

Dear reviewer,

We are grateful for the time that you put into evaluating our manuscript and for your critical comments, which are very valuable to improve the manuscript, in particular its readability. We are sorry to hear that the initial version was such a complex and dense read. Our general author comment (a separate document) provides an overview of the modifications that will be implemented in the revised version of our manuscript. We believe that the revisions have improved conceptual clarity. In this author comment we will address more specifically your major and conceptual comments. A more detailed point-by-point response will be given when submitting the revised version of the manuscript. Your comments are repeated in black while our comments are colored in blue.

Review of paper

Does the envelope of Rossby wave packets exhibit higher predictability than the underlying wave pattern?

by M. Riemer and L. Gölz

submitted to *Weather Clim. Dynam.*

This is in principle an interesting study, which investigates with a sophisticated methodological approach whether Rossby wave packets (RWPs) are better predicted than the underlying Rossby waves. The topic is of great interest to readers of WCD, and I think that after major revisions this can become a valuable contribution. The authors apply a complex set of diagnostics and honestly, I got lost several times in understanding what is done, and why, and how the results were actually obtained. Some analyses appear overly complicated and difficult to understand and should be considered to be removed from the paper or much better explained. In short, this is important work, but the paper requires major restructuring to better guide the reader.

Major comments

A) I found it challenging to understand the structure of the paper and in particular in Sect. 2 and 3 why certain concepts are introduced and how they help addressing the main question of the paper. Section 2 first led me into the climatology of RWPs ... then you introduce the envelope concept, then MSE and the idea of a saturation level, and then it felt like an abrupt jump in Sect. 3 to look at PV tendencies ... (and in L181 I wondered why this decomposition is needed – how does this contribute to answering the main questions was not clear to me!) ... before in Sect. 3.2.2 you go back to envelope errors. Could this be streamlined somehow? Maybe first discussing the envelope and envelope errors in general, and then discussing how this concept looks like in the PV framework? And explain why you need so much detail in the PV error

analysis. In the end you are not specifically interested in PV errors, right, but rather in envelope errors? To summarize this point, in Sect. 3.2.2 I felt rather lost. I tried to understand the Fourier notation and the different PV terms but had no clue how this will be related to the theme of the paper.

In the revised version, the structure will change substantially and will hopefully be much clearer. The revised version uses a standard definition of RWPs based on the wind field. This revision will make obsolete many of the discussions that prelude our main results in the original version. In addition, our notation will use the Hilbert transform explicitly, which greatly simplifies the derivation of the tendency equations for the envelope and phase. The PV framework will be introduced only after comparing the predictability of the envelope and phase and will be presented in less detail when we actually diagnose the error-growth dynamics.

- B) Figure 6 and L253: it is not clear to me how this was done: for all curves in Fig. 6, do you always consider the same grid points? And do you consider them only at times when there is a RWP or at all times? How do the curves relate to the mathematical terms introduced in the previous section? I am rather lost here how these curves were calculated. Maybe an example for one of the cases shown in Fig. 1 and 3 would help? Figure 6 leads to one of the main conclusions of the paper (L281) and therefore it is very important that the reader clearly understands how the different MSE curves were calculated. Related to this, I was completely confused in L283 (Fig. 7) because I thought that all the analysis before (Fig. 6) was done specifically for RWPs, but obviously not ...

Thank you very much for this comment, which makes us realize that the revised version of the manuscript needs to be much clearer about how we compute the MSE and the associated saturation level. In retrospective, it may seem clear that the brief introduction of the computation of the MSE and the associated saturation level in Sect. 2.24 in the original version did not suffice.

In our study, the MSE first implies a spatial average and then an average over many forecasts. The geographic region for the spatial average varies with time. In Fig. 6, the region is given by the values of the envelope in the analysis that exceed at that specific time the climatological median value. The purpose of this analysis is to focus on those regions, in which RWPs typically occur without the need to refer to specifically identified RWP objects. Object identification methods involve (a number of) free parameters and different RWP identification methods yield – albeit consistent in a general sense – to differences in the details of their respective RWP climatologies. Our original manuscript failed to make this motivation clear. Figure 7 then considers the region defined by the area of RWP objects identified in a previous study. Again, these regions for the spatial integration of the MSE vary with time. The saturation level is defined based on the same regions. In the revised version, we will extend and clarify the computation of the MSE and the saturation level along these lines.

- C) I could not understand what is shown in Fig. 8. What is the unit along the y-axis? Do these

curves relate somehow to the results shown previously in Fig. 6 or 7? I assume yes, but it is not clear to me how. Most likely references to some of the equations introduced in Sect. 2 and 3 could help. A related specific question: if the error growth curves in Fig. 8a and 8b are so similar (L308), why then are the blue and orange curves in Fig. 7a different?

Again, thank you for this comment, which makes clear to us that the introduction of the normalized error tendencies considered in Fig. 8 of the original manuscript was incomplete. In addition, there was a typo in the figure caption. The units along the y-axis should be s^{-1} . Our apologies!

The figure shows $\partial \text{MSE} / \partial t / \text{MSE}$. If the MSE grew exponentially, i.e. if

$$\partial \text{MSE} / \partial t = \text{MSE}_0 \exp(\alpha t)$$

then

$$\partial \text{MSE} / \partial t / \text{MSE} = \alpha,$$

the growth rate. In other words, the figure shows the inverse of the time by which errors grow by a factor of e (roughly: the inverse of the error doubling time).

