[Reviewer 2]

While some of the results are not new, they replicate previous findings (Moore et al. 2018, Ludwig et al. 2019). In particular, they use a coupled atmosphere-ice-ocean modelling framework that overcomes some of the limitations of the ice-ocean framework used in Moore et al. 2018 for testing the impact of sea ice thinning. The analysis regarding the increasing frequency of Polynya maker winds seems also new. I think the paper is generally well conceived and executed and makes a significant contribution to the field and I have no requests for major changes.

The paper might perhaps benefit from some reorganization that better separates the strongly supported results (lack of role of thinning) from the more "future research is needed" ones (increase in polynya frequency due to wind changes). But that's a matter of taste (and likely personal bias) and I think the paper has lingered in reviewer space for too long already so I don't think that should be a requirement.
→ We appreciate your comments. We have expanded our analysis to December and the polynya in December 1986 is now examined in the revised manuscript (see also [R1-B]). This latent heat polynya was also produced by southerly winds as strong as the one in February 2018, but its wind duration was relatively shorter. Hence, the polynya in December 1986 is the second largest one. Although the revised manuscript has become a little lengthy, there are no major changes in terms of findings and conclusions.

Details:
[R2-1] Line 80: Figure 1c. The CFS sea ice thickness looks terrible. Since you are using only above 540 mb probably not an issue for you, but maybe worth a note.
→ This is now mentioned in the beginning of the discussion.

[R2-2] Line 123… "the mean SIC was used to detect … when it dropped below 90%" I think this needs to be justified since the selection of this threshold probably affects the number of polynyas you would have detected? Does the 90% reflect some kind statistical threshold of variability or a value in the literature? This definition probably has to remain somewhat arbitrary but the sensitivity of results to this selection should be discussed somewhere in the paper.
→ It is actually the other way around. When polynya events were observed during 2010s (i.e., February 2011, 2017, & 2018), the daily averaged satellite SIC was dropped below 90% over the study region. Hence, we applied this threshold to detect any additional winter polynya events. The statement was revised accordingly.

[R2-3] Line 165. SOMs. Seems to me that the SOM analysis is a bit of an overkill in this context and may add more confusion than explanation. I think simple wind (anomaly) composites for the Polynya events would have done the job. I know, you did the work and hey, ML! but maybe save for SOM other paper?
→ We used SOM mainly to extract spatial patterns instead of using mean or anomaly fields because of a large size of RASM output: 90 days times four 6-hourly output.

[R2-4] Line 240. Looks like Figure 6 is mentioned before Figure 5 (and I don't see a reference to Figure 5). Figure 5 isn't very interesting at this scale anyhow.
→ Figure 5 is mentioned before Figure 6 in the subsection 4.1.2. Also, as the reviewer 1 suggested, Fig. 5 is revised (including color bar) to show the north of Svalbard to discuss the discrepancy of sea ice condition in February 2018; RASM overestimated its cover (see the 2nd paragraph of the discussion).

[R2-5] Figure 6. Please increase the size. This is hard to see for my aging eyes unless this is improved in production. Also, do we really need to see the full Arctic pattern or is a smaller cut out region sufficient to see what is relevant. What is the significance of the result that this temporal evolution goes through a number of SOM patterns? The key part is northward winds in the region, isn't it?
→ Fig. 6 is revised by increasing the size and the study area is zoomed in as well. The key part is to show how sea ice responds (in terms of ice divergence and convergence) to wind direction/intensity in this region.

