

Anonymous Referee #1

general comments:

This is an interesting study addressing the evolution of the interannual variability of Arctic sea ice area and its causes. The study is based on analyzing historical and RCP8.5 simulations of CESM-LE and, in part, 12 CMIP5 models. The study primarily finds an inverse relationship between the rate of sea ice retreat and the magnitude of variability. The authors further show that a sufficiently thin ice cover fosters the variability and that thermodynamical processes dominate over dynamical processes in causing this variability. Although I find the study relevant, well written and structured, I have some major concerns about the novelty of the findings, some methodological aspects and the robustness of their conclusions. I recommend publication in TC only if these major concerns will be addressed.

Thank you for the thoughtful evaluation. In this revised version, we have addressed all three of the reviewer's concerns regarding novelty, methodology, and robustness.

specific comments:

Title: The paper does not only address the future interannual variability but also the past.

Good point. The title has been changed to "Past and future interannual variability of Arctic sea ice in coupled climate models".

#1 ll. 47-48 The "important physical and societal consequences" could be given more specific.

Although there isn't room in the abstract to explain the physical and societal consequences of short-term variations in sea ice as the pack diminishes, some of these impacts are covered in the Introduction. These include enhanced marine navigation through the Arctic, amplified positive ice-albedo feedback, and increased sea ice sensitivity to ocean heat transport initiated upon short-term declines in ice cover. The components of the surrounding ecosystem will also adjust as the physical environment changes.

#2 ll. 60-62 It is not clear from the abstract what "thermodynamic processes" exactly mean. I would like to have this more specific (e.g., open water formation efficiency), especially because you specifically name the dynamic processes, which you find to be less important, but not the important thermodynamic ones.

We have added more specificity by clarifying that the thermodynamic processes involve melting (top, bottom, lateral) and growth (frazil, congelation). There is no more space to get into additional details within the short abstract.

#3 I. 106 Where does the judging statement "likely" comes from? Is this justified in the given references according to IPCC language. As you mention, the likelihood of summer ice-free conditions strongly depends on the emission scenario. For RCP8.5, it might be rather certain, for RCP2.6 it is not. "Likely" is vague here.

To avoid bogging down in the spread and plausibility of the various RCP scenarios, we have changed the sentence to read that that the Arctic "may" become seasonally ice-free within a few decades.

#4 I. 117 To me, there is no logical link between the reduction in sea-ice extent and the loss of multi-year ice. The reduction in sea-ice extent is not obviously the cause for the loss of multi-year ice. Please rephrase.

Although the reduction of sea-ice extent does not directly cause a loss of multi-year ice, past studies do show that a large portion of the reduction of sea-ice extent in the past decade is associated with the loss of multi-year ice (Kwok et al., 2010). That reference is now added. We have rephrased the sentence, "As the Arctic sea ice pack thins and retreats, multi-year ice is being lost and there is consequently a larger proportion of seasonal, first year ice."

#5 II. 122-123 How does decreased ice thickness amplifies the ice-albedo feedback? Please explain.

As demonstrated by observations (e. g., Grenfell and Maykut 1977) and applied in various parameterizations and models (e. g., Maykut 1982, Ebert and Curry 1993), the albedo of thin, first year sea ice is lower than that of thick, multiyear ice. The CICE sea ice model used in CESM-LE also lowers the surface albedo for thin ice floes (Hunke and Lipscomb 2010).

Grenfell, T. C., and G. A. Maykut, 1977: The optical properties of ice and snow in the Arctic Basin. *J. Glaciol.*, 18, 445-463.

Maykut, G. A., 1982: Large-scale heat exchange and ice production in the central Arctic. *J. Geophys. Res.*, 87, 7971-7984.

Ebert, E. E., and J. A. Curry, 1993: An intermediate one-dimensional thermodynamic sea ice model for investigating ice-atmosphere interactions. *J. Geophys. Res.*, 98, 10085-10109.

Hunke, E. C., and W. H. Lipscomb, 2010: CICE: the Los Alamos sea ice model documentation and software user's manual, version 4.1. LA-CC-06-012.