We introduced this metric in the original version to compare the *relative* importance of different error-growth mechanisms between the envelope and the underlying wave pattern, i.e., to investigate if there are qualitative differences in the error growth mechanisms. In the original version, we did not find such qualitative differences. To your specific question: Our interpretation was that the error growth dynamic is qualitatively similar and further, unfortunately tacitly, assumed that errors merely grow somewhat faster for the envelope.

In the revised version, error growth dynamics of the envelope and the phase do exhibit qualitative differences (see general author comment). In addition, there are now also clear quantitative differences. The revised version will introduce the growth rate more formally following the discussion above. In addition, the revised version will discuss the relation of the growth rate of the MSE to that of the saturation fraction. In theory, they are the same (for a saturation level that is independent of time). In practice, however, the relation holds only in an approximate sense due to inherent problems in closing the budget for an enstrophy-like variable (such as the MSE of PV) from numerical data. Besides the typical numerical inaccuracies, the omission of diabatic terms from parameterization schemes, etc., the fundamental issue is that numerical diffusion is a non-negligible sink of enstrophy, which cannot be accounted for with the output from numerical models (discussed in some detail in Baumgart et al. 2018).

Minor comments

We will respond to these comments when we submit the revised version of the manuscript.

- 1) L1: “long-standing hypothesis”: I was not aware of the quote by Lee and Held (1993), which indeed is a remarkably clear statement. However, to call something a “long-standing” hypothesis, I would expect it to appear more frequently in the literature. Maybe “a hypothesis postulated in the 1990s” would be more appropriate?
- 2) L4: try to simplify the abstract, e.g., “envelope” has not been introduced and is not clear here.
- 3) L8: I struggle with “the envelope is the wave”. I understand that both have similar error dynamics but still the envelope is a reduced concept of the entire wave.

- 4) L9: sounds obvious to me – what should have been the “large-scale dynamics” that determines the evolution of the RWP independent from its “underlying wavenumbers”?
- 5) L11: why are your results rendering the concept of large-scale RWPs a “misconception”?
- 6) Fig. 1 caption and elsewhere: write dates in the form “at 12 UTC on 6-9 August YEAR”.
- 7) Fig. 1 caption and elsewhere: refer to panels simply as, e.g., “a” not “a”).
- 8) Fig. 1 caption: not clear where these wind speed values are valid; at a certain pressure level?
- 9) L42: should read “indicated”.
- 10) L42: I don’t find this entirely convincing, changing the amplitude and/or shape of troughs and ridges in most cases directly modifies the shape of the diagnosed RWP – the existence and overall size might not be strongly sensitive to individual troughs and ridges, but the detailed shape (as shown in Fig. 1) should be. In other words, I am not too surprised that the hypothesis did not verify – but of course this does not downplay the value of this study (which is the first to quantitatively test the hypothesis), it just reduces my “degree of surprise”.
- 11) L53: should read “their Fig. 2”.
- 12) L56: what do you mean by “negates”?
- 13) L59: incorrect grammar, maybe “signify”?
- 14) L63: should read “... and Baumgart” instead of “...; Baumgart”.
- 15) L65: what are “piecewise PV tendencies”? (your 2016 paper does not use this term)
- 16) L81/82: should read “reanalyses”.
- 17) Fig. 3: I suggest including the symbol “E” in the caption, and to overlay the 20 m s^{-1} contour from Fig. 1 (this would help the comparison).
- 18) Fig. 4: which latitude band is considered here? And are grid point values weighted with latitude?
- 19) Fig. 4: do you understand why values of E are larger in summer than in winter, both inside and outside of RWPs?
- 20) L178: I don’t understand the logic here. You make the point that advection dominates the growth of PV errors, but what is the connection then to initial condition errors? And does this mean that you totally ignore the potential diabatic contributions to PV error growth, or do you capture some of them indirectly via advection?
- 21) L179: relative to what?
- 22) L239: I don’t understand the SUM+/-k notation.
- 23) L249: can you make this more explicit / intuitive to the reader: how will you apply these equations and what exactly can be learned from them?

- 24) L322: I could not really follow why this is the appropriate interpretation of the results shown.
If the envelope was the wave, the whole question of the paper would not make much sense(?).
Do you rather mean “errors in the wave directly affect errors in the envelope”?
- 25) L330: should read “These authors found ...” and “presented” in L333
- 26) L337: I don’t understand “outside of identified RWP objects where the envelope field still exceeds its median value”
- 27) L340: are you comparing here predictability inside and outside of RWP objects on the same days, or do you compare predictability on days with vs. without RWP?
- 28) L345-L358: here I was again lost with understanding what you are doing
- 29) L360: I assume that large amplitude PV anomalies occur in preferred regions (since PV variability is not the same everywhere) and I therefore wonder whether your result about predictability and the amplitude of PV anomalies says something about having higher predictability in some regions (where PV variability is larger) and smaller predictability in other regions (where no large amplitude PV anomalies occur) or do make a statement about predictability in a specific region for flow situations with small vs. large anomalies? If the latter is true, then you would say that flow situations close to climatology have the lowest predictability?
- 30) L371: should read “suggested”
- 31) L385: I wonder whether some of the conclusions here were not clear already before doing this study. For instance, “phase errors in the underlying wave pattern thereby translate into envelope errors”, isn’t this clear from basic considerations as shown in Fig. 9? Isn’t the main result that this “effect” is so strong that you get the same error growth for the envelope and the wave pattern? Similarly, “the envelope is a direct function of its underlying wavenumbers”: of course!
- 32) L401: These conclusions should be reformulated such that the reader understands the inside/outside concept (see comments above).