Response to Reviewer 1

The reviewer comments in black and our response in blue.

Reviewer 1

Review of “The first firn core from Peter 1st Island - capturing climate variability across the Bellingshausen Sea” by Thomas, et al. for Climate of the Past

Thomas et al. provide the initial results from a firn core retrieved from Peter 1st Island. The stable water isotopes and snow accumulation are compared with reanalysis data which they also ground-truth with meteorological data from an automated weather station. The authors provide evidence of some compelling connections with regional/large-scale climate variability to support further drilling campaigns in the area. My primary concerns rest with the low sample number used in the correlation analyses and the influence of substantial melt in this shallow firn core. Although the authors address these limitations to a degree, I would offer a few suggestions or clarifications that will hopefully help bolster their results and interpretation. Nevertheless, the authors demonstrate the potential that a longer core from this site could provide to understand past climate variability in the Bellingshausen Sea region.

Thank you for taking the time to review this manuscript. We agree with the limitations of short datasets and appreciate your suggestions for improvements.

Major revisions

Many of the conclusions from this manuscript are reliant on correlation analysis and standard p-values. However, only 15 years of data are available, and only 13 when two melt outliers are removed. Although I am inclined to trust the relationships discussed in the paper given the strong correlation values, I do wonder about “significance”. The authors attempt to reconcile this issue by creating a pseudo-core data. I worry about the pseudo-core based solely on the ERA5 data which is then correlated with additional variables from ERA5. Could the interpolation schemes used in ERA5 generate spuriously high correlation coefficients? For instance, are there inherent correlations among adjacent or proximal grid cells that may be coming to light here. Circumventing the inability to collect more data at this stage (although I very much hope you are able to obtain a longer core from this site in the future), perhaps applying a bootstrapping method to the correlation analysis would provide somewhat more robust p-values.

We agree that the short duration of these cores is a limitation. The intention was not to over-state the correlations for a short record, but to demonstrate the potential of this site for future deep drilling. We will add additional clarification to the text to acknowledge the limitations, including the potential to generate higher correlation coefficients when comparing ERA5 variables with each other.

Thank you for the suggestion to use bootstrapping. We have applied this to our data, using 2000 samples, to provide more robust confidence intervals. This will be updated in table 1.

The melt in the core is also a concern which relates back to the robustness of the correlation analysis. Given such a short record, the disruption to the isotope, accumulation, and chemistry records by the melt events could greatly alter the results. The authors do address this by removing the two years of greatest melt in some analyses. How much does the melt influence preceding years? For instance, the year 2012 is almost non-existent. Could its inclusion in the correlation analysis skew results? However, once you start eliminating every year potentially altered by melt, you would not have many years left in your analysis. The authors make a good point that melt events are likely more pronounced in recent years and would likely be fewer in frequency in a deeper core. This prediction supports the need for a deeper core but does not address the issues in the interpretation of the results for this manuscript. The data are what the data are; there's no changing that. Is there a way to ensure the results are more robust? In the case of melt affecting the time scale, could making a stronger case for the volcanic marker help here (see minor revisions)? Serendipitously, the second melt event in 2006 is captured by the AWS data. Could that be utilized to provide support for the time scale and subsequent analyses? I apologize that I don't have great suggestions for how to overcome these limitations.

Thank you again for suggesting improvements and acknowledging that we cannot change the data we have. We have tried to incorporate independent tie-points to strengthen the age-scale, however, with such a limited time span we don't have many (e.g., eruptions). We have restructured the section regarding volcanic tie points, so that it follows the results for the age-scale (3.1). Presenting the data together will hopefully strengthen the results.

You correctly point out that the melt is likely influencing subsequent years (e.g., 2012). However, as noted, this record is limited by the lack of data points and removing more would further reduce the statistical significance. The importance and significance of melt is discussed in a recent review [Moser et al., 2023], and we will reference this study in the text to improve our discussion regarding melt.

Minor revisions

Lines 56-60 - citation needed

Ó Cofaigh, C., Larter, R. D., Dowdeswell, J. A., Hillenbrand, C.-D., Pudsey, C. J., Evans, J., and Morris, P. (2005), Flow of the West Antarctic Ice Sheet on the continental margin of the Bellingshausen Sea at the Last Glacial Maximum, J. Geophys. Res., 110, B11103, doi:10.1029/2005JB003619.

Line 68, should be 4 instead of 3 in list

Corrected

Line 85-86, you mention GPR profiles indicated stratified layers in the upper 14m, and the GPR was set to 43 m. How did the layers appear below 14 m? Including a mention from the Thomas et al. 2021 paper here might help with demonstrating feasibility of a longer core record. I would suggest moving the bullet point from the conclusions (line 526) to this location, especially since it is not a conclusion of this manuscript's analysis.

