Review of “Medicane Zorbas: Origin of an uncertain potential vorticity streamer position and impact on cyclone formation” by Portmann et al.

Background

This revised submission documents the prediction of Medicane Zorbas’ development in the ECMWF ensemble initialized approximately three days before the event. The authors identify differences in member solutions depending on the position of the PV streamer that strongly influences the cyclone life cycle. They suggest that the differences in the streamer location over the Mediterranean are related to differences in member initial states over the Gulf of St. Lawrence. The authors also show that Zorbas’ structure is sensitive to the position and structure of the PV anomaly as the storm intensifies.

The subject of the study remains very interesting and relevant. The major changes made by the authors have made certainly improved the analysis and presentation of the results. However, there are a number of issues with this version of the submission that remain to be addressed. I hope that these comments will be useful to the authors as they further revise this work.

Reviewer: Ron McTaggart-Cowan
Recommendation: Major Revisions

General Comments

1. The main body of the text (Abstract through Acknowledgements) is pushing 10 000 words, which is >30% longer than a standard submission. This makes the study a very long read, especially for anyone who also wants to dive into the supplemental material (a further seven pages, although mostly figures). This length combines with a persistent lack of flow to make the submission a very long read. This is unfortunate, because the subject of the manuscript is interesting and it would be a shame if paper did not get the attention that it deserves simply because it is difficult to get through. Consider the sections: “Zorbas life cycle”, “PV streamer clusters”, “initial differences”, “PV streamer impact on development”, “PV streamer impact on structure”. The ordering of this set of section descriptions does not appear to follow any particular logic, leaving the reader to bounce around in both time and spatial scale (two useful ways that the content could be ordered). I think that reordering the manuscript would let the authors eliminate redundancies to reduce the length of the paper by 25% or more, a change that would be of real benefit to the study. An alternative that is discussed more in General Comment #6 is to remove section 5 completely (to be further developed into an independent study), a change that would limit length and improve the flow substantially.
2. As noted in my first review, I think that the manuscript tries to cover too much ground. This is exemplified by the main objectives of the study (L94-96), which expand quickly from two to four when the itemized list is broken apart into separable elements. As in the initial submission, it is noted that the storm’s tropical transition “is not the focus of this study” (L188); however, the CPS (section 2.2) is used only to classify core thermal structure in different members, an exercise to which section 6.2 is entirely devoted. This kind of internal inconsistency makes it seem as though the manuscript is in the midst of an identity crisis, leaving the reader with without a clear picture of the true objectives of the work. One component of the analysis that could be dropped is section 5. Although I think that identifying the source of the uncertainty is

very interesting and possibly important, it adds significantly to the volume of information presented in the text and lacks the depth required of such an analysis. Removing this section would drop almost 2.5 pages out of the submission and free up figure space for images currently found in the supplement. It would also help to keep the focus of this manuscript on Zorbas itself, rather than the state far upstream.

3. I understand that you disagree with me on the formulation of a paragraph. However, I really do think that this manuscript would be very well served by reducing paragraph lengths throughout, such that each paragraph deals with a single subject. This will help to manage the breathless pacing of the text.
4. I don't understand or approve of the use of model precipitation as "truth" (L114-115), particularly when observational estimates are readily available (i.e. GPM). All modelling systems have precipitation spin-up problems, making the use of 0-6h forecasts particularly problematic. These data are degraded to a 1o grid, hopefully using conservative remapping; however, this is unclear in section 2.1. On L482-L84, the weaknesses of model precipitation estimates are highlighted. This acknowledgement of model QPF shortcomings (especially in convective cases) is inconsistent with the use of short-range forecasts as "truth". Please use an independent, observation-based QPE as the reference for this study.
5. Why use the full standard deviation in the denominator of Eq. 1? If this were a test statistic (it looks a lot like the two-sample t statistic), then the denominator would be the pooled standard deviation (i.e. the sum of the individual sample variances). I understand that the form in Eq. 1 has been used before in the literature, but if you consider two tightly clustered samples with small internal variances and a large difference between the sample means, then use of pooled variances will make the normalized difference much larger than that estimated by the current formulation. This may serve to highlight regions of systematic differences between the classes more effectively. Such a change will also bring Eq. 1 in line with standard assessments of differences between groups by comparing between-group differences with within-group variability (i.e. ANOVA).
6. I think that section 5 should be removed from this study and form the basis of a separate piece of work instead. The analysis of the error source is potentially very interesting, but (a) distracts from the discussion about Zorbas, and (b) is not rigorous enough for inclusion without a lot of additional investigation. Terms are dropped from the equations in this section without any quantitative justification: simply stating that terms are likely to be unimportant is not acceptable. The fact that the reduced form of the equation is then used to "imply" something about the relevant dynamics is specious reasoning (L324-325 and L331-333). Similarly, the decomposition into geostrophic and ageostrophic winds for some of the terms in Eq. 2 (becoming Eq. 3) is not physically justified, any more than a decomposition into any other vector components would have been (L325-327). The standard decomposition of the wind field under such conditions would be into nondivergent and irrotational components. The secondary circulation of a jet streak also has a strong projection onto the divergent wind, so it is not clear why geostrophy is preferable here, particularly given that it is related to a lower order of balance. The choice is important, however, because the current decomposition is used to conclude that "the described amplification [is] due to the ageosterophic circulation" (L500). I do not agree that this conclusion follows from the analysis presented in section 5. Indeed, the current analysis does not really serve to advance the main objectives of this study. A full study that looks carefully at the PV error growth in this case would be a nice follow-up to the current submission.
7. Throughout the text, the authors refer to vertical alignment of PV features as implying a "barotropic" structure. In a barotropic environment, there isobaric and isothermal surfaces are aligned, so there is no vertical wind shear. Given that wind shear is being used to diagnose core

