We thank the reviewers for their helpful feedback on our manuscript. We address both of the reviewers' concerns and suggestions below. We have included clarifications and changes to the manuscript.

Review 1

Major comments:

'These results are expected, given that Weijer et al. 2017 reached similar conclusions using a fully-coupled GCM with an atmospheric model component resolution of 25km.' and 'In those above fully-coupled studies investigating open ocean polynyas, Weijer et al. 2017 exclusively looked at the atmospheric response to WSPs, and their atmospheric model component was configured at 25km horizontal resolution.'

We agree with the reviewer that (1) our results on the response of the atmosphere to the polynya are in line with expectations although this is due in part to the reviewer's interpretation that (2) Weijer et al.'s work is an investigation of the direct atmospheric response to the polynya, which we do not agree with. These 2 points are developed below.

(1) Indeed, our findings are somewhat expected. However, we argue that in itself this does not make them unworthy of publication. These expectations will remain speculations until the proper experiments are carried out to confirm expectations. This is the main goal of our study. Note that the direct response of the atmosphere to the polynya has not been explored since the studies of Dare and Atkinson (1999, 2000) and Timmermann et al. (1999). However, Dare and Atkinson did not use a full 3-dimensional Global Circulation Model, but a 2-dimensional (x,z) model spanning the boundary layer in height (1000 m) and 100 km across the polynya. Timmermann et al. (1999) assumed a warming in response to the polynya and inferred the atmospheric circulation response from thermal wind balance assuming a level of no-motion at 6000 m height. Our study using 3 GCMs at 2 resolutions each is a much extended investigation compared with previous studies, providing a 3-D, global, fully dynamical estimation of the atmospheric response to the polynya, as well as a test of the robustness of the results to model and resolution dependencies. To our knowledge, this is the most extensive AGCM study conducted for the response to the Weddell Sea Polynya (WSP).

(2) The synchronous composite analysis carried out by Weijer et al. (2017) in a coupled model does not allow to separate cause and effect. Without further information, the atmospheric signal they extracted cannot be interpreted as the direct response of the atmosphere to the polynya, but as a mix of this direct response and feedbacks. In fact, Diao et al. (2022), who used the same model as Weijer et al., argued that the polynya is the result of a coupled ocean-atmosphere mechanism. For similar reasons, the atmospheric signal extracted by Moore et al. (2002) from the NCEP-NCAR reanalysis which integrates observations of the real atmosphere and hence potential secondary feedbacks, cannot be either interpreted as the direct response of the atmosphere to the polynya. An additional issue (which does not directly pertain to the reviewer's comment but feeds into the novelty of our work) is that both Weijer et al. and Moore et al. estimate the atmospheric signal from a very small sample of polynya realisations (3

and 1, respectively), so there are large uncertainties due to the imprint of internal variability.

To separate the direct atmospheric response from the potential feedbacks in coupled models, one needs to use additional uncoupled (AGCM) simulations where feedbacks are disallowed. A recent study by the lead author highlights the need for AGCM experiments, with coupled models, in order to isolate the direct response from coupled feedbacks (Ayres et al., 2022). The use of AGCM experiments with prescribed surface boundary conditions is a standard approach which has been extensively used in the mid-latitudes and polar regions to investigate coupled interactions. As argued in point (1), our study addresses this point with state-of-the art models providing a very significant improvement compared to the previous literature.

A posteriori, in light of our results, it appears that the atmospheric signal extracted by Weijer et al. is to the first order of the direct response to the polynya, and that if there are coupled feedbacks in their model, they are small. Our results suggest that feedbacks invoked by Diao et al., are weak.

The reviewer’s comment made us realise that we have not made a good job of explaining the scope of the paper. We specifically intend to *not* use coupled models.

We have made our goals clearer in the manuscript, stating our differences to the Wiejer et al paper, and how this work complements theirs.

‘The study has mainly suffered from the experimental setup since the varying resolutions apply only to the atmosphere.’

With the clarifications above about the scope of our study, we now hope it is clear that the varying resolution being applied to only the atmosphere is not a limitation but a goal of our study.

