Response to Anonymous Referee #2

We would like to thank the reviewer for their helpful comments on our manuscript. Below we list how we've addressed each of their comments (in blue italics).

Please concisely expand the context provided to give appropriate credit to past work, which is thin in some places (e.g. Sections 2.1 & 2.2).

We have expanded these sections with more detail. Section 2.1 has had L80-L85 added:

Numerous SIs have been created for wind damage. Klawa and Ulbrich (2003) use the cube of wind gust above the local 98th percentile, with numerous other studies (e.g. Leckebusch et al. (2007, 2008); Pinto et al. (2012); Little et al. (2023)) using an SI of similar form adapted to gridded data. Heneka and Ruck (2008) presented an SI using the square of exceedances, although this assumed the damage threshold was normally distributed. Bloomfield et al. (2023) introduced a flood severity index, also using the exceedance over threshold approach, using linear exceedances of river flow data. SIs of this form are less influenced by outlier extreme events.

Section 2.2 has had L96-L103 added:

Aggregated Severity Indices (ASIs) are frequently used as a proxy for total damage. Hillier and Dixon (2020) aggregated wind gust and total precipitation over the extended winter season (October-March), concluding that extreme precipitation winters results in an uplift of aggregate extreme wind hazard for most of Europe. This compared the relative value of wind ASIs for the top and bottom thirds of winters ranked by precipitation ASI. Hunter et al. (2015) used ASIs to investigate the relationship between frequency and mean intensity of windstorms, concluding the Scandinavian Pattern was a driver of this relationship. Jones (2022) derived a framework to model the relationship between frequency and wind ASI, while Jones et al. (2024) found most of Europe have negative Pearson's correlation between wind and precipitation ASIs at high thresholds.

To increase the readability of the work and make it accessible to a wider audience, I think that an additional Figure panel is needed to illustrate the data underlying the negative correlation i.e. a scatter plot showing the ASI data points for a particular (x,y) location, the two thresholds from Fig2 a-c, different colours for dots contributing to Fig. 2a-c at the location. Consider adding a trend line for the correlation(s). And, also mark this location on Fig2 a-c. This might be a second panel in Fig. 4.

Figure 4 had 3 panels added, showing the ASI for all of France, the correlation and the framework estimate for the 3 thresholds shown in Figure 2 (new figure shown in Figure A) Figure 4 has been amended to ensure calculation of framework components reflects the methods shown in Appendix A.

We have added "Jones et al. (2024) describes how sample correlation between wind and precipitation ASIs decreases with increasing threshold, including differing behaviours between regions.". (L195)

