



EGUsphere, author comment AC1
<https://doi.org/10.5194/egusphere-2022-388-AC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Maria Shaposhnikova et al.

Author comment on "Bedfast and floating-ice dynamics of thermokarst lakes using a temporal deep-learning mapping approach: case study of the Old Crow Flats, Yukon, Canada" by Maria Shaposhnikova et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-388-AC1>, 2022

This paper used a temporal deep learning mapping approach for the purpose of classifying bedfast and floating ice in the Old Crow Flats. The authors have clearly done some work, but the manuscript needs major revision and in my opinion several aspects of the study need to be revisited before it can be considered for publication.

We would like to thank the referee for valuable comments which have substantially helped improve the clarity and quality of the manuscript and stimulated interesting and constructive discussion. *Please, refer to the supplement for all figures and tables.

Major comments:

1. Your method is different from those used in previous studies (thresholding method or data mining approach), why did you not perform a comparison between them? If it works better, it is logical to proceed by presenting the method. The authors could do a more careful job comparing the proposed method with the approaches used previously by using the same dataset. Right now, there is not enough information to understand if the new method is necessary. There is almost no information about how the training and testing data sets were selected, what is the time stamps and amount of them?

Thank you for a valuable comment.

- The goal of this manuscript was not to compare methods, but rather present a new method not previously used for lake ice mapping describing its benefits for the Old Crow Flats. The method presented in our paper could, however, be compared to other previously published approaches in a follow-up study either in the same or other study area(s). Although, this method is not necessarily better than other methods it does present advantages, namely 1) it takes advantage of the temporal evolution of ice backscatter over the entire lake ice season and is not reliant on a single value, as is the case for other methods; as new SAR platforms are being launched with higher quality, spatial resolution, and denser temporal coverage it is only reasonable to look for ways to make use of the wealth of available data; 2) wetland landscapes are dynamic in nature and the presented TempCNN method does not require a lake mask, which is necessary for all the previously presented methods; 3) a well-trained TempCNN can be applied to various SAR sensors, both HH and VV polarizations, and different

spatial resolutions. The accuracy of the method was assessed using a training set and the overall accuracy values fall in the range reported by other ice regime mapping solutions.

- The annotated dataset was created through visual assessment of the ice regime/land on the last day of the tie-series for each season. Labels were assigned based on the following three factors: 1) backscatter values, 2) value of the projected incidence angle of the SAR pulse (Claude Duguay's research team has done previous work on the relation of backscatter values and incidence angle values for the two ice types; Duguay and Wang, 2019), and 3) location of the pixel within the scene (inside or outside a lake; at the lake edges or lake center; in a larger or smaller lake).

- The amount of full and partial time stamps for each year is included in Table 1 (please, see below). Each year of data was further interpolated to 161 time stamps which corresponds to a daily frequency from October 4 to March 14. There is a total of 555 scenes and including dates for each year of data would be two lengthy. However, the full table can be found in the following thesis of the lead author: <https://uwspace.uwaterloo.ca/handle/10012/17414>.

***Please, refer to Table 1 in the supplement**

Duguay, C.R. and J. Wang, 2019. Arctic-wide ground-fast lake ice mapping with Sentinel-1. ESA Living Planet Symposium, Milan, Italy, 13-17 May.

2. The sentences that were used to describe the implementation of the deep learning mapping approach is not enough. This cannot be used to judge whether the deep learning model was implemented correctly. How to determinate the model parameter values? Did you compare the model performances when using different parameter values? What is the loss curve of your model? More details are needed here to evaluate your model.

Thank you for a valuable comment. This work closely follows the TempCNN approach presented for land cover mapping using optical imagery by Pelletier et al., 2019. In terms of model parameters, the number of convolutional units for each convolutional layer was selected to be 64 based on temporal cross validation using 4, 8, 16, 32, 64, and 128 units (Please, refer to Figure 5). Batch size was adjusted to achieve best results. In addition, the number of training epochs is controlled by an early stopping technique with a patience of 10 epochs and a validation set comprising 5% of the training set.

The training and validation loss and accuracy for the final model used for creating lake ice maps is shown in the figure below.

***Please, refer to the supplement for all figures and tables.**

Parameter selection is discussed in the subsection 3.3.3. TempCNN architecture Selection.

"Out of 18 years of data, one year was reserved for final testing, while the remaining 17 years were used to identify the best model structure utilizing a cross-validation procedure. Cross-validation allowed to determine the optimal number of convolutional units (4, 8, 16, 32, 64, or 128) while keeping other elements of the structure constant (Fig. 4). TempCNN with each number of units was trained on 16 years of data and tested on one year of data 17 times each time leaving a different year 235 out for testing. The overall accuracy summary of each of the six sets of experiments is shown in a box-plot form in Fig. 5. Based on the value of the median, interquartile range (IQR), length of the whiskers, and number of outliers, architecture with 64 convolutional units was identified as the best and was subsequently used for creating lake ice maps."

The rest of the parameters were kept the same as in the Pelletier et al., 2019 study, from which the TempCNN architecture was adopted. Further parameter testing will be done in future work.

Pelletier, C., Webb, G. I., and Petitjean, F.: Temporal convolutional neural

network for the classification of satellite image time series, Remote Sens., 11, 1–25, 2019

3. Several types of SAR images were used in your paper, but the effect of their differences (e.g., spatial resolution) was not considered in your model.

Thank you for a valuable comment. The authors believe that being trained on data of different spatial resolutions, polarizations (HH and VV), and temporal resolution the TempCNN is capable of recognizing time-series from the three C-band platforms: ERS1/2, Sentinel 1, RADARSAT 1.

4. The author still analyzed the lake ice dynamics. However, as you pointed out in the manuscript, the lake area is also changing, how to consider the effect of this on the lake ice dynamics analysis.

