Reviewer 2:

This manuscript examines a set of extreme swell events along the coast of Peru and connects them to jet characteristics in both the Southern and Northern Hemisphere. The recorded extreme swells coincide with strong wind events in either hemisphere. The circulation patterns are then related to a larger sample of events using an analogue approach with an accounting for the long-term changes in wind speed between the mid 20th century and the more recent period. The connections between the remote drivers of extreme swell events is important for planning purposes and the current work is a nice addition to this understanding with some examination of the historical changes in these drivers that could inform potential future projections of extreme swell events. I have some broad suggestions related to the

We thank the reviewer for the careful review and constructive comments. We appreciate the positive assessment of the manuscript and the recognition of the relevance of understanding the remote drivers of extreme swell events and their historical variability. Your suggestions have been very helpful in improving the quality and clarity of the manuscript. Detailed responses to all comments are provided below.

Main comments:

1. There is a relatively small sample size of these events presented and presumably available from the Peruvian Navy records. Especially in the case of the Northern Hemisphere generated events, this introduces a very large degree of uncertainty in the types of events and the background conditions. There is discussion related to the PDO and IPO, but all the events are during 2010 or later, and thus are likely not sampling a representative range of the full variability associated with these modes or their trajectories. It's not clear to me that the Southern Hemisphere analogue, where a climate change signal is detected, is representative of the swell events in general. Additional discussion related to these limitations could be added to the analogue analysis and discussion.

The observed extreme swell events are indeed all recorded after 2010, as they were identified from recent Peruvian Navy records. We fully agree that this limits the number of events available and introduces uncertainty. As explained in our response to Reviewer 1, we intentionally did not expand the observed-event sample using pre-1980 historical records or swell events reported for the Chilean coast because those records are not classified under the same homogeneous DIHIDRONAV operational warning framework used for the Peruvian coast. Including them would have increased the sample size, but at the cost of reducing the internal consistency of the event definition. However, the analogue search was performed over the full period 1950–2025, with analogues drawn from two sub-periods spanning several decades: 1950-1987 and 1988-2025. Therefore, they do sample both positive and negative states of the PDO and IPO. The core comparison in our analysis is between these historical analogues and their recent counterparts. We note that this analogue-based approach is methodologically consistent with the ClimaMeter framework (Faranda et al., 2024), which has been successfully applied to attribute individual extreme weather events to climate change using a limited number of observed events and historical analogues drawn from reanalysis data spanning several decades, supporting the validity of our approach even given the constraints of the available dataset.

We acknowledge, however, that the limited number of observed events means we are working with a small set of circulation patterns, and that some NH patterns may be more or less representative of the full range of variability. We have added a sentence to the discussion explicitly noting these limitations and cautioning against over-interpretation, particularly for NH events where the wind and SWH signals are weaker and less consistent **(L527-534)**.

For the SH events, we maintain that the detected signal is robust, as it is consistent across all events, physically coherent with the well-documented positive trend in the SAM, and now further supported by the statistically significant increases in SWH shown in the new boxplot analysis (please see the next answer).

Reference: Faranda, D., Messori, G., Coppola, E., Alberti, T., Vrac, M., Pons, F., Yiou, P., Saint Lu, M., Hisi, A. N. S., Brockmann, P., Dafis, S., Mengaldo, G., and Vautard, R.: ClimaMeter: contextualizing extreme weather in a changing climate, Weather Clim. Dynam., 5, 959–983, https://doi.org/10.5194/wcd-5-959-2024, 2024.

2. The authors do account for this somewhat with the flow analogues, but the primary novelty of the manuscript is the connection between extratropical cyclones or large-scale circulation patterns and the swells from this specific dataset. It would be ideal to make more of a connection between the analogues and the swell data where and if possible.

We thank the reviewer for this suggestion. To more explicitly connect the atmospheric circulation analogues with the swell dataset, we have added a comparison of regional mean SWH between past and present analogue days for each event.

For the SH events, where we previously identified a consistent increase in 10m wind speed across all events between the two periods, the significant wave height (SWH) analysis reveals a corresponding significant increase in all events except E2. In the case of E2, a positive signal is still present but is spatially more localized, which explains the lack of regional statistical significance.

For the NH events, the SWH results closely mirror the wind speed findings: the three events associated with a decrease in 10m winds show no significant change in SWH, while the three events with increasing winds show a statistically significant increase in SWH (two at the 5% level, one at $p \approx 0.08$). The absence of a significant SWH response to decreasing wind speeds in NH events is consistent with the nonlinear, power-law nature of the wind-wave relationship (Gao et al., 2023). This nonlinearity produces an asymmetric response: wind speed increases above a given threshold generate disproportionately larger wave height anomalies than equivalent decreases, particularly when background wind conditions are already moderate.

This consistency between the wind and SWH signals across both hemispheres and all events strengthens the physical link between the large-scale circulation patterns captured by the analogues and the observed changes in swell characteristics. We have incorporated these analyses in the result section **(L258-260, L405-406)** and the figures into the Supporting Information (Figs. S7, S18).

Reference:
Gao Y, Schmitt FG, Hu J and Huang Y (2023) Probability-based wind-wave relation. Front. Mar. Sci. 9:1085340. doi: 10.3389/fmars.2022.1085340

