Anonymous Referee #1, 18 Jul 2023

Overall comments:
The aim of this study is to investigate potential sources of European winter windstorm predictability. The overall methodology appears to make sense, but I found the article hard to read (it really needs a thorough proof-read as there are many poorly-worded or grammatically incorrect sentences; in some cases to the extent that it wasn't really clear what you mean).

In particular, I didn't fully understand the "process-based view" method, so please could you review the description of this to make sure it is clear.

Some of the figures need a bit of work to make them clearer – in particular I found the schematics in Fig 1 and 2 confusing (and also some of the text on them was too small and hard to read – not sure if this was in the conversion to pdf or an issue with the actual figures).
[LD] The current figure setting is for one column, we could increase the figure size inside the manuscript to a two-column figure, but also try to adjust the font within the figures. Another option is to remove figure 1 completely from the manuscript and move figure 2 to the appendix, this would lead to only explanations of the physical connections as text within the manuscript.

Specific comments:
Abstract: The second paragraph of the abstract seems rather vague. Please could this be rewritten – perhaps if you look at the conclusions section and summarise the key results in a more solid way.
[LD] The second paragraph can be adjusted, rephrased and clarified with more details. The 2nd abstract paragraph could be simplified by using numbering for the 3 investigation steps and give a clear statement for each investigation step as summary.

L38: what do you mean by "signal" here?
[LD] This refers to the correlation skill score used in the mentioned publications. This can be changed to a more precise word ("skill")

L 50: can you say what the equivalent-potential temperature actually is (either here or later) as I, and presumably other readers, am not familiar with this quantity.
[LD] More explanation of Theta_e can be added to the manuscript, especially the introduction. E.g. "Θe is a parameter that describes the temperature of a fully dried air parcel dry-adiabatically lowered onto a reference level, usually 1000hPa (Bolton, 1980)."

L105: do you think your method will detect sting-jet storms? These storms have some of the strongest and most damaging surface winds (e.g. the Great storm of 1987) but are relatively short-lived and affect a relatively small area compared with

other types of storms where the strong winds are due to e.g. the cold-conveyor belt winds. I wonder if your tracking method would capture these events and if so, whether they might have different dynamical drivers than other types of storms. This may be beyond the scope of your study but definitely worth considering/mentioning.

[LD] The used algorithm is not designed to detect sting jets or cyclones. The algorithm is an impact-based tracking, that is locating synoptic-scale exceedances of wind-speed. The authors are aware of the interest in sting jets, but this is not part of this study as sting jets would only be possible to detect on smaller scales.

L122 check reference format, should be in brackets
[LD] Thanks for pointing this out, this will been changed.

Throughout: you seem to use "i.a." instead of "i.e." or "e.g."
[LD] "i.a." stands for inter alia, but we can change it throughout to "e.g."

Figure 1: Caption should be "schematic" rather than "scheme". Some of the text within the figure is very small.
[LD] Please see comment above. Figures were set to one column figures, but can be changed to 2-column, so they appear bigger. And font can be adjusted.

Figure 1: In general, I didn't really understand this schematic and I think more description is needed in the figure caption and/or text. In particular:

- What do the different coloured boxes mean? In the text it says "The coloured boxes indicate in which physical view (Quasi-geostrophic Omega- and PV-theory) these factors are included" but it doesn't say anywhere what each individual colour corresponds to (this should be in the figure caption).
- What are your definitions of "cyclone" and "windstorm" that means these are considered as separate things? And for instance, why is tropical precipitation labelled as a "source of predictability" for windstorm but not for cyclone?
- What is the PV 350K bandpassfilter and how is it different from 350K?

[LD] We would think about removing this quite complex figure from the manuscript

Figure 2: Text is small/fuzzy and the text on the diagonal lines is squashed and hard to read. There is some description of this figure in the text but it wasn't clear to me how some of the text corresponded to what is in the schematic and there are many aspects of the schematic which are not explained. For example what are the small rain-cloud shapes which appear to be at the surface? I appreciate that there are lots of complex processes involved and it would be useful to summarise them in some sort of schematic, but this particular schematic does not seem to show them very clearly and requires some work.

[LD] These suggestions can easily be change by increasing the figure size, the font size and add more details into the text. (e.g. "The upper tropospheric baroclinicity

(EGR 400hPa) triggers respective upper-level divergence (peach) and hence, creates the jet stream (orange). The counterpart to this is the SST (ocean colour) which influences the low-level baroclinicity (EGR 700hPa), which impacts the MSLP gradient (light blue) and hence, the wind speed (yellow). Another process related to the potential predictability of windstorms is caused by convective tropical precipitation (dark blue) via vertical lifting, triggering a Rossby wave train (purple) formation to the North Atlantic region in higher altitudes.")

