LINE-BY-LINE RESPONSE TO TO REFEREE COMMENTS

"A climatological characterization of North Atlantic winter jet streaks and their extremes"

Mona Bukenberger[1] | **Stefan Rüdisühli**[1] | **Lena Fasnacht**[1] | **Sebastian Schemm**[1]

CONTENTS

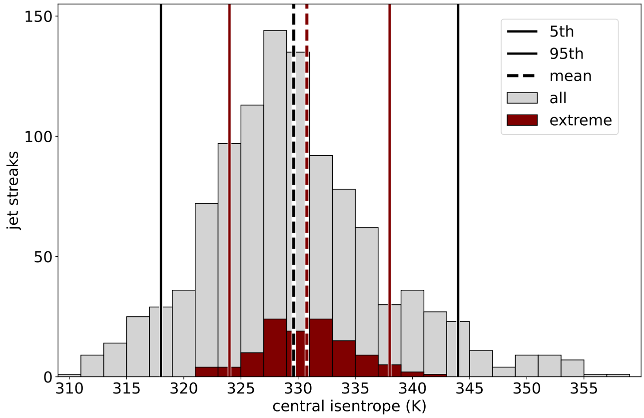

FIGURE 1 Histogram of central isentropes at peak jet streak intensity (θ_{peak}; K) for all jet streaks (grey) and extreme jet streaks, i.e. cases with wind speeds on the central isentrope exceeding 92.5 m s^{-1} on an area of at least 3.62·10^{5} km^{2}) at time of peak jet streak intensity (dark red), and vertical lines indicating the mean (dashed) and the 5th and 95th percentile (solid) of the distributions.

1 | MAJOR COMMENTS

There are some aspects of the methodological approach that might benefit from further clarification:

a. **L103: The height of the tropopause on the equatorward side of the jet can often extend as high as 350K during the wintertime, especially in cases of polar/subtropical jet superposition (Winters et al. 2020). Is there a particular motivation for cutting the analyses off at 340K and are you missing anything by not considering isentropic levels above this level say up to levels of 355K? The secondary peak in Fig. 8 at higher potential temperatures partly motivates my concern.**

We totally agree and thank you for this comment. We have rerun all of our analysis on a dataset that now includes all isentropes between 310 and 360K. The results of this new analysis (Figure 1) are largely identical with those produced on the basis of the smaller isentropic range, with two notable exceptions:

1. Central isentropes at peak jet streak intensities now reach up to 356K (See Figure 1). Interestingly, the number of jet streaks with central isentropes higher than or equal to 340K at peak jet streak intensity remains roughly the same. This indicates that we did not miss most of the subtropical jet events in our initial analysis, but rather linked their centres to a too low isentrope.
2. the relationship between jet streak duration and peak intensity (Figure 2) is even more robust with the new set of isentropes. We also find a few jet streak events less than before, and they last marginally longer, in the new analysis. Our understanding is that this is due to a higher fidelity in jet streak centre identification and merging of subtropical jet streak centres in time into one event when using the new dataset.

The composites of large-scale flow as well as the make-up of jet streak clusters are largely unaffected.

Once again, we thank you for this suggestion, which makes the study even more rigorous while also helping us to show that our algorithm works for both sub-tropical and sub-polar jet streaks.

b. L124: In many cases the flow can be highly amplified and jet streaks can be meridionally oriented rather than predominantly zonal. How does the methodology handle this common occurrence?

If our understanding of your remark is correct, it refers to cases in which the axis of the jet streak is completely aligned in the S-N direction.. We have two main reasons for thinking that our algorithm is robust:

- Through manual inspection, we confirmed that the algorithm is able to detect jet streaks that are far from zonally oriented. In fact, when analysing the distribution of all jet streak orientations around approximately 14% of have an orientation within 5° of a pure S-N orientation (at time of peak jet streak intensity). For those cases, the jet streak center is – as for most time steps – co-located with the highest wind speed, which allows us to find a tropopause-longitude intersection even for strongly tilted jet streaks.
- While it is true that especially for strong jet streaks the flow can be highly amplified, jet streak centres tend to be close to the ridge crest at time of their peak intensity, which is the time of maximum interest. If the wind speed maximum is locaated close to the ridge crest, the shape of the jet streak is zonal rather than meridional and so is its axis.

c. The selection of specific threshold values and percentiles for the analysis could benefit from more justification. For instance, the choice of 92.5 m/s for extreme peak intensities is not much different than the median of all jet streaks (i.e., its within one standard deviation according to Fig. 6). I find the extreme jet streaks category to be a bit more rigorous since there's an area criterion associated with it.

We agree with you that the category of jet streaks with 'extreme peak intensities' as we used it in the first manuscript is confusing, and since we do not use the category in any of the key analysis of the publication, we removed it. We also added a paragraph explaining the choice of the wind speed threshold in the manuscript now and how we combine wind speed and area thresholds to get to the definition of extreme jet streaks to Section 3.1. The new paragraph (Line 309ff in the revised manuscript) reads

For a robust definition of extreme jet streaks, we choose extreme events to combine high wind speed and large areas, also to avoid detection of single grid points exceeding a wind threshold. We define the wind speed threshold based on the wind speed distribution over the North Atlantic. A wind speed exceeding the 99.9th percentile for at least 90% of North Atlantic grid points on all isentropes is $92.5 ms^{-1}$*, which we therefore set as the wind speed threshold. Figure 7b shows the distribution of area covered by wind speeds exceeding* $92.5 ms^{-1}$ *at time of peak intensity on the central isentrope for the 310 jet streaks for which peak intensities exceed this threshold. The 50th percentile of area with wind speed exceeding* $92.5 ms^{-1}$ *for those jet streaks is* $1.91 \cdot 10^5 km^2$*, while the 70th and 90th percentiles are* $3.62 \cdot 10^5 km^2$ *and* $6.46 \cdot 10^5 km^2$*, respectively. We define extreme jet streaks as those events for which the area on which wind speeds exceed* $92.5 ms^{-1}$ *is larger than* $3.62 \cdot 10^5 km^2$ *at time of their peak intensity, a definition yielding 91 extreme events.*

d. L155: It is conceivable to me that the isentropic level corresponding to the maximum wind speed is likely to change throughout a single event. Is anything done to account for this as part of the analysis and is it necessary?

