

Dear Maarten Ambaum, dear anonymous reviewers,

We are very grateful for your careful reading of our manuscript and your thoughtful, critical comments. Your comments are very valuable to improve the manuscript, in particular in terms of conceptual clarity and readability. We will modify several aspects of our study following your suggestions, critiques, and comments. This author comment summarizes our revisions and provides an overview of the revised key results and interpretations. Your individual major and conceptual comments will be addressed in separate author comments. A more detailed point-by-point response will be given when submitting the revised version of the manuscript.

1) Following Maarten Ambaum's suggestion, the revised version will compare the predictability of the envelope with that of the phase of the wave packet, instead of the wavenumber-filtered anomalies, termed 'underlying wave pattern' in the original version. We will thereby follow the common approach to consider wave packets by the complex *analytical signal* Z ,

$$Z(x, t) = v_{\perp}(x, t) + i\mathcal{H}_x[v_{\perp}(x, t)]$$

where $v_{\perp}$ is the wavenumber-filtered wave signal and $\mathcal{H}_x$ is the Hilbert transform. From the analytical signal Z , the envelope E and the phase φ are computed by $E = |Z|$ and $\varphi = \arg(Z)$. We will revise Fig. 1 of our initial submission to illustrate both the envelope and phase evolution of an RWP. A draft version of two subpanels of this revised figure is presented herein as Fig. R1.

This revision adds conceptual clarity, avoids confusing notation, and simplifies the derivation of (error) tendency equations. In particular, the revision addresses comments regarding the (confusing) role of the phase of individual (monochromatic) waves and will make obsolete Fig. 9 (disputed by reviewer #3, their comment 4) and our catch phrase "The envelope is the wave." (disputed by Maarten Ambaum (his comment on L8) and by reviewer #2 (their comment 3)).

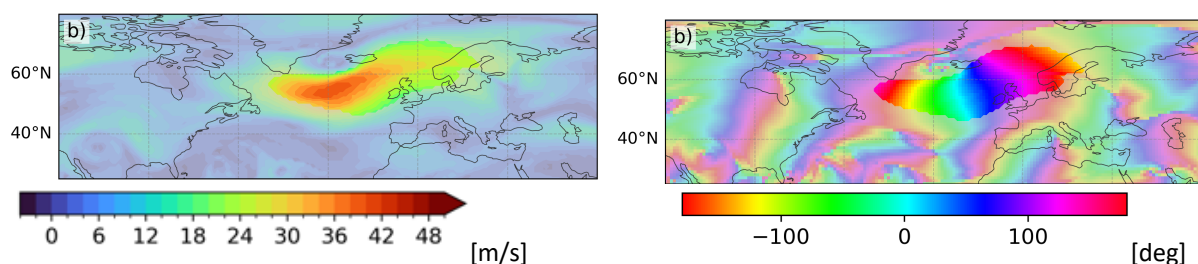


Figure R1: Draft of a revised Fig. 1 of the manuscript to illustrate the fields of the two main variables considered in the revised version: the envelope E (left panel) and the phase φ (right panel) of a wave packet, as defined in the text above. The bright colors indicate the region identified as an RWP object in previous work.

We argue that considering the RWP's phase addresses the central hypothesis that "the packet envelope should be more predictable than the individual weather systems", at least to the extent that the individual troughs and ridges of the underlying wave pattern (technically more correct: the amplitude-modulated carrier wave) constitute the individual weather systems. Maarten Ambaum has indicated this argument in one of his comments. For illustration, consider the schematic Fig. R2. An individual trough is determined by the envelope and phase information of the RWP. Differences in the predictability of individual troughs/ridges and of the envelope thus arise due to differences in the predictability of the envelope and the phase. The argument carries over to the ensemble of all troughs and ridges within the wave packet. Our analysis considers the statistical mean over many RWPs, i.e., over many envelopes and

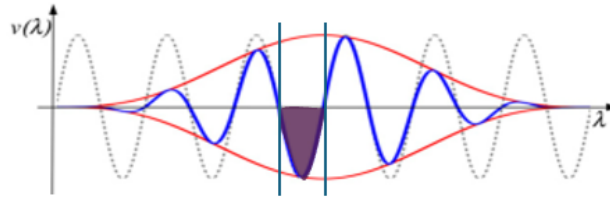


Figure R2: A schematic of a wave packet with its envelope (red), amplitude-modulated carrier wave (blue), and two phase lines of the carrier wave indicated by solid black lines. An individual trough (indicated by violet shading) combines and is determined by phase and envelope (= amplitude) information.

many (ensembles of) individual troughs and ridges. In this mean sense, we compare the predictability of individual troughs/ridges within RWPs and the predictability of the RWP envelope.