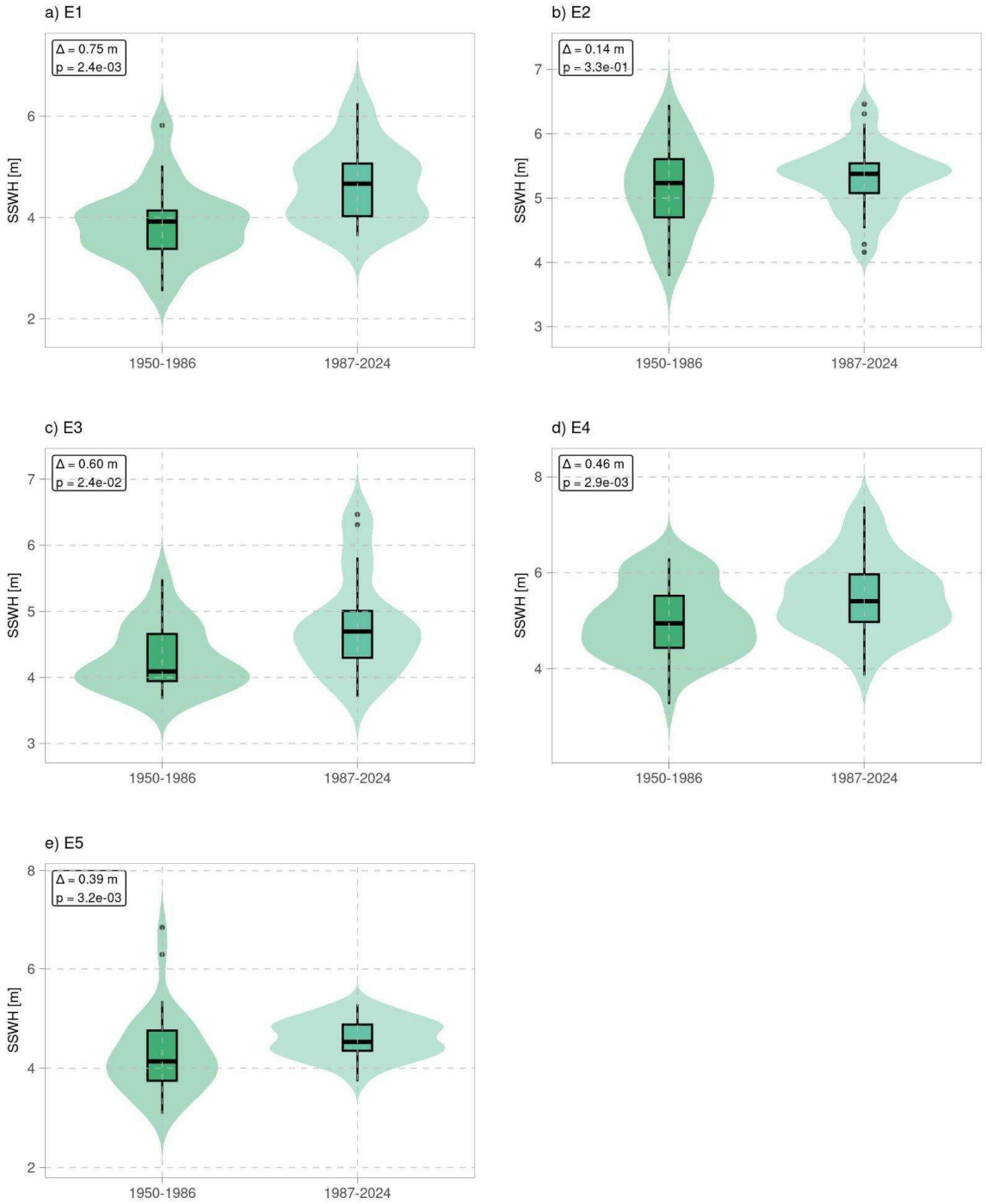

Figure S7. Violin–boxplots of regional mean significant wave height (SWH) associated with Southern Hemisphere-origin events. The distributions compare analogue days from the past period (1950–1986) with those from the present period (1987–2024). The legend shows the difference in medians between the present and past periods for SSWH for days with similar flow and its statistical significance according to the Mann-Whitney U test.

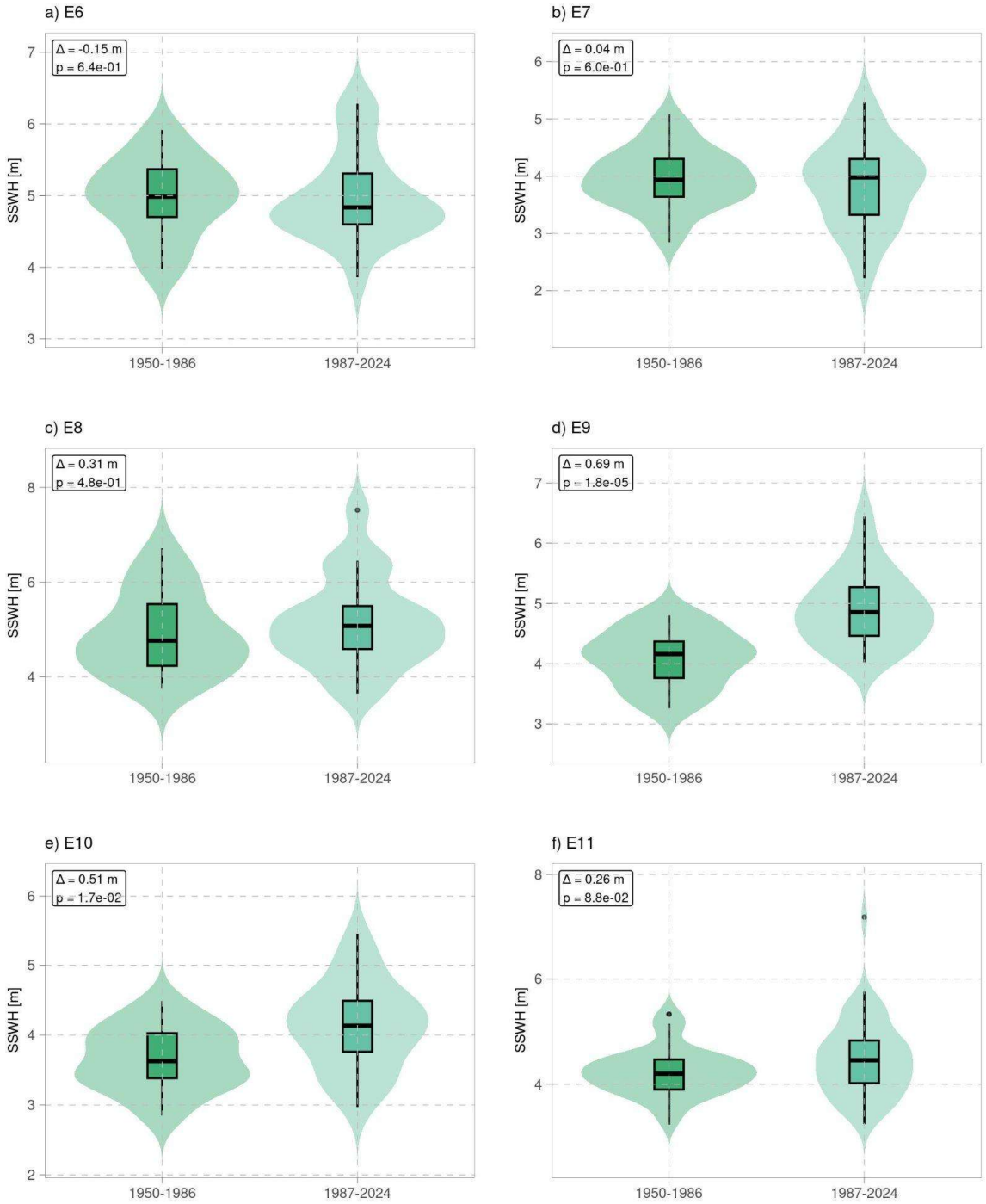

Figure S18. Same as Figure S7, but for events originating in the Northern Hemisphere.

3. While the statistical tests suggest significance, the box plots for extreme events in Figure 2 and 5 are still within the spreads of the climatological boxplots in all cases. It is not clear that there is a distinction being made between the two groups and it does not necessarily highlight any of the jet metrics being examined. These figures require revisiting or expanded discussion to make their point. The authors could consider removing these figures because the primary conclusions would not be impacted.

