nature portfolio

Peer Review File

Arctic sea-ice loss drives a strong regional atmospheric response over the North Pacific and North Atlantic on decadal scales

Corresponding Author: Dr Ivana Cvijanovic

This file contains all editorial decision letters in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Decision Letter:

** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors **

Dear Professor Cvijanovic,

Your manuscript titled "Near-term impacts of Arctic sea-ice loss" has now been seen by 3 reviewers, whose comments are appended below. You will see that they find your work of some potential interest. However, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version that fully addresses these serious concerns.

We hope you will find the reviewers' comments useful as you decide how to proceed. Should additional work allow you to
• address these criticisms (that is, either to incorporate the suggestions or provide a compelling argument why the point made by the reviewer is not valid, or relevant to the editorial threshold as outlined below)
AND
• meet our editorial thresholds as outlined below,

then we would be happy to look at a substantially revised manuscript.

In the following, we list our main editorial concerns.

****
Editorial threshold 1: Present compelling new insights on the near-term response of Northern Hemisphere climate to Arctic sea ice loss.

Editorial threshold 2: Fully integrate your findings in the existing literature (as outlined by the reviewers) and discuss differences and similarities as well as the potential role of internal variability on the impact of sea ice loss on atmospheric circulation.

Editorial threshold 3: Acknowledge caveats and uncertainties appropriately, and clearly communicate levels of statistical significance.

****

If the revision process takes significantly longer than three months, we will be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We understand that due to the current global situation, the time required for revision may be longer than usual. We would appreciate it if you could keep us informed about an estimated timescale for resubmission, to facilitate our planning. Of course, if you are unable to estimate, we are happy to accommodate necessary extensions nevertheless.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter) and any completed checklist:

https://mts-commsenv.nature.com/cgi-bin/main.plex

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Jan Lenaerts, PhD
Communications Earth & Environment
orcid.org/0000-0003-4309-4011

Heike Langenberg, PhD
Chief Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements

For your information, you can find some guidance regarding format requirements summarized on the following checklist: (https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf) and formatting guide (https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Review of "Near-term impacts of Arctic sea-ice loss" by Cvijanovic et al.

This paper explores the near-term transient response to Arctic sea ice loss imposed in three coupled models, using an albedo-perturbation method. The authors identify a weakening of the Aleutian low, and decrease in rainfall over the southwest US, concluding that this is a response that we can expect, and prepare for, in the coming decades. I have many problems with that paper, the main one being the fact that the authors do not properly recognize the existing literature on that subject, omitting several important studies that do not align with their results. Also, the sea ice forcing and physical mechanisms behind the teleconnections they identify are poorly described, and I am not sure how robust the results are, given the often low statistical significance of the anomalies. This needs much more work to be a convincing paper on an already very crowded topic. Please find hereafter a list of comments.

1) My first problem with that study is that several important papers on this very topic are not cited, and/or are not properly discussed to interpret the results and put them in perspective with what has been done previously. The study is presented as if it was the first time the transient, or near-term, response to Arctic sea ice loss is investigated (it's not), and as if the methodology used to melt the ice is novel (it's not, and like all methods, it has its own limitations).

For example, l.54, you say that"… far less attention has been given to the near-term, decadal-scale impacts" of sea ice loss. First, there is a whole Model Intercomparison Project (MIP), the Polar Amplification MIP, dedicated to understand the impact of +2°C Arctic and Antarctic sea ice loss (Smith et al. 2019). See Peings et al. (2021) for one of the first papers on these simulations, that shows the response to +2°C sea ice loss (i.e near-term) in both atmosphere-only and short coupled runs. Their coupled runs cannot really be compared to yours though, because they are 14-month runs that do not account for longer-term adjustment of the ocean. But PAMIP also includes long coupled runs in which the longer-term transient response can be investigated. Several modeling centers have already run those and some of their results should be published soon. It would be fair to mention PAMIP and the associated research because it aims exactly to answer the questions you are exploring in your paper, in a coordinated manner so different model results are more easily comparable.

Second, the transient response to sea ice loss has also been explored in Sun et al. (2018). In particular, they look at the response to RCP8.5 projected changes over the 2021-2050 period, i.e. near-term changes, like you. They do not find any significant circulation response over the North Pacific (see Z500 anomalies in their Fig. 8). The sea-level pressure response does not quite resemble the one in your three models, it looks like a wave-1 pattern that has been found in other coupled models (Screen et al. 2018). The precipitation response over the western US is also different, it looks like a north-south wet-dry dipole, but very little statistical significance.

Screen et al. (2018) is another very important paper on the response to sea ice loss in coupled models, I am surprised it is not cited. It is true there is whole range of different responses in model studies in the literature, but it tempers a little your statement of l. 56 about contradictory findings regarding the remote impacts of Arctic sea ice loss. Their multi-model study shows rather good agreement in the response to sea ice loss in coupled models.

Blackport and Kushner (2018) is cited in introduction, but not discussed later, while this is a paper that explores the response to Arctic sea ice loss in transient coupled runs too. And they also use an albedo perturbation method to melt the ice. Their results do not resemble the response you get (they are in fact opposite) : weakening of the Aleutian low, and increased precipitation over the US West coast. I am not saying that their results are more trustworthy than yours, but it illustrates the uncertainties in looking at remote regional impacts of Arctic sea ice loss.

Liu and Fedorow (2019) is another study looking at the transient response to Arctic sea ice loss in a coupled model. They also use a method to melt the ice that looks very similar to yours (albedo perturbation through alteration of ice/snow parameters). They look at the short-term (1-15 years) and longer (151-200) global response to Arctic sea ice decline. I do not know what value to give it, but they find an increase in precipitation over the Western US (their Fig. 2). Plus they also show a clear role for the AMOC decline in driving the response to sea ice loss, an aspect that is overlooked in your study, especially when you say the response is not sensitive to the fact of using a slab (no dynamics) versus coupled ocean, for example l. 185 : “employing the slab ocean instead of a fully coupled dynamic-thermodynamic ocean, does not substantially affect the simulated response to sea-ice changes on decadal timescales.”. This is in contradiction with pretty much all the studies I mention above.

2) This leads me to my second comment, and how put you put your results in perspective with previous findings in the literature. l. 280-283, you just discard previous studies that found a different response over the North Pacific (deepening of the Aleutian Low, versus weakening in your case) by saying they were looking at the long-term (one century or more) response to Arctic sea ice loss. This is simply not true. And you omit to cite them, intentionally or not (I hope not). Sun et al. (2018) and Liu and Fedorov (2019) are two of them, they look at the near-term transient response to Arctic sea ice loss. They may be more I’m missing. And do not see why the atmospheric response in the North Pacific would flip sign in the near-term vs the long-term, so the long-term argument to explain different results in other studies is questionable. I would argue that your results can also be compared to Blackport and Kushner (2018), Screen et al. (2018), and more that I may omit here. The North Pacific sea-level pressure response you get in these runs does not nearly resemble the SLP response that is commonly found in this type of experiments. See Fig. 5 of Blackport and Kushner (2018), Fig. 2 of Screen et al. (2018), Fig. 8 of Sun et al. (2019). Liu and Fedorov (2019) do not show SLP anomalies but their Fig. 2 does not show a drying over the Southwest US. In summary, this study is just one among others, that happen to find a seemingly consistent response in the North Pacific (but statistical significance is low, see one my following comment). However the overall response does not quite align with previous findings. Why is the SLP response so different from the wave-1 pattern found in previous studies, for which there seems to be more and more consensus that it is a robust feature of the atmospheric response to sea ice loss (Screen et al. 2018) ? I do not buy the argument that this is because previous studies have looked at the centennial response, or used non-conservative energy methods (see my next comment), because as I said, this is inaccurate.

3) Another argument you make to explain why previous studies may have found different results is that they were using “non-energy conserving” experimental setups. Again, not true. Many other studies have used a method similar to yours. See Sun et al. (2020) for a review of different approaches that people have used to carry out these sea ice loss experiments in coupled models. Like said in that paper, albedo reduction conserves energy but is only effective during the sun-lit portion of the year. By contrast, ghost flux and ice-flux nudging do not conserve energy but are effective in adjusting sea ice year-round. This is not discussed anywhere in your paper, and it sounds like your are coming up with a whole new method that nobody have thought about before. How do you know that your method for melting the ice gives you more trustworthy results than other ones ? Blackport and Screen (2019) suggest that most of the influence of sea ice on the winter climate comes in fact from the winter sea ice anomalies, not the autumn ones. This suggests that albedo-perturbation sea ice loss experiments may miss some important aspects of the response to sea ice loss, because they do not melt enough ice in winter. This is the case in your experiments, they do not have a great deal of ice melt in winter (Fig. 1a-c).

4) An important information that is missing is how much sea ice loss do your simulations represent. How does your sea ice loss forcing compare with previous studies ? It is a forcing that we may expect in the next decades, as projected by coupled climate scenario simulations ? Is it comparable with the near-term sea ice loss forcing created for the PAMIP exercise ? This is never discussed, despite being a very important question.

5) The physical mechanisms behind the atmospheric responses are not clearly identified. There are assumptions that the high-latitude forcing influences the tropics through mid-latitude eddy momentum flux changes and the tropical WES

feedback, mostly based on results from previous studies. But the authors do not show what is exactly happening in their runs. For example, what is the response of AMOC in their runs (Sun et al. 2019, Liu and Fedorov 2019) ? How do physical mechanisms align with previous findings in the literature ?

6) Statistical significance of the results. From the description of the experiments, I get that each of your experiment represents ~200 years of data, i.e. ~200 fall/winter seasons to analyze. Is that correct ? This should be mentioned. Internal variability is a huge factor that can explain model differences, that is not even mentioned a single time in your paper. It is more and more recognized that a lot of conflicting results between different modeling studies, as well as results from observational analyzes, can simply arise from internal variability, and the lack of robustness in the statistics or model responses that are analyzed. See Peings et al. (2021) for an illustration of how internal variability alone can create substantial differences in the remote response to a same sea ice loss forcing. I wonder how robust your results are here. 200 realizations to analyze is a fair amount, but many of the signals you discuss are at the edge of statistical significance. For example, the SLP anomalies in the North Pacific are not significant in IPSL-CM5A2 (Fig. 2f), as are the OLR anomalies of Fig. 4f for this model). Are those robust signals enough to be trusted ? Especially you use a 90% confidence level in some cases, I suspect to retain enough statistical significant anomalies, but this is quite a low restriction. Plus it is always a good idea to account for spatial correlation when assessing statistical significance of a field, using the False Discovery Rate of Wilks (2016). I suspect using this test may remove some of the “significant” anomalies you get, especially in the fields for which you only use a 90% confidence level threshold.

References

Blackport, R., & Kushner, P. J. (2017). Isolating the Atmospheric Circulation Response to Arctic Sea Ice Loss in the Coupled Climate System, Journal of Climate, 30(6), 2163-2185.

Blackport, R., &Screen, J. A. (2019) Influence of Arctic sea ice loss in autumn compared to that in winter on the atmospheric circulation. Geophysical Research Letters, 46, 2213-2221.https://doi.org/10.1029/2018GL081469

Liu, W., & Fedorov, A. V. (2019). Global impacts of Arctic sea ice loss mediated by the Atlantic meridional overturning circulation. Geophysical Research Letters, 46, 944– 952. https://doi.org/10.1029/2018GL080602

Peings Y., Z.M. Labe and G. Magnusdottir (2021) Are 100 ensemble members enough to capture the remote atmospheric response to +2°C Arctic sea ice loss? Journal of Climate, DOI:10.1175/JCLI-D-20-0613.1

Smith, D. M., Screen, J. A., Deser, C., Cohen, J., Fyfe, J. C., García-Serrano, J., Jung, T., Kattsov, V., Matei, D., Msadek, R., Peings, Y., Sigmond, M., Ukita, J., Yoon, J.-H., and Zhang, X. (2019) The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification, Geosci. Model Dev., 12, 1139–1164, https://doi.org/10.5194/gmd-12-1139-2019.

Sun L., Alexander, M., & Deser, C. (2018) Evolution of the Global Coupled Climate Response to Arctic Sea Ice Loss during 1990–2090 and Its Contribution to Climate Change, Journal of Climate, 31(19), 7823-7843.

Wilks, D. S. (2016). “The Stippling Shows Statistically Significant Grid Points”: How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It, Bulletin of the American Meteorological Society, 97(12), 2263-2273.

Reviewer #2 (Remarks to the Author):

Summary:

The paper provides results from 3 global climate models that are perturbed to extract the “sea-ice” signal on northern hemisphere atmospheric circulation characteristics. Perturbation experiments are conducted by changing model parameters to achieve substantial (summer/fall) sea ice reductions relative to their control experiments. The impact on northern hemisphere circulation features (height, slp, and precipitation) is presented. An analysis of (some) of the dynamical mechanisms underlying the proposed teleconnections is given. Results are also compared to observed (reanalysis) changes in northern and similarities between simulated responses to sea ice loss and changes over the period 2009-2018 are postulated. The results show a “shared” weakening of the Aleutian Low, ridging over the North Pacific and associated drier conditions in the southwestern United States. Wetter conditions are found over the Mediterranean.

Review:
The paper is generally well-written and easy to follow and potentially contributes to an increasing body of research on this important topic. As somebody who has followed this discussion over the past 20 years and is suffering from a fair amount of intellectual whiplash as a result, I had some hope that the paper might illuminate the topic in a new way and yield some clarity to my state of confusion. It may not be the authors fault, but this clearly didn’t happen. I am as confused as I have been. But science progresses mostly incrementally and the point at which new insights emerge from the labor of hundreds of studies, is hard to predict. As such, with some solidified and/or reduced claims, I believe the paper does represent yet another contribution to this process.
I think the authors could substantially improve the impact of their paper if they tried to put this into the context of currently

ongoing efforts to sort out this problem through coordinated experiments, specifically the CMIP6 Polar Amplification Modeling Experiment (PAMIP, Smith et al. 2019). This group has laid out a way to systematically ask questions about the importance of different models, configuration, perturbations etc. If the authors could find ways to place their experiments within that framework, it might give their simulation more context and assure that results are considered in the synthesis that will hopefully emerge from PAMIP. The present study, using multiple models (I believe all of which are already participating in PAMIP) already makes a good move in this direction. I understand that participation and coordination with other groups is not always possible and separate tracks of investigation should be possible, but I think the authors should give this some thought.

I have several major points of criticism that think the Authors should consider:
1. The overall point of differentiation from previous studies for the paper is given as “decadal” versus “centennial” for previous studies. If find this distinction not well explained. The perturbation experiments are not tied to any time scale. The “summer ice free conditions” as achieved in two of the models are still many decades away in others. They depend on the sea ice sensitivity of the particular model so tying this to a particular timescale is confusing. The comparison of the sea ice impact with respect to the last decade 2009-2018 relative to 1979-2008 is indeed decadal but I find this section actually to be the weakest part of the paper.

