**Final author comments for egusphere-2026-4755**

**Extreme hazard cyclones – concept and ERA5 climatology**

by Sven Voigt, Hanin Binder, Seraphine Hauser, and Heini Wernli

We are very grateful to both reviewers for their constructive and thoughtful feedback, which helps us further improve our manuscript. Based on their suggestions, we will implement the following major changes in the manuscript:

- We will expand the Introduction to explicitly cover tropical cyclones and extratropical transitions to clarify that our framework is applied globally to all trackable cyclones, and add a discussion in the Conclusions on the implications of combining these different cyclone types in the same analysis.
- We will test the sensitivity of our rankings to the definition of the cyclone mask by repeating the extreme event attribution using fixed circular radii (500 km and 1000 km) around the cyclone center.
- We will test a stricter attribution criterion for 72 h extremes by requiring a minimum overlap of 24 h between the cyclone mask and the extreme event location.
- We will add time-lagged composites analogous to Figs. 7 and 8 to examine the temporal evolution of the cyclones and the timing of extreme events across the four hazard categories.

Below, we provide our responses to the individual comments and outline how we plan to revise the manuscript accordingly. The reviewers' comments are in blue and our replies in **black**.

## Reviewer 1

The study on "Extreme hazard cyclones" by Voigt et al. mainly addresses the question, in which way the most intense or hazardous cyclones differ in typical properties like wind and precipitation intensity, extent, lifetime, (thermo-)dynamical structure and others compared to "normal" cyclones which they refer to as climatology. Basis to approach this is the analysis of the global ERA5 reanalysis with some deeper investigations over the North Atlantic. The paper is clearly structured, and the results are supported by the used figures and tables. The topic is highly relevant for various applications and for a better theoretical understanding what makes a cyclone more hazardous as others, for example to improve forecasts, and fells within the scope of WCD. I therefore suggest a publication after what I consider as minor revisions, as most of my comments related to a better wording and description. Please see my detailed comments below.

We thank the reviewer for their positive evaluation of our manuscript, the constructive feedback, and their suggestions for improvements. Please find a first response to the comments below.

### General comments:

1. At several points within the text, I highly encourage the authors to sharpen the text (see also details below) regarding method description or what is actually meant. This would generally help the reader which is probably not that deep into this topic to better understand what is done and

what do the authors want to say, which will also improve the readability. Currently, the reader needs more than usual comprehensive background knowledge regarding methods and cyclones.

Thank you for the comment and for pinpointing specific unclear passages in the detailed comments below. We will review and sharpen the text where appropriate and needed, particularly in the description of the methods.

2. Abstract and Introduction suggest that there will be a profound comparison between the top hazard storms and the "normal" storms, referred to as climatology. Beside the analyzed properties in Fig. 5, I do not see this kind of comparison. Most of the time, the authors just present the results of the top storms. You could add such analysis to a supplementary, for example.

We agree that the Abstract and Introduction emphasize the comparison with climatological cyclones, whereas much of the manuscript contrasts the topmost ranked cyclones across the four hazard categories, and we will adjust Abstract and Introduction accordingly. At the same time, we would like to note that beyond Fig. 5, comparisons of the top 100 cyclones to climatology or lower-ranked systems also include the seasonality of cyclone-related extreme events across all tracked cyclones in Fig. 4, the background PV anomalies in Fig. 8, and the cyclone phase space climatology in Fig. 9.

3. I have doubts that mixing tropical and extratropical cyclones in this sense is helpful. In my opinion, the tropical ones will always dominate the heavy-tail of the statistics of wind and precipitation. Maybe it would be helpful to either focus on extratropicals, only, as partly done in the focus section on the North Atlantic, or to do the analysis separately for tropicals and extratropicals.

We view the ability to evaluate all tracked cyclones within a single unified framework as a strength of our approach. From a surface-hazard perspective, a location impacted by extreme winds or precipitation experiences that hazard regardless of whether the linked system is tropical, extratropical, or undergoing extratropical transition.

Furthermore, while tropical cyclones indeed dominate the tail of the statistics for the four hazard categories on a global scale, they do not universally dominate across all regions or hazard types, and extratropical cyclones regularly appear in the highest-ranked systems on a regional scale. For instance, as shown in Fig. 5 for the North Atlantic, systems with extratropical origin account for a large share of the top 100 cyclones, particularly in the short- and long-duration wind categories (W1- and W72-tracks).

We will expand the Introduction to explicitly cover tropical cyclones and extratropical transition, frame the study as a global all-cyclone hazard classification, and discuss the physical implications of combining different cyclone types in the Conclusions (see also our response to Reviewer 2, Major Comment 1).

**Detailed comments:**

L.12: what do you exactly mean by "North Atlantic storm tracks"?

We refer here to the extratropical North Atlantic as shown in the dashed black contours in Fig. 4a. In this context, “extratropical North Atlantic” is a better fit and we will update this phrase accordingly.

L.28: "over continents and within storm tracks" This sounds like the tracks are not over the continents. Generally, you should once define what you mean by "storm tracks".
Thanks for pointing this out. We here refer to the main climatological storm track regions, which are mainly over the ocean. But indeed, the storm tracks in the North Atlantic and North Pacific may also extend over the continents. We will update this and add a definition of the storm tracks.

L.31: how do you define "intensity" of a cyclone? Is it minimum SLP?
Here, “intensity” is meant in a more general sense rather than exclusively minimum SLP. In the papers cited in this phrase specifically, Field and Wood (2007) and Pfahl and Sprenger (2016) measure cyclone intensity via (near-) surface wind speed, whereas Rudeva and Gulev (2011) use cyclone depth (the SLP difference between the cyclone center and its edge). Cyclone area averaged relative vorticity would be another metric commonly used. All these metrics for cyclone intensity are meant to be combined in that phrase.

L.35: again, what do you mean by "storm tracks". Is it the center line of the pressure minimum, or a certain area around, or . . . .
Here, “storm tracks” refers to the climatological mid-latitude regions of high extratropical cyclone frequency in the North Atlantic, North Pacific, and around Antarctica. We will state this more precisely in the revised version.

