<u>Response to Referee Comment 1: “Concurrent modes of climate variability linked to spatially compounding wind and precipitation extremes in the Northern Hemisphere”</u>

<u>Comments:</u>

The paper assesses wintertime, northern hemisphere modulation of the frequency of joint precipitation and wind events by four modes of variability. ENSO is found to be the most influential mode, combinations of modes are asserted to be essential for the co-occurrence. A doubling of the area affected by joint events is highlighted in the abstract. The analysis is probably sound although it is not currently possible to be certain, and it is currently not clear whether or not it is novel work as the Abstract and Introduction do not convincingly define the research gap being filled. Yet, it is entirely plausible that the gap is real. The manuscript would also benefit from substantive rewriting to improve its clarity. In short, I feel this has the potential to be an interesting paper if it is very significantly re-written, simplified and/or focussed and clarified.

<u>Response:</u>
We would like to thank the referee for their positive comments and the detailed questions. All the comments and our point-by-point responses are given below.

<u>Comments:</u>
Major concerns
1. As written, the abstract does not acknowledge any previous work linking joint extremes of precipitation and wind to modes such as ENSO. Linking extremes to ENSO and similar has been a common thing to do for at least 40 years, singularly and jointly. The abstract should be re-written to clearly express/define the research gap with respect to the existing literature.

<u>Response:</u>
We thank the referee for this comment. We modified the Abstract and Introduction. Please find below the modifications (in blue) of the Abstract (starting at L1 of the initial submitted article) related to this comment.

L1: “Compound wind and precipitation (CWP) extremes often cause severe impacts on human society and ecosystems, such as damage to crops and infrastructure. ~~High regional frequencies of CWP extremes across multiple regions in the same winter, referred to as spatially compounding events, can further impact the global economy and the reinsurance industry.~~ **Spatially compounding events with multiple regions affected by CWP extremes in the same winter can impact the global economy and reinsurance industry, however our understanding of these events is limited. While climate variability modes such as El Niño Southern Oscillation (ENSO) can influence the frequency of precipitation and wind extremes, their individual and combined effects on spatial co-occurrences of CWP extremes across the Northern Hemisphere have not been systematically examined.**”

<u>Comments:</u>
Also note that the use of multiple modes to understand the occurrence of extremes, particularly hurricanes, has a large literature (e.g. Vecchi, 2014). In short, by the end of the Introduction, it is still unclear what the novelty of the paper is (i.e. what is being done that

has not been done before). To be clear, there could be a gap (my knowledge of the literature is not comprehensive) but the authors have not convincingly stated what it is. Some papers that might be relevant in the re-framing are:

• Khouakhi (2017) Contribution of tropical cyclones to rainfall at global scale. i.e. ENSO and extreme rainfall.

• Bloomfield (2024) Synoptic conditions conducive for compound wind-flood events in Great Britain in present and future climates. Rain/flood and wind compound and NAO.

• Hillier (2020) [already cited] Multi-hazard dependencies can increase and decrease risk. Flooding and wind compounding and NAO.

• Vecchi (2014) On the seasonal forecasting of regional tropical cyclone activity. Introduction contains literature on use of multiple modes.

Response:

We thank the referee for this comment and the papers provided for the re-framing of the Introduction. We added the following texts in the Introduction to explicitly clarify the novelty of the paper:

L42 of the article initially submitted: "**Understanding the drivers of CWP extremes is crucial for optimising resource distribution during response efforts, yet research in this area remains limited. Several studies have explored the influence of single or combined mode of variability anomalies on precipitation and wind extremes in isolation, both at regional (e.g., Elsner et al., 2001, Abeysirigunawardena et al., 2009, Kossin et al., 2010, Grimm, 2011) and global scales (e.g., Khouakhi et al., 2017, Gao et al., 2022, Liu et al., 2024). Regarding CWP extremes, some studies have explored the influence of individual variability modes at the regional scale only (e.g., Hillier et al., 2020, Bloomfield et al., 2024), thus not allowing the assessment of the relation between concurrent climate variability modes and spatial co-occurrences of CWP extremes across several regions.**"

L76 of the article initially submitted: **"To the best of the authors' knowledge, the present study is the first to investigate CWP extremes and associated spatially compounding events across the Northern Hemisphere as mediated by multiple large-scale modes of variability. We address this research gap by using** large ensemble climate model…**"**

Comments:

2. As certain key pieces of information are not prominent early on, the manuscript (up to Results) can currently only be understood upon re-reading. Please rectify this.