2. The analysis is in large parts based on the subjective assessment of similarity of circulation patterns between different models (and ERA-5). Sometimes the similarity is labeled as “remarkable”. When pointed to specific features (e.g. the North Pacific High, Aleutian Low), I can recognize some of the similarities. But a less guided inspection could yield a very different conclusion: For example, SLP anomalies over the Arctic Ocean are vastly different. With EC-Earth showing a low, IPSL a high, and CCMS4 a bit of both? This makes me wish for a more objective measure of similarity or dissimilarity. A pattern correlation method comes to mind. Also, atmospheric circulation patterns with high spatial auto-correlation have a tendency to look meaningful, so I think a comparison should be tied to a reference to determine what constitutes a meaningful difference (level of significance). How does a difference between models compare to a random perturbation? The literature on “field significance” maybe a useful guide. Similarity measures for pattern differences between intra ensemble members (which constitute random variation) might be another way to develop a reference.

3. The explanation of the link between the high latitude perturbation and the hemispheric response seems truncated. The paper shows some credible analysis demonstrating the link from the tropics to the high(er) latitudes but the initial branch from the sea ice perturbation to the tropics is left to some references. Given the what others have termed the “pre-consensus” state of affairs in this space, I think it would be helpful if analysis specific to the model experiments presented here could confirm how the initial sea ice perturbation does propagate into the tropics. It seems particularly important also to identify which regions of sea ice loss dominate the lower latitude impact. For the comparison with the observed response, I note that the regions showing ice loss in the models have only relatively minor losses in the observations over the 2009-2018. In addition, as stated above, the Arctic Ocean, from where the lower latitude changes are supposed to be propagating shows rather different SLP anomaly patterns.

Lesser points:
I hope the figure size will be increased in production. They are very hard to see for aging eyes.

A note to what degree the experiments constitute a “sea ice only” vs. a “sea ice + SST” response might be helpful. Since SST at high latitudes can respond to the reduced sea ice by warming, it appears that the experiments show a combined SIC/SST response?

Some discussion putting the magnitude of the response into context would be useful. Notably, the SLP response is given in Pa for the model results, while in hPa for the observations (looks like a factor of 30 difference judging from the color scale). The response in terms of annual precipitation amounts (%winter or annual?) would put the chain of responses into something more tangible (eyeballing form some papers it looks like the response for California is about 10% of winter precip for example)

Appreciating the differences between models could be improved if results from other models could be over plotted in Fig 1.

Line 112 “ the most prominent”
See earlier comment on making such statements more objective
Line 136: “anomalous propagation of wave activity”
I don’t see this for EC-Earth?

Line 138: “Previous work has suggested this Rossby wave pattern”
The authors might want to cite additional literature on this (e.g. Ding et al. 2019, Gong et al. 2020). In particular, it would be helpful to put the present result into the context of work that has argued that atmospheric circulation variability in the Arctic is the response to internal SST variability in the tropical Pacific. Some words on the Author’s perspective on how their results fit into the debate for tropical vs. Arctic forcing might be useful. Seems like it doesn’t have to either/or and the present work lends some credence to the idea that those local vs remote processes could be reinforcing each other?

Line 186.
That’s a nice result!

Line 200, “Atlantic sector is featuring the largest differences”

Looks to me that there are multiple regions with large differences (e.g. Arctic Ocean). Goes to prior comment on more objective measures of similarity.

Line 213 "…Atlantic region is additionally shaped by AMOC"
Isn't this a bit of a stretch from looking at two coupled models vs. one swampy one?

Line 217 " Wetter conditions are consistently found over the western Mediterranean"
True, but there are lots of blues and reds on that map, why discuss this?

Line 225 "Recent Decade"
I understand the motivation to try to link modeled sea ice response signatures to the real world, but I find the whole section not very convincing. First, for average NH ice concentration the period from 2009-2018 hasn't been one of rapid loss (e.g. Baxter et al. 2019 , Francis et al. 2020). Moreover, regions that show sea ice loss in the models show relative weak sea ice losses in the observations (Fig 5a, 1 d,e). (This also emphasizes further the need to identify specific regions and mechanisms of how the sea ice lower latitude coupling is working in the presented simulations)
Nevertheless, despite the relatively small perturbation in the real world, the atmospheric response supposed is at least an order of magnitude larger than the modeled response. I think this section need to be reworked substantially or maybe cut altogether. The argument that the discrepancy in magnitude is expected because other factors are at work is of course true, but also a bit of an unfalsifiable "catch all". A more objective "detection" technique is probably needed to make this more convincing. I would be surprised if this were possible to show. While cutting this section may feel like a loss, I think a modeling paper by itself could be a useful contribution. A workable alternative would be to heavily caveat this section.

Supplement:
Fig 1 missing panel labels a-i. Anomaly reference period should be stated in caption.

I hope you will find these comments constructive even though they may appear very critical. I recognize that there is a lot of work behind the simulations and analysis the paper will likely make a solid contribution after some of the criticisms are addressed.

References:

Smith, D. M., J. A. Screen, C. Deser, J. Cohen, J. C. Fyfe, J. García-Serrano, T. Jung, V. Kattsov, D. Matei, R. Msadek, Y. Peings, M. Sigmond, J. Ukita, J. H. Yoon and X. Zhang (2019). "The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification." Geosci. Model Dev. 12(3): 1139-1164.

Francis, J. A. and B. Wu (2020). "Why has no new record-minimum Arctic sea-ice extent occurred since September 2012?" Environmental Research Letters 15(11): 114034.

Ding, Q., A. Schweiger, M. L'Heureux, E. J. Steig, D. S. Battisti, N. C. Johnson, E. Blanchard-Wrigglesworth, S. Po-Chedley, Q. Zhang, K. Harnos, M. Bushuk, B. Markle and I. Baxter (2019). "Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations." Nature Geoscience 12: 28-33.

Gong, T., S. B. Feldstein and S. Lee (2020). "Rossby Wave Propagation from the Arctic into the Midlatitudes: Does It Arise from In Situ Latent Heating or a Trans-Arctic Wave Train?" Journal of Climate 33(9): 3619-3633.

Baxter, I., Q. Ding, A. Schweiger, M. L'Heureux, S. Baxter, T. Wang, Q. Zhang, K. Harnos, B. Markle, D. Topal and J. Lu (2019). "How Tropical Pacific Surface Cooling Contributed to Accelerated Sea Ice Melt from 2007 to 2012 as Ice Is Thinned by Anthropogenic Forcing." Journal of Climate 32(24): 8583-8602.

Reviewer #3 (Remarks to the Author):

Review of Cvijanovic et al. 2021

Summary
The authors revisit the now frequently discussed (but largely unresolved) topic of hemispheric-scale atmospheric circulation and teleconnections to recent sea ice loss. In this study, the authors use an energy-conserving method to tackle this question in a way that allows the ocean to respond dynamically in remote regions. This approach is substantially different from prior investigations in the literature, and comes to substantially different conclusions. Here, the authors find a strong hemispheric scale response to Arctic sea ice loss that includes a weakening of the semi-permanent Gulf of Alaska low and a negative NAO-like response in the North Atlantic—modestly reducing winter precipitation in California and parts of the Southwestern US and increasing in across the Western Mediterranean. The findings are nearly opposite of those in the existing literature regarding the response to ice loss and/or climate change more broadly, and the authors discuss possible reasons for this departure—primarily focusing on a 2-step teleconnection involving tropical SST modulation and the generation of a basin-scale Rossby wave train.

Overall Assessment
As someone who has followed the "Arctic sea ice-Mid-latitude weather" debate closely over the past decade, my impression

is that this paper is topically well-targeted and offers an important new perspective (and methodology), as well as some provocative new findings. Thus, from a novelty and scope perspective, this work is well-suited to CE&E. However, there remain some important outstanding questions regarding the broader conversation to which this work contributes--as well as the context of the findings that contrast with prior work by other authors-- that should be more directly and substantively addressed than in the present draft. Additionally, while the authors focus on winter teleconnections, the primary ice loss in the Arctic is actually occurring in autumn—so some additional analysis regarding the simulated response in other calendar seasons (especially autumn) would be warranted to more clearly illustrate the seasonal progression and any potential seasonal asymmetries of the atmospheric response to ice loss.

Altogether, I am optimistic that this paper can become an excellent candidate for publication in CE&E following some substantial revisions (more specific feedback enumerated below).

Specific Comments

1. The present title is probably not specific/descriptive enough. One possible suggestion could be something like: "Near-term sea ice loss drives strong regional atmospheric response over North Pacific and North Atlantic."

2a. One key weakness of the manuscript in its current form is the relative lack of context into which these results, which are effectively an idealized sensitivity experiment testing the atmospheric response to sea ice loss in the absence of any other external/internal forcings, are placed. The authors point out that the simulated response yields anomalous North Pacific ridging (and Atlantic troughing) that reduces American Southwest (and increases Western Med) precipitation in winter—which is exactly the opposite response demonstrated in some prior sea ice sensitivity experiments as well as being contrary to the overall net response to greenhouse forcing in CMIP5-type ensembles (which show slight precipitation increases in California and stronger precipitation decreases in the Mediterranean in winter).

I believe the authors already make a sufficiently plausible argument why their results here are contrary to other sea ice sensitivity experiments (a more comprehensive level of atmosphere-ocean coupling; near-term vs long-term response, etc.). But the authors barely even mention the broader aspects of their findings in the context of the full forced response to GHG warming in multi-model ensembles. In the real world, Arctic sea ice loss is not occurring in isolation—CO_2 is also rising, and the oceans are warming more broadly (among other changes).

I think it would be very important for the authors to discuss this context, especially as it relates to the fact the even if the modeled response in this experiment turns out to be 100% correct, it is possible that the net response will be overwhelmed by other factors. This is bolstered by the facts that 1) while precipitation decreases in the present experiment are statistically robust, they are generally fairly small in magnitude, 2) that prior research has demonstrated a mean “El Nino”-like response of the tropical Pacific to CO_2 forcings that results in opposite teleconnections to the ones demonstrated here, (and 3) the much discussed idea that there is/will be an Arctic-tropical “tug of war” in a warming climate (e.g., Barnes et al. 2015). Perhaps the authors here are suggesting that the near-term balance is more likely to be won by the Arctic than the tropics, at least on a regional basis? If so, that kind of explicit framing would be useful. (In addition to discussion in the main text, this context would also be very helpful in a revised abstract.)

2b. An additional question that comes to mind: is the effect of ice loss (particularly across North Pacific) stronger in weak ENSO vs strong ENSO (either warm or cool phase) years, which are known to have strong remote effects on both North Pacific and North Atlantic circulation? This is especially relevant given that the proposed mechanism operates via remotely SST-forced Rossby wave propagation out of the tropics--similar to atmospheric teleconnections from ENSO.

Also, could this emerging evidence regarding the potentially substantial role of sea ice loss on regional hydroclimate along the North American Pacific coast comprise a new seasonal predictive factor comparable to existing predictors like ENSO? Might one expect the effect of ice loss to be more pronounced in weak ENSO conditions, and muted during strong El Nino events? If so, that could potentially offer increased seasonal predictability if the state of sea ice can be known. I raise 2b mainly as “food for thought,” but I would hope to at least see some kind of substantive response to 2a.

3. Despite the analysis here and even after reading the supplementary discussion, I am still not really clear on the hypothesized physical linkage between the actual sea ice loss and its modulation of tropical SSTs. I understand how the atmospheric response develops once the tropical SSTs are perturbed, but I still don’t understand the actual means by which the sea ice loss signal moves from the Arctic to the tropics.

4. Recent work on the “Arctic/mid-latitude” question often hinges on the existence and/or accurate simulation of the “stratospheric pathway” (whereby vertically propagation of energy triggered by sea ice loss can disrupt the stratospheric and subsequently the tropospheric polar vortex, e.g., (Hoshi et al. 2019). The authors don’t really mention this literature or its potential relevance to the present study—which I think is a conspicuous absence. (And perhaps this could be related to #3 above?)

5. Line 306: California mean annual average precipitation projections are actually clustered pretty close to zero across most multi-model ensembles (despite large projected increases in precipitation variability and seasonality (e.g., Swain et al. 2018, Dong et al. 2019))—I have no idea where these numbers are coming from. This relative lack of change in regional precip has actually been suggested to be the net result of a “tug of war” by competing wetting/drying influences—which perhaps could account for either the muted mean signal, the projected increase in variability/seasonality, or both?

6. Potentially of note: the spatial pattern of GPH change in Fig 5c (West Coast US ridging but Hudson Bay low anomaly) is

reminiscent of the “North American Winter Dipole” pattern that has increased in frequency/intensity during the recent period of accelerated Arctic sea ice loss (Wang et al. 2015, Singh et al. 2016).

7. The greatest absolute/relative magnitude of Arctic sea ice loss is actually in autumn, rather than winter, as the authors point out. But the authors focus on winter atmospheric circulation. Given the temporal lag here, and the growing interest in California/Southwestern U.S. seasonal drying, it would be extremely interesting to see a seasonal progression of GPH500 anomalies and/or precipitation anomalies from these ice loss experiments (especially for Sep-Nov). Might this explain why California is now experiencing a robust autumn drying trend (but not winter drying trend) in recent decades? (Luković et al. 2021) It’s possible that even if the winter drying trend caused by sea ice is ultimately not overwhelmed by competing forcings, it could still be relevant in the autumn (with major implications for wildfire risk, etc.)

8. Line 574: reference error.

References
Barnes, E. A., & Screen, J. A. (2015). The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it? WIREs Climate Change, 6(3), 277-286. https://doi.org/10.1002/wcc.337. https://doi.org/10.1002/wcc.337

Dong, L., Leung, L. R., Lu, J., & Song, F. (2019). Mechanisms for an amplified precipitation seasonal cycle in the U.S. West Coast under global warming. Journal of Climate. https://doi.org/10.1175/JCLI-D-19-0093.1

Hoshi, K., Ukita, J., Honda, M., Nakamura, T., Yamazaki, K., Miyoshi, Y., & Jaiser, R. (2019). Weak Stratospheric Polar Vortex Events Modulated by the Arctic Sea-Ice Loss. Journal of Geophysical Research: Atmospheres, 124(2), 858-869. https://doi.org/10.1029/2018JD029222. https://doi.org/10.1029/2018JD029222

Luković, J., Chiang, J. C. H., Blagojević, D., & Sekulić, A. (2021). A later onset of the rainy season in California. Geophysical Research Letters, n/a(n/a), e2020GL09350. https://doi.org/10.1029/2020GL090350.

