

Response to Reviewers

“Dynamical precursors to summer temperature extremes on the Antarctic Peninsula” by Dow et al, submitted to Weather and Climate Dynamics

We thank the Editor for sourcing two detailed and constructive reviews of our manuscript and thank the reviewers for their time in evaluating our work.

Reviewer 1

General comments

This manuscript is very well written and is an interesting and comprehensive survey of the area. The reference list is very impressive, almost too impressive at 10 MS pages of references. I have only a short list of fairly minor comments.

The authors thank the reviewer for their valuable comments and feedback. We address the comments in turn in red font. We hope the responses address the concerns of the reviewer and the result is an improved manuscript.

Minor comments

1. Lines 128-130: Taking an average over the whole region obscures days of foehn warming in the eastern Peninsula, surely? Those conditions have been shown to be associated with extreme high temperatures in the east and with ice shelf collapse (Kyle paper). Are your results sensitive to how the spatial averages are calculated?

We agree with the reviewer that föhn effects contribute to extreme high temperatures in some regions of the Peninsula. However, the relatively coarse spatial resolution of ERA5 means this effect is not fully resolved in our data. This motivates our focus on the role of regional extreme temperatures across the Peninsula and the role of large-scale dynamical precursors. The discussion section notes: “..however, the presence of northwesterly low-level wind anomalies in several clusters suggests there would be a central role of föhn processes in the high AP temperature events identified as found in various studies (Elvidge et al., 2014; Laffin et al. 2021; Turton et al., 2018)”. Furthermore, the k-means clustering was tested on a variety of point locations across the Peninsula to test sensitivity, the results of which showed a high degree of agreement with respect to the clustering and associated dynamical states.

2. Lines 141-142: Reference for Lloyd’s algorithm?

Thanks - included citation and reference

3. Line 189: Fair enough I suppose, you're away from the topography and the winds are strongest in the upper troposphere. But the vertical structure of temperature may well be important, and the warmest air is near the surface. Have you tested the sensitivity of your results to the pressure level chosen, or looked at vertically integrated heat flux?

Tests were conducted using winds at different levels in the troposphere (e.g. 500 hPa), with results all in general agreement but showing varying levels of orographic influence.

4. Lines 207-210: Figure 1(a) suggests the foehn effect is important, and perhaps you're picking it up anyway, even with a larger box-average. Interesting though that the percentage agreement with local extremes is lower in the east where the warming is largest, suggesting that local foehn influences are sometimes or often being missed.

We thank the reviewer for this interesting observation. We agree that the spatial structure in Figure 1a suggests that föhn-related warming may contribute to some events, particularly over the eastern AP. The lower agreement with local extremes in this region may reflect the fact that the spatially averaged event definition used here can smooth highly localised warming associated with föhn processes. As discussed in the response to comment 1, the coarse resolution of ERA5 means it likely underestimates the influence of föhn effects on local temperature extremes and supports our choice to focus on larger-scale anomalies and drivers.

5. Lines 256-257: Figure 3 is interesting, but the MSLP cluster patterns (3a) are hard to see in current format. Could you add contours, maybe have two rows of three to make the panels bigger? Panel (b) could be reduced in size.

Thanks for this suggestion. We have reformatted the plot as suggested to make the data clearer.

6. Lines 271-277: Presumably the clusters with more members contain more of a mixture of patterns and hence the centroids are weaker on average. Is the variance of the MSLP fields related to cluster size, and does that inform the persistence and AP temperature anomaly statistics?

We have now quantified within-cluster variability using the mean Euclidean distance between samples and their assigned centroid. The result is actually opposite to the reviewer's comment, insofar as clusters with more samples (e.g., cluster 2) have lower mean Euclidean distance of individual points, indicating that these states are both more frequently occurring and more dynamically coherent.

7. Lines 276-277: Cluster 2 has the largest number of events, just.

Thanks for spotting this - corrected the text.

8. Lines 302-313: I see what you mean about evolution of the spatial patterns of MSLP, but what a confusing plot Fig 5c is. It takes some thinking about and I haven't come up with a better way of showing it, though perhaps you could plot the spatial correlations between the observed MSLP anomaly pattern and the cluster centroid through time from day -5 to day +3.