[R2-6] Fig 7 (S3,S4). Couldn't those be condensed a bit to highlight the years that should be highlighted? What is the grey shading for the standard deviation? I would have expected the standard deviation of the SIC for that day in the full time series. Why is this larger for the "polynya" cases? See also above comment on defining the SIC threshold. I would have thought a reasonable definition would be 1 or 2 sigma of interannual variability (or quartiles or something like that)?
→ We would like to show that daily mean SIC for the region defined in Fig. 4a from the entire satellite records and determine potential polynya days rather than ones that we already know. We found that those polynya days coincide with when the regional mean SIC is below 90%. The standard deviation indicates spatial variability of SIC in each day. For example, if SIC is spatially homogeneous in a given day, then standard deviation is almost zero. On the other hand, when a polynya occurs, the spatial mean SIC drops down and standard deviation goes up. The figure captions of Figs. S3 and S4 are revised for clarity.

[R2-7] Line 275: Polynya periods… defined as.
How does this interact with your previous definition of 90% SIC. Needs some clarification.
→ Text has been added to clarify that the polynya period is based on the RASM simulation when dynamic sea ice loss (i.e., DVT is less than -10 km^3/day) for more than 3 days. The criterion stated above (i.e. 90% satellite SIC) was a sea ice condition, regionally averaged, when we found polynyas over the region. The corresponding statement is also revised.

[R2-8] Line 290.. This section could perhaps get a separate heading "What's driving changes in polynya Frequency".
Totally a style thing, but I would lead with testing the hypothesis that is rejected (thinning ice) and corroborates previous results (Moore et al. 2018) and follow with what the more likely hypotheses of "changes in winds". That would also allow a natural transition to a discussion of the "unanswered" question left for further research (why is the wind changing?).

I like Fig 8 but think that if the increased frequency in wind events is considered a key result, then its statistical robustness may need some additional support.
→ As suggested, a new subsection was assigned to this paragraph. However, we introduced the polynya in December 1986 in the following subsection and made a few changes in the beginning of the paragraph because the polynyas in the 2010s are not unique anymore. This paragraph is also revised by adding the reason why we used two ensembles at the end. In addition, the last paragraph of the new section 4.4. is revised to confirm the hypothesis: the role of southerly winds. Finally, we added a comment about uncertainty of causes changes in wind pattern.

[R2-9] Line 295: La Nina winters….
That idea seems to be not sufficiently developed. I think it is ok to document the increase in frequency of polynya making wind events (with some stats) and leave the global context for future research. Unless the idea can be developed better and/or supported through some results in the literature (Tropical/Arctic connections are a whole study area), I'd leave it out.
→ We agree to the idea that it is not sufficient enough to bring large-scale atmospheric connection. Hence it is removed.

[R2-10] Line 345… but none as large as 2018
This is left a bit dangling. What does this mean? Why do you think that is? Increase in strong wind frequency or statistical fluke… or you can't tell at this point?
→ This sentence is removed due to uncertainty.

[R2-11] Comments on prior review.
For independence, I didn't read the prior review in advance but since the D in TCD stands for "Discussion", I might as well. I agree with Frank's point that the delineation from existing work could be done better. It shouldn't be too hard to do this (see also above). Changing the analysis time window to include the missed 1986 polynya would of course be great but likely be a lot of work and require a complete redo of pretty much everything. Maybe a compromise is to reference Frank's unpublished analysis here (Is the discussion in TCD citable?), acknowledge the sensitivity of some results to the time window selection, and discuss the effect of the "missed polynya" on the conclusions. In my view, the existence of a 1986 polynya reinforces the conclusion that sea ice thickness change had little to do with the 2018 Polynya but may qualify the conclusion about the increasing frequency in the last decade (though one Polynya in 1986 doesn't necessarily kill this). Just an idea to consider.
→ As the reviewer 1 suggested, we decided to include December analysis in our analysis of satellite SIC, RASM hindcast, and RASM-DPLE runs. As you already pointed out, the polynya event in December 1986 further supports the idea that sea ice decline plays little role in more frequent occurrence in recent years. There are some changes in the manuscript because the polynyas in 2010s are not "new" any more, but overall structure, findings, and conclusions are basically the same. Interestingly, RASM-DPLE simulations produced more December polynyas in 1985-1995 than in 2015-2025.