Thank you for a valuable comment. Changing lake area definitely affects lake ice dynamics analysis. This is why in this work we strived to present an approach that does not need a new lake mask for each year and can capture the fluctuating lake boundaries. The reason behind using a lake mask for extracting lake ice fractions was to focus on the lake surface, rather than including all of the wetland. This will allow for future comparison to trends observed in other geographic areas of thermokarst lakes. Ideally a different lake mask had to be used for each year of data. However, due to the scope and timeframe of the project it was not feasible to create a different lake mask for each year of data. In order to establish a common baseline, the authors have decided to use a lake mask from fall of the most recent season in the dataset 2020/2021 (indication of the current status), while acknowledging that this definitely does lead to underestimation of the lake ice surface for years prior to major catastrophic drainages.

5. Most of the figure citations are short and simple, more details should be listed as in that case the readers can easily understand the figures.

Thank you for a valuable comment. We have revised the figure captions as follows:

"Figure 1. Old Crow Flats, Yukon, Canada. The background image is an RGB Landsat 8 of May 31, 2020, downloaded from USGS Earth Explorer (link: <https://earthexplorer.usgs.gov/>, accessed: July 4, 2021). Most lakes are still ice covered at this time of the year and appear white, open water surface of the river and smaller lakes, as well as some of the ice fringes appear black, tundra has a brownish shade, while areas of boreal forest appear dark green."

"Figure 2. Comparison of the three classes by sensor: (a) Sentinel-1; (b) RADARSAT-1; (c) ERS-1/2. Each class is represented by a mean and a standard deviation of a sample of 100 randomly selected pixels per sensor. Means and standard deviations are identified by solid and dashed lines, respectively: pink – floating ice; dark blue – bedfast ice; green – land."

"Figure 3. The two graphs illustrate application of 1-dimensional (1D) filters to time-series that can be used by convolutional layers of a TempCNN for extraction of temporal features. The red line represents original time-series, while the blue line denotes the filtered time-series: (a) a curve that resembles floating ice transformed by a gradient filter – the dashed red line indicates the origin of the filtered time series, where the value of the original series is increasing the filtered series has positive values, and where the value of the original series is decreasing the filtered series has negative values; (b) a curve resembling floating ice transformed by a low-pass filter."

"Figure 5. The graph illustrates temporal cross validation results using box-plots. Each box-plot contains 17 overall accuracy values and if read from left to right each box-plot corresponds to 4, 8, 16, 32, 64, and 128 convolutional units in each convolutional layer. Red highlights the best architecture with 64 convolutional units."

"Figure 6. A sample of TempCNN classification output for OCF, Yukon, Canada: (a) Sentinel-1, March 2021; (b) RADARSAT-1, March 2004; (c) ERS, March 1993. Dark blue, light blue, and grey represent bedfast ice, floating ice, and land, respectively."

"Table 2. TempCNN overall classification accuracy for 15 experiments designed to test sensitivity of the network to removing certain years of data from the training set. Runs 1-5 correspond to the 20/80% split of the entire dataset, runs 6-10 were performed by training the network on 15 years of data and testing it on 3 each from a different sensor, runs 10-15 were carried out by training the network on 17 years of data and testing it on 1 year of data that was originally reserved and was not part of the cross validation procedure for determining the best architecture. Subsequently, mean accuracy for each set of 5 runs was calculated and finally mean of the three means is shown in the last row of the table."

"Figure 7. Matching TempCNN output with lake depths collected in 2000. Horizontal lines indicate CLIMo ice thickness predictions for taiga (0.72 m) and tundra (1.21 m) environments. RADARSAT-1 SAR 1999/2000 time-series were used for ice regime classification. The colour of points corresponds to the label assigned to each location by the TempCNN: dark blue – bedfast ice, pink – floating ice, green – land; while the shape corresponds to the surrounding vegetation: circle – tundra, triangle – taiga; square – mixed assigned based on the OCF vegetation map created by Turner et al., 2014."

"Figure 8. CLIMo simulated ice thickness for OCF: (a) simulation of taiga environment; (b) simulation of tundra environment. Dark blue represents ice thickness and grey stands for snow depth. The grey and dark blue points mark the condition on March 13 (last day of time-series) for each year."

"Table 3. Field data collection in the OCF in April 2009. For each of the ten locations, UTM coordinates, ice thickness, lake depth, and ice regime collected in the field are matched with TempCNN (bedfast, floating, land) and vegetation type (tundra or taiga)."

"Table 4. Field data collected in April 2021 on a small lake located beside the drained basin of Zelma Lake. For each of the four locations, UTM coordinates, ice regime, ice thickness, snow depth, and sediment temperature are matched with the TempCNN output."

"Figure 9. Husky Lake TempCNN predicted ice regime (1993-2021). Dark blue represents bedfast ice, light blue – floating ice, grey – land. Ice regime fluctuates between floating and bedfast depending on snow conditions, water level, and air temperature."

"Figure 12. Netro Lake TempCNN predicted ice regime (1993-2021). Dark blue represents bedfast ice, light blue – floating ice, grey – land. Between 1993 and 2021 ice regime has transitioned from mostly bedfast to mostly floating."

"Figure 13. Zelma Lake TempCNN predicted ice regime (1993-2021). Dark blue

represents bedfast ice, light blue – floating ice, grey – land. You can notice a significant reduction in water surface area and a transition to mostly bedfast ice regime following the 2007 catastrophic drainage event.”

Specific comments:

Line 12 “less lakes are expected to develop bedfast ice”, Line 504 “lake ice fraction analysis from the available 18 years of data suggests that bedfast ice fraction has increased by 11% despite a warming climate”. Is there a contradiction between these two sentences?

Thank you for a valuable comment. Line 12 reflects an expectation based on studies carried out in other thermokarst regions, while line 504 discusses the results observed in through this study, so there is no contradiction.

The climate of OCF is continental with cold winters (mean January temperature of -31.1°C) and warm summers (mean July temperature of 14.6°C) (Lantz and Turner, 2015). Continental climate and glaciolacustrine deposits make this area unique and different from other thermokarst lake zones. Although the area is underlain by continuous ice-rich permafrost, unlike other regions, such as Mackenzie River Delta, even shallow bedfast-ice lakes display mean lake-bed temperatures above zero due to high summer air temperatures (Roy-Léveillé, 2014). Lakes surrounded by boreal forest tend to accumulate more snow, and therefore, develop thinner ice, while tundra lakes have a thinner, denser snow which results in ice grounding (Roy-Léveillé, 2014, Duguay et al., 2003). In light of this region's unique climate and sediment type, active thermokarst processes, and climate-driven vegetation changes, such as shrubification of the tundra area (Wang et al., 2020), the analysis carried out in this study make a valuable addition to the existing body of knowledge.

