The authors adequately addressed the comments/suggestion made to the original version of the manuscript. Please find below a few more generally minor comments I would suggest the authors to consider before the draft can be published. The line numbers refer to the new version of the document with changes highlighted.

Line 300:

“Interestingly, G1 demonstrates a higher rate of sea-ice increase from 1600 CE until the LIA, which is not apparent in the general trend (Figure 2).”

Period of 1600 until late 1800s is associated with LIA, so the statement "from 1600CE until the LIA" needs to be corrected/rephrased. Otherwise, it appears from this statement that LIA is associated with a specific time point, not a period.

Line 423:

“The satellite data, on the other hand, is the sea-ice extent, which considers the sum of the area of all grid cells covered by at least 15% of sea ice.”

Why did the authors then used the extent, while the area is also available directly from the NSIDC SSMI data? This would make the comparison way more consistent. For winter the difference between the two metrics will not be very substantial, but in the summer this is not the case. I recommend the authors to switch to similar variables.

Line 523:

“This difference is likely caused by internal climate and ecosystem dynamics where proxy-based sea-ice reconstructions were affected by complex environmental parameters, while the physical models directly responded to large-scale atmospheric forcing from variations in solar, volcanic, and orbital forcing. Furthermore, while the proxy records used here have relatively high resolution (< 100 years), the resolution is generally not enough to capture multidecadal oscillations.”

In addition and in my personal opinion, are likely most important, chronology errors of various nature (core sub-sampling itself, delta R uncertainty, depth age modelling method and the associated uncertainty etc) is also a serious obstacle when comparing various oceanic proxy series on sub-centennial time scales.

Line 531

“dominance” appears twice.

Line 531:

“…while periods with high resolution confirm the dominance of dominance of multidecadal variability …“

The authors can easily estimate how correct this statement is directly from wavelet analysis by summing the variance over the band of timescales corresponding to multidecadal variability. See example in Torrence and Compo, (1998).
It is also useful to indicate (again) that this dominance applies to the timescales over a decade, since the analysis is applied to filtered series with sub-decadal variations removed. Otherwise, annual to intra-annual variations generally contribute most to the total variance.

Line 535:
“The spatial distribution of EOF2, associated with the multidecadal to centennial-scale fluctuations…”

Better, in my opinion, refer instead to a “Spatial structure of EOF2” or a "Loading pattern"

Line 537:

“This inconsistency is probably caused by the low explained variance of PC2 (15% for CESM1 and 5% for MPI-ESM), incorporating a large degree of residual noise. Although the multidecadal and centennial variability of Arctic summer sea ice has been linked to changes in the northward Atlantic and Pacific heat transport and in the Arctic dipole pattern, there is still some significant variability between mean states of Arctic sea ice simulated by different models (Li et al., 2018).”

Another possible explanation is purely methodological, namely the orthogonality, by definition/construction, of EOFs in the basic EOF analysis. It means, in plain words, that EOF1 has a unipole spatial pattern, EOF2 - bi-pole, EOF3 4-pole etc. For the region used sea ice is mainly present on both sides of Greenalnd, this is where EOF2 will form its two poles with loadings of the opposite sign.

Line 577:

“The recent sea-ice retreat, on the other hand, was likely initiated by a recovery from the volcanic dust emissions…”

Better rephrase to "recovery after the LIA", since volcanism might have triggered a number of feedbacks that caused lasting negative SAT and positive sea ice anomalies that are now associated with the LIA manifestation. From the way it is written now, an impression of a direct liner response to continuous volcanic dust emission emerges.

And the final comment concerns the hypothesis testing in wavelet power spectra (Starting from line 527). In their response letter the authors agree that resampling to the annual scale causes unrealistically high values of the AR-1 coefficient used in hypothesis testing. The authors, however, are not correct stating in the response that “The deteriorated signal-to-noise-ratio in combination with overestimated persistence would, however, rather underestimate significance and can be assumed to be on the conservative side.” Due to a redistribution of variance towards lower frequencies in the AR1 series with a very high autocorrelation coefficient (very close to 1), the significance threshold for hypothesis testing will be underestimated in the higher frequency range, and overestimated in the lower frequency range. The actual configuration will depend on the AR(1) value of the background process that could generate the observed series. I therefore consider that the use of thresholds for significance testing calculated from the resampled data is not correct. Instead, I recommend the authors (if they would like to retain the discussion that involves significance testing from wavelet spectra of the proxy series) to use the AR(1) coefficient estimated from the original unevenly sampled data. Since the authors used the R environment in their calculations, it should not be a problem to use RedFit method (Schulz and Mudelsee, 2002) for this purpose. The implementation of RedFit can be found in dplR package (https://cran.r-project.org/web/packages/dplR/dplR.pdf). Once the AR(1) coefficients estimates are obtained, you can disable the automatic calculation of AR1 in wavelet power spectra computation/analysis

procedure and type in the RedFit estimated coefficients instead. This will provide you a way more fair view on the wavelet spectra of proxy series.

References:

Schulz, M. and Mudelsee, M. (2002) REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. Computers & Geosciences, 28(3), 421–426.

Torrence, C., and G. P. Compo, 1998: A Practical Guide to Wavelet Analysis. *Bull. Amer. Meteor. Soc.*, **79**, 61–78, https://doi.org/10.1175