L.72: "number of surface extremes associated with each cyclone along its track" I only understand this after I had read the entire paper, maybe already state here how you count extremes?
Thank you for the suggestion, we will simplify this phrase as follows: “We identify local extremes of precipitation and 10 m wind speed at individual grid points and count the total number of unique grid points experiencing an extreme within a cyclone's area over its lifetime.”

L.73: "large number", also here. Intuitively one cyclone has extreme precipitation and this would be one extreme, but you count the grid points exceeding a threshold, so probably rephrase this or define already here.
Thank you for highlighting this. With the rephrasing in L. 72 described above, the grid-point-based counting should be clear, and we will also adjust L. 73 accordingly (e.g., referring to "a large number of grid point extremes") to ensure consistency.

L.103: Why March 1950 to March 2025? ERA5 start in 1940 and depending on when you started the analysis, why not full 2025?

We chose 1950 as the starting year because observation density and reanalysis quality are lower during the 1940s (Soci et al., 2024). Starting the record in March 1950 and ending it in March 2025 avoids splitting DJF-seasons and provides a record with an equal number of all four seasons.

L.104: Did you do the interpolation? Can you provide some more details on the method you used? Furthermore, why did you choose a 0.5° grid. The original resolution is roughly 0.3°

The ERA5 data can be downloaded interpolated on a 0.5° grid directly from the Copernicus Climate Data Store. We used a 0.5° grid instead of the roughly 0.3° due to storage reasons. When looking at large-scale features like we do in this study, we expect 0.5° resolution to be sufficient.

L.110ff: The cyclone area from this approach can be quite large, right? Which can also be seen in Fig.1. Are there any further restrictions to limit the extent?

Yes, the length of the outermost closed contour is constrained to a minimum of 100 km and a maximum of 7500 km (Sprenger et al., 2017). We will add this to the revised manuscript.

L.124: the 99.9the percentile is quite high, which limits the data amount in an only 75-year-long record. Can you provide some words as justification why you used this high percentile and not a lower one (95, 99, 99.5)?

Because we use hourly time steps, lower percentiles correspond to what we regard as relatively frequent events: exceedances of the 95th, 99th, and 99.5th percentiles occur on average roughly once every 20 h, 4 d, and 8 d, respectively. We do not regard an event occurring every week as particularly “extreme”. By contrast, the 99.9th percentile corresponds to an average exceedance frequency of once every 1000 h (6 weeks), while the 75-year hourly record still yields up to 657 exceedances per grid point (except in very arid or polar regions where there are much less time steps with precipitation).

L.127: do you use a running 72h window on the hourly data or is it blockwise meaning e.g. day 1-3, day 4-6, etc.?

We use a running 72 h window shifted in 1-h steps (i.e., starting at 00:00 UTC, 01:00 UTC, and so on), which is also why consecutive 72-h extremes can partially overlap.

L.127: So you first do the averaging or accumulation and from this new values you calculate the 99.9th percentile as basis for your analysis?

Yes, this is correct.

L.146: Again, what do you mean by "intensity"?

We here mean intensity in terms of minimum SLP. This will be clarified.

Sect. 3.1. It would be nice to see an example (case study) for a 72h case (e.g. in a supplement) to get a better insight to this part of the analysis

Thank you for the suggestion. We will consider adding a case study in a supplement.

Fig.2: Especially in g-i it is hard to distinguish between the land contour and the storm track as both a black solid outlines. Could you try, e.g. gray land contours or something else that better separates the two things?
We will adjust this by using gray coastlines or shading land areas as in Fig. 2a.

L.218ff: So did I understand this correctly: you have 4 categories with each 100 storms, which is a total of 400. Removing duplicates, 274 individual storms remain, and out of these 188 are then only attributed to a single category?
Yes, this is correct. Because individual cyclones can rank in the top 100 for more than one category (and thus appear in multiple panels, such as both Fig. 4a and Fig. 4b), the 400 cyclones across the four top 100 correspond to 274 unique cyclones, of which 188 are only ranked in the top 100 in one category and 86 belong to two or more.

L.221ff: The given TCs: Do they remain TCs throughout their entire lifetime or do some of them undergo an ET?
All six cyclones named here undergo extratropical transition toward the end of their life cycle.

L.225: Do you mean, there is only one cyclone assigned to the combination of P1 and W72 or do you mean one cyclone for each category?
We mean the former: there is only one cyclone that ranks in the top 100 for both P1 and W72 simultaneously. This will be rephrased to make it unambiguous (e.g., “only one cyclone ranks in the top 100 in both the P1 and W72 categories”).

L.271: Only one time step seem rather low for me to take a storm into account for the NA region. How often does this low threshold really occur, and wouldn't it be more robust to define a combination of a certain storm area for at least one time step for consideration?
In the data, less than 2% of the selected North Atlantic cyclones are inside the domain for just a single hourly time step. We keep this one-time-step threshold to maintain consistency with the event-attribution method: because a grid-point extreme is assigned to a cyclone whenever it is covered by the cyclone mask for at least one time step, a cyclone only entering the domain for a short period can still be linked to an extreme. A stricter threshold could lead to a mismatch between cyclones in the domain and the cyclones linked to extremes.

L.290: "larger deepening rates than climatology" - do you mean mslp? Should be stated or defined somewhere. It would also be nice to have some exemplary numbers of the rates, so what actually is larger than climatology?
We here mean that extreme cyclones have larger Bergeron values than climatological cyclones. As also suggested by Reviewer 2, the definition of Bergeron will be included here.

L.305-308: Description of an additional metric - isn't this better placed in the method section?
Thank you for the suggestion, we will consider moving this phrase to Sect. 2.

L.312/313: "the majority of extremes occur during intensification phase" - where can I see that/ from which result stems this?
Thank you for catching this, that statement is indeed not directly backed by one of our figures. In the revised version, we will address this either by adding a supplementary figure showing the timing of extreme occurrences relative to the cyclone life cycle or by supporting this statement with relevant work in the literature.

