RC1: 'Comment on essd-2026-181', Mika Rantanen, 28 May 2026 reply

Review of Earth System Science Data Discussion manuscript "The Named Storms Catalogue: unlocking learnings from past events" by Gonzalez et al.

This data description manuscript presents a new named storms dataset for the UK and western Europe. The dataset contains the tracks and development and intensity metrics of all named extratropical cyclones during the ten-year period from 2015/2016 to 2025/2026.

Several catalogues and databases of European windstorms already exist. However, these datasets often include only the most intense storms or do not systematically identify storms by their operational names, making it difficult to link individual events across different data sources and reports. Thus, the main novelty of the present dataset is that it provides a comprehensive collection of all officially named UK storms during the study period, together with consistent track and intensity information. As far as I am aware, such a dataset does not yet exist.

The manuscript was really clearly written and I enjoyed reading it. The structure of the paper was good, however, the authors could pay attention to the figures, especially their font sizes and the use of abbreviations (see comments below). Overall, I can still recommend publication of the paper. I have some relatively minor comments which I hope the authors can address before the publication.

We thank the reviewer for their comments and constructive suggestions. We have worked on an improved version of the manuscript based on the reviewers' comments and we provide detailed replies to their comments below.

Minor comments:

1. At L128, it is said that operational meteorologists confirmed that the TRACK-derived tracks corresponded to the named events. However, it was not explicitly stated whether the opposite was also true. Were all named storms associated with a "trackable" extratropical cyclone? Is storm naming inherently tied to the presence of a well-defined low-pressure system, or can situations occur where the Met Office names a storm that is later not captured by TRACK? More generally, how did you ensure that TRACK identified all named storms?

Thanks for the questions. We have included further clarifications on these points in section 1 of the revised manuscript. The western storm naming group comprises of the Met Office, Met Eireann and the KNMI. Their naming procedures dictate that there should be a trackable low, but this can sometimes be well away from the impact footprint which led to storm naming (e.g. storm Desmond in December 2015) and some other named storms were triggered by wind threshold-led warnings rather than a clear low pressure system. UK storm naming started out as being windstorm-led but in extending to being multi-hazard there are occasions when the low pressure system is not easy to find (e.g., storm Christoph in January 2021) or the impacts were the result of the merging of multiple systems (e.g., storm Babet in October 2023). In this context, there may be times when TRACK does not capture a low pressure system, or when the metrics associated with the associated low pressure systems are less relevant to the description of the hazards. Nonetheless, the process applying TRACK on ERA5 followed by validation by Operational Meteorologists was very effective and successfully managed to retrieve most named storms, and all of the events named by the Met Office.

2. The manuscript is not mentioning if post-tropical storms are included in the dataset. I have the understanding that the Met Office or other met institutes do not rename post-tropical cyclones but retain the original name if the storm is forecast to cause damage in Europe. There seems to be "exOphelia" in the dataset in 2017, but are there other post-tropical cyclones? This could be mentioned in the text. In Figure 4, one track is originating from the tropics. Was that a tropical cyclone?

Thanks for the questions. We have included further clarifications on these points in section 2.3 of the revised manuscript. Post-tropical storms are not renamed in the UK, and there haven't been many since storm naming began. Ophelia was included for being the only event that was included in the Met Office list of impactful storms which have affected the UK and for which publicly available event reviews are written. These can be found at https://weather.metoffice.gov.uk/warnings-and-advice/uk-storm-centre. Experience of the co-authors indicates that this event would have been named by the Western Group had it not been named before its extratropical transition.

3. The whole dataset and thus the manuscript is based on ERA5 reanalysis. Although ERA5 is the state-of-the-art reanalysis and arguably the best possible choice for this analysis, it still has some biases, for example in wind speeds. It was mentioned at L389 that ERA5 did not capture the full severity of Storm Eunice. Do you think that the use of ERA5 may systematically underestimate (or in some cases overestimate?) some of the development or hazard-related variables analysed in the manuscript? This aspect might deserve a short discussion paragraph in the last section for the end-users, especially since ERA5 biases associated with windstorms have likely been documented in previous literature.

