General comments:
This paper focused on the lake ice mapping in the Old Crow Flats using a temporal deep learning approach from the C-band SAR time series. Lake ice maps labeled as floating ice, bedfast ice, or land Flats were created from 1993 to 2021. The created ice-map dataset could be a reference for future ice dynamics analysis. So, it is an important and interesting issue. In this paper, however, the lake ice dynamics analysis was too simple. This work can be published after some minor revisions. However, some methods, expressions and the lake ice dynamic analysis need to be clarified before considering publication. Thus, I want to recommend that the paper be published after a few (minor) modifications.

*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 tables and figures.*

Specific comments:

In the paper, the main purpose was to detect the bedfast ice and the floating ice. The two types ice was the top topic in this paper, so detail information on these is needed. In the introduction, the difference between bedfast ice and floating ice was not mentioned a lot. Please make a supplement to the bedfast ice and floating ice in the different effects on the lake and climate.

Thank you for a valuable comment. The introduction section has been modified as follows to include a more detailed description of the impact of bedfast and floating ice presence on the permafrost:

"Many shallow arctic lakes and ponds of thermokarst origin freeze to bed in the winter months, allowing lake-bottom temperatures to drop below 0°C and frost to penetrate the lake bottom sediment. Permafrost is sustained beneath the lake bottom where the freezing-degree-days at the ice-sediment interface are sufficient to counterbalance the thawing that takes place while lake-bottom temperatures are above 0°C (Roy-Léveillée and Burn, 2017). Where lake bottom conditions are too warm to sustain permafrost, for instance where ice does not reach the lake bottom or where the period of ice contact is brief, permafrost will degrade and a bulb of unfrozen ground or talik will develop and expand beneath the lake bottom. Such talik development contributes to positive feedbacks as it promotes lake deepening via subsidence of the lake bottom (Roy-Léveillée and Burn 2016), further reducing the occurrence of bottom-fast ice, and increases the ebullition of potent greenhouse gases such as methane from the thawing and decomposition of organic matter beneath the lake bottom (Arp et al., 2012; Engram et al., 2020). However, lake ice thinning and a subsequent decrease in the extent and duration of bedfast ice lakes has been noted by many researchers (Engram et al., 2018; Labrecque et al., 2009; Surdu et al., 2014). Hence, monitoring and quantifying thermokarst lake ice dynamics is critical for understanding changes in sub-lake permafrost stability and expected changes in methane ebullition patterns in thermokarst lowlands. Bedfast ice mapping, in particular, has a variety of other applications, including climate monitoring (Arp et al., 2012), permafrost studies (Arp et al., 2011), bathymetric mapping (Duguay and Lafleur, 2003; Kozlenko and Jeffries, 2000), overwintering fish habitat (Brown et al., 2010), and winter water withdrawal (Hirose et al., 2008; Jeffries et al., 1996)."

Roy-Léveillée, P. and Burn, C. R.: A modified landform development model for the topography of drained thermokarst lake basins in fine-grained sediments. Earth Surface Processes and Landforms, 41, 1504-1520, 2016. DOI: 10.1002/esp.3918

Could the TempCNN can recognize the lake from the land when there is no ice in the lake? In this paper, it is deemed the land and lake ice are seamlessly connected. Did ice fully cover lakes during the study period?

Thank you for a valuable comment. Generally speaking, land surface appears bright (high backscatter values) in SAR imagery due to roughness of its surface and vegetation volume scattering and open water surface is dark most of the time (unless it is roughened by wind) (Huang et al., 2018, Duguay and Lafleur, 2003). As such, many studies have successfully explored water-body segmentation using SAR mainly using machine learning and deep learning techniques (Guo et al., 2022). To the best of the authors knowledge, no studies have applied temporal deep learning to this task. However, applying TempCNN to extracting water bodies would probably be unnecessarily complex, due to the need to create an extensive labelled dataset.

On the other hand, mapping bedfast and floating ice in presence of open water (cracks in the ice or local melt caused by methane ebullition (Engram et al., 2020) or temporarily warming) is a challenge, as both open water and bedfast ice are characterized by low backscatter (Duguay and Lafleur, 2003). In this work, the time-series for each year of data tracked backscatter from early October to mid-March, and the final classification for each year was representative of the state of ice on the last day of the time-series (mid-March). Although, some open water was likely present in early October, mid-March end date was selected specifically to avoid open water presence (based on air temperature information).

Huang, W., DeVries, B., Huang, C., Lang, M. W., Jones, J. W., Creed, I. F., and Carroll, M. L.: Automated extraction of surface water extent from Sentinel-1 data, Remote Sens., 10, 797, 2018.

Duguay, C. R. and 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.

Guo Z, Wu L, Huang Y, Guo Z, Zhao J, Li N. Water-Body Segmentation for SAR Images: Past, Current, and Future. Remote Sensing. 2022; 14(7):1752. https://doi.org/10.3390/rs14071752

Engram, M., Anthony, K. M. W., Sachs, T., Kohnert, K., Serafimovich, A., Grosse, G., and Meyer, F.: Remote sensing northern lake methane ebullition, Nat. Clim. Chang., 10, 511–517, 2020.

Lake ice simulation is another key work in your paper. But the process of the simulation was not clear. It seems unreasonable to make the parameters unchanged as the land in the northern part and southern part are different.

