Reply to Anonymous Referee #1:

Dear Anonymous Referee #1,

we thank you very much for taking the time to review our manuscript and for providing detailed and constructive comments! In the following, we will reply to all your comments sequentially.

General comments:

The article 'Mapping potential signs of gas emissions in ice of lake Neyto, Yamal, Russia using synthetic aperture radar and multispectral remote sensing data' provides an extensive analysis of backscatter anomalies linked to possible gas emissions for Lake Neyto, Yamal, Russia. Multiple image products and processing techniques were used to support the authors' hypothesis and the results are supported by the existing literature. The study is particularly interesting due to its connection to gas emission from the warming Arctic and the multiple recent publications addressing similar elements from Alaska and northern Canada.

Reply: We are pleased to hear that our study seems interesting in the context of other works from Alaska and Northern Canada. Thank you!

The literature review provided in the article is well constructed and provides a good background. Furthermore, the discussion is well organized and outlines how the results in this study reflect and differ from similar work.

Reply: We are glad to receive positive feedback for these sections, thank you!

The methods section requires the most revision in the current manuscript. There must be further documentation of the Sentinel-1 catalog used (dates, number of images, and gaps between images). There are small concerns about the Sentinel-1 image processing done regarding how noise in the images was addressed. While many image processing techniques are used, the description given is not adequate. These techniques should be better described to ensure that the method can be replicated in future studies. Specific comments are provided regarding these issues, in addition to short comments about figures and sentence structure.

Reply: Yes, while a lot of emphasis was put on the introduction, the results, and their interpretation; we agree that the description of the methods became too short and further details are required. In case we were asked to submit a revised version of the manuscript, we would suggest listing all relevant software libraries used (including their versions) and to explicitly indicate the methods used from these libraries. For examples and replies to the other issues raised here, please see the replies to your specific comments in the following! Detailed parameters of the methods used would also be provided if we were asked to submit a revised version.

Specific Comments

Line 11: Include the actual percentage of holes mapped in the VHR data that relate to the SAR anomalies.

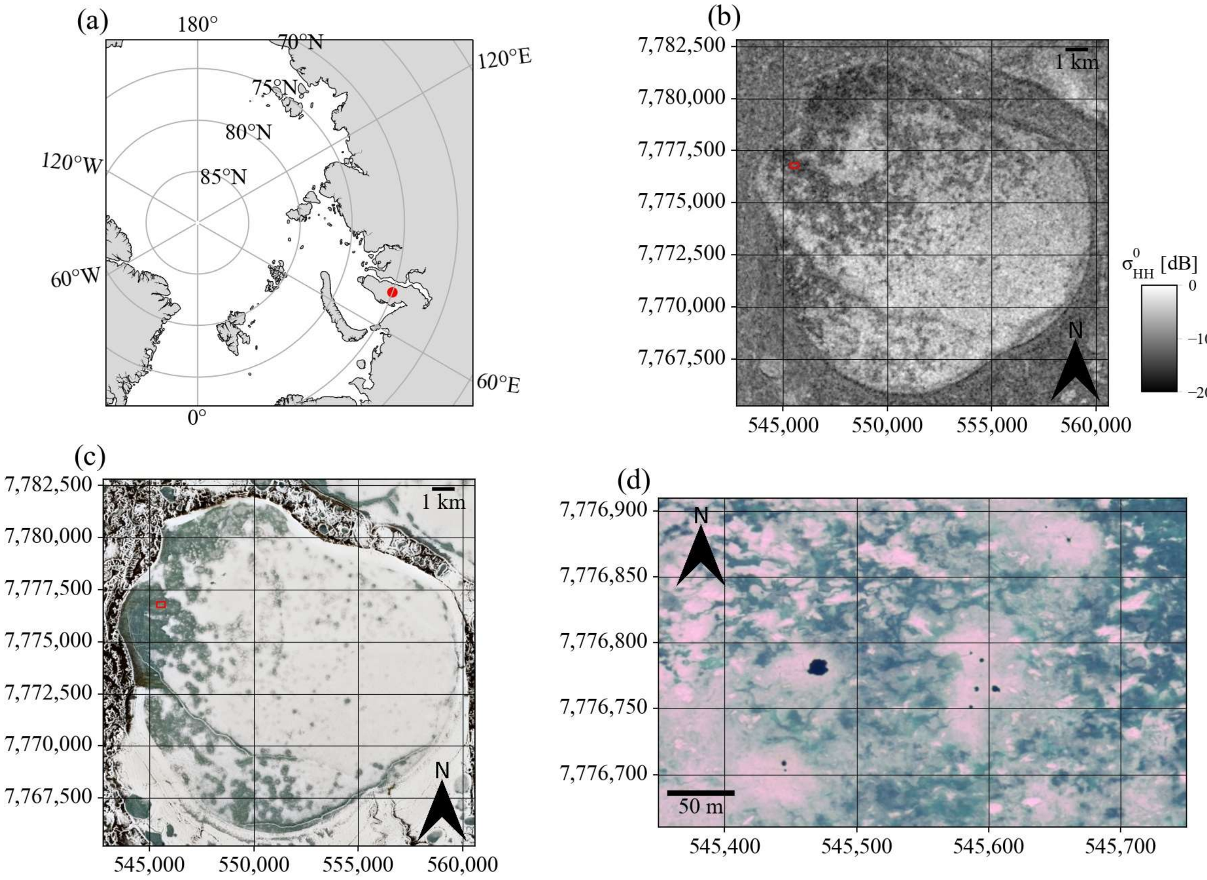
Reply: Agreed, we mentioned it in the results section (68%), but it would be good to also include it in the abstract.