2) Addressing an important comment by reviewer #3, the revised version will follow the more common approach to consider RWPs as features of the wind field and with their envelope (and phase) determined by a state-of-the-art method following streamlines of a background flow, with the wind perpendicular to these streamlines as the wave signal. Because this identification method is computationally rather demanding for the large data that we investigate, we use the envelope and phase fields that have previously been computed by Prestel-Kupferer et al. (2024) based on NOAA GEFS reforecasts, which will now be our main data instead of ERA5 reforecasts. The original version already used the RWP objects that have been identified by Prestel-Kupferer et al. (2024) based NOAA GEFS data.

Employing this much more common definition of RWPs makes obsolete the introduction of the PV framework and the discussion that justifies the use of PV envelopes early in the original manuscript, leading to a clearer structure of the material, a main concern of reviewer #2.

3) Analyzing the NOAA GEFS wind data, we realized that our original computation of the saturation level lacked bias correction, which turned out to be crucial for the GEFS wind forecasts. To avoid difficulties in determining biases of the phase, which is a circular variable, and of our conditional verification of RWPs, the saturation level in the revised version will be determined empirically by verifying forecasts with analyses that are essentially random draws from climatology. The resulting mean error sets an empirical upper limit for the climatological error, i.e., the saturation level. To account for seasonality, we use analyses from the same calendar day as the forecast, but of different years.

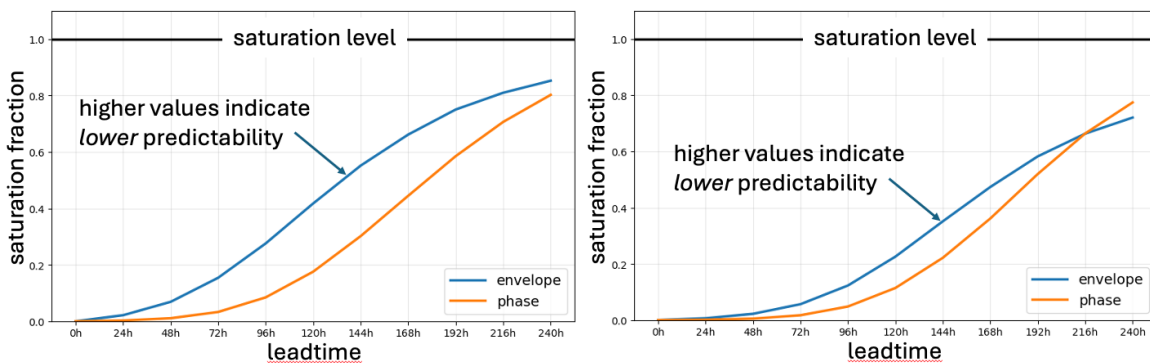


Figure R3: Revised version of Fig. 6b and Fig. 7a in the original manuscript. Comparison of the saturation fraction (our metric for predictability) of the envelope (blue) and the phase (orange) as a function of lead time for regions in which the envelope values exceed the median (left panel) and within RWP objects identified in previous work (right panel). The differences between the curves are highly statistically significant based on a two-sided, paired t-test with p-values below 10^{-6} , except for when the curves cross in the right panel at 216h.

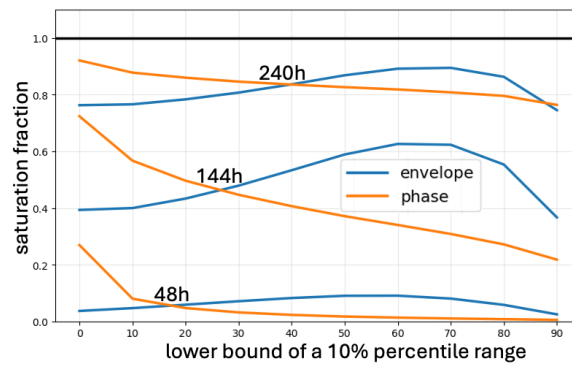


Figure R4: Saturation fraction of envelope (blue) and phase (orange) as a function of the envelope percentile range (10% ranges starting at the value given on the x axis) for three selected lead times (see labels in figure).

