SUMMARY: GENERAL COMMENTS

In light of recent extremes in Antarctic sea ice, and debate over their climatological significance and possible causes, this is a very timely, important, and thorough paper on interpretations of longer-term sea ice reconstructions. My take homes are, first, that sea ice reconstructions differ a lot pre-satellite era, and this is only partially explained by season or region being represented. Second, there is evidence that different implicit underlying atmospheric reconstructions drive these changes, in particular changes in the sea ice-pressure relationship over time and indeed in pressure trends themselves over time. There are also lots of other analysis details and interesting corollaries that will be of interest to the community. It is certainly relevant for publication, and I have no concerns with the analysis or methodology.

However, I do have major comments on interpretation and presentation in a broad sense; the text becomes overly detailed and at times the scientific reasoning that leads to conclusions is hard to follow and therefore unconvincing. Many methodological details are unnecessarily repeated in the results section, which makes it hard to follow the results; many data details are missing. The conclusion section is vague and could benefit from strengthening with some specific take-home messages. Figures 3, 4 and 12 also require reworking by the addition of correlation tables and other edits. I've given fairly detailed suggestions for all these points under 'specific comments'.

I therefore recommend this manuscript for publication after minor revisions; the revisions I suggest are large in number but would not, in my opinion, require a second round of review.

SPECIFIC COMMENTS

L26: 'has with its underlying' unsatisfactory: the proxy-based records don't explicitly have an underlying atmospheric circulation?

L29: paleo-based-> proxy-based?

L31 -> "sensitivity experiments …" This sentence is very hard to understand without having read the paper multiple times! I suggest simplifying what you try to convey: something like 'results from reconstructions based only on coral or ice core records, rather than both, imply contrasting roles of these records- and therefore of tropical versus purely local atmospheric variability- in driving different types of reconstructions' or similar.

L53: For NSIDC Sea Ice Index SIE, 2022 also a record low annual-mean, so include here.

Last line of the abstract emphasises tropical versus local drivers as a key take home but this isn't mentioned in conclusions and for me wasn't a main focus of the paper. Either change this line of the abstract or add a short discussion of this to the conclusions section!

Introduction: Some of the discussion in introduction seems to be about general Antarctic climate (e.g. line 66 'across the continent') not sea sea ice. Clarify/refine text? E.g. line 69 'Antarctica' ->'Antarctic sea ice and atmospheric variables'; similarly line 71 'climate' -> '? Temperature? Sea ice?'

L62 'near 1978-1979' -> 'in late 1978'

L65-67: Mention predicting future change as a motivation for understanding the past

L77: Add a comment on time coverage of these other instrumental observations outside Antarctica, e.g. 'dating back to as far as …'

L100 Clarify the scope of the study: mostly concerned with West Antarctic sea ice, annual focus but seasonal is examined to help understand annual.

L116-118 it took me a lot of mental effort to reconcile this with the table. Try 'We also investigate two proxy-based reconstructions of spatially limited sea ice extent, and three spatially complete proxy-based climate model reconstructions, two of SLP and one of sea ice.'

Section 2.1). Ensure each dataset (sea ice obs, sea ice recon, pressure recon) has its own paragraph-the details become blurred. A couple of lines are needed introducing the Fogt et al 2022 sea ice reconstruction. Is the best fit reconstruction used here? How is the annual mean created (move here from results).

L154: 'Equatorward extent of the' -> 'total' (equatorward extent would imply e.g. polynya area was counted as extent?)

L211: Information about observations (Met obs for Fig 12, Met obs for Fig 13, and Weddell fast ice) should be included here and any repetition removed from results. Detail should be added on the time coverage of station data in Fig 12, and how these stations were selected.

L175 and Fig 1: The black lines on Fig 1 (specifically 70-100W) are not the same as sectors in Parkinson (2019). I don't think Parkinson sectors are needed- when Raphael and Hobbs sectors are introduced it should just be mentioned that they differ from traditional sectors (with ref to their paper which describes it fully). Fig 1 should just indicate 100-70W for the Abram reconstruction. Also, decrease size of green dot slightly for more precise rendering of location.

L182-210: Explain clearly the 'fixed prior' versus forced prior that's discussed later in the section and in Results, and state which model uses what. Add this info, and whether the data is directly calibrated against obs, to Table 1. L191 'temporal variability comes only from proxies' tropical pacemaker/external forcing prior surely forces temporal variability?

L217-218 'standardised (to place on same scale) annual-mean sea ice extent' -> 'standardized annual-mean sea ice extent anomalies'. Recap why magnitudes differ (because of region definitions) and variability not magnitude matters.