Lines 100-118: These lines are more suited for a study site section, an additional section could be added before 'Data' or as a subheading of the same section to present the information. Some additional information about lake Neyto would also strengthen the description of the study site (temperature, precipitation, lake properties, distance to major settlements/coordinates). This could also be addressed in Figure 1 by adding a fourth frame that provides a geographic context.

Reply: We agree. The inclusion of a study site section was also suggested by Anonymous Referee #2. We would transfer those lines, as you suggested and suggest adding the following before the transferred lines:

"Lake Neyto (other title: Neyto-Malto), 70.073 °N, 70.350 °E, is located in the central part of the Yamal Peninsula, ca. 80 km away from the closest settlement Seyakha and ca. 80 km away from the Bovanenkovo gas field. The lake has the second biggest area (214 km^2) in Yamal after Yaroto-1 lake. The length of the shoreline is about 60 km and the lake measures approximately 17.8 km in the south – north direction and 16.5 km from west to east. The lake is relatively shallow, reaching 17 m at the north-west corner, but the average depth does not exceed 3 m, which results in a significant mixing of water masses during summer (Edelstein et al., 2017). Wide shelf areas up to 800 m can be found within the lake, whereas at the deepest part, several depressions with diameters up to 500-800 m are documented (Edelstein et al., 2017). Lake shores are mostly cliffs up to 25 m high, sometimes with tabular ground ice exposures. The ground temperature at 2 m depth in the surroundings of the lake is approximately -1.5 °C (Obu et al., 2020). The Snow Depth Liquid Water Equivalent (SDLWE) generally increases gradually in winter and spring until melt-onset and typically ranged between 15 cm and 20 cm at its maximum in recent years (Hersbach et al., 2018)."

We also think including a fourth frame would be useful. This is the suggested new figure:



Section 2.1: Further discussion on the Sentinel-1 images used is needed. While the other imagery sources use one or a handful of images, Sentinel-1 appears to be the focus of this study. Therefore, a table listing the years of data, the number of images, and the average temporal gap between imagery would be good to include. Alternatively, a calendar plot showing the dates of the study period with associated acquisitions (Sentinel-1, Worldview, PALSAR, Sentinel-2, and Landsat-8) would be a good way to convey the amount/temporal resolution of the imagery used to the reader.

Reply: We would like to follow your first suggestion and provide a table listing the years of data, the number of images, and the average temporal gap between imagery directly in the manuscript. This would be the suggested table:

Year	Number of images	Average temporal gap
2015	29	4d 7h
2016	88	1d 13h
2017	112	1d 7h
2018	52	2d 23h
2019	41	3d 14h

Anonymous Referee #2 also suggested to include a table of all acquisitions, but together with metadata (local sensing time, mean projected local incidence angle). Since altogether more than 300 Sentinel-1 acquisitions were used, we suggest to provide a detailed table including the scene ID, acquisition time and mean incidence angle as a supplement, and also indicate the exact scenes that were used for calculating the lake masks and the shelf masks. In particular, we propose to include 4 tables in total: One for all EW scenes used for calculating the time series, one for all IW scenes used, one for the EW scenes used for calculating the lake masks and one for the EW scenes used for calculating the shelf masks. What should also be included in the manuscript is that some products had to be assembled using the "slice assembly" operator in SNAP, when products have been sliced directly over the lake. We also suggest indicating these products in the tables and report a common mean incidence angle for them over the lake. If we were asked to resubmit the manuscript, we would suggest including the tables as supplement in ".csv"-format. We have already prepared these tables. Please see as an example the table for the scenes used for calculating the lake masks below:

scene ID	local time datatake start	mean projected loical incidence angle over lake	slice assembled	utc time datatake start
S1A_EW_GRDH_1SDH_20141005T020153_20141005T020253_002688_003004_6547	2014-10-05 08:01:53+06:00	35.6		2014-10-05 02:01:53+00:00
S1A_EW_GRDM_1SDH_20150922T124914_20150922T125018_007828_00AE74_FDE1	2015-09-22 17:49:14+05:00	43.4		2015-09-22 12:49:14+00:00
S1A_EW_GRDM_1SDH_20160913T122532_20160913T122632_013034_014A68_E6A2	2016-09-13 17:25:32+05:00	29.2		2016-09-13 12:25:32+00:00
S1A_EW_GRDM_1SDH_20170926T015430_20170926T015530_018540_01F3F2_4166	2017-09-26 06:54:30+05:00	40.2		2017-09-26 01:54:30+00:00
S1A_EW_GRDM_1SDH_20180926T020249_20180926T020349_023863_029AB4_C93D	2018-09-26 07:02:49+05:00	35.6		2018-09-26 02:02:49+00:00