Again, we emphasise that the purpose of our study is to infer the direct atmospheric response to the WSP through atmosphere-only models, avoiding the ambiguities around causality and feedbacks that arise in coupled models. Our study complements and helps the interpretation of the coupled model studies that assess the full atmospheric signal associated with the polynya.

By applying a boundary forcing of sea surface temperature (SST) and sea ice to the model, we infer the direct response with no secondary feedback with the ocean, a method used to infer the atmospheric response to SST and sea ice change for many decades, dating back to 1990 with (e.g., Royer et al.,1990), to the present day (e.g., Zheng et al., 2023).

We use this method because we do not have many Weddell Sea polynyas in our satellite record history, and many polynyas that occur in coupled models are not realistic (i.e., too big or occur too frequently, with poor ocean representation). We apologise if we have not made this clear in the text, and have now emphasised this further in the manuscript.

‘However, the ocean and sea ice forcing data remain the same in resolution (ERA5 data standard 31km horizontal resolution). It should have been evident to the authors that given the numerous high-resolution studies on the mechanisms of WSPs, to look at just the atmospheric response of the WSPs, one would need to use atmospheric models much higher than 25 km horizontal resolution despite the size of the WSP (the great WSPs of 1974-1976 remained ice-free throughout the austral winter with an ice-free region of ~ 250,000 sq. km).’

As discussed in the first paragraph of the methods and data section, the resolution of the boundary conditions makes little difference to the response. There is limited accurate satellite data for the 1974 WSP, hence our use of ERA5, which is one of the most accurate reanalysis products available. When compared to other reanalysis, ERA5 demonstrates the best performance in representing many processes over Antarctica (e.g., Gossart et al. 2019). We do mention limitations of ERA5 in the text, which we apologise is not clear, and have made changes accordingly.

During our initial sensitivity experiments we tested our model runs with the 2017 Maud Rise polynya (MRP) satellite data (Merchant et al. 2019), and compared it to the 2017 ERA5 forcing. The satellite and ERA5 reanalysis showed little difference for the 2017 Polynya with up to 0.1 K SST difference under sea ice, and ~0.5 K difference in the polynya region (Fig. R1). SIC within the polynya shows no difference between the two datasets. The higher ERA5 SST produces a slightly greater heat flux but we found no meaningful difference to our results beyond this.

Satellite data was not available for 1974, therefore a comparison was not possible for this event. We chose the 1974 WSP to maximise the signal-to-noise ratio due to the larger polynya that year, and the literature discussion (e.g., Cheon and Gordon 2019), that due to the smaller size and position of the 2017 MRP (80,000 km^2), it is not considered as a WSP.

In ERA5, the 1974 data is of slightly worse quality than data after 1979 as a consequence of the limited satellite observations at the time. Crucially, for this study, high accuracy of SIC and SST is not critical- that is we do not expect that change of these boundary conditions, that would impact a few in grid points bordering the polynya, will significantly affect the atmospheric response.

Additionally, for our model resolutions, an objective of this study is to explore the impact of the polynya at varying resolution AGCMs, which to our knowledge has not been done before. We could not find in the suggested coupled modelling papers by the reviewers a clear statement that the atmospheric response to the surface boundary conditions should be carried out exclusively “at much higher resolution than 25 km”. While we agree that higher resolution would be better (although this is all subject to debate), our experiments are at the forefront of what is routinely feasible. Our experiments with HadGEM3 and OpenIFS at high resolution use atmospheric resolution of about 25 km.
On a more pragmatic level, we do not have access to a global well-tuned atmospheric model with much higher resolution than 25 km that we could afford to run on large ensemble.