#6 II. 141-142 I appreciate that you specifically mention the novel aspects of your study. However, I find these aspects only partly novel. The first aspect is not truly novel. Olonscheck and Notz, 2017 (Consistently Estimating Internal Climate Variability from Climate Model Simulations. *Journal of Climate*) distinguish changes in the variability of

winter and summer Arctic sea ice area. The second aspect is also touched by Olonscheck and Notz, 2017 but your study goes beyond this by investigating the underlying processes for the model-simulated changes in CESM-LE. However, the very recent study by Massonnet et al., 2018 (Arctic sea-ice change tied to its mean state through thermodynamic processes. Nature Climate Change) covers parts of your findings. I recommend to more clearly work out the novel aspect of your study, to distinguish your results from the mentioned studies, and to discuss your results in the context of their findings.

Thank you for pointing us to these very recent papers, which are relevant for and complementary to our study. The Massonnet et al. (2018) analysis only addressed sea ice volume variability and its projected future change, which is to be a monotonic decrease, unlike the more complicated time-dependent response of ice area variability that is our focus. Our findings help resolve the findings in Olonscheck and Notz (2017), whose analysis is of future changes in sea ice variability is much coarser temporally and seasonally (the 2006-2100 time block for JAS and JFM), and less conclusive: (a) “the future internal variability of summer Arctic sea ice area possibly increases. . .”, and (b) “. . . winter and summer Arctic sea ice area show inconsistent model responses” [in CMIP5]. They also identify a decrease in future sea ice volume variability, like Massonnet et al., but that is not directly pertinent to our findings.

We have added these two new studies to the introductory material, where we highlight previous findings on this topic and distinguished our focus from their work. In addition, we now compare and contrast our results with those of Olonscheck and Notz in the final section.

#7 ll. 147-148 I don't believe that the internal variability is robustly characterized from just one model. The internal variability largely differs between the CMIP5 models. How do we know that CESM-LE is representative? I assume you mean that 40 ensemble members allow to robustly quantify the internal variability WITHIN THAT MODEL, but I don't believe that your statement is correct as it is now. Please be more precise here.

We agree that the sentence needed refining, so we have modified it to address the reviewer's point. The sentence now reads, “We analyze a large 40-member ensemble from a single GCM, which allows us to isolate internal variability, which is otherwise muddled with inter-model variability in multi-model comparisons.”

#8 ll. 162 A medium ensemble of 15 members for RCP4.5 described in Sanderson et al., 2015 (A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario. Climatic Change) and recent ensembles, e.g., for RCP2.6 described in Sanderson et al., 2017 (Community climate simulations to assess avoided impacts in 1.5 and 2C futures. Earth System Dynamics) also exist.

We have added the two Sanderson et al. papers, as suggested.

#9 ll. 170-172 For two reasons, I am not convinced by the usefulness of this selection

criteria. First, because the threshold of 20-percent error seems arbitrary to me. How is this justified? Second and more importantly, there is no reason to believe that models that fit the observations comparatively well are better than others because of the large influence of internal variability. When taking model-specific internal variability into account the sea-ice simulations of most CMIP5 models are plausible. I would like to see whether or not your basic conclusions change when using the full set of CMIP5 models. Also, as a reader I would like to know which CMIP5 models you used without having to look this up in Wang and Overland, 2015.

The model selection criteria was introduced by Wang and Overland 2009 (GRL) and then gradually accepted by the community, not only for sea ice but also for other variables. We want to use trustworthy models, which can capture the corrected physics and dynamics. This is especially important for a sea-ice model, because we are dealing with the absolute value (the sea-ice extent) instead of anomalies (e.g. global mean temperature). The systematic biases in a model can significantly contaminate the ensemble means. Besides, as Massonet et al., (2018) pointed out, models far outside the observed base state should be omitted, because they won't capture the correct physics of thermodynamic processes in the future. Then how do we define "far outside", and which ones should be omitted? The 20% threshold was chosen because it can remove the outliers yet keep a reasonable number of models for a proper ensemble size. We also tested other selection criteria and determined that 20% is a good choice for this purpose. In IPCC AR5, this model selection method was adapted with a variation for the first time in IPCC history (IPCC, 2013, Chapter12). The 20% cut-off was also followed in their modified selection scheme.