The main result of the revised version is here presented in Fig. R3, which can be compared to Fig. 6 and 7 of the original manuscript. We now omit comparison with smaller- and larger-scale fields, because we use wave-packet characteristics pre-computed in previous work and the original wind field is no longer readily available to us. For regions where the envelope value exceeds the median (left panel in Fig. R3), the envelope exhibits less predictability than the phase for all lead times, consistent with our original result. Focusing on the region within RWP objects (right panel in Fig. R3), the envelope exhibits less predictability than the phase only up to a (relatively high) saturation fraction of around 0.65. For higher saturation fractions at longer lead times, the envelope does exhibit higher predictability. This result is robustly seen in all individual seasons (not shown), with small variation of the saturation fraction (by ± 0.03) and lead time (by ± 12 h) at which the envelope becomes more predictable. Higher predictability of the envelope at long lead times was not found in our original results. Arguably, this difference arises because the state-of-the-art method identifies RWPs as a higher-amplitude anomaly (see below) than a wavenumber filter along latitude bands used in the original version, as argued by reviewer #3 in their comment 1).

To better understand the revised result, we investigate predictability as a function of wave amplitude, here in terms of percentile ranges of the envelope value (depicted for lead times of 48h, 144h, and 240h in Fig. R4). Phase predictability increases with wave amplitude and is consistently less than envelope predictability for the *lowest* percentile ranges. Envelope predictability is lowest for percentile ranges that are closest to the climatological mean value (40% - 80%; these percentile ranges do not center on the median because distribution of envelope values has a long right tail (cf. Fig. R5, left panel)). Consistent with our original result (the original Sect. 4.3), we find that anomalies with large deviation from climatology exhibit higher predictability than anomalies with small deviations. The mean envelope values within RWP objects correspond to the 95th percentile of envelope values in our data.

The above paragraph considers the climatological distribution of the full data (25N – 80N). We now consider the climatology within the region where RWP objects occur. Each RWP objects defines a certain geographic region, for which the climatological distributions of envelope and phase values for the calendar day on which the RWP occurred is determined. Aggregating these climatological distributions for all RWP object defines the *climatology within RWPs* (Fig. R5, left panel, blue distribution). The actual envelope values within RWP objects (orange distribution) still constitute a large anomaly relative to this *climatology within RWPs*, with the mean of this (orange) distribution corresponding approx. to the 85th percentile of the

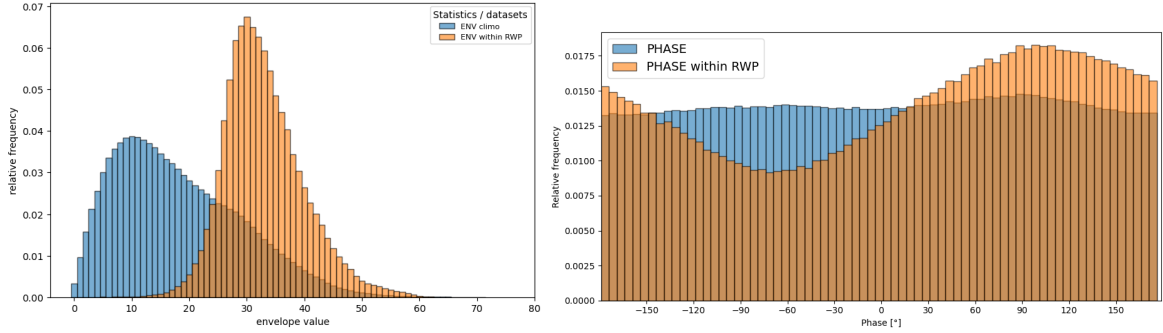


Figure R5: The climatological distribution (blue) and the distribution of actual values (orange) within RWPs for the envelope (left) and phase (right).

climatology within RWPs. The actual phase values within RWPs evidently do not constitute such a strong anomaly (right panel in Fig. R5). The revised version will use the *relative entropy* (e.g., delSole 2004) to quantify this (stark) contrast. We argue that at longer lead times, when skill has substantially decreased, the high-amplitude anomaly of the envelope exhibits a higher predictive signal than the phase, which deviates only weakly from climatology. “Low-skill” forecasts may thus better predict the RWP’s envelope than its phase.