Lantz, T. C. and Turner, K. W.: Changes in lake area in response to thermokarst processes and climate in Old Crow Flats, Yukon, J. Geophys. Res. Biogeosci., 120, 513–524, 2015.

Roy-Léveillé, P.: Permafrost and thermokarst lake dynamics in the Old Crow Flats, northern Yukon, Canada, Ph.D. thesis, Carleton University, Ottawa, Canada, 2014.

Duguay, C. R., Flato, G. M., Jeffries, M. O., Ménard, P., Morris, K., and Rouse, W. R.: Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. Hydrol. Process., 17, 3465–3483, 2003.

Wang, J. A., Sulla-Menashe, D., Woodcock, C. E., Sonnentag, O., Keeling, R. F., and Friedl, M. A.: Extensive land cover change across Arctic Boreal Northwestern North America from disturbance and climate forcing. Global Change Biology, 26, 807-822, 2020.

Line 20-21, do you have any evidences to support this point?

Thank you for a valuable comment. It has been shown by other researchers (Lantz and Turner, 2015) that drainage events are on the rise in the Old Crow Flats. Effect of fluctuations in water levels and winter snowfall on the ice thickness is discussed in, for instance (Surdu et al., 2014, Duguay et al., 2003). The discussion that the referee is referencing here suggests that based on the ice maps produced in this study, ice regime of the Old Crow Flats is influenced by the above three phenomena. For instance, partial lake drainages lead to a transition from mostly floating to mostly bedfast (e.g., Zelma Lake, Lake #1 and # 2 shown in Figure 10). Moreover, years with less snow depth (e.g., 2021) exhibit more wide-spread bedfast ice (Please, refer to Table 5).

Lantz, T. C. and Turner, K. W.: Changes in lake area in response to thermokarst processes and climate in Old Crow Flats, Yukon, J. Geophys. Res. Biogeosci., 120, 513–524, 2015.

Surdu, C. M., Duguay, C. R., Brown, L. C., and Fernández Prieto, D.: 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, 167–180, 2014.

Duguay, C. R., Flato, G. M., Jeffries, M. O., Ménard, P., Morris, K., and Rouse, W. R.: Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrol. Process.*, 17, 3465–3483, 2003.

Line 22-23, more bedfast ice?

Thank you for a valuable comment. Catastrophic drainage events lead to loss of water surface area and shallowing of the remaining lake portion. As such, lake drainage is often accompanied by a transition from a mostly floating lake to a mostly bedfast lake.

Line 31-32, here you said that a subsequent decrease in the number of bedfast ice lakes has been noted by many researchers. However, in Line 504 "lake ice fraction analysis from the available 18 years of data suggests that bedfast ice fraction has increased by 11% despite a warming climate".

Thank you for a valuable comment. The decrease in the number of bedfast ice lakes has been noted in other thermokarst regions (e.g., Arctic Alaska). However, to the best of the authors' knowledge, no ice regime trend analysis has been performed for the Old Crow Flats. Based on the results of this study an increase in the bedfast ice fraction is observed.

Line 40, as you said there are many extensive studies, you should compare your method to theirs.

Thank you for a valuable comment. The goal of this manuscript was not to compare methods, but rather present a new method not previously used for lake ice mapping describing its benefits for the Old Crow Flats. The method presented in our paper could, however, be compared to other previously published approaches in a follow-up study either in the same or other study area(s). Although, this method is not necessarily better than other methods it does present advantages, namely 1) it takes advantage of the temporal evolution of ice backscatter over the entire lake ice season and is not reliant on a single value, as is the case for other methods; as new SAR platforms are being launched with higher quality, spatial resolution, and denser temporal coverage it is only reasonable to look for ways to make use of the wealth of available data; 2) wetland landscapes are dynamic in nature and the presented TempCNN method does not require a lake mask, which is necessary for all the previously presented methods; 3) a well-trained TempCNN can be applied to various SAR sensors, both HH and VV polarizations, and different spatial resolutions. The accuracy of the method was assessed using a training set and the overall accuracy values fall in the range reported by other ice regime mapping solutions.

Line 45, 'only THREE lakes shifted from bedfast to a floating ice regime', why?

Thank you for a valuable comment. This line discusses the work presented in Arp et al., 2012. This study classifies lakes as either bedfast, floating, or intermittent (lakes with bedfast ice in one or more, but not all years) in the period between 2003 and 2011. Then the results were compared to 1980. The study reports that on the Outer Arctic Coastal Plain of northern Alaska, only three lakes bedfast in 1980 fully transitioned to a floating ice regime in 2003-2011 period, and two experience an opposite transition likely due to drainage events. Please, visit reference below for more information.

Arp, C. D., Jones, B. M., Lu, Z., and Whitman, M. S.: Shifting balance of thermokarst lake ice regimes across the Arctic Coastal Plain of northern Alaska, *Geophys. Res. Lett.*, 39, 1–5, 2012.

Line 49, 'one of the study areas', it should be more specific here.

Thank you for a valuable comment. The authors are referring to the Fish Creek region on the Inner Arctic Coastal Plain of northern Alaska. The sentence in the manuscript has a typo and has been modified as follows:

"Nonetheless, one of the study areas, namely Fish Creek region on the Inner Arctic Coastal Plain of northern Alaska, exhibited strong trends towards floating ice regimes."

Line 53, why did you choose the Old Crow Flats as your study area, more reasons are needed.

Thank you for a valuable comment. Old Crow Flats is a wetland of international significance underlain by continuous permafrost. This landscape contains over 2500 thermokarst lakes and changes in their ice regime have the potential to significantly influence the underlying permafrost and subsequently the climate of the area. However, as has been mentioned in the "Study Area" section, ice regimes of Old Crow Flats remain largely unexplored. While ice regimes of other thermokarst lake areas, such as northern Alaska, USA, Lena River Delta, Russia, and Hudson Bay, Canada, got significant attention from the research community, ice regimes of the Old Crow Flats require investigating.

