

FINAL AUTHOR COMMENTS

wcd-2022-32 Reconciling conflicting evidence for the cause of the observed early 21st century Eurasian cooling (Outten et al.)

We thank Judah Cohen and two anonymous reviewers for their time in reviewing the manuscript and their thoughtful input. Below are their comments in black with our responses in blue. We have written about changes to the manuscript in the past tense, though we realize we will not know if we are invited to submit a revised manuscript until after the handling editor has assessed this FAC. If we are invited to revise the manuscript, we will add the line numbers of the changes.

Response to Reviewer 1 - Judah Cohen

The manuscript tries to reconcile numerous observational analysis studies and model sensitivity experiments of Arctic mid-latitude linkages that offer a wide range of conclusions on whether recent observed Eurasian cooling is related to and/or in response to sea ice loss or the two are coincident in time but unrelated physically and the cooling can be attributed to internal variability of the atmosphere. The authors argue that the modeling and observational studies are not at odds or that we must necessarily conclude either/or that Arctic sea ice loss either contributes to Eurasian cooling or that the cooling is related to internal variability only. Instead, the authors argue that different conclusions can all be at least partially correct and that the cooling can be related to multiple factors at once.

I thought that the discussion of the subject and uncertainty was comprehensive and offered a novel way or at least crystallized the idea better than previous published papers on the subject to frame the debate that can help advance the discussion and how to reconcile all the disparate conclusions. Though it might be obvious that sea ice forcing and natural variability can operate simultaneously, or is not a new idea (this was the thesis of Overland et al. 2021), I think the authors expounded on this idea better than previous studies that I am familiar with. I also thought that the manuscript advanced the conversation beyond Cohen et al. (2020) where it was argued that the conclusion whether sea ice melt can force continental cooling can be grouped into observational and modeling studies and instead the authors here argue that the groups are in reality much more mixed. I also thought that the discussion around Figure 5 that even if sea ice only changed the standard deviation of surface temperatures, that alone can increase the probability of Eurasian cooling even if the sea ice doesn't directly force an atmosphere response conducive to Eurasian cooling.

One suggestion is to maybe shorten the text. I thought that it was a long windup for the punchline. I thought that the concluding remarks were well stated and valuable and I think that it would benefit the reader to get to these important conclusions sooner. But I will not go as far to suggest text to remove and I leave it totally up to the authors.

We appreciate the idea to get the punchline faster, especially for a reader who is very familiar with the field. However this is a review paper and we would like to ensure it contains all of the details so as to be useful for people who know little about the subject. We agree that some of the text is protracted and we have streamlined the text where possible throughout. We have also added suggestions in the paper description at the end of section 1, as well as signposts throughout sections 2 and 3, that readers familiar with the subject matter may skip certain parts (e.g., sections 2.1, 2.2, 3.1, 3.3) without loss of continuity, thus shortening their reading by four to five pages.

As is reflected in some of my minor comments below, I do take issue with this idea that with the inclusion of the most recent observational data, the empirical analysis has come into agreement with the modeling studies that there is no large-scale atmospheric circulation forced response to sea ice variability and that Eurasian cooling has all but disappeared. I am not raising my own paper to require that the authors cite it but rather because it is the paper that most readily comes to my mind that shows trends and the observed relationship between sea ice and large-scale circulation variability over the full reanalysis period. As seen in Cohen et al. 2021 Figure 3, the relationship between sea ice and atmospheric variability remains robust (at least in scale and based on statistical significance) and from Figure S6 Eurasian winter cooling is seen over 41 years of reanalysis and that Eurasia is a clear outlier to the widespread warming elsewhere across the Northern Hemisphere.

We note that the general impression of reviewer 1 that we have gone too far in saying that there is “no” large-scale forced response to sea ice variability is somewhat at odds with the general impression of reviewer 3, who feels we have gone too far in saying that there is some forced response. In any case, we did not mean to convey the message that including the most recent observational data leads us to the conclusion of “no” large-scale forced response, although we see that this misunderstanding was likely due to some of less than careful wording on our part (e.g., L221-222 in the original manuscript where we should not have specified we are talking about trends, and not interannual variability as shown in the correlations of Fig. 3 in Cohen et al. 2021). Hopefully, in the revisions outlined below, we have been able to clarify some of the statements that each of these reviewers object to by providing more explanation and a more balanced discussion on the studies that have backed the various viewpoints. For example, in the response to comment #1 below, we clarify why we say that the Eurasian cooling *trend* (specifically, over what time scale) has largely disappeared.

