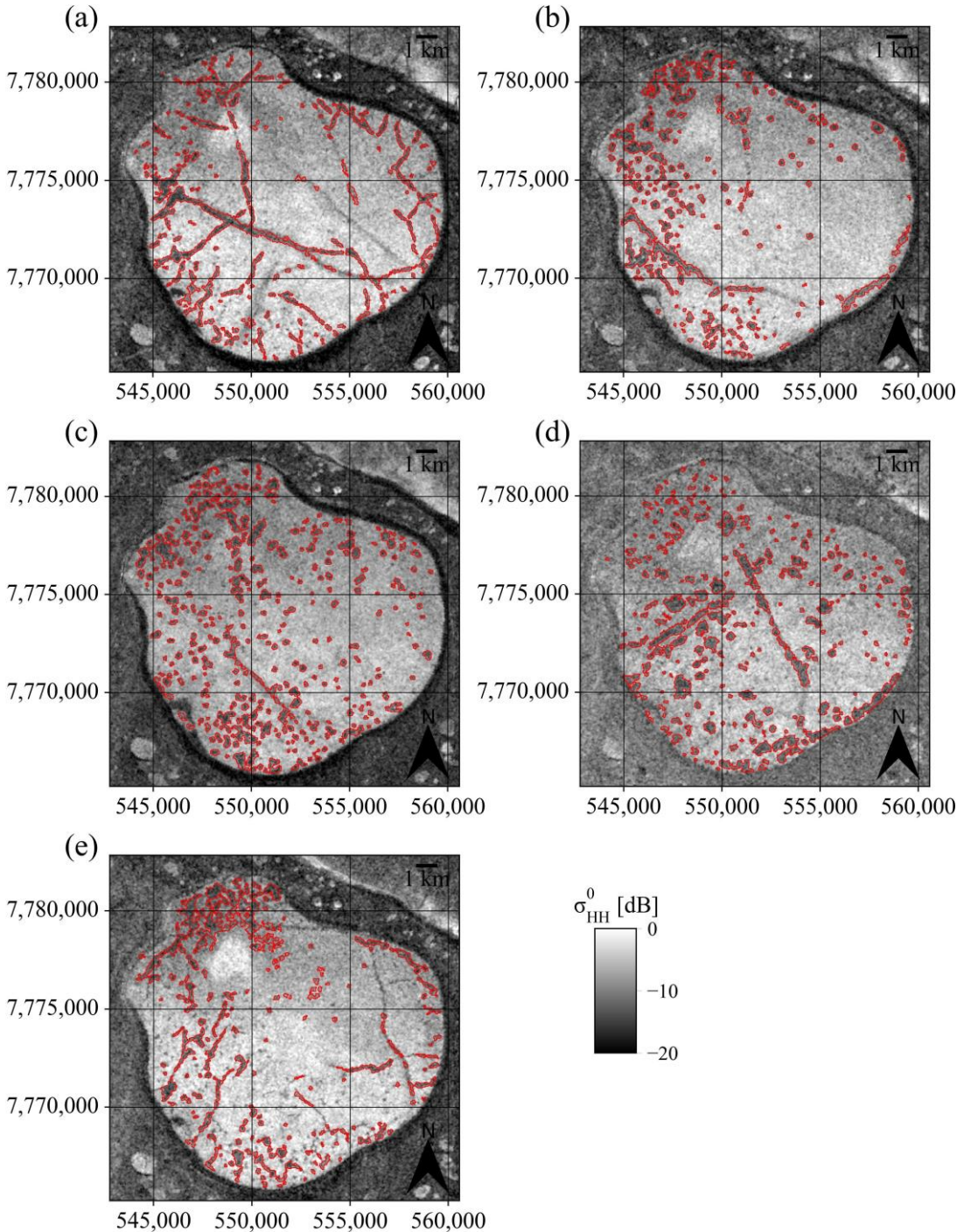
**Figure 2:** Sentinel-1 EW HH-polarized SAR images where the classified pattern area first exceeded 20 km2. Red outlines show polygonized results of the positive class from the automatic binary classification. (a) 2015, (b) 2016, (c) 2017, (d) 2018, (e) 2019.

