Replies to Reviewer #3

August 2020

We would like to thank the Reviewer for the very thorough set of comments provided. We provide a detailed reply (in red) to the individual comments (in *italics*) below.

General comments

1. *Overall, the authors refer too often to external literature for points/facts that could be briefly elaborated in this manuscript itself to make it a well rounded work, without the readers having to consult over 10 external papers to actually find out what has been done and what the basic assumptions are. Having said this, the manuscript is currently lacking a thorough assessment/explanation if the techniques that the authors use in their analysis are actually applicable to the data in their case study. i.e. if the data abides to the ideal conditions for which the dynamical systems metrics were formulated. Hence the authors need to mention what the basic mathematical and physical assumptions underlying the methods and also need to show that their data meets these criterions (which is often not the case when using measured data). Since dynamical system approaches are neither universally nor straightforwardly applicable to real-world applications, I think it is dangerous to just blindly apply methods to data without properly conveying the basic mathematical assumptions underlying the fundamental theory and calling the approach 'versatile'.*

We have taken the Reviewer's concerns on the lack of a systematic discussion of (i) the theoretical bases and (ii) the limitations of our approach very seriously. To address the first point, we have restructured Section 2.3, where we have now added details on the derivation and interpretation of the metrics (see also reply to Comment #1 by Reviewer #2) and we have added a new appendix (Appendix B) providing a thorough derivation of d and θ. To address the second point, we have added a brief mention of the limitations in applying our framework to real data in Sect. 2.3, and have provided a more detailed technical discussion in Appendix B. Finally, we have removed the term "versatile" from the concluding sentence of the study.

2. *Another shortcoming of the manuscript is the subjective qualitative inspection of how the diagnostic outputs fare in phase space, i.e. the methods applied by the authors allow only a purely descriptive and non-explanatory evaluation. Hence the authors should be more careful when assessing their outputs and avoid bold statements about their alleged ability to provide physical, process-based explanations that in the end are not proven.*

We have rephrased a number of passages in the text to this effect, following the Reviewer's specific comments. We refer the Reviewer to our replies below for the details of these changes.

3. *Additionally, the authors need to be more careful in reading/interpreting their figures, sometimes the statements made do not quite correspond to what the figure is actually showing (e.g. fig 3).*

We have revised the discussion of the figures in the text as suggested by the Reviewer in his/her specific comments. We refer the Reviewer to our replies below for the details of the changes we have implemented.

4. *The authors discuss the special case/simulation no 23 in detail, however, I think the simulation no 7 that stands out in Fig 2 as not 'fitting' the order of decreasing EKE and jet group should also be discussed briefly. I.e. where does it sit in Fig 3c/d?*

We have analysed simulation no. 7 as suggested by the Reviewer, and find that it does not appear to be anomalous in terms of its placement in Fig. 3 in the main text (see Fig. 1 below). The Reviewer may have noticed it in Fig. 2 in the main text due to the fact that it appears as a mixed jet amidst eddy-driven jets. However, we conclude that this points to the fact that the simulation may be at the boundary between the "mixed" and "eddy-driven jet" classifications (see Fig. 1d below) rather than to the fact that it displays any remarkable dynamical features.

Specific comments

1. *L4: why 'our understanding'? Do the authors mean 'their own understanding'? this is confusing...does this differ from the 'general understanding'?*

We used "our" to refer to the community studying atmospheric dynamics. We have now changed this to "the current understanding".

2. *L7: this reads as if it could be directly applied to the reanalysis data (which is not the case as the authors do some pre-processing to the data). Please make it clear that it is not the 'raw data'.*

We have removed the term "data" to avoid this misunderstanding. We further specify that the reanalysis data we use here is not "raw" in the introduction: "This approach can easily be applied to a variety of datasets, including suitably processed reanalysis data".

3. *L8: I'm not sure if this study can be called 'proof-of-concept', as there were limitations in the output. For a concept to be proven, more than a single case study is needed. Hence, please rephrase. The same applied to L 275*

We disagree with the Reviewer on this issue. A proof-of-concept, as defined by the Macmillian English Dictionary (but definitions of other dictionaries would resemble this) is: "evidence, usually obtained from a pilot project, that an idea [...] is likely to succeed". We believe that our study fits this definition, and have thus opted to retain the original phrasing in the text.