structures in this study, I believe that the authors consider this to be an “equivalent barotropic” atmosphere.

Specific Comments

1. [L22] Suggest “increasing” to “intensifying”.
2. [L27-30] This TT intro is too cursory to be useful to anyone not familiar with the concepts. Given that section 6.2 is dedicated to warm core formation, I think that this description should be improved.
3. [L30-31] The TT process involves far more than WISHE.
4. [L37] Suggest “rather” to “more”.
5. [L40] Note smaller than typical hurricanes.
6. [L42] Not “high damage”.
7. [L53-54] Strange inclusion of result in introduction.
8. [L54-55] Suggest citing Maier-Gerber et al. here.
9. [L60] Suggest “high” to “large”.
10. [L61] Referent of “they” unclear.
11. [L64-70] Awkward transitions throughout paragraph.
12. [L75] Suggest “are” to “have been”.]
13. [L88] Double period.
14. [L89] Remove “the”.
15. [L92] Reconsider use of singular throughout this sentence.
16. [L103] Suggest “previous” to “prior”.
17. [L106] “Section” misspelled.
18. [L129] Suggest, “The CPS uses three parameters to define thermal structure:”
19. [L135] Isn’t there a better reason than “simplicity”? Could it be that none of the members differ because this is a larger-scale parameter? If there are large differences in B between members, then I think that it should be considered as an important dimension of the CPS.
20. [L157] This statement assumes that variance is large when gradients are large: not necessarily the case if the state is very strongly constrained by observations.
21. [L183-184] Please provide citations at the ends of each of these sentences.
22. [L191 and L193] Are these the same “R” features described as “moderate” and “large”?
23. [L199] Should be “events associated” I think.
24. [L203-204] Suggest simplifying the first two sentences.
25. [L214-217] I think that this discussion of the storm’s vertical thermal and PV structure would really benefit from a cross-section through the centre.
26. [L222-223] It is really unclear in this discussion that the authors consider 38% a small number. Maybe this sentence should start with “Despite the prevalence of convection, only 38% ...”.
27. [L222] What area is used here?
28. [L224] How does stretching affect the cyclone path here?
29. [L227] Why is the fraction of parameterized convection at later times relevant?
30. [L227] Could also add that the DWD analysis does not show any fronts associated with Zorbas at this time (it showed them at 1200 UTC 27 September, but not thereafter).
31. [L238] Add missing “the”.
32. [L239] The previous section showed that the PV streamer was present, not that it was “crucial”.