L140: Do you do this for each ensemble member separately? So for each ensemble member you do the 10-3-10 classification, which means that when you then look at composites they are based on different sets of years for the different members and for the observations? Could you make this clearer? Also if this is the correct interpretation, could you comment on how much similarity/difference there is between the ensemble members and the obs: do the same years get picked out by most of the ensemble members as high/low storm count years? Are they consistent with the obs? This could have implications for the interpretation of the results in terms of the influencing factors.
[LD] We can add more explanation to the manuscript. Yes, we are doing each separation individually per member, as we really want to have the physics of seasons with many storms against less storms in each realization. We can add more explanation to the manuscript (like "The separation is done individually per model ensemble member to ensure that each composite compares strong vs weak storm seasons internally.")

L140: what do you mean by "at least a decade long duration"? A decade is a period of 10 years, but I think you mean 10 years of data?
[LD] Yes, this is a very good suggestions and we could take the phrasing of "10 years of data".

L144: is this the total absolute difference of individual members or of the ensemble mean?
[LD] This is based on the ensemble mean, as the forecast skill is also based on that. (could be changed to "using the absolute difference of the respective seasonally averaged factor over an individually defined region in the GloSea5 ensemble mean and ERA")

L150 Please can you give a brief description of what the ranked Tb-Kendall correlation is
[LD] A more detailed description of the used measure can be added, like "Kendall correlation is a similar measure to the commonly used Person's correlation but investigates ranked time series and is less subject to normally distributed data."

L151: "chapter" should be "section"
[LD] Can be changed throughout the manuscript to "section".

L165: This is strangely written: what do you mean by "represented in the model as derived from reanalysis"? Do you mean something like "Does the model represent the same physical connections between causal factors as the reanalysis"?

[LD] Yes, the suggestion of the reviewer is right and we would like to adopted this phrasing.

Figure 3: It's not clear to me what has been plotted here. Is it the respective quantity in the strong storm seasons minus the weak storm seasons, at each grid point? Why is the magnitude in the GloSea5 mean so much smaller than in ERA5?

[LD] Yes, it is the respective quantity as difference between strong vs. weak storm seasons. As this has been calculated per each ensemble member, it is possible that some members have the same strength of the link and others don't. Hence, an averaged over all members to present the average difference for GloSea5 could result in a smaller difference than ERA5.

Figure 4: I found this figure very confusing! Please add more description in the figure caption. Where you say "connection" I think you mean "correlation" and where you say 1st/2nd column you mean row not column. The bar graphs don't have a scale so I'm not sure how useful they really are. And I thought the colours corresponded to correlation values but then the numbers in the boxes don't seem to correspond. I'm also a bit confused by the interpretation of the figure and the conclusion you draw from it that the connection between the factors and windstorms is well represented in ensemble members as well as ensemble mean (L204/5).

[LD] The number that has been referred to is not the correlation values. The colour refers to the ERA5 correlation, the number refers to the percentile of the ERA5 correlation within the GloSea5 member distribution. This can be made more clear in the figure caption. (E.g. "The coloured-red line is the ERA5 correlation value within the GloSea5 member distribution, and the number represents the percentile of ERA5 in that distribution.")

L213: do you mean upstream of the British Isles rather than downstream?

[LD] that is correct, I am sorry for the mix-up.

Figure 6: I didn't find it clear what this is showing, please can you explain more clearly in the caption or text. My understanding: you're showing the Kendall correlation difference in windstorm forecast at each grid point, between seasons in which the various factors or processes are well- and poorly-forecast. And the boxes show the regions that you're interested in for the factors/processes. But how do you combine the boxes to determine if the season is well or poorly forecast? What if it's well forecast in one box region but not another?

[LD] I spatially average the respective quantity in the shown boxes and then use the same process with absolute difference and compare it between ERA5 and GloSea5.

This could be a potential additional sentence for clarification:

"For both views, the selected regions (which can be multiple) are spatially averaged, and well- and bad-predicted seasons are detected by the absolute difference between the resulting ERA5- and GloSea5-time series in the used regions. The regions for the factor-skill-view are all skilful for the respective factor. For the process-based view, this is not a criterion."

Figure 6: the dots and triangles are too small to be able to differentiate between them. And it's also not clear to me what they mean. What does it mean for a well predicted year or a badly predicted year to be significant?
[LD] This is supposed to be a significance measure. A dot/triangle is added if either the first or the second parameter in the difference is significant but not both.

L255 (and similar description/interpretation of results): does your definition of well forecast or badly forecast necessarily mean that the factor is well forecast in all the boxes? Where you have several boxes could the quantity be very well forecast in just one or two of the boxes but not in the others?
[LD] I have checked each box individually and the combination of all of them, for calculation details, see above.

L317: it might also be worth saying that the skill is increased over other parts of Europe – NW Europe, SW Europe?
[LD] This could be easily added to the text. Along these lines "t has been found that all main factors increase the forecast skill of winter windstorms over the British Isles and the North Sea by increasing the forecast skill of relevant factors in storm-relevant regions. SST and Θe additionally improve the windstorm forecast skill over Central Europe and Southern Scandinavia."