Model names have been added into the text.

#10 ll. 176-177 To calculate the statistics for each of the 33 ensemble members and to then average them gives a biased estimate, because models with more ensemble members have a larger weight than models with only few (or even one) members. Again, I would like to see whether your basic conclusions would change when you always use e.g. three ensemble members from a model. As it is, I don't find the approach convincing.

We use no more than 5 ensemble members from each model, even if there are more ensemble members available, to avoid overweighting certain models. In addition, we tried to make the total ensemble number as close as possible to that used in the CESM-LE. This sentence is now added to the text. It is unfortunate that some models only provided one single realization. If we had only kept one member each, then the total ensemble numbers would be too small (12).

#11 ll. 212-217 I find the analysis of the CMIP5 models rather weak. To me, it is no proof that the variability is indeed increasing as shown by e.g., Goosse et al. 2009. This is because I see no logic behind simply averaging the CMIP5 models. As you write, the timing of ice retreat is very different in the different models, so averaging them will smooth out possible signals. For instance, one could normalize the timing of

sea ice retreat before doing the analysis. I think that more analysis of the robustness of the results based on the CMIP5 models is needed.

We appreciate this suggestion and recognize that there is no ideal way to assess the collective response of the CMIP5 models. We are not sure how to normalize the timing of sea ice retreat before doing the analysis. The paper acknowledges that multi-model averaging leads to smoothing, but this procedure is standard in analyses of CMIP5 (e. g., the IPCC AR5 report Chapters 9 to 12). The purpose for presenting the CMIP5 models is to supplement our more detailed analysis of CESM-LE by showing that their first-order features are similar and therefore that our major conclusions are not model-specific. That this resemblance emerges despite the timing differences among the models is evidence that our primary conclusions are robust.

#12 ll. 230-232 Related to the previous comment, I would like to know which of the two reasons is more relevant.

The discussion in the original manuscript is a little confusing, when we say “CMIP5 model spread could also be responsible for inflated variance as models diverge in their timing of the downward trend and its rate of decline”. We were trying to explain two different things in one sentence. Here is the revised portion:

“Near the end of the 21st century, the running standard deviation also shows an increase in the CMIP5 ensembles from December to June (Fig. 2), very similar behavior to that displayed by CESM-LE. However the magnitude of the increase in the running standard deviation in the CMIP5 ensemble mean is smaller than that in CESM-LE. This is not surprising, as the timing of ice retreat varies among models, so averaging them will smooth out the possible signals. The CMIP5 models therefore provide additional evidence that increased variability is associated with decreasing sea ice cover. “

#13 ll. 219-220 See again Olonscheck and Notz, 2017.

We have added comparison of our findings to Olonscheck and Notz in a couple of places, but their analysis of the difference in sea ice area variability averaged between entire time blocks (1850-2005 vs. 2006-2100) doesn't lend itself to a direct comparison with the time series presented in Figures 1 and 2.

#14 ll. 304-322 It is not very clear to me how exactly you calculate the thermodynamic and dynamic component. For instance, do you sum up top, basal and lateral melt for the thermodynamic melt component? I think I can guess what you did, but it is not written down precisely.

The thermodynamic component is a sum of these three terms, as we have clarified in the text.

#15 I. 373 I recommend one or two introductory sentences here to guide the reader. This would also help to improve the structure of the discussion section.

Good suggestion. We have added some introductory sentences that summarize the overall study and segue to the bullet points in the rest of the section describing our major findings.

#16 I. 443 This should be "projected", instead of "predicted".

Text has been changed to "projected."