L224-228: 'Highest' implies multiple datasets being compared but at any point in this sentence there are only two (Dalaiden versus EITHER Thomas and Abram OR Abram). A different emphasis is therefore needed e.g. '[T&A] correlations with observations exceed those of Dalaiden 2021 in Ross-Amundsen, but Dalaiden exceeds Abram 2010 in both Bell-Amundsen and in Weddell'

L226 etc: 'data assimilation based' etc. gets repetitive explaining what reconstructions are every time (e.g. also line 245-246, 302,340,…) and the resultant wordiness detracts from scientific reasoning. I suggest introducing a shorthand e.g. Paleo=>'PALEO reconstruction', station based=>'STAT reconstruction', estimate from data-based reconstruction=> 'ASSIM reconstruction' and abbreviation for dataset names. E.g. L226 would become something like 'while the Dal_21 ASSIM reconstruction …'

L324 'Sectors have been adjusted' i.e. the SIE has been recalculated from Dalaiden SIC using new boundaries? Reword for clarity.

Fig 3, 4: These are very hard to decipher, not least because red is different dataset in different panels. Specify this in the caption. Please put the correlations in cross-correlation tables as a new right hand column to this figure. At the moment working out which correlation is which, reconciling with the text, and in particular comparing correlation values, is extremely challenging; a table would

greatly improve this. (Also, the correlation text at bottom of Fig 3 panels uses different layout to those of Fig 4 panels- very confusing.)

L242 Context needed: where correlations with obs 1979-present exceed cross-reconstruction correlations 1905-2020, how do cross-correlations compare over the common 1979-2020 period?

L242-259 It is hard to follow this discussion of comparison of correlations, partly due to the paragraph structure and partly due to the lack of correlation tables. Indeed I think it is misleading as the text seems to state cross-correlations are all lower than correlations with observations, then later states this isn't true for Abram 2010. To resolve, I'd use a table and address correlations in turn: first Dalaiden with pure proxy-based then Fogt with all other datasets. Also, be sure to sell the take home here; basically none of the cross-correlations are significant pre-1979 so the reconstructions really do tell us very different things!

L260 'Sudden anomalies': very intriguing in light of current narrative around records. Suggest expanding this, or commenting on implications (e.g. extremes from paleo record should be interpreted with great care- see also end of my previous comment?!)

L273: To aid with seasonal vs annual, compare Fogt SON to Abram et al (2010) ASO?

L276 'biased by accumulation at the ice core site': cite?

L286: r=0.56 isn't so different from r=0.525, similarly 0.345 vs 0.320! I'd drop line 285-289 to avoid losing the woods for the trees, and slightly rephrase L290 accordingly. Similarly line 305-307 doesn't seem to describe something that's substantially different from the annual? Is line 310-311 (better agreement in seasonal) therefore true?

L304 I have a general dislike of calling 'SON' spring for sea ice- it's the maximum! Meterologically correct, cryospherically misleading. Rephrase.

Figure 5: Indicate the relevant sector in each panel, to draw the readers eye to where one might expect the correlations to be high and positive.

L382 remove bracketed phrase: repetition

Fig 6: The satellite trends look like they start in 1969?

L398: this paragraph is crucial but a bit weak. At L399 (and 414) expansion/recap is needed on why Fig 5 implies role of atmospheric circulation? For me it's more Fig 6 (contrasting trends which we know in obs period are linked to wind patterns) that imply the circulation role? At line 406, Dalaiden pressure is linked to Fogt sea ice, but this seems a bit of a leap of reasoning?

L409: Very helpful paragraph!

L419: Split this section for readability. 3.2 'Connection to the atmospheric circulation changes' and add at L529 3.3) 'Differences in atmospheric reconstructions'

L424 'for consistency' with what?

L424-430 over-repetition of methods. Cut down and reference Methods section

L430-433. "Since the Antarctic station pressure reconstructions were generated using a similar statistical technique as the Fogt et al. (2022a) sea ice extent reconstructions, this allows for an evaluation of other estimates that are expected to provide **similar temporal variability** as the Fogt et al. (2022a) sea ice extent reconstructions." Seems at odds with the conclusions that the

relationship between Fogt sea ice and Fogt pressure changes over time and with statements at L466-469. I think this is just me being confused, so please elaborate for clarity of reasoning!

L434: "Figure 7 displays the correlations for the Weddell sector sea ice extent from Dalaiden et al. (2021) with the gridded pressure datasets in the top rows, and the Weddell sector sea ice extent reconstructions from Fogt et al. (2022a) with the same gridded datasets in the bottom rows" -> "Figure 7 displays the correlations for the Weddell sector sea ice extent from (top rows) Dalaiden et al, 2021) and (bottom rows) Fogt et al. (2022a) with three gridded pressure datasets"

L439 'panels' -> 'rows'?