Response:

We identify 1) the use of ***daily*** CWP extremes, 2) the focus on the winter defined as ***December-January-February months*** as partially missing information up to the Results section, 3) the derivation of the metrics at the seasonal time frame using seasonal counts, and 4) the focus on **increases** of CWP extremes and associated spatial co-occurrences. Changes have been made to the Abstract and Introduction to ensure that this information is highlighted correctly and at the right time to avoid re-reading. Additionally, we added a few words to the text (up to the Results section) to clarify this information throughout the reading.

Comments:

3. Throughout, more care should be taken to ground the work in related literature. Some parts are well referenced, but others are lacking (illustrative examples below).

Response:
Thanks for this comment. We included more literature throughout the article, taking care of the illustrative examples described by the referee below.

Comments:
4. The manuscript lacks a substantive discussion, e.g. about issues relating to the key findings of the paper. Perhaps, this will become clearer when the research gap and focus of the paper is more clearly defined. Perhaps separating the material in the Results into a descriptive (Results) and explanatory (Discussion) parts would help, and clarify the themes the authors wish to discuss.

Response:
Thanks for this comment. We have rewritten the Discussion Section, in addition to removing elements of discussion that were present in the Results section before, and thus were misplaced.

Comments:
Detailed Comments
Abstract:
L11 – The phrase 'combinations of modes are essential for the occurrence' is problematic. They might be necessary for a better description of the co-occurrence from this particular statistical viewpoint, but they are large-scale indicators of conditions, not processes that drive co-occurrence or not. So, please rephrase to a more precise statement.

Response:
We modified (in blue) the following sentence starting at L10 of the initial submitted article (Abstract):

"While ENSO is the most influential variability mode for such extreme spatially compounding events, ~~combinations of modes are essential for~~ the occurrence of these events ***increases further when multiple modes of variability are in anomalous phases. In particular****, combinations of modes increase both the number of regions and the population exposed to daily CWP extremes in the same winter.* ***For example, combined ENSO- and NAO+ nearly doubles*** *the number of affected regions compared to neutral conditions on average."*

Comments:
Introduction:
L25 – 'co-occurring compound' – this is tautology, remove one word.

Response:
Thanks for this comment. We suggest removing the word "co-occurring".

Comments:
L26 - (Jeong et al., 2020), this is not the only reference for this. Add 'e.g.' or/and cite another couple.

Response:

Thanks for this comment. We added three additional references:

*"For example, the combination of high winds and rainwater can result in severe damage due to the inflow of water through joints or cracks in building structures (**e.g., Blocken and Carmeliet, 2004; Mirrahimi et al., 2015; Martius et al., 2016,** Jeong et al., 2020)".*

Additional references added:

- Blocken B, Carmeliet J (2004) A review of wind-driven rain research in building science. J Wind Eng Ind Aerodyn 92(13):1079–1130
- Mirrahimi S, Lim CH, Surat M (2015) Review of method to estimation of wind-driven rain on building facade. Adv Environ Biol 9(2):18–23
- Martius, O., Pfahl, S., and Chevalier, C.: A global quantification of compound precipitation and wind extremes, Geophys. Res. Lett., 43, 7709–7717, https://doi.org/10.1002/2016GL070017, 2016.

Comments:

L44-45 (and throughout the manuscript) – Again, please use 'e.g.' when various papers could be cited. The manuscript should be improved by pairing an older reference with each of the recent singular references used. For example, convective storms were known to produce multiple hazards long before Dowdy & Catto in 2017.

Response:

Thanks for this comment. We added '***e.g.,***' to cite papers when appropriate. Also, we paired older references with recent singular ones, when possible. The list of new citations to answer this comment is provided below:

- L43: Ralph et al. (2006) for the links between atmospheric rivers and extremes.
- L45: Rappaport et al (2000) for the links between low-pressure systems and extremes.
- L45: Cerveny et al. (2000) for the links between tropical cyclones and extremes.
- L45: Raible (2007) for the links between extratropical cyclones and extremes.
- L102: Kawamura et al. (2004) for lagged effects of ENSO.
- L177: Bradley et al. (1968) for permutation tests.
- L451: Chongyin et al. (2001) for links between IOD and Asian monsoon.
- L454: Ashok et al. (2003) for the links between IOD and ENSO.
- L473: Kirov et al. (2002) for the interactions between ENSO and NAO.

Comments:

L56 – Why is an AMV abbreviation used here, but is the only mode in the abstract that is not abbreviated. Please be consistent.

Response:

Thanks for this comment. We added the AMV abbreviation in the abstract, ensuring consistency across modes in the manuscript.