**Figure 3:** Last available Sentinel-1 EW HH-polarized SAR images before snowmelt onset in each year from 2015 to 2019. Red outlines show polygonized results of the positive class from the automatic binary classification. (a) 2015, (b) 2016, (c) 2017, (d) 2018, (e) 2019.
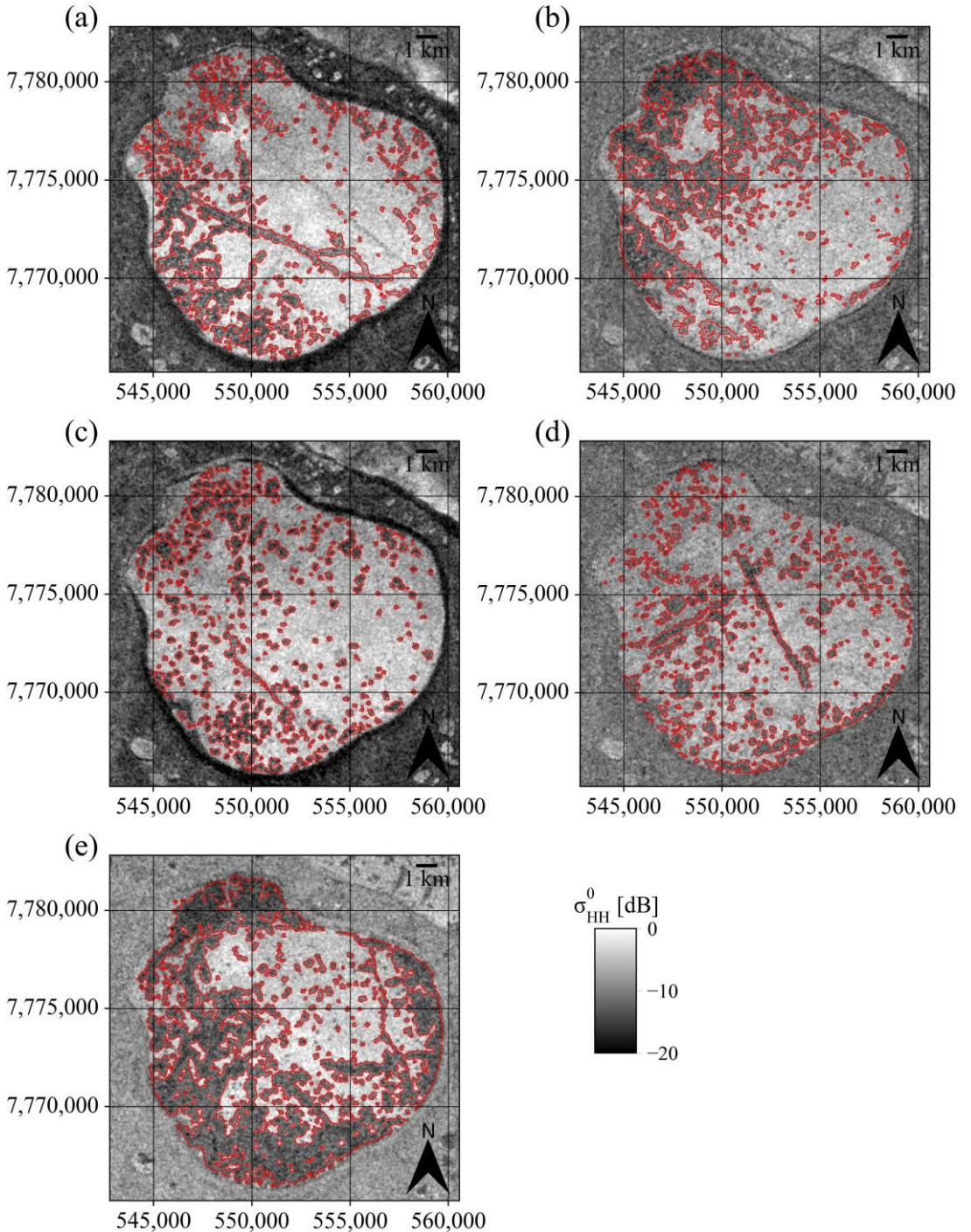
For an assessment of differences in ice conditions across the years, backscatter levels of regular floating lake ice (negative class) may be of interest. For the images in Figures 2 and 3, mean $\sigma^0$ of pixels in the negative class of each polarization channel is of similar magnitude across years (Table 1).