The products listed here were not slice assembled. In case of slice assembly, this would be indicated by a "Yes" in the "slice assembled"-column. We propose to also show a table for the other data used. By addressing a comment raised by anonymous reviewer#2, Figure 11 in the preprint (which showed Sentinel-2 images during lake ice break-up in 2019) is planned to be removed. With that, a single acquisition per satellite would remain. This is the suggested table for the sensors other than Sentinel-1 (please see below):

Platform	scene ID	local time datatake start	mean projected local incidence angle over lake	utc time datatake start
Sentinel-2A	S2A_MSIL2A_20160521T072952_N0202_R049_T42WWC_20160521T072949	2016-05-21 12:29:52+05:00	-	2016-05-21 07:29:52+00:00
WorldView-2	103001005502AD00	2016-05-22 13:03:13+05:00	-	2016-05-22 08:03:13+00:00
Landsat 8	LC08_L1TP_165011_20150406_20170410_01_T1	2015-04-06 12:03:51+05:00	-	2015-04-06 07:03:50+00:00
ALOS-2	ALOS2048741410-150418	2015-04-18 23:29:52+05:00	33.0	2015-04-18 18:29:52+00:00

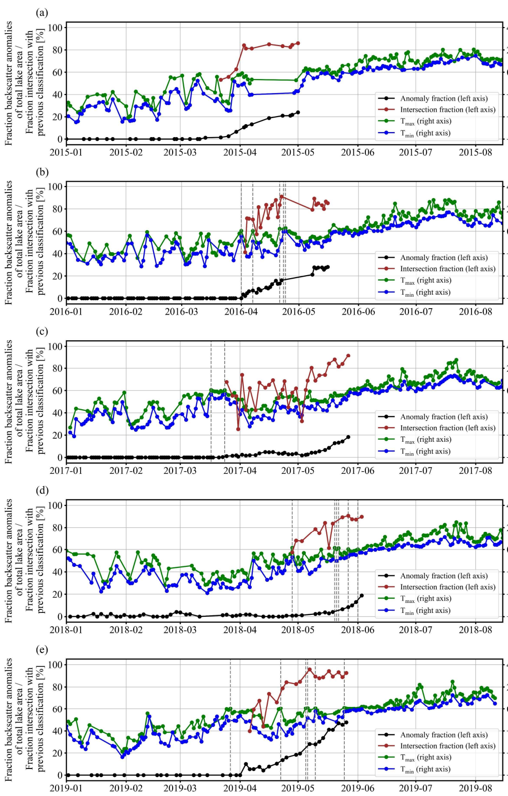
Line 192-193: No mention of speckle filtering or multi-looking is made. Was this not done? How do the authors address the issue of noise within the SAR images? The process was done for the PALSAR-2 images as stated on line 208.

Reply: For the ALOS-PALSAR-2 fully polarized data, speckle filtering was considered necessary for the polarimetric classification. The "Polarimetric Speckle Filter Operator" in SNAP was used here, as opposed to a conventional speckle filter.

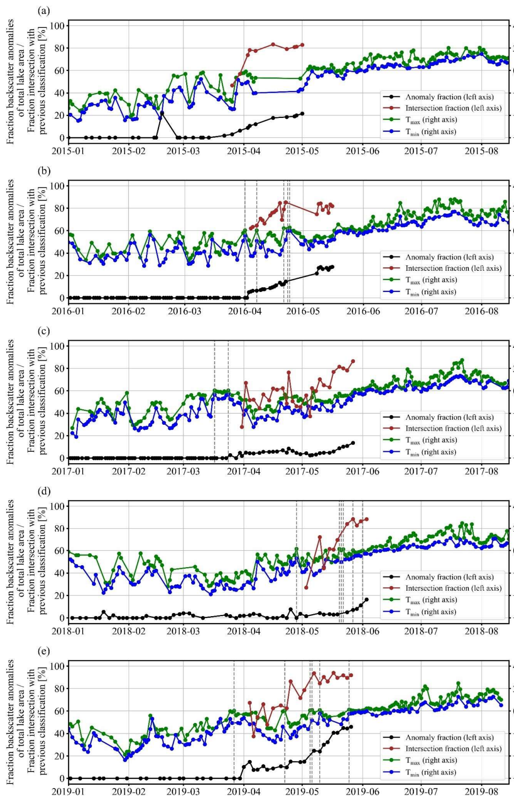
For the Sentinel-1 data, we chose a more custom approach to handle the noise as part of the classification workflow. We agree that speckle filtering is conventionally done in SAR geometry with filters specifically tuned to the theoretical statistical distribution of the speckle. However, for a flat surface (such as a lake), effects of the topography can be considered negligible and conventional speckle filters often blur the image and reduce the spatial detail a lot. With the objective of binary classification in mind, we chose to use a bilateral mean filter from the scikit-image (version 0.15.0) python library (skimage.filter.rank.bilateral_mean) to handle the noise after the pre-processing steps.