Thanks for pointing out Cohen et al. 2021, which is certainly relevant and which we missed. It is a useful reference for extreme winters and the role of the stratosphere.

I have some more minor comments below and I recommend that the manuscript be accepted pending minor revisions.

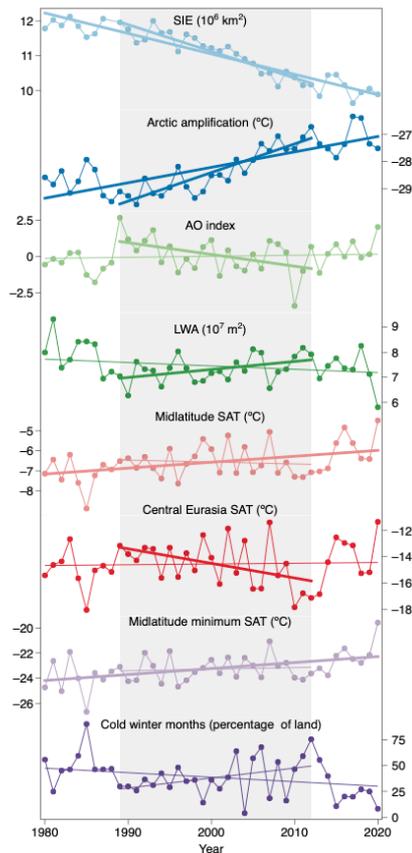
Minor comments:

1. Line 116 – I agree that the Eurasian cooling trend peaked around 2012/13 and has since dampened but I think to characterize it as passed is an overstatement. To expect a perpetual cooling trend is unrealistic given the rapid rise in global temperatures. Winter temperatures in the region of interest remain cooler relative to other regions of Northern Hemisphere and overall cooler than model forecasts. Do you know what else peaked in 2012? September Arctic Sea ice melt, would the authors claim the era of Arctic sea ice melt has “passed?”

There is no doubt that a long-term Eurasian cooling trend shows up over the reanalysis period - Figure S6 of Cohen et al. 2021 shows it for the 41-year period 1980-2021, and the rightmost column of Figure 2 in our manuscript shows it for 30-year periods from 1981 to 2020. But from the columns further to the left, it is clear that the long-term trend is dominated by strong, shorter term trends that are concentrated around the mid-1990s to mid-2010s. Our statement pertains to these strong cooling trends, and this has been clarified. As the reviewer points out, cooling trends should weaken and eventually

disappear under global warming. We have attempted to account for this, because significance is calculated compared to the Northern Hemisphere-averaged warming trend - therefore, even weak cooling trends will show up in Figure 2. This detail is mentioned in the caption but other reviewers missed the information as well, so we have made it more explicit in the text.

Regarding sea ice, we did not mean to suggest that the era of Arctic sea ice melt has passed. The difference between sea ice (top) and Eurasian temperature (third from bottom), in terms of their variability versus trend, is nicely illustrated in the attached figure of Blackport & Screen 2020 (already in the reference list). We have made sure this is clear in the text, and have referenced this figure more explicitly where appropriate.



2. Line 253 – Not sure why only the reference to GAO (2015) is listed, can the authors include a more up to date reference?

The Gao et al. paper does contain a very nice table over several pages which lists a large number of the key studies along with the models and forcings used, but it is indeed an older summary. We have rephrased the sentence to make clear that the Gao et al. paper provide a good list of studies up to 2015, but that there are many newer studies, and we provide a short list of newer references.

3. Lines 407-408 – again I feel that this statement and conclusion presented as fact – “the recent disappearance of Eurasian cooling along with its associated midlatitude circulation signals” is misleading.

This was due to a combination of our wording (not careful enough), some misunderstandings related to Fig. 2, and not indicating where support for parts of this statement come from - all of which are addressed in other parts of this response. We have revised the statement now. Please see responses to earlier comments, plus the response to reviewer #2 comment L219-220.