Comments:

Methods:
L91: Daily data used. The focus on this timeframe should be made prominently and clearly (e.g. in the Abstract). [Returning to this, I suggest my comment also highlights that the explanation needs to be clarified]

Response:
Thanks for this comment. We simply replaced some of the terms “CWP extremes” by “daily CWP extremes” in the Abstract, Introduction and Methods sections to precise the focus of the timeframe.

Comments:
L98-104: This is unclear, I’m afraid (i.e. not reproducible). E.g. is NAO the mean of DJF, and if so, which days is it applied to the Jan-Dec year? Please clarify.

Response:
Thanks for this comment. We added some clarifications to the text:

L98 of the article initially submitted: “Seasonal indices for the two oceanic (ENSO and AMV) and two atmospheric (NAO and PNA) variability modes for both CESM and ERA5 data are calculated from monthly data using the National Center for Atmospheric Research (NCAR) data package Climate Variability Diagnostics Package (CVDP, Phillips et al., 2014). The **seasonal** indices for the NAO, PNA and AMV ~~are computed for the December–February periods while for~~ **are calculated as the mean of the monthly values for December, January and February**. **For these indices, December is taken from the year *n−1*, while January and February are taken from year *n*, and the resulting seasonal index is assigned to year *n*. For** ENSO we **proceeded similarly, but used November–January averages to account** for lagged effects (Li et al., 2011; Hong Lee et al., 2023).”

Comments:
L106: First explanation that winter is defined here as DJF. This is a key piece of information, and should be prominent (e.g. in the Abstract). Also, a brief explanation of why DJF is selected is needed.

L106: First indication that this analysis is based on seasonal counts. Again, this should be prominent, because not knowing it leads to a need to re-read the sections above. Please fix this to improve the readability of the manuscript.

Response:
Thanks for this comment. We now mention that winter is defined as December-January-February in the Abstract and Introduction. We also add a brief explanation for this choice, and precise that the analysis is based on seasonal counts.

L42 of the article initially submitted: ***“To advance our understanding of CWP extremes and their drivers, this study focuses on the Northern Hemisphere due to the high population density and the severe impacts of CWP extremes in this part of the world (e.g., Liberato et al., 2014, Wahl et al., 2015, Raveh et al., 2015). Such compound extremes are most frequent in coastal regions of the Pacific and Atlantic oceans (e.g.,***

Maraun et al. 2016) and in the winter season (e.g., Greeves et al., 2007, Hansen et al., 2019). Consequently, this study focuses on seasonal frequency of daily CWP extremes during December-February."

Comments:

L109: Why the 98th percentile?

Response:

Thanks for this comment. Using percentiles to define extremes is frequently done in the literature as, for example, it allows to control the sample size for robust statistical analyses (e.g., Zhang et al., 2011, Martius et al., 2016). Although using different variables (that is, wind gusts instead of wind speed), some studies considered the local 98th percentile to investigate precipitation and wind extremes (Martius et al., 2016). Also, Klawa and Ulbrich (2003) show that the local 98th wind percentile is a damage-relevant wind threshold for wind gusts. In our study, daily data for the December-January-February months are used. It represents 90 days by season. By choosing the 98th percentile as a threshold, the expected number of exceedances per season for wind and precipitation in isolation is equal to 90*0.02 ≈ 2 events per season, which can be considered sufficient to analyse co-occurrences of wind and precipitation values above these thresholds. Choosing a percentile higher than the 98th would allow us to focus on more extreme events (e.g., Zhang et al., 2011), but it would reduce the sample size for the analyses. Note that for model evaluation, we use the 95th percentiles, which is needed given the shorter length of reanalysis data. Although CWP events exceeding the 98th percentile of wind and precipitation can be considered moderate extremes, we think that they can still be considered impact-relevant. Choosing a local and impact-relevant threshold for wind and precipitation extremes is difficult, especially for precipitation for which incorporating effects such as surface runoff, snow melt and landslides would be needed (e.g., Williams, 1978) and is thus out of the scope of this study.

We suggest to add the following explanations (in blue) to the text:

L110 of the article initially submitted: "We use the 98th percentile of wind and precipitation over the 1950–2019 period for the main analysis based on data from the CESM model. **Percentile-based thresholds are frequently used to investigate climate extremes (e.g., Zhang et al., 2011, Martius et al., 2016). Following Klawa and Ulbrich (2003) and Martius et al (2016), we chose the 98th percentile, which is a compromise to capture the most extreme events in the CESM simulations while ensuring a sufficiently large sample size for robust statistical analysis.** For model evaluation, **which involves both the CESM model and ERA5 reanalyses (Figs. S1-S5 of the Supplement only)**, we use the 95th percentiles over the 1959-2019 period -- such a lower threshold allows for a more robust evaluation. **The reason for this is that,** given the ERA5's limited period, **extremes in the reanalysis data set are more scarce and associated statistics for very extreme events are largely affected by sampling uncertainty (Bevacqua et al., 2021b). Selecting a slightly lower threshold allows us to reduce this sampling uncertainty and thus improve confidence in assessing the model's ability to simulate extremes (e.g., Bevacqua et al., 2021b, Kelder et al., 2022, Fischer et al., 2023)**."