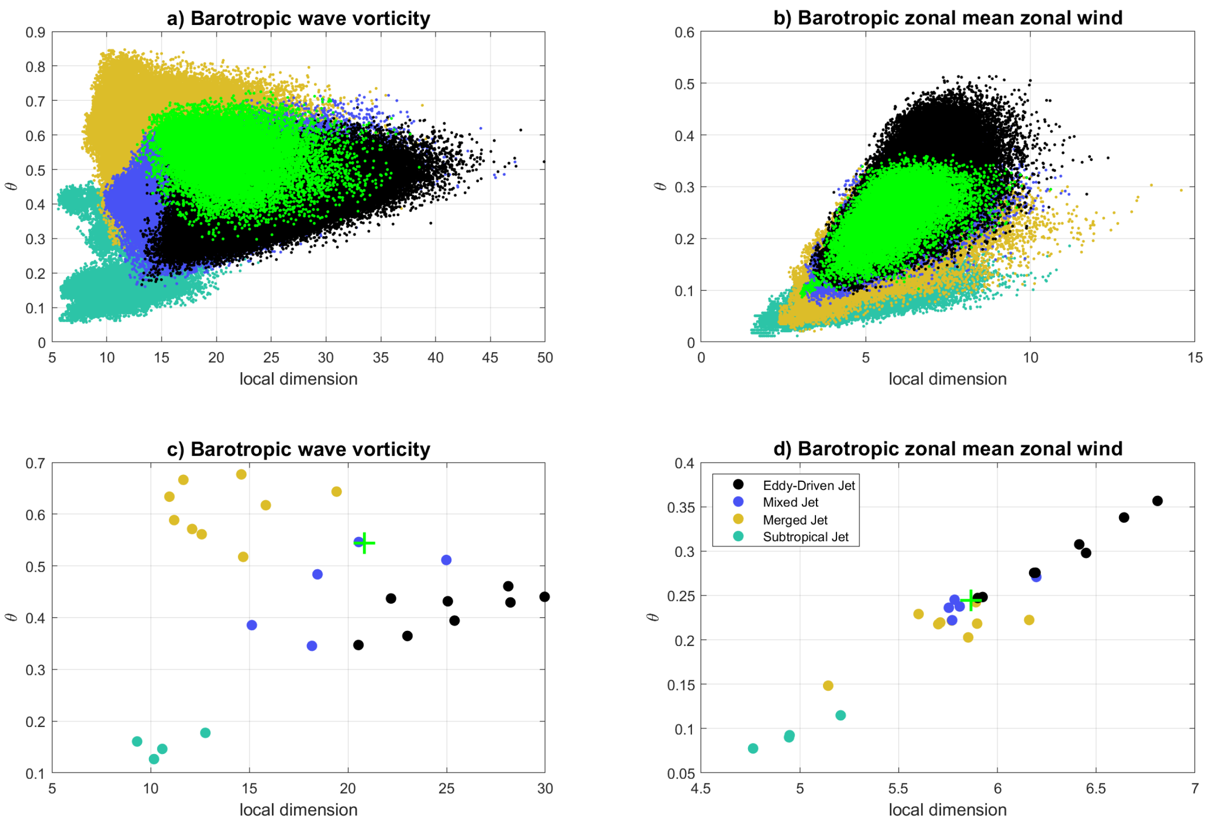

Figure 1: d–θ diagrams of barotropic wave vorticity (a,c) and barotropic zonal mean zonal wind (b,d) for all timesteps in the simulations (a,b) and centroids for each different model simulation. Colours indicate the different jet regimes. The bright green dots (+ symbols) mark the (centroids of) simulation no. 7. Note that higher (lower) θ values correspond to lower (higher) persistence, and that the axes' ranges differ across the panels.