Singh, D., Swain, D. L., Mankin, J. S., Horton, D. E., Thomas, L. N., Rajaratnam, B., & Diffenbaugh, N. S. (2016). Recent amplification of the North American winter temperature dipole. Journal of Geophysical Research: Atmospheres, 121(17), 9911-9928. http://dx.doi.org/10.1002/2016JD0251

Swain, D. L., Langenbrunner, B., Neelin, J. D., & Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. Nature Climate Change, 8(5), 427-433. https://doi.org/10.1038/s41558-018-0140-y

Swain, D. L., Singh, D., Horton, D. E., Mankin, J. S., Ballard, T. C., & Diffenbaugh, N. S. (2017). Remote linkages to anomalous winter atmospheric ridging over the northeastern Pacific. Journal of Geophysical Research: Atmospheres, n/a-n/a. http://dx.doi.org/10.1002/2017JD026

Wang, S.-Y. S., Huang, W.-R., & Yoon, J.-H. (2015). The North American winter 'dipole' and extremes activity: a CMIP5 assessment. Atmospheric Science Letters, 16(3), 338-345. <Go to ISI>://WOS:000358020000023

Version 1:

Decision Letter:

Dear Dr Cvijanovic,

Your manuscript titled "Arctic sea-ice loss drives strong regional atmospheric response over North Pacific and North Atlantic on decadal scales" has now been seen by our original reviewer 3, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment.

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication. For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/open-access

At acceptance, you will be provided with instructions for completing the open access licence agreement on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
Link Redacted
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Heike Langenberg, PhD
Chief Editor
Communications Earth & Environment

On X(Twitter): @CommsEarth

REVIEWERS' COMMENTS:

Reviewer #3 (Remarks to the Author):

Review #2 of Cvijanovic et al. 2024

I reviewed a previous version of this paper some time ago, and am happy to provide further feedback following this comprehensive revision.

As advised by the editor, I have focused on the three editorial thresholds specifically enumerated.

My overall assessment is that the authors have taken the time to complete a highly substantive and thoughtful revision that is responsive to most or all of the major reviewer points raised previously.

In particular, the authors have done a good job clearly articulating the novelty of the methodology used here (vs other sea ice/mid-latitude circulation assessments that were originally mentioned or have been completed since, including existing intercomparison projects); the mechanistic pathway by which Arctic sea ice loss propagates first to the tropics and then (as an induced Rossby train) to the North Pacific, where it exerts an influence on West Coast winter hydroclimate; the still considerable uncertainties involved, both with respect to models, simulated and observed responses, and broader implications/where these findings fit in with contemporary debates in the climate change/atmospheric circulation change literature.

For all of these reasons, I believe this paper is now a strong candidate for publication in Communications Earth and Environment. I offer more details below, but in general I think this work is now at or near publishable form. Due to time constraints, I will not be able to offer a highly detailed point-by-point assessment at the moment, but I have carefully considered whether the authors have adequately addressed my previous major comments (they have) and also assessed whether the authors have adequately addressed the key concerns raised by the other two reviewers (I cannot speak for them, of course, but from my perspective it seems they have done so as well). I believe this research will make a valuable contribution to the ongoing and still lively Arctic sea ice/mid latitude weather debate—and I look forward to seeing what further response its publication provokes in the field.

Editorial threshold 1: Present compelling new insights on the near-term response of Northern Hemisphere climate to Arctic sea ice loss.

Yes-met. The analysis described therein clearly demonstrates that the method by which one imposes sea ice loss in models (i.e., energy conserving vs not) and also the sophistication of the ocean model used both strongly affect the nature of the mid-latitude atmospheric response (in both Pacific and Atlantic sectors) to sea ice loss. The authors have further demonstrated that the timescale of analysis matters (i.e., the response at ~10-30 years most relevant to climate adaptation and planning time horizons), relative to just ~1 year or instead centuries in the future used in other studies) for the response. Additionally, the authors find a nearly opposite response to prior studies and have now clearly articulated why their answer is different, in addition to demonstrating that it is so, which is more than many published papers can claim to have done.

Editorial threshold 2: Fully integrate your findings in the existing literature (as outlined by the reviewers) and discuss differences and similarities as well as the potential role of internal variability on the impact of sea ice loss on atmospheric circulation.

Yes-met. The authors have done a much better job contextualizing their analysis into the broader field of existing literature, which was perhaps the previous manuscript's weakest point. In particular, the authors have exhaustively responded to each individual point by all three reviewers—and the revised manuscript reflects the extensive discussions that are contained in the rebuttal document. The potential additive and/or confounding role of internal variability (including that of ENSO) is now more clearly discussed.

Editorial threshold 3: Acknowledge caveats and uncertainties appropriately, and clearly communicate levels of statistical significance.
Yes, met. The methodological and interpretational caveats and uncertainties are more comprehensively articulated and quantified in this draft; figures show significance stippling where appropriate, and differences in statistical significance between models for different variables and scenarios is also noted explicitly in the text in multiple locations.

Rebuttal 1

Reviewers' response for the manuscript entitled "Near-term impacts of Arctic sea-ice loss" by Cvijanovic et al. submitted to Communications Earth and Environment

Reviewer 1

This paper explores the near-term transient response to Arctic sea ice loss imposed in three coupled models, using an albedo-perturbation method. The authors identify a weakening of the Aleutian low, and decrease in rainfall over the southwest US, concluding that this is a response that we can expect, and prepare for, in the coming decades. I have many problems with that paper, the main one being the fact that the authors do not properly recognize the existing literature on that subject, omitting several important studies that do not align with their results. Also, the sea ice forcing and physical mechanisms behind the teleconnections they identify are poorly described, and I am not sure how robust the results are, given the often low statistical significance of the anomalies. This needs much more work to be a convincing paper on an already very crowded topic. Please find hereafter a list of comments.

Response: *We have added 20 relevant references in order to integrate more existing literature. We do not have problems with any existing study and have included in the reference list all currently published papers suggested by the reviewer. We are very aware that the response we describe is different from that described in many of these studies, and as a matter of fact, this is what makes our work novel and exciting. We discuss several different factors such as the time-scales and methodologies used in order to provide a context in which different responses may arise.*

We clearly state that we use the 95% statistical level for statistical significance of our results, as done in most studies on this topic, including the ones the reviewer asked us to cite. In addition, before performing t-test, we also check for the equivalent sample size. We are worried why the reviewer finds the very same significance testing (i.e., t-test with 95% confidence level) to be adequate in all the studies we are asked to cite (Deser et al, 2015; Sun et al., 2018, Wang et al., 2018; Blackport and Screen 2019; Smith et al. 2019, Sun et al., 2020; Blackport and Screen 2021), and also in the majority of existing literature (Ding et al., 2019, Gong et a.l, 2020, Stroeve et al., 2018, Yang et al.,2016, Nakamura et al, 2015, Peings et al., 2019, to name just some), but it is strangely not acceptable in our case. This appears to be unfair treatment.

We would also like to note that Screen at al. 2018 (Nat Geoscience), Blackport and Kushner 2016 (except in Fig. 9) or Liu and Fedorov 2018, all asked to be cited by reviewer 1, do not even show any significance testing in their figures. A recent study by Hay et al., 2022 also has no statistical significance shown.

Only our precipitation panel was shown with 90% statistical significance, and we are completely transparent about this. We now additionally comment in our manuscript on the precipitation significance in case the reader had missed that in the figure caption. If the reviewer would prefer, and in order to avoid further criticism or speculation that we have given these anomalies more importance than they deserve, we can remove any significance that we have shown, like it was done for all the variables in the high impact studies by Screen at al. 2018 (Nat Geoscience), Blackport and Kushner 2016 (except in Fig. 9) and Liu and Fedorov 2018. We can discuss them accordingly, in the style of these previous studies. The key point of our study, however, is our geopotential response that is statistically significant at the 95% level.

Regardless of this, our most important features, geopotential height anomalies are extremely robust and they consistently show positive geopotential anomaly in the North Pacific in all three models included in the analysis, a feature meteorologically associated with drying over California (like it was the case with the "Ridiculously Resilient Ridge" during the 2012-2016 California drought).

The teleconnections described in our study are well known and documented, and we have referred to the original studies that have described them. This body of work of tropical to high latitude meteorology spans over the last 40 years of climate science, and it is not our aim to reinvent the tropical forcing of extratropical Rossby waves or write a review paper on the topic. We have used two additional sets of specialized simulations to demonstrate the importance of surface propagation in our mechanism, something that reviewer 1 did not appear to notice. All the key features shown are robust between the three models. If the reviewer considers that certain paragraphs merit a better or more detailed explanation, they should explicitly describe the lines so we can rewrite them carefully.

We have included the discussion of the potential role of internal climate variability on the impact of sea ice loss on atmospheric circulation, but we wish to be very clear that our results are not any less significant than the ones in other previous studies cited by the reviewer, and that our signal of North Pacific geopotential height changes is robust between three completely different models.

1) My first problem with that study is that several important papers on this very topic are not cited, and/or are not properly discussed to interpret the results and put them in perspective with what has been done previously.

Response: *In the revised version of our manuscript we added over 20 references to integrate the most relevant and recently published literature. We discuss different factors such as time-scales and methodologies used in order to provide a context in which different responses can arise, placing our study in a wider perspective. We also added a new Supplementary Table 1, in which we provide an overview of these previous papers, along with the time scales and methodologies applied in each of them. We invite the reviewer to consider this new table that provides all relevant information in a succinct manner.*

Table 1 (also added to the manuscript as Supplementary Table 1): Summary of some of the key studies focusing on Arctic sea-ice loss and its impacts on the weather and climate across the Northern Hemisphere. Listed are descriptions of the experimental setups used, as well as methods for significance testing and the type of response found.

Reference:	Model:	Ice Target:	Number of years simulated	Model spin-up length and type of response studied	Forcing	Method to constrain	Arctic or Antarctic or both	Significance estimate	North Pacific response in winter (DJF)
Deser et al., 2015	Coupled (CCSM4)	Ensemble mean of 6 historical (1980-1999) and 6 RCP8.5 (2080-2099)	1 member of 360 years	100 years, centennial response	Year 2000	ghost forcing, non-energy conserving	Arctic	Two-sided Student's *t* test 95% confidence level	Aleutian low (SLP)
Oudar et al., 2017	Coupled (CNRM-CM5)	Ensemble mean of 5 historical (1970-1999) and 5 RCP8.5 (2070-2099)	1 member of 200 year	100 years, centennial response	Year 1985	Flux adjusting, non-energy conserving	Arctic	Student's t-test at the 95% confidence level	Aleutian low (SLP)
McCusker et al., 2017	Coupled (CanESM2)	100 year PI CTRL and nearly equilibrated doubled CO 2 simulation	1 member of 301 year	100 years, centennial response	PI	Nudging, non-energy conserving	Arctic	95% level using a two-sided Student's t test	Aleutian low (SLP)
Blackport and Kushner 2016	Coupled (CCSM4)	year 2032–41 in RCP 8.5 for CTRL and 2072–81 under RCP8.5	1 member of 680 years for the CTRL and 1 member of 800 years for the reduced	270 year, centennial response and decadal response	Year 2000	Albedo modification, Energy conserving	Arctic and Antarctic	Not shown except for Fig. 9 where Student's t test with 95% confidence level	SLP decrease in century scale experiments and weak SLP increase in the North Pacific for 50 year long

			sea-ice and 8 fifty year long simulations						experiments, geopotential not shown
Blackport and Kushner 2017	Coupled (CESM1)	Ensemble mean 30 members of RCP 8.5	1 CTRL member of 725 year and 1 member of 525 year for the reduced sea-ice	300 year for CTRL and 275 year for reduced sea-ice, centennial response	Year 2000	Albedo, Energy conserving	Arctic and Antarctic	Student's t test with 95% confidence level for z500	Aleutian low (SLP and Z500)
Blackport and Kushner 2018	Coupled (CESM1)	Ensemble mean 30 members of historical and RCP 8.5 (1920-2100)	1 CTRL member of 725 year and 1 member of 525 year for the reduced sea-ice	425 year for CTRL and 275 year for reduced sea-ice, centennial response	Year 2000	Albedo, Energy conserving	Arctic and Antarctic	Student's t test	deepening of the Aleutian low, and an expansion of the Siberian high
Smith et al., 2017	Coupled (HadGEM3)	Ensemble from 1979 to 2009 with observed Arctic and Antarctic for the CTRL and sea ice representing the next couple of decades	10 members of 30 years	30 years decadal response	present-day	Nudging, non-energy conserving	Arctic and Antarctic	Results not significant at the 95% level based on a two-tailed Student's t test	Aleutian low (SLP)
Blackport and Screen 2019	Coupled (HadGEM2-ES)	Present-day CTRL ensemble forced	400 5-year realizations of	4 years? decadal	present-day	Albedo, Energy conserving	Not stated	Student t-test 95% confidence for	January geopotential high (positive

		with RCP8.5 2008–2012, 2 °C global warming above pre-industrial levels RCP8.5 forcing 2036–2040, two ensembles with reduced sea ice	present day and reduced sea-ice	response				geopotental, precipitation not shown	SLP)
Liu and Fedorov 2019	Coupled (CESM)	PI for CTRL and	10 members of 200 years	150 year, centennial response	PI	Albedo, Energy conserving	Arctic	Not given	Aleutian Low (Z850)
Sun et al., 2018	Coupled (CM3)	Transient Arctic sea-ice evolution from 1990 to 2090	5 member over 1970-2090 under RCP8.5 and 5 similar member but with sea-ice of year 1990	0 year decadal response	RCP8.5	Nudging, non-energy conserving	Arctic	two sided Student t-test 95% statistical significance	Weak Aleutian low (SLP) only weakly significant after 100 years of simulation
Sun et al., 2020	Coupled (CCSM4)	five-member ensemble CTRL 1980–1999 and five-member ensemble of year	360 years	100 years, centennial response	Year 2000	Albedo and ghost flux and nudging, Energy and non-energy conserving	Arctic	95% confidence level based on a two-sided Student's t test.	Aleutian Low (not significant)

		2080–2099 under RCP8.5 for reduced sea-ice							
Seidenglanz et al. 2021	Coupled (CMCC-SPSv3)	CTRL and perturbed sea ice cover in the BKS during November, 23 winter seasons (1993–2015), initialized on November 1, 10 ensembles per season, (230 seasonal simulations)	23 years	6 months decadal response	present-day	Energy conserving	Arctic	95% confidence level based on a test for differences of the mean for paired samples (Wilks, 2011)	geopotential high

The study is presented as if it was the first time the transient, or near-term, response to Arctic sea ice loss is investigated (it's not),

Response: We have to disagree with this statement. *In the sentence above, the reviewer has completely reverted the meaning of what we wrote in our manuscript. We kindly ask the reviewer to understand and acknowledge the following:*

We didn't write this is the first transient OR first near-term study on the impacts of sea ice loss at any place in the text. We are very clear that this is the first multi-model study applying the energy conserving setup to investigate the near-term impacts of Arctic sea- ice loss. This means all three descriptors: 'multimodel', 'near-term' (decadal) AND 'energy conserving' at the same time, not any of them separately, and we believe this is clearly conveyed in the text. Therefore, we kindly ask the reviewer to reconsider our aims and the novelty of our work.

and as if the methodology used to melt the ice is novel (it's not, and like all methods, it has its own limitations).