L.325: "at least parts if the cyclone" - can you specify this more? Is it a certain percentage?
We did not apply a specific percentage threshold. This criterion requires that at least one grid point of the cyclone mask overlaps with the North Atlantic domain at that time step. Because this is evaluated at the time step when the cyclone is linked to the largest number of extremes—at which point the cyclone centers are almost exclusively located well within the domain (as shown in the insets of Fig. 6)—large parts of the corresponding cyclone masks lie inside the domain. We will specify this in the revised manuscript.

L.330: "95th percentile" - so you define extremes by using 99.9%, and now you look at the "ensemble statistics" and analyze the 95%, right? So these are the extremes of the extremes?
Yes, this is correct for the composites showing precipitation. The mean is shown for wind speeds.

L.340: 22m/s - is it also the 95th percentile?
No, this value refers to the mean wind speed. We will make sure this is explicitly stated.

Fig.6: I think, parts of the caption are better placed within the text, especially the parts linked to the method of selection. Furthermore, you use here now a threshold of 10%, which is fairly low in my opinion. In the main text you discuss the 95th percentile, so it gets a bit confusing now.
We agree that the composite methodology, including a clear explanation of which subfigure shows which variable at which percentile, could be described more thoroughly in the main text.
Regarding the 10% threshold, this relatively low value reflects the spatial variability in where localized extremes occur relative to the cyclone center across individual cyclones. Notably, Reviewer 2 suggested rotating the composites relative to the cyclone propagation direction. We will test this, and if rotating the fields increases the spatial alignment of the extremes, we will adapt the threshold accordingly. Finally, to avoid any confusion in the main text, we will explicitly state whenever we refer to the 95th percentile (for precipitation) versus the mean (for 10 m wind speed).

Fig.7: "zonal vertical cross-section" - is this through the storm center?
Yes, the cross sections are calculated through the center of each cyclone. We will add this to the text.

Fig.9: From the caption, I don't really get what is shown in (a, b) and what is then the difference to (c, d). Can this be improved?
The two rows display the two projections of the cyclone phase space: the top row (a, b) shows thermal asymmetry (B) versus lower-tropospheric thermal wind (V^{L}_{T}), whereas the bottom

row (c, d) shows upper-tropospheric thermal wind (V^{U}_{T}) versus lower-tropospheric thermal wind (V^{L}_{T}). Across both projections, the left column (a, c) shows 1 h precipitation extremes (P1) in shading, and the right column (b, d) shows 1 h wind extremes (W1). We will revise the caption of Fig. 9 and we will also correct the typographical error where the subscript in V^{L}_{T} is not properly subscripted.

L.488: In my opinion, it is not necessary to repeat the research questions again. Maybe just say something like "with respect to the initially given RQs, the following main conclusions can be drawn from the results"

We agree that avoiding the repetition would make the concluding section more concise. We will present the main conclusions directly and only reference the research questions posed in the Introduction.

L.496: With respect to one of my general comments: You conclude here on differences to the climatology, but this was never shown.

We agree that Fig. 9 encompasses all cyclones linked to extremes rather than solely the top 100 cyclones. To ensure our conclusion reflects what is shown, we will remove the reference to climatological cyclones, generalize the statement from extreme hazard cyclones to cyclones linked to extremes, and move the revised sentence to L. 492 (e.g., "In the classical cyclone phase space, cyclones are predominantly linked to P1 and W1 extremes when they are thermally symmetric").

**Reviewer 2**

**General comments:**

The authors introduce the concept of extreme hazard cyclones, defined according to the number of extreme precipitation and surface wind events associated with individual cyclones along their tracks. Using 75 years of ERA5 reanalysis, they investigate the geographical and seasonal distributions of these cyclones at global scale, considering two timescales (1 and 72 hours). The analysis subsequently focuses on the North Atlantic, examining cyclone characteristics, life cycles, dynamical structures, and differences among the four extreme-hazard categories.

Overall, I find the manuscript well written, clearly structured, and supported by informative figures. The topic is relevant to the Weather and Climate Dynamics community. However, I have several concerns regarding the definition and interpretation of the extreme hazard cyclone concept and the potential sensitivity of the results to methodological choices. Given that the authors aim to introduce a framework that could be applied in future studies, I believe these aspects deserve particular attention. My main concerns relate to (1) the distinction between tropical and extratropical cyclones, (2) the dependence of the ranking on cyclone identification and attribution methods, (3) the definition and interpretation of long duration extremes, and (4) the physical interpretation of the differences between precipitation- and wind-related extreme hazard cyclones. Addressing some of these concerns may require additional sensitivity analyses, while others could be resolved through more detailed methodological explanations and a careful discussion of limitations. I therefore recommend major revision before the manuscript can be considered for publication.

We thank the reviewer for their positive assessment of our manuscript and are grateful for the thorough review and the many thoughtful suggestions and ideas. Please find a first response to all major and specific comments below.

**Major comments:**

***1. Scope of the study: tropical versus extratropical cyclones***

I find the scope of the analysis somewhat ambiguous regarding the cyclone types considered. The title suggests a general framework applicable to all cyclones, whereas the Introduction focuses predominantly on extratropical cyclones and their associated surface extremes. However, the methodology is applied globally, and the results indicate that many of the most extreme hazard cyclones originate in the tropics and subsequently undergo extratropical transition. This raises an important question about the physical interpretation of the global ranking. Tropical and extratropical cyclones differ substantially in their formation mechanisms, dynamical structures, geographical distributions, and lifetimes. Combining these systems in a single ranking could therefore obscure differences in the processes responsible for their associated precipitation and wind extremes. I appreciate that the authors subsequently restrict the North Atlantic analysis to extremes occurring poleward of 30° N. Nevertheless, this

geographical restriction does not necessarily guarantee that all the associated cyclones have completed extratropical transition or are dynamically extratropical at the time of the event.

I suggest that the authors clarify the intended scope of their framework and consider whether separating tropical cyclones, extratropical cyclones, and systems undergoing extratropical transition would provide a more physically meaningful comparison. At a minimum, a sensitivity analysis excluding cyclones of tropical origin from the North Atlantic composites would help establish whether the main conclusions are robust. Alternatively, the authors could explicitly frame the study as a global cyclone-hazard classification encompassing different cyclone types and strengthen the discussion of the implications of combining these systems.