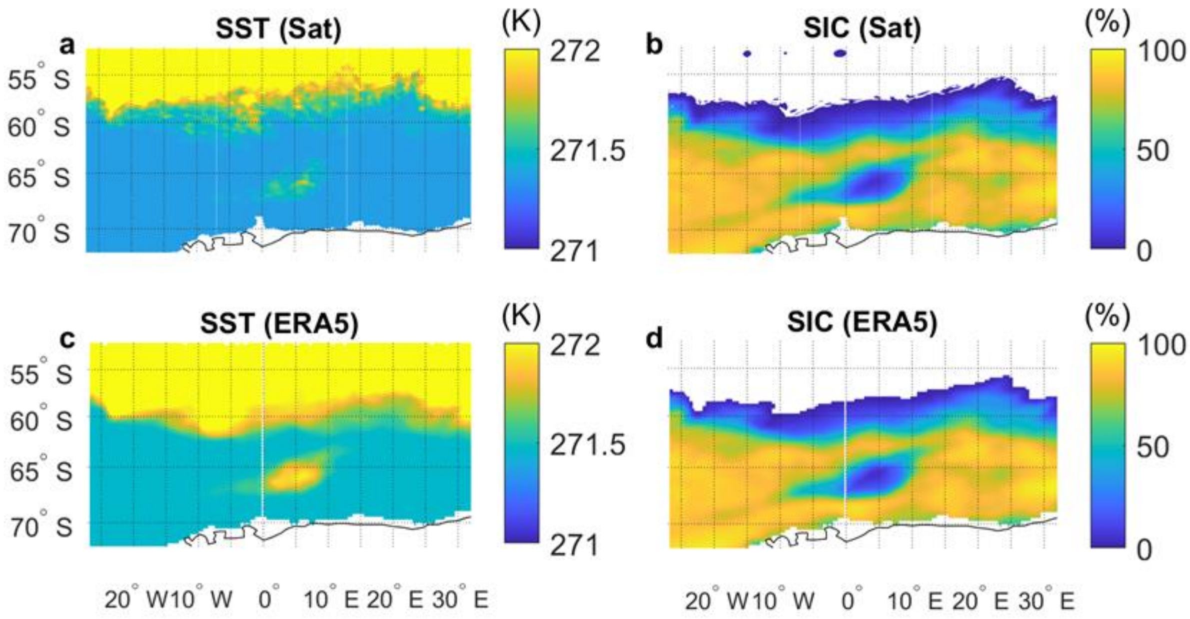

Figure R1: An experiment conducted at the start of the study, to determine the differences between satellite and ERA5 sea ice and SSTs for the 2017 Maude Rise polynya.

'It would have been valuable for the authors also to consider high-resolution fully coupled modelling studies that have simulated and investigated the cause/effects of realistic open ocean polynyas in the Southern Ocean (Dufour et al. 2017; Kaufman et al. 2020; Gutjahr et al. 2018; Stössel et al. 2015; Chang et al. 2020; Kurtakoti et al. 2021; Weijer et al.). It might help the authors to see how the representation of the WSP changes with resolution in coupled climate simulations and how the atmosphere responds to the WSPs.'

Most of the studies listed by the reviewer focus on the preconditioning and formation mechanism of the polynya, which is not the topic of our research. We are concerned by the impact of the polynya on the atmosphere once it is formed. Some of the suggested papers discuss the atmospheric signal associated with the polynya. However, as discussed in the point (1) above, there are ambiguities in interpreting the simultaneous atmospheric signal in a coupled model as the atmospheric response to the polynya. Note that this is not impossible but one needs to use sophisticated analysis techniques such as the lagged maximum covariance analysis to disentangle cause and effects (see for example, Czja and Frankignoul 2002, for an attempt to extract the atmospheric response to SST in the mid-latitudes).

Additionally, there are no models that we know of, in which WSPs are represented systematically across varying models and resolutions, where most of the suggested studies come from the same model and research group. Therefore, a direct comparison to the suggested literature is not feasible and is irrelevant to our study. Additionally, where relevant, we have already discussed these models in both the introduction, and the discussion where we discuss the potential feedbacks that these coupled models induce. We may not have emphasised enough that we aim to compliment the coupled models with uncoupled models. We have included some additional information on these studies in our manuscript.

To improve the manuscript, the authors may want to investigate how the clouds characteristics, radiative fluxes change and/or modify the cloud radiative effects over the WSP in these simulations.

We thank the reviewer for the suggestion. Unfortunately, we did not save any outputs directly relating to cloud characteristics. However, we do have net top of atmosphere (TOA) shortwave and longwave fluxes. We show below the response of these two fields to the WSP in the OIFS-HR model. The other models show similar patterns and magnitudes.