For comparison, we have now re-calculated results using the Refined Lee-filter from SNAP, which is usually considered to be good at preserving edges in the imagery. Here, for the comparison, the Refined Lee filter was applied before the terrain-correction in SNAP and the bilateral mean filter in the classification workflow was omitted. Except an obvious misclassification in February 2020, the time series appear relatively similar to those in the preprint (with bilateral mean).

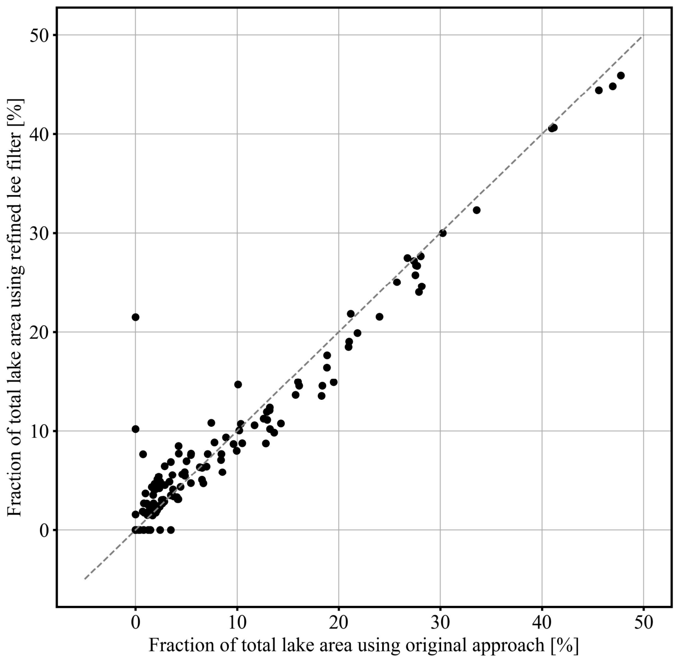
Original (with bilateral mean filter):



For comparison (with refined Lee filter):



A scatterplot for identified fractions of anomalies of total lake area (black data points in the time series plots above) using the two approaches is also illustrating similar results:



The validation metrics compared between the two approaches:

	Original (with bilateral mean filter)	**For comparison (with refined Lee filter)**
Matthews Correlation Coefficient	0.78	0.73
Cohen's Kappa Coefficient	0.78	0.73
F1 score binary	0.8	0.75
F1 score macro	0.89	0.86

Since the Matthews Correlation Coefficient, Kappa and the F1 score binary are by approximately 5% lower than for the original approach, we would prefer to keep the original workflow.

Line 194: Further explanation of the incidence angle normalization process is needed. According to Pointer et al., 2019, backscatter was normalized to 30◦ , was the same value used here? The normalization process requires further attention so that it is clear to the reader.

Reply: We agree that further explanation is required. Bartsch et al. (2017) used 2nd degree polynomial empirical functions to normalize ASAR WS HH-polarized sigma nought to a common reference incidence angle of 30 degrees. For this purpose, samples were taken manually for floating and ground-fast lake ice classes, 2nd degree polynomials were fitted to the sample data points of these two classes and the normalization function was defined as the mean of those two functions. This approach was adapted by Pointner et al. (2019) for Sentinel-1 data. This approach was also used for this study. We agree that this information should be included in the manuscript and the polynomial coefficients should also be listed.

Line 200: Line 201 states that the Sentinel-2 images were atmospherically corrected, were the Worldview-2 images also corrected?