**Table 1:** Mean backscatter coefficient σ0 of pixels in the negative class (regular floating lake ice) for the years 2015 to 2019 and the two criteria used for image selection. Images for criterion 1 are those where the area of the anomalies first exceeded 20 km²; images for criterion 2 are the last useful acquisitions in the year. The HH-channel images for both criteria are shown in Figures 2 and 3, respectively.

| Year | $\sigma^0_{HH}$ criterion 1 (Figure 2) | $\sigma^0_{HV}$ criterion 1 | $\sigma^0_{HH}$ criterion 2 (Figure 3) | $\sigma^0_{HV}$ criterion 2 |
|------|------|------|------|------|
| 2015 | -5.9 dB | -17.8 dB | -5.5 dB | -17.2 dB |
| 2016 | -5.8 dB | -18.3 dB | -6.5 dB | -18.5 dB |
| 2017 | -7.1 dB | -17.1 dB | -6.2 dB | -17.4 dB |
| 2018 | -6.8 dB | -19.7 dB | -7.6 dB | -20.3 dB |
| 2019 | -5.4 dB | -19.4 dB | -5.5 dB | -18.8 dB |

In general, a steady increase of pattern area can be observed in late winter in every year from 2015 to 2019 (Figure 4 (a)–(e)). Alongside this general trend, minor fluctuations in the classified area of low backscatter are visible for 2016, 2017, 2018 and 2019. These fluctuations are particularly apparent in early 2018. They may be caused partly by noise in the images or by imperfections in the classification method. Rare small clusters of low backscatter that revert to high backscatter may also contribute. Since no reference data are available, it is impossible to state the cause of the fluctuations with confidence, but air