We agree that the boxplots do not imply fully separated populations and, in the revised manuscript, we have softened the wording accordingly. However, we prefer to retain Figures 2 and 5 because their purpose is descriptive: they show the direction of the shift of the extreme-event sample relative to the full climatological distribution, while the statistical inference is based on non-parametric tests. To address the referee's concern, we performed an additional balanced-resampling sensitivity test in which the climatological sample was repeatedly randomly downsampled to the same size as the extreme-event sample and the Mann–Whitney and Kolmogorov–Smirnov tests were recomputed **(L236-243, L377-385)**. The key parameters highlighted in the manuscript remained significant in this balanced framework **(Table S4)**. We therefore keep the boxplots in the main text, but we revised the interpretation to emphasize a systematic shift toward preferential portions of the climatological distribution rather than a complete separation between populations (See also answer to Observation #8).

Parameter	Hemisphere	Extreme value	Climatology value	Median difference	Balanced MW p	Balanced KS p	Retained significance
PFJ Intensity	Southern	45.285 m/s	37.408 m/s	7.878 m/s	0.0329	0.0420	YES
PFJ Sharpness	Southern	24.250 m/s / °	16.838 m/s / °	7.682 m/s / °	0.0229	0.0167	YES
PFJ Departure	Southern	3.672 °	7.971 °	-4.299 °	0.0107	0.0090	YES
STJ Sharpness	Northern	33.718 m/s / °	26.001 m/s / °	7.717 m/s / °	0.0000	0.0001	YES
STJ WB Intensity	Northern	88.905 m/s	79.753 m/s	9.151 m/s	0.0004	0.0001	YES
STJ Intensity	Northern	50.856 m/s	44.653 m/s	6.203 m/s	0.0006	0.0021	YES
STJ WB Longitudinal Extent	Northern	62.500 °	47.000 °	15.500 °	0.0009	0.0022	YES
STJ EB Intensity	Northern	66.042 m/s	55.075 m/s	10.967 m/s	0.0167	0.0017	YES
STJ WB Tilting	Northern	0.071 °/°	0.012 °/°	0.059 °/°	0.0220	0.0153	YES

PFJ WB Intensity	Northern	42.631	56.910	-14.280	0.0009	0.0001	YES
PFJ WB Longitudinal Extent	Northern	20.500	32.000	-11.500	0.0065	0.0277	YES
PFJ EB Intensity	Northern	47.034	53.019	-5.985	0.0120	0.0401	YES

Table S4. *Robustness of jet-metric significance under balanced climatological resampling for Southern and Northern Hemisphere extreme-swell events.* For each parameter, the climatological sample was repeatedly randomly subsampled to match the size of the extreme-event sample, and the Mann–Whitney U and Kolmogorov–Smirnov tests were recomputed. "Median difference" denotes Extreme minus Climatology. "Retained significance" indicates parameters that remain statistically significant under the balanced framework according to the criterion defined in this study.

Specific comments:

4. Table 1: The classifications for the Southern Hemisphere events are all "Very Strong" versus only "Strong" in the Northern Hemisphere. Can you provide more information on the definition of "extreme" and "strong" swells? Please make clear the distinction between the classifications taken from the Peruvian Navy records and those in the Northern Hemisphere.

We agree that the previous version did not distinguish clearly enough between the generic term *extreme-swell events* used in this study and the operational intensity classes reported by DIHIDRONAV. In the revised manuscript, we clarify that *strong* and *very strong* are DIHIDRONAV coastal-impact categories assigned to the warnings listed in Table 1, whereas *extreme* is used here as a broader term for the selected anomalous swell episodes affecting the Peruvian coast. Accordingly, the Southern Hemisphere subset analysed in this study consists of southwesterly events classified operationally as *very strong*, while the Northern Hemisphere subset consists of northwesterly events classified as *strong*. We also revised Table 1 and Section 2.1 to make this distinction explicit and to avoid implying that both hemispheres were sampled from the same operational warning class **(L65-70)**.

5. Line 71-72: It wasn't initially clear that the swells were generated in the Northern Hemisphere and registered along the Peruvian coast. Please make this clearer.

We revised the sentence to clarify that the swell events originated in the North Pacific and were subsequently observed along the Peruvian coast. The revised text now reads:

"We similarly identified a set of extreme swell events that originated in the North Pacific and propagated to the Peruvian coast over the past ~17 years." **(L75-77)**

6. Equation 1: Nitpicking but could you make the g in Gg a subscript to make clearer it is group velocity not another gravitational constant.

Done. **(L84)**

7. Line 96-97: Were there any considerations for modifying the approach for characterizing jet stream characteristics when applying it to the Northern Hemisphere, since Collazo et al. (2024) was only focused on this Southern Hemisphere region near South America.

While the core algorithmic logic of Collazo et al. (2024) remains valid for identifying wind maxima, specific adaptations were implemented for the NH to account for its distinct geographical and dynamical features. In particular, to address geographical differences, the algorithm was modified to account for the opposite meridional configuration of the jets in each hemisphere: while in the NH the PFJ is located north of the STJ, in the SH the PFJ lies south of the STJ. Accordingly, the sign of the latitudinal threshold used to distinguish the two jets was inverted.

From a dynamical perspective, the stronger influence of land–sea contrasts in the NH led us to separate the analysis into western and eastern branches. This differs from the Southern Hemisphere, where our analysis is restricted to the central and western South Pacific and therefore did not require such a division.

We have updated the text in **lines 116-121** to clarify these methodological adjustments.

8. Line 196: Are you accounting for different numbers of samples in the MW and KS statistical tests? Qualitatively the extreme distribution doesn't look like a distinct distribution from the climatology, though the tests suggest they are and I'm concerned that is due to sampling.

Yes. The original Mann–Whitney and Kolmogorov–Smirnov tests were applied to samples of unequal size, which is statistically allowed for both tests. Nevertheless, given the small number of extreme events, we agree that an explicit robustness check is appropriate. We therefore added a balanced-resampling analysis in which the climatological sample was repeatedly randomly subsampled to match the size of the extreme-event sample, and the tests were recomputed for the same jet metrics. The main significant results reported in the manuscript remained significant under this balanced framework, indicating that the reported shifts are not explained solely by sample-size imbalance **(L236-243, L377-385)**.
Furthermore, the analysis was extended by incorporating the analogue days of each extreme event, thereby increasing the sample size and further strengthening the robustness of the results. The climatological sample was repeatedly randomly subsampled to match the size of the analogue

sample and the tests were recomputed over 1000 iterations for each jet metric, using the median p-value and the fraction of significant iterations as summary statistics. The reported differences remained robust under this framework. Most jet parameters showed significant differences between the analogue days and the climatology in both hemispheres, with the exception of the PFJ departure in the SH, and the PFJ and STJ east branch intensity in the NH **(Figs. S6, S17)**. These differences may reflect the greater variability inherent to those parameters or the limited number of events available for each hemisphere.

These additional results are now documented in new Supporting Information tables.