We appreciate this suggestion. However, each individual sample will have its own evolution of spatial correlation, so we would need to either plot the distribution of correlations, which would be difficult to interpret, or the average correlation across the samples, which is akin to showing the cluster assignment as in Fig 5c but less clearly connected to the clusters. So after some consideration we have decided to retain Fig 5c but have added an additional explainer in the text: "It should not be interpreted as the evolution of individual events, but rather as an aggregate measure of cluster coherence through time. It therefore quantifies how consistently each synoptic regime is retained relative to its day-0 classification."

9. Lines 376-380: Indeed, this sounds very reasonable.

Thanks for acknowledging this.

10. Lines 394-397: In Fig. 8, there's an impressive amount of consistency among the clusters in the heat flux patterns during days -2 to 0, suggesting that this is a dominant part of the story, as you say.

We thank the reviewer for this observation. We agree that the consistency in the heat flux anomaly patterns across the clusters during days -2 to 0 highlights the important role of anomalous heat transport in the development of these extreme warm events, despite the differing large-scale circulation structures associated with each cluster.

11. Lines 591-593: Repeated reference.

Thanks for pointing this out - the repeated reference has been removed.

Reviewer 2

This paper identifies different synoptic regimes associated with extreme high temperatures over the Antarctic Peninsula, using ERA5 data over 1979-2024. There are several significant issues in this manuscript regarding the methodology, and the figures and text contain several inconsistencies that should be addressed before publication can be considered.

We thank the reviewer for their detailed and constructive comments on various aspects of the methodology used in the manuscript. We respond to their points below in red.

On the clustering;

Days -3 to +1 are used, independently, to define the clusters. How are days before and after classified? This is never clarified in the text (though several figures show -5 to -3 and +1 to +3 lead analysis). If they are labeled into the clusters defined by the -3 to +1 days, since you're only building clusters for the extreme warm events, these clusters are not representative of the full atmospheric space (note how for example there are no circulation patterns that are zonal or cold as identified in Gonzalez et al., 2018).

This is particularly concerning for identifying circulation patterns ahead of the warm extremes outside the +3 to -1 window, especially for the longer leads (-5 days) and the end of the events.

Illustratively, in Figure 6, clusters 2 and 5 at -3 to -5 days have, on average, no discernible circulation patterns associated with them (note the lack of significant SLP anomaly regions), which I would interpret as background variability ahead of the event comprising of a mix of different states. Or put another way, if the atmospheric circulation at -5 days was truly in one of these extreme warm patterns (as implied by figure 5), then at leads -5 days, you would have warm anomalies over the AP, yet for 4 out of 5 'clusters', T anomalies at -5 day leads are close to climatology (as shown in Figure 5A).

Methodologically, this would seem to have a major impact on the interpretation of the analysis.

We thank the reviewer for this important methodological comment. Circulation states at all lags are assigned to the same set of clusters defined using days -3 to +1 - we have clarified this in the text.

The weak composite MSLP anomalies in Fig 6b and 6e are consistent with the broad distribution of cluster assignments in Fig 5cii and 5cv, indicating that, on average, the circulation anomalies at days -4 and -5 do not resemble the assigned cluster on day 0, consistent with the reviewer's comment.

We interpret this as indicating that the large-scale circulation becomes progressively less constrained by the cluster definition at longer lead times, such that these composites increasingly reflect a mixture of background states prior to the emergence of the extreme warm-event circulation anomalies. To quantify this, we evaluated the Euclidean distance between daily MSLP fields and their associated cluster centroids as a function of lag. We find no evidence of a discontinuity at the boundaries of the clustering window (i.e. between included and excluded lags on both the pre- and post-event sides). Instead, distances vary smoothly with lag, with boundary changes comparable to typical day-to-day variability. We will clarify this interpretation in the manuscript and avoid over-interpreting the physical significance of the longer-lead composites.

Added to manuscript:

To assess whether the clustering window constrains circulation states associated with warm extremes, we computed the Euclidean distance between daily MSLP fields and their associated cluster centroids as a function of lag. We tested for discontinuities at the clustering window boundaries (lag -3 to -4 and +1 to +2), which would indicate a mismatch between clustered and unclustered regimes. Instead, distances vary smoothly with lag, with boundary changes comparable to typical day-to-day variability (not shown).