Yes, as you point out correctly, the central isentrope evolves throughout jet streak evolution. Our algorithm tracks the central isentrope at each time step, such that we can also quantify the increase and decrease of isentropic levels throughout jet streak evolution (See Figure 2). We find that between the time step of peak jet streak intensification and intensity, the central isentrope changes less than 5 K for 80 % of all jet streaks.

Prior to the preparation of the manuscript, we tested both approaches, using either a varying central isentropes or a fixed isentropes (from the time of the peak of the jet streak intensity). We found that the results are rather similar and for simplicity included only the analysis using a fixed isentrope.

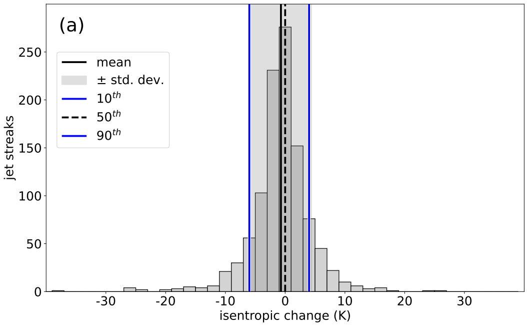

FIGURE 2 Changes in the instantaneous central jet streak isentrope between the time of peak jet streak intensification and peak jet streak intensity for the 1065 North Atlantic winter jet streaks. Histogramm in grey bars, vertical lines show the (black solid) mean, (black dashed) median, and (blue) 10th and 90th percentile (-4 and +6K, respectively) of the distribution.

e. L195: Similar to one of my earlier comments, is any methodology performed to account for the different orientations of jets when compositing (i.e., a rotation of fields so that they are aligned with the jet axis)?

We agree that this would be an additional option in particular for the second part of our study. We do not align the jet axis when compositing, but find that the clustering method does this sufficiently.

The composites as shown in Figure 16–18 are based on statistical clustering, which appears to align all jet axis within individual clusters sufficiently. This can be deduced from the figures on the PV gradient standard deviation (Figure G1), though some in-cluster variability remains.

The results sections show interesting differences and statistics regarding the characteristics of jet streaks and extreme jet streaks, but it might benefit from the application of statistical testing to verify to what extent the differences are statistically significant.

Thank you for your suggestions on statistical testing. We now applied bootstrapping analysis to all important characteristics and notified where differences are robustly shown in those analysis in the results section. We explain our bootstrapping method in the methods section.

indeed uniquely different? A similar practice could be applied to the content of Fig. 12 or Table 1.] We adopted this suggestion as proposed and noted where results are robust with respect to the subsample drawn in bootstrapping analysis. For former Figure 12 (now Figure 13), this lead to the following new sentence in Section 3.3 (Line 441ff): *A bootstrap analysis using 1000 resamples of all as well as extreme events (dots and vertical lines in Fig. 13) shows that the increased prevalence of the M regime is a robust result of the analysis. While the same is true for the under-representation of N regime, the climatological likelihood of the S regime is not altered significantly compared to climatology for both all and extreme jet streaks.*

b. In any of the composite analyses, a statistical bootstrap test could be performed to determine to what extent the characteristics of the near-jet environment during extreme jet streaks are significantly different compared to the composites of all jet streaks.

Like above, we adopted this suggestion. We conducted bootstrap analysis for all composites, but only discussed the results explicitly where they were noteworthy, in particular for the PV gradient and precipitation composites as well as the WCB outflow frequencies. This lead to multiple changes in the subsection **Zonally and anticyclonically oriented jet streak clusters** (Line 515ff). The changes can be followed in the LaTeX- diff document from line 569 onwards. The most important changes/ additions are (Line numbers in the revised manuscript:)

- Line 516f, addition: *Higher PV gradients at the jet streak centre of extreme cases are a robust feature in bootstrapping analysis (see stippling in Fig. G5 a vs. b, c vs. d).*
- old: *The difference in the composite PV gradient between extreme and non-extreme jet streaks is larger at the time of peak intensity than intensification. We hypothesize that the weaker correlation between the wind field and the PV gradient at peak intensification is due to geostrophic imbalance of 595 the rapidly accelerating flow, such that the assumptions required to directly link wind speed to PV gradients (Section 2.6) are violated..*

  new, line 531ff: *Another robust difference between extreme and non-extreme jet streaks is the more narrow region of enhanced PV gradients close to the jet streak axis for extreme cases (Fig. 16 first vs. second and third vs. fourth row).*
- Line 552ff, addition: *The maximum in cold-frontal precipitation is located below the right jet entrance and is robustly more intense for extreme jet streaks (compare the first vs. second column in Fig. 17 and see stippling in Fig. G2 a–d).*
- Line 561f, addition: *Robust differences between the WCB outflow of extreme and non-extreme C1 jet streaks are already given at time of peak jet streak intensification, but increase as they reach peak intensity (Fig. G6 a–d).*
- old: *In the warm sector of the mature cyclone, the precipitation rate for extreme jet streaks is approximately double that of non-extremes*