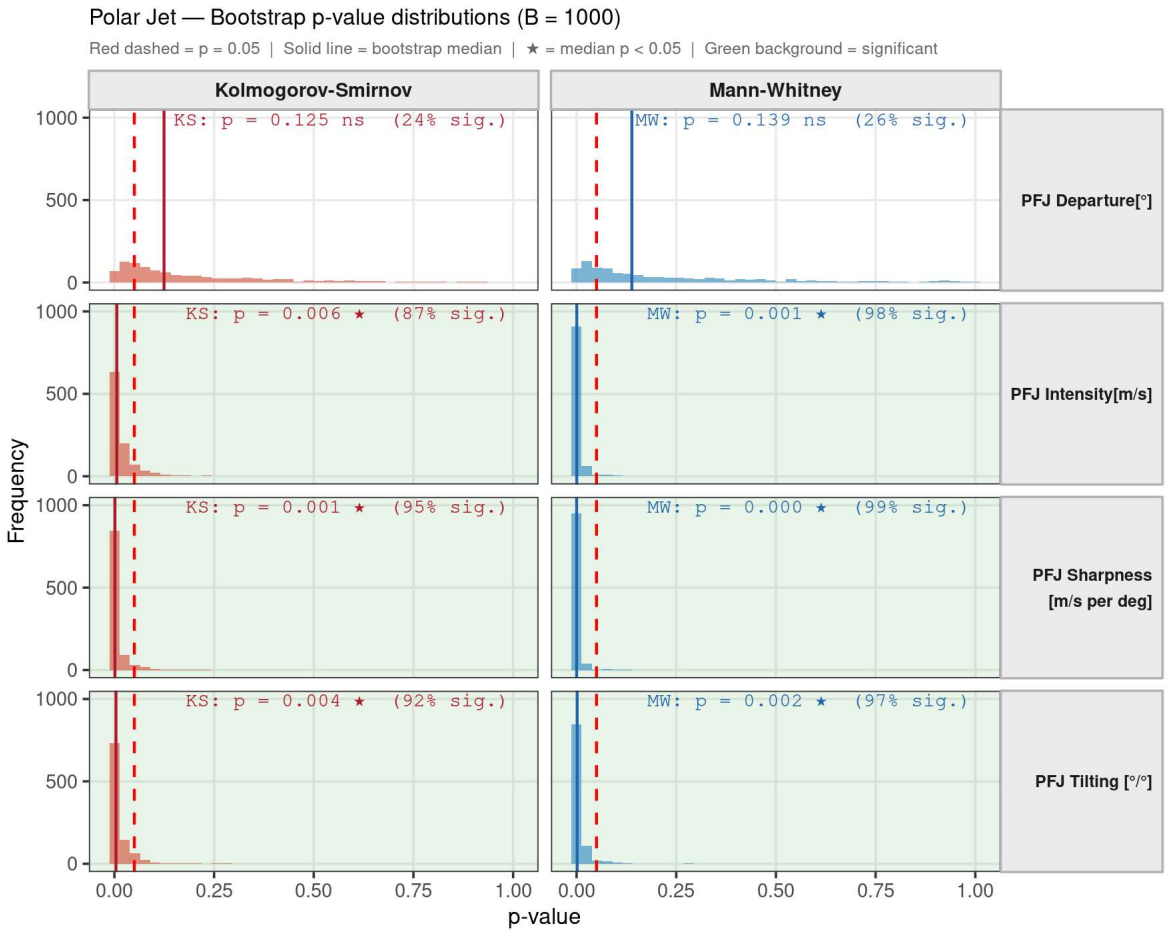

Figure S6. *Bootstrap distributions of p-values from Mann-Whitney (blue) and Kolmogorov-Smirnov (red) tests comparing polar jet parameters between analogue days of extreme events originated in the SH and the 1991–2020 climatology*, across 1000 bootstrap iterations in which a random subsample of climatological days (equal in size to the analogue set) was drawn without replacement. The red dashed vertical line marks p = 0.05; the solid vertical line indicates the median p-value across iterations. Green shading denotes parameters for which the median p-value is below 0.05 (★). Annotations show the median p-value and the fraction of iterations yielding a significant result.