The other main issue with the methodology is that it is never stated how multiday events are assigned to a single cluster. In table 1, each of the 87 events is assigned to a single cluster. But if a multi-day event has single days in different clusters, how is that event assigned to a single cluster? Is it assigned according to the cluster on the peak of the event? Or the onset? How are days after +1 (for events lasting longer than 2 days) classified? Note my earlier comment that, in theory, you're only using -3 to +1 days.

Each event is assigned to a cluster based on the circulation anomalies at day 0; this has been added to the Methods. Fig 5c shows that, for a given event, the circulation anomalies at positive and negative lags can be assigned to different clusters than at day 0; this feature in Fig 5c is discussed in the text and now emphasised further regarding the composite maps in Fig 6 showing weak anomalies at lag -3 to -5 meaning, on average, the circulation does not resemble any cluster.

In addition, there are several inconsistencies in the text and figures as detailed below.

Other comments:

L69-83. It is said that the summer SAM has trended positive in austral summer, and resulted in more northerly flow over the AP. Is this still the case when analyzing 1979-2024? Since the authors find there is no trend in warm events over the AP (L131-L132), has a different circulation mode compensated for an expected SAM-driven increase in warm events?

Also, the description in this paragraph (that a positive trend in summer SAM results in more northerly flow, i.e., warmer AP) is inconsistent with the later discussion on SAM (L442-L445), where the authors state ‘...suggestive of a negative SAM-like state, which is known to weaken the circumpolar westerlies and the storm track equatorward, thereby promoting warm air intrusions into the AP’.

Thanks for highlighting the apparent inconsistency in the description of SAM-related circulation impacts. We agree that the original wording could be interpreted as implying a fixed relationship between SAM polarity and near-surface wind direction over the Antarctic Peninsula. Following the reviewer’s comment and in line with evidence from Marshall et al. (2022), we have revised the text to emphasise that SAM modulates the probability of circulation patterns that favour either poleward or equatorward flow over the Peninsula, depending on the background state. We now clarify that the wind response to SAM variability is regionally variable and can vary depending on the background circulation state.

L90-L92 This is not what Nielsen et al found. They found the AP to be represented by three different clusters, what they call AP (northern peninsula), WA2 (southern peninsula), and SP (east coast of AP); see their Figure 3.

Thank you for the correction. These lines have been reworded to:

Nielsen et al. (2025) applied clustering to surface air temperature trends over the entire Antarctic continent and identified three-clusters that encompassed the AP and analysed their associated extremes. However, they did not specifically examine, the diversity of circulation features amongst AP temperature extremes

L130 ‘The seasonal DJF temperature trend over 1979-2024 is not significant across the AP region in ERA5 (Zhu et al., 2021)’ Typo? (how can a paper published in 2021 analyze 1979-2024 trends?)

Thanks for pointing this out - the correct time span was 1979-2019. The text has been corrected.

L132 and Fig S2. I find Fig S2 very hard to interpret. Why not plot a timeseries of the number of events per year?

We thank the reviewer for this suggestion. We have replaced the previous supplementary figure with a time series of annual event counts per austral summer, which provides a clearer representation of interannual variability and facilitates more direct interpretation of variability and trends.

L133 How sensitive are results to allowing non-extreme days in the middle of the event? Since the average event is so short, do you even need to do this?

The clustering results were insensitive to the inclusion of non-extreme days in the middle of the event. Allowing the intermittent non-extreme days which fell below the threshold prevented a single synoptic event from being artificially split into multiple events.

L170-L173 and figure S3. I am unclear as to why you select $K=5$ instead of $K=4$. The text says that two of the five clusters identified at $k=5$ merge into one at $k=4$, but comparing Fig 3A with S5, I disagree. It looks like $K=4$ clusters match clusters 1, 2, 4, and 5 for $k=5$, as follows (listing the $K=5$ clusters first, and $k=4$ clusters second: 1&4, 2&2, 4&3, 5&1). The only cluster at $K=5$ that doesn't show in $k=4$ is #3, which is the more zonal-like mode with weak meridional circulation. Can the authors clarify the advantage of using $k=4$ over $k=5$?

