Dear Referee, Thank you for the time that you have spent on our manuscript. We are happy with your positive response and grateful for your comments and suggestions. These certainly contributed to improving the quality of our manuscript.

Below you will find a summary of the changes that we have made throughout the manuscript to address all your suggestions. The replies to your comments are written in blue, while your comments are reproduced in black. Please, notice that line, page, and figure numbers mentioned in our rebuttal letter refer to the new version of the manuscript.

Yours sincerely and on behalf of all the co-authors,

Guillian Van Achter

Anonymous Referee #2

GENERAL OVERVIEW

This brief communication looks at the temporal and spatial variability of Arctic sea ice volume and thickness (respectively) using the CESM large ensemble over three multi-decadal to multi-centennial periods (pre-industrial, historical, and future). A wavelet analysis was used to explore the peak modes of variability in SIV whereas the authors used an EOF analysis to explore the spatial modes of variability in each period. The key findings of the study are that there are two peak modes of SIV variability in the pre-industrial and historical time periods centered on 8- and 16-year periods. The first of these temporal modes is shown to be related to the first spatial mode of SIT variability, which shows a strong AO signature. The relationship between the first mode of SIT and the temporal mode of SIV is made more certain by a wavelet analysis performed on the first principal component of SIT, which is found to also have a peak at 8 years. The other key finding is that both temporal modes basically vanish after 2050 when SIV reduces by 50% in winter. The results of the analysis are original in that the temporal and spatial analyses are brought together for the first time in this way. Overall, I rate the study in terms of originality, scientific quality, significance, and presentation as fair to good.
Again, we thank the referee for her/his time and the detailed revision of our manuscript. We appreciated very much her/his comments, which were all taken into account in the revised version of the paper. Below, we answer point-by-point all specific comments.

The EOF analysis is informative, but have the authors considered identifying spatial modes of variability by correlating the wavelet time series associated with the peak periods for SIV with sea ice thickness? It could be interesting to see if the SIT spatial patterns agree with the EOFs, especially since the EOFs are constrained by orthogonality.
Very interesting suggestion, we thank the reviewer for that. We computed the wavelet time series associated with the first peak of variability over the pre-industrial period. Then we correlated this time series with the spatial sea ice thickness anomalies. Figure 1a (in this document) shows the correlation coefficients with 95% of significance. The correlation is quite low, but still significant (0-0.3) and the Atlantic/Pacific dipole can be found, even if the spatial pattern has some differences with the first EOF. The Pacific part is centered in the center of the Arctic Basin and the Atlantic part is smaller, restricted to the region near the North and East coast of Greenland. Furthermore, looking at Figure 1a, we note that the highest correlation values are located in areas where sea ice drift differs the most between high and low indices of the first EOF (see Figure 3 of the manuscript). This would suggest that we are indeed showing spatial variability of the 8-yr peak that is linked to the Arctic Oscillation.

On the other hand, this method may not be the most appropriate for analysing the spatial variability. Indeed, the wavelet time series associated with the first peak captures the main part of the sea ice volume anomaly. Therefore, the wavelet associated with the first peak can be seen as an approximation of the SIV anomaly. By correlating an approximation of the sea ice volume anomaly with the sea ice thickness anomaly fields, we are not surprised to obtain a high correlation near the center of the Arctic Basin, as it is the case for the correlation map between SIV and SIT (Figure 1b).

In conclusion, the spatial pattern from the wavelet time series associated with the first peak correlation with the SIT fields does have similarities with the EOFs. But from our perspective, these spatial patterns are not comparable.