4. Lines 435-436 – I think to say “it is unfair to discount modeling results as simply wrong” is overly strong. I think a better way of saying something similar like “it is unfair to attribute differences between observed and simulated Eurasian cooling to model errors or deficiencies only.” The exact wording is not important, but I don’t think anyone would argue that the models are deficient to be useless.

We like and agree with the wording you have kindly suggested and have replaced the statement in the paper accordingly.

5. Lines 437-438 – it is my opinion that the physical mechanism can exist in the models and yet the models can still miss much if not all the Eurasian cooling forced by the ice-atmosphere mechanism especially when looking at the ensemble mean.

This is a good point and we agree. The point is raised in section 3.2, but discussion surrounding it has been expanded based on comments from reviewer #3. We have now also added a comment earlier in the subsection 4.2 that the choice of forcing, experimental setup and ensemble size are very important factors in whether a given experiment will “get” Eurasian Cooling.

Judah Cohen

References:

Smith, D.M., Eade, R., Andrews, M.B. *et al.* Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nat Commun* **13**, 727 (2022).
<https://doi.org/10.1038/s41467-022-28283-y>

Cohen, J., L. Agel, M. Barlow, C. I. Garfinkel, I. White. 2022: Arctic change reduces risk of cold extremes—Response, *Science*, **375** (6582), 729-730, DOI: 10.1126/science.abn8954.

Response to Reviewer 2

This paper analyses the extensive scientific debate around the role of Arctic warming (more specifically, localized warming from Arctic sea-ice loss) in recent Eurasian wintertime cooling trends. Although there have been many overview papers on this topic published in recent years, this one is particularly good. That may be because the authors are people who I would not place in one camp or the other. The paper is comprehensive, balanced, and reflective. As well as providing a very nice and useful synthesis of recent studies, it offers a reframing of the question that should provide a constructive way forward on what everybody agrees is an important area of scientific research. The paper correctly notes that a problem with much of the current debate is (i) the failure to acknowledge that a definitive yes-no answer is not possible given all the uncertainties involved, and (ii) the assumed dichotomy between the mean forced response and internal variability, as if they were separable. Yet the wintertime Arctic is arguably the place where the internal variability is most likely to change in response to climate change, and this separation is least defensible.

Effectively, the authors are suggesting a hypothesis that observed trends over a particular period may be primarily attributable to internal variability, but that the internal variability may have changed because of Arctic warming in such a way that the probability of such trends has increased. That is a very novel way of framing the question at hand. It will almost inevitably involve different hypotheses (or storylines) for how the internal variability might have changed, which can be compared in terms of their consistency with data. For this purpose, the proposed emphasis on the distinction between the thermodynamic and dynamical aspects of the problem will be very useful, as the different hypotheses will almost certainly be on the dynamical side. That distinction is not new in climate-change science, but I believe is new (or at least under-utilized) in this particular context.

Overall, this is an excellent and timely paper. I feel that for far too long the debate in this area has been largely sterile, encouraged by certain journals which seem to like papers with titles that are unconditional, to 'stir the pot'. This paper can help set a new tone, and lead to better science. I am happy to recommend acceptance largely as is, with just minor revisions.

Minor comments

line 6: I'm not sure that "coincidental" is the right word. From my understanding, the word can have either an inferential interpretation (by chance) or a descriptive one (at the same time). The latter has no causal implication either way, so is presumably not what is meant here. Do you really mean by chance, or rather that the two features are correlated because of a common driver (atmospheric variability)? (Correlation may not imply causation, but unless it really is by chance, it has to reflect causation somewhere in the system.)

This is correct, coincidental is a bad choice of word here. Upon revision, we have removed the last part of the sentence containing the word since it was unnecessary. Thank you for pointing this out.

