Third Review of WCD-2022-6

The authors have made significant changes to the submission that have improved the presentation of the work. I particularly appreciate the fact references to the "cyclogenesis butterfly" have been removed and that the number of sections has been reduced to six. The manuscript is certainly converging towards a publishable form.

Reviewer: Ron McTaggart-Cowan

Recommendation: Minor Revisions

General Comments

1. Description of the plotting strategy within the main body of the text disrupts the flow and will distract future readers from the main points that the figures are meant to support. Leaving details of the plot description within the caption instead of the text is a stylistic decision; however, I believe that the paper would be more effective if this convention were followed throughout.
2. Starting on L204, the growth rate of spread is used as being analogous to forecast uncertainty. But one of the key findings of this paper is that the ensemble is conditionally over-dispersed. So then the events being studied are precisely those where the ensemble spread does *not* represent uncertainty well. It seems like a more careful use of the words "spread" and "uncertainty" is warranted throughout the text. This is particularly true in relation to the LGR, which I think is formulated to represent spread growth directly, and uncertainty growth only under the assumption that the two are interchangeable (which is shown not to be entirely true during cyclogenesis events).
3. I still find the methods used in section 4 to be too complicated for the relatively simple outcome. If the goal is to demonstrate the conditional over-dispersion of the ECMWF ensemble then please reconsider dramatically simplifying the analysis. If the goal is a demonstration that a complicated spread-error decomposition can be used, then please be sure to highlight the value added by the technique over a simpler analysis.
4. The use of different shading ranges for different terms of the same equation (top two rows of Figs. 5 and 7) and equivalent plots for different sensitivity tests (Figs. 9 and 10) make these comparable panels very difficult to compare. I understand that not much will show up on the panels with small ranges, but that's useful information that the reader should be able to determine at a glance, not by looking simultaneously at the plotted structures and the colour bars simultaneously for two panels.
5. It appears that only the final paragraph of the conclusions (four lines) could really be classified as discussions or conclusions (the title of section 6). This means that ~90% of the section is actually dedicated to a thorough summary, including numerous figure and table references. As someone who read through the full manuscript I find this redundancy a missed opportunity for opening up a broader discussion of the implications of the work. Please consider either renaming section 5 to "summary and conclusion" or (better) redrafting section 5 to present a very short summary before taking a larger perspective on discussing the work.

Specific Comments

1. [L40-58] I suggest redrafting this paragraph (and removing Fig. 1, with any needed panel combined with Fig. 2) for two reasons. One is that it is hard to imagine any reader of this work who wouldn't already been very familiar with cyclogenesis in the storm track. The other is that this study isn't really about cyclogenesis itself, but rather about the growth of uncertainties related to cyclogenesis. I understand that the reader needs to know what cyclogenesis is to appreciate how errors may grow, but I think that a well-crafted literature review would be more effective at relaying this than the current case study. Describing baroclinic instability and diabatic contributions to cyclone deepening could be done briefly with relevant citations. This would be followed by an introduction to error growth on the waveguide, for example citing the recent work of Baumgart et al. (2019).

2. [L61] I think that the word "coordinate" implies too much intention here, and feels like an anthropomorphization as a result. Or maybe it's the word "act to": cyclogenesis doesn't really "act", it just happens. A similar construction appears in the subsequent sentence.

3. [L126-127] Many readers will probably know what it means to "warm start" VarBC and SPPT: please provide a brief explanation because this presumably impacts early spread growth in the forecast.

4. [L137] Does the WCD style guide cover web references? If so, it will hopefully cover citation format and include information about access date.

5. [L177] This introductory sentence is written as if clustering is the only way (or even the most obvious way) to accomplish the objective of identifying cyclones. Given that other methods for cyclone identification have been used in the literature, please consider rewording this sentence to provide a stronger introduction to the need for a cluster analysis in this case.

6. [L210] The use of the thin overline here (ensemble mean) is distinct from the thick overline in Eq. 1 (time mean). I'm not sure that the typesetting is going to be clear enough to allow readers to distinguish between these two. Please consider using a different operator (for example <>) for the ensemble mean.

7. [L248-252] This discussion makes it sound like Z250 is used throughout the remainder of the text, but the subsequent section moves back to P315. Please clarify here which sections are forced to use Z250 because of TIGGE database limitations.

8. [L259] "Cyclogenesis deepening" seems redundant given the adopted definition of cyclogenesis (i.e. cyclone deepening).

9. [L265-284] I don't understand the "it is tempting to speculate" concept here, especially when the paragraph goes on to say that these hypotheses could be confirmed (or refuted) by investigating the terms on the r.h.s. of the LGR equation. If it is tempting and confirmable, then why isn't it done? Then all of this theorizing could be replaced by a simple plot that shows what process is occurring. Between that simplification and the reduction in plot strategy description (General Comment #1), a solid analysis could be included without increase in manuscript length. This would also provide justification for the existence of the r.h.s. of Eq. 3, which is not otherwise used in the manuscript as noted by both reviewers in previous rounds of review.

10. [L292-294] There is an odd asymmetry in this discussion. The "forecast bust" reference (behaviour of a forecasting system that is entirely a property of model space) seems out of place with this discussion of physical features and phenomena. Maybe pulling out the "forecast bust"

phrase and making a separate statement would help, because then it is clear that the bust can have its origins in any of the listed features (or others).

11. [L310-315] It is unfortunate that there is no quantification of this difference. It seems like there is enough information contained in the TIGGE database for a systematic assessment of LGR for Z250 during cyclogenesis cases from different models.

12. [Sections 4.3 and 5] Does "day—2" refer to "day minus 2"? If so, please replace the em-dash with a space and a minus sign, to become "day -2" for readability and to distinguish it from a compound adjective (e.g. "the day-2 spread").

13. [L510-511] Is there really a need for further investigation to figure out how to reduce spread generated by SVs or SPPT? The options seem pretty obvious. Perhaps this would be better worded as how to reduce the associated spread growth during cyclogenesis without negatively affecting the overall well-balanced spread-error relationship.

14. [L564] I'm not really sure what this sentence means (even what "other" refers to or is distinct from) and how it follows logically from the current work.

15. [Data availability statement] The "data and code available on request" doesn't really live up to FAIR principles. Please consider at least uploading as much of the code used to generate the results shown here as possible to a public repository.