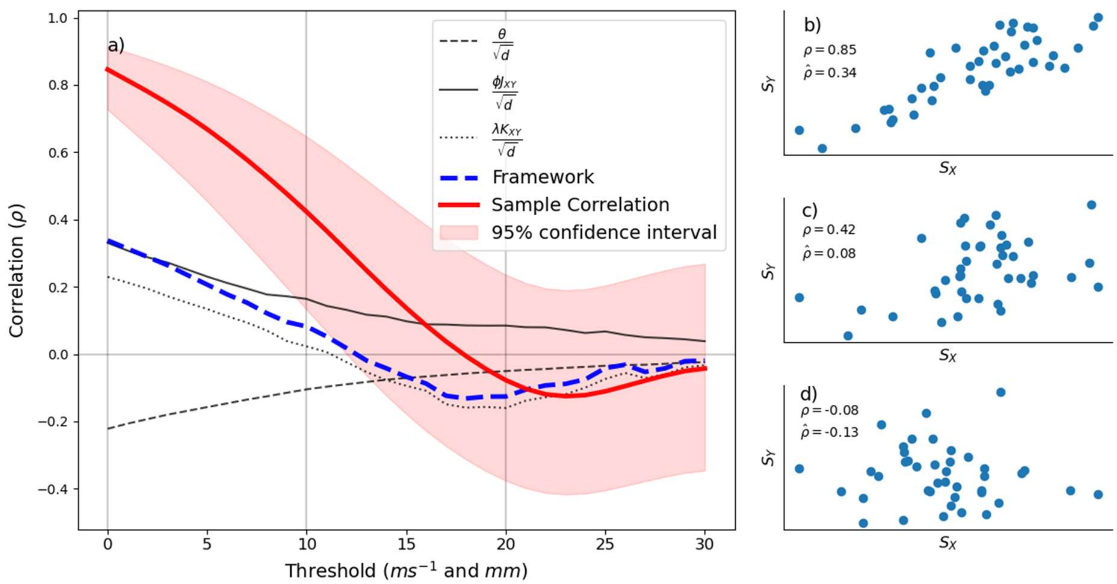

Figure A - Panel (a) shows the framework (red line) with 95% confidence interval (red shaded area). Framework C is shown by the blue dashed line. Framework components shown by the thin black lines. Thresholds considered in Figure 2 shown by vertical lines. Panels b,c,d show scatter plot of ASI values at (0m/s,0mm), (10m/s,10mm and (20m/s,20mm) respectively.

L7 – Here, please make it clear that the assumption is of the same number of events for each hazard. Why might this be appropriate for extreme wind, precipitation and flooding - particularly when they exceed a threshold? This can be discussed later. The number of impactful flooding and wind events expected in the UK are, for instance, expected to differ (e.g. see Hillier et al 2024 - doi:10.5194/gc-7-195-2024). If I have mis-understood, please use this opportunity to make the paper more widely accessible i.e. is this handled by the distribution shapes of X and Y.

For simplicity, our modelling approach uses only one count variable – the number of events where either wind or precipitation (or both) are extreme. This avoids the more complex problem of modelling the bivariate distribution of separate counts of when the wind and the precipitation are each extreme. The existing explanation (L89-92) in Section 2.2 has been amended to (now L108-112):

In this study, we shall consider events that have perils caused by two hazard variables X and Y with thresholds u_X and u_Y, respectively, resulting in annual ASI S_X and S_Y. The total number of events, N, only includes events that increase S_X or S_Y (or both) i.e. events where $X > u_X$ or $Y > u_Y$. This avoids having to model the bivariate distribution of separate counts for extremes in wind and precipitation. The count variable used in this study is an upper bound for these separate counts. For simplicity of notation, we shall refer to $X' = g(X)$ and $Y' = g(Y)$ simply as X and Y, respectively.

L14 – Please fix the grammar here. Perhaps, '... high thresholds'.

Changed to 'high thresholds'. (L14)

Response to Anonymous Referee #2

L57 – Is it important that the metrics used for 'correlation' in this paragraph are mentioned, as the difference between how these measures change in the same (or different ways) is an important to the interpretation of the models in this paper in the context of previous work. To illustrate that metric matters, it is perfectly possible for spearman's correlation to decrease as tail impact increases (i.e. see material at https://doi.org/10.5281/zenodo.12533412, Q11&12 on Task 2). Hillier et al (2025) – 10.1002/JOC.8763 – have also quantified co-occurrence and use Uplift.

Good point. The metrics used to measure association in each of the studies cited in this paragraph are now clearly stated in L97-L103. See response to the first comment for the revised paragraph.

This study uses Pearson's correlation to compute sample correlation values and as a basis to derive the frameworks. We acknowledge that Pearson's correlation is sensitive to outliers and it does not always capture the full picture of dependence. We have added L324-L325 to highlight this:

> *Pearson's correlation is used as a measure of dependency, this measure can be influenced by outlier events. Furthermore, zero correlation does not always imply independence (Embrechts et al., 2002)*

L57 – Another detail that needs to be noted to allow the reader to compare this work to previous work is how annual/seasonal are defined i.e. Oct-Mar or similar in previous work, and Jan-Dec here, effectively taking part of two storm seasons.