#17 II. 448-449 I very much appreciate that your work includes the analysis of CMIP5 models. But I question that the presented analysis is sufficiently well done to justify this statement on robustness. Especially, because the CMIP5 models are only used for section 3.1 and not in the later sections that deal with the mechanisms. The questionable (see comment #11) and generally weak inverse relationship between variability and rate of retreat that you show for the CMIP5 models does NOT necessarily imply (and also does not suggest) that the same mechanisms are at work like the ones you describe for CESM-LE. This statement is too strong. I recommend to either extend your analysis of the mechanisms to the CMIP5 models (if possible) or further weaken or delete this statement.

We have softened this sentence by stating that the physical mechanisms identified in CESM "may apply more generally" [to CMIP5 models]. We agree with the reviewer that our conclusions derived from CESM don't demonstrate robustness across all models, but we think that other simulations should exhibit a similar non-linear (parabolic) ice variability response, based on the geographic dependencies described in Goosse et al. (2009) and Eisenmann (2010).

technical corrections:

I. 331 I prefer "the variability in the thermodynamic term", rather than "the thermodynamic term variability"

The text has been changed as suggested.

Figures: I suggest to make the figures look more consistently, i.e. Figures 3 and 6 like Figures 1, 2 and 4. Also, I find the different axis labeling in Figure 6a and 6c confusing. For Figure 6, a title for each panel would increase the readability and lines at 0 percent and 100 percent in panels a and c, too.

We have made these modification to Figure 6, though it is unclear otherwise how the figures look inconsistent. The former Fig. 3 and S1 were adjusted to have matching

sequential color schemes, while Fig. 6 is displaying the data in a different way than the other figures and is presented as best as possible. These three figures are included in this discussion comment.

I. 345 frazil = frazil ice?

Yes, frazil "ice" has been added.

References: Comiso et al ... The year of publication is missing. Zhao et al., 2018 ... This reference appears twice.

These two references have been corrected. Thanks for catching these oversights.

Anonymous Referee #2

This study by Mioduszewski and co-authors is concerned with the future (and to some degree past and present) variability of the Arctic sea ice cover in GCMs. The article focuses on seasonal aspects of the variability in sea ice area, and on potential drivers of such variability. The authors find a strong correlation between ice area variability and ice thickness, and argue that thermodynamic processes have a stronger impact on variability than dynamic processes.

The study is concerned with an important topic that fits well within the scope of The Cryosphere. I agree by-and-large with the comments of the other reviewer, and would hope to see some substantial revisions of the manuscript. Furthermore, several parts of the manuscript are marked by a somewhat disappointing standard of language and presentation, in particular given the experience and seniority of the co-authors. Below I will detail concerns that I have in addition to those voiced by the other reviewer.

General comments:

1. The abstract and introduction should be thoroughly revised (see specific comments below). The writing improves from Section 2 onward.

See our responses below to the reviewer's specific comments.

2. Please consider the geographic muting effect of Eisenman (2010) in more detail. i.e. what do analogues to Fig S1 and Fig 1 look like when using Eisenman's "equivalent ice extent"? This would help quantify the role that the distribution of land around the Arctic basin plays in this context.

We have attached this analysis to this discussion comment. The adjustment for geographic constraints does result in some of the expected changes, as there are steeper declines in some of the winter and spring months which result in greater peaks in their variability. However, it shouldn't change the main results shown in the monthly

time series of ice variability in Figure 1. Furthermore, while this alternative metric does produce much larger reductions in future ice cover during winter-spring, it's not clear how this result should affect our existing interpretations. It is a theoretical construct and our purpose is to assess the variability of ice cover that actually exists and that has practical implications for societal impacts such as marine navigation.

We believe it is worth noting how the calculation of equivalent ice extent compares with our analysis since this is a relatively well-known concept in the field and some others will likely have the same question, and have made these modifications to the discussion.

3. I share the concerns of the other reviewer in that the discussion of the CMIP5 analysis is somewhat vague and incomplete. It also should be put more clearly in context with other recent work on the subject.