4. *L 8-9: 'versatility and 'computationally efficiency' have not been investigated here and should hence not be a key output mentioned in the abstract and closing sentence of the manuscript (I.e. L 276). Please rephrase.*

We agree that the use of these terms may not be warranted in the abstract, and have proceeded to remove them. We have however chosen to retain them in Section 5, where we added brief explanations to support these claims. Indeed, in Section 5 we mention these aspects in the context of a broader discussion of the approach, where we draw from both results from the present study and results from the literature.

5. *L 19: 'to a good degree'. Please be more specific (i.e. how is 'good' quantified) to allow the reader to understand the 'degree' of approximation. Or make it clear that the 'limitations are described below'*

As suggested by the Reviewer, we have now rephrased this passage to highlight that the text in the following paragraphs qualifies the use of the adjective "good" and summarises the main limitations of this view of the atmospheric jets.

6. *L 88: To allow reproducibility of the current study, please provide the model equations as used in the final model in the appendix (no derivations needed).*

We fully agree that including the model equations will make the study more self-contained, and now provide these in the expanded Appendix A.

7. *L 89: Please provide a concise summary of the main limitations of the setup that are relevant to the current study*

We have now removed the original phrasing and added a short discussion of the key limitations of the model.

8. *L 90: For reproducibility, please provide the all these details in the current manuscript.The manuscript should be a 'stand-alone' piece of work and one can not expect the readers to go to the authors past works to search for all the details.*

As recommended by the Reviewer, we now provide these details in the expanded Appendix A.

9. *L 95: Why 27? Please elaborate how this number was chosen and also how the parameter ranges and the different combinations of H and r were chosen. Additionally, why did the authors choose to have not equal numbers of simulations for each flow regime? This would allow a better comparison of the results.*

We now explain the motivation for choosing this particular set of parameter values in Sect. 2.1. Specifically, the 27 simulations issue from choosing regular increments of H and r of 0.5 km and 0.5, respectively, over the range of values which yields flows resembling those observed in the real atmosphere (i.e. without reaching the limiting case of completely stabilised eddies). We have further removed the simulation with H = 8 km and r = 2 from the analysis, because it was dominated by unrealistic regular oscillations of the eddy amplitude. Since the different flow regimes are not regularly spaced in the parameter phase-space, we opted for a regular sampling rather than sampling specific phase-space regions more densely than others.

10. *L 97-98: Please briefly elaborate on the representativity of using the mean values here.*

We added a sentence explaining the choice of using the barotropic (vertical mean) variables in the analysis. We further mention that in an early stage of the analysis, we did test using the full 3-D fields but found they provided no major additional insights while increasing the computational costs of the analysis.

11. *L 100: Please elaborate on the degree of similarity, i.e. what do the authors consider to be 'similar'*

We agree that this statement was vague. We have now rephrased this to reflect that the results allowed us to reach analogous conclusions concerning the relative differences in d and θ between the different jet regimes.

12. *L 109: 'we take these as representative'. I.e. the authors consider them as representative. But are they? Please elaborate why the authors think so.*

We have rephrased this passage to express more clearly that these are common choices (albeit not the only ones possible) in the atmospheric dynamics literature when studying flows in the upper or lower troposphere. We further provide a reference to a recent study that uses these in the context of atmospheric jets.

13. *L 110: 'To better compare...' Please elaborate why the daily data needs such a high-pass filter. i.e. elaborate on the data characteristics that make the use of such a filter necessary. How does it influence the results?*

The use of filtering to extract eddy components from atmospheric data has become widespread since the seminal works of Blackmon, Lau, Wallace and colleagues in the 1970s (see e.g. Blackmon et al., 1977). Here, we specifically apply a filter to decompose the flow field into eddy and mean components. This is a widely used conceptual decomposition , and indeed one may go so far as to state that the filtering is an integral part of the definition of atmospheric eddy fields. We now explain this briefly in the text.

