

Review of “Trends and seasonal signals in Atlantic feature-based jet stream characteristics and in weather types” by Banderier et al.

Summary

The authors assess trends and intra-seasonal variability of jet streams over the Euro-Atlantic region in summer. A wide ensemble of subtropical and eddy-driven jet streams' properties is analyzed, with the intention to utilize the know-how in follow-up studies into summer extreme weather. To better understand the spatial structure of jet streams, a self-organizing map of wind fields is trained. I consider the study a worthwhile addition to the effort of understanding atmospheric circulation. However, there are several major issues that the authors should address.

Major comments

The majority of the paper focuses on summer jets. While some results are presented on an annual scale, which effectively puts the summer findings into perspective, I believe the primary focus on summer should be more explicitly stated in the title and abstract.

Given that jet streams are highly localized features, I find it potentially problematic to train the SOMs on large-scale wind fields. It is unclear to what extent the properties of the SOM clusters reflect the jets themselves versus circulation variability far from the individual jets. I wonder how the SOM (and your subsequent results) would differ if only the identified jets were used as input instead of the entire wind fields.

The classification of a wind field into a particular SOM cluster strongly depends on the field's average wind speed. This dependence can easily introduce unwanted artifacts in any classification based on Euclidean distance—specifically, the field might be classified with the centroid that has the most similar average wind speed rather than the most similar pattern. One way to address this issue is to use a different measure of similarity, such as pattern correlation, although this is not feasible with SOMs. An alternative approach would be to remove the fields' means before training the SOM.

Was the input dataset weighted by latitude before training the SOMs?

Section 3.4, where jet properties are projected onto SOMs, appears too brief. This section integrates the two perspectives and brings new insights, but it lacks depth. For instance, a quantitative comparison between independent jet properties projected onto SOMs and those calculated from SOM centroid patterns is missing. In line 454, the authors claim that there is a good match but do not provide quantitative evidence. Can jet properties be calculated from SOM clusters (e.g., the patterns in Figure 5), and can the SOM's skill be quantified? Figure 13 highlights several cases where there is a poor match, but these are not discussed in the text. For example, why is the COM speed of the EDJ so much greater in cluster 1 than in cluster 21? Similarly, why is R16 of the EDJ identical for clusters 5 and 16? This section should also serve as the foundation for a critical assessment of the application of SOMs to analyzing jets, which is missing from the current Discussion. Finally, the chosen color scales in Figure 5 are not ideal. While the blue and red monochromatic scales are understandable, using the scale applied for the Double Jet Index would likely be more informative."

It is not clear how the persistence based on SOM nodes is calculated. The authors mention a threshold of one, allowing departures from a given node without breaking a current episode. First, defining the threshold based on the highly idealized regular 2D grid is possible but it does not reflect the real distances between the nodes (cluster centroids). Defining the threshold for instance as the median Euclidean distance between neighboring nodes in the original high-dimensional data space may be a simple way to deal with the intra-summer variability of wind fields. Second, it is not clear how the individual episodes are found. For instance, imagine one has the following sequence of clusters: 15 1 8 8 8 2 8 8 8 15. Your description suggests that the following episodes are

identified: a) 15, length=1; b) 1, length=8; c) 8, length=7; d) 2, length=4; e) possibly another 8, length=3 (?); f) 15m length=1. Is this correct? Please clarify in Sect. 2.2.3.

Several times throughout the paper (see minor comments below), the similarity of SOM patterns with modes of variability and weather regimes is mentioned. However, aside from the NAOI, this is presented only for selected patterns and not quantitatively. Linking the findings to established modes and/or regimes is an excellent idea that should be developed in greater detail.

The quality of the Discussion and Summary section could be improved. Parts of Section 4 are overly vague and difficult to follow, and the section lacks a clear summary of key results. Additionally, portions of the Discussion appear only in earlier sections (e.g., the description of annual results in lines 343–441), which disrupts the flow and focus. Consider restructuring or splitting the section to enhance clarity and coherence.