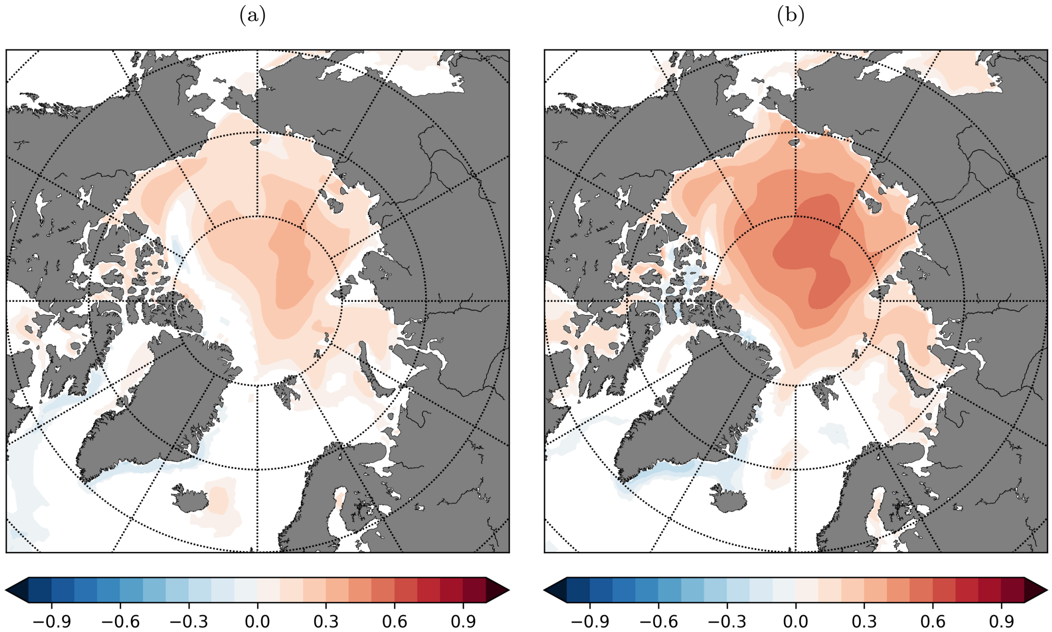


Figure 1: Correlation map between the SIT anomaly (pre-industrial period) and the wavelet time series associated with the first peak of the wavelet analysis (a). Correlation map between the SIT anomaly (pre-industrial period) and the SIV anomaly (b).

SPECIFIC COMMENTS

L20-25:

- Labe et al., 2018 did an EOF analysis comparison with PIOMAS and CESM LE monthly ice thickness that should be referenced here

  Indeed, thank you for pointing this missing reference. The paper has been added in the introduction [pg. 1, l. 21].

- The Singarayer and Bamber study didn't look at sea ice thickness, just concentration.

  Indeed, we removed it from the text.

- should state that these studies all used model-based thickness reconstructions.

  This is indeed an important information, it has been added in the text [pg. 1, l. 22].

- the EOF mode variance numbers cited were only from the Lindsay and Zhang study, so should state this. Also could note that the spatial structures of these modes and the amount of variance they explain can be sensitive to whether (and how) the SIT time series are detrended prior to performing the EOF or K-cluster analyses, the season and the time period considered, and the model.

  The paragraph has been rewritten in order to take all previous comments in account [pg. 1-2, l. 20-31].

,
L60-65:

- It needs to be clarified whether the analysis is being performed on a single ensemble member from the large ensemble or the full ensemble. If the former was done, then I would suggest at least commenting on how robust the analysis is if performed on other ensemble members. Ideally though, the full ensemble would be used. For instance, the wavelet analysis could be performed on each ensemble member and then the results of the wavelet analysis could be averaged together. It's not immediately clear to me how one would do this for EOF analysis, perhaps by appending the ensemble members to one another.

  We thank the reviewer for this comment. Since only one member spans from 1850 to 2005 we had decided to use only one historical member for the analysis.

  We tested the robustness of our one-member EOF and wavelet analysis compared to the 30 other ensemble members. Since only one historical member is spanning the 1850-2005 time period, the historical period is now 1920-2005. For this analysis, we removed the ensemble mean from each member to obtain detrended SIT anomalies. In order to apply the EOF to the 30 members, the members were appended together over time.

Figure 2 presents the first three modes of SIT variability over historical (1920-2005) period. The modes are similar to the one of the study for the historical period (1850-2005). Figure 3 presents the first three modes of SIT variability over future (2005-2050) time period. The first mode is similar, the second has the same pattern with small differences and the third has a different pattern of variability.

We conclude that our results for the EOF analysis over one-member are robust with the other ensemble members. The first and second mode that were described in the previous version of the manuscript are still present in the historical analysis over 30 members and the first one is still present in the future analysis. In the new manuscript, the EOF analyses for historical and future period has been changed from a one-member to a 30-member analysis.

(a) Historical, first mode **20%** (b) Historical, second mode **11%** (c) Historical, third mode **7%**