Reply: The WorldVIew-2 images were not atmospherically corrected. Openly available atmospheric correction algorithms for WorldView-2 such as "6S" require detailed information on atmospheric conditions at the acquisition time that are simply not available to us. As an example, we refer here to the GRASS GIS implementation of the algorithm: https://grass.osgeo.org/grass76/manuals/i.atcorr.html

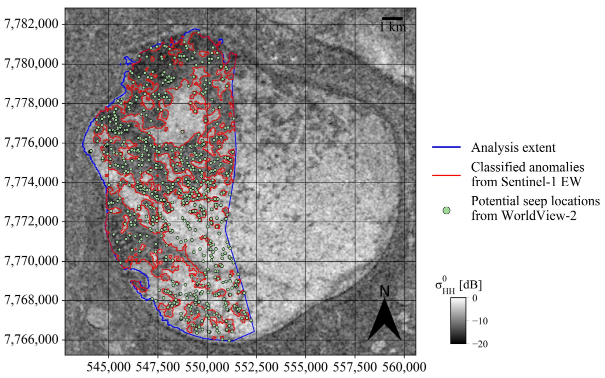
It does not include an atmospheric model for the Arctic and the aerosol concentration that is required cannot be estimated accurately for the study site and acquisition date. The reason we applied atmospheric correction to the Sentinel-2 data was that this could readily be done using the sen2cor tool. However, upon further review of literature, sen2cor also seems to fail in the automatic image-based retrieval of atmospheric parameters such as aerosol optical thickness or water vapor (König et al., 2019). Based on this, we think that sen2cor would also require parameterization based on external data to produce reliable results and as for the WorldView-2 data, this data is not available to us.

A shortcoming on our side we identified based on your comment is that we only used uncalibrated digital number (DN) data from the WorldView-2 data. We think it would be more meaningful to calibrate these data to top of atmosphere reflectance before the pan-sharpening.

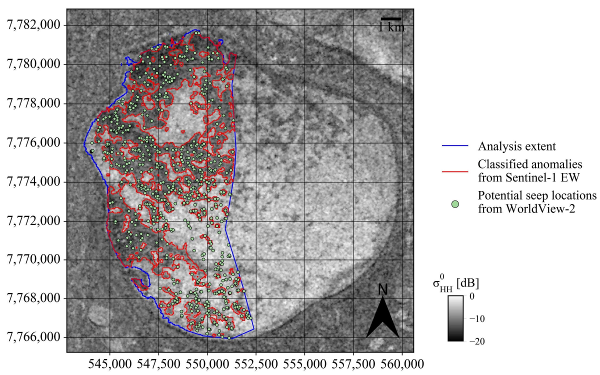
Since the Sentinel-2 images are only used for visual comparisons, we would therefore propose to use top of atmosphere reflectance for both, Sentinel-2 and WorldView-2 and show recalculated results *based on the calibrated data derived from WorldView-2 if we were asked to submit a revised version of the manuscript.*

We have already recalculated all results using TOA reflectances. Pan-sharpening has now been performed using all available bands, as this was required to address a comment by anonymous referee #2. Results are similar to the ones in the preprint. For example, please see the recalculated Fig. 7:

Old:



New:



Please find comparisons of statistics between the old (with DN) and new (with TOA reflectances and pan-sharpened using all bands) approach below:

	Old	**New**
Number of detected holes	715	718
Number of hole polygons excluded for calculating histogram using area threshold	5	10
Median hole area	4.25 m²	4.0 m²
Percentage of holes inside classified anomaly regions	68%	71%
Mean minimum distance between the points (detected holes) and the polygons (anomaly regions)	48m	38m
Median distance of all points (detected holes) lying outside the polygons (anomaly regions)	97m	67m

Line 229: A short description of the Otsu thresholding method should be included. Were backscatter values used for thresholding or were images converted to greyscale?

Reply: Agreed. Further information is required. We used backscatter values in decibels. Also, the extent of the subset is required to know which values were considered for the calculation of the Otsu threshold. The image subset extent was defined by the extent of the ArcticDEM tile 50_60_2_1_2m_v3.0. The well-know-text representation of the subset extent is: POLYGON ((70.97439396243132 70.16879530153216, 70.40771810894849 69.76495436322425, 69.22774531795417 69.9530018800304, 69.77514056883193 70.36101784242575, 70.97439396243132 70.16879530153216)). The method skimage.filters.threshold_otsu was used. We further used the method scipy.ndimage.morphology.binary_fill_holes to fill holes in the classification result, polygonized the result using gdal_polygonize.py and extracted the polygon of lake Neyto.

Here, since the incidence angle range was small for the subset region and the method determines the threshold automatically, no normalization was applied. You could now ask why we applied the normalization to the other images, since also Yen-thresholding is an automatic thresholding technique. We did this make the backscatter levels of the single images comparable to each other (e.g. for Figure 4, the boxplot).

Line 236: How were images rescaled? Was this done using a min-max normalization?

Reply: We used this method: skimage.exposure.rescale_intensity. As noted in the manuscript, the out_range interval was [-1,1] for all polarizations since this was a requirement for the other methods that were applied in the following. The choice of the in_range was more arbitrary, but since the rescaling result might be strongly affected by outliers in single images, we considered it best to use the same in_range for all images. To avoid too much clipping of the high and low values, the in_range should contain all sigma nought values that we would usually expect from the lake. On the other hand, the in_range should not be too large, to avoid too much precision loss. We decided to use a in_range of [-40 dB, 0 dB] for the co-polarized images and [-50 dB, -10 dB] for the cross-polarized images (the cross-polarized signal is usually significantly lower). We agree that this should be noted in the manuscript.