Response: *Our methodology of changing sea-ice through perturbed sea-ice physics ensemble is novel, it has been used only twice (on CCSM4 slab ocean (Cvijanovic et al. 2017) and IPSL-CM5A2 models (Simon et al. 2021)- we cited both studies) and now it has been developed for the EC-Earth3.2 model for the very first time. It is not fair to claim otherwise without providing evidence.*

We do not know what are the exact limitations that the reviewer would like us to discuss, since they have not specified them, but this methodology is energy conserving, meaning that it does not involve adding spurious energy fluxes to change the sea ice cover as commonly done. This is an important advantage to the common approach of adding artificial energy fluxes, because the energy added in this manner affects the rest of the climate system and can alter the experimental results. In contrast, our approach conserves the energy in the system and ensures that the observed response is due to Arctic sea-ice ***only****.*

In addition, ours is not just a pure 'albedo' whitening approach, but it involves changing the parameters related to snow cover ice albedo and conductivity, allowing for fine tuning of summer and winter changes at the same time based on experimental needs.

For example, l.54, you say that"… far less attention has been given to the near-term, decadal-scale impacts" of sea ice loss. First, there is a whole Model Intercomparison Project (MIP), the Polar Amplification MIP, dedicated to understand the impact of +2°C Arctic and Antarctic sea ice loss (Smith et al. 2019). See Peings et al. (2021) for one of the first papers on these simulations, that shows the response to +2°C sea ice loss (i.e near-term) in both atmosphere-only and short coupled runs. Their coupled runs cannot really be compared to yours though, because they are 14-month runs that do not account for longer-term adjustment of the ocean. But PAMIP also includes long coupled runs in which the longer-term transient response can be investigated. Several modeling centers have already run those and some of their results should be published soon. It would be fair to mention PAMIP and the associated research because it aims exactly to answer the questions you are exploring in your paper, in a coordinated manner so different model results are more easily comparable.

<u>Response:</u> *We have to disagree once again with the comments by the reviewer. We elaborate why below.*

Firstly, the majority of PAMIP simulations are run for the length of 1 year (14 months) and there are no fully coupled ocean-atmosphere experiments like ours run on decadal scales in PAMIP (see Table 1 in Smith et al 2019). As the reviewer correctly points out, the PAMIP simulations then cannot be compared to ours, because the time scale is clearly different. That said, we also cannot account for and discuss simulations that "should be published soon". It is simply not possible or in line with the publication/review process.

Secondly, the PAMIP protocol does not employ energy conserving methodology for isolating the impacts of sea ice loss, which is a key difference with our simulations.

We include the citation to the PAMIP study mentioned by the reviewer and we mention the PAMIP protocol in the new version of our manuscript as requested, and we are very clear that these are a ***completely*** *different type of simulations from ours. Please refer to lines 178-181, 398-401 and 420-424 of the revised manuscript:*

Second, the transient response to sea ice loss has also been explored in Sun et al. (2018). In particular, they look at the response to RCP8.5 projected changes over the 2021-2050 period, i.e. near-term changes, like you. They do not find any significant circulation response over the North Pacific (see Z500 anomalies in their Fig. 8). The sea-level pressure response does not quite resemble the one in your three models, it looks like a wave-1 pattern that has been found in other coupled models (Screen et al. 2018). The precipitation response over the western US is also different, it looks like a north-south wet-dry dipole, but very little statistical significance. Screen et al. (2018) is another very important paper on the response to sea ice loss in coupled models, I am surprised it is not cited. It is true there is whole range of different responses in model studies in the literature, but it tempers a little your statement of l. 56 about contradictory findings regarding the remote impacts of Arctic sea ice loss. Their multi-model study shows rather good agreement in the response to sea ice loss in coupled models.

<u>Response:</u> *Thank you for your comment. We do not claim that we are the only people to investigate the decadal responses, we say that we are first to investigate multimodel decadal response in an energy conserving setup as we highlighted earlier. The studies you mentioned substantially differ from ours and do not satisfy our criteria on energy conserving methodology and decadal response. We are trying to make methodologically clean comparisons.*

Sun et al., 2018 like many other studies used the non-energy conserving nudging methodology. They impose RCP forcing, but artificially force the ice not to melt by applying artificial energy perturbations. This is methodologically different than our approach. We have now included a citation to this study in our manuscript on line 63, and we added it to our Supplementary Table 1, but this makes no difference to our conclusions and supports our argument, because their setup is non energy conserving (they use nudging).

We have also included the citation of Screen et al. 2018 on line 63 as requested by the reviewer. This study, however, shows no measure of any significance in their figures, which makes it somewhat complicated to discuss their results in our context. If we ignore the fact that there is no significance indicated, we can see the apparent agreement found by Screen et al. 2018 is consistent with the fact that the experiments compared in their study represent centennial responses to sea ice loss (the

original studies describing the experiments used in Screen et al. 2018 are the first 6 studies listed in our new Supplementary Table 1). All of the experiments used in Screen et al. 2018 investigate the centennial responses except for Smith et al., 2017 and Blackport and Kushner 2016 who, however, consider Antarctic and Arctic sea ice impacts at the same time, so they do not make an adequate comparison to our study either. Finally, most of the experiments in Screen et al. 2018 use non-energy-conserving methodologies (see Supplementary Table 1 for details). In this way, the consistencies found by Screen at al. 2018 when comparing the centennial scale responses in the experiments with non-energy conserving setups (the study's title mentions both 'consistencies and discrepancies') are not in contradiction with anything we state in our manuscript, as we focus on different time-scales and methodologies that they did not investigate.
The experiments that consider Antarctic and Arctic forcing at the same time in Screen at al. 2018 are not comparable to our experiments too, because results are clearly expected to be different when there is sea ice loss in both the Arctic and Antarctic regions simultaneously.
In conclusion, the work presented in Screen et al. 2018 does not feature any experiments of the kind discussed in our study: energy conserving, decadal scale, and investigating the impacts of Arctic sea ice loss alone ("and" in this sentence designates that all three qualifiers "energy conserving, decadal, Arctic forcing alone" are to hold at the same time). Therefore, we cannot make meaningful comparison between our results and the results in these studies. However, we mention all of them in our revised manuscript.

Blackport and Kushner (2018) is cited in introduction, but not discussed later, while this is a paper that explores the response to Arctic sea ice loss in transient coupled runs too. And they also use an albedo perturbation method to melt the ice. Their results do not resemble the response you get (they are in fact opposite) : weakening of the Aleutian low, and increased precipitation over the US West coast. I am not saying that their results are more trustworthy than yours, but it illustrates the uncertainties in looking at remote regional impacts of Arctic sea ice loss.

Response: *We have included further discussion of Blackport and Kushner (2018) on line 406. However, this study analyzes atmosphere-only experiments so this is also not an appropriate comparison, because as they write themselves, many dynamical processes are omitted in atmosphere-only experiments, and the importance of SST changes have been stressed in our study as essential for the teleconnections we observe in our results. Another issue is as mentioned earlier that the forcing applied in their study is different to ours, because it involves Arctic and Antarctic changes at the same time.*

Liu and Fedorow (2019) is another study looking at the transient response to Arctic sea ice loss in a coupled model. They also use a method to melt the ice that looks very similar to yours (albedo perturbation through alteration of ice/snow parameters). They look at the short-term (1-15 years) and longer (151-200) global response to Arctic sea ice decline. I do not know what value to give it, but they find an increase in precipitation over the Western US (their Fig. 2). Plus they also show a clear role for the AMOC decline in driving the response to sea ice loss, an aspect that is overlooked in your study, especially when you say the response is not sensitive to the fact of using a slab (no dynamics) versus coupled ocean, for example l. 185 : "employing the slab ocean instead of a fully coupled dynamic-thermodynamic ocean, does not substantially affect the simulated response to sea-ice changes on decadal timescales.". This is in contradiction with pretty much all the studies I mention above.

Response: *The Liu and Fedorov 2019 study shows response for years 1-15, but unfortunately for this period they do not show geopotential or SLP anomalies (in the main text or in the supplementary), so it is not possible to know if our responses resemble or not. Also their study does not indicate any significance of the anomalies shown, therefore it is really difficult to derive any suitable comparisons with their precipitation anomalies that are anyway noisy in our study too. We mention this study in our revised manuscript on lines 63, and 64.*

It is extremely important to clarify that ocean processes (e.g.., AMOC changes) are ***not*** *overlooked in our study. It is incorrect to claim otherwise. We provide information on the importance of ocean changes in the decadal response to sea ice changes on lines 59-62, 64-67, 255-270, and 281-284. We employ two state-of-the-art atmospheric models coupled to full depth dynamic-thermodynamic ocean models and one with simple mixed layer (the so called slab ocean). Our results clearly demonstrate that the slab ocean simulations show the same type of response (positive geopotential height anomaly and drying over California) as the two models featuring a fully complex ocean component. Therefore, this logically shows that the deep ocean changes are not necessary for the type of response we describe in the North Pacific. Our experiments are clear scientific evidence for this. We are also very clear that this is not the case in the Atlantic sector where there are differences present, and this could be due to ocean complexity (including a potential AMOC decline). Our North Pacific geopotential response in terms of positive (and not negative) geopotential height anomaly is however consistent independently of model type or ocean model complexity.*
We also want to note that contrary to what the reviewer claims, this is not in contrast with the studies that have been mentioned, because ***all*** *of those studies either use different (i.e. non-energy conserving) methodologies for isolating the impacts of sea ice loss or consider different timescales (as we have repeated multiple times in our manuscript and in this rebuttal).*

2) This leads me to my second comment, and how put you put your results in perspective with previous findings in the literature. l. 280-283, you just discard previous studies that found a different response over the North Pacific (deepening of the Aleutian Low, versus weakening in your case) by saying they were looking at the long-term (one century or more) response to Arctic sea ice loss. This is simply not true. And you omit to cite them, intentionally or not (I hope not). Sun et al. (2018) and Liu and Fedorov (2019) are two of them, they look at the near-term transient response to Arctic sea ice loss. They may be more I'm missing.

Response: We feel that *this is a completely unfair and false conjecture. We have so far in our response to reviewer 1 discussed all studies mentioned, and explained how they differ methodologically from ours. Any scientific field, from climate to medicine, requires careful separation of methodologies when interpreting the results, and the reviewer has persistently neglected important methodological differences, and is attempting to group all the sea ice studies together independently of the fact that some focus on the response 500 years after the sea ice perturbation and others 20 years later, that some methods involve imposing completely artificial energy fluxes into the system and are, therefore, not energy conserving and others do not do that at all, that some studies made huge additional alterations of Antarctic sea ice cover in combination with Arctic sea ice loss. Grouping everything into one incomprehensive category without any attention to methods implies that the "right" answer is the one that comes from the most numerous type of studies, which are the centennial non-energy conserving studies. We firmly disagree. If the reviewer thinks that the methodology is an irrelevant factor, it is their task to prove this statement.*

The sea ice community appears to be stuck and unable to explain the different responses, and one big issue is that comparisons have not been rigorous enough. For example, if two given studies differ in *more than one factor, than it is not possible to just choose one of the differing factors and say that this is the cause of the observed difference. To take as an example Screen et al. 2018 that the reviewer persistently mentioned, the experiments used therein differ from ours in terms of: (1) the method used to isolate the impacts of sea ice loss (they are non-energy conserving), (2) the time-scales and (3) some of them even use Antarctic and Arctic forcing at the same time. Therefore, it is not possible to tease out one reason for the discrepancy and the field never attempts to do so. We are very careful to mention several possible reasons and we hope the reviewer could start to appreciate this. We worked really hard to provide clean and accurate comparison and our study is quite unique in this way.*

In any case, we have included all the studies the reviewer has suggested, although we think that at many instances they just make the manuscript longer without any added value, as we repeat again and again, decadal and energy conserving versus centennial and non-energy conserving.

Finally, our Discussion already included this explanation that the reviewer appears to have missed (now on lines 401-410):

"These study design differences include: different time-scales of the responses considered (i.e., centennial [6–17] vs. decadal [18–20]), the type of method used for isolating the impacts of sea-ice loss (energy conserving [8,19,20] versus non-energy conserving [6,7,9,18]), consideration of Arctic changes alone [6,7,9,18–20] as opposed to considering both Arctic and Antarctic changes simultaneously [8,68], model complexity (atmosphere-only [68] vs. slab ocean [19] vs. fully coupled model simulations [6,7,18,20,31]), consideration of sea-ice changes over only part of the Arctic as opposed to the whole Arctic sector [31,47], emphasizing different seasons (autumn vs winter sea ice anomalies [21], different background ocean states as well as any intermodel differences."

We also add now in continuation on lines 410-419:

"For the last point, Arctic sea-ice response could be influenced by decadal SST variability, as discussed by [5,69,70], or by internal variability even when a large number of ensemble members is available, and even for anomalies that are statistically significant based on a Student's t test [71]. In particular, ENSO is known to have strong remote effect on both Pacific and North Atlantic circulation and could potentially interfere with the Arctic-midlatitude linkages, modulating the SST-forced Rossby wave train. For example, in Supplementary Fig. 9 we show that during strong El Niño events the response in our simulations is a geopotential low in the North Pacific as expected, but a more intense geopotential high is observed during strong La Niñas in accordance with our finding that the mean response to Arctic sea ice loss is La Niña-like."

And do not see why the atmospheric response in the North Pacific would flip sign in the near-term vs the long-term, so the long-term argument to explain different results in other studies is questionable. I would argue that your results can also be compared to Blackport and Kushner (2018), Screen et al. (2018), and more that I may omit here.

Response: *From the time-scale perspective alone, it is actually perfectly reasonable to expect a different response a couple of decades and a couple of centuries later, as the deep ocean takes a long time to fully develop its response. We have already mentioned this in our study and cited the experiments by Wang et al. (2018) that nicely illustrate the development of the ocean response through time on lines 64-67:*

"Physically, "decadal" responses are dominated by faster adjustments of the climate system to sea-ice changes (e.g., atmospheric reorganizations and fast oceanic responses), while the "centennial" responses allow for slower deep ocean feedbacks to develop too"

The North Pacific sea-level pressure response you get in these runs does not nearly resemble the SLP response that is commonly found in this type of experiments. See Fig. 5 of Blackport and Kushner (2018), Fig. 2 of Screen et al. (2018), Fig. 8 of Sun et al. (2019).

Response: *Correct, none of our North Pacific response SLP, geopotential and finally precipitation is like in the studies the reviewer mentions, but they are completely consistent between three different climate models (that feature different ocean model complexities) and use energy conserving methods. This is the very reason why our study is important. We have listed above several key methodological differences for each one (including non-energy conserving approaches, different time-scales of the response, Arctic and Antarctic forcing simultaneously).*

Liu and Fedorov (2019) do not show SLP anomalies but their Fig. 2 does not show a drying over the Southwest US.

Response: *We have already answered this once above: this study does not show SLP or geopotential anomalies for years 1-15 and does not show any significance for precipitation anomalies. This does not give us anything to make a comparison to, but we do cite this study on lines 63 and 64.*

In summary, this study is just one among others, that happen to find a seemingly consistent response in the North Pacific (but statistical significance is low, see one my following comment).