temperatures close to or slightly above 0°C may play a role. However, for this study, the important aspects are the relatively steady increase of pattern area in late winter and the time of the start of this general uptrend. Variations in air temperature behave differently among the years, but are of similar magnitude, and air temperature rarely exceeds 0°C during the analysis periods of backscatter anomalies.
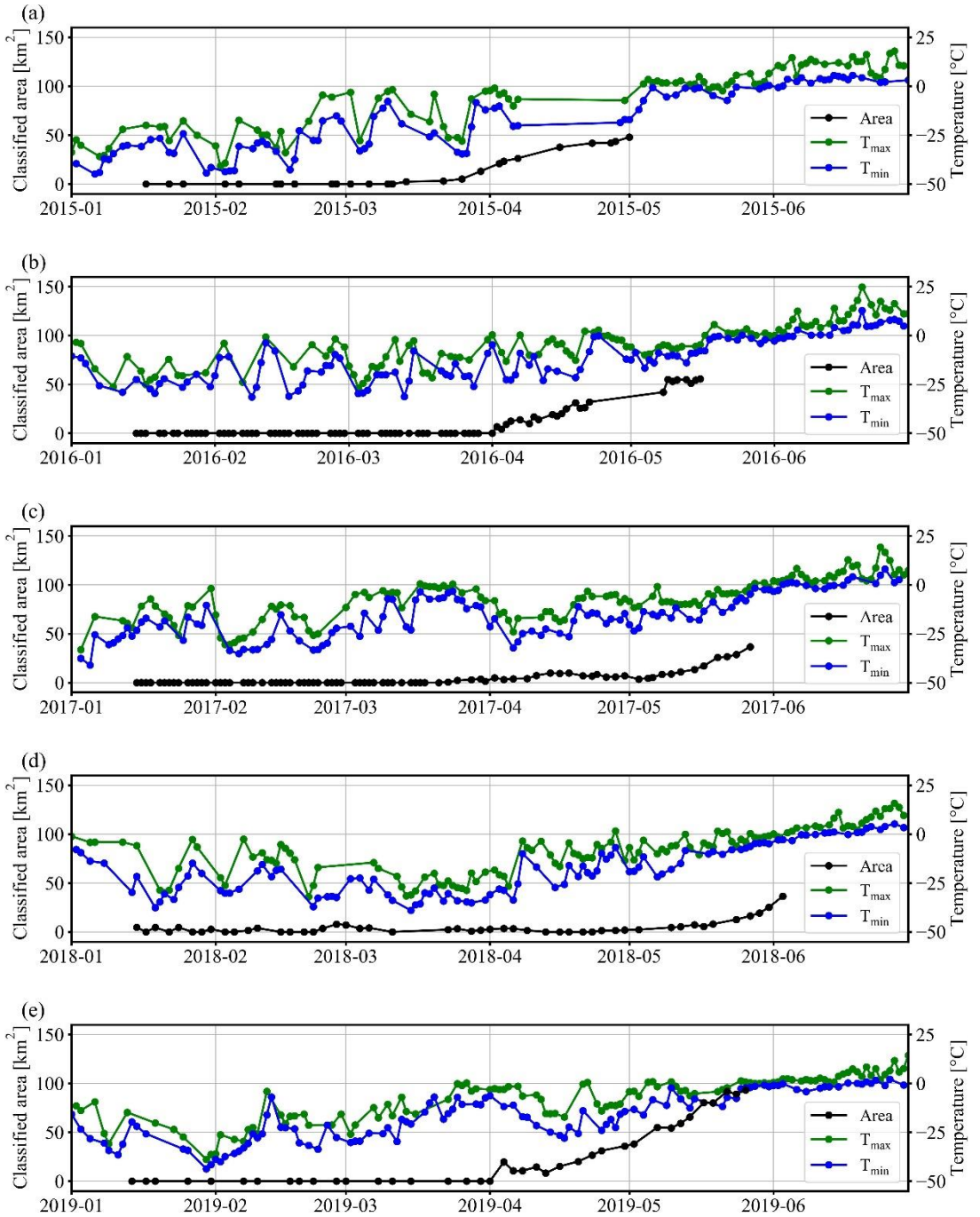
**Figure 4:** Temporal evolution, January to June, of classified area of clusters of low backscatter (black), and minimum (blue) and maximum (green) air temperatures recorded at the Seyaha weather station for (a) 2015, (b) 2016, (c) 2017, (d) 2018, (e) 2019.

## 3.2 Mean Minimum Distances

Table 2 shows the results of mean minimum distance calculations for the first criterion, where pattern area first exceeded 20 km². The mean minimum distances range from 65 metres to 227 metres, where most lie between 100 and 200 metres. Similarly, mean minimum distances for the second criterion, the last available useful acquisitions in the year, can be seen in Table 3. The mean minimum distances range from 52 to 227 metres. Of particular note is that mean minimum distances in the 2019-column in Table 3 are significantly smaller in comparison with Table 2. Smaller distances would be expected, due to the expansion of patterns as described above. Distances between consecutive years are of similar magnitude to those found over the longer period.