> Adapted so the correlation metrics for each study are mentioned and differences in timescales are highlighted (L64-65). Correlations between the ASIs for the extended winter are now shown in the appendix. Correlations computed over winter are more positive.

L70 – Applying a correlation between hazard variables in order to understand potential losses is not, as far as I am aware common. It is done in one context by Hillier (2025). Please add references to put this choice of approach in context, particularly if they've been missed by Hillier (2025).

> Line 70 (of the original manuscript) refers to the θ parameter of the framework. This correlation (between hazard variables) is a component of the overall framework which aims to capture the sample correlation ASIs. As far as we are aware, using constructing a framework for correlation between ASIs which uses the correlation between the hazards themselves is a new approach.
>
> As far as we can tell, Hillier et al. (2025) does not replicate this approach. Hillier et al. (2025) describes the ratio between two the number of events ("uplift") as correlation, comparing shorter and longer event windows.

L77 – The phrase '... following Jones et al (2024) ....' implies that Jones is the originator of this approach of SI over a threshold. Please add references and rephrase to better reflect the origin of the approach.

Response to Anonymous Referee #2

L79 – Similarly, the 20 ms-1 threshold is widely used and sensitivity tested (e.g. in Hillier & Dixon, 2020) although it long predates this. Please add references and rephrase to better reflect the origin of the approach.

> Removed "following Jones et al (2024)" and added "Klawa and Ulbrich (2003) were one of the first to use this threshold approach, noting German insurers usually pay for damages if a nearby weather station records gusts above 20ms^{-1}." on L89-90.

L83 – Similarly, for Section 2.2 concisely add some background for the use of aggregate SIs for wind and cyclones.

> We have expanded these sections to add more background (L96-103).

L105 – Fig. 1 is very helpful to improve accessibility. Thank you.

> Thank you - we are pleased to see that this has been helpful.

L115 – Please clarify in words whether Z modulates X and Y in the same way (i.e. up or down), or whether it is possible for Z to affect X and Y differently.

> We have now added "and can influence X and Y differently" to L150.

L160 – Following, or replicated from? Please clarify to make it clear whether Figure 2 is new work.

> Removed "Following Jones et al. (2024)". The sentence now reads "Figures 2a)-c) reproduce the sample correlation values at each grid point between wind and precipitation ASIs for different threshold levels shown in Jones et al. (2024)". (L200)

L163 – Please add a sentence here on whether or not the spatial pattern of high and low correlations match previous work (e.g. Martius, 2016; Hillier & Dixon, 2020). If there is a match, it will then be clear to the reader that there is agreement spatially, with the primary debate about the magnitude & sign of the correlation.

> Added *"These findings differ to existing wind-precipitation research. This study uses aggregate scores while Martius et al (2016) and Owen et al. (2021) didn't, these are computed over the calendar year rather than seasons like Hillier and Dixon (2020). This study also links wind and precipitation to tracked cyclones while Martius et al., (2016) and Hillier and Dixon (2020) use daily data."* (L208-211).

L193 – Red box for France. Is this 5° from the central point? If not, it would be better for comparison and consistency across the manuscript to make this match the approach on L157, and clarify there that it's a box not a radius.

> The France box refers to the region of interest. Any storm that comes within 5° of any land gridpoint within the box is considered. However multiple gridpoints within the box will have a SI from one storm event. To calculate one SI value for the France box, the SI for all of the gridpoints over land are summed. This is done for wind and then again for precipitation, giving each hazard one SI value for the entire region. This is now more carefully explained in the text.
>
> Changed to read: *"For each storm, the SI for the region is calculated by summing the SI from all land and sea grid points in [4.75°W-8.5°E, 42.25°-51.75°N]. For a given storm*

most gridpoint SIs within the region are zero, being >5° from the storm track or below the threshold." (L241-243)

L206 – This jet stream connection was first published in Hillier & Dixon (2020) - Dixon's idea. Please consider add the reference to this point.