Following the suggestion of the third referee, we produced a sensitivity analysis by considering the 99th and 99.5th percentile to define seasonal counts of CWP extremes. New Figures S16-S19 have been added to the Supplement and display the results we obtained for:

- Metric 1: the influence of individual and concurrent variability modes on regional wintertime CWP frequency (Figs. S16 and S17, same results as those presented in Fig. 3 but consider the 99th and the 99.5th percentile, respectively).

- Metrics 2 and 3: the influence of variability modes on spatially CWP extremes (Figs. S18 and S19, same results as those presented in Fig. 6 but consider the 99th and the 99.5th percentile, respectively).

We added some sentences in the Discussion to summarize the results of the sensitivity analyses. For more details, see our response to comments from the third referee.

"**While the 98th percentile has been used in this study to focus on extremes and is relatively well-established in the literature (e.g., Klawa et al., 2003, Martius et al., 2016), other higher thresholds could have been chosen to consider more intense extreme events (e.g., Liu et al., 2013, Schar et al., 2016, Camuffo et al., 2020). Figs. S16-S19 show results from a sensitivity analysis on the influence of variability modes on regional CWP extremes (Metric 1) and spatially compounding events (Metrics 2 and 3) with the 99th and 99.5th percentiles used as thresholds. Although there are some variations in the results compared to those for the 98th percentile, the main conclusions drawn across the different thresholds are broadly consistent for all Metrics. The magnitude of the effects of the combinations are generally consistent across thresholds, and the combinations detected at higher thresholds are generally included among those identified at lower thresholds (Figs. S16-S19). Such slight differences may be due to larger sampling uncertainty for higher thresholds limiting the ability to detect significant effects for higher thresholds rather than different physical mechanisms involved for different thresholds. While the sensitivity analyses broadly indicate the robustness of most of our findings, possible relevant differences across thresholds highlight the importance of identifying impact-relevant thresholds, though this task is challenging (Williams, 1978, Bloomfield et al., 2023).**"

References:

- Zhang, X. B., L. Alexander, G. C. Hegerl, P. Jones, A. K. Tank, T. C. Peterson, B. Trewin, and F. W. Zwiers (2011), Indices for monitoring changes in extremes based on daily temperature and precipitation data, *Wires Clim. Change*, 2, 851–870.
- Klawa, M., and U. Ulbrich (2003), A model for the estimation of storm losses and the identification of severe winter storms in Germany, *Nat. Hazard Earth Syst. Sci*, 3, 725–732.
- Williams, G. P. (1978), Bank-full discharge of rivers, *Water Resour. Res.*, 14(6), 1141–1154, doi:10.1029/WR014i006p01141.

Comments:

L111: 'more robust evaluation' – this needs to be more specific please.

Response:

We agree with this comment and we modified (in blue) the following sentences starting at L111 of the initial submitted article (Method).

L110 of the article initially submitted: "We use the 98th percentile of wind and precipitation over the 1950–2019 period for the main analysis based on data from the CESM model. **Percentile-based thresholds are frequently used to investigate climate extremes (e.g., Zhang et al., 2011, Martius et al., 2016). Following Klawa and Ulbrich (2003) and Martius et al (2016), we chose the 98th percentile, which is a compromise to capture the most extreme events in the CESM simulations while ensuring a sufficiently large sample size for robust statistical analysis.** For model evaluation, **which involves both the CESM model and ERA5 reanalyses (Figs. S1-S5 of the Supplement only)**, we use the 95th percentiles over the 1959-2019 period -- such a lower threshold allows for a more robust evaluation. **The reason for this is that,** given the ERA5's limited period, **extremes in the reanalysis data set are more scarce and associated statistics for very extreme events are largely affected by sampling uncertainty (Bevacqua et al., 2021b). Selecting a slightly lower threshold allows us to reduce this sampling uncertainty and thus improve confidence in assessing the model's ability to simulate extremes (e.g., Bevacqua et al., 2021b, Kelder et al., 2022, Fischer et al., 2023)**."