Minor comments

13 The term “seasonal cycles” is not clear to me; I suppose that the authors mean intra-seasonal variability? This appears several times in the paper.

19 abbr. 2PVU not defined, and I think that the term “flattened pressure fields” may also be unknown to many readers

24 “sudden flow transition in June” seems unclear to me at this point

52 both references are missing from the references list

77–78 The term 'data-driven approaches' does not seem appropriate to me. Perhaps terms like 'complex,' 'black-box-like,' or 'dimension-reduction-based' could be more fitting for these approaches. Neither of the approaches you use is strictly objective or subjective; they both require user/expert decisions and involve many parameters that need to be defined, which can potentially have a significant impact on the quality of the results. In the context of using SOMs in general, and specifically to study the link between circulation and extremes, I recommend referring to studies by Gibson et al. (10.1002/2016JD026256) and Stryhal et al. (10.1002/joc.7996).

82 It is not clear to me which interactions you are referring to here

89 „tool to study stationary and recurrence“

89 Please change your reference from preprint to the final paper (<https://esd.copernicus.org/articles/14/955/2023/>)

93-98 the use of language in this paragraph needs to be reviewed

93-98 assess the seasonal cycle vs focus on summer ... please reword

102 ERA5 provides much more than that; please reword

104 I suggest mentioning which levels are used exactly

110 „geopotential“

110 What kind of anomalies do you use?

129 I am not sure that I understand what the „scale parameter“ means. It is not used in Kohonen (2013) or the very well-known and cited papers on SOMs in atmospheric science, Hewitson and Crane (<https://www.int->

res.com/abstracts/cr/v22/n1/p13-26/) or Sheridan and Lee (10.1177/0309133310397582). On the other hand, parameters such as (learning) rate and (neighborhood) radius, which considerably affect the resulting map, are not mentioned.

133-134 This claim would benefit from a reference. Furthermore, instead of 'allowing,' I would suggest using 'forcing.' Arguably, it is not the similarity of nodes that causes this, but rather the tendency of SOMs to over-represent the center of the data space and under-represent its margins (i.e., extreme fields). This is compounded by the 2D constraint, where nodes lie on a plane even for multi-modal data, or more generally, in a space with much lower dimensionality than that of the data.

135 “projected trajectory” is not clear

136 I do not think that x was defined

144+Figure 1 Many readers may not be familiar with periodic boundaries, as planar topology is typically used in atmospheric science. Please provide more details on the toroidal SOM topology employed and explain why it was utilized instead.

165 reword

182 Is there a reference missing?

Fig. 4 “a, b) For each season, here JJA, ...” is a bit confusing, since the demonstration is only for JJA in these panels

268 “this framework...” is not clear, please reword the sentence

300-301 This is not clear to me

304 Please explain why COM speed is considered a measure of persistence? I am not familiar with this metric and it seems a little bit counterintuitive.

315 phase > phase space

321 maybe “edge” instead of “extremal”?

322+onwards – It is not clear to me how the SOM nodes were associated with regimes; an objective analysis is carried out only for NAOI. For example, why was node2 associated with Scandinavian blocking and node 3 was not? Why was node 22? If you feel that this is an important addition to your study, an objective analysis involving lower-order modes of variability (EOFs) and/or weather regimes should be carried out. Otherwise, one may doubt the reliability of interpretations, such as that in line 336-338

Fig. 6 Correct “b” and “c” descriptors

332-334 While I find this description mostly acceptable, there are some misconceptions and inaccuracies: 1. What would constitute a perfect alignment of SOM axes with the NAO? 2. You suggest that the SOM you are using (which is technically a 3D lattice organized on the surface of a torus and then unwrapped onto a 2D surface for visualization) should align with the leading two modes of variability (note that the second mode is not analyzed at all in your paper). Can you support this with a reference? I found a study (10.1029/2023JD039183) suggesting that SOMs behave this way only in very specific cases, but only planar grids were tested in that study. 3. It is unclear how relevant your suggestion regarding the PCs is, given that your object of interest is jets."

