The paper presents an analysis of variability in three Arctic regions using 6 metrics. An input of those metrics into the dominant modes of variability and links between those metric are discussed. Overall, I am impressed by the amount and quality of work done in this study.

Here is what I like about the paper:

- Fig 5 and 6, which show that while strong anomalies may be observed in one or two metrics, other metrics may remain close to their climatological values;
- assessment of the input of the six metrics into the main modes of variability and relationships between them;
- case studies (particularly fig. 10,11, 14) and the discussion around them. An attempt to establish a connection between the weather and seasonal anomalies is valuable;
- a wide range of metrics used in the study - not only t2m/SIC/P, but also energy fluxes, cyclone frequency, CAO and a blocking index.

However, there is a couple of major concerns that need to be addressed before the paper can be accepted for publication:

1. I am not convinced that the approach, introduced in the paper, is a good way to select extreme seasons. Despite the use of a multivariate approach, it often comes to just one metric showing a strong seasonal anomaly, which was enough to identify the season as extreme or anomalous. Thus, without applying this approach, one may simply go through all 6 metrics and select the most extreme season(s) in each of them. I don't think I saw a proof that the seasons selected with the PCA analysis were more anomalous than those that showed a strong anomaly but were not picked up by the PCA approach. The latter may be even more anomalous than those, that were selected using the PCA.

On the other hand, there are seasons that were identified as anomalous though none of the variables showed a strong anomaly. Could it be proved that they are ‘true’ anomalous
seasons and not artefacts of the method?

I am not asking for a change of the approach here, but I think more discussion around potential (dis)advantages of the proposed method is needed. In my opinion, this method identifies the dominant modes of variability and allows for assessment of the contribution of each of 6 metrics into those modes and a link between them. Section 5 explores a few seasons when one of the first two modes of variability was among the strongest.

2. My other concern is the length of the manuscript. Considering the amount of work, it is hard to make it shorter, but I think the paper will benefit from it. Some plots (especially, Fig. 3) are too busy and are difficult to interpret. Section 3 and 4, while interesting, are hard to read, particularly when plots discussed in the text are a couple of pages away (which is inevitable). Please select the most robust and/or important relationships and focus on them. I understand that each plot provides a lot of information, but, unfortunately, human beings can only keep a few facts in mind at a time.

Other comments:

Abstract: 1. The abstract is a bit long, even if there is no word limit, a page-long abstract is not ideal.

2. I think it is worth mentioning that 2016/17 winter was mostly anomalous in terms of precipitation and maybe in some other variables, otherwise, until you read the paper, it remains unclear why it was anomalous.

Sect. 2.3: For the PCA analysis, was each metric was first averaged over the corresponding region? Meaning that the special structure of those anomalies was not accounted for.

Fig. 3: As I already mentioned above, it is a very busy plot, which is hard to read. The only thing that is obvious to me that in JJA the red/blue markers can be linked to positive/negative temperature anomalies. For DJF, what is obvious is a link between T2m and P anomalies and that the low right corner has predominantly negative Es anomalies. However, regional differences, discussed in the text, are very hard to see. If you decide to keep this plot, maybe splitting into different geographical locations or the sea ice concentrations helps.

l.230: Despite good clustering in Fig. 4, this plot is again very busy. Maybe you can show the average location for each of the nine sub-regions on top of the existing plot.

Fig. 5, ‘the seasonal-mean absolute anomalies’: are these the seasonal-mean absolute daily anomalies, as in Fig. 4?

Fig. 5, 6: why Nordic seas are not shown?

l.285-287: The statement on correlation between T2m and P comes from the fact that the corresponding blue lines are close to each other?

Regarding the comment on the weather systems creating extremes in the high Arctic, I would like to agree, though none of the AR seasons across all regions in fig. 5 look particularly extreme. How about other regions that have stronger extreme seasons often just in one parameter - can they be explained by anomalous weather patterns?

l.303-307: Why the described connection between P and Rs over the sea, as well as between T2m and RL in KBM does not hold in the Kara-Barents Sea?
Section 4: The relationship between 6 metrics during cold and warm seasons, gained from the PCA analysis, is interesting. Could correlations found in this section be confirmed by using the raw data?

I. 322: “By design, extreme seasons have very large anomalies for at least one parameter... However, some anomalous seasons don’t show very strong anomalies in one particular parameter, which implies that for these seasons it is the combination of several parameters that makes them anomalous” I am not sure that the first sentence is true. Moderate anomalies in a few variables may also give an anomalous season and this is what happens in some cases.

I.367: I could not find a description of how cyclones, CAO and blocking events were defined.

I.372: Even during CAOs the temperature remained above the climatological mean, hence, I doubt that 38%-deficit in CAO can be responsible for the season being anomalous. During the first month (days1-27), there was no significant blocking events and CAOs, but T2m was well above average. To me it looks like there was a strong preconditioning. Furthermore, in the next case, shown in Fig. 10, there is a high number of CAOs but they have relatively small effect on T2m, especially during the first half of the season, brings the temperature down by only, perhaps, 2-3 deg.

I.465: A seasonal blocking anomaly over Scandinavia is probably not enough to support the statement that ‘Subsidence-induced warming [over Scandinavia] and long-range transport of warm air masses contributed to several warm episodes.’

I.498: why a persistent high does not cause subsidence warming? and why there are no blocking events during Jan 2013 at the time of a persistent high? I can also see a number of cyclones in Feb, despite the text says that Feb was also calm. I agree that probably the main reason for decreasing t2m and low P is that the High Arctic remained isolated from the lower latitudes, however, none of the metrics in this study reflect an exchange between latitudes. I am not suggesting adding such metric at this stage, but it might be something to add in the future.

I.529-534: the paragraph first describes obvious seasonal differences (higher variability in winter due to stronger gradients) and then concludes ‘hence, it is reasonable to subdivide the Arctic into several regions considering these spatial differences to study anomalous Arctic winter seasons.’ But during summer the regions were also subdivided. I am not sure if this paragraph is needed at all.

I. 541: see my major comment on the PCA approach

Minor comments:

I.61 ‘and of the feedback’: remove ‘of’

Table 1: Es should be added

Table 2 is first mentioned in section 2.3 but is only shown in section 4. Replace ‘brackets’ with ‘parentheses’

I.160: it is not the entire ERA5 period, but the entire period covered by this paper

Please use either the Kara-Barents sea or the Kara and Barents seas
406: I’d replace ‘single’ with ‘individual’

432: on this date

Fig. 13 is mentioned earlier than fig. 12.

Fig. 10, 11, 14: I suggest showing months and days of months’ along Axis X, instead of days of season, as specific dates are often mentioned in the text (e.g., 9 Jan or 17 Feb). Can SLP be added to fig. 10, 11? In fig. 14 the legend mentions CAOs, but they are not shown - could they be added?