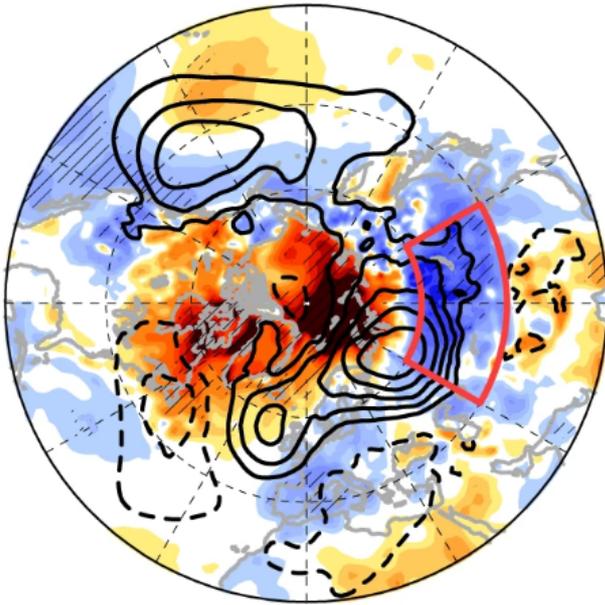
Caption to Figure 1, line 1: Wouldn't it be better to refer to WACE rather than WACC here, since it is Eurasia that is singled out?

This makes good sense and we have changed the text to read "Warm Arctic-Cold Eurasia".

Caption to Figure 1, lines 4-5: Should be "significantly different", not "significant". And not sure what is meant by "insignificant"; do you really mean that there are no trends (of at least 3 hPa/decade) anywhere else on the map?

You are correct and it has been changed to say "significantly different". Note that the figure shows trends which are significantly different from the mean NH trend, which is positive. This was written in the caption but has also been added to the text now We've double-checked the SLP trend contours and they are correct. The 15-year period over which the trends are calculated smooths things out quite a lot in most other regions. Figure 5d from Mori et al. 2019 for the period 1995-2014 DJF using ERA-Interim, pasted below, shows a similar result (contour lines -2, -1, 1, 2, 3, 4, 5... hPa/decade) .

d
ERA-Interim



e

Figure 2: I suspect that some reviewers might complain about the size of these postage-stamp images, but for me it works!

We did consider and test other ways of plotting this figure, e.g. plotting every other start year and period length, thus reducing the number of plots by a factor of four. However, after numerous attempts we decided this was the best way to plot the figure to provide a good overview of the changes in trends, which is the intention of the plot. Since many readers these days will read the paper electronically, the figure can be zoomed in to see some of the details in the individual plots if that is what they are interested in. So far, everyone has liked the postage-stamp plot, perhaps because people are already familiar with the small multiples concept from Ed Hawkins' global warming maps.