Comments:

L115: It would be beneficial to set the choice of this metric (i.e. a count-based approach e.g. Hillier 2015; Bevaqua 2021 – Guidelines paper; Owen et al 2021 in Weather & Climate Extremes, χ) in the context of the metrics/timescales in previous work (e.g. Hillier & Dixon, 2020; Bloomfield, 2023 in Weather & Climate Extremes)

Response:

Thanks for this suggestion. We added (in blue) the following sentences before the explanations on Metric 1 starting at L106 of the article initially submitted (Methods):

L106 of the article initially submitted:

"2.2.1. CWP extremes

Many techniques have been utilized to characterize CWP extremes, with the selection of a specific method being guided by the research question. For example, the correlation between wind and precipitation has been quantified at daily to seasonal timescales (e.g., Matthews et al., 2014, Deluca et al., 2017, Hillier et al., 2020, Bloomfield et al., 2023). Logistic regression models have been applied to quantify the likelihood of a precipitation extreme occurring given the presence of a wind extreme (e.g., Martius et al., 2016). Alternative approaches include examining tail dependence (e.g., Vignotto et al., 2021) and employing impact-focused metrics (e.g., Hillier et al., 2015, Hillier et al., 2020, Bevacqua et al., 2021b). The most straightforward approaches include counting extreme wind and precipitation co-occurrences above a given percentile (e.g., Martius et al., 2016, Bevacqua et al., 2021b), or using extremal dependency measures estimating the probability of one variable being extreme given that the other one is extreme (e.g., Coles et al., 1999, Hillier et al., 2015, Owen et al., 2021). Here, to investigate winter season (December–February) CWP extremes at the grid cell level, we derived seasonal counts of CWP extremes, defined as wind and precipitation values simultaneously exceeding high thresholds. This results in one count per season per grid cell, which allows for investigating the effect of seasonally-averaged climate variability modes on the counts."

Comments:

L119: Two other metrics are 'introduced'. Although I cannot recall exact papers, I struggle to believe this is the first time these sorts of metrics have been used. Again, please place in context of similar metrics and usages with a few references.

Response:

These metrics (or similar ones) may have already been used in other studies. Therefore, we replaced the word "introduce" by "use" to avoid readers thinking that these measures are totally new.

For Metric 2, we modified the following sentence (in blue) L125 of the article initially submitted:

L125: "Then, **similar to Singh et al. (2021),** the total number of affected regions during the same winter is counted."

For Metric 3, we added the following sentence (in blue) L131 of the article initially submitted:
L131: "**Population weighting is utilized here as a surrogate for the assets at risk that could experience damages due to CWP extremes (e.g., Bloomfield et al., 2023).**"

Comments:

L139: 'some confounding effect may remain'. This is a rather important statement. It is good that the authors acknowledge it, however the key question is: How much, and does this impact the key results of the paper? Either by testing with simulated/idealised data, or perhaps another statistical method, I believe that the authors need to answer this question.

Response:

Thanks for this comment. We would like to note that by fixing the non-considered modes to neutral conditions, we devised an approach that allows for taking into account confounding effects arising from the considered modes. We modified the text around this aspect in the new text as we realised that this was not clear enough.

L139: "Following Singh et al. (2021), additionally conditioning all other modes in the neutral phase in (i) **serves** for better isolating the causal effects of the individual variability mode of interest. **Specifically, such additional conditioning allows for taking into account confounding effects arising from the considered modes, still** ~~Despite such additional conditioning~~ some confounding effects may remain. **In particular,** modes in neutral states still vary within the range of neutral conditions and we do not control for them. **In addition,** further effects may arise from variability mode not considered in this study. Similarly to the analysis of single variability modes, we quantify the effect of concurrent variability modes in non-neutral phases based on the ratio between…"

Still, as with any statistical method, some limitations exist, and we wanted to acknowledge this transparently. As in all approaches, confounding effects can also arise from non-considered variables. While adding other modes could have helped in this direction, a choice on the number of considered modes has to be made, and we considered modes known for having effects on the CWP extremes. Note also that adding extra conditioning variables/modes would reduce the sample size when deriving conditional statistics, therefore increasing uncertainties and decreasing the robustness of the results. We also note that other statistical methods, such as regression, are often used to investigate causal links (e.g., Pearl, 2013), but occurrence-based methods, as we use in our study, can also be considered for such purposes (e.g., Kretschmer et al., 2021).

Overall, while we recognize inherent limitations in studies like ours with respect to r confounding effects, we believe that our study provides an important overview into the effects of the combined drivers of CWP extremes and associated spatial compounding events in the Northern Hemisphere, paving the way for further research into causal effects.