By contrast, at short lead times, when skill is generally high, forecasts of the phase provide more information with respect to the climatological distribution than forecasts of the envelope. The revised version will illustrate “high-skill” forecasts by considering forecasts at short lead times, say 48h, and use the MSE at that lead time to estimate the width of a forecast distribution. Again, we will use relative entropy as a metric to make a quantitative argument. This line of argument qualitatively explains our main result: the evolution with lead time of the saturation fraction depicted in the right panel of Fig. R3.

While illustrative, our argument for the behavior at short lead times in the paragraph above merely applies an alternative definition of predictability: the added information of the forecast (distribution) relative to climatology (e.g. delSole 2004). The remaining question is *why* the forecast distribution of the phase is sharper than that of the envelope when skill is relatively high. In other words: Why do phase errors grow more slowly than envelope errors? To address this question, we will finally consider the error-growth dynamics of envelope and phase errors in the PV framework. As in the original version, we here use what reviewer #3 termed “an approximation to the RWP envelope”, i.e., a wavenumber filter in the *zonal* direction and PV as the variable to define RWP envelope and phase. We use this approximation, because we are not aware of an extension of the state-of-the-art identification method of RWPs to PV anomalies along undulating background PV contours. At this point we consider this approximation to be justified to illustrate basic characteristics of the error-growth dynamics.

The results are presented in Fig. R6. We provide two versions. One version considers the average over the region of RWP objects (upper panels). Because the agreement between the PV envelope and the wind envelope from which the RWP objects are derived is good only in the qualitative sense (as illustrated in the original manuscript), the region of RWP objects does not only contain high values of the PV envelope. Based on Fig. R4, PV error-growth characteristics can also be expected to vary with the value of the PV envelope. A second version thus considers the average only over high values of the PV envelope (90th percentile, lower panels). For this version, however, it is not guaranteed that all situations are associated