Changed to: *Hillier and Dixon (2020) first proposed that a weaker jet stream is conducive to precipitation-only extremes. Manning et al. (2024) also concluded slow moving windstorms and a weaker jet stream are favourable for precipitation-only extremes. (L260-263)*

L235 – Fix brackets for references.

Fixed, thank you for spotting this error. (L303)

L251 – To link this work into that in the introduction, it is probably worth showing the Oct-Mar results in Supplementary Material so that readers can decide how to link this to earlier work.

Figure B has been added to the supplementary material, this shows framework estimate (a-c) for the sample correlation (d-f) for all storms with genesis times in the extended winter. This splits years from October-March by cyclone genesis time.

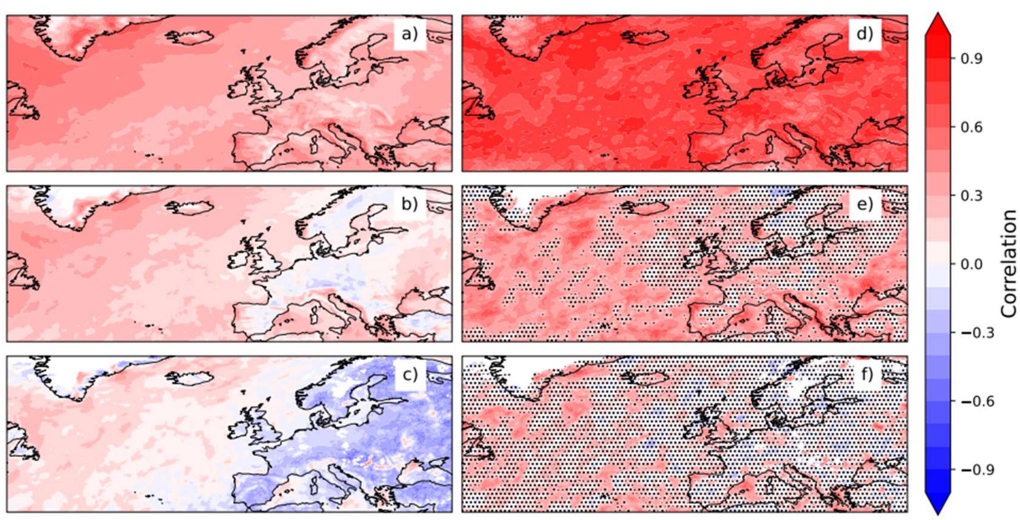

Figure B - Framework C estimate (a-c) of sample correlation (d-f) for ASIs calculated over the extended winter (1st October-31st March). Sample correlation not significant at the 5% level is shown by stippling.

Figure 2 – On a-c, please add stippling or similar to indicate areas where the observations have statistical significance (i.e. can be safely interpreted as existing). And, in the caption, decide if units are in italics or not.

This would add too much noise and reduce the message of the plot, however we have added a plot showing the framework and sample correlation with stippling to the supplementary material. The units in the captions are now in italics.

Figure 4 – Add error bars, or similar, to the sample correlation curve please.

A 95% confidence interval for sample correlation has been added.

Figure 5 – The two sets of numbers on the scale bar reduce clarity. Please find another solution. However, I like this figure; interesting similarities to Fig. 2g of Hillier & Dixon (2020) showing different wind directions for 'windy' and 'wet'.

The black bars showing days have been removed. The similarities to Fig. 2g of Hillier & Dixon (2020) are now mentioned in the text (L275-L277):

Hillier and Dixon (2020) found a similar contrast when considering wind direction on windy and wet locations days. Wind directions at a site on Scotland's east coast were south-westerly on days with extreme wind but north-easterly on days with extreme rain.

Figure 6 – Please add stippling or similar to indicate where the correlations are statistically significant (or not)

This would add too much noise and reduce the message of the plot, however we have added plots with stippling to the supplementary material.