We modified a sentence in the Discussion section:
"**Applying statistical methods such as regression techniques (e.g., Pearl, 2013, Kretschmer et al., 2021), or more advanced approaches such as causal networks (e.g., Nowack2020) may help to shed light on the complete causal pathway leading to spatially compounding events and better control potential confounding effects.**"

References:

- Kretschmer, M., S. V. Adams, A. Arribas, R. Prudden, N. Robinson, E. Saggioro, and T. G. Shepherd, 2021: Quantifying Causal Pathways of Teleconnections. *Bull. Amer. Meteor. Soc.*, **102**, E2247–E2263, https://doi.org/10.1175/BAMS-D-20-0117.1.
- Pearl, J., 2013: Linear models: A useful "microscope" for causal analysis. *J. Causal Inference*, 1, 155– 170, https://doi.org/10.1515/jci-2013-0003.
- Pearl, J., 2000: *Causality: Models, Reasoning and Inference.* 2nd ed. Cambridge University Press, 478 pp.

Comments:
L151-152: Scope limiting statement. Fair enough, but I believe this needs prominence in the paper, and am hoping it'll be so in the Discussion at least. It might also need to be in the abstract as it is a potential bias on all conclusions drawn, and so should be prominent for clarity.

Response:
We agree with the referee. This point should be clear as soon as possible in the manuscript. This point was already discussed in the originally submitted article (L461). We suggest adding the following words in the Abstract:

L4 of the originally submitted article: "Here, by combining reanalysis data and climate model simulations, we investigate how two oceanic and two atmospheric variability modes -- ENSO, the Atlantic Multidecadal Variability (AMV), the North Atlantic Oscillation (NAO) and the Pacific North American (PNA) -- **amplify the wintertime (December–February) frequency of daily CWP extremes** and associated spatial co-occurrences across the Northern Hemisphere."

We also provide such information in the Introduction:

L65 of the originally submitted article: "**In this study, we use** reanalyses **data** and large ensemble climate model simulations from the CESM General Circulation Model (Kay et al., 2015)**, to investigate the effects of ENSO, NAO, PNA and AMV modes of variability, and their combinations, on the increase in the frequency of December–January–February daily CWP extremes across the Northern Hemisphere.** [...] Specifically, we (1) analyse ~~the influence of~~ **how different** modes of variability ~~on~~ **increase** wintertime regional frequencies **(i.e., seasonal counts) of daily CWP extremes across individual regions** of ~~CWP extremes in individual regions across~~ the Northern Hemisphere."

__Comments:__

Section 2.2.4 – This approach seems reasonable, and statistically significance testing is critical in a paper like this, although I'd need to do a really careful read in a revised manuscript. Illustratively, the permutation procedure would need to account for dependency / relationships between the modes, or statistical significance of any results could be over-estimated (i.e. appear significant when they are not). And, whether this has been done is not currently clear to me.

__Response:__

Thanks for your comment. According to our interpretation of the reviewer statement "the permutation procedure should take into account the dependency/relationship between modes", our approach accounts for such dependencies.

The permutation procedure is mainly used to test, for a given CWP metric, if the average of this metric under some combination of modes is different from that during neutral conditions. We specifically focused on what we called *direct* and *combined* effects of modes, that is, for each combination under study, we additionally condition the other modes in the neutral state (see L134-L145 of the article initially submitted).

To perform the test for a given combination (e.g., ENSO+NAO-), the values of the CWP metric is collected for 1) all seasons associated with ENSO+, NAO-, (AMV and PNA being neutral), and 2) all seasons associated with all modes being neutral. Then, the ratio of averages of the CWP metric collected for 1) and 2) is computed. The obtained ratio is the "observed value" of the test statistic. **The test statistic thus is derived from combinations modes that are directly derived from original data and that, as such, accounts for the dependency/relationships between modes under study.**

Then, in a second step, the values used for the calculations of the means are randomly reassigned without remplacement to 1) seasons with ENSO+, NAO-, AMV being neutral and PNA being neutral, and 2) seasons with neutral conditions. Corresponding averages and ratios are then computed. **The permutation procedure thus assesses the significance of the relationship between [a] CWP metrics and [b] the mode combination under study by randomly breaking the relationship. Note that the relationship between the modes is not broken here, but only the relationship between CWP and modes.**

__Comments:__

Results:

L211: biases w.r.t ERA5. Fair enough, although I expect any relevant ones to be explicitly referred back to and results interpreted in light of this during the Discussion.

Response:

Thanks for this comment. Although climate model biases were initially partly discussed in the Discussion section, we chose to better highlight what the relevant biases are and how some of our obtained results should be taken with caution in light of these biases.