Response: *To make such a statement requires some really strong scientific arguments that the reviewer has not provided at all. We show a robust response between three different and independent model runs with positive geopotential anomaly developing in the North Pacific. Statistical significance used for geopotential is not low, it is 95%, the same as in any other study the reviewer asked us to cite and completely standard in our field, and better then in several studies that the reviewer keeps on mentioning (that do not show any significance, e.g. Screen at al 2018, Liu and Fedorov 2019, Blackport and Kushner 2016).*

By careful analysis of all the different factors varying between different methodologies used, we come to the conclusion that the time scale/energy conserving methodology is a common denominator for our results.Our findings are also supported by the studies by Sewall and Sloan 2004, Sewall 2005 and Seidenglanz et al. 2021.

Given the seriousness of the statement made above by the reviewer, made without reasonable explanation, we ask the reviewer to reconsider if there are any obstacles in terms of conflict of interest, personal issues with any of the authors or any other conflicts that would prevent them from providing a fair review. If any of the latter is the case, the reviewer should kindly inform the Editor to ensure that our submission receives a fair review process.

However the overall response does not quite align with previous findings.

Response: *Again, yes, that is the novelty and meaning of our study, it is different and we carefully explain why, and for this very reason these results need to become available for the community.*

Why is the SLP response so different from the wave-1 pattern found in previous studies, for which there seems to be more and more consensus that it is a robust feature of the atmospheric response to sea ice loss (Screen et al. 2018) ? I do not buy the argument that this is because previous studies have looked at the centennial response, or used non-conservative energy methods (see my next comment), because as I said, this is inaccurate.

Response: *We do not expect that the reviewer has to "buy" anything, because that would be rather rude. We would like to ask them to be more careful with their interpretations and assertions that are often misleading and not supported by evidence.*

Screen at al. 2018 uses experiments from 6 previous papers, and as we explained above these are methodologically very different from ours. In contrast to the reviewer, we consider this to be the main reason for the different results that we obtain.

This is a scientific review, and if the reviewer claims that something is inaccurate, then they should provide scientific arguments and not just "I don't believe" statements. We consider the above comment as completely inappropriate and rude.

Again, we have very carefully explained all the methodological differences between different studies. Understanding differences in method in order to understand differences in results is the essence of the scientific process.

3) Another argument you make to explain why previous studies may have found different results is that they were using "non-energy conserving" experimental setups. Again, not true. Many other studies have used a method similar to yours. See Sun et al. (2020) for a review of different approaches that people have used to carry out these sea ice loss experiments in coupled models. Like said in that paper, albedo reduction conserves energy but is only effective during the sun-lit portion of the year. By contrast, ghost flux and ice-flux nudging do not conserve energy but are effective in adjusting sea ice year-round. This is not discussed anywhere in your paper, and it sounds like your are coming up with a whole new method that nobody have thought about before.

Response: *We stress that there are several incorrect statements made in this comment. Our method is not similar to the method in which sea ice or ocean albedo reductions are imposed, the only common quality is that both approaches are energy conserving. Our method utilizes small perturbations in sea-ice physics parameter values within their user defined ranges. These are not only albedo related parameters, but also conductivity, and as such our approach (unlike the ones focusing on the reduced albedo) is not dependent only on the sun-lit portion of the year. We have carefully chosen the balance between different physical parameters we perturb to achieve the loss that imitates the one that happens in the real world – that is the response that is largest at the end of summer and smallest in the early spring. There is nothing in our methodology that forces us to focus on summer loss, if we wanted to maximize a different season (which we didn't), we could have chosen different parameter values that are more affecting conductivity and we would have increased the winter loss.*

We dare say that our method is far more advanced to the ones that merely change the albedo (paint the ocean white or ice black), because in contrast to these approaches our approach is physical. We only change physical parameter values in the sea-ice model within their expert defined ranges. These are the parameters that are typically set in a model as a single value and we use the fact that this value can be perturbed in a given range. The information which parameter to choose (which are most powerful in changing the resulting sea ice cover) has been derived from a large uncertainty quantification ensemble simulations as explained in Cvijanovic et al. 2017.

Therefore, the claim that there are many studies that used a method similar to ours is completely incorrect. Cvijanovic et al. 2017 introduced this methodology, Simon et al. 2021 compares the impacts of conductivity and albedo related parameters in the IPSL-CM5A2 model, and our proposed study employs the technique for the first time in the EC-Earth3.2 model, as well as in a multi-model context. To put it in other words, painting the ocean white is not what we have done here.

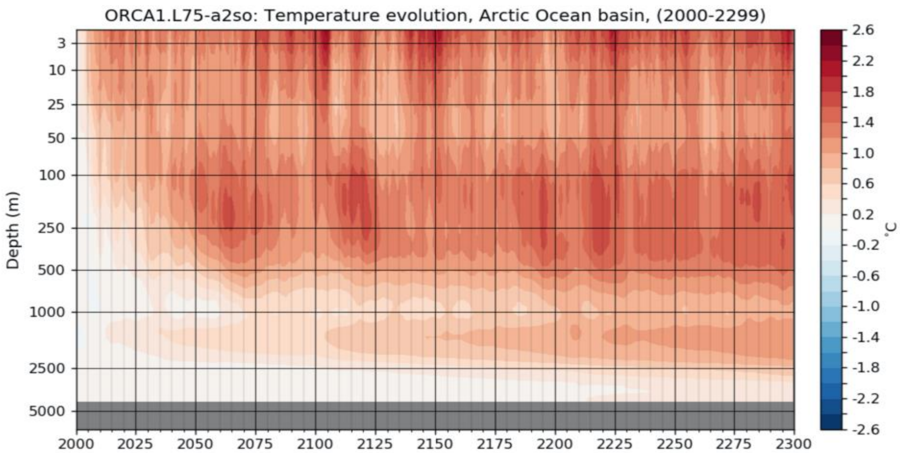
We never state that the differences arising from different studies are due to the energy conserving versus non-conserving methods alone, because it is not possible to arrive to such a conclusion. We are always very careful to compare all the factors that differ, and we find that this is at minimum the time-scale and the method together (and sometimes even unintentional insertion of Arctic and Antarctic forcing simultaneously as in Smith et al. 2017).

This is another worrisome example when the reviewer reverts what we have stated. We hope this is not done intentionally. Please also note Figure 1 below that shows that different methodologies (i.e. energy conserving versus non-energy conserving actually makes a huge difference in the ocean response.

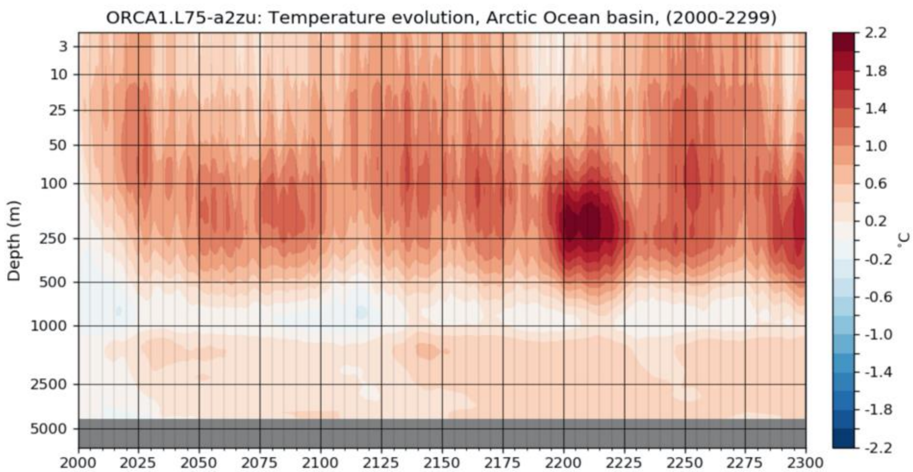
How do you know that your method for melting the ice gives you more trustworthy results than other ones ? Blackport and Screen (2019) suggest that most of the influence of sea ice on the winter climate comes in fact from the winter sea ice anomalies, not the autumn ones. This suggests that albedo-perturbation sea ice loss experiments may miss some important aspects of the response to sea ice loss, because they do not melt enough ice in winter. This is the case in your experiments, they do not have a great deal of ice melt in winter (Fig. 1a-c).

Response: *Firstly, we want to* ***emphasize*** *that we do not discuss which method gives more "trustworthy" results. It is certainly questionable when some methodologies impose artificial energy fluxes how this energy affects the overall results, but this is not the topic of our work. Especially for the reviewer, we have run the EC-Earth3.2 model for a couple of centuries with energy conserving (this study style) and energy non-conserving (PAMIP style) setups (shown in Figure 1 below) and the differences are huge as one would expect. Figure 1 will we be used in a different study that is currently in preparation, but we think it is useful to be shown here, just because the reviewer continues to dispute the importance of methodological differences.*

EC_Earth physics parameter perturbations (energy conserving)



EC-Earth extended PAMIP run (non-energy conserving)



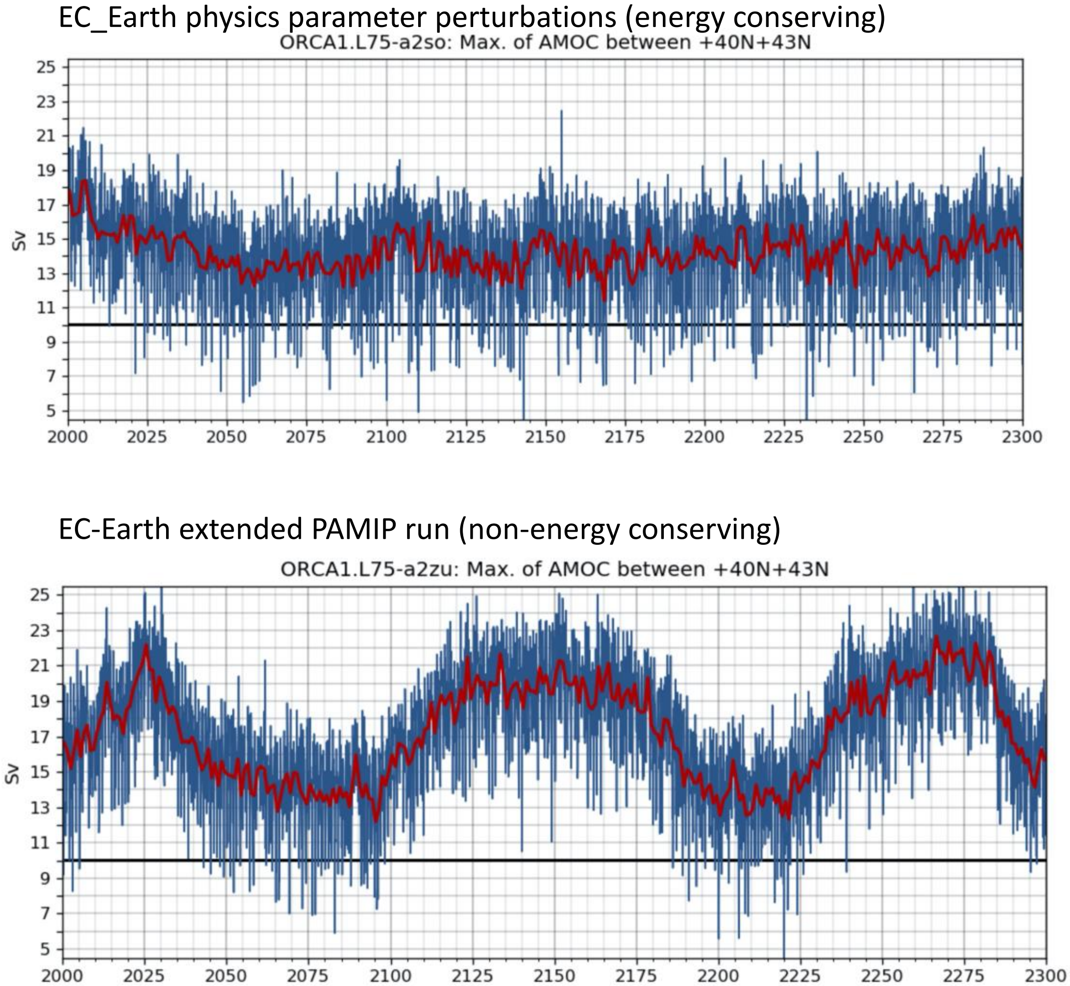


Figure 1: Comparison of energy conserving and non-conserving simulations obtained with the EC-Earth3.2 model (Levine et al. in preparation). Note the differences in the profiles of subsurface warming and the instability of the strength of AMOC in the non-conserving approach.

In Figure 1 the reviewer can see that the method indeed makes a difference as the subsurface temperature evolution of the Arctic Ocean and the AMOC responses in the northern Atlantic are rather different when applying energy conserving and non-energy conserving methods to melt the ice.

It is a widely known fact that the Arctic loses the highest amount of sea ice in late summer and the least in late winter, and this is what we have chosen to simulate in our study (as seen in Fig. 1 of the article). The claim that we 'do not melt enough ice in winter' is completely arbitrary. We have chosen our parameter perturbations to achieve the shape of sea ice loss observed in nature, we have reported how it looks like and have shown how a response to it looks like. This was our motivation and we have designed our experiments to do so.

What we have added to our manuscript (on lines 401-410) based on this discussion is the statement that different methods can also emphasize different seasons, and this could also be one of the reasons behind the different responses.

"These study design differences include: different time-scales of the responses considered (i.e., centennial [6–17] vs. decadal [18–20]), the type of method used for isolating the impacts of sea-ice loss (energy conserving [8,19,20] versus non-energy conserving [6,7,9,18]), consideration of Arctic changes alone[6,7,9,18–20] as opposed to considering both Arctic and Antarctic changes simultaneously [8,68] , model complexity (atmosphere-only [68] vs. slab ocean [19] vs. fully coupled model simulations [6,7,18,20,31]), consideration of sea-ice changes over only part of the Arctic as opposed to the whole Arctic sector [31,47] , emphasizing different seasons (autumn vs winter sea ice anomalies [21] , different background ocean states as well as any intermodel differences."

4) An important information that is missing is how much sea ice loss do your simulations represent. How does your sea ice loss forcing compare with previous studies ? It is a forcing that we may expect in the next decades, as projected by coupled climate scenario simulations ? Is it comparable with the near-term sea ice loss forcing created for the PAMIP exercise ? This is never discussed, despite being a very important question.

Response: *It is not accurate to say that this information is missing, since we show the simulated sea ice loss both in Figure 1 and Supplementary Fig. 1 and also discuss it.*
We do however now clarify that this is the type of forcing expected in the next decades on lines 53 and 124-125:
"The type of sea ice loss forcing applied in the 'low ice' experiments corresponds to the forcing expected to occur in nature in the next decades."

We also now reflect on the fact that studies using different methods may emphasize the loss in different seasons compared to ours (see the response to the previous comment).

5) The physical mechanisms behind the atmospheric responses are not clearly identified. There are assumptions that the high-latitude forcing influences the tropics through mid-latitude eddy momentum flux changes and the tropical WES feedback, mostly based on results from previous studies. But the authors do not show what is exactly happening in their runs. For example, what is the response of AMOC in their runs (Sun et al. 2019, Liu and Fedorov 2019) ? How do physical mechanisms align with previous findings in the literature ?

