Some PWRF setting questions:

1. 52-vertical level is used in this paper, is this number enough, what is the lowest model level? Like ASR v2 is 71 levels, with lowest level at 4m. I don't think re-run any model is necessary, but I would suggest authors provide more info in the table 1. PWRF default (at PMG) is 71 levels, with more levels at lower altitude. That is designed to better capture near surface conditions, and cloud formation.

   We tested higher vertical resolution in Figure 3 by increasing the number of model levels to 75. In the 52-level configuration, the lowest model level is located 12 m above ground level, whereas in the 75-level configuration the lowest level is at 4 m AGL. Increasing the vertical resolution has a negligible impact on the resulting CRE distributions. We have now explicitly listed the lowest model level in the model configuration table and clarified this point in the discussion accompanying Figure 3.

2. I might miss it. Is unified Noah or Noah MP is being used here?

   We are using Unified Noah here. We have marked this more explicitly in the model configuration table. We found that, in our use cases, Noah MP tended to lead to instabilities in the Polar WRF model.

3. Is there any nudging applied? If so, what variables for what levels?

   The full set of boundary conditions are nudged towards the ERA5 reanalysis data every 6 hours (see model configuration table). However, the domains are freely evolving. We mention this explicitly in the revised manuscript (Lines 82-86).

4. This is not a suggestion, but more like a discussion. P3 (option 50) in general provides pretty good results, especially related to super-cooled liquid water in clouds. However, we also notice sometimes it can produce large value of cloud LWP (e.g., like unrealistic large for a short period of time at WAIS compared to obs. While Thompson & Morrison did a better job). I believe Hines et al. 2019 also kinda mentioned that. I am wondering if authors encounter similar issue in Arctic.

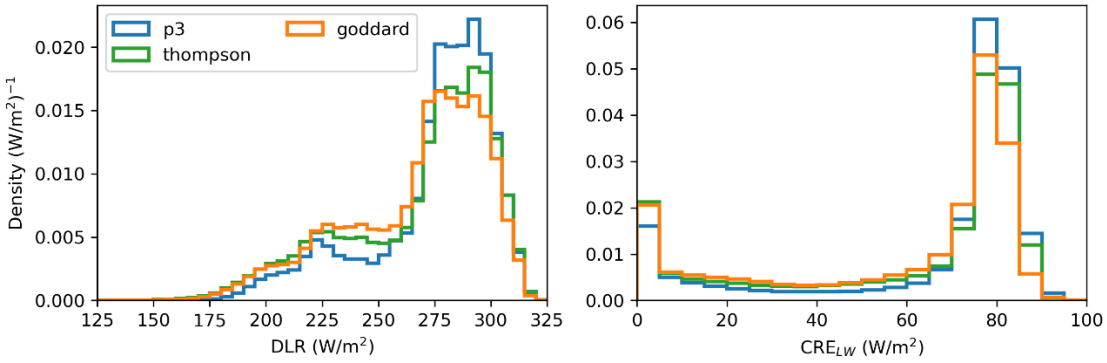
   Yes, we observed a similar issue in some cases. That said, because LW cloud radiative effects saturate relatively quickly with increasing water path (beyond ~30 g/m2, further increases in water path have little impact on CRE), this was less problematic for our analysis, which focused on comparing CRE distributions. However, this behavior may be a key reason for the overproduction of opaque cloud states, as clouds that should maintain lower LWP can unrealistically grow too rapidly.

5. I am not familiar with ERA5 default surface type for Arctic regions (aka whether it is good enough). For Antarctic region, ERA5 use a quite old land surface cover (which still have Larsen A & B ice shelf included…). Thus, a more up-to-date land-ocean description is usually needed (Like REMA). For high-resolution simulations, more accurate SST and SI observations are usually recommended to be included as initial fields. As authors mentioned in Ln 220, difference surface type matters. I am wondering if more information can be provided here regarding the simulations has been done in this study.

   We did not test a more advanced land-ocean surface representation. Our rationale was that ERA5 is widely adopted and provides a convenient "plug-and-play" option for Polar WRF (whose development is more niche than standard WRF). That said, Figure 8 somewhat ironically suggests that, in its current configuration, the model is largely insensitive to surface conditions, despite observations indicating that the frequency of these cloud states is strongly controlled by them. As a result, incorporating a dataset with a more accurate surface representation would likely require additional development of the atmosphere–surface coupling to get any benefit. This development effort is beyond the scope of this study.

6. There is no need to add any simulations. I am wondering authors have ever tested Thompson (aerosol aware) by any chance?

   Yes, we tested it briefly. We found that its behavior was similar to the Goddard scheme, in that the cloud states were less distinctly separated and tended to be somewhat blended. For reference, we include below an example from a simulation we still have archived for 1 October 2007, comparing results from P3, Thompson, and Goddard.



Minor:

Ln 15-20: Several research suggested a poleward shift of atmospheric river activities for Arctic region, which will enhance the transport of moisture, energy and warmth. I think this is worth mentioning as background change for Arctic region.

Zhe Li, Qinghua Ding ,A global poleward shift of atmospheric rivers.Sci. Adv.10,eadq0604(2024).DOI:10.1126/sciadv.adq0604

We have included this citation in our discussion of increased poleward energy transport in the Arctic

Introduction structure:

The Intro is in a good shape in general. This is just a suggestion. Introduction section has a decent number of short paragraphs, I am wondering authors ever think about merging some of those, to deliver the key messages more clearly.

As suggested, we have merged some of the shorter paragraphs into one.