We agree that there is currently a mismatch between the predominantly extratropical framing of the Introduction and the global scope of our methodology. Following the reviewer's second suggested approach, we will explicitly frame the manuscript as a global cyclone-hazard classification that encompasses all low-pressure systems, including tropical cyclones, extratropical cyclones, and systems undergoing extratropical transition.

We view the ability to evaluate all tracked cyclones within a single unified framework as a central strength of our hazard-oriented approach. From a surface-hazard perspective, a location impacted by extreme winds or precipitation experiences that hazard regardless of whether the system linked to it is tropical, extratropical, or undergoing extratropical transition. Our framework therefore asks which systems across the full spectrum of trackable cyclones are associated with the most widespread extremes, and what physical characteristics they have.

To make this scope unambiguous and address the physical implications of combining different cyclone types, we will implement the following revisions:

1. **Introduction:** We will expand the literature review to explicitly cover tropical cyclones, extratropical transition, and their roles in driving extreme precipitation and winds, stating clearly from the outset that our framework is designed to evaluate all tracked cyclones globally.
2. **North Atlantic analysis (Sect. 4):** We will clarify that restricting extreme event locations to poleward of 30° N serves as a geographical boundary rather than a cyclone-type filter, and we will emphasize the importance of cyclones with tropical origin and transitioning systems for extremes in the North Atlantic.
3. **Discussion and Conclusions (Sect. 5):** We will add a paragraph on the implications of combining tropical and extratropical systems within a single hazard-ranking framework.

### *2. Definition of extreme hazard cyclones and sensitivity to cyclone identification*

A central methodological concern is the dependence of the proposed concept on the definition of the cyclone area. The authors use the outermost closed sea-level pressure contour to construct cyclone masks and subsequently attribute local surface weather extremes to

individual cyclones. Although this is a well-established approach, alternative cyclone-tracking methods identify cyclone centers without necessarily defining spatial footprints in the same way. Consequently, the number of extremes attributed to a cyclone may depend substantially on the selected cyclone detection method, particularly on the extent of the cyclone mask. This is important because the proposed ranking is intended to provide a general framework for characterizing extreme hazard cyclones. Its applicability across different cyclone tracking algorithms, datasets, and model resolutions is therefore not immediately evident.

I would encourage the authors to assess the sensitivity of their results to the definition of the cyclone area. For example, an alternative approach could attribute extremes within a prescribed distance of the cyclone center. Hawcroft et al. (2012) used fixed-radius regions and assessed the sensitivity of cyclone-associated precipitation to this choice. A comparison between the original cyclone masks and one or two fixed-radius definitions, even for a representative subset of cyclones, would help determine whether the rankings are robust to the attribution geometry. An additional issue is that the ranking counts spatially distinct grid-point extremes accumulated along each cyclone track, rather than directly measuring the maximum intensity of the associated hazards. Large, long-lived cyclones may therefore accumulate more extreme grid-point occurrences simply because they cover larger areas or travel longer distances.

This is not necessarily a limitation of the proposed concept, but it should be clearly distinguished from a ranking based on hazard intensity. The implications of this choice deserve further discussion, particularly when comparing cyclones with substantially different sizes and lifetimes. Finally, because this framework may be useful to the wider research community, I strongly encourage the authors to provide the code used for cyclone-to-extreme attribution and ranking, or at least a sufficiently detailed and reproducible implementation.

We agree that the number of extremes attributed to a cyclone depends on how the cyclone area is defined. Following the reviewer's suggestion, we are currently repeating the extreme event attribution and ranking using fixed circular radii of 500 km and 1000 km around the tracked cyclone centers to assess how sensitive the top 100 rankings are to the attribution technique. Given the sensitivity of cyclone tracks to tracking algorithms (Neu et al., 2013), we do not expect the 500 km, 1000 km, and contour-based techniques to lead to identical top 100 lists. While a fixed radius offers a straightforward and pragmatic baseline (which hopefully gives similar results to the contour-based technique), we view the closed-contour mask as a more adaptive approach. For roughly symmetric, circular systems, contour-based masks and fixed radii likely lead to comparable attributions of extremes. Using closed SLP contours, however, allows a cyclone's size and geometry to vary during its lifetime, whereas a fixed radius assumes the cyclone's footprint to be constant across cyclogenesis, the deepening phase, and cyclolysis.

We also agree that our ranking does not directly measure the intensity of each extreme event, but only counts if an event exceeds our percentile threshold. Using a high percentile ensures that every event still represents a severe local event, even though we do not weigh events by the magnitude of exceedance above that threshold.

As rightly noted, accumulating grid-point exceedances along a track means that small or very short-lived systems cannot reach the top of the ranking, whereas large, long-lived systems have

more opportunity to accumulate extremes. At the same time, size and longevity alone are not sufficient: large but weak cyclones fail to exceed the 99.9th percentile over wide areas and therefore do not rank highly either. In the revised manuscript, we will state clearly that we do not rank cyclones based on peak hazard intensity, but rather identify the systems linked to the largest cumulative number of extremes - an approach that naturally favors cyclones that are relatively large and long-lived, alongside being intense enough to produce widespread exceedances.

We are happy to provide a reproducible implementation of the code used for the attribution of extremes to the cyclones.

### *3. Definition and attribution of 72-hour extremes*

The selection of 1-hour and 72-hour timescales is motivated by differences in the potential impacts of short- and long-duration extreme weather. I agree with the general rationale, but the specific choice of 72 hours would benefit from further justification. Is this duration motivated primarily by the timescale of particular hazards, by the typical evolution of extratropical cyclones, or by previous studies? Would the main results remain similar if an intermediate duration, such as 24 or 48 hours, were selected?

A more important concern relates to the attribution of 72-hour extremes to individual cyclones. According to the methodology, a 72-hour extreme is linked to a cyclone whenever its mask overlaps the event location at least once during the 72-hour period. This criterion seems potentially permissive. A cyclone might occupy the location for only a few hours while most of the accumulated precipitation or sustained wind conditions occur during the passage of another cyclone.

