Review - Automatic detection of Arctic polynyas using hybrid supervised-unsupervised deep learning

This is an interesting proof-of-concept study building on a recently published labelled dataset to detect Arctic polynyas using deep learning. The concept is novel and the manuscript is concise and reads smoothly. The limitations of the experiments are clearly outlined. The method, despite its limitations, demonstrates the potential to be further adapted to become a full-fledged machine learning based automatic polynya detector for future Arctic climate studies. Therefore, I consider the manuscript publishable with some revisions, especially regarding the theoretical grounding of the proposed shift in the definition of polynyas. Strengthening the assessment of the method's performance when applied on climate model output is also needed. I attach my general and more specific comments as follows.

Major comments

1. I'm missing some deeper and more systematic discussions on the definition of polynyas in relation to all the analyses done, since through the manuscript this seems to have shifted somewhat, but it is fundamental to the experiment design. An example where such discussions could help to address: to what extent is the argument 'areas where SIC strongly decreases compared to the surrounding ice can be defined as polynyas' valid, so that the 'not-that-false positives' could be deemed as 'true positives?'

   One would intuitively think that the term is geographically based and would not shift with time like the authors argue for. I appreciate that perhaps the authors are pointing towards shifting the definition to a comparative one rather than an absolute one, but this should be more clearly explained. As of now, this argument seems to be 'data-driven', i.e., a change in definition is needed as the model detects 'reduced sea ice cover surrounded by more compact ice.' However, the readers perhaps would like to first know why such a shift in definition is theoretically valid and important within the scope of sea ice/climate science.

   Some more discussions on defining polynyas in the Introduction could help set the stage for these arguments. Then, strengthening these arguments in the Results and Discussion will also be desirable.

2. Lines 200-206 & Fig. 6-7: I would be hesitant to agree with the polynyas detected along the MIZ in the Atlantic Sector, and would recommend more discussions on if these are false positives, and if similar polynya detections are common in the time series. Also, I feel Figs. 6-7 do not convey enough information to do justice to the authors' analyses. For example, the authors could present a monthly time series of a cropped-in region of interest (e.g. an area of known polynya formation) in a given year in tightly-packed sub-figures, or a time series for a given month/date for all the years.

3. It would be good to know how the method transfers to the AMSR-E based 12.5km SIC product by NSIDC. If re-running on some of these data is not within the scope of the paper, some mentioning of this in the Discussions would be helpful for the readers to understand the potential of your methods combined with this SIC product with improved spatial resolution.

Minor comments

Line 4: 'from 0.1 to 2.5%' could be confusing. Perhaps 'from 10% to 2.5%.'

Line 16: 'not begin until' is a bit awkward. Consider rewording.

Line 40: "'zooms' on the image": I suppose you mean "'zooms in' on the image"

Line 131: what are the considerations going into keeping a 'land' class rather than directly applying a land mask and classifying pack ice vs open ocean? Could using a land mask help solve the issues encountered in Lines 132-133 and in Lines 147-149?

Line 144 & in Fig. 3 caption: randomly affected → randomly assigned?

Line 157: consider moving '(Table 1, first column)' to after 'The performances' to avoid confusion.

Fig. 4 caption: '2023 stops in April' → '2023 data stops in April.' 'higher sea ice values' → 'higher values in sea ice area and extent'

Line 180: 'a polynya' → 'polynyas'

Line 180: 'strongly decreases compared to' could be confusing which leads the reader to think of the time dimension. Consider changing to e.g. 'is significantly less than'

Line 181: the sentence after 'is not perfect' is confusing to read. Consider rewording. Do you mean 'the MIZ mask filters out some polynya pixels whose surrounding ice also has relatively low SIC, while in some complex MIZ regions, the MIZ mask fails to filter out some false positives?'

Line 203: 'we do not have any truthing data for it' – if not truthing, perhaps some comparisons with observation-based polynya detections in Section 3.1?

Line 204: 'the occasional false positive pixel survives' → 'occasional false positive pixels survive'

Line 214: 'The one inconvenient' → 'One inconvenience'

Line 234: 'no false negative' → 'no false negatives'

Throughout the manuscript:

'Timeseries' should be changed to 'time series'

After the first usage of 'SIC', avoid using 'sea ice concentration' to be consistent.