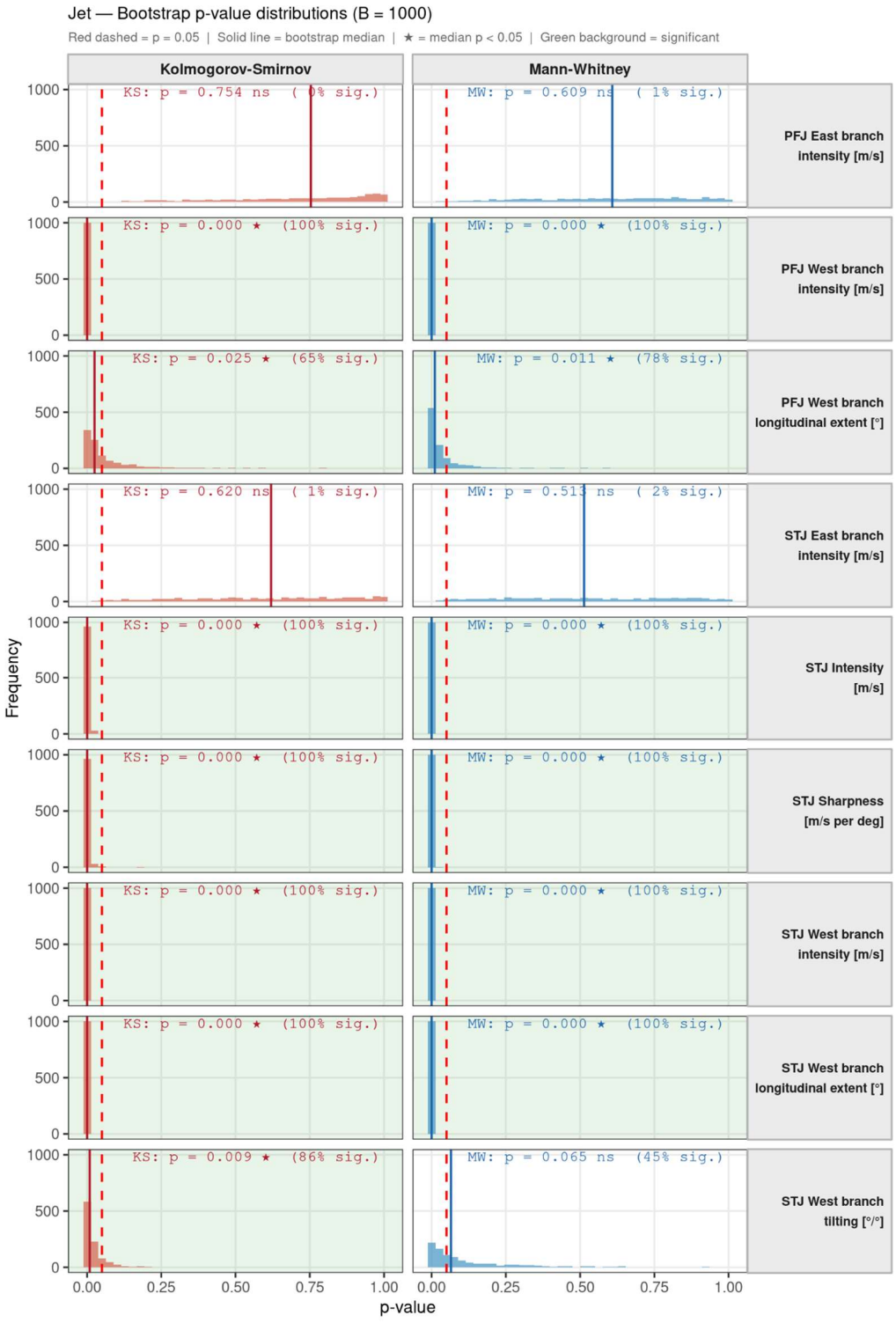

Figure S17. *Bootstrap distributions of p-values from Mann-Whitney (blue) and Kolmogorov-Smirnov (red) tests comparing subtropical and polar jet parameters between analogue days of extreme events originated in the NH and the 1991–2020 climatology,* across 1000 bootstrap iterations in which a random subsample of climatological days (equal in size to the analogue set) was drawn without replacement. The red dashed vertical line marks p = 0.05; the solid vertical line indicates the median p-value across iterations. Green shading denotes parameters for which the median p-value is below 0.05 (★). Annotations show the median p-value and the fraction of iterations yielding a significant result.

9. Line 205: How similar are the SLP patterns for the other events? The SLP patterns in Fig. 3 look quite different from those in Fig. 1, so I'm wondering how representative they are of the other events.

Figure 1 and Figure 3 are not intended to represent the same level of analysis. Figure 1 shows the mean composite pattern for all Southern Hemisphere events (E1–E5), whereas Figure 3 presents the analogue analysis for a single illustrative case (E1). Therefore, some differences in the exact position, amplitude, and shape of the SLP anomalies are expected, because the composite smooths event-to-event variability while the single-event panel retains the full structure of an individual case.
Nevertheless, E1 preserves the same large-scale ingredients identified in the composite: a deep southeast Pacific low, surrounding higher-pressure anomalies, and an enhanced meridional pressure gradient favouring strong southwesterly winds over the fetch region. We have revised the text to clarify that Figure 3 should be interpreted as an illustrative event-level example embedded within the broader composite framework, rather than as a pattern expected to reproduce Figure 1 exactly.

10. Line 216: Perhaps checking an Amundsen Sea Low index in addition to the SAM would be more targeted based on the composite patterns. It might not give substantially different interpretations, being highly correlated with the SAM, but might provide a more nuanced regional signal or predictive power.

Thank you for this valuable suggestion. We explored the use of the Amundsen Sea Low (ASL) indices as an additional regional circulation indicator. However, the available ASL index data begin in 1959 (https://scotthosking.com/asl_index), whereas our analogue search period starts in 1950; consequently, the ASL indices do not span the full study period.

Despite this limitation, we performed the analysis over the overlapping period and obtained several noteworthy results. In particular, E3, which does not show a significant relationship with the SAM, exhibits significant differences between past and present analogues in both ASL longitude and ASL Relative Central Pressure (Figs. S9-S10). The latter is defined as a regional pressure anomaly obtained by subtracting the actual ASL central pressure from the area-averaged pressure over the ASL domain, thereby isolating the local circulation signal from the hemispheric-scale SAM influence.

An opposite pattern is found for the Amundsen Sector Pressure index (Figure S11), which is strongly correlated with the SAM: the other SH events display significant differences under this index, whereas E3 does not.

These results suggest that E3 may be influenced more strongly by regional ASL dynamics than by the large-scale SAM signal, providing a more nuanced interpretation that complements the SAM-based analysis. Nevertheless, because the SAM covers the full study period and remains the most widely used large-scale circulation index in the literature, we retain it as the primary circulation metric in the main text. A brief discussion of the ASL-based results has now been incorporated into the Section 3.1.3 of the revised manuscript **(L287-289).**