**Table 2:** Mean minimum distances between objects pairwise for different years, where pattern area first exceeded 20 km2. The figures are for the mean minimum distance between classified objects in the year in the row to all classified objects in the year in the column.

| Year | 2015 | 2016 | 2017 | 2018 | 2019 |
|------|------|------|------|------|------|
| 2015 | -    | 175 m | 69 m | 118 m | 137 m |
| 2016 | 139 m | -    | 65 m | 97 m | 123 m |
| 2017 | 137 m | 210 m | -   | 117 m | 178 m |
| 2018 | 156 m | 227 m | 107 m | -  | 176 m |
| 2019 | 173 m | 170 m | 88 m | 91 m | -    |

**Table 3:** Mean minimum distances between objects pairwise for different years, for the last useful acquisitions in the year in question. The figures are for the mean minimum distance between classified objects in the year in the row to all classified objects in the year in the column.

| Year | 2015 | 2016 | 2017 | 2018 | 2019 |
|------|------|------|------|------|------|
| 2015 | -    | 60 m | 63 m | 120 m | 25 m |
| 2016 | 132 m | -   | 94 m | 140 m | 18 m |
| 2017 | 65 m | 52 m | -   | 117 m | 15 m |
| 2018 | 82 m | 77 m | 77 m | -    | 18 m |
| 2019 | 201 m | 79 m | 108 m | 161 m | -  |

## 3.3 Intersections

The intersections deduced from the results for the first criterion, where pattern area first exceeded 20 km², are generally rather low, with 30% being the maximum and most others being between 10% and 20% (Table 4). In comparison, the results of the intersection calculations for the second criterion, the last available useful acquisitions in a particular year, are displayed in Table 5. As could be expected, intersections are generally larger, but they do not exceed 50%, except for the percentages of intersections deduced using the polygon set

from 2019, which are approximately two thirds (Table 5, 2019-column). As with the mean minimum distance calculations, no clear differences between the intersections for consecutive years and those between other years can be seen.

**Table 4:** Intersections of area of objects classified in one year (row) with area of objects classified in another year (column), where pattern area first exceeded 20 km2. The results correspond to the area of intersection between the two years, divided by the area in the row-year.

| Year | 2015 | 2016 | 2017 | 2018 | 2019 |
|------|------|------|------|------|------|
| 2015 | - | 16 % | 18 % | 17 % | 14 % |
| 2016 | 17 % | - | 23% | 25 % | 30 % |
| 2017 | 15 % | 18 % | - | 18 % | 17 % |
| 2018 | 14 % | 21 % | 18 % | - | 15 % |
| 2019 | 15 % | 30 % | 22 % | 19 % | - |

**Table 5:** Intersections of area of objects classified in one year (row) with area of objects classified in another year (column) for the last useful acquisitions in the years concerned. The results correspond to the area of intersection between the two years, divided by the area in the row-year.

| Year | 2015 | 2016 | 2017 | 2018 | 2019 |
|------|------|------|------|------|------|
| 2015 | - | 45 % | 28 % | 27 % | 68 % |
| 2016 | 38 % | - | 29% | 27 % | 65 % |
| 2017 | 37 % | 45 % | - | 27 % | 66 % |
| 2018 | 36 % | 44 % | 26 % | - | 65 % |
| 2019 | 35 % | 39 % | 26 % | 25 % | - |