The net TOA shortwave response to the WSP is shown in Fig. R2 (left). There is an increase in net shortwave absorption local to the polyna, as expected when reducing the sea ice in the region. The net TOA longwave radiation response (Fig. R2, right) shows a small decrease (~ -8 Wm^{-2}), i.e. a small increase in outgoing long wave radiation. This could reflect a small increase in the cloud-top height. However the longwave effect is dwarfed by the shortwave effect, i.e. the sea ice albedo effect.
Note that we show in Fig. R2 the September-October average of the fluxes anomalies. The shortwave effect is in fact highly time dependent, varying from near zero in August (no incoming sunlight) to a maximum in November (start of the summer).

We ask the reviewer if they would suggest us including this analysis in the final manuscript?

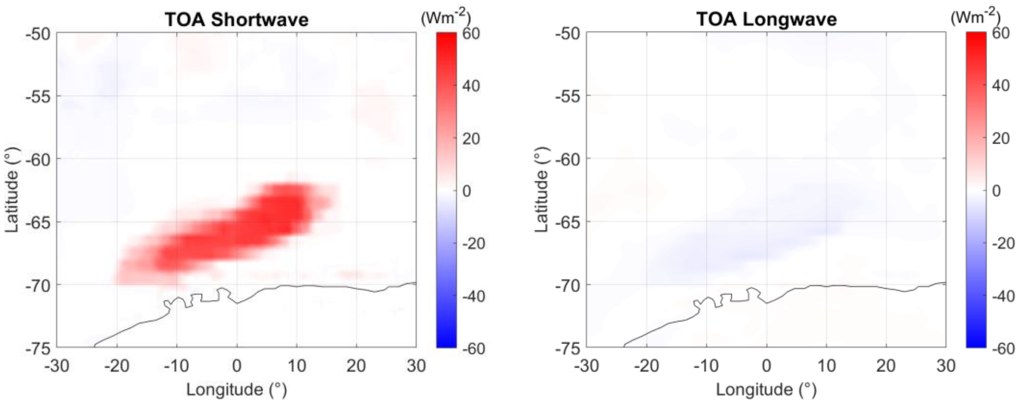

Figure R2: (left) September - October mean top of atmosphere shortwave flux response to the WSP for OIFS-HR (polynya - non-polynya). (right) as left for longwave. Positive in the down direction.

Minor Comments:
Lines18-20:Please explain this further.
We thank the reviewer for highlighting that this is unclear, we have adjusted the manuscript accordingly.

Line 34: “perhaps having only occurred once per century”.

The exact frequency of WSP in the past is not known. Studies have shown a strong link between the southern hemisphere westerlies, SAM index and WSP formation (Cheon et al. 2017; Gordon et al. 2007; Gordon 2014). Gordon et al., 2007 mentions “Gordon (1982) reports that two hydrographic stations obtained by the Argentine ship San Martin in 1961 reveal the absence of the warm deep water, similar to conditions encountered in the 1977 Islas Orcadas stations. The SAM index indicates a prolonged negative SAM in the decade prior to the possible polynya in the winter of 1960. Furthermore, except for the 5-yr period centred on

1910, a negative or neutral SAM index persisted from the 1890s into the first three decades of the twentieth century. Might the Weddell Polynya have been common then?

We thank the reviewer for these interesting suggestions but we think that such speculations on the frequency of the polynya would take the reader on a tangent and weaken our core message. The aim of our study is not the formation mechanism and frequency of polynya, but to extract the atmospheric response to the polynya to set up a solid basis on which to interpret coupled simulations and observations. Indeed we hope that our work will help with pushing further the suggestions made above.

Returning to the frequency of the polynya, we prefer to stir away from this debate and we will replace this statement by a more generic statement relying on published literature.

Review 2:

The methodology is not valid and I invite the authors to rethink how they can improve their strategy to investigate this question.