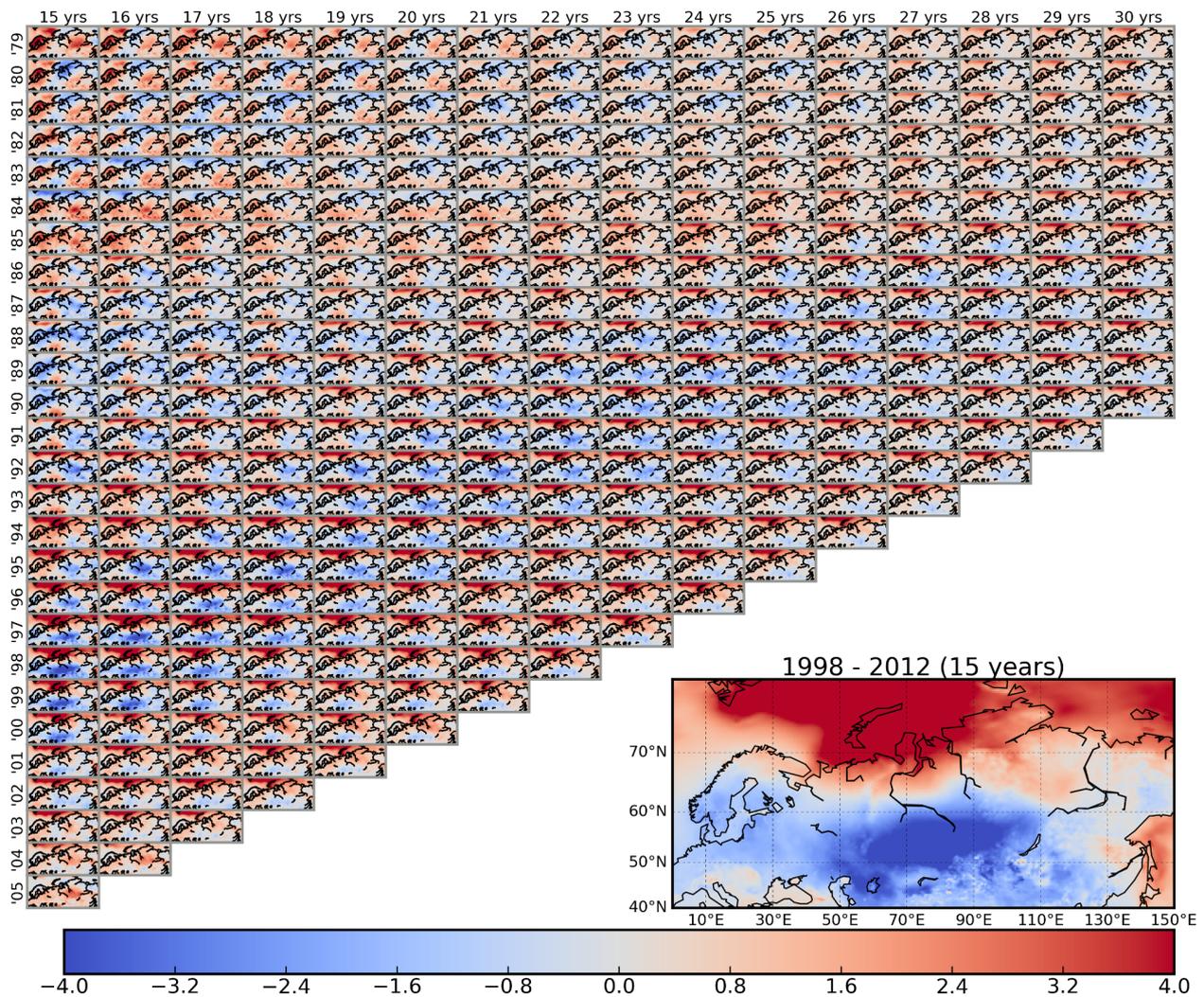
lines 175-179: I think it is only fair to refer to Kretschmer et al. (2021 BAMS) [in your reference list] here, who so far as I know were the first to point out the two rather different definitions of teleconnection in the AMS Glossary.

Thank you for pointing this out. We agree that reference should be made and have therefore added a statement and the appropriate reference.

lines 219-220: The statement "the circulation trends and Eurasian cooling itself have not continued into the most recent decade, while sea ice loss and Arctic warming unequivocally have" seems overcooked. For Arctic warming, it is contradicted by your earlier statement on line 124 that "the Arctic warming trends disappear in the period starting in 2005", which is clearly apparent in Figure 2. For sea ice, it is contradicted by your Figure S3 (as well as by other such figures which one can find on the NSIDC web site). You need to tone this paragraph down.

We did not explain the "disappearing" Arctic warming properly - it is in fact Arctic amplification that disappears. Thanks to the reviewer for noticing this. We have clarified this throughout the paragraph with the contradictory statement (originally L124) and also in other places discussing the Fig. 2. See response to reviewer #1 comment #1 for more detail. We have added a version of Fig. 2 without the significance masking to the supplementary material, and included it below for your convenience.

Regarding sea ice, while there was a peak in sea ice loss around 2012 (as with the Eurasian Cooling trend), it has indeed continued - in a much more consistent manner than EC. In fact, the statement in L219-220 is well supported by Blackport & Screen 2020's figure pasted above (see reviewer #1 comment #1), which we realize we should have referenced here as well. However, we have modified the wording - together with the clarifications described above, this will hopefully give readers a more accurate picture.



line 250: “confidence in their fidelity is high” seems overstated. Such models can give contradictory forced responses of atmospheric circulation to climate change, for reasons that are not difficult to understand (e.g. strong sensitivity to biases in background state), and indeed your later discussion says that. Your wording here suggests that the CMIP models would all give the same forced response to climate change if we only had enough ensemble members, and I think we know that is not true.

This is a good point. We actually highlighted the inconsistency of responses in the previous sentence... so we have now removed the offending phrase and clarified the previous sentence. We should note that other changes have been made in this section in response to reviewer #3's concerns.

lines 407-409: As with my earlier comment with regard to lines 219-220, this statement seems overcooked.