In the Discussion section: "**However, several results lack direct support from existing literature. While they may represent novel findings, they should be interpreted cautiously, considering the biases of CESM simulations relative to ERA5. Notably, the direct effect of ENSO+ on CWP extreme frequencies in CESM simulations exhibits some inconsistencies when compared to ERA5 over Northern Africa (Fig. S3). Better understanding and confirming the influence of climate modes on arid regions (e.g., Northern Africa), where CWP events may be less intense than in other areas, can support adaptation and mitigation policies. While CWP extremes can serve as an important source of freshwater (e.g., Berdugo2020), they also present a significant flood risk (e.g., Yin2023)**."

Comments:

Section 3.1 & 3.2: From a reader's point of view, it would be nice if this were significantly shorter, drawing out the main points of interest (i.e. that are new). Please review Section 3.1 as in a number of places it starts to discuss / explain the results to a level that is at or above the limit expected in a Results section.

Response:

Thanks for this comment. We suggest shortening Section 3.1 and 3.2 (see the modified article for the provided modifications). We tried to focus on the main points of interest (new results), despite the fact that Section 3.2 has more new interesting points than Section 3.1 (as most of the results in Section 3.1 obtained can be matched with some existing papers). We also removed identified elements of discussion and explanations of the physical mechanisms from Section 3.1 and 3.2. In addition, we also identified elements of discussion on causal links in Section 3.3, that we removed and discussed later in the Discussion section.

Comments:

L257 – 'we move to discussing' Please do not move to discussing in the results section. Please discuss in the Discussion.

Response:

In addition to changing the content of the results section, we modify the specified sentence as follows:

L258 of the article originally submitted: "**In the following, we** ~~move to discussing~~ **focus on describing the effects of a selection of mode combinations and regions in Figs. 3-5. To maintain clarity and conciseness, we do not discuss all regions and mode combinations in the text, and readers can explore specific regional effects directly in the figures.**"

Comments:
L315-326 – This seems like an expansion of or repeat of Methods. Consider moving to methods.

Response:
Thanks for this comment. It is true that this paragraph can be considered as an expansion of Methods. However, we think that it allows a smooth transition between sections. In addition to L315-326, we also identified another mention of methodology in Results (L329-333):
L329: "In general, dependencies between counts of CWP extremes in different regions can favour such spatially compounding events (Bevacqua et al., 2021) because regions connected by positive dependencies tend to experience CWP extremes at the same time. Thus, as a first step in the investigation of spatially compounding events, we analyse such dependencies. This also provides preliminary information on groups of regions that may be affected by CWP extremes during the same winters."

To address the referee's concern, we suggest removing parts of these two elements from the Results section and combined this information in Methods starting at L117 of the article initially submitted:

L117: "As Metric 1 is derived for each region individually, the influence of variability modes on the high regional frequencies of CWP extremes across multiple regions in the same winter (i.e., spatially compounding events) cannot be deduced~~. Therefore,~~ **directly. For example, based on Metric 1, we find that the variability mode phase ENSO+ modulates regionally averaged CWP extreme frequencies for North America and Central Asia. However, a possibility could be that half of the winter seasons with ENSO+ leads to increased CWP extremes for North America only, while the other half affects Central Asia, thus not simultaneously. Examining the dependencies between counts of CWP extremes in different regions can provide preliminary information on spatially compounding events (e.g., Bevacqua et al., 2021) because regions connected by positive dependencies tend to experience CWP extremes at the same time. Thus, as a first step for investigating spatially compounding events, we analyse dependencies in Metric 1 computed for different regions, so as to provide preliminary information on groups of regions that may be affected by CWP extremes during the same winters.**

Then, to examine spatially compounding events, we use two additional metrics. We employ these metrics to investigate the effects of variability modes on regional high frequencies of CWP extremes across multiple regions in the same winter (Metric 2) and on the total population of the Northern Hemisphere exposed to CWP extremes in the same winter (Metric 3)."