Thanks for the question. We have included further clarifications on these points in section 2.1 of the revised manuscript. We acknowledge that ERA5 is subject to biases but the fundamental aim of the work was to collect the catalogue based on a consistent dataset. We placed our focus on capturing the storms track and the broader scale evolutions, and these were qualitatively validated by the Operational Meteorologists. Beyond that, basing the storm metrics on a consistent dataset allows to rank and compare storms, even if the metrics are subject to biases. To precisely capture the details of each system, higher resolution and more frequently initialised datasets would be advisable, but that was not the goal of this project.

4. Tables and Figures. I find the extensive use of abbreviations in tables and figures (e.g. MxAcP, Mxppr, etc.) somewhat difficult to follow. While some abbreviations (such as Mxgust) are intuitive, many others are not immediately clear to the reader. For clarity and accessibility, I would recommend either spelling out the variables in full or providing a consistent legend/caption where all abbreviations are clearly defined. Especially in Figures 6-8.

Thanks for the comment. The figures in the revised manuscript now include more explicit variable naming to enhance the accessibility.

5. Also, especially the Figures 6-8 (and others too, such Fig. 9 subplots) have lots of small text which I was not able to read from a printed copy. Can you increase the font size of the figures?

Thanks for the comment. The figures in the revised manuscript have increased letter sizing for clarity.

Line-to-line comments:

L46: Are all windstorms in the Western naming group named in advance based on forecasts, or are some named after impacts became evident? Have some storms been named retrospectively? A brief clarification of the naming procedure (and timing relative to forecast vs. observed impacts) would be useful here. For example, in Finland, a windstorm may sometimes be named once it is clear that it is causing significant impacts, sometimes on the same or even the next day it affects the country.

Thanks for the question. Some additional clarifications have now been included in Section 1 of the revised manuscript.

In the UK, storms are named in advance based on forecasts. There is no retrospective naming. This is because storm naming is all about enhancing communication of impactful weather. Typically, a storm will be named 2-3 days in advance and, more rarely 4-5 days in advance. However, there are occasions when storms are named at (much) shorter lead times as confidence in their impact potential increases. The exact point at which a storm gets named varies slightly across the partners in the western naming group, but processes are broadly similar. At the Met Office, storms are named based on potential for impacts with a degree of confidence in these impacts occurring, i.e. the storm hitting the UK, also necessary.

At the Met Office, storms are named a) if an amber or red wind warning is issued, b) if a yellow medium or high impact wind warning issued and there is potential, as confidence increases, for this to become an amber or red wind warning or c) if there are amber or red rain or snow warnings and a NSWWS wind warning (any colour) in force. NSWWS warnings are impact-based not threshold-based so the point at which a storm may be named is not necessarily immediately obvious from meteorological data alone. The KNMI and Met Eireann make a little more use of thresholds but anecdotally the meteorological parameter space in which we'd all name a storm is very similar.

Section 2.1. I guess you used ERA5 in its native resolution (0.25°)? This is not mentioned.

Thanks. Now clarified in Section 2.1 of the revised manuscript.

Section 2.2. It was not fully clear from the text why there are two ocean datasets. Which one do you use e.g. in Fig. 12, and the figures in the database?

Thanks for the comment. The revised manuscript includes now a clarification in section 2.2. One dataset models surges and the other one models waves.

L71: Why did you track the systems with 6-hourly fields and not e.g. 3-hourly? 6-hourly data may be somewhat too coarse for representing local impacts, especially for fast-moving systems where the location and intensity of the strongest winds and precipitation can change substantially within six hours. Using 3-hourly fields could improve the representation of storm evolution by better resolving rapid deepening and the intensity peaks.