We are unsure as to why the reviewer thinks that our method is not valid, or which aspect specifically. The statement that the method is invalid is not followed by any explanation. The method used in this paper is rather standard and has been used to explore the direct response to sea ice change with the use of sea ice concentration and sea surface temperature in the polar region from back to 1990 (e.g., Royer et al.,1990) to the present day (e.g., Zheng et al., 2023).
Possibly, the reviewer misunderstood the aims of the study and hence the approach. We hope that the clarifications below help dissipate this potential misunderstanding.

- The authors could have used satellite data for sea ice instead of ERA5 in order to get an accurate coverage of sea ice during the polynya event.

The reviewer suggests using satellite data for our experiment, however, there is no gridded monthly satellite data for the 1974 polynya, which is required for our boundary conditions. If they mean for us to use satellite data for the 2017 MRP, then that is not the purpose of this study as the 2017 polynya was not classified as a WSP and is smaller than the 1974 polynya (hence issues with the signal to noise ratio).
Therefore, for the aim of this study, we could not have used satellite data, and ERA5 is the most accurate coverage we have. Note that the ERA5 sea ice concentration field is a merged product of available data (satellite, in-situ); whatever satellite data are available for 1974 would be included in the ERA5 sea ice.

- The authors interpret the anomalies in heat, temperature and precipitation as being due to the polynya. However, it has been shown that during the polynya events (2017 and 70s) there is an excess of heat and precipitation coming from the atmospheric rivers **toward** to ocean. How the authors can be sure that the values they obtained for the different parameters are solely due to ocean-to-air transfer of heat and not to the existing atmospheric conditions (atmospheric rivers and cyclones) which lasted for several days? This is critical and needs to be addressed by the authors perhaps by conducting sensitivity studies using the models and comparing one set of simulations **with** sea ice opening and one **without** sea ice opening but both with the same atmospheric conditions i.e. those occurring during the polynya events.

Our study uses an ensemble method, where the state of the atmosphere is different each year, and is averaged with statistical significance applied. We run simulations with and without the polynya in an atmospheric GCM, by design, our results (polynya minus non-polynya simulations) only show the direct atmospheric response to the polynya. This is an extremely common method for modelling studies.

We apologise that we must not have emphasised the method of ensemble averaging and the use of monthly means, in addition to significance testing to reduce the impacts of short-lived weather events. Additionally, the use of AGCMs would mean that there are no impacts of these short-lived (or any) events on the ocean or polynya, that is in fact the reasoning for our method, to eliminate secondary ocean feedbacks that one would get in a coupled model. We will make these points clearer in the manuscript.

The reviewer's suggestion of an additional computations with prescribed sea ice and atmospheric state effectively reduces to a bulk formula calculation of the fluxes (everything else would be prescribed). Effectively, the results of such a computation would be close to that of the ERA5 fluxes where the sea ice and SST are prescribed, and the atmospheric state is constrained (but not prescribed) by observations. In fact, to carry out such a computation in practice, one would need a gridded atmospheric product for surface temperature, humidity, winds etc, and thus would turn to a reanalysis product such as ERA5. In this case, the suggested computation would return exactly the air-sea/air-ice fluxes of ERA5 presented in our Fig. 8.

More importantly, we believe that such a computation would not answer the reviewer's question about causality. In addition, we are not concerned here by the effect of the atmosphere on the polynya but on the effect of the polynya and the atmosphere.

- During the 2017 event there are in-situ measurement from the SOCCOM network that can be used at least to check how the models are performing.

This suggestion is irrelevant to our study for multiple reasons. The first being that our study is based on the 1974 WSP (250,000 km^2), not the 2017 MRP (80,000 km^2), which are two different phenomena with very different in size, and thus the heat flux from the ocean to the atmosphere would be very different (Moore et al., 2002).

Second, as far as we are aware the SOCCOM data for this 2017 polynya are only ocean data. We do not have an ocean component in our model. We prescribed the ocean surface state to infer the atmospheric response. Possibly, the reviewer suggests that we use the SST sampled by SOCCOM to constraint the prescribed SST in our experiments. However, the SOCCOM sampling is extremely sparse in space and time, while we need a gridded SST field. This would be a limited data set to draw on for the atmospheric response to the polynya.