14. *L 115: 'applying a recently developed approach...' But the reverences given are from 2016 and 2017. This is not quite 'recently' as we are already in 2020... Please rephrase.*

As is often the case, our approach was developed and refined over a period of time. The bases were laid in 2016 and 2017, but for example the final version of the estimator of d was actually developed in early 2019, and further developments are ongoing. We have updated the text to clarify that the 2016 and 2017 references provide the bases for our approach, but do not necessarily represent its cristallised, final version.

15. *L 117-140: As this journal is for an interdisciplinary audience the authors should elaborate in more detail what is being done. Currently with the given description the work is not reproducible. Particularly, the reasoning of why certain theorems (and their modifications) are being used (e.g. why are they appropriate to the system under investigation and what are the basic data assumptions) are missing.*

We fully agree with the Reviewer that this was a shortcoming in the original text. We have now made a systematic effort to provide the full details of the approach by both extending Section 2.3 and compiling the new Appendix B (see also our reply to General Comment #1 above).

16. *L128: what is the 'Süveges estimator'? Elaborate on the basic information/assumptions etc.*

As suggested, we now provide a brief discussion of this estimator and its underlying assumptions in the new Appendix B.

17. *L 131: 'In our case, we deem daily data sufficient...' Why?*

The type of jet variability we are interested in, related to large-scale shifts in the flow rather than synoptic or smaller-scale flow variability, is successfully reflected by daily data. Indeed, this is the typical timestep chosen by many studies looking at the so-called jet regimes (e.g. Woollings et al., 2010; Madonna et al., 2017). We have now briefly clarified this point in the text.

18. *L 135: Please elaborate in what sense/how 'the two are partly related'.*

This point has also been raised by Reviewer #1 (see our reply to their comment #1), and we appreciate that it was not satisfactorily detailed in the original text. We have now expanded the discussion of predictability in the text, and further explain that the relation between the two amounts to the fact that they provide partly overlapping information on predictability.

19. *L 136: 'This approach has been successfully tested...' Which approach? The entire section? Please specify what has been tested before. Also elaborate how this manuscript now is different to the numerous citations of the authors previous works. What is scientifically new?*

We have rephrased the original sentence, separating the references between those focussing exclusively on climate datasets and those which explored other dynamical systems. Concerning the novelty of the study, we have added a brief clarification on this point in the introduction.

20. *L 144: For this manuscript to be a self-contained work, please briefly elaborate the main characteristics of this classification.*

We added a description of the jet regime classification at the beginning of Section 3.1, as suggested by the Reviewer.

21. *L 142-154. I think the interdisciplinary/non expert reader might find it difficult to see the points described in the figures mentioned. Some additional elaboration, guiding the reader to important points of the figure will be helpful.*

We have reorganized this paragraph and added more details regarding the figure referred to in the text, to make it clearer for the non-expert reader.

22. *L 156: Please briefly elaborate the differences in the 'temporal variability', that lead to the suggestion of 'well-separated' dynamical systems characteristics.*

We have clarified that our hypothesis issues in part from the eddy energy spectra of the different regimes shown in Fig. 2b, which underscore differences in the flows' spatio-temporal characteristics.

23. *L 166: 'clear dynamical separation...' Visually they do not look so 'clearly' separated...Please elaborate why the authors think there is a 'clear' separation in 3d*

We appereciate that there is no formal definition of "clear separation" in a numerical sense. We have therefore rephrased this passage to: "Although the spread in d and θ within each simulation is large, the centroids of the different simulations are organized in specific regions in the d-θ phase-space, according to the corresponding jet regime."

24. *L 170-171: 'display lower d than the eddy-driven jet regime, yet a higher theta...'. This is true for the centroids of the merged jet regime, however not for the centroids of the mixed cases, where there is quite an overlap of the space occupied by d and theta!*

We agree that this was an imprecise description of the results shown in the figure. What we meant to state was that *on average* the mixed jet regime shows lower ds and higher θs than the eddy-driven jet regime. Indeed, when computing the mean values we obtain ($d = 19.74$, $\theta = 0.45$) for the mixed jets and ($d = 25.55$, $\theta = 0.41$) for the eddy-driven jets. We have rephrased the text to this effect, and further explicitly mention that the mixed jet regime shows substantial overlap with the eddy-driven jet regime.

