Response to editor

Dear authors,

I have received three review reports from the referees and am pleased to tell that two of them are completely satisfied with your responses to their earlier concerns. Anonymous referee #1 raises a few minor issues; however, the referee is generally satisfied with your responses. Therefore, I have judged that this paper can be published in TC after minor revision. Please prepare the final version of this paper following the comments by anonymous referee #1.

Related to the first comment by anonymous referee #1 on the usage of "Holocene", I think adding more compelling justification of the definition of "Holocene" considered in this study at the end of Sect. 2.2 is ideal. For example, if you can add a relevant reference, which conducted similar numerical simulations to investigate the "Holocene" climate, it will be enough. If you do not find any relevant reference, please consider adding more detailed explanation why the method allows us to obtain the analogous picture of the "Holocene" climate here. I also want to know why the total simulation years of 5241 (1801 + 3440) is enough for obtaining the analogous picture of the "Holocene" climate.

I look forward to receiving the revised manuscript.

Sincerely,
Masashi Niwano

Dear Dr. Niwano,

We thank you for your thoughtful evaluation of the referees' responses.

We agree with you that use of the term "Holocene" can be misleading, and we have revised the manuscript to ensure that our use of the term "Holocene" is used less often and more carefully. We have clarified what we can (and cannot) say with existing information.

The estimates presented in this study provide an estimate of the rarity of the 1940s event in the context of natural variability, but cannot address the uncertainty associated with possible changes to that variability during the Holocene. Unfortunately, sufficiently high-resolution Holocene simulations are not available. We use the ~5,241 years of the LENS simulations to conduct this calculation because they are the best simulations available; the LENS simulations have minimal wind biases in the Amundsen Sea Embayment and sufficiently high spatial and temporal resolutions. We calculate the number of occurrences of similar events in the simulations and scale that number to report the occurrences on a scale of "per 10kyr", which is not directly analogous to occurrences throughout the Holocene (which we have made sure to clarify in our revised manuscript). Please refer to our response to Referee #1 for a more detailed explanation on these points.

Although associated with uncertainty, our results greatly advance our understanding of the importance of the 1940s event as a potential trigger for glacier retreat in West Antarctica – a topic of great interest to much of the glaciology community and, we expect, many readers of *The Cryosphere*.

Thank you for your continued consideration of our manuscript.

Sincerely,

Gemma O'Connor
on behalf of coauthors

Response to Reviewer #1

I would like to acknowledge the authors' efforts in responding to my previous comments. I appreciate the changes that have been made to make the manuscript better. However, upon my second review, some similar concerns remain in the first review to be addressed before publication.

We thank the reviewer for their constructive feedback, which we believe has improved the manuscript.

1. Upon revisiting the revised manuscript, I have to reiterate a concern I previously raised regarding the use of the term "Holocene" in the context of this study. The ensemble of pre-industrial and 20th-century experiments should not be treated as representative of the Holocene condition just because they are the same time length. Using the term "Holocene" without considering the characteristics is wrong.

We agree with the reviewer that the use of the term "Holocene" is misleading and have revised the manuscript to ensure that the term is used more carefully. We removed several instances of the term from the manuscript, which can be seen in the tracked changes manuscript. We clarify that our use of the LENS simulations provides an estimate of the rarity of the 1940s event relative to pre-industrial internal climate variability and note that this is not equivalent to a simulation of the Holocene. We further justify our use of the LENS simulations by noting that sufficiently high-resolution transient Holocene simulations are unavailable. In section 2.2, L232 of the tracked changes manuscript, we added:

"The LENS simulations allow us to quantify the rarity of the 1940s event relative to 10kyr of internal pre-industrial climate variability. We note that our calculations are not equivalent to the significance of the event relative to the Holocene, which experienced differences in variability relating to changes such as insolation and freshwater inputs. Available transient Holocene simulations are insufficient for conducting this calculation as they are only available at much lower temporal resolutions and spatial resolutions (e.g., the widely used Transient Climate Evolution of the past 21 ka (TraCE-21 ka) simulations are available only at 3.75° x 3.75° spatial resolution; He et al., 2013). The rarity calculations presented here are an imperfect analogy to the Holocene but provide a novel estimate of the significance of the 1940s event, based on the best available simulations."

The caveats are stated again in the discussion section at L597:

"Furthermore, the internal components of the LENS simulations are an imperfect analogy to the Holocene, which was subject to differences in insolation, freshwater inputs, and possibly El Niño/Southern Oscillation (ENSO) variability (e.g., Mayewski et al., 2004)."

2. The first four lines of the abstract present the research background, the following two lines articulate the study's objective, the subsequent two lines convey the findings, and the final two lines delve into the discussion. In the findings of your results, I feel it is better to say that atmospheric events in the 1940s possibly occurred about once every 100-200 years than that they are not unprecedented in 10kyr, and it may have triggered the onset of retreat in West Antarctica, together with ocean conditions. I think this way, it will be easier to understand the connection to previous studies. Linking with this comment, I suggest that the author calculate the frequency over 100 years, the 10kyr period divided by 100.