We thank the reviewer for this comment. The choice of $k=5$ is primarily based on the silhouette coefficient, which shows a local maximum at $k=5$ (Supplementary Fig. S3), indicating improved cluster separation relative to $k=4$. We have revised the manuscript to clarify this point and to better describe the relationship between the different k -values. In particular, we now emphasise that the $k=4$ solution represents a coarser partitioning of the circulation regime space identified at $k=5$. In contrast, the $k=6$ solution is associated with a noticeably reduced silhouette score and primarily results in subdivision of existing regimes (Supplementary Figs. S4 and S5). Importantly, we now explicitly highlight that the $k=5$ solution retains an additional zonal-like circulation regime that is not resolved at $k=4$, which provides additional physical interpretability and supports our choice of $k=5$.

Figure 1 How useful is this clustering mechanism for the ice shelves? Less than 50% of local warm extreme events are captured over the ice shelves along the AP east coast. Is the clustering algorithm missing out on local, east coast warm events?

We thank the reviewer for this question. We agree that the clustering does not capture all locally defined extreme warm events over the eastern Antarctic Peninsula ice shelves. However, this reflects the design choice and spatial scale of the analysis rather than a limitation of the clustering method itself. The k -means clustering is applied to large-scale mean sea level pressure anomalies associated with Peninsula-wide temperature extremes, rather than to localised grid scale variability. As such, it is primarily intended to characterise synoptic-scale circulation regimes that precede regional extreme temperature events, rather than to resolve locally driven processes such as föhn warming at individual ice shelves.

We have clarified in the manuscript that localised processes, including föhn effects, are not fully resolved in ERA5 at the spatial scale considered, and therefore may not be consistently captured within the large-scale clustering framework. Nevertheless, we note that several of the identified circulation regimes do exhibit low-level northwesterly flow anomalies consistent with föhn-favourable conditions, indicating that these processes are still embedded within a subset of the large-scale patterns. Sensitivity tests applying the clustering framework to different representative locations across the Peninsula further show

that the resulting circulation regimes are broadly consistent, supporting the robustness of the large-scale classification.

L176-L189 This approach neglects the contribution of adiabatic warming and diabatic heating to temperature extremes, (e.g., Röthlisberger and Papritz, 2023); note also that latent heat release can be a major contributor to Antarctic heat extremes (Blanchard-Wrigglesworth et al., 2023). This limitation should at a minimum be acknowledged.

We thank the reviewer for this comment. We agree that horizontal thermal advection represents only one contribution to the temperature tendency and does not provide a complete temperature budget. Our intention was not to attribute the temperature anomalies solely to horizontal thermal advection, but rather to use it as a diagnostic of the large-scale dynamical forcing associated with the identified circulation regimes. We have therefore added text to the manuscript to the methods and discussion acknowledging that adiabatic processes and latent heating may also contribute to the observed temperature anomalies and are not explicitly quantified in the analysis.

“ However, the thermal advection diagnostic presented here does not constitute a complete temperature tendency budget. Adiabatic temperature changes associated with vertical motion and diabatic heating will also contribute in part to Antarctic temperature extremes. Furthermore, latent heat release associated with moist heat fluxes has been identified as an important contributor to Antarctic heat extremes (e.g. Blanchard-Wrigglesworth et al., 2023) but is not explicitly quantified here”

Figure 2. Label the SLP contours (impossible to tell what values are, since we are not told if the zero value is plotted, or if not, if authors plot at +/- 1 mb.

Thanks for the suggestion. The plot has been updated.

L253 Isn't cluster 5 closer to a zonal wavenumber 4 pattern? The high and low SLP anomalies look to be 90 degrees longitude apart, not 120 degrees (which would be ZW3). Why aren't all longitudes plotted?