  new (Line 565ff): *In the warm sector of this cyclone, the precipitation rate for extreme jet streaks is approximately double that of non-extremes and this result is robust against resampling (Fig. G2 e–h).*
- old: *It grows and produces more precipitation as jet streaks reach peak intensity, again with more precipitation in extreme jet streak composites.*

  new (Line 572ff): *Precipitation increases as jet streaks reach peak intensity, again with robustly more precipitation in extreme jet streak composites.*
- old: *Notably, WCB frequencies in this region almost doubles between the times of peak intensification and intensity for extreme jet streaks , but barely changes for non-extreme jet streaks (Fig. 18e–h).*

  new (Line 575ff): *Similar to the anticyclonic cluster, WCB outflow on the tropospheric side of the jet is more frequent for extreme zonally oriented jet streaks (Fig. 18a–d vs. e–h), but the difference only becomes statistically robust at time of peak jet streak intensity (Stippling in Fig. G6 e–h). By this time, WCB frequencies in this region more than doubles for extreme jet streaks (Fig. 18e–h).*

In some of the composites the raw fields are used rather than anomalies (i.e., PV, sea- level pressure). Would using anomalies potentially be more effective given that the climatology of these variables can vary substantially throughout the cool season and potentially bias the analysis, especially if the jet streaks are also located in different parts of the Atlantic Basin? In addition, I might have missed it, but are the composited fields weighted at all to account for smaller distances between grid points at high latitudes compared to lower latitudes? If not, treating the input maps equally will unfairly weight the composites towards jet streaks that are located at higher latitudes.

We see the advantages and disadvantages of both methods; based on our previous research, anomalies often highlight similar regions as the full flow field, but in certain cases standardised anomalies can also de-emphasize regions that are dynamically relevant. To the first part of the comment: Concerning upper-level fields, after giving it considerable thoughts, we believe that centering on the tropopause and taking the full fields is similar effective in depicting the large-scale flow situations. Concerning sea-level pressure, the anomalies indeed show an even clearer picture of the low-level weather systems. We therefore changed the SLP- to SLP-anomaly contours in the respective figures and changed the paragraphs on low-level circulation accordingly.

To the second part of the comment: Yes, you are right, and we do use area-preserving remapping before conducting composite analysis. Fields at higher latitudes are therefore weighted such that they are not overweighted in the composites. We reformulated the description of the remapping method to clarify this point. Thank you for pointing this out. This led to the following changes:

- Sect.2.2.2, old: *After remapping the jet streak center to (0.0E, 0.0N)(Sect. 2.3:), ...*
  new (Line 209f): *After remapping a jet streak's centre to (0.0E, 0.0N) with an area-preserving coordinate transformation (Sect. 2.3), ...*
- Sect. 2.3, old: *To obtain jet streak-centred composite, the input coordinate system is rotated to position the jet streak centre at the centre for each event and time step of interest using CDO's remapping function for regular grids (Schulzweida, 2023).*,
  new (line 244f): *To obtain jet streak-centred composite, the input coordinate system is rotated to position the jet streak centre at the centre for each event and time step of interest using CDO's area-preserving remapping function for regular grids (Schulzweida, 2023).*

2 | MINOR COMMENTS – ANDREW WINTERS

2.1 | Abstract

L3–4: Consider specifying that this upper-level divergence pattern is only specific to Northern Hemisphere jet streaks, since the pattern reverses in the southern hemisphere. If looking for a more unified phrasing you could refer to the regions as "equatorward entrance region" or "poleward exit region".

Thank you for pointing this out. We modified the text according to your suggestion, such that it now reads (Line 3ff):

Upper level divergence in their equatorward entrance and poleward exit regions couples them to surface weather via vertical motion and are regions prone to precipitation formation, which feeds back on 5 the strength of upper level divergence and wind speed via diabatic heat release.

We also used this phrasing in the introductory paragraph on the 4Q-model now, to be more general.

Consider using a different term rather than "deepening" when referring to the intensification rate of jet streaks, as this is more standard when referring to cyclogenesis rather than jets. Perhaps acceleration could be a suitable alternative term?

Good point again, we modified the sentence accordingly (Line 11–13):

The peak intensity of jet streaks also increases with their lifetime and extreme jet streaks exhibit a prolonged intensification period rather than increased acceleration rates.

2.2 | Introduction

L34: The Harnik et al. (2014) study largely considers merging of the two jets from a seasonal perspective, but this also can occur on synoptic time scales and lead to some of the extreme winds observed in this study (e.g., Winters et al. 2020). It might be worth highlighting this environment as part of the introduction discussion on jet streaks, as well.

Thank you for pointing this out. We added the following lines (line 35ff) to mention merged jet events to account for this:

A merging of the two jets is an exception, although it has been observed in the past, on seasonal (Harnik et al., 2014a, e.g.) timescales. Studies investigating subtropical-polar jet superposition (e.g. Winters and Martin, 2014; Winters et al., 2020) on synoptic timescales showed that such events can be associated with extreme wind speeds and heavy precipitation.

L69: Should the cyclonically curved case favor ascent beneath the left-exit region? I believe the effects of flow curvature and speed changes should theoretically cancel in the right-exit region.