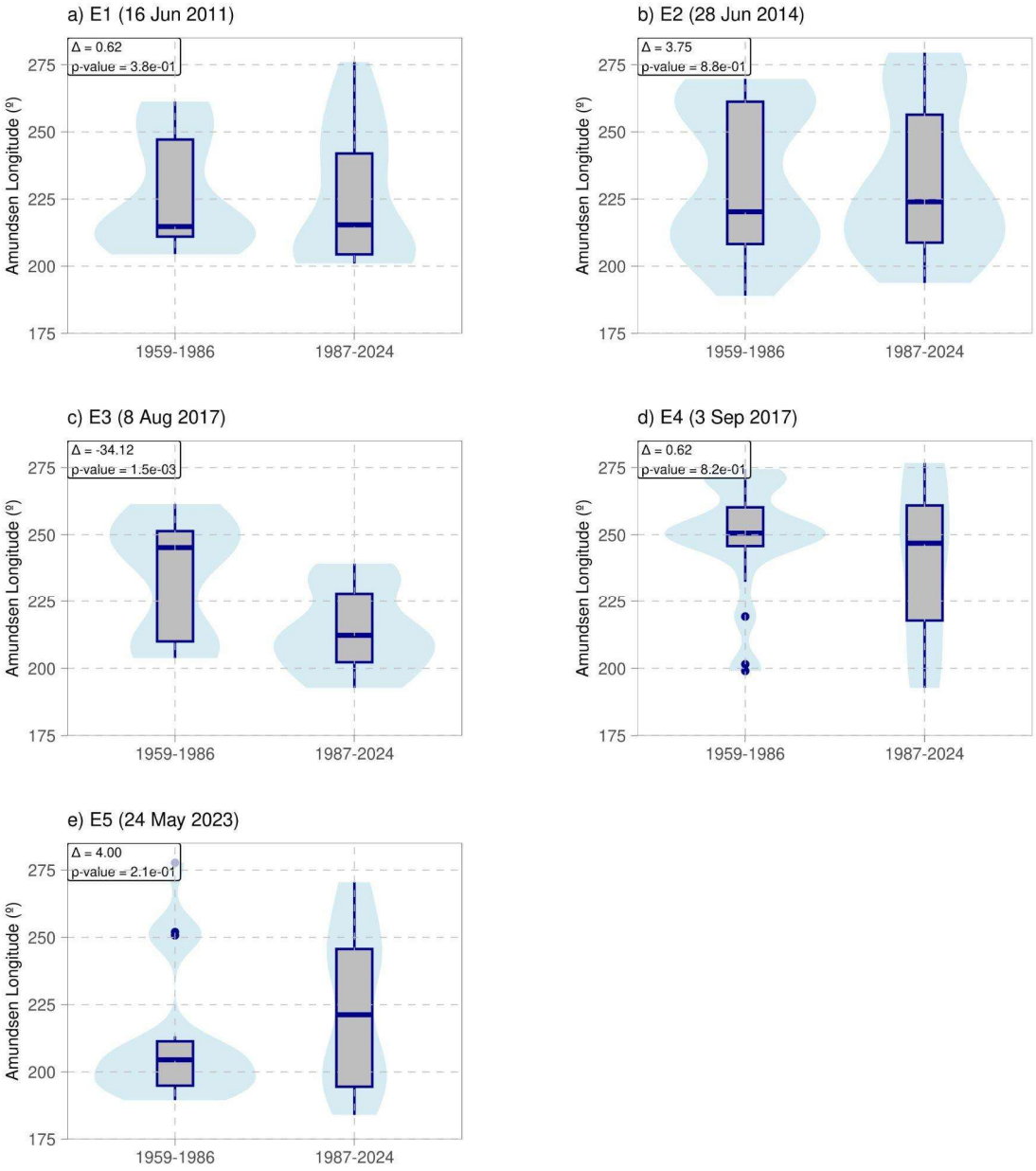

Figure S9. *Amundsen Sea Low (ASL) longitude for past and present analogues. Monthly ASL longitude index associated with days of flow analogues for the past (1959–1986) and present periods, shown separately for each event. Note that the past period is shorter than in the main analyses due to the availability of the ASL index from 1959 onward. The median difference between periods and their statistical significance are indicated within each box. Source of ASL index: https://scotthosking.com/asl_index*

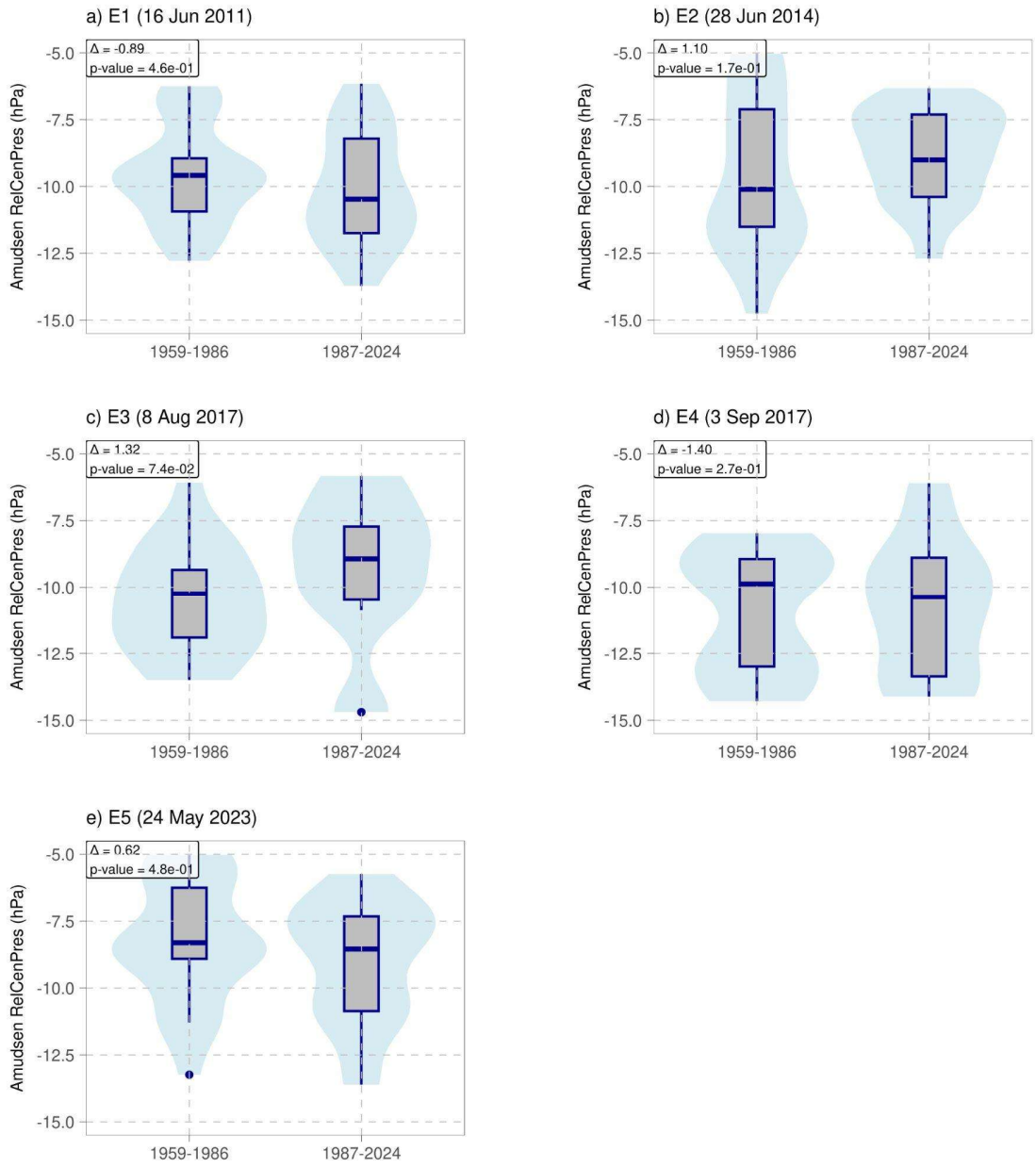

Figure S10. *Same as Figure S9 but for the ASL Relative Central Pressure, which measures the regional pressure anomaly after removing the hemispheric SAM signal.*