The authors acknowledge that a single 72-hour event can be assigned to several cyclones and later discuss cyclone clustering. Nevertheless, I wonder to what extent this attribution choice influences the resulting rankings and the interpretation of the differences between short- and long-duration extremes. Would it be possible to test a more restrictive attribution criterion, such as requiring a minimum number of hours of overlap or considering the fraction of the 72-hour accumulation occurring during cyclone passage? Such a sensitivity analysis would strengthen confidence in the conclusions and help distinguish prolonged influence by an individual cyclone from extreme conditions associated with a sequence of cyclones.

The choice of 72 hours is motivated by the goal of capturing multi-day hazards while remaining comparable to the typical lifetime of an individual cyclone. During the exploratory phase of this study, we also examined intermediate event durations of 6 hours and 24 hours. We found that cyclones linked to many 6 h events resemble the W1 and P1 systems, whereas for the 24 h extremes cyclones already exhibit similar characteristics to the W72 and P72 cyclones, albeit with a less pronounced contrast to the 1 h cyclones. We therefore selected 1 h and 72 h, though using 24 h or 48 h would result in qualitatively similar conclusions.

We agree that linking a 72-hour extreme to any cyclone whose mask overlaps the grid point for at least 1 hour is permissive, and we understand the concern that a cyclone could be linked to a 72 h event despite passing over the location only during a short time. When multiple cyclones cross a region within a 72-hour period, asking which single cyclone "caused" the multi-day extreme often has no unique answer, because each passing system contributes to the accumulated precipitation or prolonged high wind speeds. Linking the event to all overlapping cyclones explicitly preserves this multi-cyclone contribution.

Conversely, imposing a stricter overlap requirement (such as a minimum overlap of 12 or 24 hours) introduces an additional subjective threshold and systematically favors slow-moving or stationary systems, making the selected cyclones potentially "stationary by design". Nevertheless, we agree that testing how sensitive the rankings are to this choice is informative. We will conduct a sensitivity analysis using a more restrictive attribution criterion, requiring a minimum of 24 h of overlap of the cyclone mask with the location of the extreme event.

### *4. Physical interpretation of cyclone dynamics and temporal evolution*

The comparison of cyclone structures across the four hazard categories is one of the most interesting parts of the manuscript. However, I believe some of the dynamical interpretations could be supported more convincingly. In particular, the authors interpret the stronger upper-level winds associated with the W1 and W72 categories as evidence of stronger baroclinicity and dry dynamical intensification. While this interpretation is plausible, upper-level wind speed alone does not provide a direct quantitative measure of baroclinicity.

Furthermore, the comparatively weaker lower-tropospheric PV anomalies associated with wind-related cyclones do not necessarily imply that diabatic processes played a minor role in their previous intensification. Latent heating may have contributed to cyclone development at an earlier stage, while the wind extremes occur later in the cyclone life cycle. This leads to a broader question concerning the timing of precipitation extremes, wind extremes, latent heating, and cyclone intensification. The manuscript presents composites at selected times, but the temporal evolution of the dynamical structures would provide additional insight into whether the diagnosed differences reflect distinct intensification mechanisms or different stages of cyclone development. I suggest considering composites at several time lags relative to maximum cyclone intensification, minimum SLP, or the occurrence of the largest number of extremes. Even a relatively simple analysis at intervals of 6 or 12 hours could help clarify the physical interpretation. At a minimum, I recommend distinguishing more carefully between conclusions directly demonstrated by the diagnostics and plausible interpretations that would require additional analysis.

We agree that intense upper-level winds do not directly quantify baroclinicity. As also noted in our response to the reviewer's specific comment on Lines 398–399, we will rephrase these statements such that upper-level jet features are described as *consistent with* enhanced baroclinicity rather than direct evidence of it. More broadly, we will review Sects. 3, 4, and 5 to ensure a clear distinction between relationships directly demonstrated and other interpretations.

We also agree that a weak lower-tropospheric PV anomaly at the single composite time step does not exclude substantial latent heating and diabatic PV production during earlier stages, particularly if precipitation extremes tend to occur earlier in a cyclone's development while wind extremes peak later (e.g., after maximum intensification or when the minimum SLP is reached). To test this, we will add a supplementary figure showing the frequency distribution of P1, P72, W1, and W72 extremes as a function of the cyclone's life cycle. This will help clarify whether the observed differences across hazard categories partly reflect different life cycle stages at which the composites are taken. Following the reviewer's suggestion, we will also compute time-lagged composites analogous to Figs. 7 and 8 at multiple time steps (e.g., at 12 h intervals before and after the occurrence of the largest number of extremes or the time of maximum intensification/minimum SLP) for the top 100 cyclones in each hazard category.

## Specific comments

### Introduction

**Line 27:** The expression "both over continents and within the storm tracks" is somewhat ambiguous. Storm tracks can also extend across continental regions, so the intended distinction is not entirely clear. Are the authors referring to the main climatological storm-track regions, which are predominantly oceanic, versus continental regions farther away from these regions? I suggest clarifying the terminology.
We refer to the predominantly oceanic climatological storm track regions, and the intention was to highlight that cyclone-associated precipitation occurs not only within these regions, but also over continental regions farther away from them. We will clarify this in the revised manuscript.

**Line 33:** The expression "a large fraction" is rather vague. If available from the cited study, providing an approximate percentage or range would make this statement more informative.
Pfahl and Wernli (2012) found that the fraction of extreme precipitation events linked to cyclones exceeds 80% in hotspot regions such as the Mediterranean and the main storm tracks (particularly east of the Northern Hemisphere continents), while ranging between roughly 50% and 80% across the broader mid-latitudes and less than 50% in the tropics. In the revised manuscript, we will replace the vague phrasing by specifying both these hotspot regions and the approximate percentage range across the rest of the globe.

**Line 35:** The authors mention that extreme precipitation is often associated with warm conveyor belts. Given the previous discussion of frontal precipitation, it would be useful to clarify the relationship between these two mechanisms. Warm conveyor belts are closely associated with frontal structures, so these contributions are not necessarily independent. A brief clarification, ideally supported by the cited literature, would improve the physical context.
Thank you for this suggestion, we will add a brief clarification.