Thanks for pointing this out. This was a typo on our part and is corrected now. We also switched to a formulation that is valid on both hemispheres, such that the sentence now reads (line 73ff):

In an anticyclonically curved jet streak, quasi-geostrophic theory suggests that only lifting in the equatorward jet entrance and sinking in the equatorward exit prevail, while cyclonically curved jet streaks show only sinking in the poleward entrance and lifting in the poleward exit (Cunningham and Keyser, 2004; Clark et al., 2009).

2.3 | Methods and Data

L127: Could you expand a bit more as to why this percentile threshold is chosen? Was it determined empirically, and how sensitive are the results to this chosen threshold?

The threshold was determined empirically with the goal of creating a mask of very high wind speed that

- does not consist of multiple patches that are far away from each other, at least for most time steps (this give a lower bound to the percentile threshold because too small thresholds create multiple patches),

- is one connected area containing the location of maximum wind speed for most cases, and
- is large enough to ensure that, in case of a highly variable wind pattern close to the location of maximum wind speed, the jet streak center is not highly sensitive to the exact location of highest wind speed

The two latter points give an upper bound to the percentile threshold, since a very high threshold would either lead to multiple disconnect patches close to the jet centre, or the resulting mask would be very sensitive to noisy wind patterns close to the jet streak centre. We now give a bit of this reasoning in the main manuscript (Line 148ff):

The percentile threshold of 99.25% was determined empirically with the goal of creating a mask of very high wind speed that consists of a single connected patch containing the location of maximum wind speed for most time steps. Additionally, the mask should be large enough to ensure that the position of the jet streak centre is robust toward small-scale wind speed variations close to the maximum wind speed.

L134: This wind speed threshold is very defensible, but it might help to offer some citations to other studies that have used comparable thresholds for the jet.

Thanks for that hint, we included some references to better motivate the wind speed threshold (line 145ff):

This ensures that the jet streak centre is embedded in the jet stream, and our threshold is in the range of thresholds typically used to define in-jet wind speeds, as in Hartmann (2007); Eichelberger and Hartmann (2007) and Messori et al. (2021) for zonally averaged zonal wind speeds and Winters (2021) and Simmons (2022) for typical wind speeds in the North Atlantic jet stream.

L223: I'm a bit confused by the terminology, "wind speed curvature". In particular, are you referring to the curvature of the flow that would be associated with troughs and ridges, or more so describing the gradient in wind speeds present within a jet. If the latter, I might recommend rephrasing to avoid any confusion in interpretation by a reader.

It refers to the latter. Thanks for pointing this out. We now refer to this as either simply $\triangle U$, 'horizontal variations in relative vorticity', or 'horizontal Laplacian of wind speed' throughout the text, hoping that all of these are less prone to misunderstanding. The respective changes are easiest to follow in the LaTeXdiff, in lines 278, 287, 293, and 304f.

2.4 | Results

L278–280: Could more detail be provided as to why this wind speed threshold is chosen for an extreme jet streak? Why not a much larger wind speed given that this speed seems rather close to the mean of all jet streaks?

See response to your major comment on the wind threshold (Major comment 1c.).

L282–286: These trends certainly do align with the results of Shaw and Miyawaki (2023), but how much of this result can be attributed to more observations over the North Atlantic in more recent decades compared to earlier in the dataset? Discussing potential uncertainties in this result, or evaluating the significance of trends, may benefit the text.

The trends we find are consistent with Simmons (2022) findings on upward trends in wind maxima. Simmons used the background forecast for ERA5, which is less influenced by the increasing ammount of data available to the reanalysis and also finds a slight upward trend that is dominated by interannual and interdecadal variability and not statistically significant. This makes us believe that data availability might of course influence the results, but marginally rather than qualitatively. We now mention Simmons' results in this paragraph and also point out that using the 20-CR dataset and additionally using other reanalysis such as JRA and Merra-2 would be beneficial. We also clarify highlight that, so far, interdecadal variability dominates over the linear trend found in our fit (Line 335ff):

These results 335 are well in line with Simmons (2022), who also found a slight upward trend in monthly maximum wind speed (0.067 ±0.048ms^{-1} *per year) over North America and the Atlantic (see their Figure 16a). To evaluate the statistical significance and methodolical sensitivity of those trends, it would be beneficial to compare different reanalysis datasets (e.g. JRA-55 (Kobayashi et al., 2015) and NASA MERRA-2 (Gelaro et al., 2017)) and also include a dataset using a fixed* *observational basis, for example the NCEP-20C (Compo et al., 2011) reanalysis.*

L295: The bimodal distribution in the central isentrope potentially motivates extending the search range for the core isentrope to higher isentropic levels that cover the entire troposphere.

Indeed. In accordance with your suggestion in your major comment, we updated this analysis and now end up with a new distribution (see Figure 1).

L302: This trend also might relate to the strongest jet streaks being associated with superpositions of the polar and subtropical jet streams, which will feature characteristics of both polar and subtropical jets and strong wind speeds commensurate with those associated with extreme jet streaks.

Yes, good point. We added a sentence to this paragraph to mention this (Line 352ff):

The bulk of extreme jet streakscentred around 330K suggests that some of them are associated with superpositions of the polar and subtropical jet, a result in line with previous research on merged jet regimes (Harnik et al., 2014b; Winters et al., 2020).

L313: This is true for all subsets except for the jet streaks with peak intensities exceeding 105 m/s, if I am reading Fig. 9a correctly. Consider a revision to the text accordingly.