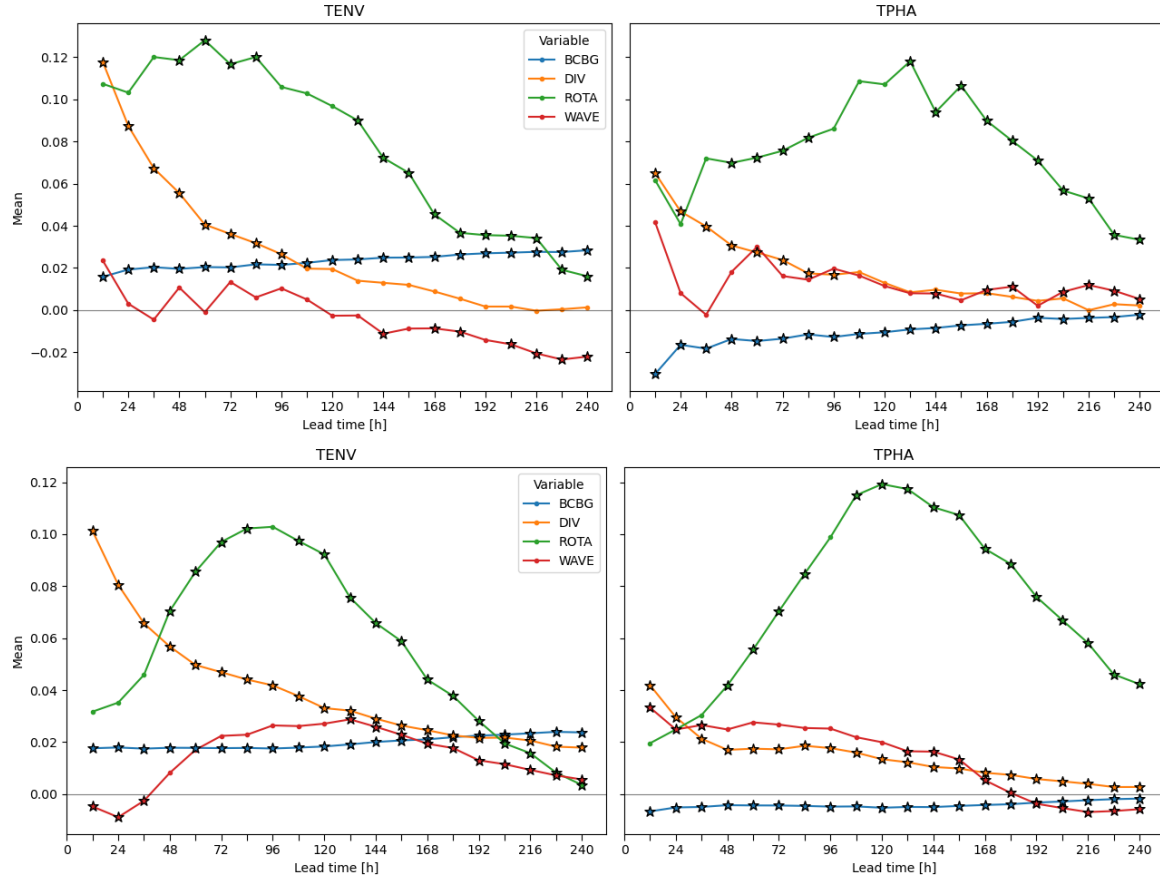


Figure R6: Error growth rates associated with individual error-growth mechanisms for envelope (left panels) and phase (right panels). The upper panels depict values averaged over the region of RWP objects and the lower panels values averaged over the region where the PV envelope exceeds the 90th percentile. The upper left panel can be compared to Fig. 8a in original manuscript. Statistical significance of differences between envelope and phase is indicated by stars, assessed by a two-sided, paired t-test with 5% confidence level.

with RWP objects. Consistent with Fig. 4, we find generally higher growth rates (at short lead times) within RWPs.

The upper-left panel corresponds to Fig. 8a in the original manuscript¹. We here focus on the divergent and nonlinear terms (denoted as DIV and ROTA in the figure). For these terms, differences between the two versions are most prominent in the first 3-4 days. Importantly, there are systematic differences between the error-growth dynamics of the envelope and the phase. These differences are consistent in the two versions. Arguably, these differences are blurred in the original manuscript for the underlying wave pattern, which can be considered a combination of both phase and amplitude information. We here combine the two “wave” terms of the original version into a single wave term, which we ignore for the purpose of this author comment. At no time does this term make a dominant contribution. In addition, the behavior of this term is not consistent in the two versions presented in Fig. R6.

¹ While very similar, the two panels are not identical. The reasons are: i) The original version used the average MSE of all years for each month, whereas the revised version considers each year separately. This change was implemented to provide more data for the statistical significance test. ii) A bug in aggregating the MSE over the years in the original version resulted in the use of approx. only half the number of years, leading to an overestimation of the growth rates by a factor of approx. 2. Note that we here present growth rates in 1/day, whereas the original version presented them in 1/s. iii) Due to a further bug, the original version averaged over the considered three theta levels first and then computed growth rates, whereas we now compute growth rates first and then average the results.

Relevant for our “high-skill” scenario above, envelope errors are substantially more affected by nonlinearity and the divergent flow than phase errors in the first 3-4 days of lead time. Divergent outflow has been linked in previous studies to both wave amplification and error growth in general. It thus seems plausible that the associated error-growth mechanism predominantly projects on amplitude rather than phase errors, consistent with our discussion in the original version of the manuscript. The impact of nonlinearities is counter to the expectations of our original manuscript. Reviewer #3 provides an insightful comment in this regard. The evolution of RWPs is associated with wave dispersion. While error growth in linear dispersion can be expected to be small, nonlinear dynamics that (partly) disrupt wave dispersion are (more) prone to error growth. Partial wave breaking, for example, strongly impacts on the amplitude of the remaining wave. Our interpretation in the revised version will follow this insightful reviewer comment. It is interesting to note that after day 5 phase errors are more strongly impacted by nonlinearities than envelope errors. Finally, the baroclinic term follows the expectations of the original manuscript and is small (even negative) for phase errors while persistently amplifying envelope errors.

Summary of revised version

Our revised results show that – for scales and wave amplitudes relevant to RWPs – the envelope does not, in general, exhibit higher predictability than the phase. For RWP objects identified by state-of-the-art methods, however, the envelope remains more predictable at longer lead times when skill is generally low. This predictability arises from the large anomaly with respect to climatology exhibited by the envelope, whereas the phase does not exhibit such an anomaly. By contrast, when skill is relatively high at earlier lead times, the envelope is more strongly affected than the phase by nonlinear and divergent error growth, the two dominating mechanisms for general error growth identified in previous work. Our results thereby provide a nuance answer to the question posed in the title of the manuscript. The argument put forth with the schematic Fig. R2 relates our results to the hypothesis by Lee and Held (1993).

References:

delSole, T. (2004). Predictability and information theory. Part I: Measures of predictability. *Journal of the atmospheric sciences*, 61(20), 2425-2440.