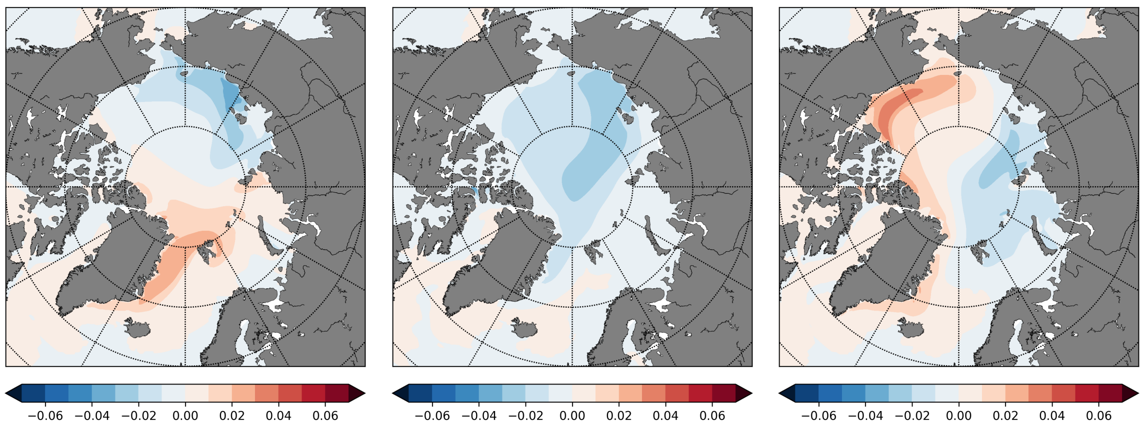


Figure 2: Modes of Arctic SIT spatial variability. First (a), second (b) and third (c) EOF of Arctic SIT over the historical period (1920-2005). EOFs are performed over 30 ensemble members by appending them over time before applying the EOF analysis.

(a) Future, first mode **20%** (b) Future, second mode **11%** (c) Future, third mode **7%**

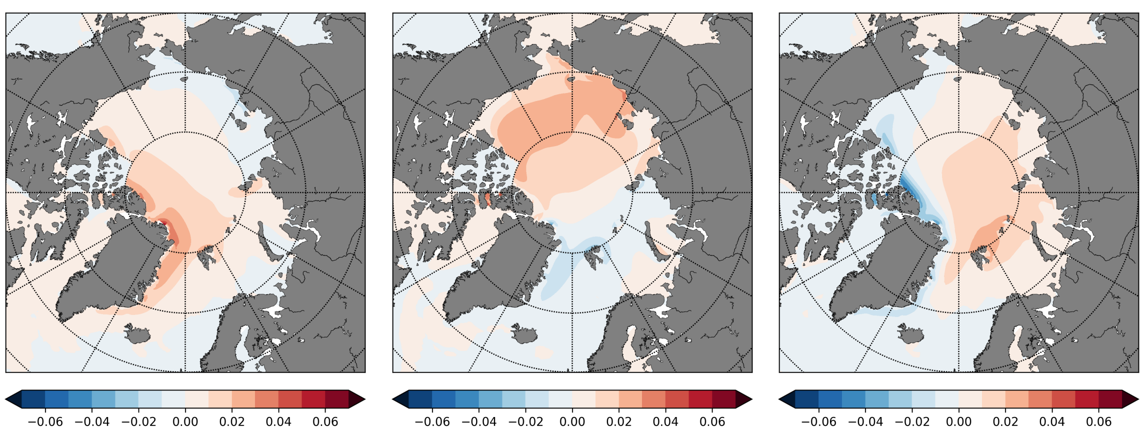


Figure 3: Modes of Arctic SIT spatial variability. First (a), second (b) and third (c) EOF of Arctic SIT over the future period (2005-2050). EOFs are performed over 30 ensemble members by appending them over time before applying the EOF analysis.

We also performed an ensemble-analysis for the temporal variability by applying the wavelet individually for each member. Afterwards, we averaged the results together (following your suggestion - Figure 4). Averaging the spectrum does not seem to be appropriate. It smoothed out any interesting information in the wavelet power spectrum (temporal-variability of the peaks), the stationarity of the SIV is not distinguishable in that form. By examining the wavelet analysis from each of the 30 members, we noticed that, for most of them (28 out of 30), 2 peaks of variability are significant and easily recognisable, one around 5-10 yrs and another around 15-25 yrs. Figure 4 presents all peaks that are significant for all ensemble members. The number of peaks (black line) shows that, for the 30 members, most of the peaks have a period of either 5-10 or 15-25 years. Because the peaks are not centered exactly at the same periods of variability, the peaks in the averaged wavelet analysis are no longer easily distinguishable.

We conclude that most of the ensemble members have two peaks of variability within the same ranges (5 to 10 and 10 to 20 years) mentioned in the paper. Because wavelet analysis over only one member has a better representation of the non-stationarity of the SIV and because the peaks are more distinguishable, we kept the one-member analysis in the study but we added a paragraph about results robustness with the ensemble members [pg. 6, l. 128-139].