We now point out the subsets of all jet streaks whose mean lifetime is outside the interquartile range of lifetime of the set of all jet streaks, such that the sentence now reads (line 363ff).

Despite a discernible trend toward longer lifetimes for more intense jet streaks, the median and mean lifetimes remain within the interquartile range of all jet streaks (30 to 72 h), for all but the weakest (peak 365 intensities below 68ms^{-1}) and strongest (peak intensities exceeding 108ms^{-1}) events.

L386–390: These are interesting statistics, but would it be possible to perform some type of significance testing, such as a bootstrap test, to evaluate the extent to which these differences in frequencies during jet streak periods are indeed significantly different than climatology?

We performed some bootstrapping to see whether the Frame regime occurrences during peak jet streak intensity differ significantly from the climatology and modified the Figure and text accordingly (line 441ff.):

A bootstrap analysis using 1000 resamples of all as well as extreme events (dots and vertical lines in Fig. 13) shows that the increased prevalence of the M regime is a robust result of the analysis. While the same is true for the under-representation of N regime, the climatological likelihood of the S regime is not altered significantly compared to climatology for both all and extreme jet streaks.

L395: Prior to discussing the transitions, it might be worth emphasizing initially that the predominant observation is that jets tend to persist in their genesis regime before discussing the transitions which are secondary in their frequency.

Thank you for pointing this out. We added the following (line 446ff):

Persistence is a key feature in both extreme and non-extreme jet streak evolution, meaning that the eddy-driven jet typically remains within the Frame regime of jet streak genesis (Fig. 14a, b). For jet streaks during whose evolution different Frame regimes occur, the following results are worth noting: [Text on transitions from before.]

L428: I am having a bit of difficulty locating the figures associated with this discussion in Section E, but they do seem associated with appendix F. References to the pertinent appendix figures to support the discussion in L428–460 would be helpful, as well, to guide the reader.

You are right, of course. Thanks for spotting this and sorry for the confusion in the first read. We corrected the reference and referenced the relevant Figures as you suggested in the paragraph that discusses the general features of all six jet streak clusters. (see section 3.4.1, lines 472ff in the new manuscript.)

L433: It might be worth emphasizing here, or elsewhere in the manuscript, that this result is not particularly surprising from a theoretical standpoint, given that supergeostrophic flow is expected at the apex of upper-tropospheric ridges from consideration of gradient wind balance.

Thank you for this valuable comment. We deemed this thought to be best placed in the conclusion section and therefore placed the following sentence in item 1. of 4.2 (Line 617ff):

This result is in line with theory, because the combination of pressure gradient and centrifugal forces balance the Coriolis force to yield a supergeostrophic wind in a ridge.

L501–503: I found this text rather repetitive with that at the end of the previous paragraph. Consider whether this small paragraph might be deleted without any loss of content.

That is a good point toward shortening the manuscript. We removed the paragraph according to your suggestion and merged the bit of information we found worth retaining into the previous paragraph, which now reads (line 550ff):

The cyclone is associated with an elongated SW-to- NE oriented band of enhanced precipitation along its cold front whose intensity changes only marginally between the times of peak intensification and intensity (Fig. 17a–d). The maximum in cold-frontal precipitation is located below the right jet entrance and is robustly more intense for extreme jet streaks (compare the first vs. second column in Fig. 17 and see stippling in Fig. G2 a–d).

L515–517: I am not so sure I agree with this conclusion, as the surface cyclone is much stronger beneath the left-exit region in the extreme subset compared to the non-extreme subset. I would presume the strong pressure gradient in the extreme subset stems from the greater intensity of both the cyclone and anticyclone relative to the non-extreme, zonally-oriented cases.

We revised the entire paragraph after switching from SLP to SLP-anomalies. The change is:

Old: *As is the case for anticyclonically oriented jet streaks, the difference in the meridional pressure gradient between extreme and non-extreme jet streaks primarily stems from the intensity of the anticyclone:(Fig. G2 :::: e–h)*

New (line 567ff): *The pressure difference within the cyclone-anticyclone pair is roughly 45 hPa during peak jet streak intensity compared to 20 hPa for non-extreme cases and is driven by both stronger cyclones and anticyclones for extreme cases (Fig. G2 e–h).*

2.5 | Appendix

L633: The reference to “westward” in this sentence appears to be incomplete and might need another word or two to complete the sentence.

We agree and revised the statement(line 694f):

For S-regime jet streaks, the composite jet at upper levels exhibits a westward displacement with respect to the jet at lower levels.

L635–640: This result might also highlight the M-regimes as featuring a greater likelihood of polar and subtropical jet stream superpositions, which would also align with the greater likelihood of extreme jet streaks in this category.

That is true. We included this idea in the new manuscript by slightly rephrasing the paragraph to now be (line 698ff):

This result points toward an increased interaction between upper and lower levels for jet streaks that peak in the M regime and might point toward an increased likelihood of merged jet states in the M regime. It is also consistent with the finding that most extreme jet streaks peak in the M regime and show enhanced lower-level-to-upper-level coupling.

2.6 | Figures and Tables

Table 2: Is there a particular reason why the cluster numbering starts at 0 rather than 1? Also, might it be useful to use more descriptive names for each cluster that are tied to their respective characteristics rather than numbers?

Not in particular. We changed the numbering to start at 1 to avoid any notion that Cluster ’0’ might be special in a particular way.