Please see the response to comment regarding L219-220.

lines 518-520: The important point made by Kretschmer et al. (2020) was that the effect of BK sea-ice loss on atmospheric circulation (there examined in terms of the strength of the stratospheric polar vortex) could be small in terms of the year-to-year variability (and thus statistically undetectable from a single realization), yet could have a first-order effect in the climate change response, because of the comparatively large magnitude of the relative change in BK sea ice from climate change. That should be somehow mentioned here as one of the important factors.

This is a good point and thank you for mentioning it. We have included this point into the paragraph.

Typos, etc.

line 73: Presumably “side” -> “wide”

line 86: Something is grammatically wrong with this sentence

line 93: “February” is misspelled

line 94: “Forecasting” -> “Forecasts” (!) [also “Medium Range” -> “Medium-Range”, since it is a proper noun]

line 107: “Eurasian” is misspelled

line 144: “Centre” -> “Center”

Caption to Figure 3: “periods” -> “period”

General: “c.f.” -> “cf.”

We greatly appreciate your pointing out these typos and grammatical errors. These have all now been fixed.

Response to Reviewer 3

This paper reviews the evidence for the causes of the 1998-2012 wintertime cooling over Eurasia. This has been a controversial topic with some studies arguing that the cooling trend was caused by sea ice loss, while others have argued that the trend was caused by internal variability and that sea ice loss has played no role. The authors review the literature, find common ground, examine where disagreements still exist, and provide some guidance for ways forward. They argue that both views can coexist and that whether the cooling trend was caused by sea ice loss or internal variability shouldn't be a 'yes-or-no' question.

Overall, I thought this was a very nice and balanced review of the literature. Although there have been quite a few reviews on the broader topic of Arctic-midlatitude links, the specific topic of Eurasian cooling is certainly worthy of its own review. I think this will be an important paper that will help to motivate future research. I do however, think there are a number of issues that need to be addressed/clarified before publication.

The authors argue that a better question to ask is "whether the cooling trend was more likely given the observed sea ice loss". Does this mean (1) more likely compared to some period prior to the large observed sea ice loss (or preindustrial period), or does it mean (2) more likely compared to a hypothetical world with global warming like we have observed, but with no sea ice loss. These are two different questions with potentially two different answers. It is entirely plausible that the answer to question (1) is no, but the answer to (2) is yes.

These questions are interchanged throughout the paper and in the wider literature. Most modelling studies that are discussed (including the *a* and *b* categories outlined in section 3) are attempting to answer question (2), or whether sea ice loss by itself makes these trends more likely (which is equivalent assuming the responses can be added linearly). However these studies are often misinterpreted as addressing question (1). I would argue that question (1) is more practically relevant.

It is important to note that the reason these questions are different is not just that thermodynamic warming from CO₂ will make the cooling trend less likely. It is also because of dynamical effects. The same models that show a high pressure response over the Urals in response to sea ice loss (sometimes with weak Eurasian cooling) also show a low pressure response (with strong Eurasian warming) in response to CO₂ warming without sea ice loss (e.g. Hay et al. 2022). These dynamical effects tend to oppose each other resulting in weak circulation responses overall in historical simulations and future projections. Thus, for the models where the answer to question (2) is yes, the answer to (1) is still likely no. Some clarification and discussion of these issues are needed.

The distinction between these two different but related questions is a very good point. While those of us who work actively in the topic often take it for granted that it's obvious which we are talking about at any given time, this is not necessarily the case - and certainly even less so for readers who are interested in the topic but not specialists. We have made small modifications throughout the paper, especially sections 3 and 4, in an attempt to clarify which results pertain to which question.

Regarding the Hay et al. 2022 paper, we were not aware of it as it only came out in April of this year, but it is clearly relevant. The setup of the dynamic versus thermodynamic response in section 4.3 as well as other related discussion have been modified accordingly.

Revisions: Since this comment has resulted in many small changes throughout the manuscript, we will not list them all here. It is easiest to check the tracked changes version for the modifications.

