"A New Look at the Jet-Storm Track Relationship in the North Pacific and North Atlantic"

Authors: Zilibotti, Wernli, and Schemm

Recommendation: Major Revisions

Overview:

The authors employ a novel averaging method for the background zonal wind and eddy kinetic energy to examine the subseasonal-to-seasonal variability of the North Pacific and North Atlantic jet streams, and their relationship to variability in storm track activity. The results are rather interesting, and highlight similar dynamical relationships that are at play in both basins, with historical theoretical relationships largely associated with different dominant timescales of variability for each jet. In total, the manuscript is very well-written and I appreciated all of the great insight offered by the authors while explaining their results and their efforts connect results to past literature/classical dynamical relationships. My only substantive comment focuses on the potential for the authors to provide a clearer discussion of their methodological approach and the sensitivity of results to that approach. Given that addressing this major comment may require some additional analysis, I am recommending the manuscript undergo Major Revisions.

Major Comments:

1. The methodological approach could benefit from the inclusion of additional details that improve the reproducibility of the study. For instance, I believe more detail could be provided regarding how the low- and high-pass filters are applied to the datasets within this study (i.e., could it be possible to add in some mathematical formulations, perhaps?). Additionally, I found myself a bit confused trying to understand how the new averaging approach is applied as a function of zonal wind, U. For example, are there certain bin sizes used to assign a timestep to a particular zonal wind speed? How are 120 timesteps selected for each wind speed, and what happens if there are not enough timesteps available for a wind speed? Is anything done to ensure timesteps selected for the new averaging methodology are temporally independent? Finally, is there any sensitivity of the results to the size and position selected spatial domains chosen? In addition to discussing these elements with greater precision in the manuscript, it might also be helpful to include a summary table that contrasts the more traditional averaging methodology with that employed in this study.

Minor / Specific Comments:

1. Introduction
Section-wide: The authors are commended for producing a very insightful and strong motivation for the forthcoming study!

2. Data and Methods
L117: Consider indicating why ERA5 data is used on a coarser 0.5-degree grid rather than its original 0.25-degree grid.

L128–130: How sensitive are subsequent results to the chosen spatial domains? For instance, the NA domain captures much more of the landmass over North America compared to the NP domain – does this potentially affect the results? Furthermore, it appears the NP domain for EKE misses the climatological maximum, whereas the NA domain is favorably located with respect to the EKE maximum. Could the authors comment on the degree to which these domain choices impact the results?

L130: How does this metric work under situations in which there may be multiple jets present at a given longitude?

L139–140: Could the authors provide some more motivation regarding these choices for intensification rate and minimum sea-level pressure? For instance, grounding these values in the cyclone climatologies for each region could be a way to provide objective support for their selection.

3. The fundamental jet-storm track relationship on two timescales
L174: The distribution during March also looks less similar to the distributions during DJF for the NA, which might be worth mentioning.

L198–206: I am having a bit of difficulty understanding the methodology for Figs. 2c,d, unfortunately. In particular, it is a bit unclear to me how the 120 timesteps are selected and averaged to produce the plots. For example, are the 120 weakest time steps selected across all years or just within each individual year? Given the varied approaches utilized within the manuscript, it might be beneficial to add either a conceptual diagram or table that summarizes the details of the different methodologies in order to help keep everything straight.

L249: I believe the wrong section is referenced at the end of this line.

L255–264: These explanations are really helpful in deducing how the averaging methodology works, and I think some of these details could be added earlier in section 3 to assist with the initial interpretation of results from Figs. 2 and 4.

L295: I believe this reference should be for panel (c) in Fig. 5.

4. The effect of jet width
L367: Are these correlations specifically calculated only for values above a certain wind speed? If so, consider specifying that here.

5. Implications of different jet states in DJF for eddy and cyclone characteristics
L427–434: Much of this information, save for the details of the statistical significance test, are provided in the Fig. 9 caption and likely could be omitted for brevity.

L450: The wavier jet structure might also argue for a stronger influence from diabatic processes in the weak jet cases too. Might there be a possibility to highlight to this effect as part of the story? Admittedly, this certainly could be an interesting avenue for future work too.

Figures and Tables:
Fig. 2: Could the correlations referenced in the text in L156–157 be calculated and plotted on Figs. 2a,b?

Fig. 5: Is there a reason why the normalized values for EKE in panel (b) do not extend to 1 for the full year, but do for the winter? Discussion of the interpretation of these normalized values being equal to 1 (or –1) might assist a reader. Also, how come the hatching in (c) extends across the colored bars? Could the statistically significant correlations be highlighted with an asterisk?

Fig. 8: These results are very interesting, but would it be possible to produce plots that show the difference between the strong and weak jet terciles and evaluate those differences for statistical significance?

Fig. 9: The difference between the violet and black stippling is a bit difficult to differentiate. Could different colors be used, perhaps?

Fig. B1: It is a bit confusing that different *y*-axes are used for terms (i) and (ii), is there a particular motivation as to why different axes are used?