Finally, we do in fact validate our models with ERA5 data in section 3.4. Here we use ERA5 data for two main purposes, one, give a direct comparison based on our model input data, and two, as mentioned in an earlier comment, despite the limitations of using reanalysis, ERA5 is one of the best products we have for the 1974 polynya event.

References

Ayres, H. C., Screen, J. A., Blockley, E. W., & Bracegirdle, T. J. (2022). The Coupled Atmosphere–Ocean Response to Antarctic Sea Ice Loss. Journal of Climate, 35(14), 4665–4685. https://doi.org/10.1175/JCLI-D-21-0918.1

Cheon, W. G., & Gordon, A. L. (2019). Open-ocean polynyas and deep convection in the Southern Ocean. Scientific Reports, 9(1), 1–9. https://doi.org/10.1038/s41598-019-43466-2

Czaja, A., & Frankignoul, C. (2002). Observed Impact of Atlantic SST Anomalies on the North Atlantic Oscillation. https://doi.org/https://doi.org/10.1175/1520-0442(2002)015%3C0606:OIOASA%3E2.0.CO;2

Dare, R. A., & Atkinson, B. W. (1999). Numerical modeling of atmospheric response to polynyas in the Southern Ocean sea ice zone. Journal of Geophysical Research Atmospheres, 104(D14), 16691–16708. https://doi.org/10.1029/1999JD900137

Diao, X., Stössel, A., Chang, P., Danabasoglu, G., Yeager, S. G., Gopal, A., Wang, H., & Zhang, S. (2022). On the Intermittent Occurrence of Open-Ocean Polynyas in a Multi-Century High-Resolution Preindustrial Earth System Model Simulation. Journal of Geophysical Research: Oceans, 127(4). https://doi.org/10.1029/2021JC017672

Gossart, A., Helsen, S., Lenaerts, J. T. M., Vanden Broucke, S., van Lipzig, N. P. M., & Souverijns, N. (2019). An evaluation of surface climatology in state-of-the-art reanalyses over the Antarctic Ice Sheet. Journal of Climate, 32(20), 6899–6915. https://doi.org/10.1175/JCLI-D-19-0030.1

Merchant, C. J., Embury, O., Bulgin, C. E., Block, T., Corlett, G. K., Fiedler, E., Good, S. A., Mittaz, J., Rayner, N. A., Berry, D., Eastwood, S., Taylor, M., Tsushima, Y., Waterfall, A., Wilson, R., & Donlon, C. (2019). Satellite-based time-series of sea-surface temperature since 1981 for climate applications. Scientific Data, 6(1), 1–18. https://doi.org/10.1038/s41597-019-0236-x

Moore, G. W. K., Alverson, K., & Renfrew, I. A. (2002). A reconstruction of the air-sea interaction associated with the Weddell polynya. Journal of Physical Oceanography, 32(6), 1685–1698. https://doi.org/10.1175/1520-0485(2002)032<1685:AROTAS>2.0.CO;2

Royer, J. F., Planton, S., & D~qu~, M. (1990). limud¢ Dynamics A sensitivity experiment for the removal of Arctic sea ice with the French spectral general circulation model*. In Climate Dynamics (Vol. 5). https://doi.org/https://doi.org/10.1007/BF00195850

Weijer, W., Veneziani, M., Stössel, A., Hecht, M. W., Jeffery, N., Jonko, A., Hodos, T., & Wang, H. (2017). Local atmospheric response to an open-ocean polynya in a high-resolution climate model. Journal of Climate, 30(5), 1629–1641. https://doi.org/10.1175/JCLI-D-16-0120.1

Timmermann, R., Lemke, P., & Kottmeier, C. (1999). Formation and Maintenance of a Polynya in the Weddell Sea.

Zheng, C., Y. Wu, M. Ting, J. A. Screen, and P. Zhang, 2023: Diverse Eurasian Temperature Responses to Arctic Sea Ice Loss in Models due to Varying Balance between Dynamic Cooling and Thermodynamic Warming. J. Climate, 36, 8347–8364, https://doi.org/10.1175/JCLI-D-22-0937.1.