The authors argue that whether the cooling trend was caused by sea ice loss or internal variability shouldn't be a 'yes-or-no' causal relationship. However, most modelling studies don't treat it as a 'yes-or-no' question. Most modelling studies just happen to find that the forced response to sea ice loss shows no or very little to cooling, but that does not mean that they did not consider that it could be a combination of both. A new perspective that authors point out is that it is possible that sea ice loss increases the magnitude of internal variability, such that the cooling trends could more likely even without impacting the forced response. This idea seems to form the bases of one of the main messages of the paper, that both the 'ice-driven' and 'internally-driven' viewpoints can coexist. This is an interesting idea that is plausible, but it is basically a hypothesis with little evidence to support it at this point.

We are aware that describing the debate as centering around a yes-or-no causal relationship is oversimplifying the issue. However, in discussing and presenting the ideas to colleagues as we were developing the review, we realized that this is the impression held by many researchers who are interested in the topic but not actively working on it. We expand on the idea in section 4, and especially in section 4.2, where we discuss the fact that modelling results from different studies are not so much at odds with each other. It is rather often the interpretations of the results as stated in the abstract or conclusions that seem to be at odds. As this statement only appears in the abstract, and as it doesn't specifically relate to modelling studies in this context, we hope the reviewer is fine with us keeping it.

Regarding the hypothesis of sea ice loss altering the magnitude of internal variability, we respond to this in the next block.

The authors do show that because of the higher standard deviation of Eurasian temperatures, the trend was made more likely during 1998-2012. However, the authors cannot attribute these changes to sea ice loss without additional evidence. Much like with seasonal mean temperatures, the standard deviations will vary from decade to decade just due to internal variability, so attributing these changes needs much more careful analysis. There is even a period in the 1960's and 1970's that had similar or even higher standard deviations relative to the recent period (Fig S4b), when sea ice was (presumably) much higher than today.

The authors should either provide additional evidence to support this idea, or make it clear that this is just hypothesis that could be considered. To explore this further, the authors could do some very simple analysis using the data that they already use. For example, do the standard deviations of Eurasian temperatures (or the probability of a large cooling trend) increase along with sea ice loss in the CESM-LE?

The paper that we have written is a synthesis and review of the literature, with a proposed viewpoint to attempt to aid in reconciling the apparent disparities in the literature. The inclusion of the AR1 model study is to help illustrate this viewpoint in a more concrete way. It should also be plainly stated that we have found no strong evidence that sea ice change is resulting in a change in the standard deviation and we leave this to future studies. To clarify these points, we have included a statement in the last paragraph of the discussion.

Although overall, I thought the interpretation of the literature was well done, there was one issue I thought was misrepresented. Throughout the text, it is portrayed that most modelling experiments find some weak Eurasian cooling in response to sea ice loss. I don't think this accurately

represents the published literature. The studies that are cited to support this all use single models and/or small ensembles. Surveying the literature of these single model studies that focus on the cooling is inadequate because it will be susceptible to selection biases. If groups run the modelling experiments and find cooling, they will focus on it, but if none is found the study will not focus on this (if it gets published at all). This could be made worse if small ensembles are used, such that the responses seen may not even be real.

To minimize these issues, large, multimodel ensembles should be used to more accurately reflect the state of the modelling evidence. These should be given much more weight than the single model studies, even if they don't specifically focus on the Eurasian temperature response. So taking a look at large multimodel ensemble studies:

Ogawa et al. 2018 : No cooling from an average of 6 models (Fig 1c) and 5 out of 6 models show no cooling (Fig 2c).

Blackport et al. 2021: No cooling from an average of 4 models (Fig 12d) and all 4 models show no cooling (Fig 13 b)

Liang et al. 2021: Average of 9 models shows no cooling over central Eurasia (Fig 7a).

Smith et al. 2022: Average of 16 models shows no cooling over central Eurasia (Fig 1b)

Hay et al. 2022: Average of 5 coupled models show no cooling over central Eurasia (Fig 6a).

Note that these last three do show some weak cooling only over East Asia, but this is not where the cooling trends are in observations. Also since these three studies do not show individual models results, we do not know how many show cooling, but if it is not seen in the average, it is unlikely that many do (but some probably do).

My interpretation of this is that most models show no cooling, but some do show weak cooling. This is just my interpretation of the literature, and the authors may disagree (maybe I am missing other large multi model studies that do show more evidence of cooling?). If so, this should be backed by evidence that does not only rely on a couple single model studies.