On the point regarding more descriptive names: We thought about doing this for all clusters but ended up with only naming the two clusters in descriptive ways that we discuss in detail. Some clusters look qualitatively similar, such that we found it most instructive to only pick the two most important and distinct clusters for descriptive naming. Note also that the other clusters are dropped from the discussion after the first paragraph, such naming them might cause anticipation that would not pay off.

Fig. 1: In panel (a), it is a bit confusing that there is one red arrow that points away from the diagram at the lowest isentropic level. What does this red arrow correspond to and how is it different from the one at the level of maximum wind speed?

This is simply the 'maximum wind speed' axis, but we see how the colour might be confusing. We removed that panel slightly from the isentropic surfaces to avoid confusion about the axis pointing 'away' from anything and transformed the colour of the axis to be black. We hope this makes the schematic more clear and are happy to incorporate more suggestions, should you have further ideas to improve it.

Figure 2: The solid black and dashed black lines that correspond to the mean and median, respectively, are a bit difficult to see against the grey histogram bars. Could different colors be used for these quantities? Similar considerations may also apply to other figures.

Thank you for this suggestion. We tried to use different colours/ linestiles in Figures for which this seemed helpful. This resulted in changes to Figures 2, 6,8, and 9. We hope it makes all the Figures easy to access.

Fig. 9: Some of the dashed lines and solid lines are a bit difficult to see in panel (b). Could they be made a bit thicker? Could the hours associated with lifetimes also be included along the x-axis in panel (a). It is a bit difficult to compare the box and whisker plot values against values along the y-axis on panel (b) – far distance for the eye to travel.

Good points, we made the vertical lines thicker and removed the lines indicating the Frame-regime related values to make the Figure more simple.

Fig. 11: I understand why the authors ordered the panels the way they have, but it seems a bit counter-intuitive that the first timestep of the evolution corresponds to panel (c) for all jet streaks rather than panel (a). Could the panel labelling conventions be changed to match the temporal evolution of the jet streak evolution? The panel labels are also incorrect for the extreme jet streaks and should be (e-h). I noticed similar errors in a few other figures, as well.

We follow the suggestion to harmonize temporal evolution and the progression of labels. We kept the ordering of the panels as is, though, to keep peak intensification and intensity side by side and at the top. We also corrected the labelling for extreme jet streaks and harmonized labelling in other figures. We hope to have caught all cases of erroneous labelling and thank you for looking at this so carefully.

Fig. 12: I might recommend a revision to the caption to not refer to the presence of a low-level jet, which can have a much different definition than the jet features considered in this study. Namely, low-level jets often correspond to isolated low-level wind speed maxima that decay in intensity by some amount with altitude. I don't believe that is the case for the jets considered here. Instead, I might just recommend referring to times in which the jet resides in the S, M, and N regimes.

Thank you for this point. We removed the mention of 'low-level' from the caption to avoid this connotation.

Fig. 13: I like this figure a lot, but it is a bit counterintuitive to me that the N regime is on the bottom and S regime is on the top. Could these regimes be reversed in their position on the plot so that the display of these jet regimes matches their characteristic location on a map?

Yes, good suggestion. We revised the Figure accordingly.

Fig. 16: Might it be possible to include a contour that corresponds to the position of 2-PVU? That way the position of precipitation and surface sea-level pressure anomalies will be easier to link to the upper-level jet structure.

Again, good point. The revised manuscript includes this suggestion for all figures showing low-level variables, including

Figures in the Appendix and those showing standard deviation of precipitation.

Fig. E1: Should the time of maximum intensification in this four-panel plot correspond to panel (b) rather than (a), since (a) is described as the start time later on in the caption?

Correct, we changed this.

Fig. F1: It appears the panel labels might be incorrect within this figure, since (a-d) is duplicated for both sets of 4 panels.

Correct, we changed this. (See comment above)

Appendix F Composite Figures: For each group of four panels that correspond to a cluster, could a label be added to more clearly identify which group of panels corresponds to a cluster (i.e., similar to what is done at the top of Fig. 15)

Yes, we revised all Figures in Appendix F accordingly.

3 | MINOR COMMENTS – R2

3.1 | Introduction

Line 23 – I would suggest: "The jet stream is a band of enhanced westerly winds in the mid- and upper troposphere found in both hemispheres."

Good point, we adopted that exactly as suggested (Line 24 in the new manuscript).

Line 24 - I would suggest adding here one of encompassing textbooks on atmospheric circulation for the benefit of wider audience, e.g., "influences daily-to-weekly weather patterns with its meanderings (Randall, 2015)."

This was adopted as suggested (line 25 in the new manuscript).

Line 40 – I would suggest adding here textbook reference: "Frame et al., 2011, Wilks, 2020). The response ..."

This was adopted as suggested (line 43 in the new manuscript).

Line 59 – I would suggest adding here a figure/plot depicting essential 4Q-model schematic for the benefit of wider audience.

Thank you for this thoughtful suggestion. We adopted schematics from the seminal works by Uccellini and Kocin (1987) and Beebe and Bates (1955) to create a simple schematic of all the relevant flow features of a straight jet streak. The Figure resulting from this and the respective caption are as depicted in Figure 3. We also referenced the figure in the introductory text to make it more easy to follow. This lead to changes in lines 63ff, which now read:

Bates (1955). In the 4Q-model, the flow is decomposed into geostrophic and ageostrophic wind components. In case of a straight jet streak (Fig. 1), the acceleration of air parcels in the jet streak entrance imply poleward ageostrophic wind and horizontal divergence/convergence in the equatorward/poleward entrance of the jet streak (Beebe and Bates, 1955; Cunningham and Keyser, 2000, 2004) (Fig. 1a). This causes rising/sinking motion below the equatorward/poleward entrance quadrants of the jet streak and vice versa for the jet exit, where equatorward ageostrophic wind prevails at jet level (Fig. 1 b for the jet entrance and Fig. 1 c for the jet streak exit).