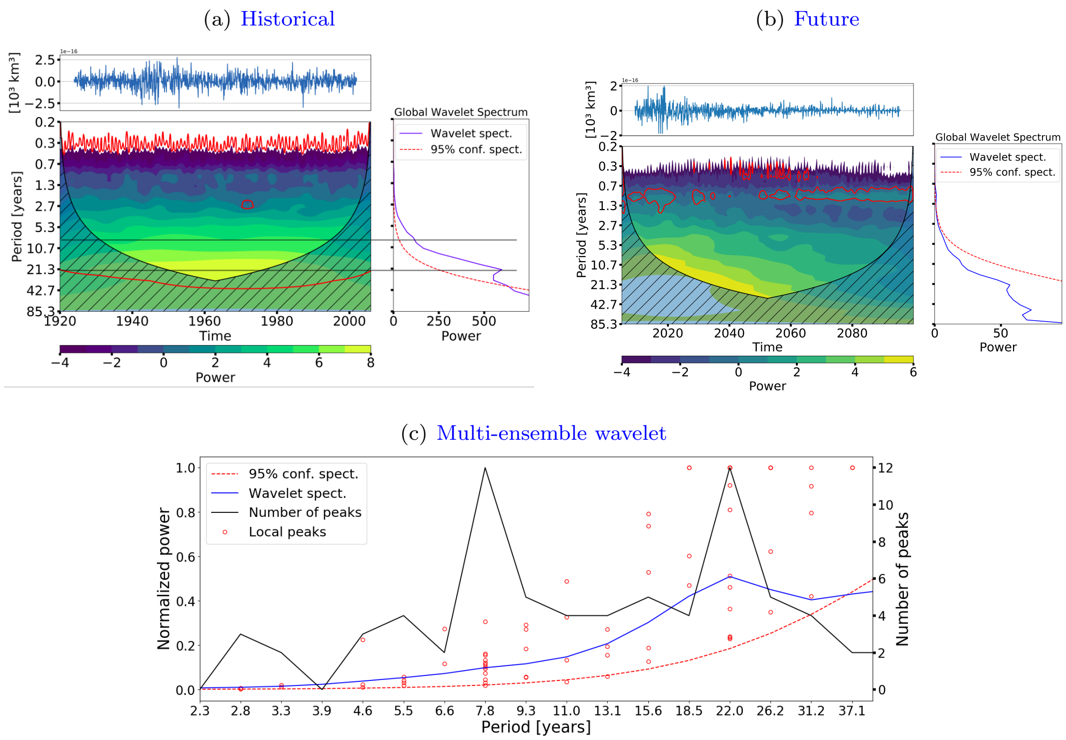


Figure 4: Wavelet analysis applied to the Arctic sea ice volume anomaly over the historical (1920-2005) (a) and future (2005-2100) (b) periods. Peaks of significant variability for 30 ensemble members (c) for the historical period. Wavelet analysis is applied to 30 ensemble members then the results are averaged together.

- One of the advantages of the large ensemble is that the externally-forced signal can be removed from the ensemble by subtracting the ensemble mean from each ensemble member. This makes detrending by fitting to a polynomial unnecessary and actually inferior since it requires an assumption about the functional form of the response to the forced signal.

  This approach is indeed better for a large number of member. We used this method when we applied the EOF and wavelet to the 30 ensemble members. It is also clarified in the manuscript [pg. 3, l. 66-70].

L70:

- Should offer some more details on how the wavelet analysis was performed [software used, wavelet function (ah I now see it says in Fig. 1 caption, but should say it here), and whether any normalization was used.]

  Details about the mother wavelet has been added in the text [pg. 3, l. 76] and the python module used to performed the wavelet analysis is now described in the Code availability Section. No normalisation is used, better description about the pre-processing of the data is given in Section 2.1 of the study.

L100:

- There is still overlap though between the occurrence of each peak so it's maybe not very accurate to say they don't occur at the same time.

  We agree with the reviewer. For more accuracy the sentence was rewritten in the text [pg. 4, l. 105].

L110:

- It might be the contouring but I'm having a hard time seeing a significant area at the 5-year period during the 2010-2025 period.

  It is more correct to say that the area is 2015-2025. We also changed the contour color [pg. 4, l. 116].

Figure 1:

- I recommend changing the units on the vertical axis from months to years since these are the units used in the text.

  We updated the figure.

- What is the yellow contour representing? Presumably it's statistical significance, but if it is then I think it would be easier to see if it were colored red as in the time integrated plots. It should also be stated.

  We changed the colour to red, and the description of the red contour has been brought to the text [pg. 4, l. 96] and in the caption of the figure.

- What are the areas of white on the contour plot representing?

  The white areas are marked by power values which are out of the range defined by the figure's color bar. For a better reading of the figure, we kept those white areas but we explained their meaning in the figure caption.