Please see our responses to Reviewer 1 about their points #9 and 10, copied below:

The model selection criteria was introduced by Wang and Overland 2009 (GRL) and then gradually accepted by the community, not only for sea ice but also for other variables. We want to use trustworthy models, which can capture the corrected physics and dynamics. This is especially important for a sea-ice model, because we are dealing with the absolute value (the sea-ice extent) instead of anomalies (e.g. global mean temperature). The systematic biases in a model can significantly contaminate the ensemble means. Besides, as Massonet et al., (2018) pointed out, models far outside the observed base state should be omitted, because they won't capture the correct physics of thermodynamic processes in the future. Then how do we define "far outside", and which ones should be omitted? The 20% threshold was chosen because it can remove the outliers yet keep a reasonable number of models for a proper ensemble size. We also tested other selection criteria and determined that 20% is a good choice for this purpose. In IPCC AR5, this model selection method was adapted with a variation for the first time in IPCC history (IPCC, 2013, Chapter12). The 20% cut-off was also followed in their modified selection scheme.

We use no more than 5 ensemble members from each model, even if there are more ensemble members available, to avoid overweighting certain models. In addition, we tried to make the total ensemble number as close as possible to that used in the CESM-LE. This sentence is now added to the text. It is unfortunate that some models only provided one single realization. If we had only kept one member each, then the total ensemble numbers would be too small (12).

4. In Sec 3.3 it seems odd to choose Sept and Dec as months to study the roles of ice retreat and expansion, respectively. First, the sea ice minimum occurs typically in mid-Sept, which means that there's substantial ice expansion in the 2nd half of the month (as remarked in L.314). Thus, if the authors want a fully retreating month, why not choose August? December, on the other hand, is fairly early in the ice expansion

phase, so if the aim is to capture as much as possible of the preceding expansion, why not choose February? Or January? To that point, in the conclusions (L.413-414) the authors relate Fig 6c,d to "Nov-Jan" variability (rather than to Oct-Dec, as used in the analysis).

We understand why the reviewer questions our choices of representative months, but we tried to balance various considerations in selecting September and December. Because there has been so much interest in September sea ice coverage as the annual minimum and previous work on interannual variability in that month (e. g., Goosse et al. 2009, Swart et al. 2015), we felt that a focus on September would be of interest to readers. In addition, our choices for selected months had to strike a balance between conditions in the present climate---in which there is some ice expansion during late September---and future conditions, when September becomes ice-free and therefore no longer has any ice expansion.

December is of particular interest because that is a month of ice expansion exhibiting the distinct three-phase evolution of interannual ice variability described in the text: essentially flat, then a pronounced peak, followed by a decline (see Figure 1). By contrast, other ice-expansion months we could have chosen, such as January-March, do not show this three-phase evolution, and March has only a modest increase in interannual variability at the tail end of the simulation (Figure 1). A benefit of choosing September and December in Figure 6 is that the analysis sheds light on the physical mechanisms responsible for the three-phase evolution in months with very different thermodynamic and dynamic processes operating.

We're sorry about the confusion regarding "November-January" in the original lines 413-414. This range of months was in reference to the slightly lower mean thicknesses coinciding with the peak in interannual ice variability in these months, as shown in Figure 4. To clarify this point in the revised text, we have added that the behavior of Nov-Jan is taken to be explained by our findings for December in Figure 6c,d as a representative month.

5. As pointed out by the other reviewer, this work needs to be put carefully in context with the very recent paper by Massonnet et al ("Arctic sea-ice change tied to its mean state through thermodynamic processes", Nature Climate Change, 2018). I appreciate that the latter study was published after this one was submitted.

We thank both reviewers for bringing the Massonnet et al. study to our attention, and we point the our response above to comment #6 of Reviewer 1 for how we put our work into that context. While sharing thematic similarities with our paper, Massonnet et al. focused exclusively on sea ice volume, rather than ice area, which is the focus of our analysis. This distinction is very important, because Massonnet et al. found that the interannual variability of Arctic sea ice volume has already peaked and that it will decline in accordance with the thinning ice pack in the future. By contrast, our study reveals a more complicated behavior in the future interannual variability of sea ice area, such that it exhibits a two-to-three phase evolution in each month: relatively steady during the

thick-ice regime of the past, a pronounced increase when the ice packs thins sufficiently, and then a decline if the ice cover diminishes sufficiently in a particular month (July-December).