Line 74 – I would suggest adding here textbook reference: "The PV Perspective (Hoskins and James, 2014) has been employed to study ..."

This was adopted as suggested. The change ammounts to:

Old: *The PV perspective has been employed to study the dynamics of extratropical cyclones and accompanying jet streaks ...*

New (line 80f): *The PV perspective (Hoskins and James, 2014) has been employed to study the dynamics of extratropical cyclones and accompanying jet streaks ...*

Line 101 – Please justify why are you not considering data from available pre-1979 period? Additional point, could 3-hourly data provide more temporarily resolved insight?

On the first question: Good point. We thought this question was also relevant for the broader audience and therefore included

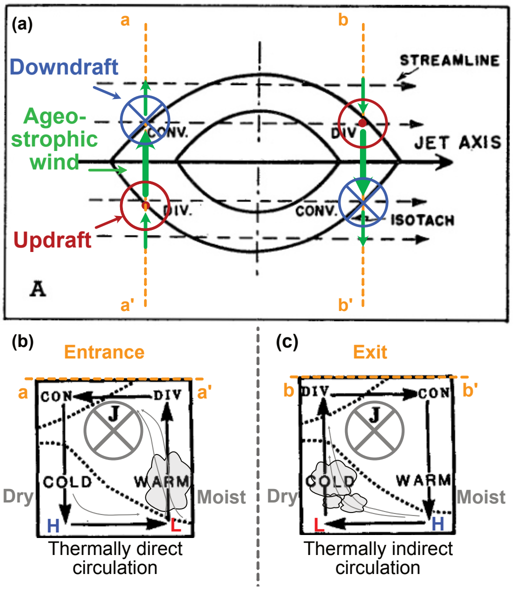

FIGURE 3 Schematics adapted from (a) Beebe and Bates (1955), their Figure 4, and (b,c) Uccellini and Kocin (1987), their Figure 3B. Panel (a) shows an idealised straight jet streak with the associated upper-level convergence and divergence and induced updraft (red dotted circles) and downdraft (blue circles with cross). Green arrows show the direction of ageostrophic wind for a straight jet streak. Orange dashed lines show the cross-sections whose transverse circulation is depicted in panels (b) and (c). In (b) and (c), black arrows show ageostrophic transverse motion around the jet. Grey areas indicate clouds, where transverse motion can induce condensation, the intensification or genesis of cyclones, and convective processes. Thin grey arrows show exemplary streamlines in such a transverse motion, and the blue **H** and red **L** indicate where transverse motion can support the formation of surface high- and low-pressure systems, respectively.

the answer in the first paragraph of the Methods and Data section. It now reads (lines 108ff):

This study uses 6-hourly ERA5 data for the winter (DJF) period in the Northern Hemisphere between 1979 and 2023 interpolated in the horizontal on a 0.5 latitude-longitude grid (Hersbach and Bell, 2020; Hersbach et al., 2023). In the vertical, the data is interpolated onto 26 isentropic levels between 310 and 360K in steps of 2K. Global satellite coverage began contributing to ERA5 from 1979 onward. This addition improved reanalysis quality and made upper-level winds and extreme speeds more reliable, especially over oceans like the North Atlantic. Hence, using data from 1979 onward ensures a high-quality and consistent basis for our analysis.

Regarding the second question: It is true that an increased temporal resolution would bring more precision to the analysis. In particular, a three-hour time step allows for a more precise analysis of the intensification of jet streaks and thereby the time of peak intensification. However, when testing the our algorithm on hourly resolved data for one winter, we found little sensitivity regarding the results concerning the lifetime and peak intensities. The times of peak intensification or intensity change only by up to three hours, such that the large-scale flow patterns at jet level, which is what we are most interested in for this study, are also robust with respect to the time step. We therefore decided to keep the analysis on six hourly time steps to be more resource efficient regarding compute time and memory. We hope this sufficiently addresses your comment.

3.2 | Methods

Line 165 – Determining K in K-means clustering approach is one of the most sensitive aspects of such clustering analysis. Hence, please justify how did you set effective/optimal K=3 in your K-means clustering analysis (why not K=2 or K=4 or ...).

Thanks for this comment. You are right that choosing K is a sensitive topic and the number of jet regimes over the North Atlantic is a debated issue. In our approach, we follow (Woollings et al., 2010), who found three preferred meridional locations for the North Atlantic jet (Woollings et al., 2010, their Figure 1.). Due to these preferred locations, choosing K = 3 for a k-means clustering of North Atlantic jet streams is motivated from a statistical point of view (see Section 2. and 3. in Woollings et al., 2010, for more details on this). This result has motivated a stream of work that considers three important states of the North Atlantic winter jet stream and the transitions between them (For example Frame et al., 2011; Ambaum and Novak, 2014). From a dynamics perspective, the central (Woollings et al., 2010; Ambaum and Novak, 2014) or M regime (how it is referred to in Frame et al., 2011) is often considered to be the background state, and the southern and northern regimes are deviations from this state induced by enhanced cyclonic and anticyclonic wavebreaking respectively. The transitions between the states are discussed as the oscillator model of the North Atlantic jet stream in Ambaum and Novak (2014). This thinking adds a physical motivation to the statistical motivation for choosing three clusters. In our approach, we follow this line of research and our clustering method is strongly inspired by that implemented in Frame et al. (2011).