Thank you for a valuable comment. The CLIMo simulation results were used to access the accuracy of lake ice maps. The CLIMo simulation was run for two scenarios: taiga and tundra to capture the fact that northern part of Old Crow Flats is characterized by polygonal tundra, while the southern part has subarctic boreal forest. Parameters were set as follows: taiga – snow depth 100%, snow density 175 kg/m3; tundra - snow depth 50%, snow density 300 kg/m3. Please refer to subsection 3.5 Accuracy assessment. The snow depth and density values used for the taiga vs tundra simulations are typical to those documented in other studies (Duguay et al., 2003; Sturm and Liston, 2003).

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.

Sturm, M. and Liston, G. E.: The snow cover on lakes of the Arctic Coastal Plain of Alaska, USA., J. Glaciol., 49, 370–380, 2003.

The figure captions are not so clear. Some lines and background lack details to make the figure hard to understand.

Thank you for a valuable comment. The figure captions have been revised and more detailed descriptions have been added to make figures more reader friendly as shown below.

"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."

Technical corrections:

Table 1: The dataset is not consistent because of the lack of data from 1996/1997 to 1998/1999. Please give detailed information about it.

Thank you for a valuable comment. In order to build a reliable time-series a more or less even coverage throughout the ice season is necessary. 18 years of data were chosen for this study as they offered a minimum of two scenes for each month throughout the ice season as is mentioned in subsection 3.1 SAR imagery. ERS-1/2 availability was quite limited and seasons 1996/1997 1997/1998, and 1998/1999 had either none or an insufficient number of scenes available.

Figure 1. please point out what the different color areas represent.

Thank you for a valuable comment. 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."

Figure 3. what's the dashed red line represent for?

Thank you for a valuable comment. Figure 3 caption has been updated as follows:

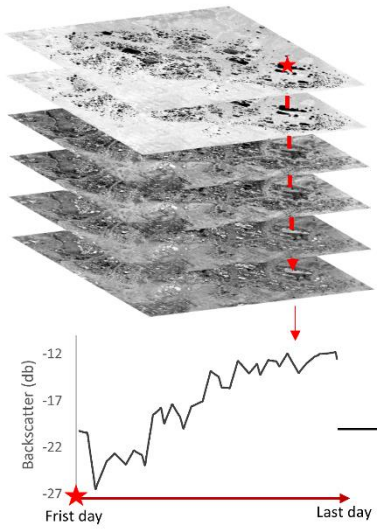
"Figure 3. The two graphs illustrate application of 1-dimensional (1D) filters to time-series that can be used by convolutional layers of a TempCNN for extraction of temporal features. The red line represents original time-series, while the blue line denotes the filtered time-series: (a) a curve that 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."

Line 140 "which cover the time period between 1992 to 2021."; line 14 "Canada over the 1993 to 2021 period"; Please unify the time of data for the paper: it's better to unify the time as 1992/1993 to 2020/2021.

Thank you for a valuable comment. The time period has been made consistent throughout the manuscript.

Line 250: please make an explanation about the interpolation of the SAR stacks.

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.

Line 268: "lake depth was specified as 2m", is it reasonable, as the average depth of lakes in OCF is 1.5m.

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 422:" Performing change detection between the first (1993) and the last years of the dataset (2021), reveals a transition of 51 km2 from bedfast to floating ice regime. However, 172 km2 of floating ice shifted to a bedfast state." How to detect the exchange between the bedfast ice and floating ice?

*Thank you for a valuable comment. The ice regime maps created by TempCNN for each year of the available data allow to detect the exchange between the two ice regime classes using the following formula: ("class on date 1" *10) + "class on date 2". The resulting GIS layer of data will be a raster where each cell contains a two-digit number with the first digit being the class on date 1 and the second digit being the class on date 2. Figure 10 is an example of a change detection between 1993 and 2021. Using the ice maps created in this study the change detection can be performed between any two years of data.*

Table 5: maybe you can add a "total ice area (bedfast ice + floating ice)" list to the table. The fraction trends analysis is better to company with an area trends analysis.

Thank you for a valuable comment. Total ice area has been added to Table 5 as is shown below. However, the authors prefer not to include it into the manuscript as the ice fraction analysis were done

using a lake mask from October 2020 and as such the total ice area for the years prior to 2020 is likely underestimated.

Table 5. Bedfast and floating lake ice fractions and area (km^2) in OCF from 1993 to 2021. A lake mask created using 2020/2021 lake extent was used to extract lake ice fractions.

Year	Floating ice fraction (%)	Bedfast ice fraction (%)	Floating ice area (km^2)	Bedfast ice area (km^2)	Total ice area (km^2)
1993	88	12	790	108	898
1994	86	14	779	127	906
1995	88	12	848	111	959
1996	92	8	909	82	991
2000	90	10	825	93	918
2001	81	19	753	175	928
2002	86	14	796	134	930
2003	79	21	746	203	949
2004	91	9	836	80	916
2005	77	23	687	207	894
2006	90	10	884	98	982
2008	89	11	826	101	927
2009	92	8	757	70	827
2010	83	17	792	160	952
2018	66	34	671	346	1017
2019	79	21	798	206	1004
2020	81	19	794	184	978
2021	75	25	762	252	1014