Specific comments:

I would suggest moving Fig S1 to the main text as Fig 1. I'd also suggest color-coding the different months sequentially in this figure so that the seasonal cycle becomes more visible.

This is a good suggestion. We have incorporated this figure into the text as Fig. 1 and changed the color scheme as suggested, which now matches Fig. 3.

L.49: I would suggest deleting "independent".

We appreciate the suggestion but feel that this term is a useful reminder to readers that the differences among the 40 CESM-LE realizations express purely internal variability within the climate system.

L.52-54: This sentence is somewhat confusing. Some months see an essentially monotonic increase and it's not immediately clear what part of the sentence refers to CESM-LE and what to CMIP5. I suggest rephrasing and/or splitting into 2 sentences.

We realize that this sentence may be ambiguous, so we have reworded into two sentences. The text now states, "Both CESM-LE and CMIP5 models project that ice area variability will indeed grow substantially, but not monotonically in every month. There is also a strong seasonal dependence in the magnitude and timing of future variability increases that is robust among CESM ensemble members."

L.55: "inversely" correlated. This is used at several points in the ms. Unless I'm mistaken, isn't the rate of retreat "directly" correlated with the variability? In other words, the larger the rate of retreat, the larger the variability (?). At L.428 the authors talk about the rate of change. Here I can see the inverse relation: the more negative the rate of change, the larger the variability.

Good point. The reviewer is correct that the variability is directly correlated with the ice retreat rate, so the text has been modified accordingly. We had been thinking in terms of the rate of change in the ice area, which is inversely correlated with the variability.

L.58: "...indicating that [for most of the years (?)] substantial future thinning ..."

We think that a qualifier is unnecessary in this sentence, because we are really just making a general statement that the ice pack needs to thin substantially before a peak in ice variability can be expected.

L.59-60 "... depends on the season, primarily due to whether ..." This could be written

more clearly.

That phrasing has been reworded to “. . . depends on the season, especially whether. . .
“.

L.98 "... reduces the [mean] thickness of the basin ice back ..."

Sentence changed as suggested

L.100 "... the [estimated] negative trend ..."

Sentence changed as suggested

L.103 "[Output from] many climate models suggest[s] ..."

Sentence changed as suggested

L.113-114 rephrase

The sentence has been rephrased as follows: "Nonetheless, navigation through the Arctic has already increased in frequency as a result of this decline (Melia 2016; Eguíluz et al. 2016), and even more trade routes associated with the increased ice-free season are expected by the end of the 21st century (Aksenov et al. 2015; Stephenson and Smith 2013)."

L.117 I agree with the other reviewer that the implied causality between reduced extent and loss of multi-year ice is misleading.

As noted above, we have rephrased this sentence: "As the Arctic sea ice pack thins and retreats, multi-year ice is being lost and there is consequently a larger proportion of seasonal, first year ice."

L.118 "Increased thin ice ...". Replace with "Overall thinner ice ..."

Sentence changed as suggested

L.121 "... ice growth and retreat rates ..." I'd argue this should either be "expansion and retreat" or "growth and melt"

The sentence has been changed to, “. . . , at least partially due to enhanced ice growth and melt.”

L.129 "relationship between ice area and its variability". Do the authors mean the "mean ice area" and the "variability in ice area"?

The wording of this topic sentence has been revised for simplicity and to accommodate the broader investigation of sea ice volume variability cited later in the paragraph. This sentence has become, "Changes in the interannual variability of sea ice have been studied only in a limited capacity, likely because they are only beginning to become visible in September in the present day."

L.130 "... it is only beginning to become visible ..." The relationship is becoming visible? Does this mean that we are starting to observe a correlation between the mean ice area and the variability of ice area? Please clarify.