**Line 38:** I suggest expanding the explanation of how complex topography can lead to precipitation extremes at considerable distances from the cyclone center. Presumably, the interaction between cyclone-induced moisture transport and orographically forced ascent is

important. Explicitly mentioning this mechanism would make the statement clearer and better motivate the potential limitations of the cyclone-mask attribution.
Thank you for the suggestion, we agree that an explicit description how topography may influence heavy precipitation will make the link much clearer. We will elaborate on this mechanism and revisit its implications in the limitations in Sect. 5.

**Data and methodology**

**Line 106:** The expression “constitutes the ideal basis” seems somewhat strong, particularly given the known limitations of reanalysis data in representing local precipitation and surface wind extremes. I suggest replacing “ideal” with “a suitable” or “a valuable” basis.
Thanks, we will replace the wording as suggested.
Regarding ERA5’s limitations for precipitation extremes, we also compared our ERA5-based precipitation extreme events against IMERG satellite data and found that, while absolute magnitudes differ as expected, the timing and location of the extremes agree well.

**Line 111:** The authors identify cyclone boundaries using closed SLP contours at intervals of 0.5 hPa. Could they clarify the rationale for this contour interval and whether it has been tested for sensitivity? Since the cyclone masks determine which extremes are attributed to each cyclone, differences in the detected footprint may affect the ranking.
The contour interval of 0.5 hPa is the standard configuration used and described in Sprenger et al. (2017). The paper introducing the tracking method (Wernli and Schwierz, 2006) evaluates the sensitivity of the algorithm to the contour interval. They demonstrate that although a smaller interval strongly increases the total number of identified cyclones, the additional systems are predominantly weak and spatially small. These weak systems are unlikely to be linked to many extreme events or end up ranked as one of the most intense cyclones in our study. For deep cyclones, using a different interval only shifts the edge of the cyclone mask. Therefore, we do not expect the choice of the pressure interval to substantially affect the attribution of extreme events to cyclones.

**Lines 113–114:** The description of cyclone tracking raises the question of how cyclone merging and splitting are handled. I understand that the authors apply a filter to avoid unrealistic jumps in cyclone position, but merging and splitting are also physically meaningful processes during cyclone evolution. How are these situations treated by the algorithm? For example, if two cyclones merge, are their tracks joined, is one track terminated, or are they treated as separate systems? A brief clarification would be helpful because the resulting track continuity may influence the number of extremes accumulated over a cyclone's lifetime.
The algorithm treats each track as a single, non-branching track.
**Cyclone splitting:** When a cyclone splits into two distinct centers, the center spatially closer to the cyclone’s position at the previous time step continues the original track, while the center farther away starts a new independent track. The two resulting systems are subsequently treated as separate cyclones.
**Cyclone merging:** When two cyclones merge into a single center, the track first processed by the algorithm is continued, while the other track is terminated. Both tracks are again treated as separate systems.
We will add a brief clarification to the revised manuscript.

**Line 125:** The authors adopt a higher percentile threshold than Pfahl (2014), defining extremes using the local 99.9th percentile. What is the main rationale for this choice? I appreciate that the study focuses on particularly severe extremes, but the selected threshold may influence both the number of identified events and the geographical distribution of the associated cyclones. A short explanation, and potentially a sensitivity test using a lower threshold such as the 99.5th percentile, would strengthen the methodology.

Several of the studies we cite use the 99$^{th}$ percentile and 6-hourly data (i.e. Pfahl and Wernli, 2012; Pfahl, 2014). Our analysis is based on data with a higher temporal resolution of 1 h. At an hourly resolution, lower thresholds correspond to relatively frequent events: exceedances of the 99th and 99.5th percentiles occur on average roughly once every 4 d and 8 d, respectively, which we do not regard as particularly extreme. By contrast, the 99.9th percentile corresponds to an average exceedance frequency of once every 1,000 hours (almost 6 weeks), while the 75-year hourly record still yields a large sample size (up to 657 exceedances per grid point, except for precipitation in very arid or polar regions).

The selected percentile naturally influences the number of identified events, and potentially also their geographical distribution and location relative to the cyclone center. To assess how sensitive our results are to this choice, we will perform a sensitivity test using the 99.5th percentile (due to computational demands potentially only with a subset of the 75 years of ERA5) and include the rationale for the 99.9th percentile threshold in the manuscript.

**Lines 127 and 216:** As mentioned in the major comments, the rationale for selecting 72 hours is not entirely evident. I recommend explaining why this timescale is particularly relevant and discussing whether the results are expected to depend strongly on the chosen duration.

A rationale will be included in the revised manuscript, possibly already after line 78. Please also refer to our response to Major Comment 3 of Reviewer 2.

**Line 132:** The attribution of extremes based on the cyclone masks is a central component of the proposed framework. Please see Major Comment 2 regarding the dependence of the results on the cyclone-area definition and the potential value of providing a reproducible implementation.

We agree that the attribution of extremes to cyclones depends on the chosen cyclone mask definition, and we are testing the sensitivity of our results by comparing the closed-contour masks against an alternative attribution using a constant radius around the cyclone center. Please refer to Major Comment 2 of Reviewer 2 for a detailed reply.

In addition, we will provide a reproducible implementation of the attribution code alongside the revised manuscript.

**Lines 135–138:** The criterion requiring only one hour of spatial overlap during a 72- hour event appears potentially permissive. I suggest quantifying how often long duration extremes are attributed to multiple cyclones and assessing how a more restrictive criterion might affect the rankings. Please also see Major Comment 3.

Thank you for the suggestion. A 72-hour extreme event can be linked either to a single (often) slow-moving or stationary cyclone, or to multiple consecutive cyclones passing over the same region. Across all attributed 72-hour extreme events globally, approximately 25% of the wind extremes and 34% of the precipitation extremes are linked to more than one cyclone.