Line 272: Similar to above, more information should be provided about the watershed segmentation. Additional settings used for the process and the software packages used to perform both blob detection and segmentation should be included.

Reply: Agreed. We used these methods: skimage.segmentation.watershed and skimage.feature.blob_log. As for the other methods mentioned above, detailed parameters could be provided if we were asked to submit a revised version of the manuscript.

Figure 4: The boxplots for 2017-05-22 and 2019-05-24 are initially confusing when you look at the plot. Could the y-axis labels be dropped on the middle frames and 2019- 05-24 frame so that there is only one shared axis? Additionally, a better demonstration that the outside frames are part of the dataset shown in the middle frame would help improve the figure.

Reply: We agree. Dropping the y-axis labels on the middle and right frames and merging all frames probably leads to a better representation. Please find the suggested new figure below:

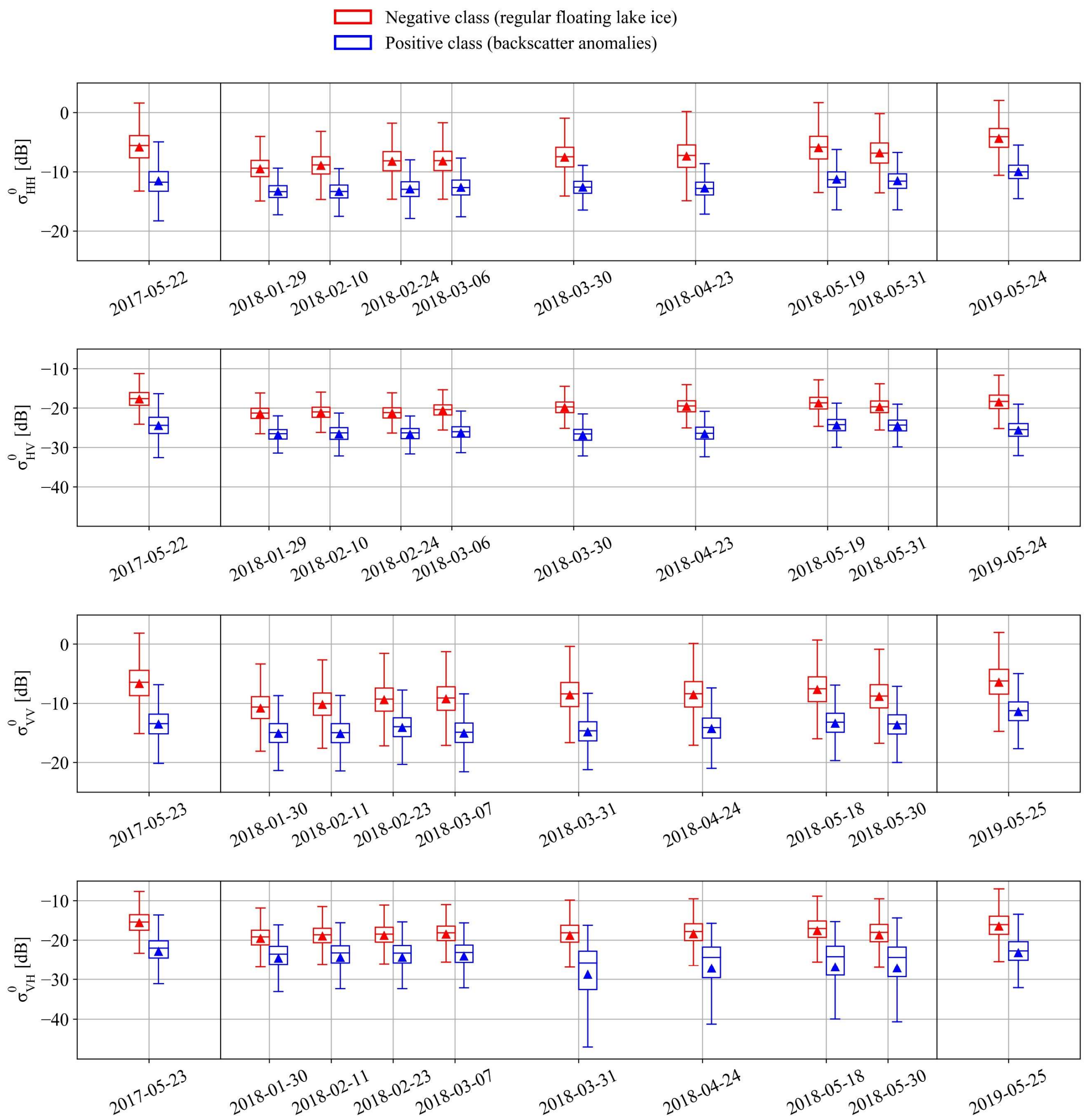


Figure 5: It would be better to show the same image/area for both a) and b) – that way the reader could see how the watershed was used to best identify the holes in the ice.