339-342 Is this “closer inspection” rooted in any of the presented analyses? I cannot find anything that would support this. Please clarify.

Figure 7 The frequency of occurrence of cluster 8 in Week 1 is striking. Do you have an explanation for this? Averaged over almost 65 summers, how is it possible that this cluster is so much more likely in Week 1 than in Week2? One would expect that the pathway will be much more gradual.

349 “Now we add a temporal dimension” is vague and inaccurate; was not the temporal dimension included before?

351-2 Figure 7 does not show empty clusters –is it possible that all clusters are visited during the whole summer? Please reword.

352 “left to right” is not clear

353 I do not think “JA” is very clear

354 “early June” is unclear; Do you speak about the first week only (based on Week 1 in Figure 7)? Moreover, the frequency of clusters varies a lot in June, so it is not clear why NAO is described only for these three particular clusters. Additionally, your current wording suggests that NAO in clusters varies week to week.

355 “these few weeks” is unclear. Previous paragraph described only the first week

356 What kind of error do you refer to?

357-8 Do the days have free will in how they choose clusters? Consider rewording.

350-359 It seems to me that in June there might still be a considerable decrease in wind speed across the whole region. Since you do not subtract this seasonal change (e.g. by removing each field’s average wind speed), the SOM badly represents early June wind patterns. A few nodes represent a significant part of the data space (in terms of its volume). Consequently, the cluster centroids are very poor models of the fields classified with the clusters, and, in turn, very likely also of the jets’ properties. This is a very common issue if one uses classifications based on the Euclidean distance, which can easily lead to misinterpretation. For instance, the high persistence of cluster 8 might be an artifact of this issue; or, the positive trend in its occurrence might be unrelated to changes in jets, but may rather reflect a minor trend in wind speed, possibly even in regions far from jets’ occurrence. Consider additional sensitivity studies to make sure that statistical artifacts do not affect your results, or at least mention in the Discussion.

361 I would suggest referring to the specific section given the length of the Methods section

363 “an next state”

366-7 I disagree with your explanation of hotspots in the corners. Given the grid topology and the values for clusters 1, 7, and 24, it seems inaccurate. I observe one hotspot in the cross-section of the SOM torus, specifically at 2-8-13-14-20 (which could represent early summer patterns), and another around 17-18 (corresponding to mid-to-late summer patterns). I suggest visualizing the data space and the SOM in 3D, for example using Sammon mapping. This is a useful tool for understanding the structure of the data and how the SOM represents it. It may also be a better alternative to your Figure 1.

370 I do not understand the term “true state persistence”. Please reword

370 To what extent is persistence only (the liner) function of the frequency of occurrence?

375 Rather than linear trends, one may utilize the SOM grid and visualize the seasonal frequencies and persistence of each node as x,y plots

382 Are you sure that cluster 6 resembles Greenland blocking? It does not seem obvious from Fig. 5. Maybe including the z500 patterns currently in Supplements would help.

384 Which clusters do you refer to?

395 JLI and JSI were already defined in 2.3.2; R16 technically wasn't defined yet; abbreviations are not used consistently in the whole section. I would suggest not using them outside of pictures to make the text clearer.

398 "this property" is not clear

412-413 Consider moving to Discussion

418-419 How? This should be discussed more and moved to Discussion

469 "clear season signal" is too vague and could be easily mistaken for summer trends; consider intra-seasonal variability or similar instead

470 I am missing a section that would show and discuss the differences between the methods and their limitations. These paragraphs read a little bit too vague and one-sided.

474-475 Did not you include the other group of nodes (that around cluster 8) in this description intentionally? If so, please discuss the reasoning?

481-482 Reword + Which features?

485 It makes sense to define the abbreviations again in the summary section – but why so late?

492 "subjective jet properties" is not clear since each method is an "objective" algorithm based on many subjective or expert choices

494-495 Why was this not quantified?

513-514 Change reference to the final paper (<https://doi.org/10.5194/wcd-5-1269-2024>).