Line 61-61, there is a distinct backscatter patterns for floating and bedfast ice, thresholding method may be enough for the lake ice classification.

Thank you for a valuable comment. Thresholding approach is indeed very useful and definitely wins over the proposed approach due to its simplicity. However, the remaining challenges of the thresholding approach are as follows: 1) need for a lake mask – which is likely different for each year of data; 2) misclassification of floating ice as bedfast ice in the deeper sections of lakes; 3) reliance on a single backscatter value (one date/one scene) that could be affected by speckle noise. In this work we have explored a temporal deep learning approach which allows to classify SAR images into three classes, eliminating the need for a lake mask. In addition, TempCNN allows to minimize misclassification in deeper portions of lakes by including such areas into the training set and allowing the network to become familiar with them. Finally, the proposed method classifies each pixel in the image based on the evolution of the backscatter throughout the ice season, and therefore is less affected by the salt and pepper effect of the SAR imagery.

Line 70, a comparison to threshold-based classification is needed to prove that your method is necessary.

Thank you for a valuable comment. The goal of this manuscript was not to compare methods, but rather present a new method not previously used for lake ice mapping describing its benefits for the Old Crow Flats. The method presented in our paper could, however, be compared to other previously published approaches in a follow-up study either in the same or other study area(s). Although, this method is not necessarily better than other methods it does present advantages, namely 1) it takes advantage of the temporal evolution of ice backscatter over the entire lake ice season and is not reliant on a single value, as is the case for other methods; as new SAR platforms are being launched with higher quality, spatial resolution, and denser temporal coverage it is only reasonable to look for ways to make use of the wealth of available data; 2) wetland landscapes are dynamic in nature and the presented TempCNN method does not require a lake mask, which is necessary for all the previously presented methods; 3) a well-trained TempCNN can be applied to various SAR sensors, both HH and VV polarizations, and different spatial resolutions. The accuracy of the method was assessed using a training set and the overall accuracy values fall in the range reported by other ice regime mapping

solutions.

Line 78-79, this accuracy is comparable to your method. So you need to clearly point out the advantage of your method.

Thank you for a valuable comment. The benefit of the proposed method consists in 1) no need for a lake mask, unlike all the other proposed methods; 2) applicability of the same model to different sensor platforms (different incidence angles and both HH and HV polarization); 3) reliance on a seasonal evolution rather than a single value, which could be easily affected by speckle noise.

Line 85, what are the remaining challenges?

Thank you for a valuable comment. The remaining challenges are as follows:

- 1. Overestimation of bedfast ice in the middle of deeper lakes, as the backscatter signature of floating ice tends to be lower in such areas.**
- 2. Misclassifications due to the presence of wet snow.**
- 3. Need for an annual lake mask as all the existing methods are able to classify only lake ice.**
- 4. Need for a new threshold value for each sensor and each scene.**

Line 88-89, does your method still work well for these lakes with different phenomena or ice properties?

Thank you for a valuable comment. Further investigation is needed to determine whether TempCNN can tackle presence of wet snow and water salinity. In this work, the last day of the time-series was selected in mid-March to avoid any occurrences of early melt. However, it is our expectation that presence of wet snow or pools of standing water would still present a challenge for TempCNN as it does for other methods. Most likely isolated instances of wet snow prior to the last day of the time-series would be classified correctly. Nonetheless, more extensive surface melt at the end of season would likely look too similar to ice grounding due to low backscatter values of both wet snow and pools of water on top of the ice. Water salinity on the other hand, can likely be tackled by TempCNN as long as brackish lakes are part of the training dataset, which is not the case for the Old Crow Flats and most thermokarst lake regions.

Line 104, did the data for different sensors/with different resolutions/acquired in different years affect the method performance? How to consider these?

Thank you for a valuable comment. A major component of this work was to create lake ice maps for the Old Crow Flats for the longest timeframe possible. The other important consideration was to have a method that would take advantage of the temporal dimension (in the light of the more recent SAR platforms supplying a wealth of regular high-quality SAR coverage) and that would allow to create lake ice maps using data from 3 different available platforms with varying spatial and temporal resolutions, and two different polarizations (HH and VV). The authors believe that the quality of the output maps matches the quality of the input SAR data.

Line 116, so your method is mainly used for shallow lakes?

Thank you for a valuable comment. That is correct. This study was mainly concerned with shallow thermokarst lakes, as landscapes underlain by permafrost are more sensitive to changes in lake ice regime. However, the authors believe that the proposed approach could be used for lakes of any size. As has been noted in the manuscript, deeper lakes often display darker (low backscatter) patches of floating ice in their centers. As such, more extensive training including such areas would be required to ensure that the TempCNN is familiar with the temporal patterns specific to the deeper lakes.

Line 125, dose the decreasing lake area affect the analysis of bedfast ice dynamics?
Thank you for a valuable comment. Changing lake area definitely affects lake ice dynamics analysis. This is why in this work we strived to present an approach that does not need a new lake mask for each year and can capture the fluctuating lake boundaries. The reason behind using a lake mask for extracting lake ice fractions was to focus on the lake surface, rather than including all of the wetland. This will allow for future comparison to trends observed in other geographic areas of thermokarst lakes. Ideally a different lake mask had to be used for each year of data. However, due to the scope and timeframe of the project it was not feasible to create a different lake mask for each year of data. In order to establish a common baseline, the authors have decided to use a lake mask from fall of the most recent season in the dataset 2020/2021 (indication of the current status), while acknowledging that this definitely does lead to underestimation of the lake ice surface for years prior to major catastrophic drainages.

Lien 128-129, 'Despite the overall trend of decreasing water surface area, most lakes are increasing in surface area'??

Thank you for a valuable comment. The authors are referring to the fact that although the Old Crow Flats landscape is dominated by catastrophic drainage events leading to an overall decrease in water surface area (1951 – 2007; further research is needed for more recent years; Lantz and Turner, 2015), lakes are subject to thermo-mechanical erosion caused by wave action (more pronounced in tundra areas of the Old Crow Flats where shores are not protected by vegetation cover) leading to rapid shoreline recession over time (Roy-Léveillé and Burn, 2015).