It appears in our attempt to summarize "all" the modelling studies, we oversimplified a few important points. We agree with the reviewer that focusing on multi-model, large-ensemble studies is useful, as it gives a better sense of the robust features. However, we have still kept some discussion of single-model studies for two reasons: 1) they are useful when trying to explain some of the details of how a certain experimental setup leads to certain interpretations, and 2) there are some well-known single-model studies that more "casual" readers are likely familiar with. Based on the reviewer's comments, we have now clarified when we are reporting on a single-model study versus a multi-model study. We have also included more information on experimental setups, and explained why it's necessary to look at large enough samples, both in terms of models and ensemble members. The multi-model studies above are all included, but separated in our discussion according to type b/c and transient/equilibrium experiments.

Small changes have been made throughout section 3.1 to more accurately report on the main messages, along with an acknowledgement that interpreting the various modelling results requires a careful look at the experimental setup and analysis. Section 3.2 includes more detail on specific studies where appropriate, and highlights differences between single-model and multi-model studies. While these modifications add length to the paper, we agree with reviewer #3 that the

details are important. Reviewer #1 commented that the paper feels a bit long already, but we hope the signposting we have added will help readers familiar with the topic to skip certain sections.

More specific comments:

L9: “with a small contribution from sea ice” Isn’t the whole point that we do not know this? It could be no contribution (as indicated by most models), or it could be a small contribution (as indicated by some models), or it could be a much larger contribution (if models substantially underestimate the response).

This has been replaced with “with some potential contribution from sea ice”, thus not specifying the magnitude of the contribution and allowing for the possibility of it being zero, of course more details are discussed later in the paper.

L15-16: This seems a bit misleading. My reading of the paper is that internal variability had a substantial contribution to the 1998-2012 cooling trend and will continue to play a large role in future trends. On the other hand sea ice loss may have altered the likelihood of the trend and may contribute to the likelihood in the future. Treating these two factors on equal footing in the abstract seems a bit misleading.

Your reading of the paper regarding the relative roles of IV and sea ice was exactly what we wanted to convey. The sentence is not intended to convey any information about the comparative magnitudes of the roles played by IV and sea ice, only that they both have a role to play, given the proposed viewpoint. We prefer to keep the sentence simple as is in the abstract and we leave the details to the reading of the paper in full.

L122-125: By ‘disappear’, do you mean that they actually disappear or do they only become non-statistically significant?

Thank you for pointing this out. The choice of words was unclear. The Eurasian Cooling becomes non-significant in the 2001-2015 panel but does continue. By the final panel of 2005-2019, the cooling has disappeared, and is replaced with warming over central Eurasia. As for the Arctic amplification, this also disappears by the 2005-2019 panel, with only small trends in the Greenland Sea and northern Kara Sea remaining. However these trends are not ‘amplified’ above the general warming trend of the NH. To allow the readers to view and consider this for themselves, we have included a new Supplemental Figure which is a copy of Figure 2 but showing the trends in all locations without hiding the non-significant trends.

L247: Related to my comments above, Honda et al. 2009 and Liu 2012 are bad examples to use here. Liu et al. 2012 used only 20 years from a single model for both their high and low sea ice runs, which is too few realizations to be meaningful. Honda et al. 2009 ran their experiments for 50 years from a single model, but “selected” the 28 years where the signal was the largest and note that the responses were much weaker if all 50 years were used.

As per our response above, we oversimplified the description in this section, resulting in a poor mismatch between what was being stated and the references being given. These sentences have been revised and new references have been included.

L248-249: “but also an inconsistency...” What is this based on? From what I have seen all sufficiently large ensembles show a consistent warming signal over recent decades (e.g. Blackport et al 2021, McCusker et al 2016, in addition to some of the studies cited later in the paragraph).

This was not clear. We did not mean to imply that the warming signal is not consistent, but rather that some of the circulation responses (those subject to tug-of-war processes) are not. We have clarified the sentence in response to this comment, and reviewer #2's comment on L250.

L508-513: Is the trend more likely because of the larger standard deviation or is the larger standard deviation because of the trend? The trend itself will cause a larger standard deviation.

For the AR1 model, the larger trend is more likely because of larger standard deviation for the "noise" term. The trend is not set explicitly.