## 3.4 Cumulative counts of occurrences

Cumulative pixel counts are visualized in Figure 5 (a) for the images where pattern area first exceeded 20km², and in Figure 5 (b) for the last useful acquisitions in the years concerned. The colourbar indicates how often a pixel was classified positively in the five images taken in the five different years (one image/year). Figure 5 (a) is clearly dominated by cumulative counts of 1 and 2, while in Figure 5 (b), where pattern area in the single images is generally larger, wider areas of counts higher than 2 can be observed, although counts of 4 and 5 are still relatively rare. Both sub-figures clearly show that the occurrence of clusters of low backscatter is a lot more frequent in the northern and western parts of the lake.
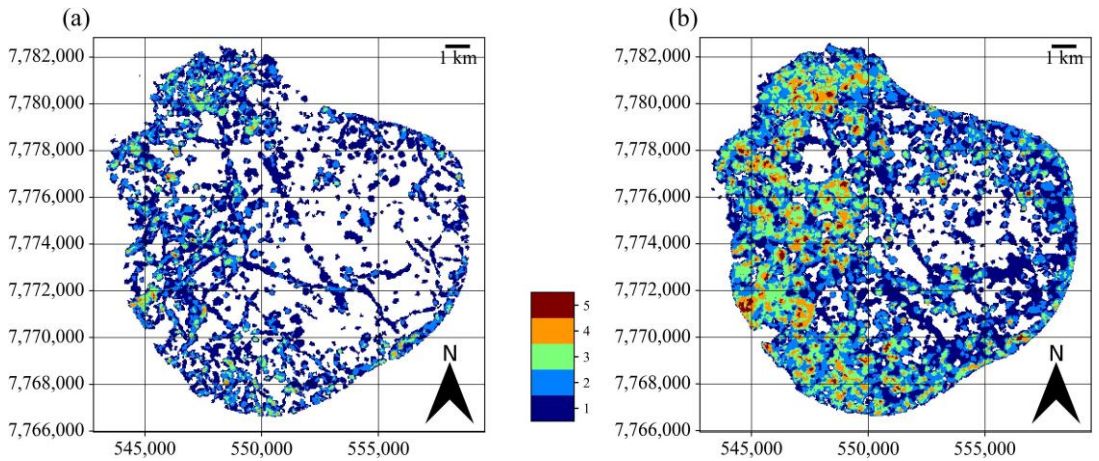
**Figure 5:** Cumulative pixel counts of backscatter anomalies identified from Sentinel-1 acquisitions from 2015 to 2019: (a) for images where classified area first exceeded 20 km2, (b) for last useful images in the years concerned.

## 4    Discussion

In this study, we base our analyses on the comparison of images from five different years selected by applying two criteria. Conventionally, time series analyses refer to day of year (DOY). Our results demonstrate (Figure 4) that the time at which the pattern area starts to increase steadily varies significantly between years. This fact supports our choice of an alternative approach – the spatial extent of the phenomenon. In general, it difficult to find a suitable criterion, since some clusters form earlier and some later in the year, and this timing may also vary spatially between years. With our approach, we tried to provide an overview of similarities and variations of locations from 2015 to 2019. The threshold of 20 km² was chosen for our analysis of the main locations from which large clusters expand. A lower threshold may not be useful due to limited spatial resolution (40-metre pixel spacing).

Binary classification results generally give a good visual impression. In situ data are, however, not available due to remoteness and safety reasons. It can be assumed that locations of anomalies are characterized by thin ice which cannot be traversed. This assumption is supported by reports from reindeer herders, who observed very thin ice on one large lake on Yamal, where ice thickness is usually more than one metre (Pointner et al., 2019). The lack of direct reference data collected on site also impedes an assessment of ice conditions in relation to the weather in different years. We could only compare our results with air temperature data recorded at a weather station located 80 km away on the coast. However, these temperatures were below 0°C throughout almost the whole of the analysis periods and variations were of similar magnitude across years, which suggests similar ice conditions in different years. This assumption is also supported by similar backscatter values reported for regular floating lake ice (negative class in our classification) across the years.

Some point-like and elongated objects can be seen, especially in Figures 2 (a), (b) and (c), which are characterized by medium contrast with the surrounding high backscatter of floating lake

ice and may belong to the class of low backscatter clusters. However, since no in-situ data are available, we used a conservative strategy and focused on mapping objects that are characterized by higher contrast.

Some objects, primarily visible in earlier acquisitions, may resemble linear fractures in the Sentinel-1 SAR images (especially in Figures 2 (a) and (e)). However, the characteristic expansion of the areas over time is the same for these objects (compare Figures 3 (a) and (e)) and fractures commonly exhibit high backscatter, which is often obscured by the high backscatter of floating lake ice in late winter (Duguay & Pietroniro, 2005).

Linear features in optical and SAR imagery of ice on the lakes on Yamal may also be attributed to leakage of methane and associated geological structures and faults (Bogoyavlensky et al., 2016, 2018). If the clusters of low backscatter for ice on lake Neyto were indeed caused by upwelling gas and accumulation under the ice layer, analysing any changes of locations could be interesting to determine the duration of seeping from a particular point source. However, the limited spatial resolution may be problematic for that purpose.