Roy-Léveillé, P. and Burn, C. R.: Geometry of oriented lakes in Old Crow Flats, northern Yukon, in: Proceedings, 68th Canadian Geotechnical Conference and 7th Canadian Permafrost Conference, 21–23, Quebec City, QC, 2015.

Figure 1, longitude and latitude are suggested to use, more detailed descriptions about this figure are needed.

Thank you for a valuable comment. The Figure 1 caption has been updated as follows:

"Figure 1. Old Crow Flats, Yukon, Canada. The background image is an RGB Landsat 8 of May 31, 2020, downloaded from USGS Earth Explorer (link: <https://earthexplorer.usgs.gov/>, accessed: July 4, 2021). Most lakes are still ice covered at this time of the year and appear white, open water surface of the river and smaller lakes, as well as some of the ice fringes appear black, tundra has a brownish shade, while areas of boreal forest appear dark green."

Line 143, how to consider this effect on your model?

Thank you for a valuable comment. The time-series was purposefully selected to end in mid-March to ensure that no melt was occurring. As such, the lake ice maps are not affected by wet snow. However, presence of wet snow or pools of standing water remain an open problem for the bedfast ice mapping as both lead to a darker signature (signal absorption and reflection, respectively).

Line 157-158, 'Adjusting the filter size allowed to account for the pixel size differences.'. According to the resolutions listed in Table1, this may be not true.

Thank you for a valuable comment. We have selected the speckle filter size to ensure that it covers approximately the same area for the three different SAR platforms. Please see the table below:

Platform	Pixel size (m)	Filter size (pixels)	Area covered by each
----------	----------------	----------------------	----------------------

			filter (m2)
Sentinel 1	30x30	7x7	44,100
ERS1/2	12.5x12.5	16x16	40,000
RADARSAT 1	50x50	4x4	40,000

Line 161, again, how to consider the effect of data inconsistency on your classification results, which may further affect the ice dynamics analysis.

Thank you for a valuable comment. This paper explores taking a temporal approach to lake ice classification from SAR. A temporal approach definitely requires a regular temporal coverage of the ice season to produce the best results possible. The aim of this work is to demonstrate the potential of the TempCNN. Going forward, higher quality, spatial, and temporal resolution SAR imagery is available through Sentinel-1 and the new RADARSAT Constellation Mission (RCM) which will allow for training a TempCNN and applying the trained model for ongoing monitoring of lake ice dynamics in the future.

Line 167, what are the spatial distributions of these labeled pixels?

Thank you for a valuable comment. Land and floating ice labeled pixels were spread out over each scene as evenly as possible. Labeled pixels with bedfast ice, on the other hand, being less prevalent were less spread out. However, we made sure to include areas in both the northern and the southern parts of the Old Crow Flats, when possible, to include both tundra and boreal forest environments.

Line 169, what is the detailed process of the visual assessment, how to deal with the mixed pixels?

Thank you for a valuable comment. Pixels at the border between floating ice and bedfast ice, as well as where land meets lake ice were avoided during the labeling process. This was done to provide TempCNN with the clearest and most representative time-series patterns to learn to distinguish between the classes. Nonetheless, the trained TempCNN outputs not only a classified map, but also probability maps for each of the three classes (land, floating ice, bedfast ice), where the value of each pixel corresponds to the probability of this pixel to belong to a specific class. As such, in the areas of mixed pixels probability values for multiple classes are high, with the final class being assigned based on the class with the highest probability. Probability maps have not been used in this work's analysis. However, their utility and potential applications will be thoroughly explored in future research.

Line 175-176, this process should be more quantified?

Thank you for a valuable comment. The authors have made utmost efforts to select representative areas of bedfast and floating lake ice, as well as land for each year of available data. Different years contained scenes with different incidence angles. In terms of incidence angles, we were guided by previous work done by Claude Duguay's research team (Duguay and Wang, 2019). Please take a look at the graph below.

***Please, refer to the supplement for all figures and tables**

Duguay, C.R. and J. Wang, 2019. Arctic-wide ground-fast lake ice mapping with Sentinel-1. ESA Living Planet Symposium, Milan, Italy, 13-17 May.

Line 181, why using early fall scenes to identify land areas?

Thank you for a valuable comment. Labels (bedfast ice, floating ice, land) were assigned to the time-series based on the condition of the ice on the last day of the time-series (mid-March). However, both land and floating ice have similar

SAR signature (high backscatter) at this time of the season. Hence, to avoid confusion between the two classes, during the labeling process the authors made sure to refer to the early October scenes where land and lake surface are more easily distinguishable. Higher contrast between the ice-covered lake surface and the surrounding land in early fall is explained by the fact that thin ice covering the lakes creates a mirror-like surface, leading to specular reflection of the SAR signal away from the sensor, and consequently lower backscatter (lakes appear darker). Later in the season as the floating ice thickens its backscatter increases due to high dielectric discontinuity and roughness of the ice-water interface. Please, see an example of an early fall scene used for identifying lake outlines during the labeling process (October 14, 2017, Sentinel 1) below. The greyscale corresponds to sigma nought(db) values. Lake surfaces appear black, while the surrounding land is grey.

***Please, refer to the supplement for all figures and tables**

Line 183-184, how many data were obtained from the interpolation?

Thank you for a valuable comment. The temporal deep learning model trained and used in this study works with time-series of a specific length. It takes a backscatter time-series as input and assigns it a specific class (bedfast ice, floating ice, or land) which corresponds to the state of the pixel on the last day of the time-series. Each time-series represents the backscatter value of a single pixel traced through time from October 4 to March 13. However, temporal resolution of SAR image stacks (# of scenes available throughout the ice season) varied significantly between years (please, refer to Table 1). As such, prior to inputting a SAR image stack into the TempCNN for classification, it had to be interpolated. The model was trained to work with time-series consisting of 161 time stamps which corresponds to a daily frequency from October 4 to March 13 – excluding February 29 for leap years. Imagine that each SAR image stack is a collection of time-series where each pixel is represented by a time-series of backscatter values starting from its backscatter value on October 4 and ending with a backscatter value on March 13 (Please, refer to portion of Figure 4 presented on the right for visual illustration). To ensure that each of the backscatter time-series for each year of data had the same length (161 values), linear interpolation was applied. Although the lake ice lifecycle is non-linear, previous studies have shown that more complex interpolation methods have little influence on classification accuracy (Pelletier et al., 2019, Valero et al., 2016). Linear interpolation was performed utilizing python programming language and the tools of pandas module. Interpolation not only filled the temporal gaps, but also replaced any missing or Not a Number (NaN) values, especially common for ERS1/2 and scene fringes. Interpolation was performed individually on every time-series (backscatter value of pixel traced through time). As a result, we obtained SAR image stacks consisting of 161 full coverage scenes, which were subsequently input into the TempCNN to perform classification.