25. *L. 197 and Fig 5: Please add to the manuscript the criteria used to determine which points are plotted in red (i.e. how was the period of day 2895 to 4076 selected?).*

We chose this interval based on a visual inspection of Fig. 5c, which we now specify in the figure caption. We note in this respect that minor changes to the chosen interval do not alter our qualitative conclusions. As example, we provide in Fig. 2 below the plots corresponding to Fig. 5a, b for the interval 3000 to 4200 days.

26. *L 201: it is not the 'dynamical systems metrics' that 'identifies this anomalous transition', but*

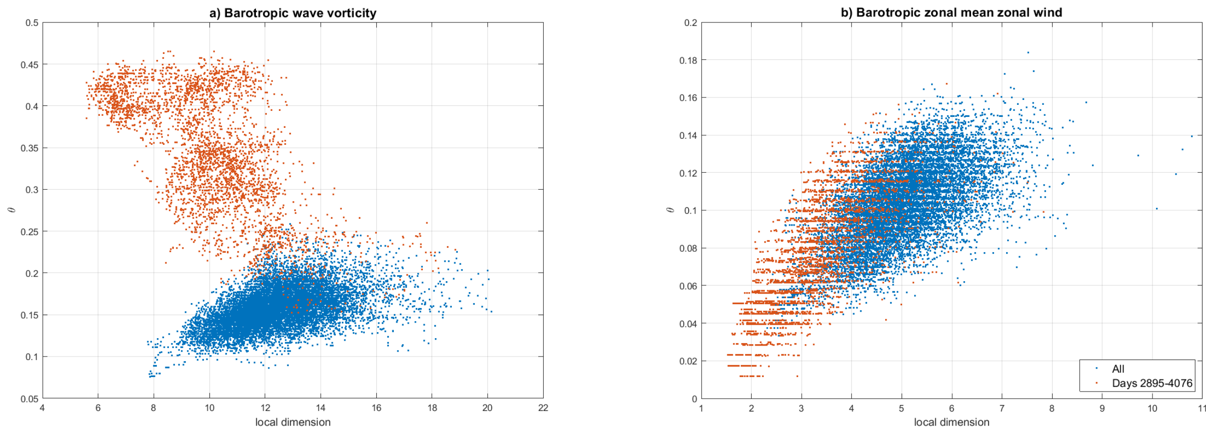

Figure 2: d–θ diagrams of barotropic wave vorticity (a) and barotropic zonal mean zonal wind (b) for all timesteps in simulation no. 23. (c) Latitude – time diagram of barotropic zonal mean zonal wind ($m\ s^{-1}$, colours) and eddy kinetic energy (EKE, $m^2\ s^{-2}$, contours) for the same simulation. The red dots in (a, b) correspond to days 3000-4200 in the simulation.

rather the authors by analysing the output! Please rephrase.

The Reviewer is absolutely correct, and we have rephrased this passage.

27. *L 215-216: Again, as mentioned above, there is no 'clear separation of regimes' in Fig 3! please rephrase. Additionally, currently it reads as if the existence of a clear separation dictates the choice of the 10th and 90th percentile... This is not a proper elaboration, please rephrase.*

This was indeed a badly formulated statement, which we have rephrased. What we meant to say was that conditioning on joint high or low percentile of the two metrics, as opposed to for example selecting high d and low θ, was dictated by the fact that for the cases where there is a clear separation between regimes in barotropic zonal mean zonal wind, this emerges along a d-θ diagonal.

28. *L 254: To me these are not 'intuitively related to the concept of predictability'. Please elaborate more on why the authors think this is the case.*

Following comment #18 above, and comment #1 from Reviewer #2, we have significantly expanded the discussion of predictability in Sect. 2.3. We now refer the readers back to that section and have further removed the term "intuitively", which we agree is highly subjective.

29. *L 255: why would be a low d be 'more predictable'? Please add explanation for readers to understand*

As mentioned above, we have expanded the discussion of the link between the dynamical systems metrics and predictability in Section 2.3, where we briefly motivate why a low d may be expected to match a more predictable state. We now refer the readers back to that section to support our claim.