L120:

- Do the authors have any idea why the leading EOF mode over the historical period explains much less variance than that found in Labe et al., even though the spatial pattern looks the same? Could it be that it's due to the use of all ensemble members in Labe et al.?

  This is indeed an interesting question. Even after applying the EOF analysis to 30 ensemble members, we still have modes that explain much less variance than the ones of Labe *et al.* The differences are that the sea ice thickness are linearly detrended in Labe *et al*, while we use a polynomial approach to detrend the SIT. This linear detrending might not be sufficient if the forced signal is accelerating over time. This would leave too much low-frequency variability in the anomalies. Another difference, but less significant in terms of explained variability differences, is that their SIT fields are weighted by the square root of the cosine of their latitude to account for converging meridians toward the pole.

- Relatedly, I'm surprised that the variance explained by the first three modes doesn't sum to a higher number. For instance, looking at Lindsay and Zhang study, the first three modes identified in their study sum closer to 60-70%.

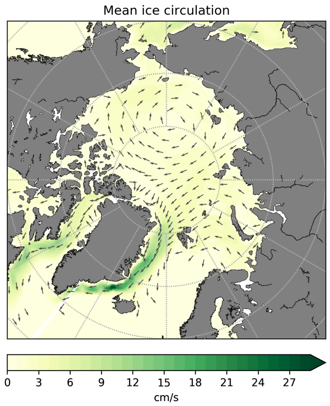
  In this case the greater estimation of the explained variability could come from the fact that Lindsay and Zhang applied EOF to annual mean sea ice thickness. By applying EOF analysis to yearly and not monthly gridded SIT, the first EOFs will represent more of the total variability. Also, the EOF are applied over a shorter time period (1953-2003). This could also increase the explained variability for each components.

L140:

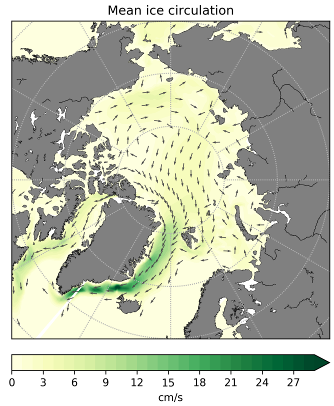
- Why was the first mode of SIT variability only compared with ice velocity and the second mode only compared with temperature? Why not compare each mode with both ice velocity and temperature?

  Thanks to your comment, we updated this section by applying the sea ice velocity analysis to the first and second modes of SIT variability. It turned out that the first mode presents AO signature, while the second mode does not. The latter has the same sea ice velocity pattern for both high and low indices (Figure 5 in this document). Then we also applied surface air temperature analysis to the first mode, which presents also differences between indices. We kept the results in the paper but we changed our conclusion. As you suggested, the surface air temperature and the sea ice thickness variabilities are associated with each other but we can not conclude anything about its causation [pg. 8, l. 179-184].

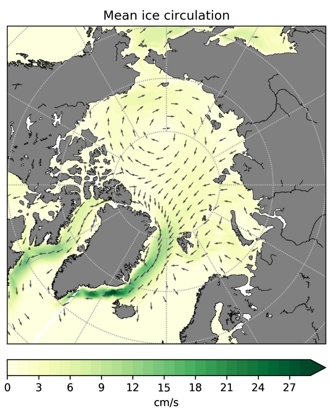
(a) High index, first mode



(b) Low index, first mode



(c) High index, second mode



(d) Low index, second mode

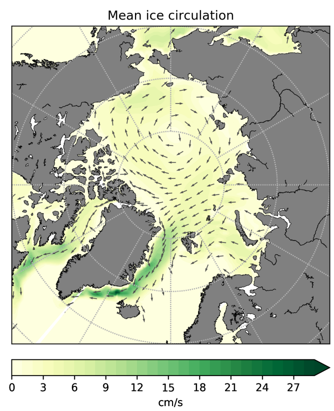


Figure 5: Mean sea ice velocity for both high and low indices of the first and second modes of SIT spatial variability. Analysis is performed over one ensemble member.

L155:

- I'm not sure it's accurate to say that the air temperatures are causing the variability in SIT from this analysis. They appear to be associated with each other, but this doesn't imply causation. For instance, a thermodynamic response from the ice thickness variability is just as plausible. Nonetheless it is at least consistent with Olonscheck et al. (as stated though, they looked at ice area not thickness).

  This has been answered in the previous point.

TECHNICAL COMMENTS

L85:

- I would suggest separating this sentence into two; it's currently a run-on (the two uses of the word "by")

  We have changed the text according to your suggestion [pg. 4, l. 89-91].