REPLY TO REFEREE 2

“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

1 | REFEREE 2 — MINOR COMMENTS

1.1 | 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 - 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 40 – I would suggest adding here textbook reference: "Frame et al., 2011, Wilks, 2020). The response ..."

This was adopted as suggested.

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 1. We also referenced the figure in the introductory text to make it more easy to follow.

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.

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 the answer in the first paragraph of the Methods and Data section. It now reads:

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 360 K in steps of 2 K. 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.

1.2 | 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

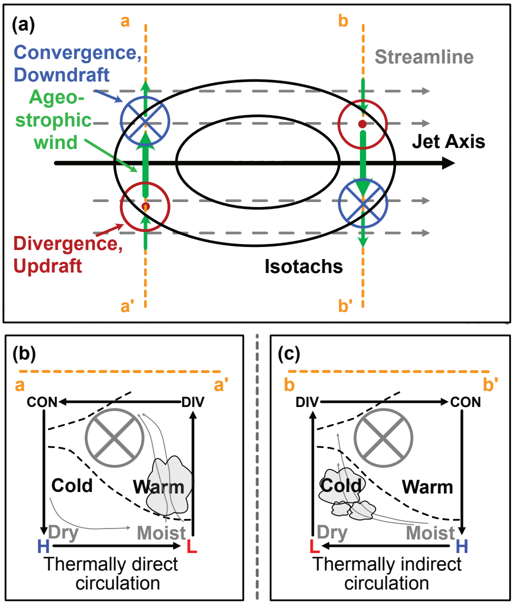

FIGURE 1 Schematics adapted from (a) Beebe and Bates (1955), their Figure 4, and (b,c) Uccellini and Kocin (1987), their Figure 3B. Panel (a) shows an idealised straight jet streak with the associated upper-level convergence and divergence and induced updraft (red dotted circles) and downdraft (blue 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.

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:

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.

1.3 | 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: *Inspired by 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 and different methods to identify low-level jet regimes would help solidify our understanding of upper- and lower-level interaction in jet streak evolution.*

Thank you for taking the time to review this study.

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.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D. and Thépaut, J.-N. (2023) Era5 hourly data on pressure levels from 1940 to present. *Tech. rep.*, Climate Change Service (C3S) Climate Data Store (CDS).

Hersbach, H. and Bell, B. e. a. (2020) The era5 global reanalysis. *Q. J. Roy. Meteor. Soc.*, **146**, 1999–2049.

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

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.