L456-L461: (overly lengthy) Change to : "However, **for** the Fogt and Connolly (2021) pressure dataset, the relationship **with** Weddell sea ice extent, regardless of the sea ice extent estimate (Dalaiden et al. (2021) or Fogt et al.(2022a)) shows a change in **sign** poleward of 60°S in the 1905-1978 period".

L471: 'this pattern' what pattern?

L503: clarify Orcadas is met data not sea ice.

L511- 'change sign' but neither sea ice time series has significant correlation with Orcadas SLP so I think not much can be read into this.

L512: 'spatial pressure change' -> 'change in the Weddell SIE-gridded SLP correlation change'

L516 and other locations: add explanation of why analysis is split at 1945.

L517 'weakly correlated' –> 'have no statistically significant correlations'

L538-552: (presentation) Overly lengthy, repeats content from methods, and implies slightly that new sensitivity runs were performed for current paper. Suggested succinct reword: 'Further, we analyse the single-proxy-based reconstructions of O'Connor (2023) to pinpoint the sensitivity of these correlations to using only tropical (coral) or Antarctic (ice core) paleo constraints'. Also this is mostly a repetition of methods section.

L552 New paragraph. Re-explain why 1956 break point?

L552-L602: This section is hard to follow and needs more careful wording/separation of 1) discussion of what's driving the patterns in the reconstructions (comparing panel patterns), 2) how well they are correlated with Fogt and Connolly and 3) what we can learn (i.e. what's driving variability in Fogt and Connolly? Or is Fogt and Connolly considered as more reliable?) Is one possible interpretation that Fogt and Connolly is biased in East Antarctica due to over-dominance of tropical here in this reconstruction? Specifically the word 'agreement' is ambiguous e.g. at line 569 'agreement improves' –agreement between assimilation reconstructions, or with Fogt and Connolly?

L556: I think this implies O'Connor 2021 has forced prior but that L456 implies it has fixed?

L563-564: rephrase to 'not surprising that the coral-only dataset of O'Connor et al is better correlated with F&C'- but this is only true pre-1957?

L566- 'is opposite that of... especially in East Antarctica' I'd rephrase as 'except in the Amundsen-Bellingshausen sector' and then include the note about teleconnections being less important for East Antarctica.

L581-L582: belongs in methods

L592 'agreement' -> 'correlation'

L619: 'changes' (implicitly in time) or 'differences' (between datasets)?

Fig 12: Again consider correlation tables. Which value refers to which dataset is very hard to interpret. The Fogt dark green and obs black are almost indistinguishable. What's the LHS? Text (L608) 'correlations with obs improve after 1945' implies it is correlation with obs but text in panels and the fact most obs aren't available until 1945 implies not. Also L619 "change in correlation in Fig 12" what's meant here?

L641 'statistically significant differences in the underlying atmospheric circulation' are there?

L660: Give Vostok longitude.

L678: The word 'incorrectly' is important; it implies Fogt and Connolly early C20 trends are in fact incorrect, which sheds information on which of all the reconsturctions is more reliable. Discuss.

L693: 'reverse' when? Previous sentence is only discussing pre-1956 but trend reversals are later?

Discussion: Byrd data limitations should be mentioned briefly in results section.

L778: "not able to determine the validity of changes in the atmospheric circulation in the early to mid 20th century." … and therefore not able to deduce which sea ice reconstructions are the best representation of reality?

L782-> (on strengths and weaknesses and lessons for users): I wonder if this could be recap some key specific conclusions from the paper? Or if the authors feel uncomfortable doing this, flag the key questions that remain to be answered to understand which datasets are more reliable.

Technical Comments
L18 minima -> lows
L20 add 'since 1900'
L27 'several'-> '5'
L113 remove 'various'. Put this sentence to end of para.
L125 link failed but https://rda.ucar.edu/datasets/ds570-0/ worked
L151 'combines of' -> 'combines'
Table 1 column 2 heading 'type' sufficient. In column 4 also ref Figs 2b) and c) where relevant
Table 1 last row/column 'interpolated' typo
L174 Remove 'a' before 'three'
L219 with->of
L224 ', and the Thomas and Abram…' -> '. The Thomas and Abram…'
L242: Remove 'however'
L246 '; part'-> '. Part'
L264 (end): 'created in'->'created with'
L279 'Compared to' -> 'in contrast with'
L428 'A blend interpolated' -> 'a blend of interpolated'
L553: 'between' -> 'for' (authors are not cross-correlating these datasets with each other)
Fg 12 Be explicit that point A = panel a)
L689 'with' -> 'between'
L691 repeated '1905-1956'
L791 'Ceter'->'Center'