Response: *We do not use assumptions, we cite and discuss some of the key studies describing the tropical to high latitude (and other way around) teleconnections, developed over the last 40 years. In Figs. 3 and 4 and Supplementary Fig. 3 we focus exactly on these teleconnections and how they play out. For this purpose we have added extra simulations with a specialized setup. Adding an extra setup and showing its result is not an assumption, it is a way to provide scientific evidence.*

One of our major findings is that the AMOC does not play an important role for the North Pacific response at decadal time scales. This is based on the fact that the slab ocean experiments that do not feature the AMOC changes arrive to the same type of response as the experiments using the full-dynamic-thermodynamic ocean model. We also discuss how there are more prominent differences in the Atlantic, that are AMOC driven on lines 304-312:

"The AMOC could play a role at decadal time scale in the Atlantic. In[20] the weakening of the AMOC due to Arctic sea-ice loss is suggested to warm the Atlantic at 0-25S and shift the ITCZ in the Atlantic southward. The fact that the atmospheric response is slightly weaker in the coupled runs compared to the slab-ocean run at decadal timescale (Fig. 2) might, therefore, be explained by the weakening of the AMOC. When the AMOC weakens, less heat is transported to the high latitudes compensating a bit the

excess of heat induced by the loss of Arctic sea ice. Thus the north-south temperature gradient is smaller for coupled runs than the slab-ocean one, and the response is diminished in the EC-Earth3.2 and IPSL-CM5A2 runs compared to CCSM4-SO."

6) Statistical significance of the results. From the description of the experiments, I get that each of your experiment represents ~200 years of data, i.e. ~200 fall/winter seasons to analyze. Is that correct ? This should be mentioned. Internal variability is a huge factor that can explain model differences, that is not even mentioned a single time in your paper. It is more and more recognized that a lot of conflicting results between different modeling studies, as well as results from observational analyzes, can simply arise from internal variability, and the lack of robustness in the statistics or model responses that are analyzed. See Peings et al. (2021) for an illustration of how internal variability alone can create substantial differences in the remote response to a same sea ice loss forcing. I wonder how robust your results are here. 200 realizations to analyze is a fair amount, but many of the signals you discuss are at the edge of statistical significance. For example, the SLP anomalies in the North Pacific are not significant in IPSL-CM5A2 (Fig. 2f), as are the OLR anomalies of Fig. 4f for this model). Are those robust signals enough to be trusted ? Especially you use a 90% confidence level in some cases, I suspect to retain enough statistical significant anomalies, but this is quite a low restriction. Plus it is always a good idea to account for spatial correlation when assessing statistical significance of a field, using the False Discovery Rate of Wilks (2016). I suspect using this test may remove some of the "significant" anomalies you get, especially in the fields for which you only use a 90% confidence level threshold.

<u>Response:</u> *For each of three models, we use 200 years of data (which is double than 100 years used under the PAMIP protocol). This is explained in our methods section where we indicate that we analyze the last 20 years of the 'low ice' simulations and we have 10 ensemble members for each model.*
We now include a short discussion about the impacts of internal variability and cite Peings (2021) on lines 410-413:
"For the last point, Arctic sea-ice response could be influenced by decadal SST variability, as discussed by[5,69,70], or by internal variability even when a large number of ensemble members is available, and even for anomalies that are statistically significant based on a Student's t test[71].

As explained earlier in response to the reviewer's first comment, we use the 95% significance level, as done in the majority of studies on this topic, including all of the ones the reviewer asked us to cite. In addition, before performing t-test we also check for the equivalent sample size. We think that it is very unusual that the reviewer finds the very same significance testing adequate in studies by Deser et al, 2015, Sun et al., 2018, Wang et al., 2018, Blackport and Screen 2019, Smith et al. 2017, Sun et al., 2020, Blackport and Screen 2021, Ding et al., 2019, Gong et al, 2020, Stroeve et al., 2018, Yang et al., 2016, Nakamura et al, 2015, Peings et al., 2019, but not in ours, and again, we would like to bring the Editor's attention to this issue. We also would like to note that Screen et al. 2018 (Nat Geoscience) and Blackport and Kushner (mentioned so many times by the reviewer) do not even show any significance in their figures. So especially when taking into account that we also additionally check for temporal autocorrelation by determining the equivalent sample size, our significance testing is completely valid by the community standards.

Only the precipitation panel was shown with 90% statistical significance and we have highlighted this in our manuscript now on lines 143 and 292 (even though this was already indicated previously in the caption of Figure 2). Our geopotential patterns are extremely robust and they consistently show a

positive *geopotential anomaly in the North Pacific in all three models, a feature meteorologically associated with drying over California (like it was the case with the Ridiculously Resilient Ridge during the 2012-2016 California drought).*

It is true that there is weaker significance of simulated changes in the IPSL-CM5A2 model (which could be a consequence of smaller sea ice changes in this model), but this does not change the fact that the main feature of our response – the geopotential anomaly, is positive and statistically significant in the North Pacific (and not negative like in the other studies mentioned by the reviewer). Consistence in dynamical response is what is the crucial element of this paper. Precipitation, which is a known noisy variable, represents a consequence of this robust dynamical response, and it is in line with what one would expect from this type of meteorological feature. We do discuss IPSL-CM5A2 anomalies significance in our revised manuscript on lines 143-147 and 266-270:

"Precipitation decrease (significant at the 90% level) over the American southwest, with cumulative precipitation for DJF decreasing by 10% (not shown), and increase in the northwest is suggested in all simulations, although the magnitude and significance of rainfall changes is weaker in EC-Earth3.2 and IPSL-CM5A2 compared to the CCSM4-SO."

"This in part explains weaker tropical Pacific convection changes (and in turn Rossby waves response) in the EC-Earth3.2 and IPSL-CM5A2 models compared to CCSM4-SO slab ocean simulations: less tropical TOA (and OLR) changes are needed for compensation since some of the energy is compensated by the surface processes."

References

Blackport, R., & Kushner, P. J. (2017). Isolating the Atmospheric Circulation Response to Arctic Sea Ice Loss in the Coupled Climate System, Journal of Climate, 30(6), 2163-2185.
Blackport, R., &Screen, J. A. (2019) Influence of Arctic sea ice loss in autumn compared to that in winter on the atmospheric circulation. Geophysical Research Letters, 46, 2213-2221.https://doi.org/10.1029/2018GL081469
Liu, W., & Fedorov, A. V. (2019). Global impacts of Arctic sea ice loss mediated by the Atlantic meridional overturning circulation. Geophysical Research Letters, 46, 944– 952. https://doi.org/10.1029/2018GL080602
Peings Y., Z.M. Labe and G. Magnusdottir (2021) Are 100 ensemble members enough to capture the remote atmospheric response to +2°C Arctic sea ice loss? Journal of Climate, DOI:10.1175/JCLI-D-20-0613.1
Smith, D. M., Screen, J. A., Deser, C., Cohen, J., Fyfe, J. C., García-Serrano, J., Jung, T., Kattsov, V., Matei, D., Msadek, R., Peings, Y., Sigmond, M., Ukita, J., Yoon, J.-H., and Zhang, X. (2019) The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification, Geosci. Model Dev., 12, 1139–1164, https://doi.org/10.5194/gmd-12-1139-2019.
Sun L., Alexander, M., & Deser, C. (2018) Evolution of the Global Coupled Climate Response to Arctic Sea Ice Loss during 1990–2090 and Its Contribution to Climate Change, Journal of Climate, 31(19), 7823-7843.
Wilks, D. S. (2016). "The Stippling Shows Statistically Significant Grid Points": How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It, Bulletin of the American Meteorological Society, 97(12), 2263-2273.

Reviewer #2 (Remarks to the Author):

Summary:

The paper provides results from 3 global climate models that are perturbed to extract the "sea-ice" signal on northern hemisphere atmospheric circulation characteristics. Perturbation experiments are conducted by changing model parameters to achieve substantial (summer/fall) sea ice reductions relative to their control experiments. The impact on northern hemisphere circulation features (height, slp, and precipitation) is presented. An analysis of (some) of the dynamical mechanisms underlying the proposed teleconnections is given. Results are also compared to observed (reanalysis) changes in northern and similarities between simulated responses to sea ice loss and changes over the period 2009-2018 are postulated. The results show a "shared" weakening of the Aleutian Low, ridging over the North Pacific and associated drier conditions in the southwestern United States. Wetter conditions are found over the Mediterranean.

Review:

The paper is generally well-written and easy to follow and potentially contributes to an increasing body of research on this important topic. As somebody who has followed this discussion over the past 20 years and is suffering from a fair amount of intellectual whiplash as a result, I had some hope that the paper might illuminate the topic in a new way and yield some clarity to my state of confusion. It may not be the authors fault, but this clearly didn't happen. I am as confused as I have been. But science progresses mostly incrementally and the point at which new insights emerge from the labor of hundreds of studies, is hard to predict. As such, with some solidified and/or reduced claims, I believe the paper does represent yet another contribution to this process.

I think the authors could substantially improve the impact of their paper if they tried to put this into the context of currently ongoing efforts to sort out this problem through coordinated experiments, specifically the CMIP6 Polar Amplification Modeling Experiment (PAMIP, Smith et al. 2019). This group has laid out a way to systematically ask questions about the importance of different models, configuration, perturbations etc. If the authors could find ways to place their experiments within that framework, it might give their simulation more context and assure that results are considered in the synthesis that will hopefully emerge from PAMIP. The present study, using multiple models (I believe all of which are already participating in PAMIP) already makes a good move in this direction. I understand that participation and coordination with other groups is not always possible and separate tracks of investigation should be possible, but I think the authors should give this some thought.

Response: *Right, this is an important point. While PAMIP experiments study even shorter timescales than ours (<1 year, initial (fast) response to sea ice forcing) we share a common goal of understanding the multimodel differences and PAMIP is one of the state-of-the-art efforts in this direction. We have added citation and discussion of PAMIP experiments throughout the manuscript, main differences and complementary points are discussed on lines 398-401 and 420-424:*

"Other than the PAMIP initiative that provides a common framework for testing the responses to sea-ice loss in a multi-model setup [67] , the existing sea-ice studies feature a number of methodological differences making it difficult to discern the exact causes of the discrepancies in simulated responses to sea-ice loss.

…
PAMIP style experiments represent a very important step in overcoming these issues. However, current PAMIP protocol focuses on the initial arrival of sea ice induced perturbations by using a large ensemble of very short (up to 1-year long) simulations. Physically, the response shown in this study represents the next phase, after the initial signal of sea-ice forcing has already arrived and got established. "

I have several major points of criticism that think the Authors should consider:
1. The overall point of differentiation from previous studies for the paper is given as "decadal" versus "centennial" for previous studies. If find this distinction not well explained. The perturbation experiments are not tied to any time scale. The "summer ice free conditions" as achieved in two of the models are still many decades away in others. They depend on the sea ice sensitivity of the particular model so tying this to a particular timescale is confusing. The comparison of the sea ice impact with respect to the last decade 2009-2018 relative to 1979-2008 is indeed decadal but I find this section actually to be the weakest part of the paper.

Response: *Thank you for bringing this to our attention, this is a really important aspect and we didn't communicate it clearly enough. The decadal and centennial distinction was intended to refer exclusively to how long the models were run after imposing the sea ice perturbation independently of the targeted sea ice conditions. This distinction is very important because of the time needed for development of different atmospheric and oceanic responses. For example, deep ocean responses likely play a more important role on centennial than on decadal time-scales, atmospheric responses happen faster than oceanic... We have added this text to our manuscript to clarify on lines 64-69:*

"Physically, "decadal" responses are dominated by faster adjustments of the climate system to sea-ice changes (e.g., atmospheric reorganizations and fast oceanic responses), while the "centennial" responses allow for slower deep ocean feedbacks to develop too. The distinction between "decadal" and "centennial" time-scales in this context is based on the time the sea-ice forcing has been applied for before studying the resulting climate response."

The Discussion section of the paper has also been majorly rewritten, so we invite the reviewer to re-consider the whole section.

2. The analysis is in large parts based on the subjective assessment of similarity of circulation patterns between different models (and ERA-5). Sometimes the similarity is labeled as "remarkable". When pointed to specific features (e.g. the North Pacific High, Aleutian Low), I can recognize some of the similarities. But a less guided inspection could yield a very different conclusion: For example, SLP anomalies over the Arctic Ocean are vastly different. With EC-Earth showing a low, IPSL a high, and CCMS4 a bit of both? This makes me wish for a more objective measure of similarity or dissimilarity. A pattern correlation method comes to mind. Also, atmospheric circulation patterns with high spatial auto-correlation have a tendency to look meaningful, so I think a comparison should be tied to a reference to determine what constitutes a meaningful difference (level of significance). How does a difference between models compare to a random perturbation? The literature on "field significance" maybe a useful guide. Similarity measures for pattern differences between intra ensemble members (which constitute random variation) might be another way to develop a reference.

Response: *Yes, we understand where this concern is coming from, we should have been more careful in clarifying our goals to justify the study design and the analysis.*

In the first place, we wanted to prove that the atmospheric response to sea ice loss, consisting of North Pacific geopotential ridge (positive geopotential height anomaly) and drying over California, found in Cvijanovic et al. 2017, Simon et al. 2020 and Sidenglanz et al. 2021 is reproducible by any model as long as the energy conserving methodology for isolating the impacts of sea-ice loss is used and the relatively short time-scales are considered. This is because the majority of other studies are showing the opposite response (negative geopotential height anomaly) and we wanted to provide the context in terms of time-scales and methodology to explain the different responses found in different sea ice studies: it is not random, it is a consequence of how the experiments were designed and run. We wanted to improve the state of knowledge so that it is not just "different studies are arriving to different conclusions and there is no consensus", but that there is a reason to why and how different studies arrive at different conclusions, and these different responses are actually applicable to different situations (decadal vs. centennial response).

So our goal number 1 is to arrive to a consistent North Pacific geopotential response. By this we mean that we are looking for the same sign of geopotential height anomaly in the North Pacific, while being aware that its exact position/strength will vary between the models, especially given the different magnitudes of sea ice forcing.

Our second goal, the second main pattern we are looking for, is a Rossby wave starting in the tropics and driving the North Pacific geopotential ridge. We needed to see if it shows up in each ensemble mean.

Our third key question is if California is drier or not in all three models.

Due to intermodel differences and also due to non-identical sea ice forcing, we know that the responses cannot be exactly the same between the three models, what we show is that the positive North Pacific geopotential height anomaly (as opposed to a negative one) appears each time for the same reason and with the similar consequences (and this occurs despite the differences in sea-ice forcing).

The pattern matching is really attractive but it is not easy to apply at this stage since our experiments vary in the amount of sea ice forcing. We have instead added the locations of key simulated features in the figure showing reanalysis data (Figure 6), so that a better visual comparison can be achieved.

We do thank the reviewer for this comment as it set a really interesting goal for our future work regarding how we could better define the matching patterns and use them as a fingerprint in detection of the sea-ice changes. For this purpose we will redesign and rerun our simulations to force them with more similar sea ice forcing and then perform the pattern matching.