30. *L 265: how do the authors define 'successfully'? As 'success' is a relative term, I suggest rephrasing to avoid ambiguity.*

As suggested, we have restated this in more neutral terms.

31. *L 269: Please also discuss the role of the filter used on the data, and if the results would have been the same if no filter would have been used.*

d and θ are computed on unfiltered data. The only filtering we apply is to separate the eddy and mean components of the atmospheric flow to produce the EKE composites. However, as we explain in the reply to comment #13 above, filtering may be taken as a part of the very definition of eddy fields. As such, the filtering itself is not a subjective choice that has an impact on the results. Rather, if we did not use filtering we would not be able to compute EKE in the first place.

32. *L 273: Please add a sentence as to why the authors think/possible reasons the method does not find a single eddy-driven jet in the reanalysis data.*

We did not express ourselves clearly in this passage. The lack of a single subtropical and a single eddy-driven jet is not due to our approach, but rather to the fact that in the real atmosphere, these two jets do not occur in isolation. Indeed, two distinct branches emerge in the Pacific sector: a subtropical jet at around 30 °S and a polar-front jet at around 60 °S, but neither occurs in isolation. We discuss this briefly in the introduction, and now refer to it also in the concluding paragraph.

33. *Figure 2a: For ease of interpretation/completeness please add the symbol x with meaning to the legend*

We have updated the figure legend as suggested.

34. *Figure 3 (and for all your future figures in publications) Please do not use red and green colours together in a figure. Such a colour choice is not 'colour-blind safe'! Please prepare figure in line with the 'manuscript preparation' guidelines of ESD (see point 7 under 'Figure composition': 'please keep colour blindness in mind and avoid the parallel usage of green and red. For a list of colour scales that are illegible to a significant number of readers, please visit ColorBrewer 2.0.'*

We fully agree with the need to make our figures accessible to those with visual impairments. As far as we can tell, the choice of a combined red-green colourscale was an issue only in Figure 3. We have now updated this figure, selecting our colours from a black-blue-yellow gradient.

35. *Additionally, for consistency, please use the same legend labelling as in Figures 2 and 4, particularly as the abbreviations used have not been introduced before.*

In the revised Fig. 3, we now spell out the different jet regimes in the legend, as was done for Fig. 2.

36. *Figure 4: Please add a legend that explains the different colours (i.e. colour scale) used in the plot.*

We have added a colourbar to each panel and updated the caption accordingly.

37. *Figure 5: For completeness, please add units to wind. I presume that it should read 'the red dots in (a, b)' not (a,c)?*

We have added the units of both zonal mean zonal wind and EKE to the caption. We further thank the Reviewer for having spotted our mistake in referring to the panels, which we have corrected.

38. *Figure 7: Pressure level please use linear y-axis scale or elaborate why a non-linear axis type was chosen.*

We have removed the non-linear pressure axis (which is often used to capture atmospheric structures that extend into the stratosphere), and now adopt a linear scale. Following comment #10 from Reviewer #2, we have further combined these two panels into the new Figure 6.

39. *Figure 8: Please adjust the colour scale and limits to match the actual absolute maximum value. In 8a, b, the darkest colour shade does not seem to appear on the map. i.e the colour scales should only show the maximum value.*

The limits of the colour scale in Figure 8 have been modified, so that now both the darkest blue and red shadings appear in at least one of the panels.

References

Blackmon, M. L., Wallace, J. M., Lau, N. C., Mullen, S. L. (1977). An observational study of the Northern Hemisphere wintertime circulation. *J. Atmos. Sci.*, **34**(7), 1040-1053.

Madonna, E., Li, C., Grams, C.M. and Woollings, T. (2017), The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Q.J.R. Meteorol. Soc*, **143**: 2960-2972. doi:10.1002/qj.3155

Woollings, T., Hannachi, A. and Hoskins, B. (2010), Variability of the North Atlantic eddy-driven jet stream. *Q.J.R. Meteorol. Soc.*, **136**: 856-868. doi:10.1002/qj.625