We then make the transition between subsections 3.2 and 3.3 shorter, which helps with following the flow of the analyses while still reducing the text, following the referee suggestion:

L315: "Results from Fig. 3 and the summary in Fig. 4 cannot be used to conclude whether the effects of concurrent variability modes lead to spatially compounding CWP extremes, that is, **to** high wintertime frequencies of CWP extreme across multiple regions during the same winter. These figures illustrate the effect of individual and concurrent variability modes on regionally averaged CWP extreme frequencies, which are derived for each region separately. ~~For example, it shows that ENSO+ significantly modulate regionally averaged CWP extreme frequencies for North America and Central Asia, but – in principle – a possibility could be that half of the winter seasons with ENSO+ leads to a significant increase of CWP extremes for North America only, while the other half affects Central Asia, therefore not simultaneously.~~ Nevertheless, the number of regions where each mode combination has significant effects in Fig. 4 (see numbers on the top of the matrix) suggests that some mode combinations may potentially lead to spatially compounding CWP extremes. For example, NAO-ENSO+ significantly enhances regional CWP extreme frequencies in eight regions, which means that if these regional effects of NAO-ENSO+ can manifest in the same winter, NAO-ENSO+ would lead to spatially compounding CWP extremes. **In the next section, we assess** whether individual and concurrent variability modes can lead to concurrent CWP extremes during the same winters across regions."

Comments:

L336-7 – This long-distance correlation is interesting. It is an example of the type of thing that could be expanded upon and discussed in a Discussion.

Response:

Thanks for this suggestion. We add in the Discussion section:

"Overall, although our aggregation in time and space may not be optimal for providing a fine-grained analysis of CWP events and additional modes may be relevant in some regions, this study provides a first comprehensive assessment of the interactions between multiple climate variability modes and the frequency of wintertime CWP extremes and associated spatially compounding events across regions of the Northern Hemisphere. **In particular, analyzing spatially compound events highlights the potential influence of variability modes that can link distant regions. These long-range relationships modulated by mode combinations can be explored in more detail, for example, with tailored experiments such as nudged atmospheric simulations."**

Comments:

L340 – 'we find that dependencies among regions overall enhance the potential for spatially compounding events' I am unsure how you can make this conclusion given that you were explicit earlier about only looking at enhancement not reduction of co-occurrence. Surely, both need to be looked at to comment on an overall effect.

Response:

Thanks for this comment. We think there is a misunderstanding here. Indeed, we refer to the fact that, in Fig. 5a, correlations between CWP extreme counts in the different regions are mainly positive. Regions that are positively correlated would tend to experience more CWP extremes at the same time than if CWP extremes in the regions were independent. Accordingly, Figure 5b shows that the dependency between the regions overall increases the number of regions that experience CWP extremes within the same winter, compared to the independent case. Thus, here the focus is on the increase of CWP extremes, compared to the independent case.

Comments:

L341 – This, and similar mentions of methodology in Results, should be put into Methods please.

Response:
Thanks. We removed the sentence and added it to the Methods. We did not find other mentions of methodology in Results that have not been already addressed.

L132 **"In addition to enabling quantifying the number of affected regions depending on variability modes, Metric 2 is also used in Fig. 5b to assess the effect of the dependencies between regions on spatially compounding events. This analysis is performed by comparing the number of affected regions (i.e., Metric 2) from the original dataset with the number obtained after breaking the dependencies via randomly shuffling regional CWP extreme counts using bootstrap in all regions in time (Bevacqua et al., 2021a)."**

Comments:
L366 – 'causal links among climate variability modes and oceanic modes exist'

Response:
It seems that there are no comments from the referee associated with this quote.

Comments:
Discussion:
L425-436 These are assertions, picking highlights from the results. These results are not discussed, i.e. reflected upon and put in the context of the literature. Suggest removing, or including in the Results.

Response:
Thanks for this comment. We moved and adapted it at the beginning of the Conclusion section.

Comments:
L437 – 445: Is a justification of the Methods, which I think is a repeat from the Methods section. Remove.

Response:
Thanks for this suggestion. We simply removed this part from the Conclusion section.

Comments:
L456 – This paragraph is a restatement of the approach, until L456 where an alignment with existing results is stated. So, it would be good to clarify what the new insights provided by this paper are.

Response:
Thanks for this comment. We suggest modifying (in blue) the following sentence starting L456 of the initial Manuscript to better clarify the new insights provided by our study.

L456: “Overall, although our aggregation in time and space may not be optimal for providing a fine-grained analysis of CWP events and additional modes may be relevant in some regions, **this study provides a first comprehensive assessment of the interactions between multiple climate variability modes and the frequency of wintertime CWP extremes and associated spatially compounding events across regions of the Northern Hemisphere**.”

Comments:
L461 – This paragraph is a caveat, which is OK, but should come after a substantive discussion.

Response:
Thanks for this comment. We let this paragraph follow elements of the discussion, following the suggestion of the referee.

Comments:
Conclusions:
The conclusions are suitable in style, but are difficult to comment while the assertions being made have not previously been discussed.

Response:
Thanks for this comment. We tried to address this issue by rewriting the Discussion section.

Comments:
Fig. 6 – It’s good to see the Bonferroni correction being used.

Response:
We thank the referee for the feedback.