Updated to demonstrate "Near continuous stratified layers were observed in the GPR profiles for the upper 43 m at this site (the maximum time window for the GPR) and bedrock was not detected at this depth. However, the full ice thickness has not been determined."

Line 101, it may be helpful to make this point in the introduction. At present, the ERA5 comparison with the AWS data appears like a separate side-project in the introduction. Making the point that reanalysis products have struggled in Antarctica and that's the motivation for the AWS comparison would justify it a bit better in my opinion.

We will add this to the introduction.

Line 131, were other thinning functions tested in light of not knowing the full depth?

Unfortunately, other thinning functions would require additional in-situ observations (e.g., measured strain rate). This is a limitation of the data that we have, and we have tried to make it clear that this is likely not appropriate for the site. Future deep drilling would require a thorough geophysical survey to determine the most appropriate thinning function.

Is there a reason for not including MSA in figure 2?

The full suite of major ions was measured for this core; however, we decided not to present all the data to avoid overcrowding the in the figure. In addition, the MSA is being utilised by a PhD student in a related project and we decided not to present the data here first. However, we can present all the data in a supplementary figure if this is considered necessary.

Line 165 - greater than 50% of hourly data is a low threshold. How might changing this threshold influence the reported warm/cold bias in ERA5? With such a high correlation between ERA5 and the AWS, I imagine the variability would not be affected too much by altering the threshold, but it could make a difference in the magnitude of temperatures. Might be worthwhile since there is considerable discussion regarding the lapse rate in reconciling these differences.

Good suggestion, we will increase the threshold to explore how the absolute values change.

Line 187 - just to clarify, the lapse rate estimate applied to the AWS greatly reduces the ERA5 warm bias, but the lapse rate estimate applied to the drill site underestimates temperature at the drill site? Is this the reported cold bias mentioned in line 380?

Apologies, there is a typo on line 380. The ERA5 data is warm biased. The temperature difference between the AWS and ERA5 data suggests that ERA slightly over-estimates the temperature.

Line 321 - I am not sure the analysis around SAM and SOI are necessary for this manuscript given how heavily caveated it is due to the short time span. I see the need to demonstrate a longer core's potential in exploring these dynamics, so I can understand its inclusion at this stage. I'm just not sure it adds much at this stage. I would suggest either paring it down by noting much of this analysis takes place in a positive SAM phase and merely mentioning the non-stationarity suggested by Figures 5g-i. Perhaps focus on the Amundsen Sea Low variability at this stage, which could be scaled up to explore SAM, SOI dynamics with a deeper core.

I appreciate that the short timespan makes it hard to draw conclusions about SAM and SOI. Considering this, we will remove the paragraph directly relating to the peter 1st data correlations with SAM and SOI (and remove from r values from table 1).

Figure 5 c and f do not match very well in the details, although in the greater scope the match is sufficient. Does the lack of strong correlation between accumulation and zonal winds extend from post-depositional effects at the drill site or small-scale wind patterns affected topography, coastal proximity, etc.?

Certainly, the post depositional effects will influence the firn core snow accumulation. We will include some text to elaborate on the potential differences as suggested. However, it is a difficult balance this with not over-interpreting such a short record. And as mentioned previously, the pseudo cores may artificially be increasing the correlation coefficients because we are comparing the ERA5 with ERA5. Including both caveats will hopefully explain the potential differences whilst highlighting that we cannot over-interpret the data. The hope is always that a longer record would overcome both these issues and provide more robust correlations.

Line 347: Regarding the sulfate peak at 4.6m, described as the volcanic eruption. Why are there accompanying peaks in Br and Na in Figure 2? I had assumed this was an issue of percolation due to melt. Is the sulfate in Figure 2 just biogenic. If so, how did you tease it out? Would air trajectories help confirm whether volcanic material was transported to your drill site?

This is a good question, and one that is a little difficult to address with the data we currently have. The inclusion of back-trajectories was considered, however, we felt this might be beyond the scope of this study. Our intention is to present the data we have (and be open about its limitations) to demonstrate that this site has potential for future drilling. Whilst leaving room for further work, including back-trajectory analysis with additional proxies. For example, we can look for evidence of tephra in the firn core, however, this would require significantly more analysis and might be best applied to a longer record.

Line 450-454 - the potential to reconstruct blocking variability and the connection with atmospheric rivers are exciting! I think this could be fleshed out a bit more, but that's merely my personal preference.

We agree that this is an exciting aspect of the site location. As mentioned previously, we don't want to overstate this and decided not to expand on the analysis here. While 2006 and 2013 do support the influence of ARs, the 2010 year behaves differently. We feel that more in-depth analysis is required to expand on this connection.