We agree that the choice of K = 3 deserves some justification in the main text. We therefore modified the first paragraph of Section 2.2.1 to now read as follows (line 185ff):

To connect jet streak life cycle characteristics with the state of the eddy-driven jet stream, we use a jet stream regime definition similar to that introduced by Frame et al. (2011). The regime definition relies on zonally averaged but meridionally varying jet profiles, denoted as $U(t, \lambda)$, which are computed by zonally and vertically averaging the zonal wind between 60° W and 0° W, and between the 700 hPa and 900 hPa. The North Atlantic winter jet stream is known to have three preferred meridional positions (Woollings et al., 2010, their Figure 1.). Since this discovery, the three jet regimes and transitions between them are discussed in terms of an oscillator model of the North Atlantic jet stream (Frame et al., 2011; Ambaum and Novak, 2014), adding physical meaning to the statistical prevalence of these positions. Therefore, we apply K-means clustering (Jain, 2010) with three degrees of freedom to the jet profiles $U(t, \lambda)$.

We hope this addresses your point sufficiently.

Line 172 - When the neighbourhood radius in SOM is set to 0, the SOM reduces to the K-means clustering (i.e., SOM can be perceived as a constrained version of K-means clustering, Hastie, et al., 2009). Please elaborate your choice of using SOM here instead of also K-means clustering (that is also unsupervised clustering technique).

Thank you for this comment. As you point out correctly, SOM reduces to K-means when the neighbourhood-radius is zero. We chose SOM over K-means for two main reasons before further reducing the number of clusters with agglomerative clustering:

- **Flexible number of clusters**: K-means requires a user to set a fixed number of clusters and we wanted more flexibility to start with. SOM allowed us to start with a high-dimensional map with small topographical and quantitative errors that captures a broad range of clusters. This map was a good basis for further simplification and reduction of the number of clusters using agglomerative clustering. We also tested K-means with varying numbers of initial clusters but did not identify an optimal number (w.r.t. the typical metrics such as the quantization error) of clusters in this approach.
- **Interpretability through distance-preserving property of SOM**: With SOM, the map structure (where $\sigma(t) > 0$) organizes similar clusters close to each other, helping visualize clusters in 2D space. This spatial distribution highlighted regions of increased density in the input data space, providing a natural basis for estimating an optimal cluster count. It revealed local event number maxima, suggesting that a final cluster count between 4 and 7 would be most interpretable.

We've revised Section 2.2.2 to clarify these points, emphasizing the advantages of SOM clustering in our case while acknowledging its limitations and similarity to K-means. We hope this response provides a clearer rationale for our approach. We refer to the LaTeX- diff document, lines 203–242 to follow the changes in detail.

3.3 | Results and conclusion

Line 550 – I would suggest pointing out that the optimal number of clusters can be substantially dependent on the selected validity index or set of applied validity indices. This should deserve follow-up study (or studies) that could also involve some other state-of-the-art atmospheric reanalysis products such as JMA JRA-3Q and NASA MERRA-2.

Thank you for pointing us toward this again. It is true that our results and the ideal number of clusters may well depend on the method of jet identification, the choice of validity indices for K-means clustering, and possibly other methodological choices. In this work, we follow Woollings et al. (2010); Frame et al. (2011) in their approach, but we acknowledge that different classifications of jet regimes could be useful, as well as using different reanalysis datasets. We therefore incorporated your suggestion into the conclusion and added the following paragraph to our conclusions in Section 4.1: Characteristic properties of extreme North Atlantic winter jet streaks.

The amended paragraph reads as follows (line 602–608 in the revised manuscript):

With Frame et al. (2011), we follow a line of research that finds three jet regimes based on vertically and zonally averaged wind profiles using K-means clustering. The ideal number of regimes could change with the reanalysis dataset at hand, other methods for jet detection and profile computation, as well as the validity indices chosen to evaluate the K-means clustering. Follow-up studies involving other state-of-the-art atmospheric reanalysis products such as JMA JRA-55 and NASA MERRA-2 (Kobayashi et al., 2015; Gelaro et al., 2017) and different methods to identify low-level jet regimes would help solidify our understanding of upper and lower-level interaction in jet streak evolution.

REFERENCES

Ambaum, M. H. P. and Novak, L. (2014) A nonlinear oscillator describing storm track variability. *Q. J. Roy. Meteor. Soc.*, **140**, 2680–2684.

Beebe, R. G. and Bates, F. C. (1955) A mechanism for assisting in the release of convective instability. *Mon. Weather. Rev.*, **83**, 1–10.

Frame, T. H. A., Ambaum, M. H. P., Gray, S. L. and Methven, J. (2011) Ensemble prediction of transitions of the north atlantic eddy-driven jet. *Q. J. Roy. Meteor. Soc.*, **137**, 1288–1297.

Jain, A. K. (2010) Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, **31**, 651–666.

Simmons, A. J. (2022) Trends in the tropospheric general circulation from 1979 to 2022. *Weather and Clim. Dynam.*, **3**, 777–809.

Uccellini, L. W. and Kocin, P. J. (1987) The interaction of jet streak circulations during heavy snow events along the east coast of the united states. *Weather Forecast.*, **2**, 289–308.

Woollings, T., Hannachi, A. and Hoskins, B. (2010) Variability of the north atlantic eddy-driven jet stream. *Q. J. Roy. Meteor. Soc.*, **136**, 856–868.