Comment on wcd-2021-18
Anonymous Referee #3

Referee comment on "Identification, characteristics, and dynamics of Arctic extreme seasons" by Katharina Hartmuth et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2021-18-RC3, 2021

The authors have investigated seasonal extremes in the Arctic using PCA of six climate variables and analysis of some key dynamical elements – cyclones, blockings, and marine cold air outbreaks – to further investigate particular extreme seasons. This is an interesting and valuable framework for understanding the various causes of seasonal extremes, and it is very well presented. I recommend the manuscript for publication with some minor adjustments. My principal concerns relate to the justification of the many choices which needed to be made in this analysis, these are detailed below.

Specific comments:

L13: “respectively” – this doesn’t quite follow when you say 2-3 extreme seasons for four seasons.

L15: I think a justification of why 2 winter seasons were chosen for the in-depth case studies is needed here.

L117: It is very nice to have these questions in the Introduction to frame the paper, but as far as I could see the synoptic systems of interest are pre-defined in the study (cyclones, blockings, and marine CAOs), so perhaps this question should be reframed to reflect this.

L131: What was the method(s) of interpolation?

L155: What is the justification for choosing these regions?
L161: Are results sensitive to the choice of definition of ice, mixed, and sea? Why were these thresholds chosen?

L174: Why choose just the first 2 PCs? This seems arbitrary, although I see later you mention that these explain a very large part of the overall variance.

L178: Why are these rescaled by their respective SDs to give equal weight to each PC? Do you not wish to identify the extremeness of a season rather than the extremeness of a season with respect to these two PCs? (ref L114) If you don’t do this rescaling do you still identify the same seasons as being extreme seasons?

Fig 8 and elsewhere: why was 10^5 km^2 chosen as the size threshold for a region?

L393: grammar – “This periods typically are...”

L408: “exemplarily shows” -> “exemplifies”

L510: “as is the case”