We will recalculate the rankings with a more restrictive criterion requiring a minimum overlap of

24 hours between the cyclone mask and the extreme event location and examine its impact on the cyclone rankings. Please refer to Major Comment 3 of Reviewer 2 for a detailed response.

**Line 143:** The authors argue that only a small fraction of cyclone-related precipitation extremes are likely missed because most heavy precipitation occurs relatively close to the cyclone center. However, the Introduction also emphasizes the importance of frontal precipitation, and elongated fronts can extend beyond the closed SLP contours used to define cyclone masks. I would therefore be cautious about assuming that the missed fraction is necessarily small.
We agree that elongated fronts can extend beyond the outermost closed SLP contour and therefore also beyond the cyclone mask, leaving associated precipitation extremes unattributed. At the same time, defining an overly large attribution area around the cyclone to include elongated fronts would increase the risk of falsely linking extremes that are not actually caused by the system. While inspecting a sample of cyclones linked to numerous extremes suggests that a large share of precipitation extremes occur relatively close to the cyclone center along the warm front or bent-back front, we agree that assuming the unattributed frontal fraction is necessarily small is overly strong. In the revised manuscript, we will rephrase it using a more cautious formulation and mention precipitation along elongated fronts outside the cyclone mask as a limitation in Sect. 5.

### Global and North Atlantic results

**Table 1:** The table illustrates an important aspect of the global ranking: many of the highest-ranked cyclones have a tropical origin and subsequently undergo extratropical transition. I suggest emphasizing this more clearly when introducing the results, particularly given the predominantly extratropical focus of the Introduction. It would help readers distinguish between the global analysis and the subsequent North Atlantic investigation.
Thank you for the suggestion. In the revised manuscript, we will expand the Introduction to also cover tropical cyclones and extratropical transition and more clearly emphasize that many of the top-ranked cyclones originate in the tropics and undergo extratropical transition in both Sect. 3 and 4.

**Lines 269–271:** Restricting extreme-event locations to latitudes poleward of 30° N does not necessarily ensure that the associated cyclones are dynamically extratropical. This issue is particularly relevant for cyclones undergoing extratropical transition. Could the authors clarify whether cyclone type was assessed at the time of the extreme event, or whether some systems may still retain tropical characteristics? Please also see Major Comment 1.
We agree that restricting the analysis to latitudes poleward of 30° N does not guarantee that the associated cyclones are dynamically extratropical. We did not assess or filter by type of cyclone as an inclusion criterion; rather, cyclones in Sect. 4 were selected solely on a geographical basis (i.e., whenever a cyclone mask overlaps at least once with a grid point within the North Atlantic study domain). Consequently, some included cyclones may retain tropical characteristics, as can also be seen in Sect. 4.3, which shows that many extreme-linked systems exhibit thermally tropical structures in the cyclone phase space. In the revised manuscript, we will clarify this and, building on the reviewer's earlier comment regarding Table 1, emphasize the important role of cyclones of tropical origin and systems undergoing extratropical transition even within the geographically extratropical North Atlantic.

Please also refer to our detailed response to Major Comment 1 of Reviewer 2 regarding cyclone classification and tropical versus extratropical systems.

**Line 291:** The unit “Bergeron” should be defined when first introduced. I recommend providing its conventional definition, including the latitude normalization, and specifying how the deepening rate is calculated in this study. This is especially important because the investigated cyclones span a relatively broad latitude range.
Thank you for the suggestion, we will include a definition in the revised manuscript.

**Cyclone composites and dynamical interpretation**

**Figure 6:** I have two comments regarding this figure. First, could the authors clarify whether the cyclone-relative composites are constructed by translating the cyclone centers to a common origin only, or whether the fields are also rotated to account for the propagation direction of each cyclone? If the composites are not rotated, variations in cyclone orientation could smooth the frontal signals, particularly those associated with warm and cold fronts. It might be informative to consider whether rotating the cyclones according to their propagation direction changes the spatial patterns, although I recognize that such an approach introduces its own limitations. Second, there appears to be an inconsistency between the figure caption and the panel titles. Panels (c) and (d), as well as panels (g) and (h), appear to have their hazard-category labels reversed in the caption. The panels identified as containing hatched regions should also be checked for consistency.
Currently, the composites in Figure 6 are constructed solely by translating the cyclone centers to a common origin without rotating the fields, meaning they preserve their true geographical orientation. Although the propagation direction can be noisy or poorly defined for slow-moving or stationary cyclones, we will assess whether rotating the fields according to each cyclone's propagation direction alters the spatial patterns.
Thank you for spotting the inconsistency. The panel titles in the figure are correct, while the labels in the caption are reversed. They should read “(c) W1 and (d) P1” and “(g) W72 and (h) P72”.

**Figure 7:** The vertical PV composites are particularly interesting. Would it be possible to complement the absolute PV fields with PV anomalies relative to an appropriate background state, analogous to the upper-level PV anomalies shown in Figure 8? This could help distinguish the cyclone-associated PV structures from background PV variations and provide further insight into differences among hazard categories. It could be considered as supplementary material if including it in the main manuscript would make the figure too complex.
Thank you for the suggestion. We will compute the vertical PV anomaly composites, constructed analogously to the anomalies in Figure 8, and include them most likely in a supplement.

**Lines 398–399**: The authors infer strong baroclinicity from the presence of intense upper-level winds associated with the W1 and W72 categories. However, it is not immediately evident from Figure 8 how baroclinicity itself is diagnosed. While strong upper-level winds may be consistent with stronger vertical wind shear and horizontal temperature gradients, they do not directly quantify baroclinicity. Could the authors provide an additional diagnostic, such as the lower-tropospheric horizontal temperature gradient, vertical wind shear, or Eady growth rate, to

support this interpretation? Alternatively, I suggest rephrasing the statement to indicate that the jet structure is consistent with, rather than direct evidence of, enhanced baroclinicity.

We agree that Figure 8 does not directly diagnose baroclinicity. Following the reviewer's suggested alternative, we will rephrase this statement to avoid presenting the jet structure as direct evidence of baroclinicity (e.g., "Extreme hazard cyclones in the W1 and W72 categories are typically located in regions with strong upper-level winds near the left exit of a jet streak (Fig. 8f, h), which is consistent with enhanced baroclinicity favoring the development of surface wind extremes.")