The manuscript briefly explains time-series interpolation in Subsection 3.2 Annotated dataset creation:

"Resampling to a daily frequency and linear interpolation were applied to compensate for the temporal irregularity of the data gearing it for the deep learning classification (Pelletier et al., 2019; Valero et al., 2016). The final labeled time-series consisted of 161 time steps (i.e., one time step per day) covering the time period 185 between October 4 and March 13."

As per the referee's request, Subsection 3.4 Creation of ice regime maps using TempCNN was extended to include more details on interpolation of SAR stacks as

follows:

"In order to transform SAR image stacks for each of the 18 years of data into lake ice regime maps using the trained TempCNN each stack had to be interpolated. Interpolation allowed to compensate for temporal resolution variability between different years such that each year's stack consisted of 161 scenes corresponding to a daily frequency from October 4 to March 13. Pixel-based linear interpolation was performed utilizing python programming language and the tools of pandas module. Although the lake ice lifecycle is non-linear, previous studies have shown that more complex interpolation methods have little influence on classification accuracy (Pelletier et al., 2019; Valero et al., 2016). Once the SAR stacks for 18 years were interpolated and each consisted of 161 scenes, the trained TempCNN model was used to create ice regime classification maps consisting of three classes: floating ice, bedfast ice, and land."

Pelletier, C., Webb, G. I., and Petitjean, F.: Temporal convolutional neural network for the classification of satellite image time series, Remote Sens., 11, 1–25, 2019.

Valero, S., Pelletier, C., and Bertolino, M.: Patch-based reconstruction of high resolution satellite image time series with missing values using spatial, spectral and temporal similarities, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2308–2311, 2016.

***Please, refer to the supplement for all figures and tables**

Line 189-190, how to consider this effect on your method?

Thank you for a valuable comment. The pattern described in the lines mentioned by the referee forms the basis of the proposed method. Training the deep learning model to recognize these temporal patterns allows it to distinguish between floating and bedfast ice and create lake ice maps.

Figure 2, the colors of two types of ice are difficult to classify.

Thank you for a valuable comment. The colours have been updated, please, see below:

***Please, refer to the supplement for all figures and tables**

Line 218-224, how to determinate the optimal parameter values for your model?

Thank you for a valuable comment. Parameter selection is discussed in a subsequent section. Please, refer to the subsection 3.3.3. TempCNN architecture Selection:

"Out of 18 years of data, one year was reserved for final testing, while the remaining 17 years were used to identify the best model structure utilizing a cross-validation procedure. Cross-validation allowed to determine the optimal number of convolutional units (4, 8, 16, 32, 64, or 128) while keeping other elements of the structure constant (Fig. 4). TempCNN with each number of units was trained on 16 years of data and tested on one year of data 17 times each time leaving a different year 235 out for testing. The overall accuracy summary of each of the six sets of experiments is shown in a box-plot form in Fig. 5. Based on the value of the median, interquartile range (IQR), length of the whiskers, and number of outliers, architecture with 64 convolutional units was identified as the best and was subsequently used for creating lake ice maps."

The rest of the parameters were kept the same as in the Pelletier et al., 2019 study, from which the TempCNN architecture was adopted. Further parameter

testing will be done in future work.

Pelletier, C., Webb, G. I., and Petitjean, F.: Temporal convolutional neural network for the classification of satellite image time series, Remote Sens., 11, 1–25, 2019.

Line 231, this subtitle seems to be subordinate to Section 3.3.2 (TempCNN architecture).
Thank you for a valuable comment. The authors prefer to leave this section as a separate subsection to avoid having the fourth level of subsections.

Line 242, why THREE SETS were used here?

Thank you for a valuable comment. The 15 experiments described in the subsection 3.3.4 TempCNN training and testing aim to test how sensitive the model is to inclusion or exclusion of certain years of data. We have used three different ways to split the data into training and testing sets: 1) a random split of all data points into 80% for training and 20% for testing; 2) three complete years (each from a different SAR platform) were left out for testing, 15 years of data were used for training; 3) training was done using 17 years of data and testing using data from the 2020/2021 ice season which was originally reserved for validation and not used when determining optimal neural network architecture.

Line 250, which dataset was used to train the TempCNN model?

Thank you for a valuable comment. Please, refer to Line 246: "The model trained on points from all years of the SAR dataset was selected for the ice regime mapping."

Line 254, what elevation data was used here?

Thank you for a valuable comment. The radiometric terrain corrected (RTC) Sentinel-1 imagery downloaded from Alaska Satellite Facility is supplemented by a digital elevation model (DEM) granule which was used in this study. Cell values indicate elevation in meters and have been resampled to a pixel spacing of 30 m by Alaska Satellite Facility. The DEM granule comes from the National Elevation Dataset at 2 arc seconds resolution (NED2) which is produced and distributed by USGS.

Line 257, what is the spatial distribution of these field data.

Thank you for a valuable comment. Figure below displays the spatial distribution of the field data. There are 51 measurement points in total.

***Please, refer to the supplement for all figures and tables**

Line 262, does the Canadian Lake Ice Model have the enough resolution for method evaluation? Because the resolution of ERA5 is quite coarse.

Thank you for a valuable comment. CLIMo does provide enough resolution. CLIMo can use data from either nearby meteorological stations (not available in this study) or atmospheric reanalysis data such as ERA5 to provide forcing atmospheric field (air temperature, wind speed, relative humidity, cloud cover and snowfall or depth).