This would indeed be possible, and we thank the reviewer for suggesting this. There is a case for using a smaller time step between tracking points; however, we used the 6-hourly data for a few reasons. First, operational synoptic analyses are available every six hours from the Met Office, which were used to visually inspect whether the tracking algorithm had identified the correct storm (we don't have analysis maps for the intervening hours). Second, the hazards are computed from an hourly dataset and so anything that occurs within the ±3-hour window around the track's points will be included in the hazard analysis. Last of all, 6-hourly data have the advantage of being the most consistent across different datasets beyond ERA5, and consistent with previous studies (e.g., Manning et al. 2024). For example, if we were to apply this method to other model data (e.g. CMIP), we may only get access to 6-hourly data. Therefore, having a baseline with 6-hourly output gives us more flexibility to apply this method beyond ERA5 and our own model data. We do acknowledge that, now we have this dataset, there is no reason we cannot extend it in the future in the way the reviewer suggests, but that is beyond the scope of the current paper.

L98: Here you write that you calculated vorticity from wind components. But vorticity is available directly at pressure levels from ERA5. Why did you calculate the vorticity fields by yourselves?

This is true and we thank the reviewer for checking in on this point. The TRACK software package can calculate the vorticity field from the winds as part of its setup, which we have retained and undertaken as a matter of process. Furthermore, spectral filtering is applied to the vorticity field to smooth it so that synoptic-scale features are retained with sub-synoptic and planetary scale fields removed. Again, this is a process applied as part of the software that we have retained. Finally, having the 850 hPa wind field is useful for looking at other processes around the cyclones. There is no reason why we could not use ERA5's vorticity; however, given the field is smoothed, it would not be the exact field that appears in the ERA5 database.

L121: Vorticity maxima may not always be associated with MSLP minima, particularly during the early development stage of cyclones. Were such situations included in the tracking procedure, and how did you deal with them?

This is a good point and there were occasions when this was noted. Nevertheless, in all cases, the cyclone was very remote from the target area of the UK, and the relevant parts of the cyclone's lifecycle (e.g. the rapid deepening phase) were always captured using the method we applied. Furthermore, there is no definition of when an extratropical cyclone becomes an extratropical cyclone (unlike tropical cyclones, for example). Therefore, the inclusion or not of these points is highly subjective. Ultimately, as it is explained in the text, we must make our own subjective decision on when a cyclone becomes a cyclone and, in these instances, the cyclones were always remote from the target area.

Table 1, bearing. I wasn't sure about how this is calculated. Here you state "6h positions of maxima". Which maxima?

Thanks for the question. The calculation is now described in more detail in Table 1 as follows: Direction is derived from assuming a great circle trajectory between the 6h displacements of the vorticity maxima and assigned to the time of the end point. Bearing is expressed in degrees using the wind direction meteorological convention.

L184: Surge residual is a new term for me. Can you explain it here a bit more?

A more detailed description is now included in Table 2. Surge residual is the difference between the observed sea level and the predicted astronomical tide at any given moment. It represents the change in water level caused by meteorological factors like strong winds and low atmospheric pressure.

L186: Section 0?

Thanks for the note. The Section hyperlink should have pointed to Section 2.2. This is now fixed in the revised manuscript.

L216: "adding 6 hours on - 3 hours for each end of the track." I can't follow here, did you add 6 hours or 3 hours to the durations?

Thanks for the question. This explanation has now been edited in the revised manuscript to improve clarity as follows: "Approximate durations of impactful wind speed and accumulated rainfall (considering track points are 6 hours apart) were derived by taking the time between the track points where the first and last exceedances occurred, and adding 6 hours to account for a possible start of the impacts in the midpoints between track locations (3 hours before the first exceedance and 3 hours after the last one).

Table 2, Maximum wind gust. Is this the single highest ERA5 grid-point value within the storm footprint? Since peak gusts are often associated with mountainous / high-elevation regions, it would be useful to clarify whether these maxima are strongly influenced by high-elevation grid points.

Yes, the 3s instantaneous wind gusts are considered over a 10-degree radius of influence of the storm track for all the points over land. These values can certainly include strong gusts for areas of topographical complexity, but these would still be representative of the hazard associated with the storm. It Is also important to highlight that the identification of storm events is in this case independent of any threshold exceedance of wind gusts or wind speeds, unlike other cases where “windstorms” are derived from the winds themselves.