**Lines 409–410:** The interpretation that W72 cyclones are primarily associated with strong baroclinicity and dry dynamical intensification may be too strong based on the presented diagnostics. Could these systems also have undergone substantial latent heating during earlier stages of their development, contributing to lower-tropospheric PV production and cyclone intensification before the occurrence of the most widespread wind extremes? The relatively weak PV tower in the composite at the selected time does not necessarily exclude an earlier contribution from diabatic processes. I suggest discussing this possibility or using a time-evolving diagnostic to distinguish between dry and diabatic contributions.

Thank you for the suggestion. We agree that because our composites show a single time step in the entire life cycle of the cyclones, we cannot exclude substantial latent heating during earlier time steps for the W72 category. We plan to include a supplementary figure illustrating when extreme events occur relative to the cyclone life cycle, and, as detailed in our response to Major Comment 4 of Reviewer 2, we are examining time-lagged composites (at earlier and later time steps than shown in Figs. 7 and 8) to better capture the temporal evolution of the top 100 cyclones across all four hazard categories.

## Conclusions

**Lines 504–505:** The authors conclude that cyclones associated with many surface wind extremes reach lower minimum SLP and intensify more rapidly than those associated with precipitation extremes. This is an interesting result, but I wonder whether part of the difference could be related to the timing of the processes involved. Latent heating associated with precipitation can contribute to lower tropospheric PV production and subsequent cyclone intensification. Consequently, intense precipitation may occur before the cyclone reaches its maximum intensity or strongest surface winds. Could the differences between precipitation- and wind related cyclones partly reflect their respective stages of development? It would be interesting to examine the evolution of precipitation and wind extremes relative to cyclone intensification and minimum SLP. For example, composites at time lags of ±6 or ±12 hours, or a longer life-cycle analysis, could provide useful evidence. This would help determine whether the identified differences reflect fundamentally distinct cyclone dynamics or the different timing of precipitation and wind extremes during cyclone development.

We agree that the observed differences could partly reflect different stages of cyclone development. To test this, we will examine the timing of extremes along the cyclone life cycle and compute time-lagged composites (see also our response to Major Comment 4 of Reviewer 2).

**Code and data availability:** The authors indicate that key cyclone tracks and extreme-event locations will be made available upon acceptance. Given that the manuscript introduces a methodological framework intended for application to other regions and climate simulations, I would strongly encourage the authors to make the corresponding cyclone-to-extreme attribution and ranking code available as well. This would substantially improve the reproducibility and transferability of the proposed framework.

Thanks for the suggestion, we will provide a reproducible implementation of the code to attribute extremes to cyclones.

### References

Hawcroft, M. K., L. C. Shaffrey, K. I. Hodges, and H. F. Dacre (2012): How much Northern Hemisphere precipitation is associated with extratropical cyclones?, Geophys. Res. Lett., 39, L24809, https://doi.org/10.1029/2012GL053866.

### References

Field, P. R. and Wood, R.: Precipitation and cloud structure in midlatitude cyclones, J. Climate, 20, 233–254, https://doi.org/10.1175/JCLI3998.1, 2007.

Neu, U., Akperov, M. G., Bellenbaum, N., Benestad, R., Blender, R., Caballero, R., Cocozza, A., Dacre, H. F., Feng, Y., Fraedrich, K., Grieger, J., Gulev, S., Hanley, J., Hewson, T., Inatsu, M., Keay, K., Kew, S. F., Kindem, I., Leckebusch, G. C., Liberato, M. L. R., Lionello, P., Mokhov, I. I., Pinto, J. G., Raible, C. C., Reale, M., Rudeva, I., Schuster, M., Simmonds, I., Sinclair, M., Sprenger, M., Tilinina, N. D., Trigo, I. F., Ulbrich, S., Ulbrich, U., Wang, X. L., and Wernli, H.: IMILAST: A community effort to intercompare extratropical cyclone detection and tracking algorithms, B. Am. Meteorol. Soc., 94, 529–547, https://doi.org/10.1175/BAMS-D-11-00154.1, 2013.

Pfahl, S.: Characterising the relationship between weather extremes in Europe and synoptic circulation features, Nat. Hazards Earth Syst. Sci., 14, 1461–1475, https://doi.org/10.5194/nhess-14-1461-2014, 2014.

Pfahl, S. and Sprenger, M.: On the relationship between extratropical cyclone precipitation and intensity, Geophys. Res. Lett., 43, 1752–1758, https://doi.org/10.1002/2016GL068018, 2016.

Pfahl, S. and Wernli, H.: Quantifying the relevance of cyclones for precipitation extremes, J. Climate, 25, 6770–6780, https://doi.org/10.1175/JCLI-D-11-00705.1, 2012.

Rudeva, I. and Gulev, S. K.: Composite analysis of North Atlantic extratropical cyclones in NCEP–NCAR reanalysis data, Mon. Weather Rev., 139, 1419–1446, https://doi.org/10.1175/2010MWR3294.1, 2011.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Villaume, S., Haimberger, L., Woollen, J., Buontempo, C., and Thépaut, J.: The ERA5 global reanalysis from 1940 to 2022, Q. J. Roy. Meteorol. Soc., 150, 4014–4048, https://doi.org/10.1002/qj.4803, 2024.

Sprenger, M., Fragkoulidis, G., Binder, H., Croci-Maspoli, M., Graf, P., Grams, C. M., Knippertz, P., Madonna, E., Schemm, S., Škerlak, B., and Wernli, H.: Global climatologies of Eulerian and Lagrangian flow features based on ERA-Interim, B. Am. Meteorol. Soc., 98, 1739–1748, https://doi.org/10.1175/BAMS-D-15-00299.1, 2017.

Wernli, H. and Schwierz, C.: Surface cyclones in the ERA-40 dataset (1958–2001). Part I: Novel identification method and global climatology, J. Atmos. Sci., 63, 2486–2507, https://doi.org/10.1175/JAS3766.1, 2006.