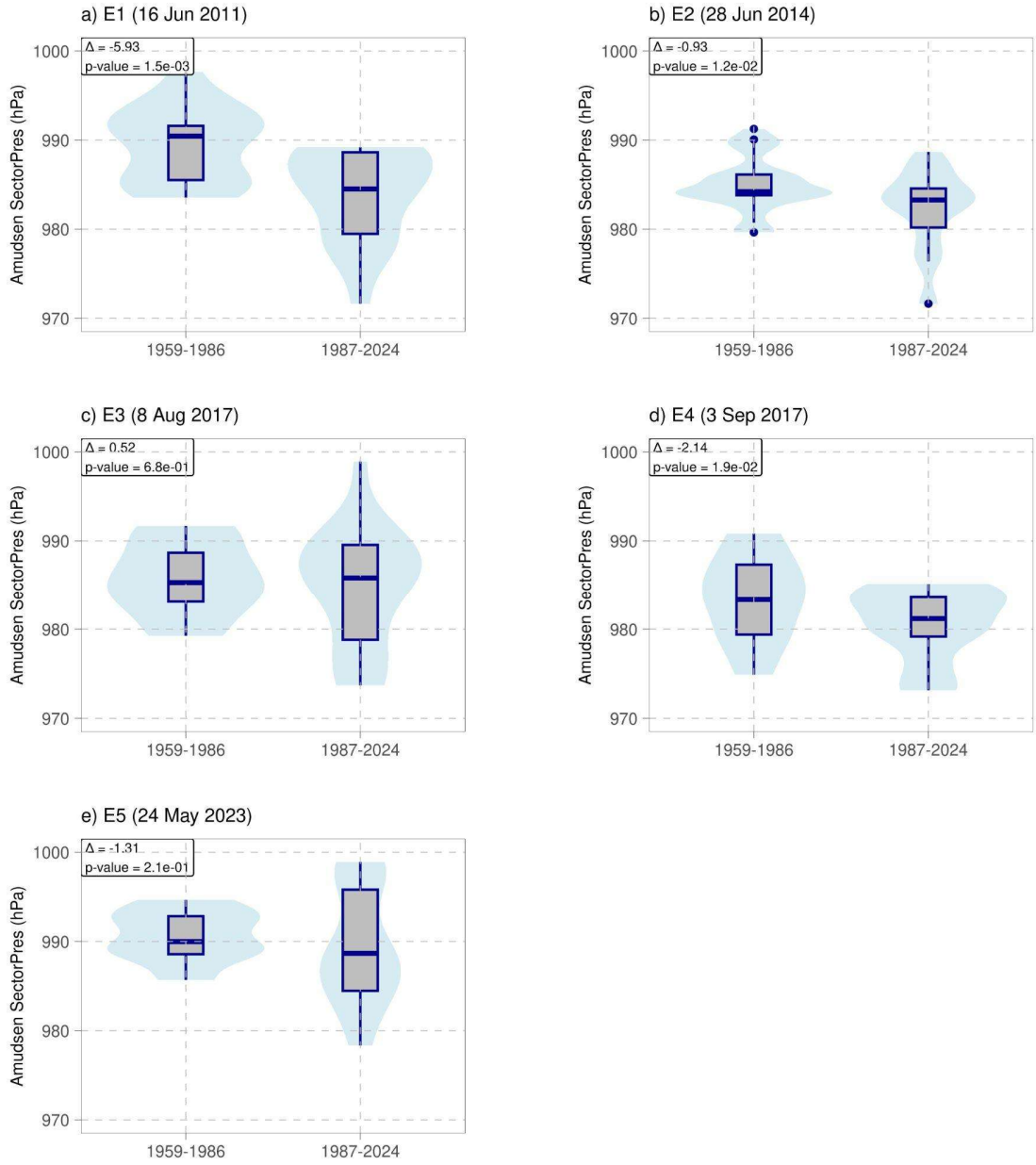

Figure S11. *Same as Figure S9 but for the ASL Sector Pressure, which represents the area-averaged pressure over the ASL domain and is strongly correlated with the large-scale SAM signal.*

11. Line 223: It would be good to introduce the counterfactual approach here in the text before discussing the results, rather than only in the figure caption.

We agree that the concept of counterfactual reconstruction should be introduced in the main text before discussing the results **(L249-252)**. We have therefore revised the paragraph to explicitly define the counterfactual approach prior to presenting the interpretation of Fig. 3b, rather than relying solely on the figure caption.

12. Line 224-225: In addition to changes in the SAM, studies have shown strengthening of the Southern Hemisphere over the satellite period that are expected to continue in the future in response to anthropogenic forcing (Priestley and Catto, 2022).

We have incorporated the point about the observed strengthening of SH extratropical cyclones and their projected continuation under anthropogenic forcing into the revised text, citing Priestley and Catto (2022).

L290-292: *"Beyond changes in the SAM, SH extratropical cyclones have also shown a strengthening trend over the satellite period, and are projected to further intensify under continued anthropogenic forcing, with stronger wind speeds throughout the troposphere and an expansion of the area of extreme winds in both winter and summer seasons (Priestley and Catto, 2022)."*

13. Line 226: The timing of the SAM changes and their relationship with the stated external forcings is somewhat difficult to gauge given the time period available in this study. But an additional factor that might be interesting in this case is the rapid decline of sea ice cover beginning around 2016. Perhaps this exposed more ocean surface and increased fetch in the region.

Regarding the SAM, we acknowledge that the observational period (2008–2025) is relatively short to robustly characterize its long-term influence. To address this limitation, we extended the analysis using analogue events identified in the longer reanalysis record. When comparing SAM values associated with months containing analogues in the past (pre-1986) versus the present period (1986–2025), we find a statistically significant difference, suggesting that the large-scale circulation background has indeed shifted toward more positive SAM conditions in recent decades. This supports the notion that SAM changes may be contributing to the increased intensity of the analysed extreme events, even if the direct causal link is difficult to establish over the available observational window.

The potential role of sea ice decline as an additional forcing is indeed a plausible mechanism, as reduced sea ice coverage could expose more open ocean surface and increase fetch in the region. We examined seasonal (MJJAS) 10-m wind composites for the periods before and after 2016 (1950–2015 vs. 2017–2025). While some differences are apparent around 60°S, these are not statistically significant over most of the study region (Fig. AR1). Furthermore, when comparing 10-m winds on analogue days in the present-climate period (1986–2025) between the pre- and post-2016 subperiods, differences are again generally not significant (Fig. AR2). The exception is E3, which shows some regions with increased wind speeds after 2016, consistent with the mechanism proposed by the reviewer, though spatially limited. For E5, no analogues were found after 2016, precluding this comparison for that event.

Overall, while we cannot rule out a contribution from sea ice loss, particularly for E3, the results suggest that this mechanism does not appear to dominate the wind signal in the analysed events at the scales considered. Given the lack of a robust and consistent signal across events, we considered that adding a dedicated discussion of this aspect to the manuscript would not be well supported by the available evidence, and have therefore kept this analysis as part of our response to the reviewer.