See the comment above.

L.143 "... monthly differences are [societally/economically?] important ..."

We have altered the text to read "societally important"

L.148 "... characterize internal variability [of CESM]" (see other reviewer's comment)

This sentence has been modified as suggested: "We analyze a large 40-member ensemble from a single GCM, which allows us to isolate internal variability, which is otherwise muddled with inter-model variability in multi-model comparisons."

L.198 "... follows [an analogous] three-phase progression ..."

Sentence changed as suggested

L.222 "inverse" see comment at L.55

Sentence changed as described above

Fig.3 I find it hard to decipher the individual curves here. What about splitting the figure into 2 panels, with panel (a) showing spring/summer months and (b) showing fall/winter months. The missing curves in each panel could shown as faint gray in the background for reference. Again, I would use a sequential color map.

We have modified Figure 3 by using a sequential color map to help readers discern the individual monthly curves, while still allowing for a direct comparison of the curves on a single graph.

L.240 "... between ice thickness and [ice area] variability ... " Otherwise it might be read as "ice thickness and ice thickness variability"

Sentence changed as suggested

L.267-272 Would it be worth showing another thickness curve (<0.2 m) in Fig 4 to illustrate the phase dependence (and different area coverage) for different ice thicknesses?

We appreciate this suggestion and have tried overlaying another thickness curve on Figure 4, but this ended up cluttering the figure and distracted from the main point we're making with these graphs.

Fig 5: The left hand side of the bounding boxes was cut off. Also, the resolution of the figure was low (jpg? Better to use png with resolution > 150 dpi). It'd be nice to add the respective decade in the top left corner of each panel.

We have resized Figure 5, so that the left-hand side doesn't get truncated during the online pdf conversion of figures, and we added the decade to each panel.

L.275: Please mention CESM-LE in the caption.

Caption changed as suggested

L.287: "... thin ice and [the variability of] inter-annual ice coverage ... "

Sentence changed as suggested

L.280-282: How much of this difference is simply due to the limited run length of the simulations? In other words, once the ice retreats further in winter and spring after 2100, would one then also see the horse-shoe pattern in those seasons? Conversely, in the summer, are the regions of high variability restricted to the Arctic boundaries during earlier decades in the simulations?

Presumably, the horse-shoe pattern of maximum ice variability would shift to the winter-spring months after 2100 if the model had been run longer, but we can't prove that with the existing simulations. We have confirmed that the regions of highest variability during summer in the earlier decades of the simulation occur along the periphery of the ice pack, where the thinnest ice exists.

L.293 "expanding" rather than growing

Sentence changed as suggested

L.304 Why is a different method used here to calculate the standard deviation?

We have clarified that the method is not actually different, but the decadal average of the running standard deviation is used in Figure 5, which results in a slightly lower amplitude of standard deviation when comparing to the time series.

L.337-340 delete "rather than melt". Split into two sentences?

As suggested, we have deleted that phrase and split into two sentences: "From the 20th century well into the 21st century, ice growth occurs in the October-December period in a similar region of maximum interannual variability as September, except slightly equatorward (Fig. S2b). Ice export plays a relatively larger role in the regions of interest in December than in September (Fig. 6c)."

L.342 "[mid] 21st century"

Sentence has been changed to "early-mid 21st century".

L.345 "frazil [ice]"

Sentence changed as suggested

Fig 6: The (a)-(d) labels are too big and bold, and the rest of the text in the figure is too small.

This figure has been modified to improve readability.

L.376 Isn't the smaller magnitude of spring variability just a result of the time series ending in 2100 (before the ice edge retreats into the Arctic basin in spring)?

Presumably, although we can't prove that. We have modified the sentence by adding "by the time the simulation ends in 2100" to avoid implying that future variability during spring will necessarily be less than in other seasons.

L.415 "... ice area variability [in winter] also coincides ..."

The sentence has been altered nearly as suggested, but we use the phrase "in these months" to refer to the Nov-Jan period cited in the previous sentence (since November isn't a winter month).