Reply: Yes, indeed. Please find the suggested new figure below:

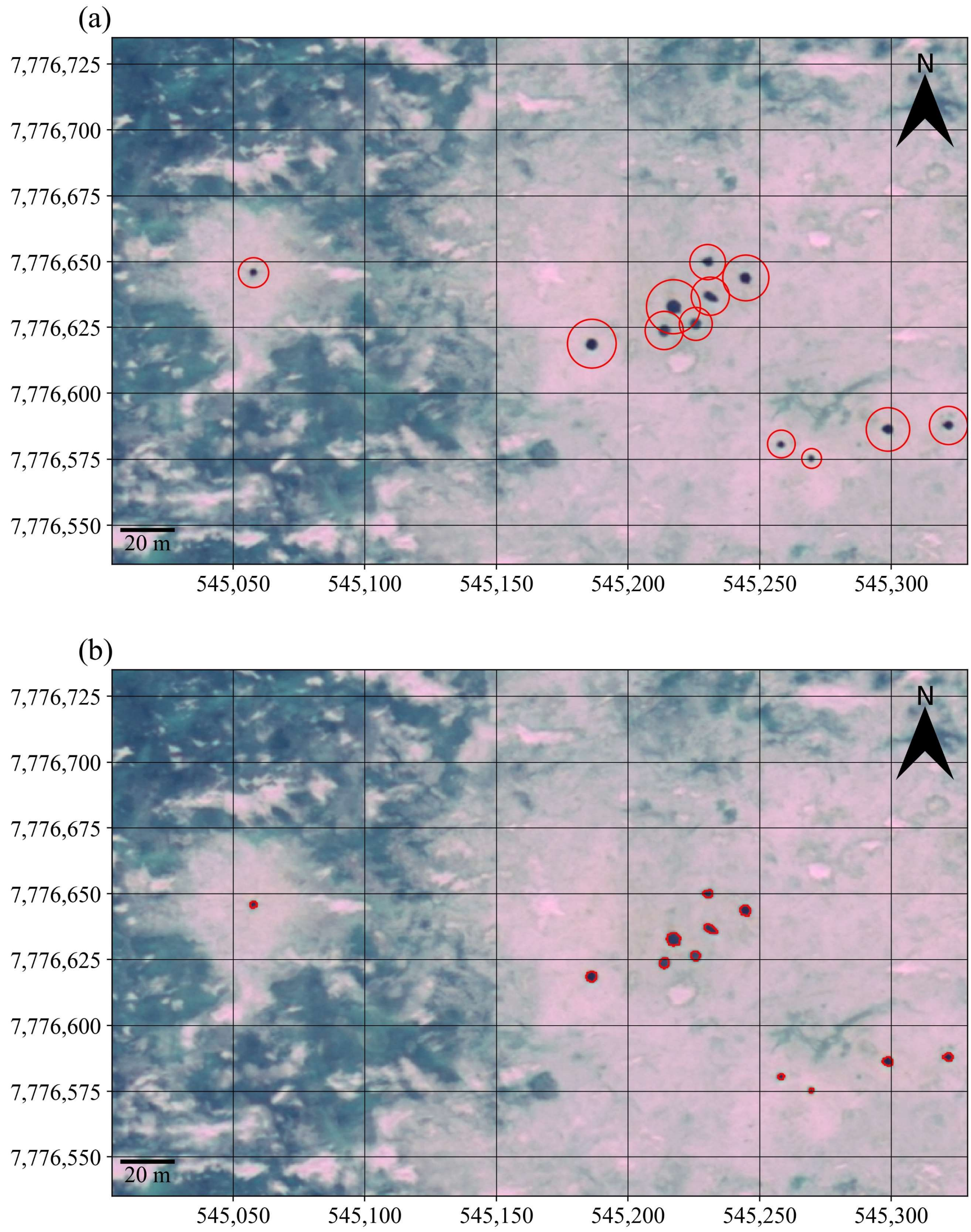


Figure 9: The intersection fraction is confusing, the explanation needs to be changed so that the metric is clear to readers. The repeated mention of 'positive class' makes the explanation wordy, possibly it could be changed to anomaly regions.

Reply: We would like to follow your suggestion to change "positive class" to "anomaly regions". Additionally, the explanation should be clearer. We suggest naming it "fraction of overlap between anomaly regions on consecutive dates".

We suggest the following changes in the text:

OLD: In order to assess the expansion of anomaly regions, the fraction of intersection of the positive class of the previous classification in time with the positive class of the classification at the timestamp indicated is shown in brown (area of intersection divided by area of the anomaly regions at the previous timestamp).