We thank the reviewer for this comment. The clustering and associated centroid diagnostics are defined over the same hemispheric domain used throughout the analysis. As such, the spatial patterns in Fig. 3 are strictly consistent with the input domain used to derive both the clusters and the zonal wavenumber spectra. We agree that visual interpretation of zonal wave structure can be sensitive to the displayed longitude range; however, extending the plots to the full 360° domain would require redefinition of the clustering input space and is therefore beyond the scope of this study. We have clarified the text at L253 to read:

“Cluster 5 exhibits a wave-like circulation pattern with a dipole structure flanking the Antarctic Peninsula. The zonal wavenumber spectrum indicates broadly enhanced power

across low wavenumbers (ZW2–ZW4), consistent with a broadband wave structure rather than a single dominant zonal mode (Fig. 3av and Fig. 3b).”

Figure 3. Make panels in top row larger, they are hard to read, add contours, and show all longitudes? Consider also adding wind vectors.

Thanks, Figure 3 has now been updated based on this feedback and similar feedback from Reviewer 1.

Are the zonal wavenumber power spectrums calculated just using the anomalies over half the longitudes (as plotted in Fig 3A), or over all longitudes?

The zonal wavenumbers are calculated over the longitudes included in the k-means clustering methodology, i.e. over the hemispheric domain. This has been clarified in the caption of Figure 3.

Figure 5 Show longer positive lags; cluster 1 is still p95 at day 3. In panel B, why are growth rates for cluster 1 negative at all positive lags, if in panel A, cluster 1 day 3 is warmer than day 2? Also, in panel A, cluster 1 goes from 2K at day -1 to 5K at day 0, so a 3K/day dT/dt . But in panel B, the growth rate for cluster 1 peaks at 2K/day. Also, it is not defined if the growth rate for day 't' is $T(t)-T(t-1)$, or $T(t+1)-T(t)$, or something else.

Growth rates were calculated using a centred finite-difference scheme method, providing an estimate of the local temporal gradient rather than a forward difference between consecutive lag days. This has now been clarified in the Methods and Figure caption.

Figure 6 show all SLP contours? Seems odd that in some panels, the higher amplitude SLP anomalies are not plotted. I would also plot a wider colorbar axis interval; the current intervals (-3 to 3C) implies that all the clusters are beyond the colorbar at the peak of the events (all clusters peak above 4C anomalies, table 1), making it impossible to differentiate between them over the AP.

We have updated Fig. 6 to extend the temperature anomaly colour scale to $\pm 5^{\circ}\text{C}$ to better capture the full range of variability and improve inter-cluster comparability. The masking of MSLP contours at non-significant grid points has been clarified in the caption to avoid ambiguity regarding discontinuities in the contour fields.

Figure 7 The near-surface wind anomalies are inconsistent with the SLP anomalies for the clusters (as shown in Figure 1 and 6). For example, for cluster 2, at the -2 to zero lag, we are told there are anomalous southerly winds over the AP (which should be associated with cold temperatures), but from the SLP pattern shown in Figure 1 and Figure 6, one would expect NW winds over the AP, not southerly winds.

In cluster 2, there are strong anticyclonic wind anomalies centered over the AP at -5 to -3 day leads (fig 7F), but no SLP anomalies associated with this wind pattern? (Fig 6B). How can that be?

Equally, in cluster 1, one would expect anticyclonic winds around the SLP high over the Bellingshausen sea, with southerly winds over the eastern Weddell (east of the high, see Fig 1), and yet, Fig 7 has strong northerly winds in the eastern Weddell. In short, the wind patterns in Figure 7 seem inconsistent with the SLP patterns.

We thank the reviewer for pointing out this discrepancy. Upon scrutiny, this was a mistake in the plotting function. This has now been rectified and the correct plots have now been produced and added to the manuscript. The results are now in line with what one would expect based on earlier figures. We should make clear that the plots in Figure 7 are of 500hPa Geopotential Height anomalies, rather than SLP anomalies, which the reviewer refers to in their comments.

References

Blanchard-Wrigglesworth, E., Cox, T., Espinosa, Z.I. and Donohoe, A., 2023. The largest ever recorded heatwave—Characteristics and attribution of the Antarctic heatwave of March 2022. *Geophysical Research Letters*, 50(17), p.e2023GL104910.

Röthlisberger, M. and Papritz, L., 2023. Quantifying the physical processes leading to atmospheric hot extremes at a global scale. *Nature Geoscience*, 16(3), pp.210-216.