We estimate that ~17 to 250 similar events occur per 10 kyr of internal climate variability (depending on the window length, simulation used, and reconstruction used), which corresponds to .17 to 2.5 occurrences per century, or .34 to 5 occurrences per 200 years. Thus, it would be misleading to say "once every 100-200 years" because of the lower estimates from the natural-prior reconstruction. We have revised the abstract to include the estimate in units of per century and have removed the statement that our results suggest the event is not unprecedented. We have reworded the final sentence about the additional factors to improve readability. The final sentences of the abstract now read:

"Climate model simulations provide evidence that events of similar magnitude and duration may occur tens to hundreds of times per 10 kyr of internal climate variability (~0.2 to 2.5 occurrences per century). Our results suggest that the 1940s westerly event is unlikely to have been exceptional enough to be the sole explanation for the initiation of Amundsen Sea glacier retreat. Additional factors are likely needed to explain the onset of retreat in West Antarctica, such as naturally arising variability in ocean conditions prior to the 1940s or anthropogenically driven trends since the 1940s."

3. Figs 5a, 5d and 6
Can you please calculate the statistical significance from the variance and shade where it is not significant?

We calculate the variance in each pacemaker simulation ensemble mean, and stippled the areas where the variance is statistically significant with 95% confidence (where the anomaly is greater than 2 standard deviations from the mean). None of the regions on the map are significant; the maps look identical to Figures 5 and 6. This makes sense, as the 1940s event -- as constrained by the reconstructions -- is the only 2-sigma event in the 20th century, and our study finds that it is driven by a combination of sources. The pacemaker simulations are constrained by SST variability from only one tropical basin, which could explain why none of these simulations shows that the 1940s event is a statistically significant event. Instead, we revised Figure 6 by stippling areas greater than 1 standard deviation from the mean. We decided to make this revision only to Figure 6, given that 5a and 5d are simply repeated in Figure 6, with different scales.

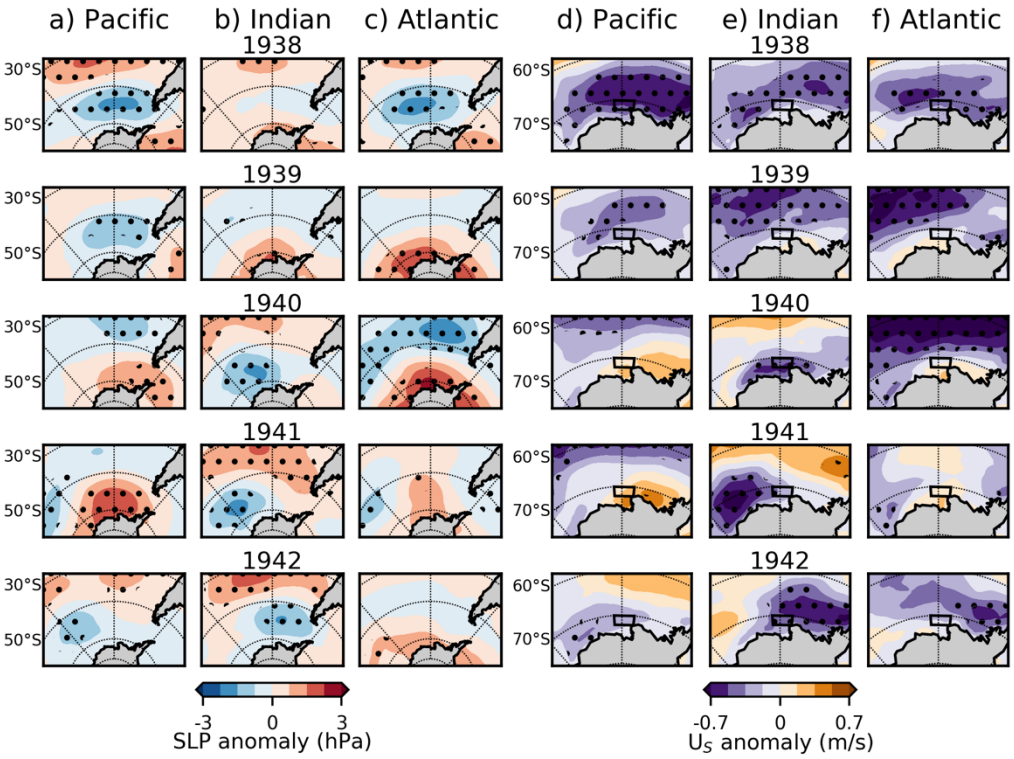

Figure 6. Modeled SLP anomalies from 1938 to 1942 in the ensemble mean of the (a) tropical Pacific, (b) Indian Ocean, and (c) North Atlantic pacemaker simulations. Stippling is shown where anomalies are greater than 1 standard deviation from the mean (no regions contain anomalies greater than 2 standard deviations). (d-f) Same as in a-c, but for U_S anomalies. Anomaly reference period is 1961-1990. We note that the color bar in this figure is smaller than that of Figure 5 due to the smaller variability in the ensemble means of the simulations (Figure 5 shows the ensemble mean and individual members from the Pacific pacemaker simulations).