NEW: In order to assess the expansion of anomaly regions, the fraction of overlap between anomaly regions on consecutive dates is shown in brown (area of intersection between classified anomaly regions on the timestamp indicated and that of the previous timestamp, divided by area of the classified anomaly regions at the previous timestamp).

OLD CAPTION: Time series of fraction of area of anomaly regions with respect to total lake area (black, (Pointner and Bartsch, 2020)), fraction of intersection of the previous classification with the classification at the timestamp indicated (brown) for the time period after no anomalies were detected for the last time in the years concerned, maximum (green) and minimum (blue) air temperature recorded at the Seyakha weather station. The left axis indicates the fraction of anomaly region areas to total lake area and fraction of intersection. The right axis indicates air temperature. Fractions of intersection were calculated as area of intersection between anomalies detected at the timestamp indicated and that of the previous timestamp, divided by the area of anomalies detected at the previous timestamp. Gray dashed lines indicate dates where maximum air temperature exceeded 0 ◦C during the analysis periods of the SAR data.

NEW CAPTION: Time series of fraction of area of anomaly regions with respect to total lake area (black, (Pointner and Bartsch, 2020)), fraction of overlap between anomaly regions on consecutive dates (brown) for the time period after no anomalies were detected for the last time in the years concerned, maximum (green) and minimum (blue) air temperature recorded at the Seyakha weather station. The left axis indicates fraction of area of anomaly regions with respect to total lake area and the fraction of overlap between anomaly regions on consecutive dates. The right axis indicates air temperature. Fractions of overlap were calculated as area of intersection between classified anomaly regions on the timestamp indicated and that of the previous timestamp, divided by area of the classified anomaly regions at the previous timestamp. Gray dashed lines indicate dates where maximum air temperature exceeded 0 ◦C during the analysis periods of the SAR data.

Minor Typography

Line 4: ‘so far’ can be removed to improve conciseness, and it should be changed to ‘due to a lack of...’

Reply: Agreed.

Line 22: ‘remain’ should be changed to are.

Reply: Agreed.

Line 28: ‘distinguish’ should be changed to ‘distinguished’

Reply: Agreed.

Line 122: ‘threads’ should be threats?

Reply: Yes, ‘threats’ was meant.

Line 361-362: “temperature is often approaching or slightly exceeding” should be changed to “often approaches or slightly exceeds”.

Reply: Agreed.

Line 375: “is by approximately”, the ‘by’ can be removed.

Reply: Agreed.

Line 404: A citation is needed for the causes of holes on Lake Baikal.

Reply: We agree and suggest adding two citations. The new line would be (text is the same, only citations added):

“Other causes of holes in lake ice were identified for lake Baikal, such as seal breathing holes, hot springs or oil seepage (Galaziy, 1987, Petrov, 2009).”

References:

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N.: ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 30-10-2020), https://doi.org/10.24381/cds.adbb2d47, 2018.

Bartsch, A., Pointner, G., Leibman, M. O., Dvornikov, Y. A., Khomutov, A. V., and Trofaier, A. M.: Circumpolar Mapping of Ground-Fast Lake Ice, Frontiers in Earth Science, 5, 12, https://doi.org/10.3389/feart.2017.00012, 2017.

Galaziy, G. I.: Baikal in questions and answers. Eastern-Siberian Publishing. (In Russian.). 380 pp., 1987.

König, M., Hieronymi, M., and Oppelt, N.: Application of Sentinel-2 MSI in Arctic Research: Evaluating the Performance of Atmospheric Correction Approaches Over Arctic Sea Ice, Frontiers in Earth Science, 7, 22, https://doi.org/10.3389/feart.2019.00022, 2019.

Obu, J., Westermann, S., Barboux, C., Bartsch, A., Delaloye, R., Grosse, G., Heim, B., Hugelius, G., Irrgang, A., Kääb, A., Kroisleitner,C., Matthes, H., Nitze, I., Pellet, C., Seifert, F., Strozzi, T., Wegmüller, U., Wieczorek, M., and Wiesmann, A.: ESA Permafrost ClimateChange Initiative (Permafrost_cci): Permafrost ground temperature for the Northern Hemisphere, v2.0, Centre for Environmental Data Analysis, 02 November 2020, https://doi.org/10.5285/6ebcb73158b14cd5a321b7c0bc6ed393., 2020.

Petrov E.A.: The Baikal seal. « ECOS », Ulan-Ude, 2009, 176 pp. (In Russian). ISBN 978-5-85693-340-5, 2009.

Pointner, G., Bartsch, A., Forbes, B. C., and Kumpula, T.: The role of lake size and local phenomena for monitoring ground-fast lake ice, International Journal of Remote Sensing, 40, 832–858, https://doi.org/10.1080/01431161.2018.1519281, 2019.