The metrics included in the catalogue do not intend to capture actual impacts, vulnerability or risk. Building them from a consistent long record dataset like ERA5, despite its known biases, allows a robust comparison between events.

Table 2, maximum wind gust. Your tracks are every 6 hours. Did you get the ERA5 wind gust data every 1 one hour to obtain the absolute maximum wind gust associated with the tracks, or are these just the instantaneous wind gusts from the 6-hourly wind gust fields? It was not entirely sure from L73 whether the wind gust data is hourly or not (it could be stated more explicitly).

Thanks for the comment. A clarification is now included in the revised manuscript: The hourly wind gusts fields were used to extract the maximum value over the period of influence of the storm over land.

L264: Full range may be misleading as it’s definitely not the physically possible “full range” due to small(ish) sample size. Maybe just the widest range?

Thanks for the comment. The revised manuscript now makes use of the expression “wider range”, but a clarification about the sample size was already included at the end of the sentence.

Figure 6, 7 and others: There is no unit in the y-axis. Also, could the titles be explicitly written, without the abbreviations, in order to ease the interpretation (and with bigger size). The abbreviations may be clear for the authors that have used them for a long time, but not for a regular reader who reads the paper for the first time.

Thanks for the comment. The figures’ legend and titles have now been modified to improve clarity in the revised manuscript.

L281: What do you mean by “corrections to account for changes in the background climate”? Do you mean that the storms in 2015 occurred in a colder climate than the storms in 2025, and this may have an impact on the hazard metrics? The effect is probably very small, given the short 10-year period involved (~0.2 °C of global warming) and the relatively limited response of extratropical cyclones to warming.

Thanks for the comment. We have now included a clarification in the revised manuscript that explains that no metric is based on detrended data. Even though the naming period is short, a decision could have been made to detrend the underlying variables based on a longer period (e.g., WMO climatological normals) to compare and rank storms. A conscious decision was made not to apply such strategy.

L282: I think it's interesting that DJF season has lots of named storms with close to zero hazard metrics (Fig. 7a–c), while this is clearly not the case for JJA. Does this imply that named JJA storms have a higher probability to cause damage? One possible explanation is that winter storms are more often named based on forecast intensity and anticipated impacts, even if their hazard metrics remain modest because strong winds are more common in winter climatology. This also seems somewhat inconsistent with the Introduction statement that naming is primarily based on impacts (L46-47), so a short discussion of this could be useful.

Thanks for the comment. The primary reason for seeing strong summer storms compared to winter storms by this metric is that it calculates the strength of the storm relative to a monthly (and location specific) climatology. This means that a storm with particular wind/rain in summer would score more highly than the same wind/rain in winter. Whilst not a perfect mapping to impacts, it is the case that the impacts of a storm relate to how unexpected the conditions are compared to usual (e.g. since more outdoor activity happens in the summer, the same rainfall is more impactful in summer than in winter).

As a result of the definition of the metric caution should be used comparing storms across very different seasons as a result. The text has been adjusted to reflect this in Section 3.1.2.

Figure 8. As commented earlier, the abbreviations could be written open. Particularly in the caption of the figure.

Thanks. This has been addressed in the revised manuscript.

L375: Citation comes before the dot.

Thanks. This has been corrected in the revised manuscript.

L387-388: Volonte et al. 2024 is missing from the reference list. How about the ERA5 resolution? Can that be one reason for the underestimation of the severity?

Thanks. The Volonte et al. 2024 reference was included in line 531. And yes, as expressed in the following sentence, ERA5's limited resolution can at least partly explain the misrepresentation.

Conclusions. Are you planning to update the dataset in the future? If yes, this could be mentioned here.

Thanks for the comment. We would like to ensure that the Zenodo catalogue gets updated at the end of each storm seasons, but we do not have guaranteed funding to do so, and we therefore prefer not to misrepresent this in the manuscript.