Line 268, why 2 m? in Line 116 you mentioned an averaged depth of 1.5 m.

Thank you for a valuable comment. The change in depth could impact the freeze-up date by at most 1 day. However, the ice thickness simulated by CLIMo for the end of the season and used in this study will not be impacted by such a small change in depth (no difference in the break-up date).

Line 272-274, figure or table are suggested to use here to help readers understand your opinion.

Thank you for a valuable comment. Figure 7 and 8, as well as Tables 3 and 4 in the Results and Discussion section illustrate the analysis referred to in Lines 272-274.

Line 284, the lake area is still changing, a single lake mask may include some uncertainties.

Thank you for a valuable comment. The referee is absolutely right, using a single lake mask is not ideal as over the years the lake area is changing, and catastrophic lake drainages are occurring. However, due to the scope and timeframe of the project it was not feasible to create a different lake mask for each year of data. The use of lake mask for extracting bedfast and floating lake ice fractions are explained by the fact that the Old Crow Flats is a dynamic wetland where areas of standing water outside of the lakes are also covered by ice in winter and get classified as such. In order to establish a common baseline, the authors have decided to use a mask to extract bedfast/floating lake ice fractions, while acknowledging that this definitely does lead to underestimation of the lake ice surface for years prior to major catastrophic drainages.

Line 292-293, what are the known lake drainage and refilling events?

Thank you for a valuable comment. Lantz and Turner, 2015 report that between 1951 and 2007 the Old Crow Flats experience a decline of ~6000 ha in total lake area (but gained 232 lakes). 38 large lakes drained. Examples include lake Zelma (catastrophically drained in 2007) and lake Netro (drained around 1972). In addition, results of our work suggest that Lakes referred to as Lake #1 and #2 appear to have lost some of their surface area in the recent years. Moreover, result of our study suggest that Lake Netro has been refilling.

Lantz, T. C. and Turner, K. W.: Changes in lake area in response to thermokarst processes and climate in Old Crow Flats, Yukon, J. Geophys. Res. Biogeosci., 120, 513–524, 2015.

Line 300, so did you apply any quality control strategies for the SAR images?

Thank you for a valuable comment. All the SAR imagery has been properly processed and speckle filtered. In addition, the interpolation process not only filled the temporal gaps but also replaced any missing or Not a Number (NaN) values, especially common for ERS1/2 and scene fringes.

Line 302, when S1 data are not accessible, what is the coverage of other data in the study area?

Thank you for a valuable comment. All the available C-band SAR platforms have been explored. ERS offers data between 1992-2011, RADARSAT 1 1996 – 2008, and Sentinel 1 2014 – present. All years of data with at least two scenes for each month of the ice season were included in the dataset. RADARSAT 2 (2007 – present) data is not freely accessible, while Envisat (2002 – 2012) did not have a reasonable coverage for any given year within the study period.

Line 316, 'the major features of the lifecycle can be lost', why?

Thank you for a valuable comment. TempCNN classifies SAR stacks based on the shape of the time-series it has learned during the training process. Generally, land in the Old Crow Flats has little variability during the ice season compared to the lake ice (please, take a look at the figure below). For Sentinel 1, which has regular temporal coverage a clear lifecycle or pattern emerges and is similar between different years of data. However, for ERS1/2 where temporal coverage is sparser and more variable for different years and interpolation is more extensive, pattern/lifecycle of the three classes becomes more "unique" for a specific year, making it harder to extract and learn a general pattern.

***Please, refer to the supplement for all figures and tables**

Figure 6, is ice area in 2021 larger than that in 2004?

The Old Crow Flats is a wetland environment and has areas of standing water outside of the lake boundaries which are classified as lake ice by TempCNN. As such, fluctuations of the landscape wetness from year to year result in some years appearing to have more lake ice than others, but could be shallow surface water rather than "lake surface".

Table 2, more detailed descriptions are need.

Thank you for a valuable comment. The description of Table 2 has been modified as follows:

"Table 2. TempCNN overall classification accuracy for 15 experiments designed to test sensitivity of the network to removing certain years of data from the training set. Runs 1-5 correspond to the 20/80% split of the entire dataset, runs 6-10 were performed by training the network on 15 years of data and testing it on 3 each from a different sensor, runs 10-15 were carried out by training the network on 17 years of data and testing it on 1 year of data that was originally reserved and was not part of the cross validation procedure for determining the best architecture. Subsequently, mean accuracy for each set of 5 runs was calculated and finally mean of the three means is shown in the last row of the table."

Line 321-323, some possible explanations are suggested to add here.

Thank you for a valuable comment. The manuscript has been updated as follows:

"The following pattern emerged when analyzing confusion matrices from all experiments: the 80/20% models had very minimal misclassifications equally distributed between the three classes; the models tested on three years had most significant misclassification of land as bedfast ice; and the models tested on one year of S1 had most significant misclassifications of floating ice as bedfast ice. The better performance of the 80/20% model is believed to be attributed to the high inter-year temporal variability of available scenes. As such when time-series from each year of data are present in both the training and the test set higher accuracy is achieved."

Line 328, why the dark spots are particularly pronounced for R1 2001, ERS 1993, 1994, are they related to sensor?

Thank you for a valuable comment. A further investigation is necessary in order to say with certainty what the nature of the dark spots is. As is mentioned in the manuscript "Some of the recent studies (e.g., Engram et al., 2020; Pointner and Bartsch, 2020) suggest that the dark spots pattern could be a result of local ice thinning or complete melt caused by methane ebullition." The authors believe that it is not related to the sensor, as the pattern appears clearly in most lakes, but not outside of them and the backscatter values could represent open water. Future research is definitely necessary to answer the referee's question and discover why the pattern appeared in specific years only.

Line 340-341, did you use the optical images in this study? Do you mean that your training dataset may contain erroneous labeling?

Thank you for a valuable comment. The optical imagery was not used for the analysis and is not part of the dataset. However, Landsat imagery was occasionally consulted throughout the labeling process. As to the "erroneous labeling", authors are referring to the fact that field observations of ice regimes in Old Crow Flats are very scarce and a more extensive field work would be beneficial to validate the ice regime patterns observed via SAR imagery analysis

only. Moreover, as any dataset created manually, human error cannot be completely eliminated. In order to avoid the confusion, the sentence has been modified as follows:

"As such, due to the lack of ground truthing, the labeled dataset used for model training was created from visual interpretation of SAR imagery and, therefore, could be subject to human error."