3. The explanation of the link between the high latitude perturbation and the hemispheric response seems truncated. The paper shows some credible analysis demonstrating the link from the tropics to the high(er) latitudes but the initial branch from the sea ice perturbation to the tropics is left to some references. Given the what others have termed the "pre-consensus" state of affairs in this space, I think it would be helpful if analysis specific to the model experiments presented here could confirm how the initial sea ice perturbation does propagate into the tropics. It seems particularly important also to identify which regions of sea ice loss dominate the lower latitude impact. For the comparison with the observed response, I note that the regions showing

ice loss in the models have only relatively minor losses in the observations over the 2009-2018. In addition, as stated above, the Arctic Ocean, from where the lower latitude changes are supposed to be propagating shows rather different SLP anomaly patterns.

Response: *The reviewer is right, this section focused too much on the processes and time periods addressed in other studies that our setup was not designed to answer. Our experiments were designed to study the decadal response (when the initial impacts of sea ice loss have already established) and not the initial arrival of sea ice forcing investigated in PAMIP.*

Responses we try to capture require at least 200 years to achieve the statistical significance. In our case this meant performing 10 ensemble members each run for 30 years and then considering the last 20 years in the analysis. In contrast, to study initial progression of sea ice impacts we would need PAMIP style experiments, so one year long simulations but 100 or 200 ensemble members (to achieve at least 100 or 200 years for the analysis). We cannot extract the information on initial propagation from our setup, as we don't have enough ensemble members to focus on a single year. The Discussion section has been rewritten to clearly separate from what is done in this study and what is out of scope.

Then we focus more deeply on the decadal scale response that should be addressed in our analysis. For this purpose we added one more figure explaining the energetics of the decadal response: we compare annual zonal TOA net flux changes and atmospheric net column flux changes. This comparison shows a very interesting difference between the slab ocean and fully coupled experiments: in slab ocean experiments where the deep ocean cannot compensate for the high latitude changes, the compensations for high latitude changes happen predominantly in the tropics, through TOA response (convection and OLR changes). In the fully coupled experiments however, compensation in ATM column flux does not happen so much from TOA process, but also due to surface processes. This explains the somewhat weaker response in Rossby waves triggered by the tropical Pacific convection changes.

Please see new Figure 5 and the discussion in regards to it on lines 236-270.

We have also included discussions of two previous studies: Levine et al. 2021 and Seidenglanz et al. 2021 that study this initial response by imposing the regional sea ice loss in their experiments: due to our experimental setup that encompasses all of the Arctic, it is not possible to clearly distinguish which region of sea ice loss is the most important in our experiments; however, the key features of our response are typical for the one produced when using Barents-Kara Seas sea ice anomalies alone (Levine et al. 2021 and Seidenglanz et al. 2021).

Lesser points:
I hope the figure size will be increased in production. They are very hard to see for aging eyes.

Response: *We believe so. And we have increased font sizes wherever possible.*

A note to what degree the experiments constitute a "sea ice only" vs. a "sea ice + SST" response might be helpful. Since SST at high latitudes can respond to the reduced sea ice by warming, it appears that the experiments show a combined SIC/SST response?

Response: *You are correct, we have added the following sentence to clarify this on lines 226-229:*

"While the sea ice loss is the initiator, the consequent tropical Pacific SST and surface flux changes are crucial for the development of the global response described in our study. In that sense, our experiments show a combined response to Arctic sea-ice and Pacific SST and surface flux changes."

Some discussion putting the magnitude of the response into context would be useful. Notably, the SLP response is given in Pa for the model results, while in hPa for the observations (looks like a factor of 30 difference judging from the color scale). The response in terms of annual precipitation amounts (%winter or annual?) would put the chain of responses into something more tangible (eyeballing form some papers it looks like the response for California is about 10% of winter precip for example)

Response: *Instead, we changed the SLP scale in Figure 2d-f to be in hPa too, to make the comparison more obvious. We have also added information about cumulative precipitation changes over the American southwest on lines 143-145.*

"Precipitation decrease (significant at the 90% level) over the American southwest, with cumulative precipitation for DJF decreasing by 10% (not shown), and increase in the northwest is suggested in all simulations..."

Appreciating the differences between models could be improved if results from other models could be over plotted in Fig 1.

Response: *We tried several combinations of colors to combine panels a,b,c into one, but each time the panel looked very crowded. We have now increased them by about 50%, this hopefully helps the overall visibility. We will also make sure that they are maximized once we are at the final production level with the publisher.*

Line 112 " the most prominent"
See earlier comment on making such statements more objective

Response: We rephrased:
"A notable winter Pacific response " on lines 130.

Line 136: "anomalous propagation of wave activity"
I don't see this for EC-Earth?

Response: That is true, we have rephrased the sentence on lines 154-157:

"*These show an anomalous propagation of wave activity from the Pacific sector towards North America in response to the reduced Arctic sea ice cover in CCSM4-SO and IPSL-CM5A2. In EC-Earth wave activity propagation is much less clear at this level.*"

Line 138: "Previous work has suggested this Rossby wave pattern"
The authors might want to cite additional literature on this (e.g. Ding et al. 2019, Gong et al. 2020). In particular, it would be helpful to put the present result into the context of work that has argued that atmospheric circulation variability in the Arctic is the response to internal SST variability in the tropical Pacific. Some words on the Author's perspective on how their results fit into the debate for tropical vs. Arctic forcing might be useful. Seems like it doesn't have to

either/or and the present work lends some credence to the idea that those local vs remote processes could be reinforcing each other?

Response: *We very much agree with the reviewer on the idea that processes could be reinforcing each other. We have added the following paragraph on lines 226-234:*

"While the sea ice loss is the initiator, the consequent tropical Pacific SST and surface flux changes are crucial for the development of the global response described in our study. In that sense, our experiments show a combined response to Arctic sea-ice and Pacific SST and surface flux changes. Some studies have suggested that atmospheric circulation in the Arctic is driven by the internal SST variability in the tropical Pacific [36,37] . The present work supports the idea that it may not have to be one or another (high latitudes driving tropical changes or tropics driving the high latitudes), but that once the changes are initiated by the tropics or high-latitudes they could potentially be reinforcing each other."

Line 186.
That's a nice result!

Response: *Thank you!*

Line 200, "Atlantic sector is featuring the largest differences"
Looks to me that there are multiple regions with large differences (e.g. Arctic Ocean). Goes to prior comment on more objective measures of similarity.

Response: *We have removed this sentence.*

Line 213 "…Atlantic region is additionally shaped by AMOC"
Isn't this a bit of a stretch from looking at two coupled models vs. one swampy one?

Response: *That is true, we made a more general statement about CCSM4-SO missing the deep ocean response on lines 288-289:*

"*It is possible that the atmospheric response over the Atlantic region is additionally shaped by emerging ocean feedbacks that are not simulated in a CCSM4-SO setup.* "

Line 217 " Wetter conditions are consistently found over the western Mediterranean"
True, but there are lots of blues and reds on that map, why discuss this?

Response: *We rephrased this sentence to justify the mention of the Mediterranean as the only region in Europe consistently significant in all three models, now on lines 291-294:*

"*Over the European continent, the only region where statistically significant (at the 90% level) precipitation changes are consistently found in all three model ensembles is the western Mediterranean, where the wetter winter conditions are suggested (Fig. 2g-i).*"

Line 225 "Recent Decade"
I understand the motivation to try to link modeled sea ice response signatures to the real world, but I find the whole section not very convincing. First, for average NH ice concentration the

period from 2009-2018 hasn't been one of rapid loss (e.g. Baxter et al. 2019 , Francis et al. 2020). Moreover, regions that show sea ice loss in the models show relative weak sea ice losses in the observations (Fig 5a, 1 d,e). (This also emphasizes further the need to identify specific regions and mechanisms of how the sea ice lower latitude coupling is working in the presented simulations)
Nevertheless, despite the relatively small perturbation in the real world, the atmospheric response supposed is at least an order of magnitude larger than the modeled response. I think this section need to be reworked substantially or maybe cut altogether. The argument that the discrepancy in magnitude is expected because other factors are at work is of course true, but also a bit of an unfalsifiable "catch all". A more objective "detection" technique is probably needed to make this more convincing. I would be surprised if this were possible to show. While cutting this section may feel like a loss, I think a modeling paper by itself could be a useful contribution. A workable alternative would be to heavily caveat this section.

Response: *We have substantially shortened and de-emphasized this section. We leave it to reviewer to decide if it is useful to keep in its current form or completely remove.*

Supplement:
Fig 1 missing panel labels a-i. Anomaly reference period should be stated in caption.
Response: *We have now included the panel labels to Supplementary Fig. 1. These are the anomalies from the simulations (low-ice minus control) and it is indicated in the caption: "Monthly mean sea-ice concentration anomalies (low ice minus control) [%] for December (first row), January (second row) and February (third row)."*

I hope you will find these comments constructive even though they may appear very critical. I recognize that there is a lot of work behind the simulations and analysis the paper will likely make a solid contribution after some of the criticisms are addressed.

Response: *We are very grateful for the comments received, they are constructive and we can clearly see that the reviewer honestly tried to help us improve our study.*

References:
Smith, D. M., J. A. Screen, C. Deser, J. Cohen, J. C. Fyfe, J. García-Serrano, T. Jung, V. Kattsov, D. Matei, R. Msadek, Y. Peings, M. Sigmond, J. Ukita, J. H. Yoon and X. Zhang (2019). "The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification." Geosci. Model Dev. 12(3): 1139-1164.
Francis, J. A. and B. Wu (2020). "Why has no new record-minimum Arctic sea-ice extent occurred since September 2012?" Environmental Research Letters 15(11): 114034.
Ding, Q., A. Schweiger, M. L'Heureux, E. J. Steig, D. S. Battisti, N. C. Johnson, E. Blanchard-Wrigglesworth, S. Po-Chedley, Q. Zhang, K. Harnos, M. Bushuk, B. Markle and I. Baxter (2019). "Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations." Nature Geoscience 12: 28-33.
Gong, T., S. B. Feldstein and S. Lee (2020). "Rossby Wave Propagation from the Arctic into the Midlatitudes: Does It Arise from In Situ Latent Heating or a Trans-Arctic Wave Train?" Journal of Climate 33(9): 3619-3633.
Baxter, I., Q. Ding, A. Schweiger, M. L'Heureux, S. Baxter, T. Wang, Q. Zhang, K. Harnos, B. Markle, D. Topal and J. Lu (2019). "How Tropical Pacific Surface Cooling Contributed to Accelerated

Sea Ice Melt from 2007 to 2012 as Ice Is Thinned by Anthropogenic Forcing." Journal of Climate 32(24): 8583-8602.

Reviewer #3 (Remarks to the Author):

Review of Cvijanovic et al. 2021

Summary
The authors revisit the now frequently discussed (but largely unresolved) topic of hemispheric-scale atmospheric circulation and teleconnections to recent sea ice loss. In this study, the authors use an energy-conserving method to tackle this question in a way that allows the ocean to respond dynamically in remote regions. This approach is substantially different from prior investigations in the literature, and comes to substantially different conclusions. Here, the authors find a strong hemispheric scale response to Arctic sea ice loss that includes a weakening of the semi-permanent Gulf of Alaska low and a negative NAO-like response in the North Atlantic—modestly reducing winter precipitation in California and parts of the Southwestern US and increasing in across the Western Mediterranean. The findings are nearly opposite of those in the existing literature regarding the response to ice loss and/or climate change more broadly, and the authors discuss possible reasons for this departure—primarily focusing on a 2-step teleconnection involving tropical SST modulation and the generation of a basin-scale Rossby wave train.

Overall Assessment
As someone who has followed the "Arctic sea ice-Mid-latitude weather" debate closely over the past decade, my impression is that this paper is topically well-targeted and offers an important new perspective (and methodology), as well as some provocative new findings. Thus, from a novelty and scope perspective, this work is well-suited to CE&E. However, there remain some important outstanding questions regarding the broader conversation to which this work contributes--as well as the context of the findings that contrast with prior work by other authors-- that should be more directly and substantively addressed than in the present draft. Additionally, while the authors focus on winter teleconnections, the primary ice loss in the Arctic is actually occurring in autumn—so some additional analysis regarding the simulated response in other calendar seasons (especially autumn) would be warranted to more clearly illustrate the seasonal progression and any potential seasonal asymmetries of the atmospheric response to ice loss.

Altogether, I am optimistic that this paper can become an excellent candidate for publication in CE&E following some substantial revisions (more specific feedback enumerated below).

Specific Comments
1. The present title is probably not specific/descriptive enough. One possible suggestion could be something like: "Near-term sea ice loss drives strong regional atmospheric response over North Pacific and North Atlantic."

<u>Response:</u> *We have modified the title to "Arctic sea-ice loss drives strong regional atmospheric response over North Pacific and North Atlantic on decadal scales."*

2a. One key weakness of the manuscript in its current form is the relative lack of context into which these results, which are effectively an idealized sensitivity experiment testing the

atmospheric response to sea ice loss in the absence of any other external/internal forcings, are placed. The authors point out that the simulated response yields anomalous North Pacific ridging (and Atlantic troughing) that reduces American Southwest (and increases Western Med) precipitation in winter—which is exactly the opposite response demonstrated in some prior sea ice sensitivity experiments as well as being contrary to the overall net response to greenhouse forcing in CMIP5-type ensembles (which show slight precipitation increases in California and stronger precipitation decreases in the Mediterranean in winter).

I believe the authors already make a sufficiently plausible argument why their results here are contrary to other sea ice sensitivity experiments (a more comprehensive level of atmosphere-ocean coupling; near-term vs long-term response, etc.). But the authors barely even mention the broader aspects of their findings in the context of the full forced response to GHG warming in multi-model ensembles. In the real world, Arctic sea ice loss is not occurring in isolation—CO2 is also rising, and the oceans are warming more broadly (among other changes).

I think it would be very important for the authors to discuss this context, especially as it relates to the fact the even if the modeled response in this experiment turns out to be 100% correct, it is possible that the net response will be overwhelmed by other factors. This is bolstered by the facts that 1) while precipitation decreases in the present experiment are statistically robust, they are generally fairly small in magnitude, 2) that prior research has demonstrated a mean "El Nino"-like response of the tropical Pacific to CO2 forcings that results in opposite teleconnections to the ones demonstrated here, (and 3) the much discussed idea that there is/will be an Arctic-tropical "tug of war" in a warming climate (e.g., Barnes et al. 2015). Perhaps the authors here are suggesting that the near-term balance is more likely to be won by the Arctic than the tropics, at least on a regional basis? If so, that kind of explicit framing would be useful. (In addition to discussion in the main text, this context would also be very helpful in a revised abstract.