33. [L241] It would be useful to say a couple of words here about the impact on Zorbas' outcomes in the members, otherwise this whole development of the streamer classifications happens without a clear motivation.
34. [L242] Suggest "the important" rather than "this significant".
35. [L258] Does not go to 73N.
36. [L259] Don't use contractions in professional writing (yes, I know).
37. [L261] Suggest "explaining" to "to explain".
38. [L276] Is this J1 or J2 in Fig. 2?
39. [L279] "Relative" to what?
40. [L280-284] Suggest deleting to avoid pre-summarizing.
41. [L306-307] Equation 1 has nothing to do with geostrophic winds, so why are the balance concepts introduced here?
42. [L314] The diabatic heating term includes effects from radiation, turbulence, etc. Why would they necessarily not be important here?
43. [L315] Does using W as the PV "master" have an impact on the results?
44. [L317] Hypothesizing about the size of terms is not acceptable.
45. [L327 and 329] Why is there a mix between normalized and non-normalized PV error between Eq. 3 and Fig. 9? It would be much simpler to pick one measure and stick to it.
46. [L336] Is it not the absolute differences (PV*) that are of most interest here?
47. [L338] How does the "cyclonic pattern around the positive PV difference" "counteract the advection of the PV difference"? By definition, a circulation centred on a feature does not advect the feature.
48. [L378] The W cutoff is over North Africa. Could turbulent mixing not also be affecting PV at this level?
49. [L390] Suggest "than" to "as".
50. [L397-399] It may be true that most ensemble members have less precipitation than the short-range forecast; however, that is not a conclusion that can be drawn from the subensemble mean. Shifts in precipitation structure will lead to a "smearing" of the ensemble mean and a reduction in peak values. The ensemble mean forecast should not look like the QPE unless there is no uncertainty in the event.
51. [L404-406] Remove this sentence after plotting track only to the common final time.
52. [L404-409] Why is the within-cluster variability important for this discussion? Consider removing the members from Fig. 11 and associated discussion in the text.
53. [L421] This is a complicated description for an introduction. A simpler introduction here would be more effective, followed by the more thorough description on L450 (including citations).
54. [L438-439] Run-on sentence.
55. [L450] Remove extra comma.
56. [L458-459] This conclusion seems to be a tautology. The formation of an anticyclonic outflow layer is essential for any tropical cyclone (or medicane). This layer will necessarily have an anticyclonic PV anomaly. So the replacement of a cyclonic PV anomaly at upper levels with an anticyclonic one is simply a requirement for tropical transition that is not case-dependent.
57. [L460-479] This member-by-member analysis is not sufficiently in-depth to be highly valuable for readers; however, it takes up a lot of space in the text. I think that it would be fair to mention that there are some members whose evolution does not follow the story line, and that these members tend to be the ones that lie on the borders between the subensembles. If this discussion is to be kept, then I think that it should be made more robust and moved to the supplemental material for (very) keen readers.
58. [L482] Reverse "storm" and "internal".

59. [L484] It is not clear what the precipitation analysis has yielded in terms of results that feed into the study conclusions.
60. [Table 1] This is not a clear caption.
61. [L502] What “process” is being referred to here?
62. [L519] Is the Fig. 7 feature really “large-scale”? It is larger than convective, but nowhere near the Rossby radius.
63. [L520] This is the first appearance of the word “energy” in reference to the results of this study. Because the reader does not know the size of wind differences that contribute to the PV feature shown in Fig. 7, it is impossible for them to assess the size of differences compared to the background.
64. [L523-526] We also showed that cyclogenesis and TT can occur in the lee of the Alps during trough passage (shameless self-promotion of McTaggart-Cowan et al. 2010a and b, but they’re highly relevant to this discussion).
65. [L531] I do not believe that making “all data available from the authors upon request” is aligned with the FAIR data policy adopted by WCD.
66. [Fig 1] Suggest adding “at 6-hourly intervals in each panel” on the second line.
67. [Fig. 3] Explain the circle in panel (e) (not referenced until L418). Note that the reference vector is above panel (d). The last sentence seems to be missing. What are the white dots in the lower row?
68. [Fig. 6] Is panel (a) really needed? All of this information is available in earlier figures, and the grey boxes in this panel appear to have been subjectively determined. Note that there are no blue lines in panel (c) as implied by the caption.
69. [Fig. 7] What are the white areas? Adding normalized differences in the inset would be useful so that the reader has a better sense of how unusual the patterns over the Gulf of St. Lawrence really are. A cross-section across the jet core might be an interesting way to show the depth of this feature.
70. [Fig. 8] Consider whether the right-hand column of plots is really needed. These figures are complicated to interpret, barely referenced in the text, and do not appear to contribute to the discussion. Suggest adding a semicolon at the end of the first line of the caption.
71. [Fig. 10] Why are the blue contours discontinuous (especially in panels (j) and (k))?
72. [Fig. 11] Suggest adding row labels as “mean”, “largest” and “smallest”. Consider stopping the track in (d) at 0000 UTC 30 September as for the other panels. This will simplify interpretation, text and caption.
73. [Fig. 12] Consider adding a full legend with three primarily classifications (“before”, “between” and “after”), each with three sub-classes (“medicane”, “weak medicane” and “no medicane”) so that all plot symbols are shown in the legend. Does the marker size scale progressively with upper-level thermal wind strength? What is the no-medicane threshold?