Line 381-392, how can the ice thickness data be used to evaluate the lake ice classification results?

Thank you for a valuable comment. Ice thickness measured in the field was accompanied by lake depth measurement (please, refer to Table 3) or a record of ice regime (floating/bedfast) (please, refer to Table 4). Where ice thickness is equal to the lake depth, lake ice regime is bedfast, whereas in areas where ice thickness is less than the lake depth, lake ice regime is floating. This way, using ice thickness and lake depth or a direct field record of lake ice regime can be compared to the classification output of the TempCNN.

Line 1.29, what about the ice whose thickness between 1.14 and 1.29 m?

Thank you for a valuable comment. The aim of the discussion that the reviewer is referring to is to draw general parallels between CLIMo predicted ice thickness for the tundra area and the ice regime classification output by TempCNN for the Husky Lake ($\sim 68.33^\circ$ N, 140.17° W). Below we are including a table which shows CLIMo simulated ice thickness for each of the years in the dataset along with the TempCNN ice regime class for the majority of the Husky Lake surface. Looking at the table below you can notice that for years where CLIMo predicted ice thickness is 1.08-1.14 m (thinner ice) the TempCNN classified ice regime is floating (highlighted in blue). For years where CLIMo simulated ice thickness is 1.29-1.40 m (thicker ice), TempCNN ice regime is bedfast (except for 1996) (highlighted in black), while the for the rest of the years with ice thickness 1.19 – 1.23 m the TempCNN ice regime fluctuates between floating and bedfast. It should be noted that CLIMo has a general uncertainty of ~ 0.05 - 0.06 m (Duguay and Lafleur, 2003).

***Please, refer to the supplement for all figures and tables**

Duguay, C. R. and Lafleur, P. M.: Determining depth and ice thickness of shallow sub-Arctic lakes using space-borne optical and SAR data, Int. J. Remote Sens., 24, 475–489, 2003.

Line 474-475, does the changing lake area affect the ice dynamics analysis?

Thank you for a valuable comment. Changing lake area definitely affects lake ice dynamics analysis. This is why in this work we strived to present an approach that does not need a new lake mask for each year and can capture the fluctuating lake boundaries. The reason behind using a lake mask for extracting lake ice fractions was to focus on the lake surface, rather than including all of the wetland. This will allow for future comparison to trends observed in other geographic areas of thermokarst lakes. Ideally a different lake mask had to be used for each year of data. However, due to the scope and timeframe of the project it was not feasible to create a different lake mask for each year of data. In order to establish a common baseline, the authors have decided to use a lake mask from fall of the most recent season in the dataset 2020/2021 (indication of the current status), while acknowledging that this definitely does lead to underestimation of the lake ice surface for years prior to major catastrophic drainages.

Line 482-483, how to reduce the effect from spatial resolution difference?

Thank you for a valuable comment. The authors are referring to the fact that in coarser spatial resolution imagery (e.g., R1) smaller bedfast ice areas are likely to be missed due to larger pixels and signal dilution by the surrounding land/floating ice. Nothing can be done about this; coarser spatial resolution will always lead to some information loss. However, going forward, higher spatial and temporal resolution C-band SAR (Sentinel-1, RADARSAT Constellation Mission (RCM)) and TempCNN method could be advantages.

Table 5, how many areas of floating ice transferred to bedfast ice? How many areas of bedfast ice \rightarrow floating ice, ice \rightarrow land, land \rightarrow ice? You can add this information into this table.

Thank you for a valuable comment. Unfortunately, we did not understand the question quite clearly. There could be many (thousands) areas that transitioned from floating to bedfast ice and vice versa. Based on reference to Table 5 we think that the referee likely meant change in square km, we will add this information to the revised manuscript.

Line 499-500, so the accuracy of lake ice maps in different periods are variable. Does this affect the final ice dynamics analysis?

Thank you for a valuable comment. To avoid confusion, the following sentence has been removed:

"However, the accuracy of the lake ice maps is heavily dependent on the temporal density and quality of the input data."

The accuracy assessment has been done for all the points in the dataset and not for different instruments separately. As such, we cannot definitively say if accuracy of one of the platforms is lower than the others. Overall accuracy for all the three platforms together is quite high and comparable to the results achieved by other methods. In lines 499-500 we were refereeing to the fact that ERS 1/2 had a much lower temporal coverage than the other two platforms, and a lot of its scenes did not cover the Old Crow Flats in their entirety. As such, we assume that accuracy of the ERS lake ice maps might be affected. Nonetheless, overall accuracy for the annotated dataset is high and comparable to other studies. Moreover, this study is the first to investigate a temporal approach to lake ice mapping and further investigation is definitely needed to determine the temporal sampling necessary to achieve optimal results.

Line 505-507, more evidences are needed here.

Thank you for a valuable comment. The manuscript has been modified as follows:

"Over the period between 1993 and 2021, lake ice fraction analysis from the available 18 years of data suggests that bedfast ice fraction has increased by 11% despite a warming climate (Fig. 10). This counter-intuitive result is linked to the increased frequency of catastrophic lake drainages affecting OCF (Lantz and Turner 2015), which result in drastic decreases in water depths and exposed portions of lake bottoms in several large lakes that were associated with a transition to bedfast ice (Fig. 13, 14). The partial refilling of drained basins, in some cases linked to the catastrophic drainage of adjacent lakes (Fig. 12), was associated with a transition to floating ice (Fig. 12). As permafrost may be sustained or aggrade under shallow water with a bedfast ice regime in OCF (Roy-Léveillé and Burn 2017), this transition is likely to have implications for permafrost sustainability and talik development beneath OCF lakes, with potential implications for methane ebullition and the regional carbon budget (Helsop et al., 2015)."

Please also note the supplement to this comment:

<https://egusphere.copernicus.org/preprints/2022/egusphere-2022-388/egusphere-2022-388-AC1-supplement.pdf>