Response: *We thank thank the reviewer for this comment. This is a well-thought point, we focused on the specific type of response without enough reflection on how it fits with other ingredients of climate change.*

We used the 'last decade' section, that needed substantial rewriting to bring our results into the real world context (now on lines 340-386):

We have also revised the abstract on lines 47-51:

"Although in the overall context of climate change sea-ice loss is just one of the factors capable of driving remote atmospheric responses, its potential to majorly affect the precipitation over the western United States should be carefully considered in mitigation and adaptation strategies, especially given the model disagreement in future projections of Californian rainfall."

2b. An additional question that comes to mind: is the effect of ice loss (particularly across North Pacific) stronger in weak ENSO vs strong ENSO (either warm or cool phase) years, which are known to have strong remote effects on both North Pacific and North Atlantic circulation? This is especially relevant given that the proposed mechanism operates via remotely SST-forced Rossby wave propagation out of the tropics--similar to atmospheric teleconnections from ENSO.

Also, could this emerging evidence regarding the potentially substantial role of sea ice loss on regional hydroclimate along the North American Pacific coast comprise a new seasonal

predictive factor comparable to existing predictors like ENSO? Might one expect the effect of ice loss to be more pronounced in weak ENSO conditions, and muted during strong El Nino events? If so, that could potentially offer increased seasonal predictability if the state of sea ice can be known. I raise 2b mainly as "food for thought," but I would hope to at least see some kind of substantive response to 2a.

Response: *We have added a new figure (Supplementary Fig. 9) showing the 250hPa streamfunction anomalies in the slab-ocean low ice simulation for strong El Niño and strong La Niña conditions, as well as a short discussion on the potential interplay between ENSO and Arctic sea-ice induced teleconnection on lines 350-369:*

"For example, El Niño-Southern Oscillation (ENSO) is known to have strong remote effect on both Pacific and North Atlantic circulation and could potentially affect the Arctic sea-ice loss induced Rossby wave train. In that sense La Niña conditions could strengthen the sea-ice impacts and El Niño weaken it. In fact, our simulations show that during a strong La Niña event the response in the low ice simulations is very similar to the mean response shown in Figure 3, but it is significantly stronger (Supplementary Fig. 9a). However, during a strong El Niño event the response in the North Pacific and North Atlantic is the opposite to the mean response shown in Figure 3 (Supplementary Fig. 9b) in agreement with the typical teleconnection pattern during warm ENSO events [62] . References [63-65] found that recent observed trends for strengthening of the Walker circulation in the tropical Pacific and of the east-to-west surface temperature gradient are due to the ocean-thermostat (OT) effect in combination with the effect of aerosols. They showed that some CMIP6 models that correctly simulate these two effects indicate an initial (20-100 years) La Niña-like mean response to global warming in the tropical Pacific, which is gradually replaced by the consensus El Niño-like response characterized by a weakened temperature gradient that will start to appear in the longer term (in 50-100 years) after the leveling off of atmospheric aerosols and the overcoming of the OT effect. This suggests that over the forthcoming decades the La Niña-like response to Arctic sea-ice loss on the western United States precipitation that we found here might be dominating."

3. Despite the analysis here and even after reading the supplementary discussion, I am still not really clear on the hypothesized physical linkage between the actual sea ice loss and its modulation of tropical SSTs. I understand how the atmospheric response develops once the tropical SSTs are perturbed, but I still don't understand the actual means by which the sea ice loss signal moves from the Arctic to the tropics.

Response: *Prompted by this comment and a comment from another reviewer, we have majorly rewritten this part of the manuscript. The issue is that our experiments were designed to study the decadal response (when the initial impacts of sea ice loss have already established) and not the initial arrival of sea ice forcing. By discussing the arrival of anomalies (citing other studies) we have confused the reader about what our setup is supposed to answer.*

Responses we try to capture require at least 200 years to achieve the statistical significance. In our case this meant performing 10 ensemble members each run for 30 years and then considering the last 20 years in the analysis. In contrast, to study initial progression of sea ice impacts we would need PAMIP style experiments, so one year long simulations but 100 or 200 ensemble members (to achieve at least 100 or 200 years for the analysis). We cannot extract the information on initial propagation from our setup, as we don't have enough ensemble members to focus on a single year. The Discussion section has been rewritten to clearly separate from what is done in this study and what is out of scope.

In the revised version we refer the reader to the studies focusing on the initial response (PAMIP, Seidenglanz 2021, Levine et al. 2021). Then we focus more deeply on the decadal scale response that should be addressed in our analysis. For this purpose we added one more figure explaining the energetics of the decadal response: we compare annual zonal TOA net flux changes and atmospheric net column flux changes. This comparison shows a very interesting difference between the slab ocean and fully coupled experiments: in slab ocean experiments where the deep ocean cannot compensate for the high latitude changes, the compensations for high latitude changes happen predominantly in the tropics, through TOA response (convection and OLR changes). In the fully coupled experiments however, compensation in ATM column flux does not happen so much from TOA process, but also due to surface processes. This explains the somewhat weaker response in Rossby waves triggered by the tropical Pacific convection changes.

Please see the new rewritten section "Compensating for the sea-ice loss induced high-latitude energy imbalance" on lines 236-270 and the new Figure 5.

4. Recent work on the "Arctic/mid-latitude" question often hinges on the existence and/or accurate simulation of the "stratospheric pathway" (whereby vertically propagation of energy triggered by sea ice loss can disrupt the stratospheric and subsequently the tropospheric polar vortex, e.g., (Hoshi et al. 2019). The authors don't really mention this literature or its potential relevance to the present study—which I think is a conspicuous absence. (And perhaps this could be related to #3 above?)

Response: *We included the citation to Hoshi et al. 2019 and acknowledge that the stratospheric pathway could be playing a role too (on lines 377-380):*

"Finally, it is useful to note that Arctic sea-ice loss has been suggested to affect the planetary-scale wave propagation during the weak polar vortex events through a stratospheric pathway [27]. Further examination is needed to clarify the possible interplay between the stratospheric and the tropical surface processes described in this study."

5. Line 306: California mean annual average precipitation projections are actually clustered pretty close to zero across most multi-model ensembles (despite large projected increases in precipitation variability and seasonality (e.g., Swain et al. 2018, Dong et al. 2019))—I have no idea where these numbers are coming from. This relative lack of change in regional precip has actually been suggested to be the net result of a "tug of war" by competing wetting/drying influences—which perhaps could account for either the muted mean signal, the projected increase in variability/seasonality, or both?

Response: *We have added a citation to previously mentioned reference by Polade et al. 2017 and included a recent citation to Petrova et al. 2024 (see lines 446-450) to make it clearer where these numbers came from. In Fig. 1B in Polade et al. 2017 you can find these intermodel ranges for CMIP5 winter precipitation and in Fig. 3 in Petrova et al. 2024 you can find the ranges for both CMIP5 and CMIP6 separated for Northern and Southern California. As you correctly stated the ensemble mean change is close to zero, but the range of possible outcomes is rather wide. We wanted to stress this fact, because it is a really major model disagreement and we think this could potentially be important in the context of sea-ice loss too (models predict very different rates of sea-ice loss and sea-ice loss affects California's rainfall).*

"...while California's precipitation, influenced by changing midlatitude and tropical circulation patterns, has been suggested to change anywhere between -30% and +75% by the end of century (2060–2089 relative to 1960–1989) [73] *, and for southern California the range is even wider -20% to +80% (2080-2099 relative to 1980-1999)* [77] *.*

6. Potentially of note: the spatial pattern of GPH change in Fig 5c (West Coast US ridging but Hudson Bay low anomaly) is reminiscent of the "North American Winter Dipole" pattern that has increased in frequency/intensity during the recent period of accelerated Arctic sea ice loss (Wang et al. 2015, Singh et al. 2016).

Response: *Thank you, this is interesting to note, we have added it to our text on lines 333-337:*

"Aloft, geopotential height increases in the North Pacific and Barents-Kara-Greenland seas regions and decreases over north-east United States and western Europe (Fig. 6c). This pattern is reminiscent of the "North American Winter Dipole" that has increased in frequency and intensity during recent decades [55,56]*."*

7. The greatest absolute/relative magnitude of Arctic sea ice loss is actually in autumn, rather than winter, as the authors point out. But the authors focus on winter atmospheric circulation. Given the temporal lag here, and the growing interest in California/Southwestern U.S. seasonal drying, it would be extremely interesting to see a seasonal progression of GPH500 anomalies and/or precipitation anomalies from these ice loss experiments (especially for Sep-Nov). Might this explain why California is now experiencing a robust autumn drying trend (but not winter drying trend) in recent decades? (Luković et al. 2021) It's possible that even if the winter drying trend caused by sea ice is ultimately not overwhelmed by competing forcings, it could still be relevant in the autumn (with major implications for wildfire risk, etc.)

Response: *Yes, while the greatest magnitude of Arctic sea ice loss is in September in our simulations, California's precipitation response peaks in winter in all three models. The autumn (or any other season's) response is mostly not statistically significant. In terms of Lukovic et al. 2021 study (that we were very excited about) an opposite response is found in our simulations - November becomes wetter (but also mostly not statistically significant). So it seems there is some other factor than the sea-ice loss that affects autumn patterns observed by Lukovic et al. 2021. An earlier study by Deser et al. 2010 also implied that the largest response in the mid-latitude atmospheric circulation to sea-ice loss is lagged and is expected to be in winter. We note in our text this discrepancy between the highest sea-ice loss and the timing of strongest impacts on lines 128-130:*

"*While the greatest magnitude of Arctic sea-ice loss in our model simulations is observed in autumn, its remote impacts occur with some delay and are most pronounced during the winter season."*

8. Line 574: reference error.

Response: *Thank you we have corrected it.*

References

Barnes, E. A., & Screen, J. A. (2015). The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it? WIREs Climate Change, 6(3), 277-286. https://doi.org/10.1002/wcc.337. https://doi.org/10.1002/wcc.337

Dong, L., Leung, L. R., Lu, J., & Song, F. (2019). Mechanisms for an amplified precipitation seasonal cycle in the U.S. West Coast under global warming. Journal of Climate. https://doi.org/10.1175/JCLI-D-19-0093.1
Hoshi, K., Ukita, J., Honda, M., Nakamura, T., Yamazaki, K., Miyoshi, Y., & Jaiser, R. (2019). Weak Stratospheric Polar Vortex Events Modulated by the Arctic Sea-Ice Loss. Journal of Geophysical Research: Atmospheres, 124(2), 858-869. https://doi.org/10.1029/2018JD029222. https://doi.org/10.1029/2018JD029222
Luković, J., Chiang, J. C. H., Blagojević, D., & Sekulić, A. (2021). A later onset of the rainy season in California. Geophysical Research Letters, n/a(n/a), e2020GL09350. https://doi.org/10.1029/2020GL090350.
Singh, D., Swain, D. L., Mankin, J. S., Horton, D. E., Thomas, L. N., Rajaratnam, B., & Diffenbaugh, N. S. (2016). Recent amplification of the North American winter temperature dipole. Journal of Geophysical Research: Atmospheres, 121(17), 9911-9928. http://dx.doi.org/10.1002/2016JD0251
Swain, D. L., Langenbrunner, B., Neelin, J. D., & Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. Nature Climate Change, 8(5), 427-433. https://doi.org/10.1038/s41558-018-0140-y
Swain, D. L., Singh, D., Horton, D. E., Mankin, J. S., Ballard, T. C., & Diffenbaugh, N. S. (2017). Remote linkages to anomalous winter atmospheric ridging over the northeastern Pacific. Journal of Geophysical Research: Atmospheres, n/a-n/a. http://dx.doi.org/10.1002/2017JD026
Wang, S.-Y. S., Huang, W.-R., & Yoon, J.-H. (2015). The North American winter 'dipole' and extremes activity: a CMIP5 assessment. Atmospheric Science Letters, 16(3), 338-345. ://WOS:000358020000023
Francis, J. A. and B. Wu (2020). "Why has no new record-minimum Arctic sea-ice extent occurred since September 2012?" Environmental Research Letters 15(11): 114034.
Ding, Q., A. Schweiger, M. L’Heureux, E. J. Steig, D. S. Battisti, N. C. Johnson, E. Blanchard-Wrigglesworth, S. Po-Chedley, Q. Zhang, K. Harnos, M. Bushuk, B. Markle and I. Baxter (2019). "Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations." Nature Geoscience 12: 28-33.
Gong, T., S. B. Feldstein and S. Lee (2020). "Rossby Wave Propagation from the Arctic into the Midlatitudes: Does It Arise from In Situ Latent Heating or a Trans-Arctic Wave Train?" Journal of Climate 33(9): 3619-3633.
Baxter, I., Q. Ding, A. Schweiger, M. L’Heureux, S. Baxter, T. Wang, Q. Zhang, K. Harnos, B. Markle, D. Topal and J. Lu (2019). "How Tropical Pacific Surface Cooling Contributed to Accelerated Sea Ice Melt from 2007 to 2012 as Ice Is Thinned by Anthropogenic Forcing." Journal of Climate 32(24): 8583-8602.
Screen, J. A., Deser, C., Smith, D. M., Zhang, X., Blackport, R., Kushner, P. J., ... & Sun, L. (2018). Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models. Nature Geoscience, 11(3), 155-163.
Francis, Jennifer A. "Why are Arctic linkages to extreme weather still up in the air?." *Bulletin of the American Meteorological Society* 98.12 (2017): 2551-2557.
Sun L., Alexander, M., & Deser, C. (2018) Evolution of the Global Coupled Climate Response to Arctic Sea Ice Loss during 1990–2090 and Its Contribution to Climate Change, Journal of Climate, 31(19), 7823-7843.
Blackport, R., & Kushner, P. J. (2017). Isolating the Atmospheric Circulation Response to Arctic Sea Ice Loss in the Coupled Climate System, Journal of Climate, 30(6), 2163-2185.
Blackport, R., &Screen, J. A. (2019) Influence of Arctic sea ice loss in autumn compared to that in winter on the atmospheric circulation. Geophysical Research Letters, 46, 2213-2221.https://doi.org/10.1029/2018GL081469

Liu, W., & Fedorov, A. V. (2019). Global impacts of Arctic sea ice loss mediated by the Atlantic meridional overturning circulation. Geophysical Research Letters, 46, 944– 952. https://doi.org/10.1029/2018GL080602

Peings Y., Z.M. Labe and G. Magnusdottir (2021) Are 100 ensemble members enough to capture the remote atmospheric response to +2°C Arctic sea ice loss? Journal of Climate, DOI:10.1175/JCLI-D-20-0613.1

Smith, D. M., Screen, J. A., Deser, C., Cohen, J., Fyfe, J. C., García-Serrano, J., Jung, T., Kattsov, V., Matei, D., Msadek, R., Peings, Y., Sigmond, M., Ukita, J., Yoon, J.-H., and Zhang, X. (2019) The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification, Geosci. Model Dev., 12, 1139–1164, https://doi.org/10.5194/gmd-12-1139-2019.

To be changed in the reference:

Simon, A., Gastineau, G., Frankignoul, C., Rousset, C., & Codron, F. (2021). Transient climate response to Arctic sea ice loss with two ice-constraining methods. *Journal of Climate*, *34*(9), 3295-3310.