The calculated mean minimum distances may not at first seem large when compared to the 40-metre pixel spacing, but keep in mind that if two polygons overlap, the minimum distance is zero. So, we still consider the calculated distances as signs of large spatial variations of occurrences. Although we argue that this metric is the most difficult to interpret, we nevertheless think it is useful for highlighting that the variations between consecutive years are similar to those between other years.

Intersection metrics are easier to interpret. Intersections calculated for the images where the classified pattern area first exceeded 20 km² (Table 4) are generally rather low, often below 20%, which shows that clusters emerge mainly from different regions of the lake in different years, although there is always some overlap. Intersections calculated for the last useful acquisitions in any one year (Table 5) are significantly higher, which can be explained partly by the fact that the classified pattern area is generally larger for the images used here. Intersections with the polygon set deduced from the 2019 image range from 65% to 68% (last column on the right in Table 5), but the classified area for the 2019 image also covered nearly half of the total lake area (compare Figures 3 (e) and 4 (e)).

Strong spatial variations of cluster locations can also be seen in Figure 5. There is a strong spatial difference in the emergence of patterns between the five years (Figure 5 (a)), although higher cumulative pixel counts can be seen in Figure 5 (b), where the classified pattern areas in the single images were generally higher. Especially interesting is the more frequent occurrence in northern and western parts of the lake. Bogoyavlensky et al. (2018) discuss the occurrence of patterns in these same parts of lake Neyto in single images. They associate the patterns (but without quantitative analyses) with a nearby gas field that stretches under these areas of lake Neyto. Our results may support their assumption that the clusters of low backscatter are caused by gas emissions.

The cumulative pixel counts (Figure 5) provide some insight into the spatial variation of backscatter anomalies during the five years. However, the analysis of spatial relationships between individual objects (mean minimum distances and intersections), as presented here in relation to backscatter anomalies of lake ice for the first time, may reveal actual variations

between the years. Results suggest that spatial changes between consecutive years are similar to changes over the entire time period, which cannot be deduced from cumulative pixel counts.

# 5    Conclusions

The purpose of this study was to examine the interannual variability of clusters of low backscatter on Sentinel-1 SAR images of lake ice on lake Neyto in northwestern Siberia. Our results show that there are significant spatial variations in occurrences of clusters of low backscatter between the years 2015 to 2019, although there is also always some overlap. Geospatial analysis reveals that variations are of similar magnitude, whether we look at consecutive years or the longer period. Linear structures and the more frequent occurrence of backscatter anomalies in the northern and western parts of lake Neyto may point to gas release as the primary cause of the anomalies. Methods commonly used for assessing spatial relationships of vector data can provide valuable insight into the phenomenon.

# Acknowledgements

# References

Atwood, D. K., Gunn, G. E., Roussi, C., Wu, J., Duguay, C., & Sarabandi, K. (2015, Nov). Microwave Backscatter From Arctic Lake Ice and Polarimetric Implications. IEEE Transactions on Geoscience and Remote Sensing, 53(11), 5972-5982.

Bogoyavlensky, V. I., Sizov, O. S., Bogoyavlensky, I. V., & Nikonov, R. A. (2016). Remote Detection of Surface Gas Shows Zones and Gas Blowouts in the Arctic: The Yamal Peninsula. *Arctic: Ecology and Economy, 3*(23), 4–15. (in Russian)

Bogoyavlensky, V. I., Sizov, O. S., Bogoyavlensky, I. V., & Nikonov, R. A. (2018). Technologies for Remote Detection and Monitoring of the Earth Degassing in the Arctic: Yamal Peninsula, Neito Lake. *Arctic: Ecology and Economy, 2*(30), 83–93. (in Russian)

Burn, C. R. (2005). Lake-bottom thermal regimes, western Arctic coast, Canada. *Permafrost and Periglacial Processes, 16*(4), 355–367.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Drusch, M., Bello, U. D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment, 120* (Supplement C), 25 -36.