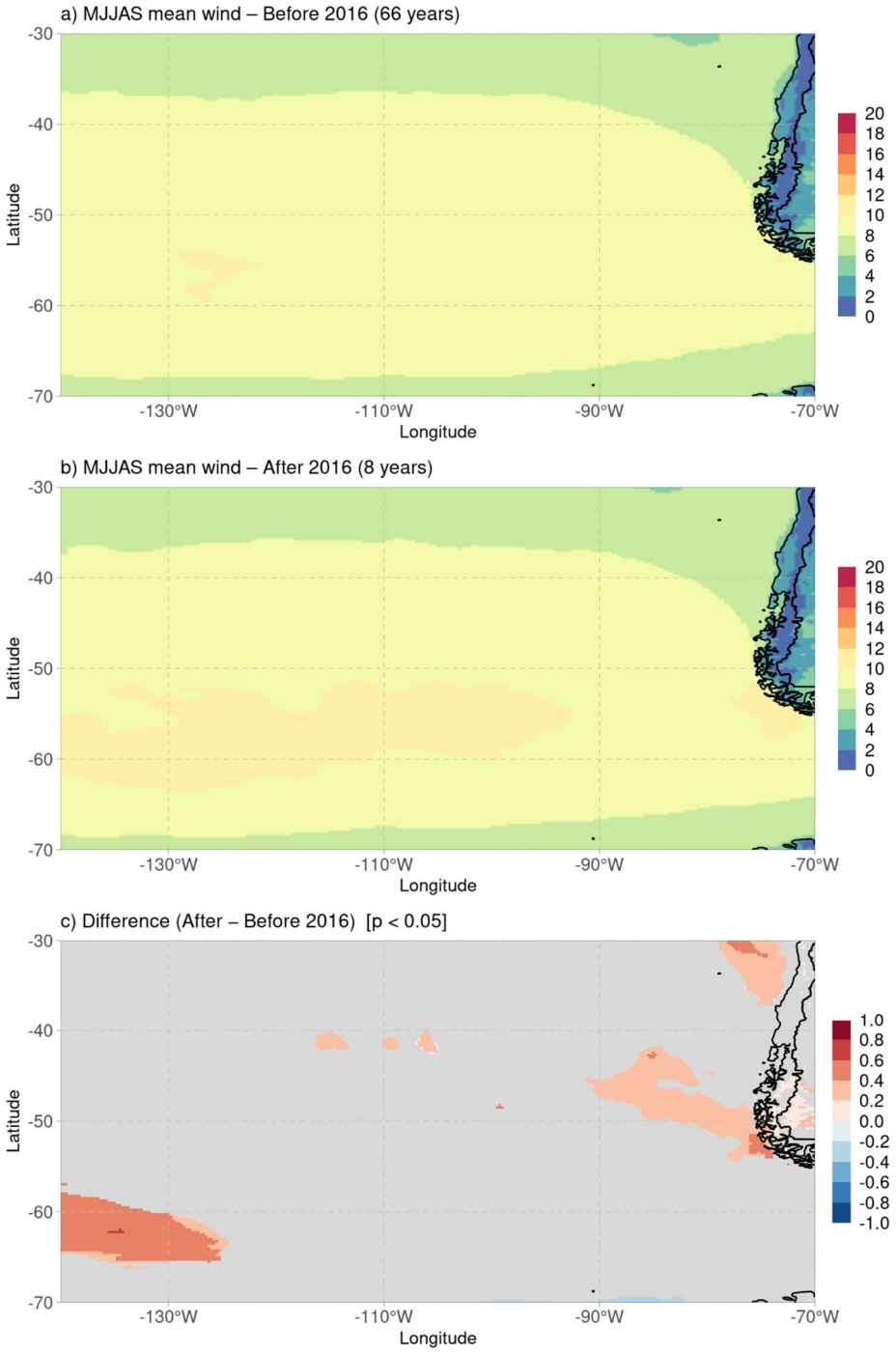

Figure AR1: Composites of seasonal (MJJAS) near-surface wind before and after 2016, together with the difference between the two periods. Only differences significant at the 5% level according to a t-test are shown.

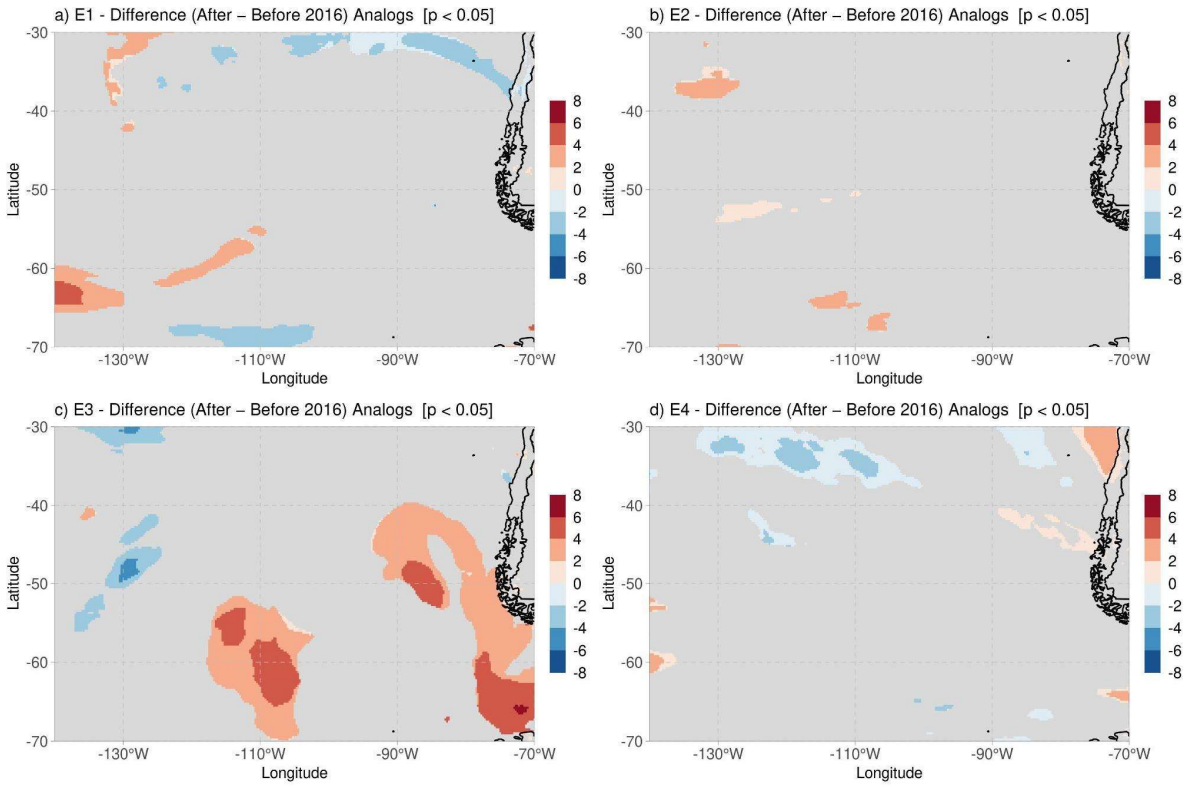

Figure AR2: Mean differences in near surface wind between days analogs in the present period (1986-2025) before and after 2016. Only differences significant at the 5% level according to a t-test are shown.

14. Line 282: The 10m wind anomalies for the Northern Hemisphere composites occur behind the SHWW anomalies, but in the Southern Hemisphere it is the opposite. Why causes this difference?

The discrepancy was caused by an error in the previous version of the figure. For the Southern Hemisphere SHWW composites, the figure mistakenly displayed conditions preceding the event, whereas the Northern Hemisphere composites showed conditions during the event itself. After correcting the Southern Hemisphere figure to use the same event timing as the Northern Hemisphere, the apparent difference in the phase relationship of the 10 m wind anomalies was resolved.

15. Lines 354-355: Looks like maybe the sentence got cut off around “intensifies”

The sentence in Lines 354–355 was indeed incomplete due to an editing issue. We have now corrected and rephrased the text to improve clarity and completeness in the revised manuscript **(L430-436)**.

L432-437: *“Extreme swell events along the Peruvian coast pose recurrent risks to coastal communities, infrastructure, and maritime activities. These events originate far offshore and exhibit strong seasonal variability in their source regions: during the austral winter they are primarily generated in the South Pacific, whereas in summer they typically develop in the western*

North Pacific. This study examines the atmospheric circulation patterns associated with extreme wave events reaching the Peruvian coast from both hemispheres, with particular emphasis on the structure and variability of the upper-level jet. In addition, it assesses the potential influence of climate change on the intensity of these events using an analogue-based methodology."

References:

Priestley, M. D. K., and J. L. Catto. Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure, Weather Clim. Dynam., 3, 337–360. 2022.