Duguay, C. R., & Lafleur, P. M. (2003). Determining depth and ice thickness of shallow sub-Arctic lakes using space-borne optical and SAR data. *International Journal of Remote Sensing, 24*(3), 475-489.

Duguay, C. R., & Pietroniro, A. (2005). *Remote Sensing in Northern Hydrology: Measuring Environmental Change*. Washington DC American Geophysical Union Geophysical Monograph Series, 163.

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

Engram, M., Anthony, K. W., Meyer, F. J. & Grosse, G. (2013). Synthetic Aperture Radar (SAR) Backscatter Response from Methane Ebullition Bubbles Trapped by Thermokarst Lake Ice. *Canadian Journal of Remote Sensing, 38*(6), 667–682.

ESA. (2013). *Sentinel-1 User Handbook*. European Space Agency.

Gillies, S. (2007). Shapely: manipulation and analysis of geometric objects. Retrieved from https://github.com/Toblerity/Shapely

Gillies, S. (2011). Fiona is OGR's neat, nimble, no-nonsense API. Retrieved from https://github.com/Toblerity/Fiona

Grunblatt, J., & Atwood, D. (2014). Mapping lakes for winter liquid water availability using SAR on the North Slope of Alaska. *International Journal of Applied Earth Observation and Geoinformation, 27*, 63–69.

Jeffries, M. O., Morris, K., Weeks, W. F., & Wakabayashi, H. (1994). Structural and stratigraphic features and ERS 1 synthetic aperture radar backscatter characteristics of ice growing on shallow lakes in NW Alaska, winter 1991–1992. *Journal of Geophysical Research: Oceans, 99*(C11), 22459–22471.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*(1), 159–174.

Menne, M. J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., . . . and others (2012). *Global Historical Climatology Network-Daily (GHCN-Daily), version 3.26*. NOAA National Climatic Data Center. (accessed on 6 April 2020)

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *Journal of atmospheric and oceanic technology, 29*(7), 897–910.

Pointner, G., Bartsch, A., Forbes, B. C., & Kumpula, T. (2019). The role of lake size and local phenomena for monitoring ground-fast lake ice. *International Journal of Remote Sensing, 40*(3), 832-858.

Surdu, C. M., Duguay, C. R., Brown, L. C., & Fernández Prieto, D. (2014). Response of ice cover on shallow lakes of the North Slope of Alaska to contemporary climate conditions (1950-2011): radar remote-sensing and numerical modeling data analysis. *The Cryosphere, 8*(1), 167–180.

Surdu, C. M., Duguay, C. R., Pour, H. K., & Brown, L. C. (2015). Ice Freeze-up and Break-up Detection of Shallow Lakes in Northern Alaska with Spaceborne SAR. *Remote Sensing, 7*(5), 6133–6159.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., ... and others (2014). scikit-image: image processing in Python. PeerJ, 2, e453. Retrieved from https://doi.org/10.7717/peerj.453

Wakabayashi, H., Weeks, W. F., & Jeffries, M. O. (1993). A C-band backscatter model for lake ice in Alaska. In *Proceedings of IGARSS '93 - IEEE International Geoscience and Remote Sensing Symposium* (Vol. 3, pp. 1264-1266).

Walter Anthony, K. M., Anthony, P., Grosse, G., & Chanton, J. (2012). Geologic methane seeps along boundaries of Arctic permafrost thaw and melting glaciers. *Nature Geoscience, 5*(6), 419–426.

Wik, M., Varner, R. K., Anthony, K. W., MacIntyre, S., & Bastviken, D. (2016). Climate-sensitive northern lakes and ponds are critical components of methane release. *Nature Geoscience, 9*(2), 99–105.

Yen, J.-C., Chang, F.-J., & Chang, S. (1995). A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing, 4*(3), 370–378.